

Universidade Federal do Rio Grande do Sul  
Faculdade de Direito  
Bacharelado em Ciências Jurídicas e Sociais

Daniele Verza Marcon

**Inteligência Artificial Autônoma:**  
Reflexos e desafios no âmbito da responsabilidade civil

Porto Alegre

2020

Daniele Verza Marcon

**Inteligência Artificial Autônoma:**

Reflexos e desafios no âmbito da responsabilidade civil

Trabalho de Conclusão de Curso apresentado como requisito parcial à obtenção do título de bacharela em Ciências Jurídicas e Sociais da Faculdade de Direito da Universidade Federal do Rio Grande do Sul.

Orientador: Prof.<sup>a</sup> Dra. Maria Cláudia Mércio Cachapuz

Porto Alegre

2020

Daniele Verza Marcon

**Inteligência Artificial Autônoma:**

Reflexos e desafios no âmbito da responsabilidade civil

Trabalho de Conclusão de Curso apresentado como requisito parcial à obtenção do título de bacharela em Ciências Jurídicas e Sociais da Faculdade de Direito da Universidade Federal do Rio Grande do Sul.

Orientador: Prof.<sup>a</sup> Dra. Maria Cláudia Mércio Cachapuz

**Aprovada em:**Porto Alegre, 21 de novembro de 2020.

BANCA EXAMINADORA:

---

Prof.<sup>a</sup> Dra. Maria Cláudia Mércio Cachapuz - Orientadora  
Universidade Federal do Rio Grande do Sul

---

Prof.<sup>a</sup> Dra. Kelly Lissandra Bruch  
Universidade Federal do Rio Grande do Sul

---

Prof.<sup>a</sup> Dra. Lisiane Feiten Wingert Ody  
Universidade Federal do Rio Grande do Sul

## AGRADECIMENTOS

Meu primeiro agradecimento não poderia deixar de ser para aqueles que são a minha base. À minha mãe, Lucélia, que me apresentou para o mundo da literatura e que lê para mim desde que eu “me conheço por gente”: obrigada por ser sempre o meu porto seguro e pelo apoio incondicional, que chega nas mais variadas formas (inclusive em marmita congelada). Ao meu irmão, Giovanni, por todas as vezes em que ouviu, de livre e espontânea vontade, as minhas palestras na cozinha sobre o tema e por opinar no desenvolvimento deste trabalho sempre que necessário. Ao meu pai, Nelson, por constantemente me lembrar que, no final, daria tudo certo. À minha “dindinha”, Fabiana, por compartilhar comigo o amor pela academia e por todos os conselhos (profissionais e de vida). E a todos eles, por entenderem os momentos em que me fiz ausente nos últimos meses.

À Faculdade de Direito da UFRGS, minha “segunda casa” pelos últimos cinco anos, e todas as professoras e professores que, com muito empenho, paciência e dedicação, me guiaram pelos caminhos das ciências jurídicas e sociais e contribuíram para que eu fosse capaz de formar juízo crítico sobre questões tão importantes que cercam nosso dia a dia, nossa universidade e nosso país: obrigada. De nada vale o conhecimento quando não compartilhado, e eu tenho muito orgulho de ter aprendido com todas e todos vocês e de concluir a minha graduação em uma universidade pública e de qualidade. Espero que um dia eu possa retribuir tudo o que pude aprender aqui – na academia, na profissão que eu exercer e, também, como cidadã.

Agradeço também a todos os técnicos administrativos que nos auxiliaram nesta caminhada e, em especial, à equipe da Biblioteca da Faculdade de Direito, pelo material cuidadosamente preparado para auxiliar na pesquisa e na escrita dos trabalhos acadêmicos e por todas as dúvidas prontamente atendidas, mesmo à distância.

Mas este trabalho – e a pesquisa por trás dele – não teria sido possível sem a ajuda de algumas pessoas.

À Profa. Maria Cláudia Cachapuz, agradeço por me mostrar a importância da pesquisa, das fontes do direito, da filosofia e do diálogo acadêmico aberto e sem melindres. Mais do que a orientadora deste trabalho, eu tive a felicidade de encontrar, logo no segundo semestre, a orientadora de todo um curso, e, para mim, é motivo de muito orgulho ter aprendido e convivido com ela ao longo dos últimos cinco anos, seja como aluna, monitora ou integrante do grupo de pesquisa “Direito Privado e Liberdades Cívicas”. Aos integrantes deste grupo, meu agradecimento pela constante troca de ideias, pelas profundas e mais variadas discussões

travadas ao longo dos últimos quatro anos e meio e pelo compartilhamento de livros e materiais de pesquisa.

Aos amigos que conheci no começo, no meio e no final do curso: Anna Laura Dal Molin, Bibiana Roenick, Eduarda Lyrio, Felipe Todeschini, Giovanna Belisário, Júlia Aita, Julia Goulart, Karina Camargo, Letícia Tedesco, Marina Radaelli, Matheus Freitas, Nicole Pretto, Olga Machado e Priscila Borges, por todos os momentos em que, longe ou perto, se fizeram presentes, contribuindo, cada qual de seu modo, para que fosse possível concluir não apenas este trabalho, como tantos outros desafios que acompanharam a nossa jornada pessoal, acadêmica e profissional.

Às minhas queridas colegas do Souto Correa Advogados, pelo apoio ao longo deste último ano, e ao Rodrigo Cantali, pelo incentivo à pesquisa, pelas provocações a respeito do tema, pelos artigos e reflexões compartilhadas nos últimos anos, que muito contribuíram para os resultados deste trabalho.

Às professoras Lisiane Ody e Kelly Bruch, pelas críticas, elogios, sugestões, e provocações na banca de avaliação, incorporadas com muito carinho à versão final do trabalho.

Por fim, aos meus anjos da guarda, Lukas Irion e Erika Donin. Escrever este trabalho demandou tempo. Tempo longe de casa, longe da família, longe dos amigos. E ao longo do ano e meio em que venho conduzindo esta pesquisa, demandou também a paciência daqueles que recebiam meus áudios, muitas vezes longos e em horários pouco apropriados, discutindo sobre robôs ou sobre algum artigo que eu havia acabado de ler – ou que, por outras tantas vezes, foram obrigados a me ouvir pessoalmente. E mesmo que poucas novidades tenham sobrado depois daqueles áudios, eles fizeram a gentileza de revisar e opinar sobre o texto escrito.

O percurso até aqui não teria sido tão precioso sem a paciência, as caronas e *playlists* elaboradas pelo Lukas (especialmente *Eazy*, trilha sonora que acompanhou a escrita de boa parte deste trabalho), que hoje conclui esta etapa comigo e que me acompanha desde o primeiro dia da graduação. Também não teria sido tão divertido sem as risadas e conselhos sempre sinceros da Erika, com quem compartilho o gosto pelas questões filosóficas mais variadas e que em muitos momentos mergulhou nos temas deste trabalho para desbravá-los comigo. A vocês dois, agradeço sobretudo pela amizade, por todas as conversas, jurídicas e não jurídicas, compartilhadas nos últimos anos, pelo toque de leveza tão fundamental que sempre trouxeram ao longo desta caminhada.

Sem a contribuição de cada uma dessas pessoas, este trabalho – e tudo o mais que o precedeu – não teria acontecido. Obrigada.

*Man shall be free and supreme; he shall have no other aim, no other labor, no other care than to perfect himself. He shall serve neither matter nor man. He will not be a machine and a device for production. He will be Lord of creation.*

Karel Čapek, *R.U.R.*, 2001

## RESUMO

O desenvolvimento de sistemas de inteligência artificial capazes de atingir objetivos independentemente da intervenção humana, a partir de processos de aprendizado como o *machine learning* e o *deep learning*, tem suscitado questionamentos sobre qual o tratamento jurídico adequado aos casos em que o uso dessa tecnologia puder causar danos a pessoas. Projetos de Lei apresentados no Congresso Nacional Brasileiro e a resolução do Parlamento Europeu de 16 de fevereiro de 2017, com recomendações à Comissão sobre disposições de Direito Civil sobre Robótica, apresentam duas possíveis perspectivas, quais sejam, considerar a inteligência artificial como objeto do direito ou como um sujeito. Neste caso, seria atribuída à inteligência artificial um estatuto jurídico próprio (personalidade eletrônica). Assim, com base no método indutivo de pesquisa e a partir da revisão bibliográfica sobre o tema, buscou-se analisar as duas perspectivas identificadas para o tratamento jurídico da inteligência artificial (inteligência artificial como sujeito e inteligência artificial como objeto), testando teorias da responsabilidade extracontratual que possam fundamentar soluções jurídicas de acordo com as regras do Código Civil Brasileiro. Rejeitou-se, *a priori*, a responsabilidade civil subjetiva, porque o elemento da culpa apresentaria desafios muitas vezes intransponíveis para se alcançar a reparação do dano. Porém, verificou-se que também as teorias de responsabilidade objetiva (baseada no risco, no fato do produto e no fato da coisa ou animal) e de responsabilidade indireta apresentam lacunas que poderiam impedir a solução do problema relacionado à utilização da inteligência artificial autônoma. Ao final, verificou-se a existência de lacunas (normativas e de reconhecimento) no ordenamento jurídico brasileiro com relação ao tratamento jurídico aplicável aos casos de responsabilidade civil envolvendo inteligência artificial, e sugeriu-se alternativas para o seu preenchimento.

**Palavras-chave:** Inteligência artificial. Algoritmos. Responsabilidade Civil. Lacunas jurídicas. Personalidade eletrônica.

## ABSTRACT

The development of artificial intelligent systems capable to achieve goals independently of human intervention, based on learning processes such as machine learning and deep learning, has raised questions about which legal treatment is appropriate in cases where the use of such technology causes damages to people. Bills presented on the Brazilian National Congress and the European Parliament resolution edited on February 16<sup>th</sup>, 2017, with recommendations to the Commission on Civil Law Rules on Robotics, provide two possible perspectives, namely, to consider artificial intelligence as an object of Law or as its subject. In this case, it would be given artificial intelligence its own legal status (electronic personality). Therefore, based on the inductive research method and bibliographic review about the issue, we sought to analyze the two perspectives identified for the legal treatment of artificial intelligence (artificial intelligence as a subject and artificial intelligence as an object), testing extracontractual liability theories that could justify legal solutions in accordance with the Brazilian Civil Code. *A priori*, fault-based civil liability has been rejected, because the fault element could present challenges that are often insurmountable to achieve damage compensation. However, it was verified that also theories of strict liability (based on risk, on product liability and on custody of animals or things) and vicarious liability have gaps that could prevent the solution of the problem related to the use of autonomous artificial intelligence. In the end, it was verified that there are gaps (normative and of recognition) in the Brazilian legal system regarding the applicable legal treatment to cases of civil liability involving artificial intelligence, and it was suggested alternatives for its filling.

**Keywords:** Artificial Intelligence. Algorithms. Civil liability. Legal gaps. Electronic personality.

## SUMÁRIO

<b>INTRODUÇÃO .....</b>	<b>10</b>
<b>1 A INTELIGÊNCIA ARTIFICIAL .....</b>	<b>14</b>
1.1 DA LITERATURA À MATEMÁTICA: O QUE É E COMO SURTIU A INTELIGÊNCIA ARTIFICIAL .....	15
1.2 ALGORITMOS, ROBÔS E O APRENDIZADO DE MÁQUINA .....	19
<b>2 INTELIGÊNCIA ARTIFICIAL NO PROCESSO DECISÓRIO: O QUE SE PODE APRENDER COM O COMPAS.....</b>	<b>27</b>
2.1 O CASO DE ERIC L. LOOMIS E A SUPREMA CORTE DE WISCONSIN .....	28
2.2 COMPORTAMENTO ENVIESADO E ESTATÍSTICA: ALGORITMOS SÃO MELHORES QUE HUMANOS? .....	30
<b>3 LACUNAS NO DIREITO E AS PROPOSTAS DE REGULAMENTAÇÃO DA INTELIGÊNCIA ARTIFICIAL NO PARLAMENTO BRASILEIRO E EUROPEU ....</b>	<b>36</b>
3.1 AS LACUNAS NO DIREITO .....	36
3.2 OS PROJETOS DE LEI NO PARLAMENTO BRASILEIRO .....	38
3.3 A RESOLUÇÃO DO PARLAMENTO EUROPEU.....	40
<b>4 NOTAS ELEMENTARES SOBRE RESPONSABILIDADE CIVIL .....</b>	<b>45</b>
4.1 A ORIGEM DA RESPONSABILIDADE CIVIL.....	45
4.2 RESPONSABILIDADE NEGOCIAL E EXTRA NEGOCIAL .....	48
4.3 RESPONSABILIDADE OBJETIVA E SUBJETIVA .....	51
4.4 PRESSUPOSTOS DA RESPONSABILIDADE CIVIL NO DIREITO BRASILEIRO ...	53
<b>5 PRIMEIRA PERSPECTIVA: INTELIGÊNCIA ARTIFICIAL COMO OBJETO.....</b>	<b>58</b>
5.1 RESPONSABILIDADE SUBJETIVA: O DESAFIO DA ANÁLISE DA CULPA E A INSUFICIÊNCIA DO REGIME PARA OS CASOS ENVOLVENDO IA .....	58
5.2 RESPONSABILIDADE OBJETIVA (OU <i>STRICT LIABILITY</i> ).....	61
<b>5.2.1 A Responsabilidade objetiva em razão de atividade de risco.....</b>	<b>63</b>
5.2.1.1 Os pontos positivos e negativos apontados na doutrina estrangeira para a responsabilidade por atividade de risco .....	63
5.2.1.2 O art. 927, parágrafo único, do Código Civil de 2002 .....	67
<b>5.2.2 Responsabilidade objetiva pelo fato do produto (<i>product liability</i>).....</b>	<b>71</b>
5.2.2.1 O debate acerca da responsabilidade fundamentada no fato do produto.....	71
5.2.2.2 O artigo 931 do Código Civil: um dispositivo de utilidade controvertida .....	73
<b>5.2.3 Responsabilidade objetiva pelo fato da coisa .....</b>	<b>76</b>

5.2.3.1	A teoria da guarda da coisa inanimada e do animal .....	77
5.2.3.2	Fato da coisa para a inteligência artificial? .....	80
5.3	RESPONSABILIDADE INDIRETA OU POR FATO DE OUTREM: A IA COMO AGENTE .....	83
5.4	RESPONSABILIDADE OBJETIVA PELO ART. 187 DO CÓDIGO CIVIL.....	88
<b>6</b>	<b>SEGUNDA PERSPECTIVA: A PESSOA ELETRÔNICA.....</b>	<b>93</b>
<b>7</b>	<b>CONCLUSÃO.....</b>	<b>104</b>
	<b>BIBLIOGRAFIA .....</b>	<b>115</b>

## INTRODUÇÃO

Publicado pela primeira vez em 1818 e considerado um dos principais romances para a origem do gênero literário de ficção científica, *Frankenstein ou o Prometeu moderno*, de Mary Shelley, narra a história de uma enorme criatura costurada a partir de pedaços de corpos humanos que é capaz de pensar e agir com autonomia e racionalidade, de compreender a linguagem humana, de se comunicar e desenvolver sentimentos. Victor Frankenstein, cientista que dá vida à criatura, é a metáfora moderna de Prometeu<sup>1</sup>, porque desafiou os deuses, dando vida à matéria morta e criando algo que, embora espelhado no homem, é totalmente irreconhecível como tal pela humanidade. Abandonada, a criatura de Frankenstein jura vingança à raça humana e persegue seu criador até o fim de seus dias (SHELLEY, 2008).

Frankenstein foi escrito para ser uma história aterrorizante – como de fato é. Além de traços marcantes da história da própria autora, o livro provoca a reflexão sobre as consequências da criação de entes não-humanos racionais que poderão, em dado momento, adquirir autonomia e agir de tal modo que seus criadores sejam incapazes de prever ou evitar seus próximos passos. Não por outro motivo, Frankenstein (além dos robôs de Asimov e de Karel Čapek) foi lembrado pelo Parlamento Europeu logo no início dos considerandos da Resolução de 16 de fevereiro de 2017, com recomendações à Comissão sobre disposições de Direito Civil sobre Robótica. Eis, aqui, a primeira razão desta pesquisa ter sido iniciada: se, em tempos de um imenso avanço da robótica, nos deparamos com textos legislativos com expressa referência à literatura da ficção, estaria a ficção se tornando realidade?

A quantidade de “novos amigos” de hoje pode indicar que sim. Dos algoritmos da Netflix e do Spotify, que indicam filmes e músicas que os usuários podem gostar, ao assistente do Google, Siri (Apple), Bixby (Samsung) e Alexa (Amazon), que auxiliam em diversas atividades do cotidiano em *tablets* e *smartphones*, é inegável que a tecnologia e, em especial, a inteligência artificial, seja a protagonista da vez.

---

<sup>1</sup> Conforme narra Thomas Bulfinch, “Prometeu era um dos titãs, uma raça de gigantes que habitava a terra antes da criação do homem. Ele e seu irmão Epimeteu foram incumbidos de fazer o homem e de dotá-lo, bem como a todos os outros animais, das faculdades necessárias à sua preservação. Epimeteu comprometeu-se a fazê-lo, e Prometeu encarregou-se de supervisionar a conclusão do trabalho do irmão. Epimeteu distribuiu aos diferentes animais os vários dons de coragem, força, agilidade, sagacidade (...) foi tão generoso ao distribuir seus recursos que, quando chegou o momento de prover o homem com faculdades que o fizessem superior a todos os outros animais, nada mais havia sobrado para legar-lhe. Perplexo, recorreu ao irmão Prometeu, que, com o auxílio de Minerva, subiu ao céu e acendeu a sua tocha na carruagem do sol, e trouxe o fogo para o homem. Com esse dom o homem tornou-se muito mais capaz que os outros animais” (BULFINCH, 2006, p. 30). O autor também comenta que Prometeu “é representado como o amigo da humanidade, aquele que por esta intercedeu quando Júpiter estava enfurecido contra os homens, aquele que também a ensinou a civilização e as artes”. (BULFINCH, 2006, p. 35)

O avanço da tecnologia tende a afetar a vida privada, as relações de emprego e mesmo a vida política. A relevância que o uso e desenvolvimento de sistemas inteligentes adquire no mundo contemporâneo e na ideia de futuro que se está construindo decorre justamente do fato de a tecnologia não deixar incólume nenhuma esfera de convívio.

Com efeito, uma análise da questão sob a teoria das esferas de Hannah Arendt pode demonstrar a amplitude que a IA representa na sociedade contemporânea. Arendt diferencia três esferas de convívio social, cada qual com uma característica principal: (i.) a privada, circunscrita às quatro paredes de convivência íntima das pessoas, em que impera a liberdade de estar só e o princípio da exclusividade; (ii.) a social, na qual se deve consagrar a igualdade material, em que a discriminação é válida no sentido de evitar a formação de sociedades massificadas e garantir a diversidade cultural e liberdade associativa dos grupos; e (iii.) a política, na qual se busca a igualdade formal e abstrata, onde a discriminação é, portanto, prejudicial e, por isso, deve ser evitada (ARENDR, 1957).

Algoritmos das redes sociais são um bom exemplo de como a tecnologia afeta essas três esferas. A pretexto de ampliar os contatos pessoais, a rede social impactou profundamente a noção atual de privacidade, agora tão facilmente “compartilhada”. Como apontou Zygmunt Bauman, “a privacidade invadiu, conquistou e colonizou o domínio público - mas à custa da perda de seu direito ao sigilo, sua característica definidora e seu privilégio mais valorizado e ardentemente defendido” (BAUMAN; LYON, 2013, p. 33). O paradoxo da “proximidade” (facilidade de aproximação das pessoas que, no mundo real, não são tão próximas vs. afastamento das pessoas muito próximas fora da realidade virtual) também permite verificar o impacto da rede social na formação de “comunidades virtuais” e padronização de relações culturais.

E como isso se dá? Um bom exemplo é referido por David Sumpter quando alertou sobre o perigo da “bolha filtro”, em que “[c]ada ação que você faz no seu navegador é usada para decidir o que lhe será mostrado depois” (SUMPTER, 2019, p. 150). Isso se aplica tanto ao conteúdo de informações disponíveis no *feed* de redes sociais quanto à interação com os “amigos”. Quanto mais *likes* os usuários derem a determinado tipo de conteúdo, maior a chance de que conteúdos similares sejam mostrados futuramente em detrimento de postagens divergentes; ou seja, algoritmos têm afetado não apenas com quem os usuários interagem, mas também o tipo de informação que recebem.

Exemplos de quão longe esse tipo de “condicionamento” pode chegar, afetando a esfera política, ocorreram recentemente nas eleições americanas em 2016 e, novamente, nas eleições brasileiras em 2018, com o uso de redes sociais para campanhas eleitorais e disseminação de

*fake news*<sup>2</sup>. Algoritmos também têm sido utilizados para a análise de desempenho de atletas, seleções de emprego, rankings universitários e a própria avaliação de risco e necessidades de pessoas respondendo a processos criminais, com base em cálculos atuariais (o COMPAS, que será abordado neste trabalho) (O'NEILL, 2016).

Embora nenhum dos exemplos acima seja necessariamente novo, todos eles demonstram que a discussão sobre a relação entre humanidade e tecnologia precisa ser enfrentada, notadamente diante da possibilidade de desenvolvimento de IA a partir de técnicas como *machine learning* e *deep learning*, em que os próprios desenvolvedores do algoritmo poderão desconhecer o seu integral funcionamento. Particularidades como essa demonstram que entender a interação entre algoritmos e humanos e compreender como o ordenamento jurídico poderá mitigar possíveis danos decorrentes do uso de IA ou assegurar a sua reparação se mostra, por isso, fundamental.

Nesse contexto, a preocupação central desta pesquisa é analisar quais as possíveis respostas para a reparação de danos causados diretamente pela IA autônoma. Isso porque o atual regramento da responsabilidade civil foi construído com base na premissa de que o ato ilícito/antijurídico será praticado por agentes humanos (ou, no máximo, sob a custódia de humanos). Assim, ainda que se admita a imputação do dever de indenizar ao Estado ou à pessoa jurídica, o ato que dá origem ao dano é praticado, direta ou indiretamente, por um ser humano, que, nestes casos, será vinculado ao ente estatal ou à pessoa jurídica responsabilizada.

O advento de algoritmos capazes de aprender a partir de experiências prévias e dados analisados, executando comandos que podem ser inclusive desconhecidos por seus desenvolvedores, pode representar uma mudança de paradigma, já que introduz um novo tipo de alguém que não é humano, mas que é capaz de praticar, de modo autônomo, atos que causam danos.

Questiona-se, assim, se a IA deve ser vista como um mero objeto do direito, dado que, até ulterior alteração legislativa em sentido diverso, o robô/algoritmo ainda é uma “coisa” que pertence e foi criado por alguém (mesmo que essa “coisa” seja diferente de tudo o que a

---

<sup>2</sup> O documentário “*The Social Dilemma*”, dirigido por Jeff Orlowski e produzido por Larissa Rhodes em 2020, ilustra o episódio. Cathy O’Neill (2016) também comenta sobre experimentos conduzidos pelo Facebook nas eleições americanas de 2010 e 2012. A ferramenta “*voter megaphone*”, desenvolvida para encorajar eleitores a votar, permitia que se compartilhasse na rede o comparecimento às urnas. Após comparar os resultados do impacto dessa ferramenta (o voto, nos Estados Unidos, é facultativo), pesquisadores estimaram que o Facebook contribuiu para aumentar o número de eleitores votantes em 340 mil (O’NEILL, 2016, p. 181). Em 2012, o Facebook alterou o algoritmo do *feed* da rede de 2 milhões de usuários, para que fossem priorizadas notícias em detrimento de postagens de lazer. A iniciativa foi liderada por Solomon Messing, pesquisador da companhia, que divulgou posteriormente uma estimativa de que a participação nas eleições dos usuários afetados pela manipulação do algoritmo aumentou de 64 para 67% (O’NEILL, 2016, p. 182).

humanidade já conheceu em termos de inteligência), ou se a IA poderia (ou deveria) responder pelos próprios atos, caso em que seria criada uma “pessoa eletrônica”, como sugeriu o Parlamento Europeu na Resolução aprovada em 16 de fevereiro de 2017. As discussões podem variar, também, a depender da origem da responsabilidade civil analisada, se negocial ou extranegocial, além, é claro, do papel da IA para a concretização do evento danoso.

Nesta pesquisa, será realizada análise da bibliografia já publicada a respeito da responsabilidade civil decorrente da utilização de inteligência artificial autônoma. Com base no método indutivo de pesquisa, busca-se compreender quais os possíveis impactos que o uso de algoritmos de *deep learning* e a IA (o “novo alguém”) representará para a análise da imputação da responsabilidade.

A exposição da pesquisa será feita em duas partes. A primeira parte, dividida em três capítulos, trata da origem histórica da inteligência artificial e a sua intersecção com a literatura, até o estágio atual, dos algoritmos de *deep learning* (capítulo 1). O segundo capítulo analisará o exemplo do COMPAS, ferramenta *evidence-based* que tem sido utilizada na análise de risco e necessidades de indivíduos acusados em processos criminais nos Estados Unidos, abordando as discussões sobre o uso de IA na tomada de decisões. Já o capítulo 3 abordará, sucintamente, as propostas legislativas voltadas à regulamentação do uso da IA no Brasil e na União Europeia, identificando os aspectos centrais relacionados à imputação do dever de indenizar.

Na segunda parte, inicialmente, serão abordados conceitos básicos da responsabilidade civil, da sua origem aos pressupostos para a sua caracterização (capítulo 4). Em seguida, os capítulos 5 e 6 ocupam-se da análise dos possíveis tratamentos jurídicos da IA sob duas diferentes perspectivas: primeiro, como um objeto do direito, de modo que responderá(ão) pelos danos o(s) humano(s) envolvido(s) na criação/controlar/utilização da IA (capítulo 5); segundo, como um sujeito a quem é reconhecida personalidade específica e independente de seu proprietário ou desenvolvedor (capítulo 6). O capítulo 7 apresenta a conclusão, analisando se o uso de IA implica, efetivamente, uma lacuna normativa que demanda regulamentação específica ou se a questão pode ser solucionada a partir das regras já existentes de responsabilidade civil do Código Civil de 2002.

## 1 A INTELIGÊNCIA ARTIFICIAL

Em 1920, o tcheco Karel Čapek escreveu a peça teatral intitulada “R.U.R. (*Rossum’s Universal Robots*)”, na qual introduziu, pela primeira vez, o conceito de “robô” (derivação da palavra tcheca *robota*, que significa “escravo” ou “trabalho compulsório”). Čapek imaginou um mundo em que todo o trabalho braçal seria feito por robôs projetados com os requisitos minimamente necessários para a atividade<sup>3</sup>. A redução de custos pela substituição dos trabalhadores humanos por robôs permitiria que tudo fosse produzido a um preço baixíssimo e acessível, “libertando” os humanos para se dedicarem às atividades de seus interesses e ao seu autodesenvolvimento. Em dado momento, contudo, os robôs concluem que são melhores que seus mestres e iniciam uma revolta que culmina com a extinção da raça humana (ČAPEK, 2001).

Incomodado com a retratação de robôs como uma ameaça à humanidade, Isaac Asimov decidiu escrever histórias em que os robôs conviveriam com humanos de modo harmônico (ASIMOV, 2014, p. 308). Surgiram, assim, robôs que cuidavam de crianças (*Robbie*), que administravam estações espaciais e tinham crenças religiosas (*Cutie*), ou mesmo que eram políticos ideais (*Stephen*). A harmonia da convivência com os humanos foi pautada no que Asimov chamou de “Três Leis da Robótica”, apresentadas no conto “*Runaround*”<sup>4</sup>, posteriormente incluído no livro “Eu, Robô”, cuja primeira edição saiu no ano de 1950.

As Leis da Robótica respeitam uma hierarquia, sendo a Primeira Lei a mais importante: “um robô não pode ferir um ser humano ou, por inação, permitir que um ser humano venha a ser ferido”. A Segunda Lei prevê que “um robô deve obedecer às ordens dadas por seres humanos, exceto nos casos em que tais ordens entrem em conflito com a Primeira Lei”, e a Terceira Lei estabelece que “um robô deve proteger sua própria existência, desde que tal proteção não entre em conflito com a Primeira ou com a Segunda Lei” (ASIMOV, 2014, p. 65). Anos depois, em 1985, no livro “Robôs e Império”, Asimov criou a Lei Zero: “um robô não

---

<sup>3</sup> No Primeiro Ato, ao explicar a sua visitante sobre os robôs da R.U.R., Harry Domin, Diretor Geral da companhia, relata que “*in ten years Rossum’s Universal Robots will produce so much corn, so much cloth, so much everything, that things will be practically without price. There will be no poverty. All work will be done by living machines. Everybody will be free from worry and liberated from the degradation of labor. Everybody will live only to perfect himself.*” (ČAPEK, 2001, p. 15).

<sup>4</sup> O conto foi publicado em português com o título “Andando em Círculos” e narra a história de uma expedição à Mercúrio. Speedy, o robô desenvolvido para explorar minério, começa a andar em círculos porque a Segunda e a Terceira Leis da Robótica entram em conflito em seu cérebro positrônico. Os humanos, Donovan e Powell, veem-se obrigados a encontrar uma solução para salvar o robô e acabam forçando-o a salvar um deles, já que a Primeira Lei (sobrevivência dos humanos) se sobreporia às demais (sobrevivência do robô) (ASIMOV, 2014).

pode fazer mal à humanidade e nem, por inação, permitir que ela sofra algum mal”<sup>5</sup> (PAGALLO, 2013, p. 23, tradução nossa).

As histórias acima criaram robôs muito mais avançados do que era concebível na época de sua publicação. Ainda assim, as fictícias Leis da Robótica de Asimov foram invocadas no texto da Resolução do Parlamento Europeu de 16 de fevereiro de 2017, com recomendações à Comissão sobre Disposições de Direito Civil sobre Robótica, como parâmetro ético para o desenvolvimento de IA<sup>6</sup>.

Com efeito, em tempos em que a literatura de ficção parece estar cada vez mais próxima de se tornar realidade, considerar as lições aprendidas nas histórias de Mary Shelley, Karel Čapek e Isaac Asimov pode não ser algo tão surreal<sup>7</sup>. A expressa menção às personagens criadas por aqueles autores na Resolução do Parlamento Europeu de 2017 indica, em verdade, que há uma linha bastante tênue separando a literatura de ficção das criações tecnológicas propiciadas por avanços na área da matemática e da ciência da computação.

O objetivo deste capítulo é, por isso, entender o que é inteligência artificial, qual o significado de conceitos como *machine learning* e *deep learning* e porque discutir sobre a necessidade (ou não) de inserção de novas regras no sistema jurídico relacionadas ao tema se tornou relevante nos últimos anos.

## 1.1 Da Literatura à Matemática: o que é e como surgiu a Inteligência Artificial

Os contos de Isaac Asimov são um exemplo de como os robôs e os programas de computador habitam o imaginário humano há décadas. Inicialmente, esses robôs eram apenas isso: ficção. Nada obstante, os passos iniciais para tornar reais os robôs de Asimov foram contemporâneos à publicação dos primeiros contos do escritor.

Foi Ada Lovelace que, em 1843, escreveu o primeiro algoritmo para ser processado por uma máquina, mas ele não foi imediatamente testado (RUSSEL; NORVIG, 1995, p. 15). Embora a criação de máquinas tenha sido relevante para desencadear a Segunda Revolução Industrial, os estudos voltados ao desenvolvimento de máquinas *inteligentes* tiveram avanço

<sup>5</sup> No original, “A robot may not injure humanity or, through inaction, allow humanity to come to harm.”

<sup>6</sup> “T. Considerando que as Leis de Asimov têm de ser encaradas como dirigindo-se aos criadores, aos produtores e aos operadores de robôs, incluindo robôs com autonomia integrada e autoaprendizagem, uma vez que não podem ser convertidas em código de máquina” (UNIÃO EUROPEIA, 2017).

<sup>7</sup> As obras de todos os autores também são referidas no corpo da Resolução de 2017, logo no primeiro considerando: “A. Considerando que desde o Frankenstein de Mary Shelley ao mito clássico do Pigmaleão, passando pela história do Golem de Praga pelo robô de Karel Čapek, que cunhou o termo, as pessoas têm fantasiado acerca da possibilidade de construir máquinas inteligentes, frequentemente andróides com características humanas”. (UNIÃO EUROPEIA, 2017)

significativo somente a partir da década de 1950, notadamente quando Alan Turing publicou o artigo “*Computing Machinery and Intelligence*”. Turing queria testar se máquinas poderiam pensar, e propôs, como teste, que uma máquina participasse do “jogo da imitação” contra um humano, fornecendo respostas escritas. O jogo da imitação tem três participantes: dois interrogados (A e B) e um interrogador (C); enquanto A tem por objetivo confundir C (interrogador), B deve ajudá-lo, falando a verdade. Turing questionou, então, se seria possível substituir o participante “A” por um computador digital e fazer com que o computador convencesse C de que B não era humano (TURING, 1950, p. 434).

Uma das primeiras objeções afastadas por Turing com relação à proposta foi a capacidade matemática. Isso porque o computador sairia em vantagem com relação ao humano se cálculos extensos e complexos fossem apresentados durante o jogo da imitação. A ideia de Turing era outra: fazer com que a máquina imitasse o *comportamento* humano, e não necessariamente o raciocínio (TURING, 1950, p. 435). O Teste de Turing, como ficou conhecido, está relacionado a uma possível definição para o que veio a ser chamado de inteligência artificial, mas não a única.

A expressão “inteligência artificial” foi cunhada somente em 1956, pelo cientista norte-americano John McCarthy, durante um seminário em Dartmouth College, Estados Unidos (RUSSEL; NORVIG, 1995, p. 17). A proposta do evento era desenvolver estudos baseados na ideia de que “todo aspecto de aprendizado e qualquer outra característica da inteligência pode, em princípio, ser tão precisamente descrita que uma máquina pode ser criada para simulá-la”<sup>8</sup> (RAJARAMAN, 2014, p. 201, tradução nossa).

A definição de IA é bastante aberta, variando conforme o trabalho de pesquisa. Silvia Díaz Alabart conceitua a inteligência artificial como “uma ciência que trata de estudar o comportamento inteligente através de modelos matemáticos”<sup>9</sup> (2018, p. 14, tradução nossa). Stuart J. Russel e Peter Norvig indicam quatro diferentes conceitos a partir da combinação das ideias de comportamento *versus* motivação e imitação humana *versus* racionalidade. Assim, uma IA pode (*i.*) agir de modo a imitar o comportamento humano ou de modo “racional”; ou, ainda, (*ii.*) pensar de modo a imitar o comportamento humano ou pensar de modo “racional”. Enquanto a imitação humana está relacionada à ciência empírica e às hipóteses de confirmação

---

<sup>8</sup> No original: “*every aspect of learning and any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it*”.

<sup>9</sup> No original, “*es una ciencia que trata de estudiar el comportamiento inteligente a través de modelos matemáticos*”.

experimental, a abordagem racional envolve a matemática e a engenharia (RUSSEL; NORVIG, 1995, p. 5).

Como exemplo de programas que agem como humanos, Russel e Norvig mencionam o Teste de Turing e o jogo da imitação (conforme já indicado, a ideia central do Teste é obter uma imitação comportamental, pois, se a máquina competisse com o humano para executar cálculos complexos, por exemplo, poderia deixar facilmente evidente quem era o humano e quem era a máquina). Os sistemas que pensam como humanos, por sua vez, dependem da ciência cognitiva e da própria compreensão de como a mente humana realiza os processos internos de decisão. Sistemas que pensam/agem racionalmente seriam baseados em silogismos e na construção de soluções lógicas formais, a partir de descrições sobre todas as coisas do mundo e a relação entre elas. Nessa abordagem, seria preciso mapear todas as possíveis premissas ligadas a determinada conclusão (o que demandaria extensa memória da máquina). De todo modo, a abordagem racional é, segundo os autores, mais vantajosa, porque tem definição clara e é completamente geral, ao passo que o comportamento humano “é bem adaptado para um ambiente específico e é o produto, em parte, de um complicado e longo processo evolutivo que ainda está longe de atingir a perfeição”<sup>10</sup> (RUSSEL; NORVIG, 1995, p. 8, tradução nossa).

Atualmente, a Universidade de Stanford mantém uma página na internet dedicada ao trabalho do professor McCarthy com um compilado de informações relacionadas ao estudo da IA<sup>11</sup>, onde consta que IA é “a ciência e engenharia de fazer máquinas inteligentes, especialmente programas de computador inteligentes. Está relacionada à tarefa similar de usar computadores para entender a inteligência humana, mas a IA não deve se limitar a métodos biologicamente observáveis”<sup>12</sup>, ao passo que “inteligência” é a “parte computacional da habilidade de atingir objetivos no mundo”<sup>13</sup> (JOHN MCCARTHY COMPUTER SCIENCE DEPARTMENT, 2007, tradução nossa).

O significado de “inteligência” é essencial para compreender a IA. Como apontaram Russel e Norvig, o Teste de Turing fornece apenas um parâmetro para conceituar um computador inteligente: a sua capacidade de imitar um ser humano. Mas esse é apenas um lado da questão; embora um computador que passe no Teste de Turing seja certamente inteligente,

<sup>10</sup> No original: “*is well-adapted for one specific environment and is the product, in part, of a complicated and largely unknown evolutionary process that still may be far from achieving perfection*”.

<sup>11</sup> A página está disponível em: <<http://jmc.stanford.edu/artificial-intelligence/>>. Acesso em: 25 ago. 2019.

<sup>12</sup> No original, “*It is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable*”.

<sup>13</sup> No original: “*Intelligence is the computational part of the ability to achieve goals in the world*”.

isso não significa que outro que não passe não o seja também<sup>14</sup>. Se a abordagem do computador que falha no Teste de Turing for de uma inteligência artificial baseada no agir racional, por exemplo, enganar o interrogador e convencê-lo de que a IA é um humano pode ser bastante difícil, mas isso não impedirá que a IA tome decisões autônomas ou seja independente de seus programadores nos resultados produzidos.

Silvia Díaz Alabart (2018) comenta que não é conveniente identificar a inteligência de uma máquina com a inteligência observada nos humanos, porque o significado de cada uma não se compara e cada uma funciona de forma muito diferente. Segundo a professora espanhola:

Os humanos têm uma inteligência que poderíamos chamar de generalista, porque lhes permite desenvolver comportamentos em todos os aspectos da vida, enquanto as máquinas de hoje são totalmente especializadas, e apenas desenvolvem comportamentos muito específicos. Os humanos têm senso comum e emoções, as máquinas carecem de ambas as coisas. Além disso, um dos problemas que surge é que não é factível traduzir à linguagem dos algoritmos os princípios éticos das pessoas, de modo que não é possível introduzi-los em um robô.<sup>15</sup> (ALABART, 2018, p. 14-15, tradução nossa).

A questão relativa ao desenvolvimento de emoções será retomada no capítulo 6, que trata do reconhecimento de personalidade jurídica própria aos robôs. Por ora, assumiremos que, seja qual for o modelo de abordagem (comportamental ou racional), desenvolver inteligência artificial significa atribuir a um programa a capacidade de aprender com a própria experiência e treinar a si próprio. E é essa capacidade de aprendizado e agir autônomos – de maneira independente dos humanos, embora à sua similaridade – que torna a IA tão revolucionária. Conforme destacam Paulius Čerka, Jurgita Grigienė e Gintare Sirbikyte:

O desenvolvimento da Inteligência Artificial é voltado a fazê-la autotreinável (a capacidade de acumular experiência pessoal) ou aprendizado de máquina. Essa característica única permite que a IA aja de modo diferente em situações iguais, dependendo das ações realizadas previamente. Isso é muito similar à experiência humana. Modelagem cognitiva e técnicas de pensamento racional dão mais flexibilidade e permitem a criação de programas que podem “entender”, *i.e.* que tem

---

<sup>14</sup> Conforme apontado no material disponibilizado na página dedicada ao professor John McCarthy: “*The Turing test is a one-sided test. A machine that passes the test should certainly be considered intelligent, but a machine could still be considered intelligent without knowing enough about humans to imitate a human. Daniel Dennett's book Brainchildren has an excellent discussion of the Turing test and the various partial Turing tests that have been implemented, i.e. with restrictions on the observer's knowledge of AI and the subject matter of questioning. It turns out that some people are easily led into believing that a rather dumb program is intelligent.*” (JOHN MCCARTHY COMPUTER SCIENCE DEPARTMENT, 2007)

<sup>15</sup> No original, “*Los humanos tienen una inteligencia que podríamos llamar generalista en cuanto que permite desarrollar comportamientos en todos los aspectos de la vida, mientras que las máquinas hoy por hoy, están totalmente especializadas, y solo desarrollan comportamientos muy específicos. Los humanos tienen sentido común y emociones, las máquinas carecen de ambas las cosas. Es más, uno de los problemas que se plantean, es que no es factible traducir al lenguaje de los algoritmos los principios éticos de las personas, por lo que no es posible introducirlos en un robot.*”

os traços de uma pessoa razoável (atividades de processo cerebral).<sup>16</sup> (2015, p. 378, tradução nossa)

Para eventual finalidade normativa, portanto, não é imprescindível estabelecer um conceito fechado para a inteligência artificial, desde que as principais características que permitem identificar uma IA sejam referidas. Nesse sentido, a Resolução do Parlamento Europeu de 16 de fevereiro de 2017 propõe que o conceito de “robôs autônomos inteligentes” considere a sua capacidade de: (i.) “adquirir autonomia através de sensores e/ou da troca de dados com o seu ambiente (interconectividade) e a análise destes dados”; (ii.) “aprender com a experiência e com a interação”; (iii.) “adaptar o seu comportamento e as suas ações ao ambiente”, bem como (iv.) “a forma do suporte físico do robô” (UNIÃO EUROPEIA, 2017).

A importância dessas características será analisada a seguir.

## 1.2 Algoritmos, Robôs e o aprendizado de máquina

É comum encontrar referências à IA por meio de termos como “algoritmo”, “programa de computador”, “*software*” e “robô”. Embora relacionados até certo ponto, esses conceitos não compartilham do mesmo significado técnico.

O conceito e a função de um algoritmo são bem explicados por Sílvia Díaz Alabart:

Os *algoritmos* são estruturas básicas para a resolução de problemas. Essencialmente, são um conjunto de instruções, detalhadas passo a passo, que tratam de obter um resultado que controle um determinado conjunto de variáveis em uma ordem específica. A ordem é fundamental para o funcionamento do algoritmo, já que afetará diretamente como o computador resolverá o problema proposto.<sup>17</sup> (2018, p. 16, tradução nossa).

Ou seja, algoritmo é uma sequência de instruções que analisa as informações de “entrada” (*input*) para produzir a “saída” (*output*). O algoritmo será lido por um “programa”<sup>18</sup>

<sup>16</sup> No original, “*The development of Artificial Intelligence is aimed at making it self-training (the ability to accumulate personal experience) or machine learning. This unique feature enables AI to act differently in the same situations, depending on the actions previously performed. This is very similar to human experience. Cognitive modeling and rational thinking techniques give more flexibility and allow for the creation of programs that can “understand,” i.e. that have the traits of a reasonable person (brain activity processes).*”

<sup>17</sup> No original: “*Los algoritmos son estructuras básicas para la resolución de problemas. En esencia, son un conjunto de instrucciones, detalladas paso a paso, que tratan de obtener un resultado que controle un determinado conjunto de variables en un orden específico. El orden es fundamental para el funcionamiento del algoritmo, ya que afectará directamente a cómo el ordenador trate de resolver el problema propuesto.*”

<sup>18</sup> A definição de programa consta do art. 1º da Lei n.º 9.609/98: “Programa de computador é a expressão de um conjunto organizado de instruções em linguagem natural ou codificada, contida em suporte físico de qualquer natureza, de emprego necessário em máquinas automáticas de tratamento da informação, dispositivos, instrumentos ou equipamentos periféricos, baseados em técnica digital ou análoga, para fazê-los funcionar de modo e para fins determinados” (BRASIL, 1998).

ou um “*software*”<sup>19</sup> que estará abrigado em um suporte físico, denominado *hardware*. A capacidade do *hardware* é outro fator relevante, já que afetará a capacidade de armazenamento de informações do sistema.

Já um robô, conforme explicam Neil Richards e William Smart é “um sistema construído que exhibe ação física e mental, mas não está vivo no sentido biológico”<sup>20</sup> ou “algo fabricado que se move pelo mundo, parece tomar decisões racionais sobre o que fazer e é uma máquina”<sup>21</sup> (RICHARDS; SMART, 2016, p. 6, tradução nossa). Essa definição, segundo os autores, “exclui inteligências artificiais totalmente baseadas em software que não exercem nenhuma ação no mundo físico”<sup>22</sup> (RICHARDS; SMART, 2016, p. 6, tradução nossa).

Ou seja, enquanto um programa é estático, o robô se movimenta no mundo físico. Ele pode ser remotamente controlado tanto por seres humanos como por programas de computador inteligentes (RICHARDS; SMART, 2016). São os robôs controlados por programas de computador inteligentes que interessam, particularmente, a esta pesquisa, já que o seu comportamento é, ao cabo, independente do controle humano.

Assim, a IA pode estar presente na sociedade tanto através de robôs, que interagem diretamente com o meio físico, como por meio de algoritmos lidos por programas de computador, cujos exemplos mais populares se encontram nas redes sociais e *smartphones*. O questionamento que fundamenta esta pesquisa considera a IA autônoma, isto é, robôs controlados por computadores e algoritmos de *deep learning*, pois, em ambos os casos, as ações do robô e do algoritmo serão tomadas de modo independente da ação humana, mas, ainda assim, serão reputadas inteligentes.

Não se pode deixar de notar que nenhum conceito até aqui apresentado seja “novo”. Ora, a introdução de computadores foi um dos principais marcos da Terceira Revolução Industrial, iniciada na segunda metade do século XX e que, como aponta Klaus Schwab, foi catalisada pelo desenvolvimento “da computação mediante servidores tipo *mainframe* (nos

---

<sup>19</sup> Há uma sutil distinção entre “programa” e “*software*”. Conforme esclarece Ian Sommerville, “Muitas pessoas pensam que software é simplesmente outra palavra para programas de computador. No entanto, quando falamos de engenharia de software, não se trata apenas do programa em si, mas de toda a documentação associada e dados de configurações necessários para fazer esse programa operar corretamente. Um sistema de software desenvolvido profissionalmente é, com frequência, mais do que apenas um programa; ele normalmente consiste em uma série de programas separados e arquivos de configuração que são usados para configurar esses programas. Isso pode incluir documentação do sistema, que descreve a sua estrutura; documentação do usuário, que explica como usar o sistema; e sites, para usuários baixarem a informação recente do produto” (SOMMERVILLE, 2011, p. 3).

<sup>20</sup> No original, “*a constructed system that displays both physical and mental agency but is not alive in the biological sense*”.

<sup>21</sup> No original, “*something manufactured that moves about the world, seems to make rational decisions about what to do, and is a machine*”.

<sup>22</sup> No original, “*excludes wholly software-based artificial intelligences that exert no agency in the physical world*”.

anos sessenta), a informática pessoal (décadas de 1970 e 1980) e internet (década de 1990)”<sup>23</sup> (SCHWAB, 2017, p. 20, tradução nossa).

O que se tem de novo é a capacidade de processamento dos computadores que desencadearam a Terceira Revolução Industrial e a multiplicidade de tarefas que tais máquinas se tornaram capazes de desempenhar. Devido ao avanço dessas tecnologias, Schwab sustenta que, atualmente, estaria em curso uma Quarta Revolução Industrial, que tem como base a revolução digital. Segundo o autor, as principais características dessa nova revolução são (i.) internet onipresente e móvel, (ii.) sensores pequenos, potentes e cada vez mais baratos e (iii.) inteligência artificial e aprendizado de máquina (SCHWAB, 2017). Todos esses fatores contribuem para um processo de digitalização e, conseqüentemente, de automatização. Em especial quanto à IA, Schwab destaca que:

A IA fez avanços impressionantes, impulsionados pelo aumento exponencial no poder computacional e na disponibilidade de grandes quantidades de dados, desde o software usado para descobrir novos medicamentos até os algoritmos que preveem nossos interesses culturais. Muitos deles aprendem com o rastro de 'migalhas de pão' que nossos dados deixam no mundo digital, resultando em novos tipos de 'aprendizado de máquina' e a descoberta automatizada que permite aos robôs e computadores 'inteligentes' se auto programarem e encontrar soluções ideais a partir de princípios básicos.<sup>24</sup> (SCHWAB, 2017, p. 25)

Há, portanto, uma enorme distinção entre (i.) a automação que marcou a Segunda Revolução Industrial<sup>25</sup>, (ii.) a computação que ascendeu na Terceira Revolução Industrial e (iii.) a automatização de processos hodiernamente, que decorre da digitalização. Até o advento da inteligência artificial, a computação consistia basicamente em descrever etapas processadas pelo programa, de modo que todas as tarefas executadas pelo computador dependeriam do algoritmo fornecido pelo programador e do conhecimento que este possuía. A consequência é que o computador poderia otimizar o desempenho humano e, em alguns casos, substituí-lo integralmente, mas as suas soluções seriam, em qualquer caso, sempre previsíveis.

---

<sup>23</sup> No original, “*la computación mediante servidores tipo «mainframe» (en los años sesenta), la informática personal (décadas de 1970 y 1980) e internet (década de 1990)*”.

<sup>24</sup> No original, “*La IA ha logrado avances impresionantes, impulsada por el aumento exponencial de la potencia de cómputo y por la disponibilidad de grandes cantidades de datos, desde el software utilizado para descubrir nuevos fármacos hasta los algoritmos que predicen nuestros intereses culturales. Muchos de estos aprenden del rastro de «migalhas de pan» que nuestros datos dejan en el mundo digital, lo cual da por resultado nuevos tipos de «aprendizaje de máquina» y el descubrimiento automatizado que les permite a los robots y ordenadores «inteligentes» autoprogramarse y encontrar soluciones óptimas a partir de principios básicos*”.

<sup>25</sup> A definição de automação pelo *Dictionary of World History* é “o uso de máquinas automáticas ou sistemas, particularmente de manufatura ou sistemas de processamento de dados que requerem pouca ou nenhuma intervenção humana na sua operação normal” (tradução nossa). No original, “*the use of automatic machinery and systems, particularly those manufacturing or data-processing systems which require little or no human intervention in their normal operation*” (2000, p. 47).

Atualmente, a realidade da computação é muito diferente. Essa distinção se deve, em especial, a dois conceitos característicos da inteligência artificial: o *machine learning* e o *deep learning*. Ambos os conceitos se referem à capacidade diferenciada da inteligência artificial de aprender a partir das informações a que teve acesso.

Em 1950, Alan Turing indicou que a máquina submetida ao jogo da imitação deveria ser baseada em processos de aprendizado, semelhantes aos de uma criança. Assim, o programador poderia treinar a máquina, que aprenderia com as experiências às quais fosse submetida para, então, enfrentar um humano no Teste de Turing. Quanto à possibilidade de aprendizado da máquina, Turing apontou que o “professor” estaria constantemente ignorante sobre o que se passa no processamento da máquina, embora pudesse prever, até certo ponto, como a máquina iria se comportar:

Isso está em claro contraste com o procedimento normal quando se usa uma máquina para fazer cálculos, o objetivo é ter uma imagem mental clara do estado da máquina em cada momento do cálculo. Este objeto só pode ser alcançado com esforço. A visão de que ‘a máquina só pode fazer o que sabemos mandá-la fazer’ parece estranha diante disso. A maioria dos programas que podemos colocar na máquina resultará nela fazendo algo que não podemos compreender (se no todo, ou o que consideramos como um comportamento completamente aleatório).<sup>26</sup> (TURING, 1950, p. 459, tradução nossa)

Se Alan Turing já apontava a possibilidade de desenvolvimento de máquinas que aprendem em 1950, também o *machine learning* não é uma novidade do século XXI, embora a sua concretização seja uma das grandes responsáveis para a criação dos novos “amigos” de hoje (Siri, Alexa, Assistente do Google, Bixby). Segundo Will Knight, o verdadeiro responsável pela “explosão” da inteligência artificial nos últimos anos é o “*deep learning*” (KNIGHT, 2017). Knight explica que o *deep learning* é um tipo de algoritmo baseado em redes neurais que se assemelha ao processo de atividade cerebral humana. Diferentemente da programação tradicional, no caso do *deep learning*, é o próprio computador que “se programa”. Trata-se, pois, de uma modalidade bastante desenvolvida de *machine learning*.

Contudo, se, de um lado, o uso de algoritmos de *deep learning* permite que a IA alcance soluções em princípio inimagináveis para o ser humano, de outro, permitir que aprendizados algorítmicos complexos sejam utilizados sem qualquer restrição pode tornar a tomada de decisão automatizada inescrutável, como uma caixa preta (em inglês, “*black-box*”) (KNIGHT,

---

<sup>26</sup> No original, “*This is in clear contrast with normal procedure when using a machine to do computations one’s object is then to have a clear mental picture of the state of the machine at each moment in the computation. This object can only be achieved with a struggle. The view that “the machine can only do what we know how to order it to do,” appears strange in face of this. Most of the programmes which we can put into the machine will result in its doing something that we cannot make sense (if at all, or which we regard as completely random behavior).*”



Curiosamente, algumas negociações entre Bob e Alice foram exitosas. Apesar disso, o Facebook decidiu desligar ambos os sistemas, já que o objetivo inicial era que eles fossem capazes de conversar com os usuários da plataforma, o que demandaria, evidentemente, uma linguagem que pudesse ser compreendida pelos humanos, e não apenas pela IA.

Outro exemplo paradigmático referente à capacidade de aprendizado da IA é o robô Gaak, desenvolvido em um experimento de “robôs vivos” na Inglaterra, voltado à confirmação do princípio da sobrevivência dos mais fortes aplicado à IA. Em 2002, Gaak, que estava sendo treinado para agir como predador, fugiu da arena de testes e foi parar no estacionamento. Conforme matéria publicada na época pelo jornal britânico *The Guardian*, a ênfase atribuída à caça na programação do robô pode ter contribuído para a sua tentativa de ver o restante do mundo (WAINWRIGHT, 2002). Čerka, Grigiene e Sirbikyte resumiram o episódio:

Em 2002, o Magna Science Center em Rotherham, Inglaterra, dirigiu um projeto chamado “*Living Robots*”. Robôs atribuídos às funções de “predadores” e “presas” foram lançados em uma arena de 2.000 metros quadrados. Usando sensores infravermelhos, a “presa” procurava comida indicada pela luz e os “predadores” as caçavam e sugavam sua energia. A Inteligência Artificial opera por meio de redes neurais. O experimento teve que confirmar o princípio da “sobrevivência do mais apto” e revelar se os robôs (e, também, a IA) eram capazes de se beneficiar da experiência adquirida, ou seja, de criar de forma independente novas técnicas de caça e autodefesa. Durante o experimento, o robô Gaak, que foi involuntariamente deixado sem vigilância por 15 minutos, conseguiu escapar. O robô abriu caminho pela parede da “prisão”, encontrou uma saída e foi para a rua. Chegou ao estacionamento, onde foi atropelado por um carro.<sup>30</sup> (ČERKA; GRIGIENE; SIRBIKYTE, 2015, p. 381)

Os dois exemplos acima ilustram a premissa já indicada anteriormente: é a capacidade da IA de atingir determinado objetivo que lhe caracteriza como *inteligente*. Uma vez criada e

---

<sup>30</sup> No original, “*In 2002, the Magna Science Centre in Rotherham, England, ran a project called “Living Robots.” Robots assigned to the roles of “predators” and “prey” were released into an arena of 2000 square meters. Using infrared sensors, the “prey” was looking for food indicated by light and “predators” were hunting them and draining their energy. Artificial Intelligence operates by means of neural networks. The experiment had to confirm the principle of the “survival of the fittest” and reveal whether the robots (and AI as well) were able to benefit from the gained experience, i.e. to independently come up with new hunting and self-defense techniques. During the experiment, a robot Gaak that was unintentionally left unattended for 15 min managed to escape. The robot forced its way through the “prison” wall, found an exit and went to the street. It reached the parking lot where it was hit by a car.*” Tathiana Cristina Fontão Pires e Rafael Peteffi da Silva utilizaram-se de descrição praticamente idêntica em artigo publicado em 2017: “Um exemplo singelo para ilustrar a capacidade da IA de aprender e responder ao meio ambiente, independentemente da vontade do seu desenvolvedor, ocorreu em 2002: o episódio da fuga do robô Gaak do *Magna Science Center*, na Inglaterra. A referida instituição realizou um projeto chamado “Robôs Vivos” (*Living Robots*), que consistiu em atribuir aos robôs os papéis de “predador” ou de “presa”, com a diretriz “caçar”, para o primeiro, e “fugir” para o segundo, lançando-os em seguida em uma arena. Usando sensores infravermelhos, a “presa” procurava alimentos indicados pela luz e o “predador” caçava e drenava a sua energia. O intuito do experimento era verificar se o princípio da sobrevivência do mais apto seria aplicável aos robôs dotados de IA e se eles poderiam se beneficiar da experiência adquirida, ou seja, se eram capazes de criar, de forma independente, novas técnicas de caça e autodefesa. Ocorre que, durante o experimento, o robô Gaak “presa”, involuntariamente deixado sem vigilância durante 15 minutos, conseguiu escapar da arena. O robô atravessou o muro da “prisão”, encontrou uma saída e foi para a rua. Chegou ao estacionamento onde foi atingido por um carro.” (SILVA; PIRES, 2017, p. 243)

instigada a se desenvolver por meio de aprendizado de máquina, a IA poderá criar soluções ou algoritmos próprios que não foram inicialmente previstos. Em razão disso, o cientista da computação Stephen Omohundro destacou que a IA poderá adquirir comportamentos perigosos se não for desenvolvida cuidadosamente.

As observações de Omohundro se mostram particularmente relevantes para a análise da responsabilidade civil relacionada à IA, já que a falta de cuidado na criação da IA pode desencadear condutas indesejadas e resultar em danos. O comportamento “perigoso” da IA está associado aos recursos que o sistema se utiliza para alcançar seus objetivos, sendo características intrínsecas à IA avançada: (i.) o desejo de autodesenvolvimento ou busca pelo próprio aperfeiçoamento; (ii.) o desejo de ser racional; (iii.) o desejo de preservar a utilidade das próprias funções; (iv.) tentativas de prevenir a falsificação de suas utilidades; (v.) autoproteção; (vi.) o desejo de adquirir novos recursos e de utilizá-los de maneira eficiente (OMOHUNDRO, 2008). Omohundro destaca que compreender essas características é essencial para que o desenvolvimento tecnológico viabilize um futuro positivo para a humanidade – e não o contrário.

As possíveis consequências negativas do desenvolvimento inadvertido de algoritmos foram ilustradas em recente documentário dirigido por Jeff Orlowski intitulado “*The Social Dilemma*”. A partir de depoimentos de profissionais envolvidos com o desenvolvimento de algoritmos e funcionalidades disponíveis em redes sociais como Facebook, Instagram, LinkedIn, Snapchat, Twitter, Google (desde a ferramenta de pesquisa e vídeos do YouTube até o próprio Gmail), ilustrados na exibição paralela da história de uma família viciada em seus *smartphones*, o documentário aborda como o uso de redes sociais e da tecnologia tem afetado aspectos importantes do convívio social, desde a educação de crianças e adolescentes até a disseminação das chamadas “*fake news*” (THE SOCIAL DILEMMA, 2020). A preocupação expressa no documentário pode ser resumida grosseiramente em duas frases, exibidas entre as entrevistas e a história da família “se você não está pagando por isso, você é o produto”<sup>31</sup> e “há apenas duas indústrias que referem-se aos seus consumidores como ‘usuários’: drogas ilícitas e *software*”<sup>32</sup> (THE SOCIAL DILEMMA, 2020, tradução nossa).

---

<sup>31</sup> No original, “*If You're Not Paying For It, You Become The Product*”. A expressão é utilizada há anos para tratar da relação das redes sociais e funcionalidades *online* com a sociedade. Parte dos temas tratados em *The Social Dilemma* foram apontadas em artigo publicado na revista americana *Forbes* ainda em 2012, disponível neste link: <https://www.forbes.com/sites/marketshare/2012/03/05/if-youre-not-paying-for-it-you-become-the-product/#2aecf9945d6e>. Acesso em 19 set. 2020.

<sup>32</sup> No original: “*There are only two industries that refer to their customers as ‘users’: illegal drugs and software.*” A frase é atribuída ao estatístico americano Edward R. Tufte (THE SOCIAL DILEMMA, 2020).

Dessa interação – bastante presente atualmente – já se pode identificar inúmeras situações em que algoritmos podem dar causa – ou contribuir – a eventos danosos. Seja na facilitação da veiculação de notícias falsas ou o notório papel das redes sociais na criação de estereótipos de beleza e sucesso pessoal e o seu impacto no aumento da depressão e ansiedade em usuários, entender como – e até que ponto ainda é possível – controlar a tecnologia tornou-se um passo imprescindível para definir se eventos vinculados ao uso de *softwares* e, especificamente, IA, poderão ser considerados danos indenizáveis e, nesse caso, a quem se deve imputar o dever de indenizá-los.

## 2 INTELIGÊNCIA ARTIFICIAL NO PROCESSO DECISÓRIO: O QUE SE PODE APRENDER COM O COMPAS

Bob, Alice e Gaak são exemplos de como a inteligência artificial pode surpreender humanos com algo totalmente inesperado – mesmo que essa surpresa esteja associada com o “treinamento” que é dado à IA. Todos esses experimentos demonstram que a IA tem potencial de afetar seus “interlocutores” a partir das decisões tomadas com base em seus processos internos de aprendizado. A fim de demonstrar possíveis implicações práticas, o objeto deste capítulo não é de um experimento, mas uma IA já implementada, cuja utilização foi analisada pelo judiciário americano em 2016.

O COMPAS ou “*Correctional Offender Management Profiling for Alternative Sanctions*” é um *software* de quarta geração, desenvolvido pela empresa americana Northpointe para auxiliar na avaliação de risco e necessidades de indivíduos respondendo a processos criminais. O COMPAS é uma das ferramentas denominadas “*evidence-based practices*”, que, como aponta Antônio Tovo (2019), “estão a amparar a dosimetria de sentenças, fixação de regimes de cumprimento e a execução de penas privativas de liberdade, dentre outras etapas do procedimento criminal”. Baseadas em cálculos atuariais, tais ferramentas “orientam o processo de decisão dos julgadores a partir de análises probabilísticas do potencial de reincidência do acusado (presumidamente aferido por dados da experiência, portanto *evidence-based*)” (TOVO, 2019).

Anne L. Washington aponta que “a literatura sobre pontuações de análise de risco geralmente afirma que elas são baseadas em dados administrativos, registros públicos, autoavaliações e entrevistas”<sup>33</sup>(WASHINGTON, 2019, p. 139, tradução nossa). Além dessas fontes, o relatório elaborado pelo COMPAS considera as respostas da pessoa sob análise a 137 questões (ANGWIN *et al.*, 2016). As informações de entrada (*input*) do algoritmo incluem: comportamento cognitivo, ambiente social, uso de drogas, educação, histórico familiar e de violência, instabilidade residencial, problemas financeiros, isolamento social ou falha de socialização e a análise de como o indivíduo lida com o lazer e o tédio (ou a sua falta de perspectivas/metapas)<sup>34</sup>. A pontuação de saída (*output*) varia entre 1 e 10, analisando tanto o risco de reincidência quanto a periculosidade do indivíduo para a sociedade. As pontuações são

<sup>33</sup> No original: “[t]he literature on risk assessment scores generally states that they are based on administrative data, public records, self-reporting, and interviews”.

<sup>34</sup> Todas as informações estão disponíveis em “*Practitioner’s Guide to COMPAS Core*”, disponível para download no website da empresa: <http://www.northpointeinc.com/downloads/compas/Practitioners-Guide-COMPAS-Core-031915.pdf>. Acesso em: 05 out. 2019.

agrupadas em três classificações: um a quatro indicam risco baixo, cinco a sete indicam risco médio e oito a dez indicam risco alto (NORTHPOINTE INC, 2015).

Em levantamento realizado pelo *Electronic Privacy Information Center*<sup>35</sup> em fevereiro de 2020, apontou-se que o COMPAS é utilizado em quatro estados americanos (Flórida, Michigan, Wisconsin e Wyoming). A ferramenta *evidence-based* que está presente no maior número de estados é a *Public Safety Assessment* (“PSA”); entre a fase de testes e a implementação total, a PSA aparece em ao menos vinte estados.

O COMPAS não é uma IA autônoma e, por ser uma ferramenta desenvolvida especificamente para a análise de risco em processos criminais, estudá-lo em um trabalho que trata exclusivamente da responsabilidade civil pode parecer inadequado. Ainda assim, o estudo do *software* americano pode ser valioso: o COMPAS foi objeto de discussão em caso paradigma relacionado aos reflexos do uso de algoritmos em processos decisórios, notadamente porque é uma das poucas ferramentas de IA efetivamente submetida ao escrutínio judicial e público, do que resultaram importantes questões a serem consideradas quando se trata de responsabilidade.

## 2.1 O caso de Eric L. Loomis e a Suprema Corte de Wisconsin

Em 2013, Eric L. Loomis foi preso após uma ocorrência de disparo de dois tiros em via pública na cidade de La Crosse, Wisconsin, e indiciado por cinco delitos. Em acordo realizado com a Promotoria, Loomis renunciou ao seu direito a julgamento, negou envolvimento com os disparos e a posse da arma de fogo e admitiu a culpa pelos delitos de tentativa de fuga de agente de trânsito e de condução de veículo sem o consentimento do proprietário; defesa e acusação acordaram que Loomis ficaria em liberdade condicional por um ano e que a sentença final deveria ser proferida pela corte distrital. Após homologar o acordo, a corte distrital ordenou a realização de uma Investigação Pré-sentença (“*Presentence Investigation*” ou “PSI”), à qual seria anexado o relatório de análise de risco do COMPAS.

Ocorre que o relatório do COMPAS considerou as respostas de Loomis ao questionário e apontou que ele apresentava alto risco para a fase de instrução, alto risco de reincidência e de

---

<sup>35</sup> O estudo intitulado “*Algorithms in the Criminal Justice System: Pre-Trial Risk Assessment Tools*” destaca que “*Jurisdictions have generally used one of three main systems, or adapted their own version of each: Correctional Offender Management Profiling for Alternative Sanctions (COMPAS), Public Safety Assessment (PSA) and Level of Service Inventory Revised (LSI-R). COMPAS, created by the for-profit company Northpointe, assesses variables under five main areas: criminal involvement, relationships/lifestyles, personality/attitudes, family, and social exclusion. The LSI-R, developed by Canadian company Multi-Health Systems, also pulls information from a wide set of factors, ranging from criminal history to personality patterns. Using a narrower set of parameters, the Public Safety Assessment, developed by the Laura and John Arnold Foundation, only considers variables that relate to a defendant’s age and criminal history*” (ELECTRONIC PRIVACY INFORMATION CENTER, 2020).

reincidência violenta. Esse relatório foi expressamente mencionado na sentença<sup>36</sup> como uma das razões para afastar a possibilidade de liberdade condicional e condenar Loomis à pena de sete anos (sendo os quatro iniciais em reclusão), pelo crime de dirigir veículo sem o consentimento do proprietário, e outros quatro anos (sendo os dois iniciais em reclusão) pelo crime de tentativa de fuga de agente de trânsito (WASHINGTON, 2019, p. 138).

O caso foi levado à Suprema Corte de Wisconsin. O argumento central do recurso de Loomis foi a alegação de violação do direito ao devido processo legal, porque a condenação teria considerado o relatório do COMPAS, mas (i.) o critério utilizado pelo algoritmo que culminou com a indicação de alto risco de violência e reincidência não foi divulgado e (ii.) embora Loomis conhecesse as perguntas e respostas fornecidas, não se sabia qual era o efetivo peso de cada uma no resultado do processo algorítmico (*output*). A defesa também sustentou (iii.) que a pontuação do COMPAS é baseada em dados de grupos, e não no indivíduo, de modo que o julgamento não estaria considerando apenas a pessoa do acusado; (iv.) a inexistência de um estudo de validação dos dados do *input* do COMPAS e os dados específicos no Estado de Wisconsin; e (v.) que o COMPAS atribuía pontuações diferentes em razão do gênero do indivíduo analisado. O Estado de Wisconsin destacou, em resposta, que o relatório do COMPAS foi apenas um dos elementos considerados na sentença, mas que não foi o fator determinante para a conclusão do magistrado<sup>37</sup>. Devido aos possíveis reflexos da decisão no *software* comercializado pela Northpointe<sup>38</sup>, a empresa também se manifestou, defendendo o sigilo do algoritmo.

O recurso de Loomis foi desprovido em 2016. A Suprema Corte de Wisconsin acolheu a alegação da Northpointe de que o algoritmo do COMPAS estava protegido por sigilo comercial e direitos de propriedade intelectual e indeferiu o pedido de sua divulgação. Destacou-se no julgado que “ferramentas *evidence-based* podem amparar o trabalho do Juiz sentenciante, mas não podem ser a única fonte de fundamentação” (TOVO, 2019), bem como “que o acusado deve receber um esclarecimento prévio sobre o uso de tais ferramentas de

---

<sup>36</sup> O trecho foi transcrito pela defesa de Loomis no pedido de *writ of certiorari* endereçado à Suprema Corte dos Estados Unidos: “*You’re identified, through the COMPAS assessment, as an individual who is at high risk to the community. In terms of weighing the various factors, I’m ruling out probation because of the seriousness of the crime and because your history, your history on supervision, and the risk assessment tools that have been utilized, suggest that your (sic) extremely high risk to re-offend*”. Disponível em: <<https://www.scotusblog.com/wp-content/uploads/2017/02/16-6387-cert-petition.pdf>>. Acesso em: 16 mai. 2020.

<sup>37</sup> Os documentos do caso e as alegações das partes estão disponíveis no site da Suprema Corte dos Estados Unidos: <<https://www.scotusblog.com/case-files/cases/loomis-v-wisconsin/>>. Acesso em: 16 mai. 2020.

<sup>38</sup> À época; atualmente, o COMPAS é comercializado pela empresa *Equivant*.

avaliação de risco e que suas respostas podem servir de parâmetro para graduação da sentença” (TOVO, 2019)<sup>39</sup>.

A defesa de Loomis peticionou perante a Suprema Corte dos Estados Unidos em outubro de 2016. Na resposta ao pedido de *writ of certiorari*<sup>40</sup>, o Estado de Wisconsin destacou que caso de Loomis era apenas o segundo no país a tratar do uso de *evidence-based practices* no processo decisório; o primeiro foi o caso *Malenchik v. Indiana*<sup>41</sup>, no qual a Suprema Corte de Indiana concluiu que as ferramentas de *risk assessment* forneciam informações apropriadas para consideração judicial no momento de sentenciar<sup>42</sup>. Nesse sentido, devido à insignificante quantidade de casos sobre a questão, seria prudente aguardar maior desenvolvimento para que a Suprema Corte dos Estados Unidos se manifestasse.

Na qualidade de *amicus curiae*, o Procurador-Geral dos Estados Unidos manifestou-se pelo indeferimento do *writ of certiorari*, o que efetivamente ocorreu, dando-se o caso por encerrado em 2017.

## 2.2 Comportamento enviesado e estatística: algoritmos são melhores que humanos?

Em março de 2016, antes que a Suprema Corte de Wisconsin julgasse a apelação de Loomis, jornalistas da ProPublica publicaram uma matéria apontando que o algoritmo do

<sup>39</sup> Tovo (2019) critica, nesse sentido, que a utilização de sistemas *evidence-based* como o COMPAS aponta para um retorno ao direito penal do autor, já que as características particulares da pessoa acusada podem se mostrar mais relevantes que os fatos do caso em julgamento.

<sup>40</sup> Diferentemente do Brasil, em que o critério específico de admissibilidade de um recurso extraordinário ao Supremo Tribunal Federal é a repercussão geral da matéria, a Suprema Corte dos Estados Unidos pode decidir se irá ou não revisar a decisão de determinado caso. Para solicitar o julgamento, a parte deve apresentar um “*writ of certiorari*”, que deve ser concedido por, no mínimo, quatro juízes. A denegação da *certiorari* impede o julgamento do caso pela Suprema Corte Americana. Ver mais em: <https://www.law.cornell.edu/wex/certiorari>. Acesso em: 29 set. 2019.

<sup>41</sup> A íntegra da decisão está disponível em: <https://www.in.gov/judiciary/opinions/pdf/06091001bd.pdf>. No caso de Anthony Malenchik, foram utilizadas duas ferramentas *evidence-based*, a LSI-R e a SASSI. A LSI-R é uma ferramenta de terceira geração que está entre as três principais dos Estados Unidos e que considera no *input*, dentre outros dados, as companhias do indivíduo sob análise, seu estado civil, histórico familiar, moradia, educação e informações emocionais/pessoais (maiores informações estão disponíveis no site da desenvolvedora: <https://storefront.mhs.com/collections/lsi-r>). Na fundamentação da decisão, a Suprema Corte de Indiana refere que o “*SASSI scoring model ‘is a brief selfreport, easily administered psychological screening measure’ that includes items ‘to identify some individuals with alcohol and other drug problems who are unwilling or unable to acknowledge’ substance abuse warning signs or symptoms*”. Mais adiante, refere-se que “*the SASSI is not designed to diagnosis a substance abuse disorder. It only seeks to identify persons with a high probability for substance dependence*” (INDIANA, 2010, p. 10). Ou seja, tal como o COMPAS, os sistemas *evidence-based* do Estado de Indiana utilizam-se de dados diversos (*input*) para mensurar, numericamente (*output*), quanto uma pessoa é ou não perigosa para a sociedade, inclusive com a probabilidade de envolvimento com drogas.

<sup>42</sup> Íntegra da conclusão da Suprema Corte de Indiana: “*We hold that the results of LSI-R and SASSI offender assessment instruments are appropriate supplemental tools for judicial consideration at sentencing. These evaluations and their scores are not intended to serve as aggravating or mitigating circumstances nor to determine the gross length of sentence, but a trial court may employ such results in formulating the manner in which a sentence is to be served*” (INDIANA, 2010, p. 14).

COMPAS apresentava viés racial (WASHINGTON, 2019, p. 149). Apesar de o COMPAS acertar em 61% dos casos, negros recebiam pontuações de alto risco duas vezes mais que brancos, embora não voltassem a reincidir (ANGWIN *et al.*, 2016).

Para comprovar as conclusões da matéria, a ProPublica disponibilizou vasto material que foi analisado por jornalistas, acadêmicos, profissionais da justiça criminal, estatísticos, cientistas políticos e especialistas em *machine learning*. A Northpointe respondeu com novas publicações e a discussão que se seguiu envolveu, em síntese, definições de justiça, explicação algorítmica e indicadores de base populacional (WASHINGTON, 2019, p. 148).

No contexto do debate encabeçado pela ProPublica e pela Northpointe, Julia Dressel e Hany Farid, especialistas em estatística de Dartmouth College, nos Estados Unidos, decidiram verificar se “algoritmos são melhores do que humanos não treinados para prever a reincidência em uma forma justa e precisa”<sup>43</sup> (DRESSEL; FARID, 2018, p. 1, tradução nossa). A pesquisa tinha por objetivo comparar o “falso-positivo” (o algoritmo prevê que a pessoa irá reincidir, mas ela não reincide) e o “falso-negativo” (o algoritmo prevê que a pessoa não irá reincidir, mas ela reincide) indicado pelos participantes leigos no assunto com os índices do COMPAS (DRESSEL; FARID, 2018).

Dressel e Farid selecionaram 1000 casos julgados no estado americano da Flórida entre 2013 e 2014 e forneceram uma breve descrição do acusado e do tipo de crime cometido a participantes entre 18 e 74 anos com nível educacional que variava entre não ter concluído o ensino médio e ter formação profissional. A descrição abordava sete características, ou seja, os humanos receberam uma quantidade significativamente menor de informação se comparada com a quantidade fornecida ao algoritmo (que trabalha com 137 características). Após analisar as informações recebidas, os participantes deveriam responder se achavam que a pessoa da descrição cometeria um novo crime nos próximos dois anos ou não. A pesquisa foi dividida em duas fases: primeiro, omitindo a raça das descrições e, depois, informando-a.

Na comparação com os grupos de humanos, o algoritmo teve a pior performance. Enquanto a média geral de acerto do COMPAS é de 65,2%, humanos leigos informados da raça da pessoa analisada acertaram em média 66,5% e humanos leigos não informados da raça, 67% dos casos. A média de falsos-negativos (previsão de não reincidência) para brancos indicada por humanos não informados da raça do indivíduo (40%) é, aliás, 7% menor que a do COMPAS (47%) (DRESSEL; FARID, 2018).

---

<sup>43</sup> No original, “*algorithms are any better than untrained humans at predicting recidivism in a fair and accurate way*”.

Os resultados da pesquisa mostraram que, apesar de dispor de uma quantidade significativamente inferior de informação, humanos “leigos” com relação às questões da justiça criminal conseguem obter desempenho semelhante ao algoritmo que tem acesso às respostas de 137 questões. Os pesquisadores concluíram, por isso, que, “coletivamente, esses resultados lançam dúvidas significativas sobre todo o esforço de previsão algorítmica de reincidência”<sup>44</sup> (DRESSEL; FARID, 2018, p. 1, tradução nossa).

Além da comparação entre o desempenho de humanos e algoritmos, outra importante preocupação surge no contexto da crescente implementação de mecanismos de IA em processos decisórios: a fundamentação. Tanto na análise dos participantes da pesquisa de Dressel e Farid quanto na análise do COMPAS, as informações de *input* eram conhecidas. A grande diferença entre eles é que os humanos poderiam, de certa forma, “explicar” porque respondiam sim ou não à pergunta ou, em qualquer caso, associar a conclusão ao elemento da vontade humana. Isso quer dizer que eventual conclusão enviesada ou, na comparação com o COMPAS, racista, que um humano apresentasse, não apenas pode ser mais facilmente identificada, como também permitiria imputar responsabilidade ao humano envolvido. O algoritmo, de outro lado, não é capaz de dar a mesma justificativa, nem de exercer algo parecido com a vontade humana.

O exemplo do algoritmo do COMPAS permite ligar os conceitos antes trabalhados que estão relacionados à IA com aspectos práticos que deverão ser considerados no âmbito da responsabilidade civil. A partir de diversas informações, processadas na sequência de instruções do algoritmo (*input*), o COMPAS chega ao *output*, que pode ser uma pontuação favorável ou totalmente desfavorável à pessoa submetida a análise de risco *evidence-based*. Mas compreender como o algoritmo chegou a essa conclusão pode ser uma tarefa difícil, que envolve desde entraves jurídicos, como eventual necessidade de ponderação entre o direito de acessar o algoritmo e direitos de propriedade intelectual e sigilo comercial, até entraves matemáticos, notadamente quando o caso envolver *machine learning* e *deep learning*, em que os próprios desenvolvedores poderão desconhecer qual é o algoritmo.

No caso do COMPAS, a discussão foi centralizada no entrave jurídico, qual seja, a existência de direitos de propriedade intelectual e sigilo comercial protegendo o algoritmo, que acabou não sendo divulgado. Ocorre que sem conhecer o algoritmo, é praticamente impossível compreender como a pontuação atribuída pode ser justificada. Com efeito, Loomis conhecia as informações que forneceu ao *software*, mas, ainda assim, não foi possível entender como cada dado, individualmente, influenciou na pontuação indicada pelo relatório.

---

<sup>44</sup> No original, “[c]ollectively, these results cast significant doubt on the entire effort of algorithmic recidivism prediction”.

Embora o guia de uso disponibilizado pela Northpoite informe que o COMPAS considera teorias criminológicas e lista diferentes fatores que são considerados no *input*, como cada uma dessas informações afeta o *output* ainda é algo desconhecido. E como os dados considerados dizem respeito tanto ao indivíduo quanto a sua família e amigos, além de indicadores educacionais e de moradia, sequer há como dizer, com certeza, o que pode ter influenciado mais ou menos no resultado.

A “falha” do algoritmo exposta pela ProPublica é ainda mais grave porque, para além da incógnita quanto aos fatores de maior ou menor peso na análise algorítmica, a apuração dos resultados demonstrou que o COMPAS beneficia significativo percentual de brancos com falsos-negativos enquanto prejudica outro significativo contingente de negros que recebem falsos-positivos. E essa “falha” não apenas deixa de ser explicada pelas teorias criminológicas do guia de uso, como também contraria dados públicos de criminalidade. Conforme informações levantadas no *Uniform Crime Reporting* do *Federal Bureau of Investigation* americano, 69,6% das prisões realizadas em 2016<sup>45</sup> foram de pessoas brancas, contra 26,9% de pessoas negras (FEDERAL BUREAU OF INVESTIGATION, 2016). Considerando essa informação, o resultado do COMPAS deveria ser exatamente o contrário.

Além da dificuldade de compreender a fundamentação, há um segundo fator relevante: o tipo de informação analisado. Em artigo voltado à questão da utilização de IA como árbitros, Maxi Scherer apontou a tendência à adoção, pela IA, de um comportamento conservador. Isso porque, em algoritmos de *machine learning* ou de *deep learning*, as informações consideradas para resolver determinado problema (*input*) estarão baseadas no passado, isto é, em dados de anos anteriores e nas experiências prévias que a IA participou, contudo, a interpretação de situações não se resume apenas ao que já se conhece. No caso da atividade jurisdicional, por exemplo, a interpretação de situações de acordo com a boa-fé exige do intérprete noções dificilmente traduzidas de forma objetiva para a leitura de uma máquina, assim como uma visão *prospectiva* de como determinada decisão poderá afetar casos futuros. Scherer lembra que:

[...] na medida em que as teorias jurídicas enfatizam o significado político do direito, bem como o fato de que os tomadores de decisão têm liberdade para 'preencher' padrões gerais, como 'boa-fé', surge a questão de como essas considerações políticas ou morais seriam gerenciadas em um modelo de IA. Quem ou o que estaria em posição de influenciar essas considerações políticas ou morais? Em um modelo de computador tradicional, pode-se apontar para o programador. No entanto, [...] em modelos avançados de IA, o algoritmo não é codificado por um programador, mas extraído de dados observáveis. Portanto, a única base para a decisão, mesmo em questões morais ou politicamente sensíveis, serão os dados anteriores. Como já apontado acima, os

---

<sup>45</sup> Utilizou-se este ano como base para se permitir uma comparação mais próxima com os dados das pesquisas da ProPublica e de Dressel e Farid, que também consideraram o ano de 2016.

modelos de IA são, assim, propensos a assumir uma abordagem conservadora, mesmo em um contexto de aprendizado de máquina de algoritmos cada vez melhores.<sup>46</sup> (SCHERER, 2019, p. 571, tradução nossa)

Ou seja, deixar que algoritmos resolvam questões afetas ao convívio social e, inclusive, político, contribui mais para a perpetuação de determinado modelo do que para a sua inovação. Um exemplo de como isso pode afetar as relações privadas (e a responsabilidade civil) pode ser eventual responsabilização do desenvolvedor de ferramentas *evidence-based* como o COMPAS perante pessoas que eventualmente demonstrarem que o resultado do algoritmo lhes prejudicou injustamente ou, também, a seleção de candidatos a determinada vaga de emprego capitaneada exclusivamente por um sistema de IA. Nesse segundo exemplo, a depender da vaga, a IA pode considerar características comuns dos ocupantes anteriores (gênero, faixa etária, raça, formação acadêmica) e deixar de selecionar alguém apto para a posição apenas porque não se encaixa no “padrão” identificado pelo algoritmo. E isso, de fato, aconteceu.

Na década de 1970, o Hospital St. George, em Londres, decidiu automatizar o processo seletivo para vagas de emprego, colocando um computador para filtrar os currículos recebidos em uma primeira etapa, reduzindo a quantidade de currículos que deveria ser analisada pelos humanos na segunda etapa (O’NEILL, 2016, p. 115–117). A ideia tinha dois objetivos: tornar a seleção mais eficiente e, também, mais justa, porque o computador não teria os mesmos preconceitos dos humanos. Como isso se daria? Pelo *input* na máquina do expressivo volume de dados já analisado pelo hospital em anos anteriores. O resultado foi desastroso. Um número significativo de currículos de médicos estrangeiros foi descartado por pequenos erros gramaticais e negros e mulheres foram severamente prejudicados na seleção e, em 1988, a Comissão de Igualdade Racial do governo britânico entendeu que a política de contratação do hospital era discriminatória, tanto em termos de gênero como de raça (O’NEILL, 2016, p. 117).

O caso confirma o que aponta Scherer: “quaisquer modelos de computador baseados em dados são tão bons quanto os dados de entrada”<sup>47</sup> (SCHERER, 2019, p. 559, tradução nossa). Conseqüentemente, há o risco de que as informações fornecidas a algoritmos utilizados em processos decisórios estejam “infectadas” com preconceitos humanos, de modo que um

---

<sup>46</sup> No original, “[...] to the extent legal theories emphasize the political significance of law, as well as the fact that decision-makers have discretion to ‘fill in’ general standards, such as ‘good faith’, the question arises how these political or moral considerations would be managed in a AI model. Who or what would be in a position to influence those political or moral considerations? In a traditional computer model, one might point towards the programmer. However, [...] in advanced AI models, the algorithm is not coded by a programmer but extracted from observable data. Therefore, the only basis for the decision, even on morally or politically sensitive issues, will be past data. As already pointed out above, AI models are thus likely to take conservative approach, even in a machine learning context of ever-improving algorithms”.

<sup>47</sup> No original, “[a]ny data-based computer models are only as good as the input data”.

algoritmo de *machine learning* poderia não apenas reproduzir os preconceitos que lhe foram apresentados, como inclusive acentuá-los em seus resultados no processo de aprendizado. Ou seja, mesmo que o algoritmo não tenha sido desenvolvido com viés, o conjunto de dados que ele processa o apresentará essa tendência.

Não seria, portanto, um “erro” do algoritmo, mas um erro de dados de *input*. Daí porque, em tempos de crescente inserção de algoritmos em processos decisórios, sejam eles autônomos ou validados posteriormente por seres humanos, a qualidade e o tipo de informação processados pelo algoritmo serão de fundamental importância para os resultados – assim como será quem seleciona, registra ou se utiliza de tais informações, especialmente porque são essas pessoas que poderão explicar os *outputs* questionados e, eventualmente, responder por falhas da IA.

Como já indicou Alan Turing (1950), haverá, por certo, determinado controle por parte de quem criou o algoritmo e inseriu as informações, mas o processo de aprendizado implicará, ao final, em significativo desconhecimento do que ocorre dentro da máquina por seu “professor”. Assim, ainda que se possa prever o *output*, ele será em grande parte desconhecido – o que já acontece com o COMPAS. Esse problema pode não ser tão grave enquanto os resultados da IA necessariamente precisarão de validação humana, em que seria possível identificar responsáveis (embora seja pertinente entender quanto o humano pôde influenciar a máquina). A questão com relação à IA autônoma é, todavia, bastante diferente.

As respostas fornecidas pela IA podem ou não estar associadas aos humanos envolvidos no processo. O que se buscou neste capítulo foi, por isso, demonstrar algumas das questões que afetam a análise de determinado resultado relacionado à IA: de onde são obtidos os dados analisados, se um algoritmo poderia, eventualmente, permanecer inacessível às partes (em razão de direitos de propriedade intelectual ou sigilo comercial), a possibilidade de *outputs* enviesados ou demasiadamente conservadores, além da precisão com que algoritmos podem fornecer resultados em comparação com humanos – e até que ponto esses resultados podem ser independentes das informações fornecidas ao algoritmo ou consolidadas anteriormente pela sociedade. Todos esses fatores podem fundamentar alguma forma de responsabilidade civil, especialmente a objetiva, e devem ser considerados.

A complexidade de se determinar quem é o responsável direto pelo resultado da IA (se o programador, o “professor”, o usuário ou o próprio sistema) pode confundir as possíveis soluções normativas existentes. Daí porque definir qual será a resposta do ordenamento jurídico é de fundamental importância para assegurar que a automatização não torne impunes condutas que, se realizadas por humanos, ensejariam a imputação do dever de indenizar.

### **3 LACUNAS NO DIREITO E AS PROPOSTAS DE REGULAMENTAÇÃO DA INTELIGÊNCIA ARTIFICIAL NO PARLAMENTO BRASILEIRO E EUROPEU**

No capítulo anterior, buscou-se conceituar a inteligência artificial autônoma e explicar como ela pode agir de maneira independente, a partir do aprendizado de máquina. Em casos tais, explicar e compreender uma decisão automatizada pode ser uma tarefa bastante difícil, o que afeta, por consequência, a identificação de responsáveis por eventual falha da IA (se humanos ou se o próprio algoritmo).

Dentre as principais discussões afetas à relação entre responsabilidade civil e IA, pode-se referir a dificuldade de se definir: *(i.)* a quem se deve imputar o dever de indenizar (ao humano ou à IA?); *(ii.)* como se justificaria essa imputação; *(iii.)* como seria resolvida a responsabilidade patrimonial; e *(iv.)* que regime de responsabilidade civil seria aplicado.

Em 16 de fevereiro de 2017, o Parlamento Europeu aprovou Resolução com recomendações à Comissão sobre disposições de Direito Civil sobre Robótica, destacando que “tendo em conta a fase alcançada no desenvolvimento da robótica e da IA, é apropriado começar por problemas de responsabilidade civil” (UNIÃO EUROPEIA, 2017). O documento é bastante amplo e aborda desde parâmetros éticos básicos para o desenvolvimento da IA até o regime eventualmente aplicado em caso de análise de responsabilidade civil.

No Brasil, igualmente há projetos de lei em trâmite no Congresso Nacional tratando do desenvolvimento e utilização da IA. Mas, diferentemente da Resolução do Parlamento Europeu antes mencionada, os projetos brasileiros preocupam-se especialmente na regulamentação geral, sem tratar especificamente sobre responsabilidade civil.

Para compreender o alcance dessas propostas e a efetiva necessidade de uma alteração legislativa especificamente para a inteligência artificial autônoma, este capítulo tem como objetivo introduzir o conceito de lacunas jurídicas (item 3.1) para, então, mapear as principais sugestões dos projetos de lei acima referidos para o tratamento jurídico da IA no âmbito da responsabilidade civil (itens 3.2 e 3.3).

#### **3.1 As Lacunas no Direito**

A rigor, a inserção lógica de novas regras no ordenamento jurídico ocorre no intuito de preencher determinadas lacunas que estariam, em tese, impedindo a solução de problemas concretos. Porém, nem toda lacuna demanda a edição específica de uma nova lei. Cabe

entender, por isso, qual tipo de lacuna está presente em determinado caso, razão pela qual convém distinguir três tipos de lacunas: a normativa, a axiológica e a de reconhecimento (MANERO, 2015, p. 48).

Há lacuna normativa quando inexistente uma regra que correlacione determinado caso relevante para o sistema jurídico com uma solução normativa, isto é, quando determinada hipótese fática deveria estar regulada, mas não está (MANERO, 2015, p. 49). Juan Ruiz Manero comenta que a constatação desse tipo de lacuna depende da ponderação dos princípios *prima facie* aplicáveis ao caso concreto, de modo que, se a conclusão alcançada pela ponderação de tais princípios for a de que uma regra não é necessária, então não há falar em lacuna a ser preenchida (2015, p. 50).

Uma particularidade deve ser apontada, todavia. A ausência de regra no âmbito do direito penal ou do direito administrativo sancionador implica na conclusão de que a conduta analisada é permitida. Já no direito civil, essa conclusão é apenas *prima facie*, porque só haverá uma permissão definitiva se, da ponderação dos princípios relevantes para o caso, o intérprete não encontrar uma regra proibitiva (MANERO, 2015, p. 51–52). Um bom exemplo de como isso ocorre pode se dar pelo art. 187 do Código Civil (BRASIL, 2002), em que determinada conduta que, *a priori*, é lícita, ao final, poderá ser reputada ilícita se contrariar a boa-fé objetiva, o fim econômico ou social ou os bons costumes (CACHAPUZ, 2017, p. 117).

Na lacuna axiológica, diversamente, a regra existe e se aplica, *a priori*, à hipótese fática, mas a solução por ela fornecida desconsidera uma particularidade que ensejaria o afastamento da consequência jurídica prevista, demandando uma solução normativa distinta (MANERO, 2015, p. 53). Aqui, há um papel relevante dos princípios. Em sua crítica ao positivismo de Herbert L. Hart, Ronald Dworkin apontou que, diferentemente das regras, que têm uma dimensão de “tudo ou nada” (2002, p. 39), os princípios têm uma dimensão de peso e ganham concretude a partir da sua aplicação a determinado caso, podendo justificar, por vezes, o afastamento de uma regra (2002, p. 40–41). Dworkin exemplifica a questão a partir do caso *Riggs v. Palmer*, em que a Suprema Corte Americana decidiu que o neto que assassinou o avô não poderia receber a herança a ele destinada em testamento, pois, embora a lei reconhecesse o direito à herança, era preciso atentar ao princípio de que ninguém pode se beneficiar dos próprios ilícitos (2002, p. 37). Assim, a lacuna axiológica é identificada em concreto e decorre da própria regra, sendo corrigida pelo intérprete.

Por fim, a lacuna de reconhecimento se verifica diante de uma indeterminação semântica que pode se originar (*i.*) da inexistência de usos linguísticos firmes ou (*ii.*) de uma falta de clareza sobre a finalidade ou o propósito da regra, isto é, sobre qual estado de coisas ela visa

proteger (MANERO, 2015, p. 55). A consequência da lacuna de reconhecimento é a dúvida sobre o alcance de uma regra diante de determinada hipótese fática. Essa lacuna decorre do que Herbert L. A. Hart chamou de “textura aberta” do direito, pois “[e]m todos os campos da experiência, e não só no das regras, há um limite, inerente à natureza da linguagem, quanto à orientação que a linguagem geral pode oferecer” (2011, p. 139). Hart menciona, a exemplo, que uma regra que proíbe a circulação de veículos no parque certamente impede a entrada de automóveis, mas pode gerar dúvidas quanto ao enquadramento de bicicletas e patinetes no conceito de “veículos” para a finalidade da regra proibitiva (2011, p. 139).

A existência de diferentes tipos de lacunas no ordenamento jurídico pode sinalizar a necessidade de diferentes soluções frente ao caso concreto, que podem ser encontradas pelo intérprete de acordo com os princípios ou motivos justificadores de determinada regra. Com efeito, o art. 4º da Lei de Introdução às Normas do Direito Brasileiro (BRASIL, 1942) autoriza que, diante de uma lacuna, se decida o caso de acordo com a “analogia, os costumes e os princípios gerais do direito”. Ou seja, nem toda lacuna demanda a introdução de uma nova regra no ordenamento jurídico. Contudo, a depender dos questionamentos que determinada lacuna provoca, a edição de novas regras pode ser, efetivamente, necessária.

A análise que se seguirá ao longo deste trabalho tem por finalidade apontar as possíveis conjecturas sobre a aplicabilidade das regras de responsabilidade civil previstas no Código Civil brasileiro aos casos envolvendo inteligência artificial. Por vezes, pode ser que a dúvida resida apenas em uma lacuna de reconhecimento, que dependerá fundamentalmente de um exercício interpretativo no significado e alcance semântico de determinada palavra. Em outros casos, pode ser que se esteja diante de um conflito entre a aplicabilidade de uma regra e outros princípios do Direito brasileiro, sendo necessária uma solução distinta. Mas é também possível que se conclua que casos de IA são relevantes e não têm solução normativa suficiente, e, então, a discussão demandaria a edição de novas regras. Antes de verificar se essas lacunas efetivamente existem, serão analisadas as propostas na mesa dos parlamentos brasileiro e europeu sobre o assunto.

### **3.2 Os Projetos de Lei no Parlamento Brasileiro**

Até junho de 2020, foram identificados quatro projetos de lei relacionados à inteligência artificial em trâmite no Congresso Nacional, dois no Senado Federal (PLS 5051/2019 e PLS 5961/2019) e dois na Câmara dos Deputados (PL 21/2020 e PL 240/2020). Todos buscam, ao cabo, regular o “uso da IA” e estabelecer princípios e diretrizes gerais para o seu

desenvolvimento ético, sem tratar de um assunto jurídico específico. De todos, apenas o PL 21/2020 (o mais completo sobre o assunto) se preocupa em definir, efetivamente, o que é IA para, então, regulamentá-la<sup>48</sup>.

Outro ponto comum entre as propostas legislativas brasileiras é a preocupação com o protagonismo humano na relação com a IA. Os projetos não detalham qual será o tratamento para o caso da IA autônoma (ou mesmo se o uso desse tipo de IA poderia ter alguma restrição no Brasil), mas ainda que esse tipo de tecnologia venha a ser implementada, há preocupação com a decisão final ser tomada ou sempre observada por um humano. Nesse sentido, o art. 4º do PLS 5051/2019 prevê que “sistemas decisórios baseados em Inteligência Artificial serão, sempre, auxiliares à tomada de decisão humana”, e o art. 9º, inciso IV, do PL 21/2020, prevê que agentes de inteligência artificial responsáveis por determinado sistema de IA devem “encerrar o sistema se o seu controle humano não for mais possível”, além de responder pelas decisões tomadas pelo sistema de IA (art. 9º, inciso V).

Daí porque, quando há menção à questão da responsabilidade por atos da IA, a imputação recairá sobre o supervisor da IA ou o “agente de inteligência artificial”<sup>49</sup> (caso do PL 21/2020). Os “agentes de inteligência artificial” podem ser agentes de desenvolvimento ou operação, a depender do tipo de atividade desenvolvida. Nesse sentido, a redação do art. 2º, inciso IV, do PL 21/2020 define:

- IV - agentes de inteligência artificial: pessoas físicas ou jurídicas, de direito público ou privado, e entes sem personalidade jurídica, assim considerados:
- a) agentes de desenvolvimento: todos aqueles que participam das fases de planejamento e design, coleta e processamento de dados e construção de modelo; de verificação e validação; ou de implantação do sistema de inteligência artificial; e
  - b) agente de operação: todos aqueles que participam da fase de monitoramento e operação do sistema de inteligência artificial. (BRASIL, 2020).

Talvez porque o protagonismo humano é o ponto central das propostas legislativas brasileiras, nenhuma delas trata da possibilidade de criar um estatuto jurídico para que a IA responda pelos próprios atos. Pela definição de um “agente de inteligência artificial”, a

<sup>48</sup> Nesse sentido, dispõe o art. 2º, “I - sistema de inteligência artificial: o sistema baseado em processo computacional que pode, para um determinado conjunto de objetivos definidos pelo homem, fazer previsões e recomendações ou tomar decisões que influenciam ambientes reais ou virtuais” (BRASIL, 2020).

<sup>49</sup> O § 2º do art. 4º do PLS 5051/2019 assim dispõe: “§ 2º A responsabilidade civil por danos decorrentes da utilização de sistemas de Inteligência Artificial será de seu supervisor”. A previsão constante do PL 240/2020 é bastante similar: “Art. 4º As soluções, programas e projetos da Inteligência Artificial devem atender: (...) IV – os Robôs e equipamentos que utilizam Inteligência Artificial devem se submeter aos seres humanos e serem operados por responsáveis técnicos e empresas que responderão por todos os resultados negativos à sociedade”. Já o PL 21/2020 determina o seguinte em seu art. 6º, que trata dos “princípios para o uso responsável de inteligência artificial”: “VI - responsabilização e prestação de contas: demonstração, pelos agentes de inteligência artificial, do cumprimento das normas de inteligência artificial e da adoção de medidas eficazes para o bom funcionamento dos sistemas, observadas suas funções”.

imputação é resolvida com a pessoa natural, jurídica ou a entidade não personificada que eventualmente tenha criado, programado ou supervisionado a operação da IA. Conseqüentemente, é o patrimônio dessas pessoas que sofrerá afetação (hipótese dos artigos 391<sup>50</sup> e 942<sup>51</sup> do Código Civil). Além da definição da imputação, o art. 6º, inciso VI, do PL 21/2020 inverte o ônus probatório, atribuindo-o aos agentes de IA o dever de demonstrar ter cumprido as normas de IA e adotado medidas eficazes para o seu bom funcionamento (uma espécie de sistema baseado na culpa presumida, sem qualquer referência à possibilidade de enquadramento na responsabilidade objetiva, por exemplo).

Por derradeiro, nenhum dos projetos analisados aborda expressamente regras afetas ao regime de responsabilidade civil. Das disposições em discussão, cabe perquirir se a IA seria considerada atividade de risco, um produto submetido ao regime das relações de consumo ou uma coisa submetida à guarda do proprietário, à semelhança do que ocorre com automóveis, a ensejar a responsabilidade objetiva dos envolvidos, ou se, a depender do caso e da posição do humano envolvido, haverá tratamento diferenciado, aplicando-se a responsabilidade subjetiva.

Portanto, ainda que se identifique uma lacuna em termos de responsabilidade civil relacionada à IA autônoma, os projetos de lei que hoje tramitam no Congresso Nacional pouco ou nada contribuem para solucionar o problema. Pelo contrário, a definição de padrões genéricos para o desenvolvimento da IA poderá dar ensejo a uma ampla margem de interpretação das situações concretas, que poderão concluir ora pela culpa presumida de agentes de IA, ora pela responsabilidade subjetiva ou, então, uma solução puramente objetiva, sendo necessário, assim, se pensar em novos projetos que contenham regras capazes de enfrentar, efetivamente, o problema posto.

### **3.3 A Resolução do Parlamento Europeu**

Ao contrário das propostas brasileiras, a União Europeia vem discutindo há alguns anos o tratamento destinado à inteligência artificial, inclusive a autônoma. A principal Resolução editada pelo Parlamento Europeu sobre o tema da responsabilidade civil e a IA é a de 16 de fevereiro de 2017, com recomendações à Comissão sobre disposições de Direito Civil sobre Robótica. Essa Resolução introduziu duas diferentes e possíveis respostas para eventuais danos

---

<sup>50</sup> Código Civil, art. 391: “Pelo inadimplemento das obrigações respondem todos os bens do devedor” (BRASIL, 2002).

<sup>51</sup> Código Civil, art. 942: “Os bens do responsável pela ofensa ou violação do direito de outrem ficam sujeitos à reparação do dano causado; e, se a ofensa tiver mais de um autor, todos responderão solidariamente pela reparação. Parágrafo único. São solidariamente responsáveis com os autores os co-autores e as pessoas designadas no art. 932” (BRASIL, 2002).

causados pela IA: tratar a IA como um produto ou serviço, de modo que o responsável seria seu fabricante, programador, proprietário ou utilizador; ou criar um estatuto jurídico próprio para a IA, isto é, uma “pessoa eletrônica”, “a casos em que os robôs tomam decisões autônomas ou em que interagem por qualquer outro modo com terceiros de forma independente” (UNIÃO EUROPEIA, 2017).

A sugestão acerca da criação de uma “pessoa eletrônica” é um dos pontos mais polêmicos da Resolução e permite questionar, por exemplo, se a pessoa eletrônica serviria única e exclusivamente para resolver o problema da imputação da responsabilidade civil quando o dano está vinculado a uma ação da IA autônoma ou se também passaria a ter direitos que, originalmente, são reconhecidos apenas aos seres humanos. Nesse caso, poder-se-ia questionar se a IA deve agir de modo a preservar a própria sobrevivência, bem como se a sua sobrevivência seria um valor a ser preservado no ordenamento jurídico (TEPEDINO; SILVA, 2019, p. 79).

Quanto à responsabilidade patrimonial, no caso da imputação ao fabricante, ao programador/operador, ao proprietário ou ao utilizador, o patrimônio afetado pelo pagamento de eventual indenização já encontra resposta no ordenamento jurídico, pois tanto as pessoas naturais como as pessoas jurídicas podem adquirir e possuir patrimônio próprio capaz de arcar com eventuais obrigações ou, então, contratar um seguro. A crítica, neste ponto, reside no fato de que a Resolução não sugeriu uma definição clara do que é um “operador” ou um “fabricante” da IA, o que pode suscitar dúvidas na interpretação sobre quem deverá ser responsabilizado a depender da situação concreta (ALABART, 2018, p. 71).

A sugestão de reconhecimento de personalidade jurídica à IA, todavia, não permite chegar a mesma conclusão sobre a responsabilidade patrimonial referida no parágrafo anterior.

Primeiro, porque, considerando o fundamento para a criação de uma pessoa eletrônica, esta não se equipara à pessoa natural ou jurídica. O que fundamenta a autonomia da pessoa eletrônica na recomendação do Parlamento Europeu é o fato de que a IA poderá tomar decisões de modo autônomo e totalmente independente dos humanos, admitindo-se, nesse caso, que o próprio sistema seja responsabilizado em caso de falhas. Trata-se de fundamento diverso daquele relacionados à personalidade da pessoa natural, que é o centro do ordenamento jurídico, e da pessoa jurídica, a qual, conforme comenta Mafalda Miranda Barbosa, “é atribuída em função de determinados interesses das pessoas que estão na base da sua constituição” (BARBOSA, 2017, p. 1486). Ao cabo, a pessoa jurídica terá uma pessoa natural como fundadora/sócia que tomará as decisões em seu nome; o mesmo não aconteceria com a IA.

Da diferença de fundamento para a personalidade da IA em comparação à personalidade da pessoa natural e jurídica decorre uma segunda preocupação relevante em matéria de

responsabilidade civil: a existência de patrimônio. Isso porque atribuir personalidade eletrônica à IA somente faria sentido se isso assegurar que eventual indenização será paga. Nesse sentido, tanto o ordenamento jurídico brasileiro como o europeu reconhecem direitos e deveres à pessoa natural e à pessoa jurídica, além da autonomia para celebrar negócios jurídicos e estabelecer obrigações entre si. Assim, quando uma pessoa natural ou jurídica é responsabilizada por determinado fato, é possível que o seu patrimônio seja afetado para o pagamento de indenização. No caso da IA, para além da personalidade, assegurar o adimplemento de eventual obrigação de indenizar demandaria o reconhecimento da possibilidade de uma pessoa eletrônica adquirir patrimônio em nome próprio, ou, então, estar vinculada a determinado conjunto de bens.

Nesse ponto, a Resolução de 16 de fevereiro de 2017 traz duas sugestões: *(i.)* a criação de um regime de seguros obrigatórios, à semelhança do que já ocorre com automóveis, que poderia ou não ser separado por categorias de robôs; ou *(ii.)* a criação de um fundo de compensação. Essas soluções não se limitariam a resolver a questão da responsabilidade patrimonial da pessoa eletrônica, viabilizando, também, a limitação da responsabilidade do fabricante, do programador, do proprietário e do utilizador, que poderiam contratar conjuntamente um seguro para garantir eventual indenização por danos causados pelo robô, ou, também, contribuir para um fundo de compensação voltado a esse propósito.

Além disso, a Resolução recomenda que o fundo de compensação não seja apenas uma garantia subsidiária, no caso de o seguro não ter sido contratado ou se mostrar insuficiente<sup>52</sup>. Também se sugere deliberar se o fundo seria geral<sup>53</sup>, para todos os robôs autônomos inteligentes, ou individual, vinculado ao robô a partir de um número de registro único, bem como se a contribuição seria pontual ou periódica (nesse caso, perdurando durante o tempo de vida do robô).

Por fim, quanto ao regime de responsabilidade, o Parlamento Europeu recomendou a aplicação da responsabilidade objetiva (princípio 54<sup>54</sup>) ou da gestão de riscos (princípio 55<sup>55</sup>).

---

<sup>52</sup> Princípio 59 (b) “Garantir que os fundos de compensação não sirvam apenas para garantir uma compensação no caso de os danos causados por um robô não serem abrangidos por um seguro” (UNIÃO EUROPEIA, 2017).

<sup>53</sup> Princípio 59 (d) “Decidir quanto à criação de um fundo geral para todos os robôs autônomos inteligentes ou quanto à criação de um fundo individual para toda e qualquer categoria de robôs e quanto à contribuição que deve ser paga a título de taxa pontual no momento em que se coloca o robô no mercado ou quanto ao pagamento de contribuições periódicas durante o tempo de vida do robô” (UNIÃO EUROPEIA, 2017).

<sup>54</sup> Princípio 54: “Observa, ao mesmo tempo, que a responsabilidade objetiva exige apenas a prova de que o dano ocorreu e o estabelecimento de um nexo de causalidade entre o funcionamento prejudicial do robô e os danos sofridos pela parte lesada” (UNIÃO EUROPEIA, 2017).

<sup>55</sup> Princípio 55: “Observa que a abordagem da gestão de riscos não se concentra na pessoa «que atuou de forma negligente» como individualmente responsável, mas na pessoa que é capaz, em determinadas circunstâncias, de minimizar os riscos e de lidar com os impactos negativos” (UNIÃO EUROPEIA, 2017).

A responsabilidade objetiva implicaria na adoção de um mecanismo de responsabilização similar ao dos produtos defeituosos, devendo-se provar o dano, o funcionamento incorreto do robô/sistema de IA (defeito) e o nexo causal entre ambos (ALABART, 2018, p. 66-67). Já a gestão de riscos envolve uma análise mais econômica do que propriamente jurídica. Silvia Díaz Alabart (2018, p. 69-70) aponta que a gestão de riscos pode ocorrer de modo preventivo ou, então, quando o risco se mostrar real, momento em que serão adotadas medidas para mitigar as consequências dele decorrentes. Esse segundo sistema, segundo a autora, se basearia mais no interesse das empresas que lidam com tecnologia do que no interesse dos consumidores.

Henrique Sousa Antunes comenta que tanto a abordagem pela responsabilidade “objetiva” quanto pela “gestão de riscos” demandam uma interpretação que exclui a responsabilidade subjetiva, de modo que a distinção sugerida pelo Parlamento Europeu estaria equivocada (2019, p. 150). Antunes admite, de todo modo, que a gestão de riscos permitiria “repensar a causalidade e, de forma concomitante, a identificação dos sujeitos responsáveis pelos danos”, o que teria especial impacto na interpretação da sugestão para a adoção de um seguro de responsabilidade civil (2019, p. 150).

Silvia Díaz Alabart, de outro lado, aponta que os dois regimes são compatíveis:

Além disso, tendo em conta a confluência de dois tipos de interesses distintos (o dos empresários e o dos consumidores), parece aconselhável utilizar ambas as soluções. Por um lado, a adoção de controle dos riscos derivados da fabricação e colocação no mercado de robôs com as características mencionadas no início deste trabalho é essencial para as empresas. Por outro lado, é prioritário garantir aos proprietários ou simples usuários dos robôs a indenização pelos danos que venham a sofrer ao utilizá-los, ou mesmo também às pessoas que, sem serem usuários ou donos dos robôs, sofrem esses danos pela circunstância de se encontrarem próximas do robô no momento do dano (comumente chamadas de *bystanter* ou “espectador”).<sup>56</sup> (2018, p. 70, tradução nossa)

Apesar de poucos pontos em comum entre os projetos brasileiros e a Resolução europeia, pode-se extrair duas possíveis perspectivas relacionadas à imputação da responsabilidade civil relacionada à IA autônoma. A primeira, da IA como objeto do direito, cujas ações serão imputadas sempre ao humano envolvido na tarefa (podendo responder, nesse caso, a pessoa natural, a pessoa jurídica ou entidades sem personalidade jurídica reconhecidas no direito brasileiro). A segunda perspectiva é, indubitavelmente, a mais inovadora e, por certo,

<sup>56</sup> No original: “*Es más, tomando en cuenta la confluencia de dos tipos de interés distintos (el de los empresarios y el de los consumidores), parece aconsejable la utilización de ambas soluciones. Por una parte, la adopción del control de los riesgos derivados de la fabricación y puesta en el mercado de robots con las características mencionadas al inicio de este trabajo, resulta esencial para las empresas. Por otra parte, es prioritario garantizar a los propietarios o simples usuarios de los robots, la indemnización por los daños que puedan sufrir al utilizarlos, o incluso también a las personas que sin ser usuarios o propietarios de los robots sufren esos daños por la circunstancia de encontrarse cerca del robot en el momento del daño (el denominado habitualmente bystander o ‘espectador’).*”

a mais polêmica: tratar a IA como um sujeito, atribuindo-lhe, para tanto, um estatuto jurídico próprio e estabelecendo-se regras aptas a fornecer respostas sobre a responsabilidade patrimonial, regime de responsabilidade e questões relevantes para a análise da imputação.

Cada perspectiva – e as críticas já feitas com relação a cada uma delas – será trabalhada nos capítulos 5 e 6. Antes disso, é importante retomar conceitos fundamentais da responsabilidade civil, bem como os seus pressupostos, estabelecendo-se as premissas para a análise que se seguirá.

## 4 NOTAS ELEMENTARES SOBRE RESPONSABILIDADE CIVIL

A responsabilidade civil tem como objetivo precípua assegurar a reparação de danos que afetam interesses particulares, seja *in natura* ou mediante uma indenização pecuniária. Esse traço que lhe é particular a difere, essencialmente, da responsabilidade penal, em que o interesse tutelado que dá origem à imposição da pena é de toda a sociedade. Além disso, por tratar de interesses privados, a obrigação de indenizar pode surgir no âmbito de um negócio jurídico (responsabilidade negocial) ou da violação a um dever geral (responsabilidade extra negocial).

Embora, hoje, tais conceitos sejam de fácil visualização, o seu desenvolvimento decorre de um longo processo, cabendo, aqui, um breve resgate histórico antes de adentrar na análise dos temas discutidos acerca da responsabilidade civil que considera a IA um objeto ou sujeito titular de direitos subjetivos. Assim, propõe-se, neste capítulo, a retomada da origem da responsabilidade civil, com a distinção entre responsabilidade negocial e extra negocial, subjetiva e objetiva, assim como a apresentação de seus pressupostos. Feito isso, será possível analisar, mais adiante, se as sugestões já implementadas ou em discussão em outras jurisdições fazem, efetivamente, sentido no contexto brasileiro.

### 4.1 A origem da responsabilidade civil

Caio Mário da Silva Pereira comenta, ao tratar da responsabilidade civil, a dificuldade em conceituá-la. Seja pela teoria que fundamenta a imputação do dever de indenizar, seja pelo aspecto social que circunda a sua função, “é difícil conter em uma frase concisa a amplitude da responsabilidade civil” (PEREIRA, 2018b). Nada obstante, ele nos fornece uma noção suficiente para delimitar o objeto de estudo:

A responsabilidade civil consiste na efetivação da reparabilidade abstrata do dano em relação a um sujeito passivo da relação jurídica que se forma. Reparação e sujeito passivo compõem o binômio da responsabilidade civil, que então se enuncia como o princípio que subordina a reparação à sua incidência na pessoa do causador do dano. Não importa se o fundamento é a culpa, ou se é independente desta. Em qualquer circunstância, onde houver a subordinação de um sujeito passivo à determinação de um dever de ressarcimento, aí estará a responsabilidade civil. (PEREIRA, 2018b)

A responsabilidade civil nem sempre teve existência autônoma e a possibilidade de reparação de danos direcionada ao próprio sujeito lesado tampouco foi desde cedo reconhecida. Inicialmente, inexistia distinção entre a responsabilidade civil e a criminal, e a noção de reparação do dano estava predominantemente atrelada à função punitiva e à vingança privada.

Embora não tenham construído uma efetiva teoria da responsabilidade civil, é do Direito romano que se extraem os principais subsídios para a formulação moderna do instituto. Conforme refere Moreira Alves, os juriconsultos romanos distinguiram os ilícitos em delitos públicos e privados. “Delito público é a violação de norma jurídica que o Estado considera de relevante importância social”, como “a *perduellio* (atentado contra a segurança do Estado), o *parricidium* (assassinio do homem livre)” (MOREIRA ALVES, 2019). Nesses casos, era imposta a *poena publica*, que poderia ser a morte do ofensor, castigos corporais ou o pagamento de uma multa ao Estado.

Delitos privados eram ofensas à pessoa ou aos bens dos indivíduos. Nesse caso, “a iniciativa da ação penal era entregue à própria vítima, que poderia ingressar com uma *actio* para buscar a imposição de *poena privata*” (SANSEVERINO, 2010, p. 20). Moreira Alves (2019) explica que, no direito clássico, a pena privada e a pública compartilhavam do mesmo caráter punitivo; já no direito justiniano, a pena privada passa a ter uma característica de ressarcimento. Os romanos conceberam quatro delitos privados: o *furtum* (furto), a *rapina* (roubo), a *iniuria* (ofensa à integridade física ou moral) e o *damnum iniuria datum* (o dano à coisa alheia, animada ou inanimada) (MOREIRA ALVES, 2019).

Dos delitos privados, interessam especialmente a *iniuria* e o *damnum iniuria datum*, que surge com a *Lex Aquilia*, um plebiscito proposto por volta do ano 286 a. C. A *Lex Aquilia* foi um verdadeiro marco para a responsabilidade civil. Conforme aponta Cario Mário:

Onde se realiza a maior revolução nos conceitos jus-romanísticos em termos de responsabilidade civil é com a *Lex Aquilia*, de data incerta, mas que se prende aos tempos da República. Tão grande revolução que a ela se prende a denominação de *aquiliana* para designar-se a responsabilidade extracontratual em oposição à contratual. Foi um marco tão acentuado, que a ela se atribui a origem do elemento “culpa”, como fundamental na reparação do dano. A *Lex Aquilia*, bem assim a consequente *actio ex lege Aquilia*, tem sido destacada pelos romanistas e pelos civilistas, em matéria atinente à responsabilidade civil. (2018B)

No mesmo sentido, observa Maria Helena Diniz:

A *Lex Aquilia de damno* veio a cristalizar a ideia de reparação pecuniária do dano, impondo que o patrimônio do lesante suportasse o ônus da reparação, em razão do valor da *res*, esboçando-se a noção de culpa como fundamento da responsabilidade, de tal sorte que o agente se isentaria de qualquer responsabilidade se tivesse procedido sem culpa. Passou-se a atribuir o dano à conduta culposa do agente. (1990, p. 10)

A *Lex Aquilia* era dividida em três capítulos: (i.) morte do escravo ou animal de rebanho alheio, em que seria devida indenização no máximo valor do escravo ou do animal no ano anterior à morte; (ii.) fraude contra credores, gerando a obrigação de pagar o valor da dívida; e (iii.) dano à coisa alheia animada ou inanimada, gerando a obrigação de pagar ao dono o valor

máximo da coisa nos 30 dias anteriores ao evento danoso (SANSEVERINO, 2010, p. 21). A segunda hipótese logo caiu em desuso, mas as duas outras são particularmente importantes.

Moreira Alves aponta três requisitos para a configuração do *damnum iniuria datum*: a *iniuria* (ato contrário ao direito), a *culpa* (ato praticado com dolo ou culpa em sentido estrito) e o *damnum* (a lesão à coisa). Posteriormente, o conceito de *damnum* foi alargado, abrangendo outras hipóteses, e o lesado pleiteava a reparação do dano pela chamada *actio legis Aquiliae*.

Nesse sentido, duas são as contribuições mais fundamentais da *Lex Aquilia* em matéria de responsabilidade civil: as bases para o surgimento de uma obrigação de indenizar independentemente da existência de uma relação negocial prévia e a introdução de um elemento subjetivo (culpa em sentido amplo), que afastou o regime coletivo e objetivista que imperava até então, essencialmente vinculado à vingança privada e à índole penal (COSTA, 2001, p. 480).

Quanto à culpa, Floriana M. Cursi aponta que há duas possíveis visões sobre o significado da *iniuria* na *Lex Aquilia*: (i.) uma base puramente objetiva de ilicitude, que somente veio a ser acrescida pelos elementos da culpa e do dolo posteriormente; e (ii.) uma base subjetiva, em que a *iniuria* expressa a *vontade* de causar dano ao proprietário da coisa (2011, p. 18). Cursi entende que o elemento subjetivo era intrínseco à *iniuria*, e não uma aquisição posterior (2011, p. 19). Segundo ela, os romanos não concebiam uma conduta ilícita (*iniuria*) sem um elemento subjetivo intrínseco (*dolus/culpa*) (CURSI, 2011, p. 20). De todo modo, Cursi conclui que foi o surgimento da culpa que permitiu aplicar a *Lex Aquilia* a um maior número de casos<sup>57</sup> (2011, p. 25) sendo este um elemento fundamental para a responsabilidade decorrente daquela lei.

Caio Mário (2018b) pontua que há divergência entre romanistas a respeito da necessidade de comprovação da culpa para que se caracterizasse o dever de reparar o dano na *Lex Aquilia*. Nada obstante, o autor menciona que a necessidade de culpa ou dolo (culpa em sentido amplo, portanto) foi expressamente referida nas Institutas de Gaio, destacando a parte final do texto, da qual se conclui “que ‘o elemento subjetivo da culpa’ foi introduzido, ‘contra o objetivismo do direito primitivo’” (PEREIRA, 2018b).

Silvio Rodrigues aponta, em sentido semelhante, que “na *Lex Aquilia*, embora a culpa levíssima incriminasse o agente causador do dano, *in Lex Aquilia et levíssima culpa venit*, certamente a inexistência de culpa era causa de exoneração da responsabilidade” (1975, p. 154).

---

<sup>57</sup> No mesmo sentido, Sanseverino observa que “[a] ‘culpa’ foi introduzida posteriormente pelos juristas romanos, reelaborando a *fattispecie* prevista na *Lex Aquilia* e permitindo gradualmente a ampliação das moralidades de danos ressarcíveis com a inclusão de condutas não tipificadas” (2010, p. 22).

Portanto, o termo “responsabilidade aquiliana”, hoje indiscriminadamente utilizado pela doutrina como referência à responsabilidade extra negocial, decorre da *Lex Aquilia* e, em sua origem, remonta à responsabilidade extra negocial *subjetiva*, não abarcando a responsabilidade objetiva.

Apesar dos avanços na matéria, os Romanos não chegaram a distinguir com precisão a responsabilidade civil da criminal, e tampouco conceberam o ato ilícito como uma figura autônoma, tal como ocorre no direito moderno (MOREIRA ALVES, 2019). Conforme mencionado acima, a palavra “delito” servia tanto para a ação de natureza pública quanto a privada – daí porque, ainda hoje, é comum que se refira à responsabilidade civil extra negocial também pelo termo “delitual”.

Com a queda do Império Romano do Ocidente e as invasões bárbaras, Sanseverino comenta que houve um retorno à noção primitiva de vingança privada, ideia superada apenas no final da Idade Média, quando começam a aparecer os princípios da responsabilidade civil moderna (2010, p. 23). É na Idade Média que se estabelecem: (i.) o “princípio de que todo dano deve ser reparado”; (ii.) “lenta separação entre os delitos penais e civis com a distinção entre pena e reparação”; e (iii.) “critérios mais objetivos para distinguir as modalidades de reparação (natural ou pecuniária) e para definir critérios de cálculo da indenização” (SANSEVERINO, 2010, p. 23).

O papel fundamental da culpa, contribuição da *Lex Aquilia*, veio a ser reconhecido em definitivo na idade moderna quando positivado no art. 1.382 do Código Civil de Napoleão, em 1804. O Código Civil francês, vale pontuar, teve grande influência perante o Código Civil brasileiro de 1916 e, embora tenha instituído a regra geral de responsabilidade fundada na culpa, igualmente previu hipóteses em que o elemento subjetivo seria dispensável para a caracterização do dever de indenizar, conforme será visto adiante.

## **4.2 Responsabilidade negocial e extra negocial**

A distinção entre a responsabilidade negocial e a extra negocial está associada à origem do ato ilícito praticado. Conforme refere Jorge Cesa Ferreira da Silva, a expressão “ato ilícito” pode assumir diferentes significados. Em um sentido amplo, o ilícito se refere a qualquer conduta contrária ao direito e, nessa linha:

[...] se costuma distinguir entre os *ilícitos absolutos* - aqueles que podem ser cometidos por qualquer um, pois representam ataques a direitos subjetivos absolutos (propriedade, direitos de personalidade) - dos *ilícitos relativos* - que representam o descumprimento de norma advinda de negócio jurídico e, portanto, limitam-se à

conduta de apenas uma e determinada pessoa, o devedor. (FERREIRA DA SILVA, 2007, p. 107)

No sentido restrito, ato ilícito diz respeito à violação de direitos absolutos, apenas. É dessa distinção entre o significado de ilícitos absolutos e relativos, na concepção ampla do termo, que se separa a responsabilidade civil em dois regimes: negocial/contratual e aquiliana/extra negocial (CAVALIERI, 2012, p. 16).

Mario Julio de Almeida Costa assim define a responsabilidade negocial ou contratual:

*A responsabilidade contratual resulta da violação de um direito de crédito ou obrigação em sentido técnico. Verificamos, portanto, que o qualificativo corrente não se mostra rigoroso, dado que, além dos contratos, existem outras fontes de tais vínculos, cujo incumprimento ocasiona essa espécie de responsabilidade civil. Podem eles, do mesmo modo, surgir de negócios jurídicos unilaterais e, inclusive, directamente da lei. Mercê da razão exposta, alguns autores preferem chamar-lhe *responsabilidade negocial* ou *responsabilidade obrigacional*. (COSTA, 2011, p. 539).*

A responsabilidade civil extra negocial, de outro lado, decorre “da violação de deveres ou vínculos jurídicos gerais, isto é, de deveres de conduta impostos a todas as pessoas e que correspondem aos direitos absolutos, ou até da prática de certos actos que, embora lícitos, produzem dano a outrem” (COSTA, 2011, p. 540). Em outras palavras, enquanto a responsabilidade negocial exsurge de um vínculo obrigacional pré-existente, a responsabilidade extra negocial tem origem na violação a um dever legal e aplicável a qualquer pessoa, independentemente de haver um pacto anterior.

Pelo cuidado terminológico, é pertinente o comentário de Jorge Cesa Ferreira da Silva a respeito da distinção da origem da responsabilidade civil, à luz do quanto já exposto acima no item 4.1:

Muito embora a designação corrente, decorre a primeira da ofensa de direito decorrente de negócio jurídico, unilateral ou bilateral. O carácter contratual, portanto, é contingente, razão pela qual a sua melhor designação é a de “responsabilidade negocial”, como muitas vezes se encontra na doutrina. Por sua vez, a responsabilidade extracontratual (*rectius*, extranegocial) é conhecida como responsabilidade “por ato ilícito” ou “aquiliana”, haja vista decorrer dos desenvolvimentos realizados a partir da *Lex Aquilia*. As duas designações, apesar de consagradas, igualmente portam problemas. A responsabilidade “aquiliana” em sentido estrito não englobaria os casos de responsabilidade objetiva, ao passo que a “responsabilidade por ato ilícito” não envolveria aquela decorrente de ato lícito. De qualquer sorte, esses títulos se consagraram, razão pela qual se faz importante, mais do que uma absoluta precisão terminológica, apenas ter a dimensão dos problemas de cada uma, para não se cair em falso juízo. (FERREIRA DA SILVA, 2007, p. 59)

Conforme refere Sérgio Cavalieri Filho, “adeptos da teoria unitária, ou monista, criticam essa dicotomia, por entenderem que pouco importam os aspectos sobre os quais se apresente a responsabilidade civil no cenário jurídico, já que os seus efeitos são uniformes” (2012, p. 16).

Não obstante, a distinção foi acolhida pelo Código Civil de 2002, na linha do que já constava do diploma civil anterior; assim, as regras de responsabilidade civil estão divididas em dois grandes grupos de artigos: um grupo trata da responsabilidade negocial (artigos 389 a 420), enquanto o outro grupo é dedicado à responsabilidade extra negocial (artigos 927 a 954).

Embora se encontrem disposições comuns a ambos os regimes, essa separação ganha importância na análise concreta da responsabilidade civil. A exemplo, pode-se mencionar a análise da culpa<sup>58</sup>. O ônus de demonstrar o elemento subjetivo incumbe, de regra, à vítima, quando se tratar de responsabilidade civil extra negocial. Na responsabilidade negocial, de outro lado, a culpa poderá ser presumida (quando se tratar de obrigação de resultado, já que a obrigação de meio igualmente impõe ao lesado o ônus de demonstrar a culpa do devedor<sup>59</sup>). A distinção da origem da responsabilidade igualmente pode afetar a quantificação do dano (seja pela possibilidade de inclusão de cláusula de limitação de responsabilidade no contrato ou a sua prévia liquidação), assim como o termo inicial para a constituição em mora do devedor<sup>60</sup> (FERREIRA DA SILVA, 2007, p. 61–63).

<sup>58</sup> Conforme explica Silvio Rodrigues (1975, p. 9), ao referir-se à questão da prova, “na responsabilidade contratual, demonstrado pelo credor que a prestação foi descumprida, o *onus probandi* se transfere para o devedor inadimplente, que terá que evidenciar a inexistência de culpa de sua parte, ou a presença de força maior, ou outra excludente da responsabilidade capaz de eximi-lo do dever de indenizar”.

<sup>59</sup> A prova da culpa é, com efeito, um dos aspectos mencionados pela doutrina para a relevância da distinção entre obrigações de meio e de resultado. Nesse sentido: “A importância prática da distinção está no ônus da prova do dano sempre que a culpa for condição da responsabilidade: nas obrigações de resultado, presume-se a culpa sempre que o resultado não é alcançado. Já nas obrigações de meio, não há presunção de culpa, é o credor quem deve provar a culpa” (MARTINS-COSTA, 2009, p. 192–193). Fábio Konder Comparato sustenta que essa distinção também se aplica à responsabilidade extracontratual: “Para a distinção que ora nos ocupa, a referência a esse dever geral nas obrigações extracontratuais é importante, pois é ele que determina afinal se a obrigação em causa é de meios ou de resultado, sendo fundamental que se identifique, no caso, a pré-existência de um dever geral por parte do autor do dano ou não” (COMPARATO, 2010). Assim, “o dever imposto a todos os membros da coletividade pode referir-se a um resultado objetivo do agir humano — ‘não matar’, ‘respeitar a integridade física de outrem’, ‘não destruir os bens alheios’ — ou ao contrário dizer respeito à própria maneira de agir: ‘não prejudicar a segurança, o sossego e a saúde dos moradores de prédio vizinho com o mau uso de sua propriedade’, ou ‘exercer concorrência desleal no comércio ou na indústria’”. Enquanto no primeiro caso o resultado objetivo já dá ensejo ao dever de indenizar, no segundo, a parte lesada teria o ônus de demonstrar o prejuízo e a culpa do agente. Mas deve-se pontuar que o artigo de Comparato foi originalmente publicado em 1967, na vigência do Código Civil de 1916, quando a legislação civil ainda não dispunha de norma similar ao art. 187 do Código Civil de 2002, que, conforme refere Maria Cláudia Cachapuz, permite reconhecer uma situação de ilicitude diante da afronta à *relação de confiança* estabelecida entre privados a partir da boa-fé, de dados empíricos da tradição a qual estão vinculadas as partes e do fim econômico ou social envolvido, de maneira independente não apenas da culpa, como também da ocorrência concreta de um dano (CACHAPUZ, 2017). Nesse sentido, a distinção entre obrigações de meio e de resultado na responsabilidade civil extra negocial para se definir o ônus probatório da culpa parece não mais ter o mesmo sentido apontado por Comparato, se consideradas as hipóteses de ilicitude do atual Código Civil e a consequente origem da obrigação de indenizar, especialmente porque o reconhecimento de uma situação de ilicitude a partir do art. 187 prescinde da prova de culpa do agente e, portanto, o ônus probatório da vítima restringe-se à demonstração do dano e do nexo de causalidade (para a tutela reparatória).

<sup>60</sup> Coube à doutrina esclarecer que a regra do art. 398 do Código Civil (“nas obrigações provenientes de ato ilícito, considera-se o devedor em mora, desde que o praticou”) se aplica apenas às hipóteses de responsabilidade civil aquiliana, e não à responsabilidade civil negocial, regrada pelo art. 397 do Código Civil, já que o termo “ilícito” referido no artigo é no seu sentido restrito, isto é, com relação ao ilícito absoluto. Assim, conforme leciona Caio

Diante de tais distinções e considerando que esta pesquisa se propõe a analisar se há uma lacuna no ordenamento jurídico para resolver os casos de responsabilidade civil extra negocial, isto é, quando o uso de algoritmos/IA violar um “dever geral” (ilícito absoluto, portanto), serão confrontadas, especificamente, o segundo grupo de regras de responsabilidade civil, isto é, o conjunto dos artigos 927 a 954 do Código Civil.

### 4.3 Responsabilidade objetiva e subjetiva

Por muito tempo se discutiu se o fundamento da responsabilidade civil seria a culpa ou o risco (PEREIRA, 2018b). Atualmente, admite-se que tanto a culpa como o risco podem dar ensejo à caracterização do dever de reparar o dano. Todavia, é interessante desmembrar como se deu a evolução da responsabilidade civil que era fundamentalmente baseada na culpa, para um regime de culpa presumida e, então, o reconhecimento de hipóteses em que o dever de indenizar decorre, objetivamente, do risco.

Conforme mencionado anteriormente, a *culpa* (em sentido amplo) é uma contribuição da *Lex Aquilia* do Direito romano, que se tornou fundamental para a responsabilidade civil moderna a partir do Código Civil Napoleônico de 1804, cujo art. 1.382<sup>61</sup> previa a obrigação de reparar o dano sempre que este fosse causado por culpa. Regra semelhante foi reproduzida no art. 159 do Código Civil de 1916 e, novamente, no art. 186 do Código Civil de 2002.

Caio Mário (2018b) sintetiza que “a teoria da responsabilidade subjetiva erige em pressuposto da obrigação de indenizar, ou de reparar o dano, *o comportamento culposo do agente*, ou simplesmente a sua culpa, abrangendo no seu contexto a *culpa propriamente dita e o dolo do agente*”<sup>62</sup>. Nos termos do art. 186, aliás, a culpa está intrinsecamente associada ao conceito de ilicitude, ao ponto de se considerar como pressupostos da responsabilidade civil decorrente desta regra o dano, a culpa do agente e o nexo de causalidade entre ambos.

Porém, como também observa Caio Mário (2018b), o regime da culpa “é hoje insuficiente para abranger todo o plano da reparação”, o que, de todo modo, não justifica a sua

---

Mário da Silva Pereira, “Considera-se que o art. 398 do Código Civil de 2002, somente é aplicável às hipóteses de responsabilidade civil extracontratual subjetiva, não incidindo nos casos de ilícito contratual e de responsabilidade civil objetiva” (PEREIRA, 2018a, p. 301).

<sup>61</sup> No original, “*Tout fait quelconque de l’homme, qui cause à autrui un dommage, oblige celui par faute duquel il est arrivé, à le réparer*” (FRANÇA, 1804).

<sup>62</sup> Caio Mário (2018b) também comenta que “A distinção entre o *delito*, cujo pressuposto é o dolo do agente, e o *quase delito*, que deste se abstém adotando como fato gerador a *culpa*, desapareceu na moderna concepção da responsabilidade. Atualmente (dentro na doutrina subjetiva) concentra-se na ideia do ato ilícito. Não que o direito desconheça e despreze o comportamento doloso. Mas, se o dolo pode agravar a condição do agente, não constitui fator etiológico da responsabilidade civil a antiga dicotomização do fato danoso em delitual e quase delitual. O que importa é a conduta do agente.”

eliminação. A principal dificuldade para a manutenção de um sistema de responsabilidade civil exclusivamente baseado na culpa decorre do fato de que, na responsabilidade extra negocial, incumbe ao lesado o ônus de demonstrar o elemento subjetivo na conduta do agente. E essa prova, em muitos casos, poderia ser demasiadamente difícil, a ponto de impedir que se alcançasse a reparação do dano.

Um dos meios técnicos desenvolvidos pela doutrina para superar o problema da prova da culpa foi a teoria da culpa presumida, bastante presente no Código Civil Português, por exemplo, e com diversos dispositivos também no Código Civil de 1916, a exemplo do regime de responsabilidade indireta dos artigos 1.521 a 1.523. Caio Mário (2018b) comenta, com o exemplo da doutrina e jurisprudência francesas, a interpretação do art. 1.384 do Código Napoleônico de 1804, com a inversão do ônus da prova e a dispensa ao lesado de demonstrar a culpa na conduta do agente a quem se imputava o dever de indenizar. Ainda assim, a culpa permanecia como pressuposto da responsabilidade. O autor refere, nesse sentido:

A presunção de culpa podia variar de intensidade. Ou assumia as características de uma *praesumptio iuris et de iure*, quando a lei a estabelecia em termos peremptórios, fazendo crer que abraçava sem rebuços a doutrina objetiva, o que ocorria em alguns casos já examinados de responsabilidade pelo fato das coisas (v. *supra*, capítulo VIII); ora permanecia no campo da presunção *iuris tantum*, partindo da indicação de um responsável cuja culpa era presumida, porém admitindo que podia ele trazer elementos de prova convincentes de sua ausência de culpa. (PEREIRA, 2018b)

Assim, se afastada a presunção da culpa pelo agente, o dano restaria sem reparação.

O movimento de objetivação da responsabilidade civil teve início na França, com Saleilles e Josserrand, e, no Brasil, especialmente com as obras de Alvino Lima e José de Aguiar Dias (PEREIRA, 2018b). Na responsabilidade objetiva, o elemento subjetivo assume papel totalmente irrelevante, porque serão pressupostos da responsabilidade tão somente o nexo de causalidade entre o dano e a conduta do agente (GAGLIANO; PAMPLONA FILHO, 2018, p. 66).

O atual Código Civil aderiu a um modelo dual de responsabilidade civil (GAGLIANO, PAMPLONA FILHO, 2018, p. 68). Há, no art. 186, a adoção da teoria da responsabilidade subjetiva, fundada no elemento da culpa em sentido amplo, assim como há, no art. 927, parágrafo único, a consagração da teoria do risco criado e a possibilidade de responsabilidade objetiva, independentemente de culpa, para os casos previstos na lei. A seguir, veremos como essa distinção de regimes conversa com os pressupostos da responsabilidade civil, notadamente o fator de imputação, para, no capítulo 5, adentrar na análise específica das hipóteses de responsabilidade subjetiva e objetiva previstas no Código Civil de 2002.

#### 4.4 Pressupostos da responsabilidade civil no direito brasileiro

Pressuposto, nas palavras de Judith Martins-Costa, é “o elemento estrutural cuja existência há de estar comprovada para que o instituto desempenhe os fins a que está voltado” (MARTINS-COSTA, 2020, p. 395). No caso da responsabilidade civil, independentemente da origem (negocial ou aquiliana), há, essencialmente, quatro pressupostos fundamentais: (i.) a existência de um dano indenizável; (ii.) um ato ilícito ou antijurídico; (iii.) o nexo de causalidade entre o dano e o ato ilícito; e (iv.) um fator de imputação do dever de indenizar (FERREIRA DA SILVA, 2007, p. 52).

A ideia de dano está relacionada a uma perda suportada por alguém (material ou não). Nem toda perda será, contudo, indenizável: somente deverá ser indenizado o dano considerado “injusto”, isto é, o dano que é certo (e não hipotético), direto (que decorre imediatamente de um evento danoso) e imputável (à alguém, o que implica na análise do sujeito a quem se imputa o dever de indenizar e do fator que dá ensejo à imputação) (FERREIRA DA SILVA, 2007, p. 152–153). Além disso, há que se considerar o legítimo interesse de quem sofreu o dano para dimensioná-lo, ainda que nos limites da imputação (MARTINS-COSTA, 2020, p. 407). Por fim, os danos indenizáveis podem ser patrimoniais (mensurados em pecúnia) ou extrapatrimoniais (comumente referidos como danos “morais”).

O segundo pressuposto da responsabilidade civil é a existência de um ato ilícito ou antijurídico<sup>63</sup>, isto é, um ato contrário ao direito<sup>64</sup>. Conforme aponta Judith Martins-Costa, há, basicamente, três formas de ilicitude:

(a) a violação de direito alheio; (b) a afronta a lei (que, direta ou reflexamente, tutela interesses de alguém ou da comunidade), então se utilizando o termo legalidade; e (c) o exercício de direito próprio, de modo manifestamente contrário à boa-fé, aos bons costumes e ao fim econômico e social do direito. (MARTINS-COSTA, 2020, p. 397)

Tais hipóteses são tratadas nos artigos 186 e 187 do Código Civil. O art. 186 pressupõe a existência de um dano e da culpa do agente, porque depende de uma ação ou omissão voluntária, negligente ou imprudente. Diz-se, por isso, que se trata de hipótese de ilicitude subjetiva. O art. 187, de outro lado, trata de hipótese de ilicitude objetiva, porque prescinde da existência de culpa ou de dano e permite identificar situações de ilicitude a partir de conceitos

<sup>63</sup> Embora tanto a “ilicitude” quanto a “antijuridicidade” refiram-se a um ato contrário ao direito, há uma distinção terminológica decorrente da confusão entre o ato ilícito e a culpa. Nesse sentido, o ato ilícito estaria também associado à culpa *lato sensu* do agente (ilicitude subjetiva), ao passo que a antijuridicidade é a mera existência do ato contrário ao direito, independentemente da culpa (MARTINS-COSTA, 2020, p. 397).

<sup>64</sup> Admite-se, todavia, hipóteses expressamente previstas na legislação em que haverá obrigação de indenizar pela prática de atos reputados *lícitos*. São as hipóteses dos artigos 188, I e II, 1.285, 1.286, 1.289, 1.293, 1.311, parágrafo único, e 1.385, § 3º, todos do Código Civil.

indeterminados e externos à conduta humana, em contrariedade a um fim econômico ou social, à boa-fé ou aos bons costumes (CACHAPUZ, 2017). Conforme Caio Mário da Silva Pereira<sup>65</sup> e Sérgio Cavalieri Filho<sup>66</sup>, o art. 187 teria consagrado no direito brasileiro a teoria do abuso de direito, segundo a qual “abusa de seu direito aquele que leva o seu exercício ao extremo de convertê-lo em prejuízo para outrem sem vantagem para si mesmo” (PEREIRA, 2017, p. 189). A verificação do abuso de direito prescinde da análise da intenção do sujeito, bastando a verificação do exercício excessivo de um direito que lhe é reconhecido pela ordem jurídica.

Maria Cláudia Cachapuz sustenta, de outro lado, que o art. 187 representa verdadeira “cláusula geral capaz de contribuir à concreção de direitos subjetivos” (CACHAPUZ, 2017, p. 115), oferecendo elementos fundamentais para o exercício de ponderação na interpretação das normas jurídicas. Não se trataria, assim, de disciplinar especificamente o abuso de direito<sup>67</sup>, pois o art. 187:

Trata-se de norma que, por meio de uma cláusula geral de restrição de liberdades – e mais, de uma efetiva *cláusula de subsidiariedade jurídico-fundamental*, na melhor aplicação da conformação descrita por Schwabe – (...) exige o reconhecimento da existência de um vínculo de confiança entre os sujeitos de direito: a ideia de que é necessário confiar-se na autoridade da lei moral como forma de qualificar-se a liberdade humana. E, nessa medida, restringi-la apenas quando – e na devida proporção – se faça necessário. (CACHAPUZ, 2017, p. 118–119)

Ou seja, a relevância do art. 187 está justamente na possibilidade de se verificar situações de ilicitude e, a partir delas, restringir determinado direito subjetivo apenas quando este estiver em situação de colisão com outro direito igualmente reconhecido pela ordem

---

<sup>65</sup> Nos comentários do jurista aos Código Civil de 2002, atualizados por Cristiano de Souza Zanetti e Leonardo de Campos Melo, assim se refere: “Consagrando a doutrina do abuso do direito, coloca-se o Código na situação contrária à tese de PLANIOL, para quem a noção mesma de abuso é incompatível com a de direito, por conter uma contradição nos próprios termos (*contradictio in adjecto*). O art. 187 oferece os extremos da caracterização pelo abuso do direito, assentando que o exercício dele há de ser limitado” (PEREIRA, 2017, p. 189).

<sup>66</sup> Afirma o doutrinador: “O art. 187 conceitua o abuso do direito nos seguintes termos: ‘Também comete ato ilícito o titular de um direito que, ao exercê-lo, excede manifestamente os limites impostos pelo seu fim econômico ou social, pela boa-fé ou pelos bons costumes.’ Como se vê, o abuso do direito está aqui definido como ato ilícito, diferente daquele conceituado no art. 186, e quem praticá-lo ficará também obrigado a indenizar pela norma do art. 927” (CAVALIERI, 2012, p. 171).

<sup>67</sup> Maria Cláudia Cachapuz refere, nesse sentido, o exemplo do art. 334 do Código Civil Português, que inspirou a regra brasileira, no qual o abuso de direito está relacionado ao regime das invalidades, o que não se aplicaria necessariamente no caso do Código Civil de 2002. No exemplo do direito contratual, “[p]assa-se pela determinação, não de invalidação, mas de construção de hipótese de ilicitude, pela combinação dos arts. 187 e 423 do CC brasileiro, com a restrição, numa medida ótima, daquilo que afronta hipótese de boa-fé, fim econômico ou social (...) na cláusula contratual discutida em concreto” (CACHAPUZ, 2017, p. 116–117). No mesmo sentido, Judith Martins-Costa sugeriu que “Com os olhos voltados às potencialidades do art. 187 para a configuração desse “homem normativo”, ousemos, pois, unindo ética e técnica civilista, impedir que a denominação prevaleça sobre a significação da figura: afastemos o exercício jurídico das trilhas subjetivistas encerradas tradicionalmente no *nomem iuris* “abuso do direito” para, recheando com renovado conteúdo a concha do marisco abandonada, redirecioná-lo a uma qualificação normativamente polarizada pela conduta segundo a boa-fé, os bons costumes e o fim econômico e social no exercício dos direitos, faculdades, posições e situações jurídicas subjetivas.” (MARTINS-COSTA, 2006, p. 38)

jurídica. E justamente por disciplinar hipóteses de colisão, oferecendo elementos interpretativos para a ponderação no caso concreto e a aplicação da teoria externa das restrições a direitos fundamentais (CACHAPUZ, 2017, p. 119), o art. 187 permite reconhecer a ilicitude também quando não estiver caracterizado o dano (dando azo à tutela contra o ilícito, portanto, e não apenas à tutela contra o dano).

A última observação que merece destaque quanto ao ato ilícito diz respeito às hipóteses de exclusão da ilicitude, previstas no art. 188 do Código Civil: a legítima defesa e o estado de necessidade. Isso porque, se constatado que o agente agiu sob o pálio de uma dessas excludentes de ilicitude, não há falar em responsabilidade civil. No caso da IA autônoma, pode-se questionar, certamente, o alcance de tais excludentes, especialmente se os parâmetros éticos de desenvolvimento da IA forem pautados nas Leis de Asimov, em que a sobrevivência do robô somente deve ser preservada se não entrar em conflito com a Primeira e a Segunda Leis, que priorizam a vida humana. Nesse contexto, supondo-se que haja uma situação em que a IA autônoma, a quem se reconheça personalidade eletrônica, se envolva em um acidente com humanos: deverá a pessoa eletrônica indenizar porque optou por preservar a própria vida em vez de evitar danos a humanos? Lembra-se que “magoar” um humano, segundo as Leis de Asimov, é também mais importante que a vida do robô, de modo que tal acidente não precisaria causar danos graves ou a morte dos humanos...

O terceiro pressuposto para a responsabilidade civil é o nexó de causalidade, isto é, a conexão que deve ser estabelecida entre o dano e o ato ilícito ou antijurídico. O nexó causal não se confunde com a *causa* do dano: para fins de responsabilização civil, deve-se identificar uma relação *necessária* entre o ato ilícito e o dano (MARTINS-COSTA, 2020). Trata-se da teoria do dano direto e imediato, reconhecida como teoria da interrupção do nexó causal pelo Supremo Tribunal Federal quando do julgamento do RE 130.764, de relatoria do Ministro Moreira Alves, que assim explicou:

Ora, em nosso sistema jurídico, como resulta do disposto no artigo 1.060 do Código Civil, a teoria adotada quanto ao nexó de causalidade é a teoria do dano direto e imediato, também denominada teoria da interrupção do nexó causal. Não obstante aquele dispositivo da codificação civil diga respeito à impropriamente denominada responsabilidade contratual, aplica-se ele também a responsabilidade extracontratual, inclusive a objetiva, até por ser aquela que, sem quaisquer considerações de ordem subjetiva, afasta os inconvenientes das outras duas teorias existentes: a da equivalência das condições e a da causalidade adequada. [...] Essa teoria, como bem demonstra AGOSTINHO ALVIM [...], só admite o nexó de causalidade quando o dano é efeito necessário de uma causa, o que abarca o dano direto e imediato sempre, e, por vezes, o dano indireto e remoto, quando, para a produção deste, não haja concausa sucessiva. Daí, dizer AGOSTINHO ALVIM (l.c.): ‘os danos indiretos ou remotos não se excluem, só por isso; em regra, não são indenizáveis, porque deixam de ser efeito

necessário, pelo aparecimento de concausas. Suposto não existam estas, aqueles danos são indenizáveis”. (BRASIL, 1992, p. 25)

A regra do art. 1.060 do Código Civil de 1916 foi reproduzida no art. 403 do Código Civil de 2002 e, muito embora tal disposição permaneça no grupo de regras que tratam da responsabilidade civil negocial, pacificou-se o entendimento de que a teoria da interrupção do nexa causal é igualmente aplicável à responsabilidade extra negocial (SCHREIBER, 2018, p. 636).

Sérgio Cavaliere Filho identifica a teoria do dano direto e imediato à teoria da causalidade adequada. Refere o doutrinador, nesse sentido, que:

[...] por causa direta, imediata, necessária ou adequada deve-se entender como sendo aquela que revela um liame de necessariedade entre a causa e o efeito e não de simples proximidade temporal ou espacial. Próxima ou remota, imediata ou mediata, a causa será adequada quando o evento danoso for efeito necessário de determinado acontecimento. O exame do nexa causal limita-se a verificar se a atividade desenvolvida pelo agente vincula-se de algum modo - próximo, direto, necessário, adequado ou eficiente - ao dano [...]. (CAVALIERI, 2012, p. 55).

Assim, quanto ao pressuposto do nexa causal, há que se identificar, de regra, o fator necessário que vincula determinada conduta ilícita ao evento danoso. Isso pode ser um fator bastante relevante no tratamento da IA, em especial diante do uso de algoritmos de *deep learning*, em que identificar a origem do dano pode ser um verdadeiro desafio, conforme será melhor exposto adiante. Há que se pontuar, por fim, as três situações em que o nexa causal, embora constatado, poderá ser excluído no caso concreto, afastando a responsabilidade do agente: caso fortuito ou força maior, fato de terceiro e fato exclusivo da vítima.

O quarto pressuposto (fator de imputação) se refere ao motivo pelo qual se atribui a alguém o dever de indenizar por determinado dano. “Imputar” é atribuir a alguém um dever; “imputabilidade” é o conjunto de elementos pessoais do agente, que lhe permitem compreender e responder pelas condutas praticadas (CAVALIERI, 2012). É a partir do fator de imputação que se define o regime de responsabilidade civil: se o fator for a culpa (*lato sensu*), então se trata de responsabilidade subjetiva; se o fator for o risco, então será aplicado o regime da responsabilidade objetiva, porquanto o “risco” estará associado não ao estado psicológico da pessoa, e sim ao tipo de atividade praticada, à coisa utilizada, etc.

A doutrina diferencia o risco em diferentes teorias: (i.) *risco-proveito*: imputa-se o dever de indenizar a quem se aproveite da atividade que deu origem ao dano, o que demanda a definição do que significa “proveito”, bem como a prova de que o agente se beneficiaria da atividade, questões que dificultam a aplicação dessa teoria; (ii.) *risco profissional*: haverá dever de indenizar sempre que o dano decorrer da atividade/profissão da vítima, razão pela qual essa

teoria se aplica, de regra, à responsabilidade por acidentes de trabalho; (iii.) *risco excepcional*: que escapa à atividade comum da vítima, como é o caso da exploração de energia nuclear, materiais radioativos, etc.; (iv.) *risco criado*: quem pratica determinada atividade e “cria” determinada situação de perigo deverá ser responsabilizado pelos danos dela decorrentes, mas, diferentemente do que ocorre com a teoria do *risco-proveito*, nesse caso, a imputação independe do fato de o agente obter ou não algum ganho com a atividade; e (v.) *risco integral*: é a teoria mais extrema, porque dispensa inclusive onexo causal entre o ilícito e o dano, restando apenas este último como pressuposto para a responsabilidade civil (CAVALIERI, 2012, p. 153).

A teoria adotada pela cláusula geral de responsabilidade objetiva prevista no art. 927, parágrafo único, do Código Civil, é a do risco criado (CAVALIERI, 2012, p. 186). Com efeito, conforme aponta Caio Mário da Silva Pereira, autor do Projeto do Código de Obrigações de 1965:

Das modalidades de risco, eu me inclino pela subespécie que deu origem à teoria do *risco criado*. Como já mencionei (Capítulo XVIII, *supra*), ao elaborar o Projeto de Código de Obrigações de 1965, defini-me por ela, no que fui seguido pelo Código Civil. Depois de haver o art. 927 deste Código enunciado o dever ressarcitório fundado no conceito subjetivo, seu parágrafo único espousa doutrina do risco criado, a dizer que, independentemente da culpa, e dos casos especificados em lei, haverá obrigação de reparar o dano “quando a atividade normalmente desenvolvida pelo autor do dano implicar, por sua natureza, risco para os direitos de outrem”. (2018b)

Assim, para fins de aplicação desse dispositivo no direito brasileiro, basta, a rigor, que o agente tenha praticado determinada atividade ou criado algo capaz de causar danos, não sendo necessário demonstrar que obteve qualquer vantagem com isso.

Deve-se destacar, todavia, que a responsabilidade civil objetiva poderá decorrer da aplicação do art. 187 do Código Civil, caso em que o fator de imputação não será o risco, e sim o exercício em excesso de determinado direito reconhecido pela ordem jurídica, conforme já mencionado acima.

Delineadas as premissas gerais da responsabilidade civil e seus pressupostos, passa-se a analisar a sua intersecção com a inteligência artificial a partir das duas perspectivas já anunciadas: a IA como objeto, no que serão consideradas as teorias da responsabilidade subjetiva, objetiva e, por fim, indireta ou vicária, e, em seguida, a IA como sujeito, no que será considerada, principalmente, a discussão sobre a atribuição de personalidade jurídica aos produtos de IA.

## 5 PRIMEIRA PERSPECTIVA: INTELIGÊNCIA ARTIFICIAL COMO OBJETO

O principal ponto de distinção entre a abordagem de responsabilidade civil da IA como um objeto ou como um sujeito de direitos está centrada na autonomia da IA. A autonomia da IA decorrerá, fundamentalmente, de processos de aprendizado (*machine learning*), a partir das experiências anteriores e dados processados. Mas mesmo esse processo remeterá, ao fim e ao cabo, à intervenção humana. Conforme pontua Silvia Díaz Alabart (2018, p. 116), um dos desafios da IA é distinguir quais atividades do robô decorrem das informações introduzidas pelo programador e quais atividades são próprias do robô (e independentes da sua intervenção direta). Porém, ainda que nesse segundo caso não se tenha nenhuma atividade humana diretamente relacionada, é fato que a aprendizagem alcançada pelo robô não seria possível sem o algoritmo e os dados introduzidos pelos humanos.

Dessa constatação é que se pode dizer que haverá, em certa medida, uma intervenção humana que justifique a imputação da responsabilidade aos humanos envolvidos e, por isso, a percepção de que robôs são objetos do direito, embora manifestem certo grau de autonomia.

Conforme comenta José Carlos Moreira Alves (2019), o “*objeto do direito subjetivo é aquilo sobre o que incide o poder de seu titular*”. Isso não significa, porém, que o objeto seja sempre uma coisa. Nesse sentido, também pessoas (ou a atuação de uma pessoa) podem ser objeto do direito, como ocorre com o pai que exerce o poder familiar sobre o filho e o credor que pode exigir determinada prestação do devedor.

Todas as hipóteses de tratamento jurídico da IA a partir das atuais teorias de responsabilidade civil partirão da premissa de que a IA é um objeto do direito. E como objeto sobre o qual incide o poder de um humano titular de direitos subjetivos (pessoa física ou jurídica), a IA poderá ser equiparada a uma coisa inanimada, a animais que, embora sencientes, não são dotados da mesma inteligência humana ou, ainda, a pessoas sob a responsabilidade de um terceiro, a quem se imputa o dever de indenizar. Analisaremos cada uma dessas hipóteses a seguir.

### 5.1 Responsabilidade Subjetiva: o desafio da análise da culpa e a insuficiência do regime para os casos envolvendo IA

A primeira regra da responsabilidade civil consta do *caput* do art. 927 do Código Civil: “aquele que, por ato ilícito (arts. 186 e 187), causar dano a outrem, fica obrigado a repará-lo”. Embora faça referência expressa apenas ao ato ilícito e ao dano, desse dispositivo se extraem

três diferentes fundamentos para a responsabilidade civil: (i.) a regra geral da responsabilidade subjetiva (art. 927, *caput*, e art. 186); (ii.) a responsabilidade objetiva decorrente do art. 187 do Código Civil (art. 927, *caput*, e art. 187, portanto); e (iii.) a responsabilidade objetiva por atividade de risco ou fato do serviço (art. 927, parágrafo único). Analisaremos, neste tópico, a primeira hipótese.

“O fundamento maior da responsabilidade civil está na *culpa*”, referia Caio Mário da Silva Pereira (2018a, p. 524). Com efeito, por muito tempo, este foi o elemento central para a responsabilidade civil e permanece sendo a regra geral, não sendo possível aplicar o regramento da responsabilidade objetiva quando não previsto expressamente em lei. A demonstração da culpa (*lato sensu*) envolve, necessariamente, uma análise subjetiva do autor do dano, relacionada a uma conduta omissiva ou comissiva reputada como imprudente, imperita, negligente (*culpa stricto sensu*) ou dolosa.

Embora a culpa também possa ser diferenciada em graus (grave, leve ou levíssima), esse fator não é, de regra, considerado para fins de reconhecimento do dever de indenizar. Exceção está no parágrafo único do art. 944 do Código Civil, que autoriza a redução do montante devido à título de indenização quando verificada “excessiva desproporção entre a gravidade da culpa e o dano”. Sérgio Cavalieri Filho pontua, todavia, que tal regra somente se aplicaria em casos de culpa levíssima em que o dano foi, efetivamente, expressivo, estando restrita aos casos de responsabilidade subjetiva (inclusive de origem negocial), não encontrando aplicação em hipótese de responsabilidade objetiva (2012, p. 40).

De todo modo, a culpa (em qualquer grau que se apresente) ou o dolo são indispensáveis para a imputação do dever de indenizar em se tratando de responsabilidade civil subjetiva, já que este é o elemento principal que a distingue da responsabilidade objetiva. Quando se trata de IA, todavia, a demonstração de culpa pode ser uma tarefa um tanto quanto árdua. Conforme aponta a professora Mafalda Miranda Barbosa:

[...] a centralidade do conceito de culpa que faz com que os sistemas delituais se mostrem insuficientes para lidar com danos causados por entes dotados de inteligência artificial. As características de autonomia e autoaprendizagem de tais entes dificultam o traçar de fronteira entre os danos que resultam de um erro humano e aqueles que são devidos ao próprio algoritmo. O comportamento imprevisível deste, que decide por si como agir, que se desenvolve como resultado de um *deep-learning*, sem controle humano, torna impossível conezionar um eventual dano que possa eclodir com uma conduta negligente do ser humano. (BARBOSA, 2020, p. 284)

Barbosa (2020) destaca, nesse sentido, que mesmo a presunção da culpa pode ser insuficiente e derrubada, já que o humano envolvido poderia demonstrar que não agiu com culpa/dolo, que os danos ocorreriam independentemente da sua participação ou mesmo que

agiu com diligência para evitar o dano<sup>68</sup>. Nesse ponto, vale lembrar o quanto já exposto com relação a algoritmos de *deep learning*, admitidos como verdadeiras “caixas pretas”, em que pouco ou nada se sabe sobre o que está sendo processado e como o algoritmo chega a determinado resultado. Trata-se de tecnologia avançada e surpreendentemente próxima da sociedade, que é empregada em ferramentas de pesquisa (como é o caso do Google) e na personalização de anúncios e conteúdos oferecidos em redes sociais, por exemplo<sup>69</sup>. Ainda que seja possível apontar os desenvolvedores e que os usuários possam, em certa medida, controlar as informações fornecidas ao algoritmo, é praticamente impossível precisar onde ocorreu determinada falha ou quem, intencionalmente ou não, pode ter sido responsável por tanto.

Tais preocupações correspondem ao atual estágio de desenvolvimento da IA. Henrique Sousa Antunes destaca, entretanto, que o regime subjetivo pode não ser um efetivo obstáculo para a responsabilidade de entes dotados de inteligência artificial superior à capacidade humana. Segundo ele:

A opção por uma responsabilidade subjetiva, dependente de um juízo de culpa do lesante, encontraria assento nesse patamar em que se descobrem no robô capacidades superiores às do ser humano. Claro que sempre restará questionar a pertinência de um juízo de culpa se o robô se apresentar despidido da dimensão ética da ação humana. (ANTUNES, 2019, p. 153)

Esse debate será mais pertinente quando dita IA surgir, todavia. Por ora, além da dificuldade de identificação de fatores e pessoas responsáveis por determinado evento relacionado ao uso de IA, Gustavo Tepedino e Rodrigo da Guia Silva indagam se seria necessário analisar o grau de intervenção do humano sobre o sistema; como se poderia adotar critérios diferentes para investigar a culpa de um usuário ou de um desenvolvedor; e, ainda, se seria possível reconhecer a incidência de excludentes de ilicitude como o estado de necessidade e a legítima defesa para os atos resultantes da utilização da IA (TEPEDINO; SILVA, 2019). Tais questões ainda não possuem resposta, mas são certamente pertinentes para a verificação da suficiência das regras atuais sobre responsabilidade civil e para a reflexão sobre a personalidade de entes dotados de inteligência artificial (próximo capítulo).

Não obstante, é justamente pelo alto grau de indefinição relacionado à adoção da culpa como fator de imputação da responsabilidade envolvendo a IA que “prepondera em doutrina o entendimento favorável à aplicação do regime objetivo de responsabilidade, usualmente apontado como mais idôneo à tutela efetiva das vítimas” (TEPEDINO; SILVA, 2019, p. 81).

---

<sup>68</sup> Essa hipótese pode obstar a responsabilização de agentes de inteligência artificial mencionados no PL 21/2020, por exemplo, a despeito da previsão de inversão do ônus da prova, conforme comentado no tópico 3.2.

<sup>69</sup> É o caso de anúncios veiculados em redes sociais como Facebook e Instagram, por exemplo, personalizados de acordo com as preferências e buscas do usuário (THE SOCIAL DILEMMA, 2020).

Nesse sentido, vale lembrar que o surgimento da responsabilidade civil objetiva está justamente atrelado ao fato de que “a deficiência de meios, a desigualdade de fortuna, a própria organização social acabam por deixar larga cópia de danos descobertos e sem indenização” (PEREIRA, 2018A, p. 524). Com efeito, há situações em que é sabidamente difícil para a vítima demonstrar o elemento subjetivo; nesses casos, fez-se necessário aderir a meios técnicos capazes de viabilizar a reparação do dano sem o fundamento da culpa.

Essa lógica é igualmente adotada na Resolução de 16 de fevereiro de 2017 do Parlamento Europeu, que indica a adoção do regime da responsabilidade civil objetiva ou pela gestão de riscos como técnicas idôneas em eventual caso de reparação de danos. Embora os projetos de lei em trâmite no Congresso Nacional não indiquem de modo claro o regime de responsabilidade, dispondo apenas que a responsabilidade recairá sobre o supervisor da IA, a melhor adequação de um regime de responsabilidade civil objetiva é apontada com farto lastro na doutrina, que discute o enquadramento da IA como atividade de risco, produto ou a equipara aos animais, a fim de viabilizar a aplicação de teorias já existentes no direito civil.

Portanto, considerando que a principal preocupação quanto ao tratamento jurídico destinado ao uso da IA reside no fato de que algoritmos de *deep learning* se apresentam como verdadeiras caixas pretas, desafiando seus próprios criadores, bem como as sugestões no sentido de aderir à responsabilidade objetiva nas propostas legislativas analisadas no capítulo anterior, pode-se descartar, por ora, a possibilidade de adoção de um regime subjetivo de responsabilidade para a utilização de IA. Resta analisar, assim, as possibilidades de fundamentação para a responsabilidade objetiva ou indireta.

## 5.2 Responsabilidade Objetiva (ou *Strict Liability*)

A discussão sobre a responsabilidade civil relacionada à inteligência artificial vem ocorrendo na doutrina estrangeira há décadas<sup>70</sup>, mas é ainda recente no direito brasileiro. Por isso, boa parte das referências analisadas neste tópico foi extraída de publicações acadêmicas norte-americanas e europeias, escritas à luz de sistemas jurídicos de *Common Law* e de tradição romano-germânica.

Daí porque dois esclarecimentos iniciais são importantes para permitir uma análise crítica das teorias de responsabilidade civil objetiva ou indireta analisadas na doutrina estrangeira à luz da sistemática brasileira, tendo em vista que a responsabilidade subjetiva,

---

<sup>70</sup> “*Legal Personhood for Artificial Intelligences*”, de Lawrence B. Solum, foi publicado ainda em 1992, por exemplo.

porque fundada na culpa, é, de regra, comum em todos os sistemas – assim como o são as críticas apresentadas pela doutrina em relação à sua abordagem envolvendo a IA.

O primeiro esclarecimento que deve ser feito é que o tratamento da responsabilidade civil objetiva no direito brasileiro se dá por meio de cláusulas gerais. Tais cláusulas preveem, em síntese, a responsabilidade objetiva: (i.) em razão do exercício de atividade de risco ou perigosa (art. 927, parágrafo único); (ii.) em razão de danos causados por produtos postos em circulação (art. 931); (iii.) por fato de terceiro (arts. 928, 932 e 933); e (iv.) por fato da coisa e do animal (arts. 936, 937 e 939). Assim, conforme observa Sérgio Cavalieri Filho “muito pouco sobrou para a responsabilidade subjetiva” (2012, p. 170).

A existência de cláusulas gerais de responsabilidade objetiva é um aspecto relevante porque essa particularidade do sistema jurídico brasileiro o distingue de outras jurisdições, como é o caso de Portugal, cujo sistema de responsabilidade civil é essencialmente subjetivo<sup>71</sup> ou, então, baseado na culpa presumida (presunção relativa), hipótese da responsabilidade pelos danos causados por coisas e animais sob poder ou a vigilância do agente ou pelo exercício de atividades perigosas<sup>72</sup>. Assim, eventuais propostas de alteração legislativa para viabilizar um regime de responsabilidade objetiva mais abrangente apresentadas por autores estrangeiros poderão se mostrar desnecessárias no contexto brasileiro.

O segundo esclarecimento diz respeito às teorias de responsabilidade objetiva e indireta do direito comparado, especialmente da *Common Law*. Isso porque as propostas doutrinárias, em geral, dividem a possibilidade de tratamento jurídico da responsabilidade civil relacionada à IA em duas vertentes: a *strict liability* (responsabilidade objetiva, que abrange a responsabilidade por danos causados por animais, por fato do produto e por atividades de risco<sup>73</sup>) e *vicarious liability* (correspondente à responsabilidade indireta ou por fato de outrem<sup>74</sup>).

<sup>71</sup> Conforme refere Mafalda Miranda Barbosa, em relação ao sistema português, “as previsões de responsabilidade objetiva são caracterizadas pelo seu âmbito circunscrito. Responsabilizar um sujeito independentemente de culpa implica a existência de uma previsão expressa do legislador que possa assimilar a intencionalidade problemática do caso. No momento da decisão judicativa, em face de danos causados por entes dotados de inteligência artificial, podemos estar desamparados, pela inexistência de uma norma que solucione o problema concreto” (2020, p. 285).

<sup>72</sup> Conforme previsto no art. 493º do Código Civil Português. Nesse sentido, Henrique Sousa Antunes compara o sistema português com o brasileiro, destacando que “[n]o direito civil português, o legislador estabeleceu um regime geral de responsabilidade com culpa presumida pelo exercício de atividades que, pela sua natureza ou pela natureza dos meios envolvidos, causem danos a terceiros (artigo 493.º, n.º 2 do Código Civil). No direito civil brasileiro, o dever de indenizar nasce com independência de um juízo de censura ao comportamento do agente [...]” (2019, p. 146), referindo-se, nesse ponto, à regra do parágrafo único do art. 927 do Código Civil de 2002.

<sup>73</sup> Nesse sentido: “*There are three main cases where strict liability applies: (a) injuries by wild animals; (b) products liability; (c) abnormally dangerous activities*” (ČERKA; GRIGIENE; SIRBIKYTE, 2015, p. 286).

<sup>74</sup> Ainda conforme Čerka, Grigiene e Sirbikyte, “*This means that for AI's behaviour vicarious liability appears to a person on whose behalf it acts or at whose disposal and supervision AI is. It can be listed as users of AI or their owners*” (2015, p. 385).

Especificamente quanto às hipóteses de “*strict liability*”, Curtis Karnow aponta que nenhuma seria suficiente para lidar com problemas de IA autônoma, pois todas têm em comum um elemento de previsibilidade que não se observa na inteligência artificial (2016, p. 52). Outros autores apontam, de outro lado, a possibilidade de utilização de tais teorias. A análise dos pontos a favor e contra cada perspectiva será feita de acordo com as hipóteses previstas no Código Civil: atividade de risco ou fato do serviço, fato do produto e fato da coisa ou animal.

A discussão sobre a responsabilidade objetiva por fato de outrem (*vicarious liability*) será abordada em tópico próprio, já que pressupõe uma aproximação da IA com a figura de um “agente”, e não puramente um objeto. Na mesma linha, deve ser abordado, ainda que sucintamente, o papel do art. 187, brevemente referido no capítulo anterior como um dos pressupostos da responsabilidade civil, para viabilizar possíveis soluções normativas à hipótese de danos causados pela IA autônoma, sendo, também este, um tópico apartado, ainda que ensejador da responsabilidade objetiva.

## 5.2.1 A Responsabilidade objetiva em razão de atividade de risco

### 5.2.1.1 Os pontos positivos e negativos apontados na doutrina estrangeira para a responsabilidade por atividade de risco

Čerka, Grigiene e Sirbikyte referem que há duas teorias principais para a origem do risco ou perigo: a teoria do objeto, em que um objeto do mundo físico que não pode ser totalmente controlado é a fonte de perigo; e a teoria da atividade, em que a atividade, em si, está associada à fonte do risco (2015, p. 386). Os autores sustentam que a IA se encaixa na teoria do objeto<sup>75</sup>, razão pela qual as atividades executadas pelo sistema inteligente deveriam ser consideradas atividades perigosas, ensejando, conseqüentemente, a responsabilidade de quem se beneficia da utilização da IA.

A teoria aplicável, nesse caso, seria a de “*deep pocket*” do direito americano: quem pratica atividade de risco lucrativa e útil para a sociedade deve ser responsável pela indenização de danos eventualmente relacionados a essa atividade (ČERKA; GRIGIENE; SIRBIKYTE, 2015, p. 386). Ainda conforme os autores, essa modalidade de responsabilidade demandaria um seguro contra responsabilidade civil, de modo a garantir a reparação do dano, o que vai de

---

<sup>75</sup> Argumentam, nesse sentido, que: “*Since AI is able to draw individual conclusions from the gathered, structured, and generalized information as well as to respond accordingly, it should be accepted that its activities are hazardous. Accordingly, the AI developer should be held liable for the actions of the greater source of danger, and, in this case, liability arises without fault*”. (ČERKA; GRIGIENE; SIRBIKYTE, 2015, p. 386)

encontro à proposta contida na Resolução de 16 de fevereiro de 2017 do Parlamento Europeu sobre a criação de um sistema de seguros obrigatórios<sup>76</sup>.

Em reforço a essa conclusão, os autores ressaltam que a IA não poderia ser equiparada a um animal porque “as atividades da IA são baseadas em um processo algorítmico e ele pode entender as consequências de suas ações”<sup>77</sup>, ao passo que o fundamento da responsabilidade do dono no animal é que este último não seria capaz de compreender as consequências das próprias ações. Quanto à responsabilidade objetiva pelo fato do produto, os autores entendem que “a vítima encontraria muita dificuldade para provar que o produto de IA era defeituoso e, especialmente, que o defeito existia quando a IA deixou seu fabricante ou desenvolvedor”<sup>78</sup> (ambos os temas serão abordados em maior detalhe nos próximos tópicos). Restaria, portanto, a responsabilidade objetiva pela atividade de risco realizada pela IA.

Mas o entendimento quanto à IA/robô ser considerada uma atividade de risco não é unânime.

Considerando a responsabilidade objetiva pela prática de atividades de alto risco (*ultrahazardous activities*), Curtis Karnow (2016)<sup>79</sup> aponta que muitos dos danos causados por robôs não serão sérios, especialmente porque robôs não deverão causar danos ordinariamente. Além disso, a previsibilidade de que o robô implique em algum risco quando posto em circulação é um fator essencial. O autor adverte, assim, que “a menos que estejamos dispostos a assumir todas as ações robóticas previsivelmente perigosas porque algumas são imprevisivelmente perigosas, esta doutrina não ajudará”<sup>80</sup> (KARNOW, 2016, p. 68, tradução nossa).

Além disso, há que se considerar que a criação de sistemas de inteligência artificial pode ter por objetivo justamente *reduzir* riscos e tornar atividades mais seguras. Nesses casos, há um evidente paradoxo no fomento ao desenvolvimento de ferramentas de IA voltadas a tornar mais

---

<sup>76</sup> “59. Insta a Comissão a explorar, analisar e ponderar, na avaliação de impacto que fizer do seu futuro instrumento legislativo, as implicações de todas as soluções jurídicas possíveis, tais como: a) Criar um regime de seguros obrigatórios, se tal for pertinente e necessário para categorias específicas de robôs, em que, tal como acontece já com os carros, os produtores ou os proprietários de robôs sejam obrigados a subscrever um seguro para cobrir os danos potencialmente causados pelos seus robôs”. (UNIÃO EUROPEIA, 2017)

<sup>77</sup> No original: “AI activities are based on an algorithmic process and it can understand the consequences of its actions” (ČERKA; GRIGIENE; SIRBIKYTE, 2015, p. 387).

<sup>78</sup> No original: “the plaintiff would find it very difficult to prove that the AI product was defective and especially that the defect existed when AI left its manufacturer or developer” (ČERKA; GRIGIENE; SIRBIKYTE, 2015, p. 387).

<sup>79</sup> No original: “Many of the injuries may not be serious, and robots are not likely routinely to pose hazards. As we will see with the other three types of strict liability, foreseeability runs a line through this doctrine of ultrahazardous as well: the essence of the liability is that harm is especially likely to happen, harm that the defendant knows about, or should have known about; that is, predictable.” (KARNOW, 2016, p. 68).

<sup>80</sup> No original, “[u]nless we are willing to dub all robotic actions foreseeably dangerous because some are unforeseeably dangerous, this doctrine will not assist”.

seguras determinadas atividades ao mesmo tempo em que a IA é considerada, ela própria, uma fonte de risco.

Tome-se o exemplo dos veículos autônomos. Apesar dos dilemas morais sobre situações de colisão de direitos envolvendo acidentes com veículos autônomos (se o veículo deve optar por atropelar um idoso ou um cachorro, uma gestante ou uma criança, salvar o ocupante do veículo ou quem estiver na rua, etc.), há um fator objetivo importante: a maior parte dos acidentes com veículos terrestres é causada pela imprudência dos motoristas<sup>81</sup>. A imprudência está associada à embriaguez ao volante, fadiga e violação às normas de trânsito, e nenhum desses fatores afeta a inteligência artificial. Ou seja, objetivamente, substituir motoristas humanos por veículos inteligentes *reduz* o risco de acidentes de trânsito.

Foi nesse contexto que David Vladeck defendeu que a IA não pode ser considerada uma atividade de risco, sendo necessária a adoção de um regime de responsabilidade objetiva que seja desvinculado da noção de risco:

Em minha opinião, nos casos em que carros sem motorista falham e causam danos a pessoas ou propriedades e não seria razoável atribuir a falha à fabricação ou ao design do veículo, o direito precisará formular uma resposta que melhor atenda aos interesses coletivos das partes afetadas. Minha proposta é construir um sistema de responsabilidade objetiva, completamente desvinculado das noções de culpa para esse seletivo grupo de casos. Um regime de responsabilidade objetiva não pode ser baseado aqui no argumento de que os veículos são "ultra perigosos" ou "excessivamente arriscados" pela simples razão de que os veículos sem motorista são provavelmente muito menos perigosos ou arriscados do que os produtos que eles substituem. De fato, é precisamente porque essas máquinas são tão tecnologicamente avançadas que esperamos que não falhem. Por essas razões, será necessário um verdadeiro regime de responsabilidade objetiva; um que não recorra a um teste de utilidade-risco ou regresse ao padrão de negligência pelo simples fato de que esses testes serão difíceis, se não impossíveis, de serem superados pela parte lesada.<sup>82</sup> (2014, p. 146, tradução nossa)

Vladeck apontou ao menos quatro argumentos favoráveis à adoção da responsabilidade objetiva (não fundada no risco) para o uso de IA: *(i.)* impor que indivíduos arquem com um dano relacionado à IA apenas porque a sua causa é inexplicável contraria noções básicas de justiça compensatória e da alocação de riscos na sociedade; *(ii.)* diferentemente das vítimas, os

<sup>81</sup> Conforme reportagem publicada no Estadão em março de 2020, 53,7% dos acidentes de trânsito são causados pela imprudência dos motoristas. Disponível em: <https://mobilidade.estadao.com.br/mobilidade-para-que/imprudencia-e-principal-caoa-de-acidentes-de-transito-no-brasil/>. Acesso em: 24 out. 2020.

<sup>82</sup> No original: *"In my view, in cases where driver-less cars fail and cause injuries to persons or property and it would be unreasonable to attribute the failure to the vehicle's manufacture or design, the law will need to fashion a response that best serves the collective interests of the affected parties. My proposal is to construct a system of strict liability, completely uncoupled from notions of fault for this select group of cases. A strict liability regime cannot be based here on the argument that the vehicles are "ultra-hazardous" or "unreasonably risky" for the simple reason that driver-less vehicles are likely to be far less hazardous or risky than the products they replace. Indeed, it is precisely because these machines are so technologically advanced that we expect them not to fail. For these reasons, a true strict liability regime will be needed; one that does not resort to a risk-utility test or the re-institution of a negligence standard for the simple fact that those tests will be difficult, if not impossible, for the injured party to overcome"*.

criadores dos veículos autônomos (hipótese do exemplo) estão em uma posição em que é possível absorver os custos da indenização; (iii.) a definição de que o regime de responsabilidade será objetivo evita maiores discussões no processo (ideia de que é melhor direcionar os recursos para indenizar a parte lesada do que para pagar custas do processo e advogados); e (iv.) um sistema legal previsível é mais benéfico à inovação do que um sistema incerto (VLADECK, 2014).

O americano sugeriu, então, que a responsabilidade objetiva seja baseada na doutrina da “*common enterprise*”<sup>83</sup>, isto é, todas as empresas envolvidas na criação, programação e comercialização da ferramenta de IA serão responsáveis solidariamente pelos danos causados pela IA, pois elas têm melhores condições de gerir os custos que o pagamento de eventuais indenizações impactará na produção e desenvolvimento da IA. Segundo Vladeck:

Uma teoria do empreendimento comum permite que a lei imponha responsabilidade conjunta sem ter que expor e lidar com os detalhes da atribuição de cada aspecto do ilícito a uma parte ou outra; basta que, em busca de um objetivo comum, as partes se envolveram no ilícito. Esse princípio poderia ser enxertado em um novo regime de responsabilidade objetiva para lidar com os danos que podem ser causados aos seres humanos por máquinas autônomas inteligentes quando for impossível ou impraticável atribuir a culpa a uma pessoa específica.<sup>84</sup> (2014, p. 149, tradução nossa)

Por ter como pressuposto justamente a exclusão do elemento do risco, a proposta de Vladeck não encontra correspondência com a hipótese de aplicação do art. 927, parágrafo único, do Código Civil. A sugestão, todavia, encontra argumentos contundentes a seu favor, porque permite um ajuste econômico dos custos de um sistema de responsabilidade civil, ao mesmo tempo que viabiliza a responsabilidade daqueles que se beneficiam da comercialização de produtos de IA e são os responsáveis por escrever os códigos-fonte do *software*, os quais, aliás, estarão protegidos por direitos de propriedade intelectual e, de regra, não serão exibidos aos usuários<sup>85</sup>, o que certamente dificultará a eventual apuração da responsabilidade, como ocorreu com o COMPAS, analisado anteriormente.

---

<sup>83</sup> A teoria da “*enterprise liability*” está associada justamente à alocação de recursos na sociedade e tem origem econômica, no equilíbrio entre oferta e demanda. No caso da responsabilidade civil, significa definir sobre quem recai o ônus da perda e como o dever de compensar determinado dano pode ser realocado, por exemplo, no preço de determinado produto (KLEMMME, 1975).

<sup>84</sup> No original, “*A common enterprise theory permits the law to impose joint liability without having to lay bare and grapple with the details of assigning every aspect of wrongdoing to one party or another; it is enough that in pursuit of a common aim the parties engaged in wrongdoing. That principle could be engrafted onto a new, strict liability regime to address the harms that may be visited on humans by intelligent autonomous machines when it is impossible or impracticable to assign fault to a specific person.*”

<sup>85</sup> Conforme comenta Sabina Cavalli (2008), ao tratar da natureza jurídica da obrigação de implantação de software, “a fim de resguardar sua criação intelectual, o fornecedor geralmente entrega ao usuário o programa em linguagem inteligível apenas à máquina, guardando para si a informação acerca do código fonte.”

### 5.2.1.2 O art. 927, parágrafo único, do Código Civil de 2002

Conforme abordado no tópico sobre os pressupostos da responsabilidade civil no direito brasileiro, é o fator de imputação que distingue a responsabilidade subjetiva da objetiva: enquanto esta decorre do *risco*, aquela dependerá da demonstração da *culpa*. O Código Civil de 2002 possui, nessa linha, uma cláusula geral de responsabilidade objetiva fundada no risco, prevista no art. 927, parágrafo único (“haverá obrigação de reparar o dano, independentemente de culpa, nos casos especificados em lei, ou quando a atividade normalmente desenvolvida pelo autor do dano implicar, por sua natureza, risco para os direitos de outrem”) (BRASIL, 2002).

Sérgio Cavalieri Filho pontua que a expressão “atividade normalmente desenvolvida” do texto legal “indica serviço, ou seja, atuação reiterada, habitual, organizada profissional ou empresarialmente para realizar fins econômicos” (2012, p. 187). Essa regra trata, assim, da responsabilidade por fato do serviço, à semelhança do que disciplina o art. 14 do Código de Defesa do Consumidor (“CDC”) no âmbito das relações consumeristas, mas com alcance mais amplo, abrangendo todos os casos não enquadrados como relação de consumo ou com regramento específico (como é o serviço público, regrado pelo art. 37, § 6º, da Constituição Federal, e o transporte de passageiros, regrado pelo art. 734 e seguintes do Código Civil).

Quanto à noção de risco, Cavalieri aponta que há na doutrina quem entenda se tratar de atividades de risco excepcional, como é o caso de serviços que envolvem materiais radioativos, explosivos ou inflamáveis, por exemplo. Todavia, o Código Civil adotou a teoria do risco criado, de modo a fundamentar a responsabilidade sempre que o risco for *inerente* ao serviço (abrangendo, assim, o risco excepcional ou não) (CAVALIERI, 2012, p. 188). No mesmo sentido, também dispôs o Enunciado 448, aprovado na V Jornada de Direito Civil<sup>86</sup>.

Assim, para que a responsabilidade objetiva nos casos relacionados ao uso de IA autônoma (*softwares* e, especialmente, robôs) com base na cláusula geral do parágrafo único do art. 927, seria preciso delimitar, de início, se a conduta relacionada à IA (*i.*) pode ser considerada uma “atividade” (ou serviço); e (*ii.*) pode ser considerado um fator de *criação* de risco. Em qualquer caso, será necessário demonstrar o nexo de causalidade entre o dano efetivamente suportado e o uso da IA.

---

<sup>86</sup> “A regra do art. 927, parágrafo único, segunda parte, do CC aplica-se sempre que a atividade normalmente desenvolvida, mesmo sem defeito e não essencialmente perigosa, induza, por sua natureza, risco especial e diferenciado aos direitos de outrem. São critérios de avaliação desse risco, entre outros, a estatística, a prova técnica e as máximas de experiência”.

Mas não basta definir se a IA pode ser enquadrada na regra que autoriza a responsabilidade objetiva: é preciso estabelecer, ainda, a quem se deve imputar o dano. Seguindo a linha do que sugere o PL 21/2020, parece sensato concluir que a regra do art. 927, parágrafo único, do Código Civil, poderia ser conjugada com a regra proposta no art. 9º do projeto de lei, ensejando a responsabilidade objetiva do agente de inteligência artificial.

De outro lado, vale observar que o art. 927, parágrafo único, não menciona nenhuma característica quanto a quem se deve imputar o dano. Nesse sentido, na eventualidade de se admitir um estatuto jurídico próprio, é igualmente possível questionar se o robô poderia responder objetivamente com base em dita cláusula, já que seria considerado, ao mesmo tempo, o “autor do dano” e a própria “atividade de risco”.

Embora essa suposição implique em evidente exercício de futurologia, especialmente nesse segundo caso, em que a IA figuraria no suporte fático da regra como o agente e a própria atividade, caberia ao legislador delimitar, de forma clara, se a IA é uma “atividade” de risco, editando uma regra específica para tanto, ainda que de aplicação conjunta com o parágrafo único do art. 927. Sugere-se isso por dois motivos: primeiro, porque, conforme apontado por David Vladeck, a IA tem sido desenvolvida para *evitar* o risco de determinadas atividades, ou seja, não se está nem no campo do risco excepcional, nem no do risco inerente; segundo, porque a IA pode dar causa a determinado evento por funcionar perfeitamente, adotando um comportamento a partir de processos de *machine learning* baseados em informações que seu próprio usuário lhe forneceu.

São fatores como esses que evidenciam a IA como algo completamente novo para as situações reguladas pelo Direito e, mais do que isso, algo completamente “sui generis” para os padrões humanos. Conforme já mencionado, o regime de responsabilidade objetiva é o mais indicado para os casos relacionados à IA autônoma, mas delegar a solução dos casos concretos a uma regra geral baseada no risco pode não ser a melhor alternativa. Nesse sentido, destaca-se posição já apresentada por Mafalda Miranda Barbosa: embora seja possível que alguns robôs ou sistemas de IA se envolvam em situações de perigo, isso pode não ocorrer com outros, de sorte que “[a] adequação do critério depende do específico desenho da norma que possa vir prever uma hipótese de responsabilidade objetiva” (2020, p. 292).

Mas há ainda um ponto – não abordado expressamente por Vladeck ou Karnow – que permite resgatar o argumento do risco: o viés algorítmico.

No capítulo 2, buscou-se apresentar a experiência do COMPAS na justiça criminal norte-americana justamente para expor que a possibilidade de resultados enviesados é um dos grandes dilemas envolvendo a automatização de processos decisórios. O exemplo igualmente

mencionado do Hospital St. George, no Reino Unido, permite afirmar que o uso de algoritmos, ainda que bem intencionado, pode ter como consequência exatamente o resultado que se buscava evitar, com a reprodução de comportamentos preconceituosos ou “conservadores” aprendidos pelo algoritmo a partir dos dados de *input* de experiências anteriores. Com efeito, no caso do Hospital St. George, a própria instituição foi responsabilizada quando se descobriu que a automatização da primeira fase do processo seletivo para contratação de médicos, baseada nas contratações anteriores e nos arquivos da instituição, resultara na exclusão de mulheres, negros e imigrantes da segunda etapa da seleção, já que o sistema aprendeu a partir de um padrão enviesado que vinha se repetindo nas últimas décadas.

O viés lá verificado pode igualmente estar presente na IA autônoma. Ora, algoritmos são conjuntos de instruções baseadas em modelos matemáticos; a própria definição de IA como “uma ciência que trata de estudar o comportamento inteligente através de modelos matemáticos”<sup>87</sup> (ALABART, 2018, p. 14, tradução nossa) deixa claro que a sua utilização não se confunde com as decisões – supostamente – morais tomadas no processo decisório humano. A opção pela utilização de sistemas inteligentes baseados em técnicas de *machine learning* e *deep learning* envolve, necessariamente, a assunção de um risco relacionado ao aprendizado da IA e como as informações fornecidas ao algoritmo serão interpretadas com o passar do tempo.

Na conclusão do livro “*Weapons of Math Destruction*”, Cathy O’Neill apresenta um argumento interessante que permitiria interpretar o elemento “risco” da responsabilidade civil como uma decorrência da possibilidade de viés algorítmico e a consequente aplicação da teoria para resolver os casos em que a IA autônoma causar danos:

Se nos afastarmos deles e tratarmos os modelos matemáticos como uma força neutra e inevitável, como o tempo ou as marés, abdicamos de nossa responsabilidade. E o resultado, como vimos, são as *WMDs*<sup>88</sup> que nos tratam como peças de máquina no local de trabalho, que excluem os funcionários e se banqueteam nas desigualdades. Devemos nos unir para policiar essas *WMDs*, para domesticá-las e desarmá-las. Minha esperança é que elas sejam lembradas, como as mortais minas de carvão de um século atrás, como relíquias dos primeiros dias desta nova revolução, antes de aprendermos como trazer justiça e responsabilidade para a era dos dados. A matemática merece muito mais do que *WMDs*, e a democracia também.<sup>89</sup> (O’NEILL, 2016, p. 218, tradução nossa)

---

<sup>87</sup> No original, “*es una ciencia que trata de estudiar el comportamiento inteligente a través de modelos matemáticos*”.

<sup>88</sup> Termo utilizado pela autora como abreviação de “*weapons of math destruction*”, em português, “máquinas de destruição matemática”.

<sup>89</sup> No original, “*If we back away from them and treat mathematical models as a neutral and inevitable force, like the weather or the tides, we abdicate our responsibility. And the result, as we’ve seen, is WMDs that treat us like machine parts in the workplace, that blackball employees and feast on inequities. We must come together to police these WMDs, to tame and disarm them. My hope is that they’ll be remembered, like the deadly coal mines of a*

O’Neill destaca que os modelos matemáticos utilizados nos produtos de IA “são construídos não apenas a partir dos dados, mas das escolhas que nós fazemos sobre quais dados prestar atenção – e quais deixar de fora”<sup>90</sup>, de modo que essas escolhas não dizem respeito apenas à “logística, lucros e eficiência”, porque são, fundamentalmente, escolhas *morais* (2016, p. 218, tradução nossa).

Ou seja, a partir da escolha de delegar determinada atividade a um produto de IA autônoma e da filtragem inicial de dados que estarão disponíveis para o *input*, há um intrínseco elemento humano que vincula a atividade do proprietário da IA, do programador ou do fabricante que dela se utiliza, à criação de um risco de tomada de decisão enviesada. Essa possibilidade de utilização de um sistema de IA com viés – ainda que a IA seja criada para reduzir outros riscos inerentes à atividade que ela irá substituir ou atuar – parece permitir uma interpretação favorável a aplicação da teoria do risco.

É preciso atentar, todavia, para a advertência de Curtis Karnow (2016, p. 68), no sentido de que se assumiria que todos os algoritmos podem ter um risco de viés apenas porque alguns irão, efetivamente, agir desse modo. Nessa linha, parece plausível questionar, na hipótese de que apenas alguns sistemas podem apresentar viés – e não todos –, se seria possível interpretar o viés algorítmico não como um “risco”, e sim como um “defeito” da IA autônoma, aplicando-se, então, a responsabilidade pelo fato do produto.

Contudo, o exemplo do COMPAS, trabalhado no Capítulo 2, pode ser novamente útil para rechaçar essa interpretação de viés algorítmico como um defeito, especialmente porque provar o viés pode ser tão difícil quanto provar a culpa do agente a quem se imputaria o dever de indenizar – e este foi, justamente, o motivo pelo qual se refutou anteriormente a possibilidade de aplicação da responsabilidade subjetiva. Lembra-se, nesse sentido, que o uso da ferramenta *evidence-based* foi considerado possível pela Suprema Corte de Wisconsin (caso *Loomis v. Wisconsin*) e, também, pela Suprema Corte de Indiana (caso *Malenchik v. Indiana*), a despeito das pesquisas acadêmicas realizadas que apontavam para a possível existência do comportamento enviesado do algoritmo.

Portanto, pode-se dizer que há uma lacuna de reconhecimento acerca da amplitude da noção de risco, cujo termo poderia ser interpretado de modo a abranger o “risco” de uma decisão enviesada do algoritmo como elemento central para a aplicação do parágrafo único do art. 927,

---

*century ago, as relics of the early days of this new revolution, before we learned how to bring fairness and accountability to the age of data. Math deserves much better than WMDs, and democracy does too.”*

<sup>90</sup> No original, “*these models are constructed not just from data but from the choices we make about which data to pay attention to – and which to leave out.*”. A autora complementa que “*Those choices are not just about logistics, profits, and efficiency. They are fundamentally moral.*” (O’NEILL, 2016, p. 218)

a despeito de o sistema ter sido desenvolvido para mitigar riscos inerentes a determinada atividade. A suficiência desta solução e a possibilidade de assumir que o viés algorítmico pode, efetivamente, ser um risco a *todos* os sistemas de IA, e não apenas a alguns, demanda, todavia, maiores discussões, motivo pelo qual passa-se a analisar se as demais teorias de responsabilidade extra negocial fornecem solução mais adequada ou completa.

## 5.2.2 Responsabilidade objetiva pelo fato do produto (*product liability*)

### 5.2.2.1 O debate acerca da responsabilidade fundamentada no fato do produto

A responsabilidade objetiva pelo fato do produto ou *product liability* é a segunda teoria discutida na doutrina como fundamento para a responsabilidade civil relacionada à IA. A responsabilidade civil por fato do produto está prevista no art. 931 do Código Civil e, também, no CDC. Conforme antes mencionado, o objetivo primordial desta pesquisa é analisar a suficiência do regime de responsabilidade civil do Código Civil de 2002. Sem prejuízo, quando pertinentes, serão expostos, especialmente nesta seção, aspectos da legislação consumerista, intrinsecamente ligada à análise do art. 931 do Código Civil.

Pois bem, os principais aspectos com relação a abordagem da responsabilidade pelo fato do produto são, em síntese, a possibilidade de classificar a IA como um produto; o ônus de provar um defeito na IA que permita estabelecer o nexo causal entre dito defeito e o dano; e a responsabilidade pelos riscos do desenvolvimento. Todas essas questões têm o condão de afetar não apenas as regras do Código Civil, como também a aplicação do CDC em eventual relação de consumo, especialmente em razão das duas últimas preocupações.

O art. 3º, § 1º, do CDC define produto como “qualquer bem, móvel ou imóvel, material ou imaterial” (BRASIL, 1990). Na perspectiva da IA como objeto do direito, tal conceito poderia, portanto, ser aplicado, já que a IA pode ser considerada um bem, seja na forma de *softwares*, seja na forma de um robô.

Com efeito, não é na caracterização da IA como um produto que reside propriamente o debate sobre a possibilidade ou não se aplicação de regras de responsabilidade objetiva por fato do produto, mas, sim, nos dois outros pontos, isto é, na preocupação com a dificuldade imposta à vítima de demonstrar que a IA possuía um defeito e que esse defeito já existia à época em que a IA foi lançada ao mercado. Conforme pontua a professora Mafalda Miranda Barbosa,

[...] a idealização do robot (programação do *software*) pode não apresentar qualquer defeito, do mesmo modo que, na fase do fabrico do mecanismo no qual se integra a

inteligência artificial, pode não ocorrer qualquer desconformidade entre o resultado final e o que era esperado pelo produtor. Os danos causados pelo robot dito inteligente são gerados pela sua atuação autónoma que, longe de ser uma marca de defeituosidade, se traduz numa sua característica intrínseca. (2020, p. 288).

O mesmo argumento foi considerado por Čerka, Grigiene e Sirbikyte para afastar a viabilidade da teoria da responsabilidade por fato do produto na solução de casos de responsabilidade civil envolvendo IA<sup>91</sup>. As duas principais objeções referem-se à dificuldade de prova de defeito e à possibilidade de responsabilização do produtor pelos riscos do desenvolvimento.

Considerando características como *machine learning* e *deep learning* sob o ponto de vista da responsabilização do fabricante pelos riscos do desenvolvimento, há, inclusive, um interessante paradoxo. De um lado, em sistemas jurídicos que não admitem a responsabilidade pelos riscos do desenvolvimento, fabricantes teriam uma certa “imunidade” em termos de responsabilização, já que o produto estará em constante modificação em razão das interações futuras e os próprios processos de aprendizado, não sendo possível prever, no momento em que se põe o produto em circulação, quais serão as ações tomadas pela IA no futuro. De outro lado, admitir que o fabricante responda por tais riscos significa imputar-lhe o ônus de arcar com comportamentos que são dependentes das informações e interações da IA com terceiros, que poderão ou não agir de boa-fé, ou seja, ações que estão completamente fora de suas possibilidades de controle e previsibilidade *razoável*.

Isso não significa que os produtores não tenham responsabilidade de acompanhar os produtos postos em circulação e tomar providências caso constatem alguma falha de segurança. Na legislação brasileira, esse dever foi positivado, no âmbito das relações de consumo, no art. 10 do CDC, que disciplina o chamado “*recall*”. Conforme destaca Leonardo Roscoe Bessa, o *recall* tem como ideia base “a prevenção de eventos danosos, ou seja, que o consumidor seja chamado para realizar o reparo ou troca do produto (ou de uma peça) antes de ocorrência do acidente de consumo” (BESSA, 2013). Embora se trate de uma regra da legislação consumerista, Mafalda Miranda Barbosa destaca, à luz do sistema jurídico português, que o produtor tem uma obrigação geral de segurança, que, por si só, já justificaria a sua responsabilidade pela falha no cuidado no âmbito civil, não de consumo (2020, p. 289).

---

<sup>91</sup> Nesse sentido: “*Also in some cases it would be difficult to apply the product liability case, because AI is a self-learning system that learns from its experience and can take autonomous decisions. Thus, for the plaintiff it would be difficult to prove an AI product defect and especially that the defect existed when AI left its manufacturer's or developer's hands. It is hard to believe that it is possible to draw the line between damages resulting from the AI will, i.e. derived from self-decision, and damages resulting from product defect; unless we would equate the independent decision-making (which is a distinctive AI feature) with a defect.*” (ČERKA; GRIGIENE; SIRBIKYTE, 2015, p. 386).

Essa questão é considerada por Ugo Pagallo, que aponta que o desempenho da IA dependerá do tipo de informação que lhe é fornecido e quais comportamentos cercam o seu meio. Nesse contexto, Pagallo sustenta que os usuários da IA estariam em melhores condições de verificar um possível comportamento ilícito da IA em comparação com os fabricantes, que tiveram contato apenas com o sistema nativo. Segundo ele:

[...] a capacidade de tais máquinas de obter conhecimento e habilidades a partir da interação com cuidadores humanos sugere que a falha raramente recairia sobre os projetistas, fabricantes ou fornecedores de tais robôs. Em vez disso, de acordo com a lógica das regras de responsabilidade objetiva, se poderia argumentar que os proprietários ou usuários de robôs estão na melhor posição para entender o que está acontecendo com a máquina, a fim de evitar seu comportamento perigoso, independentemente de a conduta do robô ter sido típica de robôs similares, razoavelmente previsível e assim por diante.<sup>92</sup> (PAGALLO, 2013, p. 126, tradução nossa)

Assim, antes dos fabricantes, faria mais sentido imputar aos usuários ou donos da IA a responsabilidade por eventual ação que tenha causado danos, já que eles seriam responsáveis pelo “treinamento” do *software*. A responsabilidade pelo fato do produto, na linha do argumento de Pagallo (2013, p. 136), só seria aplicada aos chamados “*robots-as-means*”, isto é, robôs utilizados como meios, e não os robôs mais avançados, como seria o caso da IA autônoma. Esse argumento será abordado no próximo item, todavia, porque a imputação é baseada na teoria da responsabilidade indireta, e não pelo fato do produto. Por ora, passa-se a analisar se as questões apontadas na doutrina sobre a dificuldade de se adotar como fundamento o fato do produto (prova de defeito e riscos de desenvolvimento) seriam, também no contexto do Código Civil, um empecilho para a solução da responsabilidade civil relacionada à IA com base no art. 931.

#### 5.2.2.2 O artigo 931 do Código Civil: um dispositivo de utilidade controvertida

A segunda cláusula geral de responsabilidade objetiva do Código Civil está no art. 931, que positiva a responsabilidade por fato do produto e dispõe que, “[r]essalvados outros casos previstos em lei especial, os empresários individuais e as empresas respondem independentemente de culpa pelos danos causados pelos produtos postos em circulação”. A utilidade desse dispositivo é bastante contestada na doutrina, tendo em vista que a

---

<sup>92</sup> No original: “*the capability of such machines to gain knowledge and skills from interaction with human caretakers, suggests that the fault would rarely fall on the designers, manufacturers or suppliers of such robots. Rather, according to the rationale for strict liability rules, it could be argued that owners or users of robots are in the best position to understand what is going on with the machine, so as to prevent its dangerous behaviour, regardless of whether the conduct of the robot was typical of similar robots, reasonably foreseeable and so forth.*”

responsabilidade pelo fato do produto foi regrada de modo muito mais detalhado no art. 12 e seguintes do CDC.

A favor da utilidade do art. 931, Tula Wesendonck sustenta que a norma constante do Código Civil é relevante, especialmente, por três aspectos: *(i.)* é uma cláusula geral de responsabilidade objetiva da empresa perante terceiros, consumidores ou não; *(ii.)* a sua incidência independe da existência de um defeito no produto; e *(iii.)* tal como posta, a redação do art. 931 permite reconhecer a responsabilidade civil pelos riscos do desenvolvimento, já que o que interessa, para a aplicação da norma, é a colocação do produto em circulação (WESENDONCK, 2015).

Já Sérgio Cavalieri Filho aponta que o art. 931 do Código Civil e o art. 12 do CDC “se integram e se harmonizam” (2012, p. 194), sugerindo uma leitura conjunta dos dois dispositivos. O doutrinador também destaca argumentos levantados por Tula Wesendonck (2015) no sentido de que *(i.)* o art. 931 é mais amplo que a norma consumerista, permitindo resolver, pela via da responsabilidade objetiva, demandas que antes dependiam da prova da culpa pela regra geral do Código Civil e que não se enquadravam no âmbito de aplicação do CDC; e *(ii.)* que o art. 931 dá azo à responsabilidade civil pelos riscos do desenvolvimento, mencionando o Enunciado nº. 43, aprovado na Jornada de Direito Civil promovida pelo Centro de Estudos Judiciários do Conselho da Justiça Federal em 2002<sup>93</sup> (CAVALIERI, 2012, p. 200).

Nada obstante, diferentemente do posicionamento sustentado por Tula Wesendonck, Sérgio Cavalieri Filho entende que a responsabilidade pelo fato do produto o art. 931 depende da comprovação de defeito, asseverando que “dano que não tenha por causa um defeito do produto (e isso pode decorrer de diversos outros fatores, inclusive da própria conduta do usuário) não pode ser imputado ao empresário” (CAVALIERI, 2012, p. 195). Essa divergência também reflete nas causas excludentes da responsabilidade.

Enquanto Cavalieri aponta como excludentes de responsabilidade pelo fato do produto *(i.)* a ausência de defeito do produto; *(ii.)* o fortuito externo; *(iii.)* força maior; ou *(iv.)* fato exclusivo da vítima ou de terceiro (2012, p. 198), Wesendonck aponta que “a possibilidade de exclusão da responsabilidade dá-se pela prova do caso fortuito, força maior, fato de terceiro ou da vítima, ou então pela prova de não ter colocado o produto no mercado, não havendo como excluir a responsabilidade pela prova de inexistência do defeito” (WESENDONCK, 2015).

Daniel Amaral Carnaúba ressalta que divergências como essa evidenciam a insuficiência do modelo de responsabilidade civil fornecido pelo art. 931. Carnaúba elenca três

---

<sup>93</sup> “A responsabilidade civil pelo fato do produto, prevista no art. 931 do novo Código Civil, também inclui os riscos do desenvolvimento”.

pontos que demonstram a inutilidade do art. 931 na legislação brasileira: (i.) o fato de o art. 931 não mencionar que o defeito do produto é uma condição para a responsabilização do produtor; (ii.) a ausência de uma regulamentação diferenciada da responsabilidade de comerciantes; e (iii.) a menção a um rol específico das excludentes de responsabilidade (CARNAÚBA, 2020).

A fim de sustentar a inutilidade do dispositivo, Carnaúba aponta ao menos quatro posições doutrinárias que tentaram “salvá-lo”. A primeira sustenta que o art. 931 tem aplicação subsidiária ao CDC, incidindo nos casos não considerados como relação de consumo; mas esse entendimento não se sustenta, pois o art. 17 do CDC equipara a consumidor todas as vítimas de determinado evento danoso, permitindo-se, assim, a aplicação do diploma consumerista às relações paritárias, o que já foi, inclusive, reconhecido pelo Superior Tribunal de Justiça (STJ)<sup>94</sup>.

A segunda proposta consistiria em complementar eventuais lacunas do art. 931 com as regras do CDC, tal como apontado por Sérgio Cavalieri Filho no sentido de “integração” e “harmonização” dos dois diplomas legais. A principal discussão, aqui, é se o defeito é ou não um elemento necessário para a configuração da responsabilidade civil. Conforme já apontado acima, Tula Wesendonck é uma das principais vozes na doutrina no sentido de que o defeito seria um elemento dispensável à luz do art. 931, visão criticada por Carnaúba, que afirma:

Há apenas dois tipos de riscos que os produtos podem oferecer aos seus usuários: os riscos naturais e esperados daquela espécie de produto (os riscos inerentes) e os riscos não esperados, decorrentes de falhas de segurança (os riscos adquiridos, isto é, defeitos). Por consequência, quando a autora afirma que o empresário não responde pelos riscos inerentes ao produto, ela nada mais faz do que declarar, por meio de uma formulação negativa, que ele só responderá pelos riscos decorrentes de defeitos. No fundo, portanto, sua interpretação não destoa do entendimento majoritário de que a responsabilidade do art. 931 exige a constatação de um defeito do produto. (CARNAÚBA, 2020)

Além disso, Carnaúba entende que a interpretação “complementada” do art. 931 do Código Civil com as regras do CDC lhe retira todo o seu efeito prático, uma vez que a sua incompletude daria azo, ao fim e ao cabo, justamente à aplicação do diploma consumerista, muito mais detalhado e específico.

A terceira linha em favor do art. 931 é no sentido de que este dispositivo viabilizaria a responsabilidade do empresário pelos riscos do desenvolvimento (conforme afirmam

---

<sup>94</sup> “(...) 1 - Comerciante atingido em seu olho esquerdo pelos estilhaços de uma garrafa de cerveja, que estourou em suas mãos quando a colocava em um freezer, causando graves lesões. 2 - Enquadramento do comerciante, que é vítima de um acidente de consumo, no conceito ampliado de consumidor estabelecido pela regra do art. 17 do CDC (“*bystander*”). 3 - Reconhecimento do nexos causal entre as lesões sofridas pelo consumidor e o estouro da garrafa de cerveja. 4 - Ônus da prova da inexistência de defeito do produto atribuído pelo legislador ao fabricante. 5 - Caracterização da violação à regra do inciso II do § 3º do art. 12 do CDC. (...)” (BRASIL, 2013).

Wesendonck e Cavalieri). Isso, contudo, já seria admitido no direito brasileiro ainda antes do advento do Código Civil de 2002, pela aplicação do art. 12, § 1º, do CDC, dado que a verificação de defeito se dá por "dois parâmetros objetivos: de um lado, o grau de segurança que legitimamente se esperava daquele produto; de outro, o grau de segurança que, de fato, ele apresentou", de modo que o produto será defeituoso "toda vez que este parâmetro fatural for inferior àquele parâmetro expectado" (CARNAÚBA, 2020), o que vale também para eventuais falhas que não poderiam ter sido constatadas quando da colocação do produto em circulação.

A última corrente em defesa do art. 931 sustenta que esse dispositivo viabilizaria a responsabilidade civil do comerciante. Porém, conforme comenta Carnaúba (2020), o art. 931 tem expressas ressalva quanto aos "casos previstos em lei especial" e, estando a responsabilidade do comerciante regulada de modo específico no art. 13 do CDC, este dispositivo poderia garantir a imunidade do comerciante, ainda que o art. 931 admitisse a sua responsabilidade.

A controvérsia doutrinária quanto à utilidade do art. 931 parece indicar que delegar a esse dispositivo as soluções para a responsabilidade civil envolvendo a IA pode não ser uma solução eficiente. Especialmente com relação à IA, há duas objeções relevantes que precisam ser consideradas: (i.) a divergência interpretativa quanto à necessidade de prova do defeito do produto ou não, já que a IA poderá causar danos mesmo em perfeito funcionamento; e (ii.) o art. 931 trata da responsabilidade do empresário, que poderá não ter nenhuma ingerência nos processos de aprendizado da IA a partir das interações com os usuários, os quais teriam, nesse caso, relação mais próxima com a causa do dano e, especialmente, com o processo de aprendizado da IA.

Portanto, embora forneça uma solução normativa pela via da responsabilidade objetiva, parece-nos que também a regra do art. 931 de mostra insuficiente para resolver, de modo consistente, a questão dos danos causados pela IA.

### **5.2.3 Responsabilidade objetiva pelo fato da coisa**

A terceira e última hipótese de responsabilidade objetiva mencionada pela doutrina na concepção da IA como objeto diz respeito à teoria do fato da coisa ou "teoria da responsabilidade do guarda da coisa inanimada" (RODRIGUES, 1975, p. 91). Tal como nos dois subitens anteriores, a exposição, aqui, será dividida em dois momentos: primeiro, a análise teórica e, em seguida, a possibilidade de aplicação do fato da coisa à IA.

### 5.2.3.1 A teoria da guarda da coisa inanimada e do animal

Conforme refere Justino Magno Araújo, “a expressão ‘responsabilidade pelo fato da coisa’ é gênero que engloba as espécies de danos causados por animais e por coisas inanimadas” (2011). A coisa inanimada “é a que não tem vida própria ou animal, e o exemplo clássico é a máquina, que apesar de inanimada se move por força da criatividade humana” (ARAÚJO, 2011).

A teoria da responsabilidade pela guarda da coisa inanimada surgiu na França, no fim do Século XIX, baseada “em princípio de lei que jazia morto no Código Napoleônico desde há quase um século” (RODRIGUES, 1975, p. 91–92). Conforme previa o art. 1384 do Código Civil francês, “[a] responsabilidade decorre não somente de fato próprio do agente, mas ainda de fatos de pessoas por quem se deve responder, ou das coisas que estão sob a nossa guarda”<sup>95</sup> (RODRIGUES, 1975, p. 94). Silvio Rodrigues comenta, ainda, que:

[...] tal concepção surgiu para atender à ânsia dos juristas franceses que, ante o crescente número de acidentes, resultante da crescente utilização do maquinismo, se preocupavam em proporcionar às vítimas de tais infortúnios um meio mais amplo de obter ressarcimento. Exigir da vítima a prova de culpa do agente causador do dano, o que decorria do sistema tradicional, era impor-lhe pesado encargo, por vezes invencível, hipótese em que ficava o prejuízo irreparado. A doutrina da responsabilidade presumida do guarda da coisa inanimada veio reverter o ônus da prova, além de limitar os casos em que se possibilitava a elisão da presunção acima mencionada. (1975, p. 115)

Sérgio Cavalieri Filho destaca que a responsabilidade pelo fato da coisa só deve ser invocada quando a coisa é a causa direta do dano, “sem a conduta direta do dono ou de seu preposto” (2012, p. 227). Isso não significa dizer que não haja fato próprio daquele que é reputado “guardião”. Como resume Justino Magno Araújo,

[...] fato da coisa é apenas expressão técnica de caráter convencional, pois somente o homem pode responder por, seus próprios atos ou fatos, não se concebendo possa responder por prejuízo causado pelas coisas, a não ser que exista uma relação de causa e efeito entre um fato seu (positivo ou negativo) e o fato da coisa. (ARAÚJO, 2011)

Igualmente relevante para a teoria é a noção de “guarda”. Silvio Rodrigues comenta que a jurisprudência francesa desde logo entendeu que “a simples detenção material de uma coisa não bastava para caracterizar a figura do guarda. A responsabilidade derivaria, *não da guarda*

---

<sup>95</sup> No original, “*On est responsable non seulement du dommage que l'on cause par son propre fait, mais encore de celui qui est causé par le fait des personnes dont on doit répondre, ou des choses que l'on a sous sa garde.*” O enunciado passou a constar do art. 1242 após a reforma do direito dos contratos e obrigações ocorrida em 2016 (FRANÇA, 1804).

*material da coisa, mas da guarda jurídica*<sup>96</sup> (1975, p. 116). Caio Mário da Silva Pereira aponta que recai sobre o proprietário da coisa uma “presunção de guarda”, que pode ser afastada se o proprietário provar que a guarda da coisa é exercida por terceiro com o seu consentimento ou autorização, “ou ainda quando o terceiro se apossa dela no desconhecimento ou contra a vontade do proprietário” (2018b). Trata-se, pois, de presunção *iuris tantum*, cabendo ao julgador, caso a caso, “examinar se subsiste a ‘presunção de guarda’ imposta ao proprietário, ou se, reversamente, foi ilidida a *praesumptio* como no caso de transferência da coisa, com direito à sua utilização, com suficiente independência” (2018b).

Nesse cenário, uma discussão relevante que se apresenta é se a responsabilidade do proprietário permanece mesmo quando a coisa lhe tiver sido furtada ou de alguma forma subtraída de sua guarda (RODRIGUES, 1975, p. 117).

José de Aguiar Dias defendia que o furto ou roubo da coisa não impedia a responsabilização do proprietário (ARAÚJO, 2011; CAVALIERI, 2012, p. 230). Nada obstante, prevaleceu na doutrina<sup>97</sup> o entendimento de que o proprietário responde pelos danos somente quando a perda da guarda se deu por culpa sua (CAVALIERI, 2012, p. 231). Assim, em casos de furto ou roubo, não há falar em responsabilidade do proprietário.

Embora o Código Civil de 1916 não possuísse nenhuma cláusula similar ao que previa o art. 1384 do Código Civil francês, situação que se manteve no Código Civil de 2002, aliás, a doutrina e a jurisprudência brasileira foram paulatinamente incorporando a teoria da responsabilidade pela guarda da coisa inanimada ao direito pátrio. Atualmente, a teoria encontra correspondência nos artigos 936, 937 e 938 do Código Civil<sup>98</sup>.

<sup>96</sup> Sérgio Cavalieri Filho refere-se à guarda jurídica também pelo termo “guarda intelectual” (2012, p. 228).

<sup>97</sup> “A distinção entre a guarda jurídica e a guarda material não tem fundamento sólido e é contrária à própria significação da palavra *guarda*, que supõe um poder de vigilância sobre a coisa e meios para evitar que esta venha a causar danos a terceiros. Não se compreende guarda quando o controle da coisa se torna impossível de ser exercido.” (RODRIGUES, 1975, p. 119). Na mesma linha, posiciona-se Justino Magno Araújo (2011): “se o proprietário perder a guarda da coisa, não mais responde como guardião, pois não mais conserva aquela condição. Se, entretanto, a guarda foi perdida por imprudência ou negligência, permanece responsável, aplicando-se aqui a regra do art. 159 do CC. Nesse caso, caberá à vítima o ônus da prova de que o proprietário agiu culposamente.”

<sup>98</sup> A redação atual consolidou entendimento antes controvertido sobre a natureza da responsabilidade do guardião do animal em comparação com a do guardião da coisa. Isso porque, diferentemente dos artigos. 1.528 e 1.529 do Código Civil de 1916, cuja redação foi integralmente reproduzida no Código atual, o art. 1.527 previa outras duas causas excludentes de responsabilidade, dando a entender que o dono do animal tinha mera presunção de culpa. Ao examinar a aplicabilidade do art. 1.527, bem como a redação proposta no Projeto de Código de Obrigações de 1965 e o Anteprojeto de Código Civil de 1972, Silvio Rodrigues apontou que as duas hipóteses eram, na verdade, idênticas, resultando, ao fim e ao cabo, na exclusão da responsabilidade apenas por caso fortuito ou força maior (1975, p. 146), daí porque passou-se a fazer abordagem conjunta da responsabilidade pela guarda da coisa e do animal.

Apesar do Código Civil tratar expressamente apenas da responsabilidade (*i.*) do dono ou detentor do animal<sup>99</sup>; (*ii.*) do proprietário do edifício ou da construção; e (*iii.*) daquele que habitar prédio, a aplicação a teoria da responsabilidade pela guarda da coisa não se limita a essas hipóteses. Quando do julgamento do Recurso Especial n.º 577.902/DF, a Terceira Turma do STJ entendeu pela responsabilidade solidária do proprietário de automóvel, seja na hipótese de transporte gratuito ou oneroso. Conforme destacou a Ministra Nancy Andrighi, Redatora para o acórdão, “o proprietário do automóvel é a pessoa que tem a guarda jurídica da coisa, isto é, aquele quem exerce um poder de comando em relação à coisa, de direção intelectual, de dar ordens, relativamente à coisa”. Consequentemente, “se o proprietário descarta-se da guarda de seu veículo – que é, repise-se, instrumento potencialmente muito perigoso –, entregando a sua direção à pessoa sem condições de utilizá-lo e que acaba causando um acidente, deve responder solidariamente com essa pessoa”<sup>100</sup> (BRASIL, 2006).

Nesse ponto, é pertinente o comentário de Sérgio Cavalieri Filho:

Seria ilógico, todavia - como bem observa Aguiar Dias -, responsabilizar o proprietário do animal e do imóvel e não responsabilizar, em medida igual, o guarda das demais coisas. Temos, então, que nos valer da mesma regra adotada nos arts. 936, 937 e 938, pois, até por uma questão de lógica, o princípio aplicado ao dono do animal e do edifício deve ser analogicamente estendido ao guarda da coisa em geral; ao guardião da coisa deve ser reconhecida a mesma situação do guardião do animal. E do exame daqueles dispositivos, principalmente pela redação que recebeu o art. 936, não resta a menor dúvida de que o Código de 2002 estabeleceu responsabilidade objetiva para o proprietário do animal e do edifício. Eles têm uma obrigação de resultado, isto é, estão obrigados não apenas a guardar a coisa (animal ou edifício), mas a guardá-la com segurança, de modo a que jamais escape do seu controle e, em consequência, ocasione um dano a terceiro. (2012, p. 234)

Assim, ainda que não haja previsão específica no Código Civil para todas as situações de guarda, reconhece-se a possibilidade de responsabilização civil do proprietário pelo fato da coisa inanimada, porque possui uma obrigação de resultado, da qual decorre a dispensa da demonstração de culpa. A natureza da obrigação do guarda é fundamental, porque, conforme sintetiza Silvio Rodrigues, “dentro da teoria da guarda só se permite a exoneração do guarda se lograr provar caso fortuito, força maior ou culpa da vítima, não podendo desvencilhar-se do

<sup>99</sup> Ainda quanto a esta hipótese, é interessante pontuar que “a norma objetiva os animais domésticos, só sendo aplicável aos animais selvagens que tenham sido apropriados pelo homem e estejam sob sua guarda, como animais de jardins zoológicos, circos e outros que estão vivendo em cativeiro.” (CAVALIERI, 2012, p. 243)

<sup>100</sup> Na mesma linha, o STJ igualmente sedimentou entendimento baseado na teoria da guarda da coisa inanimada para fundamentar a responsabilidade solidária do proprietário de reboque/semirreboque com o proprietário do cavalo mecânico, a despeito de inexistir previsão específica no Código Civil. Conforme restou decidido quando do julgamento do Agravo Interno no Recurso Especial n.º 1.548.722/TO, Relator: Ministro Antônio Carlos Ferreira, em 08 abr. 2019; e do Agravo Interno no Recurso Especial n.º 1.870.557/GO, Relator: Ministro Paulo de Tarso Sanseverino, em 21 set. 2020.

dever de ressarcir o dano, mediante mera demonstração de que agiu diligentemente” (RODRIGUES, 1975, p. 137).

Portanto, se adotada a premissa de que a IA é uma “coisa”, pela teoria da guarda da coisa inanimada, a imputação do dever de indenizar pelos danos causados pela IA deve recair sobre o proprietário da IA (e não o fabricante ou, eventualmente, o usuário), sob o fundamento de que é o proprietário quem tem o dever de garantir a segurança de sua utilização.

### 5.2.3.2 *Fato da coisa para a inteligência artificial?*

Compreendido o escopo da responsabilidade pelo fato da coisa no direito civil, passa-se a analisar a sua aplicabilidade especificamente aos casos relacionados à responsabilização por danos causados pelo uso de IA. Duas questões são importantes nesse sentido: (i.) a possibilidade de equiparar a IA às coisas e animais que estão sob guarda; e (ii.) se o guardião é, efetivamente, a pessoa mais adequada para responder pelos atos da IA.

Quanto à primeira questão, Gustavo Tepedino e Rodrigo da Guia Silva entendem que a responsabilidade pela guarda da coisa poderia servir como fundamento para a aplicação do regime objetivo de responsabilidade civil à IA:

Sustenta-se a existência de similar ordem de inteligência e de imprevisibilidade tanto dos animais quanto dos sistemas inteligentes. Já no que tange ao paralelo com a guarda da coisa inanimada, afirma-se que tanto as coisas já conhecidas de longa data quanto os sofisticados sistemas inteligentes consistem em bens sob custódia de uma pessoa, que por seus atos deve responder. Em qualquer dos casos – seja em relação a animais ou a coisas inanimadas –, destaca-se a opção do legislador por imputar, independentemente de culpa (ou em regime de culpa presumida), a responsabilidade civil à pessoa sob cuja custódia estejam os animais ou as coisas. (TEPEDINO; SILVA, 2019, p. 81–82)

Ugo Pagallo aponta ao menos três motivos pelos quais a IA deveria ser equiparada a animais, e não a produtos ou coisas: (i.) robôs tem capacidade de interação e percepção do ambiente e são capazes de responder a estímulos e alterar os próprios valores; (ii.) robôs são autônomos e são capazes de exercer controle sobre as próprias ações independentemente da intervenção humana; e (iii.) robôs são adaptáveis e podem se auto aperfeiçoar (2013, p. 38).

Henrique Sousa Antunes destaca a possibilidade de aplicação das regras atinentes à responsabilidade decorrente da falta de vigilância da coisa ou do animal à IA, pois “[e]nquanto ao robô faltar personalidade jurídica, a sua natureza confundir-se-á com uma dessas duas realidades” (2019, p. 147). Pelo mesmo motivo, seria “desajustado configurar a

responsabilidade do utilizador pelos danos causados por um robô no plano da responsabilidade das pessoas obrigadas à vigilância de outrem” (ANTUNES, 2019, p. 147).

Todavia, Antunes pontua que essa seria uma solução provisória, porque a “sensibilidade coletiva cedo rejeitará a aparência de escravidão que a relação descrita sugere” (2019, p. 147). Em linha semelhante, tem-se a oposição apresentada por Čerka, Grigiene e Sirbikyte quanto à equiparação da IA aos animais:

Não há fundamentos para igualar a IA a um animal porque as atividades da IA são baseadas em um processo algorítmico semelhante ao pensamento humano racional e apenas parcialmente semelhante aos instintos e sentidos como os dos animais. Presume-se que a IA pode compreender as consequências de suas ações e se distinguir dos animais. Isso leva à conclusão de que não podemos aplicar a responsabilidade objetiva que seria aplicada nos casos em que o dano é causado por um animal.<sup>101</sup> (2015, p. 386)

Embora não mencionem expressamente a equiparação da IA à uma coisa inanimada, o motivo para a inaplicabilidade da teoria parece ser o mesmo: se a IA é mais inteligente que animais sencientes, evidentemente que não poderia ser comparada à coisas que não possuem nenhuma sensibilidade e, muito menos, qualquer capacidade de ação autônoma.

As três posições acima mencionadas têm em comum o ponto de partida, qual seja, a preocupação com o nível de inteligência da IA, do qual decorrem, também, as divergências. E esse ponto de partida é de extrema relevância, pois, conforme já destacado acima, o fato nunca é “da coisa”; a “coisa é mero instrumento do dano, sendo sua causa a omissão humana, por falta de vigilância ou cuidado” (CAVALIERI, 2012, p. 227). Nesse sentido, enquanto a IA for mero instrumento dos humanos, mal não há em equipará-la à “coisa” sob vigilância.

O imbróglio surge quando a IA for evidentemente mais inteligente que animais – e mesmo humanos – sendo capaz de causar danos independentemente da vigilância de seu guardião. Nesse caso, surge o segundo questionamento inicialmente posto com relação à adequação da teoria da guarda da coisa inanimada para assegurar a solução de casos envolvendo danos causados pela IA: até que ponto o guardião da IA é capaz de garantir a sua utilização segura? Ou, em outras palavras: sendo a obrigação de guarda uma obrigação de resultado (que independe, portanto, do grau de diligência daquele que deve prestar), até que ponto ela poderá ser factível e razoavelmente desempenhada pelo proprietário da IA?

---

<sup>101</sup> No original, “*There are no grounds to equate AI to an animal because the activities of AI are based on an algorithmic process similar to rational human thinking and only partially similar to instincts and senses like those of animals. It is presumed that AI can understand the consequences of its actions and distinguish itself from animals. This leads to the conclusion that we cannot apply strict liability which would be applied in cases where the damage is caused by an animal.*”

Via de regra, o proprietário é quem adquire a coisa para seu próprio benefício. Um dos exemplos mais comuns de responsabilidade pela guarda da coisa, já mencionado neste trabalho, é o do proprietário do veículo de circulação terrestre. Essa responsabilidade é razoável porque se presume, por exemplo, que o proprietário possui habilitação para dirigir (ou conta com os serviços de alguém apto para tanto) e utilizará o veículo de acordo com as normas de trânsito. Com exceção da realidade emergente dos veículos autônomos, até então, veículos não transitam sem um condutor, isto é, um veículo tem o potencial de causar danos, mas só o fará se for utilizado por alguém.

A mesma lógica não se aplica à IA. Diferentemente de veículos (novamente, com exceção dos autônomos), a IA é capaz de agir por meio de processos algorítmicos, desempenhando atividades independentemente de um “condutor”. Algoritmos de *deep learning* são capazes de processar informações sem que seus desenvolvedores conheçam sequer a sequência de instruções que possuem. A tendência da IA é se tornar cada vez mais desenvolvida e independente, e isso a afasta sobremaneira da noção que autoriza a responsabilidade civil pelo fato da coisa, exatamente porque, no caso de uma IA autônoma, o fato será justamente “da coisa”, e não uma decorrência da omissão humana ou falha de vigilância.

Um contraponto que pode ser arguido é o argumento de Ugo Pagallo, já referido quando da análise da teoria da responsabilidade pelo fato do produto, de que os usuários da IA estariam em melhores condições de verificar um possível comportamento ilícito da IA em comparação com os fabricantes, que tiveram contato apenas com o sistema nativo (2013, p. 126). Porém, “usuário” não equivale, necessariamente, ao dono da IA e, para a teoria da guarda da coisa, o que importa é a guarda jurídica, não a guarda material. Aqui, seria necessário equiparar o regime aplicável à IA à hipótese de responsabilidade pelos danos causados por animais, em que o Código Civil prevê expressamente a responsabilidade do “detentor”.

O argumento é fraco, todavia, porque não se coaduna com o pressuposto essencial da responsabilidade pelo fato da coisa, que é a possibilidade de o guardião agir de modo a evitar o evento danoso (porquanto a responsabilidade só se afasta nas hipóteses de caso fortuito, força maior e fato exclusivo da vítima). Ocorre que, se o desenvolvimento da IA for tal, a ponto de o guardião não ser capaz de evitar ou impedir determinado evento danoso, haveria então uma causa excludente do nexo de causalidade que impediria a imputação do dever de indenizar, porque equiparada ao caso fortuito<sup>102</sup>.

---

<sup>102</sup> A definição de caso fortuito consta do parágrafo único do art. 393 do Código Civil: “O caso fortuito ou de força maior verifica-se no fato necessário, cujos efeitos não era possível evitar ou impedir.”

Nesses termos, o proprietário que adquire um robô e não possui nenhum conhecimento de computação pode estar totalmente desprovido de meios de desempenhar uma obrigação de vigilância. Entende-se que não se trata apenas de equiparar a IA aos animais, mas também de considerar se a ingerência dos humanos sobre a IA pode ser razoavelmente equiparada à ingerência sobre as coisas e animais, a ponto de autorizar a aplicação da teoria ora examinada.

Ao fim e ao cabo, a depender do nível de inteligência da IA, o fundamento da responsabilidade se torna alheio ao fato da coisa. Assim, ou se retoma a discussão sobre fato do produto (cujas deficiências foram apontadas anteriormente), ou se prossegue para o reconhecimento da IA como um ser inteligente equiparável aos humanos. Essa segunda perspectiva pode resultar em duas soluções, discutidas a seguir: (i.) a responsabilidade indireta, por fato de terceiro; ou (ii.) a responsabilidade direta da IA, como um sujeito de direitos.

### 5.3 Responsabilidade Indireta ou por fato de outrem: a IA como agente

Ao tratar da responsabilidade civil, Ugo Pagallo propõe a distinção da IA em dois tipos: “*robots-as-means*” (ou “robôs como meios”) e “*robots-as-agents*” (ou “robôs como agentes”) (2013, p. 136). Enquanto aqueles seriam equiparados a produtos, atraindo a aplicação da responsabilidade pelo fato do produto e a imputação do dever de indenizar ao fabricante; estes seriam equiparados a crianças, animais ou empregados, atraindo a responsabilidade do guardião ou empregador (PAGALLO, 2013, p. 136–337). É sobre esse segundo grupo que se analisa a teoria da responsabilidade indireta.

Para ilustrar a proposta, Pagallo menciona três exemplos de robôs de uso doméstico: um robô desenvolvido para brincar com crianças (“*robot-toy*”), um robô-babá (“*robot-nanny*”) e um robô que gerencia as contas da casa (o modelo i-Jeeves 2.0) (2013, p. 129). Como o robô está à disposição ou sob supervisão de seu dono, agindo em seu benefício, isso justificaria a responsabilidade do proprietário (ČERKA; GRIGIENE; SIRBIKYTE, 2015, p. 385).

Traçando um paralelo entre esses robôs inteligentes e empregados, Pagallo aponta que seria possível responsabilizar os humanos, de modo objetivo, pelos danos causados pelo que ele refere como uma nova geração de empregados baseados em IA (2013, p. 130). O fundamento jurídico para essa responsabilização seria a teoria da *vicarious liability*<sup>103</sup> dos países da *Common Law* e, no direito italiano, nas hipóteses de responsabilidade dos pais pelos

---

<sup>103</sup> Em sistemas de *Common Law*, a responsabilidade por fato de outrem é denominada “*vicarious liability*” (GILIKER, 2010, p. 5).

atos dos filhos menores e do empregador pelos atos dos seus empregados, previstas, respectivamente, nos artigos 2048 e 2049 do Código Civil italiano (ITÁLIA, 1942).

A responsabilidade por fato de outrem consiste na imputação do dever de indenizar a alguém que não deu causa direta ao evento danoso. Todavia, para que isso ocorra, “é preciso que esse alguém esteja ligado por algum vínculo jurídico ao autor do ato ilícito, de sorte a resultar-lhe, daí, um dever de guarda, vigilância ou custódia” (CAVALIERI, 2012, p. 204). Nesse sentido, o art. 932 do Código Civil elenca cinco hipóteses nas quais se aplica a teoria da responsabilidade indireta<sup>104</sup>, sendo duas delas de especial interesse para a análise da proposta de Ugo Pagallo (*i.*) a responsabilidade dos pais pelos filhos menores e; (*ii.*) do empregador ou comitente pelos empregados, serviçais ou prepostos.

Conforme pontua Silvio Rodrigues, a “ideia principal que justifica a responsabilidade por fato de terceiro é o anseio de garantia que deve inspirar o legislador nesse campo. Cumpre-lhe assegurar a reparação do dano causado sempre que possível” (1975, p. 68). Para Rodrigues, o risco é o fundamento que melhor se amolda à realidade no que diz respeito à responsabilidade pelo fato de outrem. Afirma o jurista que tanto o pai que põe o filho no mundo quanto o patrão que se beneficia dos serviços do empregado correm o risco de que tais pessoas deem causa a danos, razão pela qual seria “razoável que, se tal dano advier, por ele respondam solidariamente com seus causadores diretos, aqueles sob cuja dependência estes se achavam” (RODRIGUES, 1975, p. 68).

Sérgio Cavalieri Filho pontua, de outro lado, que o fundamento dessa responsabilidade não seria o risco, mas, sim, “o *dever objetivo de guarda e vigilância* legalmente imposto aos pais, tutores e curadores”, sendo que tal dever “é exigível daquele que tem autoridade sobre outrem, e enquanto o tiver em sua companhia” (2012, p. 206).

Em estudo sobre a responsabilidade indireta em países da *Common Law* e, também, no direito francês e alemão, Paula Giliker (2010, p. 230) apontou que há, com efeito, diversos fundamentos para justificar a imputação de responsabilidade a um terceiro que não deu causa direta ao dano, sendo os três principais a culpa *in vigilando* (falha na vigilância dos

---

<sup>104</sup> Paula Giliker aponta que a relação entre empregador e empregado é o exemplo mais comum da *vicarious liability* (2010, p. 101). Porém, a situação em que uma pessoa age no interesse de outra também tem ensejado outras hipóteses de responsabilidade indireta, especialmente em sistemas de *Civil Law*, em que as previsões adicionais estão previstas nos códigos (GILIKER, 2010, p. 101). Exemplo disso é o próprio art. 932 do Código Civil de 2002, que, além da responsabilidade do empregador, elenca outras quatro situações: (*i.*) os pais, pelos filhos menores que estiverem sob sua autoridade e em sua companhia; (*ii.*) o tutor e o curador, pelos pupilos e curatelados, que se acharem nas mesmas condições; (*iii.*) os donos de hotéis, hospedarias, casas ou estabelecimentos onde se albergue por dinheiro, mesmo para fins de educação, pelos seus hóspedes, moradores e educandos; e (*iv.*) os que gratuitamente houverem participado nos produtos do crime, até a concorrente quantia. (BRASIL, 2002).

subordinados) ou *in eligendo* (falha na escolha dos empregados ou prepostos)<sup>105</sup>, a compensação do dano suportado pela vítima e o risco<sup>106</sup>. Porém, a autora conclui que nenhum fundamento, isoladamente, explica a responsabilidade indireta e que “o desafio para um sistema jurídico é, portanto, encontrar a combinação correta de motivos políticos. Ao utilizar a doutrina da responsabilidade vicária, cada sistema jurídico deve decidir qual peso dar às motivações concorrentes de culpa, compensação da vítima e risco/dissuasão.”<sup>107</sup> (GILIKER, 2010, p. 244, tradução nossa)

No Brasil, houve verdadeira evolução no tratamento da matéria na comparação do Código Civil de 1916, que mantinha um regime baseado na culpa presumida. Embora o art. 932 trate das mesmas hipóteses já elencadas no art. 1.521 do código anterior (BRASIL, 1916), o art. 933 afastou qualquer discussão acerca da natureza objetiva da responsabilidade aqui discutida<sup>108</sup> (BRASIL, 2002).

Deve-se pontuar, contudo, que a responsabilidade será objetiva com relação àquele contra quem se imputa o dever de indenizar, e não com relação a quem dá causa direta ao dano (CAVALIERI, 2012, p. 206). Analisando os requisitos para a responsabilidade indireta do empregador, Silvio Rodrigues aponta: “a) culpa do empregado; b) relação de emprego ou de dependência do agente direto do dano, para com o patrão, amo ou comitente; c) que o ato danoso do preposto se dê no exercício do trabalho, ou por ocasião dele.” (1975, p. 75). O jurista destaca que “se o empregado que causou o dano não teve culpa, nem ele nem seu patrão devem responder”<sup>109</sup> (RODRIGUES, 1975, p. 75).

Na mesma linha, é o comentário sobre a responsabilidade dos pais:

Se o ato danoso do menor não foi culposo, é extremamente provável que o dano experimentado pela vítima tenha resultado de força maior ou caso fortuito. Ademais, como uma das excludentes da responsabilidade é a ausência de culpa, parece inegável que se não houve culpa do agente causador do dano, isto é, do menor, obviamente não se pode admitir a responsabilidade de seus pais. (RODRIGUES, 1975, p. 71)

<sup>105</sup> Giliker menciona que a culpa *in eligendo* é o fundamento previsto no Código Civil alemão, por exemplo, em que a responsabilidade indireta do empregador é baseada na culpa presumida, que pode ser afastada se o empregador demonstrar que escolheu o empregado com diligência (2010, p. 231).

<sup>106</sup> Quanto a este fundamento, Giliker destaca que a sua compreensão também exige que se considere qual o papel da responsabilidade civil na sociedade, dado que o fundamento do risco exige uma mudança da noção de justiça corretiva, baseada na culpa do indivíduo, para uma noção de justiça distributiva, em que objetivos comuns da sociedade justificam o afastamento da culpa como pressuposto da responsabilidade (GILIKER, 2010, p. 239).

<sup>107</sup> No original, “*The challenge for a legal system is therefore reaching the correct combination of policy rationales. In utilising the doctrine of vicarious liability, each legal system must decide what weight to give to the competing policies of fault, victim compensation and risk/deterrence.*”

<sup>108</sup> Conforme dispõe o art. 933, “as pessoas indicadas nos incisos I a V do artigo antecedente, ainda que não haja culpa de sua parte, responderão pelos atos praticados pelos terceiros ali referidos.”

<sup>109</sup> A necessidade de culpa é igualmente mencionada pela Súmula 341 do Supremo Tribunal Federal: “É presumida a culpa do patrão ou comitente pelo ato culposo do empregado ou preposto.” (BRASIL, 1963)

Porém, no caso de menores de pouca idade, incapazes de distinguir o certo do errado, Rodrigues aponta que a solução deveria ser idêntica à responsabilidade dos “amentais” (termo utilizado quando da vigência do Código Civil de 1916): “[s]e o infante causou dano a outrem, as pessoas por ele responsáveis devem ser compelidas a indenizar” (1975, p. 72). Nesse caso, a responsabilidade indireta e objetiva dos pais prescinde da demonstração de ato culposo do filho menor<sup>110</sup>. O art. 932, I, do Código Civil exige, de todo modo, que o filho menor estivesse sob a autoridade do genitor e em sua companhia (BRASIL, 2002).

Nesse contexto, a primeira dificuldade de se aderir à proposta de Pagallo à luz do Código Civil brasileiro está na questão da culpa da IA. Conforme comentado acima, a responsabilidade indireta consiste na imputação objetiva do dever de indenizar a um terceiro, sob a condição de que se demonstre a culpa de quem deu causa direta ao dano. Ocorre que, conforme já apontado no item 5.2, acima, a análise do elemento subjetivo na conduta adotada pela IA pode consistir em efetiva “prova diabólica”, seja porque implica em um ônus que a vítima poderá ter verdadeira dificuldade de superar (ČERKA; GRIGIENE; SIRBIKYTE, 2015, p. 387), seja pela dificuldade de se analisar a culpa de um robô que não possui a mesma dimensão ética dos humanos (ANTUNES, 2019, p. 153).

A alternativa, então, seria equiparar a IA ao filho menor de pouca idade ou aos absolutamente incapazes, dispensando-se a análise da culpa. Contudo, parece um tanto quanto contraditório afastar a responsabilidade pelo fato da coisa porque a IA seria intelectualmente superior a objetos e animais, mas equiparar sua inteligência a de um bebê humano, incapaz de discernir o comportamento lícito do ilícito. Ora, se a IA não é capaz de julgar a licitude de seu comportamento, qual o sentido de não a equiparar a um animal doméstico, por exemplo?

Para evitar a discussão sobre o reconhecimento de uma personalidade jurídica própria aos robôs e, ao mesmo tempo, viabilizar uma solução para casos de responsabilidade civil envolvendo IA, Jan-Erik Schirmer propôs a utilização do conceito de *Teilrechtsfähigkeit*<sup>111</sup>, definido na década de 1930 por Hans-Julius Wolff como “o estatuto aplicável a um humano ou

---

<sup>110</sup> Paula Giliker comenta que, em 1984, a Suprema Corte da França reconheceu a responsabilidade dos pais por todos os atos dos filhos, independentemente da inabilidade da criança de compreender que sua ação era errada (2010, p. 29–30). Com efeito, o art. 1384(5) do Code Civil (FRANÇA, 1804) não fala, expressamente, na necessidade da prova da culpa do causador direto do dano. Em 2004, porém, a Suprema Corte reafirmou a necessidade de prova da culpa, consolidando o entendimento antes alterado (GILIKER, 2010, p. 30).

<sup>111</sup> Schirmer (2019, p. 134) comenta que o conceito foi utilizado durante o regime nazista para justificar a exclusão de liberdades civis dos judeus ao mesmo tempo em que lhes eram impostas diversas obrigações. Um dos principais juristas a fazer uso dessa técnica foi Karl Larenz, que, após o fim da guerra, juntou-se ao grupo que defendeu a falha da personalidade “pela metade” e o retorno ao sistema binário (ou se tinha total capacidade jurídica, ou não se tinha nenhuma).

a uma associação de humanos com capacidade jurídica apenas de acordo com previsões legais específicas, mas sem possuir deveres ou direitos”<sup>112</sup> (2019, p. 134, tradução nossa).

À semelhança do que ocorreu com o reconhecimento expresso de que animais não são coisas pelo Código Civil alemão, em 1990, Schirmer sugeriu a inclusão de uma regra nos seguintes termos: “Agentes inteligentes não são pessoas. De acordo com a sua função de serviço, eles serão regrados pelas disposições aplicáveis aos agentes, com as modificações necessárias, salvo disposição em contrário”<sup>113</sup> (2019, p. 140, tradução nossa). O autor sustenta que, a partir de tal regra, (i.) seria possível evitar o dilema do reconhecimento de uma personalidade jurídica própria à IA; (ii.) agentes inteligentes conservarão capacidades jurídicas consistentes com as suas funções, identificadas pelos tribunais e pela academia; e (iii.) a regra assegura que a IA permanecerá sendo um objeto na maior parte do tempo, o que significa que ela ainda poderia ser negociada como qualquer outro bem (SCHIRMER, 2019, p. 140).

A responsabilidade civil desses robôs “parcialmente capazes” se resolveria, então, pela doutrina do *respondeat superior*, isto é, a ideia de que o “mestre” deve responder objetivamente pelos atos de seu subordinado, que, no caso, seria a IA (SCHIRMER, 2019, p. 137). Mas a doutrina do *respondeat superior*, muitas vezes associada aos fundamentos da *vicarious liability* (GILIKER, 2010, p. 228), exige a demonstração de que o subordinado agiu com negligência. Schirmer sugere, nesse ponto, a utilização de um regime inteiramente objetivo de responsabilidade civil, imputando-se o dever de indenizar ao usuário independentemente da culpa da IA (2019, p. 138).

A proposta de Schirmer de uma “capacidade jurídica parcial” baseada na doutrina do *respondeat superior* parece, todavia, equiparar robôs a escravos<sup>114</sup>, que servem aos interesses dos humanos, mas poderão ser negociados como mercadoria. Trata-se de uma medida paliativa, que pode se mostrar insuficiente no futuro, especialmente com o desenvolvimento de robôs que

---

<sup>112</sup> No original, “the status applicable to a human or an association of humans having legal capacity only according to specific legal rules, but otherwise not bearing duties and having rights”.

<sup>113</sup> No original, “Intelligent agents are not persons. Consistent with their serving function, they are governed by the provisions that apply to agents, with the necessary modifications, except insofar as otherwise provided”.

<sup>114</sup> Conforme comentam Čerka, Grigiene e Sirbikyte (2015, p. 385), “Analysis of the legal status of AI and its liability for its actions reveals a link between the status of AI and that of slaves established in the Roman law. Both AI and the slave are not subjects of law, but rather its objects. They could not apply to courts, because only free persons could participate in litigation. Assuming that the parallel between legal status of AI and that of slaves is possible, it can be stated that damages caused by the actions of AI should be compensated by its owner or AI developer the legal person on whose behalf it acts. In Roman law, this meant that the person (head of household) responsible for persons *alieni iuris* (subordinate slaves), i.e. their owner, was held liable for torts committed by the slaves.”

estarão próximos dos humanos e poderão despertar afeição<sup>115</sup>, como os robôs cuidadores de idosos e crianças, por exemplo<sup>116</sup>, ensejando um possível movimento de “libertação” dos robôs ou mesmo de ampliação dos direitos que lhes são reconhecidos.

Não se pode negar, de outro lado, que a proposta de Schirmer permite superar os dois principais óbices para a aplicação da teoria da responsabilidade indireta como fundamento para a responsabilidade civil dos entes dotados de IA: o regime completamente objetivo de responsabilidade indireta afasta o problema da prova da culpa da IA e a “capacidade jurídica parcial” permite considerá-la como um agente, mas não como uma pessoa. Ainda assim, seria inevitavelmente necessário proceder a uma alteração legislativa, restando claro, portanto, que o art. 932 do Código Civil, em sua redação original, é insuficiente para resolver a responsabilidade civil da IA.

Portanto, as atuais regras de responsabilidade indireta não resolvem os casos de responsabilidade civil relacionados à IA autônoma, havendo, assim, uma lacuna normativa. No entanto, a teoria tem méritos relevantes e permitiria reconhecer a IA autônoma como um agente, ainda que não humano, superando diversos dos óbices até aqui apontados. Restaria apenas a discussão sobre a imputação apenas sobre o proprietário ou, também, ao desenvolvedor.

#### 5.4 Responsabilidade objetiva pelo art. 187 do Código Civil

Conforme visto até aqui, todas as teorias de responsabilidade extra negocial que partem da premissa de que a IA autônoma é um objeto do direito apresentam algum tipo de lacuna que demanda alteração legislativa ou desenvolvimento dogmático para assentar a interpretação de determinado termo que é elemento necessário para a aplicação da respectiva regra. Isso não quer dizer que todos os modelos sejam insuficientes, mas apenas que a sua aderência às situações em que os danos indenizáveis sejam causados pela IA autônoma pode encontrar algumas dificuldades.

---

<sup>115</sup> Há uma observação interessante a respeito do desenvolvimento de afeição por um robô. Na década de 1970, o professor Masahiro Mori cunhou o termo “*uncanny valley*” para descrever o fato de que a afeição por robôs (especialmente andróides) tende a ser maior à medida que o robô se torna mais parecido com os humanos, até o ponto em que essa afeição se transforma em medo e repulsa (*uncanny valley*). O professor Christoph Bartneck, da University of Canterbury, Nova Zelândia, apontou, todavia, que o “*uncanny valley*” pode acabar por se tornar um “*uncanny cliff*”, já que a afeição inicial não é retomada. Disponível em: <https://spectrum.ieee.org/automaton/robotics/humanoids/what-is-the-uncanny-valley>. Acesso em: 08 nov. 2020.

<sup>116</sup> Os robôs cuidadores são mencionados também no princípio 31 da Resolução do Parlamento Europeu de 16 de fevereiro de 2017 (UNIÃO EUROPEIA, 2017). Sobre este ponto, Silvia Díaz Alabart aponta que “[s]e pide especial atención al riesgo de que las personas más vulnerables – niños, ancianos y personas con discapacidad – puedan desarrollar un vínculo emocional con los robots, y los consiguientes posibles trastornos psicológicos para los humanos. Es evidente que en el ámbito en que más fácilmente puede desarrollarse dicho vínculo emocional es en el de los robots asistenciales.” (2018, p. 52)

Nada obstante, se todos os modelos anteriores se mostrarem insuficientes, há ainda uma possível solução normativa aplicável às situações em que a IA autônoma causar danos: a responsabilidade objetiva lastreada no art. 187 do Código Civil.

É importante referir que a discussão relativa ao art. 187 desloca a tônica até aqui atribuída à *responsabilidade* para a noção de *ilicitude*. E a separação desses dois conceitos é fundamental, porque a existência de um ato reputado *ilícito*, nos termos do art. 187, não tem como consequência, necessariamente, a *responsabilidade civil* (que dependerá, ainda, da existência de um dano e do nexo de causalidade). Conforme comenta Judith Martins-Costa, “a estrutura hoje codificada mostra que a obrigação de indenizar ganhando autonomia e estando agora situada em título próprio, *liberou a reconstrução do conceito de ilicitude*” (MARTINS-COSTA, 2006, p. 19).

Daí porque Maria Cláudia Cachapuz destaca que o art. 187 surge como uma verdadeira fonte de obrigações civis e, por isso:

[...] a ilicitude no novo Código Civil alcança hipótese normativa mais ampla e com abertura suficiente para não apenas amparar a obrigação de reparação, mas formas diversas de obrigações civis - como a hipótese do estabelecimento de uma obrigação proibitiva ao exercício de determinada posição jurídica. (CACHAPUZ, 2011)

Consequentemente, o art. 187 pode servir de fundamento não apenas para a tutela reparatória, como também para tutela inibitória, mandamental o de remoção do ilícito (MARTINS-COSTA, 2006).

Conforme já mencionado no item 4.4, parte da doutrina civilista entende que o art. 187 consagrou a teoria do abuso de direito, à semelhança do art. 334º do Código Civil Português<sup>117</sup>. Houve, também, discussão quanto à necessidade ou não de um elemento subjetivo (o ato emulativo) para a configuração do abuso de direito<sup>118</sup>. Todavia, Cachapuz (2011) aponta que o enunciado deve ser interpretado a partir de uma teoria externa de restrição às liberdades, e não

<sup>117</sup> “É ilegítimo o exercício de um direito, quando o titular exceda manifestamente os limites impostos pela boa fé, pelos bons costumes ou pelo fim social ou económico desse direito.” (PORTUGAL, 1966).

<sup>118</sup> Conforme explica Judith Martins-Costa (2006, p. 9-11), “Foi só na década da virada para o séc. XXI que a jurisprudência passou a caminhar - ainda que em passos vacilantes - no sentido da objetivação do abuso. Seja em razão das grandes mudanças legislativas ocorridas desde 1988 (data da vigência da Constituição Federal) e em especial nos anos 90 (quando se aprofunda o processo de descodificação civil, principalmente pela edição do Código de Defesa do Consumidor e sua noção objetiva de abusividade contratual fundada no “significativo desequilíbrio” em detrimento do consumidor); seja, enfim, por força da criação e do papel desempenhados desde então pelo Superior Tribunal de Justiça, o fato é que começam a pipocar arestos a invocar o abuso prescindindo dos elementos subjetivos sendo digno de especial nota o acolhimento jurisprudencial da teoria da desconsideração da personalidade jurídica no caso de abuso da personalidade. Mas, mesmo então, nenhuma referência há aos elementos de concreção do abuso que viriam a ser consagrados no Código Civil de 2002, quais sejam a boa-fé, os bons costumes e o fim econômico ou social do direito, situação que até hoje perdura, invocando-se o art. 187, mas fundamentando-se o decidido nos postulados normativos da proporcionalidade e da razoabilidade a rejeição ao abuso.”

do ponto de vista interno<sup>119</sup>, razão pela qual a regra vai além da noção de abuso de direito tradicionalmente referida pela doutrina.

A ilicitude decorrente do art. 187 implicará em uma violação à boa-fé objetiva, aos bons costumes ou ao fim econômico ou social de determinada posição jurídica subjetiva, que deverá ser analisada de acordo com as particularidades do caso concreto. Quanto a este ponto, evidentemente que uma IA autônoma poderá comportar-se de tal modo a violar o que se entende por “boa-fé”, “bons costumes” ou “fim econômico ou social”. E uma vez constatada tal violação, pode-se concluir pela ocorrência de um ilícito que, resultando em um dano, dará azo à imputação da responsabilidade civil ao *titular do direito subjetivo*.

A questão que pode surgir, na linha do que já se objetou anteriormente quando da análise da responsabilidade pela guarda da coisa, diz respeito à condição de tal “titular de direito subjetivo” como responsável pela IA autônoma e até que ponto o dano causado pelo sistema ou robô poderia ser evitado pelo humano que o possui ou utiliza. Igualmente, cabe perquirir se a imputação da responsabilidade civil ao humano, nesse caso, não poderia ser afastada pela conclusão de que a sua total inabilidade para prever o comportamento da IA seria uma hipótese de caso fortuito.

Mas há uma diferença entre a obrigação de indenizar que se fundamenta na ilicitude objetiva do art. 187 e a obrigação que decorre da responsabilidade pela guarda da coisa, porque a análise da conduta do agente no caso da guarda da coisa e da ilicitude objetiva do art. 187 é diversa. Com relação ao art. 187, Judith Martins-Costa explica que:

O mais relevante, porém, é observar que nesse novo modelo a ilicitude não é apenas "ilegalidade" nem "contrariedade culposa a preceito contratual", pressupondo uma idéia de direito subjetivo que - integrante das complexas situações jurídicas subjetivas, existenciais ou patrimoniais - não é "poder da vontade", antes situando-se na *integração de liberdades coexistentes*, como algo que já nasce "conformado" no jogo de ponderações entre os diferentes princípios que se põem como vetores axiológicos fundamentais do ordenamento. (MARTINS-COSTA, 2003)

Ou seja, a partir do art. 187, há um deslocamento da “expectativa” da conduta do agente, que deve ser analisada por ótica diversa daquela aplicada à responsabilidade pela guarda da coisa. O art. 187 não reflete uma obrigação de resultado do titular do direito subjetivo, que exerce a propriedade sobre a coisa, pois o ilícito, nesse caso, é “configurado por uma situação objetiva e concreta, decorrente do exame da conduta humana a partir das condições fáticas e

---

<sup>119</sup> Isso porque “Indiscutivelmente, a opção por uma teoria interna conduz o intérprete a considerar que toda a norma contém em si uma disposição definitiva já em abstrato, insuscetível de ponderação frente a uma determinada hipótese empírica proposta à discussão. Ou há adequação da situação empírica à regra proposta em sua abrangência máxima ou, no máximo, a possibilidade de limitação ao direito pressuposto em seu conteúdo - desde logo determinado este conteúdo abstratamente.” (CACHAPUZ, 2011)

jurídicas impostas pela realidade do caso e tendentes à configuração de uma restrição à liberdade do homem” (CACHAPUZ, 2011). Parte-se não de uma relação de dominação/propriedade do humano perante a IA autônoma, analisada a partir de uma teoria interna do direito, mas, essencialmente, da premissa de um direito de liberdade da pessoa que escolhe fazer uso de uma IA e, fazendo-o, viola a boa-fé, os bons costumes ou o fim econômico/social para o qual aquele produto foi desenvolvido. A diferença decorre da leitura que Cachapuz propõe a partir de uma teoria externa de restrição de direitos fundamentais, pois:

[...] a opção por uma teoria interna conduz o intérprete a considerar que toda a norma contém em si uma disposição definitiva já em abstrato, insuscetível de ponderação frente a uma determinada hipótese empírica proposta à discussão. Ou há adequação da situação empírica à regra proposta em sua abrangência máxima ou, no máximo, a possibilidade de limitação ao direito pressuposto em seu conteúdo - desde logo determinado este conteúdo abstratamente. (CACHAPUZ, 2011)

Nesse sentido, não se parte do pressuposto de que o titular do direito subjetivo tem um dever de “guarda”, no sentido de evitar danos causados pelo produto de IA, mas sim a liberdade de fazer uso desse produto, nos limites impostos pela boa-fé, bons costumes e seu fim econômico ou social. A violação desses limites não se constata pela ocorrência do dano e a conclusão de que se violou uma obrigação de resultado, já pressuposta em abstrato, a partir de conceitos já consolidados, mas da ponderação, frente ao caso concreto, entre outras liberdades colidentes com o direito de fazer uso do produto de IA autônoma.

A eventual dificuldade que o humano pode encontrar em agir de modo a evitar uma situação de ilicitude relacionada ao uso da IA autônoma é a única objeção que poderá suscitar maiores discussões a respeito da adequação da aplicação do art. 187 como norma capaz de solucionar a hipótese de danos causados pela IA autônoma. Contudo, o fato de que o “titular de direito subjetivo” é justamente quem fornece os dados à IA e a “ensina” é um forte argumento a favor da conclusão de que os resultados da IA autônoma – especialmente quando enviesados pelos dados de *input* – decorrem do exercício irregular do direito do humano.

Evidentemente que, também aqui, seria possível discutir o papel do fabricante e do programador e se tais pessoas poderiam ser, de algum modo, responsabilizadas em conjunto com o titular do direito subjetivo, que se presume ser o proprietário ou usuário da IA. Neste ponto, talvez se possa discutir a hipótese de introdução de uma regra de solidariedade da obrigação de indenizar os danos causados pela IA autônoma.

Sem prejuízo, parece-nos que, se nenhuma teoria de responsabilidade civil extra negocial se mostrar adequada para solucionar as hipóteses de danos causados pela IA autônoma – e já se apontou que o risco, em razão do viés algorítmico, pode ser um meio de solução –, é

possível resolver os casos a partir da aplicação do art. 187 do Código Civil, conclusão que afasta, ao menos *a priori*, a constatação de uma efetiva lacuna normativa no Código Civil de 2002. O desafio, nesse sentido, é identificar, concretamente, a hipótese de ilicitude a partir da confrontação entre o exercício irregular de determinada posição jurídica subjetiva pelo seu titular e a autonomia da IA. Ou, ainda, se a própria IA poderia ser considerada um “titular de direito subjetivo” para, nos termos do art. 187, responder pelo próprio comportamento. É o que se passa a analisar no próximo capítulo.

## 6 SEGUNDA PERSPECTIVA: A PESSOA ELETRÔNICA

Conforme brevemente apontado no tópico 3.3 desta pesquisa, talvez o ponto mais polêmico da Resolução de 16 de fevereiro de 2017 seja o princípio 59 (f), no qual o Parlamento Europeu instou a Comissão Europeia a analisar a seguinte possibilidade:

Criar um estatuto jurídico específico para os robôs a longo prazo, de modo a que, pelo menos, os robôs autônomos mais sofisticados possam ser determinados como detentores do estatuto de pessoas eletrônicas responsáveis por sanar quaisquer danos que possam causar e, eventualmente, aplicar a personalidade eletrônica a casos em que os robôs tomam decisões autônomas ou em que interagem por qualquer outro modo com terceiros de forma independente. (UNIÃO EUROPEIA, 2017)

Ou seja, diante das dificuldades para se solucionar o problema da imputação da responsabilidade civil com as regras já existentes em razão das particularidades intrínsecas à inteligência artificial, como o *deep learning*, em que os algoritmos muitas vezes são escritos pela própria máquina, sugere-se que a própria IA seja responsabilizada<sup>120</sup>. O patrimônio afetado se resolveria, então, pela criação de um regime de seguros obrigatórios ou um fundo de compensação (igualmente sugeridos no princípio 59 da referida Resolução).

A sugestão tem como fundamento a autonomia adquirida por robôs a partir de processos de *machine learning*, mas essa autonomia seria voltada principalmente para a solução da responsabilidade civil, não implicando, por conseguinte, em um reconhecimento geral da personalidade jurídica dos robôs (ALABART, 2018, p. 75). Alabart aponta que “não se trataria de fazer dos robôs sujeitos de direitos de forma geral, mas apenas que com essa personalização limitada se eliminariam alguns problemas para tornar efetiva a indenização pelos danos causados”<sup>121</sup> (2018, p. 77).

Maria Cláudia Cachapuz pontua, porém, a necessidade de se compreender “em que medida o problema da responsabilização civil por atos ilícitos oriundos da atividade de produtos de Inteligência Artificial [...] parte da necessidade de visualização do conceito de pessoa que é pressuposto para esta responsabilização” (2019). Nesse sentido, a autora aponta a importância de se verificar se a autonomia de entes dotados de IA permitiria o reconhecimento, desde logo, de direitos subjetivos, ou se ainda estamos distantes dessa realidade, porque lhes faltaria outras características da personalidade (CACHAPUZ, 2019).

---

<sup>120</sup> Conforme refere Silvia Díaz Alabart, “*puede llegar un momento en que resulte complejo imputar la responsabilidad a alguna persona, ya que ninguna persona va a poder, en realidad, prever las reacciones del robot.*” (2018, p. 74)

<sup>121</sup> No original, “*No se trataría de hacer a los robots sujetos de derechos de forma general, sino que con esa personalización limitada se eliminarían algunos problemas para que se pueda hacer efectiva la indemnización por los daños causados.*”

Cachapuz alerta que, sob o ponto de vista de uma ética utilitarista, o reconhecimento de direitos subjetivos a robôs não apresenta maiores dificuldades, “na medida em que toda a construção do juízo – em termos cognitivo, volitivo, normativo – se dá por uma ética de iguais interesses” (2019). De outro lado:

Na medida em que o próprio texto da Resolução do Parlamento Europeu elege, *a priori*, as Leis de Asimov como imperativos de partida para o estabelecimento de princípios gerais a essa normatividade, é preciso que o teste pragmático de qualquer enunciado normativo que discipline sobre um novo estatuto de pessoa – e que exija considerações interespecies – passe igualmente por um reexame das questões pertinentes aos direitos de personalidade em termos universais – pela capacidade de ser válido a todos os que se dispõem ao compartilhamento desses novos espaços de liberdade –, ainda que exigido um condicionamento prévio, embora cauteloso, à aceitação de ausência de freios aos avanços tecnológicos. (CACHAPUZ, 2019)

Com efeito, a Resolução do Parlamento Europeu parece preocupar-se com uma personalidade eletrônica que decorre fundamentalmente da autonomia decisória dos entes dotados de IA. Mas é apenas da autonomia que decorre a personalidade? Aliás, dizer que “robôs tomam decisões autônomas ou em que interagem por qualquer outro modo com terceiros de forma independente” (UNIÃO EUROPEIA, 2017) significa dizer que um robô tem, ao fim e ao cabo, a autonomia de uma pessoa humana?

A discussão sobre o reconhecimento de personalidade própria à IA não é exclusiva da Resolução de 16 de fevereiro de 2017 e tampouco surgiu recentemente. Em 1992, Laurence B. Solum analisou possíveis objeções para a hipótese em que uma IA demandasse o reconhecimento de direitos civis previstos na Constituição dos Estados Unidos contra o seu dono. Solum considerou argumentos relacionados a três diferentes objeções: (i.) que apenas pessoas naturais deveriam ter direitos constitucionais; (ii.) que a IA não possui elementos essenciais da personalidade, como consciência, sentimentos, alma, etc.; e (iii.) como a IA é uma criação humana, ela nunca será mais do que uma propriedade dos humanos (1992, p. 1258).

Em 2007, Robert Alexy e Alfonso García Figuerola analisaram a possibilidade do reconhecimento de direitos humanos aos robôs a partir da história do androide Data, contada no episódio “*The Measure of a Man*” de *Star Trek: Next Generation*, exibido pela primeira vez nos Estados Unidos em 11 de fevereiro de 1989 (2007, p. 86). Nesse episódio o comandante Bruce Maddox queria iniciar a produção em massa de androides iguais a Data e, para isso, desejava desmontá-lo. Maddox admitiu que Data poderia ser destruído se a operação não desse certo e, por isso, Data se negou a participar do experimento. Diante do impasse, foi realizado um julgamento para decidir se Data era mera propriedade da frota estelar ou se teria direito à liberdade, que resultou na vitória de Data (ALEXY; FIGUEROA, 2007, p. 87–91).

A partir da história do episódio, Alexy e Figueroa propuseram-se a testar se a decisão do tribunal de reconhecer o direito à liberdade à Data estaria ou não correta. A primeira questão que se apresenta, nesse sentido, é *quem* deve ser titular de direitos humanos. Isso porque, se apenas humanos no sentido biológico terão direitos humanos, a única conclusão possível é que a decisão do tribunal estaria equivocada, já que Data é uma máquina. Mas se direitos humanos forem reconhecidos também a quem, sem ser biologicamente humano, é considerado “pessoa”, então é possível verificar se Data pode ser considerado “pessoa” e, com isso, ser titular de direitos (ALEXY; FIGUEROA, 2007, p. 93).

Essa solução, porém, deve ser analisada com cautela, para que não se coloque em risco a concepção de humanidade. Os autores apontam, nesse sentido, que a consciência, por exemplo, é uma característica essencial do conceito de pessoa, mas não é um pressuposto para o reconhecimento de direitos humanos a seres humanos; do contrário, aqueles que por alguma enfermidade foram privados de consciência acabariam privados de seus direitos humanos (ALEXY; FIGUEROA, 2007, p. 93). Alexy e Figueroa esclarecem, assim, a necessidade de se “debilitar” a tese de direitos humanos, de modo que “a compatibilidade da tese dos humanos com a tese das pessoas se elabora mantendo a parte da tese dos humanos que diz que todos os humanos são titulares de direitos e suprimindo a parte que diz que apenas humanos são titulares desses direitos”<sup>122</sup> (2007, p. 93–94, tradução nossa).

Solum, por sua vez, comenta que o reconhecimento de uma “pessoa” que não é humana poderia ser um impeditivo ao reconhecimento de direitos constitucionais à IA, porque o conceito de pessoa está intrinsecamente ligado à natureza humana – nunca conhecemos uma “pessoa” que não fosse humana<sup>123</sup> (1992, p. 1259). No entanto, houve um tempo em que nem todos os humanos eram considerados pessoas. Esse dualismo é explicado por Hannah Arendt em belíssima passagem sobre a origem da palavra pessoa, do latim “*persona*”:

*Persona*, em todo caso, referia-se em latim à máscara do ator, aquela que cobria a sua face ‘pessoal’ individual, indicando para o espectador o papel e a parte do ator na peça. Mas nessa máscara, que era criada e determinada para a peça, havia uma abertura larga no lugar da boca, pela qual soava a voz individual e sem disfarce do ator. É desse soar através que a palavra *persona* derivou originalmente: *per-sonare*, ‘soar através’, é o verbo do qual *persona*, a máscara, é o substantivo. E os próprios romanos foram os primeiros a usar o substantivo num sentido metafórico; na lei

<sup>122</sup> No original, “*La compatibilidad de la tesis de los humanos con la tesis de las personas se elabora manteniendo la parte de la tesis de los humanos que dice que todos los humanos son titulares de derechos y suprimiendo la parte que dice que sólo los humanos son titulares de estos derechos*”.

<sup>123</sup> Solum comenta que o fato de companhias terem reconhecido seu direito à propriedade e liberdade de expressão perante a Constituição Americana não poderia ser utilizado como fundamento para embasar igual reconhecimento a entes dotados de IA, porque a pessoa jurídica tem por trás justamente o interesse de pessoas naturais e o patrimônio da companhia pertence, em última instância, às pessoas naturais que são proprietárias das ações ou quotas sociais (1992, p. 1259).

romana, *persona* era alguém que possuía direitos civis, em contraste agudo com a palavra *homo*, que denotava alguém que não passava de um membro da espécie humana, diferente, sem dúvida, de um animal, mas sem nenhuma qualificação ou distinção específica, de modo que *homo*, como o grego *anthropos*, era frequentemente usado de modo desdenhoso para designar pessoas não protegidas pela lei. (ARENDDT, 2004, p. 75)

Moreira Alves (2019) aponta que “pessoa”, no direito romano, era o titular de direitos subjetivos, que poderia ser uma pessoa física ou natural (homens não escravos) ou uma pessoa jurídica ou moral (seres abstratos a quem se reconhece a titularidade de direitos subjetivos). Para que um homem fosse considerado pessoa, deveria existir para a ordem jurídica e ter *personalidade jurídica*. Há que se distinguir, nesse sentido, três relevantes conceitos: (i.) a capacidade de fato, “que é a aptidão para praticar, por si só, atos que produzam efeitos jurídicos”; (ii.) a personalidade jurídica, que é “a aptidão de adquirir direitos e de contrair obrigações” e (iii.) a capacidade jurídica, que é o limite da personalidade jurídica (um conceito relativo, portanto) (MOREIRA ALVES, 2019).

Para que fosse reconhecida a um homem a personalidade jurídica, exigia-se que o homem fosse (i.) livre (*status libertatis*); (ii.) cidadão romano (*status civitatis*) e (iii.) *pater familias* (*status familiae*). A condição de cidadão romano foi flexibilizada posteriormente, a fim de que se reconhecesse personalidade jurídica também aos estrangeiros, aos quais se limitava, todavia, a capacidade jurídica, para que fosse mais restrita em comparação àquela reconhecida aos cidadãos romanos (MOREIRA ALVES, 2019). Tratava-se de um regime bastante excludente, que garantia a um grupo seletivo o acesso à esfera pública.

Essa concepção de pessoa foi modificada na modernidade. Conforme explica Cachapuz, ao contrário da concepção dos antigos, a modernidade construiu “uma estrutura capaz de reconhecer, àquele *homo*, interesses fundamentais, inclusive independentemente da possibilidade de que, como *persona*, não pudesse alcançar mesma proteção jurídica no âmbito público” (2019). Nesse sentido, a autora destaca que:

O que o Direito Civil aplicado às questões da robótica, na contemporaneidade, se propõe a fazer é, para além da dotação de interesses fundamentais ao *homo*, investigar se há como estender esses interesses a quem esteja atrás da máscara, sem embargo de sua humanidade, reconhecendo mesma capacidade originária a qualquer espécie para que, como no verbo *personare*, possa, de forma autêntica, “soar através”. (CACHAPUZ, 2019)

Mas não é apenas a tese de direitos humanos que é “debilitada”. Alexy e Figueroa também estabelecem uma premissa de liberdade deficitária, pois “na perspectiva humana, há subjacente a compreensão de que o conceito de pessoa só se explica num ambiente de liberdade” (CACHAPUZ, 2019).

Assim, partindo da tese “debilitada” de direitos humanos e liberdade de Alexy e Figueroa, o segundo passo é analisar quais são as características que definem uma pessoa. Os autores apontam, nesse sentido, que “para que algo seja uma pessoa, deve primeiro, ser inteligente, segundo, ter sentimentos e, terceiro, ter consciência”<sup>124</sup> (ALEXY; FIGUEROA, 2007, p. 94, tradução nossa).

No caso do androide Data, a inteligência foi indiscutível. De fato, essa característica parece ser a mais simples de ser superada, já que se estaria tratando justamente de entes dotados de “inteligência” artificial, isto é, entes capazes de atingir objetivos no mundo. Nesse sentido, Cachapuz comenta que “a racionalidade há muito deixou de ser um empecilho real ao reconhecimento da condição de pessoa, observadas capacidades de desenvolvimento de juízos cognitivos e decisórios de forma racional [...] e livre de uma interferência humana.” Mas a autora também questiona a qual tipo de IA se poderia reconhecer uma racionalidade com tal traço de independência a ponto de justificar o reconhecimento de uma personalidade própria.

Alan Turing apontou que uma máquina poderia ser considerada inteligente se vencesse um humano no jogo da imitação (1950). Contudo, conforme referido por Solum, o americano John Searle criticou o Teste de Turing a partir de um experimento chamado “A Sala Chinesa”, que funcionaria do seguinte modo:

Imagine que você está trancado em uma sala. Na sala, aparecem conjuntos de escrita chinesa, mas você não sabe chinês. No entanto, você recebe um livro de regras, escrito em inglês, no qual pode consultar os fragmentos de chinês, por sua forma. O livro de regras fornece um procedimento para produzir sequências de caracteres chineses que você envia para fora da sala. Os que estão fora da sala estão jogando alguma versão do jogo de Turing. Eles estão convencidos de que tudo o que está na sala entende chinês. Mas você não sabe uma palavra em chinês, está simplesmente seguindo um conjunto de instruções (que podemos chamar de programa) com base na forma dos símbolos chineses. Searle acredita que esse experimento hipotético demonstra que nem você nem o livro de instruções (o programa) entende chinês, embora você e o programa possam simular tal conhecimento.<sup>125</sup> (1992, p. 1236, tradução nossa)

Ou seja, o fato de uma IA ser capaz de agir de modo aparentemente inteligente não significa que o programa esteja compreendendo, efetivamente, o que está fazendo ou esteja

---

<sup>124</sup> No original, “*Para que algo sea una persona, debe primero ser inteligente, segundo tener sentimientos y tercero tener conciencia*”.

<sup>125</sup> No original, “*Imagine that you are locked in a room. Into the room come batches of Chinese writing, but you don't know any Chinese. You are, however, given a rule book, written in English, in which you can look up the bits of Chinese, by their shape. The rule book gives you a procedure for producing strings of Chinese characters that you send out of the room. Those outside the room are playing some version of Turing's game. They are convinced that whatever is in the room understands Chinese. But you don't know a word of Chinese, you are simply following a set of instructions (which we can call a program) based on the shape of Chinese symbols. Searle believes that this thought experiment demonstrates that neither you nor the instruction book (the program) understands Chinese, even though you and the program can simulate such understanding.*”

*querendo* produzir aquele resultado. Ora, se o programa realiza algo apenas porque foi programado para tanto, inteligente é seu criador, não os algoritmos que o comandam.

De tudo isso, o que se pode concluir é que a definição de “inteligência” é subjetiva e dependerá do método adotado para a sua aferição. O Teste de Turing pode fornecer um parâmetro do ponto de vista comportamental, mas há, além deste, modelos de comportamento racional ou silogístico que igualmente permitirão concluir pela presença de inteligência (RUSSEL; NORVIG, 1995, p. 5). Essa dependência da análise caso a caso da inteligência aponta, todavia, que o reconhecimento de uma personalidade eletrônica é questão complexa que talvez não devesse ser tratada de modo geral a qualquer robô que pareça agir com relativa autonomia.

O segundo requisito da personalidade é a capacidade de ter sentimentos. Alexy e Figueroa apontam que as emoções são etéreas, isto é, não podem ser percebidas pelos sentidos, embora seja possível identificar o sentimento das outras pessoas a partir de pequenos gestos ou atitudes (2007, p. 95). No caso de Data, foi o fato de ter guardado um livro que ganhara de presente de Picard como recordação pela sua amizade que demonstrou a presença de sentimentos (ALEXY; FIGUEROA, 2007, p. 96).

Solum aponta que se poderia argumentar a favor do desenvolvimento de sentimentos na IA, porque:

A emoção é uma faceta da mentalidade humana, e se a mente humana pode ser explicada pelo modelo computacional, então a emoção poderia ser um processo computacional. Mais genericamente, se as emoções humanas obedecem às leis naturais, então (em teoria) um programa de computador pode simular a operação dessas leis.<sup>126</sup> (SOLUM, 1992, p. 1270, tradução nossa)

Porém, o americano destaca que mesmo quando uma IA reproduzir comportamentos humanos que denotem alguma emoção, caberá questionar se aquela emoção é real ou se é apenas uma simulação da máquina (SOLUM, 1992, p. 1271). Nesse sentido, Cachapuz aponta que há um aspecto cognitivo no sentir, um juízo de valor, “uma associação entre algo que pode ser padronizado – a resposta à determinada provocação – e algo que se encontra no campo subjetivo do desejo, sujeito à permanente correção pela experiência da própria vida” (2019). Essa “singularidade” decorrente da autenticidade do sentir e que “afasta a mecanicidade do comportamento da pessoa” (CACHAPUZ, 2019) é fundamental para diferenciar uma reação robótica daquilo que pode ser, efetivamente, considerado um sentimento.

---

<sup>126</sup> No original, “*Emotion is a facet of human mentality, and if the human mind can be explained by the computational model, then emotion could turn out to be a computational process. More generally, if human emotions obey natural laws, then (in theory) a computer program can simulate the operation of these laws.*”

Portanto, tal como ocorre com a inteligência, também a capacidade de uma IA ter sentimentos implica em uma análise subjetiva e casuística, que, parece-nos, não poderia decorrer de uma previsão genérica da lei e, de todo modo, parece estar ainda muito distante da realidade.

O terceiro requisito apontado por Alexy e Figueroa como característico de uma pessoa é a “consciência” ou a capacidade “espiritual” (2007, p. 96-97). Esse é o atributo mais complexo, que se subdivide em três tipos de reflexividade: cognitiva, volitiva e normativa (ALEXY; FIGUEROA, 2007, p. 97).

A reflexividade cognitiva se observa “quando alguém faz de si mesmo objeto de seu conhecimento”<sup>127</sup> (ALEXY; FIGUEROA, 2007, p. 97). A reflexividade volitiva é a capacidade de alguém reger o próprio comportamento mediante atos de vontade, com autodomínio, isto é, a capacidade de agir com autonomia (ALEXY; FIGUEROA, 2007, p. 98). Por fim, a reflexividade normativa diz respeito à capacidade de julgar os próprios atos e analisar a correção da própria conduta (ALEXY; FIGUEROA, 2007, p. 99).

Assim, para que seja considerada pessoa, além da inteligência e da capacidade de ter sentimentos, é preciso que a IA: (i.) tenha conhecimento sobre a própria existência e o seu lugar no mundo; (ii.) possa agir com autonomia, decidindo por si mesma; e (iii.) seja capaz de analisar os próprios atos e buscar agir com correção. Se todos esses aspectos forem observados, então outros robôs, além de Data, poderiam ser considerados pessoas e, conseqüentemente, ser titulares de direitos subjetivos.

A primeira observação que se apresenta quanto à característica da consciência é que, para que uma IA possa adquirir autoconhecimento e autonomia e desenvolver juízos de correção sobre o próprio comportamento, um pressuposto fundamental seria a sua habilidade de aprendizado (*machine learning*). Conseqüentemente, o nível de cada aspecto da consciência desenvolvido pela IA dependerá das suas interações com o mundo, com outros robôs/sistemas de IA e com os seres humanos. Assim, mesmo que uma linha inteira de robôs seja lançada ao mercado com a mesma programação inicial, cada robô se desenvolverá individualmente, a partir das próprias experiências de aprendizado, adquirindo o respectivo conhecimento sobre si mesmo ao longo do tempo.

Isso significa que robôs de uma mesma linha poderão se desenvolver de maneiras totalmente distintas, de modo que, tal como já apontado com relação às características da inteligência e do sentimento, a análise da consciência da IA só poderá ser feita caso a caso.

---

<sup>127</sup> No original, “cuando alguien hace de sí mismo objeto de su conocimiento”.

Assim, não seria possível reconhecer uma personalidade própria a uma linha inteira de robôs, porque, quando imediatamente postos no mercado, esses robôs ainda não interagiram com o mundo, não aprenderam com as próprias experiências e não desenvolveram características e habilidades que permitam compará-los com pessoas humanas no sentido de se autoconhecerem.

A segunda observação diz respeito à distinção entre a noção de autonomia que compõe a reflexividade volitiva necessária para o reconhecimento da personalidade e a noção de liberdade.

Conforme aponta Jürgen Habermas, na liberdade, “a vontade é determinada por máximas de prudência, pelas preferências ou motivos racionais, digamos, que uma pessoa tem”, de modo que “o ato da liberdade surge como parte da consciência de um único sujeito” (2013, p. 12). Já na autonomia, segundo Habermas, “a vontade se deixa determinar por máximas aprovadas pelo teste da universalização”, isto é, “a vontade de uma pessoa é determinada por motivos que deveriam igualmente ser levados em conta por todas as outras pessoas” (2013, p. 12). Assim:

Quanto à liberdade subjetiva, não é difícil imaginar que algumas pessoas possam gozar da liberdade e outras não, ou que algumas possam ser mais livres do que outras. A autonomia, ao contrário, não é um conceito distributivo e não pode ser alcançada individualmente. Nesse sentido enfático, uma pessoa só pode ser livre se todas as demais o forem igualmente. A ideia que quero sublinhar é a seguinte: com sua noção de autonomia, o próprio Kant já introduz um conceito que só pode explicitar-se plenamente dentro de uma estrutura intersubjetivista. (HABERMAS, 2013, p. 13)

A dimensão intersubjetiva da autonomia, tal como explicitada por Habermas, é fundamental para a análise da personalidade própria da IA, pois, conforme destaca Cachapuz, “[p]ara além de questões de empatia, compaixão e tolerância, debates como esse se inserem num teste da reciprocidade exigida à ideia contemporânea de autonomia” (2019).

Ocorre que a Resolução do Parlamento Europeu indica, no considerando “T”, que “as Leis de Asimov têm de ser encaradas como dirigindo-se aos criadores, aos produtores e aos operadores de robôs, incluindo robôs com autonomia integrada e autoaprendizagem, uma vez que não podem ser convertidas em código de máquina” (UNIÃO EUROPEIA, 2017). E as Leis de Asimov estabelecem uma desigualdade de tratamento entre humanos e robôs bastante evidente, determinando que um robô proteja a própria existência apenas quando isso não implicar em “magoar a humanidade” ou desobedecer às ordens dos humanos (de acordo com a Primeira e Segunda Leis da Robótica) (ASIMOV, 2014, p. 65). Daí porque Cachapuz aponta:

[...] vê-se que a autonomia desenhada para a personalidade de um produto de Inteligência Artificial, no ideal do Parlamento Europeu, apresenta-se como mitigada desde logo – e não sujeita a uma universalização *a priori* –, na medida em que ausente

um elemento de reciprocidade de tratamento entre humanos e robôs, mesmo que ressalvada a hipótese de coexistência de uma autoproteção ao robô. (CACHAPUZ, 2019)

Por isso, Cachapuz destaca que a autonomia mencionada na Resolução do Parlamento Europeu é insuficiente para fins de reconhecer uma personalidade eletrônica, servindo somente “para autorizar medidas de proteção diversas em termos de responsabilidade civil” (2019).

A terceira observação que se apresenta com relação à análise da consciência – aplicável também à análise da inteligência e da capacidade de ter sentimentos – diz respeito a como se daria o exame da presença de características humanas em uma IA. Conforme aponta Solum, os argumentos sobre a análise de consciência, sentimentos, autonomia e intencionalidade de uma IA geralmente são baseados no seu comportamento (1992, p. 1275). Porém, Solum lembra que:

[...] computadores podem simular o comportamento de muitas coisas, de terremotos e ondas à guerra termonuclear. Nós não somos tentados a dizer que uma simulação de computador de um terremoto é um terremoto – não importa quão boa a simulação seja. Por que quereríamos dizer que uma simulação de computador de uma pessoa é uma pessoa ou que uma simulação de computador de inteligência é inteligência?<sup>128</sup> (1992, p. 1275)

Solum refere que a ciência cognitiva poderia fornecer melhores subsídios para a análise da consciência da IA futuramente, pois falta-nos evidências para estabelecer os limites da personalidade para além da natureza humana (1992, p. 1285). Trinta anos depois da publicação do artigo de Solum, Mafalda Miranda Barbosa comenta que o avanço na neurociência tem dado lugar a argumentos de que as decisões humanas decorrem de processos neuronais inconscientes, captados por tomografias, em um movimento no qual o “homem sociológico” é substituído pelo “homem neurobiológico” (2020, p. 301-302). Ela lembra, porém, que só se pode entender a responsabilidade a partir de uma perspectiva cultural:

A liberdade humana é muito mais do que uma mera possibilidade de decisão. Envolve a possibilidade de decidir tendo em consideração os interesses dos outros e, portanto, a possibilidade de nos transcendermos a nós próprios. Ora, nenhum destes aspetos pode ser capturado por uma imagem tomográfica, porque envolve dimensões que ultrapassam uma decisão imediata tomada sob impulso ou de acordo com um comando. (BARBOSA, 2020, p. 303)

Liberdade (subjativa) e autonomia (intersubjetiva) (HABERMAS, 2013, p. 13) são noções fundamentais para o que se entende por pessoa que deverão ser analisadas com relação à pessoa eletrônica. Uma personalidade que decorre apenas da “autonomia” da IA – que, aliás,

---

<sup>128</sup> No original, “computers can simulate the behavior of lots of things, from earthquakes and waves to thermonuclear warfare. We are not tempted to say that a computer simulation of an earthquake is an earthquake - no matter how good the simulation is. Why would we want to say that a computer simulation of a person is a person or that a computer simulation of intelligence is intelligence?”

é bastante mitigada pelas Leis da Robótica (CACHAPUZ, 2019) – não parece ser suficiente para justificar a mudança de paradigma na noção de pessoa. Partilha-se da conclusão de Cachapuz, no sentido de que a ampliação do conceito de pessoa com o fim de resolver casos de responsabilidade civil envolvendo a IA só será possível se se ponderar, de maneira suficientemente informada, “o risco causado pela fragilização do conceito de humanidade quando associado a uma tese alargada de proteção a direitos humanos a outras espécies” (2019).

No entanto, na hipótese de se atribuir personalidade à IA, a análise de cada característica da pessoa não pode ser feita de maneira geral. Conforme apontado ao longo deste capítulo, qualquer aspecto que caracterize uma pessoa (a inteligência, a capacidade de ter sentimentos e a consciência) seria desenvolvido pela IA a partir das suas habilidades de aprendizado e interação com humanos. Assim, parece-nos que a outorga da condição de pessoa a uma IA não poderia se dar “de fábrica”, mas permanecer necessariamente atrelada à análise casuística, sob pena de se estar equiparando por completo humanos a robôs sem que haja sequer um traço real de semelhança – lembra-se da “Sala Chinesa”: o fato de um programa simular determinada habilidade não significa que ele a compreenda efetivamente.

Além disso, se superada a questão sobre a outorga de personalidade à IA, o óbice apontado para a adoção de um regime subjetivo de responsabilidade em razão do não compartilhamento dos mesmos padrões éticos entre humanos e produtos de IA parece não mais fazer sentido. Ora, se seremos capazes de identificar sentimentos, capacidade de autoconhecimento e de autonomia na IA, por certo que também será possível julgar se tais entes agiram com culpa ou dolo em determinada situação.

Nesse caso, é evidente que não basta apenas reconhecer a personalidade à IA; ela também deverá ser capaz de adimplir com as obrigações que serão reconhecidas pela ordem jurídica. Daí a necessidade de se resolver, para além da questão da imputação, também os meios pelos quais se dará a responsabilidade patrimonial (seja por meio de um seguro vinculado à IA, seja por um fundo de compensação).

Por fim, este capítulo não tem o objetivo de esgotar a discussão sobre a personalidade eletrônica – e tampouco poderia, em tão curto espaço. A análise aqui apresentada buscou, primordialmente, testar as implicações da sugestão do Parlamento Europeu para a criação de uma pessoa eletrônica como forma de solucionar casos de responsabilidade civil decorrente de danos causados diretamente pela IA autônoma e apresentar os principais pontos críticos dessa abordagem.

Não se descarta que “em um futuro em que interagiremos com IAs ou seres inteligentes de outros planetas, talvez sejamos forçados a refinar nosso conceito de pessoa”<sup>129</sup> (SOLUM, 1992, p. 1286, tradução nossa). Todavia, hoje, parece que ainda estamos distantes de uma realidade de robôs efetivamente inteligentes, sensíveis e capazes de exercer plena autonomia, à nossa imagem e semelhança. E, considerando todas as preocupações intrínsecas ao conceito de humanidade e a noção de reconhecimento de direitos humanos, que serão certamente afetados se ampliados de modo a abarcar também robôs, não parece demais questionar se a realidade em que coexistiremos com corpos metálicos mais inteligentes, mais fortes e mais duradouros que nós deveria ser tão almejada.

---

<sup>129</sup> No original, “*in a future in which we interact with such AIs or with intelligent beings from other planets, we might be forced to refine our concept of person*”.

## 7 CONCLUSÃO

A razão para ter iniciado esta pesquisa decorre de um questionamento surgido diante da Resolução do Parlamento Europeu de 16 de fevereiro de 2017: se, em tempos de um imenso avanço da robótica, nos deparamos com textos legislativos com expressa referência à literatura da ficção para o estabelecimento de parâmetros éticos de desenvolvimento tecnológico, estaria a ficção se tornando realidade?

Os exemplos de IA mencionados ao longo do trabalho (o robô Gaak, os negociadores do Facebook, Bob e Alice, o COMPAS) demonstram que a discussão sobre a relação entre humanidade e tecnologia é necessária. Hoje, há diversos questionamentos na mesa, mas poucos deles com respostas: o quão longe estamos de conviver com robôs como os de Asimov? Como será a relação entre a humanidade e robôs ou *softwares* inteligentes, que não dependem de humanos para agir ou produzir resultados? Os robôs serão uma ameaça para a sobrevivência da humanidade? Quanto essas tecnologias impactarão no modo de vida da sociedade – para além do que já vêm impactando hoje? Robôs inteligentes devem ser tratados como pessoas ou como objetos? Por que estamos desenvolvendo robôs cuja inteligência pode superar a nossa?

Nesse universo de perguntas sem respostas, há algoritmos capazes de aprender a partir de experiências prévias e dados analisados, executando comandos que podem ser inclusive desconhecidos por seus criadores, sendo desenvolvidos e introduzidos no convívio social diariamente. Diante disso, pode a IA autônoma representar, de fato, uma mudança de paradigma na responsabilidade civil, ao introduzir um agente não humano capaz de praticar, de modo autônomo, atos reputados ilícitos e causar danos? Se sim, qual é o tratamento jurídico adequado para os danos causados diretamente por uma IA que age por conta própria?

O assunto, conforme indicado nos dois capítulos iniciais, não é exatamente novo, mas ganhou relevância nas duas últimas décadas, em especial, em razão do desenvolvimento de algoritmos de *deep learning*, que são capazes de escrever outros algoritmos e executar comandos totalmente desconhecidos dos programadores humanos, como verdadeiras “caixas pretas” (KNIGHT, 2017). Em razão disso, compreender uma decisão algorítmica ou o comportamento desses sistemas de IA autônomos pode ser impossível aos humanos.

Mas o desenvolvimento de tal tecnologia e a sua interação com a sociedade não necessariamente indica que é preciso alterar a legislação sobre responsabilidade civil. Conforme apontam Gustavo Tepedino e Rodrigo da Guia Silva, “ressalvadas, por certo, as hipóteses em que a inovação legislativa se afigurar indispensável –, poderá o intérprete concluir que ao

ineditismo das questões suscitadas pelas novas tecnologias não há de corresponder necessariamente o ineditismo das soluções jurídicas” (2019, p. 72).

Daí porque se buscou delimitar, no Capítulo 3, que tipos de lacunas podem ser identificadas em dado sistema jurídico, uma vez que nem toda lacuna demanda a introdução de uma nova regra no ordenamento jurídico. Assim, a necessidade – ou não – de inserção de novas regras de responsabilidade civil para tratar de casos envolvendo inteligência artificial decorreria de uma lacuna normativa, se, da ponderação dos princípios relevantes para uma situação de ocorrência de dano, se concluir que o ordenamento jurídico não dispõe de regra para tratar do assunto e que essa regra deveria existir. Mas também se apontou que a questão pode resultar em uma lacuna axiológica, em que há regra, mas a solução conflita com outros princípios do Direito, ou de reconhecimento, em que será preciso delimitar qual o alcance do significado semântico de determinada linguagem já prevista em uma regra existente. Como a lacuna axiológica depende da ponderação dos princípios ao caso concreto, a análise em abstrato baseia-se principalmente nas lacunas normativa e de reconhecimento, sendo que esta última dependerá essencialmente de interpretação, ou seja, de outras fontes do Direito para além da lei (como a dogmática e a jurisprudência).

Estabelecida essa premissa, foram apresentadas as propostas de tratamento legislativo para a IA. O Parlamento Europeu endereçou a questão na Resolução aprovada em 16 de fevereiro de 2017, sugerindo a adoção de um regime de responsabilidade objetiva ou pela gestão de riscos e instando a Comissão Europeia a analisar a possibilidade de criação de um estatuto jurídico próprio a robôs autônomos, que são independentes dos humanos: a “pessoa eletrônica”. No Brasil, os Projetos de Lei analisados não abordaram, especificamente a responsabilidade civil, mas constatou-se uma preocupação com o protagonismo humano na relação com a tecnologia, na medida que todos os projetos mencionam como diretriz ética para o desenvolvimento da IA a dignidade da pessoa humana.

Assim, foram identificadas duas linhas distintas de tratamento para os casos de responsabilidade civil envolvendo a IA autônoma, analisadas à luz de disposições de responsabilidade civil do Código Civil brasileiro: a responsabilidade, direta ou indireta, de humanos (Capítulo 5), e a criação de um estatuto jurídico próprio para a IA – personalidade eletrônica – para que o próprio robô responda por seus atos (Capítulo 6).

A responsabilidade civil que considera a IA um objeto ou, no máximo, um “agente” não humano, para fins de responsabilidade objetiva de quem se beneficia de seu uso, utiliza-se das teorias já tradicionais sobre a matéria. Assim, no capítulo 4, iniciou-se pela apresentação dos pressupostos da responsabilidade civil e da distinção entre a responsabilidade subjetiva, que

tem como fator de imputação a culpa, e a objetiva, lastreada no risco (à exceção do art. 187 do Código Civil, que permite a responsabilidade objetiva independente do risco), bem como entre a responsabilidade negocial e a extra negocial, em que não haverá uma relação jurídica prévia entre o agente responsável por indenizar e a vítima.

Foram analisadas cinco possíveis abordagens para a responsabilidade civil envolvendo a IA: a responsabilidade subjetiva do agente humano e, em razão da dificuldade de se comprovar a culpa, as hipóteses de responsabilidade objetiva pela atividade de risco, fato do produto, fato da coisa ou do animal e responsabilidade indireta.

Quanto à responsabilidade civil subjetiva dos usuários, proprietários, desenvolvedores e fabricantes da IA, demonstrou-se que é pouco aceita na doutrina e que tampouco é recomendada na Resolução do Parlamento Europeu de 16 de fevereiro de 2017. A demonstração do dolo ou da culpa de um humano poderia ser uma tarefa difícil, senão impossível, em situação sobre a qual se dispõe de tão pouca informação precisa (inclusive para se constatar o que seria negligência na utilização de uma IA autônoma, por exemplo). Aliás, justamente por se estar diante de entes que não são dotados do mesmo senso ético dos humanos, afasta-se, *a priori*, também a responsabilização subjetiva da própria IA, já que não haveria, a rigor, um parâmetro moral comum com os robôs que permitisse identificar a malícia de uma IA, a não ser que o reconhecimento da personalidade eletrônica seja lastreado também na constatação de que a IA autônoma possui todas as características de uma pessoa (inteligência, capacidade de ter sentimentos e consciência). Nesse caso, identificar o dolo e a culpa não parece ser uma tarefa mais árdua que identificar a autoconsciência de um sistema de IA ou a sua sensibilidade.

Todavia, como essa realidade ainda parece distante, analisou-se as teorias que fundamentam a responsabilidade objetiva.

A primeira solução foi a responsabilidade pela atividade de risco, prevista no art. 927, parágrafo único, do Código Civil. Čerka, Grigienė e Sirbikyte apontam que essa solução seria possível, porque produtos de IA podem ser considerados de risco e, portanto, quem se beneficia do produto de IA deveria responder objetivamente pelo risco criado em razão da sua utilização (2015, p. 386). David Vladeck pontua, de outro lado, que há produtos de IA criados especificamente para reduzir o risco e que seria incongruente imputar a responsabilidade objetiva com esse fundamento se a eliminação ou redução do risco de determinada atividade foi justamente o objetivo para o qual a IA foi desenvolvida (VLADECK, 2014).

Com efeito, é possível rotular como um risco o uso de um veículo autônomo (sem condutor) que não sofre efeitos da fadiga, não dorme dirigindo, não fica embriagado e nem pode ter sua capacidade psicomotora afetada pelo uso de substâncias psicotrópicas? É de risco

o uso desse veículo que segue um GPS, tem velocidade controlada e é programado para respeitar as leis de trânsito? Não são justamente fatores como fadiga, embriaguez ao volante, excesso de velocidade e violação às normas de trânsito as principais causas de acidentes por condutores humanos – comportamentos que não serão reproduzidos pelo veículo autônomo? No mesmo sentido, é de risco o uso de um robô capaz de identificar com precisão o diagnóstico de doenças quando médicos humanos poderiam não enxergar determinada célula ou organismo sem o uso da tecnologia? É de risco a possibilidade de realização de cirurgias menos invasivas e mais precisas, viabilizadas por produtos de IA (atualmente, operados por humanos)?

No caso, há regra, então não se pode dizer que há lacuna normativa. Restaria verificar, em concreto, se a aplicação entraria em conflito com outros princípios do direito (lacuna axiológica) ou se o problema seria o alcance do significado da palavra “risco” (lacuna de reconhecimento).

Ocorre que, se a IA foi utilizada justamente para mitigar ou eliminar o risco da atividade, ao que parece, não se trata de uma lacuna, mas de uma hipótese de *afastamento* da regra, porque a sua aplicação estaria ocorrendo em hipótese em que o elemento central do suporte fático não está presente: a criação de um risco. E esse é, conforme apontado no tópico 5.2.1, o motivo pelo qual Vladeck sugeriu a criação de um sistema de responsabilidade civil objetiva deslocado do risco (2014, p. 146). Do contrário, caberia questionar se o risco reside tão somente no fato de se estar delegando determinada tarefa à IA autônoma, caso em que seria necessário compreender o alcance da palavra risco (lacuna de reconhecimento) para resolver casos de produtos de IA criados para trazer maior segurança a determinada atividade.

Mas esse posicionamento não subsiste se considerarmos a questão do ponto de vista do viés algorítmico. Nesse sentido, demonstrou-se, a partir dos exemplos do COMPAS e do Hospital Sr. George, no Reino Unido, que o uso de algoritmos que são baseados em modelos matemáticos pode apresentar comportamentos tendenciosos que seriam suficientes para fundamentar a noção de risco.

Portanto, a regra de responsabilidade objetiva baseada no risco não poderia ser utilizada, *a priori*, se se considerar que o risco não é inerente à IA. De outro lado, a utilização dessa teoria se justifica quando baseada na ideia de que o viés algorítmico nas decisões da IA autônoma representam, sim, um risco. A responsabilidade objetiva baseada no risco dependeria, assim, de se compreender o alcance do termo “risco” e se ele também poderia ser aplicado às hipóteses em que um produto de IA é desenvolvido para mitigar ou eliminar o risco inerente a determinada atividade quando desenvolvida por um humano.

A segunda solução analisada foi a responsabilidade objetiva pelo fato do produto, prevista no art. 931 do Código Civil.

Conforme tratado no item 5.2.2, o problema, aqui, se inicia pela própria regra do art. 931, cuja utilidade é bastante questionada na doutrina, seja porque há uma regra mais específica sobre a responsabilidade do produtor no art. 12 do CDC, seja porque a própria redação do art. 931 é dúbia quanto à necessidade ou não de prova de um defeito do produto. Tula Wesendonck (2015) aponta que, diferentemente do art. 12 do CDC, o art. 931 do Código Civil permitiria a responsabilidade do fabricante independentemente da prova do defeito do produto, bem como que haveria no dispositivo um fundamento para a responsabilidade pelos riscos do desenvolvimento. Daniel Carnaúba (2020), de outro lado, indica que a responsabilidade pelos riscos do desenvolvimento é igualmente extraída no CDC e que o art. 931 não teria utilidade, ao fim e ao cabo, porque o art. 17 do CDC autoriza a aplicação da regra da responsabilidade objetiva pelo fato do produto a outras relações para além da de consumo, sendo que o art. 12 do CDC exige, a seu turno, a prova do defeito. Assim, embora não expressamente mencionado no art. 931, o defeito do produto seria indispensável para a responsabilidade do fabricante.

A prova do defeito e a responsabilidade pelos riscos do desenvolvimento são pontos fundamentais na análise da responsabilidade pelo fato do produto. Todavia, a prova de defeito na IA pode ser impossível, notadamente porque, em diversas situações, a IA poderá causar danos a terceiros em seu perfeito funcionamento. Já quanto aos riscos do desenvolvimento, tem-se que produtos de IA baseados em algoritmos de *deep learning* produzirão resultados (e a própria programação) a partir da interação com humanos, de modo que seria impossível ao produtor prever para qual finalidade um robô autônomo seria utilizado por humanos ou que tipo de informação seria apresentada ao robô para subsidiar seu aprendizado, motivos que poderiam, inclusive, afastar a sua responsabilidade, por excluírem o nexo de causalidade entre a causa do dano (aprendizado estimulado pelo usuário) e o evento danoso.

Nesse caso, novamente, há regra, mas o seu alcance é dúbio. A lacuna é, portanto, de reconhecimento: a IA é um produto, para fins de aplicação do art. 931? Ou ela deve ser vista como uma criança pequena ou animal, atraindo outro tipo de responsabilidade? Além disso, é preciso demonstrar que a IA agiu com defeito? Como seria feita essa prova em sistemas de *deep learning*? O aprendizado proporcionado pelo usuário pode excluir o nexo de causalidade com relação ao produtor que desenvolveu o sistema baseado em técnicas de aprendizado? Há muitos questionamentos e poucas respostas conclusivas, razão pela qual essa teoria encontra pouco respaldo na doutrina, servindo apenas para produtos de IA que são utilizados como meio, isto é, que não são, efetivamente, autônomos (PAGALLO, 2013, p. 126).

Portanto, embora seja mais fácil reconhecer a IA autônoma como um produto em comparação com uma atividade “de risco”, também aqui se tem um potencial problema de *afastamento* da regra, e não de lacuna, na medida em que, sendo o defeito do produto um requisito para a responsabilidade do produtor (ponto de vista da doutrina majoritária e do CDC), não haveria a subsunção à regra dos casos de danos causados por uma IA em perfeito funcionamento. Por tais motivos, a responsabilidade pelo fato do produto, nos moldes tradicionais, se mostra igualmente insuficiente para solucionar casos envolvendo IA autônoma.

A terceira hipótese analisada foi a da responsabilidade civil pelo fato da coisa ou animal, prevista nos artigos 936, 937 e 938 do Código Civil. Nesse caso, diferentemente da responsabilidade pelo fato do produto, não é o fabricante ou desenvolvedor da IA que responderia objetivamente pelos danos, mas o “guardião” ou proprietário da IA, conforme tratado no item 5.2.4.

Um motivo a favor dessa abordagem, especialmente em face da característica de aprendizado de máquina, é que o usuário ou proprietário da IA estará mais próximo do robô e exercerá maior influência sobre os seus processos de aprendizado, tendo melhores condições de prever ou evitar determinado comportamento ilícito (PAGALLO, 2013). Porém, pode-se questionar a equiparação da IA autônoma a coisas ou animais, argumentando-se que a IA apresentaria uma inteligência superior (ČERKA, GRIGIENE E SIRBIKYTE, 2015).

A responsabilidade pelo fato da coisa não possui uma regra geral quanto ao dever de guarda sobre a coisa, tal como ocorre no art. 1.384 do Código Civil francês<sup>130</sup> (FRANÇA, 1804). O art. 936 fala da responsabilidade pelo dono ou detentor do animal, ao passo que os artigos 937 e 938 responsabilizam o dono do edifício e o habitante de prédio. Assim, conforme tratado no item 5.2.3, a aplicação dessa teoria para além dos casos descritos no Código Civil se desenvolveu na doutrina e na jurisprudência, havendo entendimento consolidado no STJ a respeito da responsabilidade objetiva do proprietário do veículo, por exemplo.

Essa situação permitiria concluir pela existência de uma lacuna normativa em caso de responsabilidade pelo fato da coisa à IA, porque a relação do proprietário/guardião do produto de IA não se assemelha ao dono ou detentor de um animal e tampouco ao dono do edifício e ao habitante do prédio. Mas, devido à interpretação doutrinária e jurisprudencial, seria também possível concluir que se trata, em verdade, de uma lacuna de reconhecimento, em que seria preciso compreender o alcance das regras sobre a responsabilidade pelo fato da coisa a produtos de IA autônoma.

---

<sup>130</sup> Após a reforma do direito dos contratos e obrigações ocorrida em 2016, a regra passou a constar do art. 1242.

A segunda questão apresentada no desenvolvimento do trabalho sobre este ponto diz respeito a imputação do dever de indenizar ao proprietário ou usuário direto da IA autônoma. Ugo Pagallo (2013) aponta que o usuário é quem interage diretamente com o sistema e tem melhores condições de acompanhar o processo de aprendizado da IA, razão pela qual faria mais sentido imputar-lhe a responsabilidade. Objetou-se, porém, que o usuário, na maior parte dos casos, provavelmente não disporá de conhecimentos técnicos necessários para identificar um erro no sistema ou corrigir um comportamento potencialmente danoso da IA autônoma. Aliás, mesmo usuários conhecedores de programação poderão não reconhecer uma situação de risco na IA autônoma quando esta for baseada em algoritmos de *deep learning*, havendo espaço para a discussão sobre a ocorrência de caso fortuito que, ao fim e ao cabo, simplesmente afaste a responsabilidade do proprietário/usuário. Assim, parece fazer mais sentido responsabilizar o programador ou o fabricante que possuem os códigos-fonte da IA e conhecem a fundo o seu sistema de base, que muitas vezes estará protegido pelo sigilo comercial e não será informado mesmo a usuários capacitados em computação.

Considerando, todavia, que a responsabilidade pelo fato da coisa não exige que a coisa apresente defeito e permite a responsabilidade objetiva do proprietário/guardião, os dois principais óbices residem na caracterização da IA como coisa ou a sua equiparação a um animal, com relação ao qual o produto de IA será tendencialmente mais inteligente, e na justificativa para se deslocar a responsabilidade do programador ou fabricante para a responsabilidade exclusiva do proprietário ou usuário. A possibilidade de se assegurar a responsabilidade pelo fato da coisa com algum mecanismo de responsabilidade solidária decorrente da lei, todavia, pode resolver o problema. Embora as regras de responsabilidade pelo fato da coisa não se amoldem exatamente às hipóteses de IA autônoma, parece razoável equiparar o produto de IA autônomo a um bem e reconhecer a responsabilidade objetiva e solidária do proprietário, do programador e do fabricante, por todos exercerem, direta ou indiretamente, a guarda da IA, porque sujeita a atualizações periódicas do *software* pelo fornecedor. Para se alcançar essa solução, todavia, seria imprescindível a edição de uma regra específica, já que o Código Civil brasileiro não prevê possibilidade similar nos artigos 936, 937 e 938 e, nos termos do art. 265 do Código Civil, a obrigação solidária decorrerá da lei ou da vontade das partes, não podendo ser presumida.

Todavia, se acolhido o argumento da inaplicabilidade da teoria da responsabilidade pelo fato da coisa à IA autônoma, por se entender que a IA possui uma inteligência superior que impede sua equiparação aos animais ou coisas, resta, ainda, a responsabilidade por fato de outrem ou, conforme doutrina desenvolvida na *Common Law*, *vicarious liability*.

Os dois principais exemplos de responsabilidade indireta decorrem da relação entre pais e filhos menores e empregadores e empregados, ambos previstos no art. 932 do Código Civil de 2002. Conforme analisado no item 5.3, a responsabilidade indireta é objetiva com relação aos pais ou empregadores, mas subjetiva com relação aos filhos menores e empregados, sendo imprescindível, especialmente neste segundo caso, a demonstração de culpa, conforme exige a Súmula 341 do STF (BRASIL, 1963).

Paula Giliker (2010) apontou diversos fundamentos para esse tipo de responsabilidade, sendo os principais a culpa *in vigilando* e *in eligendo*, presentes no Código Civil alemão, a garantia de compensação da vítima e o risco assumido pelo terceiro a quem se imputa a responsabilidade ou dissuasão de uma determinada conduta. A princípio, todas as hipóteses de responsabilidade envolvem um agente humano (o filho menor, o empregado, o pupilo ou curatelado etc.), donde surge a primeira questão sobre a aplicabilidade dessa teoria: poderia a IA autônoma ser equiparada a um desses agentes?

Ugo Pagallo (2013) entende que sim. O italiano divide a IA autônoma em duas categorias: robôs como meios, sujeitos à responsabilidade pelo fato do produto, e robôs como agentes, estes equiparados a empregados. Jan-Erik Schirmer (2019) pontua que seria possível atribuir uma capacidade jurídica parcial à IA autônoma, reconhecendo-lhe apenas alguns direitos, num estatuto jurídico *sui generis* similar ao que é hoje reconhecido aos animais. Nesse caso, a IA não seria reconhecida propriamente como uma pessoa, o que evita a discussão sobre a personalidade eletrônica.

Talvez essa seja a abordagem mais apropriada, considerando a possibilidade de desenvolvimento de sistemas de inteligência artificial que poderão superar não apenas a inteligência de animais, como também a própria inteligência humana. A percepção de robôs e andróides como agentes que efetivamente se movimentam pelo mundo físico tende a facilitar o seu reconhecimento como um “agente”, subordinado aos humanos que detém sua propriedade, e não como um objeto.

Essa solução, todavia, demandaria ainda a superação de um segundo óbice: a prova da culpa da IA. Schirmer (2019) aponta que o regime de responsabilidade deveria ser inteiramente objetivo e, com efeito, não parece haver outra possibilidade, pelos mesmos problemas apontados no tópico 5.1 sobre as dificuldades de se analisar o elemento subjetivo de um programa de computador que não compartilha dos mesmos parâmetros éticos dos humanos. Como não há nenhuma regra nesse sentido, a hipótese, aqui, é de lacuna normativa.

Portanto, a responsabilidade indireta dos humanos por atos praticados pelos “agentes” de IA autônoma só seria possível (*i.*) pela equiparação da IA a um empregado, subordinado aos

humanos; e (ii.) pela aplicação de um regime inteiramente objetivo de responsabilidade. E tal solução demanda, efetivamente, uma alteração legislativa, já que, atualmente, o Código Civil não dispõe de uma regra disciplinando a responsabilidade indireta de humanos por atos praticados por empregados que prescindia da comprovação da culpa – pelo contrário, seria necessário estabelecer que, apesar da equiparação com empregados, a Súmula 341 do STF não poderia ser aplicada a essas hipóteses específicas.

Por fim, analisou-se a responsabilidade objetiva lastreada na ilicitude objetiva do art. 187 do Código Civil (item 5.4). Nesse caso, o debate desloca-se da *responsabilidade* para a compreensão da *ilicitude*. Demonstrou-se, porém, que, a partir de uma teoria externa de restrição a direitos fundamentais, conforme sustenta Cachapuz, seria possível reconhecer situações em que o titular de um direito subjetivo exerce a posição jurídica em que se encontra de maneira irregular, concluindo-se, a partir da ponderação das circunstâncias do caso concreto, pela caracterização de um ilícito.

Essa hipótese permitiria resolver situações de danos causados pela IA autônoma, porque parte-se não da ótica de uma obrigação de resultado (como ocorre na guarda da coisa), mas, sim, de um direito geral de liberdade – no caso, de adquirir e utilizar uma IA autônoma – que, se exceder à boa-fé objetiva, os bons costumes ou o fim econômico ou social a que se presta o produto de IA, dará azo à obrigação de indenizar, quando ocorrido o dano. Resolve-se a questão pela combinação do *caput* do art. 927 com o art. 187 do Código Civil.

Assim, embora a responsabilidade objetiva se mostre mais adequada para permitir uma efetiva solução dos casos de responsabilidade civil envolvendo a IA autônoma, as teorias clássicas reconhecidas no Código Civil brasileiro não apresentam completa adesão ao problema. Em todos os casos, há um elemento que obsta uma solução adequada: a possível inexistência de risco (art. 927) ou de defeito no produto de IA (art. 931), a impossibilidade de se equiparar a IA autônoma a uma coisa ou animal devido ao seu grau de inteligência, ou mesmo a inadequação de uma responsabilidade que recaia exclusivamente sobre um proprietário leigo em questões de computação, e, por fim, o desafio de se equiparar a IA autônoma a um agente e demonstrar a sua culpa para fins de responsabilidade indireta do humano “empregador”. Resta, contudo, a solução fornecida pelo art. 187 do Código Civil, que mostrou-se, *a priori*, suficiente para resolver o problema.

Nada obstante, diante das dificuldades para se responsabilizar humanos e considerando o fato de que uma IA autônoma seria capaz de agir de modo totalmente independente, há uma segunda via: a responsabilização da própria IA autônoma, mediante a criação de um estatuto jurídico próprio, a personalidade eletrônica (capítulo 6).

Também sob essa perspectiva se identificam os desafios já apontados no capítulo 4 sobre a prova da culpa da IA, o nexos de causalidade entre a conduta da IA e o dano e, inclusive, o reconhecimento de estado de necessidade e legítima defesa para robôs (capítulo 3). Mas é preciso dar um passo atrás, e entender não apenas os pressupostos da responsabilidade civil, mas, essencialmente, o conceito de *pessoa*. E isso porque, se o ordenamento jurídico reconhece a necessidade de se reparar um dano causado à vítima, é porque reconhece, em primeiro lugar, a dignidade da pessoa humana.

Conforme apontado no capítulo 6, Alexy e Figueroa analisaram os requisitos para se estender direitos fundamentais a robôs, partindo de uma teoria debilitada de direitos humanos que permite enquadrá-los como pessoas. Para tanto, seria necessário reconhecer no robô três características: a inteligência, a capacidade de ter sentimentos e a consciência (de si mesmo, a autonomia e a capacidade de julgar os próprios atos). Há, todavia, relevantes impeditivos para essa análise. O primeiro é o modo como se irá reconhecer a inteligência ou a sensibilidade de um robô, já que, conforme apontado por Solum (1992), o fato de que uma IA pode simular um comportamento ou um sentimento não significa que ela entenda o que está fazendo. Além disso, Cachapuz (2019) lembra que a autonomia reconhecida a robôs, tendo como parâmetros as Leis de Asimov, é essencialmente mitigada.

Ademais, pontuou-se que, em razão de características como *machine learning* e *deep learning*, uma IA não poderia ser reconhecida como “pessoa”, de acordo com as características elencadas por Alexy e Figueroa, tão logo saia da fábrica. A sua auto consciência, autonomia e mesmo a capacidade de sentir só serão desenvolvidas em um ambiente de interação, sendo imprescindível, na hipótese de se reconhecer a efetiva possibilidade de extensão do conceito de pessoa a entes dotados de IA, que essa outorga seja feita de maneira casuística.

De todo modo, mesmo se reconhecendo a personalidade eletrônica, vinculada a regimes de seguro obrigatório e fundos de compensação, ainda assim haveria dificuldade de se definir o regime aplicável à IA (se objetivo ou, já que identificadas características intrínsecas ao conceito de pessoa, subjetivo) e as hipóteses em que a IA deveria ser responsabilizada. Nesse caso, há uma lacuna normativa, mas, também, potencial lacuna axiológica que precisa ser considerada quando da formulação de regras que reconheçam um traço de personalidade – até hoje reconhecido apenas a humanos – também às máquinas.

Em conclusão, considerando as lacunas de cada um dos pontos abordados ao longo do trabalho, sugere-se que uma solução adequada para a responsabilidade civil de sistemas de IA autônoma não deve implicar na imputação do dever de indenizar somente ao proprietário, dada a natureza complexa do produto e a constante necessidade de atualização do *software*. Some-

se a isso o argumento econômico mencionado por David Vladeck (2014) sobre a doutrina do empreendimento comum: quem coloca produtos em circulação tem as melhores condições de precificar os seus custos e absorver os valores eventualmente empregados no pagamento de indenizações.

Assim, deve haver a responsabilidade do proprietário, que se beneficia diretamente da IA e lhe fornece a maioria das informações. Mas a responsabilidade do proprietário não pode excluir a responsabilidade dos programadores e fabricantes, que conhecem o sistema e beneficiam-se da sua comercialização. Em qualquer caso, a IA autônoma deve ter sido a causa *necessária* para a ocorrência do dano, de acordo com a teoria da interrupção donexo causal.

Nesse panorama, conclui-se que há, quando muito, uma lacuna de reconhecimento no Código Civil brasileiro, que demanda a interpretação do alcance de termos como “risco” e “ilicitude”. Sugere-se, por isso, as seguintes soluções para os danos causados pela IA autônoma:

- (i.) Responsabilidade objetiva pela teoria do risco, que se fundamenta no viés algorítmico, com a imputação do dever de indenizar a quem se utiliza e comercializa a IA, nos termos do art. 927, parágrafo único, do Código Civil;
- (ii.) Responsabilidade objetiva pela teoria da guarda da coisa, desde que a IA seja equiparada a um objeto, mas sugere-se que seja assegurada a possibilidade de responsabilização solidária do proprietário, do programador e do fabricante, já que todos desempenham um papel na criação, manutenção e treinamento do algoritmo (com a necessidade de alteração legislativa, em razão do art. 265 do Código Civil);
- (iii.) Responsabilidade objetiva daquele que se utilizar da IA autônoma de modo contrário à boa-fé objetiva, aos bons costumes ou ao fim econômico e social, nos termos do art. 187 do Código Civil, sendo, também aqui, possível discutir-se eventual obrigação solidária ou direito de regresso do titular do direito subjetivo perante o programador ou fabricante da IA.

A operacionalização de tais sugestões evidentemente demanda profunda discussão na sociedade. A IA autônoma pode trazer inúmeros benefícios à humanidade, na mesma proporção em que traz desafios, muitos deles completamente desconhecidos. Porém, uma premissa que não pode ser esquecida na necessária análise sobre os parâmetros éticos de desenvolvimento de inteligência artificial e a responsabilidade civil para os danos decorrentes de sua utilização é que sistemas de IA autônoma são desenvolvidos por humanos e no interesse de humanos. Portanto, independentemente de como o “algoritmo” que soluciona essa questão for escrito (ou interpretado a partir das “instruções” já escritas) nos Códigos Civis, é imprescindível que se reconheça o elemento humano como parte do processo – e, por consequência, responsável.

## Bibliografia

- ALABART, Silvia Díaz. **Robots y Responsabilidad Civil**. Madrid: Reus, 2018.
- ALEXY, Robert; FIGUEROA, Alfonso García. **Star Trek y los derechos humanos**. Valencia: Tirant lo Blanch, 2007.
- ANGWIN, Julia *et al.* **Machine Bias**. 2016. Disponível em: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>. Acesso em: 25 mai. 2020.
- ANTUNES, Henrique Sousa. Inteligência Artificial e Responsabilidade Civil: enquadramento. **Revista de Direito da Responsabilidade**, Lisboa, Ano 1, p. 139–154, 2019.
- ARAÚJO, Justino Magno. Responsabilidade civil pelo fato da coisa. **Doutrinas Essenciais de Responsabilidade Civil**, São Paulo, v. 4, p. 663–674, out. 2011.
- ARENDT, Hannah. Reflections on Little Rock. **Dissent Magazine**, New York, p. 45–56, 1957.
- ARENDT, Hannah. **Responsabilidade e julgamento**. São Paulo: Companhia das Letras, 2004.
- ASIMOV, Isaac. **Eu, Robô**. 1. ed. São Paulo: Aleph, 2014.
- BARBOSA, Mafalda Miranda. Inteligência Artificial, E-Persons e Direito: Desafios e Perspetivas. **Revista Jurídica Luso Brasileira**, Lisboa, v. 6, n. 3, p. 1475–1503, 2017.
- BARBOSA, Mafalda Miranda. O futuro da responsabilidade civil desafiada pela inteligência artificial: as dificuldades dos modelos tradicionais e caminhos de solução. **Revista de Direito da Responsabilidade**, Lisboa, Ano 2, p. 280–326, 2020.
- BAUMAN, Zygmunt; LYON, David. **Vigilância Líquida**. 1. ed. Rio de Janeiro: Zahar, 2013.
- BESSA, Leonardo Roscoe. Responsabilidade pelo fato do produto: questões polêmicas. **Revista de Direito do Consumidor**, São Paulo, v. 89, p. 141–163, set./out. 2013.
- BULFINCH, Thomas. **O Livro da Mitologia: Histórias de Deuses e Heróis**. São Paulo: Martin Claret, 2006.
- CACHAPUZ, Maria Cláudia. O conceito de pessoa e a autonomia de data (ou sobre a medida da humanidade em tempos de inteligência artificial). **Revista de Direito Civil Contemporâneo**, São Paulo, v. 6, n. 20, p. 63–85, jul./set. 2019.
- CACHAPUZ, Maria Cláudia. **A obrigação pelo discurso jurídico**. Porto Alegre: Sergio Antonio Fabris, 2017.
- CACHAPUZ, Maria Cláudia. A ilicitude e as fontes obrigacionais: análise do art. 187 do novo Código Civil brasileiro. **Doutrinas Essenciais Obrigações e Contratos**, São Paulo, v. 1, p. 1273 – 1296, jun. 2011.

ČAPEK, Karel. **R.U.R. (Rossum's Universal Robots)**. New York: Dover Publications, 2001.

CARNAÚBA, Daniel Amaral. Para que serve o art. 931 do Código Civil? Considerações críticas sobre um dispositivo inútil. **Revista de Direito Civil Contemporâneo**, São Paulo, v. 22, p. 203–239, jan./mar. 2020.

CAVALLI, Sabina. Natureza jurídica da obrigação de implantação de software em contratos de licença de uso concluídos com usuário pessoa jurídica, e as consequências do seu não-cumprimento. **Revista de Direito do Consumidor**, São Paulo, v. 65, n. 17, p. 253–282, jan./mar. 2008.

CAVALIERI, Sergio Filho. **Programa de Responsabilidade Civil**. 10ª ed. São Paulo: Atlas, 2012.

ČERKA, Paulius; GRIGIENE, Jurgita; SIRBIKYTE, Gintare. Liability for damages caused by artificial intelligence. **Computer Law and Security Review**, London, v. 31, n. 3, p. 376–389, 2015.

COMPARATO, Fábio Konder. Obrigações de meios, de resultado e de garantia. **Doutrinas Essenciais de Direito Empresarial**, São Paulo, v. 4, p. 63-78, dez. 2010.

COSTA, Mario Julio de Almeida. **Direito das obrigações**. 12ª ed. Coimbra: Almedina, 2011.

COSTA, Mario Julio de Almeida. **Direito das obrigações**. 9ª Ed. Coimbra: Almedina, 2001.

CURSI, Floriana M. What Did Occidere Iniuria in the Lex Aquilia Actually Mean. **Roman Legal Tradition**, v. 7, p. 16-29, 2011. Disponível em: <https://heinonline.org/HOL/P?h=hein.journals/rltrad7&i=28>. Acesso em: 01 nov. 2020.

DINIZ, Maria Helena. **Curso de direito civil brasileiro – vol. 7. Responsabilidade civil**. São Paulo: Saraiva, 1990.

DRESSEL, Julia; FARID, Hany. The accuracy, fairness, and limits of predicting recidivism. **Science Advances**, Hanover, v. 4, n. 1, p. 1–6, jan. 2018.

DWORKIN, Ronald. **Levando os Direitos a Sério**. São Paulo: Martins Fontes, 2002.

ELECTRONIC PRIVACY INFORMATION CENTER. **Algorithms in the Criminal Justice System: Risk Assessment Tools**. 2020. Disponível em: <https://epic.org/algorithmic-transparency/crim-justice/>. Acesso em: 14 mai. 2020.

FEDERAL BUREAU OF INVESTIGATION. **2016 Crime in the United States: Arrests by Race and Ethnicity, 2016**. 2016. Disponível em: <https://ucr.fbi.gov/crime-in-the-u.s/2016/crime-in-the-u.s.-2016/tables/table-21>. Acesso em: 24 out. 2020.

FERREIRA DA SILVA, Jorge Cesa. **Inadimplemento das Obrigações**. São Paulo: Revista dos Tribunais, 2007.

GAGLIANO, Pablo Stolze; PAMPLONA FILHO, Rodolfo. Novo Curso de Direito Civil: Responsabilidade Civil. 17ª Ed. São Paulo: Saraiva Educação, 2019. *E-book*. Disponível em: <https://bridge.minhabiblioteca.com.br/#/books/9788553609529/>. Acesso em: 02 nov. 2020.

GILIKER, Paula. **Vicarious Liability in Tort: a Comparative Perspective**. Cambridge: Cambridge University Press, 2010.

GRIFFIN, Andrew. Facebook's Artificial Intelligence Robots shut down after they start talking to each other in their own language. **The Independent**, 2017. Disponível em: <https://www.independent.co.uk/life-style/facebook-artificial-intelligence-ai-chatbot-new-language-research-openai-google-a7869706.html>. Acesso em: 15 mai. 2020.

HABERMAS, Jürgen. **A Ética da Discussão e a Questão da Verdade**. 3ª ed. São Paulo: WMF Martins Fontes, 2013.

HART, Herbert L. A. **O Conceito de Direito**. 6ª ed. Lisboa: Fundação Calouste Gulbenkian, 2011.

JOHN MCCARTHY COMPUTER SCIENCE DEPARTMENT. What is Artificial Intelligence? **Stanford University**, 2007. Disponível em: <http://jmc.stanford.edu/artificial-intelligence/>. Acesso em: 25 ago. 2019.

KARNOW, Curtis E. A. The application of traditional tort theory to embodied machine intelligence. In: CALO, Ryan; FROOMKIN, A. Michael; KERR, Ian (org.). **Robot Law**. Cheltenham: Edward Elgar Publishing, 2016. p. 51–77.

KLEMME, Howard C. The Enterprise Liability Theory of Torts. **University of Colorado Law Review**, [S. l.], v. 47, n. 153, p. 153–231, 1975.

KNIGHT, Will. The dark secret at the heart of AI. **MIT Technology Review**, abr. 2017. Disponível em: <https://www.technologyreview.com/2017/04/11/51113/the-dark-secret-at-the-heart-of-ai/>. Acesso em 10 abr. 2020.

MANERO, Juan Ruiz. Sistema Jurídico: Lagunas y Antinomias. In: LAGIER, Daniel González (coord.). **Conceptos Básicos del Derecho**. Madrid: Marcial Pons, 2015.

MARTINS-COSTA, Judith. A linguagem da responsabilidade civil. In: BIANCHI, José Flávio De; GOMES, MENDONÇA PINHEIRO. Rodrigo; ARRUDA ALVIM, Teresa (org.). **Jurisdição e Direito Privado: Estudos em homenagem aos 20 anos da Ministra Nancy Andrighi no STJ**. São Paulo: Revista dos Tribunais, 2020. p. 389–418.

MARTINS-COSTA, Judith. **Comentários ao novo Código Civil. Vol. V, tomo II: do inadimplemento das obrigações**. Rio de Janeiro: Forense, 2009.

MARTINS-COSTA, Judith. **Os avatares do abuso do direito e o rumo indicado pela boa-fé**. Canela, 2006. Disponível em: <http://www.fd.ulisboa.pt/wp-content/uploads/2014/12/Costa-Judith-Os-avatares-do-Abuso-do-direito-e-o-rumo-indicado-pela-Boa-Fe.pdf>. Acesso em: 29 nov. 2020.

MARTINS-COSTA, Judith. Breves anotações acerca do conceito de ilicitude no nCC (estruturas e rupturas em torno do art. 187). **Migalhas**, 23 mai. 2003. Disponível em: <https://migalhas.uol.com.br/depeso/1726/breves-annotacoes-acerca-do-conceito-de-ilicitude-no-ncc>. Acesso em: 29 nov. 2020.

MOREIRA ALVES, José Carlos. **Direito Romano**. 19ª Ed. Rio de Janeiro: Forense, 2019. *E-book*. Disponível em: <https://bridge.minhabiblioteca.com.br/#/books/9788530984274/>. Acesso em: 01 nov. 2020.

NORTHPOINTE INC. **Practitioner's Guide to COMPAS Core**. Disponível em: [http://www.northpointeinc.com/downloads/compas/Practitioners-Guide-COMPAS-Core\\_031915.pdf](http://www.northpointeinc.com/downloads/compas/Practitioners-Guide-COMPAS-Core_031915.pdf). Acesso em 05 out. 2019.

O'NEILL, Cathy. **Weapons of Math Destruction: how big data increases inequality and threatens democracy**. 1. ed. New York: Crown, 2016.

OMOHUNDRO, Stephen M. The basic AI drives. **Proceedings of the 2008 conference on Artificial General Intelligence**, Amsterdam, v. 171, n. 1, p. 483–492, 2008.

PAGALLO, Ugo. **The Laws of Robots: Crimes, Contracts, and Torts**. Dordrecht: Springer, 2013.

PEREIRA, Caio Mário da Silva. **Comentários ao Código Civil de 2002: v. 1. Parte geral: arts. 1º a 232**. Rio de Janeiro: GZ, 2017.

PEREIRA, Caio Mário da Silva. **Instituições de direito civil: contratos: declaração unilateral de vontade: responsabilidade civil**. 22ª ed. Rio de Janeiro: Forense, 2018.

PEREIRA, Caio Mário da Silva. **Responsabilidade Civil**. 12ª Ed. Rio de Janeiro: Forense, 2018. *E-book*. Disponível em: <https://bridge.minhabiblioteca.com.br/#/books/9788530980320/>. Acesso em: 01 nov. 2020.

RAJARAMAN, V. JohnMcCarthy — Father of artificial intelligence. **Resonance**, [s.l.] v. 19, n. 3, p. 198–207, 2014. DOI: <https://doi.org/10.1007/s12045-014-0027-9>.

RICHARDS, Neil M.; SMART, William D. How should law think about robots? *In*: CALO, RYAN; FROOMKIN, A.; KERR, Ian (org.). **Robot Law**. Cheltenham: Edward Elgar Publishing, 2016. p. 3–22.

RODRIGUES, Silvio. **Direito Civil: responsabilidade civil**. São Paulo: Saraiva, 1975.

RUSSEL, Stuart Jonathan; NORVIG, Peter. **Artificial Intelligence: a modern approach**. New Jersey: Prentice Hall, 1995.

SANSEVERINO, Paulo de Tarso Vieira. **Princípio da Reparação Integral: indenização no Código Civil**. 1ª Ed. São Paulo: Saraiva, 2010.

SCHERER, Maxi. Artificial intelligence and legal decision-making: The wide open? a study examining international arbitration. **Journal of International Arbitration**, Amsterdam, v. 36, n. 5, p. 539–573, 2019.

SCHIRMER, Jan-Erik. Artificial Intelligence and Legal Personality: Introducing “Teilrechtsfähigkeit”: A Partial Legal Status Made in Germany. *In*: WISCHMEYER, Thomas; RADEMACHER, Timo (org.). **Regulating Artificial Intelligence**. Switzerland: Springer, 2019. p. 124–142.

SCHREIBER, Anderson. **Manual de direito civil contemporâneo**. São Paulo: Saraiva, 2018.

SCHWAB, Klaus. **La cuarta revolución industrial**. Barcelona: Debate, 2017.

SHELLEY, Mary. **Frankenstein: or the Modern Prometheus - The 1818 text**. New York: Oxford University Press, 2008.

SILVA, Rafael Peteffi da; PIRES, Thatiane Cristina Fontão. A responsabilidade civil pelos atos autônomos da inteligência artificial: notas iniciais sobre a resolução do Parlamento Europeu. **Revista Brasileira de Políticas Públicas**, Brasília, v. 7, n. 3, p. 238–254, dez. 2017.

SOLUM, Lawrence B. Legal Personhood for Artificial Intelligences. **North Carolina Law Review**, Chapel Hill, v. 70, n. 9, p. 1231–1287, 1992. Disponível em: [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1108671](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1108671).

SOMMERVILLE, Ian. **Engenharia de software**. 9ª ed. São Paulo: Pearson Prentice Hall, 2011.

SUMPTER, David. **Dominados pelos números**. Rio de Janeiro: Bertrand Brasil, 2019.

TEPEDINO, Gustavo; SILVA, Rodrigo da Guia. Desafios da inteligência artificial em matéria de responsabilidade civil. **Revista Brasileira de Direito Civil**, Belo Horizonte, v. 21, p. 61–86, jul./set. 2019. DOI: <https://doi.org/10.33242/rbdc.2019.03.004>.

THE SOCIAL DILEMMA. Direção: Jeff Orlowski. Estados Unidos: Netflix, 2020. Disponível em: <https://www.thesocialdilemma.com/the-film/credits/>.

TOVO, Antônio. O periculosômetro digital? **AB2L**, jan. 2019. Disponível em: <https://www.ab2l.org.br/o-periculosometro-digital/>. Acesso em 15 mar. 2020.

TURING, Alan M. COMPUTING MACHINERY AND INTELLIGENCE. **Mind, New Series**, [s.l.], v. 59, n. 236, p. 433–460, 1950. Disponível em: <http://www.jstor.org/stable/2251299>.

VLADÉCK, David C. Machines without principals: Liability rules and artificial intelligence. **Washington Law Review**, [s.l.], v. 89, n. 1, p. 117–150, 2014.

WACHTER, Sandra; MITTELSTADT, Brent; FLORIDI, Luciano; Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation. **International Data Privacy Law**, [s.l.], v. 7, n. 2, p. 76–99, 2017.

WAINWRIGHT, Martin. Robot fails to find a place in the sun. **The Guardian**, 20 jun. 2002. Disponível em: <https://www.theguardian.com/uk/2002/jun/20/engineering.highereducation>. Acesso em: 12 mar. 2020.

WASHINGTON, Anne L. How to Argue with an Algorithm: Lessons from the COMPAS-ProPublica Debate. **The Colorado Technology Law Journal**, [s.l.], v. 17, n. 1, p. 131–160, 2019.

WESENDONCK, Tula. Art. 931 do Código Civil: repetição ou inovação? **Revista de Direito Civil Contemporâneo - RDCC (Journal of Contemporary Private Law)**, São Paulo, v. 3, p. 141–160, abr./jun. 2015.

### **Legislação consultada:**

BRASIL. **Decreto-Lei n.º 4.657, de 4 de setembro de 1942**. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/decreto-lei/del4657compilado.htm](http://www.planalto.gov.br/ccivil_03/decreto-lei/del4657compilado.htm). Acesso em: 15 abr. 2020.

BRASIL. **Lei n.º 3.701, de 1º de janeiro de 1916**. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/leis/13071.htm](http://www.planalto.gov.br/ccivil_03/leis/13071.htm). Acesso em: ago. 2020.

BRASIL. **Lei n.º 8.078, de 11 de setembro de 1990**. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/leis/18078compilado.htm](http://www.planalto.gov.br/ccivil_03/leis/18078compilado.htm). Acesso em: set. 2020.

BRASIL. **Lei n.º 9.609, de 19 de fevereiro de 1998**. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/leis/L9609.htm](http://www.planalto.gov.br/ccivil_03/leis/L9609.htm). Acesso em: 15 mar. 2020.

BRASIL. **Lei n.º 10.406, de 10 de janeiro de 2002**. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/leis/2002/110406compilada.htm](http://www.planalto.gov.br/ccivil_03/leis/2002/110406compilada.htm). Acesso em: ago. 2020.

BRASIL. **Projeto de Lei n.º 5051 de 2019**. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/138790>. Acesso em: 20 jun. 2020.

BRASIL. **Projeto de Lei n.º 5961 de 2019**. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/139785>. Acesso em: 20 jun. 2020.

BRASIL. **Projeto de Lei n.º 21 de 2020**. Disponível em: <https://www.camara.leg.br/propostas-legislativas/2236340>. Acesso em: 20 jun. 2020.

BRASIL. **Projeto de Lei n.º 240 de 2020**. Disponível em: <https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=2236943>. Acesso em 20 jun. 2020.

FRANÇA. **Code Civil**. Disponível em: [https://www.legifrance.gouv.fr/codes/section\\_lc/LEGITEXT000006070721/LEGISCTA000006090271/2016-10-01/#LEGISCTA000006090271](https://www.legifrance.gouv.fr/codes/section_lc/LEGITEXT000006070721/LEGISCTA000006090271/2016-10-01/#LEGISCTA000006090271). Acesso em: 11 out. 2020.

ITÁLIA. **Codice Civile**. Disponível em: <https://www.studiocataldi.it/codicecivile/codice-civile.pdf>. Acesso em: 19 out. 2020.

PORTUGAL. **Código Civil Português**. Disponível em:

<https://www.igac.gov.pt/documents/20178/358682/C%C3%B3digo+Civil.pdf/2e6b36d8-876b-433c-88c1-5b066aa93991>. Acesso em: 29 nov. 2020.

UNIÃO EUROPEIA. **Resolução do Parlamento Europeu, de 16 de fevereiro de 2017, que contém recomendações à Comissão sobre disposições de Direito Civil sobre Robótica (2015/2103(INL))**. Disponível em: [http://www.europarl.europa.eu/doceo/document/TA-8-2017-0051\\_PT.html?redirect](http://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_PT.html?redirect). Acesso em: 10 dez. 2019.

### **Jurisprudência consultada:**

BRASIL. **Superior Tribunal de Justiça**. Recurso Especial n.º 1.288.008/MG. Recorrente: Antonio Rodrigues Cordeiro. Recorrida: Primo Schincariol Indústria de Cervejas e Refrigerantes S/A. Relator: Ministro Paulo de Tarso Sanseverino. Brasília, 4 abr. 2013. Disponível em:

[https://processo.stj.jus.br/processo/revista/documento/mediado/?componente=ITA&sequencia=1218415&num\\_registro=201102481429&data=20130411&formato=PDF](https://processo.stj.jus.br/processo/revista/documento/mediado/?componente=ITA&sequencia=1218415&num_registro=201102481429&data=20130411&formato=PDF). Acesso em: 10 out. 2020.

BRASIL. **Superior Tribunal de Justiça**. Recurso Especial n.º 577.902/DF. Recorrente: Sérgio Rocha de Faria. Recorrido: Alessandro Dehon Picinin. Relator: Ministro Antônio de Pádua Ribeiro. Redatora para o Acórdão: Ministra Nancy Fátima Andrichi. Brasília, 13 jun. 2006. Disponível em:

[https://processo.stj.jus.br/processo/revista/documento/mediado/?componente=ITA&sequencia=552813&num\\_registro=200301571792&data=20060828&formato=PDF](https://processo.stj.jus.br/processo/revista/documento/mediado/?componente=ITA&sequencia=552813&num_registro=200301571792&data=20060828&formato=PDF). Acesso em: 12 out. 2020.

BRASIL. **Supremo Tribunal Federal**. Recurso Extraordinário 130.764/PR. Recorrentes: Estado do Paraná e Ministério Público do Estado do Paraná. Recorridos: H. Kaminski & CIA. LTDA. e outros. Relator: Ministro Moreira Alves. Brasília, 12 mai. 1992. Disponível em: <http://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=AC&docID=207632>. Acesso em: 30 set. 2020.

BRASIL. **Supremo Tribunal Federal**. Súmula n. 341. Brasília: Supremo Tribunal Federal, 13 dez. 1963. Disponível em: <https://jurisprudencia.stf.jus.br/pages/search/seq-sumula341/false>. Acesso em: 19 out. 2020.

INDIANA. **Supreme Court of Indiana**. Recorrente: Anthony Malenchik. Recorrido: Estado de Indiana. Relator: Justice Dickson. Indianapolis, 9 jun. 2010. Disponível em: <https://www.in.gov/judiciary/opinions/pdf/06091001bd.pdf>. Acesso em: 15 mar. 2020.

WISCONSIN. **Supreme Court of Wisconsin**. Recorrente: Eric. L. Loomis. Recorrido: Estado de Wisconsin. Relatora: Ann Walsh Bradley. Madison, 13 jul. 2016. Disponível em: <https://caselaw.findlaw.com/wi-supreme-court/1742124.html>. Acesso em: 22 nov. 2019.