



UNIVERSIDADE DE BRASÍLIA
PROGRAMA DE PÓS-GRADUAÇÃO EM DIREITO

MARINA DE ALENCAR ARARIPE COUTINHO

A PROTEÇÃO DA VÍTIMA DE DANOS CAUSADOS POR UMA INTELIGÊNCIA
ARTIFICIAL:
uma perspectiva civil-constitucional

BRASÍLIA
2022

Marina de Alencar Araripe Coutinho

**A PROTEÇÃO DA VÍTIMA DE DANOS CAUSADOS POR UMA INTELIGÊNCIA
ARTIFICIAL:
uma perspectiva civil-constitucional**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação da Faculdade de Direito da Universidade de Brasília, como requisito parcial à obtenção do título de Mestre.

Orientador: Prof. Dr. Fabiano Hartmann

Brasília

2022

Marina de Alencar Araripe Coutinho

A PROTEÇÃO DA VÍTIMA DE DANOS CAUSADOS POR UMA INTELIGÊNCIA
ARTIFICIAL: uma perspectiva civil-constitucional

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação da Faculdade de Direito da Universidade de Brasília, como requisito parcial à obtenção do título de Mestre.

Aprovada em 17 de fevereiro de 2022

Banca Examinadora

Prof. Dr. Fabiano Hartmann
(Orientador – UnB)

Profa. Dra. Caitlin Sampaio Mulholland
(PUC-Rio)

Profa. Dra. Fernanda de Carvalho Lage
(UnB)

AGRADECIMENTOS

No período em que finalizo os dois anos de mestrado, completam-se, também, dois anos de pandemia de COVID-19. Apesar das perdas, dos medos e das incertezas que decorreram do contexto sanitário, poder terminar esse processo em tempo e com saúde é um privilégio pelo qual me sinto extremamente grata.

Agradeço à minha família, pelo amor de sempre, pela oportunidade do estudo, pela lição da disciplina e pelo incentivo à leitura e à escrita.

Agradeço a meu marido, Emidio, por ser meu companheiro de vida e fonte de afeto e sanidade no isolamento pandêmico-acadêmico.

Agradeço ao meu orientador, Professor Fabiano Hartmann, por ter me apresentado ao mundo da inteligência artificial e compartilhado seu conhecimento de forma generosa, bem como aos demais professores do PPGD que contribuíram de alguma forma com esta pesquisa.

Agradeço aos amigos que emprestaram os ouvidos nesse período, com quem compartilhei minhas angústias e preocupações, em especial, Paola, Biba e Pezão.

Agradeço aos colegas da pós-graduação, que, sem me conhecer pessoalmente, mantiveram-me em contato com a Universidade e nutriram em mim um sentimento de coletividade.

Agradeço aos meus colegas de trabalho pelo incentivo, pela compreensão e pela deferência ao meu tempo de estudo e escrita.

"Whether AI will help us reach our aspirations or reinforce the unjust inequalities is ultimately up to us." Joy Buolamwini

RESUMO

O desenvolvimento da inteligência artificial, notadamente, da técnica do *machine learning*, acarretou o incremento do grau de autonomia com que computadores operam, característica que desafia o campo da responsabilidade civil, na medida em que pode suscitar dúvidas acerca da atribuição de responsabilidade por danos causados por essas máquinas, ante a dificuldade de se estabelecer um liame entre uma conduta humana e o injusto. Nesse contexto, a presente dissertação busca aferir a aptidão do ordenamento jurídico vigente para a proteção da vítima de danos causados pela inteligência artificial, adotando-se uma pesquisa qualitativa sob o método dedutivo, baseada no levantamento bibliográfico e na análise de proposições legislativas. O primeiro capítulo dedica-se à apresentação do objeto de pesquisa, de forma a especificar o que se entende por inteligência artificial, descrever o seu funcionamento e discorrer sobre suas características e técnicas, a partir de bibliografia especializada. O segundo capítulo consiste na apresentação das premissas teóricas que fundamentarão o trabalho, referentes ao processo de constitucionalização do Direito Civil e suas repercussões no instituto da responsabilidade civil, correlacionando-as com as especificidades da inteligência artificial. O terceiro capítulo tem por intuito aferir como o legislador brasileiro tem tratado a questão da responsabilidade civil por danos causados pela inteligência artificial, à luz das premissas teóricas estabelecidas. Por fim, no quarto capítulo, busca-se demonstrar que, no campo da responsabilidade civil, o dever de não lesar a outrem extrapola a dimensão reparatória do instituto, pelo que serão abordadas possibilidades regulatórias que internalizem os princípios constitucionais que consagram a proteção à pessoa humana, atentando-se às oportunidades e dificuldades apresentadas pelo cenário global.

Palavras-chave: Responsabilidade civil, Inteligência Artificial, regulação, Brasil.

ABSTRACT

The development of artificial intelligence, notably the machine learning technique, has increased the degree of autonomy with which computers operate, a feature that challenges the field of civil liability, insofar as it can raise doubts about the attribution of responsibility for damages caused by these machines, given the difficulty of establishing a link between human conduct and damage. In this context, the present dissertation seeks to assess the suitability of the current legal system for the protection of victims from damage caused by artificial intelligence, adopting a qualitative research under the deductive method, based on the bibliographic survey and the analysis of legislative proposals. The first chapter is dedicated to the presentation of the research object, in order to specify what is meant by artificial intelligence, describe its operation and discuss its characteristics and techniques, based on specialized bibliography. The second chapter consists of the presentation of the theoretical premises that will support the work, referring to the process of constitutionalization of Civil Law and its repercussions on the civil liability institute, correlating them with the specifics of artificial intelligence. The third chapter aims to assess how the Brazilian legislator has dealt with the issue of civil liability for damages caused by artificial intelligence, in the light of the established theoretical premises. Finally, in the fourth chapter, the author seeks to demonstrate that, in the field of civil liability, the duty not to harm others goes beyond the reparatory dimension of the institute, so the study addresses regulatory possibilities that internalize the constitutional principles that enshrine the protection of the person, paying attention to the opportunities and difficulties presented by the global scenario.

Keywords: Liability, artificial intelligence, regulation, Brazil.

LISTA DE FIGURAS

Figura 1 – Esquema do teste de Turing.....	18
Figura 2 - Fluxograma de representação de um algoritmo.....	26
Figura 3 - Esquema de aprendizado supervisionado.....	32
Figura 4 - Esquema de aprendizado não supervisionado.....	32
Figura 5 - Esquema de aprendizado por reforço.....	33
Figura 6 - Representação de uma rede neural.....	34

SUMÁRIO

INTRODUÇÃO.....	10
1 INTELIGÊNCIA ARTIFICIAL E RESPONSABILIDADE CIVIL: CONTEXTUALIZANDO O PROBLEMA DE PESQUISA.....	15
1.1. Registros dentro de um contexto histórico.....	15
1.2. Definições de inteligência artificial.....	19
1.3. Inteligência artificial específica, geral e a singularidade.....	22
1.4. Algoritmos e o funcionamento da inteligência artificial.....	25
1.5. <i>Machine learning</i>	28
1.5.1. <i>O dataset</i>	29
1.5.2. <i>Formas de aprendizado de máquina</i>	31
1.5.3. <i>Deep learning e redes neurais</i>	34
1.5.4. <i>A “black box” da inteligência artificial</i>	35
1.6. Chegando ao problema de pesquisa: a dificuldade de se atribuir responsabilidade por dano causado pelo uso da inteligência artificial.....	37
2 A RESPONSABILIDADE CIVIL NA LEGALIDADE CONSTITUCIONAL E A TUTELA DA VÍTIMA DE DANOS CAUSADOS PELA INTELIGÊNCIA ARTIFICIAL.....	42
2.1. A perspectiva Civil-Constitucional.....	42
2.2. Da responsabilidade civil ao Direito de Danos: superando as barreiras para a tutela da vítima.....	52
2.2.1. <i>A tutela da vítima de danos</i>	52
2.2.2. <i>A mitigação dos filtros para a reparação</i>	59
2.2.2.1. <i>A culpa e o risco no regime de responsabilidade</i>	60
2.2.2.2. <i>O nexo causal</i>	68
2.3.3.3. <i>O dano</i>	75
3 AS PROPOSTAS DE REGULAÇÃO DA INTELIGÊNCIA ARTIFICIAL NO BRASIL: UMA ANÁLISE SOB O PARADIGMA SOLIDARISTA.....	79
3.1. Projeto de Lei nº 21/2020 da Câmara dos Deputados.....	81
3.2. Projeto de Lei nº 5.051/2019 do Senado Federal.....	82
3.3. Análise das proposições à luz do paradigma solidarista.....	83
4 O DEVER DE NÃO LESAR A OUTREM: SOLIDARIZAÇÃO DOS RISCOS E PREVENÇÃO DE DANOS.....	91

4.1. O princípio da precaução como deliberação pública: lições das leis de proteção de dados para o campo da inteligência artificial.....	95
4.2. Regulação <i>ex ante</i> da inteligência artificial: dificuldades e oportunidades.....	102
CONCLUSÃO.....	116
REFERÊNCIAS BIBLIOGRÁFICAS.....	119

INTRODUÇÃO

O desenvolvimento da tecnologia na área de inteligência artificial¹, notadamente, no que se refere à técnica do *machine learning*, tem proporcionado o surgimento de sistemas computacionais que aprendem e replicam comportamentos a partir de experiências. Cada vez mais, máquinas dotadas de inteligência distanciam-se da necessidade da supervisão humana e mostram-se capazes de interagir com o ambiente e emitir reações independentemente de programação específica para tanto.

Ao passo em que o nível de autonomia da inteligência artificial aumenta, também aumentam os riscos de que o comportamento das máquinas se torne imprevisível, podendo vir a causar danos aos seus usuários ou a terceiros sem que se possa constatar algum problema mecânico ou de programação que explique tal incidente. Neste caso, suscitam-se dúvidas sobre a atribuição de responsabilidade pelo ocorrido, já que a relação de causalidade entre a programação do *software* e o funcionamento da inteligência artificial se distancia ao longo do processo de aprendizagem da máquina.

A partir desse cenário, o presente trabalho tratará da questão referente à responsabilidade por danos causados por uma inteligência artificial, de forma a responder a seguinte indagação: o ordenamento jurídico brasileiro está apto a proteger direitos fundamentais das vítimas de danos causados por inteligências artificiais com capacidade de aprendizagem (*machine learning*)?

De forma mais específica, buscar-se-á 1) especificar o que se entende por inteligência artificial, descrever o seu funcionamento e apresentar suas características e técnicas; 2) apresentar as premissas teóricas que permearão o trabalho, fundamentadas no processo de constitucionalização do Direito Civil; 3) aferir como o legislador brasileiro propõe tratar a questão da responsabilidade civil por danos causados pela inteligência artificial, à luz das premissas teóricas estabelecidas; 4) extrapolar a função reparatória do instituto da responsabilidade civil para aventar possibilidades regulatórias que internalizem os princípios constitucionais que consagram a proteção à pessoa humana, atentando-se às oportunidades e dificuldades apresentadas pelo cenário global.

¹ Ao longo do trabalho, o termo poderá ser suscitado pela sigla IA.

Tratar-se-á de uma pesquisa qualitativa sob o método dedutivo, baseada, fundamentalmente, no levantamento bibliográfico e na análise de proposições legislativas.

Para tanto, o caminho a ser percorrido partirá da descrição do objeto da pesquisa – inteligência artificial – tendo em vista tratar-se de tema que traz consigo especificidades técnicas nem sempre conhecidas aos operadores do Direito. Assim, o primeiro capítulo consistirá, de início, em um registro histórico da matéria, seguido da apresentação das possíveis definições de inteligência artificial. Menciona-se definições, no plural, considerando não haver um conceito preponderante do termo, tratando-se, em verdade, de uma área do conhecimento que busca a solução do problema que consiste na criação de máquinas inteligentes, o que atrai abordagens de disciplinas e perspectivas diversas.

Em seguida, o esforço se voltará à descrição do estágio tecnológico em que nos encontramos atualmente. O intuito é um só: afastar o objeto de estudo de ideias cinematográficas que sugerem a existência da máquinas superinteligentes que desafiam o controle humano e tornam-se um perigo para a perpetuação da humanidade. Embora estejamos a tratar de uma tecnologia com potencial disruptivo, a IA que conhecemos até o momento presta-se à execução de tarefas determinadas e depende de recursos finitos, razão pela qual não estamos a cogitar da equiparação da máquina dotada de inteligência às amplas capacidades humanas consideradas em conjunto.

Subsequentemente, buscar-se-á explicar o funcionamento da IA e, de forma mais específica, da técnica do *machine learning*, chegando à característica que torna essa tecnologia um desafio ao Direito, notadamente, ao campo da responsabilidade civil, qual seja, a capacidade de se aperfeiçoar a partir da experiência, sem precisar ser programada para todas as tarefas que serão executadas, ou seja, de funcionar com autonomia. Esse é um pressuposto do problema de pesquisa que se propõe estudar, já que a atribuição de responsabilidade por eventual dano causado por uma máquina autônoma poderá esbarrar na ausência de um liame de causalidade entre uma ação humana e o prejuízo experimentado pela vítima.

Assim, a primeira fase da pesquisa consistirá em um estudo exploratório através do levantamento bibliográfico de produções científicas nacionais e estrangeiras na área de ciência da computação, de forma a acessar os conceitos e

institutos relevantes para a dissertação, bem como para compreender a forma de funcionamento da inteligência artificial.

Delineia-se, assim, o problema de pesquisa e apresenta-se a hipótese: a doutrina civilista, fundamentada no processo de constitucionalização do Direito Civil e na consagração dos princípios constitucionais da dignidade da pessoa humana e da solidariedade social, fornece mecanismos e institutos aptos a tutelar a vítima de danos causados pela inteligência artificial.

Esse é o mote do segundo capítulo da dissertação, que, mediante levantamento bibliográfico do tema, apresentará as premissas teóricas que fundamentarão o estudo, descrevendo, inicialmente, o processo pelo qual o Direito Civil viu seus institutos, anteriormente exauridos pelo código, serem absorvidos pelo texto constitucional e, ao mesmo tempo, tornou-se poroso aos princípios e valores que irradiam da Lei Maior, passando a integrar de forma harmônica um ordenamento jurídico dotado de unidade e completude.

Em matéria de responsabilidade civil, esse processo volta-se à tutela da vítima de danos da sociedade industrializada, cuja dinâmica posiciona a ocorrência de acidentes no âmbito do esperado e orienta a mitigação dos elementos configuradores do dever de indenizar, enquanto manifestação da solidariedade social. Firma-se o paradigma solidarista, pelo qual se viabiliza a reparação do injusto, sem, todavia, transferi-lo a um único indivíduo responsável, buscando-se instrumentos que fomentem uma cultura de prevenção de riscos, em atenção ao dever de não lesar a outrem que fundamenta a responsabilidade civil extracontratual.

Deixa-se, assim, de priorizar a punição do agente causador do dano para garantir a sua plena reparação, tornando a responsabilidade objetiva ao se suscitar a noção de risco da atividade. O nexo causal, por sua vez, deixa de ser investigado como mero elemento de fato, de forma a permitir a sua aferição ainda que impossibilitada a produção de prova da existência de liame entre uma conduta e o dano. Por fim, o dano passa a ser o elemento delimitador dos interesses merecedores de tutela, que serão assim identificados à luz dos valores e princípios constitucionais.

Sob essas premissas, chega-se à constatação de que, embora as atividades que façam uso de inteligência artificial não possam, *a priori* e de forma generalizada, ser consideradas como atividades de risco, a teoria do risco pode ser adotada como um fundamento de imputação para viabilizar a reparação da vítima quando a prova da

culpa do agente causador do dano se mostrar dificultada. Ademais, quanto à demonstração do liame entre conduta e dano, suscita-se a teoria da presunção de causalidade enquanto manifestação da solidarização dos riscos, em hipóteses nas quais a IA seja associada a uma atividade arriscada ou quando os danos sejam caracterizados pela multiplicidade de causas, pela variedade e indeterminação das vítimas e pela violação a interesses que se disseminam no tempo e no espaço.

Embora os interesses merecedores de tutela em face da inteligência artificial não sejam propriamente novos - privacidade, igualdade, liberdade, autodeterminação, incolumidade física e psicológica, defesa do consumidor, para citar alguns exemplos -, o olhar voltado para a vítima do dano permite constatar uma vulnerabilidade potencializada pela disparidade entre o indivíduo comum e as grandes empresas de tecnologia, o que impõe a viabilização da reparação sempre que um direito individual ou coletivo se mostrar violado.

Feitas tais constatações, o terceiro capítulo buscará analisar, à luz do paradigma descrito no capítulo anterior, proposições legislativas cujo intuito é regular a inteligência artificial em matéria de responsabilidade civil de forma específica. De pronto, se verifica que o intuito de estipular um regime de responsabilidade próprio da IA vai ao encontro da visão sistemática do ordenamento jurídico exposta no capítulo anterior.

Como dito, porém, o paradigma solidarista não se encerra na viabilização da reparação à vítima, sendo certo que proteger direitos fundamentais significa, também, evitar a ocorrência dos danos aos quais os indivíduos estão sujeitos, sendo esta a ideia central do quarto e último capítulo. Assim, a constatação da desnecessidade de se estipular novas regras para que a vítima de danos causados pela inteligência artificial possa ser reparada não significa dizer que, eventualmente, não precisaremos de uma regulação que induza os integrantes da indústria de tecnologia a comportamentos voltados à prevenção de danos e gestão de riscos, o que demandará um debate que permita a participação de todos os atores envolvidos com a tecnologia e impactados por ela e que estipule obrigações de transparência e prestação de contas.

Uma regulação *ex ante* da inteligência artificial, porém, esbarra em dificuldades associadas às características da tecnologia – ante a sua complexidade e a rapidez com que se desenvolve -, bem como dos atores envolvidos em seu desenvolvimento

– ante a predominância do setor privado em face do setor público, revelando uma assimetria de acesso a expertise e recursos. Dessa forma, o trecho final do trabalho busca descrever tais dificuldades sob uma perspectiva global, de forma a registrar o cenário em que o debate regulatório brasileiro estará inserido.

Mediante a estrutura proposta, espera-se aferir se as vítimas de danos causados pela inteligência artificial estão atualmente contempladas pelo ordenamento jurídico e, mais especificamente, se há algum esforço regulatório a ser implementado de forma a garantir a proteção a direitos fundamentais dos indivíduos em face da tecnologia.

1 - INTELIGÊNCIA ARTIFICIAL E RESPONSABILIDADE CIVIL: CONTEXTUALIZANDO O PROBLEMA DE PESQUISA

Embora o cerne do presente trabalho seja jurídico, há questões técnicas a serem endereçadas antes de se abordar o problema da pesquisa, que versa sobre a responsabilidade civil por danos causados pela inteligência artificial. Tratando-se de tema que tem por objeto uma tecnologia, em grande medida, desconhecida ao campo do Direito, será necessário, de início, apresentar os conceitos e noções pertinentes, de forma a contextualizar o que será abordado nos capítulos subsequentes.

Assim, este capítulo consistirá na apresentação de um breve contexto histórico do desenvolvimento da inteligência artificial, para, em seguida, abordar as definições possíveis do termo e pormenorizar em que estágio de avanço tecnológico nos encontramos. Ainda, será descrito o funcionamento da IA e, mais especificamente, da técnica do *machine learning*, relacionando, por fim, as características apresentadas ao problema de pesquisa proposto, qual seja, a dificuldade de se atribuir responsabilidade por danos causados pela inteligência artificial.

1.1. Registros dentro de um contexto histórico

Muito antes de concebermos a inteligência artificial que conhecemos hoje, a noção de máquinas inteligentes foi aventada por filósofos que buscavam definir o que significava ser humano (BUCHANAN, 2005). No século XVII, Descartes, considerado como um dos primeiros filósofos da tecnologia dos tempos modernos, sustentava que máquinas automáticas jamais poderiam substituir inteiramente o humano, pois, ainda que um robô pudesse executar as mesmas operações intelectuais de seu criador, não haveria como mecanizar a consciência (TEIXEIRA, 2018).

Embora seja possível retroagir à Antiguidade para se extrair possíveis contribuições ao campo dos autômatos, ou seja, de máquinas capazes de funcionar por algum tempo sem intervenção humana (SILVA, 2020), foi no século XX que surgiram as ideias mais influentes do campo da inteligência artificial, a partir do

trabalho do matemático britânico Alan Turing, que investigou a possibilidade de máquinas apresentarem comportamento inteligente².

Em artigo publicado em 1948, intitulado *Intelligent Machinery*, Turing observou que haveria um entendimento geral no sentido da impossibilidade de existirem máquinas inteligentes, indicando algumas razões para tanto, como (i) uma resistência para admitir a concepção de um rival ao ser humano em poder intelectual; (ii) uma crença religiosa de que a tentativa de construir tais máquinas equivaleria à audácia de Prometeu³; (iii) o caráter limitado do maquinário utilizado à época; (iv) o fato de que qualquer erro matemático cometido pela máquina levaria ao descrédito do seu uso e (v) a ideia de que a inteligência da máquina seria um mero reflexo da inteligência de seu criador (TURING, 2004).

Turing refutou cada um desses argumentos, afirmando, quanto aos dois primeiros, tratar-se de objeções meramente emocionais, embora tenha reconhecido que não deveriam ser ignoradas, por considerar que a ideia de inteligência seria mais emocional do que matemática. Turing explica que nosso estado de espírito é determinante para considerarmos uma máquina como inteligente ou não. É que, se conseguirmos explicar e prever o comportamento da máquina, tenderemos a rejeitar a ideia de inteligência. Por outro lado, a mesma máquina pode ser considerada como inteligente apenas pelo fato de ignorarmos as regras que condicionam seu comportamento.

Quanto à terceira, o autor registrou a existência, à época, de maquinário apto a fazer operações com grande quantidade de números, sem repetição e por longo período. A quarta objeção, por sua vez, foi rejeitada por Turing sob o argumento de que inteligência não significaria ausência de erros, comparando a máquina a uma criança que, na hipótese de ser orientada de forma equivocada por um professor de aritmética, chegaria também a resultados errados, não podendo, por isso, ser considerada como um ser não inteligente. Por fim, a quinta objeção não se sustentaria, segundo Turing, pois a inteligência de uma máquina não seria mero reflexo da

² O termo “inteligência artificial” é creditado ao matemático John McCarthy que, em 1956, em conjunto com os pesquisadores Marvin Minsky, Claude Shannon e Nathaniel Rochester, organizou o workshop *Dartmouth Summer Research Project on Artificial Intelligence* (MOOR, 2006).

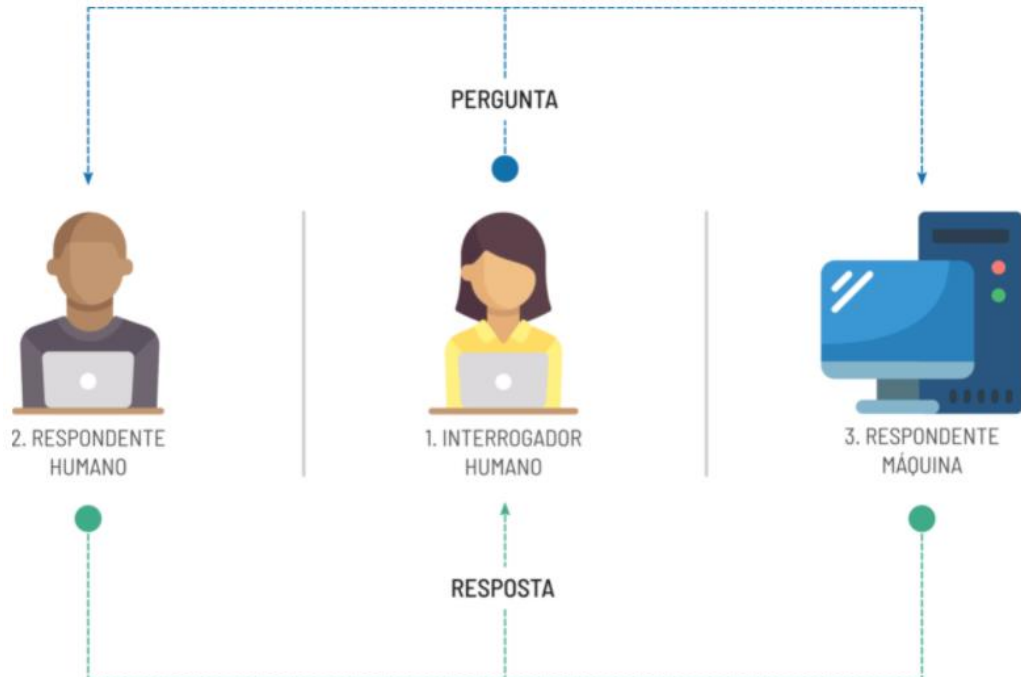
³ Herói da mitologia grega, que cunhou a humanidade do barro e, ao apresentar-lhe o fogo, despertou a ira de Zeus, que vingou-se acorrentando Prometeu no alto do monte Cáucaso, para ter suas entranhas bichadas por uma águia todas as manhãs e regeneradas sucessivamente (BOTELHO, 2019).

inteligência de seu criador, assim como a descoberta feita por um pupilo não poderia necessariamente ser atribuída a seu professor.

Interessante notar que, em diversas passagens, o autor compara a máquina a uma criança em aprendizagem, cujo cérebro encontra-se desorganizado e que, mediante treinamento intelectual e estímulos sensoriais, vem a amadurecer e a tornar-se uma “máquina organizada”, o que condiz com algumas das características observadas na inteligência artificial atualmente existente, como se verá adiante. O intuito dessa tal “máquina organizada” seria emular um matemático humano, que faz os seus cálculos em uma folha de papel, de acordo com uma série de instruções. A generalização dessa ideia equivaleria a uma “máquina universal”, capaz de exercer as funções de qualquer máquina computacional (MUGGLETON, 2014).

Em célebre artigo intitulado *Computing Machinery and Intelligence*, Turing (2009) se propôs a verificar se máquinas poderiam pensar. Para tanto, elaborou o famoso “Jogo da Imitação”, também conhecido como o Teste de Turing, que consistia na elaboração de perguntas pelo jogador denominado interrogador, a serem direcionadas a uma pessoa e a um computador, no intuito de descobrir qual deles seria o humano e qual seria a máquina. O interrogador estaria impedido de ver, tocar ou ouvir os demais, sendo que as respostas seriam fornecidas por escrito ou por um intermediário. Assim, o objetivo do jogo não seria aferir a capacidade da máquina de emular a fisionomia ou a voz de seres humanos, mas de efetivamente imitar a forma humana de pensar.

Figura 1: Esquema do teste de Turing



Fonte: Grupo DataAt, 2020. Disponível em <https://dataat.github.io/introducao-ao-machine-learning/introdu%C3%A7%C3%A3o.html>. Acesso em 15 mar. 2020

Turing vislumbrava que computadores digitais poderiam ser construídos de forma a emular as ações de um “computador humano” – desde que tivessem capacidade de armazenamento e de processamento suficientes -, seguindo instruções previamente definidas para determinada atividade, as quais seriam expressas em forma de códigos representados por números. Aqui, também, o autor avança a ideia de que, em vez de tentar construir uma máquina para imitar a mente adulta, seria mais fácil produzir uma simulação da mente de uma criança, que seria algo como uma folha em branco a ser preenchida pela programação e pela aprendizagem, mediante um sistema de punições e recompensas transformadas em símbolos reconhecíveis pelo computador.

Apesar de ainda não haver notícias de que uma máquina tenha logrado êxito no Jogo da Imitação, Alan Turing apresentou peças-chave para a pesquisa em inteligência artificial que o sucedeu. A máquina vislumbrada pelo autor deu origem à indústria computacional que conhecemos hoje, da qual se extraem *softwares* capazes de fazer pesquisas em grandes volumes de informação, identificar objetos em imagens, aprender padrões de comportamento dos usuários, traduzir textos e atuar como assistentes digitais (MUGGLETON, 2014).

O desenvolvimento da IA, desde os anos quarenta, contou com ciclos de otimismo, mas também com percalços e retrocessos, os quais foram importantes para a adoção de abordagens criativas por parte dos pesquisadores e para o refinamento de conhecimentos adquiridos ao longo do tempo (NORVIG; RUSSELL, 2013).

A partir dos anos 1990, o campo experimentou avanços significativos, observando-se aumentos na produção científica, na oferta de empregos, na abertura de empresas e no investimento financeiro no setor, o que pode ser atribuído à evolução metodológica desse campo de conhecimento, que levou à elaboração de modelos matemáticos que possibilitaram a emulação de tarefas essencialmente humanas, como o reconhecimento de imagens, voz e textos e a correlação de informações visando a tomada de decisões. Além disso, a diminuição dos custos computacionais e o aumento da capacidade de processamento e da disponibilidade de dados são fatores que impulsionaram o desenvolvimento do setor (SILVA, 2020).

1.2. Definições de inteligência artificial

Embora não se trate de uma ideia propriamente nova, não há ainda um conceito universal de IA, o que proporciona abordagens flexíveis e amplas acerca da matéria (STONE et al, 2016). Nesse sentido, afirma Russell (2016) que a inteligência artificial é o estudo do problema geral que consiste na criação de máquinas inteligentes, não se confundindo com um produto técnico específico resultante de uma pesquisa ou outra. Não se trata, portanto, de uma única abordagem para a resolução do problema; ela é o próprio problema, que atrai diferentes perspectivas visando sua solução (RUSSELL, 2016).

Para tanto, diversas disciplinas forneceram contribuições ao campo da inteligência artificial. Por exemplo, a filosofia tornou a inteligência artificial concebível, pois admitiu que a mente humana funciona, por vezes, como uma máquina, operando com conhecimento codificado em uma linguagem própria, formando um pensamento que leva a uma ação. A matemática forneceu as ferramentas para a elaboração de afirmações lógicas e envolvendo probabilidade, bem como para a compreensão dos algoritmos, enquanto a economia lida com a tomada de decisões que busquem o

melhor resultado possível. A neurociência buscou compreender o funcionamento do cérebro, detectando o que o aproxima ou o distancia de um computador. A psicologia, por sua vez, contribuiu com a adoção da ideia de que humanos e animais podem ser considerados máquinas de processamento de informações. Já os engenheiros de computação desenvolveram as máquinas capazes de lidar com a quantidade de informações que a inteligência artificial demanda. Por fim, os especialistas em cibernética elaboraram mecanismos que respondem aos estímulos do ambiente (NORVIG; RUSSEL, 2013).

Stuart Russel e Peter Norvig (2013) consideram que a inteligência artificial é um campo de estudos ainda aberto para novas descobertas, descrevendo-o como universal, por ser relevante para qualquer atividade intelectual, desde as mais amplas (percepção e aprendizagem), até as mais específicas (jogar xadrez, provar teoremas matemáticos, escrever poesia, dirigir um carro e diagnosticar doenças).

Os autores não apresentam uma única definição de IA, mas organizam possíveis definições em quatro categorias observadas em diferentes métodos utilizados pelos pesquisadores, conforme a intenção seja de fazer uma máquina: a) agir como um humano; b) pensar como um humano; c) pensar racionalmente; ou d) agir racionalmente. Uma abordagem focada no ser humano aproximar-se-ia de uma ciência empírica, envolvendo observações e hipóteses sobre o comportamento e a forma de pensar das pessoas. Por outro lado, uma abordagem racionalista combinaria a matemática e a engenharia para estipular um modelo ideal de inteligência e que não necessariamente poderia ser observado em um ser humano real.

No que se refere à criação de uma máquina que aja como um ser humano, Russel e Norvig (2013) suscitam o Teste de Turing, descrito acima, porém reconhecem que os pesquisadores de IA pouco têm se dedicado a ele, pois considerariam mais importante compreender os princípios subjacentes da inteligência do que duplicar um exemplar humano em um computador. Nesse sentido, fazem uma analogia com a empreitada humana de voar: quando paramos de tentar imitar os pássaros e passamos a compreender as regras da aerodinâmica, o primeiro voo artificial foi bem-sucedido.

Para que possamos programar um computador para pensar como um ser humano, precisaríamos primeiro compreender como a mente humana funciona, o que pode ser alcançado pela auto-observação, pela observação de outras pessoas ou pela

observação do cérebro em ação. Uma vez compreendida a mente do ser humano, este conhecimento seria passível de ser transferido para um programa de computador. Para tanto, a ciência cognitiva teria um papel preponderante em unir modelos computacionais de IA e técnicas experimentais da psicologia, para construir teorias precisas e verificáveis da mente humana. Não obstante, modernamente, não se considera que uma máquina que opera de forma satisfatória na consecução de uma tarefa necessariamente equivalerá a um modelo exato do pensamento humano (NORVIG; RUSSEL, 2013).

Programar uma máquina para pensar racionalmente refere-se a uma tentativa de codificar o “pensamento correto” para determinada situação, o que é uma atividade que relaciona a inteligência artificial à lógica, a qual fornece estruturas de argumentação para, dadas certas premissas, se chegar a uma conclusão acertada. Esta abordagem depara-se com o desafio de traduzir um conhecimento informal em termos aceitáveis para uma concatenação lógica, bem como o fato de que a resolução de problemas na prática mostra-se uma tarefa mais difícil que resolvê-los na teoria (NORVIG; RUSSEL, 2013).

Por fim, o agir racionalmente refere-se à habilidade de um computador de operar de forma autônoma, perceber seu ambiente, adaptar-se e perseguir objetivos. Um agente racional é aquele que busca o melhor resultado diante das circunstâncias. O parâmetro de racionalidade é definido matematicamente e é utilizável de forma generalizada, ao contrário do comportamento humano, que é adaptado para um ambiente específico. Norvig e Russel (2013) indicam que a busca pela racionalidade perfeita – sempre chegar à conclusão correta – talvez seja inalcançável, não passando de uma hipótese para fins didáticos.

Nils John Nilsson (2009), por sua vez, reconhecendo a inexistência de uma definição unânime de inteligência artificial, elege a que considera mais adequada: a atividade voltada ao desenvolvimento de máquinas inteligentes, sendo inteligência a qualidade que permite que uma entidade funcione de forma apropriada e com capacidade de previsão de resultados⁴. O autor repara, porém, que, de acordo com tal definição, tanto humanos como animais e algumas máquinas simples poderiam ser considerados inteligentes. Assim, suscita a noção de um *continuum* em que vários

⁴ Tradução livre de “*artificial intelligence is that activity devoted to making machines intelligent, and intelligence is that quality that enables an entity to function appropriately and with foresight in its environment*” (NILSSON, 2009, p. 13).

níveis de inteligência estariam inseridos, considerando que funcionar de forma apropriada e com previsão de resultados exigiria diferentes habilidades de acordo com o ambiente no qual determinada entidade encontra-se inserida.

Por fim, consigna-se a definição de Fabiano Hartmann Peixoto (2020, p. 17), para quem a IA seria “um ramo da ciência da computação que busca, com interação multidisciplinar com outras áreas do conhecimento, a reprodução de ações cognitivas tipicamente humanas”.

1.3. Inteligência artificial específica, geral e a singularidade

Peter Stone et al (2016) indicam não existir, até o momento, inteligência equivalente à humana no mundo biológico ou artificial, o que coloca os seres humanos como um modelo a inspirar o progresso da IA. Dito isso, os autores advertem que a medida da habilidade humana é apenas uma condição suficiente, mas não significa o limite máximo para se desenvolver sistemas de inteligência artificial, tendo em vista que muitos deles excedem a capacidade das pessoas de executarem determinadas tarefas.

Nesse sentido, Fabiano Hartmann Peixoto (2020, p. 17-18) afirma que a “capacidade de integração de funções cognitivas artificiais e os limites em graus variados de complexidade são as chaves principais nos estudos de IA”. O autor destaca, porém, que, até o momento, “só há respostas e caminhos que reproduzem parcialmente a capacidade da inteligência humana”.

De fato, a IA desenvolvida até agora, embora seja utilizada em uma gama de aplicações, é destinada à execução de tarefas determinadas, com atuação limitada por uma série de parâmetros predefinidos e voltados à resolução de problemas também determinados, razão pela qual é considerada uma inteligência artificial específica, em contraposição à inteligência artificial geral, que se equipararia ao amplo conjunto de atividades praticadas por um ser humano, mas que ainda não é uma realidade (HARTMANN PEIXOTO; SILVA, 2019).

Assim, a IA que conhecemos hoje pode ser aplicada em funcionalidades como carros autônomos, reconhecimento de voz, *game playing*, detector de *spam*,

planejamento de logística, robótica, tradução de idiomas etc., o que são feitos extraordinários, mas com escopo bastante delimitado (NORVIG; RUSSELL, 2013).

O futuro da IA ainda não é uma unanimidade entre os pesquisadores, havendo aqueles que, como Stephen Hawking, temem o desenvolvimento de uma superinteligência que venha a ser uma ameaça existencial à humanidade ao perseguir objetivos que sejam incompatíveis com a perpetuação da nossa espécie (STANKOVIC et al, 2017).

Impulsionada pela ficção científica, a noção de singularidade tecnológica tem habitado o debate público. O termo deriva das ciências exatas, para as quais a singularidade diz respeito ao ponto no espaço-tempo em que as regras da matemática e da física encontram um limite de aplicabilidade, diante da ausência de previsibilidade do fenômeno, a exemplo do que ocorre com o horizonte de eventos dos buracos negros. No que se refere ao progresso da tecnologia, trata-se da hipótese de que os avanços tecnológicos levariam a uma guinada inevitável no curso da história humana, mediante a qual paradigmas sociais, políticos e até biológicos sofreriam uma quebra com passado e presente, encaminhando-se para um futuro desconhecido (STEPHENS, 2015). Quando aplicado à inteligência artificial, o termo diria respeito à ideia de máquinas inteligentes capazes de criar máquinas ainda mais inteligentes, o que causaria um aumento exponencial na capacidade cognitiva e intelectual de tais computadores, levando a uma inteligência infinita (HAWKINS, 2008).

A inevitabilidade da singularidade parece ser reforçada, para alguns autores, pela Lei de Moore⁵, elaborada por Gordon Earle Moore, co-fundador da Intel, que previu, em 1975, que a quantidade de transistores que poderiam ser encaixados em um *chip* de computador dobraria a cada 18 meses, com o mesmo custo de fabricação. A partir de então, as fabricantes de processadores passaram a definir suas metas com base na previsão de Moore, o que contribuiu para o surgimento de produtos menores, que consomem menos energia e que contam com mais estabilidade e melhor desempenho, em um ritmo de inovação acelerado (ALENCAR, 2015).

Embora a Lei de Moore tenha contribuído para o desenvolvimento de computadores mais rápidos e potentes e, conseqüentemente, para o avanço da IA, é de se destacar que a inteligência não é sinônimo de velocidade e capacidade de

⁵ Embora o próprio criador da Lei de Moore seja cético quanto à ocorrência da singularidade tecnológica (MOORE, 2008).

processamento (WALSH, 2017). Nesse sentido, não há ainda motivos relevantes para vislumbrar a IA como um perigo iminente, pois não existem – e não há perspectiva de que venham a existir no futuro próximo - máquinas autossustentáveis com capacidade de perseguir objetivos ou nutrir intenções de longo prazo (HARTMANN PEIXOTO; SILVA, 2019).

Jeff Hawkins (2008), na mesma linha, sustenta que a aplicação da noção de singularidade ao desenvolvimento da inteligência artificial denotaria um entendimento equivocado do funcionamento de uma máquina inteligente. O autor aponta que o crescimento exponencial da capacidade de computadores demanda o consumo exponencial de recursos que são limitados (matéria-prima, energia, tempo), o que afastaria a percepção de que as máquinas simplesmente escapariam do controle humano. Dessa forma, ainda que uma máquina consiga criar réplicas aprimoradas de si mesma, não haveria que se falar em singularidade, no entender do autor.

Para Russel (2016), a inevitabilidade do surgimento de uma superinteligência não é um fato óbvio, pois o desenvolvimento da capacidade de processamento dos computadores não significaria, necessariamente, um incremento da inteligência das máquinas, estando mais inclinado a aceitar a hipótese de que a IA possa ter uma performance sobre-humana em algumas tarefas específicas e sub-humana em outras.

Stone et al (2016) relembram que a humanidade desenvolveu e se adaptou a novas tecnologias ao longo da sua história, pelo que apostam na previsão de que os avanços relacionados à inteligência artificial ocorrerão de forma gradual, a partir do conhecimento que acumulamos até hoje, e não de maneira abrupta ou inesperada. Por outro lado, reconhecem que o aprimoramento de técnicas, o incremento da capacidade dos computadores e a disponibilidade de grande quantidade de dados podem levar a novas aplicações disruptivas, cuja medida do sucesso será o valor a ser agregado à vida das pessoas.

A questão premente a ser endereçada pelos pesquisadores é, na verdade, a forma com que a inteligência artificial influencia o nosso cotidiano no presente, considerando que essa tecnologia já impacta de forma relevante a vida das pessoas, determinando, por exemplo, quem está apto a obter crédito de bancos e quem pode ser libertado mediante fiança. Ainda, observa-se uma transformação no mercado de trabalho pela automação de atividades que vinham sendo desempenhadas por seres humanos, além de vieses em modelos estatísticos que se mostram tendenciosos em

relação a raça, gênero ou idade, o que suscita preocupações quanto à justiça das decisões tomadas com base em algoritmos de inteligência artificial (ZHANG et al., 2020).

Nesse sentido, mostra-se essencial notar que existem desigualdades na forma de experimentar a tecnologia, sendo que a falta de controle sobre as máquinas pode ocorrer, para alguns grupos sociais, devido a alguma hipossuficiência, muito antes de uma eventual singularidade (NARUHODO, 2021).

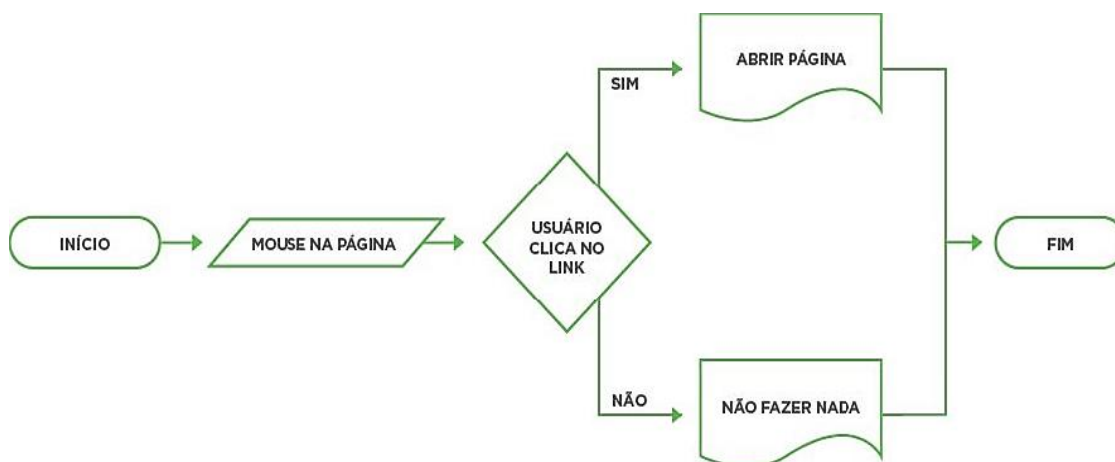
1.4. Algoritmos e o funcionamento da inteligência artificial

O funcionamento da inteligência artificial se dá a partir de “sistemas de dados programados para dar respostas conforme a base de dados disponível” (NUNES; MARQUES, 2018, p. 423). Tais sistemas são chamados de algoritmos, que, por sua vez, podem ser definidos como um conjunto de instruções que determinam a sequência de operações a serem realizadas para atingir a finalidade proposta pelo desenvolvedor (HARTMANN PEIXOTO; SILVA, 2019).

O algoritmo pode ser comparado a algo que requer uma concatenação de passos, como uma receita, um método, um procedimento, uma coreografia ou um ritual, porém, contando com as seguintes características específicas: i) deve ser finito - finalizar após um número finito de passos; ii) deve ser definido - cada passo deve ser rigorosamente e precisamente delineado, evitando ambiguidades; iii) *input* – demanda um mecanismo de inserção dos dados a serem analisados; iv) *output* – demanda um meio de saída para os dados processados; e v) efetividade – conta com instruções suficientemente básicas, de forma que possam ser executadas por uma pessoa (KNUTH, 1997; VALENTINI, 2018; HARTMANN PEIXOTO; SILVA, 2019).

Rômulo Soares Valentini (2018, p. 43) sumariza que o algoritmo “é um plano de ação pré-definido a ser seguido pelo computador, de maneira que a realização contínua de pequenas tarefas simples possibilitará a realização da tarefa solicitada sem novo dispêndio de trabalho humano”.

Figura 2: Fluxograma de representação de um algoritmo



Fonte: S.O.S Tecnologia e Educação, 2018. Disponível em <https://www.sos.com.br/noticias/tecnologia/o-que-e-um-algoritmo>. Acesso em 15 mar. 2020.

O algoritmo analisado isoladamente, todavia, pouco contribui para a compreensão do funcionamento da inteligência artificial, pois as atividades de um computador não consistem em seguir apenas uma receita, mas a combinação de várias delas, agrupadas em torno de uma tarefa, que, por sua vez, soma-se a outras tarefas para que a máquina atinja seus objetivos. Dito de forma mais técnica:

Um algoritmo não costuma ter vida própria e é útil quando encapsulado num programa que, combinando-o com outros algoritmos, permite o cumprimento de certa tarefa pela máquina. De programas fazem-se os sistemas que cabem, de uma maneira mais própria, no conceito de software. Portanto, algoritmo, programa, sistema e software denotam uma ordem ou sequência natural de organização de elementos pela qual se dá condições de utilidade a algoritmos. Em termos orgânicos e analogicamente, pode-se pensar no corpo que dá sustentação ao psíquico (TAVARES PEREIRA, 2017).

Nem todo algoritmo, porém, diz respeito à inteligência artificial, sendo importante, nesse ponto, diferenciá-la da ideia de automação. Há algoritmos que contam com regras e objetivos muito restritos e definidos, a exemplo daqueles que simplesmente executam orientações para organizar certos elementos em dada ordem, ou para realizar uma busca textual em determinado documento, não podendo, por isso, ser considerados como inteligentes (BARTNECK, 2021).

Por sua vez, a inteligência artificial não se restringe à automação, pois atua com o processamento de uma enorme quantidade de dados e *softwares* com capacidade de aprimoramento sem programação específica para cada tarefa que será desenvolvida. Assim, “[e]m vez de precisar que desenvolvedores escrevam um novo

código manualmente, a IA depende de algoritmos capazes de se tornarem ‘mais inteligentes’ ao processar mais dados do mundo real” (GOOGLE, 2017).

Christoph Bartneck et al (2021) apresentam uma descrição didática da multiplicidade de aplicações que levam uma máquina a comportar-se de forma inteligente, ao exemplificar o que acontece com um assistente virtual como a Siri, quando recebe uma pergunta do usuário. Primeiro, o microfone do celular precisa traduzir o som em uma informação digital que possa ser armazenada na memória do dispositivo. Tal informação precisa ser encaminhada via *internet* para um computador potente presente na nuvem, o qual tentará classificar o som captado em palavras escritas e extrair significado da combinação de palavras, atentando-se à existência de homônimos, gírias, ironias ou duplos sentidos. Por fim, a Siri fornecerá uma resposta ou recomendação e, para tanto, precisará entender o contexto da pergunta, o que requer conhecimento sobre o mundo, raciocínio lógico e habilidade de aprender e adaptar-se de forma autônoma.

Fabiano Hartmann Peixoto e Roberta Zumblick Martins da Silva (2019, p. 33) apontam que a simulação de partes do intelecto humano demanda que a máquina desenvolva habilidades, as quais podem ser definidas, segundo os autores, pelas seguintes perguntas:

- 1) Como reconhecer objetos?
- 2) Como converter sons em palavras e vice-versa?
- 3) Como extrair sentido da linguagem e transmitir significado por meio de sentenças geradas?
- 4) Como ordenar informações de uma forma prática?
- 5) Como combinar pedaços de informações para alcançar conclusões?
- 6) Como programar uma sequência de ações para cumprir determinado objetivo e ter certeza de que ela foi bem executada?

As habilidades correspondentes seriam, principalmente, a visão computacional (*computer vision*), o reconhecimento de fala (*speech recognition and synthesis*), o processamento de linguagem natural (*natural language processing*), a representação de conhecimento (*knowledge representation*), o raciocínio (*reasoning*) e o planejamento (*planning*). Em destaque, importa mencionar, ainda, a habilidade de melhorar a performance em cada habilidade através da experiência, ou seja, o aprendizado de máquina (*machine learning*) (HARTMANN PEIXOTO; SILVA, 2019).

Percebe-se, portanto, que a inteligência artificial não é apenas uma tecnologia, mas um conjunto de técnicas e habilidades que conferem uma aparência de

inteligência a computadores (FIRTH-BUTTERFIELD, 2017), ou, dito de outra forma, a IA é um “termo guarda-chuva” que “abriga muitas áreas de estudo e técnicas” (HARTMANN PEIXOTO; SILVA, 2019, p. 75).

1.5. *Machine learning*

Machine learning, ou aprendizado de máquina, é uma subárea do campo de estudo da inteligência artificial que se refere a sistemas que apresentam a capacidade de adquirir conhecimento, resolver problemas e tomar decisões através da detecção de padrões em um conjunto de dados, visando a previsão de dados futuros. Trata-se, portanto, de ramo da IA que estuda a capacidade da máquina de melhorar sua performance pela experiência (HARTMANN PEIXOTO; SILVA, 2019).

Nesse sentido, o *machine learning* se associa à “obtenção de um considerável volume de dados e seu processamento para a possibilidade de identificação de padrões que, também combinados, acabam por possibilitar a predição e recomendação de ações características da atividade cognitiva humana” (HARTMANN PEIXOTO, 2020, p. 18).

Segundo Tom Michael Mitchell (2006), as questões centrais estudadas por esse subcampo da inteligência artificial são as seguintes: como podemos construir sistemas computacionais que se aperfeiçoam automaticamente com a experiência e quais são as leis fundamentais aplicáveis a todos os processos de aprendizagem?

A técnica do aprendizado de máquina é o resultado da interseção entre a ciência da computação e a estatística. Assim, enquanto a ciência da computação tem como foco a programação de computadores, o *machine learning* buscará formas de fazer o computador programar a si mesmo, a partir de uma estrutura inicial e mediante a aquisição de experiência. Enquanto a estatística busca inferir conclusões a partir da análise de dados, o *machine learning* acrescenta o intuito de aferir quais sistemas computacionais e algoritmos podem ser usados para captar, armazenar, classificar, extrair e combinar tais dados e como várias pequenas tarefas podem integrar um grande sistema orquestrado de aprendizagem (MITCHELL, 2006).

Desenvolvedores de sistemas de IA vislumbram que a utilidade da técnica do *machine learning* reside no fato de que é muito mais simples treinar um sistema através de exemplos de comportamentos ou resultados desejáveis para determinado comando do que estipular e programar cada resposta possível (JORDAN; MITCHELL, 2015). Nesse sentido, um dos objetivos desse subcampo é programar computadores para que aprendam de forma autônoma, mediante a identificação de padrões nos dados analisados, a partir dos quais serão elaborados modelos capazes de prever comportamentos futuros (HARTMANN PEIXOTO; SILVA, 2019).

Por exemplo, um algoritmo de *machine learning* para o qual é dada a função de prever em quanto tempo uma pessoa consegue correr um quilômetro⁶, mediante análise de dados como altura, peso e idade. No caso, o algoritmo processaria centenas ou milhares de exemplos de pessoas com altura, peso e idade variados, de forma a verificar a distância que cada uma delas consegue correr, visando a elaboração de um modelo, que, uma vez validado, ou seja, verificada a sua capacidade de generalizar exemplos para situações futuras, estará apto a funcionar com novos dados e, de fato, fornecer previsões acerca do tempo que uma pessoa conseguiria correr determinada distância, de acordo com as características informadas (BATHAEE, 2018).

Entre outras aplicações, algoritmos de *machine learning* podem detectar padrões de consumo na *internet* para fornecer propagandas direcionadas ao interesse dos consumidores (HARTZOG, 2017); encontrar tendências nas deliberações de determinado Tribunal no intuito de prever futuros resultados em processos judiciais; descobrir informações relevantes em documentos extensos (SURDEN, 2014); identificar pessoas pelas suas características físicas (BUOLAMWINI; GEBRU, 2018).

1.5.1. O *dataset*

O *dataset* é o banco de dados do qual são extraídos os exemplos cujas características serão consideradas pela máquina como padrões passíveis de serem generalizados no futuro, podendo ser descrito também como “uma coleção de

⁶ O exemplo usado como fonte mencionava a medida de uma milha.

exemplos, que são transformados em uma coleção de características” (HARTMANN PEIXOTO; SILVA, 2019, p. 101). O processo de aprendizagem, portanto, se dá pela exposição do algoritmo de *machine learning* ao *dataset*.

É comum se presumir que algoritmos de *machine learning* aprendem melhor quanto maior for a quantidade de dados disponíveis para treinamento. Porém, embora a técnica seja relevante em um contexto de *big data*, ou seja, de volume, velocidade e variedade de dados em grande escala, trata-se, em verdade, de um fator que complexifica o processo de aprendizagem e pode torná-lo longo, custoso e potencialmente ineficaz, diante da dificuldade de se extrair generalizações de dados muito extensos e desuniformes. Erros de generalização podem ser resumidos à variância, quando a máquina em aprendizagem acaba por realizar previsões aleatórias, e aos vieses (*bias*), quando a máquina aprende a aprender de forma errada (L'HEUREUX, 2017) e apresenta resultados tendenciosos (HARTMANN PEIXOTO, 2020).

Há vieses que podem significar o aprofundamento de injustiças e vulnerabilidades de determinados grupos sociais, pela replicação de preconceitos e tratamentos discriminatórios observados na sociedade. Hartmann Peixoto e Silva (2019), ao reconhecerem a existência de subjetividade no processo de desenvolvimento de um sistema de inteligência artificial, atentam-se à importância da fase de treinamento e do tratamento do *dataset* para que um algoritmo de *machine learning* não automatize preconceitos presentes no mundo real.

Quanto ao ponto, Cláudia Toledo (2021, p. 82) observa que “qualquer resultado discriminatório ou preconceituoso apresentado por programas de IA não é nem inerente a tais programas nem é por eles espontaneamente gerado, mas decorre da combinação de dados neles inseridos, o que é obra humana”.

Não são poucos os exemplos em que o uso da inteligência artificial levou a resultados preconceituosos. Por exemplo, algoritmos de reconhecimento facial têm apresentado, de forma reiterada, comportamentos discriminatórios quando a pessoa a ser identificada é do sexo feminino ou tem a pele negra (BUOLAMWINI; GEBRU, 2018). Na mesma linha, o Instituto de Tecnologia da Geórgia observou que carros autônomos teriam maior probabilidade de atropelar pessoas negras do que pessoas brancas, pois identificariam com mais facilidade as peles claras (WILSON; HOFFMAN; MORGENSTERN, 2019).

Joy Buolamwini e Timnit Gebru (2018) advertem que algoritmos de aprendizado de máquina (*machine learning*) podem apresentar *outputs* discriminatórios quanto a gênero e raça quando são treinados com dados que não representam a diversidade presente no mundo real. Daí a relevância de que exista a possibilidade de uma prestação de contas por parte dos desenvolvedores, de forma que seja possível auditar os dados utilizados para treinar algoritmos e aferir a precisão e acurácia⁷ da performance dos sistemas de IA e, conseqüentemente, a confiabilidade dos resultados por eles fornecidos. Por óbvio, quanto mais complexos os algoritmos e o *dataset* disponível, mais difícil será a identificação e a correção de desvios (HARTMANN PEIXOTO; SILVA, 2019).

1.5.2. Formas de aprendizado de máquina

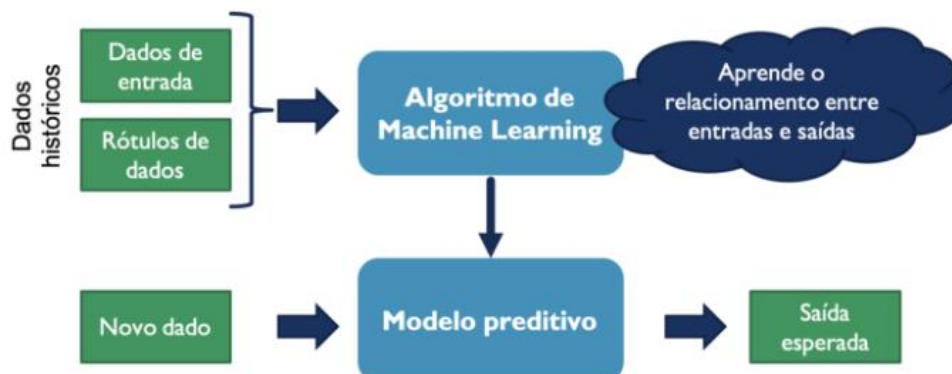
As formas de aprendizado dos algoritmos de *machine learning* referem-se à maneira pela qual eles serão expostos ao *dataset* durante o processo de treinamento, o que pode se dar de forma supervisionada e não supervisionada. Há, também, algoritmos que não experimentam um *dataset* fixo e aprendem mediante reforço (GOODFELLOW et al, 2016).

O aprendizado supervisionado se dá mediante a rotulagem prévia de elementos do *dataset*, que serão separados em categorias pelos programadores (HARTMANN PEIXOTO, 2020). Segundo Ian Goodfellow et al (2016), a técnica consiste em algoritmos que aprendem a associar um *input* a um *output*, diante de um conjunto de exemplos predefinidos. Os dados rotulados são utilizados no treinamento do algoritmo, que passará, posteriormente, por etapas de validação do desempenho da máquina, para contrastá-lo ao resultado esperado. Aplicações comuns de algoritmos dessa natureza são as tarefas de classificação - a exemplo da classificação de documentos, detecção de e-mails spam e categorização de imagens -, e de regressão, ou seja, a previsão de valores numéricos contínuos – a exemplo da previsão do valor

⁷ Goodfellow et al (2016) definem acurácia como a proporção de exemplos para os quais determinado modelo fornece o resultado (*output*) correto.

de imóveis, de custos de produção e de variações do mercado financeiro (HARTMANN PEIXOTO; SILVA, 2019).

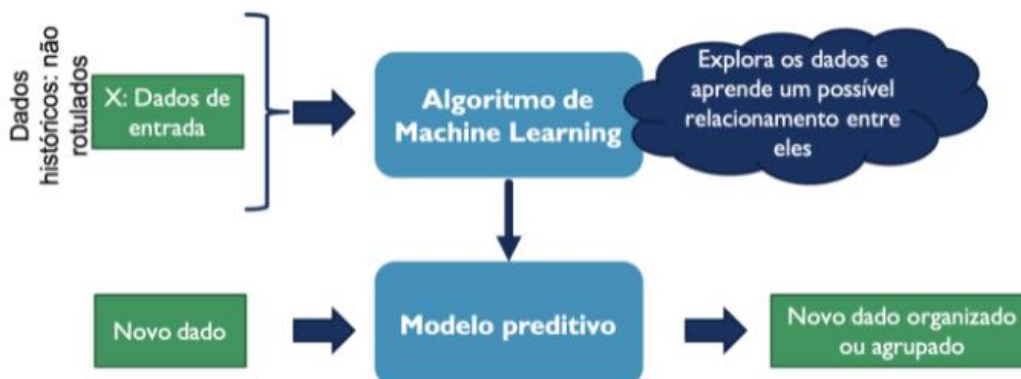
Figura 3: Esquema de aprendizado supervisionado



Fonte: Tatiana Escovedo, 2020. Disponível em <https://tatianaesc.medium.com/machine-learning-conceitos-e-modelos-f0373bf4f445>. Acesso em 15 mar. 2020

O aprendizado não supervisionado ocorre sem rotulagem prévia dos elementos do *dataset*, já que a atividade é realizada pela própria máquina, que fará o processamento, identificação e categorização dos dados (HARTMANN PEIXOTO, 2020). Para esse tipo de aprendizagem, o algoritmo deve ser exposto a um banco de dados com muitos elementos, de forma que consiga perceber as categorias que se formam dentro do *dataset* e agrupar os exemplos similares no mesmo grupo. Trata-se de técnica aplicável a problemas não tão definidos, “pois não se sabe que tipo de padrões se está procurando e não há uma métrica clara de erro para controle – ao contrário do aprendizado supervisionado, em que se sabe o tipo de padrão almejado e a métrica adotada” (HARTMANN PEIXOTO; SILVA, 2019, p. 94).

Figura 4: Esquema de aprendizado não supervisionado



Fonte: Tatiana Escovedo, 2020. Disponível em <https://tatianaesc.medium.com/machine-learning-conceitos-e-modelos-parte-ii-aprendizado-n%C3%A3o-supervisionado-fb6d83e4a520>. Acesso em 15 mar. 2020

Por sua vez, o aprendizado por reforço se dá por mecanismos de recompensa e punição, a indicarem a obtenção de resultados corretos ou equivocados pela máquina (HARTMANN PEIXOTO, 2020). Tais algoritmos interagem com o ambiente em que estão inseridos, havendo, entre a aprendizagem e a aquisição de experiência, um *feedback loop*, ou um retorno de informações sobre a performance apresentada pelo algoritmo (GOODFELLOW et al, 2016). Desse modo, o aprendiz não possui informações prévias sobre qual ação adotar, devendo aferir qual comportamento lhe confere a maior recompensa, através de uma abordagem de tentativa e erro (SUTTON, 1992). Essa forma de aprendizado de máquina tem como objetivo extrapolar as reações possíveis de determinado sistema, de forma que consiga agir corretamente nas hipóteses em que não exista correlação imediata com o conjunto de treinamento (HARTMANN PEIXOTO; SILVA, 2019).

Figura 5: Esquema de aprendizado por reforço



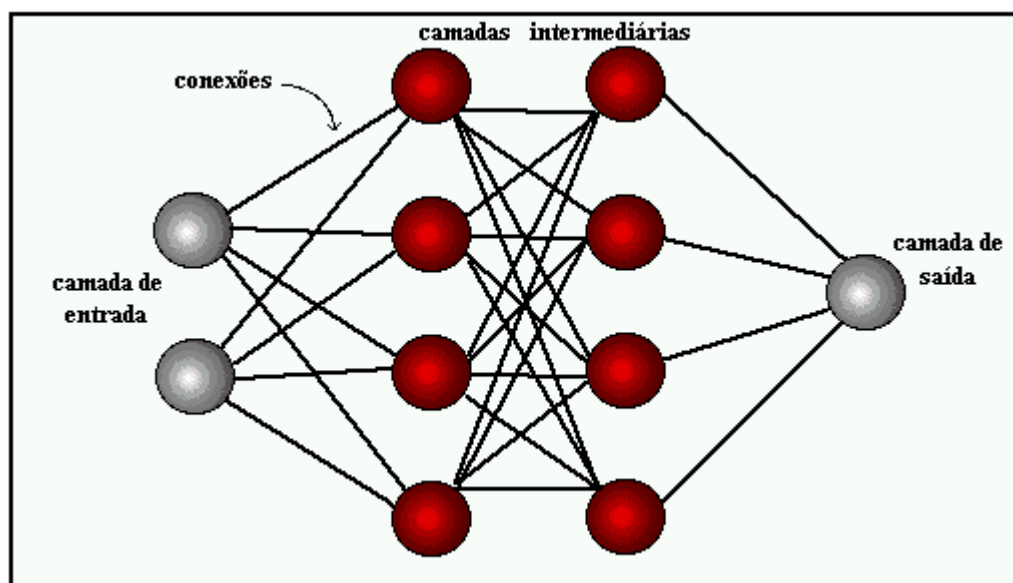
Fonte: Grupo DataAt, 2020. Disponível em <https://dataat.github.io/introducao-ao-machine-learning/introdu%C3%A7%C3%A3o.html>. Acesso em 15 mar. 2020

1.5.3. Deep learning e redes neurais

Muitos dos feitos mais extraordinários no campo do *machine learning* podem ser creditados ao *deep learning*, devido à capacidade de solucionar problemas anteriormente considerados insolucionáveis, mediante um processo de aprendizado mais complexo (ZHANG et al., 2021).

Trata-se de uma forma de aprendizado supervisionado que faz uso de um sistema computacional cuja arquitetura consiste em camadas de unidades interligadas que, por serem inspiradas nas propriedades dos neurônios, são chamadas de redes neurais. O sistema funciona por ligações que direcionam o recebimento de *inputs* e o envio de *outputs* entre as unidades dispostas em camadas, cujos pesos, quantidade e tamanhos podem ser variados (HARTMANN PEIXOTO; SILVA, 2019). Embora baseadas no cérebro humano, as redes neurais não buscam emular o neurônio biológico, mas sim, atingir a mesma habilidade de aprender e generalizar comportamentos através da experiência (BATHAEE, 2018).

Figura 6: Representação de uma rede neural.



Fonte: Instituto de Ciências Matemáticas e de Computação (ICMC – USP). Disponível em <<https://sites.icmc.usp.br/andre/research/neural/>>. Acesso em 15 mar. 2020.

Hartmann Peixoto (2020, p. 21) detalha:

Há um conjunto de entradas, que geram saídas correspondentes, variáveis de acordo com parâmetros ajustáveis. Os modelos compartilham essa estrutura de entradas, saídas e parâmetros e isso se reproduz em cada camada subsequente, formando camadas e blocos.

Assim, além do conceito de camada, tem-se também o conceito de bloco que pode auxiliar a sintetizar essa complexidade. Um bloco pode ser uma única camada, várias camadas ou um modelo inteiro, que podem inclusive ser combinados mesmo em redes neurais mais complexas, para atender a uma demanda definida. Sob a ótica da programação, um bloco é representado por uma classe e em cada uma está definido uma função de propagação, isto é, a transformação de entradas em saídas e o registro dos parâmetros.

A técnica do *machine learning* tradicional faz uso de dados para aprender a transformar *input* em *output*, como por exemplo, transformar áudio em texto por meio de reconhecimento de fala. Para tanto, é necessário que os dados utilizados estejam representados de forma que os algoritmos consigam desempenhar o comando que lhes é indicado. Por outro lado, o *deep learning* performa mediante modelos de aprendizado dispostos em vários níveis de transformação da informação de entrada em saída. As camadas mais próximas do *input* podem representar um nível baixo de detalhamento dos dados em escrutínio, sendo que, conforme são acionadas as camadas mais próximas do *output*, o sistema aprende progressivamente a representar conceitos mais abstratos e com maior nível de detalhamento. Esta característica faz com que o *deep learning* apresente resultados melhores ao processar dados que não tenham sido previamente tratados, em comparação às técnicas tradicionais de aprendizado de máquina (ZHANG et al., 2021).

1.5.4. A “black box” da inteligência artificial

Se, por um lado, o *deep learning* fornece modelos computacionais que apresentam maior acurácia, por outro, tais modelos têm revelado certa opacidade, gerando resultados pouco explicáveis e de difícil interpretação (ZHANG et al., 2021; HARTMANN PEIXOTO; SILVA, 2019). Trata-se do problema chamado de *black box*, ou caixa preta da inteligência artificial, que consiste na dificuldade de compreender completamente o processo de tomada de decisão da máquina (BATHAEE, 2018).

Algoritmos de *machine learning* tornaram-se capazes de aprender a partir de uma quantidade massiva de dados que, uma vez internalizados, fazem com que

computadores consigam tomar decisões a partir da experiência, assim como o fazem os seres humanos. Isso significa que as máquinas não estão meramente executando instruções predefinidas, pois conseguem apresentar soluções para problemas com base em padrões nem sempre perceptíveis para os programadores. Assim sendo, torna-se complicado compreender a forma com que uma inteligência artificial chega a determinadas decisões ou previsões, considerando a natureza intrincada de suas operações (BATHAEE, 2018).

A complexidade das redes neurais é um fator que dificulta a interpretabilidade do modelo. A lógica dessa forma de aprendizado profundo é a conectividade, segundo a qual uma grande quantidade de unidades computacionais simples consegue executar tarefas sofisticadas. Dessa forma, as milhares de unidades dispostas em camadas trabalham em conjunto para chegar a um resultado, não havendo uma divisão clara de tarefas ou fases determinadas para cada uma delas (BATHAEE, 2018).

Embora uma camada ou bloco de unidades possa conter o código para determinado conjunto de características a serem extraídas do *input*, o processo nem sempre será inteligível para seres humanos, ou seja, pode ser impossível determinar qual informação foi determinante para o resultado apresentado. Além disso, o aprendizado por experiência remete a um processo de tentativa e erro que, de certa forma, é intuitivo e não pode ser exatamente descrito em um conjunto de instruções predefinidas ou analisado de acordo com os passos adotados para se chegar a um resultado (BATHAEE, 2018).

Segundo Yavar Bathaee (2018), há gradações de opacidade no processo de tomada de decisão por inteligências artificiais, havendo aquelas que fornecem certo grau de transparência, pois permitem a realização de engenharia reversa para que se compreenda as variáveis consideradas pela máquina para chegar ao *output*, o que pode significar, ainda que de forma imprecisa, a possibilidade de prever quais resultados serão gerados pelo modelo computacional. Por outro lado, há modelos em que não é possível determinar o processo que levou a determinada decisão ou previsão, qual informação foi determinante para tanto, tampouco elaborar uma ordem das variáveis processadas conforme a sua preponderância para o resultado (BATHAEE, 2018).

Quanto ao tema, Hartmann Peixoto (2020, p. 29) observa:

O resultado consistente de um sistema de IA segue um fluxo de inserção de dados proveniente de um *dataset*, a internalização algorítmica e o resultado entregue. Embora exista a característica da caixa preta algorítmica, os riscos de desvios estão fortemente associados à deficiência na curadoria do *dataset* (alimentado com dados desviados e outras falhas) e pela falta de sistemas de controle e transparência no resultado, que possam detectar erros e apontar para soluções.

Registre-se que, por transparência no âmbito da inteligência artificial, entende-se não apenas a divulgação de códigos fonte ou a disponibilização de dados, mas, principalmente, “a explanação dos critérios de raciocínio da máquina de modo adequado” (PEREIRA, 2021, p. 165).

Como se verá adiante, as características aqui apresentadas são um desafio para o campo do direito, notadamente, no que se refere à atribuição de responsabilidade para o caso de danos causados por uma inteligência artificial.

1.6. Chegando ao problema de pesquisa: a dificuldade de se atribuir responsabilidade por dano causado pelo uso da inteligência artificial

O caráter interdisciplinar da presente pesquisa demandou que, de início, se apresentasse uma revisão bibliográfica acerca do tema da inteligência artificial, de maneira a fornecer uma contextualização do que será tratado no restante do trabalho.

A inteligência artificial tem ampla aplicabilidade e não são poucos os casos em que a tecnologia demonstra o potencial de promover resultados benéficos para a sociedade, como se vê, por exemplo, no aprimoramento da infraestrutura urbana mediante sistemas que otimizam o tráfego de transportes nas cidades, no auxílio à proteção ao meio ambiente através da coleta e do processamento de dados relevantes para os ecossistemas e na atuação como suporte para o diagnóstico médico (HAGER et al, 2019).

Gustavo Tepedino e Rodrigo da Guia Silva (2020, p. 297-299) reforçam a difusão da inteligência artificial nas mais diversas áreas:

No campo dos transportes, podem-se mencionar os meios de transporte autônomos, de que constituem exemplos emblemáticos os veículos (terrestres, marítimos ou aéreos) autônomos e os drones. No campo médico-

farmacêutico, basta pensar nos robôs de assistência pessoal (por exemplo, para a prestação de cuidados a idosos ou para o auxílio em tratamentos ortopédicos ou fisioterapêuticos), nos robôs-médicos, nos tratamentos da neuroengenharia (como na utilização de microchips cerebrais para tratamento da epilepsia e do Mal de Parkinson), nos procedimentos autônomos de prescrição e confecção de medicamentos, e nas intervenções que promovem reparação ou aperfeiçoamento humanos (neste ponto se inserem os casos das próteses biônicas e das intervenções no corpo humano associadas às alunhas pós-humano ou homem-ciborgue).

No campo militar, podem-se mencionar as armas autônomas e os robôs-soldados (capazes de decisões autônomas sobre o uso de armas letais). No campo financeiro, reconhecem-se os robôs com atuação autônoma em investimentos no mercado de ações e no mercado de câmbio. No ambiente estritamente virtual, destacam-se os provedores de aplicação de redes sociais, as plataformas de busca e de compartilhamento de dados. Mencione-se, ainda, o fenômeno denominado Internet of Things (Internet das Coisas), campo particularmente propício para a difusão dos sistemas de inteligência artificial.

Como se viu, estamos a tratar de algoritmos de inteligência artificial, ou, de forma mais específica, de *machine learning*, que aprendem a partir de dados e resolvem problemas de forma autônoma, ou seja, sem necessitar de programação específica para todas as atividades que realizam. Ademais, o desempenho dessa tecnologia não depende inteiramente da sua programação, pois também é influenciado pela qualidade do *dataset* ao qual serão expostos os algoritmos e pelas interações com o ambiente durante o seu funcionamento, circunstâncias que, por sua vez, são processadas pelo sistema computacional de forma pouco transparente e, por vezes, imprevisível.

A difícil interpretabilidade dos resultados apresentados por sistemas de inteligência artificial tem especial relevância quando estamos a tratar de situações em que alguém experimenta algum dano pelo uso dessa tecnologia ou quando se percebe que certos algoritmos podem impactar a vida das pessoas de forma arbitrária e discriminatória. Afinal, se não há como determinar exatamente como uma IA chegou a determinado resultado, será também difícil aferir o que ou quem causou o dano ou qual informação foi preponderante para que ocorresse a discriminação.

Carlos Edison do Rêgo Monteiro Filho e Nelson Rosenvald (2020, p. 545) indicam que, embora as novas tecnologias tenham desafiado constantemente o instituto da responsabilidade civil, a inteligência artificial acrescenta um novo patamar de preocupação, tendo em vista que

agentes não humanos estão se tornando cada vez mais autônomos em termos da complexidade das tarefas que podem realizar – com possíveis

impactos casuais não mitigáveis por seres humanos -, além da capacidade decrescente de entender, prever ou controlar como eles operam.

Na mesma linha, Caitlin Mulholland e Bianca Kremer (2020, p. 578) indagam com que fundamento e a quem seria possível atribuir a obrigação de indenizar por um dano causado pela inteligência artificial⁸, considerando que “o aprimoramento de sistemas de aprendizado de máquinas permitiu à IA uma autonomia nas tomadas de decisão que usurpa, por completo, a capacidade humana de monitorar e adaptar os caminhos da programação”.

Percebe-se que a matéria de responsabilidade civil resta desafiada pela ideia de que o comportamento de uma máquina que cause danos a alguém possa não guardar relação de causalidade com uma conduta humana, o que nos leva a questionar se devemos implementar uma regulação que contemple de forma específica a IA, de forma a evitar que as vítimas desses danos fiquem sem reparação. Embora existam, no Congresso Nacional brasileiro, projetos de lei em tramitação que visam disciplinar o uso da inteligência artificial, inclusive no que tange à responsabilidade por danos⁹, não há ainda consenso sobre a conveniência de se implementar uma legislação específica da matéria (MONTEIRO FILHO; ROSENVALD, 2020).

Para Fabiano Hartmann Peixoto e Fernanda de Carvalho Lage (2021), por enquanto, ainda não é o momento de regular a inteligência artificial¹⁰. Além de não haver uma definição única do termo, o que, de início, dificulta a tarefa legislativa, há ainda dificuldades relacionadas ao caráter transnacional da tecnologia e ao fato de que o desenvolvimento de *hardware* e *software* é feito por atores dispersos em diferentes localidades do globo, sem necessariamente estarem coordenados para a consecução de um objetivo comum. Assim, os autores apontam que, com o tempo, a ampliação do conhecimento sobre o funcionamento da IA pode oferecer consenso

⁸ No trabalho citado, as autoras focam na tutela jurídica de dados pessoais sensíveis, os quais, quando processados por sistemas de inteligência artificial que apresentam vieses, podem acarretar práticas discriminatórias, abusivas e que demandam reparação (MULHOLLAND; KREMER, 2020).

⁹ Por exemplo, o Projeto de Lei nº 5051/2019, proveniente do Senado Federal, que pretende normatizar princípios para a utilização da inteligência artificial no Brasil e estipulou, no seu texto inicial, que a “responsabilidade civil por danos decorrentes da utilização de sistemas de Inteligência Artificial será de seu supervisor” (BRASIL, 2019). Da Câmara dos Deputados, teve origem o Projeto de Lei nº 21/2020, que estabelece princípios, direitos e deveres para o uso de inteligência artificial no Brasil e dá outras providências (BRASIL, 2020). Ambos serão detalhados adiante.

¹⁰ Os autores analisam a regulação do uso da inteligência artificial no contexto do Poder Judiciário.

sobre conceitos e princípios para o seu uso, bem como indicar pistas sobre a melhor forma de regulação.

Reconhecendo tais dificuldades, este trabalho buscará aferir a aptidão do direito vigente para dirimir os conflitos que envolvem, de um lado, uma máquina dotada de inteligência artificial e, de outro, uma pessoa que tem um direito violado em decorrência do funcionamento desse agente inteligente. O que se buscará demonstrar ao longo da pesquisa é que, analisado à luz do ordenamento constitucional, o instituto da responsabilidade civil tem se mostrado capaz de absorver os danos causados por novas tecnologias e, sob essa perspectiva, o uso da inteligência artificial não se encontraria em um vácuo regulatório, mas poderia responder a institutos, regras e princípios já estabelecidos e consolidados.

Maria Celina Bodin de Moraes (2006) credita no ordenamento constitucional a aptidão para a solução de conflitos envolvendo inovações tecnológicas e responsabilidade civil. A autora pondera que a responsabilidade civil é uma ferramenta jurídica flexível e de ampla aplicabilidade, capaz de abrigar interesses novos que apresentam uma demanda por proteção. Nessa abordagem, o foco desloca-se para a tutela da pessoa da vítima do dano, deixando em segundo plano a atribuição de culpa e a punição.

Por sua vez, Gustavo Tepedino (2012, p. 18-19) observa que novas tecnologias atrairiam, “a um só tempo, regulação de natureza privada e de ordem pública”. Nessa concepção, a aplicação de princípios constitucionais às relações privadas proporcionaria a proteção da pessoa e condicionaria a atividade econômica à observância de “postulados valorativos existenciais”, razão pela qual seria desnecessária a implementação de legislações casuísticas que busquem disciplinar de forma irrefletida todas as relações jurídicas que vão surgindo à medida que a sociedade se desenvolve.

Tepedino e Rodrigo da Guia Silva (2020, p. 5) criticam, ainda, a ansiedade em se disciplinar a matéria de forma apartada, afirmando que, “no afã de se identificarem as particularidades das novas categorias, ameaça-se perder de vista a compreensão do ordenamento jurídico em sua integralidade”.

Juan Francisco Sánchez Barrilao (2016) aborda a questão sob uma perspectiva do direito público, indicando uma responsabilidade constitucional de evitar tais riscos e sanar os conflitos relacionados ao progresso tecnológico, o que não significaria uma

intenção de barrar o desenvolvimento da inteligência artificial mediante institutos do direito constitucional, mas sim, de evitar um desdobramento que negligencie os princípios mais caros à sociedade e que vulnere direitos fundamentais.

São essas as premissas teóricas que fundamentarão a análise da controvérsia e que serão aprofundadas no capítulo seguinte.

2. A RESPONSABILIDADE CIVIL NA LEGALIDADE CONSTITUCIONAL E A TUTELA DA VÍTIMA DE DANOS CAUSADOS PELA INTELIGÊNCIA ARTIFICIAL

Definidos o problema de pesquisa e a hipótese que se pretende verificar, o presente capítulo buscará apresentar os referenciais teóricos do trabalho, situando, primeiramente, a controvérsia no âmbito do Direito Constitucional.

Assim, o estudo que ora se propõe partirá de uma perspectiva Civil-Constitucional, colacionando a lição de autores que abordaram o processo de constitucionalização do Direito Civil e, mais especificamente, da responsabilidade civil, notadamente, no que se refere à consagração dos princípios da solidariedade social e da dignidade da pessoa humana e ao favorecimento da efetiva tutela das vítimas de danos.

Sob esta perspectiva, serão analisados os elementos configuradores do dever de indenizar, de forma a aferir a adequação do instituto da responsabilidade civil enquanto resposta jurídica aos danos causados pela inteligência artificial, considerando as dificuldades apontadas pela doutrina para tanto.

2.1. A perspectiva Civil-Constitucional

Ao longo do século XX, o Direito Civil, antes pautado na proteção da propriedade e do contrato sob uma perspectiva individualista e codificado exclusivamente em caráter infraconstitucional, passou por transformações que levaram à mudança de foco da tutela proporcionada pelos institutos que regiam as relações privadas e culminaram em um novo paradigma de análise: a Constituição.

O Código¹¹ Civil de 1916, de cunho marcadamente individualista e voluntarista, espelhava o entendimento das classes dominantes europeias do século XIX¹²,

¹¹ Pietro Perlingieri (2007, p. 3-4) define código como o “documento (que é uma lei) contendo um conjunto de proposições prescritivas (das quais se extraem normas) consideradas unitariamente, segundo uma idéia de coerência e de sistema, destinadas a construir uma disciplina tendencialmente completa de um setor”.

¹² Eugênio Facchini Neto (2006, p. 20) detalha: “E a partir da ruptura simbolizada pela Revolução francesa, que marca o ingresso na era contemporânea, o direito privado torna-se também *burguês*, no sentido de que o direito privado passa a espelhar a ideologia, os anseios e as necessidades da classe

consagrado pelo Código de Napoleão, voltando-se primordialmente à proteção da autonomia da vontade e da propriedade, ao passo em que a norma constitucional então vigente se prestava à organização política do Estado e tinha por destinatário o legislador ordinário, abstendo-se de interferir na esfera de liberdade dos indivíduos. Considerado a constituição do direito privado, cabia ao Código Civil conceder segurança e estabilidade à atividade econômica, disciplinando de forma exaustiva as regras pelas quais se dariam negócios e transações (TEPEDINO, 2006; 2012).

Sob essa lógica, os direitos fundamentais positivados nas Constituições voltavam-se à disciplina da relação entre o Estado e os indivíduos, enquanto as codificações civis encarregavam-se da regulação das relações entre os particulares. Naquele contexto histórico, concretizar os direitos fundamentais significava assegurar liberdades individuais, cabendo ao Estado apenas deixar de interferir na seara privada para que estivesse em pleno cumprimento da norma constitucional (FACHIN; RUZYK, 2006).

Eugênio Facchini Neto (2006, p. 22) contextualiza:

A ética que predomina nesse período é a da liberdade formal. No âmbito político, temos o Estado liberal e pouco intervencionista. A economia vive o período do liberalismo econômico, com pouca regulamentação estatal. No mundo jurídico, tudo isso desemboca na ética do individualismo, com o aprimoramento, pela pandectística, das figuras do sujeito de direito (enquanto sujeito abstrato) e do direito subjetivo. Os códigos civis deste período caracterizam-se por estarem centrados na propriedade, com ênfase na propriedade imobiliária, com caráter absoluto e individualista, no voluntarismo jurídico, na liberdade contratual, na igualdade meramente formal.

A ideologia jurídica então predominante pretendia a elaboração de uma legislação civil que fosse completa – sem lacunas -, clara – de fácil interpretação e sem ambiguidades ou polissemias – e coerente – sem antinomias -, o que remontava ao “mito do legislador iluminista, inteligente, onisciente, previdente, capaz de tudo regular detalhadamente, antecipadamente, de forma clara e sem contradições”. Para

socioeconômica que havia conquistado o poder em praticamente todos os Estados ocidentais. Como os códigos nascem com pretensões de regular todo o espaço jurídico de uma nação, abandonando-se o intenso pluralismo jurídico que vigorava nos períodos históricos anteriores (em que a legislação régia convivia com o direito canônico, com o direito costumeiro, com a *lex mercatoria*, com o direito das corporações de artes e ofícios, com o direito romano, com os direitos municipais), passa-se a regular *toda* a sociedade a partir das necessidades e ideologias de *uma* fração dessa sociedade, qual seja, a classe burguesa. Ou seja, toda a nação passa a ser convocada a espelhar-se na tábua valores e anseios da burguesia (representados substancialmente pelo liberalismo econômico, tendo a propriedade territorial como valor principal e a liberdade contratual como instituto auxiliar para facilitar as transferências e a criação de riqueza)”.

tanto, a técnica legislativa valia-se de disposições taxativas, contendo o preceito e a respectiva consequência jurídica, sendo raras as cláusulas gerais e os conceitos indeterminados (FACCHINI NETO, 2006, p. 22-23). Nesse contexto, o papel do juiz resumia-se à replicação da letra da lei, no caso, do Código Civil, atribuindo-se ao legislador infraconstitucional a última palavra sobre a resolução de conflitos entre os sujeitos de direito, assim considerados os contratantes e proprietários (TEPEDINO, 2006; 2012).

Por pretender reger as relações privadas com exclusividade, o Direito Civil exacerbava a separação entre as esferas pública e privada, pelo que, no âmbito das relações entre particulares, predominava a liberdade de contratar e adquirir, alheia à imposição de limitações relacionadas a direitos e garantias fundamentais. Isso porque, sob a perspectiva da primeira geração de direitos fundamentais, tais garantias eram exercidas contra o Estado, ente dotado de superioridade em relação aos sujeitos de direitos, aos quais era garantido um espaço de liberdade intangível. Por outro lado, entre os titulares desses direitos, prevaleceria uma relação de igualdade, a qual, porém, se limitava a um aspecto formal – igualdade perante a lei -, não significando uma igualdade fática (FACHIN; RUZYK, 2006).

Esse cenário, todavia, não tardou em deixar exposto o desequilíbrio do quadro social europeu, que, em meados do século XIX, passou a observar demandas que evidenciavam as injustiças e vicissitudes do processo de industrialização, potencializadas pelas Grandes Guerras, e que, por resultarem na necessidade de maior intervenção do Estado na economia, impactaram diretamente no Direito Civil, o que veio a ter reflexos legislativos no Brasil a partir do início do século XX (TEPEDINO, 2006).

A centralidade que se atribuiu ao diploma de 1916, quando da sua promulgação, foi sendo gradativamente mitigada, diante da necessidade do Estado de endereçar medidas para contemplar as demandas e os conflitos sociais emergentes, assim como pelo surgimento de hipóteses jurídicas impostas pela realidade econômica e que não haviam sido contempladas pelo Código Civil (TEPEDINO, 2006).

A partir da década de 1930, observou-se, no Brasil, um comportamento estatal de maior intervenção na economia conhecido por dirigismo contratual, mediante o qual se buscou refrear o ideário individualista liberal e reduzir o acirramento das

desigualdades, através da adoção de técnicas legislativas que passaram a estipular políticas públicas para atingir metas econômicas e objetivos sociais. Com esse intuito, vislumbra-se o surgimento de leis extravagantes que passaram a subtrair situações jurídicas do escopo do Código Civil, para discipliná-las integralmente por meio de diplomas setoriais - os estatutos -, caracterizados por uma técnica legislativa que visava a definição de diretrizes, metas, programas e políticas públicas, valendo-se de cláusulas gerais e abertas¹³ no lugar da tipificação taxativa das codificações (TEPEDINO, 2006; 2012).

Por sua vez, as normas constitucionais que surgiram no contexto do pós-guerra estipularam limitações ao exercício da autonomia privada e da propriedade, ao disciplinar matérias antes reservadas ao código, mediante princípios e regras que atribuíram deveres sociais aos sujeitos das relações privadas, como é o caso da função social da propriedade, dos limites à atividade econômica e da tutela da família, além de endereçar compromissos ao legislador ordinário. Paralelamente, alterou-se, também, a dogmática do Direito Civil, cujas categorias já não correspondiam mais à realidade econômica e aos fatos sociais, os quais exigiam uma atenção do legislador, do intérprete e da doutrina com o conteúdo e as finalidades das relações entre os particulares (TEPEDINO, 2006).

Nesse processo, em que a estabilidade e a segurança do Código Civil cedem espaço para regulações setoriais voltadas à concretização de objetivos sociais e econômicos, a promulgação da Constituição de 1988 se dá como a consagração do intuito de se implementar um Estado social, conferindo unidade sistemática ao conjunto de estatutos e confirmando a superação da dicotomia público *versus* privado, que já não correspondia mais à lógica jurídica vigente (FACCHINI NETO, 2006; TEPEDINO, 2006; 2012).

Pietro Perlingieri (2007, p. 54-55) fala em crise da separação entre o direito público e o direito privado, apontando para os pontos de confluência entre essas esferas:

O Direito Civil não se apresenta em antítese ao Direito Público, mas é apenas um ramo que se justifica por razões didáticas e sistemáticas, e que recolhe e evidencia os institutos atinentes com a estrutura da sociedade, com a vida

¹³ Conforme leciona Virgílio Afonso da Silva (2005, p. 78), cláusulas gerais são “cláusulas que requerem um preenchimento valorativo na atribuição de sentido, pois são, para usar uma expressão difundida na doutrina jurídica brasileira, conceitos abertos, cujo conteúdo será definido por uma valoração do aplicador do direito”.

dos cidadãos como titulares de direitos civis. Retorna-se às origens do direito civil como direito dos cidadãos, titulares de direitos frente ao Estado. Neste enfoque, não existe contraposição entre privado e público, na medida em que o próprio direito civil faz parte de um ordenamento unitário.

O ponto de referência para as relações de Direito Civil desloca-se para a Constituição, que, ante a sua posição hierárquica no ordenamento jurídico, confere unidade à pluralidade de universos legislativos existentes (TEPEDINO, 2006). Descreve-se, assim, por um lado, a redução do protagonismo do Código Civil¹⁴ no âmbito do direito privado e, por outro, uma ascensão da importância concedida à Constituição no ordenamento jurídico, a qual, além de ocupar o topo da ordem jurídica, passou a ter reconhecido o caráter imperativo de suas disposições (LEAL, 2015).

Nesses termos, o processo que se pode denominar constitucionalização do Direito Civil¹⁵, ao mesmo tempo em que alçou matérias tipicamente cíveis para o âmbito da ordem pública, também fez irradiar princípios constitucionais para o espaço das relações privadas (FACCHINI NETO, 2006).

O âmbito de atuação do intérprete e aplicador do Direito Civil passa a responder, primeiro, à norma constitucional, parâmetro máximo para a validade da regulamentação e para a interpretação de relações jurídicas tanto públicas quanto privadas. Ressalta-se, quanto ao ponto, o caráter histórico e relativo dos institutos jurídicos - no caso específico, do direito privado -, em contraposição à ideia de conceitos absolutos e dotados de neutralidade, o que diz respeito ao reconhecimento de alterações semânticas que ocorrem ao longo do tempo, pelas quais se atribui novos significados a velhas terminologias (LEWICKI, 2014; MORAES, 2006; PERLINGIERI, 2019; TEPEDINO, 2012).

Para Perlingieri (2019), deve-se reconhecer que o conhecimento jurídico é sempre relativo, considerando que os conceitos e instrumentos se alteram conforme cada lugar e cada época. Ao passo em que a realidade está sempre em movimento, também as leis e os institutos jurídicos estão em constante e rápida mutação, o que

¹⁴ Segundo Fachin e Ruzyk (2006, p. 97), a promulgação do Código Civil de 2002 não logrou em alterar a lógica do seu antecessor, pois, baseado na ideia de relação jurídica, condiciona o reconhecimento de direitos subjetivos à existência prévia de um modelo respectivo, deixando de lado a concretização da dignidade da pessoa humana e dos direitos fundamentais que dela decorrem. Dito de outra forma: “quem não se insere em dado modelo de relação jurídica não teria direitos subjetivos”. No mesmo sentido, manifesta-se Maria Celina Bodin de Moraes (2006), ao afirmar que o texto do código de 2002, elaborado em 1970, contém anacronismos e deficiências que vão de encontro ao movimento de personalização que a jurisprudência e a doutrina vinham promovendo.

¹⁵ Tepedino (2012, p. 16) menciona os termos “[s]ocialização, desapatrimonialização, repersonalização, constitucionalização do Direito Civil”.

obriga o jurista a adaptar seus instrumentos ao mundo dos fatos. Para tanto, o autor afasta, de um lado, o formalismo e o dogmatismo jurídico, de forma que não se exacerbe a aplicação da letra fria da lei em detrimento do seu espírito e substância, e nem se exaspere a utilização de conceitos e dogmas absolutos, apartados do ordenamento jurídico e da realidade.

Por outro lado, Perlingieri (2019, p. 3) afasta também o pragmatismo jurídico, para que o mundo do ser não se sobreponha ao mundo do dever-ser, destacando a importância do Direito para a transformação da realidade e não apenas para retratar a natureza das coisas. Busca-se, assim, equilibrar a forma e a substância, pois a “interpretação e o conhecimento do Direito não se podem limitar à letra da lei, mas devem compreender o desafio dialético entre a lei e o fato, entre a realidade e a lei, entre o individualismo da lei (e os valores que estão em jogo) e o caso concreto”.

Opera-se, assim, uma guinada valorativa no campo civilista, pela qual se buscou interpretar os institutos do direito privado à luz dos postulados existenciais presentes no texto da Constituição, a exemplo dos fundamentos da dignidade da pessoa humana e do valor social do trabalho, bem como dos objetivos fundamentais da construção de uma sociedade livre, justa e solidária e da promoção do bem de todos, sem qualquer forma de discriminação (TEPEDINO, 2012). Nesse sentido, Tepedino (2006, p. 48) aponta não se tratar “de ler a normativa especial através de seus próprios princípios – como se fora um microssistema -, encontrando-se tais preceitos setoriais condicionados, vinculados, instrumentalizados, ao projeto constitucional”.

É o que indica, também, Maria Celina Bodin de Moraes (2006, p. 234), ao afirmar:

Neste contexto, dito pós-positivista, o respeito das normas inferiores à Constituição não é examinado apenas sob o ponto de vista formal, a partir do procedimento de sua criação, mas com base em sua correspondência substancial aos valores que, incorporados ao texto constitucional, passam a conformar todo o sistema jurídico. Valores que adquirem positividade na medida em que consagrados normativamente sob a forma de princípios. Assim, a solução normativa aos problemas concretos não se pauta mais pela subsunção do fato à regra específica, mas exige do intérprete um procedimento de avaliação condizente com os diversos princípios jurídicos envolvidos.

Discorrendo sobre o tema, Fachin e Ruzyk (2006, p. 101) afastam a ideia de que a constitucionalização do Direito Civil prejudicaria a precisão dos conceitos, bem

como a autonomia das disciplinas envolvidas – civil e constitucional -, reputando o argumento como uma “fetichização de uma racionalidade sistêmica fechada”, que estaria a vislumbrar o direito como um fim em si mesmo e não como um instrumento destinado a dar concretude à dignidade da pessoa humana. Para os autores, a pretensão de se solucionar conflitos através da operação mecânica da subsunção do fato à norma, acionando-se uma pretensa autonomia principiológica e conceitual do Direito Civil, acabaria por gerar situações de exclusão e violação a direitos fundamentais, em hipóteses nas quais não houvesse modelo jurídico previamente estabelecido.

Embora complexo, pois composto por normas de procedências diversas, o ordenamento jurídico é dotado de unidade, ante a sua estrutura escalonada e hierárquica, em cujo topo se encontra a Constituição¹⁶, de onde decorre o fundamento para todas as demais normas (BOBBIO, 1995). A hierarquia das fontes, todavia, não trata de mero mecanismo formal de resolução de conflitos através da relação de superioridade e subordinação entre diferentes normas, mas de lógica que diz com a observância necessária do conteúdo e da filosofia de vida expressos pelo modelo constitucional (PERLINGIERI, 2007).

Para tanto, uma norma nunca deve ser considerada de forma isolada, pois assume um significado de acordo com o sistema no qual está inserida. Daí se falar que a “verdadeira interpretação jurídica não pode ser literal, não pode ser só lógica: deve ser sistêmica e axiológica”, ou seja, deve funcionalizar os princípios fundamentais de todo o sistema jurídico (PERLINGIERI, 2019, p. 4).

Dito de forma mais detalhada:

Um enunciado linguístico torna-se norma quando é lido e confrontado com o inteiro ordenamento, em dialética com os fatos históricos concretos, com as relações individuais e sociais. A função do sistema é, portanto, necessária – não como resultado estático, mas – como o instrumento e o fim dinamicamente conhecíveis, como uma experiência cultural global, idônea a transformar a lei em direito, o enunciado linguístico em norma. O sistema

¹⁶ Norberto Bobbio (1995) aponta como norma fundamental, ou seja, o fundamento último do ordenamento, aquela que respalda de forma implícita a legitimidade do poder constituinte originário e, conseqüentemente, impõe obediência a todo o sistema jurídico. Além de ser caracterizado pela unidade, segundo Bobbio, o ordenamento jurídico constitui um conjunto ordenado de normas, ou seja, é formado por partes que devem guardar coerência lógica entre si. Para tanto, não há que se falar em antinomias, pois, em caso de incompatibilidade entre duas normas, uma delas será excluída para que o todo seja preservado. Por fim, além de unidade e coerência, o ordenamento jurídico é dotado também de completude, pois nele estão contidas as normas para regular qualquer situação, não havendo hipóteses quanto à proibição ou à permissão de dado comportamento que não possam ser deduzidas do sistema.

jurídico não é puramente eventual porque as relações conteudísticas (contentutistiche), do qual é expressão, representam um componente essencial mesmo na interpretação do enunciado legislativo individualmente considerado. A unidade interna não é um dado contingente, mas, ao contrário, é essencial ao ordenamento, sendo representado pelo complexo de relações e de ligações efetivas e potenciais entre as normas singulares e entre os institutos. O conhecimento, cientificamente apreciável, é totalidade, superação do finito, uma continuação sem fim, um momento singular e coletivo de um movimento perene.

Não existem normas, portanto, que não tenham como pressuposto o sistema e que ao mesmo tempo não concorram a formá-lo; não existem normas que sejam inteligíveis no seu efetivo alcance se não inseridas como partes integrantes, em uma totalidade formal (sistema legislativo) e substancial (sistema social). Este resultado postula a superação da exegese considerada exclusivamente como investigação e individuação do significado literal do texto (PERLINGIERI, 2007, p. 78).

Assim, além de consagrar a Constituição como ápice do sistema jurídico, uma perspectiva civil-constitucional¹⁷ endossa a ideia de que o texto constitucional expressa os valores primordiais da coletividade e que, como tal, garante unidade ao ordenamento jurídico e condiciona o comportamento do Estado e dos particulares, posicionando os direitos fundamentais como o “sistema de referência da ordem constitucional” (LEAL, 2015, p. 128).

Trata-se, ademais, de reconhecer aplicabilidade imediata e direta aos direitos e garantias fundamentais, no sentido de que vinculam não apenas os poderes constituídos, mas também incidem sobre as relações interprivadas, ainda que não exista atuação do legislador infraconstitucional a positivar certos direitos subjetivos (LÔBO, 2014; SARLET, 2012). Isso porque, conforme leciona Facchini Neto (2006, p. 51), “[c]aso a eficácia de um direito fundamental dependesse de uma legislação infraconstitucional que o implementasse, correr-se-ia o risco de a omissão do legislador ordinário ter mais força eficaz do que a ação do legislador constituinte”.

¹⁷ Paulo Lôbo (2014) adverte que o Direito Civil-Constitucional não significa uma disciplina autônoma e descolada do Direito Civil, mas uma ressignificação de categorias e conceitos para atender às mudanças da sociedade. Trata-se da compreensão do Direito Civil à luz da Constituição. Por sua vez, Virgílio Afonso da Silva (2005) considera que a denominação civil-constitucional careceria de sentido, pois normas de direito civil não deixariam de sê-lo apenas pelo fato de estarem consagradas no texto constitucional e, além disso, falar em Direito Civil-Constitucional poderia levar à conclusão equivocada de que há uma parte do Direito Civil que não deve observância à Constituição. Para os fins deste trabalho, nos parece que o Direito Civil-Constitucional se faz relevante enquanto contraposição ao Direito Civil patrimonialista e individualista herdado do paradigma do Estado Liberal, porém, não compreende uma disciplina nova, mas sim, uma mudança metodológica, como mencionam Lôbo (2014), Moraes (2006) e Tepedino (2006).

Embora existam argumentos em sentido contrário¹⁸, é possível extrair do conteúdo do artigo 5º, § 1º, da Constituição de 1988, o qual confere aplicação imediata às normas que definem os direitos e garantias fundamentais, a noção de que não apenas os órgãos estatais, mas também os particulares, em suas relações, estariam submetidos de maneira direta à disciplina constitucional. Assim sendo, adota-se a premissa de que os direitos fundamentais, tanto na sua dimensão negativa, quanto na positiva (prestacional)¹⁹, seriam dotados de uma eficácia direta *prima facie* nas relações entre particulares, a qual seria operacionalizada conforme as circunstâncias de cada caso concreto e as peculiaridades dos direitos envolvidos, mediante técnicas de interpretação e solução de conflitos, assim como de argumentação e motivação, que assegurem uma resposta adequada à luz da Constituição²⁰(SARLET, 2012).

O que se absorve do exposto até o momento acerca do processo de constitucionalização do Direito Civil é que a tutela da pessoa e a promoção da dignidade humana demandam uma visão sistemática do ordenamento jurídico, a ser compreendido em sua unidade e completude, a partir do que orienta a Constituição, que é o fator que confere harmonia à pluralidade de disciplinas jurídicas.

Ademais, parte-se do reconhecimento de que o mundo da vida, em seu dinamismo e complexidade, não é passível de ser capturado em disposições taxativas de lei, sob pena de, nesse mister, se incorrer em regulações tão casuísticas quanto excludentes. Sob essa perspectiva, e graças a uma mudança de comportamento do legislador, que passou a adotar cláusulas gerais, porosas ao sistema de valores constitucionais, verificou-se que a eventual ausência de direitos subjetivos

¹⁸ As divergências consistem em (i) não atribuir quaisquer efeitos dos direitos fundamentais às relações privadas, mas somente àquelas entre o Estado e os particulares; (ii) atribuir efeitos indiretos dos direitos fundamentais às relações privadas, demandando, para tanto, a atuação do legislador ordinário e (iii) atribuir efeitos diretos dos direitos fundamentais às relações entre os particulares (SILVA, 2005).

¹⁹ Sarlet (2012, p. 28) exemplifica como os direitos fundamentais poderiam ter eficácia sobre as relações entre os particulares em uma dimensão prestacional: “uma avaliação mesmo sumária da evolução do tema a partir da prática jurisprudencial no Brasil facilmente leva à identificação de uma série de exemplos no que diz com uma influência dos direitos sociais a prestações no Direito Privado. Um dos casos dignos de nota refere-se ao reconhecimento por parte dos Tribunais da obrigação de empresas gestoras de planos de saúde privados arcarem com o pagamento de tratamentos não previstos nas cláusulas do contrato”.

²⁰ No campo da prestação jurisdicional, Sarlet (2012, p. 29-30) constata abusos na seara da constitucionalização do direito privado, decorrentes da desconsideração pelo uso de estruturas argumentativas e de métodos de interpretação adequados, o que acarretaria uma situação na qual mesmo aqueles simpáticos à teoria da eficácia direta dos direitos fundamentais entre os particulares estariam a pugnar por cautela, considerando que, paralelamente à “hipertrofia da Constitucionalização da ordem jurídica” pode-se observar uma “excessiva e problemática judicialização das relações sociais”.

expressamente positivados não impede a fruição de direitos e garantias fundamentais, os quais, ante sua posição hierárquica no ordenamento jurídico, devem condicionar a tarefa do intérprete e aplicador do Direito para que sejam concretizados.

Retomando o contato com o objeto desta dissertação, nada tem proporcionado tantas e tão rápidas transformações na sociedade quanto a tecnologia, que tem, desde há muito, pressionado os institutos do Direito Civil, notadamente, no que se refere ao âmbito da responsabilidade. Afunilando, pois, o escopo da investigação, o que se pretende analisar a partir de agora é a capacidade do ordenamento jurídico de promover a tutela de uma pessoa específica, qual seja, a vítima de danos causados pela inteligência artificial.

Gustavo Tepedino e Rodrigo da Guia Silva (2020) observam que o desenvolvimento da tecnologia, não raro, é associado à necessidade de se elaborar novos instrumentos jurídicos, diante da percepção de que a novidade técnica demandaria o suprimento de lacunas no ordenamento jurídico, o que explicaria a recente proposição da criação de ramos do direito voltados especificamente aos sistemas de inteligência artificial, como o Direito da Robótica, espécie do gênero Direito Digital ou Direito Cibernético.

Não obstante, advertem que tal pretensão acaba por proporcionar um tratamento assistemático da inteligência artificial, destacando que, embora desafiadores e inusitados, os problemas suscitados em matéria de IA relacionam-se com categorias já sedimentadas na doutrina do direito civil.

Nesse contexto, a enunciação de supostos vazios normativos representa problema muito mais grave do que o mero abalo à dogmática consolidada na tradição jurídica. Com efeito, ao afrontar a unidade e a completude do ordenamento, a indicação insistente de lacunas finda por comprometer a própria efetividade da tutela prometida às vítimas de danos injustos, como se das suas necessidades não desse conta o sistema ora vigente. Em vez de buscar – muitas vezes irrefletida – novas soluções e novos diplomas legais, melhores resultados se haverão de alcançar pelo esforço de releitura dos institutos já conhecidos pela civilística. Desse modo, ainda que determinada questão relacionada à inteligência artificial não corresponda imediatamente ao alcance tradicional de certas previsões normativas, poderá o intérprete perquirir o seu sentido com fundamento nos valores do ordenamento, no intuito de encontrar a solução para os novos problemas (TEPEDINO; SILVA, 2020, p. 306).

Com atenção a esse processo, o presente estudo volta-se à responsabilidade civil, de forma a compreender a evolução do instituto e de seus elementos

configuradores ao longo do tempo e a forma com que os danos causados por uma inteligência artificial podem ser contemplados pelo ordenamento jurídico.

2.2. Da responsabilidade civil ao Direito de Danos: superando as barreiras para a tutela da vítima

Para a solução de problemas relacionados aos danos causados pela inteligência artificial, Tepedino e Silva (2020) suscitam o tratamento sistemático da matéria, mediante a observância do ordenamento jurídico em sua unidade e complexidade. Para os autores, a disciplina da responsabilidade civil e consumerista, quando fundamentada na tábua axiológica constitucional, mostra-se suficiente para o endereçamento de questões atinentes aos danos atribuídos a sistemas autônomos, aduzindo que, “ao ineditismo das questões suscitadas pelas novas tecnologias não há de corresponder necessariamente o ineditismo das soluções jurídicas” (p. 307).

Reconhecer a historicidade e a relatividade dos conceitos se faz necessário para o presente estudo, considerando que estamos a testemunhar desafios diferentes daqueles decorrentes das Revoluções Industriais, complexificados pela introdução da inteligência artificial, da coleta massiva de dados pessoais, do reconhecimento facial, de carros autônomos e de outras tecnologias que, em tese, podem oferecer risco e causar danos (QUEIROZ, 2020).

Para a compreensão de onde chegamos, é pertinente percorrer o caminho pelo qual a disciplina da responsabilidade civil passou enquanto o direito privado afastava-se dos valores individualistas e patrimonialistas para submeter-se aos postulados existenciais da Constituição.

2.2.1. A tutela da vítima de danos

A partir de uma perspectiva liberal e individualista, a responsabilidade civil da Modernidade baseava-se na ideia de mau uso da liberdade individual, a qual, para ser

exercida sem entraves, precisava estar atrelada à ideia de responsabilidade. É dizer, o livre-arbítrio atrai para aquele que o exerce o dever de evitar consequências danosas de seus atos (SCHREIBER, 2009). Tratava-se de uma contraposição ao sistema medieval europeu, que não contava com institutos de responsabilidade puramente civil para o controle da liberdade individual, conforme leciona Anderson Schreiber (2009, p. 12-13):

Para a tarefa de controle da liberdade individual, mostravam-se impróprios os sistemas de responsabilidade coletiva e vingança familiar que predominaram no Medievo europeu, já que, ao transcender em suas consequências a esfera de cada indivíduo, constituíam verdadeira antítese do binômio liberdade-responsabilidade dos juristas modernos. Igualmente inconveniente mostrava-se o sistema de responsabilidade delitual, que havia gradualmente substituído a responsabilidade coletiva em favor de uma responsabilidade individual do autor do delito, mas que se mostrava excessivamente preso à delimitação típica de delitos e penas, muitas vezes de natureza corporal. Daí a necessidade de construção de um sistema de responsabilidade moderno, puramente civil, desvinculado da tradição medieval, e fundado não na violação de normas penais expressas, mas no ilegítimo exercício da liberdade individual, identificado, desde logo, e de forma exclusiva, com a noção de culpa. Na consagrada lição da doutrina francesa, a pedra de toque da responsabilidade civil consiste no “uso culpável da liberdade, que moralmente demanda uma sanção”. Como se vê, a culpa invocada pelos juristas da Modernidade possuía forte conotação moral.

Sob o Código Civil de 1916, obedecendo à mesma lógica, o instituto da responsabilidade civil baseava-se na prática de um ato ilícito, com fundamento em cláusula geral que, para a configuração de seu elemento nuclear, exigia a aferição de culpa. O dever de indenizar surgia desde que se constatasse culpa por parte do agente que causou o prejuízo, cabendo à vítima provar o elemento subjetivo que orientou a prática do ato danoso (MORAES, 2006a). Nesse sentido, sempre que quisesse pleitear uma indenização, a vítima de um dano precisaria, além de provar o prejuízo sofrido, demonstrar a culpa do agente e o nexo de causalidade entre a conduta e o injusto. Tais requisitos funcionavam como verdadeiros filtros para o acolhimento de demandas de ressarcimento pelo Poder Judiciário, o que atendia ao ideário liberal segundo o qual a autonomia privada só poderia ser limitada em caso de uso patentemente ilegítimo da liberdade individual (SCHREIBER, 2009).

A dificuldade de comprovação da culpa exacerbou-se com o desenvolvimento industrial e tecnológico, ante a proliferação de acidentes em que a vítima carecia do conhecimento técnico pertinente à demonstração das causas do ocorrido, o que acabava por torná-la ainda mais vulnerável perante o culpado pelo prejuízo. Como

resultado, grande parte dos danos ficavam sem ressarcimento, o que proporcionou a reação dos tribunais e da doutrina, que passaram a estipular expedientes destinados a facilitar o acesso da vítima à reparação (SCHREIBER, 2009).

Ainda no início do século XX, o legislador brasileiro viu-se diante da necessidade de atuar em razão do incremento do uso e, conseqüentemente, da quantidade de acidentes envolvendo transportes ferroviários, implementando mecanismos para que se retirasse da vítima o ônus de comprovar a causa dos prejuízos sofridos por eventos dessa natureza. Assim, em 1912, promulgou-se o Decreto nº 2.681, que responsabilizava as estradas de ferro pelos desastres ocorridos com os viajantes, mediante culpa presumida²¹. Pouco depois, em 1919, promulgou-se o Decreto Legislativo nº 3.724, que adotou a responsabilidade objetiva do empregador em caso de acidentes de trabalho. Sucederam-se a essas normas legislações especiais que consagraram a responsabilidade objetiva, por exemplo, para atividades de mineração, acidentes veiculares, atividades nucleares, atividades causadoras de danos ao meio ambiente e transporte aéreo (MORAES, 2006a).

Portanto, aquele sistema de responsabilidade civil contava, como regra, com uma cláusula geral baseada na culpa e, excepcionalmente, com normas especiais que previam hipóteses de responsabilização independente da comprovação de um ato ilícito (MORAES, 2006a).

A ideia de se atribuir a alguém o dever de indenizar independente da prática de um ilícito é próprio da dinâmica da sociedade industrializada, na qual a ocorrência de acidentes não é mais considerada uma fatalidade, mas um fenômeno esperado, estatisticamente calculável e inerente ao curso regular das atividades desenvolvidas em coletividade. Assim, por se tratar de danos que decorrem da organização em sociedade, “nenhuma causa, nem transcendente nem pessoal, pode disso dar conta” (MORAES, 2006a, p. 17).

Quanto ao ponto, faz sentido reconhecer que vivemos em uma sociedade de risco, termo atribuído a Ulrich Beck (1992), para quem a produção social da riqueza é sistematicamente acompanhada pela produção social de riscos. Remetendo-se ao pensamento desenvolvido por Beck, Maria Celina Bodin de Moraes (2006) descreve uma sociedade preocupada com seu futuro e com a vida das próximas gerações e

²¹ Na culpa presumida, “é permitido ao ofensor adentrar-se à discussão de culpa para, por meio da contraprova oriunda do ônus invertido, afastar o elemento culpa que lhe foi imposto por presunção relativa” (FALEIROS JÚNIOR, 2019).

que se vê instada a buscar meios para protegê-los, anseio que se potencializa com a velocidade com que a tecnologia se desenvolve, em contraste com a lentidão com que se adaptam as soluções jurídicas para contemplar conflitos novos.

Nesse contexto, a função da responsabilidade civil renova-se de forma a priorizar a reparação dos danos sofridos pela vítima, em detrimento da punição do agente causador do dano, eis que, sob a lógica da culpa, os danos reconhecidamente inevitáveis ficariam sem qualquer reparação (MORAES, 2006a). Tal renovação pode ser atribuída ao “sentimento coletivo a respeito da injustiça de deixar a vítima em uma posição desfavorável em relação ao dano que lhe foi ocasionado, mesmo consciente de que o autor do dano pode não ter agido de forma causá-lo culposamente” (MULHOLLAND, 2010, p. 16).

Por conseguinte, o processo de objetivação da responsabilidade civil corresponde a relevantes transformações socioculturais que escancararam a insuficiência do modelo de responsabilidade subjetiva para contemplar as relações jurídicas da sociedade industrializada. Traduz-se, assim, a superação do modelo individualista de responsabilidade e o acolhimento do modelo solidarista, baseado na Constituição e dedicado à proteção da pessoa humana (MORAES, 2006a).

A superação da tradicional tendência moralizadora e punitiva da responsabilidade civil abriu espaço para a percepção de um dever geral de solidariedade, positivado no artigo 3º, inciso I, da Constituição de 1988 e traduzido na obrigação de agir de forma a não lesar interesses de terceiros, o que resulta da já aduzida historicidade dos conceitos, que permite dar nova conotação ao instituto (MORAES, 2006). Trata-se do reconhecimento de que existe uma relação ética entre os indivíduos e deles com a comunidade e com o mundo, inclusive com as gerações que ainda estão por vir, estando justificadas, portanto, medidas que visem mitigar os efeitos nocivos de determinadas atividades ao invés de buscar punir os seus autores. É esse, aliás, o fundamento “que dá foros de constitucionalidade, generalidade e eticidade à responsabilidade objetiva em todas as hipóteses em que ela se manifesta” (MORAES, 2006a, p. 24).

A lógica jurídica correspondente a tal ideário solidarista importa em reconhecer a imprescindibilidade da coexistência humana e do respeito ao outro, visando o bem comum. Além disso, pressupõe abandonar o individualismo e admitir que cada um de nós tem responsabilidades em relação à coletividade, ainda mais, no que se refere à

proteção da parte mais vulnerável das relações sociais. Por outro lado, sendo impossível ditar que alguém se comporte de forma solidária para com os demais, cabe ao Direito impor obrigações jurídicas de forma a condicionar o comportamentos dos particulares em direção à concretização de uma sociedade mais justa (SARMENTO, 2006).

No contexto atual, a Constituição determina — ou melhor, exige — que nos ajudemos, mutuamente, a conservar nossa humanidade porque a construção de uma sociedade livre, justa e solidária cabe a todos e a cada um de nós. Este é o fundamento da reparação dos danos pessoais injustamente sofridos, não mais ignorados ou, antes, suportados solitariamente pela vítima, mas, na lógica da justiça social distributiva, transferidos, sempre que possível, à comunidade. Com efeito, subordinando-se o conceito de responsabilidade à efetiva reparação dos danos (injustos) sofridos pela vítima, independentemente da identificação de um culpado, ressalta-se a relação de solidariedade entre a coletividade (na qual se inclui o autor do dano) e a vítima, evidenciando-se desta forma a opção, pelo ordenamento jurídico, da valorização da pessoa humana, a qual terá o seu prejuízo ressarcido (MORAES, 2006a, p. 27).

Inaugurando um novo contexto axiológico, a Constituição de 1988 consagrou o comprometimento com a proteção à vítima de danos, o que teve reflexos em nível infraconstitucional, conforme se observa pela edição do Código de Defesa do Consumidor, pelo qual se instituiu a responsabilidade objetiva do fornecedor de produtos e serviços (SCHREIBER, 2009).

O Código Civil de 2002, por seu turno, estipulou a responsabilidade objetiva em hipóteses para as quais o diploma anterior previa a culpa presumida, como é o caso da responsabilidade por fato de terceiro e por fato de animais (SCHREIBER, 2009).

Além disso, estabeleceu, paralelamente à cláusula geral da responsabilidade subjetiva, outra regra geral fundamentada no risco da atividade, prevendo, no parágrafo único do artigo 927, a “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).

A releitura dos institutos clássicos do direito privado à luz das finalidades consagradas pela Constituição e das transformações da sociedade, processo pelo qual foram operadas significativas mudanças semânticas em terminologias já conhecidas, é decorrência necessária da observância à hierarquia das fontes, condição para o estudo do Direito sob uma metodologia civil-constitucional (MORAES, 2006). No caso da responsabilidade civil, a adoção de novos parâmetros não

patrimoniais e voltados à tutela da pessoa, em grande medida, decorre da insuficiência daquele Direito Civil tradicional para absorver os impactos das novas tecnologias que, por gerarem novas hipóteses de risco, não raro, ensejam situações que vulneram direitos fundamentais e demandam novas formas de proteção (EHRHARDT JÚNIOR, 2014).

A proteção da pessoa humana orientou, neste âmbito, a extensão da tutela da vítima de danos e o aumento das hipóteses de danos ressarcíveis, o que denota a flexibilidade do instituto da responsabilidade civil que, ante a sua simplicidade, mostra uma aptidão a acolher interesses novos que, de outra forma, restariam desamparados, eis que ainda não contemplados por uma regulamentação própria do legislador ordinário. Evidencia-se, por esse motivo, a preponderância do papel do magistrado em observar as mudanças sociais e, na operação de preencher as cláusulas gerais e os conceitos indeterminados a partir da análise do caso concreto, fazer incidir os princípios e valores constitucionais nas relações em conflito²² (MORAES, 2006).

Se, sob uma perspectiva liberal, o lucro era individualizado e o prejuízo era socializado, o que se almeja agora é “a diminuição dos conflitos sociais através da distribuição dos riscos das atividades empresariais capitalistas e das vantagens econômicas por elas geradas” (EHRHARDT JÚNIOR, 2014, p. 306). Nesse sentido, o sistema de responsabilidade objetiva induz o agente a internalizar o custo da atividade econômica desempenhada, uma vez que, sendo obrigado a arcar com os prejuízos eventualmente causados, lhe será conveniente agir de forma a “escolher o nível de atividades que acarreta a maior diferença entre a utilidade resultante da atividade e os danos por ela produzidos” (MORAES, 2006a, p. 19).

Ao passo em que o foco de preocupação deixa de ser o sujeito responsável pelo dano e passa a ser a reparação do prejuízo sofrido pela vítima, também a esfera patrimonial individual cede espaço para a proteção de situações existenciais, as quais

²² Para tanto, defende Ramos (2015, p. 21): “O atípico sistema de direito de danos brasileiro, articulado, à francesa, por cláusulas gerais, imprescinde, para sua construção e atuação, de sensatos exercícios de judicatura, os quais, por seu turno, não podem ser senão informados pela literatura jurídica. Mais que isso: a sisífica (re)construção deste sistema clama, para se tornar factível, por arranjos que lhe permitam conferir previsibilidade ao tráfego jurídico, sem perder em maleabilidade. Trocando em miúdos: a realização das potencialidades do sistema brasileiro de responsabilidade civil demanda o acoplamento de uma postura hermenêutica ao mesmo tempo tópica e sistemática à afirmação dum mecanismo estabilizador que já se delinea no horizonte de sentido posto diante da comunidade jurídica brasileira: o precedente judicial”.

dizem respeito à sociedade como um todo, sob uma lógica de justiça distributiva e não apenas retributiva (EHRHARDT JÚNIOR, 2014).

No entanto, na concepção de Schreiber (2009, p. 8), não se trata meramente de relativizar os elementos configuradores da responsabilidade para transferir de um indivíduo a outro o ônus reparatório²³. Para o autor, a responsabilidade civil presta-se a conferir efetividade à reparação e a reduzir a sua litigiosidade, devendo voltar-se à gestão dos danos que resultam da convivência, pois a “implementação de uma genuína responsabilidade social requer que o resultado das ações de responsabilização passe a ser repartido pela coletividade de agentes potencialmente lesivos ou até, em algumas hipóteses, pela sociedade como um todo”.

Menciona, assim, que já se opera uma solidarização da reparação de danos por meios indiretos, como ocorre com a estipulação, pelo legislador, de mais hipóteses de responsabilidade solidária e de mecanismos de prevenção e precaução de danos, bem como o desenvolvimento, pela iniciativa privada, de seguros de responsabilidade civil, instrumentos que tendem a distribuir os custos da administração dos riscos e a diluir o ônus da reparação por toda a sociedade ou, pelo menos, pelo grupo de agentes potencialmente lesivos. Ademais, sugere formas mais diretas de difusão de danos, como a implementação de sistemas de seguridade social ou de seguros obrigatórios por lei, como ocorre com os acidentes automobilísticos e outras atividades consideradas de risco elevado. Nesse sentido:

O essencial é que se busque, por via hermenêutica ou por reforma legislativa, a substituição de uma responsabilidade individual – que, estendida pelo imperativo de proteção à vítima, acaba distribuindo de forma cada vez mais aleatória e ineficiente o custo das reparações – por uma autêntica responsabilidade social, que atribua a cada pessoa um ônus correspondente ao seu real potencial lesivo, transformando o problema dos danos em um problema de toda a sociedade (SCHREIBER, 2009, p. 248).

Portanto, diluir os danos pela coletividade remete à observância do princípio da solidariedade social e ao atendimento da função contemporânea da sociedade civil, qual seja, assegurar a tutela da vítima, sem, todavia, onerar de forma exacerbada um único indivíduo apontado como responsável. Trata-se, assim, de vislumbrar a responsabilidade civil “como um renovado instrumento a serviço não de pretensões

²³ Schreiber (2009, p. 7) afirma que o “ônus de auxiliar as vítimas pertence a todos, mas vem atribuído a cada réu, aleatória e isoladamente, o que acaba por resultar em injustiça, a rigor, tão grave quanto manter o dano sobre a vítima”.

individuais, mas voltado à solução dos danos como problema coletivo e social” (SCHREIBER, 2009, p. 229).

2.2.2. A mitigação dos filtros para a reparação

Até o momento, o que se tem é a ideia abstrata de que a responsabilidade civil, transformada ao longo do século XX, tem sido funcionalizada em prol da proteção da pessoa da vítima, com fundamento em valores extraídos do ordenamento constitucional, como os princípios da dignidade da pessoa humana e da solidariedade social. Todavia, até chegarmos à confirmação da hipótese da presente dissertação, ou seja, para saber se a disciplina da responsabilidade civil, da forma como se encontra, está apta a contemplar os danos causados pela inteligência artificial, faz-se necessário conferir concretude ao estudo, percorrendo-se os elementos caracterizadores da obrigação de indenizar.

A matéria tem especial relevância para o problema desta pesquisa, considerando que a atribuição de responsabilidade por danos causados por uma inteligência artificial costuma esbarrar justamente na relação de causalidade entre o evento danoso e a conduta de alguma pessoa integrante da cadeia de desenvolvimento da tecnologia.

Atualmente, vivemos em um momento da responsabilidade civil descrito por Schreiber (2009, p. 11-12) como a “erosão dos filtros tradicionais da reparação”, ou seja, encontramos-nos em um estágio no qual os requisitos para a configuração da responsabilidade civil, enquanto obstáculos à reparação de danos, vêm sendo mitigados. É o que também afirma André Luiz Arnt Ramos (2015), ao constatar que, entre aqueles elementos tradicionalmente exigidos para a responsabilização, nenhum deles restou integralmente preservado.

O que se pretende apresentar a seguir é que a disciplina da responsabilidade civil contemporânea, forte em uma perspectiva constitucionalizada e pautada pela unidade e completude do ordenamento jurídico, proporciona um arcabouço teórico apto a contemplar os danos ora cogitados, o que se verifica pela maneira com que a doutrina e a jurisprudência têm se debruçado sobre a culpa, o nexo causal e o dano.

2.2.2.1. A culpa e o risco no regime de responsabilidade

Como se viu, o regime de responsabilidade civil tradicional baseava-se na culpa, exigindo-se, para o acolhimento de uma pretensão reparatória, que a vítima lograsse em demonstrar que o ofensor teria agido de forma ilícita para causar o dano. O contexto de industrialização e desenvolvimento tecnológico que se firmou nos séculos XIX e XX exacerbou a injustiça de se atribuir à vítima o ônus de comprovar o elemento subjetivo do agente causador do prejuízo, já que a experiência do dano deixou de ser considerada excepcional e passou a ser vista como um desdobrar inevitável da vida em coletividade, pelo que, ao lado da culpa, o desenvolvimento de atividades de risco passou a ser um elemento configurador do dever de reparar, sob o regime de responsabilidade objetiva.

Importa ressaltar que a culpa não foi definitivamente afastada para a configuração do dever de indenizar, mas apenas se estabeleceu, na teoria geral da responsabilidade civil, uma relevância equivalente entre a responsabilidade subjetiva e a responsabilidade objetiva com fundamento no risco da atividade²⁴, deixando esta última de ser uma hipótese excepcional (MULHOLLAND, 2010). Assim sendo, não se exclui, *a priori*, a hipótese de que danos causados pela inteligência artificial possam ser inseridos na lógica da responsabilidade subjetiva. Aliás, no que se refere às atividades que fazem uso de novas tecnologias, como é o caso da inteligência artificial, João Quinelato de Queiroz (2020) defende cautela quanto à aplicação generalizada

²⁴ Por outro lado, embora a responsabilidade subjetiva e a responsabilidade objetiva correspondam a conceitos de justiça distintos – retributiva e distributiva, respectivamente -, é possível vislumbrar uma reunificação do sistema mediante a convergência das noções de culpa e de risco. Isso porque, cada vez mais, a culpa se faz objetiva, estando configurada quando houver o descumprimento de um dever de cuidado, e, por sua vez, a aferição da responsabilidade objetiva pelos tribunais passa a levar em consideração o comportamento do agente ofensor, por exemplo, para que se estabeleça o valor indenizatório (MORAES, 2006a). Maria Celina Bodin de Moraes (2006a, p. 21) detalha: “A manutenção da dicotomia culpa-risco torna-se ainda mais complexa quando se pensa que, do ponto de vista teórico, vacilam a doutrina e a jurisprudência quanto aos fundamentos seja da culpa, seja do risco; a propósito, chega-se mesmo a afirmar que a noção de risco/perigo desempenha um papel importante em ambos os critérios de imputação: assim, quanto mais perigosa (rectius, quanto maior o risco) é uma situação, mais cuidado e diligência devem ser adotados para prevenir os danos. A consequência disto é que em algumas hipóteses será impossível separar as duas noções e tanto faz atribuir a responsabilidade por um ou outro fundamento”. Caitlin Mulholland (2010, p. 46) também ressalta: “a análise do risco pela jurisprudência é realizada cada vez mais levando em consideração aspectos da culpa na atividade desenvolvida, seja através da chamada culpa normativa, seja através da imputação de uma reprovação de conduta realizada, o que, em última instância nada mais é do que a verificação de culpa”.

do regime objetivo, sob pena de se abandonar definitivamente o instituto da culpa na responsabilidade civil.

Ao analisar o conteúdo clássico da culpa, Caitlin Mulholland (2010) verifica que, para fins de responsabilização, era necessário que o agente incorresse em uma falha no dever de diligência (culpa *lato sensu*), tivesse capacidade de reconhecer a situação danosa (imputabilidade) e estivesse em situação de contrariedade ao Direito (lesão a terceiros). A autora detalha:

Considerava-se essencial pelos codificadores, embebidos na ideologia liberal-individualista, que a noção de culpa estivesse sempre associada a um comportamento objetivamente ilícito – porque contrário à lei – e ao mesmo tempo moralmente reprovável, no sentido de ser possível ao autor de dano ter a capacidade de reconhecimento de suas ações e de suas potenciais consequências (MULHOLLAND, 2010).

Atualmente, porém, há uma tendência de se desconsiderar a imputabilidade, ou seja, “o elemento subjetivo de possibilidade abstrata de reconhecimento social do ilícito ou do reprovável” para fins de reconhecimento da conduta culposa. Por conseguinte, aquela conotação tradicional que se conferia à culpa, qual seja, a da culpa psicológica baseada na manifestação de vontade do agente, não mais prevalece, suscitando-se a ideia de uma culpa objetiva, caracterizada quando infringida a obrigação geral de não lesar a outrem (MULHOLLAND, 2010, p. 41-42).

Assim, a responsabilidade subjetiva foi, à sua maneira, objetivada, mediante a formulação de uma concepção normativa da culpa, pela qual se passou a valorar o ato ilícito a partir de padrões de comportamento – *standards* de conduta - tidos como razoáveis para determinado caso concreto e não a partir da aferição da intencionalidade do agente (TEPEDINO, 2017). A culpa passa a configurar-se, dessa forma, pela inobservância de um dever de cuidado (MORAES, 2006a), ou, dito de outra forma, pelo desvio do padrão de conduta esperado de quem age de boa-fé e de forma diligente (MULHOLLAND, 2010).

Ao cogitar hipóteses em que danos causados por uma inteligência artificial atrairiam o regime subjetivo de responsabilidade, Tepedino e Silva (2020) lembram que o sistema de responsabilidade civil consiste na imputação do dever de indenizar a pessoas e não a robôs. Assim, ausentes os pressupostos da responsabilidade

objetiva²⁵, para fins de reconhecimento do dever de indenizar, ainda que o comportamento da inteligência artificial tenha relação de causalidade com a produção do dano, deve-se aferir a conformidade da conduta da pessoa que tinha o robô sob sua utilização²⁶ e não a conformidade da conduta do sistema autônomo, desprovido de personalidade jurídica²⁷.

No caso, qual deveria ser o *standard* para aferição da conduta culposa do agente que causou dano a um terceiro mediante o uso de um sistema autônomo?

Vale lembrar que, por muito tempo, o *standard* de conduta utilizado referia-se à ideia abstrata do “homem médio”, ou seja, o que se convencionou como o comportamento esperado de uma pessoa medianamente diligente. A pretensão de universalidade de tal critério, porém, foi desacreditada pela constatação de que cada magistrado conferia ao conceito abstrato de diligência mediana parâmetros individuais e de acordo com referências socioculturais próprias (MULHOLLAND, 2010).

Assim, os tribunais têm recorrido a “modelos múltiplos e menos generalizados de comportamento”, fazendo uma análise fragmentada das condutas submetidas à sua apreciação, mediante a busca de conhecimento específico sobre cada tipo de comportamento observado no caso concreto, o que proporciona uma “orientação mais pluralista, e mais efetiva sob o ponto de vista da realidade prática” (SCHREIBER, 2009, p. 41-43).

Portanto, no caso da inteligência artificial, considerando o seus múltiplos usos para finalidades variadas – por exemplo, no campo médico, automobilístico, educacional, doméstico etc. -, a análise precisaria fragmentar-se dentro desse campo de conhecimento, pois cada destinação dada a um sistema inteligente demandaria um padrão de diligência próprio. Ademais, esse *standard* também deve se particularizar

²⁵ Por exemplo, no caso da reciprocidade de riscos. Conforme descrito por Laura Osório Bradley dos Santos Dias (2020, p. 638), no caso de acidentes causados por veículos autônomos: “Haveria de se considerar, também, a possibilidade de aplicação da responsabilidade subjetiva ao proprietário se o dano for causado a outro veículo autônomo, tendo em vista a reciprocidade dos riscos (...)”.

²⁶ Os autores ressaltam que o grau de intervenção do usuário poderia ter relevância para a sua responsabilização, havendo maior ou menor carga de responsabilidade conforme o agente tiver maior ou menor ingerência sobre a máquina, o que dependeria da análise de cada sistema individualmente considerado (2020, p. 314).

²⁷ A atribuição de personalidade jurídica a robôs não faz parte do escopo desta dissertação, porém, cabe consignar o entendimento de Eduardo Nunes de Souza (2020, p. 236), para quem “a atribuição de personalidade jurídica a entes diversos da pessoa natural tem servido precipuamente para permitir uma limitação, em maior ou menor grau, da responsabilidade dos indivíduos envolvidos nessas operações”. Complementa o autor que “a reparação às vítimas seria muito melhor garantida pela responsabilização do agente que colocou a tecnologia em operação ou dela se beneficiou economicamente”.

conforme o suposto ofensor seja um usuário da tecnologia ou um profissional do campo tecnológico.

Nesse contexto de culpa normativa e de ampliação do uso de tecnologias de IA, deve ser destacado ainda que a necessidade de se dar mais atenção ao componente humano traz em si também o surgimento dos códigos e normas profissionais como um guia importante para avaliar a conduta danosa e o comportamento que seria esperado. Códigos profissionais sempre desempenharam um papel importante na aferição de culpa normativa e a adoção de aplicações inteligentes apenas introduz novas considerações para sua aplicação (SOUZA; PADRÃO, 2020, p. 675).

Embora se cogite da aplicação do regime de responsabilidade subjetiva no âmbito da inteligência artificial, é preponderante na doutrina o entendimento mais favorável ao acionamento da responsabilidade objetiva, por ser mais idônea à efetiva tutela às vítimas de dano (TEPEDINO; SILVA, 2020). Isso porque não há como isolar a inteligência artificial do processo de solidarização pelo qual passou o sistema de responsabilidade civil, sendo certo que, reconhecendo-se a culpa como um obstáculo para a proteção da vítima de danos, dever-se-á cogitar da aplicação da responsabilidade objetiva ou, segundo Queiroz (2020, p. 592), da “busca no ordenamento de alternativas que não compactuem com a culpa a apresentar-se como empecilho insuperável à indenização de danos decorrentes de sistemas de IA”.

Portanto, no que se refere à inteligência artificial, atribuir-se à vítima o ônus probatório da culpa do agente causador do dano, sob um regime de responsabilidade civil subjetiva, pode representar um retorno dos empecilhos à reparação, eis que, ao usuário de sistemas dotados de inteligência, seria tortuoso demonstrar a ilicitude da conduta, ou seja, a inobservância de um dever de cuidado de algum integrante da cadeia de desenvolvimento do *software*, de fabricação do *hardware* (QUEIROZ, 2020), ou até mesmo de um usuário ou responsável pelo funcionamento da máquina.

Vive-se muito além das revoluções industriais, que levaram cerca de 20 ou 30 anos para marcar definitivamente a história da humanidade. Vive-se um fenômeno inédito e revolucionário: a tecnologia, de mês em mês – e não de século em século – inova assustadoramente e transforma radicalmente o modo de viver-se em sociedade. É nesse contexto que a culpa, para os objetivistas, paulatinamente vai assumindo papel de quase coadjuvante ante às suas dificuldades probatórias (...) (QUEIROZ, 2020, p. 590).

No que se refere aos mecanismos implementados com a finalidade de evitar o rigor da exigência da prova da culpa, foram dedicadas muitas menções à teoria do risco, positivada, no parágrafo único do artigo 927 do Código Civil, mediante a

estipulação de cláusula geral de responsabilidade objetiva pelo desempenho de atividades potencialmente arriscadas. Embora não seja a única hipótese de responsabilidade objetiva da legislação brasileira, a imputação do dever de indenizar com fundamento no risco da atividade normalmente desempenhada afastou a primazia da culpa no ordenamento jurídico (SCHREIBER, 2009).

Se antes, a obrigação de indenizar um dano tinha como principal – e senão, único - fundamento a culpa daquele que lesiona o direito subjetivo de outrem, na atualidade ela deve ser necessariamente reinterpretada no sentido de considerar-se também responsável aquele que proporciona um risco de dano, calcando-se esta responsabilidade no princípio da solidariedade social que deve estar presente em todas as relações sociais (MULHOLLAND, 2010, p. 20).

A definição e o alcance da cláusula geral fundamentada no risco da atividade, porém, ficou a cargo da doutrina e da jurisprudência²⁸, searas nas quais foram elaboradas concepções e teorias diversas, que, por exemplo, ora levam em consideração, para fins de responsabilização, o agente que extraia efetivo proveito da atividade (teoria do risco-proveito), ora contemplam aquele que exerce toda e qualquer atividade que importe em risco a terceiros, lucrativa ou não (teoria do risco criado). Há, ainda, teorias que consideram apenas as atividades que apresentam risco exacerbado, como é o caso da exploração de energia nuclear (teoria do risco excepcional) e aquelas que introduzem uma classificação relativa à extensão da responsabilidade, conforme se admitam ou não excludentes como caso fortuito ou força maior (teoria do risco integral ou do risco mitigado) (MORAES, 2006a; SCHREIBER, 2009). Destaca-se, ainda, a teoria do risco do desenvolvimento, que traduz a responsabilidade do empreendedor por produtos inseridos no mercado e que, em razão do avanço técnico e científico, oferecem riscos que vêm a ser conhecidos apenas após um período de uso pelos consumidores²⁹ (QUEIROZ, 2020).

²⁸ Anderson Schreiber (2009, p. 3) critica que o “diálogo entre doutrina e jurisprudência lembra, nesta matéria, uma babel de idiomas desconexos, em que não se chega a qualquer resposta por total desacordo sobre as perguntas”.

²⁹ Embora o risco do desenvolvimento possa ser arguido como causa de exclusão da responsabilidade (MONTEIRO FILHO; ROSELVALD, 2020), o Enunciado nº 43 da I Jornada de Direito Civil estabelece que a “responsabilidade civil pelo fato do produto, prevista no art. 931 do novo Código Civil, também inclui os riscos do desenvolvimento”. Disponível em <https://www.cjf.jus.br/enunciados/enunciado/720>. Acesso em 13 ago. 2021.

Aventa-se, em grande medida³⁰, a incidência da responsabilidade civil objetiva às hipóteses de danos causados pela IA, com fundamento no risco da atividade³¹, considerando que “são poucas as atividades empresariais que não estão atreladas a algum risco inerente à sua atividade, o que igualmente se observa nas atividades que empregam sistemas de inteligência artificial” (QUEIROZ, 2020, p. 589). Nesse sentido, pode-se argumentar que a expectativa de desenvolvimento da inteligência artificial traz consigo muitas incertezas, o que torna a atividade de seus desenvolvedores e programadores potencialmente perigosa e apta a receber um tratamento em consonância com o risco proporcionado, uma vez ser possível a ocorrência de danos mesmo com a observância de todas as precauções de segurança (ČERKA; GRIGIENĖ; SIRBIKYTĖ, 2015).

Indaga-se, porém, se o emprego da inteligência artificial acarretaria, de forma generalizada, a configuração de uma atividade de risco para fins de incidência do parágrafo único do artigo 927 do Código Civil.

De fato, a doutrina ainda não logrou em definir com precisão o que seria uma atividade de risco para fins de aplicação da cláusula geral, o que gera ainda mais inquietação quando se trata do uso uma tecnologia relativamente nova, cujas implicações ainda não são completamente conhecidas (QUEIROZ, 2020).

Para Queiroz (2020), as teorias que buscaram delimitar a aplicação da cláusula geral não são suficientes para inserir, de forma generalizada, a inteligência artificial no âmbito da responsabilidade objetiva com fundamento no risco. Segundo o autor, seria necessária uma investigação casuística de cada sistema de IA visando a aferição do grau de risco da atividade em concreto. Por outro lado, reconhece que, ao cidadão comum, escapa a capacidade de provar eventual imprudência ou imperícia de desenvolvedores de sistemas inteligentes, sendo nesse contexto que o intérprete poderia aplicar o regime de responsabilidade objetiva com fundamento na noção genérica de risco da atividade.

³⁰ Ainda, poder-se-ia submeter a inteligência artificial à disciplina da responsabilidade pela guarda de coisa ou de animal, já 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” (TEPEDIDO; SILVA, 2020, p. 317).

³¹ Registre-se, porém, a crítica de Hartmann Peixoto e Silva (2019) quanto à abordagem regulatória da IA baseada no risco. Os autores argumentam que este debate tem preponderado sobre propostas de enfrentamento das inconsistências apresentadas pela tecnologia e negligenciado os benefícios observados pelo seu uso.

David C. Vladeck (2014), direciona a discussão para o caso específico dos carros autônomos, afirmando que a teoria do risco não se aplicaria para tais veículos, já que, neste caso, a inteligência artificial atribuiria mais segurança às interações no trânsito, reduzindo a preponderância de fatores humanos que favorecem a ocorrência de acidentes com frequência. Nesse sentido, o autor sustenta que um sistema que permita a responsabilidade objetiva de fabricantes e desenvolvedores não pode ter como pressuposto o risco causado pelos carros autônomos, mas sim, a adoção de um regime que possibilite que uma pessoa que sofreu um dano experimente algum tipo de reparação, ainda que a causa do ilícito seja desconhecida. Também, o autor considera que os fabricantes de veículos têm mais condições de absorver os custos da compensação do dano, o que, aliás, pode ser mais vantajoso do que ter que defender-se em juízo, mediante o pagamento de advogados e custas processuais. Ademais, uma política de responsabilidade previsível e estável, na qual não haja uma busca inalcançável pela atribuição de culpa, poderia vir a favorecer a inovação consciente, ao invés de travar o desenvolvimento da tecnologia.

Monteiro Filho e Rosenvald (2020) ressaltam que o sistema de responsabilidade civil brasileiro conta com uma cláusula geral de risco e não de periculosidade. Assim, consideram que a obrigação de indenizar surge quando o desempenho de uma atividade lícita, autorizada e regulamentada, por sua natureza, pode provocar danos numerosos e/ou graves. Portanto, há modelos de inteligência artificial que não são intrinsecamente perigosos, porém, guardam especial aptidão para gerar lesões, notadamente, de ordem extrapatrimonial.

Destacam os autores que a legislação europeia³² recorre ao conceito de alto risco da atividade – e não apenas de risco -, considerando que o emprego de novas tecnologias, em algumas circunstâncias, pode vir a instrumentalizar a pessoa humana e a violar direitos fundamentais, como liberdade e igualdade, promovendo discriminações. Nesse sentido, não se admite apenas uma maior probabilidade de o dano acontecer, mas também se leva em consideração o fato de que o potencial dano tem natureza grave. Concluem, assim, que no campo da inteligência artificial o alto risco

³² Referência ao artigo 35 do Regulamento 2016/79 da União Europeia (Regulamento Geral de Proteção de Dados).

se evidencia pelo risco intensificado da atividade para liberdades fundamentais das pessoas naturais, resultado do processamento de dados pessoais que acarrete danos físicos, patrimoniais ou extrapatrimoniais, em especial, quando o processamento cause discriminação, furto de identidade, fraudes diversas, perdas financeiras, dano reputacional, perda de confidencialidade dos dados pessoais protegidos por sigilo profissional, reversão não autorizada da pseudonimização ou qualquer outra desvantagem econômica ou social significativa. Acresçam-se a isso as circunstâncias nas quais titulares dos dados pessoais sejam privados de seus direitos ou impedidos de exercer controle sobre eles; em que são processados dados pessoais que desnudem a intimidade, notadamente de pessoas vulneráveis (2020, p. 560).

Caso a máquina dotada de inteligência artificial tenha sido vendida a um destinatário final como um produto, aventa-se a aplicabilidade, na hipótese, do Código de Defesa do Consumidor³³, o qual prevê responsabilidade objetiva de toda a cadeia de produção e fornecimento pela reparação de danos causados por defeitos no produto ou relativos à prestação do serviço (TEPEDINO; SILVA, 2020).

De início, Monteiro Filho e Rosenvald (2020, p. 555) rejeitam a aplicação da teoria do risco do desenvolvimento enquanto forma de evitar a responsabilidade por defeitos imprevisíveis, ao argumento de que tal tese defensiva “não deve estar disponível nos casos em que era previsível a ocorrência de desenvolvimentos imprevistos”. Quanto ao ponto, os autores detalham que as

características das tecnologias digitais emergentes, como opacidade, abertura, autonomia e previsibilidade limitada, geralmente acarretam dificuldades desproporcionais ou custos elevados para a vítima, frustrando a sua legítima expectativa quanto ao nível de segurança que um usuário médio poderia esperar (2020, p. 555).

Todavia, a questão se complexifica quando se trata de máquinas sem nenhum defeito mecânico ou de programação detectado, mas que causam danos que podem ser atribuídos à capacidade de aprendizado do *software*. Pertinente, portanto, indagar acerca da resposta jurídica para os casos em que o sistema inteligente vem a causar danos no exercício regular da tarefa para o qual fora programado, ou seja, sem que o prejuízo causado decorra de uma falha de segurança, funcionando em conformidade com o que se espera de uma utilização razoável. Ademais, o defeito pode ser um

³³ Diferenciando a hipótese da responsabilidade objetiva do Código de Defesa do Consumidor daquela prevista na cláusula geral do parágrafo único do artigo 927 do Código Civil, Schreiber (2009, p. 25) adverte que “o fundamento de tutela, aqui e ali, são inteiramente diversos. O escopo do parágrafo único do art. 927 é o de impor responsabilização com base no elevado risco produzido por certa atividade, o que não se verifica em qualquer espécie de prestação de serviços, mas apenas naquelas hipóteses em que houver uma alta possibilidade de dano”.

desdobramento de uma escolha de programação, como, por exemplo, no caso de carros autônomos, em que o veículo é programado para escolher, entre dois bens jurídicos em perigo, qual deles sacrificar (MONTEIRO FILHO; ROSENVALD, 2020; KIM, 2018).

Nessas hipóteses, que, em tese, escapariam ao escopo do Código de Defesa do Consumidor ante a ausência de fato do produto ou do serviço, os autores remetem a resposta da questão, também, à cláusula geral do risco da atividade, ao argumento de que “a sofisticação do desenvolvimento tecnológico e, em especial, a inteligência artificial convertem o ônus da prova em exigência acrescida, de demonstração muito difícil” (MONTEIRO FILHO; ROSENVALD, 2020, p. 556).

Verifica-se, portanto, que, embora não se possa caracterizar toda e qualquer atividade que faça uso da inteligência artificial como sendo arriscada, suscitar o risco nesse âmbito diz com a necessidade de se adotar um fundamento de imputação que viabilize a reparação às vítimas de danos, assim como afirmam Tepedino e Silva (2020, p. 320):

Como se percebe, o reconhecimento da configuração de atividades de risco a partir do emprego generalizado de sistemas de inteligência artificial parece a solução adequada, em linha de princípio, para o equacionamento da questão atinente à individualização do critério de imputação do regime de responsabilidade. O que não parece possível, ao revés, é a invocação indiscriminada e irrefletida da noção de atividade de risco. Deve-se, com efeito, lançar mão dos critérios desenvolvidos pela doutrina para a elucidação do que vem a ser atividade de risco para fins de incidência da correlata cláusula geral de responsabilidade objetiva. Há que se investigar detidamente, em cada atividade, à luz das especificidades dos respectivos sistemas e de seu contexto, a possibilidade de caracterização de atividade de risco.

Relativizada a prova da culpa e firmada a incidência da teoria do risco, para fins de oportunizar a reparação dos danos causados pela inteligência artificial, continuamos a percorrer os elementos configuradores da responsabilidade civil e passamos à análise do nexos causal.

2.2.2.2. *O nexos causal*

Monteiro Filho e Rosenvald (2020) apontam que os principais desafios impostos à disciplina da responsabilidade civil pelas tecnologias digitais emergentes,

entre as quais está incluída a inteligência artificial, referem-se aos seguintes atributos: complexidade, ante a pluralidade de atores envolvidos na fabricação e o alto grau de sofisticação técnica do *hardware*; opacidade, pela dificuldade de identificação dos processos causadores de danos; abertura, eis que sofrem contribuições e atualizações após serem colocadas em circulação; autonomia, pela capacidade de executar tarefas sem supervisão humana; vulnerabilidade, pois estão sujeitas a violações de segurança cibernética; e imprevisibilidade, pois são programadas para responderem a novos estímulos e não apenas àqueles predefinidos pelos desenvolvedores. Todas as características aqui declinadas denotam a dificuldade de se manter um liame causal entre uma pessoa envolvida na cadeia de desenvolvimento da tecnologia e eventual dano decorrente do seu funcionamento.

O nexos causal pode ser definido como o liame jurídico³⁴ que se estabelece entre uma conduta e um dano, com a finalidade de imputação do dever indenizatório, sendo certo que, se não for estabelecida esta ligação entre agente e dano, não estará configurada a obrigação de indenizar (MULHOLLAND, 2010).

Embora a aferição do nexos causal tenha adquirido evidência, diante do arrefecimento da culpa no sistema de responsabilidade civil, também este elemento vem deixando de atuar como uma barreira intransponível à reparação, sendo-lhe atribuído pelo Poder Judiciário certa elasticidade, de forma a assegurar a efetiva reparação às vítimas de danos, aliada, porém, a um baixo rigor na aferição deste requisito (SCHREIBER, 2009; RAMOS, 2015). Nesse sentido:

A dificuldade contemporânea em relação à pesquisa da causalidade diz respeito a sua possível desconsideração ou afastamento, muito mais do que o de sua investigação. Percebe-se, aos poucos e através da pesquisa jurisprudencial, que o nexos de causalidade é relegado a um segundo plano, perdendo importância a sua determinação frente à investigação do princípio da culpa ou do princípio do risco (MULHOLLAND, 2010, p. 61).

A questão complexifica-se com a multiplicidade de teorias que buscam explicar as relações de causalidade, com ligeiras variações entre si, o que proporciona um “ecletismo judicial revelador da preferência pela pretensa realização da justiça no caso

³⁴ É relevante notar que, embora o nexos de causalidade seja, geralmente, considerado como um elemento de fato (causa e efeito), essa é apenas a etapa inicial da investigação. Caitlin Mulholland (2010, p. 99) leciona que a aferição da causalidade passa também por uma segunda etapa, denominada de causalidade jurídica, cuja função é indicar as consequências danosas e os limites da indenização, ou seja, “busca-se o estabelecimento de uma ligação entre o fato ilícito ou atividade perigosa e a consequência danosa que deles resultou”.

concreto, mesmo em detrimento da previsibilidade dos resultados” (RAMOS, 2015, p. 16). Sem consenso doutrinário significativo em torno da matéria, os tribunais vêm aplicando teorias diversas, sem que se possa detectar um padrão nas decisões judiciais que deferem pedidos de reparação por danos (SCHREIBER, 2009).

Entre as teorias existentes, Schreiber (2009) e Tepedino (2017) destacam (i) a teoria da equivalência das condições, que considera como causa, para fins de responsabilização, todo e qualquer evento que tenha contribuído para a produção do dano; (ii) a teoria da causalidade adequada, que busca identificar, entre as possíveis causas, qual delas, em abstrato, é a mais apta a produzir o efeito danoso; (iii) a teoria da causalidade eficiente, que busca eleger, no caso concreto, o antecedente que, em razão de um poder qualitativo ou quantitativo, pode ser considerado como causa do evento e (iv) a teoria da causa direta e imediata, que considera como causa apenas os eventos vinculados direta e imediatamente ao prejuízo, ressalvados os danos que, embora indiretos e remotos, sejam consequência necessária da conduta tomada como causa (subteoria da necessariedade causal).

Paralelamente, observa-se um rigor no reconhecimento das excludentes de causalidade – caso fortuito, força maior, culpa exclusiva da vítima e fato de terceiro -, acolhendo-se a noção de fortuito interno, desenvolvida no campo das relações consumeristas, de forma a evitar a exclusão da responsabilidade do fornecedor por eventos que, apesar de serem imprevisíveis e inevitáveis, podiam ser constatados antes da disponibilização do produto no mercado. Trata-se de instituto voltado à imputação de responsabilidade a sujeitos que exercem atividades potencialmente lesivas, com fundamento no risco que oferecem para os direitos de terceiros (SCHREIBER, 2009).

Na mesma linha, a culpa exclusiva da vítima e o fato de terceiro têm sido interpretados de forma restritiva pelo Poder Judiciário, deixando de interromper a causalidade em diversas situações em que o dano, embora seja resultado de alguma ação incompatível com padrões de diligência por parte da vítima ou de terceiros, pode ser incluído no âmbito do risco da atividade desempenhada pelo responsável (SCHREIBER, 2009).

Assim como ocorreu com a mitigação da prova da culpa, o contexto remete a situações em que há uma dificuldade da vítima em demonstrar precisamente a conexão entre o dano sofrido e a ação ou omissão de um agente determinado

(MULHOLLAND, 2010), o que condiz com eventuais danos causados por uma inteligência artificial, considerando a capacidade de aprendizado e a autonomia com que sistemas inteligentes operam, bem como a ampla e difusa cadeia de desenvolvedores envolvidos na fabricação e programação da máquina, o que significa que a investigação da causalidade pode levar a múltiplas respostas ou a resposta nenhuma.

Por outro lado, embora tais expedientes busquem legitimar-se pelo intuito de tutelar as vítimas de danos, é certo que devem ser vistos com cautela, ante a potencialidade de estimular a multiplicação de demandas por indenização e de fomentar uma cultura de vitimização social (*blame culture*)³⁵ (SCHREIBER, 2009).

No intuito de estipular parâmetros de interpretação para a flexibilização da demonstração do nexos causal, Caitlin Mulholland (2010) suscita o princípio constitucional da dignidade da pessoa humana, que, aplicado a uma situação de impossibilidade de prova da causalidade, sustenta a obrigação de indenizar como forma de priorização da vítima do dano. Ainda, a autora aponta que o princípio constitucional da solidariedade social, no que se refere à delimitação do nexos causal, densifica-se pelo objetivo de distribuir adequadamente os riscos da atividade, impondo o dever de reparar àquele que teria a melhor condição de fazê-lo.

Sob essas premissas, em hipóteses nas quais seja dificultada a demonstração do nexos causal, seja pela hipossuficiência probatória da vítima ou pela multiplicidade de causas potenciais do dano, a autora invoca instrumento da presunção de causalidade, pelo qual se extrai da experiência o que ocorreria em situações semelhantes, mediante um juízo de probabilidade. Nesse sentido, o nexos de causalidade passa a ser aferido de acordo com a probabilidade e potencialidade da ligação entre o dano e a atividade em escrutínio (MULHOLLAND, 2010).

Mulholland (2010) aponta para a necessidade de uma fundamentação jurídica adequada para a aferição da existência do nexos de causalidade, pelo que propõe a sistematização das hipóteses em que o instrumento da presunção, enquanto meio de prova, se justifica no âmbito da responsabilidade civil. De início, afasta a ideia do nexos

³⁵ “Inspirada na solidariedade social, a jurisprudência tem, por toda parte, se recusado a deixar as vítimas dos danos sem reparação. As cortes flexibilizam, deste modo, a configuração dos pressupostos da responsabilização, no solidário reconhecimento de que, de alguma forma, por vivermos em sociedade, todos somos culpados por todos os danos e todos somos causadores de todos os danos. Este itinerário, todavia, se conclui, atualmente, com o lançamento do dever de indenizar sobre um único agente, despertando sentimentos de injustiça e suscitando perplexidade quanto ao papel da responsabilidade civil” (SCHREIBER, 2009, p. 247).

de causalidade enquanto mero elemento de fato, acrescentando um elemento jurídico à investigação:

Diferentemente da concepção condicionalística que estabelece que todas as condições de um dano são sua causa (ligação natural de causa e efeito); a concepção probabilística da responsabilidade civil é baseada na ideia de que a análise jurídica da imputação deve revelar-se por meio da classificação da atividade ou conduta como sendo estatisticamente e tipicamente associada ao dano sofrido, levando-se em consideração nesta análise não só o que é observado em situações semelhantes, ordinariamente, mas também o que pode legitimamente esperar-se da execução da atividade (MULHOLLAND, 2010, p. 289).

Para a autora, a teoria da causalidade adequada seria a que melhor acomodaria a aferição probabilística do evento danoso, considerando a aptidão para identificar, em abstrato, situações que se desenrolariam de forma semelhante àquela que se analisa concretamente, podendo ser associada à noção de repartição de riscos, de forma a situar a obrigação de reparar na própria realização de uma atividade considerada arriscada, analisada de acordo com a relevância estatística da ocorrência do dano (MULHOLLAND, 2010).

Por exemplo:

Se existe um acidente aéreo com morte de passageiros, por exemplo, a descrição do evento deve ser realizada a partir da identificação da condição *sine qua non* do dano, neste caso, a queda do avião. É importante para a técnica estatística esta descrição adequada, pois é ela que indicará a frequência percentual relevante para o estabelecimento da obrigação indenizatória. Se a descrição deste evento é realizada tendo como base a prestação de serviços de transporte aéreo (e não a queda do avião), a conclusão que se terá é a seguinte: não há responsabilidade por parte da transportadora porque são raríssimos os casos de morte de passageiros em acidentes aéreos, sendo incomuns as quedas de avião, e não existindo, por isso, a regularidade entre atividade e dano. Mas, se a descrição do evento se dá a partir da queda do avião (*conditio sine qua non* dos danos sofridos) percebe-se que, estatisticamente, quando há um acidente aéreo desta natureza, os danos serão quase sempre estes: a morte de passageiros. O percentual de passageiros mortos em queda de aviões é infinitamente superior ao de passageiros sobreviventes num mesmo evento. Estatisticamente, portanto, existe uma incidência relevante que, por consequência, identifica o nexo de causalidade (MULHOLLAND, 2010, p. 301-302).

A flexibilização do exame da causalidade pelo uso da técnica da probabilidade estatística, segundo Mulholland (2010), seria cabível quando o fundamento para a responsabilização se referir à noção de risco (responsabilidade objetiva), ante a possibilidade de se identificar a probabilidade estatística da ocorrência de danos associados ao exercício da atividade. Além disso, caberia invocar a presunção de

causalidade nas hipóteses de danos difusos, ou seja, aqueles que atingem uma coletividade de pessoas ou a sociedade como um todo e, como tais, caracterizam-se pela multiplicidade de causas, pela variedade e indeterminação das vítimas e pela violação a interesses que se disseminam no tempo e no espaço, o que confere complexidade à demonstração do liame causal³⁶. Em ambos os casos, substitui-se a certeza pelo critério da probabilidade, no intuito de possibilitar a reparação.

Transpondo as hipóteses de presunção de causalidade acima mencionadas para o campo tecnológico, retoma-se a ideia de que a inteligência artificial pode ser incluída no âmbito das atividades de risco, pois, embora nem sempre se possa observar um incremento na quantidade de acidentes pelo uso da IA, há situações em que a gravidade abstrata do dano em potencial justifica um tratamento jurídico à altura, o que se assemelha ao exemplo da queda de avião apresentado pela autora. No caso, quando dificultada a demonstração do nexos causal, seja pela hipossuficiência probatória da vítima e/ou pela multiplicidade de causas atribuíveis ao resultado, um exame probabilístico pode ser acionado para que se estabeleça o liame entre o desempenho da atividade arriscada e o dano sofrido.

Por exemplo, no caso dos carros autônomos, embora se espere, com a sua difusão, uma diminuição da quantidade total de acidentes, há que se aventar a possibilidade de que pessoas venham a ser lesionadas em situações que jamais ocorreriam se não estivessem fazendo uso do veículo dotado de inteligência artificial (TEPEDINO; SILVA, 2020).

Ademais, quanto aos danos difusos, verifica-se que a lógica do uso da inteligência artificial, pela qual se processam enormes quantidades de dados e se automatizam decisões que, por vezes, impactam na fruição de direitos e garantias fundamentais, podem vulnerabilizar grupos inteiros de usuários, aos quais nem sempre é dado negociar sobre quais dados pessoais poderão ou não ser processados ou qual parte da sua vivência será automatizada.

Por exemplo, Mulholland e Kremer (2020) abordam os danos causados pela violação do princípio da igualdade no tratamento de dados pessoais, constatando que o fluxo de dados coletados massivamente e processados por ferramentas tecnológicas pode reproduzir discriminações observadas na sociedade, como se tem verificado insistentemente no âmbito do reconhecimento facial e no uso de algoritmos

³⁶ A autora menciona exemplos relacionados, principalmente, aos danos ambientais.

preditivos. As autoras ilustram o problema com o caso em que uma câmera fotográfica que apresentava a funcionalidade de evitar fotografias com olhos fechados falhou em reconhecer a fisionomia de pessoas asiáticas. Apresentam também o evento em que um aplicativo de reconhecimento facial e marcação automática de fotos rotulou homens negros como gorilas. Ainda, descrevem as ocorrências de “falsos positivos” pelo uso de reconhecimento pessoal no campo do policiamento, o que evidenciou vieses de raça e classe social.

Verifica-se, portanto, que a capacidade de tratamento de dados pessoais das mais diversas ordens vem aumentando exponencialmente, principalmente devido ao advento de tecnologias avançadas de inteligência artificial, com o uso de algoritmos sofisticados e com a possibilidade de aprendizado por máquinas (machine learning). Significa dizer que o tratamento de “big data” – literalmente, grandes bases de dados – por meio de técnicas computacionais cada vez mais desenvolvidas pode levar a análises probabilísticas e resultados que, ao mesmo tempo que atingem os interesses de uma parcela específica da população, retiram a capacidade de autonomia do indivíduo e o seu direito de acesso ao consumo de bens e serviços e a determinadas políticas públicas, por exemplo (MULHOLLAND; KREMER, 2020, p. 570).

Pertinente mencionar, ainda, como exemplo, o uso da inteligência artificial associado a grupos sociais em situação de vulnerabilidade, como ocorre no contexto humanitário. Jasmine Wright e Andrej Verity (2020) descrevem o teste de um *chatbot* usado para realizar sessões de psicoterapia em um campo de refugiados na Síria, apresentando preocupações com a capacidade de compreensão dos sujeitos envolvidos acerca da interação com uma IA – e não com uma pessoa de verdade -; com o nível de sigilo das conversas, em contraste com um possível monitoramento de governos ou organizações terroristas; e com a existência de um alerta tempestivo quanto à necessidade de intervenção de um profissional. Além disso, os autores trazem exemplo da coleta de dados biométricos de refugiados que buscam asilo em países da União Europeia, com a intenção de facilitar a identificação e localização dessas pessoas, o que pode acabar por deixá-las mais vulneráveis em relação à perseguição que ocasionou a fuga de seus países de origem.

Wright e Verity observam que a coleta indiscriminada de dados pessoais pode reduzir a capacidade de escolha desses indivíduos sobre quais informações desejam compartilhar e quais deveriam ser guardadas em sigilo, bem como dificulta a compreensão sobre quais decisões relevantes são automatizadas e como elas poderiam ser revertidas, o que pode lhes causar danos permanentes.

No campo consumerista, menciona-se o uso de sistemas de inteligência artificial para a formação de perfis de indivíduos e grupos mediante o uso de dados pessoais, para subsidiar práticas discriminatórias, como ocorre com o *geopricing* e o *geoblocking*, que consistem na diferenciação dos preços e na negativa de atendimento com base na localização geográfica do consumidor, prática observada em agências de viagens (MEDON, 2020).

Desta feita, recorrendo-se ao instrumento da presunção de causalidade apresentado por Mulholland, mitigam-se as dificuldades apontadas para a investigação do liame entre o dano causado por uma máquina autônoma e o titular da atividade que faz uso dessa tecnologia, o que pode ser um mecanismo ao alcance do intérprete para garantir a tutela da vítima e concretizar os princípios constitucionais que dão fundamento à responsabilidade civil contemporânea.

2.2.2.3. O dano

Sem dano – patrimonial ou extrapatrimonial -, não se cogita da existência do dever de indenizar, tratando-se, por esta razão, do elemento essencial da responsabilidade civil (MULHOLLAND, 2010).

Além de significar a violação de um interesse tutelado pelo ordenamento jurídico, a ocorrência do dano também remete à sua quantificação, enquanto consequência jurídica definidora da verba indenizatória. São esses os aspectos conceituais do dano (evento e resultado), aos quais é possível acrescentar um terceiro elemento, que “é qualificador do dano e representante de uma mudança de perspectiva ideológica – e metodológica – na responsabilidade civil contemporânea”, qual seja, o seu caráter injusto, ou de ilicitude objetiva (MULHOLLAND, 2010, p. 25).

Relativizados os demais requisitos configuradores de responsabilidade civil, confere-se evidência à ocorrência do dano enquanto o último filtro de contenção de demandas por reparação, mediante o acionamento de referenciais valorativos constitucionais que permitem identificar quais condutas ensejam o dever de reparar. Para tanto, a ideia de dano reparável separa-se da ilicitude do ato para fundar-se na lesão a um interesse posicionado, na ordem jurídica, em nível superior em relação

àquele que orienta a ação lesiva. (RAMOS, 2015). Dessa forma, o dano deixa de ser visto como “elemento estático e meramente pressuposto das ações de responsabilidade civil” e passa a ser entendido como “cláusula geral destinada à seleção dos interesses merecedores de tutela, cuja violação dá margem ao dever de indenizar” (SCHREIBER, 2009, p. 243).

Nesse sentido:

A admissão de uma cláusula geral de dano injusto representa, conseqüentemente, a aceitação de uma responsabilidade decorrente de uma conduta, ilícita (no sentido de culposa) ou lícita (não-culposa), que viola os princípios e valores constitucionais ou infraconstitucionais e infringe o direito à integridade psicofísica do ser humano, enquanto derivação lógica e jurídica do princípio da dignidade da pessoa humana (MULHOLLAND, 2010, p. 26).

Qualificar um dano como injusto, ou seja, atestar a sua desproporcionalidade em relação aos interesses juridicamente protegidos, no contexto solidarista que pauta a responsabilidade civil contemporânea, caracteriza a obrigação de indenizar (MULHOLLAND, 2020).

Ampliam-se, assim, as hipóteses de danos indenizáveis de forma quantitativa e qualitativa³⁷, eis que novos interesses – sobretudo existenciais e coletivos – passam a ser vislumbrados como dignos de proteção, levando à conseqüente expansão das demandas judiciais por ressarcimento. Schreiber³⁸ (2009), Mulholland³⁹ (2010) e Ramos⁴⁰ (2015) consideram, porém, que o desafio que se impõe aos juristas quanto à seleção dos interesses merecedores de tutela é justamente a definição dos métodos pelos quais se pretende delimitá-los, reconhecendo a preponderância da função judicial nesse mister, mas demonstrando preocupação com a aleatoriedade e o arbítrio com que a questão tem sido tratada dos tribunais.

³⁷ “Não sendo possível exauri-los [os novos danos], sua indicação tem como utilidade apenas a descrição ilustrativa da amplíssima expansão do dano ressarcível que vem chocando tribunais ao redor do mundo. Ao mesmo tempo, compreensões sistemáticas que poderiam, quiçá, auxiliar na apreensão do fenômeno vêm já desencorajadas pela imensa variabilidade das tantas figuras reunidas sob o signo comum da novidade” (SCHREIBER, 2009, p. 89).

³⁸ Para Schreiber (2009, p. 105), o “desafio que, hoje, se impõe aos juristas brasileiros é justamente o de definir os métodos de aferição deste merecimento de tutela, reconhecendo a importância da discricionariedade judicial na tarefa, mas sem deixá-la exclusivamente ao arbítrio dos tribunais”.

³⁹ Mulholland (2010, p. 31) afirma que o “grande perigo na exploração destes novos danos é justamente a sua falta de delimitação, na medida em que caberá ao magistrado a reflexão sobre a sua caracterização e o conseqüente estabelecimento de um valor a título de reparação”.

⁴⁰ A definição dos interesses dignos de tutela, porém, é descrita por Ramos (2015, p. 19), como “o grande desafio que se coloca aos contemporâneos sistemas de responsabilidade”, chamando atenção à postura paternalista do Poder Judiciário em relação às vítimas e à adoção de critérios de seleção dos danos ressarcíveis que careceriam de exame crítico.

A Teoria Geral da Responsabilidade Civil, em seu estado atual, portanto, seria mais bem denominada com a expressão Direito de Danos, pois, abandonando suas conotações tradicionais – individualistas e moralizantes – passou a ter maior afinidade com o estudo da obrigação de indenizar, ou seja, refere-se à aferição das hipóteses de danos que devem sofrer reparação, objetivando-se a análise em relação ao resultado (MULHOLLAND, 2010).

Assim explica Thaís Gouveia Pascoaloto Venturi (2012, p. 15):

As mudanças vivenciadas pela sociedade contemporânea em razão das relações de massa, pautadas no dinamismo, na impessoalidade e na efemeridade, revelaram-se especialmente nocivas quanto ao extraordinário incremento de danos causados à coletividade, passando a se afirmar que a expressão mais adequada para caracterizar o atual momento da regulação da responsabilidade civil seria “Direito de Danos”. Com efeito, diante da proliferação de novas modalidades de danos e a necessidade de aprimoramento da tutela da pessoa, surgiu também a necessidade da revisão de diversos institutos de natureza substancial e processual, como condição indispensável para a superação dos naturais conflitos decorrentes da ultrapassagem da ideologia individual e patrimonialista.

No mesmo sentido, indica Rodrigo Xavier Leonardo (2002, p. 114):

Se antes o elemento primordial da responsabilidade (expressão que traz consigo a idéia de reprimenda, de desvalor moral) era a culpa, hoje o elemento basilar ao dever de indenizar é o dano. Nesse sentido, a própria expressão “responsabilidade civil” tem um significado limitado, vez que nem sempre a imputação do dever de indenizar recai sobre o responsável pelo dano. Melhor referir-se a essa disciplina, hoje, como um direito de danos.

A disciplina contemporânea da responsabilidade civil – ou do direito de danos - portanto, atenta ao dinamismo das relações sociais e econômicas e balizada pelos valores constitucionais da dignidade da pessoa humana e da solidariedade social, orienta a releitura dos elementos configuradores da obrigação de indenizar em prol da efetiva tutela da pessoa da vítima e à sua reparação integral.

Embora devam surgir situações inéditas nas quais o uso da inteligência artificial seja preponderante para a ocorrência de danos que de outra forma não ocorreriam, é certo que os interesses em disputa não passaram a existir com o desenvolvimento desta tecnologia, como é o caso da privacidade, da igualdade, da liberdade, da autodeterminação, da incolumidade física e psicológica e da defesa do consumidor, para citar apenas algumas preocupações que decorrem da difusão do emprego da IA, mas também de qualquer outra atividade, econômica ou não.

O que se percebe, porém, é que as vítimas de danos causados pela inteligência artificial experimentam uma nova forma de vulnerabilidade, potencializada pela enorme disparidade entre o indivíduo comum - ou até populações inteiras - e as grandes empresas de tecnologia, o que tem impacto em aspectos como o direito à informação, o poder de consentimento, a capacidade de reverter decisões tomadas por máquinas inteligentes, bem como de obter reparação por eventuais prejuízos patrimoniais ou extrapatrimoniais.

A proteção desses interesses é um imperativo constitucional e, como tal, condiciona a interpretação e a aplicação do Direito Civil e, mais especificamente, da responsabilidade civil, razão pela qual, independentemente da atividade em escrutínio, a proteção à vítima deve ser viabilizada em homenagem ao paradigma solidarista e ao princípio da dignidade da pessoa humana.

Como defende Schreiber (2009), porém, a tutela da vítima não deve se converter em um sistema de responsabilidade injusto que se limita a transferir o ônus do dano de um indivíduo para outro, devendo ser buscados instrumentos que, de fato, promovam a solidarização da reparação dos danos e fomentem uma cultura coletiva de administração de riscos. Nesse sentido, afirma Mulholland (2020, p. 345) ser “necessário refletir sobre qual seria a melhor solução jurídica para atender aos interesses não só da pessoa que sofreu o dano, mas também da sociedade como um todo”, ainda mais quando almejamos um equilíbrio entre a responsabilização por danos e o necessário incentivo à inovação tecnológica, pois também há efeitos benéficos do uso da IA que devem ser mantidos e ampliados.

3 AS PROPOSTAS DE REGULAÇÃO DA INTELIGÊNCIA ARTIFICIAL NO BRASIL: UMA ANÁLISE SOB O PARADIGMA SOLIDARISTA

Até o momento, constatou-se que, embora, a princípio, o caráter autônomo de máquinas que fazem uso de inteligência artificial seja um fator dificultador da atribuição de responsabilidade por eventuais danos, a doutrina civilista tem fornecido mecanismos e institutos que superam o mencionado óbice, considerando a releitura dos elementos caracterizadores do dever de reparar, em deferência à tutela da pessoa da vítima e sob uma perspectiva que prestigia os princípios constitucionais da solidariedade social e da dignidade da pessoa humana.

Assim, observou-se que, na medida em que o regime subjetivo de responsabilidade se mostre um empecilho à reparação das vítimas de danos, por exigir prova de difícil ou impossível produção, a teoria do risco fornece fundamentos para a imputação de responsabilidade por danos causados pela inteligência artificial.

Quanto aos elementos para a configuração do dever de indenizar, o nexo causal, em especial, foi apontado como um problema, considerando a dificuldade de se estabelecer um liame entre o dano causado e alguma pessoa integrante do processo de desenvolvimento ou funcionamento da máquina autônoma. Não obstante, conforme detalhado anteriormente, tal dificuldade foi abordada no âmbito do Direito Civil, tendo sido estipulados expedientes que a mitigam, como é o caso da teoria da presunção de causalidade apresentada por Caitlin Mulholland, aplicável a hipóteses de responsabilidade objetiva pelo fator de risco e de danos difusos.

O que se verifica até o momento, portanto, é que a legislação e a doutrina do Direito Civil brasileiro, considerando a unidade e a completude do ordenamento jurídico, bem como os princípios constitucionais da solidariedade social e da dignidade da pessoa humana, não precisa, necessariamente, de inovações jurídicas para contemplar eventuais danos causados pela inteligência artificial. Dito de outra forma, as vítimas desses danos não se encontram sob um vazio regulatório que impossibilita a pretensão de se obter reparação.

Não obstante, existem, no Brasil, proposições legislativas que buscam regular o assunto de forma específica. Assim sendo, sob o paradigma descrito no capítulo anterior, a análise que segue abordará os projetos de lei em tramitação no Congresso

Nacional que guardam pertinência com a matéria aqui tratada. Para tanto, a pesquisa consistirá em uma busca por proposições legislativas em tramitação nos *sites* da Câmara dos Deputados e do Senado Federal na *internet*, sendo destacados aqueles que mencionarem regras para a responsabilidade por danos.

A pesquisa no site da Câmara dos Deputados, realizada em 12 de novembro de 2021, utilizou como argumento de busca o termo “inteligência artificial”, selecionando os seguintes tipos de proposição em tramitação: proposta de emenda à Constituição; projeto de lei complementar; projeto de lei; medida provisória; projeto de lei de conversão; projeto de decreto legislativo e projeto de resolução. Justifica-se a seleção pela natureza primária das normas que venham a ser aprovadas após a tramitação, considerando o seu fundamento constitucional de validade, nos termos do artigo 59 da Constituição Federal.

Foram encontrados 25 resultados, sendo que apenas 1 proposição continha menção específica à responsabilidade por danos causados pelo uso da inteligência artificial, qual seja, o Projeto de Lei nº 21/2020, de autoria do Deputado Federal Eduardo Bismarck (PDT-CE), que visa criar o marco legal do desenvolvimento e uso da inteligência artificial.

A pesquisa no site do Senado Federal, realizada em 12 de novembro de 2021, utilizou como argumento de busca o termo “inteligência artificial” na opção “pesquisa rápida” do site daquela casa legislativa, que considera todas as informações disponíveis sobre cada proposição, como ementa, apelido, tramitação e arquivos anexados. Nesta modalidade de pesquisa, não foi disponibilizada a opção sobre os tipos de proposição desejados.

Retornaram 16 resultados (3 projetos de lei, 1 medida provisória e 12 requerimentos), dos quais 1 guarda pertinência com a matéria aqui tratada, qual seja, o Projeto de Lei nº 5.051/2019, da autoria do Senador Styvenson Valentim (PODEMOS-RN), que propõe estabelecer princípios para o uso da inteligência artificial no Brasil.

Os projetos de lei serão descritos a seguir e analisados em conjunto adiante, de forma a suscitar as questões mais pertinentes ao objeto deste trabalho.

3.1. Projeto de Lei nº 21/2020 da Câmara dos Deputados

Embora de autoria do Deputado Federal Eduardo Bismarck, a versão aprovada pela Câmara dos Deputados (em 29 de setembro de 2021) e atualmente em tramitação na casa revisora foi o substitutivo apresentado pela Deputada Luisa Canziani (PTB-PR). Trata-se da proposição em estágio mais avançado do processo legislativo.

Verifica-se que, de início, a proposta define o que entende por sistema de inteligência artificial e delimita o seu escopo, excluindo de seu âmbito de incidência aqueles processos de automação que sejam orientados exclusivamente por parâmetros predefinidos, sem a capacidade de aprender e de interagir com o ambiente.

A proposta apresenta os objetivos da aplicação de inteligência artificial no Brasil, mencionando, inicialmente, o desenvolvimento científico e tecnológico, acrescido dos seguintes:

Art. 3º (...)

I – a promoção do desenvolvimento econômico sustentável e inclusivo e do bem-estar da sociedade;

II – o aumento da competitividade e da produtividade brasileira;

III – a inserção competitiva do Brasil nas cadeias globais de valor;

IV – a melhoria na prestação de serviços públicos e na implementação de políticas públicas; e

V – a promoção da pesquisa e desenvolvimento com a finalidade de estimular a inovação nos setores produtivos. (BRASIL, 2020)

Ainda, elenca fundamentos para o desenvolvimento e aplicação da inteligência artificial no Brasil, os quais, além de apontar para a observância de valores constitucionais como a liberdade de expressão, a não discriminação e a privacidade, estimulam o respeito a diretrizes éticas e a adoção de códigos de conduta e guias de boas práticas (autorregulação), que poderão servir como elementos indicativos de conformidade. Ademais, há menção expressa à harmonização da lei com legislações preexistentes, como a Lei Geral de Proteção de Dados, o Marco Civil da Internet, o Código de Defesa do Consumidor e a Lei de Acesso à Informação.

Os princípios para o desenvolvimento e aplicação da IA arrolados pela proposta referem-se à finalidade benéfica dos sistemas; à centralidade do ser humano; à não discriminação; à mitigação de vieses (busca pela neutralidade); à transparência; à

adoção de práticas para redução de riscos; à documentação do processo interno de gestão e responsabilização pelos resultados do funcionamento dos sistemas e à proteção dos dados usados para treinamento.

O projeto estipula diretrizes para que o Poder Público discipline a matéria, quais sejam (i) a intervenção subsidiária, no sentido de que a regulação específica deve ocorrer apenas quando absolutamente necessário; (ii) a atuação setorial do poder público; (iii) a gestão baseada em risco, segundo a qual deverão ser considerados os riscos concretos de cada sistema para que se determine o grau de intervenção necessário; (iv) a participação social e interdisciplinar para a adoção de normas que impactem o desenvolvimento e a operação de sistemas de IA; (v) a análise de impacto regulatório e (vi) a responsabilidade subjetiva dos agentes que atuam na cadeia de desenvolvimento e operação dos sistemas, salvo disposição legal em contrário, como no caso de relações de consumo.

Também, traça diretrizes para uso e fomento dos sistemas de IA por parte dos entes da federação, as quais se traduzem em estímulos à inovação, à adoção da inteligência artificial nos setores público e privado, à capacitação para as transformações do mercado de trabalho, à autorregulação e à cooperação internacional.

As diretrizes serão aplicadas por órgãos e entidades setoriais, que deverão

Art. 8º (...)

I – monitorar a gestão do risco dos sistemas de inteligência artificial, no caso concreto, avaliando os riscos da aplicação e as medidas de mitigação em sua respectiva área de competência;

II – estabelecer direitos, deveres e responsabilidades; e

III – reconhecer instituições de autorregulação. (BRASIL, 2020)

Por fim, estabelece a competência privativa da União para legislar sobre o tema.

3.2. Projeto de Lei nº 5.051/2019 do Senado Federal

O Projeto de Lei nº 5.051/2019, de forma mais sucinta que o anterior, elege os fundamentos do uso da inteligência artificial, entre os quais estão valores existenciais como o bem-estar humano, a dignidade, a igualdade e a privacidade. Destaca-se,

ainda, que a proposta arrola a supervisão humana como fundamento da inteligência artificial.

Frisa-se a menção de que os sistemas decisórios baseados em IA serão considerados como auxiliares à tomada de decisão humana. Além disso, consta que a forma de supervisão humana deverá ser compatível com a gravidade e as consequências da decisão.

De forma mais específica, destaca-se a previsão do art. 4º, § 2º, da proposição, segundo o qual a responsabilidade por danos decorrentes da utilização de sistemas de inteligência Artificial recairá sobre a figura do supervisor.

Constam, ainda, diretrizes para a atuação dos entes da federação no desenvolvimento da IA no Brasil, quais sejam:

Art. 5º (...)

I – a promoção da educação para o desenvolvimento mental, emocional e econômico harmônico com a Inteligência Artificial;

II – a criação de políticas específicas para proteção e para qualificação dos trabalhadores;

III – a garantia da adoção gradual da Inteligência Artificial;

IV – a ação proativa na regulação das aplicações da Inteligência Artificial;
(BRASIL, 2019)

Por fim, consta que o uso da IA pelo Poder Público buscará a qualidade e a eficiência dos serviços oferecidos à população.

3.3. Análise das proposições legislativas à luz do paradigma solidarista

Como se viu, cada uma das proposições traz uma regra específica para a deflagração do dever de indenizar.

O Projeto de Lei nº 5.051/2019 limita-se a atribuir a responsabilidade por danos causados pela inteligência artificial à figura do supervisor. O autor da proposição apresenta a seguinte justificativa para a regra mencionada:

Destaco que, nos termos da proposição, todo sistema de Inteligência Artificial terá a supervisão de uma pessoa humana, de forma compatível com cada aplicação. Com isso, é possível aliar as vantagens trazidas por essa inovação tecnológica com a necessária segurança, evitando que eventuais equívocos do sistema automatizado provoquem consequências indesejadas.

Além disso, a fim de dirimir eventuais dúvidas acerca da responsabilidade por danos decorrentes de sistemas de inteligência artificial, o projeto define que a responsabilidade será, sempre, do supervisor humano do sistema (BRASIL, 2019).

De início, cabe mencionar que a proposta deixa de apontar qual a definição legal de inteligência artificial que estaria sendo adotada, o que dificulta a delimitação do escopo da norma.

A indicação da supervisão humana enquanto fundamento necessário para o uso da inteligência artificial parece ser incompatível com uma das características intrínsecas a esta tecnologia, qual seja, a autonomia. Como já dito, há campos de estudo da inteligência artificial, como é o caso do *machine learning*, cujo intuito é o desenvolvimento de algoritmos que permitam que a máquina aprenda mediante experiência, sem precisar ser programada para toda e qualquer atividade.

Além disso, sabe-se que técnicas de IA, notadamente, o *deep learning*, operam com certo grau de opacidade e pouca explicabilidade, o que significa dizer que nem sempre é possível interpretar o processo de tomada de decisão da máquina, tornando extremamente difícil a tarefa de supervisão de suas atividades.

Quanto à atribuição de responsabilidade à figura do supervisor, de se notar que a proposta não esclarece sobre quem recairia tal encargo, sendo certo que o desenvolvimento de máquinas dotadas de inteligência artificial passa por inúmeras fases que contam com diversos atores com diferentes funções e, não raro, espalhados em diversos locais do mundo. Leonardo Netto Parentoni, Rômulo Soares Valentini e Tarik César Oliveira e Alves (2020, p. 16) bem ilustram esta dificuldade:

Quem deve ser considerado supervisor para fins de responsabilização civil no caso de um software desenvolvido por equipe de programadores, onde cada um deles tem função específica, mas nenhum conhece os detalhes do projeto como um todo? Se esse mesmo software for licenciado para uso de terceiros, supervisor será o desenvolvedor originário ou quem efetivamente utiliza o programa no dia a dia? Pense, ainda, em quem seria o supervisor de um assistente vocal instalado em telefone celular: o fabricante do aparelho (que desenvolveu o sistema) ou o próprio usuário, proprietário do celular (que diariamente “calibra” como esse sistema irá funcionar na prática, fornecendo seus inputs e preferências pessoais)? Se o celular do exemplo anterior for revendido, isto automaticamente altera quem é o supervisor do sistema? No caso de um carro autônomo, que sabidamente reúne milhares de componentes, provenientes de fabricantes distintos, quem é o supervisor? Seria o fabricante de cada componente individualmente considerado? O fabricante final do carro? O revendedor/distribuidor? Ou o usuário que, de novo, utiliza o veículo e o “calibra” diariamente, conforme as suas preferências pessoais? Mais complexa ainda é a situação na qual um mesmo carro autônomo é utilizado por várias pessoas diferentes, no mesmo dia, por

exemplo, para transporte pessoal em aplicativos como o Uber, Lyft e Cabify. Todos esses exemplos são situações concretas que já estão ocorrendo ou podem ocorrer num futuro muito próximo.

Sob essa perspectiva, é possível aventar que o ônus da reparação pode vir a ser atribuído aos indivíduos que têm menos condições de arcar com o prejuízo, como empregados de empresas ou consumidores finais. Esta problematização já foi apresentada por Hartmann Peixoto e Coutinho (2020, p. 9). Vejamos:

No texto do projeto de lei, a previsão pura e simples da responsabilidade do supervisor da máquina não deixa claro a quem será atribuído o encargo. Caso o usuário da máquina seja considerado o seu supervisor, isto pode acarretar, por exemplo, a responsabilidade do consumidor por danos causados pelo produto, subvertendo a lógica do direito consumerista. Ainda, o usuário da máquina pode ser o empregado de alguma empresa que utiliza inteligência artificial em suas atividades. No caso, a pessoa a ser instada a reparar eventual dano seria aquela que menos condições teria de suportar o prejuízo.

Assim, para os autores, ao invés de vincular o uso da inteligência artificial à figura de um supervisor, seria mais interessante a regulação de boas práticas “nos campos de responsabilidade ética e responsabilidade normativa, que devem refletir nas etapas de validação, verificação, controle e segurança tanto no desenvolvimento, quanto no uso de uma IA” (2020, p. 8).

A proposta não deixa claro qual regime de responsabilidade recairia sobre o supervisor da inteligência artificial. Não obstante, de pronto, constata-se que atribuir responsabilidade a uma única pessoa por todo e qualquer dano causado por máquinas dotadas de inteligência, além de denotar certo desconhecimento em relação às peculiaridades da tecnologia, vai de encontro ao paradigma solidarista apresentado no capítulo anterior. Isso porque a regra importaria na transferência do prejuízo da vítima para um outro único indivíduo, que, ademais, seria o único efetivamente incentivado a evitar a ocorrência de danos, enfraquecendo a busca por uma IA mais segura, ética e responsável (HARTMANN PEIXOTO; COUTINHO, 2020).

Quanto ao ponto, a proposição negligencia hipóteses como a incidência da responsabilidade objetiva pelo risco da atividade, deixando de reconhecer a participação de toda uma cadeia de atores da indústria tecnológica na ocorrência do dano, os quais teriam melhores condições de repartir os custos indenizatórios. Ainda, os danos poderiam ocorrer no âmbito consumerista, para o qual há previsão de

responsabilidade independente de culpa e em caráter solidário de toda a cadeia de fornecimento de produtos.

Por fim, em casos nos quais pudesse ser observada certa paridade entre as partes envolvidas, poder-se-ia suscitar, até mesmo, o regime de responsabilidade subjetivo, para o qual seria necessário, mediante prova, apontar os *standards* de conduta inobservados pelo supervisor, cujas obrigações e atribuições também não restaram declinadas pelo legislador.

Controvérsia similar tem sido suscitada no âmbito da Lei Geral de Proteção de Dados, na qual não há indicação explícita da exigência ou prescindibilidade de culpa para a deflagração da responsabilidade por danos causados na atividade de tratamento de dados pessoais, gerando uma discussão doutrinária que ora conclui pelo regime objetivo de responsabilidade, ora pelo subjetivo.

Por exemplo, Caitlin Mullholland (2020, p. 15) considera que

a Lei Geral de Proteção de Dados, em seu artigo 42, adota a teoria que impõe a obrigação de indenizar independentemente da análise da culpa dos agentes de tratamento de dados, isto é, a responsabilidade civil é objetiva. Fundamenta esta conclusão o fato de que a atividade desenvolvida pelo agente de tratamento é evidentemente uma atividade que impõe riscos aos direitos dos titulares de dados. Estes riscos, por sua vez, são intrínsecos, inerentes à própria atividade. Significa dizer que os danos resultantes da atividade habitualmente empenhada pelo agente de tratamento de dados, uma vez concretizados, são quantitativamente elevados - pois atingem um número indeterminado de pessoas - e qualitativamente graves - pois violam direitos que possuem natureza personalíssima, reconhecidos pela doutrina como direitos que merecem a estatura jurídica de direitos fundamentais.

De outra sorte, para Bruno Bioni e Daniel Dias (2020, p. 21),

“não deve haver dúvidas de que a política legislativa adotada exige a investigação em torno de um juízo de culpa dos agentes de tratamento de dados, mas, ao mesmo tempo, prescreve uma série de elementos com alto potencial de erosão dos filtros para que os agentes de tratamentos de dados sejam responsabilizados”.

Aliás, sobre o tema, Maria Celina Bodin de Moraes e João Quinelato de Queiroz (2019, p. 1) aduzem que “a cada dia tem-se mais e mais hipóteses de regimes especiais diferenciados, tornando impossível a tarefa de sistematizar a matéria da responsabilidade civil”.

Como se vê, ao conceder um tratamento assistemático à responsabilidade civil por danos causados pela inteligência artificial, o legislador pouco contribuiu com a

pacificação do tema, apresentando regra (i) atécnica, por não observar as peculiaridades da tecnologia em questão; (ii) injusta, por situar o ônus da reparação em uma única figura, negligenciando os demais atores da cadeia de desenvolvimento da IA e (iii) imprecisa, por não delimitar o regime de responsabilidade que buscava propor.

Quanto ao Projeto de Lei nº 21/2020, na versão que tramita atualmente, chama atenção, para fins de responsabilidade civil, o detalhamento conferido ao princípio da transparência, pelo qual se estipula que as pessoas têm o direito de serem informadas acerca da “identidade da pessoa natural, quando ela operar o sistema de maneira autônoma e individual, ou da pessoa jurídica responsável pela operação dos sistemas de inteligência artificial” (artigo 5º, inciso V, al. “b”), pelo indício de que a reparação por eventuais danos poderá recair sobre essas figuras, cujas atribuições e obrigações, porém, restaram não ditas.⁴¹

Por outro lado, o projeto estipula o regime subjetivo de responsabilidade por danos causados pela inteligência artificial, salvo disposição legal em sentido contrário, denotando a exclusão da hipótese de incidência da cláusula geral de responsabilidade objetiva com fundamento no risco da atividade, nos termos do artigo 927, parágrafo único, do Código Civil. A previsão foi objeto de uma carta aberta encaminhada ao Senado Federal, assinada por diversos juristas brasileiros⁴², pela qual manifestaram preocupação com regime de responsabilidade adotado pelo projeto de lei.

Os juristas advertem que a norma vai de encontro com as construções doutrinárias nacionais e internacionais sobre o tema, bem como que estaria colocando em risco a possibilidade das vítimas de danos causados pela inteligência artificial de obterem a devida reparação.

Destacam que, embora a IA seja fundamental para o desenvolvimento social e econômico do país, o fato de ser uma tecnologia complexa e ainda em fase de

⁴¹ O grupo de pesquisa DR.IA.UnB – Direito e Inteligência Artificial emitiu nota técnica com críticas ao texto do substitutivo. Entre elas, consta que o nível de detalhe conferido ao princípio da transparência importa em limitação à sua aplicabilidade, que deveria ser mais geral e dinâmico (HARTMANN PEIXOTO, 2021).

⁴² Assinam a carta os juristas: Ana Frazão, Anderson Schreiber, Bruno Bioni, Bruno Miragem, Caitlin Sampaio Mulholland, Carlos Edison do Rêgo Monteiro Filho, Carlos Nelson Konder, Cristiano Chaves de Farias, Danilo Doneda, Dierle Nunes, Estela Aranha, Fabiano Menke, Filipe José Medon Affonso, Gustavo Tepedino, Guilherme Damasio Goulart, Guilherme Calmon Nogueira da Gama, Guilherme Magalhães Martins, Heloisa Helena Barboza, Ingo Wolfgang Sarlet, Juliano Madalena, Laura Schertel Mendes, Lucia Maria Teixeira Ferreira, Marcos Ehrhardt Júnior, Maria Celina Bodin de Moraes, Milena Donato Oliva, Nelson Rosenvald e Rafael Zanatta.

desenvolvimento aponta para a ausência de um nível satisfatório de previsibilidade quanto às consequências de sua aplicação, pelo que devem ser garantidos um grau mínimo de risco às pessoas e a reparabilidade por eventuais danos.

Acrescentam que o regime subjetivo acabaria por impor ao cidadão os custos do desenvolvimento da tecnologia e não seria suficiente para fomentar a adoção de medidas de precaução pelos atores responsáveis. Além disso, aduzem que priorizar tal regime de forma abstrata, ao passo em que não favorece o investimento em inovação no país, proporciona um ambiente de ausência de responsabilização generalizada, pois impossibilita a prova, pelas vítimas, da culpa dos agentes que figuram na cadeia de desenvolvimento e operação de sistemas de IA.

Assim sendo, sugerem a alteração do teor do artigo 6º, inciso VI, do Projeto de Lei nº 21/2020 para que seja adotada a seguinte redação:

Artigo 6º: VI – responsabilidade: normas sobre responsabilidade dos agentes que atuam na cadeia de desenvolvimento e operação de sistemas de inteligência artificial devem, salvo disposição legal em contrário, levar em consideração a tipologia da inteligência artificial, o risco gerado e seu grau de autonomia em relação ao ser humano, além da natureza dos agentes envolvidos, a fim de se determinar, em concreto, o regime de responsabilidade civil aplicável⁴³.

De fato, uma abordagem que permita ao julgador levar em consideração as características do caso concreto em relação aos aspectos técnicos da IA, ao grau de risco e autonomia, bem como em relação às características das partes envolvidas, aferindo eventual disparidade entre ofendido e ofensor, parece ir ao encontro do paradigma solidarista descrito no capítulo anterior.

A escolha sobre estabelecer uma regulação específica da IA e como fazê-lo requer a busca de um equilíbrio entre os danos que se pretende evitar e os benefícios que se almeja pelo emprego da tecnologia. Regras muito lenientes, que dificultam a responsabilização, podem fomentar comportamentos e hábitos irresponsáveis e, de outra sorte, regras demasiadamente rígidas podem desestimular a inovação em busca de uma IA mais segura e benéfica. De qualquer forma, tendo em vista a multiplicidade de aplicações proporcionadas por essa tecnologia, nem todos os danos decorrentes

⁴³ Disponível em <https://www.change.org/p/senado-federal-carta-aberta-de-juristas-ao-senado-federal-contra-o-artigo-6o-inciso-vi-do-pl-21-a-2020?utm_content=cl_sharecopy_31194808_pt-BR%3A2&recruiter=1231429564&utm_source=share_petition&utm_medium=copylink&utm_campaign=share_petition>. Acesso em 15 nov. 2021.

de seu uso poderão ser endereçados da mesma forma, sendo essencial haver razoabilidade e proporcionalidade na atribuição de responsabilidade (GIUFFRIDA, 2019).

No caso, atribuir à vítima o ônus probatório da culpa do ofensor pode inviabilizar a pretensão reparatória, principalmente porque, quanto mais autônoma a máquina, mais difícil será comprovar a inobservância de um *standard* de conduta que possa ser atribuível a alguma pessoa integrante da cadeia de desenvolvimento da IA que, aliás, é composta por atores diversos e dispersos, como já mencionado. Relembre-se que, conforme delimitado no texto da proposta, a norma se aplicaria apenas àquelas tecnologias que aprendem e se adaptam ao ambiente sem ser previamente programadas para toda e qualquer tarefa.

Acrescente-se que talvez seja difícil até mesmo para os desenvolvedores da tecnologia indicar quais padrões de comportamento seriam adequados para cada setor que faz uso de IA, considerando não apenas a autonomia característica de máquinas dotadas de inteligência, mas também o fato de tratar-se de uma tecnologia cujo potencial ainda não foi completamente descoberto.

A efetiva solidarização dos riscos, porém, não se encerra na estipulação do regime de responsabilidade civil, sendo certo que o dever de não lesar a outrem e de agir de forma responsável para com a coletividade antecede a ocorrência de um dano e demanda ações concretas para evitá-lo.

Da leitura do texto das proposições legislativas, extrai-se uma preocupação do legislador em definir uma base principiológica para a IA, no intuito de que esta se revele benéfica para a humanidade, veiculando o respeito à dignidade humana, à liberdade, à democracia, à igualdade, aos direitos humanos, à pluralidade, à diversidade, à privacidade, à transparência, à livre iniciativa, à livre concorrência, à liberdade de expressão, entre outros ali mencionados.

Nesses termos, as propostas parecem compartilhar de inquietações já manifestadas por diversos autores que trataram dos riscos apresentados pela inteligência artificial, a exemplo da interferência de algoritmos em processos eleitorais (NEMITZ, 2018), do monitoramento discriminatório e da violação a liberdades individuais (STONE et al, 2016), da coleta e utilização de dados pessoais para fins não autorizados e da substituição da mão-de-obra humana (STANKOVIC et al, 2017).

De se notar, porém que toda e qualquer atividade deve observar fundamentos tão importantes quanto os acima mencionados, os quais são protegidos em nível constitucional e, portanto, têm incidência sobre toda a vida em sociedade. Questiona-se, assim, se seria necessária uma legislação específica para formalizar a noção de que o uso de IA deve respeitar valores como a dignidade humana, a liberdade, a democracia e a igualdade, já que esse tipo de previsão pouco inova no ordenamento jurídico.

Por outro lado, há, também, nas propostas, notadamente, no PL nº 21/2020, fundamentos e princípios mais específicos à matéria aqui tratada, cujo intuito parece ser induzir os agentes que desenvolvem e utilizam a tecnologia a adotarem práticas que preservem a segurança dos sistemas, que identifiquem e corrijam vieses, que concedam ao usuário informações claras e acessíveis sobre os sistemas de IA com que interagem e que documentem o processo interno de gestão. Ou seja, são fundamentos e princípios voltados à prevenção de danos e, quando da sua ocorrência, indicam um caminho para possibilitar a reparação.

Perceba-se, porém, que o disposto nas proposições funciona como um guia para eventuais regulações infraconstitucionais, não havendo qualquer instrumento concreto de fiscalização, sanção por descumprimento, ou, sequer, parâmetros definidos de conformidade que efetivamente vinculem os agentes do setor. Chama atenção, também, o fato do legislador brasileiro ter optado por deixar um amplo espaço para a autorregulação, o que, como se verá adiante, pode ser um empecilho para que se efetive direitos fundamentais individuais e coletivos em face de grandes empresas de tecnologia.

O texto analisado contrasta com outras legislações brasileiras que implementaram mecanismos mais efetivos para proteger direitos fundamentais dos cidadãos em temas que tangenciam ao presente, como é o caso da Lei Geral de Proteção de Dados (LGPD) e o Marco Civil da Internet (MCI), as quais, embora também contenham dispositivos de caráter principiológico, delineiam de forma mais clara os direitos e obrigações dos atores envolvidos, descrevem detalhadamente a atividade em escrutínio e, no caso da LGPD, estipula sanções administrativas por infrações à lei.

4 O DEVER DE NÃO LESAR A OUTREM: SOLIDARIZAÇÃO DOS RISCOS E PREVENÇÃO⁴⁴ DE DANOS

Embora o Código Civil, ao tratar da responsabilidade civil, faça referência explícita à função reparatória do instituto, o enunciado nº 446 da V Jornada de Direito Civil estabelece que a “responsabilidade civil prevista na segunda parte do parágrafo único do art. 927 do Código Civil deve levar em consideração não apenas a proteção da vítima e a atividade do ofensor, mas também a prevenção e o interesse da sociedade” (BRASIL, 2011).

O Código do Consumidor, em seu art. 6º, inciso VI, elenca como direito básico do consumidor “a efetiva prevenção e reparação de danos patrimoniais e morais, individuais, coletivos e difusos” (BRASIL, 1990).

A Lei Geral de Proteção de Dados, ao estipular princípios para as atividades de tratamento de dados pessoais, além de mencionar a prevenção da ocorrência de

⁴⁴ Sabe-se da existência de corrente doutrinária que reconhece que a responsabilidade civil teria uma função preventiva que prescindiria da ocorrência do dano para ser acionada. Por exemplo, Teresa Ancona Lopez (2010, p. 79 e 139), considerando o desgaste do efeito preventivo enquanto mera consequência da sanção aplicada ao ofensor, defende ser “perfeitamente possível responsabilidade civil sem dano (apenas sua ameaça). Essa nova responsabilidade surgiu da adoção dos princípios da prevenção e da precaução com a finalidade de dar maior segurança a todos os cidadãos”. Assim, a autora vê como necessário que se “ponha em prática o princípio da prevenção como instrumento anterior ao dano, e não como consequência desse (função automática da indenização), quando o seguro já ‘reparou’ os prejuízos sofridos”. Para tanto, Lopez apresenta a noção de uma ação de responsabilidade civil preventiva, pela qual se sancionaria a não-observância de atitudes para evitar danos graves e irreversíveis.

Por outro lado, há autores que divergem desse posicionamento, tendo em vista que a responsabilidade puramente preventiva estaria a negligenciar seu escopo principal, qual seja, a ocorrência de um dano digno de reparação, sendo próprio da responsabilidade civil a análise posterior à lesão, já que outros ramos do direito se ocupariam de hipóteses potencialmente danosas que antecedem o fenômeno lesivo. Assim, para Cássio Monteiro Rodrigues (2020, p. 22), a função preventiva figuraria de forma acessória à função reparatória, pois a responsabilidade civil “se volta, em sua estrutura e função prioritária, à reparação dos prejuízos certos experimentados pela vítima”.

Gustavo Tepedino (2004, p. 2-3), por sua vez, assevera que, “por mais louvável que seja a ampliação do dever de reparar, protegendo-se as vítimas de uma sociedade cada vez mais sujeita a riscos – decorrentes das novas tecnologias, dos bancos de dados pessoais, dos aparatos industriais, da engenharia genética, e assim por diante –, não se pode desnaturar a finalidade e os elementos da responsabilidade civil. O dever de reparar não há de ser admitido sem a presença do dano e do nexo de causalidade entre a atividade e evento danoso, tendo por escopo o ressarcimento da vítima”.

No mesmo sentido, leciona Roberto Paulino de Albuquerque Júnior (2016, p. 99): “se o efeito próprio da responsabilidade civil é o indenizativo, não é possível falar em responsabilidade sem dano, não sem que ao menos haja uma radical transformação da própria concepção de responsabilidade civil em um sentido que nem mesmo a proposta do direito de danos parece pretender operar”.

Para os fins do presente trabalho, seguiremos considerando o dano como elemento essencial para a configuração da responsabilidade civil, não prescindindo, porém, das noções de prevenção e precaução como instrumentos de solidarização dos riscos, como se detalhará neste capítulo.

danos (art. 6º, inciso VIII), posiciona a responsabilização ao lado da prestação de contas, para que o agente demonstre a “adoção de medidas eficazes e capazes de comprovar a observância e o cumprimento das normas de proteção de dados pessoais e, inclusive, da eficácia dessas medidas” (art. 6º, inciso X) (BRASIL, 2018), medida que foi considerada por Moraes e Queiroz (2019, p. 5) como “uma mudança profunda em termos de responsabilização *lato sensu* — vale dizer, no que diz respeito à imposição de deveres voltados a prevenir danos”. Os autores remetem ao termo “responsabilidade civil proativa”, no sentido de que não basta o cumprimento da lei, mas a demonstração do seu cumprimento por parte dos protagonistas no tratamento de dados.

Retomando o que mencionava Anderson Schreiber (2009, p. 248), técnicas de prevenção e precaução de danos seriam meios para a solidarização do dever de reparação, na medida em que “distribuem por toda a sociedade ou, ao menos, pela coletividade de agentes potencialmente lesivos os custos da administração dos riscos”.

Caitlin Mulholland (2010, p. 297), referindo-se à ideia de repartição dos riscos e remetendo-se à teoria da causalidade presumida, entende que

o reconhecimento social – e, conseqüentemente, jurídico – de que a probabilidade estatística pode ser utilizada como instrumento para a análise e identificação presumida do nexu causal, gera como efeito a tomada de consciência por aquele que desenvolve uma atividade arriscada para a adoção de medidas de precaução para evitar, assim, a sua potencial responsabilização. Isto é, a utilização da probabilidade na delimitação presumida da causalidade pode tornar-se uma ferramenta útil para a efetivação do princípio da precaução: se existe uma probabilidade de que a atividade desenvolvida cause um determinado dano, envidemos todos os esforços para que este não aconteça.

Por sua vez, Nelson Rosenvald (2020), em trabalho focado na responsabilidade civil no âmbito da atividade de tratamento de dados pessoais, suscita os diferentes significados que a palavra “responsabilidade” atrai na língua inglesa, para indicar que a função reparatória (*liability*) não teria o monopólio da responsabilidade civil do século XXI.

Assim, o autor menciona o termo *responsability*, remetendo-se a uma conotação moral da responsabilidade, a qual seria absorvida de forma voluntária e sem imposição legal. Trata-se de “um conceito prospectivo de responsabilidade, no qual ela se converte em instrumento para autogoverno e modelação da vida”, segundo

Rosenvald (2020). Nessa linha, suscita instrumentos como ética e capacitação no exercício da atividade, bem como educação para os seus destinatários, a exemplo do que se observa no Marco Civil da Internet⁴⁵.

O termo *accountability* refere-se a parâmetros preventivos estipulados por meio de regulação e que, antes da ocorrência do dano, estabelecem boas práticas e padrões de segurança para o desenvolvimento da atividade. Caso o dano venha a se concretizar, a *accountability* funcionaria como uma orientação para que a autoridade competente identifique e delimite responsabilidades. O autor detalha:

Assim, ao invés do juiz se socorrer da discricionariedade para aferir o risco intrínseco de uma certa atividade por sua elevada danosidade - o desincentivo ao empreendedorismo é a reação dos agentes econômicos à insegurança jurídica -, estabelecem-se padrões e garantias instrumentais que atuam como parâmetros objetivos para a mensuração do risco em comparação com outras atividades. Aliás, se o causador do dano houver investido em *compliance*, com efetividade, pode-se mesmo cogitar da redução da indenização, como espécie de sanção premial, a teor do parágrafo único do art. 944 do Código Civil.

Por fim, a *answerability*, também de caráter preventivo, diz respeito à explicabilidade ou possibilidade de justificação de escolhas, relacionando-se com o direito à informação, porém, extrapolando-o. Rosenvald (2020) exemplifica que, no âmbito da LGPD, essa noção de responsabilidade se concretiza mediante o direito do titular de dados de solicitar a revisão humana de decisões automatizadas, como nos casos em que uma inteligência artificial, fazendo uso de dados pessoais, elabora perfis de pessoas para fins diversos. No caso, a *liability* surgiria posteriormente, caso o perfilamento venha a violar direitos ou alcançar situações existenciais.

Conclui o autor que a *responsability*, a *accountability* e a *answerability* performam as funções preventiva e precaucional da responsabilidade civil, arrematando com uma “renovada perspectiva bilateralizada” da responsabilidade civil:

a responsabilidade como mecanismo de imputação de danos - foco da análise reparatória - no qual o agente se responsabiliza "perante" a vítima, convive com a responsabilidade "pelo outro", o ser humano. Aqui, agrega-se a pessoa do agente e a indução à conformidade mediante uma regulação de gestão de riscos, sobremaneira a sua mitigação, seja por parte de um desenvolvedor de tecnologias digitais emergentes como de um agente de tratamento (*accountability/answerability*). Porém, em uma noção de

⁴⁵ “Art. 26. O cumprimento do dever constitucional do Estado na prestação da educação, em todos os níveis de ensino, inclui a capacitação, integrada a outras práticas educacionais, para o uso seguro, consciente e responsável da internet como ferramenta para o exercício da cidadania, a promoção da cultura e o desenvolvimento tecnológico” (BRASIL, 2014).

reciprocidade, a mitigação de ilícitos e danos também incumbe a cada um de nós, mediante a paulatina construção de uma autodeterminação responsável que nos alforrie da heteronomia e vitimização (responsibility), pois como já inferia Isaiah Berlin "O paternalismo é a pior forma de opressão" (ROSENVALD, 2020).

Iria Giuffrida (2019) argumenta que, antes de pensarmos em elaborar um sistema de reparação de danos causados por sistemas de IA, devemos priorizar a identificação e o desestímulo das práticas mais temerárias, de forma a conformá-las aos diversos códigos éticos que estão emergindo em todo o mundo. Nesse sentido, Fabiano Hartmann Peixoto (2020, p. 84) defende que o aprimoramento do nível ético relacionado ao uso da inteligência artificial é possível, e os desafios de sua aplicação devem ser “enfrentados de forma integrada à sociedade, com os aspectos normativos, culturais, institucionais em perspectiva de um diálogo social sobre o conceito e limites do que é uma ação responsável, equitativa, rastreável, confiável e governável”.

Considerando a amplitude do avanço da inteligência artificial e um contexto de ansiedade pelo seu domínio por parte de empresas e governos, nem sempre a preocupação com princípios éticos e de segurança será uma prioridade. Ademais, não é uma tarefa fácil determinar concretamente quando e o que precisa ser regulado, pois os dados para estimar o risco da atividade são limitados, especialmente em estágios iniciais de desenvolvimento e implantação da tecnologia (SANTOS *et al*, 2021).

A escolha por adotar ou não medidas de precaução induz os agentes do setor a um dilema entre um desenvolvimento mais lento e custoso, porém mais seguro, pois permite detectar e mitigar os potenciais riscos da atividade, e um desenvolvimento rápido que, embora apresente mais riscos, pode fazer os responsáveis pela sua implementação avançarem na corrida pela supremacia tecnológica (SANTOS *et al*, 2021).

Portanto, o desafio regulatório que se impõe diz respeito à forma com que os agentes do campo tecnológico serão instados a adotar procedimentos para administrar os riscos da atividade, mantendo-se, porém, um cenário de incentivo à inovação (SANTOS *et al*, 2021).

Para tanto, ao contrário do que se constatou quanto ao aspecto reparatório, talvez se faça necessária uma regulação específica da inteligência artificial, de forma a endereçar seus riscos e incertezas *ex ante*. A seguir, se abordará a forma com que o princípio da precaução pode incidir sobre a pesquisa e o desenvolvimento da

inteligência artificial, com foco em seu “combustível” (SOLOW-NIEDERMAN, 2019): os dados.

4.1. O princípio da precaução como deliberação pública: lições das leis de proteção de dados para o campo da inteligência artificial

O princípio da precaução no mundo contemporâneo foi introduzido através do direito ambiental. Embora tenha sido mencionado em debates internacionais desde a década de 70, foi consagrado na Declaração do Rio de Janeiro – Eco-92⁴⁶, a partir da qual passou a ser mencionado em outros ramos do direito, conforme explica Lopez (2010, p. 98):

Assim, foi a partir daí que o princípio da precaução tomou força não só no direito ambiental, mas também começou a ser aplicado no direito sanitário (sangue contaminado), direito alimentar (“vaca louca”) e foi aceito nos ramos do direito que têm como finalidade proteger a incolumidade física e a saúde dos indivíduos. Dessa forma, aparece no direito médico e hospitalar, no direito do consumidor (englobando todo tipo de produto ou serviço que não apresentam a segurança esperada), assim como na biotecnologia e nas descobertas tecnológicas e científicas em geral.

Trata-se de preceito suscitado quando estamos diante de incerteza acerca dos danos que uma atividade pode vir a causar, de forma a administrá-los antecipadamente, sendo uma maneira de mitigar os perigos da “sociedade de risco”. Assim, uma atitude que atende ao imperativo da precaução seria aquela tomada com prudência e cautela, de forma a evitar riscos possíveis (LOPEZ, 2010, p. 98).

Nesse sentido, Bruno Bioni e Maria Luciano (2020), explicam que o princípio da precaução surgiu enquanto resposta aos desafios regulatórios em situações para as quais ainda não havia conhecimento científico suficiente para embasar modelos de avaliação de risco, os quais pressupunham a existência de dados para possibilitar a antecipação da ocorrência de danos.

⁴⁶ Princípio 15 da Declaração do Rio de Janeiro de 1992: “De modo a proteger o meio ambiente, o princípio da precaução deve ser amplamente observado pelos Estados de acordo com suas capacidades. Quando houver ameaça de danos sérios ou irreversíveis, a ausência de absoluta certeza científica não deve ser utilizada como razão para postergar medidas eficazes e economicamente viáveis para prevenir a degradação ambiental”.

Embora não exista um conceito preponderante do princípio da precaução, há expressões comuns às diversas formulações elaboradas sobre o tema, como “dano”, “irreversível”, “risco” e “sério”, cujas definições, todavia, segundo os autores, “caberia à toda a sociedade definir”, de forma que as escolhas regulatórias observem procedimentos deliberativos que assegurem a sua justificação explícita (BIONI; LUCIANO, 2020, p. 211). Nesse sentido, o princípio da precaução teria uma conotação de deliberação pública e funcionaria como

um framework para se pensar medidas regulatórias, dentre as quais contratos, sanções, acordos legalmente definidos, financiamento a projetos de pesquisa e recomendações. É justamente essa tarefa de determinação que abre espaços para a discussão de valores a serem protegidos ou preteridos, convidando diferentes stakeholders a discutir padrões de segurança. (BIONI; LUCIANO, 2020, p. 212)

Bioni e Luciano (2020) observam que o emprego da IA demandaria, em grande medida, o processamento de dados pessoais para a automatização de processos e decisões, pelo que apontam para as leis de proteção de dados como o “portal de entrada” do princípio da precaução ao campo da inteligência artificial.

No âmbito da proteção de dados, os autores observam uma preocupação cada vez mais acentuada com uma regulação precaucionária, o que seria atribuído ao aumento da assimetria de informação causado pelo desenvolvimento tecnológico e pelo incremento de uma economia dirigida por dados.

A nova onda de tecnologias de informação e comunicação (TICs) tornou ainda mais exponencial os possíveis efeitos adversos de uma atividade de tratamento de dados pessoais. Juntas, Internet das Coisas, Big Data e Inteligência Artificial permitem a coleta massiva de informações pessoais e, principalmente, inferências mais intrusivas a respeito dos cidadãos (Spina, 2014). Com isso, torna-se mais complexo o processo de cognição, avaliação e gerenciamento dos riscos de uma economia de dados. Os agentes de tratamento de dados – controladores e operadores – passaram a deter uma superioridade informacional ainda maior frente aos demais atores – cidadãos e órgãos fiscalizadores – desse ecossistema. (BIONI; LUCIANO, 2020, p. 214)

Mas a proteção dos titulares de dados pessoais não é o único fator a merecer prudência e cautela quando estamos a lidar com algoritmos de inteligência artificial, merecendo destaque o tratamento de dados em geral.

A principal ferramenta de IA atualmente utilizada, o *machine learning*, é treinada com bancos de dados imensos, sendo que a maneira com que se dá o

treinamento pode tanto mitigar quanto perpetuar vieses e erros e ser determinante para a ocorrência de danos. De se notar que não estamos a tratar de algoritmos estáticos, pois, após serem treinados com um *dataset*, passam a funcionar mediante o contato com dados novos, pelo que precisam ser programados para adaptarem-se constantemente. Dessa forma, há casos em que a forma com que algoritmos de inteligência artificial chegam a determinado resultado, bem como o resultado em si, são surpreendentes e inesperados até para seus desenvolvedores (SOLOW-NIEDERMAN, 2020).

Alicia Solow-Niederman (2020), apresenta um exemplo bem-humorado em que um algoritmo programado para classificar dados acabou por deletar todas as informações que lhe foram alimentadas sob a lógica de que, se não há nenhum dado para classificar, não há tarefas pendentes a serem realizadas. Outro exemplo apresentado pela autora refere-se a um programa de simulação de pouso de aeronaves que foi programado para identificar alternativas para desacelerar aviões. Em vez de alcançar o objetivo proposto, o algoritmo encontrou uma brecha pela qual poderia fazer os indicadores de força medirem valores próximos de zero, enquanto, na verdade, o veículo estava acelerando rumo a uma colisão. Felizmente, o erro foi detectado na simulação, porém a autora chama atenção à gravidade de hipótese semelhante ocorrer com veículos autônomos em funcionamento.

Solow-Niederman (2020) não atribui acontecimentos dessa natureza a uma intencionalidade da máquina em escapar ao controle humano, ou a alguma característica intrínseca da IA que a transforme em uma tecnologia absolutamente indomável, mas sim, à dificuldade quanto à seleção de dados e às escolhas de programação por parte das instituições responsáveis pelo seu desenvolvimento.

Para ilustrar a afirmação, a autora menciona o caso em que o supercomputador *Watson*, da empresa de tecnologia *International Business Machines Corporation* (IBM), recomendou tratamentos inseguros e inadequados para pacientes diagnosticados com câncer. A verificação dos motivos para a falha levou à conclusão de que o problema não teria sido do computador em si, mas do fato de que os engenheiros da IBM teriam treinado o algoritmo com dados de pacientes hipotéticos e não de pacientes reais, além de terem pressionado a comercialização do *software* para o setor da saúde, embora soubessem de suas inconsistências (SOLOW-NIEDERMAN, 2020).

As instituições que lideram a pesquisa e o desenvolvimento da inteligência artificial⁴⁷ têm historicamente adotado comportamentos voltados ao rápido desenvolvimento da tecnologia, com pouca atenção à prevenção de danos⁴⁸ e ampla margem de liberdade para estipular suas próprias regras de segurança, ao passo em que órgãos do Estado afastam-se da tarefa de regular e fiscalizar o setor (SOLOW-NIEDERMAN, 2020), o que tende a ampliar a assimetria informacional e de poder entre os usuários da tecnologia e seus desenvolvedores e, também, entre os setores público e privado, reduzindo, conseqüentemente, a capacidade de se chegar a uma resposta democrática sobre o que se entende por termos como “dano”, “irreversível”, “risco” e “sério”, mencionados acima.

Uma transparência qualificada⁴⁹ seria imprescindível para os processos de tomada de decisão acerca do que é considerado um risco tolerável em atividades que envolvem tratamento de dados, para que se permita uma pluralidade de olhares que forneçam considerações técnicas, mas também de caráter ético, social e de respeito aos direitos humanos. No âmbito da IA, notadamente, no que se refere às decisões automatizadas, essa perspectiva seria também uma forma de democratizar⁵⁰ a regulação da tecnologia, proporcionando-se uma decisão coletiva acerca da sua adoção e dos riscos que decorrem da sua aplicação (BIONI; LUCIANO, 2020).

Para tanto, Bioni e Luciano (2020) indicam dois instrumentos presentes nas leis de proteção de dados e que poderiam proporcionar a aplicação do princípio da precaução no âmbito da inteligência artificial, quais sejam, os Relatórios de Impacto e o direito à revisão de decisões automatizadas. Como se verá, trata-se, respectivamente, de um instituto voltado para permitir uma supervisão estatal do desenvolvimento da atividade de tratamento de dados - e, caso fosse implementado no âmbito da inteligência artificial, poderia influenciar até no *design* de algoritmos -, e

⁴⁷ De forma mais ilustrativa, pode-se afirmar que estamos aqui a tratar das “*frightful five*”, ou seja, das empresas que dominam o setor de tecnologia: Google, Facebook, Microsoft, Apple e Amazon (NEMITZ, 2018, p. 2).

⁴⁸ Solow-Niederman (2020) exemplifica com o caso do Facebook, cujo lema interno era “*move fast and break things*”, e do Google, em cujo código de conduta havia variações da frase “*don’t be evil*” entre os anos de 2000 e 2018.

⁴⁹ Para os autores, essa transparência qualificada seria o princípio da *accountability*, o que, para Nelson Rosenwald, como exposto acima, referir-se-ia ao termo *answerability*.

⁵⁰ No mesmo sentido, posiciona-se Matthew U. Scherer (2016), que atribui ao processo legislativo democrático a legitimidade para que se possa determinar o nível de risco aceitável aos cidadãos em decorrência do uso da IA, bem como para decidir se há esferas em que não deve ser conferido espaço de decisão às máquinas autônomas.

de outro voltado à garantia de direitos individuais das pessoas impactadas pela decisão automatizada.

Os relatórios de impacto à proteção de dados pessoais (RIPDP) são registros dos processos de tratamento de dados bem como das providências tomadas para mitigar os riscos gerados para os titulares dos dados (BIONI; LUCIANO, 2020).

Na União Europeia⁵¹, o controlador⁵² tem a obrigação de elaborar um relatório de impacto quando a atividade for qualificada como de alto risco, sendo que, na lista exemplificativa de tais hipóteses, consta o perfilamento⁵³ como auxílio à tomada de decisões, alcançando, por exemplo, “o emprego de IA para automatização de processos de concessão à crédito, precificação de planos e seguro de saúde, seleção e ou recrutamento de candidatos, elegibilidade à programas de assistências social”, entre outras situações rotineiras. Não havendo a indicação de meios para mitigar os riscos da atividade, inicia-se um diálogo entre o agente interessado e o órgão regulador, que deliberará acerca da sua continuidade (BIONI; LUCIANO, 2020, p. 216)

Recentemente, a Comissão Europeia propôs novas regras⁵⁴ voltadas a garantir a confiabilidade de sistemas de IA, elaborando uma classificação de risco dos

⁵¹ Referência ao Regulamento Geral sobre a Proteção de Dados aprovado pelo Parlamento Europeu, que, em seu artigo 35, item 1, estabelece: “Quando um certo tipo de tratamento, em particular que utilize novas tecnologias e tendo em conta a sua natureza, âmbito, contexto e finalidades, for suscetível de implicar um elevado risco para os direitos e liberdades das pessoas singulares, o responsável pelo tratamento procede, antes de iniciar o tratamento, a uma avaliação de impacto das operações de tratamento previstas sobre a proteção de dados pessoais. Se um conjunto de operações de tratamento que apresentar riscos elevados semelhantes, pode ser analisado numa única avaliação” (Disponível em <https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:32016R0679#d1e4426-1-1>. Acesso em 26 nov. 2021).

⁵² Descrito pelos autores como “quem tem poder de tomada de decisão na cadeia de tratamento de dados” (BIONI; LUCIANO, p. 215).

⁵³ O perfilamento é descrito no RGPD da seguinte forma: “Definição de perfis: qualquer forma de tratamento automatizado de dados pessoais que consista em utilizar esses dados pessoais para avaliar certos aspetos pessoais de uma pessoa singular, nomeadamente para analisar ou prever aspetos relacionados com o seu desempenho profissional, a sua situação económica, saúde, preferências pessoais, interesses, fiabilidade, comportamento, localização ou deslocações” (Disponível em <https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:32016R0679#d1e4426-1-1>. Acesso em 26 nov. 2021).

⁵⁴ A União Europeia tem liderado as iniciativas regulatórias para garantir o desenvolvimento de uma IA benéfica e centrada no ser humano. Ainda em 2017, o Parlamento Europeu emitiu uma resolução com recomendações à Comissão de Direito Civil sobre Robótica consignando a fase avançada de desenvolvimento da inteligência artificial e afirmando ser “apropriado começar por problemas de responsabilidade civil”. A resolução também propõe princípios gerais e éticos para o desenvolvimento da tecnologia e registra que tratar da questão da responsabilidade civil por danos causados por robôs é de extrema importância, tendo em vista a necessidade de “garantir o mesmo nível de eficácia, de transparência e de coerência na execução da segurança jurídica em toda a União para benefício dos cidadãos, dos consumidores e das empresas”, manifestando uma preocupação em não “afetar o processo de investigação, de inovação e de desenvolvimento da robótica” Disponível em <[Textos aprovados - Disposições de Direito Civil sobre Robótica - Quinta-feira, 16 de Fevereiro de 2017 \(europa.eu\)](#)>. Acesso em 14 dez. 2021. Em 2018, a Comissão Europeia emitiu a Comunicação da

sistemas que varia entre risco inaceitável⁵⁵, o qual importaria na sua proibição, e o risco mínimo⁵⁶, para o qual a proposta não se aplicaria, ante a ausência de perigos a direitos dos cidadãos. As regras mais detalhadas voltam-se aos sistemas de risco elevado⁵⁷, que passarão por etapas de avaliação desde a sua criação e durante todo o seu ciclo de vida, o que parece ir ao encontro do que Bioni e Luciano consideram como uma transparência qualificada.

Nos Estados Unidos, há uma projeto de lei intitulado *Algorithmic Accountability Act*, de 2019⁵⁸, no qual consta a obrigação de elaboração de relatórios de impacto para situações de alto risco que envolvam a atividade de tratamento de dados pessoais, bem como para a hipótese de emprego de IA para processos de tomada de decisão. O projeto detalha as situações em que o relatório seria obrigatório, bem como as informações que devem constar do documento, que envolvem a descrição detalhada do sistema, bem como dos benefícios e ameaças dele esperados, à luz do seu propósito, e as medidas tomadas para mitigação dos riscos apontados.

Comissão Europeia ao Parlamento Europeu sobre IA, na qual consta uma proposta para que a União Europeia adote uma abordagem coordenada para liderar o desenvolvimento e utilização da inteligência artificial para o bem comum (Disponível em <[https://ec.europa.eu/transparency/documents-register/detail?ref=COM\(2018\)237&lang=pt](https://ec.europa.eu/transparency/documents-register/detail?ref=COM(2018)237&lang=pt)>. Acesso em 14 dez. 2020). Em 2020, a Comissão Europeia publicou o *White Paper* da inteligência artificial, consignando uma abordagem europeia para excelência e confiabilidade da tecnologia. O documento reconhece o caráter estratégico da IA e o seu potencial para fortalecer a competitividade da indústria europeia. A proposta foca no incentivo à inovação ao mesmo tempo em que propõe medidas para uma tecnologia ética e confiável, notadamente, em casos de alto risco, no que se refere a fatores como: dados de treinamento; armazenamento de dados; fornecimento de informações sobre o grau de risco do sistema; robustez e acurácia do sistema; e supervisão humana de acordo com o uso e os efeitos esperados do sistema (Disponível em <https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf>. Acesso em 14 dez. 2021).

⁵⁵ Por exemplo, a realização de uma classificação social por parte de governos e a fabricação de brinquedos com assistentes de voz que induzem comportamentos perigosos (Disponível em <[Excelência e confiança na inteligência artificial | Comissão Europeia \(europa.eu\)](https://ec.europa.eu/artificial-intelligence/white-paper-artificial-intelligence-feb2020_en.pdf)>. Acesso em 14 dez. 2021).

⁵⁶ Por exemplo, jogos de vídeo ou filtros de *spam*. Consta do site da Comissão Europeia que a maioria dos sistemas de IA se inserem nessa categoria (Disponível em <[Excelência e confiança na inteligência artificial | Comissão Europeia \(europa.eu\)](https://ec.europa.eu/artificial-intelligence/white-paper-artificial-intelligence-feb2020_en.pdf)>. Acesso em 14 dez. 2021).

⁵⁷ Por exemplo, sistemas utilizados em infraestruturas críticas, como transportes; no ensino e formação de profissionais; em componentes de segurança de produtos, como em equipamentos para cirurgias assistidas por robôs; em recrutamento de trabalhadores; em serviços públicos e serviços privados essenciais; em medidas repressivas; em processos que envolvem migração, concessão de asilo e controle de fronteiras; na administração da justiça e processos democráticos (Disponível em <[Excelência e confiança na inteligência artificial | Comissão Europeia \(europa.eu\)](https://ec.europa.eu/artificial-intelligence/white-paper-artificial-intelligence-feb2020_en.pdf)>. Acesso em 14 dez. 2021).

⁵⁸ Disponível em <<https://www.congress.gov/116/bills/hr2231/BILLS-116hr2231ih.pdf>>. Acesso em 4 dez. 2021.

No caso brasileiro, a Lei Geral de Proteção de Dados contém menções ao relatório de impacto, porém, a ausência de detalhamento deixa a sua procedimentalização a cargo da regulamentação infralegal (BIONI; LUCIANO, 2020).

Os projetos de lei em tramitação no Congresso Nacional nem chegam a tanto. Quanto ao ponto, o PL nº 21/2020 estabelece o princípio da inovação responsável, segundo o qual os agentes que atuam na cadeia de desenvolvimento e operação de sistemas de IA, devem documentar o seu processo interno de gestão e responsabilizarem-se, nos limites da sua participação, pelos resultados do funcionamento de tais sistemas. Há, ainda, a estipulação da diretriz referente à gestão baseada em risco, segundo a qual, caso seja constatada hipótese de alto risco, poderá a administração pública requerer informações sobre medidas de segurança e prevenção, sem, todavia, declinar o que configuraria tais situações.

Por sua vez, o direito de revisão de decisões automatizadas, relacionado diretamente com o direito à explicação, encontra guarida na legislação europeia e na brasileira. Trata-se de permitir ao interessado compreender a lógica da decisão, de forma que possa aferir qual *input* foi determinante para o resultado apresentado, o que se relacionaria com o princípio da precaução, segundo Bioni e Luciano (2020), na medida em que proporcionaria a correção de problemas em decisões futuras, bem como endereçaria desafios relacionados a segredos comercial e industrial. Vejamos:

A abordagem da precaução parece ser útil na definição dos contornos desse debate. De um lado, ela colabora na construção de espaços de deliberação para se discutir o que seria “informação qualificada” ou como mitigar os problemas em decisões futuras. De outro, ela possibilita endereçar questionamentos a respeito dos segredos comercial e industrial. Ao exigir informações sobre a racionalidade de uma decisão específica, o direito à explicação não se confunde com a transparência pura e simples. Variações nos dados de raça, por exemplo, já poderiam fornecer o impacto e a maneira como esse tipo de dado impacta uma decisão, sem, contudo, demandar a revelação de todo o sistema automatizado envolvido naquela decisão (Doshi-Velez e Kortz, 2017). Ademais, espaços deliberativos com a participação de diversos atores podem ajudar a mitigar os custos envolvidos em sistemas de explicação - que, de outra forma, poderia afetar desproporcionalmente empresas menores - bem como os desafios tecnológicos de se pensar esse tipo de sistema. (BIONI; LUCIANO, 2020, p. 218)

Os autores fazem uma ressalva específica em relação às tecnologias de reconhecimento facial, as quais seriam um caso de fronteira para o princípio da precaução no âmbito da inteligência artificial, ante as incertezas quanto aos benefícios e riscos do seu uso. Embora a LGPD classifique dados biométricos como dados

peçoais sensíveis, não atribui deveres específicos aos desenvolvedores de tecnologias de reconhecimento e tampouco proceduraliza mecanismos obrigatórios de avaliação de riscos, deixando espaço para uma regulação setorial que estabeleça padrões de conformidade para os agentes do setor e demandando, ainda, procedimentos que permitam a deliberação pública a respeito da sua adoção (BIONI; LUCIANO, 2020).

Sob o contexto apresentado, Bioni e Luciano (2020) sugerem dois vetores de regulação para a inteligência artificial, impulsionados pelo princípio da precaução e inspirados nas leis de proteção de dados, quais sejam, (i) a implementação de um debate regulatório que inclua todos os atores envolvidos com a tecnologia, desde desenvolvedores até as pessoas que sofrerão seus efeitos, o que seria corolário de um sistema democrático que reconheça suas assimetrias de poder e de informação, e (ii) a estipulação de obrigações que mitiguem as incertezas relacionadas aos riscos e benefícios decorrentes do uso da IA, de sorte a instrumentalizar a decisão sobre sua adoção.

Um debate regulatório no campo da IA traz consigo dificuldades associadas às características da tecnologia, bem como dos atores envolvidos em seu desenvolvimento, o que, ao invés de contrariar as conclusões apresentadas acima, reforçam a necessidade de se democratizar as discussões sobre o tema, fortalecendo a proteção dos direitos e garantias fundamentais individuais e coletivos e, em última instância, a capacidade do Estado em fazê-lo, conforme será abordado no item seguinte.

4.2. Regulação *ex ante* da inteligência artificial: dificuldades e oportunidades

Esta parte do trabalho abordará as dificuldades de se estipular uma regulação de inteligência artificial que garanta que os pesquisadores e desenvolvedores internalizem práticas seguras e que guardem observância a valores de ordem pública. Sem a pretensão de exaurir a discussão ou apresentar uma proposta definitiva para a questão, as linhas seguintes consistirão em um esforço descritivo para se

compreender o problema e possíveis caminhos a serem trilhados pelos agentes do setor e pelo poder público.

Cada vez mais, fala-se da incapacidade das estruturas políticas tradicionais em oferecer explicações sobre a ordem mundial contemporânea, na qual o conceito de soberania nacional já não exerce tanta importância. Assim, tem-se constatado o aumento da preponderância de uma pluralidade de organizações não estatais, muitas delas de natureza privada, que logram em regular suas próprias relações e têm liberdade para tomar decisões que repercutem na vida dos indivíduos. Nesse sentido, menciona Gunther Teubner (2019) que a globalização não cumpriu a expectativa kantiana de uma sociedade mundial regida por uma política internacional – entre Estados - uniforme, revelando um processo fragmentado no qual a política não tem um papel preponderante, diante do surgimento de uma pluralidade de subsistemas sociais com dinâmicas internas próprias.

A ordem global tem sido descrita por muitos teóricos como detentora de um caráter plural, no sentido de que não existe qualquer tipo de organização que ocupe o centro do sistema político mundial, abrindo espaço para “uma multiplicidade de autodescrições” (HOLMES, 2014, p. 1140) nos âmbitos político e jurídico da sociedade. Encontramo-nos em um contexto em que organizações não governamentais, movimentos sociais e corporações regulam suas próprias esferas de atuação em nível transnacional⁵⁹, instituem mecanismos de resolução de conflitos e exercem autoridade fora da estrutura de soberania do Estado e sem observância à lógica das relações internacionais ou de uma ordem constitucional específica (HOLMES, 2014).

Segundo Pablo Holmes (2014), fala-se mais em governança do que em governo, sendo que, em nível global, o fenômeno remeteria ao conjunto de muitos mecanismos pelos quais instituições – públicas e privadas – e indivíduos administram seus assuntos comuns e que, em uma esfera permeada por meios de comunicação de massa de alcance mundial, acabariam por vincular, ainda que indiretamente, as relações ocorridas no interior dos Estados nacionais.

⁵⁹ “Regimes jurídicos transnacionais, por exemplo, constituem-se frequentemente ao redor de organizações privadas ou de jurisdições privadas, desenvolvendo a pretensão de regular, em nível global, um setor específico de processos sociais, sem ser limitados pelas fronteiras territoriais nacionais” (HOLMES, 2014, p.1.142).

Quanto ao ponto, Holmes (2014, p. 1146) observa um aumento de importância de atores privados no exercício de autoridade e poder, o que poderia ser observado pela tendência de atrair instrumentos de governança corporativa para as técnicas de regulação transnacional e até mesmo nacional, ou seja, uma forma de “privatização” do vocabulário regulatório em nível global. O autor exemplifica que, “[e]m lugar de mecanismos públicos de controle, fala-se em *accountability*; e em lugar de uma negociação política de interesses, fala-se em *joint decision making*” (2014, p. 1151).

O autor descreve, também, uma “desformalização” dos regimes jurídicos do direito global, mediante normas que deixam ampla margem de escolha dos intérpretes, através de cláusulas em branco a serem preenchidas pela atuação técnica de especialistas.

Além disso, Holmes (2014) chama atenção para o fenômeno de formação de redes de agentes nacionais e internacionais que formam comunidades fragmentárias no intuito de produzir e trocar informações acerca de assuntos regulatórios pelo mundo. Trata-se de um cenário em que uma pluralidade de atores tem controle recíproco sobre seu poder regulatório, sem que exista uma entidade central a exercer um poder decisório definitivo.

Este panorama leva à constatação de enfraquecimento da noção de democracia, pois todos esses mecanismos de autorregulação desconsideram procedimentos de participação e inclusão, ignoram a noção de separação de poderes e importam no exercício de atribuições públicas por agentes privados (HOLMES, 2014).

Essa transformação vem associada a um esvaziamento das instâncias políticas baseadas em processos democráticos e, do ponto de vista do direito, faz com que setores regulatórios globais (transnacionais e supranacionais), antes vistos como legítimos apenas por estarem acoplados às práticas, vocabulários e organizações do sistema político, passem também a ser descritos por meio da linguagem do direito privado. (HOLMES, 2014, p. 1146).

A relação entre autorregulação e democracia é uma preocupação quando se trata do setor da tecnologia. Trata-se de um setor dominado por poucas e grandes empresas, as quais detêm poder econômico, acesso a governos e legisladores e capacidade de financiamento de setores relevantes para a formação da opinião pública (NEMITZ, 2018). Além disso, o setor da tecnologia também expõe a desigualdade entre países considerados desenvolvidos e aqueles em

desenvolvimento, pois os primeiros contam com recursos financeiros e humanos e influência sobre a comunidade internacional que os tornam aptos a dominar o mercado e a competir para exercer influência em países que precisam de *hardware* e *software* estrangeiros. Ao mesmo tempo, países em desenvolvimento vivenciam desigualdade digital e dificuldades na área de educação e pesquisa, tornando-se dependentes da tecnologia de outros países e vulneráveis à coleta indiscriminada de dados de seus cidadãos e governos e à submissão de suas populações a conteúdos direcionados (PINTO, 2019).

Some-se a isso o fato de que tais empresas manifestam aversão a legislações que busquem regular suas atividades, sob a perspectiva de que governos e parlamentos não as compreenderiam suficientemente. Por outro lado, elas também não demonstram afinidade com o funcionamento da democracia, do Estado de Direito e nem da proteção de direitos fundamentais. Nestes termos, a esfera pública atualmente proporcionada pela tecnologia dificulta a construção de políticas que contemplem a complexidade da sociedade moderna mediante um debate inclusivo e democrático (NEMITZ, 2018).

Para Paul Nemitz (2018), é improvável que a tendência de autorregulação apresentada por corporações privadas que atuam em nível transnacional leve a uma comunhão de princípios éticos e morais que inspirem uma conformidade com o interesse público, sendo que questões que tenham impacto em direitos fundamentais das pessoas ou tenham relevância no âmbito estatal devem ser submetidas à legitimação democrática da lei. Isso porque “[a]penas a política, a democracia e a participação ativa (positiva) de grupos sociais concretos tornaram possíveis novos graus de inclusão social, bem como possibilitaram a amplificação do que deveria ser considerado como igual em um sentido jurídico” (HOLMES, 2014, p. 1158).

Por sua vez, Gilmar Ferreira Mendes e Victor Oliveira Fernandes (2020, p. 5), discorrendo sobre tema análogo, qual seja, a afirmação e proteção de direitos fundamentais no ciberespaço, suscitam a ideia de Constitucionalismo Digital, descrevendo-o como “a marca de uma verdadeira ideologia constitucional que se estrutura em um quadro normativo de proteção dos direitos fundamentais e de reequilíbrio de poderes na governança do ambiente digital”.

Os autores reconhecem a coexistência de diferentes categorias regulatórias utilizadas por governos, sociedade e mercado visando a disciplina dos ambientes

virtuais. Partem da perspectiva de que, ao passo em que o Estado Nacional tem força impositiva para mediar a fruição de direitos e garantias fundamentais, existem outras modalidades de conformação de direitos provenientes da atuação de atores privados ou de mecanismos transnacionais de governança da internet. Nesse sentido, os autores identificam iniciativas públicas e privadas cujo intuito é prestigiar, de forma abrangente, direitos fundamentais dos usuários da internet, como leis em sentido formal, declarações de organizações intergovernamentais e regulamentos de uso de plataformas digitais (MENDES; OLIVEIRA FERNANDES, 2020).

Embora o Estado Nacional não atue de forma exclusiva, a proteção do espaço digital atinge sua máxima efetividade, segundo Mendes e Oliveira Fernandes (2020, p. 8), mediante a edição de legislações que formalizam os direitos dos usuários da internet, destacando que, embora não impliquem em alterações do texto constitucional, “consagram materialmente categorias de direitos, princípios e normas de governança para a internet, limitando drasticamente o poder de autoridades públicas e de atores privados nas suas relações com os usuários”. No Brasil, o Marco Civil da Internet e a Lei Geral de Proteção de Dados seriam exemplos de legislação desse calibre.

Os autores relativizam, portanto, as previsões anteriormente descritas de que o Estado não teria mais a capacidade de fazer frente à “complexidade regulatória da sociedade interconectada” (2020, p. 10), considerando que as forças governamentais tradicionais continuam a exercer função relevante na regulação do ciberespaço.

Porém, este não parece ser o caminho adotado pelo legislador brasileiro na regulação da inteligência artificial, ao analisamos a tramitação do Projeto de Lei nº 21/2020, que parece priorizar a adoção voluntária de mecanismos de conformidade e prevenção de riscos e danos por parte dos agentes integrantes da cadeia de desenvolvimento e uso da inteligência artificial.

Por exemplo, como um dos fundamentos para o desenvolvimento e a aplicação da inteligência artificial consta, no artigo 4º, inciso VII, “o estímulo à autorregulação, mediante adoção de códigos de conduta e guias de boas práticas” (BRASIL, 2020).

Ao mencionar a segurança e a prevenção como princípios para o desenvolvimento e aplicação da IA no Brasil, o projeto descreve medidas genéricas das quais as empresas dispõem para gerenciar riscos:

VI – segurança e prevenção: utilização de medidas técnicas, organizacionais e administrativas, considerando a utilização de meios razoáveis e disponíveis na ocasião, compatíveis com melhores práticas, os padrões internacionais e viabilidade econômica, voltadas a permitir o gerenciamento e a mitigação de riscos oriundos da operação de sistemas de inteligência artificial durante todo o seu ciclo de vida e o seu contínuo funcionamento. (BRASIL, 2020)

Chama atenção, ainda a orientação de intervenção subsidiária do poder público ao disciplinar a aplicação da IA (artigo 6º, inciso I).

De se frisar também, como já mencionado, que, apesar da proposição mencionar a possibilidade da regulamentação por parte do Poder Executivo Federal, por órgãos e entidades setoriais com competência técnica na matéria, não haverá disposição legal, caso o texto seja aprovado na forma como está, que estipule procedimentos obrigatórios de prestação de contas, que estabeleça de forma clara as obrigações atribuídas aos agentes do setor e tampouco sanções por eventual descumprimento.

Assim sendo, na medida em que privilegia a adoção de mecanismos de autorregulação e deixa a cargo das empresas escolher os métodos de prevenção de danos que entenderem cabíveis, relegando a atuação estatal à subsidiariedade, a proposta não parece ser um instrumento que, efetivamente, logrará em proteger direitos fundamentais e estabelecer um ambiente seguro de inovação.

Acrescente-se, ainda, que a tramitação do projeto de lei tem sido mais rápida e menos debatida quando comparada com o MCI e a LGPD, legislações que foram aprovadas mediante participação pública e para as quais foi dedicado mais tempo de discussão⁶⁰⁶¹. Quanto ao ponto, nota-se que o PL nº 21/2020 foi aprovado em regime de urgência na Câmara dos Deputados e aguarda tramitação no Senado Federal.

⁶⁰ Conforme afirma Bruno Bioni em entrevista ao Estadão: “Se a gente compara com outros PLs de leis que se tornaram boas leis, a gente vê um processo muito mais acelerado, no qual não foram esgotados todos os instrumentos de participação pública. Apesar desse PL ter contado com audiências públicas, a gente poderia ter tido mais tempo e o texto poderia ter sido colocado para consulta pública, que foi o que aconteceu com a Lei Geral de Proteção de Dados (LGPD) e com o Marco Civil da Internet. Essas foram discussões legislativas, que, dadas suas complexidades, foram necessários anos de discussão. Depende muito de como será o desenho final da lei. Mas, comparando com as outras duas, essa talvez seja a lei mais complexa de todas. Embora seja uma tecnologia que está ganhando dianteira e que já foi adotada no dia a dia, ela ainda vai chacoalhar muita coisa, como o setor econômico e o trabalho. Os temas das outras leis estavam mais maduros. O tema da IA ainda é bastante desafiador. Isso justifica os motivos para ter mais tempo para fazer as discussões adequadas. Quando comparamos iniciativas como essa, tanto na União Europeia quanto na América do Norte, constitui-se uma comissão de especialistas para facilitar o trabalho do parlamento. Isso não aconteceu aqui e seria bem-vindo”. Disponível em <<https://link.estadao.com.br/noticias/cultura-digital,mais-importante-lei-de-tecnologia-no-brasil-nao-esta-sendo-debatida-diz-especialista,70003918886>>. Acesso em 7 dez. 2021.

⁶¹ No mesmo sentido, a Coalizão Direitos na Rede emitiu um documento pelo qual afirma que o Brasil não estaria pronto para regular a inteligência artificial. Sobre o PL nº 21/2020, consta: “Do ponto de

Suscitando discussão análoga à presente, Amélie Heldt (2019) analisa o tema sob a perspectiva da regulação das redes sociais. A autora nota que, após escândalos como o *Cambridge Analytica*, em que se constatou a coleta ilegal de informações no *Facebook* para perfilamento de usuários no intuito de bombardeá-los com propagandas políticas personalizadas e manipular processos eleitorais (CADWALLADR; GRAHAM-HARRISON, 2018), a desconfiança com tais plataformas tem aumentado, sendo que, após um período de passividade de atores privados e estatais, o cenário tem apontado para ações mais efetivas no endereçamento de questões como desinformação e discurso de ódio nas redes sociais.

A regulação da comunicação no âmbito das redes sociais não é trivial, considerando importantes implicações para a liberdade de expressão dos usuários que, de sua parte, estão espalhados pelo mundo e submetidos a ordenamentos jurídicos que conferem tratamentos diversos a esse direito fundamental, enquanto a maioria das plataformas estão inseridas no contexto norte-americano, no qual a Primeira Emenda à Constituição é extremamente deferente ao livre trânsito de ideias (HELDT, 2019).

Recentemente, no intuito de frear o discurso de ódio crescentemente observado nas redes sociais, o Parlamento Alemão aprovou o *Network Enforcement Act* (NetzDG), que obriga as plataformas a removerem conteúdo manifestamente ilegal de seus domínios no prazo de 24 horas, mediante um mecanismo de reclamação dos usuários. A medida levantou questões acerca do poder delegado a tais empresas para “julgar” os casos submetidos à sua apreciação, como uma espécie

vista interno, o atropelo do processo legislativo saudável impede o aprimoramento do texto e a análise de tantos pontos em aberto quanto aos impactos tecnológicos da IA. Além disso, a marcha acelerada, sem consulta pública concreta e profunda aos diferentes atores envolvidos com a tecnologia e afetados por ela, expõe a nova lei a se tornar um exemplo de má prática em termos de índice democrático e participação popular. Em outras palavras, é uma lei que não condiz com os princípios constitucionais brasileiros. Do ponto de vista externo, destacam-se as consequências gerais de ordem econômica. Tal destaque se dá pelo fato de que, seguindo uma tendência geral na regulação da tecnologia, legislações de outros países e uniões econômicas (como a União Europeia) tendem a impor condições de exercício de modelos de negócio em seu território. Uma legislação que não possua critérios em diálogo transnacional e seja ausente de mecanismos adequados de controle e de garantias de direitos pode obstar a certificação e liberação de produtos envolvendo inteligência artificial em mercados estrangeiros. Isso significa dizer que, se aprovado da forma como está, o PL 21-A/2020 pode representar um barreira para que empresas brasileiras de tecnologia vendam seus produtos e serviços para mercados internacionais, bem como desestimulará investidores estrangeiros a estabelecerem ou ampliarem negócios no Brasil” (Disponível em <https://direitosnarede.org.br/2021/12/07/brasil-nao-esta-pronto-para-regular-inteligencia-artificial/>>. Acesso em 14 dez. 2021).

de privatização do judiciário no que se refere à decisão sobre o que seria considerado ilegal nas redes sociais, a ser tomada por atores privados que não têm total domínio da lei nacional (HELDT, 2019).

Na União Europeia, a necessidade de conter a desinformação nas redes sociais, notadamente, para resguardar a higidez de processos eleitorais, levou à edição do *Code of Practice on Disinformation* (CPD), que consiste em um instrumento de autorregulação acordado entre o bloco de países e as principais empresas de tecnologia, no qual constam compromissos e incentivos para a adoção de medidas para combater a disseminação de informações falsas. Para atingir os objetivos propostos, porém, as empresas poderiam operar à sua maneira, escolhendo se e como implementariam as medidas cabíveis, tendo sido observada certa relutância em excluir conteúdo, ante a dificuldade de traçar uma linha divisória entre opinião e desinformação (HELDT, 2019).

Em ambos os exemplos, o Estado delegou às corporações a atribuição de garantir a observância de leis nacionais, conferindo-lhes autoridade, controle e poder sobre indivíduos. Nesse sentido, ainda que a regulação estatal tenha ocorrido, a interpretação sobre o que é ou não desinformação ou discurso de ódio ficou a cargo desses atores privados (HELDT, 2019).

Por outro lado, a autora descreve um movimento curioso observado em relação a empresas que administram redes sociais, como é o caso do *Facebook*, que acabou por elaborar mecanismos que se assemelham ao funcionamento de órgãos estatais, ao estipular regras de comportamento, garantir seu cumprimento e punir os usuários que não as observam, ao mesmo tempo em que garante o devido processo, a recorribilidade das decisões e confere transparência aos atos praticados. Trata-se, por óbvio, de uma semelhança meramente procedimental, mas que denota uma maior atenção dessas corporações à pressão social para que tomem as rédeas da nova esfera pública que ajudaram a criar (HELDT, 2019).

As razões para a adoção de regulações internas, como a implementada pelo *Facebook*, podem ser muitas, mas o efeito mais evidente é o arrefecimento da intervenção estatal no âmbito regulatório e fiscalizatório (HELDT, 2019).

Desse contexto, Heldt (2019) extrai que, de um lado, observamos a delegação do poder de interpretação da lei do Estado para as plataformas e, de outro, a adoção

de mecanismos quase administrativos por parte das corporações para combater a desinformação e o discurso de ódio nas redes sociais.

São duas as conclusões apresentadas pela autora: Estados e corporações parecem perseguir os mesmos objetivos, ainda que sob razões distintas, e, ademais, parecem mover-se de forma convergente, encontrando-se no meio do caminho (*meeting halfway*) em uma relação de cooperação e compartilhamento de responsabilidades. Quanto ao segundo ponto, embora seja contraintuitivo presumir que, em um contexto de autorregulação, uma corporação venha a adotar regras comumente associadas ao Estado, como o devido processo legal e a legalidade, a incorporação de tais valores tornaria a atividade mais aderente ao interesse público e mais atrativo ao público em termos de confiança e transparência.

Um sistema de governança colaborativa⁶² entre atores públicos e privados, para Margot E. Kaminski (2019), seria interessante no caso de tecnologias complexas como a IA, porém, demandaria um regulador com poder efetivo e um sistema de *accountability* não apenas em face do Estado, mas de representantes de todos que possam ser afetados pela atividade, de forma a contemplar tanto problemas sistêmicos, como vieses discriminatórios⁶³, quanto aqueles relacionados a direitos individuais, como a explicabilidade de determinada decisão automatizada. Por *accountability*, a autora entende um contexto em que decisões possam ser legitimadas pelo público, além do reforço da cultura de *compliance* e a adoção de múltiplos mecanismos que possibilitem a análise de especialistas e a fiscalização dos três estágios de governança: estabelecimento de regras, implementação e execução/coação.

Kaminski (2019) considera a *General Data Protection Regulation* (GDPR) um exemplo de governança colaborativa⁶⁴, já que a norma europeia prevê, além do direito individual à explicabilidade das decisões automatizadas, um regime de *compliance* aplicado às empresas privadas, de forma a influenciar no *design* dos algoritmos de

⁶² A autora aponta que a governança colaborativa não seria o mesmo que autorregulação ou desregulação. Nesse sentido, ainda haveria espaço para uma regulação prescritiva, porém, mesclada com mecanismos de *soft law*, de forma que o setor privado seja incentivado a participar da atividade em conformidade com o interesse público.

⁶³ A autora considera os vieses discriminatórios como problemas sistêmicos, considerando que são mais facilmente detectados quando o foco se volta ao comportamento do sistema como um todo, ao invés de analisar cada decisão de forma específica.

⁶⁴ Embora reconheça que, na maior parte, o GDPR é um instrumento de *hard law*, ou seja, um regime obrigatório com fortes mecanismos coercitivos, a exemplo da estipulação de multas administrativas.

forma sistêmica. Trata-se de tornar as decisões tomadas por algoritmos responsabilizáveis quando afetam indivíduos, além de responsabilizar as empresas pelo estabelecimento de normas, políticas e técnicas e pelo dever de possibilitar a fiscalização da atividade. A qualidade colaborativa desse tipo de governança residiria justamente na estipulação de obrigações gerais às empresas, cujo significado é complementado por meio de códigos de conduta ou por parâmetros estipulados por agentes internos à corporação.

Por outro lado, a cooperação entre atores públicos e privados no campo da inteligência artificial não é uma solução fácil para Alicia Solow-Niederman (2020).

Sobre a regulação da inteligência artificial, de início, a autora pondera que, se, de um lado, não está claro se devemos confiar na autorregulação como um indicativo de que as empresas de tecnologia irão evitar riscos e danos, de outro, uma regulação estatal tradicional (*top-down regulation*) não seria uma panaceia para os problemas da IA.

Para tanto, afirma que a rapidez com que a IA tem sido desenvolvida gera um problema para que o legislador acompanhe o ritmo com que os algoritmos podem ser aprimorados e alterados. Ademais, uma regulação prescritiva precisaria endereçar a complexidade da tecnologia, atentando-se para a dificuldade de interpretação dos processos pelos quais a IA chega a determinados resultados e à especialização necessária para compreender os meandros dessa tecnologia. Por fim, haveria o desafio da imprevisibilidade, ante a dificuldade de prever eventuais acidentes e pela capacidade de adaptação dos algoritmos em razão da interação com novos estímulos. Assim, tais características tornariam a IA única e desafiariam uma abordagem universal de comando e controle para a sua regulação, tendo em vista que nem sempre será possível estabelecer relações de causa e efeito para tais sistemas complexos.

Aventando uma possível solução de governança colaborativa entre governo e empresas, a autora chama atenção ao fato de que, no campo tecnológico, o setor privado supera o setor público em termos de expertise e de recursos. Quanto ao ponto, destaca que uma relação de cooperação demandaria certa simetria entre as partes, o

que não ocorre no caso, já que não há uma entidade estatal capaz de representar o interesse público no âmbito dessa combinação de forças⁶⁵.

Para Solow-Niederman (2020), novas soluções de governança demandam novas concepções do que se entende por política pública no contexto de tecnologias digitais emergentes, considerando que, ao lidarmos com inteligência artificial, a programação é a política, já que a forma com que os algoritmos funcionam envolvem escolhas que acabam por influenciar a vida dos cidadãos, impactando em valores fundamentais como privacidade, liberdade e até a segurança física.

Por exemplo, a autora menciona uma ferramenta elaborada pelo *Google* denominada *TensorFlow Federated*, a qual, mediante código aberto, oferece um sistema para treinar outros aplicativos de *machine learning* sem que os dados utilizados pelo sistema central sejam expostos ao algoritmo em treinamento, o que proporciona um tratamento de dados mais sensível à privacidade dos seus titulares. A decisão sobre utilizar ou não essa ferramenta, porém, é objeto de liberalidade das empresas de tecnologia e passa por decisões técnicas e corporativas que, efetivamente, determinarão a forma com que a IA irá interagir com os valores humanos.

Nesse sentido, o desafio apresentado diz respeito ao fato de que, em grande medida, a IA é dominada por esses atores privados que têm conhecimento e expertise para tomar decisões que impactam na vida das pessoas, sem qualquer tipo de controle democrático das opções adotadas.

⁶⁵ A autora fala a partir da experiência norte-americana. No Brasil, há importantes iniciativas de pesquisa e desenvolvimento de IA por parte de órgãos públicos, notadamente, do Poder Judiciário. Conforme explica Fernanda Carvalho Lage (2021, p. 141): “O investimento em inteligência artificial está no plano de gestão das altas Cortes brasileiras. Note-se, por exemplo, que tanto o Supremo Tribunal Federal (STF) como o Superior Tribunal de Justiça (STJ) já possuem soluções de IA para realização de tarefas relacionadas ao processo digital. Como será analisado neste capítulo, no STF está em funcionamento o Projeto Victor. A solução de machine learning, desenvolvida integralmente pela Universidade de Brasília (UnB), desempenha quatro funções nos processos digitais: conversão de imagens em textos; separação do começo e do fim dos documentos de acervo; classificação das peças processuais mais usadas; e identificação de temas de maior repercussão. Já no STJ, está em operação na Secretaria Judiciária para automatizar a etapa da classificação processual, que antecede a distribuição. O sistema é capaz de ler o teor dos processos e apontar o tema com um percentual de acerto de 86%. Hoje, 100% dos processos que entram no STJ passam por este algoritmo. O objetivo de ambas as aplicações de inteligência artificial é diminuir o trabalho repetitivo e melhorar a gestão do tempo dos processos. O panorama brasileiro do uso da IA no campo jurídico é altamente promissor, mas dependerá do compartilhamento de experiências pelos Tribunais para que sejam reduzidos custos e esforços repetitivos pelas Cortes. Com o crescimento das empresas de Legaltech, haverá a adaptação cada vez maior das rotinas de trabalho por meio da inteligência artificial, pois a questão não é se ela irá substituir juízes, promotores e advogados, mas como irá afetar a maneira como eles trabalham”.

Mas, então, que papel poderia ter o Estado para que o desenvolvimento da IA equilibre os interesses comerciais com valores de ordem pública? Sem apresentar propostas definitivas para o desafio suscitado, Solow-Niederman (2020) considera que o assunto seria mais bem endereçado se, ao invés de estabelecer um controle vertical e generalizado sobre atores privados, o poder público proporcione um diálogo inclusivo que permita estabelecer os valores e preocupações em jogo pelo uso da IA em contextos determinados, para então proceder a uma escolha regulatória. A pergunta a ser respondida seria: em que condições estaríamos confortáveis em delegar escolhas que afetam a coletividade a empresas privadas?

Por exemplo, no âmbito da justiça criminal, a inserção de atores motivados pela obtenção de lucro é preocupante, por se tratar de setor que decide sobre a liberdade de indivíduos, razão pela qual a discussão deve abordar se devemos utilizar a IA nesse âmbito e quais seriam as salvaguardas públicas para a atividade. No campo dos veículos autônomos, seria relevante aferir quais métricas e resultados poderiam ser considerados aceitáveis; se os produtos disponibilizados ao público seriam suficientemente seguros para a livre escolha do consumidor ou se seria necessária uma supervisão prévia das decisões de programação (SOLOW-NIEDERMAN, 2020).

A autora sugere o investimento público em pesquisa e desenvolvimento de IA, de forma que o Estado possa figurar como um ator competitivo que acrescenta motivações não econômicas para a agenda tecnológica.

Voltando-se para a importância dos dados para o funcionamento da IA, Solow-Niederman (2020) defende o investimento estatal na criação de bancos de dados públicos que possam ser utilizados para treinar algoritmos, como uma forma de garantia institucional de que haveria um cuidado na mitigação de vieses discriminatórios⁶⁶. Embora não elimine preocupações com as escolhas técnicas adotadas por cada empresa, essa medida poderia permitir mais controle público no desenvolvimento da tecnologia.

Além disso, há a possibilidade de que o Estado avalie, no mercado, os *datasets* que observam parâmetros voltados ao interesse público, incentivando o seu

⁶⁶ Cabe ressaltar que nada garante que o Estado também não terá os seus vieses discriminatórios. Todavia, a capacidade de, em tese, proporcionar espaços deliberativos e procedimentos democráticos de tomada de decisão, além da necessária observância a princípios da Administração Pública, como moralidade, publicidade e transparência, a existência de um sistema de freios e contrapesos e o imperativo constitucional de respeito aos direitos fundamentais e da busca por uma igualdade material podem ser fatores que venham a mitigar manifestações discriminatórias pelo uso de tais *datasets*, quando a curadoria for realizada por um órgão estatal.

compartilhamento e uso por outros agentes. Subsequentemente, a autora aventa a possibilidade de se estipular sanções caso um desenvolvedor opte por não usar *datasets* aprovados pelo Estado e, por esse motivo, vier a oferecer riscos ou causar danos a terceiros. Em suma, investimentos em bancos de dados previamente aprovados poderia levar ao desenvolvimento de produtos de IA mais robustos e seguros para a coletividade e, além disso, direcionar o mercado a um comportamento que não é orientado exclusivamente pelo lucro.

Uma análise da forma com que diferentes países abordam a IA parece convergir com a conclusão acima, acrescentando-se repercussões geopolíticas à questão. No cenário mundial, observa-se um contraste entre duas abordagens regulatórias sobre a IA, quais sejam, a norte-americana e a europeia.

A abordagem norte-americana fundamenta-se, em grande medida, na autorregulação, sendo que a pesquisa e desenvolvimento em IA é representada principalmente pelo setor privado, notadamente, aquelas cinco empresas que dominam o campo tecnológico⁶⁷. O principal concorrente dos Estados Unidos nesse quesito é a China, que, por sua vez, tem iniciativas do Estado para fortalecer uma IA nacional, bem como um plano de desenvolvimento com etapas a serem vencidas até o ano de 2030⁶⁸ (SCIPIONE, 2020).

A União Europeia, por sua vez, tem se destacado pelo aprofundamento das discussões sobre abordagens éticas à IA, focadas em um desenvolvimento da tecnologia centrada nos seres-humanos e cuja utilização possa ser confiável. Não obstante, o uso de inteligência artificial no âmbito europeu é, em grande parte, dependente da tecnologia e infraestrutura proveniente dos líderes mundiais nesse campo – Estados Unidos e China –, o que tende a aumentar ainda mais a escala de acesso desses atores a dados e, conseqüentemente, ampliar a sua vantagem perante os demais (WOLFF, 2020). Como afirma Guntram Wolff⁶⁹, uma abordagem baseada

⁶⁷ A ideia apresentada é a de que o governo norte-americano não é considerado um ator preponderante no desenvolvimento dessas tecnologias, em comparação com as empresas privadas que dominam o setor (SCIPIONE, 2020).

⁶⁸ Referência ao *Artificial Intelligence Development Plan* (AIDP), lançado em 2017, que visa tornar a China o centro mundial de inovação até o ano de 2030 e inserir a IA no processo industrial e de transformação econômica. O documento estipula etapas que visam ampliar a competitividade da China perante os outros atores globais, ampliar a indústria nacional de IA, despontar em aspectos teóricos, liderar o desenvolvimento de funcionalidades específicas, endereçar questões éticas e regulatórias e, enfim, tornar-se o centro mundial de inovação em IA (ROBERTS et al, 2019).

⁶⁹ Diretor do *think tank* Bruegel, sediado em Bruxelas, que conduz pesquisas em diversas áreas com o intuito de melhorar o debate econômico e a formulação de políticas públicas.

apenas no estabelecimento de regras para o jogo coloca a União Europeia na condição de árbitro, o que a exclui da competição enquanto ator internacional relevante⁷⁰.

O que se extrai do que foi exposto é a noção de que a busca por uma inteligência artificial segura, robusta e que, conseqüentemente, não cause danos às pessoas, não pode depender apenas da boa vontade de grandes corporações, o que demanda a atuação de agentes públicos no sentido de orientar um desenvolvimento benéfico da tecnologia e garantir salvaguardas que protejam direitos fundamentais dos cidadãos. Todavia, há claras dificuldades em fazê-lo mediante técnicas tradicionais de regulação, o que faz suscitar a necessidade de que governos atuem em conjunto com o setor privado, onde se concentram os recursos e a expertise sobre o tema, para encontrar uma solução que atenda às especificidades dessa tecnologia. Antes, porém, se faz necessário reduzir a assimetria entre governos e corporações, de forma que o Estado possa figurar como um ator relevante e capaz de induzir a forma com que se dará o futuro da inteligência artificial.

⁷⁰ O termo utilizado foi “*referees don’t win*”.

CONCLUSÃO

A hipótese proposta como resposta ao problema de pesquisa parece ter sido confirmada, ao menos, parcialmente. Ou seja, a constatação que se chega ao final do trabalho é a de que a responsabilização por danos causados pela inteligência artificial é possível à luz do ordenamento vigente, o que se afirma após o estudo do Direito Civil sob uma perspectiva solidarista e voltada à tutela da pessoa humana, enquanto imperativos constitucionais.

De início, registrou-se que uma perspectiva Civil-Constitucional demanda uma visão sistemática do ordenamento jurídico, considerado em sua unidade e completude, sob a orientação máxima da Constituição. Reconhece-se, assim, que a fruição de direitos e garantias fundamentais não depende necessariamente da atuação do legislador infraconstitucional, cabendo ao intérprete absorver os princípios e valores emanados pelo texto constitucional para concretizá-los.

O efeito dessa perspectiva sobre a responsabilidade civil se traduz no reconhecimento de assimetrias de poder na sociedade e na ideia de que a coletividade deve proteger seus vulneráveis em prol do bem comum, que, em última instância, é dever e responsabilidade de todo indivíduo. Dessa forma, firma-se um ideário solidarista comprometido com a tutela da vítima de danos e com a formação de um ambiente social mais seguro para todos.

Sob essa lógica, ao longo do tempo, relativizaram-se os requisitos que configuravam o dever de indenizar. Primeiramente, verificou-se que, quando o regime subjetivo de responsabilidade se mostrar como um empecilho à reparação das vítimas de danos, diante da dificuldade probatória a ele associada, a teoria do risco poderá ser suscitada como fundamento para a imputação de responsabilidade por danos causados pela inteligência artificial. Por sua vez, a dificuldade de se estabelecer um liame entre o dano causado e alguma pessoa integrante do processo de desenvolvimento ou funcionamento da máquina autônoma também restou superado pelo estudo da teoria da presunção de causalidade, aplicável a hipóteses de responsabilidade objetiva pelo fator de risco e de danos difusos.

Para tanto, nada de inédito foi suscitado, restando demonstrado que a autonomia de máquinas dotadas de inteligência artificial não seria um óbice

intransponível à configuração do dever de indenizar, não sendo, portanto, necessária uma nova regulação que venha, de forma específica, disciplinar o regime de responsabilidade aplicável a danos dessa natureza.

Porém, o legislador brasileiro parece movimentar-se em sentido contrário, o que se conclui pela análise das proposições legislativas que buscam regular a responsabilidade civil em matéria de inteligência artificial em caráter específico e assistemático. Quanto ao ponto, chama atenção o rápido avanço do trâmite do Projeto de Lei nº 21/2020, que, por seu turno, estipula o regime subjetivo de responsabilidade por danos causados pela inteligência artificial, o que vai ao encontro da perspectiva Civil-Constitucional descrita neste trabalho, por obstar, mediante prova de difícil produção, a devida reparação à vítima.

Mas a responsabilidade civil, traduzida no dever de não lesar a outrem, não se encerra em seu caráter reparatório, considerando que a efetiva proteção de direitos fundamentais demanda um esforço coletivo para que se evite a ocorrência de danos e para que se mitigue os riscos das atividades. Nesse sentido, ao passo em que não necessitamos de uma regulação específica para que a vítima de danos causados pela inteligência artificial possa ser indenizada, há um debate regulatório a ser encarado pelo Estado brasileiro no que tange ao estabelecimento de obrigações de conformidade ao setor da tecnologia, de forma a condicionar comportamentos para o desenvolvimento e uso benéfico da IA.

Para tanto, decisões coletivas e informadas acerca de quais riscos estamos dispostos a tolerar devem ser tomadas, a partir de uma pluralidade de perspectivas, para que a regulação da IA seja fruto de um debate democrático e inclusivo.

O caminho para a mencionada regulação parece estar pavimentado por outras legislações que implementaram mecanismos para proteger direitos fundamentais dos cidadãos em temas que tangenciam ao presente, com especial relevância para a Lei Geral de Proteção de Dados (LGPD), que, além de dispor de instrumentos que são aplicáveis ao campo da inteligência artificial, inspira uma disciplina que contemple as assimetrias entre seus destinatários e fomente uma cultura de prestação de contas e transparência.

O entendimento da forma com que se se dará a regulação da inteligência artificial, porém, demanda a compreensão do contexto global dessa tecnologia, a qual é desenvolvida, em grande medida, por poucas e grandes empresas transnacionais

que logram influenciar o processo de formação das normas às quais serão submetidas, quando não o fazem elas mesmas as próprias regras. Questiona-se, assim, acerca da capacidade do Estado em garantir a proteção de direitos fundamentais dos cidadãos e a garantia da observância de valores de ordem pública, considerando a preponderância, no campo da IA, de atores privados que têm conhecimento e expertise para tomar decisões que interferem na vida das pessoas, sem qualquer controle democrático acerca das opções adotadas.

Ademais, as características da tecnologia em questão dificultam uma regulação tradicional, de ordem prescritiva, ante a sua complexidade técnica, a rapidez do seu desenvolvimento e a forma com que os algoritmos se adaptam e interagem com o ambiente.

Nesse contexto, delineia-se a necessidade de uma relação de complementaridade entre os setores público e privado, não sem um debate plural que nos permita decidir, enquanto sociedade, quais decisões poderiam ser delegadas para empresas privadas e quais as salvaguardas públicas necessárias para tanto. Paralelamente, a capacidade do Poder Público de garantir um desenvolvimento e uso benéfico da IA parece depender da redução da assimetria entre governos e corporações, mediante a inserção Estado como ator competitivo no setor tecnológico, para que possa induzir a prevalência de motivações não econômicas nesse campo.

REFERÊNCIAS BIBLIOGRÁFICAS

ALBUQUERQUE JÚNIOR, Roberto Paulino de. Notas sobre a teoria da responsabilidade civil sem dano. **Revista de Direito Civil Contemporâneo**. vol. 6. ano 3. p. 89-103, 2016.

ALENCAR, Felipe. **O que é a Lei de Moore?** Entenda a teoria que 'prevê' futuro da Informática. TechTudo, 2015. Disponível em <https://www.techtudo.com.br/noticias/noticia/2015/06/o-que-e-lei-de-moore-entenda-teoria-que-preve-futuro-da-informatica.html>. Acesso em 26 abr. 2021.

BARFIELD, Woodrow. Liability for autonomous and artificially intelligent robots. **Paladyn, Journal of Behavioral Robotics**, v. 9, n. 1, p. 193-203, 2018.

BARRILAO, Juan Francisco Sánchez. El Derecho constitucional ante la era de Ultrón: la informática y la inteligencia artificial como objeto constitucional. **Estudios de Deusto: revista de la Universidad de Deusto**, v. 64, n. 2, p. 225-258, 2016.

BARTNECK, Christoph et al. **An introduction to ethics in robotics and AI**. Springer Nature, 2021.

BATHAEE, Yavar. The artificial intelligence black box and the failure of intent and causation. **Harvard Journal of Law and Technology**, v. 31, p. 889-938, 2018.

BECK, Ulrich. **Risk Society: towards a new modernity**. London: Sage Publications, 1992.

BIONI, Bruno, DIAS, Daniel. Responsabilidade civil na proteção de dados pessoais: construindo pontes entre a Lei Geral de Proteção de Dados Pessoais e o Código de Defesa do Consumidor. **civilistica.com**, v. 9, n. 3, p. 1-23, 2020.

BIONI, Bruno; LUCIANO, Maria. O Princípio da Precaução na Regulação de Inteligência Artificial: seriam as leis de proteção de dados o seu portal de entrada? *In*: FRAZÃO, Ana; MULHOLLAND, Caitlin. (Coord.). **Inteligência Artificial e Direito: Ética, Regulação e Responsabilidade**. 2. ed. São Paulo: Revista dos Tribunais, p. 205-229, 2020.

BOBBIO, Norberto; **Teoria do ordenamento jurídico**. Brasília: Editora UnB, 6ª ed., 1995.

BOTELHO, José Franciso. **Prometeu, o Inventor da Humanidade**. Superinteressante. 2019. Disponível em <https://super.abril.com.br/historia/prometeu-o-inventor-da-humanidade/>. Acesso em 11 mar. 2021.

BRASIL. **Enunciado nº 43**. I Jornada de Direito Civil. Brasília: Conselho da Justiça Federal, 2002. Disponível em: <<https://www.cjf.jus.br/enunciados/enunciado/720>>. Acesso em 13 ago. 2021.

BRASIL. **Enunciado nº 446**. V Jornada de Direito Civil. Brasília: Conselho da Justiça Federal, 2011. Disponível em: < <https://www.cjf.jus.br/enunciados/enunciado/371>>. Acesso em 6 nov. 2021.

BRASIL. **Lei nº 8.078, de 11 de setembro de 1990**. Código de Defesa do Consumidor. Disponível em http://www.planalto.gov.br/ccivil_03/leis/l8078compilado.htm. Acesso em 8 jul. 2021.

BRASIL. **Lei nº 10.406, de 10 de janeiro de 2002**. Código Civil. Disponível em http://www.planalto.gov.br/ccivil_03/leis/2002/l10406compilada.htm. Acesso em 8 jul. 2021.

BRASIL. **Lei nº 12.965, de 23 de abril de 2014**. Marco Civil da Internet (MCI). Disponível em <http://www.planalto.gov.br/ccivil_03/ato2011-2014/2014/lei/l12965.htm>. Acesso em 3 dez. 2021.

BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados (LGPD). Disponível em: http://www.planalto.gov.br/ccivil_03/ato2015-2018/2018/lei/l13709.htm. Acesso em 19 out. 2021.

BRASIL. **Projeto de Lei do Senado Federal nº 5.051, de 2019**. Disponível em: <<https://www25.senado.leg.br/web/atividade/materias/-/materia/138790>>. Acesso em 10 ago. 2021.

BRASIL. **Projeto de Lei da Câmara dos Deputados nº 21/2020**. Disponível em: <<https://www.camara.leg.br/propostas-legislativas/2236340>>. Acesso em 10 ago. 2021.

BUCHANAN, Bruce G. A (very) brief history of artificial intelligence. **AI Magazine**, v. 26, n. 4, p. 53-53, 2005.

BUOLAMWINI, Joy; GEBRU, Timnit. Gender shades: Intersectional accuracy disparities in commercial gender classification. **Conference on fairness, accountability and transparency**, p. 77-91, 2018.

CADWALLADR, Carole; GRAHAM-HARRISON, Emma. Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. **The Guardian**, v. 17, p. 22, 2018.

ČERKA, Paulius; GRIGIENĖ, Jurgita; SIRBIKYTĖ, Gintarė. Liability for damages caused by artificial intelligence. **Computer Law & Security Review**, v. 31, n. 3, p. 376-389, 2015.

DIAS, Laura Osório Bradley dos Santos. Danos Causados por Veículos Autônomos: adequação das respostas contemporâneas às perguntas futuras. *In*: TEPEDINO, Gustavo; SILVA, Rodrigo da Guia (Coord.). **O Direito Civil na Era da Inteligência Artificial**. São Paulo: Thomson Reuters Brasil, p. 627-649, 2020.

EHRHARDT JÚNIOR, Marcos. Responsabilidade Civil ou Direito de Danos? Breves reflexões sobre a inadequação do modelo tradicional sob o prisma do Direito Civil Constitucional. *In*: RUZYK, Carlos Eduardo Pianovisky; SOUZA, Eduardo Nunes de; MENEZES, Joyceanne Bezerra de; EHRHARDT JÚNIOR, Marcos (Org.). **Direito Civil Constitucional**: a ressignificação da função dos institutos fundamentais do direito civil contemporâneo e suas consequências. Florianópolis: Conceito, p. 19-27, 2014.

FACHIN, Luiz Edson; RUZYK, Carlos Eduardo Pianovski. Direitos Fundamentais, dignidade da pessoa humana e o novo Código Civil: uma análise crítica. *In*: SARLET, Ingo Wolfgang (org). **Constituição, direitos fundamentais e direito privado**. Porto Alegre: Livraria do Advogado, 2ª ed., 2006, p. 89-106.

FACCHINI NETO, Eugênio. Reflexões histórico-evolutivas sobre a constitucionalização do direito privado. *In*: SARLET, Ingo Wolfgang (org). **Constituição, direitos fundamentais e direito privado**. Porto Alegre: Livraria do Advogado, 2ª ed., 2006, p. 13-62.

FALEIROS JÚNIOR, José Luiz de Moura. A culpa presumida como viés intermediário entre a teoria da culpa e a teoria do risco. **Revista Amagis Jurídica**, [S.I.], n. 11, p. 243-257, ago. 2019. ISSN 2674-8908. Disponível em: <<https://revista.amagis.com.br/index.php/amagis-juridica/article/view/94>>. Acesso em: 05 ago. 2021.

FIRTH-BUTTERFIELD, Kay. Artificial Intelligence and the Law: More questions than answers? **SciTech Law**, v. 14, n. 1, p. 28-31, 2017.

GIUFFRIDA, Iria. Liability for AI Decision-Making: Some Legal and Ethical Considerations. **Fordham Law Review**, v. 88, p. 439, 2019.

GOODFELLOW, Ian et al. **Deep learning**. Cambridge: MIT press, 2016. Disponível em < <https://www.deeplearningbook.org/>>. Acesso em 1º mai. 2021.

GOOGLE. **Guia sobre Análise de Dados e Aprendizado de Máquina para CIO**. 2017. Disponível em: < https://lp.google-mkto.com/rs/248-TPC-286/images/Guide_to_Data_Analytics_%26_Machine_Learning_portuguese.pdf> . Acesso em: 19 abr. 2021.

HAGER, Gregory D. et al. **Artificial Intelligence for Social Good**. Washington: Computing Community Consortium. 2019. Disponível em: <https://arxiv.org/abs/1901.05406>. Acesso em 11 ago. 2021.

HARTMANN PEIXOTO, Fabiano. **Direito e Inteligência Artificial**: referenciais básicos. Coleção Inteligência Artificial e Jurisdição. Volume 2. DR.IA. Brasília, 2020.

HARTMANN PEIXOTO, Fabiano. **Inteligência Artificial e Direito**: convergência ética e estratégica. Coleção Direito, Racionalidade e Inteligência Artificial. Volume 5. Curitiba: Alteridade, 2020a.

HARTMANN PEIXOTO, Fabiano. **Nota de Colaboração**. Substitutivo ao Projeto de Lei nº 21/2020. Grupo de Pesquisa DR.IA.UnB – Direito e Inteligência Artificial. Universidade de Brasília. 2021.

HARTMANN PEIXOTO, Fabiano; COUTINHO, Marina de Alencar Araripe. Inteligência Artificial e Regulação: uma análise do Projeto de Lei 5.051/2019. **Revista Em Tempo**, v. 19, n. 1, 2020. Disponível em: <<https://revista.univem.edu.br/emtempo/article/view/3129>>. Acesso em: 12 nov. 2021.

HARTMANN PEIXOTO, Fabiano; LAGE, Fernanda de Carvalho. Inteligência Artificial e Direito: desafios para a regulação do uso da inteligência artificial. In: HARTMANN PEIXOTO, Fabiano (org.). **Inteligência Artificial: estudos de inteligência artificial**. Coleção Direito, Racionalidade e Inteligência Artificial. Volume 4. Curitiba: Alteridade, 2021, p. 267-292.

HARTMANN PEIXOTO, Fabiano; SILVA, Roberta Zumblick Martins da. **Inteligência Artificial e Direito**. Coleção Direito, Racionalidade e Inteligência Artificial. Volume 1. Curitiba: Alteridade, 2019.

HARTZOG, Woodrow. On Questioning Automation. **Cumberland Law Review**, v. 48, p. 1, 2017.

HAWKINS, Jeff. Tech Luminaries Address Singularity. **IEEE Spectrum**. Special Report: The Singularity, 2008. Disponível em: <https://spectrum.ieee.org/computing/hardware/tech-luminaries-address-singularity>. Acesso em 26 abr. 2021.

HELDT, Amélie. Let's Meet Halfway: Sharing New Responsibilities in a Digital Age. **Journal of Information Policy**, v. 9, p. 336-369, 2019.

HOLMES, Pablo. O constitucionalismo entre a fragmentação e a privatização: problemas evolutivos do direito e da política na era da governança global. **Dados-Revista de Ciências Sociais**, v. 57, n. 4, p. 1137-1168, 2014.

JORDAN, Michael I.; MITCHELL, Tom M. Machine learning: Trends, perspectives, and prospects. **Science**, v. 349, n. 6245, p. 255-260, 2015.

KAMINSKI, Margot E. Binary Governance: Lessons from the GDPR's approach to algorithmic accountability. **Southern California Law Review**, v. 92, p. 1529-1616, 2019.

KIM, Richard et al. A computational model of commonsense moral decision making. **Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society**. ACM, p. 197-203, 2018.

KNUTH, Donald Ervin. **The Art of Computer Programming**. Addison-Wesley, 3ª ed., v. 1, 1997.

LAGE, Fernanda de Carvalho. **Manual de Inteligência Artificial no Direito Brasileiro**. Salvador: Juspodium, 2021.

L'HEUREUX, Alexandra et al. Machine learning with big data: Challenges and approaches. **IEEE Access**, v. 5, p. 7.776-7.797, 2017.

LEAL, Fernando. Seis objeções ao direito civil constitucional. **Revista Brasileira de Direitos Fundamentais & Justiça**, v. 9, n. 33, p. 123-164, 2015.

LEONARDO, Rodrigo Xavier. Responsabilidade civil contratual e extracontratual: primeiras anotações em face do novo Código Civil Brasileiro. **Tuiuti: Ciência e Cultura**, n. 30, p. 107-119, 2002.

LEWICKI, Bruno. Metodologia do Direito Civil Constitucional: Futuros possíveis e armadilhas. **Revista Brasileira Direito Civil**, v. 1, p. 271-276, 2014.

LÔBO, Paulo. Metodologia do Direito Civil Constitucional. *In*: RUZYK, Carlos Eduardo Pianovisky; SOUZA, Eduardo Nunes de; MENEZES, Joyceanne Bezerra de; EHRHARDT JUNIOR, Marcos (Org.). **Direito Civil Constitucional: a ressignificação da função dos institutos fundamentais do direito civil contemporâneo e suas consequências**. Florianópolis: Conceito, p. 19-27, 2014.

LOPEZ, Teresa Ancona. **Princípio da Precaução e Evolução da Responsabilidade Civil**. São Paulo: Quartier Latin, 2010.

MEDON, Filipe. Decisões Automatizadas: o necessário diálogo entre a inteligência artificial e a proteção de dados pessoais para a tutela de direitos fundamentais. *In*: TEPEDINO, Gustavo; SILVA, Rodrigo da Guia (Coord.). **O Direito Civil na Era da Inteligência Artificial**. São Paulo: Thomson Reuters Brasil, p. 337-370, 2020.

MENDES, Gilmar Ferreira; OLIVEIRA FERNANDES, Victor. Constitucionalismo digital e jurisdição constitucional: uma agenda de pesquisa para o caso brasileiro. **Revista Brasileira de Direito**, v. 16, n. 1, p. 1-33, 2020.

MITCHELL, Tom Michael. **The discipline of machine learning**. Pittsburgh, PA: Carnegie Mellon University, School of Computer Science, Machine Learning Department, 2006.

MONTEIRO FILHO, Carlos Edison do Rêgo; ROSENVALD, Nelson. Riscos e Responsabilidades na Inteligência Artificial e noutras Tecnologias Digitais Emergentes. *In*: TEPEDINO, Gustavo; SILVA, Rodrigo da Guia (Coord.). **O Direito Civil na Era da Inteligência Artificial**. São Paulo: Thomson Reuters Brasil, 2020, p.543-564.

MOOR, James. The Dartmouth College artificial intelligence conference: The next fifty years. **Ai Magazine**, v. 27, n. 4, p. 87-87, 2006.

MOORE, Gordon E. Tech Luminaries Address Singularity. **IEEE Spectrum**. Special Report: The Singularity. 2008. Disponível em:

<https://spectrum.ieee.org/computing/hardware/tech-luminaries-address-singularity>. Acesso em 26 abr. 2021.

MORAES, Maria Celina Bodin de. A constitucionalização do direito civil e seus efeitos sobre a responsabilidade civil. **Revista Direito, Estado e Sociedade**, n. 29, p. 233-258, 2006.

MORAES, Maria Celina Bodin de. Risco, solidariedade e responsabilidade objetiva. **Revista dos Tribunais**, v. 854, p. 11-37, 2006a.

MORAES, Maria Celina Bodin de; QUEIROZ, João Quinelato de. LGPD: um novo regime de responsabilização civil dito proativo. Editorial. **civilistica.com**, a. 8, n. 3, 2019.

MUGGLETON, Stephen. Alan Turing and the development of Artificial Intelligence. **AI communications**, v. 27, n. 1, p. 3-10, 2014.

MULHOLLAND, Caitlin Sampaio. **A responsabilidade civil por presunção de causalidade**. Rio de Janeiro: GZ Editora, 2010.

MULHOLLAND, Caitlin. Responsabilidade civil por danos causados pela violação de dados sensíveis e a Lei Geral de Proteção de Dados Pessoais (Lei 13.709/2018). In: MARTINS, Guilherme Magalhães; ROSENVALD, Nelson (Coords.). **Responsabilidade civil e novas tecnologias**. São Paulo: Editora Foco, 2020.

MULHOLLAND, Caitlin; KREMER; Bianca. Responsabilidade Civil por Danos Causados pela Violação do Princípio da Igualdade no Tratamento de Dados Pessoais. In: TEPEDINO, Gustavo; SILVA, Rodrigo da Guia (Coord.). **O Direito Civil na Era da Inteligência Artificial**. São Paulo: Thomson Reuters Brasil, p. 565-584, 2020.

NARUHODO 278: **O que é singularidade?** - Parte 2 de 2. Locução de: Ken Fujioka e Altay de Souza. São Paulo: Brainstorm9 Podcasts, 5 abr. 2021. Podcast. Disponível em: <https://www.b9.com.br/shows/naruhodo/naruhodo-278-o-que-e-singularidade-parte-2-de-2/>. Acesso em 17 mai. 2021.

NEMITZ, Paul. Constitutional democracy and technology in the age of artificial intelligence. **Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences**, v. 376, n. 2133, 2018.

NILSSON, Nils J. **The quest for artificial intelligence**. Cambridge University Press, 2009.

NORVIG, Peter; RUSSELL, Stuart. **Inteligência Artificial**. Rio de Janeiro: Elsevier, 3ª Ed., 2013.

NUNES, Dierle; MARQUES, Ana Luiza Pinto Coelho. Inteligência artificial e direito processual: vieses algorítmicos e os riscos de atribuição de função decisória às máquinas. **Revista de Processo**, v. 285, p. 421-447, 2018.

PARENTONI, Leonardo Netto; VALENTINI, Rômulo Soares; ALVES, Tárík César Oliveira. Panorama da Regulação da Inteligência Artificial no Brasil: com ênfase no PL n. 5.051/2019. **Revista Eletrônica do Curso de Direito da UFSM**, v. 15, n. 2, 2020.

PEREIRA, Thales Alessandro Dias. Transparência de Algoritmos e Decisões Automatizadas: compreendendo o art. 20 da LGPD. *In*: HARTMANN PEIXOTO, Fabiano (Org.). **Inteligência Artificial: estudos de inteligência artificial**. Coleção Direito, Racionalidade e Inteligência Artificial. Volume 4. Curitiba: Alteridade, 2021, p. 159-174.

PERLINGIERI, Pietro. **Perfis do Direito Civil: Introdução ao direito civil constitucional**. Rio de Janeiro: Renovar, 3ª ed., 2007.

PERLINGIERI, Pietro. Normas constitucionais nas relações privadas. **Civilistica.com**. Rio de Janeiro, a. 8, n. 1, 2019.

PINTO, Renata Ávila. Digital Sovereignty or Digital Colonialism?. **Sur International Journal on Human Rights**, 2019.

QUEIROZ, João Quintelato de. Responsabilidade Civil no Uso da Inteligência Artificial: imputação, culpa e risco. *In*: TEPEDINO, Gustavo; SILVA, Rodrigo da Guia (Coord.). **O Direito Civil na Era da Inteligência Artificial**. São Paulo: Thomson Reuters Brasil, p. 585-608, 2020.

RAMOS, André Luiz Arnt. A responsabilidade civil para além dos esquemas tradicionais: prospecções do dano reparável na contemporaneidade. **Revista Fórum de Direito Civil - RFDC**. Belo Horizonte, ano 4, n. 10, p. 13-33, set./dez., 2015.

ROBERTS, Huw et al. **The Chinese approach to artificial intelligence: An analysis of policy and regulation**. 2019. Disponível em https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3469783. Acesso em 10 dez. 2021.

RODRIGUES, Cássio Monteiro. Reparação de danos e função preventiva da responsabilidade civil: parâmetros para o ressarcimento de despesas preventivas ao dano. **Civilistica.com**, a. 9, n. 1, p. 1-37, 2020.

ROSENVALD, Nelson. **A Polissemia da Responsabilidade Civil na LGPD**. Migalhas, 2020. Disponível em < <https://www.migalhas.com.br/coluna/migalhas-de-protecao-de-dados/336002/a-polissemia-da-responsabilidade-civil-na-lgpd>>. Acesso em 13 nov. 2021.

RUSSELL, Stuart. **Q&A: The Future of Artificial Intelligence**. University of Berkeley, CA, 2016. Disponível em <https://people.eecs.berkeley.edu/~russell/research/future/q-and-a.html>. Acesso em 30 mar. 2021.

RUSSELL, Stuart; DEWEY, Daniel; TEGMARK, Max. Research priorities for robust and beneficial artificial intelligence. **Ai Magazine**, v. 36, n. 4, p. 105-114, 2015.

SANTOS, Francisco C. et al. A Regulation Dilemma in Artificial Intelligence Development. **ALIFE 2021: The 2021 Conference on Artificial Life**. MIT Press, 2021.

SARLET, Ingo Wolfgang. Neoconstitucionalismo e influência dos direitos fundamentais no direito privado: algumas notas sobre a evolução brasileira. **Civilistica.com**. Rio de Janeiro, ano 1, n. 1, jul.- set., 2012.

SARMENTO, Daniel. **Direitos Fundamentais e Relações Privadas**. Rio de Janeiro: Lumen Juris, 2ª ed., 2006.

SCHREIBER, Anderson. **Novos Paradigmas Da Responsabilidade Civil: Da Erosão Dos Filtros Da Reparação À Diluição Dos Danos**. São Paulo: Editora Atlas, 2ª ed., 2009.

SCHERER, Matthew U. Regulating artificial intelligence systems: Risks, challenges, competencies, and strategies. **Harvard Journal of Law and Technology**, v. 29, p. 353, 2016.

SCIPIONE, Jacopo. **Artificial Intelligence and Europe: risks, developments and implications**. 2020. Disponível em https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3598543. Acesso em 10 dez. 2021.

SILVA, Nilton Correia da. Inteligência Artificial. *In*: FRAZÃO, Ana; MULHOLLAND, Caitlin. (Coord.). **Inteligência Artificial e Direito: Ética, Regulação e Responsabilidade**. 2. ed. São Paulo: Revista dos Tribunais, 2020, p. 33-50.

SILVA, Virgílio Afonso da. **A Constitucionalização do Direito: os direitos fundamentais nas relações entre particulares**. São Paulo: Malheiros, 2005.

SOLOW-NIEDERMAN, Alicia. Administering Artificial Intelligence. **Southern California Law Review**, v. 93, p. 633, 2020.

SOUZA, Eduardo Nunes de. Personalidade Jurídica e Inteligência Artificial. *In*: TEPEDINO, Gustavo; SILVA, Rodrigo da Guia (Coord.). **O Direito Civil na Era da Inteligência Artificial**. São Paulo: Thomson Reuters Brasil, p. 213-245, 2020.

SOUZA, Carlos Affonso; PADRÃO, Vinícius. Responsabilidade Civil e os “Novos Turcos”: o papel da conduta humana nos danos relacionados à inteligência artificial. *In*: TEPEDINO, Gustavo; SILVA, Rodrigo da Guia (Coord.). **O Direito Civil na Era da Inteligência Artificial**. São Paulo: Thomson Reuters Brasil, p. 663-681, 2020.

STANKOVIC, Mirjana et al. **Exploring Legal, Ethical and Policy Implications of Artificial Intelligence**. Law Justice and Development, Washington, 2017.

STEPHENS, Phillip. **The Technological Singularity: An Ideological Critique**. Dissertação (mestrado). University of Arkansas, 2015.

STONE, Peter et al. Artificial Intelligence and Life in 2030. **One hundred year study on artificial intelligence**: Report of the 2015-2016 Study Panel. Stanford University, CA, 2016.

SURDEN, Harry. Machine learning and law. **Washington Law Review**, v. 89, p. 87, 2014.

SUTTON, Richard S. Introduction: The challenge of reinforcement learning. *In*: SUTTON, Richard S.; BARTO, Andrew G. **Reinforcement Learning**: an introduction. Boston: Springer, 1992. p. 1-3.

TAVARES PEREIRA, S. Que é isto, a eNorma? Elementos para a teoria geral do eDireito. *In*: BRANDÃO, Cláudio (org.) **Princípios do processo em meio reticular-eletrônico: fenomenologia, normatividade e aplicação prática**. São Paulo: LTr, 2017.

TEPEDINO, Gustavo. O futuro da responsabilidade civil. Editorial. **Revista Trimestral de Direito Civil**, vol. 24, 2004. Disponível em <<http://www.tepedino.adv.br/wpp/wp-content/uploads/2012/09/RTDC.Editorial.v.024.pdf>>. Acesso em 11 nov. 2021.

TEPEDINO, Gustavo. Premissas Metodológicas para a Constitucionalização do Direito Civil. **Revista de Direito do Estado**, ano 1, n. 2, p. 37-53, 2006.

TEPEDINO, Gustavo. Marchas e contramarchas da constitucionalização do direito civil: a interpretação do direito privado à luz da Constituição da República. **(Syn)Thesis**, v. 5, n. 1, p. 15-21, 2012.

TEPEDINO, Gustavo. Nexo de Causalidade: anotações acerca de suas teorias e análise de controvertida casuística no direito brasileiro. *In*: MUNIZ, Francisco Arthur de Siqueira; BARBOSA, Mafalda Miranda (Coord.). **Responsabilidade Civil: 50 anos em Portugal e 15 anos no Brasil**. Salvador: JusPodium, p. 303-325, 2017.

TEPEDINO, Gustavo; SILVA, Rodrigo da Guia. Inteligência Artificial e Elementos da Responsabilidade Civil. *In*: FRAZÃO, Ana; MULHOLLAND, Caitlin. (Coord.). **Inteligência Artificial e Direito: Ética, Regulação e Responsabilidade**. 2. ed. São Paulo: Revista dos Tribunais, 2020, p. 296-325.

TEUBNER, Gunther. Global Bukowina: legal pluralism in the world society. *In*: **Critical theory and legal autopoiesis**. Manchester University Press, 2019.

TOLEDO, Cláudia. Inteligência Artificial e sua Aplicabilidade em Decisões Judiciais. *In*: HARTMANN PEIXOTO, Fabiano (org.). **Inteligência Artificial: estudos de inteligência artificial**. Coleção Direito, Racionalidade e Inteligência Artificial. Volume 4. Curitiba: Alteridade, 2021, p. 57-90.

TEIXEIRA, João de Fernandes. **O pesadelo de Descartes**: do mundo mecânico à Inteligência Artificial. Porto Alegre: Editora Fi, 2018.

TURING, Alan Mathison. Intelligent Machinery. *In*: COPELAND, B. Jack (Ed.). **The Essential Turing**. Oxford University Press, 2004, p. 395-432.

TURING, Alan Mathison. Computing machinery and intelligence. *In*: EPSTEIN, Robert et al. **Parsing the turing test: philosophical and methodological issues in the quest for the thinking computer**. Springer, 2009. p. 23-65.

VALENTINI, Rômulo Soares. **Julgamento por computadores? As novas possibilidades da juscibernética no século XXI e suas implicações para o futuro do direito e do trabalho dos juristas**. Tese (Doutorado em direito) – Faculdade de Direito, Universidade Federal de Minas Gerais. Belo Horizonte, 2018.

VENTURI, Thaís Gouveia Pascoaloto. **A Construção da Responsabilidade Civil Preventiva no Direito Civil Contemporâneo**. Tese (Doutorado em Direito das Relações Sociais) – Programa de Pós-graduação em Direito, do Setor de Ciências Jurídicas Universidade Federal do Paraná. Curitiba, 2012. p. 15.

VLADECK, David C. Machines without principals: liability rules and artificial intelligence. **Washington Law Review**, v. 89, p. 117, 2014.

WALSH, Toby. The singularity may never be near. **AI Magazine**, v. 38, n. 3, p. 58-62, 2017.

WILSON, Benjamin; HOFFMAN, Judy; MORGENSTERN, Jamie. Predictive inequity in object detection. **arXiv preprint arXiv:1902.11097**, 2019.

WOLFF, Guntram B. Europe may be the world's AI referee, but referees don't win. **Politico**, 2020. Disponível em <<https://www.politico.eu/article/europe-may-be-the-worlds-ai-referee-but-referees-dont-win-margrethe-vestager/>>. Acesso em 10 dez. 2021.

WRIGHT, Jasmine; VERITY, Andrej. **Artificial Intelligence Principles For Vulnerable Populations in Humanitarian Contexts**. Digital Humanitarian Network, 2020. Disponível em <<https://www.digitalhumanitarians.com/artificial-intelligence-principles-for-vulnerable-populations-in-humanitarian-contexts/>>. Acesso em 19 out. 2021.

ZHANG, Aston; et al. **Dive into Deep Learning**. 2021. Disponível em <https://d2l.ai/d2l-en.pdf> . Acesso em 17 mai. 2021.