

QUEM É O CULPADO? UMA ANÁLISE CRIMINAL DAS INTERAÇÕES HUMANAS COM A INTELIGÊNCIA ARTIFICIAL

Rogério Luís Marques de Mello

Doutor e Mestre pelo programa em Ciências Policiais de Segurança e Ordem Pública do Centro de Altos Estudos em Segurança Coronel PM Nelson Freire Terra da Polícia Militar do Estado de São Paulo – CAES. Mestre em Direito pela Universidade de Marília/SP – UNIMAR. Graduado em Direito pela Universidade de São Paulo – USP e em Ciências Policiais de Segurança e Ordem Pública pela Academia de Polícia Militar do Barro Branco/SP – APMBB. Professor da graduação em Direito da Fundação Educacional de Penápolis/SP – FUNEPE. *E-mail:* rogeriomello@funepe.edu.br.

Rafael José Nadim de Lazari

Pós-Doutor em Democracia e Direitos Humanos pelo Centro de Direitos Humanos da Faculdade de Direito da Universidade de Coimbra – Portugal. Estágio Pós-Doutoral pelo Centro Universitário Eurípides Soares da Rocha, de Marília/SP. Doutor em Direito Constitucional pela Pontifícia Universidade Católica, de São Paulo/SP. Mestre em Teoria do Estado pelo Centro Universitário Eurípides Soares da Rocha, de Marília/SP. Graduado em Direito pelas Faculdades Integradas Antônio Eufrásio de Toledo, de Presidente Prudente/SP. Professor da Graduação, do Mestrado e do Doutorado em Direito da Universidade de Marília/SP – UNIMAR. Advogado, consultor jurídico e parecerista na área de Direito Público. *E-mail:* prof.rafaeldelazari@hotmail.com.

Resumo: A inteligência artificial, tecnologia inovadora que procura imitar as habilidades humanas, tem conceito e características variáveis e fluidas, que vão desde uma inteligência atual sem autonomia em relação ao ser humano, até agentes inteligentes com total independência em relação aos comportamentos humanos. Em face de um crescente poder decisório das máquinas em razão de sistemas dotados de *machine learning* e redes neurais, o artigo pretende investigar a eventual responsabilização penal – dos humanos e das máquinas – considerando a estrutura jurídico-penal vigente no Brasil no contexto do sistema finalista da conduta. Por hipótese, sustenta-se que ações gravosas provocadas por uma inteligência artificial autônoma não podem ser imputadas aos humanos; por outro lado, máquinas inteligentes, no contexto dessa mesma dogmática penal, não podem ser autoras de delito. Diante desse vazio ou (insuficiência) de imputação criminal, são analisadas as possibilidades da responsabilização penal relacionadas à inteligência artificial, sempre considerando os direitos humanos fundamentais envolvidos.

Palavras-chave: Inteligência Artificial. Inovação Tecnológica. Direitos Fundamentais. Responsabilidade Penal.

Sumário: 1 Introdução – 2 Fundamentos da responsabilidade penal no Brasil – 3 Mente humana e inteligência artificial – 4 Responsabilidade penal da inteligência artificial – 5 Conclusão – Referências

1 Introdução

A evolução do ser humano ao longo da história é, antes de tudo, uma luta pela sobrevivência. Das espécies sobre a face da Terra, os humanos prevaleceram como criaturas especiais, dotadas de fala, postura ereta, mãos com polegares opostos, cérebro comparativamente mais desenvolvido – com capacidade de memória, aprendizado, consciência, racionalidade e criatividade –, dentre inúmeras outras vantagens evolutivas. Foi nesse contexto que o homem aprendeu a cultivar seu próprio alimento, domesticou animais, descobriu o fogo, criou ferramentas, inventou a escrita, desenvolveu cidades, criou moedas e buscou, cada vez mais, uma melhor qualidade de vida.

Dentre as várias mudanças evolutivas conquistadas pelo homem, uma nova e surpreendente revolução ocorre a partir do século XX: a chamada revolução digital, baseada na invenção dos computadores e da *internet*. Desde então, as transformações tecnológicas ocorrem cada vez mais rapidamente e com maior impacto em diversos setores da vida humana como transporte, saúde, entretenimento, educação, segurança, trabalho, bem-estar e produtividade. Uma dessas tecnologias é a inteligência artificial, criada ainda em meados do século XX com a pretensão de, simulando as características do cérebro humano através das máquinas, imprimir soluções adequadas a problemas complexos.

A partir do início do século XXI, com o desenvolvimento de processadores cada vez mais velozes, bancos de dados gigantescos e novas demandas industriais, a inteligência artificial ganha impulso vertiginoso e passa a permear todos os ramos do conhecimento humano e, conseqüentemente, da vida das pessoas. Baseada em algoritmos que determinam procedimentos às máquinas, a inteligência artificial foi aprimorada através de técnicas de *machine learning*, *deep learning* e redes neurais cogitando-se, na atualidade, a sua capacidade de decidir e realizar ações de modo autônomo, sem qualquer participação do ser humano. Foi a partir do atropelamento e morte de uma pedestre nos Estados Unidos, em 2018, durante testes de um carro autônomo,¹ que a questão da responsabilidade penal relacionada à inteligência artificial passou a ser evidenciada.

Nesse contexto, o presente artigo objetiva analisar os fundamentos e condições de uma eventual responsabilização penal relacionada à inteligência artificial tendo, por paradigma, o regime jurídico brasileiro. Baseado no método dedutivo e com emprego auxiliar dos métodos histórico e comparativo, serão investigadas fontes bibliográficas de relevância do Direito Penal e das neurociências de modo

¹ “Carro autônomo da Uber nos EUA causa primeira morte por atropelamento.” *Folha de São Paulo*, 19 mar. 2018. Disponível em: <https://www1.folha.uol.com.br/mercado/2018/03/mulher-morre-nos-eua-apos-ser-atropelada-por-carro-autonomo-da-uber.shtml>. Acesso em: 20 dez. 2020.

a entender as possibilidades de utilização do modelo jurídico-penal brasileiro em relação a possíveis danos praticados através e pela inteligência artificial.

2 Fundamentos da responsabilidade penal no Brasil

Em algum momento da sua existência, o ser humano decidiu viver em comunidade e, desde então, regras mínimas de interação foram impostas entre os conviventes visando a sua própria sobrevivência. Com a complexidade da sociedade e das relações humanas, tornou-se necessária a definição de regras claras que deveriam ser cumpridas por todos na consecução do bem comum. Para Beccaria, foram as leis que deram as condições mínimas de vida em sociedade a homens que, independentes e isolados até então, estavam “cansados de viver em contínuo estado de guerra e de gozar uma liberdade inútil pela incerteza de conservá-la”.²

Nos seus primórdios e ainda na Idade Média, durante a fase do antigo direito germânico, a responsabilidade penal era objetiva e decorria, estritamente, de uma relação de causalidade entre a conduta e o resultado. Responsabilizava-se, por isso, o coletivo, seja a família, a tribo ou as demais pessoas envolvidas com o eventual autor de crime, fomentando-se animosidades, vinganças e injustiças. A partir, principalmente, dos ideais iluministas, o Direito Penal passa a exigir uma vinculação subjetiva, portanto, individual e pessoal entre o autor da conduta e o resultado criminoso. Nesse contexto, o próprio conceito de delito foi reconfigurado impondo, como requisitos, não só a ocorrência de um injusto penal (realização de um fato típico e ilícito) mas, também, a culpabilidade como símbolo da responsabilidade penal pessoal.³

A definição do conceito de conduta tem permeado as maiores discussões ao longo da evolução do Direito Penal, uma vez que é a partir dela que são estabelecidas as teorias sobre o delito e suas estruturas componentes. A teoria causal-naturalista da ação foi elaborada por Von Liszt no final do século XIX, definindo a ação como um “[...] fato que repousa sobre a vontade humana, a mudança no mundo exterior referível à vontade do homem”.⁴ Em termos práticos, não se questiona a vontade (ou dolo) na conduta, transportando-a para o elemento culpabilidade.

Diante da impossibilidade fática de separação da vontade em relação ao seu conteúdo, Hans Welzel elaborou o conceito finalista de conduta, considerando

² BECCARIA, Cesare. *Dos delitos e das penas*. São Paulo: Revista dos Tribunais, 1999, p. 27.

³ BUSATO, Paulo César. *Direito penal*. São Paulo: Atlas, 2015.

⁴ LISZT, Franz Von. *Tratado de direito penal alemão*. Brasília: Senado Federal, 2006, p. 193. V. 1.

que o homem age, conscientemente, em direção a um fim, impondo ao desvalor do resultado um correspondente desvalor da ação e retirando da culpabilidade a análise volitiva do agente, tornando-a puramente normativa.⁵

Já a culpabilidade, segundo Zaffaroni e Pierangeli,⁶ é o conceito mais debatido da teoria do delito, justamente pela análise aprofundada que se faz da atuação humana, ainda que esta permeie, em maior ou menor medida, todo o conceito analítico de crime. Destaque-se que, enquanto o injusto penal (conduta típica e antijurídica) revela o desvalor da conduta pelo direito, a culpabilidade denota a reprovabilidade da conduta diante de uma especial condição do autor. Nestes termos:

[...] é reprovável ao autor a realização desta conduta porque não se motivou na norma, sendo-lhe exigível, nas circunstâncias em que agiu, que nela se motivasse. Ao não se ter motivado na norma, quando podia e lhe era exigível que o fizesse, o autor mostra uma disposição interna contrária ao direito.⁷

Ainda que prevaleça, atualmente no direito brasileiro, a aplicação da teoria finalista da conduta, não se desconhece a existência de outros modelos de conceituação de ação, cada um deles correspondendo à evolução da ciência de seu tempo e variando conforme as condições materiais da sociedade como o modo de produção econômica, as concepções filosóficas, jurídicas e institucionais, dentre várias outras circunstâncias.⁸

Fundamentado no sistema significativo de ação, Antón busca superar os dilemas filosóficos e verdades absolutas relacionadas à culpabilidade, atribuindo-lhe uma dimensão pragmática baseada na liberdade de ação relacionada à própria forma de vida do homem e nas normas que o regem; em outras palavras, somos livres consideradas certas margens ancoradas em nossa realidade e podemos ou não atuar de outro modo, diferente de como fazemos.⁹

⁵ WELZEL, Hans. *O novo sistema jurídico-penal: uma introdução à doutrina da ação finalista*. São Paulo: Revista dos Tribunais, 2015.

⁶ ZAFFARONI, Eugenio Raul; PIERANGELI, José Henrique. *Manual de direito penal brasileiro*. São Paulo: Thomson Reuters, 2019.

⁷ ZAFFARONI, Eugenio Raul; PIERANGELI, José Henrique. *Manual de direito penal brasileiro*. São Paulo: Thomson Reuters, 2019, p. 537.

⁸ TAVARES, Juarez. *Fundamentos da teoria do delito*. Florianópolis: Tirant lo Blanch, 2018.

⁹ ANTÓN, Tomás Salvador Vivés. *Fundamentos del sistema penal*. Valencia: Tirant lo Blanch, 1996.

3 **Mente humana e inteligência artificial**

A inteligência, entendida como a capacidade de compreender, raciocinar e solucionar problemas, sempre foi atribuída aos seres humanos. Contudo, a possibilidade de máquinas possuírem atributos similares ao intelecto humano já era discutida no final do século passado, com pródigos argumentos favoráveis e contrários à ideia de máquinas inteligentes. Partindo do pressuposto de que os seres humanos são mecanismos físicos extremamente complexos, mas dotados de processos bem compreendidos, os mais otimistas sustentam a viabilidade de, um dia, criar-se máquinas que reproduzam esses sistemas, ainda que com variações nos materiais e formas. Afinal, processos orgânicos e físicos um dia desconhecidos pela ciência foram, em algum momento, criteriosamente compreendidos e não haveria motivos para a inteligência ser uma exceção. Esse mesmo fundamento, porém, serve aos céticos para sustentarem a impossibilidade de que isso seja concretizado um dia, afinal, a mente humana é absolutamente imaterial e, portanto, incompatível com a ideia de robôs que são puramente materiais.¹⁰

Por mais que o tema inteligência artificial pareça recente, a ideia de construir uma máquina pensante que imitasse as habilidades humanas é muito antiga. Em 1950, Alan Turing, matemático inglês, descobriu os princípios de funcionamento dos computadores modernos, estabelecendo um teste para definir a capacidade de pensar das máquinas similarmente aos humanos.¹¹ Basicamente, o teste consiste em saber se a máquina consegue enganar o ser humano. Na prática, são elaboradas perguntas pelo ser humano a um sistema de inteligência artificial e a outro ser humano, sem que se saiba quem efetivamente responde; caso o interrogante não descubra quais respostas foram dadas pelo computador, a máquina passou no teste.

O conceito de inteligência artificial, ainda que extremamente variado, baseia-se na capacidade de um sistema de interpretar dados externos corretamente, aprender com esses dados e usá-los para atingir objetivos e tarefas específicos por meio de uma adaptação flexível.¹² Frazão e Mulholland definem a inteligência

¹⁰ DENNETT, Daniel C. *Consciousness in Human and Robot Minds*. In: *IIAS Symposium on cognition, computation and consciousness*. Kyoto: 1 set. 1994. Disponível em: <https://ase.tufts.edu/cogstud/dennett/papers/concrobt.htm>. Acesso em: 10 jan. 2021.

¹¹ TURING, Alan. *Computing Machinery and Intelligence*, *Mind*, V LIX, e. 236, out. 1950, p. 433-460. Disponível em: <https://academic.oup.com/mind/article/LIX/236/433/986238>. Acesso em: 05 jan. 2021.

¹² KAPLAN, Andreas; HAENLEIN, Michael. *Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence*. *Business Horizons*, p. 15-25, 2019. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0007681318301393?via%3DIihub>. Acesso em: 07 jan. 2021.

artificial como “todo sistema computacional que simula a capacidade humana de raciocinar e resolver problemas, por meio de tomada de decisões baseadas em análises probabilísticas”.¹³

Em termos gerais, pode-se dizer que a inteligência artificial utiliza os dados disponíveis em grandes bancos de armazenamento (*big data*) para a definição de regras e modelos necessários ao aprendizado de máquinas (*machine learning*), uma vez que auxiliam os computadores a aprenderem sem serem explicitamente programados. Assim, é possível conceituar *machine learning* como uma forma de inteligência artificial que permite que um sistema aprenda a partir de dados e não através de programação explícita. Em regra, esses algoritmos são treinados para, ao receberem informações através de dados, melhorarem seus modelos e oferecerem resultados mais precisos, levando a uma evolução nos tipos de associações feitas entre elementos de dados. Esse aprendizado, por sua vez, pode ocorrer em termos mais simples – como durante o uso de redes sociais ou *sites* de compras – ou extremamente complexos, como acontece no chamado *deep learning* ou aprendizado profundo, que incorpora redes neurais artificiais – algoritmos que imitam a estrutura do cérebro humano –, dispostos em inúmeras camadas com o objetivo de aprender padrões complicados em grandes quantidades de dados e que pode consistir em milhares ou milhões de nós de processamento simples densamente interconectados.¹⁴

Mesmo que se entenda que o estado da arte atual em relação à inteligência artificial não possibilite às máquinas decisões estritamente autônomas, também não se pode afirmar, peremptoriamente, que isso não acontecerá.¹⁵ Para que se considere um ser autônomo racional, sua aprendizagem deve compensar um conhecimento prévio parcial ou incorreto. Portanto, adquirida a habilidade de aprender e obtendo experiência sobre seu ambiente de atuação, o comportamento de um agente racional pode se tornar independente de seu conhecimento anterior, ou seja, “um agente que aprende pode ser totalmente autônomo”.¹⁶

A par de todas as vantagens inegáveis que a tecnologia e a sua exponencial evolução oferecem, não se desconhece a existência de problemas relacionados

¹³ FRAZÃO, Ana; MULHOLLAND, Caitlin (Coord.). *Inteligência artificial e direito – Ética, regulação e responsabilidade*. São Paulo: Revista dos Tribunais, 2019, p. 6.

¹⁴ ELIAS, Paulo Sá. *Algoritmos, Inteligência Artificial e o Direito*. E-Gov: portal de e-governo, inclusão digital e sociedade do conhecimento, Florianópolis, 2018. Disponível em: <http://www.egov.ufsc.br/portal/conteudo/algoritmos-intelig%C3%Aancia-artificial-e-o-direito>. Acesso em: 16 nov. 2020.

¹⁵ “O Estado da Califórnia, nos Estados Unidos, libera *delivery* sem motorista e entregas podem começar em 2021. Os veículos autônomos Nuro R2 usam um radar, imagens térmicas e câmeras de 360 graus para direcionar seus movimentos.” Disponível em: <https://g1.globo.com/economia/tecnologia/noticia/2020/12/25/california-libera-delivery-sem-motorista-e-entregas-podem-comecar-em-2021.ghtml>. Acesso em: 25 dez. 2020.

¹⁶ RUSSELL, Stuart J.; NORVIG, Peter. *Inteligência artificial*. Rio de Janeiro: Elsevier, 2013, p. 287.

à sua utilização e que podem repercutir em diversos aspectos relacionados aos direitos fundamentais como: vida privada, não discriminação, integridade física e até mesmo a vida humana. Os defensores da inteligência artificial sustentam que: apesar dos riscos envolvidos, o que se busca não é a perfeição, mas, sim a melhor decisão possível;¹⁷ também os humanos cometem erros e adotam decisões enviesadas, baseadas em preconceitos;¹⁸ que a análise inteligente de dados possibilita entender a sociedade em toda a sua complexidade, oferecendo uma visão mais precisa dos padrões de vida humana e, conseqüentemente, contribuindo para a solução mais adequada das suas demandas.¹⁹ Da relativa autonomia conquistada pela inteligência artificial até o momento, surgem questões práticas e éticas que desafiam a sociedade e que ganham força justamente em razão da conquista de maior autonomia desses sistemas e possível ausência do controle humano nas decisões algorítmicas, com o conseqüente incremento de riscos e ameaças, cada vez mais imponderáveis.²⁰

4 Responsabilidade penal da inteligência artificial

A verificação da responsabilidade penal do ser humano relacionado à inteligência artificial suscita, além da questão jurídica, um necessário debate inicial sobre inovação tecnológica e direitos fundamentais. Para Hallevy, uma análise ética e moral das decisões de entidades inteligentes levaria à conclusão, com base no senso comum, pela responsabilização do ser humano por trás da inteligência artificial, mas não do *software* que, afinal, nada mais é do que uma ferramenta nas mãos do seu criador.²¹ Contudo, a atribuição de responsabilidades com fundamento na moralidade não é adequada, inclusive para os seres humanos, ainda que aceitável por grande parte das pessoas. Corre-se o risco de decisões contraditórias baseadas em um mesmo fato justamente pela dificuldade de avaliação de cada caso.

A questão central que se coloca é como a sociedade prefere avaliar a responsabilidade de agentes inteligentes em caso de dano e perigo a ela própria e,

¹⁷ BRYNJOLFSSON, E.; MCAFEE, A. *The Second Machine Age: work, progress, and prosperity in a time of brilliant technologies*. NY: W.W. Norton & Company, 2016.

¹⁸ SCHÖNBERGER, Viktor Mayer; RAMGE, Thomas. *Reinventing Capitalism in the Age of Big Data*. New York: Basic Books, 2018.

¹⁹ PENTLAND, Alex. *Social Physics: How Social Networks Can Make Us Smarter*. New York: The Penguin Press, 2014.

²⁰ KAUFMAN, Dora. *A inteligência artificial irá suplantar a inteligência humana?* Barueri: Estação das Letras e Cores, 2019.

²¹ HALLEVY, Gabriel. *Liability for Crimes Involving Artificial Intelligence Systems*. Cham: Springer International Publishing, 2015.

aqui, tem-se que o Direito Penal seria a medida mais eficiente, ainda que esteja longe de ser perfeita.

4.1 Responsabilidade penal decorrente da inteligência artificial sem autonomia

Não se desconhece a possibilidade de que, de alguma maneira, agentes inteligentes possam produzir ou contribuir para a lesão a bens jurídicos tutelados pelo Direito Penal. Não são, na atualidade, incomuns as notícias que dão conta de lesões corporais,²² mortes,²³ ou ofensas à honra,²⁴ por exemplo, perpetradas através de sistemas inteligentes.

Na opinião da maioria dos autores – não só neurocientistas²⁵ ou cientistas da computação,²⁶ mas dos poucos juristas dedicados ao tema²⁷ – os sistemas de inteligência artificial não possuem autonomia decisória até o momento e, portanto, sujeitam-se ao arbítrio humano, uma vez que:

[...] mesmo um agente inteligente altamente desenvolvido é, porém, apenas tão inteligente quanto sua programação. As regras de comportamento que lhe são oferecidas com base em consideráveis simulações não podem nunca lhe preparar para todos os casos pensáveis

²² “Um operário chinês sobreviveu ao ser espetado com dez espigas de metal quando um robô funcionou mal. Zhou, 49 anos, trabalhava no turno da noite em uma fábrica de porcelana na província de Hunan quando foi atingido por um braço robótico em queda. [...] Seis hastes de aço fixadas em uma placa de aço perfuraram seu ombro e peito direito, e quatro penetraram em outras partes do corpo [...]” (TAHIR, Tariq. “Robô de fábrica empala trabalhador com espigas de aço de 10 pés de comprimento após mau funcionamento.” *The Sun*, 14 dez 2018. Disponível em: <https://www.thesun.co.uk/news/7954270/factory-robot-malfuctions-and-impales-worker-with-10-foot-long-steel-spikes/>. Acesso em: 19 dez. 2020).

²³ “Uma mulher de 49 anos morreu na cidade de Tempe, Arizona, Estados Unidos, ao ser atropelada por um veículo sem motorista operado pela Uber, segundo informou através de um comunicado a polícia local. ‘O veículo se dirigia para o norte quando uma mulher que caminhava fora da faixa de pedestres cruzou a calçada e foi atropelada pelo veículo da Uber’, assinalou a polícia em um comunicado. É o primeiro atropelamento fatal protagonizado por um carro autônomo” (CANO, Rosa Jiménez. Carro sem motorista da Uber provoca primeiro acidente fatal. *El País*, Nova York, 19 mar. 2018. Disponível em: https://brasil.elpais.com/brasil/2018/03/19/tecnologia/1521479089_032894.html. Acesso em: 19 dez. 2020).

²⁴ “Em junho de 2015, um usuário do Google Photos descobriu que o programa etiquetava seus amigos negros como gorilas. A inteligência artificial do Google não era capaz de distinguir a pele de um ser humano da dos macacos, como gorilas e chimpanzés. Esse viés racista da máquina forçou o Google a pedir desculpas e prometeu encontrar uma solução para o erro [...]” (SALAS, Javier. Google conserta seu algoritmo “racista” apagando os gorilas. *El País*, 16 jan. 2018. Disponível em: https://brasil.elpais.com/brasil/2018/01/14/tecnologia/1515955554_803955.html. Acesso em: 19 dez. 2020).

²⁵ BOSTROM, Nick. *Superinteligência: caminhos, perigos, estratégias*. Rio de Janeiro: DarkSide Books, 2018.

²⁶ NILSSON, Nils J. *The quest for artificial intelligence: a history of ideas and achievements*. Cambridge: Cambridge University Press, 2010.

²⁷ PAGALLO, Ugo. *The Laws of Robots: Crimes, Contracts, and Torts*. New York: Springer, 2015.

de mudança da vida real. Some-se a isso o fato de que os agentes inteligentes não podem, sem mais, como podem os humanos, transpor para novas situações, seja por analogia ou por intuição, as regras com as quais foram originariamente programados.²⁸

Assim, não seria possível, na atualidade, considerar as máquinas como agentes autônomos, dotados de consciência, vontade ou culpabilidade; qualquer resultado criminoso provocado através da inteligência artificial decorreria da conduta humana relacionada ao programador, ao fabricante ou ao usuário, conforme o caso. Aplica-se, nesses termos e sem maior dificuldade, as regras estabelecidas pelo Código Penal²⁹ e Código de Defesa do Consumidor.³⁰

Pode-se imaginar, por exemplo, a figura de um programador que planeje a morte do seu desafeto através do escape de gases venenosos pelo robô. O ser humano por trás da máquina, no caso, seria autor do crime de homicídio, utilizando a máquina como mero instrumento.

É fato, porém, que a responsabilidade penal relacionada ao uso da inteligência artificial caracteriza-se, normalmente, como um dano indesejado e, nesses termos, tipifica um delito culposo.

Um uso rotineiro que tem sido emprestado à inteligência artificial refere-se ao auxílio à tomada de decisão, com o uso de sistemas que processam muitas informações e oferecem conclusões orientadas ao conhecimento analisado de modo a facilitar as decisões humanas. Imagine-se, por exemplo, um sistema inteligente de diagnósticos que indique certo tratamento a um paciente; a decisão final e, consequentemente, a responsabilidade penal, continua sendo do médico que, através do seu próprio conhecimento, deve avaliar a regularidade e a confiabilidade da indicação. Aqui, não se trata necessariamente de falha do equipamento ou erro do programador; o sistema ofereceu uma resposta possível diante das informações disponíveis sendo certo que sua função evidente é de auxiliar as decisões humanas e não substituí-las.

Outras situações de lesão a bens jurídicos, entretanto, surgem em razão de falhas na concepção do algoritmo, ocasionando a respectiva responsabilização penal do ser humano por trás da inteligência artificial. Ocorre quando o programador projeta a máquina sem visualizar – ainda que devesse – os danos que ela poderá

²⁸ GLESS, Sabine; WEIGEND, Thomas. Agentes inteligentes e o direito penal. *In*: ESTELLITA, H.; LEITE, A. (Ed.). *Veículos autônomos e direito penal*. São Paulo: Marcial Pons, 2019, p. 39.

²⁹ BRASIL. *Decreto-Lei nº 2.848, de 7 de dezembro de 1940*. Código Penal. Brasília. 1940. Disponível em: http://www.planalto.gov.br/ccivil_03/decreto-lei/del2848compilado.htm. Acesso em: 15 jan. 2021.

³⁰ BRASIL. *Lei nº 8.072, de 11 de setembro de 1990*. Dispõe sobre a proteção do consumidor e dá outras providências. 1990. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l8078.htm. Acesso em: 10 jan. 2021.

causar, por manifesta desatenção em relação à situação de risco, produzindo um resultado não desejado e não previsto, ainda que previsível.

Foi o que restou apurado em relação a dois acidentes envolvendo aviões da empresa Boeing. Em 29 de outubro de 2018 um avião 737 MAX, novo modelo da companhia, caiu no mar de Java com 189 pessoas a bordo, não restando sobreviventes. Em 10 de março de 2019, o mesmo modelo de avião caiu na Etiópia, matando os seus 157 ocupantes. Segundo investigações, um novo sistema de aumento de características de manobras (MCAS), baseado em inteligência artificial, teria contribuído para os eventos ao, equivocadamente, interpretar que o avião corria risco de perder sustentação (estolar) e, isoladamente, insistiu em abaixar o nariz da aeronave enquanto os pilotos tentavam, sem sucesso, conduzir o avião à sua posição regular. Além de outras falhas relacionadas ao projeto da aeronave, constatou-se em um dos acidentes que a disputa homem-máquina relacionada ao acionamento automático do MCAS ocorreu por vinte e seis vezes, até a queda do avião.³¹ A inteligência artificial, nesse caso, foi erroneamente programada para sobrepor-se às ordens do piloto que, sem treinamento adequado, não conseguiram desligar o equipamento e impedir a preponderância da máquina. Após mais de 20 meses impedidos de voar, os aviões 737 MAX da Boeing voltaram a transportar passageiros em dezembro de 2020, após revisão completa do projeto da aeronave.³²

Outro evento em que a inteligência artificial sem autonomia provocou morte ocorreu na Flórida, Estados Unidos, em 07 de maio de 2016, quando um carro semiautônomo da Tesla não detectou um caminhão que cruzou a sua frente em razão do reflexo do sol e, por isso, não ativou a frenagem de segurança. O motorista, por sua vez, não teve tempo de acionar os freios do carro que se chocou com um poste, vindo a falecer.³³ No relatório elaborado pela *National Transportation Safety Board* (NTSB) – órgão do governo americano encarregado de investigar acidentes de transportes, identificar suas prováveis causas e emitir recomendações de segurança –, constatou-se que o excesso de confiança do motorista no sistema de condução semiautônoma do Tesla Model S foram determinantes para o evento. A falta de reação do motorista no acidente decorreu, ainda, do seu desconhecimento dos limites da tecnologia e do projeto operacional do veículo que permitiu seu prolongado desengajamento da tarefa de dirigir. Recomendou-se, a final, que

³¹ BICEGO, Bruno Ett; HIGASI, Plínio. A queda dos aviões Boeing 737 MAX: um alerta para o Direito Digital no Brasil. *Digital Rights*, 9 abr. 2019. Disponível em: <https://digitalrights.cc/blog/2019/04/09/a-queda-dos-avioes-boeing-737-max-um-alerta-para-o-direito-digital-no-brasil/>. Acesso em: 20 dez. 2020.

³² JUSTO, Gabriel. Após acidentes e aterramento, 737 MAX volta a voar. O que aconteceu? *Exame*, 7 dez. 2020. Disponível em: <https://exame.com/mundo/apos-acidentes-e-aterramento-737-max-volta-a-voar-o-que-aconteceu/>. Acesso em: 25 jul. 2021.

³³ VENTURA, Felipe. *Motorista morre em acidente ao dirigir carro da Tesla em modo semiautônomo*. Disponível em: <https://gizmodo.uol.com.br/acidente-morte-tesla-autopilot/>. Acesso em: 18 dez. 2020.

as empresas fabricantes de veículos semiautônomos incorporem proteções que limitem o uso dos sistemas automatizados de controle dos veículos bem como desenvolvam aplicativos que monitorem o necessário envolvimento do condutor na utilização dos sistemas inteligentes.³⁴

Nos dois casos – Boeing e Tesla – não se pode falar em inteligência artificial autônoma considerando a existência de nexos causal entre a conduta humana – culpa do programador, do fabricante ou do próprio usuário, por exemplo, e a possibilidade de responsabilização penal do ser humano por trás do agente inteligente. No acidente com o Model S da Tesla, contudo, não foi detectada qualquer falha no sistema que, mesmo permitindo grande autonomia ao condutor do veículo em razão da tecnologia inteligente adotada, não o eximia de colocar as mãos nos volantes em certos períodos. Dos 37 minutos de viagem em que seria necessária a participação do motorista na condução do veículo, suas mãos foram reconhecidas pelo sistema durante, apenas, 25 segundos. A Tesla, por sua vez, sempre deixou evidente aos proprietários dos seus veículos as limitações do sistema *autopilot*, bem como a necessidade de interação do motorista nos deslocamentos, eximindo-a de qualquer responsabilidade nesse caso.³⁵ Já nos acidentes relacionados ao Boeing 737 Max, não restaram dúvidas acerca das falhas de sistema que geraram as quedas dos aviões e a morte dos seus ocupantes. Nesse sentido, a empresa foi criminalmente acusada de conspiração para fraudar o governo americano, quando seus funcionários esconderam informações cruciais sobre as condições da aeronave da *Federal Aviation Administration* (FAA), responsável pela segurança aérea nos Estados Unidos. Diante dos fatos, a Boeing concordou em pagar mais de US\$2,5 bilhões, encerrando a demanda criminal com o Departamento de Justiça americano.³⁶

Não se desconhece, entretanto, as dificuldades de imputação penal no âmbito das relações de consumo no Brasil e, nestes termos, a adequada proteção dos usuários em decorrência de defeitos e consequentes danos causados pelo emprego da inteligência artificial, ainda que sem autonomia. Inicialmente, temos que a responsabilidade penal é definida pela aplicação de diversos princípios de estatura constitucional que visam evitar abusos no exercício do *jus puniendi* estatal, como a legalidade estrita, taxatividade, lesividade, intervenção mínima

³⁴ SUMWALT III, Robert L *et al.* Accident Report NTSB/HAR-17/02 PB2017-102600. *National Transportation Safety Board*. Washington, 12 set. 2017. Disponível em: <https://www.nts.gov/investigations/AccidentReports/Reports/HAR1702.pdf>. Acesso em: 25 jul. 2021.

³⁵ SHEPARDSON, David. Tesla driver in fatal 'Autopilot' crash got numerous warnings: U.S. government. *Reuters*, 19 jun. 2017. Disponível em: <https://www.reuters.com/article/us-tesla-crash-idUSKBN19A2XC>. Acesso em: 31 jul. 2021.

³⁶ UNITED STATES. Department of Justice. Boeing Charged with 737 Max Fraud Conspiracy and Agrees to Pay over \$2.5 Billion. *Justice News*, 7 jan. 2021. Disponível em: <https://www.justice.gov/opa/pr/boeing-charged-737-max-fraud-conspiracy-and-agrees-pay-over-25-billion>. Acesso em: 31 jul. 2021.

e culpabilidade. Some-se a isso a massificação do consumo, o lançamento de milhares de novos produtos e os embaraços científicos de uma clara definição dos riscos que dificultam a caracterização da autoria e donexo causal. No âmbito de organizações complexas como as atuais, a fabricação e a distribuição de produtos defeituosos não são atos isolados ou facilmente individualizáveis, uma vez que decorrem de empreitada coletiva, decisões colegiadas, execução por partes, com o envolvimento de inúmeros funcionários. Aumenta-se enormemente, nestes termos, a dificuldade de definição da autoria, a especificação do dolo ou da culpa, a fixação de limites entre o risco permitido e o ilícito previsível, a aferição do grau de culpabilidade de cada conduta, dentre vários outros aspectos problemáticos que expõem a fragilidade dos institutos dogmáticos tradicionais do Direito Penal no tratamento da matéria.³⁷

É possível e indicado, neste contexto, pensar-se na responsabilização penal da pessoa jurídica por trás da inteligência artificial sem autonomia. Não se descarta que, desde os tempos mais remotos, o direito tem oscilado entre a punição individual e coletiva. Da Idade Antiga à Idade Média, as sanções eram impostas, predominantemente, aos entes coletivos como tribos, cidades, vilas e famílias. Já após a Revolução Francesa, as sanções coletivas deixaram de ser impostas com fundamento, para uns, no pensamento iluminista que pugnava pela preservação das liberdades individuais; para outros, pelo simples fim do absolutismo que, enquanto vigorou, necessitava punir os entes coletivos que ameaçavam a sua soberania.

Este movimento pendular perdura até os dias atuais: enquanto a responsabilidade penal das pessoas jurídicas foi tratada de forma dogmática no século XIX, sem a abertura de espaços para a discussão da punição dos entes coletivos, a partir do século XX observa-se uma retomada gradativa da responsabilização criminal das pessoas jurídicas. O fato é que o comportamento criminoso viola regras sociais e merece reprimenda penal, seja praticado por pessoa física ou jurídica; afinal, se o próprio fundamento da culpabilidade individual se baseia em uma representação das coisas do mundo e da vida, é possível entender a sua aplicabilidade como juízo de reprovação a outros entes que não apenas os seres humanos.³⁸

Atualmente, países da *common law* (Estados Unidos, Austrália) reconhecem a responsabilidade penal das pessoas jurídicas, seguidos por alguns países da *civil law* que também reconhecem essa premissa (Espanha). Outros países rejeitam totalmente a imputação penal aos entes coletivos (Itália) e um terceiro grupo (Alemanha) adota postura intermediária, com a imposição de sanções às pessoas

³⁷ BOTTINI, Pierpaolo Cruz. Indagações sobre a responsabilidade penal da empresa pelo produto defeituoso. *Boletim IBCCRIM*, v. 16, n. 194, jan. 2009.

³⁸ SHECAIRA, Sérgio Salomão. *Responsabilidade penal da pessoa jurídica*. Rio de Janeiro: Elsevier, 2011.

jurídicas através do direito penal administrativo, com finalidade mais pragmática em que não se questiona a culpabilidade da empresa.³⁹

A Constituição Federal, em seu art. 225, §3º dispõe que “as condutas e atividades consideradas lesivas ao meio ambiente sujeitarão os infratores, pessoas físicas ou jurídicas, a sanções penais e administrativas, independentemente da obrigação de reparar os danos causados”.⁴⁰ O art. 173, por sua vez, dispõe que:

Art. 173 [...]

§5º. A lei, sem prejuízo da responsabilidade individual dos dirigentes da pessoa jurídica, estabelecerá a responsabilidade desta, sujeitando-a às punições compatíveis com sua natureza, nos atos praticados contra a ordem econômica e financeira e contra a economia popular.⁴¹

Os Tribunais Superiores (STF e STJ) já pacificaram o entendimento no sentido da possibilidade de responsabilização da pessoa jurídica no Brasil. Nesse sentido:

As pessoas jurídicas tornaram-se destinatárias da lei penal desde 1988, há 25 anos portanto, em decorrência de imposição expressa da norma constitucional acima transcrita. A Lei 9.605, de 12 de fevereiro de 1998, conferiu a possibilidade de concreção da Constituição, ao estipular os pressupostos e as penas aplicáveis às pessoas jurídicas [...]. Os argumentos teóricos e as concepções abstratas do modelo dogmático da ciência penal tradicional, embasados na ação do indivíduo (*societas delinquere non potest*), não convenceram o legislador constitucional originário, e, desse modo, são insuficientes para que se afirme a ilegitimidade da opção feita.⁴²

Haveria, portanto, plena e indiscutível possibilidade de responsabilização penal da pessoa jurídica por crimes econômicos, financeiros e contra a economia popular uma vez que os entes coletivos já respondem, com as punições compatíveis, pelos delitos ambientais.⁴³

³⁹ SHECAIRA, Sérgio Salomão. *Responsabilidade penal da pessoa jurídica*. Rio de Janeiro: Elsevier, 2011.

⁴⁰ BRASIL. *Constituição da República Federativa do Brasil*. Brasília. 1988. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 01 jan. 2021.

⁴¹ BRASIL. *Constituição da República Federativa do Brasil*. Brasília. 1988. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 01 jan. 2021.

⁴² BRASIL. Supremo Tribunal Federal. *Recurso Extraordinário nº 548.181*. Primeira Turma. Relator: Min. Rosa Weber. Brasília, 6 ago. 2013, p. 45. Disponível em: <http://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=7087018>. Acesso em: 19 dez. 2020.

⁴³ NUCCI, Guilherme de Souza. *Curso de direito penal: parte geral*. Rio de Janeiro: Forense, 2019.

Para a responsabilização criminal da pessoa jurídica, entretanto, alguns requisitos devem ser observados. Inicialmente, deve-se assegurar que o delito individual tenha sido praticado no interesse da pessoa jurídica: caso o interesse seja exclusivo da pessoa física, fica excluída a responsabilidade do ente coletivo. Outra exigência é que o crime praticado pela pessoa física esteja situado no âmbito normal de atividades da pessoa jurídica. Ainda, a infração penal deve ter sido praticada por pessoa que pode agir em nome do ente coletivo demonstrando estreita ligação e confluência de interesses entre a pessoa física e a pessoa jurídica. Por fim, é preciso que o delito praticado pela pessoa física tenha se servido do poderio resultante da pessoa jurídica; em outras palavras, as pessoas sob o manto protetor da pessoa jurídica, praticaram crime que, isoladamente, não seria possível.⁴⁴

Acerca de uma alegada suficiência da adoção de sanções de natureza civil e administrativa às pessoas jurídicas, Busato ressalta que o Direito Penal é reservado a intervenções graves e, nesses casos, em razão do seu viés garantista, deve ser empregado também em relação às pessoas jurídicas.⁴⁵ Caso se utilize outros ramos do direito para intervenções gravosas nos entes coletivos o que ocorrerá, em suma, é uma fraude de etiquetagem de atos de essência penal travestidos de mecanismos de garantias mínimas. Na atualidade, não se desconhece que crimes graves, por vezes, são perpetrados através das pessoas jurídicas, exigindo do direito uma reação enérgica e contundente que só o Direito Penal é capaz de oferecer. Além da sua maior eficácia, a sanção penal imprime uma reprovação estatal consistente e perene na pessoa jurídica infratora.

Em regra, a distinção entre o emprego do direito penal ou outros ramos jurídicos para a responsabilização da pessoa jurídica deve basear-se na qualificação jurídica do interesse violado e não na natureza do sujeito que cometeu a violação. Assim, qualquer pessoa, física ou jurídica, que lesionar interesses relevantes e, portanto, elevados à categoria de bem jurídico-penal, merece a reprovação penal.⁴⁶

Conclui-se, portanto, pela desejável responsabilização penal da pessoa jurídica nos casos relacionados à utilização da inteligência artificial sem autonomia em regimes empresariais complexos, considerando a dificuldade de individualização de condutas e respectivo nexos causal. Nesses termos, casos como os acidentes ocorridos com o Boeing 737 MAX, além da eventual responsabilização administrativa e civil da empresa, ensejariam a responsabilização penal da pessoa jurídica, sem prejuízo de possível imputação criminal às pessoas físicas culpadas. Considerando o disposto no art. 173, §5º da Constituição Federal, restaria a necessária elaboração de lei ordinária regulando o tema.

⁴⁴ SHECAIRA, Sérgio Salomão. *Responsabilidade penal da pessoa jurídica*. Rio de Janeiro: Elsevier, 2011.

⁴⁵ BUSATO, Paulo César. *Direito penal*. São Paulo: Atlas, 2015.

⁴⁶ SHECAIRA, Sérgio Salomão. *Responsabilidade penal da pessoa jurídica*. Rio de Janeiro: Elsevier, 2011.

4.2 Responsabilidade penal do ser humano por trás da inteligência artificial autônoma

Não se ignora a evolução exponencial da inteligência artificial nos últimos anos através, principalmente, da sua capacidade de aprendizagem. Busca-se, a cada dia, a superação de obstáculos que levem máquinas a pensarem em termos similares aos seres humanos e, se possível, superá-los na solução de problemas complexos. Para tanto, uma das características pretendidas às máquinas pelos programadores é a aptidão para, diante de casos abstrusos, responder de modo rápido, autônomo e, por isso mesmo, imprevisível.

A análise da responsabilidade penal da inteligência artificial autônoma traz consigo duas questões opostas e indissociáveis. A primeira, se o ser humano que se relaciona com esse agente inteligente, seja na sua programação, fabricação ou mero uso, pode e deve responder penalmente por danos causados pela máquina. A segunda – dependente da primeira – se a inteligência artificial autônoma pode e deve ser responsabilizada penalmente.

Considerando a possibilidade de existência futura (próxima ou remota) da inteligência artificial autônoma, importa-nos avaliar se qualquer dos humanos que interagem com a agentes inteligentes mantém a dirigibilidade de suas condutas em relação aos resultados ocorridos através da atuação autônoma da máquina, tudo de acordo com o sistema finalista adotado pelo direito brasileiro.

A conduta, no finalismo, é composta dos elementos consciência e vontade; portanto, é forçoso entender que o ser humano, ao relacionar-se com a inteligência artificial independente, tem noção clara da realidade, quer determinado resultado e atua deliberadamente nesse sentido. Como a regra é termos consciência de nossos atos, é possível reconhecê-la na conduta dos seres humanos que lidam com uma inteligência artificial autônoma, seja o programador, fabricante ou usuário.

Os impasses começam a surgir na análise dos aspectos psíquicos que conformam o conteúdo da conduta. A vontade é um querer ativo, que se movimenta em direção ao resultado e implica sempre em uma finalidade, ou seja, uma vontade sem conteúdo não é concebível. Em essência, pode-se afirmar que:

[...] a doutrina da imputação penal carece, até o momento de um instrumental adequado para decidir adequadamente sobre a atribuição de responsabilidade no âmbito da interação entre os (possivelmente menos inteligentes) seres humanos e as (possivelmente mais inteligentes) máquinas.⁴⁷

⁴⁷ GLESS, Sabine; WEIGEND, Thomas. Agentes inteligentes e o direito penal. In: ESTELLITA, H.; LEITE, A. (Ed.). *Veículos autônomos e direito penal*. São Paulo: Marcial Pons, 2019, p. 62.

Constata-se, nesse contexto, a injustiça da responsabilização penal do ser humano por danos causados pelo agente artificial autônomo através da análise dos fundamentos da teoria do delito que exige, para a sua configuração, um fato típico, ilícito e culpável. Na constituição do fato típico, não se pode atribuir dolo ou culpa ao humano por trás da máquina autônoma. Faltaria, ademais, à eventual conduta humana dolosa ou culposa, nexos de causalidade em relação ao resultado danoso provocado, a rigor, pela decisão independente do agente inteligente. Por fim e sintetizando a tese de irresponsabilidade penal do ser humano por trás da máquina inteligente, falta-lhe a decisão pelo injusto quando, no caso concreto, poderia atuar de modo diverso, que é pressuposto da culpabilidade. Não há, *in casu*, um juízo de reprovação acerca do comportamento do ser humano imputável, com potencial consciência da ilicitude efetivada pela máquina e sobre a qual pudesse fazer atuar conforme o direito.

Na análise da culpa relacionada ao ser humano por trás de uma inteligência artificial autônoma, deve-se observar que a pessoa não necessariamente incorrerá na violação de um dever de cuidado objetivo. Ao programar ou fabricar um agente inteligente, o ser humano dedica seu conhecimento e energia na elaboração de um produto que, efetivamente, tenha a capacidade de decidir de modo independente, baseado em dados complexos e com resultados, muitas vezes, imprevisíveis. Não se desconhece, afinal, as diversas vantagens das máquinas autônomas em relação aos seres humanos: a capacidade de decisões rápidas e complexas, com base em um número elevado de dados; potencial de trabalho ininterrupto, sem cansaço, sono, emoções, desvios de finalidade, fome, direitos trabalhistas, dentre outras.

A previsibilidade objetiva – “quando a previsão do advento, no caso concreto, podia ser exigida do homem normal, do *homo medius* [...]”⁴⁸ – também não se sustenta na relação homem/máquina autônoma. Considerando que a máquina decide independentemente da participação humana, qualquer delito praticado pelo agente autônomo não decorre de uma previsibilidade antecedente do ser humano envolvido com a sua criação ou seu desenvolvimento. O fato é que no complexo ambiente de processamento algorítmico de agentes inteligentes não se pode impor responsabilidade aos seres humanos programadores ou usuários em razão do resultado estocástico proporcionado pela máquina.⁴⁹ Em suma, pode-se afirmar que:

⁴⁸ HUNGRIA, Nélon. *Comentários ao Código Penal*. Rio de Janeiro: Forense, 1949, v. 1, p. 357.

⁴⁹ KARNOW, Curtis E. A. Liability for distributed artificial intelligences. *Berkeley Technology Law Journal*, v. 11, n. 1, 1996. Disponível em: <https://lawcat.berkeley.edu/record/1115611>. Acesso em: 04 ago. 2021.

[...] os agentes inteligentes avaliam autonomamente as informações tomadas do seu entorno e reagem ao produto dessa avaliação sem uma influência humana, de forma a cumprir de forma ótima a missão que lhes foi assignada [...] Isso quer dizer que uma certa imprevisibilidade dos agentes inteligentes e o correlato risco para terceiros é como que ‘pré-programada’.⁵⁰

Há, contudo, posicionamentos no sentido de que se deve adotar uma previsibilidade penal em sentido amplo, ou seja, em qualquer circunstância, o ser humano por trás da máquina tem condições de prever a ocorrência de um dano provocado pelas máquinas inteligentes e, nesses termos, responde penalmente pelo resultado. O professor de Direito Penal da Universidade do País Basco (Espanha) Carlos María Romeo Casabona, por exemplo, entende que pelo menos em curto ou médio prazo, o Direito Penal dispõe de instrumentos adequados para responsabilizar penalmente os seres humanos que intervierem na concepção, fabricação, distribuição ou utilização destes sistemas inteligentes, ainda que autônomos. Nesse sentido, assevera que deve haver o que chama de “controle humano significativo” (CHS) – mecanismos jurídicos de defesa dos humanos e não dos computadores ou algoritmos – e recomenda que esse modo de controle se torne um padrão para as novas tecnologias: “é nosso dever como espécie que de alguma forma domina o planeta e seus seres vivos, apesar de, infelizmente, abusar muitas vezes desse poder”.⁵¹ Para Hanson, qualquer que seja o papel dos objetos não humanos, não se desconhece que, em última análise, há a participação e o interesse humano envolvidos; nesse sentido, toda ação das máquinas inteligentes é atribuível aos indivíduos humanos, que devem ser responsabilizados.⁵²

Esse posicionamento, contudo, repercute em inaceitável imputação da responsabilidade penal entre os seres humanos por erro próprio do agente inteligente. Em regra, entendidos como previsíveis todos os resultados danosos produzidos pelas máquinas, o pressuposto da previsibilidade penal objetiva cai por terra, inexistindo obstáculos a “uma abrangente responsabilidade culposa do fornecedor ou operador por qualquer comportamento danoso do agente inteligente”.⁵³ Nesse

⁵⁰ KARNOW, Curtis E. A. Liability for distributed artificial intelligences. *Berkeley Technology Law Journal*, v. 11, n. 1, 1996, p. 55. Disponível em: <https://lawcat.berkeley.edu/record/1115611>. Acesso em: 04 ago. 2021.

⁵¹ ALONSO, Raúl. ¿Un robot puede ser responsable penal del delito que comete? *Mutualidad Abogacia*, 21 abr. 2021. Disponível em: <https://tudefinestufuturo.mutualidadabogacia.com/innovacion/un-robot-puede-ser-responsable-penal-del-delito-que-comete/>. Acesso em: 02 ago. 2021.

⁵² HANSON, F. Allan. Beyond the skin bag: on the moral responsibility of extended agencies. *Ethics and Information Technology*, v. 11, p. 91-99, 2009. Disponível em: https://kuscholarworks.ku.edu/bitstream/handle/1808/13802/Skin_Bag_Hanson.pdf?sequence=1&isAllowed=y. Acesso em: 06 ago. 2021.

⁵³ GLESS, Sabine; WEIGEND, Thomas. Agentes inteligentes e o direito penal. In: ESTELLITA, H.; LEITE, A. (Ed.). *Veículos autônomos e direito penal*. São Paulo: Marcial Pons, 2019, p. 56.

sentido, Matthias argumenta que, no caso das máquinas autônomas, surge uma “lacuna de responsabilidade” uma vez que o agente humano que a programou não exerce mais controle direto sobre as ações da máquina, pois são criadas para que assumam, gradualmente, suas decisões algorítmicas. Portanto, seria injusto responsabilizar os humanos por ações de máquinas sobre as quais eles não poderiam ter controle.⁵⁴

Acerca da culpabilidade, também não parece plausível falar na reprovação do ser humano por trás da inteligência artificial diante de uma lesão provocada de modo absolutamente independente da conduta humana. Aqui, como alhures, a definição é binária e excludente: caso falemos de uma inteligência artificial definitivamente autônoma, não há como se pensar em qualquer vínculo entre criatura e criador. Obviamente que, havendo qualquer tipo de nexos entre a conduta desses agentes humanos e o resultado criminoso produzido pela inteligência artificial, não se trata de entidade independente restando, nesses casos, a culpabilidade do ser humano por trás da inteligência artificial sem autonomia. Em detida análise sobre a responsabilização dos danos causados por máquinas inteligentes, Curtis Karnow, juiz do Tribunal Superior do Condado de San Francisco, nos Estados Unidos, assevera que só faz sentido responsabilizar alguém que tenha o poder de evitar a lesão provocada pela inteligência artificial autônoma.⁵⁵ Ao tentarmos usar um sistema jurídico tradicional para o tratamento das decisões algorítmicas complexas, pessoas serão injustamente punidas por danos que elas não poderiam ter evitado ou previsto.

Para Ryan Abbott e Alex Sarch, pode-se imaginar várias formas de utilização do Direito Penal existente para estabelecer responsabilidade criminal para desenvolvedores ou usuários da inteligência artificial autônoma.⁵⁶ O problema é que isso não parece equivaler a uma conduta humana reprovável, considerando a complexidade e os riscos inerentes às decisões autônomas das máquinas. Esta expansão do direito penal sufocaria a inovação e os benefícios comerciais esperados da tecnologia. Na verdade, se houvesse tal criminalização, a maioria dos primeiros desenvolvedores da *internet*, provavelmente, seriam considerados culpados dos eventos algorítmicos danosos da atualidade.

⁵⁴ MATTHIAS, Andreas. The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics and Information Technology*, v. 6, p. 175-183, 2004. Disponível em: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.456.8299>. Acesso em: 06 ago. 2021.

⁵⁵ KARNOW, Curtis E. A. Liability for distributed artificial intelligences. *Berkeley Technology Law Journal*, v. 11, n. 1, 1996. Disponível em: <https://lawcat.berkeley.edu/record/1115611>. Acesso em: 04 ago. 2021.

⁵⁶ ABBOTT, Ryan; SARCH, Alex. Punishing Artificial Intelligence Legal Fiction or Science Fiction. *University of California*, v. 53, 10 out. 2019. Disponível em: https://lawreview.law.ucdavis.edu/issues/53/1/articles/files/53-1_Abbott_Sarch.pdf. Acesso em: 05 ago. 2021.

4.3 Responsabilidade penal da inteligência artificial autônoma

Uma discussão que permeia toda a evolução do Direito Penal diz respeito à sua própria estabilidade como ciência e os métodos que adota para a sua orientação. A importância dessa discussão é incontestável e leva-nos a dois desfechos antagônicos. Desenvolvido ao longo de séculos, o Direito Penal recebe a influência de fundamentos filosóficos e concepções ideológicas que, segundo Zafaroni e Pierangeli, orientam seus métodos.⁵⁷ De um lado, diante de métodos rígidos e pouca disposição a mudanças conceituais, temos um Direito Penal mais estável, sistematizado e juridicamente seguro; essa mesma estabilidade, porém, dificulta um avanço da ciência em direção aos anseios sociais e às mudanças de paradigmas. Em outro sentido, caso o Direito Penal adote sistemática flexível, dinâmica e atualizada, ganha em capacidade de resposta às demandas sociais e inovações, mas provoca, concomitantemente, inequívoca insegurança jurídica e estrutural.

O método atualmente mais difundido é o dogmático que consiste em uma análise da letra da lei, na sua decomposição analítica em elementos e na sua reconstrução de modo coerente. O princípio básico do método dogmático, emprestado do próprio sentido da palavra “dogma”, é que o intérprete não pode alterar o conteúdo da lei. Outras disposições relacionadas ao método dogmático pretendem que a construção dos elementos em análise seja conforme a lógica, não contrarie o texto da lei e seja convenientemente simétrica, natural e não artificiosa. Assim, o jurista, como qualquer cientista, busca elaborar um sistema não contraditório de proposições em busca da verdade e, por isso, “a pretensão de rejeitar o método dogmático, por considerá-lo inútil, é um infantilismo jurídico, próprio dos pretensos teóricos gerais que jamais enfrentaram os problemas concretos de algum ramo do saber jurídico”.⁵⁸ Roxin, por sua vez, elenca as diversas vantagens do pensamento sistemático: facilita o exame do caso penal em suas diversas etapas analítica; presuppõe a aplicação uniforme do direito segundo o seu objeto; simplifica e melhora o manejo do direito; guia a elaboração do desenvolvimento do direito, dentre outras.⁵⁹

A par de todos os méritos do método dogmático, Ramos defende que “se a teoria jurídica é bela, mas não serve para solucionar os problemas concretos, deve ser descartada como inservível, já que a teoria não é o mais importante, mas sim o resultado que permite alcançar”.⁶⁰ Roxin também enumera perigos

⁵⁷ ZAFFARONI, Eugenio Raul; PIERANGELI, José Henrique. *Manual de direito penal brasileiro*. São Paulo: Thomson Reuters, 2019.

⁵⁸ ZAFFARONI, Eugenio Raul; PIERANGELI, José Henrique. *Manual de direito penal brasileiro*. São Paulo: Thomson Reuters, 2019, p. 154.

⁵⁹ ROXIN, Claus. *Funcionalismo e imputação objetiva no direito penal*. Rio de Janeiro: Renovar, 2002.

⁶⁰ RAMOS, João Gualberto Garcez. *A inconstitucionalidade do direito penal do terror*. Curitiba: Juruá, 1991, p. 73.

do pensamento sistemático: a desatenção à justiça do caso concreto, a redução da possibilidade de solução de problemas, deduções sistemáticas incompatíveis com as diretrizes político-criminais e a utilização de conceitos demasiadamente abstratos.⁶¹

Já há algum tempo, persistem controvérsias em razão de inovações e descobertas científicas e sua afetação no Direito Penal. Penalistas e neurocientistas controvêtem acerca da conformação de conceitos que dão sustentação ao instituto da culpabilidade penal; em suma, divergem os próprios estudiosos do direito acerca da aplicabilidade ou não dessas novas descobertas. Diante do interesse de alguns em repensar os dogmas da culpabilidade e responsabilidade penal, Hassemer assume posição enérgica e resoluta. Deixa claro que a ciência penal se diferencia das demais ciências, enfatizando as especificidades de métodos e interesses em cada área do saber humano.⁶² Portanto, a discussão entre neurocientistas sobre a real liberdade do ser humano para as suas decisões não deve contaminar o juízo de reprovação a ser elaborado na seara penal, uma vez que é o Direito Penal, no bojo de um processo, que deve definir as causas que bastam para a exclusão da responsabilidade penal de uma pessoa.

É, nesse contexto, que temos a atual discussão relacionada à inteligência artificial e certos conceitos da ciência jurídico-penal, mormente relacionados à conduta e à culpabilidade. Partindo do pressuposto relacionado à existência de uma máquina dotada de autonomia decisória e de ação, surge um impasse acerca da responsabilidade penal dos seres humanos em face da atuação dos agentes inteligentes. Nesse sentido, pode-se afirmar que:

A penetração de agentes inteligentes mais ou menos benéficos em nossa vida é um daqueles fenômenos que coloca a ordem jurídica e especialmente o direito penal de frente a questões que não podem ser verdadeiramente respondidas com os recursos da dogmática tradicional.⁶³

Surge, nesse ponto, um vazio de imputação penal, segundo os dogmas penais e o direito penal brasileiro em vigor: nem seres humanos podem ser penalizados por danos causados pela inteligência artificial autônoma – uma vez que não agiram com culpa ou não se verifica qualquer reprovabilidade em suas condutas –, nem

⁶¹ ROXIN, Claus. *Funcionalismo e imputação objetiva no direito penal*. Rio de Janeiro: Renovar, 2002.

⁶² HASSEMER, Winfried. Neurociências e culpabilidade em direito penal. In: BUSATO, P. C. (Ed.). *Neurociência e direito penal*. São Paulo: Atlas, 2014.

⁶³ GLESS, Sabine; WEIGEND, Thomas. Agentes inteligentes e o direito penal. In: ESTELLITA, H.; LEITE, A. (Ed.). *Veículos autônomos e direito penal*. São Paulo: Marcial Pons, 2019, p. 62.

os agentes inteligentes são capazes de conduta, portanto, são irresponsáveis perante o Direito Penal.

Desse impasse, surge a necessidade de discussão das opções viáveis para a responsabilização da inteligência artificial. Poder-se-ia, por exemplo, estabelecer uma personalidade jurídica às máquinas autônomas (*E-Person*) de modo a torná-las responsáveis pelos delitos que praticar.⁶⁴ Aqui, a teoria da realidade da pessoa jurídica idealizada por Gierke e adotada no Brasil, estabelece que:

[...] pessoas jurídicas são pessoas reais, dotadas de uma real vontade coletiva, devendo ser equiparáveis, como seres sociais que são, às pessoas físicas. Excetuando-se determinadas relações que por sua natureza são incompatíveis com tais pessoas jurídicas, sua capacidade é em tudo equivalente à do homem.⁶⁵

Os animais, assim como a inteligência artificial, são considerados coisas pelo Direito e, nestes termos, poder-se-ia pensar na adoção de idêntico tratamento penal em relação às máquinas inteligentes, mormente considerando a autossuficiência de que ambos usufruem. Ocorre, contudo, que tal aparência de similaridade não se confirma, o que implica na impossibilidade de utilização do sistema jurídico emprestado aos animais no tocante à inteligência artificial autônoma.

É fato que os homens convivem há milênios com os animais, seja em razão da sua domesticação, do aproveitamento dos seus produtos, da sua utilização na agricultura, sempre estabelecendo regras no interesse dessa convivência. Para Hallevy, essas leis, normalmente, tratam de dois aspectos: animais como propriedade dos seres humanos e o dever de mostrar misericórdia para com os animais.⁶⁶ Das regras que tratam dessa relação com animais, observa-se que nenhum sistema jurídico atribui responsabilidade penal diretamente ao animal, independentemente de sua inteligência, diante de lesões causadas a terceiros; nesses casos, eventual culpa será do ser humano proprietário do animal. A misericórdia em relação aos animais, por sua vez, não se vincula diretamente ao animal e sim em razão dos interesses humanos, possibilitando a responsabilização criminal de quem praticar crueldade contra os bichos.

Trata-se, a rigor, de uma análise relacionada a três seres dotados de autonomia: seres humanos, animais e a inteligência artificial e, se o objetivo é conferir um tratamento jurídico à inteligência artificial próximo daquele dado aos seres

⁶⁴ ESTELLITA, Heloisa; LEITE, Alaor. *Veículos autônomos e direito penal*. São Paulo: Marcial Pons, 2019.

⁶⁵ SHECAIRA, Sérgio Salomão. *Responsabilidade penal da pessoa jurídica*. Rio de Janeiro: Elsevier, 2011, p. 90.

⁶⁶ HALLEVY, Gabriel. *Liability for Crimes Involving Artificial Intelligence Systems*. Cham: Springer International Publishing, 2015.

humanos, importante apontar semelhanças que indiquem essa adequada similaridade de tratamento. Caso contrário, o regime adotado em relação aos animais seria suficiente para regulação das ações das máquinas.

Não se desconhece que o modelo chamado por Hallevy de jurídico-zoológico possa responder satisfatoriamente algumas questões relacionadas à responsabilização das ações realizadas pelas máquinas, mas não todas.⁶⁷ Assim, quando a máquina aprende a decidir por conta própria, tem-se uma vinculação que se aproxima do pensamento humano e não do dos animais. Homens e máquinas decidem em razão de raciocínios fundamentados na lógica humana enquanto os animais, em regra, não possuem esse tipo de raciocínio, ainda que atuem com razoabilidade. A emoção desempenha papel importante na atividade dos seres vivos, aproximando daí homens e animais, o que não é caso das máquinas. Porém, para a lei e especialmente para o Direito Penal, a emocionalidade raramente é levada em consideração.

O modelo jurídico-zoológico ainda teria um papel pedagógico, educando seres humanos a serem misericordiosos com os animais, o que não acontece em relação às máquinas uma vez que não podem sofrer. Enquanto seria um evidente mau-tratos um humano cortar a perna de um cachorro sem motivo relevante, o mesmo não acontece em relação aos robôs. Considerando, portanto, que a lei prefere a racionalidade para a definição de responsabilidades e que essa racionalidade se baseia no raciocínio lógico humano, pode-se concluir que eventual responsabilidade penal da inteligência artificial deve pautar-se pelos mesmos critérios de imputação definidos para os seres humanos.⁶⁸

Importante enfatizar a existência de corrente crescente em todo o mundo que visa abandonar a ideia antropocêntrica de “coisificação” dos animais. Mesmo reconhecendo que, obviamente, não são pessoas humanas, é possível atribuir-lhes a titularidade de direitos civis e constitucionais, dotando-os de uma personalidade *sui generis* própria de sua condição.⁶⁹

Outro modelo possível de tratamento penal da inteligência artificial autônoma consiste na reinterpretação de diversos institutos jurídicos da *common law*, emprestando-lhes sentido mais amplo e reforçando a similaridade de vários instrumentos

⁶⁷ HALLEVY, Gabriel. *Liability for Crimes Involving Artificial Intelligence Systems*. Cham: Springer International Publishing, 2015.

⁶⁸ HALLEVY, Gabriel. *Liability for Crimes Involving Artificial Intelligence Systems*. Cham: Springer International Publishing, 2015.

⁶⁹ TOLEDO, Maria Izabel Vasco de. A tutela jurídica dos animais no Brasil e no direito comparado. *Revista Brasileira de Direito Animal*, Salvador, v. 7, n. 11, p. 197-223, dez. 2012. Universidade Federal da Bahia. Disponível em: <https://portalseer.ufba.br/index.php/RBDA/article/download/8426/6187>. Acesso em: 24 dez. 2019.

penais também vigentes no sistema da *civil law*.⁷⁰ Não se trata, contudo, da mera transposição de eventual análise de tratamento penal da inteligência artificial no sistema da *common law* para o nosso sistema jurídico-penal. Afinal, não se desconhece as peculiaridades históricas de evolução e os fundamentos específicos de cada sistema.

Essa reinterpretação da *common law*, útil na eventual incriminação da inteligência artificial, parte do pressuposto de que a caracterização da responsabilidade penal prescinde de certas habilidades humanas tidas por irrelevantes para a imputação de um delito. Os sistemas de inteligência artificial, em regra, buscam imitar a inteligência humana e, em muitos casos, superam os seres humanos em suas capacidades e habilidades. Na verdade, os próprios seres humanos são muito distintos entre si em diversos aspectos como força física, inteligência e conhecimento, por exemplo. Assim, ao invés de se buscar indefinidamente a equiparação da inteligência artificial à inteligência humana a fim de configurar a chamada *machina sapiens*, deve-se considerar os requisitos exigidos para a responsabilização penal. Atendidas essas condições lógicas, razoáveis e quase matemáticas de imputação penal, teríamos a chamada *machina sapiens criminalis*, ou seja, uma máquina inteligente que, sem pretender equiparar-se em tudo à inteligência humana, seria passível de responsabilização penal.⁷¹

Para Hallevy, as entidades inteligentes autônomas são capazes de condutas no sentido jurídico-penal.⁷² Igualmente aos seres humanos, máquinas inteligentes exteriorizam sua conduta através de uma ação que se projeta no mundo exterior, não restando grande dificuldade nessa constatação fática. Esse mesmo entendimento pode ser atribuído à omissão, configurada quando a inteligência artificial podia e devia agir para evitar um resultado danoso.

Para entender a possibilidade de consciência das máquinas inteligentes na prática de suas ações, importa entender o processo de formação da consciência humana, que acontece em duas etapas. Inicialmente, dados factuais são absorvidos pelos sentidos e, em seguida, uma imagem geral relevante sobre esses dados é gerada no cérebro. Isto posto, é possível dizer que a análise da consciência ocorre em termos binários: ou existe plenamente ou deve ser desconsiderada. A inteligência artificial, por sua vez e à semelhança dos sentidos humanos, possui dispositivos para absorver fatos relevantes – inclusive com maior capacidade de

⁷⁰ HALLEVY, Gabriel. *Liability for Crimes Involving Artificial Intelligence Systems*. Cham: Springer International Publishing, 2015.

⁷¹ HALLEVY, Gabriel. *Liability for Crimes Involving Artificial Intelligence Systems*. Cham: Springer International Publishing, 2015.

⁷² HALLEVY, Gabriel. *When Robot Kills. Artificial Intelligence under Criminal Law*. Northeastern: Northeastern University Press, 2013.

processamento como câmeras, sensores e microfones – transformando-os em representações importantes e adquirindo, assim, conhecimento.⁷³

Ainda que consciência e vontade sejam processos internos da mente e, nestes termos, ofereçam dificuldades na sua constatação, a vontade é ainda mais difícil de provar uma vez que se refere a fatos futuros enquanto a consciência relaciona-se a fatos atuais e, portanto, reais. Diante dessa dificuldade, o Direito Penal estabeleceu substitutos probatórios como a regra da previsibilidade, consistente na presunção de que o resultado decorrente da conduta é altamente provável. No caso da inteligência artificial, nada impede que se analise a ocorrência de um resultado também através da regra da previsibilidade, ou seja, a constatação da alta probabilidade de que o evento ocorresse a partir daquela ação, sempre considerando a conduta pautada pela consciência.⁷⁴

Para Hallevy, ainda que haja grande dificuldade no reconhecimento da inteligência artificial como dotada de consciência e vontade, não se deve olvidar que os requisitos mínimos exigidos pelo Direito Penal para a caracterização da responsabilidade penal inteligência artificial, ou seja, da *machina sapiens criminalis*, restam configurados.⁷⁵ Importa, aqui, não confundir a exigência desses mesmos elementos, também relevantes em outros ramos do pensamento humano (como a psicologia ou as ciências cognitivas) e que são utilizados na busca sem fim da *machina sapiens*.

No sistema da *common law*, as chamadas *defences* ou defesas gerais são entendidas como elementos que pretendem refutar eventual responsabilidade penal do infrator. Dividem-se em dois grupos, as isenções e as justificativas e, se alegadas, devem ser provadas pelo réu, importando em ausência de culpabilidade. As isenções relacionam-se às características pessoais do agente incluindo a infância, a perda de autocontrole, a insanidade, intoxicação, erros factuais, erros legais e imunidade substantiva; as justificativas, por sua vez, dependem das características factuais do evento e tratam da autodefesa, necessidade, coação, ordens superiores e defesa de *minimis*.⁷⁶

Tratando detalhadamente de cada uma dessas espécies de defesas gerais, questiona-se se a infância – que se refere na verdade ao grau de maturidade do agente –, pode ser atribuída à inteligência artificial. Ainda que se adote uma certa

⁷³ HALLEVY, Gabriel. *When Robot Kills. Artificial Intelligence under Criminal Law*. Northeastern: Northeastern University Press, 2013.

⁷⁴ HALLEVY, Gabriel. *When Robot Kills. Artificial Intelligence under Criminal Law*. Northeastern: Northeastern University Press, 2013.

⁷⁵ HALLEVY, Gabriel. *When Robot Kills. Artificial Intelligence under Criminal Law*. Northeastern: Northeastern University Press, 2013.

⁷⁶ Consiste, entre nós, no chamado princípio da insignificância, com fundamento no brocardo “de minimis no curat pretor” – “o julgador não se preocupa com coisas insignificantes”.

idade mínima na maioria dos países para considerar um ser humano presumidamente maduro, é possível que mesmo pessoas de mais idade não possuam, por algum motivo, maturidade suficiente para entender o caráter equivocado e ilícito de suas decisões. Em relação à inteligência artificial, os sistemas autônomos desenvolvem sua capacidade de aprendizagem de máquina com o passar do tempo. Assim, similarmente ao que acontece com humanos, seria possível imaginar uma presunção absoluta de inimputabilidade das máquinas autônomas baseada no seu ciclo de aprendizagem e maturação, definindo um tempo de uso baseado em evidências antes do qual, qualquer ofensa praticada pelo robô seria considerada isenta de responsabilidade criminal. Ademais, nada impede que, no caso concreto e superado esse período, seja avaliada a imputabilidade do agente robótico como ocorre na verificação da capacidade de entendimento dos seres humanos.⁷⁷

A perda do autocontrole diz respeito a situações em que o infrator não tem domínio dos movimentos corporais, cuja defesa geral baseia-se na lógica de que a ação perpetrada não corresponde à vontade do agente e, portanto, não merece reprimenda penal. Relacionado aos seres humanos, é o que ocorre por exemplo nos casos de convulsões e sonambulismo. Assim, nada impede que a inteligência artificial tenha, em certas situações, a perda total do seu autocontrole e possa reivindicar essa defesa geral, como diante de um mau funcionamento interno e que cause movimentos descontrolados; mesmo consciente da falha, não tem controle sobre seus atos.⁷⁸

A insanidade consiste em perturbações mentais que causem danos cognitivos ou avarias volitivas e pode ser aplicada à inteligência artificial, segundo Hallevy, desde que observadas as suas peculiaridades.⁷⁹ Assim, se um defeito interno da máquina causa essas mesmas consequências, ou seja, mau funcionamento cognitivo ou volitivo do sistema, teremos a configuração da insanidade. Temos por intoxicação a perturbação mental causada por substâncias externas ao organismo e que interferem na percepção do infrator e na compreensão da realidade factual. Aqui, se não há diferença entre os efeitos produzidos por essas substâncias em humanos ou máquinas inteligentes também não há justificativa para diferenciar a aplicação da referida defesa geral.

O erro real ou factual ocorre nas situações em que há uma interferência na percepção pelos sentidos do agente, resultando em uma compreensão equivocada

⁷⁷ HALLEVY, Gabriel. *When Robot Kills. Artificial Intelligence under Criminal Law*. Northeastern: Northeastern University Press, 2013.

⁷⁸ HALLEVY, Gabriel. *When Robot Kills. Artificial Intelligence under Criminal Law*. Northeastern: Northeastern University Press, 2013.

⁷⁹ HALLEVY, Gabriel. *When Robot Kills. Artificial Intelligence under Criminal Law*. Northeastern: Northeastern University Press, 2013.

da realidade pelo nosso cérebro. O entendimento aqui é que alguém só pode ser responsabilizado criminalmente por fatos que conhecia ou acreditava conhecer, inclusive no que diz respeito às máquinas inteligentes.⁸⁰

Analisando as possibilidades de punição da inteligência artificial autônoma, Ryan Abbott e Alex Sarch sintetizam que o caminho mais óbvio seria reconhecer que máquinas são objetos inanimados e, como tal, carecem de estados mentais, de capacidade deliberativa e, portanto, de culpabilidade.⁸¹ Nesse sentido, qualquer sanção criminal à inteligência algorítmica seria injusta uma vez que não se verificam os pressupostos básicos necessários à imputação penal. Essa seria, contudo, uma visão estreita sobre o fenômeno pautada, apenas, pelos requisitos da doutrina tradicional do direito penal. Em uma visão mais ampla e abrangente, ainda que essencialmente especulativa, Abbott e Sarch reiteram o posicionamento de Hallevy, indicando a concepção de uma estrutura jurídica diretamente relacionada à responsabilização da inteligência artificial, atribuindo-lhes legalmente a ideia de culpabilidade criminal, com a necessária reinterpretação dos institutos vigentes no direito penal.⁸²

Restam, contudo, infundáveis questões ainda sem resposta acerca da responsabilidade penal da inteligência artificial como, por exemplo, quais seriam as eventuais penas impostas às máquinas. Afinal, mesmo uma punição “pessoal”:

[...] não atinge o agente da mesma forma que um ser humano - ao menos enquanto os agentes inteligentes não forem dotados de um desejo de (sobre)viver humano. Por esta razão, no estágio atual de desenvolvimento, a punição penal de agentes inteligentes fracassa porque eles não podem sentir a pena como tal.⁸³

Não se desconhece, entretanto, o imenso desafio que se descortina na discussão e definição de novos parâmetros penais relacionados à inteligência artificial, de modo a tutelar direitos fundamentais sem, contudo, inviabilizar o próprio desenvolvimento social.

⁸⁰ HALLEVY, Gabriel. *When Robot Kills. Artificial Intelligence under Criminal Law*. Northeastern: Northeastern University Press, 2013.

⁸¹ ABBOTT, Ryan; SARCH, Alex. Punishing Artificial Intelligence Legal Fiction or Science Fiction. *University of California*, v. 53, 10 out. 2019. Disponível em: https://lawreview.law.ucdavis.edu/issues/53/1/articles/files/53-1_Abbott_Sarch.pdf. Acesso em: 05 ago. 2021.

⁸² HALLEVY, Gabriel. *When Robot Kills. Artificial Intelligence under Criminal Law*. Northeastern: Northeastern University Press, 2013.

⁸³ GLESS, Sabine; WEIGEND, Thomas. Agentes inteligentes e o direito penal. In: ESTELLITA, H.; LEITE, A. (Ed.). *Veículos autônomos e direito penal*. São Paulo: Marcial Pons, 2019, p. 53.

Conclusão

Preliminarmente, conclui-se que o Direito Penal moderno foi idealizado e estruturado, em grande medida, para os seres humanos. Seus institutos atuais e a sua evolução histórica denotam a criação de teorias e conceitos visceralmente relacionados às pessoas naturais como conduta, dolo, culpa e culpabilidade, todos constituídos de elementos psicológicos como finalidade, consciência e vontade que seriam próprios dos seres humanos.

Ainda que prevaleça grande dogmatismo no Direito Penal ao longo dos séculos, não se desconhece que, diante dos novos dilemas e riscos, como aqueles relacionados à modernidade a partir do século XX, um movimento de responsabilização penal dos entes coletivos tenha ganhado força, impondo a superação de axiomas clássicos da ciência penal.

A partir do início do século XXI, a tecnologia da inteligência artificial ganhou impulso diante da existência de diversas condições favoráveis como a multiplicidade e grandiosidade dos bancos de dados (*big data*), o desenvolvimento dos processadores computacionais e novas demandas industriais e sociais. Nesse contexto, surgem algoritmos inteligentes baseados em tecnologias de aprendizado de máquinas, aprendizado profundo e redes neurais, viabilizando a criação de sistemas inteligentes que decidem, em grande medida, de modo imprevisível e, portanto, com total autonomia em relação aos seres humanos.

A responsabilidade penal relacionada à inteligência artificial sem autonomia não oferece maiores dificuldades de verificação uma vez que decorre, inexoravelmente, da conduta humana. Nestes termos, os sistemas inteligentes devem ser entendidos como meros instrumentos adotados pelos seres humanos para a realização dos seus intentos criminosos. Aqui, sem maiores percalços, institutos e dogmas sedimentados no Direito Penal brasileiro são utilizados sem qualquer questionamento. Surgem, entretanto, evidências de uma carência punitiva quando a pessoa jurídica é o agente por trás da inteligência artificial, justificando a implementação legal da sua responsabilização penal.

Por sua vez, quando se trata da real possibilidade de autonomia das máquinas inteligentes, surge um vazio de imputação criminal consistente na insuficiência de respostas do sistema finalista da conduta adotado no Brasil e do próprio dogmatismo do direito penal. Tratando-se de uma máquina programada para aprender e alterar os seus próprios algoritmos visando a produção de respostas imprevisíveis e complexas que atendam à sua finalidade, tem-se que não há como estabelecer uma relação de causalidade entre eventuais danos causados e a atuação inicial do ser humano. Insistir na existência de uma finalidade criminosa ou na falta de um cuidado pelo ser humano de modo a evitar a atuação autônoma de máquinas inteligentes parece contrariar não só a lógica, mas a própria essência da criação

desses mecanismos. Além da incoerência de uma tentativa de responsabilização do ser humano por trás da máquina a qualquer custo, cria-se intransponível barreira à inovação tecnológica e, conseqüentemente, aos avanços da humanidade na busca de uma melhor qualidade de vida.

Constatada a impossibilidade de responsabilização penal do ser humano por trás da inteligência artificial autônoma surge o problema de quem e em que condições haveria eventual imputação criminal em decorrência dos danos causados por essas máquinas. Considerando o Direito Penal como a ferramenta mais eficiente na proteção de bens jurídicos relevantes ao ser humano – em razão do seu caráter sancionatório retributivo e preventivo – surgem, basicamente, dois caminhos possíveis: a criação de um regime jurídico específico para ações desencadeadas pela inteligência artificial ou a eventual compatibilização de institutos atuais a essa nova realidade.

Nesse sentido, constatou-se que a adoção de regime jurídico penal similar àquele empregado aos animais não se conforma com as características inerentes à inteligência artificial que, por sua vez, aproxima-se mais de particularidades humanas como a racionalidade baseada em um raciocínio lógico.

A par da teoria finalista da conduta que não atende à responsabilização da inteligência artificial autônoma, outras teorias foram analisadas como a da ação significativa, de Tomás Salvador Vives Antón que, sem desconstruir a dogmática penal em todos os seus fundamentos, busca definir o crime e seus elementos através de conceitos objetivos como ação e norma. Aqui, o modelo de conduta penal relevante não se baseia em aspectos subjetivos e impenetráveis da mente humana. Além de mais pragmática e juridicamente segura, a teoria da ação significativa possibilita a idealização de uma responsabilidade penal resultante do comportamento das máquinas inteligentes, consideradas necessárias adequações constitucionais, legais e interpretativas relacionadas à natureza específica da inteligência artificial.

Escrutinou-se o pensamento de Gabriel Hallevy que, em visão ainda mais arrojada, entende que a inteligência artificial autônoma pode ser responsabilizada criminalmente através de uma adequada interpretação dos diversos institutos jurídico-penais que compõem a maioria dos sistemas penais da atualidade sempre, contudo, considerando a natureza e características das máquinas inteligentes.

Conclui-se, enfim, pela impossibilidade de responsabilidade penal do ser humano por trás da inteligência artificial autônoma de acordo com o regime jurídico-penal adotado no Brasil na atualidade. Por outro lado, considerando a necessária tutela do Direito Penal a graves lesões eventualmente causadas por máquinas inteligentes, sugere-se a adoção de alterações legislativas – inicialmente, na seara constitucional –, de modo a delinear o melhor modelo de responsabilização penal da inteligência artificial consideradas as opções ora expostas e o tratamento, ainda incipiente e seminal sobre o tema ao redor do mundo.

Who is to blame? A criminal analysis of human interactions with artificial intelligence

Abstract: Artificial intelligence, an innovative technology that seeks to imitate human abilities, has variable and fluid concepts and characteristics, ranging from current intelligence without autonomy in relation to human beings, to intelligent agents with complete independence from human behavior. Faced with the growing decision-making power of machines due to systems endowed with machine learning and neural networks, the article intends to investigate the possible criminal liability – of humans and machines – considering the legal-criminal structure in force in Brazil in the context of the finalist system of conduct. Hypothetically, it is argued that serious actions caused by autonomous artificial intelligence cannot be imputed to humans; on the other hand, intelligent machines, in the context of this same penal dogmatics, cannot be perpetrators of crime. Faced with this void or (insufficiency) of criminal liability, the possibilities of criminal liability related to artificial intelligence are analyzed, always considering the fundamental human rights involved.

Keywords: Artificial Intelligence. Technological Innovation. Fundamental rights. Criminal Liability.

Contents: **1** Introduction – **2** Fundamentals of criminal liability in Brazil – **3** Human mind and artificial intelligence – **4** Criminal liability and artificial – **5** Conclusion – References

Referências

ABBOTT, Ryan; SARCH, Alex. Punishing Artificial Intelligence Legal Fiction or Science Fiction. *University of California*, v. 53, 10 out. 2019. Disponível em: https://lawreview.law.ucdavis.edu/issues/53/1/articles/files/53-1_Abbott_Sarch.pdf. Acesso em: 05 ago 2021.

ALONSO, Raúl. ¿Un robot puede ser responsable penal del delito que comete? *Mutualidad Abogacia*, 21 abr. 2021. Disponível em: <https://tudefinestufuturo.mutualidadabogacia.com/innovacion/un-robot-puede-ser-responsable-penal-del-delito-que-comete/>. Acesso em: 02 ago. 2021.

ANTÔN, Tomás Salvador Vivés. *Fundamentos del sistema penal*. Valencia: Tirant lo Blanch, 1996.

BECCARIA, Cesare. *Dos delitos e das penas*. São Paulo: Revista dos Tribunais, 1999.

BICEGO, Bruno Ett; HIGASI, Plínio. A queda dos aviões Boeing 737 MAX: um alerta para o Direito Digital no Brasil. *Digital Rights*, 9 abr. 2019. Disponível em: <https://digitalrights.cc/blog/2019/04/09/a-queda-dos-avioes-boeing-737-max-um-alerta-para-o-direito-digital-no-brasil/>. Acesso em: 20 dez. 2020.

BOSTROM, Nick. *Superinteligência: caminhos, perigos, estratégias*. Rio de Janeiro: DarkSide Books, 2018.

BOTTINI, Pierpaolo Cruz. Indagações sobre a responsabilidade penal da empresa pelo produto defeituoso. *Boletim IBCCRIM*, v. 16, n. 194, jan. 2009.

BRASIL. *Decreto-Lei nº 2.848, de 7 de dezembro de 1940*. Código Penal. Brasília. 1940. Disponível em: http://www.planalto.gov.br/ccivil_03/decreto-lei/del2848compilado.htm. Acesso em: 15 jan. 2021.

BRASIL. *Constituição da República Federativa do Brasil*. Brasília. 1988. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 01 jan. 2021.

BRASIL. *Lei nº 8.072, de 11 de setembro de 1990*. Dispõe sobre a proteção do consumidor e dá outras providências. 1990. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l8078.htm. Acesso em: 10 jan. 2021.

BRASIL. Supremo Tribunal Federal. *Recurso Extraordinário nº 548.181*. Primeira Turma. Relator: Min. Rosa Weber. Brasília, 6 ago. 2013. Disponível em: <http://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=7087018>. Acesso em: 19 dez. 2020.

- BRYNJOLFSSON, E.; MCAFEE, A. *The Second Machine Age: work, progress, and prosperity in a time of brilliant technologies*. NY: W.W. Norton & Company, 2016.
- BUSATO, Paulo César. *Direito penal*. São Paulo: Atlas, 2015.
- DENNETT, Daniel C. Consciousness in Human and Robot Minds. In: IAS Symposium on Cognition, Computation and Consciousness. Kyoto: 1 set. 1994. Disponível em: <https://ase.tufts.edu/cogstud/dennett/papers/concrobt.htm>. Acesso em: 10 jan. 2021.
- ELIAS, Paulo Sá. *Algoritmos, Inteligência Artificial e o Direito*. E-Gov: portal de e-governo, inclusão digital e sociedade do conhecimento, Florianópolis, 2018. Disponível em: <http://www.egov.ufsc.br/portal/conteudo/algoritmos-intelig%C3%Aancia-artificial-e-o-direito>. Acesso em: 16 nov. 2020.
- ESTELLITA, Heloisa; LEITE, Alaor. *Veículos autônomos e direito penal*. São Paulo: Marcial Pons, 2019.
- FRAZÃO, Ana; MULHOLLAND, Caitlin (Coord.). *Inteligência artificial e direito – Ética, regulação e responsabilidade*. São Paulo: Revista dos Tribunais, 2019.
- GLESS, Sabine; WEIGEND, Thomas. Agentes inteligentes e o direito penal. In: ESTELLITA, H.; LEITE, A. (Ed.). *Veículos autônomos e direito penal*. São Paulo: Marcial Pons, 2019.
- HALLEVY, Gabriel. *When Robot Kills*. Artificial Intelligence under Criminal Law. Northeastern: Northeastern University Press, 2013.
- HALLEVY, Gabriel. *Liability for Crimes Involving Artificial Intelligence Systems*. Cham: Springer International Publishing, 2015.
- HANSON, F. Allan. Beyond the skin bag: on the moral responsibility of extended agencies. *Ethics and Information Technology*, v. 11, p. 91-99, 2009. Disponível em: https://kuscholarworks.ku.edu/bitstream/handle/1808/13802/Skin_Bag_Hanson.pdf?sequence=1&isAllowed=y. Acesso em: 06 ago. 2021.
- HASSEMER, Winfried. Neurociências e culpabilidade em direito penal. In: BUSATO, P. C. (Ed.). *Neurociência e direito penal*. São Paulo: Atlas, 2014.
- HUNGRIA, Néelson. *Comentários ao Código Penal*. Rio de Janeiro: Forense, 1949, v. 1.
- JUSTO, Gabriel. Após acidentes e aterramento, 737 MAX volta a voar. O que aconteceu? *Exame*, 7 dez. 2020. Disponível em: <https://exame.com/mundo/apos-acidentes-e-aterramento-737-max-volta-a-voar-o-que-aconteceu/>. Acesso em: 25 jul. 2021.
- KAPLAN, Andreas; HAENLEIN, Michael. *Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence*. *Business Horizons*, p. 15-25, 2019. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0007681318301393?via%3Dihub>. Acesso em: 07 jan. 2021.
- KARNOW, Curtis E. A. Liability for distributed artificial intelligences. *Berkeley Technology Law Journal*, v. 11, n. 1, 1996. Disponível em: <https://lawcat.berkeley.edu/record/1115611>. Acesso em: 04 ago. 2021.
- KAUFMAN, Dora. *A inteligência artificial irá suplantará a inteligência humana?* Barueri: Estação das Letras e Cores, 2019.
- LISZT, Franz Von. *Tratado de direito penal alemão*. Brasília: Senado Federal, 2006. V. 1.
- MATTHIAS, Andreas. The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics and Information Technology*, v. 6, p. 175-183, 2004. Disponível em: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.456.8299>. Acesso em: 06 ago. 2021.
- NILSSON, Nils J. *The quest for artificial intelligence: a history of ideas and achievements*. Cambridge: Cambridge University Press, 2010.
- NUCCI, Guilherme de Souza. *Curso de direito penal: parte geral*. Rio de Janeiro: Forense, 2019.

- PAGALLO, Ugo. *The Laws of Robots: Crimes, Contracts, and Torts*. New York: Springer, 2015.
- PENTLAND, Alex. *Social Physics: How Social Networks Can Make Us Smarter*. New York: The Penguin Press, 2014.
- RAMOS, João Gualberto Garcez. *A inconstitucionalidade do direito penal do terror*. Curitiba: Juruá, 1991.
- ROXIN, Claus. *Funcionalismo e imputação objetiva no direito penal*. Rio de Janeiro: Renovar, 2002.
- RUSSELL, Stuart J.; NORVIG, Peter. *Inteligência artificial*. Rio de Janeiro: Elsevier, 2013.
- SCHÖNBERGER, Viktor Mayer; RAMGE, Thomas. *Reinventing Capitalism in the Age of Big Data*. New York: Basic Books, 2018.
- SHECAIRA, Sérgio Salomão. *Responsabilidade penal da pessoa jurídica*. Rio de Janeiro: Elsevier, 2011.
- SHEPARDSON, David. Tesla driver in fatal 'Autopilot' crash got numerous warnings: U.S. government. *Reuters*, 19 jun. 2017. Disponível em: <https://www.reuters.com/article/us-tesla-crash-idUSKBN19A2XC>. Acesso em: 31 jul. 2021.
- SUMWALT III, Robert L. et al. Accident Report NTSB/HAR-17/02 PB2017-102600. *National Transportation Safety Board*. Washington, 12 set 2017. Disponível em: <https://www.nts.gov/investigations/AccidentReports/Reports/HAR1702.pdf>. Acesso em: 25 jul. 2021.
- TAVARES, Juarez. *Fundamentos da teoria do delito*. Florianópolis: Tirant lo Blanch, 2018.
- TOLEDO, Maria Izabel Vasco de. A tutela jurídica dos animais no Brasil e no direito comparado. *Revista Brasileira de Direito Animal*, Salvador, v. 7, n. 11, p. 197-223, dez. 2012. Universidade Federal da Bahia. Disponível em: <https://portalseer.ufba.br/index.php/RBDA/article/download/8426/6187>. Acesso em: 24 dez. 2019
- UNITED STATES. Department of Justice. Boeing Charged with 737 Max Fraud Conspiracy and Agrees to Pay over \$2.5 Billion. *Justice News*, 7 jan. 2021. Disponível em: <https://www.justice.gov/opa/pr/boeing-charged-737-max-fraud-conspiracy-and-agrees-pay-over-25-billion>. Acesso em: 31 jul. 2021.
- VENTURA, Felipe. *Motorista morre em acidente ao dirigir carro da Tesla em modo semiautônomo*. Disponível em: <https://gizmodo.uol.com.br/acidente-morte-tesla-autopilot/>. Acesso em: 18 dez. 2020.
- WELZEL, Hans. *O novo sistema jurídico-penal: uma introdução à doutrina da ação finalista*. São Paulo: Revista dos Tribunais, 2015.
- ZAFFARONI, Eugenio Raul; PIERANGELI, José Henrique. *Manual de direito penal brasileiro*. São Paulo: Thomson Reuters, 2019.

Informação bibliográfica deste texto, conforme a NBR 6023:2018 da Associação Brasileira de Normas Técnicas (ABNT):

MELLO, Rogério Luís Marques de; LAZARI, Rafael José Nadim de. Quem é o culpado? Uma análise criminal das interações humanas com a inteligência artificial. *Direitos Fundamentais & Justiça*, Belo Horizonte, ano 16, n. 46, p. 123-153, jan./jun. 2022.

Recebido em: 20.01.2021

Pareceres: 19.06.2021; 19.07.2021; 21.07.2021; 06.09.2021

Aprovado em: 06.09.2021