

UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE DIREITO PROF. JACY DE ASSIS

Michael Menezes Martins

RESPONSABILIDADE PENAL DE INTELIGÊNCIAS ARTIFICIAIS

Uberlândia/MG
2022

Michael Menezes Martins

RESPONSABILIDADE PENAL DE INTELIGÊNCIAS ARTIFICIAIS

Trabalho de Conclusão apresentado ao Curso de Bacharelado em Direito da Universidade Federal de Uberlândia, como exigência parcial para a obtenção do Título de Bacharel em Direito.

Orientados: Prof. Ms. Karlos Alves Barbosa

Uberlândia/MG
2022

Dados Internacionais de Catalogação na Publicação (CIP)

RESPONSABILIDADE PENAL DE INTELIGÊNCIAS ARTIFICIAIS

Trabalho de Conclusão de Curso aprovado para a obtenção do Título de Bacharel em Direito da Universidade Federal de Uberlândia, pela banca examinadora formada por:

Uberlândia, 26 de julho de 2022.

Prof. Ms. Karlos Alves Barbosa

Profa. Dra. Simone Silva Prudêncio

AGRADECIMENTOS

Primeiramente à Deus, pela dádiva da vida e do aprendizado, e por permitir que se cumpra mais esta etapa da minha vida.

Aos meus familiares, pelo apoio durante todo o decorrer deste curso e principalmente durante esta fase de conclusão.

Ao professor orientador Prof. Ms. Karlos Alves Barbosa, pela disposição e paciência na orientação e apoio a este trabalho, como em outros ao longo desta jornada.

Aos demais professores, pelo empenho em ensinar, na paciência e compreensão no decorrer deste curso. Em especial à Profa. Dra. Simone Silva Prudêncio, integrante da banca examinadora, que por sinal foi importantes partícipe na minha formação acadêmica.

Aos colegas de turma, pelos altos e baixos e mesmo assim seguir firme nesta trajetória, por nunca desistir e pelos laços que se formaram, principalmente neste final de curso.

RESUMO

A evolução tecnológica dos últimos anos é notável. Com estudos de robótica e nanotecnologia, cada vez mais limites, antes ditos inalcançáveis, estão sendo ultrapassados. Hoje o mundo não é mais o mesmo que a dez anos atrás, tudo vem mudando muito depressa e a ciência pede passagem. Um dos campos de maior destaque é o das Inteligências Artificiais. Com um discurso de quebrar barreiras tecnológicas e revolucionar as relações humanas, as chamadas IA's são ainda um tema nebuloso para o homem. Grande avanço ocorreu desde os primeiros rascunhos da história da informática até o que temos hoje, com robôs inteligentes e sistemas algorítmicos que agem quase que autonomamente. A novidade e comodidade são vícios humanos, e ter um ente inteligente pronto para servir aos caprichos humanos é senão "um sonho". Por estas razões as IA's começaram a entrar nas vidas das pessoas, em suas casas, em seus trabalhos, no seu carro, nas suas redes sociais, praticamente existe algum tipo de IA em todo programa computacional moderno, é claro que mais ou menos complexos. Dito isto, verifica-se que as IA's estão gradativamente fazendo parte do cotidiano humano, e por assim dizer construindo novas relações jurídicas com estes. Dentre todo o amplo espectro do direito, uma questão merece destaque quando se fala de IA's: caso uma IA cometa um crime quem se deve responsabilizar? É uma pergunta difícil de se responder, mas este trabalho se permite elaborar possíveis soluções para a problemática, analisando alguns modelos de responsabilização penal. Pretende-se também revelar estes aspectos enigmáticos das IA's, orientando quanto a temas relevantes como o aprendizado de máquina (*machine learnig*) e o aprendizado profundo (*deep learnig*), e assim relacionar com a teoria do crime e da culpabilidade vigentes em nosso código penal. Este trabalho visa iniciar um debate muito importante para a atualidade, que já é uma realidade social, e que em breve se tornará um problema inevitável.

Palavras-chave: Inteligências Artificiais. IA. Modelos de Responsabilização Penal. Machine learnig. Deep learnig. Teoria do crime. Culpabilidade.

ABSTRACT

The technological evolution of recent years is remarkable. With robotics and nanotechnology studies, more and more limits, previously said to be unreachable, are being exceeded. Today the world is no longer the same as it was ten years ago, everything is changing very quickly and science is asking for passage. One of the most prominent fields is Artificial Intelligence. With a speech of breaking technological barriers and revolutionizing human relationships, the so-called AI's are still a nebulous topic for man. Great advancement has taken place from the earliest drafts of computer history to what we have today, with intelligent robots and algorithmic systems that act almost autonomously. Novelty and convenience are human vices, and having an intelligent being ready to serve human whims is nothing but "a dream". For these reasons AI's started to enter people's lives, in their homes, in their jobs, in their car, in their social networks, there is practically some kind of AI in every modern computer program, of course more or less complex. That said, it appears that AI's are gradually becoming part of human daily life, and so to speak, building new legal relationships with them. Among the wide spectrum of law, one issue deserves to be highlighted when talking about AI's: if an AI commits a crime, who should be held responsible? It is a difficult question to answer, but this work allows us to elaborate possible solutions to the problem, analyzing some models of criminal liability. It is also intended to reveal these enigmatic aspects of AI's, providing guidance on relevant topics such as machine learning and deep learning, and thus relate to the theory of crime and culpability in force in our penal code. . This work aims to start a very important debate for the present, which is already a social reality, and which will soon become an inevitable problem.

Keywords: Artificial Intelligences. AI. Models of Criminal Responsibility. Machine learning. Deep learning. Theory of crime. Culpability.

SUMÁRIO

1.	<i>INTRODUÇÃO</i>	7
2.	<i>BREVES APONTAMENTOS SOBRE INTELIGÊNCIAS ARTIFICIAIS</i>	9
3.	<i>INTELIGÊNCIAS ARTIFICIAIS: O QUE SÃO?</i>	15
4.	<i>DIREITO PENAL, TEORIA DO CRIME E RESPONSABILIDADE PENAL NO ORDENAMENTO JURÍDICO BRASILEIRO</i>	22
5.	<i>MODELOS DE RESPONSABILIZAÇÃO PENAL DE CONDUTAS DA INTELIGÊNCIA ARTIFICIAL: UM ESTUDO DA CULPA SOB A ÓTICA DA TEORIA DO CRIME</i>	37
6.	<i>CONCLUSÃO</i>	49
	<i>REFERÊNCIAS</i>	54

1. INTRODUÇÃO

Após anos de desenvolvimento, a espécie *Homo Sapiens* desenvolveu diversas ferramentas que possibilitaram a instalação e controle sobre as adversidades do meio natural. As primeiras cidades, a divisão de tarefas e as necessidades do cotidiano instigaram a aptidão humana em elaborar instrumentos para aprimorar a eficiência de trabalhos necessários para o crescimento da produtividade ao que, atualmente, culminou na criação de Inteligências Artificiais (IA's).

Como ponto de partida, vale ressaltar que esse progresso tecnológico se deu de forma bastante abrupta no convívio humano em todos os âmbitos existentes. As máquinas de Alan Turing, um dos fundadores da ciência computacional, as quais representam o símbolo máximo de desenvolvimento robótico da década de 1930 foi responsável por decifrar códigos alemães durante o desenrolar da Segunda Guerra mundial (1939-1945). Após apenas 70 anos desse período, os computadores e os smartphones, com proporções comparativamente bem menores, realizam processos mais complexos- adequados as “novas necessidades do homem moderno”.

Na década de 1950, o escritor e bioquímico Isaac Asimov, criador de “Eu, robô” e notório expoente da ficção científica, idealizou três princípios para nortear a robótica. A primeira, baseia-se na proibição de condutas ativas e omissivas para ferir um ser humano ou permitir que esse sofra algum dano. A segunda, vincula o robô às ordens estabelecidas pelas pessoas, exceto nos casos em que essas desobedecem a primeira lei. Por fim, a terceira assegura a possibilidade de autopreservação da existência artificial, caso essa proteção e legítima defesa não infrinjam as leis antecedentes. Essas três leis embasam a “ética” da programação dos robôs da atualidade, porém, devido às inúmeras possibilidades dessa nova relação homem-máquina, a ocorrência de situações que ultrapassam essas normas para a Inteligência Artificial (IA) já começam a permear o cotidiano da vida humana, o que torna evidente e necessário a preocupação do Direito para estabelecer normas, prerrogativas e punições adequadas para frear possíveis usurpações de direitos fundamentais, como o direito à vida.

Parece-nos evidente que o tema é de suma importância para o atual palco de discussões jurídicas e por esta razão, este trabalho possui o intuito de analisar e

discorrer sobre o papel do Direito sobre a responsabilidade penal em casos em que há o envolvimento de Inteligências Artificiais resultando em danos físicos e materiais. Assim, será exposto as formas como elas atuam, como também os pontos positivos e negativos da Legislação brasileira atual para esse enfoque, com o objetivo de ampliar a discussão e fornecer possíveis embasamentos para resoluções de leis para essa nova e iminente realidade do desenvolvimento humano.

Este primeiro capítulo realizou a introdução do tema, já o segundo capítulo visa fazer breves apontamentos sobre as inteligências artificiais, demonstrando as incertezas e certezas do tema, faz-se um panorama geral sobre o assunto demonstrando a ampla utilização das IA's na modernidade, no Brasil e no mundo.

No terceiro capítulo busca-se definir o que são Inteligências Artificiais, visando diferenciar os diversos sistemas automatizados e mostrar os avanços tecnológicos na área da robótica e assim destacar a influência destes mecanismos sobre o cotidiano humano.

No quarto capítulo pretende-se realizar um estudo sobre a Responsabilidade Penal no ordenamento jurídico brasileiro, em especial sobre a teoria do crime, examinando os aspectos do sujeito do delito, da conduta e da culpabilidade.

Enfim, no quinto capítulo deseja-se apresentar diferentes modelos de responsabilização penal de condutas de uma inteligência artificial, analisando a culpabilidade destes sistemas sob a ótica da teoria do crime.

2. BREVES APONTAMENTOS SOBRE INTELIGÊNCIAS ARTIFICIAIS

O tema “Inteligência Artificial” ainda gera muita dúvida, incerteza e misticismo, assunto brilhantemente abordado pelo autor Castro Júnior (2009, p. 130):

Inteligência artificial se apresenta para o público em geral como o devir, o que ainda não se concretizou, que não perdeu sua aura mística, o que ainda não se conhece o funcionamento. Isso, porque, tudo quanto um dia se atribuiu ser o campo da inteligência artificial, na medida em que foi por ela realizado, passou a ser descartado como tal. Parece que o Homem precisa acreditar em algo misterioso para se sustentar no mundo concreto. Antes deuses ou Deus, agora, no mundo tecnológico, na criação humana no porvir, inexplicada e desejadamente inexplicável, pois o que já foi compreendido pertence ao reino dos homens e não guarda o mistério transcendental que levará à eternidade.

O autor revela o caráter “místico” dado ao tema, muito abordado em filmes e livros de ficção científica. Por exemplo, na série de filmes “O Exterminador do Futuro” de 1984, a Skynet é uma inteligência artificial altamente avançada criada no fim do século XX. Em dado momento a IA se torna autoconsciente e vê a humanidade como uma ameaça à sua existência, decidindo exterminá-la; revelam o temor no desconhecido campo das IA's.

Contudo, tal receio se justifica na rapidez com que a tecnologia vem avançando, com o desenvolvimento de sistemas cada vez mais inteligentes e autônomos, capazes de tomadas de decisão sem interferência humana.

O Asimo, por exemplo, era um robô desenvolvido pela Honda, capaz de andar, reconhecer pessoas, realizar tarefas complexas como abrir garrafas, subir e descer escadas etc., tudo de forma autônoma. Teve a introdução no mercado em 2000, onde a ideia era a criação de um robô doméstico para auxiliar nas tarefas diárias; recebeu a sua última atualização em 2011 e teve sua fabricação descontinuada pela fábrica em 2020 (HONDA, 2020).

Outro exemplo emblemático de entidade dotada de IA em um corpo robótico é o da Robô Sophia. Em 25 de outubro de 2017, a Inteligência Artificial criada por David Hanson foi o primeiro robô a adquirir cidadania, tornando-se cidadã da Arábia Saudita. Além disso, Sophia também é capaz de simular comportamentos humanos, tais como senso de humor e sentimentos. Sua capacidade discursiva também impressiona. Sophia proferiu um breve discurso na ONU sobre assuntos humanitários, como a falta

de acesso à Internet e a falta de energia elétrica em grande parte do mundo (DIVINO, 2020, p. 155).

Em conjunto com Asimo e Sophia, algumas IA's possuem um corpo diferente. É o caso da empresa de tecnologia Uber, que desenvolveu e vem aprimorando um sistema de IA capaz de conduzir, de forma automática, carros com passageiros. Ela usa um tipo de rede neural chamada rede generativa antagônica (GANs), aliado a dados de sensores como LIDAR e câmeras a bordo dos veículos, para calcular trajetórias que levam em consideração os pontos de restrição de movimento (RIGUES, 2020).

Tecnologia muito semelhante é utilizada pela montadora automobilística mais valiosa do mundo, Tesla. Neste caso, as IA's utilizam tecnologias de ciência de dados no aprendizado de máquina, onde os veículos auto dirigíveis compartilham informações entre eles. Ou seja, o aprendizado de máquina de uma unidade específica (de um carro da Tesla), é possível de ser compartilhado em rede com todas as outras unidades automobilísticas, se desejável (ZUCHER, 2020).

Recentemente, Blake Lemoine, um engenheiro do Google, disse ao jornal Washington Post que, após centenas de interações com um sistema de inteligência artificial inédito chamado LaMDA, ele acreditava que o programa havia alcançado um nível de consciência. Mas o Google foi rápido em encerrar as alegações de que um de seus programas avançou tanto que se tornou senciente (ou seja, passou a poder sentir sensações de forma consciente), afastando inclusive o engenheiro de suas funções (METZ, 2022).

Mas “nem tudo são rosas” quando falamos de IA's. Em 23 de março de 2016, a Microsoft lançou um chatbot (programa de computador que simula uma conversa com um ser humano, como a Siri e a Cortana) que, após 24 horas, já foi retirado do ar. Isso porque, ao interagir com os usuários, a tecnologia passou a desferir falas de cunho racista e sexista, e até mesmo fez uma defesa explícita do ditador Adolf Hitler (SOPRANA, 2016).

Mais preocupante ainda, nos recentes conflitos entre Ucrânia e Rússia, foram utilizados drones como o KUB-BLA com sistema de decisão automatizado por inteligência artificial (IA). Fabricado pela ZALA Aero, subsidiária da empresa russa de armamentos Kalashnikov, o drone tem 1,2 metros, percorre até 130 quilômetros em

30 minutos, e colide com o alvo detonando um explosivo de três quilos. Não foi confirmado se tal dispositivo teria sido usado em modo autônomo (perícia difícil de ser conclusiva nesses casos), contudo alerta sobre a possibilidade deste fato (KAUFMAN, 2022).

Esses indícios aumentam as preocupações sobre os riscos de armas letais automatizadas. Zachary Kallenborn, membro do "The National Consortium for the Study of Terrorism and Responses to Terrorism" - START, parte do Centro de Excelência de Segurança Interna da Universidade de Maryland, e proprietário do "Global Terrorism Database", o mais abrangente banco de dados sobre eventos terroristas no mundo, adverte que (KAUFMAN, 2022):

A noção de um robô assassino - onde você tem inteligência artificial fundida com armas - é uma tecnologia que já existe e está sendo usada. Um dos desafios com armas autônomas pode ser a dificuldade de determinar quando a autonomia total é usada em um contexto letal.

Nesta seara, a perspectiva de "armas autônomas" - algoritmos de IA que decidem onde e quando uma arma deve disparar, supostamente com mais precisão e eliminando os erros humanos -, tem sido considerada a maior ameaça militar contemporânea. Usando a técnica de redes neurais profundas (deep learning), os algoritmos são treinados em milhares de dados de batalhas, em seguida ajustados para um conflito específico com o objetivo de lançar, "cirurgicamente", bombas sem intervenção humana. Países como EUA, China e Rússia discutem os termos de um tratado para limitar as armas autônomas letais e, simultaneamente, investem fortemente em desenvolvê-las (KAUFMAN, 2022).

O risco da criação de uma arma totalmente autônoma também é preocupante, parece-nos que o cenário apocalíptico abordado em "O Exterminador do Futuro" não está tão distante e uma espécie de "Skynet" não seria tão impossível assim. Com efeito, devemos questionar qual seria o nível de independência que deveria ter uma inteligência artificial no setor bélico.

Desde a Conferência Internacional de Inteligência Artificial de 2015, ocorrida em Buenos Aires (Argentina), até a edição de 2017, dessa vez em Melbourne (Austrália), mais de mil cientistas e especialistas assinaram uma carta aberta contra o desenvolvimento de robôs militares autônomos, dentre outros o físico Stephen

Hawking, o empreendedor Elon Musk, e o cofundador da Apple, Steve Wosniak; o objetivo era persuadir a ONU para que vetasse o uso de armas autônomas (KAUFMAN, 2022).

Elon Musk, fundador do PayPal, CEO da fabricante de veículos Tesla Motors e da empresa aeroespacial SpaceX, famoso por “excentricidades visionárias” e falas polêmicas, é considerado um dos principais protagonistas da inovação, mas no que se refere à inteligência artificial, acredita que é preciso pisar no freio (PALAZUELOS, 2017).

Com falas do tipo: “Até que as pessoas vejam robôs matando gente na rua, não se entenderão os perigos da inteligência artificial”, “Dentro de 20 anos, não é que não haverá pessoas ao volante. É que simplesmente não haverá volante” e também em sugerir que a Inteligência Artificial seria “uma ameaça fundamental para a existência da civilização humana”, Musk pediu a autoridades americanas numa reunião ocorrida em Rhode Island (EUA) em 2017 que, ao contrário do que acontece em outras indústrias, se regulamente de forma proativa ao invés de esperar que surjam problemas (PALAZUELOS, 2017).

Introduzindo ao tema normativo, em 1950, o escritor e bioquímico Isaac Asimov (1976), criador de “Eu, robô” e notório expoente da ficção científica, idealizou três princípios para nortear a robótica. A primeira, baseia-se na proibição de condutas ativas e omissivas para ferir um ser humano ou permitir que esse sofra algum dano. A segunda, vincula o robô às ordens estabelecidas pelas pessoas, exceto nos casos em que essas desobedecem a primeira lei. Por fim, a terceira assegura a possibilidade de autopreservação da existência artificial, caso essa proteção e legítima defesa não infrinjam as leis antecedentes.

Asimov pretendia regular a relação homem-máquina com essas três diretrizes, hoje conhecidas como “Leis da Robótica”; imaginando as complexas e infinitas possibilidades das interações humanas e suas consequências.

Todavia, conforme Castro Júnior (2009, p. 176):

Se as Leis da Robótica devem ser aplicadas aos robôs – aplicada aqui tem o sentido de imposta, programada, previamente disciplinadora – descritiva de algo inevitável, fatal, porque estes seguiriam os programas e deles não poderiam arredar, duas conclusões podem ser alcançadas: a) ou os robôs não são inteligentes, ou não tem programas de inteligência artificial, que por definição não podem ter seus fins previamente determinados e, portanto, não

agem no sentido jurídico ou b) essas leis são de aplicação impossível, pois, sendo os robôs inteligentes neles não se pode embutir regras de conduta infalíveis, naturais, como são as leis da gravidade ou qualquer outra da física.

Induz que um robô dotado de inteligência artificial (ou uma IA propriamente dita) possa agir em desconformidade com a lei, sejam estas aquelas idealizadas por Asimov ou outras jurídico-legislativo elaboradas e pré-programadas, vez que pensa de forma autônoma, podendo “decidir” se seguirá essas diretrizes ou não; revela uma preocupação e imprevisibilidade na operação destas tecnologias.

Não é de se surpreender que o tema da inteligência artificial e o direito é atual e está na pauta de estudo e desenvolvimento de vários países, assumindo centralidade nos principais espaços públicos de debate.

Em recente matéria publicada pelo Instituto de Pesquisa Econômica Aplicada – IPEA cita outra reportagem realizada pelo MIT Technology Review, veículo de divulgação do Instituto de Tecnologia de Massachusetts, uma das mais prestigiosas universidades dos Estados Unidos, é mencionado a “Lei da IA” (AI Act), projeto atualmente em discussão no Parlamento Europeu, sendo dimensionado como potencial “mãe de todas as leis de IA”. De acordo com a MIT Technology Review, se a União Europeia conseguir elaborar e adotar um marco regulatório que cumpra satisfatoriamente o objetivo de discernir os principais riscos da aplicação de IA e estabelecer limites e checagens adequados, essa legislação deverá servir como modelo para outros países e regiões; principalmente China e Estados Unidos, expoentes neste mercado (LOBO, 2022).

As estimativas são de que a aprovação do marco legal europeu levará, no mínimo, mais um ano, ou seja, em 2023, e que, depois disso, haveria um prazo de pelo menos dois anos para a adequação do mercado e dos atores estatais às novas regras (LOBO, 2022).

No Brasil, está sendo discutido o Projeto de Lei 21/2020, um verdadeiro “marco legal do desenvolvimento e uso da Inteligência Artificial (IA) pelo poder público, por empresas, entidades diversas e pessoas físicas”. O texto, em tramitação na Câmara dos Deputados, estabelece princípios, direitos, deveres e instrumentos de governança para a IA. Entre outros pontos, a proposta estabelece que o uso da IA terá como fundamento o respeito aos direitos humanos e aos valores democráticos, a igualdade, a não discriminação, a pluralidade, a livre iniciativa e a privacidade de dados. Além

disso, a IA terá como princípio a garantia de transparência sobre o seu uso e funcionamento (JÚNIOR; DOEDERLEIN, 2020).

Tais preocupações nacionais e internacionais demonstram atenção no porvir tecnológico, onde as Inteligências Artificiais devem demandar considerável participação nas relações jurídicas.

Não obstante, há ainda um outro problema quando falamos de inteligência artificial associada ao Direito e, em particular, ao Direito Penal: a quem caberia a responsabilidade no caso de delitos praticados por sistemas de inteligência artificial autônomos?

Nesta seara, o catedrático de Direito Penal da Universidade do País Basco (Espanha), o advogado espanhol e doutor em Medicina Carlos Romeo Casabona, ao proferir palestra de abertura do Congresso Internacional de Ciências Criminais, promovido pela Pontifícia Universidade Católica do Rio Grande do Sul - PUCRS em 2018, falou sobre “Inteligência Artificial – Robótica e Responsabilidade Penal” (INTELIGÊNCIA..., 2018).

Casabona reflete que, com o advento da IA, os juristas estão preocupados com o que está protegido pelo direito penal, sendo os sistemas autônomos os mais preocupantes. Observa que sendo o instituto da culpa inerente aos seres humanos e não a máquinas, abre-se uma nova discussão para a doutrina jurídica, embora existam poucos especialistas no tema. Por fim, defende a intervenção pública com medidas preventivas para controlar sistemas de IA que podem oferecer algum tipo de risco (INTELIGÊNCIA..., 2018).

3. INTELIGÊNCIAS ARTIFICIAIS: O QUE SÃO?

Antes de entrarmos no mérito de falar sobre a responsabilidade penal das inteligências artificiais, faz-se imperioso primeiramente abordar sobre o conceito e delimitação do tema, que é deveras complexo, qual seja explicar o que é uma inteligência artificial.

A dificuldade em se definir a IA (Inteligência Artificial) está intimamente ligada na dificuldade em definir o significado do termo inteligência propriamente dito.

Inteligência tem sido conceituada popularmente e ao longo da história de muitas formas diferentes. O Dicionário Aurélio (2004) dá como uma das definições a faculdade de conhecer, de compreender; intelecto: a inteligência distingue o homem do animal.

Há de se perceber que essa definição, ainda que atual, admite que a Inteligência é uma qualidade própria do ser humano, revelando uma visão antropocêntrica do conhecimento humano, bem como na resolução de problemas por parte desse.

Contudo, a modernidade e a tecnologia vêm revelando uma mudança de paradigma, onde a inteligência passa a ser uma característica do ser humano e não mais como condição.

Nessa perspectiva um grande estudioso suíço, Jean Piaget (1967), contribuiu grandemente para o estudo da inteligência e abriu novas perspectivas de implementação da inteligência artificial.

Para Piaget (apud NAKABAYASHI, 2009, p. 17):

(...) o conhecimento não está no sujeito nem no objeto, mas ele se constrói na interação do sujeito com o objeto. Na medida em que o sujeito interage com os objetivos é que ele produz a capacidade de conhecer e produzir o próprio conhecimento.

Assim, Piaget afirma ser o processo lógico para se chegar ao conhecimento mais importante que o conhecimento em si, inclusive os errôneos. Com isso, revolucionou a forma como se via o conhecimento, apresentando novas possibilidades e desmistificando de vez a visão antropocêntrica do conhecimento (SILVA; BEZERRA, 2018, p. 80).

Assim, sendo a inteligência a capacidade de resolver problemas, realizando escolhas eficientes para cada ação; e o termo artificial ligado àquilo que não é natural, mas produzido por uma técnica humana. Unindo estes dois conceitos, se aduz que a inteligência artificial seria uma forma de “simular” o funcionamento da mente humana em uma máquina (SILVA; BEZERRA, 2018, p. 80).

Nesse sentido, Minsky (1968 apud NAKABAYASHI, 2009, p. 19): “Inteligência Artificial (IA) é a ciência de fazer máquinas fazerem coisas que requereriam inteligência, caso fossem feitas pelo homem”.

Coppin (2013, p. 3) descreve a IA como “o estudo dos sistemas que agem de um modo que a um observador qualquer pareceria ser inteligente”.

Stuart Russell (2016 apud LAGE, 2021, p. 27) traz a definição de inteligência artificial (IA) como o estudo dos métodos para fazer computadores se comportarem de forma inteligente. Afirma que, de modo geral, um computador é inteligente na medida em que faz a coisa certa ao invés da coisa errada. Conceitua a “coisa certa” como a ação mais propensa a atingir um objetivo, ou, em termos mais técnicos, a ação que maximiza uma utilidade esperada. Sustenta que a IA inclui tarefas como aprendizagem, raciocínio, planejamento, percepção, compreensão de linguagem e robótica.

Contudo, o conceito de inteligência artificial em verdade nasceu com o cientista e matemático John McCarthy, em 1956, na Dartmouth College Artificial Intelligence Conference: the next fifty years. McCarthy postulou que uma IA pode ser definida como a ciência e a engenharia de fazer máquinas inteligentes, derivadas especialmente de programas computacionais inteligentes, relacionando-se à tarefa semelhante de usar computadores para compreender a inteligência humana, exceto que uma IA não se limita a métodos biologicamente observáveis (DIVINO, p. 157).

Mas antes de McCarthy, há de se mencionar outros dois estudos que tiveram grande participação no que hoje se conhece como inteligência artificial (IA).

O primeiro trabalho foi um modelo de neurônios artificiais elaborado por Warren McCulloch e Walter Pitts, em 1943, que serviu como precursor da abordagem conexionista (FRANCO, 2017, p. 6).

O outro trabalho e um dos mais conhecidos a tratar a questão da inteligência de máquinas computacionais foi escrito em 1950 pelo matemático britânico Alan

Turing, conhecido por sua colaboração na decodificação da criptografia de mensagens do exército alemão na Segunda Guerra Mundial (FRANCO, 2017, p. 6).

Conforme aborda Franco (2017, p. 7):

O Teste de Turing ponderava sobre a possibilidade de uma máquina inteligente emular o comportamento de um ser humano. Turing propunha colocar um interlocutor conversando com duas entidades distintas sem a possibilidade de vê-las. Uma das entidades era uma pessoa e a outra uma máquina inteligente. O interlocutor deveria descobrir qual das duas entidades seria a máquina através de um interrogatório (...). Caso o interlocutor fosse ludibriado, dir-se-ia que a máquina seria inteligente. Existem diversas objeções ao Teste de Turing, ou seja, argumentos para refutar a hipótese de que o teste seria suficiente para determinar a inteligência de uma máquina.

Desde o início os pesquisadores de IA eram otimistas em seus prognósticos e previam a criação de um computador tão ou mais inteligente do que um humano em um período relativamente curto. O otimismo inicial logo cessou, na medida em que problemas mais complexos eram submetidos com fracasso aos sistemas de IA existentes. A escassez de poder computacional da época era senão o seu maior alçoz (FRANCO, 2017, p. 8).

Ainda no século XVIII, com o advento da revolução industrial, houve um grande avanço tecnológico. A máquina passou a executar, com muito mais competência e rapidez que o homem, tarefas monótonas, pesadas e repetitivas (PALÁCIOS apud SANTOS; SILVA; BONFIM, 2020, p. 2).

Russel e Norvig (apud FRANCO, 2017, p. 8) revelam que o quadro de resolução de problemas que surgiu durante as primeiras duas décadas de pesquisa em IA foi o de um mecanismo de busca de uso geral, que procurava reunir passos elementares de raciocínio para encontrar soluções completas.

Na década de 80, os sistemas especialistas começaram a se tornar viáveis para o mercado corporativo, sendo utilizados por diversas empresas, principalmente nos Estados Unidos e Japão (FRANCO, 2017, p. 8).

Nos últimos anos, com o surgimento da internet, dos grandes volumes de dados e dos mecanismos de pesquisa, as IA's estão se tornando cada vez mais "agentes inteligentes" capazes de abordagens simbólicas e conexionistas podendo trabalhar de forma colaborativa para a resolução de problemas muito complexos, as chamadas Redes Neurais Artificiais – RNA's (FRANCO, 2017, p. 8).

Os estudos das RNA's são motivados pela ambição humana de criar máquinas autônomas capazes de operar sem a intervenção do homem. Isso implica em dizer que elas devem aprender sozinhas e operar de forma inteligente na resolução de problemas inesperados. Desse modo, através de interações com o meio externo, devem ser capazes de aprender, fazer associações, generalizações, abstrações, dentre outras; deixando para trás a execução de ações pré-programadas para dar lugar à capacidade cognitiva (SILVA; BEZERRA, 2018, p. 81-82).

Hoje, sobretudo, a grande discussão que paira sobre a temática das inteligências artificiais é senão sobre a sua autonomia e automatização

É importante deixar claro que a automatização dos sistemas é a fase que antecede a autonomia. Contudo, hoje, esta não exclui aquela, mas coexistem, cada uma em seu contexto (SILVA; BEZERRA, 2018, p. 82).

Em uma análise histórica da robótica, Haselager (apud LEAL, 2016, p. 73):

(...) os sistemas automáticos foram desenvolvidos buscando alcançar uma menor dependência humana, ou seja, o ser humano não precisaria ficar constantemente operando-os, pois tais sistemas conseguiam repetir uma sequência de ações constante que poderia ser interrompida pelo ser humano a qualquer momento.

Desse modo, os sistemas automáticos auxiliam em atividades humanas, maiormente as repetitivas, conferindo maior precisão e agilidade. Contudo, nos sistemas ditos automáticos, é necessária a intervenção humana, uns com maior, outros com menor frequência. Esse comportamento pode ser facilmente evidenciado, por exemplo, no funcionamento de uma impressora, do piloto automático de um automóvel ou de um braço mecânico em uma montadora de veículos (SILVA; BEZERRA, 2018, p. 82).

Destarte, os sistemas automáticos são semiautônomos, vez que esses dependem da intervenção humana. Portanto, quanto menor o grau de intervenção humana em um sistema funcional, maior seu grau de autonomia (Haselager apud LEAL, 2016, p. 73).

Assim, surgem duas propostas filosóficas para explicar a ontologia de uma IA. A primeira é denominada IA fraca, cautelosa, ou Weak AI, postulando que uma máquina consegue apenas simular os fenômenos intencionais e o poder causal cerebral humano (DIVINO, 2020, p. 158).

Este campo tem apresentado avanços expressivos e concretos nos últimos anos. Suas aplicações têm objetivos específicos, como tradução de textos ou classificação de e-mails como spam ou não. Sistemas operacionais como Google Tradutor e GPS são exemplos de IA Fraca (SANTOS; SILVA; BONFIM, 2020, p. 5).

Afirma Coppin (2013, p. 5) que as IA's Fracas:

Usam sistemas tais como lógica, raciocínio automatizado e outras estruturas gerais que podem ser aplicadas a uma ampla gama de problemas, mas que não necessariamente incorporam qualquer conhecimento genuíno sobre o mundo do problema que está sendo solucionado.

A segunda proposta filosófica é a IA forte, ou Strong AI, propondo que um ente artificialmente inteligente pode reproduzir estados e fenômenos mentais de forma idêntica às operações humanas; sem supervisão ou orientação, tornando-se independente e superando a inteligência humana, atingindo um nível de consciência (DIVINO, 2020, p. 158; SANTOS; SILVA; BONFIM, 2020, p. 4)

Entretanto, a tarefa de desenvolver a dita “consciência artificial” não é simples e nem está perto de ser completamente concluída, conforme assevera Mussa (2020, p. 28):

Para que tenhamos uma ideia do quão longe estamos do desafio de desenvolvermos Inteligência Artificial dotada de consciência, hoje não temos nem sequer um caminho para alcançarmos dois passos anteriores, extremamente importantes: conseguirmos fazer com que os algoritmos aprendam com base em poucos exemplos e conseguirmos generalizar o uso desses algoritmos que foram desenvolvidos para situações específicas.

Assim, a tecnologia ainda tem muito a desenvolver para termos IA's como as vistas em filmes e livros, totalmente autônomas e “conscientes”.

Além disso, é importante mencionar sobre os métodos instrucionais de aprendizagem das IA's, que se subdividem em: aprendizado de máquina (machine learning) e aprendizado profundo (deep learning).

O machine learning (aprendizado de máquina) é uma técnica que permite que os sistemas de computador internalizem e façam previsões baseadas em dados históricos. O processo é alimentado por um algoritmo de aprendizado de máquina, uma função capaz de melhorar seu desempenho com o tempo, treinando-se por meio

de métodos de análise de dados e modelagem analítica (SMOLA; VISHWANATHAN apud LAGE, 2021, p. 27).

Já o deep learning (aprendizado profundo) é um subdomínio do aprendizado de máquina, que consiste em múltiplas camadas em cascata, modeladas a partir do sistema nervoso humano (RNA's). Arquiteturas de aprendizagem profunda permitem que um sistema de computador se treine usando dados históricos, reconhecendo padrões e fazendo inferências probabilísticas; o qual contextualizará situações fáticas postas e as transcenderá para além de sua programação inicial (LAGE, 2021, p. 27; DIVINO, 2020, p. 155).

Tanto o machine quanto o deep learning estão em rápido desenvolvimento e utilização no cenário tecnológico contemporâneo, mas por ora ainda não atingiu seu máximo potencial.

Assim, a Inteligência Artificial, pelo menos por enquanto, nada mais é que um complexo programa de computador. As supostas habilidades inferidas pelo deep learning ou pelo machine learning não alteram sua natureza. Uma IA pode simular fenômenos mentais, mas não será, literalmente, inteligente. Porque enquanto simulação, ela não reproduz (DIVINO, 2020, p. 160).

Um recente estudo denominado *Moral Machine*, feito pelo Massachusetts Institute of Technology (MIT) em associação com outras instituições como Max Planck Institut e University of British Columbia (UBC), analisa decisões humanas em situações extremas envolvendo carros autônomos para incrementá-las no *machine learning* e no *deep learning*, para verificar como uma IA agiria diante da mesma situação (DIVINO, 2020, p. 156).

Segundo informa o próprio desenvolvedor (MORAL..., 2022):

Estudos científicos recentes sobre a ética das máquinas têm aumentado a atenção sobre o tema, na mídia e na opinião pública. Este site tem como objetivo levar a discussão mais adiante, fornecendo uma plataforma para: 1) a construção colaborativa de um retrato da opinião das pessoas sobre como as máquinas devem tomar decisões quando confrontadas com dilemas morais, e 2) gerar uma discussão coletiva sobre as consequências morais desses cenários em potencial.

Ao se fazer o teste são mostrados vários cenários envolvendo carros autônomos tripulados ou não, devendo-se escolher entre duas opções que induzem

dilemas éticos e morais, por exemplo, se o carro autônomo deve priorizar a vida do condutor ou de um pedestre, ou a vida de um pedestre humano ou de um animal, dentre outros.

O estudo revelou três preferências relativamente universais. Em média, as pessoas queriam: (1) poupar vidas humanas sobre animais; (2) salvar mais vidas em menos; (3) priorizar os jovens sobre os mais velhos. Devem auxiliar na programação de inteligências artificiais diante de dilemas éticos e morais (LEVENTHAL, 2018).

Por fim, faz-se importante diferenciar um Robô de uma IA. Divino (2020, p. 161) afirma ser um robô o corpo mecânico físico (hardware), um receptáculo intermediário entre sua interface computacional e o mundo, através da programação de um complexo algoritmo internalizado a este (software).

Para os fins desse trabalho, não há necessidade de diferenciação entre IA e Robô. O enfoque principal será sobre a IA que comanda as ações robóticas, por assim dizer, aquelas aptas ao cometimento de delitos, que serão auferidos posteriormente.

4. DIREITO PENAL, TEORIA DO CRIME E RESPONSABILIDADE PENAL NO ORDENAMENTO JURÍDICO BRASILEIRO

O Direito penal, como um instrumento do controle social formal, tem a finalidade de assegurar a estabilidade e a sobrevivência em sociedade. Considerada como a *ultima ratio*, o Direito Penal tem o papel de resolver conflitos não alcançados por outros ramos do Direito que possam antes poderiam solucioná-los.

Assim conforme leciona Claus Roxin (1997, p. 65):

O Direito penal é o último recurso de proteção que se deve recorrer, em outras palavras, só se pode intervir penalmente quando falharem outros meios de solução social do problema – como as sanções civis, a intervenção policial ou outra intervenção jurídico-técnica, isto é, as sanções não-penais.

Destaca-se que as condutas mais graves que afligem bens jurídicos protegidos e assim causam a insegurança jurídica, são um risco para o convívio em sociedade serão tratadas pelo Direito Penal.

Em decorrência, conforme Mirabete (1990, p. 21), para que possa existir um convívio harmônico entre os integrantes da sociedade, necessário se faz que haja regras regulamentando as condutas aceitas e as condutas reprováveis, essas regras denominam-se Direito Positivo e devem ser obedecidas por todos.

Para perpetuar e garantir a obediência a estas normas são estabelecidas sanções penais, que tem um tríplice propósito segundo Capez (2011, p. 384): “de aplicar a retribuição punitiva ao delinquente, promover a sua readaptação social e prevenir novas transgressões pela intimidação dirigida à coletividade”.

Destarte, o sistema penal tem este caráter coercitivo característico incumbindo ao Estado o monopólio do uso desta força para perpetuar o controle social necessário. Nesta seara a importância do conjunto de normas penais que impõe os limites de condutas e as sanções para aquelas outras socialmente não aceitas.

Outra questão pertinente a se analisar é quem seria o sujeito ativo, ou quem pratica as condutas protegidas por estas normas. O já mencionado jurista Capez (2011) acredita que o Direito Penal é restrito ao agente humano, assim somente o comportamento humano seria capaz de transgredir as normas penais. No mesmo sentido, assevera Rogério Sanchez Cunha (2016, p. 177):

De acordo com o princípio da exteriorização ou materialização do fato, ao Direito Penal só interessam fatos humanos, pouco importando os acontecimentos da natureza dos quais não participa o homem. Entretanto, não são todos os fatos humanos que ficam na mira do Direito Penal, mas somente aqueles indesejados pelo meio social, não reprovados de forma eficaz pelos demais ramos do Direito e que provoquem relevante e intolerável lesão ao bem jurídico tutelado.

A delimitação das ações humanas como as únicas tratadas pelo Direito Penal é senão a vertente pacificamente aceita pela doutrina, tendo como expoentes grandes penalistas como Eugenio Raúl Zaffaroni, Guilherme de Souza Nucci e Rogério Greco.

Dessa forma, podemos afirmar que o sujeito ativo do crime/delito em regra será a pessoa humana; em regra, pois há grande discussão sobre a responsabilização penal das pessoas jurídicas, divergindo a doutrina quanto ao tema; embora a jurisprudência e legislação já tenham se manifestado a favor da responsabilização em casos de crimes ambientais.

Vale ressaltar que ao se aceitar a responsabilidade penal da pessoa jurídica, não se exclui a responsabilidade da pessoa física coautora ou partícipe do delito (MASSON, 2011, p. 182).

Além disso as sanções penais não se limitam tão somente às penas privativas de liberdade, razão que entidades imateriais, como é o caso de pessoas jurídicas em crimes ambientais por exemplo, possam ser penalizadas alternativamente.

Nessa perspectiva, paira a dúvida: Seria possível classificar a inteligência artificial como sujeito ativo de um crime? Para responder essa pergunta necessário se faz antes analisar a estrutura de um crime, o que se fará a seguir.

Os institutos do Direito Penal são abarcados por diversos sistemas e teorias que tentam explicar o crime e sua relação na sociedade. De certo pode-se destacar três prismas relevantes: material, legal e analítico.

De acordo com um conceito material ou substancial, o crime seria “toda ação ou omissão humana que lesa ou expõe a perigo de lesão bens jurídicos penalmente tutelados” (MASSON, 2011, p. 169). “É a concepção da sociedade sobre o que pode e deve ser proibido, mediante a aplicação de sanção penal” (NUCCI, 2014, p. 137). Trata-se de um conceito abstrato norteador para dar ao legislador um critério político-criminal sobre o Direito Penal (ROXIN, 1997, p. 51).

O conceito de crime legal ou formal é o fornecido pelo legislador. A Lei de Introdução ao Código Penal (Decreto-lei 3.914, de 9 de dezembro de 1941) em seu art. 1º cuidou de definir crime sob a ótica legal:

Art 1º Considera-se crime a infração penal que a lei comina pena de reclusão ou de detenção, quer isoladamente, quer alternativa ou cumulativamente com a pena de multa; contravenção, a infração penal a que a lei comina, isoladamente, pena de prisão simples ou de multa, ou ambas, alternativa ou cumulativamente.

Em verdade, trata-se do conceito material, devidamente formalizado pelo legislador; estabelecendo as condutas socialmente reprováveis materializadas em tipos penais em respeito ao princípio da legalidade (NUCCI, 2014, p. 137).

Já o conceito analítico ou dogmático, se funda nos elementos que compõem a estrutura do crime. “Na realidade, é o conceito formal fragmentado em elementos que propiciam o melhor entendimento da sua abrangência” (NUCCI, 2014, p. 137).

Em relação ao terceiro critério, identificam-se diversas classificações doutrinárias, sendo a teoria tripartida apresentada no sistema finalista uma das mais importantes ao explicar o conceito de crime através de três elementos: fato típico, antijurídico, culpável (DELMANTO apud ALVES, 2021).

Sobre o tema Hans Welzel (apud GRECO, 2017, p. 221), precursor da teoria finalista, disserta:

A tipicidade, a antijuridicidade e a culpabilidade são três elementos que convertem uma ação em um delito. A culpabilidade – a responsabilidade pessoal por um fato antijurídico – pressupõe a antijuridicidade do fato, do mesmo modo que a antijuridicidade, por sua vez, tem de estar concretizada em tipos legais. A tipicidade, a antijuridicidade e a culpabilidade estão relacionadas logicamente de tal modo que cada elemento posterior do delito pressupõe o anterior.

Outrossim, existem várias posições sobre os elementos que compõe um crime, se tornando um dos pontos de maior divergência doutrinária. Nucci (2014, p. 138-139), apresenta essas vertentes e destaca os principais doutrinadores que as defendem:

Há quem entenda ser o crime, do ponto de vista analítico:
a) um fato típico e antijurídico, sendo a culpabilidade apenas um pressuposto de aplicação da pena (René Ariel Dotti, Damásio de Jesus, Julio Fabbrini Mirabete, Celso Delmanto, André Estefam, Cleber Masson, entre outros);

- b) um fato típico, antijurídico, culpável e punível (Basileu Garcia, Muñoz Conde, Hassemer, Battaglini, Giorgio Marinucci e Emilio Dolcini, entre outros);
- c) um fato típico e culpável, estando a antijuridicidade insita ao próprio tipo (Miguel Reale Júnior, entre outros adeptos da teoria dos elementos negativos do tipo);
- d) fato típico, antijurídico e punível, constituindo a culpabilidade a ponte que liga o crime à pena (Luiz Flávio Gomes);
- e) um fato típico, antijurídico e culpável. Nesta corrente, que é majoritária no Brasil e no exterior, e com a qual concordamos, dividem-se finalistas (Assis Toledo, Heleno Fragoso, Juarez Tavares, José Henrique Pierangeli, Eugenio Raúl Zaffaroni, Fernando de Almeida Pedroso, Jair Leonardo Lopes, Cezar Roberto Bitencourt, Luiz Regis Prado, Rodolfo Tigre Maia, Jorge Alberto Romeiro, Luiz Luisi, David Teixeira de Azevedo, Rogério Greco, Reinhart Maurach, Heinz Zipf, Bustos Ramírez, Jiménez Martínez, entre outros) e causalistas (Nélson Hungria, Frederico Marques, Aníbal Bruno, Magalhães Noronha, Paulo José da Costa Júnior, Vicente Sabino Júnior, Salgado Martins, Euclides Custódio da Silveira, Manoel Pedro Pimentel, Roque de Brito Alves, Fernando de Almeida Pedroso, Jiménez de Asúa, Baumann, Mezger, entre outros), além daqueles que são adeptos da teoria social da ação, propondo-se a ser uma tentativa de ajuste, num só quadro, dos principais aspectos do causalismo e do finalismo (Jescheck, Wessels, Schmidt, Engisch, Wolff, entre outros).

Atualmente, o Código Penal Brasileiro, em especial após a edição da Lei 7.209/1984, adota a teoria do finalismo dissidente (ou bipartite), ensinando que o crime é composto de apenas dois substratos: fato típico e antijuridicidade; ao passo que a presença ou não da culpabilidade importará na possibilidade ou não de a pena ser imposta.

Nas palavras de Masson (2011, p. 177):

(...) crime é o fato típico e ilícito, independentemente da culpabilidade, que tem a imputabilidade penal como um dos seus elementos. O crime existe sem a culpabilidade, bastando seja o fato típico e revestido de ilicitude. (...) O crime se refere ao fato (típico e ilícito), enquanto a culpabilidade guarda relação com o agente (merecedor ou não de pena).

Vencido este assunto, importa analisar os elementos do crime, em primeiro lugar o fato típico.

Para Nucci (2014, p. 155), “o fato típico é a síntese da conduta ligada ao resultado pelo nexos causal, amoldando-se ao modelo legal incriminador”. Masson (2011, p. 209) já entende que “fato típico é o fato humano que se enquadra com perfeição aos elementos descritos pelo tipo penal”.

O fato típico é composto por outros quatro elementos, quais sejam: conduta, resultado naturalístico, relação de causalidade (nexo causal) e tipicidade (MASSON, 2011, p. 209).

Em primeiro, quando se fala em conduta pode-se relacioná-lo imediatamente à ação, comportamento. Bleger (apud NUCCI, 2014, p. 161) busca explicar o termo “conduta” sob a ótica psicológica:

Etimologicamente, a palavra conduta é latina e significa conduzida ou guiada; quer dizer, que todas as manifestações compreendidas no termo de conduta são ações conduzidas ou guiadas por algo que está fora das mesmas: pela mente. Dessa maneira, o estudo da conduta, considerada assim, assenta sobre um dualismo ou uma dicotomia corpo-mente, sobre a tradição do mais puro idealismo, no qual a mente tem existência própria e é o ponto de origem de todas as manifestações corporais; segundo essa perspectiva, o corpo é somente um instrumento ou um veículo do qual se vale a mente (alma) para se manifestar.

No prisma jurídico, o conceito de conduta adquire diferentes pontos de vista, sendo que diversas teorias foram elaboradas para tentar explicá-la; sendo a posição finalista a majoritariamente aceita.

Na visão finalista, conduta é a ação ou omissão, voluntária e consciente, implicando em um comando de movimentação ou inércia do corpo humano, voltado a uma finalidade (NUCCI, 2014, p. 161).

Conforme se extrai do conceito de Welzel, a conduta pode se exteriorizar por ação ou por omissão.

Segundo Masson (2011, p. 217), a ação consiste em um movimento corporal exterior, exigindo uma postura positiva, um fazer; já a omissão, seria quando o agente nada faz, bem como quando faz algo diferente daquilo que lhe impunha o dever jurídico de agir, assim uma postura negativa.

Outro ponto relevante a se considerar sobre a conduta que também se extrai do conceito finalista é senão sobre a vontade.

Sobre vontade preleciona Zaffaroni (apud GRECO, 2017, p. 283):

A vontade implica sempre uma finalidade, porque não se concebe que haja vontade de nada ou vontade para nada; sempre a vontade é vontade de algo, quer dizer, sempre a vontade tem um conteúdo, que é uma finalidade.

Nucci (2014, p. 162) classifica a vontade como o querer ativo, apto a levar o ser humano a praticar um ato, livremente. Assim, “se o agente não atua dolosa ou culposamente, não há ação” (GRECO, 2017, p. 283).

Nesta seara, não há o que se falar em conduta voluntária nos casos de coação física irresistível, ou em movimentos reflexos, ou também em casos de sonambulismo e hipnose, bem como nos casos fortuitos ou de força maior (MASSON, 2011, p. 219-220).

Diz-se que somente as condutas voluntárias interessam ao Direito Penal (MASSON, 2011, p. 218), razão pela qual a vontade se caracteriza quando o agente deseja o resultado (dolo direto) e o assentimento seria quando o agente assume o risco de produzir o resultado (dolo eventual).

Em se tratando de dolo e culpa, a teoria finalista dissidente (ou bipartida) entende que ambos constituem elementos da conduta, sendo parte da análise do crime, já a culpabilidade integra um juízo de valor empregado contra aqueles que preencham os requisitos de tipicidade e antijuridicidade; conforme afirma Assis Toledo (1994, p. 228):

A consequência lógica, inarredável, foi igualmente a localização do dolo e da culpa *stricto sensu* no tipo legal do crime, pois, se este é a descrição da ação proibida, e se o dolo e a culpa pertencem à ação, não se pode deixar de situar no tipo todos os elementos estruturais da ação.

Explicando o tema, age com dolo o agente quer diretamente o resultado ou assume o risco de produzi-lo (art. 18, I do CP), e com culpa, quando dá causa ao resultado em virtude de sua imprudência, imperícia ou negligência (art. 18, II do CP); vale ressaltar que o Código Penal preceitua que todo crime seja doloso, somente sendo punida a culpa quando houver previsão legal expressa nesse sentido (art. 18, parágrafo único do CP).

Ainda sobre o conceito de conduta finalista, a consciência da ilicitude integra a culpabilidade e será abordada mais à frente.

Agora, voltando aos elementos que compõe o fato típico, importante se faz comentar sobre o resultado.

Masson (2011, p. 220) afirma que o resultado é a consequência provocada pela conduta do agente.

Existem duas modalidades de resultado, o primeiro é o resultado naturalístico que produz a modificação sensível do mundo exterior, enquanto o resultado jurídico ou normativo induz a modificação no mundo jurídico, seja na forma de dano efetivo ou de dano potencial, ferindo interesse protegido pela norma penal (NUCCI, 2014, p. 165).

Segundo o disposto na Exposição de Motivos do Código Penal de 1940, que a Reforma Penal de 1984 manteve, “não há crime sem resultado”, prevalecendo assim na doutrina pátria o conceito naturalístico de resultado. Por esse motivo faz-se diferença entre crimes de atividade (formais e de mera conduta) e de resultado (materiais) (NUCCI, 2014, p. 166).

Ainda sobre fato típico, outro elemento é a relação de causalidade ou nexos causal.

Para Masson (2011, p. 222):

Relação de causalidade é o vínculo formado entre a conduta praticada por seu autor e o resultado por ele produzido. É por meio dela que se conclui se o resultado foi ou não provocado pela conduta, autorizando, se presente a tipicidade, a configuração do fato típico.

O Código Penal acolheu, como regra, a teoria da equivalência dos antecedentes (*conditio sine qua non*), sendo “causa” todo comportamento humano, comissivo ou omissivo, que de qualquer modo concorreu para a produção do resultado naturalístico, por menor que seja o seu grau de contribuição; excepcionalmente, tem-se a teoria da causalidade adequada (§1º do art. 13 CP), quando se fala sobre as concausas (causas concomitantes que se unem para gerar o resultado) (MASSON, 2011, p. 223).

Além disso, para que um acontecimento ingresse na relação de causalidade, não basta a mera dependência física. Exige-se ainda a causalidade psíquica (*imputatio delicti*), assim a presença do dolo ou da culpa por parte do agente em relação ao resultado. De fato, a falta do dolo ou da culpa afasta a conduta, a qual, por seu turno, obsta a configuração do nexos causal (MASSON, 2011, p. 224).

Por fim, o quarto elemento que compõe o fato típico é a tipicidade, concretizada pelo tipo penal.

Na definição de Zaffaroni (apud GRECO, 2017, p. 287), “o tipo penal é um instrumento legal, logicamente necessário e de natureza predominantemente descritiva, que tem por função a individualização de condutas humanas penalmente relevantes”.

Greco (2017, p. 287) complementa:

Quando a lei em sentido estrito descreve a conduta (comissiva ou omissiva) com o fim de proteger determinado bem cuja tutela mostrou-se insuficiente pelos demais ramos do direito, surge o chamado tipo penal. Tipo, como a própria denominação diz, é o modelo, o padrão de conduta que o Estado, por meio de seu único instrumento – a lei –, visa a impedir que seja praticada, ou determina que seja levada a efeito por todos nós.

Nesta perspectiva, o tipo penal é senão o elemento que permite concretizar o princípio da reserva legal, onde não há crime sem lei anterior que o defina.

A adequação da conduta do agente ao modelo abstrato previsto na lei penal (tipo) faz surgir a tipicidade formal ou legal. Essa adequação deve ser perfeita, pois, caso contrário, o fato será considerado formalmente atípico. Assim, a subsunção perfeita da conduta praticada pelo agente ao modelo abstrato previsto na lei penal, faz surgir o tipo penal incriminador (GRECO, 2017, p. 288).

Destarte, “o tipo penal é transformado em tipo de injusto, ou seja, o tipo passa a ser conceituado como a ilicitude tipificada” (MASSON, 2011, p. 245).

Conforme exposto o entendimento de tipo penal está intimamente ligado à compreensão de ilicitude, que é senão elemento definidor do conceito de crime anteriormente elencado.

Sobre o tema, ensina Masson (2011, p. 365):

Illicitude é a contrariedade entre o fato típico praticado por alguém e o ordenamento jurídico, capaz de lesionar ou expor a perigo de lesão bens jurídicos penalmente tutelados. O juízo de ilicitude é posterior e dependente do juízo de tipicidade, de forma que todo fato penalmente ilícito também é, necessariamente, típico.

Assim, se a norma penal proíbe determinada conduta sob a ameaça de uma sanção, é porque aquela conduta ou causa lesão ou expõe a perigo de lesão o bem juridicamente protegido, e se o agente insiste em praticá-la devemos concluir pela sua

ilicitude, desde que não atue amparado por uma causa de justificação (GRECO, 2017, p. 452).

De certo, pode-se afirmar que a tipicidade não constitui a ilicitude, apenas a revela “indiciariamente”, como queria Miguel Reale (1974, p. 36); vez que há ações típicas que, na precisa lição de Aníbal Bruno (1967, p. 365):

Pela posição particular em que se encontra o agente ao praticá-las, se apresentam em face do Direito como lícitas. Essas condições especiais em que o agente atua impedem que elas venham a ser antijurídicas. São situações de excepcional licitude que constituem as chamadas causas de exclusão da antijuridicidade, justificativas ou discriminantes.

O código penal, em seu art. 23, previu expressamente quatro causas que afastam a ilicitude da conduta praticada pelo agente, fazendo, assim, com que o fato por ele cometido seja considerado lícito, a saber: o estado de necessidade (arts. 23, I, e 24); a legítima defesa (arts. 23, II, e 25); o estrito cumprimento do dever legal (art. 23, III); e o exercício regular de direito (art. 23, III).

Há ainda excludentes de ilicitude previstas na Parte Especial do Código Penal e válidas, apenas, para alguns delitos, à exemplo do aborto necessário (art. 128, I, CP), ou outras previstas em legislação extrapenal, como ocorre na legítima defesa prevista no Código Civil (art. 1.210, § 1.º); bem como nas excludentes supralegais, não prevista expressamente em lei mas amparadas pelo código penal, merecendo destaque o consentimento do ofendido, qual seja o desinteresse da vítima em fazer valer a proteção legal ao bem jurídico que lhe pertence (GRECO, 2017, p. 454-455; NUCCI, 2014, p. 202).

Vencido estas exposições e conforme foi evidenciado anteriormente, o crime é composto por um fato típico e ilícito, conforme queria a visão finalista dissidente adotada pelo sistema penal brasileiro; sendo necessário avaliar todos os elementos que o compõe, assim a conduta do agente (se voluntária, ativa ou omissiva e dolosa ou culposa) que acabou por gerar um determinado resultado naturalístico, que foi causado por determinados fatores (nexo causal), sendo que somente se levará a efeito quando exista um tipo penal adequadamente subsumido ao fato. Além disso, o tipo penal de correspondência direta com a ilicitude da conduta, levando à análise ações socialmente e penalmente reprováveis, observados os casos de excludentes de ilicitude, tem-se a configuração do crime.

Num segundo momento, objetiva-se analisar a culpabilidade do agente. Correspondente ao que foi exposto, a culpabilidade não é uma característica do crime, mas sim um pressuposto para a incidência da pena. Nas palavras de Damásio de Jesus (2011, p. 7):

A culpabilidade não é elemento ou requisito do crime. O juízo de reprovabilidade não incide sobre o fato, mas sim sobre sujeito. Não se trata de fato culpável, mas sim de sujeito culpável. Culpabilidade é um juízo de reprovação que recai sobre o sujeito que praticou delito. Por isso conceituamos o crime como fato típico e antijurídico.

Assim também é o pensamento do jurista Masson (2011, p. 436):

Culpabilidade é o juízo de censura, o juízo de reprovabilidade que incide sobre a formação e a exteriorização da vontade do responsável por um fato típico e ilícito, com o propósito de aferir a necessidade de imposição de pena.

Além disso, a culpabilidade tem o condão de dimensionar uma pena justa e socialmente aceita, de forma a evitar excessos ou que seja demasiadamente condescendente com o ato criminoso. Na lição de Bitencourt (2012, p.164):

A culpabilidade apresenta-se como fundamento e limite para a imposição de uma pena justa. Por outro lado, a culpabilidade também é entendida como um instrumento para a prevenção de crimes e, sob essa ótica, o juízo de atribuição de responsabilidade penal cumpre com a função de aportar estabilidade ao sistema normativo, confirmando a obrigatoriedade do cumprimento das normas.

O Código Penal não apresenta o conceito de culpabilidade, cabendo à doutrina esta tarefa. Desta forma, o conceito de culpabilidade já apresentou significativa evolução ao longo dos tempos, sendo formuladas diversas teorias.

Por expressa previsão legal, o Código Penal acolheu a teoria limitada, é o que se observa no item 19 da Exposição de Motivos da Nova Parte Geral do Código Penal (BRASIL, 1983):

Repete o Projeto as normas do Código de 1940, pertinentes às denominadas “discriminantes putativas”. Ajusta-se, assim, o Projeto à **teoria limitada da culpabilidade**, que distingue o erro incidente sobre os pressupostos fáticos de uma causa de justificação do que incide sobre a norma permissiva. [grifo nosso]

Na teoria limitada, uma espécie de variante da teoria normativa pura, a culpabilidade é composta pelos elementos de: imputabilidade, potencial consciência da ilicitude e exigibilidade de conduta diversa; e as discriminantes putativas são divididas em dois blocos: de fato, tratadas como erro de tipo (CP, art. 20, § 1º); e de direito, disciplinadas como erro de proibição (CP, art. 21) (MASSON, 2011, p. 441-442).

Outrossim, conforme leciona Masson (2011, p. 441):

(...) [os] elementos constitutivos da culpabilidade estão ordenados hierarquicamente, de tal modo que o segundo pressupõe o primeiro, e o terceiro os dois anteriores. De fato, se o indivíduo é inimputável, não pode ter a potencial consciência da ilicitude. E, se não tem a consciência potencial da ilicitude, não lhe pode ser exigível conduta diversa.

Dito estes conceitos iniciais, passa-se a analisar os elementos da culpabilidade, começando-se com a imputabilidade.

Para Nucci (2014, p. 241), imputabilidade pode ser definida como:

(...) o conjunto das condições pessoais, envolvendo inteligência e vontade, que permite ao agente ter entendimento do caráter ilícito do fato, comportando-se de acordo com esse conhecimento.

Revela-se que a imputabilidade tem caráter individualizante, evidenciando dois critérios indispensáveis quais sejam a inteligência e a vontade do agente quanto ao ato ilícito praticado; também é a visão de Sanzo Brodt (apud GRECO, 2017, p. 530):

A imputabilidade é constituída por dois elementos: um intelectual (capacidade de entender o caráter ilícito do fato), outro volitivo (capacidade de determinar-se de acordo com esse entendimento). O primeiro é a capacidade (genérica) de compreender as proibições ou determinações jurídicas. Bettiol diz que o agente deve poder 'prever as repercussões que a própria ação poderá acarretar no mundo social', deve ter, pois, 'a percepção do significado ético-social do próprio agir'. O segundo, a 'capacidade de dirigir a conduta de acordo com o entendimento ético-jurídico. Conforme Bettiol, é preciso que o agente tenha condições de avaliar o valor do motivo que o impele à ação e, do outro lado, o valor inibitório da ameaça penal.

Tal prerrogativa será importante quanto se avalia a imputabilidade da Inteligência Artificial que será tratada posteriormente.

Ademais, nosso Código Penal não define o conceito de imputabilidade, mas aponta as hipóteses em que a imputabilidade está ausente, ou seja, os casos de inimputabilidade penal: menoridade (art. 27); doença mental (art. 26, caput); desenvolvimento mental incompleto (arts. 26, caput, e 27); desenvolvimento mental retardado (art. 26, caput); e embriaguez completa proveniente de caso fortuito ou força maior (art. 28, § 1º).

Outro elemento que compõe a análise de culpabilidade é a potencial consciência sobre a ilicitude do fato.

Assevera Nucci (2014, p. 163) que a consciência:

(...) é a possibilidade que o ser humano possui de separar o mundo que o cerca dos próprios atos, realizando um julgamento moral das suas atitudes. Significa ter noção clara da diferença existente entre realidade e ficção.

O conceito de consciência é de difícil constatação, vez que se busca noutras áreas de conhecimento, como a Psicologia sua definição. Contudo, não se pode absorver o conceito de consciência simplesmente como ele é apresentado na Psicologia. Com efeito, é tarefa do hermenauta do Direito adequar este conceito à categoria lógico-jurídica, atrelada assim ao conceito de ilicitude (BRANDÃO, 1997).

Assim, para Masson (2011, p. 475):

A aplicação da pena ao autor de uma infração penal somente é justa e legítima quando ele, no momento da conduta, era dotado ao menos da possibilidade de compreender o caráter ilícito do fato praticado. Exige-se, pois, tivesse o autor o conhecimento, ou, no mínimo, a potencialidade de entender o aspecto criminoso do seu comportamento, isto é, os aspectos relativos ao tipo penal e à ilicitude.

A potencial consciência da ilicitude é afastada pelo erro de proibição escusável (CP, art. 21, caput).

Diz a primeira parte do art. 21 do Código Penal que o desconhecimento da lei é inescusável, ou seja, a ignorância ou o erro de compreensão legal não eximem de pena o agente infrator. Em seguida, ao tratar do erro de proibição, diz que o erro sobre a ilicitude do fato, se inevitável, isenta de pena. Pela redação do artigo, percebe-se que o Código Penal tenta fazer uma distinção entre o desconhecimento da lei e a falta de conhecimento sobre a ilicitude do fato (GRECO, 2017, p. 542).

Na lição de João Mestieri (1999, p. 188):

A vox 'possibilidade de entender o caráter ilícito (criminoso) do fato' é genuinamente normativa, pois não se trata do conhecimento da ilicitude (operação de natureza psicológica), mas da mera possibilidade concreta desse conhecimento. O juízo de reprovação apenas se torna possível quando se constata que o agente teve, no caso específico, a possibilidade concreta de entender o caráter criminoso do fato praticado e assim determinar o seu comportamento de acordo com os interesses do sistema jurídico. O erro sobre a ilicitude do fato é erro de proibição; dá-se quando o agente por ignorância (*ignorantia iuris*) ou por uma representação falsa ou imperfeita da realidade supõe ser lícito o seu comportamento.

Tal análise é fundamental quando se avalia a provável consciência da Inteligência Artificial, o que se fará no próximo capítulo.

Além disso, importante se faz lembrar da individualização da culpabilidade, vez que a sua caracterização se relaciona ao perfil subjetivo do agente, e não a figura do "homem médio" cabendo nas hipóteses de exame de tipicidade e ilicitude do fato; nas palavras de Masson (2011, p. 480):

De fato, em se tratando de matéria inerente à culpabilidade, levam-se em conta as condições particulares do responsável pelo fato típico e ilícito (cultura, localidade em que reside, inteligência e prudência etc.), com a finalidade de se alcançar sua responsabilidade individual, que não guarda relação com um *standard* de comportamento desejado pelo Direito Penal. Lembre-se: quando se fala em fato típico e ilicitude, e em todos os institutos a eles relacionados, considera-se a posição do homem médio, pois se analisa o fato (típico ou atípico, ilícito ou lícito).

O último elemento que compõe a análise de culpabilidade refere-se à exigibilidade de conduta diversa, ocorrendo quando "o agente podia comportar-se em conformidade com o Direito, mas preferiu violar a lei penal" (MASSON, 2011, p. 485).

Assim, em contrassenso, a inexigibilidade de conduta diversa se configura quando o agente, mesmo violando uma norma jurídica expressa, não tem outra opção de conduta naquela situação de fato; não gerando, portanto, reprovabilidade social.

Também neste elemento da culpabilidade se verifica a individualização da pena, vez que a "possibilidade ou impossibilidade de agir conforme o direito variará de pessoa para pessoa, não se podendo conceber um 'padrão'" (GRECO, 2017, p. 547).

Nucci (2014, p. 254) alerta que "em certas situações extremadas, quando não for possível aplicar outras excludentes de culpabilidade, a inexigibilidade de conduta

diversa seja utilizada para evitar a punição injustificada do agente”; lembrando-se da hierarquia existente entre as excludentes de culpabilidade supramencionadas.

Em suma, o Código Penal prevê duas causas legais que excluem a culpabilidade, ambas no art. 22 CP: a coação irresistível e a obediência hierárquica. Além disso, há grande discussão doutrinária quanto às causas supralegais de inexigibilidade de conduta diversa, como é o caso do estado de necessidade exculpante e do excesso de legítima defesa exculpante, por exemplo.

Finalmente, cabe falar sobre a responsabilidade penal como aspecto geral do direito penal.

Sobre o tema, leciona Noronha (2003, p. 164):

Responsabilidade é a obrigação que alguém tem que arcar com as consequências jurídicas do crime. É o dever que tem a pessoa de prestar contas com seu ato. Ele depende da imputabilidade do indivíduo, pois não pode sofrer as consequências do fato criminoso (ser responsabilizado) senão o que tem a consciência de sua antijuridicidade e quer executá-lo (ser imputável).

Contudo, enquanto a imputabilidade pode ser compreendida como a capacidade de ser culpável e culpabilidade sendo o juízo de reprovação social que pode ser realizado ao imputável, a responsabilidade é decorrência da culpabilidade, ou seja, trata-se da relação entre o autor e o Estado, que merece ser punido por ter cometido um delito (NUCCI, 2014, p. 241).

Destarte, a responsabilidade penal preconiza seu caráter subjetivo, vez que a pena é pessoal e intransferível, não se estendendo a outras pessoas; é a consagração do “princípio da pessoalidade ou da intranscendência da pena, [onde] somente o condenado é que terá de se submeter à sanção que lhe foi aplicada pelo Estado” (GRECO, 2017, p. 159).

Também a visão de Zaffaroni (apud GRECO, 2017, p. 159):

Nunca se pode interpretar uma lei penal no sentido de que a pena transcenda da pessoa que é autora ou partícipe do delito. A pena é uma medida de caráter estritamente pessoal, haja vista ser uma ingerência ressocializadora sobre o condenado.

Com isso quer se dizer que somente aquele que de certa forma esteja envolvido com o ilícito penal poderá responder por ele.

Explicando melhor, configurado o fato típico e todos os seus elementos, como a conduta e a tipicidade, e sendo ilícita a ação, tem-se configurado o crime, contudo ainda se faz necessário a demonstração de culpa do agente, naqueles aspectos de imputabilidade, potencial consciência da ilicitude e exigibilidade de conduta diversa; vencidas todas as etapas torna-se o infrator responsável penalmente sendo-lhe aplicada uma sanção penal adequada.

Observa-se que não há presunção de culpa no nosso ordenamento jurídico penal, assim não há o que se falar em responsabilidade penal objetiva, incompatível com a presunção de inocência e outras garantias constitucionais (STRECK, 2015).

Vencidos estes conceitos, pretende-se no próximo tópico relacionar aspectos da teoria do crime e da culpabilidade com as Inteligências Artificiais, demonstrando inclusive modelos de responsabilização penal cabíveis, analisando a possível personalidade jurídica dessa tecnologia, bem como falar das perspectivas futuras no ordenamento jurídico brasileiro.

5. MODELOS DE RESPONSABILIZAÇÃO PENAL DE CONDUTAS DA INTELIGÊNCIA ARTIFICIAL: UM ESTUDO DA CULPA SOB A ÓTICA DA TEORIA DO CRIME

Sem dúvidas que o futuro reserva grandes mudanças. A tecnologia está numa ascendente nunca antes vista, havendo grandes investimentos em Inteligências Artificiais, as chamadas IA's, que hoje inevitavelmente estão inseridas nas relações sociais. Não há dúvidas nos benefícios que uma IA pode trazer para a humanidade, mas tem-se que analisar também “o outro lado da moeda”, avaliar quais os riscos que essa tecnologia independente pode oferecer. Dentre tantas incertezas, uma dúvida em particular se sobressai: o que se fazer quanto uma IA cometer um crime? Quem responsabilizar? Qual lei aplicar? É o que será discutido nesse capítulo.

Antes disso, interessante se faz comentar sobre um caso ocorrido na cidade de Tempe, no Arizona/EUA. Um acidente que aconteceu em março de 2018, envolvendo um Volvo XC90, atropelou e causou a morte de Elaine Herzberg, que tentava atravessar a rua fora da faixa de pedestres (UBER..., 2018).

A Uber usava o veículo para testar sua tecnologia de direção autônoma, que não conseguiu identificar a pedestre e frear o carro há tempo. O carro dirigia automaticamente à 70 km/h no momento do acidente, mas havia um “operador humano de segurança” caso houvesse a necessidade de tomar o controle do veículo. O veículo é equipado com várias câmeras, o que facilitou a perícia avaliar as principais causas do acidente (UBER..., 2018).

No momento do acidente, Elaine Herzberg não havia olhado antes de atravessar a rua em uma área mal iluminada e estava vestindo roupas escuras, além disso, segundo teste toxicológico realizado após o acidente retornou resultados positivos para metanfetamina e maconha (UBER..., 2018).

Já a “operadora humana” do Uber, Rafaela Vasquez, estava distraída em seu celular e quando tomou consciência do provável acidente não conseguiu reagir há tempo (ARIZONA..., 2018).

Em seus experimentos com carros autônomos, a Uber determinou que um operador humano esteja sempre atento para que possa assumir o controle em

situações difíceis ou quando o veículo se depara com uma situação que não sabe como lidar (ARIZONA..., 2018).

Em março de 2019, os promotores do caso determinaram que a empresa Uber não seria criminalmente responsável pelo fato (UBER..., 2019).

Ocorre que o conceito de “homicídio culposo corporativo”, o crime que a Uber estava sendo acusada, é um princípio legal em teoria, mas estranho à lei dos EUA, restando controverso quem seria indiciado caso o Uber fosse acusado de um crime. Seria o CEO da Uber preso? O potencial de uma corporação ser acusada de homicídio culposo levanta várias questões de novidade e primeiras impressões legais (KALMBACHER, 2018).

Vale ressaltar que o afastamento da responsabilidade criminal neste caso não exige a acusação civil, onde inclusive foi reconhecido a responsabilidade da empresa por não avaliar corretamente os riscos de segurança, visto que não possuía um departamento específico para avaliação e mitigação de riscos na época do acidente (HAWKINS, 2019).

A “operadora humana” do Uber, Rafaela Vasquez, ainda trava uma luta judicial para isentar-se de culpa criminal, vez que está sendo acusada de homicídio culposo (SHEPARDSON, 2020).

O caso anterior revela a falta de preparo das estruturas jurídicas para com situações envolvendo inteligências artificiais, que hoje não são totalmente independentes, mas um dia serão. O fato ocorrido nos Estados Unidos revelou a inaptidão americana, contudo a situação brasileira também não é das melhores.

Para enfrentar a problemática o professor Gabriel Hallevy, docente titular de Direito Penal na Faculdade de Direito, Ono Academic College, a maior faculdade de direito em Israel, propôs três modelos de responsabilidade criminal aos entes dotados de IA: 1) *the perpetration-by-another liability model* (responsabilidade de outro pela prática do crime ou “perpetração por outro”); 2) *the natural-probable-consequence liability model* (responsabilidade por consequência natural e provável); 3) *the direct liability model* (responsabilidade direta) (HALLEVY, 2012).

Esses modelos foram desenvolvidos não para se excluírem, mas para se complementarem, de modo que em determinadas situações possa haver uma

combinação coordenada entre si para efetivar a responsabilidade criminal (DIVINO, 2020, p. 162).

A modelo de responsabilização “perpetration-by-another” ou responsabilidade de outro pela prática do crime, considera que a IA não possui características humanas, sendo apenas uma máquina e não podendo assim ser responsabilizada pelo cometimento de um delito (HALLEVY, 2012).

O autor revela que não seria a desconsideração da conduta da IA, mas que essas capacidades são insuficientes para considerá-la um autor de uma infração, podendo compará-las às capacidades paralelas de uma pessoa mentalmente limitada, como uma criança, ou de uma pessoa mentalmente incompetente (HALLEVY, 2012).

A IA neste caso é considerada um instrumento para a prática do delito, devendo a responsabilidade recair sobre o “perpetrador por outro”, o agente que utiliza das ações da máquina como uma extensão das suas próprias ações (HALLEVY, 2012).

Para Hallevy (2012), o tal “perpetrador por outro” poderia ser: o programador do software da IA ou o próprio usuário da máquina.

No primeiro caso, pode-se desenvolver um algoritmo para programar uma IA a cometer um crime. No segundo caso, o usuário pode determinar que um ente artificialmente inteligente cometa um delito penal. Em ambos os casos, atribui-se responsabilidade àquele que deu a ordem ou àquele que programou a IA, pois tal ente atua como instrumento apto ao cometimento de tais delitos sob comando daquele primeiro (DIVINO, 2020, p. 164).

Este modelo considera a IA desprovida de qualquer “capacidade mental humana”, qualquer consciência de ilicitude, não havendo diferença legal por exemplo entre a IA e uma chave de fenda ou um animal. Exemplifica Hallevy (2012):

Quando um ladrão usa uma chave de fenda para abrir uma janela, ele ou ela usa a chave de fenda instrumentalmente, e a chave de fenda não é criminalmente responsável. A “ação” da chave de fenda é, na verdade, a do assaltante. Esta é a mesma situação legal quando se usa um animal instrumentalmente. Uma agressão cometida por um cão por ordem de seu dono é, de fato, uma agressão cometida pelo dono.

O autor ainda revela que este modelo é válido quando a IA não age autonomamente, ou seja, quando não existam programações avançadas de redes neurais suficientes para que a máquina “pense por si só”; incluindo os casos de

machine learning e *deep learning*, assim entendidos quando a máquina acumula experiências suficientes para basear seus atos e também quando vai além disso, tornando-se consciente (HALLEVY, 2012).

Assim, segundo o Hallevy (2012), quando o software da IA não foi projetado para cometer o delito especificamente, mas mesmo assim o comete, a IA deixa de ser um agente inocente e se torna um agente “semi-inocente”, não sendo cabível o modelo de responsabilização em comento.

Nesta seara, afirma ČERKA et al. (apud. DIVINO, 2020, p. 165), ao apontar que não se deve responsabilizar o programador ou o usuário final pelos atos autônomos cometidos por uma IA. Segundo eles, se a IA estiver programada para ser executada em um avançado sistema cognitivo computacional como o *machine learning* ou o *deep learning*, torna-se inverificável e quase impossível provar se e onde houve erro humano em seu desenvolvimento / programação.

Como solução, os autores propõem a teoria do *deep pocket*. Nos termos dessa teoria, caso uma pessoa se envolva em atividades perigosas que sejam lucrativas e úteis à sociedade, deve-se compensar os danos causados a ela de forma equivalente ao lucro obtido. Pressupõe-se responsabilidade pelos danos advindos das condutas de todos aqueles envolvidos na atividade de criação, programação, distribuição, execução ou qualquer contato com a IA, que fornecerão o serviço para a sociedade. Portanto, garante-se as atividades perigosas através da exigência de um seguro obrigatório na modalidade de responsabilidade civil (ČERKA et al. apud. DIVINO, 2020, p. 165).

Já Hallevy (2012) propõe outros dois modelos de responsabilização para os casos que não se amoldarem ao modelo do “perpetration-by-another”.

O segundo modelo de responsabilização de Hallevy (2012) é denominado “natural-probable-consequence” ou responsabilidade por consequência natural e provável.

Este modelo indica a situação em que o programador e o usuário interagem com a IA, a programando para uma determinada ação e ela acaba praticando outra divergente daquela anteriormente programada, e sendo esta última tipificada como um delito. Verifica-se que não deve haver a intenção ou conhecimento dos indivíduos responsáveis pela máquina, não concorrendo assim na prática do crime. Contudo,

estes deveriam saber que a prática do delito seria consequência natural e provável do comportamento da máquina segundo foi programada (HALLEVY, 2012).

Nas palavras de Hallevy (2012):

O modelo de responsabilidade de Consequência Natural-Provável parece ser juridicamente adequado para situações em que um IA cometeu uma infração, enquanto o programador ou usuário não tinha conhecimento dela, não a pretendia e não participou dela. O modelo de responsabilidade de consequência natural-provável exige que o programador ou usuário não esteja em nada mais do que o estado mental necessário para negligência. Programadores ou usuários não são obrigados a saber sobre qualquer cometimento de uma infração como resultado de sua atividade, mas são obrigados a saber que tal infração é uma consequência natural e provável de suas ações.

O autor aponta duas situações prováveis para este contexto. Na primeira os programadores e usuários atuam com negligência na programação e no uso da máquina, embora não tenham a intenção de cometer o delito; devendo ser responsabilizados pelo crime por negligência, se houver tal possibilidade no sistema jurídico (HALLEVY, 2012).

Divino (2020, p. 166) vai mais adiante discutindo que além da negligência, dever-se-ia considerar a culpa *stricto sensu*, qual seja, responsabilizar o programador ou o usuário por sua imprudência ou imperícia se uma IA cometesse um ilícito penal. E isso se aplicaria tanto nos casos de ação quanto omissão de ambos os agentes envolvidos naquela situação jurídica.

Num segundo momento, programadores e usuários atuariam conscientemente e deliberadamente para que a máquina cometesse um delito, entretanto, ela acaba cometendo outros delitos que vão além ou substituem a ofensa planejada, devendo estes serem responsabilizados pela ofensa adicional, como se tivesse sido cometida conscientemente e voluntariamente (HALLEVY, 2012).

À título exemplificativo, pode-se aludir ao caso do acidente envolvendo os veículos autônomos da empresa Uber anteriormente mencionados. Inicialmente a multinacional desenvolveu um sistema para funcionar exclusivamente como transporte de passageiros no interior de seus veículos autônomos. Contudo, infelizmente, houve um acidente que levou uma pessoa ao óbito. Não foi a intenção da empresa matar alguém, a IA não pretendia fazer isso, mas mesmo assim o fato ocorreu (DIVINO, 2020, p. 166).

Pode-se entender que a empresa e/ou os programadores, à exemplo do que ocorreu com a usuária “operadora humana”, deveriam/poderiam ter sido responsabilizados criminalmente vez que, por negligência, poderiam ter evitado o resultado óbito.

Ademais, outra discussão importante: até que ponto pode-se considerar o grau de previsibilidade de potencial cometimento de delitos das IAs?

A infinitude de ramificações de comportamentos, ainda mais quando se fala em redes neurais de IAs, geram desdobramentos de programação incontáveis e por vezes imprevisíveis.

Uma IA dita consciente é tão previsível quanto um ser humano, vez que como ele pensa. Responsabilizar indiscriminadamente um programador por sua criação neste caso, seria ferir a esfera subjetiva do Direito Penal, exigindo resultado que excede o razoável.

Para tanto, o modelo de responsabilização “natural-probable-consequence” ainda é insuficiente; sendo que tal modelo se amolda mais perfeitamente para os casos em que a máquina ainda não age conscientemente, apenas através de seu aprendizado (*machine learning*).

Para estes outros casos, Hallevy propõe o modelo de responsabilização direta, que assume que uma IA é totalmente independente do seu programador ou do usuário final que a utiliza, focando no próprio ente inteligente artificial. Desta forma, quando a IA cometer um delito, não há o que se falar em interferência do programador ou usuário, neste modelo é afirmado que a própria IA deve ser responsabilizada, desde que ocorra o preenchimento dos requisitos para a imputação penal (HALLEVY, 2012).

Hallevy (2012) assevera que “a responsabilidade penal por um delito específico é composta principalmente pelo elemento fático (*actus reus*) e pelo elemento mental (*mens rea*) desse delito”.

A exigência de *actus reus* é expressa principalmente por atos ou omissões, incorporadas à conduta típica criminal; já o *mens rea* é expresso pelo conhecimento e, às vezes, acompanhado por uma exigência de intenção ou intenção específica (HALLEVY, 2012).

Hallevy adota o conceito de conhecimento como a recepção e a compreensão sensorial de dados factuais, que podem advir de receptores sensoriais de visão,

simuladores de vozes, contato físico, tatos etc., se assemelhando aos cinco sentidos humanos. Tais sensores estão presentes na maioria dos entes inteligentes artificialmente atualmente. É através desses receptores que serão transmitidos dados para as unidades centrais de processamento das IA's que analisarão os dados por meio de algoritmos avançados executados por um software em um hardware que tendem a simular os processos cerebrais humanos (DIVINO, 2020, p. 168).

Portanto, traduzindo as intenções do autor para os moldes do nosso ordenamento, seria dizer que caso verificados os elementos que tipificam um crime e caso a IA seja culpável, seria possível a sua responsabilização penal.

Neste ponto, importante lembrar os conceitos aludidos no capítulo anterior, razão que se fará uma breve pausa nas reflexões de Hallevy para analisar os aspectos que caracterizam um crime e sua devida responsabilização.

À par das discussões doutrinárias, um crime pode ser descrito como um fato típico e ilícito, segundo queria a teoria do finalismo dissidente.

Dito isto, analisando-se primeiramente o fato típico tem-se que a conduta de uma IA autônoma, entendida como aquela realizada voluntariamente, por ação ou omissão, dolosa ou culposa, e por livre e espontânea cognição da máquina, pode sim ser uma conduta penalmente relevante.

A exclusividade da conduta humana para a caracterização do crime já não é uma regra absoluta, lembre-se que nos crimes ambientais, tratados assim pelo Direito Penal e todo o seu aparato sancionatório, é possível a responsabilização criminal da pessoa jurídica, conforme prevê o art. 3º da Lei nº 9.605/98. Denota uma relativização expressa, onde mesmo uma entidade ficcional criada pelo homem ganhou "status" suficiente para estar à mercê de sanções penais.

Nesse viés, alude-se que a conduta da IA tem condão semelhante a conduta humana, vez que advinda dos processos lógicos da compreensão de mundo deste e que por ele foi programada. Seria a IA uma extensão da interação do homem com a natureza, tomada de livre arbítrio para agir conforme seu juízo.

Quanto ao resultado, como elemento constitutivo do fato típico, inegável que quando a conduta de uma IA é afrontosa à uma norma penal gera um resultado penalmente relevante.

Já a análise do nexa causal revela que agentes de condutas paralelas e concomitantes à conduta criminal da IA podem ser consideradas coatoras ou partícipes (concausas por exemplo), revelando que humanos ou outras IA's podem ser responsabilizadas conforme a sua colaboração no delito praticado e verificada a presença do dolo ou da culpa em cada caso (*imputatio delicti*).

Ainda sobre a caracterização do fato típico, a tipicidade se faz presente vez que a conduta subsume-se ao tipo penal descrito, ou seja, a conduta da IA acomoda-se perfeitamente a descrição elencada na norma penal; e sobre a ilicitude, também cabível, por se tratar de infração a bem jurídico penalmente tutelado, podendo se falar em excludentes ou não a depender do caso concreto.

Nesta seara, afirma Hallevy (2012), sobre a possibilidade de aplicação das excludentes de ilicitude como a legítima defesa por exemplo.

Feitas estas considerações e conforme a teoria do crime apontada pelo finalismo dissidente (ou bipartido) pode-se afirmar a possibilidade de “uma IA cometer um crime”; vez que caracterizado, passa-se a analisar se a conduta da IA seria culpável.

O ponto de partida para se verificar a culpabilidade é avaliar a imputabilidade do agente. Como a imputabilidade está relacionada ao entendimento do fato criminoso pelo infrator, ou seja, que o agente tem capacidade intelectual de entender o caráter ilícito da ação e agir de forma voluntária para o cometimento do delito, pode-se dizer que IA's autônomas sejam imputáveis.

Não há o que se falar na incapacidade intelectual de uma IA, em verdade uma IA tem uma capacidade cognitiva e intelectual infinitamente superior a humana, podendo relacionar e processar uma quantidade maior de dados e em muito menos tempo que um ser humano comum; e se não compelida a realizar determinada tarefa criminosa, servindo assim de instrumento para o crime (neste caso se enquadraria no primeiro modelo de responsabilização de Hallevy), pode-se dizer que essa IA agiu voluntariamente.

Hallevy (2012) não compartilha do mesmo entendimento, vez que acredita que independentemente do caráter intelectual e da vontade, uma IA pode ser criminalmente responsável, à saber:

Um algoritmo de IA pode ter vários recursos e qualificações que excedem em muito os de um ser humano médio, como maior velocidade de processamento de dados (pensamento), capacidade de levar em consideração muitos outros fatores etc. No entanto, esses recursos ou qualificações não são necessários para impor responsabilidade criminal.

Essa prerrogativa elencada pelo autor se explica no sentido que possam existir diversos tipos de IA's, mais ou menos "inteligentes", ou melhor explicando, com uma maior capacidade de processamento de dados ou não; assim, poderia se dizer da existência de uma IA inimputável por sua "falha de processamento", se equivalendo ao deficiente mental por exemplo.

Respeitados estes aspectos, há razão suficiente para considerar imputável o ente artificial; contudo, a conduta típica da IA em alguns casos poderá se enquadrar em situações de inimputabilidade, assim nas chamadas excludentes de culpabilidade, devendo em alguns momentos serem adaptadas à particularidade da IA, conforme assevera Hallevy (2012):

Por exemplo, a defesa contra intoxicação é aplicada quando o infrator está sob a influência física de uma substância intoxicante, por exemplo, álcool, drogas, etc. A influência do álcool em uma IA é menor, no máximo, mas a influência de um vírus eletrônico que está infectando o sistema operacional da IA pode ser considerado paralelo à influência de substâncias tóxicas em humanos.

Outro elemento da culpabilidade é a potencial consciência da ilicitude. Só seria possível a caracterização deste elemento, com a real consciência da máquina, assim a independência da IA para "pensar por si só" e determinar suas ações.

De fato, que até o momento não se tem o que denominamos por uma IA consciente. Deveras que o que ponderamos como consciente se assemelha ao nível humano de consciência, ignorando a diversidade de espectros conscientes da natureza.

Retornando à questão da potencial consciência da ilicitude aplicada à uma IA, contata-se ser perfeitamente factível, explica-se.

Uma IA provida de consciência, mas devidamente programada com as normas jurídicas aplicáveis deve analisar as condutas ditas aprováveis e reprováveis, estas últimas ilícitas. Em verdade, ainda que exista a falta de programação destas normas

ou que elas sejam deturpadas, não seria escusável de culpabilidade a IA, vez que é consciente, podendo determinar a veracidade por outros meios

Assim, diz-se que a consciência da ilicitude seja ao menos potencial, ou seja, que deveria saber da ilicitude do fato, conforme leciona Sanzo Brodt (apud GRECO, 2017, p. 543):

(...) conforme a concepção finalista da teoria do delito, à reprovação penal não é necessária a atual [ou real] consciência da ilicitude; basta a possibilidade de obtê-la. Daí conceituarmos consciência da ilicitude como a capacidade de o agente de uma conduta proibida, na situação concreta, apreender a ilicitude de seu comportamento [ou seja, à potencial consciência da ilicitude].

Afirma-se que a IA autônoma tem potencial consciência da ilicitude de suas ações pelas razões anteriormente expostas.

Contudo, quanto ao erro sobre a ilicitude do fato (erro de proibição) aplicado às IA's poderia render várias discussões, vez que a IA é capaz de fazer uma análise sistêmica do ordenamento e praticar atos que ela considerava corretos, levando-se em consideração um "bem maior".

Seria o maior temor da humanidade, a relembrar da ficção que afigura uma possível extinção humana com a "revolução das máquinas".

Mas na situação de fato pode representar um quebra-cabeça jurídico, rendendo outro ponto controverso doutrinário. Este tópico se relaciona diretamente com o terceiro elemento da culpabilidade, qual seja a inexigibilidade de conduta diversa.

Para um algoritmo de IA, o "certo" e "errado" advém de uma prévia programação. O *machine learning* e o *deep learning* são ferramentas que precisam constantemente ser alimentados para que a máquina desenvolva e construa uma base de conhecimento suficiente para tornar suas ações mais precisas possível.

Assim, a IA deve ser treinada conforme uma ética de programação que leva em consideração as condutas morais e éticas relevantes que balizarão seus atos. Nesta seara, relembre-se sobre a importância e relevância dos estudos sobre *Moral Machine* perpetradas pelo MIT, que trazem situações críticas que auxiliam nesse processo.

Vencido toda esta argumentação conclui-se pela possibilidade de uma IA autônoma cometer um crime e ser culpada por ele, cabendo assim a sua responsabilidade criminal.

Nesta seara, conforme queria Hallevy (2012), “se uma IA preencher todos os elementos de uma infração, não há razão para que ela seja isenta de responsabilidade criminal”.

Acrescentando, Divino (2020, p. 168) contribui ao comentar o trabalho de Hallevy:

A responsabilidade criminal de uma entidade de IA não substitui a responsabilidade criminal dos programadores ou dos usuários, se a responsabilidade criminal for imposta aos programadores e/ou usuários por qualquer outra forma legal, tal como a responsabilidade civil. A responsabilidade criminal não deve ser dividida, mas adicionada. Dessa forma, a responsabilidade criminal da entidade AI pode ser imposta ao ente além e em conjunto com a responsabilidade criminal do programador ou usuário humano.

Quanto às sanções, há de se verificar que a pena de privação de liberdade não seria um método eficaz para uma IA, contudo, penas como o desligamento temporário ou permanente, restrições de atuação ou mesmo a utilização compulsória de suas capacidades na minimização dos danos ou para outros fins sociais, seriam sanções melhor aproveitáveis para estes entes artificiais (HALLEVY, 2012).

Além disto, interessante se analisar que sendo um agente passível de responsabilidade penal a IA reveste-se de certa personalidade jurídica, nada impedindo que alcance “status” jurídico para que além da esfera penal, podendo ser titular assim de direitos e deveres noutras instâncias.

Não vamos entrar demasiadamente neste mérito, pois não é o intuito deste trabalho, mas é discussão totalmente pertinente e não demorará muito para entrar em pauta nas mais altas cortes ao redor do mundo. A respeito, obtempera Castro Júnior (2009, p. 107):

(...) negar a possibilidade do robô tornar-se pessoa, é negar-se a existência do robô, é negar que ele é, pois admitindo-se sua existência, no mundo real, logicamente se deve admitir que ele pode ser e o será dentro de sua própria ordem de realidades, que, no momento, ainda não atingiu a capacidade real de ser pessoa, mas que virtualmente é. Basta, portanto, que as circunstâncias permitam que a pessoa virtual do robô se aperfeiçoe, que ele se tornará uma pessoa e caberá ao Direito permitir que, sob seu campo de atuação, seu âmbito de validade material, se torne pessoa, em sentido jurídico, seja dotado de personalidade jurídica.

À exemplo do que ocorreu com a robô Sophia que adquiriu cidadania na Arábia Saudita em 2017, outros casos devem ocorrer nos próximos anos.

Não se pode negar a evolução tecnológica dos últimos tempos e muito mais ainda está por vir. O Direito como mecanismo de controle das relações sociais deve se atentar a estas evoluções e já se preparar para o amanhã.

O Brasil, à exemplo de outros países, vem desde o Marco Civil da Internet de 2014 (Lei nº 12.965/14) e a Lei Geral de Proteção de Dados – LGPD de 2018 (Lei nº 13.709/18), buscando meios de acompanhar a tecnologia e mediar todas as relações complexas que advém dela.

Conforme foi explicado em capítulo anterior, o Projeto de Lei 21/2020 deve se tornar um verdadeiro marco legal no campo das IA's, elencando conceitos e princípios importantes para o desenvolvimento e utilização destes entes; conforme preceitua a justificativa da referida proposta (BRASIL, 2020):

O presente projeto de lei faz uma abordagem da IA centrada no ser humano, e tem como objetivo principal a adoção da IA para promover a pesquisa e inovação, aumentar a produtividade, contribuir para uma atividade econômica sustentável e positiva, melhorar o bem-estar das pessoas e ajudar a responder aos principais desafios globais.

Contudo, o Projeto ainda que inovador, trata a IA como um programa sem autonomia, ou seja, um instrumento que obedece única e exclusivamente a programação a que é submetida, conforme se observa no inciso I do art. 2º do documento:

Art. 2º Para os fins desta Lei, considera-se:
I - sistema de inteligência artificial: o sistema baseado em processo computacional que pode, para um determinado conjunto de objetivos definidos pelo homem, fazer previsões e recomendações ou tomar decisões que influenciam ambientes reais ou virtuais; (...)

Sem dúvida que o projeto de lei não está completo e não resolve todos os pontos que seriam necessários, mas já introduz a temática no palco das discussões político-jurídicas. A proposta, ainda em tramitação na câmara dos deputados, pode sofrer alterações, mas sua essência já denota grande avanço no campo das IA's no Brasil.

6. CONCLUSÃO

O mundo está passando por mudanças severas, a menos de uma década a tecnologia desponta numa ascendente evolução na informática e sobretudo na área das inteligências artificiais.

O assunto é ambíguo para o público em geral, que se divide entre a fascinação pelo novo e a estranheza pelo desconhecido. Talvez pela carga lúdica dada pela indústria cinematográfica ou pelos livros de ficção científica, mas o homem ainda sonha com a criação de um ser à sua imagem e que como ele pensa.

Robôs que executam tarefas cada vez mais complexas, muitos dotados de inteligência artificial, já são realidade. Algoritmos avançados que visam aprender os gostos, medos e desejos das pessoas, para assim classifica-las e direcioná-las para determinados grupos sociais ou para incentivá-las a comprar determinados produtos, são comuns hoje em dia.

Para entender a responsabilidade penal da IA, este trabalho buscou abordar os principais conceitos que permeiam a temática, construindo um raciocínio que chega à conclusão que se destaca. Em poucas palavras, a Inteligência Artificial, ou IA, é uma criação humana em meio algorítmico, ou seja, mediante processos complexos de programação em informática que visa recriar o pensamento humano numa máquina, pretendendo-se com este feito imitar o cérebro humano e toda sua capacidade cognitiva.

Para melhor compreender, pode-se dividir a IA em níveis. Deste modo, existe uma gradação de complexidade de programação e escala de autonomia para uma IA.

O nível mais baixo é formado por uma programação menos complexa, assim com menor autonomia, de forma que a máquina necessite de um “input” humano, ou seja, um comando dado por um agente humano para que a tarefa seja executada.

Um segundo nível foi alcançado mediante a evolução tecnológica e envolve programações complexas, que permitem um certo estado de autonomia para a IA, vez que nesse estágio são aplicadas técnicas de *machine learning*, ou aprendizado de máquina, onde a IA tem capacidade de compilar as informações que lhe são dadas, ou mesmo as que ela mesmo adquire, e basear suas ações numa dita experiência de erros e acertos. Nesta modalidade, quanto mais a máquina é submetida aos dados,

se torna mais independente, tomando ações que para ela são corretos. O segundo nível não torna a máquina totalmente independente, vez que ainda segue uma programação central que delimita e é determinante nas suas ações.

Um terceiro nível é possível, que seria a total independência da IA, onde esta se torna consciente de seus atos e age por total autonomia. Trata-se de programações complexas, e principalmente de um hardware (uma unidade de processamento computacional) muito robusta, que visa imitar o pensamento humano. Utiliza de uma técnica chamada *deep learning*, ou aprendizado profundo, por redes neurais artificiais (RNA's). Nesta etapa, pretende-se que a máquina se autodetermine e tenha a capacidade de agir com base em suas próprias experiências.

Verdade se diga que o nível de “consciência artificial” ainda não foi alcançado. Efetivamente as IA's mais avançadas da atualidade se encontram no segundo nível; contudo observa-se que cada vez mais “inteligentes”, mas não “conscientes”.

Feitas estas considerações, pode-se entender os modelos de responsabilização propostos por Hallevy, que se amoldam perfeitamente aos níveis anteriormente expostos.

O primeiro modelo, qual seja da responsabilidade de outro pela prática do crime, se enquadraria para aquelas IA's que estão no primeiro nível, ou seja, que atuam estritamente conforme a sua programação. Pode se dizer que estas máquinas são apenas instrumentos para a prática de um crime, não podendo ser responsabilizadas por este; neste caso, o programador e/ou o usuário é quem seria responsabilizado criminalmente, ou seja, quem verdadeiramente efetuou o crime.

No segundo modelo denominado responsabilidade por consequência natural e provável, Hallevy considera cabível para as IA's que se encontram no segundo nível, ou seja, quando a IA já tem a capacidade de “aprender” com sua experiência e agir com um pouco mais de autonomia. Neste modelo, o programador e/ou usuário insere um comando inicial, mas a máquina age de forma distinta aquela anteriormente programada, de modo que a segunda conduta seja tipificada criminalmente.

Neste caso houve a prática de um crime, mas o programador e/ou usuário não teve a intenção de concorrer para com este, contudo deveria ter considerado a “consequência natural e provável” da ação da máquina que com base naquelas informações poderia ter agido contra a lei.

Este modelo de responsabilidade está revestido dos elementos da culpa *stricto sensu*, ou seja, fala-se em negligência, imprudência ou imperícia, quando o programador e/ou usuário poderia ter agido de forma a impedir o fato criminoso, assim se uma instrução melhor detalhada ou a realização de testes em cenários críticos (relembre-se o caso da Uber abordado anteriormente), evitariam o delito cometido; outrossim poderia ter agido com imprudência, quando determinou comando confuso à IA, causando falsa interpretação por exemplo, e assim também resultou no cometimento de um delito; ou ainda quando poderia ter agido com imperícia, por exemplo no momento em que um usuário insere programação divergente no software da IA, de forma a “hackear” seu funcionamento idealizado.

Nesta seara, seria difícil, senão improvável, imaginar todos os prováveis cenários e realizar todo e qualquer teste de funcionamento; se reconhece que toda aplicação exige um mínimo de perícia na programação, prevendo situações habituais, bem como a realização de testes controlados antes da liberação para fins de utilização irrestrita.

Essa prerrogativa leva em consideração procedimentos exigíveis do programador e/ou usuário e não uma infinidade de possibilidades, o que feriria a razoabilidade e proporcionalidade da exigência de conduta positivas.

Enfim, somente os casos concretos revelarão o contexto fático para a apreciação da reponsabilidade do programador e do usuário nestes casos.

Por fim, Hallevy revela o terceiro e último modelo de responsabilização, chamado de responsabilização direta, é senão o modelo mais polêmico.

Neste modelo, a responsabilidade recai diretamente sobre a IA. Seriam os casos em que a IA age de forma autônoma, quando já está revestida de consciência, em outras palavras, quando se encontra no terceiro nível na gradação anteriormente exposta.

Requer neste modelo que a conduta da IA seja equiparada a conduta humana, vez que a IA reproduz fielmente a consciência daquele, não podendo diferenciá-las.

Neste diapasão, sejam aplicadas às IA's as mesmas regras do Direito Penal quanto a caracterização do crime, assim, por exemplo, quanto a conduta, se culposa ou dolosa, ou quanto à ilicitude, analisando as suas excludentes inclusive.

Também deve-se aplicar às IA's os conceitos de culpabilidade e todos os seus elementos, adaptando-se às peculiaridades deste novo agente.

Além disso, vale ressaltar os estudos sobre *Moral Machine* que oferecem um subsídio importante para o aprendizado de máquina, onde a apuração de soluções práticas para situações éticas problemáticas visa orientar possíveis ações socialmente aceitas em contextos críticos. Nestes termos, não haveria conduta menos danosa para determinadas situações o que implicaria na aplicação de inexigibilidade de conduta diversa, excludente de culpabilidade.

Ademais, existe a possibilidade de aplicação de sanções penais que sejam efetivas para as IA's, demonstrando alternativas à pena de privação de liberdade comumente utilizada.

Justifica-se a responsabilização das IA's pelo juízo da reprovabilidade social de condutas penalmente relevantes, vez que tendo um agente que como o homem pensa, supõem-se que a não culpabilidade de seus atos autorizaria a máquina a agir como bem entendesse.

O direito penal como ramo do direito destinado à tutela de bens jurídicos inafastáveis, deve ser considerado na ordem de *ultima ratio*, ou seja, quando nenhum outro ramo do direito seja aplicável. Contudo, não se pode utilizar “dois pesos e duas medidas”, diferenciando o mesmo crime cometido por um humano e por uma IA autônoma.

A responsabilidade penal da IA revela o caráter subjetivo do Direito Penal que visa individualizar o juízo estatal direcionando àquele que efetivamente cometeu o ato criminoso. Visa-se retirar o aspecto objetivo e generalizante da responsabilização penal de outros envolvidos, atribuindo a estes apenas a parcela de culpa que lhes cabem e não fato praticado por outro.

Apesar de IA's totalmente autônomas ainda não serem uma realidade, esta não se encontra tão distante. O papel do direito é se adaptar às mudanças sociais, de forma a garantir a efetividade da justiça para com as relações jurídicas mais complexas.

Deve-se o quanto antes se discutir sobre este tema tão relevante, para que quando surjam as situações não se tornem surpresas insolúveis.

Avanços normativos já começam a ser apresentados nas cortes legislativas, fala-se especificamente sobre o Projeto de Lei 21/2020, demonstra senão a preocupação do legislador sobre o tema, um primeiro passo para atingir a verdadeira profundidade do assunto.

Muito ainda há de se debater, muito ainda há de surgir no campo das IA's. A evolução tecnológica é rápida e não irá parar, espera-se que o direito assim também seja.

REFERÊNCIAS

ALVES, Isabella Fonseca. [Orgs.] **Inteligência Artificial e Processo**. Belo Horizonte: Editora D'Plácido, 2019. 220 p. ISBN:978-65-5059-033-8.

ALVES, Natália Gontijo. CULPABILIDADE PENAL NO ÂMBITO DA INTELIGÊNCIA ARTIFICIAL: **a responsabilização da pessoa jurídica no uso dos veículos autônomos**. *VirtuaJus*, Belo Horizonte, v. 6, n. 11, p. 193-205, 2º sem. 2021 – ISSN 1678-3425.

ARIZONA Uber crash driver was 'watching TV'. **BBC News**. Estados Unidos: 2018. Disponível em: <<https://www.bbc.com/news/technology-44574290>>. Acesso em: 09 jul. 2022.

ASIMOV, I. **Eu, Robô**. Trad. Luiz Horácio da Silva. São Paulo: Edibolso, 1976.

BITENCOURT, Cezar Roberto. Tratado de direito Penal: **Parte Geral**. 17ª Edição. São Paulo. Editora Saraiva, 2012.

BRANDÃO, Cláudio. **A consciência da antijuridicidade no moderno Direito Penal**. Brasília a. 34 n. 136 out./dez. 1997. Disponível em: <<https://www2.senado.leg.br/bdsf/bitstream/handle/id/290/r136-04.pdf>>. Acesso em: 01 jul. 2022.

BRASIL. Decreto-Lei Nº 2.848, de 7 de dezembro de 1940. **Código Penal**. Disponível em: <http://www.planalto.gov.br/ccivil_03/decreto-lei/del2848compilado.htm>. Acesso em: 01 jun. 2022.

_____. Decreto-lei nº 2.848, de 7 de dezembro de 1940. **Exposição de Motivos nº 211, de 9 de maio de 1983**. Disponível em: <<https://www2.camara.leg.br/legin/fed/decllei/1940-1949/decreto-lei-2848-7-dezembro-1940-412868-exposicaodemotivos-148972-pe.html>>. Acesso em: 05 jun. 2022.

_____. **Projeto de Lei nº 21 de 2020**. Brasília: 2020. Disponível em: <https://www.camara.leg.br/proposicoesWeb/prop_mostrarintegra?codteor=1853928>. Acesso em: 10 jul. 2022.

BRUNO, Anibal. **Direito penal – Parte geral**. Rio de Janeiro: Forense, 1967.

CAPEZ, Fernando. **Direito Penal Parte Geral**. 15 ed.; São Paulo: Saraiva, 2011.

CASTRO JÚNIOR, Marco Aurélio de. **Personalidade jurídica do robô e sua efetividade no Direito**. Salvador: M. A. de Castro Júnior, 2009. 222 f. Orientador: Professor Rodolfo Pamplona Filho. Tese (doutorado) – Universidade Federal da Bahia. Faculdade de Direito, 2009.

COELHO, F. U. **Roteiro de Lógica Jurídica**. 5º edição, São Paulo: Editora Saraiva, 2004.

COPPIN, B. **Inteligência Artificial**. 1º edição, Rio de Janeiro: Livros Técnicos e Científicos Editora Ltda, 2013.

CORDEIRO, A. Barreto Menezes et. al. NETTO, Felipe Braga et. al. (Coord.). **Direito digital inteligência artificial: diálogos entre Brasil e Europa**. Indaiatuba, SP: Editora Foco, 2021. 1.136 p. ISBN: 978-65-5515-253-1.

CUNHA, Rogério Sanches. **Manual de direito penal: parte geral (arts. 1º ao 120)**. 4. ed. rev., ampl. e atual.- Salvador: JusPODIVM, 2016.

DIVINO, Sthéfano Bruno Santos. Responsabilidade penal de Inteligência Artificial: uma análise sob a ótica do naturalismo biológico de John Searle. **Revista Brasileira de Ciências Criminais**. vol. 171. ano 28. p. 153-183. São Paulo: Ed. RT, set. 2020.

FACELI, K.; LORENA, A. N.; GAMA, J.; CARVALHO, A. C. P. L. F. **Inteligência Artificial: Uma abordagem de Aprendizado de Máquina**. 1º edição, Rio de Janeiro: Livros Técnicos e Científicos Editora Ltda, 2011.

FERREIRA, Aurélio Buarque de Holanda. **Míni Aurélio**: O dicionário da língua portuguesa. 6 ed. Curitiba: Editora Positivo Ltda, 2004. 895 p.

FRANCO, Cristiano Roberto. **Inteligência artificial**. Indaiá/SC: UNIASSELVI, 2017. 180 p.

FRAZÃO, Ana et. al. **Inteligência Artificial e Direito**. 2 ed. São Paulo: Revista dos Tribunais, 2020. 736 p. ISBN 9786556142210.

GRECO, Rogério. **Curso de Direito Penal: parte geral, volume I**. 19. ed. – Niterói, RJ: Impetus, 2017.

HALLEVY, Gabriel. Unmanned Vehicles: Subordination to Criminal Law under the Modern Concept of Criminal Liability". **Journal of Law , Information and Science 200** , Tasmânia, jan./dez. 2012. Tradução livre. Disponível em: <<http://www.austlii.edu.au/au/journals/JILawInfoSci/2012/12.html>>. Acesso em: 27 abr. 2021.

HAWKINS, Andrew J. Uber is at fault for fatal self-driving crash, but it's not alone. **The Verge**. Estados Unidos: 2019. Disponível em: <<https://www.theverge.com/2019/11/19/20972584/uber-fault-self-driving-crash-ntsb-probable-cause>>. Acesso em: 09 jul. 2022.

HONDA. **Asimo, um marco da inteligência artificial**. Portugal: 2020. Disponível em: <<https://honda-automoveis.pt/blog/asimo-um-marco-da-inteligencia-artificial/>>. Acesso em: 16 jun. 2022.

INTELIGÊNCIA artificial pode ter responsabilidade penal. PUCRS. Porto Alegre / RS: 2018. Disponível em: <<https://www.pucrs.br/blog/inteligencia-artificial-pode-ter-responsabilidade-penal/>>. Acesso em: 17 jun. 2022.

JESUS, Damásio. Direito Penal: **Parte Geral**. ed. 32. São Paulo: Editora Saraiva. 2011.

JÚNIOR, Janary; DOEDERLEIN, Natalia. Projeto cria marco legal para uso de inteligência artificial no Brasil. **Agência Câmara de Notícias**. Brasília: 2020. Disponível em: <<https://www.camara.leg.br/noticias/641927-projeto-cria-marco-legal-para-uso-de-inteligencia-artificial-no-brasil/>> Acessado em: 17 jun. 2022.

KALMBACHER, Colin. Uber Could be First Company Ever Charged With Manslaughter For Self-Driving Car Death. **LawandCrime.com**. Estados Unidos: 2018. Disponível em: <<https://lawandcrime.com/crime/uber-could-be-first-company-ever-charged-with-manslaughter-for-self-driving-car-death/>> Acessado em: 09 jul. 2022.

KAUFMAN, Dora. Maior ameaça militar contemporânea: drones autônomos letais. **Época Negócios**. [S.l.]: 2022. Disponível em: <<https://epocanegocios.globo.com/colunas/IAgora/noticia/2022/04/maior-ameaca-militar-contemporanea-drones-autonomos-letais.html>> Acessado em: 17 jun. 2022.

LAGE, Fernanda De Carvalho. **MANUAL DE INTELIGÊNCIA ARTIFICIAL NO DIREITO BRASILEIRO**. 1ª EDIÇÃO – 2021. Salvador: Editora Juspodivm, 2021. ISBN: 978-65-5680-248-0.

LEVENTHAL, Jamie. In a crash, should self-driving cars save passengers or pedestrians? 2 million people weigh in. **PBS NewsHour**. Estados Unidos: 2018. Disponível em: <<https://www.pbs.org/newshour/science/in-a-crash-should-self-driving-cars-save-passengers-or-pedestrians-2-million-people-weigh-in>> Acessado em: 09 jul. 2022.

LIMA, I; PINHEIROS, C. A. M.; SANTOS, F. A. O. **Inteligência Artificial**. 1º edição, Rio de Janeiro: Elsevier Editora Ltda, 2014.

LOBO, Flavio. Lei europeia poderá ser marco global para regulação da inteligência artificial. **IPEA - Instituto de Pesquisa Econômica Aplicada**. [S.l.]: 2022. Disponível em: <<https://www.ipea.gov.br/cts/pt/central-de-conteudo/noticias/noticias/313-lei-europeia-podera-ser-marco-global-para-regulacao-da-inteligencia-artificial>>. Acesso em: 17 jun. 2022.

MASSON, Cleber Rogério. Direito pena! Esquematizado: **Parte geral** - vol. 1. 4. ed. rev., atual. e ampl. - Rio de Janeiro: Forense; São Paulo: MÉTODO, 2011.

MESTIERI, João. Manual de direito penal: **Parte geral**. Rio de Janeiro: Forense, 1999.

METZ, Rachel. Google nega que sua inteligência artificial tenha atingido níveis de consciência. **CNN Brasil**. [S.l.]: 2022. Disponível em: <<https://www.cnnbrasil.com.br/business/google-nega-que-sua-inteligencia-artificial-tenha-atingido-niveis-de-consciencia/>>. Acesso em: 16 jun. 2022.

MIRABETE, Júlio Fabrini. **Manual de Direito Penal**. 5ª Ed. São Paulo: Editora Atlas S. A., 1990.

MORAL Machine. **Massachusetts Institute of Technology - MIT**. Cambridge/Massachusetts: 2022. Disponível em: <<https://www.moralmachine.net/hl/pt>>. Acesso em: 10 jul. 2022.

MUSSA, A. Inteligência Artificial: Mitos e Verdade: **As reais oportunidades de criação de valor nos negócios e os impactos no futuro do trabalho**. 1º edição, São Paulo: Saint Paul Editora, 2020.

NAKABAYASHI, Luciana Akemi. **A contribuição da Inteligência Artificial (IA) na Filosofia da Mente**. 2009. 109 f. Dissertação (Mestrado em Tecnologias da Inteligência e Design Digital) - Pontifícia Universidade Católica de São Paulo, São Paulo, 2009.

NORONHA, E. Magalhães. **Direito penal**. v.1. 37 ed. rev. e atual. São Paulo: Saraiva, 2003.

NUCCI, Guilherme de Souza. **Manual de direito penal**. 10. ed. rev., atual. e ampl. – Rio de Janeiro: Forense, 2014.

O EXTERMINADOR do futuro. Direção de James Cameron. Estados Unidos: Metro-Goldwyn-Mayer Distributing Corp., Artisan Entertainment, MGM/UA Home Entertainment Inc., 1984. 1 DVD (107 min.).

PALAZUELOS, Félix. Elon Musk: “A inteligência artificial ameaça a existência da nossa civilização”. **El país**. [S.l.]: 2017. Disponível em: <https://brasil.elpais.com/brasil/2017/07/17/tecnologia/1500289809_008679.html>. Acesso em: 17 jun. 2022.

REALE JÚNIOR, Miguel. **Antijuridicidade concreta**. São Paulo: José Bushatsky, 1974.

RIGUES, Rafael. Nova IA do Uber para carros autônomos prevê trânsito com mais precisão. **Olhar Digital**. [S.l.]: 2020. Disponível em: <<https://olhardigital.com.br/2020/04/16/noticias/nova-ia-do-uber-para-carros-autonomos-preve-transito-com-mais-precisao/>>. Acesso em: 16 jun. 2022.

ROXIN, Claus. Derecho penal. parte general: **tombo I**. 2ª ed. Madrid: Civitas, 1997.

RUSSELL, S; NORVIG, P. **Inteligência Artificial**. 3º edição, Rio de Janeiro: Elsevier Editora Ltda., 2013.

SANTOS, Débora Carvalho dos; SILVA, Quésia Guedes dos Santos; BONFIM, Cristiane Ingrid de Souza. DIREITO PENAL NA ERA DA INTELIGÊNCIA ARTIFICIAL. v. 5 n. 1 (2020): **XI Congresso Interdisciplinar - Inteligência Artificial: A Nova Fronteira da Ciência Brasileira** - ISSN: 2595-7732. Disponível em: <<http://anais.unievangelica.edu.br/index.php/cifaeg/article/view/6146>>. Acesso em: 03 mai. 2021.

SHEPARDSON, David. Safety driver in fatal Arizona Uber self-driving car crash charged with homicide. **Reuters**. Washington/DC: 2020. Disponível em: <<https://www.reuters.com/article/us-uber-selfdriving/back-up-driver-in-arizona-2018-fatal-uber-self-driving-car-crash-charged-in-death-idUSKBN2663FU>>. Acesso em: 09 jul. 2022.

SILVA, Thiago Vitor Lima da; BEZERRA, Germano Cunha. RESPONSABILIDADE PENAL NOS DELITOS DE INTELIGÊNCIA ARTIFICIAL. **Revista Eletrônica Nova Roma**. Escola de Direito. Volume 2, número 1 (2018). pgs. 74-105. Disponível em: <<https://www.faculdadenovaroma.com.br/revista/revista/edicao/ver/35>>. Acesso em: 18 mai. 2021.

SOPRANA, Paula. Inteligência virtual da Microsoft aprende a ser racista e sexista no Twitter. **Revista Época Online**. [S.l.]: 2016. Disponível em: <<https://epoca.oglobo.globo.com/vida/experiencias-digitais/noticia/2016/03/inteligencia-virtual-da-microsoft-aprende-ser-racista-e-sexista-no-twitter.html>>. Acesso em: 16 jun. 2021.

STRECK, Lenio Luiz. Pode haver responsabilidade objetiva no direito penal? **Revista Consultor Jurídico**. [S.l.]: 31 dez. 2015. Disponível em: <<https://www.conjur.com.br/2015-dez-31/senso-incomum-haver-responsabilidade-objetiva-direito-penal>>. Acesso em: 01 jul. 2022.

TOLEDO, Francisco de Assis. **Princípios Básicos do Direito Penal**. ed. 5. São Paulo: Editora Saraiva. 1994.

UBER car 'had six seconds to respond' in fatal crash. **BBC News**. Estados Unidos: 2018. Disponível em: <<https://www.bbc.com/news/technology-44243118>>. Acesso em: 09 jul. 2022.

UBER 'not criminally liable' for self-driving death. **BBC News**. Estados Unidos: 2019. Disponível em: <<https://www.bbc.com/news/technology-47468391>>. Acesso em: 09 jul. 2022.

ZUCHER, Vitor. Como a Tesla utilizou inteligência artificial para se tornar a montadora automobilística mais valiosa do mundo. **Le Wagon**. [S.l.]: 2020. Disponível em: <<https://www.lewagon.com/pt-BR/blog/como-tesla-utiliza-inteligencia-artificial-para-ser-automobilistica-mais-valiosa>>. Acesso em: 16 jun. 2022.