

**XXXII CONGRESSO NACIONAL DO  
CONPEDI SÃO PAULO - SP**

**DIREITO PENAL, PROCESSO PENAL E  
CONSTITUIÇÃO I**

**GUSTAVO NORONHA DE AVILA**

**ROGERIO LUIZ NERY DA SILVA**

Todos os direitos reservados e protegidos. Nenhuma parte destes anais poderá ser reproduzida ou transmitida sejam quais forem os meios empregados sem prévia autorização dos editores.

#### **Diretoria - CONPEDI**

**Presidente** - Profa. Dra. Samyra Haydée Dal Farra Naspolini - FMU - São Paulo

**Diretor Executivo** - Prof. Dr. Orides Mezzaroba - UFSC - Santa Catarina

**Vice-presidente Norte** - Prof. Dr. Jean Carlos Dias - Cesupa - Pará

**Vice-presidente Centro-Oeste** - Prof. Dr. José Querino Tavares Neto - UFG - Goiás

**Vice-presidente Sul** - Prof. Dr. Leonel Severo Rocha - Unisinos - Rio Grande do Sul

**Vice-presidente Sudeste** - Profa. Dra. Rosângela Lunardelli Cavallazzi - UFRJ/PUCRio - Rio de Janeiro

**Vice-presidente Nordeste** - Prof. Dr. Raymundo Juliano Feitosa - UNICAP - Pernambuco

**Representante Discente:** Prof. Dr. Abner da Silva Jaques - UPM/UNIGRAN - Mato Grosso do Sul

#### **Conselho Fiscal:**

Prof. Dr. José Filomeno de Moraes Filho - UFMA - Maranhão

Prof. Dr. Caio Augusto Souza Lara - SKEMA/ESDHC/UFMG - Minas Gerais

Prof. Dr. Valter Moura do Carmo - UFERSA - Rio Grande do Norte

Prof. Dr. Fernando Passos - UNIARA - São Paulo

Prof. Dr. Edinilson Donisete Machado - UNIVEM/UENP - São Paulo

#### **Secretarias**

##### **Relações Institucionais:**

Prof. Dra. Claudia Maria Barbosa - PUCPR - Paraná

Prof. Dr. Heron José de Santana Gordilho - UFBA - Bahia

Profa. Dra. Daniela Marques de Moraes - UNB - Distrito Federal

##### **Comunicação:**

Prof. Dr. Robison Tramontina - UNOESC - Santa Catarina

Prof. Dr. Liton Lanes Pilau Sobrinho - UPF/Univali - Rio Grande do Sul

Prof. Dr. Lucas Gonçalves da Silva - UFS - Sergipe

##### **Relações Internacionais para o Continente Americano:**

Prof. Dr. Jerônimo Siqueira Tybusch - UFSM - Rio Grande do Sul

Prof. Dr. Paulo Roberto Barbosa Ramos - UFMA - Maranhão

Prof. Dr. Felipe Chiarello de Souza Pinto - UPM - São Paulo

##### **Relações Internacionais para os demais Continentes:**

Profa. Dra. Gina Vidal Marcilio Pompeu - UNIFOR - Ceará

Profa. Dra. Sandra Regina Martini - UNIRITTER / UFRGS - Rio Grande do Sul

Profa. Dra. Maria Claudia da Silva Antunes de Souza - UNIVALI - Santa Catarina

##### **Educação Jurídica**

Profa. Dra. Viviane Coêlho de Séllos Knoerr - Unicuritiba - PR

Prof. Dr. Rubens Beçak - USP - SP

Profa. Dra. Livia Gaigher Bosio Campello - UFMS - MS

##### **Eventos:**

Prof. Dr. Yuri Nathan da Costa Lannes - FDF - São Paulo

Profa. Dra. Norma Sueli Padilha - UFSC - Santa Catarina

Prof. Dr. Juraci Mourão Lopes Filho - UNICHRISTUS - Ceará

##### **Comissão Especial**

Prof. Dr. João Marcelo de Lima Assafim - UFRJ - RJ

Profa. Dra. Maria Creusa De Araújo Borges - UFPB - PB

Prof. Dr. Antônio Carlos Diniz Murta - Fumec - MG

Prof. Dr. Rogério Borba - UNIFACVEST - SC

D597

Direito penal, processo penal e constituição I[Recurso eletrônico on-line] organização CONPEDI

Coordenadores: Gustavo Noronha de Avila, Rogerio Luiz Nery Da Silva – Florianópolis: CONPEDI, 2025.

Inclui bibliografia

ISBN: 978-65-5274-319-0

Modo de acesso: [www.conpedi.org.br](http://www.conpedi.org.br) em publicações

Tema: Os Caminhos Da Internacionalização E O Futuro Do Direito

1. Direito – Estudo e ensino (Pós-graduação) – Encontros Nacionais. 2. Direito penal. 3. Processo penal. XXXII Congresso Nacional do CONPEDI São Paulo - SP (4: 2025: Florianópolis, Brasil).

CDU: 34

# XXXII CONGRESSO NACIONAL DO CONPEDI SÃO PAULO - SP

## DIREITO PENAL, PROCESSO PENAL E CONSTITUIÇÃO I

---

### **Apresentação**

Tivemos a oportunidade de coordenar a apresentação dos excelentes trabalhos do GT Direito Penal, Processo Penal e Constituição I. Novamente, foi possível identificar o estado da arte da dogmática penal sendo discutidos por pesquisadores de Norte a Sul do país.

Inicialmente, Beatriz Azevedo e Giovanna Souza apresentaram texto sobre crimes de resultado e imputação objetiva no caso do Boeing 737 Max. A partir da categoria dos riscos proibidos, presente na referida teoria, defendem a possibilidade da imputação objetiva ainda em que atividades remotas, especialmente em contextos corporativos.

Sebastian Mello e Beatriz Azevedo discutiram os relatórios de inteligência financeira do COAF (Conselho de Controle de Atividades Financeiras) e a (i)legalidade de sua utilização. São trabalhadas a jurisprudência dos Tribunais Superiores, bem como a constitucionalidade de relatórios obtidos na informalidade.

O persistente tema da corrupção é discutido por Camila Costa e Sebastian Mello. Os autores trazem diferenciação entre as corrupções cotidianas e os esquemas de corrupção que normalmente ganham as manchetes midiáticas. São trazidas as diferenciações legais, além da discussão de casos paradigmáticos julgados no âmbito do Supremo Tribunal Federal.

As práticas laborais abusivas e sua criminalização, no âmbito internacional, são discutidas por Alexander Rodrigues de Castro, Pedro Henrique Facco, João Marcos Mariani Junior. São tratados, além do tema da política criminal, os reflexos das práticas no tocante aos direitos da personalidade e dos direitos humanos das vítimas.

A seguir, os mesmos autores, trabalham o atual tema do direito ao esquecimento, normalmente tratado de forma restrita ao direito constitucional, é analisado também em termos dos processos de criminalização. São identificados o direito à honra e intimidade como forma de prevenir futuros processos de estigmatização. Desta forma, está violada não apenas a dignidade humana do sujeito criminalizado, assim como a de seus familiares.

Tema também contemporâneo é o da lavagem de dinheiro e dos jogos de azar "online", analisado por Roberto Carvalho Veloso, Wendelson Pereira Pessoa e Monique Leray Costa. Os autores trabalham, em perspectiva comparada, com as regulamentações da Colômbia

(pioneira em normatizar a questão na América Latina) e a brasileira. Os autores defendem que, para além de regulação administrativa, é importante também a criminalização da conduta como forma de atenuar o problema.

O persistente problema do sistema prisional é discutido por Roberta Karina Cabral Kanzler , Wendelson Pereira Pessoa , Camila Kanzler Catunda da Silva. É debatida a questão da reinserção social enquanto (im)possibilidade de finalidade da pena, bem como trazida a teoria crítica da pena de Zaffaroni para o diálogo.

Os mesmos autores discutem o acordo de não persecução penal não apenas em termos dogmáticos, mas também na perspectiva político-criminal. Ao trabalhar o instituto, trazem o desenho legislativo previsto na Lei 13/964/2019 e problematizam a questão a partir do binômio eficiência x eficácia das garantias constitucionais fundamentais.

O trabalho, intitulado "PROCESSO DE CRIMINALIZAÇÃO QUATERNÁRIA: DADOS E REALIDADE SOBRE A (IN)TRANSCENDÊNCIA DA PENA PRIVATIVA DE LIBERDADE NO SISTEMA DE JUSTIÇA CRIMINAL EM IJUÍ", de autoria de Thiago dos Santos da Silva, Emmanuelle de Araujo Malgarim e Nelci Lurdes Gayeski Meneguzzi, tem como objetivo geral apresentar o papel da pesquisa acadêmica em direito sobre temas complexos, a partir da análise das condições de vulnerabilidade social e criminalização, explicitando as diversas violações dos princípios da dignidade humana e da personalidade da pena sofridas por familiares de pessoas encarceradas. A pesquisa qualitativa e exploratória questiona a efetividade do princípio da personalidade da pena no sistema carcerário brasileiro, focando em como a pena transcende o corpo do condenado, atingindo seus familiares.

A seguir, foi apresentado o texto intitulado "DESAFIOS E POSSIBILIDADES DE REINTEGRAÇÃO PELA LEITURA: UMA ANÁLISE A PARTIR DO PROJETO DE EXTENSÃO “LEITURA E EXISTÊNCIA” DA UNIJUÍ", de autoria de Thiago dos Santos da Silva, Patrícia Borges Moura e Patricia Marques Oliveski, tem como objetivo geral apresentar o projeto “Leitura e Existência” e o papel do letramento literário na reinserção social de apenados, como reforço ao direito à remição pela leitura, com foco na PMEI. O estudo analisa o papel da universidade na implementação da remição pela leitura, confirmando a hipótese de que o letramento literário fortalece a reinserção social e garante a dignidade das pessoas privadas de liberdade.

O objetivo do artigo "O CRIME DE ROUBO PRATICADO NO PERÍODO NOTURNO COMO CIRCUNSTÂNCIA JUDICIAL NEGATIVA E A ANÁLISE DA

JURISPRUDÊNCIA DO SUPERIOR TRIBUNAL DE JUSTIÇA" é analisar criticamente a possibilidade de majorar a pena-base do roubo apenas pelo horário noturno. Os autores, Yuri Anderson Pereira Jurubeba, Bruna Patricia Ferreira Pinto e Fernanda Matos Fernandes de Oliveira Jurubeba, concluem que a jurisprudência do STJ rechaça essa majoração isolada, exigindo fundamentação específica para evitar violação ao princípio da legalidade estrita e aos direitos fundamentais do acusado.

Os mesmos pesquisadores examinam os desafios processuais da Lei nº 15.123/2025, que aumentou a pena para crimes de violência psicológica contra a mulher com uso de IA. O artigo conclui que é premente a instituição de protocolos específicos de cadeia de custódia para prova digital, visando garantir a autenticidade e integridade da prova e a segurança jurídica.

André Vecchi e José Luiz de Moura Faleiros Júnior são os autores do ensaio "RESPONSABILIDADE PENAL DOS SISTEMAS AUTÔNOMOS DE INTELIGÊNCIA ARTIFICIAL: REFLEXÕES E PERSPECTIVAS JURÍDICAS FRENTE À DOGMÁTICA DO DELITO". O objetivo do trabalho é tentar vislumbrar a possibilidade de atribuir responsabilidade penal a sistemas inteligentes que causem lesões a bens jurídicos relevantes. O ensaio aborda as dificuldades de responsabilização das máquinas frente à dogmática penal atual, que se vê desafiada pelo surgimento da Inteligência Artificial.

A seguir, André Vecchi e Luciano Santos Lopes trabalham soluções para a aferição da tipicidade subjetiva no crime de lavagem de capitais, analisando se sua prática é possível apenas na modalidade dolo direto ou se também é admissível o dolo eventual. O artigo "A Imputação Subjetiva no Crime de Lavagem de Capitais" busca fixar parâmetros dogmáticos e propor soluções para as dificuldades probatórias da imputação subjetiva no processo penal.

José Guimarães Mendes Neto, Lucas Rafael Chaves de Sousa e Thiago França Sousa são os autores do trabalho "TEORIA DA PROVA NO PROCESSO PENAL E VEDAÇÃO À REVITIMIZAÇÃO: ANÁLISE DA ADPF 1107 E DOS PROTOCOLOS DO CNJ PARA JULGAMENTO COM PERSPECTIVA DE GÊNERO". O objetivo do estudo é investigar como a ADPF 1107 e os Protocolos do CNJ ressignificam a teoria da prova no processo penal, a partir da vedação à revitimização. O trabalho conclui que a tutela da dignidade da vítima se torna um novo pilar da dogmática probatória, exigindo reforma cultural dos operadores do Direito.

Wanderson Carlos Medeiros Abreu, Thiago França Sousa e Lucas Rafael Chaves de Sousa são os autores do trabalho "A ATIPICIDADE DO LINCHAMENTO NO DIREITO PENAL

BRASILEIRO: INCONGRUÊNCIAS DOGMÁTICAS E POLÍTICO-CRIMINAIS E CAMINHOS PARA O ENFRENTAMENTO INSTITUCIONAL". O objetivo do trabalho é identificar as incongruências dogmáticas e falhas político-criminais decorrentes da ausência de um tipo penal próprio para o linchamento no Brasil. O artigo propõe a reformulação do direito penal, com a criação de um tipo penal específico ou qualificadora, para oferecer uma resposta institucional mais proporcional a esse fenômeno de violência coletiva.

Em seguida foi apresentado o trabalho "A ARQUITETURA LEGISLATIVA DA PUNIÇÃO: COALIZÕES, NECROPOLÍTICA E A PRODUÇÃO SELETIVA DA POLÍTICA CRIMINAL NO CONGRESSO NACIONAL BRASILEIRO PÓS-1988", de autoria de Kennedy Da Nobrega Martins, Alexandre Manuel Lopes Rodrigues e Lucas Victor De Carvalho Gomes. O objetivo é analisar como o Congresso Nacional, pós-1988, produziu e consolidou um modelo de política criminal seletiva, atravessado por coalizões e uma racionalidade necropolítica. O artigo conclui que a seletividade penal é uma escolha política que esvazia a promessa constitucional de cidadania universal.

João Pedro Rêgo Balata, Emanuelle de Alencar Pereira e Wanderson Carlos Medeiros Abreu são os autores do artigo "A AMEAÇA DO CARÁTER SUBJETIVO DO DEPOIMENTO ESPECIAL ÀS GARANTIAS PROCESSUAIS EM CASOS DE VIOLÊNCIA SEXUAL INFANTO-JUVENIL". O objetivo é examinar como o caráter subjetivo do depoimento especial (Lei n.º 13.431/2017) tensiona garantias processuais como o contraditório e a ampla defesa. O trabalho busca evidenciar os dilemas do instituto e a necessidade de maior rigor metodológico e parâmetros claros de confiabilidade, sem perder sua função protetiva.

Deise Neves Nazaré Rios Brito e Alexandre Manuel Lopes Rodrigues Investigam como a midiática interfere na distinção entre dolo eventual e culpa consciente em casos de grande repercussão social no Brasil a partir de 2010, examinando fundamentos teóricos e propondo diretrizes de mitigação. O trabalho utiliza metodologia qualitativa com análise de casos paradigmáticos (Boate Kiss, Mariana, Brumadinho, Nardoni e Mariana Ferrer), demonstrando que a cobertura midiática dilui fronteiras dogmáticas entre institutos penais, favorece responsabilização pelo resultado e fragiliza presunção de inocência e devido processo legal, comprometendo imparcialidade judicial e segurança jurídica.

Por fim, Lucas Nacur Almeida Ricardo, Ana Carolina Letayf Campos e Luciano Santos Lopes analisam a diferenciação entre atos de preparação (impuníveis) e atos de execução (puníveis como tentativa) no iter criminis, propondo critérios interpretativos para o conceito de "iniciada a execução" mediante precedente vinculante. O artigo analisa o art. 14, II, do Código Penal, expõe teorias justificadoras da punição da tentativa, examina jurisprudência do

STJ que adota a teoria objetivo-formal e problematiza esse posicionamento por potencialmente gerar decisões desproporcionais e proteção penal insuficiente, considerando as obrigações processuais positivas do Estado de proteger bens jurídicos e vítimas, buscando equilíbrio entre legalidade e tutela efetiva.

Foi um privilégio poder acompanhar tantas discussões de excepcional nível acadêmico. Que venham os próximos encontros e debates!

São Paulo, Primavera de 2025.

Gustavo Noronha de Ávila

Rogério Luiz Nery Da Silva

# **RESPONSABILIDADE PENAL DOS SISTEMAS AUTÔNOMOS DE INTELIGÊNCIA ARTIFICIAL: REFLEXÕES E PERSPECTIVAS JURÍDICAS FRENTE À DOGMÁTICA DO DELITO**

## **CRIMINAL LIABILITY OF AUTONOMOUS ARTIFICIAL INTELLIGENCE SYSTEMS: LEGAL REFLECTIONS AND PERSPECTIVES IN THE FACE OF CRIMINAL LAW DOGMATICS**

**André Vecchi <sup>1</sup>**

**José Luiz de Moura Faleiros Júnior <sup>2</sup>**

### **Resumo**

O exponencial crescimento da Inteligência Artificial (IA), com o surgimento de máquinas inteligentes representou, efetivamente, uma quebra de paradigma, alterando, por completo, as relações sociais. Nesse compasso, o direito, em especial o direito penal, se viu desafiado a resolver questões inéditas capazes de desafiar até mesmo as suas tradicionais e consolidadas concepções. Este ensaio aborda este tema e vai adiante, tentando vislumbrar a possibilidade de atribuir responsabilidade penal a sistemas inteligentes quando estes lesarem de maneira demasiada bens jurídicos relevantes. Parte-se da compreensão do conceito de Inteligência Artificial (IA), definindo e explicando o funcionamento dos sistemas autônomos de tecnologia. Após, serão traçadas breves considerações acerca da concepção dogmática do delito, refletindo sobre cada um de seus elementos. Finalmente, elencam-se algumas dificuldades de atribuição de responsabilidades às máquinas frente a dogmática penal atual. Há, portanto, uma tentativa de compreender a resposta penal frente a lesões graves a bens jurídicos relevantes causados por sistemas autônomos de inteligência.

**Palavras-chave:** Inteligência artificial, Sistemas autônomos, Desvios incontrolados, Responsabilidade penal, Dogmática penal clássica

### **Abstract/Resumen/Résumé**

The exponential growth of Artificial Intelligence (AI), with the emergence of intelligent machines, effectively represented a paradigm shift, altering social relations. In this context, the law, especially criminal law, found itself challenged to resolve unprecedented issues capable of challenging even its traditional and consolidated concepts. This essay addresses this topic and goes further, attempting to glimpse the possibility of attributing criminal responsibility to intelligent systems when they cause excessive damage to relevant legal assets. It starts from an understanding of the concept of Artificial Intelligence (AI), defining and explaining the functioning of autonomous technology systems. Next, brief considerations will be made about the dogmatic conception of crime, reflecting on each of its elements.

---

<sup>1</sup> Mestrando em Direito pela Faculdade Milton Campos. Advogado. E-mail: andre.vecchi.lima@gmail.com

<sup>2</sup> Doutor em Direito pela Universidade de São Paulo. Professor da Faculdade Milton Campos. E-mail: josefaleirosjr@outlook.com



Finally, some difficulties in attributing responsibility to machines under current criminal dogma are listed. There is, therefore, an attempt to understand the criminal response to serious damage to relevant legal rights caused by autonomous intelligence systems.

**Keywords/Palabras-claves/Mots-clés:** Artificial intelligence, Autonomous systems, Uncontrolled deviations, Criminal liability, Classical criminal dogmatic

## 1. INTRODUÇÃO

De quando em vez, o direito penal se vê diante de situações inéditas que são capazes de desafiar até mesmo as suas tradicionais e consolidadas concepções. Tema já debatido com muito mais solidez na seara do direito civil, inegável reconhecer que a Inteligência Artificial (IA) passou a ser uma realidade em todas as áreas da vida humana, ensejando uma maior reflexão e discussão da dogmática penal.

Não existem dúvidas de que a acentuada evolução tecnológica do século XX permitiu ganhos inegáveis de eficiência, solucionando até as mais complexas mazelas humanas, antes impensáveis de solução, nos mais diversos campos, representando uma inequívoca mudança de paradigma. O surgimento de máquinas inteligentes, incluindo aplicações em *machine learning*, e *deep learning*, tem sido capaz de revolucionar todos os campos da vida humana, influenciando, até mesmo, as relações sociais.

A possibilidade de algoritmos aprenderem de forma autônoma, por meio da análise de dados, e tomarem decisões independentes demonstra o potencial disruptivo das novas tecnologias, sendo possível vislumbrar inúmeros ganhos na sua utilização. Todavia, em mesma proporção, inúmeros são os riscos gerados pelo uso desregrado e irresponsável dessas tecnologias autônomas.

Fato é, que o uso dessas novas tecnologias não pode ser desacompanhado de qualquer regulação, sendo responsabilidade do direito proporcionar e oferecer respostas para tentar coibir danos cometidos pela Inteligência Artificial (IA), bem como, para os danos que venham a ser causados por ela e que sejam incapazes de serem evitados.

Nesse contexto, a discussão de atribuição de responsabilidade civil de danos causados pela Inteligência Artificial (IA) pelos chamados “desvios autônomos” nos parece bem consolidada e capaz de proporcionar respostas, de certo modo, efetivas, para solucionar situações em que tenha ocorrido algum dano a um bem jurídico.

Contudo, o debate passa a ser norteador, por qual seria a resposta a situações em que a Inteligência Artificial (IA) venha a causar graves danos a um bem jurídicos relevantes (vida; incolumidade, tanto física quanto psíquica, dos indivíduos), parecendo insuficiente a mera responsabilização na seara civil.

Imagine uma situação em que um veículo autônomo, trafegando em rodovia, venha a perder o controle e colidir com outro veículo, causando a morte de todos os passageiros ali presentes. É de se questionar: a máquina seria passível de responsabilização penal?

O Direito Penal enfrenta, portanto, um desafio inédito: dar soluções e respostas a danos causados por sistemas de Inteligência Artificial (IA) contra um bem jurídicos penalmente

relevantes. Ainda que se mantenha distante desta realidade, sendo o último ramo do direito utilizado para dar repostas aos fenômenos sociais, até pelos seus próprios princípios (*ultima ratio*), tornou-se urgente a necessidade de o Direito Penal acompanhar a expansão tecnológica, desafiando até os seus conceitos mais antigos, a fim de propor soluções para lidar com a colisão entre inovação e proteção de bens jurídicos.

O presente estudo tem por escopo refletir acerca de situações em que sistemas autônomos de Inteligência Artificial (IA) lesam de maneira demasiada bens jurídicos relevantes, sendo insuficiente a mera responsabilização civil. Para tanto, buscar-se-á analisar se a tradicional dogmática do crime é capaz de atribuir responsabilidade penal aos sistemas autônomos de Inteligência Artificial (IA), ou, se faz necessário a criação de uma nova teoria do crime, um novo sistema de imputação próprio para as máquinas.

Nesse sentido, o objetivo do trabalho é, portanto, apresentar breves reflexões no campo dogmático, com a finalidade de reforçar o breve debate existente acerca de Direito Penal e Novas Tecnologias, verificando a resposta penal a situações de desvios autônomos causados por máquinas inteligentes.

A fim de proceder a essa análise, o trabalho partirá, inicialmente, da análise da Inteligência Artificial (IA), em especial de sistemas autônomos de tecnologia e seus desvios autônomos. Em seguida, abordar-se-á a concepção tradicional da dogmática penal, delimitando todos os elementos do crime. Por fim, irá se refletir se é possível a atribuição de responsabilidade penal a máquinas inteligentes com base na dogmática penal existente. Assim, buscará se compreender a possibilidade de reposta penal a lesões graves a bens jurídicos relevantes causados por sistemas autônomos de inteligência.

## **2. INTELIGÊNCIA ARTIFICIAL: CONCEITO E SISTEMAS AUTÔNOMOS**

### **2.1. Conceito**

Em que pese a infinidade de definições que podem ser encontradas na doutrina nacional e estrangeira para conceituar a Inteligência Artificial (IA), fato é que existe muito pouca clareza quanto a todos os conceitos relacionados a esse campo de estudo, como “algoritmos”, “robótica”, “*machine learning*” e “*deep learning*” (FALEIROS JÚNIOR, 2022).

Segundo Marvin Minsky, Inteligência Artificial é “a ciência de construir máquinas para que façam coisas que, se fossem feitas por humanos, exigiriam inteligência” (MINSKY, 2003 in MORÁN ESPINOSA, 2021). No mesmo sentido, Rouse compreende a Inteligência Artificial como “aquela inteligência que se manifesta através das máquinas, a ajuda idônea para qualquer

ser humano, um agente racional, flexível e serviçal” (ROUSE, 2017 in MORÁN ESPINOSA, 2021).

Por sua vez, Barr e Feigenbaum conceituam a Inteligência Artificial como “a parte da ciência da computação que se preocupa em desenvolver sistemas computacionais inteligentes, isto é, sistemas que exibem características que associamos à inteligência no comportamento humano (BARR; FEIGENBAUM, 1981).

Por esta razão, Túlio Felipe Xavier Januário compreende a Inteligência Artificial a partir de duas perspectivas: a primeira, como um sistema no qual estão incluídos “softwares e hardwares que, para alcançar um determinado objetivo, analisam o entorno através da correlação e interpretação de dados e tomam uma decisão com base em regras simbólicas preestabelecidas” (JANUÁRIO, 2023); a segunda, como disciplina científica que engloba “diversas técnicas e metodologias, como *machine learning*, *machine reasoning* e robótica, integrando-as em sistemas ciberfísicos” (JANUÁRIO, 2023).

Apesar de discordarmos brevemente dessa delimitação, pode-se compreender, portanto, a Inteligência Artificial não como um conceito único, mas como um ramo científico multidisciplinar (FALEIROS JÚNIOR; VECCHI; GOULART, 2025), produto da evolução tecnológica desde os anos 50 com Alan Turing (MORÁN ESPINOSA, 2021), e não como uma única tecnologia propriamente dita.

O que se coloca em debate é a delimitação dos conceitos de *machine learning*, e *deep learning*, compreendendo o que seriam os Sistemas Autônomos de tecnologia, bem como, a capacidade dessas inovações de sistematizarem dados e tomarem decisões por conta própria, a partir desses dados.

## 2.2. SISTEMAS AUTÔNOMOS

Compreendendo, portanto, a Inteligência Artificial (IA) como um ramo científico multidisciplinar (FALEIROS JÚNIOR, 2022), isto é, uma disciplina científica que engloba diversos tipos de tecnologia, faz-se necessário compreender, especificamente, quais são as tecnologias de maior potencial disruptivo, ou melhor, quais são as tecnologias capazes de tomarem decisões de forma autônoma (FALEIROS JÚNIOR; VECCHI; GOULART, 2025).

Agentes autônomos são normalmente empregados em situações em que são necessários um enorme processamento de dados e informações, demandando combinações precisas e uma eficaz reação (GLESS; WEIGEND, 2019). No entanto, os pressupostos para se considerar um agente como inteligente são tão controversos quanto à própria definição do conceito de Inteligência Artificial (IA) (GLESS; WEIGEND, 2019).

Morán Espinosa compreende quatro modalidades de sistemas de inteligência: IA Assistida, que colabora para a realização de tarefas com rapidez; IA Automatizada, que realiza tarefas cotidianas e excepcionais de forma automática; IA Aumentada, que facilita tomadas de decisão, aprendendo com a interação realizada e resultados obtidos; e por fim, IA Autônoma; aquela com capacidade de tomar decisões sem intervenção humana (MORÁN ESPINOSA, 2021).

Em uma concepção mais concisa, Gless e Weigend trabalham com três tipos de sistemas para apreciação do ponto de vista jurídico: Sistemas Simples de Processamento de Dados, nos quais o desenvolvedor do sistema definiu por completo a forma de captação e processamento de informações com base em algoritmos definidos; Sistemas Abertos, os quais, por meio de sensores, podem registrar dados do mundo exterior e processá-los conforme regras imutáveis já estabelecidas do próprio sistema; e, Sistemas Inteligentes, aqueles que não só recolhem uma enorme quantidade de dados, mas os interpretam conforme determinados padrões, reagindo, de forma autônoma (GLESS; WEIGEND, 2019).

Nessa concepção, sistemas abertos e inteligentes atuam normalmente, analisando e processando sensorialmente o ambiente. Contudo, somente sistemas inteligentes são capazes de reagir diretamente às informações recolhidas e interpretadas, armazenando-as de forma a utilizar em futuros processos de tomada de decisão (GLESS; WEIGEND, 2019).

Desse modo, seguindo uma linha adotada por Susana Aires de Souza, podemos diferenciar os sistemas entre sistemas pré-programados e orientados à realização de tarefas determinadas (“*deterministic robots*”) e sistemas de inteligência artificial capazes de autonomamente, em situações complexas, tomarem decisões de forma independente, ou em outras palavras, sem tomarem decisões pré-programadas ou que sequer foram previstas pelo programador (AIRES DE SOUSA, 2020).

Assim, compreende-se por sistemas autônomos, sistemas capazes de tomarem decisões de maneira independente, sem qualquer interferência ou precisão humana, a partir da análise e processamento de dados, ou seja, aprendem e tomam decisões de forma autônoma. São estes que, lesando interesses jurídicos protegidos pelo direito penal, constituem objeto de reflexão.

Estes sistemas autônomos operam através de bases algorítmicas capazes de produzirem informações através de “inputs” que lhe são dados (AIRES DE SOUSA, 2020). Esses algoritmos normalmente são potencializados com base em *machine learning* e *deep learning*.

*Machine learning* é uma aplicação da Inteligência Artificial que desenvolve máquinas com acesso a dados, sendo capazes de aprender por si mesmas, ou seja, algoritmos que reconhecem padrões específicos, organizam dados e permitem que a máquina aprenda a realizar

suas funções de maneira mais inteligente (MORÁN ESPINOSA, 2021). Em outras palavras, compreende-se *machine learning* como um conjunto de processos estatísticos iniciados a partir de uma base de informações e de um conjunto de algoritmos, a partir dos quais é deduzido um procedimento que permite prever uma nova informação e, com isso, a tomada de uma decisão (AIRES DE SOUSA, 2020).

*Deep learning* tem como objetivo imitar uma rede neural, simulando o modo como os seres humanos adquirem certas formas de conhecimento. Conecta-se, portanto, comunicações com outras determinadas informações, sendo uma vertente do *machine learning* que foi desenvolvido para ampliar os contextos em que era aplicado (MORÁN ESPINOSA, 2021).

A potencialização destes algoritmos com *machine learning* e *deep learning* são capazes de tornarem os sistemas ainda mais autônomos e inteligentes, construindo modelos e regras que permitem tomar decisões mais próximas dos seres humanos. Por estas razões, alguns doutrinadores defendem que os algoritmos seriam tão complexos que ultrapassariam o mero processamento estatístico de dados e se tornariam máquinas capazes de ‘pensar’ (FALEIROS JÚNIOR, 2022; DE SANCTIS, 2020).

Todavia, ainda que potencializados por *machine learning* e *deep learning*, os algoritmos são falíveis e propensos a erros de representação e assimilação (CRESPO, 2021), em razão do seu caráter exclusivamente matemático de processamento dos dados (FALEIROS JÚNIOR, 2022). Exatamente por essa característica, os sistemas autônomos são capazes de gerarem determinado “*output*” (“*act unexplainably*”) sem qualquer previsibilidade humana, podendo-se dizer que os resultados lesivos causados por desvios são exclusivamente seus (AIRES DE SOUSA, 2020).

Neste contexto, considera-se que as decisões tomadas por sistemas autônomos são opacas, sendo que nem o seu próprio desenvolvedor é capaz de prever como a tecnologia agirá em determinadas situações. Isto quer dizer, “o tratamento algorítmico dos dados, segundo uma estrutura complexa, torna opaco o processo que conduz a determinado resultado, não obstante a sua capacidade de grande precisão na determinação de nova informação” (AIRES DE SOUSA, 2020, p. 67).

Exemplo tradicional é o dos veículos automatizados, sendo o carro integralmente conduzido pelo computador e não pelo humano. Na hipótese de um acidente com um carro sem motorista, que, a partir do processamento de informações, toma uma decisão equivocada (desvio incontrolado) para realizar uma ultrapassagem em uma rodovia e acaba colidindo com outro veículo, vindo a causar a morte de todos os ocupantes deste carro, quem agiu e suporta a responsabilidade penal? É possível uma responsabilidade penal do próprio agente inteligente?

Fato é, que no campo do direito civil, a questão tem sido amplamente discutida e respostas a danos causados por sistemas inteligentes já são sólidos. Todavia, no âmbito penal, a questão está longe de solução. Não nos parece suficiente a mera responsabilização civil em situações em que a máquina atinge bens jurídicos relevantes penalmente protegidos. Para responder a tal situação, coloca-se a dogmática penal tradicional em um complexo desafio: entre a responsabilidade penal e o dano causado por um sistema autônomo.

### **3. TEORIA DO DELITO**

A imputação de responsabilidade penal aos sistemas autônomos passa, em primeiro lugar, pela compreensão do fato punível, assim como, de todos os elementos que o compõem.

Partindo de uma concepção majoritária na dogmática jurídico-penal moderna, compreende-se o fato punível como “uma ação, que deve ser típica, antijurídica, culpável e realizadora de outros eventuais pressupostos de punibilidade” (ROXIN, 2024, p. 388). O comportamento punível compõe-se, portanto, de quatro elementos, quais sejam: ação, tipicidade, antijuridicidade e culpabilidade, aos quais, em determinadas situações, se pode exigir um adicional pressuposto de punibilidade.

#### **3.1. Ação**

Em primeiro lugar, não há como se falar em fato punível sem ação. No entanto, nem toda ação é penalmente relevante, sendo necessário a vinculação de seu conceito a uma norma criminalizadora para estar ser objeto do direito penal (TAVARES; MARTINS, 2025). Zaffaroni enfatiza que, apesar de ser composto por elementos da vida humana, o conceito de ação é jurídico (ZAFFARONI; PIERANGELI, 2023), pois uma ação penalmente irrelevante não tem nenhuma importância para a teoria do delito (TAVARES; MARTINS, 2025).

Ao longo do tempo, a doutrina penal sempre se preocupou em conceituar a ação. Desde a teoria causal, evoluindo para teoria finalista até a teoria social, há uma preocupação em delimitar a ação, limitando as hipóteses de intervenção estatal.

O conceito causal de ação, desenvolvido por Litz e Beling, compreende a ação como “produção causal de um resultado no mundo exterior por um comportamento humano voluntário” (SANTOS, 2022). Nessa concepção, a ação era composta pelos seguintes elementos: 01) um movimento corporal; 02) causado por um impulso de vontade, que, por sua vez; 03) causava uma modificação no mundo exterior (MIR PUIG, 2019).

A ação era, portanto, uma mera causação do evento externo, sendo o impulso da vontade (BITENCOURT, 2023). Contudo, o conteúdo da vontade era um elemento interno, deslocado

para a culpabilidade (PACELLI; CALLEGARI, 2018). Esse conceito enfrentou problemas, pois conduzia a uma concepção insatisfatória quanto aos delitos omissivos, que não comportavam um movimento corpóreo (TAVARES; MARTINS, 2025).

O conceito finalista de ação, desenvolvido por Welzel, compreende que a ação não é causada somente por uma vontade, mas conduzida pelo conteúdo da finalidade da vontade (MIR PUIG, 2019), em outras palavras, uma conduta dirigida a um fim. Nessa concepção a ação seria composta: 01) a antecipação do fim ou objetivo; 02) a escolha, a seleção e domínio dos meios para o alcance do fim; 03) a vontade dirigente da execução; 04) a execução dos meios para conseguir o objetivo (TAVARES; MARTINS, 2025).

A ação seria, portanto, composta por um segmento externo, que comportaria os meios causais e sua execução, e por um meio subjetivo, composto pela representação do fim e pela vontade do agente (TAVARES; MARTINS, 2025). Assim, a ação seria “o comportamento humano voluntário conscientemente dirigido a um fim” (BITENCOURT, 2023, p. 277). Essa concepção não ficou isenta de críticas, especialmente, à sua não adequação satisfatória aos crimes culposos, pois não seriam abrangidos pela vontade (PACELLI; CALLEGARI, 2018).

O conceito social de ação, desenvolvido por Schmidt e, posteriormente, desenvolvido por Jescheck, compreende a ação como “uma conduta socialmente relevante” (JESCHECK, 1996 in TAVARES; MARTINS, 2025, p. 189). Buscou-se, a partir dessa concepção, encontrar um conceito comum aos delitos dolosos, culposos e omissivos, a fim de encontrar um denominador que alcançasse todas as classes de delito (MIR PUIG, 2019). Essa concepção é criticada por representar um juízo epistemológico, atribuindo ao conceito de ação um juízo de relevância (TAVARES; MARTINS, 2025).

De acordo com a doutrina majoritária, pode-se compreender a ação como “um comportamento humano relevante no mundo exterior, dominado ou ao menos dominável pela vontade” (ROXIN, 2024, p. 389). Não se consideram ações, em sentido jurídico-penal, efeitos causados pela natureza, animais e sequer os atos de uma pessoa jurídica (PACELLI; CALLEGARI, 2018).

Conforme compreende Muñoz Conde, “se chama ação todo comportamento da vontade humana. Só o ato voluntário pode ser penalmente relevante e a vontade implica sempre uma finalidade” (MUÑOZ CONDE; GARCÍA ARÁN, 2007, p. 25). A ação, enquanto elemento do delito, engloba, um fazer (comissão) quanto um não fazer (omissão).

O conceito de ação também possui uma função negativa, ou seja, compreende aqueles casos em que não há conduta (ZAFFARONI; PIERANGELI, 2023). Assim, se exclui do direito



penal todos os comportamentos que não merecem o qualificativo de ação (MUÑOZ CONDE; GARCÍA ARÁN, 2007).

### **3.2. A Tipicidade**

A ação deve ser típica, isto é, ela deve corresponder a uma norma penal incriminadora (tipo penal) (ROXIN, 2024). Essa ação típica deve se integrar com os componentes necessários a todo comportamento: a parte objetiva e a parte subjetiva (MIR PUIG, 2019).

A tipicidade é, portanto, “a adequação de um ato cometido à descrição desse ato no direito penal” (MUÑOZ CONDE; GARCÍA ARÁN, 2007, p. 55). Essa adequação deve ocorrer tanto no campo objetivo quanto do campo subjetivo, sob pena de não ocorrer a conduta típica.

Elementos objetivos são aqueles que não dependem da vontade do agente para sua existência (TAVARES; MARTINS, 2025), podendo serem identificador pelos sentidos (PACELLI; CALLEGARI, 2018). Estes se dividem em descritivos e normativos.

Os descritivos são aqueles que podem ser aprendidos sem qualquer esforço, não necessitam de qualquer juízo de valor para sua compreensão (PACELLI; CALLEGARI, 2018). Já os normativos são aqueles que exigem, para sua compreensão, um juízo de valor, seja do ponto de vista social, cultural ou jurídico (TAVARES; MARTINS, 2025).

A tipicidade objetiva abarca, portanto, o aspecto externo da ação (MIR PUIG, 2019). Só pode se falar em tipicidade objetiva quando a ação se adequa a todos os elementos objetivos descritos no tipo penal. Isto é, a conduta somente será objetivamente típica quando se adequar a todos os elementos, descritivos e normativos, descritos na lei.

Segundo a teoria final da ação, a ação ou omissão não se subsumi ao tipo em um simples processo causal cego, mas como um processo dirigido ao alcance de uma vontade dirigida a um fim (MUÑOZ CONDE; GARCÍA ARÁN, 2007). Assim, ao realizar um juízo de tipicidade, este deve compreender o conteúdo dessa vontade (BITENCOURT, 2023). A tipicidade subjetiva alcança, portanto, a expressão mental do agente, isto é, o conteúdo da vontade que dirige a ação (MUÑOZ CONDE; GARCÍA ARÁN, 2007).

O elemento subjetivo geral é o dolo, compreendido como “a consciência e vontade de realizar os elementos objetivos do tipo, tendo como objetivo final a lesão ou o perigo concreto de lesão ao bem jurídico” (TAVARES; MARTINS, 2025, p. 329).

Seguindo ainda uma concepção do dolo baseada na teoria da vontade, este é composto de dois elementos: um cognitivo (intelectual) e um volitivo.

Para atuar dolosamente, o agente deve conhecer todas as circunstâncias objetivas do tipo penal (PACELLI; CALLEGARI, 2018). Assim, o elemento cognitivo (intelectual) se refere

“aos elementos que caracterizam objetivamente a conduta como típica (elementos objetivos do tipo): sujeito, conduta, resultado, relação causal” (MUÑOZ CONDE; GARCÍA ARÁN, 2007, p. 70). Sem conhecimento, não há que se falar em dolo.

O elemento volitivo, por sua vez, se consubstancia na vontade de querer realizar a ação típica, lesando o bem jurídico protegido. O agente, conhecendo a ação típica, decide ainda sim realizá-la e age de acordo com essa decisão.

Como elemento subjetivo ainda existe a culpa. A culpa “é a inobservância do dever objetivo de cuidado manifestada numa conduta produtora de um resultado não querido, mas objetivamente previsível” (BITENCOURT, 2023, p. 358)

Assim, a conduta culposa ocorre quando o agente não deseja produzir o resultado, porém, ainda sim o produz em razão de uma imprudência, ou, em outras palavras, em razão da falta de um dever objetivo de cuidado (PACELLI; CALLEGARI, 2018). O agente confia na não produção do resultado, não desejando a realização do tipo penal, mas a conduta desatenta acaba por gerar efeitos penais.

Desse modo, para compreender a ação como típica, devem estar presentes os elementos objetivos previsto no tipo penal (tipicidade objetiva), consubstanciados pelo elemento subjetivo, dolo ou culpa (tipicidade subjetiva). Sem a tipicidade, não há que se falar em fato punível.

### **3.3. Antijuridicidade**

Toda ação compreendida em um tipo será antijurídica caso não houver uma causa de justificação (CEREZO MIR, 1976). Assim, aferida a tipicidade no caso concreto, quer dizer, demonstrado que a ação corresponde a um tipo penal, tanto em seu campo objetivo quanto subjetivo, o próximo passo, passa para atribuição de responsabilidade penal, passa pela determinação da sua antijuridicidade (MUÑOZ CONDE; GARCÍA ARÁN, 2007).

A antijuridicidade não surge do direito penal, mas da ordem jurídica em si (ZAFFARONI; PIERANGELI, 2023). Assim, a antijuridicidade é a contradição da ação com o direito (PACELLI; CALLEGARI, 2018), compreendido como uma ordem normativa e de preceitos permissivos.

Assim, a antijuridicidade passa pela “constatação de que a conduta típica (antinormativa) não está permitida por qualquer causa de justificação (preceito permissivo)” (ZAFFARONI; PIERANGELI, 2023, p. 494). Existente uma causa de justificação, não há que se falar em antijuridicidade e, conseqüentemente, não há fato punível (ROXIN, 2024).

Nessa concepção, a antijuridicidade é um conceito externo à teoria do crime, relacionando-se a ordem jurídica em seu sentido global, restringindo apenas a tipicidade e culpabilidade como categorias próprias do crime (PACELLI; CALLEGARI, 2018).

Costuma-se ainda dividir a antijuridicidade em formal e material (TAVARES; MARTINS, 2025). A antijuridicidade formal compreende a contradição do fato com a ordem jurídica, enquanto a antijuridicidade material pode ser concebida como a contradição entre o fato e as condições sociais de existência (TAVARES; MARTINS, 2025).

Compreende-se, portanto, que a realização de um tipo penal não é antijurídica se abarcada por alguma causa de justificação.

### **3.4. Culpabilidade**

A ação típica e antijurídica tem, por fim, ser culpável, isto quer dizer, o seu ator poder ser responsabilizado por ela, ou, em outras palavras, ela deve poder ser-lhe reprovada (ROXIN, 2024). Examina-se, portanto, se o agente pode ser responsabilizado pela violação da norma de determinação, ou seja, pela prática da ação ou omissão típica e antijurídica (CEREZO MIR, 1976).

Para a imposição de uma pena, principal consequência jurídica do delito, não é suficiente que o ato seja típico e antijurídico (MUÑOZ CONDE; GARCÍA ARÁN, 2007). A mera realização de uma ação típica e ilícita não acarreta a imposição de uma sanção penal, existindo casos em que o autor se livra da responsabilidade penal (MUÑOZ CONDE; GARCÍA ARÁN, 2007).

A culpabilidade, enquanto um juízo de reprovação, significa, portanto, “a valoração negativa dos princípios orientadores pelos quais o autor se deixou levar na formação de sua vontade e por isso o fato cometido por ele deve ser reprovado” (PACELLI; CALLEGARI, 2018, p. 405). Assim, culpabilidade é a reprovação da vontade do agente em descumprir a norma proibitiva, a qual deveria ter se motivado, inexistindo razão que o impedisse (ZAFFARONI; PIERANGELI, 2023).

A culpabilidade, como juízo de reprovação, tem como elementos: a) capacidade de culpabilidade; b) a consciência potencial da antijuridicidade; c) a exigibilidade de comportamento distinto (MUÑOZ CONDE; GARCÍA ARÁN, 2007). Assim, a aferição de culpabilidade deve ser realizada a partir da verificação destes elementos.

A capacidade de culpabilidade é um juízo político acerca da capacidade de entendimento do agente e sua autodeterminação frente às suas deficiências (TAVARES; MARTINS, 2025). Assim, o agente, no momento do fato, deve ser capaz de “compreender que

o fato não está autorizado e determinar-se de acordo com esta compreensão, é dizer, abster-se da realização do fato” (PACELLI; CALLEGARI, 2018, p. 410).

A consciência potencial da antijuridicidade é a capacidade de o agente conhecer o conteúdo das proibições, ou seja, se o sujeito não sabe que a conduta é proibida, não há razões para evitar sua realização (MUÑOZ CONDE; GARCÍA ARÁN, 2007). Assim, caso o agente desconheça a norma ou a interpreta falsamente, ele opera em erro de proibição<sup>1</sup>, supondo erroneamente que a conduta não é contrária ao direito (PACELLI; CALLEGARI, 2018).

A exigibilidade do comportamento distinto corresponde a exigência que o agente guie seu comportamento conforme o direito, ou seja, em obediência às normas que regem a sociedade (MUÑOZ CONDE; GARCÍA ARÁN, 2007). Assim, conhecendo que sua conduta era ilícita, o agente deveria ter decidido atuar de outro modo, em conformidade com o ordenamento jurídico (CEREZO MIR, 1976).

A atribuição de responsabilidade penal, passa, portanto, pela verificação de todos os elementos que compõe o delito: ação, tipicidade, antijuridicidade e culpabilidade. Assim, presentes todos os elementos, a conduta seria punível e o agente capaz de sofrer a imposição de uma pena.

## **4. A ATRIBUIÇÃO DE RESPONSABILIDADE PENAL À INTELIGÊNCIA ARTIFICIAL**

### **4.1. A Problemática da Ação**

A atribuição de responsabilidade penal aos sistemas autônomos encontra seu primeiro óbice na própria concepção dogmática da ação (CRESPO, 2021). Um carro autônomo que colide e vem a causar a morte de duas pessoas, ainda que esse resultado tenha sido produzido exclusivamente pela própria máquina com base em algoritmos, independentemente de qualquer controle humano, é possível dizer que essa conduta é penalmente relevante?

De início, cumpre dizer que não. Isso porque, a doutrina, a unanimidade, compreende a ação como um comportamento exclusivamente humano, sendo que eventos da natureza, os animais e atos de pessoas jurídicas não constituem ação em sentido jurídico-penal (ROXIN, 2024).

---

<sup>1</sup> **Erro sobre a ilicitude do fato**

Art. 21 - O desconhecimento da lei é inescusável. O erro sobre a ilicitude do fato, se inevitável, isenta de pena; se evitável, poderá diminuí-la de um sexto a um terço.

Parágrafo único - Considera-se evitável o erro se o agente atua ou se omite sem a consciência da ilicitude do fato, quando lhe era possível, nas circunstâncias, ter ou atingir essa consciência.

Desse modo, a dogmática exige, para considerar uma ação como penalmente relevante, que essa seja humana, não sendo possível, ao menos a *prima facie*, a imputação de uma responsabilidade penal aos sistemas autônomos, pois ausente elemento imprescindível da concepção tripartida do delito (HINCKEL TEIDER; PIVATTO DOS SANTOS, 2020).

Contudo, nos parece possível refletir brevemente acerca da concepção da ação como um comportamento exclusivamente humano. Partindo de uma concepção causalista, a qual compreendia a ação como um mero movimento corporal voluntário, causado por um impulso de vontade, que causava uma modificação no mundo exterior (MIR PUIG, 2019), pode-se pensar em uma adequação do conceito de ação aos agentes inteligentes, pois as máquinas executam movimentos, sendo capazes de movimentar seus “corpos” (GLESS; WEIGEND, 2019).

Para além, considerando que para o causalismo, o conteúdo da vontade reside na culpabilidade (BITENCOURT, 2023), só se negando a voluntariedade quando o corpo reage “automaticamente” ou quando ele é usado como objeto psíquico por outro homem, pode-se dizer que as máquinas, com a sua capacidade de tomar decisões autônomas, fogem do controle humano e tomam decisões voluntárias (GLESS; WEIGEND, 2019).

Diferentemente da teoria causal, partindo de uma concepção finalista da ação, que a compreende como uma conduta dirigida a um fim (MIR PUIG, 2019), exigindo-se, como elementos: 01) a antecipação do fim ou objetivo; 02) a escolha, a seleção e domínio dos meios para o alcance do fim; 03) a vontade dirigente da execução; 04) a execução dos meios para conseguir o objetivo (TAVARES; MARTINS, 2025), a adaptação do conceito de ação à máquina encontra óbices.

Ainda que se compreenda que a máquina consiga se autodeterminar para alcançar um objetivo, não parece ainda, que o agente inteligente reconheça o sentido e o significado de suas ações, tendo a autoconsciência necessária para estabelecer metas “deliberadamente” e se colocar de acordo com essas metas (GLESS; WEIGEND, 2019).

Ao menos por enquanto, não parece que os agentes inteligentes tenham uma determinação finalista autoconsciente, capazes de praticarem atos com a plena consciência das suas consequências e de deliberadamente traçarem metas, se colocando na posição de realizá-las.

Desse modo, ainda que a princípio o conceito de ação sequer permita considerar uma conduta praticada por uma máquina como penalmente relevante, limitando essa concepção exclusivamente a comportamentos humanos, são possíveis de se vislumbrar hipóteses de reflexão.

Do ponto de vista do causalismo, baseado em considerações meramente externas, compreendendo ação como qualquer movimento corporal voluntário que cause uma alteração no mundo exterior, é possível compreender os agentes inteligentes como seres que praticam condutas (GLESS; WEIGEND, 2019). Contudo, partindo do ponto de vista finalista, o atual estágio de desenvolvimento dos agentes inteligentes não permite essa mesma conclusão. Isso porque, não é possível se dizer que os agentes tenham autoconsciência plena, determinando objetivos. Desse modo, ante uma maior densidade do conceito de conduta, amparada em uma determinação autoconsciente, os agentes inteligentes ainda não seriam compreendidos como seres que praticam conduta.

#### **4.2. A Problemática do Elemento Subjetivo**

Ainda que realizadas reflexões acerca da ação, a fim de considerar como penalmente relevantes as condutas perpetradas pelos agentes inteligentes, a responsabilização penal das máquinas encontra óbice em outro elemento do fato punível: a tipicidade subjetiva.

A partir da concepção finalista, criada por Welzel, a finalidade do agente passou a ser considerada como elemento da tipicidade, sendo necessária a verificação do dolo (vontade/propósito) ou da culpa para a configuração do fato punível. Surge então, outro problema: as máquinas inteligentes têm intenção?

A *prima facie*, compreendemos que não. Isso porque, a própria opacidade do sistema dificulta, ou melhor, torna quase impossível a verificação do elemento subjetivo (AIRES DE SOUSA, 2020). A realização de um resultado danoso, de uma lesão a um bem jurídico por um agente inteligente, ou seja, sua tomada de decisão é baseada em um sistema objetivo, a partir de informação algorítmica, sendo impossível afirmar que há qualquer intenção ou até mesmo “descuido” por parte da máquina (AIRES DE SOUSA, 2020).

Gabriel Hallevy, em uma concepção inovadora, compreende a possibilidade de atribuição de responsabilidade penal aos agentes inteligentes a partir da verificação dos elementos externos (*actus reus*) e mentais (*mens rea*) (AIRES DE SOUSA, 2020). Para o autor, as máquinas seriam capazes de aprender com o mundo externo, se tornando capazes de prever as próprias ações, e assim, agir de acordo com um objetivo (HALLEVY, 2024). Desse modo, suas ações poderiam integrar os requisitos de algumas das formas de *mens rea*: pelo menos negligência (*negligence*) e, em última análise, também conhecimento (*knowledge*), se não mesmo intenção real” (CAPPELLINI, 2019). Seria ainda irrelevante, o fato de que as máquinas não possam provar emoções e sentimentos (CAPPELLINI, 2019).

Todavia, essa concepção encontra diversas limitações. Em primeiro lugar, admitir uma compreensão assética, esvaziada de conteúdo valorativo, confronta firmemente a atribuição de responsabilidade penal tradicional, equiparando o cometimento de um delito a responsabilidade por um dano causado, quase nos moldes da responsabilidade civil (AIRES DE SOUSA, 2020).

Ademais, a capacidade de agir antijuridicamente não se confunde com a capacidade do algoritmo de fazer opções e tomar decisões, confundindo o livre arbitrio e entendimento moral, com uma escolha pré-programada. O agente só pode ser considerado responsável por determinada conduta quando tiver consciência e capacidade de entender o conteúdo da norma enquanto destinatário da norma (AIRES DE SOUSA, 2020).

Fato é, que conforme defende Cappellini, ainda que a máquina seja capaz de tomar decisões a partir da verificação do mundo externo, a culpabilidade deve ser verificada antes da própria vontade, agindo como um pressuposto da tipicidade. Isso porque, ainda que a máquina tome decisões típicas, não há uma capacidade de autodeterminação conforme um comportamento antijurídico, ou melhor, a máquina não é dotada de capacidade de escolha e de se posicionar conforme a lei (CAPPELLINI, 2019).

Portanto, a concepção de Hallevy, ainda que baseada na verificação de elementos mentais (*mens rea*) a partir de fatores externos, encontra diversas limitações. Mesmo que suficiente para resolver problemas de imputação subjetiva, a teoria encontraria limitação no elemento da culpabilidade, sendo as máquinas incapazes de autodeterminação.

Destarte, diante da concepção tradicional, nos parece impossível admitir ou verificar que a máquina tenha vontade ou intenção. A tomada de decisão pelos agentes inteligentes ainda é extremamente objetiva, baseada somente em processamento de dados e verificação algorítmica, o que tornaria impossível a imputação subjetiva de crimes a essas tecnologias.

#### **4.3. A Problemática da Culpabilidade**

Passando então a conceber a análise da culpabilidade previamente a da própria tipicidade, a fim de compreender se os agentes inteligentes são capazes de autodeterminação, surge o terceiro problema: as máquinas podem se fazer culpáveis?

Conforme a dogmática jurídico-penal, a culpabilidade, enquanto juízo de reprovação, depende da verificação dos seguintes elementos: tem como elementos: a) capacidade de culpabilidade; b) a consciência potencial da antijuridicidade; c) a exigibilidade de comportamento distinto (MUÑOZ CONDE; GARCÍA ARÁN, 2007). Questiona-se, portanto, se é possível reprovar a máquina a partir da verificação de todos esses elementos em razão do recebimento de um *input* determinado, baseado em algoritmos, e a tomada de uma decisão.

Em primeiro lugar já seria difícil ou quase impossível falar que a máquina teria capacidade de culpabilidade, isso porque, não nos parece que a tecnologia tenha atingido patamares tão elevados de desenvolvimento que a tornem capazes de compreender por inteiro o caráter ilícito do fato praticado e assim se autodeterminar (PAULA; CORNWALL; CABRAL, 2020).

Mais complexo ainda se torna a decisão acerca da consciência potencial da antijuridicidade do fato. Para verificação desse elemento seria necessário demonstrar que as máquinas possuem capacidade para compreender que a conduta praticada é crime (ALVES, 2021). Ou seja, exigiria do agente inteligente uma compreensão de toda a ordem normativa, tendo possibilidade de conhecimento do que é proibido e permitido na sociedade, a fim de se autodeterminar de forma contrária ao ordenamento jurídico.

Por fim, ainda que a máquina atingisse patamares de desenvolvimento aptos a compreender a ordem normativa por inteiro, seria necessário exigir que ela tivesse conduta diversa daquela antijurídica, compreendendo que poderia ter agido de outro modo e não o fez (ALVES, 2021).

Como afirmam Sabine Gless e Thomas Weigend, nos termos do definido pelo BGH (Superior Tribunal de Justiça da Alemanha), “o fundamento interno da reprovação da culpabilidade reside no fato de que o homem está investido de uma liberdade de autodeterminação, responsável e moral, e, por isso, é capaz de decidir pelo justo e contra o injusto, de adaptar seu comportamento aos mandamentos jurídicos e evitar o que seja juridicamente proibido”. A partir dessa concepção tradicional, a culpabilidade do agente inteligente estaria categoricamente excluída (GLESS; WEIGEND, 2019).

Contudo, os autores propõem uma concepção da culpabilidade a partir de uma perspectiva de expectativas e imputações nas relações sociais e não na capacidade de autodeterminação, fundada em processos neurais. Assim, aquele identificado como causador de um conflito social sedimentado no ato punível (GLESS; WEIGEND, 2019). Desse modo, o agente que causa um resultado lesivo a um bem jurídico protegido é o causador de um conflito social, podendo ser culpável.

Quanto mais complexas as funções desenvolvidas pelos agentes inteligentes, mais complexas são suas estruturas de aprendizado, de modo que, a partir do fornecimento de regras de comportamento detalhadas à máquina, essa seria capaz de ajustar sua conduta a fim de não violar regras sociais (GLESS; WEIGEND, 2019).

Então, a partir de um sistema positivo-negativo, onde a máquina fosse capaz de avaliar suas missões e se autovalorar eticamente, Gless e Weigend compreendem que seria pertinente



fazer um “reproche” de culpabilidade a um agente inteligente. Ao invés da compreensão da antijuridicidade da conduta, a verificação da culpabilidade seria realizada através do reconhecimento do caráter indesejado da conduta a nível social pelo agente (GLESS; WEIGEND, 2019).

Esta proposta ressignifica, obviamente, o conceito tradicional da culpabilidade na dogmática jurídico-penal, possibilitando uma atribuição de responsabilidade penal aos agentes inteligentes. Contudo, ainda sem uma efetiva afirmação, com base somente na verificação dos elementos da culpabilidade, a sua atribuição às máquinas nos parece, ainda, impensável.

## **5. CONCLUSÃO**

O principal objetivo deste trabalho foi refletir sobre a problemática da atribuição de responsabilidade penal a agentes inteligentes, sistemas autônomos de inteligência artificial, a fim de compreender se esses são capazes de praticar ações penalmente relevantes, e, conseqüentemente, serem responsabilizados criminalmente por danos causados à bens jurídicos relevantes.

Nesse sentido, após a análise e compreensão do conceito de inteligência artificial, bem como das características dos sistemas autônomos (agentes inteligentes), traçou-se breves linhas sobre o delito, compreendendo o fundamento de cada um de seus elementos. Isso para poder compreender a possibilidade de atribuição de responsabilidade penal a essas máquinas frente ao conceito de fato punível.

Como se pode observar, novamente o direito penal se vê desafiado pelo progresso científico e tecnológico. Apesar das inúmeras contribuições da utilização destas tecnologias de potencial disruptivo, diversos são os riscos inerentes ao seu ingresso no cotidiano, impondo diversos desafios à ordem jurídica, em especial, ao campo penal.

A bem verdade, que apesar do enfrentamento já mais aprofundado no campo do direito civil, os resultados danosos causados por desvios incontrolados de sistemas autônomos de inteligência artificial ainda carecem de um debate efetivo no campo penal, sendo capaz de levar a repensar conceitos da dogmática tradicional que não são capazes de enfrentar tais questões.

As características de imprevisibilidade e autonomia dos agentes inteligentes, sendo estes capazes de captar informações do mundo externo, processá-los a partir de algoritmos e tomar decisões autônomas imprevisíveis, tornam as máquinas “opacas” por si mesmas, dificultando qualquer processo de imputação.

Como demonstrado neste trabalho, a dogmática penal tradicional parece incapaz de alcançar os resultados danosos causados por agentes inteligentes, esbarrando, logo de início, no

próprio conceito de ação. A compreensão de ação como conduta exclusivamente humana, incapaz de alcançar fatos da natureza e os animais, traz dificuldades de conceber as decisões tomadas por agentes inteligentes como condutas penalmente relevantes.

Não suficiente, também sob o aspecto da tipicidade subjetiva, a responsabilização de agentes inteligentes encontra problema. Ainda não nos parece que os agentes inteligentes tenham tamanho desenvolvimento tecnológico que permita dizer que esses tenham vontade. Todas as suas tomadas de decisão são objetivas, a partir de informação algorítmica, o que levar a crer em um vazio no elemento subjetivo, sendo impossível de se pensar em atribuição de dolo ou culpa.

Por fim, questiona-se a possibilidade de os agentes serem culpáveis. Na perspectiva atual, os agentes inteligentes também não possuem capacidade de autodeterminação, nem de se compreender a antijuridicidade da sua ação, sendo incapaz de se posicionar contra o resultado danoso. Seria impensável aceitar que os agentes inteligentes compreendam o caráter moral de suas tomadas de decisão autônomas.

Novas propostas, como as de Hallevy e de Gless e Weigend, surgem como soluções para revisitar os conceitos tradicionais de responsabilidade penal, atribuindo novos entendimentos acerca da conduta, da imputação subjetiva e da culpabilidade, a fim de que estes alcancem sistemas de inteligência artificial. Esses novos modelos não estão isentos de críticas, necessitando de maior aprofundamento e discussão, contudo, parece um início à discussão.

Portanto, a nosso ver, da perspectiva tradicional do fato punível, ainda parece impossível a atribuição de responsabilidade penal aos sistemas autônomos de inteligência artificial, mesmo em situações de violação a bens jurídicos relevantes. A solução, necessária de maior discussão, reside na possível criação de um novo direito penal para as máquinas, no qual sejam revisitados conceitos tradicionais dos elementos do crime.

## REFERÊNCIAS

AIRES DE SOUSA, Susana. *“Não fui eu, foi a máquina”: Teoria do crime, responsabilidade e inteligência artificial*. In: MIRANDA RODRIGUES, Anabella. *Inteligência artificial no direito penal*, p. 59-93, 2020.

ALVES, Natália Gontijo. *Culpabilidade Penal no Âmbito da Inteligência Artificial*. *Virtuajus*, v. 6, n. 11, p. 360-377, 2021.

BARR, Avron; FEIGENBAUM, Edward. 1981. *The handbook of artificial intelligence*. California, William Kaufmann Inc., 1981. 1 v.

BITENCOURT, Cezar. Roberto. *Tratado de direito penal: Parte Geral* (arts. 1º a 120). 29. ed. São Paulo: SaraivaJur, 2023.

CAPPELLINI, Alberto. *Machina delinquere non potest? Brevi appunti su intelligenza artificiale e responsabilità penale*. - In: DISCRIMEN. - ISSN 2704-6338. - ELETTRONICO. - (2019), pp. 1-23.

CEREZO MIR, José. *Curso de derecho penal espanol: parte general*. Imprenta: Madrid, Tecnos, 1976.

CRESPO, Marcelo. Inteligência artificial, *machine learning* e *deep learning*: relações com o Direito Penal. In: BARBOSA, Mafalda Miranda; BRAGA NETTO, Felipe; SILVA, Michael César; FALEIROS JÚNIOR, José Luiz de Moura (coord.). *Direito digital e inteligência artificial: diálogos entre Brasil e Europa*. Indaiatuba: Foco, 2021. p. 1003-1012.

DE SANCTIS, Fausto Martin. *Inteligência Artificial e Direito*. São Paulo: Almedina, 2020.

FALEIROS JÚNIOR, José Luiz de Moura. *Responsabilidade por falhas de algoritmos de inteligência artificial: ainda distantes da singularidade tecnológica, precisamos de marcos regulatórios para o tema?*. Revista de Direito da Responsabilidade, Coimbra, v. 4, p. 906-933, 2022.

FALEIROS JÚNIOR, José Luiz de Moura; VECCHI, André; GOULART, Victor Rezende. Responsabilità civile e sistemi di intelligenza artificiale: panoramica regolatoria e possibili approcci in relazione al nesso di imputazione. In: *I International Experience Perugia - Conpedi*, 2025, Perugia, Itália. GORDILHO, Heron José de Santana; LANNES, Yuri Nathan da Costa (Org.). *Inteligência Artificial: Desafios da Era Digital IV* [recurso eletrônico]. Conpedi: Florianópolis, 2025. v. 7. p. 139-157.

GLESS, Sabine; WEIGEND, Thomas. *Agentes Inteligentes e o Direito Penal*. In: ESTELLITA, Heloisa; LEITE, Alaor (organização e introdução). *Veículos Autônomos e direito penal*. 1. ed. São Paulo: Marcial Pons, 2019.

HALLEVY, Gabriel. (2024). *The Basic Models of Criminal Liability of AI Systems and Outer Circles*. In: Moura Vicente, D., Soares Pereira, R., Alves Leal, A. (eds) *Legal Aspects of Autonomous Systems*. ICASL 2022. Data Science, Machine Intelligence, and Law, vol 4. Springer, Cham. Disponível em: [https://doi.org/10.1007/978-3-031-47946-5\\_5](https://doi.org/10.1007/978-3-031-47946-5_5). Acesso em: 12 ago. 2025.

HINCKEL TEIDER, Lucas; PIVATTO DOS SANTOS, Gabriel. *Responsabilidade penal da inteligência artificial(?): a problemática relacionada ao elemento da conduta na clássica estrutura analítica do delito*: in "Derecho público y privado ante las nuevas tecnologías - Madrid :Dykinson, 2020- Casalini id: 4685145" – p. 186-195 - Permalink: Disponível em: <https://digital.casalini.it/4685551> - Casalini id: 4685551. Acesso em: 12 ago. 2025.

JANUÁRIO, Túlio Felipe Xavier. *Inteligencia artificial y responsabilidad penal de personas jurídicas: un análisis de sus aspectos materiales y procesales*. Estudios Penales y Criminológicos, v. 44, p. 1-39, 2023.

MIR PUIG, Santiago. *Fundamentos de Derecho Penal y Teoría del delito*. Editorial Reppertor. Barcelona, 2019.

MORÁN ESPINOSA, Alejandra. *Responsabilidad penal de la Inteligencia Artificial (IA). ¿La próxima frontera?*. Revista IUS, v. 15, n. 48, p. 289-323, 2021.

MUÑOZ CONDE, Francisco; GARCÍA ARÁN; Mercedes. *Derecho penal: parte general*. 7ª edición, revisada y puesta al día. Imprenta: Valencia, Tirant lo Blanch, 2007.

PACELLI, Eugenio; CALLEGARI, André. *Manual de direito penal: parte geral*. 4. ed. rev., atual. e ampl. São Paulo: Atlas, 2018.

PAULA, Alice Lima; CORNWALL, Bruno Meirelles de M.; CABRAL, Dalila M. *Breves reflexões sobre a inteligência artificial e seus impactos no campo do Direito Penal*. In: CHAVES, Natália Cristina (org.). *Direito, tecnologia e globalização*. [online]. Porto Alegre, 2019. p. 98-117. Disponível em: [https://www.direito.ufmg.br/wp-content/uploads/2019/12/direito\\_tecnologia\\_globalizacao.pdf](https://www.direito.ufmg.br/wp-content/uploads/2019/12/direito_tecnologia_globalizacao.pdf). Acesso em 15 ago. 2025.

ROXIN, Claus. *Direito penal: parte geral: tomo I: fundamentos: a estrutura da teoria do crime*. Org: Luís Greco e Alaor Leite. São Paulo: Marcial Pons, 2024.

SANTOS, Juarez Cirino dos. *Direito Penal: parte geral*. 10. ed. rev. atual e ampl. São Paulo: Tirant lo Blanch, 2022.

TAVARES, Juarez. MARTINS, Antonio. *Fundamentos de Teoria do Delito*. 5. ed. rev. e atual. São Paulo: Tirant Lo Blanch, 2025.