II ENCONTRO NACIONAL DE DIREITO DO FUTURO - II ENDIF

INTELIGÊNCIA ARTIFICIAL, DIREITO E REGULAÇÃO I

I61
Inteligência artificial, direito e regulação I [Recurso eletrônico on-line] organização II Encontro Nacional de Direito do Futuro: Escola Superior Dom Helder Câmara – Belo Horizonte;

Coordenadores: Marco Antônio Sousa Alves e Fernanda dos Santos Rodrigues Silva – Belo Horizonte: Escola Superior Dom Helder Câmara - ESDHC, 2025.

Inclui bibliografia
ISBN: 978-65-5274-403-6
Modo de acesso: www.conpedi.org.br em publicações
Tema: Justiça social e tecnológica em tempos de incerteza.
1. Direito do Futuro. 2. Justiça Social. 3. Justiça Tecnológica. I. II Encontro Nacional de Direito do Futuro (1:2025 : Belo Horizonte, MG).

CDU: 34

Dom Helder
CENTRO UNIVERSITÁRIO

II ENCONTRO NACIONAL DE DIREITO DO FUTURO - II ENDIF

INTELIGÊNCIA ARTIFICIAL, DIREITO E REGULAÇÃO I

Apresentação

O II Encontro Nacional de Direito do Futuro (II ENDIF), organizado pelo Centro Universitário Dom Helder com apoio técnico do Conselho Nacional de Pesquisa e Pós-graduação em Direito – CONPEDI, reafirma-se como um espaço qualificado de produção, diálogo e circulação do conhecimento jurídico, reunindo a comunidade científica em torno de um propósito comum: pensar, com rigor metodológico e sensibilidade social, os caminhos do Direito diante das transformações que marcam o nosso tempo. Realizado nos dias 09 e 10 de outubro de 2025, em formato integralmente on-line, o evento assumiu como tema geral "Justiça social e tecnológica em tempos de incerteza", convidando pesquisadoras e pesquisadores a enfrentar criticamente os impactos da inovação tecnológica, das novas dinâmicas sociais e das incertezas globais sobre as instituições jurídicas e os direitos fundamentais.

Nesta segunda edição, os números evidenciam a força do projeto acadêmico: 408 trabalhos submetidos, com a participação de 551 pesquisadoras e pesquisadores, provenientes de 21 Estados da Federação, culminando na organização de 31 e-books, que ora se apresentam à comunidade científica. Essa coletânea traduz, em linguagem acadêmica e compromisso público, a vitalidade de uma pesquisa jurídica que não se limita a descrever problemas, mas busca compreendê-los, explicar suas causas e projetar soluções coerentes com a Constituição, com os direitos humanos e com os desafios contemporâneos.

A publicação dos 31 e-books materializa um processo coletivo que articula pluralidade temática, densidade teórica e seriedade científica. Os textos que compõem a coletânea passaram por avaliação acadêmica orientada por critérios de qualidade e imparcialidade, com destaque para o método double blind peer review, que viabiliza a análise inominada dos trabalhos e exige o exame por, no mínimo, dois avaliadores, reduzindo subjetividades e preferências ideológicas. Essa opção metodológica é, ao mesmo tempo, um gesto de respeito à ciência e uma afirmação de que a pesquisa jurídica deve ser construída com transparência, responsabilidade e abertura ao escrutínio crítico.

O II ENDIF também se insere em uma trajetória institucional já consolidada: a primeira edição, realizada em junho de 2024, reuniu centenas de pesquisadoras e pesquisadores e resultou na publicação de uma coletânea expressiva, demonstrando que o Encontro se consolidou, desde o início, como um dos maiores eventos científicos jurídicos do país. A

continuidade do projeto, agora ampliada em escopo e capilaridade, reafirma a importância de se fortalecer ambientes acadêmicos capazes de integrar graduação e pós-graduação, formar novas gerações de pesquisadoras e pesquisadores e promover uma cultura jurídica comprometida com a realidade social.

A programação científica do evento, organizada em painéis temáticos pela manhã e Grupos de Trabalho no período da tarde, foi concebida para equilibrar reflexão teórica, debate público e socialização de pesquisas. Nos painéis, temas como inteligência artificial e direitos fundamentais, proteção ambiental no sistema interamericano, proteção de dados e herança digital foram tratados por especialistas convidados, em debates que ampliam repertórios e conectam a produção acadêmica aos dilemas concretos vividos pela sociedade.

A programação científica do II ENDIF foi estruturada em dois dias, 09 e 10 de outubro de 2025, combinando, no período da manhã, painéis temáticos com exposições de especialistas e debates, e, no período da tarde, sessões dos Grupos de Trabalho. No dia 09/10 (quinta-feira), após a abertura, às 09h, realizou-se o Painel I, dedicado aos desafios da atuação processual diante da inteligência artificial ("Inteligencia artificial y desafios de derechos fundamentales en el marco de la actuación procesal"), com exposição de Andrea Alarcón Peña (Colômbia) e debate conduzido por Caio Augusto Souza Lara. Em seguida, às 11h, ocorreu o Painel II, voltado à proteção ambiental no Sistema Interamericano, abordando a evolução da OC-23 ao novo marco da OC-32, com participação de Soledad Garcia Munoz (Espanha) e Valter Moura do Carmo como palestrantes, sob coordenação de Ricardo Stanziola Vieira. No período da tarde, das 14h às 17h, desenvolveram-se as atividades dos Grupos de Trabalho, em ambiente virtual, com apresentação e discussão das pesquisas aprovadas.

No dia 10/10 (sexta-feira), a programação manteve a organização: às 09h, foi realizado o Painel III, sobre LGPD e a importância da proteção de dados na sociedade de vigilância, com exposições de Laís Furuya e Júlia Mesquita e debate conduzido por Yuri Nathan da Costa Lannes; às 11h, ocorreu o Painel IV, dedicado ao tema da herança digital e à figura do inventariante digital, com apresentação de Felipe Assis Nakamoto e debate sob responsabilidade de Tais Mallmann Ramos. Encerrando o evento, novamente no turno da tarde, das 14h às 17h, seguiram-se as sessões dos Grupos de Trabalho on-line, consolidando o espaço de socialização, crítica acadêmica e amadurecimento das investigações apresentadas.

Ao tornar públicos estes 31 e-books, o II ENDIF reafirma uma convicção essencial: não há futuro democrático para o Direito sem pesquisa científica, sem debate qualificado e sem compromisso com a verdade metodológica. Em tempos de incerteza — tecnológica, social,

ambiental e institucional —, a pesquisa jurídica cumpre um papel civilizatório: ilumina problemas invisibilizados, questiona estruturas naturalizadas, qualifica políticas públicas, tensiona o poder com argumentos e oferece horizontes normativos mais justos.

Registramos, por fim, nosso reconhecimento a todas e todos que tornaram possível esta obra coletiva — autores, avaliadores, coordenadores de Grupos de Trabalho, debatedores e equipe organizadora —, bem como às instituições e redes acadêmicas que fortalecem o ecossistema da pesquisa em Direito. Que a leitura desta coletânea seja, ao mesmo tempo, um encontro com o que há de mais vivo na produção científica contemporânea e um convite a seguir construindo, com coragem intelectual e responsabilidade pública, um Direito à altura do nosso tempo.

Belo Horizonte-MG, 16 de dezembro de 2025.

Prof. Dr. Paulo Umberto Stumpf – Reitor do Centro Universitário Dom Helder

Prof. Dr. Franclim Jorge Sobral de Brito – Vice-Reitor e Pró-Reitor de Graduação do Centro Universitário Dom Helder

Prof. Dr. Caio Augusto Souza Lara – Pró-Reitor de Pesquisa do Centro Universitário Dom Helder

AGENTES DE IA (IN)SEGURANÇA CIBERNÉTICA: DESAFIOS SOCIAIS E REGULATÓRIOS À LUZ DO PROJETO LEI Nº 2.338/2023

AI AGENTS AND CYBER(IN)SECURITY: SOCIAL AND REGULATORY CHALLENGES IN LIGHT OF DRAFT LAW N. 2.338/2023

Gabriel Cemin Petry [1]
Murilo Haupenthal

Resumo

Agentes de IA constituem um dos avanços tecnológicos mais recentes, trazendo simultaneamente oportunidades e riscos relevantes. O problema central da pesquisa concentra-se nas implicações dessa tecnologia para a cibersegurança. Adotando o método dedutivo e a pesquisa bibliográfica, o estudo analisa conceitos técnicos e jurídicos sobre agentes de IA, os riscos cibernéticos envolvidos e os mecanismos regulatórios previstos no Projeto de Lei nº 2.338/23. Em conclusão, embora o PL estabeleça uma base relevante para a regulação, mas sua eficácia dependerá da articulação entre normas jurídicas e soluções técnicas que garantam segurança, transparência e confiança social.

Palavras-chave: Inteligência artificial, Regulação, Cibersegurança

Abstract/Resumen/Résumé

AI agents are one of the most recent technological advances, bringing both opportunities and significant risks. The central problem of the research focuses on the implications of this technology for cybersecurity. Adopting the deductive method and bibliographic research, the study analyzes technical and legal concepts about AI agents, the cyber risks involved, and the regulatory mechanisms provided for in Bill No. 2,338/23. In conclusion, although the bill establishes a relevant basis for regulation, its effectiveness will depend on the articulation between legal norms and technical solutions that guarantee security, transparency, and social trust.

Keywords/Palabras-claves/Mots-clés: Artificial intelligence, Regulation, Cybersecurity

[1] Mestrando em Direito pela Universidade Vale do Rio dos Sinos – UNISINOS. Bolsista CAPES/PROEX. Bacharel em Direito pela Universidade Feevale. Colaborador Editorial na MICHR, Itália. Advogado. E-mail: gabrielcpetry96@gmail.com.

1 INTRODUÇÃO

A emergência de sistemas agênticos, comumente chamados de agentes de Inteligência Artificial (IA), representa uma das evoluções mais significativas da tecnologia atual. Estes agentes são caracterizados pela capacidade de operar de forma autônoma, realizar tarefas complexas, adaptar-se e interagir com ambientes físicos ou virtuais, apresentando graus variados de independência. Tal novo paradigma técnico levanta desafios éticos, jurídicos e, de forma proeminente, relacionados à questão da cibersegurança, exigindo uma regulamentação que acompanhe a rapidez das inovações. No Brasil, o Projeto de Lei (PL) nº 2.338/2023 busca estabelecer uma estrutura normativa para o uso ético e responsável da IA introduzindo a figura legal dos "agentes de IA".

Neste contexto, a pesquisa orbita entorno da seguinte questão: se os agentes de IA podem influenciar a segurança cibernética e se o PL nº 2.338/2023 oferece salvaguardas ligadas à cibersegurança. O objetivo central deste trabalho é descrever os agentes de IA e analisar suas implicações na cibersegurança, sob a ótica do referido PL. Como objetivos específicos, busca-se: (i) apresentar conceitos técnicos e legais sobre agentes de IA, incluindo classificações e exemplos de uso; (ii) discutir os riscos desses sistemas, com foco em sua utilização como vetores para exploração de vulnerabilidades e ataques cibernéticos; e (iii) analisar os dispositivos do PL que abordam a segurança cibernética. A pesquisa é um produto de atividade acadêmica “Direito, Tecnologia e Inovação”, realizada no programa de Pós-Graduação em Direito da Universidade do Vale do Rio dos Sinos – UNISINOS.

O método de pesquisa é dedutivo, fundamentado em pesquisa bibliográfica, com base em fontes legislativas, doutrinária e em documentos internacionais, como o *Workshop Report “Through the Chat Window and Into the Real World: preparing for AI Agents”* do *Center for Security and Emerging Technology* (CSET). Estruturalmente, o trabalho, primeiro, examina a conceituação dos “agentes de IA”, sob as perspectivas tecnológica e jurídica, e suas capacidades e, segundo, foca na análise dos riscos cibernéticos associados a esses agentes e examina os mecanismos de governança e segurança propostos pelo PL 2.338/2023.

2 AGENTES DE IA E ATAQUES CIBERNÉTICOS: CONCEITOS, RISCOS E REGULAÇÃO DA CIBERSEGURANÇA NO PL Nº 2.338/2023

O campo da Inteligência Artificial (IA) tem evoluído para o surgimento de “agentes artificiais” capazes de planejar e executar tarefas complexas, com reduzida intervenção humana (Kolt, 2025, p. 4), operando autonomamente. O surgimento destas nova tecnologia apresenta tanto oportunidades (*v.g.* “*AI personal assistants*” ou “*virtual coworkers*”) quanto riscos significativos, especialmente no que tange à segurança cibernética das organizações (Kolt, 2025, p. 4; Gutowska, 2025). Para compreensão do termo, dois caminhos se apresentam: (a) o normativo/legislativo; (b) o tecnológico/técnico.

Sob o ponto de vista normativo, o Projeto Lei nº 2.338/2023 define "agentes de IA" como "os desenvolvedores, distribuidores e aplicadores que atuem na cadeia de valor e na governança interna de sistemas de IA" (art. 4º, inc. VIII) (Pacheco, 2023). Estes agentes são os responsáveis por criar, disponibilizar e utilizar sistemas de IA, devem cumprir obrigações específicas, conforme se depreende dos artigos 7º a 12 do PL (Pacheco, 2023) – o termo, aliás, lembra o termo “agentes de tratamento”, constantes na Lei Geral de Proteção de Dados (LGPD).

Contudo, sob o ponto de vista tecnológico, um "agente" pode ser definido como algo que percebe e age sobre seu ambiente (Russel; Norvig, 2022, p. 54), recebendo dados, realizando ações autônomas para atingir metas e melhorando seu desempenho pelo aprendizado (Floridi, 2022, pp. 49-50). Embora agentes computacionais sejam algoritmos programados, sua autonomia crescente, impulsionada por *Large Languages Model* (LLMs) e e *Multimodel Large Langueages Model* (MLLMs), gerou os denominados "Agentes de IA" (AIAs) e até "*multi-agentes*" (Braga; Henriques, 2025, pp. 2-5). A responsabilidade por danos de agentes programados recai sobre seus criadores (Braga; Henriques, 2025, p. 2).

Quatro características indicam um sistema mais agêntico: complexidade dos objetivos, complexidade do ambiente, planejamento e adaptação independente, e ação direta (Toner et al., 2024, p. 12). Gutowska (2025), por sua vez, classifica os agentes de IA em cinco tipos: agentes de reflexo simples, de reflexo baseados em modelos, baseados em objetivos, baseados em utilidade e de aprendizagem. Em síntese, tais sistemas agênticos se valem da IA para tarefas autônomas e específicas, com autonomia limitada por algoritmos programados por humanos, circunstância que implica responsabilidade dos desenvolvedores, distribuidores e aplicadores (Cf. art. 4º, inc. VIII, do PL 2.338/23) (Fang et al., 2024, p. 2).

A despeito das novas oportunidades que os agentes de IA oferecem, parece igualmente evidente que a IA também é um vetor de risco na cibersegurança. Neste sentido, Kolt (2025, p. 4) alerta para a exploração de agentes de IA por atores maliciosos para automatizar ciberataques e perpetrar fraudes. Pesquisas demonstraram que agentes de LLM podem autonomamente *hackear websites* (Fang et al., 2024, pp. 1-11) e que "times de agentes de LLM" podem explorar vulnerabilidades *Zero-Day* (Fang et al., 2024, p. 8), acelerando ataques cibernéticos – não se olvide, contudo, de que tal tecnologia também encontra oportunidades na defesa cibernética. Portanto, é evidente que, na era dos sistemas agênticos, diversas oportunidades e problemáticas se apresentam.

Vale dizer ainda que a interação de múltiplos agentes pode propagar falhas rapidamente e criar redes complexas e opacas, aumentando a "zona de exploração de vulnerabilidades" para atacantes (Kolt, 2025, pp. 16-17). Para falar em responsabilização, Braga e Henriques (2025, pp. 2 e 6) sugerem, em cenários de multiagentes, protocolos de identificação e autenticação de algoritmos, o que possibilitaria um “rastreamento” até os responsáveis pelo desenvolvimento. Seja como for, está claro que agentes de IA também são um vetor de risco significativo, ampliando superfícies de ataque e vulnerabilidades, do que decorre a necessidade de criação de salvaguardas regulatórias, normativas e técnicos para assegurar integridade, confiabilidade e responsabilização – em suma: um princípio de (ciber)segurança (Petry; Hupffer, 2023, p. 89).

A regulação, portanto, torna-se essencial. O PL 2.338/23 estabelece que o desenvolvimento da IA deve estar fundado na segurança da informação e cibersegurança (Art. 2º), de modo que, caso determinado sistema de IA (agêntico ou não) seja inseguro (ou não confiável), a responsabilidade por danos dele decorrentes recairá sobre o seu criador (Braga; Henriques, 2025, pp. 2-5). Para segurança e transparência, "agentes de IA" devem passar por uma "avaliação preliminar" (Cf. arts. 12 a 29, do PL 2.338/23) e, em casos de alto risco (especialmente em tratando-se de infraestruturas críticas – conforme o art. 14, inc. I), avaliações de impacto algorítmico (nos termos dos arts. 25 e 27). Vale destacar ainda que o PL ainda prevê exceção ao direito de informação para sistemas de IA dedicados à cibersegurança (art. 5º, inc. I), o que sucede, possivelmente, para proteger o funcionamento dos sistemas de defesa.

Para fortalecer a segurança e a responsabilização, a adoção de apoio técnico (*technical guardrails*) é crucial (Toner et al., 2024, pp. 17-22). O *Workshop Report* do CSET destaca quatro categorias de apoio: (i) visibilidade (identificação, monitoramento, logs); (ii) controle (interruptibilidade, reversibilidade, controle de acessos); (iii) confiabilidade (*human-in-the-*

loop, explicabilidade); e (iv) privacidade e segurança (codificação segura, testes adversariais, criptografia). O PL 2.338/23 incorpora medidas semelhantes, como exigência de documentação, testes de segurança e comunicação de incidentes graves, sujeitando os agentes de IA (nos termos conceituados pela norma) à responsabilização civil e a observação da legislação vigente, ainda que esparsa, que trate de matérias ligadas à cibersegurança, proteção de dados (pessoais ou não) e defesa do usuário/consumidor.

4 CONSIDERAÇÕES FINAIS

Os agentes de IA representam uma inovação técnica significativa, que propicia tanto novas oportunidades econômicas, quanto novos vetores de riscos no campo da cibersegurança. A autonomia e complexidade desses sistemas (e a interação entre eles) permitem sua exploração maliciosa para ciberataques, como invasões de sites e ataques do tipo *Zero-Day*. Essa realidade desponta na inegável conclusão de que regras de cibersegurança devem ser observadas, auxiliando na posterior responsabilização por danos causados por negligências de segurança de programadores, desenvolvedores e implementadores.

O Projeto de Lei nº 2.338/2023, ao conceituar agentes de IA como sujeitos jurídicos responsáveis, oferece uma base normativa inicial para a regulação do uso e desenvolvimento de sistemas de IA no Brasil. Embora não defina explicitamente o que são "sistemas agênticos", o PL reconhece a segurança da informação e a cibersegurança como fundamentos para o uso ético e responsável da IA. Medidas como a avaliação preliminar de risco, os estudos de impacto algorítmico, a documentação e o monitoramento contínuo demonstram um avanço regulatório focado na cibersegurança dos sistemas de IA.

Contudo, a efetividade do PL 2.338/2023 dependerá da implementação de diretrizes claras e de procedimentos fiscalizatórios, especialmente em cenários ainda não detalhados, como de sistemas agênticos. Relatórios especializados destacam que a governança de sistemas agênticos exige a integração de normas legais com soluções técnicas que garantam visibilidade, controle, confiabilidade e segurança. Tais medidas, além de fortalecerem a privacidade, a segurança cibernética e a confiança do usuário/consumidor, promovem práticas mais transparentes que auxiliam na apuração de responsabilidade dos desenvolvedores, distribuidores e aplicadores dessas tecnologias.

Assim, conclui-se afirmativamente que os agentes de IA, enquanto sistemas agênticos, podem, contribuir para a ocorrência de ciberataques, ao ampliarem as superfícies de ataque e

operarem de forma autônoma. No entanto, esses sistemas também oferecem novas oportunidades na defesa de ataques cibernéticos. O Projeto de Lei nº 2.338/2023, mesmo sem abordar os sistemas agênticos de forma específica, apresenta dispositivos que podem ser aplicados ao seu uso e desenvolvimento, sendo relevantes para a mitigação de riscos, especialmente sob a perspectiva da cibersegurança.

REFERÊNCIAS

BRAGA, Luiz Juliao; HENRIQUES, Percival. **Agentes computacionais.** 3 fev. 2025. Disponível em: https://doi.org/10.31219/osf.io/92xh6_v1. Acesso em: 29 mai. 2025.

FANG, Richard et. al. LLM Agents Can Autonomously Hack Websites, **ARXIV**, 16. Fev. 2024. Disponível em: https://arxiv.org/abs/2402.06664. Acesso em: 1 jun. 2025.

FANG, Richard et. al. Teams of LLM Agents can Exploit Zero-Day Vulnerabilities, **ARXIV**, 16. Fev. 2024. Disponível em: https://arxiv.org/pdf/2406.01637. Acesso em: 1 jun. 2025.

FLORIDI, Luciano. **Etica dell'intelligenza artificiale**: sviluppi, opportunità, sfide. Milano: R. Cortina, 2022.

GUTOWSKA, Anna. O que são agentes de IA? **IBM Think**. Disponível em: https://www.ibm.com/br-pt/think/topics/ai-agents. Acesso em 30 mai. 2025.

KOLT, Noam. Governing AI Agents. **ARXIV**, 11 fev. 2025. Disponível em: https://arxiv.org/pdf/2501.07913. Acesso em: 29 mai. 2025.

PACHECO, Rodrigo. **Projeto de Lei n°. 2.338, de 2023**. Dispõe sobre o uso da Inteligência Artificial. Senado Federal. Disponível em: https://www25.senado.leg.br/web/atividade/materias/-/materia/157233. Acesso em 30 mai. 2025.

PETRY, Gabriel Cemin; HUPFFER, Haide Maria. O princípio da segurança na era dos ciberataques: uma análise a partir do escopo protetivo da LGPD. **Revista do CNJ**, v. 7, n. 1, 2023.

RUSSELL, Stuart J.; NORVIG, Peter. **Artificial intelligence:** a modern approach. Fourth edition, global edition. Harlow: Pearson, 2022.

TONER, Helen et. al. **Through the Chat Windo and Into the Real World: preparing for AI Agents.** Georgetown: Center for security and emerging technology, October 2024. Disponível em: https://cset.georgetown.edu/publication/through-the-chat-window-and-into-the-real-world-preparing-for-ai-agents/. Acesso em: 29 mai. 2025.