

UNIVERSIDADE FEDERAL DE MINAS GERAIS
FACULDADE DE DIREITO E CIÊNCIAS DO ESTADO

PEDRO LUIZ DE JESUS BELETABLA BRAVO

**A NECESSIDADE DE INTERVENÇÃO HUMANA NOS MODELOS DE
INTELIGÊNCIA ARTIFICIAL: UMA ANÁLISE DA SEÇÃO III DO PROJETO DE LEI
Nº 2.338/2023**

Belo Horizonte
2025

Trabalho de Conclusão de Curso
Relatório de Iniciação Científica

**A NECESSIDADE DE INTERVENÇÃO HUMANA NOS MODELOS DE
INTELIGÊNCIA ARTIFICIAL: UMA ANÁLISE DA SEÇÃO III DO PROJETO DE LEI
Nº 2.338/2023**

**Relatório de Iniciação Científica apresentado como
Trabalho de Conclusão de Curso da Faculdade de
Direito e Ciências do Estado da Universidade
Federal de Minas Gerais, como requisito para a
obtenção do diploma de bacharel em Ciências do
Estado.**

Orientadora: Profa. Titular Mariah Brochado

Belo Horizonte
2025

1. INTRODUÇÃO.....	4
2. REVISÃO TEÓRICA.....	5
2.1. O Cenário de Risco e a Necessidade de Hard Law.....	5
2.2. A Arquitetura de Direitos da Seção III do PL 2.338/2023: Explicação e Intervenção Humana.....	6
2.3. O Paradigma Europeu (The Brussels Effect).....	6
3. OBJETIVOS PERCORRIDOS.....	7
3.1. Objetivo Geral: A Evolução Legislativa (Do PL 21/2020 ao PL 2.338/2023)..	7
3.2. Objetivos Específicos e Resultados.....	8
4. METODOLOGIA DA PESQUISA.....	9
5. RESULTADOS E DISCUSSÃO.....	10
5.1. Problemáticas da IA sem Intervenção Humana Significativa.....	10
5.2. O Sistema de Risco da UNESCO e o PL 2.338.....	11
5.3. Inspirações no AI Act e o Gabinete de IA Europeu.....	12
5.4. Do Direito à Filosofia: O Projeto de Mestrado.....	13
6. CONSIDERAÇÕES FINAIS.....	13
REFERÊNCIAS.....	15

1. INTRODUÇÃO

A forte presença da Inteligência Artificial (IA) nos processos decisórios contemporâneos inaugurou uma nova era na gestão da vida social, econômica e política¹. De diagnósticos médicos à concessão de crédito, passando pelo policiamento preditivo, algoritmos assumiram o papel de árbitros de direitos fundamentais. Contudo, essa “dataficação” da existência traz consigo o risco da opacidade: o fenômeno das *black boxes*² (caixas-pretas), onde decisões críticas são tomadas por lógicas matemáticas inescrutáveis, desafiando os pilares do Estado Democrático de Direito³.

A presente Iniciação Científica (IC), desenvolvida entre setembro de 2024 e outubro de 2025, dedicou-se a investigar a resposta jurídica brasileira a este desafio, concentrando-se na análise da Seção III do Projeto de Lei (PL) nº 2.338 de 2023. O foco central recaiu sobre a eficácia dos instrumentos de contestabilidade e, fundamentalmente, sobre a necessidade de intervenção humana (*human-in-the-loop*)⁴ em decisões automatizadas.

A pesquisa partiu da premissa de que a eficiência algorítmica não pode suplantar a garantia do devido processo legal. Em um cenário onde a automatização promete celeridade, a supervisão humana surge não apenas como uma barreira de segurança técnica, mas como um imperativo ético para preservar a dignidade humana contra a frieza do cálculo probabilístico.

Metodologicamente, o estudo adotou uma abordagem qualitativa de natureza bibliográfico-documental. Foram analisados textos legislativos nacionais e

¹ Como observa a literatura recente sobre IA e big data, foi sobretudo a partir da década de 2010, com o avanço do aprendizado profundo e a disponibilidade massiva de dados, que os sistemas passaram a ser empregados em decisões de alto impacto em crédito, saúde, segurança pública e gestão de pessoas, ampliando de forma inédita o alcance da automatização decisória (cf. Santos, 2024; Matias, 2024).

² Frank Pasquale, em *The Black Box Society* (2015), descreve a “sociedade da caixa-preta” para se referir a um contexto em que decisões relevantes são tomadas por algoritmos opacos, difíceis de auditar e de contestar, o que cria assimetrias de poder entre desenvolvedores, empresas e cidadãos.

³ Como enfatizam diretrizes e relatórios da União Europeia e da OCDE, o uso de IA em processos decisórios sensíveis exige salvaguardas de transparência, possibilidade de contestação e responsabilização, justamente para compatibilizar a inovação tecnológica com o Estado Democrático de Direito (European Commission, 2019; OECD, 2019).

⁴ Como destacam diretrizes éticas europeias sobre IA confiável, a presença de um humano “na”, “sobre” ou “no comando” do circuito decisório é condição para preservar agência humana, prevenir abusos e viabilizar a imputação de responsabilidade por erros e danos (High Level Expert on AI, 2019).

internacionais, documentos de organismos multilaterais (UNESCO, OCDE) e literatura especializada em Filosofia do Direito e da Tecnologia.

Os resultados alcançados demonstram que, embora o PL 2.338/2023 represente um avanço regulatório robusto ao adotar uma abordagem baseada em risco, persistem lacunas significativas quanto à operacionalização da supervisão humana. A pesquisa identificou que a “qualidade” dessa intervenção é de difícil mensuração quantitativa, o que redirecionou o estudo para uma análise mais profunda sobre a natureza qualitativa e ética do juízo humano — descoberta esta que serviu de base para a elaboração do projeto de Dissertação de Mestrado apresentado no decorrer dos últimos meses⁵.

2. REVISÃO TEÓRICA

O arcabouço teórico desta pesquisa situa-se na interseção entre o pensamento jurídico e a filosofia da tecnologia, utilizando os documentos analisados para construir uma crítica à automação irrestrita.

2.1. O Cenário de Risco e a Necessidade de *Hard Law*

Relatórios internacionais, como o estudo conjunto da Trend Micro, UNICRI e Europol (*Malicious Uses and Abuses of Artificial Intelligence, 2020*), evidenciam que a IA, se não regulada, potencializa ameaças que vão desde a manipulação de opinião pública via *deepfakes*⁶ até a amplificação de discriminações estruturais. A tecnologia deixa de ser uma ferramenta neutra para tornar-se um vetor de risco social, exigindo o que a doutrina chama de “design voltado para a proteção jurídica” (*Legal Protection by Design*). É dever do Direito regular novos acontecimentos que vem a ocorrer como desdobramentos dessa era cibernética (Brochado, 2021).

⁵ Como será detalhado adiante, a constatação de que não há critérios objetivos e consensuais para mensurar quantitativamente a eficácia da intervenção humana em processos decisórios automatizados com IA conduziu a pesquisa a uma inflexão qualitativa: mais do que perguntar se o humano “melhora” o resultado, passou-se a indagar se não seriam justamente características intrínsecas da racionalidade humana – como a capacidade ética, a prudência e o juízo moral – que tornam essa intervenção indispensável e, em certa medida, indelegável a sistemas algorítmicos.

⁶ Os chamados *deepfakes* são conteúdos sintéticos (principalmente imagens, áudios e vídeos) gerados ou manipulados por técnicas avançadas de inteligência artificial, capazes de simular de forma verossímil a aparência ou a voz de uma pessoa, o que permite, por exemplo, atribuir-lhe falas e comportamentos que nunca ocorreram na realidade, com impactos relevantes para honra, privacidade e integridade dos processos democráticos.

2.2. A Arquitetura de Direitos da Seção III do PL 2.338/2023: Explicação e Intervenção Humana

A fundamentação normativa da pesquisa, assim como seu objetivo geral, concentrou-se na análise detalhada da Seção III do PL nº 2.338/2023, intitulada “Dos Direitos das Pessoas Afetadas por Sistemas de Inteligência Artificial”. Este segmento legislativo (Arts. 9º a 11) institui o que a doutrina denomina de “devido processo legal tecnológico”⁷, estabelecendo garantias essenciais frente ao poder decisório dos algoritmos.

A pesquisa destacou dois pilares centrais nesta seção:

- a) Direito à Explicação (Art. 10): A obrigação de que os fornecedores ou operadores de IA forneçam informações claras e adequadas sobre os critérios e procedimentos utilizados na decisão automatizada, garantindo a inteligibilidade da lógica algorítmica.
- b) Direito à Contestação e Revisão Humana (Art. 11): O dispositivo mais crítico para este estudo, que assegura à pessoa afetada o direito de solicitar a intervenção humana para revisar decisões que produzam efeitos jurídicos relevantes ou impactem significativamente seus interesses.

Esta arquitetura legal busca mitigar a assimetria de informação entre o indivíduo e a máquina. Contudo, a revisão teórica apontou que a eficácia destes direitos depende da interpretação do que constitui “efeito jurídico relevante” e da capacidade técnica real de um humano reverter uma decisão baseada em correlações estatísticas complexas, ponto de tensão central investigado neste trabalho.

2.3. O Paradigma Europeu (*The Brussels Effect*)

A regulação brasileira não ocorre em um vácuo geopolítico, mas reflete um fenômeno de harmonização global descrito pela jurista Anu Bradford como o “Efeito

⁷ Parte da literatura defende que, em contextos de decisão automatizada, garantias como contraditório, ampla defesa e motivação precisam ser reinterpretadas para assegurar “explicabilidade algorítmica” e possibilidade real de contestação, sob pena de esvaziamento material do devido processo (cf. PASQUALE, 2015; HILDEBRANDT, 2020).

Bruxelas” (*The Brussels Effect*)⁸. Esta teoria explica a capacidade da União Europeia de unilateralmente exportar seus padrões regulatórios, transformando-os em normas globais de fato, dada a necessidade de interoperabilidade dos mercados digitais. No Brasil, a influência desse paradigma é visível na drástica evolução legislativa observada entre 2020 e 2023⁹.

3. OBJETIVOS PERCORRIDOS

3.1. Objetivo Geral: A Evolução Legislativa (Do PL 21/2020 ao PL 2.338/2023)

O objetivo geral desta investigação consistiu em analisar criticamente o Projeto de Lei nº 2.338 de 2023, novo marco regulatório da Inteligência Artificial no Brasil, com foco específico na sua Seção III (“Dos Direitos das Pessoas Afetadas por Sistemas de Inteligência Artificial”). A pesquisa buscou examinar a eficácia normativa dos mecanismos que regulamentam o direito de contestação e de solicitação de intervenção humana em decisões automatizadas.

A finalidade central desta análise foi verificar a imperatividade da supervisão humana (*human-in-the-loop*)¹⁰ não apenas como uma salvaguarda técnica, mas como uma exigência decorrente da natureza intrínseca do julgamento jurídico — que pressupõe a ponderação ética e contextual inacessível a modelos estocásticos.

Para compreender a densidade destes novos direitos, o estudo contextualizou a evolução legislativa da matéria. Observou-se a transição do paradigma principiológico do antigo PL nº 21/2020¹¹ — que priorizava o fomento à inovação com diretrizes genéricas — para o modelo robusto do PL nº 2.338/2023.

⁸ Anu Bradford argumenta que a União Europeia, ao estabelecer padrões regulatórios rigorosos em áreas como proteção de dados e mercados digitais, acaba projetando suas regras globalmente, na medida em que empresas e países terceiros se adaptam para manter acesso ao mercado europeu (BRADFORD, 2020).

⁹ A trajetória que vai do PL 21/2020 ao PL 2.338/2023 reflete, como ressalta a própria Comissão de Juristas do Senado, um deslocamento de uma regulação mais principiológica e voltada ao fomento para um modelo inspirado em experiências estrangeiras recentes, com maior densidade em termos de direitos e deveres (CJSUBIA, 2022).

¹⁰ A expressão *human-in-the-loop* foi utilizada inicialmente em áreas de engenharia e ciência da computação para designar sistemas em que um operador humano intervém em fases críticas do processo; no debate jurídico e regulatório, passou a designar a exigência de “supervisão humana significativa” como condição para a legitimidade e a responsabilização em decisões automatizadas de alto impacto.

¹¹ Essa aposta inicial em diretrizes éticas e *soft law* acompanhava o movimento de diversos países, mas, como apontam análises críticas sobre discriminação algorítmica e opacidade em crédito, trabalho e segurança pública, mostrou-se insuficiente para lidar com danos concretos (cf. SAINZ, 2024; CUSCIANO, 2024).

Este último, fruto do trabalho da Comissão de Juristas (CJSUBIA), adota uma abordagem baseada em riscos e direitos, institucionalizando a transparência, a *accountability* e o devido processo legal tecnológico como princípios basilares para mitigar a opacidade das “caixas-pretas” — como chama Frank Pasquale (2015) — algorítmicas.

3.2. Objetivos Específicos e Resultados

A) Analisar a Seção III do PL 2.338/2023 (Alcançado): Mapeou-se os artigos 9º, 10º e 11º, que garantem o direito à explicação e à revisão humana. Identificou-se que a eficácia destes dispositivos depende de uma “alfabetização algorítmica” dos cidadãos e de regulamentação infralegal pela autoridade competente.

B) Verificar a compatibilidade com a LGPD (Alcançado): A pesquisa identificou uma tensão normativa. O Art. 20 da LGPD garante a revisão de decisões automatizadas¹². O PL 2.338/2023, por sua vez, condiciona a revisão humana a decisões que produzam “efeitos jurídicos relevantes” ou “impacto significativo”. Concluiu-se que o PL deve ser interpretado como norma especial que complementa, e não restringe, as garantias da LGPD, sob pena de retrocesso social.

C) Investigar a métrica de qualidade da intervenção humana (Não Alcançado/Reorientado): Este objetivo revelou a principal descoberta epistemológica da pesquisa. Não foi possível encontrar ou desenvolver uma metodologia quantitativa capaz de medir a “qualidade” da intervenção humana. A eficácia do juízo humano diante de uma máquina não se mede por tempo de resposta ou taxa de cliques, mas por critérios qualitativos de independência e cognição. Essa impossibilidade técnica de quantificação redirecionou a pesquisa para a filosofia, culminando na proposta de Mestrado.

D) Formular recomendações de aprimoramento (parcialmente alcançado): Com base no *AI Act* europeu e nas lacunas identificadas, foram propostas diretrizes para mitigar o “viés de automação” (a tendência humana de aceitar cegamente a sugestão da máquina)¹³ e fortalecer a autoridade da ANPD.

¹² A doutrina brasileira tem debatido intensamente o alcance do art. 20 da LGPD, discutindo se ele implica um direito à explicação compreensível das decisões automatizadas e como esse direito se articula com práticas de crédito, consumo e trabalho apoiadas em algoritmos (cf. MATIAS, 2024; SANTOS, 2024).

¹³ O chamado *automation bias* é descrito em estudos de psicologia cognitiva e de sistemas de apoio à decisão como a tendência de pessoas operadoras a confiar excessivamente em recomendações

E) Maturação do debate legislativo no Brasil: Observou-se uma mudança de paradigma radical. O debate iniciou-se com o PL nº 21/2020, que propunha uma regulação minimalista e principiológica, focada excessivamente no fomento à inovação e com poucas travas de segurança. Com a instalação da Comissão de Juristas (CJSUBIA)¹⁴ em 2022 e a realização de audiências públicas multissetoriais, consolidou-se o entendimento de que uma regulação leve seria insuficiente. O resultado foi o PL nº 2.338/2023, que adota uma Abordagem Baseada em Risco (*Risk-Based Approach*). Ao contrário de seu predecessor, o novo texto estabelece categorias claras (risco excessivo, alto risco) e impõe obrigações Ex-Ante de governança, alinhando o Brasil às melhores práticas da OCDE e da União Europeia¹⁵.

4. METODOLOGIA DA PESQUISA

Considerando que o desafio contemporâneo é construir um pensamento sistemático capaz de dar conta dos antagonismos da era digital, a presente pesquisa adotou uma abordagem qualitativa, fundamentada na articulação de diferentes métodos que, juntos, permitiram uma análise profunda e multifacetada do objeto. As técnicas de pesquisa incluíram:

1. Revisão Bibliográfica: Análise de doutrina jurídica, filosofia da tecnologia (Hans Jonas, Mireille Hildebrandt) e relatórios técnicos (Trend Micro, UNESCO).
2. Análise Documental: Exegese comparada entre o PL 21/2020, o Substitutivo do PL 2.338/2023, a LGPD (Lei 13.709/2018) e o *AI Act* europeu.
3. Análise Institucional: Estudo dos documentos da CJSUBIA e das notas técnicas da ANPD para compreender a *mens legis* (intenção do legislador).

produzidas por sistemas automatizados, mesmo diante de sinais de erro, o que torna ainda mais importante a formação crítica e a possibilidade de contestar a saída algorítmica

¹⁴ Segundo o Relatório Final da Comissão de Juristas, a CJSUBIA foi instituída justamente para aproximar o marco brasileiro de IA das melhores práticas internacionais, após amplo ciclo de audiências públicas que evidenciou a necessidade de regras mais robustas de governança e proteção de direitos.

¹⁵ Os Princípios sobre Inteligência Artificial da OCDE, aprovados em 2019, destacam valores como respeito aos direitos humanos, transparência e robustez, influenciando iniciativas regulatórias posteriores, inclusive o *AI Act* europeu (OECD, 2019; EUROPEAN COMMISSION, 2019).

5. RESULTADOS E DISCUSSÃO

A era tecnológica em que vivemos diverge, de muitas maneiras, das que os nossos antepassados presenciaram. Estamos em uma sociedade onde não mais, necessariamente, será o humano que comandará e decidirá sobre sua vida. A automatização das decisões tomadas nos mais diversos âmbitos impactam diretamente no futuro do indivíduo, até mesmo nas mais críticas das situações. A escolha tomada por uma máquina, através da Inteligência Artificial (IA) no que tange, por exemplo, quem ocupará um leito em uma UTI, se o indivíduo pode ou não realizar um empréstimo após uma análise de créditos, se um carro automatizado deve parar ou acelerar diante de uma placa; geram consequências objetivas não somente no porvir do sujeito presente na decisão, mas em uma rede de indivíduos ao seu redor e, quiçá, na sociedade como um todo.

Essas situações precisam, portanto, serem passíveis de uma análise não apenas da máquina, senão do próprio humano, imbuído de sua natureza intrínseca principiológica, determinar o que deve, de fato, ser feito. Para tanto, o Projeto de Lei n 2.338/2023, proposto pelo Senador Rodrigo Pacheco, possui em seu art.9º, art. 10 e art. 11, medidas para que o indivíduo possa contestar decisões automatizadas e solicitar a intervenção humana. Tais artigos não são novidades no contexto da legislação brasileira, haja vista que estão presentes na Lei Geral de Proteção de Dados (LGPD) em seu art. 20, onde é dito que “O titular dos dados tem direito a solicitar a revisão de decisões tomadas unicamente com base em tratamento automatizado de dados pessoais que afetem seus interesses, incluídas as decisões destinadas a definir o seu perfil pessoal, profissional, de consumo e de crédito ou os aspectos de sua personalidade”.

5.1. Problemáticas da IA sem Intervenção Humana Significativa

A pesquisa evidenciou que a ausência de uma supervisão humana qualificada e significativa converte sistemas de Inteligência Artificial em potenciais vetores de injustiça automatizada. Ao operarem desprovidos de capacidade hermenêutica para contextualizar dados fáticos, os modelos algorítmicos tendem a perpetuar e, por vezes, amplificar discriminações pretéritas sedimentadas nas bases

de treinamento, materializando o fenômeno sintetizado pela máxima *bias in, bias out*¹⁶.

Essa dinâmica engendra severas implicações jurídico-sociais, manifestando-se primeiramente na erosão da agência humana. Neste cenário, o sujeito de direitos corre o risco de ser reduzido a um mero objeto de cálculo probabilístico, perdendo a capacidade dialética de influenciar seu próprio destino jurídico frente a uma decisão opaca e determinista. Simultaneamente, observa-se um crítico déficit de responsabilidade (*accountability*), uma vez que a autonomia da máquina tende a diluir o nexo causal em casos de dano; a presença do revisor humano, portanto, torna-se indispensável não apenas como corretivo técnico, mas como o ponto focal necessário para a imputabilidade ética e jurídica no processo decisório.

5.2. O Sistema de Risco da UNESCO e o PL 2.338

A análise do documento de consulta da UNESCO (*Consultation Paper on AI Regulation*) revelou que a estratégia regulatória brasileira reflete uma convergência estratégica com os padrões de governança global, abandonando modelos binários em favor de uma Abordagem Baseada em Risco (*Risk-Based Approach*)¹⁷. O PL nº 2.338/2023 internaliza essa lógica ao estruturar uma pirâmide regulatória que modula o rigor normativo proporcionalmente ao potencial de dano que a tecnologia impõe aos direitos humanos.

Nesse contexto, a legislação veda terminantemente as aplicações classificadas como de risco excessivo (Art. 14) — a exemplo do *social scoring* ou técnicas que manipulem o comportamento subconsciente —, considerando-as incompatíveis com a dignidade humana e o Estado Democrático de Direito. Já para os sistemas de alto risco (Art. 17), que envolvem decisões críticas como identificação biométrica, triagem em saúde ou concessão de crédito, impõe-se um

¹⁶ *Bias in, bias out* sintetiza a ideia de que sistemas de IA reproduzem e, em geral, amplificam os vieses presentes nos dados com que são treinados: se os dados de entrada carregam discriminações históricas ou assimetrias sociais (*bias in*), os resultados produzidos pelo modelo tenderão a refletir essas distorções nas decisões automatizadas (*bias out*), ainda que o algoritmo, em si, seja tecnicamente “neutro”.

¹⁷ O AI Act europeu é frequentemente citado como paradigma dessa abordagem, ao propor uma classificação de sistemas de IA em níveis de risco (proibidos, alto risco, risco limitado e mínimo) e vincular a cada categoria deveres específicos de conformidade, transparência e supervisão humana (EUROPEAN COMMISSION, 2024)

regime de governança robusto que, embora não proíba o uso, exige estritas medidas de mitigação, transparência e supervisão humana.

Além dessa classificação, a pesquisa identificou que o Brasil adota um modelo híbrido, fundindo a gestão de riscos com a Abordagem Baseada em Direitos, materializada na exigência da Avaliação de Impacto Algorítmico (AIA). Este instrumento obriga o desenvolvedor a documentar preventivamente não apenas falhas de segurança, mas também riscos de viés discriminatório, garantindo que o PL 2.338/2023 situe o Brasil no mapa da “IA Confiável” (*Trustworthy AI*) ao equilibrar o fomento à inovação com barreiras éticas intransponíveis.

5.3. Inspirações no AI Act e o Gabinete de IA Europeu

Inicialmente, o debate nacional foi pautado pelo Projeto de Lei nº 21/2020, que propunha uma regulação de caráter principiológico e minimalista (*light-touch regulation*). Aquele texto focava primordialmente no fomento à inovação, estabelecendo diretrizes éticas amplas, mas carecendo de mecanismos de *enforcement* robustos e de uma classificação clara de riscos. A premissa era evitar travas ao desenvolvimento tecnológico, todavia, a complexidade dos danos causados pela IA — desde discriminação algorítmica até a manipulação democrática — tornou evidente a insuficiência de uma abordagem meramente principiológica.

A virada de chave ocorreu com a constituição, em 2022, da Comissão de Juristas (CJSUBIA) no Senado Federal. Em diálogo direto com as tendências da OCDE e, especificamente, com o rascunho do *AI Act* europeu, a Comissão promoveu uma “europeização” do texto legal brasileiro. O resultado, consolidado no PL nº 2.338/2023, abandona o modelo generalista anterior para adotar a *Risk-Based Approach*, espinha dorsal da legislação europeia.

Esta adesão ao modelo europeu manifesta-se estruturalmente na nova lei através de três eixos transplantados do *AI Act*. A Categorização de Riscos, já que assim como na Europa, o PL 2.338/2023 estratifica os sistemas em uma pirâmide regulatória: sistemas de risco excessivo (proibidos por ofenderem direitos fundamentais, como o *social scoring*), sistemas de alto risco (permitidos sob estrita governança) e sistemas de baixo risco; a Governança Ex-Ante, onde há a exigência de Avaliações de Impacto Algorítmico (AIA) antes da colocação do sistema no

mercado, uma burocracia de segurança inspirada na certificação europeia; e o Direitos dos Afetados, que é a garantia de explicação e supervisão humana, buscando alinhar o Brasil aos padrões globais de *Trustworthy AI*.

Como sugestão de aprimoramento (Objetivo D), a pesquisa aponta para a estrutura institucional da Europa. A criação do Gabinete Europeu de IA (*AI Office*) demonstra que a lei, por si só, é insuficiente; é necessária uma infraestrutura de fiscalização técnica. Recomenda-se que o Brasil fortaleça a Autoridade Nacional de Proteção de Dados (ANPD) ou crie um corpo técnico similar, dotado de competência para auditar algoritmos e definir, via regulação setorial, o que constitui uma “supervisão humana significativa”, evitando que o humano seja apenas um “carimbador” de decisões da máquina (*rubber-stamping*).

Vale dizer, ainda, que a evolução do PL 21/2020 para o PL 2.338/2023 não foi apenas um aprimoramento técnico, mas uma decisão política de alinhar o Brasil às democracias ocidentais, garantindo que o país possua um nível de proteção de dados e direitos compatível com o mercado internacional, evitando um possível isolamento tecnológico

5.4. Do Direito à Filosofia: O Projeto de Mestrado

A constatação da dificuldade em operacionalizar métricas para a intervenção humana (Objetivo C) não encerrou a investigação, mas a elevou de patamar. A IC serviu como base empírica e legal para a proposição do Projeto de Mestrado “*Entre a Ética e a Moral: A Supervisão Humana e o Lugar do Direito na Era da Inteligência Artificial*”.

Diante da insuficiência da técnica para explicar o humano, o novo projeto busca na filosofia os fundamentos que tornam a presença humana indispensável. Argumenta-se que, enquanto a IA opera sob uma moralidade funcional e programada, apenas o ser humano detém a prudência e a capacidade de responsabilização ética. O Mestrado se dedicará, portanto, a provar que a supervisão humana exigida pelo PL 2.338/2023 é o mecanismo que impede a substituição do Direito pela matemática, garantindo que a decisão jurídica permaneça um ato de humanidade.

6. CONSIDERAÇÕES FINAIS

A execução desta Iniciação Científica permitiu diagnosticar o estado da arte da regulação da Inteligência Artificial no Brasil, evidenciando a maturação institucional que conduziu o país de uma abordagem principiológica (PL 21/2020) para um modelo robusto de gestão de riscos (PL 2.338/2023). Conclui-se que o novo marco legal é um instrumento moderno e necessário, alinhado aos padrões globais de governança, mas que carrega o risco latente de ineficácia se a “intervenção humana” for reduzida a um formalismo burocrático.

A utilidade social desta pesquisa reside na defesa de que a supervisão humana não é um mero entrave à inovação, mas a última trincheira de defesa da dignidade contra a automação total da vida. Em um país marcado por profundas desigualdades e vieses estruturais, delegar decisões críticas a algoritmos sem uma instância de revisão ética rigorosa equivale a institucionalizar a exclusão social sob o verniz da neutralidade matemática.

Por fim, o estudo encerra-se com um achado epistemológico determinante: a impossibilidade técnica de quantificar a “qualidade” da intervenção humana (Objetivo C) validou a hipótese de que o juízo jurídico é uma competência qualitativa e indelegável. Esta constatação não apenas reafirma que a tecnologia deve servir ao humano — e não o contrário —, mas pavimenta o caminho para o aprofundamento filosófico desta investigação em nível de pós-graduação, onde a *técnica* dará lugar à *ética*.

REFERÊNCIAS

- AUTORIDADE NACIONAL DE PROTEÇÃO DE DADOS (ANPD). **Análise preliminar do Projeto de Lei nº 2338/2023**, que dispõe sobre o uso da Inteligência Artificial. Brasília: ANPD, 2023.
- BRADFORD, Anu. **The Brussels Effect: How the European Union Rules the World**. Oxford: Oxford University Press, 2020.
- BRASIL. Lei nº 13.709, de 14 de agosto de 2018. **Lei Geral de Proteção de Dados Pessoais (LGPD)**. Brasília, DF: Presidência da República, 2018.
- BRASIL. Senado Federal. Comissão de Juristas responsável por subsidiar a elaboração de substitutivo sobre inteligência artificial no Brasil (CJSUBIA). **Relatório Final**. Brasília, 2022.
- BRASIL. Senado Federal. Projeto de Lei n. 21, de 2018. **Dispõe sobre o marco legal da inteligência artificial no Brasil**. Brasília, DF: Senado Federal, 2018.
- BRASIL. Senado Federal. Projeto de Lei nº 2.338, de 2023. **Dispõe sobre o uso da Inteligência Artificial**. Brasília, DF: Senado Federal, 2023.
- BROCHADO, Mariah. Prolegômenos a uma Filosofia Algorítmica Futura Que Possa Apresentar-se Como Fundamento para um Cyberdireito. **Revista de Direito Público**. Brasília, v. 18, n. 100, p. 131-170. out/dez, 2021.
- CUSCIANO, Dalton Tria. A discriminação algorítmica nas contratações laborais digitais. **Revista do Tribunal Superior do Trabalho**, Porto Alegre, v. 90, n. 3, p. 45-60, jul./set. 2024.
- HILDEBRANDT, Mireille; DE BOIS, Arno. **Law for Computer Scientists and Other Folk**. Cham: Springer, 2020.
- JONAS, Hans. **O princípio responsabilidade: ensaio de uma ética para a civilização tecnológica**. Rio de Janeiro: Contraponto, 2006.
- EUROPEAN COMMISSION. **Artificial Intelligence Act (AI Act)**. Brussels, 2024.
- GUTIÉRREZ, Juan David. **Consultation Paper on AI Regulation: Emerging Approaches Across the World**. Paris: UNESCO, 2024.
- EUROPEAN COMMISSION. **Ethics Guidelines for Trustworthy AI**. High-Level Expert Group on Artificial Intelligence. Brussels, 2019.
- EUROPEAN COMMISSION. **Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act)**. Brussels, 2021 (versão consolidada 2024).
- OECD. **Recommendation of the Council on Artificial Intelligence**. Paris: Organisation for Economic Co-operation and Development, 2019.
- MATIAS, J. L. N., & Moraes Júnior, R. A. M. de. (2024). **Discriminação algorítmica na relação de emprego: eficiência econômica, inteligência artificial e fragilidade do empregado**. *Revista Do Tribunal Superior Do Trabalho*, 90(2), 128–147.
- PASQUALE, Frank. **The Black Box Society: The Secret Algorithms That Control Money and Information**. Cambridge: Harvard University Press, 2015.
- SAINZ, Nilton Garcia; GABARDO, Emerson; ONGARATTO, Natália. **Discriminação algorítmica no Brasil: uma análise da pesquisa jurídica e suas perspectivas para a compreensão do fenômeno**. *Revista Direito Público*, Brasília, v. 21, n. 110, p. 258-289, abr./jun. 2024.
- SANTOS JÚNIOR, Ricardo Rodrigues dos. **A explicabilidade como diretriz para as decisões automatizadas e o art. 20 da Lei 13.709/18 (LGPD)**. Migalhas,

São Paulo, 19 jul. 2021. Disponível em: <https://www.migalhas.com.br>. Acesso em: 02 de dezembro de 2025.

TREND MICRO RESEARCH; UNICRI; EUROPOL. ***Malicious Uses and Abuses of Artificial Intelligence***. 2020. Disponível em: <https://www.europol.europa.eu>. Acesso em: 06 de dezembro de 2025.

UNESCO. ***Recommendation on the Ethics of Artificial Intelligence***. Paris: United Nations Educational, Scientific and Cultural Organization, 2021.