ABORDAGEM DIREITO E POLÍTICA PÚBLICA E AVALIAÇÃO REALISTA SOBRE A PROTEÇÃO DE DADOS EM PESQUISAS COM INTELIGÊNCIA ARTIFICIAL

Rodrigo de Mattos Silva Rodrigues

Universidade Federal do Estado do Rio de Janeiro - UNIRIO

Tema 3 - Abordagem DPP nos Campos Disciplinares do Direito

Resumo:

O artigo investiga os mecanismos pelos quais Políticas Públicas de Proteção de Dados (PPPD) geram, em contextos de pesquisa científica com Inteligência Artificial (IA), resultados efetivos de tutela da privacidade. A hipótese central é que a efetividade dessas políticas não deriva da mera vigência da norma, mas da forma como o arranjo jurídico-institucional ativa mecanismos psicossociais nos pesquisadores. Para testá-la, emprega-se uma Revisão Sistemática da Literatura (RSL) de alcance global, com protocolo registrado no Open Science Framework (OSF), operacionalizada pelo software Publish or Perish (PoP) e gerenciada pela plataforma Parsifal. Da análise qualitativa de 18 artigos selecionados, extraíram-se variáveis do modelo Contexto-Mecanismo-Resultado (CMO) para fundamentar a construção de uma Teoria de Médio Alcance (MRT). O artigo propõe um *framework* híbrido original que integra a Abordagem Direito e Políticas Públicas (DPP) à Avaliação Realista (AR), introduzindo a variável Mecanismo Jurídico (M_j) na equação analítica. Os resultados confirmam que a proteção de dados é maximizada quando o arranjo jurídico-institucional, Arr(J+I), é desenhado de modo a converter obrigações de *compliance* em critérios de excelência metodológica, legitimidade acadêmica e competitividade científica. A proposição oferece uma agenda concreta para o redesenho das PPPD brasileiras no âmbito da LGPD e das atividades da ANPD.

Palavras-chave: Políticas Públicas. Inteligência Artificial. Proteção de Dados. Avaliação Realista. Pesquisa Científica.

Abstract:

This article investigates the mechanisms through which Data Protection Public Policies generate effective privacy protection outcomes within the context of scientific research involving Artificial Intelligence (AI). The central hypothesis is that the effectiveness of these

policies does not derive from the mere enforcement of the legal norm, but from how the legal-institutional arrangement triggers psychosocial mechanisms in researchers. To test this hypothesis, a global Systematic Literature Review (SLR) was employed, with a protocol registered on the Open Science Framework (OSF), operationalized via Publish or Perish (PoP) software, and managed through the Parsifal platform. From the qualitative analysis of 18 selected articles, variables from the Context-Mechanism-Outcome (CMO) model were extracted to support the construction of a Middle-Range Theory (MRT). The article proposes the use of an original hybrid framework that integrates the Law and Public Policy Approach with Realistic Review, which introduces the Legal Mechanism variable into the analytical equation. The results confirm that data protection is maximized when the legal-institutional arrangement is designed to convert compliance obligations into criteria for methodological excellence, academic legitimacy, and scientific competitiveness. The proposition offers a concrete agenda for the redesign of Brazilian Data Protection Public Policies within the scope of its General Data Protection Law and the activities of the ANPD (National Data Protection Authority).

***Keywords**: Public Policy. Artificial Intelligence. Data Protection. Realistic Evaluation. Scientific Research.*

***Resumen**:*

El artículo investiga los mecanismos mediante los cuales las Políticas Públicas de Protección de Datos (PPPD) generan, en contextos de investigación científica con Inteligencia Artificial (IA), resultados efectivos de tutela de la privacidad. La hipótesis central es que la efectividad de estas políticas no deriva de la mera vigencia de la norma, sino de la forma en que el arreglo jurídico-institucional activa mecanismos psicosociales en los investigadores. Para probarla, se emplea una Revisión Sistemática de la Literatura (RSL) de alcance global, con protocolo registrado en el Open Science Framework (OSF), operada mediante el software Publish or Perish (PoP) y gestionada por la plataforma Parsifal. Del análisis cualitativo de 18 artículos seleccionados, se extrajeron variables del modelo Contexto-Mecanismo-Resultado (CMO) para fundamentar la construcción de una Teoría de Alcance Medio (MRT). El artículo propone un framework híbrido original que integra el Enfoque de Derecho y Políticas Públicas (DPP) con la Evaluación Realista (AR), introduciendo la variable Mecanismo Jurídico (M_j) en la ecuación analítica. Los resultados confirman que la protección de datos se maximiza cuando el arreglo jurídico-institucional, Arr(J+I), se diseña de modo que convierta las obligaciones

de compliance en criterios de excelencia metodológica, legitimidad académica y competitividad científica. La propuesta ofrece una agenda concreta para el rediseño de las PPPD brasileñas en el ámbito de la LGPD y de las actividades de la ANPD.

***Palabras clave**: Políticas Públicas. Inteligencia Artificial. Protección de Datos. Evaluación Realista. Investigación Científica.*

1. Introdução

Em 2025, o volume global de dados digitais alcançou aproximadamente 181 zettabytes, o equivalente a 38,5 trilhões de DVDs. Esse número não é apenas uma estatística de impacto: é a evidência material de uma ruptura paradigmática na produção científica contemporânea. A ciência orientada por dados (*data-driven science*) converteu a Inteligência Artificial (IA) no motor epistêmico central de pesquisas em Medicina, Sociologia, Direito e Ciências da Computação. Algoritmos que mapeiam prontuários hospitalares, analisam redes sociais ou predizem comportamentos humanos têm em comum uma matéria-prima singular: dados pessoais.

É precisamente nesse ponto que reside a tensão estruturante desta pesquisa. A IA alimenta-se de dados e dados pessoais são, juridicamente, fragmentos da personalidade humana. A cada modelo treinado sobre registros médicos ou históricos de navegação, o imperativo da inovação científica confronta o direito fundamental à privacidade. Esse atrito não é acidental: é sistêmico. A velocidade do desenvolvimento tecnológico supera, em progressão geométrica, a cadência da produção normativa, fenômeno descrito pela literatura como *pacing problem.*

Marcos regulatórios como o General Data Protection Regulation (GDPR) europeu e a Lei Geral de Proteção de Dados (LGPD) brasileira representam tentativas consolidadas de resposta estatal a esse descompasso. No entanto, a questão que permanece em aberto na literatura não é se essas normas existem, mas quando, para quem e sob quais condições elas efetivamente protegem dados. Estudos internacionais documentam um fenômeno recorrente: a coexistência de marcos legais sofisticados com práticas de *compliance* meramente ritualizada, em que a norma vigora, mas a proteção, nem sempre.

É esse hiato entre a norma e a prática que justifica a pergunta central deste artigo: como as Políticas Públicas de Proteção de Dados acionam mecanismos, em contextos de pesquisa científica com IA, para gerar a proteção efetiva da privacidade como resultado? Responder a

essa pergunta exige superar o formalismo jurídico que examina a lei pelo texto e adotar uma perspectiva que examine a lei pela ação dos atores que dela se apropriam - ou a ignoram.

A contribuição original deste artigo é dupla. Em termos teóricos, propõe um *framework* híbrido que integra a Abordagem Direito e Políticas Públicas (DPP) à metodologia da Avaliação Realista (AR), introduzindo uma nova variável analítica - o Mecanismo Jurídico (M_j) - na equação clássica Contexto-Mecanismo-Resultado (CMO). Em termos metodológicos, emprega uma Revisão Sistemática da Literatura de alcance global para testar duas hipóteses: ***(i)*** que Políticas Públicas são necessárias para salvaguardar dados pessoais em pesquisas com IA; e ***(ii)*** que os resultados de tais políticas dependem de contextos e mecanismos específicos.

O artigo organiza-se da seguinte forma: a seção 2 apresenta o referencial teórico; a seção 3 detalha os procedimentos metodológicos; a seção 4 desenvolve a proposição analítica central; e a seção 5 encerra com as conclusões, limitações e agenda de pesquisa.

2. Referencial Teórico

2.1 Abordagem Direito e Políticas Públicas

A análise jurídica tradicional das normas de proteção de dados tende a encerrar-se na exegese do texto legal: o que a norma proíbe, ordena ou permite. Essa leitura, por mais tecnicamente precisa que seja, permanece cega à questão decisiva: o que acontece quando a norma encontra a realidade? A Abordagem DPP, desenvolvida principalmente por Maria Paula Dallari Bucci, propõe precisamente essa transição do direito posto ao direito em ação (Bucci, 2019).

Para a Abordagem DPP, o Direito não é um conjunto estático de comandos, mas um material mobilizável (Caillosse, 2000 *apud* Bucci, 2022) que os atores acionam estrategicamente para conferir legitimidade e balizas operacionais a arranjos institucionais. Nessa perspectiva, regulamentos como o GDPR e a LGPD transcendem a dimensão de normas proibitivas para se consolidarem como verdadeiras tecnologias de governo, capazes de estruturar incentivos, definir papéis institucionais e induzir comportamentos nos diferentes atores do ecossistema científico.

Essa visão dialoga diretamente com o conceito de Direito Reflexivo de Gunther Teubner (1983), segundo o qual, diante da hipercomplexidade social, a regulação mais eficaz não é aquela que tenta prescrever cada detalhe do comportamento humano, mas aquela que estrutura processos de autorregulação e cria procedimentos que induzam os próprios sistemas sociais a

se orientarem de acordo com diretrizes públicas. No ecossistema da pesquisa com IA, isso significa projetar arranjos que levem os pesquisadores a internalizar a proteção de dados não como uma obrigação externa, mas como um componente da qualidade científica.

O conceito operacional central da Abordagem DPP para este artigo é o Arranjo Jurídico-Institucional - Arr(J+I) - definido por Bucci e Coutinho (2017) como o conjunto sistemático de normas, processos, atores e recursos que coordenam a intervenção estatal e a conduta dos agentes públicos e privados em torno de um objetivo comum. Para as Políticas Públicas de Proteção de Dados (PPPDs), esse arranjo abrange desde o texto normativo do GDPR ou da LGPD até os protocolos de Comitês de Ética em Pesquisa, os programas de Sandbox Regulatório da ANPD ou autoridades de outras jurisdições, e as exigências de auditabilidade impostas pelo *EU AI Act*, por exemplo.

2.2 Avaliação Realista e a Configuração CMO

A Avaliação Realista (AR), originalmente proposta por Ray Pawson e Nick Tilley (1997), emerge como alternativa aos modelos de avaliação de políticas que se limitam a verificar se uma intervenção funcionou, sem investigar por que funcionou ou por que falhou. Em vez de perguntar "esta política reduziu violações de dados?", a AR pergunta "o que funcionou, para quem, em que contextos e por meio de quais mecanismos?" (Pawson, 2006).

O instrumental analítico central da AR é a configuração Contexto-Mecanismo-Resultado (CMO). O Contexto (C) corresponde às condições pré-existentes à intervenção, abrangendo fatores socioeconômicos, culturais, organizacionais e institucionais que determinam como os atores recebem e respondem à política. O Mecanismo (M) não é a intervenção em si, mas o processo causal que ela desencadeia, isso é, as reações, interpretações, motivações e comportamentos dos atores que, em interação com o contexto, produzem efeitos. Já o Resultado (O) é o produto dessa interação entre contexto e mecanismos, podendo ser positivo (O+) ou negativo (O-), e caracteriza-se pela não-linearidade e pela variabilidade entre contextos (Wong *et al.*, 2014).

O modelo de causalidade da AR é generativo, não sucessivo (Pawson, 2002). Isso significa que os resultados não decorrem mecanicamente das intervenções, mas emergem da interação contingente entre mecanismos e contextos. Uma mesma PPPD pode produzir proteção efetiva em uma universidade com forte cultura de integridade acadêmica e mero *compliance* de fachada em outra, onde predomina a lógica de publicações a todo custo ("publish

or perish"). Essa percepção é fundamental: a falha não está necessariamente na norma, mas talvez na ausência das condições contextuais que a tornaria eficaz.

Por fim, o objetivo da AR é a construção de Teorias de Médio Alcance, conceito originalmente desenvolvido por Robert Merton (1968) e adaptado por Pawson (2006) ao campo das políticas públicas. Tais teorias situam-se entre as grandes teorias abstratas e as generalizações empíricas pontuais, oferecendo explicações sobre mecanismos em tipos específicos de contexto e, portanto, orientações transferíveis para o redesenho de outras intervenções. São, em síntese, uma tradução do conhecimento acumulado por meio de avaliações realistas que tem como finalidade a aplicação prática para gerar novos aprendizados.

2.3 *Framework* Híbrido: sinergia entre Abordagem DPP e AR

A integração entre a Abordagem DPP e a AR não é trivial: ambas operam em tradições disciplinares distintas. O ponto de inflexão que viabiliza essa síntese é a constatação de que o Arr(J+I) e as variáveis CMO compartilham a mesma premissa fundamental, segundo a qual os resultados de políticas públicas são dependentes de contexto e mediados pelas decisões dos atores envolvidos.

A contribuição teórica original deste artigo consiste na introdução de uma nova variável na equação da AR: o Mecanismo Jurídico (M_j). Na AR clássica, o mecanismo é definido como qualquer processo causal, seja ele psicossocial, organizacional, motivacional, que, ativado pelo contexto, produz resultados. O M_j representa especificamente a forma como o material mobilizável do Direito é interpretado e acionado pelos atores para redirecionar seus comportamentos. Conceitualmente, o M_j sintetiza as reações, escolhas e interpretações dos pesquisadores e instituições diante das normas e demais fontes jurídicas.

Por essa razão, a equação analítica resultante pode ser expressa como:

$$\mathbf{Arr(J+I) \rightarrow (C) + (M) * (M_j) = (O)}$$

Nessa formulação, o Arranjo Jurídico-Institucional funciona como o catalisador que condiciona o contexto (C), estabelecendo as regras do jogo. Dentro dessa estrutura, o M_j interage com os mecanismos psicossociais tradicionais (M), criando uma sinergia multiplicadora que gera os resultados (O) aferidos.

Em razão disso, esse *framework* supera tanto a análise dogmática - que examina a norma sem a prática - quanto a análise puramente comportamental - que examina a prática sem a

norma. A proteção de dados, compreendida por esse prisma, emerge como um fenômeno eminentemente relacional: é o produto da interação entre o que o Direito disponibiliza e o que os atores decidem fazer com isso, em determinados contextos.

3. Procedimentos Metodológicos

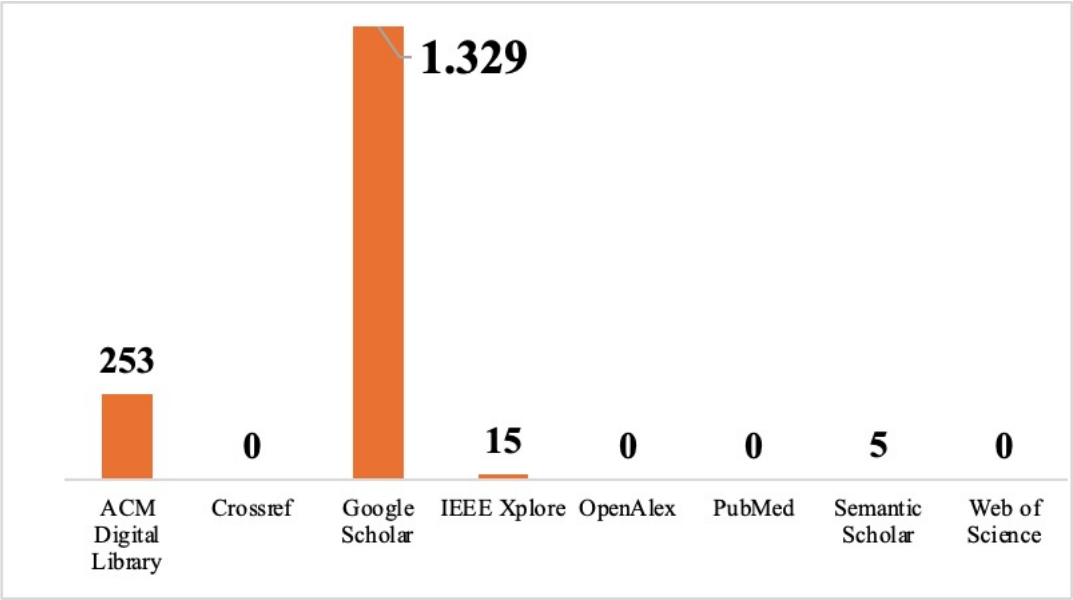
A escolha metodológica desta pesquisa deriva diretamente do problema investigado. Como a AR exige a triangulação de evidências empíricas para extrair entendimentos sobre as configurações CMO e, com isso, construir Teorias de Médio Alcance, optou-se pela Revisão Sistemática da Literatura (RSL) como estratégia de coleta de dados. Para tanto, a RSL seguiu um protocolo pré-definido, registrado publicamente na plataforma Open Science Framework (OSF), garantindo a transparência e a reprodutibilidade do percurso investigativo.

A operacionalização da busca bibliográfica utilizou o software *Publish or Perish* (PoP), que permitiu o acesso simultâneo a múltiplas bases de dados. A estratégia de busca empregou *strings* booleanas estruturadas para capturar a interseção entre três dimensões: inovação tecnológica ("artificial intelligence"), marco regulatório ("General Data Protection Regulation") e ambiente de interesse ("scientific research"). Um segundo bloco de termos estatísticos e econométricos - como "logistic regression", "ARIMA model" e "p-value" - foi adicionado para filtrar documentos com materialidade metodológica empírica, afastando ensaios puramente teóricos ou opinativos. Assim, a string de busca utilizada para o levantamento do conjunto documental que seria analisado foi a seguinte:

> "artificial intelligence" AND "General Data Protection Regulation" AND "scientific research" AND "logistic regression" OR probit OR logit OR "regression analysis" OR "econometric model" OR "negative binomial" OR Poisson OR "ARIMA model" OR "p-value" OR "normal distribution"

As bases de dados diretamente consultadas incluíram Google Scholar, Semantic Scholar, ACM Digital Library e IEEE Xplore. As bases Crossref, PubMed, OpenAlex e Web of Science retornaram resultados nulos ou incompatíveis com a *string*, sendo excluídas da amostra. Portanto, foram obtidos o quantitativo de documentos detalhado na Figura 01 em cada uma das plataforma, sendo o Google Scholar a fonte mais vultosa, com 1.329 artigos.

Figura 1. Quantidade de Documentos por Banco Bibliográfico



Fonte: elaboração própria.

No mais, a gestão do processo de triagem foi centralizada na plataforma Parsifal (*https://parsif.al*), especializada no suporte a revisões sistemáticas, que permitiu a rastreabilidade de cada decisão de inclusão e exclusão de trabalhos (Tractenberg; Struchiner, 2011).

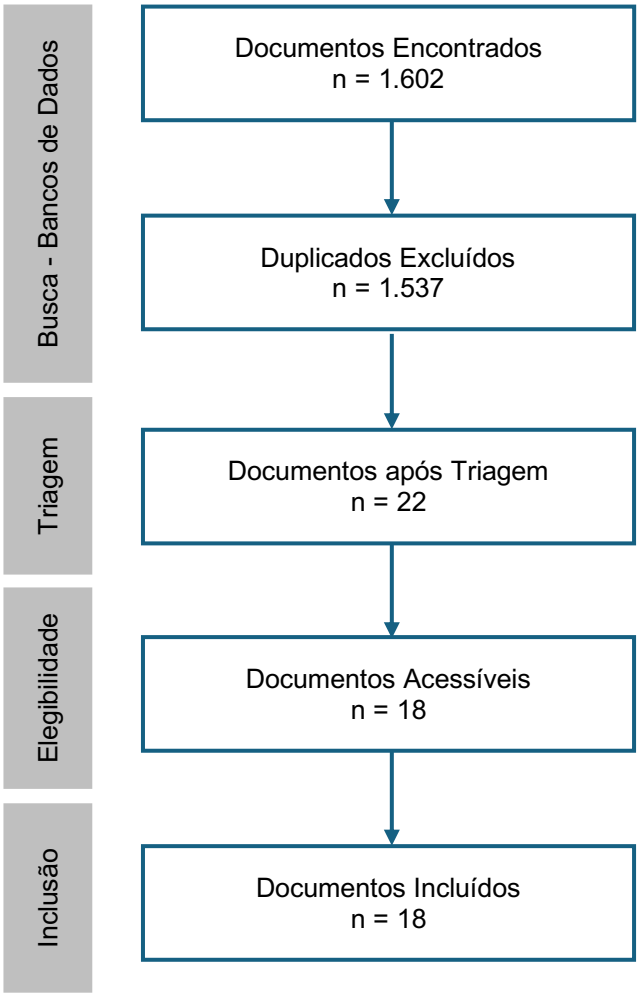
Os critérios de inclusão restringiram-se a documentos que tratassem de pesquisa científica realizada com IA e que abordassem direta ou indiretamente temas de proteção de dados pessoais ou privacidade. Por sua vez, os critérios de exclusão abrangeram quatro eixos:

(a) estudos focados em contextos puramente mercadológicos, dissociados da produção acadêmica;
(b) trabalhos puramente teóricos ou conceituais, sem evidências empíricas;
(c) documentos focados exclusivamente em aspectos técnicos de engenharia da IA, sem menção à proteção de dados; e
(d) estudos sem uso efetivo de IA ou pertencentes à literatura cinza.

A aplicação sequencial desses critérios configurou um funil de seleção que partiu de um conjunto documental bruto de aproximadamente 1.602 registros. Após a remoção de 65 duplicatas, a triagem por título e resumo e a verificação de disponibilidade integral dos textos -

sendo quatro artigos excluídos por indisponibilidade de acesso - a amostra consolidou-se em 18 artigos, os quais constituem a base empírica desta análise.

Figura 2. Fluxograma da RSL



Fonte: elaboração própria.

Por conseguinte, o framework híbrido de Abordagem DPP e AR foi aplicado sobre os 18 artigos listados na Tabela 1.

Tabela 1. Literatura selecionada na RSL.

Título do Documento	**Autor(es), ano**
Interpretable machine learning as a tool for scientific discovery in chemistry	Dybowski, 2020

Machine learning in information systems - a bibliographic review	Abdel-Karim; Pfeuffer; Hinz, 2021
Disentangling the Components of Ethical Research in Machine Learning	Ashurst *et al.*, 2022
The Automated Laplacean Demon: How ML Challenges Our Views	Srećković; Berber; Filipović, 2021
Distributed learning for optimal radiomics knowledge	Zerka, 2022
A Theory-Guided Approach to Explaining and Modeling Deep Neural Networks	Mastropietro, 2022
Responsible and Regulatory Conform Machine Learning for Medicine	Petersen et al., 2022
Knowledge Recommendation Systems with Machine Intelligence Algorithms	Protasiewicz, 2023
AI's next frontier: The rise of ChatGPT and its implications	Ramos; Márquez; Rivas-Echeverría, 2023
Prompt engineering: The next big skill in rheumatology research	Venerito *et al.*, 2024
AI-assisted prescreening of biomedical research proposals	Carbonell Cortés *et al.*, 2024
Use of AI in Social Sciences Research	Shahid; Fatima, 2024
ChatGPT for Bibliometrics: Potential applications and limitations	Torres-Salinas; Thelwall; Arroyo Machado, 2024
The Role of Artificial Intelligence for Societal Challenges	Abbonato, 2024
Leveraging AI in Anesthesiology Research	Doyal *et al.*, 2025
Methodological Foundations for AI-Driven Survey Generation	Mburu *et al.*, 2025

A sistematização da produção acadêmica reunida na RSL permitiu identificar um campo em transformação acelerada, no qual o tratamento conferido à IA migrou progressivamente de uma abordagem centrada no desempenho técnico para uma perspectiva orientada à conformidade ética e normativa. Essa transição sinaliza uma reconfiguração das prioridades epistemológicas da comunidade científica em relação ao desenvolvimento e à aplicação de modelos de aprendizado de máquina.

Do ponto de vista terminológico, as palavras-chave presentes nos documentos analisados funcionam como indicadores sensíveis dessa mudança de orientação. Nos estudos mais antigos, predominam termos de natureza técnica, como *machine learning*, *deep learning* e *artificial intelligence*, cujo emprego revela uma preocupação central com a eficiência e a arquitetura dos modelos. À medida que a produção avança, esse vocabulário divide protagonismo com expressões de ordem normativa e social, como *AI ethics*, *explainability*, *research integrity* e *GDPR*. Esse deslocamento léxico não é acidental, pois traduz a percepção crescente de que a validade de um sistema de inteligência artificial não se esgota em sua capacidade preditiva, mas exige também que seus processos internos sejam inteligíveis, auditáveis e conformes a padrões éticos reconhecidos.

A dimensão temporal da amostra, que abrange o intervalo entre 2020 e 2025, oferece uma leitura quase cronológica dos estágios de maturação desse campo. Os estudos inaugurais possuíam caráter predominantemente exploratório, dedicando-se à descrição das capacidades e limitações dos algoritmos. A partir de 2022, observa-se que os trabalhos passam a confrontar o desenvolvimento tecnológico com critérios de integridade científica, propondo *frameworks* para a mitigação de danos e para a responsabilização dos agentes envolvidos na pesquisa. O biênio seguinte, marcado pela irrupção dos modelos de linguagem no debate público, intensificou esse movimento, com estudos dedicados a examinar as implicações da IA generativa em domínios tão distintos quanto a bibliometria, a gestão da saúde e o ensino superior. Em 2025, os estudos apontaram para uma fase de institucionalização, na qual a tecnologia é incorporada estruturalmente às práticas profissionais e acadêmicas, e os requisitos de conformidade regulatória deixam de ser um desejo de futuro para se tornarem condição de legitimidade do uso da IA.

A distribuição geográfica dos estudos revela, por sua vez, uma divisão de trabalho intelectual que não é indiferente ao contexto econômico e regulatório de cada região. A

produção europeia, especialmente a proveniente de países como Alemanha, Itália, Espanha e Reino Unido, tende a tomar o Regulamento Geral de Proteção de Dados da União Europeia como referência normativa central, articulando o desenvolvimento técnico à exigência de transparência algorítmica como imperativo legal, o que condiciona o acesso ao mercado comunitário à adoção de requisitos elevados de proteção. Em contrapartida, estudos oriundos de regiões como a América Latina e a África subsaariana tendem a enfatizar as assimetrias de acesso às ferramentas digitais e os impactos sociais da adoção da IA em contextos de recursos limitados, deslocando o eixo da discussão da conformidade normativa para a equidade e a democratização tecnológica.

No que concerne às áreas do conhecimento representadas, a amostra ilustra com clareza o caráter transversal que a IA assumiu na ciência contemporânea. Da Química à Medicina, da Filosofia à Educação, passando por Sistemas de Informação e Bibliometria, todas as disciplinas presentes na RSL compartilham uma preocupação estrutural com o processamento de grandes volumes de dados e com as implicações éticas desse processamento para os sujeitos envolvidos.

Em síntese, o panorama delineado pela RSL aponta para um campo em rápida consolidação, no qual a proteção de dados pessoais se afirma como eixo estruturante do debate sobre o uso da IA na pesquisa científica, o que será aprofundado

Em síntese, o panorama delineado pela RSL aponta para um campo em rápida consolidação, no qual a proteção de dados pessoais se afirma como eixo estruturante do debate sobre o uso da IA na pesquisa científica. Essa base empírica, consolidada nos 18 artigos selecionados, fornece os insumos necessários para a próxima etapa da pesquisa: a extração qualitativa e interpretativa das variáveis de Arranjo Jurídico-Institucional, Contexto, Mecanismos e Resultado, mediante a aplicação do *framework* híbrido proposto, com vistas a fundamentar a proposição de uma Teoria de Médio Alcance que explique as condições de efetividade das políticas públicas de proteção de dados em pesquisas científicas realizadas com IA.

4. Resultados à luz do *framework* híbrido

4.1 Arranjos Jurídico-Institucionais identificados

A análise da literatura selecionada revelou um panorama regulatório global marcado pela densidade normativa e pela pluralidade de fontes. O eixo central dos arranjos identificados é o GDPR europeu, cuja influência extraterritorial - o chamado Efeito Bruxelas (Bradford,

2012) - reconfigura práticas de pesquisa em jurisdições tão distintas quanto o Equador, o Brasil e a China. Pesquisadores que visam à publicação em periódicos internacionais de alto impacto operam, na prática, sob padrões regulatórios que excedem as exigências de seus próprios países, pois a aceitabilidade editorial funciona como vetor de disseminação normativa (Zerka, 2022).

No domínio biomédico, o arranjo torna-se mais denso: ao GDPR somam-se o HIPAA norte-americano, o Medical Device Regulation (MDR) europeu e as normas da National Medical Products Administration (NAMD) chinesa, além de padrões técnicos como as ISO 13.485 e IEC 62.304 (Petersen *et al.,* 2022).

Já o EU AI Act, publicado em 2024, adiciona uma nova camada ao classificar a maioria das aplicações biomédicas como de alto risco, impondo obrigações de documentação técnica e auditabilidade que, ao exigir transparência sobre vieses algorítmicos, transformam o Direito: de freio em base para confiança epistêmica.

No contexto brasileiro, o arranjo está em maturação. A LGPD estabelece o marco normativo geral; o Guia Orientativo da ANPD para fins acadêmicos (2023) define bases legais específicas para o tratamento de dados em pesquisa; e o programa de Sandbox Regulatório, inaugurado pelo Edital nº 01/2024, representa um instrumento de experimentação controlada que permite testar arranjos de IA sem cristalizar prematuramente normas que podem se tornar obsoletas diante do ritmo da inovação.

4.2 Contextos

Os contextos identificados na literatura distribuem-se em dois eixos principais. O primeiro é o macrocontexto da ciência orientada por dados, caracterizado pela proliferação de técnicas de *Machine Learning* sem padronização ética (Abdel-Karim; Pfeuffer; Hinz, 2021) e pela fragmentação de repositórios de dados em silos institucionais, particularmente nas ciências da saúde. Esse cenário cria um ambiente de incerteza regulatória no qual as normas de proteção de dados chegam como uma novidade disruptiva, e não como um padrão cultural estabelecido.

O segundo eixo refere-se aos microcontextos institucionais, que determinam como os pesquisadores recebem as exigências do arranjo jurídico-institucional. A literatura revela dois polos. No primeiro, a cultura institucional valoriza a credibilidade e a reprodutibilidade como critérios de prestígio: nesse contexto, normas de proteção de dados são percebidas como garantias de validade científica (C+). No segundo, predomina a lógica do *publish or perish*, em que a pressão quantitativista por produtividade trata qualquer burocracia adicional como

obstáculo à carreira (C-). A mesma norma, nesses dois contextos, pode gerar resultados radicalmente distintos (Ashurst *et al.*, 2022).

4.3 Mecanismos

A análise das configurações CMO identificou dois padrões de mecanismos dominantes, que correspondem aos contextos descritos acima.

No padrão positivo (O+), o M_j de supervisão humana ou de compliance proativo é percebido como um recurso de legitimidade, não como um constrangimento. Carbonell Cortés *et al.* (2024) documentam esse processo no contexto de triagem algorítmica de projetos de pesquisa: ao incorporar o M_j da supervisão humana ao *design* do sistema de IA, os gestores ativaram um mecanismo de busca por objetividade e equidade (M+), convertendo uma exigência legal em um critério de qualidade institucional. Analogamente, Zerka (2022) demonstra que a exigência de proteção de dados sensíveis em pesquisas multicêntricas tornou-se o motor da inovação técnica em prol do aprendizado federado (*federated learning*), que permite análises colaborativas sem centralização de dados. Aqui, o M_j não apenas protegeu dados, mas redefiniu o estado da arte metodológico.

No padrão negativo (O-), o M_j é percebido como uma externalidade punitiva, desconectada dos valores da prática científica. Ashurst *et al.* (2022) descrevem ambientes em que pesquisadores encaram os protocolos de ética em IA como rituais burocráticos a serem cumpridos formalmente, sem internalização dos princípios subjacentes. O mecanismo ativado é o medo sancionatório (M-), que produz um *compliance* de fachada, pois os formulários são preenchidos, os documentos são assinados, mas as práticas efetivas de tratamento de dados permanecem inalteradas. O Direito, nesse caso, produz legitimidade aparente sem proteção real.

A variável que distingue os dois padrões não é a qualidade da norma, mas a natureza do M_j ativado, o que depende criticamente do contexto institucional no qual a norma aterrissa. Isso confirma a segunda hipótese desta pesquisa: os resultados das PPPDs são dependentes da configuração CMO específica em que operam.

4.4 Discussões e proposição de Teoria de Médio Alcance

A presente investigação buscou responder de que maneira as PPPDs acionam mecanismos específicos, em contextos de pesquisa com IA, para gerar a proteção da privacidade como um resultado efetivo. Pautada no framework híbrido proposto, a análise

demonstrou que a eficácia dessas políticas não decorre da simples vigência da norma, mas da capacidade do arranjo jurídico-institucional em ativar mecanismos que alinham a regulação aos valores epistêmicos da prática científica. Sob essa ótica, percebeu-se que a proteção é maximizada quando o arcabouço legal deixa de ser percebido como uma externalidade burocrática e passa a ser integrado como um recurso de excelência metodológica, convertendo exigências de transparência em motores para a adoção, por exemplo, de técnicas de Inteligência Artificial Explicável (XAI), as quais garantiriam simultaneamente a tutela do titular dos dados e a validação científica do próprio pesquisador.

Em contrapartida, o estudo revela que mecanismos puramente dissuasórios, baseados no medo de sanção e na aversão à burocracia, tendem a falhar em ambientes de alta pressão produtiva, resultando em um *compliance* de fachada que esvazia a finalidade protetiva da norma. Logo, a motivação central dos atores envolvidos na produção científica mostra-se vinculada a uma economia de prestígio, na qual a adesão a padrões rigorosos de proteção de dados funciona como uma estratégia de sinalização de qualidade e busca por legitimidade internacional, permitindo que pesquisadores transformem seus deveres regulatórias em diferenciais competitivos e colaborações estratégicas.

Com isso, a síntese das configurações identificadas permite propor a seguinte Teoria de Médio Alcance:

> A efetividade das PPPDs em pesquisas científicas com IA é maximizada quando os Arranjos Jurídico-Institucionais são desenhados e implementados de modo a integrar os mecanismos aos valores epistêmicos e aos incentivos constitutivos da prática científica, convertendo obrigações de *compliance* em critérios institucionais ou pessoais de excelência metodológica, legitimidade e competitividade acadêmica.

Essa teoria possui escopo delimitado, mas transferível. Ela descreve os mecanismos de efetividade em contextos de pesquisa acadêmica com IA nos campos biomédico, das ciências sociais e da computação, que são os domínios predominantes na literatura analisada, mas permite a verificação empírica de sua aplicabilidade em outros setores.

Ainda, a teoria opera em dois sentidos. No sentido diagnóstico, explica por que arranjos regulatórios similares produzem resultados distintos: em contextos institucionais diferentes, os arranjos são percebidos de forma diferente, despertando ou deixando de despertar mecanismos que promoveriam resultados. No sentido propositivo, indica que o redesenho de PPPDs mais efetivas deve atuar sobre o contexto, criando as condições que despertem mecanismos para que o arranjo seja percebido como um recursos de qualidade, e não apenas como mera burocracia.

4.5 Implicações para a PPPD brasileira

A Teoria de Médio Alcance proposta oferece uma agenda concreta para a ANPD e para o Ministério da Ciência, Tecnologia e Inovação - MCTI, pois corrobora que esses órgãos precisam transitar do modelo de fiscal externo, focado em sanções, para o de arquiteto institucional, propondo o desenho de arranjos que promovam contextos em que a proteção de dados seja vista como o caminho para obter recursos e reconhecimento científico.

Operacionalmente, entende-se que isso implica: ***(i)*** condicionar editais de fomento à apresentação de Relatórios de Impacto à Proteção de Dados (RIPD) e ao uso de modelos de IA Explicáveis (XAI), de modo que a proteção de dados se torne critério de acesso ao financiamento e não apenas exigência de conformidade; ***(ii)*** converter os Comitês de Ética em Pesquisa (CEP/CONEP) e os Encarregados de Dados em centros de assessoria técnica que auxiliem os pesquisadores na implementação prática do *privacy by design*, reduzindo o custo de aprendizado associado ao *compliance*; e ***(iii)*** expandir o uso de Sandboxes Regulatórios da ANPD para testar de forma experimental como o arranjo da LGPD se adapta a tecnologias específicas de IA, antes da promulgação de normas que poderiam rapidamente se tornar obsoletas.

Essas medidas alteram o sinal do M_j: de dissuasório ou punitivo para estratégicom uma vez que, quando o pesquisador percebe que proteger dados é o que garante financiamento, aceitação editorial e credibilidade internacional, o mecanismo ativado não é o medo, mas a aspiração à excelência. É nesse ponto que o Direito, compreendido como material mobilizável, cumpre sua função mais sofisticada: não coibir, mas orientar.

5. Conclusão

Esta pesquisa partiu de um incomodo que a dogmática jurídica tradicional não consegue resolver: por que normas de proteção de dados tão sofisticadas quanto o GDPR ou a LGPD convivem, em muitos contextos, com práticas de *compliance* que protegem mais o papel do que os dados? A resposta, elaborada ao longo deste artigo, reside na natureza contingente da efetividade jurídica.

As duas hipóteses iniciais foram confirmadas. A primeira - de que Políticas Públicas são necessárias para a proteção de dados em pesquisas com IA - foi corroborada pela evidência de que, na ausência de arranjos jurídico-institucionais estruturados, a pressão competitiva da

produção científica tende a privilegiar a eficiência e a produtividade em detrimento da integridade no tratamento dos dados pessoais. A boa vontade ética individual é insuficiente diante da complexidade técnica dos sistemas de IA e das pressões institucionais do ambiente acadêmico.

A segunda hipótese - de que os resultados das PPPD dependem de contextos e mecanismos específicos - não foi apenas confirmada, mas operacionalizada. A Teoria de Médio Alcance proposta assevera que o sucesso da proteção de dados ocorre quando o arranjo facilita a tradução de comandos normativos em soluções técnicas e incentivos epistêmicos, permitindo que o pesquisador veja a proteção de dados como parte constitutiva de seu processo científico, e não como um obstáculo à sua carreira.

Ainda, a análise de Mecanismos Jurídicos como variável analítica na equação da Avaliação Realista permitiu incorporar a dimensão jurídico-normativa sem reduzi-la a um fator de contexto. Afinal, a norma não apenas configura o ambiente, mas também provoca reações específicas nos atores, que são o cerne da mudança comportamental que qualquer política pública pretende induzir.

Outrossim, as limitações do estudo derivam principalmente do *pacing problem* que o próprio artigo discute, porque a velocidade da evolução tecnológica da IA ameaça tornar evidências rapidamente desatualizadas. A RSL capturou um momento, no mês de recorte do levantamento documental, em novembro de 2025, enquanto os sistemas de IA e a produção acadêmica sobre o tema evoluem continuamente, gerando novos contextos que demandam a avaliação de novas configurações CMO. Adicionalmente, a opacidade inerente aos sistemas de IA (black-box) impõe limites ao escrutínio empírico dos mecanismos subjacentes, tornando as inferências interpretativas inevitavelmente sujeitas a revisão.

Em síntese, este artigo defende que a proteção de dados pessoais não é um custo da pesquisa científica com IA, mas, sim, seu critério de validade epistêmica e social em uma era de algoritmos. A efetividade dessa proteção não reside na perfeição estática da letra da lei, mas na inteligência do arranjo em modular incentivos e despertar a prontidão ética dos pesquisadores.

REFERÊNCIAS

ABDEL-KARIM, Bassel Musie; PFEUFFER, Nicolas; HINZ, Oliver. Machine learning in information systems – a bibliographic review. *Electronic Markets*, v. 31, p. 959-985, 2021.

ASHURST, Carolyn et al. Disentangling the Components of Ethical Research in Machine Learning. In: *ACM Conference on Fairness, Accountability, and Transparency* (FAccT 2022). p. 124-136.

ASHURST, Carolyn; BAROCAS, Solon; CAMPBELL, Rosie; RAJI, Inioluwa Deborah. Disentangling the Components of Ethical Research in Machine Learning. *In*: ACM CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY (FAccT), 2022, Seul. *Proceedings [...]*. Nova York: ACM, 2022. p. 1-12. DOI: 10.1145/3531146.3533781.

BRADFORD, Anu. The Brussels Effect. *Northwestern University Law Review.* v. 107, n. 1, p. 1-67, 2012.

BRASIL. Autoridade Nacional de Proteção de Dados (ANPD). *Guia orientativo*: tratamento de dados pessoais para fins acadêmicos e para a realização de estudos e pesquisas. Brasília, DF: ANPD, 2023.

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.

BUCCI, Maria Paula Dallari; COUTINHO, Diogo Rosenthal. Arranjos jurídico-institucionais da política de inovação tecnológica. In: *Inovação no Brasil*: avanços e desafios jurídicos e institucionais. São Paulo: Blucher, 2017. p. 339.

BUCCI, Maria Paula Dallari. A abordagem direito e políticas públicas: aspectos metodológicos. *Revista de Investigações Constitucionais*, v. 6, n. 3, p. 791-832, 2019.

CARBONELL CORTÉS, Carla; FERNÁNDEZ-MÁRQUEZ, Jose Luis; FONT, Àngel; LÓPEZ, Ignasi; MAURI, Ester; MONTSERRAT, Maite de; OLAZABAL, Paula de; ORPÍ, Maria de la Pau; ROMÁN, Begoña. AI-assisted prescreening of biomedical research proposals: ethical considerations and the pilot case of "la Caixa" Foundation. *Data & Policy*, Cambridge, v. 6, e34, p. 1-21, 23 out. 2024.

DOYAL, Alexander S.; VANDER PLOEG, Matthew R.; WNEK, Russel; CHEN, Fei. *Leveraging Artificial Intelligence in Anesthesiology Research and Article Writing:* Opportunities, Challenges, and Best Practices. A&A Practice, [S. l.], 2025.

DYBOWSKI, Richard. Interpretable machine learning as a tool for scientific discovery in chemistry. **New Journal of Chemistry**, v. 44, p. 20914-20914, 2020. DOI: 10.1039/d0nj02592e.

MASTROPIETRO, Antonio. *A Theory-Guided Approach to Explaining and Modeling Deep Neural Networks*. 2022. Tese (Doutorado em Matemática Pura e Aplicada) - Politecnico di Torino; Università degli Studi di Torino, Turim, 2022.

MBURU, Ted K.; RONG, Kangxuan; MCCOLLEY, Campbell J.; WERTH, Alexandra. Methodological foundations for artificial intelligence-driven survey question generation. *Journal of Engineering Education*, [S. l.], 2025.

MBURU, Ted K.; RONG, Kangxuan; MCCOLLEY, Campbell J.; WERTH, Alexandra. *Methodological Foundations for AI-Driven Survey Question Generation.* [S. l.: s. n.], 2025.

MERTON, Robert K. *Social theory and social structure*. Nova Iorque: Free Press, 1968.

PAWSON, Ray; TILLEY, Nick. *Realistic evaluation*. Londres: Sage, 1997.

PAWSON, Ray. *Evidence based policy*: The promise of realist synthesis. Londres: Centre for Evidence Based Policy, 2002.

PAWSON, Ray. *Evidence-based policy*: a realist perspective. Londres: SAGE, 2006.

PETERSEN, Eike et al. Responsible and Regulatory Conform Machine Learning for Medicine. *Nature Machine Intelligence*, v. 4, p. 116-124, 2022.

PETERSEN, Eike et al. Responsible and Regulatory Conform Machine Learning for Medicine: A Survey of Challenges and Solutions. *IEEE Access*, v. 10, p. 58315-58382, 2022. DOI: 10.1109/ACCESS.2022.3178382.

PROTASIEWICZ, Jarosław. *Knowledge Recommendation Systems with Machine Intelligence Algorithms*: People and Innovations. Cham: Springer, 2023.

REĆKOVIĆ, S. et al. *The automated Laplacean demon: How ML challenges our views on prediction and explanation*. [S. l.]: Springer Nature, 2021.

SHAHID, Aimen; FATIMA, Nosheen. Use of AI in Social Sciences Research: Ethical and Methodological Perspectives. *Social Science Computer Review*, v. 42, n. 2, p. 215-234, 2024.

TEUBNER, Gunther. Substantive and reflexive elements in modern law. *Law and Society Review*, v. 17, n. 2, p. 239-285, 1983.

TORRES-SALINAS, Daniel; THELWALL, Mike; ARROYO-MACHADO, Wenceslao. ChatGPT for Bibliometrics: Potential applications and limitations. *In*: GOODING, Paul; TERRAS, Melissa; AMES, Sarah (ed.). *Library Catalogues as Data*: Research, Practice and Usage. London: Facet Publishing, 2025.

TRACTENBERG, Leonel; STRUCHINER, Miriam. Revisão realista: uma abordagem de síntese de pesquisas para fundamentar a teorização e a prática baseada em evidências. *Ciência da Informação*, v. 40, n. 2, p. 102-114, 2011.

VENERITO, Vincenzo; LALWANI, Devansh; DEL VESCOVO, Sergio; IANNONE, Florenzo; GUPTA, Latika. *Prompt engineering*: the next big skill in rheumatology research. International Journal of Rheumatic Diseases, [S. l.], 2024.

VENERITO, Vincenzo; LALWANI, Devansh; DEL VESCOVO, Sergio; IANNONE, Florenzo; GUPTA, Latika. Prompt engineering: The next big skill in rheumatology research. *International Journal of Rheumatic Diseases*, v. 27, e15157, 2024. DOI: 10.1111/1756-185X.15157.

WONG, Geof; Greenhalgh, Trish; Westhorp, Gill; Pawson, Ray. Development of methodological guidance, publication standards and training materials for realist and meta-narrative reviews: the RAMESES project. *Health Services and Delivery Research*, v. 2, n. 30, 2014.

ZERKA, Fadila; BARAKAT, Samir; WALSH, Sean; BOGOWICZ, Marta; LEIJENAAR, Ralph T. H.; JOCHEMS, Arthur; MIRAGLIO, Benjamin; TOWNEND, David; LAMBIN, Philippe. Privacy-preserving distributed machine learning from federated databases in healthcare. *In: Distributed Machine Learning for Healthcare.* [S. l.: s. n.], 2022.