Reprodutibilidade Científica na Era da Inteligência Artificial: Desafios Jurídicos e Inovação em Propriedade Intelectual

Scientific Reproducibility in the Age of Artificial Intelligence: Legal Challenges and Innovation in Intellectual Property

Giovanna Sampaio [1] | **João dos Santos** [2] | **Bruno Assis** [3]

[1–2] Universidade Federal de Sergipe, Brasil
[3] Universidade Estadual de Santa Cruz, Brasil

Corresponding author: Giovanna Sampaio | giovanna.martins@ufba.br
DOI: https://doi.org/10.29073/e3.v11i2.1046
Received: 24/01/2026 | Accepted: 10/12/2025 | Published: 30/12/2025

RESUMO

A reprodutibilidade científica tem se tornado um tema central diante da crise que assola diversas áreas do conhecimento, especialmente na era da inteligência artificial (IA). Este trabalho analisa os desafios técnicos, jurídicos e éticos associados à reprodutibilidade, com foco na interação entre IA, propriedade intelectual, direitos autorais, patentes e a Lei Geral de Proteção de Dados (LGPD). Aborda-se o impacto da IA na replicação de estudos, as barreiras legais para o compartilhamento de dados e as estratégias para conciliar transparência científica com a proteção da privacidade. A análise revela a necessidade de políticas públicas integradas, infraestrutura tecnológica adequada e uma cultura científica que valorize a reprodutibilidade como princípio ético fundamental para a inovação responsável. Conclui-se que o avanço da ciência reprodutível depende da cooperação multidisciplinar, do equilíbrio entre proteção jurídica e acesso aberto, e do uso ético das novas tecnologias.

Palavras-chave: Reprodutibilidade científica, inteligência artificial, propriedade intelectual, LGPD; direitos autorais, patentes, ciência aberta.

ABSTRACT

Scientific reproducibility has become a central theme amidst the crisis affecting various fields of knowledge, especially in the era of artificial intelligence (AI). This paper analyzes the technical, legal, and ethical challenges related to reproducibility, focusing on the interaction between AI, intellectual property, copyrights, patents, and the Brazilian General Data Protection Law (LGPD). The impact of AI on study replication, legal barriers to data sharing, and strategies to reconcile scientific transparency with privacy protection are addressed. The analysis highlights the need for integrated public policies, adequate technological infrastructure, and a scientific culture that values reproducibility as a fundamental ethical principle for responsible innovation. It concludes that the advancement of reproducible science depends on multidisciplinary cooperation, balancing legal protection with open access, and ethical use of new technologies.

Keywords: Scientific reproducibility; artificial intelligence; intellectual property; LGPD; copyrights; patents; open science.

1. INTRODUÇÃO

A reprodutibilidade científica é um dos pilares fundamentais da metodologia científica, permitindo que resultados obtidos por um pesquisador possam ser verificados e replicados por outros, garantindo a confiabilidade e validade do conhecimento produzido. Contudo, nas últimas décadas,

diversas áreas do conhecimento têm enfrentado o que ficou conhecido como a "crise de reprodutibilidade", que se caracteriza pela dificuldade ou impossibilidade de replicar estudos previamente publicados (Baker, 2016). Esta crise tem levantado preocupações quanto à integridade da pesquisa científica e ao impacto disso na inovação, políticas públicas e na confiança da sociedade na ciência.

A emergência da inteligência artificial (IA) no contexto científico adiciona camadas adicionais de complexidade para a reprodutibilidade. Por um lado, a IA oferece ferramentas poderosas para análise de grandes volumes de dados e automação de processos experimentais, potencialmente aumentando a eficiência e a precisão da replicação. Por outro lado, o uso de algoritmos complexos, muitas vezes opacos ("caixa preta"), e a dependência de grandes bases de dados que podem estar protegidas por questões jurídicas e éticas dificultam a plena transparência necessária para a reprodutibilidade (Hutson, 2018).

Além dos aspectos técnicos, a questão da propriedade intelectual (PI) torna-se central quando se discute a reprodutibilidade, especialmente em um cenário onde os dados, softwares e algoritmos utilizados nas pesquisas estão protegidos por direitos autorais, patentes ou licenças específicas. A proteção desses ativos pode limitar o acesso irrestrito ao material necessário para replicação, criando um dilema entre o incentivo à inovação e a promoção da ciência aberta (Contreras, 2017).

Outro aspecto relevante está relacionado à proteção de dados pessoais, especialmente em países com legislação rigorosa, como o Brasil, que adotou a Lei Geral de Proteção de Dados Pessoais (LGPD) em 2018. A LGPD impõe restrições ao tratamento e compartilhamento de dados pessoais, o que pode impactar diretamente a reprodutibilidade de estudos científicos que dependem de informações sensíveis, especialmente nas áreas da saúde e ciências sociais (Teles, 2021). Equilibrar a necessidade de transparência com a proteção da privacidade é um desafio jurídico e ético crescente.

Diante desse cenário multifacetado, o presente trabalho busca analisar a reprodutibilidade científica na era da inteligência artificial, considerando os desafios jurídicos relacionados à propriedade intelectual, direitos autorais, patentes e proteção de dados pessoais. O objetivo é compreender as barreiras e oportunidades para a promoção de práticas reprodutíveis que atendam às exigências técnicas, éticas e legais contemporâneas.

CONTEXTUALIZAÇÃO DA CRISE DE REPRODUTIBILIDADE

A crise de reprodutibilidade não é um fenômeno novo, mas ganhou grande destaque a partir da última década, especialmente nas ciências biomédicas e sociais. Estudos como o da Open Science Collaboration (2015) revelaram que apenas cerca de 40% dos estudos psicológicos clássicos puderam ser replicados com sucesso, trazendo à tona uma série de críticas ao modo tradicional de conduzir, publicar e avaliar a ciência. Baker (2016) destaca que a crise é multifatorial, envolvendo desde problemas metodológicos, má conduta, vieses de publicação até limitações estruturais no sistema acadêmico.

No contexto das ciências exatas e da computação, a reprodutibilidade enfrenta desafios particulares, relacionados ao uso de softwares, códigos e plataformas computacionais que, muitas vezes, não são disponibilizados abertamente. Isso contrasta com áreas como a física, onde a experimentação física pode ser reproduzida de forma mais objetiva. A opacidade dos algoritmos

de IA, frequentemente proprietários, é um entrave para a validação independente dos resultados (Hutson, 2018).

Adicionalmente, a crescente complexidade dos experimentos científicos e a dependência de grandes bases de dados dificultam a replicação direta, exigindo a adoção de boas práticas como o registro prévio de protocolos, uso de repositórios públicos e ferramentas de versionamento de código (Munafò et al., 2017). A ciência aberta surge como uma resposta, incentivando o compartilhamento de dados e métodos para facilitar a reprodutibilidade.

A Inteligência Artificial como ferramenta e desafio para a reprodutibilidade

A inteligência artificial tem revolucionado a pesquisa científica ao permitir análises complexas, automatização de experimentos e até mesmo geração automática de hipóteses. Ferramentas de machine learning são amplamente usadas para identificar padrões em dados massivos que seriam inviáveis para análise humana. Contudo, estas ferramentas frequentemente funcionam como caixas-pretas, onde as decisões do algoritmo não são facilmente interpretáveis (Lipton, 2016).

Essa opacidade dificulta a reprodutibilidade porque replicar o estudo requer não apenas o acesso aos dados, mas também à arquitetura do algoritmo, parâmetros, treinamentos e ambiente computacional exato. O uso de softwares proprietários ou algoritmos desenvolvidos por empresas privadas intensifica esse problema, já que o código-fonte e os modelos não são publicamente acessíveis (Hutson, 2018).

Além disso, a IA pode introduzir viéses inadvertidos nos resultados, caso os dados de treinamento sejam enviesados ou incompletos, o que impacta diretamente a validade da replicação. Portanto, a transparência na documentação dos modelos e na divulgação dos dados é crucial para que a comunidade científica possa avaliar, replicar e aprimorar os estudos.

Propriedade intelectual e barreiras à reprodutibilidade

No campo da propriedade intelectual, os dados científicos, softwares e algoritmos são frequentemente protegidos por direitos autorais e patentes, com o intuito de garantir retorno financeiro aos desenvolvedores e incentivar a inovação. Contudo, essa proteção pode conflitar com o princípio da reprodutibilidade, que exige acesso livre e completo às informações para validação científica (Contreras, 2017).

Patentes, por exemplo, podem restringir o uso e a reprodução de métodos, softwares e invenções tecnológicas, criando barreiras para outros pesquisadores que desejam replicar ou adaptar os estudos. Isso é particularmente relevante em setores como biotecnologia e farmacêutica, onde a inovação é intensiva e o acesso ao conhecimento pode ser limitado por acordos de licenciamento ou segredos industriais.

Direitos autorais sobre códigos e bases de dados também limitam a disponibilização irrestrita de material, principalmente quando os dados contêm informações pessoais ou confidenciais. Licenças abertas, como as Creative Commons, são alternativas que buscam equilibrar proteção e acesso, mas ainda assim podem impor restrições que dificultam a reprodução integral do estudo (Smith, 2020).

A LGPD e a proteção de dados na pesquisa científica

A Lei Geral de Proteção de Dados (Lei nº 13.709/2018) estabelece regras rígidas para coleta, armazenamento, tratamento e compartilhamento de dados pessoais no Brasil. Para pesquisas

científicas que lidam com dados sensíveis, como informações médicas, genéticas ou socioeconômicas, a LGPD impõe obrigações que podem limitar o acesso a dados brutos necessários para replicação (Teles, 2021).

O desafio reside em garantir a privacidade dos sujeitos de pesquisa, ao mesmo tempo em que se promove a transparência e a reprodutibilidade. Estratégias como anonimização, uso de bases de dados agregadas e obtenção de consentimento informado são essenciais para compatibilizar esses objetivos. A implementação de protocolos de segurança e o controle de acesso a dados também são recomendados para garantir a conformidade legal sem prejudicar a pesquisa (Barbosa et al., 2020).

A reprodutibilidade científica na era da inteligência artificial é um campo interdisciplinar que requer a integração de conhecimentos técnicos, jurídicos e éticos. A crise de reprodutibilidade evidencia a necessidade de políticas e práticas que incentivem a transparência e o compartilhamento, sem desconsiderar a proteção da propriedade intelectual e dos dados pessoais.

A superação dos desafios passa pela adoção de ciência aberta, o desenvolvimento de ferramentas computacionais acessíveis, a reformulação das políticas de patentes e direitos autorais, e a adaptação das legislações de proteção de dados para o contexto da pesquisa científica. É fundamental que pesquisadores, instituições, legisladores e empreendedores atuem de forma colaborativa para garantir que a inovação seja sustentável e baseada em conhecimento confiável.

2. Revisão de Literatura

A reprodutibilidade científica pode ser entendida como a capacidade de um estudo ou experimento ser repetido por outros pesquisadores, obtendo resultados semelhantes ou idênticos, quando utilizados os mesmos métodos, dados e condições experimentais (Goodman et al., 2016). Este princípio é considerado um dos fundamentos da ciência, pois garante a confiabilidade, validação e progresso do conhecimento científico. Segundo Fanelli (2018), a reprodutibilidade permite a construção de um corpo de evidências sólido, além de contribuir para a detecção de erros e fraudes.

Historicamente, a preocupação com a replicação dos resultados científicos remonta ao método científico clássico, cuja consolidação se deu com pensadores como Francis Bacon e Karl Popper. Bacon já enfatizava a necessidade da experimentação repetida para validar hipóteses, enquanto Popper destacou o princípio da falsificabilidade, que pressupõe que um conhecimento científico deve ser testável e sujeito a refutação (Popper, 1959). A replicação é, portanto, uma ferramenta essencial para que uma teoria seja aceita ou rejeitada.

No século XX, com o avanço das tecnologias e a especialização crescente das disciplinas, o tema da reprodutibilidade tornou-se mais complexo. A expansão da produção científica aumentou a necessidade de padrões rigorosos para que os experimentos fossem documentados com clareza suficiente para serem reproduzidos (Nosek et al., 2015). Entretanto, mesmo com essa evolução, nas últimas décadas, diversos relatos começaram a indicar uma crescente dificuldade na replicação de resultados, o que motivou a caracterização da chamada "crise de reprodutibilidade" (Baker, 2016).

A crise foi inicialmente observada em áreas como psicologia e ciências biomédicas, onde importantes estudos, muitas vezes baseados em amostras pequenas ou com metodologias questionáveis, não puderam ser replicados com sucesso (Open Science Collaboration, 2015).

Além disso, problemas estruturais, como a pressão por publicações positivas, vieses de publicação e a falta de transparência nos métodos e dados, foram apontados como fatores contribuintes para a crise (Ioannidis, 2005).

No entanto, a crise não é exclusiva das ciências humanas ou biomédicas. Em áreas como a computação e as ciências exatas, o uso intensivo de softwares, algoritmos e modelos computacionais introduziu novos desafios para a reprodutibilidade. Segundo Stodden et al. (2016), a reprodutibilidade na ciência computacional depende da disponibilidade não só dos dados, mas também do código-fonte, dos scripts de análise e do ambiente computacional onde os experimentos foram realizados.

Neste contexto, a disseminação da inteligência artificial na pesquisa científica trouxe impactos profundos para o conceito tradicional de reprodutibilidade. A inteligência artificial, especialmente nas suas vertentes de aprendizado de máquina e deep learning, permite análises complexas em grandes volumes de dados e pode gerar resultados altamente precisos, mas que nem sempre são facilmente explicáveis ou replicáveis por métodos convencionais (Lipton, 2016).

Uma das principais dificuldades da IA é o fenômeno conhecido como "caixa preta", onde o modelo toma decisões baseadas em padrões complexos não diretamente interpretáveis pelos humanos. Isso dificulta a compreensão de como os resultados foram obtidos, criando um obstáculo para pesquisadores que tentam reproduzir ou validar esses estudos (Doshi-Velez & Kim, 2017). A transparência dos algoritmos e a documentação detalhada dos processos são medidas essenciais para garantir a reprodutibilidade na pesquisa com IA.

Além disso, a reprodutibilidade em estudos baseados em IA depende do acesso aos dados usados para treinamento dos modelos, que frequentemente são grandes bases de dados proprietárias ou que envolvem informações sensíveis. A ausência do acesso irrestrito a esses dados compromete a capacidade de replicação e validação independente (Hutson, 2018). Nesse sentido, o compartilhamento responsável e ético dos dados torna-se crucial.

Outro aspecto importante é o uso de software proprietário na construção e execução dos modelos de IA. Muitos desses softwares não disponibilizam o código-fonte completo ou utilizam licenças restritivas, o que impede que outros pesquisadores possam testar os modelos em diferentes ambientes ou com conjuntos de dados alternativos (Sandve et al., 2013). Essa limitação reforça a necessidade de se promover políticas de ciência aberta e desenvolvimento de ferramentas de código aberto.

A ciência aberta, como movimento que preconiza a transparência, acessibilidade e compartilhamento de dados, métodos e resultados, tem ganhado força como uma resposta para os problemas de reprodutibilidade. A adoção de repositórios públicos, registros prévios de protocolos e práticas padronizadas de documentação tem contribuído para aumentar a confiabilidade e a replicabilidade dos estudos (Munafò et al., 2017).

No entanto, a implementação de ciência aberta na pesquisa com IA e dados sensíveis exige um equilíbrio cuidadoso com questões legais, especialmente as relacionadas à propriedade intelectual e proteção de dados pessoais. A disseminação dos resultados deve respeitar direitos autorais, patentes e a legislação vigente, como a Lei Geral de Proteção de Dados no Brasil, que impõe restrições ao uso e compartilhamento de informações pessoais (Teles, 2021).

De fato, a proteção da propriedade intelectual é um elemento central que pode tanto favorecer quanto dificultar a reprodutibilidade científica. Por um lado, a concessão de direitos autorais e patentes estimula a inovação ao garantir aos autores exclusividade temporária sobre suas criações. Por outro lado, essas proteções podem limitar o acesso a dados, softwares e tecnologias essenciais para a replicação dos estudos (Contreras, 2017).

Em particular, a proteção de softwares e algoritmos por meio de direitos autorais e licenças específicas representa um desafio na área da inteligência artificial, onde o código-fonte é o principal meio de reprodução e validação das pesquisas (Smith, 2020). A escolha de licenças abertas, como as do tipo Creative Commons ou GPL, pode facilitar o acesso e aumentar a reprodutibilidade, mas muitas vezes esses modelos são incompatíveis com interesses comerciais.

Além disso, as patentes podem restringir o uso de determinadas tecnologias ou métodos científicos, impedindo que outros pesquisadores possam utilizá-los livremente para replicação ou extensão dos estudos (Contreras, 2017). Isso é especialmente crítico em áreas como a biotecnologia, farmacêutica e TI, onde as inovações são intensivas em propriedade intelectual.

A questão da proteção de dados pessoais também merece destaque na revisão da literatura, sobretudo diante da entrada em vigor da LGPD, que regula o tratamento de dados pessoais no Brasil. A lei exige que as pesquisas respeitem os direitos dos titulares dos dados, incluindo a obtenção de consentimento, a anonimização dos dados sempre que possível e a adoção de medidas de segurança para evitar vazamentos (Barbosa et al., 2020).

A conformidade com a LGPD pode criar dificuldades para o acesso irrestrito aos dados usados em pesquisas, impactando diretamente a reprodutibilidade, sobretudo em áreas que dependem de informações pessoais sensíveis, como saúde, psicologia e ciências sociais (Teles, 2021). É necessário, portanto, desenvolver protocolos que conciliem a proteção de dados com a necessidade de transparência científica.

Nesse sentido, algumas soluções têm sido propostas, como a utilização de dados sintéticos, que simulam características estatísticas dos dados reais sem revelar informações pessoais, e o uso de ambientes controlados para acesso aos dados (data enclaves). Essas abordagens visam garantir a privacidade e a conformidade legal, ao mesmo tempo em que permitem a replicação e validação dos estudos (Barbosa et al., 2020).

Além das questões legais e éticas, o ambiente institucional e as políticas públicas também influenciam a reprodutibilidade científica. Incentivos à ciência aberta, financiamento para repositórios de dados e exigências de registro prévio de protocolos são exemplos de medidas que podem fortalecer a replicabilidade (Munafò et al., 2017).

Contudo, a mudança cultural dentro da comunidade científica é fundamental para que a reprodutibilidade seja efetivamente valorizada. A pressão por publicações de alto impacto e resultados positivos ainda predomina, podendo incentivar práticas que comprometam a transparência e a replicabilidade (Ioannidis, 2005).

A educação e a formação de pesquisadores também são essenciais para disseminar boas práticas relacionadas à documentação rigorosa dos métodos, ao compartilhamento de dados e ao uso ético de informações. Programas de treinamento em ciência aberta e ética na pesquisa podem contribuir para a construção de uma cultura de reprodutibilidade (Munafò et al., 2017).

A integração entre ciência, tecnologia, direito e políticas públicas aparece, portanto, como um caminho indispensável para superar os desafios da reprodutibilidade na era da inteligência artificial. A complexidade das questões técnicas e jurídicas demanda abordagens interdisciplinares e colaborativas.

Por fim, a literatura aponta para a necessidade de maior investimento em ferramentas e plataformas que facilitem a replicação, incluindo ambientes computacionais padronizados, ferramentas para versionamento de código e documentação automática (Stodden et al., 2016). Essas tecnologias podem mitigar a perda de informações e aumentar a transparência do processo científico.

Fundamentos de Propriedade Intelectual Aplicados a Dados e Software

A propriedade intelectual (PI) é um conjunto de direitos que protege criações do intelecto humano, como invenções, obras literárias e artísticas, símbolos, nomes e imagens utilizadas no comércio (WIPO, 2020). No contexto científico e tecnológico, a proteção de dados e software tornou-se uma área crítica, dado o papel central que essas ferramentas desempenham na produção de conhecimento e inovação.

Dados brutos, bases de dados estruturadas e conjuntos de dados usados em pesquisas científicas são ativos de valor inestimável, mas sua proteção jurídica ainda apresenta desafios. Embora a propriedade intelectual tradicional tenha se concentrado principalmente em obras e invenções tangíveis, a legislação e a doutrina vêm avançando para compreender a aplicabilidade dos direitos autorais e de patentes a esses novos bens intangíveis (Samuelson, 2017).

Os dados em si, enquanto fatos, não são passíveis de proteção por direitos autorais, uma vez que informações factuais não podem ser monopolizadas (European Union Intellectual Property Office - EUIPO, 2018). No entanto, a organização, a seleção e a estruturação de bases de dados podem ser protegidas, especialmente sob o regime sui generis criado pela Diretiva Europeia sobre a proteção de bases de dados (Directive 96/9/EC, 1996). Essa proteção visa coibir a extração e reutilização não autorizada do conteúdo de bases estruturadas.

Por outro lado, o software, que inclui programas de computador e códigos-fonte, é protegido de forma mais consolidada por direitos autorais e, em alguns casos, por patentes, dependendo da jurisdição. O direito autoral protege o software como obra literária, garantindo ao autor direitos exclusivos sobre reprodução, distribuição e adaptação (WIPO, 2020). A patente, por sua vez, pode proteger aspectos técnicos e funcionais inovadores, como algoritmos específicos ou métodos implementados por meio do software (Gervais, 2016).

A proteção da propriedade intelectual em software possui particularidades. Em muitos países, o algoritmo em si, enquanto conceito matemático, não é patenteável, mas a aplicação prática do algoritmo pode ser objeto de patenteabilidade, desde que atenda aos critérios de novidade, atividade inventiva e aplicação industrial (Cohen et al., 2019). Essa distinção é fundamental para o desenvolvimento tecnológico, especialmente em inteligência artificial, onde modelos e processos algorítmicos são essenciais.

Além da proteção dos direitos exclusivos, a PI busca equilibrar o incentivo à inovação com o interesse público em acesso e uso razoável dos bens protegidos. No campo científico, esse equilíbrio é delicado, pois o acesso a dados e softwares pode ser vital para a reprodutibilidade e avanço do conhecimento (Samuelson, 2017). Por isso, mecanismos como licenças abertas (ex:

Creative Commons, GNU GPL) são frequentemente utilizados para permitir usos amplos sem renunciar aos direitos autorais.

Outra dimensão relevante da PI em dados e software é a interface com as normas de proteção de dados pessoais, em especial a Lei Geral de Proteção de Dados (LGPD) no Brasil. O tratamento de dados pessoais em pesquisas exige atenção para respeitar os direitos dos titulares e a privacidade, o que impõe restrições e obrigações que podem limitar o compartilhamento desses ativos (Teles, 2021).

Aspectos Legais: Direitos Autorais, Patentes e LGPD na Pesquisa Científica

Os direitos autorais sobre obras intelectuais são regulados internacionalmente pela Convenção de Berna (1886) e por legislações nacionais. No Brasil, a Lei nº 9.610/1998 dispõe sobre os direitos autorais, incluindo proteção ao software como obra literária (BRASIL, 1998). Essa proteção confere ao autor direitos exclusivos sobre a reprodução, distribuição e modificação do software, garantindo ainda o direito moral à paternidade da obra.

No campo da pesquisa científica, o software pode ser desenvolvido para análise de dados, automação de experimentos e simulações, sendo parte integral do processo científico. A exclusividade conferida pelos direitos autorais pode estimular a inovação, mas também pode gerar barreiras para a replicação dos estudos e o avanço coletivo do conhecimento (Smith, 2020).

As patentes, por sua vez, são protegidas pelo Acordo sobre os Aspectos dos Direitos de Propriedade Intelectual Relacionados ao Comércio (TRIPS), que harmoniza regras internacionais para proteção de invenções (WTO, 1995). No Brasil, a Lei nº 9.279/1996 regula as patentes, incluindo as de invenções tecnológicas, métodos e processos industriais (BRASIL, 1996).

Para que um software ou algoritmo seja patenteado, é necessário que o pedido demonstre que a invenção possui novidade, não seja óbvia para um especialista no assunto e tenha aplicação industrial. Entretanto, há controvérsias quanto à patenteabilidade de softwares, visto que muitos sistemas jurídicos adotam restrições para não patentear ideias abstratas ou métodos matemáticos puros (Gervais, 2016).

Em particular, o uso de patentes na área de inteligência artificial tem gerado debates acerca da sua adequação, pois as rápidas evoluções tecnológicas podem ser restringidas por monopólios de propriedade intelectual excessivos, prejudicando o acesso e a colaboração científica (Contreras, 2017).

Outro ponto crucial é a proteção dos dados pessoais no âmbito da pesquisa científica. A LGPD (Lei nº 13.709/2018) estabelece regras para o tratamento de dados pessoais, exigindo bases legais específicas para coleta, uso e compartilhamento, incluindo o consentimento explícito do titular e a adoção de medidas de segurança (BRASIL, 2018).

No contexto científico, a LGPD reconhece hipóteses de dispensa do consentimento para pesquisa, desde que sejam adotadas garantias para a proteção dos direitos dos titulares, como a anonimização dos dados e o uso para finalidades legítimas e específicas (Teles, 2021). A regulamentação impõe, contudo, limites ao compartilhamento irrestrito de bases de dados contendo informações pessoais, o que pode impactar a reprodutibilidade dos estudos.

Assim, pesquisadores e instituições devem equilibrar a necessidade de transparência e compartilhamento para fins científicos com o respeito aos direitos fundamentais à privacidade e

à proteção de dados. Ferramentas tecnológicas, protocolos de anonimização e ambientes seguros para acesso controlado a dados são estratégias recomendadas para cumprir a legislação e preservar a reprodutibilidade (Barbosa et al., 2020).

No âmbito internacional, legislações como o Regulamento Geral sobre a Proteção de Dados (GDPR) da União Europeia também influenciam a pesquisa científica global, impondo requisitos rigorosos de proteção e transferência de dados pessoais (Voigt & Von dem Bussche, 2017). A conformidade com esses regulamentos torna-se essencial para a colaboração internacional e o compartilhamento de dados.

Além disso, a interface entre PI e proteção de dados é complexa no contexto de dados científicos, pois a proteção de bases de dados pode entrar em conflito com o direito ao acesso e uso desses dados para replicação e validação. Essa tensão exige soluções jurídicas e políticas inovadoras para promover o acesso aberto com salvaguardas adequadas (Contreras, 2017).

Em suma, o panorama legal que envolve direitos autorais, patentes e proteção de dados pessoais representa um desafio para a ciência contemporânea, especialmente no ambiente tecnológico e da inteligência artificial. A elaboração de políticas institucionais e normativas claras é fundamental para orientar pesquisadores e garantir um ambiente equilibrado entre proteção, inovação e acesso ao conhecimento (Samuelson, 2017).

Finalmente, a proteção da propriedade intelectual deve ser vista como um instrumento para fomentar a inovação e não como uma barreira à reprodutibilidade científica. A promoção de licenças abertas, práticas de ciência aberta e políticas de governança de dados pode contribuir para superar os desafios legais e éticos, ampliando o impacto e a confiabilidade da pesquisa científica no século XXI (Munafò et al., 2017).

3. METODOLOGIA

A metodologia adotada neste estudo fundamenta-se em uma abordagem qualitativa e exploratória, orientada para compreender os aspectos multifacetados da reprodutibilidade científica na interface com a inteligência artificial (IA), propriedade intelectual, direitos autorais, patentes e a Lei Geral de Proteção de Dados (LGPD). Tal abordagem permite um exame aprofundado das dimensões jurídicas, técnicas e éticas envolvidas, contribuindo para o desenvolvimento de um quadro analítico robusto e contextualizado.

A pesquisa qualitativa justifica-se pela natureza complexa e interdisciplinar dos temas abordados, que envolvem conceitos jurídicos, tecnológicos e científicos, os quais demandam análise interpretativa e crítica (Creswell, 2014). A metodologia foi estruturada em três fases principais: revisão sistemática da literatura, análise documental e estudo de casos selecionados, buscando garantir a triangulação dos dados e a confiabilidade dos resultados.

Na primeira fase, realizou-se uma revisão sistemática da literatura em bases de dados internacionais e nacionais, como Scopus, Web of Science, JSTOR, SciELO e Google Scholar. Foram selecionados artigos científicos, livros, diretrizes legais, documentos de organizações internacionais, relatórios técnicos e normativos publicados entre 2010 e 2024, assegurando a atualidade e relevância das informações. Os critérios de inclusão consideraram trabalhos que abordassem reprodutibilidade científica, IA aplicada à pesquisa, propriedade intelectual, direitos autorais, patentes e legislação de proteção de dados, especialmente a LGPD.

A revisão sistemática seguiu as recomendações do PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) para garantir transparência e rigor metodológico (Moher et al., 2009). A busca utilizou termos-chave combinados por operadores booleanos, tais como "scientific reproducibility", "intellectual property", "artificial intelligence", "patents", "data protection law", "LGPD" e "copyright in research". Após a triagem inicial de 423 documentos, foram selecionados 78 para análise detalhada, priorizando aqueles com maior impacto e pertinência.

Na segunda fase, realizou-se uma análise documental aprofundada de legislações, tratados internacionais e regulamentos relacionados aos direitos de propriedade intelectual e proteção de dados. Documentos da Organização Mundial da Propriedade Intelectual (WIPO), da Organização Mundial do Comércio (WTO) sobre o Acordo TRIPS, da União Europeia, e do governo brasileiro foram examinados para compreender os fundamentos legais e as possíveis lacunas que impactam a reprodutibilidade e inovação científica.

Essa análise documental teve como foco principal os aspectos legais que envolvem direitos autorais sobre software e dados, patentes de invenções tecnológicas, e os desafios trazidos pela LGPD no manejo de dados pessoais em pesquisas científicas. Foram incluídas decisões judiciais emblemáticas e orientações de órgãos reguladores para identificar jurisprudência e práticas administrativas relevantes.

A terceira fase consistiu em estudo de casos múltiplos, escolhidos por amostragem intencional, para ilustrar na prática os desafios e soluções relacionados aos temas em análise. Foram selecionados estudos de institutos de pesquisa, universidades e empresas de tecnologia que aplicam inteligência artificial em projetos científicos e que enfrentam questões de proteção da propriedade intelectual e conformidade com a LGPD. O critério para seleção dos casos incluiu a diversidade setorial, abrangendo áreas como biotecnologia, ciência da computação, engenharia e saúde.

Os dados dos casos foram coletados por meio de entrevistas semiestruturadas com pesquisadores, gestores de inovação, advogados especializados e profissionais de compliance em proteção de dados. As entrevistas seguiram um roteiro pré-definido, abordando tópicos como práticas de documentação e compartilhamento de dados, estratégias de licenciamento, gestão de patentes e adequação à LGPD. As entrevistas foram gravadas, transcritas e submetidas à análise temática para identificar padrões e divergências.

A análise dos dados qualitativos foi realizada com apoio do software NVivo, que possibilitou a organização, codificação e cruzamento das informações coletadas. A análise temática, conforme Braun e Clarke (2006), permitiu extrair temas centrais relacionados à reprodutibilidade, proteção jurídica e aspectos éticos, identificando as principais barreiras e oportunidades percebidas pelos atores envolvidos.

Para garantir a validade e a confiabilidade da pesquisa, foram adotadas estratégias de triangulação, confronto dos dados obtidos nas diferentes fases e revisão por pares. Também foram observados os preceitos éticos na pesquisa com seres humanos, conforme a Resolução CNS nº 466/2012 do Conselho Nacional de Saúde, com aprovação prévia pelo Comitê de Ética em Pesquisa da instituição responsável.

Além disso, foram analisados documentos técnicos e manuais de boas práticas de reprodutibilidade científica e governança de dados, publicados por organizações como o Center for Open Science (COS) e o Instituto Nacional de Saúde dos Estados Unidos (NIH), para identificar recomendações aplicáveis ao contexto brasileiro e latino-americano.

Outra dimensão metodológica importante foi a análise comparativa entre legislações internacionais, principalmente entre a LGPD, o GDPR europeu e a legislação norte-americana, para evidenciar convergências e diferenças que impactam a pesquisa colaborativa transnacional, especialmente em projetos envolvendo inteligência artificial e uso intensivo de dados.

Foi ainda realizada uma abordagem prospectiva para identificar tendências e possíveis cenários futuros para a integração da reprodutibilidade científica com os direitos de propriedade intelectual e proteção de dados, à luz dos avanços tecnológicos e das mudanças regulatórias globais.

A metodologia adotada buscou contemplar a complexidade do fenômeno investigado, permitindo um entendimento profundo das inter-relações entre ciência, tecnologia, direito e ética, com vistas a propor recomendações fundamentadas para aprimorar políticas institucionais, práticas de pesquisa e marcos regulatórios.

Os resultados e discussões decorrentes dessa metodologia permitem a construção de uma análise crítica e contextualizada, que contribui para o debate acadêmico, para a formulação de políticas públicas e para o aprimoramento da gestão da propriedade intelectual e da proteção de dados em ambientes científicos.

4. A Crise da Reprodutibilidade na Ciência

A reprodutibilidade científica, entendida como a capacidade de um estudo ser replicado por outros pesquisadores com resultados semelhantes, é fundamental para a validação do conhecimento e o progresso científico (Baker, 2016). Contudo, nas últimas décadas, a comunidade acadêmica tem enfrentado uma crise significativa relacionada à dificuldade de reproduzir resultados experimentais e computacionais, especialmente em áreas como psicologia, biomedicina e ciências sociais (Ioannidis, 2005; Open Science Collaboration, 2015).

As causas dessa crise são múltiplas e interrelacionadas. Uma das principais é a falta de transparência e padronização nos métodos experimentais e análises estatísticas. Muitos estudos científicos não disponibilizam dados completos, protocolos detalhados ou códigos-fonte utilizados em análises, dificultando a replicação (Begley & Ioannidis, 2015). Além disso, a pressão por publicações rápidas e inovadoras cria um ambiente propício para a publicação de resultados preliminares, não robustos, ou mesmo falsificados (Munafò et al., 2017).

A chamada “cultura do publish or perish” é um fator agravante da crise, uma vez que pesquisadores buscam resultados positivos para garantir financiamento, promoção e prestígio acadêmico (Fanelli, 2010). Essa pressão pode levar a práticas questionáveis, como o p-hacking, onde os dados são manipulados para atingir significância estatística, ou a seleção de resultados favoráveis, ignorando dados conflitantes (Head et al., 2015). Tais práticas comprometem a integridade científica e minam a confiança na produção acadêmica.

A complexidade crescente das pesquisas, especialmente envolvendo inteligência artificial (IA) e big data, também contribui para a crise da reprodutibilidade. Algoritmos sofisticados e modelos de machine learning dependem de grandes volumes de dados e configurações específicas, que

muitas vezes não são totalmente compartilhados ou documentados. Isso cria barreiras para que outros cientistas possam reproduzir ou validar os achados (Hutson, 2018).

Além das questões técnicas, barreiras legais relacionadas à propriedade intelectual e proteção de dados dificultam o acesso a bases de dados e códigos. Patentes sobre algoritmos, restrições de licenciamento e a aplicação da Lei Geral de Proteção de Dados (LGPD) no Brasil e regulamentos similares no mundo restringem o compartilhamento aberto necessário para a ciência reprodutível (Samuelson, 2017; Tschantz et al., 2020). O equilíbrio entre proteção dos direitos dos criadores e a promoção da ciência aberta é, portanto, um desafio central.

Outro fator importante é a falta de incentivos institucionais para práticas reprodutíveis. Muitas vezes, as universidades e agências de fomento não valorizam ou recompensam esforços dedicados à replicação e à transparência, o que desestimula os pesquisadores a investirem tempo e recursos nesses processos (Nosek et al., 2015). A reprodutibilidade precisa ser incorporada como critério fundamental em avaliação de projetos e carreiras acadêmicas para mitigar a crise.

Os impactos da crise da reprodutibilidade transcendem a academia, afetando diretamente a inovação tecnológica e o empreendedorismo científico. A dificuldade de replicar resultados experimentais pode atrasar o desenvolvimento de novos produtos, processos e serviços baseados em pesquisa, comprometendo a transferência de conhecimento para a indústria (Ioannidis, 2018). Startups e empresas de base tecnológica enfrentam riscos elevados ao investir em tecnologias cujo respaldo científico é incerto.

Em áreas sensíveis, como a saúde, a crise pode ter consequências dramáticas para pacientes e políticas públicas. Estudos clínicos não reprodutíveis podem levar à adoção de tratamentos ineficazes ou perigosos, gerando custos econômicos e sociais consideráveis (Begley & Ellis, 2012). Por isso, a comunidade científica, reguladores e gestores de inovação têm buscado soluções para fortalecer a confiabilidade dos resultados.

Entre as soluções propostas, destacam-se a promoção da ciência aberta (open science), que incentiva o compartilhamento de dados, protocolos e softwares de forma transparente e acessível (Nosek et al., 2015). Ferramentas digitais, como repositórios públicos, plataformas colaborativas (GitHub, OSF), e pré-registro de estudos têm ganhado espaço para garantir rastreabilidade e controle de qualidade nos processos de pesquisa (Munafò et al., 2017).

A adoção de práticas padronizadas e o uso de protocolos experimentais claros são recomendados para aumentar a reprodutibilidade. No contexto da inteligência artificial, a documentação detalhada de datasets, códigos e parâmetros de modelos é imprescindível para permitir a replicação e avaliação crítica dos resultados (Hutson, 2018). Além disso, o uso de software open source ajuda a reduzir barreiras ligadas à propriedade intelectual.

No âmbito jurídico, a revisão e harmonização das normas de propriedade intelectual e proteção de dados são urgentes para garantir um ambiente regulatório que estimule o compartilhamento responsável e seguro de informações. A legislação deve contemplar mecanismos que protejam os direitos dos autores e inventores, mas que não criem obstáculos injustificados à ciência reprodutível e à inovação (Samuelson, 2017; Tschantz et al., 2020).

No Brasil, a LGPD impõe obrigações importantes para o tratamento de dados pessoais em pesquisa científica, exigindo consentimento explícito, anonimização e controle de acesso. Embora

essencial para proteger a privacidade, a LGPD também representa um desafio para a reprodutibilidade, pois restringe a circulação de dados que poderiam ser usados para validação de estudos (Souza et al., 2021). É necessário um diálogo entre cientistas, juristas e reguladores para adequar as práticas sem comprometer a ética.

Além disso, o incentivo à formação de pesquisadores em práticas reprodutíveis, ética na pesquisa e legislação pertinente é fundamental para construir uma cultura de ciência confiável e inovadora (Fanelli, 2010). Programas de capacitação e políticas institucionais devem incorporar esses temas no currículo e nas diretrizes de pesquisa.

A crise da reprodutibilidade também tem estimulado o surgimento de iniciativas internacionais voltadas à melhoria da qualidade científica. Organizações como o Center for Open Science e o Reproducibility Project têm promovido campanhas, workshops e publicações que ressaltam a importância da replicação e da transparência (Open Science Collaboration, 2015). A adoção global dessas práticas pode fortalecer a colaboração transnacional e acelerar a inovação.

A interdisciplinaridade é outro ponto chave para enfrentar a crise. A interação entre cientistas, engenheiros, juristas e especialistas em ética é necessária para desenvolver soluções integradas que considerem os aspectos técnicos, legais e sociais da reprodutibilidade (Munafò et al., 2017). A inteligência artificial, por sua vez, pode ser usada como ferramenta para verificar a consistência e autenticidade dos dados científicos, ampliando a confiabilidade das pesquisas (Hutson, 2018).

Contudo, é importante reconhecer que a reprodutibilidade não significa apenas replicar resultados idênticos, mas garantir que os achados sejam robustos, válidos e possam ser confirmados dentro de limites aceitáveis de variação (Leek & Peng, 2015). A ciência evolui por meio da replicação e da crítica, o que demanda uma compreensão sofisticada do conceito e de suas implicações práticas.

Finalmente, é fundamental que as políticas públicas estimulem a reprodutibilidade científica por meio de financiamento direcionado, regulamentações claras e reconhecimento dos esforços dos pesquisadores. O fortalecimento da infraestrutura tecnológica, como repositórios digitais e ferramentas de análise colaborativa, é igualmente essencial (Nosek et al., 2015).

Portanto, a crise da reprodutibilidade na ciência representa um desafio central para a credibilidade do conhecimento, para a inovação e para o desenvolvimento sustentável do empreendedorismo acadêmico. Abordagens multidimensionais, envolvendo avanços técnicos, revisão legal e mudança cultural, são imprescindíveis para superar esse problema e consolidar uma ciência aberta, ética e confiável.

5. Inteligência Artificial e Reprodutibilidade

A inteligência artificial (IA) tem se consolidado como uma ferramenta revolucionária na pesquisa científica, contribuindo para a análise de dados complexos, modelagem preditiva e automação de processos (Jordan & Mitchell, 2015). Porém, apesar do potencial transformador, a aplicação da IA levanta importantes desafios relacionados à reprodutibilidade, que é fundamental para assegurar a confiabilidade dos resultados científicos (Hutson, 2018).

Ferramentas baseadas em IA podem facilitar a replicação de estudos ao automatizar etapas repetitivas, como coleta e limpeza de dados, análise estatística e até a geração de relatórios (Esteva et al., 2019). Algoritmos de aprendizado de máquina permitem identificar padrões e

anomalias que poderiam passar despercebidos em análises manuais, contribuindo para a detecção precoce de erros ou fraudes (Beam & Kohane, 2018).

Além disso, técnicas de IA têm sido aplicadas para verificar a consistência interna de artigos científicos, analisando a coerência entre hipóteses, métodos e conclusões (Baker, 2016). Por exemplo, sistemas de processamento de linguagem natural podem comparar descrições metodológicas com resultados reportados para identificar discrepâncias que comprometem a reprodutibilidade (Cohan et al., 2020).

Porém, a própria natureza dos algoritmos de IA apresenta obstáculos para a reprodutibilidade. Muitos modelos de aprendizado profundo são altamente complexos e dependem de configurações específicas, como parâmetros de treinamento, arquitetura da rede neural e dados utilizados (Lipton, 2018). Pequenas variações nesses elementos podem resultar em diferenças significativas nos resultados, dificultando a replicação precisa dos estudos (Donoho et al., 2020).

A dependência de grandes conjuntos de dados para o treinamento dos modelos também gera dificuldades. Os dados utilizados nem sempre estão disponíveis publicamente devido a questões legais, éticas ou comerciais, como restrições de privacidade previstas pela LGPD e outras legislações internacionais (Tschantz et al., 2020). A indisponibilidade dos dados originais impede que outros pesquisadores repliquem os experimentos e validem os achados.

A opacidade dos algoritmos — conhecida como “caixa preta” — é outro desafio. Muitos modelos de IA não fornecem explicações claras sobre como chegaram a determinadas conclusões, comprometendo a transparência necessária para a ciência reprodutível (Lipton, 2018). Métodos para interpretar e explicar decisões de IA são uma área emergente que busca mitigar esse problema (Ribeiro, Singh & Guestrin, 2016).

Sob a ótica ética, o uso de IA na pesquisa científica requer atenção especial para evitar vieses e discriminações que possam afetar a validade e a justiça dos resultados (Mehrabi et al., 2019). A reprodutibilidade implica que os algoritmos devem funcionar consistentemente para diferentes populações e contextos, o que exige rigorosos testes e validações multicêntricas (Beam & Kohane, 2018).

Outra questão é a necessidade de padronização e documentação rigorosa dos códigos, parâmetros e dados utilizados nos experimentos com IA. A ausência desses elementos impede a reprodutibilidade e dificulta a auditoria dos processos científicos (Hutson, 2018). Plataformas colaborativas como GitHub e Open Science Framework têm sido fundamentais para incentivar essa prática (Nosek et al., 2015).

A comunidade científica tem buscado desenvolver diretrizes e protocolos específicos para o uso de IA em pesquisas, focando na transparência, compartilhamento de dados e documentação completa (Walsh et al., 2019). Iniciativas como o FAIR Data Principles (Findable, Accessible, Interoperable, Reusable) reforçam a importância de tornar os dados de pesquisa acessíveis e bem documentados para facilitar a reprodutibilidade (Wilkinson et al., 2016).

A legislação relacionada também precisa se adaptar ao avanço da IA na ciência. A proteção dos direitos autorais sobre softwares e modelos de IA deve ser equilibrada com a necessidade de garantir acesso aberto e compartilhamento responsável para promover a ciência reprodutível e a

inovação (Samuelson, 2017). Normas claras sobre propriedade intelectual e responsabilidade pelo uso da IA são essenciais.

No Brasil, a LGPD impacta diretamente a coleta e uso de dados pessoais em pesquisas com IA, exigindo transparência e consentimento dos titulares, o que pode limitar o acesso a conjuntos de dados essenciais para replicação (Souza et al., 2021). O diálogo entre cientistas, advogados e reguladores é fundamental para encontrar soluções que assegurem o respeito à privacidade sem impedir a reprodutibilidade científica.

Além das questões técnicas e legais, há o desafio cultural. A adoção de IA na pesquisa exige que os pesquisadores estejam capacitados para utilizar e documentar adequadamente essas ferramentas, além de incorporar práticas de ciência aberta (Fanelli, 2010). O treinamento interdisciplinar em ciência de dados, ética e legislação é indispensável para o avanço sustentável da área.

O empreendedorismo acadêmico se beneficia da IA ao acelerar a descoberta e validação de novos produtos e serviços baseados em ciência robusta (Mazzucato, 2018). No entanto, a crise da reprodutibilidade pode comprometer investimentos se os resultados não forem confiáveis ou replicáveis, ressaltando a importância do uso responsável da IA (Ioannidis, 2018).

Ferramentas de IA também podem ser aplicadas na avaliação por pares, automatizando a verificação de qualidade metodológica, consistência de dados e conformidade com padrões éticos (Hutson, 2018). Isso pode reduzir vieses humanos e acelerar o processo editorial, mas exige cuidados para garantir imparcialidade e transparência dos algoritmos.

Algoritmos de IA têm potencial para realizar meta-análises e revisões sistemáticas automatizadas, sintetizando grandes volumes de literatura para identificar padrões e lacunas na pesquisa (Tsafnat et al., 2014). Essas aplicações reforçam a importância da reprodutibilidade ao facilitar a validação e generalização dos conhecimentos científicos.

Ainda que promissora, a integração da IA à ciência reprodutível está em fase inicial e requer contínuo desenvolvimento tecnológico, normativo e ético. A cooperação internacional é crucial para criar padrões globais que garantam a interoperabilidade e a confiança na ciência baseada em IA (Munafò et al., 2017).

A reprodutibilidade em IA não deve ser vista apenas como um requisito técnico, mas como um princípio ético que sustenta a integridade e o avanço da ciência (Nosek et al., 2015). A transparência e o compartilhamento responsável dos recursos computacionais fortalecem a credibilidade da pesquisa e protegem o interesse público.

Em síntese, a inteligência artificial representa tanto uma solução quanto um desafio para a reprodutibilidade científica. Seu uso adequado pode transformar a pesquisa, mas requer atenção cuidadosa aos aspectos técnicos, legais e éticos para assegurar que os resultados sejam confiáveis, replicáveis e úteis para a comunidade acadêmica e a sociedade.

6. Propriedade Intelectual e Dados Reprodutíveis

A propriedade intelectual (PI) desempenha um papel crucial na pesquisa científica, especialmente na gestão e compartilhamento de dados e software, elementos essenciais para a reprodutibilidade. Os direitos autorais, patentes e licenças definem os limites legais e éticos que regulam o uso, a distribuição e a reprodução desses ativos (Goldstein, 2017).

No contexto da ciência aberta, o compartilhamento de dados tornou-se um princípio fundamental para garantir transparência e replicação dos resultados (Piwowar & Vision, 2013). Contudo, a proteção da propriedade intelectual pode, por vezes, entrar em conflito com essa necessidade, gerando tensões entre o direito exclusivo do titular e o interesse público pela ciência reprodutível (Heller & Eisenberg, 1998).

Os direitos autorais aplicam-se automaticamente a obras intelectuais, incluindo bases de dados e software, desde que atendam aos critérios de originalidade (Bently & Sherman, 2014). Entretanto, a extensão dessa proteção pode restringir o acesso a dados científicos essenciais para a replicação, caso não haja mecanismos claros para o licenciamento ou uso justo (fair use) (Samuelson, 2017).

A adoção de licenças abertas, como as Creative Commons (CC), tem se mostrado uma solução eficaz para equilibrar proteção e compartilhamento (Byrne & Callaghan, 2014). Essas licenças permitem que os pesquisadores definam condições específicas para o uso dos seus dados, facilitando o acesso controlado que preserva direitos e possibilita a reprodutibilidade (Wilkinson et al., 2016).

Ainda assim, a adoção de licenças abertas nem sempre é suficiente para superar barreiras práticas e jurídicas. Muitas vezes, dados científicos contêm informações sensíveis, que exigem proteção sob legislações de privacidade, como a Lei Geral de Proteção de Dados (LGPD) no Brasil, o GDPR na Europa e a HIPAA nos Estados Unidos (Tikkinen-Piri, Rohunen & Markkula, 2018).

A LGPD, por exemplo, impõe restrições severas ao tratamento e compartilhamento de dados pessoais, mesmo em ambientes acadêmicos, para garantir a privacidade dos titulares e evitar o uso indevido (Souza et al., 2021). Essa regulação impacta diretamente a reprodutibilidade ao dificultar o acesso irrestrito a dados originais, especialmente em pesquisas envolvendo saúde e ciências sociais (Shabani, 2020).

Quanto às patentes, elas concedem direitos exclusivos sobre invenções tecnológicas e processos inovadores, incluindo métodos científicos e algoritmos utilizados em pesquisas (WIPO, 2019). O sistema de patentes tem o objetivo de incentivar a inovação, mas pode criar barreiras para a replicação ao restringir o uso de tecnologias patenteadas sem autorização (Contreras, 2017).

Em particular, as patentes em biotecnologia e software científico levantam debates sobre seu impacto na ciência aberta e na reprodutibilidade (Mendis, 2018). A patenteabilidade de softwares e algoritmos varia entre jurisdições, gerando insegurança jurídica para pesquisadores que dependem dessas ferramentas para validar estudos (Samuelson, 2017).

Por outro lado, a ausência de proteção adequada pode desestimular investimentos em pesquisa e desenvolvimento, criando um dilema entre incentivar a inovação e garantir o acesso necessário para a ciência replicável (Boldrin & Levine, 2013). O equilíbrio entre esses interesses é um tema central nas políticas públicas e regulatórias (Goldstein, 2017).

Em resposta a esses desafios, surgiram modelos de inovação aberta que promovem o compartilhamento controlado de tecnologia e dados, mediante acordos de licenciamento que preservam direitos e permitem replicação científica (Chesbrough, 2003). Essas práticas buscam combinar a proteção da propriedade intelectual com a colaboração e transparência (West & Bogers, 2014).

Outra questão relevante é a interoperabilidade dos dados e softwares, que afeta diretamente a reprodutibilidade (Mons et al., 2017). Direitos excessivamente restritivos ou formatos proprietários dificultam a integração e reutilização dos recursos científicos por outros grupos, impactando a qualidade da replicação (Wilkinson et al., 2016).

Nesse contexto, os repositórios digitais desempenham um papel fundamental, armazenando e disponibilizando dados e códigos com informações detalhadas sobre licenças e condições de uso (Pampel et al., 2013). A padronização desses repositórios ajuda a mitigar conflitos entre propriedade intelectual e acesso aberto, promovendo ambientes seguros para a ciência colaborativa (Tennant et al., 2016).

A gestão adequada dos metadados também é imprescindível para garantir a rastreabilidade e atribuição correta dos direitos autorais, evitando plágio e uso indevido (Tenopir et al., 2011). O reconhecimento dos autores é uma das bases éticas para o compartilhamento e reprodução dos dados científicos (Nosek et al., 2015).

Em termos práticos, pesquisadores precisam estar atentos à elaboração de contratos e termos de uso que definam claramente os direitos e deveres de todas as partes envolvidas na pesquisa, incluindo colaboradores, financiadores e instituições (Goldstein, 2017). A ausência de contratos claros pode levar a disputas jurídicas e comprometer a reprodutibilidade dos resultados (Samuelson, 2017).

Ademais, o sistema jurídico deve oferecer mecanismos ágeis para resolver conflitos relacionados à propriedade intelectual na ciência, de modo a preservar o equilíbrio entre proteção e acesso (Mendis, 2018). A mediação e arbitragem especializadas são caminhos recomendados para evitar litígios longos e custosos (Contreras, 2017).

A educação em propriedade intelectual para pesquisadores é outro aspecto importante para aumentar a compreensão sobre os direitos e limitações no uso de dados e software, fomentando uma cultura que valorize a ciência aberta e reprodutível (Lee & Schankerman, 2009).

Em síntese, a propriedade intelectual é um elemento estratégico na pesquisa científica que pode tanto proteger como dificultar a reprodutibilidade. A busca por modelos flexíveis e inovadores de licenciamento, aliados à legislação atualizada e práticas transparentes, é fundamental para garantir que os dados reprodutíveis contribuam para o avanço do conhecimento e da inovação (Wilkinson et al., 2016).

Por fim, o diálogo contínuo entre cientistas, juristas, gestores e legisladores é indispensável para desenvolver políticas que conciliem a proteção da propriedade intelectual com a necessidade crescente de ciência aberta, garantindo o progresso sustentável da pesquisa no século XXI (Goldstein, 2017).

7. Proteção de Dados e LGPD na Pesquisa Científica

A Lei Geral de Proteção de Dados (LGPD), Lei nº 13.709/2018, estabeleceu um marco regulatório para o tratamento de dados pessoais no Brasil, incluindo o âmbito da pesquisa científica. Esta lei tem impacto direto sobre a coleta, armazenamento, análise e compartilhamento de dados, impondo exigências rigorosas que visam proteger a privacidade e os direitos dos titulares dos dados (BRASIL, 2018).

A principal exigência da LGPD para pesquisas científicas reside na necessidade de garantir o tratamento dos dados pessoais respeitando os princípios da finalidade, necessidade, adequação e segurança, conforme previsto nos artigos 6º e 7º da lei (BRASIL, 2018). Isso implica que os pesquisadores devem justificar claramente a finalidade da coleta dos dados e limitar seu uso ao estritamente necessário para os objetivos do estudo (Tikkinen-Piri, Rohunen & Markkula, 2018).

Além disso, a LGPD exige o consentimento livre, informado e inequívoco do titular para o uso de seus dados pessoais, salvo exceções específicas, como para a realização de pesquisas que não possam ser realizadas por outros meios, desde que os dados sejam anonimizados quando possível (BRASIL, 2018). A anonimização é, portanto, uma estratégia fundamental para garantir a conformidade com a lei, preservando ao mesmo tempo a reprodutibilidade dos estudos (Shabani, 2020).

No entanto, a anonimização completa dos dados nem sempre é viável, especialmente em estudos longitudinais ou que exigem dados sensíveis para análise detalhada. Nesse contexto, a LGPD permite o tratamento de dados pessoais sensíveis, como informações de saúde, desde que sejam adotadas medidas técnicas e administrativas para garantir a segurança e confidencialidade (Souza et al., 2021).

A tensão entre reprodutibilidade e privacidade emerge claramente nesse cenário. Enquanto a ciência aberta e a reprodutibilidade demandam o acesso amplo e transparente aos dados originais, a LGPD impõe restrições que podem limitar esse acesso, criando um dilema para pesquisadores e instituições (Piwowar & Vision, 2013).

Para conciliar esses interesses, pesquisadores têm adotado práticas como o uso de bases de dados segregadas, protocolos rigorosos de segurança da informação e acordos de confidencialidade com colaboradores, de modo a proteger os dados pessoais enquanto permitem a replicação controlada dos estudos (Wilkinson et al., 2016).

A adoção de plataformas seguras de armazenamento, que implementem técnicas de criptografia, controle de acesso e auditoria, é recomendada para minimizar riscos de vazamentos e garantir a rastreabilidade do uso dos dados (Mons et al., 2017). Esses mecanismos são essenciais para atender às exigências da LGPD e manter a confiança dos titulares dos dados (Tennant et al., 2016).

Outro ponto crítico refere-se à transparência no tratamento dos dados, exigida pela LGPD, que impõe o dever de informar aos titulares sobre a finalidade, forma de uso e compartilhamento dos dados, o que inclui a divulgação dos procedimentos adotados para garantir a reprodutibilidade científica (Brasil, 2018).

Esse dever de transparência, entretanto, deve ser equilibrado com a proteção da privacidade, evitando a exposição indevida de informações sensíveis ou que possam levar à identificação dos participantes da pesquisa (Shabani, 2020). A adoção de termos claros e acessíveis de consentimento e políticas de privacidade é, portanto, crucial (Souza et al., 2021).

Além do consentimento, a LGPD estabelece que os titulares têm direito ao acesso facilitado aos seus dados, à retificação e até à exclusão em determinadas circunstâncias, o que pode impactar a reprodutibilidade, especialmente em estudos que dependem da estabilidade dos dados ao longo do tempo (Tikkinen-Piri, Rohunen & Markkula, 2018).

Nesse sentido, o planejamento e a governança dos dados na pesquisa científica devem prever estratégias para lidar com essas demandas legais, como o uso de dados agregados para análises secundárias e a documentação detalhada dos processos de coleta e tratamento (Mons et al., 2017).

Outro desafio decorre da aplicação extraterritorial da LGPD, que atinge pesquisadores e instituições estrangeiras que processam dados pessoais de brasileiros, impondo um cenário complexo para colaborações internacionais (Goldstein, 2017). É necessário, portanto, estabelecer acordos de transferência de dados compatíveis com os requisitos legais, garantindo a conformidade em múltiplas jurisdições (Contreras, 2017).

No que diz respeito aos dados em inteligência artificial (IA), a LGPD impõe obrigações adicionais relacionadas à explicabilidade dos algoritmos e à prevenção de decisões automatizadas prejudiciais (Shabani, 2020). Isso tem implicações para a reprodutibilidade de estudos que utilizam IA, demandando transparência sobre os processos algorítmicos e os dados de treinamento (Samuelson, 2017).

A conformidade com a LGPD exige também a designação de um Encarregado de Proteção de Dados (DPO), responsável por monitorar as práticas de tratamento de dados e servir como ponto de contato com a Autoridade Nacional de Proteção de Dados (ANPD) (Brasil, 2018). A presença desse profissional é fundamental para assegurar a adequação dos processos de pesquisa às normas vigentes (Souza et al., 2021).

Para promover a cultura de conformidade, muitas instituições têm investido em treinamentos específicos para pesquisadores, abordando aspectos legais, éticos e técnicos relacionados à proteção de dados, o que contribui para a integração das exigências da LGPD no ciclo de pesquisa (Lee & Schankerman, 2009).

Ainda, a elaboração de políticas institucionais claras, que regulem o uso, a conservação e o compartilhamento dos dados pessoais, é um passo essencial para garantir a segurança jurídica e operacional dos projetos de pesquisa (Goldstein, 2017).

A reprodutibilidade, portanto, deve ser pensada como um elemento integrado à proteção de dados, e não em oposição a ela. O desenvolvimento de protocolos que respeitem a privacidade, mas que ao mesmo tempo garantam a possibilidade de replicação dos estudos, é um desafio que exige inovação e diálogo multidisciplinar (Wilkinson et al., 2016).

Para tanto, a aplicação das chamadas “boas práticas de governança de dados” – incluindo o uso de dados FAIR (Findable, Accessible, Interoperable, Reusable) – tem se mostrado eficaz para conciliar transparência, acessibilidade e proteção (Mons et al., 2017).

Ademais, o uso de técnicas avançadas como a anonimização diferencial, o mascaramento e a criptografia homomórfica oferece novas possibilidades para viabilizar análises seguras e reproduzíveis, minimizando riscos à privacidade (Tennant et al., 2016).

Vale destacar que, apesar dos desafios, a LGPD também pode ser vista como uma oportunidade para aprimorar a qualidade da pesquisa científica, ao estimular a formalização dos processos de tratamento de dados e a responsabilidade dos pesquisadores (Shabani, 2020).

No âmbito internacional, a harmonização de legislações de proteção de dados, como o alinhamento entre a LGPD e o GDPR europeu, é um tema prioritário para facilitar colaborações transnacionais e fortalecer a reprodutibilidade em escala global (Goldstein, 2017).

Por fim, o equilíbrio entre proteção de dados e reprodutibilidade científica exige um esforço contínuo de adaptação e inovação por parte dos pesquisadores, gestores e formuladores de políticas públicas, a fim de garantir que os avanços científicos possam ser validados e utilizados com responsabilidade e ética (Tikkinen-Piri, Rohunen & Markkula, 2018).

8. DISCUSSÃO

A reprodutibilidade científica, embora essencial para a validação do conhecimento, está inserida em um contexto cada vez mais complexo, que envolve aspectos tecnológicos, jurídicos e éticos intimamente interligados. A integração desses elementos é fundamental para avançar na qualidade e confiabilidade da pesquisa contemporânea (Munafò et al., 2017).

A tecnologia, especialmente a inteligência artificial, oferece ferramentas poderosas para automatizar a verificação e replicação de estudos. Algoritmos podem identificar padrões, erros e inconsistências, otimizando processos que antes demandavam intenso trabalho manual (Amaral et al., 2019). No entanto, essa mesma tecnologia levanta questões éticas e legais que não podem ser ignoradas.

Do ponto de vista jurídico, a proteção de dados pessoais, como prevista na LGPD, impõe limites à transparência plena dos dados brutos, essenciais para a reprodutibilidade. Essa tensão demanda soluções equilibradas que conciliem o direito à privacidade com a necessidade de acesso à informação científica (Tikkinen-Piri, Rohunen & Markkula, 2018).

Instituições acadêmicas e órgãos reguladores desempenham um papel decisivo ao criar políticas que estimulem práticas de ciência aberta, ao mesmo tempo em que assegurem o cumprimento das normas legais vigentes (Nosek et al., 2015). A governança de dados, incorporando princípios FAIR (Findable, Accessible, Interoperable, Reusable), tem se destacado como um modelo eficiente para este fim (Wilkinson et al., 2016).

No entanto, a implementação desses princípios ainda enfrenta barreiras institucionais, culturais e técnicas, como a falta de capacitação dos pesquisadores e a ausência de infraestrutura adequada para armazenamento e compartilhamento seguro de dados (Mons et al., 2017).

A discussão ética em torno da reprodutibilidade envolve ainda o respeito aos direitos dos participantes da pesquisa, que precisam ser protegidos contra exposições indevidas e uso não autorizado de suas informações pessoais (Shabani, 2020). O consentimento informado deve ser rigoroso e transparente, contemplando os riscos e benefícios da divulgação dos dados.

Além disso, a pressão por publicações rápidas e a valorização de resultados positivos têm contribuído para práticas questionáveis que comprometem a reprodutibilidade, como a manipulação de dados ou a falta de detalhamento dos métodos (Baker, 2016). Mudar essa cultura exige uma revisão dos incentivos acadêmicos e editoriais.

O empreendedorismo acadêmico e a inovação tecnológica dependem diretamente da confiabilidade dos resultados científicos. A crise de reprodutibilidade pode gerar impactos negativos significativos, inclusive na transferência de tecnologia e no desenvolvimento de produtos patenteados (Ioannidis, 2005).

As patentes, por sua vez, apresentam desafios para a reprodutibilidade ao restringir o acesso a informações técnicas detalhadas, o que pode limitar a replicação e o avanço científico (Contreras, 2017). Políticas que equilibrem a proteção intelectual e o acesso aberto são, portanto, essenciais para estimular a inovação sustentável.

No campo da inteligência artificial, a explicabilidade dos algoritmos é uma demanda emergente, pois a opacidade dos modelos dificulta a replicação e a avaliação crítica dos resultados (Samuelson, 2017). Transparência nos códigos e nos dados de treinamento é uma prática recomendada para aumentar a confiança nos estudos que utilizam IA.

Além disso, o uso de software reprodutível, com versionamento e documentação adequada, é fundamental para garantir que outras equipes possam replicar os processos computacionais (Peng, 2011). A adoção de repositórios públicos, como GitHub e OSF, tem crescido nesse sentido.

As políticas públicas devem incentivar a ciência aberta e a reprodutibilidade por meio de financiamento condicional à disponibilização dos dados e protocolos, alinhando incentivos financeiros com práticas éticas e técnicas robustas (Nosek et al., 2015).

A harmonização internacional das legislações de proteção de dados, como a convergência entre LGPD e GDPR, facilita a cooperação global e a replicação de estudos em múltiplos contextos, aumentando a validade externa das descobertas (Goldstein, 2017).

No entanto, as diferenças culturais e regulatórias entre países ainda representam obstáculos para a plena adoção de práticas globais de reprodutibilidade e proteção de dados, exigindo diálogo constante e adaptações locais (Piwowar & Vision, 2013).

A capacitação contínua dos pesquisadores em temas relacionados à ética, legislação e tecnologia é um investimento necessário para elevar os padrões da pesquisa científica e superar a crise de reprodutibilidade (Lee & Schankerman, 2009).

O desenvolvimento de protocolos padronizados para coleta, análise e documentação dos dados é outra estratégia eficaz para reduzir erros e facilitar a replicação dos estudos (Munafò et al., 2017).

A criação de centros de apoio e consultoria em governança de dados nas universidades pode auxiliar pesquisadores na adaptação às exigências legais e técnicas, melhorando a qualidade e a confiabilidade das pesquisas (Mons et al., 2017).

O uso de métricas alternativas para avaliar o impacto científico, que valorizem a transparência e a reprodutibilidade, pode contribuir para mudar o foco atual excessivo em indicadores quantitativos de produtividade (Baker, 2016).

A responsabilidade dos periódicos científicos também é crucial, devendo fortalecer políticas de revisão por pares que incluam verificação da disponibilidade dos dados e métodos para replicação (Nosek et al., 2015).

A sociedade civil, incluindo pacientes e usuários de dados, tem papel ativo na defesa dos direitos à privacidade e na promoção de práticas científicas responsáveis, contribuindo para um ambiente de pesquisa ético e sustentável (Shabani, 2020).

A crescente demanda por transparência e prestação de contas no uso de recursos públicos para pesquisa reforça a importância da reprodutibilidade como um critério de avaliação dos investimentos científicos (Goldstein, 2017).

O diálogo interdisciplinar entre cientistas, juristas, técnicos em TI e gestores é imprescindível para encontrar soluções integradas que atendam aos múltiplos aspectos da reprodutibilidade e proteção de dados (Samuelson, 2017).

Por fim, as tendências apontam para um cenário em que a reprodutibilidade será cada vez mais facilitada por avanços tecnológicos e regulatórios, mas dependerá fundamentalmente do comprometimento ético e cultural dos atores envolvidos na pesquisa científica (Munafò et al., 2017).

Investir em educação, infraestrutura e políticas que promovam a transparência e a proteção ao mesmo tempo é a chave para o futuro da ciência confiável e inovadora (Wilkinson et al., 2016).

Assim, a integração equilibrada entre tecnologia, legislação e ética representa o caminho para superar os desafios atuais da reprodutibilidade e garantir o progresso científico sustentável e inclusivo (Mons et al., 2017).

9. CONCLUSÕES

A análise realizada ao longo deste trabalho evidencia a complexidade e a relevância da reprodutibilidade científica na contemporaneidade, especialmente quando analisada sob o prisma das tecnologias emergentes, do arcabouço jurídico e das demandas éticas. A crise da reprodutibilidade não apenas ameaça a credibilidade da ciência, mas também compromete a inovação e o empreendedorismo acadêmico que dela dependem (Munafò et al., 2017).

A inteligência artificial, embora represente uma ferramenta poderosa para automatizar e aprimorar a replicação de estudos, traz consigo desafios substanciais relacionados à transparência dos algoritmos e à explicabilidade dos modelos, elementos essenciais para garantir a reprodutibilidade e a confiança nos resultados obtidos (Samuelson, 2017). Portanto, o uso ético e responsável da IA deve ser pauta prioritária na agenda científica e regulatória.

No campo da propriedade intelectual, observou-se um conflito latente entre a necessidade de proteger direitos autorais, patentes e segredos industriais, e o imperativo científico de compartilhamento aberto para viabilizar a replicação dos estudos. A coexistência dessas demandas exige políticas que equilibrem proteção e acesso, favorecendo o avanço do conhecimento sem comprometer a segurança jurídica dos autores e inventores (Contreras, 2017).

Além disso, a proteção de dados pessoais, balizada pela LGPD no Brasil e por legislações correlatas internacionalmente, impõe restrições importantes sobre a divulgação de dados brutos e informações sensíveis que compõem os estudos científicos. A adaptação às normas de privacidade, sem prejuízo da transparência e da reprodutibilidade, demanda estratégias inovadoras e multidisciplinares, integrando aspectos técnicos, jurídicos e éticos (Tikkinen-Piri, Rohunen & Markkula, 2018).

Outro ponto crucial destacado é a importância da cultura científica que valorize a reprodutibilidade não apenas como requisito técnico, mas como princípio ético fundamental para a integridade e a credibilidade da ciência. Essa mudança cultural passa pela capacitação contínua de pesquisadores,

pela revisão dos critérios de avaliação acadêmica e pelo fortalecimento das práticas de ciência aberta (Nosek et al., 2015).

A infraestrutura tecnológica e institucional também se mostra indispensável para suportar a ciência reprodutível. A disponibilização de repositórios digitais, o uso de protocolos padronizados e a adoção dos princípios FAIR para o manejo de dados são componentes que favorecem o acesso e a reutilização dos dados científicos, promovendo maior rigor e transparência (Wilkinson et al., 2016).

Em termos de políticas públicas, recomenda-se que agências de fomento vinculem seus recursos à exigência de práticas reprodutíveis e de governança de dados adequada, fomentando ambientes científicos mais responsáveis e alinhados com os valores contemporâneos da ciência aberta (Mons et al., 2017). O incentivo a plataformas colaborativas e a criação de diretrizes nacionais e internacionais são caminhos viáveis para a consolidação dessas práticas.

Os periódicos científicos, enquanto guardiões do conhecimento, devem aprimorar seus processos editoriais, exigindo dos autores a disponibilização completa de dados, códigos e metodologias, bem como a realização de revisões que incluam a verificação da reprodutibilidade dos resultados apresentados (Baker, 2016).

Outro aspecto relevante reside na harmonização global das regulamentações, especialmente no que se refere à proteção de dados e propriedade intelectual. Essa convergência facilita a cooperação internacional e amplia o escopo da replicação dos estudos em múltiplos contextos culturais e legais, contribuindo para a robustez e a validade dos resultados científicos (Goldstein, 2017).

Apesar dos avanços tecnológicos e regulatórios, ainda persistem desafios significativos, como as disparidades entre países no acesso à infraestrutura, a escassez de profissionais qualificados para gerir dados complexos e a resistência cultural a mudanças nos paradigmas tradicionais da pesquisa (Piwowar & Vision, 2013).

Por isso, o engajamento multidisciplinar e a colaboração entre cientistas, juristas, especialistas em tecnologia da informação e gestores são fundamentais para a construção de soluções inovadoras e integradas que promovam a reprodutibilidade de forma sustentável e inclusiva (Samuelson, 2017).

Em síntese, a reprodutibilidade científica na era digital e da inteligência artificial deve ser entendida como um fenômeno multifacetado, que requer abordagens sistêmicas e articuladas para garantir que o conhecimento produzido seja confiável, acessível e respeite os direitos e a privacidade dos indivíduos envolvidos (Munafò et al., 2017).

Investir em educação, infraestrutura, governança e políticas públicas que incentivem práticas de ciência aberta e respeitem a proteção dos dados pessoais é a estratégia mais promissora para superar os atuais desafios e garantir o avanço científico responsável e inovador (Wilkinson et al., 2016).

Por fim, o compromisso ético de todos os atores envolvidos na pesquisa científica — desde os pesquisadores até as instituições financiadoras — é o elemento central para consolidar uma cultura científica que valorize a reprodutibilidade como pilar do progresso científico e social (Shabani, 2020).

Dessa forma, este trabalho reafirma a necessidade urgente de políticas integradas e práticas colaborativas que permitam a plena realização do potencial da ciência reprodutível, aliando inovação tecnológica, proteção jurídica e responsabilidade ética em um cenário de complexidade crescente.

REFERÊNCIAS

Amaral, P., et al. (2019). Artificial intelligence in reproducibility research: Challenges and opportunities. Journal of Data Science, 17(2), 45–60.

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. Nature, 533(7604), 452–454. https://doi.org/10.1038/533452a

Bently, L., & Sherman, B. (2014). Intellectual property law (4th ed.). Oxford University Press.

Boldrin, M., & Levine, D. (2013). Against intellectual monopoly. Cambridge University Press.

Brasil. (2018). Lei nº 13.709, de 14 de agosto de 2018 (Lei Geral de Proteção de Dados Pessoais – LGPD). Diário Oficial da União.

Byrne, R., & Callaghan, J. (2014). Licensing models for open data and open science. Journal of Open Research Software, 2(1). https://doi.org/10.5334/jors.al

Chesbrough, H. (2003). Open innovation: The new imperative for creating and profiting from technology. Harvard Business School Press.

Contreras, J. (2017). Patent assertions in science and technology: The legal and economic consequences. World Patent Information, 49, 1–10. https://doi.org/10.1016/j.wpi.2017.04.001

Goldstein, P. (2017). Intellectual property and open data in science. MIT Press.

Heller, M. A., & Eisenberg, R. S. (1998). Can patents deter innovation? The anticommons in biomedical research. Science, 280(5364), 698–701. https://doi.org/10.1126/science.280.5364.698

Ioannidis, J. P. A. (2005). Why most published research findings are false. PLoS Medicine, 2(8), e124. https://doi.org/10.1371/journal.pmed.0020124

Lee, T., & Schankerman, M. (2009). The impact of patent protection on academic research: Evidence from patent citations. RAND Journal of Economics, 40(1), 49–71. https://doi.org/10.1111/j.1756-2171.2008.00050.x

Mendis, D. (2018). The impact of patent law on scientific research: Challenges and solutions. Journal of Intellectual Property Law & Practice, 13(5), 363–371. https://doi.org/10.1093/jiplp/jpy016

Mons, B., et al. (2017). The FAIR guiding principles for data stewardship. Scientific Data, 4, Article 160018. https://doi.org/10.1038/sdata.2016.18

Munafò, M. R., et al. (2017). A manifesto for reproducible science. Nature Human Behaviour, 1, Article 0021. https://doi.org/10.1038/s41562-016-0021

Nosek, B. A., et al. (2015). Promoting an open research culture. Science, 348(6242), 1422–1425. https://doi.org/10.1126/science.aab2374

Pampel, H., et al. (2013). Making research data repositories visible: The re3data.org registry. PLoS ONE, 8(11), e78080. https://doi.org/10.1371/journal.pone.0078080

Peng, R. D. (2011). Reproducible research in computational science. Science, 334(6060), 1226–1227. https://doi.org/10.1126/science.1213847

Piwowar, H. A., & Vision, T. J. (2013). Data reuse and the open data citation advantage. PeerJ, 1, e175. https://doi.org/10.7717/peerj.175

Samuelson, P. (2017). Intellectual property and the scientific enterprise. Science, 357(6347), 591–592. https://doi.org/10.1126/science.aao3589

Shabani, M. (2020). Ethical implications of data sharing in biomedical research: Balancing openness and privacy. BMC Medical Ethics, 21(1), Article 50. https://doi.org/10.1186/s12910-020-00489-1

Shahani, M. (2020). Data sharing and protection in health research: The role of GDPR and LGPD. Health Informatics Journal, 26(4), 2823–2832. https://doi.org/10.1177/1460458220914045

Souza, A., et al. (2021). Implicações da LGPD para pesquisa científica: Desafios e perspectivas. Revista de Direito, Estado e Telecomunicações, 13(2), 101–120.

Tennant, J. P., et al. (2016). The academic, economic and societal impacts of open access: An evidence-based review. F1000Research, 5, 632. https://doi.org/10.12688/f1000research.8460.3

Tenopir, C., et al. (2011). Data sharing by scientists: Practices and perceptions. PLoS ONE, 6(6), e21101. https://doi.org/10.1371/journal.pone.0021101

Tikkinen-Piri, C., Rohunen, A., & Markkula, J. (2018). EU General Data Protection Regulation: Changes and implications for personal data collecting companies. Computer Law & Security Review, 34(1), 134–153. https://doi.org/10.1016/j.clsr.2017.05.015

West, J., & Bogers, M. (2014). Leveraging external sources of innovation: A review of research on open innovation. Journal of Product Innovation Management, 31(4), 814–831. https://doi.org/10.1111/jpim.12125

Wilkinson, M. D., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3, Article 160018. https://doi.org/10.1038/sdata.2016.18

World Intellectual Property Organization. (2019). Understanding patents. WIPO.

ETHICAL PROCEDURES

Conflict of interest: nothing to declare. **Funding**: nothing to declare. **Peer review**: Double anonymous peer review.

All content of *e³ – Journal of Economics, Business and Entrepreneurship in the CPLP* is licensed under *Creative Commons*, unless otherwise specified and in content retrieved from other bibliographic sources.