

Desenvolvimento de um Índice de Infração de Patentes com Apoio de Inteligência Artificial Multimodal

Development of a Patent Infringement Index Supported by Multimodal Artificial Intelligence

Alaydes Mikaelle de Moraes Mestra em Informática e Gestão do Conhecimento.
<https://orcid.org/0000-0003-2135-6578> Universidade Nove de Julho (UNINOVE) – Brasil.
alaydes.morais@gmail.com

Cleber Gustavo Dias Doutor em Engenharia Elétrica. Universidade Nove de Julho
<https://orcid.org/0000-0002-4232-2409> (UNINOVE) – Brasil. diascg@uni9.pro.br

RESUMO

A crescente demanda por agilidade e precisão na análise de infrações de patentes impulsiona o uso de soluções baseadas em Inteligência Artificial (IA). Este estudo tem como objetivo desenvolver e validar um Índice de Infração de Patentes (II), indicador que combina similaridade semântica e visual entre patentes e produtos para apoiar a triagem automatizada de possíveis violações. A metodologia integra análises textuais com o modelo PatentBERT e análises imagéticas com o modelo CLIP, comparando o claim da patente com a descrição e imagens de produtos suspeitos. O estudo de caso utilizou duas patentes norte-americanas, resultando em um II de 50,84%, o que indica sobreposição técnica relevante, mas não caracterizou infração direta. Os resultados evidenciam o potencial da IA para acelerar a verificação de patentes, manter a consistência técnica e reduzir vieses interpretativos. Conclui-se que a integração entre algoritmos de IA e interpretação humana contribui para a automação confiável da detecção de infrações no campo da propriedade industrial.

Palavras-chave: inteligência Artificial; patentes; propriedade intelectual; detecção de infrações.

ABSTRACT

The growing demand for agility and accuracy in patent infringement analysis drives the use of Artificial Intelligence (AI) solutions. This study aims to develop and validate a Patent Infringement Index (II), an indicator that combines semantic and visual similarity between patents and products to support the automated screening of potential violations. The methodology integrates textual analysis using the PatentBERT model and image analysis using the CLIP model, comparing the patent claim with the description and images of suspected products. A case study was conducted with two U.S. patents, resulting in an II of 50.84%, which indicates relevant technical overlap but does not constitute direct infringement. The results demonstrate the potential of AI to accelerate patent verification, ensure technical consistency, and reduce interpretative biases. It is concluded that the integration of AI algorithms and human interpretation contributes to the reliable automation of infringement detection in the field of industrial property.

Keywords: artificial intelligence; patents; intellectual property; infringement detection.

Recebido em 07/05/2025. Aprovado em 10/09/2025. Avaliado pelo sistema *double blind peer review*.

Publicado conforme normas da APA.

<https://doi.org/10.22279/navus.v16.2133>

1 INTRODUÇÃO

A propriedade intelectual é crucial na sociedade moderna, pois protege os direitos de inventores, artistas e criadores em diferentes áreas do conhecimento. As proteções de patentes garantem o direito exclusivo por um período definido, promovendo o progresso em tecnologia e ciência. A proteção legal incentiva a inovação recompensando inventores que fomentam a criação de novas tecnologias e conceitos. Devido ao número crescente de inovações e invenções complexas, há uma necessidade de soluções mais eficientes nos processos de análise de patentes (Barbosa, 2023).

Nessa situação, a Inteligência Artificial (IA) se destaca como uma solução promissora para melhorar o processo de concessão e supervisão de patentes. Ferramentas orientadas por IA podem conduzir análises mais rápidas e precisas de conjuntos de dados extensos, reconhecer semelhanças entre patentes e identificar prontamente possíveis violações. Esses sistemas de IA são capazes de automatizar tarefas complexas, como pesquisar patentes anteriores, contrastar informações de patentes e avaliar a exclusividade das invenções. Isso significa um avanço notável em matéria de propriedade intelectual (Garcia, 2020).

Incorporar agilidade e precisão é um aspecto vital quando se trata de modelos de IA que detectam violações de patentes. A detecção de infrações não se resume apenas a simples comparações de palavras-chave ou imagens, pois também envolve compreender as dimensões técnicas e legais das invenções. Redes neurais e modelos de aprendizado profundo têm a capacidade de reconhecer padrões e detalhes intrincados em descrições de patentes e visuais, melhorando, em última análise, a precisão do sistema. No entanto, apesar de sua eficiência, a IA encontra dificuldades para compreender contextos diferenciados que exigem aprimoramento contínuo para melhorar a precisão e a confiabilidade (Berente, 2021).

Além do progresso que a IA facilitou, existem obstáculos significativos que devem ser enfrentados. Entender as circunstâncias técnicas e legais precisas da inovação representa um grande desafio para os sistemas de IA. As patentes frequentemente são formuladas de forma vaga ou contêm termos técnicos complexos, tornando a análise automatizada mais complicada. Consequentemente, a experiência humana é essencial para garantir a precisão das análises conduzidas por sistemas de IA, particularmente em cenários complexos ou pouco claros. Outra preocupação pertinente está relacionada à garantia de transparência e explicabilidade nas decisões tomadas pela IA, especialmente em áreas como a Lei de Propriedade Industrial 9.279/96, onde as decisões podem ter consequências legais substanciais. A falta de certeza nos procedimentos de tomada de decisão pode diminuir a confiança no sistema e levar a complicações ao questionar ou avaliar decisões automatizadas (Sichman, 2021).

A integração da inteligência artificial na avaliação de patentes traz repercussões sociais e econômicas significativas. Ao aumentar a eficiência e reduzir os custos no processo de concessão de patentes, a IA tem o potencial de tornar a proteção de invenções mais acessível, especialmente para pequenos inventores ou aqueles de países em desenvolvimento. Isso pode promover a inclusão e motivar um número maior de indivíduos a buscar proteção para suas inovações. No entanto, a adoção generalizada dessa tecnologia pode beneficiar predominantemente grandes corporações com os recursos necessários para implantar essas soluções avançadas, o que representa um novo desafio em relação à justiça no acesso à tecnologia. A supervisão da utilização da IA deve, portanto, se esforçar para evitar que essas ferramentas reforcem a disparidade existente, mas, em vez disso, promover um sistema equitativo e facilmente acessível para todas as partes interessadas envolvidas (Ertel, 2024).

A qualidade dos dados usados para treinar modelos de IA é crucial para uma análise eficaz de patentes. Dados completos, precisos e representativos de várias áreas tecnológicas são essenciais para garantir a precisão na análise. Dados insuficientes ou desatualizados podem afetar a eficácia do modelo, resultando em análises imprecisas e decisões falhas. Manter dados de treinamento diversos e de alta qualidade é vital para a implementação bem-sucedida da IA em aplicações de propriedade intelectual, ao mesmo tempo em que considera sua importância junto com outros fatores críticos (Huynh-The, 2023).

A colaboração entre humanos e IA é crucial para superar as limitações da tecnologia e garantir uma tomada de decisão justa e bem fundamentada. Enquanto a IA aumenta a eficiência e a precisão, o discernimento humano continua vital para abordar questões subjetivas e ambientes intrincados. Os profissionais de propriedade intelectual desempenham um papel fundamental na verificação dos resultados da IA, interpretando as nuances do caso, garantindo a operação ética e eficiente do sistema. Integrar as capacidades analíticas da IA com a proficiência humana pode aumentar a velocidade, a precisão e a imparcialidade da análise de patentes, reforçando assim a proteção da propriedade intelectual como um todo (Barbosa, 2023).

A inclusão da IA na análise de patentes é um passo notável para aumentar a eficácia e a precisão do sistema de propriedade intelectual. No entanto, a integração dessa tecnologia exige a superação de vários obstáculos, como decifrar os aspectos técnicos e legais das invenções, garantindo processos de tomada de decisão compreensíveis e garantindo a disponibilidade equitativa da proteção de patentes. A fusão de tecnologias de ponta com uma estrutura legal aberta e justa desempenhará um papel crucial na otimização das vantagens da IA e na garantia de proteção equitativa para todos os inventores, independentemente da escala ou do histórico (Garcia, 2020).

O processo de identificação de infrações de patentes é fundamental para garantir a proteção dos direitos de propriedade intelectual. Ele é conduzido por especialistas técnicos e jurídicos que seguem uma abordagem sistemática para avaliar se um produto infringe uma patente existente. Esse fluxo de trabalho, apesar de eficaz, apresenta desafios como tempo elevado de análise, subjetividade na avaliação e dificuldade na comparação manual de informações técnicas e visuais (Park, 2014).

O fluxo segue um modelo sequencial, no qual cada etapa depende do resultado da anterior. O processo inicia-se com o recebimento do caso, no qual um especialista é acionado para analisar uma possível infração de patente. Essa solicitação pode ser feita pelo titular da patente, por advogados especializados ou por órgãos reguladores. Após o recebimento, o especialista conduz uma análise inicial, revisando a patente concedida e as informações preliminares sobre o produto suspeito (CHOI, 2014). Essa revisão busca identificar se há indícios concretos de infração, permitindo que, caso não haja justificativa para prosseguir, o caso seja encerrado nesta fase.

Caso a análise inicial indique a necessidade de investigação mais detalhada, o especialista segue para a captura de imagens e documentação do produto suspeito. Esse registro visual é essencial para permitir a comparação posterior com os diagramas e especificações contidos na patente. São realizadas fotografias em diversos ângulos e, em alguns casos, vídeos para demonstrar aspectos funcionais do produto (Liu, 2020).

A etapa seguinte consiste na identificação e listagem das partes e características do produto, estruturando todos os seus componentes de forma organizada. Esse mapeamento serve como referência para as comparações subsequentes. Com essa base estabelecida, inicia-se a comparação textual, na qual o especialista examina as descrições técnicas contidas na patente e no produto suspeito. O objetivo é identificar termos, conceitos e funcionalidades que possam indicar uma sobreposição de características (Choi, 2014).

Além da análise textual, realiza-se uma comparação visual entre as imagens capturadas do produto e os diagramas da patente. Esse exame detalhado permite identificar similaridades estruturais, funcionais e estéticas. Caso haja forte correspondência visual entre o produto e a patente, isso pode ser um forte indício de infração (MDPI, 2023).

Após a etapa de comparações, o especialista conduz uma verificação jurídica, na qual avalia a conformidade do produto suspeito com as leis de propriedade intelectual. Nesse momento, são analisadas infrações diretas (casos em que o produto replica elementos específicos da patente) e infrações por equivalência (situações em que métodos alternativos resultam na mesma funcionalidade protegida pela patente) (Instituto Dannemann Siemsen de Estudos Jurídicos e Técnicos, 2013).

Com base nessas análises, o especialista elabora um parecer técnico, detalhando todas as evidências e justificativas para sua conclusão. Esse parecer pode ser utilizado para fundamentar ações judiciais, negociações entre as partes ou arquivamento do caso. A última etapa do processo consiste no encaminhamento para decisão legal, momento em que os detentores da patente definem a estratégia a ser adotada, seja a judicialização do caso, a busca por mediação ou outras formas de proteção da propriedade intelectual (Zhu, 2020).

Embora seja um procedimento essencial para a proteção de patentes, o método tradicional apresenta desafios significativos. Primeiramente, há a demanda de tempo elevado, uma vez que a análise manual de textos, diagramas e fotografias torna o processo longo e custoso. Além disso, a subjetividade na interpretação pode levar a diferentes conclusões dependendo do especialista que conduz a análise. Outro fator relevante é a complexidade da análise comparativa, pois a verificação de similaridades entre produtos e patentes exige um conhecimento técnico avançado e pode ser influenciada por pequenas variações na estrutura e funcionalidade (Lee, 2013).

1.1 Objetivos

1.1.1 Objetivo Geral

Desenvolver e validar um Índice de Infração de Patentes com apoio de Inteligência Artificial Multimodal.

1.1.2 Objetivos específicos

Os objetivos específicos do presente estudo estão descritos a seguir:

- i-) Desenvolver um indicador de infração de patentes com o suporte de técnicas de Inteligência Artificial multimodal;
- ii-) Implementar a análise de infração computacional usando dados oriundos de textos e/ou imagens das patentes;
- iii-) Testar o uso do indicador de infração em uma área do conhecimento particular;

Aprimorar a estrutura de patentes com tecnologias como IA pode gerar vantagens econômicas substanciais ao facilitar um processo de alocação de direitos de propriedade mais rápido e preciso. Essa progressão auxilia na elaboração de melhores políticas públicas que garantam proteção imparcial e equitativa, promovendo uma atmosfera propícia à inovação sustentável e promovendo a justiça tecnológica. Além disso, a eficiência e a acessibilidade da IA podem dar suporte à ampliação do acesso

à proteção de patentes, capacitando um número maior de inovadores a receber o reconhecimento e a proteção legítimos para suas invenções.

2 REFERENCIAL TEÓRICO

A utilização de um Índice de Infração (II) na área de propriedade intelectual, especialmente na detecção de infrações de patentes, tem se mostrado uma frente promissora de pesquisa e desenvolvimento.

O estudo de Srinivas et al. (2024) apresenta o PatExpert, um framework autônomo baseado em múltiplos agentes, que visa otimizar e automatizar os fluxos de trabalho relacionados a patentes. A proposta de utilizar agentes especializados em tarefas como classificação, geração de reivindicações e análise de múltiplas patentes representa um avanço significativo na eficiência desses processos. A incorporação de técnicas como Graph Retrieval-Augmented Generation (GRAG), que combina semelhança semântica com grafos de conhecimento, aprimora a precisão e a relevância da análise. Essa abordagem, que integra mecanismos de erro e feedback, não apenas melhora a acurácia, mas também assegura transparência — aspecto essencial na aplicação de IA à propriedade intelectual.

Em outra vertente, Vesala & Ballardini (2019) exploram as implicações jurídicas do uso de IA, com foco nas redes neurais e seus impactos sobre os direitos de propriedade intelectual. O estudo examina em quais condições o treinamento de redes neurais pode configurar violação de propriedade, oferecendo um panorama crítico dos desafios legais da aplicação de IA em tarefas como a detecção de infrações de patentes. Essa análise reforça a necessidade de garantir que os modelos de IA operem dentro dos limites legais, especialmente quando aplicados em contextos sensíveis, como o da propriedade industrial. Diferentemente do presente trabalho, que propõe um indicador técnico para estimar violações, o estudo de Vesala & Ballardini (2019) se concentra nos aspectos normativos, sem apresentar soluções computacionais.

A pesquisa de Shi et al. (2024) introduz o PatentFinder, uma ferramenta voltada à detecção de infrações em patentes relacionadas a moléculas. Com o uso de componentes como MarkushParser e MarkushMatcher, a ferramenta aprimora a análise de pequenas moléculas e supera abordagens baseadas apenas em linguagem natural. Sua performance superior reforça o uso de IA em áreas especializadas como biotecnologia e farmacologia. O diferencial deste trabalho, no entanto, está na criação de um indicador multimodal que pode ser aplicado a diversas áreas tecnológicas, não se restringindo ao domínio químico-farmacêutico.

Complementando a discussão, Dos Santos et al. (2024) investigam o uso de IA generativa na redação de pedidos de patentes. Embora não aborde diretamente a detecção de infrações, o estudo evidencia como a qualidade da redação impacta etapas posteriores do processo. Um pedido de patente bem estruturado facilita a análise comparativa e, por consequência, a detecção de violações. O foco do presente trabalho, no entanto, está na aplicação de um índice baseado em similaridade semântica e visual, enquanto o de Dos Santos et al. concentra-se na fase anterior de geração textual.

A pesquisa de Lu & Ni (2019) apresenta o modelo BERT-CNN para a classificação automatizada de patentes, com acurácia de 84,3%. O uso de modelos de linguagem como o BERT é relevante para representar semanticamente termos técnicos e pode ser extrapolado para tarefas de identificação de similaridades e possíveis infrações. A proposta aqui, contudo, avança além da classificação, ao comparar patentes com produtos reais, utilizando dados textuais e visuais para estimar riscos infracionais.

O trabalho de Yasaei et al. (2021) propõe o GNN4IP, um método baseado em redes neurais gráficas para detectar pirataria de propriedade intelectual em hardware. O modelo alcança 96% de

acurácia ao analisar circuitos eletrônicos por grafos estruturais. Apesar de o foco ser o design de hardware, o conceito de similaridade estrutural pode ser adaptado para outras formas de patente, como o que se propõe neste trabalho, que utiliza uma abordagem multimodal mais ampla e aplicável a diversos setores.

No mesmo campo, o estudo de Cao et al. (2021) introduz o IPGuard, uma técnica de proteção de propriedade intelectual aplicada a redes neurais profundas, sem comprometer sua precisão. O uso de fingerprinting na fronteira de classificação das redes oferece uma nova abordagem para detectar cópias de modelos de IA. Essa proposta pode ser adaptada para fins de verificação em sistemas que analisam a similaridade entre patentes.

Este trabalho, por outro lado, não visa proteger modelos de IA, mas sim identificar automaticamente infrações de patentes reais cometidas por produtos disponíveis no mercado, por meio de um índice que possa ser utilizado por especialistas ou agentes inteligentes para estimar o grau de similaridade e o risco de violação.

A principal diferença em relação aos trabalhos revisados está no escopo e na generalização: enquanto diversas abordagens são restritas a domínios específicos — como hardware, biotecnologia ou geração textual — o presente estudo propõe um indicador multimodal adaptável a diferentes contextos técnicos, combinando análise textual e imagética com o objetivo de oferecer suporte à tomada de decisão na gestão da propriedade industrial.

Além da literatura internacional, estudos brasileiros reforçam a importância da Propriedade Industrial no contexto nacional. No que Barbosa (2025) destaca os fundamentos e desafios da proteção intelectual, Silveira (2018) apresenta um panorama abrangente do sistema de propriedade intelectual. Essas contribuições reforçam a representatividade nacional e enriquecem o debate sobre inovação e segurança jurídica no país.

Garcez Júnior, Eloy e Santos (2021), ainda no contexto brasileiro, realizaram uma investigação empírica inédita sobre a qualidade das patentes concedidas no país, a partir das ações de nulidade protocoladas — indicador claro de contestação técnica à validade das patentes. Esse quadro evidencia a necessidade de um mecanismo preventivo de análise de infrações, como o Índice de Infração (II), capaz de reforçar a confiabilidade e reduzir litígios futuros.

Nesse cenário, a proposta deste trabalho busca justamente preencher essas lacunas, ao desenvolver e validar um índice multimodal adaptável a diferentes contextos tecnológicos, concebido para apoiar de forma prática e transparente a análise de possíveis infrações de patentes.

3 METODOLOGIA

A presente pesquisa adota uma abordagem metodológica mista, integrando revisão bibliográfica, estudo de caso e desenvolvimento de um Índice de Infração (II) voltado à detecção automatizada de infrações de patentes. O objetivo principal é propor, implementar e testar um modelo multimodal capaz de estimar a probabilidade de infração com base em critérios técnicos e semânticos. A metodologia foi estruturada em etapas interdependentes, que abrangem desde o levantamento teórico até a construção do indicador de infração, culminando na análise dos resultados a partir de métricas de desempenho validadas.

A etapa inicial corresponde à revisão bibliográfica, com o intuito de mapear os principais métodos, ferramentas e abordagens adotadas na aplicação de IA à propriedade intelectual. Conforme Gil (2008), a revisão bibliográfica é fundamental para identificar lacunas no conhecimento e consolidar o referencial teórico.

A segunda etapa envolve o desenvolvimento e validação de um Índice de Infração (II), com base em um estudo de caso. Foram selecionadas duas patentes norte-americanas para simulação: a patente US4656605A, intitulada Single In-line Memory Module, concedida em 1987; e a patente US4727513A, intitulada Signal In-line Memory Module, concedida em 1988. Ambas apresentam alta similaridade estrutural e técnica, sendo a segunda utilizada como produto supostamente infrator. A escolha desse conjunto se justifica pela disponibilidade de documentos técnicos detalhados, incluindo reivindicações formais (claims) e representações gráficas das invenções.

O cerne metodológico desta pesquisa é a criação de um Índice de Infração (II), formulado para sintetizar a probabilidade de ocorrência de infração com base na média ponderada de duas comparações:

(i) a similaridade semântica entre a claim da patente supostamente infringida e a descrição do produto;

(ii) a similaridade semântica entre a claim da patente e as imagens do produto.

O valor final do Índice de Infração (II) é dado pela soma das similaridades obtidas, dividida pela quantidade total de comparações realizadas. A fórmula geral pode ser expressa conforme indicado na Equação

$$II = (\text{Similaridade Claim vs Descrição} + \sum \text{Similaridade Claim vs Imagem}) / \text{Número total de comparações} \quad (1)$$

onde:

Similaridade Claim vs Descrição: representa o valor de similaridade entre o texto dos claims da patente e a descrição textual do produto. –

$\sum$ Similaridade Claim vs Imagem: corresponde à soma das similaridades entre o texto dos claims da patente e cada uma das imagens do produto. –

Número total de comparações: equivale à soma de 1 (para a comparação textual) com o número de imagens analisadas.

De forma didática, o Índice de Infração (II) pode ser entendido como a média das similaridades entre o texto dos claims da patente e as informações do produto (descrição e imagens). O valor final, expresso em percentual, indica o grau de sobreposição técnica: valores elevados sugerem maior probabilidade de infração.

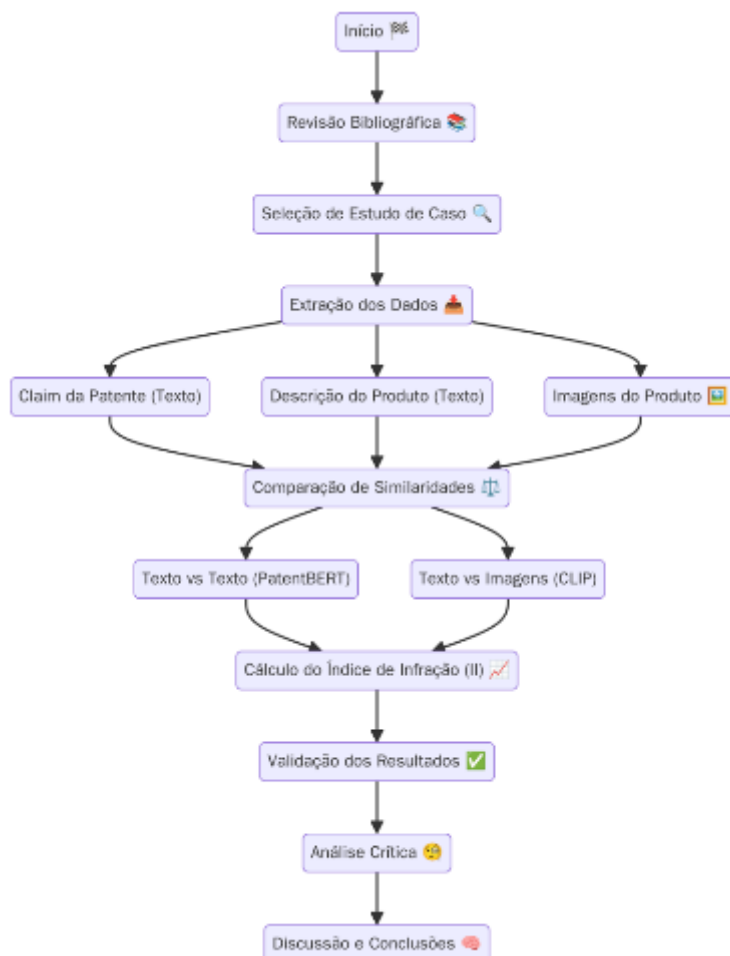
A similaridade semântica é medida por meio de embeddings vetoriais gerados a partir de modelos de linguagem pré-treinados. Para a comparação entre textos (reivindicação e descrição), utilizou-se o modelo PatentBERT (anferico/bert-for-patents); para a comparação entre texto e imagem (reivindicação e imagens do produto), utilizou-se o modelo CLIP (openai/clip-vit-base-patch32). As pontuações resultantes são normalizadas em escala percentual de 0 a 100.

Além desses dois componentes centrais, há possibilidade de incorporar pesos específicos para cada atributo, a fim de ajustar a influência relativa de cada fonte de similaridade e, assim, aprimorar a precisão do índice.

A Figura 1 a seguir apresenta o fluxo de funcionamento que integra todas as etapas operacionais do modelo:

Figura 1

Fluxograma de funcionamento e etapas operacionais



Fonte: Autores (2025).

Na etapa de validação do Índice de Infração (II), os resultados gerados pelo modelo podem ser comparados com informações e pareceres técnicos relativos ao caso selecionado. Essa comparação permite mensurar o desempenho do II com base nas métricas previamente estabelecidas.

Em seguida, realiza-se uma análise crítica dos resultados, discutindo os fatores que impactaram a acurácia do modelo, tais como ambiguidade jurídica, lacunas nas bases de dados e limitações dos modelos multimodais empregados. Essa etapa visa apontar caminhos para melhorias técnicas, ajustes no indicador proposto e possibilidades de generalização para diferentes setores tecnológicos.

Com base nessa estrutura metodológica, a pesquisa formula três hipóteses centrais:

Hipótese H1: Os modelos multimodais de Inteligência Artificial possuem acurácia comparável à de especialistas na detecção de infrações de patentes, sendo capazes de identificar similaridades entre produtos e patentes com precisão técnica adequada.

Hipótese H2: É viável criar um Índice de Infração (II) que sirva de apoio à atuação de especialistas em propriedade intelectual, automatizando, com alto grau de confiabilidade e eficiência, a análise de possíveis violações de patentes.

Hipótese nula H0: Os modelos multimodais atuais de IA não possuem precisão suficiente para complementar a análise de especialistas na identificação de infrações de patentes, apresentando limitações significativas na detecção de similaridades e na interpretação jurídica dos registros.

A partir dessas hipóteses, o presente estudo busca não apenas validar o uso da IA na análise de infrações, mas também contribuir para o desenvolvimento de metodologias mais eficazes no campo da propriedade intelectual, promovendo avanços na automação e na precisão dessas análises.

4 RESULTADOS E DISCUSSÕES

4.1 Comparativo com as soluções de IA para Detecção de Infrações de Patentes

Nesta subseção, apresenta-se uma análise comparativa entre diferentes soluções de Inteligência Artificial atualmente disponíveis no mercado para a verificação de patentes. O objetivo é compreender suas metodologias, limitações e de que forma o Índice de Infração (II) proposto neste estudo pode aprimorar o processo de identificação de violações de propriedade intelectual.

As soluções de IA empregadas na verificação de patentes adotam abordagens distintas para otimizar a identificação de infrações, cada uma com vantagens e limitações específicas. Um exemplo notável é o uso de Processamento de Linguagem Natural (PLN) e técnicas de aprendizado de máquina para comparar automaticamente textos de patentes com bancos de dados técnicos, reduzindo os erros associados à busca manual por palavras-chave (Helmets et al., 2019).

No entanto, como destaca Gollihar (2023), a aplicação da IA em contextos jurídicos, como a detecção de infrações, ainda enfrenta desafios significativos, especialmente quanto à definição de autoria, responsabilidade e confiabilidade das decisões automatizadas. O modelo de IA do INPA, por exemplo, constitui uma iniciativa relevante, porém restrita à triagem de pedidos de patentes, sem contemplar a comparação com produtos existentes.

A plataforma iPNOTE utiliza PLN para buscas textuais avançadas, permitindo localizar patentes similares. Sua limitação, entretanto, está na ausência de análise de imagens — fator crucial para identificar infrações em produtos visuais e processos industriais (iPNOTE, 2024).

Já a NeoPTO, desenvolvida com o apoio da EMBRAPPII, concentra-se na construção de um banco de dados global de patentes, empregando agentes de IA para acelerar a busca por registros relevantes. Apesar de sua utilidade, a ferramenta não realiza uma verificação detalhada de infrações, funcionando apenas como recurso auxiliar para recuperação de informação (EMBRAPPII, 2023).

A parceria entre o INPI e o CAS representa mais um esforço de integração de IA à análise de patentes, organizando dados científicos e permitindo um exame mais ágil dos pedidos. Contudo, o sistema não contempla mecanismos para verificação automática de infrações, o que limita sua aplicabilidade em contextos contenciosos (INPI & MDIC, 2025).

Por sua vez, o PATENTSCOPE, da OMPI, constitui um dos sistemas mais abrangentes para busca de patentes internacionais, mas não dispõe de funcionalidades robustas de análise preditiva ou comparação semântica aprofundada (World Intellectual Property Organization, 2025).

Diferentemente das soluções analisadas, o II proposto neste estudo adota uma abordagem multimodal, integrando análise textual e visual entre o conteúdo técnico das patentes e os produtos reais. Essa característica permite uma verificação mais precisa e automatizada, com potencial para reduzir o tempo de investigação de horas ou dias para apenas alguns minutos — sem perda de profundidade técnica.

4.2 Detecção de Infrações de Patentes em um caso prático

Para validar a eficácia do Índice de Infração (II) proposto, foi conduzido um estudo prático utilizando duas patentes norte-americanas: US4656605A (Single In-line Memory Module, concedida em 1987) e US4727513A (Signal In-line Memory Module, concedida em 1988). A primeira foi considerada como a patente protegida, e a segunda, como um produto potencialmente infrator.

Antes da aplicação do II, foram realizados testes com o objetivo de identificar o modelo de linguagem mais adequado para a verificação de similaridade semântica entre a claim da patente protegida e a descrição do produto infrator simulado, representado pela claim da segunda patente.

Tabela 1

Comparativo de Modelos de IA para Similaridade Texto x Texto

Modelo Utilizado	Tipo de Treinamento	Similaridade Texto x Texto (%)	Observações
TF-IDF + Cosine Similarity	Estatístico (termos)	94,43%	Não capta semântica profunda
MiniLM (BERT generalista)	Semântico geral (leve)	91,98%	Rápido, mas não técnico
SciBERT	Técnico-científico	87,00%	Melhor para textos técnicos, mas sem foco em patentes
PatentBERT	Específico de patentes	95,19%	Melhor desempenho para documentos jurídicos e técnicos

Fonte: Autores (2025).

TF-IDF combinado com Cosine Similarity é uma ferramenta robusta para identificar similaridade textual entre documentos de patente, sendo útil especialmente em etapas iniciais de triagem ou em verificações automáticas. No entanto, seu uso isolado é limitado quando se trata de compreender semântica complexa ou interpretar informações visuais — capacidades essenciais em casos de infrações técnicas (Shen, 2017).

O modelo MiniLM, uma versão compacta do BERT, foi desenvolvido para oferecer eficiência computacional em tarefas de Processamento de Linguagem Natural (PLN). Utilizando a técnica de deep self-attention distillation, o MiniLM reduz significativamente o tamanho do modelo e os custos de processamento, mantendo desempenho comparável ao BERT-base em tarefas como perguntas e respostas e análise de sentimentos (Wang, 2020).

O SciBERT, por sua vez, é um modelo pré-treinado baseado no BERT, voltado para a compreensão de textos científicos. Treinado com um amplo corpus de artigos acadêmicos nas áreas biomédica e de ciência da computação, o SciBERT mostra-se mais adequado do que o BERT padrão para tarefas de PLN envolvendo literatura científica — como classificação de sentenças, extração de entidades e análise semântica de textos técnicos (Beltagy, 2019).

O PatentBERT é uma adaptação do BERT treinada especificamente para tarefas de classificação e análise de patentes. Esse modelo foi desenvolvido com foco na interpretação das reivindicações (claims), demonstrando que essa seção do documento é suficiente para alcançar alto desempenho em tarefas de categorização. O modelo aplica fine-tuning sobre um BERT pré-treinado com um extenso conjunto de dados do USPTO e apresentou resultados superiores a métodos anteriores baseados em redes neurais convolucionais (CNN) e em embeddings tradicionais (Lee, 2019).

Com base nos resultados apresentados na Tabela 1, o modelo PatentBERT (repositório: anferico/bert-for-patents) foi selecionado como o componente textual do II, em razão de sua elevada acurácia e aderência ao vocabulário técnico-jurídico das patentes.

Já o CLIP (Contrastive Language–Image Pre-training), desenvolvido pela OpenAI, é um modelo que associa imagens e textos por meio de aprendizado contrastivo. Treinado em grande escala com pares de imagem e legenda extraídos da web, o CLIP consegue projetar representações linguísticas e visuais em um mesmo espaço vetorial, sendo eficaz para tarefas como busca multimodal, classificação zero-shot, análise de similaridade semântica e interpretação visual de linguagem natural (Radford, 2021).

Para a etapa de comparação visual entre a claim da patente e as imagens do produto infrator simulado, utilizou-se o modelo CLIP. Sua capacidade de representar texto e imagem em um espaço vetorial compartilhado permite realizar a avaliação de similaridade visual de forma automatizada e eficiente. A Tabela 2 apresenta os resultados combinados das comparações semânticas (textuais) e visuais realizadas.

Tabela 2

Similaridade entre claim da patente protegida e elementos do produto suspeito usando PatentBERT e CLIP

Comparação	Tipo de Análise	Similaridade (%)
Claim vs Produto (Texto)	Texto x Texto (PatentBERT)	95,19
Claim vs Imagem 1	Texto x Imagem (CLIP)	100,00
Claim vs Imagem 2	Texto x Imagem (CLIP)	8,18
Claim vs Imagem 3	Texto x Imagem (CLIP)	0,00

Fonte: Autores (2025).

Sendo a claim da patente: "What I claim is: 1. A memory module for installation on a printed circuit motherboard comprising nine data memory chips for storing digital data, each having a data input and output, control input, and an address input, and each being packaged in a plastic leaded chip carrier, wherein said ninth memory chip is for storing detection and correction information associated with the eight data memory chips, an epoxy-glass printed circuit board substrate having a length and width adequate for mounting thereon only in a single row said nine memory chips and for interconnecting the control inputs and the address inputs of the memory chips so that bytes of digital information may be input to or output from the memory chips, the substrate including thirty terminals for providing access to the data inputs and outputs, control inputs, and address inputs of the nine memory chips and to enable reading and writing of information into and out of the nine chips, support means for supporting the memory module at an angle with respect to a motherboard and decoupling capacitors mounted on said substrate and coupled to the memory chips for suppressing transient voltages. 2. The module of claim 1 wherein all nine memory chips are interconnected such that data is input to or output from the ninth memory chips when data is input to or output from the other eight memory chips".

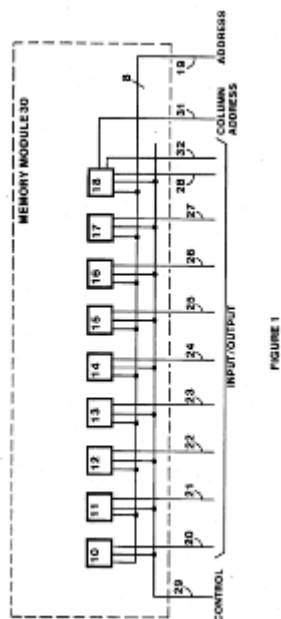
A descrição do produto infrator simulado corresponde à claim da segunda patente: "I claim: 1. A memory module for installation on a printed circuit motherboard comprising: eight data memory chips for storing digital data, each having a data input and output, a control input, and an address input, and each being packaged in a plastic leaded chip carrier; a ninth memory chip for storing error detection and correction information associated with the eight data memory chips, said ninth memory chip having a data input and output, a control input and an address input interconnected with those of the eight memory chips, and a control input to provide writing in or reading out of the ninth memory chip at times

other than when said bytes of digital information are written into or read out of the eight data memory chips to thereby facilitate said error detection and correction operation; an epoxy-glass printed circuit board substrate having a length and width adequate for mounting thereon only in a single row said nine memory chips and for interconnecting the control inputs and the address inputs of the memory chips so that bytes of digital information may be input to or output from the memory chips one at a time; the substrate including thirty terminals for providing access to the data inputs and outputs, control inputs, and address inputs of the nine memory chips to enable reading and writing of bytes of digital information into and out of the eight memory chips and to enable reading and writing of error detection and correction information into and out of the eight memory chips; support means for supporting the memory module at an angle with respect to the printed circuit motherboard when the memory module is installed thereon; and eight decoupling capacitors, mounted on said substrate and connected between the nine memory chips, for suppressing transient voltage spikes between said memory chips."

E as imagens comparadas foram as presentes nas Figura 2, Figura 3 e Figura 4:

Figura 2

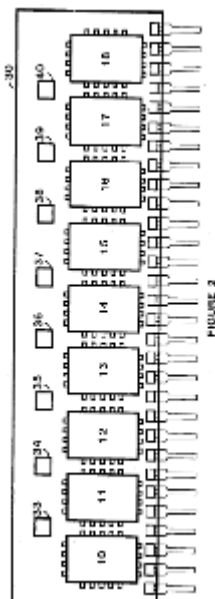
Imagem1 comparada com a Claim da Patente



Nota: Recuperado da Patente US4727513A

Figura 3

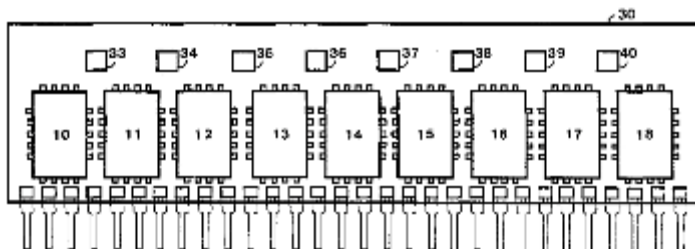
Imagem2 comparada com a Claim da Patente



Nota: Recuperado da Patente US4727513A

Figura 4

Imagem3 comparada com a Claim da Patente



Nota: Recuperado da Patente US4727513A

A média dos valores obtidos resultou em um Índice de Infração (II) de 50,84%, indicando similaridades técnicas relevantes entre os elementos analisados — sobretudo no aspecto textual — sem configurar, no entanto, uma infração direta. Esse resultado é coerente com o fato de que as patentes envolvidas representam soluções técnicas distintas, embora relacionadas.

A patente US4727513A pode ser interpretada como uma evolução funcional ou variação técnica da US4656605A, compartilhando parte de sua arquitetura, mas incorporando modificações estruturais significativas. Assim, o valor intermediário do II demonstra que, embora haja proximidade conceitual, há também evidências suficientes de diferenciação, reforçando o uso do indicador como ferramenta técnica preliminar no apoio à análise de possíveis infrações.

Apesar dos resultados promissores, o Índice de Infração (II) apresenta algumas limitações. Sua precisão depende da qualidade dos dados de entrada — descrições incompletas ou imagens pouco representativas podem comprometer a acurácia — e dos próprios modelos multimodais, que ainda enfrentam dificuldades em capturar nuances jurídico-técnicas complexas. Em contextos práticos, a

aplicabilidade do índice exigiria integração com fluxos periciais e validação junto a especialistas jurídicos, de modo a assegurar sua aceitação como ferramenta de apoio e não como substituto da análise especializada. Além disso, o estudo foi conduzido em um único domínio tecnológico, o que restringe a generalização dos achados. Futuras aplicações em setores diversos poderão ampliar a robustez do II, verificar sua adaptabilidade e contribuir para consolidar seu uso em diferentes contextos industriais.

5 CONCLUSÃO

O presente estudo evidenciou o potencial do uso de um Índice de Infração (II) como ferramenta de suporte técnico na triagem e análise preliminar de possíveis infrações de patentes, por meio da aplicação combinada de Inteligência Artificial textual e visual. Diante da crescente complexidade dos processos de patenteamento e da dificuldade em avaliar sobreposições técnicas entre invenções e produtos reais, a criação de métricas automatizadas torna-se não apenas útil, mas essencial para garantir celeridade e confiabilidade ao processo decisório.

A proposta metodológica apresentada — baseada na combinação de modelos pré-treinados de linguagem natural (como o PatentBERT) e modelos multimodais de similaridade imagem-texto (como o CLIP) — viabilizou a construção de um indicador numérico capaz de sintetizar o grau de similaridade entre o claim de uma patente e os elementos de um produto potencialmente infrator. A aplicação do modelo a um caso prático, envolvendo duas patentes norte-americanas, resultou em um II de 50,84%, valor que indica sobreposição técnica relevante, mas não absoluta. Tal resultado reforça o caráter do índice como instrumento auxiliar à interpretação especializada, e não como substituto da análise jurídica.

A combinação de análises semânticas e visuais demonstrou ganho significativo em abrangência e precisão, evidenciando que modelos contrastivos multimodais são particularmente eficazes na redução de vieses e na priorização de casos com alto grau de similaridade. No entanto, é necessário atentar para questões como a explicabilidade dos modelos, a qualidade das bases de dados utilizadas e os riscos associados à aplicação indiscriminada dessas soluções em larga escala.

Para além dos aspectos técnicos, torna-se fundamental considerar os impactos sociais e econômicos da adoção dessa tecnologia. A disseminação de ferramentas baseadas em IA deve ser acompanhada de políticas públicas e diretrizes que assegurem seu acesso equitativo, evitando que apenas grandes corporações se beneficiem dos avanços. A democratização do uso do II pode ampliar a proteção a inventores independentes e instituições menores, promovendo maior justiça tecnológica.

Como perspectivas futuras, destacam-se: a incorporação de pesos adaptativos no cálculo do II, ajustando a relevância de diferentes componentes (como texto descritivo, imagens ou partes específicas das claims), conforme a natureza da tecnologia analisada; e a validação empírica do índice junto a equipes multidisciplinares de especialistas técnicos e jurídicos, com vistas a testar sua eficácia em outros setores industriais. A aplicação do II em bases de dados reais e diversificadas permitirá refinar sua utilização como instrumento de apoio em auditorias técnicas, disputas legais e exames de patente.

Conclui-se, portanto, que a incorporação do Índice de Infração como recurso analítico automatizado representa um avanço relevante na modernização da gestão de patentes, oferecendo agilidade, precisão e suporte técnico fundamentado. Sua adoção, no entanto, dependerá da integração responsável entre tecnologia e regulação, com supervisão humana contínua e compromisso com a ética e a equidade. Nesse contexto, a Inteligência Artificial não substitui o especialista, mas o potencializa.

REFERÊNCIAS

- Barbosa, D. B. (2025). Tratado da Propriedade Intelectual – Tomo I (2ª ed., 5ª tiragem). Rio de Janeiro: Lumen Juris. ISBN 978-8551900833.
- Barbosa, L. M., & Portes, L. A. F. (2023). A inteligência artificial. *Revista Tecnologia Educacional*, (236), 16–27.
- Beltagy, I., Lo, K., & Cohan, A. (2019). SciBERT: A Pretrained Language Model for Scientific Text. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP)*.
<https://doi.org/10.18653/V1/D19-1371>
- Berente, N., Gu, B., Recker, J., & Santhanam, R. (2021). Managing artificial intelligence. *MIS Quarterly*, 45(3).
- Cao, X., Jia, J., & Gong, N. Z. (2021). IPGuard: Protecting intellectual property of deep neural networks via fingerprinting the classification boundary. *Proceedings of the 2021 ACM Asia Conference on Computer and Communications Security* (pp. 14–25).
- Choi, S. W. (2014). System and method for analysing patent infringement [Patent].
https://scispace.com/papers/system-and-method-for-analysing-patent-infringement-2jil848s2y?utm_source=chatgpt
- Dos Santos, J. W., de Oliveira, A. L., da Silva, R. G., & Pimentel, M. T. (2024). Previsão tecnológica para uso de inteligência artificial generativa na redação de pedidos de patentes. *Revista Eletrônica Científica Inovação e Tecnologia*, 15(37), 41–55.
- Dos Santos, M. J. P., Schaal, F. M. M., & Goulart, R. (2024). Propriedade intelectual e inteligência artificial. *Almedina Brasil*.
- Dos Santos, M. V. R., Alves, G. M. R., & De Andrade Maciel, A. M. (2022). Benchmarking de sistemas AutoML open-source. *Revista de Engenharia e Pesquisa Aplicada*, 7(3), 19–28.
- Ertel, W. (2024). *Introduction to artificial intelligence* (2nd ed.). Springer Nature.
- EMBRAPII. (2023). EMBRAPII apoia sistema de Inteligência Artificial para agilizar processo de patente industrial. <https://embrapii.org.br/embrapii-apoia-sistema-de-inteligencia-artificial-para-agilizar-processo-de-patente-industrial/>
- Garcez Júnior, S. S., Eloy, B. R., & Santos, J. A. B. (2021). A qualidade dos privilégios patentários concedidos no Brasil sob a ótica das ações judiciais de nulidade de patentes. *Revista Direito GV*, 17(1), e2116.
<https://doi.org/10.1590/2317-6172202116>
- Garcia, A. C. B. (2020). Ética e inteligência artificial. *Computação Brasil*, (43), 14–22.
- Gollihar, J. (2023). Patent Infringement by Development and Use of Artificial Intelligence Systems, Specifically Artificial Neural Networks. *MPI Studies on Intellectual Property and Competition Law*.
https://doi.org/10.1007/978-3-662-65974-8_20
- Helmets, L., Horn, F., Biegler, F., Oppermann, T., & Müller, K.-R. (2019). Automating the search for a patent's prior art with a full text similarity search. *PLOS ONE*. <https://doi.org/10.1371/journal.pone.0212103>
- Hu, W., Qin, Y., Jin, H., Han, L., & Qian, Y. (2024). Bliva: A simple multimodal LLM for better handling of text-rich visual questions. *Proceedings of the AAAI Conference on Artificial Intelligence* (pp. 2256–2264).

- Huynh-The, T., Hua, B. S., Ngo, T. D., & Lee, G. (2023). Artificial intelligence for the metaverse: A survey. *Engineering Applications of Artificial Intelligence*, 117, 105581.
- Instituto Dannemann Siemsen de Estudos Jurídicos e Técnicos. (2013). *Comentários à Lei da Propriedade Industrial e correlatos* (3ª ed., rev. e atual.). Rio de Janeiro: Renovar.
- INPI e MDIC. (2025, março 24). INPI e MDIC avançam para agilizar patentes e marcas no Brasil. <https://www.gov.br/inpi/pt-br/central-de-conteudo/noticias/inpi-e-mdic-avancam-para-agilizar-patentes-e-marcas-no-brasil>
- iPNOTE. (2024). AI Patent Search: Transforming the Future of Intellectual Property. <https://ipnote.pro/en/blog/ai-patent-search-transforming-the-future-of-intellectual-property/>
- Lee, C., Song, B., & Park, Y. (2013). How to assess patent infringement risks: A semantic patent claim analysis using dependency relationships. *Technology Analysis & Strategic Management*, 25(1), 23–38.
- Lee, J.-S., & Hsiang, J. (2019). PatentBERT: Patent Classification with Fine-Tuning a Pre-trained BERT Model. *arXiv: Computation and Language*. https://scispace.com/papers/patentbert-patent-classification-with-fine-tuning-a-pre-1bxqiytdzu?utm_source=chatgpt
- Li, Z., Tang, K., Pan, J., Zhang, Y., & Tang, J. (2024). Monkey: Image resolution and text label are important things for large multi-modal models. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 26763–26773).
- Liu, W., Liu, X., Kong, Y., Yang, Z., & Qiao, W. (2020). Game Theory Based Patent Infringement Detection Method. *Database and Expert Systems Applications* (pp. 134–145). Springer. https://doi.org/10.1007/978-3-030-59051-2_11
- Lo, H.-C., Yang, Y., Zhu, Y., & Wu, Y. (2024). Large language model informed patent image retrieval. *arXiv preprint arXiv:2404.19360*.
- Lu, X., & Ni, B. (2019). BERT-CNN: A hierarchical patent classifier based on a pre-trained language model. *arXiv preprint arXiv:1911.06241*.
- Lu, Y., Wang, H., Zhang, T., & Zhao, Z. (2020). Research on classification and similarity of patent citation based on deep learning. *Scientometrics*, 123, 813–839.
- Lv, Z. (2023). Generative artificial intelligence in the metaverse era. *Cognitive Robotics*, 3, 208–217.
- MDPI. (2023). Measuring Patent Similarity Based on Text Mining and Image Information. <https://www.mdpi.com/2079-8954/11/6/294>
- Park, I., & Yoon, B. (2014). A semantic analysis approach for identifying patent infringement based on a product–patent map. *Technology Analysis & Strategic Management*, 26(8), 855–874.
- Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., ... & Sutskever, I. (2021). Learning Transferable Visual Models From Natural Language Supervision. *arXiv:2103.00020*. <https://doi.org/10.48550/arXiv.2103.00020>

- Shen, Y.-C., Lin, G. T. R., Lin, J.-R., & Wang, C.-H. (2017). A Cross-Database Comparison to Discover Potential Product Opportunities Using Text Mining and Cosine Similarity. *Journal of Scientific & Industrial Research*.
- Shi, Y., Yang, Y., Yang, Y., Xie, Y., & Wang, B. (2024). Intelligent system for automated molecular patent infringement assessment. *arXiv preprint arXiv:2412.07819*.
- Sichman, J. S. (2021). Inteligência artificial e sociedade: avanços e riscos. *Estudos Avançados*, 35, 37–50.
- Silveira, N. (2018). Propriedade intelectual: Propriedade industrial, direito de autor, software, cultivares, nome empresarial, título de estabelecimento, abuso de patentes (6ª ed.). São Paulo: Manole. ISBN 978-8520457528.
- Srinivas, S. S., Vaikunth, V. S., & Runkana, V. (2024). Towards automated patent workflows: AI-orchestrated multi-agent framework for intellectual property management and analysis. *arXiv preprint arXiv:2409.19006*.
- Teli, J. S., Rai, A., & Lin, Y.-K. (2024). Abnormal returns to artificial intelligence patent infringement litigations. *Journal of Management Information Systems*, 41(2), 422–452.
- Vesala, J., & Ballardini, R. (2019). AI and IPR infringement: A case study on training and using neural networks. In R. Ballardini & J. Välimäki (Eds.), *Regulating Industrial Internet through IPR, Data Protection and Competition Law*. Kluwer Law International.
- Yasaei, R., Basu, A., Raghunathan, A., & Jha, N. K. (2021). GNN4IP: Graph neural network for hardware intellectual property piracy detection. In *2021 58th ACM/IEEE Design Automation Conference (DAC)* (pp. 217–222). IEEE.
- Wang, W., Wei, F., Dong, L., Bao, H., Yang, N., & Zhou, M. (2020). MiniLM: Deep Self-Attention Distillation for Task-Agnostic Compression of Pre-Trained Transformers. *NeurIPS*. <https://scispace.com/>
- World Intellectual Property Organization. (2025). PATENTSCOPE. <https://www.wipo.int/en/web/patentscope>
- Zhu, D. (2020). Bibliometric analysis of patent infringement retrieval model based on self-organizing map neural network algorithm. *Library Hi Tech*, 38(2), 479-491.