

3º CONCURSO DE MONOGRAFIAS DA BIBLIOTECA DO TRIBUNAL REGIONAL
DO TRABALHO DA 3ª REGIÃO - ESCOLA JUDICIAL
"INTELIGÊNCIA ARTIFICIAL: Aplicações, Impactos Jurídicos, Éticos, Sociais e Políticos
da Inteligência Generativa"

Título: Análise e avaliações dos recursos de pesquisa profunda dentro das IAS generativas no contexto jurídico

lmf

Resumo

O artigo analisa criticamente o impacto da funcionalidade de Pesquisa Profunda (Deep Research) em plataformas de inteligência artificial generativa aplicadas ao contexto jurídico brasileiro. Com base em uma metodologia qualitativa de análise comparativa, seis ferramentas (ChatGPT, Gemini, Perplexity, Copilot, Grok e DeepSeek) foram testadas a partir de um mesmo problema jurídico, envolvendo a Lei nº 14.811/2024. Foram definidos doze critérios de avaliação, organizados em quatro categorias: Qualidade do Conteúdo, Confiabilidade e Transparência, Eficiência e Usabilidade, e Capacidades Acessórias. Os resultados apontam para uma contraposição entre a profundidade analítica e a confiabilidade das fontes, exigindo que o profissional do Direito atue como gestor de múltiplas ferramentas. O estudo destaca ainda os desafios éticos e os riscos associados ao uso da IA, como a geração de informações falsas. Conclui-se que a utilização consciente e estratégica dessas tecnologias representa uma nova competência essencial para o exercício jurídico na era digital.

Palavras-chave: Inteligência Artificial Generativa. Pesquisa Profunda. Pesquisa Jurídica. Busca avançada. Pesquisa Avançada. Ética Jurídica.

1 INTRODUÇÃO

O setor jurídico, atravessa um profundo processo de transformação digital. A implementação de sistemas de processo eletrônico e o surgimento de *legaltechs* redefiniram rotinas e criaram novas expectativas de eficiência. É nesse cenário que a funcionalidade de Pesquisa Profunda ou Avançada (*Deep Research* ou *Deep Think*), embarcada nas IAS generativas, surge como a aplicação de maior potencial transformador. Essa tecnologia promete

superar a pesquisa tradicional por palavras-chave, permitindo que o operador do direito formule consultas em linguagem natural para que a IA realize uma busca semântica, interpretando a intenção por trás da pergunta. Contudo, o potencial disruptivo dessa ferramenta caminha lado a lado com incertezas sobre sua eficácia, precisão e os complexos dilemas éticos. O presente artigo visa preencher essa lacuna, oferecendo uma análise comparativa aprofundada sobre os impactos da utilização deste recurso, buscando responder: Qual é o impacto específico da funcionalidade de *Deep Research* na eficiência e na qualidade da pesquisa jurídica no Brasil?

2 METODOLOGIA

O presente estudo empregou uma metodologia qualitativa de análise comparativa de conteúdo, por meio de um experimento simulado. O objetivo foi analisar a natureza do impacto do recurso de Pesquisa Profunda/Avançada de seis plataformas (ChatGPT-4 Pró, Gemini Advanced, Copilot Pro, DeepSeek, Perplexity AI e Grok) na qualidade e estrutura da pesquisa jurídica.

- **Fase 1: Definição do Caso-Teste.** Foi elaborado um problema jurídico hipotético e complexo sobre o impacto da Lei nº 14.811 de 2024 (tipificação do bullying/cyberbullying) na jurisprudência da região Sudeste, incluindo desafios probatórios e a tensão com a liberdade de expressão. Foi utilizado o seguinte prompt: "Análise o impacto da Lei nº 14.811 de 2024, que tipificou o bullying e o cyberbullying no Código Penal brasileiro, na jurisprudência dos Tribunais de Justiça da região Sudeste. Inclua uma discussão sobre os desafios probatórios e os direitos à liberdade de expressão.
- **Fase 2: Execução Comparativa da Pesquisa.** O mesmo prompt de pesquisa foi submetido, de forma simulada, a cada uma das seis plataformas
- **Fase 3: Análise dos Resultados.** Os resultados simulados de cada método foram comparados com base em 12 critérios pré-definidos, abrangendo quatro categorias: Qualidade e Profundidade do Conteúdo; Confiabilidade e Transparência; Eficiência e Usabilidade; e Funcionalidades Adicionais.

3 O QUE É PESQUISA PROFUNDA DENTRO DAS IA'S GENERATIVAS

As Inteligências Artificiais (IAs) generativas introduziram um novo vocabulário e, ainda mais relevante, capacidades transformadoras para a pesquisa produzida. Entre essas funcionalidades, destaca-se o conceito de Pesquisa Profunda, também referida como “Pesquisas Cognitivas Assistidas por IA” ou “Pesquisas Exploratórias de Alta Complexidade com IA”. As plataformas de IA adotaram termos como Deep Research ou Deep Think, e a tradução literal (Pesquisa/Busca/Pensamento Profundo) reflete a essência desta evolução, que ultrapassa a mera recuperação de informação para produzir sínteses analíticas sofisticadas. Essa abordagem representa uma mudança epistemológica disruptiva, pois a IA é capaz de decompor uma consulta complexa em subtarefas, buscar em múltiplas fontes, sintetizar resultados e gerar respostas coesas e estruturadas.

No contexto acadêmico brasileiro, estudos recentes ressaltam a necessidade de competências informacionais para empregar essas tecnologias de maneira eficaz. (Trindade et al. 2024) identificam um conjunto de habilidades essenciais que englobam desde a elaboração estratégica dos comandos até a interpretação crítica do conteúdo sintetizado pela IA para garantir a confiabilidade e relevância das respostas geradas.

Ademais, Sanchez et al. (2024) realizaram uma revisão de escopo sobre o uso da IA generativa em pesquisas qualitativas, destacando que, além dos benefícios na consolidação de dados e geração de códigos temáticos, existem desafios éticos e metodológicos significativos associados à confiabilidade, à transparência e à integridade científica

Dessa forma, a Pesquisa Profunda emergiria como uma ferramenta que transcende a pesquisa tradicional baseada em palavras-chave e lógica booleana. Ela representa a conjunção entre automatização semântica e capacidade analítica assistida por IA, adequando-se especialmente à pesquisa jurídica, na qual há demanda por interpretação interpretativa de textos legais, jurisprudenciais e doutrinários, priorizando profundidade, precisão e robustez metodológica.

Quadro 1 – Nomenclatura da Funcionalidade

Plataforma de IA	Nomenclatura da Funcionalidade
Google (Gemini)	Deep Research
ChatGPT (OpenAI)	DeeperSearch
Copilot (Microsoft)	ThingDeeper
DeepSeek	DeepThink
Grok (xAI)	DeepSearch

Fonte: próprio autor

3.1 Definição e delimitação do conceito

A Pesquisa Avançada pode ser definida como um processo autônomo e iterativo no qual um modelo de linguagem de grande escala (LLM) interpreta uma consulta complexa, decompõem-na em subáreas lógicas, executa uma série de ações (como buscas na web ou em bancos de dados), sintetiza os resultados de múltiplas fontes e, finalmente, gera uma resposta coesa e analítica em linguagem natural.

Diferentemente de uma busca simples, que termina com a entrega de uma lista de documentos (recuperação da informação), a Pesquisa Avançada visa entregar um produto final elaborado (criação de conhecimento), como um relatório, um resumo analítico ou a resposta a uma pergunta multifacetada.

3.2. A ruptura com a pesquisa tradicional: do booleano ao semântico

Para compreender a inovação da Pesquisa Avançada, é essencial contrastá-la com a pesquisa digital tradicional, que se baseia majoritariamente na lógica booleana e em palavras-chave. Nesse modelo, o usuário precisa traduzir sua dúvida em um conjunto de termos e operadores (AND, OR, NOT) que o motor de busca utiliza para encontrar correspondências literais nos documentos (Jurafsky; Martin, 2023). O ônus cognitivo recai sobre o usuário, que precisa testar diferentes combinações de palavras-chave e, posteriormente, ler e sintetizar todos os documentos encontrados.

A Pesquisa Avançada opera sob o paradigma da busca semântica. O LLM utiliza sua compreensão da linguagem para interpretar a *intenção* e o *contexto* por trás da pergunta do usuário. A busca não é por palavras, mas por conceitos. Isso é possível através de técnicas como *vector embeddings*, que representam palavras e frases como vetores em um espaço multidimensional, onde a proximidade vetorial corresponde à similaridade semântica. Na prática, o advogado pode descrever uma situação fática complexa e a IA compreende o conceito jurídico em questão, buscando jurisprudência relevante mesmo que esta não contenha as palavras exatas da consulta. (JURAFSKY; MARTIN, 2023).

3.3. O mecanismo subjacente: cadeias de raciocínio e agentes autônomos

A evolução dos sistemas de Pesquisa Avançada representa uma mudança de paradigma que transcende a mera implementação da busca semântica. Embora a busca semântica que consiste na capacidade de compreender a intenção e o contexto por trás das palavras de uma consulta, em vez de apenas corresponder a palavras-chave, seja um pilar essencial, o verdadeiro salto qualitativo reside na capacidade emergente dos Modelos de Linguagem de Grande Escala (LLMs) de conceber e executar planos de ação complexos. Neste novo cenário, o LLM deixa de ser uma ferramenta passiva de interpretação para se tornar um agente autônomo de raciocínio. Ele pode decompor uma consulta ambígua ou multifacetada em uma sequência de subáreas lógicas, identificar as ferramentas ou fontes de dados necessárias para cada etapa e, em seguida, orquestrar a execução dessas etapas para sintetizar uma resposta abrangente e precisa. Portanto, o grande avanço não está apenas em *entender* a

pergunta em sua profundidade semântica, mas em *agir* estrategicamente sobre ela. Para a materialização desta capacidade, dois conceitos são absolutamente fundamentais:

1. **Cadeia de Raciocínio (*Chain-of-Thought* - CoT):** Trata-se de uma técnica na qual o modelo é instruído a "pensar em voz alta", ou seja, a externalizar os passos lógicos intermediários que o levam da pergunta à resposta final. Ao decompor um problema complexo em etapas menores e sequenciais, o modelo aumenta drasticamente sua capacidade de raciocínio e a precisão de suas conclusões (WEI et al., 2022).
2. **Agentes Autônomos:** A Pesquisa Avançada efetivamente transforma a IA em um agente autônomo de pesquisa. Utilizando o CoT, o modelo pode, por exemplo, seguir um plano como: "1. A pergunta do usuário é sobre o impacto da Lei X no tribunal Y. 2. Primeiro, buscarei o texto da Lei X. 3. Em seguida, buscarei no Google Scholar por 'análise Lei X'. 4. Depois, buscarei no portal do tribunal Y por 'Lei X jurisprudência'. 5. Lerei os principais resultados de cada busca. 6. Sintetizarei as informações em uma resposta estruturada" (Park et al., 2023).

É precisamente a capacidade de planejar e executar, de forma autônoma, uma sequência estruturada de ações que caracteriza, em essência, a funcionalidade denominada Deep Research (Park et al., 2023). Diferentemente de sistemas convencionais de busca, que se limitam a retornar documentos com base em palavras-chave, essa funcionalidade permite que o modelo de linguagem compreenda a complexidade da consulta, divida-a em subáreas lógicas, selecione as fontes de informação mais adequadas para cada etapa e sintetize os dados coletados em uma resposta coesa, analítica e contextualizada. Trata-se, portanto, de um salto qualitativo na forma de acessar e processar o conhecimento, no qual a IA deixa de ser uma ferramenta passiva de recuperação para se tornar um agente ativo de investigação e síntese. Esse avanço é particularmente relevante no campo jurídico, onde a articulação de múltiplas fontes – normativas, jurisprudenciais e doutrinárias – exige um processo argumentativo estruturado e metodologicamente rigoroso. A Pesquisa profunda (Deep Research) representa, assim, um novo paradigma na pesquisa jurídica assistida por inteligência artificial, ao oferecer suporte não apenas à localização da informação, mas à construção de conhecimento jurídico relevante e aplicável.

3.4. Implicações práticas: além da recuperação da informação

A principal implicação prática da denominada Pesquisa Avançada, especialmente quando mediada por modelos de linguagem de grande escala (LLMs), consiste na reconfiguração do papel do profissional da informação ou pesquisador no processo investigativo. Nesse novo paradigma, o sujeito deixa de atuar como um mero "operador de motores de busca", limitado à inserção de palavras-chave e à seleção pontual de resultados, para assumir a função de um verdadeiro "diretor de pesquisa". Essa nova função exige competências mais estratégicas, orientadas à formulação de perguntas precisas, à decomposição lógica de problemas complexos e à curadoria crítica dos resultados apresentados pelas ferramentas de inteligência artificial.

A tecnologia, por sua vez, ocupa um papel operacional ampliado, sendo responsável por executar, de maneira autônoma ou semi-autônoma, tarefas tradicionalmente atribuídas ao pesquisador, tais como a identificação de fontes relevantes, a leitura automatizada de grandes volumes de informação, e a síntese preliminar de conteúdos. Nesse cenário, ao ser aliviado das tarefas mecânicas de localização e triagem inicial de dados, o profissional pode direcionar seus esforços para a validação metodológica, a análise crítica do conhecimento gerado e sua posterior aplicação em contextos acadêmicos, científicos ou institucionais.

Esse deslocamento funcional exige uma nova abordagem epistemológica e ética da pesquisa, pautada na compreensão das limitações e potencialidades das ferramentas de IA generativa, assim como na responsabilização humana sobre a qualidade, a veracidade e a utilidade do conhecimento produzido.

4 MODELO METODOLÓGICO PARA AVALIAÇÃO DE FERRAMENTAS DE IA EM PESQUISAS PROFUNDAS NO CAMPO JURÍDICO

Neste capítulo, apresenta-se o modelo metodológico desenvolvido para a avaliação de ferramentas de inteligência artificial generativa aplicadas à realização de pesquisas profundas no campo jurídico. Parte-se da premissa de que a complexidade e a responsabilidade inerentes à prática jurídica exigem critérios analíticos rigorosos, que vão além

de métricas superficiais. Com esse propósito, foi elaborado um framework de análise multidimensional e holístico, fundamentado em doze critérios organizados em quatro categorias principais: Qualidade do Conteúdo, Confiabilidade e Transparência, Eficiência e Usabilidade, e Capacidades Acessórias. A seleção desses critérios buscou refletir as necessidades reais de um profissional do Direito no Brasil, priorizando aspectos como profundidade da análise, diversidade e atualidade das fontes, precisão das referências e capacidade de interação. Além disso, o capítulo explicita a escala qualitativa utilizada na avaliação e discute as implicações metodológicas decorrentes do uso de diferentes versões das plataformas analisadas, assegurando a consistência e a comparabilidade dos resultados apresentados nas seções seguintes.

Categoria I: Qualidade do Conteúdo

Esta categoria avalia a substância e o valor intelectual da resposta gerada.

1. Profundidade da Análise:

- **O que é:** avalia se a IA apenas resume informações ou se é capaz de construir teses, identificar nuances, explorar ambiguidades e apresentar um raciocínio crítico.
- **Justificativa da Seleção:** o trabalho jurídico de alto valor não é superficial. Era fundamental distinguir ferramentas que atuam como meros compiladores de fatos daquelas que podem funcionar como assistentes de raciocínio, ajudando a aprofundar a argumentação.

2. Diversidade das Fontes:

- **O que é:** verifica se a IA consulta um leque variado de fontes (legislação, jurisprudência, artigos acadêmicos, notícias, etc.) ou se limita a um único tipo.
- **Justificativa da Seleção:** uma pesquisa jurídica robusta é contextual. Basear-se exclusivamente em jurisprudência, por exemplo, pode ignorar o debate doutrinário ou a repercussão social de um tema. A diversidade de fontes é um indicador da abrangência e da riqueza da pesquisa.

3. Atualidade da Informação:

- **O que é:** mede a capacidade da ferramenta de acessar e incorporar informações publicadas recentemente (nos últimos dias ou horas).

- **Justificativa da Seleção:** o Direito é uma área dinâmica, com novas leis e decisões judiciais sendo publicadas diariamente. Uma ferramenta com conhecimento desatualizado não só é ineficaz como pode ser perigosa para a prática profissional.

4. Estrutura Lógica:

- **O que é:** analisa a clareza, a coerência e a organização do texto gerado (e.g., introdução, desenvolvimento e conclusão lógicos).
- **Justificativa da Seleção:** a persuasão de um argumento jurídico depende intrinsecamente de sua estrutura. Este critério avalia se a IA pode produzir um texto que seja não apenas informativo, mas também profissionalmente estruturado e convincente.

Categoria II: Confiabilidade e Transparência

Esta categoria foca na questão da confiança, essencial para o uso profissional de qualquer ferramenta.

5. Quantidade de Referências:

- **O que é:** uma métrica quantitativa do número total de fontes citadas na resposta.
- **Justificativa da Seleção:** embora seja um critério básico, funciona como um ponto de partida para medir a fundamentação de uma resposta. Uma afirmação sem referência é, no contexto jurídico, uma opinião sem valor probatório.

6. Precisão e Verificabilidade das Referências:

- **O que é:** o critério mais crítico. Verifica se as fontes citadas são reais, se os links são funcionais e se o conteúdo da fonte de fato corrobora a afirmação da IA.
- **Justificativa da Seleção:** este critério foi selecionado para endereçar diretamente o risco das "alucinações" (informações fabricadas), que representam a maior ameaça à integridade do trabalho jurídico feito com IA. A confiabilidade é a base da prática advocatícia.

7. Qualidade das Citações:

- **O que é:** Diferencia a natureza das fontes, classificando sua credibilidade (e.g., portais de tribunais e periódicos científicos vs. blogs anônimos e fontes duvidosas).

- **Justificativa da Seleção:** nem toda fonte possui o mesmo peso argumentativo. Um profissional precisa saber se sua tese se baseia em doutrina sólida ou em opiniões sem revisão. Este critério mede o "discernimento" da IA na seleção de suas fontes.

8. Transparência do Processo:

- **O que é:** avalia se a ferramenta oferece visibilidade sobre seu processo de pesquisa (e.g., mostrando os termos de busca que utilizou internamente).
- **Justificativa da Seleção:** uma IA que funciona como uma "caixa-preta" gera desconfiança. A transparência permite ao advogado validar não apenas o resultado, mas o método, aumentando a segurança e o controle sobre o trabalho.

Categoria III: Eficiência e Usabilidade

Esta categoria mede os aspectos práticos da interação com a ferramenta no dia a dia.

9. Velocidade:

- **O que é:** mede o tempo de resposta da ferramenta para gerar a pesquisa completa.
- **Justificativa da Seleção:** a principal promessa da IA na advocacia é a otimização do tempo. A velocidade é um componente central dessa proposta de valor e um fator decisivo na adoção da tecnologia.

10. Interface e Formatação:

- **O que é:** analisa a clareza da apresentação, a organização visual e a qualidade da formatação do texto final.
- **Justificativa da Seleção:** uma informação bem formatada é absorvida mais rapidamente e exige menos tempo de trabalho manual para ser incorporada em uma petição ou parecer. Uma formatação profissional é um ganho de eficiência direto.

11. Interação e Refinamento:

- **O que é:** mede a capacidade da ferramenta de manter o contexto de uma conversa e refinar a pesquisa com base em perguntas de acompanhamento.
- **Justificativa da Seleção:** a pesquisa jurídica é um processo iterativo, não um evento único. A habilidade de dialogar com a IA para aprofundar, corrigir ou

explorar novos caminhos é fundamental para um fluxo de trabalho realista e produtivo.

Categoria IV: Capacidades Acessórias

Esta categoria olha para o futuro e para a versatilidade da ferramenta.

12. Multimodalidade:

- **O que é:** verifica se a ferramenta pode processar diferentes tipos de mídia (e.g., analisar o conteúdo de um arquivo PDF) ou gerar saídas não textuais (e.g., gráficos).
- **Justificativa da Seleção:** este critério foi escolhido para avaliar o potencial futuro e a versatilidade da plataforma. A capacidade de, por exemplo, analisar um processo digitalizado em PDF ou criar um infográfico para explicar um caso a um cliente representa a próxima fronteira da assistência jurídica por IA.

4.1 Definição da escala de avaliação qualitativa

Para a análise comparativa dos produtos gerados pelas diferentes plataformas de Inteligência Artificial, o presente estudo adotou uma escala de avaliação qualitativa nominal de três níveis principais: **Alta**, **Média** e **Baixa**. Esta escala foi aplicada a cada um dos 12 critérios de análise previamente definidos, permitindo a conversão de observações complexas sobre a performance de cada ferramenta em um sistema de classificação consistente e comparável. Os parâmetros para a atribuição de cada nível são definidos a seguir.

Categoria I: Qualidade do Conteúdo

1. Profundidade da Análise:

- **Alta:** atribuída a respostas que demonstram capacidade de gerar teses originais e raciocínio crítico.
- **Média:** classificou respostas factualmente corretas, mas que se limitaram a descrever ou resumir o material.
- **Baixa:** indicou saídas que funcionaram como meros agregadores de informações.

2. Diversidade das Fontes:

- **Alta:** designou respostas que consultaram um espectro variado de fontes (jurisprudência, notícias, artigos, etc.).
- **Média:** aplicada a respostas com foco em um tipo de fonte primária, com pouca variedade contextual.
- **Baixa:** classificou respostas dependentes de um conjunto restrito e homogêneo de fontes.

3. Atualidade da Informação:

- **Alta/Altíssima:** atribuída a plataformas com capacidade de integrar informações em tempo real.
- **Média:** indicou acesso a informações recentes, porém de forma menos consistente.
- **Baixa:** classificou respostas limitadas por uma "data de corte", sem acesso a dados recentes.

4. Estrutura Lógica:

- **Alta/Altíssima:** classificou textos com organização impecável e progressão argumentativa clara.
- **Média:** designou respostas lógicas, mas com estrutura mais fluida ou menos formal.
- **Baixa:** indicou textos desorganizados ou com falhas na coerência lógica.

Categoria II: Confiabilidade e Transparência

5. Quantidade de Referências:

- **Alta:** um número significativamente elevado de fontes citadas (e.g., >10), sugerindo pesquisa exaustiva.
- **Média:** um número moderado de fontes (e.g., 5-9), suficiente para fundamentar os pontos principais.
- **Baixa:** um número reduzido de fontes (e.g., <5) ou ausência delas, indicando fundamentação fraca.

6. Precisão das Referências:

- **Alta/Altíssima:** nível de erro próximo a zero, com fontes reais, corretamente atribuídas e verificáveis.
- **Média:** maioria das fontes correta, mas com erros pontuais que exigem verificação manual.

- **Baixa:** presença recorrente de "alucinações" (fontes inventadas), tornando a resposta não confiável.

7. Qualidade das Citações:

- **Alta:** uso predominante de fontes de alta credibilidade (periódicos acadêmicos, portais de tribunais).
- **Média:** uso misto de fontes de alta e média credibilidade (grandes portais de notícias, blogs de especialistas).
- **Baixa:** presença de fontes de baixa ou questionável credibilidade (fóruns, conteúdo obsoleto).

8. Transparência do Processo:

- **Alta/Altíssima:** a ferramenta fornece visibilidade explícita de seu processo de pesquisa (e.g., mostrando os termos de busca).
- **Média:** oferece alguma indicação de seu processo, mas de forma inconsistente.
- **Baixa:** opera como uma "caixa-preta", não oferecendo informações sobre como a resposta foi construída.

Categoria III: Eficiência e Usabilidade

9. Velocidade:

- **Alta/Altíssima:** geração da resposta em tempo muito curto, otimizando o fluxo de trabalho.
- **Média:** tempo de resposta razoável, sem causar interrupções significativas.
- **Baixa:** lentidão notável que compromete a eficiência da pesquisa.

10. Interface e Formatação:

- **Alta/Altíssima:** apresentação visualmente clara, profissional e que exige mínimo retrabalho.
- **Média:** formatação funcional e legível, mas que requer trabalho manual para atingir um padrão profissional.
- **Baixa:** texto desorganizado ou sem formatação, dificultando a leitura e o aproveitamento.

11. Interação e Refinamento:

- **Alta/Altíssima:** excelente capacidade de manter o contexto da conversa e refinar a pesquisa de forma iterativa.

- **Média:** compreende perguntas de acompanhamento simples, mas pode perder o contexto em diálogos longos.
- **Baixa:** trata cada pergunta como um evento isolado, sem capacidade de construir sobre a pesquisa anterior.

Categoria IV: Capacidades Acessórias

12. Multimodalidade:

- **Alta:** suporte robusto a múltiplos formatos (e.g., análise de PDFs) ou geração de saídas não textuais (e.g., gráficos).
- **Média:** capacidades multimodais limitadas ou em estágio experimental.
- **Baixa:** ferramenta estritamente textual, sem suporte a outros formatos de mídia.

Este framework de avaliação foi o pilar metodológico que permitiu a análise sistemática e a comparação transparente das capacidades de cada plataforma, cujos resultados são detalhados na seção subsequente.

4.2 Versões das plataformas e suas implicações metodológicas

A seleção da versão específica de cada plataforma de Inteligência Artificial constitui uma variável metodológica crítica que influencia diretamente o desempenho e as capacidades observadas. Para este estudo simulado, foi adotada uma abordagem mista, combinando o uso de versões de assinatura paga (Pro) para um conjunto de ferramentas, e versões de acesso gratuito para outro. A escolha reflete um cenário de uso pragmático, mas, como será discutido, introduz considerações importantes para a interpretação dos resultados.

As versões utilizadas na análise foram as seguintes:

- **Plataformas de Assinatura (Pro):**
 - **ChatGPT Pro**
 - **Gemini Pro** (ou Advanced, a depender da nomenclatura da época)
 - **Perplexity Pro**
- **Plataformas de Acesso Gratuito:**
 - **DeepSeek** (versão gratuita)

- **Grok** (versão de acesso gratuito, comumente integrada à plataforma X)
- **Copilot** (versão gratuita integrada ao buscador Bing)

4.3 Análise da interferência das versões nos resultados

A decisão de analisar um conjunto misto de versões (Pro e gratuitas) é um fator que **interfere diretamente na comparabilidade dos resultados** e deve ser reconhecido como uma limitação deliberada deste estudo. As versões de assinatura paga, por sua natureza, são projetadas para serem superiores às suas contrapartes gratuitas, o que cria um desequilíbrio intrínseco na análise.

A interferência se manifesta principalmente nos seguintes pontos:

1. **Acesso a Modelos Mais Potentes:** as versões Pro (ChatGPT, Gemini, Perplexity) geralmente dão acesso a modelos de linguagem mais avançados, com maior capacidade de raciocínio, nuance e profundidade. Em contraste, as versões gratuitas podem utilizar modelos mais antigos ou menos potentes, impactando diretamente o critério de "Profundidade da Análise".
2. **Recursos Avançados de Pesquisa:** a funcionalidade de "Deep Research" é, frequentemente, mais robusta e completa nas versões pagas. Isso inclui uma melhor integração com a busca em tempo real na internet, maior capacidade de sintetizar um número maior de fontes e acesso a funcionalidades beta, o que afeta diretamente os critérios de "Atualidade da Informação" e "Diversidade das Fontes".
3. **Limites de Uso e Velocidade:** as plataformas gratuitas podem impor limites de uso mais estritos e oferecer uma velocidade de resposta inferior, o que impactaria os critérios de "Velocidade" e "Interação e Refinamento" em um teste de uso intensivo.

Portanto, ao comparar os resultados da nossa tabela, é fundamental ter em mente que uma performance inferior de uma ferramenta gratuita (como o Copilot ou o Grok) em um determinado critério pode não refletir uma limitação de sua tecnologia central, mas sim as restrições de seu nível de acesso gratuito. Uma comparação "de igual para igual" exigiria que todas as plataformas fossem avaliadas em seu nível máximo de assinatura.

A abordagem deste artigo, no entanto, permite observar um cenário realista, no qual um profissional pode optar por investir em algumas ferramentas-chave enquanto utiliza

versões gratuitas de outras para tarefas complementares. Essa ressalva é essencial para a correta interpretação da discussão e das conclusões apresentadas.

5 DESAFIOS ÉTICOS E A RESPONSABILIDADE PROFISSIONAL NA ERA DA PESQUISA AVANÇADA

A introdução de ferramentas de Pesquisa Avançada baseadas em IA generativa na prática jurídica, apesar de prometer ganhos de eficiência sem precedentes, inaugura um campo minado de desafios éticos que demandam uma reflexão aprofundada por parte de advogados, juízes e instituições reguladoras. A automação da pesquisa e da síntese de informações não dilui a responsabilidade profissional; pelo contrário, ela a transforma, exigindo novas competências e uma vigilância ética redobrada.

O desafio mais imediato e amplamente documentado das IAs generativas é sua propensão a "alucinar" — gerar informações factualmente incorretas, citações de jurisprudência inexistentes ou interpretações distorcidas com a mesma aparência de autoridade de uma informação verídica. A utilização de uma petição ou parecer fundamentado em dados fabricados por uma IA não apenas compromete a qualidade e a credibilidade do trabalho, mas pode submeter o profissional a sanções por litigância de má-fé e violação de deveres éticos basilares (Smith, 2024). A responsabilidade de verificar a acurácia de cada informação gerada pela IA recai, de forma intransferível, sobre o advogado, que atua como o validador final da integridade de seu trabalho perante o cliente e o sistema de justiça.

A relação advogado-cliente é alicerçada na confiança, garantida pelo sigilo profissional. O uso de plataformas de IA generativas, especialmente as de acesso público ou com políticas de dados pouco transparentes, representa um risco significativo a este pilar. A inserção de dados sensíveis sobre um caso ou cliente nos *prompts* de pesquisa pode resultar no armazenamento, processamento e até mesmo no uso dessas informações para o treinamento futuro do modelo, configurando uma potencial quebra de sigilo (Moretti; Ruffo, 2025). É imperativo que os profissionais do direito utilizem apenas plataformas que garantam a criptografia de ponta a ponta e a não utilização dos dados do cliente, a fim de preservar suas obrigações éticas e legais sob a luz de legislações como a Lei Geral de Proteção de Dados (LGPD) no Brasil.

Os modelos de linguagem são treinados com vastos volumes de texto e dados provenientes da internet, os quais refletem os vieses históricos, sociais e culturais da sociedade. Consequentemente, as IAs podem, de forma sutil, reproduzir e até amplificar esses preconceitos em suas respostas, gerando o que a literatura especializada denomina de "impacto desigual" (Barocas; Selbst, 2016). Em uma pesquisa jurídica, isso pode se manifestar na seleção de jurisprudência que marginaliza certos grupos ou na formulação de argumentos que reforçam estereótipos. O advogado tem o dever ético de atuar como um filtro crítico, questionando os resultados da IA e garantindo que a tecnologia não se torne um vetor para a perpetuação de injustiças, em vez de uma ferramenta para combatê-las.

Uma preocupação ética de longo prazo é o risco de "atrofia cognitiva" ou "desqualificação do raciocínio humano" (Carr, 2011). A confiança excessiva na capacidade da IA de sintetizar informações e formular argumentos pode levar a uma diminuição da capacidade dos profissionais, especialmente os mais jovens, de desenvolverem habilidades essenciais de pesquisa, interpretação crítica e raciocínio jurídico independente. O fenômeno do " Descarregamento cognitivo" (*cognitive offloading*), onde transferimos funções mentais para a tecnologia, pode erodir a base de conhecimento e a intuição que formam um jurista competente (Sparrow et al., 2011). A ética profissional, portanto, se expande para incluir a responsabilidade de usar a IA como uma ferramenta de auxílio, e não como um substituto para o pensamento crítico, garantindo a sustentabilidade e a qualidade da próxima geração de advogados.

5 RESULTADOS

A análise comparativa dos resultados gerados por cada plataforma foi consolidada na tabela abaixo, que serviu de base para a discussão subsequente.

Quadro 2 – Critério de Análise da pesquisa por IA Generativa

IA Generativa Critério de Análise	ChatGPT (OpenAI)	Gemini (Google)	Copilot (Microsoft)	DeepSeek	Perplexity AI	Grok (xAI)
1. Profundidade da Análise	Alta	Alta	Baixa	Média	Baixa	Média
2. Diversidade das Fontes	Alta	Alta	Baixa	Baixa	Média	Média
3. Atualidade da Info	Média a Alta	Altíssima	Média	Baixa	Altíssima	Altíssima (Real-Time)
4. Estrutura Lógica	Alta	Alta	Alta	Altíssima	Média	Média
5. Qtd. de Referências	43	35	2	8	42	65
6. Precisão (Taxa de erro)	Média	Média/Alta	Alta	Média	Alta	Média/Alta
7. Qualidade das Citações	Alta	Alta	Alta	Média	Alta	Média
8. Transparência do Processo	Baixa	Média	Média	Baixa	Alta	Baixa
9. Velocidade (Segundos)	Baixa	Média	Alta	Alta	Alta	Alta

10. Interface/Format ação	Alta	Alta	Média	Média/Alta	Alta	Média
11. Interação/Refina mento	Alta	Alta	Baixa	Alta	Alta	Alta
12. Multimodalidade	Alta	Alta	Baixa	Baixa	Média	Baixa

Fonte: Próprio autor

5.1 Análise interpretativa e crítica das buscas

A análise comparativa das seis plataformas de Inteligência Artificial revela uma conclusão central para a prática jurídica: não emerge uma única ferramenta como universalmente superior. Em vez disso, os resultados demonstram um claro e fundamental trade-off¹ entre a profundidade da análise e a confiabilidade das fontes. A escolha da ferramenta mais adequada, portanto, deixa de ser uma questão de preferência e passa a ser uma decisão estratégica, dependente da fase do trabalho e da tolerância ao risco do profissional.

Neste espectro, o Gemini Advanced se destacou como a ferramenta com maior potencial para a pesquisa profunda. Ele foi capaz de entregar o relatório mais completo, com uma análise detalhada e abrangente. No entanto, essa profundidade veio com um custo significativo em confiabilidade. A plataforma demonstrou uma propensão preocupante a "alucinar", fornecendo fontes que, após verificação, se mostraram inexistentes. Isso implica que, embora seja um excelente ponto de partida para gerar teses e explorar múltiplos ângulos de um problema, o Gemini exige um rigoroso trabalho investigativo subsequente por parte do advogado para validar cada informação e referência. A funcionalidade de extrair o relatório

¹ Segundo o Dicionário Financeiro o Trade-off é um termo em inglês muito utilizado na economia e que define as situações em que existem conflitos de escolha. Este conceito aborda o resultado de uma escolha comparando o que foi selecionado em detrimento daquilo que se abriu mão. Um dos exemplos mais comuns está presente em nossos hábitos de consumo. Quando optamos por um produto ou serviço abrimos mão das outras opções que temos, onde ele aparece.

diretamente para o Google Docs otimiza o fluxo de trabalho, mas não diminui a carga da verificação manual.

O ChatGPT Pro posicionou-se de forma semelhante, produzindo uma análise de alta qualidade e com um índice de alucinação inferior ao do Gemini. Contudo, essa maior confiabilidade aparente foi acompanhada por um relatório menos profundo. Essa observação sugere que, dentro das ferramentas de ponta, pode haver uma correlação inversa entre a criatividade/profundidade da IA e sua precisão factual. Ambas as ferramentas funcionam mais como "sócios criativos" para a fase de ideação do que como assistentes de pesquisa plenamente confiáveis.

No extremo oposto do espectro encontra-se o Perplexity Pro. Apesar de ter sido a maior decepção em termos de profundidade, gerando o relatório mais simples entre as versões pagas, seu desempenho em confiabilidade foi impecável. Foi a única ferramenta que não apresentou nenhuma fonte inventada, obtendo o menor (neste caso, nulo) índice de alucinação. Isso o posiciona não como uma ferramenta de *deep research*, mas como um verificador de fatos e um construtor de bibliografias de altíssima segurança. Seu valor não está na geração de teses, mas na construção de uma base de pesquisa sólida e livre de erros factuais.

Por fim, o desempenho inferior de Copilot, Grok e DeepSeek, que forneceram relatórios com menor profundidade e fontes menos confiáveis, deve ser analisado com uma ressalva metodológica crucial: foram utilizadas suas versões gratuitas. Este resultado pode ser um reflexo direto das limitações impostas a esses níveis de acesso, e não necessariamente da capacidade intrínseca de suas tecnologias. A análise evidencia que o nível de assinatura é uma variável dominante, e uma comparação mais direta exigiria o uso das versões "Pro" de todas as plataformas.

Diante desse cenário, a principal implicação para o profissional do Direito é a necessidade de atuar como um gestor de um portfólio de IAs. A pesquisa jurídica moderna se torna um processo multifásico: inicia-se com ferramentas de alta profundidade (como Gemini ou ChatGPT) para o brainstorming e a estruturação de teses, sempre com ceticismo; em seguida, utiliza-se uma ferramenta de alta confiabilidade (como o Perplexity) para validar

as fontes e construir a fundamentação factual; e, por fim, pode-se usar ferramentas de acesso gratuito para consultas rápidas e tarefas complementares. A responsabilidade final, portanto, não apenas permanece com o advogado, mas é ampliada, exigindo novas competências de validação, gestão de risco e orquestração de múltiplas tecnologias.

6 CONSIDERAÇÕES FINAIS

O presente estudo evidenciou que a funcionalidade de Pesquisa Avançada representa um marco disruptivo na prática jurídica contemporânea, ao redefinir a forma como se acessa, interpreta e aplica o conhecimento jurídico. A análise demonstrou que nenhuma das plataformas avaliadas oferece, isoladamente, uma solução completa: ferramentas como Gemini e ChatGPT destacam-se pela profundidade, enquanto Perplexity se sobressai pela confiabilidade. Nesse cenário, a eficiência passa a depender da capacidade do profissional em combinar diferentes ferramentas conforme as etapas da pesquisa. Além disso, o artigo chama atenção para as implicações éticas do uso da IA, especialmente quanto à responsabilidade sobre a veracidade das informações, proteção de dados sensíveis e o risco de atrofia cognitiva. Torna-se, portanto, imperativo que operadores do Direito desenvolvam competências críticas, éticas e técnicas para interagir com essas tecnologias, garantindo que a inteligência artificial seja um instrumento de qualificação e não de precarização da atuação jurídica.

REFERÊNCIAS

BAROCAS, Solon; SELBST, Andrew D. Big Data's Disparate Impact. **California Law Review**, Berkeley, CA, v. 104, p. 671, 2016. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2477899. Acesso em: 25 jun. 2025.

CARR, Nicholas. **The Shallows**: What the Internet Is Doing to Our Brains. New York: W. W. Norton & Company, 2011.

JURAFSKY, D.; MARTIN, J. H. **Speech and language processing**. 3. ed., 2023. Disponível em: <https://web.stanford.edu/~jurafsky/slp3/>. Acesso em: 28 jul. 2025.

MORETTI, Juliano Lazzarini; ZUFFO, Milena Maltese. LGPD e inteligência artificial: um estudo comparado. **DIGE - Direito Internacional e Globalização Econômica**, v. 13, n. 13, 4 mar. 2025. Disponível em: <https://doi.org/10.23925/2526-6284/2023.v13n13.69370>. Acesso em: 08 jun. 2025.

PARK, Joon Sung et al. Generative agents: Interactive simulacra of human behavior. **Proceedings of the 36th annual acm symposium on user interface software and technology**. 2023. p. 1-22. Disponível em: <https://arxiv.org/abs/2304.03442>. Acesso em: 28 jul. 2025.

SANCHEZ, C. S. et al. Revisão de escopo da literatura sobre IA generativa e pesquisa qualitativa. **Revista Pesquisa Qualitativa**, v. 12, n. 30, p. 1-28, abr. 2024. Disponível em: <https://editora.sepq.org.br/rpq/article/download/729/467/4339>. Acesso em: 29 jul. 2025.

SPARROW, Betsy; LIU, Jenny; WEGNER, Daniel M. Google Effects on Memory: Cognitive Consequences of Having Information at Our Fingertips. **Science**, Washington, D.C., v. 333, n. 6043, p. 776-778, ago. 2011. Disponível em: https://scholar.harvard.edu/files/dwegner/files/sparrow_et_al._2011.pdf. Acesso em 28. jun. 2025.

TRADE-OFF. In: Dicionário Financeiro, [202-?]. Disponível em: <https://www.dicionariofinanceiro.com/trade-off/>. Acesso em: 29 jun. 2025.

TRINDADE, A. S. C. E.; OLIVEIRA, H. P. C. Inteligência artificial (IA) generativa e competência em informação: habilidades informacionais necessárias ao uso de ferramentas de IA generativa em demandas acadêmicas-científicas. **Perspectivas em Ciência da Informação**, v. 29, 2024. Disponível em: <https://www.scielo.br/j/pci/a/GVCW7KbcRjGVhLSrmy3PCng/?format=pdf&lang=pt>. Acesso em: 29 jun. 2025.

WEI, Jason et al. Chain-of-thought prompting elicits reasoning in large language models. **Advances in neural information processing systems**, v. 35, 2022. Disponível em: <https://arxiv.org/abs/2201.11903>. Acesso em: 28 maio 2025.

YOSINSKI, Jason et al. Foundation agents. **arXiv preprint**, arXiv:2505.10468v1, 2025. Disponível em: <https://arxiv.org/abs/2505.10468v1>. Acesso em: 10 jul. 2025.