Revista Ibero-Americana de Humanidades, Ciências e Educação — REASE

doi.org/10.51891/rease.v12i5.26075

O USO DE INTELIGÊNCIA ARTIFICIAL GENERATIVA POR PROFISSIONAIS DO DIREITO E O DEVER DE PROTEÇÃO DE DADOS CONFIDENCIAIS

THE USE OF GENERATIVE ARTIFICIAL INTELLIGENCE BY LEGAL PROFESSIONALS AND THE DUTY TO PROTECT CONFIDENTIAL DATA

EL USO DE INTELIGENCIA ARTIFICIAL GENERATIVA POR PROFESIONALES DEL DERECHO Y EL DEBER DE PROTEGER DATOS CONFIDENCIALES

Michele Cristine Soares Campos Volpe[1]
Giovanna Mantovani Salim[2]

RESUMO: O presente artigo analisa os riscos jurídicos decorrentes da utilização de sistemas de inteligência artificial generativa por profissionais do direito, com enfoque na tensão entre a eficiência operacional proporcionada por essas ferramentas e o dever de sigilo profissional e de proteção de dados pessoais. A partir da compreensão dos fundamentos técnicos dos modelos de linguagem de grande escala (LLMs) e dos mecanismos de armazenamento, retenção e reutilização de dados, o artigo examina como a inserção de informações confidenciais de clientes em plataformas abertas de IA pode configurar violação ao sigilo profissional previsto no Estatuto da Advocacia e ao regime de proteção da Lei Geral de Proteção de Dados (LGPD). A análise percorre casos documentados de exposição de dados, os deveres de governança informacional aplicáveis ao advogado na qualidade de controlador de dados e as hipóteses de responsabilidade civil decorrentes do uso negligente dessas tecnologias. Conclui-se que a adoção responsável da IA generativa na advocacia exige competência digital, governança informacional efetiva e proteção real do cliente como condições indissociáveis do exercício ético da profissão no século XXI.

Palavras-chave: Inteligência artificial generativa. Sigilo profissional. Proteção de dados. Responsabilidade civil do advogado. Governança informacional.

ABSTRACT: This article analyzes the legal risks arising from the use of generative artificial intelligence systems by legal professionals, focusing on the tension between the operational efficiency provided by these tools and the duty of professional confidentiality and personal data protection. Based on an understanding of the technical foundations of large language models (LLMs) and the mechanisms of data storage, retention, and reuse, the article examines how the insertion of clients' confidential information into open AI platforms may constitute a violation of professional secrecy as established in the Statute of the Legal Profession, as well as of the data protection framework set forth by the General Data Protection Law (LGPD). The analysis explores documented cases of data exposure, the informational governance duties applicable to lawyers in their role as data controllers, and the hypotheses of civil liability arising from the negligent use of such technologies. It concludes that the responsible adoption of generative AI in legal practice requires digital competence, effective informational governance, and the genuine protection of clients as essential conditions for the ethical practice of law in the 21st century.

Keywords: Generative artificial intelligence. Professional secrecy. Data protection. Lawyer's civil liability. Informational Governance.

[1] Mestranda em Direito Comercial pela Pontifícia Universidade Católica de São Paulo - (PUC-SP).
[2] Mestranda em Direito Civil pela Pontifícia Universidade Católica de São Paulo - (PUC-SP).

RESUMEN: El presente artículo analiza los riesgos jurídicos derivados del uso de sistemas de inteligencia artificial generativa por parte de profesionales del derecho, con especial énfasis en la tensión entre la eficiencia operativa que dichas herramientas proporcionan y el deber de secreto profesional y de protección de datos personales. A partir de la comprensión de los fundamentos técnicos de los modelos de lenguaje de gran escala (LLMs) y de los mecanismos de almacenamiento, retención y reutilización de datos, el artículo examina cómo la introducción de información confidencial de los clientes en plataformas abiertas de IA puede constituir una violación del secreto profesional previsto en el Estatuto de la Abogacía, así como del régimen de protección de datos establecido por la Ley General de Protección de Datos (LGPD). El análisis abarca casos documentados de exposición de datos, los deberes de gobernanza informacional aplicables al abogado en su condición de responsable del tratamiento de datos y las hipótesis de responsabilidad civil derivadas del uso negligente de dichas tecnologías. Se concluye que la adopción responsable de la IA generativa en la práctica jurídica exige competencia digital, una gobernanza informacional efectiva y la protección real del cliente como condiciones indispensables para el ejercicio ético de la profesión en el siglo XXI.

Palabras clave: Inteligencia artificial generativa. Secreto professional. Protección de datos. Responsabilidad civil del abogado. Gobernanza informacional.

1. INTRODUÇÃO

A incorporação da inteligência artificial generativa na prática jurídica deixou de ser uma possibilidade futura para se tornar realidade cotidiana, em que advogados utilizam hoje modelos de linguagem de grande escala, em sua maioria, sem qualquer conhecimento técnico sobre a ferramenta, para redigir minutas contratuais, analisar documentos complexos, pesquisar precedentes e estruturar estratégias processuais, tarefas que, até poucos anos atrás, demandavam horas de trabalho técnico especializado, portanto a eficiência operacional é inegável, porém um problema surge no que o profissional precisa entregar ao sistema para conseguir o que deseja: narrativas fáticas completas, documentos do cliente, dados pessoais sensíveis, vulnerabilidades da parte adversa e estratégias confidenciais.

Essa prática coloca em tensão direta dois pilares que o ordenamento brasileiro protege com rigor extremo: a modernização tecnológica da advocacia e a inviolabilidade do sigilo profissional. O dever de confidencialidade é pilar fiduciário da relação advogado-cliente e foi concebido para um mundo analógico, ele não foi desenhado para um cenário em que informações estratégicas são enviadas para servidores de terceiros, frequentemente localizados fora do Brasil, processadas em arquiteturas opacas de nuvem e sujeitas a políticas de retenção, logging e possível reutilização dos dados que o usuário muitas vezes desconhece, já que a leitura na íntegra dos termos de uso e condições de serviços online é extremamente reduzida.

A lacuna regulatória e doutrinária também se impõe, afinal o Código de Ética da Ordem dos Advogados do Brasil (OAB) e o Estatuto da Advocacia não tratam expressamente do uso

de IA generativa e a Lei Geral de Proteção de Dados (LGPD) regula o tratamento de dados, mas não responde diretamente à posição do advogado como controlador que delega operações de tratamento a operadores estrangeiros sem governança prévia. Embora a OAB tenha emitido recomendação nos últimos anos sobre o uso de IA e o Conselho Nacional de Justiça (CNJ) tenha disciplinado o uso de IA no Judiciário, a responsabilidade civil do advogado individual que, sem avaliação técnica ou consentimento do cliente, insere informações confidenciais em plataformas de terceiros permanece terreno pouco explorado pela doutrina nacional.

É nesta lacuna que o presente artigo se insere, cuja problemática central investigada é "Em que medida a utilização de sistemas de inteligência artificial generativa por profissionais do direito, especialmente mediante a inserção de informações confidenciais de clientes em plataformas tecnológicas operadas por terceiros, pode caracterizar violação ao dever de sigilo profissional e aos deveres jurídicos de proteção de dados e governança informacional no exercício da advocacia?"

Para demonstrá-la, o artigo segue método hipotético-dedutivo com revisão bibliográfica e análise de casos documentais, observando trajeto que examina primeiro o funcionamento técnico dos modelos de linguagem de grande escala (LLMs) e os riscos informacionais inerentes à sua arquitetura (Capítulo 2), reconstrói o dever de confidencialidade em suas dimensões ética e protetiva (Capítulo 3), analisa os riscos concretos de exposição com casos documentados (Capítulo 4), traduz os deveres éticos em obrigações da LGPD e governança (Capítulo 5), e por fim, examina as hipóteses de responsabilidade civil por violação de sigilo e falha no dever de diligência (Capítulo 6).

O uso da IA é uma realidade, de forma que não é razoável compreender pela proibição da tecnologia, já que seria apenas uma maneira fechar os olhos para a realidade, porém faz-se necessário um chamado urgente à sua adoção responsável, com governança informacional efetiva, competência digital mínima e proteção real do cliente como condição sine qua non do exercício ético da advocacia no século XXI.

2. INTELIGÊNCIA ARTIFICIAL GENERATIVA E PROCESSAMENTO DE INFORMAÇÃO

A adequada compreensão dos desafios jurídicos associados ao uso de sistemas de inteligência artificial (IA) generativa pressupõe, como condição metodológica indispensável, a delimitação de seus fundamentos técnicos essenciais. As percepções imprecisas sobre o funcionamento desses sistemas, frequentemente tratados como "caixas pretas", dificultam a

formulação de respostas jurídicas proporcionais e eficazes, especialmente no que concerne à confidencialidade de informações.

Diferentemente dos softwares tradicionais de processamento determinístico, a IA generativa opera por meio de modelo estatístico probabilístico que identifica padrões em grandes volumes de dados para inferir probabilidades de formações linguísticas no contexto pretendido. Conforme Russell e Norvig (2013), o objetivo desses sistemas consiste em maximizar a probabilidade e determinadas saídas com base em dados previamente observados, sendo o aprendizado um processo contínuo de ajustes de parâmetros internos do modelo.

De modo que o sistema de IA generativa passa a produzir novos conteúdos a partir de padrões de treinamentos, não sendo a mera recuperação de informações armazenadas, mas a geração probabilística de respostas com algum nível de aleatoriedade, o que implica que a qualidade e o volume de dados de treinamento são a resposta chave para a performance do modelo, de forma que as informações inseridas pelos usuários podem ser processadas por múltiplas camadas, desde o armazenamento temporário, até eventual utilizada para aprimoramento do sistema.

Essa arquitetura, aliada à computação em nuvem e o processamento distribuído, transfere o controle efetivo dos dados para terceiros, o que reforça a necessidade de análise jurídica sobre o regime de tratamento de dados aplicado.

2.1 Funcionamento básico de modelos de linguagem de grande escala (LLM)

Os Large Language Models ou linguagem de grande escala (LLMs) representam uma das mais sofisticadas aplicações do aprendizado de máquina, estruturando-se na modelagem probabilística da linguagem para prever sequências de palavras a partir de um contexto fornecido. Sua arquitetura fundamental foi introduzida no artigo seminal "Attention is All You Need" publicado em 2017, que substituiu as redes neurais recorrentes (RNNs/LSTMs) pelo mecanismo de self-attention (atenção auto-referencial), cujo mecanismo permite que o modelo processe toda a sequência de entrada em paralelo, capturando dependências de longo alcance entre palavras ou tokens, independentemente da distância entre elas, sendo esta a arquitetura Transformer utilizada amplamente nos modelos de IA generativa atualmente existentes.

Conforme a abordagem de Russell e Norvig (2013), a linguagem presente na IA generativa pode ser tratada como fenômeno estatístico, no qual a probabilidade de uma palavra ou sequência depende de sua frequência e de suas relações com outros elementos linguísticos

em grande conjunto de dados. O processo técnico da IA generativa neste contexto ocorre em duas etapas principais:

I. Pré-treinamento: O modelo é exposto a bilhões ou trilhões de tokens extraídos da internet, livros, códigos e textos públicos. A tarefa é preditiva: dado um contexto o sistema prever o próximo token (unidade mínima de processamento). O modelo ajusta internamente bilhões de parâmetros (também chamados tecnicamente de pesos sinápticos) por meio de gradiente descendente, aprendendo padrões estatísticos dessas sequências; e

II. Alinhamento pós-treinamento (também denominado RLHF – Reinforcement Learning from Human Feedback ou simplesmente Aprendizado por Reforço): O modelo é refinado para gerar respostas úteis, seguras e alinhadas a preferências humanas.

Neste contexto, o treinamento envolve a exposição a volume massivo de dados, incluindo textos, por meio dos quais o sistema ajusta bilhões de parâmetros internos para otimizar a previsão da próxima unidade linguística (token) em uma sequência, ou seja, o modelo não compreende o conteúdo no sentido cognitivo da palavra, mas estabelece relações estatísticas entre padrões linguísticos.

Apesar disso, a ausência de compreensão semântica não elimina a capacidade do sistema de reproduzir ou inferir informações sensíveis presentes nos dados de treinamento ou fornecidas como entrada pelo usuário, o que pode ocorrer por generalização de padrões, o que é comum e necessário neste tipo de tecnologia.

Na fase de uso dessa tecnologia, a interação com LLMs ocorre por meio de entradas manuais dos usuários, chamadas de *prompts*, que são informações textuais que constituem o principal mecanismo de entrada de dados no sistema. Esse *prompt* é tokenizado, convertido em vetores e processado por múltiplas camadas de atenção na arquitetura transformer, de modo que a qualidade e especificidade das instruções contidas no *prompt* (e eventuais documentos adicionais enviados com eles) influenciam diretamente o resultado gerado como saída do sistema, conferindo a engenharia de *prompt* relevância significativa em contexto profissionais nos quais a precisão e a confidencialidade são essenciais.

Após o processamento do *prompt*, o modelo gera uma saída token por token, escolhendo sempre a sequência mais provável com base nos dados utilizados pelo sistema (e no sistema), não havendo armazenamento permanente da conversa no modelo em si, estando o "conhecimento" armazenado nos parâmetros de treinamento, apesar disso todo o conteúdo do *prompt* é processado (via de regra) em servidores do provedor alheios ao controle, guarda e arquivo do usuário.

Neste contexto, a capacidade de generalização e memorização parcial do LLM, é capaz de gerar risco jurídico de informações confidenciais inseridas inadvertidamente pelo usuário poderem ser reproduzidas e inseridas em outras interações com o mesmo usuário ou outros usuários da mesma tecnologia.

Um aspecto técnico de especial relevância jurídica, frequentemente negligenciado na literatura jurídica sobre o tema, é a distinção entre as diferentes modalidades de acesso aos sistemas de IA generativa e seus respectivos regimes de tratamento de dados. O uso de interfaces públicas gratuitas ou de planos individuais pagos, como o ChatGPT em suas versões de consumo, opera sob política de tratamento que, em regra, permite ao provedor utilizar as interações para aprimoramento do modelo, salvo configuração específica pelo usuário.

Em contraste, o acesso via application programming interface (API) corporativa ou planos empresariais dedicados normalmente prevê que os dados submetidos não sejam utilizados para treinamento, e estabelece janelas de retenção definidas e controláveis. Essa distinção não é meramente técnica, pois ela é juridicamente determinante para a avaliação do risco informacional e da responsabilidade do advogado, pois o nível de exposição dos dados do cliente varia substancialmente conforme a modalidade de acesso adotada pelo profissional, aspecto que a maioria dos usuários ignora integralmente ao escolher a ferramenta que irá utilizar em sua rotina.

2.2 A lógica de entrada de dados e os riscos de exposição de informações confidenciais

Diferentemente do que ocorre em interfaces de sistemas e softwares tradicionais com campos estruturados e processamento determinístico em que o sistema possui locais com campos especificados para preenchimento de informações ou seleção dentro de uma estrutura determinada, os sistemas de IA generativa operam com entrada por meio de *prompt* flexíveis e altamente contextuais, cujo número de informações e detalhamento de contexto pode ser longo e complexo.

Por conseguinte, em contextos profissionais especialmente na advocacia, consultoria empresarial e estratégica e gestão de dados corporativos, é comum que os *prompts* contenham informações estratégicas, documentos confidenciais ou dados pessoais, afinal carregam narrativas fáticas completas, trechos de documentos, depoimentos, estratégias processuais e dados pessoais sensíveis.

Como a lógica de funcionamento desses sistemas, baseada na inferência probabilística a

partir de dados de entrada, isso implica que o conteúdo fornecido pelo usuário terá o poder de influenciar diretamente a saída gerada, sendo essencial para isso que o texto usado na entrada como *prompt* seja objeto de tratamento adicional, totalmente dependente da arquitetura do sistema e das políticas do provedor (geralmente instalados nos Estados Unidos ou Europa).

Em determinadas configurações, os dados inseridos pelo usuário podem ser registrados para fins de monitoramento, auditoria, segurança ou aprimoramento do modelo, o que amplia o espectro de risco associado à sua utilização.

A consequência direta é que as informações faticas e juridicas compartilhadas pelo cliente, quando compartilhadas via *prompt* pelo usuário com o modelo de IA generativa aberto ao público, poderá acarretar o processamento destes dados via provedores terceiros, de forma, que tais dados passam a integrar (ainda que não permanentemente) a base de conhecimento da tecnologia.

Como resultado, essa prática, embora funcional do ponto de vista operacional, transfere conteúdo protegido para ambientes tecnológicos externos ao controle do usuário, de forma que, na perspectiva jurídica, isso suscita questionamentos sobre a responsabilidade do usuário pela inserção de informações confidenciais em sistemas de terceiros e os deveres de diligência no tratamento de tais dados sensíveis, especialmente na advocacia com o sigilo profissional, tutelado por normas éticas e disposições legais específicos.

2.3 Armazenamento, retenção e reutilização de dados em sistemas de IA generativa

Embora a experiência do usuário nas plataformas de IA frequentemente transmita a impressão de interação efêmera, a realidade técnica revela múltiplas camadas de processamento que podem envolver desde o armazenamento temporário até registros mais persistentes. Os sistemas de IA operam em arquitetura distribuída, nas quais os dados são transmitidos e processados em serviços controlados pelo provedor, ou seja, o modelo dissocia o local de inserção da informação do ambiente efetivo de seu processamento introduzindo complexidade jurídica quanto a sua governança e proteção.

A depender das políticas do provedor, as interações podem ser objeto de registro para fins de segurança e auditoria ou aprimoramento do desempenho (como é o caso do logging por exemplo).

Em alguns casos, dados inseridos pelo usuário podem ser utilizados, de forma agregada e anonimizada, para o treinamento ou ajuste fino (também chamado de *fine-tuning*) dos

modelos, conforme leciona Russell e Norvig (2013), os sistemas de aprendizado de máquina dependente da disponibilidade dos dados para ajustar seus parâmetros e melhorar seu desempenho ao longo do tempo, sendo esta uma característica estrutural do sistema, que reforça a tendência de utilização de dados provenientes da interação com usuários como insumo para o desenvolvimento contínuo dos modelos.

Dessa forma, ainda que o modelo seja estruturado para mitigar o risco de identificação individual da informação, não se pode afastar completamente a possibilidade de reidentificação ou inferência indireta de informações sensíveis que tenham sido incluídas no modelo pelo usuário (de forma propositada ou não). Nesta linha, verifica-se pluralidade relevante de IAs generativas sendo utilizadas por usuários advogados em suas rotinas, através de versões gratuitas ou versões individuais pagas que permitem que as informações sejam utilizadas para aprendizagem do modelo via aprendizagem de reforço, através do loop de feedback naturalmente existentes por estas plataformas, sendo que o uso destes dados, em regra, é ignorado pelos usuários que desconhecem os Termos de Uso e Privacidade da IA generativa que está sendo utilizada.

Sob a ótica normativa, esse cenário dialoga diretamente com os regimes de proteção de dados pessoais, como a Lei Geral de Proteção de Dados (LGPD), exigindo a observância de bases legais, princípios de finalidade, necessidade e segurança, além de transparência quanto às práticas do controlador. No âmbito corporativo, a inserção de informações estratégicas pode comprometer segredos industriais e know-how, especialmente se tais dados forem armazenados ou reutilizados em contextos que escapem o controle efetivo das empresas e escritórios de advocacia. A ausência de clareza sobre os regimes de tratamento de dados aplicáveis, aliada à opacidade técnica de muitos sistemas de IA, agrava esse risco, tornando indispensável a adoção de políticas internas de governança e de uso seguro dessas ferramentas.

Dessa forma, a análise dos mecanismos de armazenamento, retenção e reutilização de dados revela que o uso de IA generativa envolve, necessariamente, a inserção de informações em ambientes tecnológicos complexos, nos quais o controle do usuário é limitado, e tal constatação impõe a necessidade de reconfiguração dos parâmetros jurídicos tradicionais de confidencialidade, de modo a compatibilizá-los com a realidade técnica desses sistemas e a garantir a proteção efetiva de informações sensíveis em um contexto de crescente digitalização e automação.

A dimensão técnica do risco de retenção e extração de dados foi empiricamente

demonstrada pela pesquisa de segurança computacional, assim demonstraram que, em modelos de LLMs treinados em bases de dados privadas, é possível realizar ataques de extração de dados de treinamento por meio de consultas ao modelo, recuperando sequências dos dados utilizados no treinamento (Carlini et al., 2021), incluindo informações de identificação pessoal como nomes, números de telefone e endereços de e-mail.

Este mesmo estudo, identificou, ainda, que modelos maiores são proporcionalmente mais vulneráveis a esse tipo de ataque do que modelos menores, o que é especialmente relevante dado que são exatamente os modelos de maior porte os mais amplamente utilizados por profissionais do direito em razão de suas capacidades superiores de compreensão e geração textual.

O dever de sigilo é um dos pilares da relação entre cliente e o advogado, sendo simultaneamente, um direito fundamental do cidadão e uma obrigação profissional inafastável do advogado.

Esse achado científico reforça que o risco de exposição informacional não é teórico ou especulativo, mas empiricamente verificado e inerente à arquitetura dos sistemas mais populares, argumento que, no plano jurídico, afasta qualquer alegação de que o risco era imprevisível ou alheio ao dever de diligência do profissional que adota a ferramenta.

3. O DEVER DE CONFIDENCIALIDADE NA ATIVIDADE JURÍDICA

A confidencialidade não é um mero atributo da profissão, mas uma condição *sine qua non* para o seu exercício pleno e para a efetivação do acesso à justiça. Ela viabiliza a necessária relação de confiança que permite ao cliente expor, de forma irrestrita e sem reservas, todos os fatos e circunstâncias pertinentes à sua causa, assegurando ao advogado o conhecimento integral dos elementos indispensáveis à elaboração de uma defesa técnica adequada e eficaz.

A obrigação de manter sigilo profissional é complexa, pois para o cliente ela funciona como um direito, oponível contra todos, especialmente contra o Estado e seus agentes. Esse direito aparece, por exemplo, (i) na garantia de inviolabilidade do escritório do advogado, de seus arquivos, dados, correspondências e comunicações, conforme o artigo 7º, II, da Lei 8.906/94 (“Estatuto da OAB”), (ii) na comunicação pessoal e reservada com clientes recolhidos presos, conforme artigo 7º, III do Estatuto da OAB; e (iii) na recusa em depor como testemunha em processo em que funcionou ou deva funcionar, bem como sobre fatos relacionados com cliente ou ex-cliente, ou que constituam sigilo (art. 7º, XIX, do Estatuto da OAB).

Já para o advogado, o sigilo é um dever profissional, e descumpri-lo pode gerar punições

disciplinares, civis e até penais. Por isso se diz que o sigilo profissional é um múnus público: uma obrigação imposta ao advogado para proteger o interesse da sociedade e assegurar o bom funcionamento da Justiça. Ele não pode ser renunciado, salvo em situações muito específicas previstas em lei e nas normas éticas da profissão.

Quebrar esse sigilo não prejudica apenas o cliente envolvido, mas também compromete a credibilidade de toda a advocacia e enfraquece a confiança da sociedade na atuação do advogado como peça essencial da justiça.

3.1 Sigilo profissional na advocacia como dever fiduciário inerente à prestação de serviços jurídicos

A relação entre cliente e advogado é baseada essencialmente na confiança. O cliente não espera apenas que o advogado trabalhe de forma técnica e competente; ele também entrega ao profissional seus segredos, interesses e aspectos íntimos da vida, muitas vezes relacionados ao patrimônio ou até à própria liberdade. Por isso, o sigilo profissional é a principal expressão da lealdade que deve existir nessa relação de confiança.

Esse sigilo não é um detalhe do contrato de prestação de serviços. Ele é uma parte essencial da atuação do advogado e condição para que o mandato seja válido e eficaz. A confiança exige que o advogado atue sempre com honestidade e boa-fé, colocando o interesse do cliente acima de qualquer outro, inclusive acima de seus próprios. Assim, o sigilo é uma consequência natural e indispensável dessa obrigação.

Violar essa confiança, quebrando o sigilo, significa negar a própria essência da advocacia.

A legislação deixa isso muito claro em diversos dispositivos. O Estatuto da OAB, no artigo 34, inciso VII, classifica como infração disciplinar violar o sigilo profissional sem justa causa. A regra é rígida porque o legislador reconhece a importância desse dever. A penalidade pode ir desde uma simples censura até a exclusão do advogado dos quadros da OAB, o que demonstra que o sigilo não é só uma questão ética, mas sim uma obrigação jurídica obrigatória.

Ainda, o artigo 35 do Estatuto da OAB determina que o advogado deve guardar sigilo sobre todos os fatos de que tenha conhecimento no exercício da profissão. Esse dever se estende a depoimentos, informações obtidas em negociações e até confidências feitas por testemunhas e terceiros. Esse conjunto de normas mostra que a confidencialidade é parte integrante do papel do advogado e serve como uma garantia essencial para quem busca proteção de seus direitos.

O dever de sigilo também alcança todo o escritório, incluindo sócios, associados, estagiários e funcionários administrativos. Cabe ao advogado responsável implementar

controles, treinamentos e práticas que assegurem que todos respeitem a confidencialidade.

Ainda, o sigilo não se restringe ao tempo de prestação dos serviços, afinal, de acordo com o artigo 19 do Estatuto da OAB, o advogado, mesmo que ao postular contra ex-cliente ou ex-empregador, deverá manter o sigilo sob quaisquer informações reservadas ou privilegiadas que lhe tenham sido confiadas à época da prestação de serviços.

Esse cuidado se torna ainda mais importante no mundo digital. Ao usar serviços em nuvem, aplicativos de mensagens e ferramentas de IA, o advogado precisa agir como guardião dos dados do cliente. A escolha de qualquer tecnologia deixa de ser algo meramente operacional e passa a ser uma decisão ligada ao dever fiduciário, exigindo avaliação rigorosa sobre segurança e conformidade com as regras de sigilo e proteção de dados.

A proteção da confidencialidade na advocacia não se limita a normas deontológicas ou infraconstitucionais, encontrando fundamento direto na Constituição Federal de 1988. O artigo 5º, inciso X, assegura a inviolabilidade da intimidade, da vida privada, da honra e da imagem, garantindo o direito à indenização por danos decorrentes de sua violação. O inciso XIV do mesmo dispositivo, ao proteger o sigilo da fonte, permite, por interpretação sistemática, a extensão dessa garantia ao sigilo profissional do advogado, como condição necessária ao exercício de função essencial à justiça, conforme previsto no artigo 133 da Constituição. Mais recentemente, a Emenda Constitucional nº 115/2022 acrescentou o inciso LXXIX ao artigo 5º, elevando a proteção de dados pessoais (inclusive em ambiente digital) à condição de direito fundamental autônomo, com tutela constitucional expressa.

Em resumo, o dever de sigilo existe para garantir que a confiança entre cliente e advogado seja preservada da forma mais ampla possível. Ele não é um privilégio do advogado, mas uma garantia essencial do cidadão e um elemento indispensável para o funcionamento da justiça.

3.2 Proteção de informações e dados de clientes no exercício da advocacia

A obrigação de sigilo profissional, que sempre fez parte da advocacia, ganhou uma nova dimensão com a entrada em vigor da Lei Geral de Proteção de Dados (LGPD). Essa lei criou um sistema completo de regras sobre como tratar dados pessoais, impondo obrigações a qualquer pessoa ou organização que lide com informações como coleta, uso, armazenamento, compartilhamento ou eliminação de dados.

Na advocacia, lidar com dados pessoais faz parte da rotina. Advogados tratam grandes volumes de informações, muitas vezes sensíveis. Por isso, seguir a LGPD não é opcional: é uma necessidade inseparável da prática jurídica atual.

De acordo com a LGPD, advogados e escritórios atuam, na maior parte do tempo como controladores de dados, ou seja, são eles que decidem como e por que os dados dos clientes, funcionários e até de partes adversas serão tratados, essa posição traz uma série de responsabilidades que vão muito além do sigilo tradicional.

A lei exige uma gestão ativa, organizada e documentada dos dados, com base em princípios como finalidade, necessidade, transparência, qualidade, prevenção e responsabilização. O advogado deve demonstrar que usa apenas os dados necessários e para objetivos legítimos, específicos e informados ao titular.

O artigo 7º da LGPD lista as bases legais que autorizam o tratamento de dados pessoais. Na advocacia, várias delas se aplicam diretamente: (i) execução de contrato, como o contrato de mandato entre cliente e advogado; (ii) cumprimento de obrigação legal, por exemplo, guardar documentos exigidos pela lei fiscal ou previdenciária; e (iii) exercício regular de direitos em processos, que autoriza o uso dos dados necessários para atuar judicial, administrativa ou arbitral.

Ainda assim, mesmo autorizado por bases legais, o advogado deve seguir todos os princípios da LGPD e garantir transparência e segurança no tratamento de dados.

A situação se torna mais delicada quando envolve dados sensíveis, como informações sobre saúde, convicções religiosas, opinião política, vida sexual, dados biométricos ou origem racial. Em várias áreas do direito, é comum lidar com esse tipo de dado. Nesse caso, a regra é ainda mais restrita: para a advocacia, normalmente o tratamento só é permitido para o exercício regular de direitos em processo, conforme o artigo 11 da LGPD.

A responsabilidade do advogado é máxima. Ele deve adotar medidas técnicas e administrativas para proteger esses dados contra acessos indevidos, perda, alteração ou qualquer forma de uso inadequado. Ferramentas como criptografia, controle rigoroso de acesso aos sistemas, descarte seguro de documentos e treinamentos frequentes da equipe deixam de ser boas práticas e passam a ser obrigações legais.

Essas medidas se somam ao dever de sigilo profissional, reforçando a proteção que o advogado deve garantir aos dados que lhe são confiados, não apenas aqueles classificados como sigilosos ou confidenciais.

4. USO DE IA GENERATIVA E O RISCO DE EXPOSIÇÃO DE INFORMAÇÕES CONFIDENCIAIS

A análise do uso de sistemas de IA generativa no exercício da advocacia exige o deslocamento do debate tecnológico para o plano jurídico-normativo, pois não se trata apenas de compreender como esses sistemas operam, mas de examinar as consequências jurídicas decorrentes de sua utilização no tratamento de informações protegidas pelo sigilo profissional. Nesse contexto, o elemento central não reside na existência de vazamento efetivo de dados, mas na alteração estrutural do regime de controle informacional que tradicionalmente caracteriza a atividade advocatícia.

A utilização de ferramentas de IA generativa implica, como premissa operacional, a interação com ambientes tecnológicos que não integram a esfera de domínio do profissional do direito, afinal diferentemente dos sistemas internos ou de infraestrutura controlada tradicional dos softwares jurídicos tradicionais dos escritórios de advocacia por exemplo, essas plataformas são operadas de forma ampla e aberta por terceiros, submetidas a regimes próprios de governança e frequentemente localizadas em jurisdições diversas.

Esse deslocamento altera o paradigma clássico de proteção da informação jurídica, baseado na custódia direta e na inviolabilidade dos meios de comunicação do advogado e, sob essa perspectiva, a questão central não é apenas a confidencialidade em sentido estático, mas a perda de controle sobre o ciclo de vida da informação.

Uma vez que para uso da IA generativa em sua máxima capacidade (atual) é necessário fornecer via *prompt* uma série de informações relevantes do contexto fático e estratégica pretendida, a partir dessa dinâmica a inserção de dados em sistemas de IA generativa introduz novas camadas de risco, relacionadas não apenas à exposição direta da informação confidencial recebida, mas à possibilidade de circulação, processamento e reutilização da informação em contextos não plenamente conhecidos ou controlados pelo usuário.

Nesse cenário, a utilização dessas ferramentas passa a exigir reinterpretação dos deveres jurídicos tradicionais, especialmente no que se refere ao sigilo profissional, à diligência técnica e à governança informacional.

4.1 Inserção de dados e narrativas do cliente em sistemas de IA

A inserção de informações em sistemas de IA generativa deve ser compreendida, juridicamente, como um ato de externalização informacional qualificada. No contexto da

advocacia, essa externalização não se limita à transmissão de dados, mas envolve a transferência de conteúdo protegido por sigilo profissional para um ambiente tecnológico de terceiro.

Essa distinção é fundamental, pois diferentemente do compartilhamento de informações com auxiliares ou parceiros sob regime de confidencialidade, a inserção de dados em sistemas de IA não pressupõe, necessariamente, a existência de vínculo jurídico direto entre o advogado e o provedor da tecnologia que assegure níveis equivalentes de proteção. Trata-se, portanto, de uma forma de comunicação indireta com terceiro, mediada por infraestrutura tecnológica, cuja extensão e limites são frequentemente definidos por termos contratuais padronizados e de difícil controle efetivo pelo usuário.

Sob a perspectiva jurídica, esse fenômeno pode ser interpretado como ampliação do círculo de acesso à informação confidencial. Ainda que não haja divulgação pública ou acesso direto por terceiros humanos, o simples processamento da informação por sistemas externos pode ser suficiente para caracterizar sua exposição a ambiente não controlado.

A governança de dados contemporânea reforça essa compreensão ao destacar que a proteção informacional não se limita à restrição de acesso, mas envolve o controle sobre o processamento e a circulação dos dados, assim compreendido como a necessidade de rastreabilidade necessária e a responsabilidade dos agentes em cada etapa do ciclo de vida dos dados, conforme Guia Orientativo da Autoridade Nacional de Proteção de Dados (ANPD).

O processamento dos dados e sua circulação em ambiente informativo digital como ocorre na IA generativa é um aspecto fundamental, uma vez que a governança dos dados impor controle rígido e contínuo sobre como os dados são processados e como circulam, de forma que a inserção de informações confidenciais em sistemas de IA generativa representa uma ruptura com o modelo tradicional de custódia da informação, no qual o advogado mantém domínio sobre os meios e canais de comunicação utilizados.

Essa ruptura ganha especial relevância quando se considera a natureza fiduciária da relação advogado-cliente, pois a confiança depositada pelo cliente pressupõe não apenas a não divulgação das informações, mas também a sua preservação em ambientes seguros e controlados, de forma que a utilização de sistemas de IA generativa sem avaliação prévia das condições de tratamento dos dados pode, portanto, configurar violação do dever de diligência, ainda que não haja vazamento efetivo.

4.2 Riscos de vazamento, reutilização e reconstrução de informações

A compreensão dos riscos jurídicos associados à IA generativa exige superar a visão tradicional de vazamento de dados como evento pontual e identificável, vez que nos modelos de LLM, o risco informacional assume natureza mais difusa, decorrente da própria arquitetura tecnológica e das formas de processamento de dados.

A literatura técnica demonstra que modelos de LLM podem apresentar comportamentos de retenção parcial de informações, fenômeno associado à memorização de padrões específicos durante o treinamento ou à incorporação indireta de dados nas respostas geradas. Estudos realizados por pesquisadores de instituições como Google, Stanford, UC Berkeley e OpenAI (Carlini et al., 2021), evidenciam como que a IA generativa, apesar de sua construção como método probabilístico, em determinadas condições e com interações específicas, podem reproduzir sequências de dados , ou seja, poderia haver exposição de dados exatos, incluindo dados pessoais e dados confidenciais decorrentes do treinamento do modelo, de forma que a fronteira entre processamento e retenção não é rigorosamente delimitada.

Além disso, essa mesma pesquisa indica a possibilidade de extração ou reconstrução de dados a partir da interação com o modelo, por meio de técnicas que exploram padrões probabilísticos aprendidos pelo sistema. Esse tipo de vulnerabilidade desloca o risco jurídico do plano da divulgação direta para o plano da inferência e da reconstrução indireta.

Sob a perspectiva jurídica, esse cenário exige revisão do conceito de exposição de dados, pois a informação confidencial não precisa ser explicitamente divulgada para que haja risco juridicamente relevante, uma vez que a simples possibilidade de que ela possa ser inferida, reconstruída ou reutilizada já poderia comprometer a integridade do sigilo profissional.

Esse entendimento encontra respaldo na teoria da proteção da informação, análise do risco jurídico associado ao uso de sistemas de IA generativa exige o abandono de uma concepção restrita de confidencialidade, limitada à ideia de divulgação indevida, conforme sustenta Danilo Doneda (2014), o tratamento de dados pessoais, especialmente quando realizado por processos automatizados, constitui uma atividade intrinsecamente marcada pelo risco, na medida em que possibilita a organização, sistematização e cruzamento de informações que permitem reconstruir, com elevado grau de precisão, aspectos da vida pública e privada do indivíduo.

4.3 Casos de exposição de dados no uso da IA generativa

A análise dos riscos de exposição informacional não deve permanecer no plano abstrato.

A experiência global acumulada desde 2023 documenta casos concretos em que dados inseridos em sistemas de IA generativa abertos foram expostos, extraídos por meios maliciosos ou revelados por vulnerabilidades da própria infraestrutura dessas plataformas. Importa sublinhar, desde logo, que esses incidentes não constituem exceções: o risco associado ao uso de IA generativa não decorre de falhas pontuais ou eventos contingenciais, mas da própria estrutura e lógica de funcionamento dessas tecnologias. Trata-se de um risco inerente ao tratamento automatizado de dados, na medida em que a capacidade de organização, processamento e inferência amplia significativamente o potencial de exposição e uso indevido das informações. Pode-se, portanto, falar em risco estrutural, ou seja, aquele que emerge da arquitetura da IA generativa, e não de desvios ocasionais em sua utilização.

O primeiro episódio de relevância sistêmica foi documentado pela própria OpenAI em comunicado oficial de 24 de março de 2023, no qual noticiou-se uma falha em uma biblioteca de código aberto utilizada pelo ChatGPT permitiu que usuários visualizassem títulos do histórico de conversas de outros usuários ativos, e investigação posterior revelou que o mesmo defeito pode ter causado a exposição não intencional de informações de pagamento de 1,2% dos assinantes do ChatGPT Plus que estavam ativos durante uma janela específica de nove horas. Naquele período, tornava-se possível visualizar nome completo, e-mail, endereço de cobrança, tipo de cartão de crédito e os últimos quatro dígitos do número do cartão de outro usuário ativo.

O episódio demonstra que o risco de exposição não deriva apenas do comportamento do usuário, mas da fragilidade estrutural da própria infraestrutura do provedor, elemento completamente fora do controle de quem utiliza a ferramenta. Qualquer dado inserido por um advogado em sessão da plataforma, narrativas fáticas, nomes de clientes, estratégias processuais, estava sujeito ao mesmo mecanismo de exposição cruzada de sessões durante aquela janela, sem que o usuário pudesse saber ou prevenir.

O segundo caso paradigmático ocorreu em março de 2023 na divisão de semicondutores da Samsung Electronics. Em três incidentes separados, funcionários inseriram dados corporativos sensíveis no ChatGPT: o primeiro submeteu código-fonte defeituoso de um banco de dados interno em busca de solução para um erro; o segundo inseriu código de identificação de defeitos em equipamentos para otimização; e o terceiro converteu a gravação de uma reunião interna em arquivo de texto e a submeteu ao sistema para geração de ata. Após a ocorrência, a Samsung comunicou a seus funcionários que os dados inseridos eram impossíveis de recuperar, uma vez que passaram a estar armazenados nos servidores da OpenAI.

O aspecto juridicamente mais relevante não reside na má-fé dos envolvidos, que inexistiu, mas no fato de que a exposição decorreu de comportamento funcional e rotineiro no uso da ferramenta exatamente para o fim a que se destina. A transposição para a advocacia é imediata: um advogado que insere em sistema aberto narrativa fática de um processo, extrato de depoimento ou dados de saúde de cliente em demanda previdenciária pratica conduta estruturalmente idêntica, com a agravante de que a informação exposta é protegida por sigilo profissional de ordem pública.

O terceiro vetor documentado é o mais grave sob a perspectiva da segurança informacional, por não depender de falha acidental nem de comportamento descuidado do usuário. Pesquisadores do Google DeepMind, da Universidade de Washington, de Cornell, Carnegie Mellon, UC Berkeley e ETH Zurich (Nasr, 2023) demonstraram ser possível extrair vários megabytes de dados de treinamento do ChatGPT mediante consultas ao modelo no valor de aproximadamente duzentos dólares, utilizando um ataque de divergência consistente em solicitar ao sistema que repita indefinidamente uma determinada palavra. Os dados recuperados incluíam informações de identificação pessoal (endereços de e-mail, números de telefone e assinaturas digitais de indivíduos reais), além de trechos de código proprietário e passagens de literatura protegida. De forma, que o estudo concluiu que as técnicas de alinhamento com feedback humano, apresentadas pelos provedores como salvaguardas contra o regurgitamento de dados, não eliminam a memorização.

Além da extração via ataque direcionado, o risco engloba a reconstrução de dados por técnicas de inversão de embeddings: em sistemas que utilizam geração aumentada por recuperação, os dados fornecidos pelo usuário são convertidos em representações vetoriais e armazenados em bases externas, frequentemente geridas por terceiros, de forma que, uma vez obtido acesso a essas representações, é possível reconstruir, com grau variável de precisão, o conteúdo original dos documentos. Esse fenômeno evidencia que a exposição de informações pode ocorrer em camadas técnicas não imediatamente visíveis ao usuário, ampliando o escopo do risco informacional para além das formas tradicionais de divulgação.

A dimensão quantitativa do problema é confirmada por base empírica robusta. O "*Cloud and Threat Report 2026*", elaborado pela *Netskope Threat Labs* com telemetria de milhões de usuários corporativos, documenta que o número de incidentes de violação de políticas de dados em plataformas de IA generativa mais que dobrou no último ano, sendo que os três tipos de dados mais frequentemente inseridos em violação de políticas foram código-fonte (42%), dados

regulados (32%) e propriedade intelectual (16%), categoria que inclui expressamente contratos, documentos internos e pesquisas proprietárias. O relatório aponta ainda que 47% dos usuários de ferramentas de IA generativa no ambiente corporativo as utilizam por meio de contas pessoais e aplicações não gerenciadas pela organização, o que afasta qualquer controle institucional sobre os dados transmitidos.

Para a advocacia, esse dado tem relevância direta: o perfil de uso descrito por profissionais que submetem contratos, documentos internos e informações estratégicas a sistemas públicos de IA via contas pessoais, sem supervisão institucional, por exemplo, é precisamente o perfil de risco que caracteriza a conduta do advogado que utiliza plataformas abertas para auxiliar na prestação de serviços jurídicos.

No Brasil, não há, até a data de fechamento do presente artigo, decisões disciplinares publicadas pela OAB nem incidentes notificados à ANPD relacionados especificamente à exposição de dados de clientes por advogados mediante uso de IA generativa. Essa ausência, porém, não deve ser interpretada como segurança. O uso de IA em nuvem, sem governança, pode expor dados a riscos de vazamento, reuso indevido, transferência internacional irregular e treinamento indevido de modelos, e o advogado responde pessoalmente perante a OAB por violação de sigilo, independentemente de o dano ao cliente ter se materializado de forma perceptível. A ausência de precedentes disciplinares é sinal de invisibilidade, não de inexistência, e os casos internacionais já tornaram o risco empiricamente inegável.

5. PROTEÇÃO DE DADOS E GOVERNANÇA INFORMACIONAL NO USO DE IA NA ATIVIDADE JURÍDICA

A progressiva integração de sistemas de inteligência artificial, notadamente as de natureza generativa, no ecossistema da prática jurídica contemporânea impõe a reavaliação crítica dos paradigmas de governança informacional e proteção de dados.

A advocacia, atividade intrinsecamente dependente da manipulação de informações de alta sensibilidade, encontra-se em um ponto de inflexão. Ao mesmo tempo em que há a promessa de eficiência e otimização do trabalho, também existe o risco informacionais de magnitude inédita, que tensionam os deveres fundamentais do advogado e a estrutura normativa de proteção da privacidade.

Nesse contexto, a análise da aplicação da legislação de proteção de dados e da responsabilidade do profissional pela segurança das informações tratadas transcende o mero debate técnico, ascendendo a uma discussão sobre a própria sustentabilidade ética e jurídica da

profissão em face da inovação tecnológica.

A ausência de uma governança robusta e de uma compreensão aprofundada sobre as implicações do tratamento de dados por plataformas de terceiros pode converter a busca por eficiência em uma fonte de vulnerabilidade jurídica e reputacional, com consequências danosas para o cliente e para o próprio advogado.

5.1 Aplicação da Lei Geral de Proteção de Dados ao tratamento de informações em atividades jurídicas

A atividade jurídica, por sua natureza, constitui um campo fértil e sensível para o tratamento de dados pessoais.

O advogado, inevitavelmente coleta, processa, armazena e compartilha um vasto espectro de informações que permitem a identificação de pessoas naturais, configurando-se, para todos os efeitos, como um agente de tratamento de dados.

A incidência da Lei Geral de Proteção de Dados (LGPD) sobre o exercício da advocacia é, portanto, direta e inafastável.

A relação entre advogado e cliente, pautada na confiança e no sigilo, potencializa a importância dos princípios e das regras estabelecidas pelo diploma legal. A base legal para o tratamento de dados no contexto advocatício frequentemente se ampara na execução de contrato ou de procedimentos preliminares relacionados a contrato do qual seja parte o titular, a pedido do titular dos dados, bem como para o exercício regular de direitos em processo judicial, administrativo ou arbitral.

A utilização de sistemas de IA generativa para a elaboração de peças processuais, análise de documentos, pesquisa de jurisprudência ou gestão de casos não altera a natureza dessa relação jurídica, mas introduz uma nova complexidade.

Ao alimentar uma plataforma de IA com dados de um caso, o advogado está, em essência, realizando uma operação de tratamento que deve, obrigatoriamente, observar a LGPD.

Surge, assim, a necessidade de garantir ou reformular que o tratamento se restrinja ao mínimo necessário para o cumprimento de suas finalidades, com transparência perante o titular dos dados sobre como suas informações são utilizadas e, fundamentalmente, com a adoção de medidas de segurança aptas a proteger os dados de acessos não autorizados e de situações acidentais ou ilícitas de destruição, perda, alteração, comunicação ou difusão.

A responsabilidade primária por assegurar a conformidade dessas operações recai sobre

o advogado ou a sociedade de advogados, na qualidade de controladores dos dados.

A eventual alegação de que a tecnologia de IA opera como uma mera ferramenta, análoga a um software de edição de texto ou a um motor de busca convencional, revela-se juridicamente frágil.

A arquitetura de muitos sistemas de IA generativa, especialmente aqueles baseados em nuvem, implica o compartilhamento dos dados inseridos com o provedor da tecnologia, que pode utilizá-los para treinamento de seus algoritmos ou outros fins não diretamente relacionados à prestação do serviço ao advogado. Essa transferência de dados a um terceiro, que passa a figurar como operador ou até mesmo como um novo controlador, a depender da estrutura contratual, exige uma rigorosa análise.

Essa transferência de dados a um terceiro, que passa a figurar como operador ou até mesmo como um novo controlador, a depender da estrutura contratual, exige uma rigorosa análise, uma vez que configura na maioria dos casos, como transferência internacional de dados. A transferência internacional de dados, inerente ao uso de plataformas de IA generativa com infraestrutura majoritariamente localizada no exterior, submete-se às exigências dos artigos 33 a 36 da LGPD. Sua licitude depende da verificação de garantias específicas, como o reconhecimento de nível adequado de proteção pela ANPD, a adoção de cláusulas contratuais equivalentes ou o consentimento específico do titular. Nesse contexto, o advogado que insere dados de clientes em provedores estrangeiros sem observar tais requisitos realiza transferência irregular, sujeitando-se a sanções administrativas e à responsabilização civil, ainda que inexistente vazamento de dados.

Assim, o advogado tem o dever de compreender e validar a política de privacidade e os termos de serviço da ferramenta, assegurando que o tratamento subsequente dos dados pelo provedor seja compatível com as finalidades que justificaram a coleta original e que ofereça as garantias de segurança exigidas pela LGPD.

A complexidade se aprofunda quando o volume de dados inclui não apenas dados pessoais comuns, mas também dados pessoais sensíveis, cuja presença é recorrente em litígios de diversas naturezas, como em ações previdenciárias, trabalhistas ou de direito de família.

O tratamento de dados sensíveis submete-se a requisitos ainda mais estritos, demandando uma proteção qualificada. A introdução indiscriminada de tais informações em plataformas de IA cujas políticas de tratamento são insuficientes configura uma falha grave do agente de tratamento, passível de atrair a imposição das sanções administrativas previstas na

lei, sem prejuízo da responsabilidade civil por eventuais danos causados aos titulares.

A governança informacional na advocacia, portanto, deixa de ser um diferencial competitivo para se tornar um pressuposto de conformidade legal e de sustentabilidade da prática profissional na era digital.

5.2 Segurança da informação e dever de controle sobre o tratamento de dados por terceiros

O dever de segurança da informação, inerente à prática advocatícia e potencializado pela LGPD, impõe ao advogado a obrigação de adotar medidas técnicas e administrativas que garantam a integridade, a confidencialidade e a disponibilidade dos dados sob sua custódia.

Esse dever não se exaure nos limites físicos do escritório ou nos servidores locais. Na era da computação em nuvem e dos serviços de software como serviço, nos quais se inserem muitas ferramentas de IA generativa, o dever de segurança se expande, abrangendo a fiscalização e o controle sobre o tratamento de dados realizado por terceiros.

O advogado, na posição de controlador, responde não apenas por seus próprios atos, mas também pela escolha e pela supervisão dos operadores que contrata para auxiliá-lo no tratamento dos dados. Configura-se, aqui, uma responsabilidade decorrente da culpa in eligendo e da culpa in vigilando.

A seleção de uma plataforma de inteligência artificial não pode ser guiada unicamente por critérios de eficiência, custo ou popularidade. É imperativo que o advogado realize uma devida diligência para avaliar a maturidade da governança de dados do provedor do serviço.

O advogado deve analisar a robustez de sua política de segurança da informação, a existência de certificações de mercado, a localização dos servidores onde os dados serão armazenados, a clareza das cláusulas contratuais sobre o tratamento e a eventual reutilização dos dados inseridos, e a existência de mecanismos eficazes para o atendimento aos direitos dos titulares.

O dever de diligência impõe ao profissional o ônus de buscar a assessoria técnica necessária para tomar uma decisão informada ou, alternativamente, abster-se de utilizar tecnologias cujos riscos ele não é capaz de compreender e mitigar. Caso seja necessário, o advogado deve buscar assistência especializada antes de utilizar a inteligência artificial em seu escritório.

A responsabilidade do controlador é solidária em relação aos danos causados pelo operador quando este descumpre as obrigações da legislação de proteção de dados ou quando

não tiver seguido as instruções lícitas do controlador.

Portanto, a negligência na escolha ou na fiscalização de um provedor de IA que venha a causar um vazamento de dados ou outro tipo de dano aos titulares implicará a responsabilização direta do advogado, que terá, posteriormente, o direito de regresso contra o operador, se for o caso.

O dever de controle é, em última análise, uma manifestação do princípio da responsabilidade proativa, que exige que os agentes de tratamento sejam capazes de demonstrar a adoção de medidas eficazes e capazes de comprovar a observância e o cumprimento das normas de proteção de dados pessoais.

6. RESPONSABILIDADE CIVIL DO ADVOGADO PELO USO DE SISTEMAS DE IA GENERATIVA

A adoção de tecnologias de IA generativa na advocacia, ao mesmo tempo que inaugura um horizonte de possibilidades, introduz um novo paradigma de riscos que desafia as noções tradicionais de responsabilidade profissional.

A responsabilidade civil do advogado, historicamente associada a erros de direito processual ou material, expande seu escopo para abarcar uma nova classe de falhas: aquelas decorrentes da interação negligente com sistemas automatizados. A violação do sigilo profissional pela exposição de dados confidenciais a plataformas de IA, a falha no dever de diligência pela confiança acrítica nos resultados gerados por algoritmos e a consequente materialização de danos ao cliente formam um novo tripé de preocupações para a doutrina da responsabilidade civil.

A obrigação do advogado é de meio, e não de resultado, mas a utilização de ferramentas tecnológicas não pode servir de pretexto para a degradação dos meios empregados. Pelo contrário, a complexidade da ferramenta exige um plus de diligência, cuja inobservância configura a culpa, elemento central para a imputação da responsabilidade de indenizar.

6.1 Violação do dever de sigilo profissional e exposição indevida de informações confidenciais

O sigilo profissional é um pilar fundamental da advocacia, essencial para a garantia do direito de defesa e para a própria administração da justiça. Trata-se de um dever ético e jurídico, cuja proteção transcende o interesse meramente contratual entre advogado e cliente, ostentando natureza de ordem pública.

A utilização de sistemas de IA generativa, especialmente aqueles disponibilizados por

terceiros e operados em nuvem, cria uma tensão direta e perigosa com esse dever basilar. Ao inserir dados de um caso, narrativas fáticas, estratégias processuais ou documentos sigilosos em uma plataforma externa, o advogado está, objetivamente, comunicando informações confidenciais a um terceiro.

A natureza da tecnologia de IA generativa agrava o risco. Diferentemente de um software instalado localmente, cujo processamento ocorre no ambiente controlado pelo profissional, os modelos de linguagem mais avançados operam em servidores remotos. Os dados submetidos, denominados "*prompts*", podem ser incorporados pela IA, significando que as informações confidenciais do cliente passam a integrar o repositório de conhecimento da IA, podendo ser, em tese, acessadas, utilizadas ou até mesmo reproduzidas em respostas futuras para outros usuários.

Ainda que os provedores de tecnologia afirmem adotar medidas de anonimização ou de proteção, a simples transferência da informação para um ambiente fora do controle direto e absoluto do advogado já representa uma vulnerabilidade e, a rigor, uma quebra da cadeia de custódia da confidencialidade.

A alegação de que a transferência se dá para uma "máquina" e não para uma pessoa não descaracteriza a violação. O dever de sigilo não se limita a impedir a revelação a outros seres humanos: ele impõe a obrigação de manter a informação em um ambiente seguro e restrito, imune à exposição e ao acesso indevido, seja ele humano ou algorítmico. A infração ao dever de sigilo se consuma com a simples exposição indevida, independentemente da ocorrência de um dano subsequente. Trata-se de uma obrigação de não fazer, cujo descumprimento é instantâneo.

A conduta de submeter dados sigilosos a uma IA de terceiros sem a autorização expressa e informada do cliente e sem garantias contratuais e técnicas robustas de que tais dados não serão armazenados, aprendidos ou reutilizados configura, em si, um ato ilícito profissional e uma falha grave no dever de lealdade.

A responsabilidade do advogado, nesse cenário, pode ser aferida tanto na esfera disciplinar, perante a Ordem dos Advogados do Brasil, quanto na esfera cível.

A quebra do sigilo, ao frustrar a legítima expectativa de confiança do cliente, pode gerar danos de ordem moral, passíveis de compensação pecuniária, mesmo que nenhum prejuízo material imediato seja demonstrado.

Se a exposição indevida resultar, por exemplo, no vazamento de um segredo comercial, na antecipação de uma estratégia processual pela parte contrária ou em qualquer outra forma de

prejuízo patrimonial, a responsabilidade civil do advogado se torna ainda mais evidente, devendo ele reparar integralmente os danos causados por sua conduta imprudente. A conveniência ou a eficiência obtida pelo uso da ferramenta não são justificativas oponíveis à violação de um dever tão cardeal para a profissão.

6.2 Falha no dever de diligência na utilização de ferramentas tecnológicas

O dever de diligência constitui uma das obrigações centrais do mandato advocatício, exigindo que o profissional empregue todo o seu zelo, conhecimento técnico e prudência na condução dos interesses de seu cliente. A introdução de ferramentas de inteligência artificial no fluxo de trabalho não revoga nem mitiga esse dever, ao contrário, ela o requalifica, exigindo novas competências e uma postura de ceticismo metodológico.

A confiança cega e a aceitação acrítica de resultados gerados por sistemas de IA configuram uma manifestação contemporânea de negligência profissional, representando uma grave falha no dever de diligência. A tecnologia deve ser vista como um assistente ou uma ferramenta de apoio, e não como um substituto para o juízo crítico e a responsabilidade final do advogado.

A responsabilidade pela veracidade das informações, pela correção das teses jurídicas e pela adequação da estratégia processual permanece sendo, integral e indelegavelmente, do advogado. Um sistema de IA generativa, por sua própria arquitetura, não "pensa" nem "raciocina" no sentido humano; ele gera texto com base em padrões probabilísticos identificados em seu vasto banco de dados de treinamento. Isso o torna suscetível a produzir informações factualmente incorretas, citações de julgados inexistentes ou argumentos juridicamente equivocados, ainda que apresentados em um formato eloquente e persuasivo. A utilização de uma peça processual, de um parecer ou de uma linha de pesquisa gerada por IA, sem uma rigorosa checagem e validação por parte do profissional é um ato de manifesta imprudência.

A diligência no contexto tecnológico impõe ao advogado uma série de obrigações correlatas. Primeiramente, o dever de compreender, em um nível funcional, as capacidades e as limitações da ferramenta que utiliza.

Em segundo lugar, o dever de supervisionar ativamente o trabalho realizado pela IA, tratando cada "output" como um rascunho preliminar a ser auditado.

Em terceiro lugar, o dever de verificar todas as fontes primárias, como leis e decisões judiciais, citadas pelo sistema, para confirmar sua existência, sua vigência e sua aplicabilidade

ao caso concreto. A falha em qualquer uma dessas etapas configura a culpa profissional. O advogado não pode transferir para o algoritmo a responsabilidade que a lei e a ética lhe atribuem.

O argumento de que o erro foi produzido pela "máquina" é juridicamente irrelevante para eximir a responsabilidade do profissional perante seu cliente. A escolha de utilizar a ferramenta e a decisão de incorporar seu resultado no trabalho final são atos do advogado. A falha no dever de diligência, portanto, não reside no uso da tecnologia em si, mas em seu uso negligente, automatizado e desprovido da indispensável supervisão humana qualificada.

6.3 A configuração do dano e a responsabilidade profissional em contexto de risco informacional

A responsabilidade civil do advogado pelo uso inadequado de sistemas de IA generativa se materializa a partir da comprovação de três elementos clássicos: a conduta culposa, o dano, o nexo de causalidade entre a conduta e o dano.

A conduta culposa, como analisado, pode se manifestar como uma violação do dever de sigilo (imprudência na exposição de dados) ou como uma falha no dever de diligência (negligência na validação dos resultados da IA). O desafio, muitas vezes, reside na demonstração e na quantificação do dano e no estabelecimento de um nexo causal direto entre a falha do advogado e o prejuízo sofrido pelo cliente.

O dano pode assumir múltiplas formas. O dano material ou patrimonial é o mais evidente, podendo se configurar, por exemplo, pela perda de um prazo processual em razão de informação incorreta fornecida pela IA e não checada pelo advogado; pela sucumbência em uma demanda devido à fundamentação da petição em jurisprudência inexistente; ou pelos custos associados à remediação de um incidente de segurança causado pelo vazamento de dados confidenciais inseridos na plataforma.

Nestes casos, o prejuízo é diretamente mensurável e a obrigação de indenizar abrange tanto os danos emergentes (o que o cliente efetivamente perdeu) quanto os lucros cessantes (o que ele razoavelmente deixou de ganhar). A teoria da perda de uma chance também encontra campo fértil, sendo aplicável quando a conduta negligente do advogado, ainda que não seja a causa única do insucesso, retira do cliente a oportunidade séria e real de obter um resultado mais favorável.

Para além do dano material, o dano moral (extrapatrimonial) assume especial relevo no contexto de risco informacional. A simples quebra do sigilo profissional, com a exposição de informações sensíveis e privadas do cliente a uma plataforma de terceiros, por si só, viola

direitos da personalidade, como a intimidade e a privacidade, gerando um abalo psíquico e uma angústia que independem de qualquer repercussão patrimonial.

O nexo de causalidade, por sua vez, exige a demonstração de que o dano não teria ocorrido, ou não teria ocorrido da forma como ocorreu, não fosse a conduta culposa do advogado ao utilizar o sistema de IA. Um contra-argumento comum poderia ser a tentativa de imputar a culpa exclusivamente a terceiro (o provedor da IA) ou de alegar um caso fortuito (uma falha imprevisível do sistema).

De toda forma, a responsabilidade do advogado perante o cliente é contratual e direta. A falha do provedor pode, no máximo, gerar um direito de regresso para o advogado, mas não rompe o nexo causal primário entre a sua negligência (na escolha, na supervisão ou na validação) e o dano sofrido pelo cliente. O risco informacional associado ao uso de IA não é um evento imprevisível, mas sim um risco inerente à atividade, que o profissional assume e tem o dever de gerenciar.

O respaldo normativo para essa responsabilização encontra-se consolidado na confluência de três diplomas legais. No plano geral da responsabilidade civil, o artigo 186 do Código Civil estabelece que aquele que, por ação ou omissão voluntária, negligência ou imprudência, causar dano a outrem comete ato ilícito, ao passo que o artigo 927 impõe a obrigação de reparar o dano a quem o causar. O artigo 951 do mesmo Código, por sua vez, estabelece responsabilidade específica para os profissionais que causem dano ao cliente no exercício de atividade técnica, por imperícia, negligência ou imprudência, dispositivo diretamente aplicável ao advogado que adota ferramenta tecnológica sem a competência digital mínima necessária ao seu uso seguro.

No plano estatutário, o artigo 32 do Estatuto da Advocacia prescreve que o advogado é responsável pelos atos que, no exercício profissional, praticar com dolo ou culpa, respondendo pelos danos causados a seus clientes. A combinação desses fundamentos revela que a responsabilidade civil do advogado pelo uso negligente de sistemas de inteligência artificial generativa não depende de construção doutrinária inovadora: ela decorre da aplicação direta e combinada do direito civil comum com o estatuto profissional, bastando a demonstração da conduta culposa, do dano ao cliente e do nexo causal entre a adoção irresponsável da tecnologia e o prejuízo verificado.

A responsabilidade profissional, portanto, se adapta para abranger os novos riscos criados pela tecnologia, reafirmando que, no centro da atividade jurídica, a prudência, o sigilo e

a diligência continuam sendo obrigações indelegáveis do ser humano que exerce a advocacia.

7. CONSIDERAÇÕES FINAIS

O uso de IA generativa na advocacia não é uma questão meramente operacional ou de produtividade, mas um problema jurídico de primeira ordem, que reposiciona o profissional do direito diante de obrigações éticas e legais que o ordenamento brasileiro já contempla, mas que ainda carecem de interpretação sistemática para o ambiente digital.

A premissa central que emerge da investigação é a de que a ignorância tecnológica não é uma escusa juridicamente admissível. O advogado que utiliza sistemas de IA generativa sem compreender minimamente sua arquitetura de funcionamento, a forma como os dados são processados, retidos, potencialmente reutilizados e submetidos a regimes de governança alheios ao seu controle, age em flagrante violação ao dever de diligência que lhe é exigido pelo Estatuto da OAB e pelo Código de Ética Profissional. A complexidade da ferramenta, longe de funcionar como atenuante, opera como fator de agravamento da culpa: quanto mais sofisticada a tecnologia empregada, maior o ônus de compreensão que recai sobre aquele que a adota no exercício de função de confiança pública.

Nesse sentido, a leitura criteriosa dos termos de uso e das políticas de privacidade das plataformas de IA generativa deixa de ser um comportamento recomendável para se tornar uma obrigação profissional inafastável. É nesse documento - frequentemente ignorado pela esmagadora maioria dos usuários - que reside a definição concreta do regime de tratamento de dados ao qual o advogado submete, sem perceber, as informações confidenciais de seus clientes: se os dados poderão ser utilizados para aprimoramento do modelo, por quanto tempo serão retidos, em que jurisdição serão armazenados e quais mecanismos de segurança efetivamente protegem o conteúdo transmitido. A ausência dessa leitura não isenta o profissional de responsabilidade, ao contrário, ela configura, por si mesma, a negligência que pode fundamentar tanto a sanção disciplinar perante a OAB quanto à responsabilidade civil perante o cliente lesado.

A dimensão filosófica do problema não pode ser negligenciada, como bem observado pelo filósofo Luciano Floridi (2013) ao desenvolver sua teoria da ética da informação, sustenta que os agentes que operam em ambientes informacionais carregam responsabilidades proporcionais ao poder que exercem sobre o fluxo e a integridade dos dados, e que a degradação do ambiente informacional constitui, em si, um dano moral autônomo, independente de

consequências patrimoniais imediatas. Essa perspectiva ressoa diretamente com a situação do advogado que insere dados sensíveis de clientes em plataformas de terceiros: a simples transferência da informação para um ambiente não controlado já representa uma ruptura com a integridade do ecossistema informacional protegido pelo sigilo profissional, ainda que nenhum vazamento explícito venha a se materializar.

Na mesma linha, o filósofo belga Mark Coeckelbergh (2020) ao examinar as implicações éticas da delegação de tarefas a sistemas automatizados, alerta que a mediação tecnológica não dissolve a responsabilidade moral do agente humano, ela a transforma, exigindo novas formas de competência e vigilância. Para o advogado, isso significa que delegar à IA a produção de peças processuais, a análise de documentos ou a gestão de informações estratégicas não transfere ao algoritmo qualquer parcela da responsabilidade profissional, afinal, o advogado permanece o único responsável perante seu cliente e perante a ordem jurídica por todos os atos que realiza com ou sem auxílio da tecnologia.

O chamado à governança informacional, portanto, não é retórico, ele se traduz em medidas concretas e exigíveis, como a adoção de políticas internas de uso seguro de IA nos escritórios de advocacia, a preferência por planos corporativos nos modelos escolhidos pelos profissionais, modelos seguros dedicados a atividade jurídica e APIs empresariais que ofereçam garantias contratuais explícitas de não utilização dos dados para treinamento, além da obtenção de consentimento informado do cliente antes de qualquer externalização de suas informações para plataformas de terceiros e o investimento contínuo em capacitação técnica relevante dos profissionais que operam essas ferramentas, sendo a governança uma condição de sua legitimidade.

De forma que o século XXI impõe ao advogado uma dupla competência: a jurídica, que sempre lhe foi exigida, e a digital, que agora se torna igualmente indispensável. Não se trata de transformar o jurista em engenheiro de sistemas, mas de assegurar que o profissional que detém a confiança do cliente e o múnus público da advocacia compreenda as implicações reais das ferramentas que escolhe utilizar.

A tecnologia avançará inevitavelmente, e deve avançar. O que não pode retroceder é o compromisso ético e jurídico com a proteção daquele que, ao buscar a tutela de seus direitos, deposita no advogado não apenas um mandato, mas sua confiança mais íntima.

REFERÊNCIAS

AUTORIDADE NACIONAL DE PROTEÇÃO DE DADOS (ANPD). ***Tratamento de dados pessoais pelo Poder Público: guia orientativo***. Versão 2.0. Brasília: ANPD, 2023.

ARBEX, Sergei C.; ZAKKA, Rogério M. ***Estatuto da advocacia: prerrogativas e ética***. Barueri: Manole, 2012.

BRASIL. ***Constituição da República Federativa do Brasil de 1988***. Brasília: Presidência da República, 1988.

BRASIL. ***Lei nº 10.406, de 10 de janeiro de 2002***. Institui o Código Civil. Brasília: Presidência da República, 2002.

BRASIL. ***Lei nº 13.709, de 14 de agosto de 2018***. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília: Presidência da República, 2018.

BRASIL. ***Lei nº 8.906, de 4 de julho de 1994***. Dispõe sobre o Estatuto da Advocacia e a Ordem dos Advogados do Brasil (OAB)

CARLINI, Nicholas et al. ***Extracting training data from large language models***. Proceedings of the 30th USENIX Security Symposium, Berkeley (CA), v. 30, n. 1, p. 2633–2650, ago. 2021.

CAVALIERI FILHO, Sergio. ***Programa de responsabilidade civil***. 16. ed. São Paulo: Atlas, 2023.

COECKELBERGH, Mark. ***AI Ethics***. Cambridge: The MIT Press, 2020.

CONSELHO NACIONAL DE JUSTIÇA (CNJ). ***Resolução nº 615, de 11 de março de 2025. Estabelece diretrizes para o desenvolvimento, utilização e governança de soluções desenvolvidas com recursos de inteligência artificial no Poder Judiciário***. Brasília: CNJ, 2025.

DONEDA, Danilo. ***A proteção de dados pessoais como um direito fundamental***. Revista de Direito do Consumidor, São Paulo, v.9, p. 127-151, 2014.

FLORIDI, Luciano. ***The Ethics of Information***. Oxford: Oxford University Press, 2013.

LOBO, Paulo. ***Comentários ao Estatuto da Advocacia e da OAB***. 15. ed. São Paulo: Saraiva Jur, 2023.

MALDONADO, Viviane Nóbrega; BLUM, Renato Opice (Coord.). ***LGPD: Lei Geral de Proteção de Dados comentada***. 4. ed. São Paulo: Thomson Reuters Brasil, 2022.

NASR, Milad et al. ***Scalable Extraction of Training Data from (Production) Language Models***. arXiv preprint arXiv:2311.17035, 28 nov. 2023.

NETSKOPE THREAT LABS. ***Cloud and Threat Report: 2026***. Santa Clara: Netskope, jan. 2026. Disponível em: https://www.netskope.com/resources/cloud-and-threat-reports/cloud-and-threat-report-2026. Acesso em: 13 abr. 2026.

ORDEM DOS ADVOGADOS DO BRASIL. ***Código de Ética e Disciplina da OAB***. Brasília: Conselho Federal da OAB, 2015.

ORDEM DOS ADVOGADOS DO BRASIL. ***Conselho Federal. Recomendação n. 001/2024: diretrizes para o uso ético de inteligência artificial generativa na advocacia***. Brasília: OAB, 11 nov. 2024. Disponível em: https://diario.oab.org.br/pages/materia/842347. Acesso em: 13 abr.2026.

OPENAI. ***March 20 ChatGPT outage: here's what happened***. Blog OpenAI, São Francisco, 24 mar. 2023. Disponível em: https://openai.com/index/march-20-chatgpt-outage. Acesso em: 13 abr. 2026.

PALMER, Danny. ***Samsung workers made a major error by using ChatGPT***. TechRadar, s.l., 6 abr. 2023. Disponível em: https://www.techradar.com/news/samsung-workers-leaked-company-secrets-by-using-chatgpt. Acesso em: 13 abr. 2026.

PALMER, Danny. ***Personal LLM accounts drive shadow AI data leak risks***. Infosecurity Magazine, s.l., 7 jan. 2026. Disponível em: https://www.infosecurity-magazine.com/news/personal-llm-accounts-drive-shadow/. Acesso em: 13 abr. 2026.

RUSSELL, S; NORVIG, P. ***Artificial Intelligence: A Modern Approach***. 3. ed. Harlow: Pearson, 2013.

TEFFÉ, Chiara Spadaccini de; MORAES, Maria Celina Bodin de (Coord.). ***Lei Geral de Proteção de Dados Pessoais (LGPD): desafios e perspectivas***. Rio de Janeiro: GZ, 2021.

VASWANI, Ashish; SHAZEER, Noam; PARMAR, Niki; USZKOREIT, Jakob; JONES, Llion; GOMEZ, Aidan N.; KAISER, Łukasz; POLOSUKHIN, Illia. ***Attention is all you need***. Advances in Neural Information Processing Systems, v. 30, p. 5998-6008, 2017.

VOLPE, Michele Cristine Soares Campos. ***Direito Autoral e aprendizagem de IA generativa: análise comparada Brasil e Estados Unidos no uso de obras protegidas para treinamento de IA***. Revista DCS, [S. l.], v. 22, n. 84, p. e3880, 2025.