



UNIVERSIDADE DE BRASÍLIA – UNB

Faculdade de Tecnologia

Departamento de Engenharia Elétrica

Programa de Pós-Graduação Profissional em Engenharia Elétrica (PPEE)

SEGURANÇA E PRIVACIDADE DE DADOS EM TEXTOS DE CIRCULAÇÃO RESTRITA

Paulo Florêncio de Barros¹

Curso de Pós-Graduação Lato Sensu em Privacidade e Segurança da
Informação

Professor Orientador
Prof. Dr. William Ferreira Giozza

Brasília – DF

¹ Pós-graduado em Direito Público pela Universidade Anhanguera (UNIDERP), e-mail: paulo.florencio@gmail.com – WhatsApp: (61) 98415-8258.

RESUMO

Este artigo analisa criticamente as interações entre segurança da informação, privacidade de dados e o emprego de sistemas de inteligência artificial generativa (IA) no processamento de documentos de circulação restrita em organizações públicas e privadas. Adota-se uma abordagem qualitativa, fundamentada em revisão bibliográfica crítica e análise documental comparativa de marcos normativos nacionais e internacionais: a Lei Geral de Proteção de Dados Pessoais, ou LGPD (Lei nº 13.709/2018), o Regulamento Europeu de Inteligência Artificial, ou *AI Act* (Regulamento (UE) 2024/1689) e, no plano setorial brasileiro, a Resolução CNJ nº 615/2025, aplicável especificamente ao Poder Judiciário. No contexto regulatório dos Estados Unidos, observa-se a substituição, em janeiro de 2025, da Ordem Executiva 14110 pela Ordem Executiva 14179, de caráter não-regulatório, o que evidencia a instabilidade do arcabouço normativo norte-americano sobre o tema. A presente análise identifica lacunas regulatórias e operacionais relacionadas à proteção de informações sensíveis, especialmente quanto a riscos de exfiltração, retenção não autorizada de dados, ausência de rastreabilidade e vulnerabilidades inerentes a Modelos de Linguagem de Grande Porte – em inglês, *Large Language Models* (LLMs). Para ilustrar essas vulnerabilidades, são examinados três cenários representativos: a elaboração de memorandos sigilosos por LLMs, a síntese automatizada de atas de reuniões restritas e a tradução de documentos classificados por ferramentas externas. Os resultados indicam que, embora os marcos regulatórios atuais representem avanços relevantes na normatização do desenvolvimento, da disponibilização e do uso geral de sistemas de inteligência artificial, permanecem insuficientes para estabelecer controles operacionais específicos voltados ao uso interno de IA sobre documentos de circulação restrita. Como contribuição, o estudo propõe diretrizes de governança alinhadas a referenciais internacionais de gestão de risco em IA, como o NIST AI RMF, com vistas a mitigar os riscos identificados em ambientes organizacionais sensíveis.

Palavras-chave: Privacidade de dados. Segurança da informação. Inteligência artificial generativa. Documentos de circulação restrita. Governança de IA. LGPD. AI Act. Modelo de linguagem de grande porte (LLM).

ABSTRACT

This article critically analyzes the interactions between information security, data privacy, and the employment of generative artificial intelligence (AI) systems in the processing of restricted-circulation documents in public and private organizations. The research adopts a qualitative approach, founded on a critical bibliographic review and a comparative documentary analysis of national and international normative frameworks: the General Personal Data Protection Law, or LGPD (Law No. 13,709/2018), the European Artificial Intelligence Regulation, or AI Act (Regulation (EU) 2024/1689), and, at the Brazilian sectoral level, CNJ Resolution No. 615/2025, applicable specifically to the Judiciary. In the regulatory context of the United States, the replacement, in January 2025, of Executive Order 14110 by Executive Order 14179, of a non-regulatory character, is observed, which evidences the instability of the North American normative framework on the subject. The present analysis identifies regulatory and operational gaps related to the protection of sensitive information, especially regarding risks of exfiltration, unauthorized data retention, absence of traceability, and vulnerabilities inherent to Large Language Models (LLMs). To illustrate these vulnerabilities, three representative scenarios are examined: the elaboration of confidential memorandums by LLMs, the automated synthesis of restricted meeting minutes, and the translation of classified documents by external tools. The results indicate that, although current regulatory frameworks represent relevant advances in the standardization of the development, provision, and general use of artificial intelligence systems, they remain insufficient to establish specific operational controls aimed at the internal use of AI on restricted-circulation documents. As a contribution, the study proposes governance guidelines aligned with international AI risk management frameworks, such as the NIST AI RMF, with a view to mitigating the identified risks in sensitive organizational environments.

Keywords: Data privacy. Information security. Generative artificial intelligence. Restricted-circulation documents. AI governance. LGPD. AI Act. Large language model (LLM)..

1. INTRODUÇÃO

A adoção de ferramentas de inteligência artificial generativa no ambiente corporativo e governamental avança em um ritmo tal que supera a capacidade de resposta dos marcos regulatórios existentes, reconfigurando processos de trabalho em setores públicos e privados. Se por um lado a automação de tarefas como redação, revisão e síntese de documentos representa ganhos expressivos de eficiência, por outro, introduz vetores de risco, antes inexistentes, no ciclo de vida de informações classificadas ou de circulação restrita.

O problema central abordado neste trabalho reside em uma contradição estrutural: as organizações submetidas a obrigações legais de sigilo — quer sejam impostas pela Lei n.º 12.527/2011 (BRASIL, 2011), denominada de Lei de Acesso à Informação (LAI), quer sejam decorrentes de regimentos internos ou de normas setoriais de conformidade (*compliance*) — passaram a utilizar ferramentas de Inteligência Artificial (IA) cujo funcionamento técnico implica, em graus variados, o envio de dados a servidores externos, a retenção de conteúdo para fins de treinamento ou suporte, e a ausência de auditabilidade sobre o destino das informações inseridas como *input*.

Agrava esse cenário a ausência, em muitas organizações, de critérios claros para distinguir ferramentas de IA generativa de uso público daquelas desenvolvidas para ambientes corporativos ou governamentais, configuradas para atender a requisitos específicos de governança, controle de dados e segurança da informação. Na prática, usuários e até mesmo organizações podem não perceber claramente essa distinção, uma vez que versões públicas e empresariais frequentemente apresentam interfaces e funcionalidades semelhantes, embora adotem modelos distintos de tratamento, retenção e proteção das informações processadas.

Essa diferenciação não é exclusividade de um fornecedor específico; é observada em diversas plataformas de IA generativa que disponibilizam versões destinadas ao público em geral e, ao mesmo tempo, modalidades empresariais dotadas de mecanismos específicos de governança, conformidade e proteção de dados.

As soluções disponibilizadas pela Microsoft, por exemplo, ilustram essa distinção. Enquanto o Microsoft 365 Copilot foi concebido para operar em ambientes organizacionais sob mecanismos de identidade, conformidade e proteção de dados do ecossistema Microsoft 365 — incluindo o *Enterprise Data Protection* (EDP), que assegura que os comandos (*prompts*) e as respostas dos usuários não sejam utilizados para treinamento dos modelos fundamentais e que respeita os controles de acesso, auditoria, retenção e rotulagem de sensibilidade da organização —, o Microsoft Copilot destinado ao público em geral está sujeito a políticas de tratamento de dados distintas, evidenciando que ferramentas com interfaces semelhantes podem oferecer níveis significativamente diferentes de governança e proteção das informações processadas (MICROSOFT, 2026a; MICROSOFT, 2026b).

Este trabalho parte da hipótese de que os principais marcos normativos sobre IA e privacidade — no Brasil, na União Europeia e nos Estados Unidos — foram concebidos para regular a

inteligência artificial como produto ou serviço colocado no mercado, e não como ferramenta de processamento interno sobre documentos restritos. Essa distinção não é meramente técnica: ela demonstra a ausência de obrigações específicas de confidencialidade, rastreabilidade e controle de acesso aplicáveis ao uso cotidiano dos modelos de linguagem de grande porte (*Large Language Model* - LLM) em fluxos documentais sensíveis.

No Brasil, essa lacuna se manifesta de forma distinta da observada na União Europeia: uma vez que a competência para legislar sobre proteção de dados e diretrizes gerais de tecnologia é predominantemente federal, o risco central não é a fragmentação horizontal entre entes federativos, mas a fragmentação vertical entre normas gerais como a Lei Geral de Proteção de Dados (LGPD - BRASIL, 2018), normas setoriais como a Resolução CNJ n.º 615/2025 (CNJ, 2025), restrita ao Poder Judiciário, e a ausência de uma política nacional específica para o uso interno de IA sobre documentos de circulação restrita.

A metodologia adotada neste trabalho é qualitativa, mediante combinação da análise documental dos marcos normativos, revisão bibliográfica em direito digital, segurança da informação e governança de IA, e a construção de cenários ilustrativos que materializam os riscos identificados. O recorte temporal privilegia o período 2022–2025, de maior densidade normativa nas três jurisdições analisadas, e compreende desde a publicação da LGPD e seus desdobramentos regulatórios até a consolidação do *AI Act* e da Resolução CNJ n.º 615/2025.

O trabalho está estruturado em seis seções: a Seção 1 com esta introdução; as Seções 2 e 3 apresentam, respectivamente, o quadro conceitual e o panorama normativo comparado; a Seção 4 analisa as lacunas identificadas; a Seção 5 descreve os cenários ilustrativos; e a Seção 6 apresenta as conclusões e recomendações.

2. CONCEITOS FUNDAMENTAIS

A compreensão dos riscos associados ao emprego de sistemas de inteligência artificial generativa no processamento de documentos de circulação restrita exige, inicialmente, a delimitação dos conceitos que fundamentam este estudo.

A caracterização de um documento como sendo de circulação restrita não decorre exclusivamente do suporte documental ou do conteúdo isoladamente considerado, mas também da combinação entre a sensibilidade das informações nele contidas, os impactos potenciais decorrentes de sua divulgação indevida e o conjunto de normas jurídicas, regulatórias ou institucionais aplicáveis ao seu acesso, processamento e compartilhamento.

Igualmente, a restrição de circulação não constitui uma característica intrínseca do documento em si, mas uma condição atribuída pela organização responsável pela informação e que considera fatores como relevância estratégica, obrigações legais, proteção de dados pessoais, compromissos contratuais ou requisitos de segurança da informação.

Esta Seção apresenta os conceitos fundamentais relacionados aos documentos de circulação restrita, à inteligência artificial generativa, aos modelos de linguagem de grande porte (LLMs) e às distinções entre privacidade, confidencialidade, integridade e sigilo, a fim de estabelecer a base conceitual necessária para a análise normativa e técnica que será desenvolvida nas seções posteriores.

2.1 Documentos de circulação restrita e sua proteção jurídica

Entende-se por documento de circulação restrita todo aquele cujo acesso, processamento, compartilhamento ou divulgação seja limitado a um conjunto determinado de pessoas ou entidades em razão da sensibilidade das informações nele contidas, independentemente do suporte (físico ou digital), formato ou meio de armazenamento.

Essa restrição pode decorrer de diferentes fundamentos: obrigações legais, regulamentações setoriais, contratos, políticas internas de governança da informação ou critérios institucionais de classificação e controle de acesso.

No âmbito da Administração Pública brasileira, a classificação formal de informações sigilosas observa os graus previstos no artigo 24 da Lei nº 12.527/2011, ou Lei de Acesso à Informação (LAI) (BRASIL, 2011):

- reservado, com prazo máximo de restrição de acesso de 5 (cinco) anos;
- secreto, com prazo máximo de 15 (quinze) anos;
- ultrassecreto, com prazo máximo de 25 (vinte e cinco) anos, admitida uma única prorrogação por igual período.

É importante destacar que a classificação formal prevista na LAI se refere às informações protegidas por interesse público ou estatal e não abrange, automaticamente, todos os documentos de circulação restrita existentes em organizações públicas ou privadas.

No setor privado, a restrição de circulação decorre predominantemente de obrigações contratuais, regulamentações específicas, requisitos de governança corporativa e políticas internas de classificação da informação. Instituições financeiras e organizações reguladas, por exemplo, estão sujeitas a normas específicas relacionadas à segurança cibernética, gestão de riscos tecnológicos e proteção de informações sensíveis.

A proteção jurídica e técnica desses documentos ocorre por meio de múltiplas camadas complementares. A LAI estabelece critérios para classificação e acesso a informações sigilosas no âmbito da Administração Pública; a Lei Geral de Proteção de Dados Pessoais, ou LGPD (Lei nº 13.709/2018) é aplicada quando houver tratamento de dados pessoais; o Decreto nº 9.637/2018, que institui a Política Nacional de Segurança da Informação, estabelece princípios e diretrizes de governança para proteção das informações no âmbito federal; normas setoriais e padrões internacionais de segurança da informação complementam os mecanismos de proteção.

Sob a perspectiva deste trabalho, documentos de circulação restrita devem ser compreendidos como unidades informacionais submetidas simultaneamente a requisitos de:

- controle de acesso;
- confidencialidade;
- integridade;
- rastreabilidade;
- proteção de dados pessoais;
- governança durante o ciclo de vida da informação.

Essa abordagem amplia a visão tradicional baseada exclusivamente em classificação jurídica e, portanto, permite analisar os impactos decorrentes da utilização de sistemas de inteligência artificial generativa sobre informações sensíveis.

2.2 Inteligência artificial generativa e modelos de linguagem de grande porte

A inteligência artificial generativa corresponde ao conjunto de técnicas capazes de produzir novos conteúdos – incluindo textos, imagens, códigos, áudios e outros formatos digitais – a partir de padrões aprendidos durante o processo de treinamento dos modelos.

Entre essas tecnologias, destacam-se os Modelos de Linguagem de Grande Porte (*Large Language Models* – LLMs), sistemas treinados sobre extensos conjuntos de dados textuais capazes de executar tarefas como geração, resumo, tradução, classificação e revisão de documentos.

A arquitetura baseada em transformadores (*Transformers*), apresentada por Vaswani et al. (2017), representou um avanço significativo para esses modelos ao permitir maior capacidade de compreender relações contextuais e produzir respostas coerentes a partir de instruções em linguagem natural, ou comandos (*prompts*).

Apesar dos benefícios relacionados à automação e produtividade, a utilização de LLMs sobre documentos de circulação restrita introduz novos desafios de segurança da informação. Entre os principais riscos, destacam-se:

a) Risco de exfiltração

Ocorre quando informações inseridas nos *prompts* ou documentos processados são transmitidas para ambientes externos controlados pelo fornecedor da solução de IA e pode envolver mecanismos de armazenamento, processamento ou retenção conforme a arquitetura adotada.

b) Risco de inferência

Relaciona-se à possibilidade de obtenção indireta de informações sensíveis por meio dos resultados produzidos pelo modelo, especialmente quando padrões estatísticos ou informações contextuais permitem reconstruir elementos presentes nos dados originais.

c) Risco de rastreabilidade

Surge quando inexistem mecanismos adequados de registro, auditoria e monitoramento capazes de comprovar quais informações foram processadas, por quais usuários, em quais sistemas e sob quais condições.

d) Risco à integridade da informação

Decorre da possibilidade de geração ou alteração de conteúdos incorretos, incompletos ou não verificáveis, característica associada às limitações atuais dos modelos generativos, como alucinações (*hallucinations*) e produção de informações que, embora semanticamente plausíveis, são factualmente incorretas.

Nesse contexto de risco, também é fundamental distinguir as diferentes arquiteturas de implantação. Modelos disponibilizados em ambientes de nuvem normalmente envolvem processamento externo ou compartilhamento controlado de informações com o provedor da tecnologia. Por outro lado, modelos implantados localmente (*on-premise*) ou em ambientes privados controlados pela organização podem reduzir significativamente o risco de exfiltração para terceiros.

Entretanto, essa arquitetura não elimina outros riscos, tais como: ameaças internas, comprometimento da infraestrutura tecnológica, privilégios excessivos de acesso, falhas operacionais e ataques direcionados.

A avaliação desses riscos deve considerar referenciais contemporâneos de governança e segurança de inteligência artificial, como o NIST *AI Risk Management Framework* (AI RMF) (TABASSI, 2023), a ISO/IEC 42001:2023 (ISO, 2023a), a ISO/IEC 23894:2023 (ISO, 2023b) e o OWASP *Top 10 for Large Language Model Applications* (OWASP, 2023), que destacam desafios relacionados à segurança, confiabilidade, transparência e controle de sistemas baseados em IA.

2.3 Privacidade, confidencialidade, integridade e sigilo: distinções operacionais

Embora frequentemente utilizados como conceitos próximos, privacidade, confidencialidade, integridade e sigilo representam dimensões distintas da proteção da informação e possuem aplicações específicas no contexto de sistemas de inteligência artificial.

A privacidade refere-se ao direito do titular de controlar o tratamento de seus dados pessoais, encontrando fundamento constitucional nos incisos X e XII do artigo 5º da Constituição Federal

e sendo regulamentada pela Lei Geral de Proteção de Dados Pessoais, ou LGPD (Lei nº 13.709/2018).

A confidencialidade corresponde ao princípio segundo o qual informações devem ser acessíveis somente por indivíduos, processos ou entidades devidamente autorizados, constituindo um dos pilares tradicionais da segurança da informação.

A integridade representa a garantia de que a informação permaneça correta, completa e protegida contra alterações indevidas durante todo o seu ciclo de vida.

O sigilo, especialmente no contexto da Administração Pública, refere-se à proteção jurídica atribuída a determinadas informações cujo acesso é formalmente restringido em razão de interesses públicos, estratégicos ou institucionais, conforme estabelecido pela Lei de Acesso à Informação (LAI).

O processamento de documentos de circulação restrita por sistemas de inteligência artificial generativa pode afetar simultaneamente essas quatro dimensões.

Pode ocorrer violação de privacidade quando dados pessoais são tratados sem fundamento jurídico adequado; violação de confidencialidade quando informações são compartilhadas com terceiros não autorizados; comprometimento da integridade quando conteúdos são alterados ou gerados sem validação adequada; e violação de sigilo quando informações formalmente protegidas são submetidas a sistemas não homologados.

A análise normativa desenvolvida neste estudo utiliza essa perspectiva integrada para avaliar em que medida os atuais instrumentos regulatórios oferecem mecanismos suficientes para disciplinar o uso interno de inteligência artificial generativa em documentos de circulação restrita.

3. MARCOS NORMATIVOS E REGULATÓRIOS SOBRE INTELIGÊNCIA ARTIFICIAL, PRIVACIDADE E GOVERNANÇA DE DADOS

A regulação da inteligência artificial encontra-se em processo de consolidação em diferentes jurisdições e apresenta modelos distintos conforme características jurídicas, institucionais e econômicas de cada país ou bloco regional. Enquanto algumas iniciativas adotam uma Abordagem Baseada em Risco (ABR), outras priorizam mecanismos voluntários de governança ou regulamentos setoriais específicos.

Esta Seção apresenta uma análise comparativa dos principais marcos normativos relacionados à inteligência artificial, proteção de dados pessoais e segurança da informação no Brasil, na União Europeia, nos Estados Unidos e na Austrália, a fim de identificar em que medida esses instrumentos oferecem mecanismos adequados para disciplinar o uso interno de sistemas de IA generativa no processamento de documentos de circulação restrita.

A análise considera quatro dimensões principais:

- a natureza jurídica do instrumento regulatório;
- as obrigações impostas aos desenvolvedores, fornecedores e usuários de sistemas de IA;
- os mecanismos de proteção de dados, segurança e rastreabilidade;
- a existência — ou ausência — de requisitos específicos para governança de documentos sensíveis processados por sistemas de inteligência artificial.

3.1 Brasil: LGPD, Projeto de Lei (PL) nº 2.338/2023 e a Resolução CNJ nº 615/2025

No Brasil, o tratamento jurídico da inteligência artificial encontra-se em fase de transição entre normas gerais de proteção de dados, iniciativas legislativas específicas e regulamentações setoriais. Diferentemente da União Europeia, que estruturou um marco regulatório abrangente para IA, o modelo brasileiro ainda combina instrumentos jurídicos distintos, com níveis variados de abrangência e obrigatoriedade.

A Lei Geral de Proteção de Dados Pessoais, ou LGPD (Lei nº 13.709/2018), constitui o principal instrumento normativo relacionado à proteção da privacidade e ao tratamento de dados pessoais no país. A LGPD estabelece princípios como finalidade, adequação, necessidade, transparência, segurança e responsabilização, impondo aos controladores e operadores obrigações relacionadas à legitimidade do tratamento, proteção contra acessos não autorizados e adoção de medidas técnicas e administrativas adequadas.

Embora a LGPD seja aplicável ao uso de sistemas de inteligência artificial quando houver tratamento de dados pessoais, seu texto original não foi estruturado especificamente para lidar com características próprias da IA generativa, como modelos probabilísticos, treinamento baseado em grandes volumes de dados, geração automatizada de conteúdo e riscos associados ao processamento de informações inseridas como *inputs*.

Essa ausência de disciplina específica gera desafios interpretativos quando o objeto tratado por sistemas de IA não corresponde somente a bases estruturadas de dados pessoais, mas também a documentos completos contendo informações estratégicas, jurídicas, administrativas ou institucionais.

O Projeto de Lei nº 2.338/2023 (BRASIL, 2023) representa a primeira iniciativa legislativa abrangente destinada a estabelecer um marco geral para inteligência artificial no Brasil. O texto adota abordagem baseada em risco e incorpora princípios como transparência, supervisão humana, explicabilidade, segurança e proteção de direitos fundamentais.

Inspirado parcialmente no modelo europeu, o projeto busca estabelecer obrigações proporcionais ao nível de risco dos sistemas de IA. Entretanto, embora avance significativamente na definição de direitos e deveres relacionados ao desenvolvimento e utilização de sistemas de inteligência artificial, sua estrutura concentra-se principalmente na governança de sistemas disponibilizados no mercado e nas relações entre fornecedores, usuários e pessoas afetadas.

Nesse sentido, permanece uma lacuna relacionada ao uso interno de ferramentas de IA generativa por organizações públicas e privadas no processamento cotidiano de documentos de circulação restrita, especialmente quanto a requisitos específicos de:

- classificação de informações antes do processamento;
- controle de acesso aos modelos;
- registro de operações realizadas;
- retenção ou eliminação de dados submetidos;
- auditoria dos fluxos documentais.

No âmbito setorial brasileiro, destaca-se a Resolução CNJ nº 615/2025 (CNJ, 2025), aplicável ao Poder Judiciário, como uma das iniciativas normativas mais específicas relacionadas ao uso institucional de inteligência artificial.

A resolução estabelece requisitos de governança, transparência, supervisão humana, segurança e proteção de dados, incluindo a adoção de princípios como *privacy by design* e *privacy by default*, além de medidas relacionadas à anonimização e proteção de informações sensíveis.

Entretanto, seu alcance permanece limitado ao Poder Judiciário, não abrangendo diretamente órgãos do Poder Executivo, Poder Legislativo, governos estaduais, municípios ou organizações privadas.

Dessa forma, o cenário brasileiro apresenta avanços relevantes, porém ainda fragmentados, combinando uma lei geral de proteção de dados, uma proposta legislativa ampla de IA e normas setoriais específicas, sem uma política nacional consolidada para governança do uso interno de IA generativa sobre documentos de circulação restrita.

3.2 União Europeia: AI Act e GDPR

A União Europeia apresenta atualmente o modelo regulatório mais estruturado de inteligência artificial entre as jurisdições analisadas.

O Regulamento Europeu de Inteligência Artificial (*AI Act* — Regulamento (UE) 2024/1689) (UNIÃO EUROPEIA, 2024) estabelece o primeiro marco jurídico abrangente destinado especificamente a sistemas de IA ao adotar a Abordagem Baseada em Risco.

O regulamento classifica sistemas de inteligência artificial em diferentes níveis:

- risco inaceitável;
- alto risco;
- risco limitado;
- risco mínimo ou inexistente.

Essa classificação determina diferentes níveis de obrigações para desenvolvedores, fornecedores e usuários, incluindo requisitos de governança, documentação técnica, transparência, avaliação de conformidade, supervisão humana e gestão de riscos.

Os sistemas classificados como alto risco estão sujeitos a requisitos mais rigorosos, incluindo mecanismos de controle, monitoramento e documentação durante seu ciclo de vida.

No campo da proteção de dados pessoais, o *AI Act* opera de forma complementar ao *General Data Protection Regulation* (GDPR) (UNIÃO EUROPEIA, 2016), que desde 2018 estabelece regras abrangentes para tratamento de dados pessoais na União Europeia.

O GDPR exige fundamento jurídico adequado para tratamento de dados pessoais e prevê a realização de Avaliação de Impacto à Proteção de Dados (*Data Protection Impact Assessment* — DPIA) quando houver operações capazes de representar elevado risco aos direitos e liberdades dos titulares.

Entretanto, apesar de sua abrangência, o *AI Act* possui forte orientação regulatória relacionada à segurança, conformidade e disponibilização de sistemas de IA no mercado, aproximando-se da lógica tradicional de segurança de produtos regulados.

Essa característica cria uma lacuna quando a análise se desloca para o uso interno de sistemas de IA generativa em organizações, especialmente no processamento de documentos de circulação restrita.

O artigo 78 do *AI Act*, por exemplo, trata da confidencialidade no contexto das informações compartilhadas entre autoridades competentes e organismos regulatórios, não estabelecendo regras específicas sobre retenção de *prompts*, controle interno de documentos sensíveis ou governança operacional de informações restritas processadas por sistemas de IA.

Assim, embora o conjunto *AI Act* + GDPR represente o regime regulatório mais avançado entre os analisados, nele ainda permanece o desafio de disciplinar adequadamente o uso organizacional interno de IA sobre documentos sensíveis.

3.3 Estados Unidos: Ordens Executivas, NIST AI RMF e fragmentação regulatória

O modelo regulatório norte-americano caracteriza-se por maior fragmentação institucional quando comparado à União Europeia.

Após a publicação da Ordem Executiva (ou *Executive Order*) 14110 (UNITED STATES, 2023), em 2023, durante a administração Biden, foram estabelecidas diretrizes federais relacionadas à segurança, confiabilidade e avaliação de riscos associados a sistemas avançados de inteligência artificial.

Posteriormente, mudanças políticas resultaram na substituição dessa orientação pela Ordem Executiva 14179 (UNITED STATES, 2025), publicada em janeiro de 2025, com foco maior em inovação tecnológica e redução de barreiras regulatórias.

Esse movimento evidencia a instabilidade do modelo federal norte-americano, segundo o qual as mudanças administrativas podem alterar significativamente a orientação governamental sobre governança de IA.

Diferentemente da União Europeia, os Estados Unidos não possuem atualmente uma legislação federal abrangente equivalente ao GDPR ou ao *AI Act*.

A proteção de dados pessoais permanece distribuída entre diferentes normas setoriais, incluindo:

- HIPAA, para informações de saúde;
- FERPA, para dados educacionais;
- GLBA, para informações financeiras.

Além disso, diversos estados norte-americanos possuem legislações próprias de privacidade, como a *California Consumer Privacy Act* (CCPA) (CALIFORNIA, 2018) e a *California Privacy Rights Act* (CPRA) (CALIFORNIA, 2020).

No campo da governança de IA, destaca-se o *NIST AI Risk Management Framework* (AI RMF) (TABASSI, 2023), publicado em 2023, cujo conteúdo oferece orientações de caráter voluntário para identificação, avaliação e mitigação de riscos associados a sistemas de inteligência artificial.

Embora seja uma referência técnica internacionalmente reconhecida, esse *framework* não possui força normativa obrigatória para a maioria das organizações privadas.

Para órgãos federais, mecanismos como o *Federal Risk and Authorization Management Program* (FedRAMP) (UNITED STATES, 2024) e o *Federal Information Security Management Act* (FISMA) (UNITED STATES, 2014) estabelecem requisitos relacionados à segurança de sistemas utilizados pelo governo.

No setor privado, entretanto, permanece significativa heterogeneidade quanto a requisitos de:

- retenção de dados submetidos a LLMs;
- auditoria de operações;
- governança de *prompts*;
- rastreabilidade de processamento;
- controle sobre informações sensíveis.

Dessa forma, os Estados Unidos apresentam menor uniformidade regulatória entre as jurisdições analisadas, especialmente pela ausência de um marco federal abrangente de proteção de dados e de obrigações nacionais específicas.

3.4 Austrália: abordagem baseada em risco e governança híbrida

A Austrália representa um modelo intermediário entre a abordagem abrangente europeia e o modelo fragmentado norte-americano.

O país ainda não possui um regulamento geral equivalente ao *AI Act* europeu, mas vem desenvolvendo uma estrutura baseada em princípios, gerenciamento de riscos e responsabilidade organizacional.

O *Privacy Act* 1988 (AUSTRALIA, 1988), complementado pelos Princípios Australianos de Privacidade, ou *Australian Privacy Principles* (APPs) (AUSTRALIA, 1988), constitui o principal instrumento de proteção de dados pessoais no país. Embora anterior ao desenvolvimento da IA generativa moderna, seus princípios relacionados a coleta, uso, segurança e divulgação de informações pessoais aplicam-se ao uso de sistemas automatizados.

Em 2024, o governo australiano publicou iniciativas voltadas ao fortalecimento da segurança e confiabilidade da inteligência artificial, incluindo padrões voluntários de segurança para sistemas de IA generativa.

A abordagem australiana enfatiza:

- transparência;
- supervisão humana;
- avaliação de riscos;
- segurança durante o ciclo de vida;
- responsabilidade organizacional.

Embora essa abordagem ofereça maior flexibilidade para inovação, também apresenta limitações semelhantes às observadas em outros modelos quando aplicada especificamente ao uso interno de IA sobre documentos de circulação restrita.

A ausência de requisitos obrigatórios relacionados a retenção de dados, auditoria de *prompts* e classificação documental demonstra que, mesmo em modelos regulatórios baseados em princípios, permanece o desafio de governar o processamento interno de informações sensíveis por sistemas generativos.

Tabela 1 – Panorama comparativo entre marcos regulatórios de IA

A Tabela 1 apresenta um panorama comparativo dos marcos regulatórios de inteligência artificial e privacidade na União Europeia, nos Estados Unidos, no Brasil e na Austrália. São considerados o status regulatório, a abordagem adotada, a relação com a proteção de dados, o tratamento de documentos e informações de circulação restrita, as sanções e as principais lacunas identificadas em cada jurisdição. A comparação permite visualizar diferentes modelos de governança e seus respectivos mecanismos de proteção diante da utilização de sistemas de IA em contextos que envolvem informações restritas.

Marco Regulatório de IA e Privacidade: União Europeia, Estados Unidos, Brasil e Austrália

Panorama comparativo: status, abordagem, privacidade, documentos, sanções e principais lacunas

	UNIÃO EUROPEIA (EU) AI Act (Reg. 2024/1689)	ESTADOS UNIDOS (US) EOs 14110 → 14179 · NIST AI RMF	BRASIL (BR) PL 2.338/2023 + LGPD + CNJ 615/2025	AUSTRÁLIA (AU) Guidance for AI Adoption (AI6) + Privacy Act 1988
STATUS	✓ Em vigor — ago/2024 ✓ Faseado até 2027	⚠ Sem lei federal — EOs dominam (EO 14110 revogada; EO 14179 em vigor desde jan/2025) ⚠ +1.000 bills estaduais em 2025	▲ PL aprovado no Senado (dez/2024) ▲ Na Câmara — sem data de votação	▲ Sem lei federal de IA dedicada ▲ Guidance for AI Adoption (AI6) substituiu o VAISS voluntário (out/2025): guardrails obrigatórios descartados no National AI Plan (dez/2025)
ABORDAGEM	Regulação baseada em risco (4 níveis: inaceitável, alto, limitado, mínimo). Proibições absolutas desde fev/2025. Obrigações de conformidade e registro para sistemas de alto risco.	EO 14110 (2023, Biden) fixou diretrizes de segurança, confiabilidade e avaliação de risco para IA avançada; substituída pela EO 14179 (jan/2025), voltada à inovação e à redução de barreiras regulatórias — evidenciando a instabilidade do modelo federal. NIST AI RMF (2023): framework técnico voluntário, sem força normativa para o setor privado.	Inspirado no AI Act europeu mas centrado em direitos da pessoa humana. Prevê direito à determinação humana em decisões relevantes, à explicabilidade e à não discriminação. ANPD como coordenadora do Sistema de IA (SIA).	Abordagem tecnologicamente neutra: apoia-se em leis setoriais já existentes (Privacy Act, Consumer Law, Copyright Act) em vez de uma lei específica de IA. O Australian AI Safety Institute (AIS), operacional desde início de 2026, faz análise de lacunas e recomendações — sem poder de fiscalização.
RELAÇÃO COM PRIVACIDADE	Complementa o GDPR mas não o substitui; o AI Act é uma lei de segurança do produto: o GDPR preenche as lacunas de direitos individuais. Interação obrigatória entre as duas normas.	Sem lei federal de privacidade. Proteção distribuída entre normas setoriais — HIPAA (saúde), FERPA (dados educacionais), GLBA (informações financeiras) — e leis estaduais como CCPA/CPRA. Fragmentação extrema entre jurisdições.	LGPD (Lei 13.709/2018) já vigente. A ANPD regula incidentes desde a Resolução 15/2024. Especialistas apontam insuficiência para modelos generativos em larga escala.	Privacy Act 1988 + Privacy and Other Legislation Amendment Act 2024 (Royal Assent 10/dez/2024): novo tort civil por invasão grave de privacidade e reforço do APP 11 (medidas técnicas e organizacionais). Sem lei de IA dedicada; segunda tranche da reforma de privacidade ainda pendente.
DOCUMENTOS E SIGILO	Art. 78 trata da confidencialidade no âmbito regulatório, mas não endereça o uso interno de IA em documentos restritos de organizações.	Sem regulação federal específica para uso de IA em documentos internos. FedRAMP e FISMA impõem requisitos de segurança a sistemas de órgãos federais; no setor privado, heterogeneidade quanto à retenção de dados em LLMs,	CNJ 615/2025 introduz privacy by design e privacy by default para IA no Judiciário. Lacuna para o Executivo e órgãos privados.	Sem regulação específica para uso de IA em documentos. A DTA (Digital Transformation Agency) define políticas obrigatórias de uso responsável de IA para o setor público federal; setor

	UNIÃO EUROPEIA (EU) AI Act (Reg. 2024/1689)	ESTADOS UNIDOS (US) EOs 14110 → 14179 · NIST AI RMF	BRASIL (BR) PL 2.338/2023 + LGPD + CNJ 615/2025	AUSTRÁLIA (AU) Guidance for AI Adoption (AI6) + Privacy Act 1988
		auditoria, governança de prompts e rastreabilidade.		privado sem exigências uniformes.
SANÇÕES	Até €35 mi ou 7% do faturamento	FTC Act — caso a caso. Variam por estado.	Previstas no PL: LGPD; até R\$ 50 mi	Privacy Act: até AUD 50 mi, 30% do faturamento ajustado ou 3x o benefício obtido (o que for maior)
LACUNAS / PONTOS DE ATENÇÃO	Foco regulatório em sistemas, não no uso interno de documentos restritos. Lacunas sobre IA generativa e prompts. Orientações ainda em consolidação pela Comissão Europeia.	Ausência de lei federal unificada; troca de administração já reverteu a orientação de governança (EO 14110 → 14179). Conflitos entre estaduais. Pouca regulação de privacidade de IA e uso de documentos corporativos.	PL ainda em tramitação e sujeito a mudanças. Capacidade operacional da ANPD em escala ainda limitada. Diretrizes para IA generativa ainda incipientes.	Guardrails obrigatórios abandonados (dez/2025): sem lei de IA dedicada nem regulador central. AISI tem função apenas consultiva e de monitoramento, sem poder de fiscalização. Segunda tranche da reforma da Privacy Act ainda pendente.
LEITURA GERAL — A União Europeia lidera a regulação baseada em risco e proteção de direitos; os Estados Unidos priorizam inovação e competitividade com forte fragmentação e instabilidade federal entre administrações; o Brasil avança em alinhamento com padrões internacionais, mas ainda enfrenta lacunas regulatórias e operacionais, especialmente para IA generativa em documentos restritos; a Austrália opta por uma abordagem de baixa intervenção, apoiada em leis setoriais existentes e em um instituto de segurança sem poder normativo, tendo abandonado a proposta de guardrails obrigatórios. Fontes: Reg. (UE) 2024/1689; EOs 14110 e 14179; NIST AI RMF (2023); PL 2.338/2023; LGPD (Lei 13.709/2018); CNJ 615/2025; National AI Plan da Austrália (dez/2025); Guidance for AI Adoption / AI6 (out/2025); Privacy and Other Legislation Amendment Act 2024 (AU). Elaboração própria — jul/2026.				

3.5 Análise comparativa das lacunas regulatórias

A comparação entre Brasil, União Europeia, Estados Unidos e Austrália demonstra que nenhuma das jurisdições analisadas apresenta, atualmente, um modelo regulatório completo para o uso interno de inteligência artificial generativa aplicada a documentos de circulação restrita. A Tabela 2 resume as principais lacunas regulatórias de cada jurisdição.

Tabela 2 – Comparação das lacunas regulatórias de cada jurisdição

Jurisdição	Modelo regulatório	Proteção de dados	Governança de IA	Uso interno sobre documentos restritos
Brasil	Em consolidação	LGPD	Parcial	Lacuna significativa
União Europeia	Baseado em risco	GDPR + AI Act	Avançada	Parcialmente endereçado

Jurisdição	Modelo regulatório	Proteção de dados	Governança de IA	Uso interno sobre documentos restritos
Estados Unidos	Fragmentado	Setorial/estadual	NIST AI RMF voluntário	Lacuna elevada
Austrália	Híbrido baseado em princípios	<i>Privacy Act</i>	Orientações voluntárias	Lacuna moderada
Fontes: Reg. (UE) 2024/1689 (AI Act); Regulamento (UE) 2016/679 (GDPR); PL 2.338/2023; LGPD (Lei 13.709/2018); NIST AI Risk Management Framework (2023); Privacy Act 1988 (Cth), Austrália; Guidance for AI Adoption / AI6 (out/2025); National AI Plan da Austrália (dez/2025). Elaboração própria — jul/2026.				

A análise evidencia que os atuais marcos regulatórios enfatizam predominantemente a criação, disponibilização e utilização geral de sistemas de inteligência artificial e, por sua vez, não contemplam de forma suficientemente detalhada a governança operacional necessária quando essas tecnologias são incorporadas aos fluxos internos de processamento documental.

Essa lacuna envolve aspectos técnicos e organizacionais, tais como:

- classificação prévia da informação;
- definição de ambientes autorizados de processamento;
- controle de identidade e privilégios;
- registro e auditoria das operações;
- políticas de retenção e descarte;
- avaliação contínua de riscos.

Portanto, embora os instrumentos analisados representem avanços relevantes na construção de uma governança global de inteligência artificial, permanece uma oportunidade de desenvolvimento normativo e técnico voltado especificamente ao uso seguro de IA generativa em documentos de circulação restrita.

4. LACUNAS CRÍTICAS NOS MARCOS NORMATIVOS

Embora os marcos regulatórios analisados na Seção 3 anterior representem avanços significativos na disciplina da inteligência artificial, da proteção de dados e da governança digital, a análise comparativa evidencia limitações relevantes quando o objeto de proteção deixa de ser o sistema de IA e passa a ser a informação processada por esse sistema. Em outras palavras, a maior parte das normas concentra-se na regulação do desenvolvimento, da disponibilização e da utilização de sistemas de inteligência artificial, dedicando pouca atenção às salvaguardas necessárias para o uso interno dessas tecnologias sobre documentos de circulação restrita.

As lacunas discutidas nesta Seção não são resultantes da inexistência de normas, mas da ausência de dispositivos específicos capazes de disciplinar o processamento de informações sensíveis por modelos de linguagem de grande porte (LLMs) no contexto organizacional. Essas insuficiências são examinadas sob as perspectivas conceitual, regulatória, jurídica e técnico-operacional.

4.1 O *prompt* como dado não regulado

Uma das principais lacunas identificadas refere-se ao tratamento jurídico do *prompt* — nomenclatura dada à instrução ou conjunto de informações fornecidas pelo usuário ao sistema de IA. Embora diferentes normas estabeleçam regras para a proteção de dados pessoais, documentos classificados e informações estratégicas, nenhuma delas reconhece expressamente o *prompt* como uma categoria autônoma de ativo informacional.

Na prática, um *prompt* pode conter dados pessoais, segredos comerciais, pareceres jurídicos, informações estratégicas, documentos classificados ou qualquer outro conteúdo sujeito a restrições de acesso. Entretanto, quando essas informações são inseridas em um modelo LLM, deixam de ser tratadas pelo ordenamento jurídico como documento protegido e passam a ser interpretadas apenas como entrada (*input*) para processamento computacional.

Essa lacuna produz consequências relevantes. Um servidor público que copie o conteúdo de uma ata de reunião reservada e o insira em uma ferramenta pública de IA generativa poderá não encontrar vedação normativa expressa que trate especificamente dessa ação, embora o documento original permaneça protegido pela legislação aplicável. Em outras palavras, a proteção jurídica atribuída ao documento não acompanha, necessariamente, o seu conteúdo quando este é reutilizado como *prompt*.

Esse problema manifesta-se de forma semelhante nas jurisdições analisadas. Nem a LGPD, nem o *AI Act*, nem os atuais instrumentos regulatórios norte-americanos ou australianos estabelecem regras específicas para o tratamento do *prompt* como ativo informacional dotado de valor jurídico próprio, permanecendo essa questão praticamente sem disciplina normativa.

4.2 Regulação da IA versus governança do uso organizacional

A segunda lacuna decorre da distinção entre normatizar os sistemas de inteligência artificial e disciplinar sua utilização pelas organizações.

Os principais instrumentos normativos analisados — incluindo o *AI Act* europeu, o Projeto de Lei n.º 2.338/2023 e as Ordens Executivas norte-americanas — concentram-se, predominantemente, nas obrigações aplicáveis aos desenvolvedores, fornecedores, distribuidores e implantadores de sistemas de IA. Essa abordagem busca assegurar que os produtos disponibilizados ao mercado atendam a requisitos mínimos de segurança, transparência e gestão de riscos.

Entretanto, uma vez adquirido ou disponibilizado o sistema de IA, poucos desses instrumentos disciplinam detalhadamente o seu uso cotidiano dentro das organizações. Como consequência,

uma empresa ou órgão público pode utilizar um modelo de linguagem para processar memorandos internos, pareceres jurídicos, contratos ou documentos estratégicos sem que existam regras específicas sobre classificação da informação, autorização para processamento, registro das operações, rastreabilidade dos acessos ou mecanismos de auditoria.

Forma-se, assim, uma assimetria regulatória: enquanto o fornecedor da tecnologia está submetido a obrigações relativamente detalhadas, a organização que utiliza a ferramenta para processar seu próprio patrimônio informacional permanece sujeita, em grande medida, apenas a normas gerais de proteção de dados e segurança da informação.

É justamente nesse espaço regulatório — o uso operacional interno da IA — que se concentram os maiores riscos identificados neste estudo.

4.3 Fragmentação entre os regimes jurídicos de proteção da informação

Outra limitação relevante decorre da ausência de integração entre os diferentes regimes jurídicos que disciplinam a proteção da informação.

No Brasil, documentos de circulação restrita podem estar simultaneamente submetidos à Lei de Acesso à Informação, à Lei Geral de Proteção de Dados Pessoais, à Política Nacional de Segurança da Informação, a normas setoriais e a políticas internas de governança documental. Cada um desses instrumentos protege aspectos distintos da informação, utilizando conceitos, objetivos e mecanismos próprios.

Quando esses documentos passam a ser processados por sistemas de IA generativa, entretanto, surge uma questão para a qual o ordenamento jurídico ainda não oferece resposta uniforme: quais requisitos devem prevalecer durante esse processamento?

A Lei de Acesso à Informação disciplina os graus de sigilo e os procedimentos para classificação documental, mas não trata do emprego de inteligência artificial sobre documentos classificados. A LGPD protege dados pessoais, porém não disciplina especificamente o tratamento de documentos institucionais que contenham informações estratégicas não enquadradas como dados pessoais. Da mesma forma, a Política Nacional de Segurança da Informação estabelece diretrizes gerais de proteção, sem definir controles específicos para o uso de modelos generativos.

Essa fragmentação produz insegurança jurídica e dificulta a construção de políticas organizacionais consistentes para o uso seguro da inteligência artificial sobre documentos de circulação restrita.

4.4 Ausência de requisitos auditáveis para retenção e tratamento de prompts

Outra lacuna importante refere-se à retenção das informações submetidas aos modelos de IA.

Diversos fornecedores afirmam, em seus contratos ou políticas de privacidade, que os prompts enviados pelos usuários não serão utilizados para treinamento dos modelos ou poderão ser excluídos mediante determinadas configurações contratuais. Entretanto, essas garantias normalmente baseiam-se em declarações do próprio fornecedor, não existindo mecanismos técnicos padronizados que permitam às organizações verificar, de forma independente, se tais informações foram efetivamente descartadas ou deixaram de ser utilizadas.

Sob a perspectiva da governança da informação, isso representa um problema significativo. Organizações que processam documentos de circulação restrita necessitam demonstrar conformidade, rastreabilidade e capacidade de auditoria sobre todo o ciclo de vida da informação. Quando essas evidências dependem exclusivamente da confiança depositada no fornecedor da ferramenta, a capacidade de comprovação torna-se limitada.

Nenhuma das jurisdições analisadas estabelece atualmente um padrão obrigatório, auditável e tecnicamente verificável de não retenção de *prompts* aplicável de forma geral aos sistemas de IA generativa. A ausência desse requisito reduz a confiança no uso dessas tecnologias em ambientes que tratam informações sensíveis ou classificadas.

4.5 União Europeia: limites operacionais do *AI Act* e do GDPR

O modelo europeu representa, atualmente, a estrutura regulatória mais abrangente para a inteligência artificial. A combinação entre o *AI Act* e o Regulamento Geral sobre a Proteção de Dados (GDPR) estabelece obrigações robustas relacionadas à classificação de riscos, avaliação de conformidade, transparência, supervisão humana e proteção de dados pessoais.

Entretanto, essa arquitetura normativa permanece fortemente orientada à regulação dos sistemas de IA e das condições para sua disponibilização e utilização no mercado europeu. Embora o *AI Act* imponha deveres relevantes aos provedores e implantadores de sistemas de IA, ele não disciplina de maneira detalhada os procedimentos internos que uma organização deve adotar ao utilizar esses sistemas para processar documentos de circulação restrita.

Da mesma forma, o GDPR oferece instrumentos importantes para a proteção de dados pessoais, como a Avaliação de Impacto à Proteção de Dados (DPIA), mas não estabelece controles específicos para documentos institucionais que contenham informações estratégicas, confidenciais ou classificadas sem envolver necessariamente dados pessoais.

Assim, mesmo no ambiente regulatório europeu, permanece aberta a lacuna relacionada ao processamento interno de informações sensíveis por modelos generativos.

4.6 Austrália: salvaguardas obrigatórias e desafios de implementação

A Austrália vem adotando uma estratégia regulatória intermediária entre o modelo europeu e o norte-americano. Além da revisão em curso do *Privacy Act* 1988, o governo australiano propôs a implementação de *Mandatory Guardrails* para aplicações de IA consideradas de alto risco,

estabelecendo requisitos relacionados à gestão de riscos, transparência, responsabilização e testes de segurança.

Essa abordagem representa avanço importante ao reconhecer que a governança da IA exige medidas específicas para reduzir riscos tecnológicos antes da implantação dos sistemas. Entretanto, assim como ocorre nas demais jurisdições analisadas, as salvaguardas propostas concentram-se na segurança e na governança dos sistemas de IA e não disciplinam, de forma suficientemente detalhada, o processamento cotidiano de documentos de circulação restrita pelas organizações.

Persistem, portanto, questões fundamentais relacionadas à classificação da informação, ao controle de acesso, à rastreabilidade dos prompts, à auditoria das operações realizadas e à comprovação técnica de que informações sensíveis não foram retidas ou reutilizadas pelos modelos.

5. CENÁRIOS ILUSTRATIVOS DE RISCO

As lacunas normativas identificadas na Seção 4 tornam-se mais evidentes quando analisadas em situações concretas de uso organizacional da inteligência artificial generativa. Embora os marcos regulatórios examinados estabeleçam obrigações gerais relacionadas à proteção de dados, à segurança da informação e à governança de sistemas de IA, sua aplicação prática revela zonas de incerteza quando documentos de circulação restrita são processados por modelos de linguagem de grande porte (*Large Language Models* – LLMs).

Os cenários apresentados a seguir não correspondem a casos reais, mas foram construídos a partir de situações plausíveis identificadas na literatura especializada, em documentos técnicos e nas arquiteturas atualmente empregadas por organizações públicas e privadas. O objetivo é demonstrar como as lacunas regulatórias anteriormente identificadas podem materializar-se em diferentes contextos organizacionais, evidenciando riscos técnicos, jurídicos e operacionais que extrapolam as abordagens tradicionalmente contempladas pelos marcos regulatórios analisados.

5.1 Cenário 1: Processamento de documento reservado por LLM externo em órgão público

Uma assessoria jurídica de órgão da Administração Pública Federal utiliza um modelo de linguagem acessado por navegador para auxiliar na redação de um memorando classificado como reservado, relacionado a processo administrativo disciplinar em andamento. Para obter uma minuta mais consistente, o servidor insere no *prompt* informações detalhadas, incluindo a identificação do investigado, descrição dos fatos apurados, fundamentos jurídicos e proposta de decisão.

Sob a perspectiva jurídica, esse cenário pode envolver o tratamento de dados pessoais e informações protegidas por sigilo administrativo sem que existam diretrizes específicas disciplinando sua submissão a um sistema de IA hospedado em infraestrutura externa. Dependendo da natureza das informações processadas, poderão incidir simultaneamente disposições da Lei Geral de Proteção de Dados Pessoais, da Lei de Acesso à Informação, da Política Nacional de Segurança da Informação e das normas internas de classificação documental do órgão.

Sob a perspectiva técnica, entretanto, os riscos são ainda mais amplos. O conteúdo submetido ao modelo poderá ser transmitido para infraestrutura sob controle do fornecedor, armazenado temporariamente para fins operacionais, submetido a mecanismos de monitoramento de segurança, utilizado em processos de suporte técnico ou integrar componentes auxiliares da arquitetura do serviço, conforme as políticas contratuais adotadas pelo provedor. Além disso, inexistem mecanismos padronizados que permitam ao órgão comprovar, de forma independente, a não retenção dessas informações.

Também se observam riscos relacionados à exfiltração de dados, comprometimento de credenciais, acesso administrativo privilegiado pelo fornecedor, ausência de segregação lógica entre clientes (*multi-tenancy*) e limitações na rastreabilidade do tratamento realizado.

Esse cenário evidencia que as normas atualmente vigentes estabelecem obrigações gerais para proteção da informação, mas permanecem insuficientes para disciplinar operacionalmente o uso cotidiano de LLMs sobre documentos administrativos de circulação restrita.

5.2 Cenário 2: Síntese automatizada de ata de reunião estratégica

Uma companhia aberta utiliza uma ferramenta de transcrição automática baseada em inteligência artificial para produzir a ata resumida de reunião do comitê de divulgação, durante a qual foram discutidos resultados financeiros ainda não divulgados ao mercado, estratégias de aquisição e alterações relevantes no planejamento corporativo.

A gravação é processada por um modelo de linguagem executado em infraestrutura de nuvem pertencente a fornecedor internacional.

Além das questões relacionadas à privacidade e à confidencialidade das informações corporativas, esse cenário ativa o regime jurídico das informações privilegiadas previsto na Lei n.º 6.385/1976 e na Resolução CVM n.º 44/2021. Ainda que não ocorra negociação de valores mobiliários com base nas informações processadas, a transmissão de conteúdo estratégico para ambiente tecnológico externo pode representar falha relevante na gestão da informação privilegiada.

Do ponto de vista técnico, o principal risco não decorre apenas da possibilidade de vazamento integral da reunião. Mesmo que apenas o resumo seja disponibilizado, modelos generativos podem produzir saídas que permitam inferir informações estratégicas não públicas, como

resultados financeiros, planos de fusão, alterações societárias ou decisões corporativas ainda não anunciadas oficialmente.

Essa capacidade de inferência representa um vetor de risco frequentemente negligenciado pelas organizações, pois informações aparentemente inofensivas podem, quando correlacionadas, revelar conteúdos protegidos por sigilo empresarial ou regulatório.

Embora os marcos regulatórios analisados imponham obrigações relacionadas à proteção de dados e à governança de IA, nenhum deles disciplina de forma específica o risco de inferência decorrente da síntese automatizada de documentos estratégicos produzidos por modelos generativos.

5.3 Cenário 3: Tradução de documento classificado por ferramenta externa

Um ministério necessita traduzir uma nota diplomática classificada para participação em reunião internacional realizada em curto prazo. Diante da indisponibilidade imediata de tradutor credenciado, um servidor utiliza uma ferramenta de tradução baseada em IA hospedada em infraestrutura internacional.

Sob a perspectiva operacional, a decisão produz ganho significativo de eficiência e reduz o tempo necessário para preparação da documentação. Entretanto, o processamento do documento em ambiente externo implica a transmissão de informações diplomáticas sensíveis para infraestrutura localizada fora do ambiente tecnológico controlado pelo Estado.

Além dos riscos tradicionais de confidencialidade, surge uma questão adicional relacionada à soberania digital. Dependendo da localização física dos servidores e da legislação aplicável ao fornecedor, o conteúdo processado poderá estar sujeito a diferentes regimes jurídicos nacionais, solicitações governamentais de acesso ou obrigações legais impostas ao operador do serviço.

Embora a legislação brasileira discipline a classificação das informações e estabeleça regras para seu acesso, não existem dispositivos específicos que definam critérios objetivos para utilização de ferramentas generativas externas na tradução de documentos classificados.

Essa ausência de disciplina normativa cria uma zona de incerteza em que decisões operacionais passam a depender exclusivamente da avaliação individual do usuário ou das políticas internas da organização.

5.4 Cenário 4: Assistente corporativo baseado em RAG para consulta documental

Uma organização implanta um assistente corporativo baseado na técnica de *Retrieval-Augmented Generation* (RAG), integrado ao SharePoint, ao OneDrive, aos sistemas de gestão documental e aos repositórios internos de conhecimento. O objetivo é permitir que colaboradores consultem rapidamente normas internas, contratos, pareceres jurídicos e documentos técnicos utilizando linguagem natural.

Embora a solução permaneça dentro do ambiente corporativo, sua segurança depende diretamente da correta implementação dos mecanismos de autenticação, autorização e controle de acesso.

Suponha que um usuário solicite ao assistente:

"Liste todos os contratos relacionados ao Projeto Alfa e apresente os principais riscos jurídicos identificados."

Caso a arquitetura de autorização não respeite rigorosamente as permissões documentais existentes, o modelo poderá recuperar informações pertencentes a áreas às quais o usuário jamais teve acesso autorizado, produzindo uma forma sofisticada de vazamento interno.

Nesse cenário, o problema não está necessariamente no modelo de linguagem, mas na integração entre mecanismos de recuperação documental, indexação semântica e controle de identidade.

Além dos riscos tradicionais de confidencialidade, surgem vulnerabilidades relacionadas ao excesso de privilégios, à recuperação indevida de documentos (*over-retrieval*), à ausência de segregação de contexto, ao *prompt injection* indireto e à manipulação das bases utilizadas pelo mecanismo RAG (*data poisoning*).

Esses riscos encontram correspondência em categorias descritas pelo OWASP Top 10 for LLM Applications, especialmente aquelas relacionadas à divulgação de informações sensíveis (*Sensitive Information Disclosure*), à injeção de prompts (*Prompt Injection*) e ao tratamento inseguro das saídas produzidas pelos modelos (*Insecure Output Handling*).

Entretanto, os marcos regulatórios analisados concentram-se predominantemente na governança dos sistemas de IA, oferecendo poucas diretrizes específicas sobre controles técnicos para arquiteturas corporativas baseadas em recuperação aumentada por geração.

6. CONCLUSÕES E RECOMENDAÇÕES

A análise desenvolvida neste trabalho permitiu examinar criticamente a interação entre segurança da informação, privacidade de dados e o emprego de sistemas de inteligência artificial generativa no processamento de documentos de circulação restrita em organizações públicas e privadas. A partir de uma revisão bibliográfica e documental dos principais marcos regulatórios nacionais e internacionais, verificou-se que o avanço tecnológico das ferramentas de IA ocorre em ritmo significativamente superior à evolução das normas destinadas a disciplinar sua utilização em ambientes organizacionais sensíveis.

Os resultados obtidos confirmam a hipótese central do trabalho. Embora os instrumentos regulatórios analisados — notadamente a Lei Geral de Proteção de Dados Pessoais (LGPD), o Regulamento Europeu de Inteligência Artificial (*AI Act*), a Resolução CNJ n.º 615/2025 e os

referenciais regulatórios norte-americanos e australianos — estabeleçam importantes requisitos para o desenvolvimento, a disponibilização e a utilização de sistemas de inteligência artificial, permanecem insuficientes para disciplinar, de forma específica, o processamento interno de documentos de circulação restrita por modelos de linguagem de grande porte (LLMs).

A primeira conclusão refere-se à existência de uma lacuna regulatória comum entre as diferentes jurisdições analisadas. Em maior ou menor grau, todas concentram seus mecanismos de controle na governança dos sistemas de IA e nas responsabilidades de seus fornecedores, dedicando atenção limitada às salvaguardas operacionais necessárias quando essas tecnologias passam a integrar fluxos documentais internos de organizações públicas e privadas. Essa constatação evidencia que os riscos mais relevantes atualmente decorrem menos da disponibilidade da tecnologia e mais da forma como ela é utilizada no cotidiano institucional.

A segunda conclusão diz respeito ao papel desempenhado pelos instrumentos normativos brasileiros. Embora o país ainda não disponha de uma legislação específica sobre inteligência artificial em vigor, a Resolução CNJ n.º 615/2025 representa importante avanço ao incorporar princípios como *privacy by design*, *privacy by default*, rastreabilidade, supervisão humana e governança de dados em projetos de IA no âmbito do Poder Judiciário. Todavia, sua aplicação permanece restrita a esse segmento institucional, inexistindo instrumento equivalente com alcance transversal para a Administração Pública em geral ou para o setor privado.

A terceira conclusão decorre da ausência de reconhecimento jurídico do *prompt* como ativo informacional sujeito ao mesmo regime de proteção aplicável às informações nele contidas. Ao longo da pesquisa verificou-se que documentos classificados, informações estratégicas, dados pessoais e segredos comerciais podem perder parte de sua proteção normativa quando reutilizados como entrada (*input*) para sistemas de IA generativa. Essa lacuna evidencia a necessidade de evolução tanto das normas quanto das práticas organizacionais de governança da informação.

Além dessas constatações, os cenários ilustrativos demonstraram que os riscos associados ao uso de IA generativa extrapolam a proteção tradicional de dados pessoais, abrangendo aspectos relacionados à confidencialidade institucional, à soberania digital, à rastreabilidade das operações, à segurança documental e à governança do ciclo de vida da informação. Destacam-se, nesse contexto, ameaças como exfiltração de *prompts*, retenção indevida de informações, inferência de conteúdo sensível, *prompt injection*, *data poisoning* e vulnerabilidades em arquiteturas baseadas em *Retrieval-Augmented Generation* (RAG), temas ainda insuficientemente contemplados pelos atuais marcos regulatórios.

Como contribuição, este estudo propõe uma abordagem integrada para análise dos riscos decorrentes do emprego de IA generativa sobre documentos de circulação restrita, articulando dimensões jurídicas, tecnológicas e organizacionais que, em geral, são tratadas de forma isolada pela literatura especializada. Ao considerar o documento de circulação restrita como uma unidade informacional cujo ciclo de vida deve permanecer protegido independentemente da

tecnologia utilizada para seu processamento, o trabalho amplia a discussão sobre governança da informação em ambientes baseados em inteligência artificial.

À luz dos resultados obtidos, recomenda-se que organizações públicas e privadas adotem medidas preventivas de governança capazes de antecipar exigências regulatórias ainda inexistentes. Entre elas destacam-se:

- a implementação de políticas formais de classificação de *prompts*, vinculando seu nível de proteção ao documento ou à informação processada;
- a utilização preferencial de modelos executados localmente (*on-premise*) ou em ambientes de nuvem que ofereçam garantias técnicas auditáveis de não retenção de dados;
- o registro integral das operações realizadas por sistemas de IA sobre documentos de circulação restrita, incluindo identificação do usuário, data e hora da operação, modelo empregado, classificação da informação processada e finalidade do tratamento;
- a revisão dos contratos firmados com fornecedores de IA, incorporando cláusulas específicas sobre confidencialidade, rastreabilidade, retenção de dados, auditoria independente e conformidade com a LGPD, o *AI Act* e demais referenciais internacionais aplicáveis;
- a adoção de estruturas formais de governança alinhadas a referenciais internacionais, como o *NIST AI Risk Management Framework*, a ISO/IEC 42001 e a ISO/IEC 23894, integradas aos programas já existentes de segurança da informação e proteção de dados.

Este estudo apresenta, contudo, algumas limitações. A pesquisa possui natureza predominantemente qualitativa e fundamenta-se em uma revisão bibliográfica, análise documental e construção de cenários ilustrativos, não contemplando estudos empíricos, avaliações quantitativas ou experimentos envolvendo organizações específicas. Além disso, o rápido processo de evolução tecnológica e regulatória da inteligência artificial poderá alterar parte do panorama normativo descrito, exigindo atualização contínua das análises desenvolvidas.

Como perspectivas para trabalhos futuros, recomenda-se a realização de estudos empíricos sobre o uso organizacional de IA generativa em ambientes que tratam documentos de circulação restrita, bem como o desenvolvimento de modelos de classificação automática de *prompts*, mecanismos auditáveis de não retenção de dados, métricas de maturidade em governança de IA e modelos de conformidade específicos para arquiteturas baseadas em agentes inteligentes e sistemas de *Retrieval-Augmented Generation*.

Conclui-se, portanto, que a regulação atualmente existente constitui condição necessária, mas ainda insuficiente para enfrentar os desafios impostos pelo uso cotidiano da inteligência artificial generativa sobre documentos de circulação restrita. Enquanto novas normas não forem

consolidadas, caberá às organizações adotar mecanismos preventivos de governança capazes de assegurar que inovação tecnológica, segurança da informação e proteção de dados evoluam de forma coordenada, preservando a confidencialidade, a integridade, a disponibilidade e a rastreabilidade das informações sob sua responsabilidade.

REFERÊNCIAS BIBLIOGRÁFICAS

AUSTRALIA. Department of Industry, Science and Resources. *Introducing mandatory guardrails for AI in high-risk settings: Proposal paper*. Canberra, 2025. Disponível em: <https://consult.industry.gov.au/supporting-responsible-ai/mandatory-guardrails-for-ai-in-high-risk-settings>. Acesso em: 16 jul. 2026.

AUSTRALIA. Department of Industry, Science and Resources. *Voluntary AI Safety Standard*. Canberra, 2024. Disponível em: <https://www.industry.gov.au/publications/voluntary-ai-safety-standard>. Acesso em: 16 jul. 2026.

AUSTRALIA. Privacy Act 1988. Act No. 119, 1988. Canberra: Federal Register of Legislation, 1988. Disponível em: <https://www.legislation.gov.au/C2004A03712/latest>. Acesso em: 16 jul. 2026.

BIONI, Bruno Ricardo. *Proteção de dados pessoais: a função e os limites do consentimento*. 3. ed. Rio de Janeiro: Forense, 2024.

BRASIL. Constituição (1988). *Constituição da República Federativa do Brasil de 1988*. Brasília, DF: Presidência da República, 1988. Disponível em: https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 16 jul. 2026.

BRASIL. Lei nº 12.527, de 18 de novembro de 2011. *Regula o acesso a informações previsto no inciso XXXIII do art. 5º da Constituição Federal*. Diário Oficial da União: Brasília, DF, 18 nov. 2011. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm. Acesso em: 16 jul. 2026.

BRASIL. Lei nº 13.709, de 14 de agosto de 2018. *Lei Geral de Proteção de Dados Pessoais (LGPD)*. Diário Oficial da União: Brasília, DF, 15 ago. 2018. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. Acesso em: 16 jul. 2026.

BRASIL. Decreto nº 9.637, de 26 de dezembro de 2018. *Institui a Política Nacional de Segurança da Informação*. Diário Oficial da União: Brasília, DF, 27 dez. 2018. Disponível em: https://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2018/Decreto/D9637.htm. Acesso em: 16 jul. 2026.

BRASIL. Senado Federal. Projeto de Lei nº 2.338, de 2023. *Dispõe sobre o uso da Inteligência Artificial*. Brasília, DF, 2025. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/157233>. Acesso em: 16 jul. 2026.

CALIFORNIA. California Consumer Privacy Act of 2018. Sacramento: State of California, 2018. Disponível em: <https://www.oag.ca.gov/privacy/ccpa>. Acesso em: 16 jul. 2026.

CARLINI, Nicholas et al. *Extracting Training Data from Large Language Models*. In: 30th USENIX Security Symposium, 2021. Disponível em: <https://www.usenix.org/system/files/sec21-carlini-extracting.pdf>. Acesso em: 16 jul. 2026.

CONSELHO NACIONAL DE JUSTIÇA (Brasil). Resolução nº 615, de 11 de março de 2025. *Dispõe sobre diretrizes para o desenvolvimento, utilização e governança de soluções de inteligência artificial no Poder Judiciário*. Brasília, DF, 2025. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/6001>. Acesso em: 16 jul. 2026.

DONEDA, Danilo. *Da privacidade à proteção de dados pessoais: fundamentos da Lei Geral de Proteção de Dados*. 3. ed. São Paulo: Thomson Reuters Brasil, 2024.

EUROPEAN DATA PROTECTION BOARD (EDPB). *Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models*. Brussels, 2024.

Disponível em: https://www.edpb.europa.eu/system/files/2024-12/edpb_opinion_202428_ai-models_en.pdf. Acesso em: 16 jul. 2026.

EUROPEAN UNION. *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. *Official Journal of the European Union*, 2024. Disponível em: <https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng>. Acesso em: 16 jul. 2026.

FERNANDES, Marcos Antônio de Souza. *Repositório seguro e o impacto gerado pela Lei Geral de Proteção de Dados Pessoais (LGPD)*. 2022. Dissertação (Mestrado em Ciência da Informação) — Universidade de Brasília, Brasília, 2022. Disponível em: <http://repositorio.unb.br/handle/10482/45357>. Acesso em: 16 jul. 2026.

MENDES, Laura Schertel; DONEDA, Danilo. Reflexões iniciais sobre a nova Lei Geral de Proteção de Dados. *Revista de Direito do Consumidor*, São Paulo, v. 120, p. 469-483, 2018. Disponível em: <https://revistadedireitodoconsumidor.emnuvens.com.br/rdc/article/view/1116>. Acesso em: 16 jul. 2026.

MICROSOFT. *Enterprise data protection in Microsoft 365 Copilot and Microsoft 365 Copilot Chat*. Microsoft Learn, 2026. Disponível em: <https://learn.microsoft.com/en-us/microsoft-365/copilot/enterprise-data-protection>. Acesso em: 16 jul. 2026.

MICROSOFT. *Privacy FAQ for Microsoft Copilot*. Microsoft Support, 2026. Disponível em: <https://support.microsoft.com/en-us/microsoft-copilot/privacy-faq-for-microsoft-copilot>. Acesso em: 16 jul. 2026.

NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (NIST). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. Gaithersburg, MD: NIST, 2023. (NIST AI 100-1). Disponível em: <https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf>. Acesso em: 16 jul. 2026.

UNITED STATES OF AMERICA. Executive Order 14148: Initial Rescissions of Harmful Executive Orders and Actions. *Federal Register*, 2025. Disponível em: <https://www.federalregister.gov/documents/2025/01/28/2025-01901/initial-rescissions-of-harmful-executive-orders-and-actions>. Acesso em: 16 jul. 2026.

UNITED STATES OF AMERICA. Executive Order 14179. *Removing Barriers to American Leadership in Artificial Intelligence*. *Federal Register*, Washington, DC, 31 jan. 2025. Disponível em: <https://www.federalregister.gov/documents/2025/01/31/2025-02172/removing-barriers-to-american-leadership-in-artificial-intelligence>. Acesso em: 16 jul. 2026.

UNITED STATES OF AMERICA. Executive Order 14365. *Ensuring a National Policy Framework for Artificial Intelligence*. *Federal Register*, Washington, DC, 16 dez. 2025. Disponível em: <https://www.federalregister.gov/documents/2025/12/16/2025-23092/ensuring-a-national-policy-framework-for-artificial-intelligence>. Acesso em: 16 jul. 2026.

VASWANI, Ashish et al. *Attention Is All You Need*. In: *Advances in Neural Information Processing Systems* (NeurIPS 2017). Long Beach, CA, 2017. Disponível em: <https://arxiv.org/abs/1706.03762>. Acesso em: 16 jul. 2026.