

Automação de triagem e classificação de documentos em contencioso de massa e *eDiscovery*: uma abordagem integrada da tecnologia de Processamento de Linguagem Natural

Welisângela Cardoso da Mata

Mestre em Direito pela UNESA.

Pós-Graduação em Legal Operation, Dados, Inteligência Artificial e Alta Performance Jurídica pela PUC-PR.

Pós-Graduação em Direito Tributário pela UniDF.

RESUMO

O artigo examina como as técnicas de Processamento de Linguagem Natural (NLP, na sigla em inglês) e modelos de *machine learning* transformam os processos de triagem e classificação de documentos jurídicos em ambientes de contencioso de massa e descoberta eletrônica (*eDiscovery*). As análises indicam que ferramentas baseadas em inteligência artificial conseguem reduzir significativamente o tempo de revisão de documentos, aumentar a precisão na categorização e identificar padrões em grandes volumes de dados jurídicos. Este trabalho sintetiza as tecnologias subjacentes, suas aplicações práticas no setor jurídico e os desafios associados à implementação responsável dessas soluções, estruturando uma análise que parte das tecnologias gerais de NLP para as implementações específicas em ambientes de litígio massificado.

Palavras-chave: Processamento de Linguagem Natural (NLP). Automação de Triagem. Contencioso de Massa. e-Discovery Jurídico.

ABSTRACT

This article examines how Natural Language Processing (NLP) techniques and machine learning models transform the screening and classification of legal documents in mass litigation and electronic discovery (e-discovery) contexts. Analyses indicate that AI-powered tools can significantly

reduce the time required for document review, improve the accuracy of categorization, and identify patterns in large volumes of legal data. This study synthesizes the underlying technologies, their practical applications in the legal sector, and the challenges associated with the responsible deployment of these solutions, structuring an analysis that begins with general NLP technologies and proceeds to specific implementations in mass-litigation environments.

Keywords: Natural Language Processing (NLP). Screening Automation. Mass Litigation. Legal e-discovery.

Introdução

O sistema judiciário brasileiro, como muitos sistemas de Justiça ao redor do mundo, enfrenta um crescimento exponencial no volume de processos judiciais. Esse fenômeno, particularmente intenso nas áreas de Direito consumerista, bancário, varejo e administrativo, criou um desafio estrutural tanto para as Cortes quanto para os operadores jurídicos: como gerenciar, revisar e analisar milhões de documentos com eficiência, precisão e custo-benefício adequados¹. O contencioso de massa, definido como um serviço jurídico executado em formato de linha de produção para empresas com passivos judiciais volumosos, representa uma realidade permanente nas grandes organizações².

Paralelamente, o desenvolvimento de procedimentos de descoberta eletrônica (*eDiscovery*) agregou complexidade adicional aos processos jurídicos. O *eDiscovery* refere-se ao processo de identificação, coleta e produção de informações armazenadas eletronicamente (ESI) em resposta a solicitações de produção em litígios ou investigações^{3 4}. À medida que documentos deixaram de ser predominantemente em papel para se concentrarem em formatos eletrônicos, como e-mails, planilhas, comu-

¹ AURUM. *O que é contencioso de massa e como funciona na prática!*. 2022. Disponível em: <https://www.aurum.com.br/blog/contencioso-de-massa/>. Acesso em: 20 mar. 2026.

² SERVICENOW. *What is eDiscovery?*. [s.d.]. Disponível em: <https://www.servicenow.com/br/products/legal-service-delivery/what-is-eDiscovery.html>. Acesso em: 22 mar. 2026

³ CDS LEGAL. *The basics: what is e-discovery*. [s.d.]. Disponível em: <https://cdslegal.com/knowledge/the-basics-what-is-e-discovery/>. Acesso em: 22 mar. 2026

IBM. *Natural language processing (NLP)*. [s.d.]. Disponível em: <https://www.ibm.com/br-pt/think/topics/natural-language-processing>. Acesso em: 22 mar. 2026.

⁴ FIRST LEGAL. *What is technology assisted review?*. [s.d.]. Disponível em: <https://www.firstlegal.com/what-is-technology-assisted-review/>. Acesso em: 22 mar. 2026

nações em redes sociais, áudio e vídeo, o volume de dados examinados em cada caso aumentou exponencialmente, tornando revisões puramente manuais impraticáveis tanto financeira quanto operacionalmente.

É neste contexto que o Processamento de Linguagem Natural (NLP, na sigla em inglês) emerge como ferramenta transformadora. O NLP permite que computadores e dispositivos digitais reconheçam, entendam e gerem texto e fala ao combinar linguística computacional com aprendizado de máquina⁵. Em ambientes de contencioso de massa e *eDiscovery*, o NLP possibilita a automação de etapas críticas do fluxo de trabalho jurídico: desde a triagem inicial de documentos até a classificação por relevância, a identificação de privilégios cliente-advogado e a extração de informações estruturadas de textos não estruturados.

Os benefícios comprovados dessa automação são substanciais. Estudos demonstram que ferramentas de revisão assistida por tecnologia (*Technology-Assisted Review*, TAR) baseadas em NLP e *machine learning* podem reduzir o tempo de revisão de documentos em até 80%, representando economias de custos equivalentes a 50 vezes maiores do que a revisão manual⁶. Para grandes escritórios e departamentos jurídicos corporativos, implementar essas tecnologias significa não apenas responder de forma mais rápida e precisa a demandas de litígio, mas também potencialmente transformar a área jurídica de um centro de custos em um parceiro estratégico que gera valor mensurável para a organização.

Todavia, a adoção de NLP no contexto jurídico também suscita questões fundamentais relacionadas à responsabilidade ética, transparência algorítmica, privacidade de dados e conformidade regulatória. A Lei Geral de Proteção de Dados (LGPD) brasileira, regulamentações internacionais como o Regulamento Geral de Proteção de Dados (GDPR) europeu, e normativas emergentes, como a Lei de Inteligência Artificial da União Europeia, estabelecem requisitos rigorosos sobre como dados pessoais podem ser processados por sistemas de IA^{7 8}. Adicionalmente, a comunidade jurídica tem demons-

⁵ KPMG. *Privacy in the new world of AI*. 2023. Disponível em: <https://assets.kpmg.com/content/dam/kpmg/br/pdf/2023/12/Privacy-inthenew-world-of-AI-Final.pdf>. Acesso em: 22 mar. 2026.

⁶ UPLEXIS. *Privacidade de dados*. [s.d.]. Disponível em: <https://uplexis.com.br/blog/artigos/privacidade-de-dados/>. Acesso em: 22 mar. 2026.

⁷ IBM. *Natural language processing (NLP)*. [s.d.]. Disponível em: <https://www.ibm.com/br-pt/think/topics/natural-language-processing>. Acesso em: 22 mar. 2026

⁸ IBM. *Natural language processing (NLP)*. [s.d.]. Disponível em: <https://www.ibm.com/br-pt/think/topics/natural-language-processing>. Acesso em: 22 mar. 2026

trado cautela legítima quanto à confiabilidade de decisões automatizadas, especialmente quando essas decisões impactam direitos fundamentais ou quando algoritmos perpetuam vieses preexistentes nos dados de treinamento.

Este artigo estrutura uma análise compreensiva dessa transformação tecnológica no setor jurídico, partindo das tecnologias gerais de NLP e *machine learning*, passando pelas aplicações específicas em triagem e classificação de documentos, e finalizando com uma reflexão sobre implementação responsável e desafios futuros. A estrutura segue um funil argumentativo, descendo do geral para o específico, permitindo que leitores menos familiarizados com NLP compreendam as bases tecnológicas antes de explorar aplicações jurídicas especializadas.

1 Fundamentos tecnológicos do Processamento de Linguagem Natural e automação de documentos jurídicos

1.1 Conceitos fundamentais de Processamento de Linguagem Natural

O Processamento de Linguagem Natural (NLP), conforme conceituado pela IBM, é um ramo da inteligência artificial que permite aos sistemas computacionais reconhecer, interpretar e gerar linguagem humana de forma inteligível⁹. Diferentemente de sistemas baseados exclusivamente em regras rígidas, o NLP moderno combina técnicas de linguística computacional com algoritmos sofisticados de aprendizado de máquina, permitindo que máquinas capturem não apenas o significado literal de palavras, mas também contexto, nuance, intenção e relacionamentos semânticos entre conceitos.

O processamento de textos em NLP segue tipicamente um *pipeline* estruturado, começando com o pré-processamento de texto¹⁰. Esta fase inicial transforma dados brutos em formatos que os modelos de *machine learning* conseguem processar eficientemente. O pré-processamento envolve várias operações complementares. A tokenização divide textos complexos em unidades menores significativas, como palavras, frases ou sentenças, quebrando estruturas complexas em partes manejáveis¹¹. A conver-

⁹ IBM. *Natural language processing (NLP)*. [s.d.]. Disponível em: <https://www.ibm.com/br-pt/think/topics/natural-language-processing>. Acesso em: 22 mar. 2026

¹⁰ IBM. *Natural language processing (NLP)*. [s.d.]. Disponível em: <https://www.ibm.com/br-pt/think/topics/natural-language-processing>. Acesso em: 22 mar. 2026

¹¹ IBM. *Natural language processing (NLP)*. [s.d.]. Disponível em: <https://www.ibm.com/br-pt/think/topics/natural-language-processing>. Acesso em: 22 mar. 2026

são para minúsculas padroniza o texto, garantindo que variações do mesmo termo como “Maçã” e “maçã” sejam tratadas identicamente¹². A remoção de palavras-chave irrelevantes (*stop words removal*) elimina termos frequentemente utilizados, como “é” ou “o”, que agregam pouco significado semântico¹³.

Complementando essas operações, o *stemming* ou lematização reduz palavras às suas formas raiz, como transformar “correndo” em “correr”, facilitando análises que agrupem diferentes variações morfológicas¹⁴. A limpeza de texto remove elementos indesejados, como pontuação, caracteres especiais e números que podem prejudicar análises posteriores¹⁵. Apenas após essas transformações o texto está adequadamente preparado para etapas subsequentes de análise semanticamente rica.

A extração de características (*feature extraction*) constitui o próximo estágio crítico, convertendo texto pré-processado em representações numéricas que sistemas de aprendizado de máquina conseguem analisar e interpretar¹⁶. Esse processo transforma linguagem natural em dados estruturados através de técnicas como *Bag of Words* e TF-IDF (*Term Frequency-Inverse Document Frequency*), as quais quantificam a presença e importância relativa de palavras em documentos¹⁷. Métodos mais sofisticados como *word embeddings* utilizam redes neurais para aprender representações densas, em que palavras e documentos são transformados em vetores de números reais, capturando relacionamentos semânticos complexos de forma compacta e matematicamente manipulável.

Após esses estágios preparatórios, técnicas de análise textual mais sofisticadas tornam-se viáveis. A marcação de partes do discurso (*Part-of-Speech tagging*) identifica à qual categoria gramatical cada palavra pertence, determinando se “dever” funciona como verbo em “ele vai dever para o banco” ou como subs-

¹² IBM. *Natural language processing (NLP)*. [s.d.]. Disponível em: <https://www.ibm.com/br-pt/think/topics/natural-language-processing>. Acesso em: 22 mar. 2026

¹³ IBM. *Natural language processing (NLP)*. [s.d.]. Disponível em: <https://www.ibm.com/br-pt/think/topics/natural-language-processing>. Acesso em: 22 mar. 2026

¹⁴ IBM. *Natural language processing (NLP)*. [s.d.]. Disponível em: <https://www.ibm.com/br-pt/think/topics/natural-language-processing>. Acesso em: 22 mar. 2026

¹⁵ IBM. *Natural language processing (NLP)*. [s.d.]. Disponível em: <https://www.ibm.com/br-pt/think/topics/natural-language-processing>. Acesso em: 22 mar. 2026

¹⁶ IBM. *Natural language processing (NLP)*. [s.d.]. Disponível em: <https://www.ibm.com/br-pt/think/topics/natural-language-processing>. Acesso em: 22 mar. 2026

¹⁷ ENCORD. *Named entity recognition*. [s.d.]. Disponível em: <https://encord.com/blog/named-entity-recognition/>. Acesso em: 22 mar. 2026

tantivo em “eu preciso fazer meu dever de casa”¹⁸. O Reconhecimento de Entidades Nomeadas (*Named Entity Recognition*, NER) identifica e classifica expressões específicas em categorias predefinidas, como pessoas, organizações, localizações, datas e quantidades monetárias¹⁹. Por exemplo, NER consegue identificar “Apple” como uma organização (ORG), “U.K.” como uma entidade geopolítica (GPE) e “\$1 bilhão” como um valor monetário (MONEY) dentro de textos²⁰.

O Entendimento de Linguagem Natural (*Natural Language Understanding*, NLU), um subconjunto especializado do NLP, concentra-se especificamente na análise do significado subjacente às frases²¹. O NLU permite que software identifique significados semanticamente similares em frases diferentes ou processe palavras que possuem múltiplos significados distintos, baseando-se em contexto. A modelagem de tópicos, outra capacidade relevante, identifica temas ou tópicos subjacentes dentro de textos ou coleções de documentos, permitindo descobrir estruturas temáticas não aparentes na leitura superficial.

1.2 Arquiteturas e modelos de *machine learning* para classificação de textos

A implementação prática do NLP em ambientes jurídicos depende fundamentalmente de arquiteturas e modelos de *machine learning*. Historicamente, o processamento estatístico de linguagem natural extraía, classificava e rotulava automaticamente elementos de dados textuais e de voz, atribuindo uma probabilidade estatística a cada possível significado, dependendo de aprendizado de máquina para análise sofisticada da linguística, como marcação de partes do discurso²².

Modelos *sequence-to-sequence* (seq2seq), baseados em redes neurais recorrentes (RNNs), revolucionaram certos tipos de processamento, sendo particularmente úteis para tarefas de tra-

¹⁸ ENCORD. *Named entity recognition*. [s.d.]. Disponível em: <https://encord.com/blog/named-entity-recognition/>. Acesso em: 22 mar. 2026

¹⁹ IBM. *Natural language processing (NLP)*. [s.d.]. Disponível em: <https://www.ibm.com/br-pt/think/topics/natural-language-processing>. Acesso em: 22 mar. 2026.

²⁰ IBM. *Natural language processing (NLP)*. [s.d.]. Disponível em: <https://www.ibm.com/br-pt/think/topics/natural-language-processing>. Acesso em: 22 mar. 2026.

²¹ IBM. *Natural language processing (NLP)*. [s.d.]. Disponível em: <https://www.ibm.com/br-pt/think/topics/natural-language-processing>. Acesso em: 22 mar. 2026.

²² ENAP. *[Artigo científico]*. 2023. Disponível em: <https://bibliotecadigital.enap.gov.br/bitstream/1/7716/1/8091-Texto%20do%20Artigo-31686-1-10-20230807.pdf>. Acesso em: 22 mar. 2026

dução automática, convertendo frases de um domínio, por exemplo, a língua alemã em frases de outro domínio como o inglês²³. Contudo, RNNs convencionais apresentam limitações significativas quando processam sequências muito longas, pois o gradiente de erro tende a se dissipar ou explodir durante o treinamento.

A introdução de mecanismos de atenção e, posteriormente, de arquiteturas baseadas em Transformers representou um avanço decisivo. Os Transformers, particularmente o BERT (*Bidirectional Encoder Representations from Transformers*), desenvolvido pelo Google, empregam múltiplas camadas de *Transformers Encoder* bidirecional, utilizando uma arquitetura de rede neural mais profunda e pré-treinada com *corpus* de texto massivo²⁴. O BERT aprende representações textuais lexicais através de previsão de palavras condicionadas ao seu contexto, permitindo capturar relações bidirecionais entre palavras de maneira muito mais eficaz que modelos anteriores²⁵.

Um aspecto crucial do BERT é sua arquitetura de pré-treinamento e ajuste fino (*fine-tuning*). Na fase de pré-treinamento, o modelo é treinado com *dataset* massivo não rotulado desde o início²⁶. Como o BERT utiliza *transfer learning*, é possível usar um modelo pré-treinado como ponto de partida, carregando esse modelo pré-treinado para ajuste fino, no qual todos os parâmetros são sintonizados usando um *dataset* específico de tarefas especializadas²⁷. Essa abordagem provou-se particularmente poderosa para tarefas jurídicas, em que dados rotulados são frequentemente escassos, pois o conhecimento aprendido em *corpus* de texto geral transfere-se eficazmente para domínios legais especializados.

Para contextos jurídicos especificamente, modelos especializados como LegalBERT, CaseLaw-BERT e outras variantes treinadas

²³ ENAP. [Artigo científico]. 2023. Disponível em: <https://bibliotecadigital.enap.gov.br/bitstream/1/7716/1/8091-Texto%20do%20Artigo-31686-1-10-20230807.pdf>. Acesso em: 22 mar. 2026

²⁴ ENAP. [Artigo científico]. 2023. Disponível em: <https://bibliotecadigital.enap.gov.br/bitstream/1/7716/1/8091-Texto%20do%20Artigo-31686-1-10-20230807.pdf>. Acesso em: 22 mar. 2026

²⁵ ENAP. [Artigo científico]. 2023. Disponível em: <https://bibliotecadigital.enap.gov.br/bitstream/1/7716/1/8091-Texto%20do%20Artigo-31686-1-10-20230807.pdf>. Acesso em: 22 mar. 2026

²⁶ ARXIV. [Artigo científico]. 2025. Disponível em: <https://arxiv.org/html/2504.00725v1>. Acesso em: 22 mar. 2026

²⁷ ENAP. [Artigo científico]. 2023. Disponível em: <https://bibliotecadigital.enap.gov.br/bitstream/1/7716/1/8091-Texto%20do%20Artigo-31686-1-10-20230807.pdf>. Acesso em: 22 mar. 2026.

em *corpora* jurídicos demonstraram desempenho superior, comparado a modelos genéricos²⁸. Um estudo avaliando modelos especializados em documentos jurídicos identificou que BERT com ajuste fino utilizando trechos de leis brasileiras (BERTIMBAU-2) alcançou precisão de 94% em tarefas de Reconhecimento de Entidades Nomeadas, Similaridade Textual em Sentenças e Reconhecimento de Entidade Textual²⁹. Esse desempenho superior de modelos especializados ocorre porque o treinamento em linguagem jurídica captura terminologia técnica, estruturas argumentativas e conceitos jurídicos que modelos genéricos desconhecem.

Complementando Transformers, modelos *Long-formers* e *BigBird* foram desenvolvidos especificamente para processar documentos longos, aumentando o comprimento máximo de sequências de entrada de 512 subpalavras em BERT para até 4.096 subpalavras, abordando diretamente o desafio de documentos jurídicos que frequentemente excedem os limites de processamento de modelos anteriores³⁰. Mesmo modelos de *sparse-attention* como *Longformer*, entretanto, ainda truncam severos textos em muitos cenários jurídicos complexos, tornando necessárias estratégias adicionais, como segmentação hierárquica de documentos ou abordagens que combinam modelos transformers com representações TF-IDF mais tradicionais³¹.

A escolha entre diferentes arquiteturas repousa em considerações práticas e teóricas. *Random Forests*, histórico algoritmo de *ensemble learning*, que constrói múltiplas árvores de decisão treinadas em subconjuntos aleatórios dos dados de treinamento, consegue manejar dados altamente dimensionais e não lineares, oferecendo interpretabilidade moderada e ressaltando a importância relativa de características³². *Support Vector Machines* (SVMs) otimizam para margem máxima, oferecendo decisão clara de fronteiras, mas com menor interpretabilidade e potencial

²⁸ ARXIV. [Artigo científico]. 2022. Disponível em: <https://arxiv.org/abs/2211.00974>. Acesso em: 22 mar. 2026.

²⁹ ARXIV. [Artigo científico]. 2022. Disponível em: <https://arxiv.org/abs/2211.00974>. Acesso em: 22 mar. 2026.

³⁰ GEEKSFORGEES. *When to use random forest over SVM and vice versa*. [s.d.]. Disponível em: <https://www.geeksforgeeks.org/machine-learning/when-to-use-random-forest-over-svm-and-vice-versa/>. Acesso em: 22 mar. 2026.

³¹ GEEKSFORGEES. *When to use random forest over SVM and vice versa*. [s.d.]. Disponível em: <https://www.geeksforgeeks.org/machine-learning/when-to-use-random-forest-over-svm-and-vice-versa/>. Acesso em: 22 mar. 2026.

³² GEEKSFORGEES. *When to use random forest over SVM and vice versa*. [s.d.]. Disponível em: <https://www.geeksforgeeks.org/machine-learning/when-to-use-random-forest-over-svm-and-vice-versa/>. Acesso em: 22 mar. 2026.

complexidade computacional, aumentada em *datasets* muito grandes³³. Para *datasets* pequenos a médios com padrões não lineares moderados, SVMs frequentemente superam *Random Forests*, enquanto *Random Forests* brilham em *datasets* grandes com padrões altamente complexos e não lineares³⁴.

1.3 Abordagens de validação e métricas de desempenho

A validação robusta de modelos de NLP constitui uma etapa crítica antes da implementação em ambientes de produção jurídica de alto risco. Validação de modelo é o processo de avaliação de desempenho de modelos treinados para garantir cumprimento de padrões desejados de precisão, confiabilidade e imparcialidade³⁵. Envolve testar modelos em dados não vistos para avaliar a capacidade de generalização e identificar potenciais problemas como sobreajuste ou enviesamento³⁶.

A validação cruzada (*k-fold cross-validation*) emerge como metodologia fundamental. Esse método divide dados em *k* subconjuntos e treina o modelo *k* vezes, utilizando cada vez um subconjunto diferente para validação³⁷. Os valores comuns para *k* variam entre 5 e 10³⁸. *Leave-One-Out Cross-Validation* (LOOCV) representa uma variante extrema, em que uma única observação é retida para validação, enquanto o modelo treina no restante dos dados, repetindo esse processo para cada ponto de dados³⁹. Embora computacionalmente custoso, LOOCV oferece

³³ FOCALX. *Validação de modelos de IA: garantir a exatidão e a fiabilidade*. [s.d.]. Disponível em: <https://focalx.ai/pt-pt/inteligencia-artificial/validacao-de-modelos-de-ia-garantir-a-exatidao-e-a-fiabilidade/>. Acesso em: 22 mar. 2026.

³⁴ FOCALX. *Validação de modelos de IA: garantir a exatidão e a fiabilidade*. [s.d.]. Disponível em: <https://focalx.ai/pt-pt/inteligencia-artificial/validacao-de-modelos-de-ia-garantir-a-exatidao-e-a-fiabilidade/>. Acesso em: 22 mar. 2026.

³⁵ FOCALX. *Validação de modelos de IA: garantir a exatidão e a fiabilidade*. [s.d.]. Disponível em: <https://focalx.ai/pt-pt/inteligencia-artificial/validacao-de-modelos-de-ia-garantir-a-exatidao-e-a-fiabilidade/>. Acesso em: 22 mar. 2026.

³⁶ HACKING ANALYTICS. *Resampling: separando os dados entre treino, validação e teste*. 2016. Disponível em: <https://hackinganalytics.wordpress.com/2016/09/04/resampling-separando-os-dados-entre-treino-validacao-e-teste/>. Acesso em: 22 mar. 2026.

³⁷ FOCALX. *Validação de modelos de IA: garantir a exatidão e a fiabilidade*. [s.d.]. Disponível em: <https://focalx.ai/pt-pt/inteligencia-artificial/validacao-de-modelos-de-ia-garantir-a-exatidao-e-a-fiabilidade/>. Acesso em: 22 mar. 2026.

³⁸ BRAINS. *Medidas de performance de modelos de classificação*. 2023. Disponível em: <https://brains.dev/2023/medidas-de-performance-modelos-de-classificacao/>. Acesso em: 22 mar. 2026.

³⁹ LABELF. *What is accuracy, precision, recall and F1 score*. [s.d.]. Disponível em: <https://www.labelf.ai/blog/what-is-accuracy-precision-recall-and-f1-score>. Acesso em: 22 mar. 2026.

estimativas menos enviesadas de desempenho, particularmente valiosa em contextos jurídicos críticos.

A matriz de confusão fornece visualização fundamental dos erros e acertos de modelos de classificação. Esta matriz tabulava previsões do modelo *versus* resultados reais, permitindo identificar falsos positivos e falsos negativos. Verdadeiros Positivos (TP) representam casos em que o modelo previu corretamente a classe positiva, enquanto Verdadeiros Negativos (TN) indicam previsões negativas corretas. Falsos Positivos (FP) representam erros em que o modelo previu incorretamente a classe positiva, e Falsos Negativos (FN) indicam erros em que o modelo previu incorretamente a classe negativa⁴⁰.

Com a matriz de confusão disponível, diversas métricas derivadas permitem quantificar aspectos específicos do desempenho de modelos de classificação. A acurácia, por exemplo, responde à pergunta: “De todas as predições realizadas, quantas foram corretas?”⁴¹. Embora intuitiva, essa métrica pode ser enganosa em conjuntos de dados desbalanceados, nos quais uma classe predomina substancialmente.⁴² Já a precisão reflete: “De todas as predições positivas realizadas, quantas foram corretas?”⁴³, calculada pela fórmula TP/(TP + FP)⁴⁴. Essa métrica ganha relevância crítica em cenários nos quais o custo de falsos positivos é elevado, como diagnósticos médicos ou na classificação de documentos como importantes, quando, na realidade, são irrelevantes, o que implica desperdício de recursos em revisões desnecessárias⁴⁵.

O Recall, também denominado Sensibilidade ou Taxa de Verdadeiros Positivos, mede “De todos os casos que deveriam ser

⁴⁰ LABELF. *What is accuracy, precision, recall and F1 score*. [s.d.]. Disponível em: <https://www.labelf.ai/blog/what-is-accuracy-precision-recall-and-f1-score>. Acesso em: 22 mar. 2026.

⁴¹ BRAINS. *Medidas de performance de modelos de classificação*. 2023. Disponível em: <https://brains.dev/2023/medidas-de-performance-modelos-de-classificacao/>. Acesso em: 22 mar. 2026

⁴² LABELF. *What is accuracy, precision, recall and F1 score*. [s.d.]. Disponível em: <https://www.labelf.ai/blog/what-is-accuracy-precision-recall-and-f1-score>. Acesso em: 22 mar. 2026.

⁴³ LABELF. *What is accuracy, precision, recall and F1 score*. [s.d.]. Disponível em: <https://www.labelf.ai/blog/what-is-accuracy-precision-recall-and-f1-score>. Acesso em: 22 mar. 2026.

⁴⁴ BRAINS. *Medidas de performance de modelos de classificação*. 2023. Disponível em: <https://brains.dev/2023/medidas-de-performance-modelos-de-classificacao/>. Acesso em: 22 mar. 2026.

⁴⁵ LABELF. *What is accuracy, precision, recall and F1 score*. [s.d.]. Disponível em: <https://www.labelf.ai/blog/what-is-accuracy-precision-recall-and-f1-score>. Acesso em: 22 mar. 2026.

preditos como positivos, quantos foram corretamente preditos?”⁴⁶. Recall é essencial quando perder casos positivos é mais custoso que falsos positivos, como em contextos jurídicos em que deixar passar documentos criticamente importantes pode prejudicar uma estratégia legal. Existe frequentemente um *trade-off* fundamental entre Precisão e Recall: aumentar Recall tipicamente reduz Precisão e vice-versa⁴⁷.

O F1-Score harmoniza Precisão e Recall, calculado como $2 \times (\text{Precisão} \times \text{Recall}) / (\text{Precisão} + \text{Recall})$ ⁴⁸. F1-Score fornece medida única balanceada particularmente útil em *datasets* desbalanceados, em que uma métrica isolada seria insuficiente⁴⁹. A Área sob a Curva ROC (AUC-ROC) quantifica o desempenho geral de classificadores através de diferentes limiares, respondendo à questão “Com qual frequência o modelo rankeia uma observação positiva aleatória superior a uma observação negativa aleatória?”⁵⁰.

Para cenários jurídicos específicos, questões pragmáticas devem guiar a seleção de métricas. Na revisão de contencioso de massa, na qual perder documentos potencialmente relevantes poderia prejudicar significativamente a estratégia, Recall pode ser priorizado. Em contextos e-Discovery, em que o custo de revisar manualmente documentos classificados como relevantes, mas que não é extremamente elevado, Precisão pode ser preferível. Frequentemente, uma abordagem híbrida buscando equilíbrio através de F1-Score oferece solução pragmática razoável.

1.4 Tratamento de dados desbalanceados e enviesamento algorítmico

Um desafio técnico importante emerge em aplicações jurídicas reais: dados desbalanceados. Na classificação desbalance-

⁴⁶ LABELF. *What is accuracy, precision, recall and F1 score*. [s.d.]. Disponível em: <https://www.labelf.ai/blog/what-is-accuracy-precision-recall-and-f1-score>. Acesso em: 22 mar. 2026.

⁴⁷ BRAINS. *Medidas de performance de modelos de classificação*. 2023. Disponível em: <https://brains.dev/2023/medidas-de-performance-modelos-de-classificacao/>. Acesso em: 22 mar. 2026.

⁴⁸ LABELF. *What is accuracy, precision, recall and F1 score*. [s.d.]. Disponível em: <https://www.labelf.ai/blog/what-is-accuracy-precision-recall-and-f1-score>. Acesso em: 22 mar. 2026.

⁴⁹ LABELF. *What is accuracy, precision, recall and F1 score*. [s.d.]. Disponível em: <https://www.labelf.ai/blog/what-is-accuracy-precision-recall-and-f1-score>. Acesso em: 22 mar. 2026.

⁵⁰ BRAINS. *Medidas de performance de modelos de classificação*. 2023. Disponível em: <https://brains.dev/2023/medidas-de-performance-modelos-de-classificacao/>. Acesso em: 22 mar. 2026.

ada, o número de exemplos distribui-se desigualmente entre classes, significando que o modelo treina em mais exemplos de uma classe que de outras⁵¹. Em contencioso de massa consumerista, por exemplo, pode existir proporção 60% de processos relacionados a cobranças, 25% relacionados a vício de produtos e apenas 15% relacionados a outras questões, e, por consequência, treinar modelos em dados desbalanceados, o que tipicamente resulta em viés sistemático, em que o modelo aprende a favor da classe majoritária.

Técnicas de reamostragem enfrentam esse desafio. Subamostragem aleatória elimina aleatoriamente exemplos da classe majoritária, reduzindo proporcionalidade de desbalanceamento, mas potencialmente perdendo informação valiosa⁵². Sobreamostragem baseada em SMOTE (*Synthetic Minority Over-sampling Technique*) replica sinteticamente exemplos da classe minoritária através de interpolação, aumentando efetivamente o tamanho da classe minoritária sem simplesmente duplicar exemplos⁵³. Técnicas híbridas combinam subamostragem da classe majoritária com sobreamostragem sintética da classe minoritária, frequentemente produzindo desempenho superior.

Além do desbalanceamento estatístico, enviesamento algorítmico representa preocupação ética e prática crítica. Modelos de IA conseguem, inadvertidamente, perpetuar preconceitos presentes nos dados de treinamento, conduzindo a resultados discriminatórios^{54 55}. Se dados históricos de decisões judiciais usados para treinar modelos refletem discriminação passada contra grupos específicos, o modelo treinado nesses dados replicará esse padrão discriminatório em novas predições. Testes sistemáticos

⁵¹ DATACAMP. *Classification in machine learning*. [s.d.]. Disponível em: <https://www.datacamp.com/pt/blog/classification-machine-learning>. Acesso em: 22 mar. 2026.

⁵² DATACAMP. *Classification in machine learning*. [s.d.]. Disponível em: <https://www.datacamp.com/pt/blog/classification-machine-learning>. Acesso em: 22 mar. 2026.

⁵³ DATACAMP. *Classification in machine learning*. [s.d.]. Disponível em: <https://www.datacamp.com/pt/blog/classification-machine-learning>. Acesso em: 22 mar. 2026.

⁵⁴ KPMG. *Privacy in the new world of AI*. 2023. Disponível em: <https://assets.kpmg.com/content/dam/kpmg/br/pdf/2023/12/Privacy-in-the-new-world-of-AI-Final.pdf>. Acesso em: 22 mar. 2026.

⁵⁵ UPLEXIS. *Privacidade de dados*. [s.d.]. Disponível em: <https://uplexis.com.br/blog/artigos/privacidade-de-dados/>. Acesso em: 22 mar. 2026.

de viés devem avaliar o desempenho de modelos em diferentes grupos demográficos ou cenários para garantir equidade^{56 57}.

2 Aplicações especializadas em triagem, classificação de documentos jurídicos e eDiscovery

2.1 Revisão assistida por tecnologia e redução de custos em eDiscovery

A descoberta eletrônica (*eDiscovery*) constitui processo crítico em litígios modernos, envolvendo identificação, coleta e produção de informações armazenadas eletronicamente em resposta a solicitações legais^{58 59}. Diferentemente de episódios de descoberta tradicionais baseados em papel, *eDiscovery* deve contabilizar plenamente o arquivamento sem papel, processando emails, documentos digitais, publicações em redes sociais, arquivos multimídia, bancos de dados e uma variedade de formatos eletrônicos⁶⁰. O volume de dados envolvido é frequentemente astronômico: investigações de grande escala envolvem dezenas de milhões de documentos.

Neste contexto, a Revisão Assistida por Tecnologia (*Technology-Assisted Review*, TAR), também denominada codificação preditiva ou revisão assistida por computador, emerge como inovação transformadora^{61 62}. TAR utiliza inteligência artificial para desenvolver software que continua aprendendo e tomando decisões melho-

⁵⁶ FOCALX. *Validação de modelos de IA: garantir a exatidão e a fiabilidade*. [s.d.]. Disponível em: <https://focalx.ai/pt-pt/inteligencia-artificial/validacao-de-modelos-de-ia-garantir-a-exatidao-e-a-fiabilidade/>. Acesso em: 22 mar. 2026.

⁵⁷ PÁDUA, Sérgio Rodrigo de; LORENZETTO, Bruno Meneses. O DIREITO FUNDAMENTAL À EXPLICABILIDADE DA INTELIGÊNCIA ARTIFICIAL UTILIZADA EM DECISÕES ESTATAIS. *Revista da AGU - Brasília-DF - v. 23 - n. 02 - jun/2024*, 349-371. <https://revistaagu.agu.gov.br/index.php/AGU/article/view/3480/2934>.

⁵⁸ SERVICENOW. *What is eDiscovery?*. [s.d.]. Disponível em: <https://www.servicenow.com/br/products/legal-service-delivery/what-is-eDiscovery.html>. Acesso em: 22 mar. 2026.

⁵⁹ CDSLEGAL. *The basics: what is e-discovery*. [s.d.]. Disponível em: <https://cdslegal.com/knowledge/the-basics-what-is-e-discovery/>. Acesso em: 22 mar. 2026.

⁶⁰ SERVICENOW. *What is eDiscovery?*. [s.d.]. Disponível em: <https://www.servicenow.com/br/products/legal-service-delivery/what-is-eDiscovery.html>. Acesso em: 22 mar. 2026.

⁶¹ FIRST LEGAL. *What is technology assisted review?*. [s.d.]. Disponível em: <https://www.firstlegal.com/what-is-technology-assisted-review/>. Acesso em: 22 mar. 2026.

⁶² THOMSON REUTERS. *How predictive coding makes e-discovery more efficient*. [s.d.]. Disponível em: <https://legal.thomsonreuters.com/en/insights/articles/how-predictive-coding-makes-e-discovery-more-efficient>. Acesso em: 22 mar. 2026.

res, enquanto acelera significativamente o processo de revisão, economizando tempo e dinheiro⁶³. O processo começa com treinamento de software usando um *seed set* (ou conjunto de sementes, uma amostra de documentos) extraída do grupo total que requer revisão⁶⁴. Revisores, frequentemente advogados altamente qualificados com ampla experiência, codificam cada documento nesse *seed set* como responsivo (relevante ao caso) ou não responsivo (não relevante)⁶⁵.

Conforme o treinamento progride, a inteligência artificial permite que o software aprenda e realize decisões progressivamente melhores e mais rápidas⁶⁶. O sistema desenvolve compreensão dos critérios de relevância aplicados pelos revisores humanos ao analisar características dos documentos *seed set*, depois generaliza esse conhecimento para classificar automaticamente o restante do *corpus* documental. Essa abordagem provou-se bastante eficaz: um estudo seminal⁶⁷ de Maura R. Grossman e Gordon V. Cormack (GROSSMAN; CORMACK, 2011) demonstrou que a Revisão Assistida por Tecnologia (TAR) pode produzir economia de custos de magnitude muitas vezes superior à revisão manual, reduzindo de forma drástica o tempo necessário para revisão de documentos⁶⁸ (GRIMM, GROSSMAN e CORMACK, 2021).

Os benefícios práticos de TAR manifestam-se em múltiplas dimensões. Do ponto de vista operacional, TAR permite que equipes jurídicas estabeleçam estratégias eficientes tanto antes quanto durante a revisão. Porque os modelos preditivos conseguem ranquear documentos por probabilidade de relevância, o

⁶³ THOMSON REUTERS. *How predictive coding makes e-discovery more efficient*. [s.d.]. Disponível em: <https://legal.thomsonreuters.com/en/insights/articles/how-predictive-coding-makes-e-discovery-more-efficient>. Acesso em: 22 mar. 2026.

⁶⁴ THOMSON REUTERS. *How predictive coding makes e-discovery more efficient*. [s.d.]. Disponível em: <https://legal.thomsonreuters.com/en/insights/articles/how-predictive-coding-makes-e-discovery-more-efficient>. Acesso em: 22 mar. 2026.

⁶⁵ THOMSON REUTERS. *How predictive coding makes e-discovery more efficient*. [s.d.]. Disponível em: <https://legal.thomsonreuters.com/en/insights/articles/how-predictive-coding-makes-e-discovery-more-efficient>. Acesso em: 22 mar. 2026.

⁶⁶ THOMSON REUTERS. *How predictive coding makes e-discovery more efficient*. [s.d.]. Disponível em: <https://legal.thomsonreuters.com/en/insights/articles/how-predictive-coding-makes-e-discovery-more-efficient>. Acesso em: 22 mar. 2026.

⁶⁷ GROSSMAN, Maura R.; CORMACK, Gordon V. *Technology-assisted review in e-discovery can be more effective and more efficient than exhaustive manual review*. *Richmond Journal of Law and Technology*, Richmond, v. 17, n. 3, 2011. Disponível em: <http://jolt.richmond.edu/v17i3/article11.pdf>. Acesso em: 22 dez. 2025.

⁶⁸ GRIMM, Paul W.; GROSSMAN, Maura R.; CORMACK, Gordon V. *Artificial intelligence as evidence*. *Nw. J. Tech. & Intell. Prop.*, v. 19, p. 9, 2021.

software traz à atenção dos revisores, primeiro, os documentos com maior probabilidade de serem significativos, enquanto documentos com baixa probabilidade de relevância podem ser excluídos completamente da revisão humana⁶⁹. Essa priorização elimina o tempo desperdiçado examinando manualmente documentos claramente irrelevantes.

Adicionalmente, TAR possibilita análise de documentos em base contínua, de modo que equipes jurídicas não necessitam aguardar até que toda coleta de dados relevantes esteja completa, antes de iniciarem análise de como certos documentos serão significativos ao caso⁷⁰. Esta abordagem *rolling* reduz cronogramas críticos de litígio. Quando o outro lado do processo, então, produz documentos em resposta ao processo de descoberta, a aplicação de codificação preditiva a essa documentação oposta permite que equipes jurídicas rapidamente decidam se moções adicionais precisam ser apresentadas⁷¹.

Desse modo, a Revisão Assistida por Tecnologia (TAR) não apenas reduz custos e tempo de revisão, como demonstrado por Grossman e Cormack (estudo seminal), mas também supera a revisão manual exaustiva em precisão e recall, conforme validação imparcial recente (Cormack *et al.*, SIGIR 2024), que, em um estudo de caso envolvendo 13 milhões de registros de um importante litígio de seguros, demonstra que o modelo automatizado superou a precisão de advogados experientes. Essas descobertas sugerem que metas de recall fixas são menos úteis do que a validação comparativa para comprovar que um esforço de descoberta de provas é juridicamente razoável e proporcional. Grimm *et al.* (2021) enfatizam admissibilidade sob Daubert: testes independentes, taxa de erro baixa e padrões controlados validam TAR como evidência confiável, mitigando vieses via amostragem estratificada.

No Brasil, essa perspectiva alinha-se à Resolução CNJ 332/2020 e promove transparência algorítmica e proporcionalidade (CPC art. 139), transformando *eDiscovery* de custo em estratégia eficiente.

⁶⁹ FIRST LEGAL. *What is technology assisted review?*. [s.d.]. Disponível em: <https://www.firstlegal.com/what-is-technology-assisted-review/>. Acesso em: 22 mar. 2026.

⁷⁰ FIRST LEGAL. *What is technology assisted review?*. [s.d.]. Disponível em: <https://www.firstlegal.com/what-is-technology-assisted-review/>. Acesso em: 22 mar. 2026.

⁷¹ FIRST LEGAL. *What is technology assisted review?*. [s.d.]. Disponível em: <https://www.firstlegal.com/what-is-technology-assisted-review/>. Acesso em: 22 mar. 2026.

2.2 Ferramenta Berna: IA para identificação de litigância abusiva e demandas em massa

Um caso de aplicação específica brasileira ilustra como NLP transforma operações jurídicas em escala institucional. A ferramenta Berna (Busca Eletrônica Recursiva usando Linguagem Natural) foi desenvolvida pelo Tribunal de Justiça de Goiás e disponibilizada através do Jus.br e da Plataforma Digital do Poder Judiciário via Conecta, uma iniciativa do Programa Justiça 4.0⁷². A Berna demonstra como NLP automatiza etapas críticas de processamento de demandas judiciais massificadas.

A Berna recebe documentos de petições iniciais em formato digital, como PDFs ou DOCs, e extrai informações automaticamente⁷³. Em seguida, utiliza métricas para cálculo de similaridade e compara o conteúdo das petições com um banco de dados de casos anteriores⁷⁴. Com base nessas comparações, identifica conjuntos de petições iniciais com alta similaridade, agrupa processos semelhantes e classifica a probabilidade de ocorrência de demandas abusivas e demandas em massa⁷⁵.

A ferramenta foi desenvolvida como agnóstica e compatível com qualquer sistema processual, utilizando IA para analisar e agrupar petições iniciais que chegam às Cortes, apontando padrões de repetição e indícios de uso inadequado da Justiça⁷⁶. Essa solução visa automatizar etapas de processos, transformar a tramitação processual e garantir agilidade à prestação jurisdicional⁷⁷. Na PDPJ-Br, a Berna é acessível através de interface web ou API (Interface de Programação de Aplicações), permitindo listagem de grupos de processos

⁷² MIGALHAS. *IA que identifica litigância abusiva é disponibilizada para tribunais*. 2025. Disponível em: <https://www.migalhas.com.br/quentes/446889/ia-que-identifica-litigancia-abusiva-e-disponibilizada-para-tribunais>. Acesso em: 22 mar. 2026.

⁷³ MIGALHAS. *IA que identifica litigância abusiva é disponibilizada para tribunais*. 2025. Disponível em: <https://www.migalhas.com.br/quentes/446889/ia-que-identifica-litigancia-abusiva-e-disponibilizada-para-tribunais>. Acesso em: 22 mar. 2026.

⁷⁴ MIGALHAS. *IA que identifica litigância abusiva é disponibilizada para tribunais*. 2025. Disponível em: <https://www.migalhas.com.br/quentes/446889/ia-que-identifica-litigancia-abusiva-e-disponibilizada-para-tribunais>. Acesso em: 22 mar. 2026.

⁷⁵ MIGALHAS. *IA que identifica litigância abusiva é disponibilizada para tribunais*. 2025. Disponível em: <https://www.migalhas.com.br/quentes/446889/ia-que-identifica-litigancia-abusiva-e-disponibilizada-para-tribunais>. Acesso em: 22 mar. 2026.

⁷⁶ MIGALHAS. *IA que identifica litigância abusiva é disponibilizada para tribunais*. 2025. Disponível em: <https://www.migalhas.com.br/quentes/446889/ia-que-identifica-litigancia-abusiva-e-disponibilizada-para-tribunais>. Acesso em: 22 mar. 2026.

⁷⁷ MIGALHAS. *IA que identifica litigância abusiva é disponibilizada para tribunais*. 2025. Disponível em: <https://www.migalhas.com.br/quentes/446889/ia-que-identifica-litigancia-abusiva-e-disponibilizada-para-tribunais>. Acesso em: 22 mar. 2026.

similares por tribunal, por processos, por demandas abusivas e/ou por demandas em massa⁷⁸.

A Berna represente um avanço significativo porque endereça diretamente o problema de litigância abusiva que paralisa tribunais em várias regiões do Brasil. O Superior Tribunal de Justiça, em recurso repetitivo recente (Tema 1.198), fixou precedente estabelecendo que juízes conseguem, de modo fundamentado e com observância à razoabilidade do caso concreto, exigir emenda da petição inicial a fim de demonstrar interesse de agir e autenticidade da postulação para coibir litigância abusiva⁷⁹. A Berna oferece ferramenta tecnológica para realizar esta triagem inicial em escala, identificando padrões indicativos de litigância abusiva, que provavelmente escapariam à identificação manual.

2.3 Casos de sucesso em contencioso de massa com automação

Implementações reais de automação em contencioso de massa documentam ganhos tangíveis. Um caso em setor educacional demonstrou que, com automação da análise inicial e uso de IA para classificação de ações por perfil de risco e jurisprudência dominante, foi possível reduzir 62% o tempo médio de resposta, com economia anual estimada em R\$ 1,8 milhão em honorários advocatícios externos, mantendo qualidade⁸⁰. Uma instituição financeira enfrentando volume médio de 1.500 novos processos por mês em Juizados Especiais implementou IA para geração de petições de baixa complexidade e automatização da triagem processual, resultando em redução de 30% no custo por demanda, com economia superior a R\$ 6 milhões anuais⁸¹. Um

⁷⁸ MIGALHAS. *Contencioso de massa com IA: eficiência com responsabilidade*. [s.d.]. Disponível em: <https://www.migalhas.com.br/depeso/435551/contencioso-de-massa-com-ia-eficiencia-com-responsabilidade>. Acesso em: 22 mar. 2026.

⁷⁹ SUPERIOR TRIBUNAL DE JUSTIÇA. *Corte Especial decide que juiz pode exigir documentos para coibir litigância abusiva (REsp 2021665)*. 2025. Disponível em: SUPREIOR TRIBUNAL DE JUSTIÇA. *Corte Especial decide que juiz pode exigir documentos para coibir litigância abusiva (REsp 2021665)*. 2025. Disponível em: <https://www.stj.jus.br/sites/porta/p/Paginas/Comunicacao/Noticias/2025/20032025-Corte-Especial-decide-em-repetitivo-que-juiz-pode-exigir-documentos-para-coibir-litigancia-abusiva.aspx>. Acesso em: 22 mar. 2026. (REsp 2021665)

⁸⁰ MIGALHAS. *Contencioso de massa com IA: eficiência com responsabilidade*. [s.d.]. Disponível em: <https://www.migalhas.com.br/depeso/435551/contencioso-de-massa-com-ia-eficiencia-com-responsabilidade>. Acesso em: 22 mar. 2026.

⁸¹ MIGALHAS. *Contencioso de massa com IA: eficiência com responsabilidade*. [s.d.]. Disponível em: <https://www.migalhas.com.br/depeso/435551/contencioso-de-massa-com-ia-eficiencia-com-responsabilidade>. Acesso em: 22 mar. 2026.

conglomerado varejista e e-commerce com milhares de ações sobre vício de produto e cobranças indevidas de frete padronizou argumentação via modelos parametrizados por IA, obtendo redução de 41% no número de condenações e economia de R\$ 3,4 milhões em indenizações evitadas, além de maior uniformidade na tese institucional⁸².

Esses casos demonstram que a automação não funciona em abstrato, mas quando implementada com critério rigoroso baseado em modelos híbridos, em que a automação é utilizada com segurança de dados e validação humana criteriosa, resulta em valor mensurável. Nos casos citados, nenhuma petição crítica foi protocolada sem validação humana, mesmo quando gerada por IA⁸³. Dados processuais e de consumidores foram tratados em conformidade com a legislação de proteção de dados, e todos os fluxos automatizados mantiveram-se auditáveis e controláveis por comitês internos e auditorias⁸⁴.

2.4 Extração de informação e metadados em Gestão de Documentos

Além de triagem e classificação, NLP capacita extração sofisticada de informação estruturada a partir de textos não estruturados. O Reconhecimento de Entidades Nomeadas consegue identificar atores jurídicos críticos como partes litigantes, juízes, órgãos governamentais, datas de audiências e valores em disputa⁸⁵. Sistemas NLP conseguem varrer documentos para identificar cláusulas específicas em contratos, termos críticos em acordos e inconsistências entre documentos relacionados. Essa capacidade de extração automatizada reduz significativamente o tempo gasto em revisão manual de características críticas de documentos.

A extração de metadados representa extensão importante dessa capacidade. Metadados englobam informações estruturadas sobre documentos, como número do caso, nome do cliente, tipo de do-

⁸² MIGALHAS. *Contencioso de massa com IA: eficiência com responsabilidade*. [s.d.]. Disponível em: <https://www.migalhas.com.br/depeso/435551/contencioso-de-massa-com-ia-eficiencia-com-responsabilidade>. Acesso em: 22 mar. 2026.

⁸³ MIGALHAS. *Contencioso de massa com IA: eficiência com responsabilidade*. [s.d.]. Disponível em: <https://www.migalhas.com.br/depeso/435551/contencioso-de-massa-com-ia-eficiencia-com-responsabilidade>. Acesso em: 22 mar. 2026.

⁸⁴ MIGALHAS. *Contencioso de massa com IA: eficiência com responsabilidade*. [s.d.]. Disponível em: <https://www.migalhas.com.br/depeso/435551/contencioso-de-massa-com-ia-eficiencia-com-responsabilidade>. Acesso em: 22 mar. 2026.

⁸⁵ ENCORD. *Named entity recognition*. [s.d.]. Disponível em: <https://encord.com/blog/named-entity-recognition/>. Acesso em: 22 mar. 2026.

cumento, advogado responsável, data de criação, palavras-chave e parâmetros críticos⁸⁶. Esses elementos servem como balizador navegacional, guiando profissionais jurídicos a documentos relevantes prontamente⁸⁷. O gerenciamento competente de metadados é fundamental para alcançar a excelência em gestão de documentos jurídicos⁸⁸.

O uso de NLP para extração de metadados oferece vantagens significativas. Ferramentas automatizadas e tecnologia de Reconhecimento Óptico de Caracteres (OCR) extraem automaticamente metadados de documentos jurídicos, reduzindo erros de entrada manual de dados e acelerando o processo⁸⁹. A incorporação de Processamento de Linguagem Natural para análise de conteúdo de documentos permite extração autônoma de palavras-chave relevantes, tópicos jurídicos e sentimentos, enriquecendo os metadados do documento para capacidades melhoradas de busca e recuperação⁹⁰. Técnicas de NLP conseguem aplicar taxonomias ou sistemas de classificação customizados para enriquecer metadados, organizando documentos sistematicamente baseado em caso, cliente ou matéria jurídica.

2.5 Jurimetria preditiva e análise de tendências

Além de triagem e classificação operacionais, NLP capacita análises preditivas sofisticadas que transformam decisões jurídicas. A jurimetria utiliza dados estatísticos para prever decisões judiciais e avaliar tendências em tribunais⁹¹. Ferramentas como Lex Machina e plataformas brasileiras oferecem análises detalhadas de decisões

⁸⁶ PAGELIGHT PRIME. *Metadata extraction in legal document management*. [s.d.]. Disponível em: <https://www.pagelightprime.com/blogs/metadata-extraction-legal-document-management/>. Acesso em: 22 mar. 2026.

⁸⁷ PAGELIGHT PRIME. *Metadata extraction in legal document management*. [s.d.]. Disponível em: <https://www.pagelightprime.com/blogs/metadata-extraction-legal-document-management/>. Acesso em: 22 mar. 2026.

⁸⁸ PAGELIGHT PRIME. *Metadata extraction in legal document management*. [s.d.]. Disponível em: <https://www.pagelightprime.com/blogs/metadata-extraction-legal-document-management/>. Acesso em: 22 mar. 2026.

⁸⁹ PAGELIGHT PRIME. *Metadata extraction in legal document management*. [s.d.]. Disponível em: <https://www.pagelightprime.com/blogs/metadata-extraction-legal-document-management/>. Acesso em: 22 mar. 2026.

⁹⁰ PAGELIGHT PRIME. *Metadata extraction in legal document management*. [s.d.]. Disponível em: <https://www.pagelightprime.com/blogs/metadata-extraction-legal-document-management/>. Acesso em: 22 mar. 2026.

⁹¹ CAUSOS E CAUSAS. *O uso de inteligência artificial no Direito*. [s.d.]. Disponível em: <https://causosecausas.com/o-uso-de-inteligencia-artificial-no-direito-analise-de-documentos-automacao-de-tarefas-e-jurimetria/>. Acesso em: 22 mar. 2026.

judiciais, permitindo que advogados elaborem estratégias mais assertivas⁹². A aplicação de *machine learning* com bancos de jurisprudência tem permitido identificação rápida de decisões similares⁹³. Essa funcionalidade não só economiza tempo, mas também melhora precisão e consistência em argumentações jurídicas⁹⁴.

Um exemplo notável é o uso de Lex Machina em processos de propriedade intelectual nos Estados Unidos, onde a ferramenta analisa padrões de decisão de juizes identificando probabilidades de vitória com base em casos anteriores⁹⁵. No Brasil, a Digesto (agora Jusbrasil Soluções) foi utilizada para mapear decisões sobre a Lei Geral de Proteção de Dados, ajudando empresas a se prepararem para ações judiciais⁹⁶. Essas análises demonstram como NLP e *machine learning* transformam jurimetria de ferramenta acadêmica em recurso prático e competitivo que qualifica decisões baseadas em dados sólidos⁹⁷.

2.6 Modelos Linguísticos Grandes (LLMs) na análise de documentos jurídicos

A evolução mais recente na aplicação de NLP ao direito envolve Modelos Linguísticos Grandes (*Large Language Models*, LLMs), como GPT-4, Llama 2 e Legal-Pegasus⁹⁸. Os LLMs representam sistemas baseados em Transformers treinados em *corpus* tex-

⁹² CAUSOS E CAUSAS. *O uso de inteligência artificial no Direito*. [s.d.]. Disponível em: <https://causosecausas.com/o-uso-de-inteligencia-artificial-no-direito-analise-de-documentos-automacao-de-tarefas-e-jurimetria/>. Acesso em: 22 mar. 2026.

⁹³ NETWORKS ONLINE. *Inteligência artificial aplicada ao Direito: tendências para 2025*. 2025. Disponível em: <https://www.networksonline.com.br/2025/01/21/inteligencia-artificial-aplicada-ao-direito-veja-tendencias-para-2025/>. Acesso em: 22 mar. 2026.

⁹⁴ NETWORKS ONLINE. *Inteligência artificial aplicada ao Direito: tendências para 2025*. 2025. Disponível em: <https://www.networksonline.com.br/2025/01/21/inteligencia-artificial-aplicada-ao-direito-veja-tendencias-para-2025/>. Acesso em: 22 mar. 2026.

⁹⁵ CAUSOS E CAUSAS. *O uso de inteligência artificial no Direito*. [s.d.]. Disponível em: <https://causosecausas.com/o-uso-de-inteligencia-artificial-no-direito-analise-de-documentos-automacao-de-tarefas-e-jurimetria/>. Acesso em: 22 mar. 2026.

⁹⁶ CAUSOS E CAUSAS. *O uso de inteligência artificial no Direito*. [s.d.]. Disponível em: <https://causosecausas.com/o-uso-de-inteligencia-artificial-no-direito-analise-de-documentos-automacao-de-tarefas-e-jurimetria/>. Acesso em: 22 mar. 2026.

⁹⁷ DEEPLEGAL. *Como o jurídico pode deixar de ser custo e virar centro de lucro*. [s.d.]. Disponível em: <https://www.deeplegal.com.br/blog/a-verdade-como-o-juridico-pode-deixar-de-ser-custo-e- virar-centro-de-lucro>. Acesso em: 22 mar. 2026.

⁹⁸ ARXIV. [Artigo científico]. 2025. Disponível em: <https://arxiv.org/html/2504.00725v1>. Acesso em: 22 mar. 2026.

tual massivo, que conseguem realizar múltiplas tarefas de NLP *downstream*, com notável proficiência após instrução apropriada. LLMs aplicados ao contexto legal têm potencial de otimizar significativamente processos como revisão contratual, predição de decisões judiciais, recuperação de jurisprudência e geração automática de resumos legais⁹⁹.

Entretanto, desafios persistem. Alucinações técnicas constituem problema significativo, em que LLMs geram informações incorretas apresentadas como verdade¹⁰⁰. Vieses nos dados de treinamento afetam desempenho e podem perpetuar discriminação¹⁰¹. Dificuldades na interpretação de contextos jurídicos específicos limitam confiabilidade em domínios altamente especializados¹⁰². Um dos focos principais de pesquisa atualmente é a utilização de técnicas de engenharia de prompts para melhorar precisão e relevância de respostas de LLMs em contextos jurídicos¹⁰³.

Comparações empíricas entre diferentes arquiteturas revelam insights importantes. Modelos jurídicos especializados como LegalBERT e Sabiá frequentemente superam modelos genéricos, como GPT-3.5 e GPT-4, especialmente em tarefas envolvendo terminologia técnica¹⁰⁴. Isso ocorre porque modelos treinados especificamente em domínio jurídico desenvolvem compreensão mais profunda de conceitos legais e terminologia técnica. Técnicas como *Few-shot Learning* e *Zero-shot Learning*, em conjunto com *Chain-of-Thought prompting*, demonstraram eficácia para melhorar a interpretação de textos legais¹⁰⁵. Esses avanços sugerem que a configuração apropriada de modelos linguísticos para tarefas jurídicas específicas oferece caminhos viáveis para otimização prática.

⁹⁹ ARXIV. [Artigo científico]. 2025. Disponível em: <https://arxiv.org/html/2504.00725v1>. Acesso em: 22 mar. 2026.

¹⁰⁰ ARXIV. [Artigo científico]. 2025. Disponível em: <https://arxiv.org/html/2504.00725v1>. Acesso em: 22 mar. 2026.

¹⁰¹ ARXIV. [Artigo científico]. 2025. Disponível em: <https://arxiv.org/html/2504.00725v1>. Acesso em: 22 mar. 2026.

¹⁰² ARXIV. [Artigo científico]. 2025. Disponível em: <https://arxiv.org/html/2504.00725v1>. Acesso em: 22 mar. 2026.

¹⁰³ ARXIV. [Artigo científico]. 2025. Disponível em: <https://arxiv.org/html/2504.00725v1>. Acesso em: 22 mar. 2026.

¹⁰⁴ ARXIV. [Artigo científico]. 2025. Disponível em: <https://arxiv.org/html/2504.00725v1>. Acesso em: 22 mar. 2026.

¹⁰⁵ ARXIV. [Artigo científico]. 2025. Disponível em: <https://arxiv.org/html/2504.00725v1>. Acesso em: 22 mar. 2026.

3 Desafios, responsabilidade ética e implementação em ambientes jurídicos de alto risco

3.1 Conformidade regulatória e privacidade de dados

A implementação de NLP em contextos jurídicos coloca a conformidade regulatória como preocupação central. A Lei Geral de Proteção de Dados (LGPD) brasileira, que entrou em vigor em 2020, implementou proteção de dados incluindo medidas como anonimização de dados pessoais¹⁰⁶. O Regulamento Geral de Proteção de Dados (GDPR) da União Europeia contém disposições específicas sobre tratamento automatizado de informações¹⁰⁷. A Lei de Inteligência Artificial recentemente adotada pela União Europeia estabelece diretrizes de privacidade, transparência e equidade, exigindo certificações que atestem conformidade de sistemas de IA¹⁰⁸.

Documentos jurídicos frequentemente contêm informações altamente sensíveis e confidenciais¹⁰⁹. Dados pessoais de litigantes, informações financeiras sensíveis, antecedentes criminais e comunicações privilegiadas entre cliente e advogado constituem apenas alguns exemplos. Sistemas de IA devem ser projetados com medidas de segurança robustas para proteger essas informações contra violação de dados¹¹⁰. A conformidade com leis de proteção de dados é essencial para manter a confiança de clientes¹¹¹.

Uma abordagem eficaz é o conceito de *Privacy by Design*, que incorpora privacidade como elemento fundamental desde a concepção de sistemas de IA¹¹². Isso requer análise constante de riscos em

¹⁰⁶ UPLEXIS. *Privacidade de dados*. [s.d.]. Disponível em: <https://uplexis.com.br/blog/artigos/privacidade-de-dados/>. Acesso em: 22 mar. 2026.

¹⁰⁷ UPLEXIS. *Privacidade de dados*. [s.d.]. Disponível em: <https://uplexis.com.br/blog/artigos/privacidade-de-dados/>. Acesso em: 22 mar. 2026.

¹⁰⁸ UPLEXIS. *Privacidade de dados*. [s.d.]. Disponível em: <https://uplexis.com.br/blog/artigos/privacidade-de-dados/>. Acesso em: 22 mar. 2026.

¹⁰⁹ POWERPATENT. *Role of AI in legal document categorization*. [s.d.]. Disponível em: <https://powerpatent.com/blog/role-of-ai-in-legal-document-categorization>. Acesso em: 22 mar. 2026.

¹¹⁰ POWERPATENT. *Role of AI in legal document categorization*. [s.d.]. Disponível em: <https://powerpatent.com/blog/role-of-ai-in-legal-document-categorization>. Acesso em: 22 mar. 2026.

¹¹¹ POWERPATENT. *Role of AI in legal document categorization*. [s.d.]. Disponível em: <https://powerpatent.com/blog/role-of-ai-in-legal-document-categorization>. Acesso em: 22 mar. 2026.

¹¹² POWERPATENT. *Role of AI in legal document categorization*. [s.d.]. Disponível em: <https://powerpatent.com/blog/role-of-ai-in-legal-document-categorization>. Acesso em: 22 mar. 2026.

todas as etapas de desenvolvimento e implementação, zelando pela conformidade com normas vigentes¹¹³. O Framework LGPD, desenvolvido como solução completa para auxiliar empresas a se adequarem à legislação, fornece metodologias comprovadas no mercado, como NIST, COBIT e ITIL, para assegurar que organizações atendam a requisitos legais de forma eficiente e contínua¹¹⁴.

3.2 Transparência, explicabilidade e *accountability*

Um desafio ético fundamental emerge em sistemas de IA jurídica: como podem operadores jurídicos e tribunais explicar decisões tomadas ou apoiadas por algoritmos quando o funcionamento interno desses sistemas permanece opaco mesmo para seus designers?. A explicabilidade da IA, definida como capacidade de decifrar por que determinada observação produziu determinado resultado, representa direito fundamental em esferas públicas¹¹⁵. Explicabilidade deve fornecer informações que facilitem entendimento de entradas, saídas e processos algorítmicos fundamentais, que levam aos resultados processados que influenciam decisões¹¹⁶.

A distinção entre transparência e explicabilidade é crítica. Transparência refere-se à capacidade de descrever e justificar processos que calculam parâmetros do modelo e produzem resultados¹¹⁷. Explicabilidade, em contraste, refere-se especificamente à obrigação de tornar os resultados de sistemas de IA compreensíveis para seres humanos¹¹⁸. Interpretabilidade, terceira dimensão, descreve a

¹¹³ UPLEXIS. *Privacidade de dados*. [s.d.]. Disponível em: <https://uplexis.com.br/blog/artigos/privacidade-de-dados/>. Acesso em: 22 mar. 2026.

¹¹⁴ DISRUPTX. *Framework LGPD*. [s.d.]. Disponível em: <https://disruptx.ai/framework-lgpd/>. Acesso em: 22 mar. 2026.

¹¹⁵ PÁDUA, Sérgio Rodrigo de; LORENZETTO, Bruno Meneses. *O direito fundamental à explicabilidade da inteligência artificial utilizada em decisões estatais*. Revista da AGU, Brasília, v. 23, n. 2, p. 349-371, jun. 2024. Disponível em: <https://revistaagu.agu.gov.br/index.php/AGU/article/view/3480/2934>. Acesso em: 22 mar. 2026.

¹¹⁶ PÁDUA, Sérgio Rodrigo de; LORENZETTO, Bruno Meneses. *O direito fundamental à explicabilidade da inteligência artificial utilizada em decisões estatais*. Revista da AGU, Brasília, v. 23, n. 2, p. 349-371, jun. 2024. Disponível em: <https://revistaagu.agu.gov.br/index.php/AGU/article/view/3480/2934>. Acesso em: 22 mar. 2026.

¹¹⁷ PÁDUA, Sérgio Rodrigo de; LORENZETTO, Bruno Meneses. *O direito fundamental à explicabilidade da inteligência artificial utilizada em decisões estatais*. Revista da AGU, Brasília, v. 23, n. 2, p. 349-371, jun. 2024. Disponível em: <https://revistaagu.agu.gov.br/index.php/AGU/article/view/3480/2934>. Acesso em: 22 mar. 2026.

¹¹⁸ PÁDUA, Sérgio Rodrigo de; LORENZETTO, Bruno Meneses. *O direito fundamental à explicabilidade da inteligência artificial utilizada em decisões estatais*. Revista da AGU, Brasília, v. 23, n. 2, p. 349-371, jun. 2024. Disponível em: <https://revistaagu.agu.gov.br/index.php/AGU/article/view/3480/2934>. Acesso em: 22 mar. 2026.

possibilidade de entender o modelo e apresentar como ele toma decisões de forma compreensível para o ser humano¹¹⁹. Esses conceitos, embora interligados, operam em níveis diferentes de abstração.

A Resolução nº 332/2020 do Conselho Nacional de Justiça ressalta a importância de transparência em IA, assegurando direito de cidadãos à explicação satisfatória e auditável por autoridades humanas em relação a decisões sugeridas por modelos de IA, garantindo assim *accountability*¹²⁰. As Recomendações da UNESCO (2021) enfatizam a importância de estabelecer requisitos claros de transparência e explicabilidade para sistemas de IA, especialmente na tomada de decisões por entidades públicas, como Judiciário e Administração Pública, abordando comportamento dos sistemas, incluindo algoritmos e dados envolvidos¹²¹. Este é vital porque decisões do Judiciário e da Administração impactam direitos fundamentais.

Técnicas de Inteligência Artificial Explicável (XAI) oferecem caminhos para abordar estas questões. O LIME (*Local Interpretable Model-agnostic Explanations*) permite que modelos sejam explicados de forma local e agnóstica, podendo gerar explicações para previsões específicas, sem necessidade de compreender completamente o modelo global¹²². SHAP (*SHapley Additive exPlanations*) oferece framework para explicar contribuição de cada *feature* para uma previsão específica, baseada em teoria de jogos cooperativos¹²³. Essas técnicas conseguem transformar *black-box models*

¹¹⁹ PÁDUA, Sérgio Rodrigo de; LORENZETTO, Bruno Meneses. *O direito fundamental à explicabilidade da inteligência artificial utilizada em decisões estatais*. Revista da AGU, Brasília, v. 23, n. 2, p. 349-371, jun. 2024. Disponível em: <https://revistaagu.agu.gov.br/index.php/AGU/article/view/3480/2934>. Acesso em: 22 mar. 2026.

¹²⁰ PÁDUA, Sérgio Rodrigo de; LORENZETTO, Bruno Meneses. *O direito fundamental à explicabilidade da inteligência artificial utilizada em decisões estatais*. Revista da AGU, Brasília, v. 23, n. 2, p. 349-371, jun. 2024. Disponível em: <https://revistaagu.agu.gov.br/index.php/AGU/article/view/3480/2934>. Acesso em: 22 mar. 2026.

¹²¹ PÁDUA, Sérgio Rodrigo de; LORENZETTO, Bruno Meneses. *O direito fundamental à explicabilidade da inteligência artificial utilizada em decisões estatais*. Revista da AGU, Brasília, v. 23, n. 2, p. 349-371, jun. 2024. Disponível em: <https://revistaagu.agu.gov.br/index.php/AGU/article/view/3480/2934>. Acesso em: 22 mar. 2026.

¹²² MANAGEMENT SOLUTIONS. *Explainable artificial intelligence*. [s.d.]. Disponível em: <https://www.managementsolutions.com/sites/default/files/minisite/static/22959b0f-b3da-47c8-9d5c-80ec3216552b/iax/pdf/explainable-artificial-intelligence-pt.pdf>. Acesso em: 22 mar. 2026.

¹²³ MANAGEMENT SOLUTIONS. *Explainable artificial intelligence*. [s.d.]. Disponível em: <https://www.managementsolutions.com/sites/default/files/minisite/static/22959b0f-b3da-47c8-9d5c-80ec3216552b/iax/pdf/explainable-artificial-intelligence-pt.pdf>. Acesso em: 22 mar. 2026.

em systems, em que decisões conseguem ser racionalizadas para humanos com rigor matemático.

3.3 Viés algorítmico e equidade em decisões jurídicas

Um desafio que permeia aplicações de IA em direito é a perpetuação de viés histórico através de algoritmos. Modelos de *machine learning* treináveis em dados contêm frequentemente preconceitos refletindo discriminação passada¹²⁴. Se sistemas são treinados utilizando dados históricos de decisões judiciais que refletem discriminação contra grupos específicos, baseado em raça, gênero, classe ou outras características protegidas, os modelos aprendem e replicam esses padrões discriminatórios em novas predições¹²⁵. Esse fenômeno, denominado viés algorítmico, representa uma ameaça significativa à equidade jurídica.

Exemplos históricos ilustram o perigo. Sistemas de reconhecimento facial com taxa de erro significativamente mais alta para indivíduos de pele escura conduziram a detenções injustas¹²⁶. Algoritmos de pontuação de risco utilizados em sistemas *criminal justice* produziram vieses contra réus de cor¹²⁷. Filtros de contratação desenvolvidos por grandes corporações discriminaram sistematicamente candidatas mulheres¹²⁸. Esses exemplos demonstram como o viés não é meramente questão técnica abstrata, mas resulta em dano concreto real a indivíduos e grupos.

Endereçar viés algorítmico requer múltiplas abordagens complementares. Primeiro, dados de treinamento devem ser cuidadosamente examinados por vieses preexistentes, com esforços feitos para balancear representação de grupos. Segundo, métricas de equidade devem ser definidas explicitamente e monitoradas durante desenvolvimento e *deployment*. Terceiro, testes sistemáticos devem avaliar desempenho de modelos em diferentes grupos

¹²⁴ UPLEXIS. *Privacidade de dados*. [s.d.]. Disponível em: <https://uplexis.com.br/blog/artigos/privacidade-de-dados/>. Acesso em: 22 mar. 2026.

¹²⁵ KPMG. *Privacy in the new world of AI*. 2023. Disponível em: <https://assets.kpmg.com/content/dam/kpmg/br/pdf/2023/12/Privacy-inthenew-world-of-AI-Final.pdf>. Acesso em: 22 mar. 2026.

¹²⁶ KPMG. *Privacy in the new world of AI*. 2023. Disponível em: <https://assets.kpmg.com/content/dam/kpmg/br/pdf/2023/12/Privacy-inthenew-world-of-AI-Final.pdf>. Acesso em: 22 mar. 2026.

¹²⁷ KPMG. *Privacy in the new world of AI*. 2023. Disponível em: <https://assets.kpmg.com/content/dam/kpmg/br/pdf/2023/12/Privacy-inthenew-world-of-AI-Final.pdf>. Acesso em: 22 mar. 2026.

¹²⁸ KPMG. *Privacy in the new world of AI*. 2023. Disponível em: <https://assets.kpmg.com/content/dam/kpmg/br/pdf/2023/12/Privacy-inthenew-world-of-AI-Final.pdf>. Acesso em: 22 mar. 2026.

demográficos, identificando disparidades¹²⁹. Quarto, ferramentas de XAI conseguem revelar quando modelos tomam decisões diferentes para indivíduos similares, sinalizando potencial viés¹³⁰.

3.4 Validação humana e supervisão em sistemas críticos

Um princípio fundamental na implementação responsável de IA jurídica é que a automação deve permanecer uma ferramenta que aumenta as capacidades humanas, não substitui julgamento humano em questões críticas¹³¹. No contencioso de massa, em que as decisões afetam direitos de consumidores e integridade de estratégias legais, nenhuma petição crítica deveria ser protocolada sem validação jurídica consciente¹³². A automação não deve gerar ruído argumentativo nem despersonalização da defesa¹³³. Tom, tese e linguagem devem permanecer alinhados com a reputação do cliente e o perfil do juiz¹³⁴.

Essa responsabilidade pessoal em tempos de automação é imperativa. O advogado ou responsável técnico que autoriza ações baseadas em recomendações algorítmicas mantém a responsabilidade final pela qualidade, precisão e conformidade legal¹³⁵. Além disso, o risco de decisões judiciais excessivas, assim como o risco de decisões equivocadas são uma realidade inerente ao Sistema de Justiça. Tal risco deve ser controlado caso a caso, sem se tornar obstáculo à

¹²⁹ FOCALX. *Validação de modelos de IA: garantir a exatidão e a fiabilidade*. [s.d.]. Disponível em: <https://focalx.ai/pt-pt/inteligencia-artificial/validacao-de-modelos-de-ia-garantir-a-exatidao-e-a-fiabilidade/>. Acesso em: 22 mar. 2026.

¹³⁰ MANAGEMENT SOLUTIONS. *Explainable artificial intelligence*. [s.d.]. Disponível em: <https://www.managementsolutions.com/sites/default/files/minisite/static/22959b0f-b3da-47c8-9d5c-80ec3216552b/iax/pdf/explainable-artificial-intelligence-pt.pdf>. Acesso em: 22 mar. 2026.

¹³¹ ADVTECHPRO. *Como este advogado usou IA para vencer um caso*. [s.d.]. Disponível em: <https://blog.advtechpro.ai/como-este-advogado-usou-a-ia-para-vencer-um-caso-de-us-15-milhao/>. Acesso em: 22 mar. 2026.

¹³² MIGALHAS. *Contencioso de massa com IA: eficiência com responsabilidade*. [s.d.]. Disponível em: <https://www.migalhas.com.br/depeso/435551/contencioso-de-massa-com-ia-eficiencia-com-responsabilidade>. Acesso em: 22 mar. 2026.

¹³³ MIGALHAS. *Contencioso de massa com IA: eficiência com responsabilidade*. [s.d.]. Disponível em: <https://www.migalhas.com.br/depeso/435551/contencioso-de-massa-com-ia-eficiencia-com-responsabilidade>. Acesso em: 22 mar. 2026.

¹³⁴ MIGALHAS. *Contencioso de massa com IA: eficiência com responsabilidade*. [s.d.]. Disponível em: <https://www.migalhas.com.br/depeso/435551/contencioso-de-massa-com-ia-eficiencia-com-responsabilidade>. Acesso em: 22 mar. 2026.

¹³⁵ MIGALHAS. *Contencioso de massa com IA: eficiência com responsabilidade*. [s.d.]. Disponível em: <https://www.migalhas.com.br/depeso/435551/contencioso-de-massa-com-ia-eficiencia-com-responsabilidade>. Acesso em: 22 mar. 2026.

adoção de boas práticas na condução judicial, conforme já decidido pelo Superior Tribunal de Justiça.¹³⁶

Conclusão

A automação de triagem e classificação de documentos através de Processamento de Linguagem Natural representa uma transformação profunda na prática jurídica contemporânea, particularmente em contextos de contencioso de massa e descoberta eletrônica. Os fundamentos tecnológicos desta transformação repousam em algoritmos sofisticados de *machine learning*, arquiteturas neurais avançadas, como Transformers e BERT, e abordagens híbridas que combinam capacidades de múltiplos paradigmas computacionais. A evidência empírica demonstra que a implementação responsável dessa tecnologia produz ganhos tangíveis: redução de tempo de revisão de até 80%, economias de custos na ordem de milhões de reais anuais, e melhoria na qualidade e consistência de argumentação jurídica.

Todavia, benefícios tecnológicos não devem obscurecer desafios fundamentais relacionados à responsabilidade ética, à conformidade regulatória e à equidade. Conformidade com LGPD, GDPR e normativas emergentes sobre IA exige que as organizações implementem *Privacy by Design* desde a concepção de sistemas. Transparência, explicabilidade e *accountability* devem permanecer como valores centrais, garantindo que decisões influenciadas por IA consigam ser racionalizadas e compreendidas por humanos com rigor apropriado. O viés algorítmico permanece uma ameaça real que requer vigilância constante, testes rigorosos e revisão humana contínua.

Os casos de sucesso em setores educacional, bancário e varejista documentam que o modelo híbrido no qual a automação é combinada com a validação humana rigorosa oferece um caminho viável. Ferramentas como Berna demonstram como a IA consegue identificar padrões em escala que escapariam ao processamento humano, aumentando a eficiência institucional, enquanto apoia objetivos de justiça. Modelos linguísticos especializados em domínios jurídicos superam contrapartes genéricas, sugerindo que o investimento em modelos customizados para contextos jurídicos específicos oferece retorno significativo.

Importante ressaltar que a incorporação de tecnologias de inteligência artificial no contencioso de massa insere-se no con-

¹³⁶ REsp 2021665

texto mais amplo de transformação do sistema de justiça, especialmente à luz da terceira onda renovatória de acesso à justiça, que demanda a adequação do método à natureza do conflito. Nesse cenário, a literatura especializada evidencia que o *Technology-Assisted Review* (TAR) constitui um instrumento apto a enfrentar a litigiosidade massificada com ganhos relevantes de eficiência e racionalidade decisória. O estudo seminal de Grossman e Cormack (2011) já demonstrava que a revisão assistida por tecnologia pode superar, em termos de precisão e custo-benefício, a revisão humana exaustiva, sobretudo em universos documentais de grande escala, e pesquisas mais recentes avançam ao propor métodos de validação estatística independentes, capazes de aferir, com maior rigor, a qualidade dos resultados obtidos pelo TAR, indicando que tais sistemas podem atingir um desempenho equivalente ou superior ao de revisores humanos, dentro de parâmetros de proporcionalidade processual.

Não obstante, a adoção dessas ferramentas no âmbito da advocacia pública exige cautela metodológica e normativa, uma vez que a utilização de sistemas algorítmicos na produção e análise de provas suscita questões relacionadas à validade, confiabilidade, transparência e explicabilidade das decisões automatizadas. Nesse sentido, a literatura contemporânea enfatiza que a legitimidade do uso da IA no Direito não decorre apenas de sua eficiência, mas também da sua capacidade de operar sob parâmetros de governança compatíveis com o devido processo legal e com os direitos fundamentais. Assim, no contexto da Justiça Multiportas, a inteligência artificial deve ser compreendida não como substituta do julgador ou dos métodos consensuais, mas como uma ferramenta estratégica de triagem, organização e qualificação da informação, capaz de viabilizar soluções mais adequadas, céleres e proporcionais aos conflitos massificados que envolvem a Administração Pública.

Perspectivas futuras incluem evolução de modelos linguísticos ainda mais sofisticados, integração de xAI (Inteligência Artificial Explicável) em sistemas críticos de produção, desenvolvimento de frameworks regulatórios mais claros para IA em contextos jurídicos, e criação de *datasets* jurídicos robustos e bem documentados que suportem pesquisa contínua. À medida que o volume de processos judiciais continua crescendo e a pressão por eficiência aumenta, organizações jurídicas que conseguem integrar NLP de forma ética, responsável e sofisticada estarão melhor posicionadas para servir os clientes, reduzir os custos e manter a qualidade técnica. A tecnologia, quando apropriada-

mente implementada com critério humano e supervisão rigorosa, representa uma aliada poderosa na construção de sistema jurídico mais eficiente, equitativo e acessível.

Referências

- AMERICAN PSYCHOLOGICAL ASSOCIATION. Exemplos de referências bibliográficas segundo o estilo APA. Lisboa: Faculdade de Ciências da Universidade de Lisboa, [s.d.]. Não paginado. Disponível em: https://ciencias.ulisboa.pt/sites/default/files/fcul/unidservico/bib/APA_FCULv2.pdf. Acesso em: 22 dez. 2025.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. Normas para apresentação de referências bibliográficas. Rio de Janeiro: ABNT, [s.d.]. Não paginado. Disponível em: <https://www.todamateria.com.br/referencias-abnt/>. Acesso em: 30 mar. 2026.
- BRAINS.DEV. Medidas de performance: modelos de classificação. [S.I.]: Brains, 2023. Não paginado. Disponível em: <https://brains.dev/2023/medidas-de-performance-modelos-de-classificacao/>. Acesso em: 22 dez. 2025.
- BRASIL. Superior Tribunal de Justiça. Recurso Especial n. 2.021.665. Brasília, DF, julgamento em 2025. Não paginado. Disponível em: <https://www.stj.jus.br/sites/porta/p/Paginas/Comunicacao/Noticias/2025/20032025-Corte-Especial-decide-em-repetitivo-que-juiz-pode-exigir-documentos-para-coibir-litigancia-abusiva.aspx>. Acesso em: 30 mar. 2026.
- CAUSOSECAUSAS. O uso de inteligência artificial no direito: análise de documentos, automação de tarefas e jurimetria. [S.I.]: Casos e Causas, [s.d.]. Não paginado. Disponível em: <https://causosecas.com/o-uso-de-inteligencia-artificial-no-direito-analise-de-documentos-automacao-de-tarefas-e-jurimetria/>. Acesso em: 06 dez. 2025.
- CDS LEGAL. The basics: what is e-discovery. [S.I.]: CDS Legal, [s.d.]. Não paginado. Disponível em: <https://cdslegal.com/knowledge/the-basics-what-is-e-discovery/>. Acesso em: 17 dez. 2025.
- CORMACK, Gordon V. et al. Unbiased validation of technology-assisted review for eDiscovery. In: Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2024. p. 2677-2681.
- DATACAMP. Classification machine learning: guia completo. [S.I.]: DataCamp, [s.d.]. Não paginado. Disponível em: <https://www.datacamp.com/pt/blog/classification-machine-learning>. Acesso em: 17 dez. 2025.
- DEEPLEGAL. A verdade: como o jurídico pode deixar de ser custo e virar centro de lucro. [S.I.]: Deep Legal, [s.d.]. Não paginado. Dispo-

nível em: <https://www.deeplegal.com.br/blog/a-verdade-como-o-juridico-pode-deixar-de-ser-custo-e-virar-centro-de-lucro>. Acesso em: 17 dez. 2025.

DISRUPTX. Framework LGPD: solução completa para conformidade de dados. [S.l.]: DisruptX, [s.d.]. Não paginado. Disponível em: <https://disruptx.ai/framework-lgpd/>. Acesso em: 06 dez. 2025.

ENCORD. Named entity recognition: guia completo. [S.l.]: Encord, [s.d.]. Não paginado. Disponível em: <https://encord.com/blog/named-entity-recognition/>. Acesso em: 30 mar. 2026.

ESCOLA NACIONAL DE ADMINISTRAÇÃO PÚBLICA. BERT: bidirectional encoder representations from transformers. Brasília: ENAP, 2023. p. 1-20. Disponível em: <https://bibliotecadigital.enap.gov.br/bitstream/1/7716/1/8091-Texto%20do%20Artigo-31686-1-10-20230807.pdf>. Acesso em: 06 dez. 2025.

FIRSTLEGAL. What is technology-assisted review? [S.l.]: First Legal, [s.d.]. Não paginado. Disponível em: <https://www.firstlegal.com/what-is-technology-assisted-review/>. Acesso em: 12 dez. 2025.

FOCALX. Validação de modelos de IA: garantir a exatidão e a fiabilidade. [S.l.]: FocalX, [s.d.]. Não paginado. Disponível em: <https://focalx.ai/pt-pt/inteligencia-artificial/validacao-de-modelos-de-ia-garantir-a-exatidao-e-a-fiabilidade/>. Acesso em: 12 dez. 2025.

GEEKSFORGEEEKS. When to use random forest over SVM and vice-versa. [S.l.]: GeeksforGeeks, [s.d.]. Não paginado. Disponível em: <https://www.geeksforgeeks.org/machine-learning/when-to-use-random-forest-over-svm-and-vice-versa/>. Acesso em: 30 mar. 2026.

GRIMM, Paul W.; GROSSMAN, Maura R.; CORMACK, Gordon V. Artificial intelligence as evidence. *Nw. J. Tech. & Intell. Prop.*, v. 19, p. 9, 2021.

GROSSMAN, Maura R.; CORMACK, Gordon V. Technology-assisted review in e-discovery can be more effective and more efficient than exhaustive manual review. *Richmond Journal of Law and Technology*, Richmond, v. 17, n. 3, 2011. Disponível em: <http://jolt.richmond.edu/v17i3/article11.pdf>. Acesso em: 22 dez. 2025.

HACKINGANALYTICS. Resampling: separando os dados entre treino, validação e teste. [S.l.]: Hacking Analytics, 2016. Não paginado. Disponível em: <https://hackinganalytics.wordpress.com/2016/09/04/resampling-separando-os-dados-entre-treino-validacao-e-teste/>. Acesso em: 30 mar. 2026.

IBM. Natural language processing (NLP): definição, aplicações e exemplos. [S.l.]: IBM, [s.d.]. Não paginado. Disponível em: <https://www.ibm.com/br-pt/think/topics/natural-language-processing>. Acesso em: 22 mar. 2026.

KPMG. Privacy in the new world of AI. [S.l.]: KPMG, 2023. p. 1-15.

Disponível em: <https://assets.kpmg.com/content/dam/kpmg/br/pdf/2023/12/Privacy-intheneu-world-of-AI-Final.pdf>. Acesso em: 22 mar. 2026.

LABELFI. What is accuracy, precision, recall and F1-score? [S.l.]: LabelF, [s.d.]. Não paginado. Disponível em: <https://www.labelfi.ai/blog/what-is-accuracy-precision-recall-and-f1-score>. Acesso em: 30 mar. 2026.

MANAGEMENT SOLUTIONS. Explainable artificial intelligence (XAI). [S.l.]: Management Solutions, [s.d.]. Não paginado. Disponível em: <https://www.management-solutions.com/sites/default/files/minisite/static/22959b0f-b3da-47c8-9d5c-80ec3216552b/iax/pdf/explainable-artificial-intelligence-pt.pdf>. Acesso em: 06 dez. 2025.

MIGALHAS. IA que identifica litigância abusiva é disponibilizada para tribunais. [S.l.]: Migalhas, [s.d.]. Não paginado. Disponível em: <https://www.migalhas.com.br/quentes/446889/ia-que-identifica-litigancia-abusiva-e-disponibilizada-para-tribunais>. Acesso em: 12 dez. 2025.

MIGALHAS. Contencioso de massa com IA: eficiência com responsabilidade. [S.l.]: Migalhas, [s.d.]. Não paginado. Disponível em: <https://www.migalhas.com.br/depeso/435551/contencioso-de-massa-com-ia-eficiencia-com-responsabilidade>. Acesso em: 12 dez. 2025.

NETWORKS ONLINE. Inteligência artificial aplicada ao direito: veja

tendências para 2025. [S.l.]: Networks Online, 2025. Não paginado. Disponível em: <https://www.networksonline.com.br/2025/01/21/inteligencia-artificial-aplicada-ao-direito-veja-tendencias-para-2025/>. Acesso em: 17 dez. 2025.

PÁDUA, Sérgio Rodrigo de; LORENZETTO, Bruno Meneses. O direito fundamental à explicabilidade da inteligência artificial utilizada em decisões estatais. Revista da AGU, Brasília, DF, v. 23, n. 2, p. 349-371, jun. 2024.

PAGELIGHTPRIME. Metadata extraction in legal document management. [S.l.]: PageLight Prime, [s.d.]. Não paginado. Disponível em: <https://www.pagelightprime.com/blogs/metadata-extraction-legal-document-management/>. Acesso em: 17 dez. 2025.

POWERPATENT. Role of AI in legal document categorization. [S.l.]: PowerPatent, [s.d.]. Não paginado. Disponível em: <https://powerpatent.com/blog/role-of-ai-in-legal-document-categorization>. Acesso em: 17 dez. 2025.

PROCURADORIA-GERAL DA REPÚBLICA. Explicabilidade e responsabilidade da inteligência artificial. Brasília: PGR, [s.d.]. Não paginado. Disponível em: <https://revista.agu.agu.gov.br/index.php/AGU/article/view/3480/2934>. Acesso em: 22 mar. 2026.

SERVICENOW. *eDiscovery*: descoberta eletrônica. [S.l.]: ServiceNow, [s.d.]. Não paginado. Disponível

em: <https://www.servicenow.com/br/products/legal-service-delivery/what-is-eDiscovery.html>. Acesso em: 17 dez. 2025.

SOCIEDADE AURUM. Contencioso de massa: tudo que você precisa saber. [S.l.]: Aurum, [s.d.]. Não paginado. Disponível em: <https://www.aurum.com.br/blog/contencioso-de-massa/>. Acesso em: 22 mar. 2026.

SUPERIOR TRIBUNAL DE JUSTIÇA. Corte Especial decide em repetitivo que juiz pode exigir documentos para coibir litigância abusiva. Brasília: STJ, 2025. Não paginado. Disponível em: <https://www.stj.jus.br/sites/portalp/Paginas/Comunicacao/Noticias/2025/20032025-Corte-Especial-decide-em-repetitivo-que-juiz-pode-exigir-documentos-para-coibir-litigancia-abusiva.aspx>. Acesso em: 20 dez. 2025.

THOMSON REUTERS. How predictive coding makes e-discovery more efficient. [S.l.]: Thomson Reuters Legal, [s.d.]. Não paginado. Disponível em: <https://legal.thomsonreuters.com/en/insights/articles/how-predictive-coding-makes-e-discovery-more-efficient>. Acesso em: 30 mar. 2026.

UPLEXIS. Privacidade de dados: proteção e conformidade com LGPD e GDPR. [S.l.]: Uplexis, [s.d.]. Não paginado. Disponível em: <https://uplexis.com.br/blog/artigos/privacidade-de-dados/>. Acesso em: 21 dez. 2025.