

DIFICULDADES TÉCNICAS E ESTRUTURAIS NO ACESSO A DADOS DE JULGADOS NOS TRIBUNAIS DE JUSTIÇA BRASILEIROS: UMA ANÁLISE COMPARATIVA COM ÓRGÃOS DE DADOS PÚBLICOS ABERTOS

Cleonis Viater Figueira; Lais Gomes Bergstein (Dra.)

Resumo

Este estudo analisou as barreiras técnicas e operacionais no acesso a dados de julgados dos Tribunais de Justiça (TJs) brasileiros, contrastando-as com as práticas de órgãos benchmark em dados abertos, IBGE e DATASUS. A metodologia consistiu em análise documental e comparativa de portais, sistemas e relatórios oficiais. Os resultados identificaram a predominância de formatos não estruturados (PDF), a ausência de APIs públicas, a fragmentação da governança de dados e a ambiguidade no licenciamento como entraves centrais. Concluiu-se que estes fatores comprometem a transparência efetiva, a pesquisa jurimétrica e a inovação em legaltechs, corroborando diagnósticos do CNJ. A discussão aponta que a superação deste cenário é viável e urgente, exigindo a implementação de APIs, padronização nacional de metadados e a correta aplicação de técnicas de anonimização para conciliar a LGPD com o interesse público.

Palavras-chave: Dados Abertos, Poder Judiciário, Transparência, Jurimetria, APIs

INTRODUÇÃO

A disponibilização de dados públicos é um pilar da transparência e da governança democrática. No campo jurídico, o acesso a decisões judiciais é essencial para o controle social, a pesquisa empírica e o desenvolvimento da jurimetria. No Brasil, contudo, os Tribunais de Justiça (TJs) operam como ilhas de dados digitais, onde a informação, embora nominalmente pública, encontra-se frequentemente aprisionada em formatos inadequados e sem padronização. Esta realidade contrasta com os princípios da Lei de Acesso à Informação (Lei nº 12.527/2011) e com as práticas de órgãos como IBGE e DATASUS, referências em dados abertos. Um relatório do Conselho Nacional de Justiça (CNJ) confirma essa assimetria, destacando os custos e a sobrecarga gerados pelos métodos atuais de acesso (LAWGORITHM, 2020). Diante disso, este trabalho teve como objetivo analisar criticamente os

entraves técnicos e estruturais no acesso a dados judiciais no Brasil, propondo um diálogo construtivo com setores públicos que já superaram tais desafios.

MÉTODOS

A pesquisa foi conduzida por meio de análise documental e comparativa. Foram analisados os portais de divulgação de dados e os sistemas de consulta de uma amostra não probabilística de Tribunais de Justiça brasileiros, com foco em seus aspectos técnicos e de usabilidade. Esta análise foi contrastada com a documentação técnica, as interfaces de programação (APIs) e as políticas de dados abertos de órgãos públicos consagrados na área, namely o Instituto Brasileiro de Geografia e Estatística (IBGE) e o Departamento de Informática do Sistema Único de Saúde (DATASUS). Adicionalmente, o estudo incorporou as evidências e recomendações contidas no "Relatório sobre Acesso a Dados de Processos Judiciais no Brasil", elaborado por Grupo de Trabalho do CNJ (LAWGORITHM, 2020), para embasar e validar os achados. Os critérios de análise, elencados no Quadro 1, foram selecionados com base nos princípios de dados abertos da Open Knowledge Foundation (OKFN, 2025), que preconizam dados disponíveis, utilizáveis, redistribuíveis e interoperáveis.

RESULTADOS E DISCUSSÕES

A análise revelou um abismo técnico-operacional entre a disponibilização de dados pelos TJs e pelos órgãos de referência. Enquanto IBGE e DATASUS operam na lógica da interoperabilidade, os TJs permanecem ancorados em um modelo de mera disponibilização formal.

O cerne do problema reside na predominância do formato PDF não editável para decisões judiciais, que impede a extração automatizada de informações. Este cenário é agravado pela ausência total de APIs públicas na esmagadora maioria dos tribunais, forçando pesquisadores e desenvolvedores a recorrer a técnicas de *web scraping*, prática identificada pelo relatório do CNJ como custosa e sobrecarregadora para os sistemas dos tribunais (LAWGORITHM, 2020, p. 6).

A descentralização do Judiciário manifesta-se digitalmente como uma fragmentação radical de sistemas (PJe, e-SAJ), terminologias e estruturas de dados. Cada TJ opera como um ecossistema isolado, inviabilizando a construção de um *corpus* nacional de jurisprudência sem um investimento proibitivo. Outro déficit crítico é a

inexistência de documentação técnica pública, contrastando com os manuais e dicionários de variáveis oferecidos pelo IBGE. A ambiguidade no licenciamento dos dados gera profunda insegurança jurídica, inibindo projetos de pesquisa e inovação.

Quadro 1. Comparação de quesitos técnicos estruturais de acesso a dados.

Quesito Técnico	Tribunais de Justiça	IBGE / DATASUS
Formato dos dados	não estruturado, .pdf	estruturado, .csv, .xls, .json
Disponibilidade de API	não	sim
Atualização periódica	irregular	programada
Documentação técnica	inexistente ou limitada	completa
Acessibilidade por máquina	baixo	alto
Licenciamento de uso	ambíguo	aberto
Padronização nacional	ausente	presente
Governança de dados	fragmentada	centralizada

Fonte: elaborado pelas autoras.

Para além das barreiras técnicas, persistem obstáculos institucionais, como uma cultura de sigilo, a governança fragmentada e uma interpretação restritiva da LGPD, frequentemente invocada para bloquear o acesso, ignorando-se que seu Art. 4º, III, exclui expressamente os tratamentos para fins acadêmicos e que técnicas de anonimização permitiriam a disponibilização segura (LAWGORITHM, 2020, p. 7). A análise do portal de dados abertos do Superior Tribunal de Justiça (STJ) revelou um esforço pioneiro, com a oferta de dados em formatos estruturados (.csv, .json) e documentação básica. No entanto, suas limitações reforçam o diagnóstico geral: a ausência de uma API pública força o download de *data dumps* massivos, a atualização não é em tempo real e os conjuntos carecem dos textos integrais das decisões, informação mais valiosa para a pesquisa.

CONCLUSÕES

Concluiu-se que a ausência de dados jurídicos estruturados e acessíveis compromete gravemente a transparência ativa do Judiciário, limita o controle social, engessa a pesquisa empírica em direito e inviabiliza o desenvolvimento em escala de ferramentas de jurimetria e inteligência artificial. A superação deste cenário de opacidade técnica é viável e urgente, exigindo: (i) a implementação de APIs

públicas, conforme recomendado pelo CNJ; (ii) a adoção de formatos abertos e a criação de uma política nacional de metadados; e (iii) uma interpretação da LGPD que equilibre a proteção de dados com o interesse público, utilizando a anonimização como mecanismo central. A transformação digital do Judiciário brasileiro só se completará quando suas decisões forem tão abertas, acessíveis e utilizáveis quanto os dados que fundamentam as políticas públicas em outras áreas do Estado.

REFERÊNCIAS

BRASIL. Instituto Brasileiro de Geografia e Estatística - IBGE. Portal de APIs do IBGE. Rio de Janeiro: IBGE, 2025. Disponível em: <https://servicodados.ibge.gov.br/api/docs/>. Acesso em: 12 set. 2025.

BRASIL. Ministério da Saúde - MS. Departamento de Informática do Sistema Único de Saúde - DATASUS. Informações de Saúde - TabNet. Brasília: Ministério da Saúde, 2025. Disponível em: <https://datasus.saude.gov.br/informacoes-de-saude-tabnet/>. Acesso em: 12 set. 2025.

LAWGORITHM. Relatório sobre Acesso a Dados de Processos Judiciais no Brasil. Grupo de Trabalho do Conselho Nacional de Justiça. Portaria CNJ/63/2019. 31 de julho de 2020.

OKFN. Open Knowledge Foundation. We are building a world open by design where all knowledge is accessible to everyone. 2025. Disponível em: <https://okfn.org/en/>. Acesso em: 17 set. 2025.

STJ. Supremo Tribunal de Justiça. Dados abertos do STJ. 2025. Disponível em: <https://dadosabertos.web.stj.jus.br/>. Acesso em: 17 set. 2025.