

LegalXplain: Referências para o Legal RAG e boas práticas de publicidade e proteção de Direitos Autorais

2025.2

Lucas Thevenard

Tópicos

- Quais referências e fontes de dados podemos utilizar para melhorar a análise do Legal RAG?
- Os REs podem ser **disponibilizados publicamente** na plataforma? Quais REs são públicos e quais não são? Quais seriam as boas práticas em termos de publicidade e proteção de dados para lidar com documentos processuais e jurídicos dessa natureza?
- E quanto a questões de **Direito Autoral**? Esses documentos são protegidos? Quais seriam as boas práticas para evitar violações do Direito Autoral?

Referências jurídicas para o Legal RAG

Fontes do Direito

- Tema clássico de Teoria do Direito. De onde surge o Direito?
- **Fontes** são as **origens ou formas de manifestação** do Direito.
- Indicam **de onde o Direito provém** e **como se expressa** no ordenamento jurídico.
- Dividem-se tradicionalmente em:
 - **Fontes materiais (ou substanciais)**: fatores sociais, econômicos, políticos e culturais que influenciam a criação do Direito.
 - **Fontes formais**: formas de produção das normas jurídicas reconhecidas oficialmente pelo aparato estatal e pelos aplicadores do Direito.

Principais fontes formais no Direito brasileiro:

- **(1) Lei**
 - Expressa pela atividade do Poder Legislativo.
- **(2) Jurisprudência**
 - Interpreta e aplica a lei aos casos concretos.
- **(3) Costume**
 - Prática social reiterada com convicção de obrigatoriedade.
- **(4) Doutrina**
 - Contribuição dos estudiosos e teóricos do Direito.
- **(5) PGDs, analogia e equidade**
 - Métodos frequentemente utilizados para preencher lacunas jurídicas.

Contratos como fonte inter-subjetiva

Em sentido estrito, as fontes do Direito são aquelas que criam normas jurídicas gerais e abstratas, aplicáveis a todos os casos semelhantes — como leis, costumes e precedentes obrigatórios.

Mas os contratos produzem **normas jurídicas intersubjetiva**. Embora não sejam fontes formais gerais, os contratos têm natureza normativa: ao serem celebrados, criam obrigações, deveres e direitos com força de lei entre as partes.

Publicidade dos documentos utilizados

Regra geral: publicidade dos recursos extraordinários

Os recursos extraordinários, como qualquer outro processo judicial, estão sujeitos ao **princípio da publicidade processual** (art. 93, IX, da Constituição Federal e art. 11 do CPC).

Isso significa que **as decisões e os autos processuais são públicos**, especialmente quando tramitam em tribunais superiores (STF).

Hipóteses de sigilo e restrição de acesso

- **(1) Segredo de justiça (CPC, art. 189):**
 - Casos que envolvem:
 - Direito de família (divórcio, filiação, alimentos);
 - Adoção ou tutela de menores;
 - Doenças graves ou dados médicos de partes;
 - Casos com sigilo bancário, fiscal ou telefônico;
 - Crimes sexuais.
 - Mesmo no STF, o acesso aos autos é restrito às partes e advogados habilitados.

Hipóteses de sigilo e restrição de acesso

- **(2) Proteção de dados pessoais (LGPD, Lei nº 13.709/2018):**
 - Dados pessoais identificáveis (nomes, CPFs, endereços, e-mails, etc.) devem ser anonimizados antes de uso em pesquisa ou divulgação.
 - Dados sensíveis (origem racial, convicção religiosa, opinião política, saúde, vida sexual, filiação sindical) exigem tratamento restrito e base legal específica (art. 11 da LGPD).

Hipóteses de sigilo e restrição de acesso

- **(3) Sigilo judicial decretado por decisão fundamentada:**
 - O tribunal pode decretar sigilo por razões de segurança nacional, investigações em curso, ou proteção de direitos individuais.
- **(4) Dados de menores ou incapazes:**
 - Mesmo que o processo não esteja formalmente em segredo de justiça, é recomendável anonimizar nomes e detalhes que permitam identificação.

Sugestões de boas práticas

a) Anonimização:

Remover nomes de partes, advogados, procuradores, testemunhas, e quaisquer dados pessoais.

Substituir por identificadores neutros (ex: Parte_A, Adv_01).

b) Filtragem prévia:

Excluir REs marcados como “segredo de justiça” no sistema de origem (muitos portais informam isso no metadado do processo).

Verificar se os dados foram coletados via APIs públicas (como a do STF), que já sinalizam o status de sigilo.

Sugestões de boas práticas

c) Uso para fins determinados e éticos:

Indicar claramente as finalidades para as quais os dados são usados, que devem ser finalidades éticas (pesquisa acadêmica, aprimorar a prática jurídica).

Evitar reproduzir textos ou petições originais em interfaces públicas.

Referenciar metadados com mais frequência, tomar cuidado com referências públicas ao conteúdo.

d) Política de transparência da base:

Explicitar critérios de coleta, anonimização e exclusão de casos sigilosos.

Disponibilizar documentação de conformidade com a LGPD e o CPC.

Questões de Direito Autoral

Natureza jurídica dos textos judiciais

A Lei de Direitos Autorais (Lei nº 9.610/1998, art. 8º, incisos I e IV) estabelece que atos oficiais — como decisões judiciais, despachos, votos e acórdãos — não são protegidos por direito autoral.

Art. 8º. Não são objeto de proteção como direitos autorais de que trata esta Lei:

I – as ideias, procedimentos normativos, sistemas, métodos, projetos ou conceitos matemáticos como tais;

(...)

IV – os textos de tratados ou convenções, leis, decretos, regulamentos, decisões judiciais e demais atos oficiais.

Tipo de documento	Titularidade autoral	Pode ser divulgado livremente?
Decisões, despachos, acórdãos	Domínio público (ato oficial)	 Sim
Peças das partes (petições, recursos, contrarrazões, memoriais etc.)	Autor: advogado, parte ou procurador	 Depende — há proteção autoral
Pareceres técnicos externos (peritos, especialistas)	Autor: perito ou instituição contratada	 Protegido, salvo cessão contratual
Documentos anexados (contratos, relatórios, laudos, obras intelectuais)	Autor: quem produziu o documento	 Protegido por direito autoral

Hipóteses que garantem usos permitidos em pesquisa e plataformas digitais

A Lei 9.610/98 contém exceções e limitações que permitem determinados usos de obras protegidas sem autorização do autor, especialmente para fins científicos, didáticos e informativos, desde que:

- não haja finalidade comercial direta; e
- seja preservado o direito moral de paternidade (citação do autor, se identificável).

Hipóteses que garantem usos permitidos em pesquisa e plataformas digitais

Algumas exceções relevantes na Lei de Direito Autoral:

- Art. 46, II: é permitido o uso de trechos de obras, para fins de estudo, crítica ou polêmica, “na medida justificada pelo objetivo a atingir”.
- Art. 46, VIII: é permitido o “retrato ou reprodução de obra que esteja permanentemente em local público”, o que se estende a atos públicos.
- Art. 46, XII: é permitido o uso de obras quando “indispensável à produção de prova judicial ou administrativa”.

Conclusão

Portanto:

- Análises empíricas e estatísticas (text mining, NLP) que não envolvam reprodução literal e integral do texto são compatíveis com o direito autoral.
- A divulgação de trechos anonimizados ou trechos mínimos ilustrativos é aceitável.
- A publicação integral de petições (como REs) em uma plataforma pública sem autorização dos autores (advogados) pode configurar violação de direito autoral, especialmente se houver monetização, exposição comercial ou prejuízo à reputação.

Boas práticas de conformidade ao Direito Autoral

a) Limitar-se a uso analítico, não reproduutivo

Armazenar e processar internamente os textos (para NLP, embeddings, classificação, etc.); Expor publicamente apenas resultados agregados, métricas ou representações vetoriais, não o texto integral.

b) Publicar apenas decisões judiciais integrais

O texto das decisões (acórdãos e votos) é de domínio público e pode ser reproduzido.

Boas práticas de conformidade ao Direito Autoral

c) Anonimizar e referenciar sem divulgar o texto

Referenciar o processo (número, classe, tribunal, data, relator);

Se citar trechos, fazer dentro do limite de citação (pequenos excertos) e com indicação da fonte.

d) Obter consentimento quando necessário

Em caso de publicação integral de petições ou memoriais (por exemplo, para fins didáticos), solicitar autorização do autor (ou da instituição).