

Análise Computacional de Decisões dos Tribunais Superiores Brasileiros

Mayara Cassia de Melo
(mayara.cassia.melo@usp.br)

Orientador: Prof. Dr. Marcelo Finger
(mfinger@ime.usp.br)



IME-USP

O Supremo Tribunal Federal (STF) e o Superior Tribunal de Justiça (STJ) proferem decisões definitivas que servem como referência para as demais instituições judiciais. Grande parte dessas decisões são públicas e estão disponíveis na Web. Este trabalho gera estatísticas e propõe métricas para investigar os mecanismos que regem o terceiro poder.

Acórdãos

Os acórdãos são as decisões deliberadas por um órgão colegiado. No STF e no STJ, são os processos julgados em conjunto pelos ministros. Fichas resumos dos acórdãos estão disponíveis para consulta nos sites dos tribunais. Cada ficha contém informações básicas sobre o processo e também indicam citações a outros acórdãos. Esta rede de citações pode ser modelada como um grafo dirigido.

Extração e persistência dos dados

A ferramenta Scrapy foi utilizada para o Web Scraping das fichas dos acórdãos entre 2001 e 2012. As fichas são organizadas em seções que possibilitaram a identificação e extração de informações, como o relator, data de julgamento, unidade federativa, ementa e citações dos acórdãos. As informações coletadas são persistidas em um banco não relacional orientado a documentos usando MongoDB como sistema gerenciador.

Algoritmo de Floyd-Warshall

Encontra os caminhos mais curtos entre os vértices em um grafo dirigido.

A cada vértice k , o algoritmo tenta encontrar um caminho mais curto, passando por k , para todos os pares de vértices. Seja $d(i,j)$ o tamanho do caminho mais curto entre os vértices i e j , a cada iteração calcula-se:

$$d(i,j) = \min \{ d(i,j), d(i,k) + d(k,j) \}$$

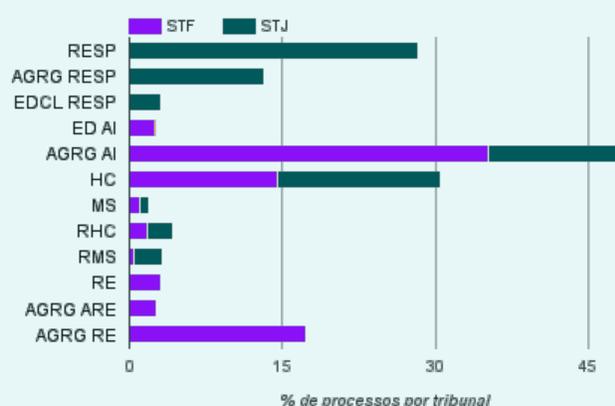
Uma versão modificada do algoritmo foi implementada para detectar ciclos no grafo de citações dos acórdãos.

Page Rank

Algoritmo que realiza um ranqueamento, por relevância, dos vértices em um grafo dirigido baseado na quantidade e na qualidade das arestas que chegam no vértice. A qualidade das arestas é proporcional a relevância e ao número de arestas que saem do seu vértice de origem. Para o grafo de citações dos acórdãos, aqueles mais citados são mais bem classificados, porém as citações dos acórdãos que fazem várias citações e/ou são poucos citados tem pouca influência. Também foi implementado uma variação do algoritmo de Page Rank na qual a qualidade das citações não depende do número de citações do acórdão

Resultados

Distribuição dos tipos de processos



Ranking dos acórdãos por Page Rank:

Page Rank original	Page Rank modificado
Grafo STF e STJ	Grafo STF e STJ
Grafo STF	Grafo STF
ERESP 435835	ADC 4
RESP 106888	HC 82959
RESP 198807	RE 415454
RESP 271214	RE 416827
RESP 248893	ADI 1662
RESP 407097	RE 290079
RESP 298191	RE 258980
AGRG AG	ADI 3105
HC 82959	HC 81611
	HC 102088
	HC 96202
	HC 102080
	HC 111727
	HC 110926
	HC 111123
	HC 110850
	RHC 89624
	HC 112319

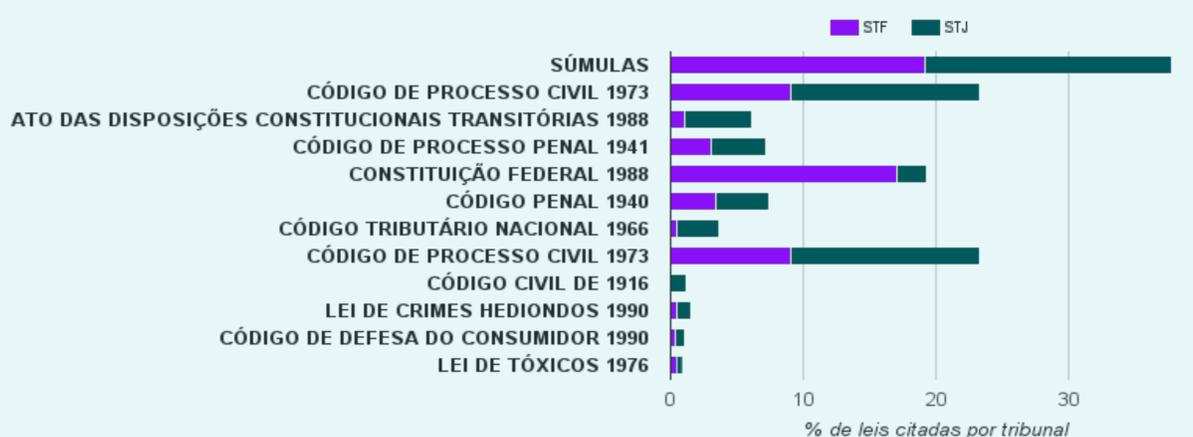
Palavras mais frequentes na indexação

STF	STJ
AUSÊNCIA	HIPOTESE
DESCABIMENTO	CARACTERIZAÇÃO
RECURSO EXTRAORDINÁRIO	DECORRENCIA
CÍVEL	OBJETIVO
VOTO VENCIDO	PREVISÃO
IMPOSSIBILIDADE	IMPOSSIBILIDADE
DECORRÊNCIA	NECESSIDADE

Citações a acórdãos

total de acórdãos: 51 mil STF + 285 mil STJ
total de citações: 1.007.163
número de acórdãos que citam: 208 mil
média de citações por acórdão: 3
média de citações por acórdãos que citam 4,8
acórdão que mais cita: RESP 1061530 - 113 citações
acórdão mais citado: ERESP 435835 - 1489 citações

Leis mais citadas



Citações a legislação

nº de citações a legislação
STF: 122.834 STJ: 512.899

média de citações de leis por acórdãos STF:
2,35 STJ: 1,8

nº de leis diferentes citadas
STF: 8.187 STJ: 10.718