

## Manifestos

Associação Brasileira de Jurimetria

09 de novembro de 2017

# Manifestos

Esse documento contém os manifestos da ABJ.

# Capítulo 1

## Manifesto por dados abertos

### 1.1 Introdução (motivação)

Não se muda aquilo que se ignora

É estratégico (tribunais e hospitais)

É grande

Os dados são públicos

Mesmo assim, temos problemas de privacidade

Queremos usar para fazer pesquisa, para adm de tribunais e políticas públicas.

### 1.2 Desafios

Existem inúmeras ferramentas públicas e privadas para busca e recuperação de processos. Os sistemas são eficazes, mas são todos voltados para a busca de informações individuais. Se uma pessoa tiver o número identificador, ela achará informações do processo. Se precisar uma lista de processos, poderá utilizar ferramentas de busca.

Cientistas de dados, no entanto, precisam ter a possibilidade de exportar os dados completos ou algum recorte da população para planilhas. Existem muitos exemplos de páginas úteis para cientistas de dados, como IpeaData, Datasus, IBGE, entre outros.

Também é possível utilizar *Application Programming Interfaces* (APIs) para obter dados de *tweets*, publicações no Facebook, entre outros. O importante é notar que os sistemas voltados para análise de dados são em sua maioria voltados para extração de informações de muitos indivíduos. Os dados são organizados para análise e não para consulta individual. Muitas vezes é necessário limpar a base, mas isso faz parte do fluxo da ciência de dados (Wickham and Grolemond, 2016).

Na área do direito, o pesquisador fica numa situação complicada, pois precisa de dados da população ou de uma amostra, com linhas e colunas, numa planilha padronizada. No entanto, tudo o que consegue encontrar são

documentos individuais, listagens de processos e arquivos em formato fechado, como o *Portable Document Format* (PDF).

Muitas vezes os dados estão disponíveis em páginas web mas é muito demorado buscar todos os casos que precisamos manualmente ou através de ofícios. Por isso, é usual construir *web scrapers*, que são robôs que baixam as páginas automaticamente e depois as transformam em dados estruturados.

Atualmente, a utilização *web scrapers* é indispensável em estudo jurimétricos. As pesquisas realizadas pela ABJ foram fortemente influenciadas por essas ferramentas.

Contudo, são raros os profissionais que dominam esse conhecimento. A ABJ disponibiliza abertamente todo seu aparato técnico<sup>1</sup>, mas as ferramentas não são capazes de resolver todos os problemas. Além disso, os sistemas dos Tribunais colocam impedimentos técnicos de acesso, dificultando a execução de pesquisas que poderiam ser benéficas para os próprios Tribunais.

Há muitos exemplos em que simplesmente não é possível obter as informações que necessitamos. Em muitos casos a única forma de acessar os dados é a partir da Lei de Acesso à Informação. Apesar da LAI ser um grande avanço, utilizá-la para todas as demandas é ineficiente pois congestionam os setores administrativos e técnicos dos Tribunais.

A solução mais eficaz para o problema de acesso aos dados envolve modificar os sites dos Tribunais, permitindo extrações de dados e disponibilizar APIs que permitam os pesquisadores de buscar as informações públicas de maneira segura e organizada.

Essas ferramentas são simples de construir para entidades como os Tribunais, que geralmente têm equipes de Tecnologia da Informação de altíssima qualidade. A solução não causaria impactos negativos nos sistemas; pelo contrário: ao permitir que os dados sejam baixados de forma programática, é possível controlar o volume de dados transferido por unidade de tempo, evitando que os servidores fiquem sobrecarregados.

## 1.3 Desafio

Em teoria, a solução da API é perfeita. O problema é que falta uma padronização concreta para que

Desafio: como conectar sistemas, sem necessariamente modificar a estrutura dos tribunais?

O erro está em querer alinhamento de sistemas, e não alinhamento focado em pesquisa.

As pesquisas readequam o sistema

## 1.4 Solução

Partir do processo de realização de pesquisas. Ciclo da ciência de dados?

A pesquisa no direito com fins de elaboração de políticas públicas ou administração de tribunais seguem padrões específicos. Na história da ABJ, executamos pesquisas em praticamente todas as esferas do direito. Já pesquisamos sobre direito tributário, processos de adoção, direito do consumidor, processos de homicídio, falências e recuperações, entre muitos outros. Com base nessa experiência e com base no que observamos na Society of

---

<sup>1</sup>Disponíveis nos links: <https://github.com/abjur> (códigos genéricos da ABJ), <https://github.com/courtsbr> (web scrapers), e <https://github.com/decryptr> (ferramentas para quebrar CAPTCHAs).

Empirical Legal Studies, montamos uma lista do que acreditamos ser os objetos de pesquisa mais comuns no direito.

### **1.4.1 Contagens e valores por tema/assunto**

Trata-se do tema mais comum em pesquisa do direito. Queremos contar processos por tipo, por valor, por tribunal etc. Por exemplo, Justiça em Números.

### **1.4.2 Tempos e complexidade processual**

### **1.4.3 Resultados e reformas dos processos judiciais**

Para entender o que é o direito, precisamos entender o que os juízes decidem. Queremos saber se as decisões variam muito dentro de um mesmo tema, se existe discriminação e se é possível

### **1.4.4 análise de impacto regulatório**

O que essas pesquisas tem em comum?

Estrutura de dados: infos básicas, partes, movimentações, decisões, documentos.

## **1.5 Propostas**

Não queremos que os tribunais mandem, queremos poder extrair dos tribunais

Proposta de como fazer: API pública, camada de APIs

Passo 1: Disponibilização de API pública de infosbásicas, partes e movimentacoes

Dados do Selo Justiça em números!

Passo 2: Diário oficial digital no CKAN - arrumar distribuicao

Passo 3: input/output: O que foi pedido, o que foi concedido?

Passo 4: Ligando processos, partes, advogados, teses e magistrados

Passo 5: Integração com polícia, defensoria e ministério público, entre outras (para segpub)

## **1.6 Aplicações**

Parágrafo introdutório

- Para o judiciário (assuntos processuais e padronização)
- Para o cidadão (acesso à justiça)
- Para empresas (acompanhamento processual)

- Para legaltechs (acelerar fluxos de mediação)
- Para magistrados (tomada de decisão, conexão com outros sistemas)
- Para advogados (acompanhamento processual e jurisprudência)

Economia de bilhões. Como? Criar uma conta

## 1.7 Próximos passos

A solução precisa ser comandada pelo DPJ-CNJ

Juntar com Selo Justiça em Números

Os tribunais já recebem diversas propostas. É preciso alinhar e limpar

1. Coordenação: DPJ
2. equipe técnica
  - pesquisadores da ABJ e do DPJ
  - arquitetos de TI dos sistemas atuais (saj, pje)
3. equipe articuladora
  - representantes da OAB, GEAL-CRASP, conselheiros do CNJ
  - Representantes dos tribunais superiores
  - universidades: Insper, GV, PUC, USP

Parágrafo de finalização

# Referências Bibliográficas

Wickham, H. and Golemund, G. (2016). R for data science.