

Laboratório de jurimetria: documentação

José de Jesus Filho

2019-04-02

Contents

1	Visão geral	5
2	Introdução	7
3	Requerimentos	9
3.1	Hardware	9
3.2	Network	9
3.3	Software	9
4	Instalação	11
4.1	Arquitetura	11
4.2	Implementação	11
4.3	11
5	Configuração do sistema	13
5.1	13
6	Configuracao do Myrthes	15
7	Busca com Myrthes	17
7.1	Busca básica	17
7.2	Autocompletar	17
7.3	Filtragem por datas	17
7.4	Facetas	17
7.5	Destaque (highligting)	17
7.6	Busca avançada	17
7.7	Alertas	17
7.8	Favoritos	17
7.9	Configuração do Solr no Myrthes	17
8	Dashboard	19
8.1	Criando dashboards	19
8.2	Tabelas	19
8.3	Gráficos de barra	19
8.4	Gráficos de linha	19
8.5	Mapas	19
8.6	Textos	19
8.7	Incluindo facetas e filtros no dashboard	19
8.8	Partilhando o dashboard	19
9	FAQ	21

Chapter 1

Visão geral

Este documento contém tudo que se refere ao Myrthes. Nele há seções úteis tanto para desenvolvedores quanto para usuários (estudantes e professores de direito).

Chapter 2

Introdução

Diante do volume de processos judiciais em versão eletrônica atualmente disponíveis nas páginas dos tribunais, o acesso a esses dados tornou-se possível. Atualmente se vislumbra a possibilidade de baixar processos inteiros e analisá-los. Essa possibilidade, porém, implica em desafios.

A comunidade do direito não é muito afeita à computação, bem como, aos métodos estatísticos e de aprendizado de máquina que atualmente avançam rapidamente e são liderados principalmente por pessoas provenientes da ciência da computação. Acontece que, embora o pessoal da ciência da computação e da estatística tenha liderado a implementação dos métodos de automação de coleta, extração, organização e análise de processos judiciais, ela carece de conhecimento substantivo do direito e isso tem causado erros e equívocos nas coletas e análises de dados.

A necessidade de se implementar um laboratório de jurimetria na faculdade de direito que contemple a interdisciplinariedade se impõe, visto que um laboratório como este exige conhecimentos em várias áreas:

- 1 - Implementação e manutenção de sistemas distribuídos para lidar com o alto volume de dados;
- 2 - Conhecimentos específicos sobre processamento de linguagem natural;
- 3 - Conhecimentos específicos sobre estatística e machine learning;
- 4 - Conhecimento aprofundado de cada uma das áreas do direito aplicado diariamente no tribunais de justiça do Brasil.

Chapter 3

Requerimentos

3.1 Hardware

Processador : Mínimo 1GHZ, recomenda-se 2+GHZ Quad Core RAM (exclusive SWAP) : Com ELK: Min 16 GB, Preferível 32 GB

3.2 Network

As portas 443, 8080, 5432, 9200, 5601, 8983 devem estar abertas. Com exceção da 443, as demais podem estar fechadas, desde que por trás de proxy reverso.

3.3 Software

Ubuntu 18.04+

JAVA 8 JRE ou JDK com JAVA_HOME configurada no /etc/environment

Python minimal instalado

Nodejs (será usado com o prometheus e grafana)

Locale configurado para pt_BR.UTF-8

Exigências:

Apache ManifoldCF Apache Tika Apache Tomcat Apache Zookeeper Apache Solr ELK (Elasticsearch, Logstash e Kibana) Prometheus + Grafana Apache Cassandra Postgresql 11

Chapter 4

Instalação

4.1 Arquitetura

```
knitr::include_url("https://www.draw.io/?lightbox=1&highlight=0000ff&edit=_blank&layers=1&nav=1&title=a
```

```
## PhantomJS not found. You can install it with webshot::install_phantomjs(). If it is installed, please
```

4.2 Implementação

A tabela abaixo, que deve ser atualizada conforme adicionamos novos elementos, lista as implementações, bem como o seu estágio.

4.3

Table 4.1: Implementações

Recurso	Implementado
scraper tribunais	parcial
crawler djes	sim
indexação para Solr	sim
tika-ocr	sim
stopwords	parcial
anotação	não
fuzzy search	sim
regex textos	sim
parseamento htmls	parcial

Chapter 5

Configuração do sistema

5.1

Chapter 6

Configuracao do Myrthes

Na configuração do sistema é importante levar em consideração o consumo da RAM e da CPU. A indexação consome bastante CPU. Por sua vez, a busca consome muita RAM. Abaixo seguem as configurações mínimas de RAM, de CPU e swap.

Chapter 7

Busca com Myrthes

- 7.1 Busca básica
- 7.2 Autocompletar
- 7.3 Filtragem por datas
- 7.4 Facetas
- 7.5 Destaque (highlighting)
- 7.6 Busca avançada
- 7.7 Alertas
- 7.8 Favoritos
- 7.9 Configuração do Solr no Myrthes

Chapter 8

Dashboard

8.1 Criando dashboards

8.2 Tabelas

8.3 Gráficos de barra

8.4 Gráficos de linha

8.5 Mapas

8.6 Textos

8.7 Incluindo facetas e filtros no dashboard

8.8 Partilhando o dashboard

Chapter 9

FAQ