

HappyHillary – PP-2Q/2016

Lenin Cristi
github.com/lnncrs
RA-21028214

lenin.cristi@aluno.ufabc.edu.br

Gustavo Murayama
github.com/gmurayama
RA-21028214

g.murayama@aluno.ufabc.edu.br

RESUMO

Projeto de análise de dados dos e-mails vazados da candidata à presidência dos EUA, Hillary Clinton, em 2016. O objetivo da análise é cruzar notícias das grandes mídias, redes sociais e

CCS Concepts

• Theory of computation → Design and analysis of algorithms → Data structures design and analysis → Pattern matching

Keywords

análise de dados; big data; hillary clinton; scheme; leak.

1. INTRODUÇÃO

Programa para analisar dados e encontrar uma relação entre eles, mais especificamente sobre os e-mails vazados da candidata à presidência, Hillary Clinton.

2. DEFINIÇÃO DO PROBLEMA

Programa feito na linguagem Scheme para encontrar um padrão entre os *e-mails* vazados e a repercussão nos meios de comunicação *online* existentes ao cruzarmos as datas (tanto da liberação dos e-mails quanto das reportagens) e as palavras mais usadas nas notícias. Visa retirar informações úteis em meio a dados aparentemente difíceis de serem analisados.

3. ARQUITETURA

```
/data
  /leak
  /sent
  /sources
/docs
/lib
  utils.scm
/modules
  /io
    reader.scm
    writers.scm
  /workers
    counter.scm
  testes.scm
/report
  /_models
  /_original
  /bower_components
    /bootstrap
    /charjs
    /font-awesome
    /jqcloud2
    /jquery
  /content
  /data
  /node_modules
```

/scripts

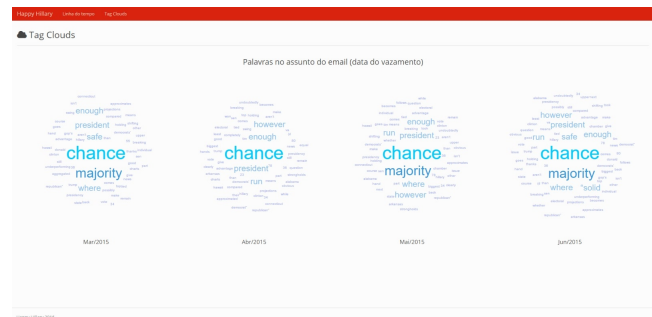


Figura 1: Tag clouds index.html

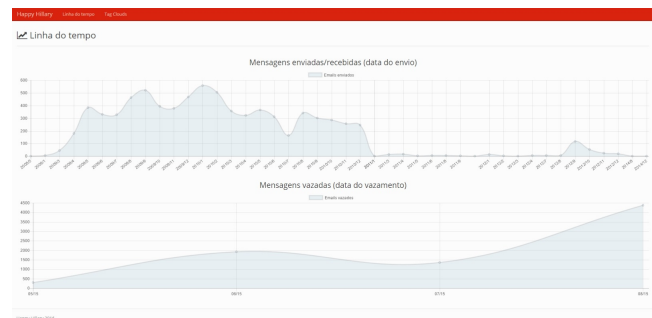


Figura 2: Gráficos

O arquivo **index.html**, localizado na pasta **report** é a tela de visualização de gráficos e as tags clouds geradas pelos arquivos **.scm**. Através do **testes.scm**, o arquivo consome os dados presentes na pasta **data** e retorna vários valores em formato **.js**.

4. SCREENSHOTS

5. DISCUSSÃO

O projeto realizado é complexo, pois visa analisar dados que podem ter ou não relação entre si, de forma a extrair uma informação. A visualização de tais dados, portanto, é essencial para uma análise precisa e completa da situação. Trabalhar com a linguagem *Scheme* foi um grande desafio, principalmente por causa de sua sintaxe e grande gama de possibilidades para realizar uma mesma tarefa, sendo difícil optar pela melhor maneira, uma vez que esta não é tão clara.

6. LINK

<https://github.com/Inncrs/HappyHillary>