

## Introdução

Contexto: O Brasil está vivenciando uma operação da polícia federal, denominada Lava Jato, que investiga um esquema bilionário de desvio e lavagem de dinheiro envolvendo a Petrobras, empreiteiras e políticos brasileiros.

Objetivo: Coletar twitters de usuários que abordem assunto relacionado com a operação Lava Jato realizada pela polícia federal.

## Análises das Palavras Coletadas

A primeira coleta foi utilizada os termos: “Lava Jato, lava-jato e Operação Lava Jato. Com essa coleta podemos verificar os principais termos ditos pelos usuários do twitter. A figura 1 ilustra a nuvem de TAGs gerada com a frequência simples das palavras mais ditas no twitter.

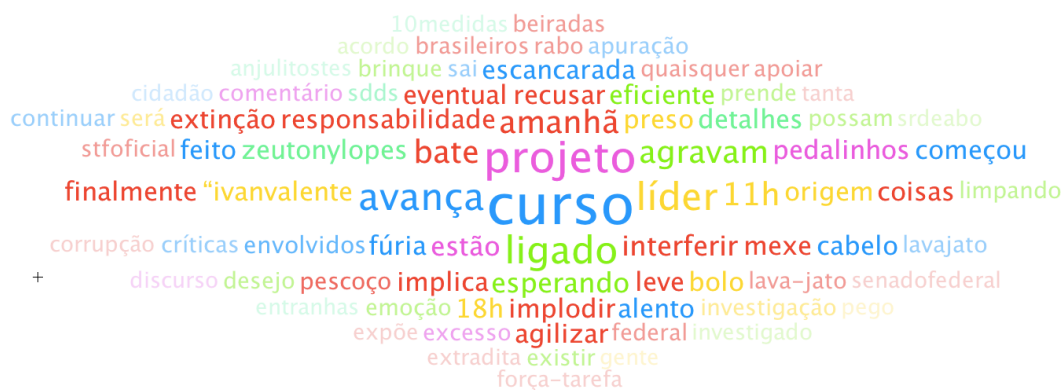


Figura 1. Nuvem de Tag com frequência simples das palavras

Para validar e verificar os resultados dessa primeira análise utilizamos uma abordagem diferente para contabilizar as frequências das palavras. Para realizar essa demanda utilizamos um extrator de palavras chaves relevantes. Esse extrator diferente da técnica de contagem das frequência das palavras, a abordagem é baseada em gráfico descrito em "KeyGraph: indexação automática por Co-ocorrência baseado na construção de Connstruction Metaphor", de Yukio

### Figura 3. Resultado Pesquisa

Conforme ilustrado podemos observar que o primeiro se trata de um usuário do Twitter que é bem reconhecido pelas suas postagens conforme pode ser visto na figura 4. Para melhor visualização, ilustramos o nível alto e baixo de pessoas que referenciam o usuário. Já o segundo se trata das 10 medidas que o ministério público federal criou para aprimorar a prevenção e o combate à corrupção e à impunidade.

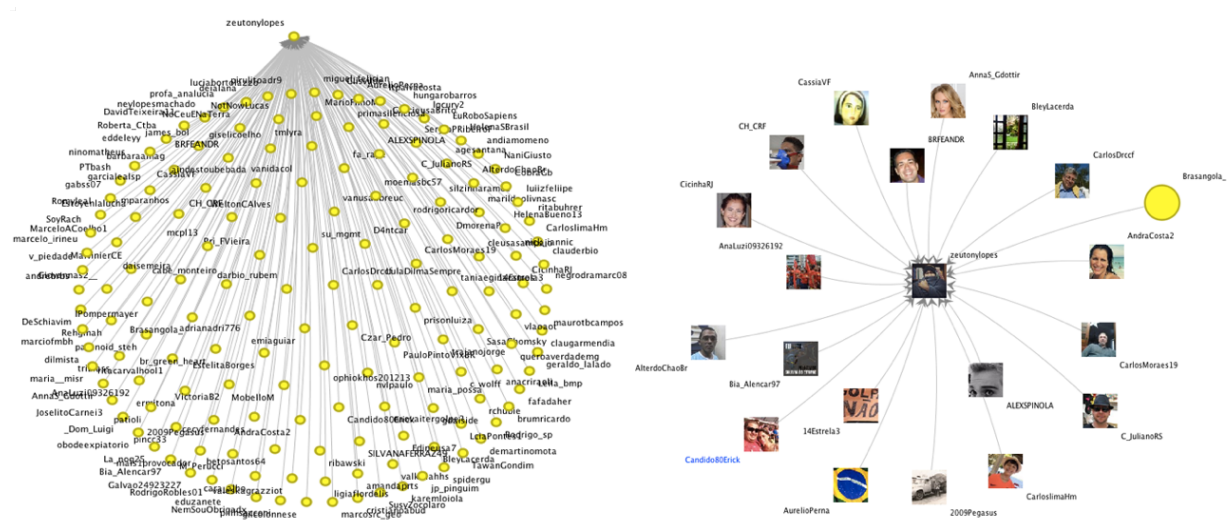


Figura 4. Pessoas que fazem referência ao Zeutonylopes

## Utilização do KNIME - Fluxos Desenvolvidos

Knime é uma ferramenta opensource utilizada para análises de dados. O fluxo apresentado na figura 5 e 6 mostram os passos criados na ferramenta para criação das nuvens de TAGs e relacionamento entre os usuários respectivamente,

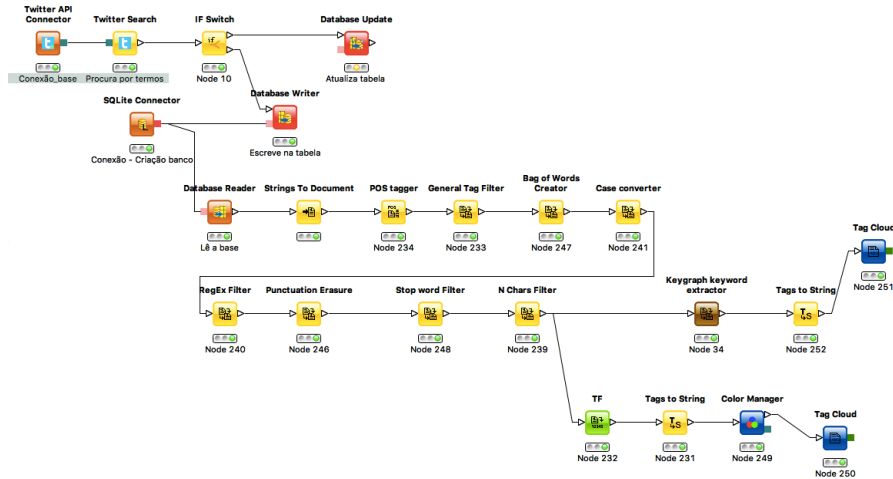


Figura 5. Fluxo desenvolvido no Knime – Criação Nuvem de TAGs

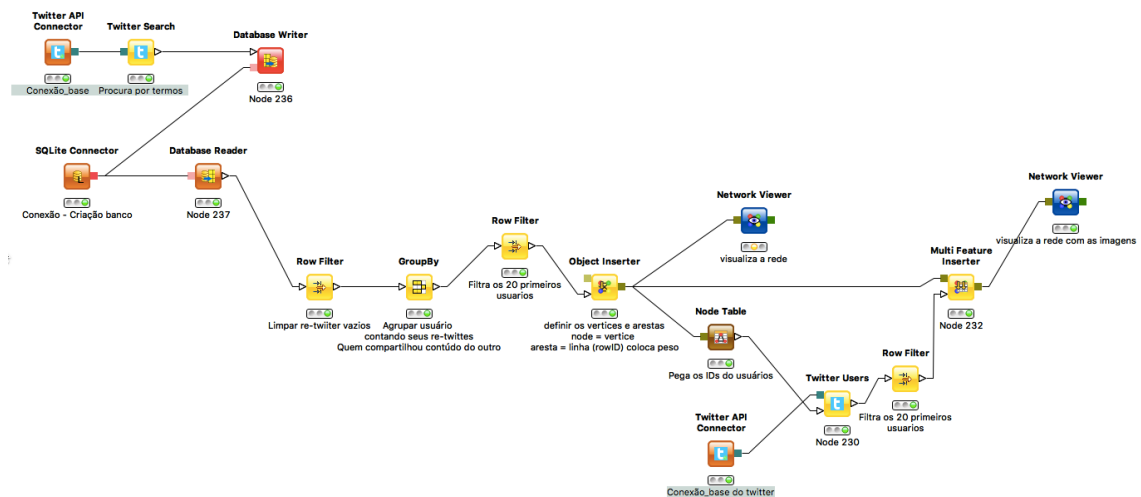


Figura 6. Fluxo desenvolvido no Knime – Relação entre usuários

## Análises de Sentimentos

Com base na nuvem de tags gerada na primeira análise, nosso segundo passo foi criar uma nuvem de tags com análise de sentimentos sobre o a força tarefa ilustrada na figura 2. De acordo com o conhecimento prévio do assunto, a força tarefa, é uma equipe investigatória do ministério publico federal formada por procuradores da República que estão na linha de frente da investigação na primeira instância da Justiça Federal do Paraná. A figura 7 e 8 ilustra os resultados dessa pesquisa.



Figura 7. Nuvem de TAGs frequência das palavras – Tema: Força tarefa



Figura 8. Nuvem de TAGs frequência das palavras – Tema: Força tarefa

Como realizado na pesquisa anterior, utilizamos dois modos de extração dos dados. Podemos observar na Figura 7, através da frequência de TAGs relacionadas a “CGU” que a usuários do twitter estão descontentes com o fim da CGU, órgão responsável pela prevenção e combate à corrupção no funcionalismo. A figura 9 ilustra a pesquisa realizada na internet sobre o tema ilustrado na nuvem de TAGs.

**CGU — Ministério da Transparência, Fiscalização e Controle**  
[www.cgu.gov.br/](http://www.cgu.gov.br/) ▼  
 Site da Controladoria-Geral da União (CGU), órgão do Governo Federal responsável pela defesa do patrimônio público, transparência e combate à corrupção.

**Extinção Da CGU É Classificada Como "Retrocesso" Pelos ...**  
<https://www.brasildefato.com.br/.../extincao-da-cgu-e-classificado-como-retrocesso-pe...> ▼  
 19 de mai de 2016 - A extinção da Controladoria Geral da União (CGU), órgão responsável pela

Figura 9. Resultados da pesquisa referente a CGU

Por outro lado, a Figura 8 ilustra palavras que os usuários estavam retratando referente a um projeto de lei que o senador Renan Calheiros, que pode atrapalhar as investigações da força tarefa da lava jatos. A Figura 10 demonstra a reportagem publicada na internet sobre o tema.

folhapolitica.org

sábado, 9 de julho de 2016

## Força-tarefa da Lava Jato pede ajuda contra projeto de lei de Renan

 Curtir 9,7 mil  78  Tweet  Pin it



O procurador e coordenador da força-tarefa da Lava Jato, Deltan Dallagnol, usou as redes sociais para fazer um apelo à população. Dallagnol pede que a população vote contra o projeto de lei sobre abuso de autoridade que Renan Calheiros desembuchou para atrapalhar a Lava Jato. Os cidadãos podem votar no site do Senado, [neste link](#).

Leia abaixo o texto publicado por Deltan Dallagnol:

Imagem: Reprodução / Redes Sociais

Ajude a Lava Jato na votação online do Senado e compartilhe!!O Projeto de Lei 280/2016, que visa disciplinar o abuso de autoridade, é um cavalo de Tróia que pretende calar investigadores e juízes e obstruir as investigações da Lava Jato. Embora fosse legítimo discutir aumento de pena para abuso de autoridade e criar novas hipóteses, o conteúdo, o momento e a "urgência" da proposta mostram claramente para que veio. Vote contrariamente a esse projeto de lei, **no link abaixo**, garantindo um Ministério Público forte e um Judiciário independente, essenciais para a democracia e para o combate à corrupção de poderosos. A lei deve valer para todos. Deltan, Andrey, Welter, Athayde, Carlos Fernando, Diogo, Isabel, Januário, Jerusa, Julio, Laura, Orlando e Roberson (procuradores da Força Tarefa da Lava Jato)

Figura 10. Pesquisada de acordo com dados ilustrados na nuvem de TAG



## **Dificuldades e Desafios**

A grande dificuldade encontrada na realização do trabalho foi primeiramente descobrir o tema a ser analisado, bem como, coletar as informações sobre o assunto. Quando fazemos essas coletas, muitas das vezes, coletamos “lixos” palavras e twitter que não são de interesse da pesquisa da coleta. Outra dificuldade encontrada foi o que analisar depois de coletado os dados. Como extrair o conhecimento daquele assunto, como tentar representar graficamente o conteúdo coletado. Por fim, o resultado analisado na segunda pesquisa com o tema “força tarefa” apesar de satisfatória e trazendo uma informação importante para o interessado, não foi possível realizar a análise de sentimentos da forma que gostaria ter realizado. Desse modo, houve uma dificuldade para gerar a análise de sentimentos com as palavras no vocabulário no idioma português brasileiro.

## **Conclusão**

O presente trabalho demonstrou que é possível extrair informações importantes de dados coletados no twitter, bem como descobrir quais pessoas que contribuem mais com a disseminação dos dados e sua rede de influências. Também mostrou efetivo de uma forma visual, demonstrar as principais palavras que os usuários do twitter estão dizendo na rede. Essas análises geradas fornece ao interessado informações importantes para tomar alguma decisão e tentar atuar em um determinado ponto abordado. No Estudo de caso abordado nesse trabalho, a operação lava jato, podemos levantar alguns aspectos que estão sendo recorrentes no Brasil e observar que as informações coletadas são coerentes com a situação atual brasileira.