Summary

[Youtube Project – Relatório do foi feito 1](#_Toc25151943)

[Qual a importância de analisar as mídias sociais? 1](#_Toc25151944)

[Criação de um banco de dados com uma grande quantidade de informação. 1](#_Toc25151945)

[Referências 2](#_Toc25151946)

# Youtube Project – Documentação

O objetivo geral desse projeto, é analisar o cenário de discussão política e a dinâmica das narrativas nas mídias sociais a partir de comentários em vídeo do Youtube referente ao ano de 2018 e 2019. Essa análise foca em 4 pilares básico, sendo eles:

1. Criação de um banco de dados com uma grande quantidade informação.
2. Análise dos trabalhos realizados na área.
3. Desenvolvimento e aplicação de algoritmos para tratamento de dados.
4. Desenvolvimento e aplicação de técnicas de mineração de dados.
5. Reflexão das técnicas aplicadas e dos resultados obtidos.

Para gerenciar o projeto foi usado o repositório GitLab para colocar os algoritmos, e o Google Drive para armazenar o banco de dados criado.

## Qual a importância de analisar as mídias sociais? (em andamento)

Na parte das eleições, as mídias sociais se tornaram relevantes desde 2008, quando o Barack Obama foi eleito conquistando boa parte do seu público usando as redes, a partir daí, foi relatado o interesse em usar as mídias como parte da estratégica política para ganhar voto [em andamento].

## Criação de um banco de dados com uma grande quantidade de informação.

Essa tarefa se divide em duas partes, a primeira é selecionar os vídeos em que os comentários serão coletados, e a segunda é coletar os comentários. Nesta etapa foi usado a API do Youtube Data v3, na linguagem Python com notebook Jupyter.

Para fazer a seleção de vídeo foi definido um conjunto de palavras chaves, para cada palavra chave foi selecionado 50 vídeos, organizado do mais visto ao menos visto

A seleção de vídeos foi feita usando as seguintes palavras chaves, Amoedo, Alckmin, Bolsonaro, Boulos, Ciro, Manuela, Haddad, Mourão, Lula, e para cada consulta a API retorna no máximo 50 vídeos, ou seja, no total foram selecionados 450 vídeos.

Além das palavras chaves outros critérios foram adicionados, como a data de publicação do vídeo tem que ser do ano referente ao que foi proposto neste projeto, isso é, um só com vídeos de 2018 e outro só com vídeos de 2019, e ter o código da região BR.

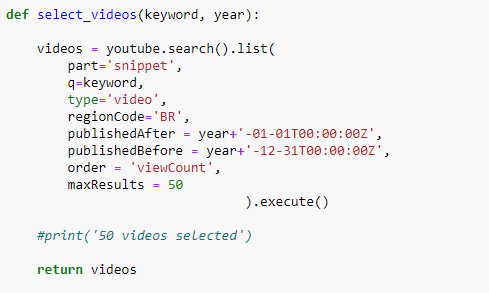


Figura 1 - Função que faz a seleção dos vídeos por palavra chave

Cada vídeo coletado segue a seguinte estrutura JSON:

{

'kind': 'youtube#searchResult',

'etag': '"p4VTdlkQv3HQeTEaXgvLePAydmU/EJxKkgAmsuSsRl9cVoyWW8iBneY"',

'id': {'kind': 'youtube#video', 'videoId': 'PDRbhocduI8'},

'snippet': {'publishedAt': '2018-10-05T11:00:09.000Z', 'channelId': 'UC-6xqzMBF2CXTImn\_a4aCVg', 'title': 'Marcelo Adnet imita Boulos, Amoêdo, Meirelles, Alvaro Dias, Eymael, Vera Lucia e João Goulart Filho', 'description': "Agora o Marcelo Adnet foi longe demais! No nono episódio da série 'Tutorial dos Candidatos' o humorista imita todos os presidenciáveis que faltavam.", 'thumbnails': {'default': {'url': 'https://i.ytimg.com/vi/PDRbhocduI8/default.jpg', 'width': 120, 'height': 90}, 'medium': {'url': 'https://i.ytimg.com/vi/PDRbhocduI8/mqdefault.jpg', 'width': 320, 'height': 180}, 'high': {'url': 'https://i.ytimg.com/vi/PDRbhocduI8/hqdefault.jpg', 'width': 480, 'height': 360}}, 'channelTitle': 'Jornal O Globo', 'liveBroadcastContent': 'none'}

}

Das informações do vídeo que a API disponibilizou, nós vamos pegar o ID, o título, a data de publicação, e a descrição. As informações foram armazenadas em arquivos JSON.

Uma imagem contendo captura de tela

Descrição gerada automaticamente

Figura 2 - Função que coleta as informações e armazena em um arquivo JSON

Das informações do vídeo que a API disponibilizou, nós vamos pegar o ID, o título, a data de publicação, e a descrição. Os vídeos foram selecionados no dia 25/12/2019, as informações dos vídeos foram armazenadas em arquivos JSON.

## Referências