

# Election Campaign Dynamics

Updates, Website, Future additions

---

Anahita Poulad

Emre Canbazer

Amir Jamali

Melina Chioutakou

December 10, 2021

# OUTLINE

1. Updates

2. Website

3. Future additions

# Updates

---

# UPDATES

- Updates for the YouTube Dataset
  - speech split into time sequences to modelize topics along a long speech.
  - challenge: auto-generated subtitles

REPLAY - Crise sanitaire: Marine Le Pen dénonc...	🇫🇷 Nous avons besoin de vous dès maintenant ...	2021-07-04 14:03:32	50:25	10.650	nous avons été obligés en cette période
REPLAY - Crise sanitaire: Marine Le Pen dénonc...	🇫🇷 Nous avons besoin de vous dès maintenant ...	2021-07-04 14:03:32	50:25	12.480	étrange de composer avec les exigences
REPLAY - Crise sanitaire: Marine Le Pen dénonc...	🇫🇷 Nous avons besoin de vous dès maintenant ...	2021-07-04 14:03:32	50:25	14.670	sanitaires
REPLAY - Crise sanitaire: Marine Le Pen dénonc...	🇫🇷 Nous avons besoin de vous dès maintenant ...	2021-07-04 14:03:32	50:25	15.920	je salue donc les personnes qui nous

# UPDATES

- Updates for the website
  - from C# to Flask
  - challenge: installing required Python packages to the Microsoft Azure server
  - temporarily local server

# Website

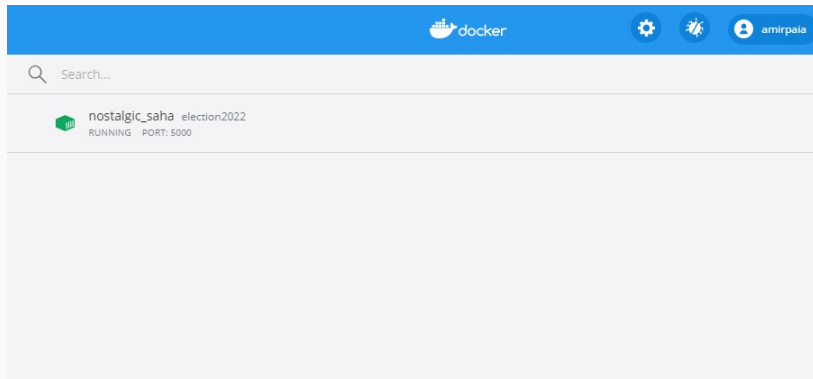
---

# WEBSITE WITH FLASK

- Previous Plan: Using two services in C# and Python, which talk together (It has unnecessary complexity)
- Now: we developed the same website with Python Flask

# BUILDING DOCKER IMAGE

A docker image has been built locally



Election2022



# DEPLOYMENT ON AZURE

This task is already started, but will be finished for the next presentation

Microsoft Azure

Search resources, services, and docs (G+)

Home > App Services >

**election2022**  
App Service

Search (Ctrl+/)

Overview

Activity log

Access control (IAM)

Tags

Diagnose and solve problems

Security

Events (preview)

Deployment

Quickstart

Deployment slots

Deployment Center

Settings

Stopped

Essentials

JSON View

Resource group (Move)	: amirpaia_rg_linux_southcentr...
Status	: Stopped
Location	: Central US
Subscription (Move)	: Azure subscription 1
Subscription ID	: 2f82af9b-1904-49d8-9b97-39...
URL	: https://election2022.azurewe...
Health Check	: Not Configured
App Service Plan	: ASP-amirpaiaLinuxouthce...
FTP/deployment username	: No FTP/deployment user set
FTP hostname	: ftp://waws-prod-dm1-199.ft...
FTPS hostname	: ftps://waws-prod-dm1-199.ft...

## **Future additions**

---

# BERTOPIC

Add customizable dynamic topic modeling with BERTopic:

- updated preprocessing (from NLTK lemmatizer to spaCy)
- export all topics over time

**Figure 1:** Topics over time .csv

1	Topic Words	Frequency	Timestamp	Name
2	0 populaire, conventionpopulaire, voter, campagne, conventionpopulaire	55	24:10.5	0_pouvoir_élection_voter_sondage
3	0 élection, vote, voter, présidentiel, liberté	38	38:50.9	0_pouvoir_élection_voter_sondage
4	0 falloir, mort, présidentiel, candidat, débat	36	52:39.8	0_pouvoir_élection_voter_sondage
5	-1 jesoutiensmelenchon, france, débat, culture, contre	34	52:39.8	-1_france_macron_guerre_contre
6	0 vote, sondage, assemblée, pouvoir, politique	30	41:52.4	0_pouvoir_élection_voter_sondage
7	0 sondage, participation, élection, liberté, politique	29	59:34.2	0_pouvoir_élection_voter_sondage
8	-1 france, macron, euro, monsieur, kiosque	28	24:10.5	-1_france_macron_guerre_contre
9	-1 guerre, avenir, afghanistan, maritime, politique	27	38:50.9	-1_france_macron_guerre_contre
10	3 élève, éducation, vacciné, lycée, interdire	23	48:46.8	3_radiopeyi_école_élève_enseignant
11	-1 jlmescp, france, avion, allemagne, combat	20	34:57.9	-1_france_macron_guerre_contre
12	-1 melenchonguyane, france, scolariser, monde, contre	18	48:46.8	-1_france_macron_guerre_contre
13	1 alimentaire, milliard, million, fiscal, pauvreté	18	24:10.5	1_alimentaire_million_milliard_payer
14	-1 france, étranger, débat, confrontation, algérie	17	59:34.2	-1_france_macron_guerre_contre
15	0 police, média, stratégie, dire, lutte	17	45:45.3	0_pouvoir_élection_voter_sondage
16	1 alimentaire, million, besoin, social, pays	17	52:39.8	1_alimentaire_million_milliard_payer
17	0 pouvoir, liberté, syndicaliste, président, condamner	16	48:46.8	0_pouvoir_élection_voter_sondage
18	-1 contre, urgence, pandémie, net, euro	14	45:45.3	-1_france_macron_guerre_contre
19	-1 débat, contre, république, france, révolutionnaire	14	10:21.6	-1_france_macron_guerre_contre
20	0 réconciliation, publier, sondage, voter, journeemondialesanimau	13	17:16.1	0_pouvoir_élection_voter_sondage
21	1 argent, payer, capitalisme, milliard, société	13	45:45.3	1_alimentaire_million_milliard_payer
22	4 nucléaire, centrale, énergie, renouvelable, besoin	13	52:39.8	4_nucléaire_énergie_réacteur_plan

Problem with BERTopic: Really time-consuming

# KEYBERT

```
In [ ]: 1 from keybert import KeyBERT
        2 kw_model = KeyBERT()

In [36]: 1
        2 df_candidate_tweets['keywords_list'] = df_candidate_tweets['preprocessing_spacy2'].apply(kw_model.extract_keywords)

In [18]: 1 df_candidate_tweets['keywords'][0]

Out[18]: [('surveillance', 0.4894),
          ('terroristes', 0.4589),
          ('palpation', 0.3328),
          ('affaires', 0.2849),
          ('bien', 0.2825)]

In [40]: 1 def get_keywords(key_list):
        2     keys=''
        3     if key_list:
        4
        5         for key,score in key_list:
        6             if score>0.3:
        7                 keys = keys+' '+key
        8
        9     return keys.strip()

In [42]: 1 df_candidate_tweets['keywords'] = df_candidate_tweets['keywords_list'].apply(get_keywords)
```

**Figure 2:** Implementing KeyBERT

# KEYBERT

```
In [42]: 1 df_candidate_tweets['keywords'] = df_candidate_tweets['keywords_list'].apply(get_keywords)

In [43]: 1 tweets_k = df_candidate_tweets.keywords.to_list()

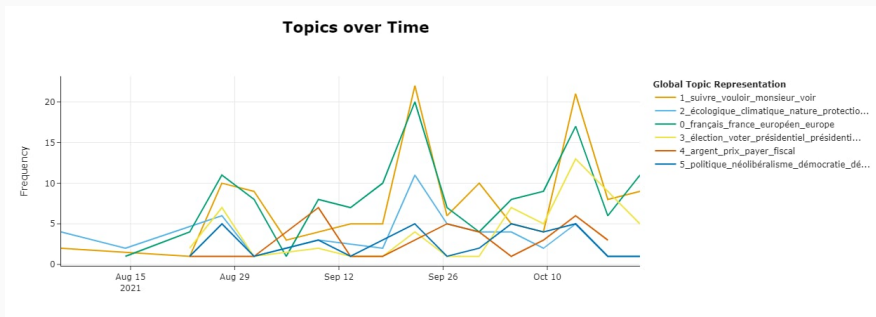
In [44]: 1 topic_model = BERTopic(language="French")
        2
        3 topics, probs = topic_model.fit_transform(tweets_k)

In [50]: 1 type(topics_over_time)
        2 topics_over_time.to_csv('keyword_topics_overtime.csv', index = False)

In [45]: 1 topics_over_time = topic_model.topics_over_time(tweets_k, topics, timestamps,
        2                                                    global_tuning=False,
        3                                                    evolution_tuning=True,
        4                                                    nr_bins=20)
        5 # topics_over_time = topic_model.topics_over_time(tweets, topics, timestamps, nr_bins=20)
        6 # visualization
        7 topic_model.visualize_topics_over_time(topics_over_time, top_n_topics=6)
```

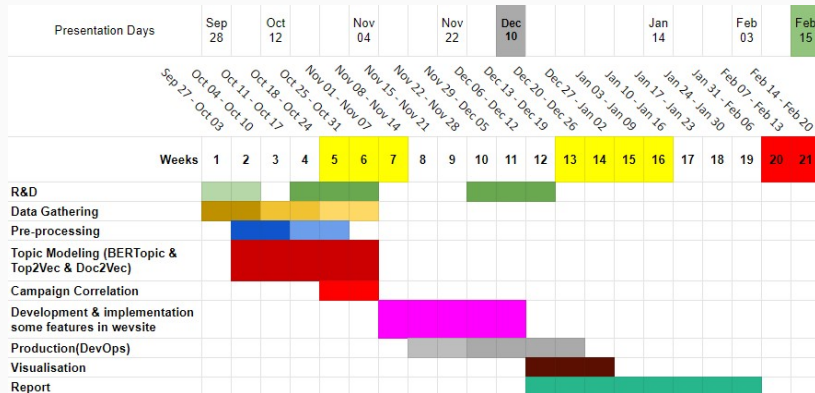
**Figure 3:** BERTopic on keywords

# KEYBERT



**Figure 4:** BERTopic on keywords

# TIMELINE



**Thank you!**