

LLMs for political speech analysis

the use case of the Greek
National Elections 2023

Pavlos Sermpezis
Researcher @ Datalab, AUTH

Datalab at a glance



- Data Analytics
- Social Network Analysis
- Core Data Management
- Sensing Data Exploration
- Massively Parallel Solutions
- Sentiment Analytics
- AI & Machine Learning
- Decentralized Technologies



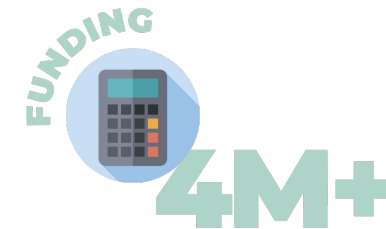
- Advanced Analytics Algorithms
- Health & Life Sciences
- IoT & cross-cloud data management
- Multi-domain data phenomena detection
- Industry 4.0 & Smart Cities
- Predictive maintenance
- Anomaly detection
- Internet of Everything & Sensing data



Strong Research team
with multidisciplinary
background



170 scientific
publications in the
last 5 years



Over 4M EURO received
in the last 5 years



EU & National

The project

lab.imedd.org/en/elections-2023

Analyzing the pre-election campaign speeches of Greece's main political leaders



*Partnership with iMedD Lab

The project

lab.imedd.org/en/elections-2023

Analyzing the pre-election campaign speeches of Greece's main political leaders

Project goals / questions

- What **issues** do political leaders raise in the public discourse?
- What **sentiments** do they convey?
- To what extent can **polarization and populism** be detected in their rhetoric?

The project

lab.imedd.org/en/elections-2023

Analyzing the pre-election campaign speeches of Greece's main political leaders

What we analyzed:

- **6 politicians**
- **171 speeches**
- **4,920 paragraphs**
- **675,582 words**

The project

lab.imedd.org/en/elections-2023

Analyzing the pre-election campaign speeches of Greece's main political leaders

<https://lab.imedd.org/en/elections-2023>



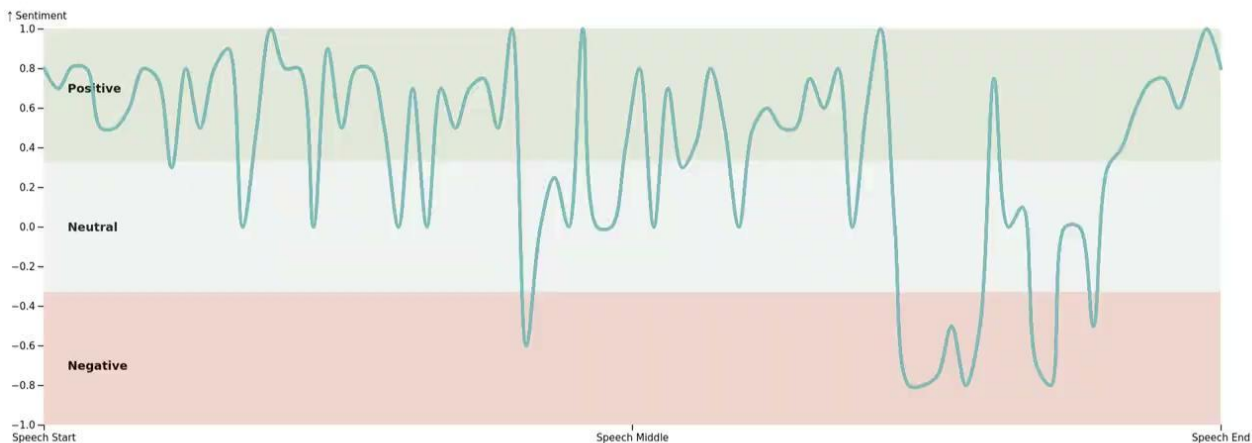
Political Discourse

lab.imedd.org/en/elections-2023

SENTIMENT EVOLUTION DURING TALK

Kiriakos Mitsotakis

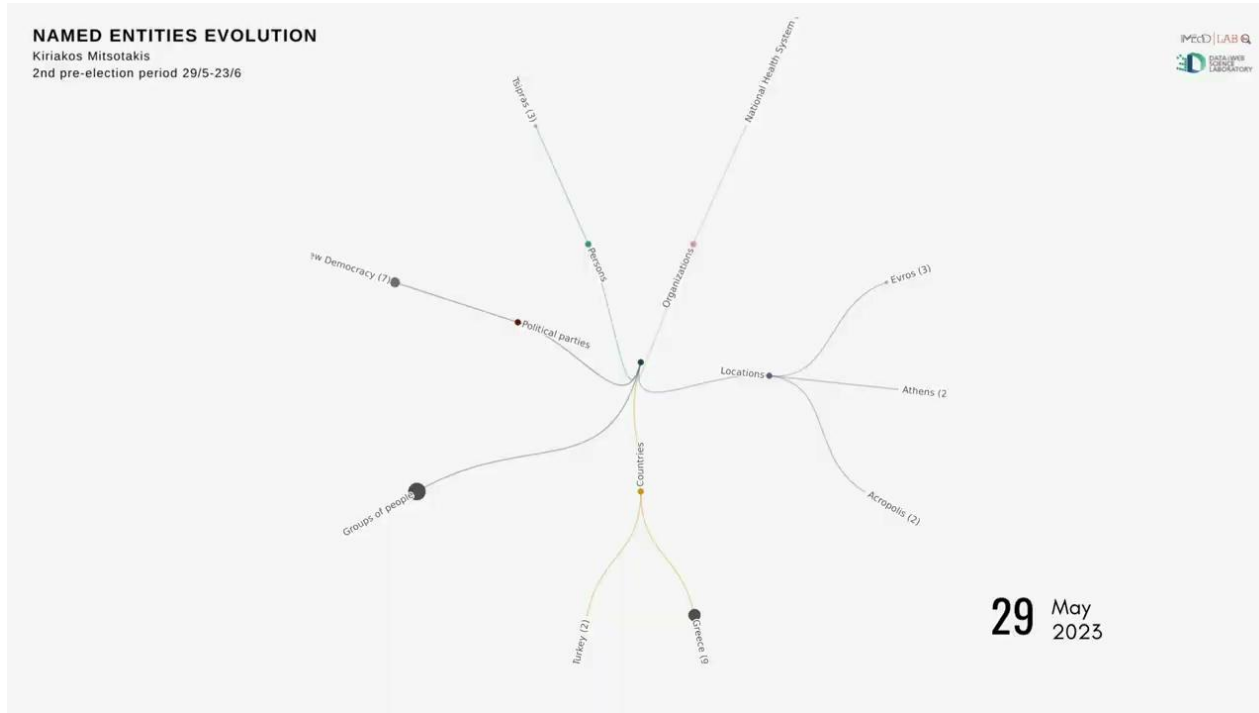
1st pre-election period 22/4-19/5



22 April 2023 CHANIA

Political Discourse

lab.imedd.org/en/elections-2023



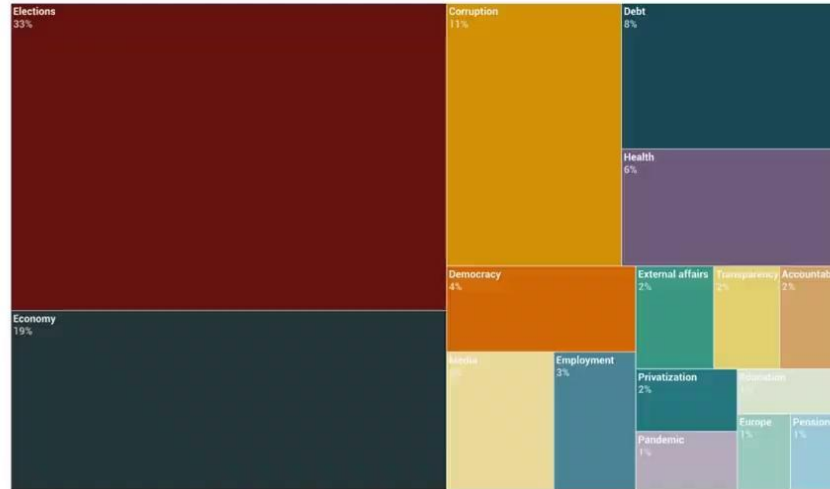
*Partnership with iMedD Lab

Political Discourse

lab.imedd.org/en/elections-2023

TOPICS DISCUSSED DURING TALK

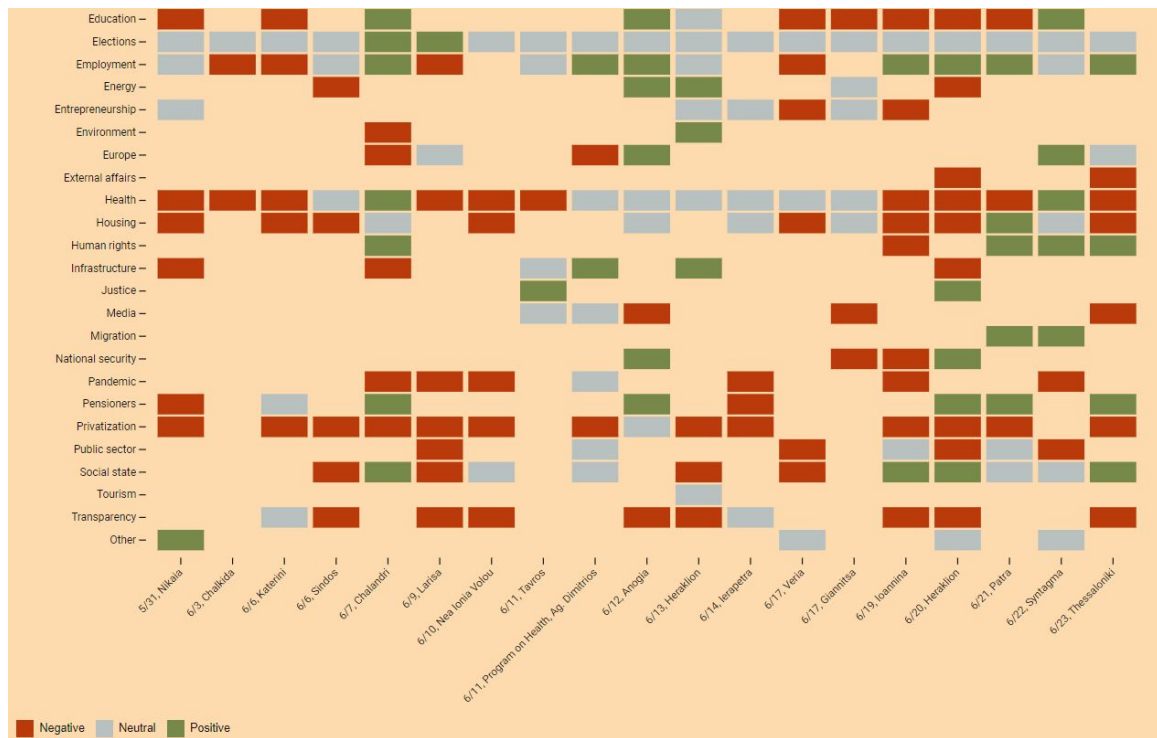
Alexis Tsipras
1st pre-election period 22/4-19/5



24 April 2023 KOS

Political Discourse

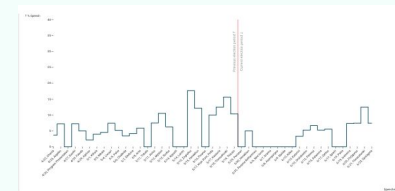
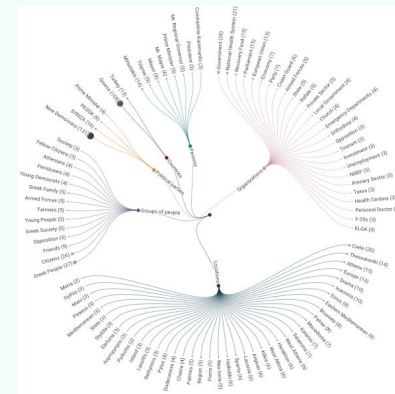
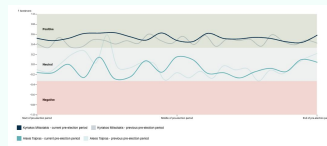
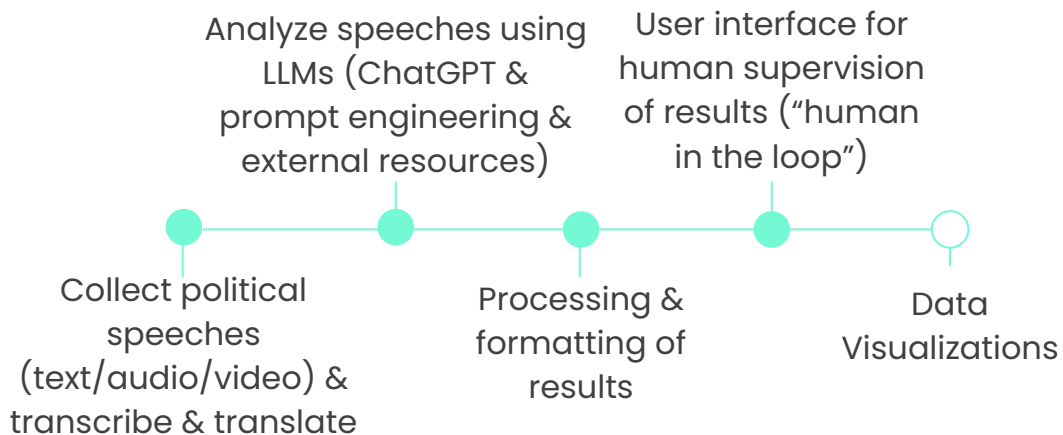
lab.imedd.org/en/elections-2023



*Partnership with iMedD Lab

Political Discourse

How it works?



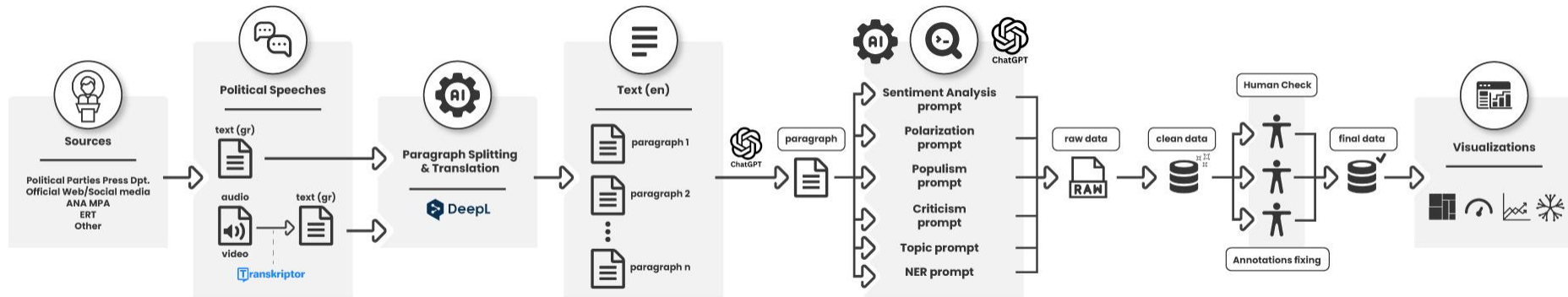
*Partnership with iMedD Lab

Why ChatGPT?

(vs. open pre-trained NLP models)

Task	Did NLP models work?
topic detection (per paragraph)	does not work
semantic similarity (per paragraph or speech)	so and so (not very reliably)
text similarity	does not work
named entity recognition	so and so (not very reliably)
hate speech detection	so and so (not very reliably)
populism detection	does not work
toxicity detection	works (more or less)
sentiment analysis	works (more or less)
polarization detection	does not work

The methodology.. in a nutshell



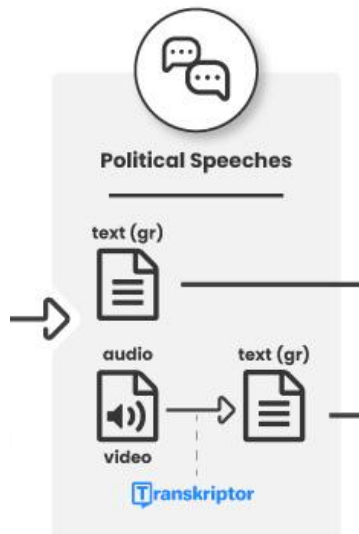
The methodology: sources



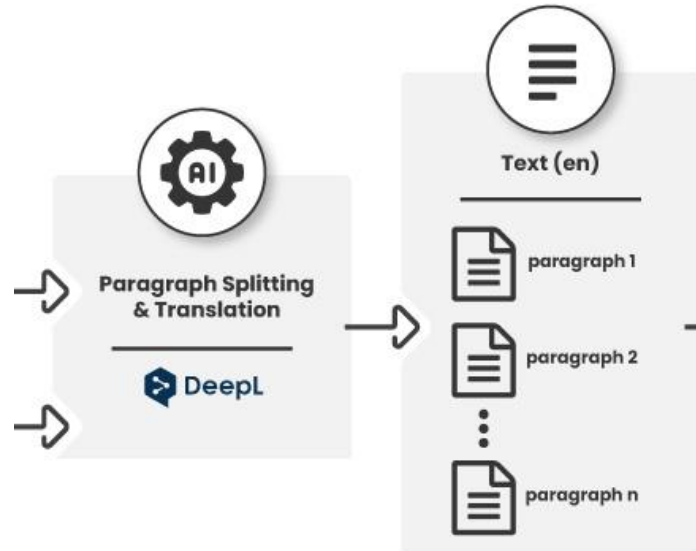
Sources

Political Parties Press Dpt.
Official Web/Social media
ANA MPA
ERT
Other

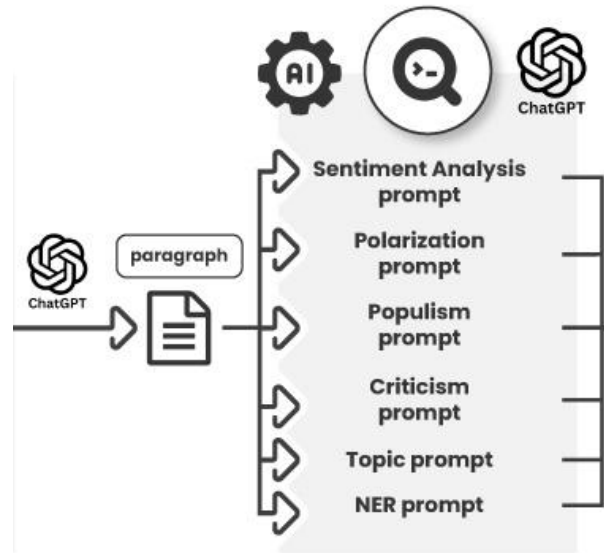
The methodology: ... to text



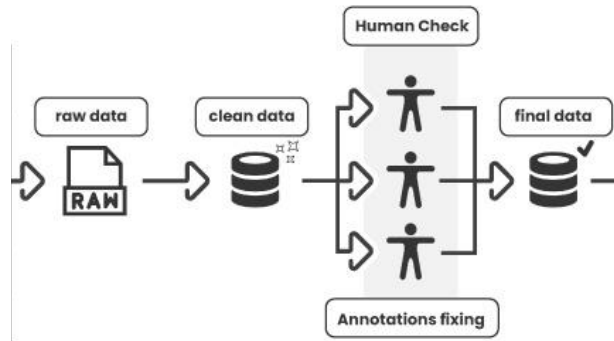
The methodology: paragraph splitting & translation



The methodology: analysis with ChatGPT



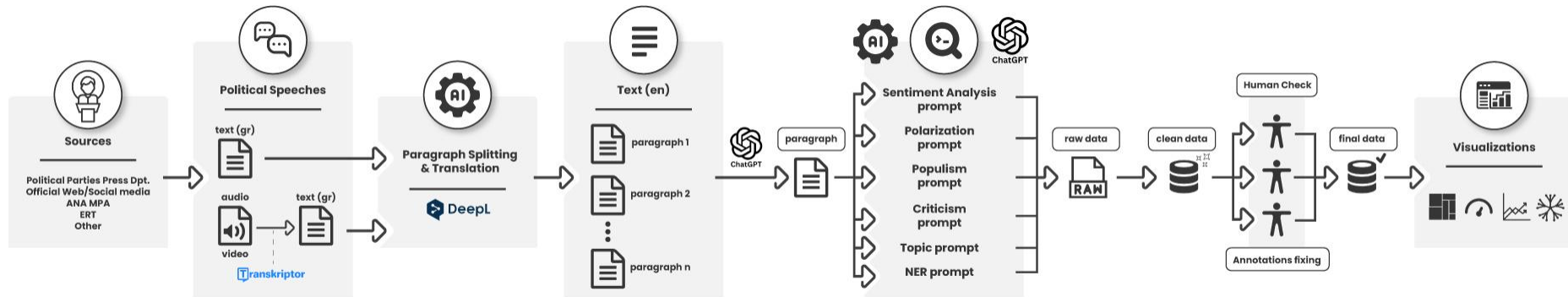
The methodology: from ChatGPT to validated data



The methodology: visualizations



The methodology.. in a nutshell



ChatGPT accuracy

(vs. manual annotation of humans)

Task	Most prevalent category (%)	Accuracy
sentiment	Positive (40%)	94.2%
topics	Election (26%)	61.7%
polarization	Zero/Low (86%)	87.4%
populism	Zero/Low (96%)	89.8%
criticism or agenda	Political agenda (61%)	89.3%

Results with other models: open LLMs

Model Name	Num of Parameters	Model Size (GB)
gemma	2b	1.4
gemma	7b	4.8
llama 2	7b	3.8
llama 2	13b	7.3

Results with other models: open LLMs

	Accuracy	Balanced Accuracy	Precision	Recall	F1
ChatGPT (fine-tuned)	0.69	0.62	0.71	0.69	0.7
ChatGPT	0.59	0.57	0.8	0.59	0.65
gemma:2b	0.32	0.26	0.54	0.32	0.35
gemma:7b	0.29	0.27	0.55	0.29	0.33
llama2:7b	0.09	0.08	0.40	0.09	0.10
llama2:13b	0.44	0.39	0.62	0.45	0.48

Thank You!

 <https://datalab.csd.auth.gr>

 info-datalab@csd.auth.gr