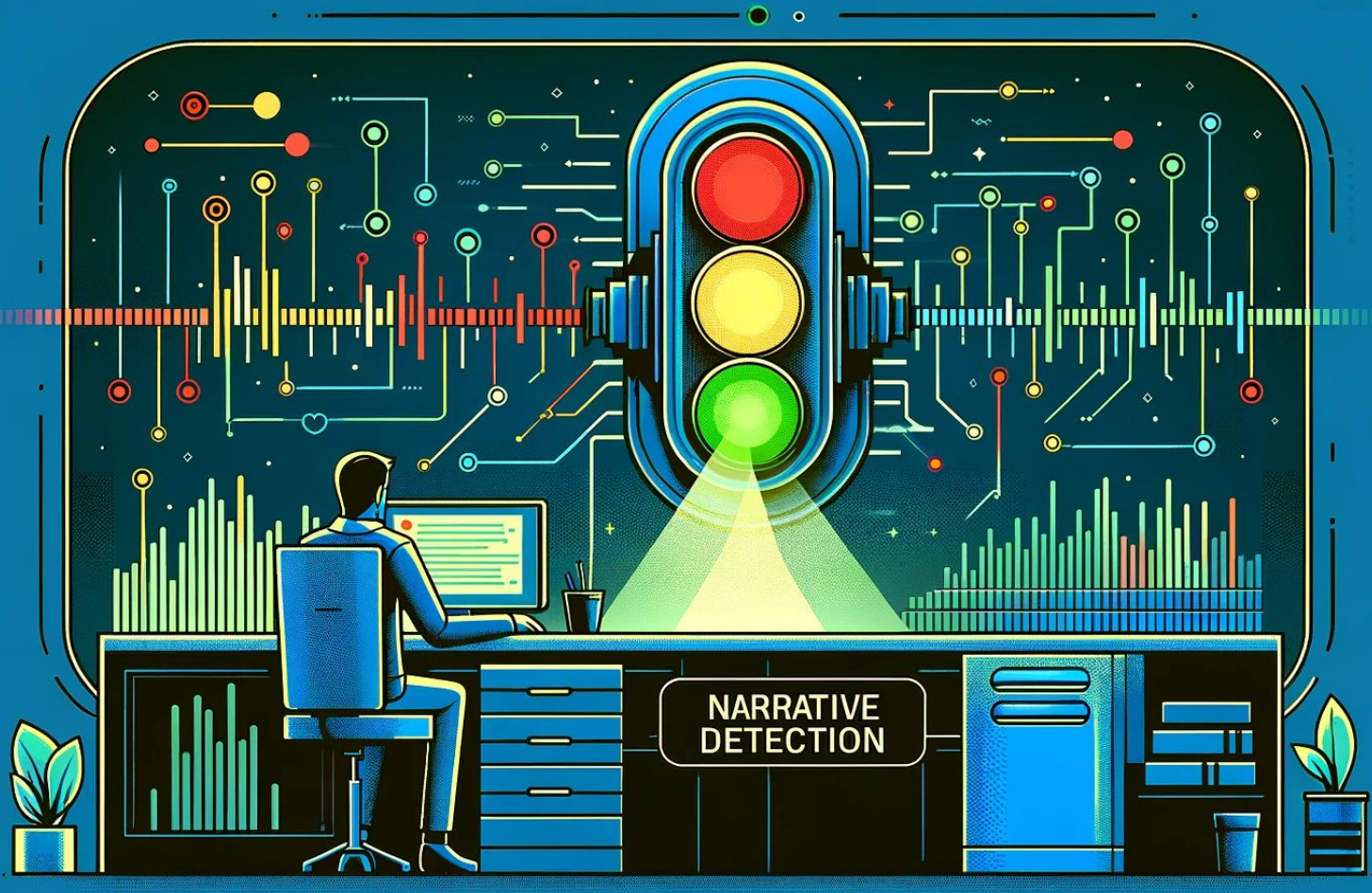

NARRATIVE
DETECTION

Table of Contents

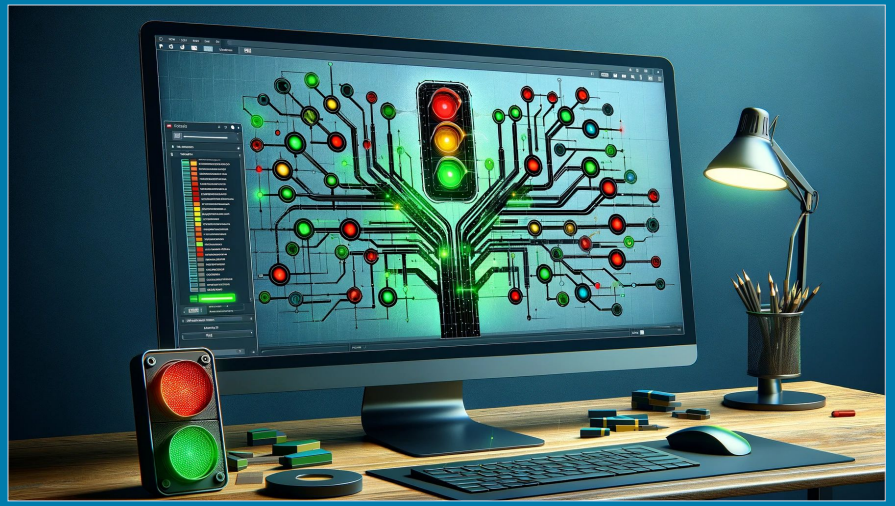
I. Research Goal
II. Data
III. Framework
IV. Evaluation
V. Results & Future Research
VI. Discussion

Research Goal

- **Design method that helps to get an overview of fact-checked claims about certain narratives.**

- **Approach:**
  Group claims into subsets and reorder them in narrative trees.

- **Background:** Utilizing clustering and visualisation techniques from the Topic-Detection And Tracking (TDT) and Newsfeed organization research fields.



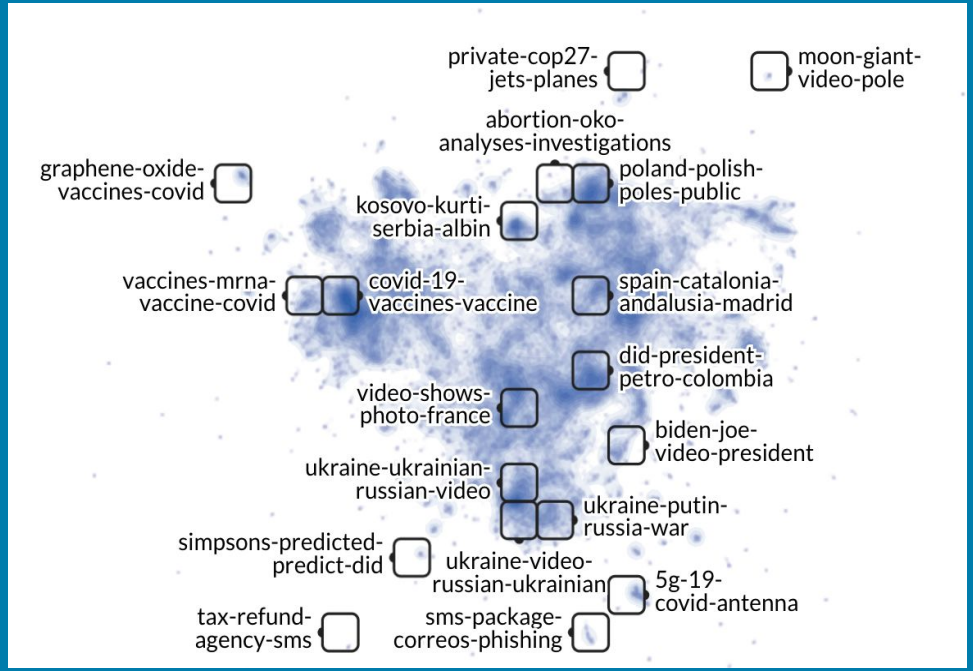
Illustrative Picture of a Generated Story Tree by DALL-E

Dataset features

“Biden's climate requirements will cut 90% of red meat from diet to a max 4 lbs per year and one burger per month.”

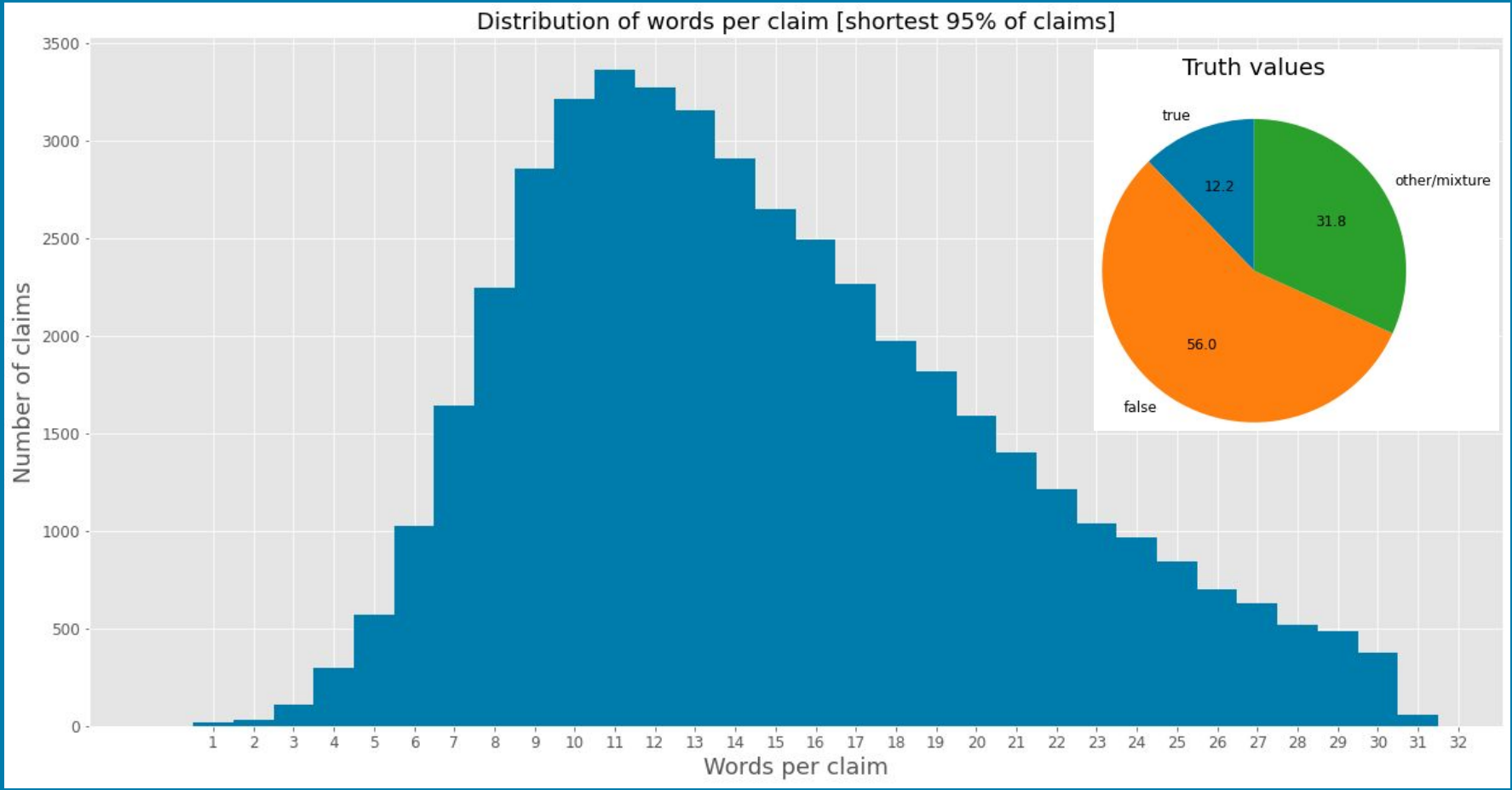
- Dataset: ClaimsKG by GESIS
- Previously fact-checked claims gathered from 13 fact-checking organizations in different languages
- 72k claims and counting (actively maintained)
- Why?
  - Standardized truth ratings (True, Mixed, False and Other)
  - Automated Entity Recognition and Linking applied

Source: https://data.gesis.org/claimskg/

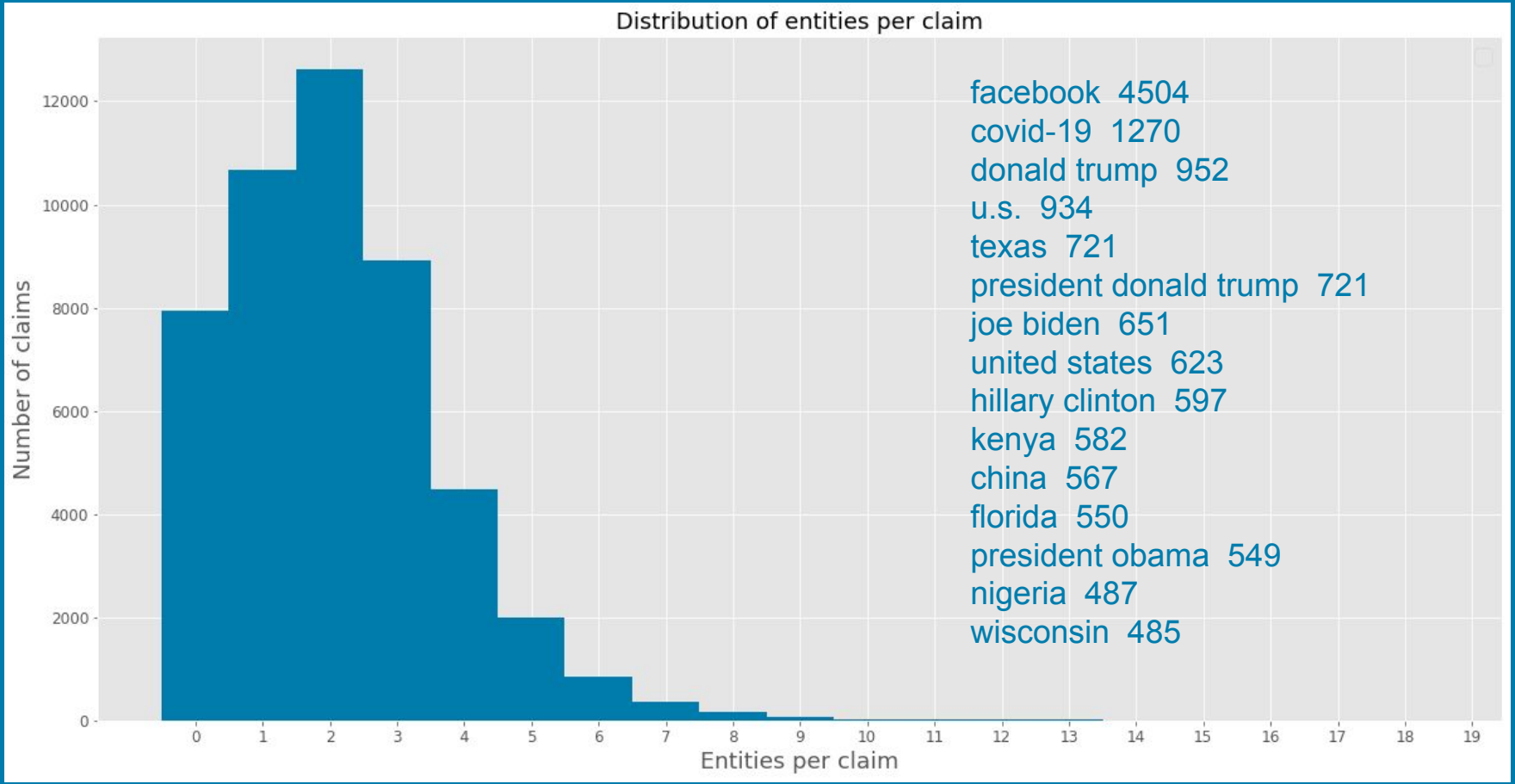


WizMap: Visualization of Embeddings

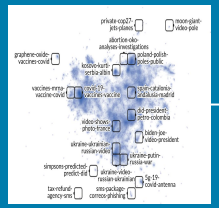
Rather Short and rather False



With Entities



Framework - Overview



ClaimsKG

Context: *Biden*

Topic 1: *Climate*

Topic 2: *Health*

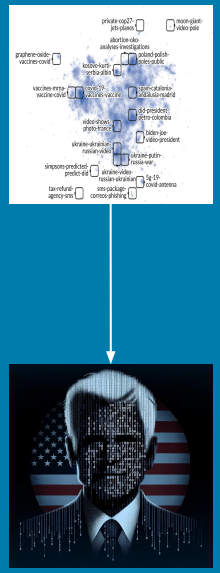
Narrative 1:
Meat Regulation

Narrative 2:
Energy Transition

Narrative 1:
COVID

Narrative 2:
Vaccine for hurricane protection

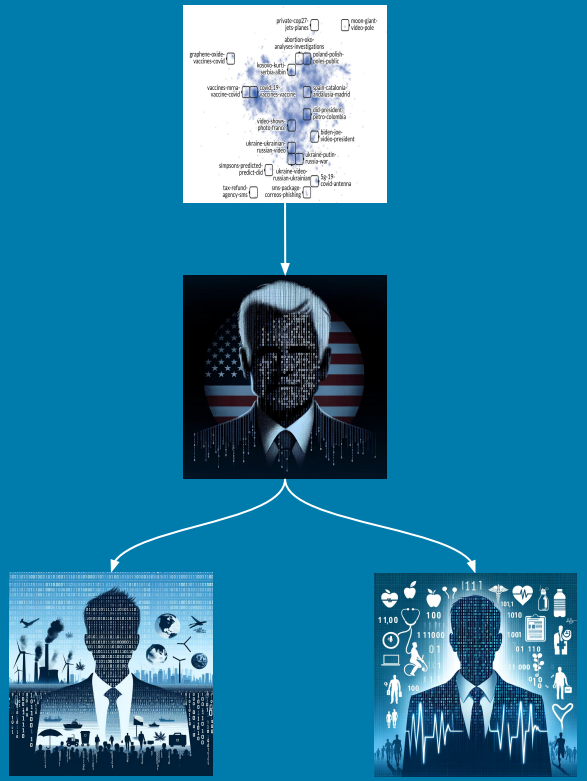
Part I - From Claims (Corpus) to Contexts



Context: *Biden*

- Capture Context by defining a set of Entities
- Why?
  - Common approach in communication/political science research to define a set of keywords
  - Subset selection to allow multiple contexts for a claim
- Case Study:
  - Biden Context: {“Biden”, "Sleepy Joe","Uncle Joe"}
  - 1596 claims

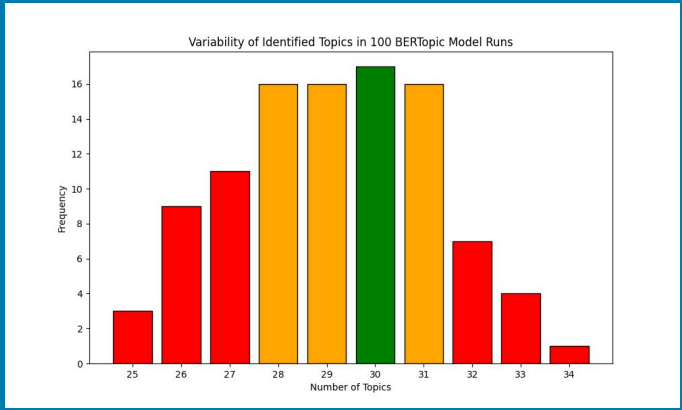
Part II - From Context to Topics



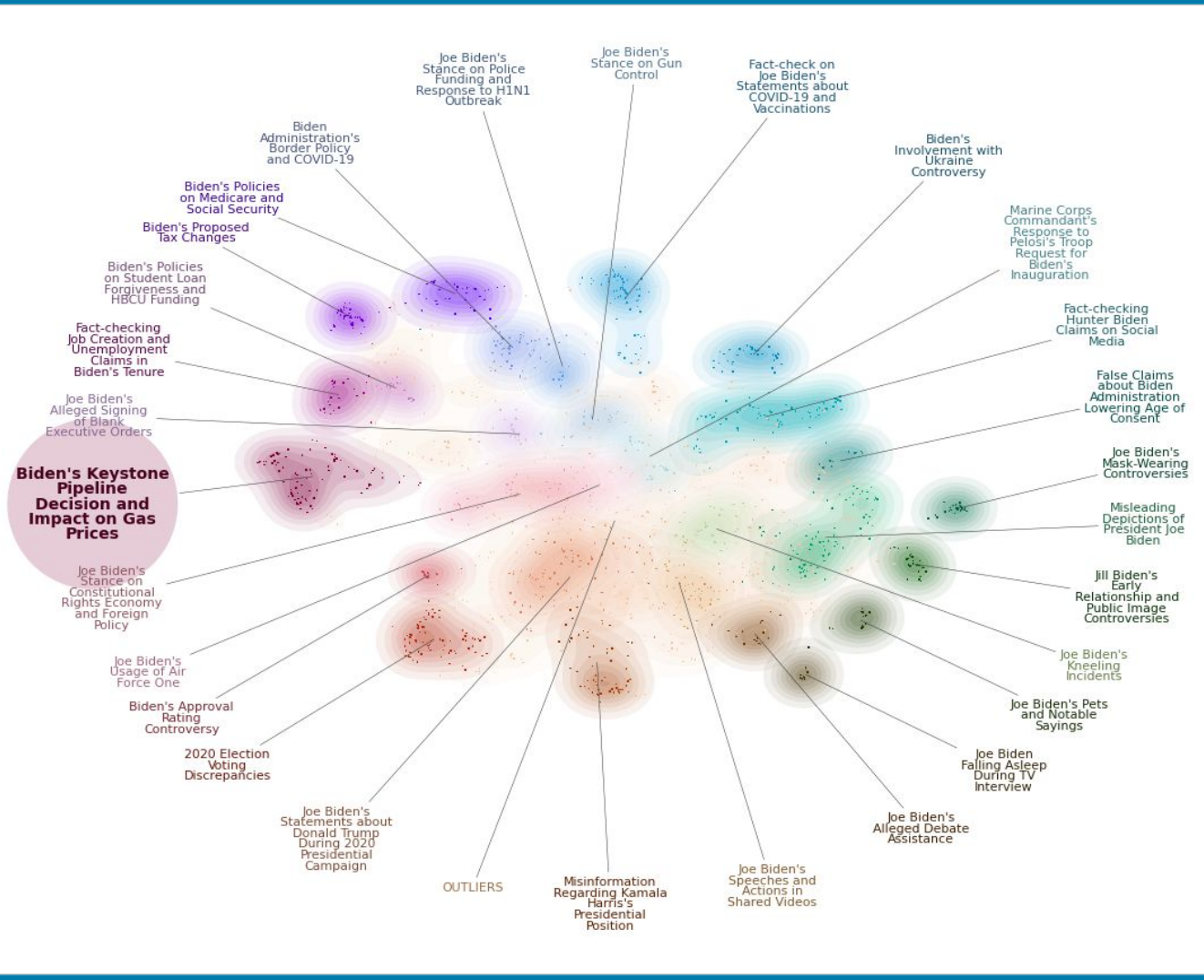
Topic 1: *Climate*

Topic 2: *Health*

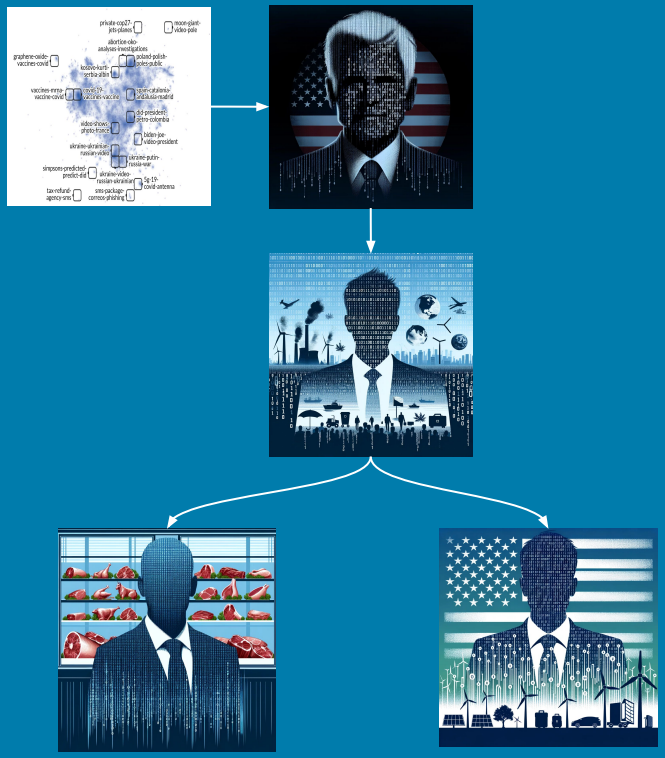
- BERTopic
- Assumption: a claim can occur only in a single topic
- Automatic Topic Label Generation using GPT-4
- 28 Topics for the Biden Context



BERTopic Distribution of Number of Topics across 100 run


Joe Biden's Stance on Police Funding and Response to H1N1 Outbreak
Joe Biden's Stance on Gun Control
Fact-check on Joe Biden's Statements about COVID-19 and Vaccinations
Biden Administration's Border Policy and COVID-19
Biden's Involvement with Ukraine Controversy
Biden's Policies on Medicare and Social Security
Marine Corps Commandant's Response to Pelosi's Troop Request for Biden's Inauguration
Biden's Proposed Tax Changes
Biden's Policies on Student Loan Forgiveness and HBCU Funding
Fact-checking Hunter Biden Claims on Social Media
Fact-checking Job Creation and Unemployment Claims in Biden's Tenure
False Claims about Biden Administration Lowering Age of Consent
Joe Biden's Alleged Signing of Blank Executive Orders
Joe Biden's Mask-Wearing Controversies
Biden's Keystone Pipeline Decision and Impact on Gas Prices
Misleading Depictions of President Joe Biden
Jill Biden's Early Relationship and Public Image Controversies
Joe Biden's Stance on Constitutional Rights Economy and Foreign Policy
Joe Biden's Kneeling Incidents
Joe Biden's Usage of Air Force One
Joe Biden's Pets and Notable Sayings
Biden's Approval Rating Controversy
Joe Biden Falling Asleep During TV Interview
2020 Election Voting Discrepancies
Joe Biden's Alleged Debate Assistance
Joe Biden's Statements about Donald Trump During 2020 Presidential Campaign
OUTLIERS
Misinformation Regarding Kamala Harris's Presidential Position
Joe Biden's Speeches and Actions in Shared Videos

Part III - From Topics to Narratives



Narrative 1:
Meat Regulation

Narrative 2:
Energy Transition

- Basic Assumption: a claim occurs in multiple narratives
- Tree Construction based on Cosine Similarity and Threshold. Select a claim: “Are here any similar claims”?
- Automatic Narrative Label Generation with GPT-4

Context: Biden Topic: Biden's Impact on Gas Prices, Keystone Pipeline Cancellation and Climate

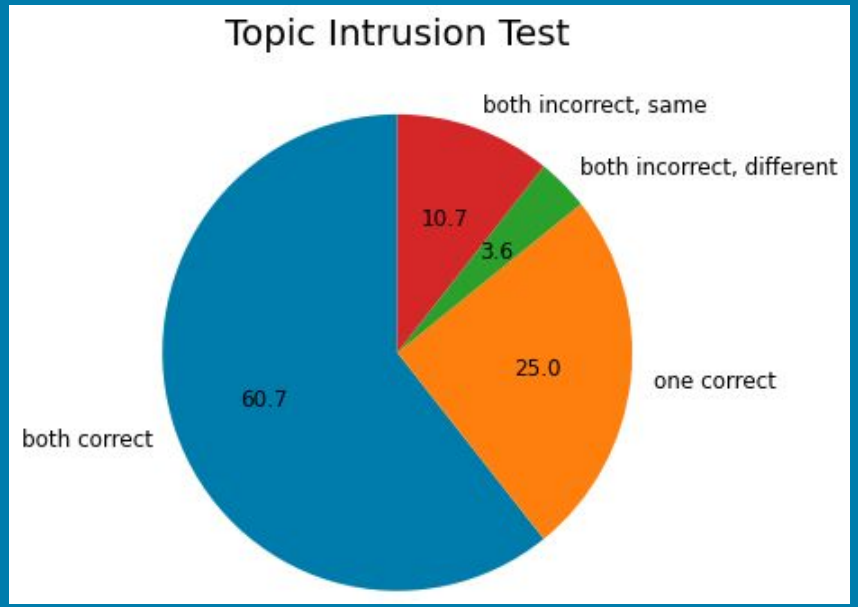
Evaluation: Intrusion Tests

>>> Please select the claim which fits the least into the group

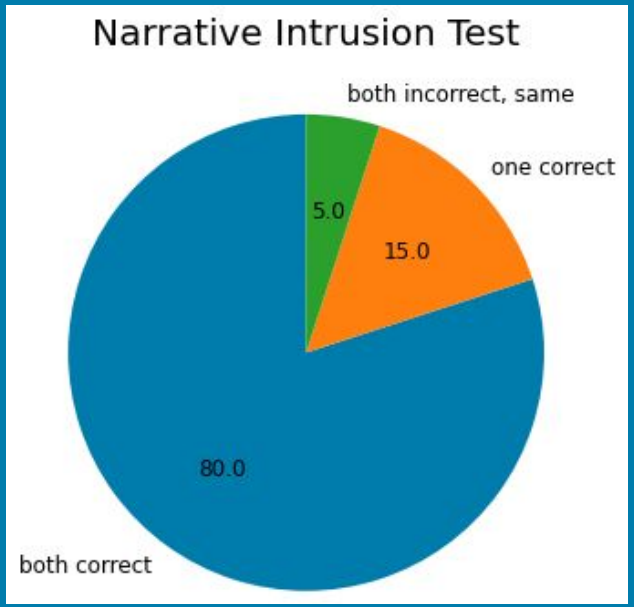
1. Biden's climate requirements will cut 90% of red meat from diet to a max 4 lbs per year and one burger per month.
2. Biden climate plan will cut American meat consumption
3. Says Joe Biden is going to control how much meat you can eat.
4. CNN reports Putin delays invasion until Biden sends weapons to Ukraine
5. Says President Joe Biden is trying to send taxpayer dollars to manufacturers overseas that do not abide by the same (carbon emission) standards we do at home.
6. Joe Biden said there is no climate crisis

Evaluation: Results

28 Topics, 5 True, 1 Intruder, 2 Raters



20 Narratives, 2-5 True, 1 Intruder, 2 Raters



Results, Limitations and Future Research

Results:

- Visualizing as Narrative Trees makes data quickly accessible
- Clustering into Topics and Narratives works well
- Framework is cheap to run and scales well

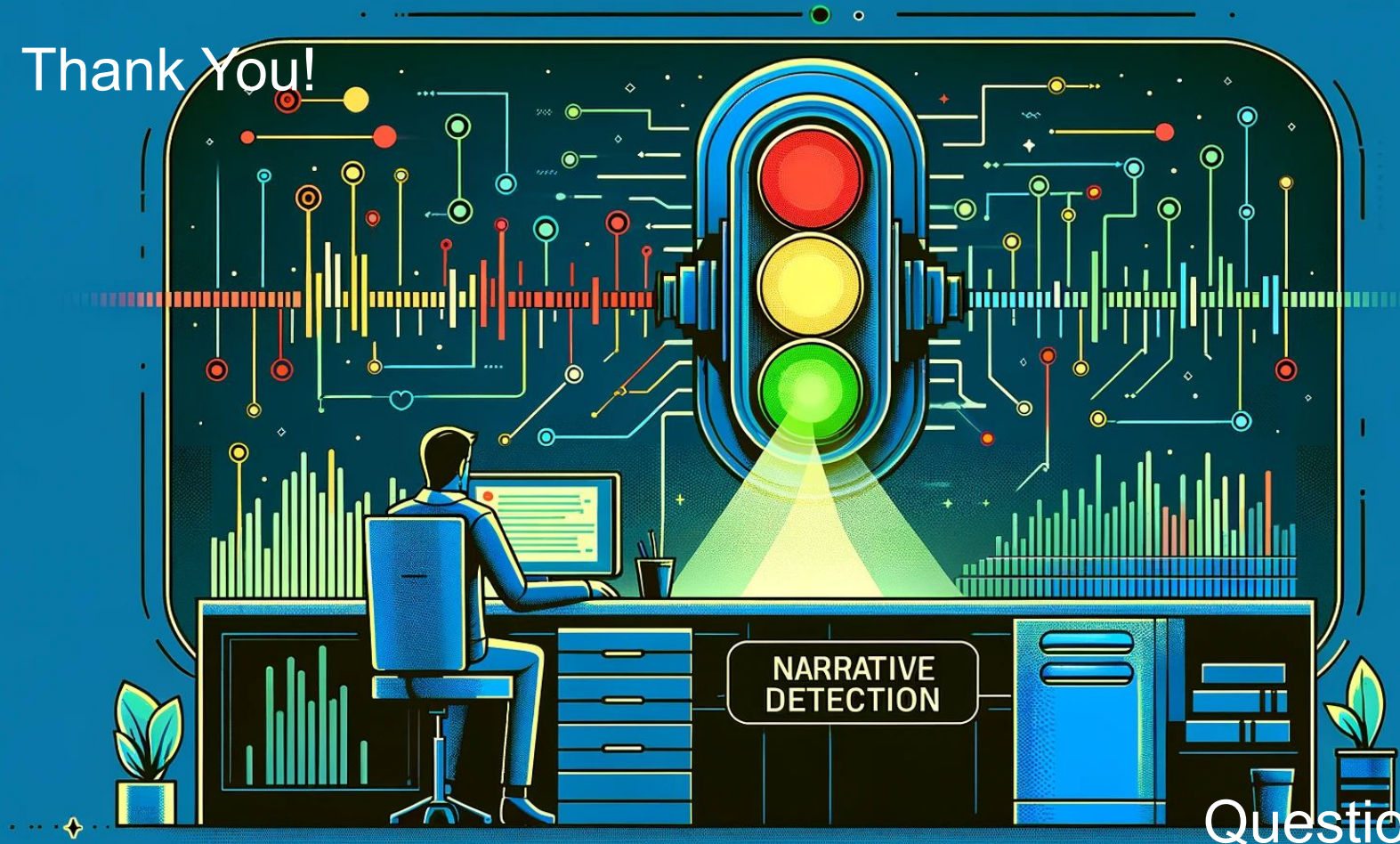
Limitations:

- Number of evaluated Case Studies
- Dependency on the sentence-transformers
- 1 document is assigned only to one topic.
- Only English language
- Outliers
- Replicability:
  - BERTopic - if the model is saved it’s possible to replicate, otherwise it’s not
  - Label generation - GPT-4 label generation is not replicable

Future Research:

- Outliers reduction & embeddings improvement with fine-tuning
- Multilingual Narrative Detection
- Integrating new claims
- Open source Models for Automatic Data Labelling (Topic, Narrative)

Thank You!



Questions?