



Measuring the Semantic Structure and Evolution of Conspiracy Theories

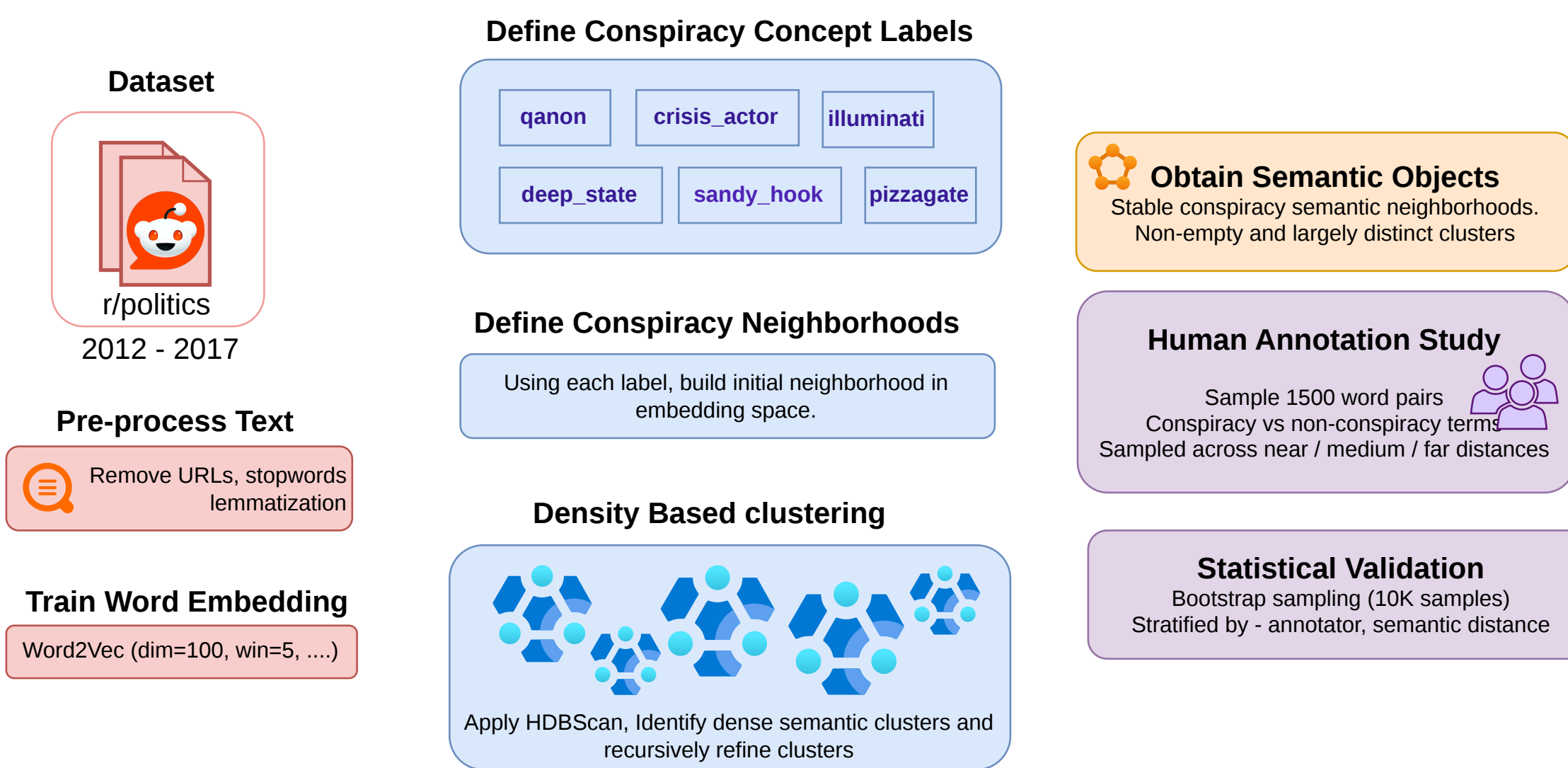
Manisha Keim, Sarmad Chandio, Osama Khalid, Rishab Nithyanand
Department of Computer Science, University of Iowa



RQ1: Is Conspiratorial Language semantically distinguishable?

Whether conspiracy-related discourse occupies a coherent and distinguishable region that is meaningfully separable from non-conspiratorial discourse?

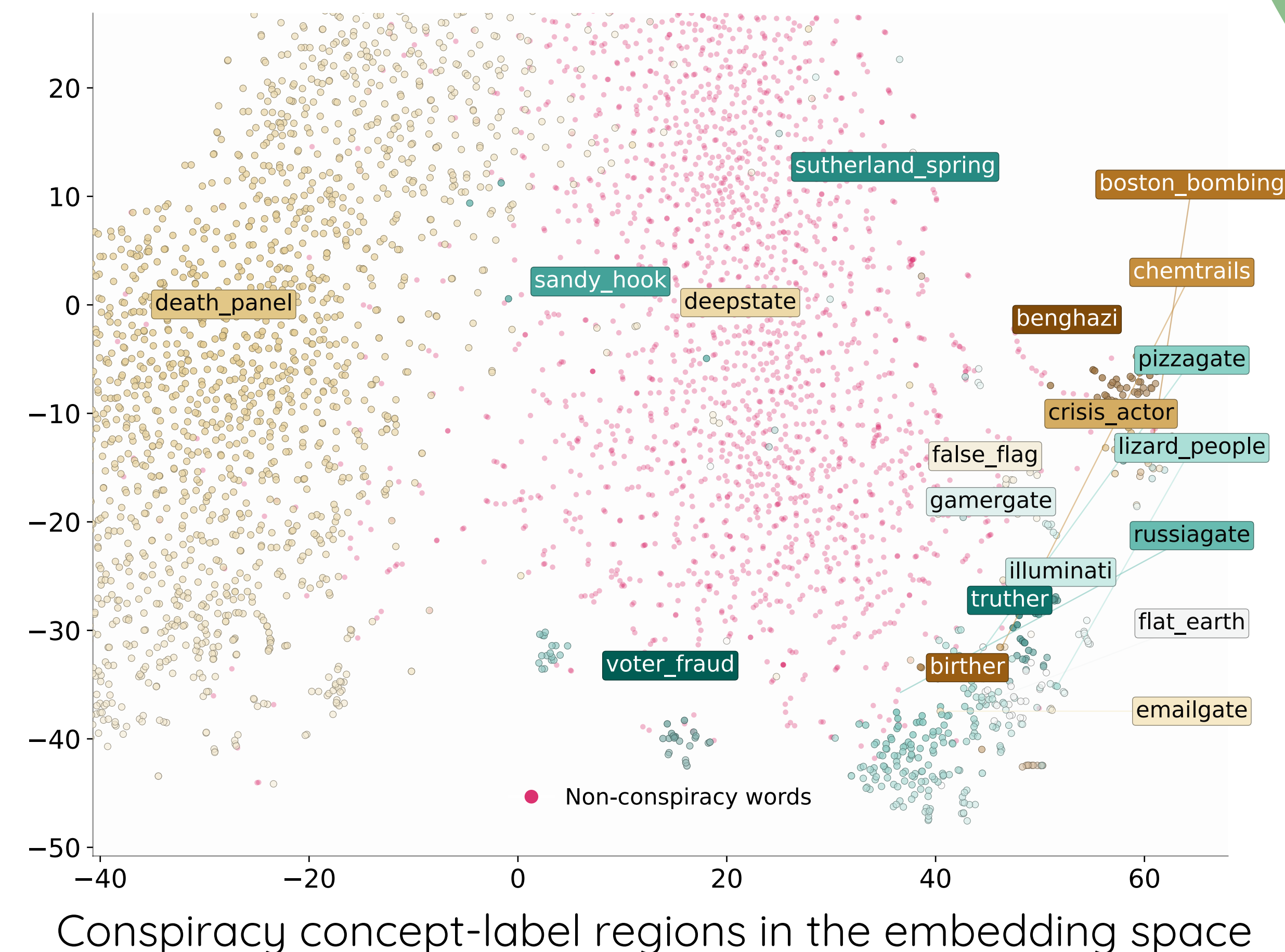
Methodology



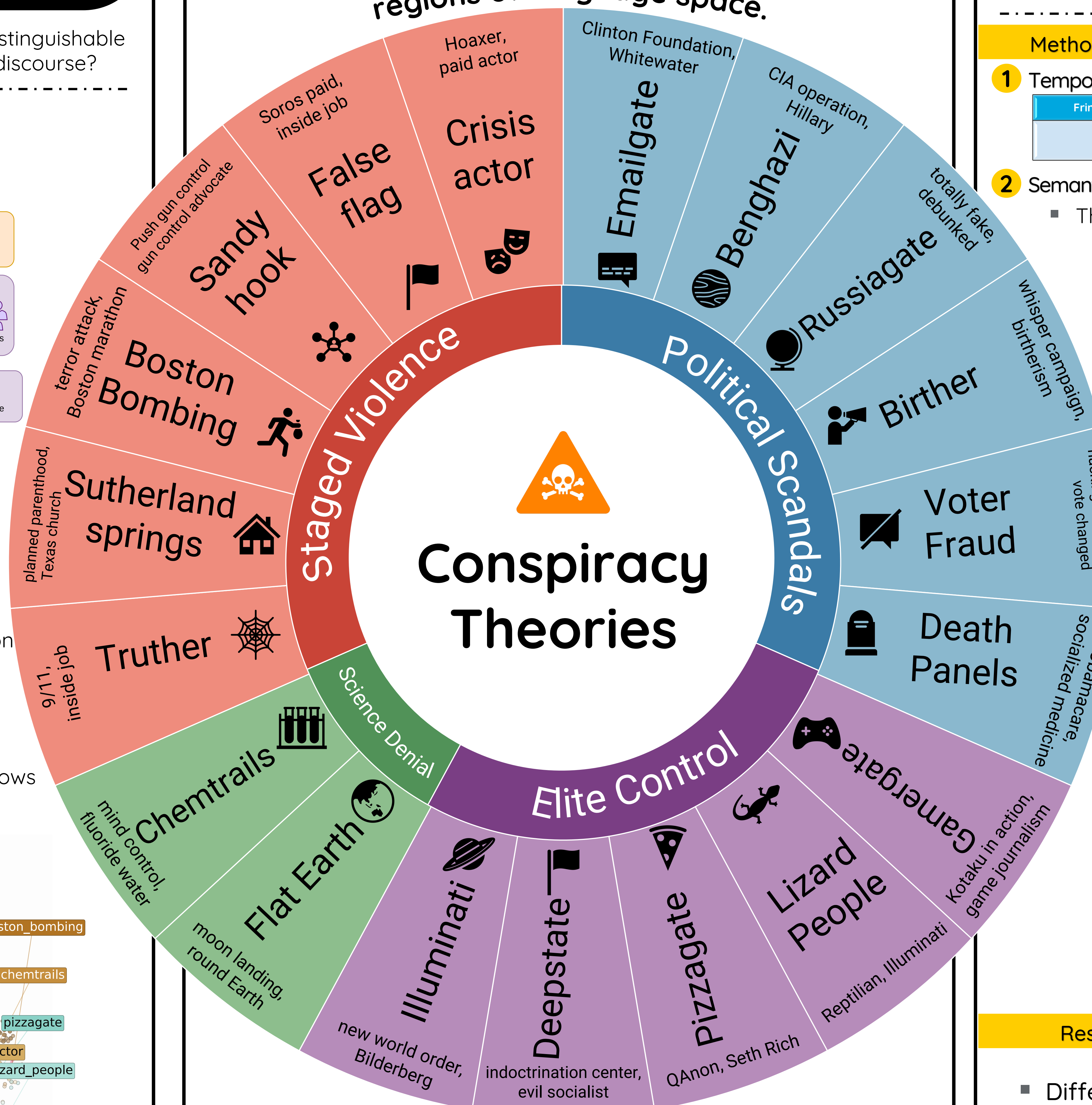
- Dataset: 169.9M comments from Reddit's r/politics subreddit spanning 2012-2022.
- We construct semantic objects anchored on 19 distinct conspiracy theories using HDBSCAN.
- We conducted and validated the words using human annotation study and performed statistical validation.

Results

- The clustering process produced 19 coherent clusters, which shows that conspiracy discourse is not randomly scattered.



Conspiracy-related language forms coherent and semantically distinguishable regions of language space.



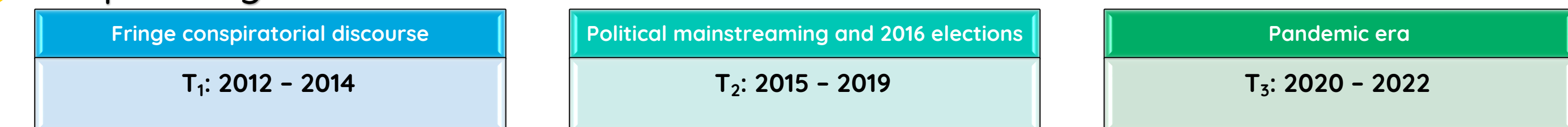
Conspiracy narratives are linguistically fluid. Understanding them requires monitoring their meaning space and not just their vocabulary.

RQ2: How do Conspiracies evolve?

How do the semantic objects associated with conspiracy theories evolve over time semantically and lexically?

Methodology

1 Temporal Alignment



2 Semantic Metrics

- The paper measures semantic evolution using three metrics:

Historical Association Divergence (HAD):	Counterfactual Semantic Divergence (CSD):	Neighborhood Coherence Divergence (NCD):
<ul style="list-style-type: none"> It measures that the terms that used to define a specific conspiracy theory still related to it with the same strength. Computed using: Cosine similarities, JS divergence High HAD: historical associations weakened 	<ul style="list-style-type: none"> It measures how semantically different the current meaning of conspiracy theory is from historic meaning of the conspiracy. Computed using average pairwise distance (APD). High CSD: Meaning shifted 	<ul style="list-style-type: none"> It captures whether the conspiracy theory has gained multiple semantically distinct meanings over time. Computed using kurtosis. High NCD: fragmentation Negative NCD: De-fragmentation

3 Lexical Evolution

Lexical Overlap (LO): change in vocabulary

Jaccard Similarity between neighborhood vocabularies.
Low Similarity: High turnover in vocabulary | High Similarity: Lexical stability

4 Classification of Evolutionary Patterns

Characterizing evolutionary patterns				
Pattern	HAD	CSD	NCD	LO
Semantic Stability	Low	Low	Low	High
Semantic Narrowing	High	Low	Low	*
Semantic Replacement	High	High	*	*
Semantic De(fragmentation)	*	*	High	High
Lexical Replacement	Low	Low	Low	Low

* indicates that the metric may take any value.

Results

- Different categories of conspiracies evolved differently.
- Political-scandal conspiracies underwent semantic replacement.
- Elite control conspiracies developed multiple diverging narratives.
- Staged violence conspiracies exhibit high lexical churn.
- Science denialism is anchored by a stable epistemic stance.
- The conspiracy landscape experience relative evolution.