

Crisis Observatory: Extracting Credible Signals During a Crisis in the Age of LLMs

Kuan-Chieh Lo*, Pranav Maneriker*, Sriram Sai Ganesh*, Dominik Winecki*, Kelly Garrett*, Ayaz Hyder*, Arnab Nandi*, Valerie Shalin+, Shannon Bowen+, Amit Sheth+, Srinivasan Parthasarathy*

Correspondence: lo.311@osu.edu, srini@ce.ohio-state.edu

Overview

We built **Crisis Observatory**, an LLM-powered system designed to help emergency responders and decision-makers extract credible, actionable information from social media during crises. Crisis Observatory provides interactive geospatial visualizations to explore crisis information across temporal, spatial, and topical dimensions for rapidly changing emergency situations.

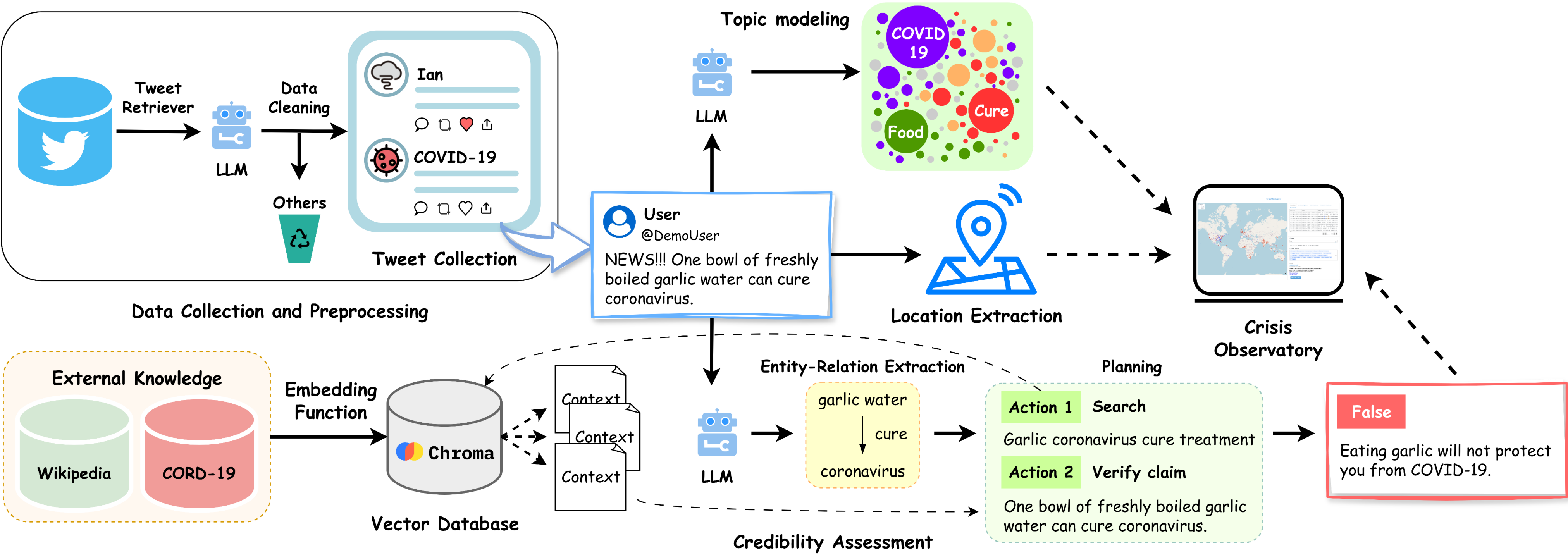
Motivation

- During crises, social media provides critical citizen-sensed insights, but distinguishing genuine needs from misinformation is challenging.
- Existing crisis response tools focus on data collection and need identification but lack credibility assessment and thematic organization necessary for effective sensemaking.
- While LLMs offer potential for crisis response, they face critical limitations: hallucination and knowledge cutoff.
- Crisis Observatory bridges these gaps by augmenting LLMs with external reliable sources and structured verification processes.

Key Contributions

- Address the challenge of extracting dependable, actionable information from social media during emergencies.
- Provide a comprehensive solution integrating credibility assessment and topic detection using LLMs.
- Enable real-time exploration of crisis information with geographic, temporal, and topical filtering using a web-based interactive dashboard.

System Overview



Topic modeling:

- BERTopic^[1] generates seed topics
- TopicGPT^[2] refines and assigns topics

Geospatial Analysis:

- Extract locations from tweet text, author profile, and geotags
- Build bounding box for location resolution

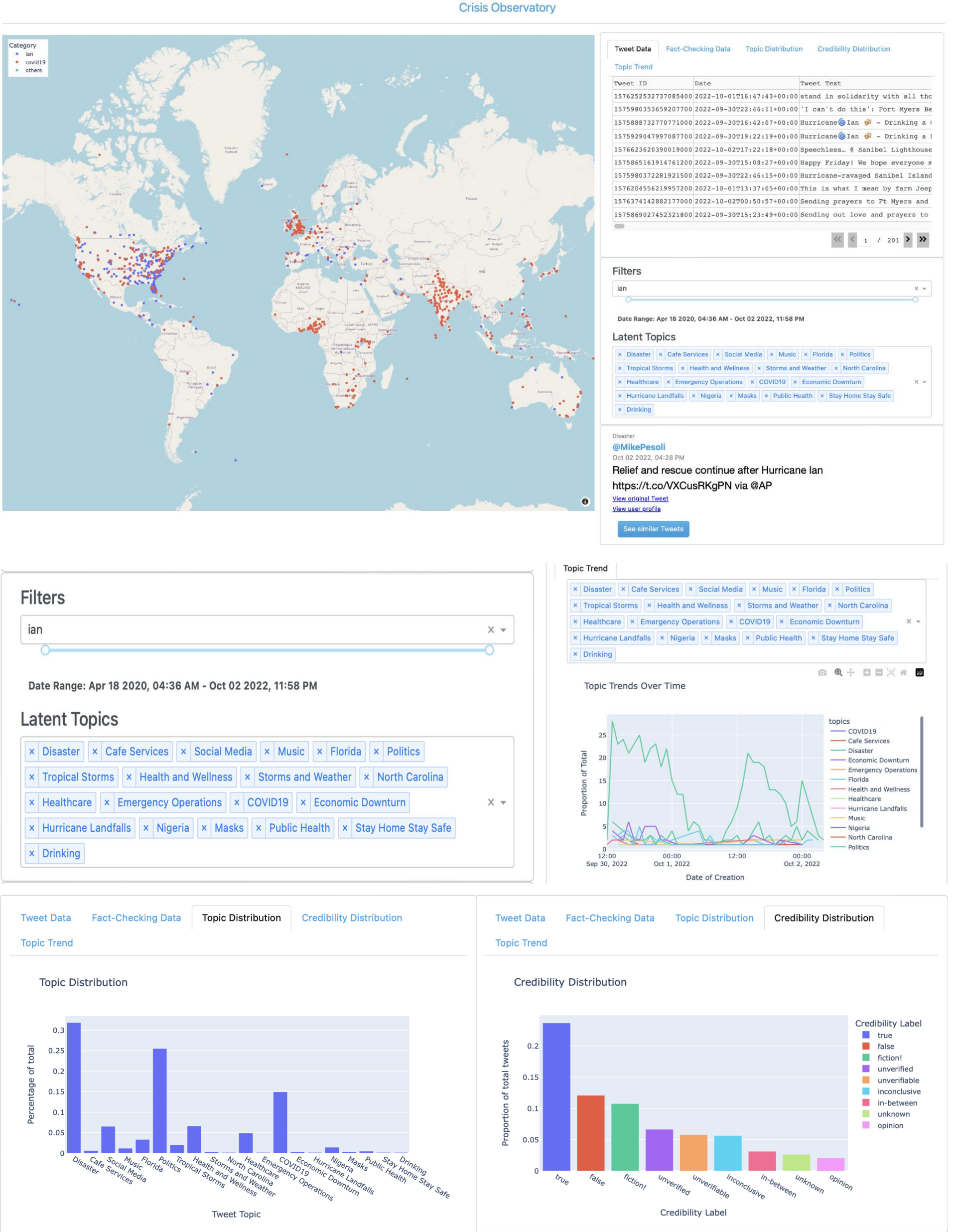
Credibility Assessment:

- Entity-Relation Extraction: Parse tweet contents to extract entities and relations.
- Verification planning: Identify verifiable claims and generate verification actions
- Evidence Retrieval and Assessment: Query vector database and synthesize evidence using RAG

Selected Examples

Topic	Tweet	Themes	Credibility	Explanation
Ian	It's been a crazy week here in Florida.I got really lucky and feel really blessed to have escaped any kind of storm damage.Others have not been as fortunate.	Disaster; Disaster Relief; Damage	True	The statement accurately reflects the situation in Florida during the recent storm.
Ian	Amazon Worker Delivers to 172 People During Hurricane Ian: 'I Hate Y'all'.	Disaster Relief; Family; Love	Unknown	Without additional context or information, it is not possible to determine the accuracy of this statement.
Covid19	Someone that would have been reported dead from UK if only Africa was hit with corona.	Conspiracy Theories	False	There is no information in the context provided that supports the claim that someone would have been reported dead from the UK if only Africa was hit with corona.
Covid19	1st Major plan after the MCO and this whole COVID thing should be "a peaceful weekend in Penang." Or maybe a bit further, Singapore sound great too !!	Family; Leisure	Opinion	The statement is an opinion and does not contain any factual claims to be evaluated for accuracy.

Interactive Dashboard



Acknowledgement

This work was supported by National Science Foundation grant ITE-2137806 and compute resources from the Ohio Supercomputing Center.

References

- [1] M. Grootendorst, "Bertopic: Neural topic modeling with a class-based tf-idf procedure," arXiv preprint arXiv:2203.05794, 2022.
- [2] C. M. Pham, A. Hoyle, S. Sun, and M. Iyyer, "TopicGPT: A prompt-based topic modeling framework," 2023.