

Walkthrough - Contextual News Risk Signal Analysis

This walkthrough explains the pipeline created to analyze news articles for risk signals using NLP and graph techniques.

Step 1: Data Collection

- News articles are collected from 4 RSS sources (BBC, Reuters, Al Jazeera, CNA).
- Titles, summaries, links, and publication dates are saved into a CSV file.

Step 2: NLP Processing

- Using spaCy and TextBlob, we extract entities (ORG, PERSON, GPE) and sentiment scores from summaries.
- These are stored in JSON format for graph construction.

Step 3: Graph Construction

- Nodes represent article titles and entities.
- Edges represent relationships such as 'mentions'.
- Graph is built using NetworkX.

Step 4: Graph Analysis

- Degree centrality is calculated to find influential entities.
- A subgraph of top connected nodes is visualized.

Below is a visualization of a sample graph.

