

# news\_signals:

## An NLP Library for Text and Time Series

- NLP + time series is rare!
- **news\_signals**: Python library for creating datasets with **text inputs** and **time series outputs**
- *Signals* are centered around real-world entities
- Collecting text and time-series from 3<sup>rd</sup> party sources about an entity
  - News articles
  - Time series of news volume
  - Time series of Wikipedia article page views
  - ...
- Exploration tools: pandas-like interface, plots
- Dataset enrichment tools: anomaly detection, summarization
- Large-scale dataset generation; Docker container, K8 config

