**Extraction Pipeline**

Our pipeline extracts events from news articles using NLP technology such as Stanford CoreNLP and TextRazor’s API.

1. We split the article into sentences and words.
2. We identify verbs - these are the Events.
3. Any named entities become Actors, Places, or Times depending on their type.
4. We parse grammatical dependencies to relate Events to Actors, Places, and Times.
5. Together, the verb and its dependent named entities form an event as shown in the graph above. We use the Simple Event Model (SEM) as a schema.

The process is repeated for every article and sentence in each article. The resulting events are exported as RDF or JSON for further usage in visual analytics applications.

**Visual Analytics**

Geographical overviews help answer questions like: What is the intensity of activist activity in certain areas? What is the chronological spread from one area to another?

Event type clouds help answer questions like: What is the current state of a conflict that activists are involved in? Are there primarily things being said, or are more concrete actions undertaken?

---

**Extractivism: Extracting & Visualising Activist Events**

**Goal**

Support organisational scientists studying activist organisations through extracting, aggregating, and visualising the events they participate in.

**Why**

Activists could have a significant role in shaping social views and opinions, e.g. Occupy, Arab Spring.

**Problem**

Incomplete, incorrect, uncontextualized, and biased information.

**Solution**

Extract event mentions from different news sources that could complement, contradict, or verify each other. Visualize the aggregated results to support analysis and discovery.

**Who**

Computer Science: Thomas Ploeger, Chun Fei Lung, Lora Aroyo

Organization Sciences: Maxine Kruijt, Frank de Bakker, Iina Hellsten

www.mona-project.org