

Time-Aware Mining and Visualization of International Alliances

Xavier Tannier

LIMSI, CNRS, Univ. Paris-Sud, Université Paris-Saclay, Orsay, France (xavier.tannier@limsi.fr)

Objective: Relation Extraction and Aggregation

1. Traditional relation extraction at sentence level

Japan and China agree to reduce tensions over Senkaku islands.

→ POS(*Japan, China*)

Obama, Merkel warn Russia against intervention.

→ NEG(*U.S.A., Russia*), → NEG(*Germany, Russia*),

→ POS(*U.S.A., Germany*)

Serbia prepares hero's welcome for Putin.

→ POS(*Serbia, Russia*)

Russia backs Ukraine rebel vote.

→ NEG(*Russia, Ukraine*)

British teens detained en route to Syria, police say.

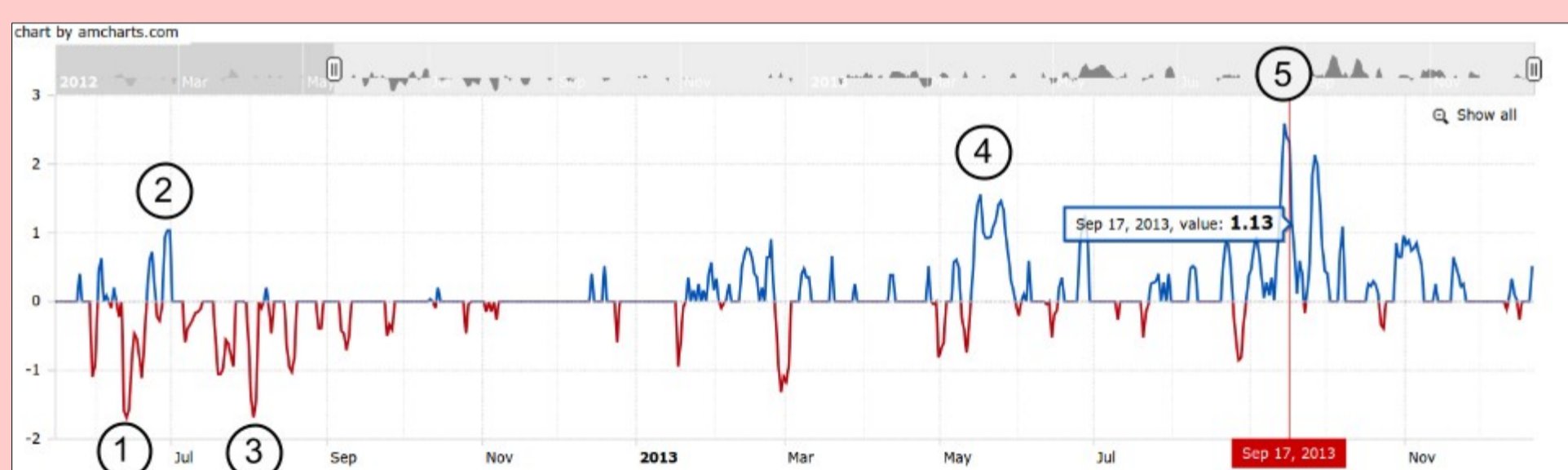
→ NEU(*Great Britain, Syria*)

2. Time-aware Aggregation

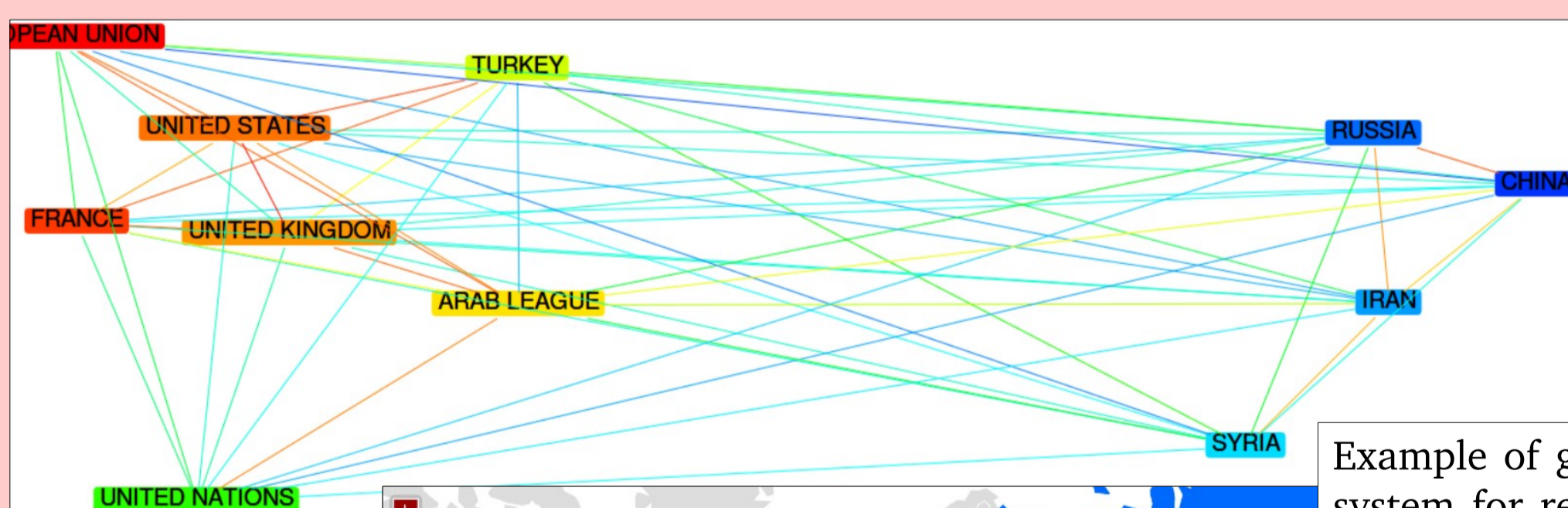


* TTIP = Transatlantic Trade and Investment Partnership

3. Visualization

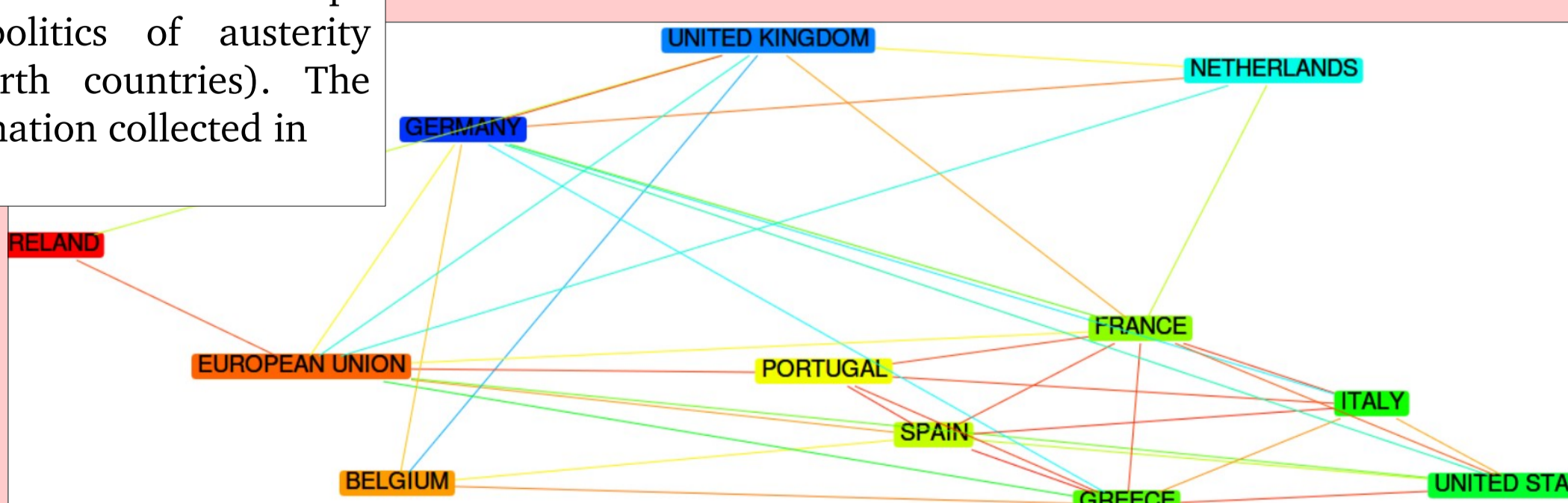


Example of plot produced by the system for bilateral relations between United States and Russia on the query "Syria". The bottom left frame shows sentences corresponding to the user-selected date (Sep. 17, 2013). Circled numbers have been manually added to the screenshot. They correspond to: ① Mutual accusations of supplying arms to Syrian authorities or opposition (bad relation, $sw(d) < 0$); ② Planning of a meeting to discuss the problem (better relation, $sw(d) > 0$); ③ Vetoes of China and Russia for United Nations resolutions; ④ Announcement of a peace conference; ⑤ Agreement at this conference.



Example of graph and map produced by the system for relations between different states on the query "syria", for the year 2012. The graph is based on information collected in 18,582 sentences. Edge colors indicate the kind of relation (from dark red for strong alliance to dark blue for strong opposition), and vertice colors reflects proximity of countries with each other.

Example of graph produced by the system for relations between different states on the query "austerity", from 2012 to the end of 2013 (where countries from South of Europe fought against the politics of austerity advocated by the North countries). The graph is based on information collected in 553 sentences.



We would like to thank Agence France Presse (AFP) for providing the corpus and allowing the distribution of the training dataset.

Method

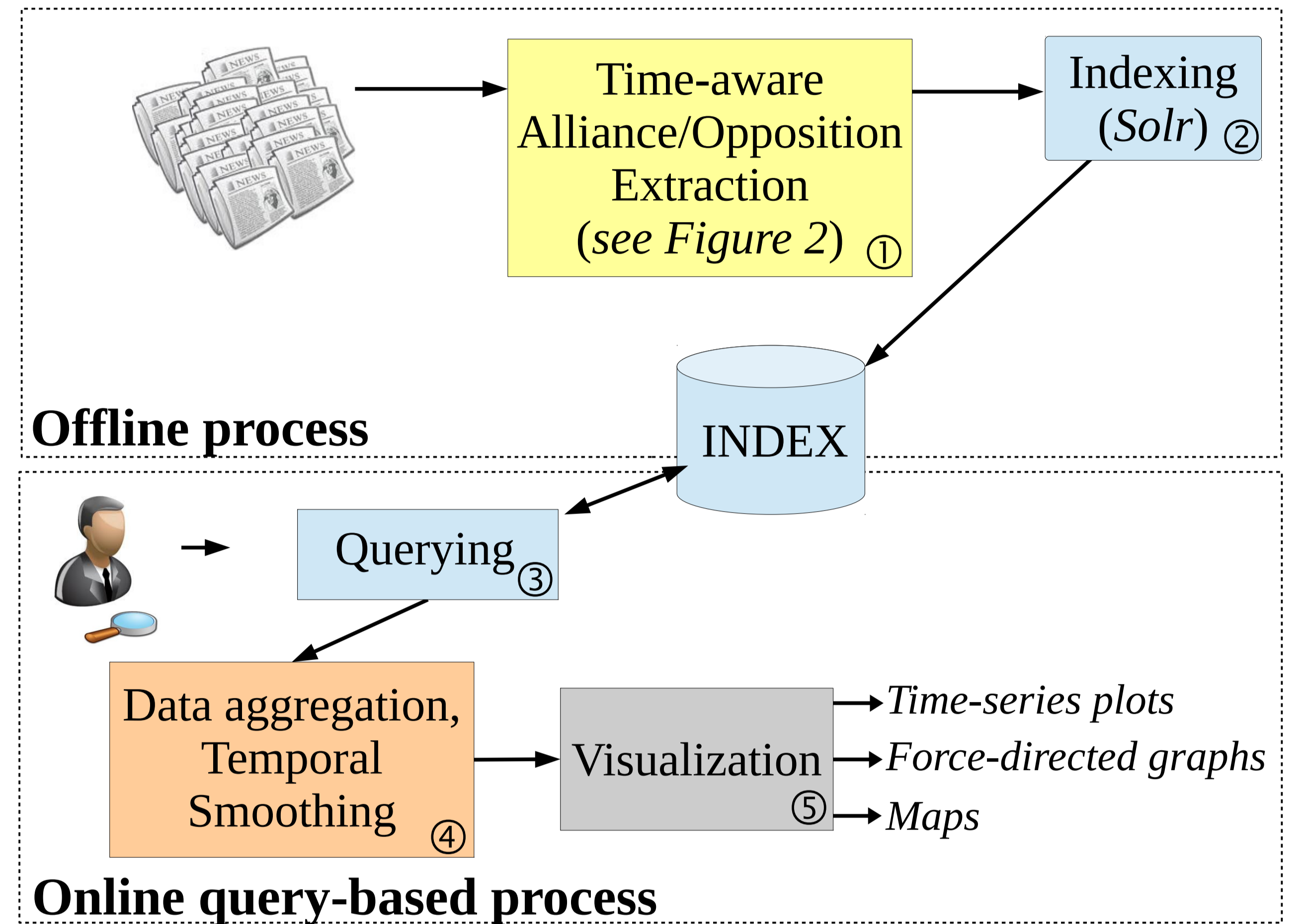


Figure 1: System overview.

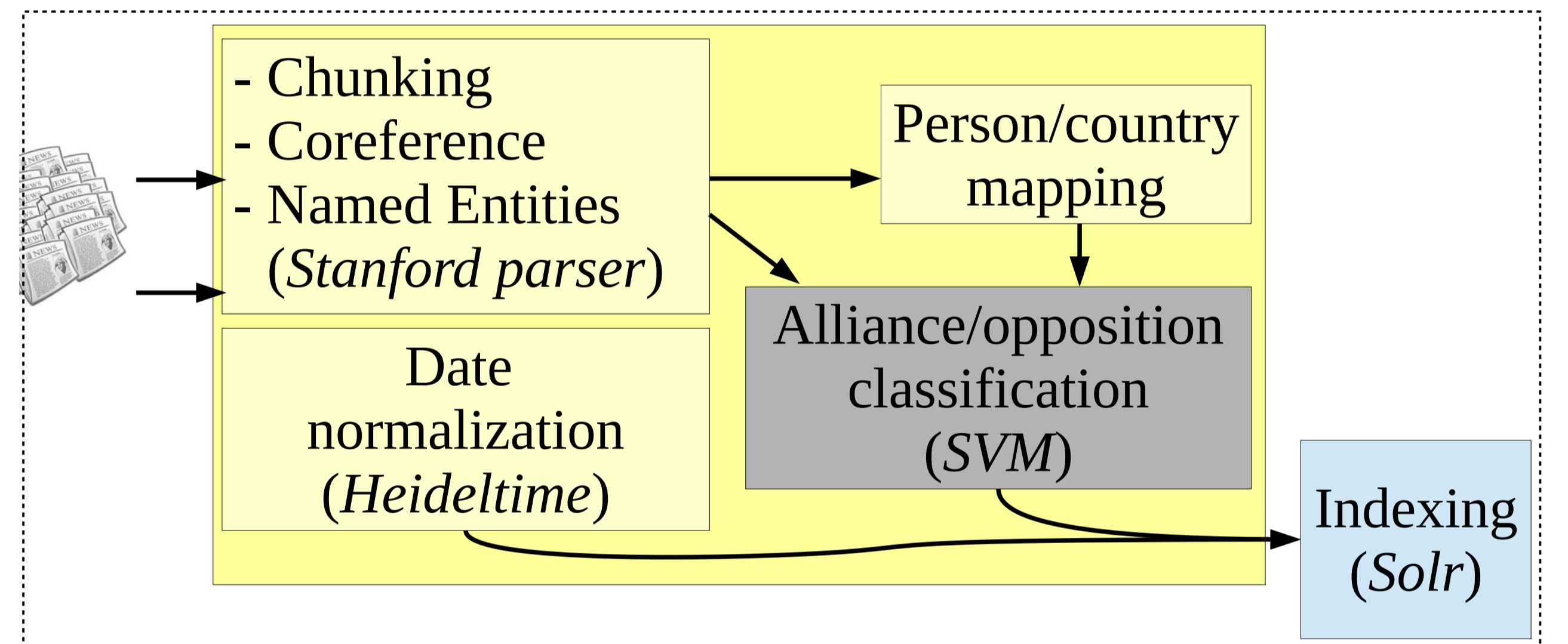
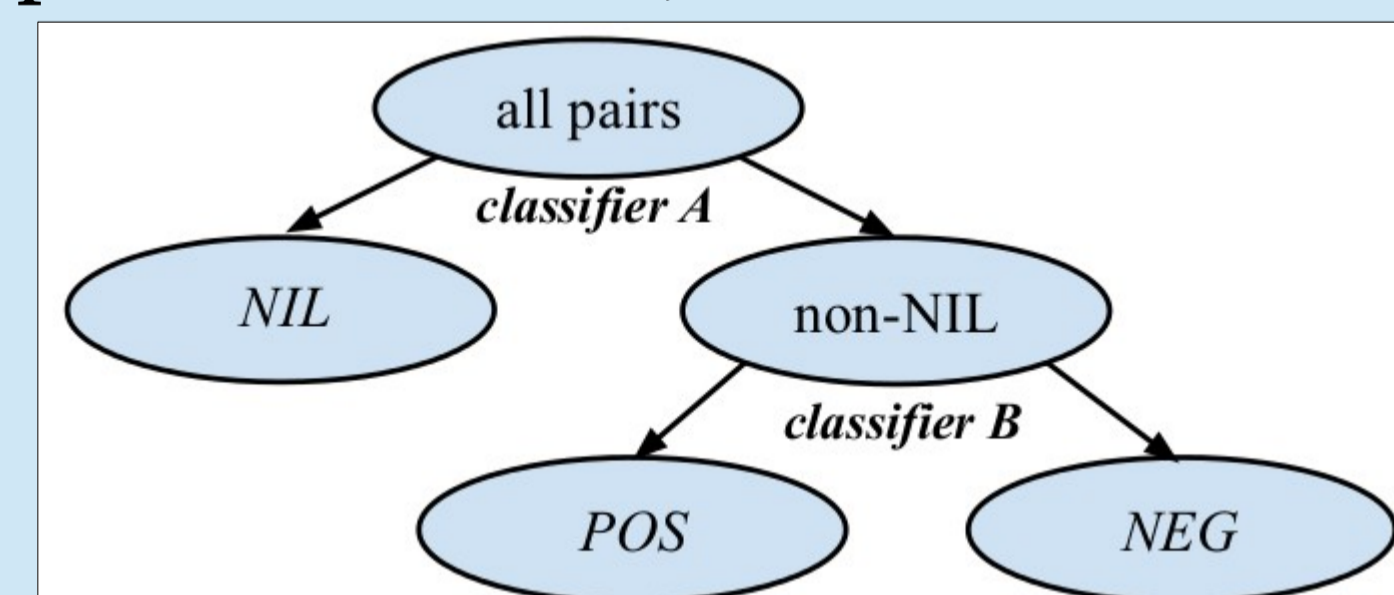


Figure 2: Offline processing details.

1. Relation classification at sentence level

Classification with SVM :

- Precision-oriented:
 - Recall will be increased at aggregation-level by the volume of data.
 - Consider only sentences containing verbs from a manually built list (*agree, oppose, ...*).
- 2105 annotated pairs (349 NEG, 293 POS)
- Two steps:



2. Time-Aware Aggregation

330,222 relations extracted from 1,79 million articles (AFP corpus).

- A date in a sentence time-stamps this sentence
- A sentiment weight is assigned to a triple (country1, country2, day d):

$$w(c_1, c_2, d) = \log \frac{1 + \text{POS}(c_1, c_2, d)}{1 + \text{NEG}(c_1, c_2, d)}$$

- Each weight is turned into a distance between 0 and 1 where taking strong stands is encouraged:

$$w'(c_1, c_2, d) = 1 - \frac{1}{1 + e^{-w(c_1, c_2, d)}}$$

