Integrating Interactive and Computational Approaches for News Understanding

Liang Wu, Fred Morstatter, J. Daniel Baird, Justin Sampson, Kai Shu, and Huan Liu

Contact: wuliang@asu.edu

Motivation

- GDELT provides "a global database of society", with daily updates of events around the world. These events are tagged with entities, and interactions between them.
- The GDELT project is a one-of-a-kind dataset in both content and scale.
- Due to its impressive size, it is impossible to read the individual records to understand trends.
- To address this challenge we propose a system to summarize news from GDELT.

Goals

- Design a system that helps researchers quickly and easily understand news trends from news data such as GDELT.

- Use the system to find underlying trends within the data.

- Enable researchers to compare topics from different countries.

System Diagram

Frontend
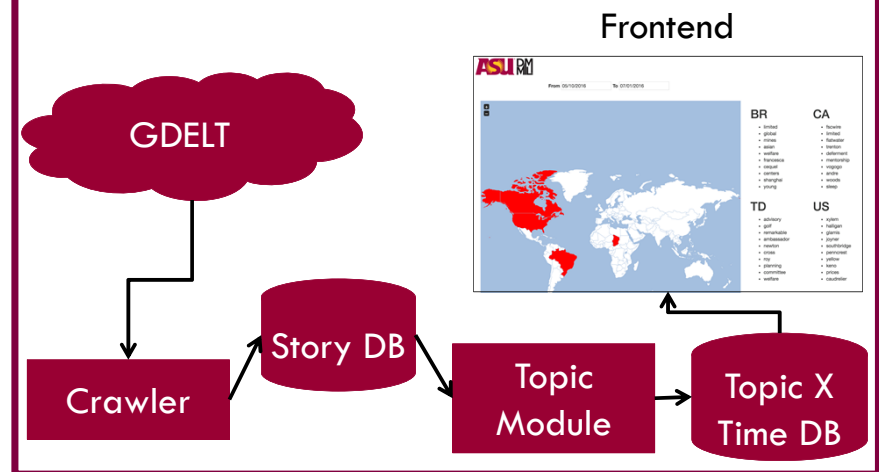
Date selection: Allows user to choose time range.

Country selection: Allows user to choose countries to compare.
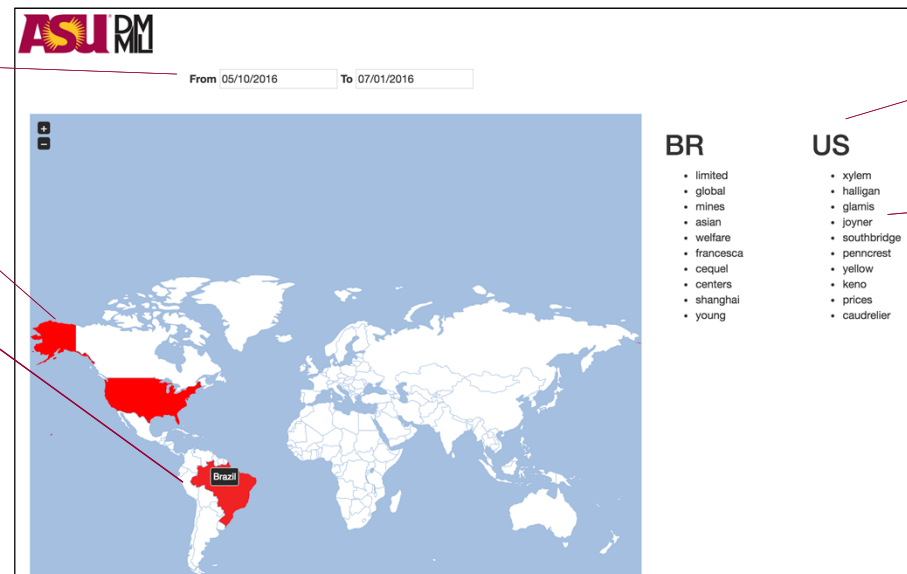
Country X Topic Window: Shows emergent topics from each country.

A topic consists of the top words emerging from a country.

References

[1] Leetaru, Kalev, and Philip A. Schrodt. "Gdelt: Global data on events, location, and tone, 1979–2012." ISA Annual Convention. Vol. 2. No. 4. 2013.

[2] Blei, David M., Andrew Y. Ng, and Michael I. Jordan. "Latent dirichlet allocation." Journal of machine Learning research 3.Jan (2003): 993-1022.