GDELT: Global Data on Events, Location and Tone

April 2013: Researchers have released the GDELT: Global Data on Events, Location and Tone dataset which promises to be an important addition to the social science research community.

In a blog post entitled, The Future of Political Science Just Showed Up, Jay Ulfelder, a Poli Sci researcher, writes, “I suspect this is going to be the data set that launches a thousand dissertations.”

The full dataset contains more than 200 million global events from 1979 to mid-2012 and is encoded in the CAMEO — Conflict and Mediation Event Observations — coding scheme.

For people interested in exploring this data now, U Penn has released a reduced set of the data. Whereas the full data set contains more than 60 attributes, the reduced set offers only 12 attributes. The full list of attributes can be found in the appendix to Philip Schrodt and Kalev Leetaru’s paper: GDELT: Global Data on Events, Location and Tone, 1979-2012.

The reduced data set and some sample Python and R code can be downloaded here http://eventdata.psu.edu/data.dir/GDELT.1979-2012.reduced.zip

---

Revision #1
Created 2013-04-22 23:17:11 UTC by Williamson, James
Updated 2013-04-22 23:17:11 UTC by Williamson, James