

## **The GDELT Project: An Open Platform For Computing on the Entire World**

Kalev Hannes Leetaru

*Yahoo! Fellow, Edmund A. Walsh School of Foreign Service, Georgetown University  
Council Member, Global Agenda Council on the Future of Government, World Economic Forum*

The GDELT Project (<http://gdeltproject.org>) monitors the world's broadcast, print, and web news from nearly every corner of every country in over 100 languages from 1979 to present day and identifies the people, locations, organizations, counts, themes, sources, and events driving our global society every second of every day, creating a free open platform for computing on the entire world. Two datasets are created, one codifying physical activities around the world in over 300 categories from riots and protests to peace appeals and diplomatic exchanges, and one recording the people, places, organizations, emotions, and themes underlying those events and their interconnections. The event database contains more than a quarter-billion global events, coded temporally to the day and geographically to the city level, with around 150,000 new events added every 24 hours. Each event is broken into 59 attributes codifying everything from the ethnic group of the perpetrator to the geographic origin of the victim. The network database offers a graph structure over the world's news media each day, connecting the actors, contexts, and worldviews behind those global events. Today the GDELT Project is one of the largest open spatio-temporal graph datasets in existence and pushes the boundary of current analytic techniques in scale, multidimensionality, and complexity. Fundamentally new algorithms and methodologies are required to fully unlock the insights captured in GDELT's archives, yet already in just its first year the database has become the gold standard for global societal-scale study and is widely deployed within the US Government and NGOs across the world, and is demonstrating substantial promise both for operational watchboarding and in dramatically increasing the accuracy of conflict forecast models. This presentation will survey the current state of the GDELT Project and present a cross-section of both current applications and new directions for the future.