

Cultural Knowledge Encoding

ERDC

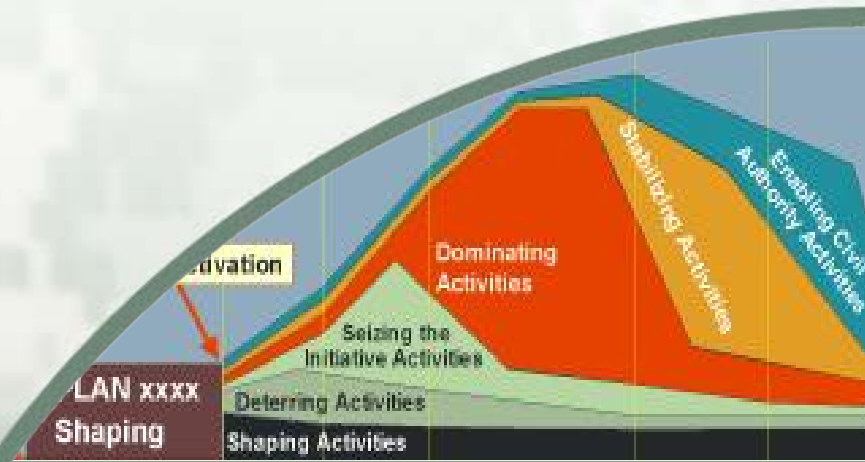
Engineer Research and Development Center

**Kalev Leetaru, Past Adjunct Faculty,
Georgetown University**

kalev.Leetaru5@gmail.com

**Chris Rewerts, PhD, Research Engineer,
U.S. Army ERDC**

Chris.Rewerts@erdc.dren.mil



**US Army Corps
of Engineers®**

Sociocultural Capabilities, Methods, Interests

Addressing Socio-Cultural Challenges to Military Operations

“The force lacks sufficient capability to integrate knowledge of the operational environment, such as terrain, infrastructure, and society...contributing to a less holistic assessment of the operational environment.”

*TRADOC Pam 525-8-5 (February 2014),
U.S. Army Functional Concept for Engagement*

- **Population characterization**
- **Urban change**
- **Infrastructure & people**
- **Sociocultural systems**
- **Exploiting & sharing knowledge**



BUILDING STRONG®

ERDC

Innovative solutions for a safer, better world

The GDELT Project

GLOBAL DATABASE OF EVENTS, LANGUAGE, AND TONE



A Global Societal-Scale Real-time Dashboard

<http://gdeltproject.org>



GDELT

[GDELT Main Website](#)

[GDELT Blog](#)

[About GDELT](#)

[Contact](#)

About GDELT

GDELT is an initiative to construct a catalog of human societal-scale behavior and beliefs across all countries of the world, connecting every person, organization, location, count, theme, news source, and event across the planet into a single massive network that captures what's happening around the world, what its context is and who's involved, and how the world is feeling about it, every single day.

Download the Data

[Daily Event Records](#)

[Historical Event Records](#)

[Reduced Event Dataset](#)

[Global Knowledge Graph](#)

Reports/Analyses

[Daily Conflict Trends Report](#)

[Daily World Leaders Index](#)

The GDELT Project

GLOBAL DATABASE OF EVENTS, LANGUAGE, AND TONE

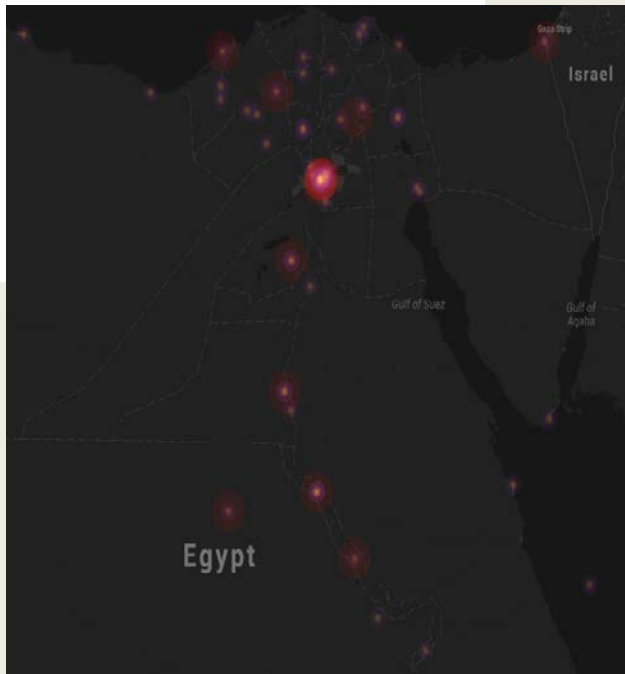


Global Event Database

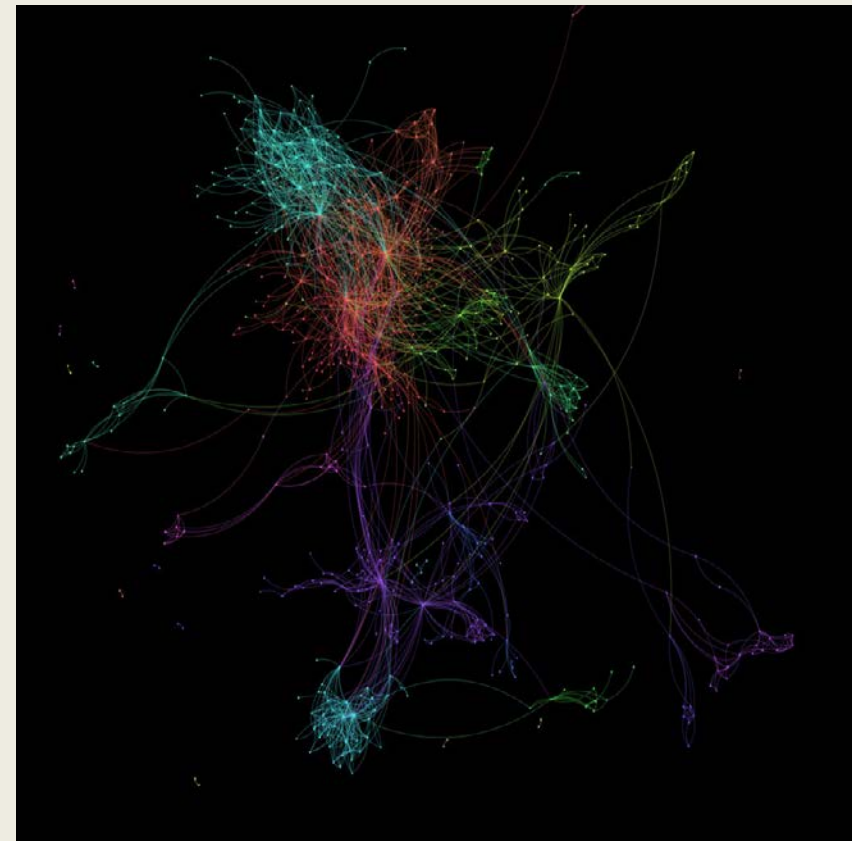
BAGHDAD. Iraqi leaders criticized Turkey on Monday for bombing Kurdish militants in northern Iraq with airstrikes that they said had left at least one woman dead.

Event Code: 111
Source: IRQ GOV
Target: TUR

Event Code: 223
Source: TUR
Target: IRQKRD REB

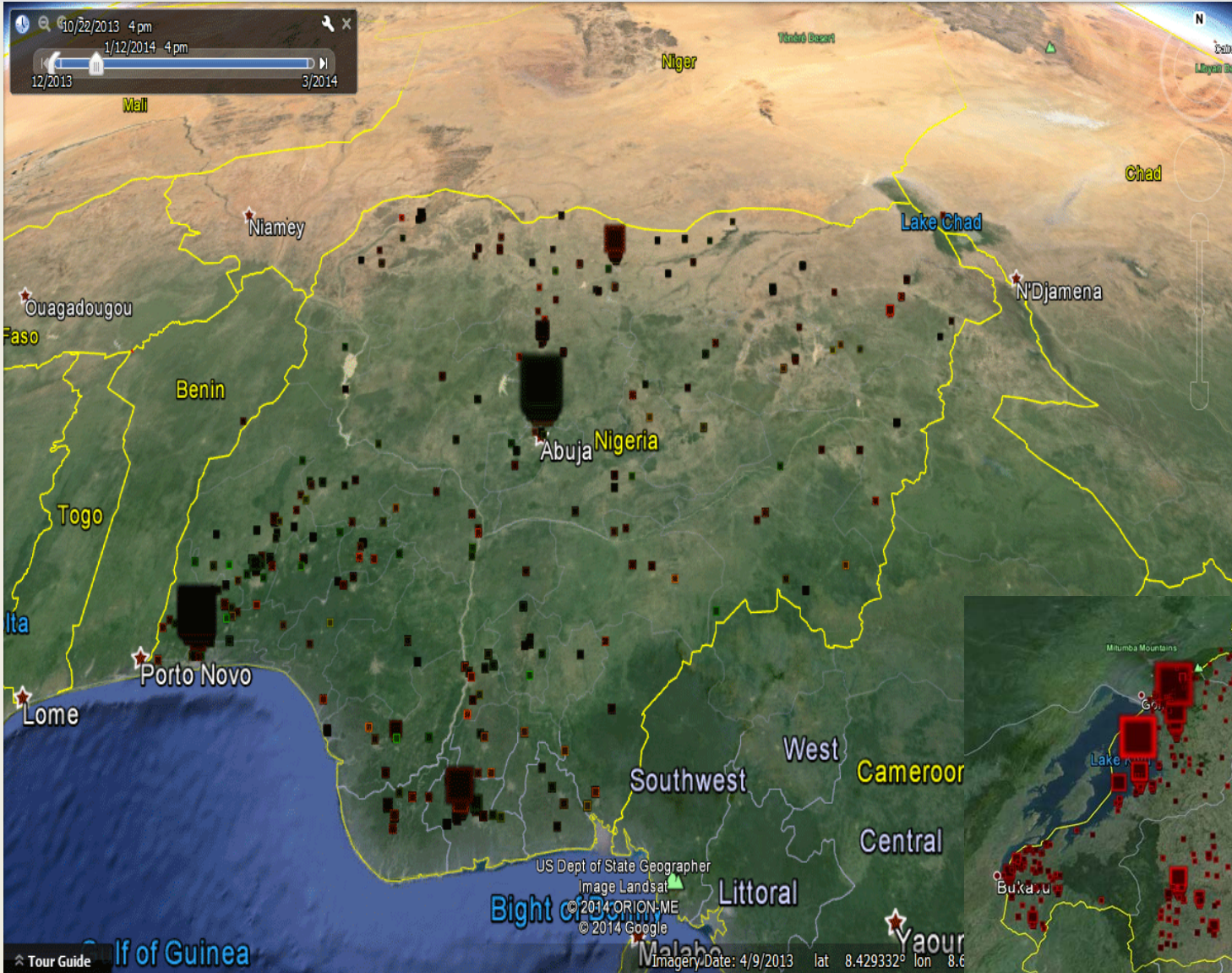


Global Knowledge Graph



The GDELT Project

GLOBAL DATABASE OF EVENTS, LANGUAGE, AND TONE



THE GLOBAL CONVERSATION

WHAT HAPPENS WHEN YOU TAKE SIX MONTHS OF news coverage from around the world and compile a list of every person mentioned and the people they were mentioned alongside? You get a network of 3 million nodes connected by 42 million links. Based on the new Global Knowledge Graph—a massive compilation of the world's people, organizations, locations, themes, emotions, and events created by Kaveh Lestani, one of FP's Global Thinkers, and his colleagues Philip Schmidt and Patrick Brandt—this visualization is a snapshot of the 25,000 newsmakers mentioned most frequently from April to October 2013 and the 100,000 connections among them.

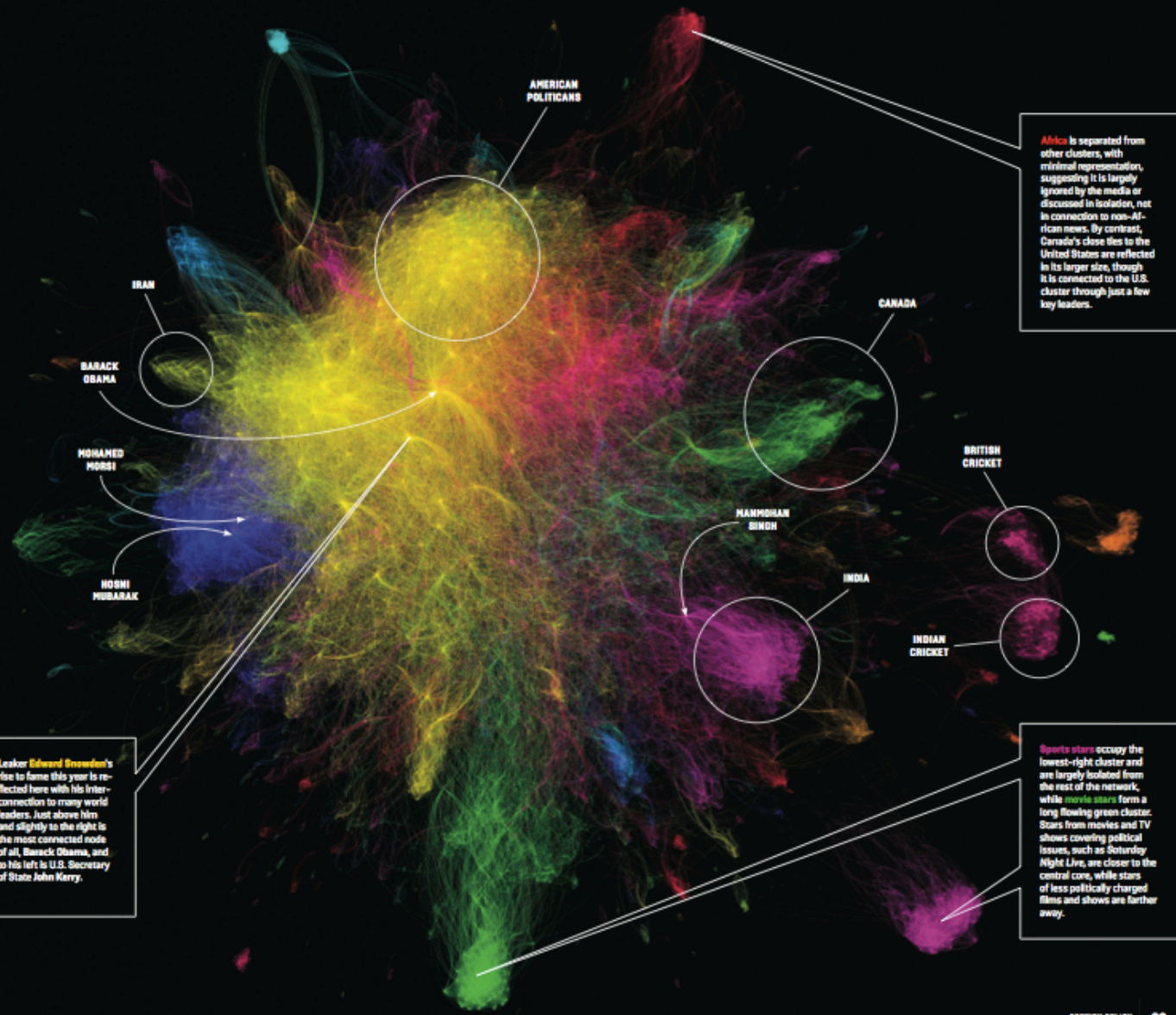
The yellow cloud portrays coverage involving the United States, with President Barack Obama near its center. The left portion encompasses world leaders most heavily interacting with Washington, from Russian President Vladimir Putin to Syrian President Bashar al-Assad, while the top portion comprises major American politicians, like House Speaker John Boehner and Senate Majority Leader Harry Reid.

Geographical regions make up many of the larger colored clusters: Egypt (left—most blue), Africa (top-right red), Canada (right—most green), India (pink). The relative centrality of the regions, their size, and the strength of their ties to the United States and each other give a sense of how prominent each has been in news coverage and, by inference, the degree of importance world leaders give them. Egypt's cluster, for example, is quite large, reflecting the detailed coverage of its struggles. By contrast, Iran's far smaller cluster is actually an extension of the U.S. grouping, reflecting the narrow focus on Tehran's nuclear ambitions, which are seen almost completely in the context of its relationship with Washington.

One of the most notable aspects of the visualization is the isolation of some subjects that receive a lot of news coverage but that have few ties to anything else happening. For example, the bottom-most green cluster focuses on movie stars, while the lower-right pink cluster focuses on sports. In a sign of just how influential the scandals over leaked intelligence documents have been this year, actor Benedict Cumberbatch, who plays hacker Julian Assange in *The Fifth Estate*, forms one of the primary connections between Hollywood and the rest of the network.

This visualization is, in essence, a snapshot of the global conversation—not only whom we are talking about and how much, but how each separate discussion is connected to every other discussion and the greater whole. It's a new way of gauging what matters to us, and it's just one of the ways in which big data is changing the way we see the world.

Leaker **Edward Snowden's** rise to fame this year is reflected here with his inter-connection to many world leaders. Just above him and slightly to the right is the most connected node of all, Barack Obama, and to his left is U.S. Secretary of State John Kerry.



Africa is separated from other clusters, with minimal representation, suggesting it is largely ignored by the media or discussed in isolation, not in connection to non-African news. By contrast, Canada's close ties to the United States are reflected in its larger size, though it is connected to the U.S. cluster through just a few key leaders.

Sports stars occupy the lowest-right cluster and are largely isolated from the rest of the network, while movie stars form a long flowing green cluster. Stars from movies and TV shows covering political issues, such as *Saturday Night Live*, are closer to the central core, while stars of less politically charged films and shows are farther away.

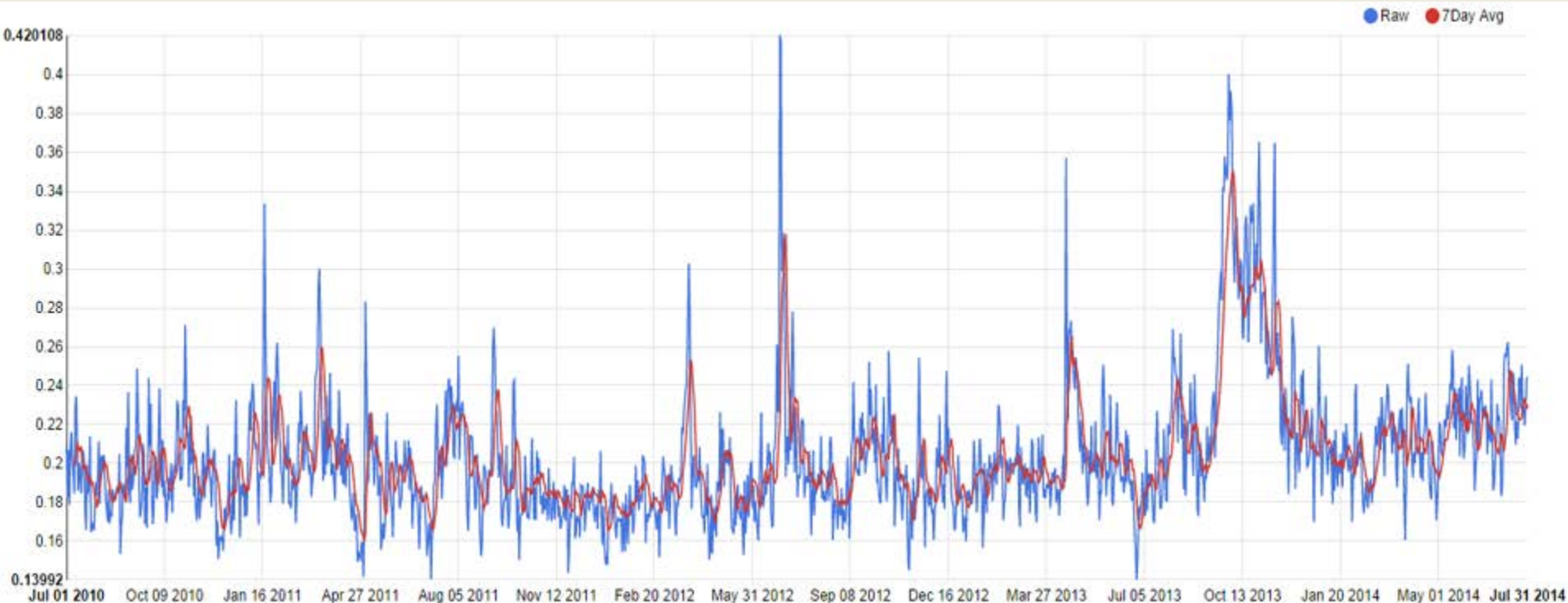
The GDELT Project

GLOBAL DATABASE OF EVENTS, LANGUAGE, AND TONE



Global Emotions & Themes

2,238 emotions and themes from every article



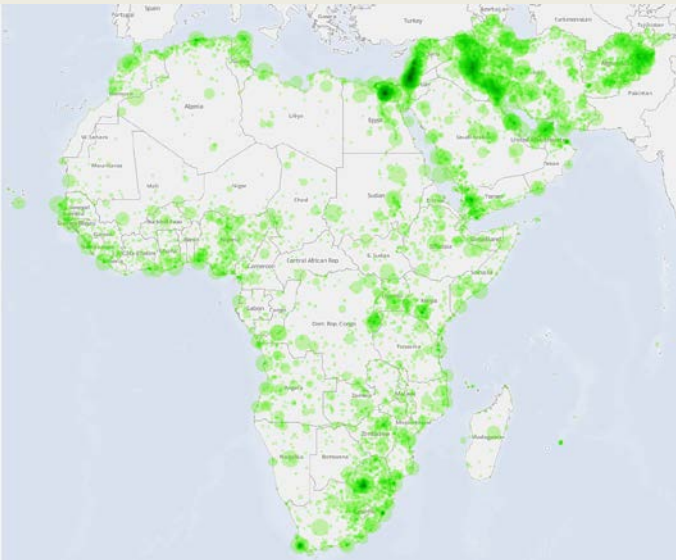
The GDELT Project

GLOBAL DATABASE OF EVENTS, LANGUAGE, AND TONE

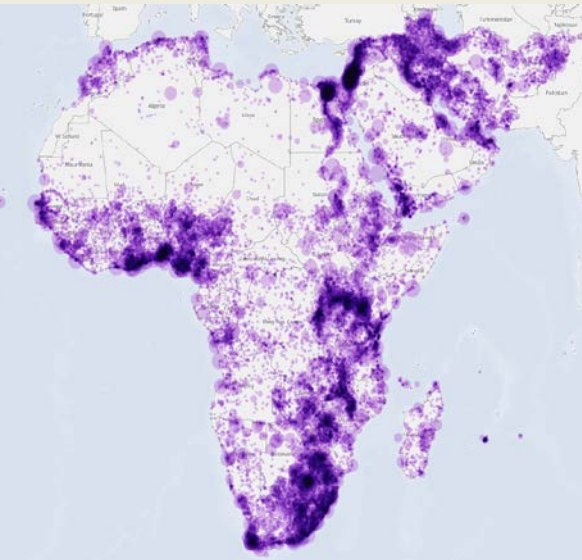


Academic Literature (US Army ERDC)

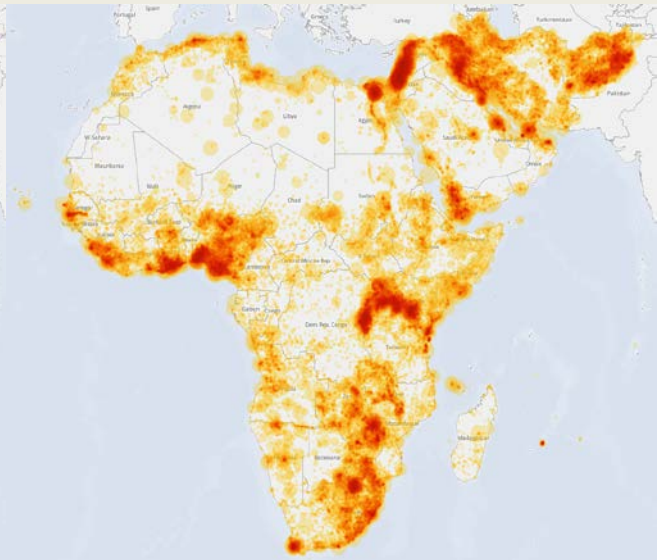
- 21 billion words of academic literature, unclassified/declassified reports, and dissertations on Africa and Middle East (JSTOR, DTIC, CIA, CORE, CiteSeerX, Internet Archive) 1945-present, Global Knowledge Graph representation available Summer 2014



USGOV



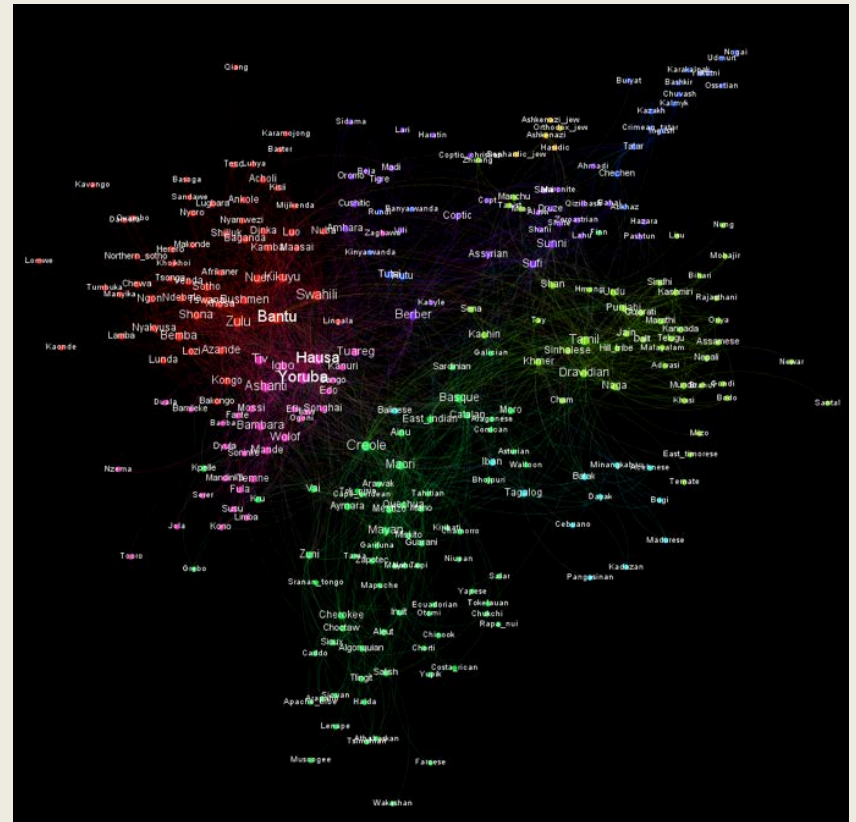
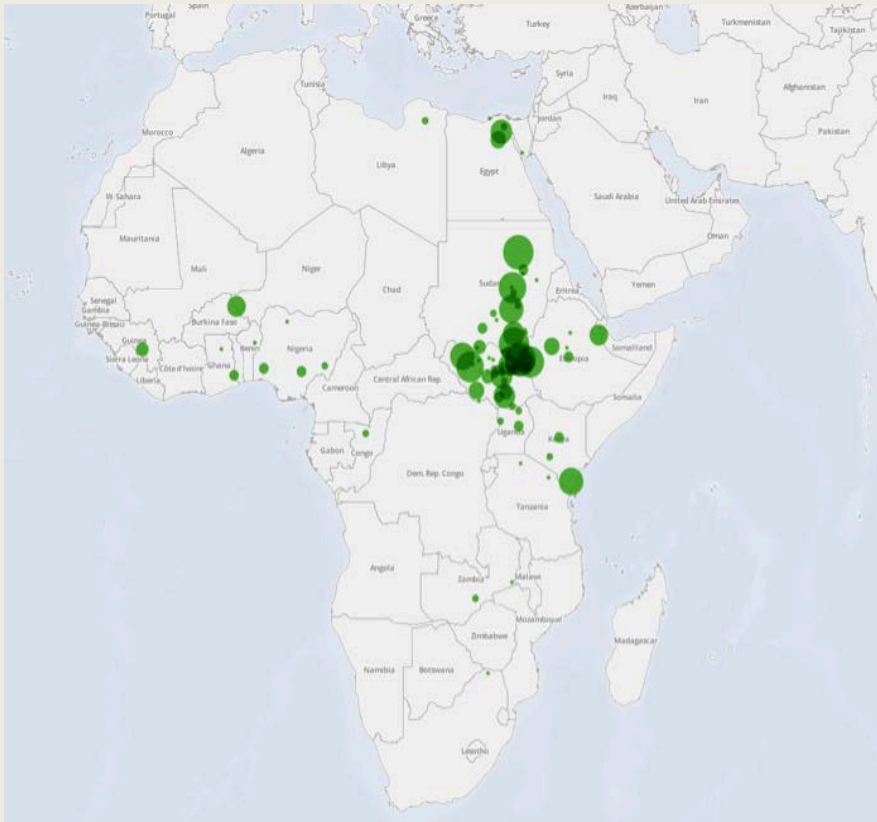
Academic Lit

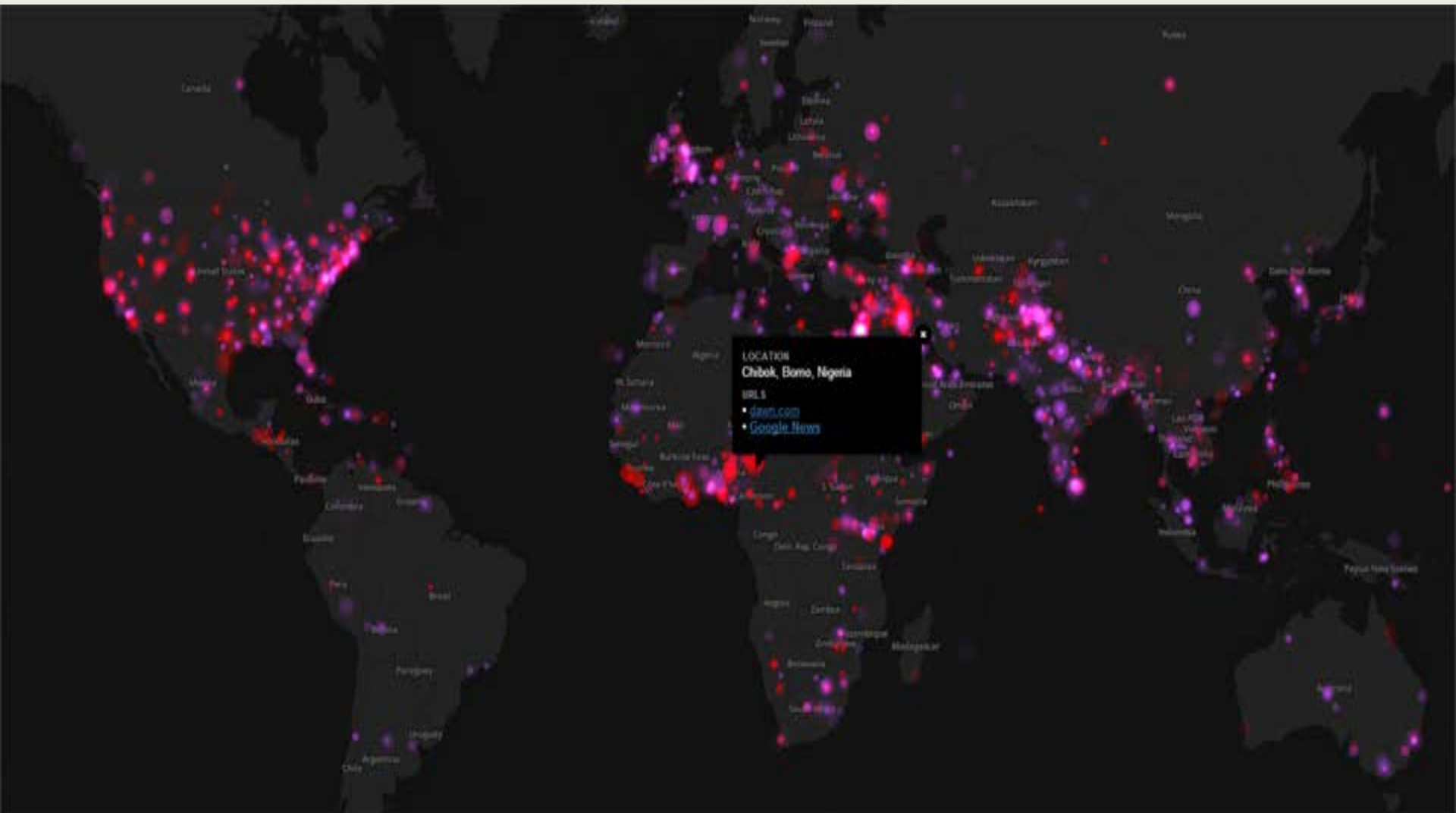


News Baseline

The GDELT Project

GLOBAL DATABASE OF EVENTS, LANGUAGE, AND TONE







Cultural Knowledge Encoding



Kalev Leetaru, Past Adjunct Faculty, Georgetown University

kalev.Leetaru5@gmail.com

Chris Rewerts, PhD, Research Engineer, U.S. Army ERDC

Chris.Rewerts@erdc.dren.mil