



# Challenges and Opportunities of Publicly Available Information (PAI)

Elizabeth Bodine-Baron  
March 20, 2019

# Overview

- 1 Publicly available information
- 2 Understanding a rapidly changing information environment
- 3 Measuring the information environment
- 4 Challenges

# Publicly Available Information (PAI)

“Information that has been **published or broadcast for public consumption**, is **available on request** to the public, is **accessible on-line** or otherwise to the public, is **available to the public by subscription or purchase**, could be seen or heard by any casual observer, is made available at a meeting open to the public, or is obtained by visiting any place or attending any event that is open to the public. Publicly available information includes information generally available to persons in a military community even though the military community is not open to the civilian general public. “

DOD MANUAL 5240.01

PROCEDURES GOVERNING THE CONDUCT OF DOD INTELLIGENCE ACTIVITIES

AUGUST 2016

# Global connectivity

**JAN  
2019**

## DIGITAL AROUND THE WORLD IN 2019

THE ESSENTIAL HEADLINE DATA YOU NEED TO UNDERSTAND GLOBAL MOBILE, INTERNET, AND SOCIAL MEDIA USE



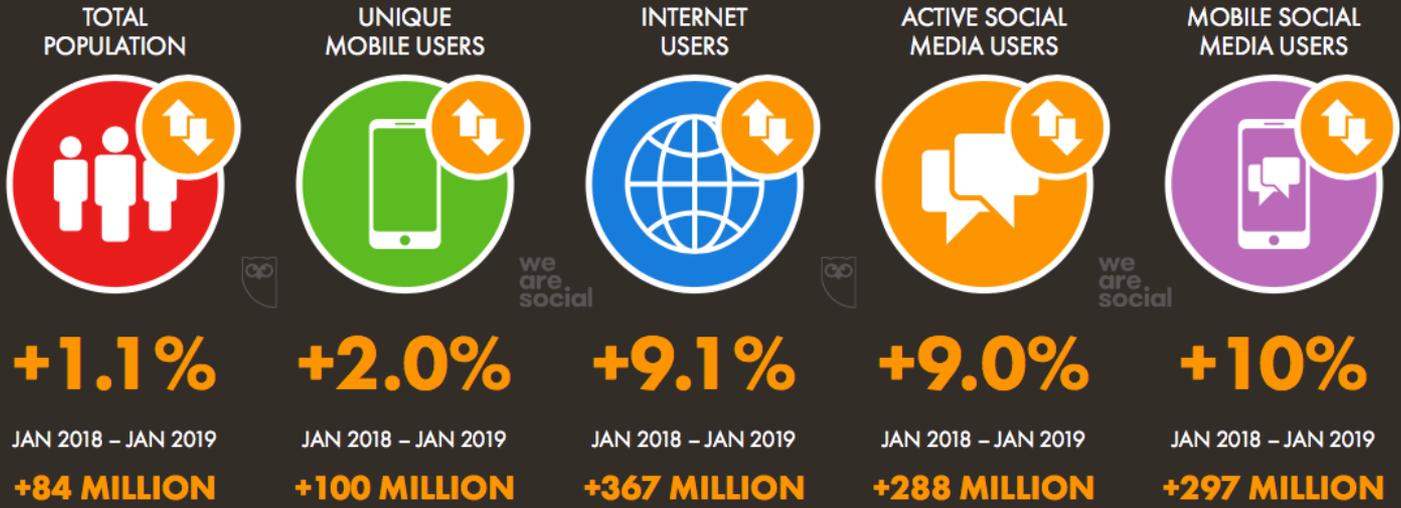
**SOURCES:** POPULATION: UNITED NATIONS; U.S. CENSUS BUREAU. **MOBILE:** GSMA INTELLIGENCE. **INTERNET:** INTERNETWORLDSTATS; ITU; WORLD BANK; CIA WORLD FACTBOOK; EUROSTAT; LOCAL GOVERNMENT BODIES AND REGULATORY AUTHORITIES; MIDEASTMEDIA.ORG; REPORTS IN REPUTABLE MEDIA. **SOCIAL MEDIA:** PLATFORMS' SELF-SERVE ADVERTISING TOOLS; PRESS RELEASES AND INVESTOR EARNINGS ANNOUNCEMENTS; ARAB SOCIAL MEDIA REPORT; TECHRASA; NIKI AGHAEI; ROSE.RU. (ALL LATEST AVAILABLE DATA IN JANUARY 2019).

# Global connectivity growth

**JAN  
2019**

## ANNUAL DIGITAL GROWTH

THE YEAR-ON-YEAR CHANGE IN KEY STATISTICAL INDICATORS



**SOURCES:** POPULATION: UNITED NATIONS; U.S. CENSUS BUREAU. MOBILE: GSMA INTELLIGENCE. INTERNET: INTERNETWORLDSTATS; ITU; WORLD BANK; CIA WORLD FACTBOOK; EUROSTAT; LOCAL GOVERNMENT BODIES AND REGULATORY AUTHORITIES; MIDEASTMEDIA.ORG; REPORTS IN REPUTABLE MEDIA. SOCIAL MEDIA: PLATFORMS' SELF-SERVE ADVERTISING TOOLS; PRESS RELEASES AND INVESTOR EARNINGS ANNOUNCEMENTS; ARAB SOCIAL MEDIA REPORT; TECHRASA; NIKI AGHAEI; ROSE.RU. [ALL LATEST AVAILABLE DATA IN JANUARY 2019].

# Social media use continues to grow

**JAN  
2019**

## **SOCIAL MEDIA OVERVIEW**

BASED ON MONTHLY ACTIVE USERS OF THE MOST ACTIVE SOCIAL MEDIA PLATFORMS IN EACH COUNTRY / TERRITORY

TOTAL NUMBER  
OF ACTIVE SOCIAL  
MEDIA USERS



**3.484**  
**BILLION**

ACTIVE SOCIAL MEDIA  
USERS AS A PERCENTAGE  
OF TOTAL POPULATION



**45%**

TOTAL NUMBER OF ACTIVE  
SOCIAL USERS ACCESSING  
VIA MOBILE DEVICES



**3.256**  
**BILLION**

ACTIVE MOBILE SOCIAL  
USERS AS A PERCENTAGE  
OF THE TOTAL POPULATION



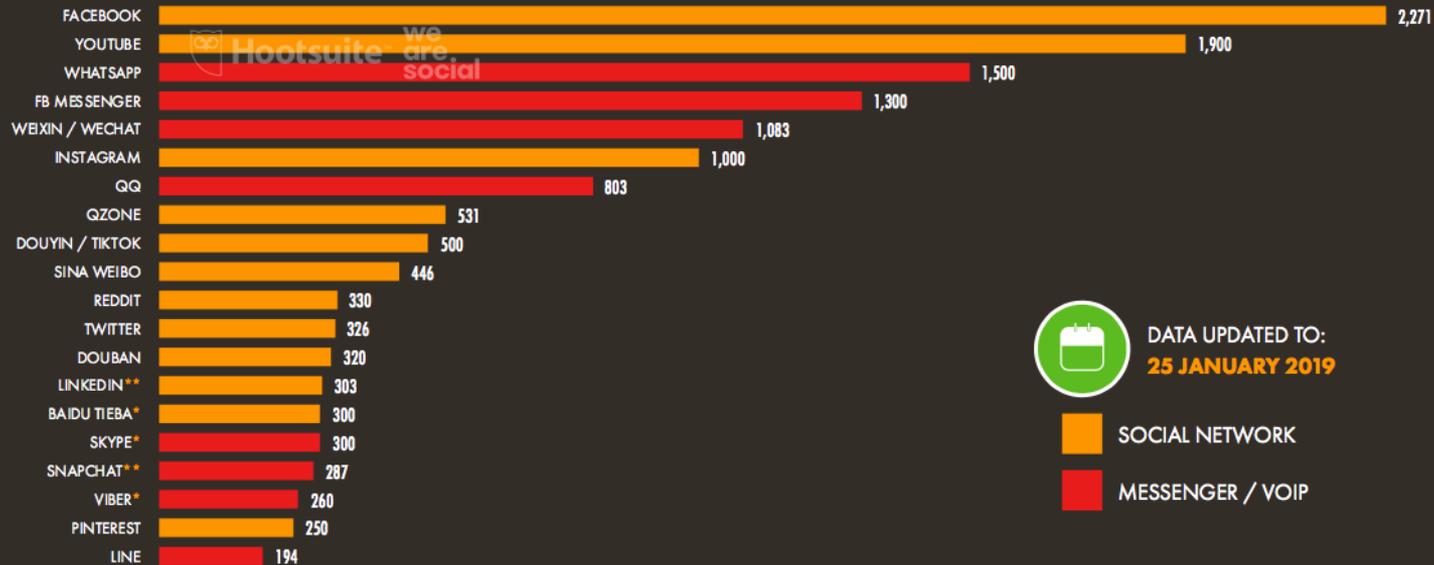
**42%**

# Social media is diverse...

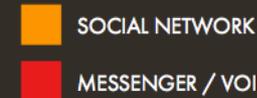
JAN  
2019

## SOCIAL PLATFORMS: ACTIVE USER ACCOUNTS

BASED ON MONTHLY ACTIVE USERS, USER ACCOUNTS, OR UNIQUE VISITORS TO EACH PLATFORM, IN MILLIONS



DATA UPDATED TO:  
25 JANUARY 2019

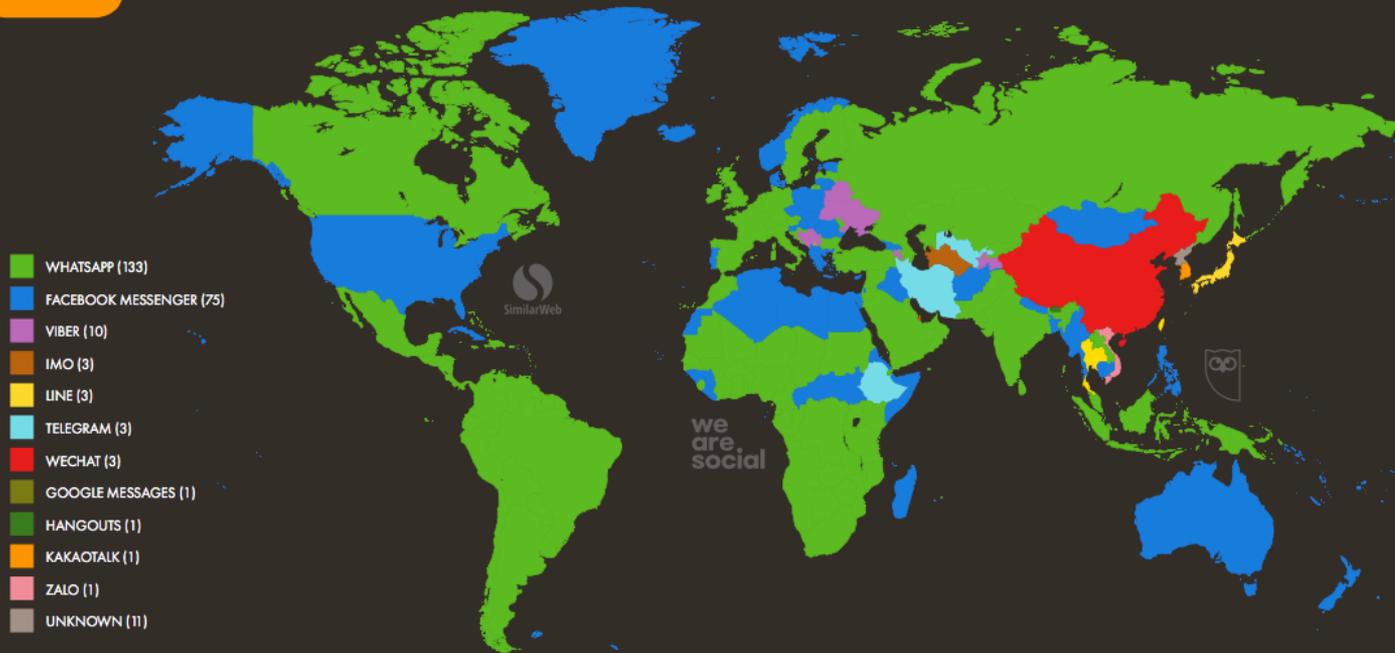


# ... And global

JAN 2019

## TOP SOCIAL MESSENGERS AROUND THE WORLD

THE MOST POPULAR MESSENGER APP BY COUNTRY / TERRITORY IN DECEMBER 2018

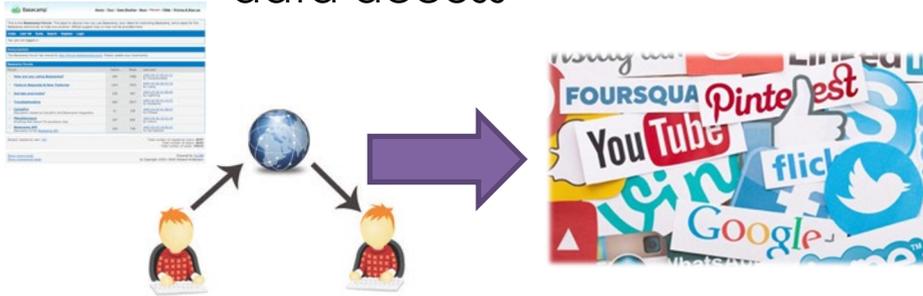


# Overview

- 1 Publicly available information
- 2 Understanding a rapidly changing information environment
- 3 Measuring the information environment
- 4 Challenges

# Understanding the information environment

- Complication:
  - Information environment is rapidly growing in size and complexity
  - Data is becoming less available through the growth of “closed” platforms and “open” platforms restricting data access



- Devices
  - Laptop/computer, smartphone, mobile, tablet..
- Channels
  - Google, Facebook, FB Messenger, YouTube, WhatsApp, Twitter, Instagram, Snapchat, Google+, Skype, Viber...
- Media
  - Text, Images, Videos, Voice
  - Often embedded (memes)

# One approach applies "old" technology to new data

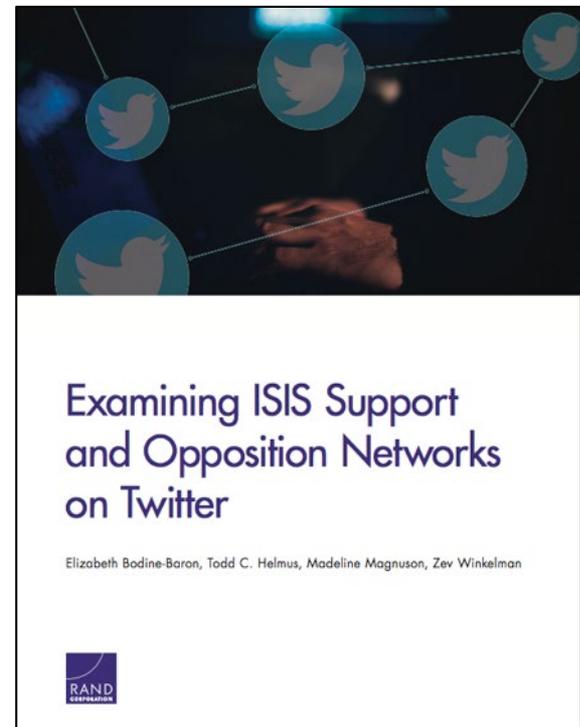
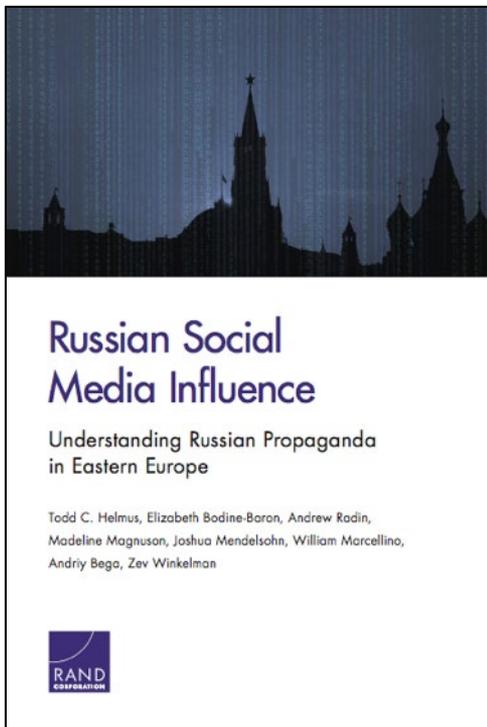
Lexical Analysis  
+  
Network Analysis  
+  
Social Media Data  
=  
Community Lexical Analysis

**Social network analysis** draws on research in sociology and graph theory to discover communities created by conversations on social media

**Lexical analysis** uses statistical analysis of text from the field of corpus linguistics to understand communities and determine who they are and what they care about

**Combining and applying these methods** to new data from social media allows for quick understanding of complex information environment

# Two example applications of CLA to map information environment



# Step 1: Collect data



Three-month sample  
(May–July 2016)

**23M**

Russian-language  
tweets

**500K**

Unique  
accounts



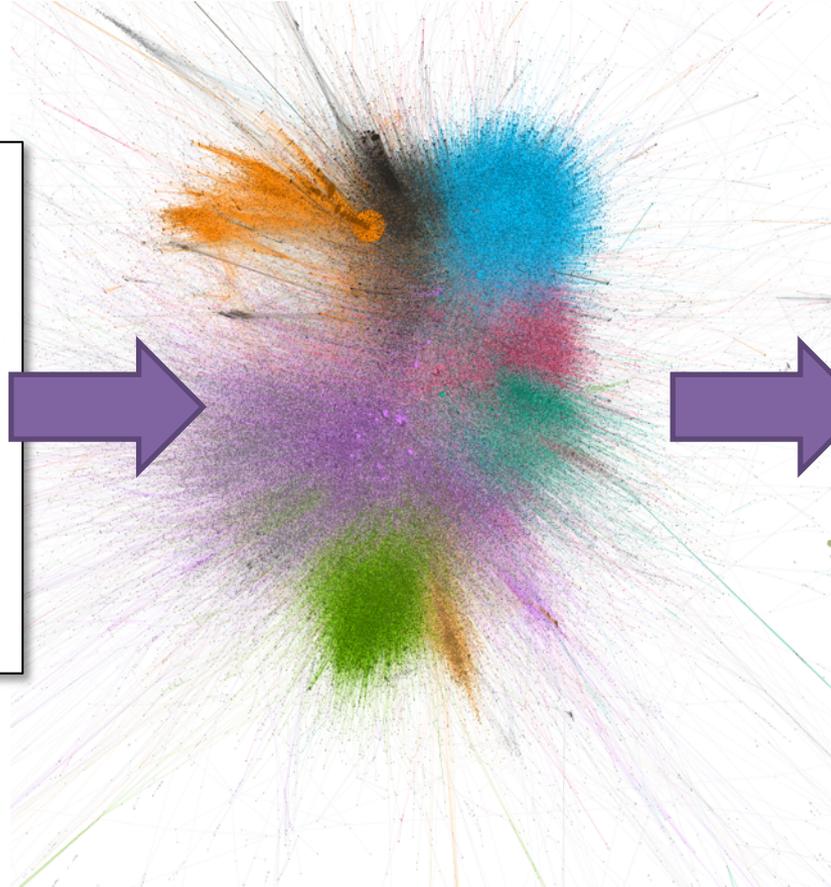
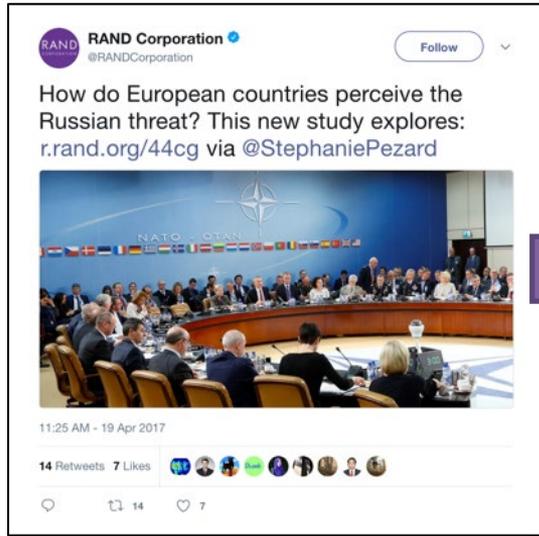
## Russian Social Media Influence

Understanding Russian Propaganda  
in Eastern Europe

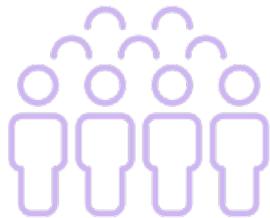
Todd C. Helmer, Elizabeth Bodine-Barron, Andrew Radin,  
Madeline Magnuson, Joshua Mendelsohn, William Mancillina,  
Andriy Begs, Zer Winkelman



# Step 2: Create mentions network and find communities



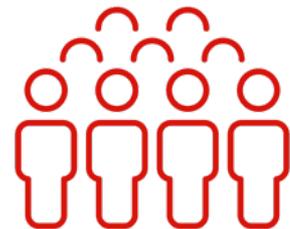
# We found two sides of an information war



Ukrainian  
activists

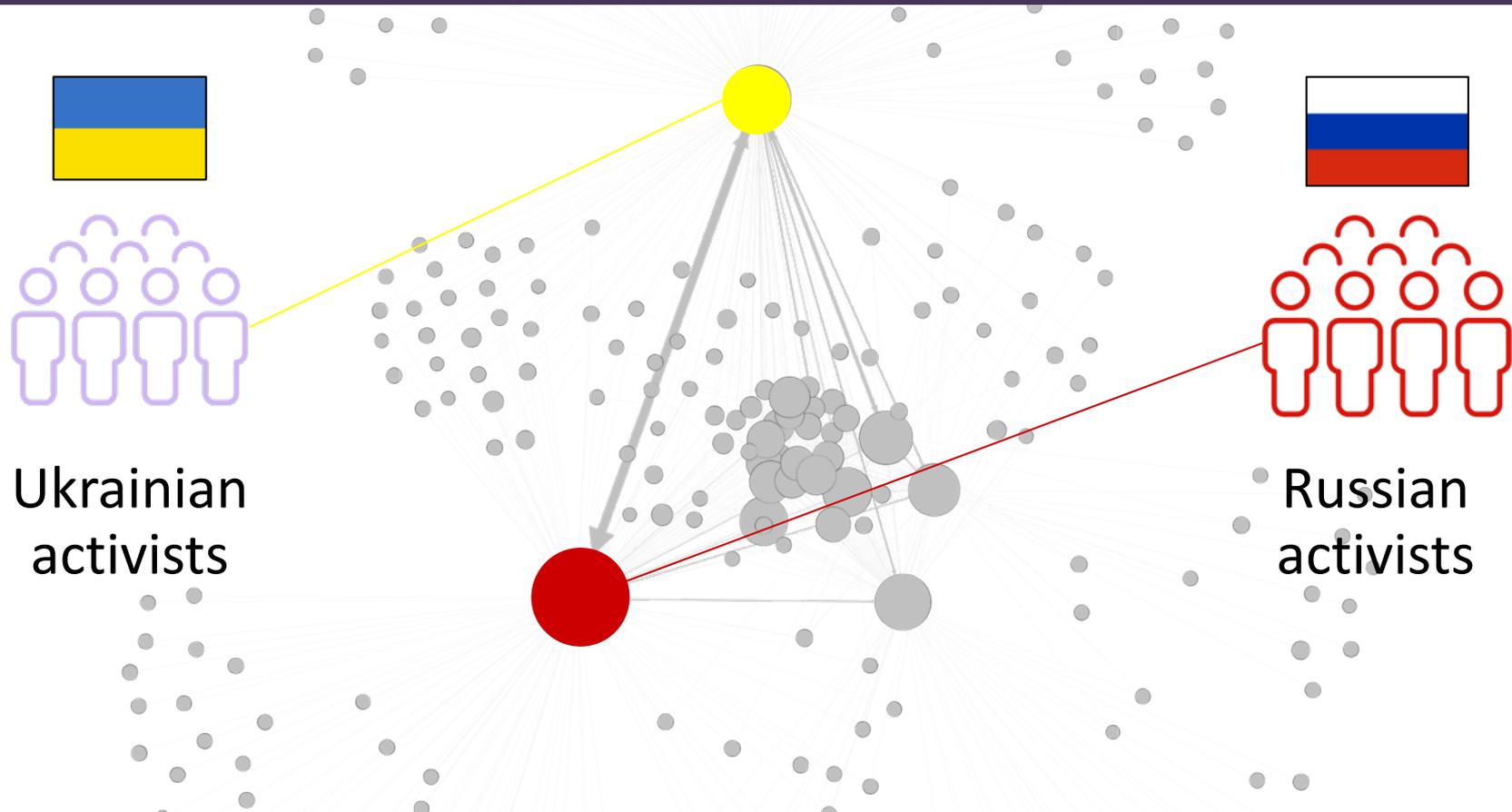
40,000

Members in each group

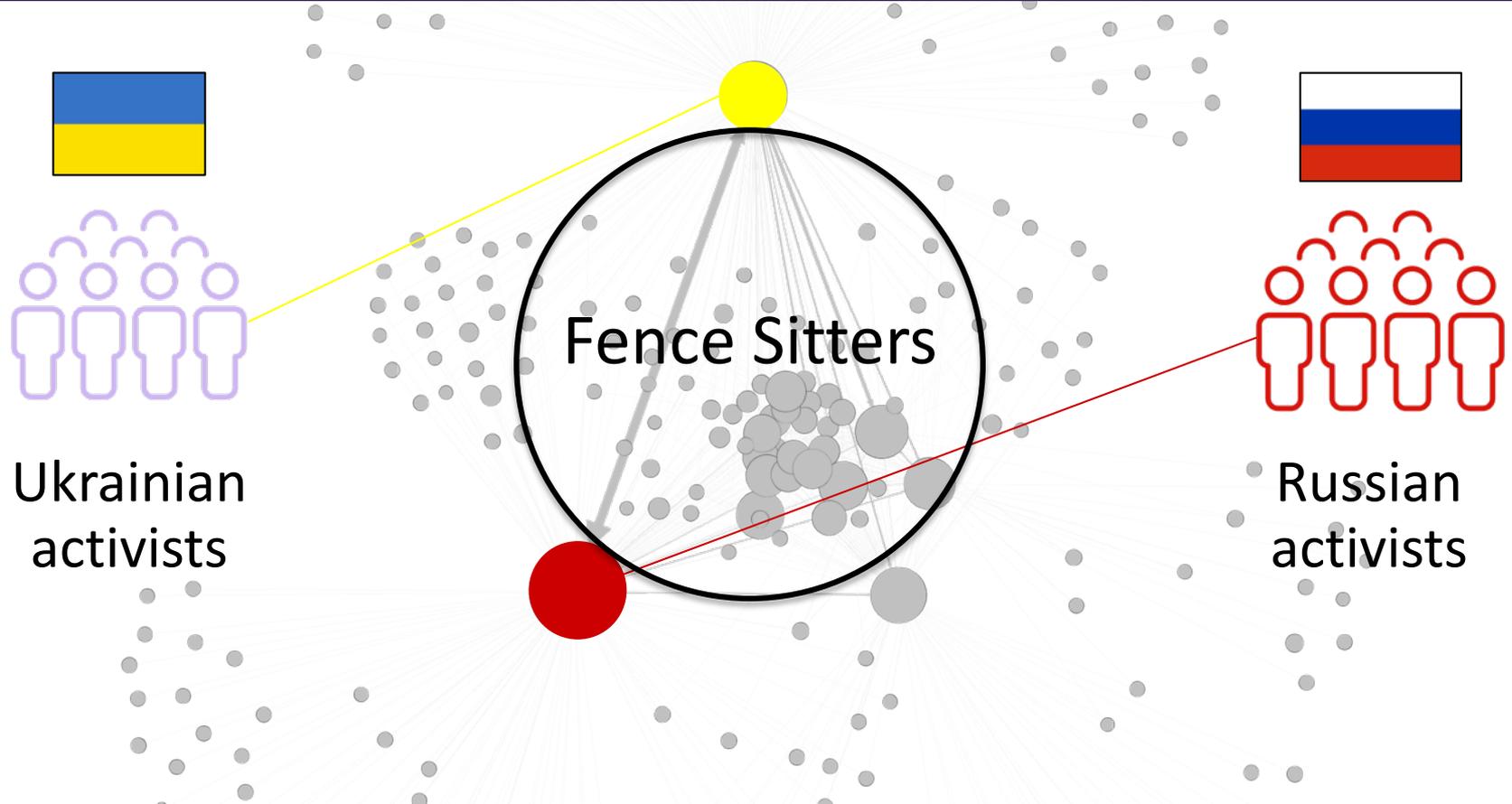


Russian  
activists

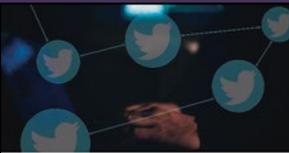
# Two opposing poles in the community network



# Several groups connected to both communities



# ISIS supporters and opponents on Twitter



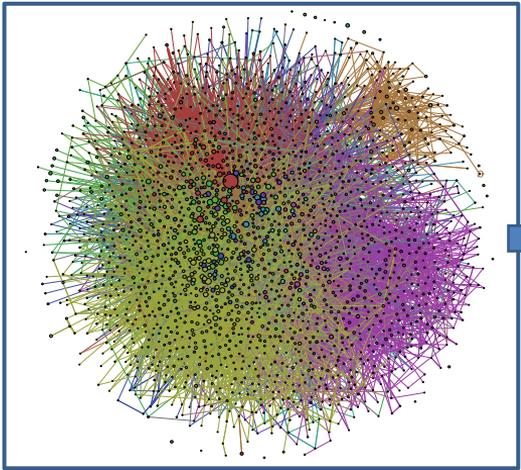
Examining ISIS Support and Opposition Networks on Twitter

Elizabeth Bohrer-Soren, Todd C. Helmus, Madeline Magnuson, Zhe Wu

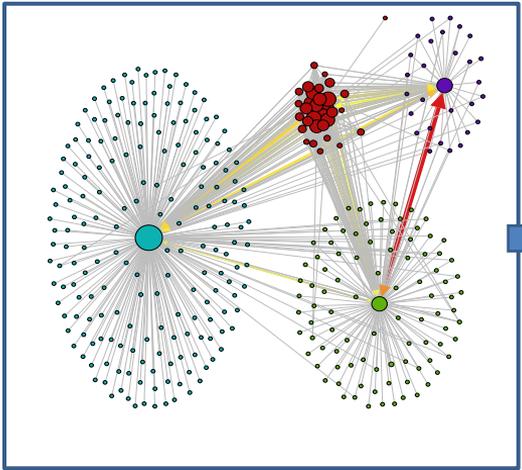
Data: Twitter data from July 1, 2014 – April 30, 2015 using Arabic search terms/variations for “Islamic State”

“Daesh”

## Create Mentions Network



## Detect Communities



## Characterize Communities with Lexical Analysis

Sunni	Saudi-specific concerns, condemnation of ISIS, distrust of West & Iran
Pro-ISIS	<i>Rawafidh/Safavid/Nusayri; religion &amp; belonging themes</i>
Syrian Muj	Negativity towards ISIS and International Coalition
Shia	Links ISIS to Saudi Arabia, positivity towards Christians/West

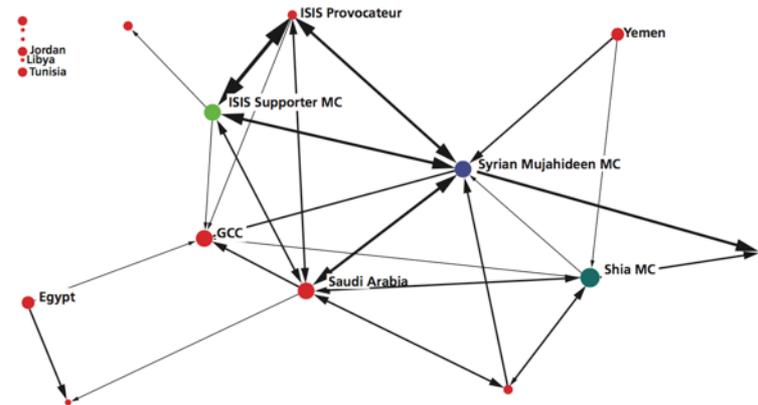
This approach uncovered a battle of ideas between socially meaningful groups

# Understand each community and how they are connected

Sunni community was comprised of several disparate nationalist groups – each needing tailored narrative and message themes

Community	Potential Message Themes
Saudi Arabia	Appeals to secular nationalism and desire to be part of the international community; counter ISIS support and expansion in SA, while addressing the threat to Islam posed by Iranian Shiism
Egypt	Appeals to nationalism and ISIS opposition, while addressing mistrust of Brotherhood and frustration with U.S. policies
Jordan	Appeals to nationalism, taking advantage of support for Moath al-Kasabeh and international air campaign
Libya	Appeals to nationalism and ISIS opposition, while addressing distrust of Libyan politicians, militants and west
Yemen	Enhance opposition to ISIS, while acknowledging population has substantial numbers of ISIS supporter and is overwhelmingly critical of KSA intervention

For messaging, choose which community to target based on network position

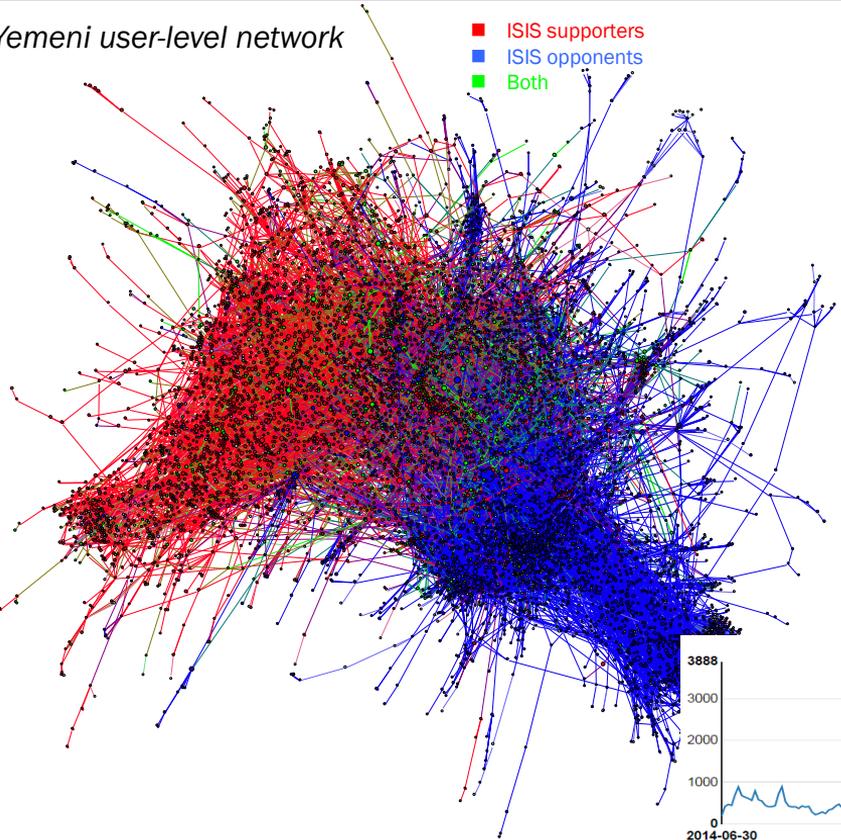


SOURCE: RAND analysis, Twitter data from July 2014 to May 2015.

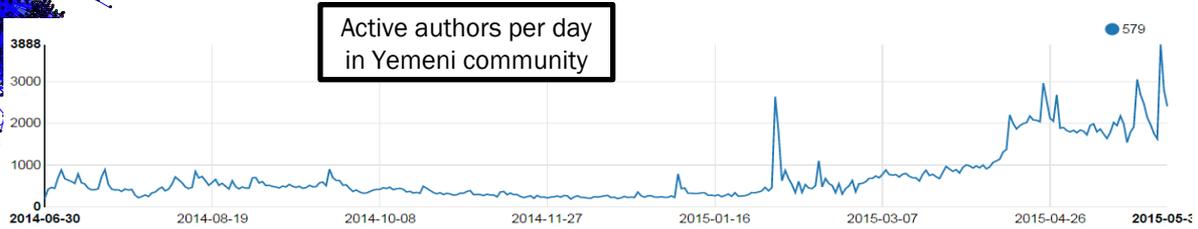
# Understand the conversation at the user level

Yemeni user-level network

- ISIS supporters
- ISIS opponents
- Both



Metrics to identify key influencers in online conversation



# Overview

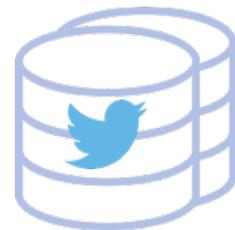
- 1 Publicly available information
- 2 Understanding a rapidly changing information environment
- 3 Measuring the information environment
- 4 Challenges

# Measuring the information environment: Resonance analysis

## Step 1: Create a lexical “fingerprint”



Propaganda Source



Baseline text

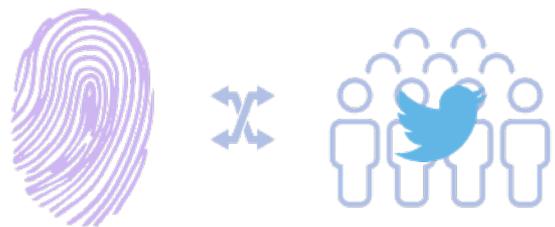
Analysis can assess a population’s affinity for propaganda by source, geographic area, and over time



“Fingerprint”= linguistic patterns in propaganda corpus, compared to a baseline text

# Measuring the information environment: Resonance analysis

## Step 2: Find matching patterns in target population



Matches = RESONANT

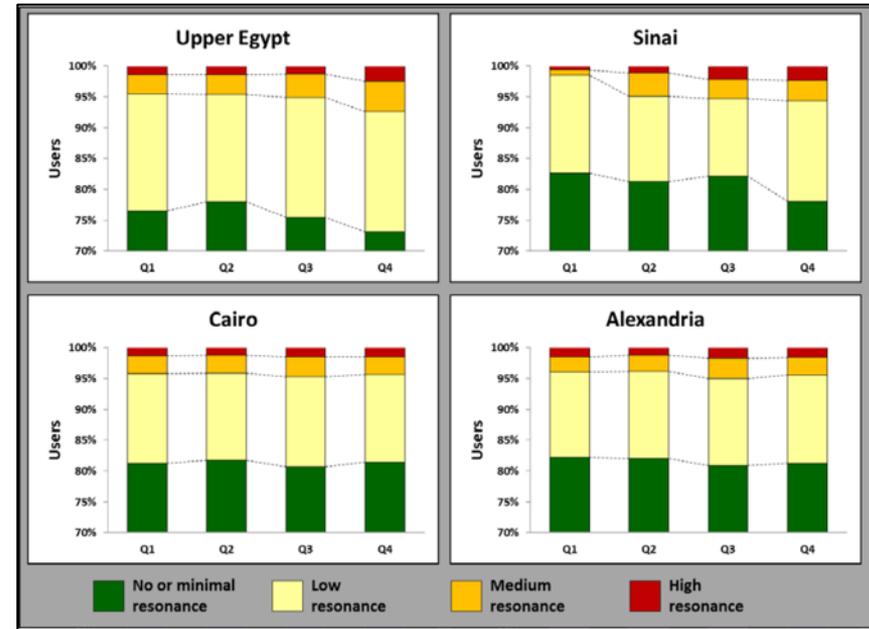
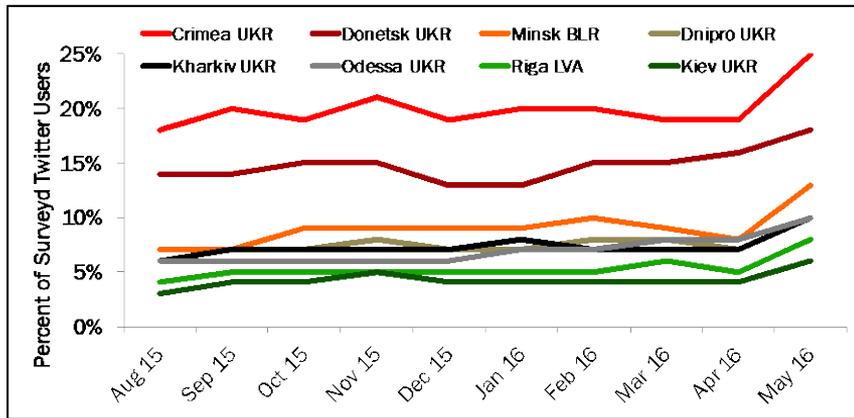
Method searches for similar word-choice patterns in target population to assess “impact” of propaganda



# Measuring the information environment: Resonance analysis

## Assessing the impact of propaganda and influence campaigns

*Russian propaganda signature is strongest where support for Russia is high*



*ISIS resonance increased eastward; ISIS gained “market share” where Muslim Brotherhood lost*

# Overview

- 1 Publicly available information
- 2 Understanding a rapidly changing information environment
- 3 Shaping and measuring information operations
- 4 Challenges

# Problems will remain, but some are solvable in the near term

1. Data continues to grow exponentially
  - Numbers but also content diversity: images/videos over text
  - Apply machine learning to leverage limited human resources
2. Developing *rigorous, repeatable* measures of effectiveness (MOEs) of information operations is incredibly difficult
  - Move beyond measures of performance (MOPs)
  - Must be able to rapidly assess changes in the information environment at the appropriate level of detail
3. Data is becoming less accessible for measurement and monitoring
  - Increase in data privacy restrictions (GDPR)
  - Unwillingness of private companies to work with government clients

