

# The digital evolution of Occupy Wall Street

{ Emilio Ferrara – SOIC – CNetS – IU Bloomington  
ferrarae@indiana.edu

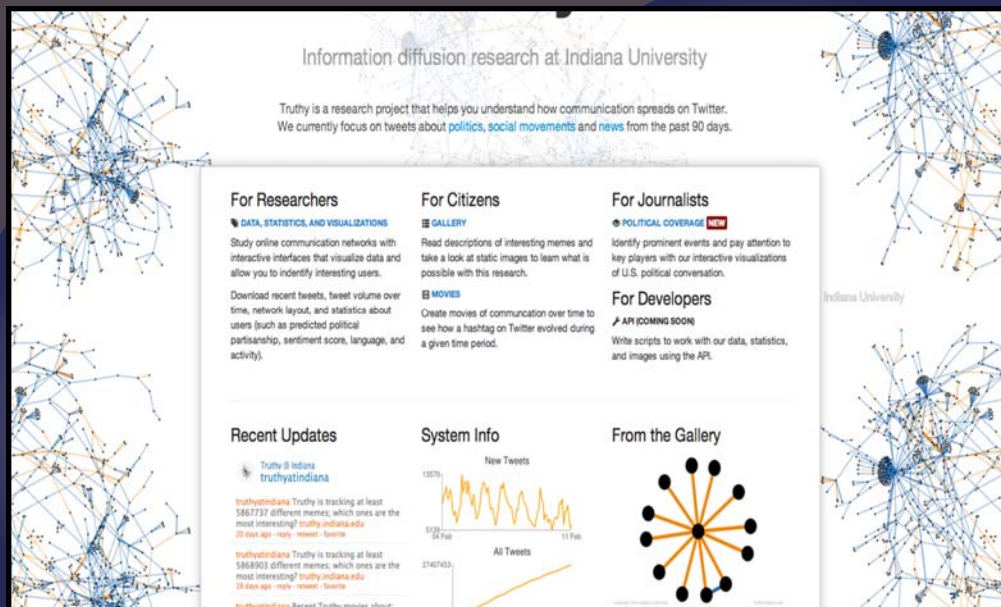
Joint work with:  
MD Conover, C Davis, K McKelvey, F Menczer and A Flammini



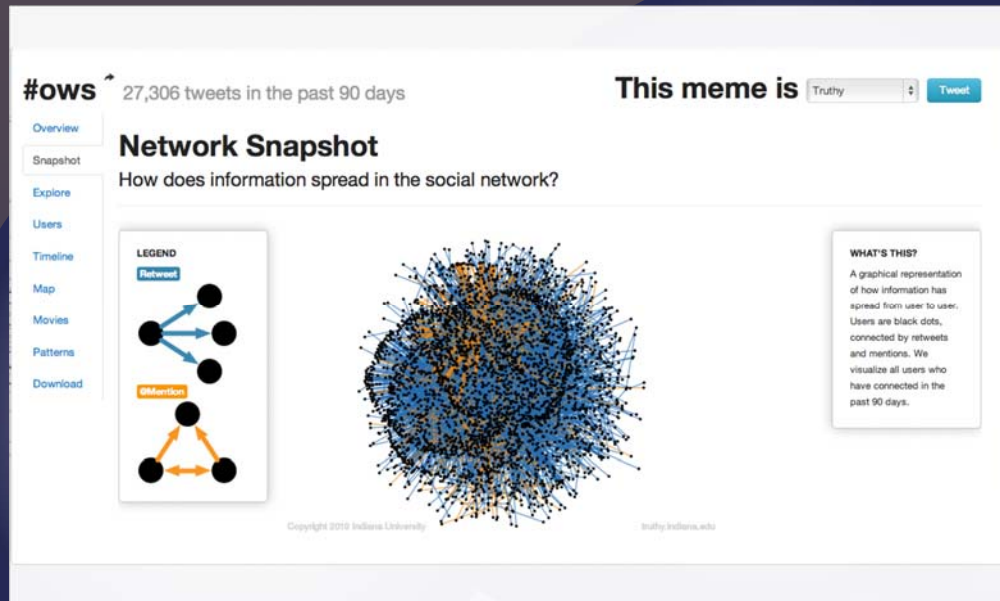
Center for Complex Networks  
and Systems Research



SCHOOL OF INFORMATICS  
AND COMPUTING  
INDIANA UNIVERSITY  
Bloomington



## The truthy project



## Meme diffusion on Social Media

## Occupy movement

- ↳ Motivations of the protest
  - ⌘ Social and wealth inequalities, taxation disparities
  - ⌘ Non sustainable capitalistic market models
  - ⌘ Political corruption, corporate influence of government
- ↳ Mottos and Motives
  - ⌘ "We are the 99%"
  - ⌘ Pacifist protest
- ↳ Key dates and events
  - ⌘ #occupywallstreet Adbusters
  - ⌘ Sep 17<sup>th</sup>, 2011 – Zuccotti park encampment
  - ⌘ May 1<sup>st</sup>, 2012 – Attempt to revive the movement
  - ⌘ Massive on-the-ground protests in NY, DC and California <sup>[1]</sup>



**YOUR LOSS, THEIR GAIN**

How much income have you given up for the top 1 percent?

INCOME GROUP	TOTAL LOSS/GAIN IN ANNUAL INCOME*	AVERAGE LOSS/GAIN PER HOUSEHOLD PER YEAR*
<b>TOP 1%</b>	\$673 billion more	\$597,241 more
96-99	\$140 billion more	\$29,895 more
95-95	\$39 billion more	\$4,912 more
81-90	\$43 billion less	\$3,733 less
61-80	\$194 billion less	\$8,598 less
41-60	\$224 billion less	\$10,100 less
21-40	\$189 billion less	\$8,582 less
<b>BOTTOM 20%</b>	\$196 billion less	\$5,623 less

\* Compared to what incomes would have been had all income groups seen the same growth rate in 1979-2009 as they did during previous decades. Source: Jacob Hacker, Yale University; Paul Pierson, UC-Berkeley



[1] [http://en.wikipedia.org/wiki/Timeline\\_of\\_Occupy\\_Wall\\_Street](http://en.wikipedia.org/wiki/Timeline_of_Occupy_Wall_Street)

# Research Outline

↳ Research questions:

- ⌘ How did protesters use Online Social Media (OSM)?
- ⌘ How much localized the diffusion of Occupy-related information on OSM is?
- ⌘ What type of users Occupy involved in OSM discourse?
- ⌘ What's the demographics of Occupy users on OSM?
- ⌘ Did Occupy change online users' behaviors, connectivity and interests?



## What's the impact of Occupy on Online Social Media?

# Data Collection

- ⌘ “Geography” dataset [1]:
  - ⌘ Occupy corpus
    - ⌘ From Jul 3<sup>rd</sup>, 2011 to Mar 12<sup>th</sup>, 2012
    - ⌘ Any tweet containing #ows or #occupy{\*}
    - ⌘ Total of 1.5M tweets [676K RT] produced by 257K distinct users
  - ⌘ Domestic politics corpus (baseline)
    - ⌘ Any tweet containing #tcot (top conservatives on Twitter) or #p2 (Progressives 2.0)
    - ⌘ Total of 825K tweets [259K RT] produced by 68K distinct users
- ⌘ “Temporal” dataset [1]:
  - ⌘ Occupy corpus
    - ⌘ From Sep 1<sup>st</sup>, 2011 to Aug 31<sup>st</sup>, 2012
    - ⌘ Any tweet containing #ows or #occupy{\*}
    - ⌘ Total of 1.82M tweets produced by 447K distinct users
  - ⌘ Random sample
    - ⌘ From Jun 1<sup>st</sup>, 2011 to Aug 31<sup>st</sup>, 2012 [3 months before ows]
    - ⌘ Any tweet produced by a random set of 25K ows users
    - ⌘ Total of 7.74M tweets

[1] Data collected from Twitter garden-hose (10% sample rate)

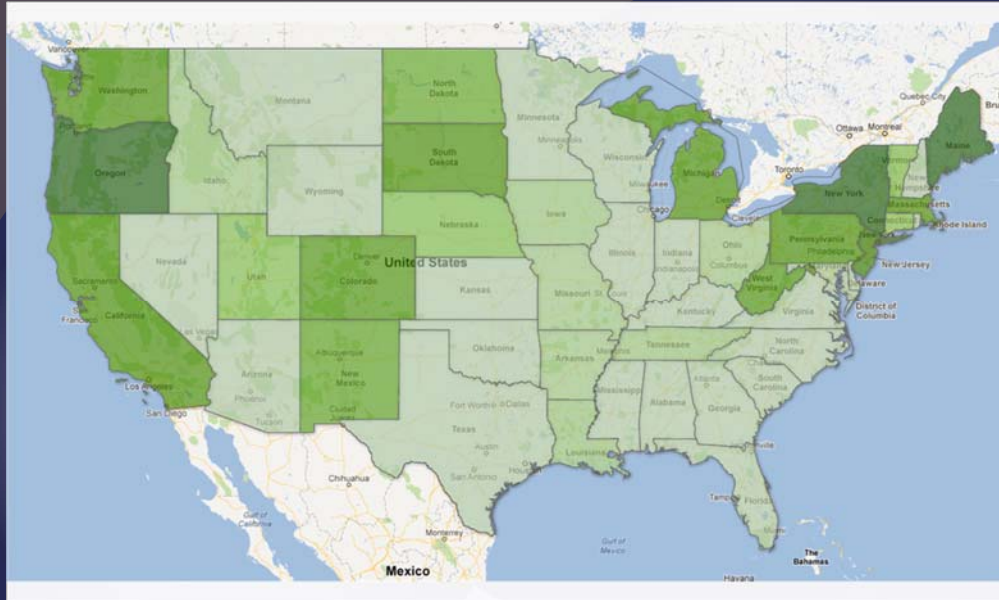
# Geocoding locations

- ⌘ Only ~1% of total tweets contains geo-data (e.g., GPS coordinates)
- ⌘ People in encampments and during marches were committed to spread news on OSM
- ⌘ from users’ self-reported locations
- ⌘ Goal: geocoding w/ high accuracy level high-volume Twitter traffic
- ⌘ Methods: Bing API + blacklist + whitelist + fuzzy string matching
- ⌘ Result: 55.7% and 29.3% of Occupy and Domestic politics users geocoded



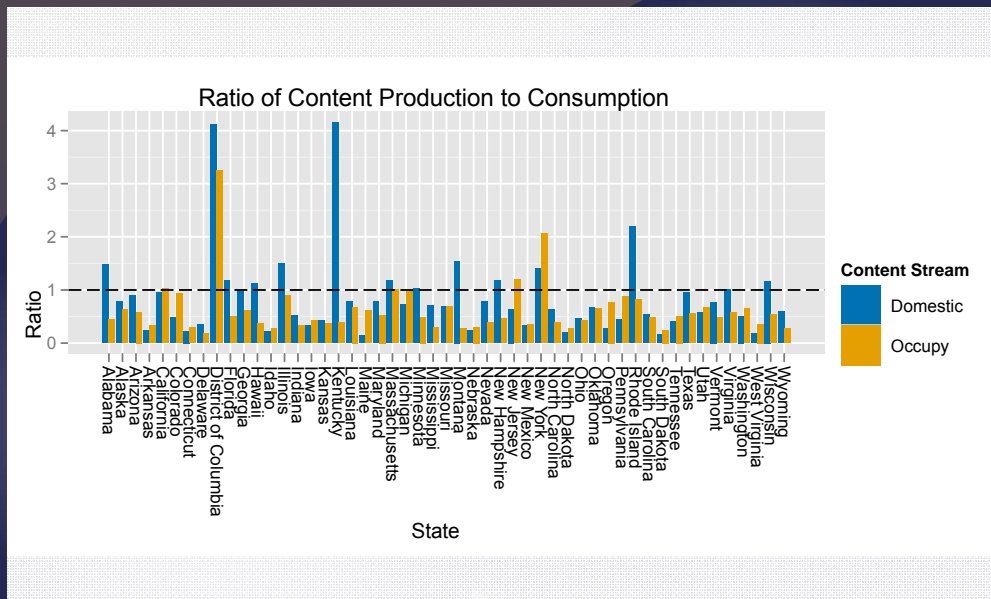






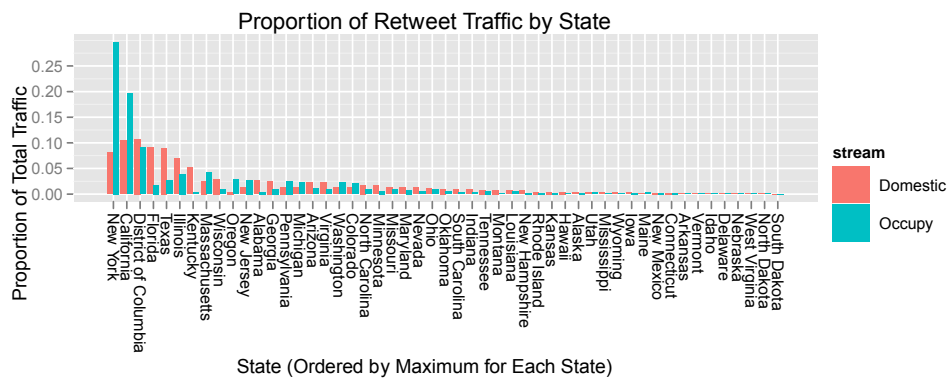
## Deviation from the baseline

The intensity in the color represents how much the amount of Occupy-related traffic deviates from that of domestic politics per state.

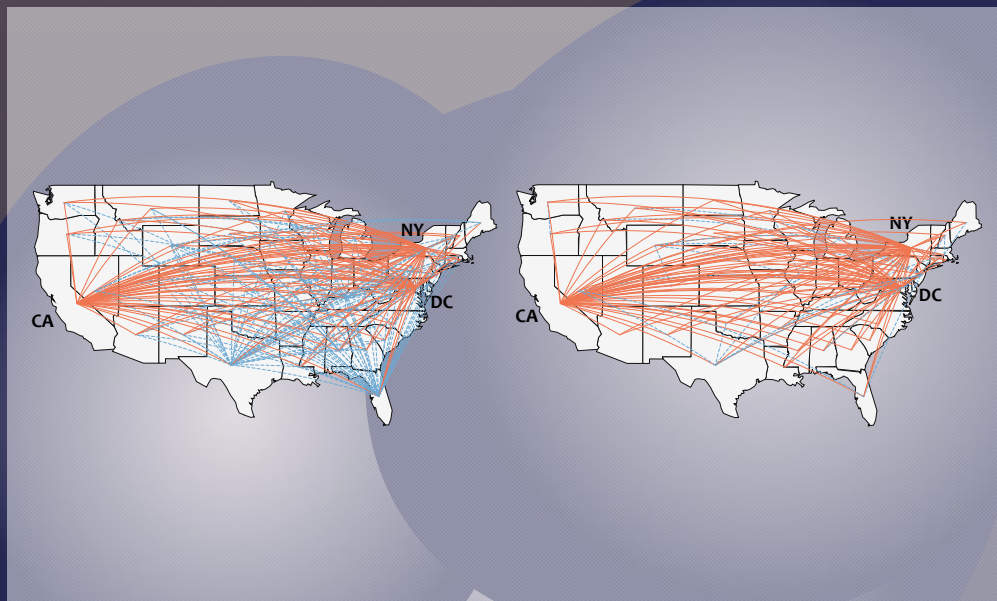


## Content production and consumption

Ratio = # RT originating from users in the state / # RT retweeted by users in the state



# Content diffusion



# Interstate communication

Occupy-related discourse show a prominent hub-and-spoke structure differently from domestic politics (on the left).  
Multiscale backbone extraction – confidence level  $\alpha = 0.15$

Collective framing

Interstate		Intrastate	
Token	Ratio	Token	Ratio
wall	.590	city	2.254
nyc	.600	tonight	1.737
street	.699	march	1.669
news	.718	join	1.494
99%	.756	solidarity	1.387
bank	.763	day	1.354
don't	.782	square	1.333
media	.837	please	1.243
peaceful	.845	park	1.220
nypd	.847	now	1.179

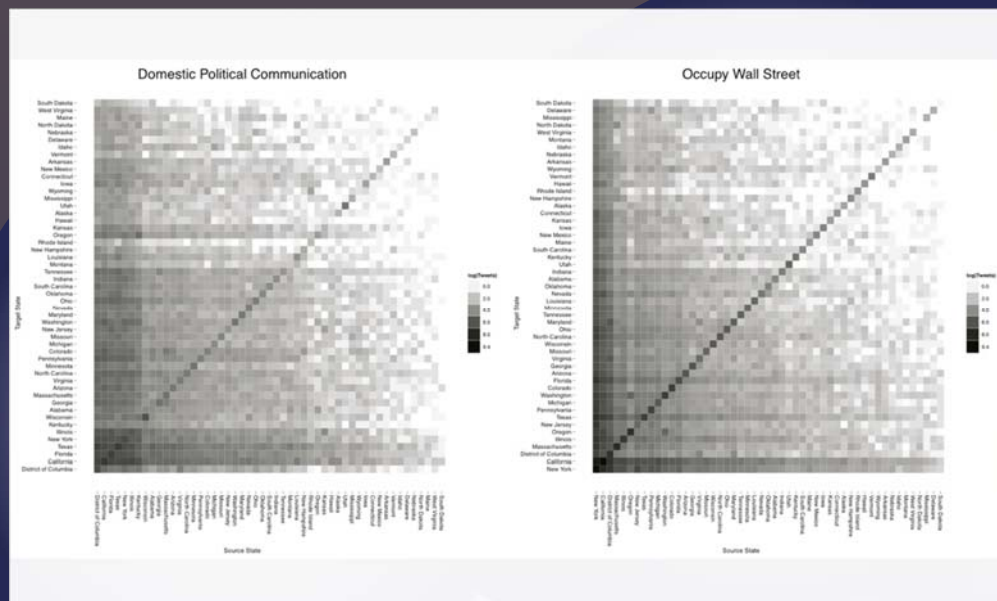
Resource mobilization

$$\text{Ratio} = \frac{P(\text{Token} | \text{Intrastate})}{P(\text{Token} | \text{Interstate})}$$

# Language and framing

**Collective framing:** the social processes whereby movement participants negotiate the shared language and narrative frames that help define the movement's identity and goals.

**Resource mobilization:** the work to marshal the physical and technological infrastructure, human resources, and financial capital necessary to sustain ongoing activity.



# Local vs. global communication

Occupy communication patterns exhibit heightened diagonal activity (more than 3 times) than domestic politics ones.



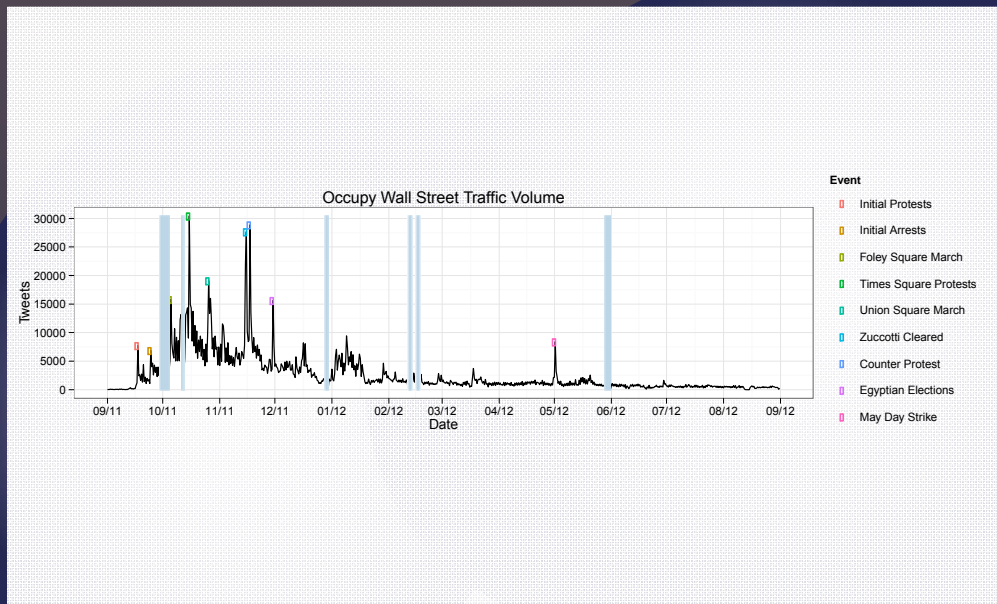
## Summary of geo-analysis

- ↳ Occupy discourse on Twitter has highly localized geospatial structure: a large amount of traffic is produced and consumed locally per state. This might be explained by **Resource Mobilization**.
- ↳ Interstate communication is driven by high-profile locations acting as information broadcasters. This might represent the **Collective Framing** process.
- ↳ Proximity to on-the-ground events plays a big role: users from **NY, DC** and **California** are the main actors of the discourse. They produce much more Occupy-related information than that they consume, unlike other states.



## The evolution of Occupy

This sign was put in front of a media tent at the Occupy Toronto camp in St. James Park (Toronto). Image by Hillary Burrridge.



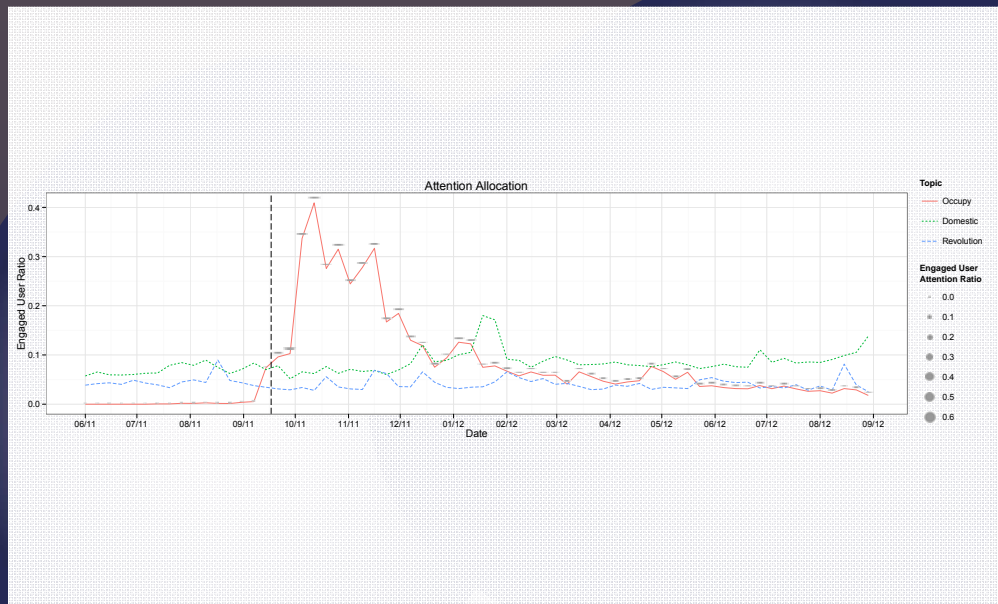
# Activity volume

On-the-ground events (circles) and Twitter data-stream outages (blue bands) are highlighted. Bins have 12-hours length.

Domestic Politics	Social Movements
#tcot	#syria
#p2	#bahrain
#teaparty	#egypt
#gop	#yemen
#anonymous	#libya
#obama	#tahrir
#tlot	#wiunion
#jobs	#iranelection
#ronpaul	#assange
#romney	#wikileaks
#sopa	#jan25
#ndaa	#14feb
#obama2012	#assad
#ocra	#greece
#twisters	#damascus
#sgp	#gaddafi
#politics	#feb14
#solidarity	#scaf
#gop2012	#antisec
#p21	#arabspring
#topprog	#tunisia
#obamacare	#noscaf
#mapoli	#syrian

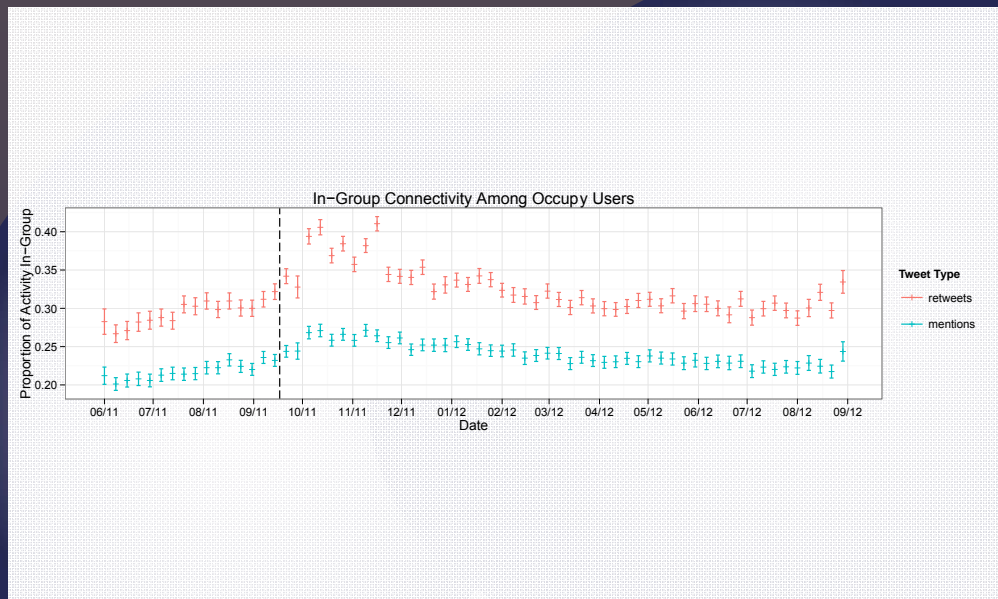
# Politics and protest keywords

The top hashtags adopted by the 25K random users related to domestic politics and foreign social movements.



# Attention allocation

Engaged User Ratio describes the proportion of active users in each time-step who produced at least one topically-related tweet. Engaged User Attention Ratio describes, among these users, the share of average attention allocated to each topic.



# In-group connectivity

Reported values represent means and 95% confidence intervals for each time-step.



## Demography of supporters

### Summary of temporal analysis (I)

- ✎ Occupy discourse on Twitter is highly correlated in time with on-the-ground events. Spikes and high volume of traffic coincide with relevant protests and police actions, marches and strikes.
- ✎ The phase of explosive growth has exhausted few weeks after the beginning of the protest.
- ✎ The volume of Occupy-related traffic has diminished by orders of magnitude in the latest observation period.



## Summary of temporal analysis (II)

- ⌘ Occupy captured the attention of users with pre-existing interests in politics and social protests.
- ⌘ Occupy users' interests and amount of attention dedicated to politics and social discourse did not exhibit any remarkable variation over time.
- ⌘ Occupy users were already highly interconnected before the movement's start.
- ⌘ The extent to which Occupy users interacted before the movement's start and at the end of the observation period is substantially unchanged.

## Current and future work

- ⌘ How groups of users create, interact and evolve in Online Social Media?
- ⌘ Can we generalize some features observed for Occupy (e.g., explosive growth, fast decay, etc.) to other classes of topics discussed on OSM?
- ⌘ How the social and the topical space interact each other to determine the success of certain topics of discussion and the failure of the vast majority of the others?

## &References

- MD Conover, C Davis, E Ferrara, F Menczer, A Flammini. The Geospatial Characteristics of a Social Movement Communication Network. PLoS ONE (2013)
- MD Conover, E Ferrara, F Menczer, A Flammini. The Digital Evolution of Occupy Wall Street. PLoS ONE (in preparation)



Fil Menczer



Sandro Flammini



Alex Vespignani



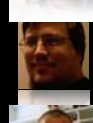
Johan Bollen



Emilio Ferrara



Jacob Ratkiewicz



Mark Meiss



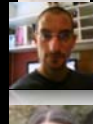
Mike Conover



Mohsen Jafari-Asbagh



Bruno Gonçalves



Alex Rudnick



Karissa McKelvey



Onur Varol



Lilian Weng



Luca Aiello



Przemek Grabowicz

# Thanks!

James S. McDonnell Foundation

