VAST 13

Visual Analysis of Topic Competition on Social Media

Panpan Xu[1], Yingcai Wu[2], Enxun Wei[2], Tai-Quan Peng[3], Shixia Liu[2], Jonathan J.H. Zhu[4], Huamin Qu[1]

1 Hong Kong University of Science and Technology

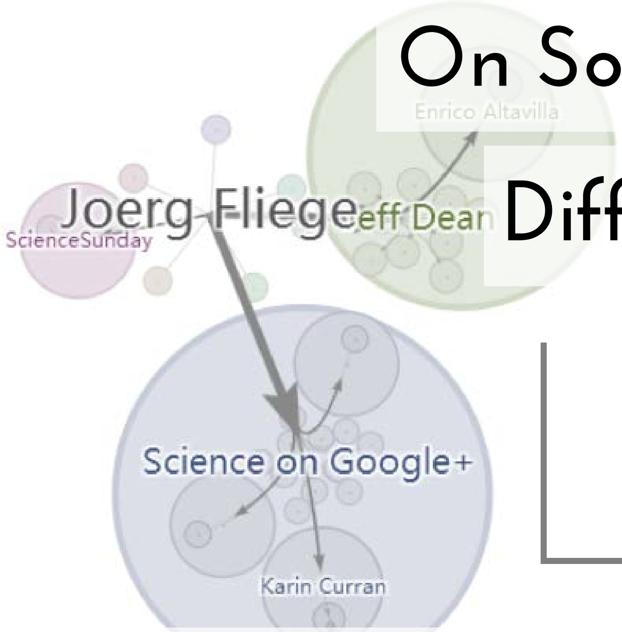
2 Microsoft Research Asia

3 Nanyang Technological University

4 City University of Hong Kong

On Social Media:

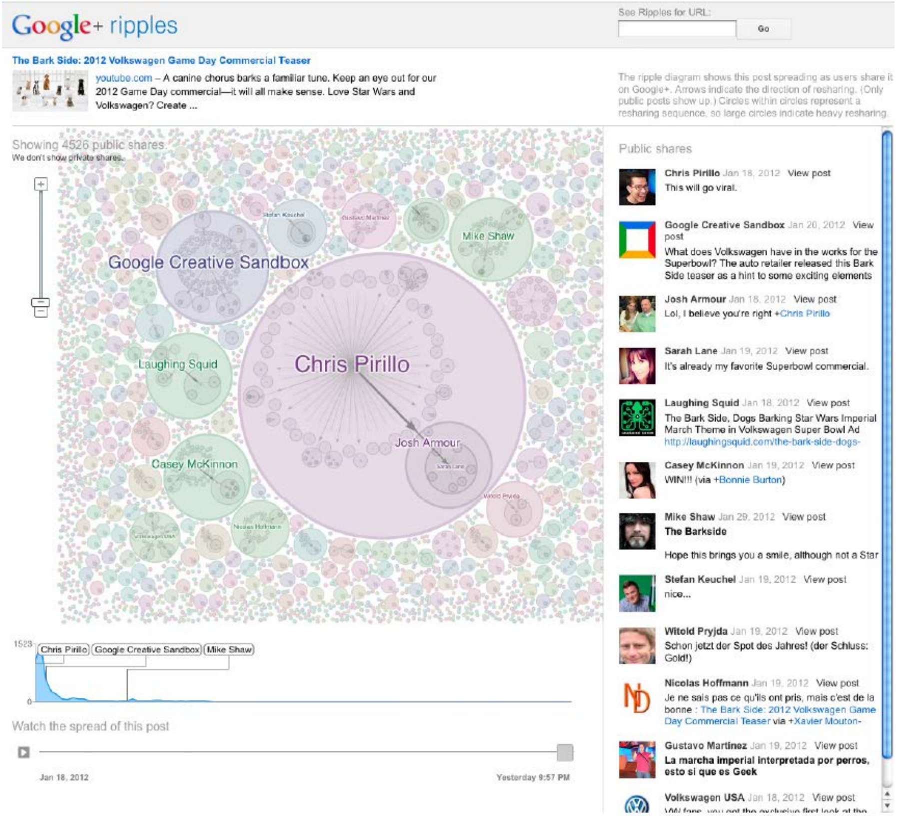
Diffusion of multiple topics



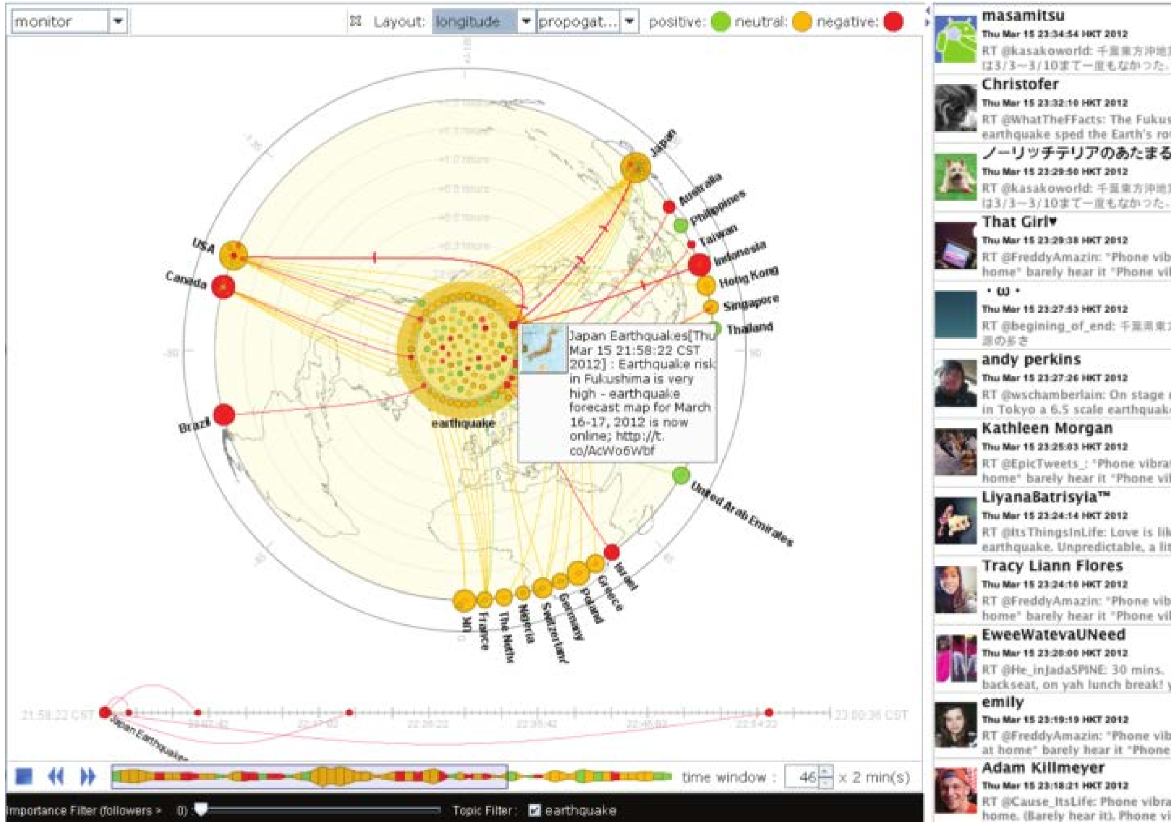
Google Ripples

The Interaction: Do people get distracted away from some topics when something more “eye-catching” is happening?

The Influence: How do the opinion leaders (influential users) affect the interaction by recruiting the public attention for some topics?



Google Ripples [F. Viégas et al. 11]



Whisper [N. Cao et al. 12]

INTRO / SYSTEM / MODEL / DESIGN / CASE STUDY

Agenda-setting

[M. E. McCombs and D. L. Shaw 72]

The ability of the news media (e.g. TV and newspaper) to influence the salience of topics on the public agenda.

Topic competition

[J. Zhu 92]

The addition of any new topic onto the public agenda comes at the cost of other topic(s).

Two-step information flow

[S. Wu et.al 11]

The information reaches the masses via intermediaries.

Agenda-setting

[M. E. McCombs and D. L. Shaw 72]

The ability of the news media (e.g. TV and newspaper) to influence the salience of topics on the public agenda.

Topic competition

[J. Zhu 92]

The addition of any new topic onto the public agenda comes at the cost of other topic(s).

Two-step information flow

[S. Wu et.al 11]

The information reaches the masses via intermediaries.

Agenda-setting

[M. E. McCombs and D. L. Shaw 72]

The ability of the news media (e.g. TV and newspaper) to influence the salience of topics on the public agenda.

Topic competition

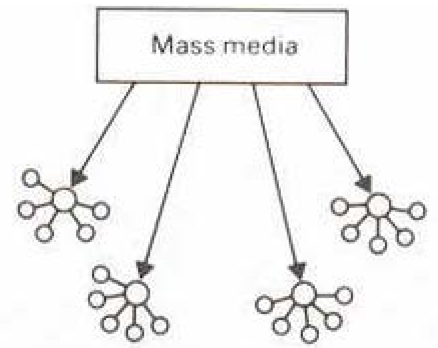
[J. Zhu 92]

The addition of any new topic onto the public agenda comes at the cost of other topic(s).

Two-step information flow

[S. Wu et al. 11]

The information reaches the masses via intermediaries (opinion leaders).



INTRO / SYSTEM / MODEL / DESIGN / CASE STUDY

Combine **quantitative modeling** and **interactive visualization**

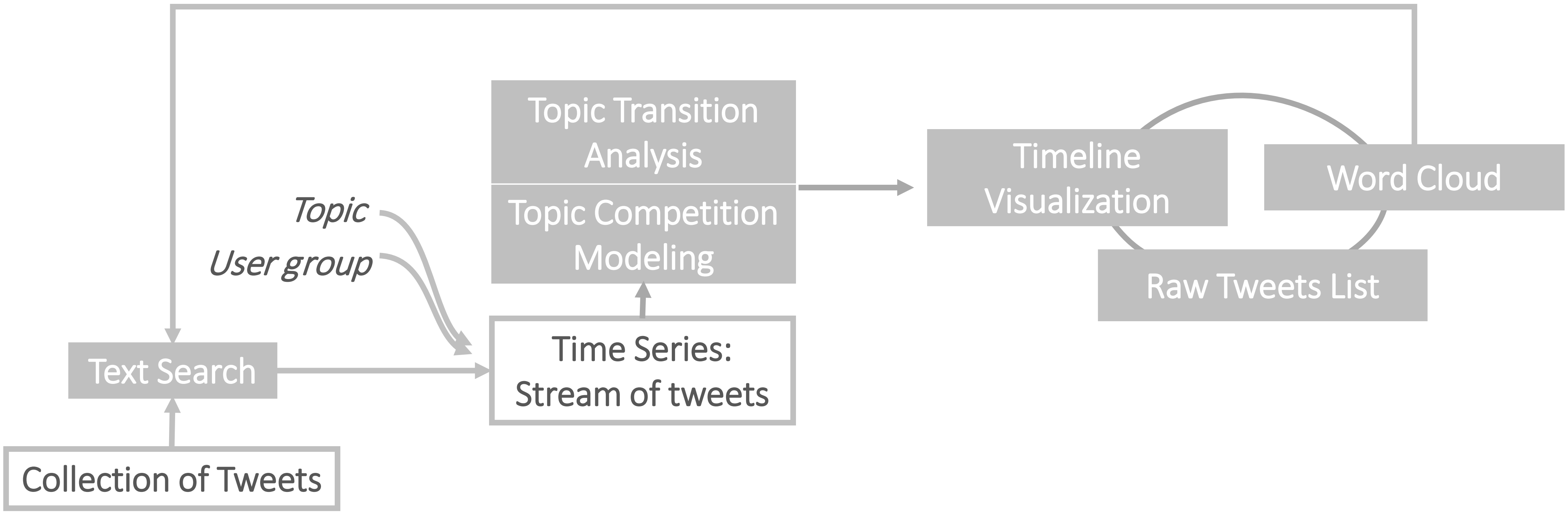
Extract time varying measurements on

- **topic competitiveness**
- **each opinion leader group's influence** on each topic
- **topic transition trend** of each opinion leader group

Visualize

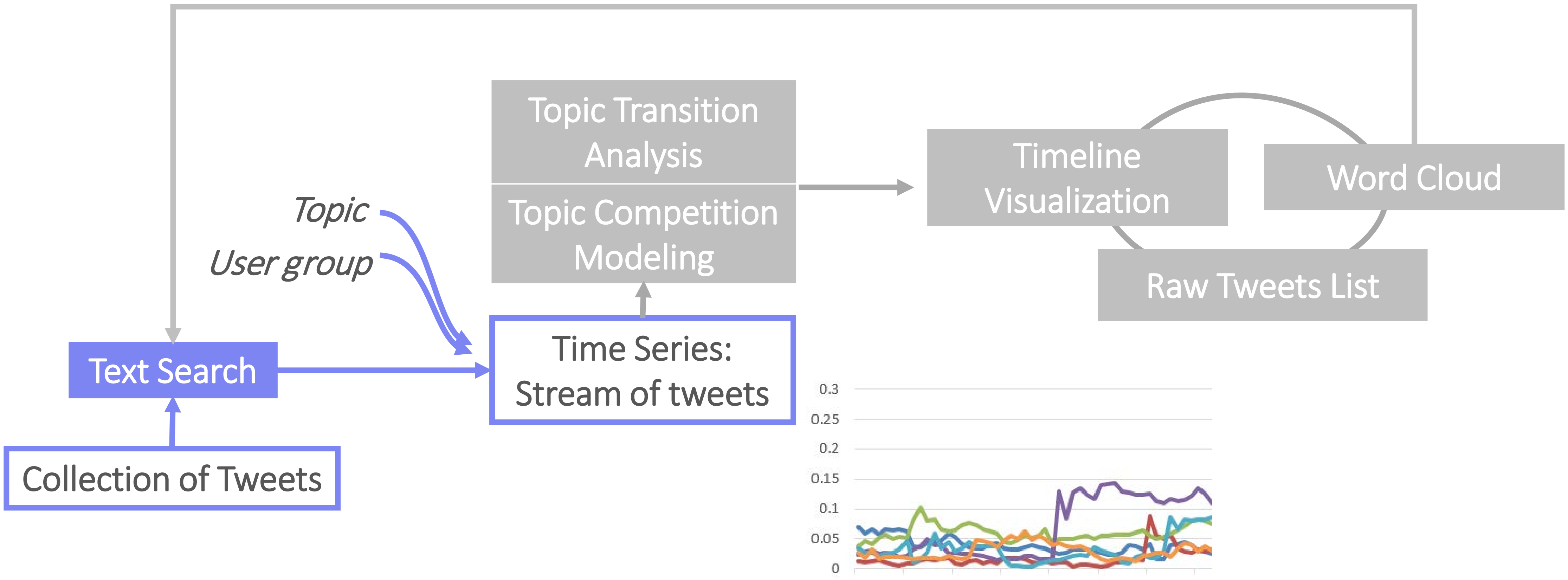
- the dynamic relation between **topics** and **opinion leader groups**
- **textual contents** of the posts

Combine **quantitative modeling** and **interactive visualization**

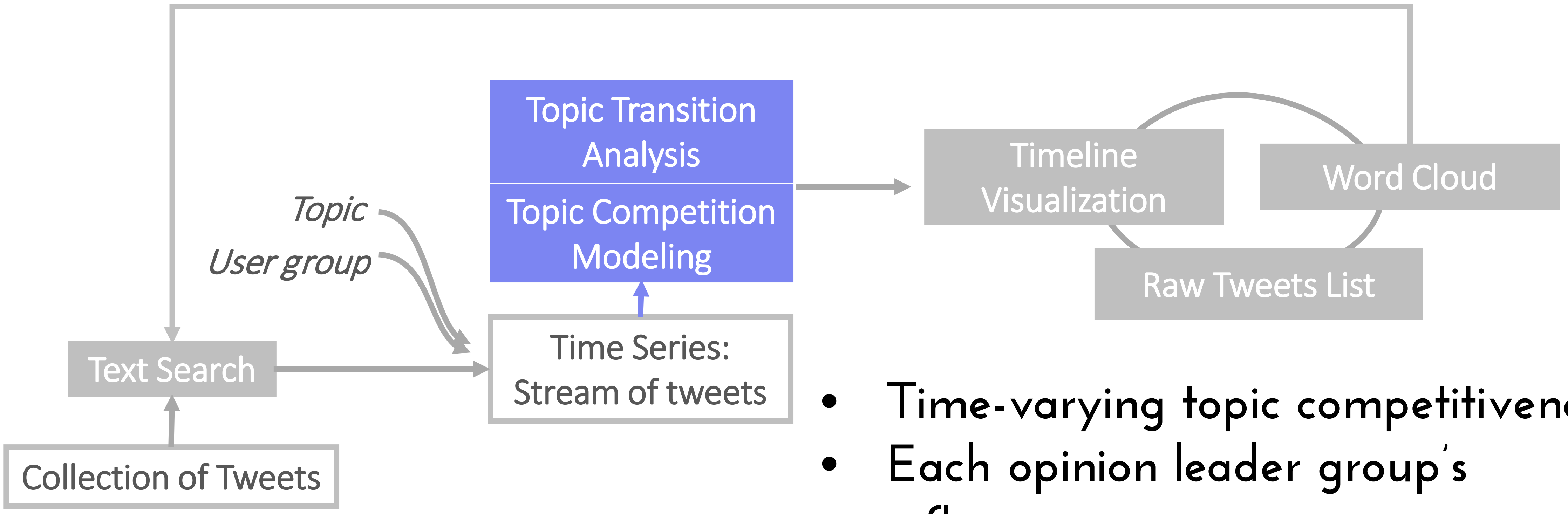


INTRO / SYSTEM / MODEL / DESIGN / CASE STUDY

Combine **quantitative modeling** and **interactive visualization**

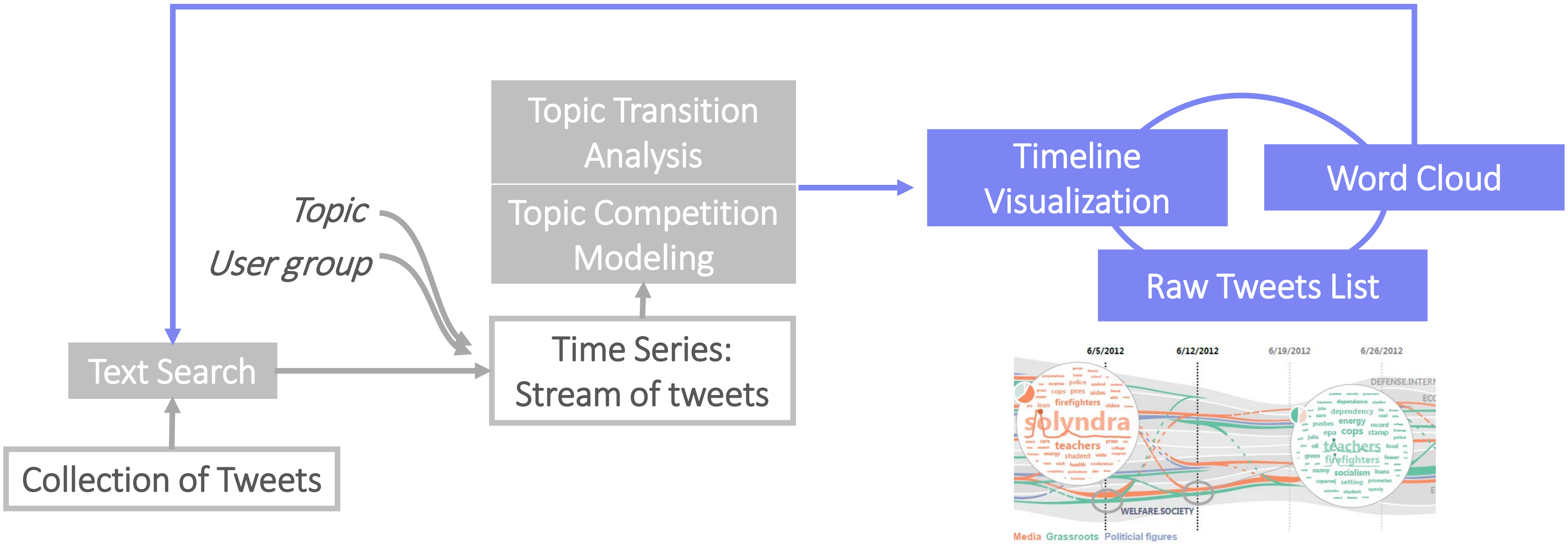


Combine **quantitative modeling** and **interactive visualization**



- Time-varying topic competitiveness
- Each opinion leader group's influence
- Topic transition trend

Combine **quantitative modeling** and **interactive visualization**



Topic Competition Model for traditional media:

change of public attention on topic i

recruiting effect

distraction effect

$$\boxed{\Delta p_i^t} = \boxed{m_i^{t-1} \sum_{j=1, j \neq i}^{k} \beta_{ij} p_j^{t-1}} - \boxed{p_i^{t-1} \sum_{j=1, j \neq i}^{k} \beta_{ji} m_j^{t-1}}$$

[J. Zhu 92]

Topic Competition Model for traditional media:

recruiting effect

change of
public attention
on topic i

$$\boxed{\Delta p_i^t} = \boxed{\boxed{m_i^{t-1}} \sum_{j=1, j\neq i}^{k} \beta_{ij} \boxed{p_j^{t-1}}} - p_i^{t-1} \sum_{j=1, j\neq i}^{k} \beta_{ji} m_j^{t-1}$$

media coverage
on *topic* i

population
on *other topic* j

[J. Zhu 92]

Topic Competition Model for traditional media:

distraction effect

change of
public attention
on topic i

$$\boxed{\Delta p_i^t} = m_i^{t-1} \sum_{j=1, j \neq i}^{k} \beta_{ij} p_j^{t-1} - \boxed{\boxed{p_i^{t-1}} \sum_{j=1, j \neq i}^{k} \beta_{ji} \boxed{m_j^{t-1}}}$$

media coverage
on *other topic* j

population
on *topic* i

[J. Zhu 92]

The Extended Topic Competition Model:

Two step information flow

Heterogeneous influence (news media, grassroots)

$$\Delta p_i^t = \boxed{m_i^{t-1}} \sum_{j=1, j\neq i}^{k} \beta_{ij} p_j^{t-1} - p_i^{t-1} \sum_{j=1, j\neq i}^{k} \beta_{ji} m_j^{t-1}$$

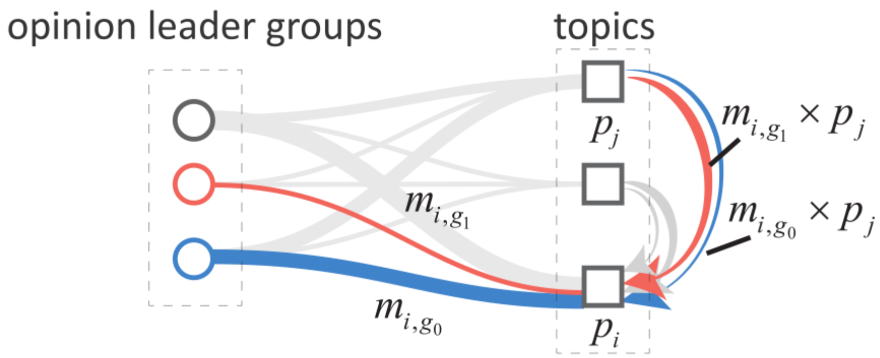
$$\boxed{\sum_{g=1}^{n} m_{i,g}^{t-1}}$$

INTRO / SYSTEM / MODEL / DESIGN / CASE STUDY

The Extended Topic Competition Model:

Two step information flow

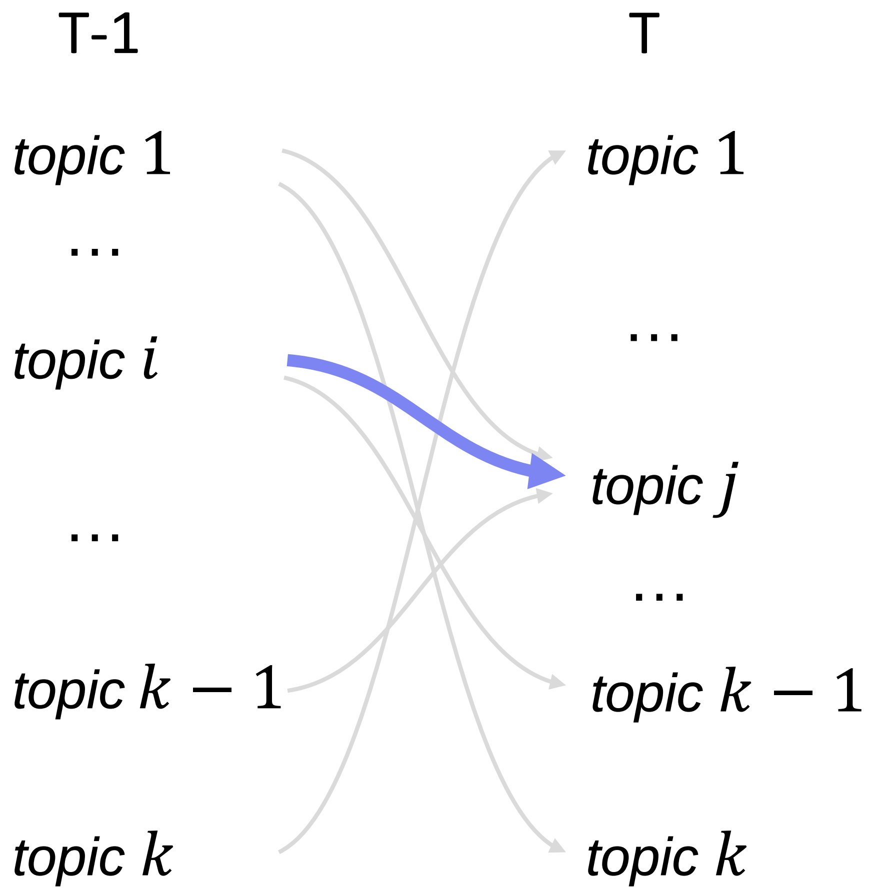
Heterogeneous influence (news media, grassroots)



$$p_i^t = a_i p_i^{t-1} + \underbrace{\sum_{g=1}^{n} m_{i,g}^{t-1} \sum_{j=1, j\neq i}^{k} \beta_{i,j,g} p_j^{t-1}}_{\text{recruiting effect}} - \underbrace{p_i^{t-1} \sum_{j=1, j\neq i}^{k} \sum_{g=1}^{n} \beta_{j,i,g} m_{j,g}^{t-1}}_{\text{distraction effect}}$$

Topic competiveness & opinion leader's influence through R^2 decomposition

Topic Transition Estimation



Transition matrix $A_{k\times k} = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \cdots & \cdots & \cdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}$

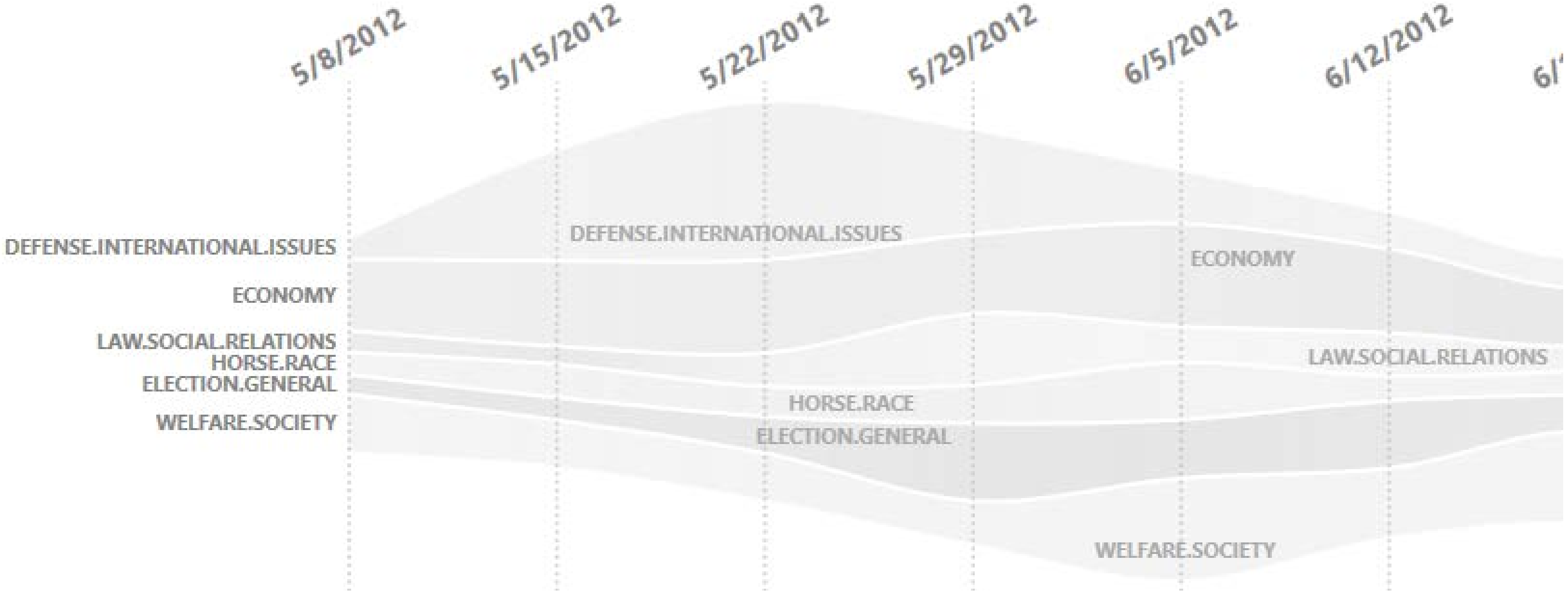
$$\min \sum_{l} \omega_l \left\| m_l^{t-1} A - m_l^t \right\|^2$$

$$subject\ to: \sum_{j=1}^{k} a_{ij} = 1\ and\ a_{ij} \geq 0$$

Output of Analysis and Modeling Step:

Time varying **competitiveness** of each topic
Time varying **opinion leader groups' influence** on each topic
The **topic transition trend** of the opinion leader groups between adjacent time stamps.

Topic competitiveness

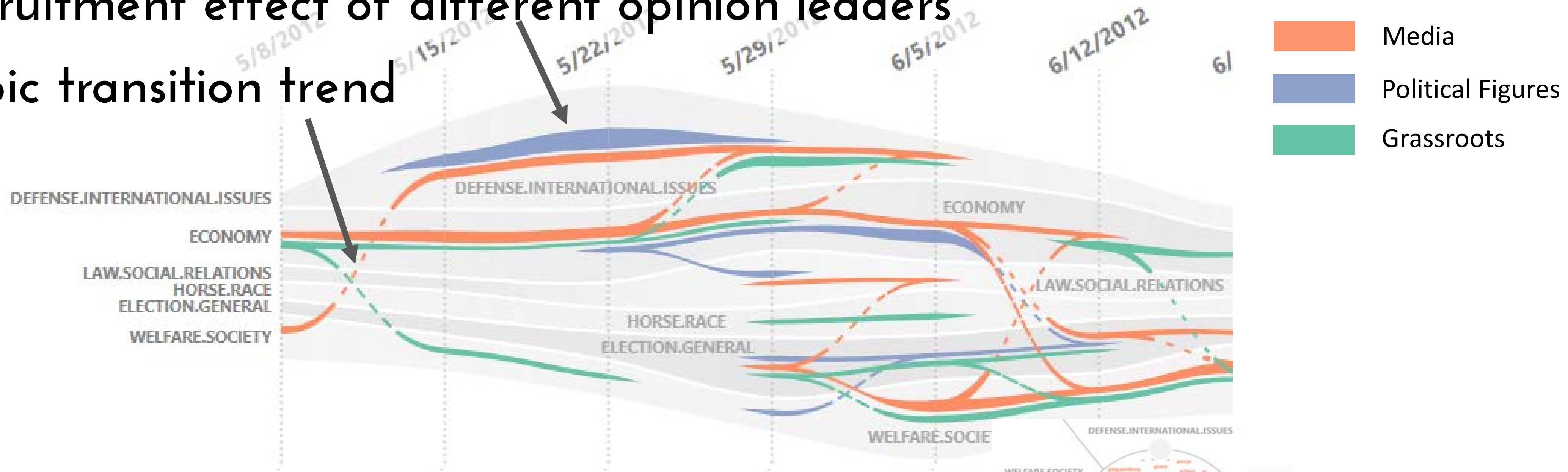


Timeline view

INTRO / SYSTEM / MODEL / **DESIGN** / CASE STUDY

Topic competiveness

+ Recruitment effect of different opinion leaders
+ Topic transition trend



Timeline view

Word Cloud

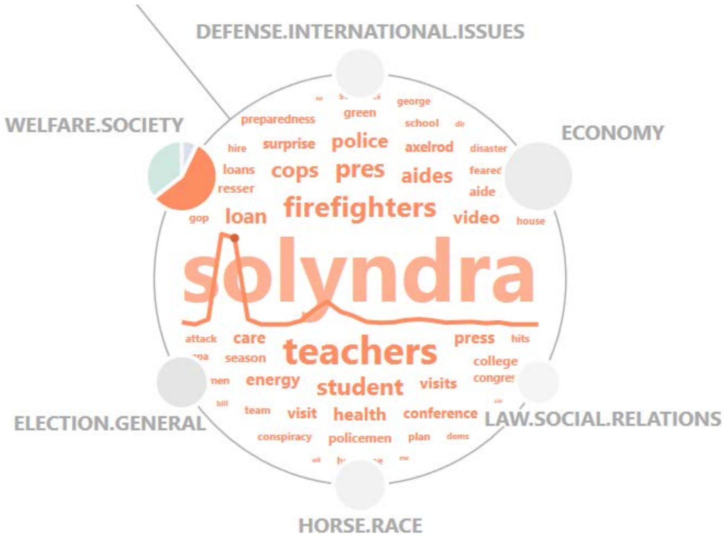
INTRO / SYSTEM / MODEL / DESIGN / CASE STUDY

Word cloud filterable by:

- Topic
- Time interval
- Opinion leader group

Sparkline:

- Time varying saliency of a word

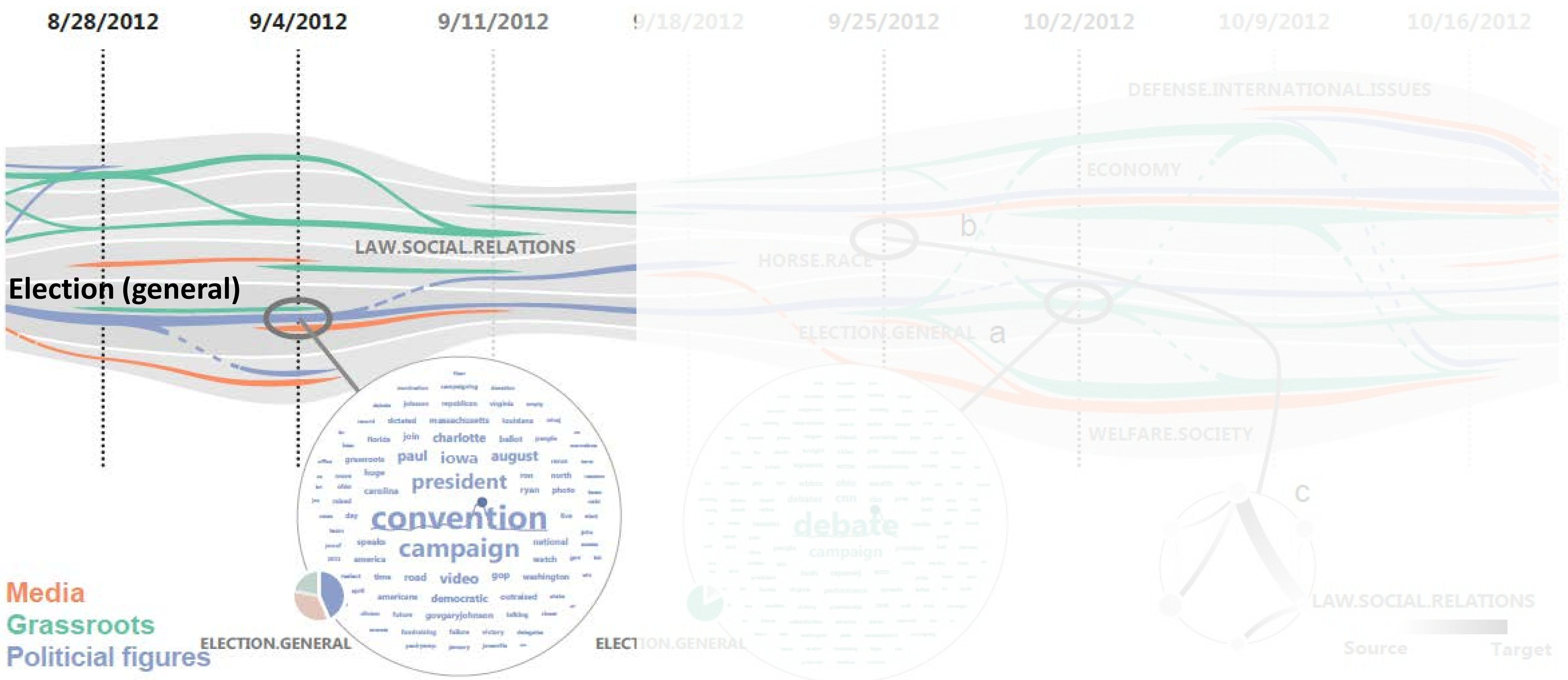


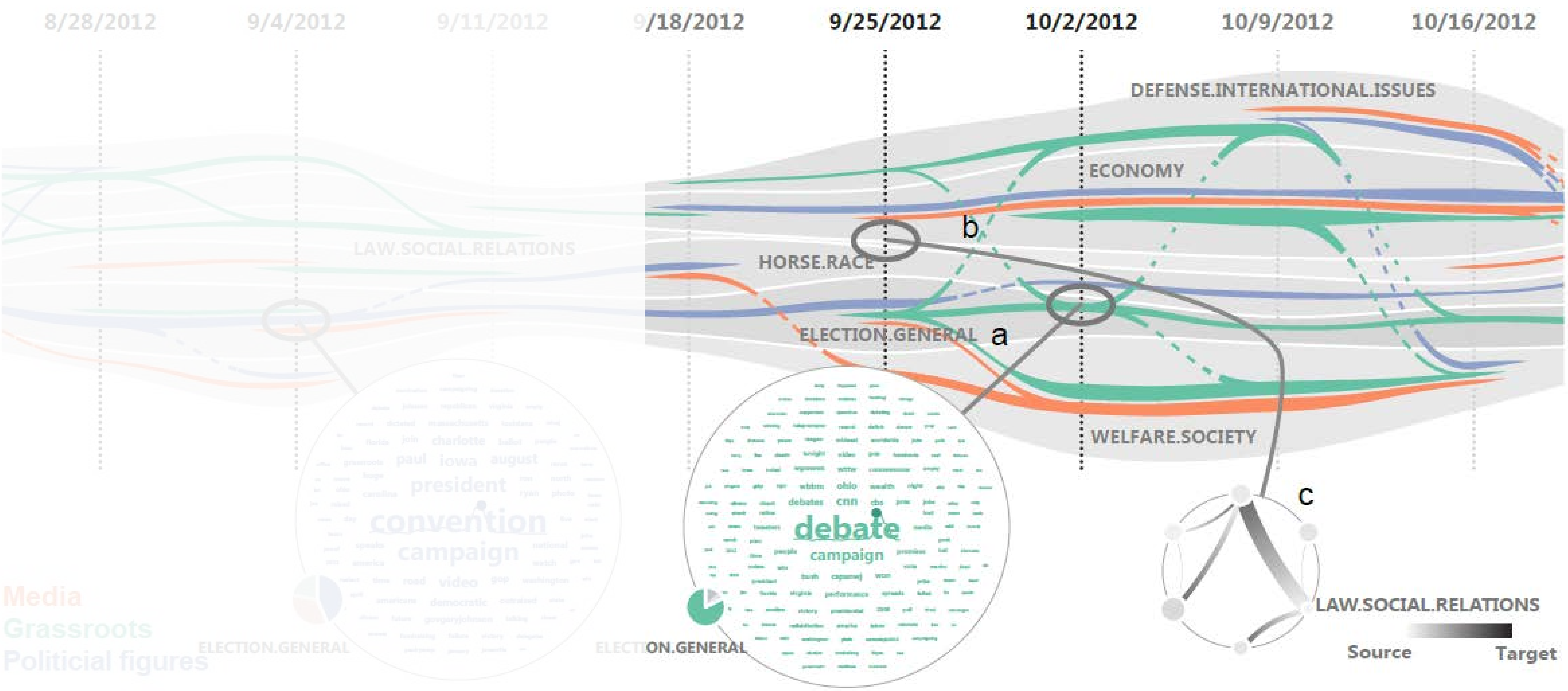
Dataset:

2012 Presidential Election; **89, 174, 308** tweets; May 01 – Nov 10
6 general topics : welfare/society, defense/international issues, economy, election (general), election (horse race), law/social relations *
3 opinion leader groups: media , political figures, and grassroots *

*identified collaboratively with media researchers

INTRO / SYSTEM / MODEL / DESIGN / CASE STUDY


8/28/2012
9/4/2012
9/11/2012
9/18/2012
9/25/2012
10/2/2012
10/9/2012
10/16/2012
DEFENSE.INTERNATIONAL.ISSUES
ECONOMY
LAW.SOCIAL.RELATIONS
HORSE.RACE
b
Election (general)
ELECTION.GENERAL
a
WELFARE.SOCIETY
c
convention
campaign
president
Media
Grassroots
Politicial figures
ELECTION.GENERAL
ELECTION.GENERAL
debate
LAW.SOCIAL.RELATIONS
Source
Target



INTRO / SYSTEM / MODEL / DESIGN / CASE STUDY

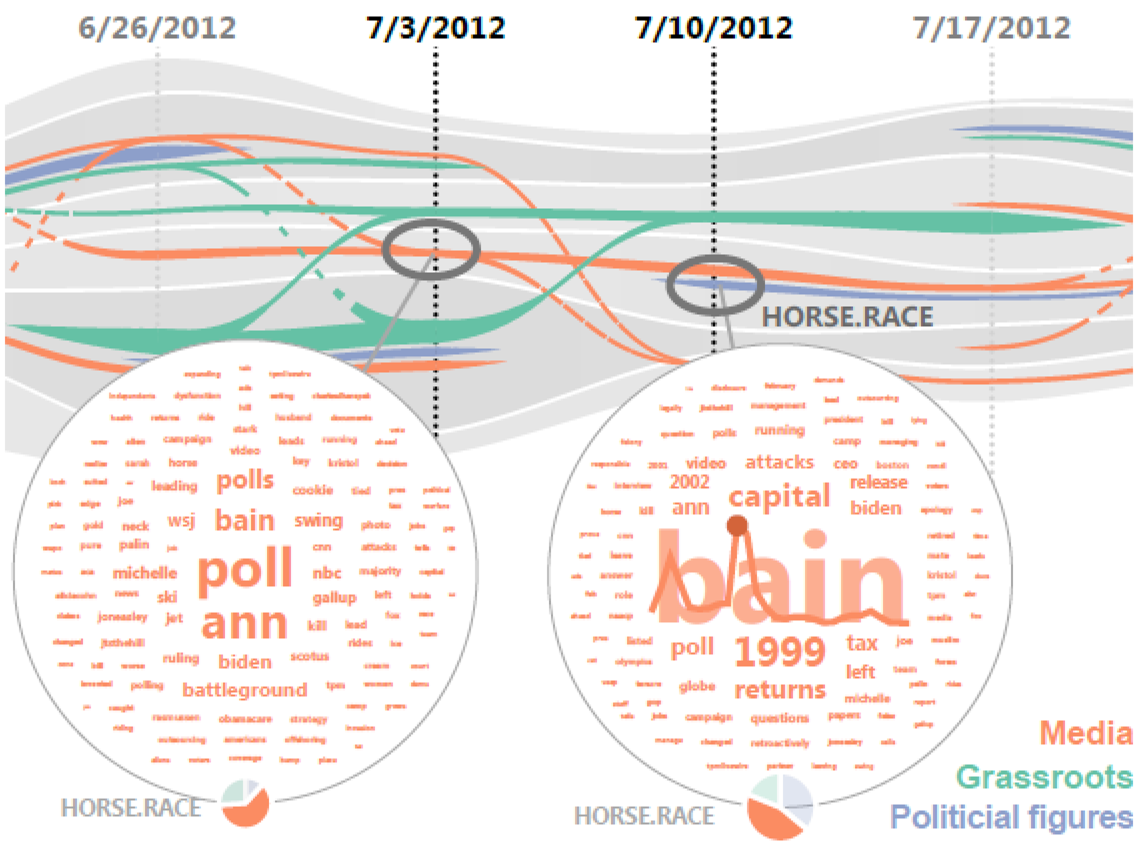


Fig. 6. A long time of influence exerted by the *media* on the topic *election (horse race)*, although with very different trending keywords.

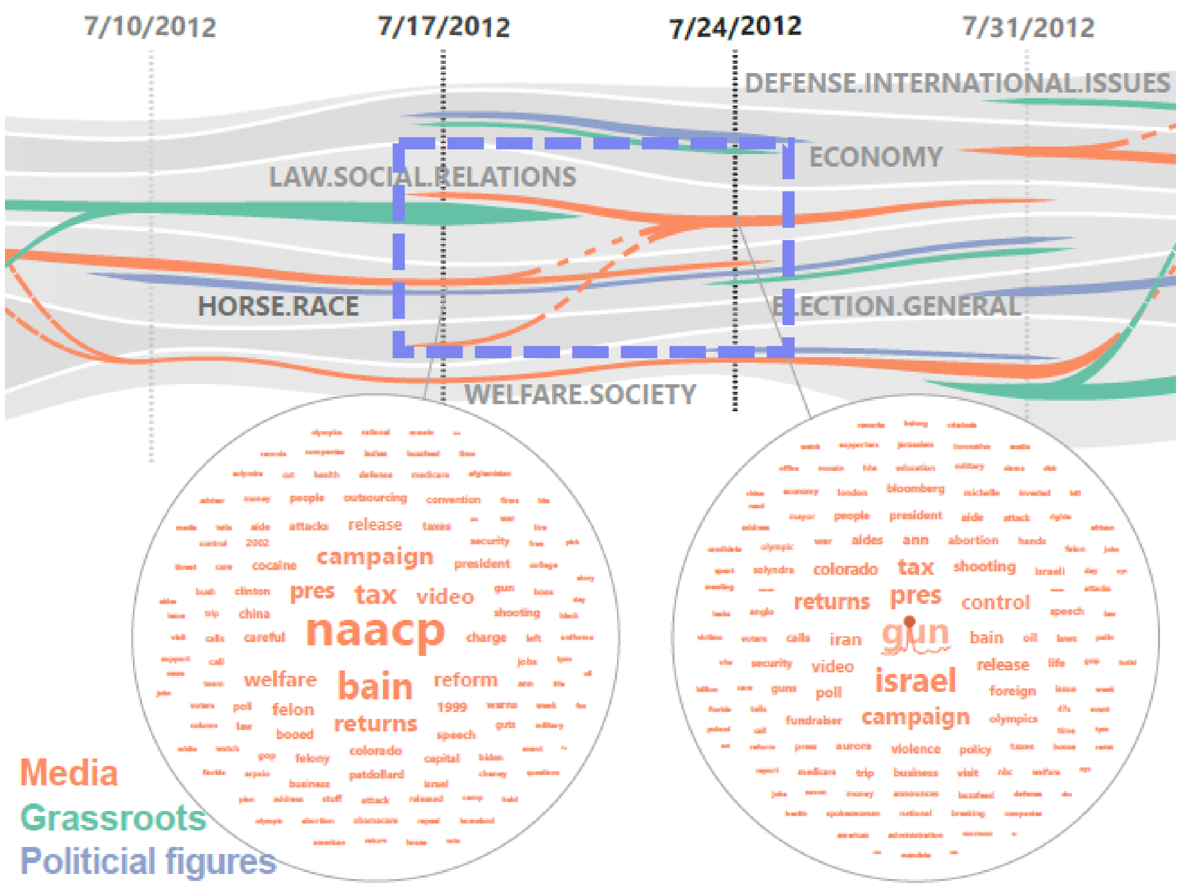


Fig. 7. Transition of topical focus of the *media* from multiple topics to *law / social relations* around July 24th. The keyword "gun" had an increasing importance when the word clouds based on all the tweets posted by the *media* were compared before and after the transition.

INTRO / SYSTEM / MODEL / DESIGN / CASE STUDY

Visual analysis framework:

Model the topic competition on social media, the influence of opinion leader groups, and the topic transition trends.

Visualize the results of the models and allow for further exploration to form explanations.

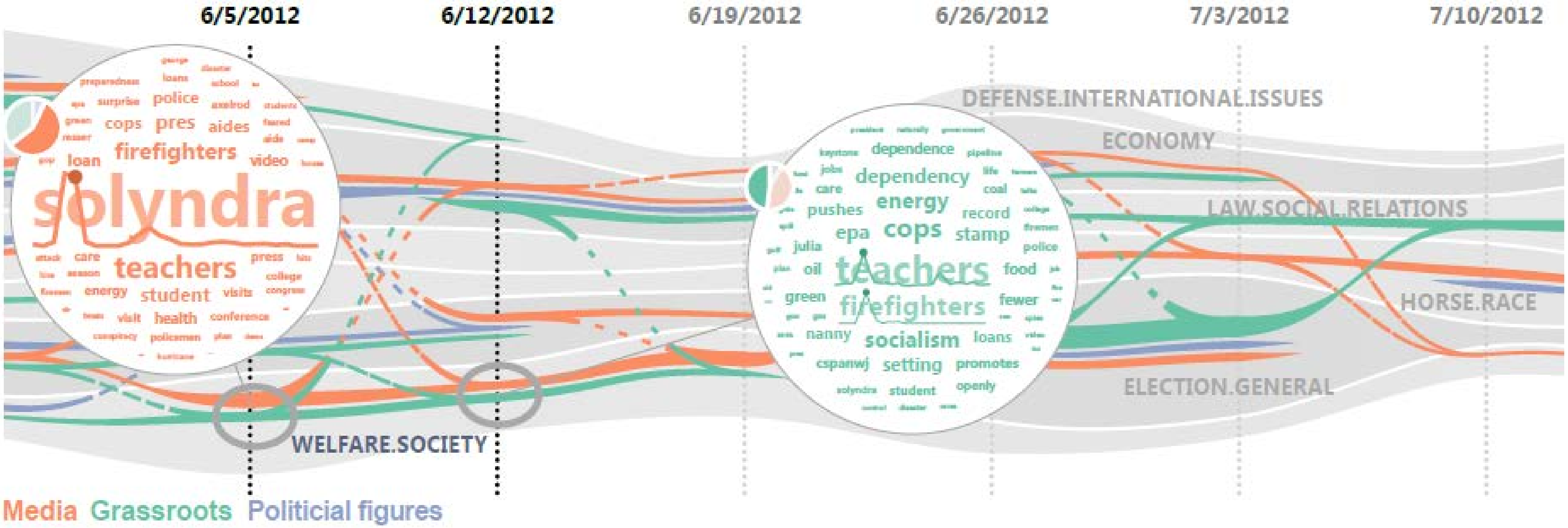
SUMMARY / LIMITATIONS & FUTURE WORK

Manual process to collect keywords and categorize opinion leaders
more efficient ways?

Time series modeling
+ the structural factors of social network ?

Competition & cooperation
other modes of interaction among topics?

SUMMARY / LIMITATIONS & FUTURE WORK



Thank You for Attention !

Model Validation

Table 1. Evaluation of the model against three common measures in time series data analysis shows that the model is highly effective and robust. The table shows the average and the standard deviation (in parentheses) of the measures when applying a moving window estimation for the *2012 presidential election data*.

	Economy	Horse Race	Election General
R^2	0.98 *(0.01)*	0.98 *(0.01)*	0.99 *(0.00)*
$se_{\hat{y}}$	0.02 *(0.006)*	0.02 *(0.008)*	0.02 *(0.006)*
DW - d	2.13 *(0.18)*	2.17 *(0.18)*	2.14 *(0.16)*

	Defense / International	Law / Social Relations	Welfare & Society
R^2	0.97 *(0.02)*	0.96 *(0.03)*	0.95 *(0.04)*
$se_{\hat{y}}$	0.02 *(0.006)*	0.01 *(0.008)*	0.02 *(0.012)*
DW - d	2.18 *(0.18)*	2.11 *(0.17)*	2.11 *(0.18)*

Data Processing

Opinion leaders: defined by number of retweets. 200 users are selected for Election data.

Keywords: collected iteratively