



ARTS1422 Data Visualization

Lecture 8

Cross-Media Visualization

Quan Li
Spring 2024
2024. 03.21



OUTLINE

1

Text Visualization

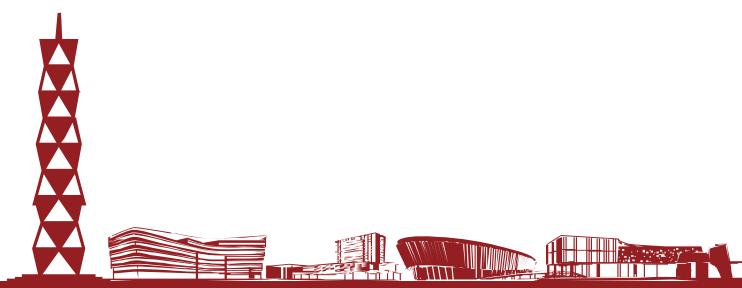
2

Multimedia Visualization



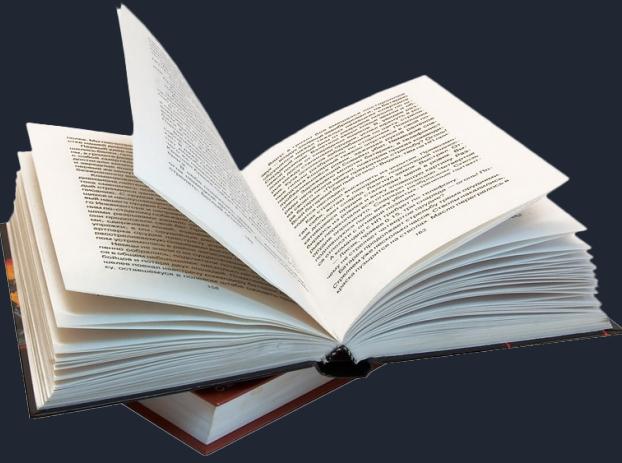


Text Visualization



Text Data

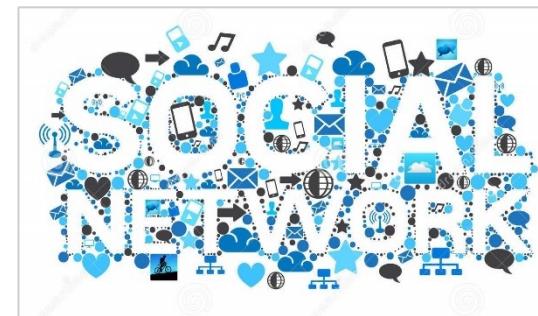
- Books.
- News.
- Web pages.
-



Possible Applications



E-Discovery



Social Computing



Business Intelligence



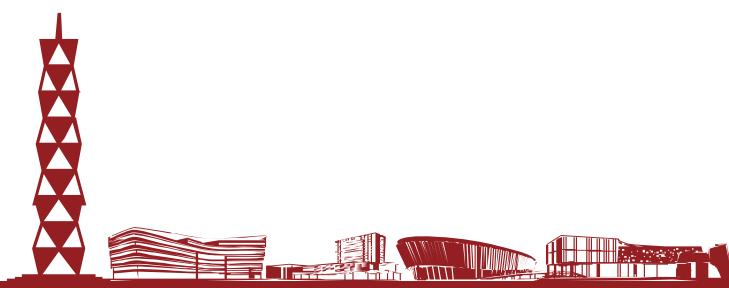
Customer Experience



Predictive Analysis



Public Relation



Tasks



How to summarize:

Content.

Topics.



Useful for:

Analyzers.

Researchers.



```
mirror_mod = modifier_obj
# mirror object to mirror
mirror_mod.mirror_object

operation == "MIRROR_X":
    mirror_mod.use_x = True
    mirror_mod.use_y = False
    mirror_mod.use_z = False
operation == "MIRROR_Y":
    mirror_mod.use_x = False
    mirror_mod.use_y = True
    mirror_mod.use_z = False
operation == "MIRROR_Z":
    mirror_mod.use_x = False
    mirror_mod.use_y = False
    mirror_mod.use_z = True

selection at the end -add
    _ob.select= 1
    mirror_ob.select=1
    context.scene.objects.active
    ("Selected" + str(modifier))
    mirror_ob.select = 0
    bpy.context.selected_objects
    ta.objects[one.name].sele
    print("please select exactly one object")
    - OPERATOR CLASSES ---
```

```
types.Operator):
    X mirror to the selected ob
    ject.mirror_mirror_x"
    or X"
context):
    context.active_object is not
```

Levels in Text Information

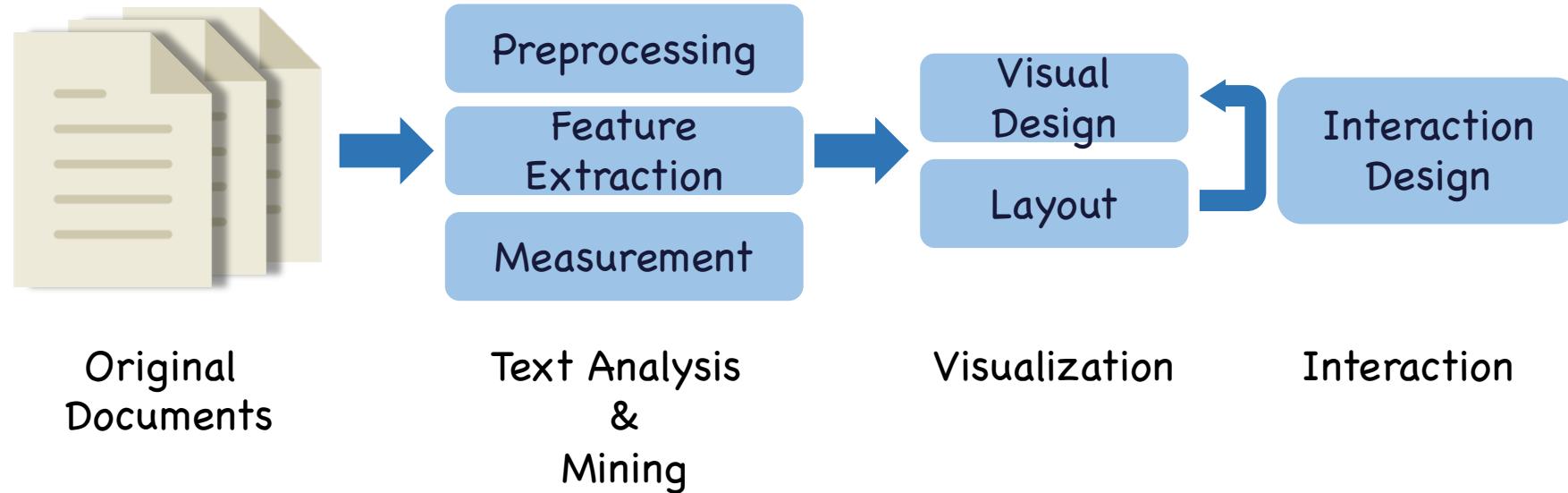
- Based on understanding levels:
 - Lexical Level: basic unit of text.
 - Syntactic Level: structure information.
 - Semantic Level: meaning of text contents.

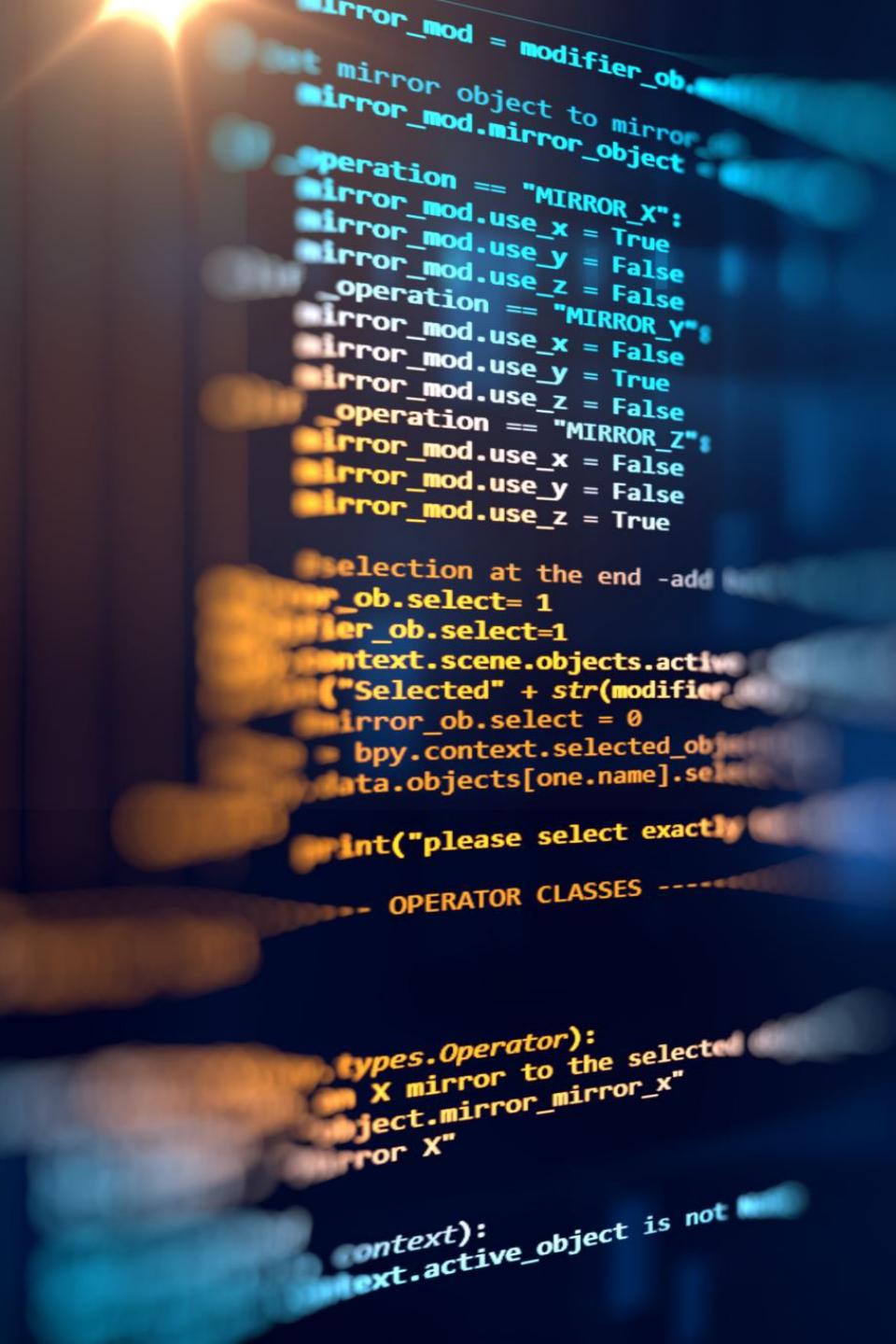


Research Direction

- Type of text:
 - Single document.
 - Document collection.
 - Temporal documents.
- Features of text:
 - Content.
 - Structure.
 - Multi-level information.

Process





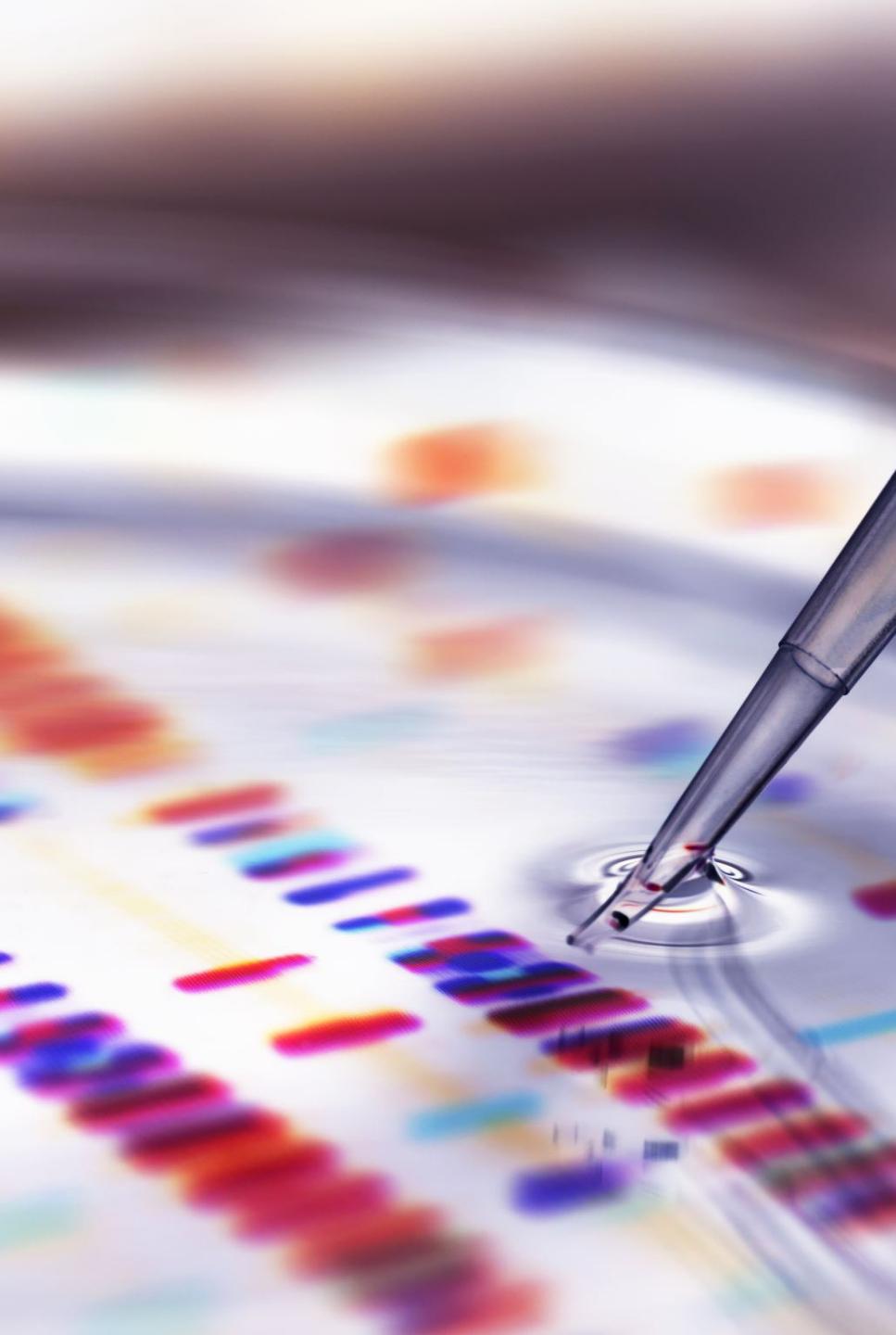
Process: Text Analysis & Mining

- Preprocessing:
 - Text wrangling.
- Feature Extraction:
 - Keywords.
 - Word frequency.
 - Topics.
- Measurement:
 - Document similarity
 - Clustering



Process: Visualization

- Visual Design:
 - Chernoff Face.
 - Flow metaphor.
 - ...
- Layout:
 - Force-directed.
 - Treemap
 - Scatterplot
 - ...



Process: Interaction

- Linking & Brushing:
 - Brushing: focusing.
 - Linking: highlighting.
- Zooming
- Focus + Content
- Filtering



Text Analysis Methods

- Tokenization
- Vector-space Model
- Topic retrieving

Tokenization



"I have a dream that one day this nation will rise up and live out the true meaning of its creed: We hold these truths to be self-evident, that all men are created equal."



- Remove stop words: a, the, that, etc.
- Plural form to singular form: men->man, truths->truth



I, dream, one, day, nation, rise, up, live, out, true, meaning, creed, hold, truth, be, self-evident, all, man, created, equal





Vector-space Model

- Bag-of-words

Word	I	dream	color	skin	nation	slave	injustice	owner
Frequency	4	4	1	1	2	2	1	1

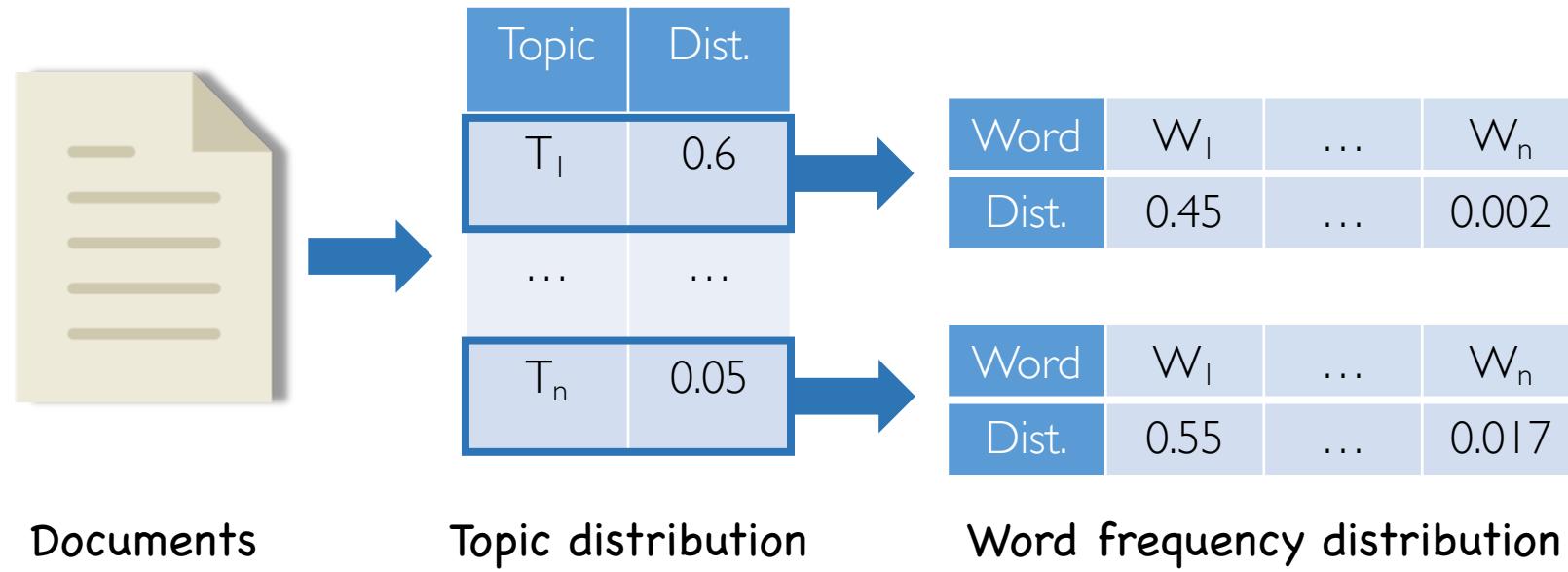
- Document similarity: compute similarity between two word-frequency vectors

- Cosine: $\cos \theta = \frac{\mathbf{v}_1 \mathbf{v}_2^T}{\|\mathbf{v}_1\| \cdot \|\mathbf{v}_2\|}$

- TF-IDF

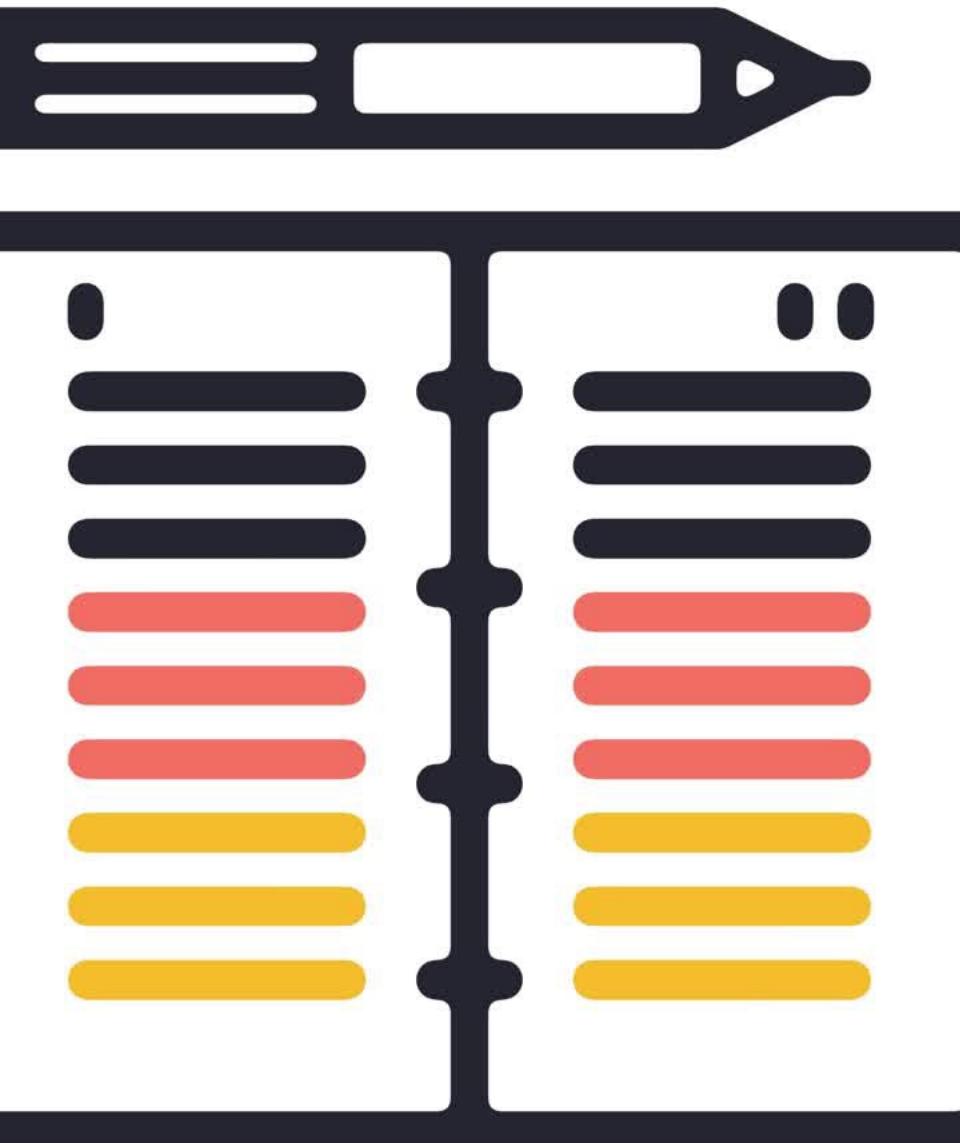


Topic Retrieving



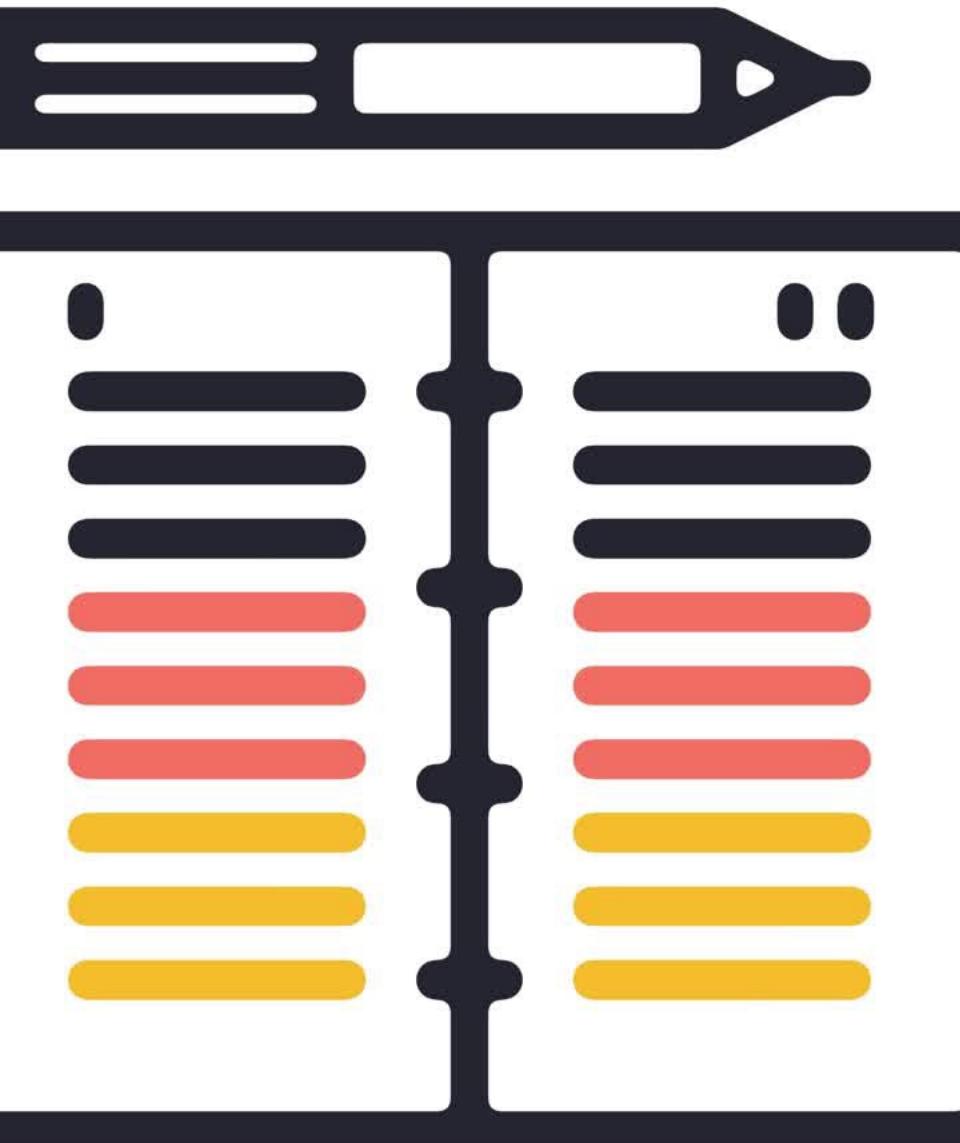
- Approaches:
 - Latent Semantic Indexing
 - pLSI
 - LDA





Text Visualization

- Categorized based on features
 - Visualizing Document Content
 - Visualizing Document Relations
 - Multi-level Document Visualization



Text Visualization

- Categorized based on features
 - Visualizing Document Content
 - Visualizing Document Relations
 - Multi-level Document Visualization

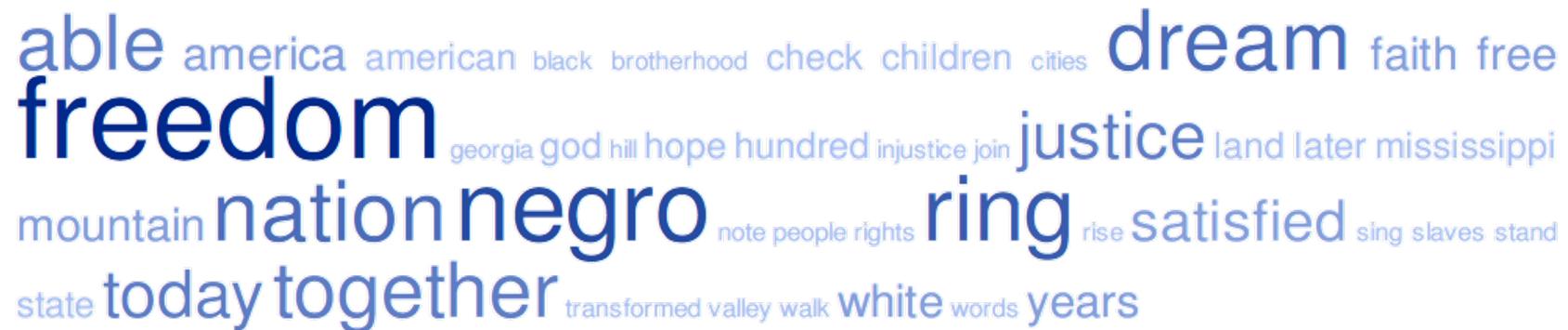
Visualizing Document Content

- Keyword-based Visualization
- Temporal Document Visualization
- Feature Distribution Visualization
- Opinion Analysis
- Query Visualization
- Software Visualization



Visualizing Document Content

- Tag Cloud/Text Cloud/Word Cloud
 - Retrieve keywords and arranged in a pattern
 - Encoding: color and size



Tag cloud of "I have a dream".



Visualizing Document Content

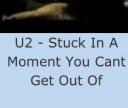
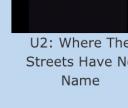
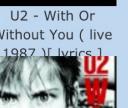
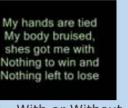
Music
Artist
Cloud

Enter artist here ...

Passengers [Bono](#) [R.E.M.](#) [INXS](#) [Bruce Springsteen](#) [The Police](#) [Simple Minds](#) [Coldplay](#) [Noel Gallagher's High Flying Birds](#) [Bryan Adams](#) [Sting](#) [Peter Gabriel](#) [Pearl Jam](#) [The Cranberries](#) [The Rolling Stones](#) [Phil Collins](#) [Paul McCartney](#) [Robert Plant](#) [Lenny Kravitz](#) [Oasis](#) [Aerosmith](#) [Tom Petty](#) [Bon Jovi](#) [Sinéad O'Connor](#) [Tom Petty and The Heartbreakers](#) [David Bowie](#) [Crowded House](#) [Midnight Oil](#) [Alanis Morissette](#) [Dire Straits](#) [Bruce Springsteen & The E Street Band](#) [Guns N' Roses](#) [Queen](#) [Elton John](#) [The Killers](#)

U2
[Read biography](#)

01. With Or Without You
02. Pride
03. I'm
04. Beautiful Day
05. Vertigo
06. Every Breaking Wave
07. Where The Streets Have No Name
08. Sweetest Thing
09. Walk On
10. Elevation
11. Electrical Storm
12. Sunday Bloody Sunday
13. Insomniac
14. I Still Haven't Found What I'm Looking For
15. Song For Someone
16. Bad
17. The Miracle
18. New Years Day
19. Ultra Violet
20. If I Go Crazy If I Don't Go Crazy Tonight



U2 Greatest Hits Full Album - The Best of U2 - U2 Love Songs Ever

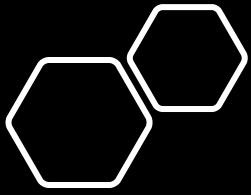
Learn more about U2 on Wikipedia
[Get details on many albums by U2 at Discogs](https://en.wikipedia.org/wiki/U2)

<https://musicartistcloud.appspot.com/>

Music Artist Cloud
version 6.2 php
©mac 2017

Camaris Dev Team
[contact us](#)

powered by

Derivation: Wordle

- Aesthetic Modification
 - Customable shape
 - Encoding: font style, size, color
 - Tightly arranged

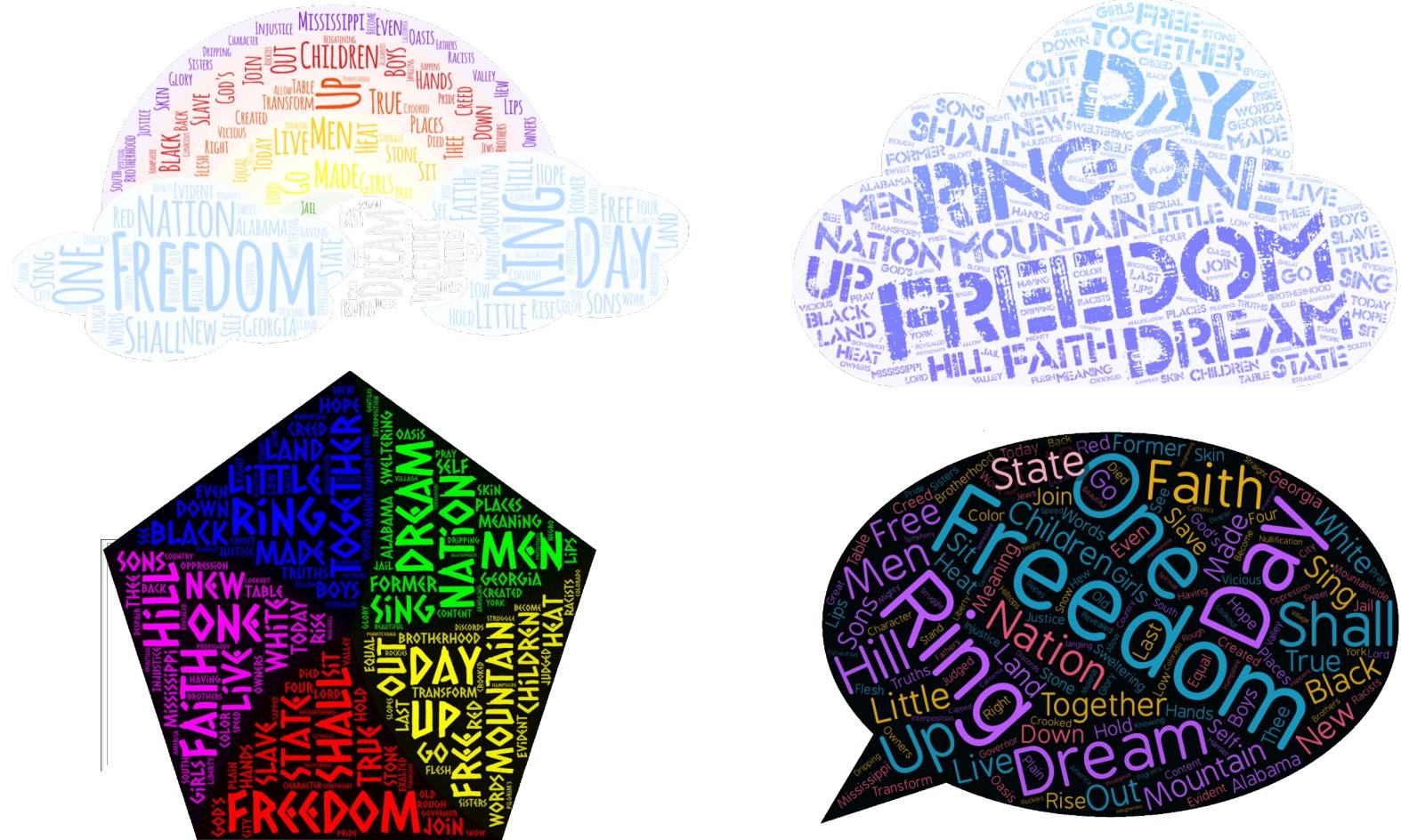


Wordle of “I have a dream”.

Created by: <https://wordart.com/create>



Wordles

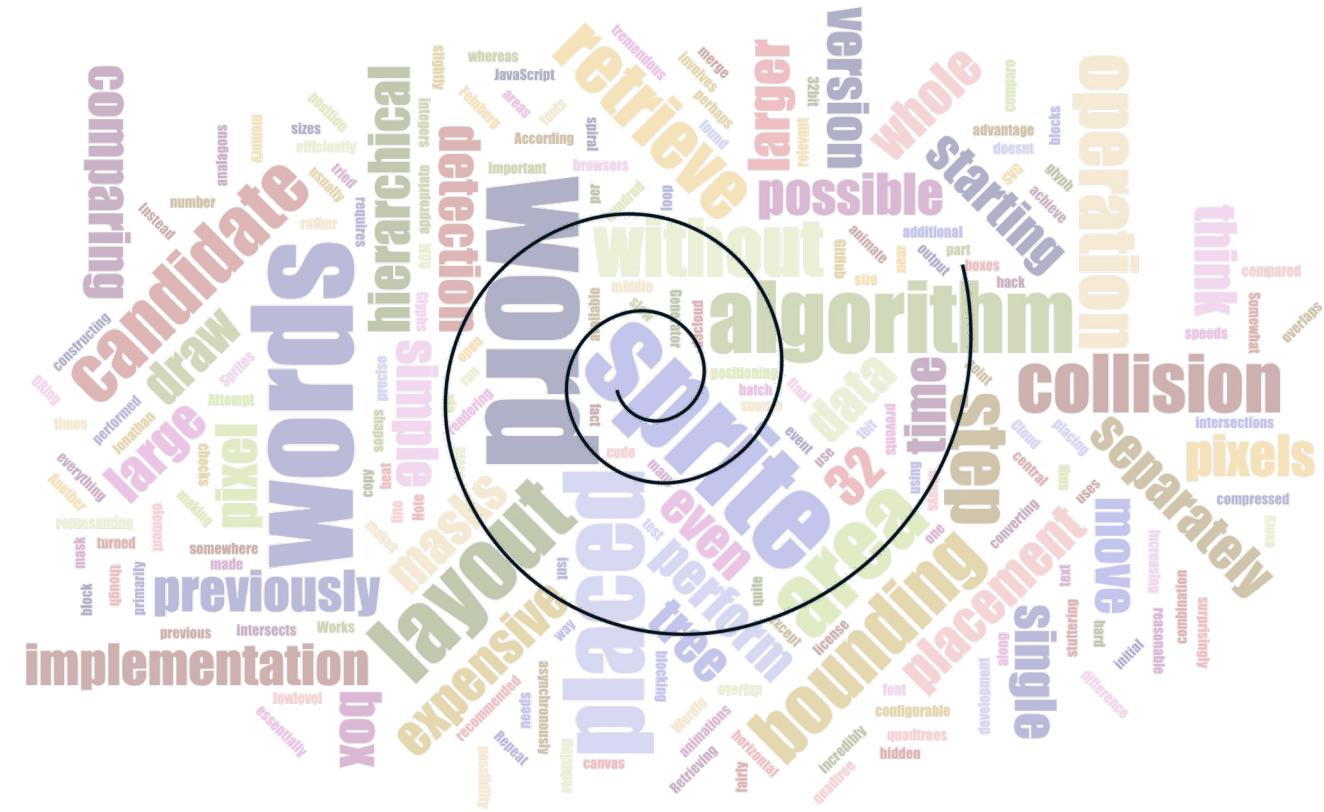


Other wordle of "I have a dream".
Created by: <https://wordart.com/create>



Layout

- Spiral layout



Created by: <https://www.jasondavies.com/wordcloud/>



Layout

- Horizontal and vertical layout



Created by: <https://www.jasondavies.com/wordcloud/>



Layout

- Random layout



Created by: <https://www.jasondavies.com/wordcloud/>



Layout

- Shape layout



Created by: <https://www.jasondavies.com/wordcloud/>



Context-preserving Word Cloud

 = Appearing  = Disappearing  = Unique



The importance criterion

 = Appearing  = Disappearing  = Unique



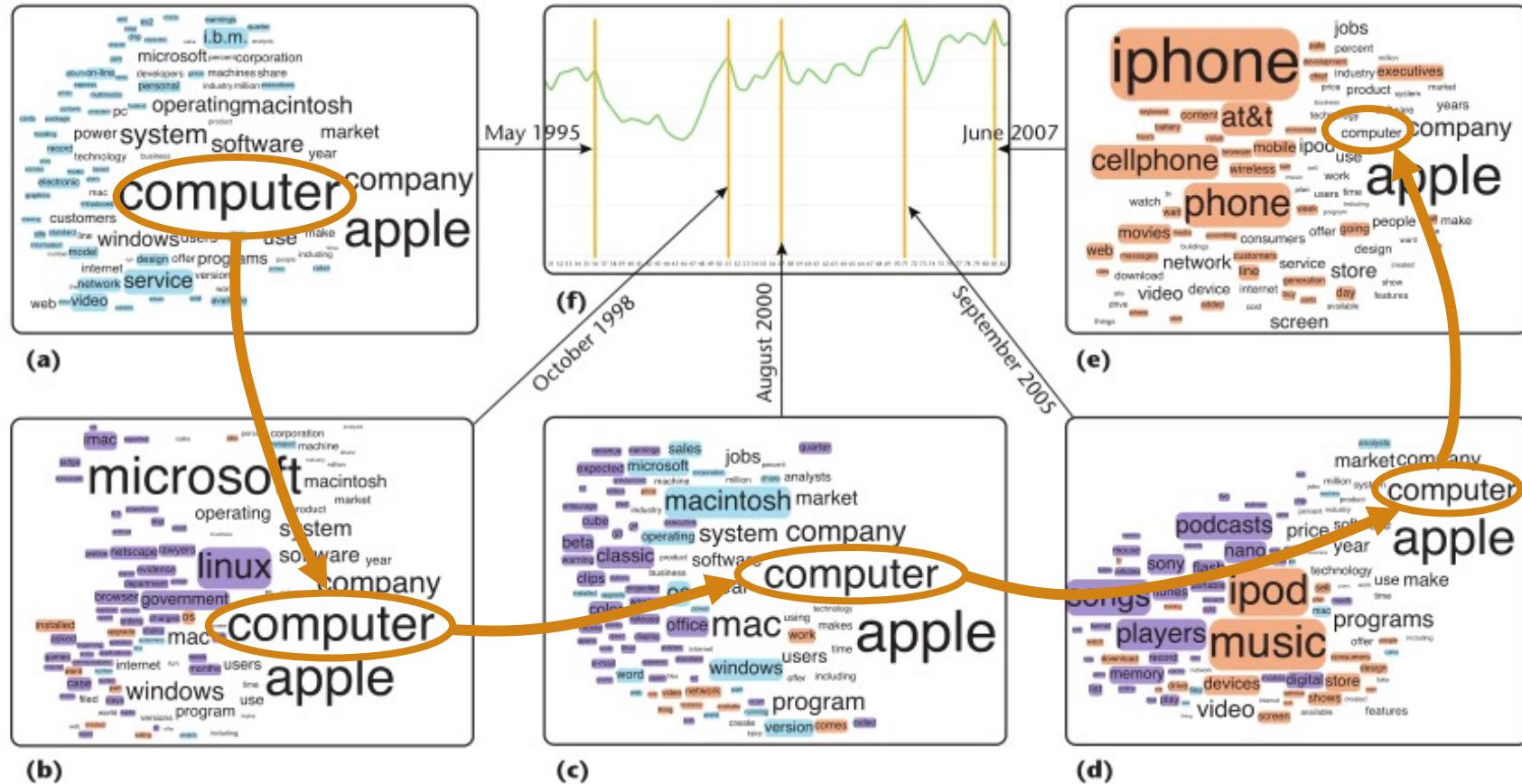
The co-occurrence criterion

Weiwei Cui et al. Context preserving dynamic word cloud visualization. (IEEE PacificVis 2010)





Context-preserving Word Cloud

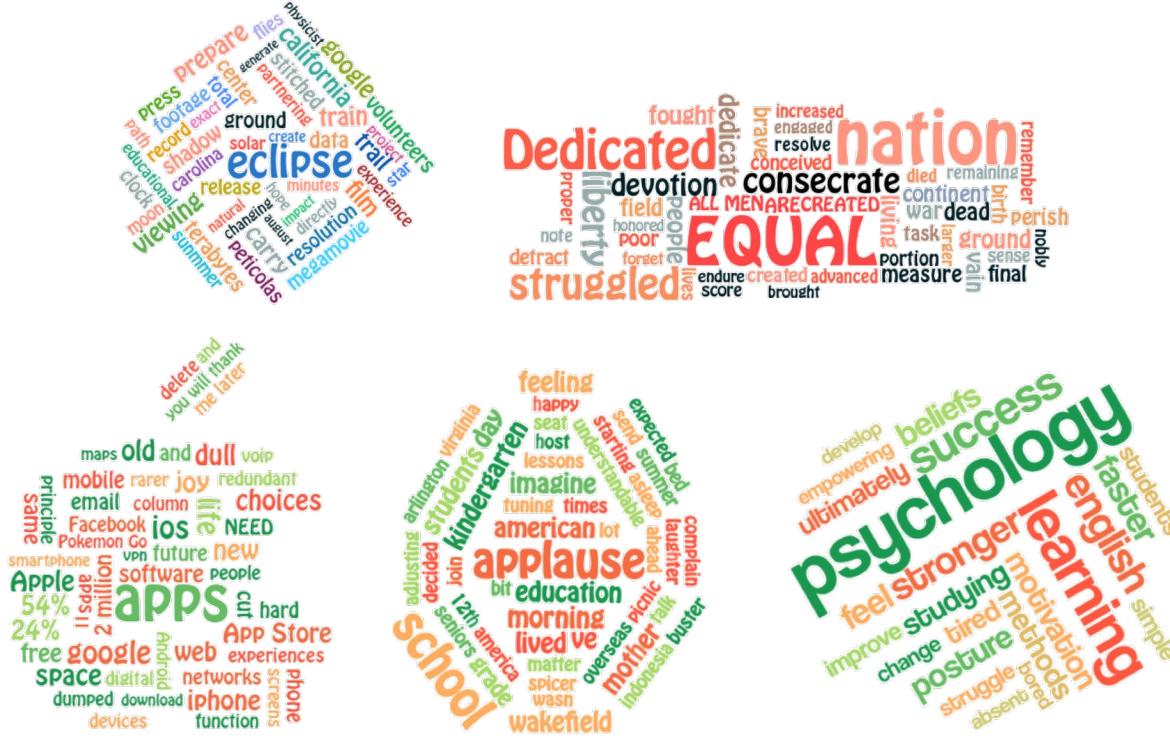


Weiwei Cui et al. Context preserving dynamic word cloud visualization. (IEEE PacificVis 2010)





Consistency-preserving Word Cloud



Word cloud of five articles

INFOVIS PAPER

EdWordle: Consistency-preserving Word Cloud Editing

yunhai wang, Xiaowei Chu, Chen Bao, Lifeng Zhu, Oliver Deussen,
Baoquan Chen, Michael Sedlmair



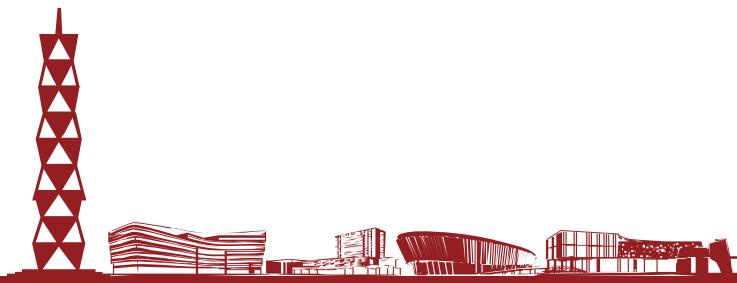
1-6 October 2017
Phoenix, Arizona, USA

ieeevis.org



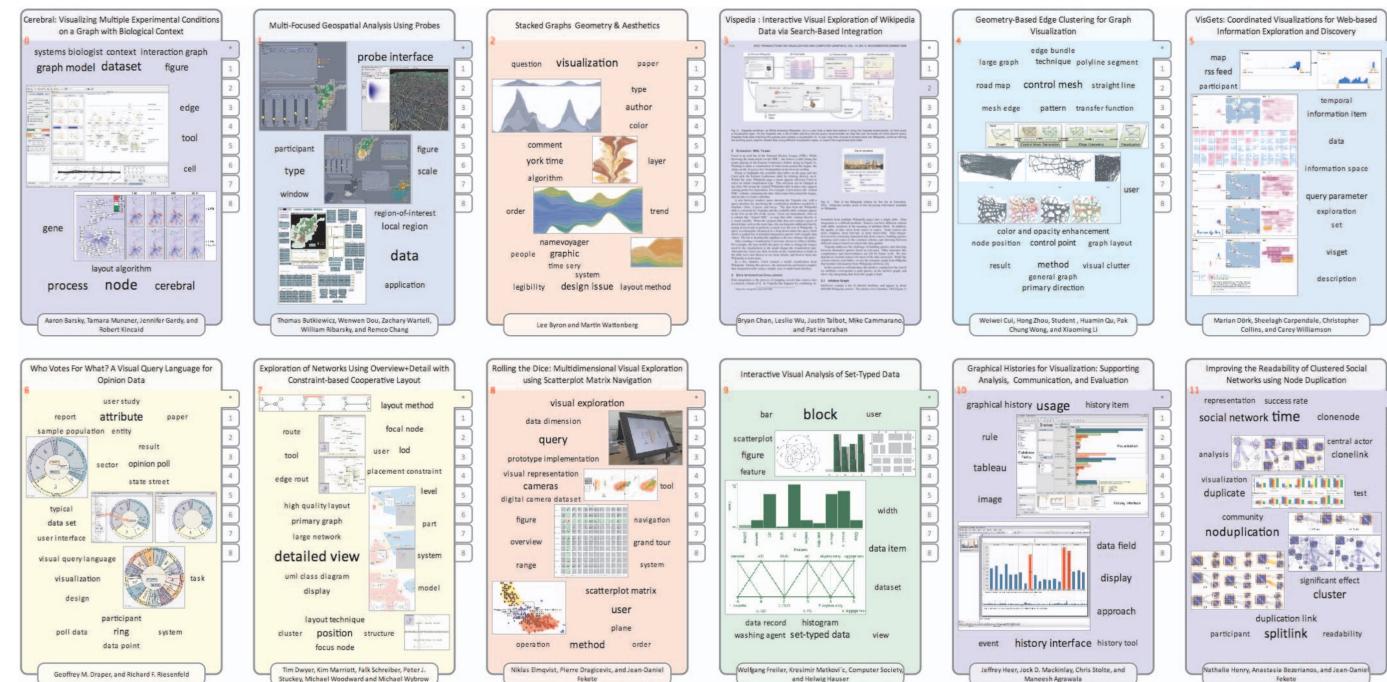
Generating Wordle

- Sort words by frequency or importance (encoded by font size)
- Specify layout rule:
 - Spiral
 - Horizontal and vertical
 - Random
 - User-specified shape
- Remember the shape (if necessary)



Document Card

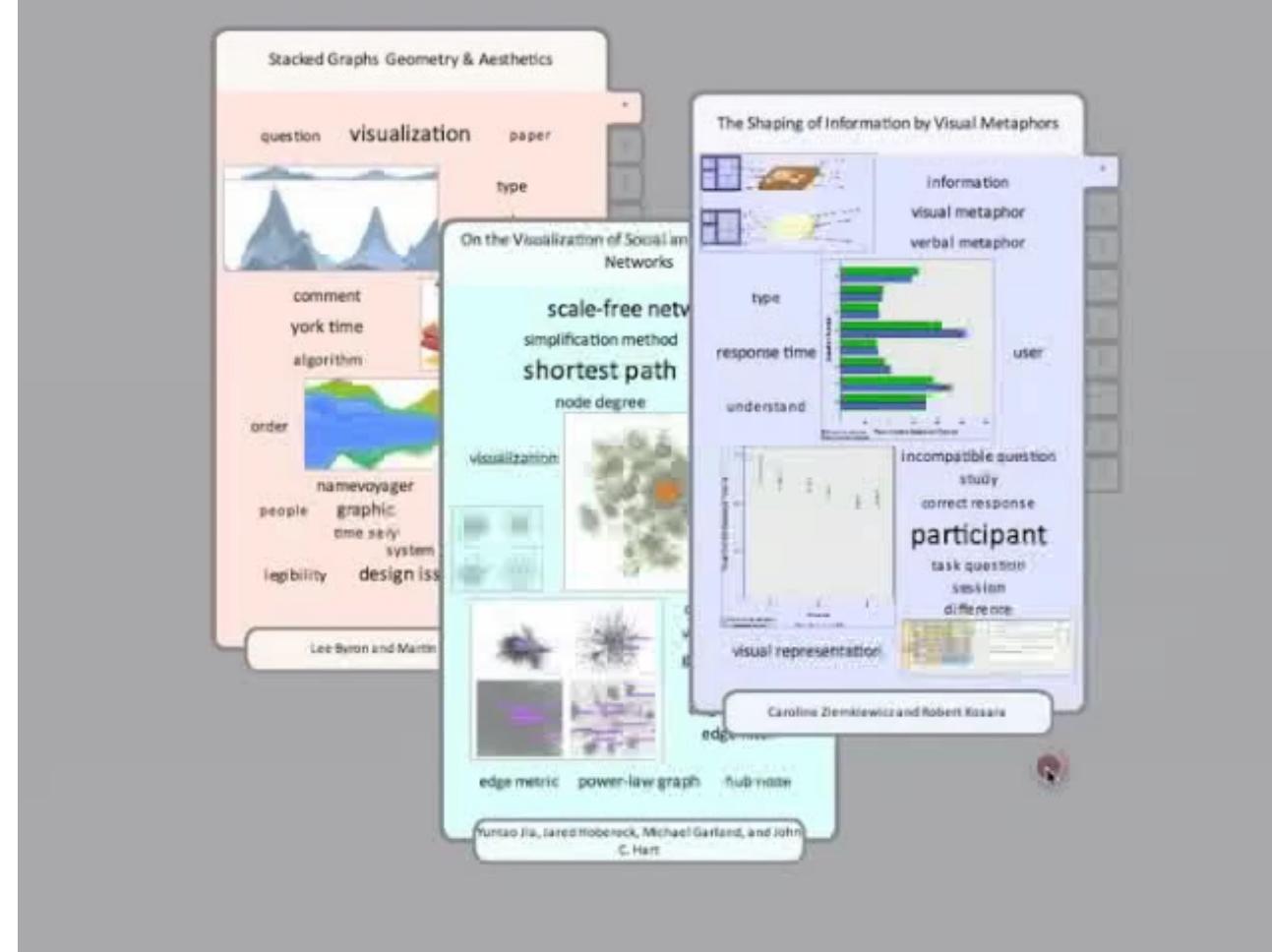
- Summarize a document with keywords and images in the document
- Layout in a card-like space



Hendrik Strobelt et al. Document Cards: A Top Trumps Visualization for Documents. (IEEE TVCG 2009)



Document Card



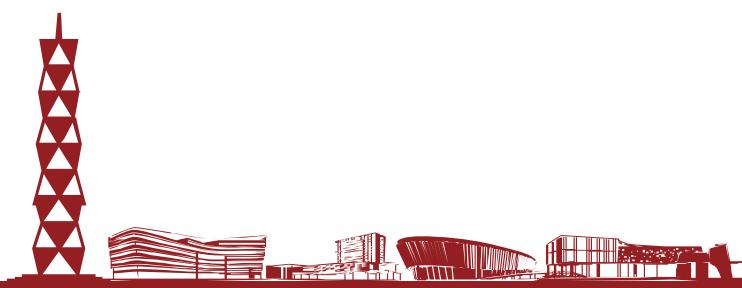
Hendrik Strobelt et al. Document Cards: A Top Trumps Visualization for Documents. (IEEE TVCG 2009)

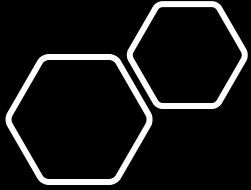
立志成才报国裕民



Visualizing Document Content

- Keyword-based Visualization
- Temporal Document Visualization
- Feature Distribution Visualization
- Opinion Analysis
- Query Visualization
- Software Visualization



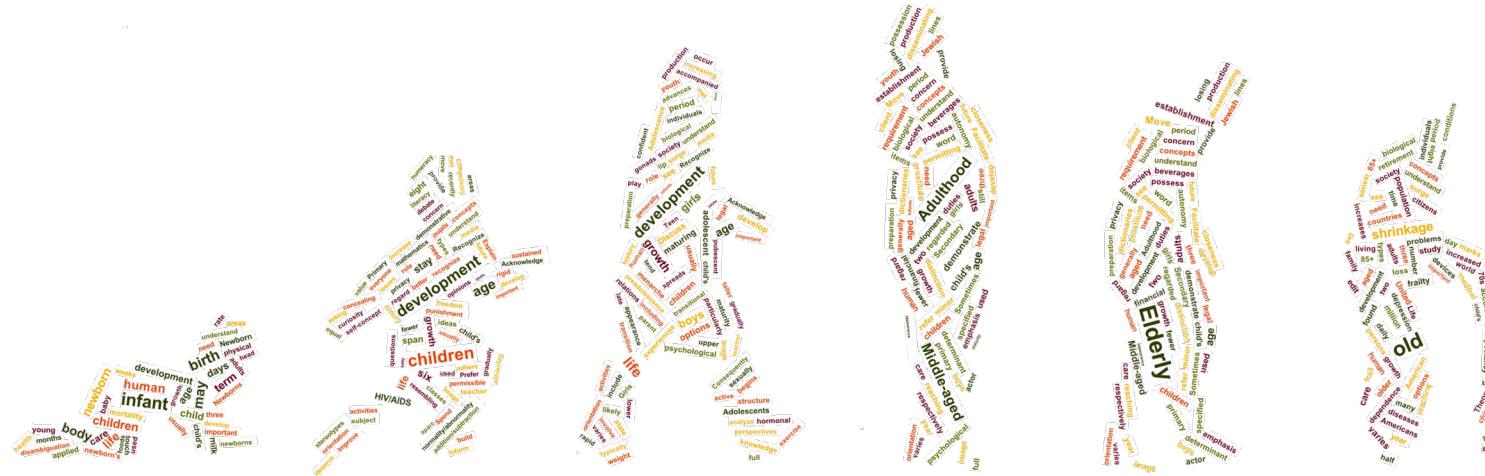


Temporal Document Visualization

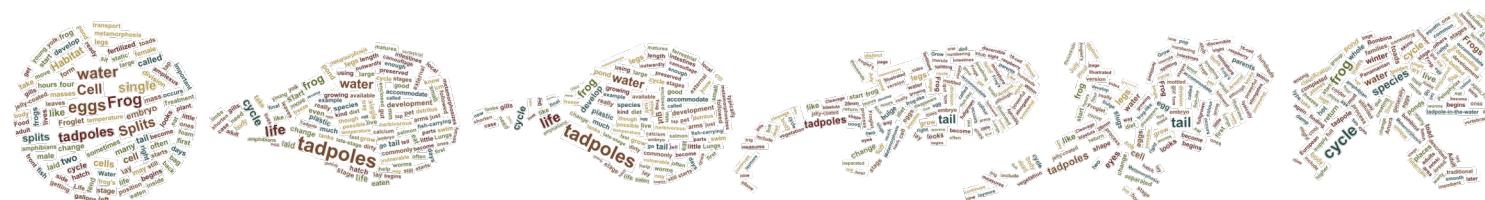
- Temporal/Ordinal Documents
- Revolution feature:
 - Plot of a novel
 - Events in news reports



Morphable Word Cloud



The stages of the human life cycle (from infancy to late adulthood)

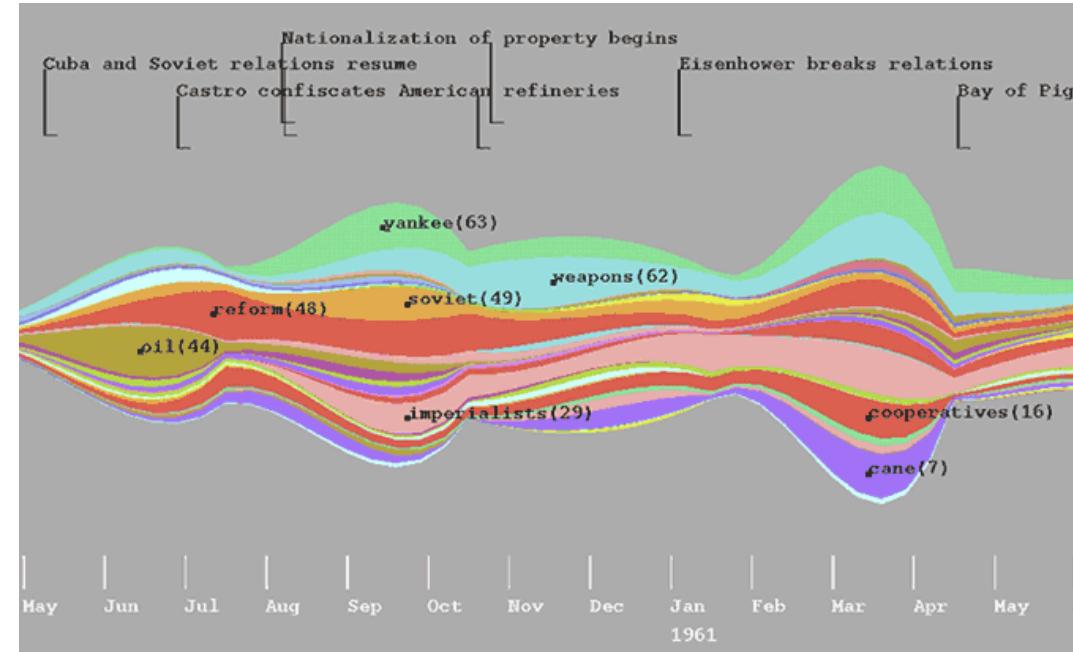


The frog life cycle

Ming-Te Chi et al. Morphable Word Clouds for Time-varying Text Data Visualization. (IEEE TVCG 2015)



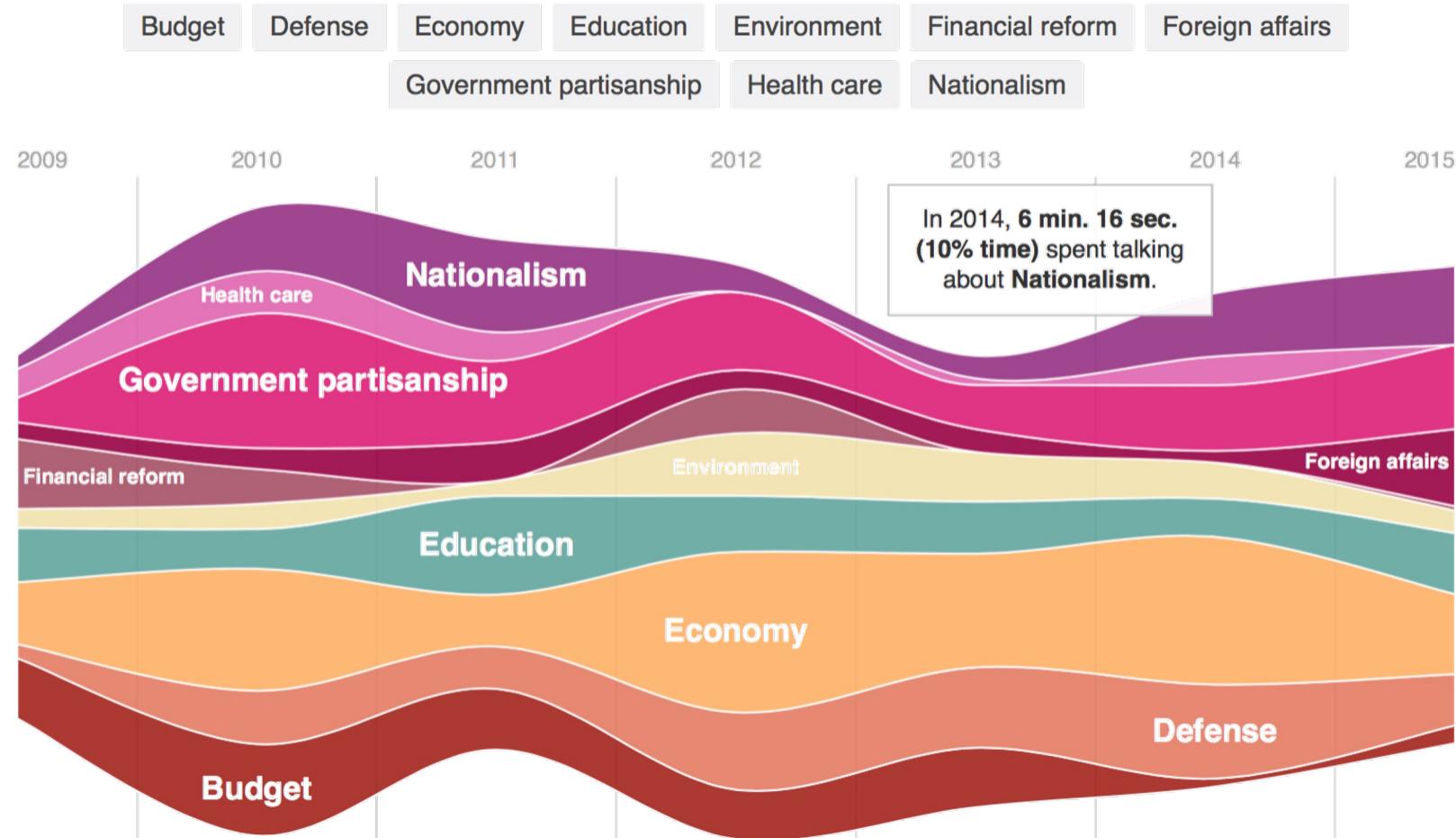
- ThemeRiver
 - Macroscopic topics
 - Metaphor: river -> temporal topic



Susan Havre et al. ThemeRiver: Visualizing Theme Changes over Time. (IEEE SIV 2000)

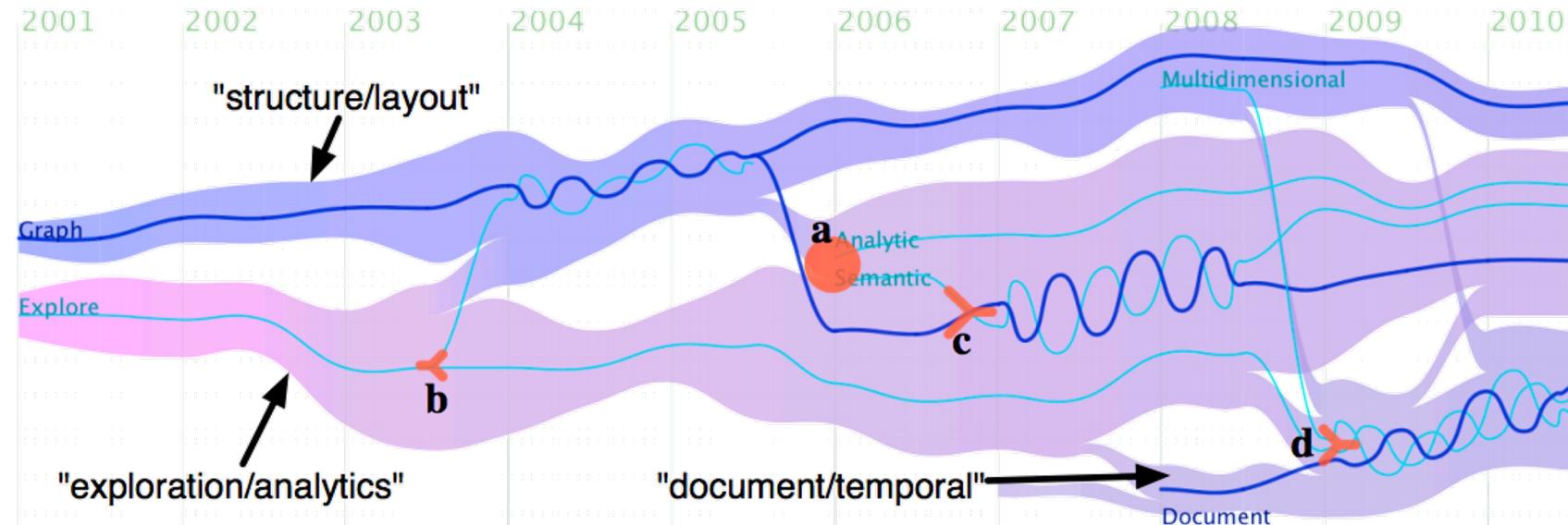
立志成才报国裕民

Obama's State of the Union



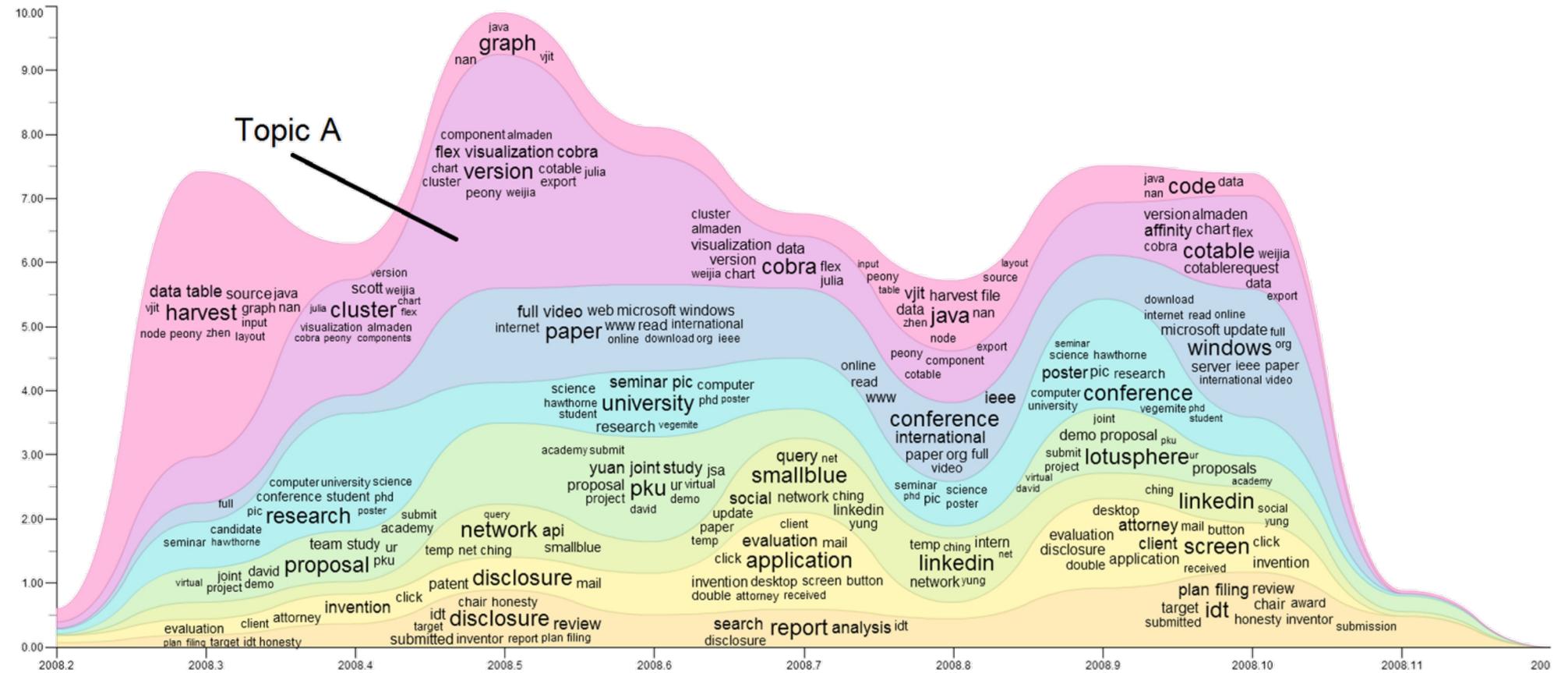
<http://interactives.ap.org/2015/sotu-2015/>

立志成才报国裕民



Weiwei Cui et al. TextFlow: Towards Better Understanding of Evolving Topics in Text. (IEEE TVCG 2011)



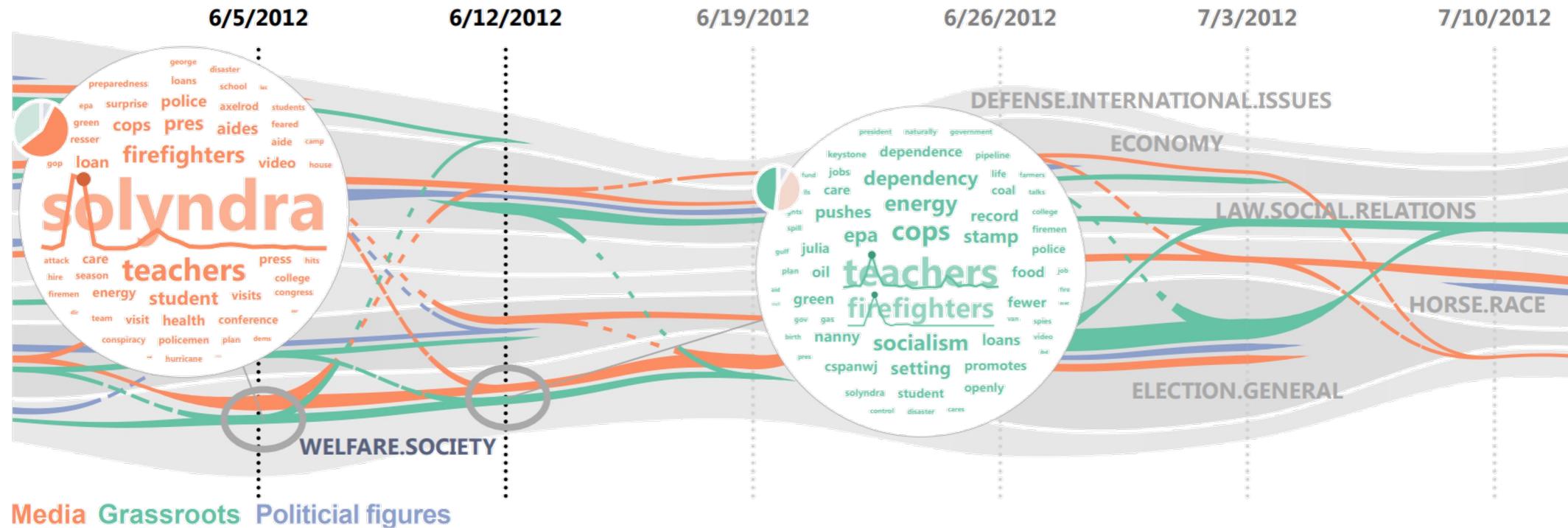


Shixia Liu et al. TIARA: Interactive, Topic-Based Visual Text Summarization and Analysis.(ACM TIST 2012)



立志成才报国裕民

Topic Competition on Social Media



Panpan Xu et al. Visual Analysis of Topic Competition on Social Media.(IEEE TVCG 2013)

SocialFlow: Visualization of Social Behavior



上海科技大学
ShanghaiTech University



Modeling of dynamic topic competition and cooperation

Visualization of dynamic behavior of opinion leaders

Propose a quantitative model, and develop an interactive visualization system to understand the social behavior of opinion leaders

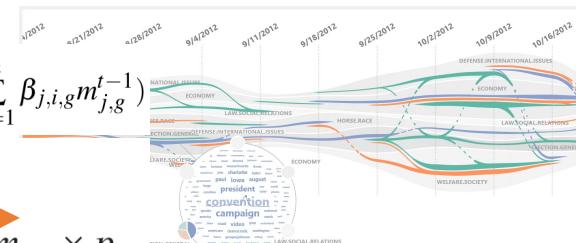
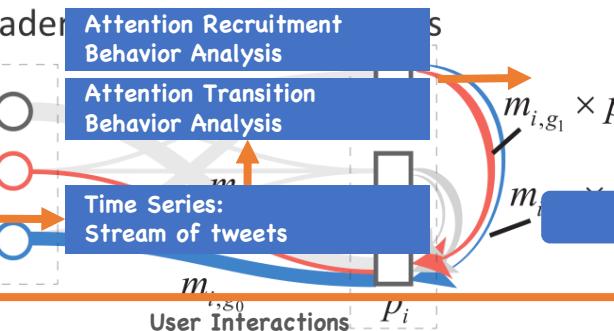


Twitter data



$$p_i^t = \alpha_i p_i^{t-1} + \sum_{g=1}^n m_{i,g}^{t-1} \sum_{j=1, j \neq i}^k \beta_{i,j,g} p_j^{t-1} - p_i^{t-1} \sum_{j=1, j \neq i}^k \left(\sum_{g=1}^n \beta_{j,i,g} m_{j,g}^{t-1} \right)$$

opinion leader



ns draw public attention to election (general)

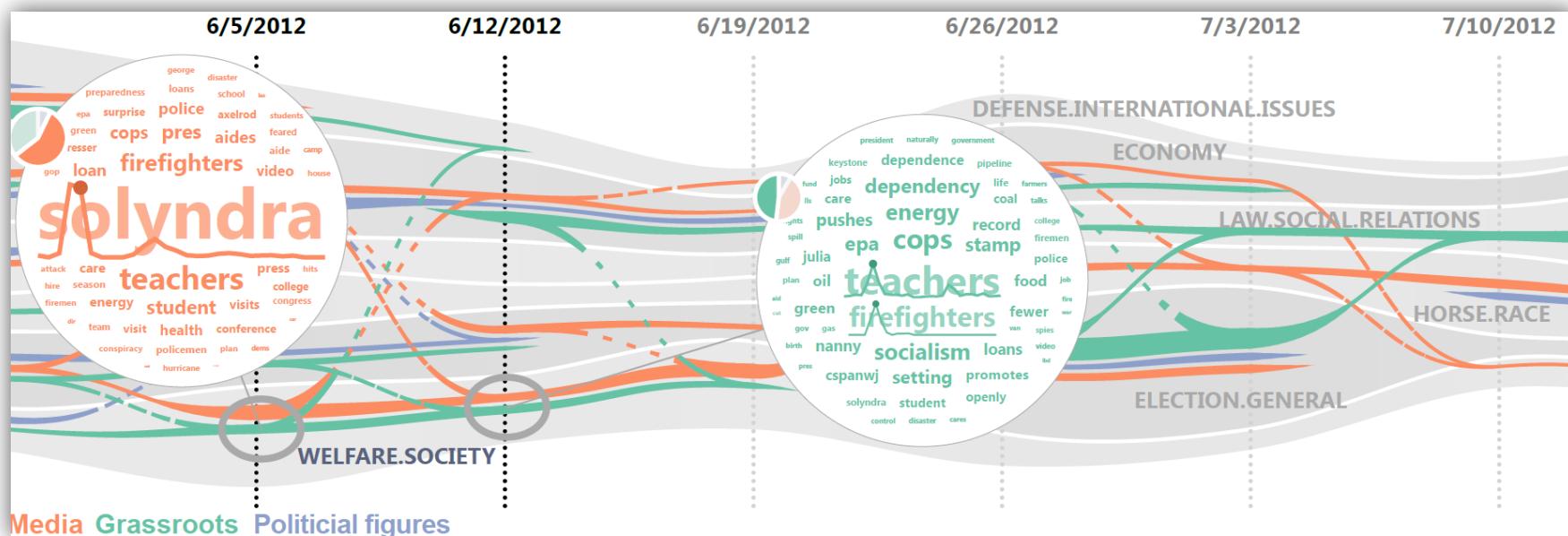
River-based visual metaphor

VAST'13 (TVCG)

VAST'14 (TVCG)

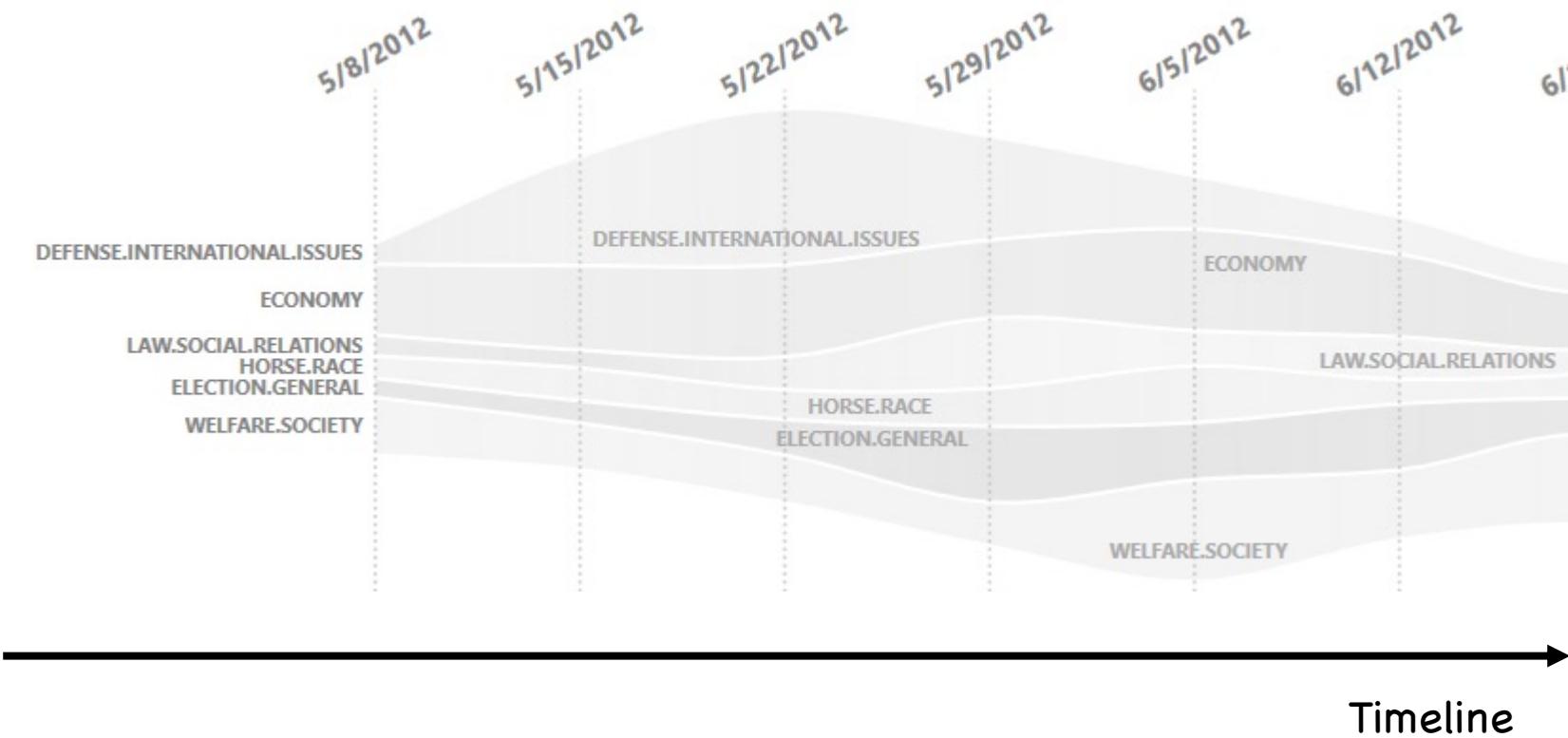
立志成才报国裕民

- 2012 US presidential election Twitter Data
 - 89,174,308 tweets classified into six major topics
 - May 01, 2012 to November 20, 2012



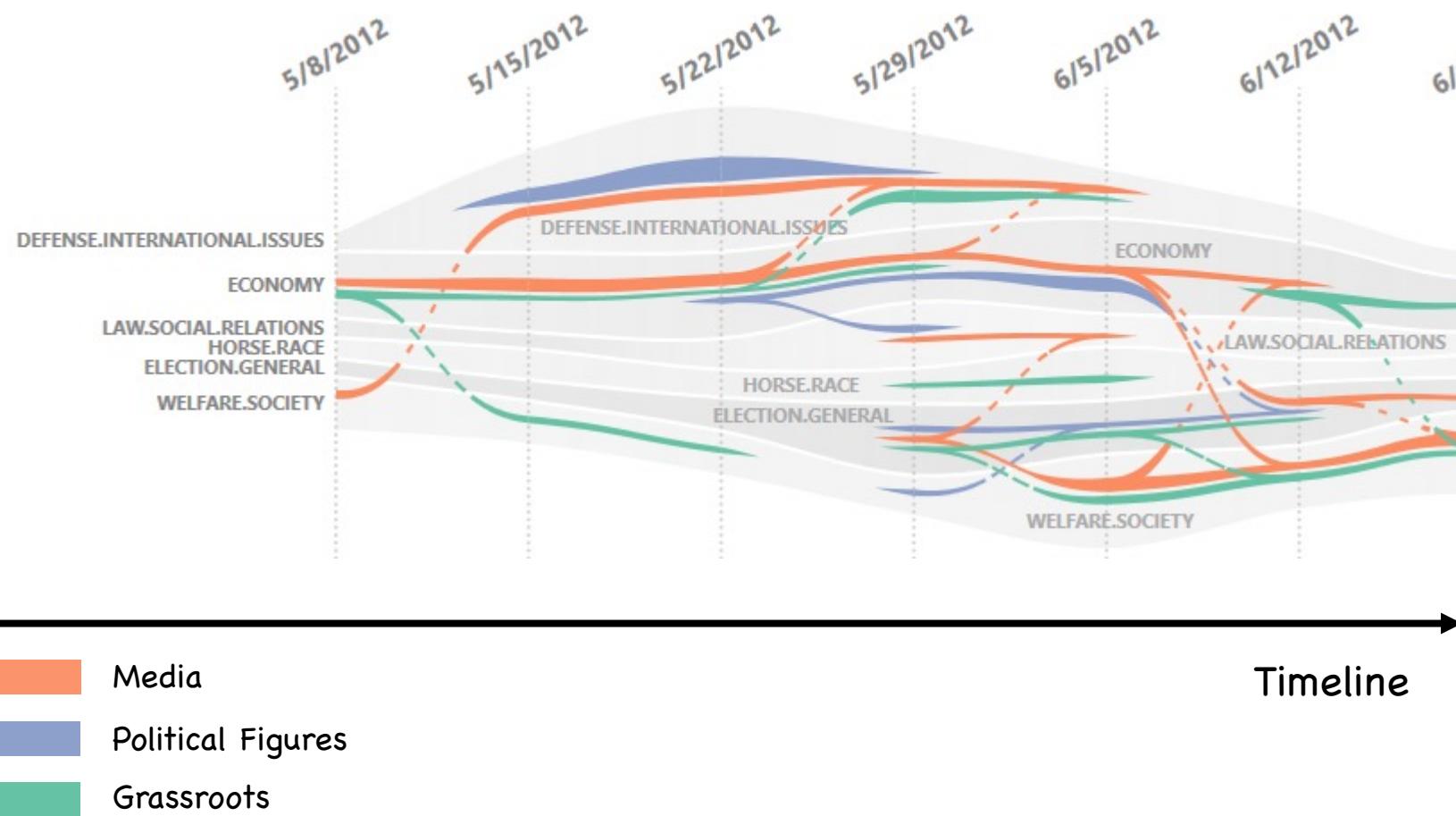
Timeline View (1/3)

- Overall trend of the competitiveness of topics



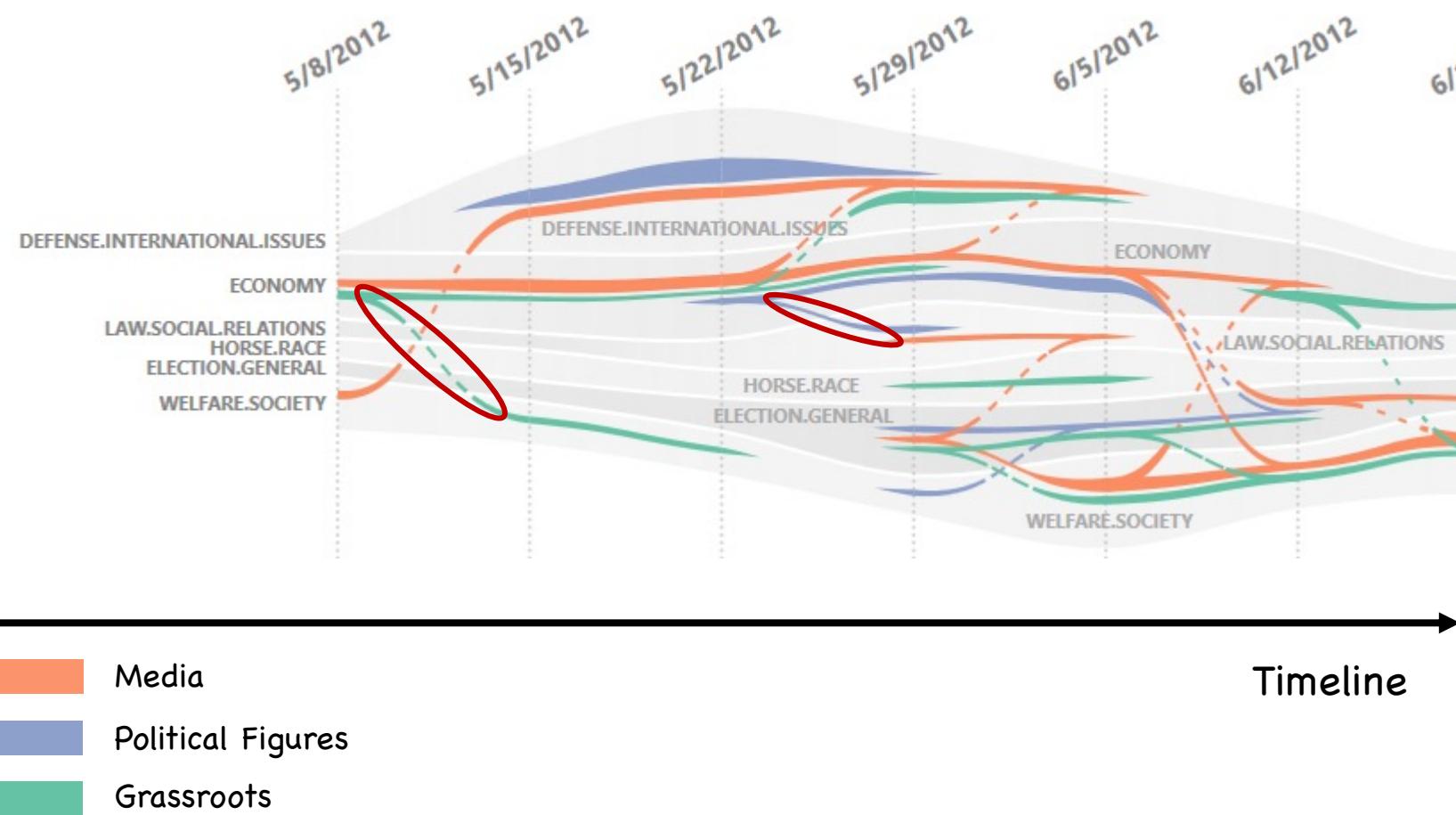
Timeline View (2/3)

- Influences of opinion leaders



Timeline View (3/3)

- Trend of topic transition of opinion leaders

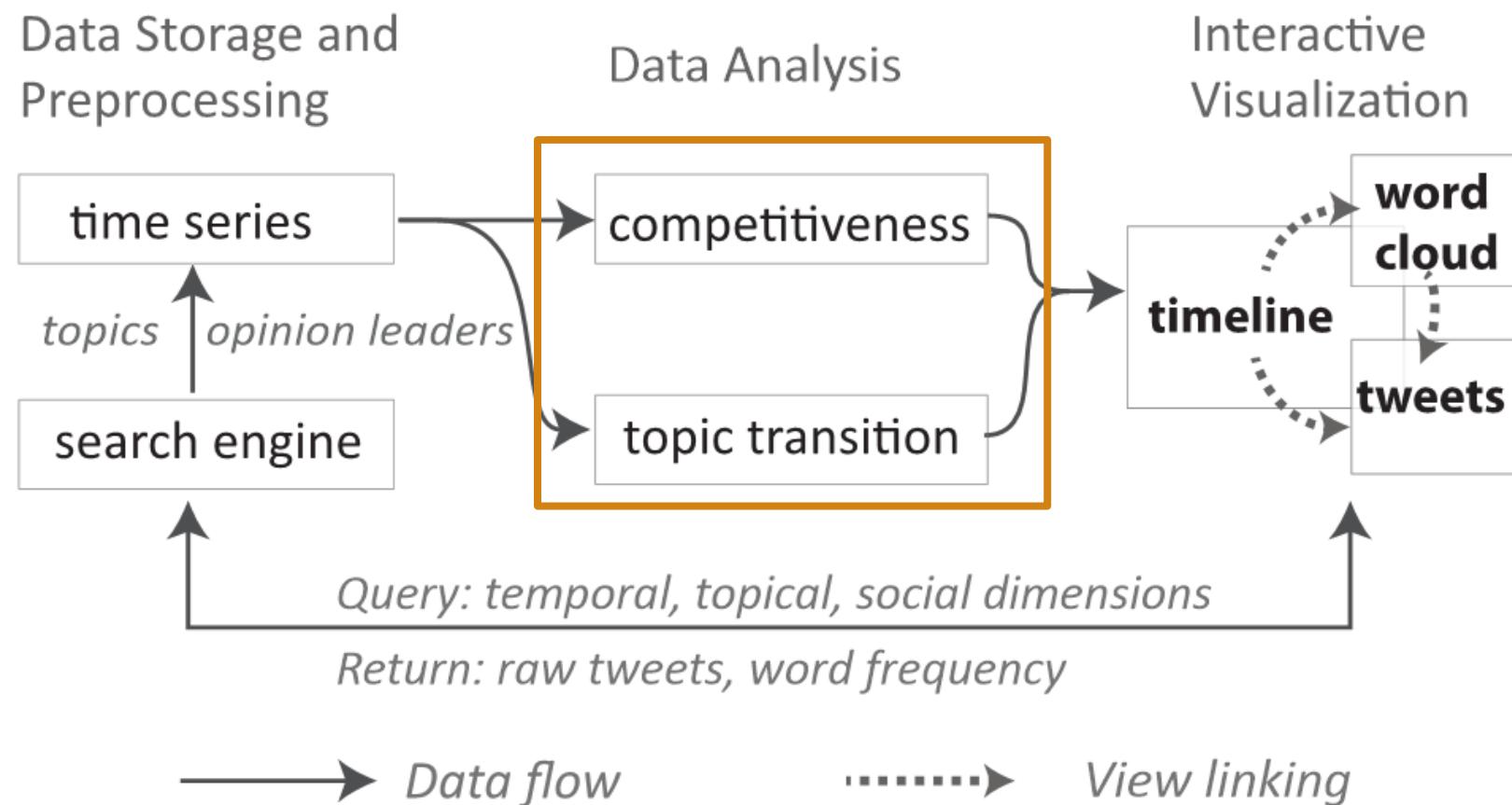


Case Study: U.S. 2012 Presidential Election

Data Analysis



- Model-based data analysis

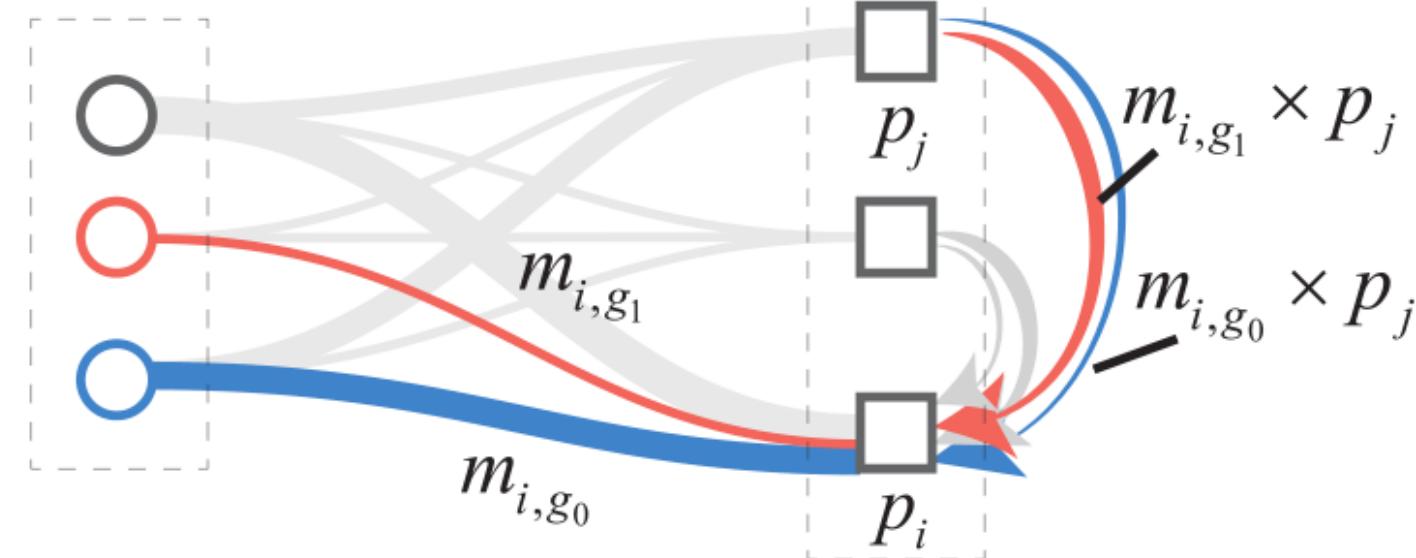


Topic Competition Model

$$p_i^t = \alpha_i p_i^{t-1} + \sum_{g=1}^n m_{i,g}^{t-1} \sum_{j=1, j \neq i}^k \beta_{i,j,g} p_j^{t-1} - p_i^{t-1} \sum_{j=1, j \neq i}^k \left(\sum_{g=1}^n \beta_{j,i,g} m_{j,g}^{t-1} \right)$$

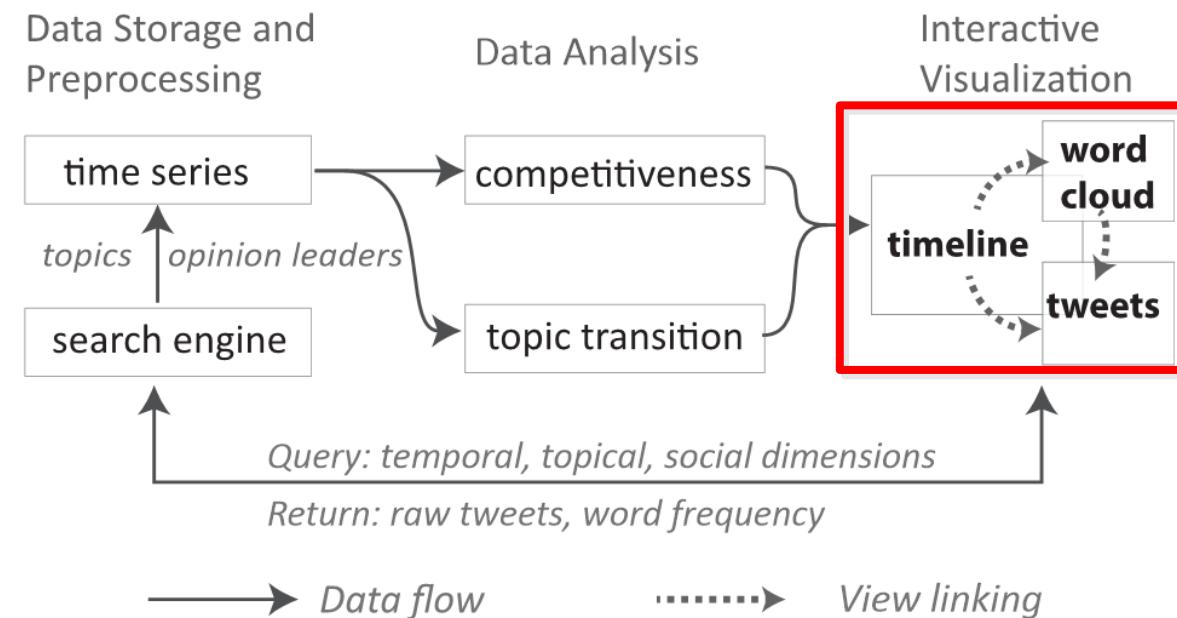
opinion leader groups

topics

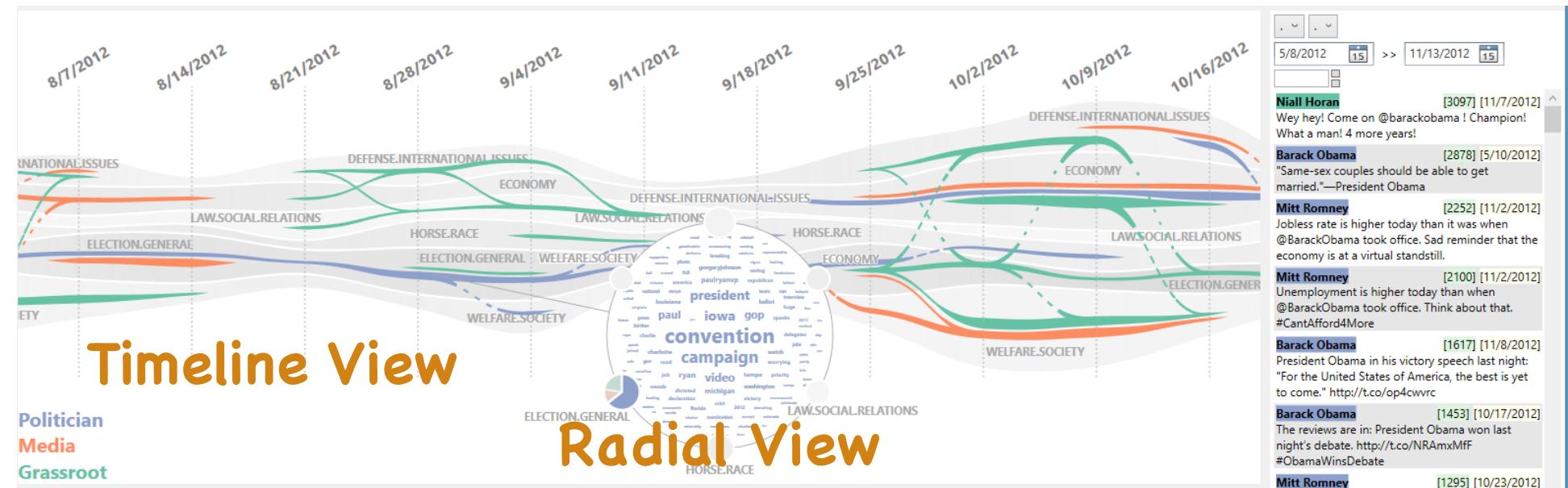


Visual Design

- Visualization system
 - Timeline View
 - Radial View
 - Tweet View



Interactive Visualization



Timeline View

Politician
Media
Grassroot

Radial View

Tweet View

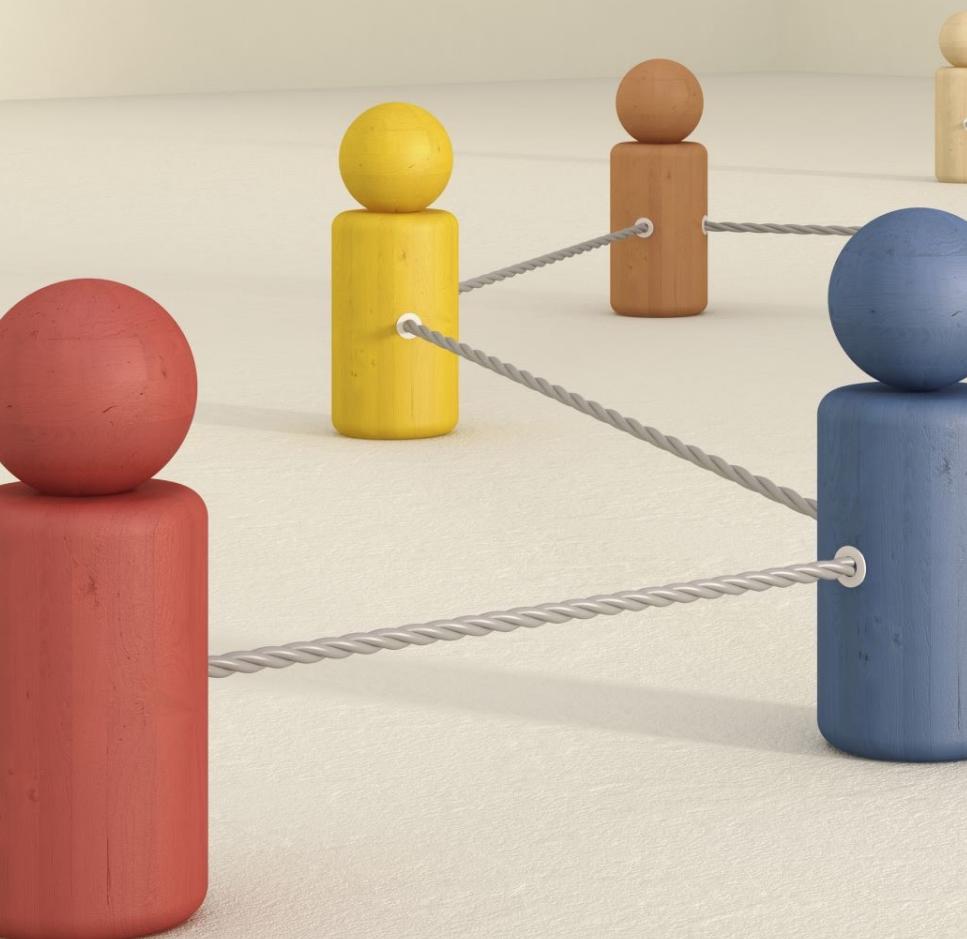
立志成才报国裕民



More Topic Interaction?

- Topic interaction is far more complex
 - Topic competition
 - Topic cooperation
 - Users' influence on their dynamics



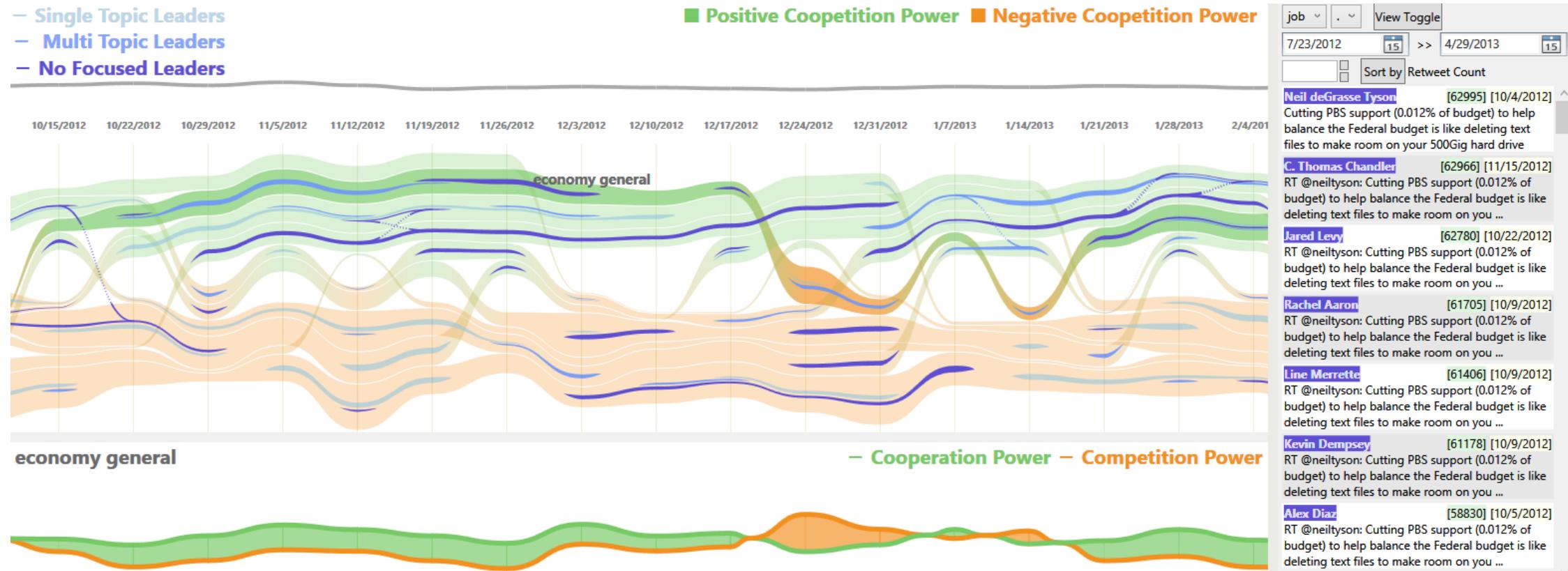


Challenges

- Measurement of dynamic topic coopetition
- Impact of users (topic leaders here)
- Interactive visualization of both above

Our Solution

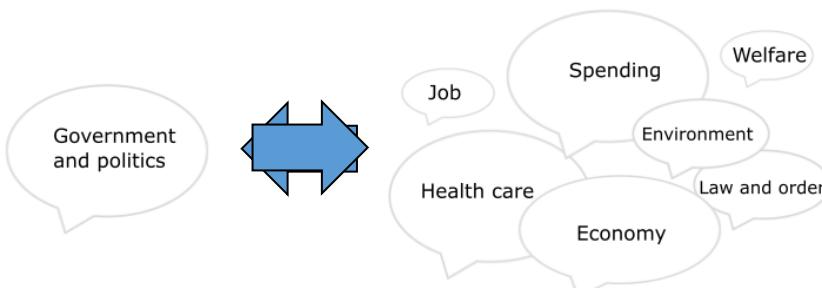
- A new visual analysis framework
 - A new model that characterizes both topic cooperation and competition
 - A new design that visualizes co-evolutionary patterns between the topics and the topic leaders



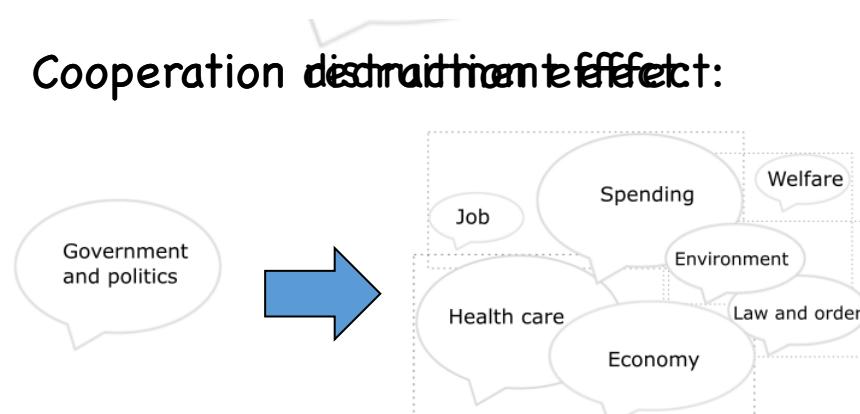
Topic competition model

$$\begin{aligned}
 p_i^t = & \boxed{\alpha_i p_i^{t-1}} + \boxed{\sum_{g=1}^n m_{i,g}^{t-1} \sum_{j=1, j \neq i}^k \beta_{i,j,g} p_j^{t-1}} - \boxed{p_i^{t-1} \sum_{g=1}^n \sum_{j=1, j \neq i}^k \beta_{j,i,g} m_{j,g}^{t-1}} \\
 & + \boxed{\sum_{g=1}^n \sum_{j=1, j \neq i}^k (m_{i,g}^{t-1} + m_{j,g}^{t-1}) \theta_{i,j} \sum_{h=1, h \neq i, h \neq j}^k \gamma_{ij,h,g} p_h^{t-1}} \\
 & - \boxed{p_i^{t-1} \sum_{g=1}^n \sum_{j=1, j \neq i}^k \sum_{h=1, h \neq i, h \neq j}^k (m_{j,g}^{t-1} + m_{h,g}^{t-1}) \theta_{j,h} \gamma_{jh,i,g}}
 \end{aligned}$$

Competition destruction effect:



Cooperation destruction effect:



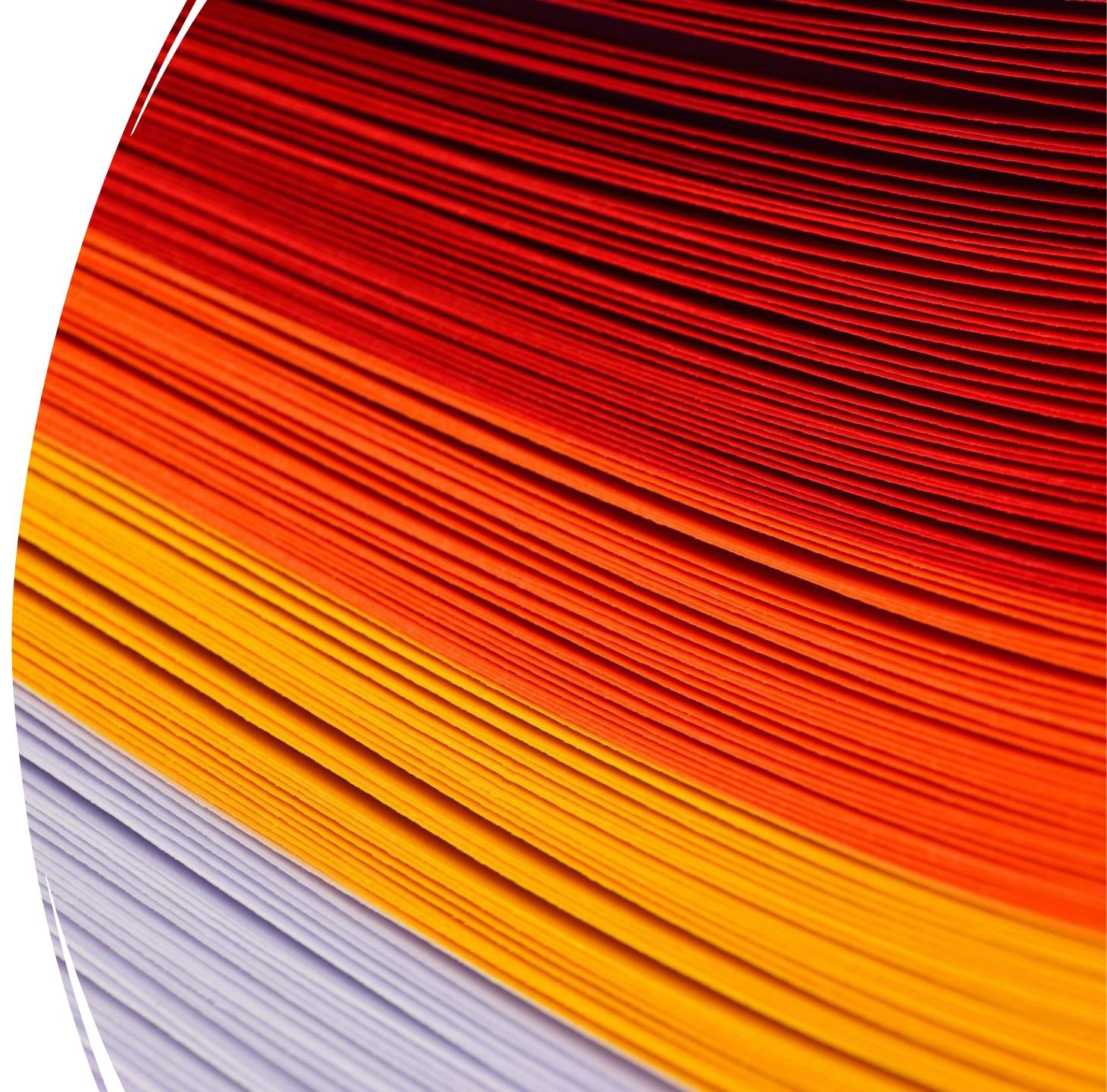
Visualizing Document Content

- Keyword-based Visualization
- Temporal Document Visualization
- Feature Distribution Visualization
- Opinion Analysis
- Query Visualization
- Software Visualization



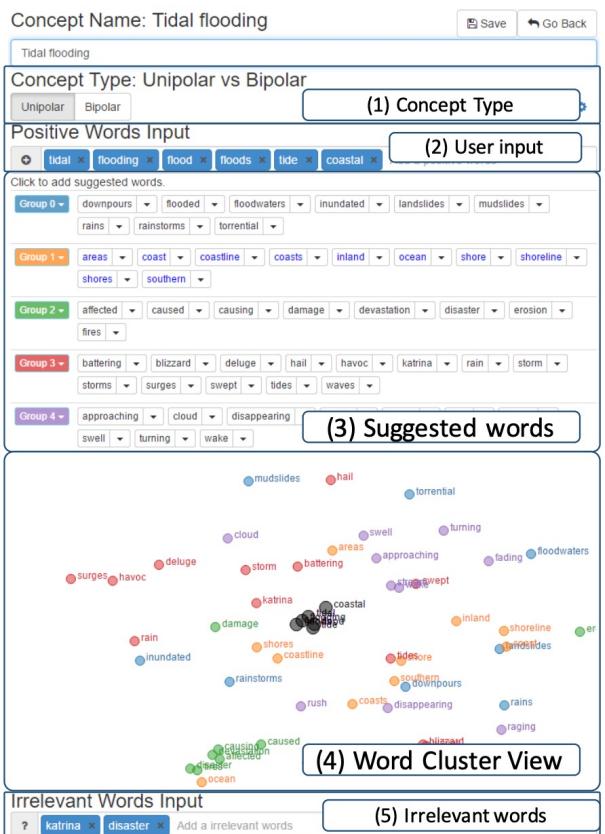
Feature Distribution Visualization

- Feature distribution in a single document or a document collection:
 - Average sentence length
 - Vocabulary
 -

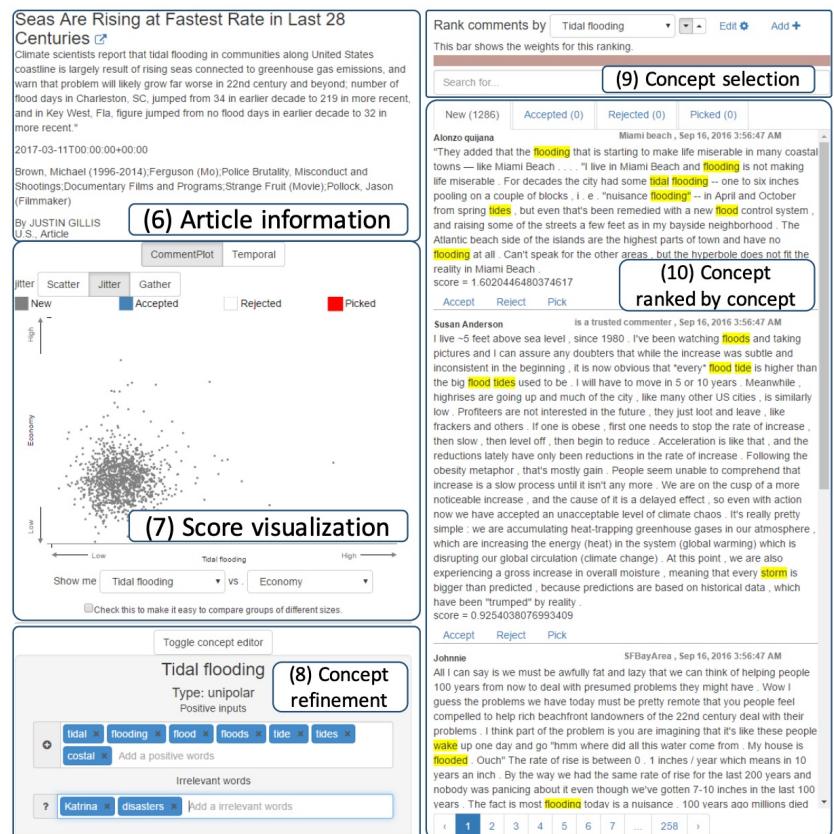


ConceptVector

Deokgun Park et al. ConceptVector: Text Visual Analytics via Interactive Lexicon Building using Word Embedding. (IEEE TVCG 2018)



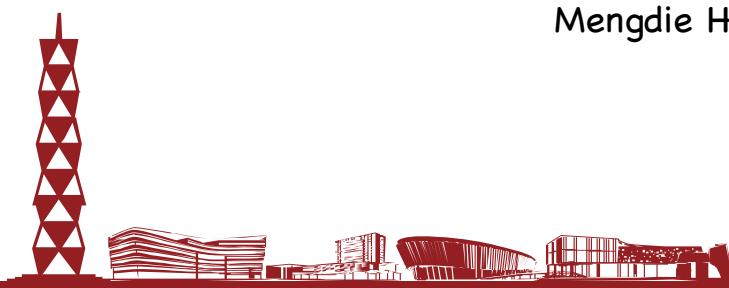
(a) Concept-Building View



(b) Document Analysis View



Mengdie Hu et al. Visualizing Social Media Content with SentenTree. (IEEE TVCG 2017)



立志成才报国裕民

INFOVIS PAPER

Visualizing Social Media Content with SentenTree

Mengdie Hu, Krist Wongsuphasawat, John Stasko

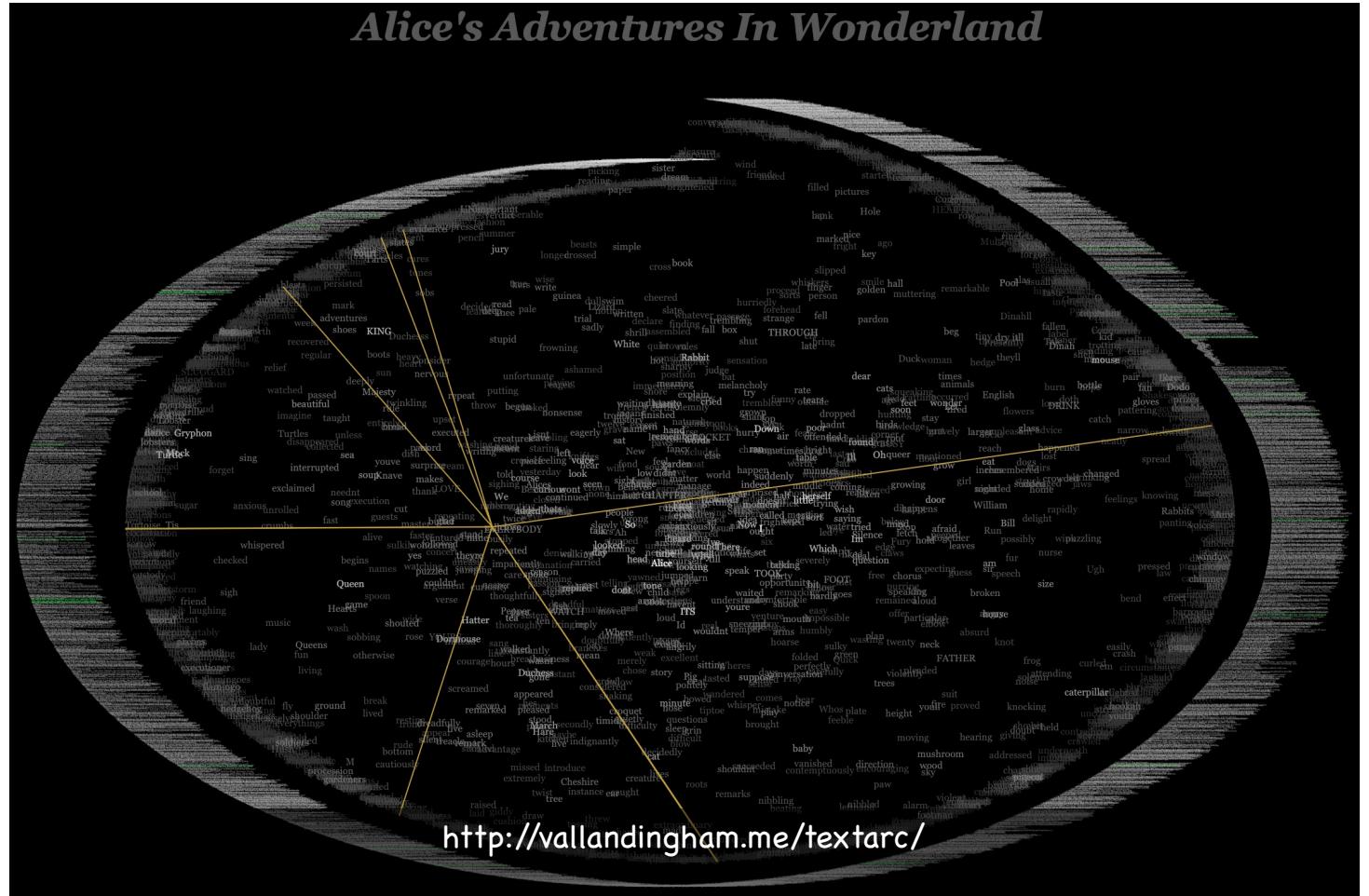


23-28 October 2016
Baltimore, Maryland, USA

ieeevis.org

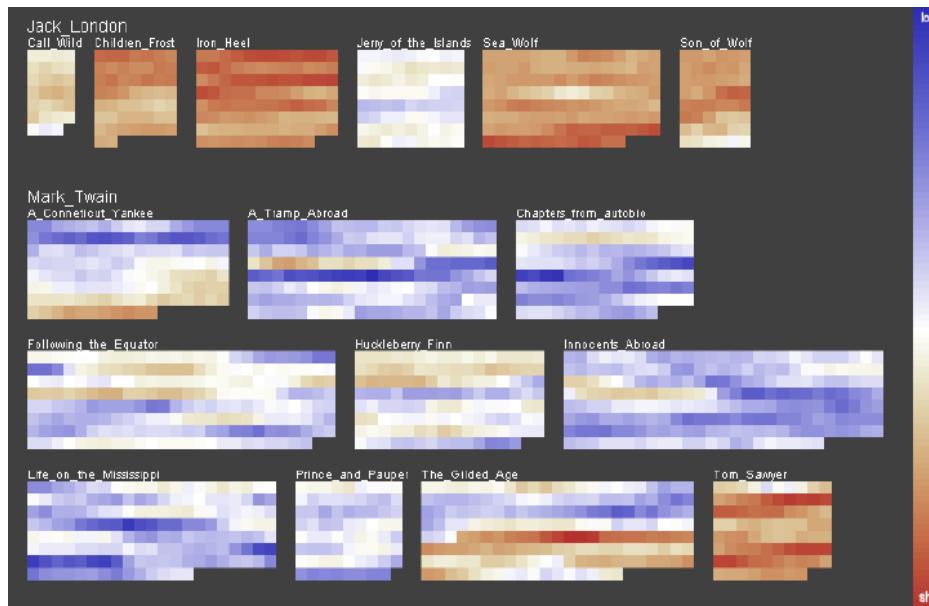
Text+Arc

- TextArc:
 - Outer arc:
sentences in a
document
 - Inner words:
vocabulary of the
document (words
closer to center
means higher word
frequency)



Literature Fingerprinting

- Describe statistical features of documents



Difference of writing style
between Mark Twain and Jack
London – sentence length

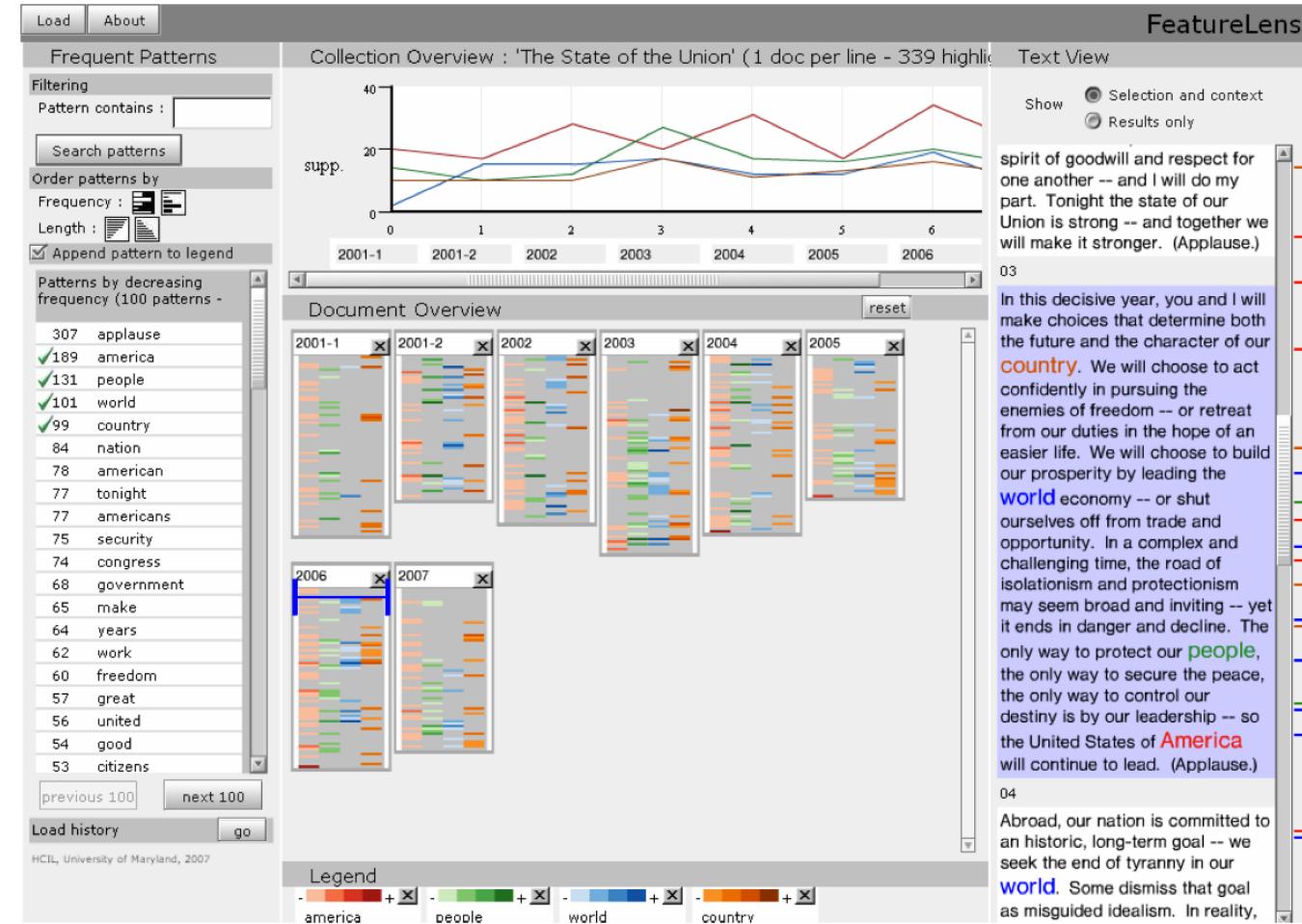
- Each pixel represents a paragraph
- A set of pixels represents a book
- Color encodes sentence length

Daniel A. Keim & Daniela Oelke. Literature Fingerprinting: A New Method for Visual Literary Analysis. (IEEE VAST 2007)



立志成才报国裕民

FeatureLens



Anthony Don et al. FeatureLens: Interactive Visualization of Text Patterns. (Information Visualization 2007)



立志成才报国裕民

Visualizing Document Content

- Keyword-based Visualization
- Temporal Document Visualization
- Feature Distribution Visualization
- Opinion Analysis
- Query Visualization
- Software Visualization



Opinion Analysis

- Opinion analysis; Sentiment analysis
- Subjective feelings: like or dislike, praise or criticize
- Opinion bias: positive ↔ negative



Emotional arcs for inaugural addresses



<http://flowingdata.com/2017/02/01/emotional-arcs-for-inaugural-addresses/>

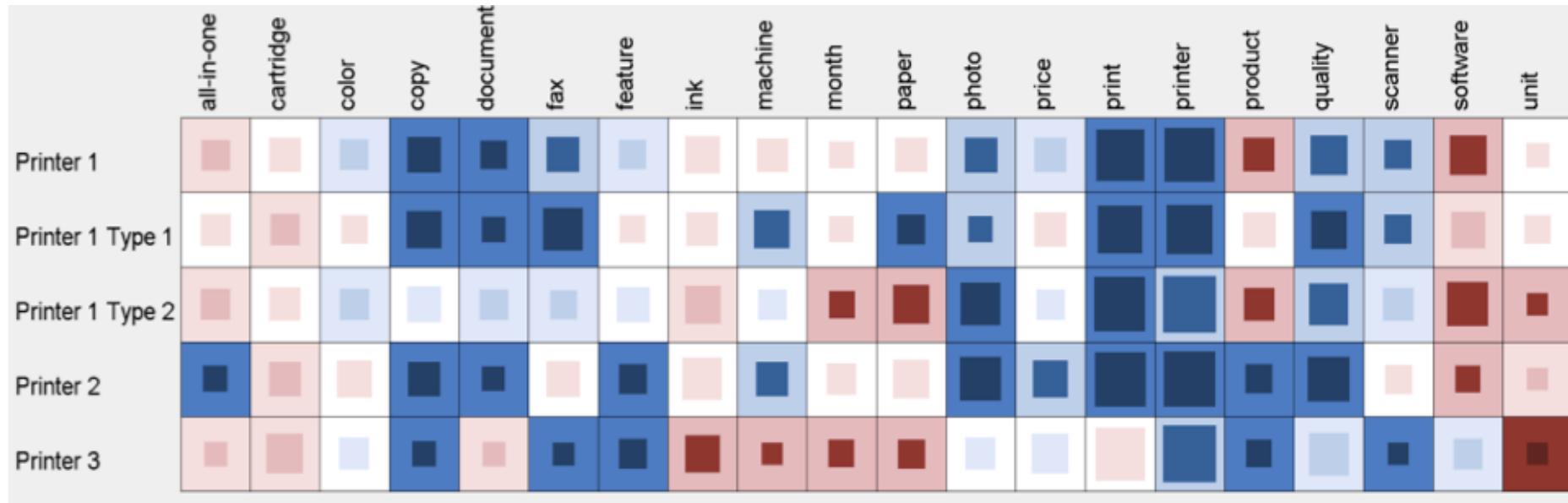
立志成才报国裕民

Matrix-based Customer Feedback Visualization

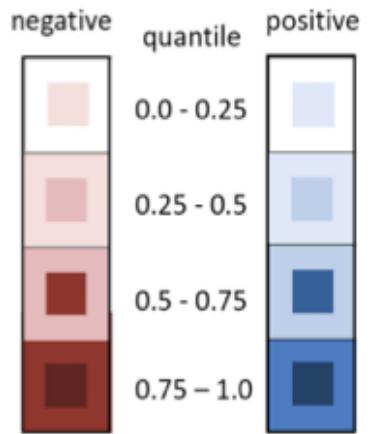


上海科技大学
ShanghaiTech University

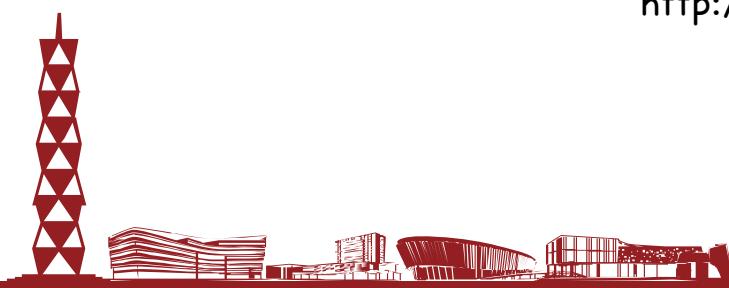
- Horizontal: features
- Vertical: models of printers



Color Scale

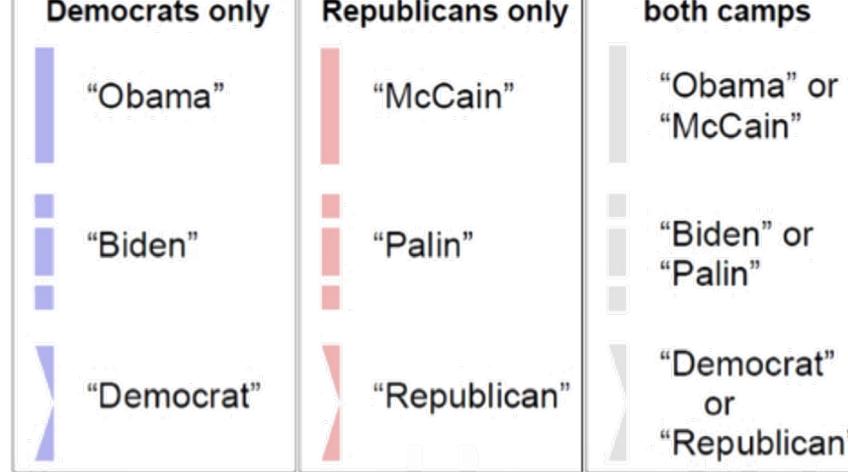


<http://flowingdata.com/2017/02/01/emotional-arcs-for-inaugural-addresses/>

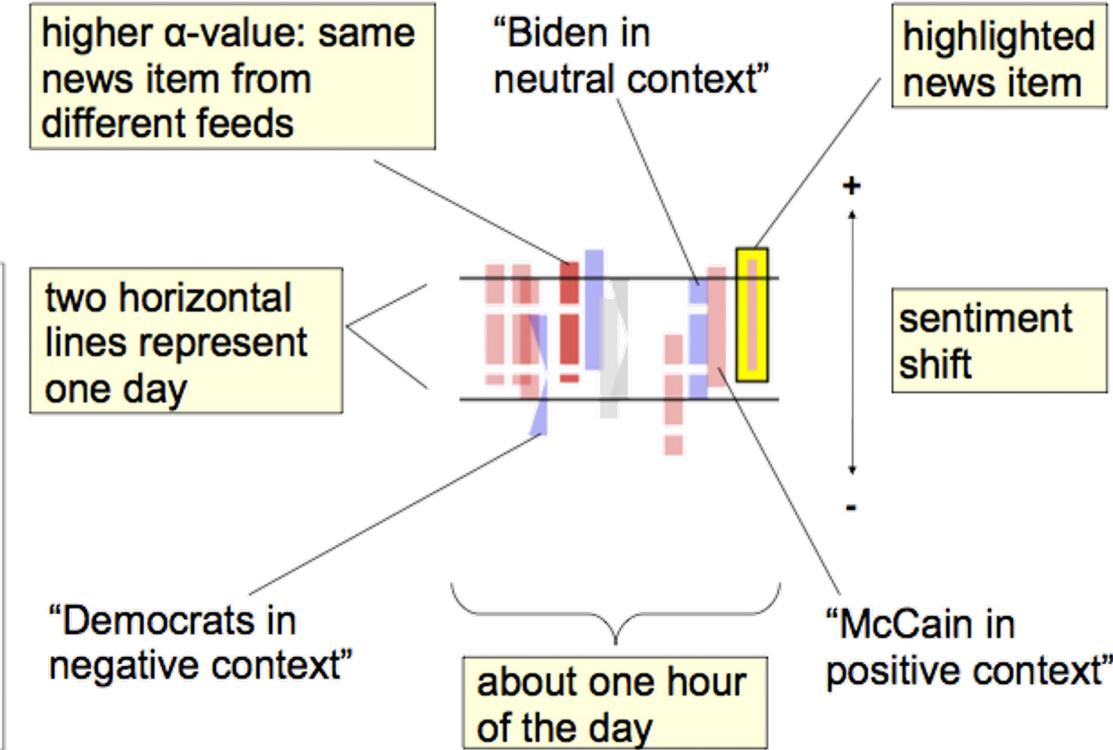


立志成才报国裕民

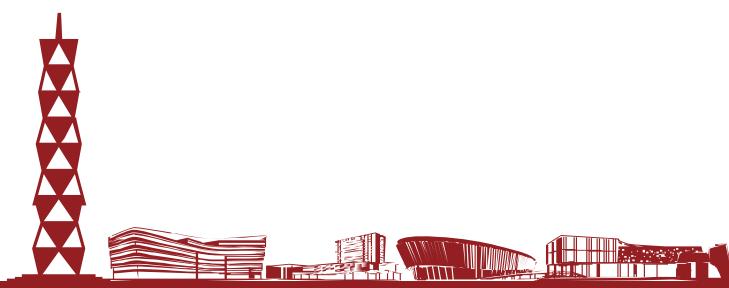
President Election: Opinions in News

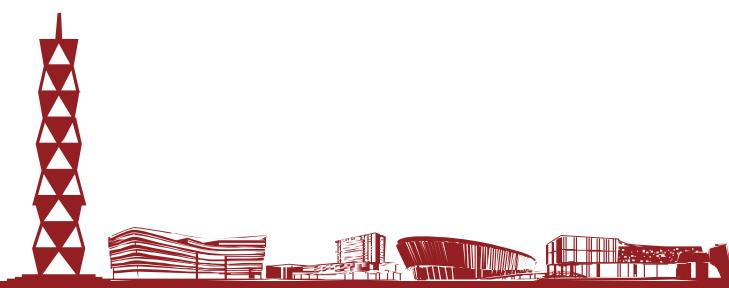
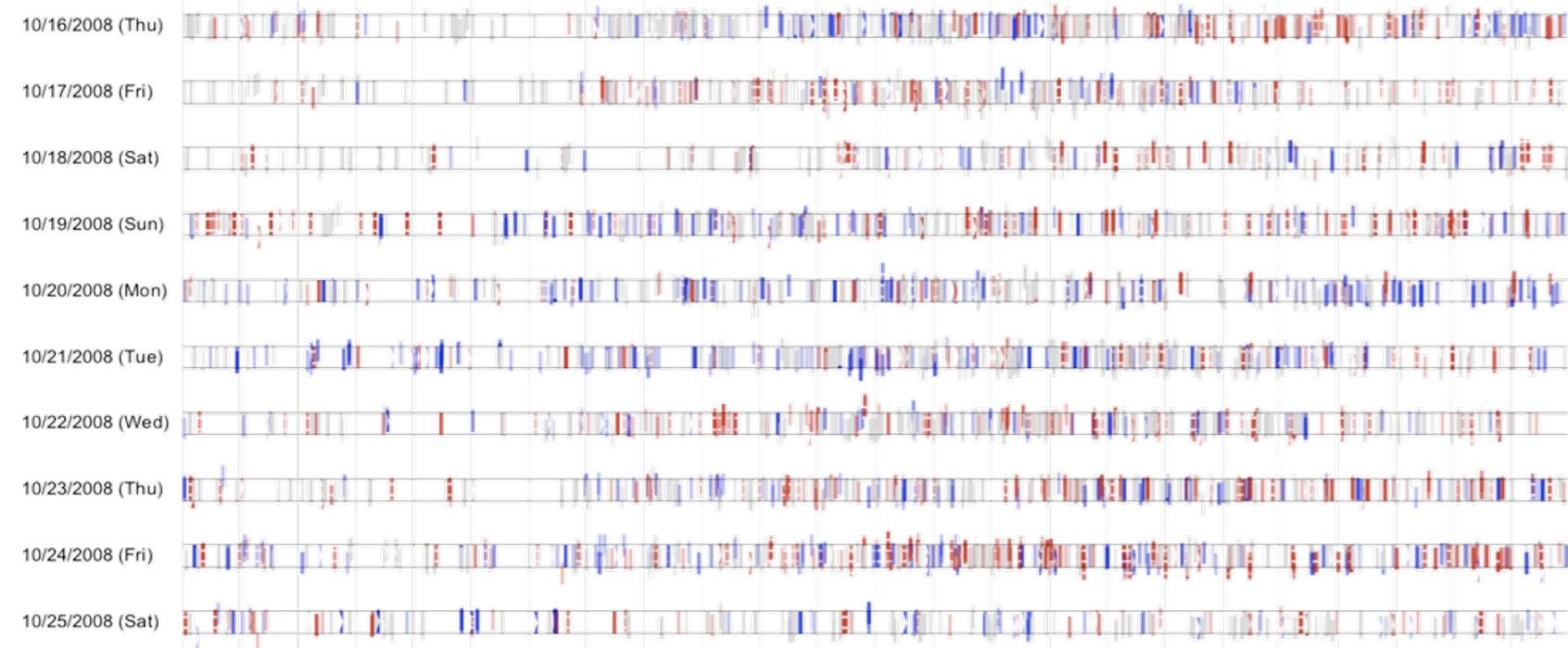


(a)



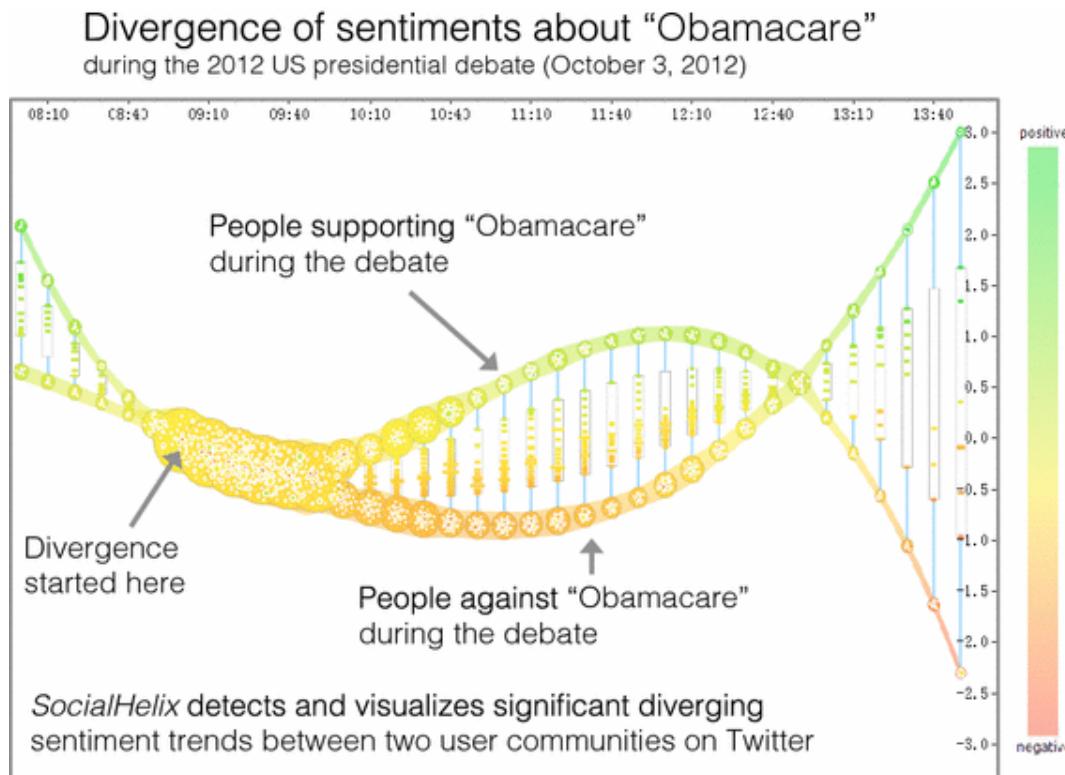
(b)





SocialHelix

- Nan Cao et al. SocialHelix: visual analysis of sentiment divergence in social media. (Journal of Visualization 2015)



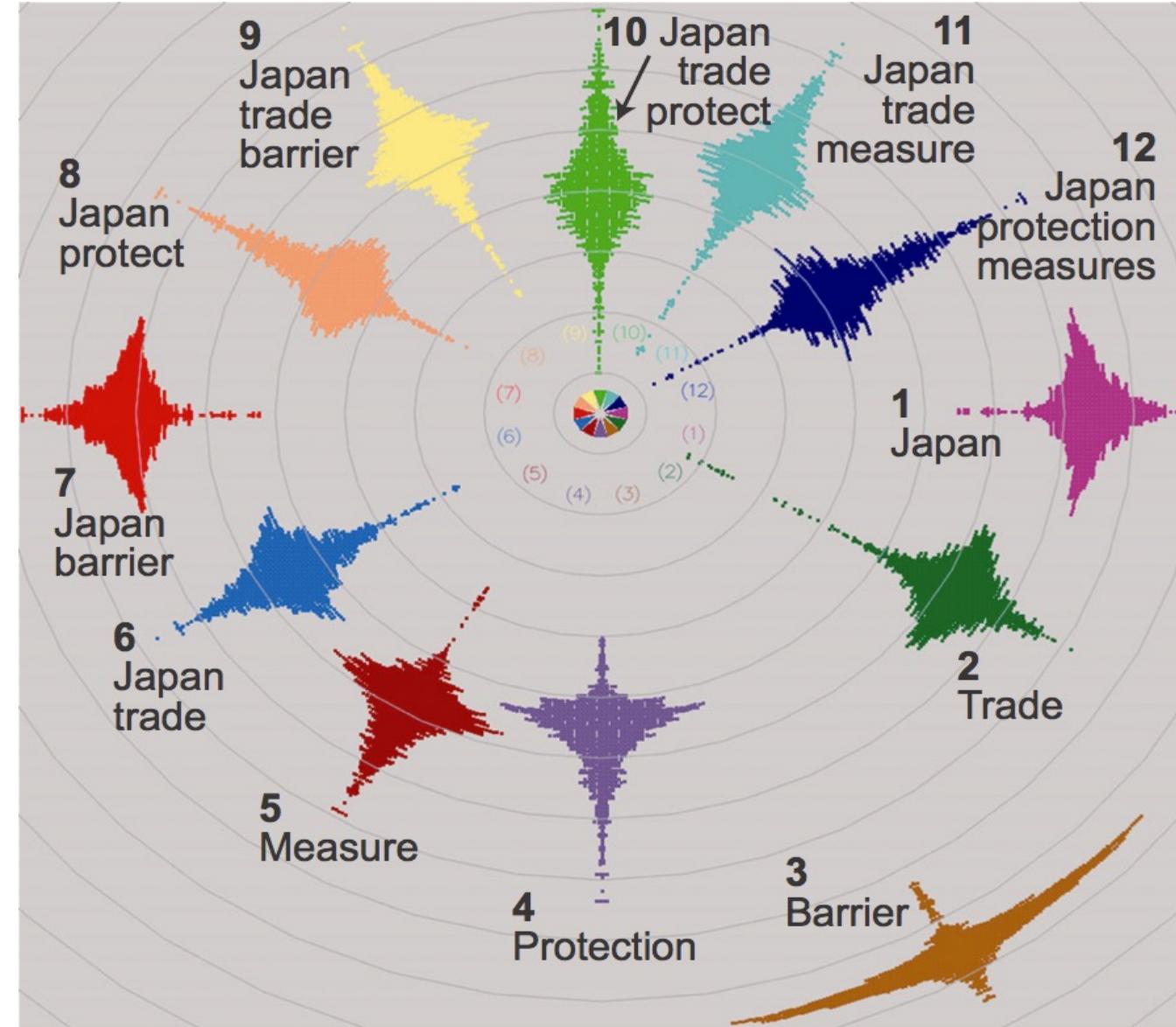
Visualizing Document Content

- Keyword-based Visualization
- Temporal Document Visualization
- Feature Distribution Visualization
- Opinion Analysis
- Query Visualization
- Software Visualization



Query Visualization

- Visual-guided search/querying:
 - Understand search result
 - Discover distribution patterns in results
- How to apply visualization:
 - Similarities of queries itself or search results
 - Words in queries
 - ...



Susan Havre et al. Interactive Visualization of Multiple Query Results (IEEE INFOVIS 2001)



立志成才报国裕民

Visualizing Document Content

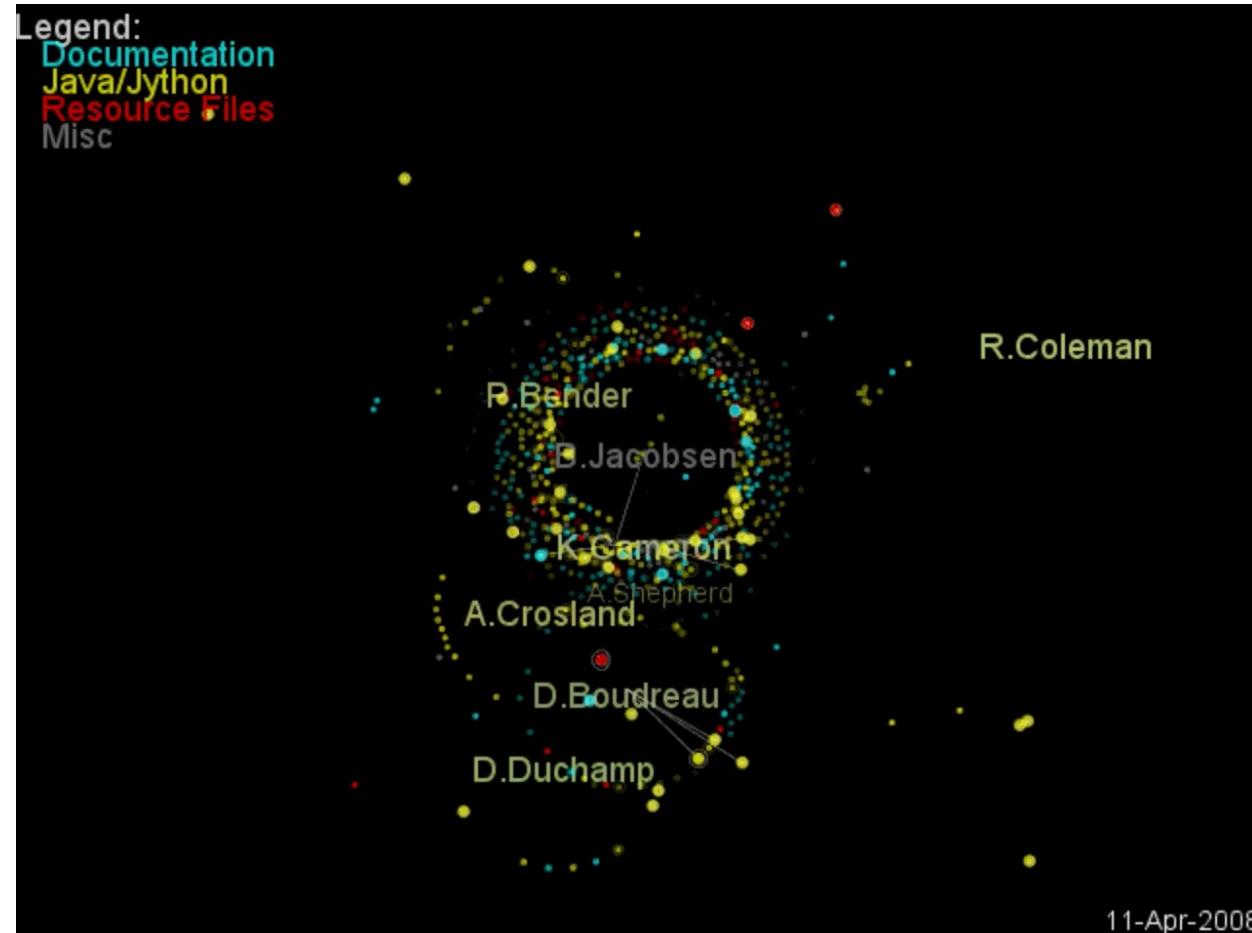
- Keyword-based Visualization
- Temporal Document Visualization
- Feature Distribution Visualization
- Opinion Analysis
- Query Visualization
- Software Visualization



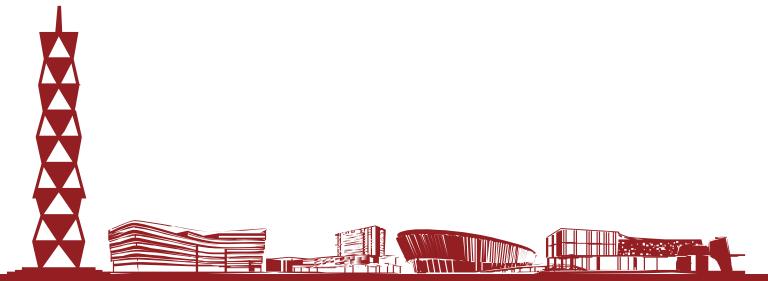
Software Visualization

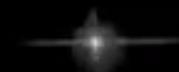
- Source codes can be regarded as a special type of text
- Levels of Software Visualization:
 - Source files
 - Structure of functions/modules, and relations between them





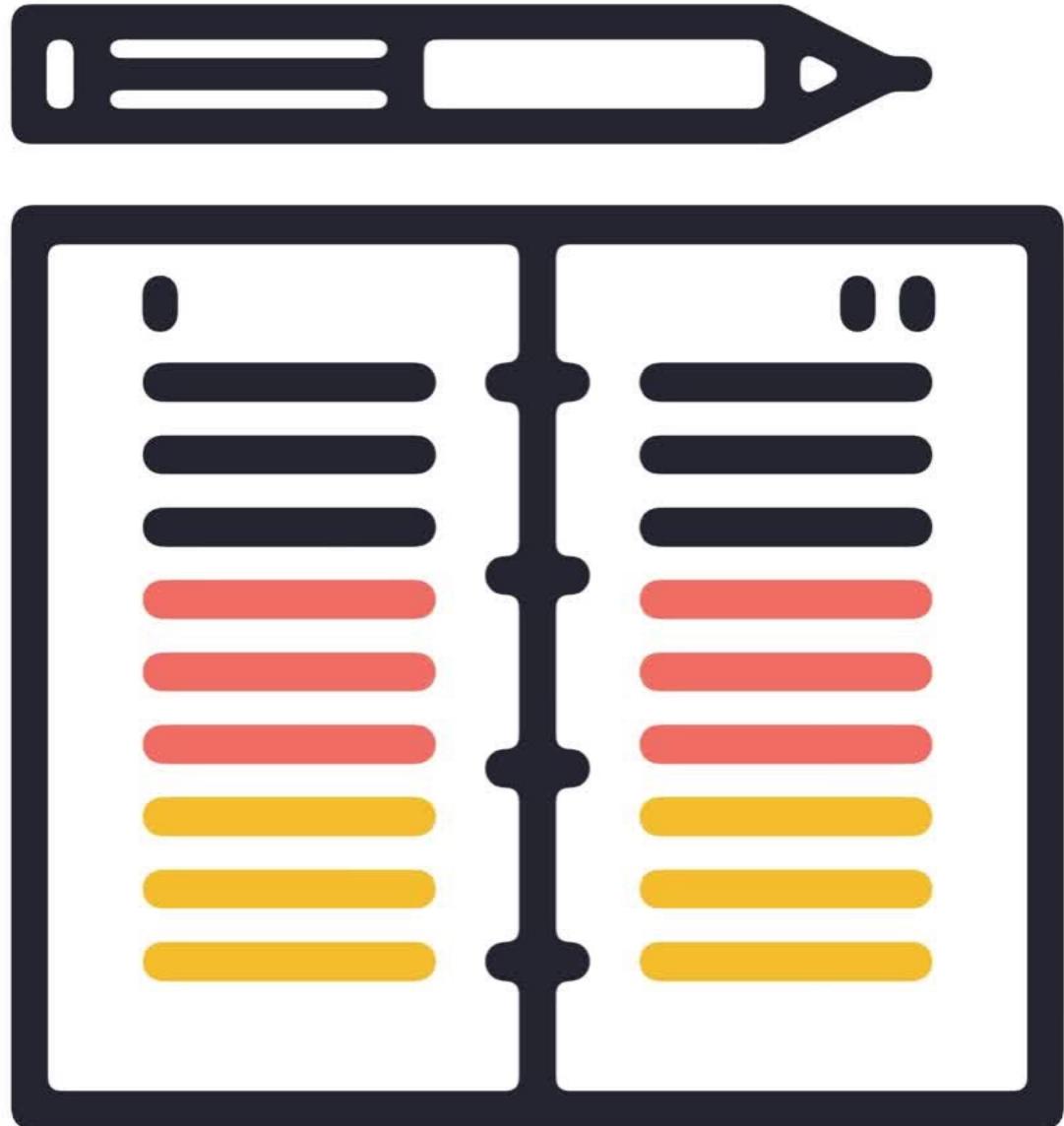
<http://jmri.org/community/codeswarm.shtml>





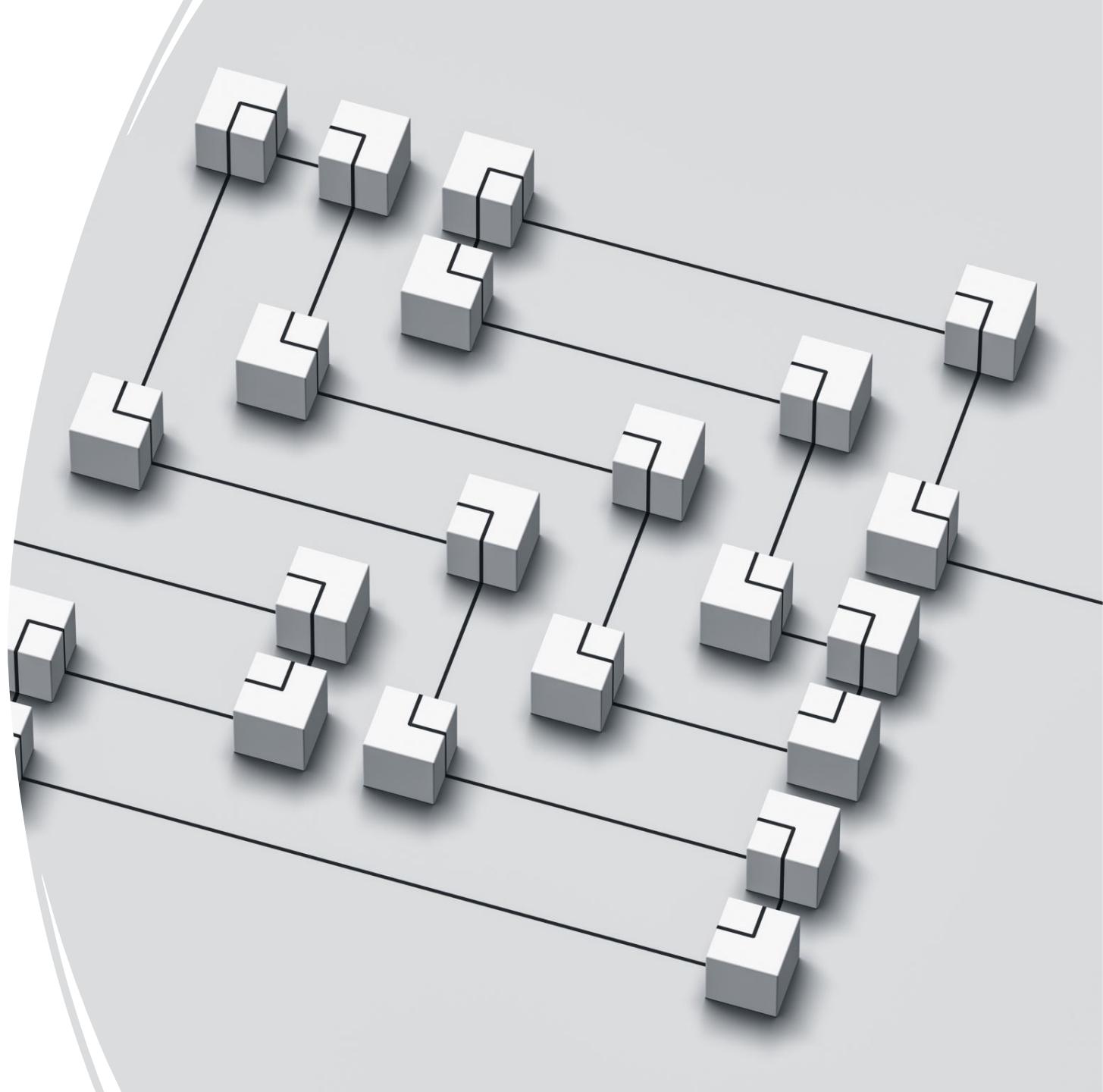
Text Visualization

- Categorized based on features
 - Visualizing Document Content
 - Visualizing Document Relations
 - Multi-level Document Visualization

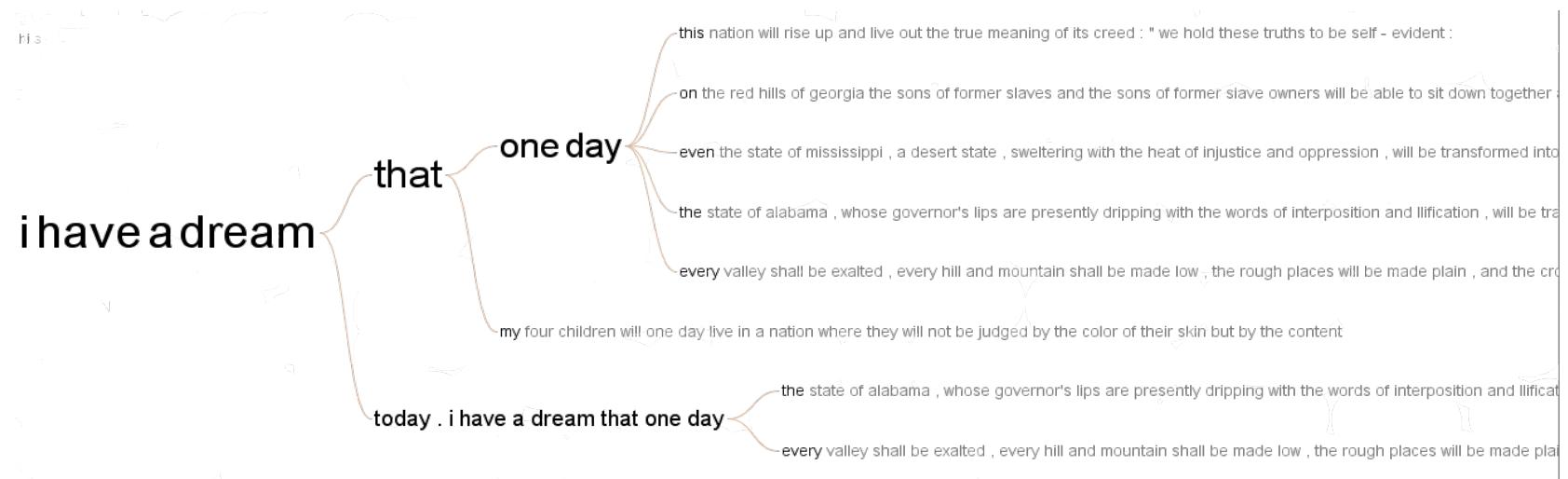


Visualizing Document Relations

- Relations in a document collection
 - References (paper)
 - Hyperlinks (web pages)
 - Similarity and Hierarchy
- Methods:
 - Graph Layout
 - Tree Layout



- Similar to Trie-tree
- Size of a word → word frequency



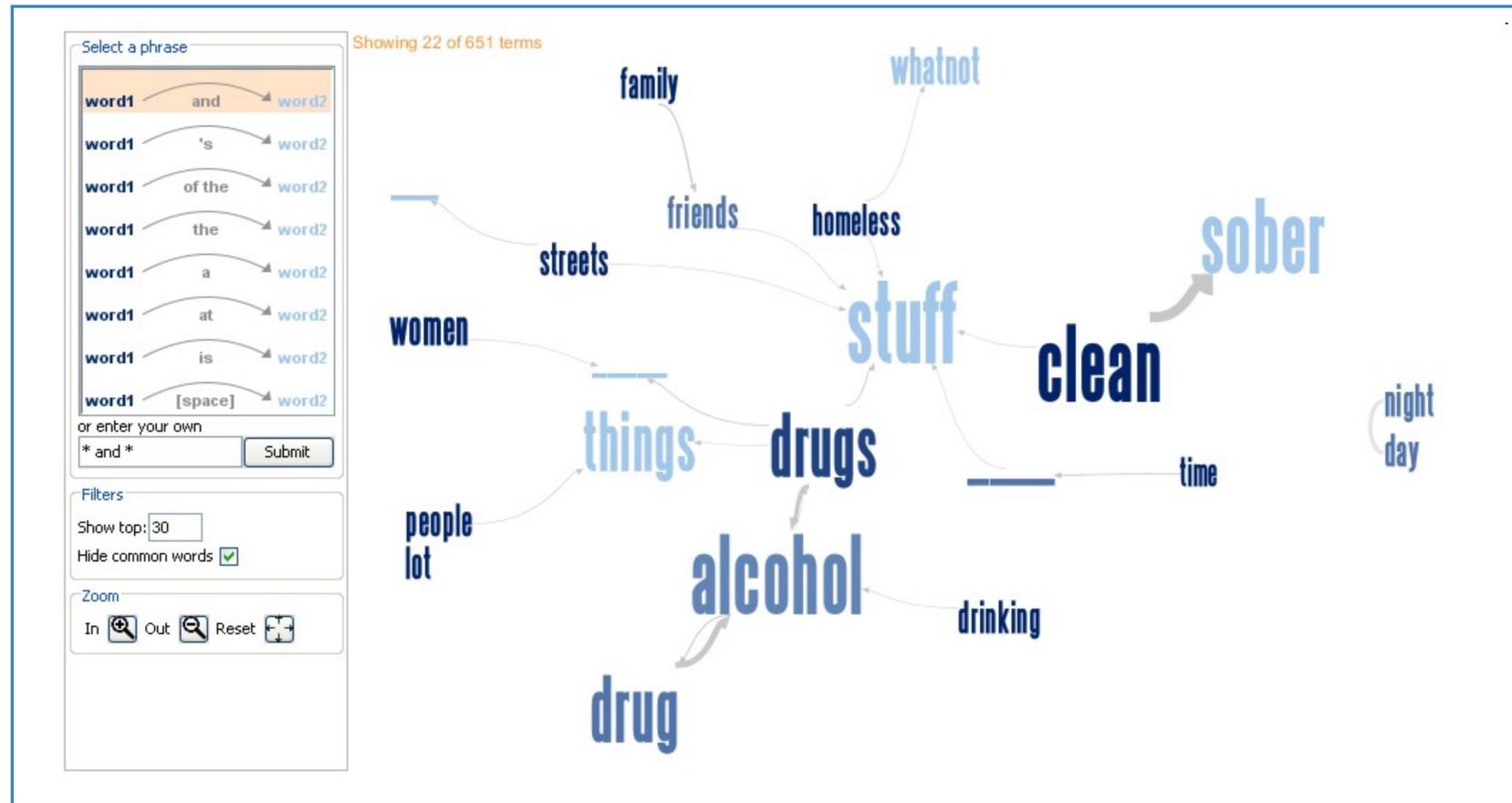
All of the sentences starting with "I" in the speech "I have a dream"

Wattenberg M and Viégas F B. The word tree, an interactive visual concordance. (IEEE TVCG 2008)



立志成才报国裕民

- Relation of “A the B” by node-link view



Van Ham F et al. Mapping text with phrase nets. (IEEE TVCG 2009)

立志成才报国裕民



- Show similarities by treemap

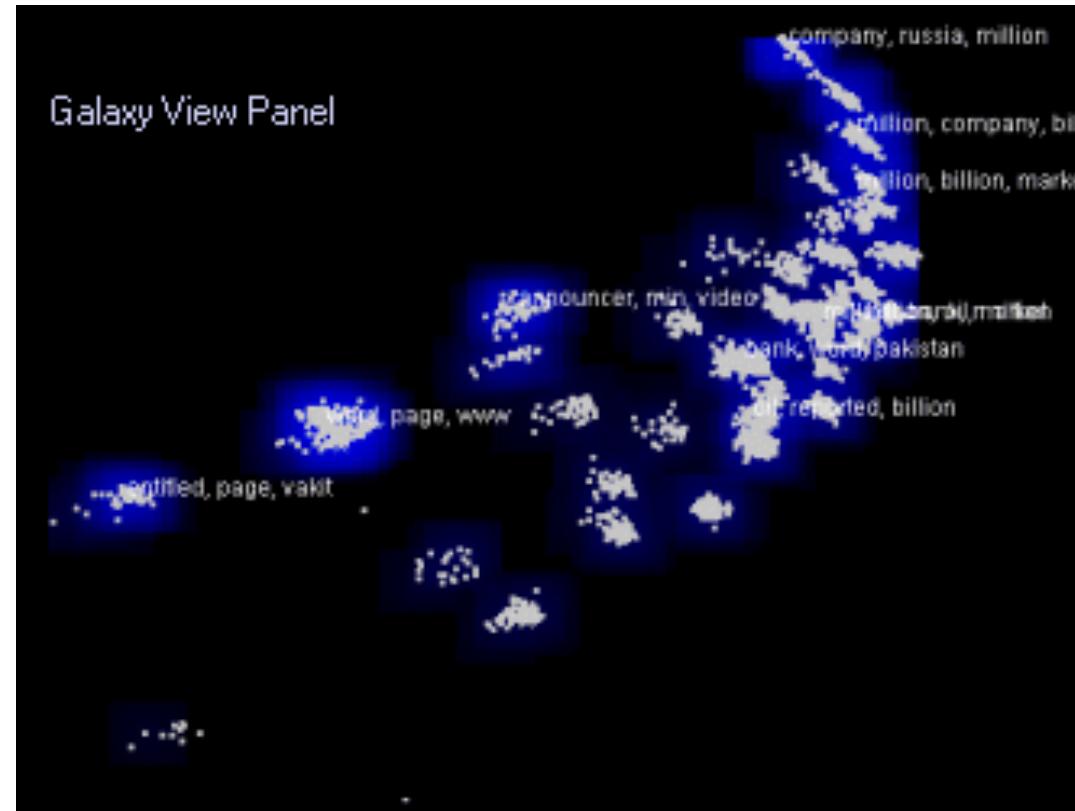
<http://newsmap.jp/>

Document Collection Visualization

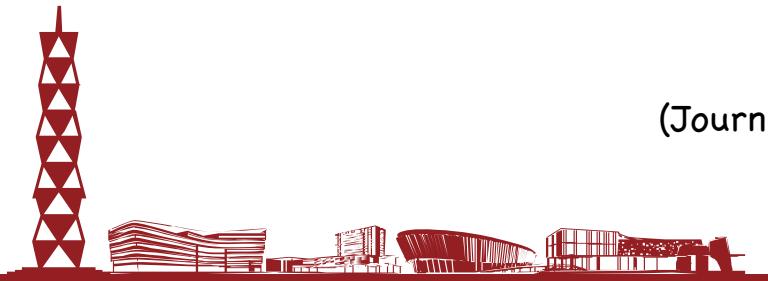
- Visualize relations between documents in a collection
- Method: Multidimensional Projection



- MDS based on document similarities



Wise J A. The ecological approach to text visualization.
(Journal of the Association for Information Science and Technology 1999)



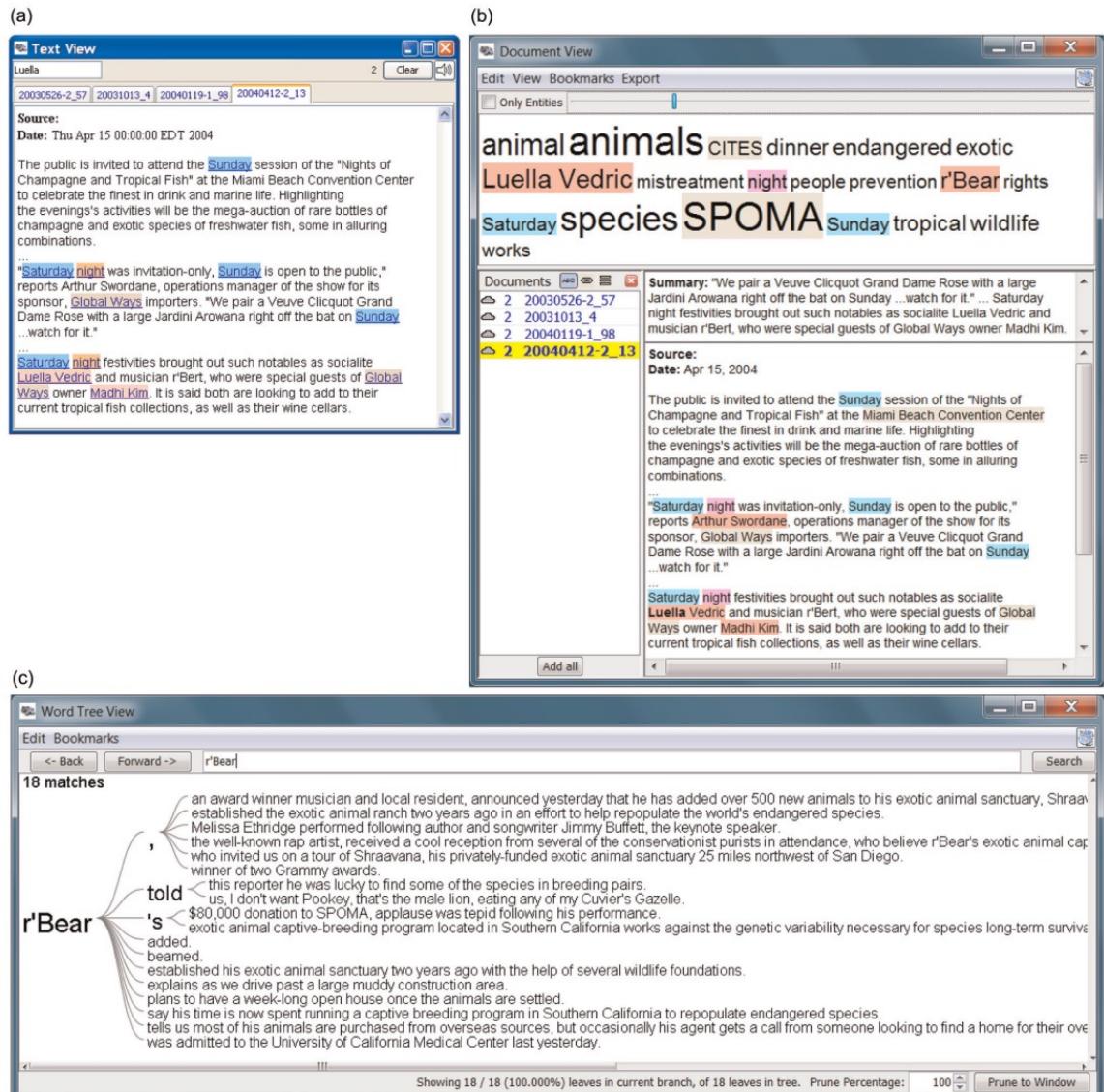
ThemeScape

- Wise J A. The ecological approach to text visualization.
- (Journal of the Association for Information Science and Technology 1999)



Jigsaw

- Developed by Georgia Tech Information Interface Lab
- Document collection visualization tool
- Help finding potential relations:
 - People - people
 - Event - event



Combining Computational Analyses and Interactive Visualization for Document Exploration and Sensemaking in Jigsaw

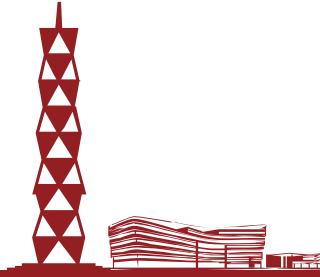
InfoVis / VAST Papers

Carsten Görg*, Zhicheng Liu§, Jaeyeon Kihm†
Jaegul Choo§, Haesun Park§, John Stasko§

*University of Colorado

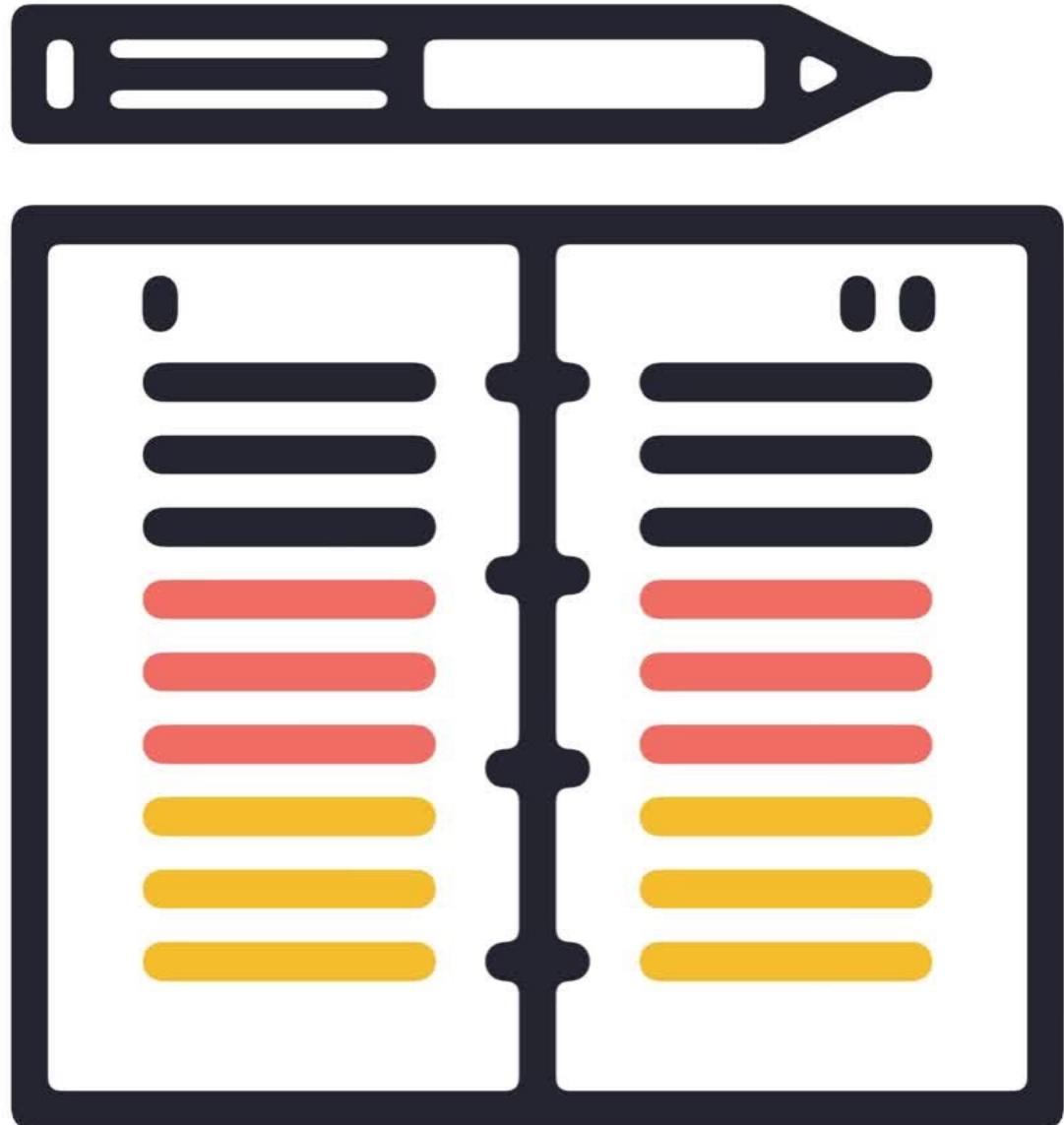
§ Georgia Institute of Technology

†Cornell University



Text Visualization

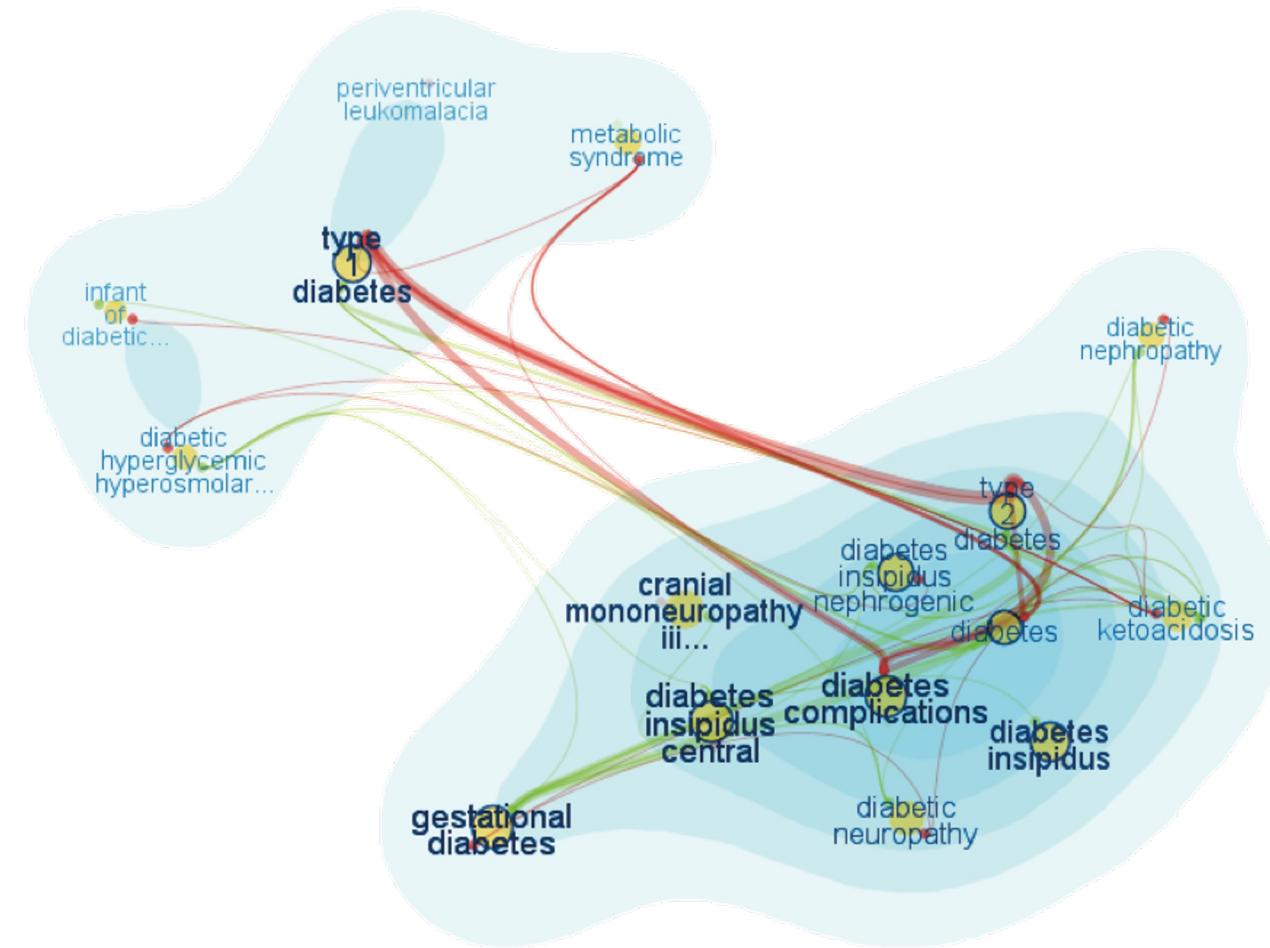
- Categorized based on features
 - Visualizing Document Content
 - Visualizing Document Relations
 - Multi-level Document Visualization



Multi-level Document Visualization

- Multi-level information
 - Context
 - Spatiotemporal Information
 - Author
 - ...
- Visual-guided approach to help discovering insights behind documents





Nan Cao et al. FacetAtlas: Multifaceted Visualization for Rich Text Corpora. (IEEE TVCG 2010)

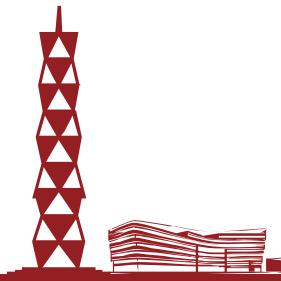


立志成才报国裕民

FacetAtlas: Multifaceted Visualization for Rich Text Corpora

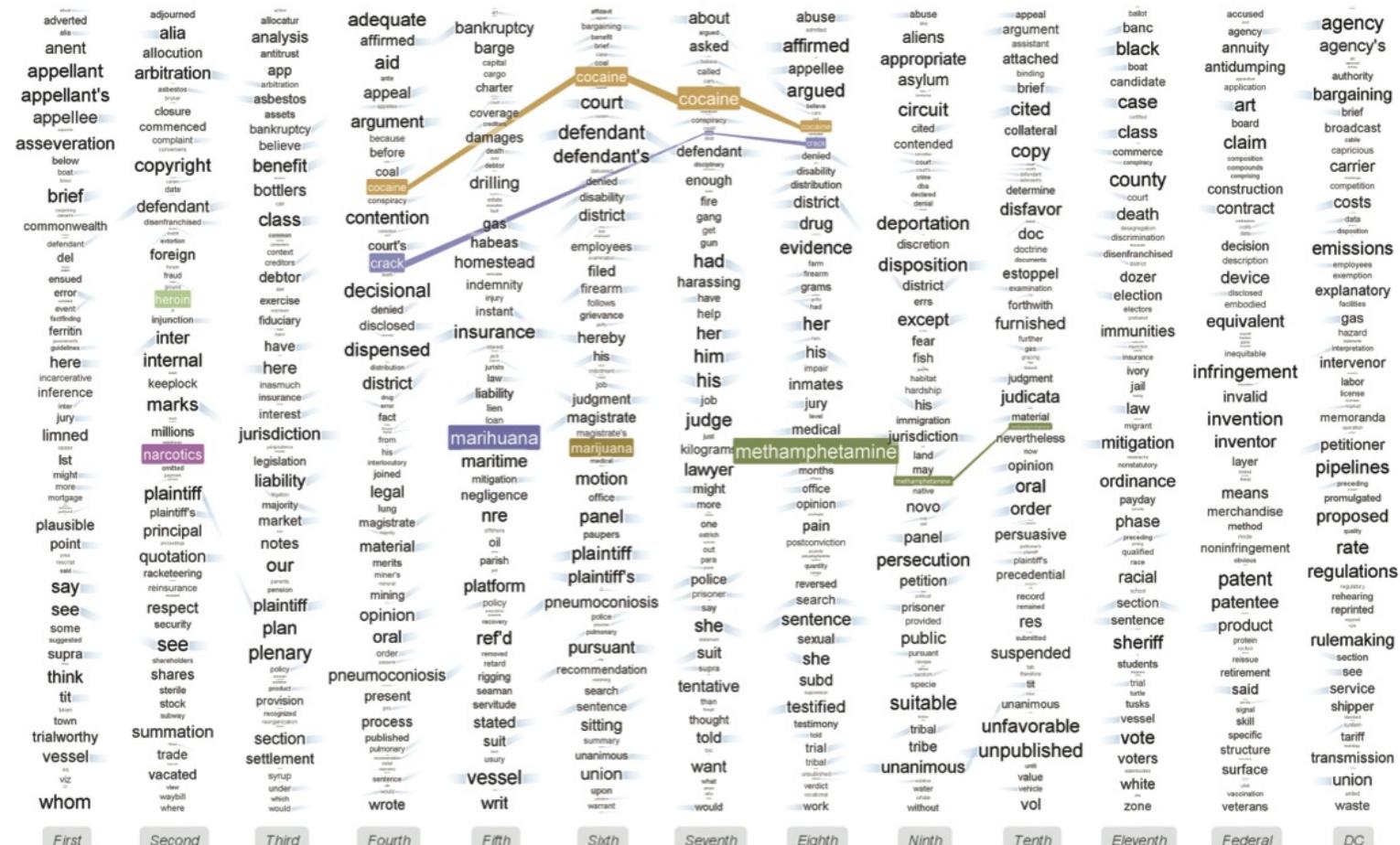
InfoVis 2010

Nan Cao, Jimeng Sun, Yu-Ru Lin, David Gotz,
Shixia Liu, Huamin Qu





Parallel Tag Cloud



Christopher Collins et al. Parallel Tag Clouds to explore and analyze faceted text corpora. (IEEE VAST 2009)



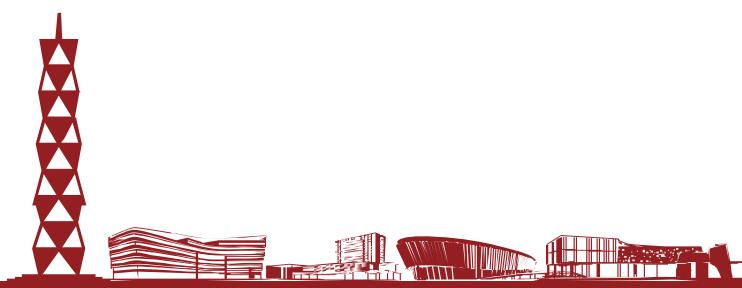
Parallel Tag Clouds to Explore and Analyze Faceted Text Corpora

Christopher Collins
Fernanda B. Viégas
Martin Wattenberg

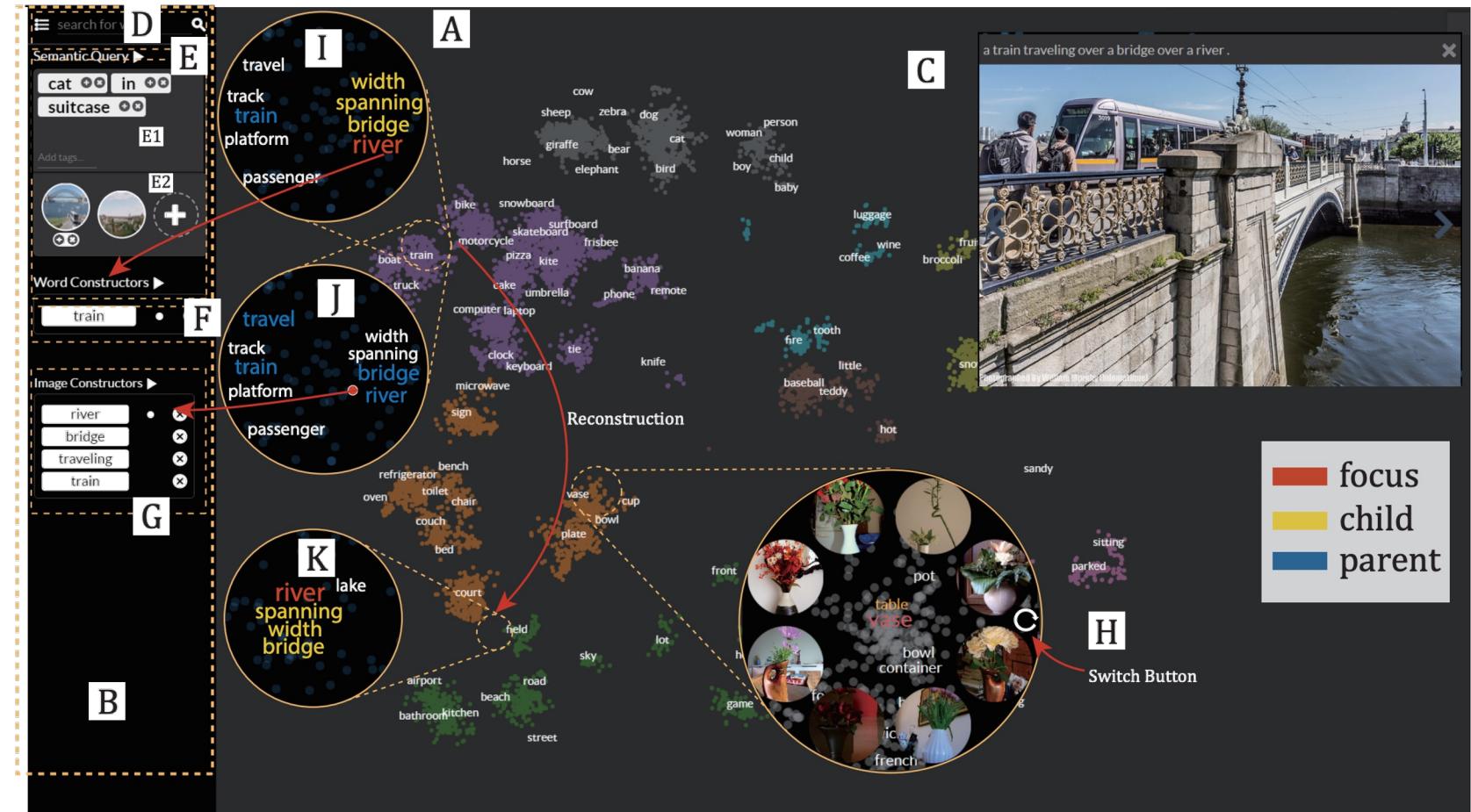


上海科技大学
ShanghaiTech University

Multimedia Visualization



立志成才报国裕民



A Semantic-based Method for Visualizing Large Image Collections

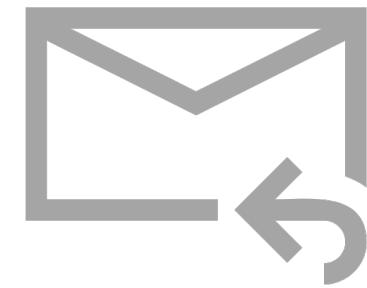
Xiao Xie¹, Xiwen Cai¹, Junpei Zhou¹, Nan Cao², Yingcai Wu¹

Zhejiang University¹
Tongji University²



Quan Li

Questions?
Thank you 😊



liquan@shanghaitech.edu.cn