

# **INTRODUCTION TO INTERNET-BASED RESEARCH**

**FALL 2021**

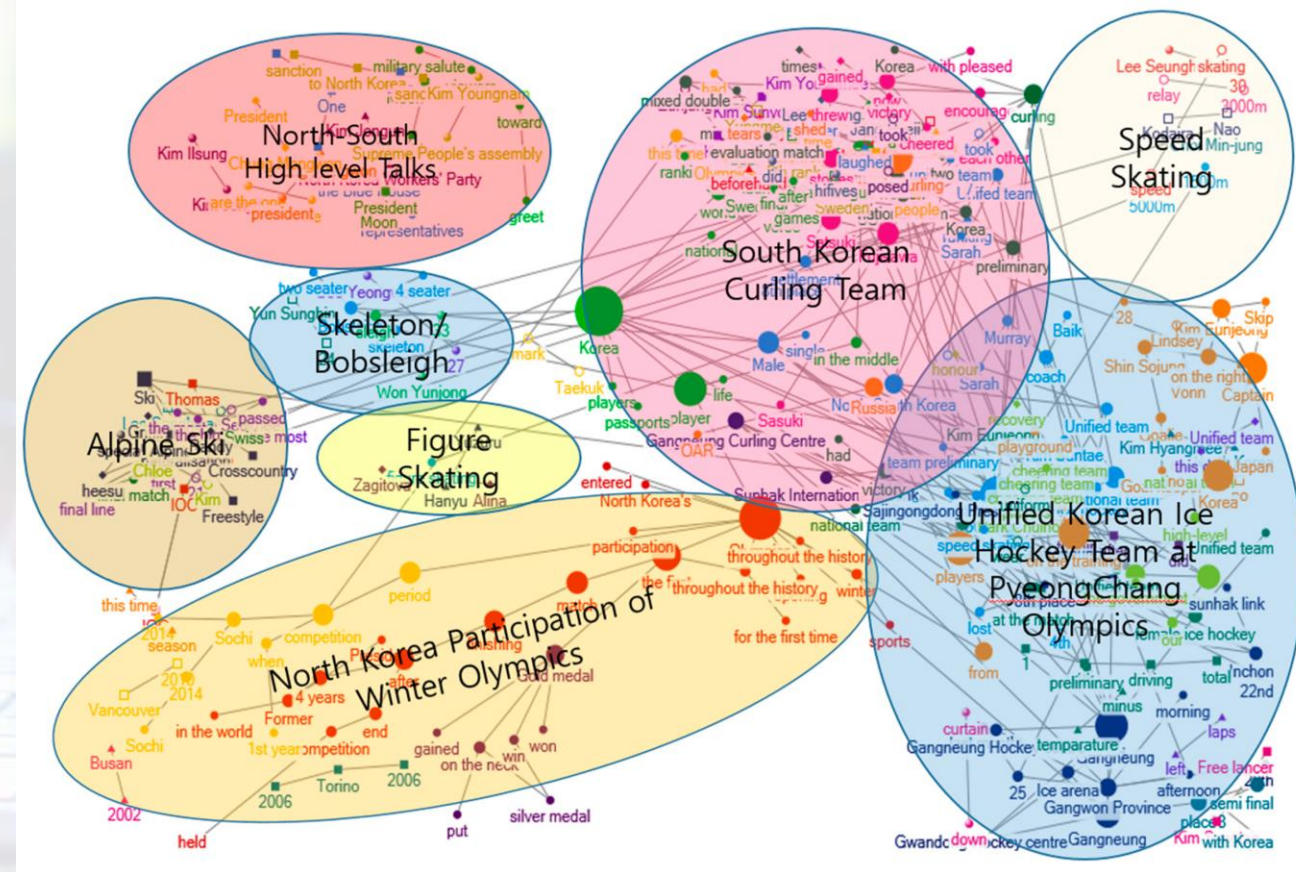
**Week 7: Semantic Network Analysis**

**Maya Higgins. PhD**

# On The Agenda For Today

2

- Introduction
- User-generated Content
- Political Content
- Semantic Networks
- Semantic Network Analysis
- Case Study



# Introduction

3

- Main types of **textual information: Facts and Opinions**
  - Factual statements can imply opinions
- Most current text information processing methods (e.g., web search, text mining) work with **factual information**
- Sentiment/discourse analysis, opinion-mining
  - Computational study of opinions, sentiments, emotions expressed in text
- **Why opinion mining now?** Huge volumes of opinionated text + important for political scientists



# User-generated Content

- **Why are opinions important?**
  - Influence our decisions (booking a hotel, voting for a candidate)
  - In the **past** =>
    - Individuals: Sharing opinions amongst friends, family
    - Businesses: Surveys, focus groups, consultants ...
- **Word-of-mouth on the Web**
  - User-generated content: Opinions are shared in reviews, forums, discussion groups, blogs, social media, etc.
  - Opinions on a global scale, no longer limited to one's circle of friends/small scale surveys



17 hours ago

## Huge let down in service

I ordered a case of dessert wine for a New Years party. Despite using the holiday delivery schedule on their site, not only did nothing turn up in time, but it took nearly a week and three separate customer service inquiries to get an explanation from the company about what had happened. Eventually I got short email stating they'd taken my money despite not having the item in stock! The shop then took another nine days to process my refund.

The company should adjust their system to let people know whether items are in stock or not so they don't create false expectations that goods will arrive on time. There should also be a shorter processing time for refunds and better customer service.

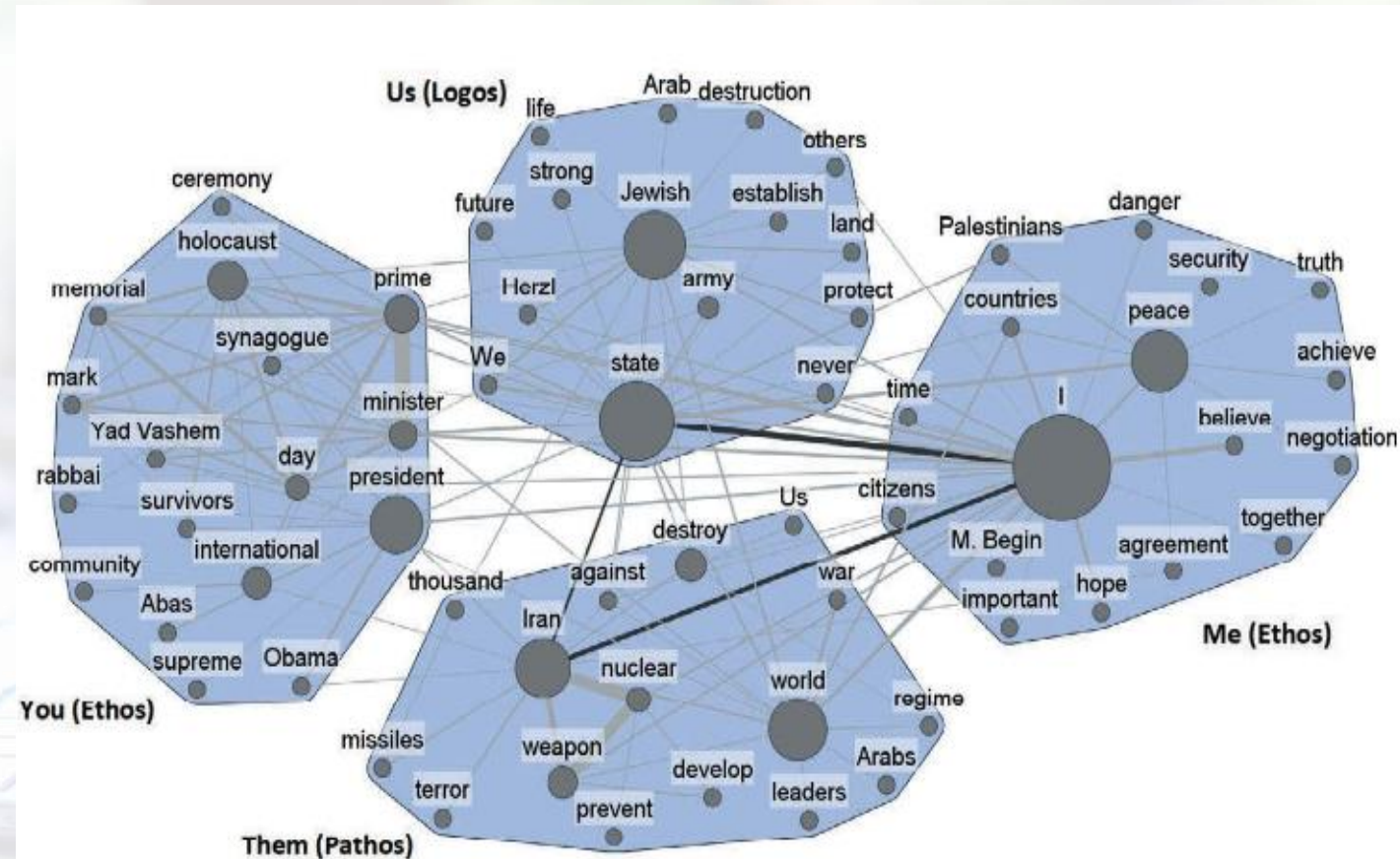
# An Example Review

- *“I bought an iPhone a few days ago. It was such a nice phone. The touch screen was really cool. The voice quality was clear too. Although the battery life was not long, that is ok for me. However, my mother was mad with me as I did not tell her before I bought the phone. She also thought the phone was too expensive and wanted me to return it to the shop. ...”*
- What do we see?
  - **Opinions, targets of opinions, opinion holders**



# User-generated Content

- **Extract and analyse** political opinions
  - Candidates, issues (e.g., policies)
  
- **Compare** opinions across cultures, time on the same issue or topic, e.g., Internet diplomacy
  
- **Opinion propagation** in social contexts



# Political Content

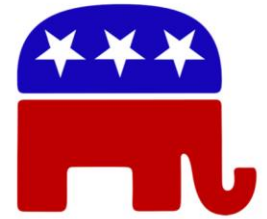
- Manifestos/party platforms
- Public content created by the **political elite** (presidential speeches, press releases, etc.)
- **Why are these important?**
  - Influence our decisions
  - To hold them accountable!
  - Discourse/rhetoric analysis (academic research, practical applications)



## DEMOCRATIC PARTY PLATFORM TOP 10

- 1. IMMIGRATION REFORM:** Defends President Obama's executive action to help DREAMers, parents of citizens and lawful permanent residents avoid deportation. pg 17
- 2. LGBT RIGHTS:** Supports the Supreme Court decision that all Americans, including LGBT persons, have the right to marry an adult of their choice. pg 19
- 3. Supports expanding VOTING RIGHTS** including early voting, automatic voter registration, ending racial gerrymandering and making Election Day a national holiday. pg 24
- 4. CAMPAIGN FINANCE REFORM:** Supports a constitutional amendment to overturn Citizens United putting an end to secret unaccountable money in politics. pg 25
- 5. Committed to protecting and advancing REPRODUCTIVE HEALTH RIGHTS** and continue to oppose – and seek to overturn – Federal and State laws that impede a woman's access to abortions. pg 37
- 6. UNIVERSAL HEALTHCARE:** Democrats believe healthcare is a right not a privilege and will continue the fight until every American has access to quality, affordable healthcare. pg 34
- 7. Supports common sense GUN CONTROL LAWS** such as expanding and strengthening background checks, closing dangerous loopholes and keeping weapons of war off the streets. pg 39
- 8. Recognizes the importance of CLIMATE LEADERSHIP** and is committed to getting 50% of our electric energy from clean energy sources within a decade. Committed to closing the Halliburton loophole that stripped the EPA of its ability to regulate fracking. pgs 27, 28
- 9. Supports a LIVING WAGE** of at least 15 dollars per hour, Equal Pay for women and the Right to Collective Bargaining, including joining or forming a union. pgs 3 & 4
- 10. DEBT-FREE COLLEGE:** Calls for debt-free community colleges and relief from crushing student debt. pg 31

[www.demconvention.com/platform/](http://www.demconvention.com/platform/)



## REPUBLICAN PARTY PLATFORM TOP 10

- 1. Calls for BUILDING A WALL** along the southern border of the country (a length of nearly 2,000 miles). Opposes any form of amnesty and supports mandatory 5-year prison sentences to anyone returning to the US after being deported. pgs 25, 26
- 2. REJECTS SUPREME COURT RULING ON MARRIAGE EQUALITY** and defines marriage as the union between a man and woman only. Calls for allowing businesses and organizations to discriminate based on religious beliefs. pgs 11, 31, 32
- 3. Calls for PROOF OF CITIZENSHIP** when registering to vote and a secure voter ID when voting. pg 16
- 4. Calls for THE REPEAL OF FEDERAL RESTRICTIONS ON CAMPAIGN CONTRIBUTIONS** (McCain/Feingold) pg 12
- 5. Supports replacing FAMILY PLANNING PROGRAMS** for teens with an abstinence-only approach. pg 34
- 6. Seeks to FILL THE SUPREME COURT WITH ULTRA-CONSERVATIVE JUDGES** who will overturn Roe v Wade and the Affordable Care Act. pg 10
- 7. OPPOSES ANY RESTRICTIONS TO GUN RIGHTS:** Calls for the Right to Carry in all 50 States. pgs 12, 13
- 8. CALLS FOR THE OPENING OF FEDERAL LANDS** and outer continental shelf to oil exploration and an expansion of the coal and fracking industries. Supports the Keystone pipeline. pg 19, 20
- 9. REJECTS SEPARATION OF CHURCH AND STATE** in allowing the Display of the Ten Commandments in public areas. Calls for Bible study in schools. Calls for States to offer a Bible curriculum as part of school civics program pgs 12, 33
- 10. Calls for a so-called "Human Life" Constitutional Amendment** that would apply 14th Amendment protections to unborn children. Supports a ban on Federal funding for abortions and healthcare plans that cover abortion rights. pgs 13, 37

[www.gop.com/the-2016-republican-party-platform/](http://www.gop.com/the-2016-republican-party-platform/)

Whose values reflect the issues you care about?  
Get registered, request your ballot – [www.votefromabroad.org](http://www.votefromabroad.org) – and vote Democrat!



**DEMOCRATS ABROAD**  
[www.democratsabroad.org](http://www.democratsabroad.org)

Paid for and authorized by Democrats Abroad.

# Example: Political Content

- Party platform of the republic party (USA), 2016 => 56 pages
- **Map the Discourse:**
  - What topic are mentioned? Emphasized?
  - Which concepts/topics are intertwined?
  - Language (level, keywords, etc)
- **Exploratory** 'first step' for further study of ...

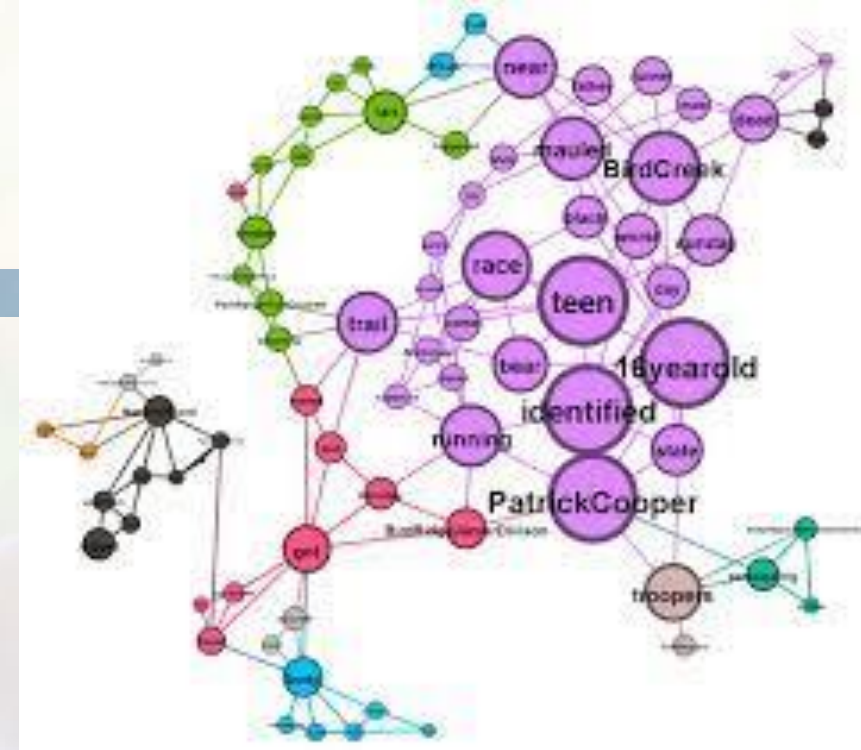
The image shows the cover of the 'Republican Platform 2016' document. The top half is a dark red textured background with the words 'REPUBLICAN PLATFORM' in white, serif, all-caps font, separated by thin white horizontal lines. Below this, the year '2016' is written in a large, white, serif font. The bottom half of the cover features a dark blue background with a pattern of white, five-pointed stars, reminiscent of the American flag.

REPUBLICAN  
PLATFORM  
2016



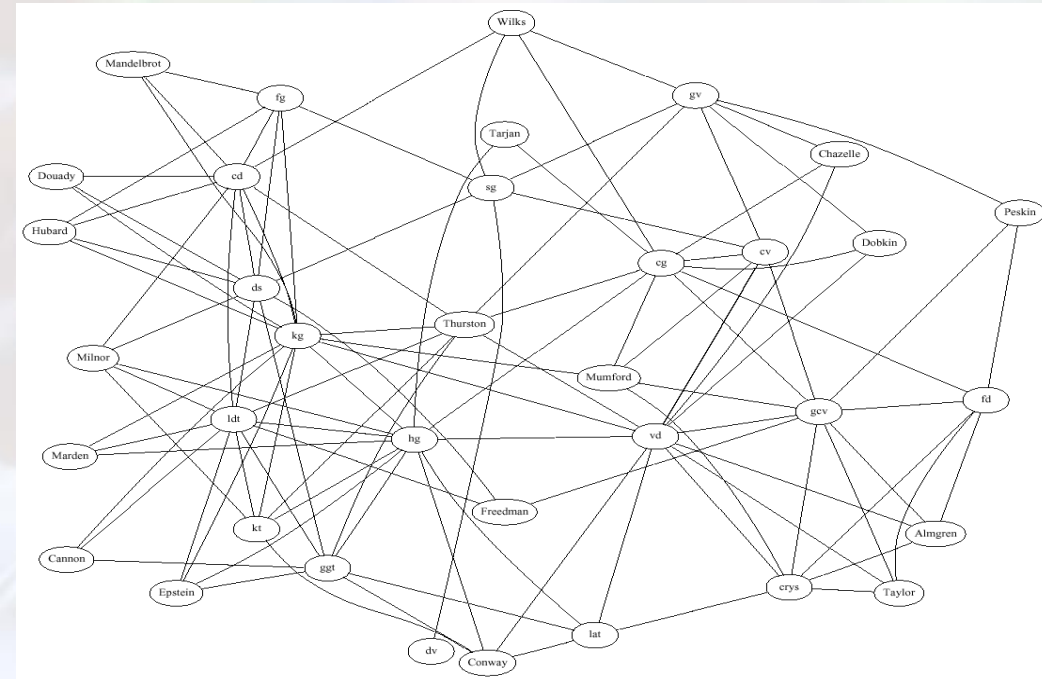
# Semantic Networks

- First introduced in the late 1960s'
  - M. Ross Quillian. "Semantic Memories", In M. M. Minsky (ed), *Semantic Information Processing*, pages 216-270. Cambridge, MA: MIT Press, 1968
- Used to be very **discipline-specific**
- **Semantic network** is a simple representation scheme, which uses a graph of labeled nodes and links to encode knowledge



# Semantic Networks

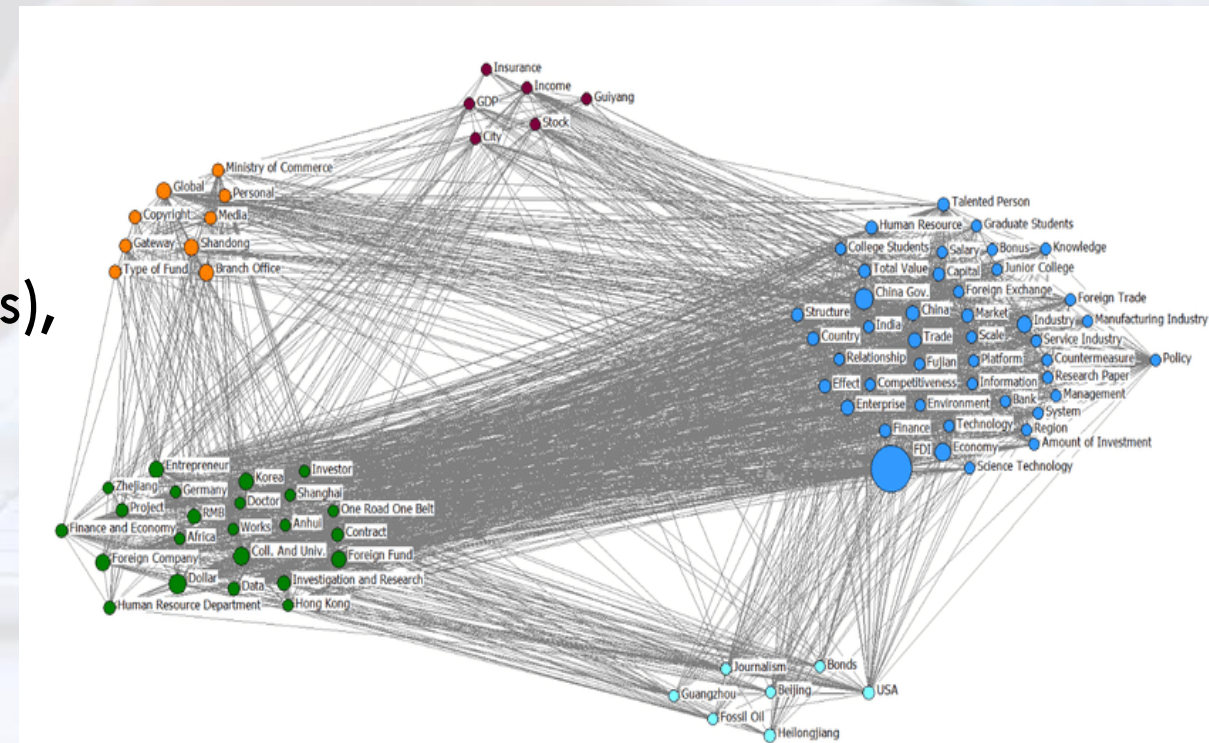
- Visual representation of data
- Each **node** represents an object (proposition/concept/event)
- Each **link** represents a relationship between two objects
- **Goal:**
  - Help humans understand, map the discourse: What terms exist? relation to other terms



# Semantic Network Analysis

## Semantic/Textual Network Analysis

- Transforms unstructured text into a 2D network of words (nodes) + correlations (links)
- Most influential keywords
- Clusters (densely connected communities), representing the main topics in the text
- Study of media coverage, political discourse



# Case Study

12

## The Political Discourse Regarding the War in Afghanistan: A Comparative Analysis Between American Presidents Bush and Obama

- The war in Afghanistan was commenced by US President Bush (2001)
- Obama took office (2009) => A promise of a shift in foreign policy
- Did Obama deliver on his promise to change the narrative regarding the war in Afghanistan
- **Main Research Question: Did a substantial discursive shift has occurred?**
- **Hypothesis: Nope**



# Case Study

13

## The Political Discourse Regarding the War in Afghanistan: A Comparative Analysis Between American Presidents Bush and Obama

- **'The longest war in American history'** => High financial costs + human suffering
- How can the war be justified?
  - Waging war requires public support and legitimacy
- Political discourse considerably affects public discourse and shape public opinion
  - Presidential discourse in the US



# Case Study

14

## The Political Discourse Regarding the War in Afghanistan: A Comparative Analysis Between American Presidents Bush and Obama

- **Research questions:**
  - What are the **main themes** governing the political discourse regarding the War in Afghanistan as depicted in speeches delivered by Bush and Obama (separately)?
  - What are the main **differences/similarities** between the discourses?
  - Can a significant **shift** in the political discourse regarding the War in Afghanistan (from Bush to Obama) be observed?
- **Data:** Presidential speeches
- **Methodology:** Semantic Network Analysis



# Case Study

15

## How were the networks created?

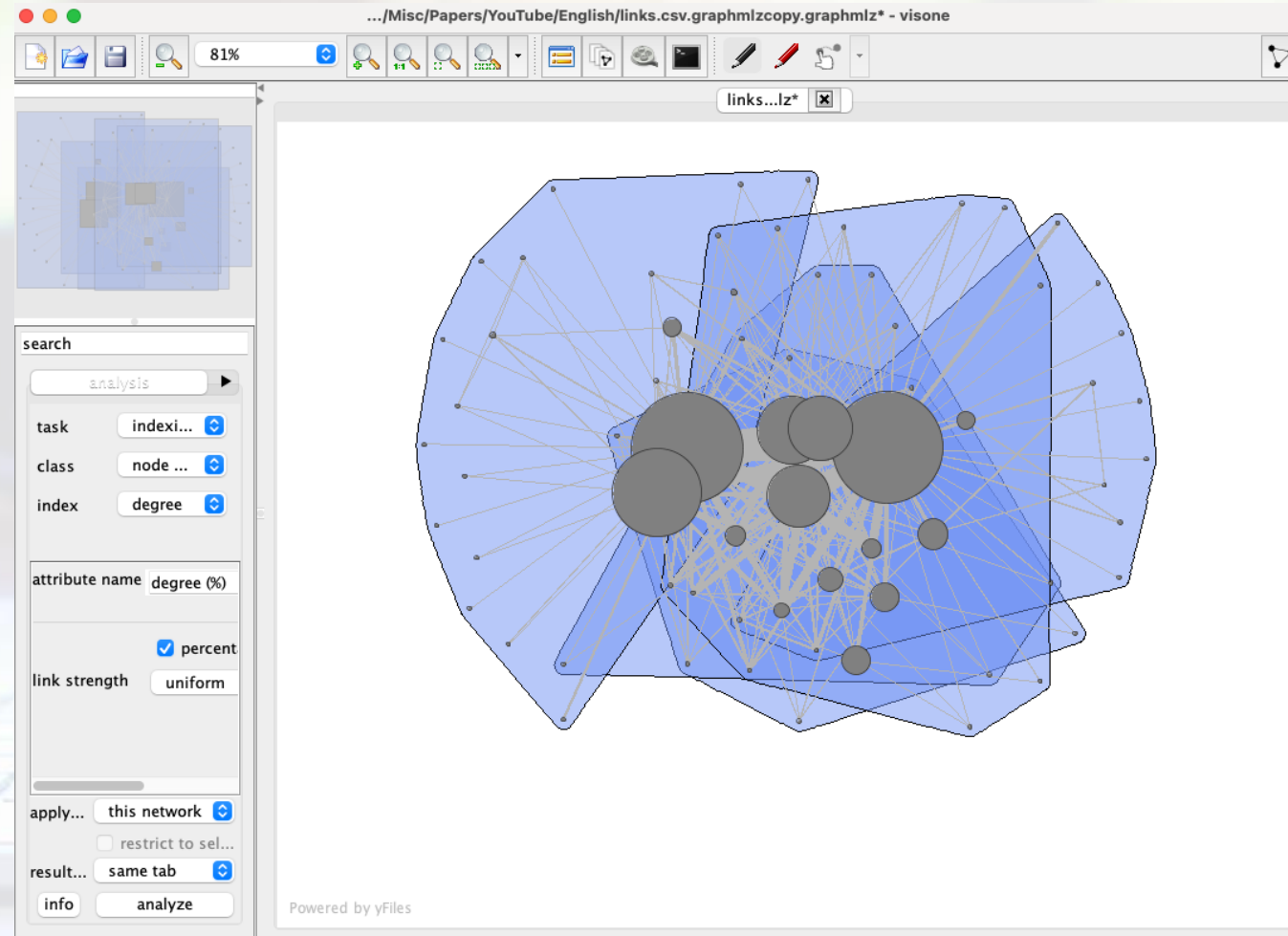
- The most frequently occurring words were identified by a web-based tool
- The top 200 most frequent words were copied into an excel file + **manual** check
- Excluding 'Afghanistan'
- **Stemming** => Different forms of a word were **merged** into a single form (e.g., American, Americans, America's, America => America)
- Words bearing the same meaning were **merged** (e.g., United States & America, America, country, nation)

# Case Study

16

## How were the networks created?

- **Sentences** as the unit of analysis: If two words from the list co-occurred within a sentence, a link was established
- The link-list was then fed into **'Visone'**
- **Cluster analysis** was conducted using the Louvain modularity

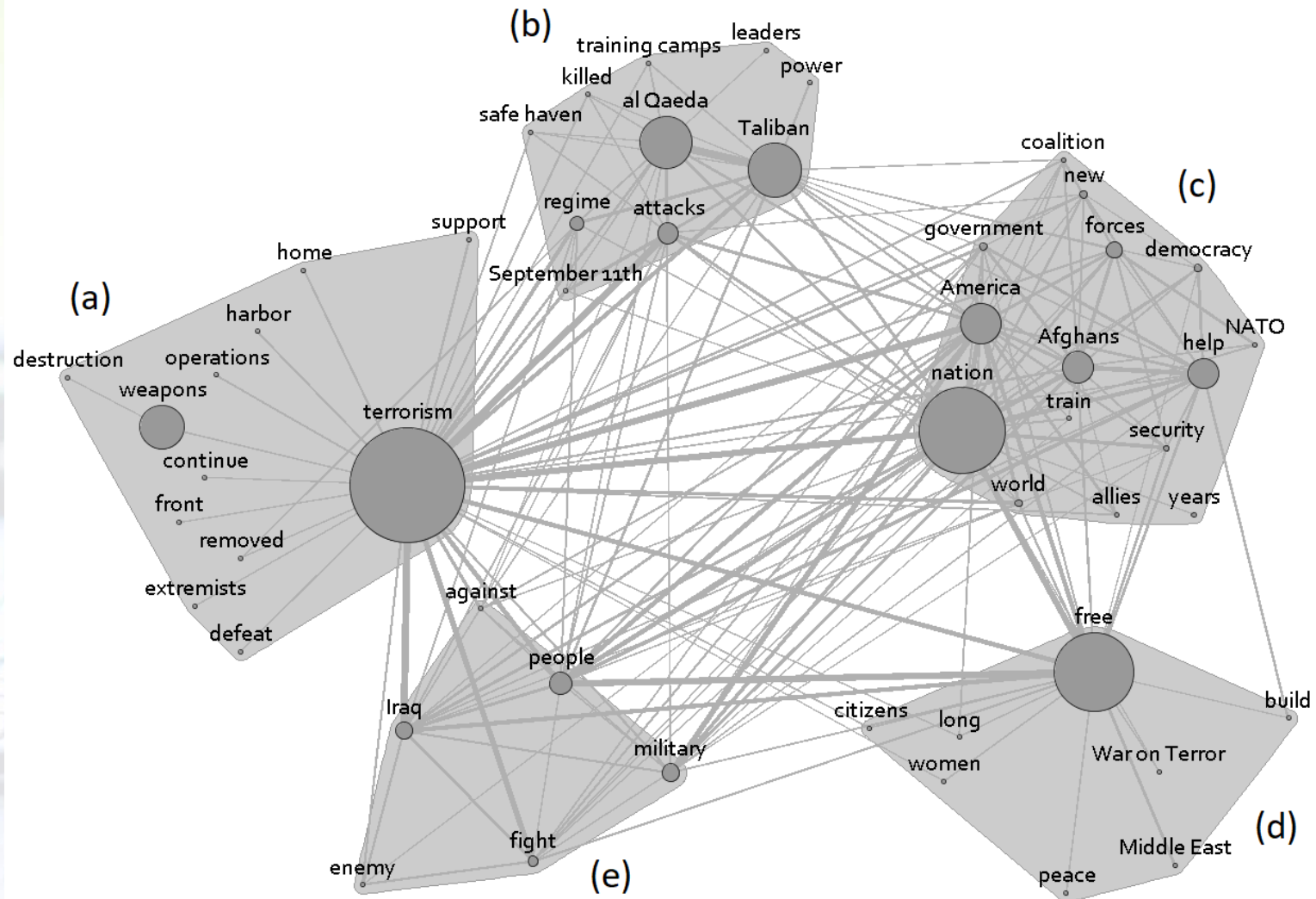




# Case Study

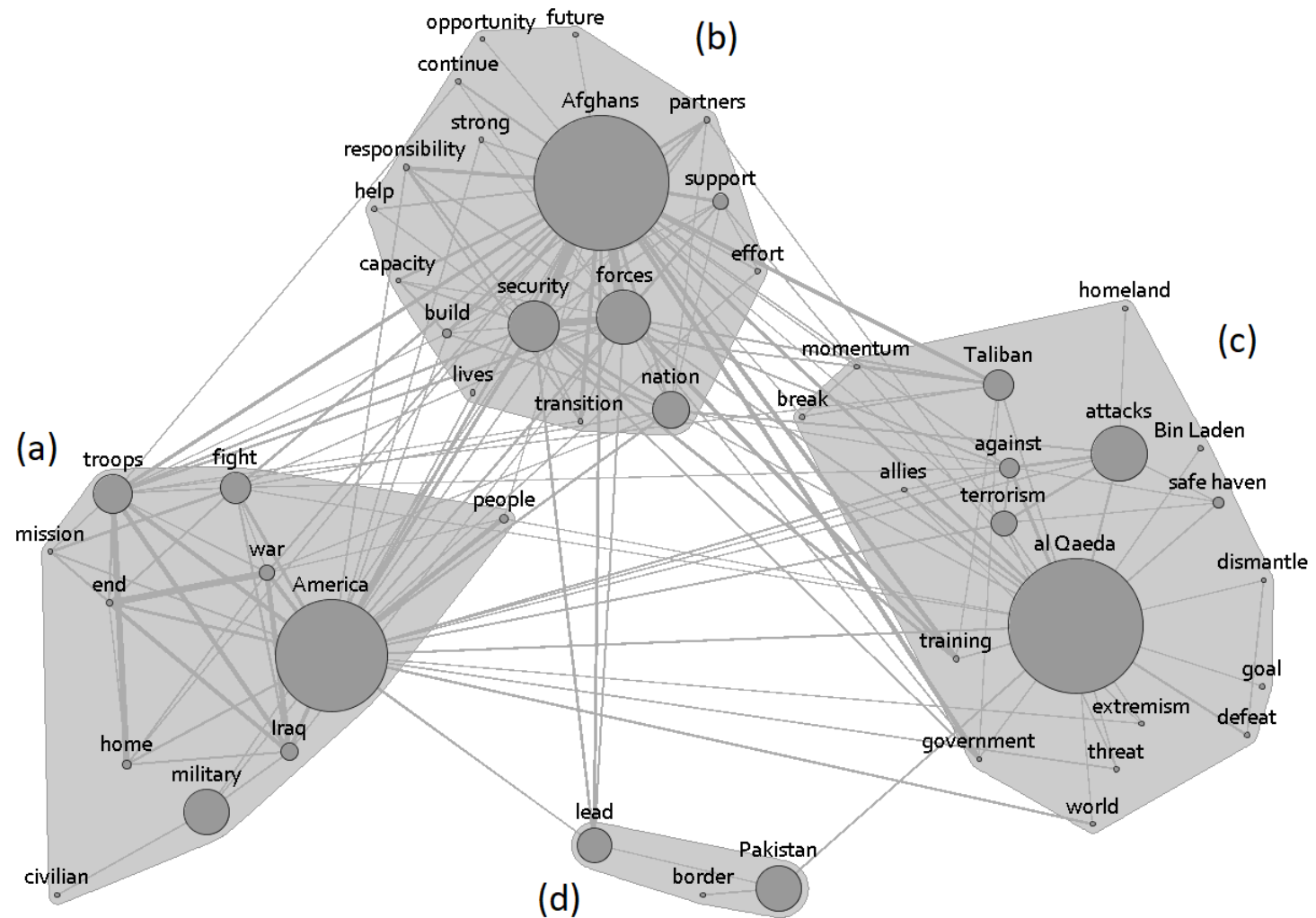
- Nodes are represented according to their 'betweenness' score => The bigger the node, the higher the score

## The Semantic Network of Bush's Speeches Mentioning the War in Afghanistan



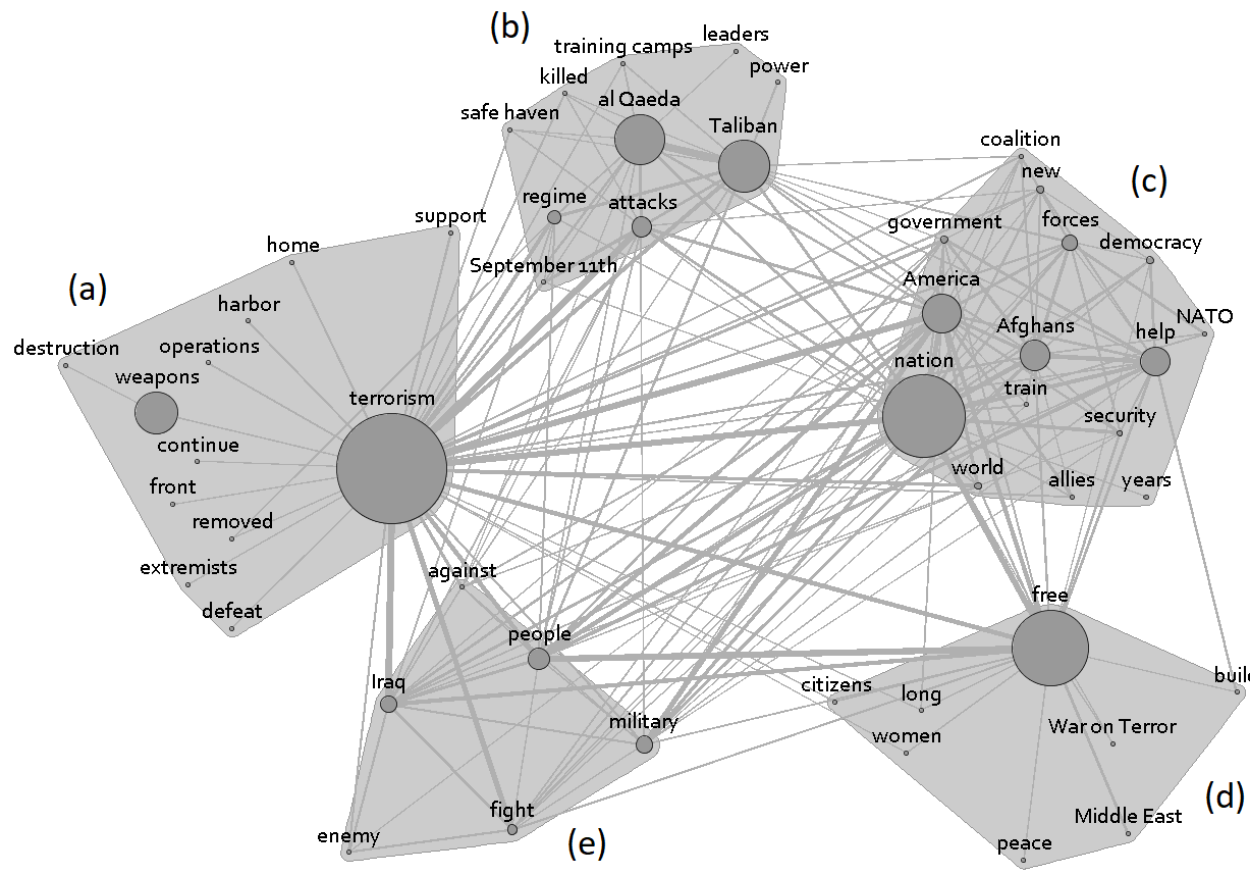
# Case Study

## The Semantic Network of Obama's Speeches Mentioning the War in Afghanistan

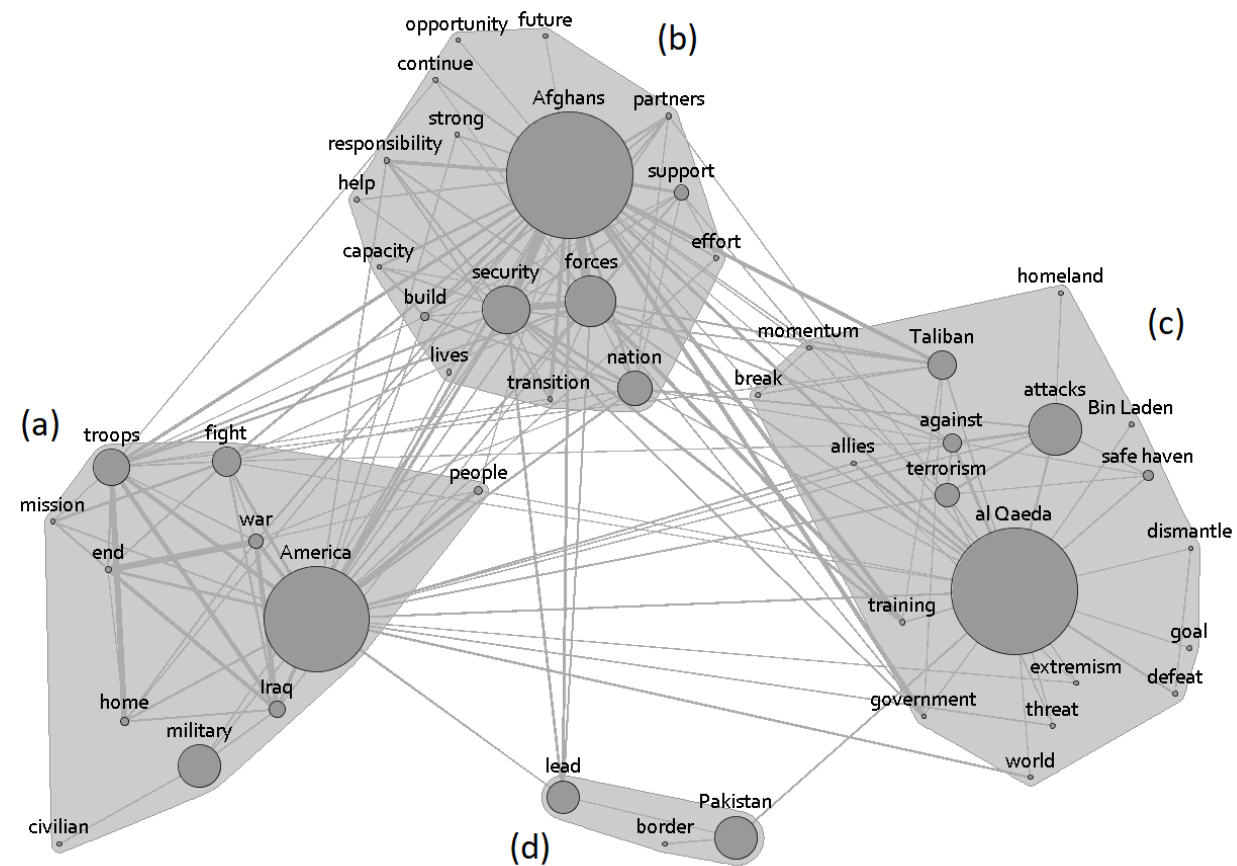


# Case Study

## Bush



## Obama



# Advantages of Semantic nets

- **Easy to visualize**
- **Formal definitions** of semantic networks have been developed
- Related knowledge is **easily clustered**
- **Efficient in space** requirements
  - Objects represented only once
  - Relationships handled by pointers

# Disadvantages of Semantic nets

- **No standard definition** of link and node names => Can be difficult to understand the network, **consistency?**
- Only **binary relationships** can be represented in this model
- **Heuristically weak**
  - Statements such as “Some books are more interesting than others”, “No book is available on this subject”, “If a fiction book is requested, do not consider books on history, health and mathematics” **cannot** be represented in a semantic network
  - Irony is missed

# Next Session...

22

- **Mining Social Networks**

**Thank You For Your Attention!**

**Questions???**