

TASK: Mark on a paper whether the tweets are positive or negative

MARK FOLLOWING TWITTER MESSAGES:

**+ (positive)**

**- (negative)**

**0 (neutral)**

Last night, Donald Trump said not paying taxes was "smart."  
You know what I call it? Unpatriotic.<https://t.co/t0xmBfj7zF> 1.

---

Both candidates were asked about how they'd confront racial  
injustice. Only one had a real answer.<https://t.co/sjnEokckis> 2.

---

When Donald Trump goes low...register to  
vote: <https://t.co/tTgeqxNqYm> <https://t.co/DXz9dEwsZS> 3.

---

Great afternoon in Little Havana with Hispanic community leaders.  
Thank you for your support! #ImWithYou<https://t.co/vxWZ2tyJTF> 4.

---

"I love this country. I'm proud of this country. I want to be a leader  
who brings people together." —Hillary #LoveTrumpsHate 5.

---

"What kind of a person would want to root for 9 million families losing  
their homes? One who should never be president." —Hillary on Trump 6.

---

"I'm really glad my dad never had a contract with Donald Trump." —  
Hillary <https://t.co/un2YiRLEY> 7.

---

# Whats your rate?

## Whats your rate a minute?

.... a day?

.... a month?

.... a period you could devote to the researcher?



Usable data is too big to handle with our amount of time

- Improve your efficiency
- Reduce data size
- Forward analysis process

Data Collection

Data Analysis

# Option 1. improve researcher's efficiency

We are making certain steps, which we are repeating over and over again.

If we automate them, or make them faster, we could accomplish the tasks faster.

## Automatization

Efficiency	
Data Collection	Data Analysis

# Coding the text

Using **physical pen and paper** slower than using **MS Word**.

Using **MS Excel** faster than **MS Word**

Using **Atlas.ti** or **Nvivo** faster than **MS Excel**  
(USUALLY)




Efficiency	
Data Collection	Data Analysis

## Option 2: hire assistants

	<b>Human Assist</b>
Data Collection	Data Analysis

# Cohen's Kappa – Inter-rater reliability

Agreement among researchers for categorizing content entities when categories are known

-  Positive message
-  Negative message
-  Neutral message

	Leading Rater	Assisting Rater
Message 1	+	+
Message 2	0	-
Message 3	0	+
Message 4	-	-
Message 5	+	+
Message 6	0	0
Message 7	+	-
Message 8	+	+

Data Collection

Human Assist

Data Analysis





Positive message



Negative message



Neutral message

5 - CORRECT

3 - WRONG

	Leading Rater	Assisting Rater
Message 1	+	+
Message 2	0	-
Message 3	0	+
Message 4	-	-
Message 5	+	+
Message 6	0	0
Message 7	+	-
Message 8	+	+

$$\text{Simple agreement rate} = \frac{\text{Correct}}{\text{All}} = \frac{5}{8} = 0.625$$

Data Collection

Human Assist

Data Analysis

# Calculate agreement of your Messages

Data Collection

Human Assist

Data Analysis

# Choose the sides (Leading rater and Assisting rater)

**Leading rater** gives directions and explains based on what he/she classified the message

**Assisting rater** captures **Leading rater** directions and tries to classify messages based on them in the future

	Human Assist
Data Collection	Data Analysis

When you work hard, you should not be living in poverty. <a href="https://t.co/86sTOCAkkq">https://t.co/86sTOCAkkq</a>	8.
Well, now they're saying that I not only won the NBC Presidential Forum, but last night the big debate. Nice!	9.
"We may have a record-setting turnout in this election. We could have the biggest turnout we've ever had." <a href="https://t.co/tTgeqxNqYm">https://t.co/tTgeqxNqYm</a>	10.
"Donald just criticized me for preparing for this debate. And you know what else I prepared for? I prepared to be... <a href="https://t.co/QjB01msPaT">https://t.co/QjB01msPaT</a>	11.
Hillary was the winner of the first presidential debate—and newspapers across America agree. #SheWon... <a href="https://t.co/oDiUHd0eXO">https://t.co/oDiUHd0eXO</a>	12.
Today is #NationalVoterRegistrationDay - I was beaten, left bloody & unconscious so that every person could register and vote. Do your part.	13.
Such a great honor. Final debate polls are in - and the MOVEMENT wins! #AmericaFirst #MAGA #ImWithYou... <a href="https://t.co/DV1BKMwHEM">https://t.co/DV1BKMwHEM</a>	14.
"Register. Vote. Go out and get engaged!" — @JoeBiden #NationalVoterRegistrationDay <a href="https://t.co/tTgeqxNqYm">https://t.co/tTgeqxNqYm</a>	15.

How to include randomness or agreement by Chance?

## Agreement Matrix

	Leading Rater	Assisting Rater
Message 1	+	+
Message 2	0	-
Message 3	0	+
Message 4	-	-
Message 5	+	+
Message 6	0	0
Message 7	+	-
Message 8	+	+

		Leading Rater		
		Positive	Neutral	Negative
Assisting Rater	Positive	3	1	0
	Neutral	0	1	0
	Negative	1	1	1

Data Collection

Human Assist

Data Analysis

Expected by chance agreement

		Leading Rater		
		Positive	Neutral	Negative
Assisting Rater	Positive	3	1	0
	Neutral	0	1	0
	Negative	1	1	1

+

Positive message Expected by chance =  $\frac{(3 + 1 + 0) * (3 + 0 + 1)}{3+1+0+0+1+0+1+1+1} = 2$

○

Neutral message Expected by chance =  $\frac{(0 + 1 + 0) * (1 + 1 + 1)}{3+1+0+0+1+0+1+1+1} = 0.375$

—

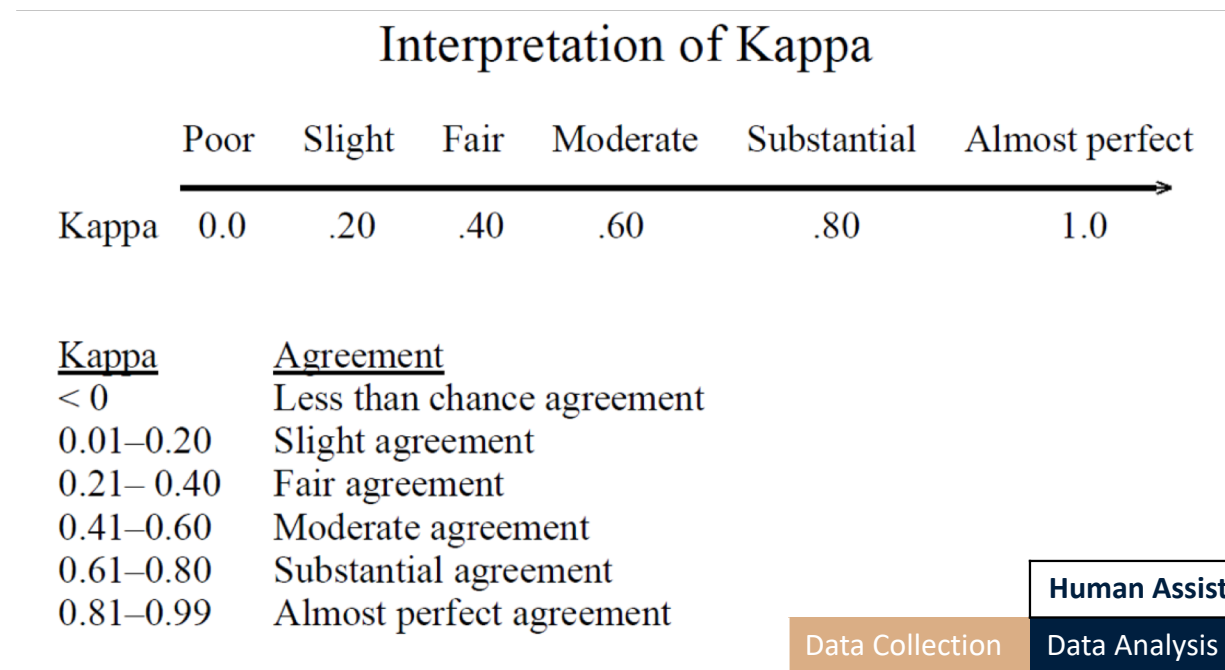
Negative message Expected by chance =  $\frac{(1 + 1 + 1) * (0 + 0 + 1)}{3+1+0+0+1+0+1+1+1} = 0.375$

↖ Total Sum of Matrix

**Expected Agreement by Chance = Positive** expected by chance + **Neutral...** + **Negative...** = 2 + 0.375 + 0.375 = **2.75**

(All agreements – by chance) / (number of cases – by chance)= **0.43**

	Leading Rater	Assisting Rater
Message 1	+	+
Message 2	0	-
Message 3	0	+
Message 4	-	-
Message 5	+	+
Message 6	0	0
Message 7	+	-
Message 8	+	+



# Hire computer as assistant

## Artificial intelligence

Computers ability to do tasks traditionally in the domain of humans

But how do the computers  
solve the tasks, and comes  
the explanations of the  
world?



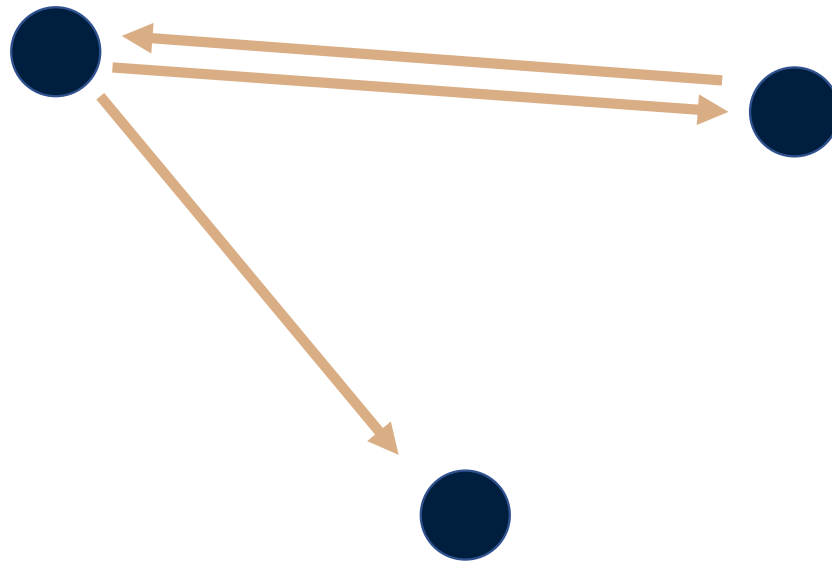


# Social Network analysis (SNA)



	<b>SNA</b>
	Comp. Assist
Data Collection	Data Analysis

Social Networks is a combination of **NODES**, and **LINKS** between them



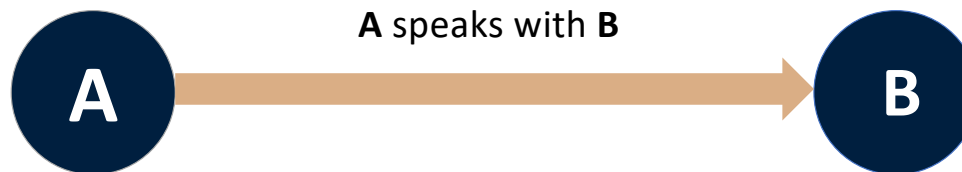
Person/book/word/product/website...

Talking/coupling/Friends List

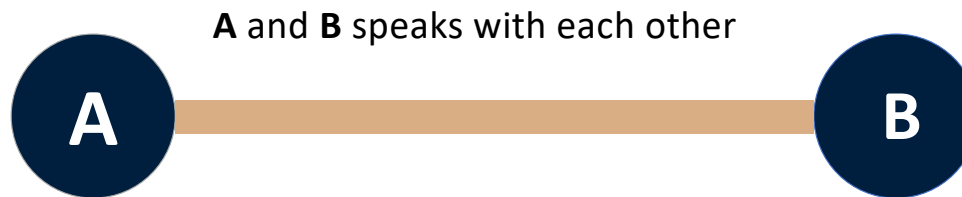
<http://cvcedhlab.hypotheses.org/106>

Data Collection	SNA
	Comp. Assist
	Data Analysis

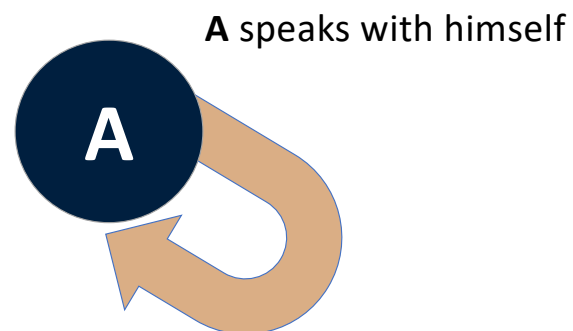
Directed link



Undirected link



Self-directed link



SNA

Comp. Assist

Data Collection

Data Analysis

The main goal of SNA is detecting and interpreting patterns of social ties among actors

	<b>SNA</b>
	Comp. Assist
Data Collection	Data Analysis



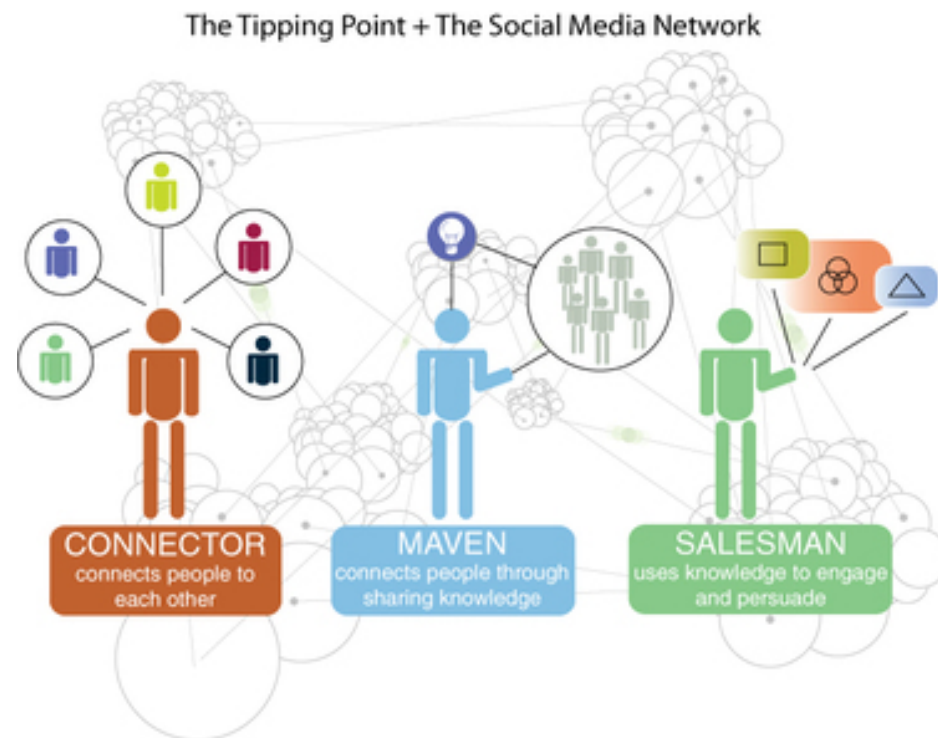
To understand networks and their participants, we evaluate the location of actors in the network.

Data Collection	SNA
	Comp. Assist
	Data Analysis

SNA measures give us insight into the various roles and groupings in a network --

Who are the...

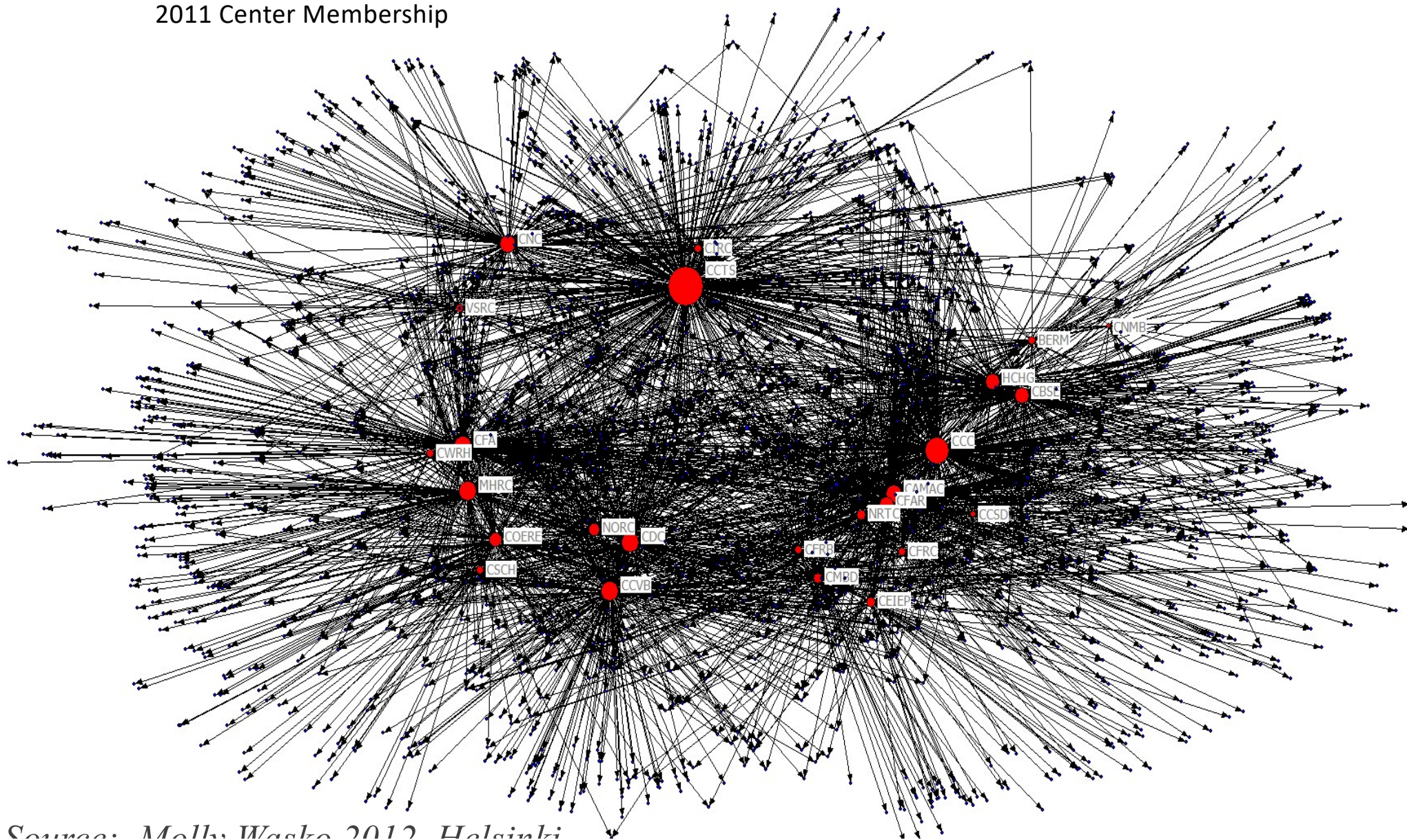
- **Connectors**
- **Mavens**
- **Salesmen**
- **Leaders**
- *Isolates*



SNA	
Comp. Assist	
Data Collection	Data Analysis



## 2011 Center Membership



Source: *Molby, Wasko 2012 Helsinki*



Using quantitative metrics instead of the eyes to evaluate / analyze the network

**Network reach** – how short is path to reach any member of the network (in perfect situation - 2 steps)

**Degree centrality** – how many connections each node has

**Betweenness centrality** – how many nodes have loose connection (disconnecting other nodes from the network)

**Closeness centrality** – which are the nodes that have shortest path to reach all other nodes (nodes that need shortest distance to reach others)

<http://cvcedhlab.hypotheses.org/106>

	SNA
	Comp. Assist
Data Collection	Data Analysis

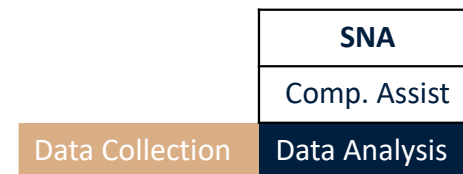
## Two major ways to supply data to Social Network Analysis tool

Adjacency matrix

	_unit2_	Martin	Peter	Beer	The	Jane	Jason
1	Martin	3	1	2	1	2	1
2	Peter	1	1	1	0	0	0
3	Beer	2	1	2	1	1	0
4	The	1	0	1	1	1	0
5	Jane	2	0	1	1	2	1
6	Jason	1	0	0	0	1	1

Edge List

In	Out	Strength
Martin	Peter	1
Martin	Beer	2
Jane	Jason	1
...	...	...



# Further readings and references

- Cohen's Kappa:
  - <http://psych.unl.edu/psycrs/handcomp/hckappa.PDF>
- Social Network Analysis:
  - Definitions <http://cvcedhlab.hypotheses.org/106>
  - Free tool Ucinet <https://sites.google.com/site/ucinetsoftware/home>