

Social Network Analysis

SICSS 2019 | Network analysis: then and now | Dr. Oliver Posegga



Social network analysis has been around for a while – much of early the seminal work dates back to the early 1900s

EMOTIONS MAPPED BY NEW GEOGRAPHY

Charts Seek to Portray the Psychological Currents of Human Relationships.

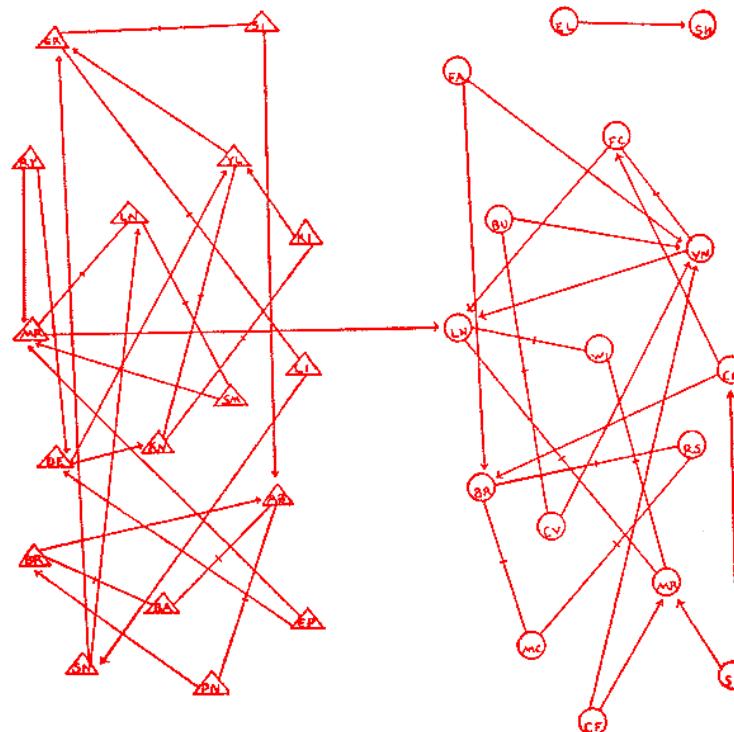
FIRST STUDIES EXHIBITED

Colored Lines Show Likes and Dislikes of Individuals and of Groups.

MANY MISFITS REVEALED

Dr. J. L. Moreno Calculates There Are 10 to 15 Million Isolated Individuals in Nation.

A new science, named psychological geography, which aims to chart the emotional currents, cross-currents and under-currents of human relationships in a community, was introduced here yesterday at the scientific exhibit of the Medical Society of the State of New York, which opens its 127th annual meeting here today at the Waldorf-Astoria.

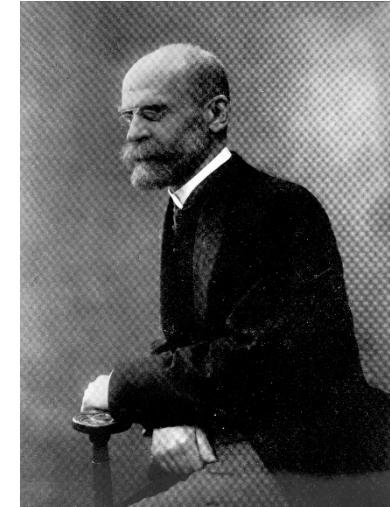
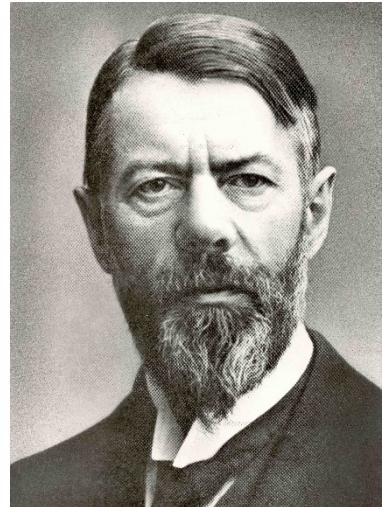
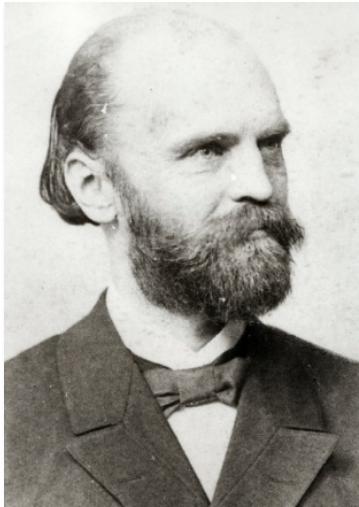
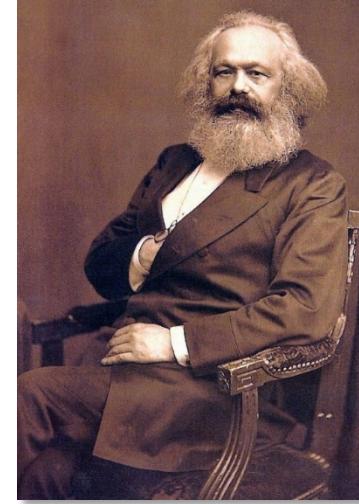


Source: New York Times, 3. April 1933

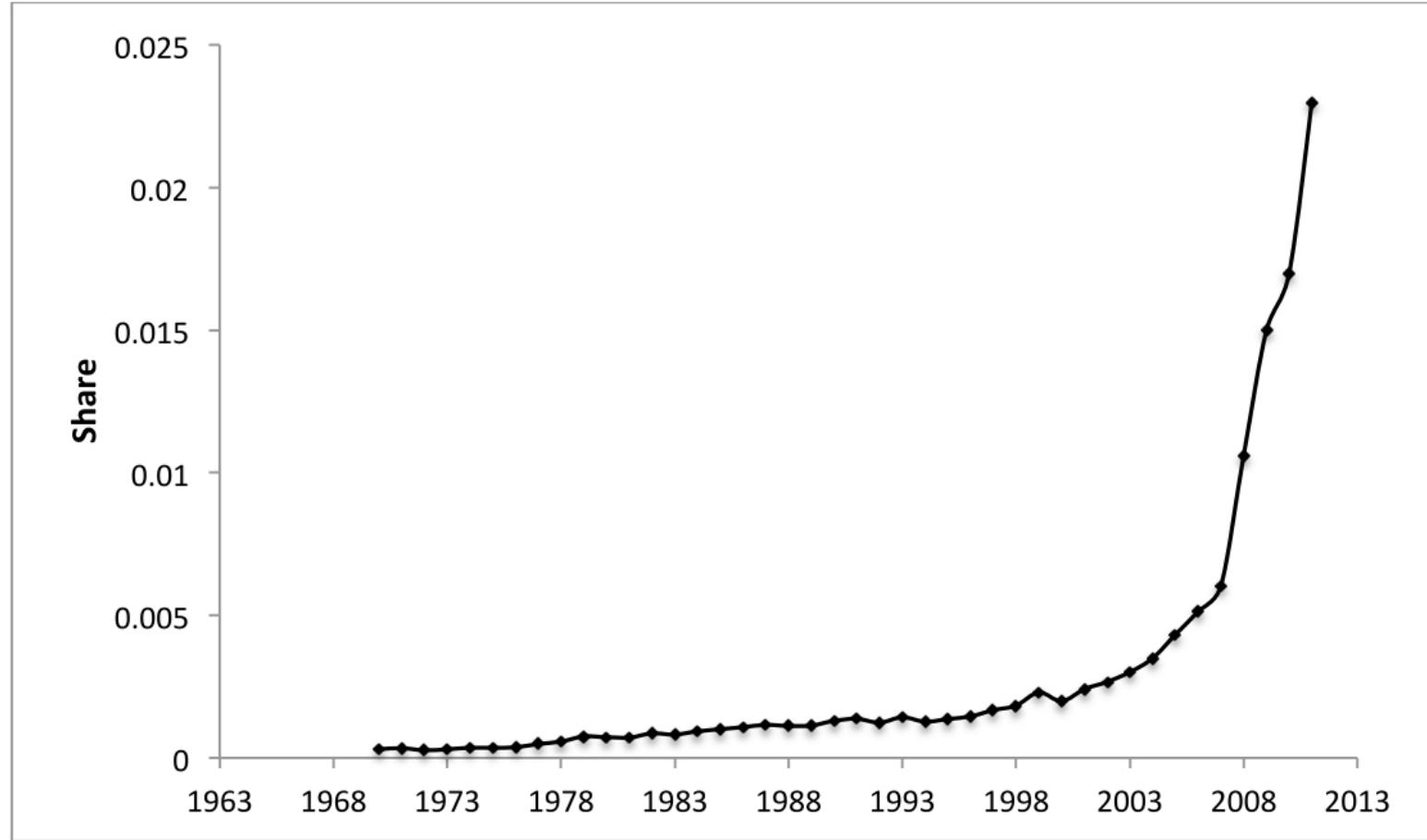
It's origins and fundamental influences date back even further

„Die Gesellschaft besteht nicht aus Individuen, sondern drückt die Summe der Beziehungen, Verhältnisse aus, worin diese Individuen zueinander stehn.“

– Marx, K. (1857/58) *Grundrisse der Kritik der politischen Ökonomie.*



Network analysis has gained an increasing amount of interdisciplinary attention over the last 20 years

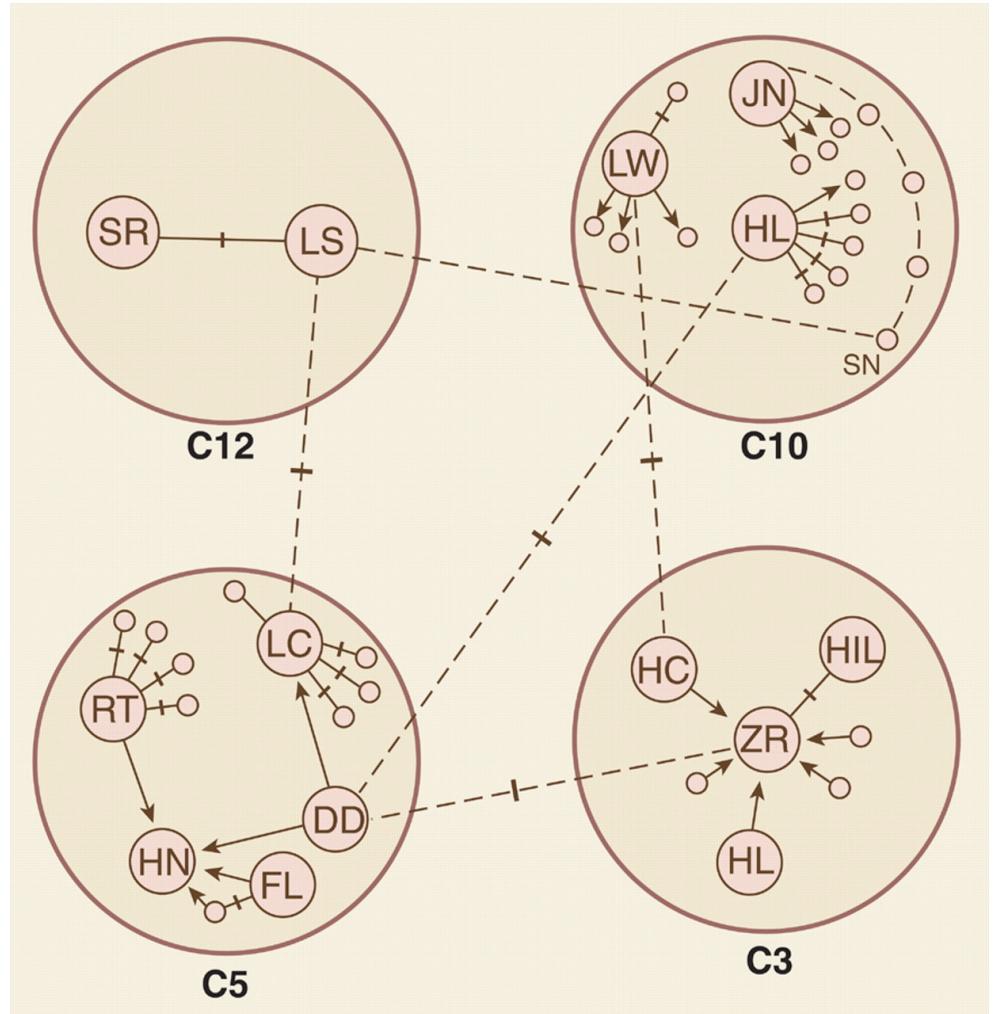


Source: Google Scholar

Fast forward: Networks today



Something happened...



(Borgatti et al., 2009)



Something happened...

Data collection then...

Name five people you would ask for help in case of an emergency:

1. _____
2. _____
3. _____
4. _____
5. _____

Data collection now...

Get the active user's friends.

```
friends = graph.get_connections(id='me', connection_name='friends')
```

Traditional methods of data collection

- Surveys and questionnaires
- Interviews
- Archives
- Observations
- Experiments
- Others

NETWORK DATA AND MEASUREMENT

Peter V. Marsden

Department of Sociology, Harvard University, Cambridge, Massachusetts 02138

KEY WORDS: network, data, survey, measurement

Abstract

Data on social networks may be gathered for all ties linking elements of a closed population (“complete” network data) or for the sets of ties surrounding sampled individual units (“egocentric” network data). Network data have been obtained via surveys and questionnaires, archives, observation, diaries, electronic traces, and experiments. Most methodological research on data quality concerns surveys and questionnaires. The question of the accuracy with which informants can provide data on their network ties is nontrivial, but survey methods can make some claim to reliability. Unresolved issues include whether to measure perceived social ties or actual exchanges, how to treat temporal elements in the definition of relationships, and whether to seek accurate descriptions or reliable indicators. Continued research on data quality is needed; beyond improved samples and further investigation of the informant accuracy/reliability issue, this should cover common indices of network structure, address the consequences of sampling portions of a network, and examine the robustness of indicators of network structure and position to both random and nonrandom errors of measurement.

INTRODUCTION

Progress in the study of social networks has been rapid over the past two decades. The network approach, developed out of analytical insights from social anthropology and methodological leads from sociometry (Shulman 1976, Wellman 1983) conceives of social structure as patterns of specifiable relations joining social units—including both individual actors and collectives

Traditional issues: asking people about their relationships is difficult

"people do not know, with any acceptable accuracy, to whom they talk over any given period of time."

(Bernard et al., 1981, p. 15)

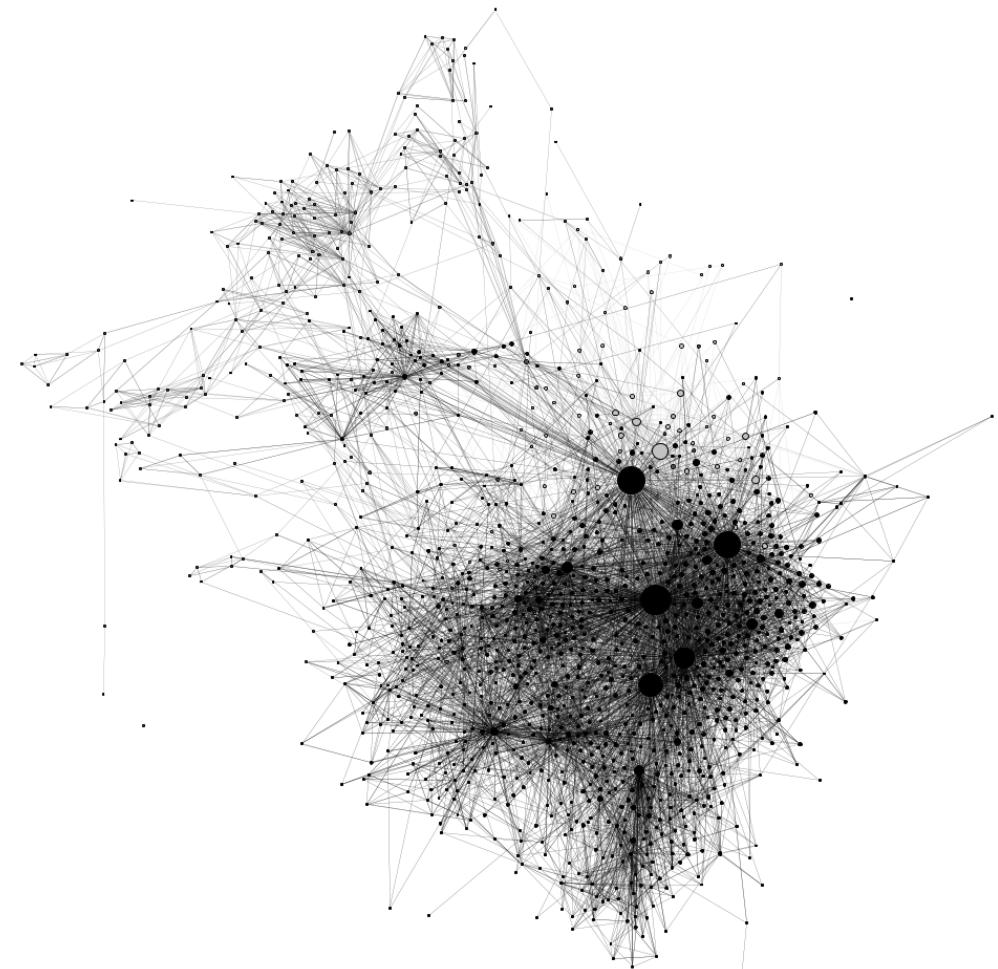
Traditional issues: *Observing social relationships is difficult*

“monitoring a communications network, especially one of any size and spatial distribution, is almost impossible for one or even a few observers.”

(Kandel, 1974)

Some other limitations

- Assumption of quasi-objective existing ties beyond respondent recognition
- Subjectivity of self-reported network data
- Reliability of recall and recognition data
- Digitally enabled human interaction and social relationships are difficult to account for
- Limitation to small to medium sized networks
- Overemphasis of cross-sectional data and static relationships
- ...



Digital trace data to the rescue!

Definition

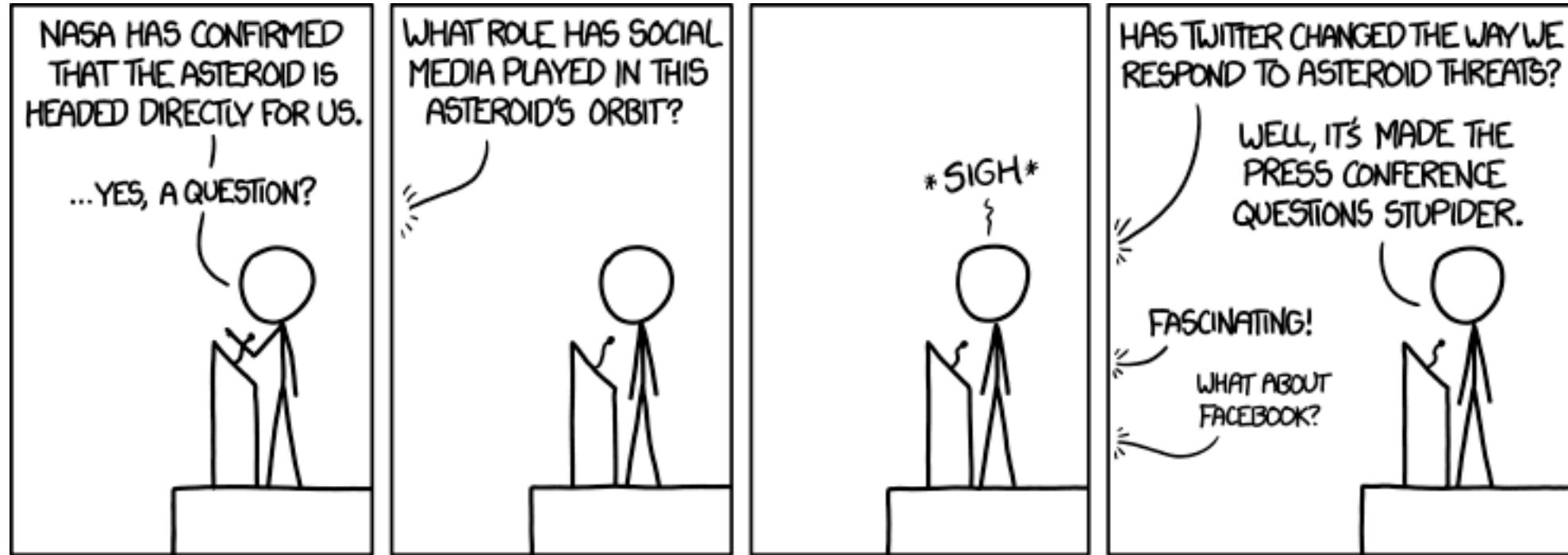
“... records of activity (trace data) undertaken through an online information system (thus, digital). A trace is a mark left as a sign of passage; it is recorded evidence that something has occurred in the past.” – (Howison et al., 2011)

Unique characteristics

- It is found rather than reported
- It is event-based rather than summarized
- It is longitudinal rather than cross-sectional



The measurement revolution



<https://xkcd.com/1239/>

The measurement revolution

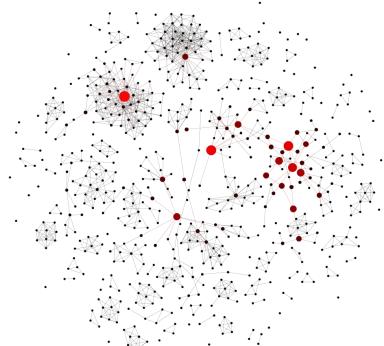
The measurement revolution ...

"[W]e are witnessing a revolution in the measurement of collective human behavior and the beginnings of a new research area - one that analyzes and builds theories of large social systems by using their reflections in massive datasets. " – Kleinberg (2008)

... is as exciting as it is challenging.

"A measurement revolution is an exciting time, but it is also a time that calls for reflection; with opportunities come risks, especially when methods developed in one context are applied in new contexts." – Howison et al. (2011)

Information
System



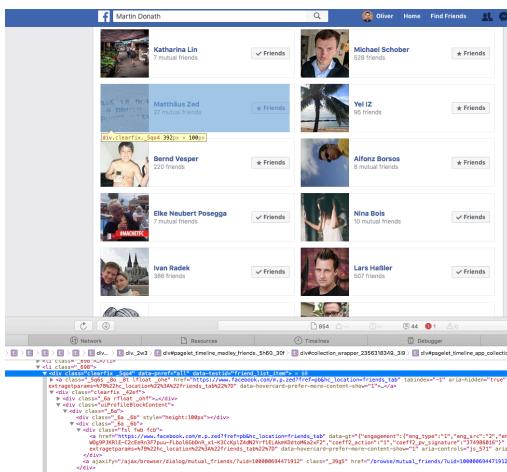
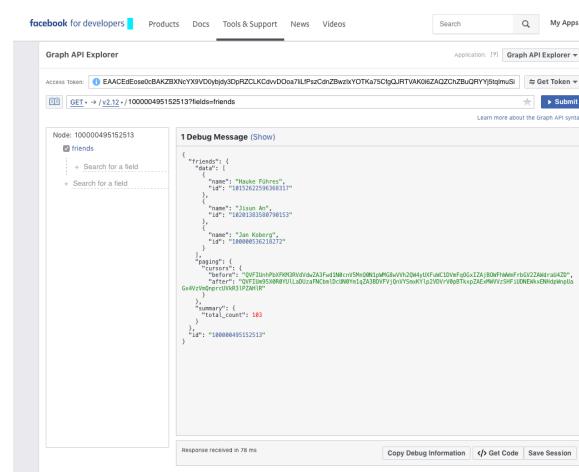
Observing social relationships using mobile sensor platforms

- Sociometric badges
- Mobile phones
- Wearables



Observing social relationships using social media data

- Screen Scraping
- Application Programming Interfaces (APIs)
- Applications

The screenshot shows the Facebook Graph API Explorer interface with a 'Debug Message (Show)' section. It displays a JSON object representing friend data, including names and IDs.

```

{
  "friends": [
    {
      "name": "Katharina Lin",
      "id": "100000495152513"
    },
    {
      "name": "Michael Schöber",
      "id": "100000495152513"
    },
    {
      "name": "Matthias Zed",
      "id": "100000495152513"
    },
    {
      "name": "Yel IZ",
      "id": "100000495152513"
    },
    {
      "name": "Bernd Vesper",
      "id": "100000495152513"
    },
    {
      "name": "Alfonz Borsos",
      "id": "100000495152513"
    },
    {
      "name": "Elke Neubert Posegga",
      "id": "100000495152513"
    },
    {
      "name": "Nina Bois",
      "id": "100000495152513"
    },
    {
      "name": "Ivan Radek",
      "id": "100000495152513"
    },
    {
      "name": "Lars Häßler",
      "id": "100000495152513"
    }
  ]
}
  
```

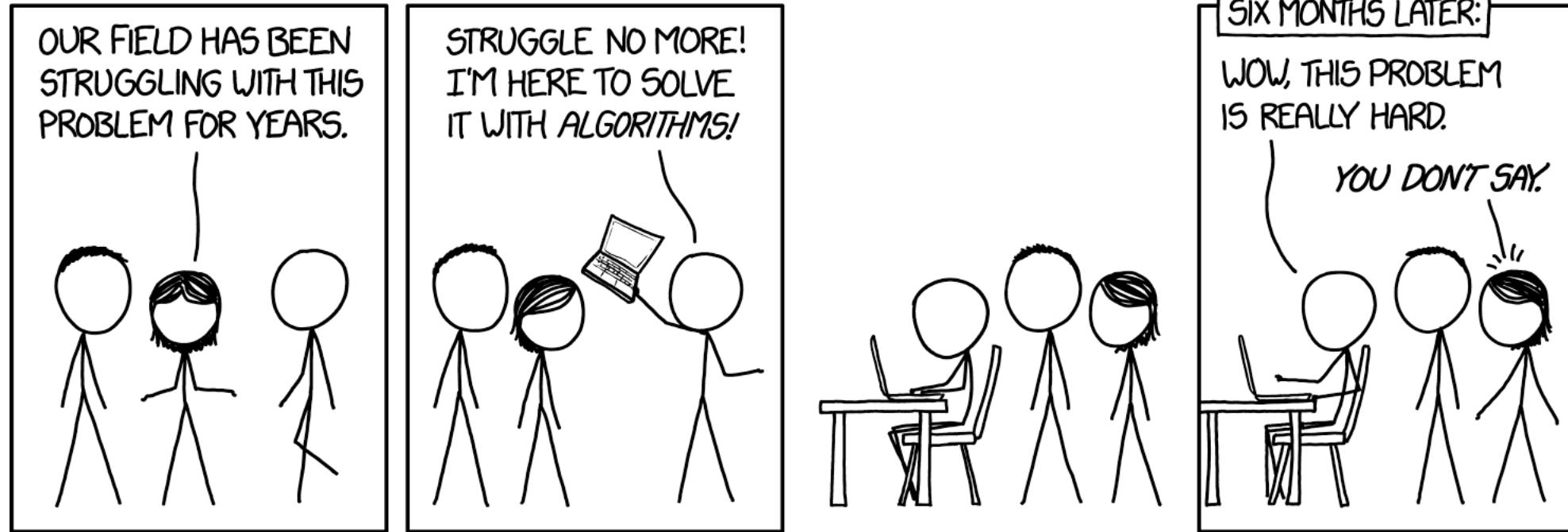


The top part shows the official BMW Twitter account (@bmw) with a banner image of a winding road through mountains. Below the banner, it displays 452 tweets, 48 followers, and 302,000 following users. The bottom part shows a timeline of tweets from January 2014, including posts about the #BMW X1, M4/M3 Walkaround Videos, and the #BMW i3.

Tweets **452** **FOLGE ICH** **48** **FOLLOWER** **302 TSD**

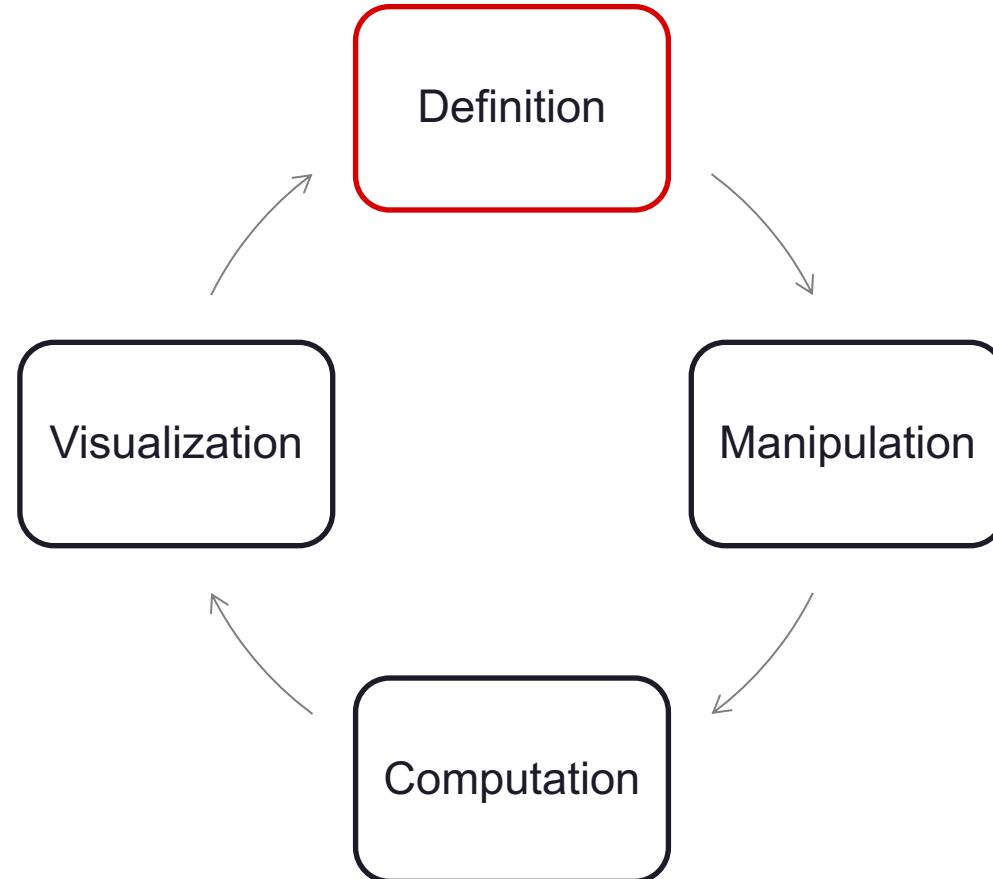
Tweet Content	Date	Retweets	Favorites
BMW @bmw · 24. Jan. A new year, a new style. The #BMW #X1 is already dressed for spring: bit.ly/1mD34dR	24. Jan.	0	0
BMW @bmw · 23. Jan. Busy for awhile. RT @bimmerpost #bmw M4 / M3 Detailed Walkaround Videos from Detroit Auto Show bit.ly/KAK5Vr	23. Jan.	21	18
BMW @bmw · 23. Jan. It's got the look. yc	23. Jan.	0	0
BMW @bmw · 23. Jan. Busy for awhile. RT @bimmerpost #bmw M4 / M3 Detailed Walkaround Videos from Detroit Auto Show bit.ly/KAK5Vr	23. Jan.	0	0

All problems solved...?

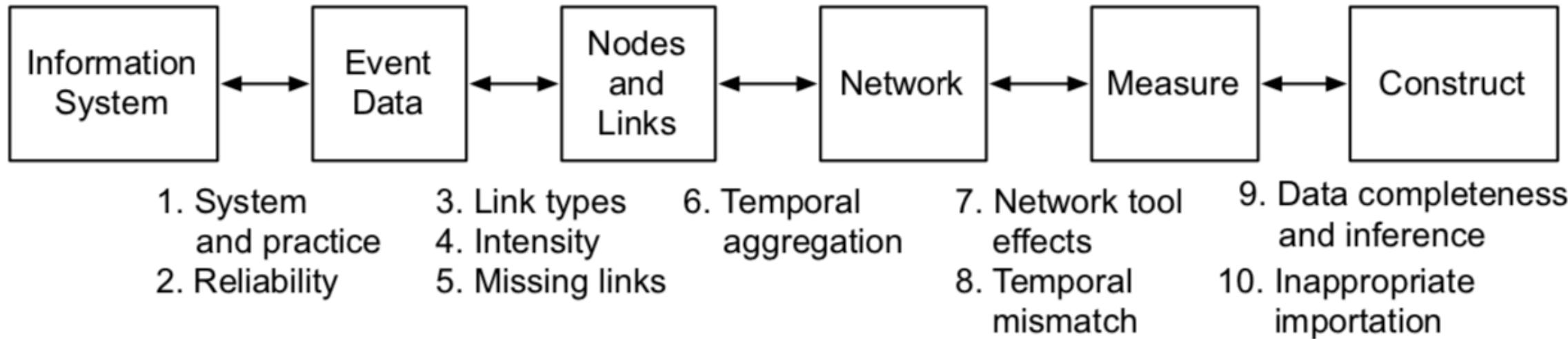


<https://xkcd.com/1831/>

DTD require is to make a lot of assumptions during the modeling process, which determine the outcome of all subsequent analyses



It is important to make those assumptions consciously to avoid validity issues



Other challenges

- Old theories in a novel context
- Limited access to proprietary data
- Interdisciplinarity
- Complexity
- Validity
- Big Data
- Privacy



„Quarks and cells neither mind when we discover their secrets nor protest if we alter their environments during the discovery process.“
– Adamic et al. (2009)

„Social phenomena involve the interactions of large (but still finite) numbers of heterogeneous entities, the behaviours of which unfold over time and manifest themselves on multiple scales [...] – even string theorists don't have it that bad!“
– Watts (2009)

Some key takeaways from successful studies

- Combine qualitative and quantitative methods
- Use multiple data sources
- Integrate traditional and digital methods of data collection
- Make use of location and context aware technology
- Analyze the *structure, content, and context* of social interaction

