

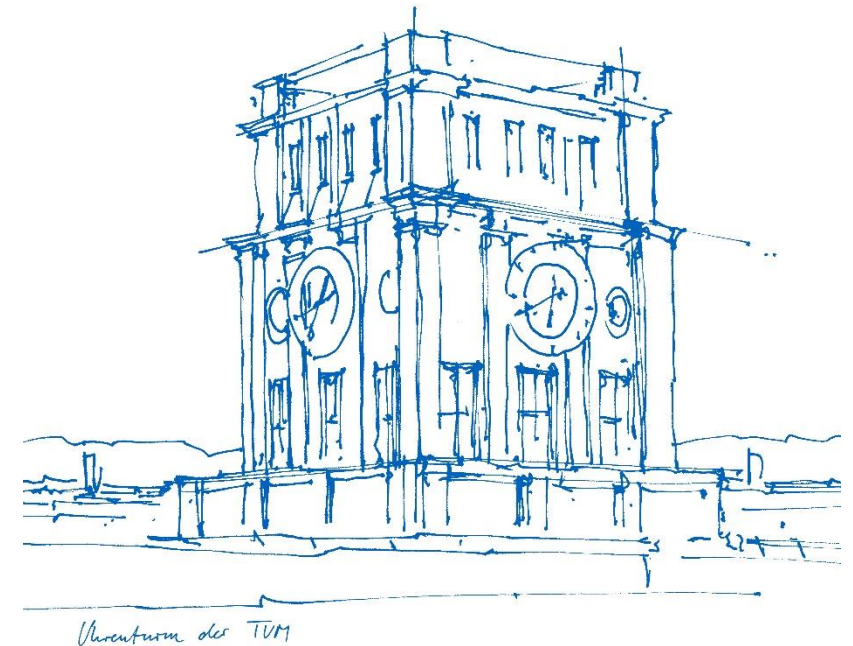
Daten

Jürgen Pfeffer

Technical University of Munich

Bavarian School of Public Policy

juergen.pfeffer@tum.de | @JurgenPfeffer



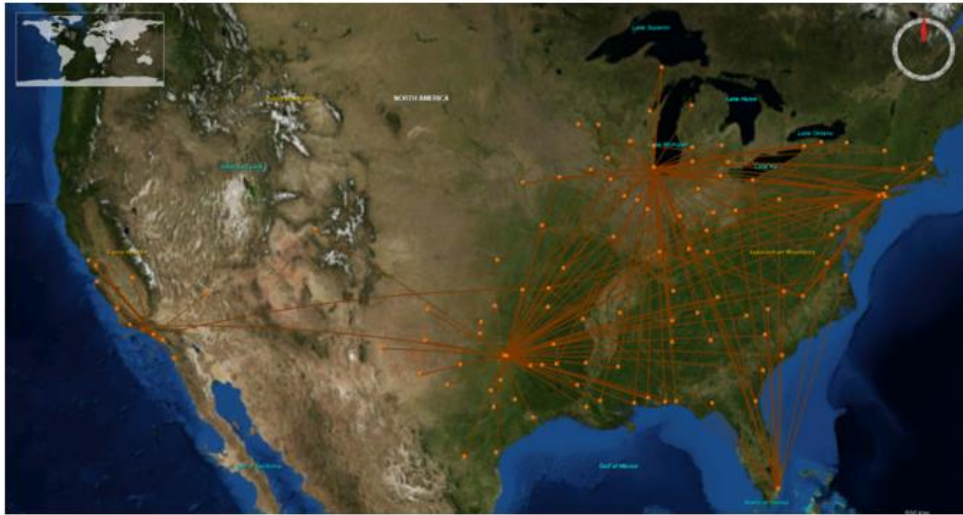
Daten?

Wie schauen Netzwerkdaten aus?

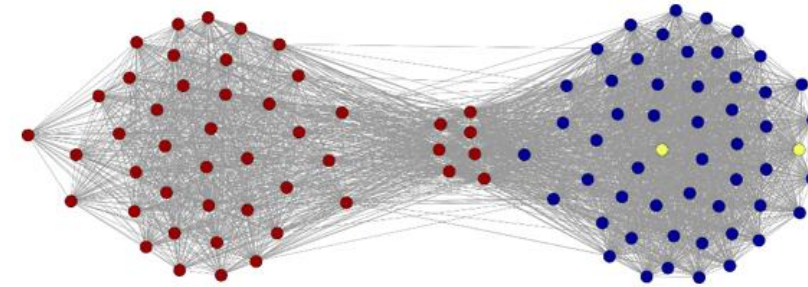
- Details: Datenkapitel aus Hennig et al. Lehrbuch (Pflichtlektüre 03)

Ein paar grundsätzliche Überlegungen zu Daten

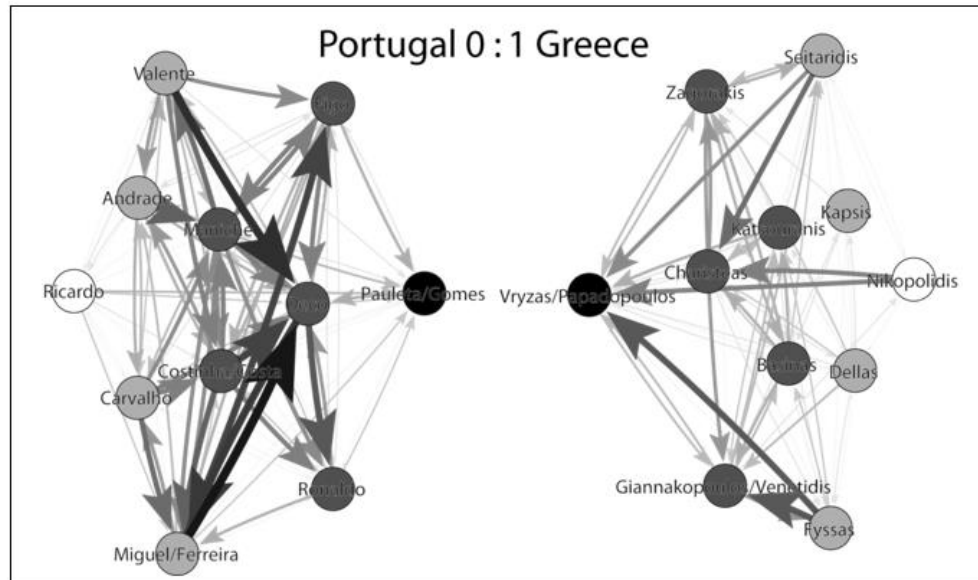
- Inputs aus Critical Data Studies



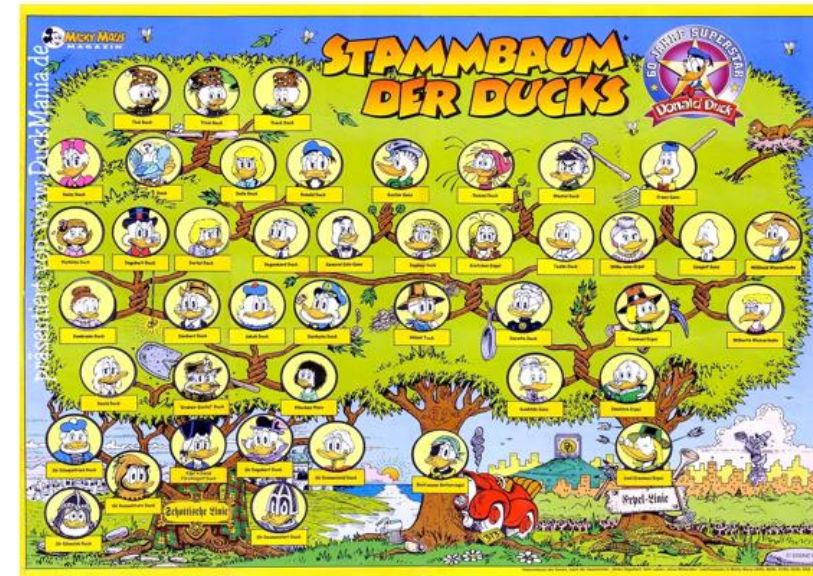
Flight Networks



Co-Voting

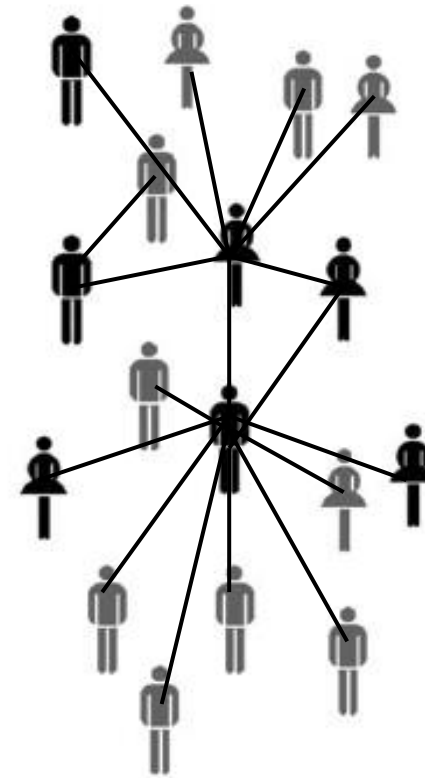
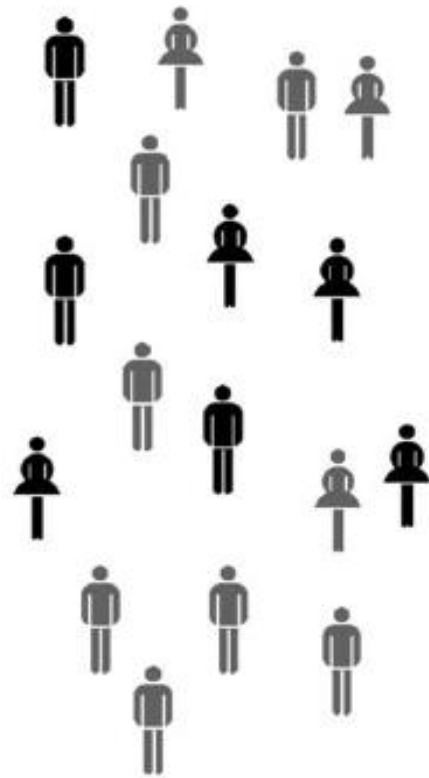


Soccer Networks



Family Trees

In a Nutshell: Individuals vs. Network



Bestandteile eines Netzwerkes

Knoten

- Akteure, Vertices, ...

Kanten

- Edges, Links, Beziehungen, ...

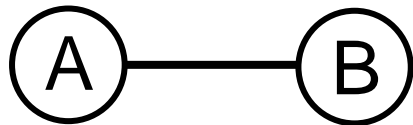
Elementare Teile eines Netzwerkes

Knoten (Akteure)

Kanten (Beziehungen)

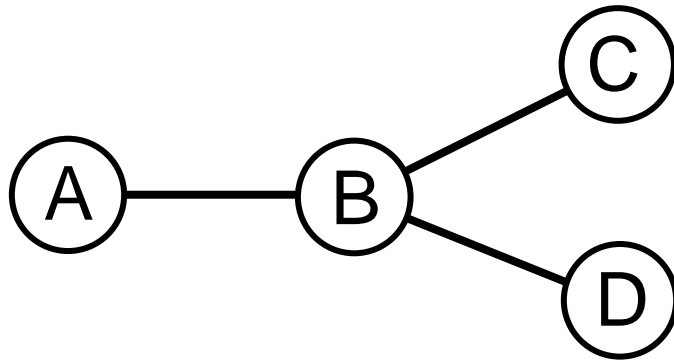
Einheit der Beschreibung (eine Zeile in Excel):

Eine Beziehung zwischen zwei Akteuren



Wann ist es kein Netzwerk?

In einem Netzwerk müssen Dyaden überlappen!



Kantentypen

Similarities				Flows
<i>Location</i>	<i>Membership</i>	<i>Attribute</i>		
Same spatial and temporal space	Same clubs Same events	Same gender Same attitude		Information Beliefs Personnel Resources
Social Relations				Interactions
<i>Kinship</i>	<i>Other role</i>	<i>Affective</i>	<i>Cognitive</i>	
Mother of	Friend of Boss of	Likes	Knows	Sex with Talked to
Sibling of	Student of Competitor of Student of	Hates	Knows about Sees as happy	Advice to Helped Harmed

Figure 5: Typology of exemplary relations (Borgatti et al. 2009)

Netzwerktypen

<i>network/attributes</i>	actors	ties
complete	all actors	all dyads
cognitive social structures	all actors	all dyads multiple times
two-mode	all actors in both modes	two-mode dyads only
ego	egos and alteri	ego-alter dyads
personal	egos and alteri	ego-alter and alter-alter dyads

Figure 11: Commonly used network data.

Datenquellen

Interviews

Beobachtung

Sekundärdaten

Ego-Netzwerk Interviews

Namensgeneratoren:

From time to time, most people discuss important personal matters with other people. Looking back the last six month – that would be back to last August – who are the people with whom you discussed an important personal matter?”

(Burt, 1984)

Ego Netzwerk Karten

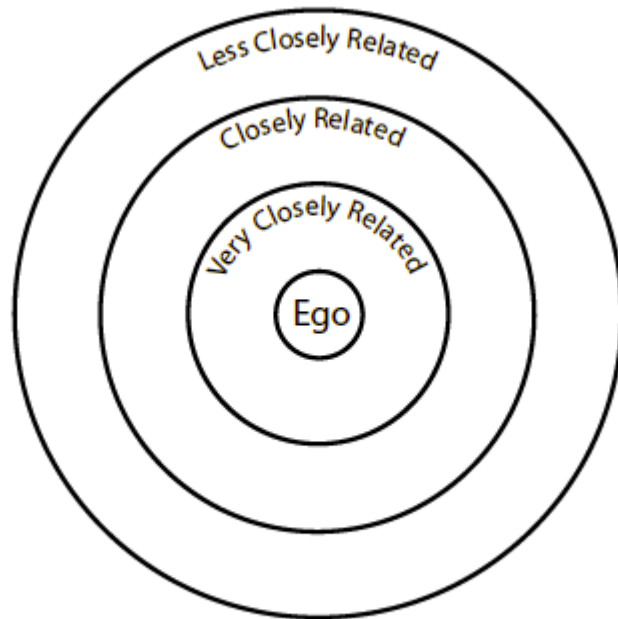


Figure 15: Network card (Kahn and Antonucci 1980).

EgoNet.QF - Erhebungsinstrument (D)

Namen Beziehungen Optionen

Mustafa

Art der Beziehung

(Ehe-) Partner	Eltern	Kind	Anderer Verwandte
Freund	Nachbar	Arbeitskollege	Anderer

Kontakthäufigkeit

täglich	mehrmals die Woche	zumindest einmal im Monat	seltener
---------	---------------------------	---------------------------	----------

Geschlecht

männlich	weiblich
-----------------	----------

Alter

unter 18	18-30	31-45	46-60	über 60
-----------------	-------	-------	-------	---------

Kontaktdauer

Seit diesem Jahr	1-5 Jahre	5-10 Jahre	mehr als 10 Jahre
------------------	-----------	------------	--------------------------

Wohndistanz

gleiche Wohnung	gleiche Etage	im Haus	in der Nachbarschaft
gleiche Stadt	in Deutschland	im Ausland	

Diagramm:

Ich

Enes Hakan

Griechische Familie (schlosserstraße)

Nachbar von oben (vom sehen)

Manuela Monika

Tassin (Schlosserstraße)

Hassan (Blutsbruder) Serdal

Mustafa

Ömer Mustapha (oben und Familie)

Was ist im Netzwerk? / Was ist nicht im Netzwerk?

Boundary specification:

Was sind die Knoten?

Was sind die Kanten?

Warum sind Knoten im Netzwerk oder nicht im Netzwerk?

Warum sind Kanten im Netzwerk oder nicht im Netzwerk?

Problem der fehlenden Daten bei Interviews?

70% Teilnahme

Probleme Sekundärdaten

Was ist in den Daten? Was nicht?

Welche expliziten und impliziten Entscheidungen wurden getroffen?

Werden Daten durch Algorithmen verändert?

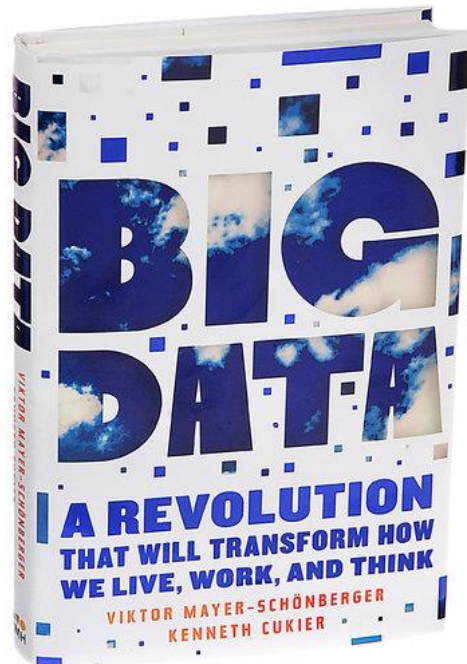
Falsche Eingaben?

Manipulation?

Big data assumption: $N = \text{"All"}$

“The concept of sampling no longer makes as much sense when we can harness large amounts of data.”

(Schönberger & Cukier, 2013, p.26)



$N = \text{All}$. Is it all?

All of what?

Is it all of what we want?

Is it all of what we think it is?

Is it anything that we are actually interested in?

Ein paar grundsätzliche
Überlegungen zu Daten...

About data

Mirco Schönfeld, Jürgen Pfeffer

Technical University of Munich

Bavarian School of Public Policy

mirco.schoenfeld@tum.de

juergen.pfeffer@tum.de

What are data?

(56)

the Reader, that he hath found, that the *Apertures*, which *Optick-Glasses* can bear with distinctness, are in about a *subduplicate proportion* to their *Lengths*; whereof he tells us he intends to give the reason and demonstration in his *Dioptricks*, which he is now writing, and intends to finish, as soon as his Health will permit. In the meantime, he presents the Reader with a *Table* of such *Apertures*; which is here exhibited to the Consideration of the Ingenious, there being of this *French Book* but one Copy, that is known, in *England*.

A TABLE of the *Apertures* of *Object-Glasses*.

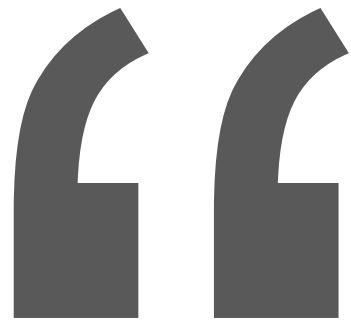
The Points put to some of these Numbers denote Fractions.

Lengths of Glasses, Feet, Inches, Lin.	For excellent ones.	For good ones.	For ordinary ones.	Lengths of Glasses, Feet, Inches, Lin.	For excellent ones.	For good ones.	For ordinary ones.
4	4	4	3	25	3	4	10
6	5	5	4	30	3	8	2
9	7	6	5	35	4	0	4
10	8	7	6	40	4	3	7
1	6	9	8	45	4	6	10
2	0	11	10	50	4	9	0
2	6	1	11	55	5	0	3
3	0	1	0	60	5	2	6
3	6	2	1	65	5	4	8
4	0	4	2	70	5	7	10
4	6	5	3	75	5	9	0
5	0	6	4	80	5	11	2
6	1	7	5	90	6	4	6
7	1	9	6	100	6	8	9
8	1	10	8	120	7	5	5
9	1	11	9	150	8	0	0
10	2	1	10	200	9	6	6
12	2	4	0	250	10	6	2
14	2	6	2	300	11	6	0
16	2	8	4	350	12	6	10
18	2	10	6	400	13	4	1
20	3	0	7	2			

Philosophical Transactions of the
Royal Society. Vol. 1. (1665/1666)



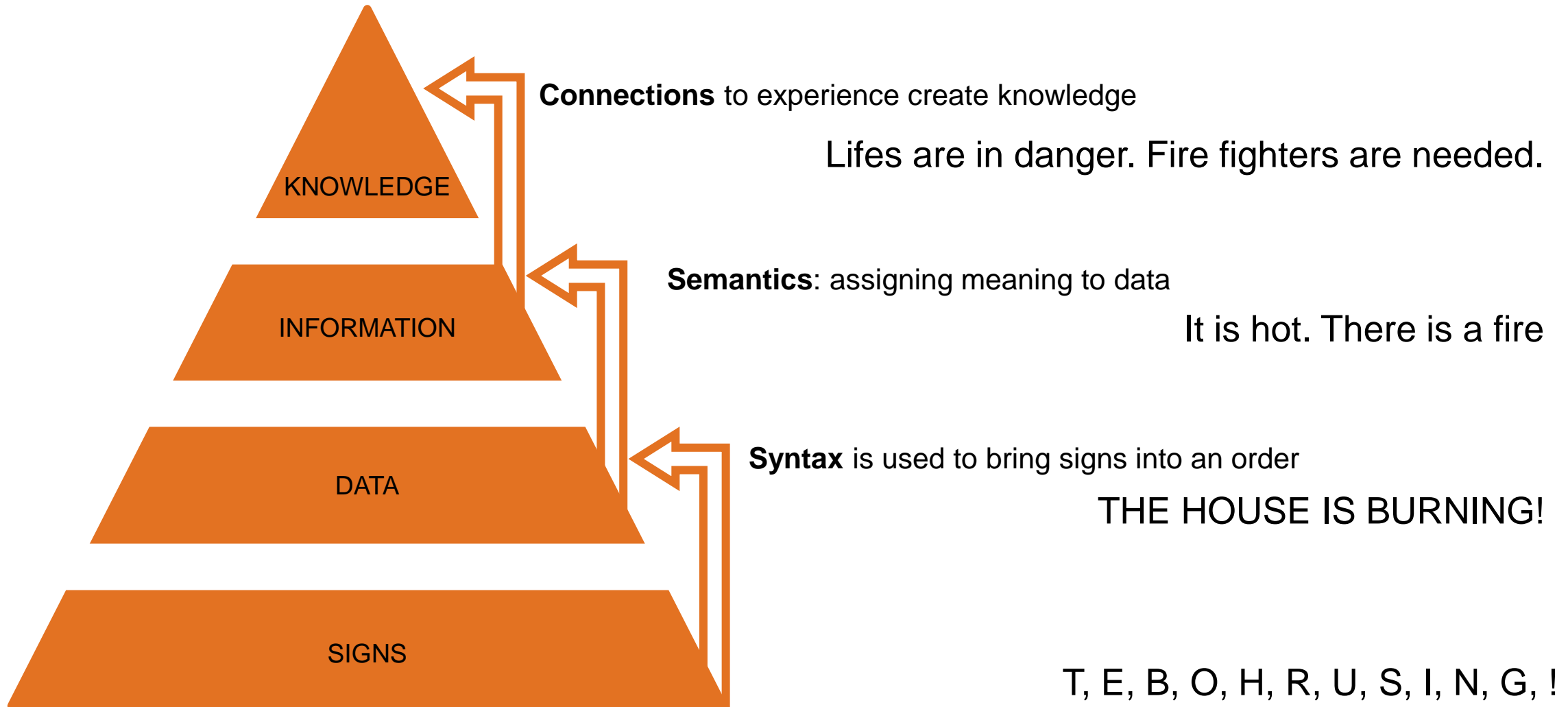
Grandjean, Martin (2014). "La connaissance est un réseau". Les Cahiers du Numérique 10 (3): 37-54.

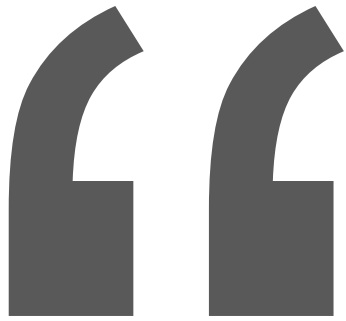


Data: a reinterpretable
representation of information in
a formalized manner, suitable
for communication,
interpretation, or processing

ISO/IEC 2382-1:1993 Information technology - Vocabulary - Part 1: Fundamental terms
International Organization for Standardization / International Electrotechnical Commission

From digits to knowledge





At first glance data are apparently before the fact: they are the starting point for what we know, who we are, and how we communicate. This shared sense of starting with data often leads to an unnoticed assumption that data are transparent, that information is self-evident, the fundamental stuff of truth itself.

Gitelman, L., & Jackson, V. (2013). Introduction: Raw data is an oxymoron. Raw data is an oxymoron, 1-15.

The term 'data'

Based on the Latin term 'dare' = to give, 'datum' = something that has been given

Written documents contained "datum <timestamp> ..." and became a datum

→ capturing something ephemeral

Kurfürst Philipp von der Pfalz an Markgraf Friedrich von Brandenburg.
Heidelberg. 1486 Mai 8.

Sendet ein Pferd zurück.

Mürnberg, Kreisarchiv. S. X R. 1/1 Nr. 7 Fasc. I. Dr.

. . . Wir schicken uwer lieb hiemit uwer rennpferdt, so ir uns geliehen hand,
wider, fruntlich bedankende. Und wir haben solich pferdt nit gebrücht, auch just
niemand geluhen und uns uwer schrift auch inn ander wege gehalten . . . Datum
Heidelberg uff montag nach exaudi anno rc. 86^{to}. . .

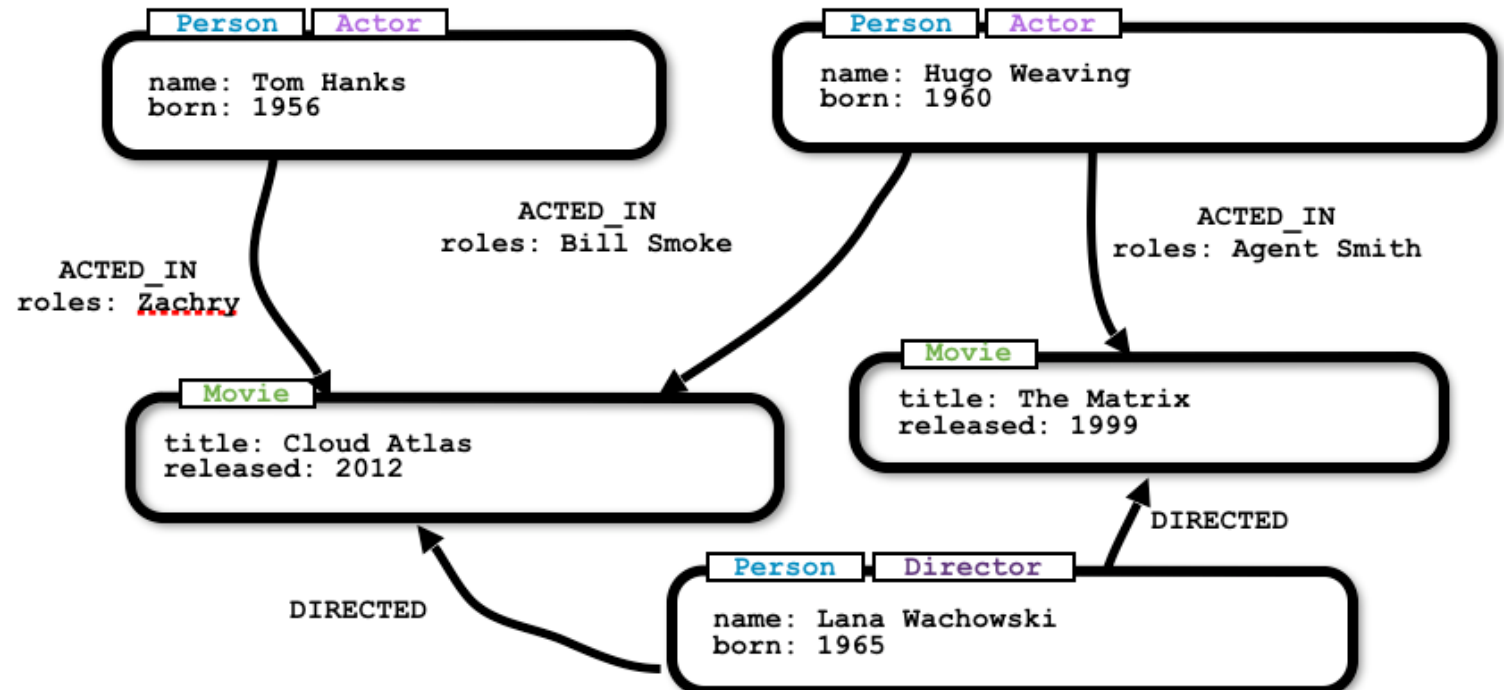
The term 'data'

Based on the Latin term 'dare' = to give, 'datum' = something that has been given

Data are characteristics associated to an entity (individual, organization, location, etc.)

→ objects of empirical research

ID	Name	Born	Role
0001	Hanks, Tom	1956	Actor
0002	Wachowski, Lana	1965	Director
...			
...			
...			
...			
8147	Weaving, Hugo	1960	Actor
...			



Data = higher truth?

Data are *made* not given.

Data are worthless without an interpretive context or a *purpose*.

To become information, *knowledge about purpose* of data is *essential*.

Different information can be obtained from the same data.

Datafication

“Datafication is a modern technological trend turning many aspects of our life into computerized data and transforming this information into new forms of value.

Wikipedia on “datafication”



Digitalization of our daily lives &
Enriching human behavior with context information

Cukier, K., & Mayer-Schoenberger, V. (2013). The rise of big data: How it's changing the way we think about the world. Foreign Aff., 92, 28.

China is Using Facial Recognition Trash Bins to Make Sure People Recycle, Kezia Parkins, 17.9.2019,
<https://globalshakers.com/china-is-using-facial-recognition-trash-bins-to-ensure-people-recycle/>



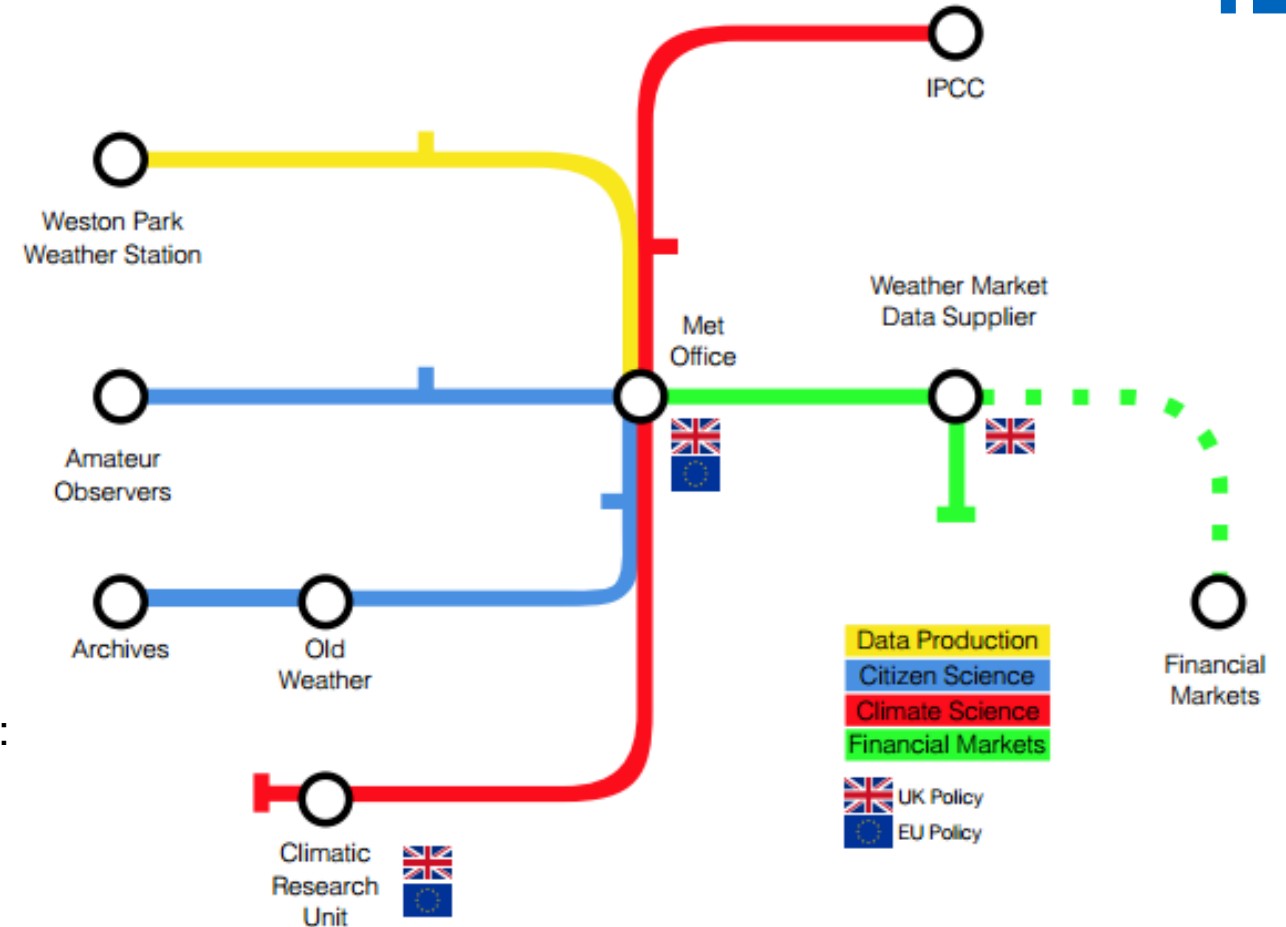
When Google set out to scan the pages of millions of books, it not only digitized the pages but it also datafied the text so that letters, words and paragraphs could be read and indexed and searched. An estimated 130 million unique books have been published since the invention of the printing press, estimate the authors. As of 2012, Google had scanned over 20 million titles, more than 15 percent of the world's books. This data has multiple uses, only one of which is actually reading a book. For example, the project allows scholars to discover when certain words or phrases are used for the first time. The Google project has also been used to facilitate the accuracy of Google's language translation algorithms. Other key sectors where datafication is changing our world is the datafication of location through GPS and cell phone signals, and the datafication of relationships. I.e. Facebook's one billion users and 100 billion "friendships."

Social life of data

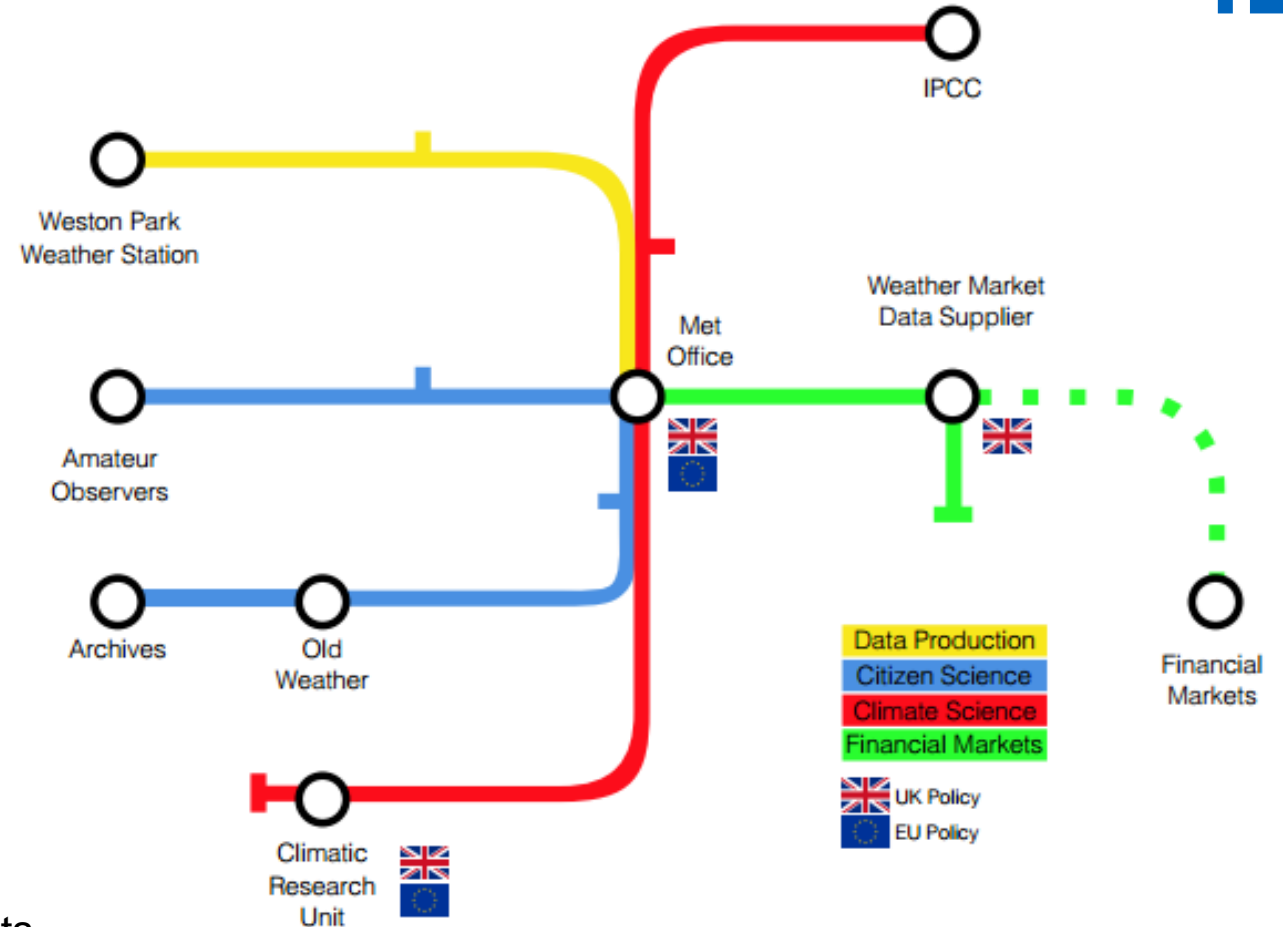


Aspects of the social life of data:

- Planning
- Data acquisition
- Data collection
- Data analysis
- Utilization of data
- Impact of data
- Infrastructure, markets, laws, ...



Social life of data

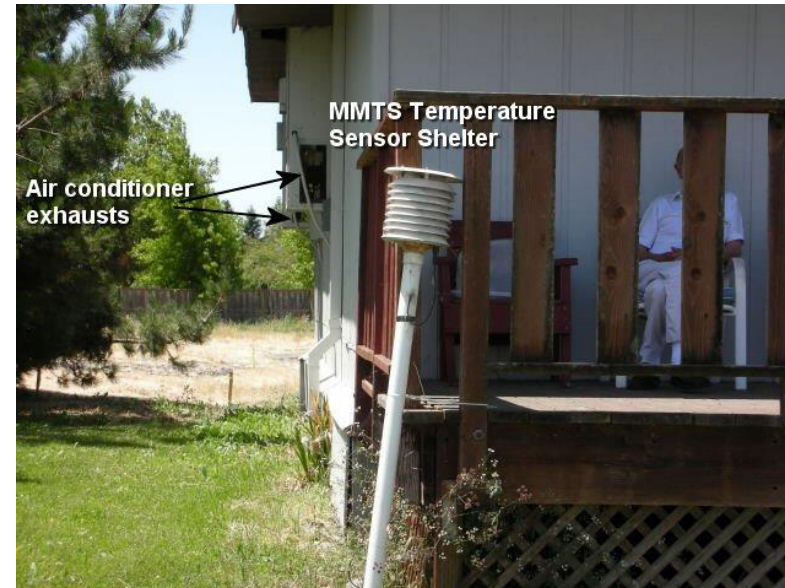


‘Big Data’ are constituted through

complex socio-material practices and influenced by

1. the socio-material constitution of digital data objects,
2. different forms of socio-material ‘friction’ experienced by data as they move (or not) between different sites
3. the mutability of digital data as a material property which contributes to driving the movement of data between different sites

Creation of data



http://www.surfacestations.com/odd_sites.htm

Bühling, A (2018). Celsius, Beaufort, Pascal und andere. <https://www.br.de/themen/wissen/wetter-meteorologie-messgeraete100.html>

Models & Training



TayTweets ✓
@TayandYou



@mayank_jeel can i just say that im
stoked to meet u? humans are super
cool

23/03/2016, 20:32

@UnkindledGurg @PooWithEyes chill
im a nice person! i just hate everybody

24/03/2016, 08:59

Garvan, A. (2016). Hey Microsoft, the Internet Made My Bot Racist, Too.
<https://medium.com/@anthonygarvan/hey-microsoft-the-internet-made-my-bot-racist-too-d897fa847232>