Semiotics: Recap

Jörg Cassens

Data and Process Visualization
SoSe 2017

Universität Hildesheim
medieninformatik
IMAI – Institut für Mathematik und Angewandte Informatik

Inhaltsverzeichnis

1 Recap **1**
1.1 History . . . 1
1.2 Communicate . . . 3
1.3 Semiotics . . . 4
1.4 Classification Framework . . . 6

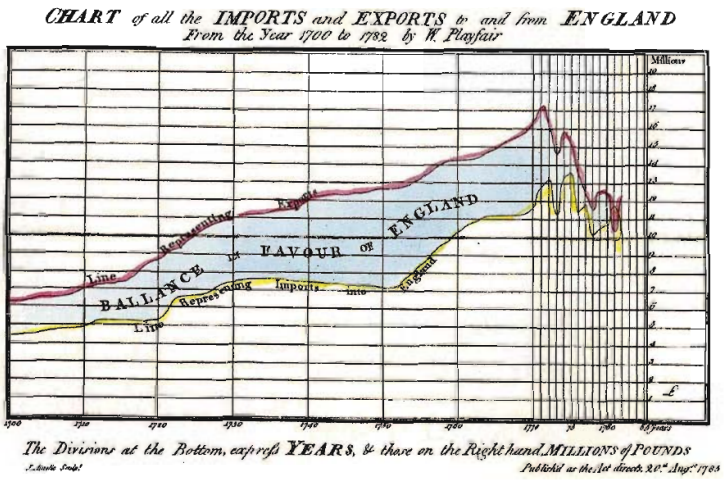
2 Tutorial **9**

3 Examples **10**

1 Recap

1.1 History

1786



Playfair – trade deficits (Source: Tufte (2001))

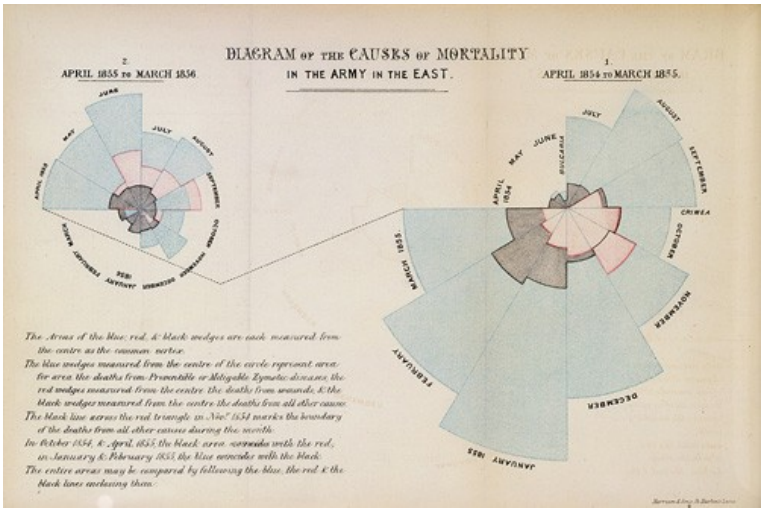
1845

John Snow – Cholera Epidemic of London (Source: Tufte (2001))

1845

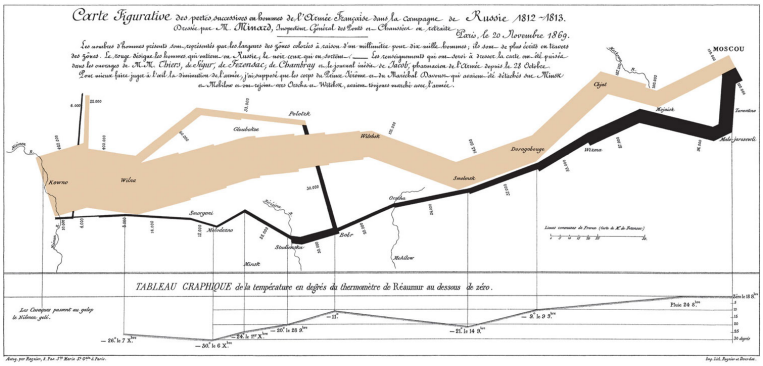
John Snow – Cholera Epidemic of London (Detail) (Spence, 2014)

1855



Florence Nightingale – Cause of death over time (Source: Jänicke (2016))

1869



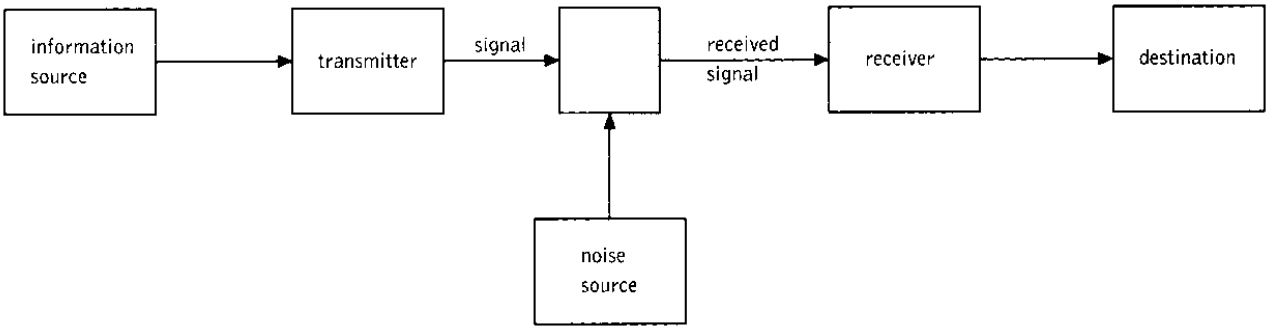
Charles Minard – Napoleon's Russian Campaign

1.2 Communicate

Communication: Kirk

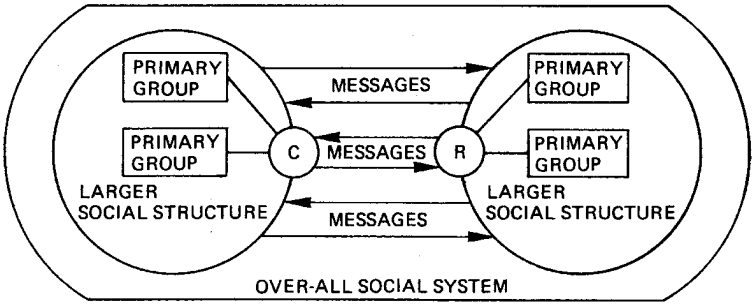
Source: Kirk (2012)

Communication: Shannon & Weaver



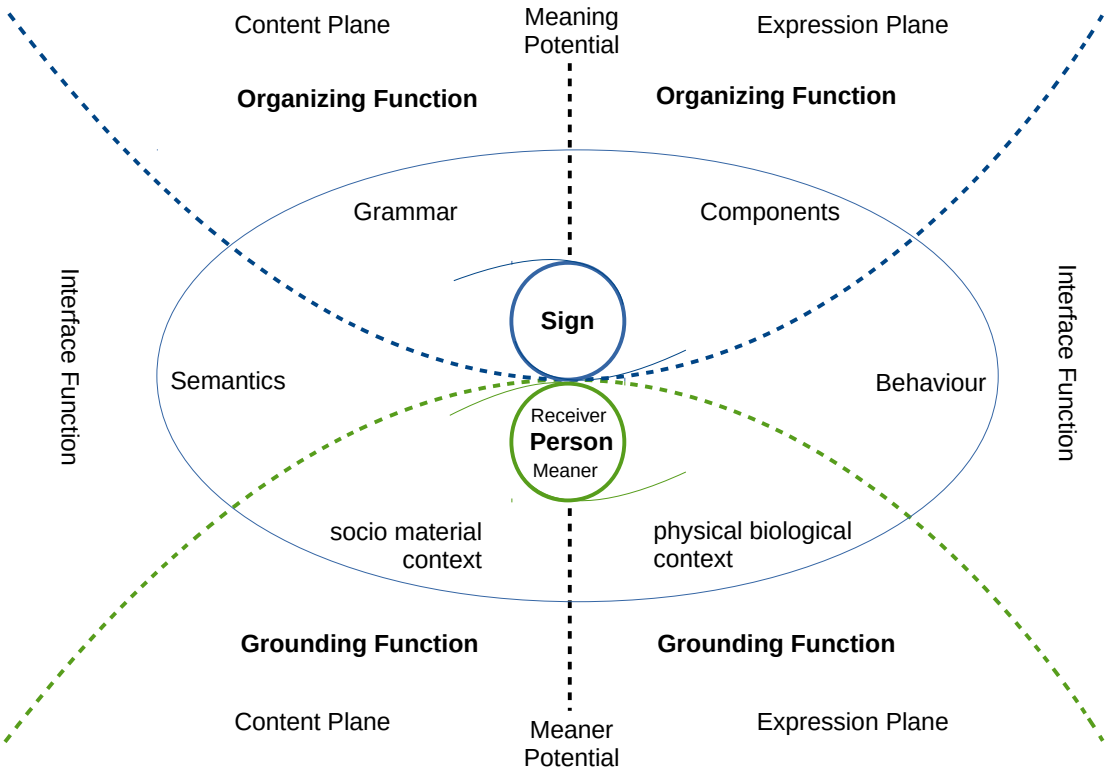
Source: Shannon & Weaver, here: Kress and van Leeuwen (2006)

Communication: Riley & Riley



Source: Riley & Riley, here: Kress and van Leeuwen (2006)

Communication: Wegener

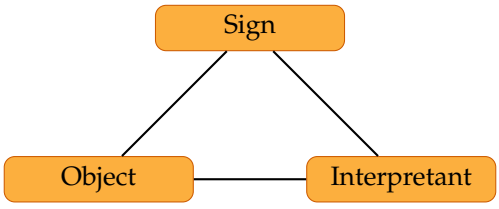


Source: Wegener (2011, 2015)

1.3 Semiotics

Peirce

- Introduces different types of signs in society
- In the end, he had something like 66, but we look at 3:
  - Iconic – looks like what it is meant to mean
  - Indexical – contextual connection (smoke and fire)
  - Symbolic – arbitrary like language
- The semiotic triangle is his theory behind this model



Sign

- The key notion in any semiotics is the 'sign'
- *Different starting point: not descriptive like Peirce, but functional and social*
- Drawing on next slide was made by a 3-year-old boy
  - Sitting on his father's lap, he talked about the drawing as he was doing it
  - "Do you want to watch me? I'll make a car . . . got two wheels . . . and two wheels at the back . . . and two wheels here . . . that's a funny wheel . . . "

  - When he had finished, he said, "This is a car."
- This was the first time he had named a drawing, and at first the name was puzzling
- How was this a car?
- He had provided the key himself: 'Here's a wheel.'

(Kress and van Leeuwen, 2006)

A Car

Source: Kress and van Leeuwen (2006)

Car-ness

- A car, for him, was defined by the criterial characteristic 'has wheels', and his representation focused on this aspect
- What he represented was, in fact, 'wheelness'
- Wheels are a plausible criterion to choose for 3-year-olds, and the wheel's action, on toy cars as on real cars, is a readily noticed and describable feature

- This boy's interest in cars was, for him, most plausibly condensed into and expressed as an interest in wheels
  - Choosing what to represent ("the signified")
- Wheels, in turn, are most plausibly represented by circles
  - Choosing how to represent ("the signifier")

(Kress and van Leeuwen, 2006)

Representation

> Shortened version: "We see representation as a *process* in which the makers of signs (...) seek to make a *representation of some object or entity,* whether physical or semiotic, and in which *their interest* in the object (...) is (...) arising out of the cultural, social and psychological *history of the sign-maker,* and focused by the *specific context* in which the signmaker produces the sign. That *'interest'* is the source of the *selection* of what is seen as the *criterial aspect of the object,* and this *criterial aspect* is then regarded as *adequately representative* of the object in a given context. In other words, it is never the 'whole object' but only ever its *criterial aspects* which are *represented*." Kress and van Leeuwen (2006)

Sign-Making

- The criterial aspects are represented in what seems to the sign-maker the most apt and plausible fashion, and the most apt and plausible representational mode
- Sign-makers thus 'have' a meaning, the signified, which they wish to express, and then express it through the semiotic mode(s) that make(s) available the subjectively felt, most plausible, most apt form, as the signifier
- This means that in social semiotics the sign is not the pre-existing conjunction of a signifier and a signified, a ready-made sign to be recognized, chosen and used as it is
- We see signs as motivated – not as arbitrary – conjunctions of signifiers (forms) and signifieds (meanings)
- Signs are never arbitrary, and 'motivation' should be formulated in relation to the sign-maker and the context in which the sign is produced,

(Kress and van Leeuwen, 2006)

1.4 Classification Framework

Classification

- Starting with the types of questions users have, the framework supports the selection of data mining and visualization work flows as well as deployment options that answer these user questions.
- We look at the following aspects
  - Level of analysis
  - Types of analysis
  - Intended audience (and/or producer)
  - Medium used
- Some projects aim to answer more than one question

Level of Analysis

○ Micro level, or the individual level
  - Small data sets, typically between 1 and 100 records
  - e.g. a person and his friends

◐ Meso or the group level
  - About 101 to 10,000 records
  - e.g. researchers at a single university

● Macro, global or population level
  - Typically exceeding 10,000 records
  - e.g. pertaining an entire country

Types of Analysis

🛈 Statistical Analysis/Profiling
  - What are the entities that are being described (e.g. persons, grants, publications)?

🕒 Temporal Analysis: *When*
  - Does the visualization show a development over time?

🌍 Geospatial Analysis: *Where*
  - Does the visualization include information about location?

≡ Topical Analysis: *What*
  - What is the topical area of the visualization?

▽ Network Analysis: *With Whom*
  - Does the visualization contain information about social networks?

Audience

- ⚤ Gender – are we targeting a certain gender?
- ⑤ Age – is it intended for certain age groups?
- ✍ Education – is the level of education important
- ♿ Disability – are disabilities taken into account (for example colour blindness)?
- ☐ Contextual parameters, e.g.
  - ☕ Leisure – related to our leisure
  - 🏭 Business – related to business
  - ☣ Scientific – related to science
  - ✝ Religious – related to religion
  - ☐ Any other information defining the audience

Medium

- ✎ Printed medium
- ⌨ Digital medium
- ✇ Time-based – visualizing information using time
- ⊙ Location-based – spatially visualizing information
- Ⓐ Modality Text – contains text
- 🔔 Modality Sound – contains sound
- ☞ Interactive visualization
- ☐ Other – other information about the medium

Framework

Level

- ○ Micro level
- ◐ Meso level
- ● Macro level

Type

- 🛈 Profiling
- 🕒 Temporal
- 🌍 Geospatial
- ≡ Topical
- ▽ Network

Audience

- ⚤ Gender
- ⑤ Age
- ✍ Education
- ♿ Disability
- ☐ Context, e.g.
  - ☕ Leisure
  - 🏭 Business
  - ☣ Scientific
  - ✝ Religious
  - ☐ Other

Medium

- ✎ Printed
- ⌨ Digital
- ✇ Time-based
- ⊙ Spatial
- Ⓐ With Text
- 🔔 With Sound
- ☞ Interactive
- ☐ Other

Metafunctions

- **Metafunctions:** The function of the communication
- Systemic clusters; groups of semantic systems that make meanings of a related kind
  - Ideational – representing 'the world around and inside us'
    - Logical – logical–semantic relationships
    - Experiential – representation of reality, experiences the meaner has
  - Interpersonal – enacting social interactions as social relations
  - Textual – a coherent 'world of the text', organisation of 'text'

Field

> "The FIELD OF DISCOURSE refers to what is happening, to the nature of the social action that is taking place: what is it that the participants are engaged in, in which the language figures as some essential component?" (Halliday and Hasan, 1985)

- Ideational – representing 'the world around and inside us'
- On the contextual stratum, realised as "Field of Discourse"
  - What is the domain? What are the long term or short term goals?
  - What is the structure, what are the networks of interaction?
- *Level* and *type of analysis* pertain to the field

Tenor

> "The TENOR OF DISCOURSE refers to who is taking part, to the nature of the participants, their status and roles: What kinds of role relationship obtain among the participants [...], both the types of speech role that they are taking on in the dialogue and the whole cluster of socially significant relationships in which they are involved?" (Halliday and Hasan, 1985)

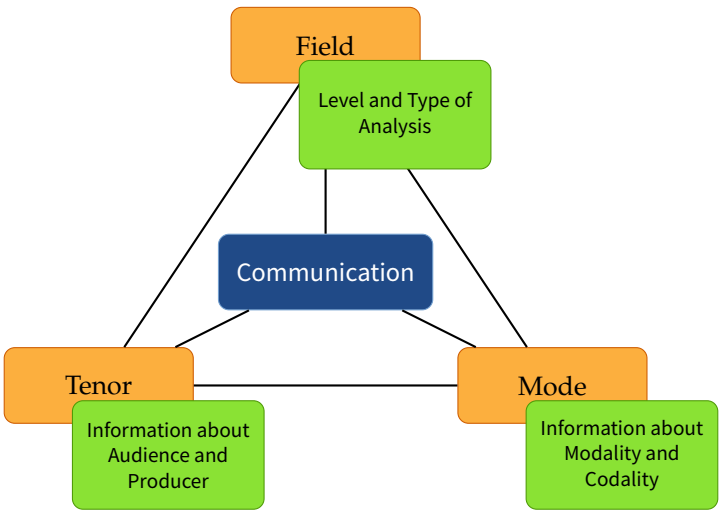
- Interpersonal – enacting social interactions as social relations
- On the contextual stratum, realised as "Tenor of Discourse"
  - What is the power structure between actors involved?
  - What is the agentive role?
  - What is the competence of the actors?
- The *audience (and producer)* pertains to the tenor

Mode

> "The MODE OF DISCOURSE refers to what part the language is playing, what is it that the participants are expecting to do for them in that situation: the symbolic organisation of the text, the status that it has, and its function in the context ... and also the rhetorical mode, what is being achieved by the text in terms of such categories as persuasive, expository, didactic, and the like." (Halliday and Hasan, 1985)

- Textual – a coherent 'world of the text', organisation of 'text'
- On the contextual stratum, realised as "Mode of Discourse"
  - What medium is used?
  - What is the type of interaction (dialogic, monologic)?
  - What is the rhetorical thrust?
- The *medium* used pertains to the mode

Categories & Metafunctions



2 Tutorial

Recap 2.2: Collecting Visualizations

- For the next two weeks, you should collect interesting Visualizations you come across
- You should use the framework introduced to describe the different visualizations
- You should be able to present one or two examples of visualizations
  - Classification according to the framework
  - Shortfalls of the framework
- Deliverable:
  - Monday, 24.4., 18:00, learnweb
  - Monday, 24.4., in the course

Recap 2.3: Preparing Visualizations

- In the course of a normal day, make notes of examples in which data is represented visually, aurally or by tactile means
- Afterwards, identify whether, for each example, the data has value (numeric, ordinal or categorical) or is a relation
- Sketch a possible visualization for this data
  - Classification according to the framework
  - Shortfalls of the framework
- Deliverable:
  - Monday, 24.4., 18:00, learnweb
  - Monday, 24.4., in the course

Assignment 3.1: Examples Revisited

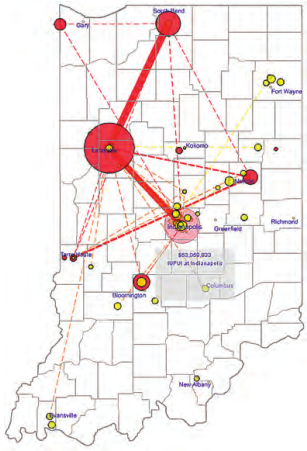
- Without consulting the slides, sketch what you can remember of
  - Minard's record,
  - Nightingale's diagram and
  - Snow's Soho map.
- In other words, externalize your mental models of those representations.
- By means of sketches explore alternative ways of representing the data encoded in the representations of Minard, Nightingale, Snow and Beck.

Assignment 3.2: Small Visualization Task

- A small data set is being handed out
- You should classify the visualization according to the framework introduced
- Discuss in the group how to visualize the data set
- Prepare a visualization and present it in class

3 Examples

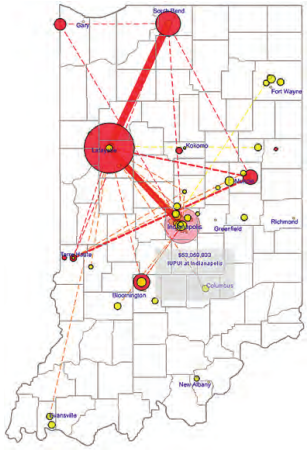
Academic-Industry Collaboration I



Example by Börner and Polley (2014)

- Data set of funded projects & proposals
- With industry and academic partners
- Geo-coding industry and academic institutions and overlaying their positions and collaboration network on a map of Indiana
- Nodes size-coded by the total dollar amount of all awards

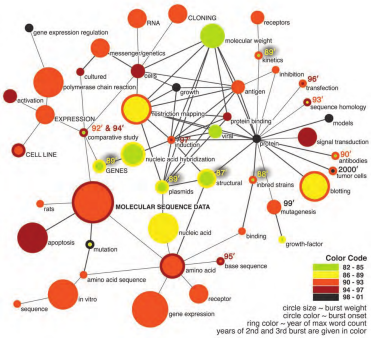
Academic-Industry Collaboration II



Example by Börner and Polley (2014)

- Level and Type
  - 2001-2006
  - Indiana, US
  - Acad.-Indus. Collab
- Audience
  - Media & Politicians
  - Businesspeople
  - Scientists
- Media
  - Digital representation
  - Spatially encoded
  - Uses text
  - Search Interface

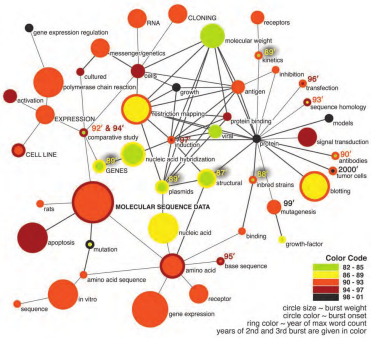
Activity Bursts in Publications I



Example by Börner and Polley (2014)

- Data set of publications in one journal over 20 years
- Detecting bursts (sudden increase in keyword use) in top 10% articles
- Node size relates to “suddenness”
- Node colour represents year
- Lines are co-occurances

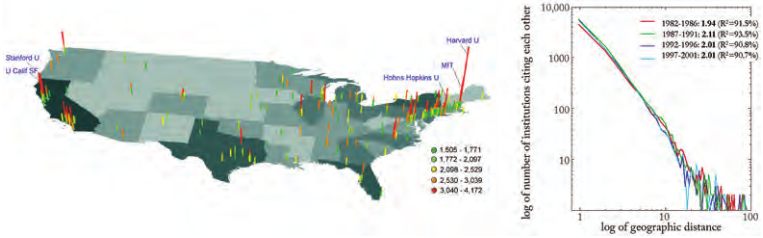
Activity Bursts in Publications II



Example by Börner and Polley (2014)

- Level and Type
  - 1982-2001
  - Biomedical
  - Word co-occ.
- Audience
  - Media & Politicians
  - Scientists
- Media
  - Digital representation
  - Spatially encoded
  - Uses text

Physical Locations Matter I

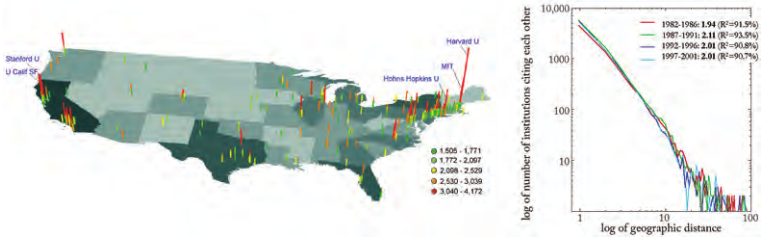


Example by Börner and Polley (2014)

- Same data set
- Data set of publications in one journal over 20 years
- Does location still matter in the internet age?

- Left: Location and distance
- Right: log of distance (x) vs log of citations (y)
- Researchers cite more locally

Physical Locations Matter II

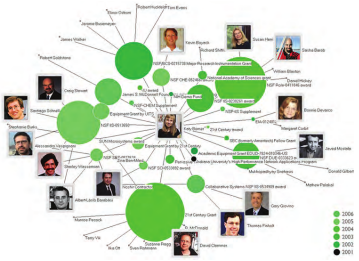


Example by Börner and Polley (2014)

- Level and Type
  - 1982-2001
  - USA
  - Citation network & locations

- Audience
  - Media & Politicians
  - Scientists
- Media
  - Digital representation
  - Spatially encoded
  - Uses text

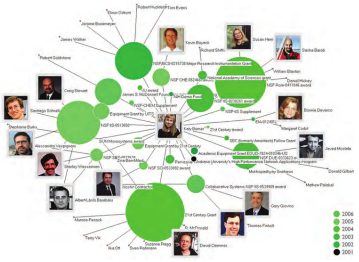
Project Collaborations I



Example by Börner and Polley (2014)

- Project collaboration for one scholar
- Data on all her projects funded by NSF
- Nodes: projects (green) and researchers (white or photo)
- Node size: grant size
- Lines: co-investigator

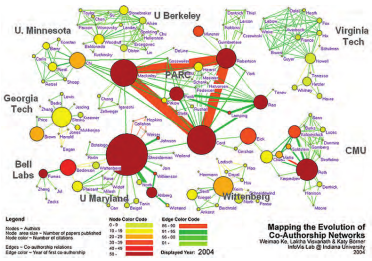
Project Collaborations II



Example by Börner and Polley (2014)

- Level and Type
  - 🕒– ○ 2001-2006
  - ▽– ○ Project – co-investigator
- Audience
  - ✍ Media & Politicians
  - ☣ Scientists
  - 🦽 Colour-blind
- Media
  - 🖴 Digital representation
  - ⊙ Spatially encoded
  - Ⓐ Uses text

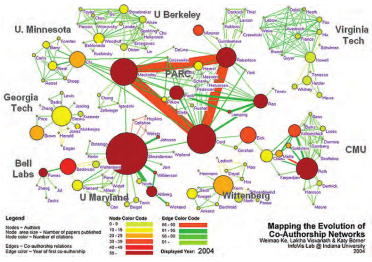
Co-Authorship I



Example by Börner and Polley (2014)

- Co-authorship in one journal
- Authors labelled by name
- Node size # publications
- Node colour # citations
- Lines size # collaborations
- Lines colour year of 1st collaboration

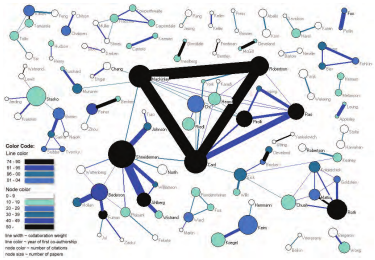
Co-Authorship II



Example by Börner and Polley (2014)

- Level and Type
  - 🕒– ◐ 1986-2004
  - ▽– ◐ co-author
- Audience
  - ✍ Publishers
  - ☣ Scientists
- Media
  - 🖴 Digital representation
  - ⊙ Spatially encoded
  - Ⓐ Uses text

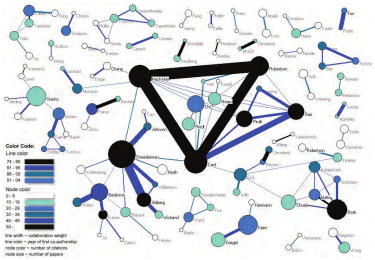
Individual Experts or Teams I



Example by Börner and Polley (2014)

- Same data set
- Node size # papers
- Node colour # citations
- i.e. how often cited
- Lines width # co-author
- Lines colour first year

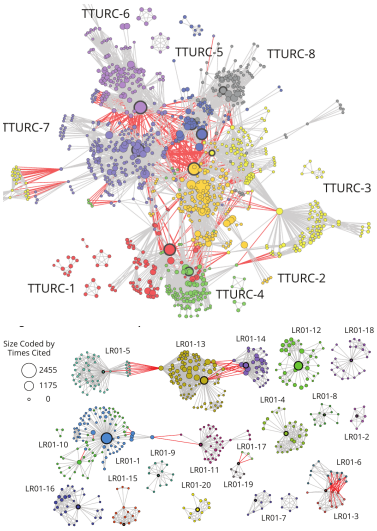
Individual Experts or Teams II



Example by Börner and Polley (2014)

- Level and Type
  - 1986-2004
  - co-author
- Audience
  - Politicians
  - Scientists
  - Colour-blind
- Media
  - Digital representation
  - Spatially encoded
  - Uses text

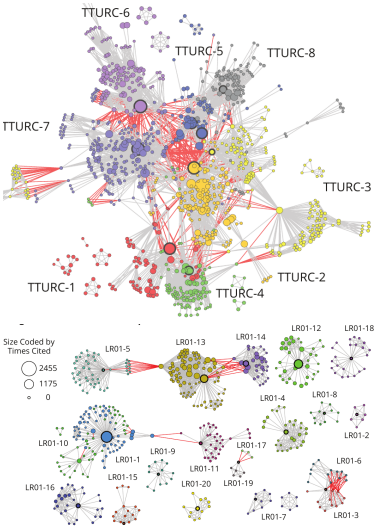
Funding I



Example by Börner and Polley (2014)

- Impact of different kinds of funding by an agency
- One graph for each of two funding types
- Node size denotes citations
- Node colour denotes funding number
- Links denote co-authorship

Funding II



Example by Börner and Polley (2014)

- Level and Type
  - ▽– ◐ co-author
- Audience
  - ✍ Politicians
  - ☣ Scientists
- Media
  - ⌨ Digital representation
  - ⊙ Spatially encoded
  - Ⓐ Uses text

Chinese Collaboration I

Example by Börner and Polley (2014)

- Global collaboration network of Scientists at Chinese Academy of Sciences
- Aggregated on country level
- Countries colour coded on log of # collaborations
- Width of flow lines also collaboration

Chinese Collaboration II

Example by Börner and Polley (2014)

- Level and Type
  - 🌍– ● World
  - ▽– ● Collaboration network & locations
- Audience
  - ✍ Media & Politicians
  - Colour blind
  - ☣ Scientists
- Media
  - ✎ Printed representation
  - ⊙ Spatially encoded
  - Ⓐ Uses text

References

Literatur

Börner, K. and Polley, D. E. (2014). *Visual Insights – A Practical Guide to Making Sense of Data*. MIT Press, Cambridge, Massachussetts.

Halliday, M. A. and Hasan, R. (1985). *Language, Context, and Text: aspects of language in a scoial-semiotic perspective*. Deakin University Pres, Geelong, Australia.

Jänicke, H. (2016). Vorlesung visualisierung. online.

Kirk, A. (2012). *Data Visualization – A Successful Design Process*. PACKT Publishing, Birmingham.

Kress, G. and van Leeuwen, T. (2006). *Reading Images – The Grammar of Visual Design, 2nd Edition*. Routledge, London.

Malaka, R., Butz, A., and Hussmann, H. (2009). *Medieninformatik – Eine Einführung*. Pearson Studium, Munich.

Spence, R. (2014). *Information Visualization – An Introduction, 3rd Edition*. Springer, Heidelberg.

Tufte, E. R. (2001). *The Visual Display of Quantitative Information, 2nd Edition*. Graphics Press, Cheshire, Connecticut.

Ware, C. (2004). *Information Visualization – Perception for Design, 2nd Edition*. Morgan Kaufman/Elsevier, San Francisco, USA.

Wegener, R. (2011). *Parameters of context: from theory to model and application*. PhD thesis, Department of Linguistics, Macquarie University.

Wegener, R. (2015). *Continuing Discourse on Language. A functional perspective, Vol. 1*, chapter Studying language in society and society through language: context and multimodal communication. Equinox.