# Introduction

Jörg Cassens

Data and Process Visualization
SoSe 2017

Universität Hildesheim
medieninformatik
IMAI – Institut für Mathematik und Angewandte Informatik

## Inhaltsverzeichnis



## 1 Examples

**Video 2.1: Transporting Meaning**

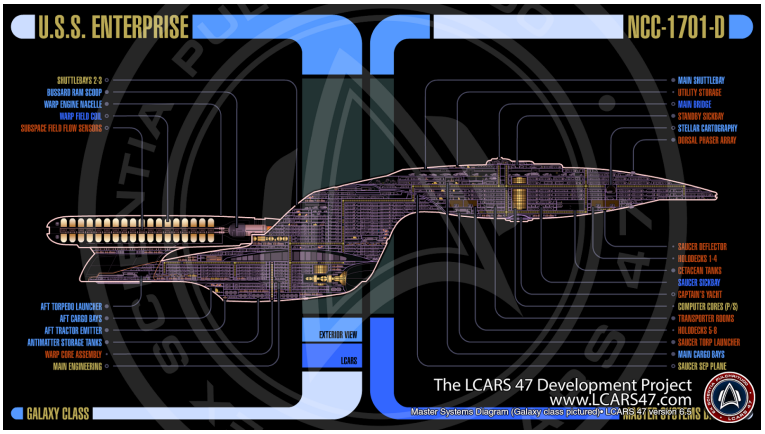

☞ ST:VOY S06E20 – The Good Shepherd
2:50

### The Good Shepherd

## Discuss

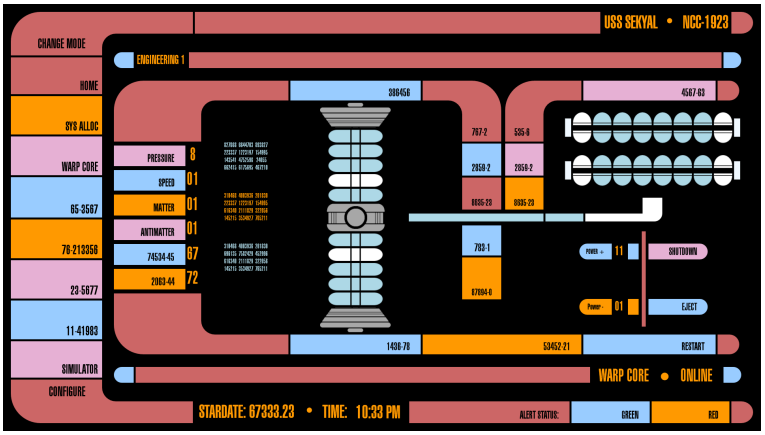
- Visualization of the chain of command
- What else?

### LCARS: Ship Diagrams



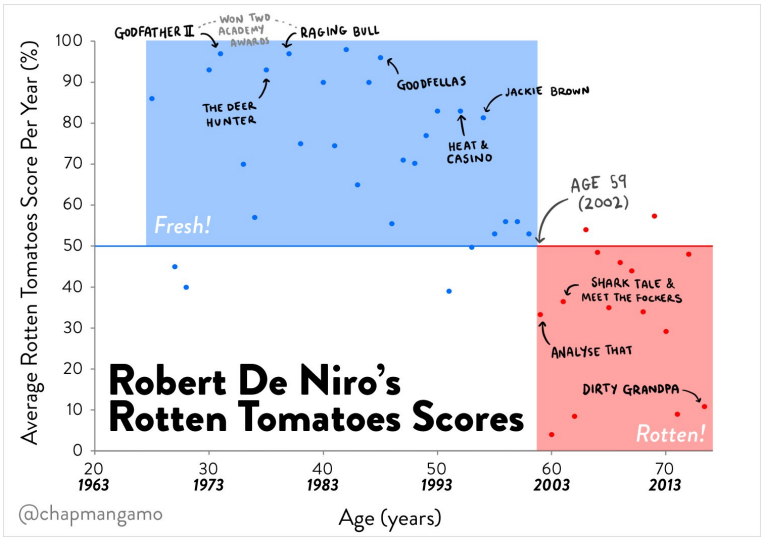

☞ LCARS project

### LCARS: Warpcore Controls

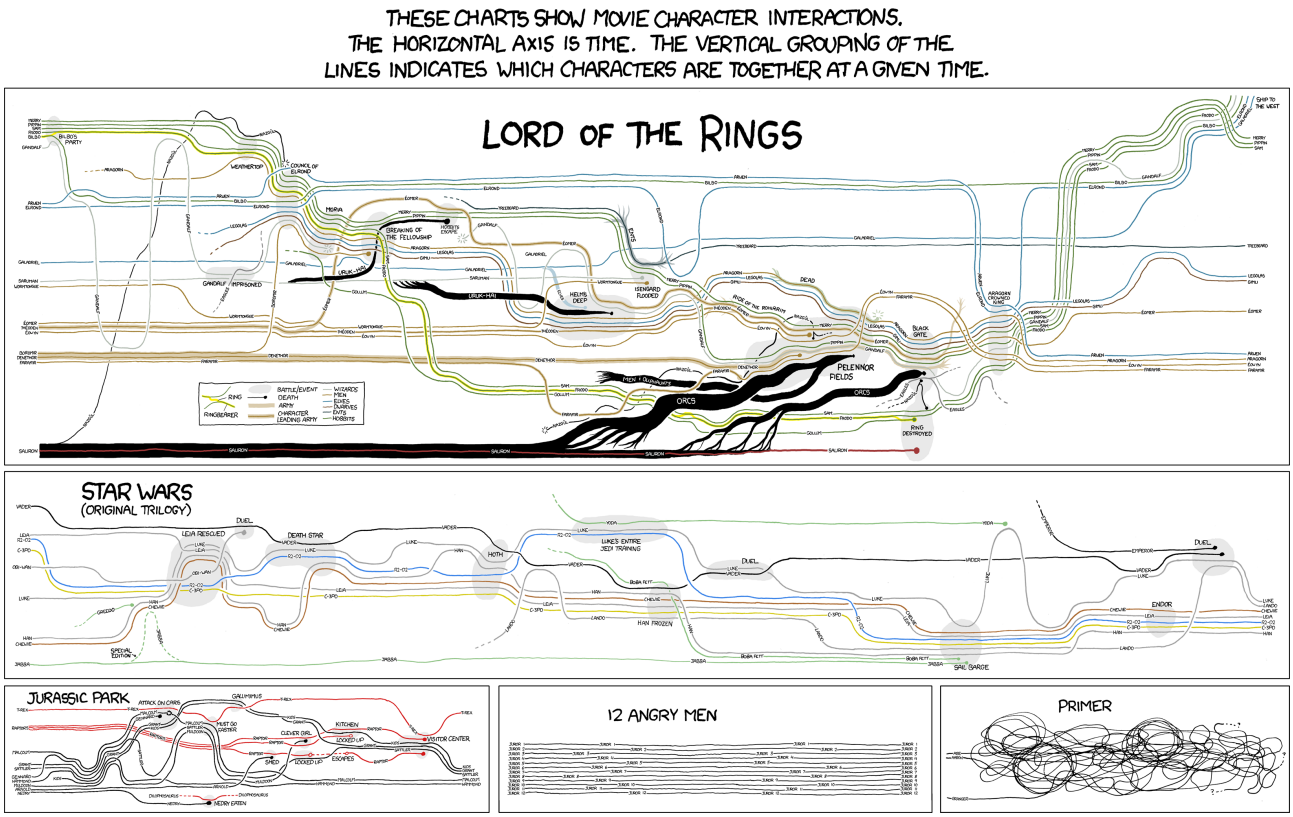



☞ throughthepanes.wordpress.com

### Robert de Niro & Rotten Tomatoes



## Movie Narrative Charts



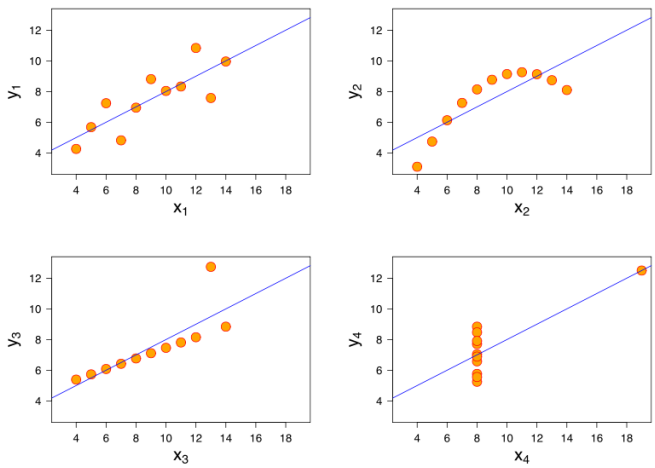

☞ xkcd: Movie Narrative Charts

## Data?

| x1 | y1 | x2 | y2 | x3 | y3 | x4 | y4 |
|---|---|---|---|---|---|---|---|
| 10 | 8.04 | 10 | 9.14 | 10 | 7.46 | 8 | 6.58 |
| 8 | 6.95 | 8 | 8.14 | 8 | 6.77 | 8 | 5.76 |
| 13 | 7.58 | 13 | 8.74 | 13 | 12.74 | 8 | 7.71 |
| 9 | 8.81 | 9 | 8.77 | 9 | 7.11 | 8 | 8.84 |
| 11 | 8.33 | 11 | 9.26 | 11 | 7.81 | 8 | 8.47 |
| 14 | 9.96 | 14 | 8.1 | 14 | 8.84 | 8 | 7.04 |
| 6 | 7.24 | 6 | 6.13 | 6 | 6.08 | 8 | 5.25 |
| 4 | 4.26 | 4 | 3.1 | 4 | 5.39 | 19 | 12.5 |
| 12 | 10.84 | 12 | 9.13 | 12 | 8.15 | 8 | 5.56 |
| 7 | 4.82 | 7 | 7.26 | 7 | 6.42 | 8 | 7.91 |
| 5 | 5.68 | 5 | 4.74 | 5 | 5.73 | 8 | 6.89 |

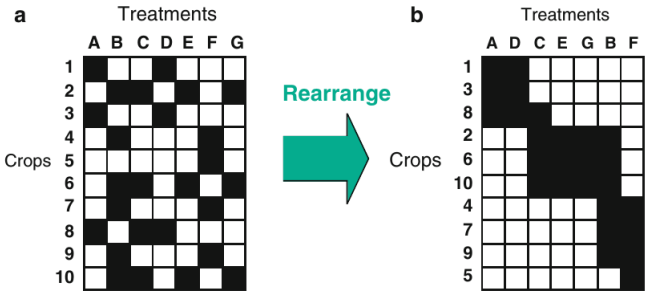
Sources: Anscombe, "Graphs in Statistical Analysis", as cited by Kirk (2012); Tufte (2001)

## Insight




Sources: Anscombe, "Graphs in Statistical Analysis", as cited by Kirk (2012); Tufte (2001)

### Effort




Reordering gives insight (Spence, 2014)

### Assignment 2.1: What is Visualization?

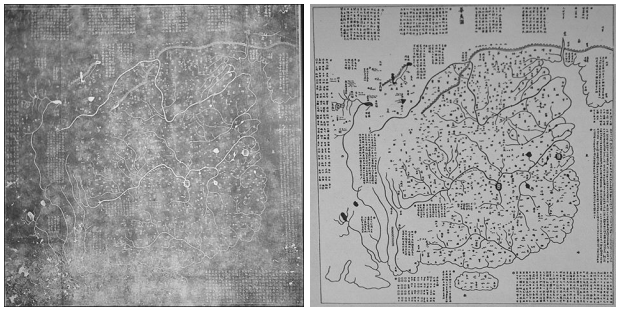
- Group work in class
- Please discuss: what is a visualization?
  - Are there characteristic elements?
    - What is the "visulizationess" of an image?
  - What does it depend on?
  - What is it good for?
- Can you come up with a definition in 3 sentences?

## 2 History

### Upper Paleolithic

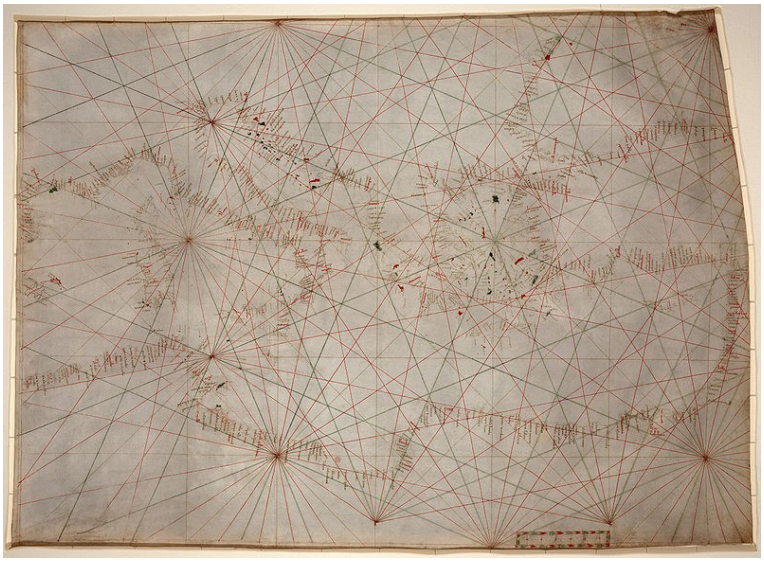
☞ Caves of Lascaux – wikipedia

### 1137



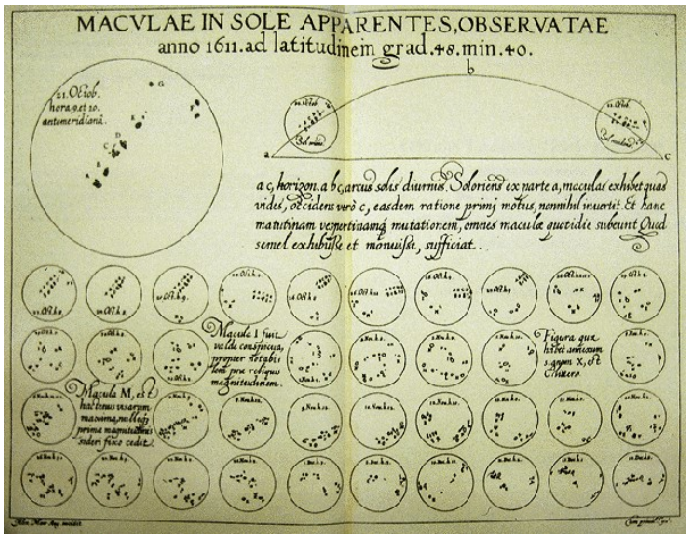
Map of China with rivers and villages (Source: Jänicke (2016))

**ca 1330**



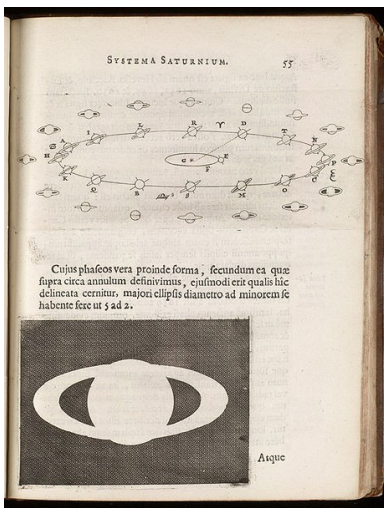
Map of the Mediterranean (Source: Jänicke (2016))

**1615**



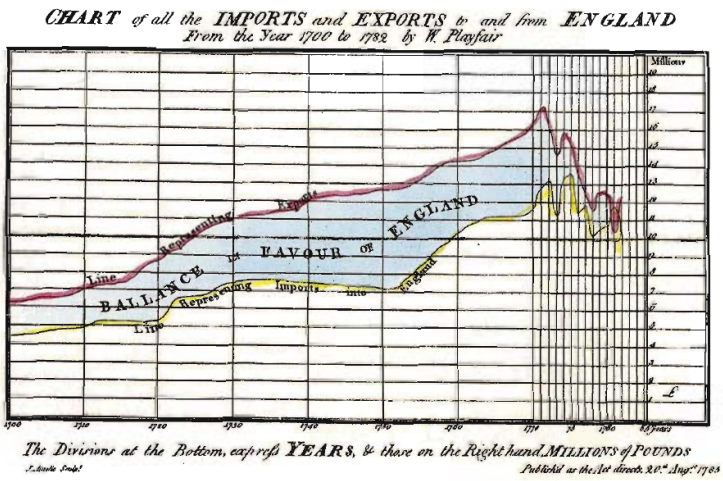

Galileo Galilei: Solar Spots (Source: Jänicke (2016))

**1659**



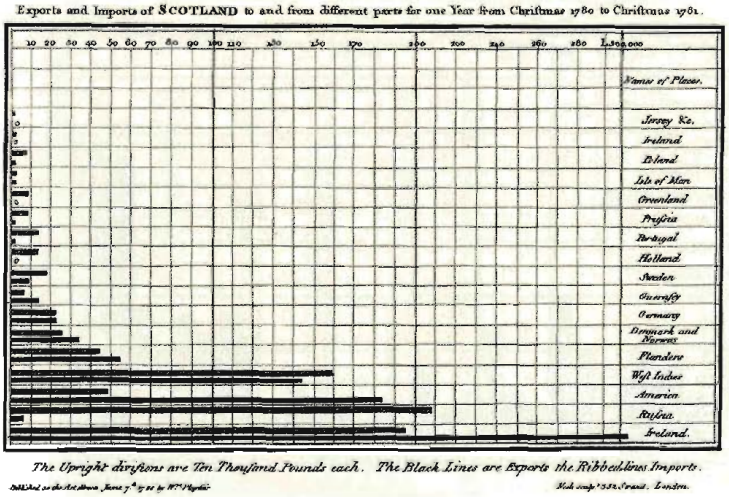

Christiaan Huygens – Systema Saturnium (Source: Jänicke (2016))

**1786**



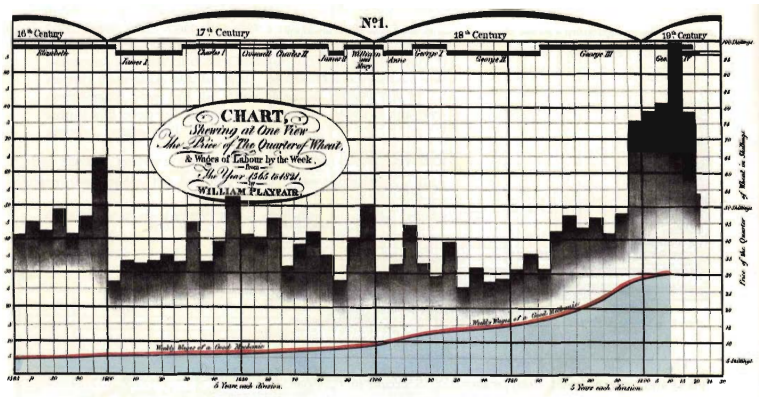

Playfair – trade deficits (Source: Tufte (2001))

**1786**




Playfair – Scottish im- and exports (Source: Tufte (2001))

**1821**




Playfair – prices, wages, royal reigns (Source: Tufte (2001))

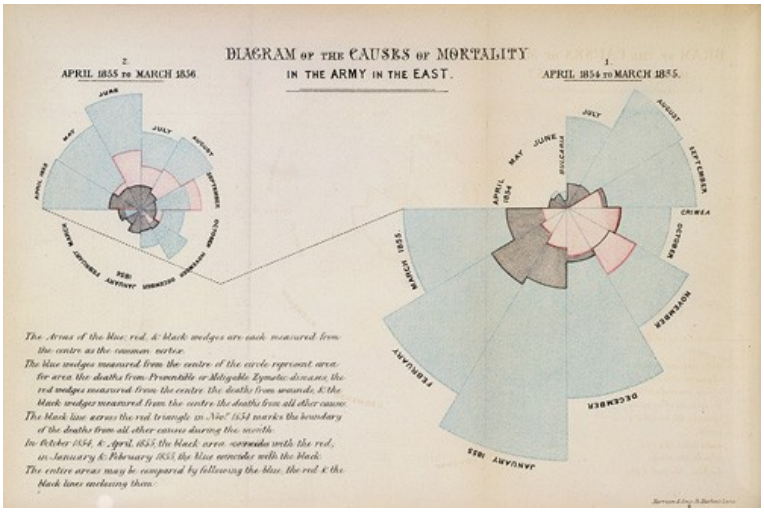
**1845**


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

**1845**

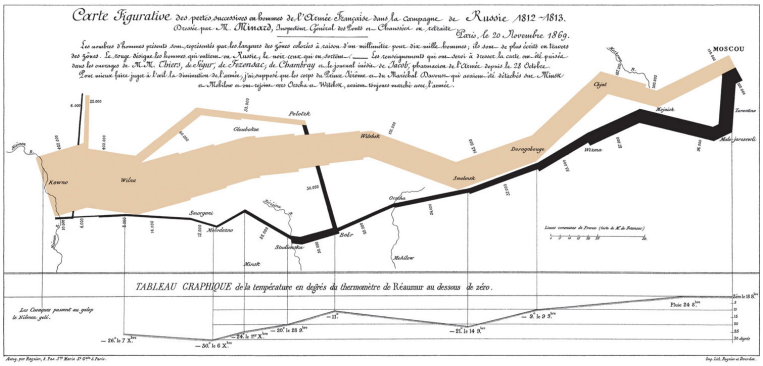

John Snow – Cholera Epidemic of London (Detail) (Spence, 2014)

**1855**



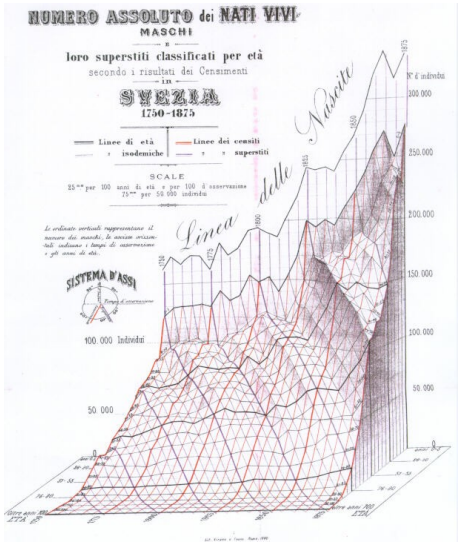

Florence Nightingale – Cause of death over time (Source: Jänicke (2016))

**1869**



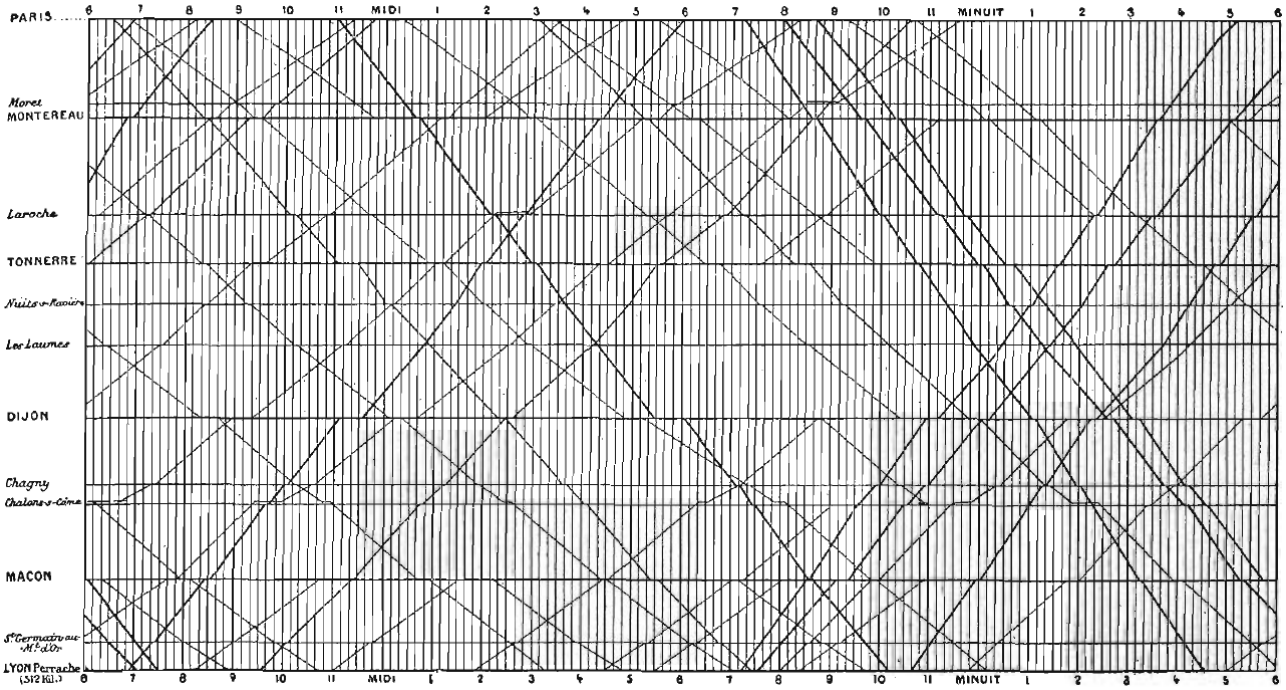

Charles Minard – Napoleon's Russian Campaign

**1880**



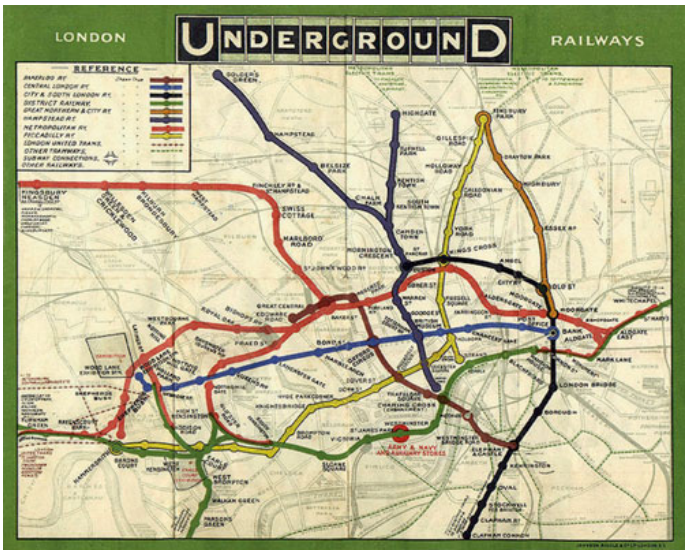

Luigi Perozzo – 3D-Visualizations (Source: Jänicke (2016))

**1885**




Marey – train timetable (Source: Tufte (2001))

**pre 1931**

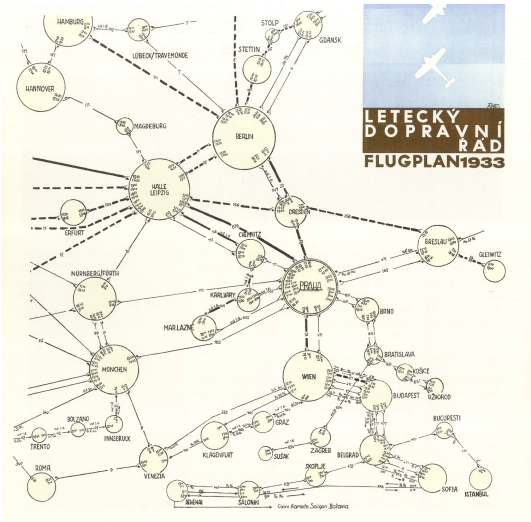



Map of the London Underground (Spence, 2014)

**1931**

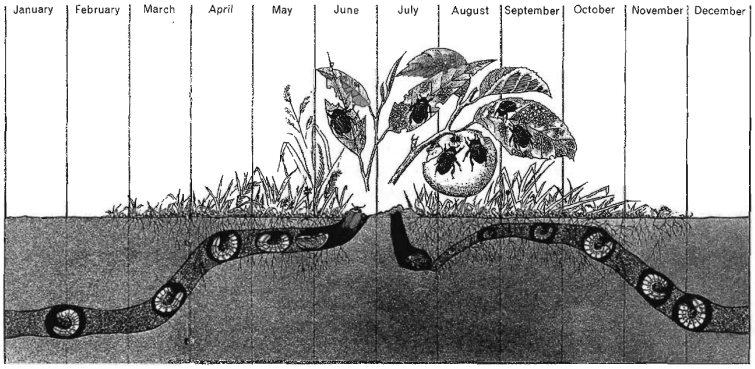

Harry Beck's new map of the London Underground (Spence, 2014)

**1933**




Brochure with flight connections (Source: Jänicke (2016))

**1965**




Newman – Man and Insects: Tufte (2001))

**1982**

Maya Lin – Vietnam War Memorial (Source: Jänicke (2016))

**Process**

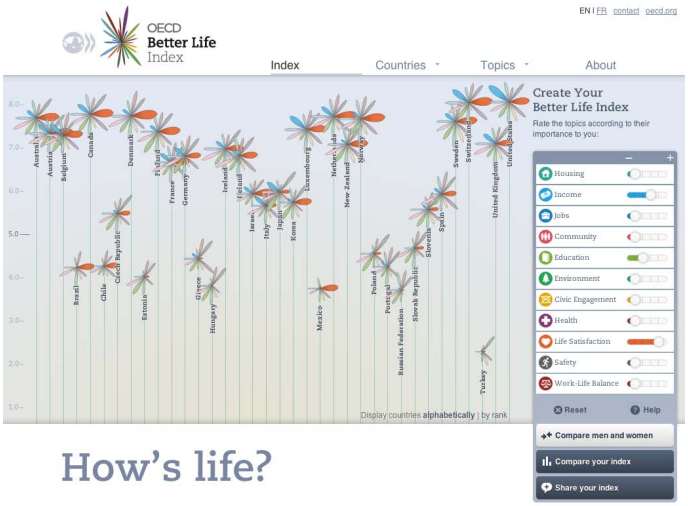
- But where are the Computers?
- And what *is* visualization?

  "visualization: the activity of forming a mental model of something" (Spence, 2014)

- Visualization is then, by definition, a human activity
  - Nevertheless, it can be enhanced immensely by means of computers

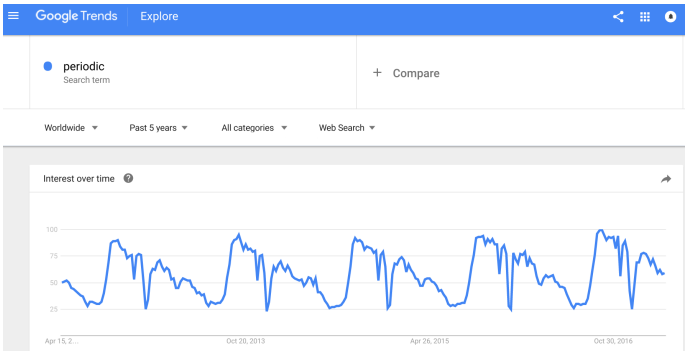
"The ability to take data—to be able to understand it, to process it, to extract value from it, to visualize it, to communicate it—that's going to be a hugely important skill in the next decades." Hal Varian, Google

**Modern Visualization I**




☞ OECD Better Life Index

### Modern Visualization II




☞ Google Trends: Periodic

## 3 Communicate

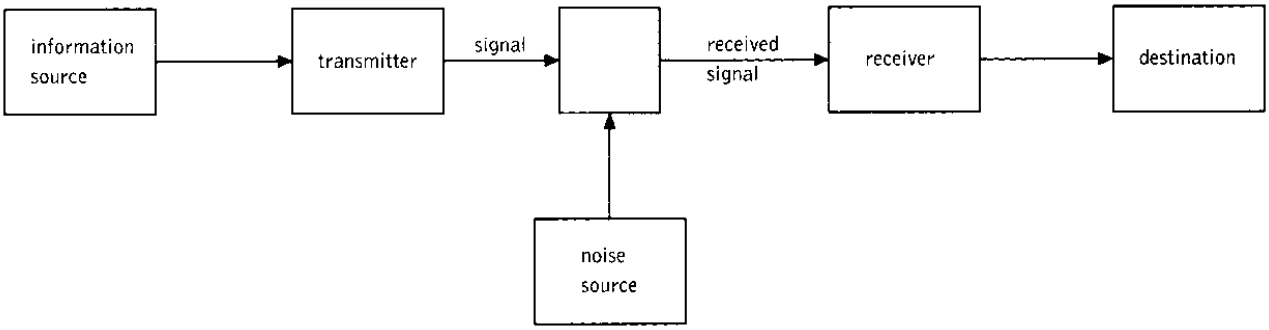
### Field of Visualization

- Draws from many fields
- Requires a deep and broad knowledge across several traditionally discrete subjects, including cognitive science, semiotics, statistics, graphic design, cartography, and computer science
- Goal: Communication

### Communication: Kirk

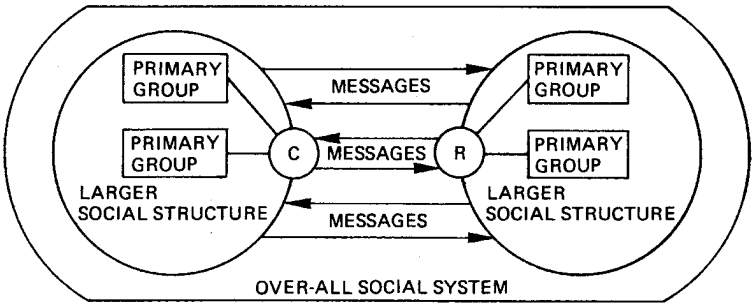

Source: Kirk (2012)

**Communication: Shannon & Weaver**



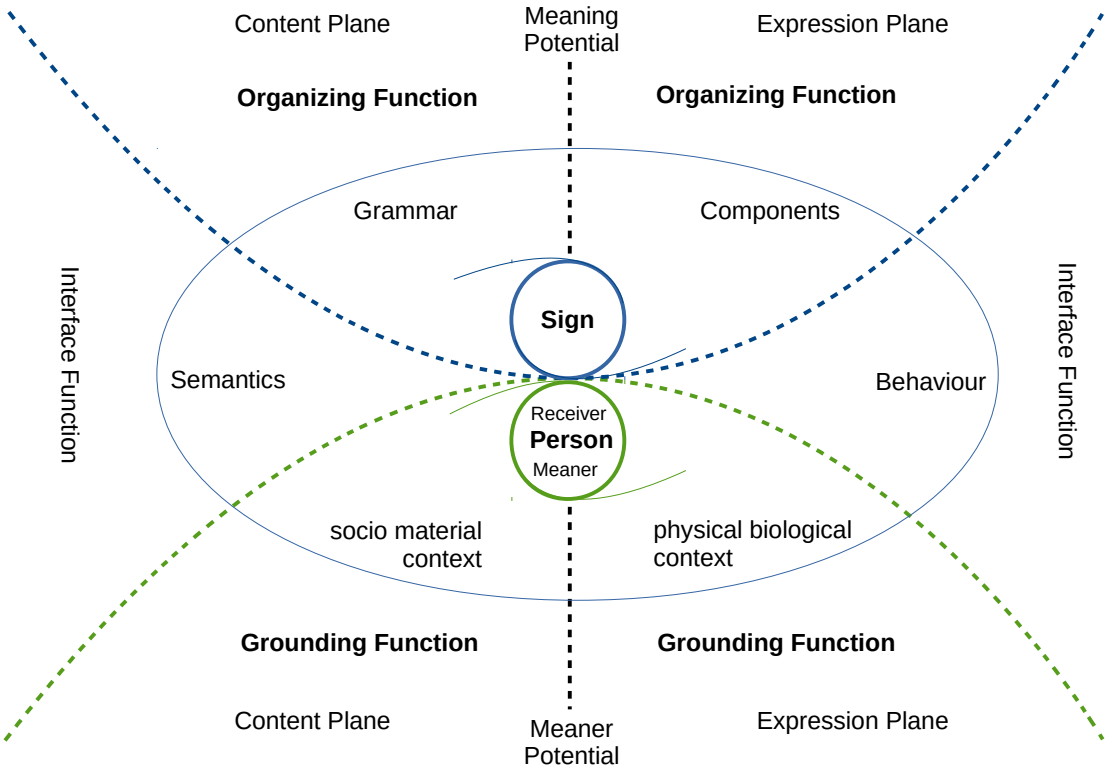

Source: Shannon & Weaver, here: Kress and van Leeuwen (2006)

**Communication: Riley & Riley**




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

**Communication: Wegener**




Source: Wegener (2011, 2015)

# 4 Semiotics

**Usefulness**

- Long history of Visualization
- We seem to be able to read the information even from old visualizations
  - But we cannot figure out all the details of the paintings in the caves of Lascaux
- What makes visualizations work?
  - Or: What makes communication work?
- This is one of the questions that are being looked at in the field of semiotics

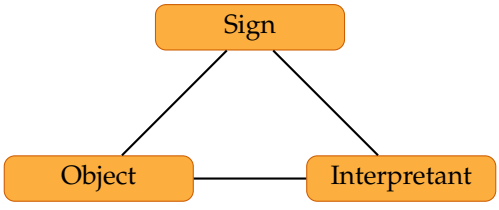
## 4.1 Concepts

**Semiotics**

- Semiotics is the science of signs or the study of sign systems (Fawcett, 1992).
- Semiotics, or semeion, was originally peculiar to medicine, referring to inference on the basis of some outward manifestation of state (or sign) (Eco, 1984).
- We can think of semiotics as a perspective, as a means of looking at anything from the point of view of how it generates meaning (Halliday, 1992).
- Semiotics deals with understanding sense making processes and sense making systems.
  - Interaction is a process of exchanging and interpreting signs, symbols referring to and standing for something else.
  - The users of a computer system see their interaction with the system against this background.

**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




**The World as Semiotic**

- If we take the perspective of looking at anything from the point of view of how it means, we are in the position of viewing all artifacts as potentially meaning bearing (Fawcett, 1992).
- Systemic Functional Linguistics (SFL) is a social semiotic theory that sets out from the assumption that humans are social beings that are inclined to interact (Halliday, 1978).
- In addition, Halliday states that human communication is inherently multi-modal.

## 4.2 Systemic-Functional Theory of Language

### Systemic Functional Theory of Language

- Halliday combines the strengths of the approaches of Saussure (1966), Peirce (1904) and Voloshinov (1973) (Cassens and Wegener, 2008).
- Saussure: the tradition of relational thinking.
- Pierce: the understanding that different modalities have consequences for the structure of meanings
- Voloshinov: the insistence that the sign is social.

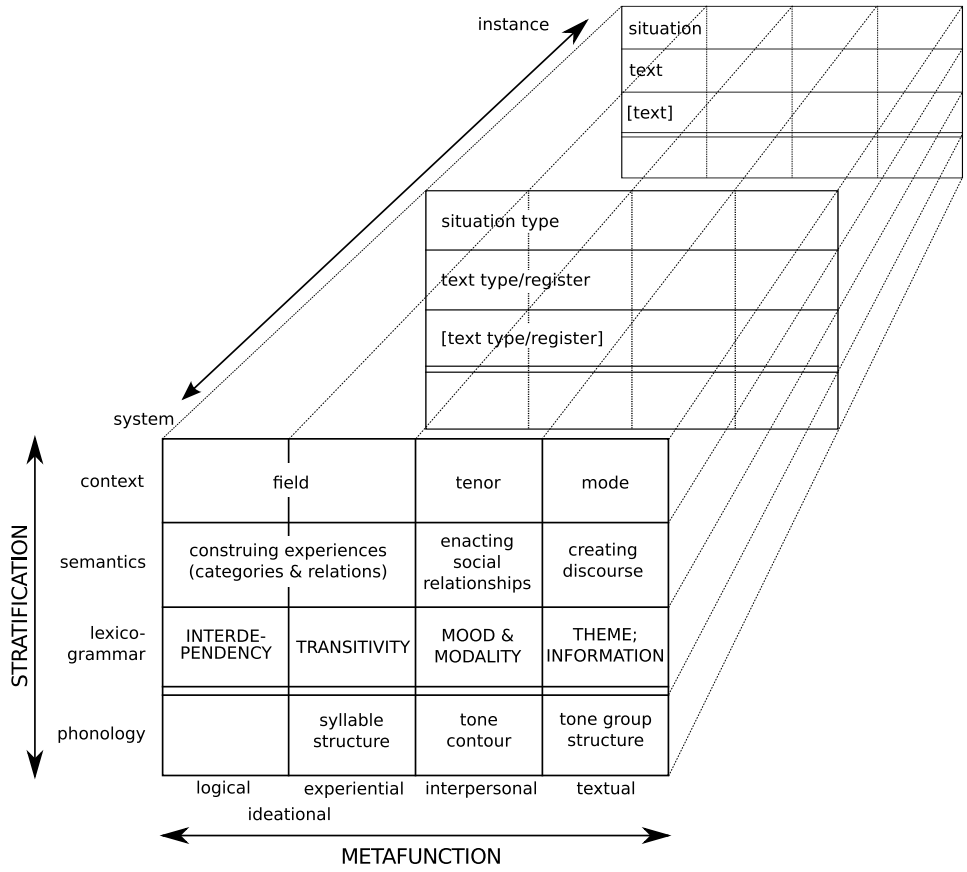
### Stratification & Register

- **Stratification:** A stratified model of language systems including:
  - Sound Systems – phonetics, phonology, gesture, pixels etc.
  - Lexicogrammar – lexis/grammar; or wording and structure
  - Semantics – the meaning system
  - Context – culture and situation; elements of the social structure as they pertain to meaning
- **Register:** Dialectic relation of system and instance
  - System – at the level of context the culture
  - Instance – at the level of context the situation that we are in
  - Register – dialectic relation
    - Abstraction of instances which typically share a similar structure
    - Concretisation of parts of the system

### 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'

### Dimensions of Language




Adapted from Halliday and Matthiessen (2004)

## 4.3 Visual Semiotic

### Culture

- Visual language is not transparent and universally understood; it is culturally specific
- Western visual communication is deeply affected by our convention of writing from left to right
  - Writing directions of cultures vary: from right to left, from left to right, from top to bottom, in circular fashion from the centre to the outside
- Consequently different values and meanings are attached to key dimensions of visual space
- Unity of Western visual communication does not exclude the possibility of regional and social variation
- Theoretical framework: 'social semiotics'

(Kress and van Leeuwen, 2006)

### Sign

- The key notion in any semiotics is the 'sign'
- 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
- Wheels, in turn, are most plausibly represented by circles

(Kress and van Leeuwen, 2006)

**Representation**

> "We see representation as a process in which the makers of signs, whether child or adult, seek to make a representation of some object or entity, whether physical or semiotic, and in which their interest in the object, at the point of making the representation, is a complex one, 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)

**Role of Sign-Maker**

- 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 (e.g. drawing, Lego blocks, painting, speech)
- 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
- In the process of sign-making, the signifier (the form) and the signified (the meaning) are relatively independent of each other until they are brought together by the sign-maker in a newly made sign

(Kress and van Leeuwen, 2006)

**Constitution**

- The process of sign-making is the process of the constitution of a sign/metaphor in two steps: 'a car is (most like) wheels' and 'wheels are (most like) circles'
- The sign-maker's interest at this moment of sign-making has settled on 'wheelness' as the criterial feature of 'car'. He constructs, by a process of analogy, two metaphors/signs: first, the signified 'wheel' is aptly represented by the signifier 'circle' to make the motivated sign 'wheel'; second, the signified 'car' is aptly represented by the signifier 'many wheels' to make the motivated sign 'car'.
- The resulting sign, the drawing called 'this is a car', is thus a motivated sign in that each conjunction of signifier and signified is an apt, motivated conjunction of the form which best represents that which is to be meant

(Kress and van Leeuwen, 2006)

**Naturalization**

- The sign is the result of a double metaphoric process in which analogy is the constitutive principle
- Analogy, in turn, is a process of classification: ***x*** *is like* ***y*** (in criterial ways)

Which metaphors (and, 'behind' the metaphors, which classifications) carry the day and pass into the semiotic system as conventional, and then as naturalized, and then as 'natural', neutral classifications, is governed by social relations of power

- It follows that we see signs as motivated – not as arbitrary – conjunctions of signifiers (forms) and signifieds (meanings)

(Kress and van Leeuwen, 2006)

**Arbitrariness**

- In 'semiology' motivation is usually not related to the act of sign-making, but defined in terms of an intrinsic relation between the signifier and the signified
- Contrasting to a common interpretation of Peirce
  - The 'icon' is the sign in which 'the signifier-signified relationship is one of resemblance, likeness (Dyer, 1982)' – i.e. objective likeness, rather than analogy motivated by 'interest'
  - The 'index' is the sign in which 'there is a sequential or causal relation between signifier and signified (Dyer, 1982)' – that is, a logic of inference, rather than analogy motivated by 'interest'
  - The 'symbol' is related to sign production, as it 'rests on convention, or contract (Dyer, 1982)', but this very fact makes it 'arbitrary', 'unmotivated', a case of meaning by decree rather than of active sign-making

(Kress and van Leeuwen, 2006)

**Motivation**

- In contrast, signs are never arbitrary, and 'motivation' should be formulated in relation to the sign-maker and the context in which the sign is produced,
  - Not in isolation from the act of producing analogies and classifications
- Sign-makers use the forms they consider apt for the expression of their meaning, in any medium in which they can make signs
  - Children treat a cardboard box as a pirate ship
  - They do so because they consider the material form (box) an apt medium for the expression of the meaning they have in mind (pirate ship), and because of their conception of the criterial aspects of pirate ships (containment, mobility, etc.)

(Kress and van Leeuwen, 2006)

**Application**

- The visual semiotic introduced by Kress and van Leeuwen (2006) is not purely theoretic
- The practical application is a descriptive framework for visual analysis
- For example, what type of images or graphs is best suited for conveying the intended meaning
- Another example is in critical discourse analysis', i.e. a critical look at meaning and how visual representation convey power and status (a "soft skill" in computer science)
- As visual communication becomes more an more a specialist activity – this course – this knowledge is becoming important in more and more areas
- They propose a descriptive framework based on Halliday's Metafunctions – we will incorporate some of those ideas

# 5 Framework

**Framework**

- We are introducing a coarse framework to get a better understanding of the visualizations
- grouping to develop organizational frameworks
- various ways to group visualizations
  - e.g. by user insight needs, by user task types, or by the data to be visualized
- Here: a pragmatic approach, extending the framework introduced by Börner and Polley (2014) with aspects from Kress and van Leeuwen (2006)

## 5.1 Classification

**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?
- 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

## 5.2 Semiotic Functions

**Relation to Metafunctions**

- The framework for classification we have just introduced relates to the functions of language
- In particular, we can relate them to what Halliday and Matthiessen (2004) call the Metafunctions
- We will look at the three different Metafunctions and their relation to the different aspects of our classification system
  - Field of discourse
  - Tenor of discourse
  - Mode of discourse

**Field**

- Level and type of analysis pertain to the 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)

- We are talking about **ideational** aspects.
  - What is the domain? What are the long term or short term goals? The experiential domain?
  - What is the structure, what are the networks of interaction?

**Tenor**

- The audience (and producer) pertains to the 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)

- We are talking about **interpersonal** aspects.
  - What is the power structure between actors involved?
  - What is the agentive role?
  - What is the competence of the actors?

**Mode**

- The medium used pertains to the 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)

- We are talking about **textual** aspects.
  - What medium is used?
  - What is the type of interaction (dialogic, monologic)?
  - What is the rhetorical thrust?

# 6 Tutorial

**Assignment 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

**Assignment 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

**References**

# Literatur


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

Cassens, J. and Wegener, R. (2008). Making use of abstract concepts – systemic-functional linguistics and ambient intelligence. In Bramer, M., editor, *Artificial Intelligence in Theory and Practice II – IFIP 20$^{th}$ World Computer Congress, IFIP AI Stream*, volume 276 of *IFIP*, pages 205–214, Milano, Italy. Springer.

Dyer, G. (1982). *Advertising as Communication*. Methuen, London.

Eco, U. (1984). *Semiotics and the philosophy of language*. Macmillan, Basingstoke, London.

Fawcett, R. P. (1992). Book reviews: A theory of computer semiotics: Semiotic approaches to construction and assessment of computer systems. *Computational Linguistics*, 18(4).

Halliday, M. A. (1978). *Language as a Social Semiotic: the social interpretation of language and meaning*. University Park Press.

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

Halliday, M. A. and Matthiessen, C. M. (2004). *An Introduction to Functional Grammar, Third edition*. Arnold, London, UK.

Halliday, M. A. K. (1992). New ways of meaning: the challenge to applied linguistics. In Putz, M., editor, *Thirty Years of Linguistic Evolution*. John Benjamins Publishing, Philadelphia/Amsterdam.

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.

Peirce, C. S. (1904). New elements (kaina stoicheia). In Eisele, C., editor, *The New Elements of Mathematics by Charles S. Peirce*, volume 4, Mathematical Philosophy, pages 235–263.

Saussure, F. d. (1966). *Course in General Linguistics*. McGraw-Hill.

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.

Voloshinov, V. N. (1973). *Marxism and the Philosophy of language*. Seminar Press, New York.

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.