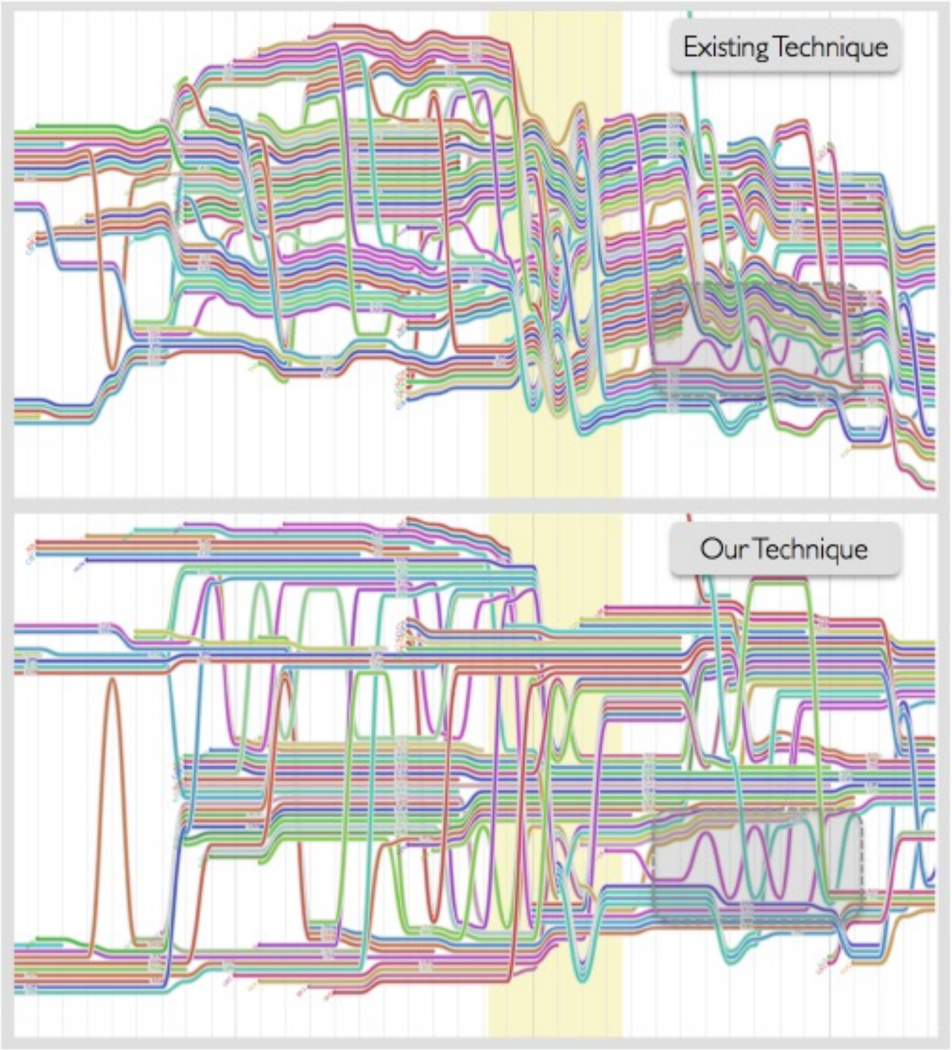
Information Visualization

(TNM111)



[Tanahashi and Ma, InfoVis 2012]

Instructor

- **Prof. Dr. Andreas Kerren**
  - Head of the iVis research group
  - https://liu.se/en/research/ivis

INFORMATION VISUALIZATION @ LiU

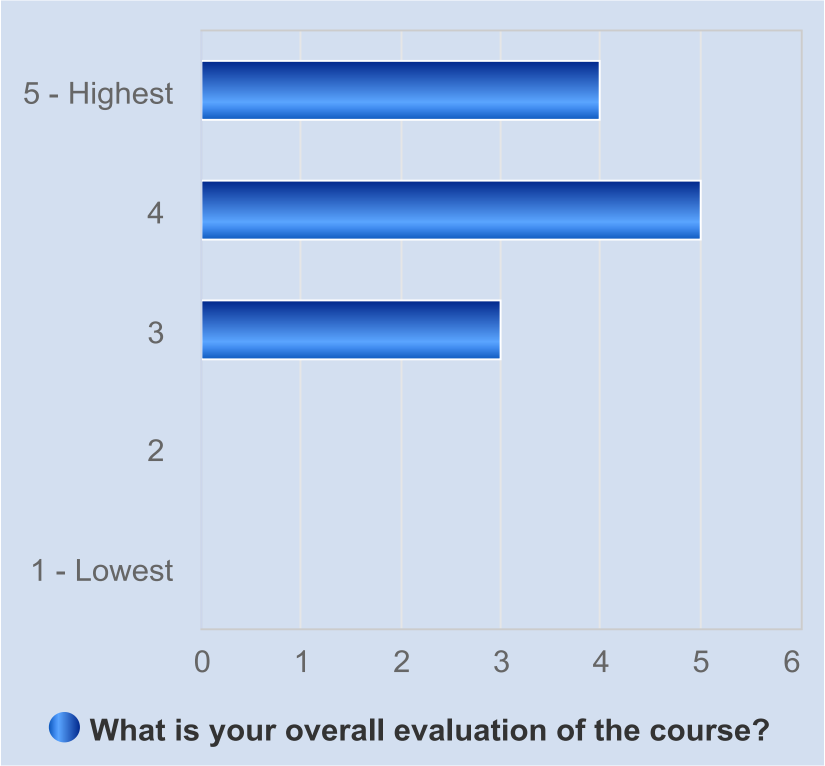
- Linköping University
  Department of Science and Technology
  Kopparhammaren 2, Room 3028
  SE-602 33 Norrköping
- Email: andreas.kerren@liu.se
- Personal Webpage: https://www.itn.liu.se/~andke01/
- Tel.: 011 36 34 06

Objectives

- Classify typical user tasks for visualizations in context of the domain situation
- Define and explain the visualization techniques and tools discussed in the course
- Describe validation and evaluation methods for information visualization tools and approaches
- Represent data by expressive and effective visualizations using state-of-the art methods, software, and tools
- Critically reflect on visualization and interaction approaches in the light of current theories and research
- Make well-grounded design choices in the context of various tasks and data constraints

Previous Course Evaluation

- VT 2023
  - 36 students of which 12 responded (~33%)
  - Summary rating: 4.08 (of 5.00)
    - Very good results in general especially for the relevance of the course for their studies, and content
    - Teaching material and methodology was acknowledged



Previous Course Evaluation

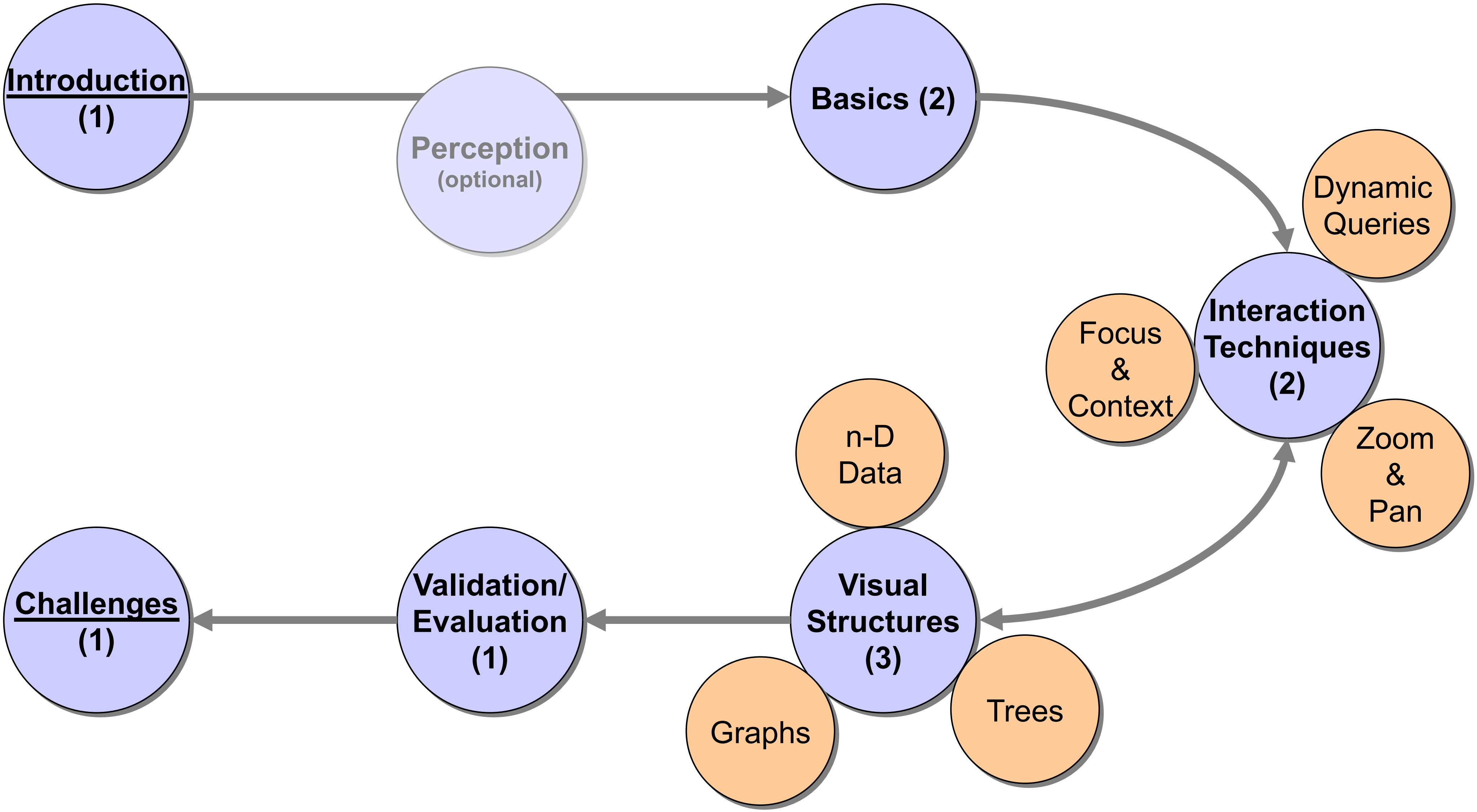
- This year, VT2024, there will be only a few changes
  - Based on student feedback, discussions to better align related program courses against each other, but also to secure a better foundational and conceptual knowledge
  - Most important changes (see corresponding document on Lisam)
    - Lectures are very similar compared to last year, but some topics have been slightly modified and updated
    - Perception lecture remains in TNM093 (added to Lisam as supplemental material)
    - Assignment structure remains more or less the same
    - More time to prepare the assignments and to implement stuff
    - More help sessions will be offered by the TAs which all are optional (see course schedule)

TNM111: Content

1. Introduction and Motivation (2024 01 15)
2. InfoVis Basics I (Reference and Design Models) (2024 01 16)
3. InfoVis Basics II (Data/Task Abstractions) (2024 01 18)
4. Interaction I (Dynamic Queries) (2024 01 19)
5. Interaction II (Zoom&Pan, O&D, F&C) (2024 01 23)
6. 1D, 2D, 3D, and Multidimensional Data (2024 01 26)
7. Hierarchies (Trees) (2024 02 01)
8. Networks (Graphs) (2024 02 02)
9. Validation and Evaluation Methods (2024 02 08)
10. InfoVis Research Challenges (2024 02 15)

- URL: **https://liuonline.sharepoint.com/sites/Lisam_TNM111_2024VT_XS**

TNM111: 10 Lectures



TNM111: Assignment Part

- Five assignments (3 credits in total)
  - Theoretical & Practical Exercises
  - Readings, minor implementations and seminar presentations
  - Each assignment must be passed to get the lab credits
    - If an assignment is failed, the students will get an additional chance to improve it
    - Deadlines are strict, i.e., missing deadlines will lead to a U for the assignment in question
    - Grading criteria are listed on the Lisam page
  - U/G grading (during lab sessions), groups of 2 students (via Lisam)
- All important information (dates, procedure, rooms, …) can be found on the Lisam page of the course!
- <u>Teaching Assistants:</u>
  - **Elias Elmqvist + Peilin Yu**

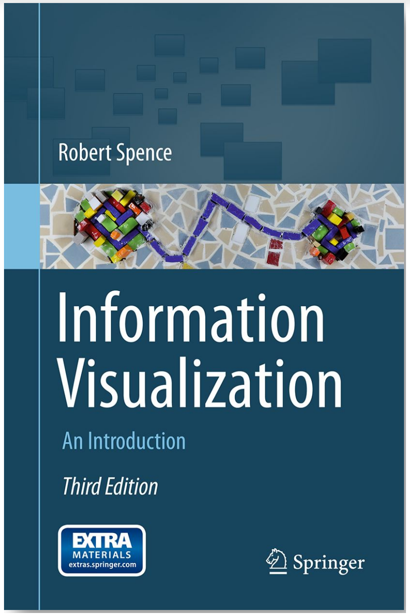
TNM111: Exam Part

- Oral examination (3 credits)
  - About 25 minutes for each student group (2 students)
  - Will happen in weeks 11 and 12 of 2024
  - Individual dates/times per exam will be chosen via Doodle
  - See Lisam course page later on for more information and exact dates

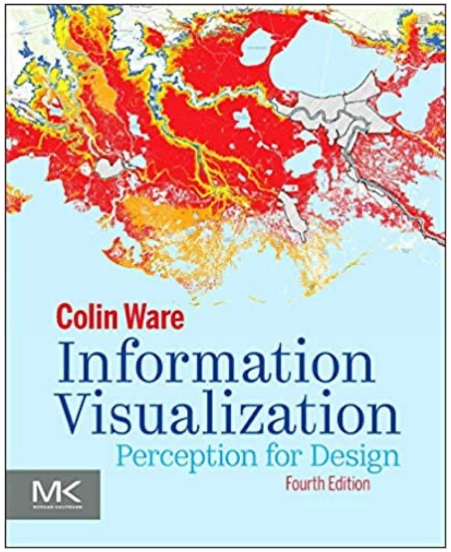
- To pass the course
  - You have to pass both parts
  - Final grade is the oral exam grade

Books

- [Spe] R. Spence. Information Visualization – An Introduction. 3rd Edition, Springer, ISBN 978-3-319-07340-8, 2014.

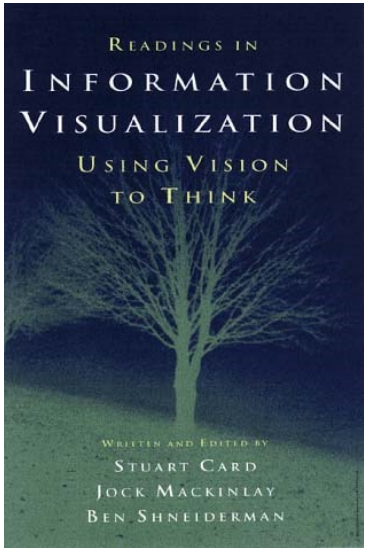


- [War] C. Ware. Information Visualization: Perception for Design. 4th Edition, Morgan Kaufman, ISBN 978-0128128756, 2019.

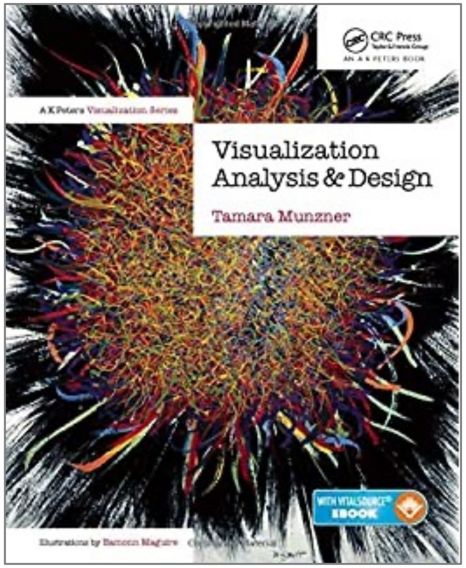


Books

- [CMS] S. K. Card, J. Mackinlay, and B. Shneiderman. Readings in Information Visualization: Using Vision to Think. Academic Press, ISBN 1-55860-533-9, 1999.

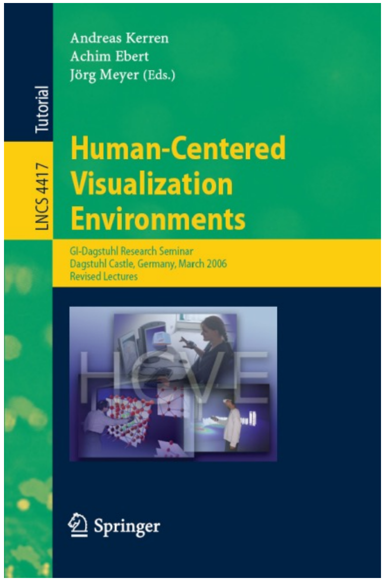


- [Mun] T. Munzner. Visualization Analysis and Design. 1st Edition, A K Peters/CRC Press, ISBN 978-1466508910, 2014.

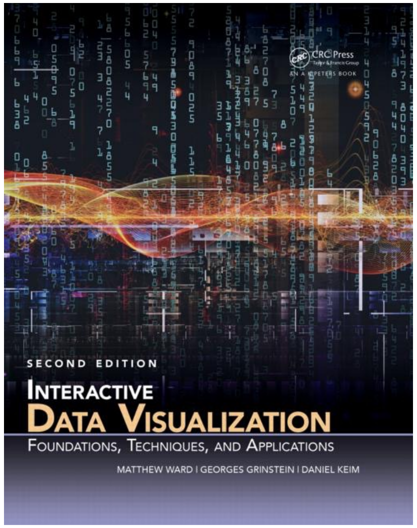


Books

- [KEM] A. Kerren, A. Ebert, and J. Meyer. Human-Centered Visualization Environments. LNCS 4417, Springer, ISBN 978-3-540-71948-9, 2007.

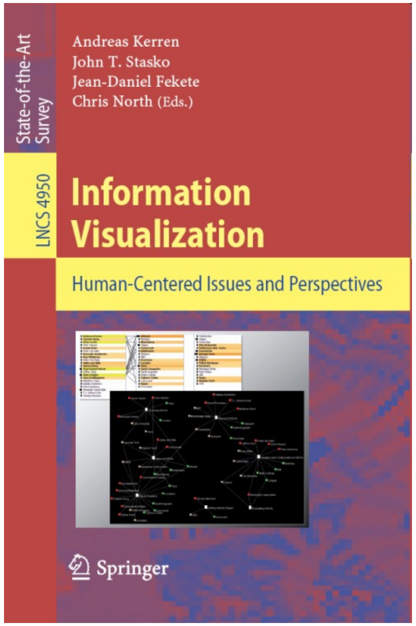


- [WGK] M. Ward, G. G. Grinstein, and D. Keim. Interactive Data Visualization − Foundations, Techniques, and Applications. 2nd Edition, A K Peters/ CRC Press, ISBN 978-1-482-25737-3, 2015.

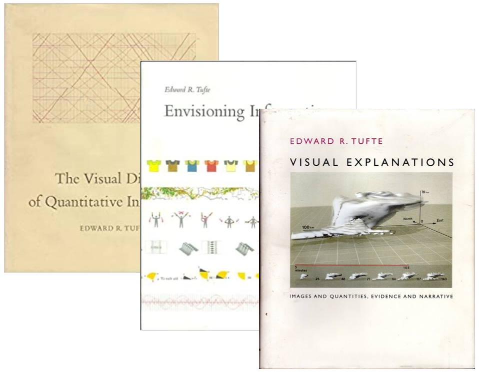


Further Literature

- [KSFN] A. Kerren, J.T. Stasko, J.-D. Fekete, and C. North. Information Visualization. LNCS 4950, Springer, ISBN 978-3-540-70955-8, 2008.



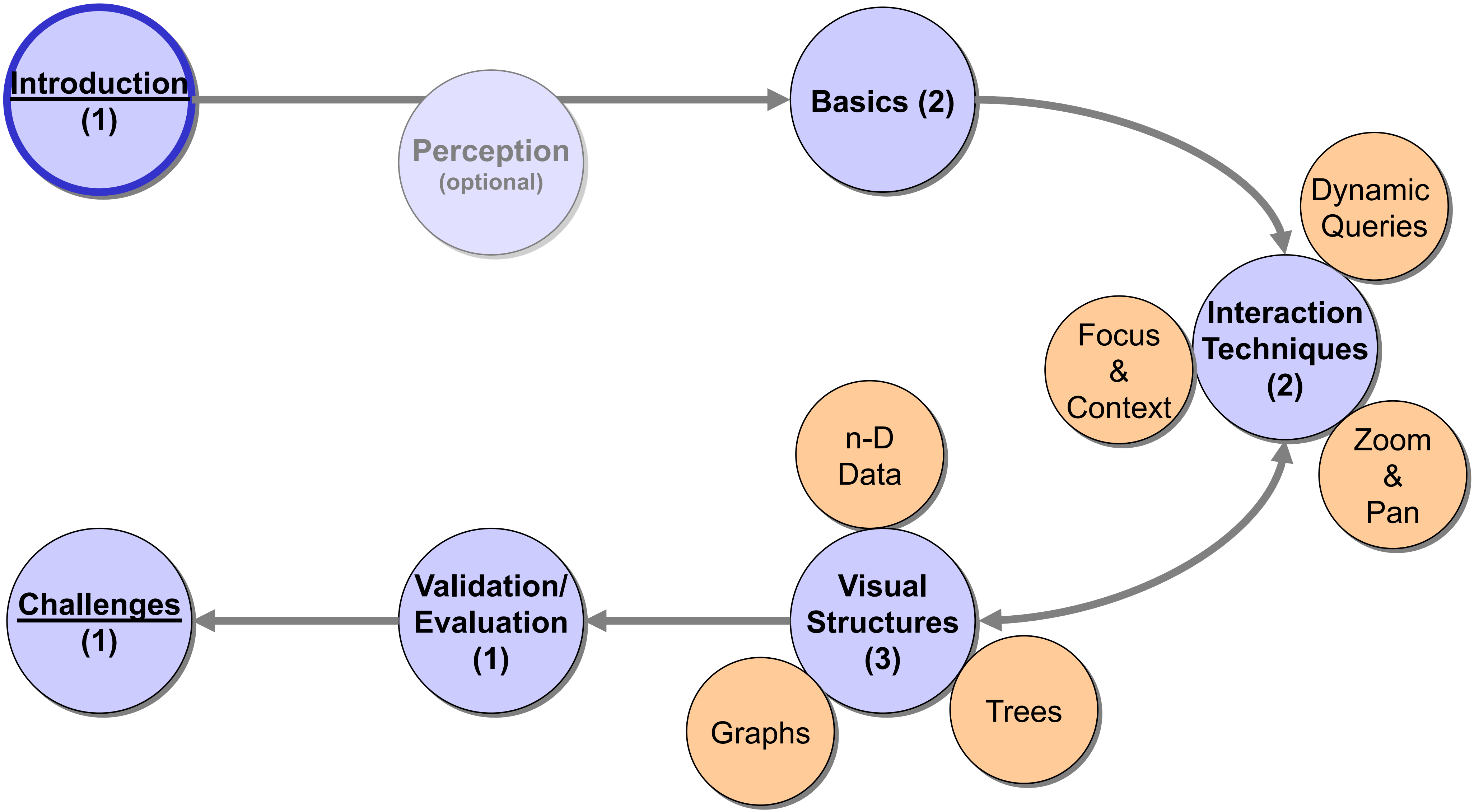
- Conference Proceedings
  - IEEE Conference on Information Visualization (InfoVis)
  - EG/VGTC Conference on Visualization (EuroVis)
  - IEEE Pacific Visualization Symposium (PacificVis), …
- Journals
  - Information Visualization, SAGE, …
- Books of Edward R. Tufte
  - [Tuf1] The Visual Display of Quantitative Information, 1983
  - [Tuf2] Envisioning Information, 1990
  - [Tuf3] Visual Explanations, 1997



Information Visualization

1. Introduction

TNM111: 10 Lectures

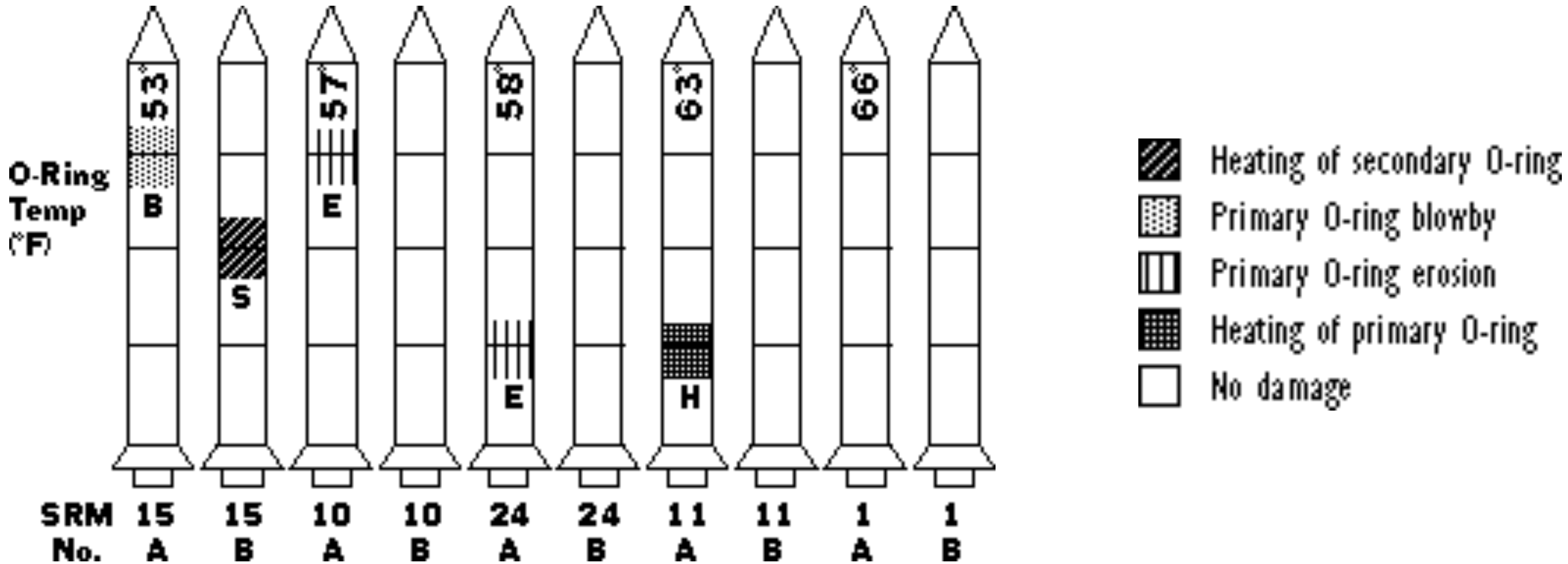


1.1 Motivation

- With the help of **suitable** InfoVis-techniques, the Challenger-catastrophe could eventually be to avoid. Why?

1.1 Motivation

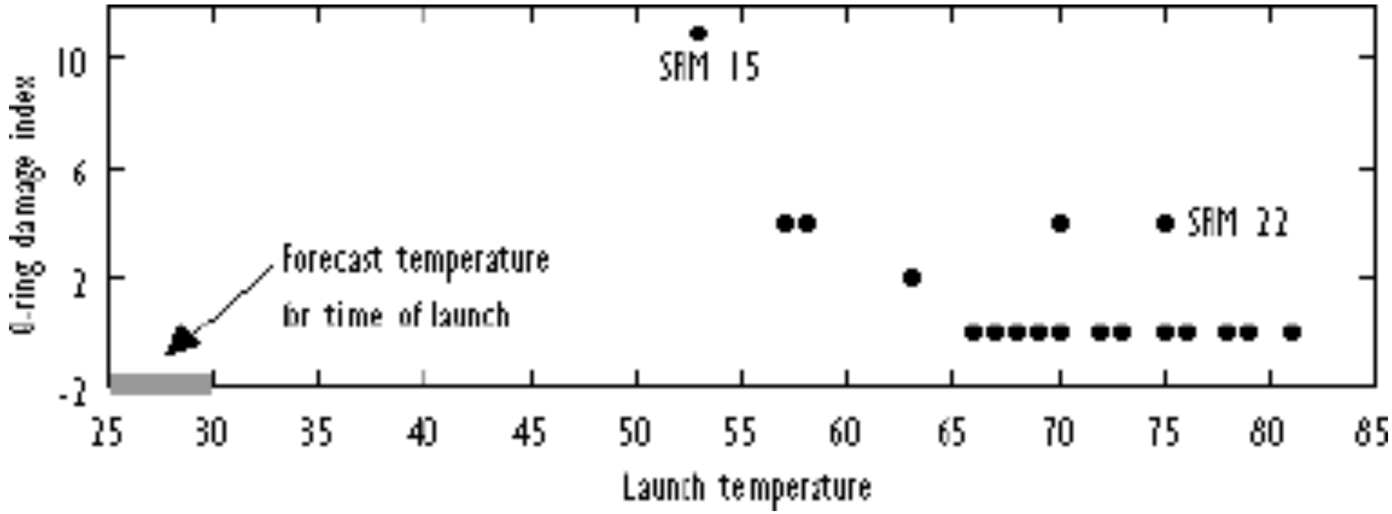
- The problem was a damaged O-ring of the boosters. The start on a very cold day was decided due to the following diagrams:



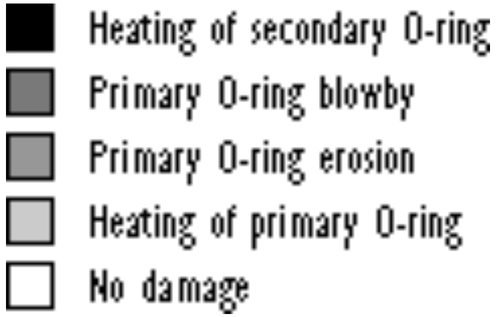
- Result: Patterns are hard to perceive. Temperature is only textually presented. Damage degrees have no “natural” graphical scale!

1.1 Motivation

- Simple solution through more clever representations
  - Scatter Plots



  - Better scale of previous damages should be added



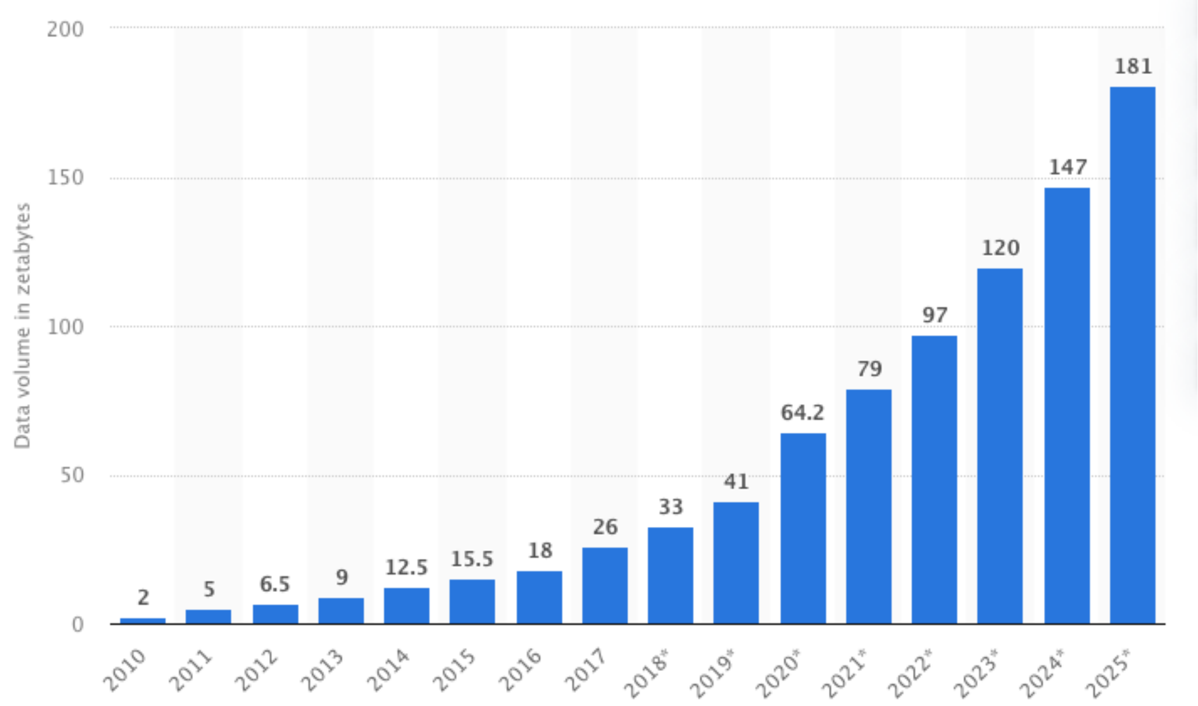
cp. [Tuf3]

1.1 Motivation

- Data explosion
  - Society will become more and more complex
  - Computers and Internet give people access to huge data sets (text, pictures, movies, …)
    - News, sports, economics, shopping, …
  - People collect and use larger data sets privately
  - Academia and industry are more and more data-driven
    - "Data is the new oil"
    - Information Overload Problem (IOP) or *Big Data*
      - Analysts are overwhelmed by data size and/or complexity (4 Vs)
      - Data is stored without filtering or refinements for future use
      - Time and money are wasted, …

1.1 Motivation

- How much data/information is created/stored?
  - Approximately, 97 Zettabyte (= $97 \cdot 10^{21}$ bytes) of new information in 2022 worldwide (estimated)
  - For comparison:
    - ~10 Zettabyte, 2008
    - ~5 Exabyte (= $5 \cdot 10^{18}$ bytes), 2002
      http://www.sims.berkeley.edu/research/projects/how-much-info-2003/
    - 1-2 Exabyte, 1999



https://www.statista.com/statistics/871513/worldwide-data-created/

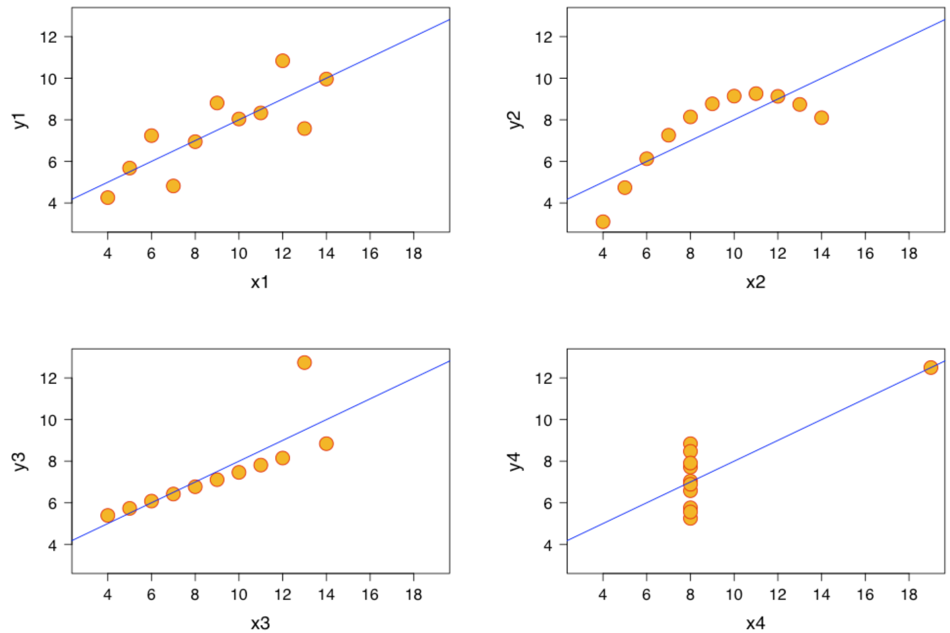
1.1 Motivation

- In many cases, an automatic analysis is *not* or *only partly* possible!

Anscombe's Quartet: Raw Data

	I		II		III		IV	
	x	y	x	y	x	y	x	y
	10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
	8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
	13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
	9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
	11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
	14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
	6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
	4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
	12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
	7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
	5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89
mean	9.0	7.5	9.0	7.5	9.0	7.5	9.0	7.5
var.	10.0	3.75	10.0	3.75	10.0	3.75	10.0	3.75
corr.	0.816		0.816		0.816		0.816	

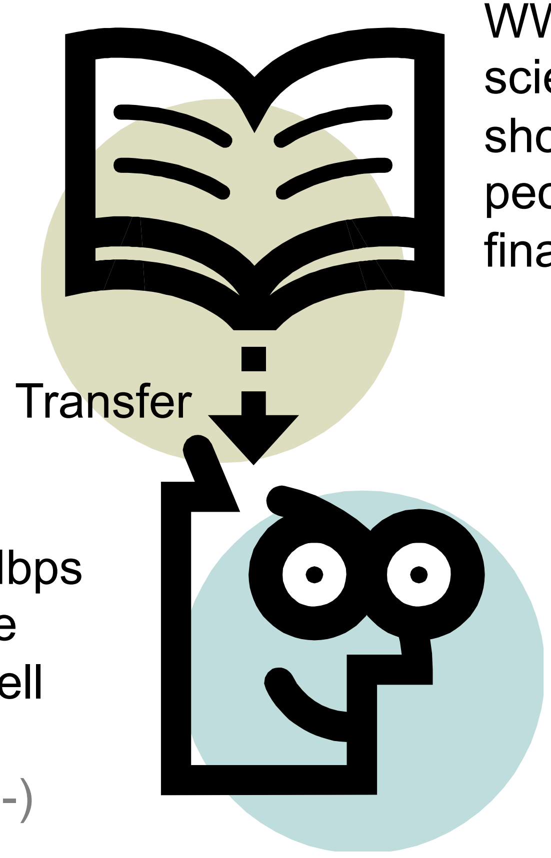
+ identical linear regression



1.1 Motivation

- Problem: How to make use of the data?

WWW, books, article, scientific data, biotech, shopping, data on specific people, economics (stock, finance), news, …



Vision: 10 Mbps
Haptic/tactile
Hearing/Smell
Taste
Telepathy? ;-)

Inspired by J. Stasko's course

http://www.eurekalert.org/pub_releases/2006-07/uops-prc072606.php

1.1 Motivation

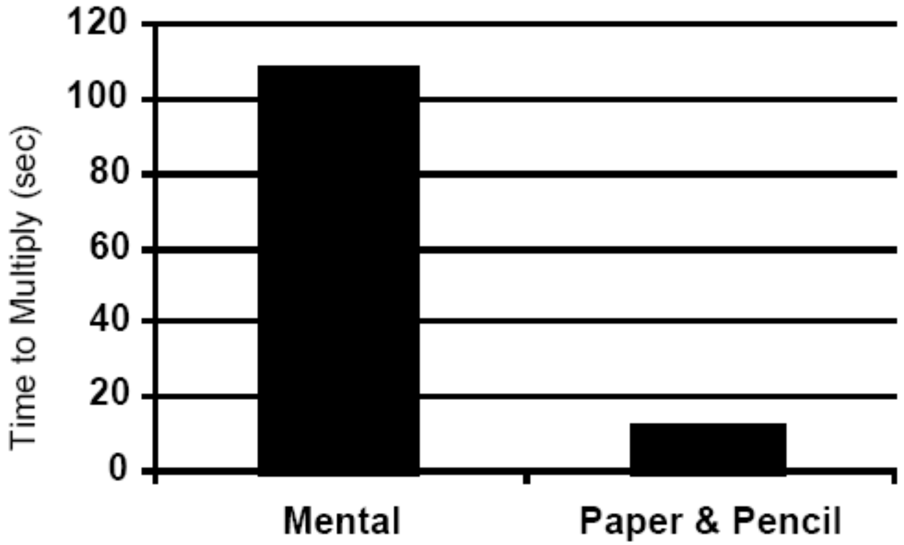
- Human Vision
  - Sense with highest bandwidth
  - Fast, parallel
  - Pattern recognition
  - Preattentive perception (ca. 60ms latency time)
  - Many people think visually

- Challenge
  - Transform the *data* into *information* (understanding, insight) and make them useful (available, analyzable) to people!

1.1 Motivation

- Visualizations help to think
- Strategy: use “external world”!
  - Internal and external representations/processes are weaved together in thought (external cognition)
  - Multiplication experiment:

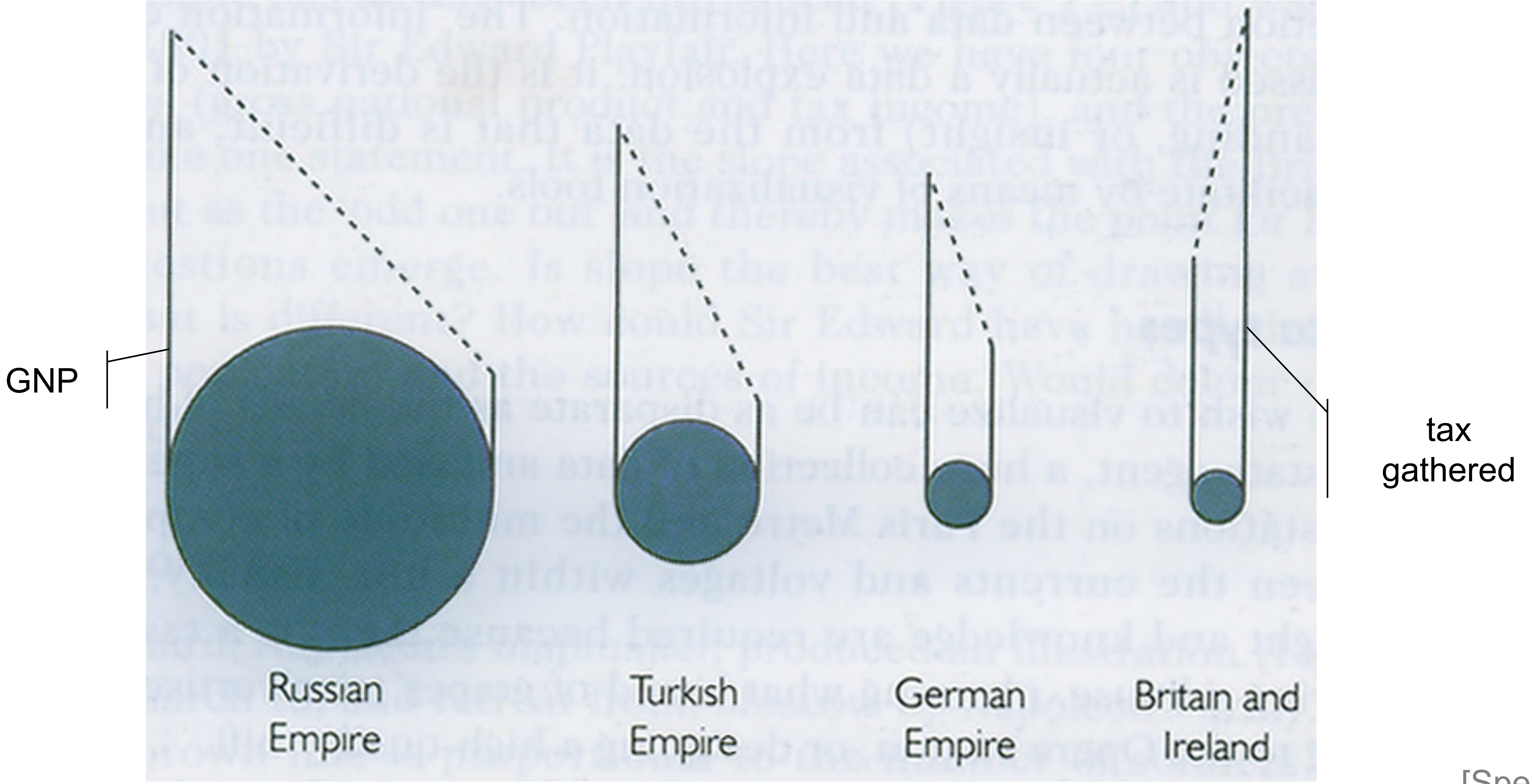
```
  34
x 72
  68
   2
2380
  1
2448
```



[CMS:1ff]

1.1.1 Classical Examples

- Sir Edward Playfair's Circles (1801)



[Spe:3f]

1.1.1 Classical Examples

- Charles Joseph Minard (1781-1870), a French engineer, already shows in a graphical representation of the march to Moscow more then 130 years ago, how to graphically show multivariate data.

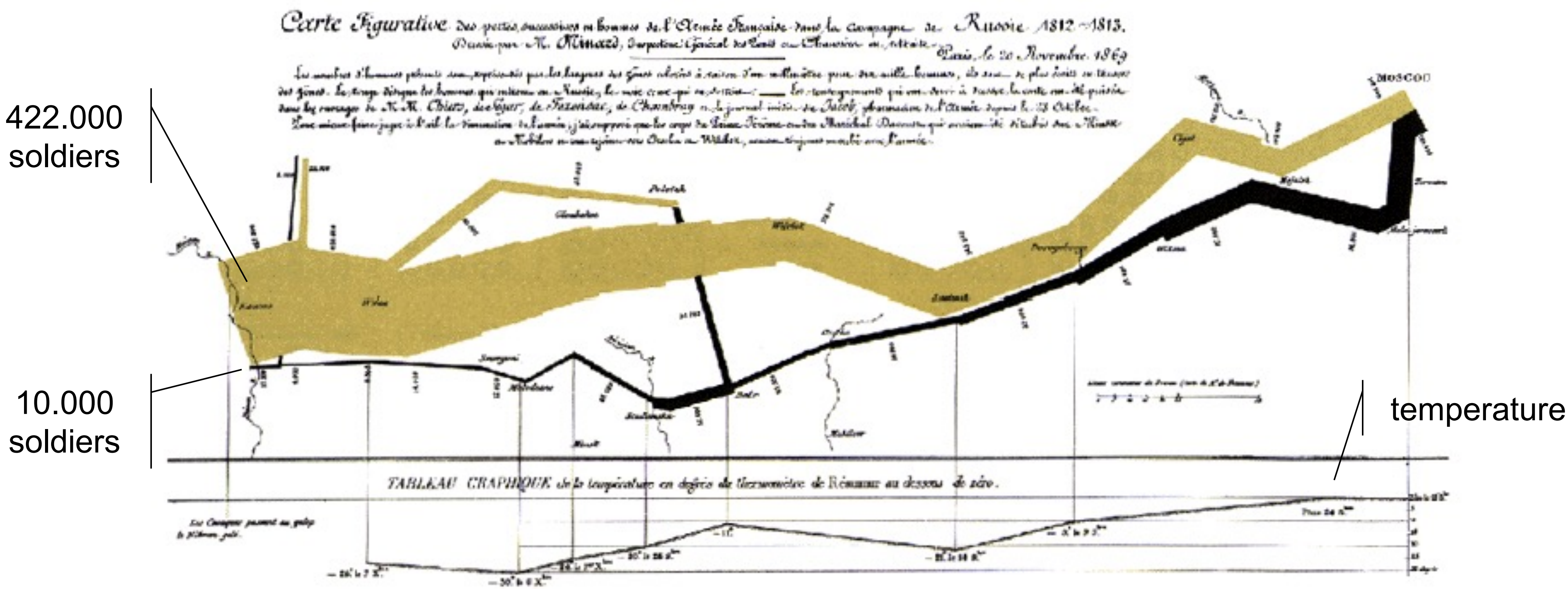
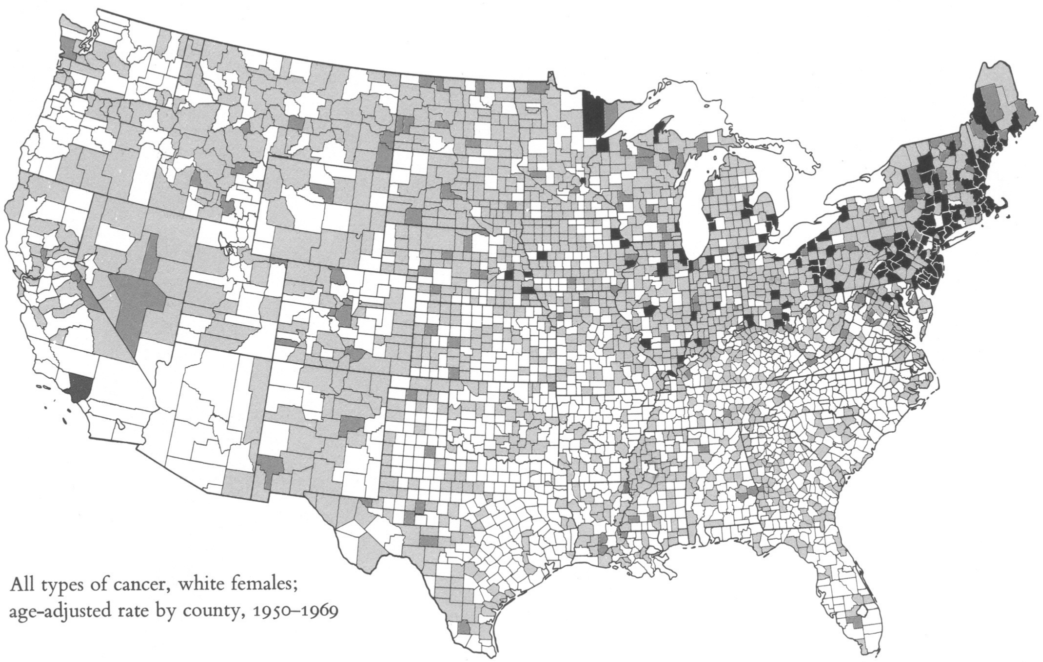


Bild 1. Joseph Minard: Karte des Rußlandfeldzuges von Napoleon von 1812-1813. [Spe:5f]

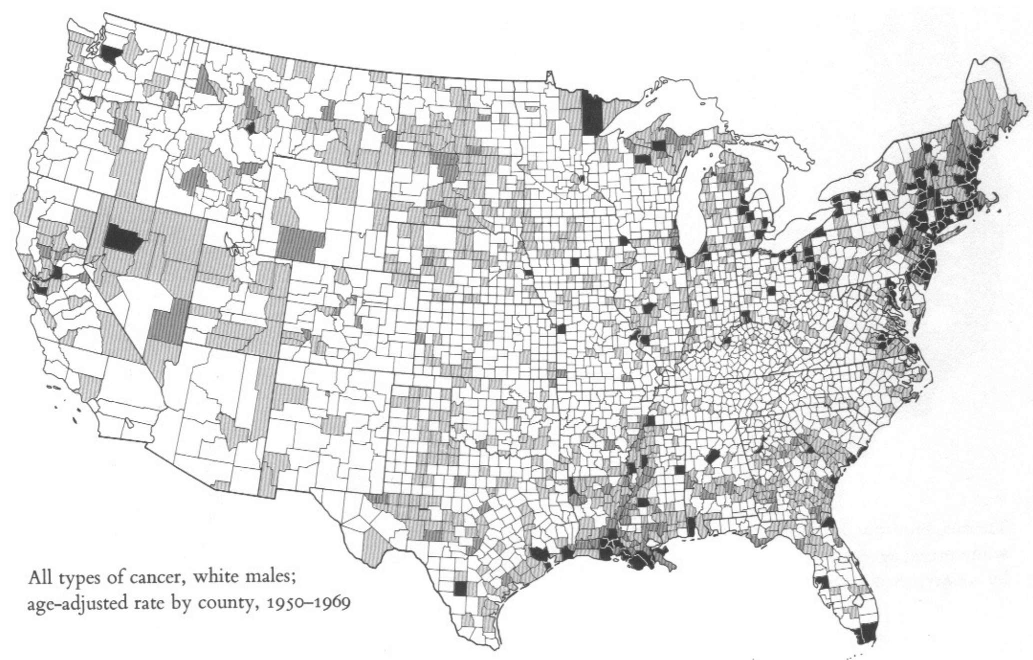
1.1.1 Classical Examples

Examinations of cancer in form of so-called data maps

- Each of the 3.056 values was positioned in 2D, and at least 4 values specify the county, i.e., around 21.000 values in one image!



All types of cancer, white females; age-adjusted rate by county, 1950–1969



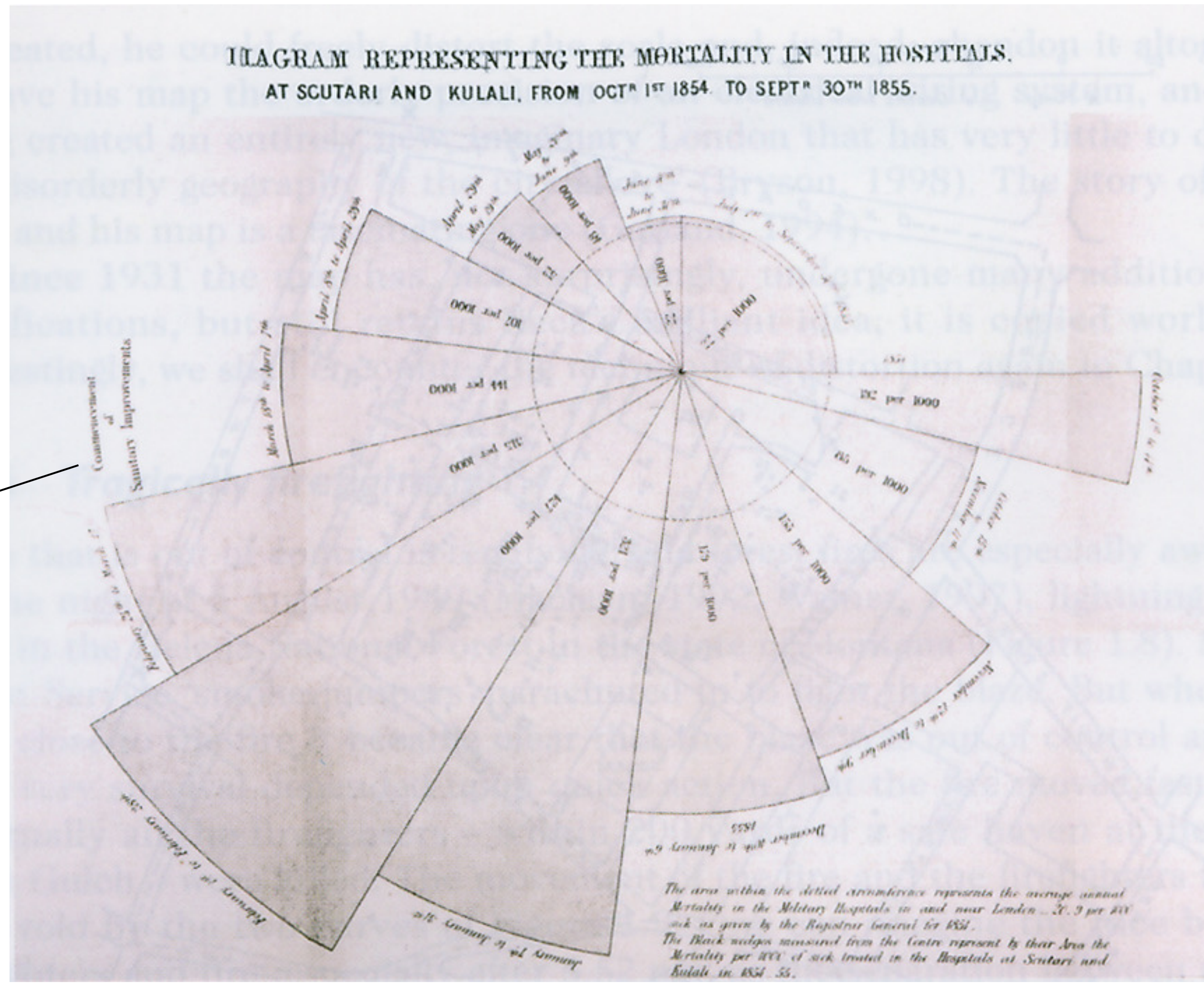
All types of cancer, white males; age-adjusted rate by county, 1950–1969

[Tuf1]

1.1.1 Classical Examples

- Death rate in hospitals (Nightingale‘s Diagram, 1858)

new regime



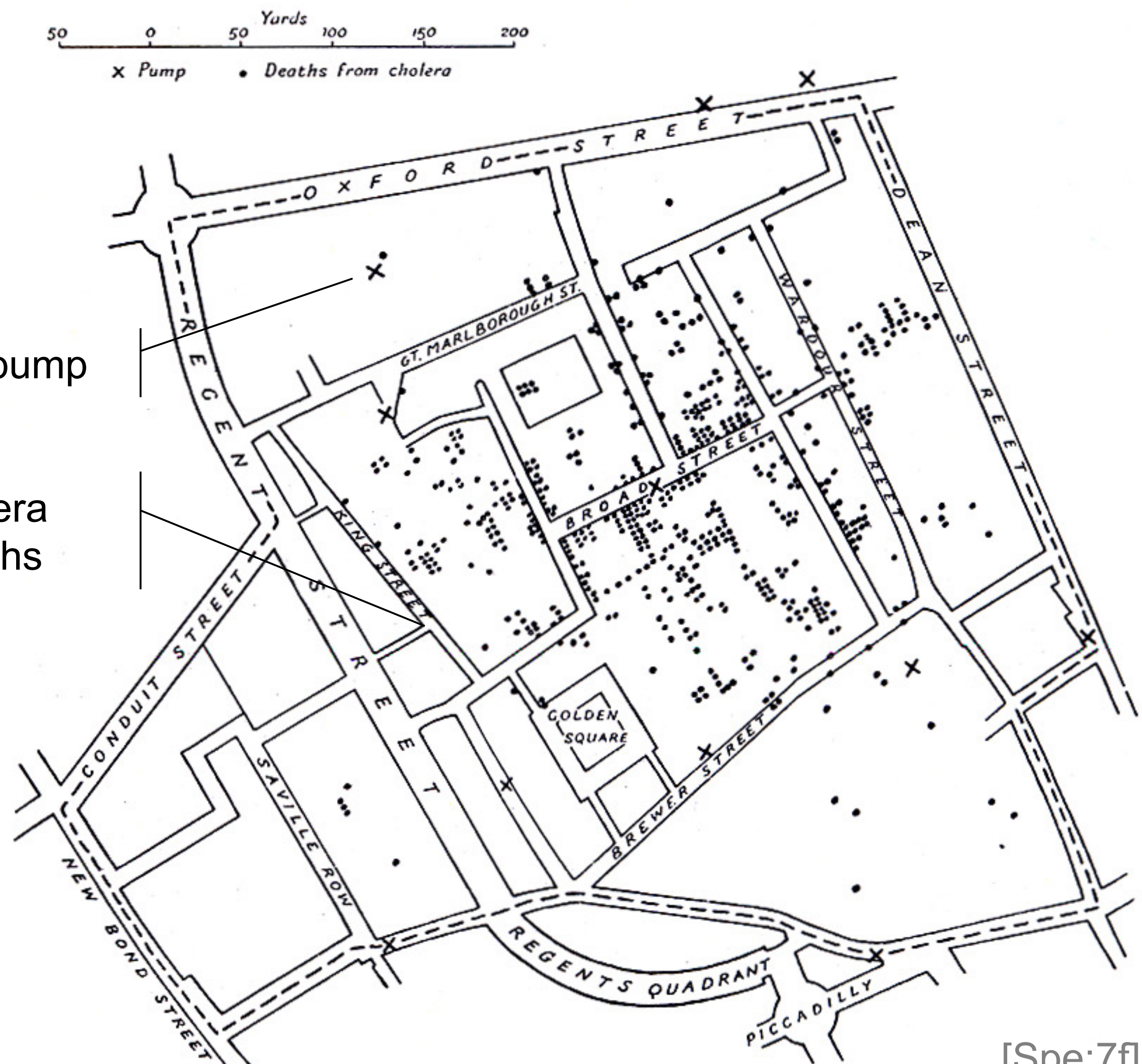
[Spe:7f]

1.1.1 Classical Examples

- Cholera epidemic, 1854, Dr. John Snow

water pump

cholera deaths

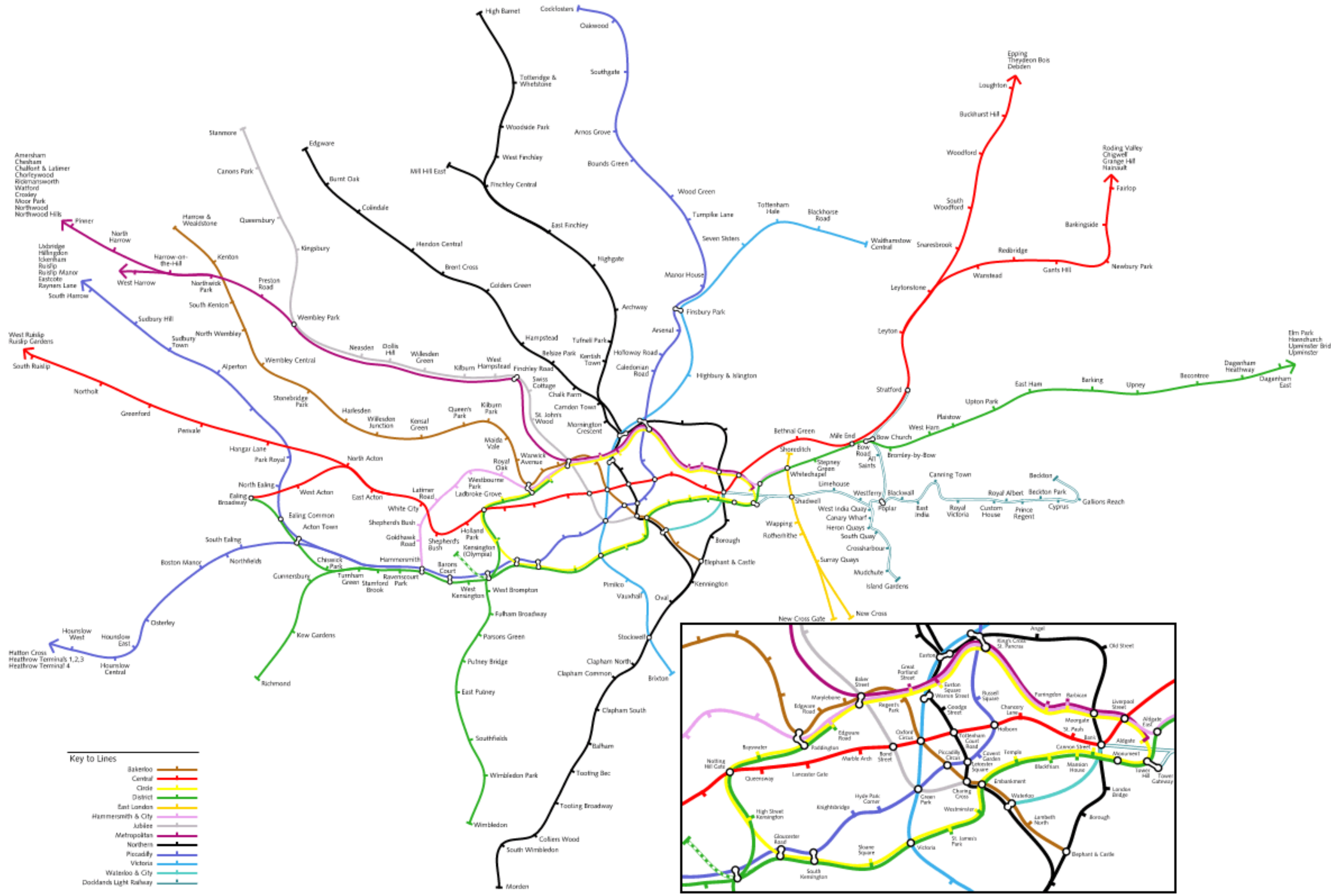


[Spe:7f]

1.1.1 Classical Examples

- The Tube – Geography

http://www.kottke.org/plus/misc/images/tubegeo.gif



1.1.1 Classical Examples

- The Tube – Beck‘s Map, 1931

http://content.tfl.gov.uk/walking-tube-map.pdf

[Spe:2ff]

1.2 Definitions

What is a *visualization*?

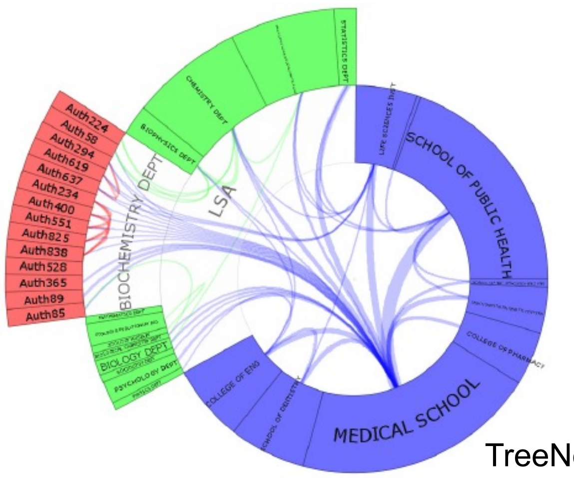
- Often defined as a process that produces a graphic, diagram, or an image
- In fact, it is a cognitive process that
  - produces a mental image or a mental model in the human brain
  - supports a better understanding/insight
- „The purpose of visualization is insight, not pictures“ [CMS: 6f]
  - Aims of this *insight*: discovery, decision making, and explanation

1.2.1 Visualization

- „The use of computer-supported, interactive, visual representations of data to amplify cognition“ [CMS:6]
  - *computer-supported*: new medium
  - *interactive*: direct manipulation
  - *visual representations*: use human perception
  - *data*: application specific
  - *amplify cognition*: helps people to think

1.2.2 Information Visualization

- „The use of computer-supported, interactive, visual representations of **abstract** data to amplify cognition“ [CMS:7]
  - *abstract:* no spatial correlation between the data, i.e., it exists no inherent physical space! The mapping into the physical worlds is in many cases **not** possible and often not useful
  - n-dimensional data
  - Examples
    - Sports statistics, stock trends, …
    - Configuration of cars or laptops, connections between criminals or terrorists, data on customer behavior, …

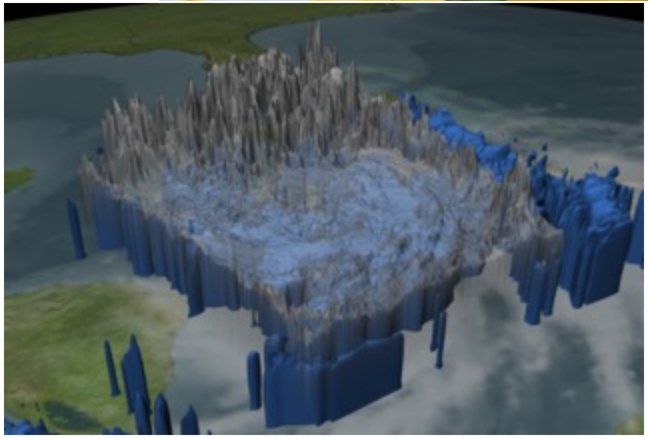


TreeNetViz

1.2.2 Information Visualization

- Information Visualization is **NOT** Scientific Visualization (SciVis)
  - SciVis uses spatial data **with** physical space
  - 2 or 3-dimensional data
  - Areas
    - Volume Visualization
    - Flow Visualization [nice 2D example: http://hint.fm/wind/]
  - Examples
    - Air flow over a wing
    - Weather in Sweden
    - Medical data from MRTs, …
  - In single cases, InfoVis and SciVis can overlap

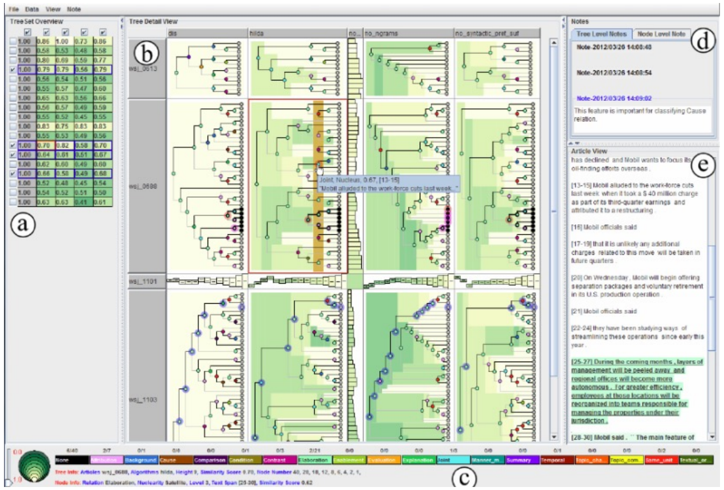
Hurricane Wilma



http://www.nasa.gov/

1.2.2 Information Visualization

- Application areas or data sources
  - Statistics of public or private sources
  - Links between Internet pages
  - Texts and document collections
  - Software (⇨ SoftVis)
  - Gene- and protein sequences
  - Biochemical networks
  - Train- or flight schedules
  - Weather-, climate- and astronomical data
  - …



Discourse Analysis

1.2.2 Information Visualization

- Typical Tasks
  - Search
    - Find a specific information
  - Browse
    - Look casually und search interesting things
  - Analyze
  - Monitor
  - Collect
  - Insight

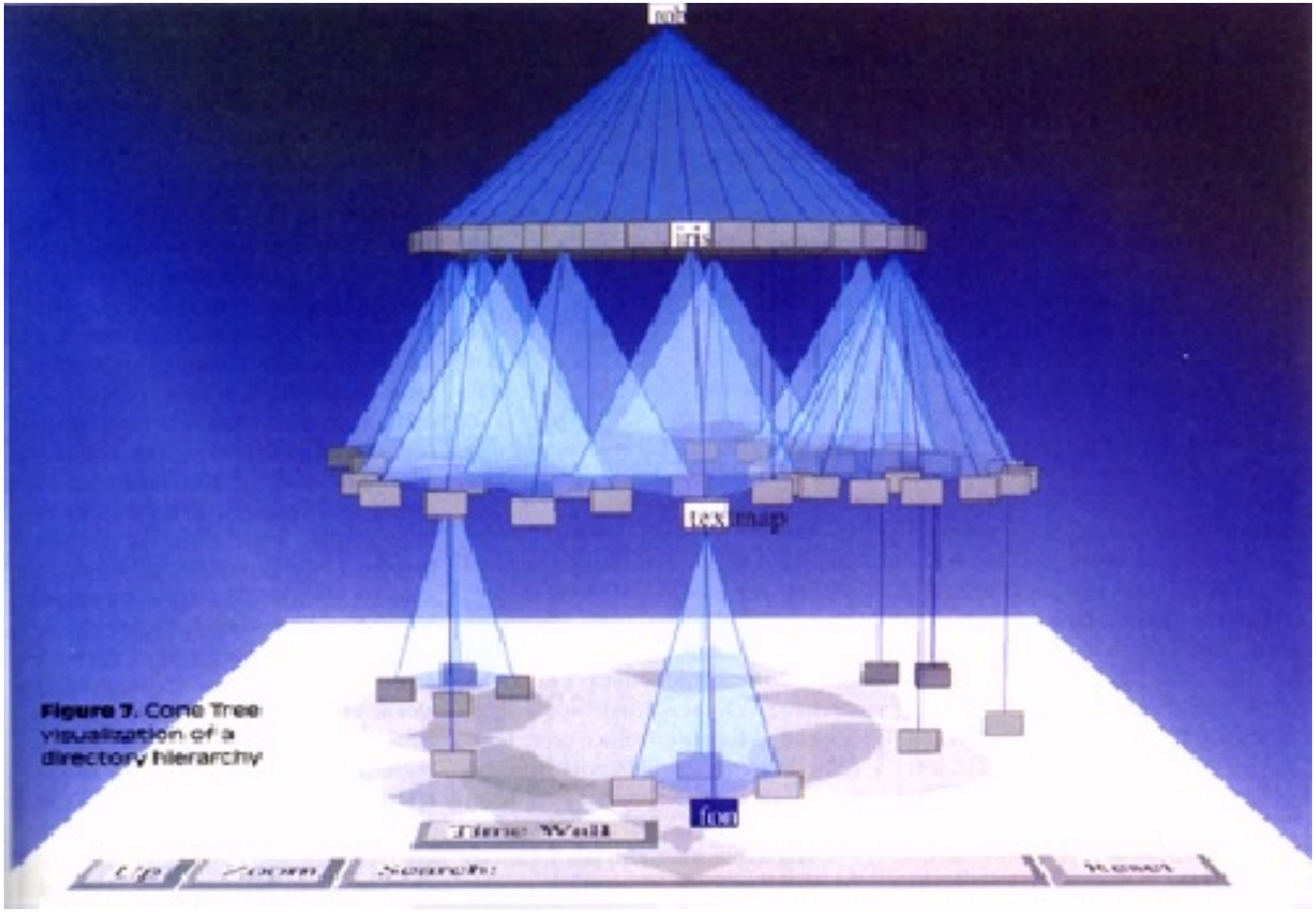
1.2.2 Information Visualization

- Benefit of InfoVis
  - Static and dynamic data can be made available for many people and groups
  - Interaction allows to build hypotheses with the help of „What if?“-questions and to verify them
  - Patterns and correlations can be discovered
  - Monitoring of a huge data set in real time and time pressure (e.g., Homeland-Security)
  - Handling of missing or wrong data
  - …

1.3 Examples

- Cone Tree

Visualization of a (hierarchical) directory structure. The cones can be rotated interactively. Thus, hidden elements in the visualization can be brought in the foreground.



[G. Robertson, J. Mackinley and S. Card. Cone trees: Animated 3D Visualizations of Hierarchical Information. In Proc. of the ACM SIGCHI '91 Conf. on Human Factors in Computing Systems, New Orleans, Louisiana, pp. 189—194, 1991]

1.3 Examples

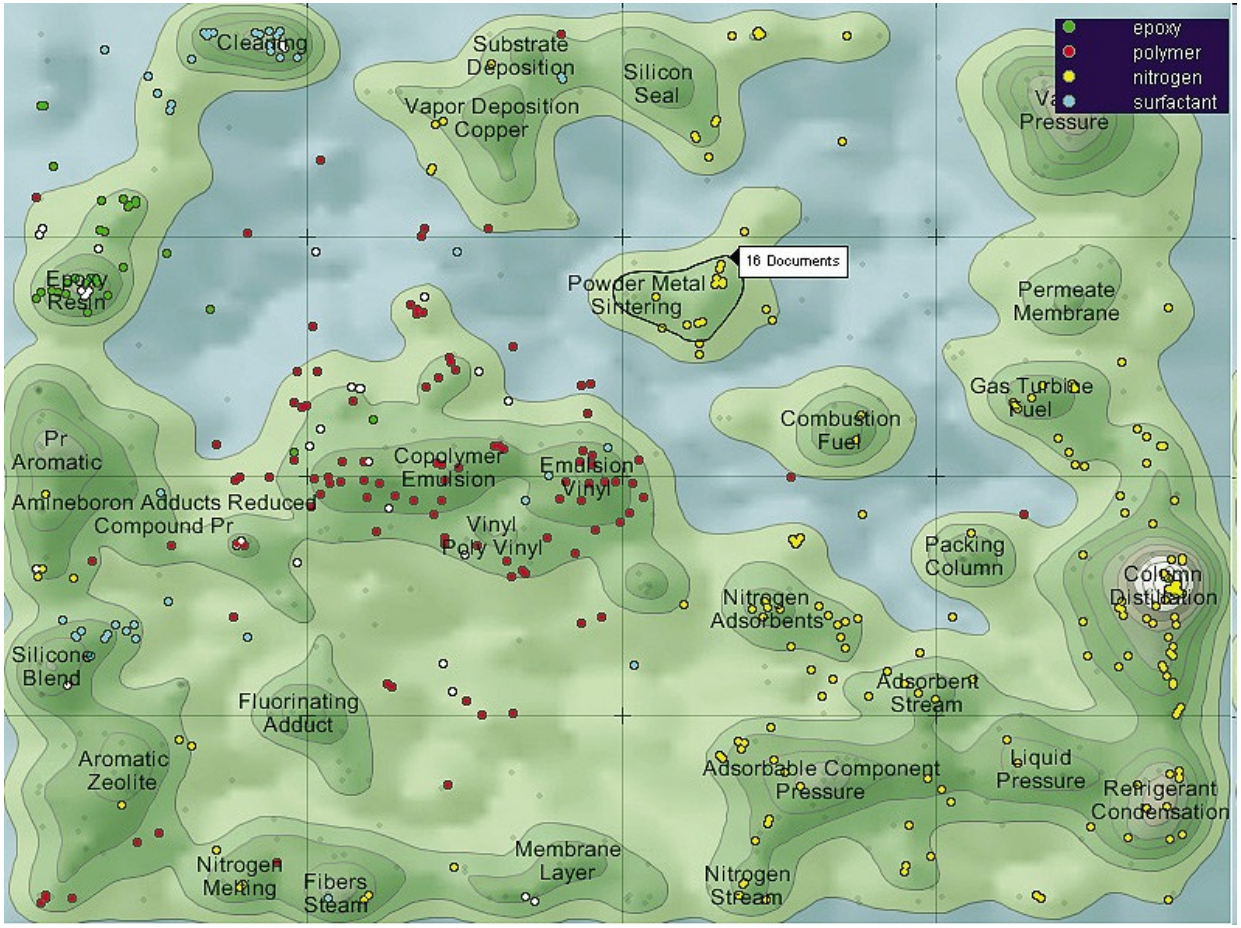
- Newsmap

Web-based visualization of actual news with the help of a *treemap*; classification is done according to categories and importance

1.3 Examples

- ThemeScape Map

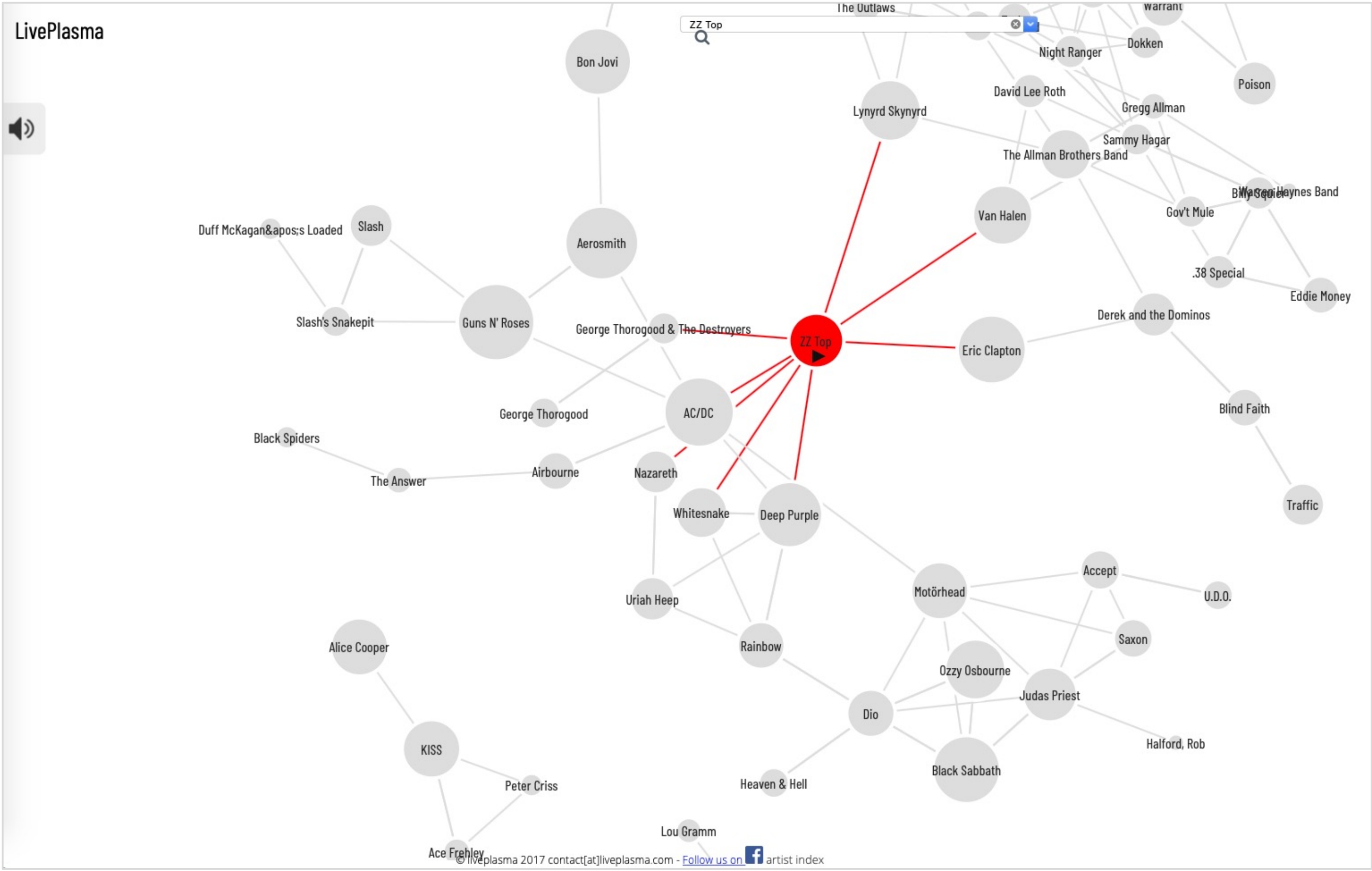
Individual patents are categorized and mapped into a map. Mountains represent patent concentrations within a thematical context; height lines connect similar conceptual domains.



1.3 Examples

- Liveplasma

Explore similar music and movies interactively!

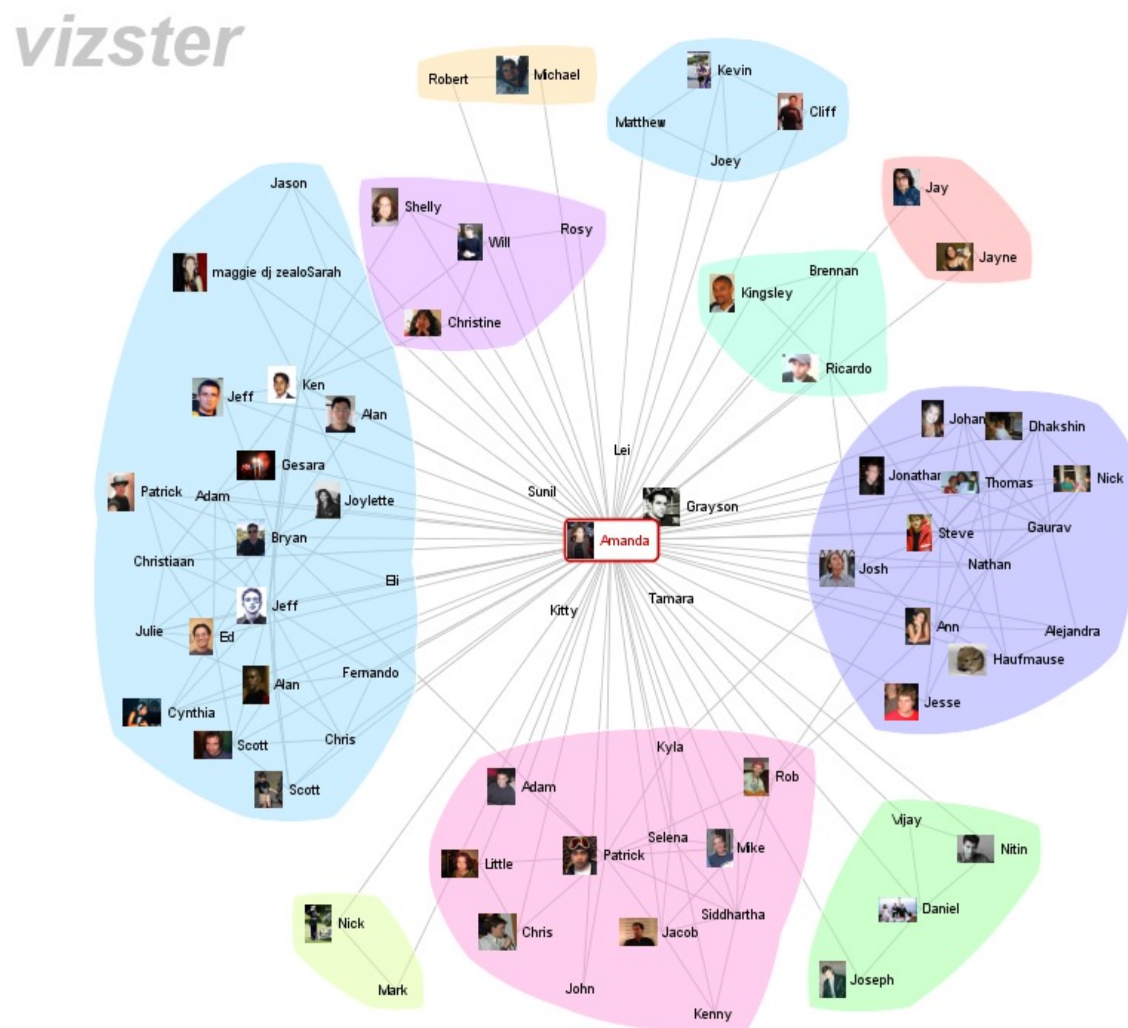


http://www.liveplasma.com/

1.3 Examples

- Networks

VIZSTER visualizes social networks. Many interaction possibilities allow to build cluster from tight relations, to show sub-networks etc.



https://homes.cs.washington.edu/~jheer//projects/vizster/