TECHNISCHE UNIVERSITÄT DRESDEN

Computergraphik und Visualisierung

INTERACTIVE MEDIA LAB DRESDEN

Faculty of Computer Science
Institute of Software and Multimedia Technology

1 Data Visualization

Introduction

Prof. Dr. Stefan Gumhold
Prof. Dr. Raimund Dachselt

Introduction
Lecture Outline

- History
- Foundations
- Examples
- Data
- Summary and Outlook

1 Introduction to Data Visualization

History

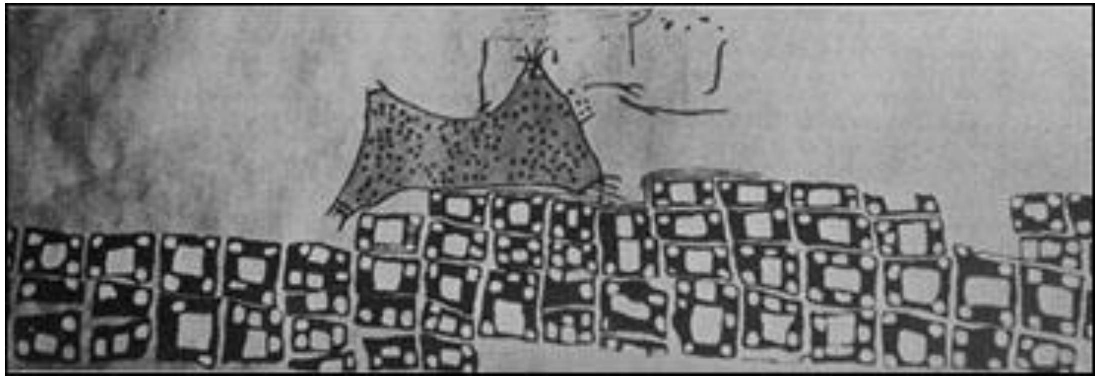
Visualization History
The Origins

Town Map of Çatalhöyük: 6200 BC

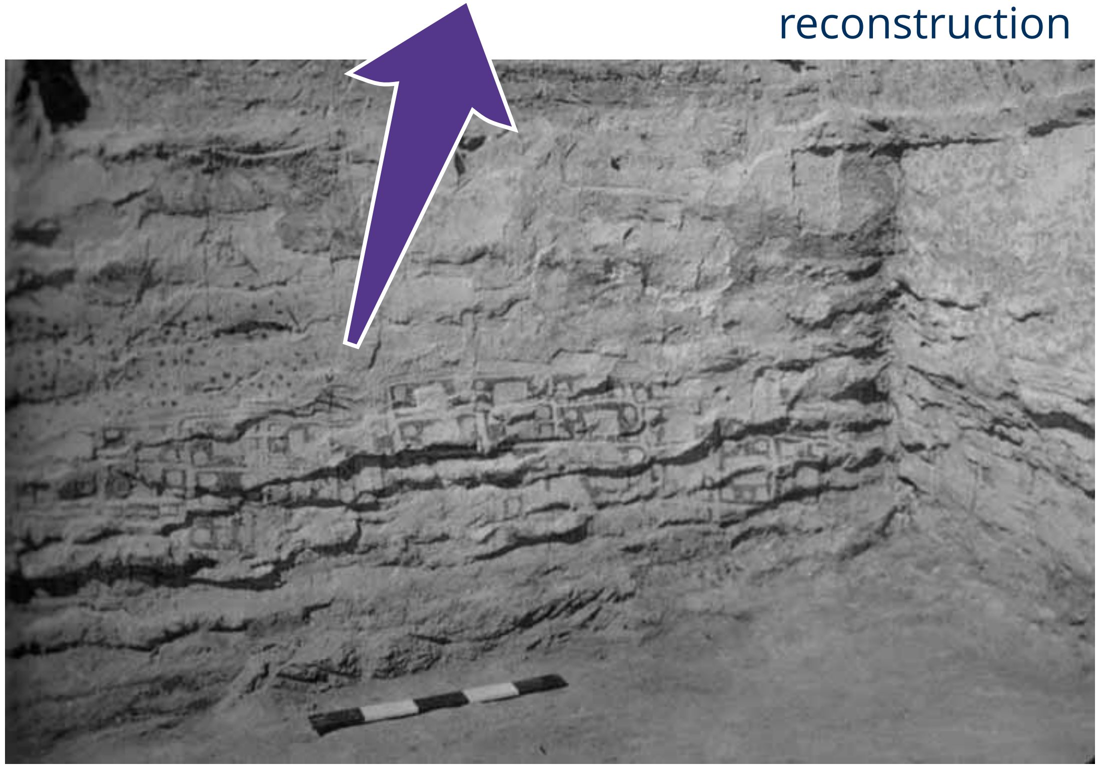
http://www.catalhoyuk.com/

- Not generally recognized as the oldest map.
- Found in 1963 in Anatolia at Çatal Höyük by James Mellaart.
- 9 foot long wall painting representing 80 houses possibly of Çatal Hyük itself.
- The C14 method determined the age to be 6200 BC.
- The settlement is a terraced city, with a volcano erupting in the background, probably the 10,672-foot Hasan Dag.

view onto Hasan Dag in google maps: link



reconstruction



excavation

Visualization History
The Origins

Nuzi Map 2200-2400 B.C. (now Yorgan Tepe, Irak)

Some information from Wikipedia:

“The Nuzi map is actually one of the so-called Gasur texts, and predates the invasion of the city of Gasur by the Hurrians, who renamed it Nuzi.“

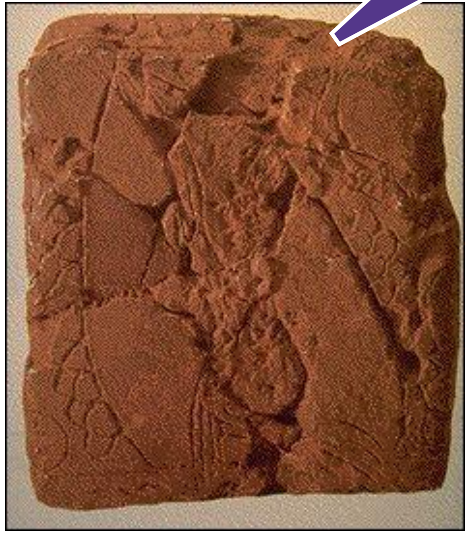
“The cache of economic and business documents among which the map was found date to the Old Akkadian period ~2360-2180 BC.”

“The tablet is approx. 6 x 6.5 cm”

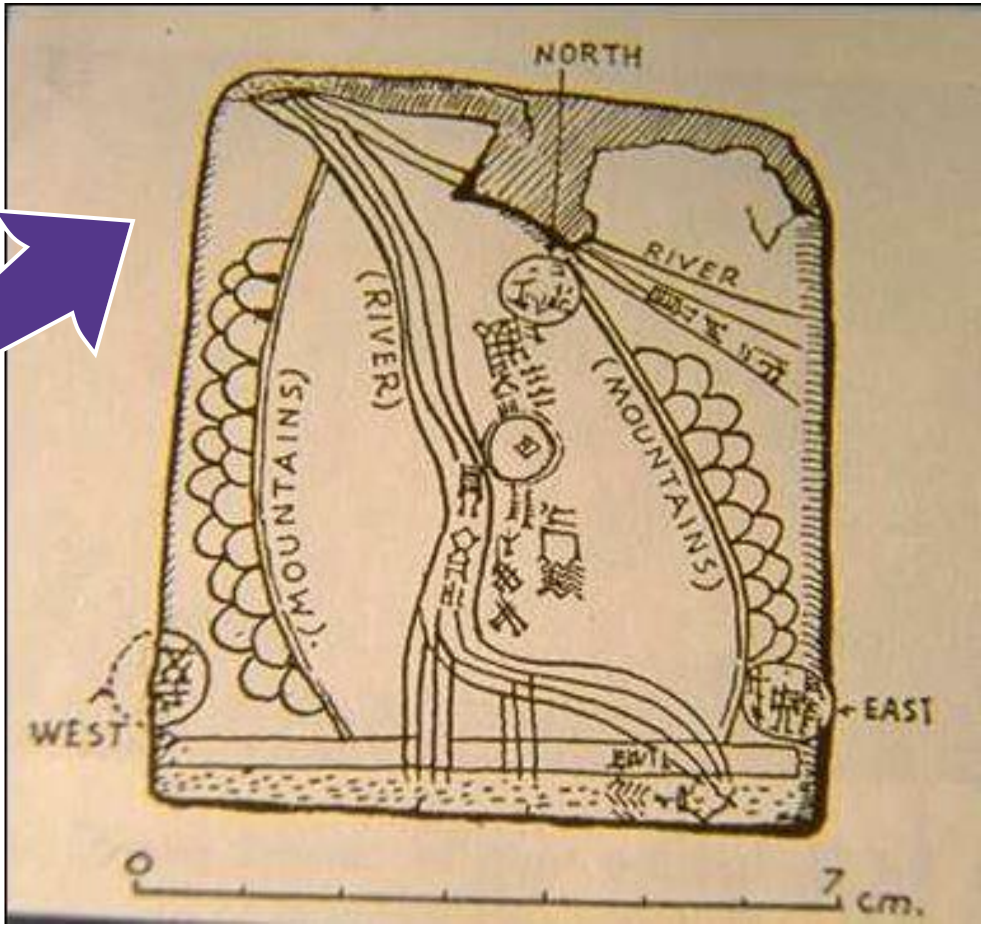
Content is not completely clear

Visual coding of mountains and rivers common in that time

Probably, map is upside down

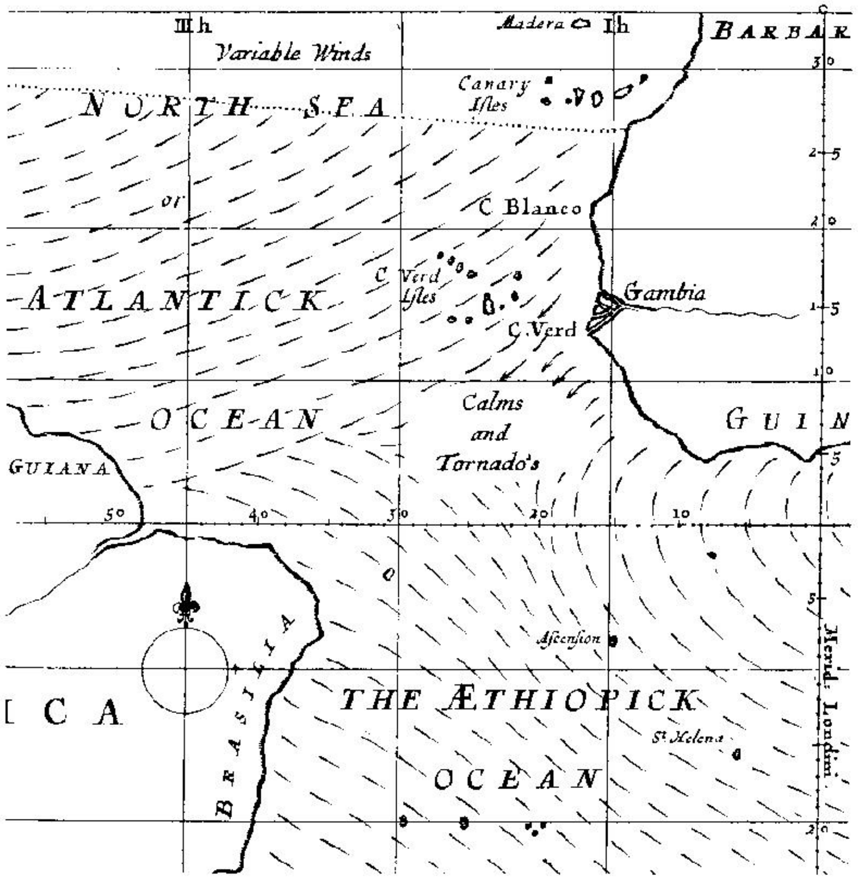


excavation

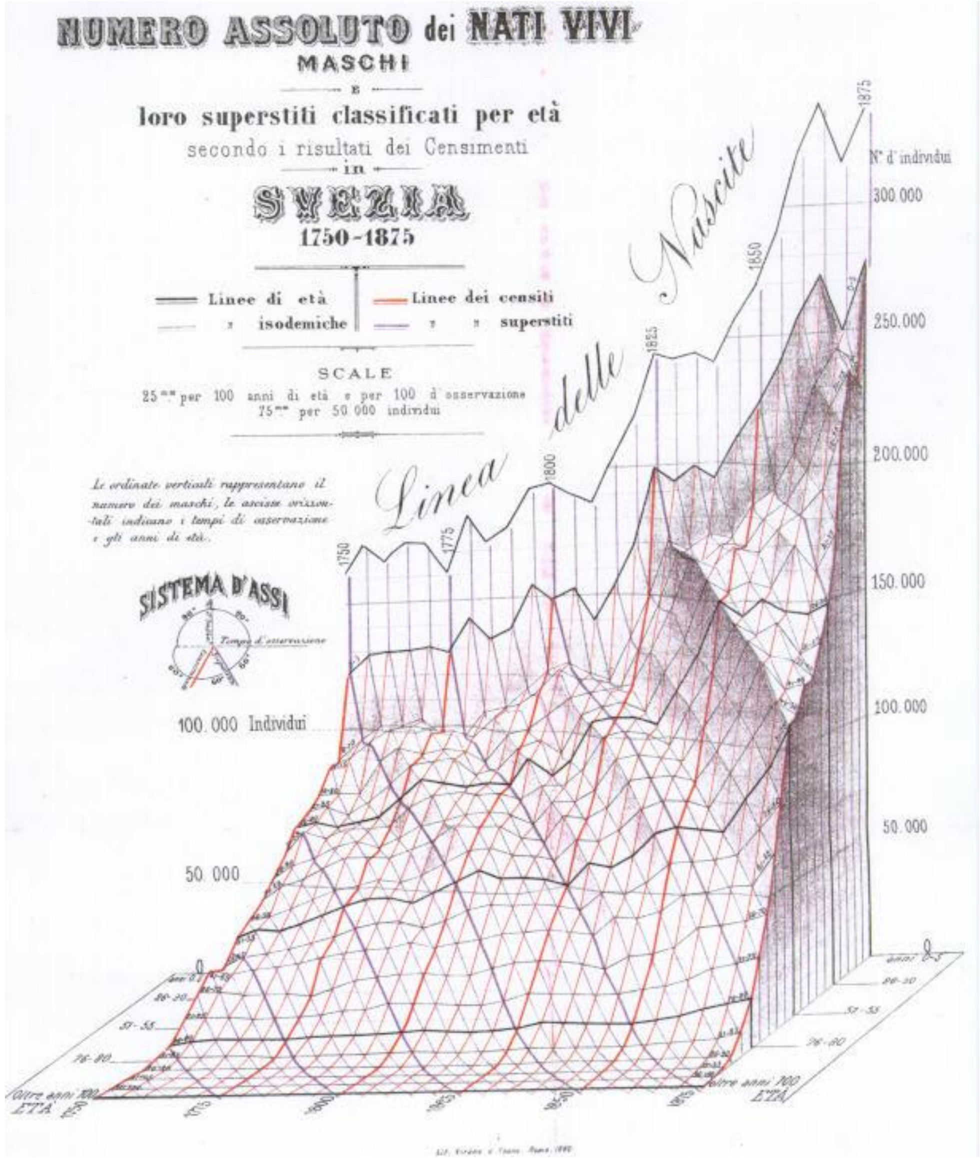


reconstruction

Visualization History
The Origins

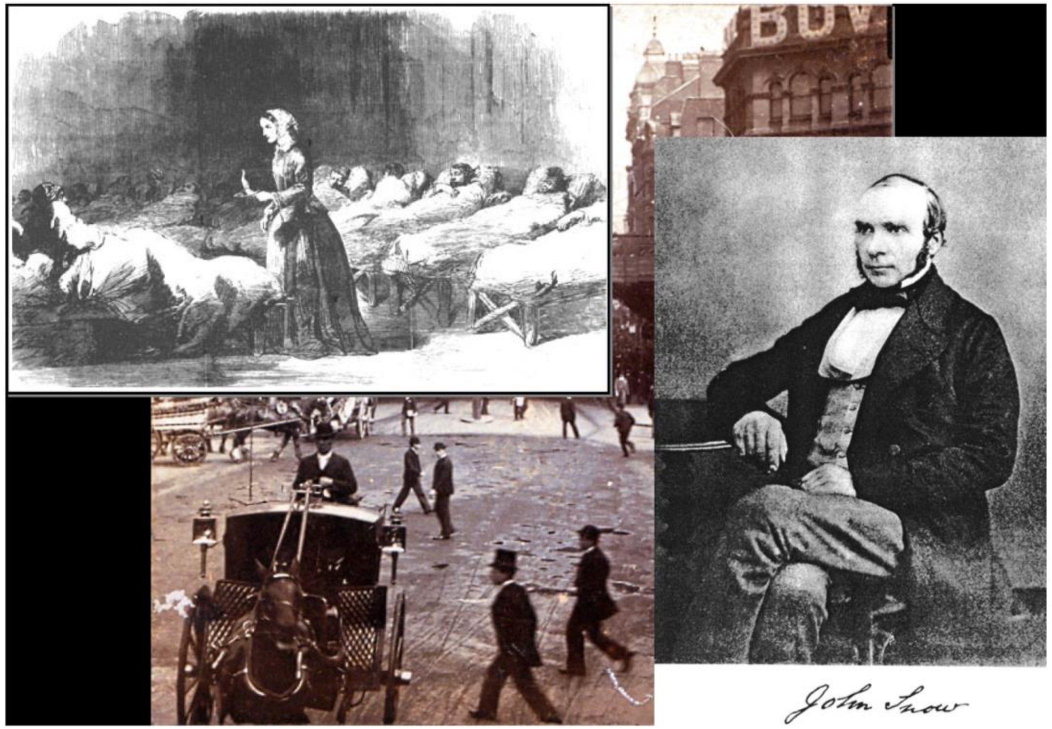


Edmond Halley's Trade Wind Visualization (1686)
Arrow plots along Streamlines show wind variation in Atlantic Ocean around equator (large map here)



Census data visualization by Italian Luigi Perozzo (1879)
Height field visualizes population in Sweden over year (left to right) and age (back to front)

Visualization History
Question: What caused a cholera epidemic in London in 1854?



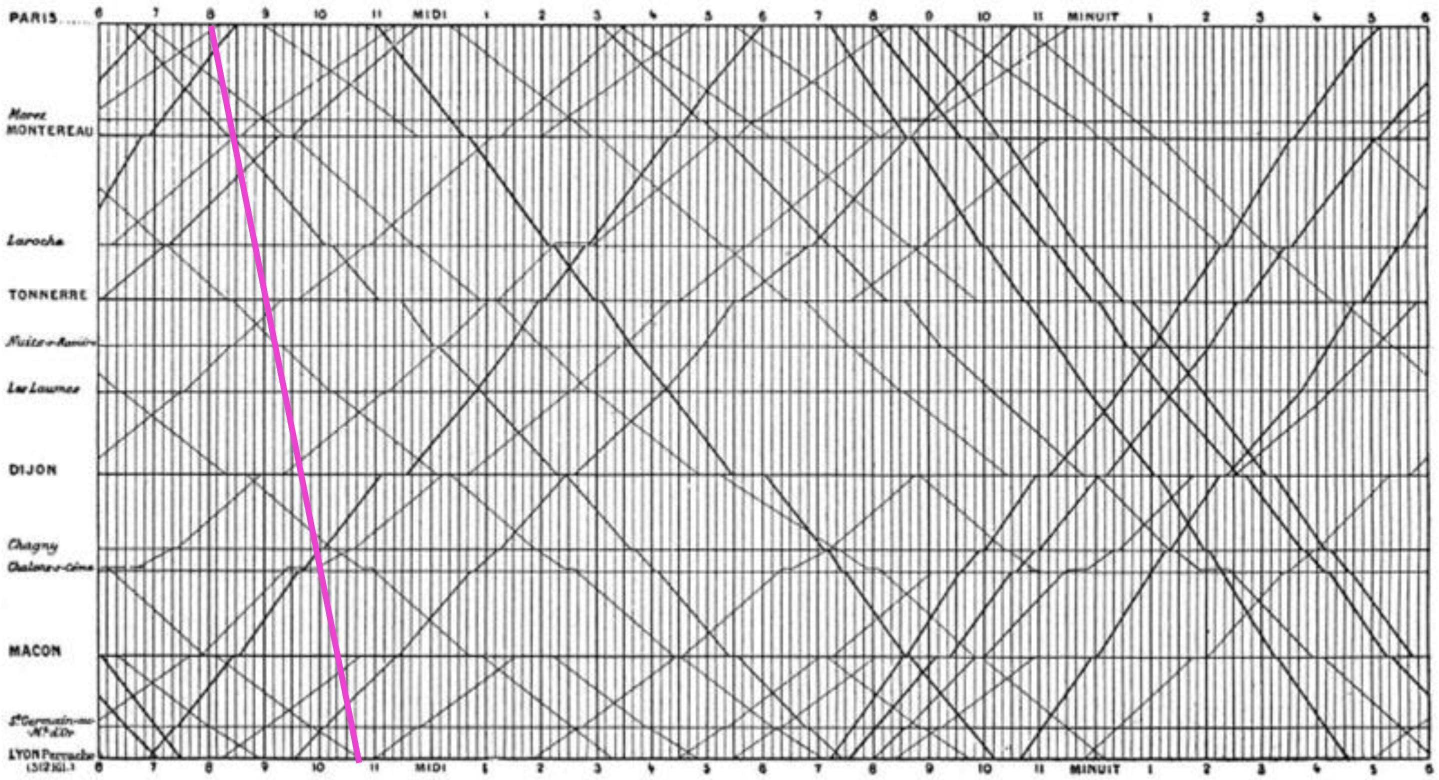
[51°30'48"N 0°08'12"W]

Solution found with Visualization:
John Snow, around 1854
Map of London's Soho district
Conclusion: epidemic was caused by a "bad water pump"

[Tufte97]

Visualization History

Timetable Visualization by Etienne-Jules Marey in the 1880s



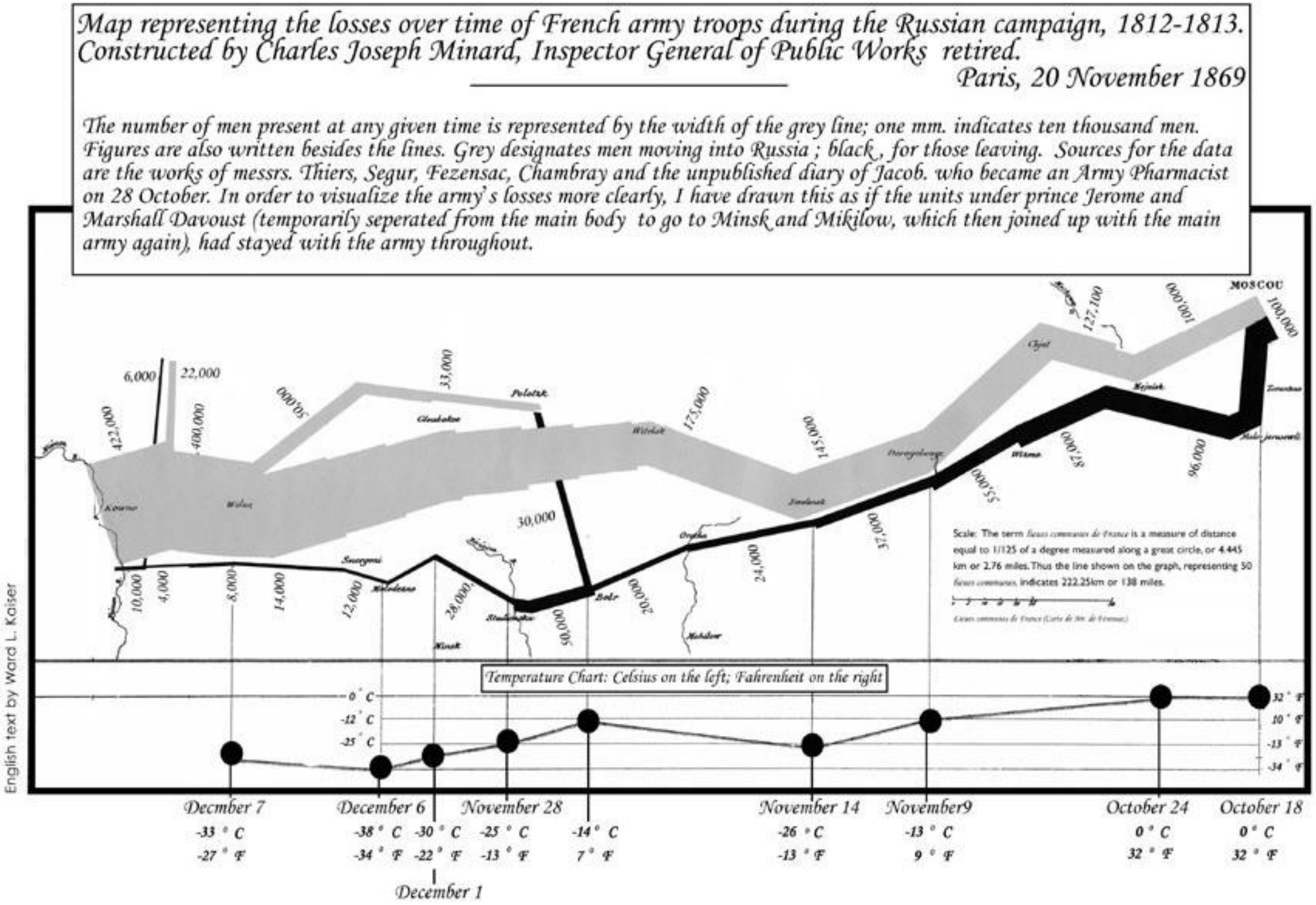
2D diagram: Train schedule Paris-Lyon

- Daytime in hours on X-axis with vertical grid lines at 10min intervals
- Travelled distance on Y-axis with one horizontal line per station
- Individual trains are represented by diagonal lines from top left to bottom right (and vice versa for Lyon-Paris)
- Slope corresponds to travel speed (the steeper, the faster)
- Stops yield horizontal sections

Benefits of visualization

- Line density represents frequency
- Shows complex information at a glance, while accurate detail analysis is also possible

Visualization History
Historic space-time vis – Map of Napoleon‘s Russian Campaign



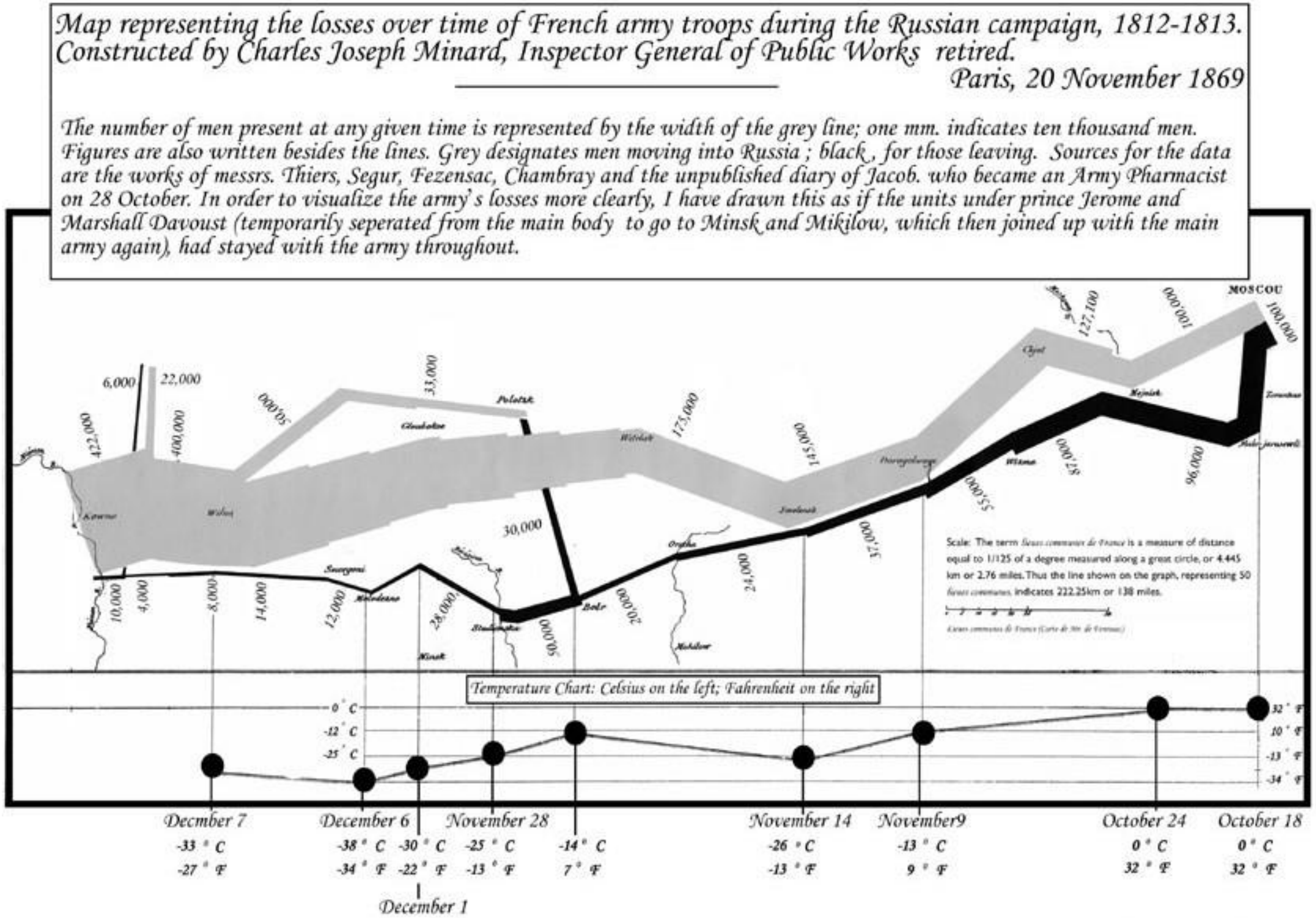
Visualization History
Historic space-time vis – Map of Napoleon‘s Russian Campaign

- Beginning left at the Polish-Russian border;
- Width of the band symbolizes troop strength at any time
- The path of Napoleon's retreat from Moscow is symbolized by the darker lower band linked to the temperature and data display in the lower part of the diagram
- Single events + entire story highlighted at the same time
  - For example, the crossing of the Berezina river was a catastrophy
- The army finally arrived in Poland with only 10,000 soldiers
- Minard's account tells a rich, coherent story using multivariate data
  - Much brighter than e.g. just a value that would change over time
- Illustrated variables:
  - Time
  - Size of the army
  - Your position on the two-dimensional symbolic map
  - The direction of movement of the army
  - Temperatures at different times during the retreat

Visualization History
Historic space-time vis – Map of Napoleon‘s Russian Campaign

What does the map show us?

- Forces visual comparisons
- Shows causality
- Captures multivariate complexity
- Integrates text and graphic into a coherent whole
- Illustrates high quality content
- Place comparisons adjacent to each other, not sequentially
- Use the smallest effective differences



2 Introduction
Terms and Foundations

Terms and Foundations
Definitions of the term „[Data] Visualization”

- **Oxford dictionary:** *“visualization: 1. the act of forming a picture of somebody/something in your mind, 2. the act of making something able to be seen by the eye.”*
- **Descartes 1637:** *“Imagination or visualization, and in particular the use of diagrams, has a crucial part to play in scientific research.”*
- **McCormick ’87:** *“Transformation of the symbolic into the geometric”*
- **Bertin ’83:** *“… finding the artificial memory that best supports our natural means of perception.”*
- **Hearst ’03:** *The depiction of information using spatial or graphical representations, to facilitate comparison, pattern recognition, change detection, and other cognitive skills by making use of the visual system.”*
- **Foley and Ribarsky ’94:** *„… binding of data to representations that can be perceived. The types of bindings could be visual, auditory, tactile, etc., or a combination of these.“*
- **Card ’99:** *“The use of computer-supported, interactive, visual representations of data to amplify cognition.”*

Terms and Foundations
Visualization Disciplines

Data Visualization is a discipline of computer science, which is heavily used in mechanical engineering, chemistry, physics, and life sciences

> Diehl 07: *"Visualization is more than a method of computing. Visualization is the process of transforming information into a visual form, enabling users to observe the information. The resulting visual display enables the scientist or engineer to perceive visually features which are hidden in the data but nevertheless are needed for data exploration and analysis."*

Scientific Visualization (SciVis)

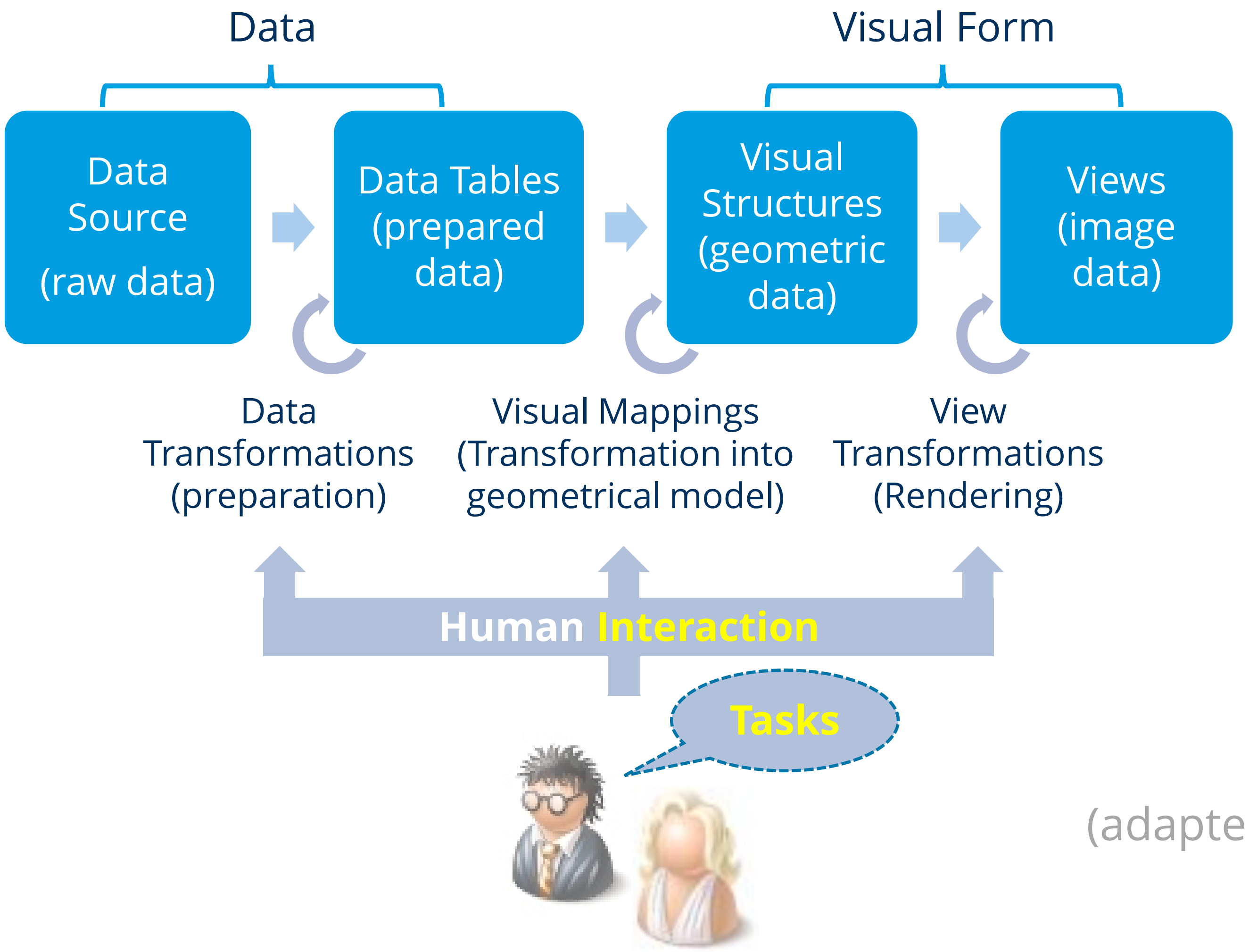
- Physical data, e.g. medical, volume or flow visualization
- Spatialization is given (i.e. inherently spatial)
- SciVis is informative, too

Information Visualization (InfoVis)

- Abstract data, e.g. hierarchy or software metrics visualization
- Spatialization is chosen
- InfoVis is scientific, too

Eyes have by far highest Perceptual Bandwidth!

Terms and Foundations
Visualization Pipeline / Reference Model for Data Visualization



(adapted from [Card et al. 99])

Terms and Foundations
Visualization Pipeline – Overview

Visualization Pipeline - Important Steps

- Collecting Data
  - Various data sources (observed, sensors, system data ...)
  - Methods for extracting and collecting relevant data from these sources are as diverse as the data sources themselves
- Transforming Data
  - Data volume can not be presented to the user all at once
  - Data is reduced by filtering, analysis, statistical methods
  - Focus on important parts
- Visualizing Data
  - Resulting data is mapped onto a visual model
  - Translated into geometric and graphical information
  - Displayed on screen or other medium as single image or image series
- Gaining insights and knowledge
  - Through interactive manipulation

Terms and Foundations
Visualization Goals and Phases

Three Phases of the Visualization Process

1. Exploration Phase
   - Visualize all data
   - Search for hypotheses
2. Validation Phase
   - Use special visualization techniques to validate hypotheses
3. Presentation Phase
   - Perspicuous presentation of results to other stakeholders

Three Goals of Visualization

- Explore/Calculate
  - Analyze
  - Reason about Information
- Communicate
  - Explain
  - Make Decisions
  - Reason about Information
- Control
  - Interactive control over application, simulation or cyber-physical system
  - Gain knowledge for decision making as soon as possible

Another Purpose of DataVis: Decorate

Terms and Foundations
Visualization Modes

Static mode (pictures, info graphics)

Animation mode (design animation and create video offline)

Interactive Visual Analysis

- support interaction with preprocessing, mapping and rendering parameters of visualization
- demands for real-time rendering

In-Situ Visualization

- stream to be visualized data directly from simulation or sensor network
- demands for efficient data processing & streaming

Interactive Steering

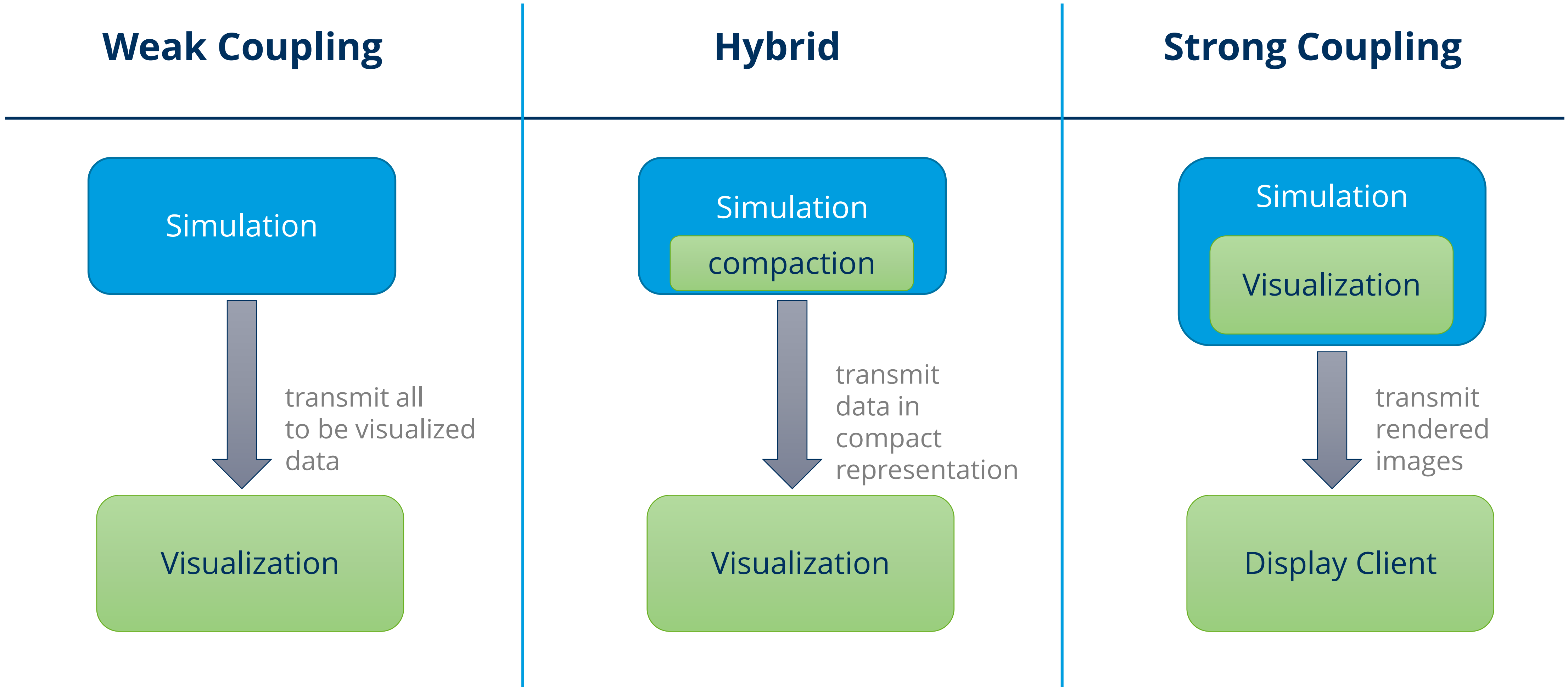
- provide interaction controls to change simulation parameters/boundary conditions or to control entities like vehicles, drones, robots, etc. in simulated environments

Visual Analytics

- combines automated analysis techniques with interactive visualizations for an effective understanding, reasoning and decision making on the basis of very large and complex data sets.

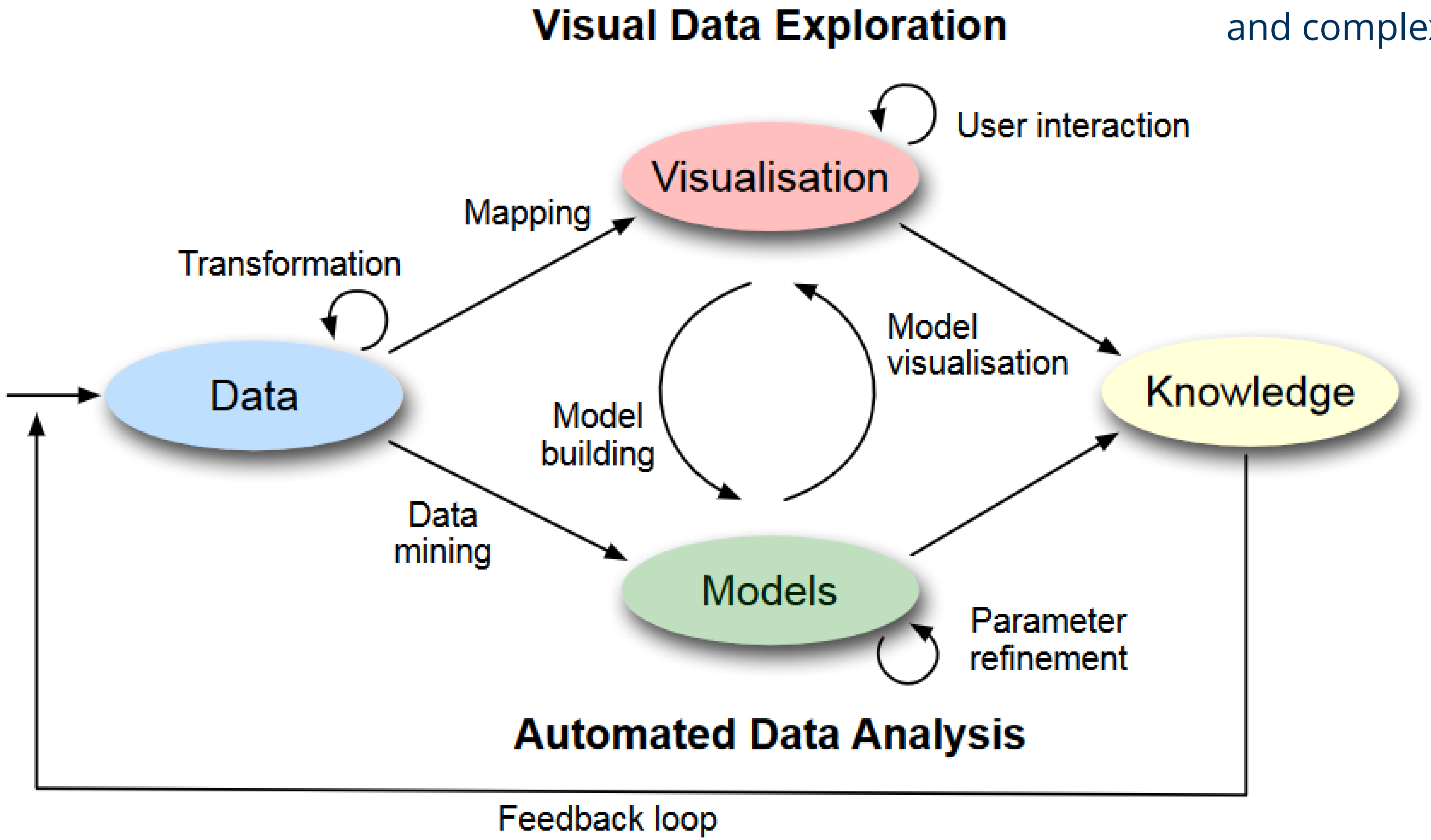
In Situ Visualization – crowd simulation example

In Situ Visualization



Terms and Foundations
Visual Analytics

Visual Analytics combines automated analysis techniques with interactive visualizations for an effective understanding, reasoning and decision making on the basis of very large and complex data sets.



Webpage and Book: https://www.visual-analytics.eu/

Terms and Foundations
Navigation and Interaction

Shneidermans visual information search mantra: [Shne05]

- Overview first, zoom and filter, then details on demand.

Seven General Interaction Tasks by Shneiderman

1. **Overview:** Gaining an overview of the entire information space, recognizing global patterns
2. **Zoom:** Zooming in on interesting information objects, viewing a smaller subset
3. **Filter:** Filtering out uninteresting data objects, selecting a subset based on v. attributes
4. **Details-on-demand:** Selection of a data object or a group of objects? To get details, display of attributes of data objects after selection
5. **Relate:** Consideration of relationships between data objects, comparison of values
6. **History:** Recording the actions to undo them (undo)
7. **Extract:** Extraction of subsets of the information space and request parameters

Interaction Task Categories
- **Select**
- **Explore**
- **Reconfigure**
- **Encode**
- **Abstract/Elaborate**
- **Filter**
- **Connect**

(inspired by Yi et al. 2007)

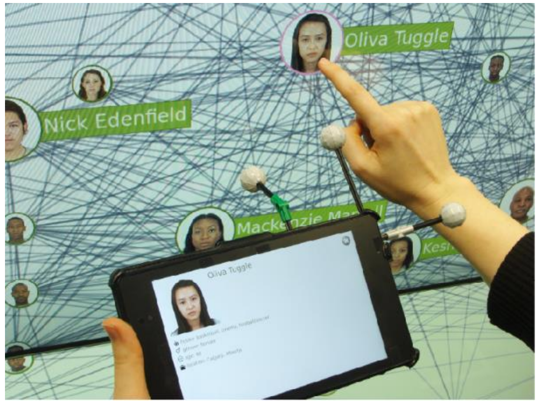
Terms and Foundations
Navigation and Interaction

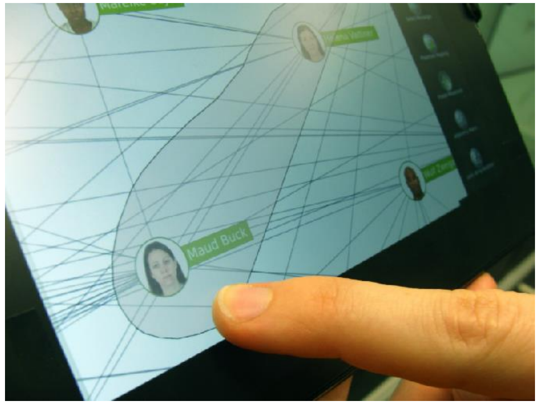
Basic Interaction techniques

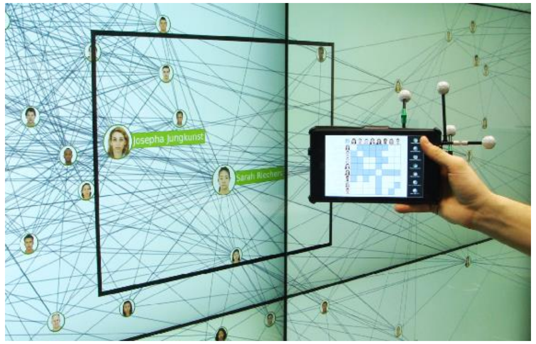
- Selection
  - Mouseover / hover / tooltip
  - Mouseclick, touch, special controller
  - Lasso / Drag, pointing, raycasting etc.
- Re-Arrange
  - move, sort, delete

Advanced Interaction Techniques

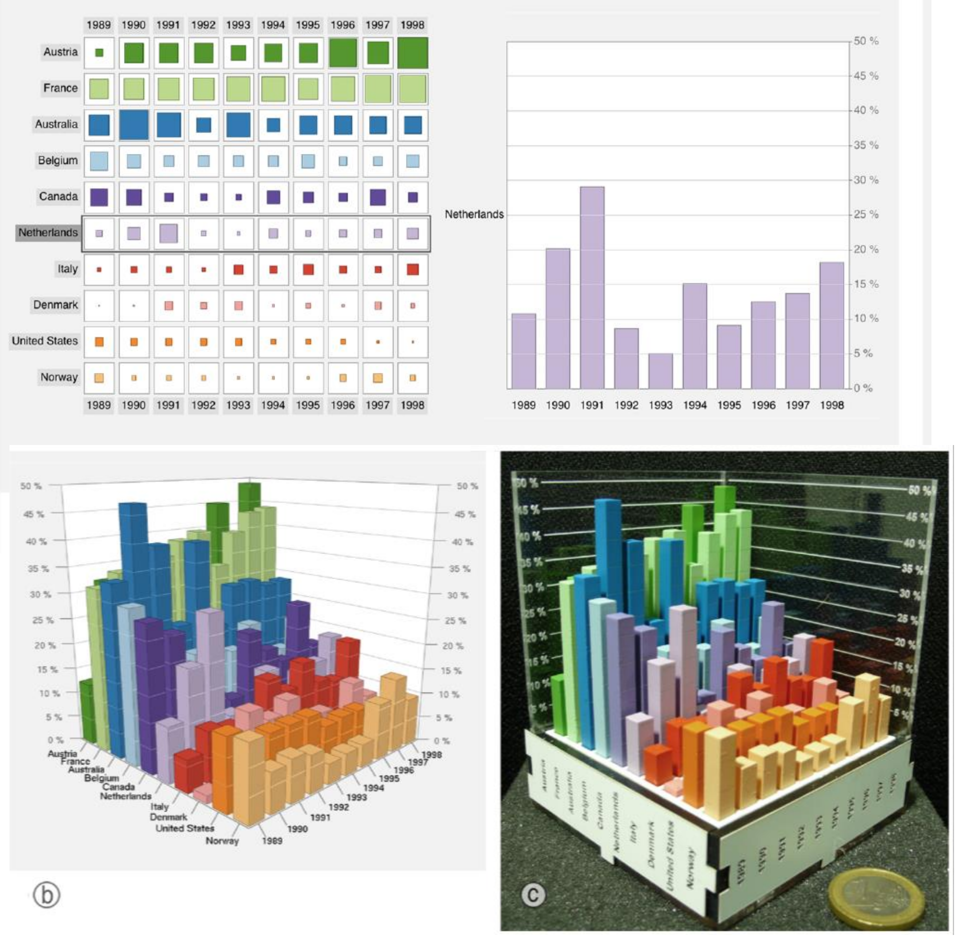
- Filtering
- Overview + Detail
- Zooming and Panning
- Focus + Context
- Brushing and Linking, Probing
- 3D Navigation, slicing etc.







Terms and Foundations
Focus on Task



Dimensions of information visualizations

- Tasks – why is the visualization needed?
- Target audience – who will use this visualization?
- Source – which is the data source to represent?
- Representation – how is it represented (encoded)?
- Medium – how is the visualization presented?

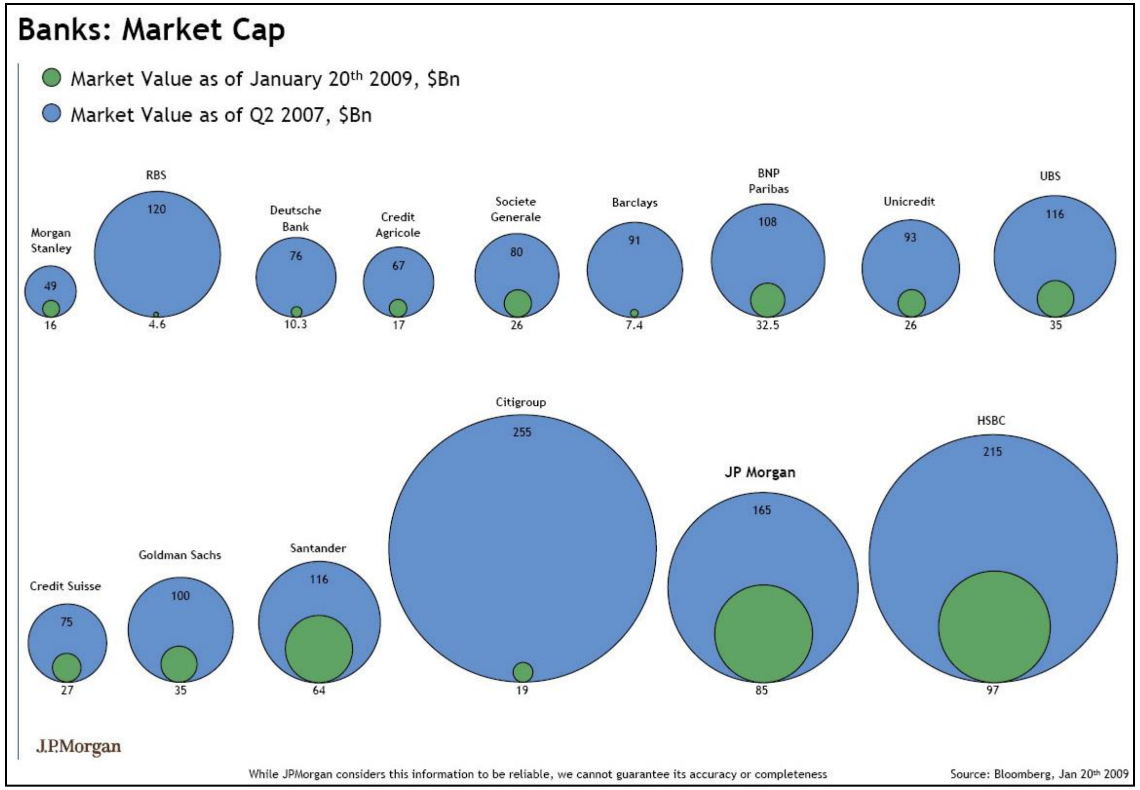
Focus on Tasks

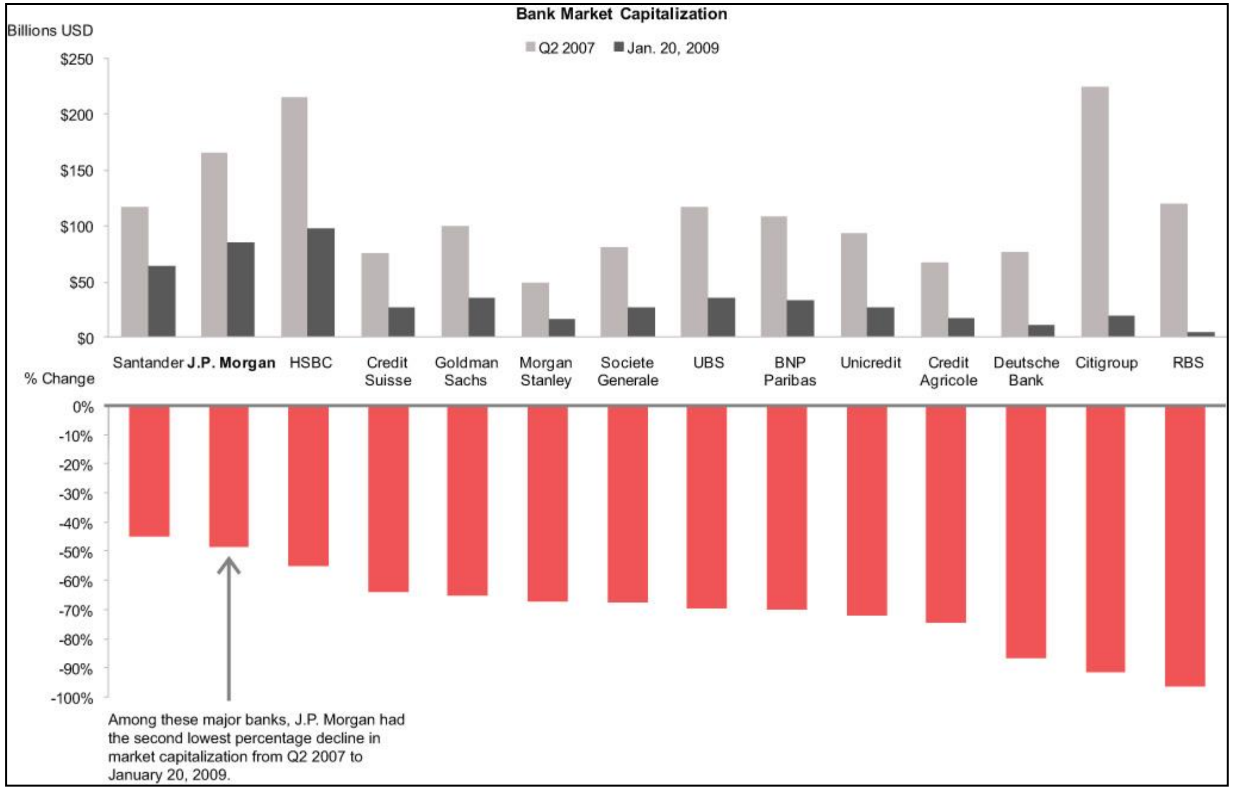
- Is that what a visualization is used for?
- Is that what the developers of a tool had in mind when they designed it?
- Each visualization emphasizes certain information and neglects others
- Suitable visualizations are always task-dependent - different tasks require different information

3 Introduction

Examples

Watch out for Pitfalls or even Vislies!

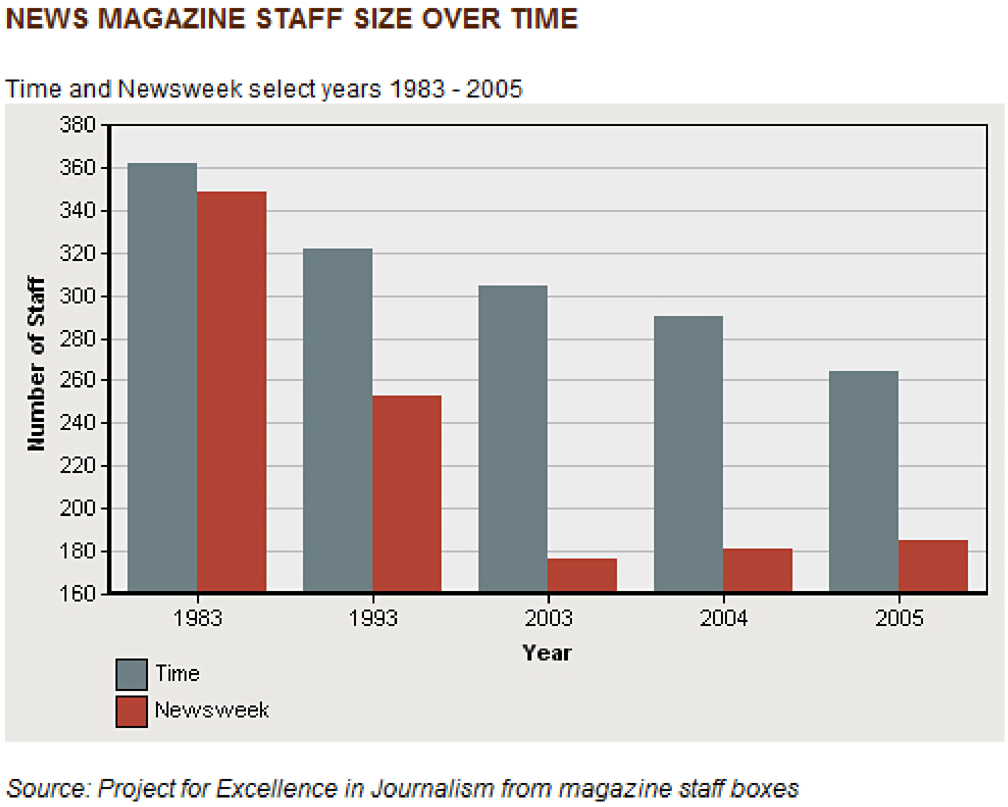


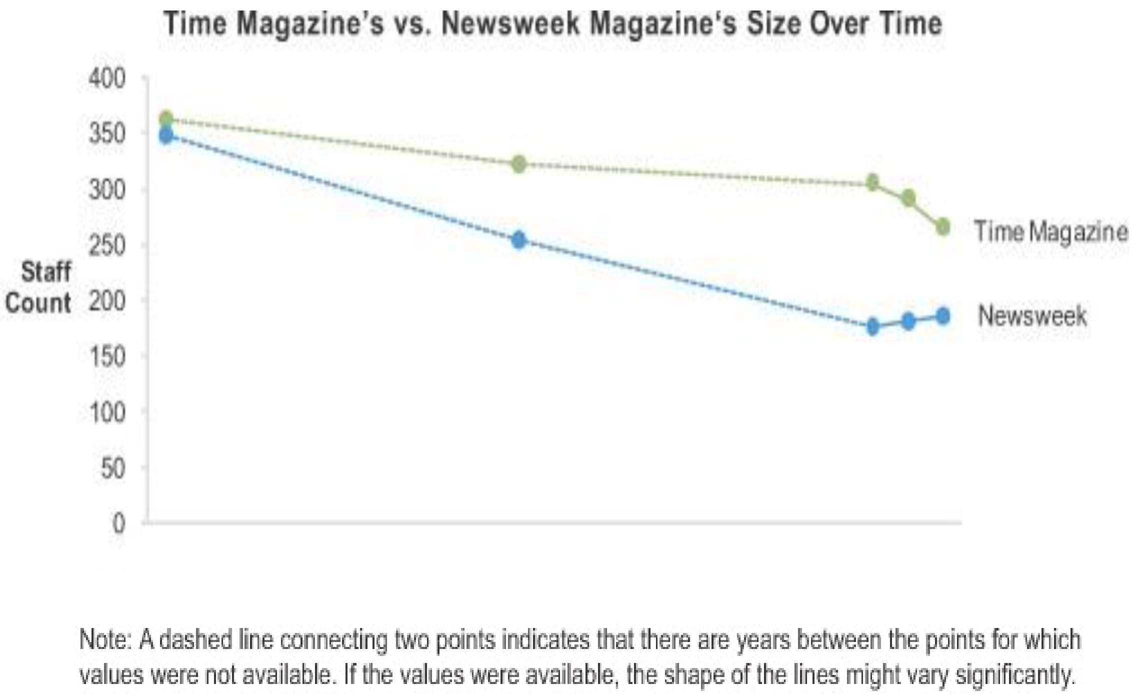


Bad: linear scaling of radius does not correspond to linear scaling of perceived size

Better: use bar charts as length is perceived linearly and additionally shows percentual change

Watch out for Pitfalls or even Vislies!





Bad: non-zero baseline, different time intervals between data points and missing continuity

Better: use zero baseline and line chart to show continuity and time intervals correctly

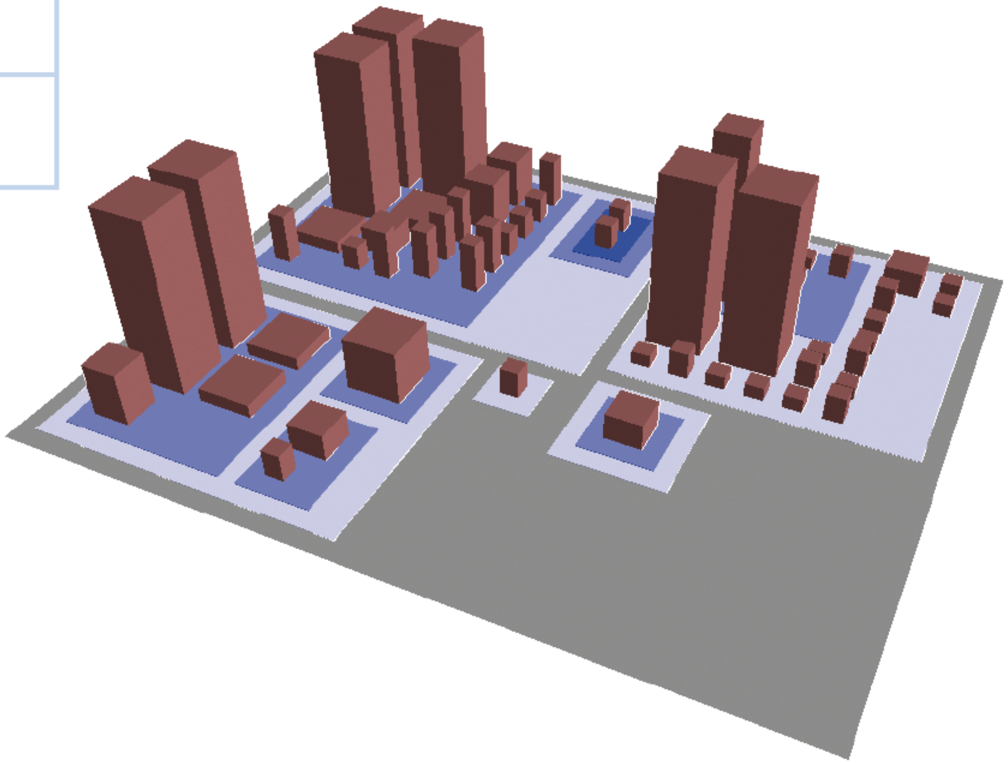
InfoVis Example

Software Visualization – Code City

software	representation
classes	buildings
packages	districts
system	city

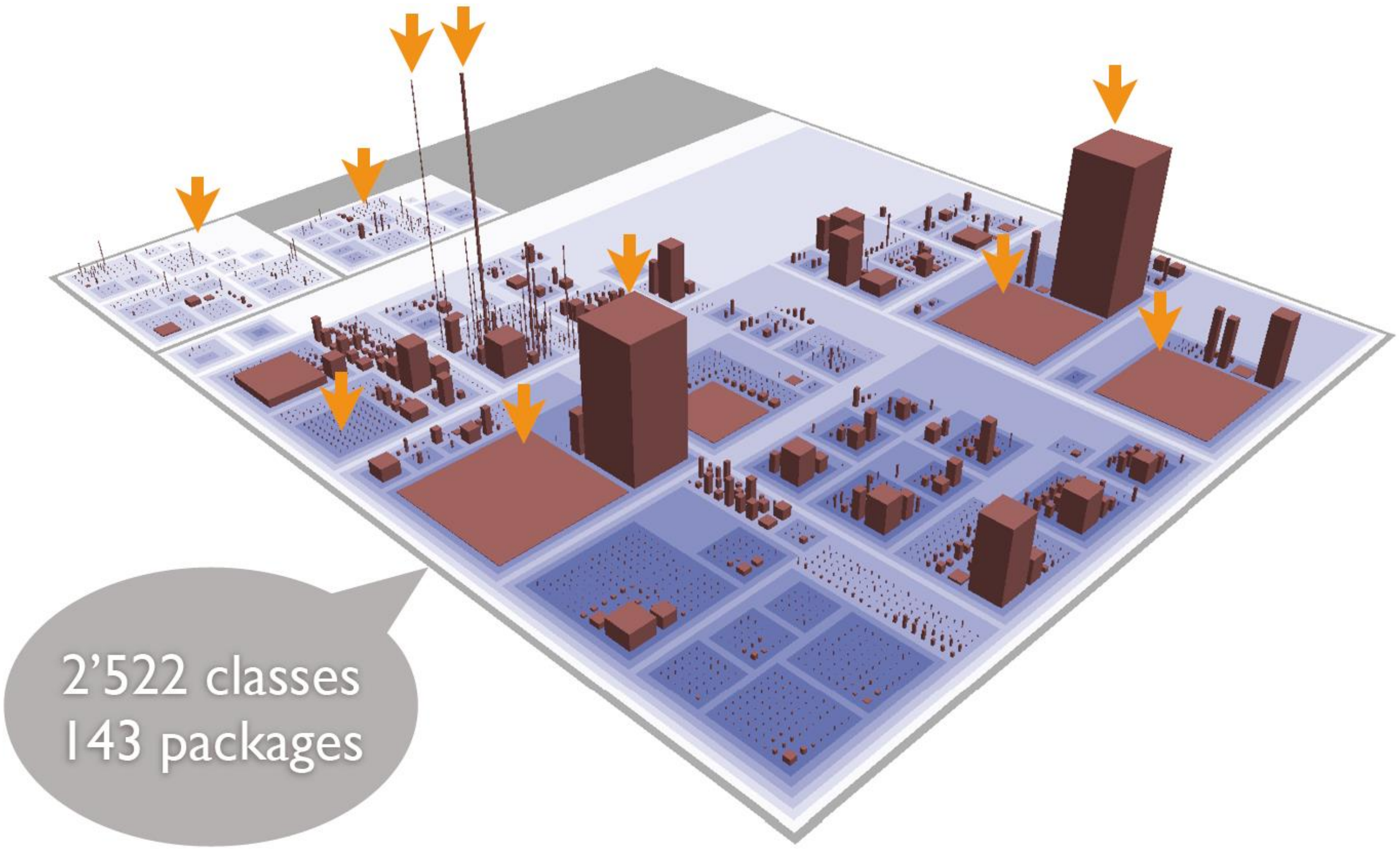
class metric	building property
number of methods	height
number of attributes	width
number of attributes	length

package metric	district property
nesting level	color saturation



InfoVis Example

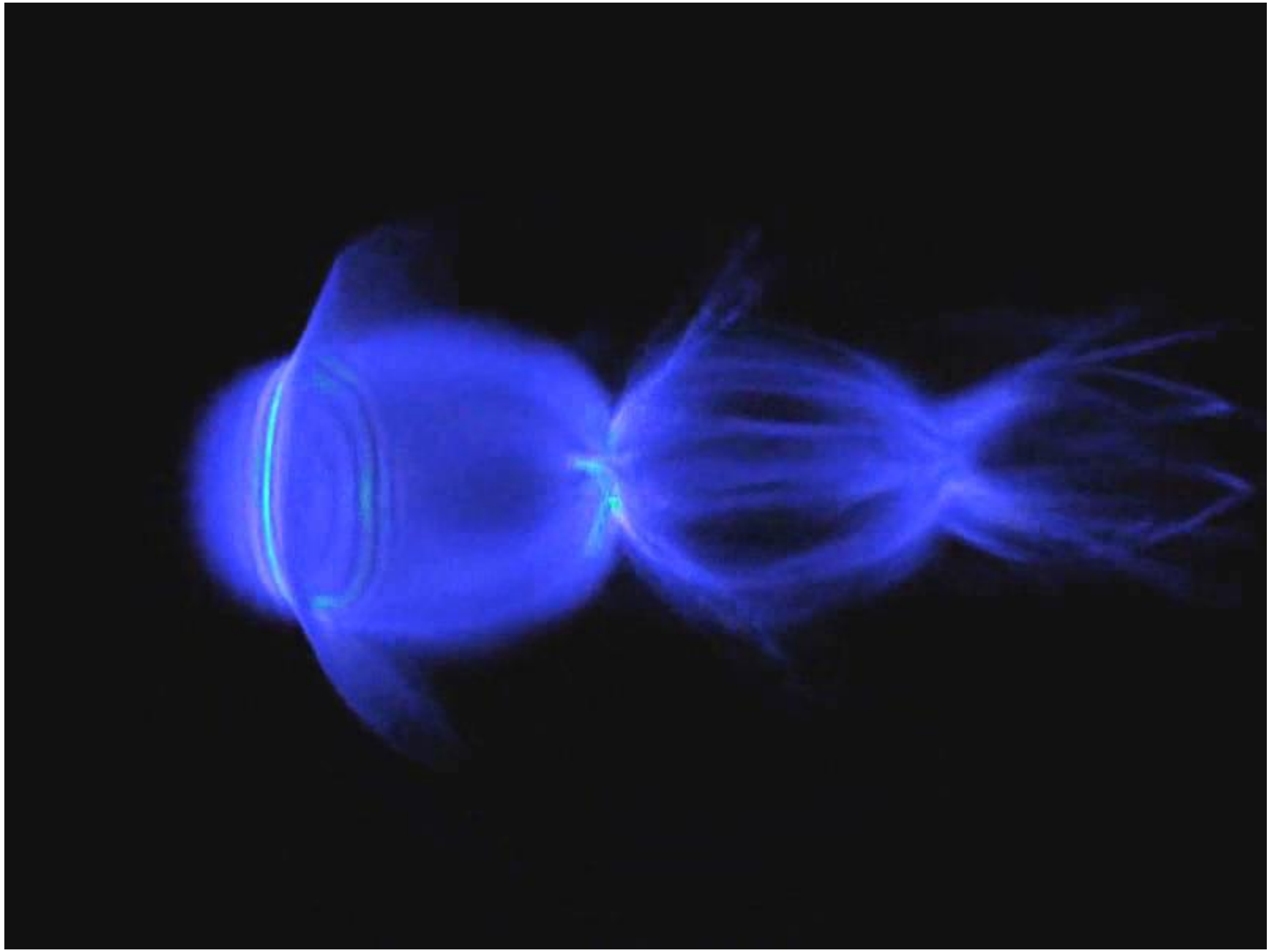
Software Visualization – Code City – Applied to ArgoUML



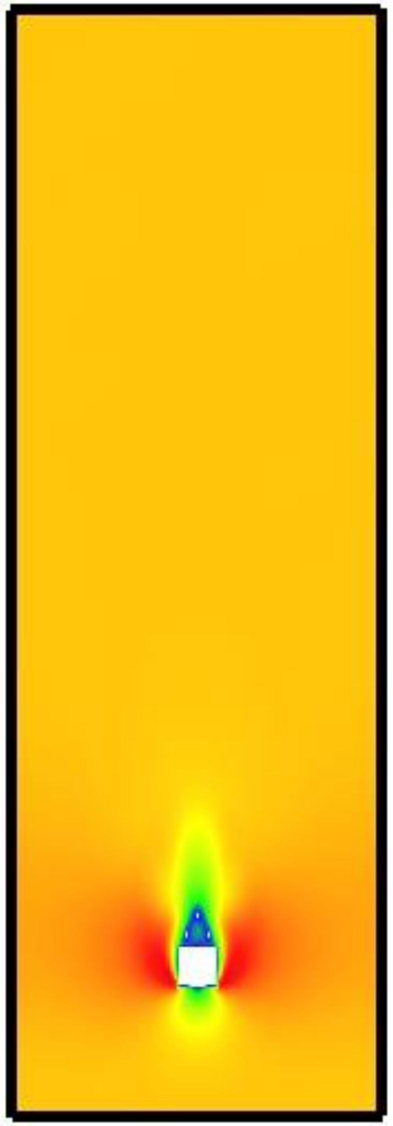
InfoVis Example
Multiple Coordinated Views, collaboratively used at large wall

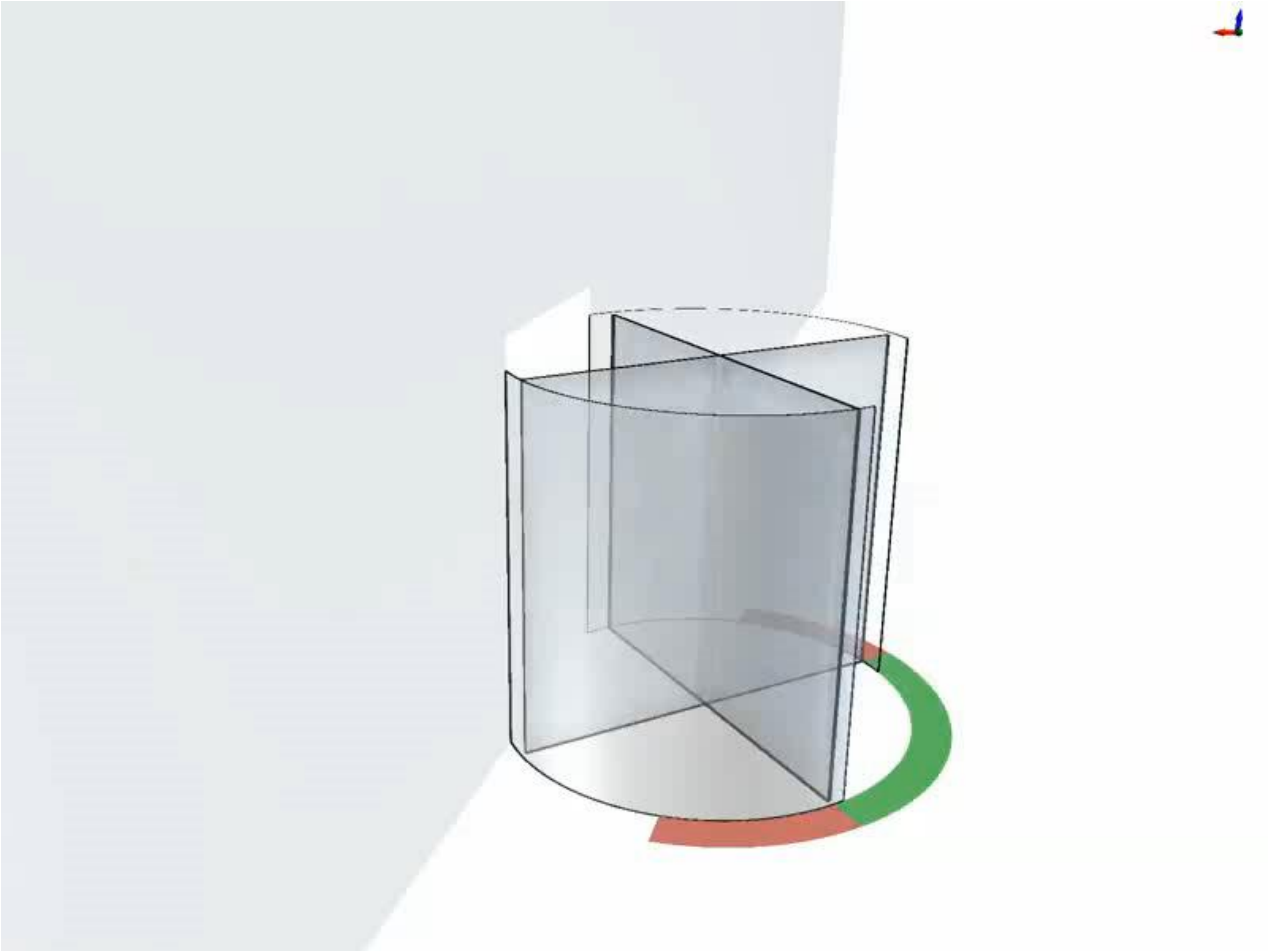
Their sheer display size also ***supports collaboration***,
so multiple users can interact with data and views at the same time

SciVis: Direct Volume & Flow Visualization

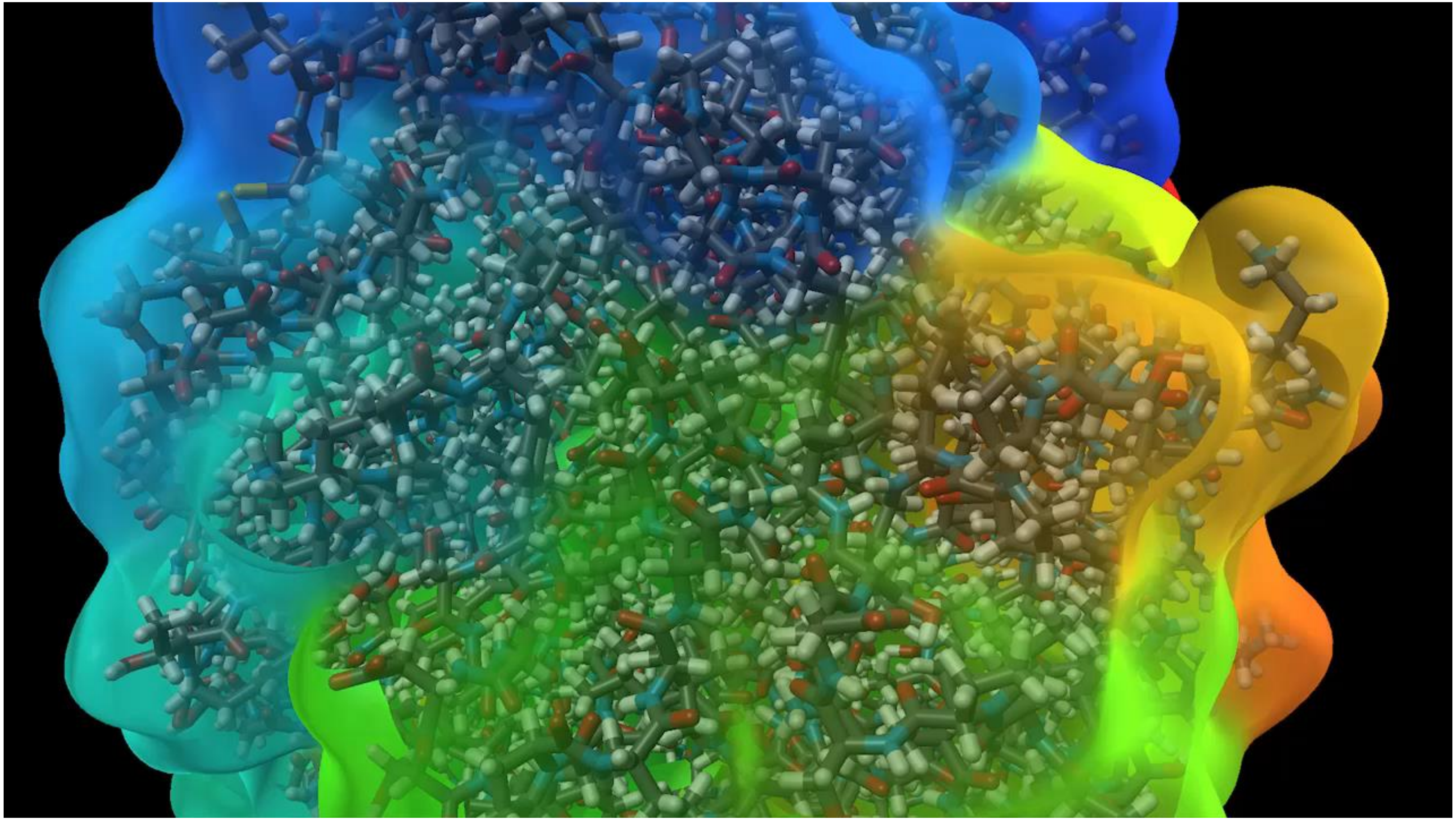


SciVis: Flow Visualization





BioVis: Protein Visualization



SciVis: Special Relativity Theory

4 Introduction

Data

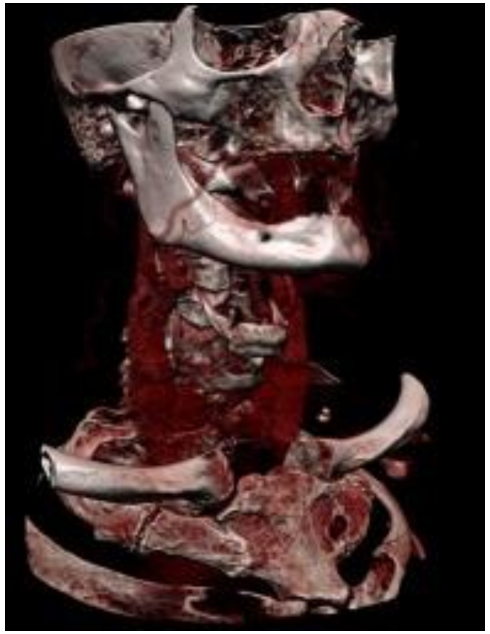
Data
Sources

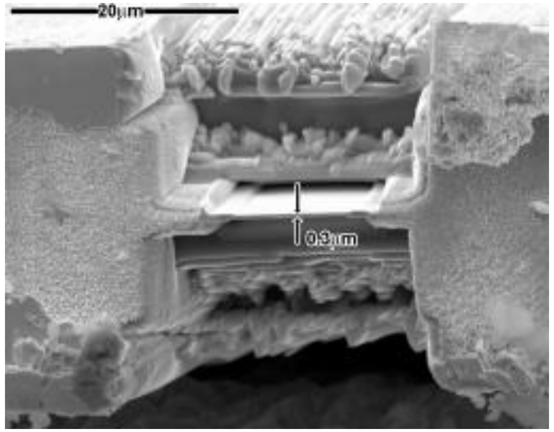
Real-world sources

- Medical Imaging (MRI, CT, PET)
- Geographical Information Systems (GIS)
- Microscopy [electron]
- Climate Data (Satellites)
- Crystallography
- High Energy Physics
- Astronomy

Virtual sources

- Simulation
- Web & social networks
- Software Development
- Trade
- Entertainment





Data
Models

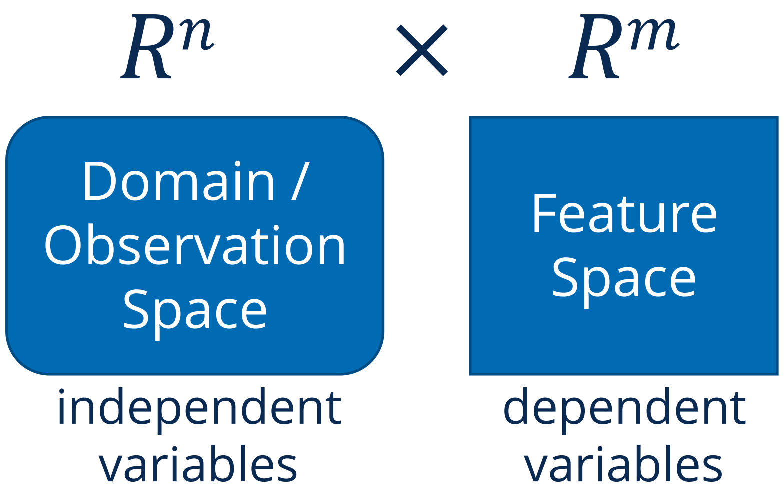
Data Models describe certain aspects of the world based on

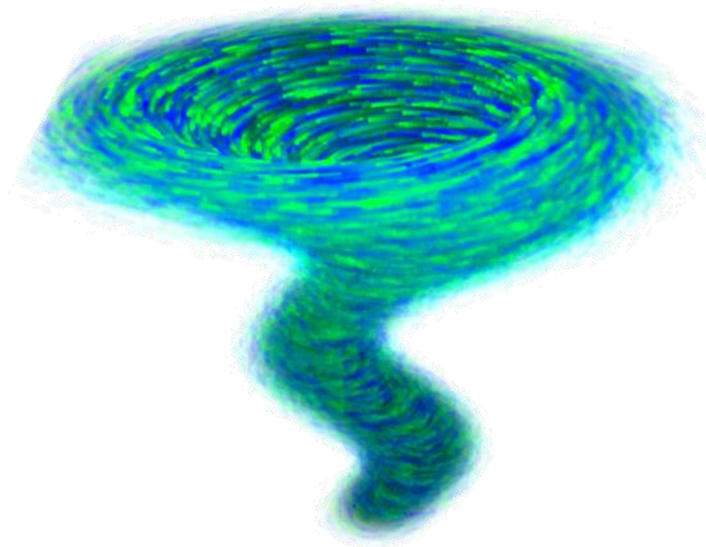
- domain or **observation space** – often a subset of R^n
- information **objects** (humans, cars, cities, particles, …)
- **features**, attributes or values (properties of objects), the **feature space** combines all feature dimensions into a multi-dimensional space R^m
- **relations** or clusters (bring several objects together and can be hierarchical)

Scientific Datasets are often very homogeneous and can then be described as subset of R^{n+m}

Example: Airflow - in a tornado

- Domain: $R^{n=3}$ (3D grid – x, y, z)
- Feature space: $R^{m=3}$ (Velocity – vx , vy, vz)





Data
Dimensionality

Examples for different dimensionalities:

A: Gas stations along a road

B: Cholera map of London

C: Simulation of temperature along rod

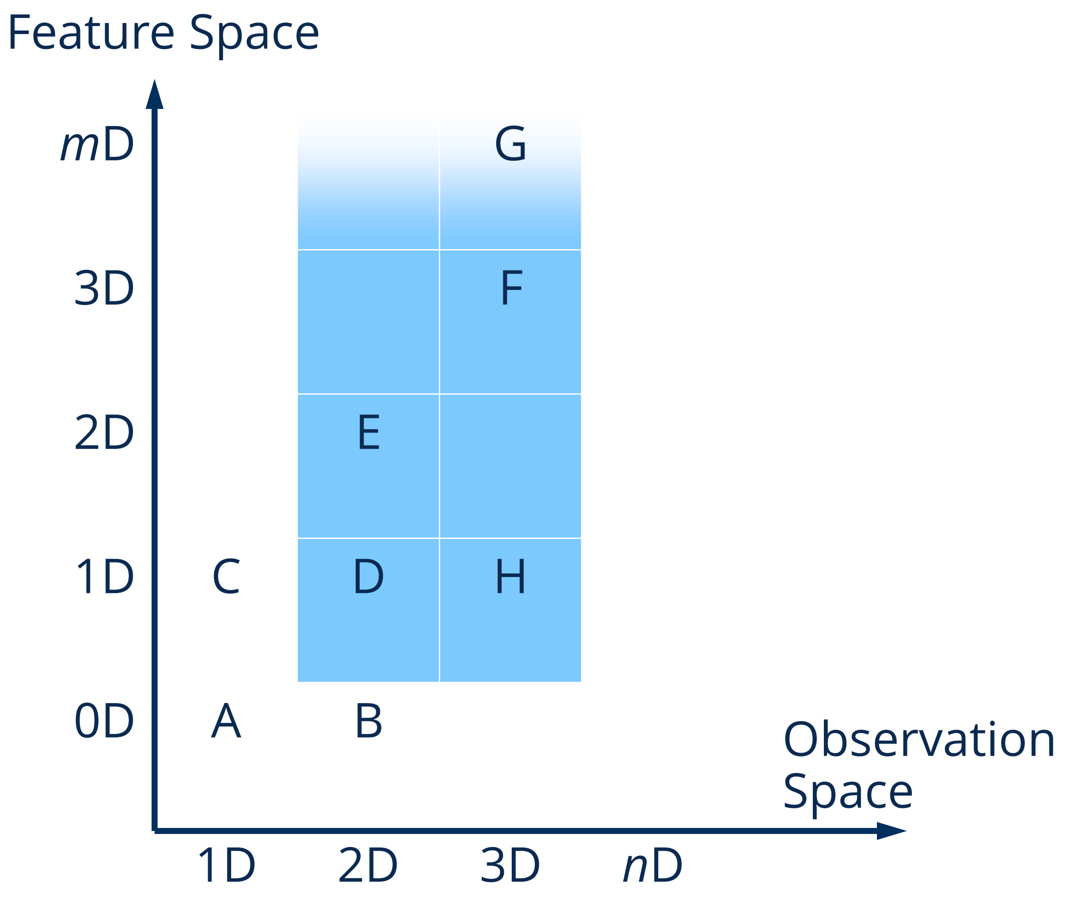
D: Height field of a continent

E: 2D airflow

F: 3D atmospheric airflow

G: Stress tensor in a work piece

H: Ozone concentration in the atmosphere



Data
Observation Space – Area of Influence

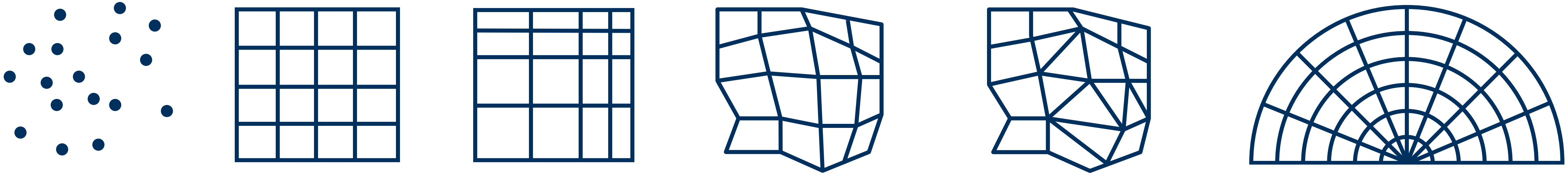
Area of Influence

- Reference point
  - The data values refer only to individual observation points (e.g. fractals, chaotic phenomena)
- Local reference
  - The data values refer to local regions of the observation space (e.g., temperature values as a function of position).
- Global reference
  - The data values apply to the entire observation space (e.g., background radiation, total energy).

Visualization Rules

- Reference point
  - The data values may only be assigned to individual points in the graphical representation.
- Local reference
  - Values at points where no value is given can be estimated from neighboring observations.
  - Suitable interpolation methods need to be defined.
- Global reference
  - Data values are assigned to a global area in the graphical representation.

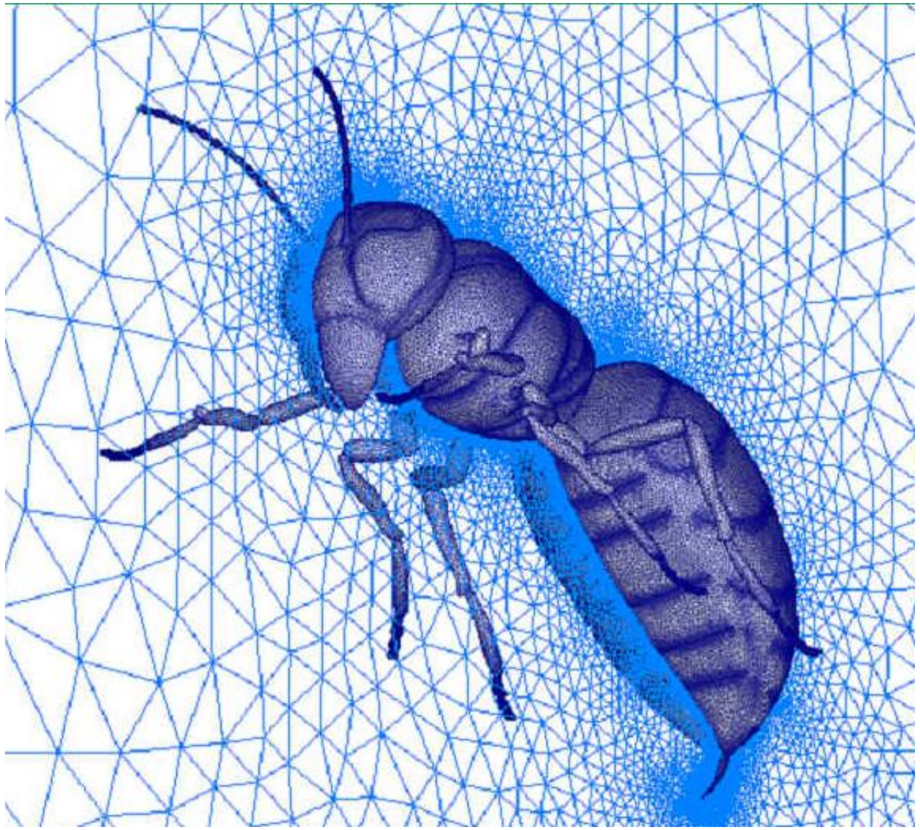
Data
Observation Space – Grid / Mesh Type



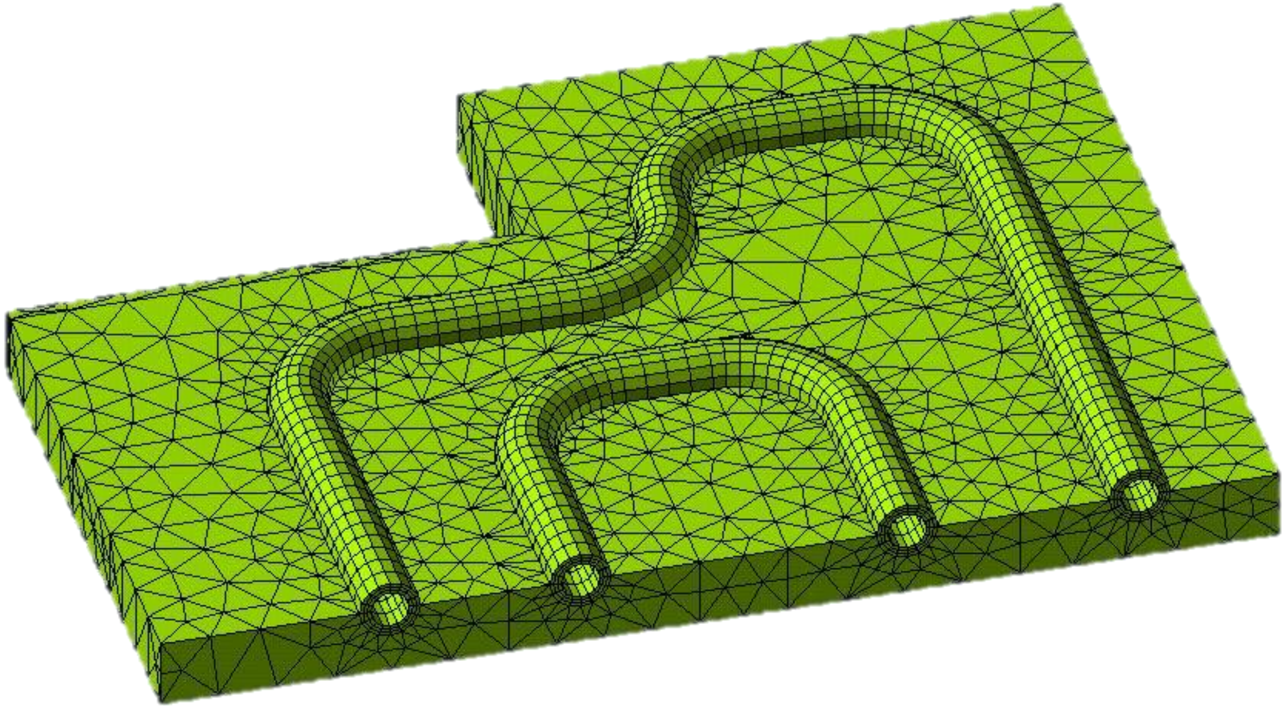
gridless | regular | block structured rectilinear | irregular | unstructured | curvilinear

grid type	abbreviation	geometry	connectivity
gridless	set/ord-set	explicit	X
regular grid	reg-grid	implicit	implicit
rectilinear grid	block-grid	explicit	implicit
curvilinear grid / mesh	curvi-grid	explicit	implicit
irregular grid / mesh	struct-grid	explicit	implicit
unstructured grid / mesh	unstruct-grid	explicit	explicit
hybrid grid			

Data
Observation Space – Grid / Mesh Type



Unstructured grid can locally **adapt** to data – here to surface inside of volumetric domain



Hybrid grid composed of unstructured tetrahedral grid (material) and structured hexahedral grid (pipes)

Data
Feature Space Description

Dimensionality (m)

Value range

- qualitative: no metric defined

  nominal: no order characteristics
  - typically names for class elements
  - only test for equality possible
  - one can impose order to convert to ordinal data

  ordinal: order relation is defined
  - first, second, third, …
  - freezing, cold, tepid, warm, hot…
  - equality test and trend analysis possible
- quantitative or metric: metric available, arithmetic operations (average, variance,…)
  - range can be discrete or continuous
  - quantization converts to ordinal data

Data types

- a scalar feature has 1D values, over domain it can be interpreted as sampled function $f(x_1, \dots, x_n): R^n \rightarrow R$ in the n independent variables x_i.
- a vectorial feature has nD values and can represent direction and magnitude of a vector. The feature function is defined as

$$\vec{\boldsymbol{v}}(x_1, \dots, x_n) = \begin{pmatrix} v_1(x_1, \dots, x_n) \\ \vdots \\ v_n(x_1, \dots, x_n) \end{pmatrix} : R^n \rightarrow R^n.$$

- a tensorial feature is a kth-order tensor with n^k dimension. values indexed over k indices: $\boldsymbol{T}(x_1, \dots, x_n) = \left(t_{i_1, i_2, \dots, i_k}(x_1, \dots, x_n)\right) : R^n \rightarrow R^{n^k}$; (scalars/vectors are 0th/1st order tensors)

5 Introduction
Summary and Outlook

Summary

You have learned something about

- Data-, Information- and Scientific Visualization: what it is and how it is defined
- The visualization pipeline and how to use it to gain knowledge
- Basic data models to describe aspects of the world
- The importance of interaction and the consideration of tasks

Literature

- A. Telea: ***Data Visualization***, A K Peters Ltd, (1st ed. 2007, 2nd ed. 2015)
- M. Watt, G. Grinstein, D Keim: ***Interactive Data Visualization: Foundations, Techniques, and Applications***, A K Peters Ltd, 2010
- C.D. Hansen, C.R. Johnson (eds.): ***The Visualization Handbook***, Elsevier, 2005
- C. Ware: ***Information Visualization: Perception for Design***, Elsevier/Morgan Kaufmann, (1st ed. 2000, 2nd ed. 2004, 3rd ed. 2012)
- Spence, R.: ***Information Visualization: Design for Interaction***, Prentice-Hall, Englewood Cliffs (2nd ed. 2007)
- Mazza, Riccardo: ***Introduction to Information Visualization***, Springer London (2009)
- Preim, Dachselt: ***Interaktive Systeme*** – Part „Interaktive Informationsvisualisierung“, Springer (2nd ed. 2010)
- Tufte, E.: **Envisioning Information**. Graphics Press, Cheshire (1990).
- G.M. Nielson, H.Hagen, H. Müller: ***Scientific Visualization,*** IEEE Computer Society Press, Los Alamitos, 1997
- K. Engel, M. Hadwiger, J. M. Kniss, C. Rezk-Salama, D. Weiskopf: ***Real-time Volume Graphics***, A K Peters, 2006
- D. Weiskopf: ***GPU-Based Interactive Visualization Techniques***, Springer, 2006
- W.J. Schroeder, K.W. Martin, B. Lorenson: ***The Visualization Toolkit: An Object-Oriented Approach to 3D Graphics***, Kitware, Clifton Park