Visual Analytics - Introduction

Eduard Gröller

Institute of Computer Graphics and Algorithms

Vienna University of Technology

Outline

- Introduction to visual analytics
- Definition of visual analytics
- Technical challenges and agenda
- Application areas

- Some slides courtesy of
  - Silvia Miksch
  - Daniel Keim / Jim Thomas

Motivation: Main Problems

Data Unmanageable – Loss of Overview

Missing Integration of

Various (Heterogeneous)
Information Sources

Various
Interdisciplinary Methods

Missing Involvement of

Users and their Tasks

Challenge of the Information Age

Konstanz University

- 100 million FedEx transactions per day
- 150 million VISA credit card transactions per day
- 300 million long distance calls in AT&T's network per day
- 50 billion e-mails worldwide per day
- 600 billion IP packets per day DE-CIX backbone

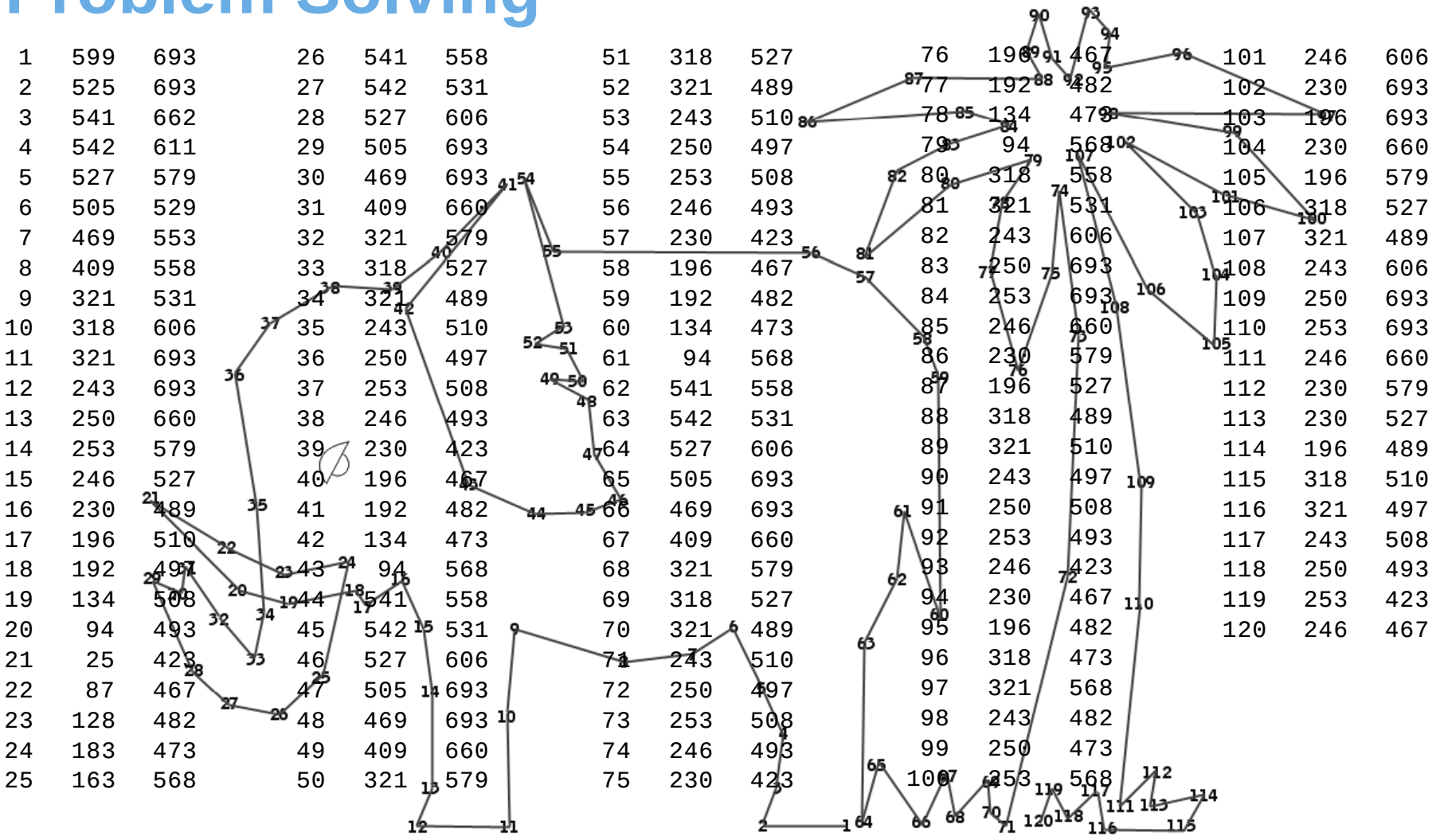
New Requirements Summary

Konstanz University

- Volume of data, orders of magnitude larger and different levels of abstraction
- Complexity of information spaces into very high dimensions, 200 the norm
- Information often out of context, incomplete, fuzzy
- Information in all media types: text, imagery, video, voice, web, sensor data
- Time and temporal dynamics fundamentally change the approach
- Spatial, yet non-spatial abstract data
- Multiple ontologies, languages, cultures

For many applications:
we now turn to data-intensive visual analytics

Visualization for Problem Solving



1	599	693	26	541	558	51	318	527	76	196	467	101	246	606
2	525	693	27	542	531	52	321	489	77	192	482	102	230	693
3	541	662	28	527	606	53	243	510	78	134	473	103	196	693
4	542	611	29	505	693	54	250	497	79	94	568	104	230	660
5	527	579	30	469	693	55	253	508	80	318	558	105	196	579
6	505	529	31	409	660	56	246	493	81	321	531	106	318	527
7	469	553	32	321	579	57	230	423	82	243	606	107	321	489
8	409	558	33	318	527	58	196	467	83	250	693	108	243	606
9	321	531	34	321	489	59	192	482	84	253	693	109	250	693
10	318	606	35	243	510	60	134	473	85	246	660	110	253	693
11	321	693	36	250	497	61	94	568	86	230	579	111	246	660
12	243	693	37	253	508	62	541	558	87	196	527	112	230	579
13	250	660	38	246	493	63	542	531	88	318	489	113	230	527
14	253	579	39	230	423	64	527	606	89	321	510	114	196	489
15	246	527	40	196	467	65	505	693	90	243	497	115	318	510
16	230	489	41	192	482	66	469	693	91	250	508	116	321	497
17	196	510	42	134	473	67	409	660	92	253	493	117	243	508
18	192	497	43	94	568	68	321	579	93	246	423	118	250	493
19	134	508	44	541	558	69	318	527	94	230	467	119	253	423
20	94	493	45	542	531	70	321	489	95	196	482	120	246	467
21	25	423	46	527	606	71	243	510	96	318	473			
22	87	467	47	505	693	72	250	497	97	321	568			
23	128	482	48	469	693	73	253	508	98	243	482			
24	183	473	49	409	660	74	246	493	99	250	473			
25	163	568	50	321	579	75	230	423	100	253	568			

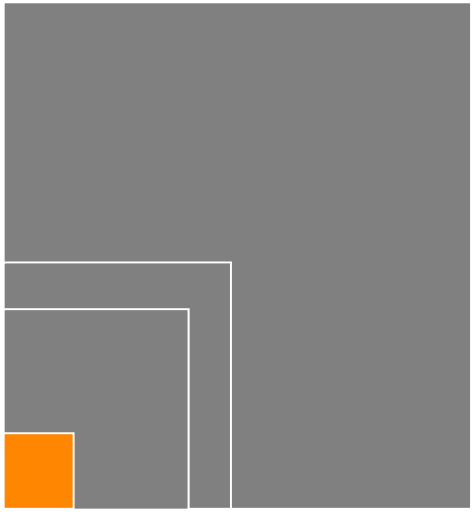
Analytical Methods

Screen Resolution: 1024 * 768 = **786.432**

Measurements of Water Level in LA Every Year: **5.256.000**

Number of Cellular Phones in Austria (2005): **8.160.000**

Transmitted Emails Every Hours (World-Wide): **35.388.000**



Whole Data often not Presentable

1. Applying Analytical Methods (*Data Reduction*)
2. Visualization of Most Important Data and Information

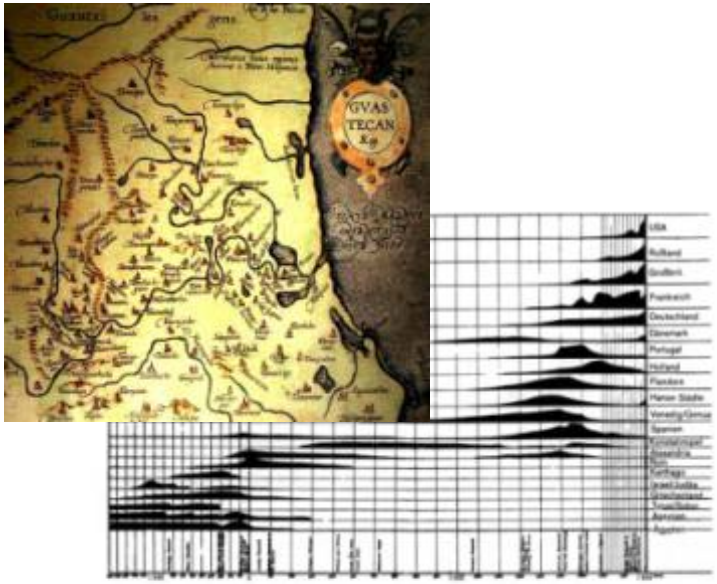
Analytical Methods

Statistics, Machine Learning & Data Mining

Interactions

Past

- Only passive Observations
- Representation not Changeable
- „one fits all“



Today

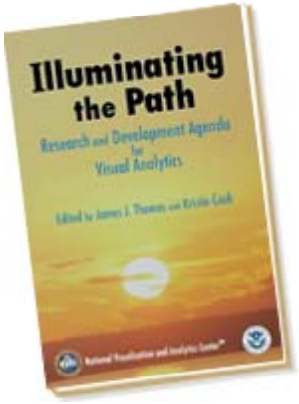
- Active Examination with Visualizations
- Dynamically Adaptable and Modifiable
  - → Different Users, Tasks and Aims

Visual Analytics – What is it?

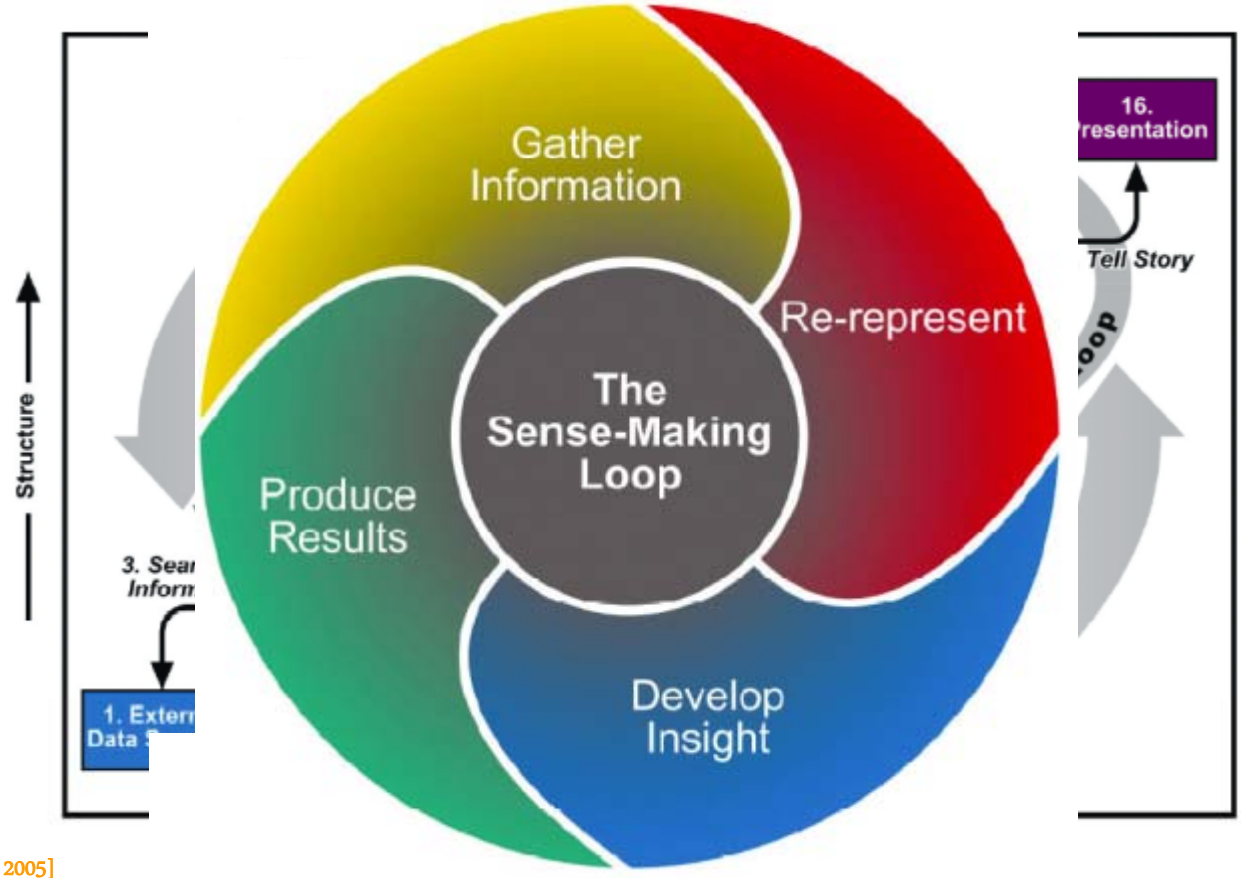
James Thomas & Kristin A. Cook:

NVAC (National Visualization and Analytics Center), Seattle, USA

„Visual Analytics is the science of analytical reasonin facilitated by interactive visual interfaces"



[Thomas & Cook 2005]



Visual Analytics Definition

Konstanz University

Visual Analytics is the science of analytical reasoning facilitated by interactive visual interfaces.

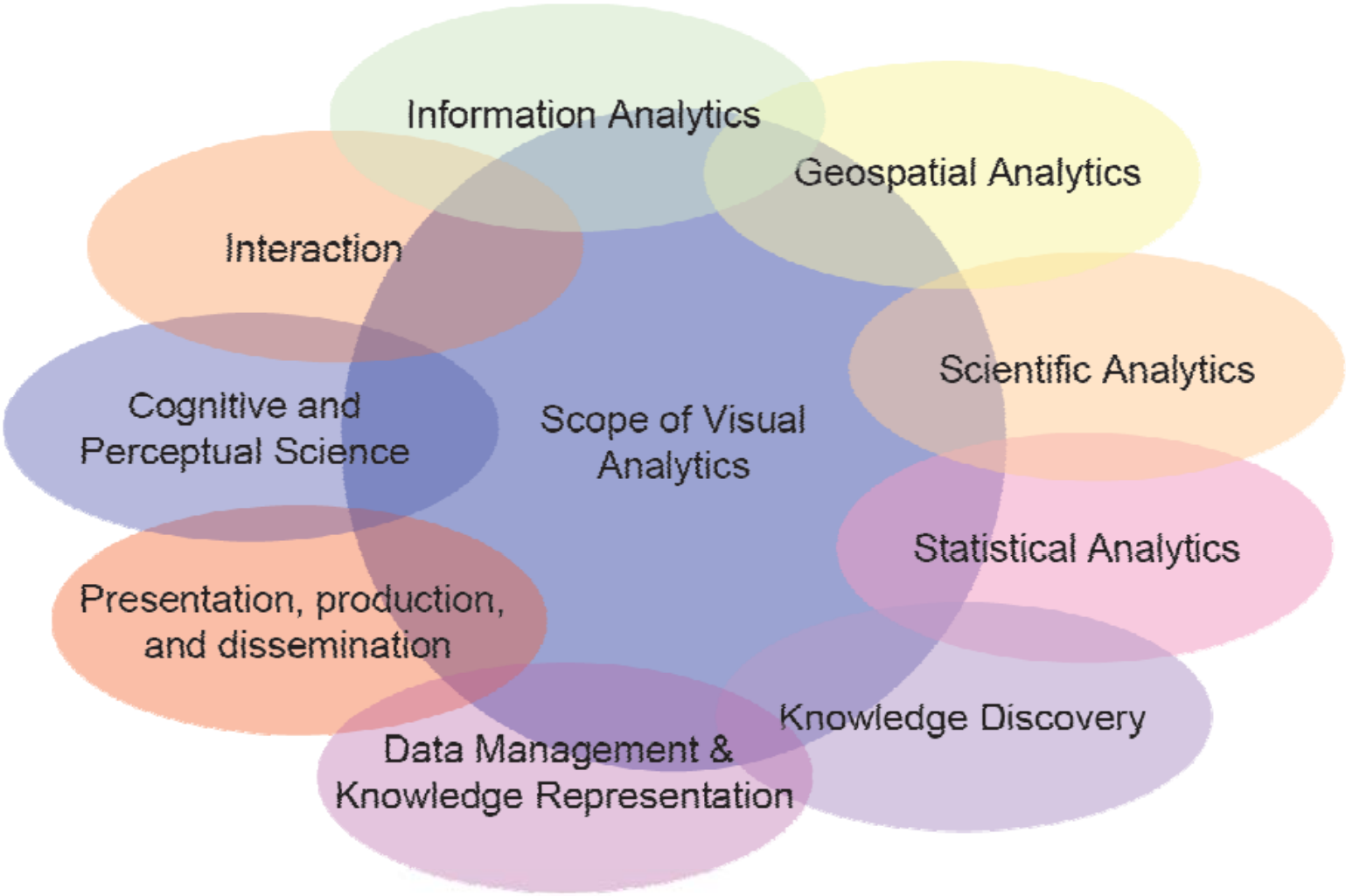
People use visual analytics tools and techniques to

- Synthesize information and derive insight from massive, dynamic, ambiguous, and often conflicting data.
- Detect the expected and discover the unexpected.
- Provide timely, defensible, and understandable assessments.
- Communicate assessment effectively for action.

"The beginning of knowledge is the discovery of something we do not understand."
~Frank Herbert (1920 - 1986)

Research Areas Related to Visual Analytics

Konstanz University



Why is the topic highly relevant today?

Konstanz University

- Very Large Data Collections are available in Databases and Data Warehouses
- On the Basis of the Data Complex Decisions have to made in a timely fashion
- Pure Visualization Methods (Information Visualisation) do not work for Billions of Data Records
- Full Automatic Knowledge Discovery Approaches only work for well-defined and clearly specifiable problems.
- Especially for adversarial situations:

  Fraud, Viruses, SPAM, Attacks, Competition, …

What is new ?

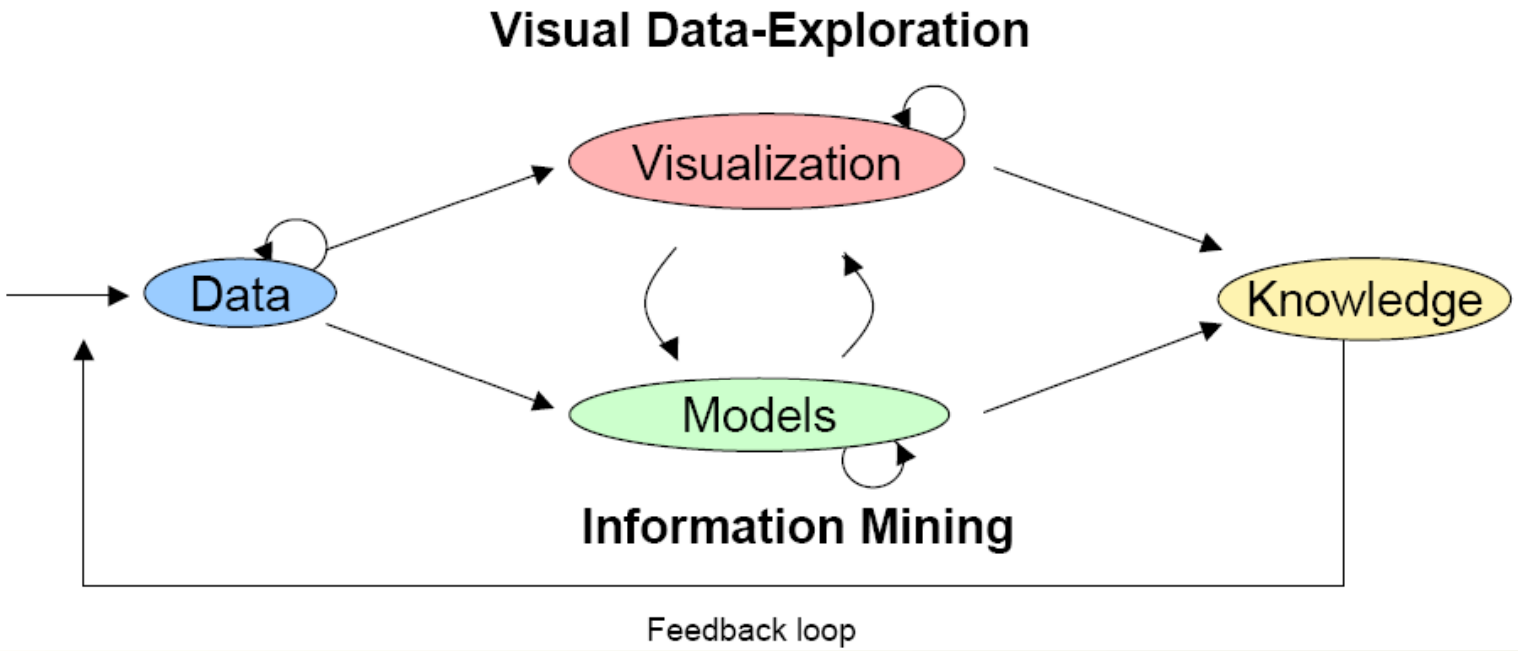
Konstanz University

What do we have?

- Automatic Knowledge Discovery & Information Mining
- Interactive Visual Data-Exploration

What do we need?

Tight Integration of Visual and Automatic Data Analysis Methods with Database Technology for a Scalable Interactive Decision Support



Technical Challenges

Konstanz University

Real-time Analysis of

- very large, complex, dynamic information
- from many diverse data sources
- in diverse formats and resolutions
- in uncertain, potentially life-threatening, and time-critical situations.

"Discovery consists of seeing what everybody has seen and thinking what nobody has thought."
~Albert von Szent-Gyorgyi (1893 - 1986)

Technical Challenge: Scalability

Konstanz University

Scalability w.r.t.

- Amount of Data and Dimensionality
- Number of Data Sources and Heterogeneity
- Data Quality and Data Resolution
- Dynamicity and Novelty
- Data Representation and Visual Resolution
- User Interface and Interaction
- Display Devices

> "All truths are easy to understand once they are discovered; the point is to discover them."
> ~ Galileo Galile (1564-1642)

Visual Analytics Agenda

[Thomas & Cook 2006]

The Science of Analytical Reasoning

*"... enable **users** to obtain **deep insights** that directly support assessment, planning, and decision making. "*

Visual Representations & Interaction Technologies

*"... take advantage of human eye's broad bandwidth pathway into the mind to allow users to **see**, **explore**, and **understand large amount of information** at once."*

Data Representations & Transformations

*"... **covert** all types of conflicting and dynamic data in ways that support visualization and analysis."*

Production, Presentation, & Dissemination

*"... **communicate** information in the appropriate context to a **variety of audience**."*

Application Areas

Economic & Business Data

- Business Intelligence
- Market Analysis

Medicine & Biotechnology

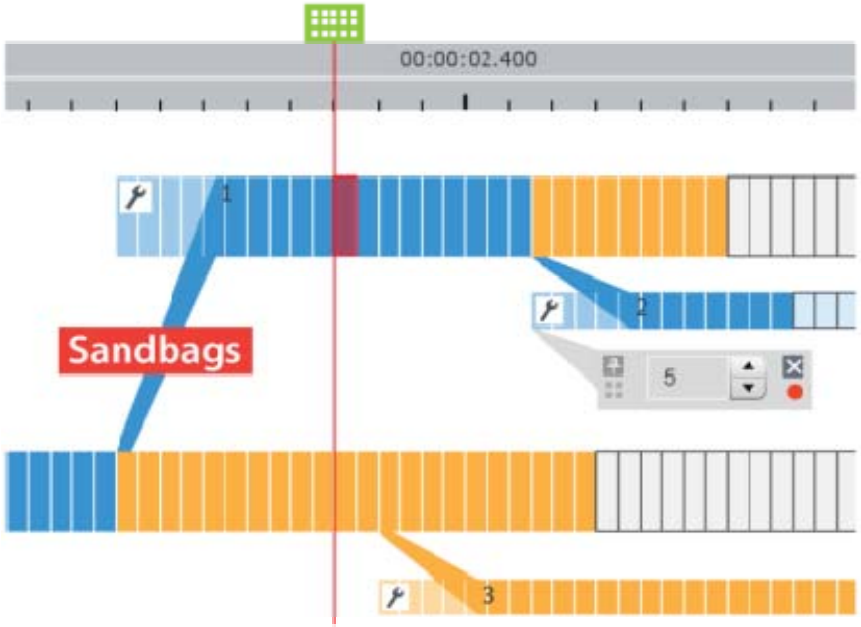
- Patients' Data Management
- Epidemiology
- Genetics

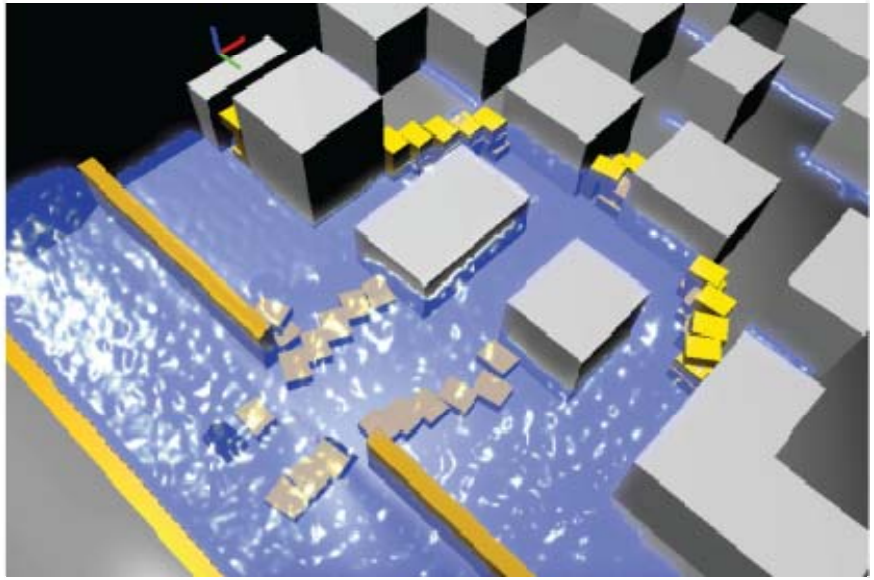
Security & Risk Management

- Disaster Management
- Computer Networks
- Transportation
- Reducing Crime and Terror Rate
- Fraud Detection

Environment & Climate Research

etc.





Visual Steering to Support Decision Making in Visdom

Jürgen Waser

http://www.cg.tuwien.ac.at/research/publications/2011/waser_2011_VSD/

Visdom

vrvis

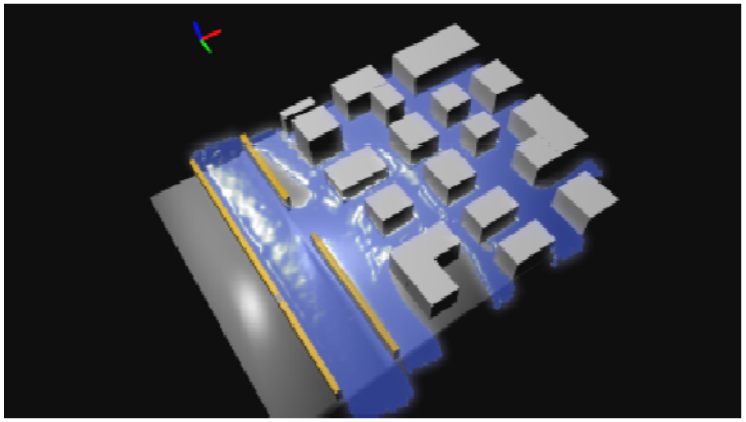
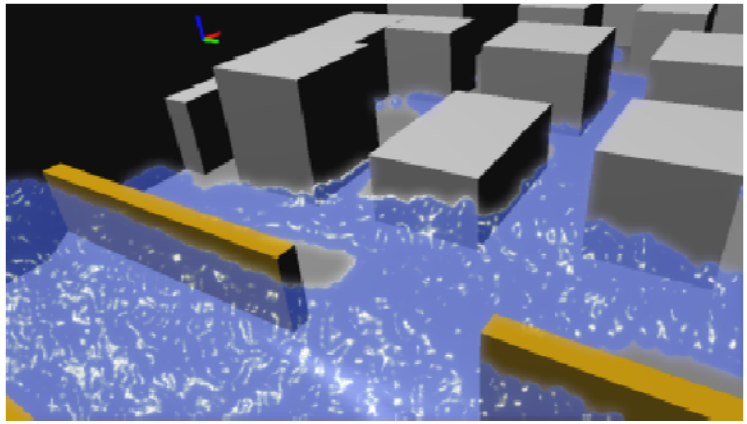
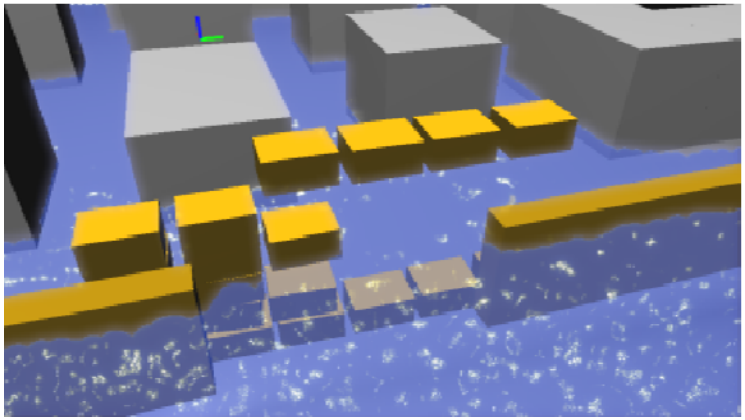
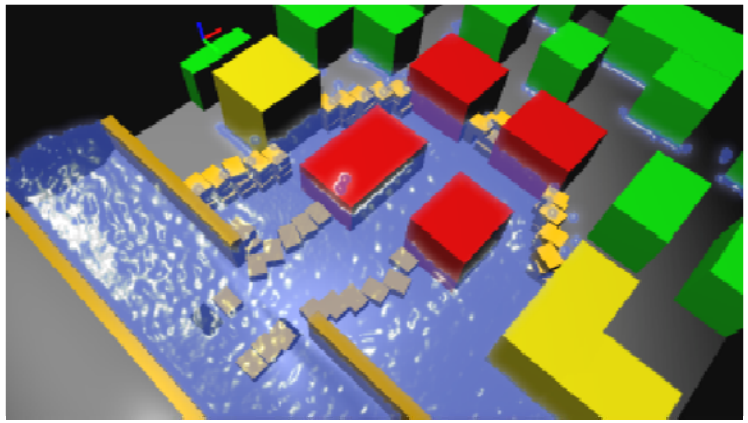
Flood emergency assistance

- **New Orleans 2005: 17th canal levee breach**

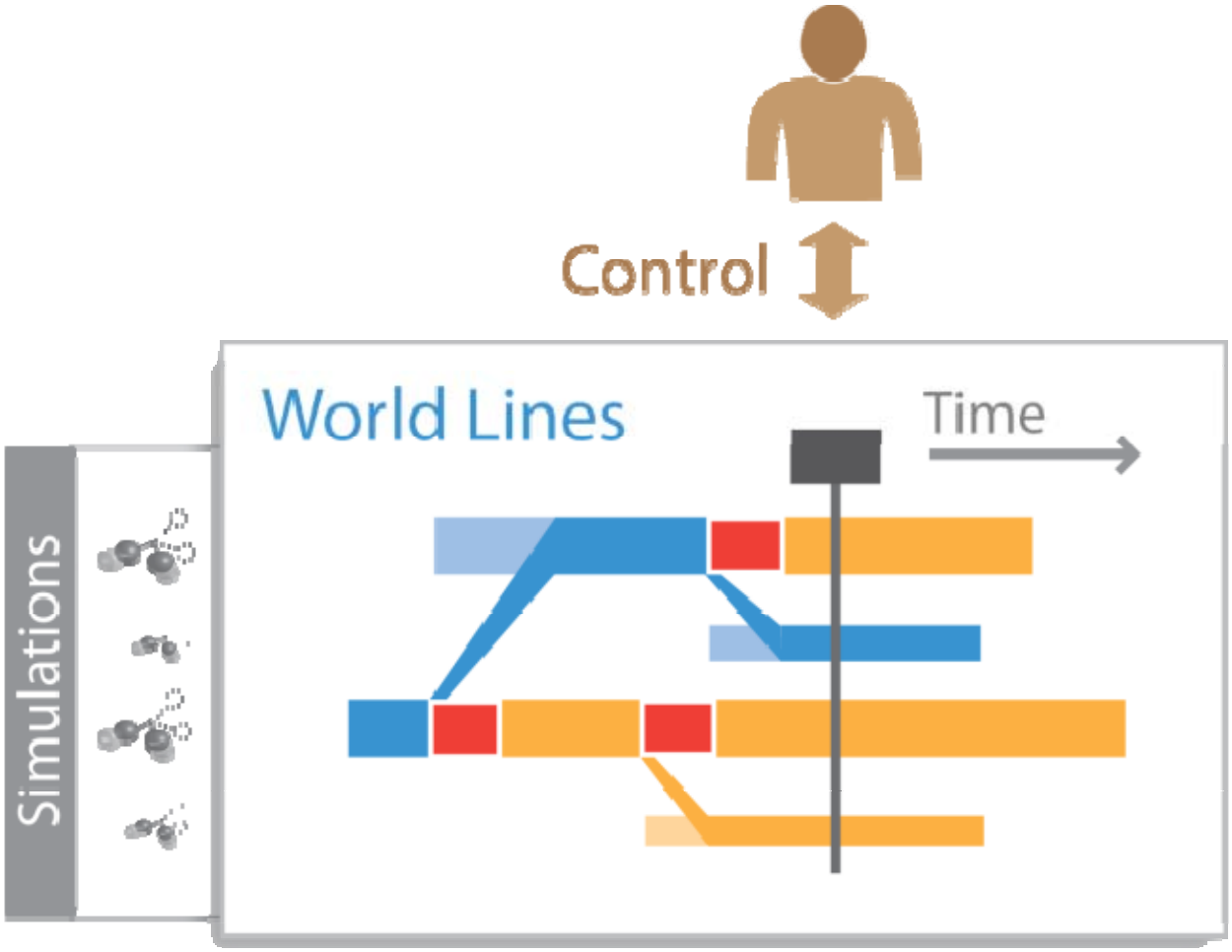
Image courtesy of USACE, US Army Corps of Engineers

Flood emergency assistance

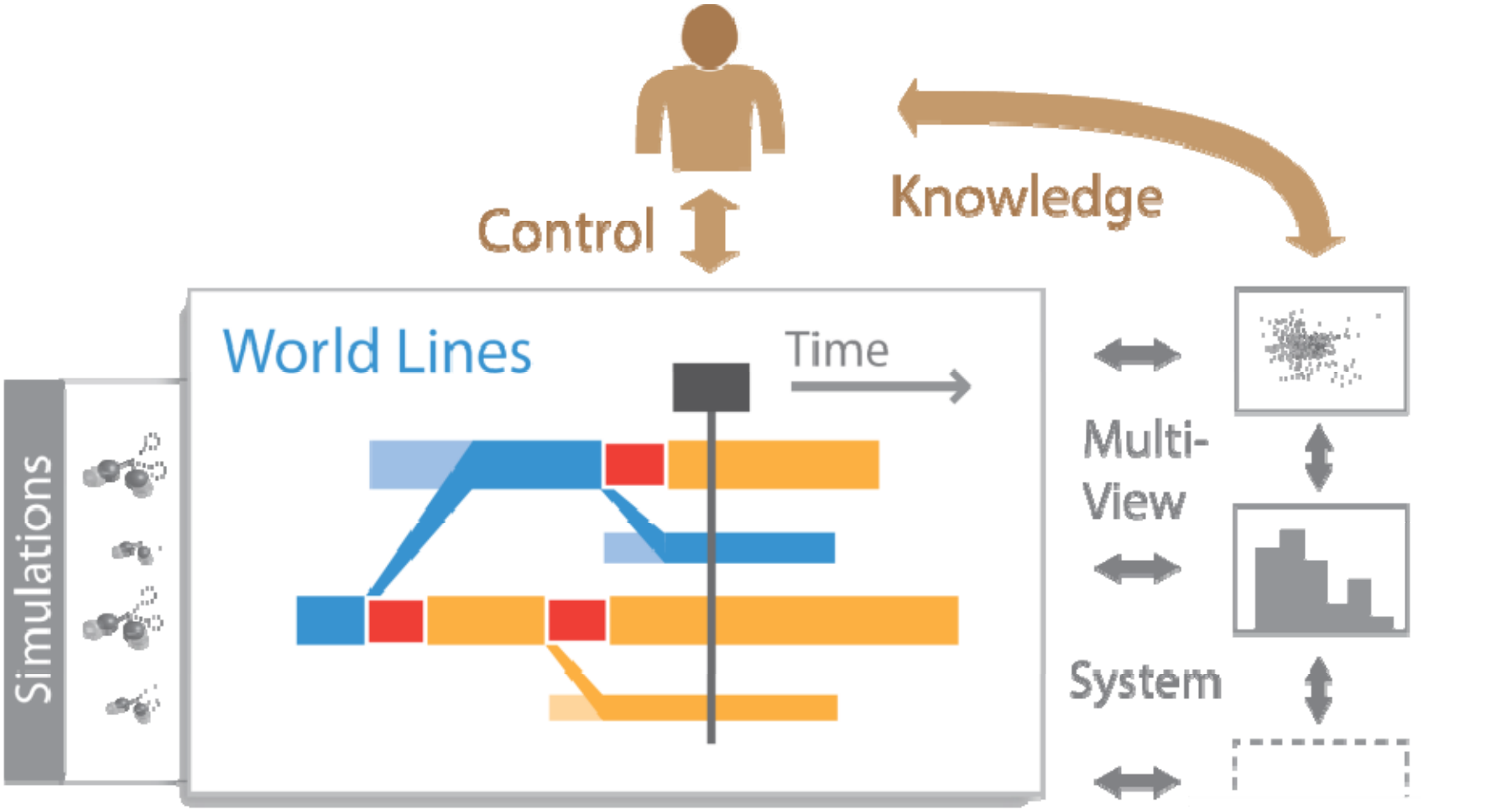
- **Testing sandbag configurations in a virtual environment**

Solution: World Lines

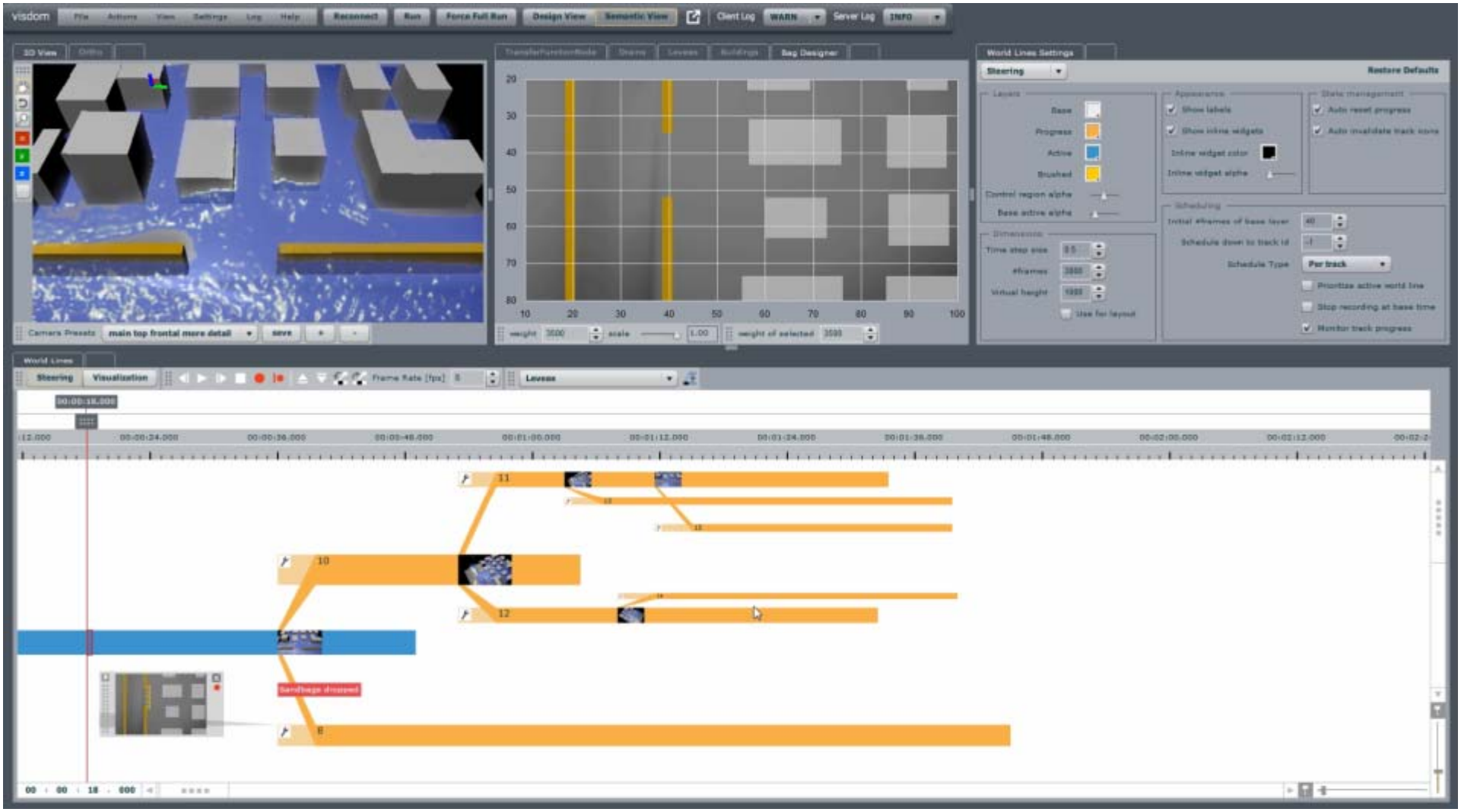


Solution: World Lines

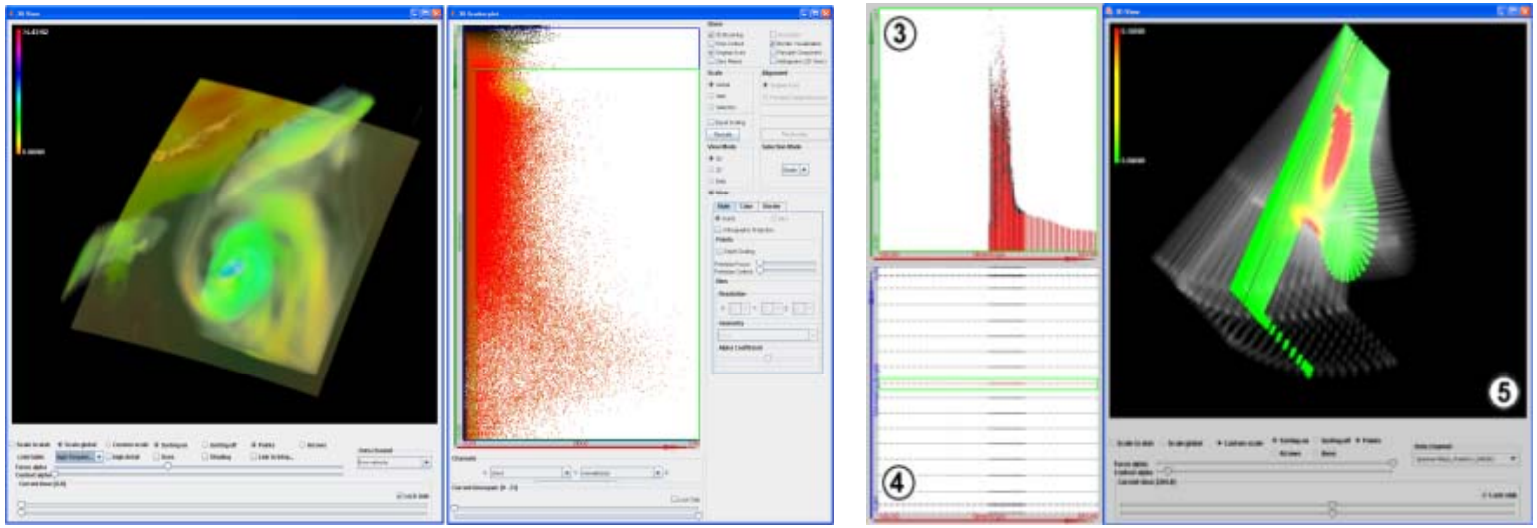


Video

Worldlines – Multiple Linked Views



SimVis: Interactive Visual Analysis of Large & Complex Simulation Data



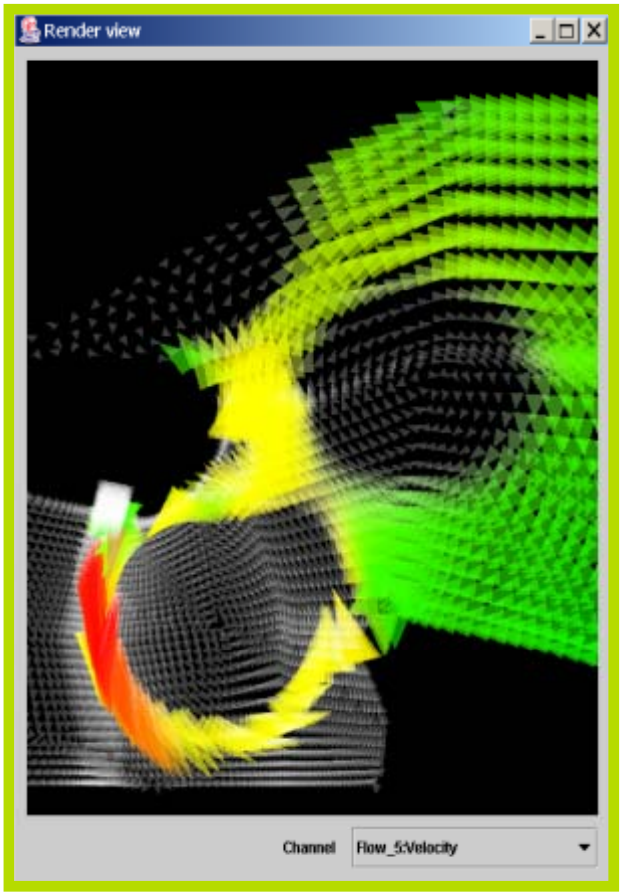
Dr. Helmut Doleisch
VRVis Research Center

http://www.VRVis.at/

vrvis

Motivation

- large data sets from simulation
- **goal**: support **exploration** and **analysis** of results
  - analyze n-dim. data interactively
  - use 3D visualization
  - overview, zoom and filter, detail on demand (Shneidermans' information seeking mantra)
- challenge:
  - occlusion
  - interactive data handling



Interactive Data Handling

- sample data set size:
  - 540 million data items
  - currently working to expand to billions

cells	timesteps	attributes	cells * timesteps	cells * timesteps * attributes
704.900	20	16	14.098.000	225.568.000
150.124	600	6	90.074.400	540.446.400
7.680.000	288	15	2.211.840.000	33.177.600.000

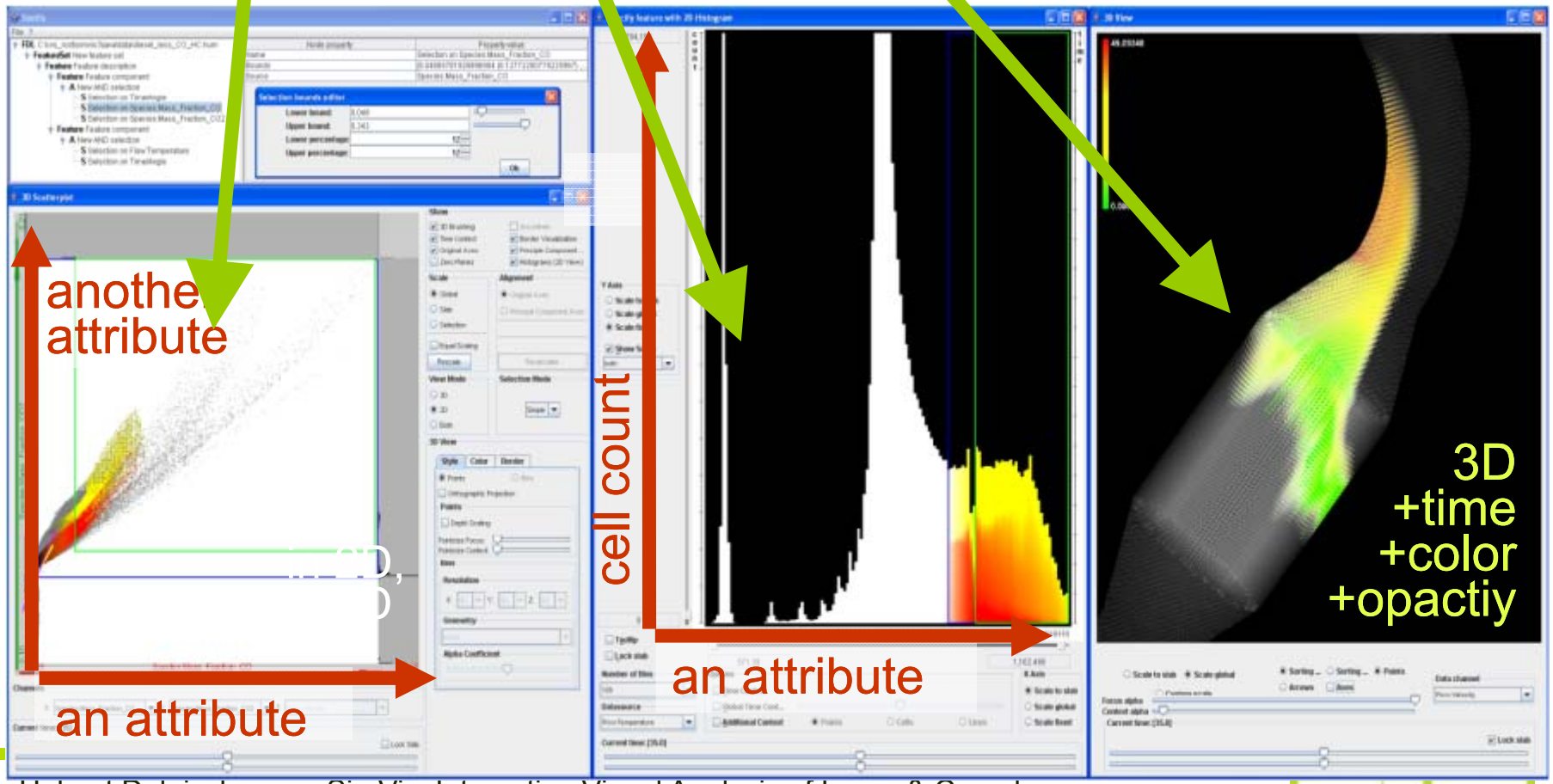
SimVis

- VRVis´ solution for these challenges
- Feature-based visualization framework

- SimVis key features:
  - Multiple, linked views
  - Interactive feature specification
  - Focus+Context visualization
  - Smooth feature boundaries
  - Explicit feature representation
  - On-the-fly attribute derivation

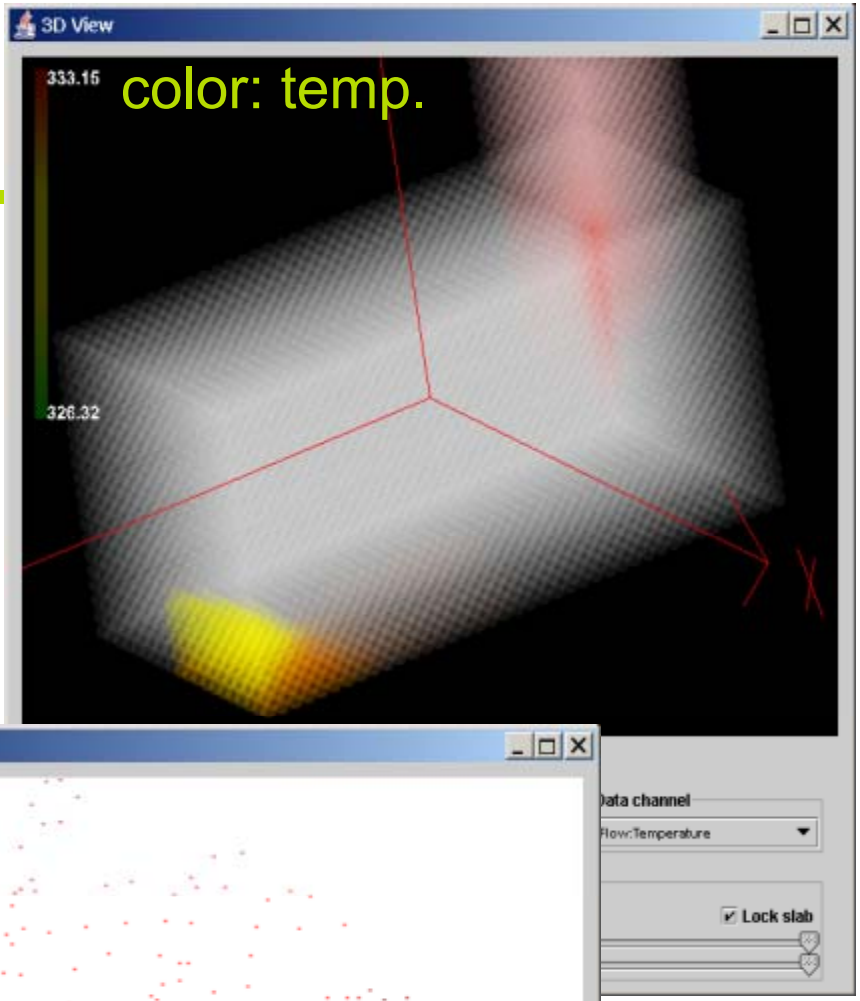
SimVis: Multiple Views

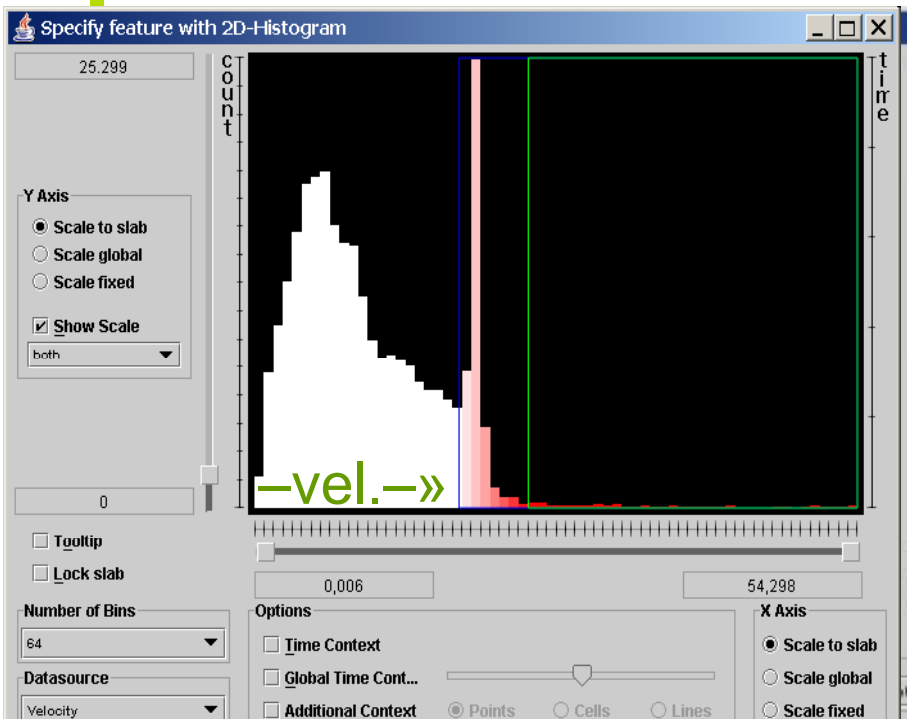
- Scatterplots, histogram, 3D(4D) view, etc.

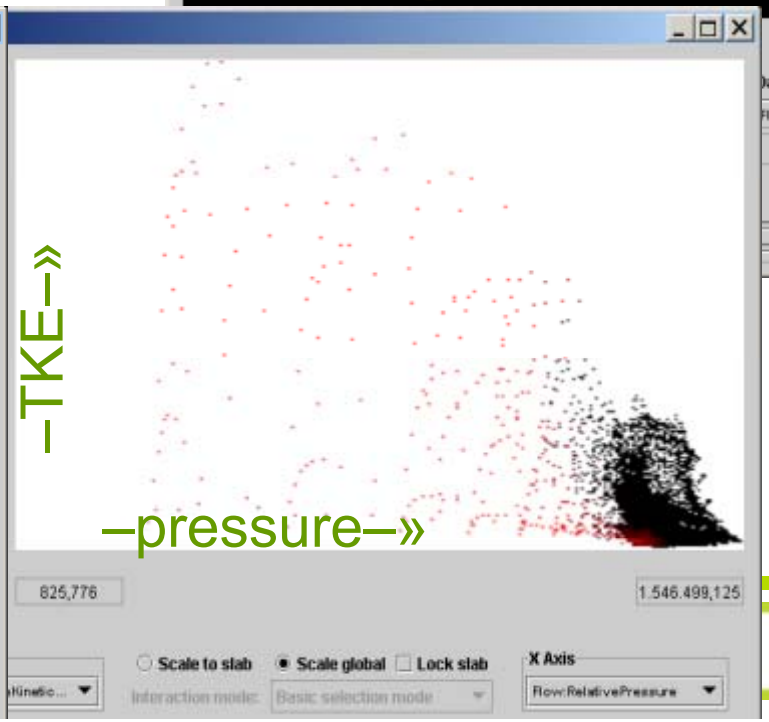


Brushing

- Move/alter/extend brush interactively
- Update linked F+C views in real-time







vrvis

Conclusions

Konstanz University

- Visual Analytics is an opportunity worth considering
- Collaboration between academia, industry, national laboratories, and government (national and international) is key
- For each of you:

The best is yet to come…