Visual Text Analytics for Technology and Innovation Management

h_da
HOCHSCHULE DARMSTADT
UNIVERSITY OF APPLIED SCIENCES

Conference

OPENRHEINMAIN 2019

23 September 2019

Kawa Nazemi and **Dirk Burkhardt**

Human-Computer Interaction & Visual Analytics
Darmstadt University of Applied Science

dirk.burkhardt@h-da.de
www.vis.h-da.de

Presentation

Motivation (I)

- Until 2025 around 163 Zettabytes (the number has 21 zeros) will be generated → 500 Mio. times of Netflix
- Around 80-90% of the available information are stored in unstructured form, many of that is raw text
- Due to information and social media platforms the daily (textual) information growth is still accelerating
- Additionally a number of new (Open-) Data initiatives, strategies or platforms came on the market
- Processing data to support decision making is standard, but almost on structured data

Motivation (II)

- Due to digitalization, rapidly changing markets, massive innovation investigations, market observations and trend recognition gets extremely important
- Particularly for the ICT domain/enterprises the technology and innovation management plays and important role
- How to make use out of all these textual (open/public) data?
- What requirements or principles have to be considered to retrieve valuable insights?

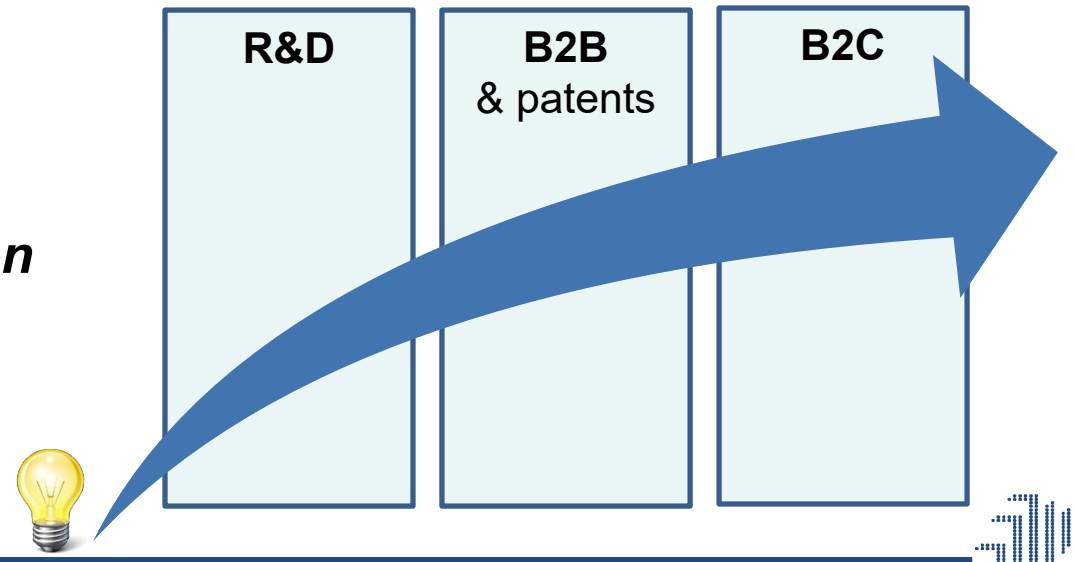
Motivation (III)

- The early awareness of emerging technological and innovation trends is essential for analyzing the potential of future technologies

- With Visual Text Analytics data can be analyzed graphically, based on automated analysis techniques in the background

→ ***Visual enables detecting emerging technological and innovation trends and discovering their potentials***

→ ***Stages of technological awareness demands on the business domain***

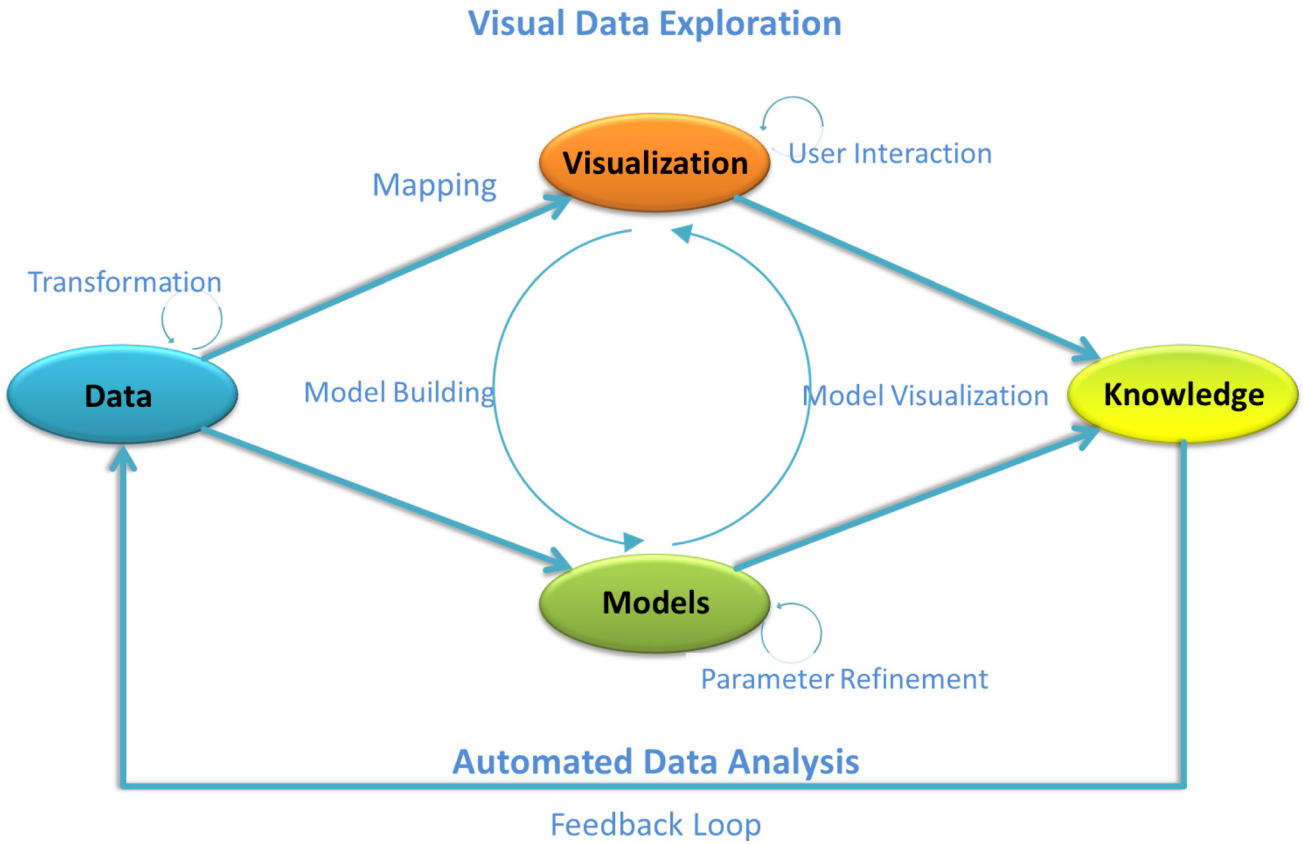


Background: Visual Analytics in Technology Management

- **Questions to be answered:**
  - **When** have technologies emerged and when established?
  - **Where** are the key-players and key-locations?
  - **Who** are the key-players?
  - **Which** technologies are relevant?
  - **How** will the technologies probably evolve in the next years?

Background: Visual Analytics

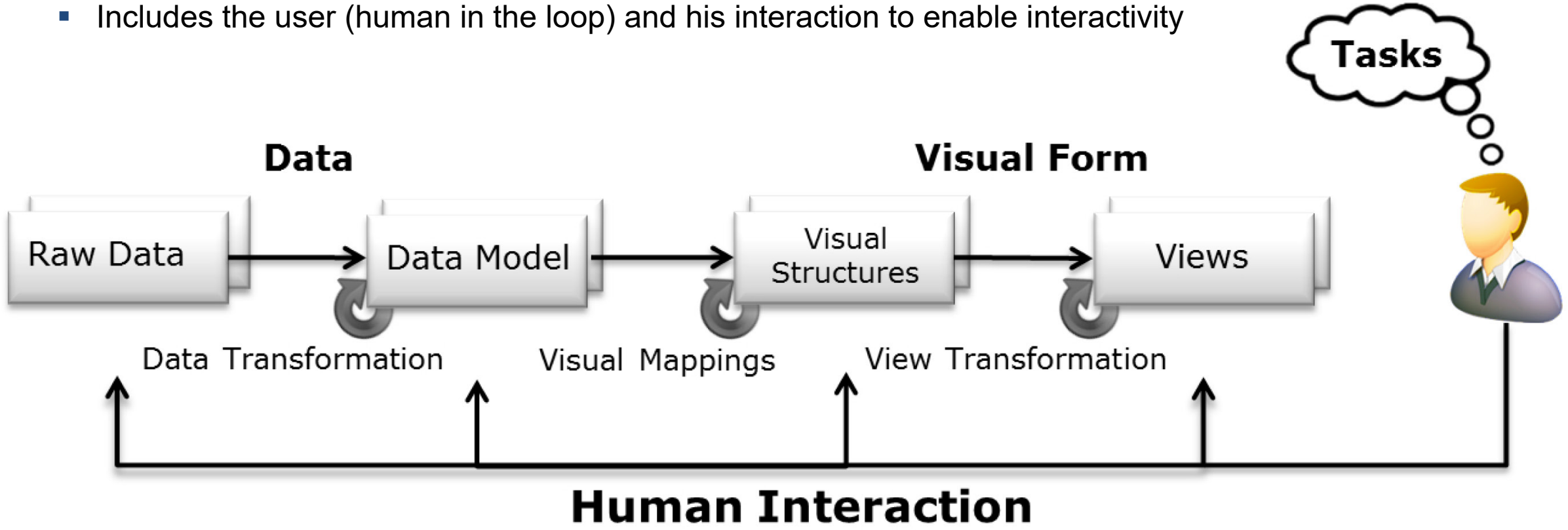
- Visual Analytics Process
  - Combines automated analysis techniques with interactive visualizations → strong binding of visualization and model



Keim D., Kohlhammer J., Ellis G., Mansmann F.: Matering the Information Age Solving Problems with Visual Analytics. Eurographics Association, 2010.

Background: Information Visualization

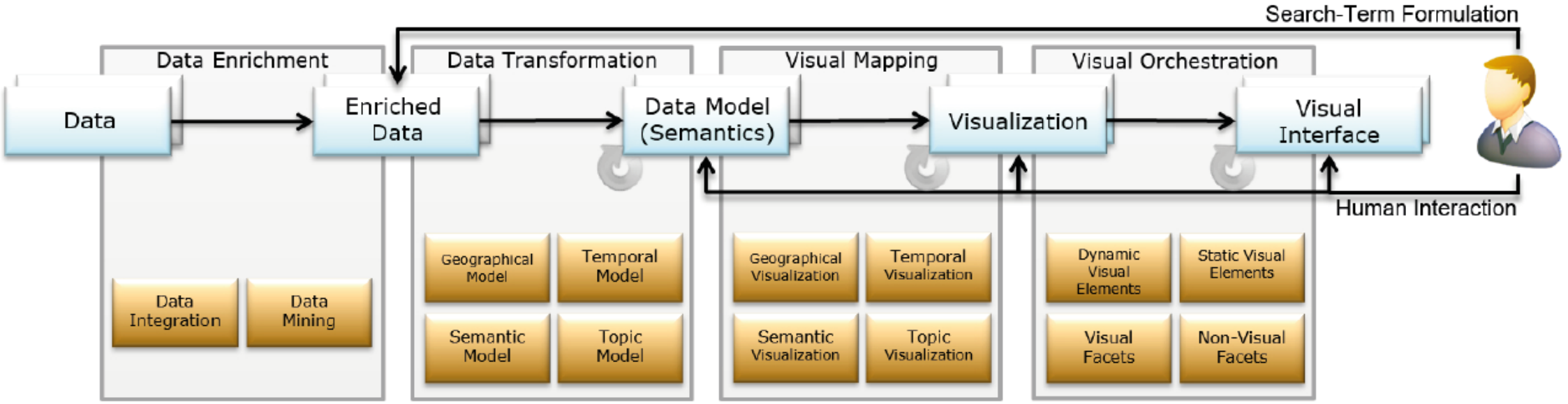
- Information Visualization Model
  - Transformation pipeline from data to views
  - Includes the user (human in the loop) and his interaction to enable interactivity



Card S. K., Mackinlay J. D., Shneiderman B.: Readings in Information Visualization: Using Vision to Think, 1. ed. Morgan Kaufmann, February 1999.

Visual Text Analytics Model

- Visual Text/Trend Analytics Model that convers a variety of processing stages
  - In each stage a number of specific actions can be performed
  - The interaction of the user can impact and manipulate any step and stage



K. Nazemi, R. Retz, D. Burkhardt, A. Kuijper, J. Kohlhammer, and Dieter W. Fellner. 2015. Visual trend analysis with digital libraries. In *Proceedings of the 15th International Conference on Knowledge Technologies and Data-driven Business*. ACM, New York, NY, USA. doi: 10.1145/2809563.2809569
K. Nazemi and D. Burkhardt, "Visual Analytics for Analyzing Technological Trends from Text," *2019 23rd International Conference Information Visualisation , IEEE*, Paris, France, 2019, pp. 191-200. doi: 10.1109/IV.2019.00041

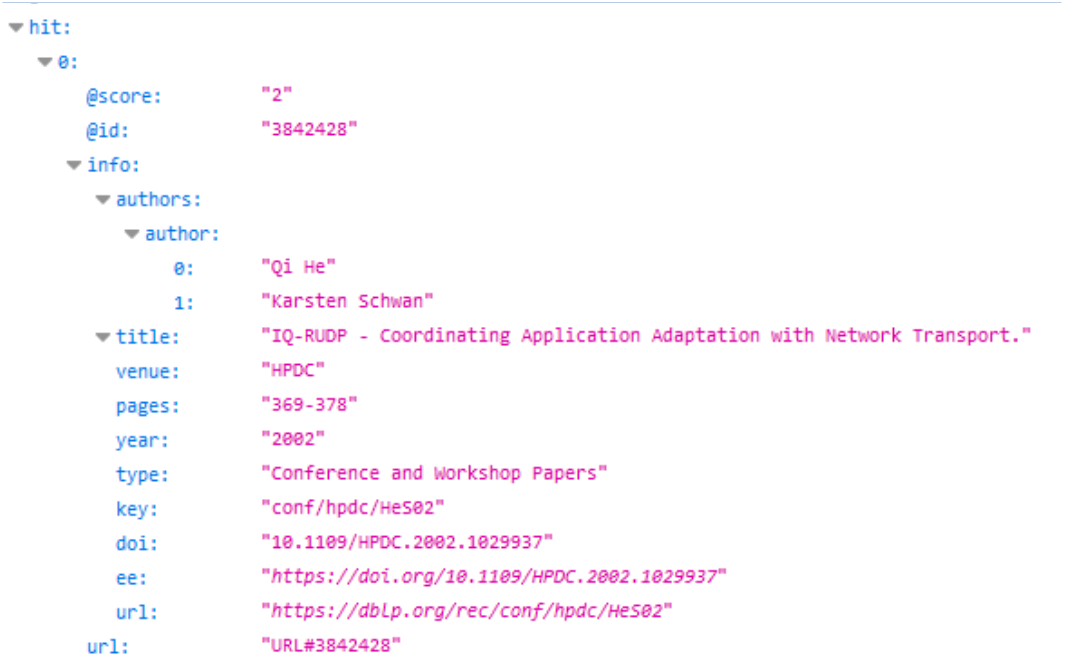
Visual Text Analytics Stages: Data (I)

- As data almost any textual source is suitable to analyze as far as it owns technological or innovation information:
  - Patent data → DPMA, EPO
  - Research data (publication) → CrossRef, DBLP, Springer, IEEE, ACM
  - News (from market players and news sites)
  - Social Media → LinkedIn, Twitter, ResearchGate
  - Internal (organizational) data → reports, studies

- Third-level data to complete/cleanse/verify/merge information:
  - Wikipedia/DBpedia
  - APIs such as from the national library

Visual Text Analytics Stages: Data (II)

- Gathering initial data (e.g. DBLP)
- Filter data by new entries
- Indexing the initial data
- Storing metadata including unique identifier (e.g. DOI)

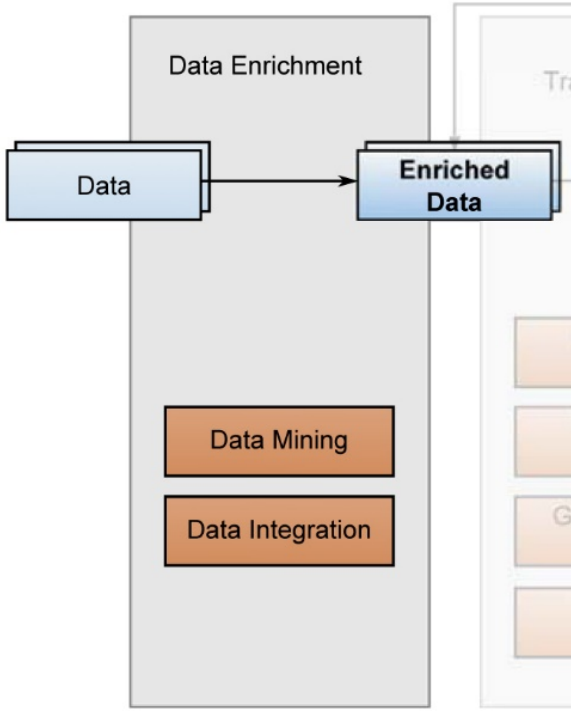


Visual Text Analytics Stages: Data Enrichment (I)

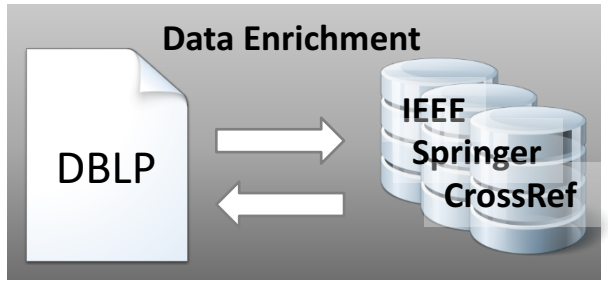
- Gathering additional data from web e.g. Springer, IEEE, IEEE Computer Society, CrossRef to complete DBLP

- Enriched Data includes
  - abstracts (and full-texts) for topic extraction
  - country (authors' origin)
  - affiliation of the authors
  - ...

- Generating information directly from given text
  - Entity extraction / Distant Supervision
  - Data mining: topics e.g. through Latent Dirichlet Allocation (LDA)



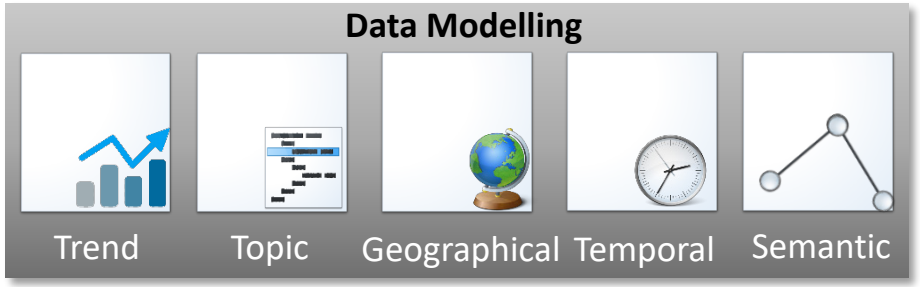
Visual Text Analytics Stages: Data Enrichment (II)

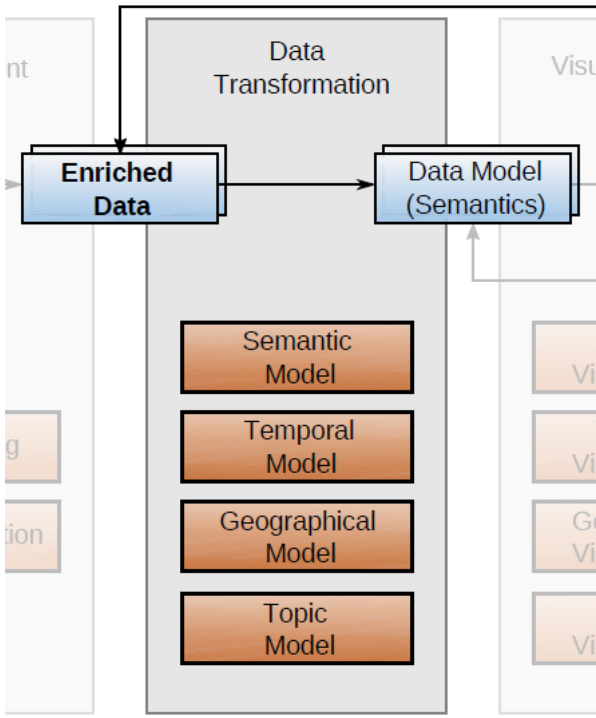


affiliation:
0: "Coll. of Comput., Georgia Inst. of Technol., Atlanta, GA, USA"
thesaurusterms: […]
uncontrolledterms: […]
abstract: "Our research addresses the efficient transfer of large data across wide-area networks, focusing on applications l
application-level with network transport-level adaptations of data communication. Specifically, complementing prev
violating fairness in network resource usage. The approach is evaluated with the IQ-ECho middleware, which impleme
underlying IQ-ECho's communications. The application- to network-layer exchange of information necessary for such
pub:
0: "High Performance Distributed Computing, 2002. HPDC-11 2002. Proceedings. 11th IEEE International Symposium on"
RAWcomputerorg: {…}

Visual Text Analytics Stages: Data Transformation

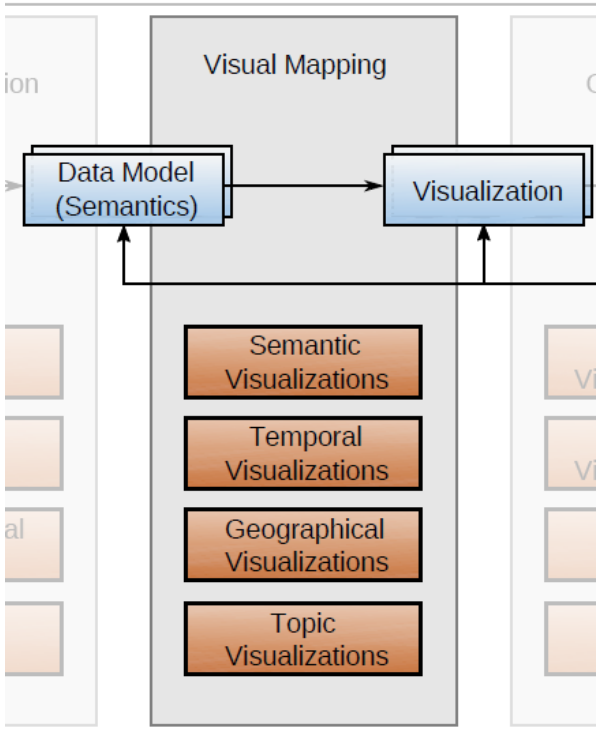
- Generating aspect-oriented data models with the particular aspects of data, e.g. temporal or geographical
- semantic data model serves as the primary data model for holding all information
- data models are the baseline for the visual structures



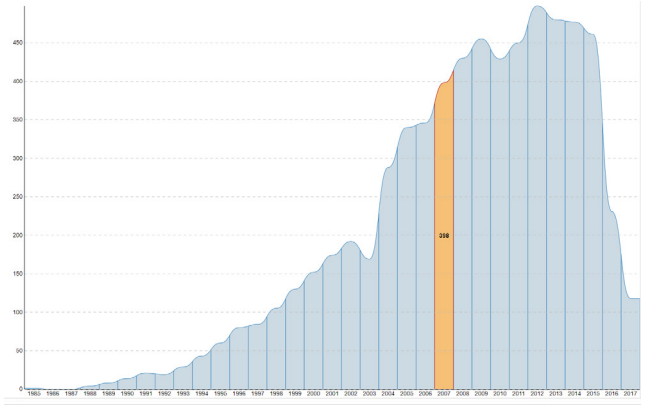


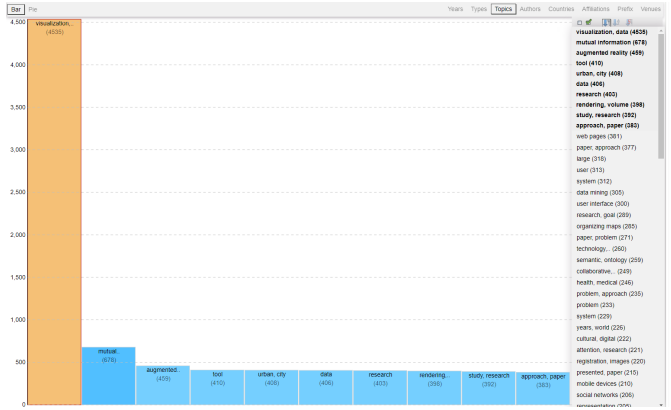
Visual Text Analytics Stages: Visual Mapping (I)

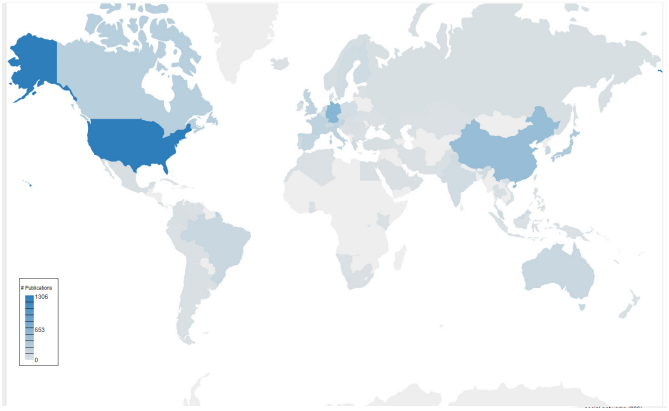
- Various visual structures for the underlying data models
- Visual structures leads to interactive visualizations
- Interactive Visualization for the different data models and tasks
- According to the data models

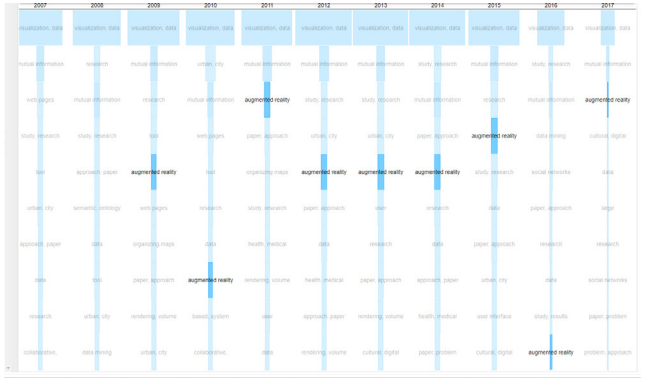


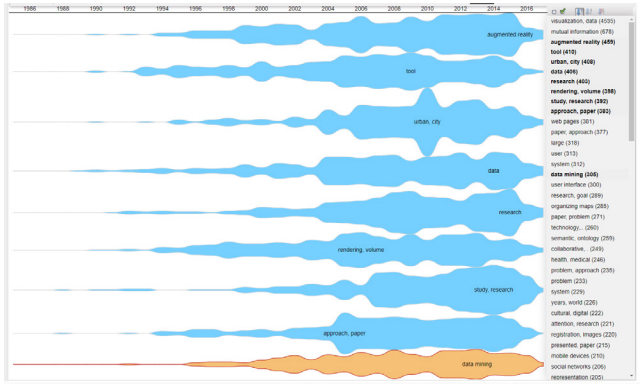
Visual Text Analytics Stages: Visual Mapping (II)

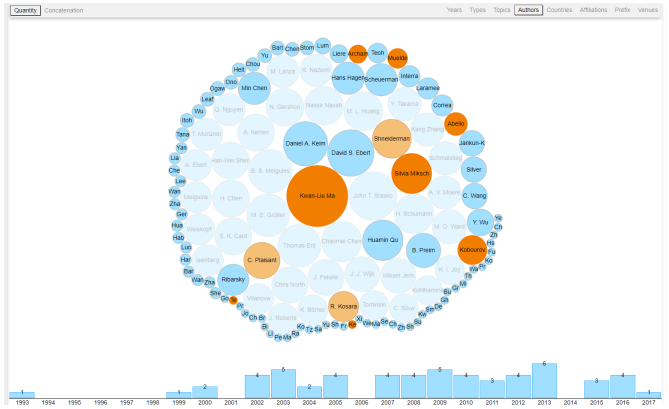






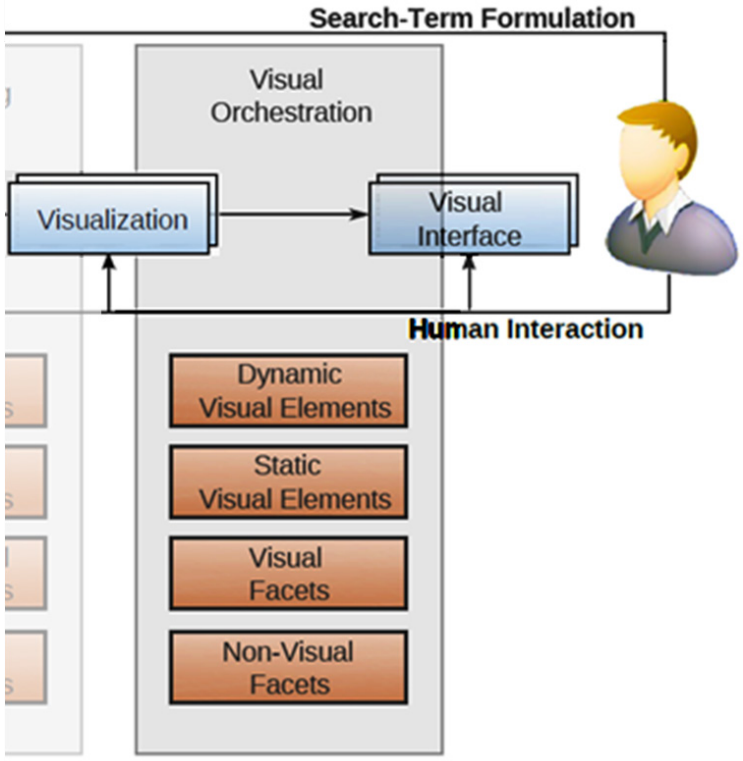






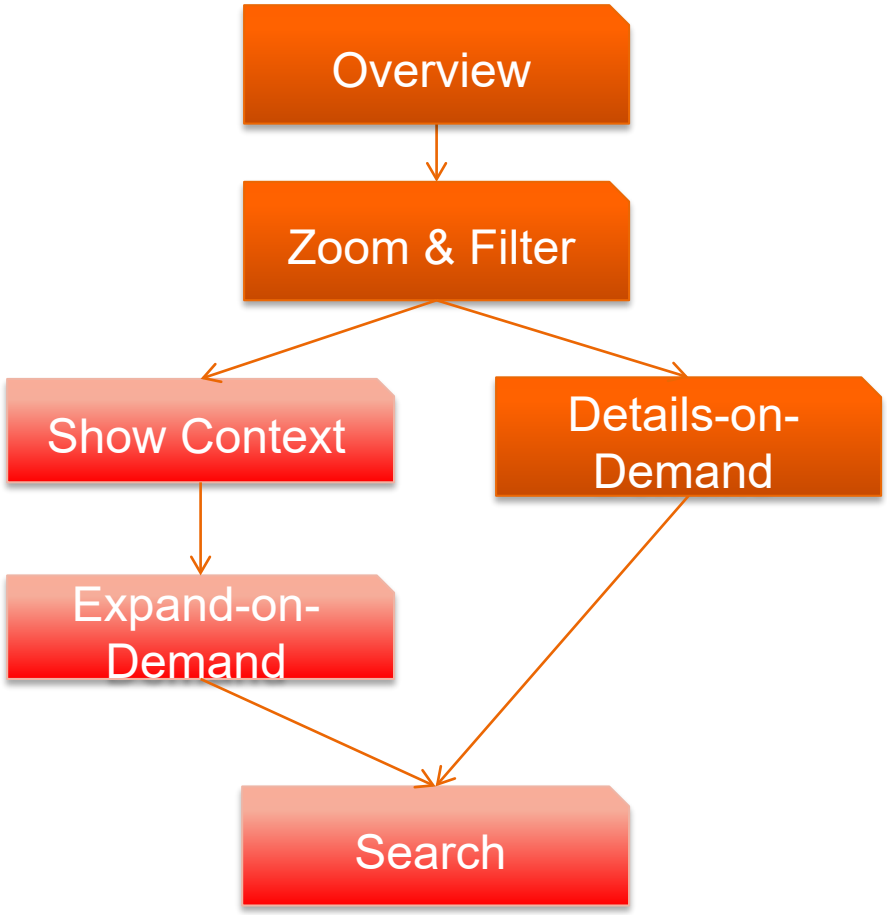
Visual Text Analytics Stages: Visual Orchestration (I)

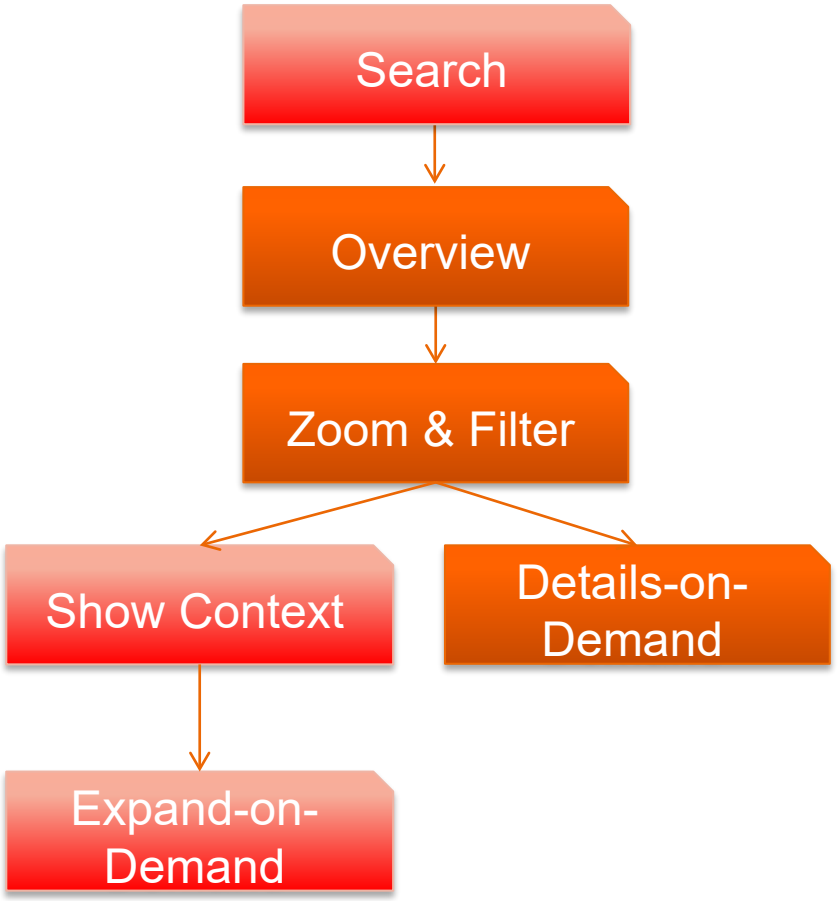
- Result of visual mapping are basically visual layouts, which represent just a set of possible interactive visualization that should enable solving analytical tasks
- With our proposed approach varieties of visual interfaces can be integrated for various tasks
- Visual Orchestration can also include corporate identity aspects or embedding constraints in existing portals and software



Visual Text Analytics Stages: Visual Orchestration (II)

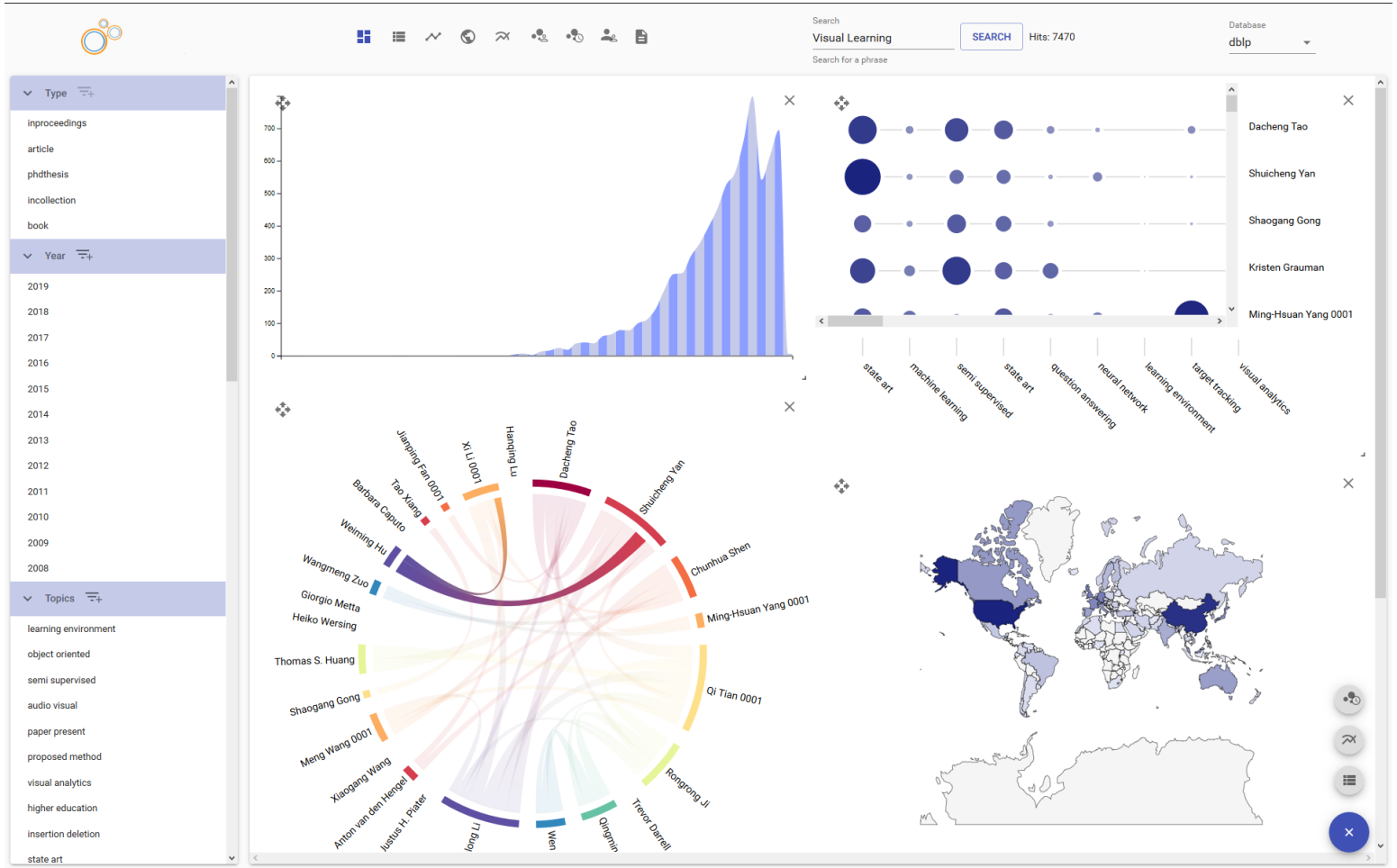
- Includes also interaction logic and procedure definitions, e.g. analytical information search





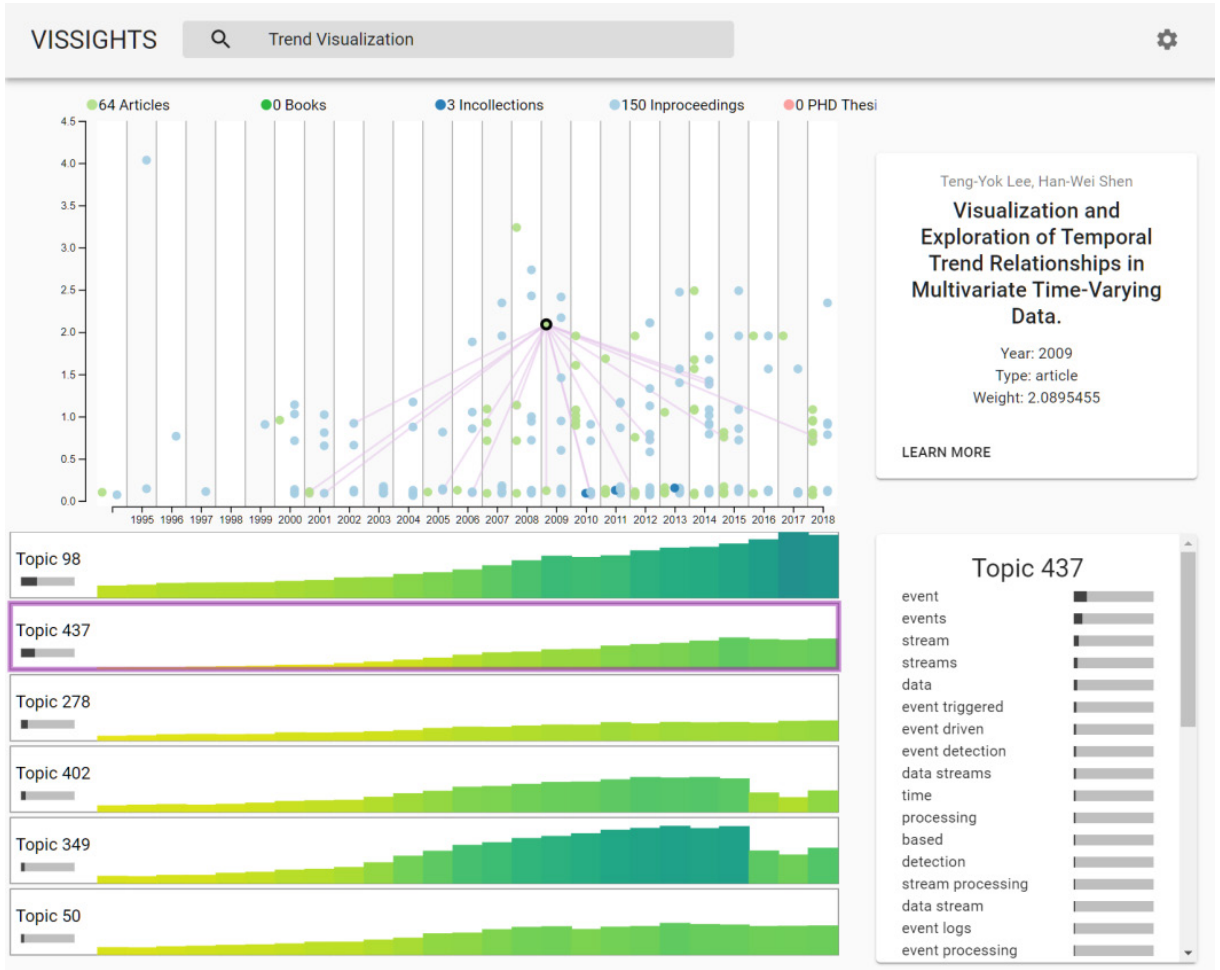
Visual Text Analytics Stages: Visual Orchestration (III)

- Example: Expert / Headhunting analysis for specialists in certain technologies or research fields



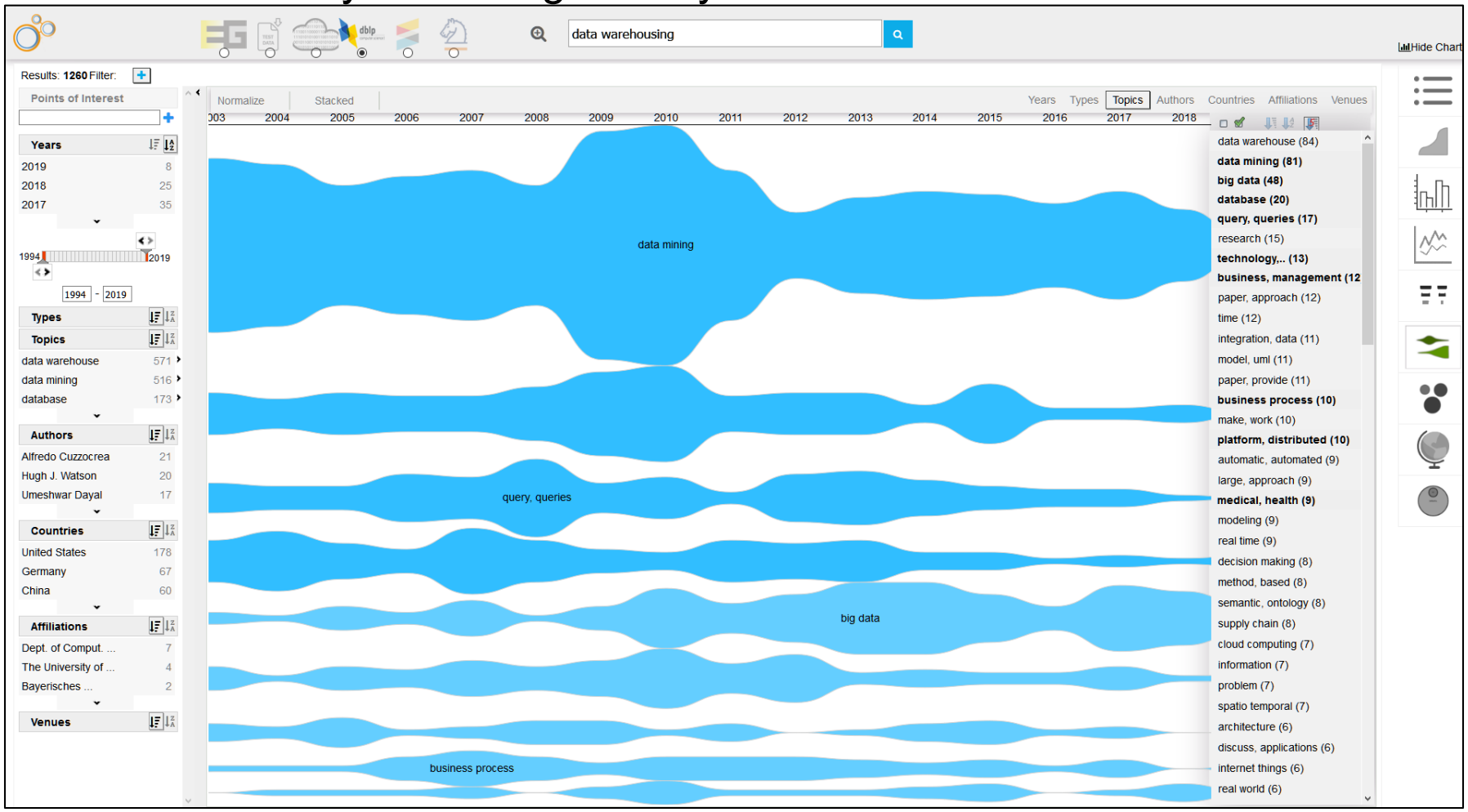
Visual Text Analytics Stages: Visual Orchestration (IV)

- Example: Document Explorer to find similar and related other documents



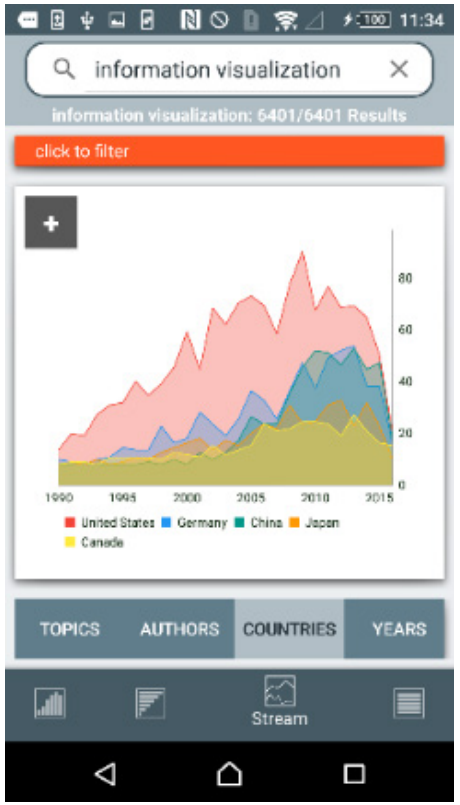
Visual Text Analytics Stages: Visual Orchestration (VI)

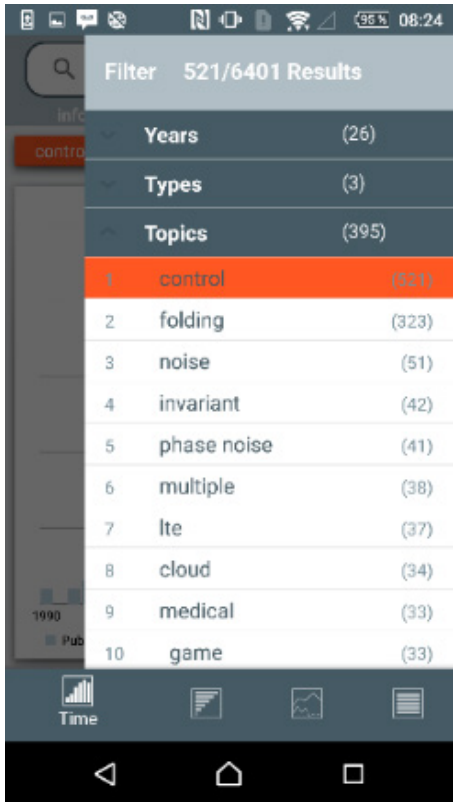
- Example: Visual Trend Analytics – insight analysis

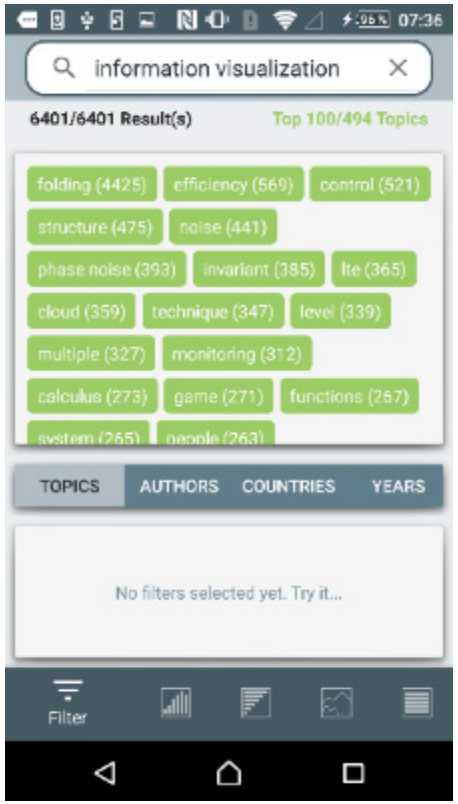


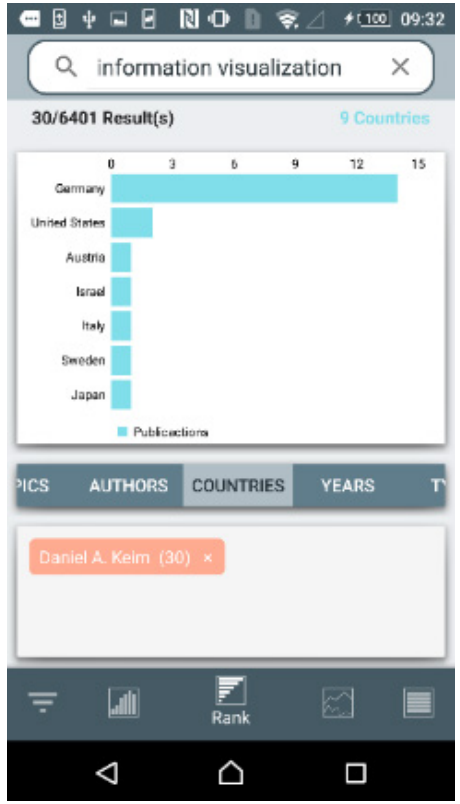
Visual Text Analytics Stages: Visual Orchestration (VI)

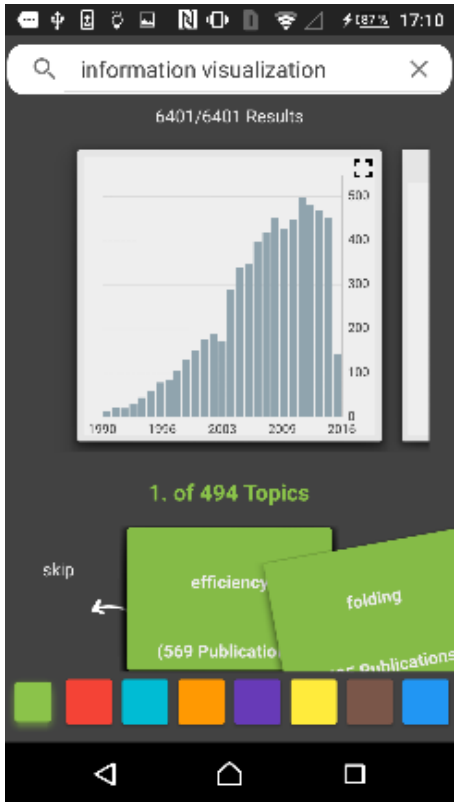
- Example: Visual Trend Analytics – mobile insight analysis





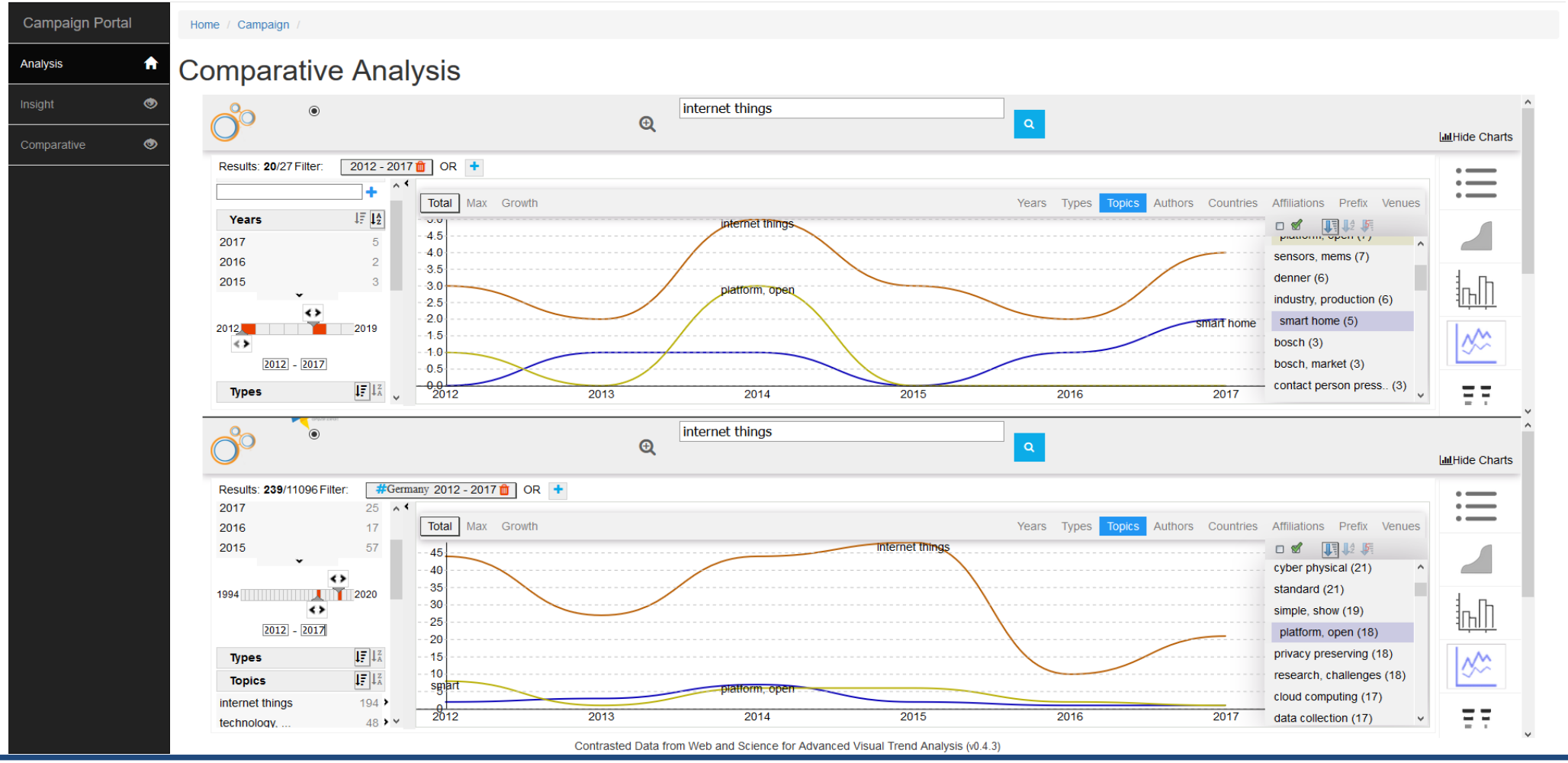






Visual Text Analytics Stages: Visual Orchestration (VII)

- Example: Visual Trend Analytics – comparative analysis & trend verification



Thank you for your attention!

Questions?

Please find a video of the system on
http://s.vis.h-da.de/video-visual-insights

More Information on
https://www.vis.h-da.de