

Global Scientific Data Infrastructures: The Role of Data Visualization

Capri, 12.5.2011



TECHNISCHE
UNIVERSITÄT
DARMSTADT



Prof. Dr. Dieter Fellner
Dr. Tatiana von Landesberger

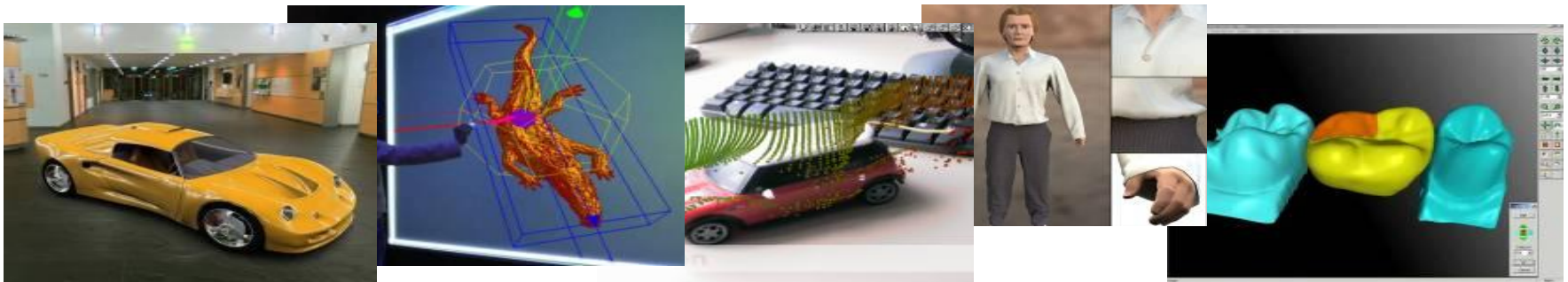
Technische Universität Darmstadt & Fraunhofer IGD
Fraunhoferstraße 5
64283 Darmstadt

Email: tatiana.von_landesberger@gris.tu-darmstadt.de
<http://www.gris.informatik.tu-darmstadt.de>



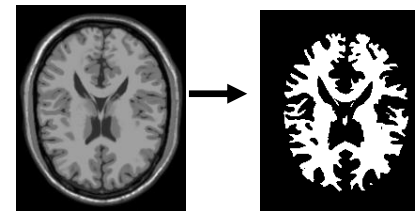
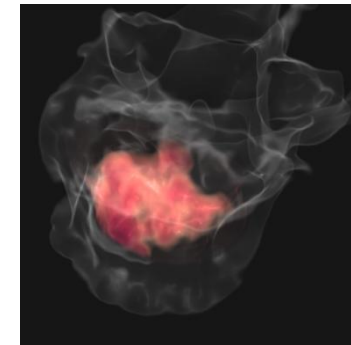
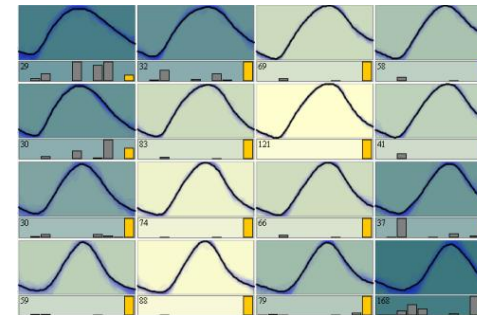
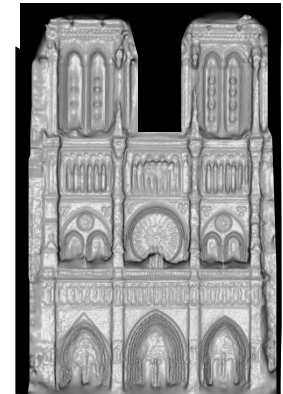
Fraunhofer Institute for Computer Graphics Research (IGD)

- **Research and Development in the area of Computer Graphics**
 - Director: Prof. D. Fellner, 180 Researchers (FTE), 17 Mio. € Turnover
 - 4 locations: Darmstadt, Rostock (Germany), Singapore and Graz (Austria)
- **Applications in Different Domains**
 - Product development (automotive, aerospace, architecture, garment industry ...)
 - Medical IT
 - Traffic Engineering and Telematics
 - Ambient Intelligence
 - Financial Sector (Information Visualization, Information Mining)
 - e-Application, e-Learning, e-Business, Digital libraries



Technische Universität Darmstadt, Germany – Interactive Graphic Systems Group

- Head: Prof. Fellner
- Research Areas
 - Visual Inference (Prof. Roth)
 - Capturing Reality (Prof. Goesele)
 - Medical Computing (Dr. Wesarg)
 - Image and Shape Analysis (Doc. Kuijper)
 - Visual Analysis and Search (Dr. v. Landesberger)
- Close cooperation with
 - Fraunhofer IGD
 - Graz University of Technology



Overview

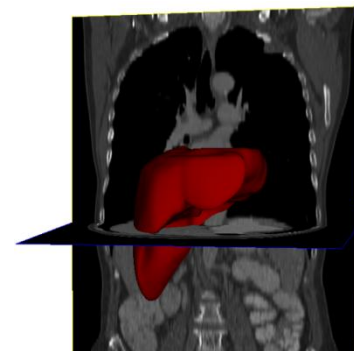
1. Introduction
2. Visual Support for Data Workflows
 1. Data acquisition
 2. Data curation
 3. Data analysis
 4. Data publishing
3. Conclusion

Overview

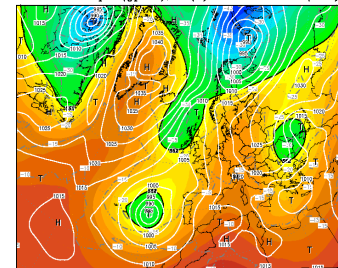
1. Introduction
2. Visual Support for Data Workflows
 1. Data acquisition
 2. Data curation
 3. Data analysis
 4. Data publishing
3. Conclusion

1. Scientific Data

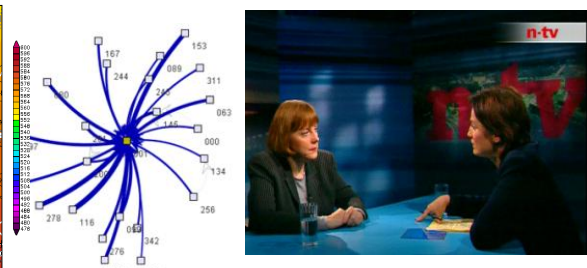
- Massive amounts data arise in research projects
 - Climate Research, Earth Observation, Nuclear and High Energy Physics, etc.
 - Biomedical research
 - ...
- Various types of data
 - Spatial and abstract data
 - Time series, multivariate data, spatial data, graphs,...
 - Multimedia data
 - Pictures, videos, 3D models,...



Int: Fri,13Oct2008 06Z Valid: Tue,17Oct2008 06Z
500 hPa Geopot.(gpm), T (C) und Bodendr. (hPa)



Daten: GFS-Modell des amerikanischen Wetterdienstes
(C) Wetterzentrale
www.wetterzentrale.de

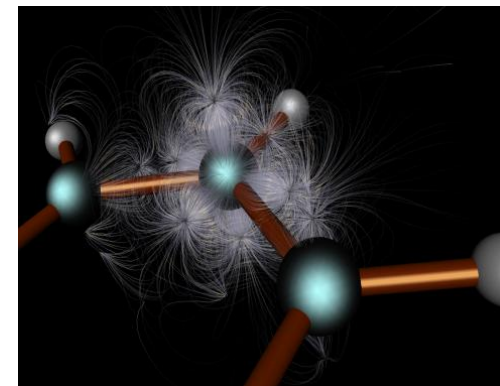
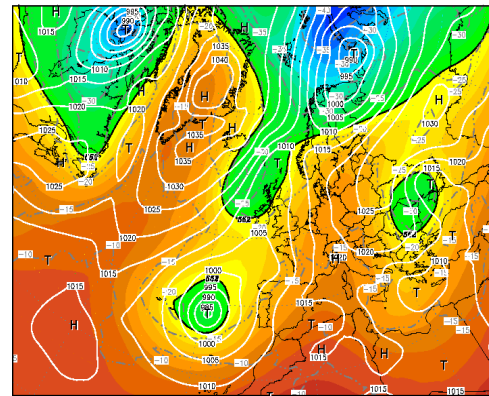
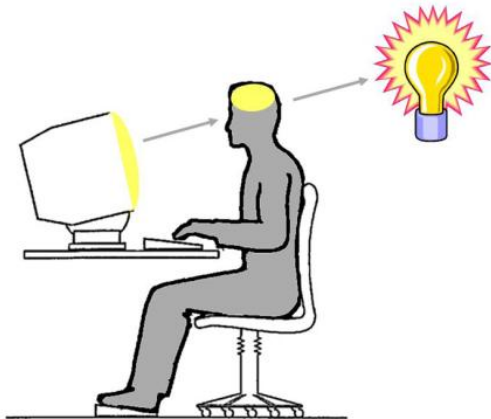
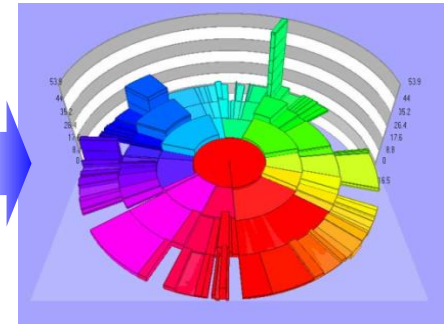


1. Introduction to Data Visualization

■ Visualization role

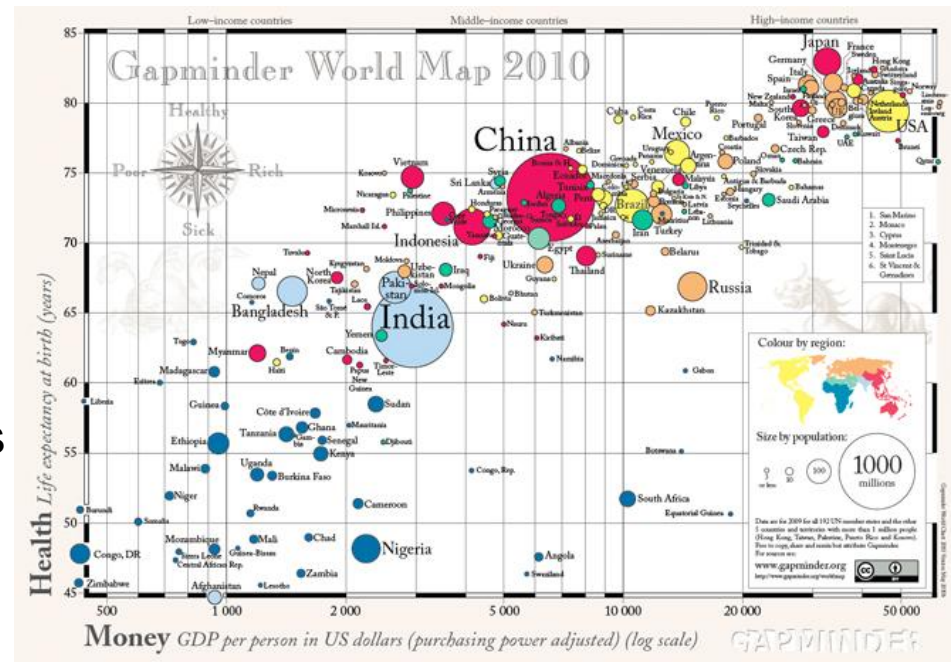
- “A picture is worth a thousand words”
 - makes it possible to absorb large amounts of data quickly
- “Using vision to think”
 - Natural human sense to identify complex / new patterns

ID	HOAX	EMAIL_GIVEN	PHONE_GIVEN	DOB_GIVEN	ID_VALID
49917	no	yes	no	17.1.1973	yes
49919	no	yes	yes	8.12.1970	no
49923	no	yes	no	3.4.1972	yes
49924	no	no	yes	1.8.1966	yes
49927	no	yes	yes	21.12.1969	yes
49928	no	no	no	20.1.1975	no
49929	no	yes	no	3.2.1978	yes
49930	no	no	no	21.7.1985	yes
49931	no	yes	no	21.5.1953	no
49933	yes	no	no	2.10.1978	yes
49935	no	no	no		no
49937	no	no	no	9.7.1965	yes
49938	no	yes	no		yes
49939	no	no	no	8.9.1964	yes
49940	no	yes	no	22.8.1964	no
49941	no	no	no	10.5.1965	no
49942	no	yes	no	28.10.1974	no
49943	no	yes	no	25.1.1971	yes
49944	no	yes	no	20.10.1974	no
49943	no	yes	no	25.1.1971	yes
49942	no	yes	no	28.10.1974	no
49943	no	yes	no	25.1.1971	yes



1. Introduction to Data Visualization

- **Visualization role**
 - “A picture is worth a thousand words”
 - makes it possible to absorb large amounts of data quickly
 - “Using vision to think”
 - Natural human sense
- **Interaction**
 - Exploration of the data from various sides
 - Focus on interesting parts



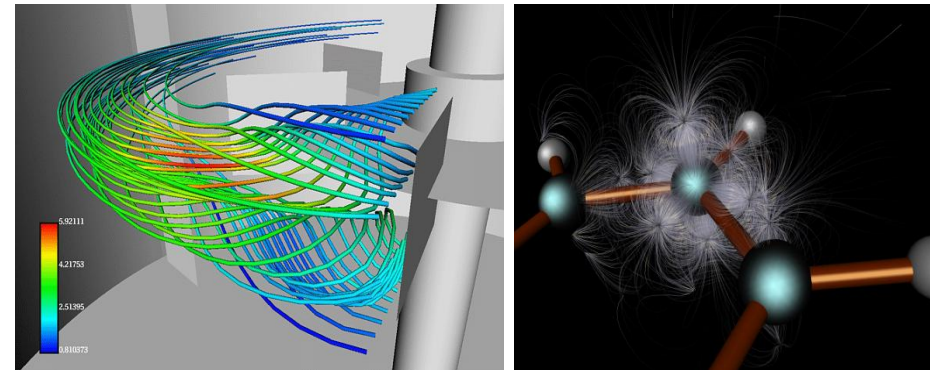
→ **Human-driven data understanding**

www.gapminder.org

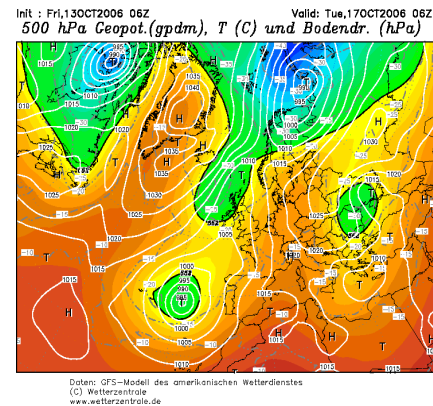
1. Data Visualization – a Traditional View

■ Visualization

- “Nice pictures” of the available data
- Present scientific results
- Understand scientific results



Can visual computing
do more?



Overview

1. Introduction
2. Visual Support for Scientific Data Workflows
 1. Data acquisition
 2. Data curation
 3. Data analysis
 4. Data publishing
3. Conclusion

2. Visual Support for Scientific Data Workflow

- Visualization as an integral part of scientific data workflows



- Visual support for **all stages** of scientific data workflow
 - Visualization as a **key way** of producing, examining and accessing scientific data
 - Scientific data includes “**visual documents**”
 - visual computing methods

2. Visual Support for Scientific Data Workflow

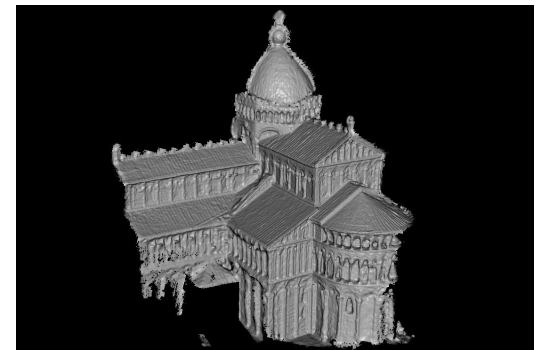
Acquire

Curate

Analyze

Publish

- Automatic data capturing (e.g., video)
 - data processing to gain more information
- Manual data collection/creation
 - Individual
 - Community-based



2. Visual Support for Scientific Data Workflow

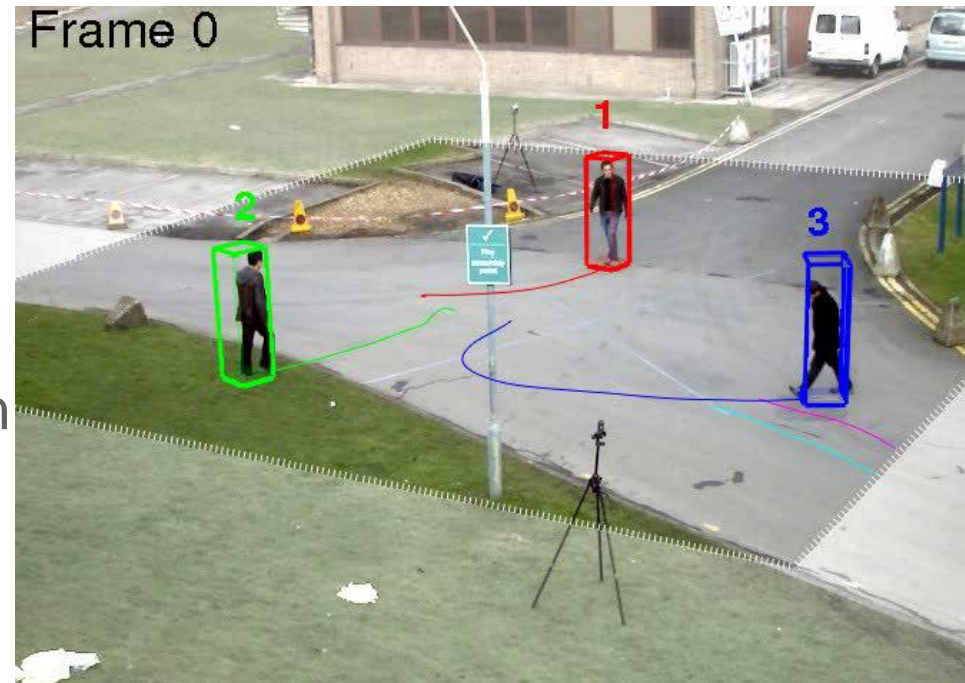
Acquire

Curate

Analyze

Publish

- Automatic data capturing (e.g., video)
 - data processing to gain more information
- Manual data collection/creation
 - Individual
 - Community-based



Human tracking in complex environments

[Andriyenko et al. 2011]

2. Visual Support for Scientific Data Workflow

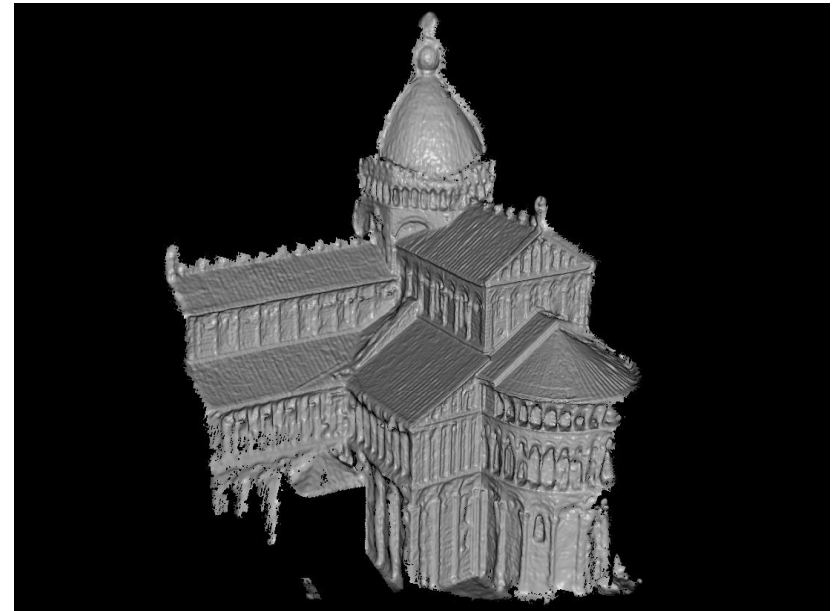
Acquire

Curate

Analyze

Publish

- Automatic data capturing (e.g., video)
 - data processing to gain more information
 - Manual data collection/creation
 - Individual
 - Community-based



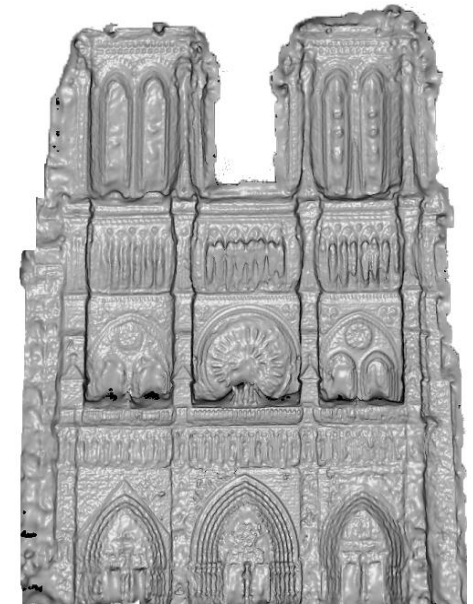
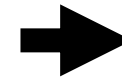
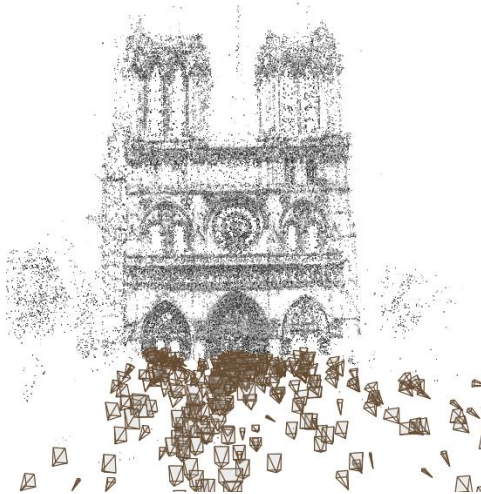
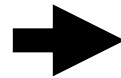
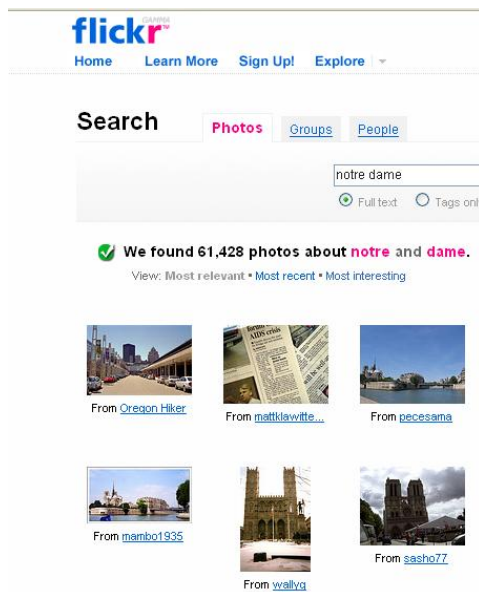
2. Visual Support for Scientific Data Workflow

Acquire

Curate

Analyze

Publish



- Reconstruction of a 3D Model using multi-view stereo [Goesele et al. 2007]
- Notre Dame facade based on 653 photos from Flickr captured by 313 persons

2. Visual Support for Scientific Data Workflow

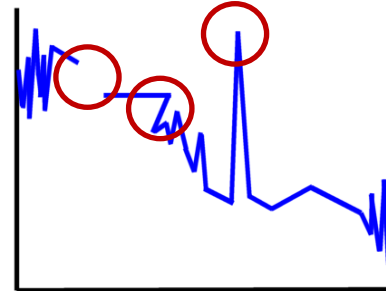
Acquire

Curate

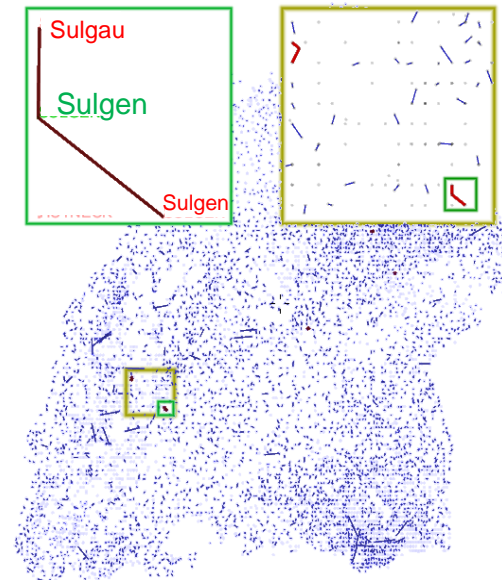
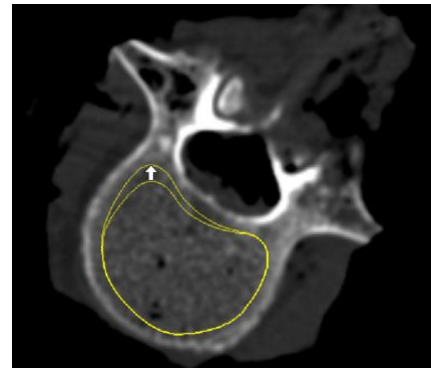
Analyze

Publish

- Show acquired data
 - Data quality information
 - Decision on further processing



- Visual data curation
 - Visual data editing



[Sanftmann et al. 2009]

[Wesarg et al.2010]

2. Visual Support for Scientific Data Workflow

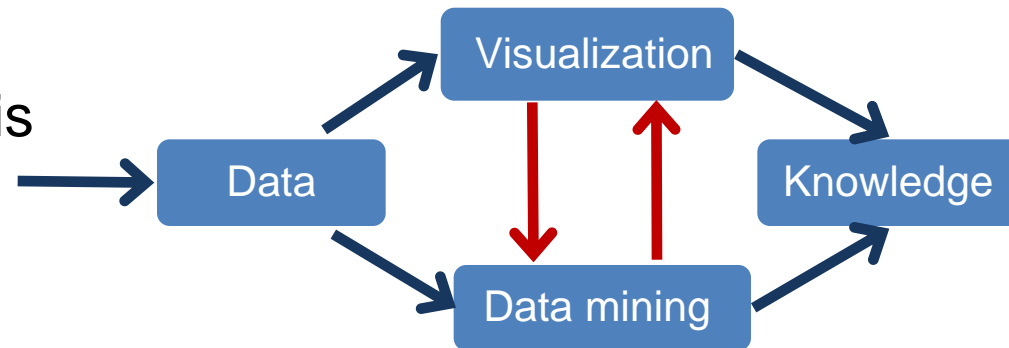
Acquire

Curate

Analyze

Publish

- Visual support for data analysis
 - Visual data mining
 - Visual feedback



- Visual support for collaborative data analysis



2. Visual Support for Scientific Data Workflow

Acquire

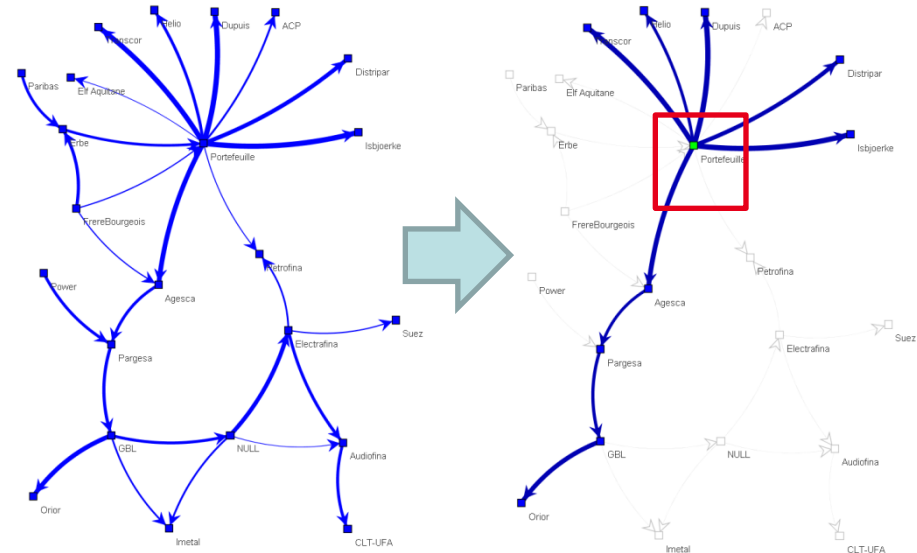
Curate

Analyze

Publish

- Visual support for data analysis
 - Visual data mining
 - Visual feedback
- Visual support for collaborative data analysis

Visual graph analysis



[von Landesberger et al. 2008]

2. Visual Support for Scientific Data Workflow

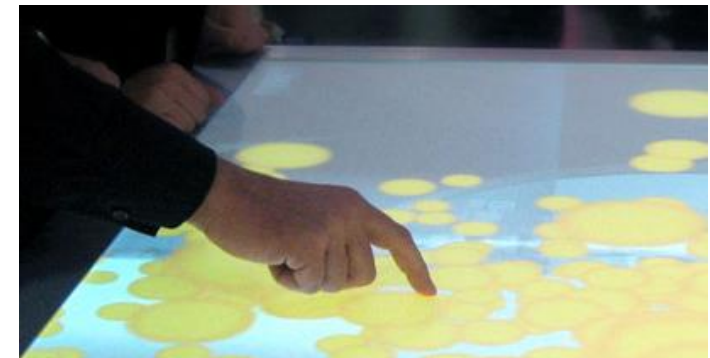
Acquire

Curate

Analyze

Publish

- Visual support for data analysis
 - Visual data mining
 - Visual feedback
- Visual support for collaborative data analysis
 - New output devices



2. Visual Support for Scientific Data Workflow

Acquire

Curate

Analyze

Publish

- Provide interactive visual interfaces for
 - Exploring the available data
 - Searching for available data
 - Define search query
 - Show search results
- Use data content not only the index structure



2. Visual Support for Scientific Data Workflow

Acquire

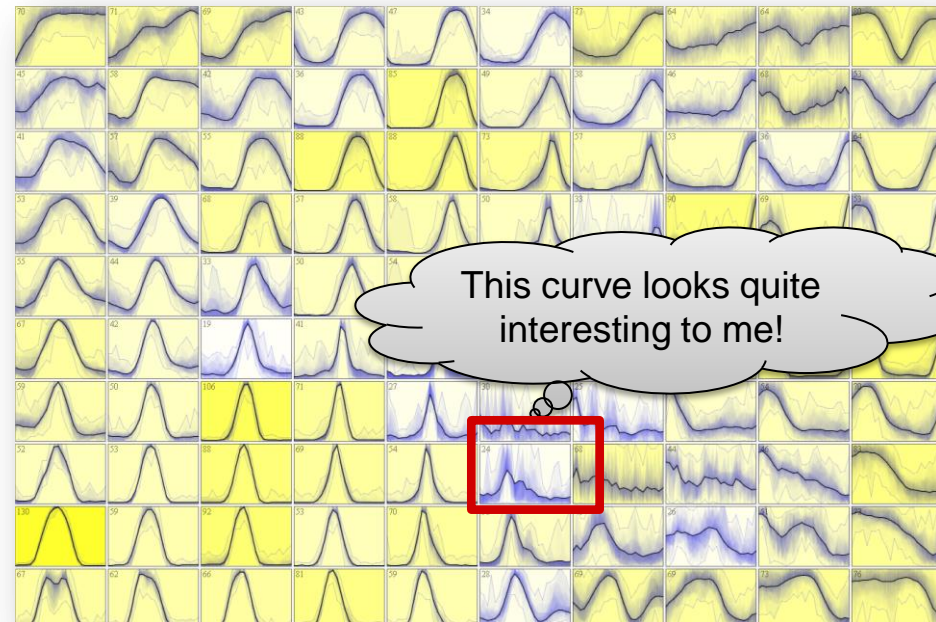
Curate

Analyze

Publish

- Provide interactive visual interfaces for
 - **Exploring** the available data
 - Visual catalogue
 - **Searching** for available data
 - Define search query
 - Show search results
- Use **data content** not only the index structure

Visual data catalogue provides overview of available scientific data



[Bernard et al. 2010]

2. Visual Support for Scientific Data Workflow

Acquire

Curate

Analyze

Publish

Metadata-based search in scientific databases

- Provide interactive visual interfaces for
 - Exploring the available data
 - Visual catalogue
 - Searching for available data
 - Define search query
 - Show search results
- Use **data content** not only the index structure

GetInfo
FIND THE WORLD OF
SCIENCE AND TECHNOLOGY

TIB | GERMAN NATIONAL LIBRARY OF
SCIENCE AND TECHNOLOGY

German National Library for all areas of
engineering as well as architecture,
chemistry, information technology,
mathematics and physics.

► Contact ► Deutsch

Search

You can conduct an interdisciplinary search in the stocks of the German National Library for Science and Technology, the German National Library of Medicine as well as of other specialised databases.

Example: (gear* OR Getriebe*) AND Hain

Advanced Search
Database selection

Limit or expand your search space by selecting one or more databases.

GetInfo is the portal for science and technology. GetInfo provides access to more than 135 million data sets from specialised databases, publishers and library catalogues.

Welfengarten 1B • 30167 Hannover • +49 (0)511 - 762 89 89

2. Visual Support for Scientific Data Workflow

Acquire

Curate

Analyze

Publish

- Provide interactive visual interfaces for

- Exploring the available data

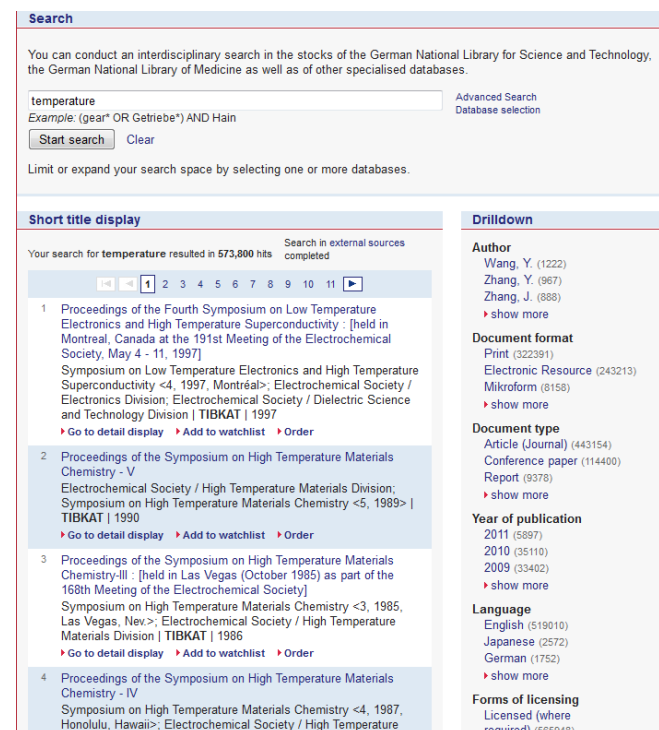
- Visual catalogue

- Searching for available data

- Define search query
 - Show search results

- Use data content not only the index structure

Metadata-based search result



Search

You can conduct an interdisciplinary search in the stocks of the German National Library for Science and Technology, the German National Library of Medicine as well as of other specialised databases.

temperature
Example: (gear* OR Getriebe*) AND Hain
Start search Clear

Advanced Search Database selection

Limit or expand your search space by selecting one or more databases.

Short title display

Your search for temperature resulted in 573,600 hits Search in external sources completed

1 Proceedings of the Fourth Symposium on Low Temperature Electronics and High Temperature Superconductivity : [held in Montreal, Canada at the 191st Meeting of the Electrochemical Society, May 4 - 11, 1997]
Symposium on Low Temperature Electronics and High Temperature Superconductivity <4, 1997, Montréal>: Electrochemical Society / Electronics Division; Electrochemical Society / Dielectric Science and Technology Division | TIBKAT | 1997
Go to detail display Add to watchlist Order

2 Proceedings of the Symposium on High Temperature Materials Chemistry - V
Electrochemical Society / High Temperature Materials Division; Symposium on High Temperature Materials Chemistry <5, 1989> | TIBKAT | 1990
Go to detail display Add to watchlist Order

3 Proceedings of the Symposium on High Temperature Materials Chemistry-III : [held in Las Vegas (October 1985) as part of the 168th Meeting of the Electrochemical Society]
Symposium on High Temperature Materials Chemistry <3, 1985, Las Vegas, Nev.>: Electrochemical Society / High Temperature Materials Division | TIBKAT | 1986
Go to detail display Add to watchlist Order

4 Proceedings of the Symposium on High Temperature Materials Chemistry - IV
Symposium on High Temperature Materials Chemistry <4, 1987, Honolulu, Hawaii>: Electrochemical Society / High Temperature

Drilldown

Author
Wang, Y. (1222)
Zhang, Y. (967)
Zhang, J. (888)
show more

Document format
Print (322391)
Electronic Resource (243213)
Mikroform (8158)
show more

Document type
Article (Journal) (443154)
Conference paper (114400)
Report (9376)
show more

Year of publication
2011 (5897)
2010 (35110)
2009 (33402)
show more

Language
English (519010)
Japanese (2572)
German (1752)
show more

Forms of licensing
Licensed (where applicable) (656018)

2. Visual Support for Scientific Data Workflow

Acquire

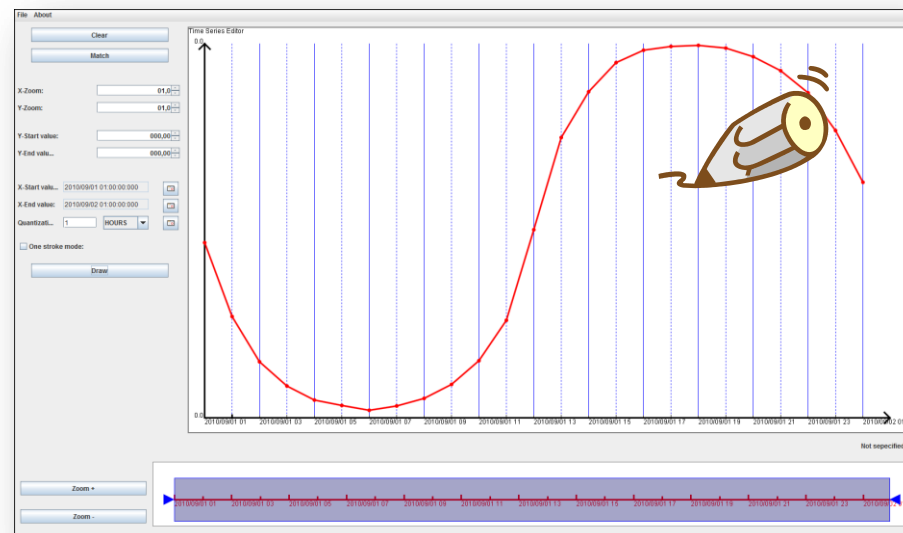
Curate

Analyze

Publish

- Provide interactive visual interfaces for
 - Exploring the available data
 - Visual catalogue
 - Searching for available data
 - Define search query
 - Show search results
- Use **data content** not only the index structure

Content-based search
- visual interactive query definition



[Bernard et al. 2011]

2. Visual Support for Scientific Data Workflow

Acquire

Curate

Analyze

Publish

- Provide interactive visual interfaces for

- Exploring the available data

- Visual catalogue

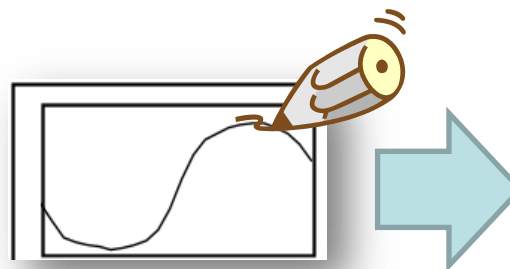
- Searching for available data

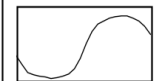
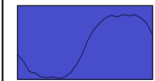
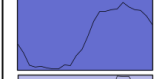


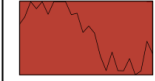
- Define search query

- Show search results

- Use **data content** not only the index structure

Content-based search
result visualization



	Query Specification
	Result 1 Description: ... Metadata: ...
	Result 2 Description: ... Metadata: ...
	Result 3 Description: ... Metadata: ...
	Result 4 Description: ... Metadata: ...
⋮	
	Result n Description: ... Metadata: ...

[Bernard et al. 2011]

2. Visual Support for Scientific Data Workflow

Acquire

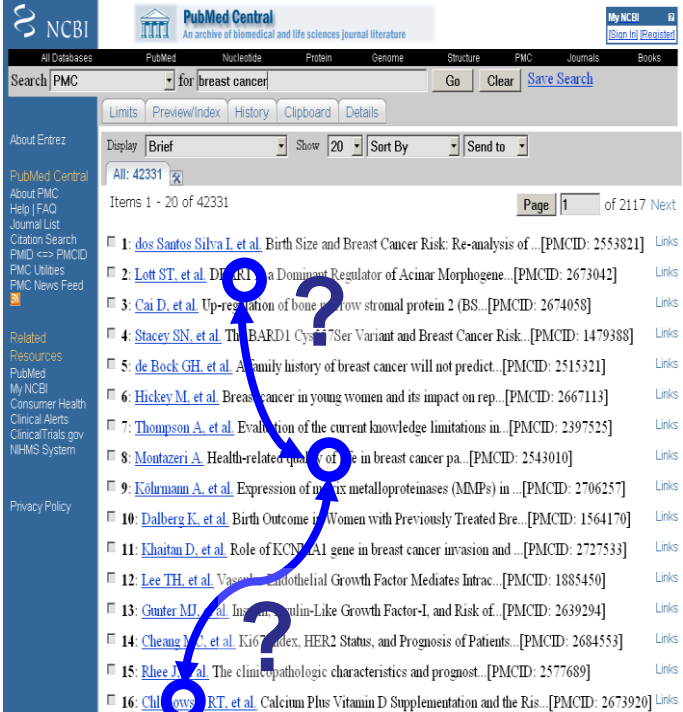
Curate

Analyze

Publish

- Provide interactive visual interfaces for
 - Exploring the available data
 - Visual catalogue
 - Searching for available data
 - Define search query
 - Show search results
- Use **data content** not only the index structure

Search result presentation without visible connections between documents



The screenshot shows the PubMed search results for 'breast cancer'. The search results are listed in a table with columns for item number, title, and links. A blue line with question marks highlights the lack of visible connections between the documents.

Item	Title	Link
1	dos Santos Silva I. et al. Birth Size and Breast Cancer Risk: Re-analysis of ... [PMCID: 2553821]	Links
2	Lott ST. et al. DEAF1 is a Dominant Regulator of Acinar Morphogene... [PMCID: 2673042]	Links
3	Cai D. et al. Up-regulation of bone marrow stromal protein 2 (BS... [PMCID: 2674058]	Links
4	Stacey SN. et al. The BARD1 Cys77Ser Variant and Breast Cancer Risk... [PMCID: 1479388]	Links
5	de Bock GH. et al. A family history of breast cancer will not predict... [PMCID: 2515321]	Links
6	Hickey M. et al. Breast cancer in young women and its impact on rep... [PMCID: 2667113]	Links
7	Thompson A. et al. Evaluation of the current knowledge limitations in... [PMCID: 2397525]	Links
8	Montazeri A. Health-related quality of life in breast cancer pa... [PMCID: 2543010]	Links
9	Köhmann A. et al. Expression of matrix metalloproteinases (MMPs) in... [PMCID: 2706257]	Links
10	Dalberg K. et al. Birth Outcome in Women with Previously Treated Bre... [PMCID: 1564170]	Links
11	Khaitan D. et al. Role of KCNQ1 gene in breast cancer invasion and ... [PMCID: 2727533]	Links
12	Lee TH. et al. Vascular Endothelial Growth Factor Mediates Intrac... [PMCID: 1885450]	Links
13	Gunter MJ. et al. Insulin-Like Growth Factor-I, and Risk of... [PMCID: 2639294]	Links
14	Cheang Y. et al. Ki67 Index, HER2 Status, and Prognosis of Patients... [PMCID: 2684553]	Links
15	Rhee J. et al. The clinicopathologic characteristics and prognost... [PMCID: 2577639]	Links
16	Chilcows RT. et al. Calcium Plus Vitamin D Supplementation and the Ris... [PMCID: 2673920]	Links

2. Visual Support for Scientific Data Workflow

Acquire

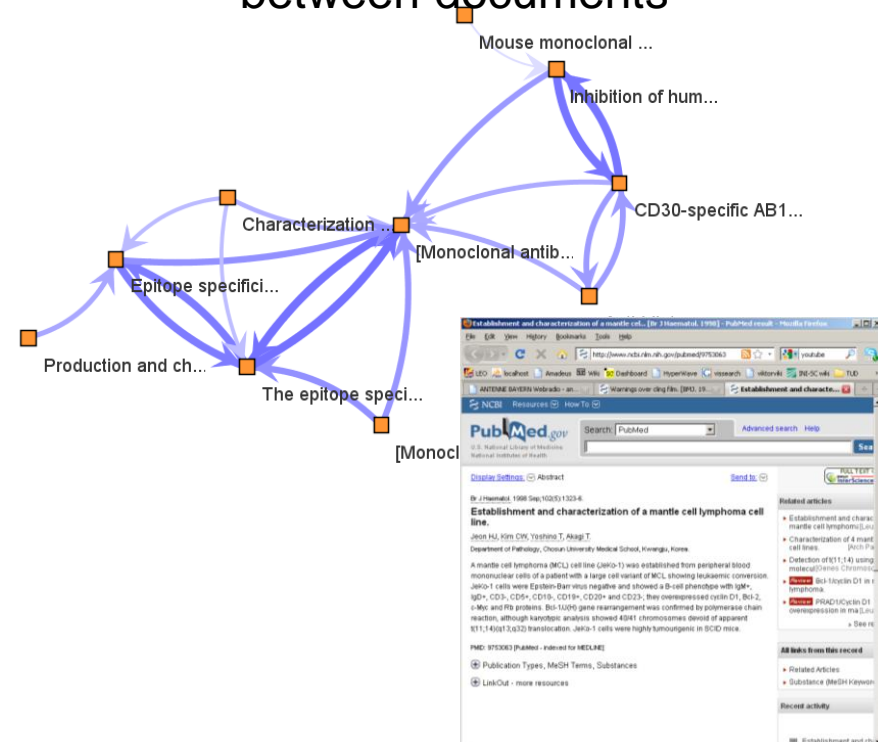
Curate

Analyze

Publish

Exploration of connections between documents

- Provide interactive visual interfaces for
 - Exploring the available data
 - Visual catalogue
 - Searching for available data
 - Define search query
 - Show search results
- Use data content not only the index structure

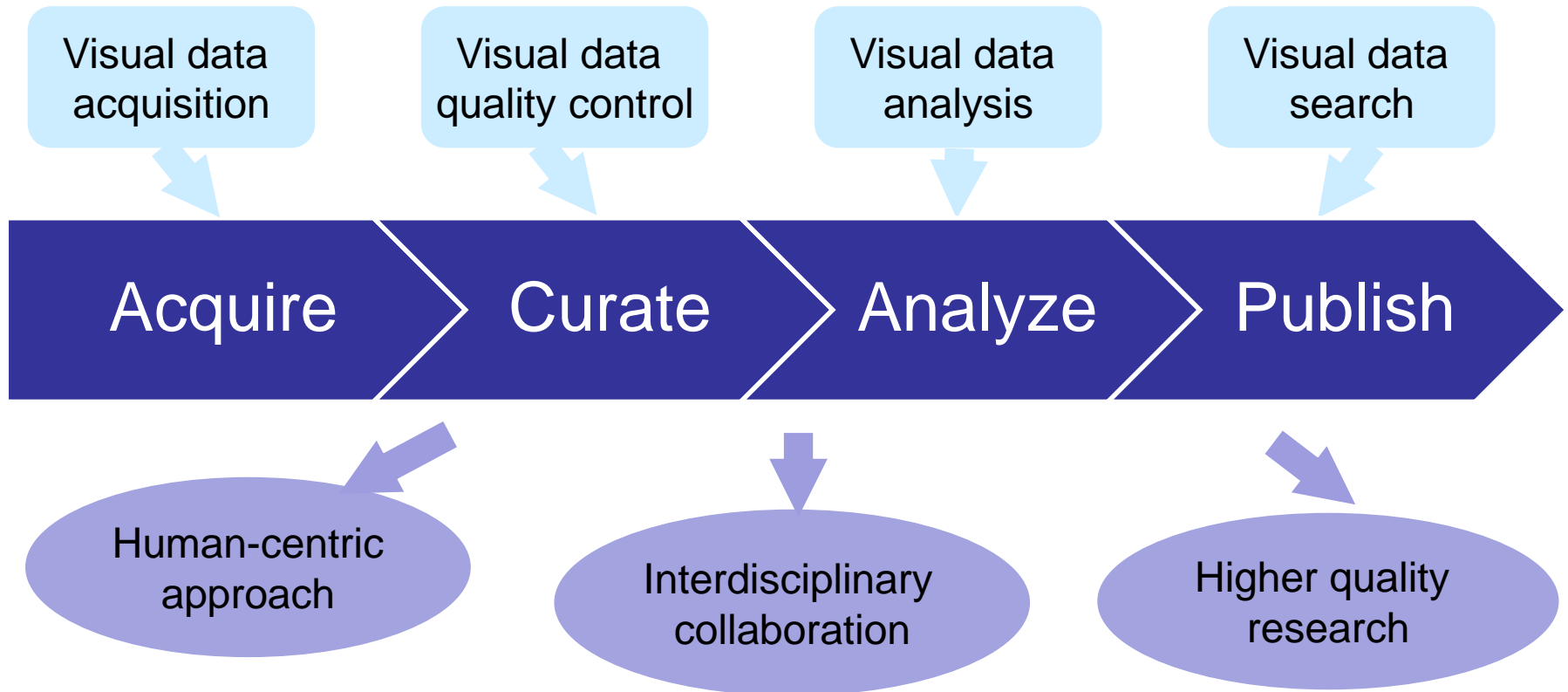


Contents

1. Introduction
2. Visual Support for Data Workflows
 1. Data acquisition
 2. Data curation
 3. Data analysis
 4. Data publishing
3. Conclusion

3. Conclusion

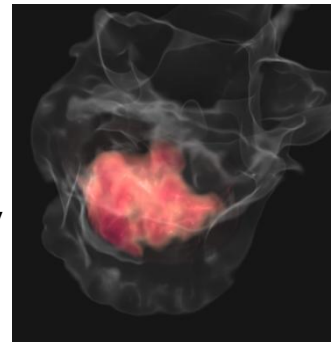
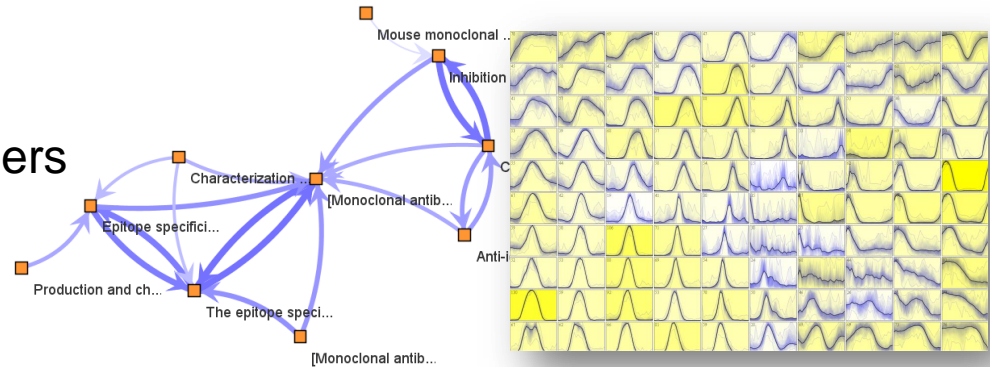
Scientific Data Infrastructures and Visual Computing:
high potential for mutual benefits



3. Challenges



- Visual solutions
 - Close cooperation with data users and data providers
- Input data
 - Common data standards
 - Common data semantics
- System interoperability
- Visualization scalability and interactivity



End

Thank you for your kind attention

