10241 Executive Summary
Information Visualization
— Dagstuhl Seminar —

Andreas Kerren[1], Catherine Plaisant[2], and John T. Stasko[3]

[1] Linnaeus University, SE
andreas.kerren@lnu.se
[2] University of Maryland - College Park, US
plaisant@cs.umd.edu
[3] Georgia Institute of Technology, US
stasko@cc.gatech.edu

Abstract. Information Visualization (InfoVis) focuses on the use of visualization techniques to help people understand and analyze data. While related fields such as Scientific Visualization involve the presentation of data that has some physical or geometric correspondence, Information Visualization centers on abstract information without such correspondences.
The aim of this seminar was to bring together theoreticians and practitioners from the field with a special focus on the intersection of InfoVis and Human-Computer Interaction. To support discussions that are related to the visualization of real world data, researchers from selected application areas also attended and contributed. During the seminar, working groups on eight different topics were formed and enabled a critical reflection on ongoing research efforts, the state of the field, and key research challenges today. This document summarizes the event.

Keywords. Information Visualization, Visualization, Data Visualization, Collaboration, Display Technologies, Human-Computer Interaction

1 Introduction

Information Visualization (InfoVis) is a research area that focuses on the use of visualization techniques to help people understand and analyze data. While related fields such as Scientific Visualization involve the presentation of data that has some physical or geometric correspondence, Information Visualization centers on abstract information without such correspondences, i.e., it is not possibile to map this information into the physical world in most cases. Examples of such abstract data are symbolic, tabular, networked, hierarchical, or textual information sources. The ever increasing amount of data generated or made available every day confirms the urgent need for suitable InfoVis tools. As prerequisite for building a successful visualization, InfoVis combines several aspects of different

research areas, such as Computer Graphics, Graph Drawing, Data Mining, Information Design, Cognitive Psychology, and Human-Computer Interaction (HCI), among many others.

One main goal of this second Dagstuhl Seminar on Information Visualization was to bring together theoreticians and practitioners from the addressed research areas with a special focus on the intersection of InfoVis and Human-Computer Interaction. Many researchers are active in both of these fields, thus the seminar was especially attractive to those people. To support discussions that are related to the visualization of real world data, we also invited researchers from selected application areas, such as Bioinformatics and the GeoSciences.

1.1 Seminar Topics

The following themes were discussed during the seminar:

- **Collaboration within Information Visualization:** Collaboration is becoming increasingly important in InfoVis and it can occur in collocated or distributed locations and be synchronous or asynchronous. The development of novel interaction techniques, suitable visual representations, social components, and special display technologies are only a sample of the important issues that were discussed.
- **The Importance of Interaction:** The representational aspects of information visualization often receive the most focus, but the interactive capabilities of an InfoVis system are just as important. The interactive dialog between the human user and the visualization system allows the user to gain new perspectives on the data and ask questions not initially present. Accordingly, what makes for an effective, interactive system? What do powerful interaction capabilities add to a specific visualization? Which interaction techniques best accomplish different analytical goals?
- **The Influence of Display Technologies on InfoVis:** Large displays with high-resolution are one possibility to present increasingly large data sets. On the other hand, small-scale displays, especially in the context of mobile phones or PDAs, are becoming more available and important. The size and the type of a display have a large influence to the user interaction and visual representation within InfoVis.
- **InfoVis for the Masses:** In addition to the typical single-analyst, deep-dive analytical nature of InfoVis, a growing focus of research is examining how to allow large numbers of people to produce, view, and discuss information visualizations as well. This topic emerged during the first Dagstuhl Seminar on InfoVis but insufficient time was available to thoroughly discuss it.
- **Multimodal User Interaction:** Multimodality is referred to as being the combination of several modalities, such as the visual, sensory, or auditory modality. The extension of InfoVis with other modalities is not very well explored, and many open questions exist. For example: when we should use multiple modalities to present a particular data set and which ones? Or, how should they be combined?

- **Prior Knowledge of Users:** Visualization tools should be sensitive to the prior knowledge of their users with respect to both application-specific knowledge and visualization knowledge. Choosing optimal levels of visual abstraction and finding effective visual metaphors are two challenges in this area, where the term "optimal" clearly depends on the user. A visualization system should adjust to the user's needs.
- **InfoVis Aesthetics:** An aesthetically appealing visualization can be more successful than a more efficient but unappealing visualization that presents the same data. A closer look into such phenomena could improve our only vague understanding of this intersection of InfoVis and the Visual Arts, and thus, it could improve the success of information visualization techniques in practice.

The seminar allowed attendees to critically reflect on current research efforts, the state of field, and key research challenges today. Participants also were encouraged to demonstrate their system prototypes and environments relevant to the seminar topics. As a result, further topics emerged and were the focus of deeper discussions:

- **Visualization of Text and Documents:** Textual data is widespread and of importance to many people. While the visualization of text and documents is often treated as a general problem, the problem and suitable solutions can differ strongly depending on the target audience of the visualization and the task support desired.
- **Comparison in Information Visualization:** The visual comparison of data is important for many analysis tasks. A deeper discussion of the underlying models, goals, and challenges helped to facilitate a better understanding of this issue.
- **Data Wrangling:** The transformation of data into alternate forms to enable analysis is an important step in the broad analytical process. Related issues include data quality, how to handle missing data, data cleaning, and normalization, among many others.
- **Analysis Process:** The true analysis process is often significantly different than many assumptions widespread in the information visualization literature. Who are the key stakeholders in this process? How do analysts usually work? What is important to them?

2   Participation and Program

48 people from 11 countries participated in this seminar. Most attendees were from the US and from Germany, but others came from Canada, Australia, Israel, and other European countries, as shown in Figure 1.

The program aimed to generate lively discussions. Presenters were asked not to give talks solely focusing on their own specific research. Instead, the group began the seminar by collecting important themes and then selecting eight specific topics to be discussed in later breakout sessions. The resulting
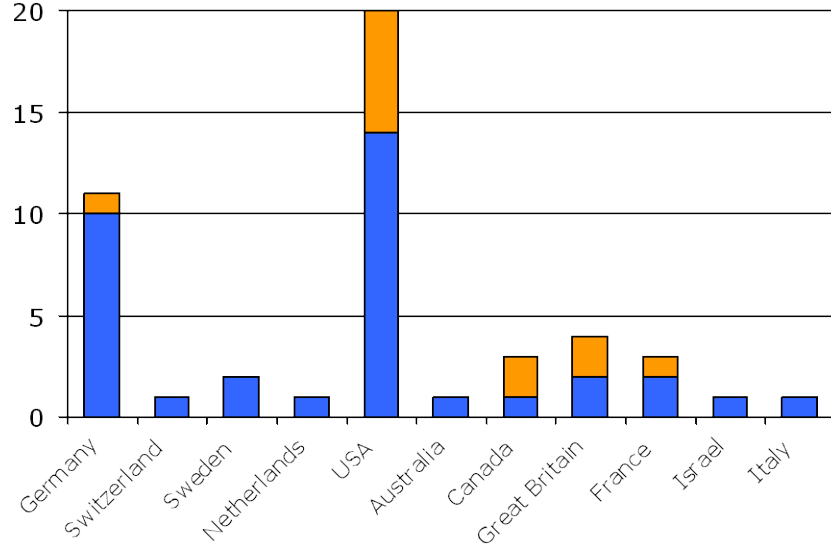
Fig. 1. Attendee Statistics of Seminar #10241. Blue colored bars represent male and orange colored bars female participants.

Table 1. Final structure of the seminar. Topics of Tuesday/Wednesday were Collaboration, Aesthetics, Display Technologies and Text/Documents. Topics of Thursday/Friday were Comparison, Data Wrangling, Interaction and Analysis Process.

Monday	Tuesday	Wednesday	Thursday	Friday
Welcome & Organization	Provocation Talks (4 topics)	Demo Session	Provocation Talks (4 topics)	Group Discussion (writing up)
Self Presentations	Group Discussions (4 topics)	Group Discussions (writing up)	Group Discussions (4 topics)	Sum Up & Closing Remarks
Topic Discussion	Presentations (4 topics)	Social Event (Trier)	Presentations (4 topics)	
Future of the InfoVis Conference	Reporting Session (4 topics)		Reporting Session (4 topics)	

eight group discussions were preceded by provocation talks, and accompanied by scientific talks, a demo session and report sessions given to the entire audience. Table 1 provides an overview of the final seminar schedule.

3   Discussions and Outcome of the Seminar

As already mentioned above, the program included breakout sessions on eight specific topics, i.e, eight working groups discussed one topic at time; four groups in two parallel sessions. Before the actual group discussions took place on a day, a common session for all attendees with four provocation talks occurred, one talk per topic. The idea of these provocation talks was to generate issues for further discussion, provoke the attendees to think broadly about the issues involved, and to help the attendees decide which working group to attend.

Furthermore, a small set of speakers were selected to give a scientific talk on a theme that fit with the group topic. In sum, ten talks were given during the seminar:

- *Tim Dwyer:* 2D or not 2D? Was 3D Information Visualization just a fad?
- *Michael Gleicher:* Aesthetics (or why Andrew wanted me to show pictures of molecules)
- *Petra Isenberg:* Collaborative Information Visualization
- *Stephen North and Jean-Daniel Fekete:* Display Technologies
- *Jeffrey Heer:* Using Topic Models to Visualize the Evolution of Academic Departments
- *John Stasko:* Adding Computational Analysis to Jigsaw
- *Niklas Elmqvist:* GraphDice: A System for Exploring Multivariate Social Networks
- *Jean Scholtz:* The Impact of Prior Knowledge on the Design, Use and Evaluation of Visualization Tools
- *Harald Reiterer:* New forms of Human-Computer Interaction for Visualizing Information
- *Jessie Kennedy:* Old & New Experiences of BioVis

The content of these talks, given for all seminar attendees after the lunch break, raised further key issues and helped to prepare the group for the final reporting session. This last session of the day reviewed results from the individual working groups.

3.1   Demo Session

Seven tools and/or prototype implementations were demonstrated during the demo session of the seminar:

- *Tim Dwyer and Natalie Henry-Riche:* Sets in graphs.
- *Jason Dykes:* Timely information for citizens: place survey prototype.
- *Jing Yang:* PIGVIS – a multidimensional visualization approach to visualizing graphs.

- *Manuel Freire-Moran:* ManyNets – visualizing many networks simultaneously.
- *John Stasko:* Jigsaw – supporting investigative analysis through interactive visualization.
- *Matt Ward:* Visualizing collections of heterogeneous models – two prototype applications.
- *Yarden Livnat:* Exploration using dynamic tag clouds and loosely coordinated multiple views.

3.2   Working Group (WG) Results

This section describes results from each of the eight working groups and identifies the attendees contributing to each group. The names of those people who wrote the summary report for the working group discussions are underlined.

WG: Aesthetics in InfoVis (*Provocation talk:* Andrew Vande Moere and Sheelagh Carpendale. *Group members:* Andreas Kerren, Andrew Vande Moere, Enrico Bertini, <u>Helen Purchase</u>, Jack Van Wijk, Jason Dykes, Michael Gleicher, Sheelagh Carpendale, Stephan Diehl, Tim Dwyer).

In creating visualizations of data, an aesthetic choice needs to be made, even if the choice is a "minimalist" one. We concluded that (1) "Aesthetics" is a loaded term; perhaps we should talk instead about "visual style and interaction"; (2) we should take a "multiple aesthetics" perspective, emphasizing that aesthetic design is an explicit choice; (3) it is important that the process of aesthetic design be described and justified; (4) we should consider the "truth" of the data and whether the aesthetic design choice helps or hinders the representation of the "truth" of the data; (5) evaluation in graphic literacy is key to the development of good visual style. While we did not answer the broad question of "What is Aesthetics and why do we need it?", we made progress in identifying key issues in the development of Information Visualization visual styles.

WG: Collaboration in InfoVis (*Provocation talk:* Niklas Elmqvist and Petra Isenberg. *Group members:* Daniel Cernea, Falk Schreiber, Hans Hagen, Heidi Lam, Jean Scholtz, Kwan-Liu Ma, Niklas Elmqvist, <u>Petra Isenberg</u>).

The working group discussed the emerging research area of Collaborative Information Visualization (CIV), the differentiation between CIV and CSCW (Computer-Supported Collaborative Work), and the vision for CIV. Several key questions were identified, such as: Who are the collaborators? How can we communicate the benefits of collaborative visualization tools in relation to thought and work processes of our end users? How can tools encourage people to improve their work processes? How can we structure research in collaborative visualization? Collaborative visualization requires research on two main aspects: (1) Technical issues: interfaces, toolkits to make collaboration effortless, and different solutions across the space/time matrix. (2) Social aspects of group work: how much overlap exists in the space/time matrix? We need to understand how these spaces relate and differ.

WG: Display Technology (*Provocation talk:* Bongshin Lee, Jonathan C. Roberts, and Nathalie Henry-Riche. *Group members:* Jessie Kennedy, Dominique Brodbeck, Jean-Daniel Fekete, Nathalie Henry-Riche, Bongshin Lee, Achim Ebert, Charles D. Hansen, Stephen North, Heidi Lam, Jonathan C. Roberts, T.J. Jankun-Kelly, Paolo Buono, Harald Reiterer).

Group members first identified open questions to start discussion on applications for large scale displays, properties, and working style. Some issues emerged including the question of whether managing display technologies is "just" a technical matter, whether the the InfoVis reference model needs to be updated for large display visualization (likely not), and the importance of InfoVis representations being usable from iPhones to wall-size displays. Key challenges remain in *Input Management.* For example, what are the right input techniques and devices for the different display technologies? How can we support visualizations at different scales? And how can we support effective collaboration on these devices?

WG: Text and Documents (*Provocation talk:* Christopher Collins. *Group members:* Adam Perer, Catherine Plaisant, Chris Weaver, Christopher Collins, Daniel Keim, Frank van Ham, John Stasko, Hendrik Strobelt, Ilir Jusufi, Jeffrey Heer, Manuel Freire-Moran, Matt Ward, Yarden Livnat, Tamara Munzner).

The discussion centered around enumerating pitfalls, successes, and challenges in visualizing text and documents. In these discussions the group found that text and document visualization is too often treated as a general problem, when in fact, the problem (and appropriateness of the solution) differs significantly by the target audience of the visualization and the task support desired. Thus, the group produced a categorized list of text visualization user populations and tasks to help frame issues in the area. Several broader questions facing the InfoVis research community also were identified: How can visualization researchers not familiar with natural language processing (NLP) techniques gain proficiency in that important related area? What are the best NLP techniques and their relative accuracies? How can we evaluate systems for text and document visualization when the tasks supported by text visualization often lack a ground truth to evaluate against?

WG: Analysis Process (*Provocation talk:* Jean Scholtz and Tamara Munzner. *Group members:* Christopher Collins, Enrico Bertini, Heidi Lam, Jason Dykes, Jean Schlotz, Petra Isenberg, Sheelagh Carpendale, Tamara Munzner).

The group found that those who have engaged deeply with analysts now perceive that actual analysis processes often significantly differs from the models and assumptions widespread in the information visualization literature. The group concluded that in order to find common ground across different domains we need specific data points from observational studies. Such studies are currently being published in external venues, but not in InfoVis forums. Participants noted that a big barrier to publication at our own venues is the (mis)perception of reviewers that such studies need to be generalized with design implications in order

to be worthwhile as a standalone paper. The group argued that applying the standards of formal experimental studies with quantitative measures, or even qualitative measures, is not always appropriate for observational studies. To advance this agenda, three action items must be addressed. First, we should refine the current taxonomy of keywords used to match papers and reviewers to include Observational Studies as a specific methodology. Second, we should write a "call for arms" position paper advocating observational studies in an InfoVis context. Third, we should bring the external literature on this topic to the attention of the InfoVis community through a Beyond Infovis Showcase at VisWeek.

WG: Comparison (*Provocation talk:* Michael Gleicher. *Group members:* Adam Perer, Charles Hansen, Fabian Beck, Ilir Jusufi, Jing Yang, Jonathan C. Roberts, Kwan-Liu Ma, Michael Gleicher, Stephan Diehl, Tim Dwyer).

A model for comparison in information visualization was discussed that maps the design space into several categories: juxtaposition, separation, & small multiples, which use the memory of a user to make visual connections; overlay &superposition that allow the user to make visual connections, as well as fusion & difference objects, which are derived representations and use algorithms to provide the correlation and differences between data. This model highlights several challenges that offer multiple opportunities for further research: including depiction of context, incorporating heterogeneous data, complexity, and comparison across different forms and partiality. Orthogonal to this model is the understanding of who performs comparison: whether the effort is put to the user or whether it is an analytic process of the computer.

WG: Data Wrangling (*Provocation talk:* Catherine Plaisant and Jeffrey Heer. *Group members:* Bongshin Lee, Catherine Plaisant, Chris Weaver, Dominique Brodbeck, Frank van Ham, Helen Purchase, Hendrik Strobelt, Jeffrey Heer, Jessie Kennedy, Paolo Buono, Stephen North).

The working group defined data wrangling as a process of iterative data exploration and transformation to enable analysis. The output of the wrangling process is not just data, it is a set of transformations and an understanding of data issues. Hereby, the data triage, exploration, cleaning and integration should be integrated and iterative. Suitable visual representations allow us to see data quality issues and can be an input device for transformations. The difference between the semantic (e.g., reformat data to fit a visualization tool) and syntactic cleaning (e.g., how missing data is imputed) is often unclear. Interesting open questions in this context include: Are there benefits/pitfalls to combining different levels of operations together? And could they be meaningfully separated?

WG: Interaction (*Provocation talk:* Matt Ward and T. J. Jankun-Kelly. *Group members:* Achim Ebert, Andrew Vande Moere, Daniel Cernea, Jack van Wijk, Jean-Daniel Fekete, John Stasko, Matt Ward, Niklas Elmqvist, T.J. Jankun-Kelly, Yarden Livnat).

After a triage of well-known InfoVis systems, the group identified a set of characteristics of systems with "good" interactions. Especially important for InfoVis seems to be fluidity and animated transitions, no interference with visual representations, immediacy, as well as visible and minimal state. Often, more interaction power means more complexity and more challenge. Thus, one goal is to increase the power without increasing the complexity in reality. These thinking processes lead to a set of questions which have to be answered: If we simplify the interface too much, will we lose the power we need? How do we interact with graphs and complex data better? Do we really need cross-cutting and complex queries? Or what makes for an effective interaction in information visualization?

WG: Multimodal, Tangible and Multi-Touch Interfaces (This group was formed as a subgroup of the interaction working group. *Group members:* Dominique Brodbeck, Harald Reiterer, Jessie Kennedy, Jonathan C. Roberts, Paolo Buono).

The multi-touch/multimodal subgroup discussed aspects of multimodal, tangible and multi-touch interfaces for Information Visualization. There is a need in InfoVis to provide interfaces that allow the information to be fused and integrated over time, i.e., there is much benefit in allowing the user to control the input of the visualization with two modalities. Such environments enable collaboration too. Several research questions came up that are listed in the following. (1) Theory: What perception issues are important when different modalities are integrated? (2) Design: What new interaction capabilities are possible and useful to interact with and control information visualizations? How do we develop InfoVis solutions that integrate, fuse or blend several different technologies together? How do the modalities fuse together to present a coherent and unambiguous story? (3) Perception: Which technique is best for integrating different modalities into devices? Is it possible to reinforce the information presented in one modality or do they contradict? Are there design models/patterns to help?

3.3   Outcome

The organizers and participants decided to organize a special issue of the *Information Visualization Journal (IVS)*[4] published by Palgrave Macmillan. The possibility of this special issue was confirmed by the Editor-in-Chief before the end of the seminar. Working groups have been invited to submit an article building on their discussions and findings, and writing is underway. The papers are to be submitted by November 15th 2010, with a planned publication date of October 2011 (volume 10, issue 4).

Topics Remained Open Not all the topics identified during the seminar could be addressed in the working groups and might be considered for a future Dagstuhl Seminar. They include the following:

---

[4] http://www.palgrave-journals.com/ivs/index.html

- Scalability
- Perception/Cognitive Science
- Multidimensionality vs. Graph Visualization
- Impact
- Applications
    - Bioinformatics, Health Care, Software Engineering
- Theory
- Evaluation
- Education
- Time/Trend
- Challenges for the next ten Years
- Uncertainty
- InfoVis Toolkits
- Visual Analytics

Acknowledgments

We would like to thank all participants of the seminar for the lively discussions during the seminar as well as the scientific directorate of Dagstuhl Castle for giving us the possibility of organizing this event. Ilir Jusufi gathered the abstracts for the abstract collection and the talks of all presenters. These talks can be found on the materials site of the seminar. In addition, many attendees agreed to take notes during the breakout sessions. These notes were the basis for writing this executive summary and are also available for download on the Dagstuhl web page of the seminar. Last but not least, the seminar would not have been possible without the great help of the staff of Dagstuhl Castle. We would like to acknowledge all of them and their assistance.