VISUAL ANALYTICS
RESEARCH WITH IMPACT

CATHERINE PLAISANT
HUMAN-COMPUTER INTERACTION LAB
UNIVERSITY OF MARYLAND

VAC consortium
May 2011
NEW HCI Masters Program!

Beginning Fall 2011, the University of Maryland's iSchool will offer a Masters degree in Human-Computer Interaction which includes internships, work with HCIL researchers, interdisciplinary classes, and more! Learn more about this exciting new program, or visit Maryland's iSchool website for application information for the Fall 2011 semester.

HCIL 28th Symposium
May 25th - 26th, 2011

Registration now open!
• Approaches?
  – Try MILCs studies (multi-dimensional, in-depth, longitudinal, case studies)
• Resources?
  – IRB for MILCs
  – Review paper on 7 Scenarios for evaluations
  – VA Benchmark Repository
    • Document YOUR use of the benchmarks
    • Continue support → new challenge competitions
EVALUATION METHOD

- **Multi-Dimensional**
- **In-depth**
- **Long-term**
- **Case studies**

MILCs

Refinements of tools
Report on usage and achievements of expert users’ professional goals,

Shneiderman, B., Plaisant, C.
Strategies for Evaluating Information Visualization Tools: Multi-dimensional In-depth Long-term Case Studies
Proc. of Beliv’06 Workshop
HIERARCHICAL CLUSTERING EXPLORER

- 5 then 3 expert users
- 6 weeks
- 1 hour per week
- Help given
- Improvements to tools

→ Scientific discoveries
  (e.g. a strong association between a specific gene and body composition)

Seo and Shneiderman., IEEE TVCG 06
GUIDELINES FOR PRACTICAL MILCs

- Identify 3-5 users
- Document current methods
- Determine what would determine professional success (Discovering a new drug? A scientific paper?)
- Schedule of observations (long sessions, then shorter but always be flexible)
- Instrument tool + extras → Logs
- Train / Modify
- Avoid bias - Encourage to use the best tool
- Document successes and failures
- Close collaboration with 1 group of users
- Later (when software advanced enough) MILC studies with 3-5 more
• Approaches?
  – Try MILCs studies
    http://portal.acm.org/citation.cfm?id=1168158

• Resources?
  – IRB for MILCs
  – Review paper on 7 Scenarios for evaluations
  – VA Benchmark Repository
    www.cs.umd.edu/hcil/varepository
    • Document YOUR use of the benchmarks
    • Continue support → new challenge scenarios
Seven Guiding Scenarios for Information Visualization Evaluation

Heidi Lam  Enrico Bertini  Petra Isenberg  Catherine Plaisant  Sheelagh Carpendale

Abstract—We take a new, scenario based look at evaluation in information visualization. Our seven scenarios, evaluating visual data analysis and reasoning, evaluating user performance, evaluating user experience, evaluating environments and work practices, evaluating communication through visualization, automated evaluation of visualizations, and evaluating collaborative data analysis were derived through an extensive literature review of over 800 visualization publications. These scenarios are described through their goals, the types of questions they embody and illustrated through example studies. Through this broad survey and the distillation of these scenarios we make two contributions. One, we encapsulate the current practices in the information visualization research community and, two, we provide a different approach to reaching decisions about what might be the most effective evaluation of a given information visualization. For example, if the research goals or evaluative questions are known they can be used to map to specific scenarios, where practical existing examples can be considered for effective evaluation approaches.

Index Terms—Information visualization, evaluation

1 INTRODUCTION

Researchers and practitioners in the field of information visualization (infovis) have long identified the need to evaluate visual data representations, interaction techniques, and visualization systems. Yet, the difficulty of conducting these infovis evaluations remains a common topic. For instance, in addition to the general evaluations challenges of choosing evaluation questions, methods, and correctly executing them, the infovis focus on data and its exploratory analysis processes pose still further challenges, since both make it difficult for visualization researchers and practitioners to find the most appropriate approaches to achieve their evaluation goals. Another aspect of the difficulty is the lack of literature guidelines—while some guidelines are available to create and analyze visualization systems, and to evaluate visualizations, these two sets of literature are disparate as discussions on evaluation are mostly “structured as an enumeration of methods with focus on how to carry them out, without prescriptive advice for when to choose between them.” ([51, p.1], author’s own emphasis). We extend this by taking a different tack—we offer advice...
SEVEN SCENARIOS

- Extensive literature review of over 800 visualization publications.
- Evaluating environments and work practices
- Evaluating visual data analysis and reasoning
- Evaluating user performance
- Evaluating user experience
- Evaluating communication through visualization
- Automated evaluation of visualizations
- Evaluating collaborative data analysis
Benchmarks contains **datasets and tasks**, as well as materials describing the **uses** of those benchmarks (the results of analysis, contest entries, controlled experiment materials etc.) Most benchmarks contain ground truth described in a solution provided with the benchmark, allowing accuracy metrics to be computed.

**List of Benchmarks**

<table>
<thead>
<tr>
<th>Provenance</th>
<th>Benchmarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>VAST Challenge 2010</td>
<td><strong>Grand Challenge 2010</strong>&lt;br&gt;<strong>Text Records - Investigations into Arms Dealing</strong>&lt;br&gt;<strong>Hospitalization Records</strong>&lt;br&gt;<strong>Genetic Sequences</strong></td>
</tr>
<tr>
<td>VAST Challenge 2009</td>
<td><strong>Badge and Network Traffic</strong>&lt;br&gt;<strong>Grand Challenge 2009</strong>&lt;br&gt;<strong>Video Analysis</strong>&lt;br&gt;<strong>Social Network and Geospatial</strong></td>
</tr>
</tbody>
</table>

Total number of benchmarks: 25
Total number of uses: 254
Total number of papers: 48
LaBRI, INRIA Bordeaux Sud-Ouest  
**Award:** Representation of uncertainty in rules & in visualization  
2009

Leonard - EAKOS 2009  
**Award:** Good Use of competing hypotheses  
2009

University of Maryland-Mindlab  
2009

Palantir Technologies  
2009

Stanfield Systems Incorporated - VIM Toolkit  
2009

SZTAKI-graphVis  
**Award:** Good Analytical debrief  
2009

Taburiente/UMD - Slice and Dice  
2009

TCS-InnovationLab, Delhi  
2009

Beijing University of Posts and Telecommunications-HumanDynamicVis  
2009

University of Konstanz  
**Award:** Good analytic debrief  
2009

VIS(US) Stuttgart  
**Award:** Innovative analytic tool  
2009

Peking University-VisualMen  
2009

---

**Papers**

Total number of papers: 7

**Citation**

Simonetto, Paolo; Koenig, Pierre-Yves; Zaidi, Faraz; Archambault, Daniel; Gilbert, Frederic; Phan-Quang, Trung-Tien; Mathiaut, Morgan; Lambert, Antoine; Dubois, Jonathan; Sicre, Ronan; Brulin, Mathieu; Vieux, Remy; Melancon, Guy; Solving the traffic and flutter challenges with tulip, Visual Analytics Science and Technology, 2009. VAST 2009. IEEE Symposium on, vol., no., pp.247-248, 12-13 Oct. 2009


Bosch, Harald; Heinrich, Julian; Muller, Christoph; Hoferlin, Benjamin; Reina, Guido; Hoferlin,
Authors and Affiliations:
Dr. Peter Bak, University of Konstanz, bak@dbvis.inf.uni-konstanz.de [PRIMARY contact]
Stefan Moritz Koch, University of Konstanz, stefan.2.koch@uni-konstanz.de [author, analyst]
Simon Butscher, University of Konstanz, simon.butscher@uni-konstanz.de [author, analyst]

Tool(s):
In order to solve the challenge we used a combination of tools. To preprocess the data, we relied on a small PHP script. To visualize the network data we used Pajek, a popular network analysis program [http://pajek.imfm.si/doku.php]. Also Pajek has a lot of functionality we used only a small part of it, mainly the force directed layout algorithms, the degree filter, measures like centrality, or walks with limited length. Beside these tools, we developed a small java tool to help us analyze the network data according to the constraints of the network structure that were given.

Video:

[Video.wmv]

ANSWERS:

MC2.1: Which of the two social structures, A or B, most closely match the scenario you have identified in the data?
A

MC2.2: Provide the social network structure you have identified as a tab delimited file. It should contain the employee, one or more handler, any middle folks, and the localized leader with their international contacts. What are the Flitter names of the persons involved? Please identify only key connections (not all single links for example) as well as any other nodes related to the scenario (if any) you may have discovered that were not described in the two scenarios A and B above.

[Flitter.txt]

MC2.3: Characterize the difference between your social network and the closest social structure you selected (A or B). If you include extra
Benchmark Details

Title: **Social Network and Geospatial**  Provenance: **VAST Challenge 2009**

Description:
Embassy employees are known to have use the social networking/micro-blogging tool, Flitter, to communicate with colleagues and friends. The Flitter network may provide a connection to a criminal ring that may have recruited an employee. We have been provided with Flitter data that we may analyze.


Ground truth present: **yes**
Solution present: **yes**  Link to solution
Creation date: 2009
Datatype: table

Contact Information:
Georges Grinstein, University of Massachusetts Lowell
Catherine Plaisant, HCIL, University of Maryland
Jean Scholtz, Pacific Northwest National Laboratory
Mark Whiting, Pacific Northwest National Laboratory

---

**Benchmark Uses**

Total number of uses: **17**

<table>
<thead>
<tr>
<th>Used by:</th>
<th>Description</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>University of Ulm and derivo-VIScover</td>
<td><strong>Award:</strong> Novel visualization of effect of rule application</td>
<td>2009</td>
</tr>
<tr>
<td>DRDC</td>
<td></td>
<td>2009</td>
</tr>
<tr>
<td>Georgia Institute of Technology</td>
<td></td>
<td>2009</td>
</tr>
</tbody>
</table>
Login

Email * 
Password *

Submit

To request a login please mail us at plaisant@cs.umd.edu or sreddy@umd.edu, in return we will email you, your login information.

A service of SEMVAST Project
Hosted by the Human Computer Interaction Lab of the University of Maryland
• Approaches?
  – Try MILCs studies
    http://portal.acm.org/citation.cfm?id=1168158

• Resources?
  – IRB for MILCs
  – Review paper on 7 Scenarios for evaluations
  – VA Benchmark Repository
    www.cs.umd.edu/hcil/varepository
    • Document YOUR use of the benchmarks
    • Continue support → new challenge scenarios