Multi-Perspective Information Visualization of Linked Data for Humanities Research

First Advisor: Prof. Dr. Claudia Müller-Birn
Second Advisor: Prof. Dr. Lutz Prechelt
Supervisor: Jesse Benjamin
Structure

➔ Motivation

➔ Chapters
◆ Analysis
◆ Design
◆ Technical implementation
◆ Evaluation

➔ Summary and future work
Motivation

Linked Data (LD) is about using the Web to connect related data that wasn’t previously linked, or using the Web to lower the barriers to linking data currently linked using other methods. More specifically a term used to describe a recommended best practice for exposing, sharing, and connecting pieces of data, information, and knowledge on the Semantic Web using URIs and RDF.

[Retrieved from LD community - linkeddata.org, 2018]
Problem definition

➔ Limitless possibilities but also many difficulties [Dadzie and Rowe, 2011]

➔ The need for more intuitive and easy to interact tools is urgent [Dadzie and Pietriga, 2016]
Goal

➔ Easier and intuitive interaction with LD

➔ Visual representation and analytics of complex data
Approaches

➔ An Iterative User Centered Design

➔ Visual Analytics
Analysis

Research and comparison of related tools

Mainly 3 types of visualization tasks [Pelzer, 2017]
- Exploring
- Data analysis
- Querying
## Analysis - Research and comparison

<table>
<thead>
<tr>
<th></th>
<th>EXPLORE</th>
<th>DATA ANALYSIS</th>
<th>Querying</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LOD Live</td>
<td>ResXplorer</td>
<td>SynopsViz</td>
</tr>
<tr>
<td>LMDB Dataset</td>
<td>✔</td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>Visual overview</td>
<td>✔</td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>Multiple perspectives</td>
<td>✔*</td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>Coordinated views</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Data overview</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Detail on demand</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Zoom in/out</td>
<td>✔*</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Visual Analytics</td>
<td></td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>Approach</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Visual Querying</td>
<td>✔</td>
<td>✔*</td>
<td>✔</td>
</tr>
<tr>
<td>Reusable output</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Target-Lay-users</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Target-Tech-users</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>RDF URIs</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
</tbody>
</table>
Analysis - Research and comparison
Analysis - User interviews

➔ Survey

◆ Demographics:
  ● 14 responses
  ● Average age: 28
  ● 14% PhD students, 26% Master’s or Bachelor’s graduates and 60% currently pursuing Bachelor's degree
  ● Study programs: Film studies and Art history, Computer science, Medicine as well as Physiotherapy sciences.
  ● Technical skills: All over Level 6 (Scale 1-10 With 10 being the highest)
Analysis - User interviews

→ Survey
◆ Questions

Can you find the name of a **certain** movie or actor if you only know the name of one of the film’s characters?

Who is the favorite actor or actress of a **certain** director?
Design - Concept

➔ Easy access over the web
➔ See brief explanation of the app and its features on demand
➔ Display visualizations enabling visual analytics approach
➔ Switch between coordinated visualizations
➔ Require no additional training
➔ Appropriate and accurate data delivery
Design

➔ Low-fidelity prototyping
Technical Implementation - Technology

- React
- Node.js
- GitHub
Technical Implementation - Development
Avoid Information overload and user confusion

Buttons for switching between perspectives

Informing users
Technical Implementation - Development

→ High-Fidelity Prototype - DEMO VIDEO
Evaluation

➔ Usability and user experience tests

◆ Demographics:
  • Total number of participants: 10
  • Students currently pursuing bachelor's or master’s degree
  • Average age: 25
  • Study programs: Computer science, Electrical engineering, Medicine and Physiotherapy sciences
  • Technical skills: Level 7 or more (Scale 1-10, with 10 being the highest)
Evaluation

➔ Usability and user experience tests

◆ Goal
  ● Can users utilize the selected tools to interact with LD
  ● Do some problems or issues arise
  ● Improve performance
  ● Get feedback
Evaluation

➔ Usability and user experience tests

◆ Criteria
  ● Effectiveness
  ● Efficiency
  ● Subjective satisfaction
Evaluation

➔ Usability and user experience tests

◆ Tasks and execution

- **Querying**: Find movie information based on the subject of a movie
- **Data analysis**: Find favorite actor or actress of a certain director
- **Data analysis**: Find all actors/actresses as well as number of times they have appeared together in a movie with a certain actor
- **Exploration**: Find movie information based on film character name.
Evaluation - Quantitative results

Success Distribution By Task

- Task 1: 100%
- Task 2: 100%
- Task 3: 100%
- Task 4: 90%

Average Time On Task By Task

- Task 1: 341.18 seconds
- Task 2: 236.4 seconds
- Task 3: 185.3 seconds
- Task 4: 162.4 seconds
Evaluation - Quantitative results

![Average Mouse Clicks By Task](chart1)

- **Task 1**: Lod Live: 135.9, Rhizomer: 47.5, LMDB-Viz: 48.5
- **Task 2**: Lod Live: 123.2, Rhizomer: 52.9, LMDB-Viz: 56.3
- **Task 3**: Lod Live: 71, Rhizomer: 59.1, LMDB-Viz: 54
- **Task 4**: Lod Live: 110.5, Rhizomer: 56.2, LMDB-Viz: 54

![Average Mouse Movement By Task](chart2)

- **Task 1**: Lod Live: 4583.49, Rhizomer: 4138.46, LMDB-Viz: 4773.37
- **Task 2**: Lod Live: 23382, Rhizomer: 38511.43, LMDB-Viz: 38657.95
- **Task 3**: Lod Live: 22420, Rhizomer: 39022.64, LMDB-Viz: 33000.88
- **Task 4**: Lod Live: 27675.1, Rhizomer: 39144.8, LMDB-Viz: 38504.5
Evaluation - Qualitative results

<table>
<thead>
<tr>
<th>Scale</th>
<th>Mean</th>
<th>Comparison to benchmark</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pragmatic Quality</td>
<td>-2</td>
<td>Bad</td>
<td>In the range of the 25% worst results</td>
</tr>
<tr>
<td>Hedonic Quality</td>
<td>0.3333333333</td>
<td>Bad</td>
<td>In the range of the 25% worst results</td>
</tr>
<tr>
<td>Overall</td>
<td>-0.83</td>
<td>Bad</td>
<td>In the range of the 25% worst results</td>
</tr>
</tbody>
</table>

**Lod Live**

<table>
<thead>
<tr>
<th>Scale</th>
<th>Mean</th>
<th>Comparison to benchmark</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pragmatic Quality</td>
<td>2.475</td>
<td>Excellent</td>
<td>In the range of the 10% best results</td>
</tr>
<tr>
<td>Hedonic Quality</td>
<td>1.675</td>
<td>Good</td>
<td>10% of results better, 75% of results worse</td>
</tr>
<tr>
<td>Overall</td>
<td>2.04</td>
<td>Excellent</td>
<td>In the range of the 10% best results</td>
</tr>
</tbody>
</table>

**Rhizomer**

<table>
<thead>
<tr>
<th>Scale</th>
<th>Mean</th>
<th>Comparison to benchmark</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pragmatic Quality</td>
<td>-0.2</td>
<td>Bad</td>
<td>In the range of the 25% worst results</td>
</tr>
<tr>
<td>Hedonic Quality</td>
<td>-0.5</td>
<td>Bad</td>
<td>In the range of the 25% worst results</td>
</tr>
<tr>
<td>Overall</td>
<td>-0.35</td>
<td>Bad</td>
<td>In the range of the 25% worst results</td>
</tr>
</tbody>
</table>
Summary and Future work

➔ Conclusion

➔ Future work

◆ Reorganize and improve categories
◆ Additional data set(s)
◆ Expert search
◆ Add more interactive visualizations
Thank you!
References

Linked Data community - [Linked Data (LD)]

Keim, Andrienko, Fekete, Görg - [Visual analytics: Definition, process, and challenges, 2008]

Dadzie and Pietriga - [Visualisation of Linked Data – Reprise, 2016]

Pelzer - Ontologie - basierte und relationale Datenvisualisierung im Vergleich - Eine Fallstudie 2017

Dadzie, Rowe - [Approaches to visualising Linked Data: A survey, 2011]

Kräutli, Boyd Davis - [Digital Humanities Research Through Design, 2016]

Hassanzadeh, Consens - [Linked Movie Database (LMDB), 2009]

Images [User Centered Design], [ReactJS], [NodeJS], [Github], [SgVizlerJS]

Prototype available at [Github Pages]
Q & A