TopicLens, and More!

John O’Donovan

Four Eyes Lab, Department of Computer Science, University of California, Santa Barbara.
RecSys: Inspectability and Control

**Personal Characteristics (PC)**
- Familiarity with recommenders
  - \( 0.166 (0.077)^* \)
- Music expertise
  - \(-0.332 (0.088)^{***} \)
- Trusting propensity
  - \( 0.375 (0.094)^{***} \)
  - \( 0.205 (0.100)^* \)
  - \( 0.257 (0.124)^* \)

**Objective System Aspects (OSA)**
- Control
  - item/friend vs. no control
    - \( \chi^2(2) = 10.61^{**} \)
      - item: \(-0.181 (0.097)^{**} \)
      - friend: \(-0.389 (0.125)^{**} \)
  - Inspectability
    - full graph vs. list only
      - \( 0.288 (0.091)^{**} \)

**Subjective System Aspects (SSA)**
- Understandability
  - \( R^2 = .153 \)
  - \( 0.459 (0.148)^{**} \)
- Perceived control
  - \( R^2 = .311 \)
  - \( 0.377 (0.074)^{***} \)
- Perceived recommendation quality
  - \( R^2 = .512 \)
  - \( 0.770 (0.094)^{***} \)
  - \( 0.323 (0.031)^{***} \)

**User Experience (EXP)**
- Satisfaction with the system
  - \( R^2 = .696 \)
  - \( 0.955 (0.148)^{***} \)
  - \(-0.152 (0.063)^* \)

**Interaction (INT)**
- Inspection time (min)
  - \( R^2 = .362 \)
  - \( 0.695 (0.304)^* \)
- Number of known recommendations
  - \( R^2 = .344 \)
  - \( 0.695 (0.304)^* \)
- Average rating
  - \( R^2 = .508 \)
  - \( 0.067 (0.022)^{**} \)
Our recent work with RS interfaces:

<table>
<thead>
<tr>
<th>System</th>
<th>Type</th>
<th>API</th>
</tr>
</thead>
<tbody>
<tr>
<td>SmallWorlds</td>
<td>Music / Movies</td>
<td>Facebook</td>
</tr>
<tr>
<td>TasteWeights</td>
<td>Musical Artists / Jobs</td>
<td>Facebook, Twitter, DBPedia, LinkedIn</td>
</tr>
<tr>
<td>TopicLens</td>
<td>Twitter users and topics / Movies</td>
<td>Static / Twitter API</td>
</tr>
<tr>
<td>WigiPedia</td>
<td>Semantic Labels</td>
<td>DBPedia / MediaWiki</td>
</tr>
<tr>
<td>TopicNets</td>
<td>People, Documents, Topics</td>
<td>PDF Documents / Structured RDF documents</td>
</tr>
</tbody>
</table>
## Inspectability and Control Elements:

<table>
<thead>
<tr>
<th>System</th>
<th>Inspectability</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>SmallWorlds</td>
<td>Column Graph, Circular Graph, List View</td>
<td>Node-repositioning Drop-down menus</td>
</tr>
<tr>
<td>TasteWeights</td>
<td>List Views, Slider positions, Background Opacity, On-hover edges, Provenance view for re-ranking</td>
<td>Item/user sliders, Locks, domain sliders.</td>
</tr>
<tr>
<td>TopicLens</td>
<td>Graph and River View, 3D view, Many on-hover actions. Zoom</td>
<td>Side panel controls (buttons and sliders). Graph “spinning”, node clicks, Sorting. ( UI only No data-level</td>
</tr>
<tr>
<td>TopicNets</td>
<td>Graph view: Zoom, Click, Drag, List views, Table views, Charts</td>
<td>Huge amount of control. 10+ panels of functions. Full graph interaction, Layout algorithms etc.</td>
</tr>
</tbody>
</table>
## Inspectability Elements:

<table>
<thead>
<tr>
<th>Inspectability Mechanism</th>
<th>Advantage</th>
<th>Disadvantage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Node-Link Graph</td>
<td>Good provenance. Easy to inspect paths, neighbor links etc.</td>
<td>Scales badly, gets cluttered quickly (abstraction / clustering can help)</td>
</tr>
<tr>
<td>List Views</td>
<td>Simple, can be reranked with provenance annotations.</td>
<td>Hard to display connectivity</td>
</tr>
<tr>
<td>Interactive (hover, click, zoom etc)</td>
<td>Can handle lots of information. Create a “game-like” feel. Keep user interested</td>
<td>Hidden functionality. Usually needs some training / learning curve, or good annotation/help tools</td>
</tr>
<tr>
<td>Tabular Views</td>
<td>Easier to understand than a graph.</td>
<td>Hard to display complex connectivity / provenance</td>
</tr>
<tr>
<td>Text-based</td>
<td>Simple, Lots of detail</td>
<td>Boring? Does not scale well.</td>
</tr>
</tbody>
</table>
## Control Elements:

<table>
<thead>
<tr>
<th>Control Mechanism</th>
<th>Advantage</th>
<th>Disadvantage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Node-Link Graph (rating using node-drags)</td>
<td>Communicates impact of user input very well</td>
<td>Not initially intuitive, difficult to rerank vertically (crossed edges)</td>
</tr>
<tr>
<td>Node-Link Graph (for data selection)</td>
<td>Very useful for selecting a subset from a general overview</td>
<td>Edges cause clutter quickly esp. for large graphs.</td>
</tr>
<tr>
<td>Slider List Views</td>
<td>Clean look, Users are familiar with slider input, can be reranked easily with provenance data shown</td>
<td>Difficult to resize, less freedom than node-link views.</td>
</tr>
<tr>
<td>Right-click</td>
<td>Useful for node-specific functionality</td>
<td>Hidden functionality. Usually needs some training / learning curve, or good annotation/help tools</td>
</tr>
<tr>
<td>Control panels (buttons, sliders etc)</td>
<td>Easier to understand than a graph, can be labeled more easily.</td>
<td>Can get cluttered quickly depending on system complexity.</td>
</tr>
</tbody>
</table>
TopicLens: Exploring Content and Network Structure in Parallel
(Devendorf, O’Donovan, Hollerer)

Hybrid Network Views

- River and Graph representations displayed in parallel.
Visual Analysis of Dynamic Topics and...
LDA “Topic Models” useful for understanding relations in large volumes of text. Visualization and Interaction can help a user gain insights into topic modeled data. LDA can be iteratively applied to tailor the information space to a users requirement.

Gretarsson, O’Donovan et al. 2011 (ACM Trans. On the Web)
TopicLens is a General solution: New
Showing Credibility in the Underlying
View Inversion (Skeleton)
TopicLens as a Recommender System
Dynamic Thresholds

Minimum Documents Per Topic

1.00

11.00

21.00
2D/3D Views, Labeling Choices,
Supplementary Slides Follow