A System for the Evaluation of Digital Library Access

Norbert Fuhr
University of Duisburg Essen, Germany
Talk structure

- Digital library systems
- Architecture levels of information retrieval / digital library systems
- Levels of experimentation
- The ezDL system
- ezDL experimentation
- Conclusion
Digital library systems
Digital Library Model

Content — Uses — System
# Experimentation: Variation of Systems

<table>
<thead>
<tr>
<th>System</th>
<th>Uses</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Standard evaluation campaign
Experimentation: Variation of Content

Tests on different collections
Experimentation: Variation of Uses

- Topic types
- Task types
- User groups
Architecture levels of information retrieval / digital library systems
Architecture Levels of IR Systems

- Physical level
- Logical level
- Syntactic level
- Semantic level
- Pragmatic level
- Functional level
- User interface
Physical level
Logical level

- Physical level
- Logical level
- Syntactic level
- Semantic level
- Pragmatic level
- Functional level
- User interface

Vector space
Boolean
Fuzzy
Probabilistic
Language model

Learning to rank

\[ P(R|q,d) \propto \prod \lambda \cdot P(w|d) + (1 - \lambda) \cdot P(w|C) \]
Syntactic level

user interface
functional level
pragmatic level
semantic level
syntactic level
logical level
physical level

informat.* retrie[f|v].*
Semantic level

- User interface
- Functional level
- Pragmatic level
- Semantic level
- Syntactic level
- Logical level
- Physical level

- Thing
  - Creature
  - Artifact
Pragmatic level

Problem

Work task
Functional level

SELECT ... WHERE ...

API

Starting, Chaining, Browsing, Monitoring,
Differentiating, Extracting, Networking
User Interface

- Physical level
- Logical level
- Syntactic level
- Semantic level
- Pragmatic level
- Functional level
Levels of experimentation
Levels of Experimentation

- User interface
- Functional level
- Pragmatic level
- Semantic level
- Syntactic level
- Logical level
- Physical level
Levels of Experimentation: Physical

Performance experiments:
- Efficiency of data structures & algorithms
- Efficiency vs. effectiveness
Levels of Experimentation: Logical

Test of IR models
  • Focus on effectiveness
Levels of Experimentation: Syntactic/Semantic

Test of representations
- Focus on effectiveness
- Efficiency vs. effectiveness
Levels of Experimentation: Pragmatic

System support for
- problem solving
- work tasks

Methods:
- Field studies
- Simulated work tasks
Levels of Experimentation: Functional

- user interface
- functional level
- pragmatic level
- semantic level
- syntactic level
- logical level
- physical level

Test of search/cognitive functions

- user experiments
  - lab
  - living labs
Levels of Experimentation: User Interface

- user experiments
  - lab
  - living labs
The ezDL system
ezDL Architecture
ezDL Backend Architecture
ezDL – Lower Levels

Wrappers to IR Backends

- Amazon books (SOLR)
- Medicine: PubMed,…
ezDL

- higher level search functions: tactics, stratagems
- proactive support
- session support
- support of the digital library lifecycle
User Interface

- User Interface
  - Functional Level
  - Pragmatic Level
  - Semantic Level
  - Syntactic Level
  - Logical Level
  - Physical Level
ezDL experimentation
Evaluation web app.

1. Start
2. Training Task
3. Pre-Exp Questionnaire
4. Pre-Task Questionnaire
5. Task
6. Post-Task Questionnaire
7. Post-Exp Questionnaire
8. End

Log Viewer

Analysis Tool

System

Database

Collected Data
Log Data
Session Data
Questionnaire & Session Data
TPDL 2011 Workshop: User-Oriented Evaluation of Digital Library Interfaces

Pre-Experiment Questionnaire

Please fill out questionnaire

Age: 0

Gender: [ ] Male [ ] Female

First language:

Current Occupation:

Which diplomas/degrees have you been awarded?

- [ ] High School
- [ ] Bachelor or equivalent
- [ ] Master or equivalent
- [ ] PhD, Doctoral equivalent
- [ ] Professional degree (medicine, law etc.)

If you are a student, which university degree are you in the process of completing?

- [ ] Bachelor or equivalent
- [ ] Master or equivalent
- [ ] PhD, Doctoral equivalent
- [ ] Professional degree (medicine, law etc.)

Overall, how many years have you used the Web or other computerized resources to search for information? [ ] years.

How often do you search? (1: Never, 2: Once/twice a year, 3: Once/twice a month, 4: Once/twice a week, 5: once/twice a day)

- Digital libraries of scholarly articles (e.g. institutional repositories, ACM Digital Library) [ ]
- Web search engines (e.g. Google) [ ]
- Digital bookstore, such as e.g. Amazon [ ]

If you already have used a digital bookstore, what was your search purpose?

- [ ] to find a known book for purchase
- [ ] to explore bookstore for leisure purposes
- [ ] to explore bookstore for study/work purposes
TPDL 2011 Workshop: User-Oriented Evaluation of Digital Library Interfaces

Please start the system

Task 7 (task7)
Try to find books about a specific topic or of a certain type, but do not look for a specific title you already know.
## Logging

### Log Viewer

#### Session overview

<table>
<thead>
<tr>
<th>Session Id</th>
<th>User Id</th>
<th>Start</th>
<th>End</th>
<th>Log Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>cfe03fbb-6d4d-3e46-4ec7-f86c2b907a7e</td>
<td>oslo-12</td>
<td>Wed, 31 Aug 2011 15:54:34 +0200</td>
<td>Wed, 31 Aug 2011 16:04:41 +0200</td>
<td>XML</td>
</tr>
<tr>
<td>58ae76d3-51e0-4661-b2f5-c5390476f89d</td>
<td>oslo-11</td>
<td>Wed, 31 Aug 2011 15:35:40 +0200</td>
<td>Wed, 31 Aug 2011 15:50:47 +0200</td>
<td>XML</td>
</tr>
<tr>
<td>e2045aef-70b8-4f19-95f4-4a6c35eb0a44</td>
<td>oslo-12</td>
<td>Wed, 31 Aug 2011 15:25:12 +0200</td>
<td>Wed, 31 Aug 2011 15:35:15 +0200</td>
<td>XML</td>
</tr>
</tbody>
</table>
- <log>
  - <event name="start" timestamp="1314799370195" timestampClient="1314799409473">
    <user>oslo-13</user>
    <version>a</version>
  </event>
  - <event name="taskselected" timestamp="1314799370263" timestampClient="1314799409473">
    <task>task3</task>
  </event>
  - <event name="clientstart" timestamp="1314799370519" timestampClient="1314799410020"/>
  - <event name="desktopinfo" timestamp="1314799370551" timestampClient="1314799410020">
    <screenX>1680</screenX>
    <screenY>1050</screenY>
    <os>Windows 7</os>
    <java>1.6.0_26</java>
  </event>
  - <event name="search" timestamp="1314799395637" timestampClient="1314799435519">
    <solrQuery>4.racial AND 4.discrimination</solrQuery>
    <textQuery/>
    <titleQuery>racial discrimination</titleQuery>
    <authorQuery/>
    <yearQuery/>
    <abstractQuery/>
    <reviewQuery/>
    <searchType>content</searchType>
    <startDocNum>0</startDocNum>
    <endDocNum>19</endDocNum>
    <sorting order="descending">9998</sorting>
  </event>
  - <fields>
    <field>7</field>
    <field>1003</field>
    <field>4</field>
    <field>31</field>
    <field>8029</field>
    <field>1018</field>
    <field>8030</field>
  </fields>
  <projection>compact</projection>
</event>
- <event name="results" timestamp="1314799396417" timestampClient="1314799436174"/>
Eye Tracking: Heatmap
Eye Tracking: Areas of Interest (AOI)
Eye Tracking: AOIs vs. Time

The diagram illustrates the distribution of AOIs (Areas of Interest) over time for various categories such as Rating, Query History, Related Terms, Basket, Result List, Controls, Advanced Search, and Search Field. Each category is represented by a vertical bar that indicates the time duration during which the AOI was active.

The x-axis represents time in milliseconds, ranging from 0 to 800,000 milliseconds. The y-axis lists the different AOI categories. The bars are color-coded to indicate the intensity or frequency of AOI engagement throughout the observed period.
Conclusion
Conclusion

- Easy performance of IIR experiments
- Flexible configuration
- Extensible system
- Distributed system
- Centralized collection of experimental data
- Flexible analysis of data