USER STUDIES IN DIGITAL LIBRARIES: Models, Standards, Methods, Questions

DR. MILENA DOBREVA, UNIVERSITY OF STRATHCLYDE & UNIVERSITY OF PORTSMOUTH

PIERLUIGI FELICIATI, UNIVERSITY OF MACERATA
WHY DO WE NEED TO KNOW MORE ABOUT THE USERS?

- Providing better services
- Coping with information overflow
- Addressing issues of value and impact in the context of information society and knowledge economy
- Refining charging models (cost per view, subscription services)
MYTHS ABOUT USERS
MYTH 1: USERS? OF COURSE WE KNOW THEM!

- Mostly based on knowing ourselves 😊

**Issues**

- Beyond Humanities and Arts?
- Multilinguality / Multicultural use
- Refinement of user needs and expectations
- Digital objects reuse and enrichment
- From informing users to rich user experiences
MYTH 2: IF WE BUILD IT THEY WILL COME

- Supply-driven logic
- Still many projects start with the idea “this has to be of interest to a wider community” but do not check carefully what the community really wants

**Issues**

- Demand-driven
- Moving target
- Groups vs individuals – personalisation
MYTH 3: “THE DIGITAL MCDONALDS”

- Offering several options makes everyone happy

Issues

- Knowledge about groups vs knowledge about individuals
- Personalisation / recommender systems
MYTH 4: USER STUDIES? ERR... THIS IS THE SAME AS EVALUATION... OR USABILITY?

- Not all methods for evaluation of DLs involve [end] users
- User studies also aim to understand better the user (e.g. information behaviour studies)
- Usability is only one aspect of those

Issues

- Users are often forgotten!
MYTH 5: FOR USERS, QUALITY MEANS INNOVATION

- Assumption that all users want the latest technological gadgets and services.

**Issues**

*User satisfaction does not depend entirely on innovation!*

**ISO/IEC 9126-1 - Information Technology. Software product quality: quality model**

- Quality: the capability of the software product to enable specified users to achieve specified goals with effectiveness, productivity, safety and satisfaction in specified contexts of use.
TOP FIVE REASONS NOT TO DO USER TESTING/STUDIES

1. We don’t have the time.
2. We don’t have the money.
3. We don’t have the expertise.
4. We don’t have a usability lab.
5. We wouldn’t know how to interpret the results.
DIGITAL LIBRARIES: USERS’ PLACE IN KEY MODELS AND STANDARDS
The model does not provide an extensive set of roles: testing and evaluation are not included in the roles, even if they are essential in the DL life-cycle.

In addition, besides human actors, there are also bots, intelligent agents and other machine actors.
EXAMPLE: EUROPEANA USERS
A representation of user needs could be done combining two points of view:

- what Scenarios are most typical for the users,
- what properties of a Society specify a particular community of users (like the designed community of OAIS model)
DIGITAL LIBRARIES AND THEIR IMPACT/VALUE

Figure 1: Business Model

Harry Verwayen, EDL Foundation
EVALUATION OF DIGITAL LIBRARIES
BASIC EVALUATION ISSUES

• Quality
• Usability
• Accessibility

• Methods
  • Empirical (involve users)
  • Analytical
    • Heuristic evaluation
    • Cognitive walkthrough
    • Claims analysis
  • Attribute by attribute
    • Interaction triptych model
    • CASSM (Concept-based Analysis of Surface and Structural Misfits)

• Pivotal role of experts (intermediaries)
INFORMATION BEHAVIOUR STUDIES
INFORMATION BEHAVIOUR RESEARCH

- Change of focus
- Searching vs foraging
- Numerous theories
Information needs are not fundamental but secondary order needs arising from the desire to satisfy primary needs.

Classification of information needs (Taylor 1968):

- Visceral need – the actual, but unexpressed need
- Conscious need – the recognized need at a cognitive level.
- Formalized need – a formal statement of the need.
- Compromised need – the question (query) as presented to the information system or intermediary.
WILSON’S MODEL (2000)

Information Behavior is the totality of human behavior in relation to sources and channels of information, including both active and passive information seeking, and information use. Thus, it includes face-to-face communication with others, as well as the passive reception of information as in, for example, watching TV advertisements, without any intention to act on the information given.

Information Seeking Behavior is the purposive seeking for information as a consequence of a need to satisfy some goal. In the course of seeking, the individual may interact with manual information systems (such as a newspaper or a library), or with computer-based systems (such as the World Wide Web).
Information Searching Behavior is the ‘micro-level’ of behavior employed by the searcher in interacting with information systems of all kinds. It consists of all the interactions with the system, whether at the level of human computer interaction (for example, use of the mouse and clicks on links) or at the intellectual level (for example, adopting a Boolean search strategy or determining the criteria for deciding which of two books selected from adjacent places on a library shelf is most useful), which will also involve mental acts, such as judging the relevance of data or information retrieved.

Information Use Behavior consists of the physical and mental acts involved in incorporating the information found into the person's existing knowledge base. It may involve, therefore, physical acts such as marking sections in a text to note their importance or significance, as well as mental acts that involve, for example, comparison of new information with existing knowledge.
Human information behavior interacting with various forms of information through all channels for both active and passive information seeking and user.

Information seeking behavior, seeking for information in response to goals and intention by interacting with systems and humans.

Information searching behavior, actions involved in interacting with information search systems.

Information Systems at all levels including document-based, organizational, market and social system.

Information seeking systems, including other humans and information and communication technology.

Information retrieval systems, typically computer systems for documents and multimedia.
METHODS FOR USER STUDIES IN DL
BACKGROUND

• Anneli Sundqvist (2007): “the general knowledge of user behaviour is a mixture of common sense, presumptions and prejudices” in a study of digitised archives.

• The Institute of Museum and Library Services: “The most frequently-used needs assessment methods do not directly involve the users” (2003).

• Michael Khoo et al.: “In the case of digital library researchers, the focus of research is often on technical issues (e.g., information retrieval methods, software architecture, etc.) rather than on user-centered issues. When these researchers turn to user based evaluations, they therefore often lack the necessary expertise to develop robust Human Computer Interaction (HCI) experiments, and their goals are typically limited to "proof of concept" tests, rather than prescribing user motivations or cognitive impacts.” (2009).
NEW MATURITY

• Isolated studies
• Supply-driven logic
• Expert opinions are dominant

• Mixed methods
• Justification of users required by some funding agencies
• Personalisation

--2005

2005--

now

NEW

EXPANSION

MATURITY

• Growth of number of studies
• Introducing new methods
• Isolated studies – no benchmarks
• Typical users/user communities
KEY QUESTIONS - METHODS

• How user studies help to understand better the needs in digital resources and their use?

• What questions could be answered by different types of studies?

• How to construct a study?
WHAT DATA CAN WE GATHER?

Wide range...

- Both through questioning and observation – direct, indirect
  - Quantitative
  - Qualitative
- Growing role of evidence-based research
HOW THE OUTCOMES COULD LOOK LIKE?

Ease of resource discovery

- Most participants found that the resource is very easy to use for resource discovery, “easy to find what I was looking for” (P6); “it comes across as very well structured and provides searching flexibility for the user” (P8). There are participants which found the search “very simple and what else could I expect…. Although you do get a lot of hits on the first search, the vast majority of people like to type something in and then advance if they want to.” (P3); a similar view was expressed by (P9). However, the multiple results are seen as beneficial in the teaching context “Initially you might get a lot of hits, but in the context of teaching you can never have too many.” (P4). The resource as “a good research tool for students and academics” (P5). One interviewee noted that the appearance of the search term in the results should also be highlighted: “I like the highlighting of the search term in the key word search but I have no idea why I am retrieving information in the index search” (P10).

- Google Analytics indicates that 4,210 of recorded visits, representing 44.16% of the total number of visitors, were directed through 10 search engines. It is worth noting what users had been searching for: 11 researchers who are searching for the Stormont Papers web site; 16 researchers who are searching for an online copy of the Stormont parliamentary debates, but are not necessarily aware or looking for the Stormont Papers; 23 searched for terms that returned as a result a resource from Stormont papers.

(SPHERE project, 2011)
HOW THE OUTCOMES COULD LOOK LIKE?

Ease of resource discovery

- Most participants found that the resource is very easy to use for resource discovery. "It comes across as very well structured and provides searching flexibility for the user" (P8). There are participants which found the search "very simple and what else could I expect... Although you do get a lot of hits on the first search, the vast majority of people like to type something in and then advance if they want to." (P3); a similar view was expressed by (P9). However, the multiple results are seen as beneficial in the teaching context "Initially you might get a lot of hits, but in the context of teaching you can never have too many." (P4). The resource as "a good research tool for students and academics" (P5). One interviewee noted that the appearance of the search term in the results should also be highlighted: "I like the highlighting of the search term in the key word search but I have no idea why I am retrieving information from the index search" (P10).

- Google Analytics indicates that 4,210 of recorded visits, representing 44.16% of the total number of visitors, were directed through search engines. It is worth noting what users had been searching for: 11 researchers who are searching for the Stormont Papers website; 16 researchers who are searching for an online copy of the Stormont parliamentary debates, but are not necessarily aware or looking for the Stormont Papers; 23 searched for terms that returned as a result a resource from Stormont papers. (SPHERE project, 2011)
HOW THE OUTCOMES COULD LOOK LIKE? (CONT’D)

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interface design</td>
<td>56%</td>
</tr>
<tr>
<td>Familiarity from previous use</td>
<td>54%</td>
</tr>
<tr>
<td>Currency (e.g., publication date)</td>
<td>54%</td>
</tr>
<tr>
<td>URL (e.g., Web domain)</td>
<td>49%</td>
</tr>
<tr>
<td>Author’s credentials</td>
<td>48%</td>
</tr>
<tr>
<td>Heard about site before</td>
<td>44%</td>
</tr>
<tr>
<td>Linkage (if links exist)</td>
<td>43%</td>
</tr>
<tr>
<td>Different viewpoints acknowledged</td>
<td>42%</td>
</tr>
<tr>
<td>Chart quality (if they exist)</td>
<td>39%</td>
</tr>
<tr>
<td>Author credits others for ideas</td>
<td>32%</td>
</tr>
<tr>
<td>Bibliography included</td>
<td>23%</td>
</tr>
<tr>
<td>Mentioned by librarian</td>
<td>11%</td>
</tr>
</tbody>
</table>

Example: Criteria for evaluating Web content by students, Head, Eisenberg, 2011
Example: Criteria for evaluating Web content by students, Head, Eisenberg, 2011
HOW THE OUTCOMES COULD LOOK LIKE? (CONT’D)

Average values of either $\text{Precision}_{[0\div1]}$ or $\text{Precision}_{[0\div2]}$ bare no significant difference for either library.

<table>
<thead>
<tr>
<th></th>
<th>$0 \div 2$</th>
<th>$0 \div 1$</th>
<th>$\sigma_{0 \div 2}$</th>
<th>$\sigma_{0 \div 1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>JeromeDL</td>
<td>31.44</td>
<td>33.18</td>
<td>0.17</td>
<td>0.17</td>
</tr>
<tr>
<td>DSpace</td>
<td>30.97</td>
<td>33.49</td>
<td>0.19</td>
<td>0.19</td>
</tr>
</tbody>
</table>

$\Delta$:

- 1.54%  
- 0.91%

Example: Kruk et al., 2008
HOW THE OUTCOMES COULD LOOK LIKE? (CONT’D)

Example: Kruk et al., 2008
HOW THE OUTCOMES COULD LOOK LIKE? (CONT’D)
HOW THE OUTCOMES COULD LOOK LIKE? (CONT’D)
HOW THE OUTCOMES COULD LOOK LIKE? (CONT’D)

This is Europeana - a place for inspiration and ideas. Search through the cultural collections of Europe, connect with other user pathways and share your discoveries.
HOW THE OUTCOMES COULD LOOK LIKE? (CONT’D)

Quantitative
“Maria is a School teacher, comfortable with computers and the internet. Happily Googles but also frequently having a specific target for her searches as she prepares for work. She uses her mobile to update her Facebook status, but mostly for calling and texting. Her aim is often to prepare for classes, but also to find new ways of motivating her pupils” - Personas (short version) from EuropeanaConnect.
"Maria is a School teacher, comfortable with computers and the internet. Happily Googles frequently having searches as she uses her mobile to update her Facebook status, but mostly for calling and texting. Her aim is often to prepare for classes, but also to find new ways of motivating her pupils" - Personas (short version) from EuropeanaConnect.
CHOICES, CHOICES, CHOICES…

- Qualitative
- Quantitative
- Qualitative
- Quantitative
- Quantitative, IR metrics
- Summative
AND EVEN MORE CHOICES…
USER STUDIES IN THE DL CONTEXT

- In most cases the current studies are evaluating existing DLs; DLs in development are addressed less frequently.
- Many studies are “stand-alone”; they address a specific DL or a small group of DLs and to be able to compare DLs we need to know how to compare the user experiences.
- The studies focus mostly on specific aspects such as usability; more work needs to be done to contextualise better specific DL user studies and information behaviour as well as user experience studies.
- In many cases the studies address a limited set of user communities but in the WWW this is not sufficient.
SUMMARY OF METHODS

• Direct user involvement
  • Questionnaires
  • Focus groups
  • Diaries
  • Observation

• Indirect observation
  • User logs
  • Eye tracking

• Personae
• Ethnographic studies
• Use scenarios

• Growing use of mixed methods
TYPICAL FAULTS: INDIVIDUAL

Developer watching videotape of usability test.

TYPICAL FAULTS: GROUPS

"You call this a focus group?!!"

Source: http://smallbiztrends.com/2009/03/when-focus-groups-are-not-well-focused.html
Summary of methods

<table>
<thead>
<tr>
<th>SUMMATIVE</th>
<th>INDIRECT</th>
<th>DIRECT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personalisation models</td>
<td>Diaries</td>
<td>Interview</td>
</tr>
<tr>
<td>Use scenarios</td>
<td>Mid-size log analysis</td>
<td>Focus groups</td>
</tr>
<tr>
<td>Personas</td>
<td>Deep log analysis</td>
<td>Questionnaires</td>
</tr>
</tbody>
</table>

Type:
- Indirect
- Direct

Population:
- Individual
- Samples
- Population

Involvement:
- Summary of methods
  - Deep log analysis
  - Questionnaires
  - Expert evaluation
  - Personalisation models
  - Use scenarios
  - Diaries
  - Personal log analysis
  - Interview
  - Eye tracking
  - Focus groups
  - Ethnographic studies

48
Cost of methods

<table>
<thead>
<tr>
<th>INDIRECT</th>
<th>DIRECT</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUMMATIVE</td>
<td></td>
</tr>
<tr>
<td>Personal log analysis</td>
<td>Mid-size log analysis</td>
</tr>
<tr>
<td>Diaries</td>
<td>Use scenarios</td>
</tr>
<tr>
<td>Personalisation models</td>
<td>Expert evaluation</td>
</tr>
<tr>
<td>Questionnaires</td>
<td>Focus groups</td>
</tr>
<tr>
<td>Interview</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Eye tracking</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TYPE</th>
<th>POPULATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>INDIRECT</td>
<td>INDIVIDUAL</td>
</tr>
<tr>
<td>DIRECT</td>
<td>SAMPLES</td>
</tr>
<tr>
<td>SUMMATIVE</td>
<td>POPULATION</td>
</tr>
</tbody>
</table>

Involvement
Timeframes for methods

- **Personalisation models** (Individual)
  - **Diaries** (Indirect)
  - **Interview** (Direct)
- **Use scenarios** (Samples)
  - **Personal log analysis** (Indirect)
  - **Eye tracking** (Direct)
- **Personas** (Population)
  - **Mid-size log analysis** (Indirect)
  - **Focus groups** (Direct)
- **Expert evaluation** (Population)
  - **Deep log analysis** (Indirect)
  - **Questionnaires** (Direct)

**Type**: Summative
**Indirect**: Personal log analysis
**Direct**: Eye tracking

**Involvement**: Deep log analysis

**Timeframes**: Questionnaires, Ethnographic studies
## TIMING OF STUDIES

<table>
<thead>
<tr>
<th>Type</th>
<th>What is it used for</th>
</tr>
</thead>
<tbody>
<tr>
<td>Front-end involvement</td>
<td>Users can take part in assessment on a variety of technical requirements, e.g. resolution, dimensions of digital objects, preferred formats for use. At this stage users can also take part in exploratory research, e.g. needs in new resources and defining requirements, as well as rationale for selection, appraisal and prioritisation of content.</td>
</tr>
<tr>
<td>Normative evaluation</td>
<td>This type of evaluation usually takes form of iterative circles of process-and-evaluation when implementing digitisation of collections. Most typically such evaluation will focus on usability, e.g. interfaces and presentation of digitised resources; coverage of identified needs for specific audiences.</td>
</tr>
<tr>
<td>Summative evaluation</td>
<td>Here the focus is the final output and the accordance to the expectations and requirements of target communities/organisation structures/the wider disciplinary domain.</td>
</tr>
<tr>
<td>Direct engagement in the digital resource creation</td>
<td>Direct user engagement can utilise social media tools which allow users to contribute their own digital objects or to take part in the enrichment of resources – e.g. supplying full texts, or metadata.</td>
</tr>
</tbody>
</table>

### Type
- **Front-end involvement**: Users can take part in assessment on a variety of technical requirements, e.g. resolution, dimensions of digital objects, preferred formats for use. At this stage users can also take part in exploratory research, e.g. needs in new resources and defining requirements, as well as rationale for selection, appraisal and prioritisation of content.

### Normative Evaluation
- This type of evaluation usually takes form of iterative circles of process-and-evaluation when implementing digitisation of collections. Most typically such evaluation will focus on usability, e.g. interfaces and presentation of digitised resources; coverage of identified needs for specific audiences.

### Summative Evaluation
- Here the focus is the final output and the accordance to the expectations and requirements of target communities/organisation structures/the wider disciplinary domain.

### Direct Engagement in the Digital Resource Creation
- Direct user engagement can utilise social media tools which allow users to contribute their own digital objects or to take part in the enrichment of resources – e.g. supplying full texts, or metadata.
OPEN QUESTIONS...

• Why and when to involve users when digitising collections: front-end, normative, summative evaluation

• How to involve users?

• How to address needs of future users?

• How to evaluate impact and value of digital resources?
AND THIS IS NOT ALL…

Crowdsourcing
Do we have the models?

Benchmarking
When could we really start comparing?

Awareness
What are we going to adopt for OUR DL?
Knowledge about users is part of the professional expertise in digital libraries.

However, we still face multiple misconceptions about users.

We need:

- Research agenda
- Benchmarking
- Wider professional discussion
Thank you!

User Studies for Digital Library Development (forthcoming)
Milena Dobreva, Andy O'Dwyer and Pierluigi Feliciati (eds)
Facet Publishing
ISBN: 978-1-85604-765-4