How We are Searching Cultural Heritage

A qualitative analysis of search patterns and success in The European Library

Vivien Petras, Juliane Stiller, Maria Gäde

Lida 2014, 17.06.2014
# Surrogates & Metadata

## Onverbreekbare banden

**Harvard-Style Citation**

UFA (Berlin), n.d. Onverbreekbare banden. (c.l.): (s.n.)

**Description**

Zarah Leander

**Creator**

UFA (Berlin)

**Contributor**

UFA (Berlin)
Mette, Frans
Mees & Bronkhorst

**Source**

BG E173 (affiche), 150 jaar Nederlandse reclame, ReclameArsenaal

**Publication**

22-08-2004

**Resource type**

Still image

**Extent**

1000 x 650 mm

**Medium**

text/xml

**Format**

affiche
## Query Statistics

<table>
<thead>
<tr>
<th>Search Term</th>
<th>Total Unique Searches</th>
<th>Results Pageviews / Search</th>
<th>% Search Hits</th>
<th>% Search ReRefinements</th>
<th>Time after Search</th>
<th>Average Search Depth</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. &quot;&quot;</td>
<td>3,948 (2.12%)</td>
<td>7.15</td>
<td>53.09%</td>
<td>6.60%</td>
<td>00:04:58</td>
<td>0.68</td>
</tr>
<tr>
<td>2. &quot;photo&quot; OR &quot;fotó&quot; OR &quot;photographie&quot; OR &quot;photography&quot; OR &quot;fotografie&quot; OR &quot;fotografía&quot;</td>
<td>984 (0.53%)</td>
<td>3.06</td>
<td>47.26%</td>
<td>34.25%</td>
<td>00:02:15</td>
<td>0.10</td>
</tr>
<tr>
<td>3. &quot;photo&quot; OR &quot;fotó&quot; OR &quot;photographie&quot; OR &quot;photography&quot; OR &quot;fotografie&quot; OR &quot;fotografía&quot;</td>
<td>604 (0.32%)</td>
<td>2.99</td>
<td>60.60%</td>
<td>43.21%</td>
<td>00:01:41</td>
<td>0.08</td>
</tr>
<tr>
<td>4. what:&quot;map&quot;</td>
<td>470 (0.25%)</td>
<td>6.07</td>
<td>62.98%</td>
<td>6.80%</td>
<td>00:06:05</td>
<td>0.27</td>
</tr>
<tr>
<td>5. what:&quot;TEXT&quot;</td>
<td>330 (0.18%)</td>
<td>1.31</td>
<td>27.27%</td>
<td>12.70%</td>
<td>00:00:43</td>
<td>0.99</td>
</tr>
<tr>
<td>6. europeana_collectionName:90402</td>
<td>326 (0.18%)</td>
<td>5.92</td>
<td>59.20%</td>
<td>7.04%</td>
<td>00:02:52</td>
<td>0.30</td>
</tr>
<tr>
<td>7. what:&quot;image&quot;</td>
<td>260 (0.14%)</td>
<td>1.73</td>
<td>28.46%</td>
<td>15.78%</td>
<td>00:01:14</td>
<td>0.93</td>
</tr>
<tr>
<td>8. 1</td>
<td>243 (0.13%)</td>
<td>1.01</td>
<td>98.77%</td>
<td>1.22%</td>
<td>00:00:01</td>
<td>0.01</td>
</tr>
<tr>
<td>9. what:&quot;text&quot;</td>
<td>223 (0.12%)</td>
<td>1.60</td>
<td>23.77%</td>
<td>17.21%</td>
<td>00:01:39</td>
<td>1.02</td>
</tr>
<tr>
<td>10. what:&quot;Photo&quot;</td>
<td>222 (0.12%)</td>
<td>1.58</td>
<td>28.36%</td>
<td>14.00%</td>
<td>00:01:17</td>
<td>1.14</td>
</tr>
<tr>
<td>11. what:&quot;Text&quot;</td>
<td>193 (0.10%)</td>
<td>1.44</td>
<td>24.87%</td>
<td>25.27%</td>
<td>00:01:35</td>
<td>1.03</td>
</tr>
<tr>
<td>12. what:&quot;photograph&quot;</td>
<td>178 (0.10%)</td>
<td>1.60</td>
<td>22.47%</td>
<td>14.04%</td>
<td>00:01:20</td>
<td>1.34</td>
</tr>
<tr>
<td>13. what:&quot;d day&quot; OR &quot;normandy landings&quot;</td>
<td>170 (0.09%)</td>
<td>3.06</td>
<td>93.53%</td>
<td>1.54%</td>
<td>00:02:08</td>
<td>0.13</td>
</tr>
<tr>
<td>14. who:(Massi-Lehre)</td>
<td>147 (0.08%)</td>
<td>1.17</td>
<td>77.55%</td>
<td>1.16%</td>
<td>00:00:19</td>
<td>0.21</td>
</tr>
</tbody>
</table>
Beyond Statistics: Semantics

- Information Need
- Goals
- Categories
- Language
- Patterns
- Success
So far...
Query Categories

**Named Entities**
- Person
- Work title
- Geographic
- Event
- Organization

**Non-Named Entities**
- Topical
- Other
QUERIES IN CONTEXT
How We Are Searching Cultural Heritage

The European Library
Connecting knowledge

Designed to meet the needs of the research community worldwide, our online portal offers quick and easy access to the collections of the 48 National Libraries of Europe and leading European Research Libraries.

Users can cross-search and reuse over 24,121,397 digital items and 140,394,972 bibliographic records.

To facilitate further research, links are also provided to other websites in the Europeans group.

National Library of Denmark
This beautiful building that is office to the National Library of Denmark is known as The Black Diamond. Image courtesy; National Library of Denmark.
Case Study

Information Need
Goals
Language
Actions
Patterns
Success

http://www.affiliateblog.de/wp-content/uploads/2010/02/Lupe.jpg
Query

Goals

Sessions

Queries

Actions

Information needs

Reformulations

Hard indicators

Soft indicators
Information Needs & Reformulations

Information Need 1

Query → Query

Reformulation

Information Need 2

Query
Query Reformulation Patterns

Specialization

“caricature”

“Caricature philipon”

Generalization

“partituras”

“Musica”

Reformulation

“Romeo and Juliet”

“Shakespeare”
Reformulations in Digital Libraries

<table>
<thead>
<tr>
<th>English</th>
<th>German</th>
<th>French</th>
<th>Greek</th>
<th>Dutch</th>
<th>Italian</th>
<th>Spanish</th>
<th>Arabic</th>
<th>Hindi</th>
<th>Chinese</th>
<th>Japanese</th>
</tr>
</thead>
<tbody>
<tr>
<td>sword</td>
<td>schwert</td>
<td>fencing</td>
<td>πεονιδεσ</td>
<td>Marseille</td>
<td>peonides</td>
<td>Marseille</td>
<td>2753700303</td>
<td>shakespeare</td>
<td>Romeo and Juliet</td>
<td>peonides</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO</td>
<td>GO</td>
<td>GO</td>
<td>GO</td>
<td>GO</td>
<td>GO</td>
<td>GO</td>
<td>GO</td>
<td>GO</td>
<td>GO</td>
<td>GO</td>
</tr>
</tbody>
</table>

- Monolingual synonyms
- Multilingual parallel search
- Multilingual synonyms
- Correction
- Category change
- Transliteration
- Exact field transformation
Session Success

- Sessions
- Queries
- Information needs
- Reformulations
- Actions
Success Rate

45% of sessions are successful

Sessions with **multiple queries** are **more successful** than the ones with only one query
Success and Reformulations

Multilingual Parallel Search
Monolingual synonyms
Broader queries
Category Change
Narrower query
Spelling variants / corrections
Simple to fielded search
Transliteration
Query reduction

1. Likely to increase the amount of retrieved objects
2. Likely to show more of the content of the DL
3. Likely to show content from different angles
Implications for Design

• Support user with query translation & query suggestion
• Match queries to vocabulary terms in query suggestion

Dominant and prevailing search paradigm
Diving into Cultural Heritage

User needs are hardly researched

Extent of collections is often elusive

Majority of content stays hidden
How to Surface Information

Let the user explore and discover

Guide them through collections

Support new users – information tourist

“Making the best of everything” by aristocrats-hat | CC BY-NC-SA 2.0
How We Are Experiencing Cultural Heritage!

User
Engage
Browse
Search
Content
Thank you

vivien.petras
juliane.stiller mariagae@gmail.com
maria.gaede
@ibi.hu-berlin.de