Redesigning our Combine Harvester

Experimenting with Metalib’s X-server: ideas, attempts and realisations

Ane W. van der Leij
Dpts of Information & Collection Development / Digital Library Facilities
University Library
University of Groningen (RUG), the Netherlands
Outline

• Introductory
• Our Metalib
• Statistics–modules

Metalib Statistics
(SFX–statistics)
• Experimenting with the X–server: LiveTrix
Demo
Summarizing: Front features
Summarizing: Rear features
Under the hood
• Concluding

(if time allows)
• Extras
Introductory

University of Groningen –
Rijksuniversiteit Groningen (RUG)

- founded 1614
- 22,000 students, 6,000 staff
- 494.2 mln EUR expenditure (2005) (632 mUSD, 4.5 mdSEK)

University Library
- Central library, in addition to 10 faculty libraries
- 130 staff
- Collection: 2.2 mln printed vols., 24,000 e–journals, 450 licenced electronic databases
- 1.36 mln EUR expenditure on collection development (incl. co–financing of electronic resources by faculty libraries) (1.74 mUSD, 12.53 mSEK)
Our Metalib: RUG Combine

- Introduced v2 as pilot 2004, Faculty of Medicine and University Medical Centre (Hospital staff)
- Migrated to v3 in January 2005, completing introduction to the whole of the University September 2005
- NOT only gateway (OPAC, Guide, Lists); usage modest so far

Our SFX: RUGlinks

- Introduced 2002, v3 since 2005
Statistics (1)

2 Modules: RUGlinks and RUG Combine

- Built upon logfiles, stored in MySQL-database
- Home made search software
- Available through HTML-webapplication
- Extract data that meet our needs and wishes
- Database and software on separate server
- Not interfering with production environment for Metalib and SFX
- Online webinterface; moved away from admin-module
- Username/Password access for library staff and management
Statistics (2)

What we found

15% errors:
technical measures > reducing to 1.5% errors/flaws

50% zero or false results:

- Misspellings and typos in search terms
- Picking databases at random
- Unable to understand QuickSearch, MetaSearch, Find Database
- Using the wrong search keys
- Using search keys wrong
- Using Dutch search terms in English language databases
- Using non-specific terms, phrases that are too broad
- Lack of understanding of Boolean logic or database peculiarities

Metalib statistics
## Statistics about specific searches

Generate statistics for specific searches. The report can be filtered by date range, the full name of the resource, the status of the search and/or the number of documents retrieved.

### Date

**From**: 2006-10-01  **To**: 2006-10-15

### Resource name:

- AATA online: art and archaeology technical abstracts
- Academic Search Premier
- AccessMedicine
- ACL Anthology
- ACM Digital Library

To select multiple databases: keep Control key (Ctrl) pressed and click with the mouse on each database.
To select all databases: select first database, keep "Shift" key pressed and simultaneously press "End" key.

### Status:

- [ ] All searches
- [ ] Successful searches
- [ ] Unsuccessful searches

### No. documents:

- [ ] Zero

**Send**
What if...

• we could users make to understand the work flow
• we could clear away misspellings
• we could suggest other or related search terms
• we could make users pick the right databases
  or better still…
• users wouldn’t have to bother which databases to choose

We can!

By building new applications around Metalib’s XML-server (X-server)
X–server

If Metalib is the car, then the X–server is the engine.

We like the engine!

It’s a Volvo!

The car itself, though...

... a DAF?

We could have power steering, ABS, a navigation system, cruise control, keyless entry, air bags all around us and a nice car stereo. We build our own car!
The Swedish dreams of Ingmar Bergman

1999 L. Axelsson

Wild strawberries
1993 P. Frencz, K. Frencz
Bergman's, Ingmar Bergman, Ingrid
ISBN: 0-85170-481-6

Three post-Inferno Strindberg plays which Strindberg did not write
1994 T. Bredsdorf
Forf. jämför Smultronstallet med En dromslöv och Den starkare med Personas och Tribunernas nat.

Long day's journey into night
1994 E. Törnqvist
Birghem, Ingmar Bergman, Ingmar
Strindberg's 3 Post Inferno Plays / Editor

Drommens klaara bilder
1993 E. Törnqvist
Om förhållandet mellan filmmanuskriptet och den färdiga filmen
Filmdoktoren Ingmar Bergman

Ingmar Bergman's opus 18 & 27
1992 J. H. Jekowsk
Birghem, Ingmar Bergman, Ingmar
Smultronstället och dödens ekipage

<table>
<thead>
<tr>
<th>By:</th>
<th>Wimark, Margareta</th>
</tr>
</thead>
<tbody>
<tr>
<td>In year:</td>
<td>1998</td>
</tr>
<tr>
<td>Full-text lookup</td>
<td>No citation information</td>
</tr>
</tbody>
</table>

Other actions:
- Print record
- View raw XML record, or this view, in new window.
- Search for article title in: A9, Google Scholar

Free-form notes:
- Birghman, Ingmar Bergman, Ingmar

Keywords:

Note: The meaning, use and usefulness of the fields below varies. Mouse-overing the field name may provide hints.
The sources to search in are guessed based on statistics. To select a set yourself, click here.
Rembrandt's anatomy lesson as a metaphor for education

By: De Ville, K.A. Kopelman, L.M.

From: ISSN: 0149-7944
Current Surgery Mar 2003

In year: 2003

Journal title: Current Surgery
ISSN: 01497944 Volume
Year: 2003 Issue
Page: ...

More information or results incorrect...

Wait for lookup...
- Elsevier ScienceDirect Complete (full text or best guess) (create tinyurl)
- RUG Catalogue (catalogue lookup, of journal) (create tinyurl)
- Citation Information and URL (formatted citation)

The following may (also) work:
- URI's support (our lookup and/or links are usually preferable, there are no guarantees to even what this is)
The anatomy lesson / Philip Roth

By: Roth, Philip, 1933-


The following items are related somehow (according to eISBN). These links are to Amazon:

- 1-4025-7365-0
- 0-224-02960-6
- 0-224-02960-6
- 0-224-02960-6
- 1-4025-0766-6
- 4-08-124086-8
- 0-09-947661-4
- 0-449-20614-9
- 0-679-74902-0
**Query analysis**

**Unknown strings**

The following strings are **not (yet) known**. This may simply mean they had not been searched for yet, they are rare, misspelled, or ignored because it is a stopword or very long.

- anatomy lessons - could you have meant `anatomy lesson` or `some lessons` or `anybody response`? (click to search in same sources)

**Likely sources**

Statistics show that the following will also be **responsive to the words in your search** (though not necessarily the search itself):

- Top few: COPAC, Library of Congress, SCIRUS (Elsevier), Nationalbibliografin - Böcker, LIBRIS: det nationella biblioteksdatasystemet, CancerLit, (click to search in this set)

**Substrings**

**Shorter strings** strings from your search are generally more likely to yield results. The following list is sorted; most likely to give results (usually because it is more general) come first. (Click to search each in sources that seem to represent them best)

- lessons
- anatomy
X-server: LiveTrix Front

- Ticking one or several Subject Areas (if you want)
- Primary, Secondary (and Tertiary) resources (if you want)
- Suggest-mode (if you want)
- Select resources (if you want)
- Help texts, search hints, explanation
- No Meta, no Quick
- Filling in search fields; no Simple, no Advanced
- Spelling check
- Graphical representation

LiveTrix
X-server: LiveTrix Rear

- Search progress
- Presentation of hits
- Immediate access, even during retrieval process
- Help texts, search hints, explanation
- Short (table) view

- Long view, featuring: Book covers, Inline SFX services (link resolving to full text, inline catalogue lookup, searching Google, Google Scholar, A9, export to RefMan), Generate persistent tiny URL, Delivering extra metadata from external resources, e.g. local library systems
LiveTrix: under the hood

Search terms are sent to all databases in the background.
The first 30 titles from each database are retrieved. These titles are indexed both as a string and word for word, as well as counted. A relation network is built up. From then on, the system knows in which databases these terms occur.

In the next query, suggestions about databases (“likely sources”) are given on the basis of these ever expanding indexes.

Dito for suggestions about word relations.

The indexes are also used for correcting typos and word/phrase suggestions.

Suggestions about similar sounding phrases are given upon phonetic matches (“Soundex”).
Concluding

How do we continue?
Uncertain
- Our funding
- Our manpower
- Maybe seeking cooperation/partners (inter)nationally
- What does ExLibris come up with?

Aim to make it a working system for end users.
- Completing it with personalization, alerting tool, solutions for “link-only” resources
- Improving the interface
- Usability tests
Thank you for your attention!
Tack för uppmärksamheten!

Ane van der Leij
a.w.van.der.leij@rug.nl

Project group:
André Keyzer (UL)
Tom Blikslager (Library of Arts)
Guus van den Brekel
(Central Medical Library)
Bart Alewijnse (UL – trainee)
Roxana Popistasu (UL – trainee)
Peter van Rees (UL)

University of Groningen Libraries: http://www.rug.nl/bibliotheek
RUG Combine: http://combine.ub.rug.nl
LiveTrix: http://livetrix.ub.rug.nl