MULTI-SEARCH - A SEARCH TOOL FOR LITERATURE RETRIEVAL

Hans Åhlfeldt, Erik Lanhammar, Ove Wigertz
Medical Informatics, S-581 85 Linköping University, SWEDEN
E-mail: Hans.Ahlfeldt@ami.liu.se
WWW: http://www.imt.liu.se/

Abstract

This paper reports on the design and implementation of a search tool for access to the Elhill family of databases. The architecture is in principle similar to other existing client/server based search applications for literature retrieval, but offers a series of important features in relation to access to Swedish databases as well as to major international databases. The system offers a uniform search form which is initialised with search elements and access methods depending on the chosen database structure. The search form could be adapted according to the users search experience and utilises in its full power all available access elements and search strategies offered in the different databases. The paper also discusses similarities and differences compared to other search tools and indicates possible future directions.

Introduction

Two general questions in relation to information retrieval are where relevant information sources can be found, and how a specific search is to be performed once a set of potentially valuable information sources have been identified. From the perspective of the Swedish health care sector and related research and educational institutions, the Medical Information Centre (MIC) at the Karolinska Institute in Stockholm is a major provider of bibliographical databases. MIC offers access to the Elhill family of databases developed by the National Library of Medicine (NLM) in USA, and to a newly introduced Ovid server currently offering Medline, Current Contents and a collection of 15 full text journals. A general trend over the last years is that end-users more and more tend to perform their own searches [1, 2] via search tools such as the well known Grateful Med [3]. Several studies on the problems encountered by end-users and the quality of search results have also been performed [4, 5, 6]. This paper reports on the design and implementation of an end-user search tool for access to the Elhill family of databases at MIC or NLM. The architecture is in principle similar to other existing client/server based search applications for literature retrieval, and builds upon the experiences from the development of MultiLink [7] but offers a series of new features e.g. in relation to MeSH navigation, available access elements, search statement formulation and reformulation and post-processing.

System description

Following a client/server architecture, remote information servers can be accessed through a front-end, making heterogeneous systems appear as a homogenous system to the end-user. MultiSearch provides access to the following databases: Aidsline, Alconline, Arbline,

Cancersdi, Cats, Cisilo, Drugline, Lokat, Medline, Nioshtic, Nordguide, Nordser, Nurseline, Riskline, Sdiline, Serline, Spriline, Swemed. Figure 1 shows the design of the search form of MultiSearch. The general layout of the form is the same for all databases, but the six search methods (six buttons on the left side of the screen) can be initiated with different access elements and associated search methods. Following the example, the first button provide the user with available search methods associated with MeSH such as explode, subheadings and restrict to focus. The second button is initiated with methods for non-MeSH searches such as abstract and keyword (see figure 2). Button five is not initiated at all in the example, while the others hold search methods for title-, publication- and text word search. The approach of linking access elements with associated search methods including screen dialogues such as pop-up lists, radio check buttons, help windows etc. to the search form, gives the user freedom to tune the system according to experience and knowledge. The experienced user can then include all available access elements and also use an on-line search window, allowing direct interaction with the database. This gives the opportunity for reformulation or application of Boolean operators on search statements "on the fly" without the need to go back to the search form for a new session. The function is supported with appropriate help windows. MeSH support is available through the Meta Thesaurus developed within the UMLS project [8, 9] or through on-line navigation in the MeSH tree on the server. The Meta Thesausus, provided with an alphabetic browser, is loaded in the client application. The thesaurus contains all MeSH concepts together with non-MeSH synonyms and related clinical concepts, giving the user the opportunity to freely browse through the thesaurus and collect suitable search terms without concern of server cost.

Figure 1. Design of the search form in MultiSearch (Swedish version). The six search buttons/fields are initiated with access elements and associated search methods according to the settings of the programme. Available access elements with associated search methods for Spriline can be seen in figure 2.

Cat.	Access Element	Status
AB	Abstract	TW
AU	Author	*
CA	Call Number	
CN	Corporate Author	*, TW
CS	Corporate Name as Subject	*, TW
DT	Document Type	*
ED	Edition	
IC	Imprint I Collation	
IS	ISBN I ISSN	*
KW	Keyword	*, TW
LA	Language	*
MH	MeSH Heading	*
NT	Note	
SE	Series Title	*, TW
SO	Source	
TA	Title Abbreviation	*
TI	Title	TW
UI	Unique Identifier	*, R
YR	Year	*, R

Figure 2. Access elements in the Spriline database. The search methods are given in the Status column; *=direct searchable, TW=textword search, R=rangeable. Different databases have different configurations, but MultiSearch allows the user to take full advantage of all the different search elements. The search form is automatically initiated with associated search methods including end-user dialogues according to the settings of the programme.

Browsing in the MeSH tree on the server is also possible as can be seen in figure 3. The tree structure of MeSH gives the experienced user the possibility to select terms on appropriate levels in the tree depending on preference between recall or precision. Based on such a navigation, MultiSearch could also provide correct lists with subheadings and scope notes via easy-to-operate dialogues. All search methods in relation to MeSH is supported by MultiSearch such as explode, pre-explode, restrict to focus, and connection of subheadings to MeSH terms.

MultiSearch is implemented with HyperCard and runs on Macintosh. Network access is available through TCP/IP or modem dialup. MultiSearch can save and load search profiles, format search results, save, print or send found references via e-mail. General features of client/server based search tools are that they save money by preparation of search profiles off-line, and that the user can forget complicated command languages. MultiSearch allows also the search forms to be highly personalised according to experience and personal preferences in relation to search interest and search strategies. Help is available in the different stages of the retrieval process, and integration with other applications through e.g. cut and paste is available in ordinary Macintosh style.

```
▤▦▦▦▦▦▦▦▦▦▦▦▦▦ tree medical records ▦▦▦▦▦▦▦▦▦▦▤▦
Records  E5.318.308.940                                          ⬆
   Medical Records  E5.318.308.940.968                           ▤
      Medical Record Linkage  E5.318.308.940.968.500
      Medical Records, Problem-Oriented  E5.318.308.940.968.750
      Trauma Severity Indices  E5.318.308.940.968.875 (+)

Records  G3.850.520.308.940
   Medical Records  G3.850.520.308.940.968
      Medical Record Linkage  G3.850.520.308.940.968.500
      Medical Records, Problem-Oriented  G3.850.520.308.940.968.750
      Trauma Severity Indices  G3.850.520.308.940.968.875

Records  L1.280.900
   Medical Records  L1.280.900.968
      Medical Record Linkage  L1.280.900.968.500
      Medical Records, Problem-Oriented  L1.280.900.968.750
      Medical Records Systems, Computerize  L1.280.900.968.775
      Trauma Severity Indices  L1.280.900.968.875 (+)

Records  N4.452.859
   Medical Records  N4.452.859.564
      Medical Record Linkage  N4.452.859.564.550
      Medical Records, Problem-Oriented  N4.452.859.564.600    ⬇
      Medical Records Systems, Computerize  N4.452.859.564.650  ▣
⬅▥▦▦▦▦▦▦▦▦▦▦▦▦▦▦▦▦▦▦▦▦▦▦▦▦▦▦▦▦▦▦▦▦▦▦▦▦▦▦▦▦▦➡▣
```
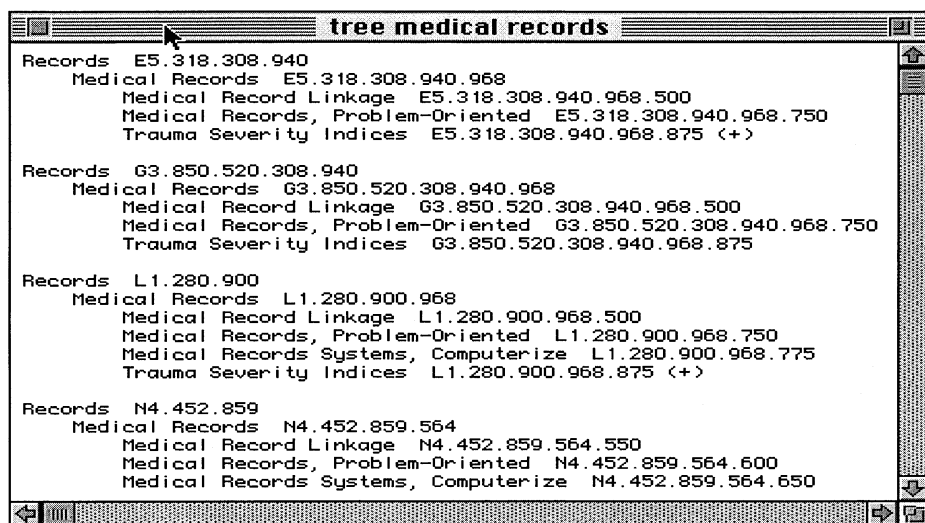
Figure 3. Window for navigation in the MeSH tree, giving the user a sense of the hierarchical structure of MeSH together with the position of the different MeSH terms which allows for correct mapping of subheadings to a selected search term.

Discussion

The MultiSearch system is based on a client/server architecture offering uniform and easy-to-operate search forms to all available databases on the server. Other products based on a client/server architecture are KR ProBase™ (Knight-Ridder Information) and the Ovid™ client/server system. Both providers are currently offering clients running under Windows. The Ovid system allows access to Z39.50-compliant servers, and available databases are Medline, Current Contents and the Ovid full text biomedical collection, currently including 15 journals. KR ProBase offers access to all DataStar databases (including Medline) and integration of DIALOG databases are announced during mid 1996. Grateful Med™ [9] offers both Mac and PC clients to the Elhill databases provided by the National Library of Medicine. MultiSearch has as a comparison a number of advantages and limitations. The system is so far limited to Elhill databases, but offers on the other hand a client/server based approach to the Scandinavian databases (such as Spriline, Swemed, Drugline, Nordguide, Nordser) which are not covered by Ovid or KR ProBase. The system takes full advantage of the Elhill databases, regardless if the chosen database is MeSH indexed or not. It offers access to MeSH through the MetaThesaurus [8, 9] loaded in the client application, as well as on-line navigation in the MeSH tree. It handles all available search elements and search strategies in the Elhill system, and includes features for initialisation of the programme according to the users search experience. An important feature for experienced users is the on-line search window, allowing the user to continue to search manually, e.g. to apply different Boolean operators on search statements or otherwise reformulate search statements. This on-line mode differs from Grateful Med [3], another search tool from NLM, which retrieves only in batch mode without the possibility for the user to directly interact with the database. Possible directions for the future include conceptual support of the information

retrieval process via access to the Swedish SPRI-term database [10] which might offer mappings from Swedish clinical concepts to e.g. MeSH. Internet and the World Wide Web technique [11] might also provide a truly distributed environment for search applications such as the one reported in this paper.

Acknowledgement

References

1. Mischo W, Lee J. End-user searching of bibliographic databases. Annual Review Information Science Technology 1987; 22: 227-263.
2. Laufmann S, Blumenthal R, Thompson L, Bowen B. Direct end-user access to remote information. SIGOIS bulletin 1991; 12: 16-28.
3. Kingsland L., Syed E, Lindberg D. Caoch: an expert searcher program to assist Grateful Med users searching MEDLINE. Proceedings MEDINFO 92; 382-386.
4. Walker C. Problems encountered by clinical end users of MEDLINE and Grateful Med. Bull Med Libr Assoc 1991; 79: 67-69.
5. Mitchell J, Johnson E, Hewett J, Proud V. Medical students using Grateful Med: analysis of failed searches and a six-month follow-up study. Comp Biomed Research 1992; 25: 43-55.
6. Haynes B et.al. Performance of 27 Medline systems tested by searches with clinical questions. JAMIA 1994; 1: 285-295.
7. Wu G, Ahlfeldt H, Wigertz O. MultiLink - an intermediary system for multi-database access. Methods Information Medicine 1993; 32: 82-89.
8. Lindberg D, Humpreys B, McCray A. The unified medical language system. Methods Information Medicine 1993; 32: 281-291.
9. Lindberg D, Humphreys B. The UMLS knowledge sources: tools for building better user interfaces. Proceeding SCAMC 1990; 121-125.
10. Junfors A, Wallin S-B, Thurin A. A conceptual model for a terminological system for the healthcare sector of Sweden. Submitted to MIE96.
11. Cimino J, Socratous S, Clayton P. Internet as clinical information system: application development using the WWW. JAMIA 1995; 2: 273-284.