Alternative interfaces for PubMed searches

Isabelle de Kaenel, Pablo Iriarte
Medical Library, University Hospital ; CHUV BH-08, CH-1011 Lausanne, Switzerland
isabelle.de-kaenel@chuv.ch ; pablo.iriarte@chuv.ch

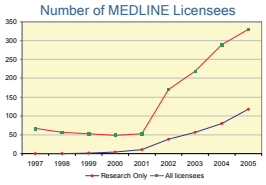
Introduction : Over the last 5 years, more than 30 new search interfaces for **MEDLINE/PubMed** have been made publicly accessible. This wave surged in the early 2000's with the launch of the PubMed web services, the « Entrez Programming Utilities ». Along with the traditional commercial redistributors (Ovid, CSA, Dialog and many others), academic institutions, bioinformatics services, pharmaceutical enterprises and software companies leased the MEDLINE/PubMed data for research purpose and devised alternative interfaces. These applications draw on data mining techniques, linguistic analysis, knowledge management resources, statistical methods, and new web technologies. The new search tools were examined by reference librarians at the Lausanne University Hospital, in order to decide whether they should be used as alternate search engines for specific requests and even be introduced in training sessions for students and clinicians.

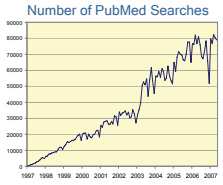
NLM Mission : To collect, organise, and **disseminate** the world's health-related and biomedical information

MEDLINE

1990's : CD-ROM + Online → 2000's : XML + web services

MEDLARS Centers → Commercial distributors → Researchers & developers





Method : To get an overview of these different retrieval systems, literature and web searches were conducted. The sites listed in the « NLM's registry of licensees who use MEDLINE/PubMed data for research purposes » and in the MEDLINE section of the « Metadatabase for the biological sciences » on *Neurotransmitter.net* were also systematically explored. The main innovations of these interfaces where then classified along specific criteria and functions.

Information retrieval

Search

- Graphical search interface : **PubMed Interact** presents slider bars to control the traditional PubMed search limits and parameters.
- Free text / natural language query tool : **AskMEDLINE** processes clinical questions written in plain English.
- Guided search form : **PICO Linguist,** a multilingual interface, helps in building a structured clinical query by using the PICO framework.
- Cross-language tool : queries can be submitted to **BabelMeSH** in 8 different languages. Search terms are then mapped to a multilingual MeSH.
- Similarity quest : the input of at least 8 PMIDs allows **Pubfinder** to return a hit list of references in order of relevance.

Screen

- Relevance ranking : **ReleMed** defines 8 levels of «relevance» to sort the records in a descending order based on the estimated scores.
- Graphical display and concept maps : **AliBaba** provides a graphical and interactive view of the extracted information.
- Frequency tables : **Pubreminer** generates hyperlinked tables after statistical analysis of the result sets.
- Clusters : **MESHPubMed** or **ClusterMed** dynamically create clusters of references classified by topic, author, MeSH headings or date.
- Tag-clouds : **iHOP** produces tag clouds based on MESH terms. The size of each cloud is proportionate to the frequency of occurrence of the keyword.

Knowledge retrieval

Semantic scan

- Analysis of word co-occurrence : **XplorMed** parses abstracts and calculates a degree of relatedness between words.
- Entity recognition : using the Gene Ontology, **GoPubMed** identifies relevant biological concepts and shows possible hidden correlations between genes, drugs, and diseases.
- Text & data mining : **LitMiner** annotates key terms in abstracts and predicts relationships between key biomedical terms in 4 categories : genes, chemical compounds, diseases and organs.
- Information extraction : **AliBaba** scans abstracts for biological objects and creates associations with descriptions of the relationships.
- Semantic networking tools : **PubFocus** performs terminology extraction and couples the information with external parameters.

Save & structure

- Control over the results list : **PubMed Interact** enables users to label, delete and add from within the existing list of citations.
- Personal annotations : **Hubmed** offers many functions that allow to add personal information, tags and links.
- Easy & direct export : **PubMed Assistant** exports metadata in multiple formats to reference management tools.
- Integration of third party content and web services : **PubFocus** integrates external indicators gathered from journal rank lists and forward referencing database.
- Quick links to outside applications : **WiKiPDF** includes links to descriptive articles in Wikipedia.

PubMed PubReMiner · XplorMed · PubFocus · gopubmed · PubMed · Whatizit · HubMed: pubmed rewired · PUBMED GOLD · PubMed Assistant · PubMed Interact (Beta) · eTBLAST · TEXTMED THE MEDICAL SEARCH ENGINE · PICO Linguist (Beta) · ReleMed · ClusterMed · <BabelMeSH>< · TermSciences · meshpubmed · iHOP · PubFinder · WormBase · MEDIE

Conclusion : These new interfaces address many different needs and user profiles : clinicians working in private practices in Europe could use the multilingual tools such as **BabelMeSH** and **PICO Linguist** ; researchers that need to scan vast amount of literature will turn to advanced methods that **mine information** and bring out hidden relations and associations. As for the students, they may be attracted by social tagging and **web 2.0** technologies also used in recreational web sites. In short, these new tools have to be introduced in training sessions, at least to appreciate the interest they raise in the different groups of users. These developments reveal the transition from « **resource centric** » to « **user centric** » systems, and from « **information retrieval** » to « **knowledge retrieval** ».

- Siadaty M, Shu J, Knaus W. **Relemed: sentence-level search engine with relevance score**. *BMC Medical Informatics and Decision Making* 2007 Jan;7(1):1. [DOI:10.1186/1472-6947-7-1]
- Eaton AD. **HubMed: a web-based biomedical literature search interface**. *Nucleic Acids Res* 2006 Jul 1;34(Web Server issue):W745-W747. [DOI:10.1093/nar/gkl037]
- Liu F, Ackerman M, Fontelo P. **BabelMeSH: development of a cross-language tool for MEDLINE/PubMed**. *AMIA Annu Symp Proc* 2006;1012. [PMID:17238631]
- Muin M, Fontelo P. **Technical development of PubMed Interact**. *BMC Medical Informatics and Decision Making* 2006 Jan;6(1):36. [DOI:10.1186/1472-6947-6-36]
- Doms A, Schroeder M. **GoPubMed: exploring PubMed with the Gene Ontology**. *Nucleic Acids Res* 2005;33:W783-W786. [DOI:10.1093/nar/gki470]
- Fontelo P, Liu F, Ackerman M. **askMEDLINE: a free-text, natural language query tool**. *BMC Medical Informatics and Decision Making* 2005 Jan;5(1):5. [DOI:10.1186/1472-6947-5-5]
- Goetz T, von der Lieth CW. **PubFinder: a tool for improving retrieval rate of relevant PubMed abstracts**. *Nucleic Acids Res* 2005;33:W774-W778. [DOI:10.1093/nar/gki429]
- Perez-Iratxeta C, Bork P, Andrade AM. **XplorMed: a tool for exploring MEDLINE abstracts**. *Trends Biochem Sci* 2001;26:573-5. [DOI:10.1016/S0968-0004(01)01926-0]

Further information :

- **CHUV Library homepage :** http://www.chuv.ch/bdfm
- **MEDLINE/PubMed at the CHUV library WiKi :** http://www.bium.ch/wiki/doku.php?id=medline_et_pubmed
- **Licensee Research Use of NLM MEDLINE/PubMed Data :** http://www.nlm.nih.gov/bsd/licensee/reports/
- **Metadatabase for the Biological Sciences :** http://www.neurotransmitter.net/metadb/index.php?catid=65