

Supplementary notes 2:

CONTENTS:

- A. Additional notes on search tools
- B. Additional notes on tools that may improvise the search options of pubmed.
- C. Additional notes on tools that process the citations to extract biological entities and their relations, with or without graphical display of the summary.
- D. Notes on other databases/search engines related to clinical activities:
- E. List of other information retrieval tools:

An important note: As we were preparing the manuscript, several tools kept changing the features and new tools were reported frequently. After some time, we had to stop following up the changes to update in this review. GoPubMed is one such program that has undergone improvements. We have listed a few new tools at the end of this document. This shows the flux in the literature search domain. The corresponding author is also developing a new tool to aid very specific aspect of citation retrieval from PubMed!

A Brief Overview of Literature Search Tools:

PubMed, despite being very popular search engine, cannot fulfil all the needs of scientists or physicians. Hence, many new tools have attempted improvements in some aspects of the literature search.

Other summary scanners: Most of the summary scanners interface with MEDLINE/PubMed. PubMed Interact (Muin, M., Fontelo, P., Liu, F. & Ackerman, M., *SLIM: an alternative Web interface for MEDLINE/PubMed searches - a preliminary study. BMC Med. Inform. Decis. Mak.* 5, 37 [2005]; Muin, M. & Fontelo, P., *Technical development of PubMed interact: an improved interface for MEDLINE/PubMed searches. BMC Med. Inform. Decis. Mak.* 6, 36 [2006]) is an enhanced query interface for MEDLINE/PubMed that allows setting query limits easily with the help of slider bars. Literature search using askMEDLINE (Fontelo, P., Liu, F. & Ackerman, M., *askMEDLINE: a free-text, natural language query tool for MEDLINE/PubMed. BMC Med. Inform. Decis. Mak.* 5, 5 [2005]) is easy for those who may not be acquainted with specialized terms and/or the complicated features of advanced search programs, as it allows MEDLINE/PubMed queries in the form of simple questions. eTBLAST (Lewis, J., Ossowski, S., Hicks, J., Errami, M. & Garner, H.R., *Text similarity: an alternative way to search MEDLINE. Bioinformatics* 22, 2298-2304 [2006]; Errami, M., Wren, J.D., Hicks, J.M. & Garner, H.R., *eTBLAST: a web server to identify expert reviewers, appropriate journals and similar publications. Nucleic Acids Res.* 35, W12-W15 [2007]) can help to perform literature search using abstracts of relevant articles as query components. Like eTBLAST, PubFocus (Plikus, M.V., Zhang, Z. & Chuong, C.M., *PubFocus: semantic MEDLINE/PubMed citations analytics through integration of controlled biomedical dictionaries and ranking algorithm. BMC Bioinformatics* 7, 424 [2006]) also provides a ranked list of citations that can assist in finding the authors and journals of related interests. ReleMed (Siadaty, M.S., Shu, J. & Knaus, W.A., *ReleMed: sentence-level search engine with relevance score for the MEDLINE database of biomedical articles. BMC Med. Inform. Decis. Mak.* 7, 1 [2007]) is specially built for sorting articles based on relevance.

Many programs search and extract the citations from MEDLINE/PubMed and attempt to enhance the utility of biological information contained in titles and abstracts by useful display of the bio-entities and their relations. These tools include GoPubMed (*Doms, A. & Schroeder, M., GoPubMed: exploring PubMed with the Gene Ontology. Nucleic Acids Res., 33, W783-W786, [2005]*), BioIE (*Divoli, A. & Attwood, T.K., BioIE: extracting informative sentences from the biomedical literature. Bioinformatics 21, 2138-2139 [2005]*; *Kim, J.J. & Park, J.C., BioIE: retargetable information extraction and ontological annotation of biological interactions from the literature. J. Bioinform. Comput. Biol. 2, 551-568 [2004]*), EBIMed (*Rebholz-Schuhmann D. et al, EBIMed—text crunching to gather facts for proteins from MEDLINE. Bioinformatics 23, e237-e244 [2007]*; *Rebholz-Schuhmann D, et al, Protein annotation by EBIMed. Nat. Biotechnol. 24, 902-903 [2006]*), ALIBABA (*Plake, C., Schiemann, T., Pankalla, M., Hakenberg, J. & Leser, U., AliBaba: PubMed as a graph. Bioinformatics, 22, 2444-2445 [2006]*), BioAsk, PubMed Assistant (*Ding, J., Hughes, L.M., Berleant, D., Fulmer, A.W., Wurtele, E. S., PubMed Assistant: a biologist-friendly interface for enhanced PubMed search. Bioinformatics 22, 378–380, [2006]*), Anne O’Tate (*Smalheiser, N.R., Zhou, W. & Torvik, V.I., Anne O’Tate: A tool to support user-driven summarization, drill-down and browsing of PubMed search results. J Biomed Discov Collab. 3, 2 [2008]*), CiteXplore, PubWindows, ClusterMed and XplorMed (*Perez-Iratxeta, C., Bork, P. & Andrade, M.A., XplorMed: a tool for exploring MEDLINE abstracts. Trends Biochem. Sci. 26, 573-575 [2001]*; *Perez-Iratxeta C., Bork P. & Andrade M.A., Exploring MEDLINE abstracts with XplorMed. Drugs Today (Barc). 38, 381-389 [2002]*; *Perez-Iratxeta, C., Pérez AJ, Bork P, Andrade MA. Update on XplorMed: A web server for exploring scientific literature. Nucleic Acids Research. 31, 3866-3868 [2003]*).

While CiteXplore and PubMed Assistant can list and/or high-light the MeSH terms and other entities per citation, PubWindows helps to initiate a search on a chosen MeSH topic by allowing an easy scan of MeSH terms. XplorMed, Anne O’Tate and ClusterMed permit searches with regular query terms and can group the resulting abstracts based on the MeSH terms within them. ClusterMed categorizes PubMed results into different folders and subfolders and thus allows researchers to navigate through the results quickly. The results are also clustered by author, date and several topics, including the mesh headings.

EBIMed is a search engine from European Bioinformatics Institute with many useful attributes, particularly the query options, automatic processing of citations and the display features.

GoPubMed extracts citations from PubMed and classifies them based on Gene Ontology (GO) terms and MeSH terms. GO is a hierarchically structured vocabulary for molecular biology with about 20000 terms. Some of the very user-friendly aspects of the tool include: (a) A list of several terms (biological entities, research topics, techniques etc) occurring in the abstract and a ‘term-hierarchical tree’, along with a crisp description of each term. (b) The terms used in the query as well as related GO and MeSH terms are high-lighted within the abstracts. (c) A ‘hot topics’ utility provides detailed statistical analysis of search results including the graphical view of rate of publications over time, distribution of publications across journals, frequent authors and geographical locations of researchers publishing on the query-topic. (d) There are also useful links. The tool can sort the citations within chosen categories but this sorting is restricted to the top 10000 hits only.

BioIE can retrieve sentences within PubMed abstracts that refer to structure, function, diseases and therapeutic compounds, localization, or familial relationships of biological entities (mainly proteins). The tool searches for the query terms and retrieves a specified number of abstracts (based on user's inputs). It also permits extraction of useful sentences within the set of abstracts. There is considerable flexibility in the choice of topic for such information extraction: Apart from the pre-defined options like structure, function, diseases etc, user can enter a new topic of own interest.

BioAsk also has interesting novel features. It can categorize citations based on the biological entities present within and indicate the distribution of such entities across different categories.

More information:

TOOLS THAT HAVE TRIED TO IMPROVISE THE SEARCH OPTIONS OF PUBMED:

CiteXplore: performs literature search based on user's query and then performs text mining (by Whatizit, internally, and iHOP, externally) on the resulting citations. Simple and advanced search options are available.

Links to full text versions are provided where available. The tool also provides options to high-light the query terms or other bio-entities within the retrieved citations. CiteXPlore offers the option to highlight bioentities in the abstracts. When possible, it is best to derive the query set from PubMed and use it in CiteXplore, as the latter doesn't provide the quality of query options similar to the former program.

PubMed Interact: is an interactive, time-saving interface for MEDLINE/PubMed and follows the search and syntax rules of PubMed. This interactive site from NCBI presents 'slider bars' to enable quicker settings and re-settings of different parameters involved in the search strategy. The parameters include date of publication and journal subset (those which provide free abstract, text, MEDLINE set, core clinical journals only etc). It also eliminates the need to go back and forth between pages to set improved search parameters. There is an option to sort articles by marking the resulting citations in the output according to relevance. The tool still seems to be in its beta version but has certain limitations, including absence of an option to search within title of articles and discrepancies in the number of citations retrieved, as compared to PubMed.

askMEDLINE: allows MEDLINE/PubMed queries in the form of simple questions. Literature search using this tool is easy for those who may not be acquainted with specialized terms and/or the complicated features of advanced search programs. The tool attempts to translate a question into an efficient search. An automated multi-round query strategy is involved where the terms used are broadened or narrowed down depending on the number of citations obtained in the initial rounds of the search. The results in text-mode can be a special advantage for those using wireless handheld devices or a low-bandwidth connection in remote locations. The overall efficiency in retrieving journal articles is shown to be ~ 97% in one study*. In a period of one year (2005-06), the tool

received more than 15,000 queries and most of these pertained to ‘therapy’ (*Fontelo P, Liu F, Ackerman M, Schardt C, Keitz S. askMEDLINE: a Report on a Year-Long Experience. AMIA Annu Symp Proc.:923 [2006]*). The tool also provides links to PICO (Patient, Intervention, Comparison, Outcome) interface & Babel MesH - (Multilanguage search for MEDLINE/PubMed).

eTBLAST: is a more recent online system that offers the easiest and most effective method of published literature search with abstracts, rather than words and phrases as query components. The set of utilities at eTBLAST can hierarchically list articles with high similarity to the query paragraph as well as the journals and authors frequently associated with such related articles. The search system, which is under consideration for further improvements, can be very useful to find appropriate journals for publishing and expert researchers for reviewing or collaboration.

A query comprising of set of key words, without Boolean operators, could also be input in eTBLAST. PubMed could retrieve more specific hits than eTBLAST and the results were more precise while eTBLAST had several unrelated hits. But an artificial abstract created by combining relevant sentences from selected abstracts and used as the query in eTBLAST, obtained lesser and more specific hits.

PubFocus: can also provide ranked list of citations and assist in deciding authors and journals of related interests. PubFocus also indicates distribution of NCI ontology and gene orthology terms. But it lacks a complete help guide and is very slow, particularly when the number of hits is high.

ReleMed: can search for query terms within sentences and it helps to avoid the task of creating multiple phrases while querying. For example, in PubMed, to search on ‘transcription initiation’, one may have to either painstakingly list all possible phrases made by the terms ‘transcription’ and ‘initiation’, when using PubMed (e.g., transcription initiation, transcriptional initiation, initiation of transcription, initiating transcription) or use individual terms (e.g., ‘transcription* AND initiat*’), which usually leads to dilution of the results with unwanted hits. But in relemed, using individual terms can still fetch required citations on priority.

ReleMed categorizes each retrieved citation into 8 different levels of relevance, depending on the frequency of occurrence of the search terms within the title, sentences of abstract & MeSH.

ReleMed can also pull out the relevant sentences and display them with highlighted keywords. These features help to save time by allowing a quick-scan for selecting or eliminating of citations. The ATM is active by default. The translations for the query allow the user to check the usefulness of the translated terms. There is an option to turn off the ATM completely or selectively. The tool also provides an easy link to other search engines for every article.

ReleMed lacks most of the advanced features present in PubMed. Some of the other major limitations of the relemed include (a) absence of the 'history' feature equivalent to that of PubMed (b) the need to connect to PubMed even for the abstract of the citation (c) the information on the free availability of the free full text article is not readily displayed as it is done in PubMed (d) The total number of hits, if at all indicated in the results page, is very approximate (e.g., more than 1000) (e) The number of hits shown per page is fixed and low (ten only).

The tool was in its beta version until very recently and has been fully functional with many modifications since its inception.

MiSearch Adaptive PubMed Search Tool*: MiSearch is very recently built software to rank the retrieved citations based on an estimated likelihood of a user choosing to view them. User's citation choices are acquired by the program during browsing. The tool lacks an efficient query system. AVAILABILITY: <http://misecond.ncibi.org> (*States DJ, Ade AS, Wright ZC, Bookvich AV, Athey BD. MiSearch Adaptive PubMed Search Tool. Bioinformatics. 2008 Mar 11; [Epub ahead of print]*)

MedKit*: We were unable to access this program. The tool seems to be developed to overcome the following limitations of PubMed: PubMed imposes an upper limit of 10,000 for downloading PMID list or citations; and MEDLINE files are too large for most off-the-shelf XML parsers. AVAILABILITY (didn't work for us): <http://metnetdb.gdcb.iastate.edu/medkit>
Ding J, Berleant D. MedKit: a helper toolkit for automatic mining of MEDLINE/PubMed citations. Bioinformatics. 2005 Mar 1;21(5):694-5.

MeSH Speller + askMEDLINE*: This software claims to provide MeSH term and Specialist Lexicon Vocabulary suggestions as it is typed on a search page.
(Fontelo P, Liu F, Ackerman M. MeSH Speller + askMEDLINE: auto-completes MeSH terms then searches MEDLINE/PubMed via free-text, natural language queries. AMIA Annu Symp Proc. 2005:957)

PubMed Informer* : is a tool for monitoring topics of interest from MEDLINE/PubMed using five tracking methods: Web access, e-mail, Short Message Service (SMS), PDA downloads and RSS feeds. *(Muin M, Fontelo P, Ackerman M. PubMed Informer: monitoring MEDLINE/PubMed through e-mail alerts, SMS, PDA downloads and RSS feeds. AMIA Annu Symp Proc. 2005: 1057).*

BioProber*: A program that claims to discover relations between various biomedical entities within the PubMed citations and provide visual display of relations.
(Jang H, Lim J, Lim JH, Park SJ, Lee KC. BioProber: software system for biomedical relation discovery from PubMed. Conf Proc IEEE Eng Med Biol Soc.1:5779-82 [2006]).

PubWindows*: allows to browse PubMed citations with MeSH terms. It allows easy set up of RSS feeds based on subject searches and easy citation exporting in popular formats: RIS/EndNote, BibTex, RDF, MODS, RefWorks. AVAILABILITY: http://www.neurotransmitter.net/PubMed_browser.php?topheight=40

MEDSUM* : provides a simple way to count PubMed citations under different categories like the time of publication, occurrence of MeSH terms, authors, journals and other PubMed pre-defined subsets. AVAILABILITY: <http://webtools.mf.uni-lj.si/public/medsum.html>

*Not studied in detail in the current analysis.

TOOLS THAT PROCESS THE CITATIONS TO EXTRACT BIOLOGICAL ENTITIES AND THEIR RELATIONS, WITH OR WITHOUT GRAPHICAL DISPLAY OF THE SUMMARY.

BioIE

The selected abstracts can be individually retrieved in text mode. There are also options to view the frequency-statistics of search terms within the citations. This output would have been more useful with appropriate links to the citations from the table that displays the statistical details.

EBIMed

A "Fuzzy search" feature permits automatic recruitment of other terms that could be similar to the query term. It is also possible to set a limit to the extent of similarity of such terms to be recruited. One can look for the co-occurrence of two terms separated by a predefined no. of terms with "proximity search". The tool enables the user to stress ('boost') the importance of a specific query term. Another feature, 'range searching', allows setting upper and lower limits for numeric values as well as certain text (like author names etc).

EBIMed finds abstracts in the PubMed-way, analyses them and presents a complete overview of associations between the biological entities found within the abstracts. The types of entities are protein/gene names, cellular components, biological processes, drugs, species and molecular functions. These entities are obtained from Uniprot, gene ontology and different web applications of NCBI. The results, summary of co-localized terms or hit-pairs, are displayed in a table-format. The first column contains the type of the reference-entity, which is protein/gene by default but can be altered by user. Every row displays one specific reference-entity in the first column and the co-localized entities in following columns. Along with the specific entity/term in each cell, the number of abstracts is shown and this number serves as a link to the corresponding source sentences from the abstracts. The abstracts, MeSH terms used and the biomolecules mentioned in each title/abstract, can be extracted. When using EBIMed, we missed the ability to select and save the citations as well as the history-based search refinements available in PubMed. But the conceptual overview of the results is unique and very useful.

Whatizit

Whatizit is a sister application from EBI that can be useful in a wide variety of advanced searches. This engine, which may also be accessed programmatically, can search specific terms in molecular biology in given texts and link them to publicly available databases.

BioAsk (beta version)

BioAsk searches MEDLINE, categorizes the citations based on biological entities identified within and provides a graphical output indicating the distribution and links of results across different categories. There is a filtering utility to help the user to modify the level of relevance for sorting. A unique feature of BioAsk is that it allows multiple query inputs at a time. The tool also has links to different databases, and provides options to export, print, email or permanently save articles of interest. The visualization tool provides an easy and interactive representation of biological concepts based on the sentence level occurrences of the bio-entities across the citations. But such useful display of the results is limited to the top 500 hits only.

ALIBABA

ALIBABA (needs java 1.5 or higher) accepts query in the form of PubMed syntax as well as uniprot IDs, interfaces with PubMed, analyses a set of retrieved abstracts and presents the extracted information on biomedical objects, as well as their relationship, an easily perceptible graphical format. The 'confidence limit' setting can be used to alter the stringency of relationships between objects in the graph and hence the level of complexity of the graph. Some of the other aspects represented in the results include proteins or genes discussed in the context of diseases, mentions of drugs together with diseases, and relations between diseases and tissues. The output includes links to the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. The program searches for simple co-occurrences of terms in the same sentence and thereby establishes many relations. Specific sentences that could have relevance to the user are displayed by the tool. However, the tool does not currently consider relations between every possible pair of objects. A unique time-saving feature of ALIBABA is that it can be started without a browser from the user's computer.

The default numbers that can be visualized is 20 but higher number of hits and unfiltered searches slowed down the process. This could be because, unlike BioAsk that focuses on broader grouping of bio-entitites, ALIBABA tries to reveal more number of specific entities.

PubMed assistant

PubMed Assistant is a Java application (requires Java 1.4.2 or later versions, with web-start) that interfaces with MEDLINE. PubMed assistant has several useful functions: (a) One can select a few citations from the results and ask the tool to 'auto-refine' the query based on these selections. (b) The MeSH terms and 'chemicals' (bio-molecules) are listed for each citation selected. (c) Additional tools to enable quick and easy connections to other frequently used applications. (d) A visual query editor However, the PubMed assistant was slow. Another major disadvantage is that the tool, as of now, can handle a maximum of 5000 hits only

XplorMed

XplorMed is a server that allows query in the PubMed format, as a file or a database entry. The tool can be useful to group abstracts based on the associations between the words within them. The output will first show the PMIDs of the first 500 abstracts categorized under 16 MeSH categories. One or more of these can be selected to explore the relationships between different biological terms. The result will be a hierarchically arranged list of the words; a word with highest association with other words, among the first 500 citations, will appear on top of the list. One can click on required word for a display of extracted sentences that contain the specific word.

Anne O'Tate

The tool organizes the results according to pre-defined categories. The categories include the biological terms found in titles or abstracts, topics, journals, authors, publication years and affiliations. One can obtain all the hits containing a specific item by clicking on it. Navigation through the sub-categories helps to easily find citations relevant to specific topics. (*Smalheiser NR, Zhou W, Torvik VI. A tool to support user-driven summarization, drill-down and browsing of PubMed search results. J Biomed Discov Collab. Feb 15;3(1):2 [2008]*)

PubFocus has a very flexible query system and provides statistical information mainly related to authors but can also indicate the NCI ontology and Gene Orthology term distribution across the citations. The tool was too slow to use with higher number of hits.

There are many other search tools with various specializations

PubWindows can be useful for MeSH based collection of citations in a specific broad subject area.

We found several other tools with potential special applications but many of them were either slow/non-responsive or didn't have enough information for a new user.

LitMiner extracts, and performs statistical analysis of co-occurrence of annotated entities (genes, compounds, disorders and tissues), from the abstracts retrieved in response to a query using one of these entities.

A list of many other free, online biomedical text mining tools is available at different websites. However, we were not able to access or use several of these programs due to a) lack of proper user guide, b) they were too slow to respond or c) they were non-responsive. Such tools included Hubmed, PubMed PubReMiner, Concept Link and PubMed Gold.

HubMed (<http://www.hubmed.org>) has been created to provide an alternative interface to the PubMed database that can be more efficient in literature search. It allows ranking and visualization (touchpad) of relations among the selected articles. For some reason, the tool didn't function effectively, particularly the touchpad feature, when we tried. When the tool is functional, one may be able to export citation data in multiple formats, receive daily updates of publications in specific areas, access links to full text and other related resources. It is also possible to retrieve data from formatted bibliography lists, navigate citation links and store annotated metadata for articles of interest. The coverage seems to

be equal to PubMed; the search for the phrase “Transcription Initiation Site” yielded 3249 hits, same as in PubMed. Hubmed also allows same search term settings as PubMed.

There is an option to refine the search terms. It displays the number of hits for every term input by the user and allows deletion of any of them or addition of new terms, after every search. But it lacks the robust operations that can be done with the ‘history’ option of the PubMed. There is a history button in HubMed but it refers to records of full text link-outs made by the user.

There is a function to help the user to identify important keywords in the abstracts: One identifies ‘Gene Ontology’ terms and adds links from each term to the Gene Ontology, UniProt and MEDLINEPlus respectively. The second filter adds links to the appropriate Wikipedia pages (*Alfred D. Eaton HubMed: a web-based biomedical literature search interface Nucleic Acids Research, 2006, Vol. 34, Web Server issue W745–W747*). HubMed also creates easy links to some other important search engines.

FULL-TEXT SEARCH TOOLS:

Google Scholar (GS), Highwire Press (HWP), PubMed Central (PMC) and Scirus can access the selected types of publications from different resources and scan the complete documents rather than scanning only the title, abstract (when available), author, affiliation fields and the journal details. A Google powered and hosted search engine customized for life sciences, (brij.in) can also scan complete documents from different sources.

Paid search tools:

Except Scopus (*Burnham, J.F., Scopus database: a review. Biomed Digit Libr. 3,1 [2006]*), which was providing free trial access earlier, we were unable to use the paid programs for this comparative study. Such programs included Web of Science (*Tomasulo P., Thread your way through ISI's Web of Science <<http://www.webofscience.com>>. Med Ref Serv Q. 20, 49-59 [2001]*) and Ovid (*Fremer, E. & Larsson, B., SPIRS, WinSPIRS, and OVID: a question of free-text versus thesaurus retrieval? Bull Med Libr Assoc. 85, 57-8 [1997]*; *Schoonbaert, D., SPIRS, WinSPIRS, and OVID: a comparison of three MEDLINE-on-CD-ROM interfaces. Bull Med Libr Assoc. 84, 63-70 [1996]*). A few other studies (*Bakkalbasi, N., Bauer, K., Glover, J., & Wang, L., Three options for citation tracking: Google Scholar, Scopus and Web of Science. Biomed. Digit. Libr. 3, 7 [2006]*; *Falagas, M.E., Pitsouni, E.I., Malietzis, G.A. & Pappas G. Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses. FASEB J. 22, 338-42 [2008]*; *De Groote, S.L., PubMed, Internet Grateful Med, and Ovid: a comparison of three MEDLINE Internet interfaces. Med Ref Serv Q. 19:1-13 [2000]*; *Gall, C. & Brahmī, F. A., Retrieval comparison of EndNote to search MEDLINE [Ovid and PubMed] versus searching them directly. Med Ref Serv Q. 23, 25-32 [2004]*) have compared certain aspects of some of the paid tools.

OTHER DATABASES/SEARCH ENGINES RELATED TO CLINICAL ACTIVITIES:

Allied and Complementary Medicine (AMED): *The database covers a selection of journals in complementary (alternative) medicine, palliative care and other professions allied to medicine.* It is produced by the Health Care Information Service of the British Library.

<http://www.bl.uk/collections/health/amed.html>

Can also accessed through: <http://www.ovid.com/site/catalog/DataBase/12.jsp>

BIOSIS: BIOSIS information solutions are provided by **Thomson Scientific**. They have databases for finding life sciences information. More than 600,000 new entries are made each year.

<http://www.biosis.org/>

CancerLit: An interface to PubMed provided by the National Cancer Institute of USA.

http://www.cancer.gov/search/search_cancertopics.aspx

CAMbase: It allows easy search of literature of Complementary and Alternative Medicine (CAM) in different resources.

http://cambase.dmz.uni-wh.de/opencam/start_en.html

Cochrane Central Register of Controlled Trials (CENTRAL): contains approx. 500,000 records towards “independent evidence to inform healthcare decision-making” (from Cochrane and other systematic reviews, clinical trials etc).

http://www.mrw.interscience.wiley.com/cochrane/cochrane_clcentral_articles_fs.html

Cumulative Index to Nursing and Allied Health (CINAHL): This database covers nursing and allied health.

<http://www.cinahl.com/prodsvcs/prodsvcs.htm>

EMBASE: This is a gateway for pharmacological, biomedical literature (more than 18 million records)

<http://www.embase.com/>

metaRegister of Controlled Trials (mRCT): *meta*Register of Controlled Trials aims to promote the availability and exchange of information about randomised controlled trials in all areas of healthcare.

<http://www.controlled-trials.com/mrct/>

An international database that combines registers of information about ongoing and some completed registers of controlled trials throughout the world. This mRCT is not comprehensive in itself but has links to other sites.

<http://www.lib.gla.ac.uk/Resources/Databases/metaregister.shtml>

National Research Register (NRR) Archive: A searchable copy of the records that were earlier in the National Research Register (NRR) Projects Database; created from the final issue of the NRR published in October 2007 (records collected up to September 2007). NRR had a public, searchable list of research activity in the UK funded by, or of interest to, the National Health Service.

Introduction: <https://portal.nihr.ac.uk/Pages/NRRArchive.aspx#what>

Search page: <https://portal.nih.ac.uk/Pages/NRRArchiveSearch.aspx>

A compilation of major tools and their features are available at:

<http://myuminfo.umanitoba.ca/index.asp?sec=857&too=100&dat=7/3/2007&sta=3&wee=1&eve=8&npa=12437>

http://biocreative.sourceforge.net/bionlp_tools_links.html

LIST OF OTHER INFORMATION RETRIEVAL TOOLS:

PAID:

1. **NextBio** is a paid web-based data search engine.

URL: <http://www.nextbio.com/b/home/home.nb>

2. **OvidSP** and other paid tools from Ovid.

URL: <http://www.ovid.com/site/products/tools/index.jsp?top=2&mid=3&bottom=8>

FREE:

3. **LitMiner** extracts, and performs statistical analysis of co-occurrence of annotated entities, from the abstracts retrieved in response to a query using one of these entities.

URL: <http://andromeda.gsf.de/litminer?choice=home>

4. **Textpresso** is an organism specific literature search tool that enables the user to search for one or a combination of categories of biological concepts and classes, and/or keywords within an entire literature.

URL: <http://www.textpresso.org/>

5. **Pubgene** can retrieve associate information on genes, proteins, biological terms [chemicals (based on Pubmed chemicals), molecular processes, biological functions and cellular components) based on GO & diseases (based on MeSH)].

URL: <http://www.pubgene.org/>

6. **iHOP** (information hyperlinked over proteins) (*Hoffmann, R., Valencia, A. A Gene Network for Navigating the Literature. Nature Genetics 36, 664 [2004]*), which allows searching the literature with a gene name and extracts the related biological information.

URL: <http://www.ihop-net.org/UniPub/iHOP/>

7. iProLink (integrated Protein Literature, INformation and Knowledge) (*iProLINK: an integrated protein resource for literature mining. Hu ZZ, Mani I, Hermoso V, Liu H, Wu CH. Comput Biol Chem. 2004 Dec;28(5-6):409-16*) tries to provide curated data sources for text mining research in the areas of bibliography mapping, annotation extraction, protein named entity recognition, and protein ontology development. It provides multiple start options and provides easy links to existing information.

URL: <http://pir.georgetown.edu/iprolink/>

8. Chilibot can search PubMed abstracts for specific relationships between proteins, genes, or keywords. It can also show the results as a graph.

URL: <http://www.chilibot.net/>

9. Bitola is a biomedical discovery support system that aims assist the biomedical researchers to detect potential relations between biomedical concepts (currently contains MeSH and HUGO, the Human Genome Organisation, <http://www.hugo-international.org/>).

URL: <http://www.mf.uni-lj.si/bitola/>

10. G2D (candidate Genes to [2] inherited Diseases) is a platform that allows searching human genomic regions for genes related to an inherited diseases.

URL: http://www.ogic.ca/projects/g2d_2/

11. Prolinks is a collection of inference methods for predicting functional linkages between proteins with search options.

URL: <http://128.97.39.94/cgi-bin/functionator/pronav>

12. STRING is a database and tool *for the retrieval of direct and indirect* protein-protein interactions associations derived from genomic context, high-throughput experiments, (conserved) co-expression and previous knowledge.

URL: <http://string.embl.de/>

13: eBiosci is a portal general biological information, including news, career, gene, drug discovery etc.

URL: <http://www.bioportfolio.com/> (link non-functional occasionally)

14. PubReMiner: PubReMiner can query pubmed with the query, retrieve citations and generate frequency tables based on journals, authors, words in T/A, MeSH headings etc. But we haven't been able to use this tool successfully, except for gene name search.

URL: <http://bioinfo.amc.uva.nl/human-genetics/pubreminer/>

15. PreBIND: Supposed to be searching for a protein or PMID.

URL: <http://prebind.bind.ca/> - not accessible/functional

Several gene information analysis tools related to microarray are offered by the **Genomics and Bioinformatics group of NCI, NIH:** <http://discover.nci.nih.gov/tools.jsp>

SpliceCenter - to check the impact of gene splice variation on common molecular biology technologies including RT-PCR, RNAi, expression microarrays, and peptide-based assays.

CIMminer - to generate color-coded Clustered Image Maps (CIMs) (“heat maps”) representing gene expression profiles & other similar high-dimensional data sets.

LeFEminer - to interpret gene microarray data

AffyProbeMiner - to re-define chip definition files (CDFs) for Affymetrix chips in view of the most recent genomic sequence information.

SpliceMiner -for an intuitive non-redundant display of splice variants

CellMiner - for molecular profile information on the NCI 60 human cancer cell lines and the DU145/RC0.1 prostate cancer cell line pair.

AbMiner - for information on antibodies screened specificity against the NCI-60 cancer cell lines.

MatchMiner – to translate gene identifier types for large number of genes (GenBank accession numbers, IMAGE clone IDs, common gene names, HUGO names, gene symbols, UniGene clusters, etc)

GoMiner –to obtain gene ontology information after microarray

High-Throughput GoMiner –for automatic analysis of multiple microarrays and integrate results across all of the microarrays.

SmudgeMiner – to highlight regional biases and other artifacts on Affymetrix and other microarrays to enable quality assessment.

MedMiner – to search and organize the biomedical literature on genes, gene-gene relationships, and gene-drug relationships.

We have only provided an indicative list of the bio-literature search tools of general and specialized nature.

Related articles published when the manuscript preparation was complete/being completed:

1. Rzhetsky, A., Seringhaus, M. & Gerstein, M., Seeking a new biology through text mining. *Cell* **134**, 9-13 (2008).
Note: The article reviews the significance of text-mining methods
2. Huang ZX, Tian HY, Hu ZF, Zhou YB, Zhao J, Yao KT. GenCLiP: a software program for clustering gene lists by literature profiling and constructing gene co-occurrence networks related to custom keywords. *BMC Bioinformatics*. **13**, 9(1):308. [Epub ahead of print] 2008
3. Jelier R, Schuemie MJ, Veldhoven A, Dorssers LC, Jenster G, Kors JA. Anni 2.0: a multipurpose text-mining tool for the life sciences. *Genome Biol*. 2008 Jun 12;9(6):R96. [Epub ahead of print]
Note: URL: <http://biosemantics.org/anni>
4. Theodosiou T., Darzentas, N., Angelis, L., Ouzounis, C.A., PuReD-MCL: a graph-based PubMed document clustering methodology *Bioinformatics*. [Epub ahead of print] Jul 1 (2008).

Note: PuReD-MCL clusters the documents, highlights important clues, and allows to visualize the clusters as well as all relevant

information. We failed to access the URL provided <http://tartara.cs-d.auth.gr/~theodos> for source code.

5. Hur J, Wild DJ. PubChemSR: A search and retrieval tool for PubChem. Chem Cent J. 2008 May 15;2:11.
6. Nakazato T, Takinaka T, Mizuguchi H, Matsuda H, Bono H, Asogawa M. BioCompass: a novel functional inference tool that utilizes MeSH hierarchy to analyze groups of genes. In Silico Biol. 2008;8(1):53-61.
7. Yu W, Clyne M, Dolan SM, Yesupriya A, Wulf A, Liu T, Khoury MJ, Gwinn M. GAPscreener: an automatic tool for screening human genetic association literature in PubMed using the support vector machine technique. BMC Bioinformatics. 2008 Apr 22;9:205.
8. Bracke PJ, Howse DK, Keim SM. Evidence-based Medicine Search: a customizable federated search engine. J Med Libr Assoc. 2008 Apr;96(2):108-13.
9. Richter M, Lombardot T, Kostadinov I, Kottmann R, Duhaime MB, Peplies J, Glöckner FO. JCoast - a biologist-centric software tool for data mining and comparison of prokaryotic (meta)genomes. BMC Bioinformatics. 2008 Apr 1;9:177.
10. States DJ, Ade AS, Wright ZC, Bookvich AV, Athey BD. MiSearch Adaptive PubMed Search Tool. Bioinformatics. 2008 Mar 11. [Epub ahead of print]
11. Gladki A, Siedlecki P, Kaczanowski S, Zielenkiewicz P. e-LiSe--an online tool for finding needles in the '(Medline) haystack'. Bioinformatics. 2008 Apr 15;24(8):1115-7. Epub 2008 Mar 5.
12. Campagne F. Objective and automated protocols for the evaluation of biomedical search engines using No Title Evaluation protocols. BMC Bioinformatics. 2008 Feb 29;9:132.
13. Caporaso JG, Deshpande N, Fink JL, Bourne PE, Cohen KB, Hunter L. Intrinsic evaluation of text mining tools may not predict performance on realistic tasks. Pac Symp Biocomput. 2008:640-51.
14. Guemes-Careaga I, Gutiérrez-Ibarluzea I. Biomedical literature search protocols: consensus statement from the documentation units of the Spanish health technology assessment agencies. Int J Technol Assess Health Care. 2008 Winter;24(1):104-11.

Other related literature:

Listed under following sub-headings:

- A. Tutorials or studies of search tool features**
- B. Literature search** (comments on search process and/or its significance)
- C. Comparison of search options/tools**
- D. New resources (tools or databases)**

A. Tutorials or studies of search tool features:

1. Chang AA, Heskett KM, Davidson TM. Searching the literature using medical subject headings versus text word with PubMed. Laryngoscope. 2006 Feb;116(2):336-40.

2. Coletti MH, Bleich HL. Medical subject headings used to search the biomedical literature. *J Am Med Inform Assoc.* 2001 Jul-Aug;8(4):317-23. Erratum in: *J Am Med Inform Assoc* 2001 Nov-Dec;8(6):597.
3. Motschall E, Falck-Ytter Y. Searching the MEDLINE literature database through PubMed: a short guide. *Onkologie.* 2005 Oct;28(10):517-22. Epub 2005 Aug 19. Review.
4. Ravenscroft EF. What can bibliographic software offer you? *CANNT J.* 2000 Oct-Dec;10(4):43-4.
5. Shultz M. Mapping of medical acronyms and initialisms to Medical Subject Headings (MeSH) across selected systems. *J Med Libr Assoc.* 2006 Oct;94(4):410-4.
6. Sood A, Erwin PJ, Ebbert JO. Using advanced search tools on PubMed for citation retrieval. *Mayo Clin Proc.* 2004 Oct;79(10):1295-9; quiz 1300.
7. Sood A, Ghosh AK. Literature search using PubMed: an essential tool for practicing evidence- based medicine. *J Assoc Physicians India.* 54, 303-8 (2006)

B. Literature search (comments on search process and/or its significance)

1. Chaussabel D. Biomedical literature mining: challenges and solutions in the 'omics' era. *Am J Pharmacogenomics.* 2004;4(6):383-93.
2. Cooper CP, Mallon KP, Leadbetter S, Pollack LA, Peipins LA. Cancer Internet search activity on a major search engine, United States 2001-2003. *J Med Internet Res.* 2005 Jul 1;7(3):e36.
3. Daumke P, Markü K, Poprat M, Schulz S, Klar R. Biomedical information retrieval across languages. *Med Inform Internet Med.* 2007 Jun;32(2):131-47.
4. de Bruijn B, Martin J. Getting to the (c)ore of knowledge: mining biomedical literature. *Int J Med Inform.* 2002 Dec 4;67(1-3):7-18.
5. Flammini A, Menczer F, Vespignani A. Topical interests and the mitigation of search engine bias. *Proc Natl Acad Sci U S A.* 2006 Aug 22;103(34):12684-9. Epub 2006 Aug 10.
6. King RV, Murphy-Cullen CL, Mayo HG, Marcee AK, Schneider GW. Use of computers and the Internet by residents in US family medicine programmes. *Med Inform Internet Med.* 2007 Jun;32(2):149-55.
7. Lau AY, Coiera EW. Do people experience cognitive biases while searching for information? *J Am Med Inform Assoc.* 2007 Sep-Oct;14(5):599-608. Epub 2007 Jun 28.
8. Lee BY, Chen EH. Health care information provided by internet search engines. *Fam Med.* 37(5):312 (2005).
9. Leung A, Ko P, Chan KS, Chi I, Chow N. Searching health information via the web: Hong Kong chinese older adults' experience. *Public Health Nurs.* 2007 Mar-Apr;24(2):169-75.
10. Müller M, Markó K, Daumke P, Paetzold J, Roesner A, Klar R. Biomedical data mining in clinical routine: expanding the impact of hospital information systems. *Stud Health Technol Inform.* 2007;129(Pt 1):340-4.
11. Nenadic G, Spasic I, Ananiadou S. Terminology-driven mining of biomedical literature. *Bioinformatics.* 2003 May 22;19(8):938-43.

12. Patrick LJ, Munro S. The literature review: demystifying the literature search. *Diabetes Educ.* 2004 Jan-Feb;30(1):30-4, 36-8.
13. Rein DC. Developing library bioinformatics services in context: the Purdue University Libraries bioinformationist program. *J Med Libr Assoc.* 2006 Jul;94(3):314-20, E193-7.
14. Rodgers RP. Searching for biomedical information on the World Wide Web. *J Med Pract Manage.* 2000 May-Jun;15(6):306-13.
15. Scherf M, Epple A, Werner T. The next generation of literature analysis: integration of genomic analysis into text mining. *Brief Bioinform.* 2005 Sep;6(3):287-97.
16. Shanahan MC. Transforming information search and evaluation practices of undergraduate students. *Int J Med Inform.* 2008 Aug;77(8):518-26. Epub 2007 Nov 26.
17. Shatkay H, Feldman R. Mining the biomedical literature in the genomic era: an overview. *J Comput Biol.* 2003;10(6):821-55.
18. Siegel ER, Logan RA, Harnsberger RL, Cravedi K, Krause JA, Lyon B, Hajarian K, Uhl J, Ruffin A, Lindberg DA. Information Rx: Evaluation of a new informatics tool for physicians, patients, and libraries. *Inf Serv Use.* 2006;26(1):1-10.
19. Silfen EZ, Patel CO. Trusted and relevant medical knowledge: the promise of information retrieval in biomedicine. *J Healthc Inf Manag.* 2006 Fall;20(4):74-82.
20. Singh PM, Wight CA, Sercinoglu O, Wilson DC, Boytsov A, Raizada MN. Language preferences on websites and in Google searches for human health and food information. *J Med Internet Res.* 2007 Jun 28;9(2):e18.
21. Stronge AJ, Rogers WA, Fisk AD. Web-based information search and retrieval: effects of strategy use and age on search success. *Hum Factors.* 2006 Fall;48(3):434-46.

C. Comparison of search options/tools:

1. Aguillo IF. [Advanced tools for searching for medical information on the web] *Aten Primaria.* 2002 Mar 15;29(4):246-52. Spanish
2. Campbell R, Ash J. Comparing bedside information tools: a user-centered, task-oriented approach. *AMIA Annu Symp Proc.* 2005:101-5.
3. Caro-Rojas RA, Eslava-Schmalbach JH. Agreement between Medline searches using the Medline-CD-Rom and Internet Pubmed, BioMedNet, Medscape and Gateway search-engines. *Rev Salud Publica (Bogota).* 2005 May-Aug;7(2):214-26.
4. Dong P, Loh M, Mondry A. Relevance similarity: an alternative means to monitor information retrieval systems. *Biomed Digit Libr.* 2005 Jul 20;2:6.
5. Hauser SE, Demner-Fushman D, Ford GM, Jacobs JL, Thoma G. Preliminary comparison of three search engines for point of care access to MEDLINE citations. *AMIA Annu Symp Proc.* 945 (2006).
6. Kagolovsky Y, Moehr JR. A new look at information retrieval evaluation: proposal for solutions. *J Med Syst.* 2004 Feb;28(1):103-16.

7. Lorence D, Abraham J. Analysis of semantic search within the domains of uncertainty: using Keyword Effectiveness Indexing as an evaluation tool. *Int J Electron Healthc.* 2006;2(3):263-76.
8. Neveol A, Zeng K, Bodenreider O. Besides precision & recall: exploring alternative approaches to evaluating an automatic indexing tool for MEDLINE. *AMIA Annu Symp Proc.* 2006:589-93.
9. Sampson M, Barrowman NJ, Moher D, Clifford TJ, Platt RW, Morrison A, Klassen TP, Zhang L. Can electronic search engines optimize screening of search results in systematic reviews: an empirical study. *BMC Med Res Methodol.* 2006 Feb 24;6:7.
10. Sutton VR, Hauser SE. Preliminary comparison of the Essie and PubMed search engines for answering clinical questions using MD on Tap, a PDA-based program for accessing biomedical literature. *AMIA Annu Symp Proc.* 1128, (2005).
11. Van Mulligen EM, Van Der Eijk C, Kors JA, Schijvenaars BJ, Mons B. Research for research: tools for knowledge discovery and visualization. *Proc AMIA Symp.*
12. Weeber M, Kors JA, Mons B. Online tools to support literature-based discovery in the life sciences. *Brief Bioinform.* 2005 Sep;6(3):277-86.
13. Wren JD. 404 not found: the stability and persistence of URLs published in MEDLINE. *Bioinformatics.* 2004 Mar 22;20(5):668-72. Epub 2004 Jan 22.
14. Zweigenbaum P, Darmoni SJ, Grabar N, Douyère M, Benichou J. An assessment of the visibility of MeSH-indexed medical web catalogs through search engines. *Proc AMIA Symp.* 2002:954-8.

D. New resources (tools or databases)

1. Atamas SP. FCP (<http://fibro.biobitfield.com/fcp.php>): a bioinformatic tool assisting in PubMed searches for literature on fibrosis-related cytokines. *Arthritis Rheum.* 2003 Jul;48(7):2083-4.
2. Bajic VB, Veronika M, Veladandi PS, Meka A, Heng MW, Rajaraman K, Pan H, Swarup S. Dragon Plant Biology Explorer. A text-mining tool for integrating associations between genetic and biochemical entities with genome annotation and biochemical terms lists. *Plant Physiol.* 2005 Aug;138(4):1914-25.
3. Becker KG, Hosack DA, Dennis G Jr, Lempicki RA, Bright TJ, Cheadle C, Engel J. PubMatrix: a tool for multiplex literature mining. *BMC Bioinformatics.* 2003 Dec 10;4:61. Note: URL: <http://pubmatrix.grc.nia.nih.gov/>
4. Bohne-Lang A, Lang E, Taube A. PMD2HD--a web tool aligning a PubMed search results page with the local German Cancer Research Centre library collection. *Biomed Digit Libr.* 2005 Jun 27;2:4.
5. Can AB, Baykal N. MedicoPort: a medical search engine for all. *Comput Methods Programs Biomed.* 2007 Apr;86(1):73-86. Epub 2007 Feb 22.
6. Corney DP, Buxton BF, Langdon WB, Jones DT. BioRAT: extracting biological information from full-length papers. *Bioinformatics.* 2004 Nov 22;20(17):3206-13. Epub 2004 Jul 1.
7. Corsi C, Ferragina P, Marangoni R. The BioPrompt-box: an ontology-based clustering tool for searching in biological databases. *BMC Bioinformatics.* 2007 Mar 8;8 Suppl 1:S8.

8. Donaldson I, Martin J, de Bruijn B, Wolting C, Lay V, Tuekam B, Zhang S, Baskin B, Bader GD, Michalickova K, Pawson T, Hogue CW. PreBIND and Textomy--mining the biomedical literature for protein-protein interactions using a support vector machine. *BMC Bioinformatics*. 2003 Mar 27;4:11. Epub 2003 Mar 27.
9. Douyère M, Soualmia LF, Névéal A, Rogozan A, Dahamna B, Leroy JP, Thirion B, Darmoni SJ. Enhancing the MeSH thesaurus to retrieve French online health resources in a quality-controlled gateway. *Health Info Libr J*. 2004 Dec;21(4):253-61.
10. Dracos A. [SIBIL: an information tool for the information retrieval on bioethics] *Ann Ist Super Sanita*. 2004;40(3):283-6. Italian.
11. Fontelo P, Liu F, Leon S, Anne A, Ackerman M. PICO Linguist and BabelMeSH: development and partial evaluation of evidence-based multilanguage search tools for MEDLINE/PubMed. *Stud Health Technol Inform*. 2007;129(Pt 1):817-21.
12. Gaudinat A, Joubert M, Aymard S, Falco L, Boyer C, Fieschi M. WRAPIN: new generation health search engine using UMLS knowledge sources for MeSH term extraction from health documentation. *Stud Health Technol Inform*. 2004;107(Pt 1):356-60.
13. Goetz T, von der Lieth CW. PubFinder: a tool for improving retrieval rate of relevant PubMed abstracts. *Nucleic Acids Res*. 2005 Jul 1;33(Web Server issue):W774-8.
14. Hanauer DA. EMERSE: The Electronic Medical Record Search Engine. *AMIA Annu Symp Proc*. 2006:941.
15. Hearst MA, Divoli A, Guturu H, Ksikes A, Nakov P, Wooldridge MA, Ye J. BioText Search Engine: beyond abstract search. *Bioinformatics*. 2007 Aug 15;23(16):2196-7. Epub 2007 Jun 1.
16. Hersh WR, Donohoe LC. SAPHIRE International: a tool for cross-language information retrieval. *Proc AMIA Symp*. 1998:673-7.
17. Ide NC, Loane RF, Demner-Fushman D. Essie: a concept-based search engine for structured biomedical text. *J Am Med Inform Assoc*. 2007 May-Jun;14(3):253-63. Epub 2007 Feb 28.
18. Ingsriswang S, Pacharawongsakda E. sMOL Explorer: an open source, web-enabled database and exploration tool for Small MOLEcules datasets. *Bioinformatics*. 2007 Sep 15;23(18):2498-500. Epub 2007 Jul 27.
19. Jang H, Lim J, Lim JH, Park SJ, Lee KC. BioProber: software system for biomedical relation discovery from PubMed. *Conf Proc IEEE Eng Med Biol Soc*. 2006;1:5779-82.
20. Kahn CE Jr. A digital library of radiology images. *AMIA Annu Symp Proc*. 2006:972.
21. Kingsland LC 3rd, Prettyman MF, Shooshan SE. The NLM Gateway: a metasearch engine for disparate resources. *Stud Health Technol Inform*. 2004;107(Pt 1):52-6.
22. Lahm A, de Rinaldis E. PatentMatrix: an automated tool to survey patents related to large sets of genes or proteins. *Source Code Biol Med*. 2007 Sep 6;2:4.

23. Lee H, Yi GS, Park JC. E3Miner: a text mining tool for ubiquitin-protein ligases. *Nucleic Acids Res.* 2008 Jul 1;36(Web Server issue):W416-22. Epub 2008 May 15.
24. Liebel U, Kindler B, Pepperkok R. 'Harvester': a fast meta search engine of human protein resources. *Bioinformatics.* 2004 Aug 12;20(12):1962-3. Epub 2004 Feb 26.
25. Lin SM, McConnell P, Johnson KF, Shoemaker J. MedlineR: an open source library in R for Medline literature data mining. *Bioinformatics.* 2004 Dec 12;20(18):3659-61. Epub 2004 Jul 29.
26. Matthiessen MW. BioWareDB: the biomedical software and database search engine. *Bioinformatics.* 2003 Nov 22;19(17):2319-20.
27. Moskovitch R, Hessing A, Shahar Y. Vaidurya--a concept-based, context-sensitive search engine for clinical guidelines. *Stud Health Technol Inform.* 2004;107(Pt 1):140-4.
28. Oliver DE, Bhalotia G, Schwartz AS, Altman RB, Hearst MA. Tools for loading MEDLINE into a local relational database. *BMC Bioinformatics.* 2004 Oct 7;5:146.
29. Page RD. A Taxonomic Search Engine: federating taxonomic databases using web services. *BMC Bioinformatics.* 2005 Mar 9;6:48.
30. Phillips J, Chilukuri R, Fragoso G, Warzel D, Covitz PA. The caCORE Software Development Kit: streamlining construction of interoperable biomedical information services. *BMC Med Inform Decis Mak.* 2006 Jan 6;6:2.
31. Poulter GL, Rubin DL, Altman RB, Seoighe C. MScanner: a classifier for retrieving Medline citations. *BMC Bioinformatics.* 2008 Feb 19;9:108.
32. Riethof DA, Balakrishnan R. LabVelocity: online tools for life science products, protocols, technical information, MEDLINE searches, and laboratory calculations. *Biotechniques.* 2001 Jun;30(6):1310-5.
33. Textpresso: an ontology-based information retrieval and extraction system for biological literature. Müller HM, Kenny EE, Sternberg PW. *PLoS Biol.* 2004 Nov;2(11):e309. Epub 2004 Sep 21.
34. Tu Q, Tang H, Ding D. MedBlast: searching articles related to a biological sequence. *Bioinformatics.* 2004 Jan 1;20(1):75-7.
35. Viksna J, Celms E, Opmanis M, Podnieks K, Rucevskis P, Zarins A, Barrett A, Neogi SG, Krestyaninova M, McCarthy MI, Brazma A, Sarkans U. PASSIM--an open source software system for managing information in biomedical studies. *BMC Bioinformatics.* 2007 Feb 9;8:52.
36. Wilson RJ, Goodman JL, Strelets VB; FlyBase Consortium. FlyBase: integration and improvements to query tools. *Nucleic Acids Res.* 2008 Jan;36(Database issue):D588-93. Epub 2007 Dec 26.
37. Xiang Z, Zheng W, He Y. BBP: Brucella genome annotation with literature mining and curation. *BMC Bioinformatics.* 2006 Jul 16;7:347.
38. Xuan W, Wang P, Watson SJ, Meng F. Medline search engine for finding genetic markers with biological significance. *Bioinformatics.* 2007 Sep 15;23(18):2477-84. Epub 2007 Sep 6.
39. Yamamoto Y, Takagi T. OReFiL: an online resource finder for life sciences. *BMC Bioinformatics.* 2007 Aug 6;8:287

40. Yoneya T. PSE: a tool for browsing a large amount of MEDLINE/PubMed abstracts with gene names and common words as the keywords. *BMC Bioinformatics*. 2005 Dec 10;6:295.
41. Zhang CB, Li YJ, Luo JB, Lü XS. [Search engine of genetics journals] *Yi Chuan*. 2005 Mar;27(2):262-70. [Article in Chinese]
42. Zheng H, Shi J, Fang X, Li Y, Vang S, Fan W, Wang J, Zhang Z, Wang W, Kristiansen K, Wang J. FGF: a web tool for Fishing Gene Family in a whole genome database. *Nucleic Acids Res*. 2007 Jul;35(Web Server issue):W121-5. Epub 2007 Jun 21.
43. Zhou Q, Zhang H, Geng M, Zhang C. A real-time and dynamic biological information retrieval and analysis system BIRAS). *Genomics Proteomics Bioinformatics*. 2003 Feb;1(1):74-7.