



Finding similar: Analyzing similarity functions in Ovid Medline and PubMed

Rebecca S. Graves & Gabriel M. Peterson

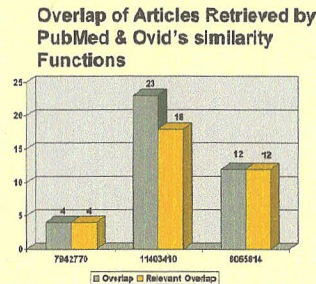
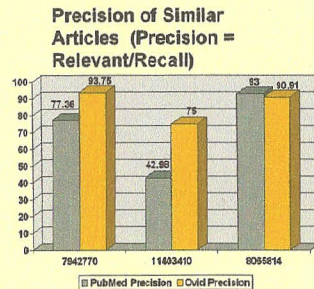
University of Missouri-Columbia J. Otto Lottes Health Sciences Library

Objectives

- While database producers offer functions that allow searchers to automatically find related articles, how well these functions work or compare is unknown.
- This project explored how the PubMed's Related Articles and Ovid's Find Similar similarity functions work and how they compare to each other.

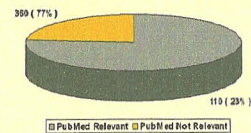
Methods

- A search was run for an article on both Ovid and PubMed and the similarity function was executed on the same citation on each system.
- Articles were restricted to 1965-2005 and English language.
- Sets retrieved by the similarity function were judged for relevance only if they returned between 15-300 hits on the systems.
- Relevance was evaluated by a pair of judges and arbitrated by a third.
- Overlap, precision and relative recall were calculated.
- Relative recall was calculated by combining the relevant items for both systems and treating it as total recall.



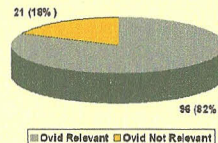
Relevance of Similar Articles in PubMed

(PubMed finds similar articles using Global & Local weighted terms drawn from title, abstract & MeSH fields.)

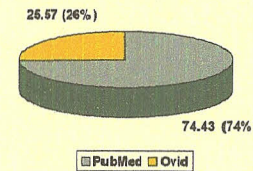


Relevance of Similar Articles in Ovid

(Ovid finds similar articles using a proprietary natural language algorithm drawing terms from the title field.)



Relative Recall of Similarity Functions in PubMed & Ovid



- Among 135 searches run, a total of 3 queries returned 15 to 200 related articles from PubMed & Ovid.
- These three searches retrieved a total of 117 articles from Ovid and 470 articles from PubMed.
- Among the 117 Ovid Articles, 96 (82%) were relevant.
- Among the 470 PubMed articles, 360 (76%) were relevant.
- Relative recall was 74% for PubMed and 26% for Ovid.
- Thus, when finding similar articles, Ovid offers higher precision, but PubMed offers higher recall.

Results and Conclusions

- The two similarity functions differ in their processing and presentation of results.
- When three sets were compared, overlap was found to be low (between 4%-11%).
- Precision varied between 75%-94% on Ovid and between 43%-93% on PubMed.
- Relative recall was 50%-100% for PubMed and 12%-50% for Ovid.
- When a user requires high recall, PubMed is obviously the system to use.
- However, when a user requires only a few relevant items, the lower recall and comparable or better precision with Ovid suggests it might be a better choice.

*Analyzing Similarity Functions in Two Systems, OVID and PubMed**

By

Rebecca Graves, MLS, Education Librarian, J Otto Lottes Health Sciences Library, MU Libraries, University of Missouri, Columbia, MO 65211, gravesr@health.missouri.edu

Gabriel Peterson, Ph.D. Assistant Professor, North Carolina Central University, Durham, NC, gpeterson@nccu.edu

MaryEllen Cullinan Sievert, Ph.D., Professor Emerita, School of Information Science and Technology, Research Consultant, J Otto Lottes Health Sciences Library, MU Libraries, University of Missouri, Columbia, MO, sievertm@missouri.edu

*A version of this paper was presented as a poster at MLA '07 the Annual Conference of the Medical Library Association, May, 2007, Philadelphia, Pennsylvania

INTRODUCTION

In the early days of online information retrieval when most searches were mediated, librarians often asked their patrons to provide them with a known relevant paper. The thinking was that this paper could give the librarian clues to find similar papers. It was often a successful method to structure the first stop of an interactive search. Now, however, with so many end users doing their own searching, vendors have begun to provide a similarity function so the user can just click a button to find other articles once he/she has found a relevant article. The principles behind these functions are like those used by librarians doing a mediated search beginning with a known relevant paper.

Several vendors offer such similarity functions. For this paper we are restricting our research to two major vendors available at our library—MEDLINE/PubMed and MEDLINE/OVID. These two vendors name their functions differently (“Related articles” for PubMed and “Find similar” for OVID) but the underlying idea is the same for both.

BACKGROUND

Those wanting to know how the function works can find out easily via PubMed. A complete explanation of the algorithm can be found at <http://www.ncbi.nlm.nih.gov/entrez/query/static/computation.html>. In brief, the similarity feature uses words from the title, abstract and MeSH fields to determine related articles. Stop words are discounted (for a complete list go to

[http://www.ncbi.nlm.nih.gov/books/bv.fcgi?](http://www.ncbi.nlm.nih.gov/books/bv.fcgi?highlight=stopwords&rid=helppubmed.table.pubmedhelp.T43)

[highlight=stopwords&rid=helppubmed.table.pubmedhelp.T43](http://www.ncbi.nlm.nih.gov/books/bv.fcgi?highlight=stopwords&rid=helppubmed.table.pubmedhelp.T43)). Words are calculated for both local and global weights. Local weight is calculated from how often the word appears within the citation in relation to the total number of words. Global weight is calculated from the number of different documents within the MEDLINE database that contain the word. Relatedness is pre-calculated daily

OVID, a private company, provides a partial explanation of their similarity algorithm on their web site. Additional information was supplied by a staff person at Ovid. The Ovid similarity feature uses words from the title field only. These terms are supplemented with terms from the UMLS thesaurus and other dictionaries. There are no stop words. Relatedness is calculated on the fly and ranked on a scale of 1-5.

Two important features to know about the OVID similarity function are their use of rankings and their elimination of long titles. Only citations ranking 4 or 5, i.e. the two highest relevance ranks in relatedness are returned to the searcher. Similarity is not calculated for citations with titles of 22 or more words. Combined, these last two features can account for the high return of null sets.

METHODS

The high number of null sets retrieved by OVID's similarity function made the selection of a sample difficult. We began with a set of 135 records retrieved from a regular search. Then we used the similarity functions on PubMed and OVID to create our sample. Only 3 records (PMID 7942770; PMID 11403410; PMID 8065814) were selected for further

analysis. For the remainder of this paper we refer to these as *Item 1*, *Item 2* and *Item 3* respectively. We set 2 criteria for a record to qualify for the further study. On both systems it must have retrieved at least 15 related citations but no more than 300. These numbers were set to give us feasible sets to work with. During the selection process we found many sets returning 10 or fewer on one system and 500+ on the other. In addition to these criteria we imposed restrictions by date and language to insure that we were retrieving from the same database. Because PubMed and OVID update their systems at different times it was advisable to use the restriction that the related records were published between 1965 and 2005. Further because we know that some of the items frequently added several month or years after publication were those in some languages using non-Roman alphabets, we added a restriction to the English language.

To analyze the results for comparative purposes we needed to make relevance judgments on the retrieved sets. The three authors discussed what would be relevant for each of the sets and formulated a question for each Item. The first 2 authors then independently judged each set. For records about which they disagreed the third author, who had been part of the discussion, served as adjudicator.

Once the relevance judgments were completed, we were able to compute two measures of retrieval success: precision and relative recall. We also computed the overlap between the retrieved sets for each of the three Items in our sample

To compute precision we used the traditional formula:

$$\frac{\text{\# of relevant items retrieved}}{\text{total \# of documents retrieved.}}$$

Since traditional recall depends on knowing all the relevant documents in the database we used a form of relative recall. We calculated the ratio of the number of relevant documents retrieved by 1 system divided by the number of relevant documents retrieved by both systems. We felt that the limitations we had imposed on our sets meant that any relevant document could have been retrieved on either system. That is, for OVID we computed the number of relevant records retrieved by the similarity function on OVID and divided that by the number of relevant records retrieved by both OVID and Pub/Med and the reverse nominator for PubMed.

of relevant items retrieved by OVID

of relevant items retrieved by both OVID and PubMed

Overlap was computed for each record in our sample as the number of relevant items retrieved by both systems. Our overlap analysis, like our relevant recall denominator, depends on the assumption that any record retrieved by one system could have been retrieved by the other.

RESULTS

The three Items retrieved a total of 117 related articles in Ovid MEDLINE and 470 related articles in PubMed. Item 1 retrieved 106 records from PubMed and 16 from OVID. Item 2 retrieved 121 records from PubMed and 68 from OVID. Item 3 in our sample retrieved 243 records from PubMed and 33 from OVID.

Table 1 displays the precision and relative for each of the Items in the sample. PM1 indicates the results for searching for Item 1 on PubMed and OV1 is the same originating citation searched on OVID. The number of items possible reflects the discussion above. That is the sum of the relevant items retrieved by each system.

The average precision was .71 for PubMed and .87 for OVID. The average recall was .75 for PubMed and .26 for OVID.

Overlap between the sets was found to be low. For Item 1 there were 4 relevant items retrieved by both systems and those were the only records from either system. For Item 2 there were 23 records retrieved and 18 of these were relevant. For Item 3 all 12 relevant records were retrieved by both systems. The records in the overlap, therefore, were generally relevant even though there were few items retrieved by both systems.

DISCUSSION AND CONCLUSIONS.

We noted during our searching, that articles with acronyms in the title "confused" the Ovid similarity function. That is, an article on ALS, meaning amyotrophic lateral sclerosis, was just as likely if not more so to retrieve articles on Advanced Life Saving. (PMID 15642900) This problem would be eliminated if the Ovid system included MeSH terms in their calculations.

Our conclusion is that with its relative recall of 74% and relevancy rate of 76%, PubMed's Related Article function is a helpful tool for users, especially those not trained in information science. Ovid with a relevancy rate of 82%, found only a quarter of the possible articles. Given the large return of null sets as discussed earlier and the large

number of relevant articles not found, we recommend significant changes in the Ovid feature and that searchers use their own skills or the PubMed similarity function. When we presented a version of this paper at the Annual Meeting of the Medical Library Association in May 2007 several people told us that when they found a particularly relevant article on OVID they then entered this record into PubMed and used the similarity function to improve their recall.

Considering the disparity in the performance of such high profile literature retrieval systems, we intend to extend this research to examine other prominent sources of health literature. The Medline database is utilized as a source for numerous information retrieval resources; it would be informative to measure the relative effectiveness of these other commonly used systems in identifying related documents from a common source. Further, a comparison of the effectiveness of retrieval systems in finding related documents need not be restricted to the United States; a comparison with prominent literature databases such as the European system Embase would provide potentially valuable information to an increasingly international body of searchers

MCMLA Award Winners

Darrel Willoughby, MLS

Honors and Awards Chair, 2006-2007

VA Medical Library

VA Medical Center

Omaha, Nebraska


darrel.willoughby@va.gov

MCMLA Awards were presented at the opening reception of the 2007 annual meeting, which was held at the McGoogan Library of Medicine.

Deb Ward, the Director of the J Otto Lottes Health Sciences Library at the University of Missouri-Columbia nominated the winner of the Barbara McDowell Award for Excellence in Hospital Librarianship, Susan Centner, the Learning Resource Coordinator for the Missouri Area Health Education Center Mid-Missouri. Susan spent a great deal of time and talent developing a health information access and education program for the Missouri Area Health Education Centers that has since been used as a model for many other Area Health Education Centers around the country.

Siobhan Champ-Blackwell, the NNLM Community Outreach Liaison for the Mid-Continental Region, nominated the winner of the Bernice Hetzner Award for Excellence in Academic Medical Librarianship, Teresa Hartman. Teresa is the Head of Education at McGoogan Library of Medicine of the University of Nebraska Medical Center. She demonstrated an uncanny dedication to the educating of people inside and outside of the University concerning information resources, even as she battled cancer this past year.

Nominated by Jim Bothmer, the winner of the MCMLA Outstanding Achievement Award, Wayne Peay, recently retired after 30 years of service to the Eccles Health Sciences Library at the University of Utah where he started as a student in the bookbinding section and spent 23 years as the Director. Wayne has been a leader in promoting the librarian profession and a tireless fighter for free access to information. He has done

much to promote the MCMLA chapter and the MLA as professional organizations. 

Research Award Winners Earn License

Jeanne Le Ber, MLIS

Research Committee Chair, 2006-2007

Spencer S. Eccles Health Sciences Library

University of Utah


Salt Lake City, Utah

jeannele@lib.med.utah.edu

The MCMLA Research Committee encourages chapter members to engage in research activities by providing mentorship and support. As part of this mission, the Committee advertises the availability of an annual award for the best research paper and poster that is presented at the annual chapter meeting. Research paper and poster presentations are reviewed by the Committee and those judged to be works of research are considered for the \$100 award.

This year the Committee is pleased to acknowledge the research work done by our colleagues at the McGoogan Library of Medicine, University of Nebraska Medical Center-Omaha and the J. Otto Lottes Health Sciences Library, University of Missouri, Columbia. The Research Awardees were recognized at the 3:00 cookie and milk break on Monday, October 15 at the joint-chapter conference.

The best paper was awarded to Lisa Grabenbauer, M.S., Nancy Woelfl, Ph.D., John R. Windle, M.D. and Roslyn S. Fraser-Maginn, M.A. for their paper title *A Qualitative Analysis of Physician Administrator Attitudes Towards Health Information Technology* (presented on Sunday, October 14).

The best poster was awarded to Rebecca S. Graves, MLS, AHIP and Gabriel M. Peterson for their poster titled *Analyzing Similarity Functions in Ovid MEDLINE and PubMed* (presented on Monday, October 15). 



MCMLA Research Award 2007

license to learn; license to lead

Best Research Poster

Awarded to:

Rebecca S. Graves, MLS, AHIP

Gabriel M. Peterson

*Analyzing Similarity Functions in
Ovid MEDLINE and PubMed*