How Similar is Similar? An evaluation of "related articles" applications among health literature portals.

 By Gabriel Peterson, PhD. (North Carolina Central University, School of Library and Information Sciences), and

Rebecca S. Graves, MLS (University of Missouri, J. Otto Lottes Health Sciences Library)

Abstract:
This study is a comparison of the relative effectiveness of similarity-based search features from multiple online health information portals. Similarity searching is the practice of retrieving potentially related information online by identifying a known relevant datum and taking advantage of a utility whose function is to find "related" online information. In the case of online health information resources, use of these applications produces varied results. Online literature retrieval services offer users the ability to identify articles in the biomedical literature that are deemed to be similar or related in some way. Among such portals, there appears to be considerable variability in the degree of precision and recall among documents retrieved using this technique. Retrieval of related articles from Medline through various interfaces produced non-identical lists of articles. This research describes the systematic comparison of similarity functions offered by multiple online health information resources and calculates precision, recall and overlap in order to identify the most effective similarity-searching application using ANOVA analysis.

 Introduction:

Before the advent of widespread online searching, literature searches were frequently mediated by a librarian who would use a document of established relevance as a starting point. With end users doing their own searching, vendors have begun to provide an application that allows a user to find related articles by simply clicking a button. Several vendors offer such similarity functions. For this paper we are restricting our research to four major vendors available at our libraries—MEDLINE/PubMed, MEDLINE/OVID, MEDLINE/EBSCO and Google Scholar. These vendors name their functions differently (e.g. "Related articles" for PubMed and "Find similar" for OVID) but the underlying concept is the same. This study is a comparison of the relative effectiveness of these search features.   Briefly, NLM's related article service harvests words from the title, abstract and MeSH fields to identify related articles. Term scores 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. For more information about how Medline's related articles feature works, readers are directed to the NLM's PubMed helpdesk. (1) OVID, a private company, relies on a proprietary algorithm to identify related articles. The Ovid similarity feature uses words from the title field only. The algorithm uses terms from the UMLS thesaurus and other dictionaries to look for related concepts. There are no stop words. Relatedness is calculated on the fly and ranked on a scale of 1-5. EBSCO compares the major subjects of the article with other articles and selects those with keywords in the text that match those subjects. (2)

Methods:

Researchers performed a comparison of related articles based on the following Medline portals: PubMed, OVID, EBSCO and Google Scholar. A pool of candidate articles was selected and entered into the four selected search portals. In order to be included in the pilot study, articles needed to have between 15 and 300 related articles identified by the search interface. (In instances in which more than 300 results were retrieved, the set was limited to the first 300 citations.) Furthermore, article retrieval was restricted to the period 1998-2008 and limited to documents in the English language. Our criteria for analysis were relevance and overlap. Researchers then recorded the lists of related articles and stored them in a bibliographic database for ease of manipulation. Relevance was established through expert review and mediation by a third participant should adjudication prove necessary. Once relevance judgments were completed we were able to compute precision and relative recall. The statistical test ANOVA determines which of the Finding Similar/Related Articles functions is the best based on these criteria. We also computed the overlap between the sets. Because calculation of recall traditionally depends on knowing all of the relevant documents in the repository, we calculated relative recall by dividing the number of relevant articles retrieved from one database by the total number of relevant articles retrieved.   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.

Preliminary Results:

In an initial study, results suggest considerable differences in precision and recall scores for article sets retrieved by PubMed and Ovid. The results indicate that PubMed's "related articles" feature is a useful tool, offering a large number of relevant articles, with PubMed generating a relative recall of 74% and a precision rate of 76%, compared with a relative recall of 26% and a precision score of 82 for Ovid. Ultimately, it was determined that larger data sets comparing the relative effectiveness across more portals would yield useful information, especially as literature searches on this subject have found no comparable research.

1) http://www.ncbi.nlm.nih.gov/bookshelf/br.fcgi?book=helppubmed&part=pubmedhelp#pubmedhelp.Computation_of_Relat
2) http://hsclibrary.uchsc.edu/handouts/database-help/cinahl.pdf
3) Mark D. Smucker "Find-similar: similarity browsing as a search tool SIGIR'06"
4) Jimmy Lin , et al "Navigating information spaces: A case study of related article search in PubMed" Information Processing and Management: an International Journal, v.44 n.5, p.1771-1783, September, 2008