Unanswerable Questions: What Do You Do with Challenging Health Reference Questions?

Carolyn Dennison, MA, MLIS, AHIP[1] and Angela Lee, MLS, MSW[2]

[1]University of Hawai'i at Manoa (UHM) Library, 2550 McCarthy Mall, Honolulu, HI 96822; [2]Health Sciences Library of the John A. Burns School of Medicine, University of Hawai'i at Mānoa, 651 Ilalo Street, MEB 101, Honolulu, HI 96813

Problem Statement

Medical librarians usually encounter difficult reference questions that are unanswerable using conventional reference techniques. This may be due to a lack of knowledge, resources, know-how or time to answer them. What are those difficult questions? Is there a way to identify and define them so that we can address them when they appear on our desk. A number of past studies about difficult clinical questions have been explored by the medical profession (Ely et al., 2005). Sometimes those questions are passed onto to us to answer. Can we learn from questions that appear to challenge our standard practices?

For medical librarians, PubMed and other health databases are the go-to site for many health sciences questions, but the literature cannot serve as the source of all answers. In a study of Oregon physicians, only 46% of questions arising in practice could be answered using the medical literature (Gorman, 1994).

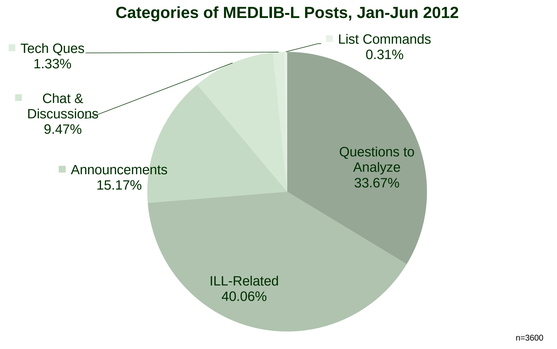
The purpose of this exploratory study is to better understand these unanswerable questions. Our objectives are to: (1) identify the types of difficult questions, (2) the response to those questions, and (3) user satisfaction. This poster reports on the preliminary results of the first phase of our study, which is to identify the types of questions that are regarded as difficult.

Method

MEDLIB-L was chosen as a test case because it has a long history of being the go-to-source for submission of difficult questions. Questions appearing on the listserv are posed only after librarians have exhausted all avenues of research. What kind of questions are there? Are they really difficult questions? Have they been answered in the listserv? Can we learn anything from them? And how might our understanding about the types of questions on listserv help inform and improve our practice? Although a study was done in 2004 that analyzes messages and traffic counts in Medlib-L (Smith), our study focuses specifically on a content analysis of reference questions.

We retrieved weekly MEDLIB-L log files over a six-month period (January to June 2012) from the MEDLIB-L server. We also received a copy of the MEDLIB-L change log from the listserv administrator. Both the MEDLIB-L log files and change log were used to set up and populate the fields on an Excel spreadsheet. There were a total number of 3600 listserv messages. Based on the subject line, messages were sorted into the following categories: questions, announcements (e.g., job openings), chats & discussions, interlibrary loan requests, job openings, technical questions related to the availability and operations of online resources, and list commands posted in error. The chart shows the distribution of these categories for the posted messages:

Method (continued)

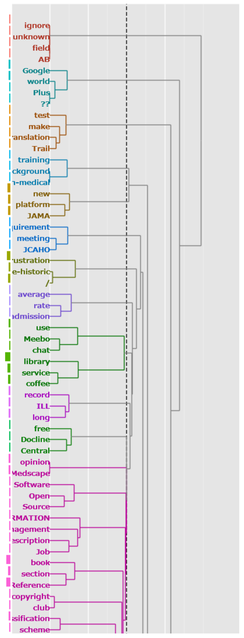


After filtering the messages, we focused our analysis on messages that were labeled as reference questions (REF Q). Of the total number (3600) of listserv messages, 1212 were identified as REF Qs. Of those, 675 were actual questions and 537 were responses to the questions.

The messages were then cleaned and prepared for entry into two open source content analysis software programs, KH Coder and TextSmart. TextSmart provides a quick read of word frequencies and concordances. KH Coder offers advanced features that permits co-occurrence of words (collocation) to identify concepts or themes.

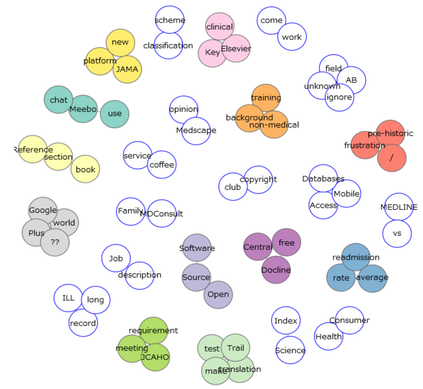
Results

Word frequencies and word collocation informed us about some of the prevailing themes. Below is a tree diagram used to illustrate the hierarchical arrangement of a cluster of terms from the posted messages based on their degree of similarity or association.

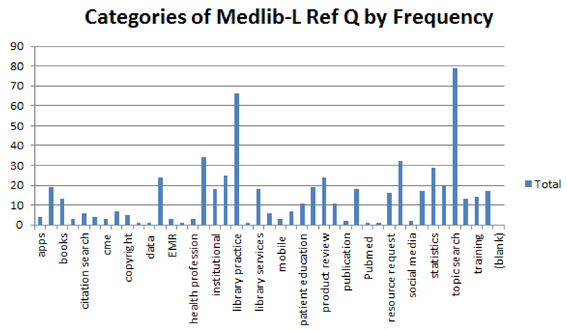


Results (continued)

Another view of the data shows a co-occurrence network, which provides a graphic visualization of potential relationships between the entities and/or concepts within the text messages:



While the computer analysis provided the initial term clusters, we found overall that the word groupings were too specific for adequate pattern discovery and categorization. Hence, we manually categorized the messages to better capture the contents of the postings. The chart below shows the categories of difficult questions by frequency.



As indicated in the chart, many of the categories are quite broad and encompassed a diverse range of questions. For example, "library services" covers a host of traditional services (instruction, reference, consultation, etc.). The chart also shows a high count of questions in the "topical search" category. At face value these question appear to typical reference questions; however, they are so specific and granular that it is difficult to establish any patterns of relationships among them.

Conclusion

What is a difficult question? Preliminary results of our content analysis of the listserv's REFQs indicate the following:

1. The question may or may not be intrinsically difficult, but the librarian simply lacked the resources to access the answer.
2. The question may be domain specific. The librarian did not appear to have subject expertise to do a proper search. For example, there were a number of business-related questions that did not fall within scope of the medical or health fields proper.
3. The question may only be answered through an internal document within an institution.
4. The question may involve proprietary information.
5. The question may involve media types (such as audio, video, and graphics) that do not lend themselves to be searched easily.
6. The question may not be a typical library question because it falls outside the scope of a published literature search. For instance, many of the questions dealt with locating people, programs, products, or services.
7. The question may depend on multiple sources to arrive at an answer. This may explain why some questions received partial or incomplete responses.
8. The question may not have an answer because no research or study has been done on it.
9. Finally, it is worth noting that many of the questions remained unanswered. The number of answers (not the responses per say) that directly addressed the questions were significantly smaller than the number of questions that were posed.

While this first phase of the exploratory study presents only a snapshot of the questions that appear in MEDLIB-L, it provides important information on what types of questions are difficult and why. The results are only preliminary and cannot be generalized at this stage. We may need to do another iteration of the study by expanding the data time period. An issue that we encountered was the problem of finding meaningful patterns in the data. This was due to the infrequency in the occurrence of question types. We attempted to broaden the categories to address this issue but the results were not very informative. Extending the coverage period may give us a more adequate sample size for analysis. Secondly, we need to evaluate the validity and reliability of the categories. Thirdly, we need to review the response to the questions and to determine whether there are any emerging trends in reference practice.

Acknowledgements

Kate Corcoran (Medical Library Association) for assisting us with obtaining files from MEDLIB-L.

References

Ely, J.W. et al. (2005). Answering physicians' clinical questions: Obstacles and potential solutions. Journal of the American Informatics Association, 12, 217-224.

Gorman, P.N., Ash, J., & Wykoff L. (1994). Can primary care physicians' questions be answered using the medical journal literature? Bulletin of the Medical Library Association, 82:140-146.

Smith, C. (2004). A longitudinal study of the culture of MEDLIB-L. Journal of Hospital Librarianship, 4: 29-43.

TEMPLATE DESIGN © 2008
www.PosterPresentations.com