

Frequency is determined by the number of occurrences of each tag and is represented using font size. Recency is based on the average publication date for the abstracts in which a tag appears and is illustrated with different font colours ranging from bright red for the most recent to dark grey for the oldest.

Only tags that are at least 10% as frequent as the most frequent tag are displayed. On each tag, a mouse-over feature displays a list of words that share the same prefixes and a hyperlink links to the set of PubMed abstracts containing it.

3. USABILITY EVALUATION

To assess the usability of PubCloud versus PubMed, we designed a series of questions for users to answer using both interfaces. All questions in the survey were about general topics of plant development. To answer the questions, the users were provided with the output from both PubCloud and PubMed to queries for “terminal flower” and “Arabidopsis”. The participants answered the questions in either the abstract-view of PubMed (with all hits in one single page) or the tag cloud view of PubCloud. The survey was run with 20 people of equal number of genders, all of whom were not experts in plant development but had some background in biology. One half of this group answered the questions with PubMed and the other half PubCloud. The number of correct answers was counted for each question, as well as the time spent answering the question.

3.1 Results

The metrics for each question in the survey were summarized in three major categories and are discussed below.

3.1.1 Correctness and quality of the answer

For questions that fall into the descriptive class, for example, “Is TFL a transcription factor?”, the quality of the answers was generally higher in PubCloud than in PubMed. Conversely, questions that required the user to identify relationships between multiple concepts, for example, “Name three other genes involved in this process”, were better answered using PubMed.

3.1.2 Time spent answering the question

Overall, users spent less time using PubMed to answer questions than using PubCloud. However, similar to the trends observed in the correctness and quality of answers, PubCloud users were able to answer descriptive questions faster than PubMed users. For relational questions, however, PubCloud users spent almost twice as long answering than those who used PubMed.

3.1.3 Degrees of helpfulness, satisfaction and understanding

In addition to the biological questions, users were asked to give their overall impression of the two interfaces. Users ranked PubCloud lower than PubMed regarding the degree of helpfulness. Although they spent more time and found the interface less helpful, the users ranked PubCloud with higher level of satisfaction. In terms of the level of understanding, the users ranked both PubMed and PubCloud equally.

4. DISCUSSION AND CONCLUSION

By summarizing the text content of web pages returned by a query, tag clouds offer not only an overview of the knowledge represented in the entire response, but also an interface that

enables users to navigate to potentially relevant information hidden deep down in the response list.

Figure 2 provides a comparison of PubMed and PubCloud along six axes, including answering descriptive questions, answering relational questions, speed, understanding, satisfaction, and helpfulness.

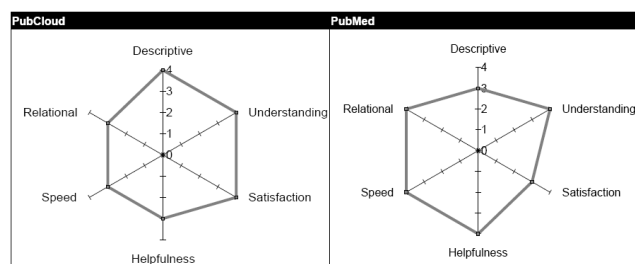


Figure 2. The ‘descriptive’ and ‘relational’ axes describe the correctness and the quality of the answers provided by the users. The ‘speed’ axis was generated based on the amount of time users took to answer each question. The ‘understanding’, ‘satisfaction’, and ‘helpfulness’ axes were based directly on users’ responses to questions about the interfaces.

As our results suggested, tag clouds are not a panacea for the summarization of web search results. Although they do provide some improvements in terms of summarizing descriptive information, tag clouds do not help users in identifying relational concepts, and in fact, slow them down when they need to retrieve specific items. Nevertheless, the widespread adoption of tag clouds and the generally positive impression of PubCloud from our users indicate that tag clouds fulfill an empty niche in the current ecosystem of Web interfaces.

Future improvements to PubCloud may include the use of clustering methods to group similar tags, and allowing more flexible visualization controls over font size, colours, hyperlinks, and location on the page. Since the development of tag clouds is ongoing and their application to the summarization of text content is still not well understood, we expect rapid improvement and increased utilization in the future.

5. ACKNOWLEDGMENTS

We thank the volunteers who participated in the evaluation study. BYLK is funded by iCAPTURE Centre for Cardiovascular and Pulmonary Research. TH is supported by CIHR/MSFHR Strategic Training Program in Bioinformatics. BMG is supported by an award from Genome British Columbia. MDW is supported in part by awards from Genome Canada and Genome British Columbia.

6. REFERENCES

1. Connotea: free online reference management for clinicians and scientists, <http://www.connotea.org/>.
2. Tag cloud - Wikipedia, the free encyclopedia, http://en.wikipedia.org/wiki/Tag_cloud.
3. Welcome to Flickr!, <http://www.flickr.com/>.
4. McEntyre, J. and Lipman, D. PubMed: bridging the information gap. *Cmaj*, 164 (9). 1317-1319.
5. Porter, M.F. An algorithm for suffix stripping. *Program*, 14 (3). 130-137.