MetaTest: Evaluation of Metadata from Generation to Use

Elizabeth D. Liddy, Eileen E. Allen, Christina M. Finneran
Center for Natural Language Processing, School of Information Studies, Syracuse University
{liddy,eeallen,cmfinner}@mailbox.syr.edu

Geri Gay, Helene Hembrooke, Laura A. Granka
Human-Computer Interaction Group, Cornell University
{gkg1,hah4,lag24}@cornell.edu

1. Introduction

Syracuse University's Center for Natural Language Processing and Cornell University's Human-Computer Interaction Group are collaborating to evaluate the utility of metadata within the Science, Technology, Engineering, and Mathematics (STEM) education domains. We are studying metadata from its initial generation to its use in accessing desired educational resources. With a testbed of lesson plans and activities, we are comparing the manually and automatically generated metadata for their retrieval effectiveness (i.e. ability to retrieve the most relevant resources); conducting a subjective evaluation of manually and automatically generated metadata as representations of the resource as judged by subject matter experts, and; conducting studies of users' search and navigation behavior when accessing the digital library. These evaluations successfully combine what we believe are necessary foci on how and whether metadata affects the user and system performance.

2. Information Retrieval Experiment

The information retrieval experiment compares the precision and recall of manually generated metadata to automatically generated metadata. The retrieval experiment is conducted on resources which have both manually and automatically assigned metadata elements. Users' queries are run on the two collections and the results are merged before presentation. Users in the STEM education domain judge the relevance of the retrieved lesson plans or educational activities.

3. Qualitative Experiment

We are extending the pilot study from our earlier digital library project that showed minimal differences in users' satisfaction with automatically generated versus manually generated metadata. In the current experiment, pre-service and in-service STEM teachers evaluate how well the automatically generated metadata and manually generated metadata represent the resources. Specifically, there are two conditions: the first measures the user's degree of satisfaction with the metadata after first viewing the resource and the second measures how well the user believes the metadata predict the actual contents of the resource, having viewed the metadata before the resource. Each experiment is a blind test in which users do not know how the metadata they are evaluating were generated – either manually or automatically.

4. User Study

The user study is evaluating the extent to which individuals use metadata to search for relevant resources, and will elucidate which metadata elements are most frequently used. Through a series of experiments involving eye-tracking and information search, we are investigating metadata use under various conditions and as a function of different demographic variables. The eye-tracking technology enables us to explore how the various ocular indices can serve to inform methods of metadata extraction. This technology also allows us to visually represent the entire scan path of individual users as they search documents for relevant information. From this we hope to be able to identify patterns or configurations of successful and unsuccessful searches, and to determine which parts of the metadata and resources are most informative.

5. Conclusion

Overall, our project goal is to assess the relative utility of metadata for retrieval and for helping the user access and browse resources. We will be sharing our preliminary findings on all experiments in our poster.

6. Acknowledgements

This project is funded by the National Science Foundation under the NSDL program (NSF Award Number: 0226312).

Proceedings of the 2003 Joint Conference on Digital Libraries (JCDL'03)
0-7695-1939-3/03 $17.00 © 2003 **IEEE**