THE USER-CENTERED DESIGN OF A NON-SPECIALIST METADATA TOOL AND INTERFACE FOR THE INTERNET PUBLIC LIBRARY

Michael Khoo, Jung-Ran Park, Xia Lin
The iSchool at Drexel University, Philadelphia PA 19104, USA

Abstract
Good quality metadata support rich user interactions with a repository, while poor quality metadata hide resources, produce poor search results, and negatively affect user satisfaction (Barton et al. 2004; Beall 2005; Geisler et al. 2002). Good quality metadata can be resource-intensive to produce, and as a result there is a ‘metadata generation bottleneck’ between the growing numbers of digital resources requiring description and the limited numbers of metadata specialists who can create such descriptions (Liddy et al., 2001). This can restrict metadata production and reduce the utility of repositories (Wilson, 2007). Approaches to addressing this bottleneck include training more metadata specialists; automatic metadata generation (Liddy et al., 2001; Ochoa & Duval n.d.); and also the solution discussed in this poster, the user-centered design and development of a tool to support student volunteers and interns to engage in the creation and quality control of Dublin Core metadata records for the Internet Public Library (IPL: http://www.ipl.org/).

Non-Specialist Metadata Tools
Non-specialist metadata tools have had mixed success (Crystal & Greenberg, 2005). This mixed success is partly related to the efficacy with which these tools convey complex metadata concepts to non-specialists. Case studies have described how, despite good intentions, metadata tools can be hard for users to understand, leading in turn to low levels of metadata production (e.g. Greenberg et al. 2003, Kastens et al. 2005, Khoo 2005, Wilson 2007). Crystal and Greenberg (2005) observe that “succinctly explaining the fields of a standardized schema such as Dublin Core to metadata novices appears to be a key challenge for designers,” and that “developing a conceptual understanding of metadata records and their use in retrieval was found to be challenging for users.”

The Internet Public Library (IPL)
The IPL receives approximately 1 million visits a month from over 200 countries and territories around the globe. It is currently crosswalking its metadata to Dublin Core. To support this work, the IPL is developing a tool for the creation, editing, and quality control of Dublin Core metadata. Wilson (2007) identifies three common metadata quality concerns that are also of concern to the IPL: completeness (the number of fields utilized), accuracy (the structural correctness of the data entered, e.g. for names, dates, etc.), and intelligibility (the semantic content of the metadata in the field). For instance, with regard to completion, an initial analysis of the IPL shows that a few fields (e.g. Title, URL, Abstract) are heavily used while other fields are lightly used (c.f. Moen et al. 2006, Ward 2003, Wilson 2007). With regard to the accuracy and intelligibility, while the original IPL records were created by trained volunteers, there are still variations in format (George Washington vs. Washington, George) or content (George Washington) (c.f. Barton et al. 2004; Beall 2005). These issues will need to be addressed.

As advocated by Crystal and Greenberg (2005) the development of the IPL metadata tool is following a user-centered ‘interaction design’ process (Norman 2002; Sharp et al. 2007) that considers not just the functionality of the tool, but also users’ needs. The work began with the elicitation of initial design requirements from students in HCI classes at Drexel University. The students reviewed IPL metadata records using existing prototype tools, critiqued these tools, and then used a range of HCI techniques, including hierarchical task analyses, personas, and scenarios, to develop basic functional paper prototypes. A common theme of these prototypes was the integration of various task windows (e.g. for the resource, the catalog record, the catalog rubric, the evaluation form, etc.) into one window. The prototypes were evaluated with heuristic evaluations, cognitive walkthroughs, and think-alouds, and the
evaluation data supported the development of a functional prototype that will be subject to further user testing.

The accuracy and efficacy of the final tool will be assessed through a number of means, including (a) a review of the accuracy of a random sample of completed cataloging tasks, (b) a comparison of completed forms by different users to test for inter-user reliability, and (c) feedback from initial users (i.e. student volunteers and interns).

**Outcomes**

The new IPL metadata tools will increase the quality of the library’s metadata, positively impact the search service, and directly benefit the IPL’s millions of global users. The research findings will be useful to other libraries involving users in their metadata work, and will have a direct relevance to the digital library and HCI research communities. The activities are integrated into courses taught at the Drexel iSchool, and so we will also report on the significant curriculum and pedagogical outcomes and implications of this research.

**Acknowledgement**

This research is supported by a grant from the OCLC-ALISE Library and Information Science Research Grant Program (OCLC LISRG). 

**References**


http://www.dlib.org/dlib/november05/kastens/11kastens.html.


