

A Study of Annotations for a Consumer Health Portal

Lili Luo, David West, Gary Marchionini, Catherine Blake
School of Information and Library Science
University of North Carolina, Chapel Hill

luolili@email.unc.edu, dwest@email.unc.edu, march@ils.unc.edu, cablake@email.unc.edu

ABSTRACT

This paper presents a study of annotations made by cataloguers of consumer health websites in order to better understand the website cataloging process.

Categories and Subject Descriptors

H.3.5 [Information Systems]: Online Information Systems-
Web-based Services

J.3 [Computer Applications]: Life and Medical Sciences-
Health

General Terms

Design, Reliability, Human Factors

Keywords

Annotation, Cataloging, Web Services.

1. INTRODUCTION

Catalogers of websites for a digital library face unique challenges. There are no well-established rules for cataloging the less-structured and constantly-changing information object. Identifying problems arising from website cataloging process will provide insights in designing better cataloging systems to support the process. One of the approaches of exploring the problems and how they are dealt with is examining the notes that catalogers made during the website cataloging process. In this study, catalogers' notes, or annotations, of cataloging a consumer health portal was examined in an attempt to find out the problems and issues involved in the cataloging process.

2. NC HEALTH INFO

NC Health Info is a collection of web-based references to health care related resources in North Carolina (<http://nchealthinfo.org/>). Websites of consumer health services are reviewed and categorized primarily based on the type of medical service provided and the geographic scope of the service. The cataloging interface of this website directory allows catalogers to make detailed notes during the cataloging process, either for themselves or to share with others.

3. METHODOLOGY

A random sample of 464 website catalog records was selected from the complete body of over 2,700. The "note" field of each catalog record was extracted for analysis. A "note" field was composed of substantive messages made by catalogers and non-substantive messages created by the system itself. For example, each cataloged website needs to be reviewed every six months and the cataloging system generates a message each time the site is reviewed. All the non-substantive messages in each "note" field were removed, which yielded 371 substantive messages.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

JCDL'05, JUNE 7-11, 2005, Denver, Colorado, USA
Copyright 2005 ACM 1-58113-876-8/05/0006...\$5.00.

These messages were considered as annotations of the cataloging process. They were analyzed and grouped to three facets and eleven categories. (see Table 1).

Table 1. The Categorizing Schema

Facet	Category
Content	Website Navigation: issues involved in navigating and accessing the website
	Categorization of Geographic Scope : issues involved in defining the geographic scope of the website
	Categorization of Topical Scope: issues involved in defining the topical scope of the website
	Miscellaneous: about issues related to website cataloging that fall into none of the above categories
Format	Question: Any comments that are questions along with anything that can reasonably be inferred to be a question.
	Answer: Any comment explicitly in response to a question and comments which probably answer unasked questions
	Statement: Declarative statements
Function	Log of Action: A statement of an action taken in the past
	Reminder: A statement to remind catalogers of actions that should or should not be taken in the future and relevant information that they should notice in the future
	Reach Consensus: A statement made in the process of reaching an agreement on a disputed point.
	Action Request: A comment that request a cataloger to take an action or provide information

4. RESULTS

The content analysis indicated that most of the notes related to establishing the topical scope of a website (n=192) and website navigation (n=147). A large number (n=266) of notes took the form of a statement while 109 were posed as questions and 97 as answers. As for functions, 99 were simply logs of what the cataloger did, 29 were reminders for the cataloging team, 181 were requests for other catalogers to take an action or provide information, and 174 were messages exchanging ideas and reaching consensus on solving a particular problem arising from the cataloging process.

The findings also indicated that ninety-seven of the 464 note fields containing at least one round of discussion with regard to properly cataloging the website. Such consensus building is necessary to avoid low levels of inter-rater reliability with respect to the final catalog decision. Thus, software tools that support collaboration between catalogers would enable catalogers to reach consensus in on- or off-line environments.

Our analysis revealed two challenges that appear to be specific to an on-line environment. The first concerns assigning a topic and geographic information to an entire site or the sub-domains. The second concerns the dynamic nature of websites that requires regular reviews by catalogers. Software tools that detect the change of a webpage would enable catalogers to target areas of change and thus increase their efficiency of the manual review process.

This research was supported by a gift from Microsoft Research.