Website Labels Construction Based on Thesaurus Concepts

Chih-Hwei Lu, National Taiwan Normal University, Taiwan
Jiann-Cherng Shieh, National Taiwan Normal University, Taiwan
Chieh Hsiao, National Taiwan Normal University, Taiwan

Abstract
Website labels are the identifiers used to represent chunks of website contents after organizing information. Website labels provide visual and cognitive cues that are the efficient retrieving channels between users and website contents, thus critically affect whether users can acquire the required information efficiently and intuitively. In this study, we try to apply thesaurus concepts, including revealing relevant vocabulary, to promote the terms of indexed labels clear and understandable, and to help in retrieving their corresponding target contents. There are five steps were proposed to approach the purpose of this study. First, analyzing the contents of library website and picking out their labels. Second, the randomly selected participants discuss the website labels extracted from the first step and rename the labels of those being not easy to understand. We call the original extracted library website labels the A labels group and the one consisted of renamed labels the B labels group. According to the above A and B labels groups and the discussions from the participants, researchers construct the library website labels using thesaurus concepts to generate another website labels the C labels group, which is combined the original library labels and has the characteristics of thesaurus. At the fourth step, card sorting method is applied to different labels groups A, B and C, in order to construct three website tree structures. Finally, we conduct the findability experiment to evaluate the tree different websites structures and compare their differences to verify the effectiveness of this research.

Keywords: Website Labels, Information Architecture, Card Sorting, Thesaurus, Findability

Introduction

Recently, a regular practice for people is to search the Internet for obtaining information. Some web sites provide logical structures that help us find answers and complete tasks. Others lack any intelligible organization and frustrate our attempts to navigate through them. The design of information architecture affects whether users can efficiently obtain the information they need (Morville & Rosenfeld, 2007). One prominent component that facilitates communication between a website and its users is the adopted website labels. The quality of website labels determines whether users can obtain information efficiently and accurately (Kalbach, 2007). Hence users' thoughts and suggestions must be incorporated during the design process to construct website labels that best meet their needs.

Thesaurus is a dictionary of controlled vocabulary, which prevents the situation of confusing the vocabulary of natural languages. The inter-relation between words and phrases are properly labeled to help users better understand their meanings (American National Standards Institute, 2005). In this study, the concepts of thesaurus were used to assist university library to construct website labels that meet the requirements of users.

Literature Review

1. Website labels

Rosenfeld and Morville (2007) classified information architecture into four systems: organization, labeling, navigation, and search. In this research, the topic for in-depth study was the labeling system. A label can be a word or short phrase that provides an effective means of summarizing a topic or action (Motive, 2004). Website labels represent the information contained on the site. Their purpose is to effectively communicate information and to provide users with the accurate links without utilizing much web space. A successful website label will often draw on a user's existing, contextual understanding, their mental model of a topic-area or process (Motive, 2004). This would inform them that they were on the correct path and would be able to find the information that they needed.

Toub (2000) proposed that labeling refers to the name or icon of a content object, such as the title of a page, or the title of a category or heading. Web designers must decide on the type of labels that best suit the users' cognition. In this study, website labels specifically refer to the indexing term for the contents found on a page within the website tree structure. Rosenfled and Morville (2007) pointed out that the users of website are the best references when constructing the labels.

2. Thesaurus

Thesaurus is a dictionary created for classification purposes. When the standard for classification is based on the semantic relations between words and phrases, thesaurus must be compiled to define those relations. The ANSI/NISO Z39.19 highlights that a controlled vocabulary is required in thesaurus because a natural language has the following special qualities: two or more words/phrases can be used to express the same concept or object and two or more words with the same spelling can represent

different concepts, contents, or objects. Thesaurus defines the semantic relations between words/phrases in three ways: (i) equivalency between concepts or usages, (ii) hierarchy between superordinate and subordinate concepts, and (iii) cross reference to remind users that there exist an association between one word and another.

Website labels consist of words or short phrases. The hierarchical structure of a website is formed by a combination of the semantic relations of words/phrases that form those labels. The controlled vocabulary function of the thesaurus is used to avoid the situation of confusion, which easily arises due to the vocabulary of natural languages. The inter-relations among words/phrases help users to understand the significance of website labels in a simple manner.

3. Card sorting

Spencer and Garrett(2009) stated that card sorting is a user-centered design methodology that enhances system findability. It can be seen as a tool that understands the target users, rather than a navigation-based design method. Furthermore, card sorting helps in understanding users cognitions (especially their opinions on categorization or website labels), thereby affirming their inclinations.

Card sorting can be implemented on physical or virtual platforms. The former refers to being desk-bound, which has the advantage of facilitating inter-personal communication. The latter makes use of computer software, including OptimalSort and EZsort. Although this method is constrained by screen size, there are no spatial or temporal limitations (Martin & Kidwell, 2001).

For the number of cards to select, Spencer and Warfel (2004) proposed that the minimum and maximum should be 30 and 100, respectively. Any lesser number will result in the categorization being incomplete, whereas any greater will make the subjects feel fatigued about the process. Hence, 30–100 cards are ideal. Kaufman (2006) believed that categorization should generally involve 20–50 cards. However, as many as 200 cards can be used if time is not a constraint, or for contents that are highly complex. Researchers can adjust the number of cards on the basis of their specific research topic and purpose, as well as the implementation method.

Regarding the number of participants to use, Spencer and Warfel (2004) proposed that a suitable number is 7–10. However, if groups are the basic test unit, better results will be obtained using five groups, with each group comprising three members (i.e., a total of 15 participants). After the data is collected through card sorting, depending on the research needs or planning scope, either qualitative or quantitative analysis may be used to generate suitable, logical, and useful analytical results (Ahlstrom & Allendoerfer, 2004).

4. Findability

When conducting research and analysis on users of website information architecture, the general approach is to do so from the perspectives of usability and findability. The latter concept was first proposed by Morville (2005), who believed that findability was more important than usability within the Internet environment. This is because if

a user cannot find a website or information in the first place, its usefulness will actually not be a matter for consideration.

Morville defined findability in two aspects: searching for information from outside a website and searching for information on the website. In this study, findability refers specifically to the latter. In the research on the modified-Delphi card sorting method by Paul (2007), the participants who assessed findability comprised 7 users. They were given 10 questions on the terms for website contents. The purpose of the findability assessment was to understand whether the categories in the users' cognition and that adopted by the website tree structure were consistent.

Research Design and Implementation

This study was the library website of the National Taiwan Normal University (NTNU), specifically, the website labels for the Chinese version of the website dated December 2010. Although the website had many user groups, the majority was the students of the university. Hence, the identity of our research participant was restricted to current NTNU students. Because the impact of different disciplines was not a research parameter, no limitation was placed on the faculty or department to which the participants belonged.

The research was carried out in five steps, which are elaborated below.

1. Analysis of website contents

Before implementation of card sorting, a detailed understanding of the contents found within the library's website had to be established and a list of items to be categorized must be summarized and prepared. The analysis of the website contents revealed that text-based labels were used predominantly. In addition to the use of contents analysis to understand the current website labels available for the selection of existing resources, screening rules were applied to eliminate unnecessary website items. In the end, 67 website labels were selected as names of the card items, that is the A labels group.

2. Discussion on website labels and gathering of participants' opinions

After the list of card items had been prepared, the next step was to hold discussions on the website labels. This was performed using the focus groups method to avoid individual participant being unable to express his/her opinions on the labels and to prevent the process from becoming too subjective. The intention was to have the mutual discussions stimulate the thoughts of the participants, and for suggestions to be made in a more objective manner. Each group comprised 3–5 members, with the researchers participating in and hosting the discussions.

The participants studied the cards for the contents represented by the various website labels, proposed the significance represented by each, and then exchanged views. They also commented on the inter-relations between labels, and were encouraged to jointly propose labels that they deemed more appropriate. Thus, some labels in the A labels group are replaced according to their discussions to form the B labels group.

3. Establishing website labels based on the thesaurus concepts

Next, the semantic relations between the words and phrases of all the new labels were determined with reference to the rules for preparing thesaurus stated in NISO/ANSI Z39.19. This process was undertaken by three subject experts in the field of library and information science: one scholar and one graduate student from the discipline and one librarian. The various relations between the words and phrases were then attached to the original library labels ("A labels group") for the participants' reference. This led to the construction of website labels based on the thesaurus concepts ("C labels group").

4. Card sorting

The number of participants for card sorting was determined on the basis of the recommendations of Spencer and Warfel (2004) and Kaufman (2006). Groups of 3–5 members (total of 15) were involved. According to Hawley (2008), card sorting by groups has the advantages of simultaneously obtaining the quantitative data for the process, as well as the participants' insights and reasons for the way they sorted the cards. Through discussions between the participants, the advantages and disadvantages of the various card sorting methods can be uncovered.

After group card sorting, the participants constructed three different tree structures using B and C labels groups (sorted out in the previous step), as well as the original A labels group (library website labels). The tree structures a, b, and c corresponded to the A, B, and C labels groups, respectively.

5. Findability test

The questionnaire survey was used in conjunction with the task assigned to participants to find the labels when browsing the website. Three different tree structures were created after the three groups of labels underwent card sorting. The participants had to search for the top 10 website labels based on the library's network traffic statistics and click-through rates. However, the repetition rate for the electronic resources category was too high. To avoid excessive overlap in terms of location of the selected labels within the tree structures, the survey items that were eventually selected for purpose of assessing findability included three labels with relatively high network traffic and another three labels with greater variability in terms of their locations.

The questionnaire was designed with a five-point Likert scale, which was provided to the participants together with diagrams on the various tree structures. The participants were asked to check the diagrams on the basis of the sequence of website labels through which they would click when browsing. Other than evaluating the degree of difficulty in finding the labels, the participants were also asked to score the appropriateness of each website label. The scores ranged from 1 to 5, representing very inappropriate, inappropriate, neutral, appropriate, and very appropriate, respectively. The SPSS statistical software was then used to conduct single-factor analysis of variance for dependent samples to understand the differences between the scores ascribed to the various tree structures.

Results and Analysis

1. Results for discussion on website labels

During the second step, the participants browsed through the A labels group, discussed these in groups, and then proposed novel and more appropriate terms or phrases for the website labels that they believed were easier for other users to understand. When several suggestions were made, the group voted to determine the new label that could directly replace the original. The opinions of the four groups were then consolidated. On the basis of a majority vote, the label with the highest repetition rate directly replaced the original. Eventually, the alternative (B labels group) comprised 66 labels, generating an equivalent number of cards for the subsequent card sorting experiment.

2. Establishing website labels based on the thesaurus concepts

The group discussions during the second step resulted in the collection of many different new labels. With the assistance of experts in the field of library and information science and based on the thesaurus concepts, these labels were organized into a lexical list of inter-related semantic relationships. Except for the number of words used, there was actually little difference between the A and B labels groups. Reference was made to the rules for preparing thesaurus as stated in NISO/ANSI Z39.19 (American National Standards Institute, 2005): labels with more detailed descriptive contents had clearer explanations, and which contained more specific implications were classified as hyponyms or narrower terms (NT), while the rest were hypernyms or broader terms (BT).

If there was a huge difference between the terms for the B labels group compared to the A labels group, both would not be treated as being affiliated in the same layer and hence, were marked as related terms (RT). When a label in the B labels group had the same meaning as that in the A labels group but with different wordings, these were treated as being equivalent and called non-descriptors (indicated as UF to stand for "use for"). The C labels group were generated by combining the A labels group with words/phrases containing semantic relationships derived using the thesaurus concepts. Those words/phrases were indicated on each card for the participants to directly refer to when card sorting, thus saving them the time and effort on checking the RTs during the sorting process.

3. Results of card sorting

Following the experimental design, the number of participants for card sorting was based on the recommendations made by Spencer and Warfel (2004) and Kaufman (2006). 15 participants were divided into four groups (3–5 members each) for card sorting of A, B, and C labels groups. The four participant groups separately categorized all three groups of labels, producing four types of categorization results for each group. Since the sample size was small, the qualitative method was adopted for analysis of the categorization results.

Based on the card sorting of the three groups of labels, participants preferred that the content items in the library's website be categorized to highlight the services being

provided. Further categorization could then be made according to the contents of the various services. In addition, the participants thought that the tree structure of the website should not be too complicated, and preferably with two or less levels of clicking through. Otherwise, it was likely that information contained in the lower levels would not be browsed by users. On the other hand, the participants did not want too many terms to be attributed to the same category. Otherwise, the important information might be lost among the excessive number of terms.

When participants in the various groups had different understanding of the terms for labels stated on the cards, their doubts over categorization were addressed through discussions with fellow group members. However, when the term of a label was too broad or general, the participants would tend to place its card in a roughly appropriate category, without prior investigation of its contents in detail. Separately, some participants indicated that when they were unsure of the exact meaning of a label, they referred to the additional notations on the C labels group and became more confident about categorizing that label.

4. Findability analysis of survey results

The survey questionnaire for findability made use of the five-point Likert scale. The participants were asked to search for six label terms from three different website tree structures, evaluate the degree of difficulty in finding those items, as well as give their personal scores on whether the labels was appropriate.

The findability analysis involved 25 participants from 10 different departments. From the total scores, the mean and standard deviation for each of the tree structure were calculated. The scores were rounded to two decimal points, with the third decimal point dropped if its value was less than four, but carried over if it was five or larger. The final results are shown in Table 1. The total and mean scores for tree structure c (598 and 23.92, respectively) were higher than that for both tree structures a and b. Tree structure a ranked second with a total of 555 and mean of 22.2.

Table 1: Scoring table for findability of website tree structure

Website tree structure ⟍ Serial no. of subject	Tree structure		
	a	b	c
Total	555	526	598
Mean	22.2	21.04	23.92
Standard Deviation	2.77	2.42	3.12

SPSS software was used to conduct single-factor analysis of variance for dependent samples. The results of the Mauchly spherical test were observed: it can be seen from Table 2 that the verification value of Mauchly's W is .937. The approximate value after chi-square transformation is 1.498. The degree of freedom is 2, $p = .473 > .05$ (significance not reached), indicating that the statistical data were consistent with a spherical assumption.

Table 2: Mauchly spherical test

Items on within-subjects effect	Website tree structure
Mauchly's W	0.937
Approximate chi-square distribution	1.498
Degree of freedom	2
Significance	0.473
Epsilon(a)	
Greenhouse–Geisser	0.941
Huynh-Feldt value	1.000
Lower limit	0.500

For verification of the within-subjects effect, since the analysis of variance for dependent samples complied with the spherical test, the data on the first row ("assumed to be spherical") was directly examined when analyzing the information. The information in Table 3 indicate that all between-group F-values are 7.865, p = 0.001 < 0.05 (significance reached), indicating that significant differences existed between the scorings for the three website tree structures.

Table 3: Verification of items on within-subjects effect

Source	Assumption	Type III sum of squares	Degree of freedom	Mean sum of squares	F-test	Significance
Website tree structure	Assumed to be spherical	104.987	2	52.493	7.865	0.001
	Greenhouse–Geisser	104.987	1.881	55.803	7.865	0.001
	Huynh-Feldt value	104.987	2.000	52.493	7.865	0.001
	Lower limit	104.987	1.000	104.987	7.865	0.010
Error (website tree structure)	Assumed to be spherical	320.347	48	6.674		
	Greenhouse–Geisser	320.347	45.153	7.095		
	Huynh-Feldt value	320.347	48.000	6.674		
	Lower limit	320.347	24.000	13.348		

Paired comparisons between the website tree structures are shown in Table 4. It can be seen that the scores for Tree Structure c are significantly different from that of Tree Structures A and B. Its mean score is also significantly better than that of the other two tree structures. This indicates that in terms of findability, the proposed website tree structure categorized using cards with characteristics of thesaurus concepts is significantly superior to the other two tree structures.

Table 4: Pairwise comparisons

(I) Web site tree structure	(J) Web site tree structure	Difference between means (I–J)	Standard error	Significance (a)	95% confidence interval for the difference (a)	
					Lower limit	Upper limit
a	b	1.160	0.639	0.082	−0.160	2.480
	c	−1.720*	0.799	0.042	−3.369	−0.071
b	a	−1.160	0.639	0.082	−2.480	0.160
	c	−2.880*	0.745	0.001	−4.417	−1.343
c	a	1.720*	0.799	0.042	0.071	3.369
	b	2.880*	0.745	0.001	1.343	4.417

An interesting result was observed: comparing Tree Structures A and B, the findability of the latter was rated poorer than the former. This was despite Tree Structure B containing website labels that were renamed by users. This indicates that the renamed labels were suggested based on the participants' literal understanding of the definition of the terms, but that understanding was not translated to one that other users could comprehend. The original website tree structure constructed by the library was found to be more professional. This could be because many users regularly made use of that website and hence, were already familiar with the contents of the original website labels.

This corroborates with the opinions of the participants made during the discussions. They felt that the new website labels were too vernacular and did not seem sufficiently professional. At the same time, they felt that some of the labels used on the original library website were over professional, making it hard for general users to understand. Nevertheless, since the participant was the website of university library, it was still appropriate for such relatively more professional words/phrases to be used.

Conclusion

The aim of this research was to study website labels construction based on the thesaurus concepts and check whether these would help users to improve their understanding of labels found in university library's website. We also wanted to assist users to carry out better card sorting to achieve the goal of improving the website's findability. The paper proposed five steps to accomplish the research. First, the library's website labels were discussed using the focus groups method. After the participants had proposed new labels, these were collected and compiled. Labels based on the thesaurus concepts were then established by referring to the rules for thesaurus. This led to three groups of labels: A labels group contained labels originally prepared by the library, B labels group contained renamed labels that directly replaced the original, and C labels group contained those with thesaurus concepts notations. The three labels groups underwent card sorting to compare the way labels with different characteristics affected the participants' categorization process. Eventually, three website tree structures were constructed for findability analysis and to determine variations in the degree of findability. The conclusions of this research are as follows: labels constructed based on the thesaurus concepts could

indeed help users to better understand the university library's website and the process of card sorting. The questionnaire survey on findability verified that the website tree structure constructed using the C labels group was indeed superior to that using the A and B labels groups. It can thus be concluded that a website tree structure constructed using labels based on the thesaurus concepts before undergoing card sorting better fits the needs of users when they are browsing a website for information. Consequently, users would be able to obtain the information that they require more efficiently.

References

Ahlstrom, V. and Allendoerfer, K. (2004). Information organization for a portal using a card-sorting technique. Retrieved from http://hf.tc.faa.gov/technotes/dot-faa-ct-tn04-31.pdf

American National Standards Institute. (2005). Guidelines for the construction, format, and management of monolingual controlled vocabularies.
Retrieved from
http://www.niso.org/kst/reports/standards/kfile_download?id:ustring:iso-8859-1=Z39-19-2005.pdf&pt=RkGKiXzW643YeUaYUqZ1BFwDhIG4-24RJbcZBWg8uE4vWdpZsJDs4RjLz0t90_d5_ymGsj_IKVaGZww13HuDlYn5U74YdfA-3TffjxYQ25QrtR8PONuJLqxvo-l0NIr5

Hawley, M. (2008). Extending card-sorting techniques to inform the design of web site hierarchies. Retrieved from
http://www.uxmatters.com/mt/archives/2008/10/extending-card-sorting-techniques-to-inform-the-design-of-web-site-hierarchies.php

Kalbach, J. (2007). Designing web navigation. Sebastopol, CA: O'Reilly.

Kaufman, J. (2006). Card sorting: An inexpensive and practical usability technique. Retrieved from http://www.stc.org/intercom/PDFs/2006/200611_17-19.pdf

Martin, S. & Kidwell, D.K. (2001). A case study in cluster analysis for intranet organization. EMAT '01 Proceedings of the 2nd International Workshop on Engineering Management for Applied Technology (pp.57-64). Washington, DC: IEEE Computer Society.

Morville, P. & Rosenfeld, L. (2006). Information architecture for the World Wide Web (3re ed.), Sebastopol, CA: O'Reilly.

Morville, P. (2005). Ambient findability: Libraries at the crossroads of ubiquitous computing and the internet. Online, 29(6), 16-21.

Motive. (2004). The motive web design glossary. Retrieved from
http://www.motive.co.nz/glossary/label.php

Paul, C.L. (2008). A modified Delphi approach to a new card sorting methodology. Journal of Usability Studies, 4(1), 7-30.

Spencer, D. & Warfel, T. (2004). Card sorting: a definitive guide. Retrieved from http://www.boxesandarrows.com/view/card_sorting_a_definitive_guide

Spencer, D., & Garrett J. J. (2009). Card sorting: Designing usable categories. Brooklyn, N.Y. : Rosenfeld Media.

Toub, S. (2000). Evaluating information architecture. Retrieved from http://argus-acia.com/white_papers/evaluating_ia.pdf
Contact email: jcshieh@ntnu.edu.tw