



Date : 09/06/2006

Facet Analyses of Categories Used in Web Directories: A Comparative Study

Kyung-Sun Kim

Sei-Ching Joanna Sin

Soo-Jin Park

Xiaohua Zhu

Jom Polparsari

School of Library and Information Studies

University of Wisconsin-Madison

Madison, USA

Meeting:	136 Classification and Indexing
-----------------	--

Simultaneous Interpretation:	No
-------------------------------------	----

[WORLD LIBRARY AND INFORMATION CONGRESS: 72ND IFLA GENERAL CONFERENCE AND COUNCIL](http://www.ifla.org/IV/ifla72/index.htm)
20-24 August 2006, Seoul, Korea
<http://www.ifla.org/IV/ifla72/index.htm>

Abstract

Faceted classification is believed to be suitable for organizing digital information resources. Based on a faceted classification model suggested for Web resources (Zins, 2002), the current study analyzed popular Web directories from different Asian countries/areas and examined cultural differences reflected in their classification systems. Three popular Web directories from four countries/regions (China, Hong Kong, Korea, and Thailand) were selected and their classifications were analyzed and compared: a local Yahoo and two home-grown Web directories from each country/region. Based on the findings, the study suggests a model that might be more suitable to Asian culture.

INTRODUCTION

Information available through the Web is becoming more diverse and its amount is also growing rapidly. As of January 2005, it is estimated that the indexable Web has reached to at least 11.5 billion pages (Gulli & Signorini, 2005). Presently, these voluminous and ever changing resources on the Web are often organized in mostly hierarchical classification systems. However, such hierarchical systems might not be as suitable for organizing cross-cultural resources for a culturally and linguistically diverse groups, as there might be cultural differences in categorization (Segall, 1990; Malt, 1995).

This study examines and analyzes classes/categories used in twelve popular Web directories from four different Asian countries/regions. Based on the findings, the study suggests and discusses a faceted classification model that could help organize Asian Web resources. The aim of this study is to take into account unique as well as common characteristics found in Web classifications of the respective

countries/regions, and to propose a faceted classification model that could facilitate access to Web resources for the culturally and linguistically diverse users in Asia.

FACETED VS. HIERARCHICAL CLASSIFICATIONS

To facilitate learning and information retrieval, we classify things by grouping similar things together. Most classification systems we use traditionally are hierarchical systems where subjects are arranged in a hierarchy of classes, divisions, and subdivisions. In such systems, a topic is pigeonholed in one location that can be reached through one path. Unless users have a good understanding of the structure of the hierarchy, and know exactly which one of the main classes (and divisions, subdivisions and so on) the particular subject is classed under, it would require of the users a series of trial-and-error attempts in locating the subject in the system.

The sameness, similarities and dissimilarities that Aristotle invoked and most hierarchical classification systems assume are not always universally accepted relationships (Olson, 2001). They are significantly influenced by our society, and may vary among groups with different cultures and needs. Besides, people may disagree on which sameness is important and which should be used to categorize things on the top level of the hierarchy. A primary problem of hierarchical classification systems is that people do not always want to choose the same sameness first: "we do not always want what the citation order gives us" (Olson, 2001). It is necessary to consider the diversity of users and to apply a range of optional, partial, and local solutions.

Unlike hierarchical systems that assign fixed slots to subjects in sequence, faceted classification systems assign to subjects clearly defined, mutually exclusive, and collectively exhaustive aspects, properties or characteristics of a class or specific subject (Taylor, 2000). Faceted classification provides more than one path to locate a subject. By using different aspects/facets of a subject, users can narrow down the search and locate the subject more easily. Although each of hierarchical and faceted systems has its own strengths and weaknesses (Kwasnik, 1999), traditional hierarchical structure is increasingly seen to be ineffective in organizing resources on the Web. Recently, there is growing recognition that the faceted approach is more effective and suitable for organizing Web resources (Ellis & Vasconcelos, 2000; Bates, 2002; Broughton, 2002).

Compared to hierarchical classification, faceted classification is flexible and hospitable in accommodating new categories. The faceted approach is also more suitable in representing complex ideas/objects. It is particularly useful for organizing complex items and materials in a multi-disciplinary environment (Broughton, 2002). In their examination of the use of the faceted approach in online databases, Ingwersen and Wormell (1992) discovered that the faceted approach excels in providing structured, multi-dimensional entry points into the documents/objects. Such multiple entry points would be beneficial to the organization of Web resources, as Web users are quite diverse in background. It is highly likely that different users would try to locate/browse a topic from different entry points. Traditional hierarchical systems, such as DDC and LCC, have been criticized as being rigid and sometimes ethnocentric (Olson, 1998). Examining the kinship terms in different classification schemes including LCC and DDC, Kwasnik and Rubin (Kwasnik & Rubin, 2003) highlighted the difficulty in mapping and translating the terms in these classification schemes to represent the kinship concepts accurately in different cultures and languages. A faceted approach, flexible and hospitable to multiple perspectives, should be more suitable for classifying cross-cultural and multilingual resources (Kwasnik, 1999).

Faceted classification systems also have some weaknesses. A rather complex or lengthy notation which users are not familiar with is one of them (Iyer, 1995). For Web resources, however, this issue would not be of concern, as documents/items can be accessible through hyperlinks without using notation. Recently, several applications have been developed, exploring efficient ways of implementing faceted indexing, browsing, and searching (Devadason, Intaraksa, Patamawongjariya, & Desai, 2001; Yee, Swearingen, Li, & Hearst, 2003; Ross & Janevski, 2005).

This study aims to investigate the potential of a faceted classification approach to organizing Web resources in different Asian counties, and to propose a faceted classification model that might be suitable

to the Asian culture. In the following sections, we shall first describe the methodology used in the study. Then, we shall present the findings of the study. A faceted classification model, developed based on the findings, will also be introduced. The paper will end with discussion of the findings and their implications.

STUDY

The goal of the study is to build a preliminary model for facet-based classification for organizing Asian Web resources, based on the analysis of the categories currently used in popular Web directories in four Asian countries/regions. Objectives of the study include: (1) to test and refine Zins' faceted model for organizing Web resources; and (2) to expand the model by adding sub-facets that might be useful.

Sample

A total of twelve Web directories, widely used in four different Asian countries/regions (including China, Hong Kong Special Administrative Region, Korea, and Thailand), were selected and their classification schemes were analyzed in the study. Web directories selected for the study include: Sina Beijing and Sohu (China); Sina HK and Tim Way (Hong Kong); Empas and Naver (Korea); and Hunsa and Sanook (Thailand). For each country/region, local *Yahoo* was also selected. In addition, the top-level categories used in Yahoo, Google Directory, and Librarians' Index to the Internet of the United States were also analyzed and compared with others in order to provide some insights on similarities and dissimilarities between Asian and American countries/regions.

Data Collection

For each Web directory, all the pages displaying the top- and second-level classes or categories were captured on July 10, 2005. Classes/categories used in directories of each country/region were translated and analyzed.

RESULTS

Top-Level Categories in Local Yahoo Web Directories

When local Yahoos were compared, some minor differences were found (See Table 1). Yahoo of Thailand, for example, had a 'Travel' category that was not found in others. In Korea, the way in which subjects were grouped was slightly different from others. Humanities, usually presented with arts, was grouped together with social sciences in Korea. In China and Hong Kong, politics was added to the 'Government' category. Except for these few variations, the top-level classes were almost identical. Such similarity across different local versions of Yahoos was not surprising as Yahoo in the United States might have some control over the local versions.

Table 1. Main Categories in Local Yahoos

Yahoo United States	Yahoo China	Yahoo Hong Kong	Yahoo Korea	Yahoo Thailand
Arts & Humanities	Y	Y	Arts	Y
Business & Economy	Y	Y	Y	Y
Computers & Internet	Y	Y	Y	Y
Education	Y	Y	Education/ <u>Learning</u>	Y
Entertainment	Y	Y	Y	Y
Government	Government & <u>Politics</u>	Government & <u>Politics</u>	Y	Y
Health	Health & <u>Medicine</u>	Health & <u>Medicine</u>	Health & <u>Medicine</u>	Y
News & Media	Y	Y	Y	Y
Recreation & Sports	Recreation & <u>Life</u>	Recreation & <u>Life</u>	<u>Leisure Life</u>	Recreation/ Sports
Reference	Y	Y	Y	Y
Regional	Y	Y	Y	Provinces & Districts
Science	Y	Y	Y	Y
Social Science	Y	Y	Humanities & Social Science	Y
Society & Culture	Y	Y	Y	Y
New Additions	N/A	N/A	N/A	N/A
				<u>Travel</u>

Top-Level Categories in Local Yahoo Web Directories Based on Zins' Model

Analyzing the top-level main categories used in nine Internet gateways including commercial portals (e.g., AOL, Yahoo), Zins (Zins, 2002) identified eight commonly used facets, based on which Web resources can be organized: (1) subjects, (2) objects, (3) applications, (4) users, (5) locations, (6) reference, (7) media, and (8) languages.

Here, the 'subjects' facet characterizes the subject or broad discipline areas (e.g., business; education; science). The 'objects' facet refers to any types of real-world objects or entities, including organizations and people. The 'applications' facet includes any resources that are related to specific applications (e.g., chats, e-mails). The 'users' facet indicates the prospective end users. The 'locations' facet covers any resources related to geographic locations. The 'reference' facet refers to different reference resources including almanacs, dictionaries, encyclopedias, etc. The 'media' facet points out any media by which the content of the resource are expressed (e.g., graphics, radio). Finally, the 'language' is related to the language used in the Web sites.

When Zins' model was used for analyzing the top-level classes, it was found that most classes belong to the 'subject' facet (See Table 2). 'Reference', 'location', 'media' and 'applications' facets were also present. As all the Web directories analyzed here were for individual countries and their own citizens, 'language' facet was not present. 'Users' and 'objects' facets were not found either.

Table 2. Main Categories in Local Yahoos: Facet Analysis Based on Zins' Model

Facets	Subjects	Locations	Reference	Media	Applications
Categories	Arts & Humanities Business & Economy Computers & Internet Education Entertainment Government Health News & Media Recreation & Sports Science Social Science Society & Culture	Regional	Reference	News & Media	Travel (Thailand)

Top-Level Categories in Home-Grown Web Directories

Next, all the home-grown Web directories from four Asian countries/regions were compared in terms of the top-level classes. Local Yahoos were not included in this analysis as their categories were already analyzed and found to be similar across nations. All the main classes from two home-grown directories of each country/region were analyzed and compared: *Sina Beijing* and *Sohu* (China); *Sina HK* and *Tim Way* (Hong Kong); *Empas* and *Naver* (Korea); and *Hunsa* and *Sanook* (Thailand). As a reference, top-level classes from *Yahoo*, *Google Directory*, and *Librarians Index* from the United States were included and used in this analysis. As a result, top-level main categories from a total of eleven Web directories were analyzed, based on Zins' model. Results are presented in Table 3.

Table 3. Main Categories in Home-Grown Web Directories

	United States	China	Hong Kong	Korea	Thailand
S	Arts & Humanities (Y)(L) Arts (G)	Arts (H)(N) Literature (H)(N)	Arts (T) Arts world (N)	Culture, <u>Arts</u> , & Religion (E) Culture & <u>Arts</u> (N)	Society, <u>Arts</u> & Culture (H)
S	Business & Economy (Y) Business (G)(L)	Business & Economy (H)(N)	Finance/ Property (T) Business Buy & Sell (T) Business & Economy (N)	Business & Shopping Mall (E) Business & <u>Shopping</u> (N) Economy & Money-Making Know-How (E)(N)	Business (S) Business & Economy (H)
S	Computers & Internet (Y) Computers (G)(L)	Computers & Internet (H)(N)	Computers (T) Computers & Internet (N)	Computers & Internet (E)(N)	Computer (S) Internet (S) Computer & Internet (H)
S	Education (Y)(L)	Education & Employment (N)	Education (T) Education &	Education & Schools (E)	Education (S) Science &

		Education & Training (H)	Employment (N)	<u>Education & Sciences</u> (N)	<u>Education</u> (H)
S	Entertainment (Y)	Entertainment & Recreation (H)(N)	Entertainment (T)	Entertainment (E)(N)	<u>Entertainment & Recreation</u> (S) Entertainment (H)
S	Games (G)		Games & Hobbies (N)	Games (E)(N)	
S	Government (Y)(L)	Politics/Law/Military (H)(N)	Government & Politics (T) (N)	Government Agencies & Society (E) Society & Politics (N)	<u>Government & Organization</u> (S)(H)
S	Health (Y)(G)(L)	Medicine & Health (N) Health (H)	Health (T) Health Care & Medical Treatment (N)	Health & Medicine (E)	Medical Science & Health (S) Health & Medical Science (H)
S	Home (G)(L)				
S	Law (L)	Politics/Law/Military (H)(N)			
S	News & Media (Y) News (G)	News & Media (H)(N)	Mass Media (T) News & Mass Media (N)	News & Media (E)(N)	News & Media (S) Media & News (H)
S	Recreation & Sports (Y) Recreation (G)(L)		Recreation/Leisure (T)	<u>Leisure & Sports</u> (N)	Entertainment & Recreation (S) Recreation (H)
S	Science (Y)(G)(L)	Science & Technology (H)(N)	Science (T) Science & Technology (N)	Sciences (E) Education & Sciences (N)	Science & Education (H)
S	Social Science (Y)	Social Science (H)(N)			
S	Society & Culture (Y) Society (G)	Society & Culture (H)(N)	Society (T) Society & Culture (N)	<u>Culture, Arts, & Religion</u> (E) <u>Society & Politics</u> (N)	People, Society & Culture (S) <u>Society, Arts & Culture</u> (H)
S	Sports (G)	Sports & Physical Training (H)(N)	Sports (T) (N)	Travel & <u>Sports</u> (E) Leisure & <u>Sports</u> (N)	Sports (S)(H) Travel & <u>Nature</u> (H)
S		Job Application & Recruitment (N)			
S		<u>Life Service</u> (H)(N)	Life/Living (T) Life Info (N)	Lives & Hobbies (E) Life Styles (N)	
S			Eating (T)		
S					Activities & Events (S)

S					Knowledge/ Wisdom & Info (S)
S					Automobile (S)(H)
S					Real Estate (S)
A			Travel (N)	<u>Travel</u> & Sports (E) World & <u>Travel</u> (N)	Travel (S) <u>Travel</u> & Nature (H)
A	Shopping (G)		Business <u>Buy & Sell</u> (T)	Business & <u>Shopping</u> (N)	Shopping (S) Shopping & Service (H)
L	Regional (Y)(G)(L)	Counties & Regions (H)			
L	World (G)			Korea & World (E) <u>World</u> & Travel (N)	
M		<u>Personal Homepage</u> (N)	Personal Homepages (T)		
M	News & Media (Y) Media (L)	News & Media (H)(N)	Mass Media (T) News & Mass Media (N) Film, TV, & Music (N)	News & Media (E)(N)	News & Media (S) Media & News (H)
O		<u>Companies</u> (H)	Service Organizations (T)		Banks & Financial Inst. (S) Government & <u>Organization</u> (S)(H)
O	People (L)			People (N)	
O			Stars (T)		
R	Reference (Y)(G) Ready Ref. & Quick Facts (L)	Reference (N)		Dictionaries & Reference (E)	
U			Adult/Sex (N)		
U	Kids & Teens (G)	Kids (N)		Junior Naver (N)	
	New Additions (Y)				Others (H)

* US: Yahoo [Y], Google Directory [G], and Librarians Index [L]; China: Sina Beijing [N] and Sohu [H]; Hong Kong: Sina HK [N] and Tim Way [T]; Korea: Empas [E] and Naver [N]; Thailand: Hunsu [H] and Sanook [S].

Results showed that most of the main classes (69%) were of the 'subject' facet. All other facets except 'language' were present. It is interesting to note that some differences in categorization seem to exist between American and Asian cultures. "Humanities", for example, is often grouped with "arts" in the US directories, whereas Asians tend to either separate "humanities" from "arts," or class "humanities" under "culture" or "science." Unlike Americans that have "government" as a category, most Asian directories tend to have a combined category "government and politics." "Travel" and "life/living" were some of the popular categories in Asian directories except Chinese.

Some Asian countries seem to share similar categorization with Americans. In American and Chinese directories, "social science" and "law" categories were both present. The "game" category was prominent in American, Hong Kong and especially Korean directories. American and Korean directories also had the "world" category.

Some categories seem to be uniquely popular in certain countries. "Automobile", for example, was one of the prominent main categories in Thailand, but was not common in other countries. The "home" category was found only in the US directory. It seems that resources under "home" were classed under "life," which is a prominent category in Asian directories.

Second-Level Categories in Asian Web Directories

In order to identify subfacets that are common and would be useful for organizing Asian Web resources, the second-level categories used in all directories from the four Asian countries were analyzed. For each country, three Web directories including one local *Yahoo* and two home-grown directories were analyzed. A total of twelve directories were included for analysis.

After the preliminary analysis based on Zins' model, we discovered some issues that needed to be resolved before further analyses. One of the issues was related to the 'user' facet. We found that it is often difficult to distinguish between people as 'object' and people as 'users.' For example, the category "kids & teens" contains not only the Web sites that can be used by children, but also those that are about children, created by children, and used by teachers / parents for children. It became clear that 'people' would be much easier to use. We decided to delete the 'user' facet, and create 'people' instead in order to cover resources related to people as user, creator, object, and so on. As the 'objects' facet was supposed to cover 'organizations' and 'people' (which was already covered by the new facet 'people'), 'objects' was eliminated. To cover organizations originally under the 'objects' facet, 'organizations' facet was created.

We found that the 'reference' facet was unclear especially in the Web environment. For example, Web sites providing factual information about countries – as gazetteers or encyclopedia do – are often classed under specific headings, such as 'country profile' or 'country information,' without specifying a reference type (e.g., gazetteers, encyclopedia). This kind of heading was quite prevalent and it seemed necessary to address this issue. 'General information' was created to cover such resources. The 'media' facet was also found to be problematic. Zins (2002) indicated that graphics, pictures and radio are examples of 'media.' It seems to us that his 'media' covers the type of information as well as the type of media. As a wide variety of different types, formats, media, and tools are used to deliver information, and the distinction among them is often challenging to users, we decided to create a rather broad facet that can cover all. 'Media/types/formats/tools' replaced 'media' in order to cover different delivery media (e.g., radio, TV), formats (e.g., audio, video), types (e.g., quotations, encyclopedia), and tools (e.g., calculator, search engine).

Zins' model was modified to have the following eight facets: (1) subjects, (2) applications, (3) locations, (4) media/types/formats/tools, (5) organizations, (6) people, (7) general information, and (8) languages. Although the Web directories we analyzed did not have the 'language' facet, we concluded that language is an important facet to keep, especially for directories targeting a diverse group of users. If the Web directories were developed for a nation consisting of many immigrants (such as the US) or international users, 'language' would undoubtedly be one of the key facets.

Table 4. Zins' and Modified Faceted Classification Model for Asian Web Resources

Zins' Model	Modified Model
Subjects	Subjects
Applications	Applications
Locations	Locations
Media	Media//Types/Formats /Tools
Objects	Organization
	People
Users	
Reference	General Information
Languages	Languages

Using the modified set of facets, we proceeded and analyzed the second-level categories used in twelve different Web directories. Based on the analysis of the used categories, we tried to identify subfacets that would be useful to organize resources under each of main facets (see Table 5).

Table 5. Subfacets Identified From the Second-Level Categories

Facets	Example Categories Under Facets	Subfacet Groups
Subjects	Automobile Basketball Business law Collectibles College entrance exam Computer science Consumer rights Contemporary literature Critical theory Cultural studies Diet Emotion, marriage and love Environment & nature Hair care Health and medicine History of literature Human rights International relations Investment and money-making knowhow Korean characters and computers Linguistics and human languages Myth and fable Nano-technology Packaging Pet health Petroleum/natural gas Popular music Teaching & learning Textile printing & dyeing Thai traditional music Toys	<ul style="list-style-type: none"> • Concrete Thing, Entity (e.g., automobile; toys) • Action & Process (e.g., textile printing & dyeing; packaging) • Type of Agent (e.g., environment & nature; human rights; pet health; petroleum/natural gas) • Time [Recency] (e.g., Thai traditional music; nano-technology) • Purpose [Function] – for business; everyday info.; education; entertainment; (e.g., business law; hair care; college entrance exam; popular music) • Academic/Popular
Applications	Blogs Chats	<ul style="list-style-type: none"> • Function – for business; everyday life

	<p>Consumer services@ Downloads E-cards Financial services Forums/BBS Mobile phone logos and ring tones download Online games Professional supplies and services Sharing files Shopping</p>	<p>(e.g., financial services; e-cards; chats)</p> <ul style="list-style-type: none"> • Online/Offline • Fee/Free • One-way/Multiple way communication (e.g., downloads; chats; blogs) • Products/Services
Locations	<p>Asia Bangkok By Nations By Region Chinese County/ City/ Division Industrial and Economic Zones Korea Oceans Provinces and Districts Regions and Countries Rural life Urban life</p>	<ul style="list-style-type: none"> • Physical & Geographical division (e.g., oceans; Asia) • Political division (e.g., nations; provinces and districts) • Economic division (e.g., industrial and economic zones) • Degree of rurality (e.g., urban; rural) • Scale of area (e.g., nation, region)
Media/Types/ Formats/Tools	<p>Classifieds Contests, Surveys and Polls Diaries Directories Magazines Maps MP3@ Personal Websites Photo gallery Ratings and rankings Reviews Transportation Timetable Weather Forecast</p>	<ul style="list-style-type: none"> • Form of publication (e.g., books; journals; maps) • Type of reference (e.g., directories; reviews) • Form of information (e.g., text; image; audio) • Formal/informal • Functions (e.g., calculator; search engines; polls) • Media (e.g., radio, TV, Web)
Organization	<p>Business schools Commercial banks Embassies and consulates Fan clubs Foreign companies Government agencies International organizations Labor/Labor organization Research institutes Student organizations</p>	<ul style="list-style-type: none"> • Formal (e.g., government agencies) / Informal (e.g., fan clubs) • Function • With/without physical space (e.g., business schools; fan clubs)
People	<p>Artists Asian chess players Children Consumers Cultures and Groups Educators Families Gay and Lesbian His Majesty the King and The Royal Family Korean musicians Lawyers National & Party Leaders People</p>	<ul style="list-style-type: none"> • Age (e.g., children) • Culture/Ethnicity (e.g., Korean musicians) • Education level (e.g., higher education) • Experience level (e.g., experts) • Sexual orientation (e.g., men; gay and lesbian) • Profession (e.g., artists; lawyers) • Status (e.g., royal family;

	Scholars Women Women, housewives	CEO • Groups (e.g., clubs) • Role (e.g., parents; housewives)
General Information	Information Resources Information for medium and small business Wedding information	• Scope (e.g., general vs. specific)
Languages		• Geographical division

The list of subfacets was developed based on commonly found categories in the second level of classification systems used in Web directories. It is by no means exhaustive and comprehensive, but relevant to the categories used in the selected directories.

DISCUSSION & FUTURE STUDY

The results from the analysis of top-level categories suggested that Web directories developed in Asian countries share some similarities with one another across countries, and also with those developed in the United States. Most of the top-level categories used in American directories were present in Asian Web directories. At the same time, local directories incorporated some unique categories that could be useful for understanding the current society, culture and values. However, it should also be noted that most of Web directories analyzed in this study are commercial ones, and some categories might have been intentionally included for business purposes. Hence, we decided not to make any remarks related to the culture of each country, although some unique patterns emerged from the data collected. Our discussion will focus instead on a faceted model that could be useful for organizing Asian Web resources in particular.

Zins' model used for our preliminary analysis was helpful as it provided a rather simple framework to start with. However, we had to make some changes to accommodate unique characteristics of resources available on the Web and to make the model more user friendly. Although a pure form of faceted classification requires exhaustive and mutually exclusive facets, we decided not to seek for such a pure faceted system as it could be impractical. That is, a pure faceted system might contain facets that users cannot really understand or use. As our goal was to develop a model that facilitates the organization and retrieval of Web resources, we tried to find a more practical, clear model based on a faceted approach.

Our model of faceted classification included eight facets and over thirty subfacets that are still being developed: (1) subjects (concrete thing, entity; action and process; type of agent; recency; purpose and function; academic/popular); (2) applications (function; online/offline; fee/free; one-way/multiple-way; products/services); (3) locations (physical & geographical division; political division; economic division; degree of rurality; scale of area); (4) media/types/formats/tools (form of publication; type of reference; form of information; formal/informal; functions; media); (5) organization (formal/informal; function; with/without physical space); (6) people (age; culture/ethnicity; education level; experience level; sexual orientation; profession; status; groups; role); (7) general information (scope), and (8) languages (geographical division).

For certain facets, identifying common subfacets was relatively easy. However, other facets, such as 'subjects,' it was difficult and challenging. The 'subject' facet covers a wide variety of different topics in different contexts and disciplines. Although some subfacets that we thought might be useful have been identified and included in the model, it is clear that more in-depth analysis and research is needed to develop a more complete set of clear and useful subfacets. Dahlberg's Information Coding Classification (Dalberg, 1994), a universal ordering system for disciplines and domains, might provide some insights and help identify some subfacets for grouping different domains. It was noteworthy that the 'function' subfacet appeared under at least four different facets. Although types of functions under each facet vary, the variance is minimal. It is perhaps legitimate to make Function as one of main facets. Further

verifications and testing will be needed to determine whether it would be one of the useful main facets, and to make this facet more clear and practical.

Based on the analysis of twelve Asian and three U.S. Web directories, the current study proposed a facet approach to classifying Web resources. A total of eight main facets were developed based on Zin's model. The paper also extended the analysis by identifying subfacets for each main facet. This is the first step to a faceted classification system that could be used across different Asian countries. To make the system more complete and useful, further research is needed. Further studies could be developed along three areas. First, empirical studies could be conducted to test whether this model is applicable and useful for classifying Web resources from other Asian countries. The flexibility of a faceted approach should allow for a more culturally hospitable classification system, and this model could also be tested to identify whether it is applicable to other non-Asian countries. Second, a future study could focus on refining the list of facets and subfacets. Further analyses of lower-level categories will be necessary to identify any additional facets. Third, a prototype of a faceted browsing interface based on the proposed model can be built and tested. User testing can be conducted to evaluate the usability and performance of the proposed model. Labeling of facets will also need to be tested and made clearer to users. A faceted model can help organize Web resources in a way that facilitates access to culturally and linguistically diverse users. Users' input would be beneficial to achieving such a goal.

REFERENCE

- Bates, M. J. (2002). After the dot-bomb: Getting web information retrieval right this time *First Monday*, 7(7) retrieved on November 8, 2005, from: http://firstmonday.org/issues/issue7_7/bates/index.html
- Broughton, V. (2002). *Facet analytical theory as a basis for knowledge organization tools in a subject portal*. Proceeding of the Seventh International ISKO Conference, Granada, Spain, 135-141.
- Dalberg, (1994). Domain interaction: theory and practice. In: Albrechtsen, H. & Oernagar, S. (eds.). Knowledge organization and quality management: Proceeding of the third ISKO conference, Copenhagen, Denmark, p. 60-71.
- Devadason, F., Intaraksa, N., Patamawongjariya, P., & Desai, K. (2001). Search interface design using faceted indexing for web resources. In *Asist 2001: Proceedings of the 64th Asist annual meeting*, vol 38, 224-238.
- Ellis, D., & Vasconcelos, A. (2000). The relevance of facet analysis for World Wide Web subject organization and searching. *Journal of Internet Cataloging*, 2(3/4), 97-114.
- Gulli, A., & Signorini, A. (2005). *The indexable Web is more than 11.5 billion pages*. Proceedings of the 14th International Conference on World Wide Web, Chiba, Japan, 902-903.
- Iyer, H. (1995). *Classificatory structures : Concepts, relations and representation*. Frankfurt/Main: Indeks Verlag.
- Kwasnik, B. H. (1999). The role of classification in knowledge representation and discovery. *Library Trends*, 48(1), 22-47.
- Kwasnik, B. H., & Rubin, V. L. (2003). Stretching conceptual structures in classifications across languages and cultures. *Cataloging & Classification Quarterly*, 37(1/2), 33-47.
- Malt, B. C. (1995). Category coherence in cross-cultural-perspective. *Cognitive Psychology*, 29(2), 85-148.
- Olson, H. A. (2001). Sameness and difference: A cultural foundation of classification. *Library Resources & Technical Services*, 45(3), 115-122.
- Ross, K. A., & Janevski, A. (2005). Querying faceted databases. In *Semantic web and databases*, Vol. 3372, 199-218.
- Segall, M. H. (1990). *Human behavior in global perspective: an introduction to cross-cultural psychology*. New York: Pergamon Press.
- Yee, K.-P., Swearingen, K., Li, K., & Hearst, M. (2003). *Faceted metadata for image search and browsing*, Ft. Lauderdale, FL, United States.
- Zins, C. (2002). Models for classifying internet resources. *Knowledge Organization*, 29(1), 20-28.