Title	Task-based navigation of a taxonomy interface to a digital repository
Author(s)	Khoo, Christopher S. G.; Wang, Zhonghong; Chaudhry, Abdus Sattar
Citation	Khoo, C. S. G., Wang, Z., & Chaudhry, A. S. (2012). Task-based navigation of a taxonomy interface to a digital repository. Information research, 17(4).
Date	2012
URL	http://hdl.handle.net/10220/24120
Rights	

Contents | Author index | Subject index | Search | Home

Task-based navigation of a taxonomy interface to a digital repository

Christopher S.G. Khoo
Wee Kim Wee School of Communication & Information, Nanyang Technological University, Singapore
Zhonghong Wang
Cataloging Maintenance Center, Illinois Heartland Library System, USA
Abdus Sattar Chaudhry
Department of Library & Information Science, College of Social Sciences, Kuwait University, Kuwait

Introduction. *This is a study of hierarchical navigation; how users browse a taxonomy-based interface to an organizational repository to locate information resources. The study is part of a project to develop a taxonomy for an library and information science department to organize resources and support user browsing in a digital repository.*
Method. *The data collection was carried out using task-based navigation exercises with twenty-two participants. A cognitive framework of hierarchical navigation is proposed, involving the cognitive process of matching context, topic and/or resource type concepts to taxonomy categories.*
Analysis. *The analysis was mainly qualitative, supplemented with simple statistics and measures of prevision and recall, and error analysis.*
Results. *Though users often use the topic concept in making navigation choices, they sometimes make use of context and resource-type concepts. Users infer a variety of relationships between a task concept and a taxonomy category, including the application area, associated tool, associated process/procedure/technique, associated institution and*

academic discipline. **Conclusions.** *Users prefer to use common or generic associations in selecting categories to browse, rather than formal disciplinary relations. Some users prefer to search by people groups, contexts and institutions, rather than by subject categories. Users have difficulty distinguishing between various kinds of document and resource types.*

Introduction

This study is part of a series of studies we are carrying out on hierarchical navigation, to investigate how people locate information resources by browsing or navigating a taxonomy-based interface or a hierarchical menu system. Such interfaces are organized based on a taxonomy of terms/categories, which are used to tag resources on a website, portal or some kind of digital repository.

Taxonomies are increasingly being used to organize content within organizations and to support navigation of Web portals and digital repositories (Gilchrist & Kibby 2000; Kremer *et al.* 2005). However, not much is known about how users navigate or browse a taxonomy-based interface, the cognitive processes involved,

and how to evaluate a taxonomy developed to support navigation. Lee and Olson ([2005: 10](#)) noted that '*research on how users utilize classification or classification-like arrangements in information seeking has been scant.*' Many papers have been written on how to develop a good taxonomy (e.g., [Lambe 2007](#)), but the guidelines and procedures are based mainly on opinion, informal observations and technical considerations. There is an urgent need for more user studies of navigation and browsing of hierarchically-organized menus and interfaces, given the tremendous amount of end-user browsing taking place on websites, portals and institutional repositories.

This paper reports an evaluation study of an organizational taxonomy developed to tag and organize resources in a digital repository of a library and information science department. The evaluation was designed as task-based navigation of a hierarchical menu based on the taxonomy. Though this was originally designed as an evaluation of the taxonomy, we attempt to draw insights about how users would navigate or browse such taxonomy-based interfaces to locate information resources.

The taxonomy, called *Information Studies Taxonomy*, was developed to organize resources in a digital repository at the Division of Information Studies at the Nanyang Technological University, Singapore. The Division offers three Master's programmes in information studies, information systems and knowledge management, as well as a research Master's programme and a Ph.D. programme. The taxonomy was designed to support students and faculty in navigating/browsing the repository to locate information resources to accomplish tasks related to teaching, learning and research. The taxonomy did not cover administrative activities and technical support.

The evaluation was carried out using scenario-based navigation exercises, supplemented with interviews of participants. Each scenario contains a description of the context and one to five search tasks. The context description can be considered a representation of the *simulated work task situation* ([Borland 2000](#)), or simply the *work task* ([Li and Belkin 2008](#)). Each search task specification includes a topic and a form or resource type. An example scenario is given in Figure 1. In the example, the *context* or work task is an assignment in a course CI6124, and the first search task specifies the topic *data mining* and the form or resource type *books*. Each search task involves finding an information resource on a particular topic.

Scenario 1

Work task: Assume you are a Master's student by research. In the 1st year of study, you are taking the CI6124 *Data mining and machine learning* course. The CI6124 course is one of the Group B electives of the information systems program. An assignment of the CI6124

An example scenario

The assumption is that the user's information need has been translated into an explicit representation comprising a context, a topic and a resource type (or form). This representation may be the user's own formulation or interpretation of his or her need, or a task assigned by another person, for example the instructor of a course. Navigating a menu system or taxonomy to locate the desired resource involves a cognitive process of matching the context, topic and/or resource type to the taxonomy categories, and identifying the most likely navigation path or leaf category (bottom-most category in the hierarchy).

Several researchers have developed models of the overall information search process (e.g., Kuhlthau 1993), and others have developed models of interactive information retrieval (e.g., Spink 1997) and of information search strategies (e.g., Thatcher 2008). In these models, browsing is often treated as a simple strategy or activity that does not need further modeling or analysis. We have not come across any model of hierarchical navigation and browse searching of taxonomies and hierarchical menus. We propose the following cognitive processing steps in user hierarchical navigation:

- Step 1: Interpret the terms (labels) in a particular level of the taxonomy displayed on the screen (menu), i.e. figure out the semantics of the terms (categories) and relate the categories to the user's knowledge structure. A term may trigger a framework, schema or mental model in the user's mind.
- Step 2: Relate the task concepts to the categories; i.e. identify a potential relationship between the task concepts with each taxonomy category.
- Step 3: Hypothesize the kinds of resources likely to be located within this category, and estimate the likelihood that the desired resource will be found here.

The user then clicks on a likely category to view lower level categories and repeat the process. We shall use this framework to interpret the results of the evaluation study. As each task specification involves three main task concepts, context, topic and resource type, the user can opt to use one or more of the task concepts in the navigation. We assume that the task concept chosen by the user to match the taxonomy categories is the most salient or important one for the user.

Literature review

Like classification schemes and thesauri, taxonomies are composed of a set of categories represented by terms organized in a hierarchical structure (Gilchrist 2001; Chaudhry and Saeed 2001; Gilchrist and Kibby 2000). Taxonomies differ from classification schemes in a number of ways. Taxonomies focus on organizations or user groups and their needs, whereas classification schemes focus on disciplines or subject areas (Wyllie 2005; Chaudhry and Saeed 2001). The subject coverage of organizational taxonomies depend more on the activities of the organizations and might not follow widely accepted subject areas or domains. Wyllie (2005) noted that taxonomies focus more on corporate knowledge. Kremer, Kolbe and Brenner (2005) pointed out that taxonomies were more often used to organize content in corporate portals for the purpose of knowledge management.

Taxonomies mainly support browsing and site navigation, though they may have other applications and roles. Gilchrist (2004) pointed out that front-end navigation systems are the most common applications of taxonomies. Taxonomies have simpler structures to support user browsing and navigation, and are often constructed from multiple facets composed of sub-taxonomies to accommodate multiple perspectives in the organization.

Researchers have found that browsing is a common information searching activity, and is highly used under the right conditions. Koch, Golub and Ardö (2006) carried out a Web-log analysis of the navigation behaviour of the users of a Web portal called Renardus, which provided a common search and browse interface to the metadata records of major subject gateways (subject directories) in Europe. The browse structure was based on the Dewey Decimal Classification system. The authors found that 60% of the user activities comprised directory-style browsing using that structure. This was partly because the browse pages were indexed by search engines and many users were referred to those pages from search engine search results. Among the users who started at the homepage of Renardus, 57% opted to browse. While a majority of users limited themselves to ten or fewer steps (clicks) in the browse sequence, many did a substantial amount of browsing with up to eighty-six steps, and explored multiple branches of the hierarchy.

There are few studies of user navigation of hierarchically-organized menus and interfaces. The studies have focused on the general characteristics of the taxonomy, for example narrow and deep hierarchies versus wide and shallow hierarchies, and on the presentation of the taxonomy or interface design. Chen, Magoulas and Dimakopoulos (2005) investigated the relationship between users' cognitive styles and their preference for different kinds of hierarchical structures for browsing Web directories. The compared two cognitive styles; *field independence* and *field dependence.* Based on their study of three Web directories, Google, Alta Vista and Lycos, they found that field dependent users preferred a wide and shallow hierarchical structure, and preferred the main categories and

subcategories to be presented on different screens. In contrast, field independent users preferred a narrow and deep hierarchy, and preferred the main categories and subcategories to be displayed on the same screen.

Researchers in the field of human-computer interaction have investigated various ways of presenting taxonomy interfaces, and carried out usability studies to compare alternative interface designs. Hearst ([2006](#)) has carried out a series of studies to investigate various designs for faceted taxonomy interfaces, implemented as hierarchical faceted metadata. In a faceted taxonomy (such as the one used in our study), the top-level categories represent different facets and the categories in each facet are organized into a mini-taxonomy. English and colleagues ([2003](#)) found that users preferred a matrix design with a simultaneous or parallel display of multiple facet hierarchies, allowing users to select categories from multiple facets to form an implicit Boolean query. The matrix design was preferred over a single-tree design which allows browsing of only one facet hierarchy at a time. In a related study, Yee, Swearingen, Li and Hearst ([2003](#)) found that users preferred the matrix design over a keyword search interface.

More recently, Uddin and Janecek ([2007](#)) reported a usability evaluation of a faceted taxonomy developed for the website of an academic institution. From their task-based navigation experiments, they found the faceted taxonomy interface to be more effective and usable than the baseline system. The users appreciated the parallel display of multiple facets, and the facility to select categories from multiple facets. On average, the participants selected three facets to complete the tasks. However, some participants (especially the non-expert users) had difficulty understanding the use of facets.

Hutchinson, Druin and Bederson ([2007](#)) compared two kinds of interface designs for a children's digital library, a flat, simultaneous interface where all the leaf categories are presented simultaneously on the main screen, versus a hierarchical, sequential interface where the subcategories are presented on subsequent screens and the user can navigate only one branch of the taxonomy at a time. Users browsed the faceted taxonomy to select categories for a Boolean query. In the experiments, the flat, simultaneous interface was found to be more effective: the participants created more Boolean queries using it, were faster and also expressed preference for this interface design.

In our study, because of system limitation, the faceted taxonomy was displayed as a single-tree menu system: the branches in the hierarchy could only be explored sequentially, with the main categories and subcategories presented on different screens. Our study is not focused on the interface design and usability, but on the taxonomy, users' interaction with taxonomy categories and the resulting navigation paths.

Fang and Holsapple ([2007](#)) compared the effectiveness and usability of a subject

taxonomy versus a *usage-oriented* taxonomy, which was based on how the categorized information resources could be used. Task-based navigation experiments were carried out using an experimental website containing information on the subject of Production and Operations Management taken from a textbook on the subject. The top-level of the usage-oriented taxonomy had the categories *concepts, events, publications, organizations* and *practices*. The usage-oriented taxonomy as well as a combined usage plus subject taxonomy were found to be more effective and received higher user satisfaction and ease-of-use ratings than the purely subject taxonomy. The authors suggested that this was because the subject taxonomy represented a discipline-specific knowledge structure which the user might not be familiar with, whereas usage, functions and procedures were more similar across domains and more likely to be familiar to users. In our study, the non-subject facets can be considered to form a usage-oriented taxonomy.

The user studies described above have investigated some general characteristics of browse taxonomies, the structure (breadth and depth), the types of categories included (e.g., non-subject facets) and the presentation (sequential or parallel). However, they did not carry out an in-depth analysis of user browsing behaviour , for example users' selection of categories and navigation paths, the cognitive processes involved and the difficulties encountered. Lee and Olson ([2005](#)) noted that such studies were rare, and proceeded to do a small study of how twenty-four library and information science students used *Yahoo! directories*. They found that to understand the hierarchical relationship requires knowledge of the concepts in the domain and their relationships. It was sometimes tricky for the participants to select the categories at the right level of specificity. The participants also sometimes selected the facets in a different citation order than the conventional order. Advantages of hierarchical navigation most frequently mentioned by the participants included speed and ease of use, high precision, and showing relationship between topics. Disadvantages most frequently mentioned included low recall, requiring knowledge of a particular subject hierarchy, and requiring understanding of the concept of hierarchical arrangements.

Some of the in-depth studies of hierarchical navigation were of children carrying out browse searches on an online library catalogue, digital library or Web directory. Behesthi, Large and Tam ([2010](#)) examined the transaction logs for a children's portal (Web directory) on Canadian history, and found that 42% of the transactions involved the use of the hierarchical browse interface which was based on a subject taxonomy, compared to 18% for the keyword and advanced search.

Borgman, Hirsh, Gallagher and Walter ([1995](#)) compared children's searching on two kinds of online catalogue interfaces, a hierarchical browse interface based on the Dewey Decimal Classification, and a keyword search interface. They found that children were able to use different versions of the hierarchical interface effectively and quickly. The authors noted that domain knowledge was needed when interacting with the browse interface to decide which category to check first. The

children performed better for the science domain, where the logical structure of the classification scheme was clearer, compared to the technology domain.

Clearly, a lot of cognitive processing takes place during browsing and hierarchical navigation, and we hope the results of this study will provide more insights on the process.

The information studies taxonomy

The procedure used in this study for constructing the taxonomy has been described in detail in (Wang *et al.* 2010). The classificatory structure and categories were constructed on the basis of many sources:

- *Sources from the school*: including course syllabi, research proposals of research students, PhD and Master's theses, publications of students and faculty, the school Website and intranet.
- *Sources from the community*: *Guidelines for professional library/information educational programs - 2000* (Daniel *et al.* 2000), course descriptions from other library school websites.
- *Domain taxonomies*: the information science taxonomy (Hawkins *et al.* 2003), two information systems taxonomies (Mentzas 1994; Doke and Barrier 1994), and categories in the area of knowledge management suggested by Cheung, Lee and Wang (2005);
- *General classification scheme and domain thesauri*: Dewey Decimal Classification and three domain thesauri (Library and Information Science Abstracts, American Society for Information Science and Technology, and the Educational Resources Information Center (ERIC) were used as sources for the subject facet.

A faceted organization scheme was selected to structure the taxonomy. Five major facets had been identified from an analysis of existing resources, interviews with stakeholders and an analysis of the user tasks that the repository was meant to support:

- Courses
- Research groups
- Resource types (course material types, document types, snf reference types)
- Information types
- Topics (the subject facet).

The first version of the *Information studies taxonomy*, used in this study, comprised seven facets and about 540 categories. Table 1 lists the facets and the main categories in the subject facet. An outline of the taxonomy with example categories for each facet is given in the Appendix. The subject facet (*Topics*) was the largest with twelve main categories and more than 440 categories. The hierarchical structure of the twelve main categories varied from two to nine items

in width and two to five levels in depth.

The taxonomy was implemented in the University e-learning platform using the TLE-Equella software. The taxonomy was deployed in a way that did not allow the participants to visualize the whole tree structure of the taxonomy. The interface only allowed the participants to navigate the taxonomy top-down, displaying one level at a time. The user had to click on a category to view the subcategories on a new screen.

Facets		Sub-categories	Width*	Depth*
Courses		15	3	3
Research groups		4	4	1
Course materials types		10	6	1-2
Document types		14	10	1-2
Reference types		15	15	1
Information types		40	35	1-2
Topics	Information science and peripheral fields	28	2	2-3
	Information institutions	17	3	2
	Information and knowledge management	29	6	1-3
	Collection management and user services	30	3	2-4
	Information and knowledge organization	56	6	2-4
	Information searching and retrieval	51	4	2-4
	Information technologies	104	9	2-5
	The information society	37	5	2-4
	The information industry	28	8	2-3
	The information profession	10	6	1-2
	Education and training	18	9	1-2
	Research methodologies and scholarly writing	34	7	1-3
Note: * Width = the number of categories at the top level of the sub-hierarchy, and depth = the number of levels in each navigation path from a top level category to a leaf category.				

Table 1: Top level of the *Information studies taxonomy*, and its size

Task-based taxonomy evaluation

Eighteen students from various programmes and four instructors participated in the navigation exercises. Twenty-two scenarios were designed for the study. The scenarios were constructed so that they were relevant to the roles of the participants who were assigned the scenario. They covered course assignments; research-related tasks such as literature review, data analysis, and data collection; academic paper writing tasks such as creating publications such as journal papers; and teaching activities such as updating course lecture slides. Each scenario contained a description of the context (the work task) and one to six search tasks. The twenty-two scenarios contained a total of seventy-five search tasks.

Each participant was assigned two scenarios and each scenario was assigned to two participants. Participants were allowed to navigate or select multiple paths and branches that they thought might lead them to the desired resource. Since the taxonomy did not have any actual resources attached to the categories (the taxonomy had not yet been used to tag resources in the repository), the

participants did not have to stop on finding a desired resource and could select as many navigation paths as they thought appropriate and sufficient for the task. This allowed us to find out the range of likely navigation paths that users might select.

For each of the seventy-five tasks, the expected selections of facets and navigation paths (expected answers) were prepared based on the task concepts. The tasks had 137 expected selections of facets (top-level category) and 143 navigation paths. (Since each task was assigned to two participants, a total of 274 selections of facets and 286 navigation paths were expected.) Table 2 lists the number of tasks with one, two or three expected facets (top level categories) and one, two or three expected navigation tasks. The majority of tasks had two expected facets and two expected navigation paths. For each facet, there can be multiple expected navigation paths because of multiple branches that the user can take. Table 3 lists the number of tasks for which each facet was expected to be selected. The table shows that sixty-six of the tasks involved the subject facet.

No. of expected facets	No. of tasks
One facet	19 (25.3%)
Two facets	50 (66.6%)
Three facets	6 (8.0%)
Total	75 (100%)
No. of expected navigation paths	No. of tasks
One path	18 (24.0%)
Two paths	46 (61.3%)
Three paths	11 (14.6%)
In total	75 (100%)

Table 2: Number of tasks with 1 2 or 3 expected facets and navigation paths

Facets	No. of tasks
Courses	8
Course material types facets	8
Research groups	4
Document types	32
Reference types	7
Information types	12
Topics (the subject facet)	66

Table 3: For each facet, the number of tasks having the facet as the expected facet or navigation starting point

Evaluation results

The twenty-two participants provided in total 150 responses (75 tasks x two participants). Each response comprised one or more selections of facets (top-level categories) and navigation paths. Table 4 lists the number of expected and actual responses. As shown in the table, about 70% of the responses involved more than one facet and more than one navigation path. Twenty per cent selected four or more paths. The participants selected more facets (top-level categories) and more

paths than expected. For example, for the task of *MARC format standard* for the H6613 course (the context), one participant selected two additional facets and four additional navigation paths. The participant could focus on the *context* (*H6613* course), the *form* (*standard*) or the *subject* (*MARC format*). The participant selected all three concepts. For *form*, the participant thought that *standard* could be both a *Reference type* or an *Information type*.

No. of facets	No. of actual responses	No. of expected responses
One facet	44 (29.3%)	38 (25.3%)
Two facets	69 (46%)	100 (66.6%)
Three facets	27 (17.3%)	12 (8%)
Four facets	10 (7%)	0
In total	150	150
No. of navigation paths	**No. of actual responses**	**No. of expected responses**
One path	34 (22.6%)	36 (24%)
Two paths	49 (33.3%)	92 (61.3%)
Three paths	37 (24%)	22 (14.6%)
Four paths	24 (16%)	0
Five paths	3 (2%)	0
Six paths	3 (2%)	0
In total	150	150

Table 4: Number of facets and navigation paths selected in responses

Table 5 lists the number of expected and actual selections of the seven facets, and their precision and recall measures. These measures are not meant to be used as retrieval effectiveness measures, but as an indication of how close the users' category selections are to those that the researchers expected. A high precision indicates that the users made the expected selections. If this is coupled with a low recall, it suggests that the users didn't *work hard* to identify more navigation paths. A high recall indicates that the users identified most of the selections anticipated by the researchers. If this is coupled with a low precision, it indicates that users made more selections than expected. Indeed, the participants provided more selections than expected in all the facets, except for *topics* and *document types*.

Facets	No. of expected selections	No. of actual selections*	Precision+	Recall#
Courses	16	34 (11 + x23)	32.3%	68.7%
Course Material Types	16	23 (10 + x13)	43.4%	62.5%
Research Groups	8	10(0 + x10)	0%	0%
Document Types	64	61(54 + x7)	*88.5%*	*84.3%*
Reference Types	14	24(10 + x14)	41.6%	71.4%
Information Types	24	33(18 + x15)	54.5%	75%
Topics (the subject facet)	132	118(115 + x3)	*97.4%*	*87.1%*
In total	274	303 (218+x85)		
Notes:				

**Table 5: Expected and actual selections of facets, and
their precision and recall scores**

Precision

The *Topics* facet and *Document types* facet had the highest precision (as well as recall). As expected, the participants usually matched the subject term to the *Topics* facet and the form to the *Document types* facet. There were, however, three instances where the participant matched the context information to the *Topics* facet.

The lower precision for the other facets was because the participants selected more facets than expected. Most of the additional selections of the *Courses* facet were because of the course information given in the scenario. The course title in the context was thus a salient concept for the participant to search.

On the other hand, four participants looked for subject terms under *Courses* that might cover the topic. For example, for the task *Websites of Internet programming languages such as ASP and JavaScript*, a participant selected the following paths:

- Courses > information systems programme > group A electives > CI6206 Internet programming
- Courses > information studies programme > group A electives > H6614 Internet & Web technologies

Clearly, courses were salient concepts to the participants, and were deemed relevant for searching.

The *Research group* facet was clearly a problem with 0% precision and recall, indicating that the participants used the facet rather differently than the way the researchers intended or expected. Most of the students were not affiliated with research groups and did not understand how the research groups were relevant to them. The *Research groups* facet was listed above the *Topics* facet in the interface, and so participants explored this category first to locate subject terms, thinking that subjects were associated with particular research areas or people groups (i.e., researchers on the subject).

The *Course material types, Reference types* and *Information types* facets obtained lower precision because the participants had their own interpretations of these facets and selected them more often than expected. These facets are all related to

resource types or form. Participants looked for certain types of teaching and learning materials under *Course material types*. For example, for the task *Information visualization examples and software*, a participant selected *Course material types > course document types > tutorials.* For the task *Government publications on information literacy standards*, a participant matched *standards* to:

- Reference types > government publications
- Information types > policies

The participant claimed in the interview that *information literacy standard*s were a kind of *policies*.

Clearly, the participants had difficulty distinguishing between the facets that relate to form. *Document types* appeared to be the most familiar to the participants, but in the post-exercise interviews 73% said they had difficulty distinguishing *Document types* from *Course materials types* and *Reference types.* 77% had difficulty understanding *Information types*.

Recall

We analysed the seventeen cases where participants did not select a *Topics* facet as expected:

- Six could not find a matching subject category under *Topics*.
- Four participants had selected *Research groups* and decided that was sufficient.
- Two gave up on the *Topics* facet because of its large number of categories and complicated structure. They just selected *Document types -> books*!
- Three preferred *Courses* and *Course material types*
- Two preferred *Information types.*

For the two facets of *Courses* and *Course material types*, 77% of the participants had difficulty distinguishing between them and suggested combining them, such as inserting course material types under each course titles.

For the *Reference types* facet, different groups of participants had different interpretations. Information studies students interpreted the facet to cover general references such as dictionaries and encyclopedias. However, students from the knowledge management programme expected the facet to cover all materials (references) related to their studies, other than course materials. 32% of the participants, most of them with no library science background, were not clear about the meaning of *Topics*!

The topics facet: main categories

We now examine the *Topics* (subject) facet more closely. There were 138 expected

navigation paths involving the *Topics* facet, and 153 actual navigation paths.

The first step in navigating the *Topics* facet is to select the main (top-level) category. The *Topics* facet has twelve main (top-level) categories, as listed in Table 1. The participants selected different (unexpected) main categories for seven tasks (nine navigation paths). For example, for the task *Metadata format*, a participant selected *Information storage and retrieval* rather than *Information and knowledge organization.* She explained in the interview that she had learnt the *metadata* concept in an *information retrieval systems* course. She perceived a type of relationship between *metadata* and *information retrieval systems*, and the association was stronger or more salient than with the expected *Information and knowledge organization.*

Thus, some participants selected different relationships between the task concept and the taxonomy categories than expected. They selected:

- A software tool (under Information technologies), rather than an academic discipline (Information science)
- An academic discipline (Information science), rather than types of Institutions
- An application (Information retrieval system, Digital libraries, Automatic classification), rather than an academic discipline (Information and knowledge organization, Machine learning)
- A different academic discipline: Computer science rather than Computer graphics
- A process or technique (Indexing and abstracting, Research method), rather than an application (text processing).

The interviews revealed that the participants preferred frameworks that they were familiar with as the main categories. Forty per cent of the participants complained that they had to spend time to be familiar with the main categories before making choices. They suggested using frameworks such as the three Master's programmes in the taxonomy, and widely accepted disciplines such as library science, information science, and computer science.

The topics facet: lower-level categories

We now examine the cases where the participants selected the expected main category but different lower categories (branches). The choices appear to reflect the different academic contexts where the participant had encountered the task concept. For example, the students associated *organizational culture* and *leadership* with knowledge management and knowledge organizations (they mistakenly thought 'Information and knowledge *organization*' referred to a type of institution, rather than organization of knowledge).

Users may also select more generic contexts that they are familiar with rather than

a more specialized context that is harder to locate in the taxonomy:

- For *HTML and other markup languages*, a participant selected *Networks > Wide-area networks > World Wide Web* and *Computer software > Computer programming > Computer programming languages* rather than the more specific context of *Multimedia > Hypertext > Markup languages.*
- For *Internet programming languages*, a participant selected the generic *Computer software > Computer programming > Computer programming languages* rather than *Multimedia > Hypertext > Internet programming languages.*

Conclusion

The lessons we have learnt from the evaluation study can be summarized as follows:

- Users are prepared to explore multiple navigation paths to locate a resource. Some users explored various top-level categories to understand them, before performing the tasks. Similar behaviour was observed by Large *et al.* ([2009](#)), who found that several children using the Canadian history portal tried all the top-level categories.
- Users are creative in inferring a variety of relationships between a task concept and a taxonomy category. The relationships include application area, associated tool, associated process, procedure or technique, associated institution and academic discipline. It is not easy to predict which relationship a user will find salient and which top-level category the user will find relevant. Large *et al.* ([2009](#)) found that students sometimes had trouble selecting the top-level entry point to the taxonomy. *Canoes* was occasionally sought under *Aboriginal peoples* rather than *Transport*, and *Vaccines* under *Everyday life* rather than *Science and technology.*
- Users associate topics with the contexts (e.g., courses) in which they encountered the topic and may not understand the formal disciplinary relations found in subject classification systems. Users prefer to use common or generic associations.
- Some users are not familiar with browsing a subject classification system, and may prefer to search by people groups, contexts and institutions.
- Users have difficulty distinguishing between various kinds of document types, resource types and formats. Certain resource types are associated with particular scenarios or contexts, for example *Course material* types with *Courses*. Rather than providing separate facets for different kinds of resource types, they should be used as subdivisions for the associated contexts.
- Some users are *lazy* and will not explore complex structures or long lists of items.

We have assumed that a user's information need can be represented as a context-

topic-resource type triple, and that navigating a taxonomy-based interface involves a cognitive process of matching the task concepts to the taxonomy categories. The user can opt to use either the *context*, *topic* or *resource-type* concept for the matching and navigation. Though users most often used the *topic* concept in making navigation choices, they did make use of the *context* and *resource-type* concept quite frequently. For some users, the *context* or *resource-type* concept was somehow salient or seemed a good top-level category to start the navigation. When and why users decide to use the *context*, *topic* or *resource-type* is not known, and merits further study.

The decision may be influenced by the top-level categories offered on the main menu. Users may select a particular top-level category if they think they understand the organization structure underlying the category (i.e., can predict what the subcategories are likely to be). It is not known which comes first, the selection of task category to search or the examination of the top-level categories to select. Does the user mentally select a task concept first and then look for a matching category on the menu, or first examine the category choices offered to see which task concept they invoke in short-term memory?

Which category the user selects may also depend on how strongly each menu category is associated with a task concept, and the type of relationship between menu category and the task concept. In this study, two of the top-level categories (*courses* and *research groups*) refer to the context, and four top-level categories represent resource types. Users do not always match the *topic* concept in the task to a *topic* category in the menu, or a *context* concept to a *context* category. They may, for example, look for a *topic* in a *context* facet, or a *context* concept in the *topic* facet. Even when users look for a task topic in the topic taxonomy, the relationship between the task topic and the taxonomy category may not be the expected subsumption (*is-a*) relationship, but some other relationship such as *application area* or *associated tool*. The taxonomy category may represent an academic or topical context in which the user has previously encountered the task concept (e.g. a course the user has taken before).

In future work, we plan to identify the main types of contexts that task concepts can be associated with, the types of relationships that often occur between task concepts and taxonomy concepts, and investigate how these associated contexts and relationships can be identified during taxonomy construction.

About the authors

Christopher Khoo is an Associate Professor in the Wee Kim Wee School of Communication & Information, Nanyang Technological University, Singapore. He received his PhD in Information Transfer from Syracuse University and MSc in Library & Information Science from the University of Illinois at Urbana-Champaign. He can be contacted at: chriskhoo@pmail.ntu.edu.sg.

Zhonghong Wang is a cataloguer at the Illinois Heartland Library System, Edwardsville, Illinois, USA. She received her PhD from the Nanyang Technological University, Singapore, and her Masters in Information Management from Peking University. She can be contacted at: jwang@illinoisheartland.org.

Abdus Sattar Chaudhry is the Programme Director of the Master of Library and Information Science programme at the College of Social Sciences, Kuwait University. He received his PhD from the University of Illinois at Urbana-Champaign, and a Master's from the University of Hawaii. He can be contacted at: abdusattar.chaudhry@ku.edu.kw.

Behesthi, J., Large, A. & Tam, M. (2010). Search patterns on a children's portal. *Annual CAIS/ACSI Conference Proceedings*, **38**. Toronto: Canadian Association for Information Science. Retrieved January 18 2012, from http://www.cais-acsi.ca/proceedings/2010/CAIS059_Beheshtietal_Final.pdf (Archived by WebCite® at http://www.webcitation.org/6BctVgAuh).

Borland: (2000). Experimental components for the evaluation of interactive information retrieval systems. *Journal of Documentation*, **56**(1), 71-90.

Borgman, C.L., Hirsh, S.G., Gallagher, A.L., & Walter, V.A. (1995). Children's searching behavior on browsing and keyword online catalogs: The Science Library Catalog Project. *Journal of the American Society for Information Science*, **46**(9), 663-684.

Chaudhry, A. S. & Saeed, H. (2001). Taxonomies applications for leveraging organizational knowledge resources. *Singapore Journal of Library & Information Management*, **30**(1), 45-52.

Chen, S.Y., Magoulas, G.D. & Dimakopoulos, D. (2005). A flexible interface design for Web directories to accommodate different cognitive styles. *Journal of the American Society for Information Science & Technology*, **56**(1), 70-83.

Cheung, C. F., Lee, W. B. & Wang, Y. (2005). A multi-facet taxonomy system with applications in unstructured knowledge management. *Journal of Knowledge Management*, **9**(6), 76-91.

Daniel, E., Lazinger, S. & Harbo, O. (2003). *Guidelines for professional library/information educational programs - 2000*. (3rd rev. draft, 2003). The Hague, The Netherlands: International Federation of Library Associations and Institutions.

Doke, E. R. & Barrier, T. (1994). An assessment of information systems taxonomies: time to be re-evaluate? *Journal of Information Technology*, **9**(2) 149-157.

English, J., Hearst, M., Sinha, R., Swearingen, K. & Yee, K.-P. (2003). Flexible search and navigation using faceted metadata. Berkeley, CA: University of Berkeley, School of Information Management and Systems.

Retrieved October 8, 2012, from http://flamenco.sims.berkeley.edu/papers/flamenco02.pdf (Archived by WebCite® at http://www.webcitation.org/6Bct6KfvT).

Fang, X., & Holsapple, C.W. (2007). An empirical study of web site navigation structures' impacts on web site usability. *Decision Support Systems*, **43**(2), 476–491.

Gilchrist, A. & Kibby: (2000). *Taxonomies for business: access and connectivity in a wired world*. London: TFPL Ltd.

Gilchrist, A. (2001). Corporate taxonomies: report on a survey of current practice. *Online Information Review*, **25**(2), 94-103.

Gilchrist, A. (2004). The taxonomy: a mechanism, rather than a tool, that needs a strategy for development and application. In Gilchrist, A. & Mahon, B. (Eds.), *Information architecture: designing information environments for purpose* (pp. 192-198). London: Facet Publishing.

Hawkins, D. T., Larson, S. E. & Caton, B. Q. (2003). Information Science Abstracts: tracking the literature of information science. Part2: A new taxonomy for information science. *Journal of the American Society for Information Science and Technology*, **54**(8), 771-781.

Hearst, M. (2006). Design recommendations for hierarchical faceted search interfaces. In ACM SIGIR Workshop on Faceted Search 2006. Retrieved October 8 2012, from http://flamenco.sims.berkeley.edu/papers/faceted-workshop06.pdf (Archived by WebCite® at http://www.webcitation.org/6BctwbKF8).

Hutchinson, H.B., Druin, A. & Bederson, B.B. (2007). Supporting elementary-age children's searching and browsing: design and evaluation using the international children's digital library. *Journal of the American Society for Information Science and Technology*, **58**(11) 1618-1630.

Koch, T., Golub, K. & Ardö, A. (2006). Users browsing behaviour in a DDC-based web service: A log analysis. *Cataloging & Classification Quarterly*, **42**(3/4) 163-186.

Kremer, S., Kolbe, L. M. & Brenner, W. (2005). Towards a procedure model in terminology management. *Journal of Documentation*, **61**(2) 281-295.

Kuhlthau, C. C. (1993). *Seeking meaning: a process approach to library and information services.* Norwood, NJ: Ablex.

Lambe: (2007). O*rganizing knowledge: taxonomies, knowledge and organizational effectiveness.* Oxford: Chandos Publishing.

Large, A., Beheshti, J., Clement, I., Tabatabaei, N. & Tam, M.T.Y. (2009). Visualizing a hierarchical taxonomy in a children's Web portal: user evaluations of a prototype. *The Canadian Journal of Information and Library Science*, **33** (3/4), 255-282.

Lee, H. & Olson, H.A. (2005). Hierarchical navigation: An exploration of Yahoo! directories. *Knowledge Organization*, **32**(1) 10-24.

Li, Y. & Belkin, N.J. (2008). Types of tasks & facets of tasks: a faceted approach to conceptualizing tasks in information seeking. *Information*

Processing and Management, **44**(6) 1822–1837.

Mentzas, G. (1994). A functional taxonomy of computer-based information systems. *International Journal of Information Management*, **14**(6), 397-410.

Spink, A. (1997). Study of interactive feedback during mediated information retrieval. *Journal of the American Society for Information Science*, **48**(5), 382–394.

Thatcher, A. (2008).Web search strategies: the influence of Web experience and task type. *Information Processing & Management*, **44**(3) 1308–1329.

Uddin, M.N. & Janecek: (2007). Performance and usability testing of multidimensional taxonomy in web site search and navigation. *Performance Measurement and metrics*, **8**(1) 18-33.

Wang, Z., Chaudhry, A.S. & Khoo, C. (2010). Support from bibliographic tools to build an organizational taxonomy for navigation: use of a general classification scheme and domain thesauri. *Knowledge Organization*, **37**(4) 256-269.

Wyllie, J. (2005). *Taxonomies: frameworks for corporate knowledge* (2nd ed.). London: Ark Group, in association with Inside Knowledge.

Yee, K-P., Swearingen, K., Li, K. & Hearst, M. (2003). Faceted metadata for image search and browsing. In *CHI '03: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 401-408). New York, NY: ACM Press. Retrieved 8 October 2012, from http://people.ischool.berkeley.edu/~hearst/papers/flamenco-chi03.pdf (Archived by WebCite® at http://www.webcitation.org/6BctlgqVW).

Find other papers on this subject

Check for citations, using Google Scholar

Appendix. Outline of the Information Studies Taxonomy (Version 1)

Courses

- Information studies programme
- Core courses
- Group A electives
- Group B electives
- ...
- Information systems programme
- ...

Research groups

- Digital libraries and information Retrieval
- Information and knowledge management
- ...

Course material types

- Course documents
- Course outlines
- Lectures
- ...
- Assignments
- ...

Document types

- Books
- Dissertations & theses
- Essays
- ...

Reference types

- Almanacs
- Bibliographies
- Biographies
- ...

Information types

- Checklist
- Components/structures
- Concepts/terms

- ...

Topics

- Information science and peripheral fields
- Information institutions
- Information and knowledge management
- Collection management and user services
- ...

---

103   © the authors, 2012.
Last updated: 3 December, 2012