Refereed article received
4 July 2006
Approved for publication
12 September 2006

The implementation of faceted classification in web site searching and browsing

Mohammad Nasir Uddin
Department of Library and Information Science, University of Rajshahi, Rajshahi, Bangladesh, and
Paul Janecek
Computer Science and Information Management Program, Asian Institute of Technology, Pathumthani, Thailand

Abstract
Purpose – The aim of this paper is to develop and implement a multidimensional classification system in the web that can provide an alternative but convenient structure for organising and finding information content.

Design/methodology/approach – A prototype system is developed following the views of Ranganathan's faceted classification, which is to provide multiple classifications of the web documents through content oriented metadata organised under different facets (orthogonal groups of categories).

Findings – Based on an architectural framework this study demonstrates a prototype faceted classification system (FCS) that is integrated into a general open-source content management system and populated with a sample collection of institutional web pages/documents.

Originality/value – The study provides significant grounds for the IR community to improve interface structure for easy access, management, and retrieval of web information. In addition, the integration of content management tools with multidimensional taxonomies can be a new instance of a corporate web system for easy content creation, organisation, and navigation.

Keywords Classification, Content management, Worldwide web

Paper type Technical paper

Introduction
In seeking information from the web the most common usability problem is the poorly organised search results and poor information architecture affecting most of the corporate sites (King, 2001). The rapid growth and heterogeneity of web information makes it increasingly difficult for web site designers to organise information coherently and for users to find information easily. Most users face difficulty on a corporate site because the site organisation and the search result do not clearly describe the content of the site. In addition, the absence of well-suited information retrieval tools creates inefficiency in searching, which often forces the user to sift through long lists of results. As a result, a major portion of corporate information remains unused.

To alleviate these problems this study proposes the use of a multidimensional classification system, which is useful for both organising the information in a web site as well as search results. Specifically, the approach that has been developed in

Online Information Review
Vol. 31 No. 2, 2007
pp. 218-233

1468-4527
DOI 10.1108/14684520710747248

this study classifies web documents through content-oriented metadata organised under multiple orthogonal groups of categories, called facets, which are based on Ranganathan's (1960) faceted classification. The use of facets and subsequent metadata isolates the classification data on a particular item and provides a multidimensional taxonomy that exposes the essential categories of domain content where users may find information. Thus, classifying and organising documents by means of faceted metadata provides ease in browsing and searching with a more relevant search result. The approach is illustrated in Figure 1.

This paper provides an architectural framework and demonstrates a prototype faceted classification system (FCS) which are integrated into a general open source content management system and populated with a sample collection of institutional web pages/documents. By combining faceted classification and content management tools this approach is not only suitable for quick finding of relevant information from the whole collection of an organisation's web site through multiple points, but also for managing and organising the collection efficiently. The approach is different from the usual searching and browsing of a web site's one-dimensional hierarchical information structure: instead it provides an alternative structure of arranging and finding content through multidimensional taxonomy to achieve consistent, effective and flexible searching and browsing.

In the following sections we first introduce the concept and importance of faceted classification. After that a short description of the related works is provided to show the current initiatives. We then provide our framework and describe the implementation details of the prototype faceted classification system, followed by a brief discussion of the evaluation result. Finally, we present our conclusion with a summary of the contribution and future research.

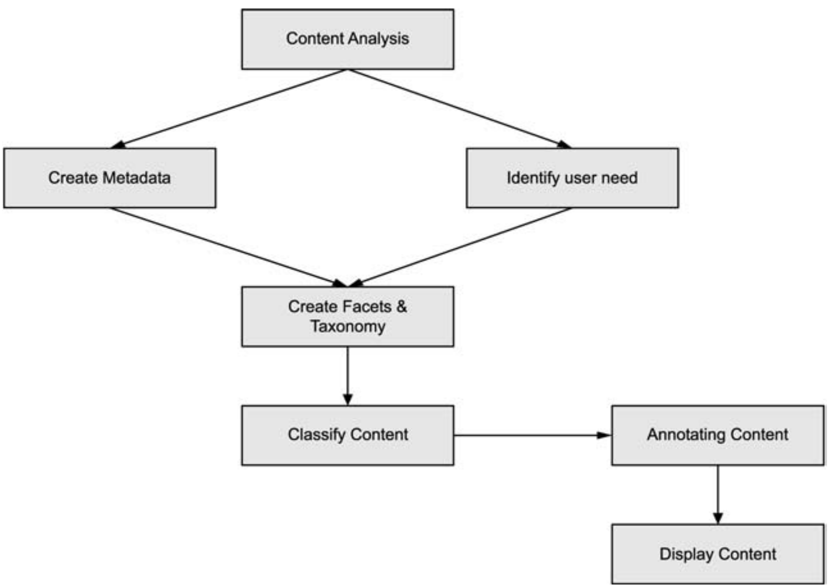


Figure 1.
Classifying documents through faceted metadata

Faceted classification

In organising information most of the classification schemes used by web sites follow simple top-down hierarchical organisations, where entities are listed by broader categories and become more specific in the lower level. In such structures each object is typically located in only one category. This is not a limitation of hypertext; it is a limitation of the tools developers use to manage large web sites. In contrast faceted classification is a bottom-up approach, where each object is tagged with a set of attributes and values, i.e. facets, and the organisation of these objects determines how a user may choose to access them. Thus, objects in a faceted collection are not limited to a single location, as it offers multiple independent taxonomies to identify a resource.

In short, faceted classification is a method of multidimensional description and arrangement of information resources by their concept, attributes or "aboutness". It addresses the fact that users may look for a document resource from any number of angles corresponding to its rich attributes. By encapsulating these distinct attributes or dimensions as "facets", the classification system may provide multiple facets, or main categories of information, to allow users to search or browse with greater flexibility (Louie *et al.*, 2003; Rao, 2002).

Each of the facets is a mutually independent category and may contain any number of isolates (content-oriented metadata/concepts) or subcategories arranged in a hierarchy. The category hierarchy can then be used to describe, organise and access the resource by browsing or querying. For example, the facet within the domain of a shopping site could be "Item type", "Brand", "Fashion" and "Price", and the positioning of this particular facet isolate might be "T-shirt", "Denim", "Summer" and "$50-80". Hypothetically, someone might be price conscious and want to start there; another might have a specific fascination with a brand and want to begin with that. Whatever facet they choose, it will lead them to discover the desired item. Thus, faceted classification can fit very well in the web information architectures that allow for searching and navigation directed by the user. Through this classification a large dataset can be filtered progressively, following the user's various choices until arriving at a manageable set to meet the user's basic criteria.

The logical and predictable structure of the faceted system makes it compatible with the requirements of classifying and organising electronic documents in a way that enumerative and pre-coordinated systems are not, and can serve as the basis of all methods of information retrieval (Broughton, 2006). In the 1930s Ranganathan introduced the idea of decomposing and organising complex subjects by their elemental concepts which he called facets. For subject classification he provides five facets in his Colon Classification (CC) scheme, such as Personality (what the object is), Matter (what something is made of), Energy (how something changes, is processed, evolves or is actioned), Space (where something is) and Time (when it happens). For example, if a document discusses "the design of wooden furniture in eighteenth century America", the facets would be as follows:

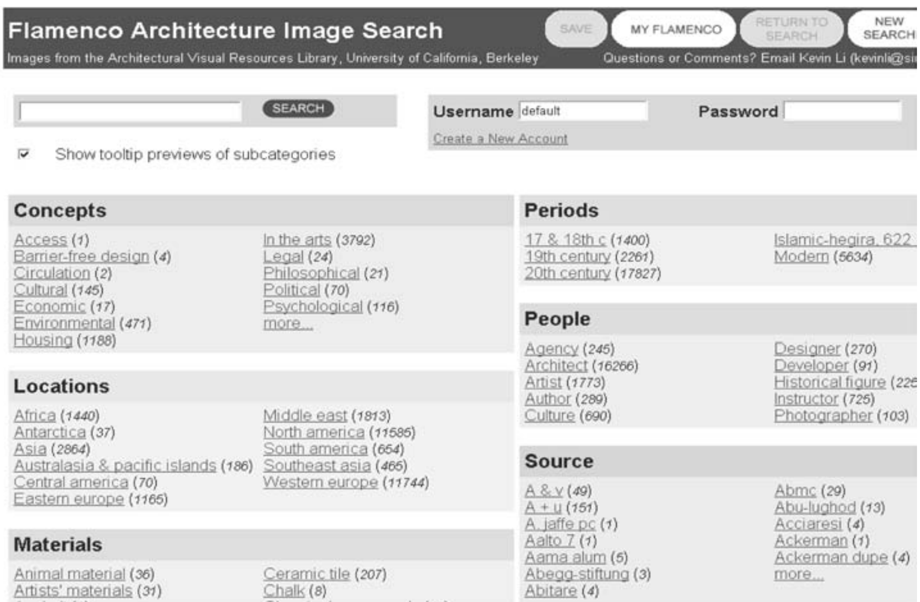
- Personality – furniture;
- Matter – wood;
- Energy – design;
- Space – America; and
- Time – eighteenth century.

The document resource is described by aggregating and combining the information attributes under each facet. "Wood" is a piece of that description which covers an area that none of the other pieces cover (Taylor, 1999).

Related works

The idea of implementing faceted classification in subject directories and search engines became significant by the end of 1990s while some researchers were addressing the issue in corporate web searching for quick retrieval of text and images from large database collections (Ellis and Vasconcelos, 2000; Yee *et al.*, 2003). Our work adopted the ideas of some of these research findings (Chen and Dumais, 2000; Yee *et al.*, 2003; Pratt *et al.*, 1999; Chen *et al.*, 1998), where authors suggested that users often like to see the search results within a predictable organisation of category hierarchies arranged with content-oriented metadata in several dimensions. Some researchers have applied hierarchical clustering to retrieval results to dynamically generate categories (Hearst, 2000; Hearst and Pedersen, 1996; Maarek and Wecker, 1994; Allen *et al.*, 1993). However, studies show that users often find the results of clustering difficult to interpret and prefer predictable hierarchies of metadata. They like to have "browsing the shelves" experience in online searching much the same as in physical libraries (Borgman, 1996; English *et al.*, 2003).

More recently, some web sites have been implementing multidimensional classification for search and navigation. For example, Flamenco (Figure 2) (Flamenco Fine Arts Search, available at: http://orange.sims.berkeley.edu/cgi-bin/flamenco/ accessed 21 October 2005) is a web-based image browser that uses hierarchical metadata as facets to organise the image collection in multiple ways



Source: http://orange.sims.berkeley.edu/cgi-bin/flamenco.cgi/famuseum/Flamenco

Figure 2. Faceted classification example from Flamenco image search

simultaneously, such as by concepts (e.g. economic, psychological), material (e.g. clay, ceramic), date (e.g. nineteenth century), and location (e.g. Asia).

Another example is Epicurious (Figure 3) (available at: http://www.epicurious.com, accessed 12 September 2005), a cooking web site that provides search and browsing over a large collection of recipes using hierarchical metadata facets. The browse interface consists of facets such as main ingredients (e.g. beans), cuisine (e.g. Asian), and preparation method (e.g. bake). These examples allow users to access every item in the collection through each of the facets.

However, these examples demonstrate the usefulness of faceted search over collections of the same type of item (e.g. images, recipes). Our research investigates the suitability and effectiveness of faceted classification in an institutional web site, where the content is often diverse and the appropriate structure unclear.

Proposed system framework

A framework as shown in Figure 4 is developed to implement the Faceted Classification System (FCS) within the context of a content management system (CMS) for an institutional web site. It aims at enhancing the organisation, sharing, and access to the content of the CMS through a dynamic multidimensional navigation and browsing structure. The capabilities of the proposed system framework depend on the

Source: http://www.epicurious.com/recipes/find/advanced/

Figure 3.
Use of facets in the Epicurious web sites

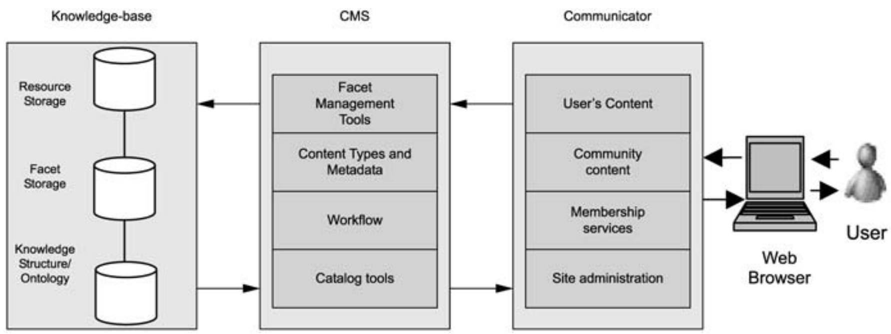


Figure 4.
The system framework

following three components: knowledge base; content management system; and communicator.

Knowledge base

The knowledge base is the core component of the architecture, as it stores the facets, the knowledge (i.e. domain thesaurus used to construct the facets) and an index to resources external to the CMS, which is primarily a collection of document resources stored both locally and on the internet.

The facet storage contains the facets and the subcategory of metadata (isolate) to classify a resource within the domain. More specifically, each of the categories of facets consists of terms or metadata concepts arranged in a generic relation to denote resources. Domain ontologies are used to provide semantically integrated information and can be represented by using relational or XML databases.

Content management system (CMS)

The CMS is the web platform used to manage and publish the information content created by authorised users. A CMS can provide organised workflow, cataloguing and metadata tools and can separate the content from its presentation layout and design by a template builder. Thus, it is important in reducing the workload of managing a large web site, such as the tedium in creating, moving, deleting and organising the contents by many members. In the proposed architecture, facet management tools can be developed within the CMS for adding, deleting and editing the facets and classification of document resources. For rapid implementation a CMS can be chosen from the open source community that can be customised and integrated with the capability of FCS.

Communicator

The communicator is the user's interface to the FCS. The communicator interacts with the user, receives requests from the user and then generates processing for concerned components. In designing the site interface, content structure and navigation should be developed based on the metadata from facet storage. Consideration can be given to show hierarchical metadata from facets containing one or two levels of depth.

Implementation

The prototype FCS is implemented by using a collection of 65 web documents culled from a typical higher education and research institute, the Asian Institute of

Technology (AIT) in Thailand. This academic web site is used not only to demonstrate the suitability of implementing faceted classification in an educational system but also to show its effectiveness with diverse content, given that very little research has focused on developing the information architecture of institutional sites that cover a wide variety of content.

Based on the proposed framework, the system is developed by relational database using python and MySQL integrated within "Plone", an open source Content Management System (CMS). Plone is an application built on the Zope Content Management Framework (CMF). Plone/Zope provides a variety of features, including a workflow that allows many members to add, edit and delete content. The finalised system architecture is provided in Figure 5.

The facets, metadata isolates, and classification of all documents are stored in the MySQL database, which runs externally to the CMS. The objects within Zope and Python script contain methods that provide an interface between the Zope content management tools and the data. The document content is stored in the directory of the file system of Zope. All components of the system interfaces are dynamically generated based on the facets and facet values defined in the database. Query previews (an indicator of the number of expected results) are generated using the SQL "group by" operator to count the number of items that fall into each subcategory.

In a typical scenario of the prototype FCS system, students, visitors and other external users seek information from the site and play the role of unregistered user. To create the collection of contents/resources, a number of authorised students and staff can be the creator and editor of the content and can play the role of members. The system administrator plays the managerial role and will review the content, decide whether to publish and then classify it with the appropriate facet to organise the

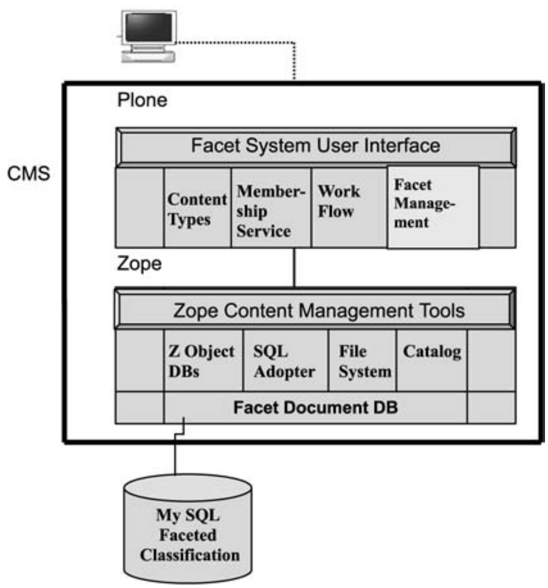


Figure 5.
System architecture

information and present it to the general users. Figure 6 illustrates the roles and processes concerned in the typical scenario in which FCS can work. In this case the administrator acts as a Manager and Template Builder, and the authorised content creators (student, staff) act as Members.

Facet and taxonomy design

The facets and metadata isolates of the prototype system are extracted by analysing the content collection of the institutional sites – specifically, by identifying groups of representative concepts and sub-concepts, which are able to describe the context as well as the user need. Facets are mainly constructed by term aggregation ("Faculty", "Staff" and "Student" can be aggregated in the facet, "People") and by term composition ("By school" and "By degree" are composed in a facet, "Area", resulting in the concept of places). Term composition is distinguished from term aggregation. Composed terms, generated by the combination of single terms, form a new concept, such as "PhD student", whereas terms which are simply aggregated, such as "Staff"

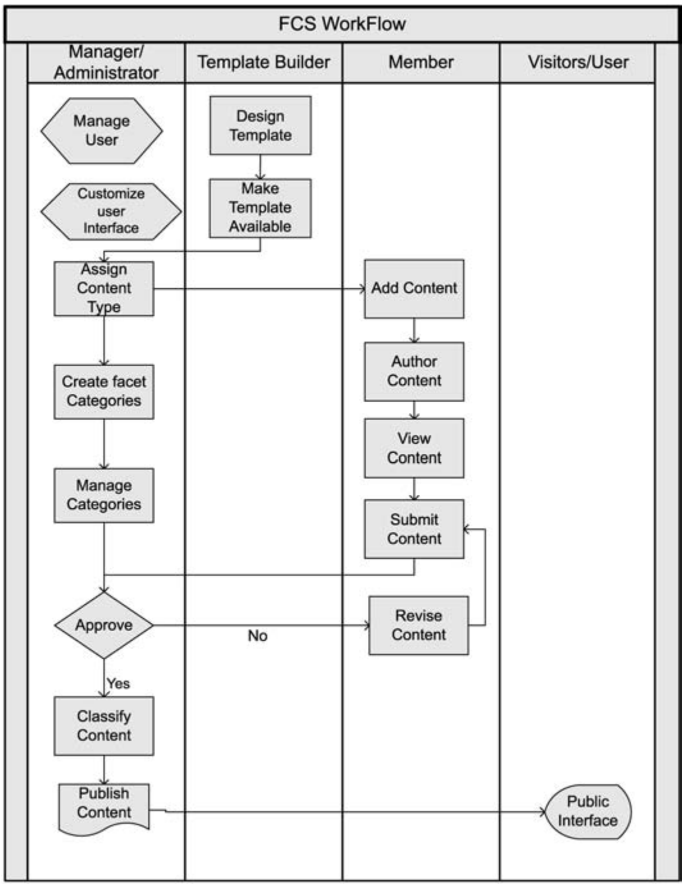


Figure 6.
Workflow in the FCS

and "Student", can co-occur in a document but do not form a concept called "Staff Student" (Priss and Jacob, 1999). However, in creating facets and taxonomy, the following procedures are followed:

- analysing the domain, the content, and the users;
- identifying and aggregating the concepts and sub-concepts covered by aforesaid analysis;
- constructing mutually independent facets by term aggregation and term composition;
- using a controlled vocabulary and taxonomic structure to fill each facet with standard metadata (isolates) that represent the concepts; and
- categorising each document according to the facets.

Thus, for the prototype implementation the following four facet categories are identified to represent all the information objects of institutional systems:

(1) *Purpose* – why the document/information will be used;

(2) *Topic* – "aboutness" of the document/information object, subject addressed;

(3) *People*– person concerned/described by the document; and

(4) *Area* – space or places of the document.

These facets are mutually exclusive because an entity cannot express a topic and people at the same time, though they can be combined with respect to the concept, "institutional document". An institutional document may have at least one topic, purpose of using, indicated people, or an area to be placed. Therefore, a document can be located under any of the four facets.

In the facet isolate relationship, a term is called an isolate/element of a facet if it is an element of the set of terms/metadata in that facet. Further, a facet is called a sub-facet of another facet if it is used as a component in constructing the facet. Figure 7 shows a part of the relationship used in the FCS.

Database design

The FCS database is developed to provide a dynamic facet system structure. It classifies URI of the document resources and links these resources to their facet classification. Moreover, the resources are annotated to describe the content that can be displayed dynamically in user interfaces. The database is constructed by two sets of tables – facet system tables and resource tables – with an aim of providing networked concepts to minimise the retrieval times. The facet system consists of the following tables: facets, synonyms, facet synonyms, and reference. The resource tables are: resources and facet resources. The facets table holds the records of facets and sub-facets/isolates, which uses both the nested set model and adjacency list model to represent a tree structure (as a kind of directed graph) by applying the techniques of SQL tree (Celko, 1996). This tree structure, with left and right values, is very useful to set a parent-child relationship among the facets and isolates and allows quick retrieval of the descendants. Figure 8 shows a part of the system facet tree, and an example of representing the tree is provided in Table I.

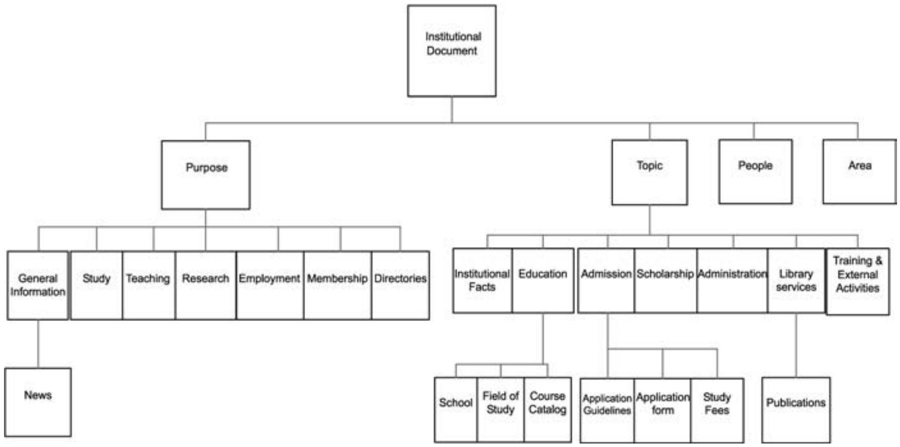


Figure 7.
Example of facet and sub-facet relation hierarchy

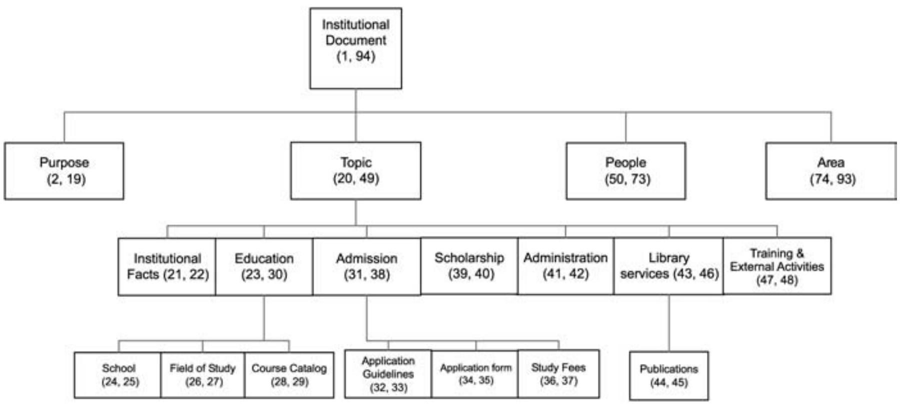


Figure 8.
Facet tree with nested set numbering

Id	Title	Parentid	Left	Right
1	Facet root	0	1	94
2	Purpose	1	2	19
3	Topic	1	20	49
4	People	1	50	73
5	Area	1	74	93
6	Institutional facts	3	21	22
7	Education	3	23	30
8	School	7	24	25
9	Field of study	7	26	27
10	Course Catalog	7	28	29

Table I.
Representation of nested set model in the database

End-user interfaces

The prototype system provides three interfaces for general users to search for institutional information: browser interface, advanced search interface and basic search interface.

Browser interface: The browser interface has a dynamically updated query preview (an indicator of the number of results to expect) that allows all current and prospective users to browse and navigate information through metadata isolates under four categories of facets, as shown in Figure 9. By selecting categorical metadata from a facet users can see a dynamic preview of the result set; for example a user looking for research information may select "Research" under the "Purpose" facet and would then see a preview of the relevant result items in the right side, which are presented by predefined metadata such as titles, content summary, reference to similar pages, file type and page ranking. Determining the need from the summary, users can click on the titles or similar pages to go to their desired pages. Users can also switch to another facet to change their search from the result page. The prototype system further simplifies the browsing by limiting each facet category to only one level of hierarchy. For example the "Faculty" category under the "People" facet has only one level to show regular and visiting faculty, and selecting "Regular Faculty" will narrow the result compared with "Faculty". Finally, the user can see the history of selected facets and metadata in a "breadcrumb" shown at the top of the results page.

Advanced search interface: An advanced search interface where users can combine their searching within the four categories of facets through multiple selection boxes. For example, if users look for research information within a specific school, they can select "Research" from the "Purpose" facet, as well as a school such as SET from the "Area" facet to combine the search. It is noted that SET is a subcategory of "By school". However, users can combine all the four facets or can exclude any of the facet

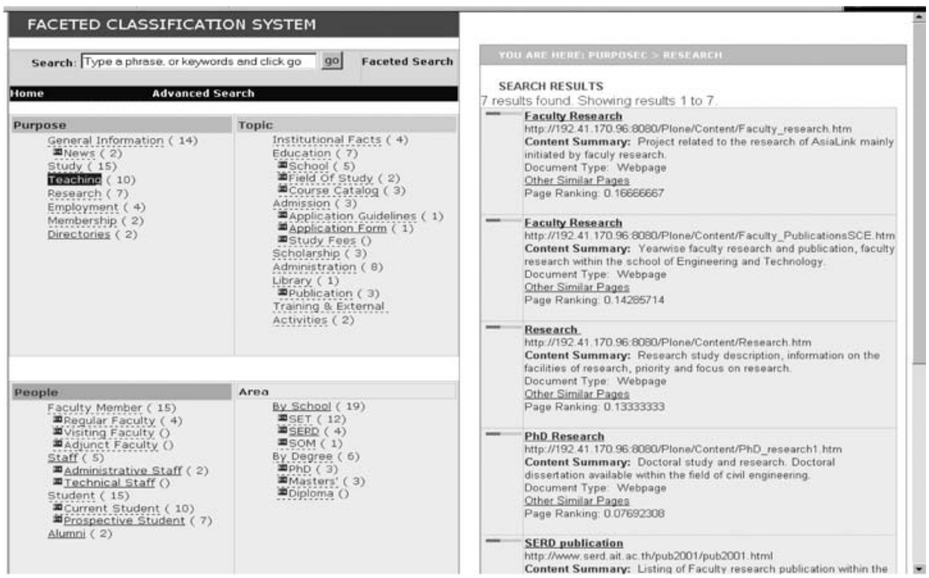


Figure 9.
Browse interface with result preview in the right side

categories to make their searches. The advance search interface with result preview is provided in Figure 10.

Basic search interface: A basic keyword search form is also provided that allows entering any number of keywords or phrases. It supports Boolean queries and presents the results with relevance ranking.

Usability study

The FCS implementation has a number of usability features that should make it more acceptable to users: it will provide more efficient access and searching, more understanding of information content, more convenient navigation and browsing, and overall more search satisfaction and relevant search results. We conducted an empirical study with 19 participants to evaluate the usability and performance issues of the prototype FCS interfaces compared with the current institutional web system, which is referred to as standard system (SS). Regarding the number of participants in usability testing, studies have shown that 80 percent of the usability problem can be identified by four to five participants (Virzi, 1992; Nielsen, 1994) when conducted by real users. But a small sample size may be too small to indicate significant differences between groups; therefore, a true experiment should be conducted with a minimum of 10 to 12 participants (Spyridakis, 1992; Ahmed *et al.*, 2005). However, in our study the participants were selected from the staff and students of AIT, where the focus was placed on computer education and internet experience to group them as expert ($n = 10$)

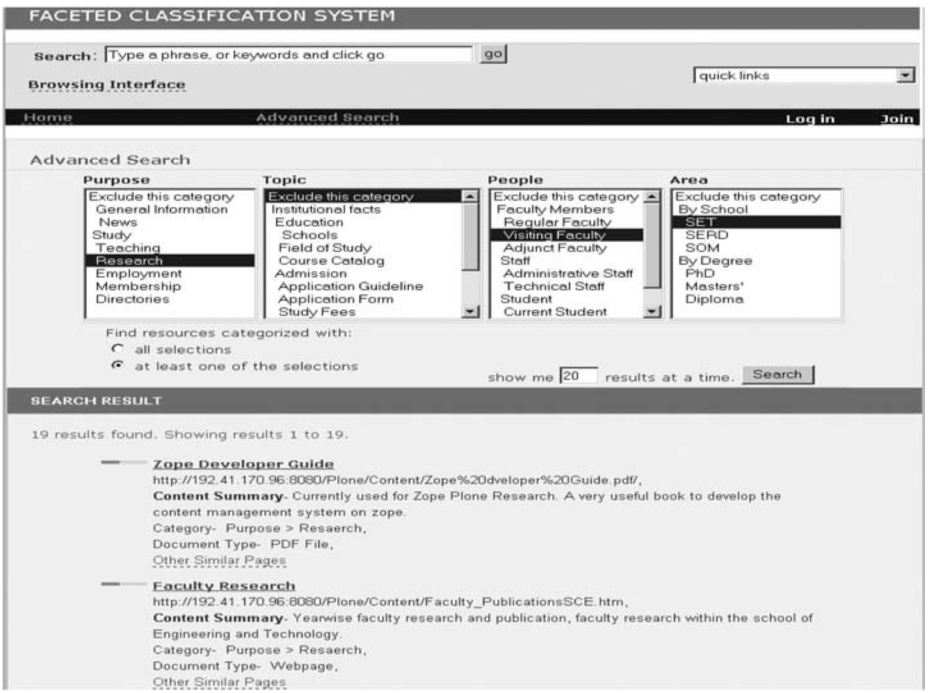


Figure 10.
Advanced search with result preview at the bottom

and non-expert ($n = 9$) users to distinguish the differences in their evaluation result. The participants consisted of five staff, five postgraduate and nine graduate students, among which 12 are male and seven female. Our experiment measured the ability of users to understand and use the multidimensional categories of the FCS, as well as performance. The null hypotheses were developed in accordance with the driven propositions of the study, which are tested through Paired-samples T test and ANOVA with post-hoc analysis. Paired-samples *t*-test was used to compare two scores (FCS vs SS) from the same participants, while ANOVA tested the differences of the result among the subsets of user groups.

The results found statistical validity in some of the areas. For example FCS significantly increases the access to information with an increased success score and satisfaction ($t(18) = 6.530, p < 0.000$). Besides, it significantly increases the understanding of information content ($t = 6.088, p < 0.000$), flexibility in searching ($t = 6.119, p < 0.000$) and relevance in the result ($t = 4.083$, $p < 0.001$).The result also revealed that in general users like the multiple category-based searching. Table II lists the overall satisfaction of users with FCS compared with SS interfaces. The features of FCS found to be the most useful were the capability to switch from one facet to another, the preview of the result set, the ability to combine facets in searching, using content summaries, and breadcrumb (the representation of the path from the home page to the current information).

However, some general users found it difficult to browse under multiple categories and showed less understanding and use of the facets in the FCS interfaces compared with the expert users with more Internet experience and formal computer education.

Table II.
Satisfaction level of the user

					95% Confidence interval for mean	
	User type	*N*	Mean	SD	Lower bound	Upper bound
Understanding facet categories	Nu	10	6.30	1.16	5.47	7.13
	Eu	9	7.22	1.48	6.08	8.36
	Total	19	6.74	1.37	6.08	7.40
FCS browsing interface	Nu	10	6.80	1.69	5.59	8.01
	Eu	9	6.89	1.62	5.65	8.13
	Total	19	6.84	1.61	6.07	7.62
FCS advance interface	Nu	10	7.40	0.70	6.90	7.90
	Eu	9	7.67	1.00	6.90	8.44
	Total	19	7.53	0.84	7.12	7.93
FCS keyword search	Nu	10	5.80	2.04	4.34	7.26
	Eu	9	7.00	1.22	6.06	7.94
	Total	19	6.37	1.77	5.52	7.22
Overall FCS	Nu	10	7.10	0.88	6.47	7.73
	Eu	9	7.44	0.88	6.77	8.12
	Total	19	7.26	0.87	6.84	7.68
Overall SS	Nu	10	5.80	0.79	5.24	6.36
	Eu	9	6.56	0.73	6.00	7.11
	Total	19	6.15	0.83	5.76	6.56

Notes: 1= Very unsatisfied; 5= Average; 9= Highly satisfied; Nu= Non-expert user; Eu= expert user

Thus, some assumptions were not statistically proved, as the results showed no significant differences between FCS and SS in terms of ease of using, browsing, and simplicity of the interfaces. The most likely reason for this is that the understanding of new interfaces with multiple categories takes more time for some users, especially when the participants are familiar with the conventional system and make a comparative evaluation within time constraints in a laboratory environment.

Conclusion and future work

This study addresses an important web-related information search problem: namely, that online users often confront and are confused by imperfect site structure and poorly designed search results. Our approach for the solution is to classify web documents through content-oriented metadata organised under different facets. We have developed a framework and implemented a prototype faceted classification system in the domain of the higher academic information system, which differs from conventional systems in many ways, such as providing multiple taxonomy, query preview of expected result, preview of the result set, annotating the result with metadata and changing the result by switching from one facet to another through a simple point-and-click interaction.

Currently, the system can be used as a thesaurus and a query processor where in the keyword search the given terminology may not match the set of words wanted by users. Our future work includes building up domain ontology for matching thesaurus terms into the search query, and use of the techniques from adaptive user interface research such as a relevance feedback mechanism. The ontology may not only provide semantically integrated information by class, subclass, property and relationship definition within the domain but also provide a structure to mark up the instance data that can be used to search the system through user query.

The outcomes of the study can provide significant grounds for the information retrieval community to improve interface structures for easy access, management and retrieval of web information. Besides, the integration of content management tools with multidimensional taxonomies can be a new instance of a corporate web system that provides easy content creation, organisation and navigation. The developed FCS architecture is also easily applicable to other web systems, such as the ability of commercial sites to adopt the whole architecture and database schema to organise their product information in multiple categories, based on which users may browse or conduct a preference-based search by adding or refining their options through checkboxes or multiple selections. The only task required to adapt to a new domain is to replace the current facets and metadata taxonomies from the database with new kinds of facets and taxonomies derived from the analysis of proposed domain, content and user.

References

Ahmed, S.M.Z., McKnight, C. and Oppenheim, C. (2005), "A study of learning and retention with a web-based IR interface", *Journal of Librarianship and Information Science*, Vol. 37 No. 1, pp. 7-16.

Allen, R.B., Obry, P. and Littman, M. (1993), "An interface for navigating clustered document sets returned by queries", *Proceedings of ACM COOCS: Conference on Organizational Computing Systems, Milpitis, CA, November 1993*, pp. 166-171.

Borgman, C.L. (1996), "Why are online catalogs still hard to use?", *Journal of the American Society for Information Science*, Vol. 47 No. 7, pp. 493-503.

Broughton, V. (2006), "The need for a faceted classification as the basis of all methods of information retrieval", *Aslib Proceedings: New Information Perspectives*, Vol. 58 Nos 1/2, pp. 49-72.

Celko, J. (19965), *A Look on the SQL Tree*, available at: www.dbmsmag.com/9603d06.html (accessed 10 December 2005).

Chen, H. and Dumais, S. (2000), "Bringing order to the web: automatically categorizing search results", *Proceedings of CHI 2000: Conference on Human Factors in Computing Systems, The Hague, Netherlands, April 2000*, pp. 145-152.

Chen, H., Houston, A.L., Sewell, R.R. and Schatz, B.R. (1998), "Internet browsing and searching: user evaluations of category map and concept space techniques", *Journal of the American Society for Information Sciences*, Vol. 49 No. 7, pp. 582-603.

Ellis, D. and Vasconcelos, A. (2000), "The relevance of facet analysis for worldwide web subject organization and searching", *Journal of Internet Cataloging*, Vol. 2 Nos 3/4, pp. 97-114.

English, J., Hearst, M., Sinha, R., Swearingen, K. and Yee, P. (2003), "Examining the usability of web site search", available at: http://bailando.sims.berkeley.edu/papers/epicurious-study.pdf (accessed on 12 January 2006).

Hearst, M.A. (2000), "Next generation web search: setting our sites", *IEEE Data Engineering Bulletin*, Vol. 23 No. 3, pp. 38-48.

Hearst, M.A. and Pedersen, J.O. (1996), "Re-examining the cluster hypothesis: scatter/gather on retrieval results", *Proceedings of the 19th Annual International ACM SIGIR: Conference on Research and Development in Information Retrieval (SIGIR), Zurich, Switzerland, June 1996*, pp. 76-84.

King, M. (2001), "A tangled web", available at: www.etailersdigest.com/resources/Specials/Tangled_Web.htm (accessed on 25 September 2005).

Louie, A.J., Washington, W. and Maddox, E.L. (2003), "Using faceted classification to provide structure for information architecture", available at: http://depts.washington.edu/pettt/presentations/conf_2003/IASummit.pdf (accessed 16 September 2005).

Maarek, Y.S. and Wecker, A.J. (1994), "The librarian's assistant: automatically assembling books into dynamic bookshelves", *Proceedings of Recherche d'Information Assistee par Ordinateur RIAO '94; Intelligent Multimedia Information Retrieval Systems and Management, New York, October 1994*.

Nielsen, J. (1994), "Heuristic evaluation", in Nielsen, J. and Mack, R.L. (Eds), *Usability Inspection Methods*, John Wiley, New York, NY, pp. 25-62.

Pratt, W., Hearst, M. and Fagan, L. (1999), "A knowledge-based approach to organizing retrieved documents", *Proceedings of 16th Annual Conference on Artificial Intelligence (AAAI 99), Orlando, Florida, July 1999*, pp. 80-85.

Priss, U. and Jacob, E. (1999), "Utilizing faceted structures for information systems design", *Proceedings of the 62st Annual Meeting of the American Society for Information Science, Washington DC, November 1999*, pp. 203-212.

Ranganathan, S.R. (1960), *Colon Classification Basic Classification*, 6th ed., Asia Publishing House, New York, NY.

Rao, R. (2002), *Information Flow*, Vol. 4, available at: www.ramanarao.com/informationflow/archive/2002-08.html (accessed 3 October 2005).

Spyridakis, J.H. (1992), "Conducting research in technical communication: the application of true experimental design", *Technical Communication*, Vol. 39 No. 4, pp. 607-24.

Taylor, A.G. (1999), *The Organization of Information*, Libraries Unlimited, Englewood, CO.

Virzi, R.A. (1992), "Refining the test phase of usability evaluation: how many subjects is enough?", *Human Factors*, Vol. 34 No. 4, pp. 457-68.

Yee, K., Swearingen, K., Li, K. and Hearst, M.A. (2003), "Faceted metadata for image search and browsing", *Proceedings of the CHI, Florida, Lauderdale*, available at: http://bailando.sims.berkeley.edu/papers/flamenco-chi03.pdf (accessed 8 February 2006).

Corresponding author
Mohammad Nasir Uddin can be contacted at: nasiruddin26@gmail.com