KO KNOWLEDGE ORGANIZATION

Contents

KNOWLEDGE ORGANIZATION

KO

Contents pages

Abstract: The purpose of this work is to develop an ontology-based framework for developing an information retrieval system to cater to specific queries of users. For creating such an ontology, information was obtained from a wide range of information sources involved with brain tumour study and research. The information thus obtained was compiled and analysed to provide a standard, reliable and relevant information base to aid our proposed system. Facet-based methodology has been used for ontology formalization for quite some time. Ontology formalization involves different steps such as identification of the terminology, analysis, synthesis, standardization and ordering. A vast majority of the ontologies being developed nowadays lack flexibility. This becomes a formidable constraint when it comes to interoperability. We found that a facet-based method provides a distinct guideline for the development of a robust and flexible model concerning the domain of brain tumours. Our attempt has been to bridge library and information science and computer science, which itself involved an experimental approach. It was discovered that a faceted approach is really enduring, as it helps in the achievement of properties like navigation, exploration and faceted browsing. Computer-based brain tumour ontology supports the work of researchers towards gathering information on brain tumour research and allows users across the world to intelligently access new scientific information quickly and efficiently.

Abstract: Creative professionals in advertising, marketing, design and journalism search for images to visually represent a concept for their project. The main purpose of this paper is to present search facets derived from an analysis of documents known as briefs, which are widely used in creative industries as requirement documents describing information needs. The briefs specify the type of image required, such as the content and context of use for the image and represent the topic from which the searcher builds an image query. We take three main sources—user image search behaviour, briefs, and image search engine search facets—to examine the search facets for image searching in order to examine the following research question—are search facet schemes for image search engines sufficient for user needs, or is revision needed? We found that there are three main classes of user search facet, which include business, contextual and image related information. The key argument in the paper is that the facet "keyword/tag" is ambiguous and does not support user needs for more generic descriptions to broaden search or specific descriptions to narrow their search—we suggest that a more detailed search facet scheme would be appropriate.

Abstract: Multilingual controlled vocabularies are rare and often very limited in the choice of languages offered. TIIARA (Taxonomy for Image Indexing and RetrievAl) is a bilingual taxonomy developed for image indexing and retrieval. This controlled vocabulary offers indexers and image searchers innovative and coherent access points for ordinary images. The preliminary steps of the elaboration of the bilingual structure are presented. For its initial development, TIIARA included only two languages, French and English. As a logical follow-up, TIIARA was translated into eight languages—Arabic, Spanish, Brazilian Portuguese, Mandarin Chinese, Italian, German, Hindi and Russian—in order to increase its international scope. This paper briefly describes the different stages of the development of the bilingual structure. The processes used in the translations are subsequently presented, as well as the main difficulties encountered by the translators. Adding more languages in TIIARA constitutes an added value for a controlled vocabulary meant to be used by image searchers, who are often limited by their lack of knowledge of multiple languages.

Abstract: The internet in its Web 2.0 version has given an opportunity among users to be participative and the chance to enhance the existing system, which makes it dynamic and collaborative. The activity of social tagging among researchers to organize the digital resources is an interesting study among information professionals. The one way of organizing the resources for future retrieval through these user-generated terms makes an interesting analysis by comparing them with professionally created controlled vocabularies. Here in this study, an attempt has been made to compare *Library of Congress Subject Headings* (*LCSH*) terms with LibraryThing social tags. In this comparative analysis, the results show that social tags can be used to en-

KO   KNOWLEDGE ORGANIZATION

hance the metadata for information retrieval. But still, the uncontrolled nature of social tags is a concern and creates uncertainty among researchers.

Ridenour, Laura. "Boundary Objects: Measuring Gaps and Overlap Between Research Areas." *Knowledge Organization* 43 no. 1: 44-55. 35 references.

Abstract: The aim of this paper is to develop methodology to determine conceptual overlap between research areas. It investigates patterns of terminology usage in scientific abstracts as boundary objects between research specialties. Research specialties were determined by high-level classifications assigned by Thomson Reuters in their Essential Science Indicators file, which provided a strictly hierarchical classification of journals into 22 categories. Results from the query "network theory" were downloaded from the Web of Science. From this file, two top-level groups, economics and social sciences, were selected and topically analyzed to provide a baseline of similarity on which to run an informetric analysis. The Places & Spaces Map of Science (Klavans and Boyack 2007) was used to determine the proximity of disciplines to one another in order to select the two disciplines use in the analysis. Groups analyzed share common theories and goals; however, groups used different language to describe their research. It was found that 61% of term words were shared between the two groups.

AlQenaei, Zainab M. and Monarchi, David E. "The Use of Learning Techniques to Analyze the Results of a Manual Classification System." *Knowledge Organization* 43 no. 1: 56-63. 19 references.

Abstract: Classification is the process of assigning objects to pre-defined classes based on observations or characteristics of those objects, and there are many approaches to performing this task. The overall objective of this study is to demonstrate the use of two learning techniques to analyze the results of a manual classification system. Our sample consisted of 1,026 documents, from the ACM Computing Classification System, classified by their authors as belonging to one of the groups of the classification system: "H.3 Information Storage and Retrieval." A singular value decomposition of the documents' weighted term-frequency matrix was used to represent each document in a 50-dimensional vector space. The analysis of the representation using both supervised (decision tree) and unsupervised (clustering) techniques suggests that two pairs of the ACM classes are closely related to each other in the vector space. Class 1 (Content Analysis and Indexing) is closely related to Class 3 (Information Search and Retrieval), and Class 4 (Systems and Software) is closely related to Class 5 (Online Information Services). Further analysis was performed to test the diffusion of the words in the two classes using both cosine and Euclidean distance.