KO   KNOWLEDGE ORGANIZATION

Contents

KNOWLEDGE ORGANIZATION   KO

Contents pages

Abstract: Traditional knowledge organization approaches struggle to make large user-generated collections navigable, especially when these collections are quickly growing, in which currency is of particular concern, for which professional classification design is too costly. Many of these collections use folksonomies for labelling and organization as a low-cost but flawed knowledge organization approach. While several computational approaches offer ways to ameliorate the worst flaws of folksonomies, some user-generated collections have implemented a human judgment-centered alternative to produce structured folksonomies. An analysis of three such implementations reveals design differences within the space. This approach, termed "curated folksonomy," presents a new object of study for knowledge organization and represents one answer to the tension between scalability and the value of human judgment.

Abstract: The purpose of this paper is to determine semantic types for tags in social tagging systems. In social tagging systems, the determination of the semantic type of tags plays an important role in tag classification, increasing the semantic information of tags and establishing mapping relations between tagged resources and a normed ontology. The research reported in this paper constructs the semantic type library that is needed based on the Unified Medical Language System (UMLS) and FrameNet and determines the semantic type of selected tags that have been pretreated via direct matching using the Semantic Navigator tool, the Semantic Type Word Sense Disambiguation (STWSD) tools in UMLS, and artificial matching. And finally, we verify the feasibility of the determination of semantic type for tags by empirical analysis.

Abstract: Notations are systems of symbols that can be combined according to syntactical rules to represent meanings in a specialized domain. In knowledge organization, they are systems of numerals, letters and punctuation marks associated to a concept that mechanically produce helpful sequences of them for arranging books on shelves, browsing subjects in directories and displaying items in catalogues. Most bibliographic classification systems, like *Dewey Decimal Classification*, use a positional notation allowing for expression of increasingly specific subjects by additional digits. However, some notations like that of Bliss *Bibliographic Classification* are purely ordinal and do not reflect the hierarchical degree of a subject. Notations can also be expressive of the syntactical structure of compound subjects (common auxiliaries, facets etc.) in various ways. In the digital media, notation can be recorded and managed in databases and exploited to provide appropriate search and display functionalities.

Abstract: It is argued that data should be defined as information on properties of units of analysis. Epistemologically it is important to establish that what is considered data by somebody need not be data for somebody else. This article considers the nature of data and "big data" and the relation between data, information, knowledge and documents. It is common for all these concepts that they are about phenomena produced in specific contexts for specific purposes and may be represented in documents, including as representations in databases. In that process, they are taken out of their original contexts and put into new ones and thereby data loses some or all their meaning due to the principle of semantic holism. Some of this lost meaning should be reestablished in the databases and the representations of data/documents cannot be understood as a neutral activity, but as an activity supporting the overall goal implicit in establishing the database. To utilize (big) data (as it is the case with utilizing information, knowledge and documents) demands first of all the identification of the potentials of these data for relevant purposes. The most fruitful theoretical frame for knowledge organization and data science is the social epistemology suggested by Shera (1951). One important aspect about big data is that they are often unintentional traces we leave during all kinds of activities. Their potential to inform somebody about something is therefore less direct compared to data that have been produced intentionally as, for example, scientific databases.

Abstract: Briefly describes the new features of the recently released 22nd edition of the *Sears List of Subject Headings*. Introduces its new publisher Grey House Publishing, Amenia, N.Y., which recently acquired comprehensive rights for the Sears along with

its aboriginal publisher, the H W Wilson from the Ebsco Information Services. The system is available both in print and online versions. Names a few new subject headings in areas like science, technology, engineering and medicine (STEM). In this edition, there are a total of 1,600 new headings making it a total of 12,000+ preferred headings meant for subject access in small and medium sized libraries. This unprecedented increase of about 1,600 headings is mostly due the complete incorporation of the Canadian Sears last published independently in 2006. Also critically examines inconsistencies in a few headings. Concludes to say the new edition in resplendent, hard binding maintains its stellar reputation of a handy list of general subject headings both for applications and a teaching resource.