Vol. 34 (2007) N°4

Knowledge Organization

International Journal

Devoted to Concept Theory, Classification, Indexing, and Knowledge Representation

CONTENT

ERGON-VERLAG

ISSN 0943-7444 Knowl.Org.

KNOWLEDGE ORGANIZATION KO

Official Quarterly Journal of the International Society for Knowledge Organization ISSN 0943 – 7444

International Journal devoted to Concept Theory, Classification, Indexing and Knowledge Representation

Contents

Knowl. Org. 34(2007)No.4

KO KNOWLEDGE ORGANIZATION

Official Quarterly Journal of the International Society for Knowledge Organization ISSN 0943 – 7444

International Journal devoted to Concept Theory, Classification, Indexing and Knowledge Representation

Contents pages

KNOWLEDGE ORGANIZATION KO

Official Quarterly Journal of the International Society for Knowledge Organization ISSN 0943 – 7444

International Journal devoted to Concept Theory, Classification, Indexing and Knowledge Representation

of digital documents. The most frequent approach to automated classification is machine learning. It, however, requires training documents and performs well on new documents only if these are similar enough to the former. We explore a string-matching algorithm based on a controlled vocabulary, which does not require training documents–instead it reuses the intellectual work put into creating the controlled vocabulary. Terms from the Engineering Information thesaurus and classification scheme were matched against title and abstract of engineering papers from the Compendex database. Simple string-matching was enhanced by several methods such as term weighting schemes and cut-offs, exclusion of certain terms, and enrichment of the controlled vocabulary with automatically extracted terms. The best results are 76% recall when the controlled vocabulary is enriched with new terms, and 79% precision when certain terms are excluded. Precision of individual classes is up to 98%. These results are comparable to state-of-the-art machine-learning algorithms.

Knowl. Org. 34(2007)No.4

KO KNOWLEDGE ORGANIZATION

Official Quarterly Journal of the International Society for Knowledge Organization ISSN 0943 – 7444

International Journal devoted to Concept Theory, Classification, Indexing and Knowledge Representation

Founded under the title *International Classification* in 1974 by Dr. Ingetraut Dahlberg, the founding president of ISKO. Dr. Dahlberg served as the journal's editor from 1974 to 1997, and as its publisher (Indeks Verlag of Frankfurt) from 1981 to 1997.

KNOWLEDGE ORGANIZATION KO

Official Quarterly Journal of the International Society for Knowledge Organization ISSN 0943 – 7444

International Journal devoted to Concept Theory, Classification, Indexing and Knowledge Representation

Publisher

ERGON-Verlag, Grombühlstr. 7, GER-97080 Würzburg
Phone: +49 (931) 280084; FAX +49 (931) 282872
E-mail: service@ergon-verlag.de; http://www.ergon-verlag.de

Editor-in-chief (Editorial office)

Dr. Richard P. SMIRAGLIA (Editor-in-Chief), Palmer School of Library and Information Science, Long Island University, 720 Northern Blvd., Brookville NY 11548 USA.
Email: Richard.Smiraglia@liu.edu

Instructions for Authors

Manuscripts should be submitted electronically (in Word, WordPerfect, or RTF format) in English only to the editor-in-chief and should be accompanied by an indicative abstract of 100 or 200 words. Submissions via email are preferred; submissions will also be accepted via post provided that submissions are accompanied by a 3.5" diskette encoded in Word, WordPerfect, or RTF format.

A separate title page should include the article title and the author's name, postal address, and E-mail address, if available. Only the title of the article should appear on the first page of the text. To protect anonymity, the author's name *should not* appear on the manuscript, and all references in the body of the text and in footnotes that might identify the author to the reviewer should be removed and cited on a separate page. Articles that do not conform to these specifications will be returned to authors.

Criteria for acceptance will be appropriateness to the field of the journal (see Scope and Aims), taking into account the merit of the contents and presentation. The manuscript should be concise and should conform as much as possible to professional standards of English usage and grammar. Manuscripts are received with the understanding that they have not been previously published, are not being submitted for publication elsewhere, and that if the work received official sponsorship, it has been duly released for publication. Submissions are refereed, and authors will usually be notified within 6 to 10 weeks. Unless specifically requested, manuscripts and illustrations will not be returned.

The text should be structured by numbered subheadings. It should contain an *Introduction*, giving an overview and stating the purpose, a *main body*, describing in sufficient detail the materials or methods used and the results or systems developed, and a *conclusion* or *summary*.

Reference citations within the text should have the following form: (author year). For example, (Jones 1990). Specific page numbers are optional, but preferred when applicable, e.g. (Jones 1990, 100). A citation with two authors would read (Jones & Smith, 1990); three or more authors would be: (Jones et al., 1990). When the author is mentioned in the text, only the date and optional page number should appear in parenthesis – e.g. According to Jones (1990), …

References should be listed alphabetically by author at the end of the article. Author names should be given as found in the sources (not abbreviated). Journal titles should not be abbreviated. Multiple citations to works by the same author should be listed chronologically and should each include the author's name. Articles appearing in the same year should have the following format: "Jones 2005a, Jones 2005b, etc." Issue numbers are given only when a journal volume is not through-paginated.

Examples:

Dahlberg, Ingetraut. 1978. A referent-oriented, analytical concept theory for INTERCONCEPT. *International classification* 5: 142-51.

Howarth, Lynne C. 2003. Designing a common namespace for searching metadata-enabled knowledge repositories: an international perspective. *Cataloging & classification quarterly* 37n1/2: 173-85.

Pogorelec, Andrej and Šauperl, Alenka. 2006. The alternative model of classification of belles-lettres in libraries. *Knowledge organization* 33: 204-14.

Schallier, Wouter. 2004. On the razor's edge: between local and overall needs in knowledge organization. In McIlwaine, Ia C. ed., *Knowledge organization and the global information society: Proceedings of the Eighth International ISKO Conference 13-16 July 2004 London, UK*. Advances in knowledge organization 9. Würzburg: Ergon Verlag, pp. 269-74.

Smiraglia, Richard P. 2001. *The nature of 'a work': implications for the organization of knowledge.* Lanham, Md.: Scarecrow.

Smiraglia, Richard P. 2005. Instantiation: Toward a theory. In Vaughan, Liwen, ed. *Data, information, and knowledge in a networked world; Annual conference of the Canadian Association for Information Science … London, Ontario, June 2-4 2005.* Available http://www.cais-acsi.ca/2005proceedings.htm.

Footnotes are not permitted; all narration should be included in the text of the article.

Illustrations should be kept to a necessary minimum and should be submitted electronically when possible. Photographs (including color and half-tone) should be scanned with a minimum resolution of 600 dpi and saved as .tif files (Tagged Image File Format preferred). Tables and figures should be embedded within the document or, alternatively, saved as separate files with clear instructions indicating their placement in the text. Tables should contain a number and title at the top, and all columns and rows should have headings. All illustrations should be cited in the text as Figure 1, Figure 2, etc. or Table 1, Table 2, etc. Illustrations submitted in hard copy only should be marked to indicate their placement in the text.

Upon acceptance of a manuscript for publication, authors must provide a wallet-size photo and a one-paragraph biographical sketch. The photograph should be scanned with a minimum resolution of 600 dpi and saved as a .tif file (Tagged Image File Format).

Advertising

Responsible for advertising: Dr. H.-J. Dietrich, ERGON-Verlag, Grombühlstr. 7, 97080 Würzburg (Germany).

KO is published quarterly by ERGON-Verlag.
The price is € 115,00/ann. including airmail delivery.

Knowl. Org. 34(2007)No.4

KO KNOWLEDGE ORGANIZATION

Official Quarterly Journal of the International Society for Knowledge Organization ISSN 0943 – 7444

International Journal devoted to Concept Theory, Classification, Indexing and Knowledge Representation

Scope

The more scientific data is generated in the impetuous present times, the more ordering energy needs to be expended to control these data in a retrievable fashion. With the abundance of knowledge now available the questions of new solutions to the ordering problem and thus of improved classification systems, methods and procedures have acquired unforeseen significance. For many years now they have been the focus of interest of information scientists the world over.

Until recently, the special literature relevant to classification was published in piecemeal fashion, scattered over the numerous technical journals serving the experts of the various fields such as:

philosophy and science of science
science policy and science organization
mathematics, statistics and computer science
library and information science
archivistics and museology
journalism and communication science
industrial products and commodity science
terminology, lexicography and linguistics

Beginning in 1974, KNOWLEDGE ORGANIZATION (formerly INTERNATIONAL CLASSIFICATION) has been serving as a common platform for the discussion of both theoretical background questions and practical application problems in many areas of concern. In each issue experts from many countries comment on questions of an adequate structuring and construction of ordering systems and on the problems of their use in opening the information contents of new literature, of data collections and survey, of tabular works and of other objects of scientific interest. Their contributions have been concerned with

(1) clarifying the theoretical foundations (general ordering theory/science, theoretical bases of classification, data analysis and reduction)
(2) describing practical operations connected with indexing/classification, as well as applications of classification systems and thesauri, manual and machine indexing
(3) tracing the history of classification knowledge and methodology
(4) discussing questions of education and training in classification
(5) concerning themselves with the problems of terminology in general and with respect to special fields.

Aims

Thus, KNOWLEDGE ORGANIZATION is a forum for all those interested in the organization of knowledge on a universal or a domain-specific scale, using concept-analytical or concept-synthetical approaches, as well as quantitative and qualitative methodologies. KNOWLEDGE ORGANIZATION also addresses the intellectual and automatic compilation and use of classification systems and thesauri in all fields of knowledge, with special attention being given to the problems of terminology.

KNOWLEDGE ORGANIZATION publishes original articles, reports on conferences and similar communications, as well as book reviews, letters to the editor, and an extensive annotated bibliography of recent classification and indexing literature.

KNOWLEDGE ORGANIZATION should therefore be available at every university and research library of every country, at every information center, at colleges and schools of library and information science, in the hands of everybody interested in the fields mentioned above and thus also at every office for updating information on any topic related to the problems of order in our information-flooded times.

KNOWLEDGE ORGANIZATION was founded in 1973 by an international group of scholars with a consulting board of editors representing the world's regions, the special classification fields, and the subject areas involved. From 1974-1980 it was published by K.G. Saur Verlag, München. Back issues of 1978-1992 are available from ERGON-Verlag, too.

As of 1989, KNOWLEDGE ORGANIZATION has become the official organ of the INTERNATIONAL SOCIETY FOR KNOWLEDGE ORGANIZATION (ISKO) and is included for every ISKO-member, personal or institutional in the membership fee (US $ 55/US $ 110).

Rates: From 2006 on for 4 issues/ann. (including indexes) € 115,00 (forwarding costs included). Membership rates see above.

ERGON-Verlag, Grombühlstr. 7, GER-97080 Würzburg;
Phone: +49 (931) 280084; FAX +49 (931) 282872; E-mail: service@ergon-verlag.de; http://www.ergon-verlag.de

The contents of this journal are indexed and abstracted in ***Referativnyi Zhurnal Informatika*** and in the following online databases: ***Information Science Abstracts, INSPEC, Library and Information Science Abstracts (LISA), Library Literature, PASCAL, Sociological Abstracts, and Web Science & Social Sciences Citation Index.***