Subject indexing of articles from LIDA proceedings using ASIS&T Thesaurus

Ivana Tanfara, MA in Information Sciences
University of Zadar, University library, Email: itanfara@gmail.com

Drahomira Gavranović, PhD student
University of Zadar, Department of Information Sciences, Email: dgavrano@unizd.hr

Abstract

The subject of this poster is subject indexing of papers from the LIDA proceedings using ASIS&T thesaurus (3rd edition). Purpose of the poster is to explore the ability of applying indexing terms from the ASIS&T thesaurus on chosen corpus of papers from the LIDA conference proceedings in order to improve subject indexing and retrieval. The research questions are: 1.) to test sufficiency of current version of ASIS&T thesaurus for subject indexing of LIDA conference papers; 2) to explore if indexing terms (descriptors) from ASIS&T thesaurus can express subject coverage of the papers; 3) to propose ASIS&T thesaurus to be updated with new terms (descriptors) concerning “digital” and “digital libraries” and descriptors around those concepts. The beginning hypothesis is that the ASIS&T thesaurus has not full coverage of subjects represented in LIDA conference proceedings due to changes in the field of library and information sciences.

Key words: LIDA proceedings papers, ASIS&T thesaurus, subject analysis, subject indexing
Introduction

Thesaurus is defined as a dictionary of key words and a mean of control of the terminology that is used to translate natural language into a language for indexing. Some thesauri are designed for a particular profession. Thesaurus consists of a list of descriptors, links between descriptors and secondary meanings of words and the explanations. ASIS&T thesaurus of Information Science is published by the American Society for Information Science and Technology. Thesaurus covers the field of information science and library science and related fields such as linguistics, computer science, and behavioural and cognitive science if they are associated with information science and librarianship. The thesaurus also partially covers the area of education and economics. It is primarily used as a tool for indexing and searching within the field of information science and librarianship. It is designed for students and researchers as a guideline for terminology specific area. The thesaurus has a hierarchical facet structure which allows making changes in the structure as well as the introduction of new terms. Therefore is chosen for this short research study for testing appropriateness of its terms for indexing papers covering topics in the field of digital libraries.

Methodology:

During this research five steps were taken.

1. **First step.** Selection of key words from the sample of papers from the LIDA conference proceedings (a. using key words created by authors of the papers and b. where needed, adding appropriate additional key words for describing subjects of the paper more accurate).
2. **Second step.** Selected key words were matched with descriptors available in the ASIS&T thesaurus to determine whether the thesaurus has sufficient number of descriptors for subject indexing e.g. in the growing field of digital libraries.
3. **Third step.** Comparison of indexing terms (descriptors) with key words used for describing the content of papers with special attention to “digital” and “digital libraries” and descriptors around those concepts.
4. **Fourth step.** Creating and proposing new potential descriptors and their scope notes for missing concepts from key words extracted from papers or created by authors of the poster.
5. **Fifth step.** Providing list of new proposed concepts for introducing them in next edition of ASIS&T thesaurus to improve subject indexing in the area of digital libraries.

Findings and results (examples)

Example 1: DIGITAL HERITAGE


**Key words:** online digital heritage, born digital heritage, digital preservation

**ASIS&T thesaurus descriptors:** digital libraries, digitization

**Suggestions:** born digital heritage, digital preservation, digitization


**Key words:** cultural heritage, cultural heritage digitization, cultural heritage digitization education, library and information science

**ASIS&T thesaurus descriptors:** digital libraries, digitization

**Suggestions:** born digital heritage, digital preservation, digital object preservation, digital collections preservation, digital libraries education
Discussion

Analysis of the research showed that in many cases there is no corresponding/matching term in the thesaurus for most of the selected key words/concepts described in papers. Only few descriptors in the thesaurus related to terms “digital” and “digital libraries” which clearly showed that existing descriptors are not sufficiently hierarchically developed to meet the needs of the subject indexing of the papers from the LIDA conference proceedings. Also, further analysis has shown that there are many outdated definitions within the scope notes. In many examples, relations between existing descriptors have not been elaborated detailed enough.

The faceted structure of ASIS&T thesaurus provides sufficient flexibility which allows the implementation of new terms in the thesaurus as well as the making necessary changes, without breaking the structure of the original thesaurus.

Conclusions

There is a need for introduction of new terms in thesaurus, especially concerning concepts “digital” and “digital libraries”. Outdated definitions inside scope notes should be replaced with modern definitions and relations between descriptors in thesaurus must be linked with similar terms and/or replaced with new terms.

There is significant number of examples where new concepts should be introduced in the ASIS&T thesaurus (3rd edition) to better accommodate changes in the literature from the field of library and information sciences.

Reference & Citations


Curriculum Vitae

Ivana Tanfara graduated from the University of Zadar in 2012. Currently on internship in the University library at University of Zadar.

Drahomira Gavranović is working as a research and teaching assistant at University of Zadar, Department of Information Sciences since 2007. She is doing her PhD in information sciences in the area of classification. Her research interests are information organization, classification, subject indexing and ontologies in semantic web.