

ANNA PET'OVÁ – SLOVENSKÁ NARODNÁ KNIŽNICA

SUBJECT HEADINGS IN THE SLOVAK NATIONAL LIBRARY – PAST, PRESENT AND FUTURE

Introduction

The subject and classifying processing as a result of the content analysis of a document is the important instrument to satisfy search questions of library users. So, it is the important instrument of information retrieval and easy access to it. Our aim in the near future is an attempt to create relatively uniform access to information on the Internet. We will make the attempt to integrate information language of the subject headings in the Slovak National Library into the present process of retrieval. We suppose this subject authority file will contribute to that situation as in many other countries, too.

New information technologies and their influence on information retrieval

Information retrieval in the past thirty years underwent huge changes which are reflection of the development in the department of information technologies (we have high level computers, better computer programmes, better quality of communication instruments).

Present days are *stigmatized by the Internet* which is known for different information display and allows access to information in different ways. When we look at the online catalogues of libraries – these offer their users not only peruse in bibliographical databases, but also access to full text documents. So, nowadays we are witnesses of the inception of a very heterogeneous milieu from the view of source display. The most serious problem that is needed to be solved is sources description and access in the Internet milieu. The quality of sources of a description provides the quality of their access. It is a paradox that in this

heterogeneous milieu the only unifying element of retrieval are data of a subject description which are very problematic and still under discussion. To create uniform access to information is not easy (and is it possible?). The present days offer a possible solution – to standardize concepts and produce files of subject and name (bibliographical) authorities. The controlled dictionaries are now coming at information scene.

Searching on the Internet

The professional literature states that the most frequently used instrument allowing effective retrieval of information in the Internet belongs to subject directories which are nowadays called information portals or gates, composed of hierarchical ordered subject categories and subcategories which make the first level of retrieval. The second level of retrieval is made of searching and retrieval of media which have to find out information with the help of keywords or combination of them with the help of Boolean algebra, phrase etc.

Traditional information-retrieval systems

As I mentioned above, the libraries use the network of the Internet and display their collections via their websides and library portals. The content of library collections can be described by some of the traditional information-retrieval systems. Among them we include the system of subject headings, subject and polysubject thesaurus or classified systems for example UDC.

We have to state that the milieu of the Internet from the view of displaying and searching for information and also a system of instruments (today's OPAC is a method of keyword searching) appeal to us to evaluate the information-retrieval system used in the Slovak National Library. Therefore we have a task to integrate this traditional retrieval system into the present process of searching for information. We suppose that we should not surrender it. Only, as it seems, what we have to do is to 'impose' or 'modernize' the information language of subject headings.

The Czechoslovak standard – The Creation of the Subject Headings – ČSN 010188 was valid in Slovakia until the end of the year 2002. It defined the rules of making subject headings. The main principle of previously mentioned standard were free created subject headings strictly only on the base of substantive inversion of the first case of plural. The only exception were expressions which specified unique kinds, types or subjects and nominatives occur only in singular or if the usage of plural was not so usual.

The subject heading is composed of input – a bear element and other parts – subheadings. The input element and subheadings create a structural subject chain or subject sentence which is ordered due to given before syntactic rules (as

it was in that standard). This standard had a tendency to decompose the concept (for example: the term *psychology of a child* – decomposed into two individual, separate concepts: a child – psychology). These are basic characteristics of creation subject heading due to the Czechoslovak standard.

This standard caused a retreat from using the broad subject sentences which characterized the period of time until its reception. It was a very important tool of indexation of the documents in time of the manual, card catalogue when the emphasis on a word had a great importance. The subject headings with long chains were effective in the usual surrounding of card catalogues and pressed bibliographical indexes.

- Their maintenance is very difficult
- It is not easy to standardize
- The application syntax is complicated

The search for OPACs done in the Slovak libraries reassured us and also used practically in the Slovak National Library where the complicated, detailed syntactic rules of the Czechoslovak standard are not kept. Although this standard has a key to solve homonymy and synonymy of subject headings, nor this – in the form of see – reference and see – also – reference – is not applied in practice, or only exemptously. What is kept, is the decomposition of terms, mainly in the second case and substantive inversion.

The information language of subject headings created in such a way is not appropriate in the electronical milieu and we think that it is necessary to transform subject headings to a modern information language of subject headings, if we want to go to meet our users. In the same time, the integration of this information language should be done in the present process of searching on the Internet.

Transformation and integration of subject headings in the Internet conditions

Transformation and integration of subject headings in the Internet conditions should happen in three resorts according to the traditional retrieval system of subject headings which is used by the Slovak National Library. The aim of this shaped retrieval system is the more effective search strategy in the new technical conditions. These resorts are: the shape and choice of lexical units, simplyfying of applicated syntax and unification of concepts.

1. Shape and choice of lexical units. The role of the information retrieval language is different in full text systems and in a bibliographical databasis. The choice of lexical units needs to be done in such a way that it should allow a perspective direct link on the full text documents. There are two conditions – to keep the natural word order and compactivity of terms. These two demands

are clearly supported by systematic instruments of today's OPACs – keyword searching.

- In done practice – free created subject headings will withdraw standardizing of individual concepts in the natural language. The lexical units in the natural language (natural wordorder) in individual parts of a subject heading will be given in the first case of plural. In the singular they will be given only when there is just this necessity (in the case of plurale tantum) or if the plural is not usual and its usage changes the meaning of concepts. Usually, there are abstract concepts.

- Let's abandon decomposition of multiwords terms. In the most possible scale, we will respect the term as a whole (theory of sets, geography of sites, psychology of a child etc.)

- In the equivalent word connections we will use a conjunction 'and' (theory of state and law, state and church).

- We will introduce using of prefix word connections: work with reader, jumps into the water.

- A spelling (orthographical) form is proceeded due to a valid standard of Slovak language, in case of spelling variants, a perspective form has a priority.

- The transliteration (metagraphy) uses the valid standards.

- The expressions of foreign languages can be used, if there is no the Slovak equivalent accepted by the Slovak ortographical (spelling) standard.

In the case of synonym to determine the priority we usually use the Slovak synonymous dictionary, 1995. Though there are also exceptions to the view of perspectivity of given concepts in a particular branch. Into this category of words belong to synonymous expressions (rhetoric – oratory – platform mastery) and spelling variants (in Slovak manažment – management or application – see appliqué) which are related to words of foreign languages or archaisms.

It is demanded that the content of lexical units – preferred terms will be approaching in the most possible way in which content of terms is used in the professional literature and communication.

In the universal sorting systems (the Slovak National Library owns, proceeds and makes an accessible universal collection and the sorting system has to reflect the universality of its collection). We often come into the contact with inaccuracy of a given term. Therefore in this case it comes to so called *mixing* of different terminologies of different branches and synonymous (accurate) terms anchored in individual branches becomes a homonym (e.g. – analysis).

A homonymous expression of a caption in a subject heading is differed by a qualifier / relator in an unshortened version in parentheses: aerobic (physical exercises), aerobic (dance). We can use a phrasal caption made with the help of an adjective for example: mathematical analysis, chemical analysis, physical analysis. We can also use an explanatory complement, for example: soul – religious aspects, soul – philosophical aspects.

2. Simplifying of an applicated syntax. The subject heading will have a simple or even minimal structure. It demands the proper electronic milieu according to possibility of application in different MARC formats (dimension, edition) and because long chains of subject headings are not convertible in conversion in one version, esp. from one system to another. The main condition is the user of our OPACs, who, according to the research of the National Library in Czech Republic, searches mainly for words and while so called browsing – takes in perceive only the first two or three positions.

Individual ‘traditional’ parts of the thematic subject heading which serve as precisions (accuracies) of the input element have their own fixed order (place):

- a topical subheading
- a geographical subheading
- a chronological subheading
- a subheading of a form

Such order of the subheadings is recommended by the methodical materials for making of authoritative records: *LCSH. Principles and Application*, p. 86, 1995; *Principles Underlying Subject Heading Languages*, p. 117, 1999; UNIMARC/ Authority, p. 51, 1996 and finally also ČSN 010188 I/9.

The subheadings or saing in another way, thematic accuracies in the subject authority file become an inseparable (precoordinate) part of the subject heading or they can be free connected with the input element of the heading (postcoordinate). From the view of a processor and user this way is more welcoming. The research, realized by the National Library of Czech Republic in the year 1997, showed us that in the field of information-retrieval languages it has the absolute primacy in searching for using of isolated lexical units (it means free created keywords). Precoordinated subject headings with the developed structure are in the absolute minority. In that situation it seems to be more suitable to use authorization to individual segments of the subject heading in a field of the applicated syntax of subject heading and choose principle of postcoordination.

The subject heading will gain rather the vertical level than horizontal one. And here arises a new rule – it is better to have higher occurrence of the subject heading, more the input element than to create one long chain. Finally, UNIMARC alone supports this tendency in the way that subfield – \$a – is in a boundary of one occurrence not repeatable.

For example :

Digitalization as a Method of Preservation /Hartmut Weber, Marianne Dorr.
– Amsterdam : European Commission on Preservation and Access, 1997

The horizontal level of the subject heading used nowadays:

6060 \$\$a collections archival \$\$x digitalization (method of collectiones' protection) \$\$x materials ECPA

Tendencies of LCSH – absolute *verticality* of the subject heading. It has an influence also on the subheadings which are shown in separate occurrence of individual fields and not as a part of chain structure of one field.

Vertical display

6060 \$\$a archival collections
 6060\$\$a digitalization
 6060\$\$a protection of collections
 608 \$\$a materials ECPA

Every piece of important information from the view of the content of the document should be given in the input element (in the caption of the subject heading). For allocation of the subject headings runs the principle of so called summarized access for the description of the monographies in institutions with universal character. One subject heading shall cover twenty percent of the document content. So, we can see that one heading will have maximum of five subject headings. But it is not a strict policy. In processing, the whole information value should be taken in account which is offered by the record. Many data are involved in the name description and therefore it is not necessary to create duplications with the help of the subject headings.

3. Unification. Free created subject headings will be continuously replaced by unified – standardized shapes of the subject authorities. The standardization is the necessary instrument of managing the heterogeneous milieu of the Internet (changing the bibliographical indexes with full texts) and also the instrument of searching in a union catalogue.

The supportive instrument for making subject headings is the unified subject authority file . Nowadays in the Slovak National Library in the system milieu of ALEPH 300, we create the local database – the Slovak authority file – a part of subject headings of the Slovak National Library in which concepts are given and written in format UNIMARC / Authorities.

The subject headings and their accuracies (subheadings) are still the subject of standardization. The headings are connected with English equivalents and notation in UDC MRF. The subject headings authority file has approximately 2500 concepts. Among lexical units in the subject authority file we state also semantic relations – equivalentious, hierarchic relations of subordination and relation of association, preferred and not preferred concepts.

The important part of the records are definitions which determine expansion of the concept in the universal system.

We use the following standard in making the subject heading authority file:

- Guidelines for Subject and Reference Entries. IFLA-UBCIM Publication
- UNIMARC/Authorities – format defines the structure of the record of the subject authority

- LCSH (in Principle)

The subcommittee on Metadata and Subject Analysis of the Association of Library Collections and Technical Services of the American Library Association recommends in the year 1999 the adaptation for a general collections – Library of Congress Subject Headings

- UDC MRF, the second Slovak middle issue

The sources of authoritative forms of the subject authorities for us are the Slovak translation of UDC MRF and the basis of the subject authorities in the National Library in Prague.

After setting down the technical conditions which are closely tied to the implementation of ALEPH 500 we plan to incorporate into the subject authority file also geographical names and subheading forms of document. For the first time we had concentrated while doing the subject headings of the Slovak National Library, on removing synonymia.

After the rise of draft material The Principles of making authoritative subject headings and after the official acceptance of LCSH as the third standard in dimension of CASLIN (Czech and Slovak Library Information Network) we had standardized almost six hundred subject headings which are included into the local database of books and are easy accessible via www.snk.sk under a bequest reference of subject processing. The file shown in the Internet will be updated every quarter of the year. In fact, it is a provisional (temporary) solution which will serve us until the implementation of ALEPH 500. This version of ALEPH offers standard technological solution of the authority files.

The experimental sample of subject authorities has been published since October 2001. During this period of time we have not recognized any negative response and this allows us to close the discussion on the subject processing. In May of this year we have prepared a course named *The Change of subject processing under influence of the new information technologies*. The aim of this course was to make known the libraries in public in the field of the changes of subject processing. In the technical milieu of ALEPH 500 we plan from the beginning of the year 2003 to implement a new way of subject processing.

In addition, I would like to notice the fact that it will be ideal if every specialized library in Slovakia creates each own branch – the specialized subject authority file and make it accessible via the Internet. The Slovak National Library according to universality of the collection will standardize concepts which belong to its database.

The cooperation with other libraries on the creation of the National Authority File is possible on the basis of consulting. The Slovak National Library play a role in standardization of the subject headings of its own database which reflects the universal collection of the Slovak National Library.

Conclusions

We can state that subject and classifying processing undergoes with influence of the new information technologies which changes radically. It is necessary to accept them if we do not want to destroy access for our user into the information world.

Slovakia accepted AACR2 and we are preparing the shift from UNIMARC to MARC 21. The principles of LCSH, as the internationally accepted standard for subject processing, will be our next standard document which will serve us as a background.

This shows that controlled access to information is possible to build up by suitable integration of traditional information-retrieval systems. But we think that transformation of these traditional instruments in the mentioned resorts is necessary.

Bibliography

Balíková M., *Problematika věcného pořádku informací a jejich zpřístupnění*, "Národní knihovna", 12 (2001) č. 3, s. 175-186.

Balíková M., *Změna koncepce věcného zpracování v Národní knihovně ČR*, 1998

ČSN 010188. Tvorba předmětových hesel. – Praha: Úřad pro normalizaci a měření, 1982, s. 33.

CASLIN 2001: Document Description and Access: A New Challenge.-Praha: Knihovna Akademie věd ČR, 2001. – 157 s.

Guidelines for Subject Authority and Reference Entries. Munchen : K. G. Saur, 1993

Chan, Lois Mai. Subject Categorization of Web Resources. Pracovní materiály. – In: CASLIN 2001. – Roč. 8. (2001), s. 47-54.

Chan, Lois Mai. LCSH. Principles and Application. Englewood, Colorado : Libraries Unlimited, INC., 1995

Makulová, Soňa: Vyhľadavanie informácií v roku 2000 alebo od kvantity ku kvalite, 2000

UNIMARC/Autority. Martin, Matica slovenská – Slovenská národná knižnica 1996.

Library of Congress Subject Headings, 1992