Conspectus categorization scheme based on UDC classification used in subject gateways in Czechia

Marie Balíková
National Library of the Czech Republic

ABSTRACT: This paper explores the role of classification and other terminological subject systems in subject gateways showing the Czech experience. The potential of combining a top level classification such as Conspectus with the application of further classified sets of subject access points using UDC is explained. Some examples are given of special subject gateways developed under a general gateway, all interlinked and providing similar interfaces and functionalities. The potential of a common strategy of subject access tools is emphasized, to further improve subject retrieval across the different gateways in a consistent and transparent manner.

KEYWORDS: Subject gateways; Conspectus; UDC; Czech Republic

Introduction

Subject gateways are usually characterised in the professional literature as specific internet services offering qualitatively evaluated (intellectually selected and assessed) information resources and systematically supporting information retrieval and research in a given field, or in a given semantic domain. They focus on a certain group of users, offering adequate research approaches and methods. The information resources made available are organised into categories based on various criteria (typological, formal, thematic), the thematic criteria usually being considered as the most important. This is understandable as users, finding themselves overwhelmed by an abundance of miscellaneous documents, more often select access approaches to information resources based on their content characteristics, i.e. subject data.

For this reason, when creating subject gateways there is a growing need to pay great attention to the tools used for subject access, at least until effective automated systems to ensure both the selection of quality information resources and their retrieval are fully available. In terms of access, we mean primarily information-retrieval systems based upon semantic technologies, systems that use cluster methods for searching in text documents etc., which, as we all wish, will in the future be able to ensure quality access to the heterogeneous information resources offered by subject gateways.

Subject access tools in subject gateways

The main issue in providing integrated subject access across collections is the lack of standardized approaches and guidelines for selecting the best option for resource discovery. For instance, there is a wide variety of subject tools to choose from: universal or specialized classification schemes; controlled vocabularies (subject heading lists, thesauri) as well as non-controlled subject terms. In addition, the amount of data supporting subject access in resource collections may differ and, apart from controlled subject terms (subject headings, descriptors), bibliographic and other kinds of metadata records may include tables of contents, abstracts and summaries, annotations, reviews. Last but not least, subject gateways may include full text collections, thus allowing not only metadata but also full-text searching. It is, however, widely accepted as a good practice that all the resources referenced in a subject gateway should be classified according to

the same classification scheme and their intellectual content analysed and indexed according to the same controlled vocabulary. This, for instance requires, that resources that are primary indexed by one system are also indexed by or mapped to the searching vocabulary chosen as a gateway standard. This strategy increases the rate of success in retrieval by means of parallel search as well as by browsing the heterogeneous information resources accessible through the gateway, using both systems.

Classification systems in subject gateways

Classification systems, universal as well as subject specific, are often criticised for their structural rigidity and for the fact that they use notation for class representation with which users may not be familiar. Nevertheless, these schemes are very useful for the organisation of information resources and their access, because they allow the systematic sorting of the information resources into corresponding thematic categories, thus supporting the browsing method and making it easier for less experienced users to become oriented in the tools offered by the subject gateway.

The advantages of universal classification schemes such as Universal Decimal Classification (UDC) or Dewey Decimal Classification (DDC), for instance, are many. They cover all fields of knowledge, can be applied universally and their hierarchical structure allows easy implementation of subject browsing. This structure also enables the narrowing of the scope of a search through query contextualisation and easy selection of thematic categories. The numerical and language independent notation used in classifications facilitates interoperability by allowing multilingual access points to be attached to the scheme.

Obvious disadvantage of decimal classifications, however, is the fact that the entire universe of knowledge is forced under general top ten classes and some fairly broad disciplines, therefore subjects are positioned on the third or even fourth level from the top. In many cases users may not know in which of top classes to look for their subject of interest which quite complicates subject browsing.

The Conspectus scheme

Conspectus is an international scheme that provides a common framework for collection assessment and description of strengths, introduced initially in the 1980s to serve as an infrastructure for coordination among research libraries. Conspectus method is nowadays used not only to support coordination of collection development but also to improve access based on content characteristics of holdings (Bushing, 2001). Classification is one of the components of the Conspectus resource categorisation scheme which is why Conspectus was found particularly suitable for the creation and development of the subject-oriented universal gateway (UIG – Uniform Information Gateway) as well as other specialist subject information gateways (KIV /Library and Information Science/, MUS, TECH, ART) in the Czech Republic.

The categorizing scheme involving Conspectus comprises three hierarchical levels:
- 24 basic groups, the so-called Subject Categories
- 584 subordinate subcategories, the so-called Conspectus groups/categories linked to the UDC classmarks
- thematic terms of the subject authority file (Czenas) linked to the UDC classmarks

The highest hierarchical level of this scheme – 24 subject categories along with further related subcategories – serves in the Uniform Information Gateway as subject crossroads. We may

assume that 24 specialised subject gateways will gradually developed from these subject cross-roads under the auspices of the Uniform Information Gateway, through further hierarchical levels of the Conspectus scheme that will be implemented to a greater or lesser extent.

Figure 1 – **Main subject categories of Conspectus at the UIG**

Figure 2 - **Example of UDC classmarks contained in one of the top Conspectus categories**

Controlled vocabularies in subject gateways

The recommendation that a uniform standardised indexing system – of subject headings, or descriptors – be used within subject gateways or groups of subject gateways has already been mentioned. The importance of these systems for parallel searching and browsing has also been noted, emphasizing the significant role of files of controlled subject terms (authority files, subject heading lists and thesauri) in subject gateways.

Examples of subject gateways in Czechia

The Uniform Information Gateway (UIG) provides simultaneous searching in different Czech and foreign resources (library catalogues, union catalogues, full text databases, etc.) through a single user interface. Resources available in UIG are either freely accessible or licensed (the access is possible from authorized IP addresses only).

Figure 3 – **The UIG interface (http://www.jib.cz)**
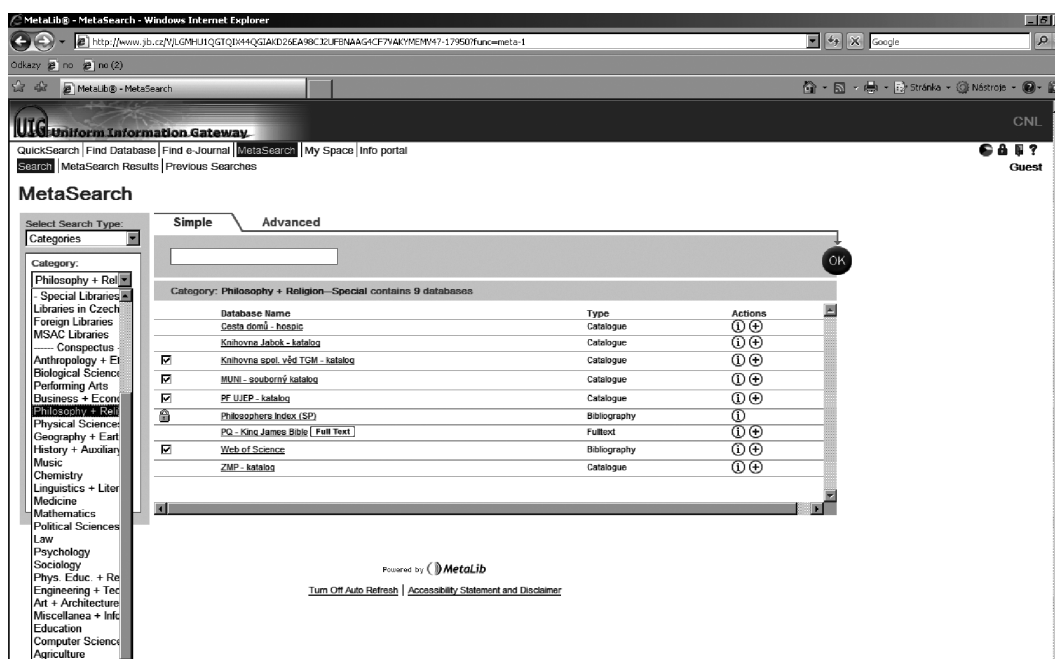
The KIV Subject Gateway is part of the network of subject gateways existing under the umbrella of UIG – Uniform Information Gateway. UIG users and users of other UIG related subject gateways, are provided with a uniform environment and search tools that are well known to them. The KIV Subject Gateway is, above all, intended to meet information needs of librarians, information specialists, teachers, educators and students of Library and Information Science (LIS). KIV is part of the services of the Librarianship Institute of the National Library of the Czech Republic and is virtually connected with the Librarianship Institute Portal called Information for Libraries.

Figure 4 – **The KIV interface (http://kiv.jib.cz/kiv-search/metalib)**
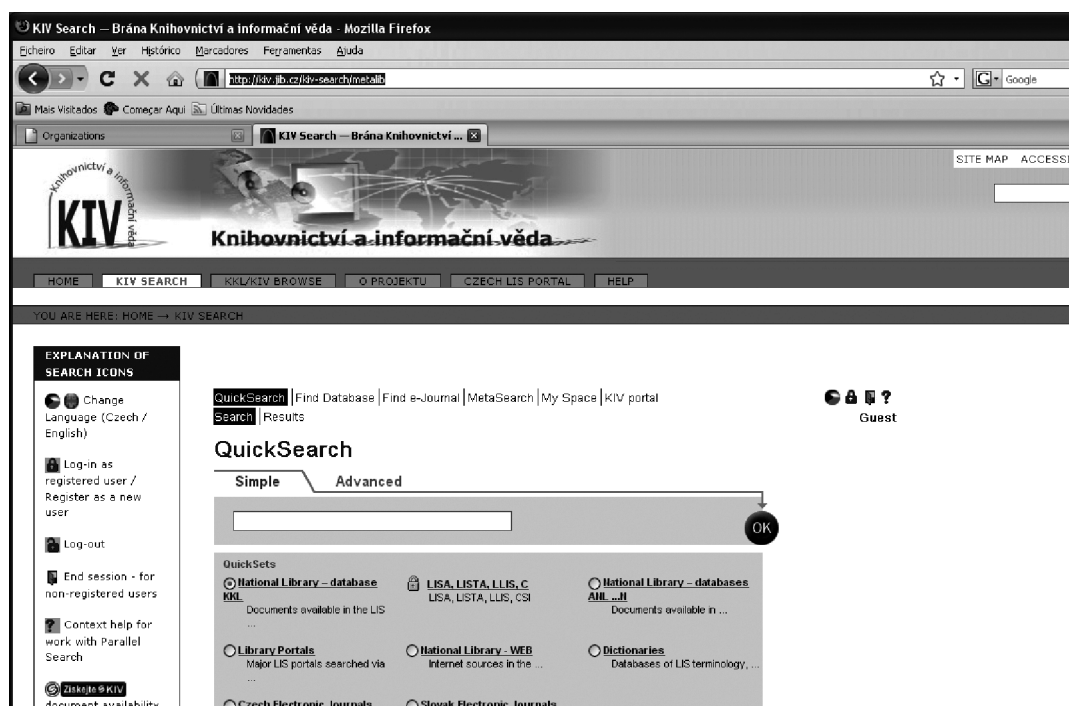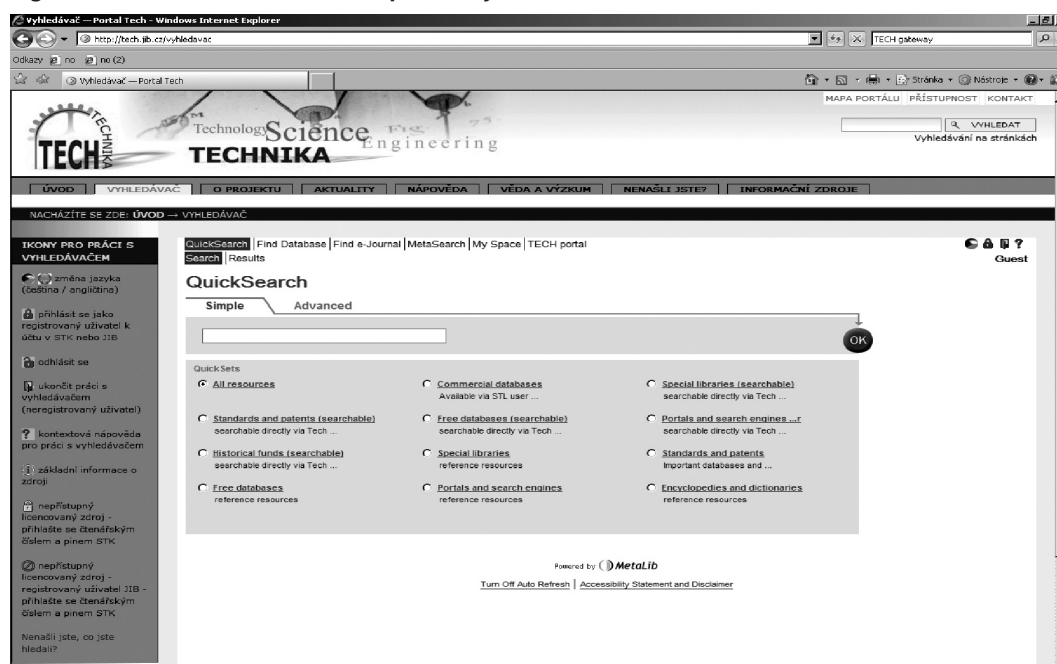

Musica Subject Gateway

The Musica Subject Gateway (MUS) provides integrated access to information resources in the fields of music and musicology. MUS is also part of the wider network of subject gateways covered by the Unified Information Gateway (UIG), thus being structured in a similar way and offering users an environment, tools and search options familiar to them.

Figure 5 – **The MUSICA interface (http://mus.jib.cz/vyhledavac)**
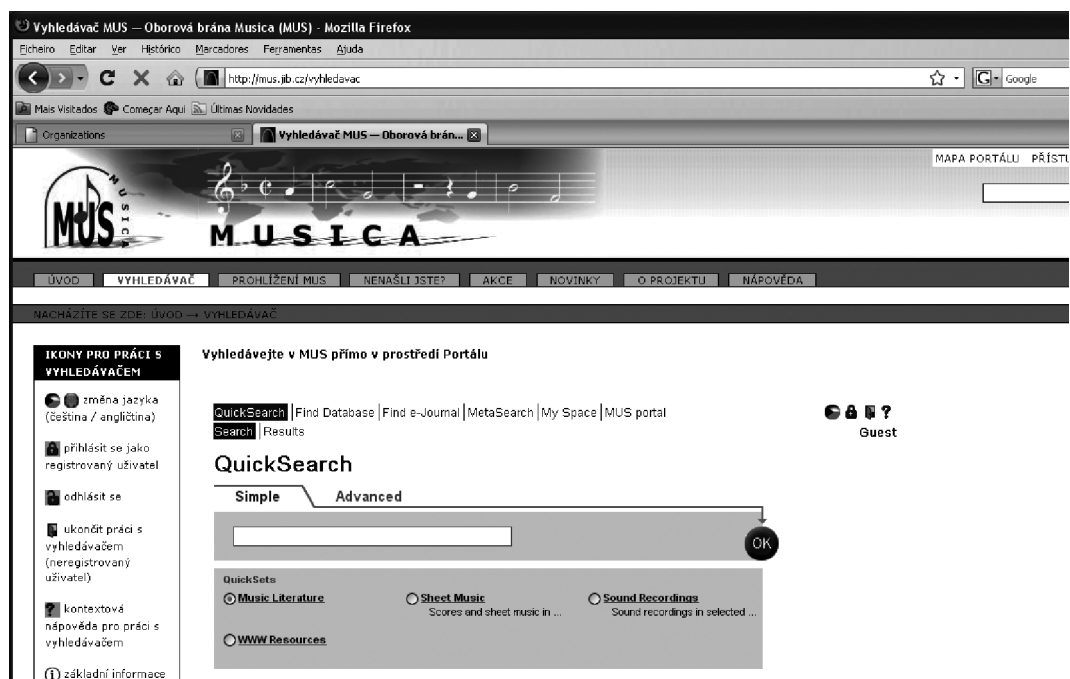
TECH Subject Gateway

The TECH Subject Gateway provides a single search environment for various remote Czech and foreign information resources (library catalogues, databases, subject portals) in the fields of science, technology, and engineering. The classification applied in this database was originally used in the STM Portal, the predecessor of TECH.

Figure 6 – **The TECH interface (http://tech.jib.cz/metalib)**

The subject gateway ART provides users with a single interface to simultaneously search in different Czech and foreign resources (library catalogues, union catalogues, full text databases etc.) from the fields of art, architecture and related subjects. Resources available in ART include both free of charge and licensed materials. Free of charge resources are available to every user, no matter where he/she is using these resources from or if he/she is registered in the library. Licensed resources are available to users registered in the library owning or having roght to access these resources, through computers validated by IP addresses. and at the same time the users have to work on computers with admissible IP addresses.

Figure 8 – **The ART interface (http://art.jib.cz/metalib)**

Conclusion

The purpose of subject gateways created within the UIG portal is to allow users to browse various Czech as well as foreign resources (library catalogues, union catalogues, full-text databases etc.) in parallel, from the respective semantic domains by means of a single search interface and from a single entry point. The KIV and MUS subject gateways allow also a cross-gateway browsing. As indicated in this article, it is, however, possible to further improve subject retrieval by introducing a uniform standardised system of subject access – classification schemes, controlled vocabularies – applied in the subject gateways with the same strategy, thus contributing to a more effective access to the information resources.

Reference

Bushing, M. C. (2001) *A short history of Conspectus developments*. Paper presented at CASLIN 2001, the 8th international seminar of the Czech and Slovak Library Information Network. Available at: http://klement.nkp.cz/Caslin/caslin01/sbornik/conspectus.html.