Usability and the user interfaces of classical information retrieval languages

Abstract: This paper examines some traditional information searching methods and their role in Hungarian OPACs. What challenges are there in the digital and online environment? How do users work with them and do they give users satisfactory results? What kinds of techniques are users employing? In this paper I examine the user interfaces of UDC, thesauri, subject headings etc. in the Hungarian library. The key question of the paper is whether a universal system or local solutions is the best approach for searching in the digital environment.

1. Introduction
The possibilities of integrated systems mean not only automated processes within libraries, but also shared catalogues linking different library systems and extending resources considerably. For users, shared catalogues are the realization of the distributed library.

1.1 When integrated library systems appeared, these questions arose:
- How do earlier methods like classification and retrieval systems apply in new information environments?
- How can users avoid the confusion arising from more user interfaces in the OPAC environment?
- How can maximum satisfaction be obtained from the management of knowledge in organizations?
- Do we keep any or all parts of earlier information retrieval systems or abandon them?
- Different search techniques (conceptual, object-based, browser-based) apply to various levels in the same record. Is it worth separating these different techniques into different fields of the record? Or would it be better to compare them and establish “new” information retrieval language dictionaries from the separate segments?
- On the list above is one apparently unreal question: Do we keep any or all parts of earlier information retrieval systems or abandon them?

It would seem we must keep them, because as things stand we cannot manage information effectively. There are three possibilities:

- to compare the outcomes of different search techniques and establish “new” information retrieval language dictionaries from the separate segments of concepts.
- to transform and reconstruct existing systems according to current needs.
- to change information seeking technology and devise new types of search engines. For instance one could combine the UDC codes (or any hierarchically structured, universal system) with some new technological solution. The Totalzoom technology is only one possibility, because it would be able to map and spatially display hierarchy, tables, codes, and common and special auxiliaries. This method not only would manage the UDC codes in OPACs, but it also could integrate other structured databases, especially hierarchically structured ones. Naturally we can use other methods that are able to visualize information systems. All possibilities should be given a trial.
2. Why the UDC?

I have studied only the UDC in different OPACs and Internet databases, although any similar structured, hierarchical and universal system would be suitable. The use of classical classification methods is a strong tradition in Hungary. One of the most widely used systems has been (and still is) the UDC.

*What advantages has the UDC?*

- Universal system
- Meaningful notation
- Clarity and transparency
- Rich network of relationships
- Well-defined categories
- Ability to describe special and general concepts with free movement between the different levels
- Efficient retrieval, relevant hits
- A long-time tradition, found widely in many libraries - In consequence of Szabó Ervin’s activity, librarians knew and used the UDC very early, and most libraries use it today.
- The concept system is our common cultural heritage and value
- Significant potential
- Standardization
- Well-developed hierarchies, able to visualize information and conceptualize it independently within its structure [Hajdu Barát, 2004 A]

Table 1: structure of Ural-Altaic language group in the UDC  [Egyetemes 2002, p.20]
There is marked interest in the UDC’s potential to assist growing numbers of Internet users. The UDC can play a role of integration in knowledge organization. Thus the answer to the earlier “unreal question” is “Definitely, yes!” We should keep the UDC. However, this answer brings with it some other questions:

- Are there any methods that can search hierarchies while changing levels easily?
- Will the UDC codes become more user-friendly?
- Can we utilize the powerful structure of UDC in OPACs or other electronic environments?
- Should UDC codes become non-terminal scores, and can that structure show the way to retrieval? [Hajdu Barát, 2004 B, 175]

Minimum expectations are:

- Users should navigate easily and unequivocally in permanently variable circumstances.
- Not only librarians but users should be enabled to work and search with UDC codes.
- A user-friendly and user-oriented system is needed.
- The expertise, craft, knowledge, and practice of librarians, professionals and scientists should remain important in the UDC system and UDC MRF.
- User satisfaction is a general expectation and top-most priority. [Hajdu Barát, 2004 B, 175]

3. **Visual Imagery + visualization = usability**

User interfaces are “communicational surfaces” or channels between human information researchers and information retrieval systems. Relevant and irrelevant, clear and nonclear, sufficient and adequate …– these are the worlds to which information-seekers are accustomed, particularly on the Internet. When users approach a known or unknown information system they often feel a muddy and fuzzy understanding of the system’s basic operation. They may be satisfied without knowing advanced functionality like archiving their results. They cannot see the whole cake, but rather only one small piece.

Visual imagery plays an important mental or intellectual role, quite like information (or data) processing, memory, learning, abstract thinking, and linguistic comprehension. Visual perception is a complex process. Visualization begins with the sensation of physical stimuli, but after that it becomes quite individualized. Perception depends, for example, on experience, knowledge, cognition, or one’s system of symbols. The process is an explicit, multilevel and symbolic work of the mind.

People are easily attracted to images and visual information. Pictures, graphics, menus, icons, buttons, graphs… etc. can help users understand, navigate, and query information systems. Information-retrieval systems equipped with computer graphics provide more accessible interfaces.

Hereinafter I focus on the Hungarian user interface and the role of UDC in the various integrated library systems (ILS) in use in Hungary. The role for visualization with UDC takes many forms.
3.1 Simple Levels

3.1.1 Searching UDC codes is impossible

Users can query title, author, place of publication, publisher, subjects…etc. UDC code is not searchable, but hits show UDC codes in the record. This is the case with the TINLIB system in use at Pázmány Péter Catholic University.

3.1.2 Searching UDC codes is possible

In the system of the National Széchényi Library, AMICUS software is in use, and users can search for concrete UDC codes, usually by exact, complex and high standards. The elements and fields searched appear highlighted in red (Table 2). Software vendors have not incorporated all levels of the classification hierarchy into the integrated library systems. For instance, if we are looking for 51 Mathematics, the hits show only the records with the 51 codes accurately. Hit lists exclude 510 Fundamental and general considerations of mathematics, 511 Number theory, 512 Algebra, 514 Geometry …519.1 Graph theory, 519.878 Search theory. This is unsatisfactory and in opposition to the philosophy of UDC, because hierarchy disappears from this solution.
There are some OPACs that use UDC codes with the complete hierarchy to a maximum length (3-4 numbers, 1-2 auxiliaries).

3.2 Translation for subjects or index forms

3.2.1 Simple translation

In some databases and catalogues, UDC codes are translated into searchable subject terms. Although this is very convenient for users, the terms lose every advantage of UDC classification.

3.2.2 Structured Translation

Corvina software is used in the National and University Library of Debrecen (DEENK). In its “Subject Category System” there are many graphically structured sites and sub-sites with UDC concepts in natural language form. Users can search several databases and several types of documents together. They can see their hits before clicking the links to full descriptions of records. This solution helps users to refine their searches by narrowing or broadening them. This is a clearly useful capability given subjects that are not highly esoteric.
3.3 Lead the users to continue their searching with links

In this system, users navigate subject relationships via links and interactive forms. Searchers can combine different subject terms, classification codes, descriptors and bibliographic elements. They have the flexibility to refine strategy or revise their search completely as they go.

3.3.1 National Document Supply System (ODR)

This system uses relatively simple, searchable UDC codes. The search history that is displayed following the actual hits helps guide searchers toward more relevant results. Unfortunately not every field contains searchable links, including the UDC codes. Searching UDC codes results in translations to subject terms, which can then be fed back into the search to refine it. One can combine keywords and subjects in the same search, for example Classification - books and Classification, Universal Decimal. Locations for hits are displayed as well.

Table /: Database of ODR

| Table 8: Hit in database of ODR |

3.3.2 University Library of University of Szeged (SZTE)

With Corvina software, UDC codes are searchable here, too. The codes are very relevant, complex and well-done, as this institute has a great tradition in the field of classification and has specialized in using the UDC system in catalogues. One drawback is that refining the search using UDC codes is not available. There are two parallel screens, the first column with hits and the second with the full record in different formats. UDC codes participate in only the MARC and XML forms and are not in every type of record. There are links for refining the search but the UDC codes are not searchable at this step or after, although subjects, authors, titles etc. are.
3.3.3 Library of the Hungarian Academy of Science (MTA)

One of the best applications of UDC codes in an electronic environment in Hungary uses the Aleph system. High-standard, searchable UDC codes are used. One can refine the search with the help of hits and the flexibility to combine different elements of descriptions. One can browse and search with UDC codes as well as with classification terms. The searched parts of records are visible and lead the searcher to further steps. Naturally, locations for hits are displayed.

3.4 One solution outside Hungary - Catalogus Openbare Bibliotheken Antwerpen
http://bibliotheek.antwerpen.be/MIDA/

This system helpfully guides users to narrow or broaden their concepts and topics. Subject hierarchy is apparent and formatted attractively. The structure (classes, subclasses and subdivisions) is clear and classical. This solution is expressive, keeps the UDC tradition, but depends upon the users’ patience and/or knowledge. One feature is simultaneous searching in separate databases.
4. Conclusion

The UDC is an artificial information-retrieval language. Users who do not know the semantics of classification codes encounter difficulties, and, earlier, reference librarians were the chief aides to users and visitors in information seeking. Today searchers also use catalogues and databases unmediated via the Internet; therefore, the databases themselves should help users in information retrieval. In the extreme, databases might take over the tasks of reference work by exploiting the strengths of visualization (structure, methods, different search engines etc.).

Most solutions involve OPACS. There is a wide variety of ways of using UDC codes in combination with subject terms and other elements of descriptions. We should study these varieties and discover which methods, especially visualization-related, will make OPACS more user-friendly and effective. This paper raises only the classical possibilities. However, the relationship between usability and visualisation is fundamental. From among all the varieties of visualisation methods, librarians and information specialists should intensively seek the best and most effective.

References: