



World Library and Information Congress: 69th IFLA General Conference and Council

1-9 August 2003, Berlin

Code Number: 054-E
Meeting: 157. Bibliographic Control - **Workshop**
Simultaneous Interpretation: -

The relationship between subject gateways and national bibliographies in international context

Bohdana Stoklasová

Deputy for Library Management
National Library of the Czech Republic
Prague, Czech Republic

Marie Balíková

Subject Cataloguing Department Head
National Library of the Czech Republic
Prague, Czech Republic

Ludmila Celbová

Remote Electronic Resources Head
National Library of the Czech Republic
Prague, Czech Republic

ABSTRACT

The paper examines the relationship between subject gateways and national bibliographies together with general principles of universal bibliographic control in the broader context of the need for integration of heterogeneous information resources. The examples from the Czech Republic's experience are used to illustrate general problems connected with integration of heterogeneous resources from different countries covering different subjects. Main obstacles for information resources integration are described. The paper concludes with recommendations for improvement of bibliographic control.

Introduction

The current IFLA Bibliography Section Strategic Plan <http://www.ifla.org/VII/s12/annual/sp12.htm> strongly urges the development of guidelines, with examples and references, to help national

bibliographic agencies (NBAs) start or improve bibliographic services. This paper is aimed at provoking discussion resulting in some guidelines for both subject gateways and national bibliographies in international context.

We will use examples from our own country to illustrate general questions and problems. The Czech Republic is a country that has had to change and improve its national bibliography dramatically during the last decade (after the 1989 revolution resulting in freedom and opportunity for data sharing and exchange). Where the guidelines were missing and/or unclear, we were not able to avoid blind alleys and we lost much time and resources, both human and financial. Nowadays, our national bibliography is comparable with those of countries that served as examples and references for us. Standing at the starting line in subject gateways building, we would like to avoid blind alleys, to get answers for our questions and share them with others as guidelines or probable trends.

The Czech National Bibliography: brief characteristics and history

The Czech National Bibliography (CNB) has been published systematically since 1922. This allowed for the relatively easy retrospective conversion of all the records representing our national output published during the 20th century. All the resulting MARC records are available in the CNB

http://aip.nkp.cz/tornado/webtor.cgi?MainPage=%2E%2E/cnb_en/cnben_main%2Ehtm

and the National Library database

http://sigma.nkp.cz:4505/F/4GVLVLI3BEUJ8YSF5THI2PU2R17SUMGBFBA6ADJFLYSMRRKF36-04203?func=file&file_name=find-b&CON_LNG=ENG

and also in WorldCat.

The National Library of the Czech Republic (NL) has found it relatively easy to define coverage for the purposes of the National Bibliography: we speak one language, Czech (the majority of the records occur in this language) and use one script, Latin. The territory of the nation has changed over time, but it is quite feasible to define and cover it for bibliographic purposes. When it comes to standards, however, we see a completely different and far more difficult situation. During the fifty years of totalitarian rule no free data exchange was possible in our country. We had almost no practical experience with international standards until the revolution in 1989. To be able to co-operate at that time, we had to translate the standards and understand not only their language but also their concept, which was often new for us. Czech cataloguers have had a hard time since 1989. They have had to learn three formats (Czech Exchange Format in 1989, UNIMARC in 1994 and MARC 21 just now, with migration scheduled for 2004); new cataloguing rules (AACR2 translated and introduced in 1994); new subject headings based on the Library of Congress Subject Headings (the CNB records have also their English equivalents, highly appreciated by foreign users and cataloguers downloading our records mainly from WorldCat); and also the new version of the UDC (the new Master Reference File – MRF) and the top level of the DDC (concordance DDC-UDC table for subject gateways and the WLN Conspectus application). We had to introduce all these new standards in a very short time compared to other countries if we wanted to be able to share both bibliographic and authority records worldwide. The new concept of national authorities for both names and subjects also required considerable time and human resources. In the meantime, our printed national output doubled (about 15,000 titles have been published annually in our country during the last few years) and remote electronic resources requiring not only bibliographic control but also transformation of our library services appeared. Like other libraries, we had to find the answer for the difficult question of how to handle all these new duties and the increasing number of documents including remote electronic materials without increasing our budget and with ever-decreasing library staff.

The Czech National Bibliography and remote electronic resources

With these problems in mind, we attended the International Conference on National Bibliographic Services (ICNBS) in Copenhagen in 1998, and afterwards deeply analysed all the recommendations

and evaluated our adherence to them. When discussing the recommendations concerning the coverage of the National Bibliography, especially the *Recommendation no. 4*:

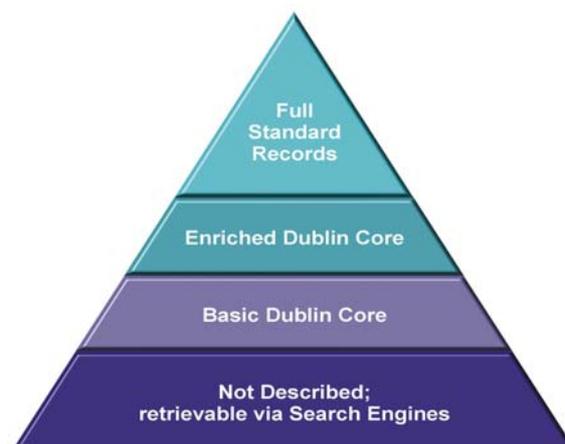
National bibliographies should include the current national output, and where practicable, they should also provide retrospective coverage. When necessary, selection criteria should be defined and published by the national bibliographic agency, 1)

we considered including remote electronic materials. There were not many of these documents waiting for bibliographic control in our country in 1998, and we decided not to include them, but instead to monitor the situation in our country and also the development of standards and co-operation abroad, and come back to this issue within one or two years. To be frank, we were aware of the fact that the documents already in existence but not included would disappear. However, we did not have enough experience and resources either for development of selection criteria and analysis of different standards used for their cataloguing or for their cataloguing itself. With our limited financial and human resources we were afraid to start a project about which nobody could predict future demands and results. Our situation could be perfectly described by the first sentence of Michael Gorman's paper presented at the IFLA conference in 2001:

The great irony of our present situation is that we have reached near-perfection in bibliographic control of "traditional" library materials at the same time as the advent of electronic resources is seen by some as threatening the very existence of library services – including bibliographic control. 2)

Besides this not very optimistic statement, Gorman presented his model of the pyramid, which tried to classify the chaos to make bibliographic control of selected electronic resources feasible. The same subject was discussed and the same model used by John Byrum in 2002. 3)

Gorman's Practical Model for Bibliographic Control of Remote Resource



This pyramid concept has been extremely helpful for all our discussions connected with the different possibilities of getting Czech remote electronic resources under bibliographic control.

Registration, archiving and permanent accessibility of Czech electronic remote resources has been an important part of the research plan of the NL since 2000. The goal of our first R&D project, called WebArchiv, was to analyze the achievements in this area of the most advanced countries and to start activities in our country by building on results achieved abroad, and, of course, in adherence to standards developed abroad. The NL is active in solving legislation problems under the umbrella of international organizations such as CDNL. It has established very fruitful co-operation with individual foreign libraries, mainly from the Nordic countries. This co-operation is focused mainly on technological development, and has resulted in the availability of programme tools, developed under

the umbrella of European Union projects and with our participation in development of these programmes. It was a great help for us to have the results of different foreign projects available; we could see the ways they tried to “climb” the pyramid, and their efforts at co-ordination. All these examples served as guidelines for us and allowed us to solve the problems of harvesting, selection criteria and metadata selection fairly fast. As soon as the guidelines and tools were ready, another difficult question appeared: who will build the two top levels of the pyramid? The NL established a small Department for Remote Electronic Resources, which is responsible for all the activities connected with these resources, including negotiations with publishers. Now there is not much time left for creating Enriched Dublin Core and Full Standard Records and the top of the pyramid is very small. We are trying to find resources and ways of co-operation that should improve the situation.

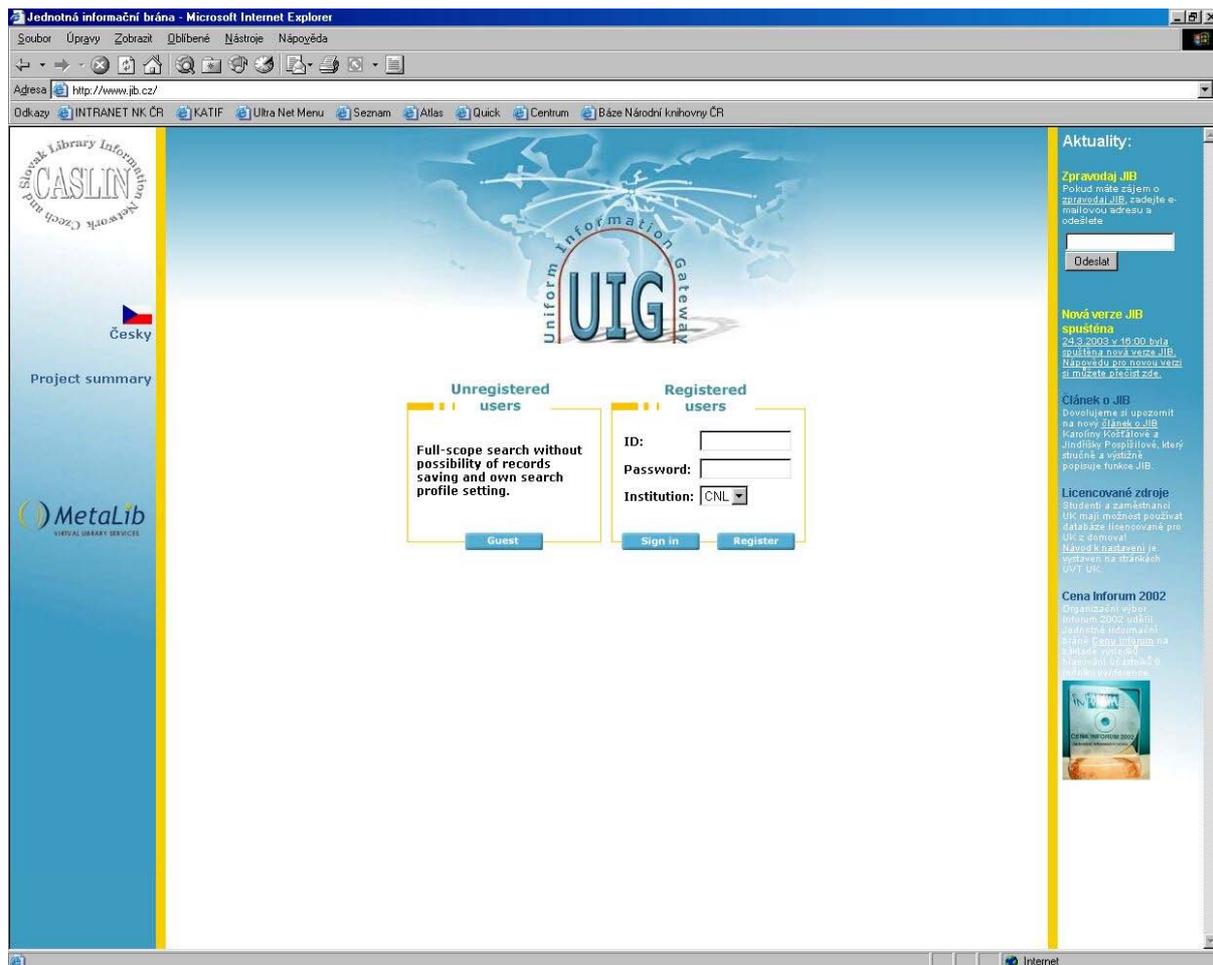
The Czech National Subject Gateway Project

When the first activities connected with the Czech National Subject Gateway Project got under way, it was quite logical to start thinking about dividing some responsibilities for building the top levels of the pyramid according to subjects. The flood in Prague in 2002 resulted in different state budgetary priorities and extremely limited funding for the Czech National Subject Gateway Project in both 2002 and 2003. But the guidelines for creating the Czech subject gateways – based on international standards and foreign experience -- were completed and published. Many European, Australian and other subject gateways were analysed. Special attention was devoted to their content, selection criteria and management. Complicated tables compared metadata sets, mainly descriptive, subject and administrative ones. Those used most frequently were included in the recommendations. We were also very much interested in the classifications and subject headings used in different projects. The Renardus project and its results were extremely useful for us; however, we went further to compare them with solutions in countries and projects outside Renardus. The Czech National Subject Gateway Project is connected with the Uniform Information Gateway Project. This is a very complex project integrating heterogeneous information resources (including full texts and different digital objects) from different countries. Its complexity shows both national bibliographies and subject gateways in broader context and it will be useful to mention it.

The Uniform Information Gateway and integration of heterogeneous resources

The Uniform Information Gateway (UIG) is a fully running national project. It can be examined and tested at:

<http://www.jib.cz>.



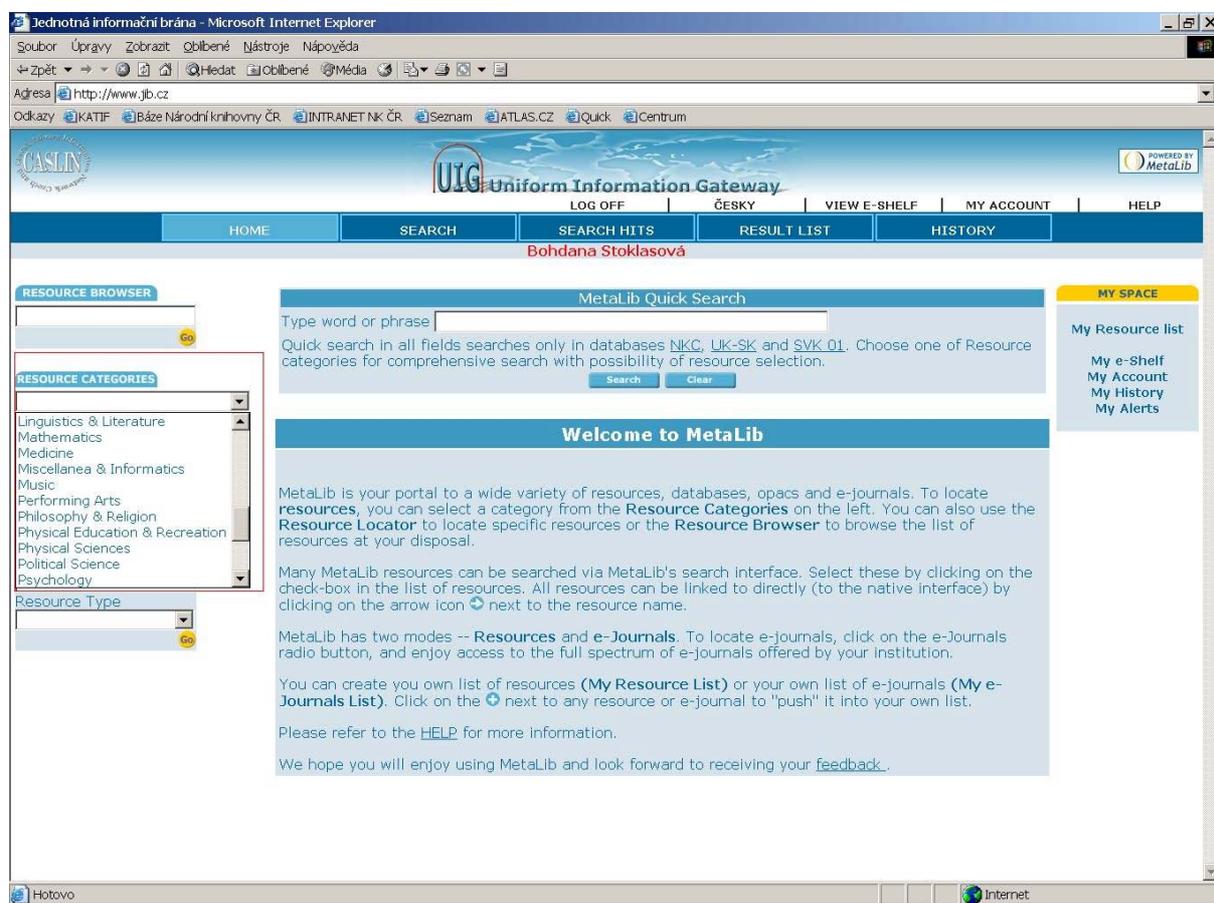
The UIG project has four main goals:

1. To integrate and provide qualitatively new access to heterogeneous information resources: integrated access to information on documents and to the primary documents themselves; uniform user-friendly interface; uniform query input and display of results; merge and de-duplication of results, authentication and authorization-based services; configuration of a personalized environment, including a list of frequently accessed resources and a personal cataloguing area with a possibility of changing the format; save queries and results options; no need for repeated logins and logouts (single sign-in to the array of resources); extended services for holding status, full texts, abstract, electronic document delivery and ILL, information on authors; related documents according to the subject and title.
2. Collection level description and collection development: unified, clear, standardized and internationally accepted description of resources available in the Czech Republic (i.e. "thematic map"); co-operation in collection development and sharing.
3. Subject access transformation: national authorities – subjects; English equivalents, DDC-UDC concordance tables.
4. Czech National Subject Gateway: creating of uniform guidelines for Czech subject gateways building based on international standards; distributing of responsibilities for covering particular subject; national co-ordination.

When we started with the UIG project in 2001, we expected serious difficulties in the areas mentioned under Goal 1. Goals 2, 3 and 4 were expected to be much easier. However, real life resulted in experience very different from what was expected. "Technical" integration with all its listed aspects was surprisingly easy. The UIG project is powered by MetaLib and SFX. These tools met all our demands and sometimes even offered more than expected. There are other similar tools available on the market offering similar functionality. This is just to say that "technical" aspects of integration are manageable provided that the information concerning resources to be integrated is prepared and presented in adherence to international standards – both library and technical ones -- and provided that

you know which are available and can be integrated for some country, subject etc. As mentioned above, the UIG integrates not only information on documents, but also documents themselves, mainly full texts. We expected that integration of information on documents, prepared, organised and offered for many years by libraries, would be easy, while the integration of full texts offered only during recent years by different information brokers and aggregators would be much more complicated. Surprisingly, integration of services offered by information brokers and aggregators proved to be often much easier to accomplish than integration of those offered by libraries. Compared to libraries, these organisations seem to be more flexible and more aware of the fact that for more and more users, information which cannot be integrated simply does not exist.

Let us look at the Goals 2,3 and 4. There exists a strong connection among them. From the outset, UIG users can choose some of the broad subject categories. We decided to use the 24 WLN Conspectus divisions as the “top subject crossroads” and organise different resources under the umbrella of these broad categories.



Theoretically, this would seem to be quite an easy job – just to identify and connect libraries with strong collections in particular subject areas with subject gateways concentrating on these subjects and other resources, including e-journals associated with these subjects. However, in real life this is often far from easy, mainly for libraries. Some libraries have a very good collection level description and collection development strategy available online and understandable internationally. These libraries sometimes build subject gateways covering their strong subject areas. This was not the case in our country, which did not allow us to connect Czech libraries properly. Building a unified, clear, standardized and internationally accepted description of resources available in the Czech Republic (i.e., a “thematic map”) has only started, based on the WLN Conspectus method. Libraries having strong collections in particular subjects will be responsible for building subject gateways focused on

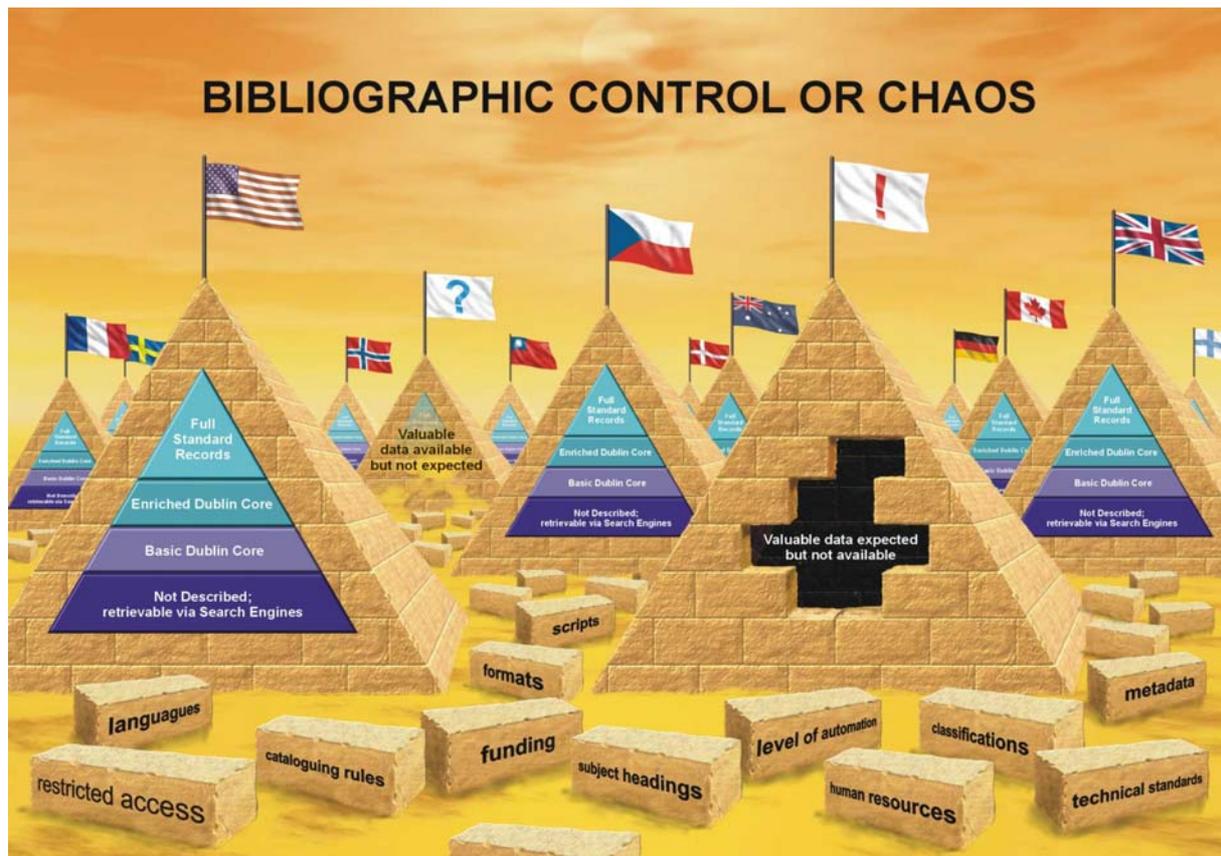
these subjects. We would like to concentrate on producing high quality descriptions of Czech resources so that other countries can share them, and just connect in those catalogued abroad – not catalogue them again. Unfortunately, because of different standards, many documents (printed as well as electronic ones) are still catalogued many times in different countries while many of them lack any bibliographic control.

For remote electronic resources we wanted to apply the same model as we have done for other documents listed in the CNB. We changed cataloguing standards for both bibliographic description and subject cataloguing, which allowed us to contribute our national output to the WorldCat database and vice versa – that is, to download records for our foreign acquisitions from WorldCat. We decided to add the LCSH in English to our records. This was mainly to make sharing of our records easier. There was much discussion as to whether we can afford to prepare this “added value” for foreign cataloguers or not. Nowadays we see that this was an extremely wise decision, not only for foreign cataloguers but mainly for our users. When they start their subject search in the UIG in the CNB or NL catalogue in Czech, they can see the English equivalents. These are extremely helpful for them for parallel searching of foreign resources and/or extended services based mainly on English.

Changing of existing standards during recent years and implementing of the new ones was far from easy for us. But we can hardly imagine the integration in the UIG without adherence to them. It is much better to have the seamless connection based on the same standards than to build bridges to overcome differences. However, we could not avoid one bridge. There is a long UDC tradition in our country and millions of existing records containing UDC. We had to solve the problem of connecting the large amount of resources and ever-increasing number of subject gateways based on the DDC. We decided to use the UDC-DDC concordance tables as a bridge. They are based on the WLN Conspectus categories (approximately 500). There was some very general background prepared by the WLN that was completed by the NL and is available there. We are testing this bridge just now.

Conclusion

Integration of heterogeneous resources is not a luxury; it should be a standard service of modern libraries if they want to be successful in the information market—which is now worldwide. Integration of heterogeneous resources is very much connected with subject gateways and bibliographic control in general. It is easy to integrate everything which is under control, but it is not feasible to integrate chaos. When we start to integrate, we are often in a similar situation to the conquerors of the Egyptian pyramids. Some obstacles not allowing us to get to the expected content are the same as thousands of years ago; namely, misleading navigation to places where one expects something valuable, but only an empty pyramid is found, and, on the other hand, very well hidden pyramids, containing treasures. Also there are different languages and scripts, the necessity of knowing a “password“, etc. Some obstacles are typical of the 21st century (different formats, different cataloguing rules, different metadata, different classifications and subject headings, different technical standards, lack of financial and human resources wasted for overcoming all these differences).



Can this picture be changed? We strongly believe so. Different languages and scripts have to be protected and developed. Some documents will always have restricted or limited access and the level of automation will remain different from nation to nation for a long time. However, more co-operation and co-ordination in the area of standardisation – both librarian and technical -- might result in stronger application of the UBC principles in real life and in saving the human and financial resources so much needed to reduce the chaos. Provided that we are able to agree on the same standards, and different countries are really able to register their national output in strict accordance to these standards while allowing the others to share their results, the picture might become very different and the word chaos might be removed from the title!

- 1) The final recommendation of the International Conference on National Bibliographic Services : Copenhagen, 25-27 November 1998. Available at <<http://www.ifla.org/IV/3/icnbs/fina.htm>> Recommendation no. 4.
- 2) Gorman, Michael, “Bibliographic control or chaos: an agenda for National Bibliographic Services in the 21st century“. IFLA Journal 27 5/6: pp. 312.
- 3) Byrum, John. D., “Challenges of providing bibliographic access to remote electronic resources in national bibliographies : problems and solutions – an overview [Paper presented at the 68th IFLA Council and General Conference, Glasgow, 2002]