

CLASSIFICATION AT A CROSSROADS

multiple directions to usability

The Hague, 29-30 October 2009

Rosa San Segundo

Using MARC classification format for UDC and mappings to other KO systems for an enriched authority file

The USMARC classification format, which was developed in the early 90s for the DDC and LCC systems

Was also conceived to house other classification systems such as the UDC

Applying the MARC classification format to the UDC

- The advantage of a **MARC classification format for the UDC** would entail ensuring that the Master Reference File (MRF) converges towards an authority and would also include notation combinations

MARC classification format for the UDC

- A priority objective for library catalogues has been **subject interoperability between traditional classification systems** such as UDC, DDC, LCC and subject headings

Interoperability between classification systems

- The one which had most impact was the ***Dewey for Windows 22***
- Offers interoperability between the classification numbers, the alphabetical index of the tables and the LCSH
- The **LCC Plus**, in its electronic version, has the LCSH linked to it, too

Dewey for Windows 22

- The creation of subject gateways from the **authorities of a catalogue** will be the instrument for the creation of **interoperability between classification systems** and subject headings

Interoperability between classification systems

- These interoperability gateways are the structure which will be the backbone of **more universal access to hypertextual navigation structures** supported by classification systems and, as is proposed for the UDC

More universal access to hypertextual navigation structures

- A great limitation in current library systems is the difficulty in retrieval systematic classification, since classification-related data cannot be integrated into them since the difficulty of undertaking searches of **the different elements that constitute a notation of UDC**

Introduction to the application of the MARC classification format

- As current library catalogues are not designed to contain data in such as complex form
- The use of the MARC classification format enables a chain of **symbols of a notation** to be loaded in a MARC classification field and to be retrieved

Symbols of a notation

- MARC formats are formats for the online representation and communication of bibliographical data
- A **MARC authority** is valid in bibliographical records and catalogues for authority files, whether they are proper names, subject headings or others
- The MARC classification format has been created for **classification authority records**

Classification authority records

- The **Library of Congress** began developing the MARC format for classification numbers owing to the need to manage the control of authorities in its catalogue

Library of Congress

- It was begun on associating the classification number **with subject headings** assigned in the authorities file, as well as the alphabetical terminology contained in the tables of the classification systems used

With subject headings

- The USMARC Classification Format was conceived for the **DDC and LCC systems**, although its creators also masterminded it so that other classification systems such as the UDC could be incorporated and the **UDC was included** in its plans

The **DDC and LCC systems**
UDC was included

- The application of the MARC classification format offers the possibility of developing a retrieval by UDC, as the notation will be broken down by authorities

UDC authorities

- The application of MARC classification enables retrieval via the notations of a classification system and it will operate like a **Classification Authorities File**, where the classification data will provide the authority for classification records and may be retrieval broken down in a way which is similar to other MARC authorities formats

The MARC classification format as a **Classification Authorities File**

- Once it has been introduced into an authority, a UDC notation will operate in a similar way to **other authorities files** and their retrieval will mean obtaining better performance from the CDU

UDC notation

- it is essential to have a file with systematic classification with a classification authorities file and **a friendly interface** which enables visualisation and navigation

A friendly interface

- **The objectives of the classification format** are numerous as it is fundamental for it to be used in systematic retrieval by users
- And also management
- And online **multilingual indexes**

The objectives of the classification format and Multilingual indexes

- The Library of Congress Library, a pioneer in **implementing MARC formats**, is attempting to convert the whole **LCC** to MARC format, a process which was begun with the HF and HG tables
- It is also addressing the conversion of the CDD; this conversion is more viable and has been based on the twentieth edition

LCC

- The application of the **MARC classification format to the UDC** has still not been undertaken

MARC classification format to the UDC

- In order to manage, maintain and interchange the classification data of the UDC, a format will be needed which houses it differentiating the following elements:
- special auxiliary number(s), and other auxiliaries

UDC: differentiating the auxiliary number

- Also the relationship between **elements within a compound number of the CDU**; broader hierarchies; references, scope notes, application notes, and the previous record of notation terms and indexes

Elements within a compound number of the CDU

- The application of the MARC classification enables good retrieval via the notations of a classification system as the notations will be in a **Classification Authorities File**

- A priority objective for library catalogues was subject interoperability and especially **interoperability between traditional classification systems** (UDC, DDC and the LCC)
- Mapping

Application and types of subject interoperability

- The creation of subject gateways is the instrument for the creation of interoperability between classification systems, subject headings and thematic descriptions which provide a fundamental structure that is going to support digital libraries which, at all times, point towards the universal virtual library or the **universal bibliographical catalogue** projected by Otlet and La Fontaine

Universal bibliographical catalogue

- The **interoperability** of contents is very broad and can be established between different classification systems, and between different languages within the same classification system, between alphabetic and systematic systems and others

Interoperability

- The ideal thing would be that the user formulated a single search, instead of formulating it in different ways and in different languages, for which the ideal thing would be to make different controlled vocabularies

**Classification systems
interoperable**

- Interoperability between subject headings in the same language, the most noteworthy is that implemented between the **Library of Congress Subject Headings (LCSH)** and the **Medicine Subject Headings (MeSH)**, where an attempt has been made to integrate online catalogue subject headings which establish relations between authority files

1. Interoperability between subject headings in the same language

- Making use of the MARC format, relationships have been established between authorities file of over 10,000 records, although the entire synthetic **LCSH structure** cannot be done completely as it has to be done manually in order to finish it
- This case of interoperability has been exported methodologically for other subjects

LCSH structure

- The **UMLS Unified Medical Language System**
- has made over **thirty** subject heading vocabularies, thesaurus terms and classification systems interoperable, using for the latter processing techniques based on lexical units

UMLS Unified Medical Language System

- **Multilingual subject headings** serve to set up a system of linguistic equivalences or relations between terms in different languages.

- This task of comparing terms is not easy

2. Interoperability between multilingual subject headings

- The European multilingual project **MACS, Multilingual Access to Interoperability** Subjects
- Switzerland, as a multilingual country, was especially interested in the articulation of this project UMLS Unified

MACS, Multilingual Access to Interoperability

- It establishes interoperability between three subject heading lists:
 - Library of Congress Subject Headings (**LCSH**)
 - **RAMEAU**
 - **SWD** Deuthe Wisdon, enabling access to a common list of the three, SHL or Subject Headings Languages
- MACS, Multilingual Access to Interoperability**

- The search can be formulated in the four national libraries, either in one or in all of them: the **Swiss** National Library, the **French** National Library, the **German** Library and the **British** Library
- Searches with SHLs

MACS, Multilingual Access to Interoperability

- As regards **multilingual subject headings**, the multilingual database on French monumental heritage, **MERIMÉE**, is also noteworthy
- The latter encompasses religious, civil, school, military and industrial architecture and aims to articulate interoperability between vocabularies controlled in different languages

MERIMÉE

- Thus the MERIMEE database is linked to some databases on adding the

- THESAURUS
- PALISY
- MEMOIRE
- and ARCHIDOC
- databases.

MERIMEE

- Another European project on interoperability **between multilingual subject headings is the HEREIN European Project**, the European Information Network of Cultural Heritage Policies
- It was sponsored by: Spain, France, Ireland, the United Kingdom, Norway and Hungary and Belgium later on.

HEREIN

- The multilingual thesaurus included in the HEREIN project aims to offer a **terminological standard** on national policies which address the subject of architectural and archaeological heritage HEREIN

HEREIN

- Interoperability between subject headings and classification systems in the same language

3. Interoperability between subject headings and classification systems in the same language

- The DDC for Windows 22
- Interoperability is established between the classification **numbers** and the alphabetical **headings of the tables**, to their notations, and to the Library of Congress Subject Headings (**LCSH**)
- The latter has already been applied to the two latest editions of the tables

DDC 22 ed

- The LCC and **the LCSH**, Library of Congress Subject Heading
- Called **LCC Plus** (Library of Congress Classification) which is available in the Web Classification

LCC Plus
electronic

- As regards interoperability with the UDC, a new product exists which is based on the abbreviated edition of the UDC and its correlation and translation to the General **Finish Subject Headings**.
- In **Spain**, interoperability between the UDC and the Subject Headings List of the National Library of Madrid from the authorities file is being undertaken, a project which is sponsored by AENOR

UDC in Spain whit subject Headings

- Between classification systems
- One of the most noteworthy is the British teaching thesaurus project, **HILT High Level Thesaurus Project HILT**

4. Interoperability between classification systems

- Project originated in Great Britain, it also aims to encompass Australia, Canada and the USA
- Different controlled vocabularies interoperable, for subject search and navigation
- The interoperability it presents **encompasses the LCSH, the DDC and the UNESCO Thesaurus, UDC and AAT, the Art and Architecture Thesaurus**

HILT High Level Thesaurus Project HILT

- The DARPA project, through **controlled vocabularies and metadata** vocabularies
- All of the above is articulated in the DARPA project

DARPA

- The European **RENARDUS** project
- between specific classification systems, different vocabularies controlled in different languages
- It is a programme based on technical models and computer tools, which is why interoperability is essentially based on

RENARDUS

- Polish interoperability project between controlled vocabularies which converges several controlled vocabularies such as
 - The Subject Headings Language (**SHL**)
 - The Thesaurus of Common Topics (**TCT**)
 - The **UDC**
 - The Polish Thematic Classification (**PTC**)

POLAND

- Likewise, the **American Mathematics Society (AMS)** is working on interoperability between
- Mathematics Subject Classification (**MSC**)
- DDC, 510 class related to Mathematics
- Developed at the New York State University

**American Mathematics Society
(AMS)**

- **Swedish Classification System (SAB) and the DDC**
- The project is financed by the Royal Library of Sweden and the table conversion is published by the library and can be found online
- It aims to articulate a new collaboration framework on the subject of heritage in these countries

Swedish Classification System (SAB) and the DDC

- It was not based on any already existing thesaurus, but it attempted to shape a specialised **multilingual thesaurus** in three languages: English, French and Spanish and it may be extended to other languages

Swedish Classification System (SAB) and the DDC

- The advantage of a classification format for the CDU also entails being able to ensure that the **Master Reference File (MRF) or the matrix file of the UDC tables**, from which all the translations, version and adaptations of the UDC emanate, converges as an **authority in the MARC classification format**

Proposal for applying the **MARC Classification Format to MRF with subject interoperability**

- The classification format applied to the CDU means that **the authority of the numbers and the notations** can also be managed using the MRF and it is integrated as such

Using the MRF

- For a synthetic classification such as the CDU:
- The most important advantage of a classification format would be the possibility of using already implemented **new notations** and seeking each compound number in broken-down format, in addition to using natural language terms to seek classification numbers

Synthetic classification such as the CDU

- Classification numbers will provide the **authority for classification records**, which can be saved and maintained as a separate file, similar to other MARC authority formats

Authority for classification records

- Their application to the CDU may bring with it a tool for all the catalogues using it
- The classification data incorporated into online systems will be a tool with great **potential for the UDC** which has yet to be tackled

Potential for the UDC

Thank you



for your attention

rosa.sansegundo@uc3m.es

CLASSIFICATION AT A CROSSROADS

multiple directions to usability

The Hague, 29-30 October 2009