The Functional Requirements for Bibliographic Records and Knowledge Organization

Abstract: Functional Requirements for Bibliographic Records (FRBR, 1998), the study commissioned by IFLA, brings revolutionary changes in the way we see modern computer catalogues. The catalogue is not seen as a sequence of bibliographic records and a copy of a card catalogue, but as an interconnected network of related information. Implications of the new model for the future development of catalogues are discussed. Special attention is given to access points and relationships between entities and the changes those will bring into both the formal and subject cataloguing, and authority files.

1. Introduction

The last three decades of the last century brought unprecedented changes to both the ways libraries and documentation centres operated and to their users' needs and expectations. Library automation, development of large bibliographic databases, union catalogue systems and shared cataloguing, new forms of publishing etc. were some of the revolutionary developments. At same time libraries and bibliographic agencies were faced with the need to reduce the high costs of operation, particularly for cataloguing and indexing.

In 1990 IFLA sponsored the Seminar on Bibliographic record in Stockholm, which addressed these issues. The Seminar acknowledged the need of libraries to reduce the cost of cataloguing, and saw as focus for further research: meeting user needs associated with the use of various types of materials and the broad range of eventual new requirements for bibliographic records. One of the resolutions of the Seminar was therefore to commission a study “to define the functional requirements for bibliographic records”. The terms of reference for the study stated as its purpose and scope “to delineate in clearly defined terms the functions performed by the bibliographic record”.

In 1992 the Standing Committee of the IFLA Section on Cataloguing accepted the terms of reference and appointed a study group. The final report of the study was accepted at the IFLA Conference in Copenhagen in 1997 and published the following year (FRBR, 1998, p.2-3).

The study group described as the aim of the study “to produce a framework that would provide a clear, precisely stated, and commonly shared understanding of what it is that the bibliographic record aims to provide information about, and what it is that we expect the record to achieve in terms of answering user needs.” (FRBR,
The study had two objectives. “The first is to provide a clearly defined, structured framework for relating the data that are recorded in bibliographic records to the needs of the users of those records. The second objective is to recommend a basic level of functionality for records created by national bibliographic agencies.” (FRBR, 1998, p.7).

In this paper we discuss the first objective only. To fulfill this objective the working group developed an entity-relationship model of a bibliographical database. This is quite revolutionary. For the first time the problem discussed is not: how can we make a more or less adequate surrogate of the title page, augmented with a few extra access points, or how can we automate our existing cataloguing process, but how can a computer and appropriate software help to give access to information? Of course the surrogate of the title page and the existing catalogue practice were adequate to solve the access problem as long as the only bearer of information were books or articles. Finding the books or journals meant finding the information. Nowadays the connection between content and its bearer is much looser and we have difficulties to impress upon our students the idea that some information is not on the Internet, but only in old-fashioned books or journals.

2. The entity-relation model

In the entity-relationship model a set of objects of interest or entities is defined, and the relations between the entities are listed. Further the important characteristics or attributes of each entity are identified.

Three groups of entities are defined in the FRBR (1998). Group 1 entities (with no common name, but we propose the term ‘bibliographical entities’) include work, expression, manifestation, and item. These entities represent the information traditionally reflected in formal cataloguing part of bibliographic records. Group 2 entities (‘name entities’) comprise of persons and corporate bodies responsible for the intellectual or artistic content, the physical production and dissemination or the custodianship of bibliographic entities. The third group (‘subject entities’) represents the subject of works and it includes concept, object, event, and place. Also the entities of the first and second group can be subject of works.

Relationships serve as a link between entities and enable the user to navigate within the bibliographic database and beyond. Relationships can link entities of different types and instances of entities of the same type. There is e.g. a relation between a work and all the expressions derived from it, or a relationship between the author (name) and a work as examples of entities of different type. The relationship between all expressions of one work (e.g. translations) is a relationship of entities of the same type.

The study identifies four generic tasks that the users perform while searching: “to find entities that correspond to the user’s stated search criteria (i.e., to locate either a single entity or a set of entities in a file or database as the result of a search using an attribute or relationship of the entity); to identify an entity (i.e., to confirm that the entity described corresponds to the entity sought, or to distinguish between two or more entities with similar characteristics); to select an entity that is appropriate to the user’s stated search criteria (i.e., to choose an entity that meets the user’s requirement with respect to content, physical format, etc., or to reject an entity as being in appropriate to the user’s needs); and to obtain access to the entity described (i.e., to acquire an entity through purchase, loan, etc., or to access an
entity electronically through an online connection to a remote computer)." (FRBR, 1998, p.82).

3. Formal and subject cataloguing: differences and similarities

In its traditional meaning formal cataloguing has as main function to enable the access to *works* in a collection when formal characteristics, which are normally present in or at the *manifestations* ("books") catalogued, are known. Subject cataloguing has as main function to enable to find *manifestations* or *items* of which the content (the *expressions of work*) fulfill certain expectations.

There are two groups of problems in these traditional views. In many cases the user wants a *manifestation* of an *expression*. If I go to a library to get a copy of the latest novel of Patricia Cornwell, it is irrelevant if it is a hardback or a paperback, but the question if it is a Slovenian or an English expression can be highly relevant. Looking for a sound recording of the Ninth Symphony of Ludwig von Beethoven many music lovers will be far more interested in a given expression than in a given manifestation. It is impossible to say beforehand if a user is searching for *works*, for *expressions*, for *manifestations*, or even a given *item*.

The second problem is that the distinction between formal and subject cataloguing is not always easy to make. The user can use the same access points to search for a known *manifestation* and a *manifestation* on a given subject. The well-known example is title words as access points. A keyword search with the name of a person in a bibliographic database gives both *manifestations* of *works* created by the person, and *manifestations* of *works* about the person.

4. Description and access points A bibliographic record consists of three parts: a *description*, a number of *access points* and administrative data like the date the record was created. Bibliographic records are built following detailed rules. An example of such a rule set is the UNIMARC Manual (UNIMARC 1994). Many (sub)fields of the records are filled with values defined in other manuals, e.g. field 200 of UNIMARC contains the title and statement of responsibility. The values of this field is regulated in a *International Standard Bibliographic Description (ISBD)* and cataloguing rules like the *Anglo-American Cataloguing Rules, 2nd edition (AACR2)*. Figure 1 gives an example, taken from the Slovenian National Bibliography.
Figure 1: A Unimarc record

The description is contained in the fields 010, 2-, and 3-. The 5-, 6- and 7- fields contain access points. The other fields are administrative fields. Note that there is a lot of redundancy in this record: often the same data are given twice in the record, once in the description and once as access point. The name of the author is given in field 200 and again in field 700. The difference between the two mentions is the form of the name: S. Stanic in 200 and Stane Stani in field 700. Only the form given in field 700 can be found in the index of the Slovenian National Bibliography as Siani, Slane. A search with Sianie gives no result. In the Dutch National Bibliography the book is found with Sianie but not with Stani! Of course it should be found under both forms of the names: the correct Slovenian form, and the form the Dutch publisher put on the title page.

To achieve this goal the different forms of the name should be connected with each other. There are two possibilities: put the different forms in the bibliographic record (e.g. in the 9-fields of UNIMARC), or use authority files. The first possibility would mean that for each possible form of the name a new field in the bibliographic record should be filled. With authority files it would be possible to have in the bibliographic record only the form as is given on the title page. In one record of the authority file all different forms of the name of the same person are collected. Searching with a name of a person starts in the authority file to find all forms of the name, which then are used as search terms for the bibliographic
169 records (as long as the user does not stipulate that she only wants to search with the form given by her). In this view it is not necessary to declare one of the forms of a name as the uniform form. The uniform heading was used in card and list catalogues to get descriptions of authors and contributors together at one place in the alphabetical order (Verona, 1963). And the alphabetical order is a necessary evil in card and list catalogues to make searching possible. The FRBR don’t require uniform headings, but it is difficult to see how the model can be adopted without authority files, at least for names.

Titles are a different matter. The identify task mentioned in the FRBR makes it almost compulsory to mentioned the original title or other alternatives titles in the description. This extra titles in the descriptions can be indexed and used for searching in the same way as the main title, which is given only once in the bibliographic record.

5. Access points to more characteristics and more access points per characteristics

The FRBR gives for each entity a list of characteristics that should be included in the relevant records. Some of them are only relevant for the identify, select or obtain tasks, but most characteristics are for at least some users very usable as access points. A few examples: a teacher of Dutch at the University of Bucharest can be interested if and which books in Dutch are present in the collection of the University Library. Only if he can search with language as access point his curiosity can be satisfied. For others it could be interesting to be able to search for books published in given cities, or printed before a given year. Given access to more different characteristics is expensive when one has to make an extra card catalogue or printed index, e.g. for printers or places of publications. For the present computers and database programs the extra costs can be neglected, and in many catalogues the data are present. Often however they are not available for searching or even viewing.

Another point that is a typical relict of the card and list catalogues is the restriction of the number of access points for one characteristic. When the number of authors of a work is more then three often it can only be found with the name of the first author (Braun, 1969). The AACR2 states that when there four or more persons or bodies are involved in a particular instance, an access point should be made only for the one named first, although it also said that if an extra entry is required in the context of a given catalogue, it should be made. This “rule of three” is discussed at the moment by the Joint Steering Committee for Revision of Anglo-American Cataloguing Rules. It should remembered that the “rule of three” is not in accordance with the FRBR’s statement that the catalogue should assist the user to do at least the following: “Find all manifestations embodying: the works for which a given person or corporate body is responsible ...”, nor was it ever in accordance with Cutter’s often cited requirement that the objective a catalogue is to enable a per son to find a book of which either the author, the title or subject is known (Cutter, 1876).
6. Challenges and conclusions

For the future we need systems, which will "meet user needs associated with the use of various types of materials and the broad range of eventual new requirements for bibliographic records". There are several questions that have to be answered before such systems can be answered. We mention only a few:

1. What to do with the distinction between the content, the information and its carrier?

2. What is the basis of the cataloguing: the work, the expression, or the manifestation?

3. Connected with these two questions is the question how bibliographic databases should be built and, particularly, presented to the user.

The first question is more complex than it looks at the first sight. In most cataloguing rules there are separate rules for different types of what is called bibliographic forms. The IFLA published several International Standards for Bibliographic Description (ISBDs). They exist for Monographs, Serials, Non-Book Materials, Printed Music, Electronic Resources, Cartographic Materials, and for Antiquarian Books. The IFLA also published an ISBD(G) meant to unify the more specialized ones, and Guidelines for the cataloguing of entities that are parts of greater entities, like journal articles. The AACR2 has a chapter "General Rules for Description" followed by 12 chapters giving instructions for the description of more specialized bibliographic forms, like Book, pamphlets and printed sheets, Music, Sound recordings, Computer files, Manuscripts and Microforms. These bibliographic forms are a strange mixture of content, physical carrier, way of publishing, etc. This gives more and more problems in a time when mixture of these forms are more and more common. There have been pleas for rules, that are not organized according to bibliographic forms, but based on the different fields recognized in ISBD(G), like title, edition, imprint and so on (Delsey, 1998). In 2001 Zlata Dimec and her collaborators published Slovenian rules based on this principle (PREKAT, 2001). In our view this is the only practicable way, especially when a strict separation of data of work and expression (both content) and manifestation and item (both carrier) can be established.

The second question is and will be the subject of much discussion. A simple answer does not exist. The existing cataloguing is based on the idea that the manifestation is described and access points are chosen for the work. This principle works all right as long as the manifestation contains one work only and the title of manifestation and work is the same. When the manifestation contains more works or title of manifestation and work are not the same the difficulties can start. In the case of a translation what should be the title access point? Should it be the title of the original work, or the title of the translation (an expression) or both? The principle says the title of the work, many rules say the title of the manifestation and if wanted add an extra access point for the title of the work. In practice often both titles can be used in searching. In most instances the title of the manifestation is the only title access point in the case of more works in one manifestation, especially if there are more then three works involved.

It is possible that a user is looking for a particular expression of a given work, in a particular format, e.g. the Violin Concerto in A minor, BWV 1041 of Johann Sebastian Bach, played by Emmy Verhey in the Beurs van Berlage, Amsterdam, in 1992, and that expression on an audio-CD. One manifestation of this
concerto can be found on a CD called Violon Concertos. How can it be found? And the user is content with any CD on which the Concerto is placed. In this case an expression-oriented cataloguing will work much better than a work or manifestation oriented cataloguing! At least composer, title (in which form?) and performers should be access points. There is a change that in many existing catalogues of CDs the answer to the question is very difficult to find or not of all. Some existing bibliographic databases offer a keyword search, and then searching with BWV 1041 can help, if all the works on a CD are mentioned in an annotation.

This brings us to the third question. Future bibliographic databases should give access to all for the finding task relevant characteristics of the bibliographic entities work, expressions, manifestations, and items, with access to authority files for the name and subject entities. These authority files should enable the user to find the correct form of access points especially in difficult cases as notations of classifications and name-title access points when there are name or title variants. In this situation the bibliographic database consists of a network of connected records, e.g. a record for a work is connected with a number of records for expressions of that work and with records for manifestations in which one or more of the expressions is present. Each of these records can be connected with one or more records from authority files, which in return each can be connected with several records for bibliographic entities. In this way navigation from record to record is possible, from an authority record to records for bibliographic entities or to other authority file records.

For this network it is necessary that also for authority file records entity-relationship models will be developed. An IFLA working group is engaged in this work (Bourdon, 2001). Another task that is still waiting completion is the development of guidelines for display formats based on the networks described. Another IFLA working group is working on this.

However we can conclude that the FRBR will have a great impact on future knowledge organization. Maybe it is not going as fast as many may want, but there is a lot of discussion going. We are learning how computers can help to find information more easily and along paths Cutter, and even the participants of the Conference on Cataloguing in Paris (1963) could only dream of, or may be not even that.

Notes
1. In our view uniform headings are also not necessary for displays in cases where a search gives a great number of hits, like a search for Bibles or all editions and versions of Shakespeare’s Hamlet in the Library of Congress catalogue. Reformulating a question into e.g. Bibles AND Slovenian is faster than scroll through the surrogate card catalogue on the screen till Bible.Slovenian is reached. It is ironic to read in the FRBR that parallel titles should not be mandatory in national bibliographies any more, what is not helping those whose mother tongue is a “small” language. Abandoning the superfluous uniform (and main?) heading brings far more gain in labour productivity.

References
Bourdon, F. (2001). Functional Requirements and Numbering of Authority Records (FRANAR): to what extent authority control can be supported by technical means.


