The objectives of cataloguing, the technology and the user interface

Knut Hegna

31st May 2007

Abstract

This paper discusses the traditional objectives of library cataloguing and holds them up against the possibilities that the new technology opens up. An user interface for an FRBR based catalogue is proposed.
1 Introduction

I am very honoured of being invited to speak at this conference held at a university with traditions far back in European history. My own university – even though the oldest in Norway – was established 450 years after this one.

Key note speeches are supposed to be very visionary. This is somewhat scary and I am afraid I will disappoint you. But may be I can give you some new ideas for the library catalogue.

I am also a bit worried, as I know absolutely nothing about the Polish library community and the discussions going on here.

To make this speech up to date I would have had to talk about hot stuff like bloggs and library 2.0. This I will not do, even though it would be interesting enough. Somehow I feel that Library 2.0 is about letting the technology decide our activities, more than figuring out what the technology can do for us.

My concern is focused elsewhere. I will focus on some old and new objectives of library catalogues and on how the technology can be used to achieve these objectives globally.

2 Universal bibliographic control

I would like to start nearly 200 years after the establishment of this university, in the year 1545 with the work by the Swiss polyhistorian Conrad Gesner. His "Bibliotheca Universalis"[4] represents an attempt to do what we still are trying to do today, establishing "Universal bibliographic control" (UBC):

- including all writings in three languages, Latin, Greek, Hebrew
- existing and non-existing
- old and new
- scholarly and non-scholarly
- published and hidden in the libraries

But there is more to it than that, he didn’t do this as an intellectual exercise. The catalogue is also a union catalogue, a means of locating where to obtain the items of interest.

He states two other objectives. First:

necessary when a establishing public and private libraries,

I understand this both as distribution and sharing of cataloguing data and as a means of selecting items for your library. So, buy this book, mark up your items and you have your own catalogue. We are still trying to do this today, but we can’t quite agree on the rules and the formats.

The second objective:
but very useful for all that studies any art or science

this is the user perspective.

In Gesner’s catalogue there are 18 000 works by 3 000 authors. He certainly must have travelled a lot, visiting libraries and made notes of their catalogues, merging it all into one large catalogue. This merging represented a problem, he had to normalize the data from the catalogues. He did this by translating both titles and author names into Latin.

He did something else, he constructed an index of inverted names. To my knowledge this is the first time someone creates an inverted index of personal names.

He had a reason for doing this and it is found in Latin text on the top of the page, I quote

To help the users that remembers only one of several names of an author. If you know one of the names you can look into this index and find out where in the catalogue this authors works will be located, that is: under their first names.

There is another reason: Gesner states that this index also collocates authors with the same family name, and this is nice to know.

His principle here is to have the users in mind when making a catalogue and he constructs the catalogue with that principle as a foundation. The objectives is to make it possible to look up authors names in the catalogue and the rules are stated here in the introduction of the inverted index.

His users are both students and libraries. Read more about Gesner and his work in Havus paper[5].

3 Objectives and rules

It is important to differentiate between these three concepts:

- principles
- objectives
- rules

Having the users in mind is an important principle when making a catalogue. Principles should deal with the character of our rules. Like not having more rules than are necessary and sufficient. That means we must have the rules necessary to achieve the objectives, but not more rules than that. Another principle would be to adhere to international standards. This is to assure that we achieve the objective of reusing and exchanging bibliographical data.

You might hear someone say: It is a principle that we put down the author name in the catalogue. But is that really so? In my opinion putting down the author name is following a rule we make to achieve the objective of finding the works of one author or another.
The principle in this connection is that we enter the author name in the same form in the catalogue and collocating the works under this form, just like Conrad Gesner did when normalizing names to Latin. We use authority files to achieve the same effect.

To illustrate the difference between objectives and rules, let us go 100 years back. Everybody is citing Charles Cutter[3] and his Rules for a dictionary catalog these days. His objectives, after 28 years of discussions were – in 1904 – stated like this:

1. To enable a person to find a book of which either
   A. the author
   B. the title
   C. the subject

   is known.

2. To show what the library has
   D. by a given author
   E. on a given subject
   F. in a given kind of literature

3. To assist in the choice of a book
   G. as to its edition
   H. as to its character

The first point describes the finding objectives. The second is the collocating objective. The last point could be called the selecting (or choice) objective.

When formulating the objectives in this way, you get operational objectives, you see what rules you must have. Cutter summarizes the rules like this:

1. For the objectives A and D we need author-entry with the necessary references
2. For B we need title-entry og title-reference
3. For C and E, we need subject-entry, cross-references and classed subject-table
4. For F: form-entry and language-entry (for F)
5. For G: giving edition and imprint, with notes when necessary (for G)
6. For H: Notes

As you see, the rules or means are closely connected to the objectives.
4 The traditional catalogue

The old card catalogue had some qualities that were lost when converting to the online catalogues. Some say that these qualities are now unneeded because we can access with free text search every item in a catalogue record. I do not quite agree.

James Duff Brown published two works on libraries in the beginning of the 20th century. The first one is entitled Manual of library economy, the second one from which this is cited, is entitled Library classification and cataloguing[1].

This is his proposal for filing cards in the catalogue the works of an author with extensive production.

- Single works in chronological order of publication
  - Originals, including manuscripts
  - Reprints
  - Translations in foreign languages
  - Parodies
  - Criticism, etc., of single works

- Collected works in chronological order of publication
  - Author’s editions
  - Editor’s editions
  - Selections
  - Paraphrases and condensed version
  - Dramatic versions of works
  - Musical settings of works
  - ...

- Bibliography

- Index

Now if you were to pull out a drawer in the card catalogue with this structure, you would be presented to the authors whole life and works at a glance, with the help of guide cards.

The intellectual work put down in the filing according to specific rules disappeared when converting to the online data base catalogue. When the card catalogue represented structure, the online catalogue was just a huge amount of single records.

When you perform a search in an online catalogue the result list is often without structure even though it is based on the same cataloguing rules as the card catalogue.

Some years after the 2nd world war Seymour Lubetzky started out as a librarian at the Library of Congress. At that time the cataloguing rules of the library was a mess - a spaghetti - with even its own rule for cataloguing works about a certain building in Brasil.

Lubetzky began an examination of the rules to see which was necessary and sufficient to meet the central objectives. His work was the basis of the international
conference in Paris in 1961 where there was consensus on what is called the Paris principles, but as you now know, it should be called the Paris objectives.

To decide

1. Whether the library contains a particular book
   specified by
   
   a. its authors and title, or
   b. if no author is named in the book, its title alone, or
   c. if author or title are inappropriate or insufficient for identification, a suitable substitute for the title

2. a. which works by a particular author
   b. which editions of a work
   are in the library.

If you examine these objectives you see that they resemble the objectives of Charles M. Cutter, but they differ in that they don’t say anything about finding and collocating works by subject.

These objectives are the basis for the international cataloguing rules formulated in the ISBDs and these again for the implementation of the various national and international cataloguing rules, like AACR.

But there are some other phrasing that needs a comment or two. Both Cutter and the Paris principles are talking about local collections. The scope is the physical collection of a single library.

Today, we are in another situation. We must think globally, also when it comes to library catalogues.

5 The FRBR conceptual model


After ten years I suppose that everybody in the library community know something about the FRBR report or at least have heard somebody mention it. It is hazardous to try to summarize it, but I need to say a few words about it.

The language of FRBR is from the data base theory, it is called the entity-relationship model, or just ER. A very short description of the FRBR-model shows that the model consists of three elements: entities, attributes of the entities, relation between entities.

entities of which there are ten, divided in three groups:

- product entities: works, expressions of the works, manifestations of the expressions (documents), and finally items (copies) of the manifestations.
Works and expressions are abstract. Manifestations and items relate to physical publications.

- **responsibility entities**: persons and corporate bodies.
- **subject entities**: concepts, places, events and objects. The other entities might also appear as subject entities.

**attributes** the entities are described through the attributes they have. Some examples: a person has a name and birth date; a work has a title; a manifestation has a date of publication and an identification number (e.g. ISBN); an item (copy) has a call number. There are, of course, several more attributes than those mentioned here.

**relations** are defined as links or references between the entities. A work is created by a person. *is created by* is a relation between the work and the author of the work, so the author is not an attribute of the work. Some other relations: a work *is realized through* one or more expressions; an expression *is embodied in* one or more manifestations; a manifestation *is exemplified by* one or more items.

Let me give you a short example of some entities, attributes and relations from the FRBR model to give you an idea of what this is all about. The starting point is the novel by Joseph Conrad: *Heart of darkness*.

## 6 The objectives of FRBR

The objectives of FRBR are stated like this in the report:

**to find** entities that correspond to the user’s stated search criteria

**to identify** i.e. to confirm that the entity described corresponds to the entity sought

**to select** an entity that is appropriate to the user’s needs or to reject an entity as being inappropriate

**to obtain** the entity described through purchase, loan or through electronic access

This isn’t much about documents, it could have been about cats or penguins. These objectives have been criticized for being too abstract. Elaine Svenonius gives in her book *The intellectual foundations of information organization*[7] a more concrete or specific formulation of the objectives:

1. **To find/locate**
   - a single document
   - sets of all documents
     - belonging to the same work
belonging to the same edition
- by a given author
- on a given subject
- defined by other criteria

2. To navigate in a bibliographic database

- to find works related to a given work by generalization, association and aggregation
- to find attributes related by equivalence, association and hierarchy

These objectives do not in fact differ much from the objectives of the FRBR. But Svenonius is more clear when it comes to the objective of collocating. Neither of them limit the scope of the objectives to local collections.

This points in the direction that it is time to have global perspective on the objectives, we want to achieve universal bibliographic control. Can we get there now and still hold on the objectives or are we in the hands of Google-like solutions for the catalogue?

From my limited point of view the libraries can outdo Google if we stick to our objectives and use the technology.

7 The standards and the technology

We have in our hands access to the catalogue of every other library on the net, especially the national libraries who are in charge of cataloguing the nations documents and in many cases also in charge of maintaining authority files.

We have standard protocols to retrieve data from these catalogues and authority files like Z39.50 or SRU (Search/Retrieve via URL) and SRW (Search/Retrieve Web service).

We can retrieve the data in standardized machine readable formats, like MARC or XML.

We have character set standards covering most alphabets of the world (Unicode 5.0).

For nearly ten years we have had the conceptual model of the Functional requirements of bibliographic records (FRBR).

We have the conceptual model for authority information (names, places, FRBR entities).

We have a golden opportunity to reach universal bibliographic control. Just like Gesner managed to do, all by himself, almost five hundred years ago.

What we now need is cataloguing rules that both supports the old clear and simple objectives and also takes into account the new possibilities.

The Anglo-American Cataloguing Rules have been under revision for quite some time. The plan is to complete the revision during the year 2008, now under a new name Resource Description and Access (RDA). One would expect that these revised
rules would meet the new challenges. I am sad to say that I do not think they are. One of reasons is pointed out in an article by Coyle and Hillman[2]:

*the desire to continue the AACR tradition while acknowledging that a greater change is needed.*

In other words it is the old rules that decides, not the new challenges and the new situation.

One may wonder why they called the revision RDA and not just AACR3. The fact is that the “new” cataloguing rules, RDA, do not meet the requirements that stems from the changes in access to information that results from the growing amount of digital – on line accessible – material.

They do not even meet the requirements from the conceptual model of bibliographic records (FRBR) which originates from the library community itself.

This is very disappointing. My hope is that the discussions will not stop, but develop our understanding of what is needed today and in the future.

In order to contribute a bit to the discussions I would like to end this speech by presenting an OPAC user interface based on the FRBR model.

Going back to the *Heart of darkness*–example, we see, even though the structure might seem clear to us, that it is too complicated to present as a search result.

8 **OPAC displays guidelines**

In the IFLA-report on guidelines for online public access catalogue displays[8] the principles for the user interface of library catalogues are categorized in three groups:

- User needs principles
- Content and arrangement principle
- Standardization principle

I have already made some comments on the first and third principle. The report specifies the second principle such:

- *display what is asked for and needed for further action*
- *display records in an order meaningful to the user, rather than in a random order, when several records are retrieved*
- *support navigation from the displayed information to related information*

It is strange that we need to say these things out loud, but at least in our local systems in Norway this not always fulfilled. These arrangement principles focuses on the bibliographic records, but I think that in the age of FRBR we should focus on structure and relations.
9 Another OPAC

I will end this speech by presenting an idea for a new user interface for a library catalogue based on the FRBR structure and global communication.

The first screen presents a search box where you enter your criteria. The next step is to decide in what mode you want the search to be performed, is it by somebody, is it about something or somebody or may be you are looking for a specific work. This kind of functionality requires some processing of the query behind the scene, analyzing it, looking up in the relevant indices and so on. This idea is not mine, it has been implemented at library of Roskilde University in Denmark (see: http://www.rub.ruc.dk/).

In this example we try the by-function searching for persons with family name beginning with *Conrad*. The search results in a list of personal names going into the screen along a horizontal axis, as in figure 1. Each name is located on the top of a “card”. You may scroll back and forth in this list using the navigational tools in the lower right corner and navigate backwards and forward in the session with the tools in the upper left corner.

On each card you find the works related to the person along a vertical axis.

By clicking on any card you bring it to the front.

The navigational links are highlighted with red-coloured text, so you might focus on both a person or a work.

In the front here on the Joseph Conrad card is a list of his novels, but he has also written short stories, essays and there are collections of his letters, and he has created translations.

In this user interface you just click on the button on the right side of the card to turn it around to investigate his activity in full, without leaving the hit list. You keep the context. If there is not enough room on the back side, we can turn it round once more to find another back side. We can have as many back sides which is necessary to show the person’s full activity.

Now, as this card focuses on the person entity of Joseph Conrad you will also find a link to works about Joseph Conrad, that is, looking at him as a subject entity.

What would one expect to happen if we focus on Joseph Conrad by clicking on his name? Logically there are two options here, first we could present information from the authority file or we could get focus on his works. I have chosen this option.

We get list in a similar interface, as in figure 2. This time the focus is on the works of Joseph Conrad, one work on each card going along the horizontal axis and the expressions of the work along the vertical axis. We can imagine that the different expressions have been supplied by communication with national libraries all over the world.

You also see here that we have relational links to works about Heart of darkness and links to works based on Heart of darkness.

If the list was a result of a search for title, that is, if we used the works-button searching one or more words in title, the horizontal axis would be a mixture of work, expression and manifestation titles. But through the links you would be guided into the FRBR structure.
Figure 1: Hit list, performing a truncated search for “conrad”

Figure 2: Focusing on the work *heart of darkness*
What the hit list should look like must both be a function of what kind of search
the user has performed and an analysis of the hits that the search results in, so this
interface should be very flexible, but lead the user into a structure where navigation
is possible.

In the spirit of this interface: If we now choose to focus on the work Heart of
darkness what would you expect the next screen to be like?

The expressions of Heart of darkness is now along the horizontal axis and the
manifestations along the vertical axis, as in figure 3.

Now we have come to a point where we have sufficient information to decide
which manifestation you want to read and order.

This is my contribution to the discussion on how a user interface for library
catalogue might look like, taking into account bibliographical structures, relations
and elements from the FRBR report.

10 Conclusion

In 1904 Charles Cutter[3] summed up 28 years of the discussions on cataloguing
with these word:

Still I can not help thinking
that the golden age of cataloging is over,
and that the difficulties and discussions
which have furnished an innocent pleasure to so many
will interest them no more.

Another lost art.

I do not think he was right in this respect. I do hope that the discussions will continue and that they will furnish a pleasure to many people and last, but least, I hope it will not take 28 years.

Thank you.
References


