How can BIBSYS benefit from FRBR?

Ole Husby

BIBSYS

Lund, April 2002

BIBSYS today

The data model is extremely non-normalized,

with excessive replication of data.

Authority control is possible in just a few data
elements.

Why?

Because

Data replication favours search speed.

Insufficient understanding of bibliographic elements and structure, attributes, and relations.

The conserving force of millions of MARC records.

Need for exchanging records with other systems.

Data element content still targeted at human eyes.

Consequences

The search results are often confusing,

not collecting what should be collected,

not separating what should be separated.

And like the much discussed web search engine problem: Far too many hits.

Herbjørg Wassmo in BIBSYS

Search for **au=wassmo,herbjørg** : 182 hits

```
 1 De zevende ontmoeting              / Wassmo, Herbjørg   2001
 2 Het boek Karna : roman             / Wassmo, Herbjørg   1999
 3 La veranda cieca                   / Wassmo, Herbjørg   1989
 4 La véranda aveugle                                     1997
 5 La chambre silencieuse                                 1998
 6 Det stumme rommet [lydopptak]      / Wassmo, Herbjørg   1995
 7 Lykkens søn                        / Wassmo, Herbjørg   1994
 8 Gefühlloser Himmel : Roman         / Wassmo, Herbjørg   1992
 9 Huset med den blinde glassveranda  / Wassmo, Herbjørg   1984
10 Hudløs himmel                      / Wassmo, Herbjørg   1997
11 Huset med den blinda glasverandan  / Wassmo, Herbjørg   1992
12 Det stumma rummet                  / Wassmo, Herbjørg   1993
13 Sohn des Glücks : Roman            / Wassmo, Herbjørg   1994
14 Hudløs himmel : roman              / Wassmo, Herbjørg   1990
15 Pimeda klaasverandaga              / Wassmo, Herbjørg   1995
```

Herbjørg Wassmo in BIBSYS

Search for **au=wassmo,herbjørg** : 182 hits – **sort by title**

```
 1 Auf der Reise : vier Erzählungen  / Wassmo, Herbjørg   1996
 2 Ciel cruel                                             1998
 3 Ciel cruel : roman                / Wassmo, Herbjørg   1998
 4 Ciel cruel : roman                / Wassmo, Herbjørg   2001
 5 Das Buch Dina : Roman             / Wassmo, Herbjørg   1989
 6 Das Buch Dina : Roman             / Wassmo, Herbjørg   1994
 7 Das Buch Dina : Roman             / Wassmo, Herbjørg   1999
 8 Das Haus mit der blinden Glasver  / Wassmo, Herbjørg   1984
 9 De stille kamer                   / Wassmo, Herbjørg   1996
10 De zevende ontmoeting             / Wassmo, Herbjørg   2001
11 Der stumme Raum : Roman           / Wassmo, Herbjørg   1985
12 Det sjuende møte : roman          / Wassmo, Herbjørg   2000
13 Det sjuende møte : roman          / Wassmo, Herbjørg   2001
14 Det sjuende møte : roman          / Wassmo, Herbjørg   2001
15 Det sjunde mötet : roman          / Wassmo, Herbjørg   2001
```

Herbjørg Wassmo in BIBSYS

Search for **au=wassmo,herbjørg AND lan=nor**: 63 hits – **sort by title**

```
 1 Det sjuende møte : roman            / Wassmo, Herbjørg    2000
 2 Det sjuende møte : roman            / Wassmo, Herbjørg    2001
 3 Det sjuende møte : roman            / Wassmo, Herbjørg    2001
 4 Det stumme rommet                   / Wassmo, Herbjørg    1983
 5 Det stumme rommet                   / Wassmo, Herbjørg    1984
 6 Det stumme rommet                   / Wassmo, Herbjørg    1990
 7 Det stumme rommet                   / Wassmo, Herbjørg    1997
 8 Det stumme rommet : roman           / Wassmo, Herbjørg    1984
 9 Det stumme rommet : roman           / Wassmo, Herbjørg    1996
10 Det stumme rommet [lydopptak]       / Wassmo, Herbjørg    1995
11 Dinas bok                           / Wassmo, Herbjørg    1992
12 Dinas bok                           / Wassmo, Herbjørg    1993
13 Dinas bok                           / Wassmo, Herbjørg    1996
14 Dinas bok : roman                   / Wassmo, Herbjørg    1989
15 Dinas bok : roman                   / Wassmo, Herbjørg    1991
```

The Herbjørg Wassmo problem

is perhaps not our most important problem to solve ???

It is also urgent to find a bibliographic method for treating manifestation-level diversity like the mixture of printed and digital publication media. The area of electronic vs printed journals is one example.

FRBR in short

Functional Requirements for Bibliographic Records is the title of a report from an IFLA study group, presenting a model for bibliographic data based on the entity-relationship data model.

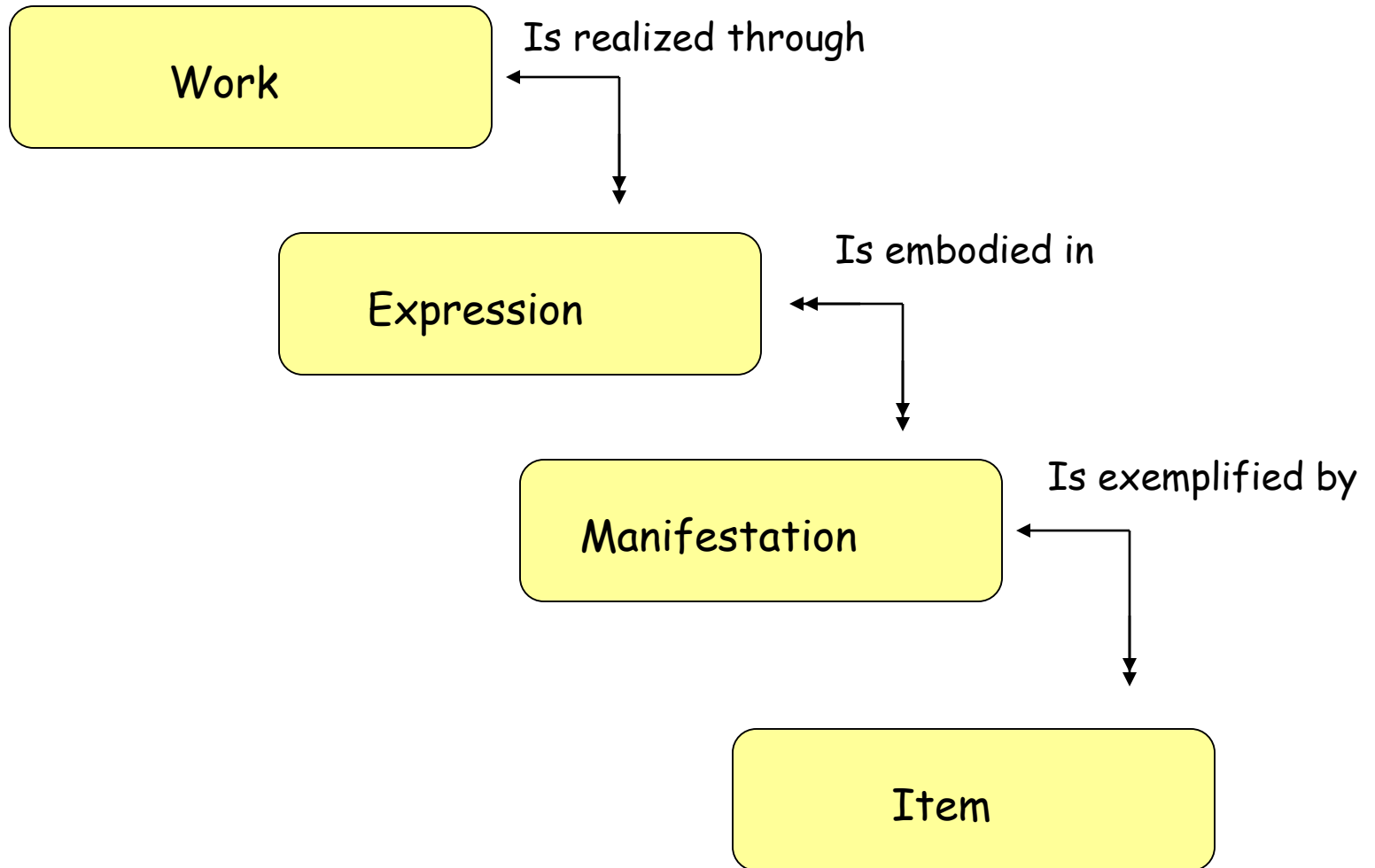Three different types of entities are defined: products, responsibilities and subjects.

FRBR entities

Group 1: the products of intellectual or artistic endeavour that are named or described in bibliographic records: *work*, *expression*, *manifestation*, and *item*.

Group 2: entities responsible for the intellectual or artistic content, the physical production and dissemination, or the custodianship of such products: *person* and *corporate body*.

Group 3: entities that serve as the subjects of intellectual or artistic endeavour: *concept*, *object*, *event*, and *place*.

Product entities

Manifestation level cataloguing

BIBSYS (and most other library catalogues) is based on manifestation level cataloguing.

An **object record** contains the 3 upper levels.

A **document record** for each item represents the lowest level.

There are unsufficient facilities for linking together the records that belong to the same work, and these belonging to the same expression.

Manifestation level cataloguing

Data elements are widely replicated, as work and expression specific data are stored for each manifestation.

The user interface becomes unnecessary complex, and the database vulnerable from a data integrity point of view.

Wassmo: works

Det stumme rommet
Dinas bok
Flotid
Hamsun : en prolog og en epistel
Hemmelig torsdag i treet
Hudløs himmel
Huset med den blinde glassveranda

Wassmo: works

Det stumme rommet
Dinas bok
Flotid
Hamsun : en prolog og en epistel
Hemmelig torsdag i treet
Hudløs himmel
Huset med den blinde glassveranda

Wassmo: expressions

Det stumme rom
Dinas bok
Flotid
Hamsun : en prolo
Hemmelig torsda
Hudløs himmel
Huset med den blinde glassveranda

De stille kamer [flamsk]
Der stille Raum [tysk]
Det stumma rummet [svensk]
Det stumme rommet [norsk]
Det stumme rum [dansk]
Hääletu ruum [estisk]
La chambre silencieuse [fransk]

Wassmo: expressions

Det stumme rom
Dinas bok
Flotid
Hamsun : en prolo
Hemmelig torsda
Hudløs himmel
Huset med den blinde glassveranda

De stille kamer [flamsk]
Der stille Raum [tysk]
Det stumma rummet [svensk]
Det stumme rommet [norsk]
Det stumme rum [dansk]
Hääletu ruum [estisk]
La chambre silencieuse [fransk]

Wassmo: manifestations

Det stumme rom
Dinas bok
Flotid
Hamsun : en prolo
Hemmelig torsda
Hudløs h
Huset m

De stille kamer [flamsk]
Der stille Raum [tysk]
Det stumma rummet [svensk]
Det stumme rommet [norsk]
Det stumme rum [dansk]
Hääletu ruum [estisk]
ansk]

Oslo : Gyldendal, 1983
Oslo : Gyldendal, 1984
[Stabekk] : Den norske bokklubben, 1984
Oslo : Gyldendal, 1990 : Norbok
Melhus : Lydbokforl., 1995
[Oslo] : Bokklubben nye bøker, 1996
Oslo : Gyldendal, 1997

Mapping FRBR to MARC

Attributes (and relations) in many cases appears in different MARC data elements - up to 5 different was found by Delsey.

Attributes for different levels sometimes are mixed / concatenated within one MARC data element.

Textual descriptions are expected in several data elements, where the use of codes might be more appropriate.

It's a MARC problem, not a FRBR problem ? ? ?

FRBR level	Attributes and relations	MARC fields
Work	Title	
	Content, form ...	008
	Date	
	Other separating characteristica	130 $q
	Continuing publication	008
	Responsibility relations	100/110/700 etc
	Subject relations	6xx
Expression	Title	245 $a, 241 $a
	Form, genre	
	Date	
	Language	008
	Other separating characteristica	130 $q, 250, 500
	Rights attributes	096 $f, 500
	Periodicity	
	Responsibility relations	100/110/700 etc

This is a quite inaccurate sketch of FRBR / MARC relationships!
See Delsey for a thorough mapping.

FRBR level	Attributes and relations	MARC fields
Manifestation	BIBSYS Record identifier	001
	Title	245 $a
	Edition	250 $a, 500
	Place of publication	260 $a
	Date	260 $c
	Series information	4xx, 022
	Physical form	245 $h, 008
	Extent	300
	Other identifiers	020, 024 etc
	"Electronic" attributes	500, 856
	Responsibility relations	100/110/700 etc
Item	BIBSYS Item identifier	
	Ownership	
	Location / holdings	
	Availability	
	Subscription data	

The ambitious plan

Build a physical and logical data model based on FRBR, with 4 different types af records, each with a set of unique identifiers.

Convert the existing database into the new model by automatic means.

Build a user interface that utilizes the model for what it's worth. (Collecting, separating, expanding, inheriting)

The ambitious plan

Make the model explicit to the cataloguers, and **educate them**.

It might be possible to keep the present user interface for cataloguing, hiding the FRBR structure for the cataloguers. But hardly wise.

Plan B: Just-in-time FRBR

Keep the data model and the principle of manifestation level cataloguing, but use just-in-time conversion to synthesize a FRBR-lookalike user interface. Rework the database indexes for this purpose.

Include as many linking fields as necessary.

Improve authority control and extend the use of coded data values.

Automatic FRBR extraction

Typical problems with automatic extraction of FRBR structures:

- Lack of role qualifier in 700
- Inconsistent treatment of original titles
- Lack of language codes
- Lack of authority control facilities

(Hegna & Murtomaa, 2002)

Hegna & Murtomaa:

The analysis shows that even though the information in the MARC records holds attributes relevant for identifying the work, expression and manifestation entities, the accuracy and formal syntax are too simple to be properly handled by programs.

March 2002.

BIBSYS goals

FRBR should be used as a framework for all BIBSYS thinking about bibliographic data.

The data model for periodicals should be based completely on the 4 FRBR levels.

This model should be explicitly visible in the user interface.

As a short term measure, new linking fields should be introduced to link different expressions of the same work and different manifestations of the same expression.

BIBSYS goals

BIBSYS-MARC should be extended by including nearly all "linking entry fields" in MARC 21.

Use of note fields should as far as possible be replaced by links.

One should consider if retrospective linking is possible, based on other field values, like "continuation links".

The user interface for searching should be based on the model, such that the user will see only one entry on each level, with the possibility to expand the lower levels.

Further reading

Knut Hegna and Eeva Murtomaa.

Data mining MARC to find: FRBR? March 2002.

http://folk.uio.no/knuthe/dok/frbr/

Tom Delsey.
Functional analysis of the MARC 21 Bibliographic and Holdings Formats. Technical report, Library of Congress, 2001.

http://www.loc.gov/marc/marc-functional-analysis/home.html

IFLA Study group on Functional requirements for bibliographic records. Final report. Technical report, IFLA, 1998.

http://www.ifla.org/VII/s13/frbr/frbr.pdf

Austlit. Australian Literature Gateway.
AustLit Data Models

http://www.austlit.edu.au/about/metadata