Towards a FRBR compliant catalogue: prospects for the City of Cape Town Public Library Management System

Niels Ruppelt
Collection Development Unit, Cataloguing Section, City of Cape Town Library and Information Services, Cape Town, South Africa.
E-mail address: Niels.ruppelt@capetown.gov.za

Copyright © 2015 by Niels Ruppelt. This work is made available under the terms of the Creative Commons Attribution 3.0 Unported License: http://creativecommons.org/licenses/by/3.0/

Abstract:

The City of Cape Town Library and Information Services (CCT LIS) currently operates the Brocade Library Management System (LMS), which uses cloud (web-based) software developed by the University of Antwerp in Belgium, and is distributed in South Africa by the State Information Technology Agency (SITA), with additional support from CiBLis (CIPAL Bibliotheca Library information system). Brocade makes use of the Apache Lucene search engine to search its catalogue.

CCT LIS operates 104 libraries which support a local population with an ever increasing literacy rate with growing information needs. The advent of the Internet and the emergence of new formats of items stocked in libraries have exposed a need for LMSes to restructure their online search results.

IFLA initiated the development of the Functional Requirements for Bibliographic Records (FRBR) conceptual model. Its work, expression, manifestation and item level entities, together with its relationship framework equip LMSes to restructure their catalogues.

For CCT LIS the retrospective manual FRBRization option would be too labour intensive and time consuming, while a retrospective programmatic FRBRization process could prove very costly. Also, CCT LIS has a unique relationship with the Western Cape Library and Archive Services (WCL&AS). The former is an independent municipal institution which cooperates with the latter Provincial institution with regard to catalogue records, which could have further implications for the FRBRization process.

Brocade in its current configuration does allow for the implementation of the FRBR conceptual model for newly acquired works. By utilising an existing cataloguing option currently available on Brocade, cataloguing can be done on an expression level leading down to the manifestation and item levels.

The degree of FRBRization of the CCT LIS catalogue is dependent on software vendor support and the relationship with WCL&AS from where a sizable number of catalogue records are imported programmatically.
1. Introduction

1.1 Brocade Library Management System (LMS)

The City of Cape Town Library and Information Services (CCT LIS) currently operates the Brocade Library Management System (LMS), also known as SITA Library Information Management System; SITA LIMS; SLIMS; or Brocade Library Services.

Brocade is a LMS using cloud (web-based) software developed by the University of Antwerp and was launched in Belgium in 2002, where it is supported and distributed by CIPAL (Centre of Informatics in the Province of Antwerp and Limburg). Kirsch (2014) explains that “in cloud computing, you the client therefore interface through a web browser (Internet) to a SaaS (Software as a Service) service provider (Brocade) via a computer, laptop, PDA, mobile device, tablet, etc. Brocade web application is a communication between user and library system via a standard web browser (Internet Explorer, Mozilla Firefox, Google Chrome, etc.).”

In 2007, the South African Provincial Library Services together with the State Information Technology Agency (SITA) made the decision to migrate from the then current PALS
SITA’s choice for this replacement LMS was Brocade, and from December 2008 they became South Africa’s domestic distributor of the new LMS, supported by CIPAL and subsequently, from May 2013 by CiBLis (CIPAL Bibliotheca Library information system), a spinoff of CIPAL Brocade Library Services.

1.2 A unique situation

CCT LIS was in the rather unique situation where it concurrently operated two separate LMSes, namely PALS and BookPlus (a command driven LMS developed in Australia), a consequence of the reorganisation of metropolitan municipal boundaries in the post-1994 period. The result was that CCT LIS now manages 104 libraries and had two LMSes to migrate to Brocade as well as continuing a complex relationship with the Western Cape Library and Archive Services (WCL&AS).

The WCL&AS comprises 331 libraries of which 104 libraries fall under the jurisdiction of CCT LIS. Library materials are supplied to these 104 libraries by both WCL&AS and CCT LIS. The complication arises from the fact that both CCT LIS and WCL&AS operate separate independent Brocade LMSes and databases, with WCL&AS catalogue records regularly and programmatically migrating to the CCT LIS Brocade LMS. All these factors will have interesting consequences for the possible future FRBRization of the CCT LIS LMS (see chapter 10).

1.3 The Online Public Access Catalogue (OPAC)

Brocade makes use of Apache Lucene (Query Parser Syntax) as the search engine to search its catalogue. Emma Bezuidenhout, ICT Officer of CCT LIS, describes Lucene as “similar to Google and contains the familiar, simple and advanced searches and serves to help patrons obtain search results of their choice. It is a technology suitable for nearly any application that requires a full text search, especially cross platform (that is why CiBLis uses it, as it is compatible with the Brocade LMS)” (internal email communication).

Interestingly, Fitch (2006) notes that on Lucene “related records are grouped into a FRBR-like structure”. This will be fully investigated in chapter 7.

A huge plus point for the Brocade LMS / Lucene search engine combination is pointed out by Odendaal (2009) who argues that “… due to the web interface, users can search the OPAC with ease from their library (from a public computer) or from their home via the Internet.”

1.4 The user community of CCT LIS

The South African National Census of 2011 found the population of Cape Town to be 3 740 026 with 24.8% of the population being under 15, while 65 year-olds and older make up 5.5%.

These people live in 1 068 573 households with 78.4% being formal structures (houses and apartment buildings), while 20.5% reside in informal structures (shacks). Electricity for lighting is used by 94% of households.
46.6% of the population have at least a grade 12 education. In the 5-25 age group, 67.8% attend an educational institution either at one of the approximately 380 government primary or secondary schools, 55 private schools, roughly 68 colleges (offering academic, commercial and technical qualifications) or 3 public universities.

There are currently 996 589 people in possession of a library card, of which 266 302 were active in 2014. A 2013 CCT LIS survey of 9 176 adult users found that 80.07% take out books, 29.78% DVDs and 18.18% music CDs. 88.98% reported that they found what they were looking for in their library. Correspondingly, of the 4 045 children surveyed, 84.80% borrowed books, 31.92% DVDs, 21.58% music CDs, while 88.83% successfully found what they were looking for.

33.26% of the adults surveyed use the Internet, compared to 35.07% of the children surveyed. The Census 2011 municipal report: Western Cape established that 37.9% of the Cape Town population possessed a computer and 49.3% accessed the Internet by means of a computer or cell phone.

2. What the user expects from an OPAC search

The National Library of Australia (NLA) contends that “users expect to be able to identify a particular item of intellectual or artistic content”. In so doing, they need to be able to confirm that the catalogue record they have found represents the item they are looking for or that they can distinguish between the catalogue records of two or more items with the same titles (Presentation 3 of RDA background information sessions - NLA). Also, users want to identify the works of prolific authors “that exist in many versions and/or formats” (Croissant 2012 : 15).

Teixeira (2010 : 94) notes that there is “a great change taking place concerning users’ behaviours. Library end-users became expert [at] using search engines and other Internet resources and were demanding new capabilities of online systems. Ranking, organising and clustering were the features they got used [to] in Google and other search engines”. The search results library users are currently faced with in online OPACs are “long lists of unstructured and unrelated records”. Teixeira goes on to argue that “technological advances caused losses of bibliographic structure, absence of hierarchical display and use of relationships. The flat structure of current [LMS] databases cannot provide a hierarchical display of search results.” (2010 : 18) This is indeed also the case with CCT LIS’ current Brocade / Lucene LMS (see chapter 7).

3. The What and the Why of FRBR

By the late 1980s it was apparent to professionals in the library world that major changes were developing in the library environment (Teixeira 2010 : 12). The emergence of the Internet brought along with it network access and new web resources; library catalogues were becoming electronically automated; new formats of library materials were being developed and electronic publishing was becoming a real future reality. The net result was that the existing catalogue components of LMSes began to lose their structure and hierarchical organisation. These concerns led to the 1990 Stockholm Seminar on Bibliographic Records, sponsored by the IFLA Division of Bibliographic Control and by the IFLA Universal Bibliographic Control and International MARC (UBCIM) Programme. Out of the nine
resolutions adopted in Stockholm one specifically led to a study to define the Functional Requirements for Bibliographic Records - **FRBR**.

Croissant (2012 : 6) explains that “FRBR (pronounced ‘furbur’) is an attempt to develop a conceptual model that can express common international understanding of what bibliographic records should be and what they should be expected to accomplish”.

In being a conceptual model, as opposed to a data model, vastly different FRBRized systems have been implemented (Teixeira 2010 : 25) which indicates that “cataloguing rules and online catalogues can be based on an understanding and interpretation of the FRBR model” (Presentation 3 of RDA background information sessions - NLA).

### 4. The FRBR entities

Group 1 is made up of the “bibliographic” entities known as Work, Expression, Manifestation and Item. The second group contains the “people” entities - individual persons and corporate bodies. The “subject” entities make up the third group, being concepts, events or places as well as any entities from groups one and two.

The four bibliographic entities from group one stand at the core of the FRBR model Croissant (2012 : 8), and it is hoped that “eventually library systems and OPACs will evolve to take full advantage of the data created using RDA [Resource Description and Access - the new cataloguing guidelines gradually replacing the AACR2 cataloguing rules] with its underlying FRBR structure of **Work, Expression, Manifestation** and **Item**” (Presentation 3 of RDA background information sessions - NLA). In this regard, Croissant notes the increased usage of the acronym **WEMI** in FRBR when referring to the four group one entities, namely **Work, Expression, Manifestation** and **Item** (2012 : 8).

“A **Work** has no material presence and can only be recognised through realisations and expressions of the work” (Presentation 3 of RDA background information sessions - NLA). An **Expression** is the realisation of a Work in various formats, e.g. the printed word, sound recording, motion picture, etc. The attributes of the **Manifestation** are what we would find in a typical catalogue record, such as statement of responsibility, edition statement, publisher, etc. Finally, an **Item** is the physical object. Its attributes being, e.g., the barcode and spine mark the library has pasted on it.

### 5. CCT LIS and the FRBR WEMI entities

The question is: do the catalogue records in the CCT LMS embody any or all of the four WEMI entities? Below is an OPAC view of a catalogue record.
Identifying the WEMI entities within the catalogue record via a colour code:
We can see that the WEMI entities are represented. However, just like Teixeira generally observed (2010: 96), CCT LIS catalogue records also have all four WEMI entities in one record, whereas the new FRBR model is based on the WEMI entities each having their own catalogue record entries. To FRBRize this catalogue record, it needs to be split into four records which should be accessible at the different FRBR retrieval levels (see chapter 9).
6. Relationships

Having confirmed that the WEMI entities exist in the current CCT LIS catalogue records, the potential exists for the next step of the FRBR entity-relationship model to be applied, which Croissant (2012: 10) describes as “defining the relationships that can exist between any two bibliographic entities, as well as the relationships that can exist between a bibliographic entity (a group 1 entity) and any group 2 entity (a person or corporate body”).

Tillet (2005: 4) refers to this relationship model as content relationships that “can be viewed as a continuum from Works / Expressions / Manifestations / Items.” It can also be described as a “Family of works” (Presentation 3 of RDA background information sessions - NLA).

The continuum moves from left to right and begins with the original work and related works that are considered Equivalent, e.g. reprints and facsimiles.

Next come the works / expressions / manifestations related via a Derivative relationship. These comprise of a range of new expressions, e.g. translations, motion pictures, audiobooks.

Finally, as the continuum moves to the far right, we find Descriptive relationships that involve new works, e.g. a literary criticism and analysis.

7. The current search results in CCT LIS OPAC

A FRBR compliant catalogue should deliver a search result that would bring expressions of a work together, display the various relationships and enable the user to navigate to the required level of interest (Hickey and Vizine-Goetz 2002: 4).

The results for a search of the work Gone with the wind on CCT LIS OPAC: …
The search resulted in 128 bibliographic records, which included various expressions including books, a translation, DVDs and CDs. Also retrieved were two websites specially created for this study proving that Brocade is a good vehicle for guiding users to quality websites. By clicking on the link for the website in the OPAC view of the bibliographic record, the user is sent straight to the website.

<table>
<thead>
<tr>
<th>Publication</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Title</strong></td>
</tr>
<tr>
<td><strong>Note</strong></td>
</tr>
<tr>
<td><strong>E-info</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subject Headings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Geographic Name</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Available in the following libraries</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTL-Internet</td>
</tr>
<tr>
<td>Electronic resources - Online database adult - -</td>
</tr>
</tbody>
</table>

Identification       c:dt:1036583
Already, the OPAC search result displays some FRBR features. Firstly, the use of icons serves, as Tillet notes (2005: 6), as “an additional means for conveying information about the mode of expression”. The problem with Brocade, in this regard, is that not all modes of expression display an icon as yet, e.g., expressions and manifestations for websites. Also, books display a cover image only if Brocade is able to match the ISBN of the manifestation to a cover that is available for automated downloading from the Internet. If no cover exists no generic icon is generated. Secondly, the search can be refined by options available on the right hand side of the screen by Author, Branch (library), Language, Collection (membership / genre), Format (media type / carrier / expression) and Year (of publication). This feature is also available on WorldCat, which Teixeira (2010: 27) considers currently to be one of the best examples of a FRBR compliant catalogue (see chapter 10).

When refining the search by author, in this case Margaret Mitchell, 25 out of the 128 catalogue records originally retrieved are displayed, even though the refinement option indicates 26 records being available for Margaret Mitchell. The reason for this discrepancy is that one of the records has the author entered twice in the Author field. Teixeira (2010: 43) warns that “FRBR exposes errors and cataloguing inconsistencies which are hidden” and “that inconsistencies in the bibliographic records … [are] a serious impediment to identifying expressions” (see chapter 9).

With only 25 of the 128 original Gone with the wind hits being written by Margaret Mitchell, the remaining 102 search results could be “noise” as Teixeira calls records that have the words making up the desired title somewhere in the catalogue record (2010: 75). That was also the case in this search result with many catalogue records featuring the jazz super hit Gone with the wind which has no relationship to Mitchell’s work. Alarmingly, the search
also put together a combination of the words **gone, with, the, wind** randomly from different songs in the contents note of an unrelated music DVD.

Scrolling further through the list of 128, it was found that 27 records that had a relationship with Mitchell’s work were lost. These relationships were either *Derivative* or *Descriptive* (see chapter 6).

The 27 records that were lost with the author refinement search had the words *Gone with the wind* in the title or note fields. However, catalogue records for a derivative work, such as *Rhett Butler’s people* by Donald McCaig which did not have the title represented anywhere in the records would also be lost in a *Gone with the wind* title search - two out of four *Rhett Butler’s people* records were lost for this reason.

### 8. Keeping it all together - possible options

The FRBR model does allow for a form of computer manipulation regarding a “related work” situation (Croissant 2012 : 14-15). For example, when creating a relationship between Mitchell’s work, *Gone with the wind*, and an expression, such as a motion picture adaptation, “the two records could be connected by a link with a defined value: HasAdaptation/IsAdaptationOf. Such a combination of entity records and links can be understood and used by a computer”. This process would require input and intervention by the LMS software developer and would incur extra costs.

Croissant has also suggested the use of what is called a “persistent identifier” (2012 : 13) and goes on to explain that “a record for a particular entry, whether from Group 1 or Group 2, can be assigned a persistent identifier. From then on, there is a stable one-to-one relationship between the entity and its identifier … Each entity record can be linked with others using a controlled set of relationships, each of which also has its own persistent identifier”.

A suitable, practical way of interpreting the concept of the persistent identifier for CCT LIS is on the broad relationship level. For the vision of the now-retired CCT LIS Collection Development Officer, Brenda Kyle, to be realised, where a user searching for information on District Six, a historical Cape Town apartheid-era suburb, the “catalogue would return results for books, DVDs, music CDs, links to articles and images in the [in-house] pamphlet collection and [Internet] links to, e.g., the District 6 Museum, SA History Online, Cape Tourism, etc.”, a persistent identifier could be inserted into each catalogue record that involves District Six - no matter what the expression, fact or fiction.
For the purposes of this study, a persistent identifier, *District Six IFLA*, was created (with *District 6 IFLA* as a SEE reference) and inserted into catalogue records. This could also be equated to RDA’s “preferred title” concept, which replaces the Uniform title of AACR2. The search result successfully retrieved the 32 records into which the persistent identifier was inserted.

However, this will involve a huge amount of retrospective work to bring together all the current records that have relationships - certainly a labour intensive exercise for which CCT LIS does not have the staff and resources.

9. **FRBRization of existing catalogue records**

FRBR requires a “decentralised” or “distributed” data structure based on relationships and on the entities, implying that a catalogue record should be created for each of the Group 1 WEMI entities (Croissant 2012 : 4).

This FRBRization process of a catalogue can be undertaken manually or programmatically.

### 9.1 Manual FRBRization

Teixeira attempted a manual FRBRization of a work on the Portuguese Union Catalogue, PORBASE (2010 : 78-81). The work chosen was *Os Maias* by Eça de Queirós. The work was represented by various expressions, large amounts of manifestations and “a relevant number of new works about the work *Os Maias*”.

The initial title search retrieved 191 records. This was reduced to 114 after an author/title combination search refinement, removing “noise” records. These records were then exported...
to an Excel spreadsheet. In each record the WEMI entities were identified. Once this had been achieved, the respective clusters (groups) were created. The manual groupings were converted to html lists of bibliographic control numbers, i.e. catalogue record numbers, which represented clusters according to WEMI entities.

Teixeira concluded that the manual FRBRization was “reasonably accomplished”. Complications encountered included difficulty in clearly distinguishing the WEMI entities from one another and that cataloguing inconsistencies and errors hampered efforts to retrieve the complete required set of records for FRBRization.

9.2 Programmatic FRBRization

For CCT LIS initiating an arduous, time-consuming, manual FRBRization process is out of the question, according to the Collection Development Specialist, Akeela Gaibie, who heads the Collection Development Unit of which the cataloguing section forms a part.

The only viable option for the department therefore is an automatic FRBRization process. The major challenge in this respect is “to find an algorithm that is effective with less [than] perfect data” (Teixeira 2010 : 48). A computer program has to be written to tell a computer what to do, step by step. This is called an algorithm (Computer.howstuffworks.com).

A FRBR Work-set algorithm has been developed by an OCLC (Online Catalogue of the Library of Congress) research group. Teixeira explains that the algorithm is based on a simple concept “that all manifestations of a given work share a common author and title, therefore a combination of the author and title should be enough to identify all manifestations” (2010 : 37). The CCT LIS OPAC does reveal elements of this algorithm in its search results.

10. FRBRization - the reality of the now

WorldCat can be used as a standard against which to measure the degree of FRBRization of a catalogue. The user can choose from the manifestation records for which expression they wish to see a more detailed search result. They can click on a hyperlink, View all forms and languages, to view more manifestations of the chosen expression (a sort of reverse FRBRization).
Brocade, in its current configuration, allows for the implementation and interpretation of the FRBR conceptual model. This can be done by utilizing the **Cataloguing component parts** cataloguing option. The idea is to have one main record and link related records to it.

Firstly, one creates a new catalogue record - a record for the expression, e.g. for a fiction novel. The mode of expression is indicated in [ ] after the title.

The way Brocade is currently configured, a location genre has to be linked to this catalogue record for it to link to the Lucene OPAC. CiBLis would need to create a generic expression genre in this case.

Once the expression-level record has been created, the cataloguer will catalogue the newly acquired manifestation record in the **Cataloguing Component Parts** option. The expression record is inserted in the **Specify catalogue entrie** box and a 1 is entered into the **Number of parts** box for one manifestation record to be created.

The next intervention required from CiBLis is to create more options in the **Data for the parts** drop-down list. Only the CD option is available.
On submission, the title is made available, via which the full manifestation record can be accessed to catalogue the remaining data.

Once the rest of the data is entered, the OPAC view of the expression with its attached manifestation looks like this:

By clicking on the manifestation title in the Contents in the expression record, the user opens to the item level of the manifestation catalogue record.
As different manifestations arrive over time, these can be linked by the cataloguer to the same expression record by following the same procedure as above.

A search for the work, *Titanium sun* in the OPAC gives this result: a rather jumbled list of expression records mixed with manifestation records.
CiBLis will need to adjust the parameters to display only the expression level records as in the following manually manipulated search result for *Another red flag*.

For the cataloguer, a small change is required in procedure. Once a work is received for cataloguing, a catalogue search must be done to ascertain if an expression-level bibliographic record already exists.
Adopting an in-house created FRBRization process such as this one would only be viable if the same FRBRization process is adopted by WCL&AS on their Brocade server (see chapter 1.2 for explanations).

11. Conclusion

The CCT LIS’s growing collection serves a user community with increasing information needs. With improving Internet access city-wide, these users are becoming familiar with ordered Internet search results, as opposed to the unstructured search result displays of online LMSes.

The CCT LIS catalogue, with its many expressions and manifestations of works is an ideal candidate for FRBRization. Retrospective manual FRBRization, with its identification and separation of the WEMI entities within one catalogue record, is far too labour-intensive and time-consuming to be considered. Retrospective programmatic FRBRization is the only viable option available at the moment, despite this requiring close cooperation between the software vendor and customer when trying to identify a suitable computer algorithm to carry out the process and the extra cost this would incur for the customer.

The way Brocade is currently configured allows for a measure of FRBRization when adding newly acquired works to the catalogue. By creating expression level catalogue records for a work and linking manifestation / item level records to these, a degree of FRBRization can be achieved. Software vendor intervention would be minimal, e.g., the addition of extra icons and drop-down menu options, thereby keeping extra costs in reasonable check.

However, CCT LIS’s unique relationship with WCL&AS (e.g. importing new catalogue records programmatically from WCL&AS) will be the determining factor as to the level of FRBRization to be adopted, as both LMSes have to display identical levels of FRBR.
Acknowledgments
Thank you to…

Lynne Burger, my cataloguing colleague, for typing the document and laying out the images, as well as her meticulous proof-reading

Esther Meyer, another cataloguing colleague, for assisting with creating dummy catalogue records

Emma Bezuidenhout, CCT LIS ICT officer, for information regarding Brocade and Lucene

Akeela Gaibie, CCT LIS Collection Development Specialist, and Gretchen Smith, Knowlead Consulting and Training Director, for editorial advice

The members of the CCT LIS cataloging team at the Collection Development Unit for their input and support

References

Access points for works and expressions in Generation RDA : IGBIS workshop, 5-6 August 2013, Centre for the Book, NLSA, Cape Town.


CROISSANT, Charles R. FRBR and RDA : what they are and how they may affect the futures of libraries in Theological librarianship : an online journal of the American Theological Library Association. Volume 5, number 2 (July 2012).


http://staff.oclc.org/~vizine/CNI/OCLCFRBR_files/frame.htm


Appendix: Practicalities

1. FRBR allows search results to be displayed in a structured way.

2. Catalogues containing works that display many expressions and manifestations are ideal FRBRization candidates.

3. FRBR is a conceptual model and therefore allows different interpretations of the model to be developed to suit each individual catalogue’s needs and requirements.

4. Retrospective manual FRBRization of current catalogue records is a too labour-intensive, time-consuming process to be practically viable.

5. Retrospective programmatic FRBRization of current catalogue records requires cooperation between the software supplier and the customer to create a suitable computer algorithm, which will incur extra expense to the customer.

6. The current configuration of the Brocade LMS allows a degree of FRBRization by cataloguing newly acquired works in one of the available cataloguing options.