Simon Edwards, Gordon Dunsire, Alan Danskin

RDA international

RDA has always been a continually evolving standard that aims to reflect the requirements of the cataloguing community. Simon Edwards will highlight the steps towards a further internationalisation and exploration of wider cultural heritage description communities. Gordon Dunsire will point out the potential implementations of RDA data in various database structures and describes the possibilities for further work with RDA and linked data scenarios in international communities. The third part of the article reflects the experience of the British Library in applying RDA in the last years. Alan Danskin gives an overview of the transition period from the project organization until training.

RDA, als ein von Beginn an auf Veränderung konzipierter Standard, hat das Ziel, sich den Anforderungen der Erschließung anzupassen. Simon Edwards beschreibt vor diesem Hintergrund die anstehenden Veränderungen hinsichtlich der Internationalisierung des Standards und der Einbeziehung weiterer Kultureinrichtungen. Gordon Dunsire zeigt das Potenzial von nach RDA erfassten Daten in verschiedenen Datenbankstrukturen auf und skizziert die Möglichkeiten der Weiterarbeit mit RDA in Linked-Data-Szenarien in einem internationalen Umfeld. Der dritte Teil des Beitrags schildert die Erfahrungen der British Library im Umgang mit RDA in den letzten Jahren. Alan Danskin gibt darin einen Überblick über die Implementierungsphase von der Projektorganisation bis zu den Schulungen.

Ensuring the further internationalisation of RDA through governing structures (Simon Edwards)

Background

RDA has always been a continually evolving standard that aims to reflect the requirements of the cataloguing, metadata and description community. As we continue to develop the standard it is essential that we broaden the range of perspectives applied to the development of the standard to ensure that it reflects a wide range of different cultural perspectives. The benefit of doing this was envisioned as RDA was developed. As more organisations develop rich and compatible data sets about their holdings, these can be actively shared across the globe to open up and increase the discoverability of collections for the benefit of the users we serve.

RDA has now reached a critical point in this development and the key to its continued success is a firm commitment to further internationalisation and exploration of wider cultural heritage description communities. The Committee of Principals (CoP) have agreed on a new governance model to which it will begin to move over the next 3–4 years that emphasises this need for wider representation. This article highlights the steps that have and will be taken to move towards this model.

Vision and Strategy

At their 2014 meeting the CoP created a new vision and agreed four Strategic Priorities that formed the basis of their Strategic Plan 2015–2020.

Vision: RDA: the global standard enabling discovery of content.

Strategic priorities 2015–2020:

1. Make RDA an internationally recognised standard.
2. Increase the adoption of RDA internationally.
3. Develop a sustainable business model.
4. Develop a relevant governance structure.

To ensure that the vision is realised and strategic priorities are delivered, the CoP undertook a review of existing governance structures. The review consisted of desk research and wide ranging consultation with key stakeholders. This identified a number of key themes:

- Strong support for the new Vision and Strategy.
- Internationalisation was the key priority highlighted and this must be reflected across the entire governance structure.
- The CoP will need to assess each of the wider cultural heritage communities to identify whether the likely outcome is adoption or alignment.
- The term ›constituencies‹ has both positive and negative connotations.
- There is support for leveraging international communities but the CoP needs to consider how closely they are able to align.
- There is a need to ensure that continuity within the Joint Steering Committee (JSC) is balanced with expertise.
- Working groups are important for getting the work done and for succession planning.
- The need for a coordinated RDA outreach programme.
- The JSC should consult wider stakeholders on the work plan and publish it.
- There is a need to manage the workload of the JSC.

New Governance Model

A new model has been created to reflect the key areas of feedback from the consultation; particularly reflecting internationalisation. This has been designed within the principles agreed by the CoP in 2014. As a result, both the CoP and the JSC will move from specific constituents to representation from the 6 inter-

Foto: CILIP

Simon Edwards

Foto: privat

Gordon Dunsire

Foto: British Library

Alan Danskin

strategic priorities

national regions identified by the United Nations (Africa, Latin America and the Caribbean, North America, Asia, Europe, Oceania). To qualify for selection, the representative must work for an institution that has implemented RDA.

RDA Board

The CoP will be renamed and will consist of:

- 3 Association Representatives (ALA, CILIP, CLA) to represent the interests of the co-owners and co-publishers.
- Chair of the RDA Steering Committee as liaison and adviser to the Board.
- 6 National Institution Representatives who will be selected from each of the international regions (these will rotate and serve a 3 year term).
- 2 co-opted members will be chosen by the Board based on communities they would like to see represented. In the first instance the CoP is considering a representative from the Archives community.
- A representative from ALA Publishing.

new structure in place by 2019

RDA Steering Committee

The JSC will be renamed and will consist of:

- 6 Regional Community Representatives nominated by appropriate bodies representing each of the international regions.
- 5 Rotating Representatives each of whom will serve for a 4 year term.
  - Chair of the Steering Committee (leadership of the committee and liaison with the RDA Board).
  - Examples Editor (responsible for RDA Examples).
  - Technical Liaison (responsible for liaising with the Technical working group).
  - Translations Team Liaison (responsible for facilitating learning from the Translations teams and representing their wider views).
  - Wider Community Engagement (responsible for engaging with user groups on the development plan).
- A representative from ALA Publishing and the Chair of the RDA Board will be in attendance.

They will also be supported by the paid post of the Steering Committee Secretary. The RDA Steering Committee will continue to use working groups to support the development of RDA; providing essential opportunities for a wider group of institutions and representatives to get involved.

working groups to support the development of RDA

Implementing the new Governance Model

The CoP has agreed a 4–5 year programme of activity to enable the current structure to move, over time, towards the new governing structure. The key principles governing the transition include:

- Current key stakeholders will be actively engaged in the change process and will be tasked with helping to shape future structures and support mechanisms to ensure they are fit for purpose.
- We will carefully balance retaining current skills and expertise and bringing new skills and representation on board in order to ensure that the stability of RDA as a tool does not suffer.
- The new structure will evolve gradually rather than one big change.
- The committee aims to have the new structure in place by 2019.

In the first instance, stakeholders will see a rebranding of the infrastructure and changes to wording used to describe supporting structures. This will be officially launched at the JSC meeting in November 2015.

The CoP assures current representatives that the new model continues to offer them multiple ways to remain influential in the development of RDA at Board, Steering Committee and working group levels. It is absolutely critical that current representatives continue to remain involved; their expertise in the development of RDA is essential for its continued success.

Over the next year, the RDA Board will:

- Make contact with current constituents to discuss the impact of Governance changes, the work that needs to be done to prepare for these changes and timescales involved.
- Develop proposals for the selection of National Institution Representatives on the Board.
- Develop new governance documentation.

Over the next year, the RDA Steering Committee will:

- Consider how working processes can be streamlined.
- Consider what new supporting structures are required.
- Develop new governance documentation.

The new RDA Board will use their annual report back from their meeting to report on progress to stakeholders.

RDA data in an international environment (Gordon Dunsire)

Background

»RDA is a package of data elements, guidelines, and instructions for creating library and cultural heritage resource metadata that are well-formed according to international models for user-focused linked data applications.«[1]

RDA: Resource Description and Access has focused on the production of data describing library and related resources from an early stage in its development. In 2007 the Joint Steering Committee for Development of RDA (JSC) discussed a paper by the RDA Editor on encoding RDA data.[2] It presented a roadmap for the creation of an element set and application profile that specify a schema for RDA metadata. It discusses the use of a proxy syntax such as MARC 21 for encoding RDA data until the RDA metadata element set was created and registered. At the same time, the RDA Editor prepared a paper on RDA Database Implementation Scenarios to illustrate some of the potential implementations of RDA data in various database structures. The paper was updated in 2009, and presented three scenarios.[3]

Scenario 1 covers a relational or object-oriented database structure that mirrors the FRBR and FRAD conceptual models on which RDA is based. The structure assigns a table or record to each primary entity in the models, in this case *Work*, *Expression*, *Manifestation*, and *Item* from FRBR, and *Person*, *Family*, and *Corporate Body* from FRAD. The scenario makes a distinction between descriptive and access point data for a resource, with descriptive data stored with the appropriate FRBR entity, and access point data stored with the FRAD entities. Relationship data are stored as links between the entities.

Scenario 2 covers linked bibliographic and authority records that reflect a »conventional« structure that assigns a table or record to a *Resource* entity and to each authority entity such as *Name* and *Name/Title*. Access point data are duplicated from the authority records for display in the bibliographic record, but the bibliographic record is also linked to the authority records for additional access points. This scenario is implemented in many current cataloguing management systems.

Scenario 3 covers a »flat file« database structure. It is essentially the same as Scenario 2, but with no data links between the bibliographic description and the authority records. This scenario is typically found in a card catalogue, where the user navigates from card to card by manually matching authorized access point data or »headings«.

The RDA instructions are intended to produce data that are compatible with these scenarios, and variations on them. The instructions therefore specify different ways of recording information about relationships between the entity being described and a related RDA entity. Four methods, a »four-fold path«, can be used to identify the related entity: unstructured description; structured description; authorized access point; identifier. An unstructured or structured description of the related entity can be embedded within the metadata for the local entity to accommodate Scenario 3. An authorized access point can be used in Scenario 2, and the use of identifiers is confined to Scenario 1.

Linked data scenario

a schema for RDA metadata

The implementation scenarios were developed before the meeting that took place in London in 2007 between members of the RDA and other metadata communities.[4] The meeting agreed that RDA should develop representations of its elements and vocabulary encoding schemes (VES) in Resource Description Framework (RDF) syntax for use by Semantic Web applications. This requires a new scenario for implementing RDA as linked data.

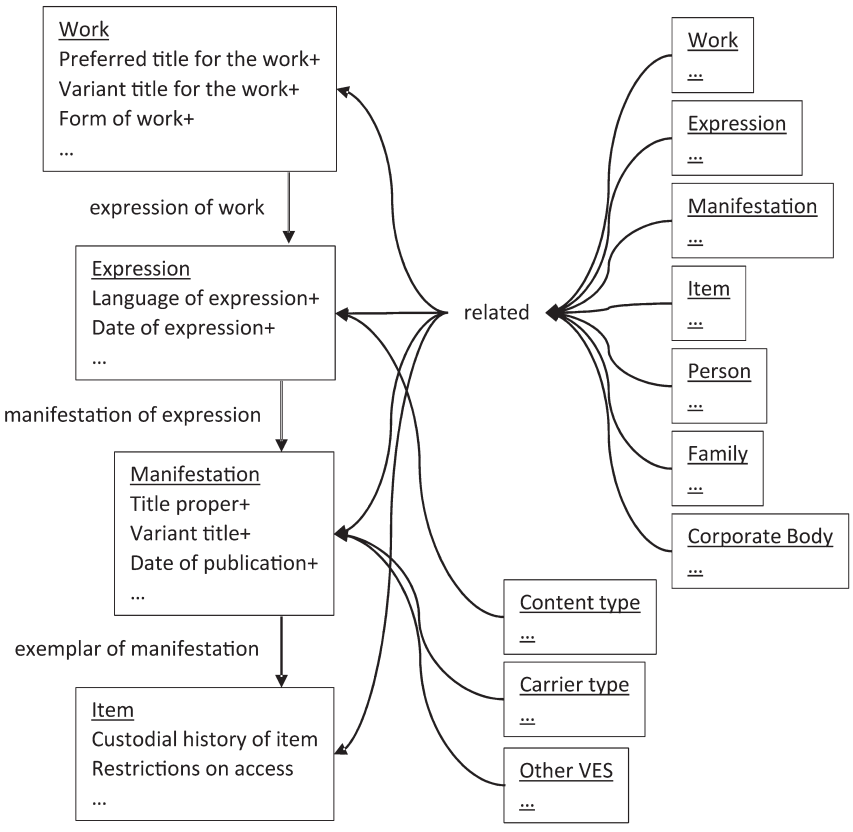

Figure 1: RDA implementation scenario for linked data

Figure 1 shows a basic linked data implementation scenario for RDA. It is a refinement of Scenario 1. Scenario 1 is in turn a refinement of Scenario 2, in turn a refinement of Scenario 3.[5] The linked data scenario can be considered therefore to be Scenario 0.

each entity has a unique machine-readable URI

Data are clustered around entities that are linked through relationships. Each entity has a unique machine-readable Uniform Resource Identifier (URI). There is no need to use any other method to record relationship data because it is built into the syntax of RDF, where data is stored as triples consisting of subject URI, predicate or property URI, and object URI or literal. Human-readable data must be stored in an object literal. In Figure 1, the data in each box are a cluster of triples with object literals and a common subject URI. A link is an RDF property between a subject and an object URI, with the arrow pointing to the object. Each property has an inverse, allowing the data to be linked in the opposite direction.

The RDA instructions specify a controlled terminology or VES for many elements, for example *carrier type* and *content type*. A VES is represented in RDF as a value vocabulary using the Simple Knowledge Organization System (SKOS) element set. Each term is assigned a URI, allowing it to be linked to the relevant entity instead of being embedded as a literal in the entity data cluster. This allows additional information about the term, for example a definition or scope note, to be linked to the entity using the same path. A significant advantage for RDA is that translated data can be treated in the same way, clustered around a common URI for the term.

»constrained« properties are mapped to the Unconstrained properties

In Figure 1, elements marked + are currently treated as literal values in RDA, but are likely to be treated as relationships with entities specified in the new consolidated Functional Requirements model, FRBR-LRM. These entities include *Nomen*, which encompasses titles, names, access points, and identifiers, *Place*, and *Timespan*, which covers dates.[6] The trend is towards greater dispersal of data between related entities, to a point where the assumption of a universal, standardized record becomes untenable. The »record« displayed in an application is local to that application and may be assembled from a larger pool of data that can support many variations in the content. It is possible for an entity to have no literal data in direct association »in the box«, just relationships to other entities that are linked to the human-readable data required by applications for display and browsing.

vocabularies in multiple RDF serializations

RDA Registry

A significant destination on the roadmap was reached in 2014 when the RDA Registry was implemented.[7] The main purpose of the Registry is to provide RDF representations of the RDA element sets and value vocabularies.

The RDA element sets contain an RDF class for each of the current RDA entities, and an RDF property for each RDA element and sub-element. An RDA element sub-type is represented as an RDF sub-property. Relationship designators are treated as RDF sub-properties of the high-level relationship element property. The result is a total of over one thousand RDF elements. This is inconvenient to manage as a single set, so it is divided into smaller element sets based on the domain of each property. There are element sets for properties associated with the *Work*, *Expression*, *Manifestation*, *Item*, *Person*, *Family*, and *Corporate Body* entities, as well as meta-properties and the RDF classes.

The Registry also offers an Unconstrained element set that parallels the RDA elements but does not make any reference to the RDA classes. That is, for each of the properties there is no domain or range, and a definition that refers to »resource« instead of Work, Expression, Manifestation, or Item, and to »agent« instead of Person, Family, or Corporate Body. This removes the FRBR and FRAD entities from the semantics of the properties, and allows the properties to be used in non-FRBR and non-RDA applications. The »constrained« properties are mapped to the Unconstrained properties using the RDFS sub-property relationship.

There is no special significance to the choice of element set scope, and it has to be flexible to accommodate some RDA elements. For example the inverse of *subject relationship* has no domain because *subject relationship* itself does not declare a range; the inverse property is included ad hoc in the Work element set rather than the Unconstrained element set.

The RDA VESs with complete definitions for all terms are represented as value vocabularies in the Registry. Terms, definitions, and scope notes are recorded in RDF/SKOS.

The Registry offers downloads of the latest versions of the element sets and value vocabularies in multiple RDF serializations, including RDF/XML and Turtle, as well as HTML, JSON-LD, and Microdata. The Registry website also provides examples of RDA linked data, and alignments and maps between RDA elements and concepts and the vocabularies used in communities.

Version control for all files is managed via the RDA-Vocabularies GitHub project.[8] This provides effective roll-back to any previous version of a file. Vocabulary releases are numbered using semantic versioning to warn applications if there is a break in seman-

tic coherency with previous releases.[9] This has been necessary already as a result of updates to the RDA instructions as well as correction of errors in the Registry itself. GitHub also provides a system for raising and tracking issues associated with the project, and this has also been very useful for improving the functionality of the Registry and its contents.

Semantics

RDA linked data is intended for use in semantic applications. RDA properties are generally assigned a domain for the RDA entity to which the RDA element applies. It is therefore not necessary to make an explicit declaration of the class or type of a data URI, because it can be inferred or entailed from the property used in a data triple.

The RDA Registry uses the RDF Syntax property *rdfs:subPropertyOf* to relate two RDA properties that have hierarchical semantics. This allows semantic »dumb-down« from a triple that uses the sub-property to an entailed triple that uses the super-property. For example, data that uses a relationship designator can generate data that uses the broader relationship element: the statement »this expression – has editor – that person« implies »this expression – has contributor – that person«. The same method is used to map an RDA property to an unconstrained property: »this expression – has editor – that person« implies »a thing – has editor – a thing«. This allows the RDA semantics, based on FRBR, to be stripped from the data so that they can be used in more general applications, for example where the first thing is explicitly declared to be a data-stream, and the second thing is declared to be a software program. This is de-FRBRization, and it can be completely automated.

International communities

RDA is intended to meet the needs of international communities using multiple languages and scripts. The Uniform Resource Identifiers (URIs) for RDA classes, properties, and concepts are opaque: They are based on arbitrary numbers to avoid language bias and the problems caused when lexical labels change, as has happened already with several RDA relationship designators. This is nevertheless inconvenient for application developers, so the Registry is developing the use of lexical aliases for the opaque, canonical URIs. For example, the compact URI *rdaw:P10145* has the English lexical alias *rdaw:absorbedByWork.en*. The Registry hopes to provide aliases in the other languages of full translation of RDA, including German, for example *rdaw:darinAufgegangenWerk.de*. A developer can use lexical aliases in application software because they can be automatically directed to the canonical URIs.

The RDA Registry is able to exploit the ability of SKOS to accommodate value vocabulary labels, definitions, and scope notes in multiple languages under a single URI. For example, the compact URI *rdact:1013* has the preferred label »spoken word« in English, and also the preferred label in German, »gesprochenes Wort«. This allows an application to switch between languages whenever required, giving it and its users better control over the lexical environment.

Open license

The Co-Publishers of RDA have defined »RDA Reference« to include all RDA elements, their definitions and any related scope notes, and all value vocabulary terms and definitions. This represents most of the information available in the Glossary and relationship designator appendices in RDA Toolkit. RDA Reference data are made freely available in the RDA Registry, and are intended for use by developers of applications using data conforming to the RDA instructions. The data are published on the RDA Registry under an open Creative Commons license that permits commercial re-use and requires attribution only.

RDA Reference data are made freely available

Example

The following listing is a Turtle (terse triple language) serialization of an RDF graph describing a published book. The data is derived from a MARC 21 record downloaded from the online catalogue of DNB and imported into RIMMF, an RDA data editor.[10] The RDF triples exported from RIMMF have been amended to show future possibilities. Line numbers are added for reference.

1. @prefix rdaa: <http://rdaregistry.info/Elements/a/> .
2. @prefix rdaco: <http://rdaregistry.info/termList/RDAContentType/> .
3. @prefix rdact: <http://rdaregistry.info/termList/RDACarrierType/> .
4. @prefix rdae: <http://rdaregistry.info/Elements/e/> .
5. @prefix rdam: <http://rdaregistry.info/Elements/m/> .
6. @prefix rdami: <http://rdaregistry.info/termList/modeIssue/> .
7. @prefix rdamt: <http://rdaregistry.info/termList/RDAMediaType/> .
8. @prefix rdata: <http://rimmfdata.com/r/> .
9. @prefix rdaw: <http://rdaregistry.info/Elements/w/> .

needs of international communities using multiple languages and scripts

```
10. @prefix skos: <http://www.w3.org/2004/02/
    skos/core#> .
11. rdata:zfbb00002524
12.   rdam:P30156 "Basiswissen RDA" ;
13.   rdam:P30142 "Einführung für deutschsprachige
      Anwender" ;
14.   rdam:P30105 "Heidrun Wiesenmüller und Silke
      Horny" ;
15.   rdam:P30111 "Berlin ; Boston, Mass. : De Gruyter
      Saur, 2015." ;
16.   rdam:P30088 "Berlin" ;
17.   rdam:P30088 "Boston, Mass." ;
18.   rdam:P30176 "De Gruyter Saur" ;
19.   rdam:P30011 "2015" ;
20.   rdam:P30157 "De-Gruyter-Saur-Studium" ;
21.   rdam:P30003 rdami:1001 ;
22.   rdam:P30004 "ISBN 9783110311464" ;
23.   rdam:P30004 "ISBN 3110311461" ;
24.   rdam:P30004 "(DE-599)DNB1029258813" ;
25.   rdam:P30004 "(OCoLC)827079430" ;
26.   rdam:P30002 rdamt:1007 ;
27.   rdam:P30001 rdact:1049 ;
28.   rdam:P30181 "xii, 300 Seiten@de" ;
29.   rdam:P30181 "xii, 300 pages@en" ;
30.   rdam:P30169 "28 cm" ;
31.   rdam:P30139 rdata:zfbb00002525 .
32. rdata:zfbb00002525
33.   rdae:P20001 rdaco:1020 ;
34.   rdae:P20214 "2015" ;
35.   rdae:P20006 "Deutsch"@de ;
36.   rdae:P20006 "German"@en ;
37.   rdae:P20231 rdata:zfbb00002526 ;
38.   rdae:P20059 rdata:zfbb00002524 .
39. rdata:zfbb00002526
40.   rdaw:P10223 "Basiswissen RDA" ;
41.   rdaw:P10219 "2015" ;
42.   rdaw:P10256 "DDC: 025.3 DE-101 22/ger" ;
43.   rdaw:P10256 "DDC: 020 DE-101 22sdnb" ;
44.   rdaw:P10256 "020 DE-101 sdnb" ;
45.   rdaw:P10256 "(BISAC Subject Heading)
      LAN025000" ;
46.   rdaw:P10061 rdata:zfbb00002528 ;
47.   rdaw:P10061 rdata:zfbb00002527 ;
48.   rdaw:P10078 rdata:zfbb00002525 .
49. rdata:zfbb00002528
50.   rdaa:P50117 "Wiesenmüller, Heidrun" ;
51.   rdaa:P50121 "1968" ;
52.   rdaa:P50195 rdata:zfbb00002526 .
53. rdata:zfbb00002527
54.   rdaa:P50117 "Horny, Silke" ;
55.   rdaa:P50195 rdata:zfbb00002526 .
56. rdaco:1020
57.   skos:prefLabel "Text"@de ;
58.   skos:prefLabel "text"@en ,
59. rdact:1049
60.   skos:prefLabel "Band"@de ;
61.   skos:prefLabel "volume"@en .
62. rdami:1001
63.   skos:prefLabel "Einzelne Einheit"@de ;
64.   skos:prefLabel "single unit"@en .
65. rdamt:1007
66.   skos:prefLabel "ohne Hilfsmittel zu benutzen"@
      de ;
67.   skos:prefLabel "unmediated"@en .
```

Lines 1–10 declare the QNames, abbreviations for http namespaces, used to shorten the URIs in the serialization. Lines 11–31 are triples with the manifestation as subject. Lines 32–38 are triples with the expression as subject, and lines 39–48 are triples about the work. Two agents related to the work are described in lines 49–52 and 53–55. Lines 56–67 are triples containing data from RDA VESs for the carrier type, content type, mode of issuance, and media type values related to the manifestation and expression.

The domain of every property in the RDA namespace »rdam« is *Manifestation*, so lines 12–31 entail *rdata:zfbb00002524 isA rdac:C10004*: »this thing – is a – manifestation«. Similarly, line 51 implies »this thing – is a – person« because the domain of *rdaa:P50121* is *Person*. This further implies »this thing – is an – agent« because *Person* is declared to be a sub-class of *Agent* in the Registry.

Line 46 states »this thing – has author – that thing«, which entails *rdata:zfbb00002526 rdaw:P10065 rdata:zfbb00002528*: »this thing – has creator – that thing«. This follows because *(has) author* is declared to be a sub-property of *(has) creator* in the Registry, representing the hierarchical relationship between the relationship designator and relationship element.

Line 15 duplicates data recorded in lines 16–19. Line 15 is an aggregated statement, composed from the values of sub-elements using a specific syntactical encoding scheme. The scheme in this case is International Standard Bibliographic Description (ISBD), used for examples in RDA Toolkit. There is no common method agreed in linked data communities for representing the structure of an aggregated statement in RDF.

Lines 56–58, etc. illustrate the multilingual data that can be accommodated in RDF value vocabularies. An application can select the @de (German) or @en (English) language qualifiers to filter the data selected for display without changing the data links. The RDA instructions allow data from other VESs that may be specified in local application profiles. Lines 28–29 and 35–36 also show parallel English and Ger-

man versions of the same data created by different cataloguing agencies. Lines 28–29 are statements containing data aggregated from finer sub-elements that are not completely specified in RDA. Lines 35–36 are examples of recording the term or preferred label from a VES as a direct data value, rather than as a linked URI for the concept from an RDF value vocabulary, similar to lines 56–58, etc. Both of these methods for recording multilingual data are extensions of the »four-fold path«.

It is trivial to link this data to data for a related resource. If the work is translated into English, then a link from the work to a new expression and manifestation is required:
rdata:zfbb00002526 rdaw:P10078 rdata:zfbbX .
rdata:zfbbX rdae:P20006 »Englisch«@de .
...

If the German expression is published in a new format, then a link from the expression to a new manifestation is required:
rdata:zfbb00002525 rdae:P20059 rdata:zfbbY .
rdata:zfbbY rdam:P30001 rdact:1018 .
...
rdact:1018 skos:prefLabel »Online-Ressource«@de .
...

If the work is made into a motion picture, »RDA: the movie!«, then a link can be made from the monograph work to the moving image work, and so on.

Future development

Many of the issues raised in this paper are under current investigation by the JSC and its working groups. Several papers submitted for discussion at the JSC meeting in November 2015 are related to these topics. The specific issues raised in lines 28–29 of the example are discussed in a paper submitted by the American Library Association.[11] The four-fold path and its extension to cover the linked data scenario are discussed in a paper by the JSC Technical Working Group.[12]

The impact of the consolidated FRBR-LRM model will be significant if it is followed by RDA. Elements marked + in Figure 1 will require augmented instructions and new element sub-types to allow the choice of paths to be extended to the new entities. In the example, the values of lines 16 and 17 may be replaced by URIs for a *Place*, line 18 may be replaced by the URI for a *Corporate Body*, and line 19 may be replaced by a URI for a *Timespan*. Lines 12, 40, 50, and 54 may be replaced by URIs for a *Nomen*. The boundary of the »record« in the example will remain consistent if the triple clusters containing the preferred labels are added:
[Replace 16.] rdam:P30088 placeVES:X ;
[Replace 17.] rdam:P30088 placeVES:Y ;
[Add] placeVES:X skos:prefLabel »Berlin (Germany)« .
[Add] placeVES:Y skos:prefLabel »Boston, Mass. (USA)« .

The ratio of data triples with literal object values to triples with URI object values will fall as more entities and VESs are added to the model and traditional attributes are replaced with relationships. In other words, the density of triple clusters will be diluted by linked triple chains better suited to international, global scale information services. This proliferation of machine-readable data is designed to exploit the processing power available in the web of machines, but it benefits resource description and access by providing fine-grained data for describing the rich and complex relationships found between cultural heritage resources.

development of RDA for authority data

The development of RDA for authority data in the context of FRBR-LRM is discussed in another paper by the JSC Technical Working Group.[13] The JSC Aggregates Working Group is investigating the modelling of aggregate resources in RDA, including anthologies, augmented expressions and manifestations, and serials. Preliminary results appear to establish compatibility between RDA and the published object-oriented FRBR model, FRBRoo, and opens a pathway for extending RDA for use by museum communities via the CIDOC-CRM.[14]

The JSC recently updated the RDA elements table.[15] The JSC is developing a basic application profile, specifying the elements expected in a »core record« and their repeatability, associated VES, and other information about the intended application of RDA data. This will provide part of the answer to the question raised by this paper, *What is an RDA record?*[16]

the consolidated FRBR-LRM model

The next stage on the original RDA roadmap, after the publication of the RDA element sets and application profiles, is to move beyond the encoding of RDA data in proxy syntaxes and develop services that can use pure RDA data to gain the maximum benefit from the application of the RDA instructions and the underlying models. RDA is not the only standard facing this stage of development; all library linked data communities are waiting for operational systems that can deliver the promise of RDF and the Semantic Web. Whatever path is taken, the RDA Registry will support as much technical interoperability and data convergence as possible.

RDA in the British Library (Alan Danskin)

Introduction

The British Library formally implemented RDA on 1st April, 2013. The original version of this article was written later that year for publication in an RDA themed issue of *Catalogue & Index*[17]. A shorter version was published in the proceedings of Bibliothekartag 103[18]. As well as updating the article, I have taken advantage of the perspective obtained by two years of experience in applying RDA to reflect on lessons learned and future opportunities.

development of RDA

Context

The British Library is the National Library of the United Kingdom. It is a legal deposit library, responsible for preserving the national printed archive. In 2014/15 Legal Deposit intake totalled 405,000 items.[19]

Legal Deposit	Actuals	Target
Monographs	107,554	103,525
Serial issues	192,719	174,515
Maps and atlases	1,905	1,227
Music scores	1,305	1,719
Newspaper issues	125,733	123,952
Play scripts	190	202
Serials Titles received	36,478	38,723

licensing the RDA Toolkit

The Library also acquires resources for the collection by purchase, exchange and donation. The Library has two main sites, located 200 miles apart, in London and West Yorkshire. The Library employs approximately 1500 staff.[20] The main processing teams are based in Boston Spa, but many specialist staff based in London also process collection items.

Prior to RDA, the Library used AACR2 for cataloguing printed resources; supplemented by DCRM and Cartographic Materials for specialist collections. Archival and manuscript materials are in a separate content stream and were ruled out of scope for RDA because traditionally we have followed archival standards for those streams. Sound Recordings and Chinese and Japanese collections were also out of scope because they are hosted on other systems and use different descriptive standards and schema.

expert groups

From the early 1990s, the Library's cataloguing strategy has been based on deriving or reusing catalogue records where possible, with minimal intervention. For this reason, the bibliographic standards we used have been closely aligned with Library of Congress.

Business Case

The Library is a partner in the development of RDA. The development of RDA is intended to satisfy short term and long term requirements. In the short term, RDA is superior to AACR2 with regard to description of non-print resources. RDA enhances access to all types of resources by providing comprehensive instructions for authority control and relationships. In the longer term, RDA is one strand of a strategy to update the infrastructure and enable libraries to exploit the potential of Web technology to expose the content of their collections.

It was estimated that implementation of RDA would result in a shortfall of 20–25k items on the Key Performance Indicator for collection processing. This was mitigated by accelerating production before training started and by outsourcing additional tranches.

We estimated that the annual cost of licensing the RDA Toolkit is greater than the annualised cost of replacing printed AACR2 every 5–7 years. We mitigated the cost, as much as possible, by taking advantage of deals offered by the publisher (2 for the price of one). A report available via the Toolkit enables subscribers to monitor the number of concurrent users and maximise the value for money from each license. Our initial estimate of the number of licences we would need was very conservative and we now manage with fewer than thirty.

The automatic update of RDA in the Toolkit means that cataloguers no longer have to manually annotate or paste amendments into their rules, as was the case with interim revisions of AACR2. On the other hand, the rate of change in RDA is much higher.

Project Organization

There was no blueprint for implementing a new cataloguing code. The Library last changed cataloguing codes when it adopted AACR2, in 1981. We reviewed other large projects, such as MARC 21 implementation, to see what worked and to estimate scope, timescales and impacts.

Overall responsibility for the implementation was within my portfolio as Metadata Standards Manager. A member of my team, Thurstan Young, was appointed as Project Coordinator. We set up expert groups in Boston Spa and St. Pancras. The expert groups, drawn

Timeline	RDA Implementation at BL	External RDA Milestones
2005		JSC Announces RDA project
2007	RDA Expert Groups formed in BSp and StP to comment on RDA Drafts	Early drafts published BL, LAC, LC, NLA announce intention for coordinated implementation
2008	Implementation Project established Staff Awareness	Full Draft of RDA published for comment
2009	Thurstan Young Appointed RDA Implementation Project Coordinator RDA Implementation Environment set up on Aleph Staff Awareness	Final Text sent to publishers in June
2010–11	Toolkit evaluation Aleph Configuration & Testing Staff Awareness Training, workflow and policy development Train the trainers Batch Upgrade testing St Pancras Expert Group re-established following restructuring	RDA Toolkit Published US RDA Test LC announces Jan. 2013 as earliest implementation Date European RDA Interest Group formed (Chaired by BL) Deutsche Nationalbibliothek announces intention to implement RDA
2012 Q1-2	RDA Name authority records contributed to NACO Redistribution of derived RDA bibliographic records to customers	Deutsche Nationalbibliothek joins JSC National Library of Scotland defers implementation to 2014
2012 Q3-4	RDA original cataloguing piloted Authority Control Training Catalogue Support and Acquisitions Training Bulk Toolkit License purchase FRBR/RDA Refresher and Toolkit session in BSpa	NACO Phase 1 Changes distributed Asbestos in BSp: training replanned
2013 Q1	Cataloguing in Publication switched to RDA Cataloguer training in BSp and StP PRIMO changes	NACO Phase 2 changes distributed
April 2013	British Library announces RDA implementation	Implementations announced: Library of Congress National Library of Medicine Cambridge University Oxford University

from experienced cataloguers and team leaders, provided the vanguard of trainers and practitioners. We identified the capability to produce RDA records as the deliverable of the project. To achieve this we would need to reconfigure systems, revise documentation and train staff. We therefore worked closely with colleagues in the Aleph Support Team, the Metadata Systems Team and the Documentation and Training Team. Although we eventually set a target date for implementation, it was never conceived of as a big bang. RDA was phased in gradually by adding capability to systems and by increasing the number of staff working in RDA over the course of 2012–13.

reconfigure systems – revise documentation – train staff

Systems

The Library's main cataloguing platform is Aleph, in which we created an RDA implementation environment. The RDA environment was used to test configuration changes and enable expert groups to practice without affecting the production database.

MARC 21 changes required numerous configuration changes to Aleph. The Library is able to make these changes directly in the Aleph tables, without involving Ex Libris. The new elements required definition of new indexes, including those for Content, Media and Carrier Types and for many new authority attributes. These changes were documented and have been made available on our Website.[21] A great deal of work was also done to review and update the templates and macros used by staff throughout the workflows, including selection and acquisition as well as cataloguing.

training plan for cataloguers

Another major task was to review the impact on the Batch Upgrade system. This process automatically updates BL & LC CIP records after the item is received and a full record is imported from Library of Congress. The introduction of a new set of cataloguing rules meant additional checks and decisions during the matching and merging process. Batch Upgrade makes a significant contribution to achievement of our production targets, so this work had high priority and was completed by June 2012.

There were similar considerations in relation to our validation tool, MARC Report. Validating MARC records against two standards is more complicated than against AACR2 only. The Library worked closely with the developers, TMQ[22], to ensure that MARC Report continued to meet our needs. TMQ's solution was to enable validation of AACR2 only, RDA only, or mixed files of records.

minor changes to the configuration of the public discovery interface

We made minor changes to the configuration of Explore the British Library, our public discovery interface. Explore is an implementation of PRIMO. RDA offers a lot of potential for enriching discovery, but the current volume of RDA data within our catalogue is too small to warrant substantial changes to PRIMO. Future development of PRIMO will also be dependent on enhancing the consistency and content of our legacy data.

Training

Different training plans were developed for different groups of staff.

Trainers

Trainers received training in how to deliver training. They also received introductory training in RDA based on the training modules developed by Library of Congress. They were given the time and encouragement to practice and to raise questions and issues, which the project team tried to resolve and feed back into the training materials. An issues log was used to track this activity.

Processing staff

Acquisitions staff, copy cataloguers, finishers and shelf markers, received job focused training and their task-centred documentation was revised. Training was mainly delivered in the team setting. The team managers received cataloguer training to facilitate communication between the processing staff and cataloguers.

Cataloguers

The training plan for cataloguers was delivered in three stages.

Stage 1: Orientation: introduction to FRBR model and RDA. Introduction to new terminology and concepts.

Stage 2: Familiarisation: reminder of FRBR & RDA structure; changes from AACR2; RDA vocabulary; navigating the Toolkit.

Stage 3: Implementation: Modular training course based on Library of Congress materials; hands on practice; review period.

Each team received training over a week, delivered by their manager and another member of the expert group. Each module was followed by hands-on practice. During the review period work was checked and gradually increased in complexity. This period varied between a few days and a couple of weeks, until the team manager was satisfied. Specialist staff received additional modules, but all cataloguers received the

same basic training. Following implementation, teams were also offered the option of attending »FRBR for the terrified« and RIMMF (RDA in Many Metadata Formats)[23] sessions, which looked forward to future application of RDA in a world without MARC.

227	Total Staff trained
121	Cataloguers trained
106	Processing Staff trained

Training raised many questions and resolving them was not always easy. The experience gained by implementers and practitioners over the last two years is now being reflected in RDA updates and releases.

Application of RDA in MARC can be difficult and it is this aspect of RDA that has probably caused most issues in the training I have subsequently delivered, at the National Library of Iceland and University of Amsterdam. With this experience, we have updated and refined our training packages and developed *RDA in a Day* and a longer course, which can be run over 3 days. These packages will boost training capacity in the UK, but a few hours or days training do not deliver expertise; that develops through application and reflection.

Documentation

In addition to documentation created for training, we reviewed and amended existing policy documentation. At an early stage in planning we decided to make use of the Workflow tool provided with the RDA Toolkit. A small group was set up to work on different content streams. The workflows were envisaged as a mediation layer, on top of the RDA instructions, which would guide cataloguers through the record creation workflow for different content streams. The workflows highlight decision points and provide links to RDA instructions, British Library or LC/PCC policy decisions and the MARC manual. They took a lot of work to create and the have to be maintained, but they have reinforced training and consistent application of RDA. The workflows are published in the RDA Toolkit[24] for the wider community.

Application

Our experience with RDA has been positive. Cataloguers on the whole seem to like RDA and appreciate what it is trying to do. However, there have been a few productivity concerns.

The main issue has been around authority control. We have a »team« authority control measure which was 20 % prior to implementation, but climbed to 33 %. This means in effect, that staff, who spent one day per week carrying out name authority work on behalf of colleagues before RDA was implemented, were spending closer to two days a week after implementation. This was not surprising, as RDA rightly puts much greater emphasis on authority control than AACR2 and has defined many more specific elements for identification of persons, families, and corporate bodies. Nevertheless, it was not a sustainable rate within our productivity targets. Analysis showed that more time was being taken to process conference proceedings and compilations. Much of this additional cost was attributed to additional authority work needed to identify multiple contributors.

RDA updates and releases

AACR2 imposed an arbitrary cap (the »rule of three«) but RDA provides more logical alternatives (RDA 0.6). As the national bibliographic agency, we encouraged cataloguers to record relationships to all persons, families and corporate bodies so long as it was not *onerous*. We wanted cataloguers to exercise their judgment; therefore we did not define onerous. When we realised that the lack of guidance was having an impact on productivity, we introduced local policies to limit the number of authorised access points assigned to these types of resources.

Productivity levels for RDA are otherwise comparable with the levels for AACR2. The cataloguing daily rate in October 2012 was 10.66; in October 2013 it was 10.41. The daily rate for copy cataloguers in October 2012 was 39.02; in October 2013 it was 39.09.

policy documentation

There are many other factors, including electronic legal deposit and the impact of budget cuts which make it impossible to isolate the impact of RDA. These additional pressures mean that we will rely more and more on derived data and automated processes. In this context collaboration and accurate identification of entities will be fundamental to future efficiencies and workflows.

RDA today and tomorrow

In this context of doing more with less, we need RDA to be a flexible tool and as widely adopted as possible. In this context, the translation not only of RDA, but also of the metadata elements and vocabularies should make it easier to reuse metadata from other cultures and other communities. There are many barriers to be overcome and in this final section I want to reflect on some of these issues.

many barriers to be overcome

Implementing FRBR

In 2013 the British Library implemented RDA but we have not implemented FRBR. Our implementation was a pragmatic response to a major change in the bibliographic standards landscape, but it is unlikely that users of our products, services or catalogues have no-

MUSIKSAMMLUNGEN IN DEN REGIONALBIBLIOTHEKEN DEUTSCHLANDS, ÖSTERREICHS UND DER SCHWEIZ
Hrsg. von Ludger Syré
2015. 446 Seiten, gebunden, Fadenheftung
ISBN 978-3-465-04235-8
ZfBB Sonderband 116

Vor dem Hintergrund der gegenwärtigen Diskussion um die Funktion der Regionalbibliotheken als Forschungsbibliotheken hat sich die Arbeitsgemeinschaft der Regionalbibliotheken der Aufgabe angenommen, die in ihren Einrichtungen überlieferten Bestände, und hierbei insbesondere die Sondersammlungen, in ihrer Bedeutung für die Forschung darzustellen. Der Fokus richtet sich dabei besonders auf jene Bestände, die im »Handbuch der historischen Buchbestände« als Sondersammlungen bezeichnet werden. Unter den dort aufgezählten Materialarten finden sich zahlreiche Nicht-Buch-Materialien, die eine hohe Relevanz für die Forschung besitzen.

Nach den Dichternachlässen und den Kriegssammlungen greift die Arbeitsgemeinschaft mit den Musikalien nun einen weiteren Sonderbestand auf, der zu jenen Quellen zählt, die aus der Sicht der Wissenschaft den besonderen Wert der Sammlungen der Regionalbibliotheken ausmachen. In den Aufsätzen dieses Bandes wird die Relevanz der jeweiligen Musiksammlung für musikwissenschaftliche und musikhistorische Forschungsansätze, aber auch als Quelle für die regionale und lokale Musikgeschichtsschreibung sowie für die historische Aufführungspraxis sichtbar. Die in diesem Band versammelten Aufsätze zeigen: Nicht zuletzt über die Musik kam »die Welt« in die Region.

VITTORIO KLOSTERMANN

ticed any difference. To realise the potential offered by RDA several further developments are required, only some of which are within our own control.

Metadata Quality

Consistent, sufficient metadata is the *sine qua non* of the FRBRized catalogue. Unfortunately, our metadata is neither consistent nor sufficiently uniform to identify all the FRBR entities. We are therefore enhancing our metadata programmatically. This is a long term, multi-stranded activity.

Subsequent on the implementation of Aleph in 2004, we invested heavily in retrospective alignment of name headings in our catalogue with the LC-NACO authority file. A similar process is necessary to identify and align works and expressions. This is in its infancy; we have begun some testing and analysis of small data sets.

The process will highlight other deficiencies that have to be addressed. For example, how do we identify expressions if the language of expression is not explicitly recorded in the legacy metadata data? We know that this is the case for a lot of our older records and article records. In many cases, it will not be possible to enhance the metadata at reasonable cost so we will prioritise and seek pragmatic solutions. For example, create »undifferentiated« expression records as a bridge between works and manifestations.

Different Standards and models

As I noted at the beginning of this article, several important collections are not covered by RDA implementation. These include our archival and manuscript collections. The British Library is also an archive and a museum! Archivists and curators have evolved standards and models that serve the needs of their communities and are different from library standards and models. However useful these distinctions have been to us in the past, the resulting silos are a barrier to our audiences. The value of offering a FRBR interface to our collections is diminished unless it can also include non-library materials. Therefore our internal metadata strategy is to aim for convergence of standards.

This does not mean that one size should fit all, but our audiences can only have a coherent view of our collections if there is a coherent underlying framework of models and standards. I am therefore excited by the initiatives being taken by IFLA and RDA to engage with special materials communities and with archivists.

Technical Infrastructure

Our technical infrastructure has not significantly changed since we implemented RDA. Some aspects

of RDA are more complex than others and can be difficult to grasp, particularly in a MARC environment. We have found that RIMMF can help cataloguers visualise RDA's intentions, which in turn enables them to determine how best to proceed in MARC, but what we need are new systems and new schemas.

We are just beginning a landscape study to analyse the technical and standards landscape in which a future architecture would be implemented. It is really too early to comment on this and the only certainty is that the future won't be dull.

analyse the technical and standards landscape

1 Committee of Principals Affirms Commitment to the Internationalisation of RDA. 2015. Available at: www.rda-jsc.org/node/235
2 Delsey, Tom. Encoding RDA data. 2007. Available at: www.rda-jsc.org/archivedsite/docs/5editor3.pdf
3 Delsey, Tom. RDA Database Implementation Scenarios. 2009. Available at: www.rda-jsc.org/archivedsite/docs/5editor2rev.pdf
4 Data Model Meeting: British Library, London 30 April–1 May 2007. Available at: www.bl.uk/bibliographic/meeting.html
5 Dunsire, Gordon. A short history of the evolution of the catalogue record. 2009. Available at: www.gordondunsire.com/pubs/pres/EvolCatRec.ppt
6 Riva, Pat, and Maja Žumer. Introducing the FRBR Library Reference Model. 2015. Available at: http://library.ifla.org/1084/1/207-riva-en.pdf
7 RDA registry. Available at: www.rdaregistry.info/
8 RDARegistry/RDA-Vocabularies. Available at: https://github.com/RDARegistry/RDA-Vocabularies
9 About the RDA Registry and vocabularies: Versions and Releases. Available at: www.rdaregistry.info/rgAbout/versions.html
10 RIMMF3 home. Available at: www.marcofquality.com/wiki/rimmf3/doku.php?id=rimmf
11 Glennan, Kathy. Machine-Actionable Data Elements for Measurements, Extent of the Carrier, Pagination and Foliation, Dimensions, Extent of the Content, and Duration – Discussion Paper (2015). 2015. Available at: www.rda-jsc.org/sites/all/files/6JSC-ALA-Discussion-5.pdf
12 Dunsire, Gordon. RDA accommodation of relationship data. 2015. Available at: www.rda-jsc.org/sites/all/files/6JSC-TechnicalWG-6.pdf
13 Dunsire, Gordon. RDA models for authority data. 2015. Available at: www.rda-jsc.org/sites/all/files/6JSC-TechnicalWG-5.pdf
14 Dunsire, Gordon, and Deborah Fritz. RDA and FRBRoo treatment of aggregates. 2015. Available at: www.rda-jsc.org/sites/all/files/6JSC-AggregatesWG-1.pdf
15 JSC. RDA Elements Table and Registry updated. 2015. Available at: http://rda-jsc.org/node/246
16 Dunsire, Gordon. What is an RDA record? 2014. Available at: www.gordondunsire.com/pubs/pres/WhatRDARecord.pptx
17 Danskin, Alan. RDA implementation case study. In: *Catalogue & Index*. December 2013, Issue 173, pp 24–29. Available at: http://tinyurl.com/oyryfkp
18 Danskin, Alan. RDA implementation and application: British Library. In: *o-bib.* 2014, Bd. 1, Nr. 1. DOI: http://dx.doi.org/10.5282/o-bib/2014H1S187-191
19 The British Library Annual Report KPIs and Statistics: 2014/15. Available at: www.bl.uk/aboutus/foi/pubsch/pubscheme6/Annual%20Report%20KPIs%20and%20Statistics%2014-15%20FINAL%20incl%202%20versions%20RR%20KPIs.xlsx
20 The British Library Annual Report KPIs and Statistics: 2014/15. Available at: www.bl.uk/aboutus/annrep/2014to2015/annual-report2014-15.pdf
21 British Library. Collection Metadata: Standards. Available at: www.bl.uk/bibliographic/catstandards.html#rda
22 TMQ Web Page: www.marcofquality.com/
23 TMQ RIMMF: www.marcofquality.com/wiki/rimmf/doku.php
24 RDA Toolkit: www.rdatoolkit.org/

Die Verfasser

Simon Edwards, Director of Professional Services at Chartered Institute of Library and Information Professionals, Chair of RDA Committee of Principles, The Chartered Institute of Library and Information Professionals, 7 Ridgmount Street, London, WC1E 7AE, Tel.: +44 (0)20-7255-0605, E-Mail: simon.edwards@cilip.org.uk

Gordon Dunsire, Chair of Joint Steering Committee for Developing RDA, Tel.: +44 (0)131-346-1026

Alan Danskin, Collection Metadata Standards Manager at the British Library, Member of the Joint Steering Committee for Developing RDA, British Library, Boston Spa, Wetherby, West Yorkshire, LS23 7BQ, United Kingdom, Tel.: +44 (0)1937546669, E-Mail: alan.danskin@bl.uk