Extension of the RDA/ONIX Framework

Draft proposal to JISC, September 2008

(Note: The project name was changed to the Vocabulary Mapping Framework once JISC funding was agreed)

Introduction

This is a proposal for extending the existing **RDA/ONIX Framework** for Resource Categorization and creating a downloadable **vocabulary interchange tool** to support the interchange of metadata content between major standards from the publisher/producer, education and bibliographic/heritage communities relevant to the Information Environment. This tool can be used in a variety of ways to **enable automated reuse of metadata** from different sources and schemas, to improve the quality and access and reduce the cost of metadata in the JISC community.

The proposal has six sections:

- Section 1: Overview, outlines the proposal and its rationale
- Section 2: Deliverables and cost
- Section 3: Project plan
- Section 4: Use cases
- Section 5: The RDA/ONIX Framework, describes the extended schema in more detail
- Section 6: Further information, gives more background and detail on various aspects of the project

A number of Appendixes support various sections.

The project is detailed in three stages, the second and third of which are optional.

- Stage 1: extended RDA/ONIX Framework - £48,645 (inc. VAT)
- Stage 2 (optional): automated generation of mappings - £15,000 – 20,000 (inc. VAT)
- Stage 3 (optional): maintenance of the Framework – max. £10,500 (inc. VAT) for one year only, thereafter to be self-supporting

1. Overview

1.1 Background

This proposal follows on from one made in June 2008 under the PALS3 call. That proposal was not thought suitable for funding in that form under that call, but the reviewers and members of the PALS committee felt there was sufficient merit in it to suggest that it be re-worked and that funding perhaps be sought elsewhere within JISC. This document is the result of that re-working. We are grateful for the two reviewers' comments, copies of which are included in the Appendix, including our comments on how issues raised have been addressed in this revised proposal.

1.2 Context

1.2.1 Trends in vocabularies

Metadata vocabularies form a major part of metadata schemas, and carry a substantial part of the meaning of metadata. Vocabularies are growing rapidly, both in number and in size. Vocabularies are becoming increasingly complex and granular. Vocabularies of relators (terms that describe relationships between entities) are growing in importance. *See 6.1 for further analysis of these points.*

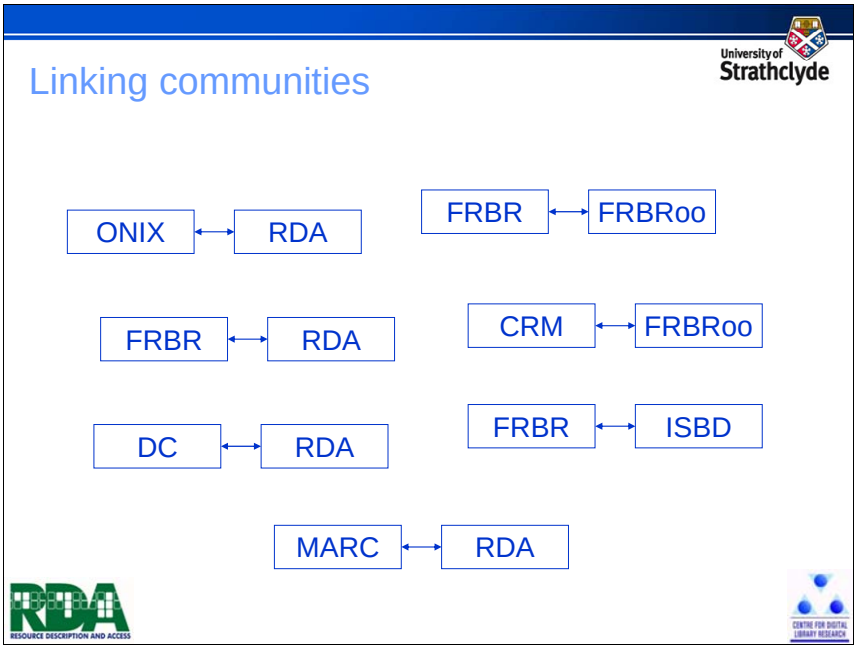
1.2.2 Trends in JISC community metadata

We make these three assumptions:

1. Bibliographic and heritage metadata is becoming increasingly diverse and complex and will require increasing interoperability for re-use and discovery.

Current extensive vocabulary developments within RDA, MARC, FRBR/FRBRoo, CIDOC CRM and Dublin Core, and the interactions between them, are the primary evidence of this[1]. Four of these five initiatives have come into being in the last decade (the exception being MARC, where there has also been a significant growth of variations) and driven mainly by the impact of RDA and FRBR there are now joint working and liaison groups between various combinations of them. The slide below, from a recent RDA presentation (August 2008) by Gordon Dunsire at an IFLA conference, illustrates the point graphically, where each pair represents significant integration activity:

[1] See Table 5 in the Appendix for a summary of, and links to details of these standards.



Each of these relationships (except FRBR-to-FRBRoo) requires vocabulary mapping. These developments greatly increase the amount of schema-to-schema mappings that are required. This slide contains only one publisher/producer standard (ONIX) and none from education or other sectors, so the impact of adding others in terms of a potential "combinatory explosion" of mappings is evident.

2. Members of the JISC community have increasingly diverse, complex and unpredictable metadata requirements.

At a basic level this trend is summarized by one of the reviewers of the original PALS proposal, when discussing the take-up of the proposed RDA "MyRDA" tool:

"Different libraries have different requirements. A public library might only need the general rules. A music library would want to include all rules relating to music scores, music recordings, as well as books and serials. An academic library is more likely to need the rules for e-journals. It is difficult to predict how quickly this functionality will be offered by vendors and exactly which aspects will be taken up."

The rapid growth of the development of complex multimedia resources for education at all levels means that more automated acquisition and integration of metadata from more diverse sources is a growing requirement.

3. Metadata from producers/providers/publishers[2] will become increasingly important as a substantial component of metadata in the JISC community.

It is now quite common that metadata originated in ONIX and IEEE LOM and in some proprietary formats like Crossref[3] to populate library and VLE systems (transformations of metadata like this are routinely provided by agencies such as OCLC, Nielsen and the Library of Congress). This will extend across all media types in time, and any authoritative domain standards such as DDEX (music industry) or PRISM (magazines) can be expected in due course to provide metadata to the JISC community.

[2] Traditionally this has been referred to as "publisher metadata" but the description is broadened here to reflect the multimedia environment.
[3] http://www.crossref.org

Again the rationale for this is put well by one of the reviewers of the original PALS proposal, seen in this case from the perspective of RDA:

"Using complementary standards means less maintenance for RDA. RDA is designed to be applied by a range of communities, and recognises the usefulness of using existing term sets and vocabularies, rather than creating an RDA set; for example, role terms from MARC 21".

ONIX has led the way, but such standards are being introduced and beginning to gain use in many domains[4] and so the availability of good quality producer/provider/ publisher metadata in other standards is both required and expected.

1.2.3 The RDA/ONIX Framework

The **RDA/ONIX Framework**[5] is a vocabulary matrix developed by a joint working group in late 2006 as a tool for creating well-formed and interoperable vocabularies for the library and content-provider communities. It has been successfully used by the RDA to support the creation of resource categories, and the intention has been for the framework to be extended, but to date there has not been an appropriate motive and opportunity. *See 5.1 for more details of the existing Framework.*

1.3 Summary of deliverables

The proposal is for a major expansion of the existing **RDA/ONIX Framework** to create a web tool which will support the automated mapping of vocabularies from major metadata standards of use to the JISC community. Vocabularies to be analysed are from RDA, MARC, ONIX, DDEX, FRBR, IEEE LOM, Dublin Core, DOI and the CIDOC CRM. *See table 5 in the Appendix for details of source standards.*

The Framework supports "one-to-many" mapping, so that each vocabulary is mapped once to the central Framework ontology, and optimal mappings between different vocabularies may then be computer-generated. *See 6.3.1 for analysis of the benefits of "hub-and–spoke" mapping.*

The project includes a proposal for self-supporting governance for the ongoing maintenance and development of the vocabulary, with namespace hosting by the IDF.

The framework is extensible to include other vocabularies as and when required.

1.4 Benefits

The value to be derived from the development of this tool comes in:

- **Efficiency**: fewer manual crosswalks and transformation required, and less work in creating and maintaining them
- **Quality**: providing authorised mappings with the minimum loss of meaning
- **Access**: opening up more possible metadata sources to community members
- **Timeliness**: dynamic updating means an assured, quicker and automated response to additions to vocabularies

[4] The DDEX message set, published earlier this year and now the subject of several pilot implementations involving major record companies, copyright societies and online music providers is an obvious example of how the multimedia web market is driving the use of metadata standards. Although transaction-data-led at this point, DDEX includes messages supporting rich bibliographic data with substantial vocabularies.

[5] RDA and ONIX are named as they were the sponsors and main focus of this work, but the goal was to create a Framework suitable for mapping vocabularies for any standard and so the Framework has no bias to any one standard or approach.

A further possible value to the community as a whole is the creation of a low-maintenance point of contact (the proposed policy governance group) between diverse metadata standards groups that are using and contributing to a common tool.

The proposal is both strategic and immediately practical. The use cases provided (*see section 4*) demonstrate that the tool has potential applications in the making of metadata crosswalks, mapping of local bespoke schemas, cross-searching from institutional repositories and in support of preservation metadata.

1.5 Risks

The project is low-risk as there is no significant technical invention required, and only a reasonable level of consensus and completeness is essential for success.

Possible risk	**Mitigation**
"R&D" unknowns	There is no significant technical invention involved: although specialized and quite complex, the technical work uses established tools and techniques.
Incompatibility of schemas	The underlying data models of the most of the schemas[6] have already been found to be compatible through previous work, borne out by the success of the first release of the RDA/ONIX Framework.
Insufficient consensus	We have reason to expect a good level of co-operation and support (*see 6.4.1*) from the authorities of the source standards but because we are working from published materials, active support is not essential. If for any reason we encounter opposition to the mapping of any particular schema or vocabulary we will simply omit it at this point.
Insufficient time to complete, or local complexity.	There is complexity, idiosyncrasy and imprecision in some of the vocabularies, and there may be some terms in some standards which prove resistant to an effective mapping in the time available. In that case they will simply be left to a later time. Provided the majority of vocabularies are mapped the tool will fulfil its function. As noted above, sufficient work has already been done in one-to-one mappings and in the existing Framework to eliminate any risk of serious "semantic roadblocks".
Too complex for implementation	We are using established Semantic Web standards. There will be those who still find their use challenging, but RDF and OWL tools and schemas are now proven, if not yet commonplace.

[6] The majority of the source schemas (RDA, FRBR, CIDOC, ONIX, DDEX and DOI) are based on or compliant with three underlying data models: FRBR, CIDOC CRM and the <indecs> framework, and these three have compatible entity models. The other three (MARC, DCMI and SCORM) are not model-based but existing mapping work has

2. Deliverables and costs

The project is detailed in three stages, the second and third of which are optional.

2.1 Stage 1: extended RDA/ONIX Framework

The deliverable of stage 1 will be a **major expansion of the existing RDA/ONIX Framework**, based on and incorporating relevant vocabularies from nine main bibliographic and supply chain metadata standards (the "source standards") of interest to the JISC community. *See table 5 in the Appendix for details of the source standards.*

The Framework (which currently supports the categorisation of resource content and carriers) will be extended to support the categorization of:

- works
- parties
- relators between resources, and relators between parties and resources

See 6.2 for further details of the introduction of relators to the Framework.

See table 3 in the Appendix for a description of a provisional schema, with an example in table 4.

The Framework will be represented in RDF/OWL for computer access and in Word/PDF formats for human readability, and declared as a SKOS[7] (semantic web) vocabulary.

The result will be an open source tool which developers can then use to create or incorporate automatic, contextual term mappings into crosswalks, transformations or search technology. Other representations of the Framework, in part or whole (for example as an XML schema) may be created by transformation from the RDF/OWL as required.

For the vocabulary to be authoritative, mappings require the approval of editors or other appropriate governors of the source standards. As far as possible that will be achieved in the course of the project. Until such time as mappings are formally approved all mappings will be shown as *provisional.*

The results will be presented to interested parties at a one-day conference/workshop, hosted at the British Library, at which representatives of the source standards will also have the opportunity to present developments which may encourage supply chain metadata reuse.

A mechanism for governance for the RDA/ONIX Framework will be established in Stage 1, with the IDF expected to host its namespace. *See 2.3 ongoing maintenance of the Framework.*

The proposed cost of Stage 1 is £48,645, representing 54 project days (including VAT). *See section 3.1, table 1 for the workplan*.

2.2 Stage 2 (optional): automated generation of mappings

The RDA/ONIX Framework ontology created in Stage 1 will support the **automatic generation of term-to-term mappings** between any pair of mapped schemes using reasonably straightforward ontological inference.

This task may be left entirely to users of the Framework, or an **initial set of inference rules and mappings** may be created as part of this project to make the tool immediately of greater value.

[7] The W3C Semantic Web "Simple Knowledge Organization Systems" RDF vocabulary, http://www.w3.org/2004/02/skos/

Stage 2, if funded by JISC, would result in the creation of a set of inference rules and the automated creation of scheme-to-scheme mappings for a selection of scheme pairs between whose vocabularies have been included. The inference rules would be available as part of the Framework for others to adapt or apply.

See 6.3.3 for further information on automated vocabulary mappings using the Framework.

Stage 2 is optional and does not require commitment either way from JISC at this point, with an estimated cost of £15-20,000.

2.3 Stage 3: maintenance of the Framework

The RDA/ONIX Framework will require ongoing maintenance in three respects:

a. a web host;

b. additions to existing and the addition of new vocabularies; and

c. governance.

The cost of (a) is expected to be met by the IDF, so no funding for this is sought.

The cost of (b) should in due course be met by a registry model, where those registering vocabularies bear any costs of update. However, the Framework will need a chance to demonstrate its value, and so we are proposing a provision covering one year following the completion of Stage 1 to support published changes to existing vocabularies within the Framework (but not the addition of new ones).

For (c), policy-making (an **RDA/ONIX policy and review group**) should be funded by those willing to participate, so no funding for (c) this is sought beyond Stage 1.

The proposed cost of this to JISC of Stage 3 (for task (b) only) is expected to be no more than £10,500, representing up to 15 project days for maintenance and convening the review group.

A detailed governance model, with support from participating standards where possible, will be produced in Stage 1 (*see 2.1*).

3. Project plan

3.1 Tasks and resources

The project will be managed by Rightscom and the principal research, analysis, semantic and technical work will be undertaken by Godfrey Rust (who will also be project manager) and Steffen Lindek (who will also be project administrator) of Rightscom, and Gordon Dunsire, of the University of Strathclyde and currently co-chair of the RDA/DCMI Task Group.

Some additional provision is made for payment and/or expenses for experts for half-day or day consultations in relation to individual standards in which they are specialist consultants. The majority of source information is available in published form, or further advice will be available from individuals for whom it is a normal part of their employment to provide it.

There will be a one-day expert workshop (similar to that which produced the RDA/ONIX Framework) attended by the three principals and three or more others (among those invited we expect to be David Martin of EDItEUR/ONIX, Martin Doerr of CIDOC CRM and another expert from the bibliographic domain).

The results will be presented at a one-day conference, probably at the British Library[8]. A sponsor will be sought to cover the administrative costs of this event. Provision is made to pay for someone appropriate to chair the event.

All other work will be desk work or using email, web and phone communication.

[8] The Library offered to host this as part of the original proposal and will be approached again.

Table 1: Tasks and resources (Stage 1 only)

Task	Description of activity	Resources (days)				Comments
		GR	SL	GD	Oth	
1	**Extend Framework structure**					
1.1	Express existing Framework in RDF/OWL structure and add structure for relators, works and parties.	1	2	1		1. A draft of the abstract schema for the extended Framework is outlined in table 3 in the Appendix. 2. For convenience we expect to use another open source ontology tool (Sigma) and language (SUO-KIF) to create the relationships, and then output them in RDF and OWL when complete. This has no impact on cost or future dependency.
1.2	Gather and add source standard vocabularies into Framework structure.		3			1. Brings candidate vocabularies from the source standards into the Framework format, assigns "surrogate URI" identities where needed, adds names and descriptions, etc. 2. Does **not** include mapping of vocabulary terms to Framework attributes at this point. 3. Candidate vocabularies are indicated in table 7 in the Appendix.
2	**Extend Framework content**					
2.1	(a) Prepare candidate attribute, category, relator and verb terms; (b) consult with sources, create first draft deliverables, add vocabularies and create categories and relators.	6	12	4		1. Vocabularies included may be proposed or provisional, but these are included only as input for developing framework terms, and no vocabulary will be mapped if it is not part of a published standard. This is desk work with email and teleconference communication when needed. 2. Risk mitigation: we hope to map all the candidate vocabularies to the Framework, but we will prioritize on the basis of immediate usefulness, and complete as much as budgeted time allows.
2.2	Provision for consultation by independent experts on specific vocabularies (in the course of task 2.1).				2	We will seek input from other experts on specific standards and expect some to be able to give their time as part of their existing responsibilities. We have made this provision to pay some experts who are consultants such as David Martin of ONIX or Martin Doerr of CIDOC CRM.
2.3	1 day workshop to review partial result and resolve issues (in the course of task 2.1).	1	1	1	3	Face to face meeting of the team with key experts to review issues.
2.4	Drafts distributed to interested groups, and deal with responses.	2	2	2		Risk mitigation: if issues arise with mappings that cannot be resolved in the time available, mappings will be omitted at this stage.
3	**Finalise deliverables**					
3.1	Prepare final representations: RDF/OWL, Word/PDF/HTML, SKOS declaration (including namespace and URIs).	1	3	3		
3.2	1-day conference: present the extended Framework to interested parties	1	1	1	1	Hopefully to be hosted at the British Library. We will seek a sponsor to cover the administrative and marketing costs of the event. "Other" is provision for a chair for the day.
	Total Stage 1	12	24	12	6	

Project management and administration time is included in each section.

3.2 Personnel

Gordon Dunsire (GD)
Gordon Dunsire has been Depute Director of the Centre for Digital Library Research at the University of Strathclyde since 2002. His main research interests are in metadata for information retrieval in distributed, networked, heterogeneous information environments. He has been leader of several relevant research projects including HaIRST: Harvesting Institutional Resources in Scotland Testbed (http://hairst.cdlr.strath.ac.uk/) and Resource Discovery Infokit (http://cdlr.strath.ac.uk/rdinfokit/index.htm), and has been closely involved in many others such as CC-interop: COPAC/Clumps Continuing Technical Cooperation (http://ccinterop.cdlr.strath.ac.uk/) and SPEIR: Scottish Portals for Education, Information and Research (http://speir.cdlr.strath.ac.uk/).

He has also been a consultant on various metadata issues to several organisations, including the IFLA FRBR Review Group and is a member of a number of relevant groups:

- Member of the IFLA Section on Classification and Indexing
- Co-Chair of the DCMI/RDA Task Group
- Member of the DCMI Advisory Board
- Member of the RDA Outreach Group
- Co-ordinator of the European DDC Users Group Technical Issues Working Group
- Member of the Ligue des Bibliothèques Européennes de Recherche Task Group on MARC Harmonization in Europe
- Member of the CILIP-BL Committee on AACR
- Member of the CILIP Committee on DDC
- Member of the Cultural Technical Group of the Common Infrastructure Standards Advisory Group (of the Scottish Government)
- Chair of the Cataloguing and Indexing Group in Scotland

Gordon was facilitator of the RDA/ONIX Framework working group in 2006.

Steffen Lindek (SL)
A science Ph.D based in Heidelberg, since 2000 Steffen Lindek has been working as a freelance database and metadata consultant for various academic and commercial projects, mostly in the UK and in Germany. One of Steffen's foci has been the field of the Digital Object Identifier (DOI), and he worked as a consultant for the International DOI Foundation and for several of its members.

In recent years he has worked closely with Godfrey Rust and with Rightscom on the development and use of some of the source standards relevant to this project, such as DDEX and ONIX for Licensing, and on several projects (including the JISC “TIME” project) involving mapping or use of other source standards including MARC and Dublin Core. He is responsible for the ongoing ontology work on the DDEX project, and works mainly with Godfrey on ontology-based systems for commercial clients.

Godfrey Rust (GR)
Godfrey Rust has been a Director and Chief Data Architect of the UK-based consultancy Rightscom since 2003. He has thirty years experience in information management in content industries. He played a leading role in the computerization of the UK music charts, the establishment of the UK National Discography and the copyright societies’ Common Information System plan and the indecs metadata framework and was director of Data Services for the UK copyright societies MCPS and PRS.

His main current activities are the development of ontology-based systems for commercial clients but continues to participate from time to time in standards development. Specifically relevant to this proposal:

- Initiator and first draft of ISWC (International Standard Work Code)
- Technical lead for the <indecs> metadata framework
- Co-editor MPEG21 21000-6 (Rights Data Dictionary)
- Member of the RDA/ONIX Framework working group
- Data architect for JISC “TIME” metadata interoperability testbed (MARC, ONIX, LOM, DC)
- Contributor to initial design of ONIX for Licensing Terms
- Designer of DDEX Data Dictionary architecture
- Metadata consultant to International DOI Foundation

3.3 Project duration and timing

The minimum elapsed time of Stage 1 of the project would be four calendar months. Ideally the extended Framework would be available as early as possible in 2009 so that it can support the crosswalk activity that is bound to follow the publication of RDA and MARC vocabularies and their anticipated registration with SKOS.

4. Use cases

This section provides five use cases showing the value of the extended Framework to the JISC community.

4.1 Use case 1: metadata crosswalk

This is a short extract from an existing published crosswalk[9] between two schemes, which is used by others in creating transformations from ONIX to MARC records for library use:

6 *type of record*
If tag = ProductForm or b012
Media type = first character of the data in value

media type	*type of record*
a (audio)	*i (non-musical sound recording)*
b (book)	*a (language material)*
c (map)	*e (map)*
d (digital)	*m (computer files)*
f (film)	*g (projected media)*
v (video)	*g (projected media)*
w(mixed)	*p(mixed materials)*

If media type = 'p'
If second character = 'i' (sheet music)
Type of record = 'c' (printed music)

If the media type is not found in the above list, set the type of record to 'a'.

7 *bibliographic level*
If any of the following fields are present for the item, it signifies the item is a series, so set the bib level to 's' (serial).

Tag name	*Description*
SeriesISSN or b016	*ISSN of series*
PublisherSeriesCode or b017	*Publisher series code*
TitleOfSeries or b018	*Title of series*
ItemNumberWithinSeries or b019	*Number within series*
YearOfAnnual or b020	*Year of annual*

If none of the above items are found for the product, set the bib level to 'm' (monograph).

This extract contains two ONIX vocabularies (lists on the left hand side) with the corresponding MARC terms for use in specific records (on the right). Some 'special cases' are indicated below the lists. In the first case the mapping is from one vocabulary to another, with codes on both sides in this example; in the second case the mapping results in the creation of a text description in MARC.

This use case raises two fundamental problems:

1. What happens when a **change** or addition is made to either vocabulary on either side?
2. On whose **authority** is the mapping made?

The answer to question 1 (**change**) is that someone has to manually update the crosswalk whenever a new change is published in one of the standards. This is already a significant issue, and will become a major problem with proliferation of schemes and crosswalks. We have seen that ONIX vocabularies are growing at the rate of up to 10 percent a year, and we expect MARC to publish some significant changes in 2009. There are no guarantees that any published crosswalks will be maintained, and there are no formal mechanisms in place to do so.

[9] BobPearson (OCLC) mapping at http://www.editeur.org/onixmarc.html

The output of this project will enable a creator of a crosswalk guide, an XSLT or similar transformation, to embed links to the RDA/ONIX Framework which will enable the crosswalk to call on (or, if preferred, to generate) the latest term-to-term mappings from the RDA/ONIX namespace at the time that it is accessed, instead of having to manually encode the mappings and maintain them thereafter.

The answer to question 2 (**authority**) is unknown. The status of such crosswalks is almost always informal. As it is published on the EDItEUR site, we may presume it has the support of the ONIX publishers; its status in relation to MARC is not clear.

This is the case with most existing crosswalks: mappings are entirely a matter of individual judgment, and require re-analysis of the schemes and vocabularies on each occasion. Whereas a standard has governance, mappings almost universally are proprietary or "open source", so for the user it is simply a matter of "which do you trust" or even just "what is available". A standard may be carefully agreed and published, but its editors or governors rarely have any authority over how it is represented through mapping in a transformation.

The RDA/ONIX Framework proposal provides the opportunity for the governing authorities of a standard to participate in establishing an authorised mapping of its vocabularies into the ontology, and therefore into other mapped vocabularies.

Participation of authorities in the mapping process also improves the quality of the results. The success of this has already been demonstrated in a limited way in the initial creation of the RDA/ONIX Framework involving experts who were editors or representatives of a number of the participating standards. The process itself can also help authorities to identify and amend weaknesses or gaps in their standards.

These two issues – change management and authority – are key quality issues for metadata, and this project will put in place the means of achieving them, which can apply to existing crosswalks and to new ones. Both RDA and MARC are expected to publish substantial sets of vocabularies in 2009, and there will inevitably be a flurry of mapping and crosswalk activity in relation to one another and to some of the other standards in the scope of this project. It will be very beneficial to have the RDA/ONIX vocabulary in place to meet that challenge to support and enable the next generation of crosswalks.

This use case also identifies examples of common mapping issues which can be managed in the RDA/ONIX Framework mapping process:

- mapping to literal values rather than vocabularies (ONIX tag name to MARC description) *(see 6.3.3)*
- qualified mappings ("if second character i"…, "if any of the following are present…") *(see 6.2.4*

4.2 Use case 2: mapping of local, bespoke metadata schemes

The Scottish Collections Network (SCONE) uses a bespoke metadata scheme for collection-level description, based on the entity-relationship analysis of Michael Heaney. The JISC-funded cc-interop project[10] delivered a comparison of the SCONE schema with other local and national/international schemas in use at the time, with recommendations for augmenting national/international schemas such as the IESR[11]. The SCONE schema is richer than any national or international schema; this is required for specific operational purposes in the SCONE, Scottish Library and Information Resources online, and Research Collections Online services. The same

[10] CC-interop: COPAC/Clumps Continuing Technical Cooperation project. Available at: http://ccinterop.cdlr.strath.ac.uk/
[11] Extending the SCONE collection descriptions database for cc-interop : report for work package B of the cc-interop JISC project. Available at: http://cdlr.strath.ac.uk/pubs/dunsireg/CCIExtendSCONE.pdf

project also delivered mappings from the SCONE schema to various output formats, including IESR, Dublin Core Collections Application Profile (DCCAP), RDF/XML, and MARC21[12].

All of these schemas, including SCONE, have undergone significant development since this work was carried out. The mappings therefore require updating to remain effective. This currently requires an analysis and comparison of changes in each of the output schemas against the local SCONE schema. The issue is similar to that illustrated in Use Case 1, except here the scheme being mapped is not a published standard but a local scheme. A more sustainable and effective approach would be to map the SCONE vocabularies to a single cross-walk "spine" such as the extended RDA/ONIX Framework. Interoperability between vocabularies from SCONE, IESR, DCCAP, MARC21, etc. sources would then be assured if the relevant national/international vocabularies were also mapped to the spine, with mappings maintained by the schemas themselves.

This use case is applicable to any local collection- or item-level metadata schema which includes attributes for content, carrier or entity relators.

4.3 Use case 3: complex term mapping between domains

There are significant differences in the criteria used for identifying and categorizing resources in commercial and bibliographic metadata. This use case introduces the general problem, then looks at a specific issue for a JISC-sponsored example.

Partly as a result of the FRBR initiative, libraries are increasingly identifying the underlying *works* and *expressions* of different *manifestations*. The main driver for this “FRBR-isation” is to enable users to find different expressions of the same works (or parts of works), and different manifestations of the same expressions, typically for resource discovery.

The publisher/producer communities make similar distinctions, but for different reasons, and so often with quite different results. Rights ownership, rather than bibliographic or educational criteria, is typically a major reason for differentiating resources in the commercial domain.

For example, a record producer may recognize a new “arrangement” of a musical work, despite there being negligible musical differences to the original, because the work is out of copyright and the arrangement copyright may be claimed without contention. On the other hand a musically “original” arrangement of (say) a Lennon and McCartney song will typically not be recognized because no rights would be likely to be granted. A bibliographic or academic approach will use quite different, typically musicological criteria. From the same set of resources a producer organization may recognize, say, 200 distinct “work” versions of “Silent Night”, but only one of “Yesterday”, while a music library might recognize twenty of each. Such mismatches are common in music, and with the establishment of the ISTC it is likely to become more common in the textual domains.

This issue arises in relation to the JISC SWAP (Scholarly Works Application Profile) for Dublin Core. This Application Profile requires vocabulary for relators between drafts and versions of refereed articles in the ‘workflow’ process of the creation of scholarly works, from author through to publication, including the aggregation of works with other content (binding, illustrations, introductions etc); and through the de-composition of the same publications when content is made available to libraries etc. The British Library in particular has an interest in the latter.

It will be essential for such relationships to be clearly identified for the “FRBR-isation” of resources in library systems. The IFLA community is recognising that the definition of the granularity of works is contextual, and publisher and library views are functionally different. The identity and nature of a work in these circumstances is best determined by good relator metadata. Specific

[12] Output formats for collection-level descriptions from the SCONE database : report for Work Package B of the JISC CC-interop project. Available at: http://cdlr.strath.ac.uk/pubs/dunsireg/CCISCONEOutput.pdf

Relator terms are needed on both sides (publisher and library), and detailed mapping between them such as can be provided by the RDA/ONIX Framework.

4.4 Use case 4: Cross-searching metadata from institutional repositories

Many institutional repositories in the Higher Education sector in the UK are using item-level metadata schemas based on extended Dublin Core application profiles. Two such schemas are UKETD_DC[13] and SWAP[14], developed respectively by the JISC-funded ETHOS[15] and Eprints Application Profile[16] projects. Although the UKETD_DC schema is intended for electronic theses and dissertations, it is being used by some repositories to describe other institutional research outputs such as journal papers and conference presentations, for which the SWAP schema is intended. The schemas are disjoint; for example, UKETD has the attribute "Sponsor" (a relator term) which is not present in SWAP. Dumbing-down both schemas to Dublin Core is lossy, and is not likely to support the functionality required for cross-search services such as that being developed for the JISC-funded Institutional Repository Infrastructure for Scotland (IRIScotland) project[17]. This is discussed in the Draft IRIScotland metadata agreement[18], a deliverable of the project (note that this document will be updated as part of the JISC-funded extension to the project). An option for IRIScotland is to develop its own application profile, which would require mappings to UKETD and SWAP to support institutional contribution to other OAI-PMH communities. Using the extended RDA/ONIX Framework would improve sustainability and maintenance. Mappings to the Framework for publisher-generated metadata would also improve the ability of cross-search services to include published versions of theses, papers, conference proceedings and other scholarly research output, versions which are not necessarily identical copies of institutional versions.

SWAP is in active development by the DCMI Scholarly Communications Community[19], so the sustainability and maintenance of mappings to similar schemas is of current concern.

4.5 Use case 5: preservation metadata

A well-known case where the item base carrier is important for preservation or digitisation prioritisation is nitrate film stock, which is highly flammable and extremely fragile, requiring transfer of the content to a more stable polymer base or digital medium. The Framework can allow identification of the extent and location of items requiring priority treatment from multiple heterogeneous metadata records through mapping of schemas to the carrier ontology. It is unlikely that a common way of describing the characteristics of carriers will be developed, even within emerging preservation metadata schemas for digital content (which can be carried on non-digital media such as CD and DVD). For example, the PREMIS data dictionary does "… not attempt to define metadata for detailed documentation of media or hardware."[20] Specific types of carrier may require highly-specialised (and often expensive) equipment to digitise their content. More generally "The method of digitisation is dictated by the nature of the material selected"[21]. The Framework carrier ontology is therefore likely to become a very useful tool for digitisation programmes funded by JISC and partners such as the Strategic Content Alliance.

---

[13] UKETD_DC: The metadata core set recommended by EThOS. Available at: http://ethostoolkit.cranfield.ac.uk/tiki-index.php?page_ref_id=47

[14] Scholarly works application profile. Available at: http://www.ukoln.ac.uk/repositories/digirep/index/Scholarly_Works_Application_Profile

[15] Electronic Theses Online System (EThOS). Available at: http://www.ethos.ac.uk/

[16] Scholarly Works Application Profile (SWAP). Available at: http://www.ukoln.ac.uk/repositories/digirep/index/Eprints_Application_Profile

[17] IRIScotland. Available at: http://www.iriscotland.lib.ed.ac.uk/

[18] Draft IRIScotland metadata agreement: standards and guidelines for institutional repositories. Available at: http://cdlr.strath.ac.uk/pubs/dawsona/irismetadatadraft.pdf

[19] DCMI Scholarly Communications Community. Available at: http://dublincore.org/groups/scholar/

[20] PREMIS data dictionary for preservation metadata. Version 2.0. Available at: http://www.loc.gov/standards/premis/v2/premis-2-0.pdf

[21] Freeze frame: historic polar images. Available at: http://www.jisc.ac.uk/whatwedo/programmes/digitisation/polar

4.6 Other Use Cases for relators

The need for a standard framework of relators to describe relationships between resources identified in different metadata and identification schemes was demonstrated in a set of uses cases produced for an ISO TC46/SC9 working group (Bide 2005[22], *Use cases for interoperability of ISO TC46/SC9 identifiers*). It was this work that provided the initial impetus for the original JISC PALS proposal.

[22] www.collectionscanada.gc.ca/iso/tc46sc9/docs/sc9n417.pdf

5. The RDA/ONIX Framework

This section introduces the current Framework, and specifies the extensions to be made in this project.

5.1 The existing RDA/ONIX Framework

The RDA/ONIX Framework for Resource Categorization[23] was developed in 2006 by a joint working group of experts from the library and publishing industries. Its goal was defined as follows:

> *"The objective is…a framework for categorizing resources in all media that will support the needs of both libraries and the publishing industry and will facilitate the transfer and use of resource description data across the two communities."*

The Framework allows for resource categories to be created out of a matrix of pre-defined attributes (a "pizza menu" approach where different ingredients are selected and the resulting combination given a name as a new category). The Framework therefore enables categories in different schemes to be defined and mapped securely to one another irrespective of naming differences. This method is a common form on ontology known as "Formal Concept Analysis"[24].

The framework was produced in a relatively short time by a small working group of experts in a number of major metadata standards. It was well received upon publication and has been used as a tool for defining the three proposed resource category lists in RDA (Media Category, Type of Carrier and Content Category)[25].

As yet none of the detailed lists from ONIX or other commercial standards have been incorporated, although that was and remains the intent[26]. A proposal for resource categories based on the RDA/ONIX Framework was recently drawn up by Rightscom at the request of the International DOI Foundation.

The current project would extend the structure of the RDA/ONIX Framework to include relators, extend the matrix to cover the scope of the selected standards, and then populate it with the selected vocabularies.

5.2 The extended RDA/ONIX Framework

This project will extend the structure and content of the Framework and express it in a Semantic Web format which will make it accessible for computer processing, including inference.

5.2.1 Format change and web declaration

The Framework is current expressed only in a human readable, tabular form. In this project it will be expressed in the Semantic Web description language RDF[27] including the ontology language OWL[28]. It will be declared as a SKOS[29] vocabulary, with a namespace and with URIs issued for each term. This combination will enable developers to reference the Framework directly as part of their software, and additions to Framework contents can be incorporated dynamically in other systems.

5.2.2 Extensions to structure

[23] http://www.loc.gov/marc/marbi/2007/5chair10.pdf
[24] http://www.upriss.org.uk/fca/fca.html
[25] http://www.collectionscanada.gc.ca/jsc/docs/5rda-parta-categorization.pdf
[26] This has been awaiting a suitable funding opportunity.
[27] http://www.w3.org/TR/WD-rdf-syntax-971002/
[28] http://www.w3.org/TR/owl-features/
[29] http://www.w3.org/2004/02/skos/

A provisional schema (without RDF syntax) is set out in table 3 in the Appendix, with an example of a term represented in the schema in table 4.

The principle extensions to the existing Framework for each term will be:

- Add **descriptors**: any number of names (eg full name, code, tag) and annotations (eg definition, comment) for each term, with type and language for each (all information derived published standards).
- Assign a **URI** to each term: taken from its schema's namespace if there is one, or an RDA/ONIX URI if not.
- Add **type** *(see table 2 below).*
- Optionally, add **schema relationships**: to establish the hierarchy of a term within a vocabulary within an element within a schema, essential for contextual crosswalk mapping *(see 6.3.2).*
- Optionally, add **ontological relationships**: for example, subclass and subrelator hierarchies respectively for attributes and relators.
- Optionally, add **mapping relationships**: for example, "is same as", "best mapped to" .
- Add **authority statement** to mappings: for example, "authorised", "provisional" according to whether the mapping is endorsed by the source schema.

Table 2: Types of term

This table shows the proposed types of terms to be included in the RDA/ONIX Framework.

Term type	Description
schema	A metadata schema in which a vocabulary is used (for example, *MARC21, ONIX for Books*). A schema is a specification of a set of data elements and their relationships. A schema may be represented as an XMS or database schema, or as a set of abstract terms and relations set out in a Word document. A metadata schema is a specific representation of metadata, and so one standard may include multiple different metadata schemas (for example, the DDEX standard includes several distinct XML schemas representing different messages). A set of schemas may also be described as a schema (for example, *ONIX* is set of schemas including *ONIX for Books, ONIX for Serials* etc which may use the same vocabularies).
element	A data element in a schema which may have different values in specific documents or messages which conform to the schema (for example, *EditionTypeCode* in ONIX for Books).
vocabulary	A set of defined terms (for example, *CodeList21* in ONIX for Books) each of which may be used as values of an element (for example, *CodeList21* is the vocabulary used for the element *EditionTypeCode*).
vocabulary term	A defined term which is a member of a vocabulary (for example, *Unabridged* in CodeList21 in ONIX for books).
attribute	A vocabulary term representing a concept which may be an attribute of an entity.
category	A vocabulary term representing a type of entity according to the combination of its attributes.
relator	A vocabulary term representing a type of relationship between two entities.
verb	A vocabulary term representing an action or state.

Each vocabulary term is defined within a hierarchical nesting of terms of four different types, for example:

scheme (eg *ONIXForBooks*)

element (eg *EditionTypeCode*)

vocabulary (eg *CodeList21*)

category (eg *Unabridged*)

See 6.3.2 contextual mapping for the rationale for this hierarchy in relation to term mapping.

5.2.3 Extensions to content

The existing content of the Framework is currently a set of Carrier and Content Vocabularies containing attributes and some exemplary categories. A number of categories have been defined in RDA vocabularies using the Framework but these have not been formally added to it as there is no mechanism to do so.

The extensions to the content will be as follows:

- Add selected attribute, category and relator vocabularies from source standards
- Extend Framework attribute types and lists to support new concepts in additional vocabularies, creating subclass and subrelator hierarchies where needed
- Map source standard attributes to equivalent Framework attributes
- Map source standard categories and relators to Framework attributes by Formal Concept Analysis, creating additional Framework relators, categories and attributes where needed: each term in a mapped vocabulary will have a corresponding identity in the RDA/ONIX Framework
- Add verbs to support relators

5.2.4 Criteria for selecting vocabularies to be added

Vocabularies to be added to the Framework will be taken from the source standards, and will be limited to those which describe or are required to support:

- resource categories based for content and carrier attributes
- resource-to-resource relationships
- party-to-resource relationships

The criteria cover relationships which are either permanent (as with a translation of an original work) and dynamic (as with the relationship of a web content to a web site).

The aim is to include all vocabularies from the source standards which meet these criteria, but priority will be given to those which are most immediately valuable to the JISC community (for example, RDA vocabularies will take precedence over DDEX).

The Framework will also not attempt to be exhaustive in relation to the source standards: highly specialized vocabularies which are referenced only in one standard will not necessarily be mapped, although they will be referenced in the Framework. For example, ONIX contains a number of Code Lists for a large number of detailed categorizations of Bibles which do not, it appears, map to any other source standard at this point.

5.2.5 Future extensibility

In principle there are no limitations on the addition of other vocabularies to the Framework in the future, from the source standards, other standards or proprietary schemas; nor is there a limitation on adding other types of attribute, category or relationships (for example, the relationships of resources with events, states, places and times).

6. Further information

This section contains more detailed background and rationale for aspects of the project. Each of these sections is referenced in the previous sections.

6.1 Metadata vocabularies

6.1.1 Description of a vocabulary

A vocabulary is a set of defined terms. It may be known as a *controlled vocabulary, code list, allowed value set, XML enumeration list* or by other names. Ideally each term in a vocabulary is clearly defined, although quite often (especially in older standards) they simply rely on a user's understanding of a single word or phrase.

Metadata schemas may be said to have four main components: *syntax, structure, element types* and *value types*. The last three all contain some of the meaning of the metadata. Vocabularies are one of the two main value types (the other being literal strings). Vocabularies therefore provide a critical part of the meaning of metadata, but not all of it.

Some metadata standards (such as ONIX, MARC or RDA) have dozens of vocabularies covering hundreds or thousands of terms. Others (like METS or PREMIS) are concerned with structure and elements and have few explicit vocabularies.

Vocabularies are invaluable for accurate and consistent searching, querying and categorization

There is a tendency across all metadata developments now to use vocabularies wherever possible rather than relying on uncontrolled literal strings whose meanings are inconsistent and normally require individual human interpretation.

Increasingly, metadata schemas enable users to use vocabularies from different schemas, often using XML namespaces and URIs to identify them. ISO Language, Currency and Territory Codes are used by many different schemas. The Library of Congress and Dewey classification systems.

6.1.2 Growth of vocabularies

The number and size of standard vocabularies is growing steadily.

New vocabularies appear regularly. For example, DDEX, the music industry message standards just entering implementation, contains 63 different vocabulary lists containing approximately 580 distinct terms. We identify over 40% of these are of bibliographic interest[30] (as shown in Appendix 1).

RDA will be publishing a substantial set of new vocabularies in 2008 or 2009, as will MARC.

Existing vocabularies are also growing steadily. For example, the ONIX vocabularies ("Onix Code Lists") grew 10% between 2005 and 2007 and a further 10% in the last year (now 2457 terms in 99 vocabulary lists).

Not all metadata standards include many vocabularies. For example, standards such as METS[31] and PREMIS[32] are concerned principally with structure, and although they support vocabularies they not define many but rely on vocabularies from other schemes. Other standards such as ISAD[33] and to some extent MARC and Dublin Core have relied more on literal descriptive values which inhibit automated interoperability. The trend in metadata is towards using vocabularies wherever possible.

[30] We use"bibliographic interest" and "bibliographic metadata" as general terms to refer to metadata of value to the JISC community.
[31] http://www.loc.gov/standards/mets/
[32] http://www.loc.gov/standards/premis/
[33] http://www.icacds.org.uk/eng/standards.htm and http://www.ica.org/en/node/30000

There are trends for both *convergence* and *divergence* in the development and use of vocabularies. The RDA and DC initiatives, for example, are tending to convergence, with the creation and consolidation of substantial vocabularies for widespread use in particular domains.

On the other hand, the ever-increasing growth and specialization of multimedia content, format and delivery methods means that the creation and growth of new, domain- and function-specific standards and vocabularies is inevitable. The number of standards (ISO and other) incorporating vocabularies has exploded in the last decade (as indicated by the tables in the appendix) and there is no reason to expect anything different in the next. The growth of variety and complexity in metadata naturally mirrors the growth of variety and complexity of domains and technology.

6.1.3 Increasing complexity and granularity of vocabularies

The rapid expansion of digital multimedia has brought a corresponding increase in the complexity and granularity of vocabularies, for example:

- more granularity (identity and relation of parts, versions, excerpts, modifications)
- more types of creations (abstract, digital, physical)
- more technical formats and details
- more types of aggregation, sets, collections
- more rapid change
- more internationalisation (multiple languages, formats)
- more links, more metadata contributors
- more complex rights metadata

and so on.

6.1.4 The growing importance of relator vocabularies

Relators (also known variously as *properties, relations* or sometimes *predicates*) describe the relationships between the things identified by referents. For example, the relator *IsLimitedEditionOf* may describe the relationship of two books identified with ISBNs:

ISBN XXXXXXXXXX *isLimitedEditionOf* ISBN YYYYYYYYYY

Metadata standards have traditionally been based on classes or types attached to a single data element (for example *LimitedEdition* as a category of a book). The trend is now towards using entity relationships defined by relators wherever possible. In particular, Semantic Web metadata is relationship-based through RDF.

We understand that RDA and MARC are each producing significant relator vocabularies. CIDOC CRM (and therefore FRBRoo) is predominantly based on relators.

6.1.5 The value of ontology for vocabularies

Ontology techniques and tools, such as CFA used in the RDA/ONIX Framework or the Semantic Web standard OWL language, enable relationships between vocabulary terms to be formally expressed and computed on in ways that are valuable for efficiency and accuracy in metadata use.

For example, when searching for the works of a particular person, a user may wish to include all kinds of creations to which they have contributed, or to limit the search to specific types (say, books of all kinds but not music or audiovisual, or non-fiction but not fiction) or to particular roles played (for example, all writer roles but not producer or director roles). There may be dozens or even hundreds of possible categories or roles at different levels of detail. How does the user know

which to include or exclude? If these are organized in a hierarchical matrix such as RDA/ONIX, then a system can enable a user easily to select groups of related hierarchical terms and achieve the most complete but refined searches possible. Without it, results are piecemeal, with omissions and with unwanted inclusions (like an “Amazon.com” search on “John Smith” as Author).

Such hierarchical techniques are common in search and query tools, but their effectiveness is entirely dependent on the availability of the underlying vocabulary structure.

An ontological approach is becoming the norm for robust, complex metadata schemas, for example:

- FRBR (in FRBRoo) has adopted the CIDOC CRM as its underlying data model. The CIDOC CRM is an object model whose vocabularies are developed using ontological principles.
- The DDEX message standard is maintained and developed in a formal ontology.
- The ONIX for Licensing Terms standard used an ontology as its start point.

6.2 Relators in the RDA/ONIX Framework

6.2.1 Relator hierarchies

The relator vocabulary will be hierarchical, supporting a relatively small number of high level general relationships (eg *IsPartOf*) and their more specialized children (eg *IsChapterOf*). The hierarchy will allow for multiple parentage to deal with complex relations (eg *isCompressedAudioClipOf* might be a child of both *isAudioClipOf* and *isCompressionOf*).

6.2.2 Relator definition

Relators will be defined in relation to the categories of their *domain* (subject) and *range* (object) classes, and by the *verbs* which define their underlying action or state. For example, a relator *isAdaptationOf* may defined as the relator between one *text* and another which it *adapts.*

All required domain and range classes will therefore be added to the category vocabulary. In addition, ISO standard identifier types (ISBN, ISRC, ISSN etc) will be added where possible as categories so that the Framework can support the definition of relationships between resources identified with these.

6.2.3 Relator names

Relators will be named for both directions (eg *isAdaptationOf* and *hasAdaptation*).

6.2.4 Relators between resources

The Framework will cover any kind of relationship between two types of resource referenced in the main source standards. We expect that there will be 20-30 high level relators and 100-200 more specialized. Examples of types of resource relationships are shown in table 7 in the Appendix.

As the definitions of relators are dependent upon the attributes of the things they link (see 6.2.2), additional resource categories will be defined in the Framework to support roles required to support relators. To say, for example, that a relator *isRecordingOf* links a *work* to a *recorded performance* requires that a *work* and a *recorded performance* are already defined.

6.2.5 Relators between resources and parties

Parties are defined as individuals[34], groups of individuals or organizations. As the vocabulary is resource-centric, relators covered in the vocabulary will be those between parties and resources, where a party plays a role as a creator, contributor, publisher, owner, collector or otherwise affects a resource.

As with resource relators, we expect there will be 20-30 important general relators (such as *hasCreator, hasAuthor, hasContributor, hasTranslator, hasDirector, hasProducer, hasPerformer, hasPublisher, hasSupplier, hasDistributor, hasRightsController)* and several hundred more specialized relators lower in the hierarchy for more refined use. Examples of types of party-to-resource relationships are shown in table 7 in the Appendix.

6.2.6 Verbs

The meaning of a resource relator (for example, *isAdaptationOf*) may be directly linked to the meaning of a relator between a party and a resource (for example, *hasAdaptor*) and both come (directly or indirectly) from the same underlying verb or verbs *(adapt* in this example). This principle is explicit in both the <indecs> and CIDOC reference models. The task of defining relators therefore includes the definition of the underlying verbs and their hierarchies. Examples of possible verbs are shown in table 7 in the Appendix. The MPEG21 RDD set of verbs will be taken a start point.

6.3 Schema-to-schema mapping with the RDA/ONIX Framework

There are numerous "crosswalks", mappings or transforms available for schema-to-schema mapping[35]. Some of these are usable tools, while others are specifications or guides. The RDA/ONIX Framework as proposed will be a tool, but **not** to provide a complete transformation, only for the mapping of vocabularies between schemes. It can be incorporated as a part of a complete crosswalk or transform.

6.3.1 The value of "hub-and-spoke" mapping

Two common, practical problems with schema and vocabulary mapping are

- **combinatory growth**: if you try and map (say) five schemas to each other, you will need 10 different "one-to-one" schemas, and any change to one schema will require changes to four mappings; add just one schema and you need five more mappings, and so on;
- **semantic loss**: any two vocabularies being mapped will have limited semantics, and there will be obvious gaps.

However, if each vocabulary is mapped to a central schema like the RDA/ONIX Framework which is specifically structured for mapping to accommodate more or less any concept, then these problems can be overcome as far as is possible. This is sometimes known as a "hub-and-spoke" mapping approach. The Framework will enable the computation of the "best fit" mapping for any term in a mapped schema with any other mapped schema, using the Framework's ontological relationships. This can solve the "semantic loss" problem as well as possible (no Framework can compensate for meaning which is not there in the original, or is not derivable from it), and will

[34] Typically meaning individual human beings, but party can include animals (for example, performers such as Lassie) and fictional characters (such as Homer Simpson or Kermit the Frog) who may be credited with performer or other roles in relation to resources.

[35] For example, Onix to Marc XML, Library of Congress, http://www.loc.gov/marc/onix2marc.html and http://www.editeur.org/onixmarc.html
Dublin Core to Marc 21, Library of Congress, http://www.loc.gov/marc/dccross.html
Marc 21 to Dublin Core, Library of Congress, http://www.loc.gov/marc/marc2dc.html
Dublin Core to IEEE LOM http://www.imsproject.org/metadata/imsmdv1p2p1/imsmd_bestv1p2p1.html#1242547
Marc to IEE LOM http://cat.inist.fr/?aModele=afficheN&cpsidt=15811705 (French)
List of crosswalks from previous JISC project http://www.ukoln.ac.uk/metadata/interoperability

eliminate the issue of combinatory growth, as each schema requires only a single mapping to the Framework to enable mappings to any other mapped schema

6.3.2 Contextual mapping

It is not sufficient to map the terms of a vocabulary to the Framework without reference to the context in which they are used: specifically the element of the schema to which the vocabulary is being applied. The reason for this is that a single vocabulary may be applied to several elements in a schema, and it may have a different meaning, and therefore require a different mapping, in each case.

For example, in a particular scheme[36] a vocabulary of contributor roles (*author, editor* etc) may be used in one element to show a the role played by a party in relation a particular resource, and in another element to categorise a party according to the roles with which they are commonly associated. For example, Aldous Huxley may be described as the *author* of the specific work *A Brave New World*, and he may also be categorised as an *author* in general, using the same vocabulary. In the first case, the scheme term *author* is mapped to an RDA/ONIX relator (such as *is author of*), and in the second case to an RDA/ONIX category (for example, *a creator of words*).

For this reason, each term in the RDA/ONIX Framework is defined by a nesting

6.3.3 Deriving schema-to-schema mappings automatically

The Framework supports the generation of mappings between terms from different schemes using basic ontological inference on subclass and subrelator hierarchies. Where two terms from different schemas are mapped to the same Framework term, an *isSameAs* equivalence can be discovered simply. Where two terms are mapped to terms which are hierarchically related, a “best fit” mapping can be discovered (for example, *AudioCD* in one schema may have a best fit mapping to *CD* in another). On occasion there may be multiple possible “best fit” mappings.

There are of course occasions where no mapping is possible. Where there is no equivalent or “best fit” map, the Framework relations can be used to determine the closest mappings where there is some but not complete commonality of attributes.

Where the target scheme for a mapping uses literal text rather than a vocabulary for a particular element, the names of the vocabulary terms from the source scheme can be used as literals. The second pair of lists in use case 1 provides an example of this.

6.3.4 Mapping local and proprietary vocabularies

Although the Framework initially includes only standard vocabularies, it is open to anyone to include proprietary or “de facto standard” vocabularies to support transformations. These can be registered publicly in the Framework for use by third parties, or simply created for local use in the users’ own copy of the Framework. The structure of the Framework schema and its facility for allowing different authorization types would allow a local user to supplement the existing Framework while keeping the standard and local vocabularies and mappings distinct from one another.

6.4 Support from source standard authorities

We expect to have support for the extension of the Framework, whether active and passive, from those who are governing or developing the source standards:

ONIX: The ONIX editor David Martin was one of the four members of the original RDA/ONIX Framework working group. ONIX’s custodians, EDItEUR (through Brian Green), supported the original PALS proposal.

[36] This example is taken from a proprietary scheme known to Rightscom.

DDEX: Rightscom has developed and now manages the DDEX standards, including their vocabularies.

DOI: Norman Paskin of the International DOI Foundation sponsored the original PALS proposal, and will propose that IDF hosts the RDA/ONIX Framework if this project goes forward.

RDA: Tom Delsey, Editor of RDA and principal contributor to FRBR, was one of the four members of the original RDA/ONIX Framework working group and supports its further development though is unlikely to be personally available to participate because of other commitments. The RDA Committee of Principals expressed its support for the original PALS proposal. Gordon Dunsire is co-chair of the DCMI/RDA Task Group.

FRBR: See RDA.

DCMI (Dublin Core): Gordon Dunsire is Co-Chair of the DCMI/RDA Task Group.

MARC: We have had no direct contact with MARC governance as yet. The British Library supported and hosted the original RDA/ONIX Framework and sponsored Godfrey Rust's participation.

CIDOC CRM: Martin Doerr, one of the editors of the CIDOC CRM, would have been a member of the original RDA/ONIX Framework development were it not for illness.

SCORM (IEEE LOM): We have had no direct contact as yet.

Appendix

Table 3: Provisional extended RDA/ONIX Framework schema

This table shows a proposed model for the triples required to express the extended RDA/ONIX Framework in RDF. Some modification will surely be made to this in the course of the project: it is included here for illustration. The cells highlighted in grey are headers for this table only. An example of a term represented in this schema follows (*see table 4*). The triple model (domain-relator-range) shows the structure of triples with which the details of terms may be described in an RDF representation of the Framework (the RDF syntax itself is not shown. Names preceded by a ? are names of variables.

Triple ID	Domain "Subject": a URI or a 'blank node' ID.	Relator Defines the meaning of the triple.	Range "Object": datatype is shown on the right.	Occ	Range Datatype	Comments (including suggested AllowedValues)
	Term descriptions Triples containing names and textual annotations of a term					Each term is identified with a URI. If a mapped schema has no URIs for terms, surrogates will be issued and replaced when possible.
?URI1	?TermURI1	HasName	?TermName	1-n	URI	Each term will have at least one name.
?URI2	?TermName	HasValue	?NameValue	1	Literal	Contains the text string representing the Name.
?URI3	?TermName	HasType	?NameType	1	AllowedValue	eg fullName, code, tag, URI
?URI4	?TermName	HasLanguage	?NameLang	0-1	AllowedValue	ISO Language Codes
?URI5	?TermURI1	HasAnnotation	?Annotation	0-n	URI	Each term may have any number of annotations in the form of definitions or comments drawn from the standard documentation.
?URI6	?Annotation	HasValue	?AnnValue	1	Literal	Contains the text string representing the annotation.
?URI7	?Annotation	HasType	?AnnType	1	AllowedValue	eg definition, comment
?URI8	?Annotation	HasLanguage	?AnnLang	0-n	AllowedValue	ISO Language Codes
	Term classifications and context Triples linking different term types according to their structure within a metadata schema					
?URI9	?TermURI1	HasType	?TermType	1	AllowedValue	AllowedValues: schema, element, vocabulary, attribute, category, relator, verb (*see table 1 in section 3 for definitions*).
?URI10	?TermURI1	HasSchema	?SchemaURI	1-n	URI	Linking the Term to the schema to which it belongs. This is implicit in the TermURI, but the explcit triple simplifies computing for users
?URI11	?TermURI1	HasVocabulary	?TermURI2	1-n	URI	Linking an attribute or category to the vocabulary to which it belongs.

?URI12	?TermURI1	HasElement	?TermURI3	0-n	URI	Linking a vocabulary to the schema element to which it applies. The element determines the semantic context of a vocabulary and its Terms. The meaning (and therefore the mapping) of a Term may vary according to the element to which it is applied, and so a vocabulary may have a distinct URI for each element application.
?URI13	?TermURI1	HasSubClass	?TermURI4	0-n	URI	This identifies hierarchical relationships between attributes, categories or verbs.
	Category mappings Triples linking a category to the RDA/ONIX attributes which it possesses.					**These apply only to Terms of type "category".**
?URI14	?TermURI1	HasAttribute	?TermURI5	1-n	URI	These are the "Formal Concept Analysis" mappings as shown in the existing RDA/ONIX Framework which determine the unique semantics of a category.
	Relator mappings **Ontology**					**These apply only to Terms of type "relator".**
?URI15	?TermURI1	HasDomain	?TermURI6	1	URI	This identifies the category to which the first element of a Triple must belong (for example, a "HasAuthor" relator may require that its Domain has textual content).
?URI16	?TermURI1	HasRange	?TermURI7	1	URI	This identifies the category to which the first element of a Triple must belong (for example, a "HasAuthor" relator may require that its Range is a Person.
?URI17	?TermURI1	HasSubRelator	?TermURI8	0-n	URI	This identifies hierarchical relationships between relators.
?URI18	?TermURI1	HasVerb	?TermURI9	0-1	URI	The identifies the verb from which the relator derives its core meaning.
	Term mappings					**Applies to attribute, category. These values show mappings between Terms from different schemas. Mappings involving RDA/ONIX terms are manually created. Other mappings are computed from the ontology.**
?URI19	?TermURI1	IsSameAs	?TermURI10	0-n	URI	Establishes that two Terms in different schemas are functionally identical.
?URI20	?TermURI1	HasBestMapping	?TermURI11	0-n	URI	Establishes the best mapping (typically a SuperClass or SuperRelator) between two elements where no direct equivalence exists.
?URI21	?TermURI1	HasPossibleMapping	?TermURI12	0-n	URI	Establishes a mapping that is valid in some circumstances, where it is not possible to determine
?URI22	?URI19	HasMapping Status	?StatusValue	1	AllowedValue	Establishes the authorisation status of a mapping triple, eg : Authorized, Provisional, Proprietary

Table 4: Example of a term in the provisional Framework schema

This table gives an example of triples for a term from an ONIX vocabulary expressed according to the schema above, using "fake" URIs for illustration (note: correct HTTP syntax for URIs has not been used to avoid creating spurious hyperlinks in this document!).

Triple ID	Domain	Relator	Range
	Schema: ONIX		
//rdaonix/2	//onix	HasName	//onix/acronym
//rdaonix/3	//onix/acronym	HasValue	**ONIX**
//rdaonix/4	//onix/acronym	HasType	**Acronym**
	Element: ONIX/**Product_EditionTypeCode**		
//rdaonix/6	//onix/productEditionTypeCode	HasName	//onix/productEditionTypeCode/name
//rdaonix/7	//onix/productEditionTypeCode/name	HasValue	**Product_EditionTypeCode**
//rdaonix/8	//onix/productEditionTypeCode/name	HasType	**Name**
//rdaonix/9	//onix/productEditionTypeCode	HasType	**element**
//rdaonix/10	//onix/productEditionTypeCode	HasSchema	//onix
	Vocabulary: ONIX/Product_EditionTypeCode/**CodeList21**		
//rdaonix/12	//onix/productEditionTypeCode/codelist21	HasName	//onix/productEditionTypeCode/codelist21/nam
//rdaonix/13	//onix/productEditionTypeCode/codelist21/name	HasValue	**CodeList21**
//rdaonix/14	//onix/productEditionTypeCode/codelist21/name	HasType	**Name**
//rdaonix/15	//onix/productEditionTypeCode/codelist21	HasType	**vocabulary**
//rdaonix/16	//onix/productEditionTypeCode/codelist21	HasSchema	//onix
//rdaonix/17	//onix/productEditionTypeCode/codelist21	HasElement	//onix/productEditionTypeCode
	Category: ONIX/Product_EditionTypeCode/CodeList21/**Unabridged**		
//rdaonix/19	//onix/productEditionTypeCode/codelist21/ubr	HasName	//onix/productEditionTypeCode/codelist21/ubr/fullName
//rdaonix/20	//onix/productEditionTypeCode/codelist21/ubr/fullName	HasValue	**Unabridged**
//rdaonix/21	//onix/productEditionTypeCode/codelist21/ubr/fullName	HasType	**FullName**
//rdaonix/22	//onix/productEditionTypeCode/codelist21/ubr/fullName	HasLanguage	**ENG**
//rdaonix/23	//onix/productEditionTypeCode/codelist21/ubr	HasName	//onix/productEditionTypeCode/codelist21/ubr/tag
//rdaonix/24	//onix/productEditionTypeCode/codelist21/ubr/tag	HasValue	**UBR**
//rdaonix/25	//onix/productEditionTypeCode/codelist21/ubr/tag	HasType	**Tag**
//rdaonix/26	//onix/productEditionTypeCode/codelist21/ubr	HasAnnotation	//onix/productEditionTypeCode/codelist21/ubr/description
//rdaonix/27	//onix/productEditionTypeCode/codelist21/ubr/description	HasValue	**Where a title has also been published in an abridged edition; also for audiobooks, regardless of whether an abridged audio version also exists.**

//rdaonix/28	//onix/productEditionTypeCode/codelist21/ubr/description	HasType	**Description**
//rdaonix/29	//onix/productEditionTypeCode/codelist21/ubr/description	HasLanguage	**ENG**
//rdaonix/30	//onix/productEditionTypeCode/codelist21/ubr	HasType	**category**
//rdaonix/31	//onix/productEditionTypeCode/codelist21/ubr	HasSchema	//onix
//rdaonix/32	//onix/productEditionTypeCode/codelist21/ubr	HasVocabulary	//onix/productEditionTypeCode/codelist21
//rdaonix/33	//onix/productEditionTypeCode/codelist21/ubr	IsSameAs	//marc21/deweyNumberInd1/typeOfEdition/fullEdition
	Schema: MARC21		
//rdaonix/35	//marc21	HasName	//marc21/acronym
//rdaonix/36	//marc21/acronym	HasValue	**Marc21**
//rdaonix/37	//marc21/acronym	HasType	**Acronym**
	Element: MARC21/**DeweyDecimalClassificationNumber_Ind1**		
//rdaonix/39	//marc21/deweyNumberInd1	HasName	//marc21/deweyNumberInd1/name
//rdaonix/40	//marc21/deweyNumberInd1/name	HasValue	**DeweyDecimalClassificationNumber_Ind1**
//rdaonix/41	//marc21/deweyNumberInd1/name	HasType	**Name**
//rdaonix/42	//marc21/deweyNumberInd1	HasType	**Element**
//rdaonix/43	//marc21/deweyNumberInd1	HasSchema	//marc21
	Vocabulary: MARC21/DeweyNumber_Ind1/**TypeOfEdition**		
//rdaonix/45	//marc21/deweyNumberInd1/typeOfEdition	HasName	//marc21/deweyNumberInd1/typeOfEdition/name1
//rdaonix/46	//marc21/deweyNumberInd1/typeOfEdition/name1	HasValue	**Type of Edition**
//rdaonix/47	//marc21/deweyNumberInd1/typeOfEdition/name1	HasType	**Name**
//rdaonix/48	//marc21/deweyNumberInd1/typeOfEdition	HasType	**vocabulary**
//rdaonix/49	//marc21/deweyNumberInd1/typeOfEdition	HasSchema	//marc21
//rdaonix/50	//marc21/deweyNumberInd1/typeOfEdition	HasElement	//marc21/deweyNumberInd1
	Category:MARC21/EditionType**/FullEdition**		
//rdaonix/52	//marc21/deweyNumberInd1/typeOfEdition/fullEdition	HasName	//marc21/deweyNumberInd1/typeOfEdition/fullEdition/fullName
//rdaonix/53	//marc21/deweyNumberInd1/typeOfEdition/fullEdition/fullName	HasValue	**fullEdition**
//rdaonix/54	//marc21/deweyNumberInd1/typeOfEdition/fullEdition/fullName	HasType	**fullName**
//rdaonix/55	//marc21/deweyNumberInd1/typeOfEdition/fullEdition	HasName	//marc21/deweyNumberInd1/typeOfEdition/fullEdition/tag
//rdaonix/56	//marc21/deweyNumberInd1/typeOfEdition/fullEdition/tag	HasValue	**0**
//rdaonix/57	//marc21/deweyNumberInd1/typeOfEdition/fullEdition/tag	HasType	**tag**
//rdaonix/58	//marc21/deweyNumberInd1/typeOfEdition/fullEdition	HasType	**category**
//rdaonix/59	//marc21/deweyNumberInd1/typeOfEdition/fullEdition	HasSchema	//marc21

Table 5: Source Standards

This table lists the metadata standards from which relevant vocabularies will be included in the Framework.

Short name	Name	Standards body/ governance	Standard type[37]	Primary domains/ Resource category	Relevant vocabularies (approx no of terms) (C)=resource categories (R)=resource relators (P)=party relators
CIDOC CRM	CIDOC Conceptual Reference Model (ISO 21127:2006) (cidoc.mediahost.org)	ISO	Object model	Museums and archives, now Libraries	(C) Classes and Events *(tba)* (R) Properties (16) (P) Properties (36)
DCMI	Dublin Core Metadata Initiative (www.dcmi.org)	DCMI	Metadata scheme	Libraries, web metadata	(C) Type (10) (R) Qualified Relators (6)
DDEX	Digital Data Exchange message standards (www.ddex.net)	DDEX	Messaging scheme	Sound recordings and music	(C) ResourceType (8) (C) SoundRecordingType (6) (C) MidiType (2) (C) VideoType (19) (C) MusicalWorkType (23) (C) ImageType (14) (C) TextType (5) (C) SoftwareType (4) (C) AudioCodecType (10) (C) VideoCodecType (11) (C) TextCodecType (5) (C) ImageCodecType (5) (C) ReleaseType (45) (C) DistributionChannelType (10) (P) MusicalWorkContributorRole (17) (P) ResourceContributorRole (17) (P) ArtistRole (31) (P) RightsControllerRole (5) (P) AdministratingRecord CompanyRole (4)
DOI	Digital Object Identifier (www.doi.org)	IDF, ISO (TC46 SC9)*	Resource ID	Digital resources	(C) Kernel metadata proposed categories (50)
FRBR	Functional Requirements for Bibliographic Records (www.ifla.org/VII/s13/wgfrbr)	IFLA	Metadata scheme, data model (FRBRoo)	Libraries	*(tba)*
LOM	Learning Object Metadata elements and allowed values (ltsc.ieee.org/wg12/)	IEEE	Metadata scheme	Education	(C) Structure (5) (C) Status (4) (C) LearningResourceType (15) (R) Kind (6)

[37] The term "Metadata scheme" is used very broadly to cover cataloguing and database standards, frameworks, dictionaries and functional specifications which define, in whole or part, formal metadata relationships.

					(P) Role (14)
MARC21	Machine-Readable Cataloging (www.loc.gov/marc/)	MARC	Metadata scheme	Libraries	(C) *(tba)* (R) *(tba)* (P) Relator Code List (225)
ONIX	ONIX message standards (www.editeur.org).	EDItEUR	Messaging scheme(s)	Text publishing	(C) List 7 Product Form (100) (C) List 10 ePublicationType (17) (C) List 11 ePublication Format (10) (P) List 17 Contributor role (90) (C) (R) List 21 Edition Type (25) (C) (R) List 25 Illustration/other content type (28) (C) List 28 Audience (7) (C) List 34 Text format (10) (C) List 38 Image/Audio/Video file type (10) (C) List 39 Image/Audio/Video file format (7) (C) (R) List 42 Text item type (10) (P) List 45 Publisher role (8) (P) List 42 Supplier role (7) (R) List 51 Relation (16) (R) List 73 Website Role (18) (C) List 78 Product form detail (9) (C) List 80 Product packaging type (25) (C) List 81 Product content type (C) List 99 Product form feature
RDA	(www.collectionscanada.gc.ca/jsc/rdaprospectus.html)	RDA Joint Steering Committee	Cataloguing standard	Libraries	(C) Media Type (8) (C) Carrier Type (6) (C) Content Type (10) Applied Material Base Material Capture Detail Coverage of Content Digital File Characteristics Form of Work Generation Intended Audience Mount Nature of Content Notational System Polarity Production Method Projection Characteristics Sound Characteristics System of Organization Technique of Image Version *Relator vocabularies tba*

Table 6: Secondary Source Standards

This table lists some other standards parts of which may be reviewed to support formal concept analysis, some of which are candidates to be included in the Framework in future.

TC46 SC9	080117 v1.0 TC46 SC9 Identifier Interoperability Scope. TC46/SC9 Glossary of Terms (www.collectionscanada.ca/iso/tc46sc9/standard/glossary.htm)	Identifier standards	IDs	Various	
ISAN	International Standard Audiovisual Number	ISO (TC46 SC9)	Resource ID, metadata scheme	Audiovisual works	(P) Participants (10+) *Others tba*
ISBN	International Standard Book Number	ISO (TC46 SC9)	Resource ID	Books	
ISCI	International Standard Collection Identifier	ISO (TC46 SC9)	Resource ID, metadata scheme	Collections	
ISMN	International Standard Music Number	ISO (TC46 SC9)	Resource ID	Printed music	
ISRC	International Standard Recording Code	ISO (TC46 SC9)	Resource ID, metadata scheme	Sound recordings	
ISSN	International Standard Serial Number	ISO (TC46 SC9)	Resource ID	Serials	
ISTC	International Standard Text Code	ISO (TC46 SC9)	Resource ID, metadata scheme	Textual works	
ISWC	International Standard Work Code	ISO (TC46 SC9)	Resource ID , metadata scheme	Musical works	
ACAP	Automated Content Access Protocol (www.the-acap.org)	ACAP	Metadata scheme	Online permissions and policies for media access and usage (all publishing sectors, other media)	
ARK	Archival Resource Key	California Digital Library	Resource ID	Archives	
CAE/IPI	Compositeur/Editeur/Auteur Number, Interested Party Identifier	CISAC	Party ID	Musical Work Contributors and Owners	
CODEN	Serial Number	ASTM	Resource ID	Copyright societies, music publishers	
DUNS	Data Universal Numbering	Dun & Bradstreet	Party ID	Companies	

	System				
EAN	European Article Number	GS1	Resource ID	Products	
FRAD	Functional Requirements for Authority Data (associated with FRANAR, www.ifla.org/VII/d4/wg-franar.htm)	IFLA	Metadata scheme	Libraries	
GRID	Global Release Identifier (www.ifpi.org/content/section_resources/grid.html)	IFPI	Resource ID	Recording industry	
GS1	GS1 Company Prefix	GS1	Party ID	Companies	
GTIN	Global Trade Item Number	GS1	Resource ID	Products	
Handle	Handle System	CNRI	Resource ID	Digital Resources	
Indecs	The <indecs> Model of Making (page 26 & 27 of Rust & Bide (2000) *The <indecs> metadata framework: Principles,model and data dictionary*)	Content industries	Metadata specification	Content industries, libraries	
INFO	INFO URI scheme	NISO	Resource ID	Web resources	
ISNI	International Standard Name Identifier	ISO (TC46 SC9)*	Party ID	Creators, Publishers	
LCCN	Library of Congress Control Number	Library of Congress	Resource ID	Libraries	
Loc Name Authority	Library of Congress Name Authority	Library of Congress	Party ID	Creators	
MPEG-21 (part 6)	Rights Data Dictionary (www.iso21000-6.net)	ISO	Metadata scheme	Content rights, Digital items	MPEG21 verbs will be used to develop the candidate relator set.
MPEG-7	Moving Picture Experts Group (www.chiariglione.org/mpeg/)	ISO	Vocabularies	Digital items	
MWLI	Musical Work Licensing Identifier (see www.cisac.org)	CISAC	Agreement ID	Music industry	
NewsML	(www.newsml.org)	NewsML	Messaging scheme	News	
PII	Publisher Item Identifier	None	Resource ID	Text publishing	
PRISM	PRISM messaging standard (www.prismstandard.org)	IDEAlliance	Messaging scheme	Magazines and Journals	PRISM has nine vocabularies which will be reviewed.
PURL	Persistent Uniform Resource Locator	OCLC	Resource ID	Libraries	
RDA/ONIX	(www.collectionscanada.gc.ca/jsc/rdaonixann.html)	RDA, EDItEUR	Metadata framework	Libraries, content providers	
SICI	Serial Item and Contribution	ANSI/NISO	Resource ID	Journals	

	Identifier				
SMPTE	(www.smpte.org)	SMPTE	Metadata dictionary	Television	
UMID	Unique Material Identifier	SMPTE	Resource ID	Television	
UPC	Universal Product Code	GS1	Resource ID	Products	
URL	Uniform Resource Locator	IETF	Resource ID	Web locations	
URN	Uniform Resource Number	IETF	Resource ID	Web-located resources	
XRI	Extensible Resource Identifier	OASIS	Resource ID	Digital Resources	

Table 7: Illustrative resource and party relators

This table is illustrative of the scope of categories and relators to be added to the Framework. It shows examples of "high level" verbs and corresponding categories and relators, and examples of more specialized terms that may appear in the hierarchies below them. These are illustrative only: the scope of high level terms, actual term names and meanings will only be determined when the project is completed.

Verb	Resource relators	Party relators
Create		hasCreator *hasAuthor* *hasComposer* *hasIllustrator* *…etc*
Derive	hasDerivation *hasVersion* *hasAdaptation* *hasArrangement* *hasTranslation* *hasEdition* *hasRemix* *…etc*	hasDeriver *hasVersionCreator* *hasAdaptor* *hasArranger* *hasTranslator* *hasEditor* *hasRemixer* *…etc*
Aggregate	hasComponent *isCompiledFrom* *isCollectionPart* *…etc*	hasAggregator *hasCompiler* *hasCollector* *…etc*
Extract	hasExcerpt *hasFragment* *…etc*	hasExcerptor *hasFragmenter* *…etc*
Express	isManifestationOf *isRecordingOf* *..etc*	IsExpresserOf *isRecorderOf* *…etc*
Abstract	hasAbstraction	hasAbstractor
Direct		hasDirector *hasChoreographer* *hasDirectorOfPhotography* *…etc*
Perform	hasPerformance	hasPerformer *hasActor* *hasVocalEnsemble* *…etc*

Produce	hasProduct	hasProducer *hasProductionManager* *…etc*
CreateWithTool		hasEngineer *hasCameraOperator* *hasEngraver* *hasManufacturer* *creditsBy* *…etc*
Publish	Publication	hasPublisher *hasOriginalPublisher* *hasDistributor* *…etc*
Have		isPossessedBy *hasRightsController* *hasOwner* *…etc*
BeAbout	hasSubject *refersTo* *isReviewOf* *…etc*	hasSubject *refersTo* *isBiographyOf* *…etc*
Contain	hasPart *containsTrack* *containsChapter* *containsIllustration* *containsPhotograph* *…etc*	
ContainPartOf	containsPartOf *containsClipFrom* *containsMaterialFrom* *…etc*	