Implementing the Scholarly Works Application Profile in DSpace

A metadata collision analysis for the MIT Open Access Initiative

Robert Wolfe, MIT Libraries

29 August 2009

Version 2.1

Introduction

This project is undertaken in support of an effort to fulfill MIT's mandate to build an open access repository of the peer-reviewed scholarly publications of its faculty. The MIT Libraries intend to deposit publications gathered under this mandate into their DSpace@MIT institutional repository, In preparing the repository for the inclusion of this new material, an opportunity has been recognized to audit and improve the DSpace@MIT metadata tables. The goal of this audit is to ensure that metadata applied to the MIT Open Access Initiative content conforms to applicable metadata standards and application profiles from the open access scholarly publications domain. A number of best practices have been developed and have acquired wide usage in this domain since the DSpace@MIT metadata tables were last audited. It is to the advantage of the DSpace@MIT repository to become compliant with the latest open access scholarly works application profiles. The best candidate application profile for adoption, incorporation, or mapping to DSpace DC metadata is the Scholarly Works Application Profile. The Scholarly Works Application Profile is a Dublin Core Application Profile develop by JISC for use with its Eprints repository software. It is fully conformant with the Dublin Core Abstract Model and all of its extension elements have been declared in an appropriate namespace. It is a robust profile and it is targeted towards the material that we hope to include in DSpace@MIT via this OA mandate.

This document will provide the following information:

- A list of Required, Recommended, and Optional elements for the MIT OA Initiative collection, mostly chosen from the SWAP
- A collision analysis of the mapping from DSpace metadata fields to SWAP with recommendations for amendments to the DSpace DC metadata tables
- A list of Vocabulary Extension Schema to adopt with the SWAP attributes.

Important Considerations

Domain Modeling

Mapping the Scholarly Works Application Profile to DSpace begins with the mapping of entities from the SWAP's domain model to the DSpace content model. In this case the most important mapping is:

- SWAP:Expression equals Dspace:Item.

All vital metadata should be attached to the DSpace Item, even if the SWAP assigns it to a different entity in its domain model.

Metadata Element Sources

The mapping table accompanying this analysis contains a list of elements identified and named by MIT Libraries staff independent from any formal metadata standard or application profile. There are some elements in the table that currently are not defined in the SWAP. Similarly, there are elements that are not yet defined in the DSpace Metadata Tables.

Compatibility with current DSpace collections

Any amendments to the DSpace metadata tables must be backwards compatible. They must not overload an element with values that reflect two separate semantic definitions for the element. Also, they must not create a scenario where two elements share the same semantic definition and split the values for what should be a single element between them. No currently declared elements in the DSpace metadata tables will be removed or redefined. This analysis will likely recommend the addition of elements to the DSpace metadata tables.

Creation and submission of metadata conforming to this recommendation

The set of elements that are here recommended for use in the MIT Open Access repository is much larger than are usually provided for a DSpace Item. We will not use every element in this recommendation for every Item, but it is likely that we will want to use metadata fields that are currently not part of the web submission process. Alternative submission processes will likely need to be employed. Much of the metadata will be provided by publishers, also necessitating alternative submission processes. We will need to design metadata workflows and cataloging systems for these processes. It is crucial that workflows for creating or processing metadata be designed so as to provide the cataloger the ability to edit any of the metadata fields listed in this document.

Metadata Recommendations

Required Elements

Element Name	SWAP Attribute	DSpace Metadata Field
Author	Creator	dc.contributor.author
Title	Title	dc.title.none
Embargo	Date Available	dc.date.available
Type (DSpace Type Vocabulary)	Type	dc.type.none
Type (Eprints Type Vocabulary)	Type	dc.type.uri
Language	Language	dc.language.iso
Identifier	Identifier	dc.identifier.uri
MIT Affiliation	Affiliated Insitution	dc.contributor.department
Peer-reviewed Flag	Status	eprint.status
Entity Type	Entity Type	dc.type.uri
MIT Author	n/a	dc.contributor.mit
Rights Context	n/a	metsrights.contextClass metsrights.otherContextType
Rights Declaration	n/a	dc.rights.none dc.rights.uri
Rights Basis	n/a	metsrights.rightsCategory metsrights.otherCategoryType
Rights Holder	n/a	metsrights.rightsHolder.rightsHolderID metsrights.rightsHolder.rightsHolderName

Recommended Elements

Element Name	SWAP Attribute	DSpace Metadata Field
Publication Date (pub. version)	n/a	dc.date.issued
Date Submited for Publication	n/a	dc.date.submitted
Citation	Bibliographic Citation	dc.identifier.citation
Abstract	Abstract	dc.description.abstract
Publisher	Publisher	dc.publisher.none
Identifier (of Scholarly Work)	Identifier	dc.relation.isversionof
Version	Version Number or String	eprint.version
Materials Description	n/a	Bitstream Description
Journal	n/a	dc.relation.journal

Optional Elements

Element Name	SWAP Attribute	DSpace Metadata Field
Identifier – Others	Identifier	dc.identifier.other dc.identifier.none dc.identifier.govdoc dc.identifier.isbn dc.identifier.issn dc.identifier.sici dc.identifier.ismn
Subjects/Keywords	Subject	dc.subject.none dc.subject.classification dc.subject.ddc dc.subject.lcc dc.subject.lcsh dc.subject.mesh dc.subject.other
Description	Description	dc.description.none
Has Version	Has Version	dc.relation.hasversion
Editor	Editor	dc.contributor.editor
Has Translation	Has Translation	eprints.hasTranslation
Supervisor	Supervisor	dc.contributor.advisor
Sponsor	Funder	dc.description.sponsorship
Grant Number	Grant Number	eprint.grantNumber
Title of Published Version	Title	dc.title.alternative
References	References	dc.relation.references
Provenance	n/a	dc.description.provenance
Series	n/a	dc.relation.ispartofseries
Source	n/a	dc.source.uri dc.source.none

Elements that are Not Recommended

Element Name	SWAP Attribute	DSpace Metadata Field
Metadata Source	n/a	mitdspace.metadataSource
DSpace Deposit Licensen/a		DSpace Deposit License

The Collision Analysis

Elements that Require No Adjustment to DSpace DC Metadata Tables.

Though some of the paired elements/attributes have dissimilar names, all pairs have functionally equivalent semantic definitions.

Elements

SWAP Attribute	**DSpace Metadata Field**
Creator	dc.contributor.author
Title	dc.title.none
Date Available	dc.date.available
Abstract	dc.description.abstract
Publisher	dc.publisher.none
Language	dc.language.iso
Identifier [of DSpace Item]	dc.identifier.uri
Identifier [of Scholarly Work]	dc.relation.isversionof
Identifer [other Identifiers]	dc.identifier.other dc.identifier.none dc.identifier.govdoc dc.identifier.isbn dc.identifier.issn dc.identifier.sici dc.identifier.ismn
Bibliographic Citation	dc.identifier.citation
Subject	dc.subject.none dc.subject.classification dc.subject.ddc dc.subject.lcc dc.subject.lcsh dc.subject.mesh dc.subject.other
Affiliated Institution	dc.contributor.department
Description	dc.description.none
Has Version	dc.relation.hasversion
Editor	dc.contributor.editor
Supervisor	dc.contributor.advisor
Funder	dc.description.sponsorship
Title [of Published Version]	dc.title.alternative

Recommendation For These Elements

Adopt these elements for use in describing Scholarly Works deposited as part of the MIT Open Access mandate. Do not change the metadata tables, use the DSpace DC elements as currently defined.

Elements that are Not Declared in the SWAP, but are Declared in DSpace DC

Elements

Element Name	**DSpace Metadata Field**
Publication Date (of published version)	dc.date.issued
Date Submitted for Publication	dc.date.submitted
Series	dc.relation.ispartofseries

Recommendation For These Elements

Adopt these elements for use in describing Scholarly Works deposited as part of the MIT Open Access mandate. Do not change the metadata tables, use the DSpace DC elements as currently defined.

Elements that are Not Declared in DSpace DC, but are Declared in SWAP

Recommendations for these elements must be made on an element-by-element basis.

Status Element

SWAP Attribute	**DSpace Metadata Field**
Status	eprint.status

This element records whether or not a member of the collection has been peer-reviewed. It should be added to the tables. Unfortunately, there is no good dublin core element to map it to. It is the first candidate element to be added to the DSpace tables from a namespace other than the DSpace DC namespace.

Recommendation For This Element

Adopt this element for use in describing Scholarly Works deposited as part of the MIT Open Access mandate. Make this a required element for this collection. Add this element to the DSpace metadata tables. Create a new namespace in the DSpace metadata table (eprint) and put the element (status) in that namespace instead of the DSpace DC namespace. See: http://www.ukoln.ac.uk/repositories/digirep/index/Scholarly_Works_Application_Profile#Status

Has Translation Element

- **SWAP Attribute** **DSpace Metadata Field**
- Has Translation eprint.hasTranslation

There is no DSpace DC element reserved for this information. Translations are currently declared in DSpace via the dc.relation.hasversion element, which is technically correct, but imprecise. Though this element is optional for DSpace Items deposited via the OA initiative, implementation of this new element in the eprints namespace will provide highly valuable to future multi-language collections deposited in DSpace.

Recommendation For This Element

Adopt this element for use in describing Scholarly Works deposited as part of the MIT Open Access mandate. Add this element to the DSpace metadata tables. Put it in the eprint namespace (See http://www.ukoln.ac.uk/repositories/digirep/index/Scholarly_Works_Application_Profile#Has_Translation) instead of the DSpace DC namespace.

References Element

- **SWAP Attribute** **DSpace Metadata Field**
- References dc.relation.references

This element exists in qualified Dublin Core, but not in the DSpace DC tables. It should be added to the tables as soon as possible.

Recommendation For This Element

Adopt this element for use in describing Scholarly Works deposited as part of the MIT Open Access mandate. Recommend but do not require this element. Add this element to the DSpace DC tables in the DspaceDC namespace.

Elements that Pose Interesting Problems

Recommendations for these elements must be made on an element-by-element basis.

Version Element

• **SWAP Attribute**	**DSpace Metadata Field**
• Version Number or String	eprint.version

This element is not intended to define a relationship between the DSpace Item (which is a SWAP:Expression) and some other entity such as a SWAP:Manifestation or SWAP:ScholarlyWork. Therefore, the DSpace Metadata Fields dc.relation.isversionof and dc.relation.hasversion do not apply, as they record relationships between the DSpace Item and other entities.. This SWAP attribute is defined as a number or string that names the particular version of a SWAP:ScholarlyWork that the DSpace Item (SWAP:Expression) represents. There is no established practice in capturing this information in DSpace DC. The following elements all represent potential candidates, none of them are optimal:

- dc.identifier.other
- dc.relation.ispartofseries
- dc.relation.isversionof
- dc.relation.hasversion

Recommendation For This Element

Adopt this element for use in describing Scholarly Works deposited as part of the MIT Open Access mandate. Recommend but do not require this element. Add this element to the DSpace metadata tables. Put it in the eprint namespace (See http://www.ukoln.ac.uk/repositories/digirep/index/Scholarly_Works_Application_Profile#Version_Number_or_String) instead of the DSpace DC namespace. This information is vital to the success of the OA Initiative and DSpace should seek to insert some clarity into its metadata by separating these metadata values into their own element.

Define Vocabulary for this element. (e.g. preprint, postprint, published) – Contacted Robin Wendler about Harvard's vocabulary for this element.

Grant Number Element

- **SWAP Attribute** **DSpace Metadata Field**
- Grant Number eprint.grantNumber

This is another element for which there is mixed practice. The following elements have all been used, none of them are optimal:

- dc.relation.ispartof
- dc.description.sponsorship
- dc.identifier.other

Recommendation For This Element

Adopt this element for use in describing Scholarly Works deposited as part of the MIT Open Access mandate. Add this element to the DSpace DC metadata tables. Put it in the eprint namespace (See http://www.ukoln.ac.uk/repositories/digirep/index/Scholarly_Works_Application_Profile#Version_Number_or_String) instead of the DSpace DC namespace. Consider migrating all grant number metadata that is currently declared in other fields to the eprint.grantNumber field. The fields currently used to capture this information are catch-all fields that are overloaded with metadata values that ought to belong to separate elements with different semantic definitions. This information is vital to the success of the OA Initiative and DSpace should seek to insert some clarity into its metadata by separating grant numbers into their own element.

Materials Description Element

- **Element Name** **DSpace Metadata Field**
- Materials Description Bitstream Description

This element is intended to capture a description for every file that is uploaded to DSpace as part of an Item. This description is displayed in the "File in this item" section of the simple and full records for a DSpace Item. If used properly, it helps identify the correct file to download for a particular Item. This information is entered during the upload section of the DSpace Item submission process. It does not become part of the Item metadata record, rather it is stored in the Bitstream metadata.

Recommendation for this Element

Adopt a best practice of providing this information where possible. Recommend but do not require this information.

Type Elements

SWAP Attribute	DSpace Metadata Field
Type (DSpace Type Vocabulary)	dc.type.none
Type (Eprints Type Vocabulary)	dc.type.uri
Entity Type (Eprints EntityType Vocabulary)	dc.type.uri

SWAP overloads the dc:type element with two separate semantic definitions:

1) Entity Type (Values conform to the Eprints EntityType Vocabulary Encoding Scheme)

2) Type (Values conform to the Eprints Type VES)

In addition, DSpace has it's own

3) Type Vocabulary (Custom Controlled Vocabulary – String Values)

that is very similar to the Eprints Type Vocabulary. Currently we have one type element in the DSpace metadata tables--dc.type.none. If we use dc.type.none for each of these elements, then the field could be repeated three times for each DSpace Item, each repetition containing a value from a different encoding scheme. This is a legal use of Dublin Core, but it is not recommended in this case. Using the same field for values from two different vocabularies works well when the values are well-established, persistent URIs. The URIs will indicate the vocabulary encoding scheme to which they belong. It is important that we are be able to deduce which vocabulary a particular value for this element represents. The values of the two Eprints Vocabulary Encoding Schemes are URIs, while the values of the DSpace Type Vocabulary are Strings. It is important then to separate the strings from the URIs.

Recommendation For These Elements

Adopt these three elements for use in describing Scholarly Works deposited as part of the MIT Open Access mandate. Make them all required. Add a new element, "dc.type.uri" to the DSpace metadata tables, putting it in the DSpace DC namespace. Use dc.type.uri for the two Eprints Vocabulary Encoding Schemes. Constrain the values of the dc.type.none to the DSpace Type vocabulary. It is important to require both the Eprints Type URI and the DSpace Type String. One to conform with the SWAP and the other to make sure that the OA content plays nicely with the other items in DSpace.

Journal Element

• **Element Name**	**DSpace Metadata Field**
• Journal	dc.identifier.citation dc.relation.ispartofseries dc.relation.sponsorship dc.publisher.none dc.description.none dc.identifier.issn dc.relation.journal

This element records a relationship between a member of the collection and the journal in which another version of the scholarly work was published. This element is similar to the MIT Author and MIT Affiliation elements. The Journal element ought to be an identifier, preferably a URI. An element for the identifier of a publication series or journal exists in the DSpace Dublin Core – dc.identifier.issn. Not every journal will have an issn. There may also be a need to record just the name of the journal as a text string. There are five fields in the DSpace metadata record where series and journal information is placed. In the table below series and journal metadata for three different kinds of scholarly works is compared. Sample content from these fields is shown.

	Working Paper	Thesis	Published Article
dc.identifier.citation			Van Evera, Stephen. "Strategy for the Terror War". Newsday, October 4, 2001
dc.relation.ispartofseries	MIT-CEEPR;09-005WP		
dc.relation.sponsorship	Massachusetts Institute of Technology. Center for Energy and Environmental Policy Research.		
dc.publisher.none		Massachusetts Institute of Technology	Newsday
dc.description.none		Thesis (B.S.)--Massachusetts Institute of Technology, Dept. of Aeronautical Engineering, 1929.	

The dc.relation.sponsorship and dc.publisher.none fields contain very similar information for working papers and theses. This is unfortunate, but not relevant as the information is not quite a Series or Journal name. These fields are meant to capture the names of organizations, not their publications. Unfortunately, some published articles have put the names of Journals into the dc.publisher.none field. This is semantically incorrect, though occasionally the name of the publisher and the name of the journal are identical. The dc.relation.ispartofseries and dc.description.none field both contain series statements. These statements include volume and sequence information, which is more information than is wanted for the journal element. The dc.identifier.citation element likewise contains more information than is wanted.

Harvard University has defined a new element for this information in their DASH metadata element set. This element is dc.relation.journal. I have asked Robin Wendler, Metadata Analyst for the Harvard University Libraries to share with us their definition for this element. I am waiting for her reply. I hope that we can share the same practice for this element.

Recommendation For This Element

Adopt this element for use in describing scholarly works that are part of the MIT Open Access Initiative. Recommend this element but do not require it. Add this element to the DSpace DC tables in the DspaceDC namespace.

Best practice recommends the use of an authorized heading for entries in the dc.relation.journal field.

Provenance Element

- **Element Name** — **DSpace Metadata Field**
- Provenance — dc.description.provenance

DSpace uses the dc.description.provenance field to automatically record changes to the chain of custody of the DSpace Item, including changes to the Item. DSpace record only that part of the history of the scholarly publication that occurs within the DSpace system. Any pertinent content or chain of custody changes that occur outside the DSpace system should also be recorded in this field.

Recommendation For This Element

Adopt this element for use in describing scholarly works deposited as part of MIT's Open Access mandate. DSpace has named this field dc.description.provenance, mistaking provenance for a qualifier of the description element. In fact, provenance does not refine the description element and the field should be renamed to dc.provenance.none. This element is optional.

Depositing Author Element

- **Element Name** — **DSpace Metadata Field**
- [Depositing Author/Approving Author]

Ask Ellen about necessity of having this metadata in DSpace Item record.

Recommendation For These Elements

Waiting to discuss this element with Ellem Duranceau. Will make a recommendation after that discussion.

MIT Author Element

• **Element Name**	**DSpace Metadata Field**
• MIT Author	dc.contributor.mit

The purpose of this field is to be able to distinguish authors of scholarly works that are affiliated with MIT. The need for this field is the desire to provide a service that tracks and publishes information on the publication histories of MIT affiliated scholarly authors. The MIT Open Access Initiative Repository records affiliations in DSpace Item records at the department level.

There are a few considerations for the implementation of this element. One is the fact that affiliations change. An author may join or leave MIT. Is it more appropriate to identify an author as MIT affiliated on all of his publications, or just those that were published during the time that she was affiliated with the Institute?

Another consideration is the relationship of this element to the Affiliated Institution element. That element records the department name of authors that have MIT affilations. It is important that the relationship of author to department be recorded. It is possible that the recording method of this relationship may also be used to distinguish the MIT authors of a scholarly work without the necessity for an additional element in the metadata record.

A third consideration is the fact that another metadata element has already been defined to store author identifications (dc.contributor.author). It is vital that author names appear in that field. Only entries in that element are indexed to enable browsing by author names. It is also very bad practice to use two elements to store metadata that has the same semantic definition. This leaves us with the disagreeable prospect of duplicating some of the names in the dc.contributor.author field in this new field.

This element could be defined by overloading one of the existing contributor fields in the DSpace dc namespace, creating a new field in that namespace (perhaps dc.contributor.mit), or creating a new element in an entirely new namespace (perhaps one for custom fields that Libraries have added to the DSpace metadata tables).

The contents of the field could either be the name of the author, or an identifier for the author. Entering the name in this field duplicates information in two separate fields in the record. It is preferable to use an identifier for each author.

Ideally we would create URIs for MIT affiliated authors as part of an authorities service for agents at MIT, a component of the service to track and publish information on the publication histories of MIT affiliated authors. These URIs would be used to create networked information about the authors including their affiliation to a particular department. The departments would also be given URIs. They are a kind of agent crucial to the function of the authorities and publication history services. It is possible that the SWAP Agent Entity definition could be used as an ontology for these author and department entities.

Recommendation for this element

Adopt this element for use in the metadata record of DSpace Items in the MIT OA Repository. Make this element required. Add a new field to the existing DSpace DC namespace (dc.contributor.mit) to capture this information. Create URIs for MIT affiliated authors and record these URIs in this field.

Metadata Source Element

• **Element Name**	**DSpace Metadata Field**
• Metadata Source	mitdspace.metadataSource

This element does not exist in the DSpace Metadata tables. There are a few issues with the implementation of this element.

First, the element does not describe the DSpace Item. It describes the metadata record for the DSpace Item. If you added it to the metadata record for the DSpace Item, the record would describe itself. The DSpace Item metadata record is only intended to describe one thing. It is not what the SWAP calls a description set. It is not a collection of records that each describe a different thing. Adding this element to the metadata record is not recommended.

Second the purpose of this metadata element is almost certainly only administrative in nature. The DSpace Item record's main purpose is discovery and display. There are a few elements in the record that are administrative and meant only for machine consumption. But those elements all serve a functional purpose in DSpace. It is hard to imagine what function of the DSpace repository would make use of this element. This element may be more appropriate for a separate metadata record that is deposited in the DSpace Item as a bitstream.

Finally, there is no element of this type in the Dublin Core or Eprint namespaces. In fact, there are few examples of this element type in any metadata schema. Best practice would recommend that we create and maintain our own namespace within which to define this element. This is preferable to adding yet another element to the DSpace Dublin Core namespace that is not found in the proper Dublin Core namespace.

Recommendation for this element

Capture and store this information in a separate metadata record from the DSpace Item metadata. If it becomes necessary to store this metadata in the DSpace Item Record, add the element to the table in a separate namespace from the DSpace Dublin Core. This element is not recommended.

Source Element

• **Element Name**	**DSpace Metadata Field**
• Source	dc.source.uri dc.source.none

The DSpace Dublin Core metadata tables contain two elements for this kind of information-- dc.source.uri and dc.source.none. The scope notes for these elements read “Do not use; only for harvested metadata.” I have asked Richard what information is placed in these fields.

Recommendation for this element

If possible, adopt these two elements to record the source of items in the MIT Open Access repository. Whenever possible record a URI for a source. This element is optional.

Rights Elements

Element Name	DSpace Metadata Field
Rights User Group	metsrights.contextClass metsrights.otherContextType
Rights Declaration	dc.rights.none dc.rights.uri
Rights Basis	metsrights.rightsCategory metsrights.otherCategoryType
Rights Holder	metsrights.rightsHolder.rightsHolderID metsrights.rightsHolder.rightsHolderName

There are many ways that rights associated with a scholarly publication can be defined. Some of the more common are by copyright statement, license, statute, or contract. These rights can be very complicated. They can be bought and sold, licensed exclusively and non-exclusively, there can often be many claims on one publication. What is of concern to the MIT Open Access repository are the rights that it has been granted and the rights that are passed on to the users of the repository. It may not always be possible to indicate every fine detail of the rights associated with the publication, but there are a few essentials that will allow users to discover their rights.

- Always clearly distinguish between declarations of the rights of the repository and the rights of the user. Indicate the group of users for which the rights apply.
- Always indicate the basis of the rights.
- Always provide access to the contents of a rights statement, whatever its basis. Either reference the contents via an identifier or include the contents within the metadata record.
- Always provide a reference to a rights holder, some agent that can be contacted for more information.
- Whenever possible try to provide more information about the circumstances in which the rights apply (jurisdiction, dates of applicability)
- Whenever possible try to explain to the user her rights.
- Whenever possible indicate the date upon which the determination of rights was made and the agent who made the determination.
-

There are four required elements to capture:

- The group of users for whom the rights apply
- The basis of the rights (copyright, license, statue, contract)
- The content of the rights statement or its identifier
- The rights holder

The other three essentials usually require more than one element apiece.

- The circumstances under which the rights apply
- What the user can do
- The circumstances of the determination of the rights

The best practice in capturing these rights metadata elements is to group them into metadata records according to 1) the bitstream to which they apply and 2) the user group for whom they apply. This practice needs to adopted for DSpace. We have already created Items with multiple bitstreams that have different rights declarations. In our thesis collection we put two pdfs in each item, one that is available free to the public, and another that is available free to the MIT community or available for purchase by the public. There should be a rights metadata record for each bitstream. Similarly, the second bitstream described above should have two rights metadata records, one for each group of users for whom rights apply (the MIT community and the general public) The rights for each group differ. This brings us to three rights statements necessary for one DSpace Item. It is not reasonable or recommended to put all of these rights metadata records in the DSpace Item metadata. It would be more appropriate to put these rights statements in the bitstream metadata that is recorded in DSpace. It would also be appropriate to create separate xml files for each rights metadata record and deposit them as bitstreams in the DSpace Item record.

At most, one rights metadata record should be included in the DSpace Item metadata. The rights that are most important to record in the DSpace Item metadata record are those that apply to the general public and whose rights statement is the closest to a statement of Open Access. This metadata describes the rights associated with the DSpace Item that best capture the spirit of the repository. In cases like MIT Theses, where there are two rights statements for the general public, both could conceivably be included by duplicating the dc.rights.none or dc.rights.uri fields. Each statement would need to clearly identify the bitstream to which it applies.

In choosing a rights metadata schema several were considered. Examples are included in the final section of this recommendation. There are two types of rights schemas that were considered. One is simple rights metadata element sets that characterize rights for human consumption. The other is Rights Expression Languages that model rights as networked information with enough detail and complexity that this information can be consumed by Digital Rights Management systems used to enforce access to content. The use of a Rights Expression Language as expressive as ODRL is probably beyond what is possible or necessary for these scholarly publications. It is unlikely that DSpace will soon develop a Digital Rights Management system that would use rights encoded in such a language.

The SWAP and Dublin Core efforts to define rights metadata are inadequate for our needs. The PREMIS and METS communities have developed rights metadata schemas that are remarkably similar and fit our needs. The METS Rights schema is a draft schema whose maintenance is doubtful. The PREMIS Rights Entity is part of the widely adopted PREMIS schema—a stable, well-maintained standard. The PREMIS Rights Entity, unfortunately does not provide methods for identifying the rights holder or the group for whom the rights apply.

Recommendation for these elements

The rights of the repository, the DSpace Deposit License, is not the primary concern of this metadata recommendation. DSpace has an established system to record and store this information. The DSpace Item metadata record should contain information about the rights of the users of the repository. There is no need to record any metadata about the DSpace Deposit License.

Adopt the following elements from the METS Rights schema. Add these elements to the DSpace metadata tables in a new namespace "metsrights" . Make these elements required:

• *Element Name*	*DSpace Metadata Field*
• *Rights Context*	*metsrights.contextClass* *metsrights.otherContextType*
• *Rights Declaration*	*dc.rights.none* *dc.rights.uri*
• *Rights Basis*	*metsrights.rightsCategory*
•	*metsrights.otherCategoryType*
• *Rights Holder*	*metsrights.rightsHolder.rightsHolderID* *metsrights.rightsHolder.rightsHolderName*

The Rights Context fields should be used together, metsrights.otherContextType only applicable if the value of metsrights.contextClass is OTHER. METS has defined a value set for this element that we should adopt. Those values are:

- *ACADEMIC USER*
- *GENERAL PUBLIC*
- *REPOSITORY MGR*
- *MANAGED GRP*
- *INSTITUTIONAL AFFILIATE*
- *OTHER*

The Rights Declaration fields should be used to provide access to the contents of either the Copyright statement, license, statute or contract, whichever is applicable. Either or both fields should be used. One allows for the reference to a rights declaration by its identifier, the other for the inclusion of the contents in the metadata record. These elements exist in the DSpace Dublin Core metadata tables.

The Rights Basis fields should be used together, metsrights.otherCategoryType only applicable if the value of metsrights.Category value is OTHER. METS has defined a value set for this element that we should adopt. Those values are:

- *COPYRIGHTED*
- *LICENSED*
- *PUBLIC DOMAIN*
- *CONTRACTUAL*
- *OTHER*

The Rights Holder fields should be used together to provide the option to indicate the rights holder by name or by ID. Either or both fields should be used.

The other attributes and elements in the METS Rights schema ought to be added to the DSpace metadata tables in their own namespace. These elements should be available, but optional

Requested Changes to the DSpace Metadata Tables

In the existing DSpace DC namespace:

Additions

- dc.type.uri
- dc.relation.references
- dc.relation.journal
- dc.contributor.mit

Changes to element names

Current Element Name	**Recommended Name**
dc.description.provenance	dc.provenance.none

In a new Eprints namespace:

Additions

- eprint.status
- eprint.version
- eprint.hasTranslation
- eprint.grantNumber

In a new Mets Rights namespace:

Additions

- metsrights.contextClass
- metsrights.otherContextType
- metsrights.rightsCategory
- metsrights.otherCategoryType
- metsrights.rightsHolder.rightsHolderID
- metsrights.rightsHolder.rightsHolderName

New Encoding Schemes

- SWAP:Type

see: http://www.ukoln.ac.uk/repositories/digirep/index/Eprints_Type_Vocabulary_Encoding_Scheme

- SWAP:Entity Type

see: http://www.ukoln.ac.uk/repositories/digirep/index/Eprints_EntityType_Vocabulary_Encoding_Scheme

- SWAP:Status

see: http://www.ukoln.ac.uk/repositories/digirep/index/Eprints_Status_Vocabulary_Encoding_Scheme

- SWAP:Access Rights

see:
http://www.ukoln.ac.uk/repositories/digirep/index/Eprints_AccessRights_Vocabulary_Encoding_Scheme

Some Examples of rights metadata schemas in increasing order of complexity

SWAP Elements

- eprint.copyrightHolder
- dcterms.accessRights
- dcterms.license

DC Elements

- dcterms.rights
- dcterms.accessRights
- dcterms.dateCopyrighted
- dcterms.license
- dcterms.rightsHolder

METS Rights Elements

- RightsDeclarationMD (RIGHTSDECID, RIGHTSCATEGORY, OTHERCATEGORYTYPE)
  - RightsDeclaration
  - RightsHolder (RIGHTSHOLDERID, CONTEXTIDS)
    - RightsHolderName
    - RightsHolderComments
    - RightsHolderContact
      - RightsHolderContactDesignation
      - RightsHolderContactAddress
      - RightsHolderContactPhone (PHONETYPE)
  - Context (CONTEXTCLASS, CONTEXTID)
    - UserName
    - Permissions (DISCOVER, DISPLAY, COPY, DUPLICATE, MODIFY DELETE PRINT)
    - Constraints (CONSTRAINTTYPE, OTHERCONSTRAINTTYPE)
      - ConstraintDescription

PREMIS Rights Entity Metadata

- Rights Entity
  - rightsStatement (O, R)
    - rightsStatementIdentifier (M, NR)
      - rightsStatementIdentifierType (M, NR)
      - rightsStatementIdentifierValue (M, NR)
    - rightsBasis (M, NR)
    - copyrightInformation (O, NR)
      - copyrightStatus (M, NR)
      - copyrightJurisdiction (M, NR)
      - copyrightStatusDeterminationDate (O, NR)

        - copyrightNote (O, R)
    - licenseInformation (O, NR)
        - licenseIdentifier (O, NR)
            - licenseIdentifierType (M, NR)
            - licenseIdentifierValue (M, NR)
        - licenseTerms (O, NR)
        - licenseNote (O, R)
    - statuteInformation (O, R)
        - statuteJurisdiction (M, NR)
        - statuteCitation (M, NR)
        - statuteInformationDeterminationDate (O, NR)
        - statuteNote (O, R)
    - rightsGranted (O, R)
        - act (M, NR)
        - restriction (O, R)
        - termOfGrant (M, NR)
            - startDate (M, NR)
            - endDate (O, NR)
        - rightsGrantedNote (O, R)
    - linkingObjectIdentifier (O, R)
        - linkingObjectIdentifierType (M, NR)
        - linkingObjectIdentifierValue (M, NR)
    - linkingAgentIdentifier (O, R)
        - linkingAgentIdentifierType (M, NR)
        - linkingAgentIdentifierValue (M, NR)
        - linkingAgentRole (M, NR)
- rightsExtension (O, R)