

# Metadata Guidelines for Contributors to NZResearch.org.nz

March 2012



[www.natlib.govt.nz](http://www.natlib.govt.nz)

## Document Control

### Revision history

Revision	Date	Author	Reason for Change
1.0	6 May 2011	IR Meeting Contributions compiled by Leonie Hayes	Contributions from the IR Meeting held 2/5/11
1.1	20 July 2011	Following review by Otago, AUT and Waikato	Split document into relevant sections and simplify
1.2	15 March 2012	Emerson Vandy	Edits to reflect transition to production service delivery on Shared Search Infrastructure (SSI)

## Table of Contents

<b>1. INTRODUCTION.....</b>	<b>4</b>
▪ <i>Additional Metadata Guidelines for NZ Repositories .....</i>	<i>Error! Bookmark not defined.</i>
<b>2. GENERAL PRINCIPLES AND ASSUMPTIONS.....</b>	<b>4</b>
2.1. CONTRIBUTORS - INSTITUTIONAL REPOSITORIES SCOPE .....	4
2.2. ELIGIBILITY .....	4
2.3. INTEROPERABILITY STANDARDS .....	4
2.4. METADATA ELEMENTS.....	4
2.4.1. <i>Recommended Metadata Elements for SSI Contributors .....</i>	<i>4</i>
2.4.2. <i>Additional Metadata Guidelines for NZ Repositories .....</i>	<i>5</i>
<b>3. METADATA GUIDELINES FOR CONTRIBUTORS TO THE SSI ENVIRONMENT - RECOMMENDED ELEMENTS.....</b>	<b>6</b>
3.1. IDENTIFIER - MANDATORY .....	6
3.2. TITLE - MANDATORY .....	7
3.3. AUTHOR (CREATOR) - MANDATORY .....	7
3.4. ABSTRACT (DESCRIPTION) - OPTIONAL .....	7
3.5. DATE – MANDATORY .....	8
3.6. TYPE – MANDATORY .....	9
3.7. THESIS AND DISSERTATION TYPE VOCABULARY - MANDATORY.....	10
3.8. RIGHTS – OPTIONAL .....	11
3.9. COVERAGE – OPTIONAL .....	11
3.10. RELATION - OPTIONAL .....	12
3.11. SOURCE - OPTIONAL.....	13

## 1. Introduction

This document contains a set of metadata guidelines for research-focused repositories that wish to contribute metadata to the Shared Search Infrastructure (SSI) service delivered by the National Library of New Zealand. The SSI service showcases NZ Research that was previously made available in the service known as Kiwi Research Information Service (KRIS).

This document should be read in conjunction with:

- Additional Metadata Guidelines for Contributors to NZResearch.org.nz
- SSI Aggregator Guidelines for Harvesting

## 2. General principles and assumptions

### 2.1. Contributors - Institutional repositories scope

- Institutions are advised to develop policy around the scope of their repository and to align themselves with these metadata guidelines to ensure harvesting by the SSI will be uniform.
- Metadata quality is the responsibility of the contributing institution.

*Note: Metadata from other meta-aggregators like ARO or Trove should not be included because of overlap, quality and duplication issues for New Zealand contributors.*

### 2.2. Eligibility

If an institution can meet **all** of the following criteria it will be eligible for harvesting in the SSI “Research” category:

- Comply with the recommended requirements of the Metadata Guidelines – section A
- Content is stored in a repository that exports structured metadata for harvesting
- Is from one of the following Institutional types
  - NZ University
  - NZ Polytechnic
  - NZ Research Institution – eg CRI
  - NZ Government entity engaged in Research activities

### 2.3. Interoperability standards

International standards and best practices will be followed wherever possible to improve metadata re-use in global repositories and search engines.

### 2.4. Metadata Elements

The levels of importance used in this document are:

#### 2.4.1. Recommended Metadata Elements for SSI Contributors

- **Mandatory:** The field is required. If an institution cannot supply metadata to comply with the mandatory requirements the institution they will not be eligible for harvesting in the SSI “Research” category
- **Optional:** The fields are optional.

### **2.4.2. Additional Metadata Guidelines for NZ Repositories**

- Designed to support Metadata creation for New Zealand repository managers, for best practice
- Optional: All the metadata fields are optional

### 3. Metadata Guidelines for Contributors to the SSI Environment - Recommended Elements

This section lists the recommended minimum set of metadata elements required to enable an effective aggregator service. Some are listed as optional in this section as SSI aggregator service may be able to harvest these elements.

#### 3.1. Identifier - Mandatory

DC definition (Identifier): An unambiguous reference to the resource within a given context.

DC definition (Bibliographic Citation): A bibliographic reference for the resource.

Each metadata record must provide a unique HTTP URL that will serve as the primary link for the resource, and will be used to refer searchers to the metadata record at the originating repository. Multiple identifiers, including multiple HTTP URLs, may also be provided (for example, if a document is stored in multiple formats or in several parts). However, only one URL will be the primary link. This should be a “handle” if this is present.

Name	Importance in KRIS	QDC Encoding	Encoding Scheme	Format	Notes
URL	Mandatory	Identifier	URI	URL.	Each metadata record must at least include a unique HTTP URL that will be the Primary Link for the resource, used to refer users back to the source repository.
Handle <sup>1</sup>	Mandatory	Identifier	URI	Handle format (e.g. hdl:123.456).	
Citation	Optional	Identifier.Bibliographic Citation	-	Free text citation	

---

<sup>1</sup> <http://www.handle.net/> (Accessed 2011 July.)

### 3.2. Title - Mandatory

DC definition (Title): A name given to the resource.

Name	Importance in KRIS	QDC Encoding	Encoding Scheme	Format	Notes
Title	Mandatory	Title	-	Free text.	The full title of the document at item level, including any subtitle.

### 3.3. Author (Creator) - Mandatory

DC definition: An entity primarily responsible for making the resource.

Name	Importance in KRIS	QDC Encoding	Encoding Scheme	Format	Notes
Author	Mandatory (Repeatable)	Creator	-	Surname, Firstnames/Initials Corporate Author	Every document must have at least one author. Allow for the use of authors in freetext format who do not conform to Surname, Firstname/Initials

### 3.4. Abstract (Description) - Optional

DC definition (Description): An account of the resource.

DC definition (Abstract): A summary of the content of the resource.

Name	Importance in KRIS	QDC Encoding	Encoding Scheme	Format	Notes
Description	Optional	Description	-	Free Text	Plain text is the preferred format. (A blank line can be used to indicate a paragraph break.)

Name	Importance in KRIS	QDC Encoding	Encoding Scheme	Format	Notes
Abstract	Optional	Description.abstract	-	Free Text	If no abstract is available, omit this field (do not use a placeholder like “No abstract”). Plain text is the preferred format. (A blank line can be used to indicate a paragraph break.)

### 3.5. Date – Mandatory

At least one date should be present. A point or period of time associated with an event in the lifecycle of the resource.

Name	Importance in KRIS	QDC Encoding	Encoding Scheme	Format	Notes
Date issued	Mandatory	Date.Issued	W3CDTF	yyyy	Date of formal issuance (e.g., publication) of the resource. Equivalent to publication date
Date	Optional	Date	W3CDTF	W3CDTF profile of ISO 8601 <sup>2</sup> (eg: yyyy, yyyy-mm, yyyy-mm-dd, yyyy-mm-ddThh:mm+12:00, or yyyy-mm-ddThh:mmZ)	

<sup>2</sup> <http://www.w3.org/TR/NOTE-datetime>



### 3.6. Type – Mandatory

DC definition (Type): The nature or genre of the resource. For the purpose of the KRIS, the primary use of Type is to describe the type of research output (i.e. whether it is a journal article, conference proceeding, or thesis).

The EPrints Type Vocabulary (in conjunction with DCMI Type Vocabulary) is recommended because it is a standard, has well-defined terms, and is designed for use with scholarly publications. Equivalences are outlined in SSI Aggregator Guidelines for Harvesting.

Name	Importance in KRIS	QDC Encoding	Encoding Scheme	Format	Notes
EPrints Type Vocabulary	Mandatory	Type	EPrintsType	Controlled vocabulary as text (eg. Journal Article)	Designed for describing scholarly publications.
			URI	Controlled vocabulary as URI (eg. <a href="http://purl.org/eprint/type/ScholarlyText">http://purl.org/eprint/type/ScholarlyText</a> )	

Scheme	Maintainer	Size	Notes
EPrints Type Vocabulary Encoding Scheme	JISC Digital Repositories Programme	15 expandable to 27	<p>15 terms: Scholarly Text, Book, Book Item, Book Review, Conference Item, Conference Paper, Conference Poster, Journal Item, Journal Article, News Item, Patent, Report, Submitted Journal Article, Thesis or Dissertation, Working or Discussion Paper.</p> <p>Formal definitions are supplied.</p> <p>The list is a refinement of the DCMI term "Text", and is used with the DCMI Types list to produce a combined list of 27 descriptors.</p> <p>The EPrints authors are also members of the DCMI, so there is the potential parts of it may be promoted for adoption by DCMI. Also, the EPrints project is aiming to embed it in future releases of IR applications (DSpace, EPrints, Fedora, Digital Commons) in the same way that DC is native to these applications currently.</p> <p><a href="http://www.ukoln.ac.uk/repositories/digirep/index/Eprints_Type_Vocabulary_Encoding_Scheme">http://www.ukoln.ac.uk/repositories/digirep/index/Eprints_Type_Vocabulary_Encoding_Scheme</a> (Accessed 2011 July)</p>

### 3.7. Thesis and Dissertation Type Vocabulary - Mandatory

Metadata fields are based on the ETD-MS standard <http://www.ndltd.org/standards/metadata/etd-ms-v1.00-rev2.html>

Name	Importance in KRIS	QDC Encoding	Encoding Scheme	Format	Notes
thesis.degree.name	Optional		ETD-MS	Name of the degree associated with the work as it appears within the work	(example: <i>Masters in Operations Research</i> )
thesis.degree.level	Mandatory		ETD-MS	Level of education associated with the document.	<b>See below</b>
thesis.degree.discipline	Optional		ETD-MS	Area of study of the intellectual content of the document	
thesis.degree.grantor	Optional		ETD-MS	Institution granting the degree associated with the work	

Institutions that supply thesis metadata MUST use the vocabulary below to indicate the academic level of a thesis or dissertation

Term	Token	Description	Currently Used by
Post-doctoral	Postdoctoral		VUW
Doctoral	Doctoral	Work submitted towards a Doctorate or equivalent degree.	UoC, UoA, VUW, UoO
<i>UF: Awarded Doctoral Thesis</i>		<i>PBRF category</i>	
Masters	Masters	Work submitted towards a Masters or equivalent post-graduate degree.	UoC, UoA, VUW, UoO
<i>UF: Awarded Research Masters Thesis</i>		<i>PBRF category</i>	
► Dissertation	Dissertation	Includes all non-research post-grad. output: Coursework Masters, Postgraduate Diploma, Masters Research Report	UoO
Undergraduate	Undergraduate	Work submitted towards an undergraduate degree.	UoA
► Bachelors	Bachelors		VUW
► Honours	Honours		UoO, UoA

### 3.8. Rights – Optional

The EPrints AccessRights Vocabulary is recommended because it is a maintained standard, has well-defined terms, and is designed for use with scholarly publications. It is generally assumed that resources in the contributing repositories will be open-access.

[http://www.ukoln.ac.uk/repositories/digirep/index/Eprints\\_AccessRights\\_Vocabulary\\_Encoding\\_Scheme](http://www.ukoln.ac.uk/repositories/digirep/index/Eprints_AccessRights_Vocabulary_Encoding_Scheme) (Accessed 2011 July)

Name	Importance in KRIS	QDC Encoding	Encoding Scheme	Format	Notes
Rights	Optional	Rights	-	Plain text.	Information about rights held in and over the resource.
			URI	URL format	
Access Rights	Optional	Rights.AccessRights	EPrintsAccessRights	OpenAccess RestrictedAccess ClosedAccess	Information about who can access the resource or an indication of its security status.
			EPrintsAccessRights	Controlled vocabulary as URI (eg. <a href="http://purl.org/eprint/accessRights/OpenAccess">http://purl.org/eprint/accessRights/OpenAccess</a> )	This field is optional, and by default resources are assumed to be OpenAccess. If a resource is not open access, it is recommended a short, plain-text statement is included describing the rights restrictions.
			-	Free text.	

### 3.9. Coverage – Optional

Name	Importance in KRIS	QDC Encoding	Encoding Scheme	Format	Notes
Coverage	Optional	Coverage	-	Free text.	The spatial or temporal topic of the resource, the spatial applicability of the resource, or the jurisdiction under which the resource is relevant.

Name	Importance in KRIS	QDC Encoding	Encoding Scheme	Format	Notes
Coverage	Optional	Coverage.Spatial	-	Free text.	Spatial characteristics of the resource including location
Coverage	Optional	Coverage.Temporal	-	Free text.	Temporal characteristics of the intellectual content of the resource.

### 3.10. Relation - Optional

DC definition: A related resource.

Name	Importance in KRIS	QDC Encoding	Encoding Scheme	Format	Notes
Relation	Optional	Relation	URI	URL. Citation. ISSN. or ISBN.	
		Relation.hasFormat			
		Relation.hasPart			
		Relation.hasVersion			
		Relation.isFormatOf			
		Relation.isPartOf	URI	URN format (eg. urn:issn:1234-5678, urn:isbn:1234-5678)	The Journal (or Book) that an article (or chapter) is published in.
		Relation.isReferencedBy		Free text.	The bibliographic Identifier of the corresponding Catalogue record.
		Relation.isReplacedBy			
		Relation.isRequiredBy			
		Relation.isVersionOf	URI	URL.	Working papers and preprints may use to reference published equivalents.
		Relation.References			
		Relation.Replaces			
		Relation.Requires			

### 3.11. Source - Optional

DC definition: The resource from which the described resource is derived.

Name	Importance in KRIS	QDC Encoding	Encoding Scheme	Format	Notes
Source	Optional	Source	URI	URL. Citation. ISSN. or ISBN.	