

# Proposed Metadata for ZULC e-theses Repository (unpublished draft)

---

## 1. Introduction

This paper seeks to propose a metadata schema for the ZULC e-theses Repository. Institutions dealing with electronic theses and dissertations have all developed their own standards or adapted existing metadata standards. These metadata standards all attempt to describe the author, the work, and the context in which the work was produced in a way that will be useful to the researcher as well as the librarians and/or technical staff maintaining the work in its electronic form. The proposed metadata schema is informed by the ETD-MS metadata schema set (the one endorsed by Virginia Tech), the discussion on Interoperability Metadata Standard for Electronic Theses and Dissertations as found on the NDLTD website (<http://www.ndltd.org/standards/metadata/etd-ms-v1.00-rev2.html>), the metadata schema used by Georgia Tech Institutional Repository for theses (<http://smartech.gatech.edu>), the discussion on the difficulty of harvesting non-Dublin Core elements via OAI on Dspace-general mailing list, in particular the thread <http://listserv.vt.edu/cgi-bin/wa?A2=ind0506&L=etd-l&P=1551>, and, of course, taking into account of our requirements.

## 2. Extensions to the default DSpace's Dublin Core registry

The DC registry that ships with DSpace is generic. For ETD implementation, the typical practice is to introduce thesis-specific metadata elements to the default registry to capture information that pertains to **degree, department, advisor, type of document** (level of degree). The [NDLTD's discussion](#) introduces the non-DC element called theses, with the following qualifiers:

- ❖ **thesis.degree.name** - Name of degree
- ❖ **thesis.degree.level** -Level of education associated with the degree i.e. Undergraduate, Masters, PhD
- ❖ **thesis.degree.discipline** -Area of study, i. e. name of program or department
- ❖ **thesis.degree.grantor** -Institution granting the degree

As have been cited by Susan Wells Parham on the [thread](#), it is advisable not to travel too far from the Dublin Core (DC) registry originally suggested in DSpace. The Georgia Tech instance of DSpace for ETD extends the available DC elements to capture the above metadata fields.

- ❖ **description.degree** -Name of degree e.g. Master of Science, etc.
- ❖ **contributor.department** -Area of study, i. e. name of program or department
- ❖ **description.advisor** -supervisors
- ❖ **type** - For the type of document (thesis, dissertation)
- ❖ **Publisher**- Institution granting the degree

Georgia Tech implementers pointed out that they could not use the **contributor.advisor** as it had been noted to give problems by early implementers. Since the idea is to come up with a registry that is OAI-PMH compatible, it would be prudent to opt for the Georgia Tech's registry.

An important point to note is that institutions are at liberty to choose what fields to use to capture the data they think is pertinent for their ETD implementations. However, why re-invent the wheel? Just follow the footsteps of successfully ETD implementers! Given the above discussion, the proposed registry for ZULC e-theses metadata is given on the next page.

### 3. Proposed DSpace Dublin Core registry for ZULC e-theses Repository

<b>Element/Field</b>	<b>Purpose/Description</b>	<b>Notes</b>
<b>dc.contributor.author</b>	Name of the author (Surname and First name)	Mandatory. Not repeatable for a thesis
<b>dc.contributor.registrationnumber</b>	Registration Number for a student. Has been included for future use e.g. interfacing with the MIS,LDAP etc.	Mandatory
<b>dc.contributor.authoremail</b>	Email of the author (if available )	Optional
<b>dc.contributor.advisor</b>	The name of the advisors	Repeatable, mandatory
<b>dc.contributor.institution</b>	The parent institution e.g. University of Zimbabwe.	Mandatory
<b>dc.contributor.institutionemail</b>	Email of the institution unit responsible for custodianship of the theses e.g. Special Collection Unit	optional
<b>dc.contributor.faculty</b>	The Name of the Faculty e.g. Faculty of Science	Mandatory
<b>dc.contributor.department</b>	The Name of the Department e.g. Computer Science	Mandatory
<b>dc.description.abstract</b>	The abstract for the thesis	Mandatory
<b>dc.description.note</b>	Additional information regarding the thesis or dissertation. <i>Example: acceptance note of the department</i>	Optional, Repeatable
<b>dc.description.sponsorship</b>	Information about sponsoring agencies, individuals, or contractual arrangements for the item.	Optional, Repeatable
<b>dc.subject</b>	Keywords	Mandatory, Repeatable
<b>dc.title</b>	Title for the thesis	Mandatory
<b>dc.title.alternative</b>	alternative title of the thesis or dissertation	Optional, Repeatable
<b>dc.type</b>	Undergraduate Thesis/Masters/DPhil etc. mandatory	Mandatory
<b>dc.type.degree</b>	Degree type and Name e.g. Master of Science in Tropical Resource Ecology	Mandatory
<b>dc.language</b>	A language of the intellectual content of the resource. This should be the primary language in which the work is recorded	Optional
<b>dc.format</b>	The digital manifestation of the resource. In the case of an ETD, this should contain a list of the electronic format(s) in which the work is stored and/or delivered e.g. PDF.	Optional, Repeatable
<b>dc.rights</b>	Information about rights held in and over the resource. Typically, this describes the conditions under which the work may be distributed, reproduced, etc., how these conditions may change over time, and whom to contact regarding the copyright of the work.	Optional, Repeatable

<b>dc.rights.availability</b>	Scope of accessibility i.e. intranet(restricted) or Internet (unrestricted)	
<b>dc.identifier.uri</b>	An unambiguous reference to the thesis. Using DSpace Handles ( <a href="http://handle.net">http://handle.net</a> ) are recommended.	Mandatory
<b>dc.date.accessioned</b>	Date DSpace takes possession of item.	For systems use
<b>dc.date.available</b>	Date or date range item became available to the public.	Mandatory
<b>dc.date.defense</b>	Date of Defense (may be the year)	Mandatory
<b>dc.date.issued</b>	Date of publication or distribution.	Mandatory (system-generated)

#### 4. Comments

The design of DSpace is to use the element dc.publisher for capturing the name of the publisher in which an article was previously published. Thus, the option to capture publisher is given only if a submitter has indicated that the article has been previously published. It also gives the option to capture the year of previous publication and any citation. In any effort to circumvent this, we experimented with capturing the parent institution in the qualified element dc.contributor.institution and the email of the unit responsible for custodianship using the qualified element dc.contributor.institution .