



**Dictionary of Texas Digital Library Descriptive Metadata for Electronic
Theses and Dissertations, v. 2**

September 2015

Prepared by the TDL ETD Metadata Working Group

Sarah Potvin, Chair (Texas A&M University)

Kara Long (Baylor University)

Colleen Lyon (University of Texas)

Kristi Park (Texas Digital Library)

Monica Rivero (Rice University)

Santi Thompson (University of Houston)

Note: This dictionary should be read in tandem with the *Report for Texas Digital Library Descriptive Metadata for Electronic Theses and Dissertations, v. 2*.

Dictionary of Texas Digital Library Descriptive Metadata for Electronic Theses and Dissertations, v. 2

September 2015

Table of Contents

Explanation of Guidelines	3
Changes to TDL ETD standard	4
Elements	5
Title (M).....	5
Name of Author (M).....	6
Name of Thesis Advisor or Committee Chair (O).....	6
Name of Committee Member(s) (O).....	7
Type of Resource (M).....	8
Genre (M).....	8
Date (M).....	9
Language (O).....	10
Format (O).....	10
Abstract (M).....	11
Description (O).....	12
Subject (O).....	13
Identifier (M).....	14
Author Identifier - ORCID (O).....	14
Rights (O).....	15
Degree Information (M).....	16
Embargo (O, S).....	17
Provenance (S).....	18
Appendix A: Quick Reference Mapping Table	20
Appendix B: Full example of ETD encoded in TDL ETD DC.....	22
Appendix C: Crosswalks	24
Appendix E: ORCIDS and Name Disambiguation	26

Explanation of Guidelines

This document is intended to support the internal needs of TDL and also the needs of the contributing institutions. Internally, this guide will inform Vireo development and the production of ETD records in repositories, which present ETDs to their local communities and the public at large. Contributing institutions can use this document to guide and shape local metadata practices for describing ETDs. Accompanying this document, and laying out more detailed explanations for process and findings, is the *Report for Texas Digital Library Descriptive Metadata for Electronic Theses and Dissertations, v. 2*. These documents are designed to be used in tandem, and we encourage you to consult the *Report* while using this *Dictionary*.

Our work is heavily indebted to the 2008 TDL ETD standard, which laid out mappings to a newly-created "TDL ETD DC" schema, as well as to the NDLTD ETD-MS standard. We have adapted the 2008 TDL ETD standard approach of providing definitions, examples, and mappings for each element, which we see as an elegant and useful way of presenting this information.

TDL ETD DC standard defines the following 18 elements. Included in this document is an overview of each TDL ETD DC element, their qualified elements, and characteristics, and their use for describing electronic theses and dissertations. The document is structured according to the following list of main elements:

1. Title (mandatory)
2. Name of Author (mandatory)
3. Name of Thesis Advisor (optional)
4. Name of Committee Member(s) (optional)
5. Type of Resource (mandatory)
6. Genre (mandatory)
7. Date (mandatory)
8. Language (optional)
9. Format (optional)
10. Abstract (mandatory)
11. Description (optional)
12. Subject (optional)
13. Identifier (mandatory)
14. Author Identifier - ORCID (optional)
15. Rights (optional)
16. Degree Information (mandatory)
17. Embargo (optional)
18. Provenance information (system-generated)

Changes to TDL ETD standard

The 2015 standard incorporates the following changes:

- TDL ETD schema is now based on qualified Dublin Core rather than MODS. The Dublin Core standard more closely aligns with current practices of Texas Digital Library members. (See note below for further explanation.)
- The switch to DC from MODS necessitates changes in some element names; for example instead of “Origin Information,” use “Date” as top level element. Other changes include “Format” instead of “Physical Description,” the movement of “Location (URL)” to the “Identifier” element, and the removal of the “Record Information” section (as this information is recorded automatically by the repository system as <dc.date.accessioned> and <dc.description.provenance>).
- New elements and encoding guidelines have been added for rights metadata, author identifiers (ORCID), and description (non-abstract) information. Recommendations have been made for “Text” to be mapped to dc.type.dcmi instead of dc.type.material.
- To build greater awareness around best practices for aggregation of metadata and improving discoverability, additional mappings have been added (e.g. OAI-ETD and Google Scholar’s Highwire Press tags).

For a more detailed explanation of drivers for these changes, please see accompanying *Report* section on Case Studies and Recommendations.

A note on MODS

As the Introduction to the Report noted, both the 2005 and 2008 guidelines were centered around MODS application profiles, with the 2008 guidelines including flat, key-value paired Dublin Core and a thesis schema (known collectively as “TDL DC”) only for crosswalking to meet the Networked Digital Library of Theses and Dissertations (NDLTD) ETD-MS exchange standard. The current working group observed that MODS files were problematic in the context of DSpace repositories. MODS files, which are maintained as separate bitstreams, are not actionable in DSpace: the system cannot read or write to them. These separate, distinct metadata files pose concurrency problems: changes made to the flat metadata are not necessarily reflected in the MODS, and vice versa.¹ Both working group and task & review members observed that they had moved away from a practice of maintaining MODS in DSpace. Those that do produce MODS for ETDs pointed either to catalog workflows (with MODS derived from MARC rather than created in tandem with ETD submission) or to the usage of MODS exclusively in preservation repositories.

The working group recognizes that MODS remains a standard for descriptive metadata, which may smooth migrations of metadata between systems and can provide a strong basis for the creation of linked data. However, given institutional particularities and platform constraints, which essentially require MODS to be maintained separately from actionable TDL DC metadata, the working group

¹ DSpace is capable of crosswalking and exposing flat metadata as dynamically-produced MODS, but these files typically do not match the natively-created and richer MODS files that can only be maintained as bitstreams.

decided not to include an update to MODS mappings in the Guidelines. This omission does not constitute a recommendation against the use of MODS, but a pragmatic weighing of working group time and expertise and current practice of TDL repositories.

Elements

The following conventions are used to express guidelines for each metadata element:

Name of Element: The name of Dublin core element. M=Mandatory O=Optional S=System-generated.

Definition of Element: Definition of element based on ETD-MS Guidelines².

Mandatory and Optional Practice: Guidelines for describing, including whether elements or qualified elements are mandatory or optional.

Recommended data values: May include references to appropriate content standard, authority file, thesaurus, encoding standard, etc. to guide data value entry.

Example of Use: Provides examples of preferred data values within elements.

Mapping to ETD-MS: Provides encoding analog between elements in TDL ETD DC and ETD-MS. A quick reference mapping table appears in Appendix A.

Mapping to OAI-ETD: Provides encoding analog between elements in TDL ETD DC and OAI-ETD. These are suggested crosswalks to support better interoperability and global sharing of metadata outside local repositories. A quick reference mapping table appears in Appendix A.

Mapping to Highwire Press Tags³: Provides encoding analog between elements in TDL ETD DC and Highwire Press tags (Google Scholar). These are recommended “meta-tags” encoding in item page HTML header. Only tags required for inclusion in Google Scholar as of August 2015 are provided. A quick reference mapping table appears in Appendix A.

Title (M)

Definition: A name given to the resource.

Mandatory practice: Encode the complete title information in <dc.title> element.

Optional practice: n/a

Recommended data values: The title of the work as it appears on the title page proper or equivalent.

Example of Use:

<dc.title>Critical processes and performance measures for patient safety systems in healthcare institutions: a Delphi Study</dc.title>

Mapping to ETD-MS

<dc.title>Critical processes and performance measures for patient safety systems in healthcare institutions: a Delphi Study</dc.title>dc.c

² <http://www.ndltd.org/standards/metadata>

³ <http://scholar.google.com/intl/en/scholar/inclusion.html#indexing>

Mapping to OAI-ETD

<dc.title>Critical processes and performance measures for patient safety systems in healthcare institutions: a Delphi Study</dc.title>

Mapping to Highwire Press Tags (Google Scholar)

<meta name="citation_title" content="Critical processes and performance measures for patient safety systems in healthcare institutions: a Delphi Study">

Name of Author (M)

Definition: An entity primarily responsible for making the content of the resource; the author of the work.

Mandatory practice: Encode name of the author in the <dc.creator> element.

Optional practice: For additional information about the author, may link name to ORCID. For encoding of ORCIDs see <dc.identifier.orcid> element.

Note: The use of birth dates to distinguish author names is a traditional cataloging practice, but is sometimes considered invasive of personal privacy in born-digital environments. The use of this type of information may be contingent on local institutional policy. Adopting and implementing ORCID for ETD authors would provide an alternative to other forms of name disambiguation. For discussion of recommendations for institutional practice around name disambiguation, see case study section. For discussion of ORCID, see Appendix E.

Recommended data value: Expressed in the authorized form of the author's name if possible.

Example of Use:

<dc.creator>Akins, Ralitsa B., 1967-</dc.creator>

Mapping to ETD-MS

<dc.creator>Akins, Ralitsa B., 1967-</dc.creator>

Mapping to OAI-ETD

<dc.creator>Akins, Ralitsa B., 1967-</dc.creator>

Mapping to Highwire Press Tags (Google Scholar)

<meta name="citation_author" content="Akins, Ralitsa B., 1967-">

Name of Thesis Advisor or Committee Chair (O)

Definition: Name of thesis/dissertation advisor or committee chair as it appears on title page or equivalent.

Mandatory practice: N/A

Optional practice: Encode information about the thesis advisor in the <dc.contributor.advisor> element. The element <dc.contributor.advisor> is repeatable for thesis advisors.

Recommended Data Value: Expressed in the authorized form of the advisor's name if possible.

Example of Use:

<dc.contributor.advisor>Cole, Bryan R.</dc.contributor.advisor>

Mapping to ETD-MS

<dc.contributor>Cole, Bryan R.</dc.contributor>

<dc.contributor.role>Thesis advisor</dc.contributor.role>⁴

Mapping to OAI-ETD

<dc.contributor>Cole, Bryan R.</dc.contributor>

Mapping to Highwire Press Tags (Google Scholar)

N/A

Name of Committee Member(s) (O)

Definition: Name of committee member(s) found on title page or equivalent.

Mandatory practice: N/A

Optional practice: Encode name of committee members in <dc.contributor.committeeMember> element. The element <dc.contributor.committeeMember> is repeatable.

Recommended Data Value: Expressed in the authorized form of the committee member's name if possible.

Note: Some institutions elect to include other roles, such as department heads, supervisors, directors of research, in their metadata records. Vireo users should consult documentation for more information on roles collected and exported in that system.

Example of Use:

<dc.contributor.committeeMember>Davis, Eddie J.</dc.contributor.committeeMember>

<dc.contributor.committeeMember>Lincoln, Yvonna S.</dc.contributor.committeeMember>

<dc.contributor.committeeMember>Smith, Elvin E.</dc.contributor.committeeMember>

Mapping to ETD-MS

N/A

Note: Because of the lack of semantic precision for this element, we've elected not to map this element to ETD-MS.

⁴ As this example illustrates, ETD-MS specifies that contributor names be provided as values for dc.contributor and their roles specified in dc.contributor.role. Given the inability, with flat metadata, to specify which contributor is attached to which role when multiple contributors or roles are specified, we have chosen to only map advisor names to dc.contributor in ETD-MS.

Mapping to OAI-ETD

N/A

Mapping to Highwire Press Tags (Google Scholar)

N/A

Type of Resource (M)

Definition: Characteristic and general type of content of the resource.

Mandatory practice: Encode the type of resource in a <dc.type.dcmi> element. The element <dc.type.dcmi> is repeatable for ETDs with multiple files.

Optional practice: N/A

Recommended Data Value: Use the DCMI Type vocabulary.⁵ For most ETDs the type will be “Text” (First letter is capitalized, use singular form).

Example of Use:

```
<dc.type.dcmi>Text</dc.type.dcmi>
```

Mapping to ETD-MS

```
<dc.type>Text</dc.type>
```

Mapping to OAI-ETD

```
<dc.type>Text</dc.type>
```

Mapping to Highwire Press Tags (Google Scholar)

N/A

Genre (M)

Definition: A term that designates a category characterizing a particular style, form, or content. Describes the nature or genre of the content of a resource.

Mandatory practice: Encode the genre in the <dc.type.genre> element.

Optional practice: N/A

Recommended Data Value: Use the string “Thesis”. First letter is capitalized, use singular form for term.

Example of Use:

⁵ <http://dublincore.org/documents/2000/07/11/dcmi-type-vocabulary/>

<dc.type.genre>Thesis</dc.type.genre>

Mapping to ETD-MS

<dc.type>Thesis</dc.type>

Mapping to OAI-ETD

<dc.type>Thesis</dc.type>

Mapping to Highwire Press Tags (Google Scholar)

N/A

Date (M)

Definition: A date associated with an event in the life cycle of the resource. Information about the origin of the resource; in this case, specifically, the date of creation of the resource and the date of publication of the resource.

Mandatory practice: Use the primary date associated with the thesis. This is defined as the date the student graduates or the date the degree is conferred, expressed in YYYY-MM format. Encode this date in the <dc.date.issued> element. The publication date is defined as the date the ETD is released to the public. This date is provided in the <dc.date.available> element.

Optional practice: A natural language version of the graduation date may be supplied in the <dc.date.created> element. Date values are set up in the Vireo System (e.g. May or December).

System generated: Dates associated with life cycle events may be automatically added during the submission process (e.g. date deposited). The <dc.date.accessioned> and <dc.date.available> are automatically generated by the repository system (DSpace). The <dc.date.accessioned> element is the date the repository takes possession of an item. The <dc.date.available> element is the date an item became available to the public. When an item's embargo is lifted, that date is recorded for the <dc.date.available> element. These dates are expressed in extended date/time format per ISO 8601 standard.

Recommended Data Value: Express dates as specified in ISO 8601⁶.

Note! Dates as defined in this standard require Vireo system updates. For example, remapping <dc.date.issued> to graduation date instead of approval date. Please see accompanying report section on use cases for explanation of recommended changes.

Example of Use:

<dc.date.issued>2004-12</dc.date.issued>

<dc.date.created>December 2004</dc.date.created>

<dc.date.accessioned>2005-03-15T19:44:57Z</dc.date.accessioned>

<dc.date.available>2005-03-15T19:44:57Z</dc.date.available>

⁶ http://www.iso.org/iso/catalogue_detail?csnumber=40874

Mapping to ETD-MS

```
<dc.date>2004-12</dc.date>
```

Mapping to OAI-ETD

```
<dc.date>2004-12</dc.date>
```

Mapping to Highwire Press Tags (Google Scholar)

```
<meta name="citation_publication_date" content="2004">
<meta name="citation_online_date" content="2005/03/15">
```

or use single tag: <meta name="citation_date" content="2004">

Language (O)

Definition: A designation of the language in which the content of a resource is expressed.

Mandatory practice: Encode language information in the <dc.language.iso> element. The language element is repeatable.

Optional practice: N/A

Recommended Data Value: Use the ISO 639-2⁷ code for the representation of the name of the language.

Example of Use:

```
<dc.language.iso>eng</dc.language.iso>
```

Mapping to ETD-MS

```
<dc.language>eng</dc.language>
```

Mapping to OAI-ETD

```
<dc.language>eng</dc.language>
```

Mapping to Highwire Press Tags (Google Scholar)

```
<meta name="citation_language" content="eng"/>
```

Format (O)

Definition: The digital manifestation of the resource and/or a designation for the source (physical) of the digital file important to its creation, use and management.

Mandatory practice: N/A

⁷ https://www.loc.gov/standards/iso639-2/php/code_list.php

Optional practice: Format element is repeatable. Use qualifiers to express different format information, e.g. MIME type, physical origin, or extent.

Recommended Data Value: Use the Internet Media MIME types⁸ standard to express the electronic format(s) in which the work is stored and/or delivered. To indicate the source of the digital file the following values⁹ may be used: “born digital” or “reformatted digital.”

Note: Some format information can be automatically generated by the repository software (such as MIME types) at the bitstream-level. It’s debatable whether this information should then be also added to the item-level metadata. One benefit of recording this information at the item-level is that this format information can then be included in exports of metadata (such as OAI).

Example of Use:

```
<dc.format.mimetype>application/pdf</dc.format.mimetype>
<dc.format.digitalOrigin>born digital</dc.format.digitalOrigin>
<dc.format.extent>103 pp</dc.format.extent>
```

Mapping to ETD-MS

```
<dc.format>application/pdf</dc.format>
<dc.format>born digital</dc.format>
<dc.format>103 pp</dc.format>
```

Mapping to OAI-ETD

```
<dc.format>application/pdf</dc.format>
<dc.format>born digital</dc.format>
<dc.format>103 pp</dc.format>
```

Mapping to Highwire Press Tags (Google Scholar)

N/A

Abstract (M)

Definition: An account of the content of the resource.

Mandatory practice: Encode the abstract in the <dc.description.abstract> qualified element.

Optional practice: N/A

Recommended Data Value: The full text of the abstract of the thesis or dissertation.

⁸ <http://www.iana.org/assignments/media-types/media-types.xhtml>

⁹ Additional values may be taken from the MODS User Guidelines, <http://www.loc.gov/standards/mods/userguide/physicaldescription.html#digitalorigin>

Note: Mandatory if available. If an abstract is not available, do not use the phrase "Not available." Leave the abstract element empty. Do not start the field with the word "Abstract" as a declaration of content.

Example of Use:

```
<dc.description.abstract>This dissertation study presents a conceptual framework for implementing and assessing patient safety systems in healthcare institutions. The conceptual framework consists of critical processes and performance measures identified in the context of the 2003 Malcolm Baldrige National Quality Award (MBNQA) Health Care Criteria for Performance Excellence.</dc.description.abstract>
```

Mapping to ETD-MS

```
<dc.description.abstract>This dissertation study presents a conceptual framework for implementing and assessing patient safety systems in healthcare institutions. The conceptual framework consists of critical processes and performance measures identified in the context of the 2003 Malcolm Baldrige National Quality Award (MBNQA) Health Care Criteria for Performance Excellence.</dc.description.abstract>
```

Mapping to OAI-ETD

```
<dc.description.abstract>This dissertation study presents a conceptual framework for implementing and assessing patient safety systems in healthcare institutions. The conceptual framework consists of critical processes and performance measures identified in the context of the 2003 Malcolm Baldrige National Quality Award (MBNQA) Health Care Criteria for Performance Excellence.</dc.description.abstract>
```

Mapping to Highwire Press Tags (Google Scholar)

```
<meta name="citation_abstract_html_url" content="http://hdl.handle.net/1969.1/1042"/>
```

Description (O)

Definition: Additional information regarding the thesis or dissertation.

Mandatory practice: N/A

Optional practice: May use unqualified or qualified elements (e.g. <dc.description.sponsorship>). Notes information is repeatable.

Recommended Data Value: N/A

Note: The current ETD-MS standard, upon which OAI-ETD is based, specifies that abstract is equated with description and that unqualified dc.description can therefore only be filled with abstract values. Other descriptions are relegated to dc.description.note. We recommend using dc.description.abstract for the abstract, and modifying OAI-ETD to recognize dc.description.abstract (see Appendix C). We caution users that using unqualified dc.description for values other than the abstract violates ETD-MS and may confuse mappings. The 2008 TDL Guidelines did not address description.

Example of Use:

```
<dc.description>page 29 is missing from hardcopy</dc.description> <dc.description.sponsorship>This thesis was funded in part by the U.S. National Science Foundation (NSF)</dc.description.sponsorship>
```

Mapping to ETD-MS

```
<dc.description.note>This thesis was funded in part by the U.S. National Science Foundation
```

(NSF)</dc.description.note>

Note! Per the ETD-MS standard, dc.description is interchangeable with the qualified element dc.description.abstract. Therefore unqualified dc.description is ignored in the direct ETD-MS mapping but is included in the OAI-ETD output.

Mapping to OAI-ETD

<dc.description.note>page 29 is missing from hardcopy</dc.description.note>

<dc.description.note>This thesis was funded in part by the U.S. National Science Foundation (NSF)</dc.description.note>

Mapping to Highwire Press Tags (Google Scholar)

N/A

Subject (O)

Definition: A term or phrase representing the primary topic(s) on which a work is focused.

Mandatory practice: N/A

Optional practice: Subject element is repeatable. Controlled subject headings may be encoded by using a qualifier to indicate the controlled vocabulary (e.g. lcsh, tgm, mesh).

Recommended Data Value: The form of the heading may be taken from a standard or local thesaurus, such as the Library of Congress Subject Headings (LCSH). Keywords or subjects supplied by the author may be encoded in unqualified element. Multiple subjects should be encoded by repeating the subject element for each subject.

Example of Use:

<dc.subject>healthcare</dc.subject>

<dc.subject>patient safety</dc.subject>

<dc.subject.lcsh>Medical care—Quality control—United States—20th century</dc.subject.lcsh>

Mapping to ETD-MS

<dc.subject>healthcare</dc.subject>

<dc.subject>patient safety</dc.subject>

<dc.subject.lcsh>Medical care—Quality control—United States—20th century</dc.subject.lcsh>

Mapping to OAI-ETD

<dc.subject>healthcare</dc.subject>

<dc.subject>patient safety</dc.subject>

<dc.subject>Medical care—Quality control—United States—20th century</dc.subject>

Mapping to Highwire Press Tags (Google Scholar)

<meta name="citation_keywords" content="healthcare; patient safety; Medical care—Quality control—United States—20th century">

Identifier (M)

Definition: A locally-defined, unique standard number, code, or handle that distinctively identifies a resource

Mandatory practice: Encode the uniform resource locator (URL) of the resource in <dc.identifier.uri> qualified element. Only one handle identifier should be encoded per resource.

Note: A handle location (persistent URL) is automatically generated in DSpace upon ingestion.

Optional practice: Other valid identifiers may be used. Other valid identifiers should be encoded in qualified elements and such elements are repeatable.

Recommended Data Value: Use system-supplied handle URI or persistent URL for <dc.identifier.uri> element . For any unique identifier meaningful to the institution, e.g. OCLC number, Proquest number, bib number, handle, etc., it is recommended to define the type with an qualifier that describes the unique identifier. For example, "oclc", "proqst", "doi" or "bib."

Example of Use:

```
<dc.identifier.uri>http://hdl.handle.net/1969.1/1042</dc.identifier.uri>
<dc.identifier.proqst>742126641</dc.identifier.proqst>
```

Mapping to ETD-MS

```
<dc.identifier>http://hdl.handle.net/1969.1/1042</dc.identifier>
```

Mapping to OAI-ETD

```
<dc.identifier>http://hdl.handle.net/1969.1/1042</dc.identifier>
```

Note: Only map the URI to the location where the work can be viewed or downloaded.

Mapping to Highwire Press Tags (Google Scholar)

```
<meta name="citation_pdf_url" content="http://oaktrust.library.tamu.edu/bitstream/handle/1969.1/1042/etd-tamu-2004B-EDAD-Akins-3.pdf">
```

Note: Google Scholar guidelines specify to map the location of the full text or absolute URL of the PDF file.

Author Identifier - ORCID (O)

Definition: Unique standard number or code to aid in author disambiguation.

Mandatory practice: N/A

Optional practice: Non-repeatable. Use qualifier.

Recommended Data Value: Use a valid and authenticated 16-digit number from the Open Researcher

and Contributor ID registry, called an ORCID¹⁰ identifier.

Note: ORCID implementation is emergent, and best practices are not yet established. Please see Appendix E for more information on ORCID and name disambiguation.

Example of Use:

<dc.identifier.orcid>0000-0000-0000-0000</dc.identifier.orcid>

Mapping to ETD-MS

N/A

Mapping to OAI-ETD

<dc.identifier.orcid>0000-0000-0000-0000</dc.identifier.orcid>

Note: Though author IDs are not part of the ETD-MS standard, it is recommended to export this data if provided.

Mapping to Highwire Press Tags (Google Scholar)

N/A

Rights (O)

Definition: Information about rights held in and over the resource. Typically, this describes the conditions under which the work may be distributed, reproduced, etc.

Mandatory practice: N/A

Optional practice: Repeatable element. Use qualifiers to specify different types of rights.

Recommended Data Value: Open text field for simple rights statement. This may be boilerplate as determined by institutional policy. Additional values may be assigned such as link to a Creative Commons license or explicit rights holder statement. An access level may be assigned, and we recommend using a consistent vocabulary, for example: "Publicly accessible"; "Access restricted to campus"; or "Access restricted due to embargo. Release date XXXX-XX-XX."

Example of Use:

<dc.rights>The authors of the theses and dissertations are the copyright owners. The Digital Library and Archives has their permission to store and provide access to these works.</dc.rights>

<dc.rights>Copyright © is held by the author. Digital access to this material is made possible by the University Libraries. Further transmission, reproduction or presentation (such as public display or performance) of protected items is prohibited except with permission of the author.</dc.rights>

<dc.rights>Attribution 3.0 United States</dc.rights>

<dc.rights.uri>http://creativecommons.org/licenses/by/3.0/us/</dc.rights.uri>

<dc.rights.rightsHolder>Copyright 2013 © John Smith</dc.rights.rightsHolder>

¹⁰ <http://orcid.org/>

<dc.rights.accessRights>Access restricted to campus</dc.rights.accessRights>

Mapping to ETD-MS

<dc.rights>The authors of the theses and dissertations are the copyright owners. The Digital Library and Archives has their permission to store and provide access to these works.</dc.rights>

<dc.rights>Copyright © is held by the author. Digital access to this material is made possible by the University Libraries. Further transmission, reproduction or presentation (such as public display or performance) of protected items is prohibited except with permission of the author</dc.rights>

<dc.rights>Attribution 3.0 United States</dc.rights>

<dc.rights><http://creativecommons.org/licenses/by/3.0/us/></dc.rights>

<dc.rights>Copyright 2013 © John Smith</dc.rights>

<dc.rights>Access restricted to campus</dc.rights >

Mapping to OAI-ETD

<dc.rights>The authors of the theses and dissertations are the copyright owners. The Digital Library and Archives has their permission to store and provide access to these works.</dc.rights>

<dc.rights>Copyright © is held by the author. Digital access to this material is made possible by the University Libraries. Further transmission, reproduction or presentation (such as public display or performance) of protected items is prohibited except with permission of the author.</dc.rights>

<dc.rights>Attribution 3.0 United States</dc.rights>

<dc.rights><http://creativecommons.org/licenses/by/3.0/us/></dc.rights>

<dc.rights>Copyright 2013 © John Smith</dc.rights>

<dc.rights>Access restricted to campus</dc.rights>

Mapping to Highwire Press Tags (Google Scholar)

N/A

Degree Information (M)

Definition: Information about the conferred degree, including name, level, discipline of degree, and the name of institution and academic department granting the degree associated with the work.

Mandatory practice: Encode information with the appropriate qualified element using the <thesis.degree._> schema. Degree name, level, grantor elements are mandatory.

Optional practice: Other “degree Information” (such as department, college, major) only populate if available.

Recommended Data Value: For degree name: Use the fully spelled-out form of the degree name, e.g. Doctor of Philosophy. For degree level: Use the level of education associated with the document, e.g. Masters, Doctoral, Postdoctoral, etc. For degree granting institution: use the form of the name authorized by the Library of Congress Name Authority File. For degree discipline: Use the area of study of the intellectual content of the document or program name. For degree department: Use the academic department granting the degree. It is recommended for institutions to maintain a list of disciplines and departments relevant to their institution to provide consistency in values. Degree major is the main area of focus as supplied by student.

Note: Mandatory if available. If any values are not available, do not use the phrase "Not available." Leave the element empty.

Note: Degree name, level, grantor, and discipline are prescribed in the ETD-MS standard. Other elements that are included as "degree information" (department, college, major) are unique to the Vireo Electronic Thesis and Dissertation submission system. Vireo enables drop-down lists for department and college fields. Degree major is a free text field in Vireo populated by the author, where multiple values are allowed.

Example of Use:

```
<thesis.degree.name>Doctor of Philosophy</thesis.degree.name>
<thesis.degree.level>Doctoral</thesis.degree.level>
<thesis.degree.grantor>Texas A&M University</thesis.degree.grantor>
<thesis.degree.discipline>Engineering</thesis.degree.discipline>
<thesis.degree.department>Materials Science and Engineering</thesis.degree.department>
<thesis.degree.college>College of Engineering</thesis.degree.college>
<thesis.degree.major>Molecular dynamics</thesis.degree.major>
```

Mapping to ETD-MS

```
<thesis.degree.name>Doctor of Philosophy</thesis.degree.name>
<thesis.degree.level>Doctoral</thesis.degree.level>
<thesis.degree.grantor>Texas A&M University</thesis.degree.grantor>
<thesis.degree.discipline>Engineering</thesis.degree.discipline>
```

Mapping to OAI-ETD

```
<thesis.degree.name>Doctor of Philosophy</thesis.degree.name>
<thesis.degree.level>Doctoral</thesis.degree.level>
<thesis.degree.grantor>Texas A&M University</thesis.degree.grantor>
<thesis.degree.discipline>Engineering</thesis.degree.discipline>
```

Mapping to Highwire Press Tags (Google Scholar)

```
<meta name="citation_dissertation_institution" content="Texas A&M University">
```

Note: thesis.degree.grantor may be considered equivalent to dc.publisher element for recording the entity responsible for making a resource available. Using thesis.degree.grantor allows for a specific mapping of the entity's name in Google Scholar tags that will not conflict with non-ETD collections that may also appear in a repository's holdings.

Embargo (O, S)

Definition: An embargo is a temporary access restriction placed on metadata or bitstreams.

Mandatory practice: These fields are used to automate the restriction of embargoed ETDs and removal of such restrictions. They are both automatically populated with the date to remove an embargo as set in the Vireo system. Once an embargo date has past, the value is removed from <dc.embargo.lift> element by the repository system. The exact embargo terms are defined in the Vireo system; options may include: None, Six month, One year, Two year embargo periods.

Note: For embargoed items, the metadata may be publicly viewable while the bitstreams (PDF and any supplementary files) are not. This depends on local configuration.

Optional practice: n/a

Recommended data values: n/a

Example of Use:

```
<dc.embargo.lift>2015-12-01</dc.embargo.lift >
<dc.embargo.terms>2015-12-01</dc.embargo.terms>
```

Note: Some institutions elect to add text to an item page (such as a Header or footer) to better communicate to potential reader that the PDF is restricted. However this information is not part of the formal metadata and therefore not sharable to aggregators. It is recommended to use dc.rights.accessLevel to record any embargoed information of the thesis.

Mapping to ETD-MS

N/A

Mapping to OAI-ETD

N/A

Mapping to Highwire Press Tags (Google Scholar)

N/A

Provenance (S)

Definition: The history of the custody of the item since its creation, including any changes that may be significant in terms of authority, integrity, and interpretation.

Mandatory practice: System-generated data should not be edited manually.

Optional practice: N/A

Recommended Data Value: For Vireo deposits, tracking of key steps in the ETD submission process are documented in the <dc.description.provenance> element. This element is usually not viewable by the general public but only viewable to users with collection admin permissions.

Example of Use:

```
<dc.description.provenance>Submission original under the a zero month embargo labeled
'None'.</dc.description.provenance>
<dc.description.provenance>The student, John Smith, accepted the attached license on 2015-01-22 at
11:17.</dc.description.provenance>
<dc.description.provenance>The student, John Smith, submitted this document for approval on 2015-01-22 at
13:19.</dc.description.provenance>
<dc.description.provenance>This document was approved for publication on 2015-04-14 at
15:12.</dc.description.provenance>
<dc.description.provenance>DSpace METS Submission Ingestion Package generated from Vireo submission #1234
```

on 2015-06-25 at 14:52:51</dc.description.provenance>
<dc.description.provenance> Submitted by Vireo Thesis (nobody+thesis@school.edu) on 2015-06-25T19:53:17Z
No. of bitstreams: 2 SMITH-THESIS-2015.pdf: 54866104 bytes, checksum: 1978c60b86426c67df8f82a8fdae554a
(MD5) LICENSE.txt: 2672 bytes, checksum: 03a8517b633fb40e0b3f41bb7693ef03
(MD5)</dc.description.provenance>
<dc.description.provenance>Made available in DSpace on 2015-06-25T19:53:17Z (GMT). No. of bitstreams: 2 ...
</dc.description.provenance>

Mapping to ETD-MS

N/A

Mapping to OAI-ETD

N/A

Mapping to Highwire Press Tags (Google Scholar)

N/A

Appendix A: Quick Reference Mapping Table

M=Mandatory O=Optional S=System generated

TDL ETD DC Element	ETD-MS Element	OAI-ETD [†]	HighWire Press Tags
Title Information (M) <dc.title>	<dc.title>	<dc.title>	citation_title
Name of Author (M) <dc.creator>	<dc.creator>	<dc.creator>	citation_author
Name of Thesis Advisor (M) <dc.contributor.advisor>	<dc.contributor>	<dc.contributor>	N/A
Name of Committee Member (O) <dc.contributor.committeeMember>	N/A	N/A	N/A
Type of Resource (M) <dc.type.dcmi>	<dc.type>	<dc.type>	N/A
Genre (M) <dc.type.genre>	<dc.type>	<dc.type>	N/A
Date (M) <dc.date.created> <dc.date.issued> <dc.date.accessioned> <dc.date.available>	N/A <dc.date> N/A N/A	N/A <dc.date> N/A N/A	N/A citation_publication_date N/A citation_online_date
Language (O) <dc.language.iso>	<dc.language>	<dc.language>	N/A
Format (O) <dc.format.extent> <dc.format.mimetype> <dc.format.digitalOrigin>	<dc.format> <dc.format> <dc.format>	<dc.format> <dc.format> <dc.format>	N/A N/A N/A
Abstract (M) <dc.description.abstract>	<dc.description.abstract>	<dc.description.abstract>	citation_abstract_html_url
Description (O) <dc.description> <dc.description.qualifier>	N/A <dc.description.note>	<dc.description.note> (a) <dc.description.note>	N/A N/A
Subject (O) <dc.subject> <dc.subject.lcsh>	<dc.subject> <dc.subject>	<dc.subject> <dc.subject>	citation_keywords citation_keywords
Identifier (M) <dc.identifier.uri> <dc.identifier.qualified>	<dc.identifier> N/A	<dc.identifier> N/A	citation_pdf_url N/A
Author Identifiers ORCID (O) <dc.identifier.orcid>	N/A	<dc.identifier.orcid> (b)	N/A

TDL ETD DC Element	ETD-MS Element	OAI-ETD [†]	HighWire Press Tags
Rights (O) <dc.rights> <dc.rights.uri> <dc.rights.rightsHolder> <dc.rights.accessRights>	<dc.rights> <dc.rights> <dc.rights> <dc.rights>	<dc.rights> <dc.rights> <dc.rights> <dc.rights>	N/A N/A N/A N/A
Degree Information (M) <thesis.degree.name> <thesis.degree.level> <thesis.degree.discipline> <thesis.degree.grantor> <thesis.degree.department> <thesis.degree.program> <thesis.degree.college>	<thesis.degree.name> <thesis.degree.level> <thesis.degree.discipline> <thesis.degree.grantor> N/A N/A N/A	<thesis.degree.name> <thesis.degree.level> <thesis.degree.discipline> <thesis.degree.grantor> N/A N/A N/A	citation_dissertation_name N/A N/A citation_dissertation_institution N/A N/A N/A
Embargo (O, S) <dc.embargo.lift> <dc.embargo.terms>	N/A N/A	N/A N/A	N/A N/A
Provenance (S) <dc.description.provenance>	N/A	N/A	N/A

Notes:

[†] OAI-ETD format

a) Given the common usage of unqualified description element to record non-abstract information, we elected to map the <dc.description> element to <dc.description.note> for OAI-ETD output. As noted in these guidelines, the preferred method for recording abstract information is in the <dc.description.abstract> element.

b) Given the movement in the scholarly community by publishers and academic institutions to enable ORCIDs, we recommend including this identifier as part of the OAI-ETD output. See Appendix E for discussion of ORCID for ETDs.

Appendix B: Full example of ETD encoded in TDL ETD DC

```

<?xml version="1.0" encoding="UTF-8"?>
<dublin_core schema="dc">
  <dcvalue element="title">Critical processes and performance measures for patient safety systems in healthcare institutions:
  a Delphi Study</dcvalue>
  <dcvalue element="creator">Akins, Ralitsa B., 1967-</dcvalue>
  <dcvalue element="contributor" qualifier="advisor">Cole, Bryan R.</dcvalue>
  <dcvalue element="contributor" qualifier="committeeMember">Davis, Eddie J.</dcvalue>
  <dcvalue element="contributor" qualifier="committeeMember">Lincoln, Yvonna S.</dcvalue>
  <dcvalue element="contributor" qualifier="committeeMember">Smith, Elvin E.</dcvalue>
  <dcvalue element="type" qualifier="dmci">Text</dcvalue>
  <dcvalue element="type" qualifier="genre">Thesis</dcvalue>
  <dcvalue element="date" qualifier="created">December 2004</dcvalue>
  <dcvalue element="date" qualifier="issued">2004-12</dcvalue>
  <dcvalue element="date" qualifier="accessioned">2005-03-15T19:44:57Z</dcvalue>
  <dcvalue element="date" qualifier="available">2005-03-15T19:44:57Z</dcvalue>
  <dcvalue element="language" qualifier="iso">eng</dcvalue>
  <dcvalue element="format" qualifier="medium">electronic</dcvalue>
  <dcvalue element="format" qualifier="mimetype">application/pdf</dcvalue>
  <dcvalue element="format" qualifier="digitalOrigin">born digital</dcvalue>
  <dcvalue element="description" qualifier="provenance">Ralitsa B. Atkins graduated in December 2004.</dcvalue>
  <dcvalue element="description" qualifier="abstract">This dissertation study presents a conceptual framework for
  implementing and assessing patient safety systems in healthcare institutions. The conceptual framework consists of critical
  processes and performance measures identified in the context of the 2003 Malcolm Baldrige National Quality Award
  (MBNQA) Health Care Criteria for Performance Excellence. Methodology: The Delphi technique for gaining consensus from a
  group of experts and forecasting significant issues in the field of the Delphi panel expertise was used. Data collection included
  a series of questionnaires where the first round questionnaire was based on literature review and the MBNQA criteria for
  excellence in healthcare, and tested by an instrument review panel of experts. Twenty-three experts (MBNQA healthcare
  reviewers and senior healthcare administrators from quality award winning institutions) representing 18 states participated in
  the survey rounds. The study answered three research questions: (1) What are the critical processes that should be included in
  healthcare patient safety systems? (2) What are the performance measures that can serve as indicators of quality for the
  processes critical for ensuring patient safety? (3) What processes will be critical for patient safety in the future? The identified
  patient safety framework was further transformed into a patient safety tool with three levels: basic, intermediate, and
  advanced. Additionally, the panel of experts identified the major barriers to the implementation of patient safety systems in
  healthcare institutions. The identified "top seven" barriers were directly related to critical processes and performance
  measures identified as "important" or "very important" for patient safety systems in the present and in the future. This
  dissertation study is significant because the results are expected to assist healthcare institutions seeking to develop high
  quality patient safety
  programs, processes and services. The identified critical processes and performance measures can serve as a means of
  evaluating existing patient safety initiatives and guiding the strategic planning of new safety processes. The framework for
  patient safety systems utilizes a systems approach and will support healthcare senior administrators in achieving and sustaining
  improvement results. The identified patient safety framework will also assist healthcare institutions in using the MBNQA Health
  Care Criteria for Performance Excellence for self-assessment and quality improvement.</dcvalue>
  <dcvalue element="subject">healthcare</dcvalue>
  <dcvalue element="subject">patient safety</dcvalue>
  <dcvalue element="subject" qualifier="lcsch">
    Medical care—Quality control—United States—20th century</dcvalue>
  <dcvalue element="identifier" qualifier="proqst">742126641</dcvalue>
  <dcvalue element="identifier" qualifier="uri">http://handle.tamu.edu/1969.1/1042</dcvalue>
  <dcvalue element="rights">The authors of the theses and dissertations are the copyright owners. The Digital Library and
  Archives has their permission to store and provide access to these works.</dcvalue>
  <dcvalue element="rights">Copyright © is held by the author. Digital access to this material is made possible by the University
  Libraries. Further transmission, reproduction or presentation (such as public display or performance) of protected items is
  prohibited except with permission of the author.</dcvalue>
  <dcvalue element="rights">Attribution 3.0 United States</dcvalue>
  <dcvalue element="rights" qualifier="uri"> http://creativecommons.org/licenses/by/3.0/us/</dcvalue>

```

```
<dcvalue element="rights" qualifier="rightsHolder">Copyright 2014 © Ralitsa B Akins</dcvalue>
<dcvalue element="rights" qualifier="accessRights">Access restricted to campus</dcvalue>
  <dublin_core schema="thesis">
    <dcvalue element="degree" qualifier="grantor">Texas A & M University</dcvalue>
    <dcvalue element="degree" qualifier="department">Educational Administration and Human Resource
Development</dcvalue>
    <dcvalue element="degree" qualifier="name">Doctor of Philosophy</dcvalue>
    <dcvalue element="degree" qualifier="level">Doctoral</dcvalue>
    <dcvalue element="degree" qualifier="discipline">Educational Administration</dcvalue>
    <dcvalue element="degree" qualifier="department">Educational Psychology</dcvalue>
    <dcvalue element="degree" qualifier="college">College of Education and Human Development</dcvalue>
  </dublin_core>
</xml>
```

Appendix C: Crosswalks

The fundamental concept of [Open Archives Initiative Protocol for Metadata Harvesting](#) (OAI-PMH) is to provide a simple approach for sharing metadata globally. This promise of universal integration is challenged by lack of consistency in metadata practices. Communities of practice have sought to alleviate these challenges by customizing OAI output to more closely match community-developed metadata standards. The [Networked Digital Library of Theses and Dissertations](#) (NDLTD) community has published and maintains a metadata standard ([ETD-MS v1.1: an Interoperability Metadata Standard for Electronic Theses and Dissertations](#)) which establishes a common set of elements for use by the international community, with a goal of facilitating exchange. A [customized OAI format for ETDs](#), designed to comply with ETD-MS, was released with DSpace version 3.0 in 2011. However, this OAI ETD format is significantly out of date with the latest version of NDLTD's standard (ETD-MS v1.1). The list below highlights recommended changes to update the OAI ETD format to better align with the latest standard version and addresses changing practices in qualified Dublin core by the community at large.

Authors. The NDLTD standard (ETD-MS v1.1) recommends using <dc.creator> to record author names. The existing crosswalk only exposes the deprecated <dc.contributor.author> element.

Degree Info. The NDLTD standard (ETD-MS v1.1) recommends using specialized schema for recording degree information in syntax: <thesis.degree.qualifier>. The existing crosswalk uses syntax <thesis.qualifier>. The missing "degree" element prevents the display of this data in the output. The existing crosswalk is also missing mapping for grantor (<thesis.degree.grantor>) element.

Dates. There are many dates associated with a thesis or dissertation, including system-generated dates and dates in non-ISO format. The existing ETD crosswalk outputs all dates, though ETD-MS recommends using only one date as <dc.date>. The updated crosswalk will limit mapping of dates to <dc.date.issued> (graduation date)¹¹, to ensure compliance with both ETD-MS and best practices maintained by both DSpace and ETD metadata communities.

Descriptions. The ETD-MS standard reserves both unqualified description <dc.description> and qualified description <dc.description.abstract> for abstract information. The updated crosswalk will map specific qualified description elements to <dc.description.note> to better align with the current standard (ETD-MS v1.1) recommendation. Given the common practice of using unqualified description element for sundry information, this will also be mapped to <dc.description.note>.

Formats and Types. A growing practice among repositories is to use qualified type and format elements. The current crosswalk only maps simple <dc.type> and <dc.format>. The updated crosswalk will include output for qualified elements.

Rights. To better align with the current NDLTD standard (ETD-MS v1.1), access levels are displayed in a separate tag <dc.rights.accessLevel> to distinguish this information from other rights data. All other rights data should be mapped to <dc.rights>.

Example code: https://github.com/TexasDigitalLibrary/DSpace/blob/dspace-5_x-TDL-ETD/dspace/config/crosswalks/oai/metadataFormats/etdms1-1.xsl

¹¹ Please see discussion of date mapping updates in accompanying *Report for Texas Digital Library Descriptive Metadata for Electronic Theses and Dissertations*, v. 2.

Appendix D: Suggested metadata for supplemental files and complex objects

Metadata about supplemental files or complex digital objects may be recorded to aid the discoverability and preservation of non-textual documents. Most basic of these are type elements using standard vocabulary such as <dc.format.mimetype>¹² and <dc.type.dcmi>¹³. A benefit of assigning type elements is that many digital archives now are browsable by type terms making navigating to items with these sorts of files more visible. Assigning type information may also help administrators in identifying non-standard or unsupported file formats, a possible first step in either migrating or converting files to ensure long-term access.

Another basic practice is to identify which files are supplemental. One method is to use the bitstream description field, found in DSpace repository software. Recording the words “Supporting Document” next to the file itself, will clearly differentiate the primary thesis or dissertation document from the supplemental file(s) found in the item. Though this DSpace bitstream description field is not harvestable metadata in the simple Dublin core/OAI PMH sense, a file caption may provide meaning for a potential reader.

For complex objects, especially those including non-proprietary file types, it is recommended to include a README.txt file documenting key aspects of these documents and providing a manifest of any zipped or compressed files. Minimum metadata to include in a README.txt file is information on data formats, file types, file structures, rights status, and any special software needed to render the files.

For related datasets created by the author and deposited in an external repository, a link to this data may be provided in the <dc.description> element. Though this citation may already be included in the bibliography of the thesis or dissertation, providing the direct link as part of the metadata record can raise awareness and visibility of unique datasets. Any supplemental files or datasets should include a statement of rights status.

¹² <http://www.iana.org/assignments/media-types/media-types.xhtml>

¹³ <http://dublincore.org/documents/2000/07/11/dcmi-type-vocabulary/>

Appendix E: ORCIDS and Name Disambiguation

The application of ORCIDs in ETD environments remains an emergent area. At this time, best practices have not been established that stipulate the use and exposure of ORCIDs in ETD records.

We encourage graduate school professionals, institutional repository administrators, and librarians to consider ORCID implementation at their institutions. The library is well-situated to serve as a leader in this effort, and full and effective ORCID implementation will require a coordinated effort between administrators, staff, faculty, and students across the institution. Academic institutions, granting agencies, publishers, and national and international consortia are implementing ORCID more widely.¹⁴ We anticipate that ORCID adoption and integration into local systems for tracking scholarly output will continue to increase, but we are not at a point where we can dictate practice.

Our name disambiguation case study is limited to the Vireo environment. We discovered highly individualized practices for name disambiguation at different institutions. Some of these practices go beyond the scope of Vireo but directly or indirectly impact decisions around name disambiguation in ETD records:

- **Cataloging and MARC**
At some institutions, the practice of creating MARC records for theses and dissertations was discontinued after the introduction of ETDs. Other institutions are continuing to create MARC records for ETDs (resulting in one repository record and one MARC record per ETD). And, other institutions have MARC records for some ETDs but not all, or have discontinued the creation of MARC records with the adoption of an integrated discovery layer. Institutions that continue to create MARC records may choose to rely on RDA-formatted and NACO controlled name headings in MARC as the authoritative form of a name and allow repository records to remain uncontrolled.¹⁵ For these institutions, adoption of ORCID for the purposes of name disambiguation may be a low priority.¹⁶
- **Birth Dates and Privacy**
In both MARC and non-MARC records, birth years have been employed as part of the author's name as a means of name disambiguation. The use of birth years for this purpose has long been a practice in library cataloging. Some ETD authors, however, have asked that their birth years be removed from the publicly viewable record as a matter of privacy. In our study of name

¹⁴ Jisc in the UK, ANVUR, and CRUI in Italy have adopted ORCID. There is currently work to develop an Australian ORCID Consortium Model as well. See the ORCID blog for updates from these institutions: <http://orcid.org/category/newsletter/blog>

¹⁵ The transition from AACR2 to RDA resulted in changes in LC/NACO authority files. You can find a detailed explanation of these changes in the summary report from Library of Congress and the Program for Cooperative Cataloging here: http://www.loc.gov/aba/rda/pdf/lcnaf_rdaphase.pdf

¹⁶ The transition out of MARC to BIBFRAME and a linked data environment may further complicate this issue.

disambiguation, these requests have always been honored. An ORCID would disambiguate the author without the need to provide a birth year, middle name, or middle initial. From the perspective of author privacy, ORCID is a better solution than “traditional” methods of name disambiguation.

- Legacy data

Adoption of ORCID in ETD records would obviate the need for “traditional” name disambiguation practices in a repository record, but it would not necessarily be helpful or possible with legacy data. For institutions that are digitizing print theses and dissertations, there is no way to guarantee that the author has or will register for an ORCID. Librarians and repository administrators may need to employ alternative means of name disambiguation for these records. Other scenarios involving legacy data - such as repository clean-up, or cross walking of MARC records into the repository - will present their own complications.

Using ORCID for new and incoming ETDs may prove to be the most practical solution to the issue of name disambiguation, but other acceptable alternatives must be available for dealing with legacy data, for institutions unable to implement ORCID, and for other, unanticipated scenarios.

We strongly recommend that name disambiguation in ETDs continue to be an area of development and that Vireo continue to capture ORCID and other information used for name disambiguation purposes.