

Electronic Submission to UMI using FTP

As the digital era progresses, UMI Dissertation Publishing has been encouraging our university publishing partners to move from paper submissions to electronic submissions. UMI currently supports two options for electronic submission:

1. The [UMI ETD Administrator](#), an online system we developed; and
2. Electronic submissions via FTP.

The intent of this document is to outline the parameters around FTP submission of graduate works.

Overview of Criteria and Parameters

In order for UMI Dissertation Publishing to accept graduate work submissions via FTP, there are certain criteria and parameters our publishing partners must adhere to. By adhering to these standards, our ability to publish individual graduate works more quickly, and address problem submissions in a timely manner, is significantly increased. Below is a brief overview of the criteria:

1. **Adhere to UMI Dissertation Publishing XML DTD:** At UMI, we have developed an XML Document Type Definition (DTD) that defines how to provide metadata about the graduate work and the author. Each FTP submission must include an .xml file that includes all the required metadata fields - and preferably, all the optional metadata fields – mapped to our DTD. The file should be labeled as follows: "*authorlastname_authorfirstname_DATA.xml*", where "authorlastname" and "authorfirstname" are that of the author.
2. **Provide full text of graduate work in PDF format:** The PDF of the graduate work should be labeled as follows: "*authorlastname_authorfirstname.pdf*", where "authorlastname" and "authorfirstname" are that of the author. Each PDF should have fonts embedded, and the security settings for the document should not preclude printing, inserting additional pages (in this case, the UMI front matter), or extracting text (to create a preview that is made freely available with the bibliographic record). In addition, multimedia should not be embedded in the PDF. This significantly increases the file size, making it more difficult to access by researchers, and will make future digital preservation efforts much more difficult. Instead, multimedia files should be included separately, in a media folder (see Item 3 below).
3. **Include supplementary files in their own file folder:** Any supplementary files, like spreadsheets and video clips, that support the dissertation itself should be included, too. They should be included in a folder labeled "*authorlastname_authorfirstname_media*," and the file structure should adhere to the structure in which the author submitted them.
4. **1 zip file per submission:** Each individual submission must be uploaded to our FTP server as a single zip file. The zip file must contain the XML file of metadata, the PDF of the full text, and any supplementary files in a file folder, as necessary. The zip file itself should follow this convention: "*upload_authorlastname_authorfirstname.zip*", and contain files that adhere to the naming convention described in Items 1, 2, and 3 above. A typical submission might look like this:
 - o Upload_barbarino_vinny.zip
 - barbarino_vinny_DATA.xml
 - barbarino_vinny.pdf
5. **Notify us** when you have uploaded submissions by emailing us at dissepubl@proquest.com.
6. **Include a "packing list":** In the email you send to dissepubl@proquest.com to alert us that you have uploaded new submissions, please include a document that lists each author

whose work is included in the latest batch of submissions. We will check what we find on the FTP site against this list, to make sure we have received the correct set of submissions.

7. **Author Publishing Agreement:** Each author will need to sign a paper copy of the ProQuest/UMI Publishing Agreement, which must be mailed to UMI Dissertation Publishing at:

Dissertations Acquisitions
ProQuest
789 E Eisenhower Pkwy
PO Box 1346
Ann Arbor, MI 48106-1346

Related Documents

There are several support documents and files UMI will provide to you, should you decide to use our FTP submission process:

1. The UMI Dissertation Publishing DTD.
2. A countries.xls file that includes our accepted country codes (use for populating the <DISS_country> and <DISS_citizenship> elements).
3. A degree_codes.xls file that includes all accepted degree codes (use for populating the <DISS_degree> element).
4. A language_codes.xls file that includes all accepted language codes (use for populating the <DISS_language> element).
5. A state_codes.xls file that includes all accepted U.S. state codes (use for populating the <DISS_st> element).
6. A umi_subjects.xls file that includes all accepted UMI subject categories (use for populating the <DISS_cat_code> and <DISS_cat_desc> elements).
7. The ProQuest/UMI Publishing Agreement for FTP submissions.

DTD Directions

The following directions describe how specific data elements and attributes should be formatted in the XML metadata file. Most are also covered under comments in the DTD itself. Those elements that are generally self explanatory are not described below.

1. **<DISS_submission> element:** <DISS_submission> is the root element. One XML file structured as described in the DTD equals one graduate work.
 - a. **publishing_option attribute:** An attribute of the <DISS_submission> element, publishing_option can be 0 (Traditional Publishing) or 1 (Open Access Publishing). Open Access publishing requires an extra \$95 fee, in addition to the Traditional Publishing fee, and ProQuest will make the graduate work freely available in our [PODT Open](#) repository.
 - b. **embargo_code attribute:** An attribute of the <DISS_submission> element, the embargo_code attribute can be:
 - i. 0 - corresponds to no embargo
 - ii. 1 - 6 months embargo

- iii. 2 - 1 year embargo
 - iv. 3 - 2 year embargo
- c. **third_party_search attribute:** An attribute of the <DISS_submission> element, the third_party_search attribute can have a Y or N value. Y indicates the author would like to have his or her metadata indexed by 3rd party search engines like Google Scholar. N indicates the author does not want his or her graduate work metadata indexed by outside search engines.
2. **<DISS_authorship> element:** Used to collect metadata describing at least one primary, and optionally, additional authors.
 3. **<DISS_author> element:** Describes one person claiming authorship of the dissertation.
 - a. **type attribute:** An attribute of the <DISS_author> element, describes the type of author – most often will be “primary”, and possibly “secondary”, if there multiple authors. “claimant” attribute identifies a designated owner if the author of the graduate work has passed away. note that the author can specify multiple addresses, and the contact element has a “current” or “future” attribute, letting us know where to contact the author subsequently
 4. **<DISS_citizenship> element:** A child element of <DISS_author>, this element is used to capture the primary’s author country of citizenship. It is optional.
 5. **<DISS_affiliation> element:** A child element of <DISS_name>, this element is used to capture an individual’s school affiliation.
 6. **<DISS_advisor> element:** Parent element used to capture advisor names. One advisor is required per submission.
 7. **<DISS_cmte_member> element:** Parent element used to capture committee member’s names. Not required; can provide up to 8 committee members per submission.
 8. **<DISS_name> element:** This element, and its child elements – <DISS_surname>, <DISS_fname>, <DISS_middle>, and <DISS_suffix> – are used to capture an individual’s name (author, advisor, and committee members). It can be child element of <DISS_author>, <DISS_advisor>, and <DISS_cmte_member>
 9. **<DISS_contact> element:** Child element of <DISS_author>. Along with its child elements, <DISS_contact> is used to provide address, phone, and email for authors.
 - a. **type attribute:** used to indicate whether address is current or future. Value should be “current” or “future”.
 10. **<DISS_contact_effdt> element:** Child element of <DISS_contact>, the value should be the effective date for a given address. It should be formatted as “mm/dd/yyyy”.
 11. **<DISS_phone_fax> element:** A child element of <DISS_contact>. Along with its child elements - <DISS_cntry_cd>, <DISS_area_code>, <DISS_phone_num>, and <DISS_phone_ext> - used to capture the author’s phone number.
 - a. **type attribute:** used to indicate a phone or fax number. P is Phone number, F is Fax number (note: UMI is no longer interested in capturing fax numbers; just phone numbers).

12. **<DISS_email> element:** Child element of <DISS_contact>, this element is intended to capture a “permanent” email address for the author – i.e., a Yahoo! email address, not the university email address used while a student.
13. **<DISS_address> element:** A child element of <DISS_contact>. Along with its child elements - <DISS_addrline>, <DISS_city>, <DISS_st>, <DISS_pcode>, <DISS_country> - this element is used to provide a mailing address for the author.
14. **<DISS_description> element:** Along with its child elements - <DISS_title>, <DISS_dates>, <DISS_degree>, <DISS_institution>, <DISS_advisor>, <DISS_cmte_member>, and <DISS_categorization> - this element describes the graduate work being submitted for both processing and access purposes.
 - a. **page_count attribute:** The number of pages in the graduate work
 - b. **type attribute:** graduate work type – “masters” or “doctoral”
 - c. **external_id attribute:** a unique id assigned by the transmission service. For FTP sites, it should begin with a lower case school identifier such as hkust or emory, followed by a colon, and then the school's own unique identifier (i.e., “emory:31”).
 - d. **apply_for_copyright attribute:** indicate if the author want us to apply for copyright. There is an additional fee for copyright registration, and for FTP sites, the author must submit a signed paper form with the additional payment.
15. **<DISS_dates> element:** Used to capture dates related to the graduate work. Has two child elements - <DISS_comp_date> and <DISS_accept_date>.
16. **<DISS_comp_date> element:** Used to capture the year the degree was conferred. Should be provided as a year, in “yyyy” format.
17. **<DISS_accept_date> element:** Used to capture the date the manuscript was completed. Should be provided in “mm/dd/yyyy” format. If only year is available, default to “01/01/yyyy”.
18. **<DISS_degree> element:** The acronym for the degree granted (i.e., “Ph.D.”). Should be drawn from the degree_codes.xls file referenced above under Related Documents.
19. **<DISS_institution> element:** Along with its child elements - <DISS_inst_code>, <DISS_inst_name>, and <DISS_inst_contact> - this element used to capture institutional affiliation. <DISS_inst_code> is your university school code, which ProQuest can provide, and <DISS_inst_name> is the full name of the degree-granting institution. <DISS_inst_contact> is used to capture the department in the university to which the author belonged. It is an optional element.
20. **<DISS_categorization> element:** Along with its child elements - <DISS_category>, <DISS_keyword>, <DISS_ISBN>, and <DISS_language> - are used describe the graduate work in order to enhance discoverability. Note that <DISS_ISBN> is optional, and should only be included if an ISBN has already been assigned to the graduate work in question. Otherwise, UMI will assign an ISBN as part of the publishing process.
21. **<DISS_category> element:** Along with its two child elements - <DISS_cat_code> and <DISS_cat_desc> - this element is used to assign UMI-defined subject categories to each graduate work. At least one category must be assigned by the author, and the author should have the option to add up to two additional categories as well. The values for <DISS_cat_code> and <DISS_cat_desc> must conform to the category codes and descriptions provided in the umi_subjects.xls file, described above under Related Documents.

22. **<DISS_keyword> element:** This element is optional. Authors can use it to assign up to 6 "free form" search descriptors that further describe the content of the graduate work.
23. **<DISS_language> element:** This element captures the code for the language in which the graduate work is written. The default is "EN" for English. FTP sites must conform to the list of language codes provided in the language_codes.xls file, as described above under Related Documents.
24. **<DISS_content> element:** This element and its two child elements - <DISS_abstract> and <DISS_binary> - are used to capture information about the full text of the graduate work itself. It is required.
25. **<DISS_abstract> element:** A child element of <DISS_content>, this element is used to capture a plain text version of the abstract. Each paragraph of the abstract should be included in the child <DISS_para> element.
26. **<DISS_binary> element:** A child element of <DISS_content>, this element is used to capture the format of the full text (must be a PDF) and the file name of the full text document. The "type" attribute is used to capture the format, and should always be "PDF". The file name should be something like "barbarino_vinny.pdf".

The End Result

With the metadata and full text PDF in hand, our Dissertations Operations staff will begin the publishing process for each FTP submission your university sends to us. Once it is published, a citation record will be available in the ProQuest Dissertations & Theses (PQDT) database:

Databases selected: Dissertations & Theses

Document View [« Back to Results](#)

[Print](#) | [Email](#) | [Copy link](#) | [Cite this](#) | Mark Document
 Translate abstract from:

Development of an expansible taxonomy of software requirements nomenclature: Viewing the elicitation process as communicative discourse

by [D'Ambrosio, Donna Maria](#), Ph.D., Long Island University, C. W. Post Center, 2006, 338 pages; AAT 3265667

Abstract (Summary)
 This research was undertaken to achieve an understanding of the mental conceptualizations of systems analysts for requirements nomenclature (expressions) used in research and practice. New approaches for requirements elicitation are necessary due to ambiguous requirements cited as the cause for software failures. The primary objective of this research is to develop an expansible taxonomy of requirements nomenclature using personal conceptualizations of the meaning of requirements nomenclature as a guide in the development of a requirements taxonomy, which can be expanded as nomenclature and nature of a particular domain changes. The techniques used to gather data included a combination of questionnaire to obtain background information from participants, discourse analysis of an email list used by practitioners and literature to obtain requirements expressions, repertory grid to obtain mental conceptualizations of the requirements expressions and term sorting which revealed an overall conceptual understanding of the relational information about language in use during requirements elicitation.

It was postulated for this research that common nomenclature is used during requirements elicitation and specification discourse, but that the mental conceptualizations of nomenclature meaning, although similar among the participants, will differ in usage and reflect variations affected by education and work experience. These results, including noted differences, provided the guidelines for the development of an unnumbered, three-level taxonomy of common nomenclature.

The taxonomy developed into a hierarchical-faceted conceptual structure representing emergent themes discovered from the repertory grid content analysis. The emergent categorical themes (Level A labeled "Category") were developed during content analysis of each participant's constructs. The constructs represent the way in which the individual construes the expressions through the repertory grid process. Level B, which is labeled "Facets", represents the actual constructs given by the participants. The constructs, in turn, were elicited during the repertory grid interviews using the common requirements expressions used in discourse. These expressions represent Level C and are labeled "Basic-level expressions". Although relatively few Level-A categories emerged, it was found that the same basic-level expressions that comprise the every-day discourse were used in different categorical areas by the same individual as well as by different individuals. That is, for some individuals, an expression, such as "quality requirement" would be used in the context of testing and measurement, point-of-view or when defining user needs. This shows a subtlety of understanding by an individual, which could hinder the requirements development process.

With this knowledge and the use of a structured interviewing technique like repertory grid to develop a taxonomic conceptual aid that can be changed and expanded for the contextual needs of the work domain, the true conceptualizations of the stakeholders and systems analysts can be revealed to afford a deeper understanding of the domain language of the domain in which the systems development is taking place.

Indexing (document details)

Advisor:	Chu, Heting
School:	Long Island University, C. W. Post Center
School Location:	United States -- New York
Keyword(s):	Expansible taxonomy , Software requirements , Nomenclature , Elicitation , Communicative discourse
Source:	DAI-A 68/05, Nov 2007
Source type:	Dissertation
Subjects:	Library science , Information systems
Publication Number:	AAT 3265667
ISBN:	9780549044934

Other available formats:

[24 Page Preview](#)
 [Full Text - PDF \(18 MB\)](#)
 [Order a Copy](#)

Find more documents like this:

Subjects:

Library science

Information systems

Author(s):

[D'Ambrosio, Donna Maria](#)

Degree Date:

2006

Advisor:

Chu, Heting

School:

Long Island University, C. W. Post Center

Keyword(s):

Expansible taxonomy

Software requirements

Nomenclature

Elicitation

Communicative discourse

Publication title:

ProQuest Dissertations and Theses

[close ↑](#)

ProQuest Dissertations & Theses (PQDT): Citation View

In addition to the bibliographic record, a digital copy (i.e., the PDF) is stored in our digital vaults, and a microform copy is created and stored in our microform vaults for long term preservation.

A unique, persistent URL for each graduate work

Each citation has a unique and persistent URL based on the OpenURL syntax. The OpenURL has two components, and they are easy to construct. The bulk of the string is the same from URL to URL, with the unique component being the UMI Publication Number appended onto the end (minus the "AAT" prefix). Here is what one looks like:

- http://gateway.proquest.com/openurl?url_ver=Z39.88-2004&res_dat=xri:pqdiss&rft_val_fmt=info:ofi/fmt:kev:mtx:dissertation&rft_dat=xri:pqdiss:7706369

The two components are as follows:

- Component 1: http://gateway.proquest.com/openurl?url_ver=Z39.88-2004&res_dat=xri:pqdiss&rft_val_fmt=info:ofi/fmt:kev:mtx:dissertation&rft_dat=xri:pqdiss : [the same from URL to URL]
- Component 2 (UMI Publication Number): 7706369 [unique for each different URL]

The citation, abstract, and a 24 page preview will be available to any user that follows a properly formed graduate work URL. If an end user is authenticated as having rights to the full text, that end user will also have free access to the full text PDF as well.

Questions?

Contact us at scholarlypublishing@proquest.com with any questions you might have.