

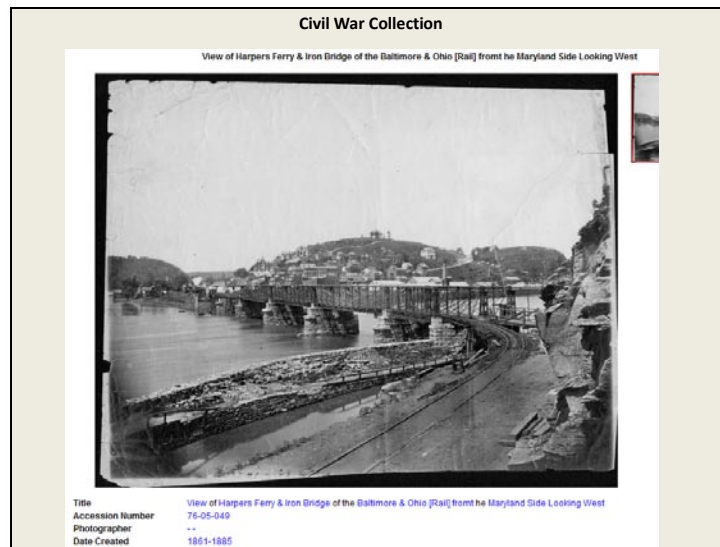


Electronic Thesis and Dissertations in Acquisitions

UMBC has some ContentDM Digital Collections and includes with them types of materials that fit better in the Intuitional Repository Category. However, UMBC doesn't consider ContentDM an Institutional Repository software, and on investigation, we found it unsuited to serve as one. University of Utah does use Contentdm as a repository, but only with programming added on to it to make it work as one. OCLC is currently working with user groups to identify features lacking that prevent it from being used as one, but has not made a commitment to adding features for that purpose.

A primary emphasis of UMBC's Special Collections is photography and the history of photography, and we had digital images only before we purchased Contentdm in 2006. Contentdm was originally purchased for digitized photos and we didn't intend to use it for other purposes but have done so since it's the platform we currently have available. I'm going to show you some images from our original collections which went live in the fall of 2009 so you'll understand what our main purpose and use for ContentDM.

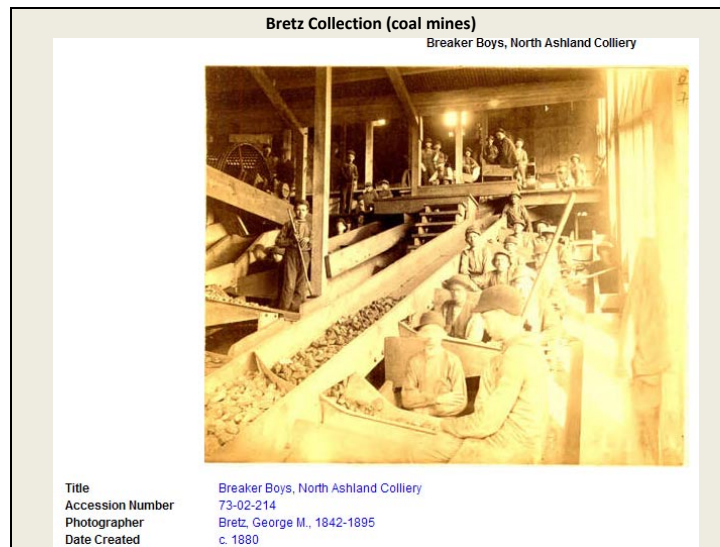
Slide 2



Civil War Collection

This is an image from our Civil War Collection of Harper's Ferry during the Civil War.

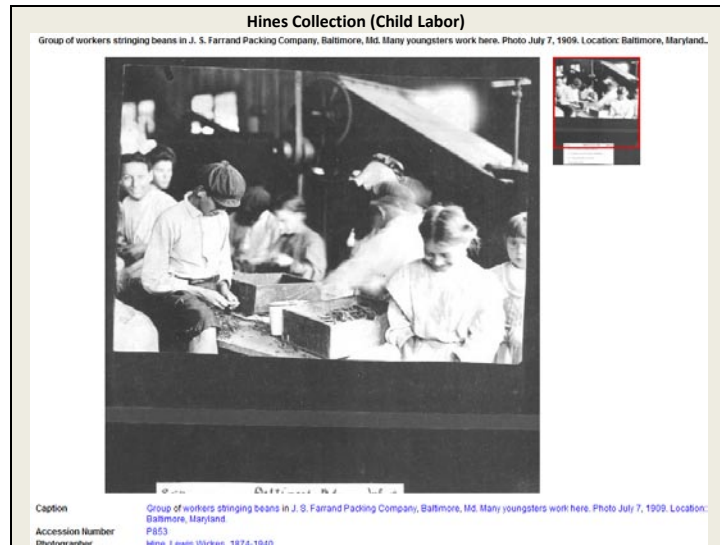
Slide 3



Bretz Collection

This is our Bretz Collection, photographs of coal mines and coal miners.

Slide 4



Hines Collection

This is an image from our Hines Collection, which are photographs of child labor. There are children working for a packing company.


All of our photos were digitized in Special Collection. Many were digitized and available in html before we purchased contentdm. Staff responsible for bringing up our digital collections ContentDM are Susan Graham, Special Collections Librarian, and Lindsey Loeper, University Archivist, And in Technical Services, cataloger Vicki Sipe who helped with the metadata. Web Designer Janet Hack designed our public Contentdm website which went live in 2009.

Slide 5

University Archives


The University Archives holdings contain historical records of UMBC — including the campus, staff, students, and alumni. Department records, campus photographs, university publications, and faculty papers are available. Selections included in the UMBC Digital Collections include photographs, illustrations, publications, promotional materials, reports, and select correspondence as digitization efforts allow. The items included here do not represent the full extent of our holdings; please email specialcollections@umbc.edu for further information.

University Photographs
[Browse Photographs >>](#)




Pulled from the University Archives holdings, these photographs provide a visual history of UMBC. The photographs date to the school's original construction in 1965 and they highlight the development and construction of the campus, university events, and student life.

Theatre Department Production Materials
[Browse Materials >>](#)




Theater programs, posters, photographs, and newspaper clippings are available for productions from the UMBC Theatre Department and include production materials for the Maryland Stage Company and Shakespeare on Wheels, dating from 1968 to 2007. The contents are grouped by the title of play. Not all items are currently available in the UMBC Digital Collections and the productions are being added in order by date beginning with the 1990s. Please contact Special Collections staff if you do not see the production you are looking for. Supplementary production records are available in the Special Collections department.

University Publications
[Browse Publications >>](#)



This growing collection of historical university material is primarily born-digital research findings, newsletters, and annual reports produced by UMBC departments and centers. A large number of non-digital university publications are also available in Special Collections; more information about browsing the full collection is available on the University Archives webpage: <http://iuk.lib.umbc.edu/specialcollections.php>

The Retriever Weekly
[Browse The Retriever Weekly \(1968-1977, 2002-2010\) >>](#)
[Browse Volume 44 \(2009-2010\) >>>](#)
[Browse Volume 43 \(2008-2009\) >>](#)



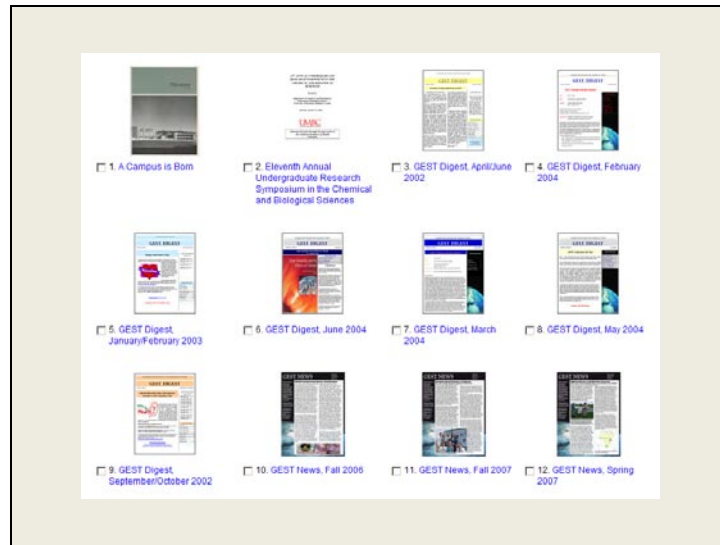
The Retriever Weekly is the student newspaper for the University of Maryland, Baltimore County (UMBC). The premier issue was published for students and staff on September 15, 1968, during the first semester of classes. That first UMBC issue, our newspaper began the tradition of reporting student and campus life as UMBC continued to grow.

[Digital Collections Home](#)
[Photography Collections](#)
[Archives](#)
[Electronic Theses & Dissertations](#)
[Other Digital Resources](#)
[Feedback](#)

University Archives

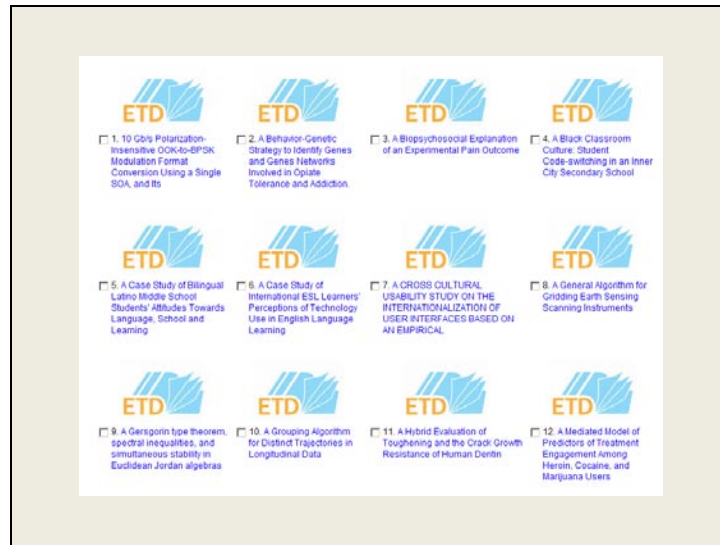
We also have digital University Archives, part of which is digitized print collections, but some is being sent to us in a digital format for inclusion in the Digital Archives.

Slide 6



GEST (Goddard Earth Science Technology Center) publications are examples materials being sent to us in a Digital Format for inclusion in our Digital Collections. These types of materials straddle our definitions of Digital Collections and Institutional Repository, in that they are already in our collection in print, or would be included in print, but they are materials being submitted to us electronically by the author for us to provide access and preservation.

Slide 7



We also have Electronic Thesis and Dissertations in our Contentdm Digital Collections, although most usually ETD's are considered part of an IR. I've talked to a library that currently utilizes both Contentdm and an IR software, and they've indicated that there are materials that are particular suited to one or the other of the systems, but in instances where the line is blurred between the two, they chose based on what is easiest and works best.

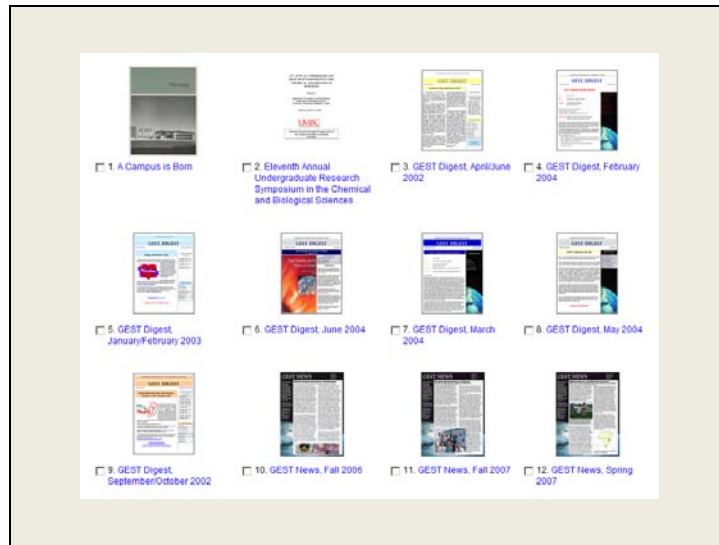
Transfer-Related Services
or
Digital Transfer Service

- Acquiring digital content
- Conducting quality review of digital
- Moving digital content between systems
- Inventorying
- Manipulating
- And ingesting it into the appropriate application or repository.

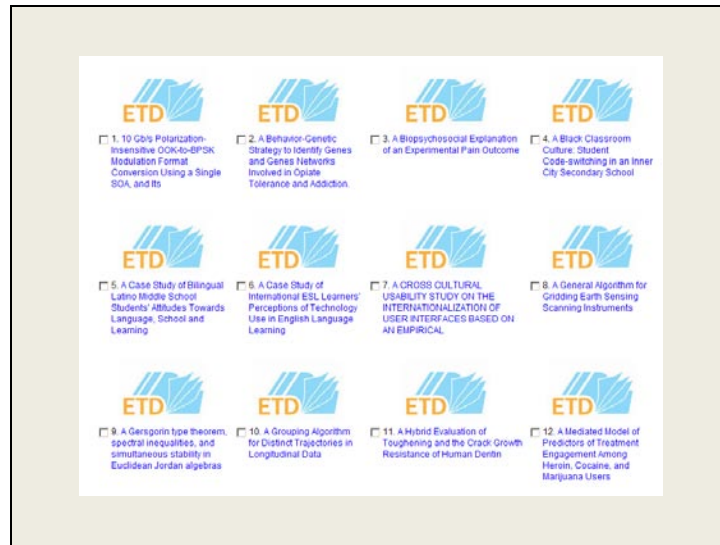
As we began to bring our Digital Collections up In February 2009, I proposed that my Acquisitions Unit take responsibility for Transfer Related Service, which I'd read about in a blog post. I later reconceived this as Digital Transfer Services. I began talking with Special Collection regarding materials coming into ContentDM and support Acquisitions might be able to provide. My concept of what Digital Transfer services is includes:

- Acquiring digital content
- Conducting quality review of digital content
- Moving digital content between systems
- Inventorying
- Manipulating
- And ingesting it into the appropriate application or repository.

Slide 9



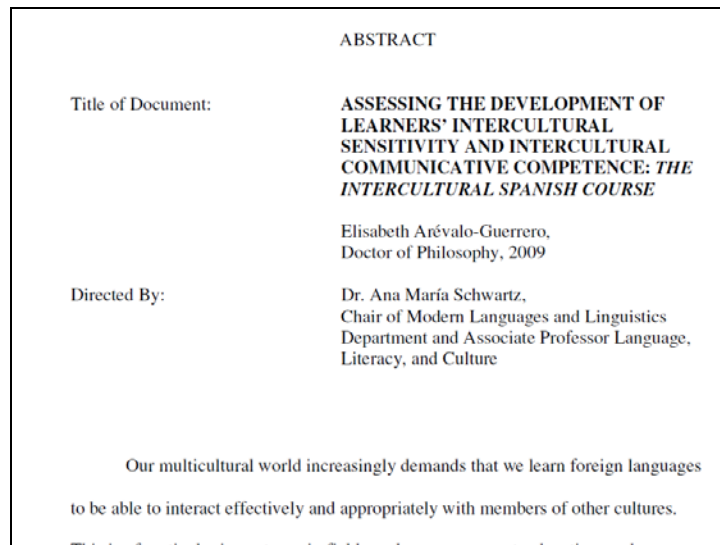
I had in mind that we could work with the the e-serial files being periodically acquired by receiving the files, inventorying them, checking them, and then loading them into ContentDM.



But the problematic workload at the time was ETD's, which were receive from Proquest along with metadata for them, and this was what they wanted to give me. The librarian who had been doing them had resigned, the programming that reformatted the metadata had broken and we no longer had a programmer to fix it. Special Collections took the work temporarily and was having a student assistant manually reformat the data. The work was supposed to shift to back to the position it had come from when that position was filled but the workload for that position was already heavy so the work came to Acquisitions instead.

Assessing the development of learners' intercultural sensitivity and intercultural communicative competence: The Intercultural Spanish Course	
Access this item.	
Title	Assessing the development of learners' intercultural sensitivity and intercultural communicative competence: The Intercultural Spanish Course
Author	Arvalo, Elisabeth
Advisors	Schwartz, Dr. Ana Maria
Program	Literacy and Culture
UMBC Department	Language, Literacy & Culture (LLC)
Document Type	dissertation
Sponsors	Digital Repository at the University of Maryland, Baltimore County
Keywords	Assessing intercultural competence; intercultural communicative competence; Intercultural sensitivity; Intercultural Spanish course
Date Issued	2009-01-01;
Abstract	Our multicultural world increasingly demands that we learn foreign languages to be able to interact effectively and appropriately with members of other cultures. This is of particular importance in fields such as government, education, and commerce. Effective intercultural interaction entails developed intercultural sensitivity, knowledge of cultural aspects as well as the ability to communicate in the language. This is known as Intercultural Communicative Competence (ICC). The study of effective intercultural communication should be taught alongside the linguistic aspect of language. Yet the teaching of culture in foreign language classes is often superficial. There is no consensus with regard to how to develop and assess intercultural sensitivity and ICC in the foreign language classroom. This study hypothesizes that an intercultural approach to teaching foreign languages facilitates the development of the learners' intercultural communicative skills and linguistic proficiency while raising cultural awareness, developing contextualized cultural knowledge, and fostering intercultural attitudes of respect and understanding for one's own culture and the cultures of others. The study analyzes and assesses the development of intercultural sensitivity and ICC in a university advanced conversation and composition Spanish course called here the Intercultural Spanish Course. The curriculum for this course was designed specifically for this study and integrates linguistic and intercultural communication skills and knowledge. A mixed methods design was used: A) Quantitative instruments: 1. Hammer & Bennett's (2001) Intercultural Development Inventory (IDI-B) Qualitative instruments: 1. Pre/ post learners' questionnaires 2. Pre/ post learners' journals Bennett's (1993) Developmental Model of Intercultural Sensitivity (DMIS) was used to analyze the IDI pre/ post results. For the qualitative data analysis themes were categorized following Byram's (1997) Model of Intercultural Competence (MICC). The results of the study show that 50% of the learners developed their intercultural sensitivity toward an ethnocentric orientation whereas 50% moved toward an ethnocentric orientation. Learners developed their ICC as they (1) gained deeper cultural knowledge of Spanish-speaking countries; (2) developed critical thinking skills and skills to interpret cultural differences in a more ethnocentric perspective; (3) showed intercultural attitudes of understanding other perspectives, interest in engaging in conversation with Spanish speakers and willingness to question their worldviews; (4) raised self-cultural awareness; and (5) improved their language proficiency.
Identifier	10107
Format	application/pdf
Language	en
Collection	UMBC Theses and Dissertations

What I knew about workflow when I agree to take it over was that we had ETD's in ContentDM with metadata that looked like this,



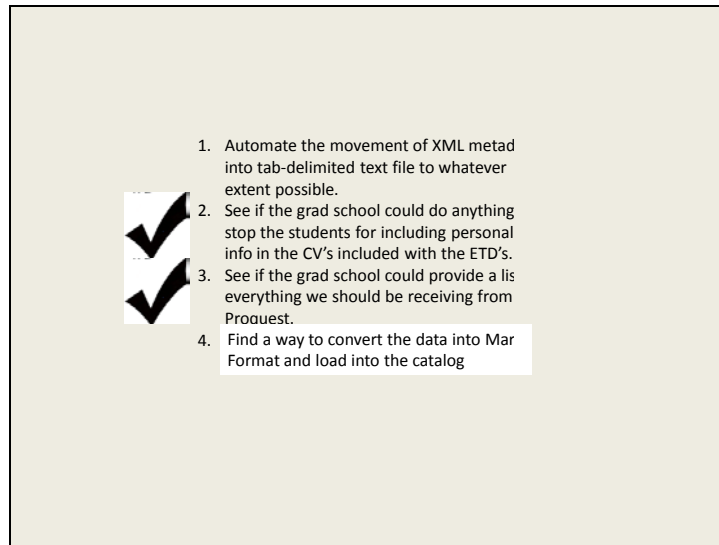
And ETD's in PDF format that looked like this. I went in and looked at the collection with a mind for understanding what was going on with them, how they would work, and how patrons access and view them. I found some issues and Lindsey Loeper, our archivist doing the work, made me aware of some other issues.

1. Automate the movement of XML metadata into the format needed for contentdm, tab-delimited text files.
2. See if the grad school could do anything to stop the students from including personal info in the CV's included with the ETD's.
3. See if the grad school could provide a list of everything we should be receiving from Proquest.
4. Find a way to get links into the records in the catalog.

Initial Tasks

- Automate the movement of XML metadata into the format needed for contentdm, tab-delimited text files.
- See if the grad school could do anything to stop the students from including personal info in the CV's included with the ETD's.
- See if the grad school could provide a list of everything we should be receiving from Proquest.
- Find a way to get links into the records in the catalog.
- Correct ETD values to match those of other collections where appropriate.





1. Automate the movement of XML metadata into tab-delimited text file to whatever extent possible.

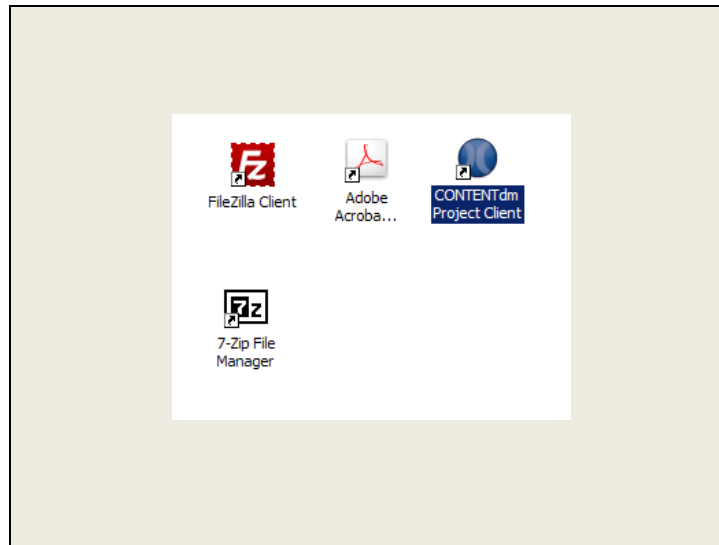
2. See if the grad school could do anything to stop the students from including personal info in the CV's included with the ETD's.

3. See if the grad school could provide a list of everything we should be receiving from Proquest.

4. Find a way to convert the data into MARC format and load into the catalog

Two large black checkmarks are positioned to the left of the list items, indicating that the first two items have been completed.

So it appeared that two of the items on my list were completed, and I added a new task of take over the receipt and payment of hardcopy thesis and dissertations. However, neither of the things agreed to with the grad school have occurred yet. I'm hopeful changes will occur soon, but it may take some follow-up.



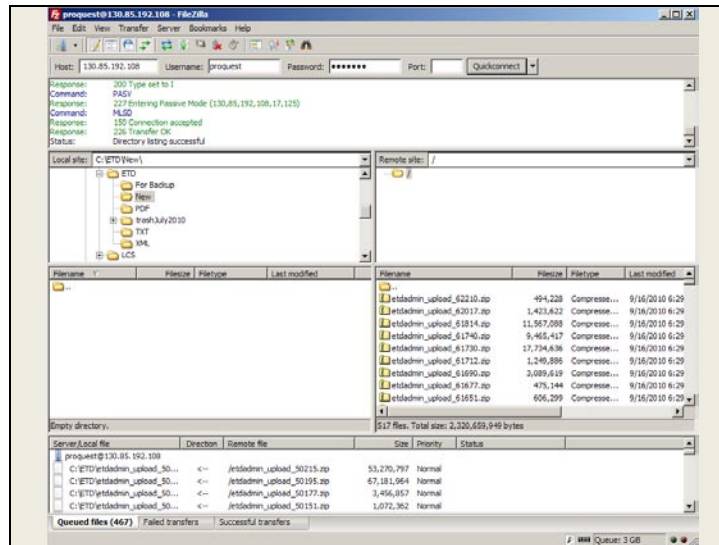
Next I needed to add 3 new pieces of software to my computer, and I added a 4th later when I began looking at the existing procedures. The software needed was FileZilla, an FTP tool for downloading the files, 7-Zip File Manager, a utility for unzipping and moving zip files, Adobe Acrobat, a tool for manipulating .pdf files, and the ContentDM client, used for working with materials to load onto the ContentDM server. The ContentDM also needed to be configured, which was initially done by Lindsey, our Archivist, and Lindsey also provided me with a written procedure outlining how the ETD's were being processed at the time, and walked me through the procedure once.

Slide 17



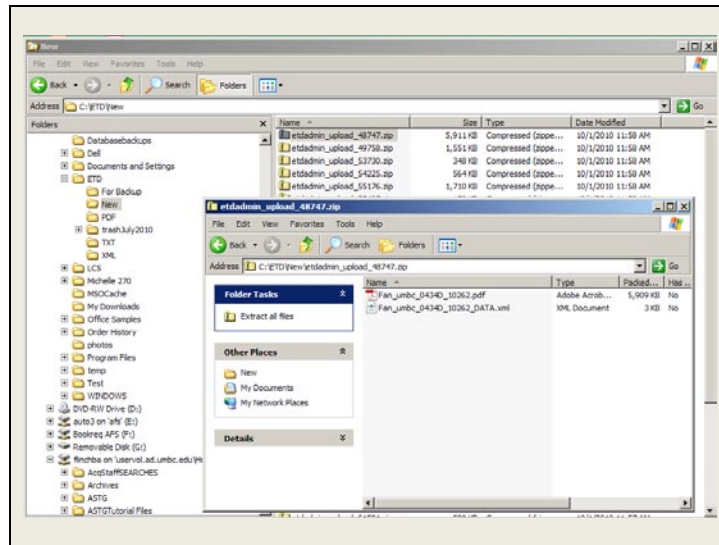
The process works like this. We receive e-mailed notices from Proquest that look like this that we have new files. If you look up at the top of this notice, you'll note that indicates that 29 of 53 files delivered successfully, and 24 of 53 failed to deliver. Usually, the files successfully load, but other times they don't, and thereafter they sometimes come in small bits. This can make it challenging to figure to figure out if you received all of the files or not except by detailed checking for each of the failed loads item-by-item in later notifications.

Slide 18

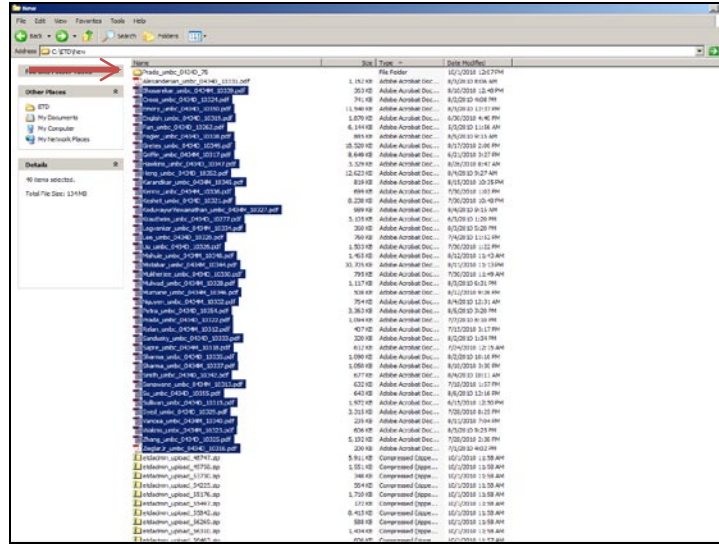


When all the files are on the server, we download using Filezilla, an FTP utility. This is a simple drag and drop utility.

Slide 19



The downloaded files are now on my computer in zip folders. The procedure had been to open each individually and drag and drop each pdf and xml file to the appropriate folder to separate them, but this is what I got the unzip program for.



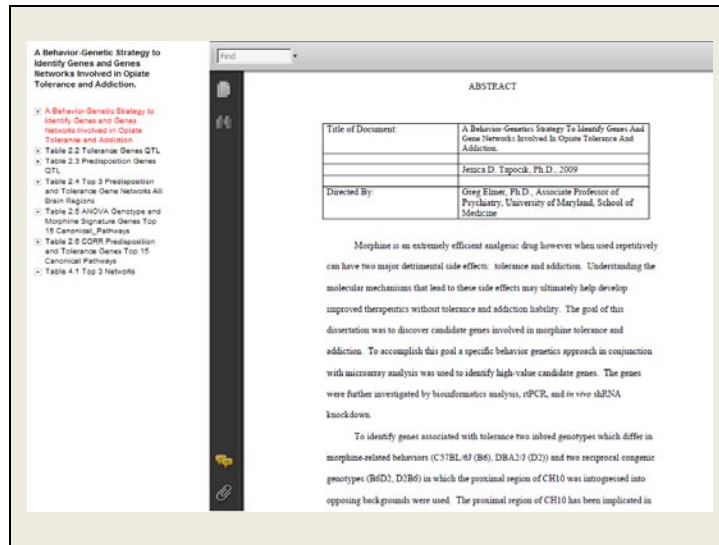
7-Zip and Supplemental Material

In 7-Zip I can find the directory the zip directories are in, highlight them all, then click extract. When you go back to look at the directory and refresh, the pdf and xml files are all in the main directory, so the files can be sorted by type then dragged and dropped as a group. As we look through and sort the files received, most have .xml file and a .pdf file, but others come with the authors name on them and include and extra files that came with the .xml and .pdf files. These may be approval sheets or vitesas, which are usually included in the main .pdf, in separate files. In other cases, they may be supplemental materials, sometimes that aren't included in the main .pdf already, but sometimes they are included already. Sometimes supplemental material files will be in MACOS. If supplemental materials aren't included, files need to be combined into the main .pdf using adobe, or alternatively, if this won't work because of the nature or format, they need to be attached to the main object in ContentDM.

Gene	Symbol	Gene Name	Physical Map Location	Chromosome	Marker ID	Phenotype	Reference	Other
ANKRD12	ANKRD12	ankrd12	chromosome 11p15.5	11	ANKRD12	ankrd12	ANKRD12	ANKRD12
ANKRD13	ANKRD13	ankrd13	chromosome 11p15.5	11	ANKRD13	ankrd13	ANKRD13	ANKRD13
ANKRD14	ANKRD14	ankrd14	chromosome 11p15.5	11	ANKRD14	ankrd14	ANKRD14	ANKRD14
ANKRD15	ANKRD15	ankrd15	chromosome 11p15.5	11	ANKRD15	ankrd15	ANKRD15	ANKRD15
ANKRD16	ANKRD16	ankrd16	chromosome 11p15.5	11	ANKRD16	ankrd16	ANKRD16	ANKRD16
ANKRD17	ANKRD17	ankrd17	chromosome 11p15.5	11	ANKRD17	ankrd17	ANKRD17	ANKRD17
ANKRD18	ANKRD18	ankrd18	chromosome 11p15.5	11	ANKRD18	ankrd18	ANKRD18	ANKRD18
ANKRD19	ANKRD19	ankrd19	chromosome 11p15.5	11	ANKRD19	ankrd19	ANKRD19	ANKRD19
ANKRD20	ANKRD20	ankrd20	chromosome 11p15.5	11	ANKRD20	ankrd20	ANKRD20	ANKRD20
ANKRD21	ANKRD21	ankrd21	chromosome 11p15.5	11	ANKRD21	ankrd21	ANKRD21	ANKRD21
ANKRD22	ANKRD22	ankrd22	chromosome 11p15.5	11	ANKRD22	ankrd22	ANKRD22	ANKRD22
ANKRD23	ANKRD23	ankrd23	chromosome 11p15.5	11	ANKRD23	ankrd23	ANKRD23	ANKRD23
ANKRD24	ANKRD24	ankrd24	chromosome 11p15.5	11	ANKRD24	ankrd24	ANKRD24	ANKRD24
ANKRD25	ANKRD25	ankrd25	chromosome 11p15.5	11	ANKRD25	ankrd25	ANKRD25	ANKRD25
ANKRD26	ANKRD26	ankrd26	chromosome 11p15.5	11	ANKRD26	ankrd26	ANKRD26	ANKRD26
ANKRD27	ANKRD27	ankrd27	chromosome 11p15.5	11	ANKRD27	ankrd27	ANKRD27	ANKRD27
ANKRD28	ANKRD28	ankrd28	chromosome 11p15.5	11	ANKRD28	ankrd28	ANKRD28	ANKRD28
ANKRD29	ANKRD29	ankrd29	chromosome 11p15.5	11	ANKRD29	ankrd29	ANKRD29	ANKRD29
ANKRD30	ANKRD30	ankrd30	chromosome 11p15.5	11	ANKRD30	ankrd30	ANKRD30	ANKRD30
ANKRD31	ANKRD31	ankrd31	chromosome 11p15.5	11	ANKRD31	ankrd31	ANKRD31	ANKRD31
ANKRD32	ANKRD32	ankrd32	chromosome 11p15.5	11	ANKRD32	ankrd32	ANKRD32	ANKRD32
ANKRD33	ANKRD33	ankrd33	chromosome 11p15.5	11	ANKRD33	ankrd33	ANKRD33	ANKRD33
ANKRD34	ANKRD34	ankrd34	chromosome 11p15.5	11	ANKRD34	ankrd34	ANKRD34	ANKRD34
ANKRD35	ANKRD35	ankrd35	chromosome 11p15.5	11	ANKRD35	ankrd35	ANKRD35	ANKRD35
ANKRD36	ANKRD36	ankrd36	chromosome 11p15.5	11	ANKRD36	ankrd36	ANKRD36	ANKRD36
ANKRD37	ANKRD37	ankrd37	chromosome 11p15.5	11	ANKRD37	ankrd37	ANKRD37	ANKRD37
ANKRD38	ANKRD38	ankrd38	chromosome 11p15.5	11	ANKRD38	ankrd38	ANKRD38	ANKRD38
ANKRD39	ANKRD39	ankrd39	chromosome 11p15.5	11	ANKRD39	ankrd39	ANKRD39	ANKRD39
ANKRD40	ANKRD40	ankrd40	chromosome 11p15.5	11	ANKRD40	ankrd40	ANKRD40	ANKRD40
ANKRD41	ANKRD41	ankrd41	chromosome 11p15.5	11	ANKRD41	ankrd41	ANKRD41	ANKRD41
ANKRD42	ANKRD42	ankrd42	chromosome 11p15.5	11	ANKRD42	ankrd42	ANKRD42	ANKRD42
ANKRD43	ANKRD43	ankrd43	chromosome 11p15.5	11	ANKRD43	ankrd43	ANKRD43	ANKRD43
ANKRD44	ANKRD44	ankrd44	chromosome 11p15.5	11	ANKRD44	ankrd44	ANKRD44	ANKRD44
ANKRD45	ANKRD45	ankrd45	chromosome 11p15.5	11	ANKRD45	ankrd45	ANKRD45	ANKRD45
ANKRD46	ANKRD46	ankrd46	chromosome 11p15.5	11	ANKRD46	ankrd46	ANKRD46	ANKRD46
ANKRD47	ANKRD47	ankrd47	chromosome 11p15.5	11	ANKRD47	ankrd47	ANKRD47	ANKRD47
ANKRD48	ANKRD48	ankrd48	chromosome 11p15.5	11	ANKRD48	ankrd48	ANKRD48	ANKRD48
ANKRD49	ANKRD49	ankrd49	chromosome 11p15.5	11	ANKRD49	ankrd49	ANKRD49	ANKRD49
ANKRD50	ANKRD50	ankrd50	chromosome 11p15.5	11	ANKRD50	ankrd50	ANKRD50	ANKRD50

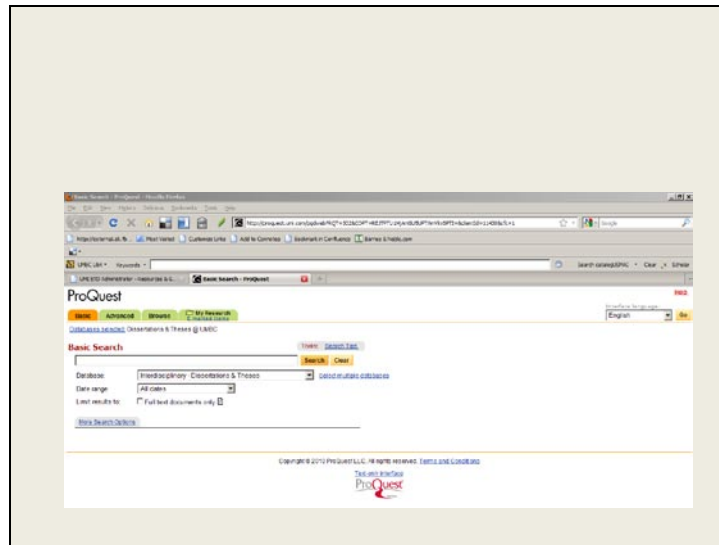
Excel Appendixes Spreadsheet

In this instance this dissertation included datasets in a number of Excel files with genetic codes and I couldn't find a way to convert this into a usable .pdf .



Excel Appendixes Spreadsheets in Contentdm

So I ended up attaching in Excel format for user download. In an instance like this where this isn't a clear and documented method for handling, the staff doing this work sends all of the files to me to handle. In simpler instances where materials that are usually together in the main .pdf but are separated, or there is a supplemental .pdf that can be combined with the main .pdf. procedures are in place for staff to handle.



Missing ETD's


Once we have the extra files figured out, we have to open all of the ETD's to delete the CV's because of the personal information in them. In some instances when we've opened the PDF file to delete the CV, we've found that there is no thesis or dissertation included in the file, but only the approval sheet. Initially I searched these in Dissertation Abstracts to confirm that they all had embargoes on them, but there were also missing files that weren't embargoed. At this point I contact Proquest for more information and found that I could determine if there were embargoes on items from the metadata, as well as the date when the embargoes ends. Additionally I found out that the missing files without embargoes were too large for the authors to upload to Proquest so had been sent on disk via mail, in which case we don't get the files automatically but have to go to a Proquest Website to retrieve them.

Embargos


```
</ui><abstract>
<DISS_binary type="PDF">Seabottom_1458M_10000.pdf</DISS_binary>
</DISS_content>
- <DISS_restriction>
<DISS_restriction code="1" @remove="11002011" />
</DISS_restriction>
</DISS_submission>
```

sales_restriction code =
"1" – Not Available for Sale
"2" – Sell to Author Only/Copyright
"3" – Sell to Author Only
REMOVE – Lift Date of the Restriction

Keywords
Cadherin, Microtubule... (1)



1. MECHANISTIC INSIGHTS INTO SECONDARY NEURULATION



2. Mechanistic Insights Into Secondary Neurulation

[select all](#) : [clear all](#) : [add to favorites](#)

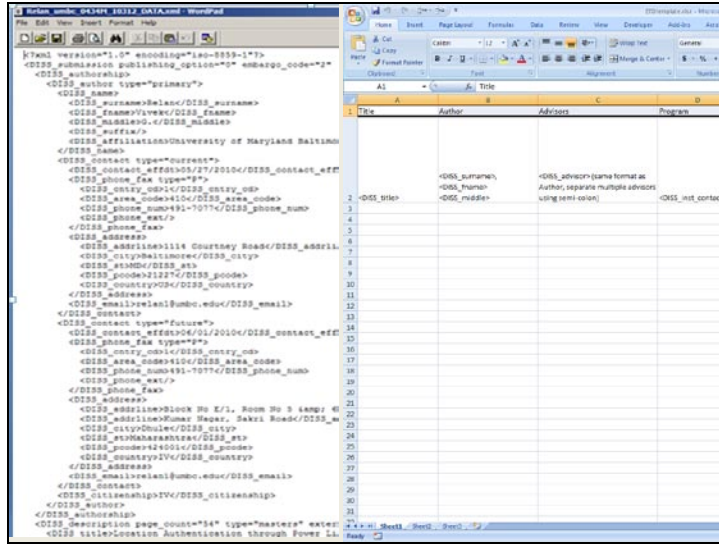
results 1-2 of 2 item(s) page 1 of 1 : (<< 1 >>) : [previous](#) : [next](#)

Embargoes

The embargoed document have an abstract which we load the embargoed documents into ContentDM without the document, and expect to receive and load the full document later. We maintain a list of embargoed ETD's to check on them later. When I checked on this one I found that we had received and loaded the full ETD, so that the partial one now needed to be deleted.

- Received notice from Proquest that files are available
- FTP'd files
- Unzipped files
- Sorted files into folders for XML and for PDF's
- Figured out what to do with extra files
- Noted embargoed documents on our "embargo list"

So at this point, we've Received notice from Proquest that there are files available, FTP'd files, Unzipped files, Sorted files into folders for XML and for PDF's, Figured out what to do with extra files, Noted embargoed documents on our "embargo list," and we've retrieved missing files from the UMBC Proquest website.



Next we need to get the metadata we need from the XML files into an Excel spreadsheet that can be saved as a text-delimited file for import into ContentDM. This was being done by manually locating each needed piece of data in the individual XML file then copying and pasting it into the spreadsheet. I imagined this to be very time-consuming and tedious for just 1 file much less for 50 or 60,

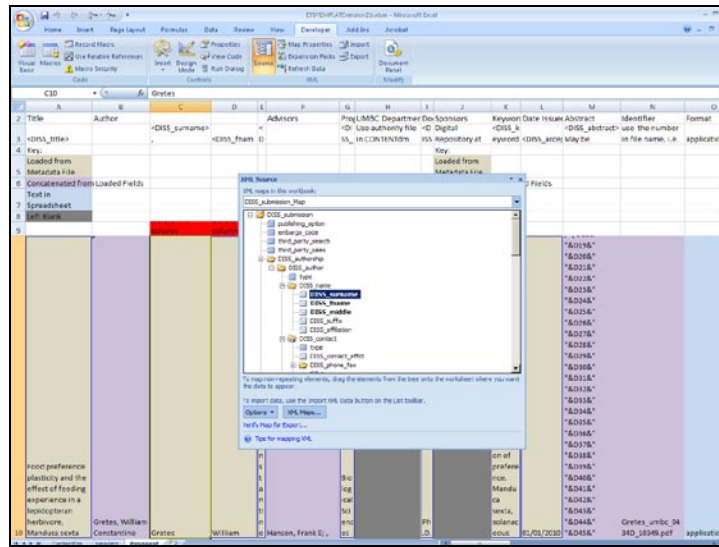
Slide 27

Title	<DISS_title>
Author	<DISS_surname>, <DISS_fname> <DISS_middle>
Advisors	<DISS_advisor> (same format as Author, separate multiple advisors using semi-colon)
Program	<DISS_inst_contact>
UMBC Department	Use authority file in CONTENTdm
Document Type	<DISS_degree> use dissertation or thesis
Sponsors	Digital Repository at the University of Maryland, Baltimore County
Keywords	<DISS_keyword> Separate keywords by semi-colon. Example: Education, Technology (0710); Education, Bilingual and Multicultural (0828); Computer-assisted Language Learning; CALL; Technology Integration; English as a Second Language; ESL; Intensive English Programs; IEP
Date Issued	<DISS_accept_date>
Abstract	<DISS_abstract> May be separated by <DISS para> - if so combine into one field and remove the <DISS para> and </DISS para> tags.
Identifier	use the number in file name, i.e. 10065 for Angara_umbc_0434D_10065.pdf
Format	application:pdf
Language	<DISS_language> (probably "en")
Collection	UMBC Theses and Dissertations
Digital - Image Notes	Notes about the files. Include reference to supplemental files if available (mpeg, jpeg, etc & storage location)
Source	File name, <DISS_binary> Include file extension (.pdf)

Metadata Map

So I utilized the materialize Lindsey had given me to make an easy to read map for the metadata and it's relationship to the XML and began to automate using Excel advanced operations and Excel XML Tools and Macro Developer Utilities. While the process didn't end up quite as effortless as I wanted, it came up vastly better way than copying and pasting bit by bit. The metadata conversion process works like this:

Slide 28



Excel will automatically map metadata for you when import the first XML file of a particular type. It maps all of the metadata to the spreadsheet, and in the order of the XML. We don't need all of it, and needed it in the ContentDM order, so I modified Excel's mapping. This is done by clicking source. In the XML source map, you click any bolded element to see where it's mapped to. The yellow highlighted field is where the author surname maps to. By right clicking an element you can remove it, and you can drag and drop any element to any cell on the spreadsheet and it will mapped to that spot. All of the brown fields in the template spreadsheet or directly mapped from the XML.

Slide 29

Description

The **CONCATENATE** function joins up to 255 text strings into one text string. The joined items can be text, numbers, cell references, or a combination of those items. For example, if your worksheet contains a person's first name in cell A1 and the person's last name in cell B1, you can combine the two values in another cell by using the following formula:

```
=CONCATENATE(A1," ")B1
```

The second argument in this example (" ") is a space character. You must specify any spaces or punctuation that you want to appear in the results as an argument that is enclosed in quotation marks.

Syntax

```
CONCATENATE(text1, [text2], ...)
```

The **CONCATENATE** function syntax has the following arguments:

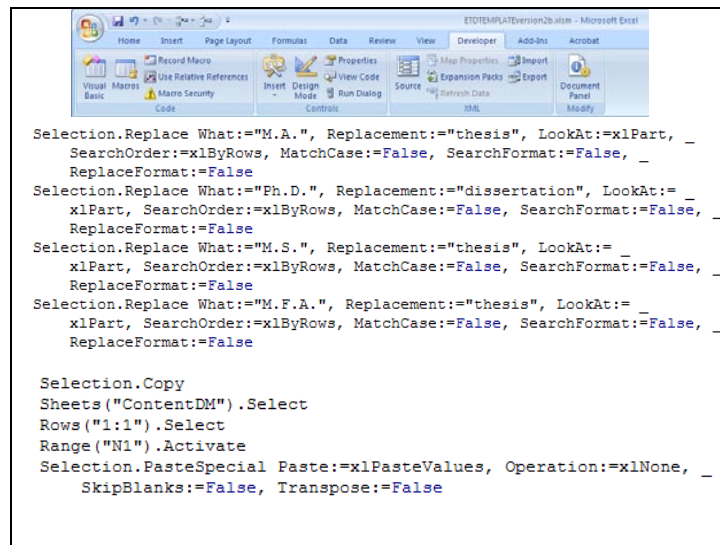
- text1** Required. The first text item to be concatenated.
- text2** Optional. Additional text items, up to a maximum of 255 items. The items must be separated by commas.

NOTE: You can also use the ampersand (&) calculation operator instead of the **CONCATENATE** function to join text items. For example, =A1 & B1 returns the same value as =CONCATENATE(A1, B1).

Example

=CONCATENATE(A18,"",&B18&","&C18)	Hanson, Frank E
-----------------------------------	-----------------

In many instances data is separated in the XML that needs to be combined for contentDM. An important function for doing this is the Concatenate function. This can be done either using a function that begins =concatenate as shown above, or with an & to join text terms. Characters you want to appear between the text terms are put in quotes. For example, the above function puts together XML mapped metadata for author last name, first name, and middle initial, inserting a comma between the last and first names and a blank space between the first name and middle initial.



```

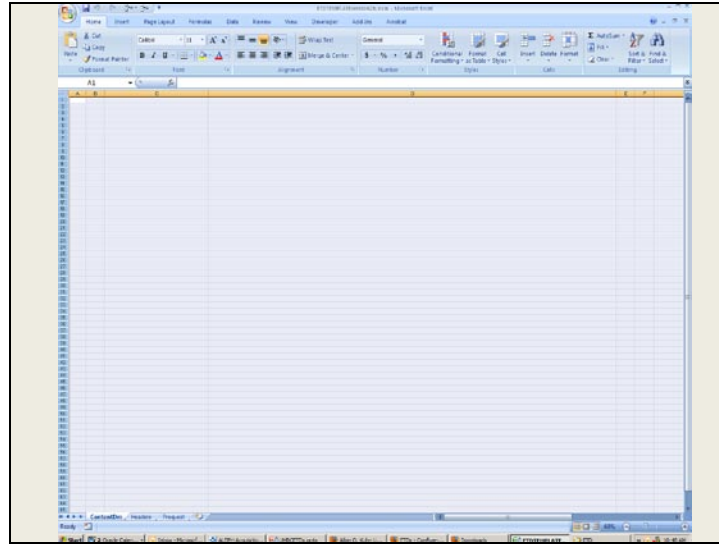
Selection.Replace What:="M.A.", Replacement:="thesis", LookAt:=xlPart, _
SearchOrder:=xlByRows, MatchCase:=False, SearchFormat:=False, _
ReplaceFormat:=False
Selection.Replace What:="Ph.D.", Replacement:="dissertation", LookAt:=
xlPart, SearchOrder:=xlByRows, MatchCase:=False, SearchFormat:=False, _
ReplaceFormat:=False
Selection.Replace What:="M.S.", Replacement:="thesis", LookAt:=
xlPart, SearchOrder:=xlByRows, MatchCase:=False, SearchFormat:=False, _
ReplaceFormat:=False
Selection.Replace What:="M.F.A.", Replacement:="thesis", LookAt:=
xlPart, SearchOrder:=xlByRows, MatchCase:=False, SearchFormat:=False, _
ReplaceFormat:=False

Selection.Copy
Sheets("ContentDM").Select
Rows("1:1").Select
Range("N1").Activate
Selection.PasteSpecial Paste:=xlPasteValues, Operation:=xlNone, _
SkipBlanks:=False, Transpose:=False

```

I used macros to change values and manipulate data in many different ways. Most commonly I used find and replace macros or copy and paste macros as shown above. Macros can be recorded and written in visual basic. I generally use a combination of the two. I'll begin by recording, then edit to add on to that macro. For things like find/replaces, it's very simple to copy a find/replace macro and change the values to run an additional find and replace. Other macros may be more difficult to edit such as the copy/paste at the bottom with a pastespecial with variables set appropriately. Sometimes, instead of completely recording or editing, I'll re-record the snippet that needs correction, then copy the visual basic for that snippet and paste it over the part of the macro that needs to be fixed.

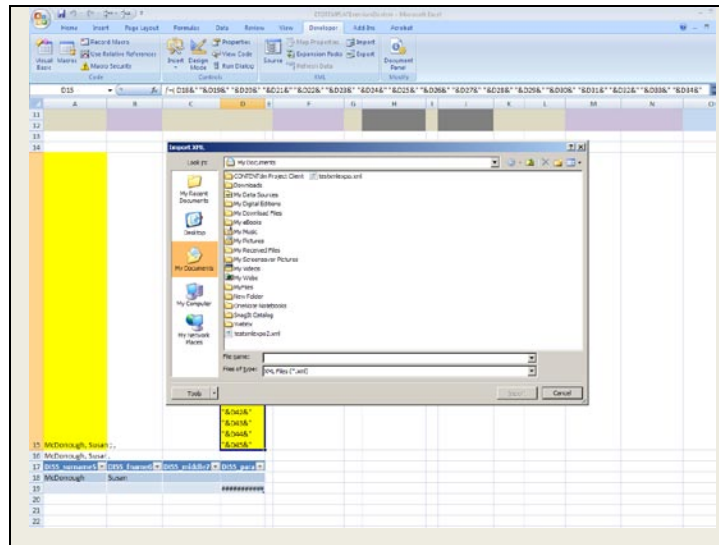
Slide 31



Excel Deletions

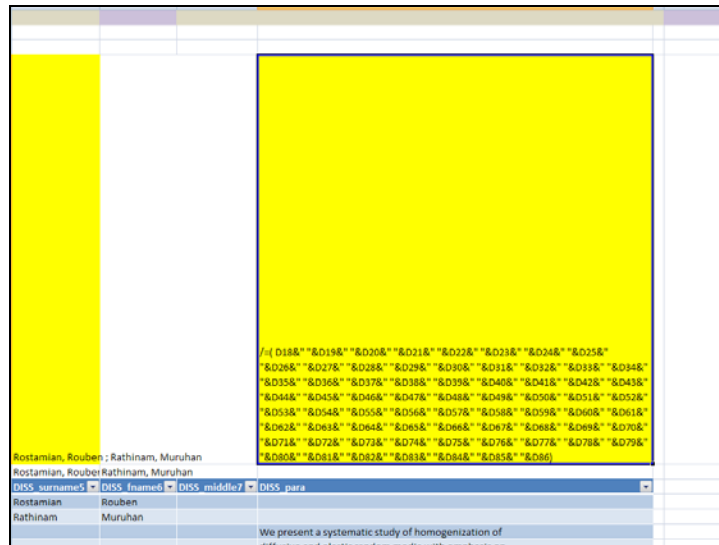
Open the ETD Metadata Template Excel file. Go the ContentDM tab and push CTRL-X to remove all old content...

Slide 33



S-30--XML Import

Click Developer and import and select the first XML file to import



Concatenate

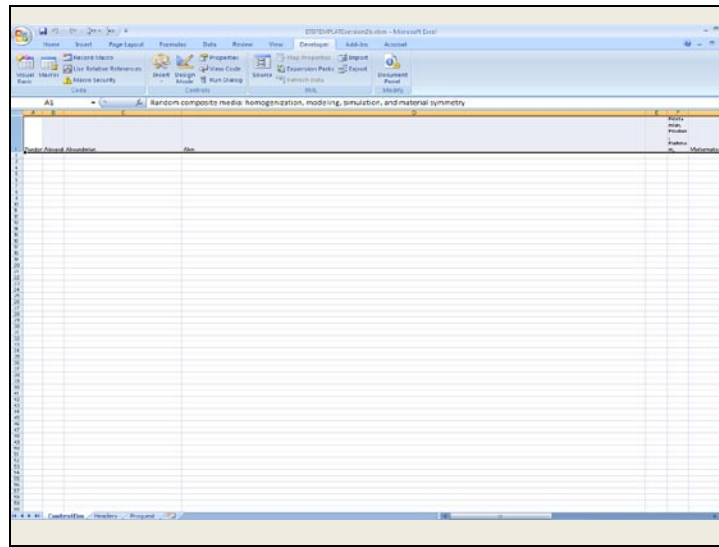
At the bottom of the spreadsheet, cells with concatenate functions and macros combine multiple advisor and dissertation paragraph fields without the user doing anything



Move

At the top of the spreadsheet, a variety of simpler Excel functions reformat data in a variety of ways, again without the staff person doing anything. The staff person just runs a “mover” macro by pushing CTRL-M

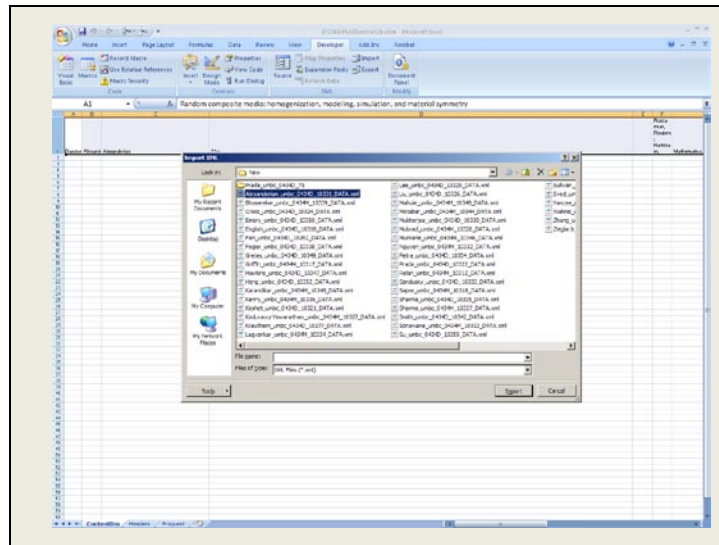
Slide 36



Into Content DM Worksheet

That moves the data into the ContentDM sheet

Slide 37



Move and Move and Move

While the user remains in the Proquest sheet continuing to Import, then CTRL-M each file. When all the files are loaded and moved, the user then goes to the complete ContentDM load file.

Slide 38

1	1	1	1
2	2	2	2
3	3	3	3
4	4	4	4
5	5	5	5
6	6	6	6

S-35--Contentdm Sheet Ctrl-R Reformat

And pushes CTRL-R to further reformat the data to be compatible with our ContentDM data.

ID	Title	Author	Advisor/Program	UMBC	Document Type	Keywords	Core strand
8	Learn/Last Name, First Middle				Thesis	Machine Learning, Machine Learning	
9	Food/Fruit	Haroon	Biological Sciences		Thesis	Feeding behavior, induction of oligophagy, induction of preference, Manduca sexta, zooplankton	090920
10	A state	Fugler, T. Nancy	Neurology		Thesis	home- and community based services, long term care, Medicaid, nursing home	090920
11	The EF	Fu, Qa, Anshou	Information Systems		Thesis	Strategic Cognitive Process, Comments, Eye tracking, Program Completion	090920
12	Learn/Last Name, First Middle				Thesis	language learning, machine learning, etc.	090920
13	Vibrato	Limoy, J. Vinson	Lighting/Mechanical		Thesis	Damage Detection, Elevator, Testing, Engineering Education, Inverse Problems, Modal Impact, Vibration	090920
14	Subject	Cross, J. (Dabki)	Geontology		Thesis	Health, Health Behavior, Middle, Neighborhood, Socioeconomic Status, Subjective Social Status	090920
15	Predict	Bhattacharya, Nishita	Computer Science		Thesis	Hollywood, Internet Movie Database (IMDb), Movie genre, Oscar Award Nomination, Predictive model, Support Vector Machine	090920
17	Random	Assandri, Ivan	Mathematics, Applied		Thesis	Group representation, Homomorphism, Isotropy, Mutual symmetry, Random matrix, Schur's Lemma	090920

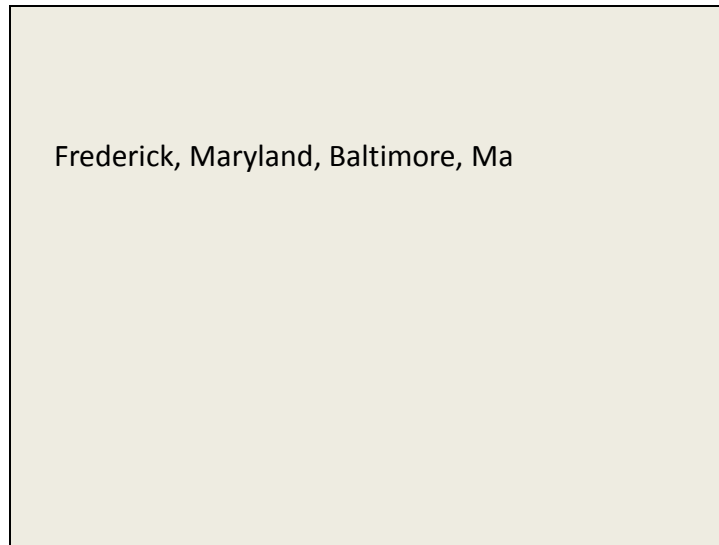
Formatted Data except Keywords

All data ends up in the proper form without additional staff intervention with the exception of the keywords

Keywords
Separate keywords by a semicolon
feeding behavior, induction of oligophagy, induction of preference, Manduca sexta, solanaceous
home- and community-based services, long-term care, Medicaid, nursing home
Beacon, Cognitive Process, Comments, Eye tracking, Program Comprehension
language learning, machine learning, nlp
Damage Detection, Elevator Testing, Engineering Education, Inverse Problems, Modal Impact Device, Vibration
Health, Health Behavior, Midlife, Neighborhood, Socioeconomic Status, Subjective Social Status
Hollywood, Internet Movie Database (IMDb), Movie script, Oscar Award Nomination, Predictive model, Support Vector Machine
Group representation, Homogenization, Isotropy, Material symmetry, Random media, Schur's Lemma

Keywords

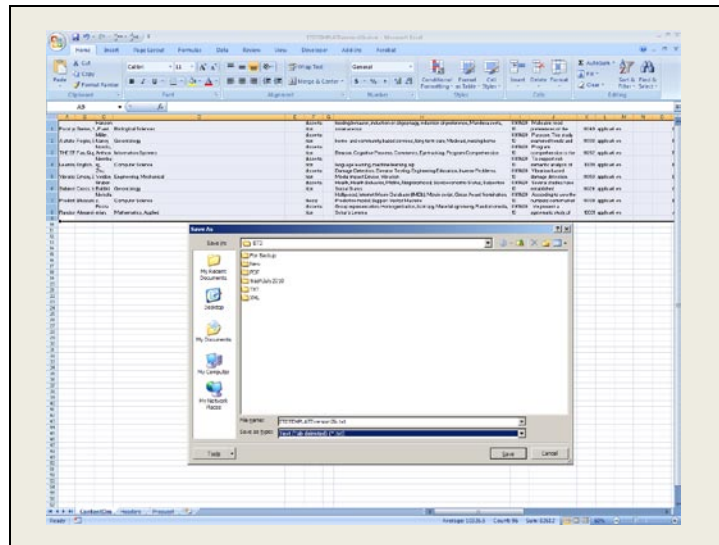
Which come from proquest separated by commas. These are user input, so sometimes the keywords include comas in them. For ContentDM, we need the keywords separated by semi-colons, but for the commas within a single keyword to remain commas. So we find an replace to do this, with staff checking each replace before doing it to insure that it's appropriate. If keywords include a value like



Keyword Examples

For example, if the keywords contain this, the staff member clearly leave the commas between Frederick and Maryland and Baltimore and Maryland, while changing the comma between Maryland and Baltimore to a semicolon to define these as two separate keywords.

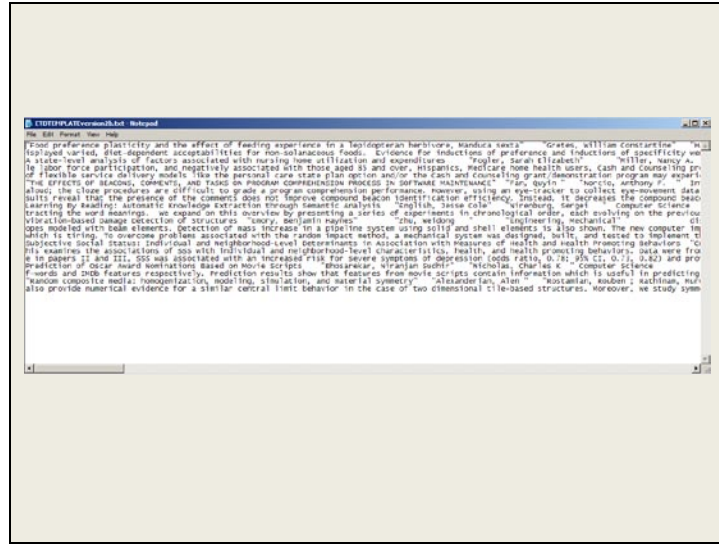
Slide 42



Delete Headers

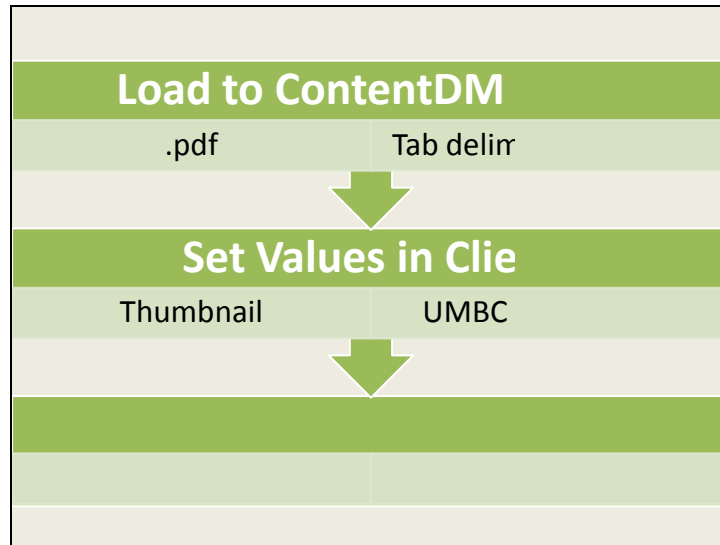
After the keywords all in order the colored headers are deleted and the file is saved a text-delimited file

Slide 43

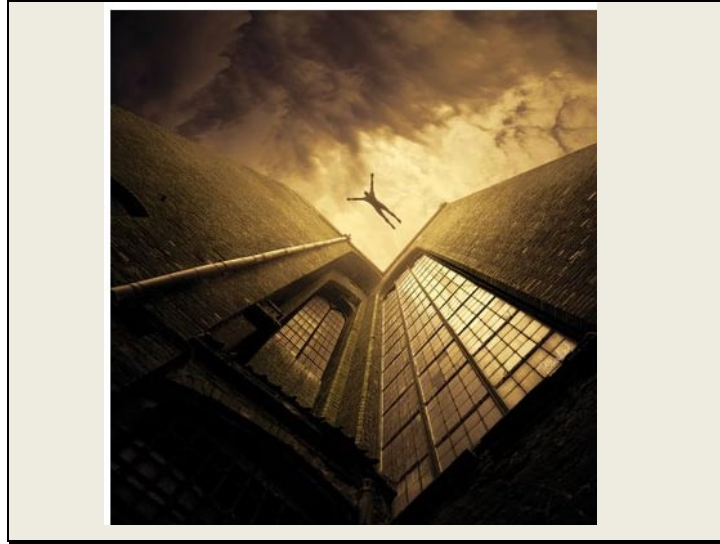


S-40—Text File

The text delimited file has to be opened and all quotes found and replaced with nothing, and any extra space deleted from the end. The file and the .pdf are then uploaded into ContentDM. Departments are added from a list, and the correct thumbnail is selected. They are then uploaded to the ContentDM server, approved, and indexed, but we have substantive difficulties with the server and the approval and indexing processes.



Then we load the .pdf's and tab delimited metadata into the contentDM client. We set 2 values in the client, the thumbnail and the UMBC department from an authoritative list. Finally, we load the files to the ContentDM serve and approve and index them.



What I've Learned

This project has been a substantial learning opportunity for me. I've gained a lot new Excel skills as well new skills in manipulating .pdf's. I've also learned how to use ContentDM. Additionally I found that a project may entail much than is apparent at first and that that when accepting work utilizing software that I'm unfamiliar with, it would be a good idea to investigate the quirks of that software and whether there will be sufficient support to troubleshoots problems.

