**Archived Documentation for the UF Digital Library Center (approx. 2007-2011):**

**Optical Character Recognition (OCR)  
Interactive & Performance Media Preservation   
Layered Media Preservation  
Basic Digital Curation /Data Curation  
Preparation of Project in lieu of Thesis  
Aerials (2009 LSTA grant workflow notes)  
Bulk Loading Internet Archive Files**

### Optical Character Recognition (OCR)

After digitally imaging text documents, and conducting the necessary processing for image quality and skew, the DLC uses Prime OCR for optical character recognition, and for image zoning if the target data is arranged in[columns or tables](http://digital.uflib.ufl.edu/technologies/documentation/zoning.htm). The OCR process recognizes the text within the images and creates plain-text files from the TIFF image files.

#### Named Entity Recognition (NER)

The textual content is scanned against different database tables, and hits are tagged accordingly: geographic names; personal names; and company names. For example, in this text exerpt, the program tagged place names (geograhical hits) with HTML to highlight them in red:

earl Hld WMOM

Crossing, for the site at which the Florida

Souwh was completed froalatka to Gaisesville in   
188 co e e1eninsular Railroad. families of Calvin

Waits and James Hawthorn. A diverse economy

stimulated the development of Hawthorne.

The addition of an "e" to the town's name

1. Wafts-Baker House - (606 Sid Martin

Highway, U.S. 301) The original log house

on this site, which was occupied in the

nineteenth century by Hawthorne founder

Calvin Waits and his family, had a sepa-

rate kitchen and dining room to minimize

the possibility of fire. Members of the

Baker family have lived in it since 1909

when farmer and rural mail carrier R. B.

Baker moved into the house with his new bride.

2. Hawthorne State Bank -

(2 N. Johnson

St.) This building housed Hawthorne's first

bank, which was organized in 1911 and

not long after advertised that it had assets

of $15,000. Francis J.

Hammond, leading

merchant and grandson of town founder

James M. Hawthorn, donated the land, and

A. L. Webb, proprietor of a successful gen-

eral store in Hawthorne, served as first presi-

dent of the institution.

3. Moore House - (207 W. Lake Ave.) This q

house, which was built in 191 1, still looks

much as it did not long after Glen D.

Moore purchased the house in 1913. A

sleeping porch, casement windows, and a

bathroom were added by Moore, whose

father, William Shepard Moore, had arrived

in

Hawthorne from Tennessee in 1882.

4. Mahin House - (301 W. Lake Ave.) The broad

porch on three sides of this tum-of-the-century

dwelling provided a perfect place for occu-

pants to sit and catch the breezes. Lottie

Mahin, widow of a businessman influen-

tial in the town during the second decade

of the twentieth century, lived in the house |

in the 1920s and rented part of it as apart-

ments.

5. State Historical Marker - (On the church

grounds, comer N. Johnson St. and N.W. 3rd

Ave.) A brief history of the town of ^

Hawthorne is provided.

6. First Baptist Church - (Comer N. Johnson

St. and N.W. 2nd Ave.) The First Baptist

Church in Hawthorne was organized in

1853. This building was erected in 1900.

Gothic style windows punctuate the white

horizontal clapboard siding that covers the

exterior.

**About & Running PrimeOCR™**

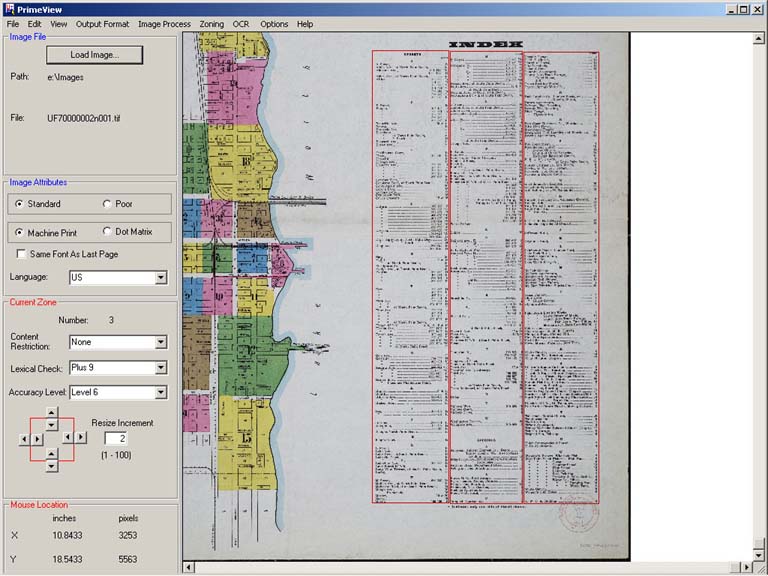
PrimeOCR features: six OCR engines (WordScan; TextBridge; Recore; TypeReader; OmniPage; FineReader); voting and weighing of the engine results; 11 Western languages; fault tolerant architecture, automatic engine recovery; image pre-processing; 1 to 6 CPU's; 16 output format formats including .PRO: metadata on location, confidence, etc., per character; and features for developers including an application programming interface (API) with 40 documented calls, dynamic link libraries (DLLs), and configurable initialization files (INIs).

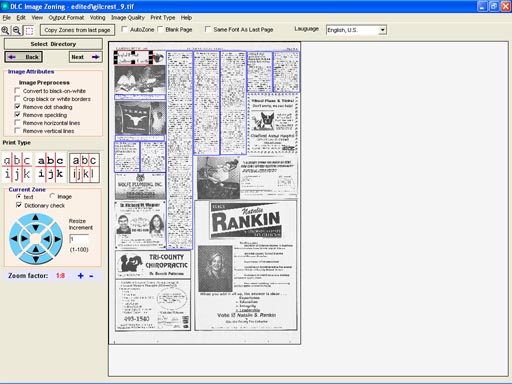
Job Server (Batch OCR; Prioritized jobs; Network aware; Job file)Prime Recognition Job FileVersion=3.90 1   
E:\Prdev\Images\UF70000002n001.tif   
E:\Prdev\Templates\UF70000002n001.ptm  
Document Template file:Prime Recognition Document Template   
Version=3.90   
0,-1 1,0,1,0,10,0,64,1,0,0   
4   
0,0,1,999999,0,0,0,0,   
1,0,1,999999,0,0,0,0,   
2,0,1,999999,0,0,0,0,   
3,0,1,999999,0,0,0,0,

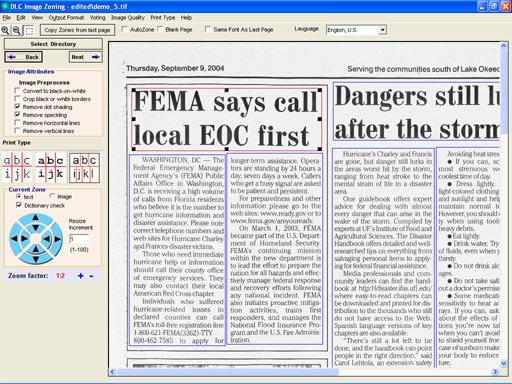
Running PrimeOCR™: Write a job file to the Job folderThe Job Server is a Prime Recognition™ program always running on the remote server reads the locations of the TIFF image and the document template from the job file. It processes the TIF file according to the template, and outputs selected file types: PDF+Image and TXT, leaving the original, archival TIF unchanged in its folder.  
Prime Recognition Job FileVersion=3.90 1  
\\Smathersnt19\ScanTech\ScanQC\Complete\UF00016972\00116.tif \\Smathersnt19\Prdev\Templates\UF70000002n001.ptm

**Image Zoning**

PrimeOCR is also bundled with PrimeView™ image zoning, which features a GUI to draw zones on the image:







# Image Higlighting

**DRAFT PAGE FOR NOTES TOWARDS IMAGE HIGHLIGHTING**

* [Example using a JPEG2000 source from the Aware server](http://128.227.223.117/imgserver/SobekCM_ImageServer.aspx?file=/UFDC/UF/08/09/08/71/00001/00005.jp2&w=512&h=512&z=5&r=0&ax=95&ay=95&at=FillRectangle&aw=50&ah=16&ac=blue)
* [Example with JPG Source](http://128.227.223.117/imgserver/SobekCM_ImageServer.aspx?file=http://ufdcweb1.uflib.ufl.edu/UFDC1/UF/08/09/08/71/00001/00005.jpg&w=512&h=512&z=5&r=0&ax=176&ay=177&at=FillRectangle&aw=94&ah=32&ac=blue)
* [Example with both the JPEG2000 ( with an incorrect name ) and a JPEG as the alternate.](http://128.227.223.117/imgserver/SobekCM_ImageServer.aspx?file=/UFDC/UF/08/09/08/71/00001/00005t.jp2&alt=http://ufdcweb1.uflib.ufl.edu/UFDC1/UF/08/09/08/71/00001/00005.jpg&w=512&h=512&z=5&r=0&ax=95&ay=95&at=FillRectangle&aw=50&ah=16&ac=blue)   
  Tested with a good JPEG2000 name, and with the Aware server down (disabled Apache) and the static JPEG file was subsituted. Still need auto-email sent to admin to alert to cycle Apache. Whenever it fails to the alternate name, the added feature is not included, since the coordinates will probably be wrong.
* Additional Examples
  + [Example 1](http://128.227.223.117/imgserver/?file=/UFDC/UF/01/02/07/76/00007/12001_1937_6_96.jp2&w=512&h=512&z=5&r=0&ax=100&ay=100&ah=40&aw=40&at=DrawEllipse&ac=Red)
  + [Example 2](http://128.227.223.117/imgserver/?file=/UFDC/UF/01/02/07/76/00007/12001_1937_6_96.jp2&w=512&h=512&z=5&r=0&ax=100&ay=200&ah=40&aw=40&at=DrawEllipse&ac=Red)
  + [Example 3 on scale](http://128.227.223.117/ufdc/ufdc.aspx?a=flap&m=ld8&i=76709&coord=30.297922,-81.577337,,)
  + [Example 4 on scale](http://128.227.223.117/ufdc/ufdc.aspx?a=flap&m=hd11&i=54054&coord=30.297922,-81.577337,,)
* [Zoomed-in example](http://128.227.223.117/imgserver/?file=/UFDC/UF/01/02/07/76/00007/12001_1937_6_96.jp2&w=512&h=512&z=3&x=0&y=0&r=0&ax=90&ay=115&ah=40&aw=220&at=DrawRectangle&ac=Yellow)

**Options**

* Possible types are FillRectangle, DrawRectangle, FillEllipse and DrawEllipse. Whenever fill is used, opacity is applied.
* Possible colors are blue, red, yellow, black, and white.
* URL query string parameters are as follows:
  + file = filename for the primary file ( either JP2 on Aware server or JPEG with full URL )  
    alt = alternate filename if the primary file fails.  
    w = viewport width ( for JPEG2000 )  
    h = viewport height ( for JPEG2000 )  
    r = rotation ( for JPEG2000 )  
    z = zoom ( for JPEG2000 ). Hightest number is furthest out.  
    x = x-location of viewport ( if zoom is not all the way out )  
    y = y-location of viewport ( if zoom is not all the way out )  
    ax = x-location of added feature  
    ay = y-location of added feature  
    at = type of added feature  
    aw = width of added feature  
    ah = height of added feature  
    ac = color of added feature

**Notes**

The imgserver will have to programmed to read the ALTO file for a page image ( with alto=xxx ) and perhaps the PRO file as well ( with pro=xxx ). And then to include the user's search as well.

**Interactive & Performance Media Preservation**

Standards for digital preservation of interactive and performance digital media are available through the Electronic Literature Organization (ELO):

* ["Acid-Free Bits: Recommendations for Long-Lasting Electronic Literature"](http://eliterature.org/pad/afb.html)
* ["Born-Again Bits: A Framework for Migrating Electronic Literature."](http://eliterature.org/pad/bab.html)

**Resources**

* Berkeley Art Museum, et. al. ["Archiving the Avant-Garde: Documenting and Preserving Digital / Variable Media Art"](http://www.bampfa.berkeley.edu/about/avantgarde)
* Collection of articles and perspectives on archiving methods for variable media, [*The Variable Media Approach: Permance through Change*](http://www.variablemedia.net/e/preserving/html/var_pub_index.html)
* [Digital Preservation at the Library of Congress](http://www.digitalpreservation.gov/)
* [Variable Media](http://variablemedia.net/)
* ["Toward an Organic Hypertext"](http://wordcircuits.com/connect/ht2000.htm) on Connection Muse, a system for creating standards based hypertext poetry and fiction

**Layered Media Preservation**

Standards for digital preservation of interactive and performance digital media are available from texts on preservation, and the following is adapted from Erich J. Kesse's "Strategies for Microfilming Scrapbooks and Layered Objects" in *RLG Archives Microfilming Manual.*

Scrapbooks, diaries, and unconventional book structures often contain pages with layered or dimensional objects. Such pages include parts which display one of the following characteristics:

* Overlapping parts (e.g., page layered with memorabilia).
* Enveloped parts (e.g., a diary containing letters in pasted-down envelopes).
* Folded or hinged parts (e.g., a page containing folded newspaper clippings).
* Movable or pop-up parts (e.g., an "artist's book" or a child's pop-up book).

Steps for Digitizing:

* First, film an image of the page as found, with all facets or openings on the page closed.
* Then, film each facet of each item in logical sequence for reading order (so this should vary as applicable): left to right and top to bottom.
* Within this order, overlaying items should be filmed before underlaying items, and envelopes should be filmed prior to filming their contents.
* Every facet of an item should be filmed before the filming of another item begins, and every item on a page should be filmed before the filming of another page begins.
* If a layered or dimensional item can be safely detached, remove it and film it separately from the page on which it is found. Otherwise, a blank sheet of ivory or black paper can be laid behind the individual item to mask it from the rest of the page.

**Resources**

* The [Jackson Henson MacDonald Scrapbook](http://ufdc.ufl.edu/UF00001625&v=00001) and the [Visionaires Scrapbook (1938-1948)](http://ufdc.ufl.edu/UF00024336&v=00001) are examples of basic scrapbooks - no movable parts, all images on pages, and no separation of the parts of the images on the pages instead having only the pages with images scanned.
* The [Bardin Scrapbook](http://ufdc.ufl.edu/UF00078910&v=00001) is a slightly more indepth example because each scrapbook page scanned, followed by high resolution scans of each of the individual images on the pages.
* "[Creating Online Historical Scrapbooks with a User-Friendly Interface](http://www.dlib.org/dlib/november07/zhang/11zhang.html)" article on using Flash page flip and embedded SWF files for scrapbooks
* Testing [OpenLaszlo for pop-ups](http://library.gameology.org/2008/02/05/grebo-mask-part-ii/) and other interactive books in DHTML and Flash

**Digital & Data Curation: Digital Curation Processing Workflow**

DLC

* runs virus scan
* reviews for file types;  normalizing any unusable formats and/or identifying needed software
* maintains 1 complete copy until copy-in-process is complete and archived
* sends files and any processing notes/information on software needed or recommended to SASC

SASC

* reviews for organization, maintaining any existing organization (folders, if intellectually grouped and significant contents), organizing other materials as appropriate based on underlying structure, and then placing other materials into a general folder(s)
* decides whether this accession is closed/1 set or if it can be added to/expanded with future donations (processing may change based on this)
* creates a finding guide/spreadsheet for all materials at the applicable level

DLC

* imports the spreadsheet
* renames all folders into BIB\_VID format
* normalizes all files and file names
* properly places all dark items for dark processing (note: multiple versions of same file are normally not kept for print collections unless literary collections because of space concerns, multiple versions will be retained for digital-only collections unless size/cost is an issue)

**Other Processing Notes**

In the past, SASC received only small portions of collections as digital and so would preserve materials by printing files. In the future, all digital files should go to DLC for digital preservation. The DLC will then also print the files for SASC as wanted by SASC (and this will likely be the case for any collections with only small portions that are digital, where having the material printed is needed for ease of use for the collection as a whole).

**Data Curation**

Around the world, key scientific data are at risk of being lost, either because they are held on fragile or obsolete media or because they may be destroyed by researchers who are unaware of their value. Now a team of scientists is planning to scour museums and research institutes to draw up a global inventory of threatened data. Launched on 29 October, shortly after the biennial conference of the Committee on Data for Science and Technology in Stellenbosch, South Africa, the project aims to publish the inventory online in 2012. (<http://www.nature.com/news/2010/101102/full/468017a.html>)

The digital files used in research are endangered and “at risk” data from all fields, including the humanities, arts, and social sciences. Modern research requires digital tools and the research process includes the selection, collection, organization, and analysis of digital files. Those digital files - spanning all types of content including scientific data, images of archival letters, interactive media files, and more - are all too often held in precarious ways. As such, the potential for total loss exists and this is in addition to the lost possibilities from the materials being inaccessible.

The UF Digital Collections include digital objects created through the digitization of source materials as well as born digital. All materials in the UF Digital Collections are curated to support optimal access and ensure long-term preservation. Please [contact us](mailto:ufdc@uflib.ufl.edu) for more information on the UF Digital Collections and services.

**Data Management**

NSF requires plans for data management to be included with grant proposals. See the full information from [NSF here](http://www.nsf.gov/bfa/dias/policy/dmp.jsp), which also links to the requirements by directorate, office, division, program, and other NSF units.

The NSF grant proposal preparation instructions ([online here](http://www.nsf.gov/pubs/policydocs/pappguide/nsf11001/gpg_2.jsp#IIC2j)) are copied below for ease of reference.

Plans for data management and sharing of the products of research. Proposals must include a supplementary document of no more than two pages labeled "Data Management Plan". This supplement should describe how the proposal will conform to NSF policy on the dissemination and sharing of research results (see [AAG Chapter VI.D.4](http://www.nsf.gov/pubs/policydocs/pappguide/nsf11001/aag_6.jsp#VID4)), and may include:   
  
1. the types of data, samples, physical collections, software, curriculum materials, and other materials to be produced in the course of the project;   
  
2. the standards to be used for data and metadata format and content (where existing standards are absent or deemed inadequate, this should be documented along with any proposed solutions or remedies);   
  
3. policies for access and sharing including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements;   
  
4. policies and provisions for re-use, re-distribution, and the production of derivatives; and   
  
5. plans for archiving data, samples, and other research products, and for preservation of access to them.   
  
Data management requirements and plans specific to the Directorate, Office, Division, Program, or other NSF unit, relevant to a proposal are available at: http://www.nsf.gov/bfa/dias/policy/dmp.jsp. If guidance specific to the program is not available, then the requirements established in this section apply.   
  
Simultaneously submitted collaborative proposals and proposals that include subawards are a single unified project and should include only one supplemental combined Data Management Plan, regardless of the number of non-lead collaborative proposals or subawards included. Fastlane will not permit submission of a proposal that is missing a Data Management Plan. Proposals for supplementary support to an existing award are not required to include a Data Management Plan.   
  
A valid Data Management Plan may include only the statement that no detailed plan is needed, as long as the statement is accompanied by a clear justification. Proposers who feel that the plan cannot fit within the supplement limit of two pages may use part of the 15-page Project Description for additional data management information. Proposers are advised that the Data Management Plan may not be used to circumvent the 15-page Project Description limitation. The Data Management Plan will be reviewed as an integral part of the proposal, coming under Intellectual Merit or Broader Impacts or both, as appropriate for the scientific community of relevance.

**Resources**

* [Born Digital Collections: An Inter-Institutional Model for Stewardship (AIMS)](http://www.digitalcurationexchange.org/?q=aims)
* [Paradigm workbook on digital private papers](http://www.paradigm.ac.uk/workbook/index.html)
* [Acid-Free Bits](http://www.eliterature.org/pad/afb.html)
* [NSF: Data Management for Engineering Directorate Proposals and Awards](http://nsf.gov/eng/general/ENG_DMP_Policy.pdf)
  + Investigators are expected to share with other researchers, at no more than incremental cost and within a reasonable time, the primary data … created or gathered in the course of work under NSF grants. ([Section VI.D.4.b.](http://www.nsf.gov/publications/pub_summ.jsp?ods_key=gpg))
  + What data are included? Research data are formally defined as “the recorded factual material commonly accepted in the scientific community as necessary to validate research findings” by the U.S. Office of Management and Budget (1999).
  + The basic level of digital data to be archived and made available includes (1) **analyzed data** and (2) the **metadata** that define how these data were generated. These are data that are or that should be published in theses, dissertations, refereed journal articles, supplemental data attachments for manuscripts, books and book chapters, and other print or electronic publication formats.
    - **Analyzed data** are (but are not restricted to) digital information that would be published, including digital images, published tables, and tables of the numbers used for making published graphs.
    - **Necessary metadata** are (but are not restricted to) descriptions or suitable citations of experiments, apparatuses, raw materials, computational codes, and computer-calculation input conditions.
  + **What data are not included at the basic level?**
    - The Office of Management and Budget statement (1999) specifies that this definition does not include “preliminary analyses, drafts of scientific papers, plans for future research, peer reviews, or communications with colleagues.” Raw data fall into this category as “preliminary analyses.” More stringent data management requirements may be specified in particular NSF solicitations or result from local policies and best practices at the PI’s home institution.
  + **Expected Data:**
    - The Data Management Plan (DMP) should describe the types of data, samples, physical collections, software, curriculum materials, and other materials to be produced in the course of the project. It should then describe the expected types of data to be retained.
  + **Period of data retention:**
    - The DMP should describe the period of data retention. Minimum data retention of research data is three years after conclusion of the award or three years after public release, whichever is later. Public release of data should be at the earliest reasonable time.
  + **Data formats and dissemination:**
    - The DMP should describe the specific data formats, media, and dissemination approaches that will be used to make data available to others, including any metadata. Policies for public access and sharing should be described, including provisions for appropriate protection of[privacy, confidentiality](http://www.nsf.gov/bfa/dias/policy/), security, intellectual property, or other rights or requirements.
  + **Data storage and preservation of access:**
    - The DMP should describe physical and cyber resources and facilities that will be used for the effective preservation and storage of research data. In collaborative proposals or proposals involving sub-awards, the lead PI is responsible for assuring data storage and access.

**Example of Aerial Photographs Pre-Curation**



# Preparation of Project in lieu of Thesis

The Digital Library Center accepts projects-in-lieu-of-theses for digitization and inclusion in the University of Florida Digital Collections (UFDC), within the [University of Florida Institutional Repository](http://ufdc.ufl.edu/?c=ufir). For materials to be processed into the University of Florida Digital Collections, all program requirements apply as well as several technical requirements listed below. Please note that non-project theses and dissertations go through the [Graduate Editorial Office](http://gradschool.ufl.edu/editorial/) and are submitted electronically through an alternate process.

Requirements for projects-in-lieu-of to be processed into the University of Florida Digital Collections:

1. Project-in-lieu-of must be turned into the graduate advisor for the department.
2. Project-in-lieu-of must be turned in with a signed [Internet Distribution Permissions form](http://www.uflib.ufl.edu/digital/procedures/copyright/GrantofPermissions.doc). All materials submitted to the University of Florida Digital Collections and the Institutional Repository are completely Open Access.
3. If copyrighted materials are included, the project-in-lieu-of must be turned in with a signed [letter of permission to quote or reproduce copyrighted material](http://digital.uflib.ufl.edu/technologies/documentation/SampleLetterofPermissiontoQuoteorReproduceCopyrighted%20Material.doc). Or it must include a list of materials for which copyright permissions were not granted so that the Digital Library Center can ensure that those materials are not included with the online files.
4. File formats should conform to the [acceptable ETD formats list](http://www.uflib.ufl.edu/committees/etd/policyandprocedures/accept_ETD_formats_0307.pdf).
5. Recommended: please turn in all files for the project on CD or DVD. If you are turning your project in on paper, please use non-glossy paper to make the scanning more accurate if possible. For paper-only submissions, please follow the binding guidelines ([architecture](http://digital.uflib.ufl.edu/technologies/documentation/Arch-MRP-guidelines.doc) and [music](http://digital.uflib.ufl.edu/technologies/documentation/Musicspecific.doc)).

**Printable guide:** [draft one page handout](http://digital.uflib.ufl.edu/technologies/documentation/PILO.doc)  
**Example:** "[The End and How it Looks](http://ufdc.ufl.edu/uf00090020&v=00001)" and "[Film+Architecture](http://ufdc.ufl.edu/?c=ufir&b=UF00090056)" are examples of "project-in-lieu-of a thesis" digitized by the Digital Library Center.  
**Questions:**Please contact your departmental advisor or the [Digital Library Center](http://digital.uflib.ufl.edu/technologies/documentation/dlc@uflib.ufl.edu) with any questions.

## ****Processing**** of Project in lieu of Thesis

* The IR Coordinator liaises with advisors and acquires digital files from departmental advisor when theses approved along with copyright form.
* The IR Coordinator uses the metadata provided by advisor to create a brief record and then processes the electronic files into one digital package and loads the files onto UFDC.
  + Any projects without the copyright clearance form go through the printed and binding process.
  + For any projects with materials that cannot be digitized for issues related to copyright or materiality, the IR Coordinator will turn those files over to Preservation with all existing metadata for processing following the [normal binding procedures](http://www.uflib.ufl.edu/preserve/binding/thesis/studentinfo.html).
* IR Coordinator notifies CatMet that the item with brief metadata is ready for full cataloging using the DLC metadata editor or using the MARCXML generated by UFDC (whichever is easier--the DLC can work with either for our workflows).
* Abstracts aren’t included in the PILO catalog records. For standard ETDs, the abstract text comes to cataloging in the metadata from FCLA. Thus, the cataloging process for PILO catalog records will include harvesting and tidying the OCRed text from the Greenstone metadata.

# Aerials

This information is for the Florida Aerials (LSTA grant, phase III, 2009).

**Processing Stages**

[FIPS codes for Florida](http://ufdc.ufl.edu/?c=flap&m=hbbindex)

* Map Library will select and prepare 1,000 photographs per week.
  + Map Library will clean all marks from the photographs.
  + Map Library will create a check in sheet for the batches with an inventory for the 1,000 tiles.
  + Map Library will send the check in sheet and inventory list to the DLC electronically and print copies that will accompany the physical materials.
  + [Sample check in sheet and inventory](http://digital.uflib.ufl.edu/technologies/documentation/Box2Alternate.xlsx)
* DLC picks up, digitizes, and returns the photos.
  + DLC will pick up each batch of the 1,000 tiles and accompanying documentation from the Map Library.
  + When picking up new batches, DLC will return old batches, only if entire batch is complete. DLC will continue to pick up new batches on a weekly schedule even if old batches have not been returned.
  + DLC checks batch for all flights and items.
  + Carrie assigns imaging by flight to students.
  + Students scan items to local external hard drives, using the naming convention explained below.
  + Carrie or Traveler move images remotely to another drive.
  + Note: PreQC on the aerials must be run using Adobe Photoshop for the derivative creation. For the aerial BIBIDs, PreQC includes a standard levels adjustment.
  + Quality control is done by Jane or a student she assigns to it, and quality control is for image quality and correct tile numbering schema.
  + Aerials are loaded (no OCR) and archived.
  + DLC checks loaded and archived batches.
  + DLC prepares batches for return to the Map Library, ensuring batches are complete, photos are in numerical order, and photos face the same direction.

**Saving Aerial Images**

All aerials will use the already established BIBID for the county. Each image will be named in this format: FIPScountyCode\_YearOnThePhoto\_FlightNumber\_tilenumber.tif

**Example:**12001\_1937\_1F\_2.tif

Example is for Alachua County (FIPS 12001), year of the flight on the photo is 1937, flight number is 1F, and the tile number is 2

A printable sign with naming conventions for students is [here.](http://digital.uflib.ufl.edu/technologies/documentation/SavingAerials.doc) The Map Library page on Aerial Photography coverage is [here](http://www.uflib.ufl.edu/maps/Aerials/MAPNEWAERIAL.HTML).

**Bulk Loading Internet Archive Files**

Script to grab [Internet Archive files](http://code.google.com/p/bhl-bits/source/browse/trunk/utilities/grabby/grabby.sh?r=33)   
Remove LZW with ImageMagick: mogrify –compress none \*.tif   
Convert all JP2s in a dir to TIF with ImageMagick, using [mogrify](http://www.imagemagick.org/Usage/basics/#mogrify): mogrify -format tiff \*.jp2