

# Data Accessioner

*Developing a Tool for  
Mediating Migration  
of File-based Electronic  
Records*

DUKE UNIVERSITY  
  
LIBRARIES

University Archives

Beat things with *rocks*...



... until they *work*.



In the beginning...



## Duke University Rare Book, Manuscript, and Special Collections Library

### Digital Media/ Electronic Records Action Guidelines

The following guidelines have been drafted to provide guidance to Duke University Rare Book, Manuscript, and Special Collections Library staff on the workflow procedures for handling digital media items (such as word processor and computer diskettes, tapes, CD-ROMs, etc.) found in manuscript collections acquired by the library. The procedures described are intended to facilitate the prompt action required for preserving the information carried on short-lived magnetic and optical media, and present the various options for doing so. The guidelines are based on contributions from members of the several RBMSCL departments, and their successful implementation requires similar ongoing cooperation. While each section title lists the staff members most directly involved with the subsequent series of operations, the various steps of the described process may require a collaborative cross-departmental effort.

- 1.0 Acquisition: Collection Development Staff
  - 1.1 Discover Information About Digital Media Items Within Collection
  - 1.2 Record Information About Digital Media Items Within Collection in RBMSCL Database
- 2.0 Accessioning: Technical Services Staff Accessioner/ Processing Archivist
  - 2.1 Find Digital Media Items Within Collection and Remove
  - 2.2 Record Information about Removed Digital Media Items in Container List and RBMSCL Database
  - 2.3 Migrate information from original physical item to Special Collections server.
  - 2.4 Create a Disk Directories log.
  - 2.5 Determine if the number of disks and files warrants further processing.
  - 2.6 Change access permissions.
- 3.0 Accessioning: Reformat word processing digital items into Rich Text and Text formats and create a Digital Files Locator Form
  - 3.1 Reformat the digital items into Rich Text and Text formats.
  - 3.2 Create a Digital Files Locator Form.
  - 3.3 Change access permissions
- 4.0 Accessioning: Disposition of Original Digital Media Items: Technical Services Staff Accessioner/ Processing Archivist
  - 4.1 Discard or Retain Original Digital Media Items
  - 4.2 Relocate, Indicate Location of, and Impose Restrictions on Retained Items
- 5.0 Full Processing: Create a Digital Files Locator Form and reformat word processing digital items into Rich Text and Text formats
  - 5.1 Create a Digital Files Locator Form.
  - 5.2 Change access permissions.
  - 5.3 Reformat the word processing digital items into Rich Text and Text formats.
  - 5.4 Change access permissions.
- 6.0 Processing: Providing User Access Points to Migrated Digital Files: Technical Services Staff Processing Archivist
  - 6.1 Place Separation Sheets in Appropriate Locations During Processing
  - 6.2 Note the Presence of Digital Files in Finding Aid
  - 6.3 Encode the finding aid.

### Directions for migrating digital media items from 3.5" floppies and CDs to the Special Collections Server

1. Open a My Computer window.
2. Navigate to S:\Accessioned.
3. Create a new folder using the last name, first name of the accession and the accession number.
4. Within that folder, create new sub-folders for each disk in the accession (Disk 001, Disk 002, Disk 003, etc.)
5. Open a second My Computer window.
6. Put Disk 001 in the a:\ or e:\ drive.
7. In the second My Computer window, navigate to the disk in the a:\ or e:\ drive.
8. Copy the contents from the a:\ or e:\ drive to the Disk 001 folder in the first My Computer window.
9. Repeat steps 6-8 for all remaining disks. Use View → Refresh to change the contents of the My Computer window when you've inserted a new disk into the a:\ or e:\ drive.
10. Any MacIntosh format, 5" floppy disks, zip disks, or other formats will require different procedures.

### Directions for creating a disk directories log for any newly migrated digital media accession

*Copying the disk directory into a text document:*

1. In a My Computer window, navigate to S:\Logs.
2. Create a new folder using the last name, first name of the accession and the accession number. The name should be the same as the folder you created in S:\Accessioned.
3. Open a command window.
4. At the c: prompt, use these commands to navigate to the disk directory in S:\Accessioned you wish to log:

```
s:
cd accessioned
cd [name of accession]
```

5. Once you have reached the accession you want to log, use this command to copy the entire directory (the contents of the accession) to a newly created document.

```
dir /s/s:\logs\[last name of accession.txt]
example: dir /s/s:\logs\shor.txt
[the ">" means "copy to" and then you define the file path]
```

If you need to, you may also append or edit the document you've just created by using this command:

```
dir [of whatever disk]>>s:\Logs\[name of accession]\[last name of accession.txt]
[the ">>" means "append output to" and then you define the file path]
```

6. Other commands you might need are:
  - a. `cd ..` takes you one level higher in the directory hierarchy

Ruth Bryan, 7/1/01; 2/4/02; 3/13/02. Updated by jk 2/9/07.

# Paper based process ...

# SEPARATION SHEET

for

## Digital Media Items (Diskettes, CD-ROMs, etc.)

Item Number: \_\_\_\_\_  
(Example: "Disk 003")

Item Name: \_\_\_\_\_  
(From Item Label)

COLLECTION  
Main Entry \_\_\_\_\_  
Accession Number \_\_\_\_\_

DIGITAL ITEMS: copy and paste this section for each additional digital item

ITEM ATTRIBUTES  
New Item Name and Location Path (<specoll/Archived Disks/COLLECTION FOLDER/ ITEM FOLDER>)  
~~specoll/Archived Disks/~~ \_\_\_\_\_  
Item Label (if any) \_\_\_\_\_  
Hardware Type (PC, Macintosh, Stand-alone Word Processor [i.e., Brother], etc.) \_\_\_\_\_  
Media Type (3.5" Diskette, 5.25" Floppy Disk, CD-ROM, etc.) \_\_\_\_\_  
Access Restrictions (Indicate "See Individual File" if only on particular files) \_\_\_\_\_

Physical Item Retained? No \_\_\_\_\_ Yes \_\_\_\_\_ Location: \_\_\_\_\_  
Note: If retained, the original digital media item is closed to use by researchers.

ITEM CONTENT  
General Description (Correspondence, Email, Drafts of . . . , etc.) \_\_\_\_\_

FILES  
File Name \_\_\_\_\_  
File Content Description if not Evident from File Name \_\_\_\_\_  
Original File Format (WP program, etc.) \_\_\_\_\_  
New File Formats: (X all conversion methods that apply): .txt \_\_\_\_\_ .rtf \_\_\_\_\_  
Original File Date/ Data Migration Date \_\_\_\_\_ / \_\_\_\_\_  
Access Attempt Notes (Problems, etc.) \_\_\_\_\_

File Name \_\_\_\_\_  
File Content Description if not Evident from File Name \_\_\_\_\_  
Original File Format (WP program, etc.) \_\_\_\_\_  
New File Formats: (X all conversion methods that apply): .txt \_\_\_\_\_ .rtf \_\_\_\_\_  
Original File Date/ Data Migration Date \_\_\_\_\_ / \_\_\_\_\_  
Access Attempt Notes (Problems, etc.) \_\_\_\_\_

File Name \_\_\_\_\_  
File Content Description if not Evident from File Name \_\_\_\_\_  
Original File Format (WP program, etc.) \_\_\_\_\_  
New File Formats: (X all conversion methods that apply): .txt \_\_\_\_\_ .rtf \_\_\_\_\_  
Original File Date/ Data Migration Date \_\_\_\_\_ / \_\_\_\_\_  
Access Attempt Notes (Problems, etc.) \_\_\_\_\_

File Name \_\_\_\_\_  
File Content Description if not Evident from File Name \_\_\_\_\_  
Original File Format (WP program, etc.) \_\_\_\_\_  
New File Formats: (X all conversion methods that apply): .txt \_\_\_\_\_ .rtf \_\_\_\_\_  
Original File Date/ Data Migration Date \_\_\_\_\_ / \_\_\_\_\_  
Access Attempt Notes (Problems, etc.) \_\_\_\_\_

with paper forms ...

Bogart, Leo. Leo Bogart papers  
2006-0087

NOTE: some diskettes were not formatted. Disk 042 contained a virus and was not migrated.

Volume in drive S is DMM  
Volume Serial Number is 9803-EEEE6

Directory of S:\Accessioned\Bogart, Leo 2006-0087

```
04/24/2007 03:20 PM <DIR>      .
04/24/2007 03:20 PM <DIR>      ..
04/24/2007 03:20 PM <DIR>      Disk 001
04/24/2007 03:20 PM <DIR>      Disk 002
04/24/2007 03:20 PM <DIR>      Disk 003
04/24/2007 03:20 PM <DIR>      Disk 004
04/24/2007 03:20 PM <DIR>      Disk 005
04/24/2007 03:20 PM <DIR>      Disk 006
04/24/2007 03:20 PM <DIR>      Disk 007
04/24/2007 03:20 PM <DIR>      Disk 008
04/24/2007 03:20 PM <DIR>      Disk 009
04/24/2007 03:20 PM <DIR>      Disk 010
04/24/2007 03:20 PM <DIR>      Disk 011
04/24/2007 03:20 PM <DIR>      Disk 012
04/24/2007 03:20 PM <DIR>      Disk 013
04/24/2007 03:20 PM <DIR>      Disk 014
04/24/2007 03:20 PM <DIR>      Disk 015
04/24/2007 03:20 PM <DIR>      Disk 016
04/24/2007 03:20 PM <DIR>      Disk 017
04/24/2007 03:20 PM <DIR>      Disk 018
04/24/2007 03:20 PM <DIR>      Disk 019
04/24/2007 03:20 PM <DIR>      Disk 020
04/24/2007 03:20 PM <DIR>      Disk 021
04/24/2007 03:20 PM <DIR>      Disk 022
04/24/2007 03:20 PM <DIR>      Disk 023
04/24/2007 03:23 PM <DIR>      Disk 024
04/24/2007 03:23 PM <DIR>      Disk 025
04/24/2007 03:23 PM <DIR>      Disk 026
04/24/2007 03:23 PM <DIR>      Disk 027
04/24/2007 03:23 PM <DIR>      Disk 028
04/24/2007 03:23 PM <DIR>      Disk 029
04/24/2007 03:23 PM <DIR>      Disk 030
04/24/2007 03:23 PM <DIR>      Disk 031
04/24/2007 03:23 PM <DIR>      Disk 032
04/24/2007 03:23 PM <DIR>      Disk 033
04/24/2007 03:23 PM <DIR>      Disk 034
04/24/2007 03:23 PM <DIR>      Disk 035
04/24/2007 03:23 PM <DIR>      Disk 036
04/24/2007 03:23 PM <DIR>      Disk 037
04/24/2007 03:23 PM <DIR>      Disk 038
04/24/2007 03:23 PM <DIR>      Disk 039
04/24/2007 03:23 PM <DIR>      Disk 040
04/24/2007 03:23 PM <DIR>      Disk 041
04/24/2007 03:23 PM <DIR>      Disk 042
04/24/2007 03:23 PM <DIR>      Disk 043
04/24/2007 03:23 PM <DIR>      Disk 044
04/24/2007 04:59 PM <DIR>      Disk 045
04/24/2007 04:59 PM <DIR>      Disk 046
04/24/2007 04:59 PM <DIR>      Disk 047
04/24/2007 04:59 PM <DIR>      Disk 048
04/24/2007 04:59 PM <DIR>      Disk 049
04/24/2007 04:59 PM <DIR>      Disk 050
04/24/2007 04:59 PM <DIR>      Disk 051
04/24/2007 04:59 PM <DIR>      Disk 052
04/24/2007 04:59 PM <DIR>      Disk 053
04/24/2007 04:59 PM <DIR>      Disk 054
04/24/2007 04:59 PM <DIR>      Disk 055
04/24/2007 04:59 PM <DIR>      Disk 056
04/24/2007 04:59 PM <DIR>      Disk 057
04/24/2007 04:59 PM <DIR>      Disk 058
04/24/2007 04:59 PM <DIR>      Disk 059
04/25/2007 08:55 AM <DIR>      Disk 060
04/25/2007 08:55 AM <DIR>      Disk 061
```

Directory of S:\Accessioned\Bogart, Leo 2006-0087\Disk 001

Item Label:  
infohwy  
premises  
preface (new 5/94)  
intervus  
foreign  
strategy 2/27/94  
strat  
bogota  
sfstate  
letters  
presstim  
10/28/93  
4/11/94  
Hardware Type: PC  
3.5" diskette  
Document Type: Microsoft word documents

```
04/24/2007 03:20 PM <DIR>      .
04/24/2007 03:20 PM <DIR>      ..
04/03/1994 06:17 PM          56,479 BOGOTA.DOC
01/22/1994 12:59 PM          73,459 FOREIGN.DOC
11/29/1993 03:00 PM          38,098 INFOHWY.DOC
12/09/1993 05:59 PM          66,037 INTERVUS.DOC
05/23/1994 02:20 PM          171,648 LETTERS.DOC
12/27/1993 03:16 PM          146,617 PREFACE.DOC
10/28/1993 03:32 PM          63,498 PREMISES.DOC
04/28/1994 12:08 PM          27,399 PRESSTIM.DOC
04/21/1994 02:25 PM          41,018 SFSTATE.DOC
02/28/1994 10:49 AM           8,071 STRAT.DOC
03/02/1994 04:16 PM          132,402 STRATEGY.DOC
11 File(s)                   824,726 bytes
```

Directory of S:\Accessioned\Bogart, Leo 2006-0087\Disk 002

Item Label: Commentaries  
Hardware Type: PC  
3.5" Diskette  
Document Type: Microsoft word document

```
04/24/2007 03:20 PM <DIR>      .
04/24/2007 03:20 PM <DIR>      ..
10/22/2004 11:42 AM          987,136 commentaries3.doc
1 File(s)                   987,136 bytes
```

Directory of S:\Accessioned\Bogart, Leo 2006-0087\Disk 003

NOTE: Diskette not formatted

Item Label: contact  
fictious  
intervus  
intro  
otto  
research  
residue  
times  
(several illegible)  
Hardware Type: PC  
3.5" Diskette  
Document Type: Microsoft word documents

```
04/24/2007 03:20 PM <DIR>      .
04/24/2007 03:20 PM <DIR>      ..
04/24/2007 03:29 PM          18 DISK NOT FORMATTED.txt
1 File(s)                   18 bytes
```

Directory of S:\Accessioned\Bogart, Leo 2006-0087\Disk 004

and DOS directory listings.

# Goals

- Migrate disks to \_\_\_\_\_
- Capture metadata
- Integrate 3<sup>rd</sup> party tools  
(adaptability/extensibility)
- Simple to use



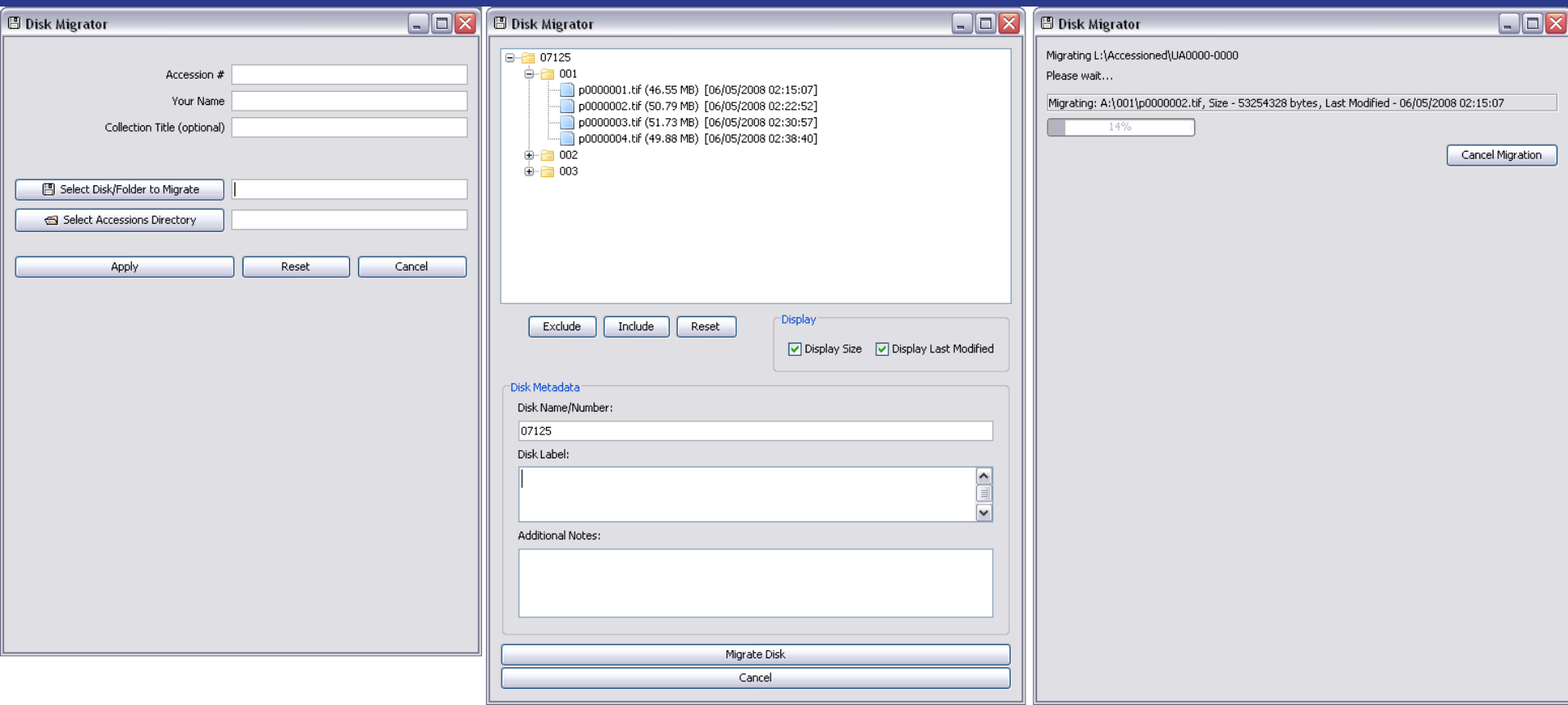
# Environment Scan

A green-tinted image of a microscope on a circular platform with a grid and glowing lines, symbolizing environmental scanning.

- Journals
  - Conference Proceedings
  - Project Websites

# Round One ~2 weeks

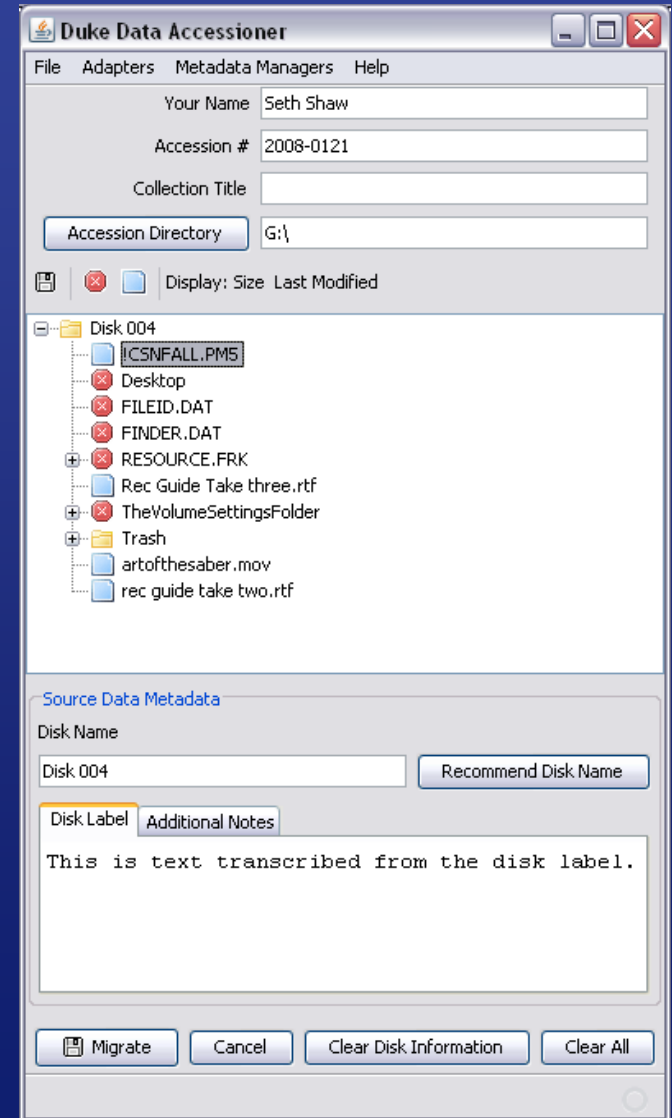
- Initial design
- Proof of concept
- Initial testing
- Local deployment



# Round Two

- Bug fixes
- Integrated interface
- Configurability
- Misc. small changes

Sporadic work over 6 months  
as time permitted



# Sharing...

- Informal Communication
- Demos
- Partner Sites
- Website
- Workshop



## ... spurs change

- Additional bug fixes
- Formal plug-in architecture
- Better documentation
- Improved configurability
- More ideas



# Outcomes

- It works
- People use it

