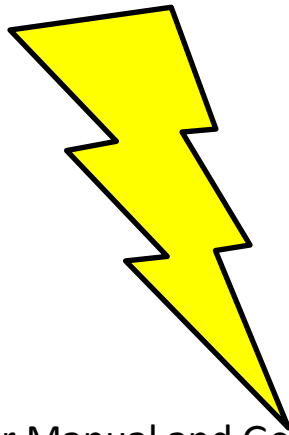


BENTLEY  
HISTORICAL  
LIBRARY

## PROCESSING HYBRID AND DIGITAL COLLECTIONS



An AutoPro User Manual and General Reference  
Version 2.0

Prepared by Michael Shallcross, Division of Digital Curation  
Reviewed by Nancy Deromedi, Head, Division of Digital Curation

1. Introduction .....	1
1.1 Institutional Context .....	1
1.2 Digital Processing as a Concept and Approach at the Bentley Library .....	1
1.3 Overview of the Automated Processing Workflow .....	1
1.4 Notes on the Windows Command Prompt .....	4
2. Initiate a Processing Session .....	6
2.1 Launching AutoPro .....	6
2.2 Logging In .....	7
2.2.1 Directory confirmation.....	7
2.2.2 Identification of Processing Archivist.....	8
3. Main Menu—Selecting Processing Options.....	9
4. Guide to Individual Procedures.....	11
4.1 Virus Scan.....	11
4.2 Backup Content.....	11
4.3 File Extraction.....	12
4.3.1 Disk Image Extraction.....	12
4.3.2 Archive File Extraction .....	15
4.4 Identify Missing File Extensions .....	16
4.5 Format Conversion.....	18
4.6 Personally Identifiable Information (PII) Scan .....	19
4.7 Appraisal, Arrangement, and Packaging for Deposit.....	21
4.7.1 In-Depth Review of Content .....	22
4.7.1.1 The “Characterize Deposit” Menu .....	22
4.7.1.1.1 Directory Hierarchy .....	22
4.7.1.1.2 Relative size of directories .....	23
4.7.1.1.3 File Extensions.....	24
4.7.1.1.4 Age of Files .....	24
4.7.1.1.5 Duplicate content.....	25
4.7.1.2 The “Review Content” Menu .....	25
4.7.1.2.1 Quick View Plus:.....	25
4.7.1.2.2 IrfanView:.....	26
4.7.1.2.3 Inkscape: .....	27

4.7.1.2.4 VLC Media Player: .....	27
4.7.2 Separations .....	28
4.7.2.1 AutoPro-separations-onlySelectedItems.bat.....	28
4.7.2.2 AutoPro-separations-extensionRemoval.bat.....	29
4.7.3 Arrangement .....	30
4.7.4 Packaging Content for Deposit and Defining Deep Blue ‘Items’ .....	30
4.7.4.1 Overview of Deep Blue .....	30
4.7.4.2 Determining How to Package Content .....	32
4.7.4.3 Example Packaging and Deposit Plan .....	33
4.7.4.4 Packaging Content with <i>AutoPro-zipOneFolder.bat</i> .....	34
4.7.4.5 Packaging Content with <i>AutoPro-zipMultipleFilesFolders.bat</i> .....	35
4.7.5 Completing the Appraisal and Arrangement Procedures .....	36
4.8 Extract Technical Metadata .....	37
4.9 Add Descriptive and Administrative Metadata.....	38
4.9.1 Make Sure that Intellectual Arrangement and Deposit Plan are Established .....	39
4.9.2 Add Deposit-Level Metadata .....	39
4.9.3 Check for Pre-Completed Metadata Spreadsheet .....	40
4.9.4 Create a New Item .....	42
4.9.5 Add Item-Level Metadata .....	43
4.9.6 Add Content to the Item .....	45
4.9.7 Enter Path to File Associated with Item.....	45
4.9.8 Describing the File.....	46
4.9.9 Original Creation or Use Dates.....	46
4.9.10 Other Information.....	47
4.9.11 Complete the Packaging and Description of an Item .....	47
4.9.12 Review an Item .....	48
4.9.13 Finalize the Deposit.....	50
4.10 Transfer to Long-Term Storage .....	50
4.10.1 Transfer to Deep Blue Deposit Folder.....	50
4.10.2 Transfer to Dark Archives.....	51
4.11 Clean Up.....	52
5. Version History.....	54

This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0/>.



# **1. Introduction**

## **1.1 Institutional Context**

Established in 1935 by the University of Michigan Regents, the Bentley Historical Library serves as the official archives of the university and documents the history of the state of Michigan and the activities of its people, organizations and voluntary associations. The library has successfully managed and preserved digital content since the 1997 accession of former University President James J. Duderstadt's digital desktop. Given the steep increase in born digital and digitized content accessioned by the library in recent years, archivists have sought more efficient and standardized processing procedures. The Andrew W. Mellon Foundation-funded MeMail Project (2010-2011) provided the library with resources to establish a workflow and corresponding policies for the ingest and processing of archival email, but a similar solution was needed for mixed digital content (i.e. Office documents, PDFs, audio and video files, images, etc.). Archivists in the library's Digital Curation Division have advanced the work of the MeMail Project in developing the AutomatedProcessor (or AutoPro), a series of inter-dependent scripts that automates key steps in preparing digital content for long-term preservation and access.

## **1.2 Digital Processing as a Concept and Approach at the Bentley Library**

Archival processing in the digital era requires traditional steps such as appraisal, arrangement, and description in addition to procedures that ensure the authenticity, integrity, and security of content. "Digital processing" therefore corresponds to the "generate AIP" function of the Open Archival Information System (OAIS) Reference Model's Ingest entity. After a Submission Information Package (SIP) has been assigned an accession record, digital processing permits archivists to assume intellectual control, establish the integrity of materials, and perform preservation events (i.e. scans for viruses and personally identifiable information, conversion to preservation formats, recording of descriptive and technical metadata, etc.) that transform the SIP into an Archival Information Package (AIP).

Bentley archivists initially developed a manual workflow with more than 40 discrete steps that required the operation of numerous stand-alone applications and saving tool output to various log files. In addition to being highly labor intensive and introducing numerous opportunities for operator error, this approach was daunting for staff without technical expertise. Given these challenges, the Division of Digital Curation developed AutoPro to fulfill two goals: (1) to make digital processing more efficient by automating key workflow steps and (2) to reduce technical barriers and thereby permit archivists to focus their energies on the traditional archival functions of appraisal, arrangement, and description.

## **1.3 Overview of the Automated Processing Workflow**

AutoPro is comprised of 28 Windows CMD.EXE shell scripts that move content through an 11 step workflow and thereby simplify the operation of more than 20 applications and command line utilities. The Windows Command Prompt and Explorer windows function as the main interfaces, a feature that may be unique to staff more familiar with Graphical User Interfaces (see section 1.4 for tips on using the CMD.EXE console). In addition to providing a framework to guide archivists through the workflow, AutoPro tracks the current processing status, generates log files for all operations, and records PREMIS

preservation metadata that will be stored alongside the processed content in a preservation environment. Archivists must approve the successful completion of each step and may stop at any point in the workflow and resume their work at a later time. A brief overview of the workflow (with information on component applications) follows below.

Immediately after content is accessioned and deposited in the Bentley Library's interim repository (a secure Windows file server), an archivist starts AutoPro to run a virus scan (the University of Michigan employs Microsoft Forefront Endpoint Protection on all work stations) and create a working backup so the SIP can be restored in case of an error or accidental data loss.<sup>1</sup>

For the next step in the workflow, AutoPro searches for disk image and archive files (.ISO, .AFF, .ZIP, .TAR, .RAR, etc.); if any are found, a script employs 7-Zip to extract the contents to a directory named after the archive file, with the original file paths preserved.<sup>2</sup> After verifying the extraction's success, AutoPro moves the archive file to a separations directory and records the operations in a log file. The newly extracted content is then searched for additional archive files, from which the contents are extracted, if necessary. Please note that some disk image formats (.AFF, .AD1, .E01, etc.) will require the processor to employ FTK Imager to manually extract content.<sup>3</sup>

AutoPro then runs DROID to search for files with missing or mismatched extensions and the archivist may then append correct file extensions using information generated by the TrID File Identifier utility and collected from the PRONOM format registry.<sup>4</sup> If the archivist determines that an extension should be added or corrected, AutoPro will document the action in a log file.

In transforming the SIP to an AIP, the Bentley Library relies upon file format conversion as a primary preservation strategy. Based upon the Library of Congress's work on the "Sustainability of Digital Formats" and documentation from the Florida Center for Library Automation and other peer institutions, the library has identified a number of at-risk (i.e. proprietary or potentially obsolete) file formats and developed conversion pathways to sustainable formats with various open source and freeware tools.<sup>5</sup> AutoPro searches for these at risk formats (based upon extension) and then employs the following tools (with digital media and target format in parentheses): ImageMagick (raster images to .TIFF), Ghostscript (.PS, .EPS and .PDF to .PDF/A; JHOVE verifies if the original PDF meets PDF/A

---

<sup>1</sup> See UM ITS FAQ pages at <http://safecomputing.umich.edu/antivirus/faq.php>. Microsoft antivirus information may be found at: <http://www.microsoft.com/en-us/server-cloud/system-center/endpoint-protection-2012.aspx>.

<sup>2</sup> 7-Zip is an open source file archiving application. For more information see <http://www.7-zip.org/>.

<sup>3</sup> FTK Imager is digital forensics software produced and freely distributed by AccessData. For more information, see the FTK Imager manual release notes and manual at <http://www.accessdata.com/support/product-downloads>.

<sup>4</sup> DROID is a file identification tool developed by the National Archives (U.K.). For more information, see <http://droid.sourceforge.net/>. TrID is a freely distributed utility that identifies file types based upon a library of over 4,800 binary signatures. For more information, see <http://mark0.net/soft-trid-e.html>. PRONOM is an "on-line information system about data file formats." For more information, see <http://www.nationalarchives.gov.uk/help/PRONOM/faq.htm>.

<sup>5</sup> For more information on the Library of Congress's "Sustainability of Digital Formats" and the FCLA's format recommendations, see <http://www.digitalpreservation.gov/formats/index.shtml> and [http://fclaweb.fcla.edu/fda\\_format\\_landing\\_page](http://fclaweb.fcla.edu/fda_format_landing_page), respectively. For an overview of sustainable formats and conversion strategies at the Bentley Historical Library, see <http://deepblue.lib.umich.edu/handle/2027.42/93307>.

specifications), Inkscape (vector images to .SVG), ffmpeg (audio to .WAV; video to MP4 with H.264 encoding), Aid4Mail (various email formats to .MBOX), and Microsoft Office File Converter (Office files to Open Office XML).<sup>6</sup> These preservation versions are stored alongside the original and denoted by a suffix consisting of ‘\_bhl-’ and (where possible) the CRC32 hash of the original file (i.e. oralHistoryProject\_bhl-0fbc2cc7.wav). AutoPro also creates a log of all file conversions, including the original and new filenames, timestamp, and conversion software.

In order to protect the identities of record creators and limit its exposure to risk, the Bentley Historical Library has established policies in regard to personally identifiable information (PII) such as credit card numbers and U.S. Social Security numbers. AutoPro thus employs Identity Finder DLP Endpoint to scan for PII.<sup>7</sup> Archivists then use the Identity Finder interface to verify search results and—if true positive hits are found—redact the PII (from Open Office XML and plain text files) or assign appropriate access restrictions to the content. A record of identity matches and corresponding archival intervention is maintained with the log files.

Archivists then proceed to a more in-depth appraisal and arrangement of content. AutoPro loads data visualizations (such as the distribution of file extensions, date range of content, relative size of directories, etc.) produced by TreeSize Professional to better characterize and launches Quick View Plus (a file viewing program) to rapidly review a wide range of file types for description in finding aids.<sup>8</sup> While reviewing content with Quick View Plus or the Windows Explorer, archivists use a batch file in the right-click context menu to remove superfluous files or folders to a separations directory. Every effort is made to retain the original order of materials, but archivists may group unorganized content in directories or package content in .ZIP files to simplify the management and storage (with such actions recorded in log files). Archivists also develop a plan as to how content will be deposited in Deep Blue in a manner that is both convenient to end users and in accordance with the intellectual arrangement of material in the finding aid.

Once the arrangement is established, AutoPro calls DROID to extract technical metadata and generate an MD5 checksum for all content (including files in .ZIP archives). Archivists then use the AutoPro interface to apply descriptive and administrative metadata to materials. This step produces a Dublin Core XML file and Excel spreadsheet used to deposit material in Deep Blue, the University of Michigan’s

---

<sup>6</sup> ImageMagick (<http://www.imagemagick.org/script/index.php>) is an open source raster image editor. Ghostscript (<http://www.ghostscript.com/>) is an open source interpreter for the PostScript language and PDF documents that may be used to convert the latter documents to PDF/A. Inkscape (<http://inkscape.org/>) is an open source vector graphics editor. ffmpeg (<http://ffmpeg.org/>; for Windows builds, <http://ffmpeg.zeranoe.com/builds/>) is freely available software used for audio and video recording and conversion. Aid4Mail (<http://www.aid4mail.com/>) is a proprietary email conversion program. Microsoft File Converter (<http://www.microsoft.com/en-us/download/details.aspx?id=11454>) is part of the freely available Office Migration Planning Manager.

<sup>7</sup> Identity Finder Data Loss Prevention (DLP) Endpoint (<http://www.identityfinder.com/us/Business/IdentityFinder/EnterpriseClient>) is proprietary software that can identify potentially sensitive information..

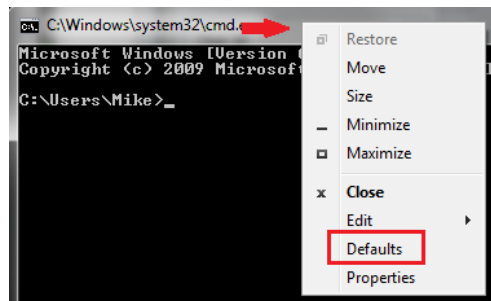
<sup>8</sup> TreeSize Professional is a proprietary hard disk space and file manager and Quick View Plus is a file viewing utility. For more information, see <http://www.jam-software.com/treesize/> and <https://avantstar.com/>, respectively.

DSpace repository.<sup>9</sup> Finally, AutoPro employs BagIt to transfer a copy of all material (and metadata) to a secure dark archives.<sup>10</sup> At the conclusion of processing, AutoPro deletes the working directory, restore point and temporary files and the archivist records the completed digital deposit in the Bentley's collections management database.

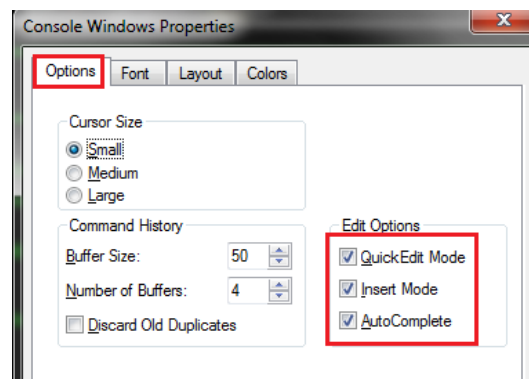
This basic workflow and the component software is subject to change, as the Division of Digital Curation actively tracks the development of standards and professional best practices. It is furthermore recognized that unique features of digital deposits and material may require additional steps to process and record metadata.

## 1.4 Notes on the Windows Command Prompt

This section of the User's Manual provides additional information on features and functionality of the Windows CMD.EXE console (also referred to as the command prompt). The properties for the CMD.EXE console on your work station should have been configured when AutoPro was installed but you may want to check to make sure that the following options are set. Open a CMD.EXE console window (enter CMD.EXE into the Start Menu's search box), right-click on the border of the window, and select "Defaults" from the context menu.



When the "Console Windows Properties" window opens, make sure that the boxes for the following items are checked under the "Options" tab:



<sup>9</sup> Explore the Bentley Historical Library's archival community in Deep Blue at: <http://deepblue.lib.umich.edu/handle/2027.42/65133>.

<sup>10</sup> BagIt is part of an open source set of transfer tools developed by the Library of Congress. For more information, see <http://sourceforge.net/projects/loc-xferutils/>.



- “QuickEdit Mode” allows you to highlight text with the mouse and then hit “Enter” to copy it to the clipboard
- “Insert Mode” allows you to paste text from the clipboard by right-clicking where you would like to insert text
- “AutoComplete” allows you to hit the “Tab” key to complete the entry of folder and file names when entering path information

Another great feature is the “Command History” (familiar to Linux/Mac terminal users): use the up and down arrow keys to browse through information previously entered into the CMD.EXE console. This feature will be particularly useful when compiling administrative and descriptive metadata!

#### Resources:

AutoPro relies on a number of CMD.EXE utilities and Windows batch file syntax to move content through its work flow. If you’d like to learn more about using the CMD.EXE console and batch files in general, the following sites can provide some basic information:

- *An A-Z Index of the Windows CMD Command Line* (<http://ss64.com/nt/>)
- *(DOS) commands and their usage in batch files*  
(<http://www.robvanderwoude.com/batchcommands.php>)
- *Command-line Reference A-Z*  
(<http://www.microsoft.com/resources/documentation/windows/xp/all/proddocs/en-us/ntcmds.mspx?mfr=true>)

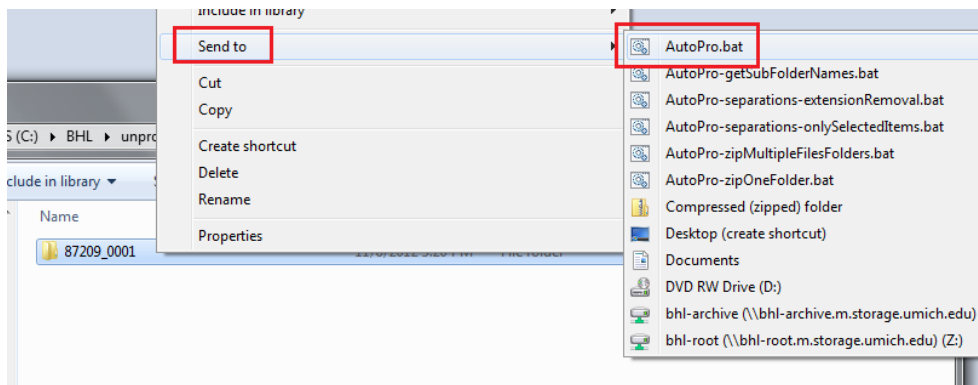
## 2. Initiate a Processing Session

While you may be able to process a deposit of digital content in one session, you may need several sessions over a series of days to completely process and package content so that it is ready for storage in a long-term repository. Follow the steps below to initiate all processing sessions for a given deposit.

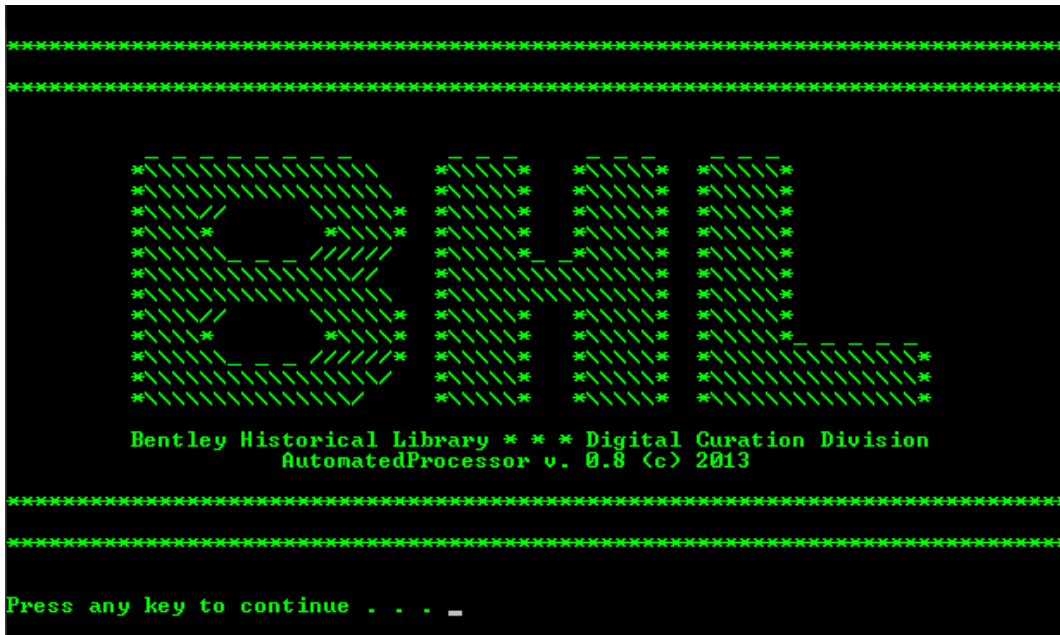
### 2.1 Launching AutoPro

Go to the 'Unprocessed' directory and locate the appropriate directory.

1. Right-click on the processing directory (i.e. \87209\_0001\).
2. Select *AutoPro.bat* from the "Send to" context menu.



The AutoPro application will now open; hit any key to proceed.



## 2.2 Logging In

You will now be required to confirm the processing directory and enter your name so that an audit trail may be generated for the processing session.

### 2.2.1 Directory confirmation

In addition to making sure that you have selected the right processing directory, verify that the Collection ID for the deposit is correct. This ID will be the string of numbers prior to the underscore in the BHL digital deposit number (i.e. the processing directory's name). For example, deposit number 87209\_0001 has a Collection ID of 87209.

```
*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.6 (c) 2012
*****

CURRENT PROCESSING DIRECTORY:
--> C:\BHL\unprocessed\87209_0001
NOTE: Take the time to verify that the collection ID is correct! For
example, UARP 'Records' and 'Publications' have different collection
numbers, as do other formats that are described in their own finding
aid.

If you aren't 100 percent sure, please check. Seriously.
Correct? (Y / N / Q to QUIT): _
```

If this is a new collection, the ID should have been correctly entered by Mike or Nancy. However, if the deposit represents an addition to an existing collection, check the online finding aid to make sure that the Collection ID is the same. The Collection ID will be the first string of digits in the Call Number. **This simple step can save a lot of time and energy down the line!**

Call number: 87209 Bimu C448 2  
Language: The materials are in English.  
Repository: Bentley Historical Library University of Michigan  
1150 Beal Ave.  
Ann Arbor, MI 48109-2113  
Phone: 734-764-3482  
Fax: 734-936-1333  
e-mail: bentley.ref@umich.edu  
Home Page: <http://www.bentley.umich.edu/>

Enter Y if the processing directory (and Collection ID) is correct, N if a different directory needs to be selected, or Q to quit the application.

### 2.2.2 Identification of Processing Archivist

To fully document preservation events and provide a full audit trail of digital processing activities, AutoPro will record the identity of the processing archivist each time someone logs in to work on the deposit.

```
*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.6 (c) 2012
*****

PROCESSING ARCHIVIST
At the prompt, enter your name <Last name, first name--i.e. Doe, Jane>.
Shallcross, Michael
```

After entering your name, you will be asked if it is correct. Enter Y to proceed or N to re-enter your name.

```
*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.6 (c) 2012
*****

PROCESSING ARCHIVIST
At the prompt, enter your name <Last name, first name--i.e. Doe, Jane>.
Shallcross, Michael
Is your name entered correctly? <Y / N> Y_
```

On subsequent sessions, AutoPro will retrieve the name of the last processing archivist from the main log file and ask you to confirm it (see below). If you are taking over from another processor, respond with N and enter your own.

```
*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.7 (c) 2012
*****

PROCESSING ARCHIVIST
Shallcross, Michael
Is your name entered correctly? <Y / N>
```

Once the processing directory and your identity have been verified, AutoPro will set a number of variables and, if this is the first processing session for a deposit, create directories for log files, temporary files, a backup restore point, and separations in addition to generating an initial manifest of the content.

### 3. Main Menu—Selecting Processing Options

After verifying the processing directory and processor name, AutoPro will open its main menu. This screen is divided into four parts:

1. Main Menu of Procedures: This section lists all eleven steps in the AutoPro workflow. Steps should be completed in the order provided, but some content types (such as archived websites) may require variations.
2. Statistics: This section provides up-to-date information on the number of files (and size thereof, in bytes) in the processing directory.
3. Processing Status: This section lists all the procedures completed within the processing workflow. Upon first opening the application, it reports that a “new batch processing sequence” has been initiated.
4. Option Entry: This section permits the user to enter the number of the procedure to run next. All newly initiated workflows must begin with “1. Virus scan.” If you are returning to a previously started workflow, you should enter the next procedure (in numerical order) after the most recently completed step, as listed in the Processing Status section. After selecting a procedure, AutoPro will immediately take you to the selected procedure.

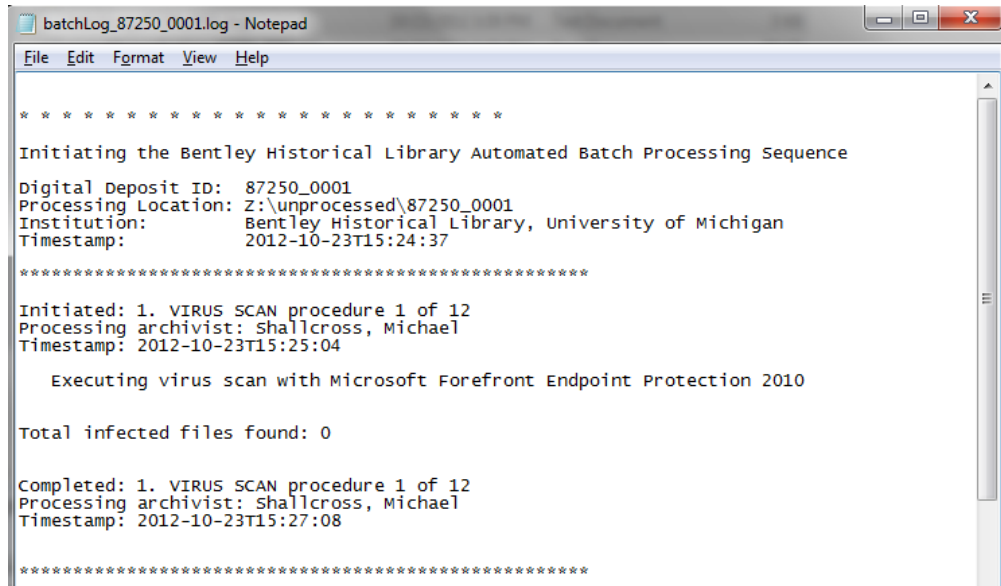
```
*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.8 (c) 2013
*****
1
MAIN MENU OF PROCEDURES:
1. VIRUS SCAN
2. BACKUP CONTENT
3. FILE EXTRACTION
4. IDENTIFY MISSING FILE EXTENSIONS
5. FORMAT CONVERSION
6. PII SCAN
7. APPRAISAL, ARRANGEMENT, and PACKAGING FOR DEPOSIT
8. EXTRACT TECHNICAL METADATA
9. ADD DESCRIPTIVE and ADMINISTRATIVE METADATA
10. TRANSFER to LONG-TERM STORAGE
11. CLEAN UP

Q: Quit the BHL Automated Batch Processing Sequence
*****

AutoPro may be completed in one session or may be stopped at any point.
Quitting an automated procedure mid-operation may result in loss of data.
Please complete a procedure before exiting the sequence.
*****
2
CURRENT STATISTICS--DEPOSIT 9999_0001:
FILES: 3
VOLUME: 6813076 bytes
*****
3
PROCESSING STATUS--DEPOSIT 9999_0001 (as of Tue 04/09/2013):
Completed: 1. VIRUS SCAN procedure 1 of 12
*****

Enter an option: 4
_____
```

At each stage of the workflow, AutoPro will record the procedure name, processing archivist, initiation and completion timestamps, and additional information in the main batch processing log file (see example below).



In addition, all ‘preservation events’ (that is, activities performed by the archivist that impact the provenance, authenticity, and/or integrity of the content) are recorded in a PREMIS (PREservation Metadata: Implementation Strategies) spreadsheet.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	eventType	eventId	eventIdentifierV	eventDateTime	eventDetail	eventOutcome	linkingAgentI	linkingAg	linkingAg	linkingAg	linkingAgentI	linkingAgentRole	
2	virus scan	UUID	d76cfe99-9706-4f	2012-07-27T15:50:2	identification a	action complete	MARC21 Cod	MiU-H	Executor	tool	Microsoft Ant	software	
3	decompression of file	UUID	84c36d87-d675-4e	2012-07-27T15:54:3	decompression action	complete	MARC21 Cod	MiU-H	Executor	tool	7-Zip 9.20	software	
4	personally identifiable	UUID	f119bf41-1b56-47	2012-07-27T15:58:2	identify and re	action complete	MARC21 Cod	MiU-H	Executor	tool	Identity Find	software	
5	message digest calcul	UUID	82262b07-9155-4c	2012-07-27T16:09:3	calculation of N	action complete	MARC21 Cod	MiU-H	Executor	tool	md5deep 4.0.	software	
6													
7													

From left to right, this document records (A) the event type, (B) the event identifier type (here, a Universal Unique Identifier, or UUID), (C) an event identifier value, (D) a time stamp, (E) a description of the event, (F) the outcome, (G) the agent identifier type (here, the MARC21 institution code), (H) the value or identifier for agent (here, MiU-H, the Bentley’s MARC21 identifier), (I) the type of agent, and then information on the software ‘agent’ employed (J-L).

At the conclusion of this (and subsequent) procedures, you will have the opportunity to enter A to ADVANCE to the procedure, M to return to the MAIN MENU or Q to QUIT AutoPro.

```
Options:
  -A: ADVANCE to next step (3. FILE EXTRACTION)
  -M: return to AutoPro MAIN MENU
  -Q: QUIT
<A / M / Q>
```

If you choose to quit, you will resume processing at the point you left off.

## 4. Guide to Individual Procedures

The following pages will provide an explanation of and guidance for running the individual procedures in the AutoPro workflow. **Be sure to take your time!** While processing should be conducted in a timely and efficient manner, there is no race. Do not try to get through the workflow as fast as you can—instead pay careful attention to detail, especially in those steps that require user interaction, review of content, and metadata entry.

Also, despite extensive testing and troubleshooting, unique files or deposit conditions may result in an unrecoverable error or application failure. If you have any issues with AutoPro, please immediately share them with Mike so that a workaround can be developed in an efficient and timely manner.

### 4.1 Virus Scan

AutoPro runs Microsoft Antimalware Service Command Line Utility (a component of Microsoft's Forefront Endpoint Protection 2010) on each file in the deposit. If a virus or malware is detected, the antivirus software will delete the file and record the deletion in a log file.

If one or more infected file cannot be removed, you will be alerted by AutoPro and be permitted to view a listing of such files.

```
*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.6 (c) 2012
*****

ANTIVIRUS SCAN
Scan complete; checking log file for issues...

  Total files scanned: 17
  Total infected files found: 4

User must address outstanding issues; see Digital Curation for assistance.
Hit "Enter" to see a list of files that could not be cleaned or removed.
You may wish to consult the antivirus scan log for more details:
--> C:\BHL\logs\87209_0001\virusScan_87209_0001.log

Press any key to continue . . . _
```

If this occurs, please consult with Mike; upon review of the antivirus log, it may be necessary to manually delete the remaining infected files.

### 4.2 Backup Content

AutoPro creates a working backup copy of all content in the processing directory so that you have a restore point in the event that an irrevocable error or data loss occurs at some stage of the workflow. If the backup encounters any errors, AutoPro will alert you—please consult with Mike to resolve the issue.

If content in the processing directory exceeds 20 GB, you will receive an error message (see below).

```
*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.6 (c) 2012
*****

Initiated: 2. BACKUP CONTENT
*****

This procedure will create a temporary backup copy of the processing
directory that will be deleted at the end of the AutoPro workflow.

  DIRECTORY IS TOO LARGE FOR BHL BACKUP AREA.
  --> Maintain backup copy on original media or in alternate location.

Press any key to continue . . . _
```

To conserve hard drive and interim repository space, we will need to make a backup of the content in another location (such as external hard drive, local desktop, or other location). Please consult with Mike or Nancy as to the appropriate place to backup content in such cases.

### 4.3 File Extraction

This procedure consists of the identification of disk image (such as .E01, .IMG, .AFF, etc.) and archive files (such as .ZIP, .TAR, .RAR, etc.) extraction of content (with its original directory structure) with FTK Imager and 7-Zip, respectively.

#### 4.3.1 Disk Image Extraction

Upon initiation of the procedure, AutoPro will search for common disk image file format. If none are found, the resource will continue on to search for common archive files (see section 4.3.2 below). Please note that given the wide variety of disk image file formats this step may not detect all formats. If additional disk images are detected in the Appraisal and Arrangement workflow step, they may be extracted at that time.

If disk image files are detected, AutoPro will open a text file with the full paths to each file; save locally or refer to it in its location in the 'decompress' folder in the deposit's /tmp/ directory.

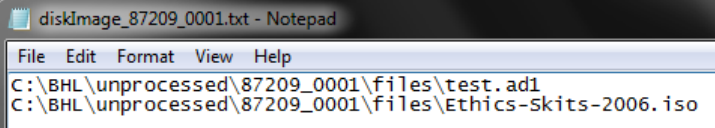
```
This procedure will search for disk image files (.ISO, .AFF, .DMG, etc.)
and archive files (.ZIP, .GZIP, .TAR, etc.) so that content may be
extracted.

Notepad has opened a list of common disk images found in the deposit.
Coordinate with Digital Curation to extract content from the Disk Images
with FTK Imager.

Please note that the contents of the disk images must be scanned for any
viruses or malware.

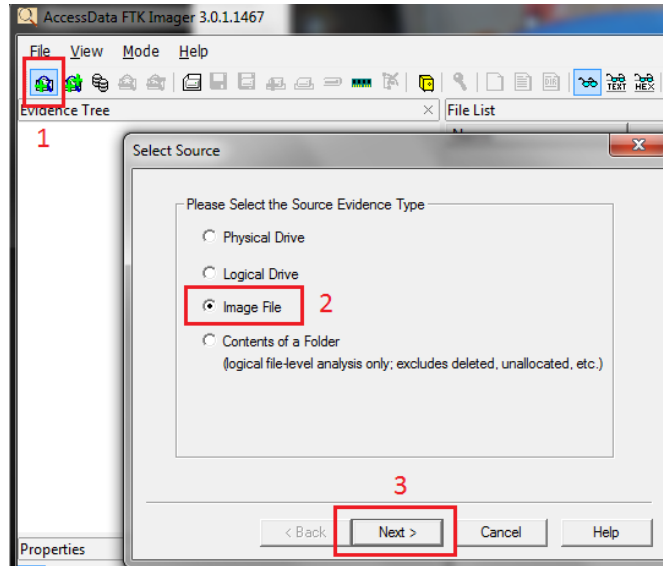
Options:
  -C: CONTINUE processing; AutoPro will conduct a virus scan again. You may
      then return to the main menu and initiate this procedure again.
  -Q: QUIT and resume later

<C / Q>
```

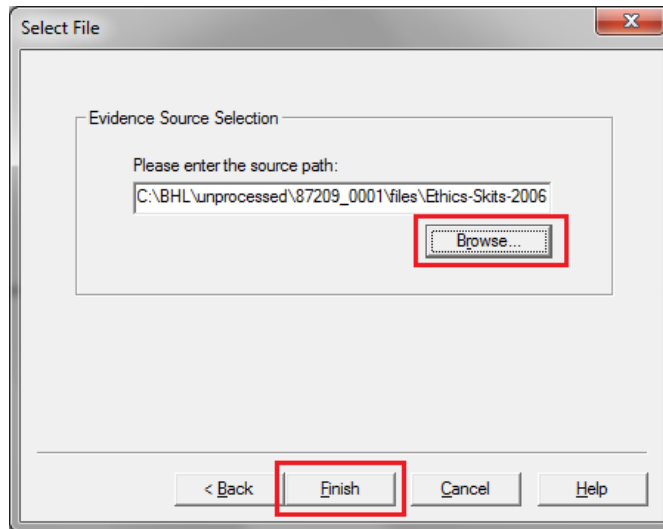




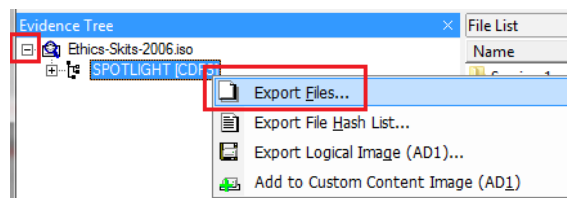
After opening FTK Imager, (1) click the “Add Evidence” icon (or go to File > Add Evidence Item), (2) check the radio dial next to “Image File,” and (3) click “Next.”



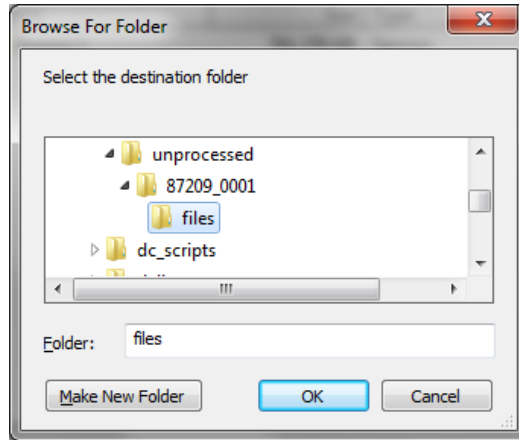
You may then copy and paste the path to a disk image file in the new window’s text box (or browse to its location) before clicking “Finish.”



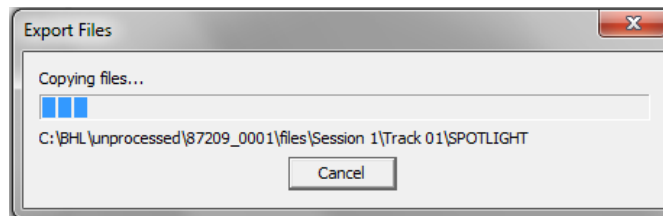
FTK Imager will now show the disk image in the upper-left “Evidence Tree” panel. Expand the disk image file (by clicking on the ‘+’ next to its name), right click on directory immediately within, and select the “Export Files...” option.



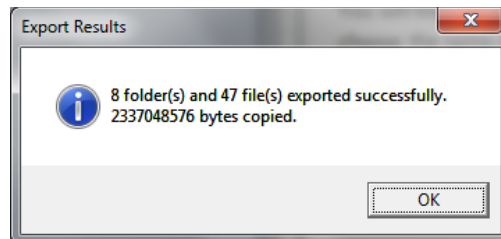
You will then be prompted to select a destination for the extracted files; to maintain original order, choose the same directory that holds the original disk image file.



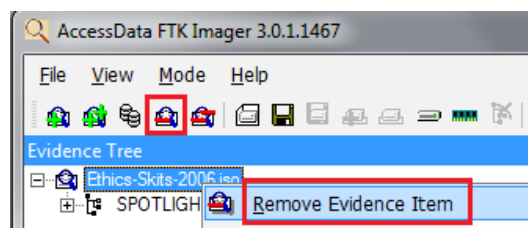
FTK Imager will immediately begin to extract files from the disk image:



When the operation is successfully completed, you will receive a notice:



You may now remove the disk image (or “evidence item”) from FTK Imager; select the image file and either click the “remove evidence item” from the toolbar or right-click and select the option from the menu. You are now ready to proceed to the next disk image!



Once you are done extracting content from the disk image file(s), you may enter *C* into the AutoPro interface to COMPLETE this step. If you need to quit and resume file extraction at a later date, enter *Q* and recommence this procedure the next time you log into AutoPro.

```
Options:
-C: CONTINUE processing; AutoPro will conduct a virus scan again. You may
  then return to the main menu and initiate this procedure again.
-Q: QUIT and resume later
<C / Q>
```

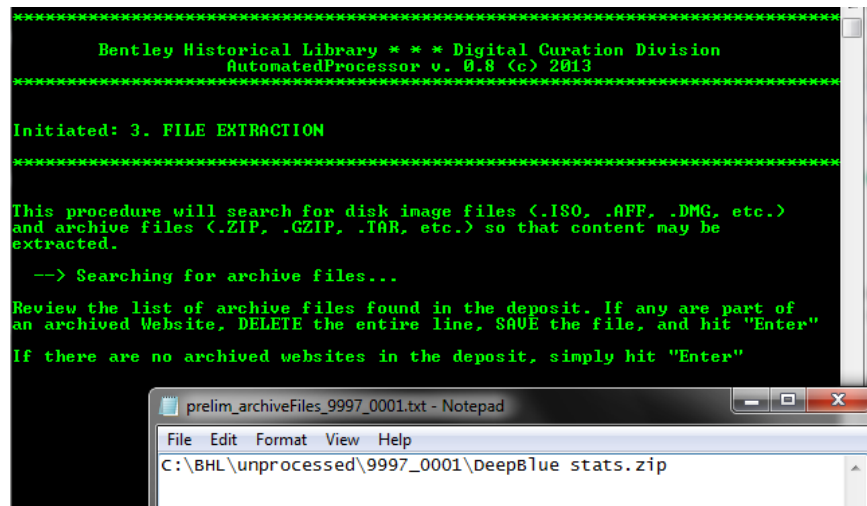
Upon completing the step, AutoPro will rerun a virus scan to verify that the newly extracted contents do not contain a virus or malware. The program will then return to the Extraction procedure to search for other archive files.

### 4.3.2 Archive File Extraction

AutoPro will search for a variety of common archive file formats and, if any are found, extract the contents of each into a newly created folder that bears the name of the original archive file. To avoid name collisions and identify content extracted by the Bentley Historical Library, a suffix (consisting of “-bhl\_” and the eight-character CRC32 hash of the archive file) will be applied to the name of this new directory.

If an archive file cannot be extracted with 7-Zip, you will have an opportunity to manually extract content using the Windows “Compressed (zipped) folder” utility. If the archive file is corrupted, it may be impossible to retrieve content; consult with Mike if issues arise.

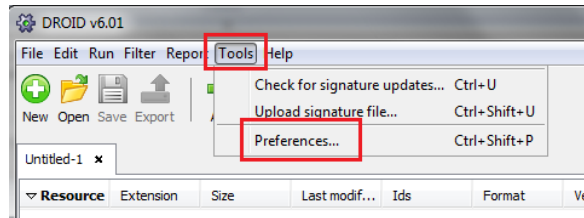
If AutoPro identifies one or more archive files in its initial search, you will be prompted to review a list and to delete any archive files that form part of an archived Website. We will not extract content from archive files in archived Websites because doing so might compromise the overall integrity of the site.



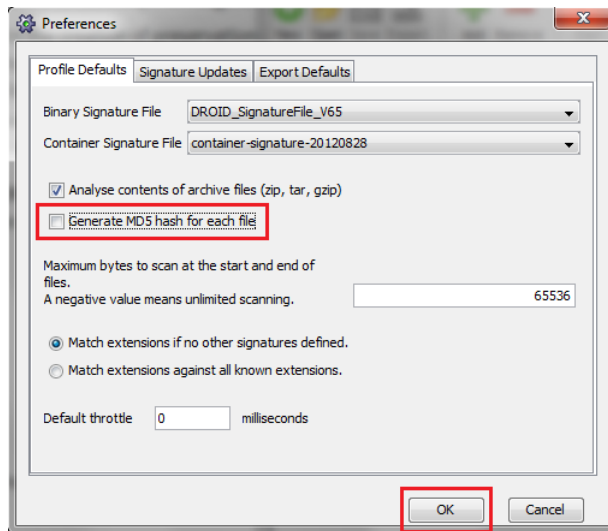
Once you are reasonably sure that no archive files on the list are within an archived Website, hit “Enter” to continue with the procedure.

## 4.4 Identify Missing File Extensions

To aid in the creation of preservation copies and facilitate the eventual use of content by patrons, AutoPro will identify missing or incorrect file extensions. Upon initiation of this procedure, AutoPro will launch DROID and provide instructions to adjust the application's settings so that MD5 checksums are **not** calculated at this point. It is *very* important to do so! If checksums are calculated, this procedure may take a very long time to complete and the resulting spreadsheet of information will not be parsed correctly by AutoPro. After DROID opens, it will check for any updates; if any are available, approve their installation. Next, click on the *Tools* item in the top navigation menu and select *Preferences*.



Make sure that the box next to “Generate MD5 hash for each file” is **NOT** checked. Click OK and then close the DROID window.



You will then be prompted to enter S once these settings have been saved and DROID is closed. AutoPro will launch DROID to produce a spreadsheet that identifies files with missing or mismatched file extensions.

Once this operation is complete, AutoPro will inform you of the number of files identified as having ‘mismatched’ extensions. A text file of these files will then open; each line will contain the full path, PRONOM unique identifier (PUID), and mime type (if the latter two were identified).

```

*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.8 (c) 2013
*****

FILE EXTENSION IDENTIFICATION
8 files identified as having an extension mismatch.
Review the list of files that have been identified by DROID as having a
file extension mismatch.
If you identify files that have been mistakenly included, delete the full
line containing each file from the list. Be sure to save when you are done
Options:
-C: CONTINUE with file extension identification
-N: NO file extension identification is necessary. Finish this procedure
<C / N>
candidates_9873_0001.txt - Notepad
File Edit Format View Help
"X:\unprocessed\9873_0001\Curriculum Development\New Course Material\tiger.ps" "fmt/122" "applicati
"X:\unprocessed\9873_0001\Curriculum Development\New Course Material\BachCPE+SonataAmin_2" "fmt/132"
"X:\unprocessed\9873_0001\Back Up\Curriculum Development\New Course Material\BachCPE_SonataAmin_2" "fr
"X:\unprocessed\9873_0001\Back Up\Curriculum Development\New Course Material\tiger.ps" "fmt/122" "ai
"X:\unprocessed\9873_0001\Back Up\Executive Committee Meeting Materials\Agendas\budget" "fmt/61" "ai
"X:\unprocessed\9873_0001\Back Up\Executive Committee Meeting Materials\Agendas\appraisal recommendations i
"application/msword"
"X:\unprocessed\9873_0001\Executive Committee Meeting Materials\Agendas\budget" "fmt/61" "applicati
"X:\unprocessed\9873_0001\Executive Committee Meeting Materials\Agendas\appraisal recommendations removabl
"application/msword"

```

Review this file and completely delete those lines that do not require extension identification (this may require reviewing content with Quick View Plus—see section 4.7.1.2.1 of this manual). Be sure to save any changes you make to the file! You may then continue with the file extension identification or quit the procedure if you determine that it is unnecessary.

AutoPro will then loop through this list of files, using the TrID file identification utility and the PRONOM file format registry to identify appropriate file extensions.

```

*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.6 (c) 2012
*****

MANUAL FILE EXTENSION IDENTIFICATION
Filename: C:\BHL\unprocessed\87209_0001\Meeting-Materials\jhove.txt
TrID output: probability of file-type association based on binary signature
Warning: file seems to be plain text/ASCII
TrID is best suited to analyze binary files!
64.6% .XMP Adobe Extensible Metadata Platform
27.8% .RDF Resource Description Framework
7.5% .XML Generic XML

PRONOM Format registry information:
Mime type: text/plain
Suggested extension(s) from the PRONOM Format registry:
--> xml

*****

OPTIONS:
- Enter an appropriate 3 character file extension based upon the TrID and
PRONOM output
- Hit "ENTER" to do nothing and leave the file as-is (recommended if file
cannot be identified by TrID/PRONOM or by archivist's review)
Extension (or "ENTER" to do nothing):

```

Based upon these results (and, if necessary, your review of the file) you may choose to adopt one of the proposed extensions. Enter the three-digit extension without a period; AutoPro will apply it to the file and record the operation in a log file. If you are unsure of what the extension should be, simply hit

“Enter”; AutoPro will leave the file unchanged and proceed to the next one on its list. AutoPro will continue until all the files on the list have been reviewed, with the addition of any extensions recorded in a log file.

## 4.5 Format Conversion

The creation of preservation copies for content in at-risk formats is one of the Bentley Historical Library’s primary preservation strategies. AutoPro searches for at-risk formats (using the matrix developed by the Division of Digital Curation) and creates temporary files listing all files associated with the following formats and media types (conversion software noted in italics):

- Raster images (.BMP, .PSD, .PCD, .PCT, and .TGA to .TIF): *Image Magick*
- Raw digital camera image files (.3FR, .ARW, .CR2, .DCR, .MRW, .NEF, .ORF, .PEF, .RAF, .RAW, .X3F to .JPG): *Image Magick*
- Vector images (.AI, .WMF, .EMF to .SVG): *Inkscape*
- PostScript files (.PS to .PDF/A): *Ghostscript*
- Encapsulated PostScript files (.EPS to .PDF/A) : *Ghostscript*
- Portable Document Format (.PDF to PDF/A) : *Ghostscript*
- Audio files (.WMA, .RA, .AU, .SND to .WAV): *ffmpeg*
- Video files (.FLV, .WMV, .RV, .RM, .MTS to .MP4 with H.264 encoding): *ffmpeg*
- DVD-encoded video (VIDEO\_TS folders to .MP4 with H.264 encoding): *Handbrake*
- Office Documents (.DOC, .PPT, .XLS to Open Office XML): *MS Office File Convertor*
- Email (.PST, .MSG, .EML, etc. to .MBOX): *Aid4Mail*

After compiling listing of each format/media type, AutoPro calls subroutines that kick off the various conversion applications, using preset parameters to create preservation copies. Each conversion involves the following steps:

- Check to see that a preservation copy has not already been created
- Generation of the preservation copy.
- Validation of conversion success
- If successful, the conversion is recorded in a log file.

These sustainable versions of content are stored alongside the original bitstream and are differentiated by the addition of a suffix to the filename (“-bhl\_” plus the 8 character CRC32 hash of the original file). Thus, a standard PDF file named “BHL-LogoBlack.pdf” will yield a preservation copy named “BHL-LogoBlack.\_bhl-f14fe8e6.pdf.” For the conversion of PDF files to PDF/A, AutoPro employs JHOVE to verify if the original file meets the PDF/A 1-a or 1-b specifications.

The conversion process is largely automated, with the exception of the email and DVD-encoded video routines. In each case, AutoPro will guide you through operations with detailed instructions. Please consult with Mike if any issues arise!

## 4.6 Personally Identifiable Information (PII) Scan

Personally identifiable information (PII) includes Social Security numbers (SSN), credit card numbers, bank account information, and passwords, any of which may be used to steal an individual's identity or perpetrate fraud in some manner. To ensure that the Bentley Historical Library is aware of the presence of PII, AutoPro uses Identity Finder DLP Endpoint to search for such material. Identity Finder uses regular expressions as well as validation of information (via Luhn algorithms for credit card numbers and dictionaries of known SSN number patterns) to reduce the occurrence of false positive results.

AutoPro immediately initiates Identity Finder; the scan may take a long period of time, especially if the deposit includes large PDFs or email accounts. (AutoPro is configured to skip the scans of image, audio, and video files to reduce processing time). At the scan's conclusion, you will be asked if Identity Finder produced any search results:

```
*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.6 (c) 2012
*****

Identity Finder QUICK START GUIDE

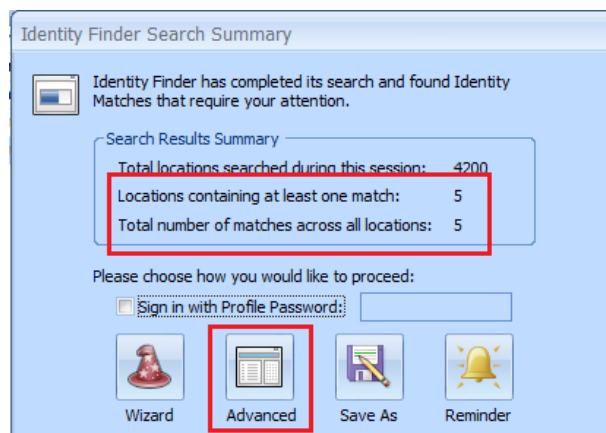
1. Identity Finder is now searching each file in the processing directory
   for Personally Identifiable Information (PII).

2. When the search is complete, an Identity Finder "Search Summary" window
   will appear with details on the number of files (i.e. "locations")
   searched and the number of identity "matches" found, if any.

Options:
-N: there are NO identity matches.
-Y: YES, there are identity matches.

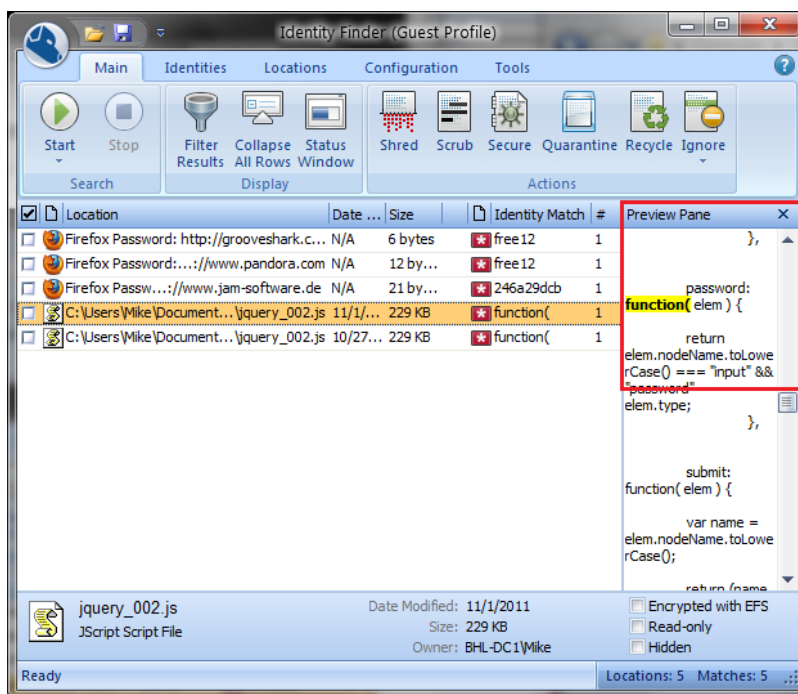
<Y / N>
```

The "identity matches" will be displayed in an Identity Finder Search Summary window. The following example has five matches; the number would read '0' if no matches had been identified.

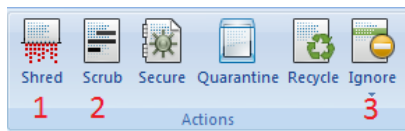


For no matches, you will simply enter *N* at the AutoPro prompt; if there are matches, enter *Y*, click the "Advanced" option and then follow the step-by-step instructions in AutoPro.

If there are identity matches, you will need to review each one to determine if the match is legitimate or merely represents a “false positive.” All files containing identity matches will appear in the *Location Pane* on the left-hand side of the Identity Finder interface. Please note that some of these files may contain multiple matches. If you click on a file (or an identity match), the potentially sensitive information will be displayed in the right-hand *Preview Pane*.



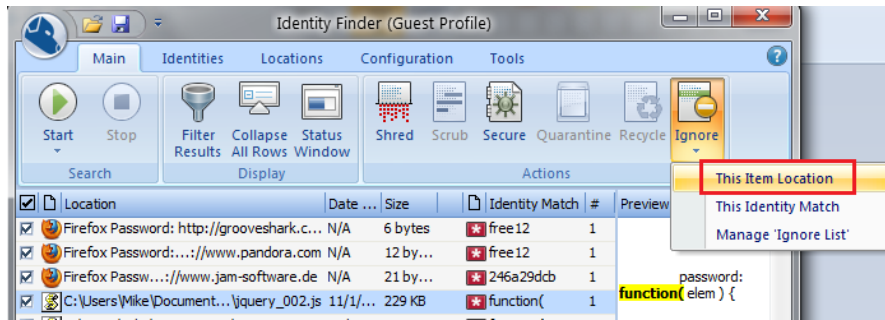
There are three possible actions you can take with content that has been found to contain an identity match:



1. If the content is in non-unique or routine business documents (such as business expense reports or P-Card logs), you may use the "SHRED" option to securely delete content. Check the boxes next to the target filenames, click the “Shred” action, and then confirm your choice. Please note that this option will permanently delete the files—please consult with Nancy, Mike, or division heads if you have any uncertainty!
2. If the identity match is in a plain text file or MS Office Open XML document (i.e. .DOCX, .PPTX, or .XLSX), you may use the “SCRUB” action to redact the PII from file. This option replaces the PII with a string of X’s. Please note that if a preservation copy of a MS Office 1997-2003 file has been created, you may redact the PII from the preservation copy (in OOXML) and then SHRED the original. Please consult with a division head before taking this step.



- For false positives, check the boxes next to the affected files, click on the “IGNORE” action, and then select “This Item Location” from the options.



Finally, if the identity match is in a significant record that must be preserved but cannot be redacted, it may be necessary to impose an access restriction. Please consult with Nancy, Mike or a division head to determine the appropriate restriction.

#### 4.7 Appraisal, Arrangement, and Packaging for Deposit

This procedure provides you with an opportunity to gain full intellectual control of the content so that it can be meaningfully described in a finding aid and packaged in a manner that will facilitate its long-term preservation, management, and access. **Take time to review the files!** Familiarity with the content is essential for the production of rich metadata and informative finding aids; it is therefore *highly recommended* that you (1) *take notes* throughout this procedure to help with separations and gaining intellectual control of content and (2) *establish the intellectual arrangement* before completing any arrangement or packaging of files for deposit.

While digital processing should be done as efficiently as possible, it is very important that you do not rush through this step. If your collection lacks an appropriately detailed finding aid and descriptive metadata, researchers may not be able to discover valuable information therein. You may quit this procedure at any time and return to complete it at a later date.

At the beginning of this procedure, AutoPro will display detailed instructions in a separate CMD.EXE window in regards to the in-depth appraisal and review of content, separations, arrangement, and packaging of files.

```

*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.7 (c) 2012
*****

Initiated: 8. APPRAISAL AND ARRANGEMENT
*****

This stage will help archivists appraise and arrange content in the digital deposit. In
completing this step, the archivist should also record the intellectual and "physical"
arrangement of material in a finding aid.

1. APPRAISAL AND REVIEW OF CONTENT

Archivists may get more information about the size, age, and variety of content by
selecting an option from the below list. In addition, archivists may review content

```

After reading through the instructions, you may hit “Enter” and then choose from the following options to complete this procedure:

```

*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.6 (c) 2012
*****

APPRAISAL AND ARRANGEMENT

Please select from the following options:

CHARACTERIZE DEPOSIT                                REVIEW CONTENT
1: Directory hierarchy                               Q: Quick View Plus
2: Relative size of directories                       U: ULC Media Player (audio/video)
3: File extensions / formats                         IRF: IrfanView (raster images)
4: Age of files                                     INK: Inkscape (vector images)
5: Duplicate content

OTHER
C: COMPLETE procedure
I: view INSTRUCTIONS again
QUIT: QUIT AutoPro and resume appraisal and arrangement at a later time

Enter an appraisal option:

```

#### 4.7.1 In-Depth Review of Content

AutoPro has two sub-menus of options to facilitate the appraisal and review of digital materials.

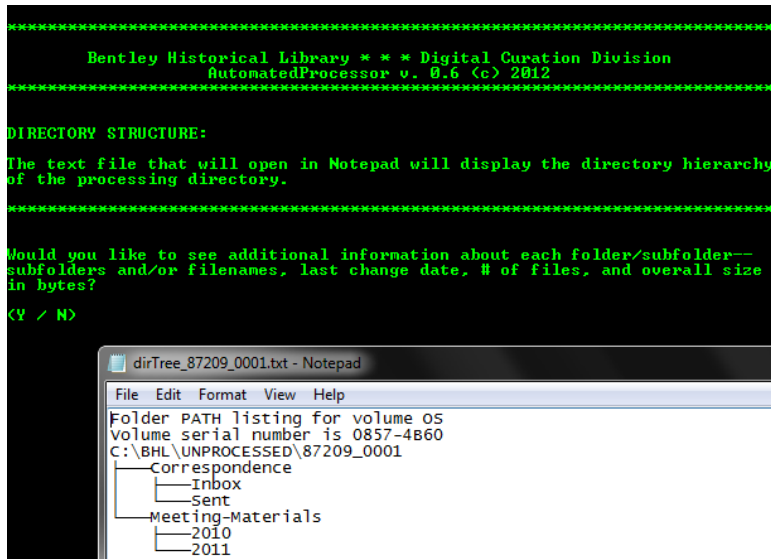
##### 4.7.1.1 The “Characterize Deposit” Menu

Items within the “Characterize Deposit” employ TreeSize Professional and Windows utilities to provide visualizations of various aspects of the deposit.

##### 4.7.1.1.1 Directory Hierarchy

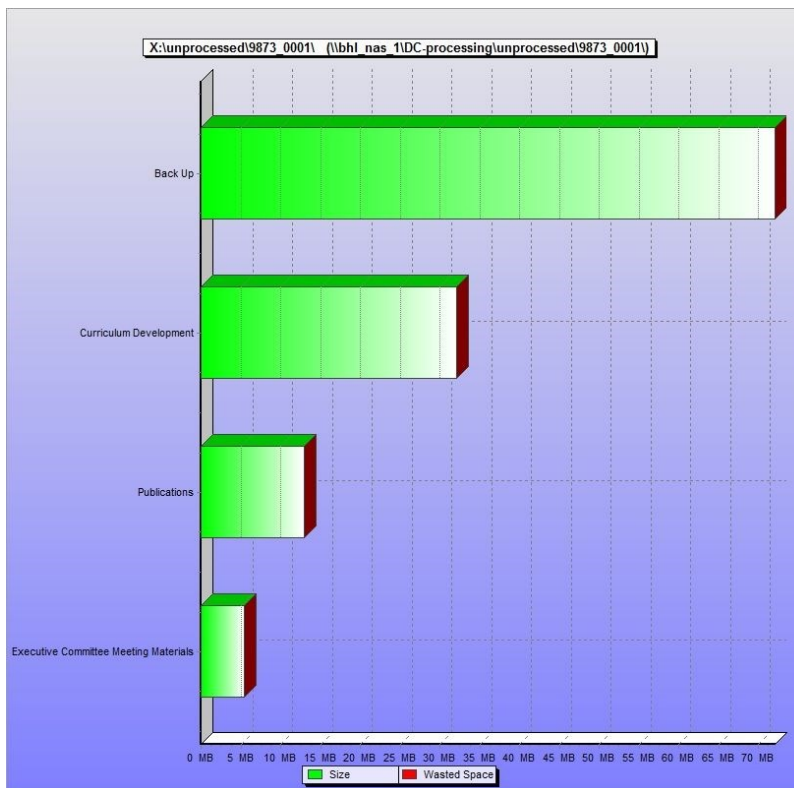
Reviewing the directory hierarchy may aid in the appraisal of the deposit and help you to understand the breadth and scope of content transferred to the library. Folder naming conventions and directory structure may also reveal organizational principles (if any) and areas of particular interest.

The “Directory Hierarchy” displays the structure of folders within the deposit:



#### 4.7.1.1.2 Relative size of directories

The “Relative size of directories” will produce a bar graph comparing the sizes of the folders in the main processing directory.



To examine the relative size of the contents of subfolders, simply complete the folder path in the AutoPro interface or enter C to complete the step.

```

*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.6 (c) 2012
*****

DIRECTORY SIZE:

This operation will produce a bar chart representing the relative size of
folders within the processing directory.

To view the relative size of subfolders/content within a particular
directory, complete the path with the appropriate folder name.

--> Enter C to COMPLETE this step.

C:\BHL\unprocessed\87209_0001\

```

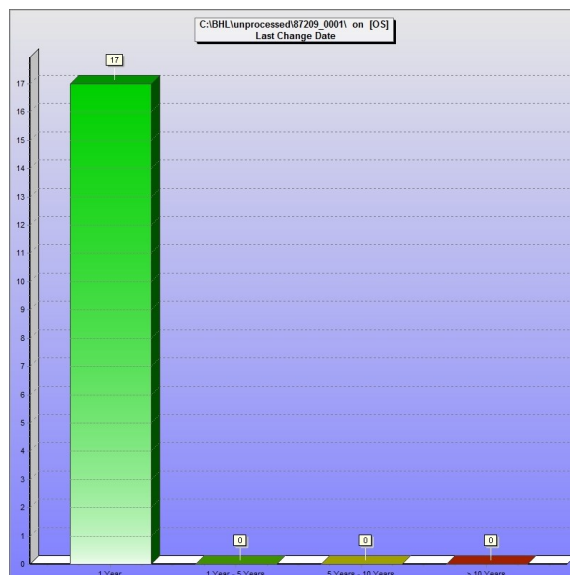
#### 4.7.1.1.3 File Extensions

The “File extensions” option produces a list of file extensions grouped by format type (images, video, Office files, etc.), with the number of files for each extension included. This information can be useful in understanding the dominant types of materials as well as for noting unique file types.

	A	B	C	D	E	F	G
1	Extension	Size	Allocated	Percent	Files	Description	
2	Container Files	241.4 MB	241.4 MB	51.2%	3	Compressed Archives and disk images	
3	zip	169.3 MB	169.3 MB	35.9%	2	Compressed (zipped) Folder	
4	tar	72.1 MB	72.1 MB	15.3%	1	TAR File	
5	Mail Files	153.0 MB	153.0 MB	32.5%	18	Email messages and files of email clients	
6	.pst	74.6 MB	74.6 MB	15.8%	1	Outlook Data File	
7	.mbox	73.1 MB	73.1 MB	15.5%	8	MBOX File	
8	.eml	5.3 MB	5.3 MB	1.1%	7	E-mail Message	
9	.msg	0.0 MB	0.0 MB	0.0%	2	Outlook Item	
10	Graphic Files	39.2 MB	39.2 MB	8.3%	8	Files containing pictures, images or mouse cursors	
11	.arw	14.3 MB	14.3 MB	3.0%	1	ARW File	
12	.orf	11.6 MB	11.6 MB	2.5%	1	ORF File	
13	.psd	5.0 MB	5.0 MB	1.1%	1	IrfanView PSD File	
14	.bmp	4.1 MB	4.1 MB	0.9%	1	IrfanView BMP File	

#### 4.7.1.1.4 Age of Files

The “Age of files” option generates a bar graph representation of the age range for content, identifying material produced within the year, from 1 to 5 years ago, from 5 to 10 years, and then older than 10.



#### 4.7.1.1.5 Duplicate content

The “Duplicate content” option uses MD5 checksum comparison to identify duplicate material and then produce a .CSV report. The Bentley Library is not going to do mass de-duplication due to the difficult and labor-intensive analysis required to identify the ‘record’ version of content. To maintain original order and prevent the potential loss of record versions of content, you may allow duplicate content to remain in the deposit.

	A	B	C	D	E
1	Name	Path	Size	Last Change	Last Access
2	todo.txt	(2 Files)	22.4 KB	8/27/2012 5:58 AM	8/27/2012 5:58 AM
3	todo.txt	C:\Users\Mike\Desktop\test\FITS\openplanets-fido-ae19806\	11.2 KB	8/27/2012 5:58 AM	8/27/2012 5:58 AM
4	todo.txt	C:\Users\Mike\Desktop\test\FITS\vido\openplanets-fido-ae19806\	11.2 KB	8/27/2012 5:58 AM	8/27/2012 5:58 AM
5	ARL SPEC Kit.zip	(2 Files)	30.0 KB	10/3/2012 1:06 PM	[multiple]
6	ARL SPEC Kit.zip	C:\Users\Mike\Desktop\test\9999_0001\Back Up\Executive Committee Meeting Materials\Correspondence\	15.0 KB	10/3/2012 1:06 PM	11/1/2012 12:46 PM
7	ARL SPEC Kit.zip	C:\Users\Mike\Desktop\test\9999_0001\Executive Committee Meeting Materials\Correspondence\	15.0 KB	10/3/2012 1:06 PM	2/6/2012 9:53 AM
8	Updated_Agenda.eml	(2 Files)	36.6 KB	[multiple]	[multiple]
9	Updated_Agenda.eml	C:\Users\Mike\Desktop\test\9999_0001\Back Up\Executive Committee Meeting Materials\Correspondence\L	18.3 KB	#####	11/1/2012 12:46 PM
10	Updated_Agenda.eml	C:\Users\Mike\Desktop\test\9999_0001\Executive Committee Meeting Materials\Correspondence\Literature I	18.3 KB	10/3/2012 1:06 PM	10/28/2011 1:37 PM
11	Agenda_for_News_and	(2 Files)	37.2 KB	[multiple]	[multiple]
12	Agenda_for_News_and	C:\Users\Mike\Desktop\test\9999_0001\Back Up\Executive Committee Meeting Materials\Correspondence\L	18.6 KB	#####	11/1/2012 12:46 PM
13	Agenda_for_News_and	C:\Users\Mike\Desktop\test\9999_0001\Executive Committee Meeting Materials\Correspondence\Literature I	18.6 KB	10/3/2012 1:06 PM	10/28/2011 1:36 PM
14	ccn.doc	(2 Files)	38.0 KB	10/3/2012 1:06 PM	[multiple]
15	ccn.doc	C:\Users\Mike\Desktop\test\9999_0001\Executive Committee Meeting Materials\Agendas\	19.0 KB	10/3/2012 1:06 PM	11/1/2011 1:02 PM
16	ccn.doc	C:\Users\Mike\Desktop\test\9999_0001\Back Up\Executive Committee Meeting Materials\Agendas\	19.0 KB	10/3/2012 1:06 PM	11/1/2012 12:46 PM
17	CCE_assignment.doc	(2 Files)	49.0 KB	10/3/2012 1:06 PM	[multiple]
18	CCE_assignment.doc	C:\Users\Mike\Desktop\test\9999_0001\Back Up\Executive Committee Meeting Materials\Agendas\	24.5 KB	10/3/2012 1:06 PM	11/1/2012 12:46 PM
19	CCE_assignment.doc	C:\Users\Mike\Desktop\test\9999_0001\Executive Committee Meeting Materials\Agendas\	24.5 KB	10/3/2012 1:06 PM	11/1/2011 1:02 PM

At the same time, if the duplicate content report reveals a high concentration of duplicate content—or entire folders that were used to backup material—you may separate this content with the methods described below.

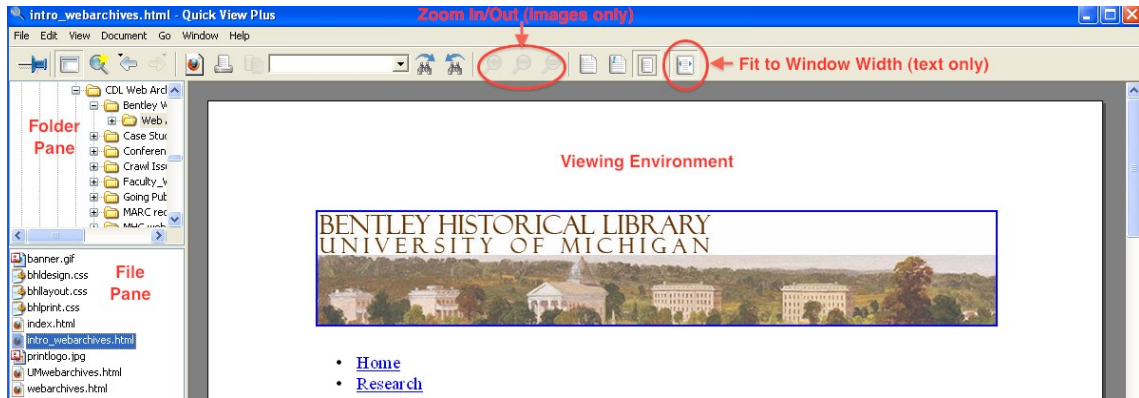
#### 4.7.1.2 The “Review Content” Menu

AutoPro also brings together various tools that can be used to review materials; of these, Quick View Plus (QVP) is ideal for browsing through a wide variety of Office files, images, PDFs, some email formats (including .PST and .MBOX), and other common file types. IrfanView’s thumbnail view allows for quick browsing of large image galleries (although QVP is able to view many of the same files) while Inkscape should be used for vector images and VLC Media Player for audio and video content. The following will provide a brief overview of how to operate these various applications after opening them via AutoPro.

##### 4.7.1.2.1 Quick View Plus:

The QVP interface is divided into three main parts in addition to the navigation menu and ribbon at the top of the application window. The right portion of the interface holds the *Viewing Environment* while the left-hand side is divided between the *Folder Pane* on the top and the *File Pane* on the bottom.

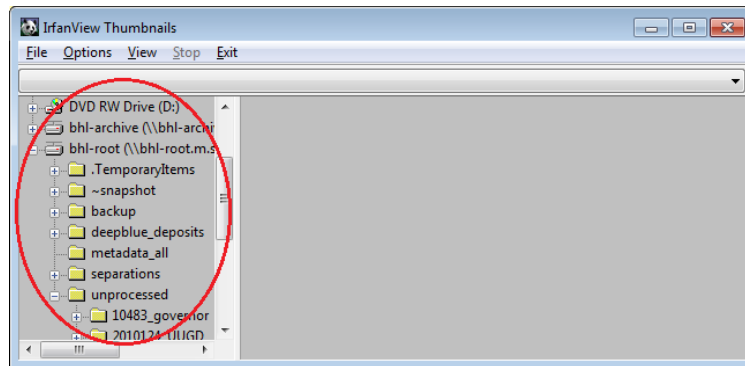
After QVP opens, use the mouse or arrow keys (right and left arrows may be used to expand/collapse subfolders) to navigate to the appropriate directory in the *Folder Pane*. Once the appropriate folder has been selected, a list of its contents (both subfolders and files) will be displayed in the *File Pane*. You may use the mouse or the tab key to move to the *File Pane*; then, whatever file is highlighted will appear in the *Viewing Environment*. Please note that very large files (especially email) may take longer to open.



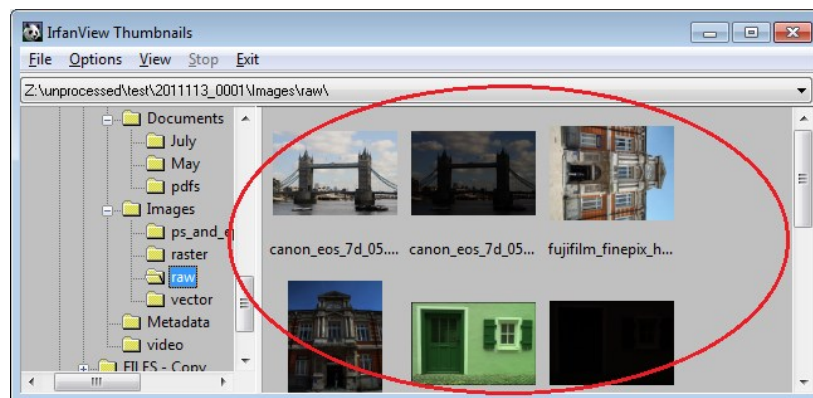
While viewing a file, you may search for text in the ribbon’s text box, zoom in/out (for images only) and adjust the document size so that it fits entirely in the view area or is fit to the area’s width.

#### 4.7.1.2.2 IrfanView:

When IrfanView is selected, AutoPro will open the application in “thumbnail” view, with the main processing directory selected in the left-hand folder pane. It will be necessary for you to navigate down into the directory hierarchy to the folder(s) that contain image galleries.

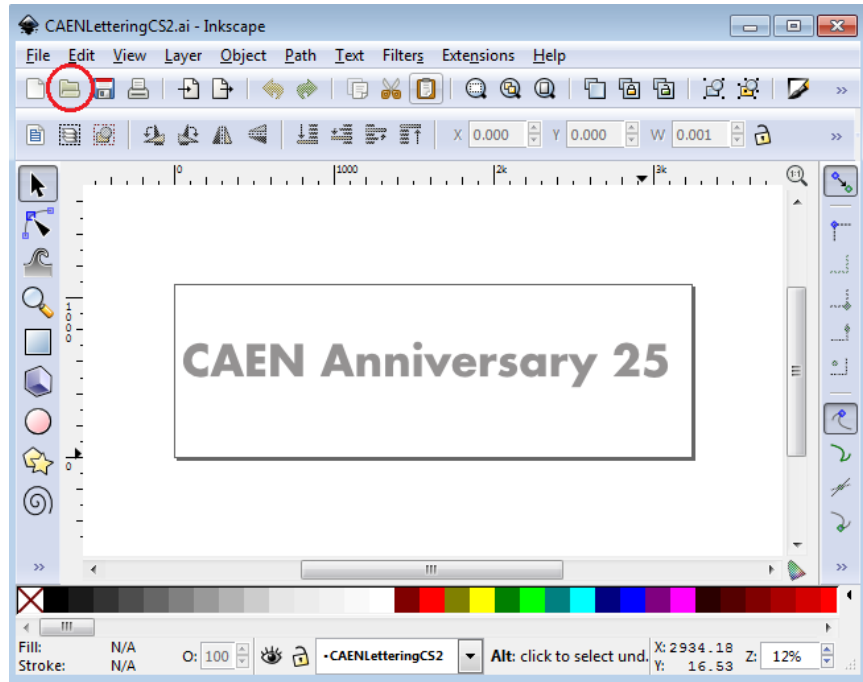


Once you have selected a folder with images, IrfanView will automatically load thumbnails of all image files into the viewing pane. You may then click on a thumbnail to take a closer look at the full-scale or use the navigation pane to browse to other folders.



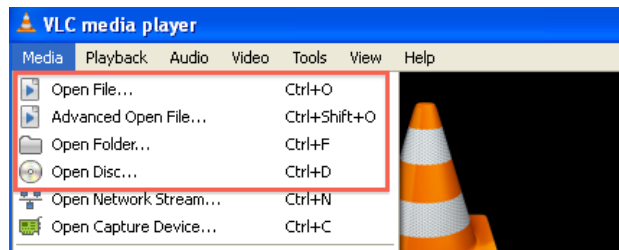
#### 4.7.1.2.3 Inkscape:

Inkscape may be used to view vector image files (such as .SVG, .AI, .WMF, etc.) that cannot be opened in IrfanView or Quick View Plus. AutoPro will open Inkscape and you will then need to click on the open file icon (or use *File > Open* in the navigation menu) to access vector image files.



#### 4.7.1.2.4 VLC Media Player:

AutoPro will open VLC Media Player after which you may use the *Media* menu item to open a single file or an entire folder (in which case, all audio/video files within the folder will be added to a playlist).



Play controls are located at the bottom of the Media Player window; in addition to *Play*, *Pause*, and *Stop* buttons, the archivist may fast forward or reverse progress by adjusting the slider on the progress bar.



If there are multiple files in a playlist, you may click the arrow keys to move to the next/previous item in a playlist.

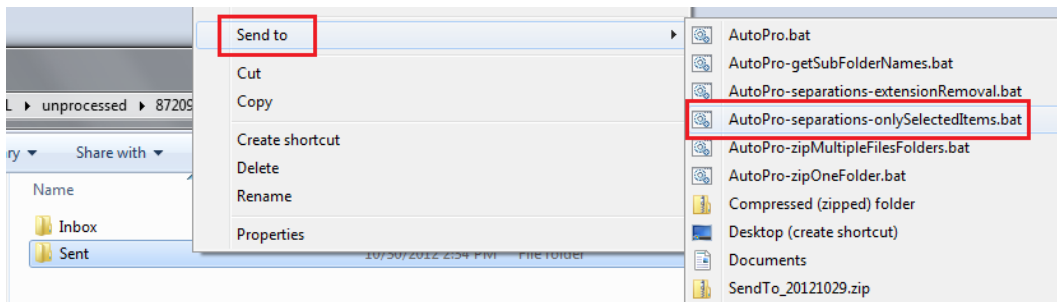
### 4.7.2 Separations

The appraisal and review process may reveal content that should be separated from the collection prior to its deposit in a long-term repository. This may include certain file types and content deemed to be superfluous or outside the collecting scope of the library. When content is moved to the separations directory, AutoPro will recreate the folder structure of the deposit so that separated materials retain their original position and context.

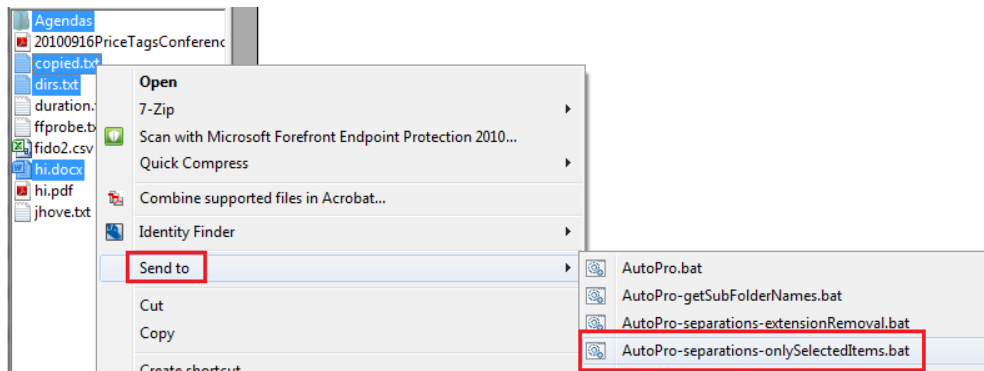
Upon initiation of the Appraisal and Arrangement procedure, AutoPro will search for and separate a number of common files generated by operating systems. These include thumbs.db and .LNK file shortcuts on Windows and .DS\_STORE and resource fork (.\_\*) files produced by Macs. If you would like to move additional files to the deposit's separations directory, you have two options, both of which are batch files found in the "Send to" section of the right-click context menu. Please note that each option is available when you are reviewing content with Quick View Plus or a Windows Explorer window.

#### 4.7.2.1 AutoPro-separations-onlySelectedItems.bat

Use the batch file "AutoPro-separations-onlySelectedItems.bat" to remove only those items you have selected to the separations directory. As the following example illustrates, this option may be used on a single file or folder:



This option may also be used on multiple files and/or folders, selected by clicking on the left mouse button while holding down the "Shift" key (to separate an entire range of files) or the "Ctrl" key (to choose a select number of files and/or folders for separation):



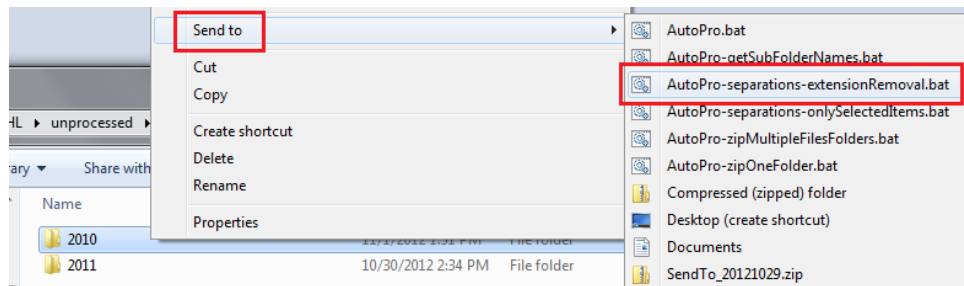


Please note that when selecting multiple files and/or folders for separation, all content must reside in the same parent directory.

#### 4.7.2.2 AutoPro-separations-extensionRemoval.bat

In the second option, the archivist may remove all files of a selected extension from a given folder—and all subfolders within that directory. **Use this option with care!** If used at an upper level of the processing directory hierarchy, all files of the chosen extension will be removed from each subfolder in that branch of the deposit.

To use, this option, right-click on a folder and then click on the batch file “AutoPro-separations-extensionRemoval.bat” from the "Send to" context menu.



A new CMD.EXE window will open and prompt you to enter the extension for the files you wish to remove. Be sure to enter the appropriate extension (as used in Windows) and to precede it with a period (i.e. “.ini”).

```
*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.7 (c) 2012
*****
SEPARATE ALL FILES OF A SPECIFIC FILE EXTENSION
At the prompt, enter the file extension (with leading period) you wish to
remove from this directory. USE CAUTION! The program will remove every file
with this extension to the separations directory.
Example: .ini
Enter file extension with a leading period: .txt
All files with this extension will be removed from this directory: .txt
Is this extension entered correctly? (Y/N)
```

AutoPro will ask you to verify your choice; enter *N* to identify a different extension or *Y* to proceed with the separation. Once AutoPro has an affirmative, all files with that extension will be removed to the separations directory.

At the conclusion of the Appraisal and Arrangement procedure, AutoPro will generate a manifest of all files that have been separated from the deposit (with figures on the number of files and volume) and store a copy of this file with other log files.

### **4.7.3 Arrangement**

The Bentley Historical Library strives to respect provenance and maintain the original order of content in order to preserve important contextual information found in the structure of directories and the associations of different files. Given the structure that many record creators impose on their files and the importance of original order, **additional arrangement will be unnecessary in most cases.**

At the same time, a basic assumption in our digital processing workflow is that folders in the top level of a processing directory are analogous to series. You may therefore need to create a top-level folder in order to organize files (or subfolders) into series. It may also be necessary to impose some organizational principle if the files were copied pell-mell from their source location with no structure. All arrangement must be manually performed via a Windows Explorer window. If you create a folder avoid using spaces and/or non-alphanumeric characters in the names (i.e. "John\_Doe\_Audio\_Files").

**Be sure to complete the intellectual arrangement before you move any files!** Once you have identified series and more precise levels of arrangement, you will have a defined strategy of how to 'physically' arrange material. AutoPro produces a record of the initial arrangement the first time that the application is run on a digital deposit. It will create a final record of arrangement at the close of this procedure—be sure that you have completed all separations and packaging before completing the procedure.

### **4.7.4 Packaging Content for Deposit and Defining Deep Blue 'Items'**

To ensure that content is packaged correctly the first time, you **MUST**:

- Complete your intellectual arrangement.
- Prepare a draft of your finding aid and submit it for review.
- Discuss with Digital Curation how the content will be deposited and presented in Deep Blue. Develop a deposit plan to guide your packaging.

Finally, be sure to complete all packaging before proceeding to the next procedure. There are two options for packaging content, both of which are Windows batch files available via the right-click "Send to" menu.

#### **4.7.4.1 Overview of Deep Blue**

A deposit plan will help you translate the intellectual and 'physical' arrangement of material into convenient and appropriately 'items' in our Deep Blue repository. Understanding Deep Blue's unique nomenclature will be helpful:

- A collection corresponds to a Bentley Library record group or manuscript collection:

The Bentley Historical Library was established in 1936 by the University of Michigan to serve as the official archives of the University and to document the history of the people, organizations and voluntary associations.

**Sub-communities within this community**

- Faculty Archives

**Collections in this community**

- A. Alfred Taubman College of Architecture and Urban Planning
- Andrew Babicki Collection
- Bentley Historical Library (University of Michigan) Records
- Center for Chinese Studies (University of Michigan) Records
- College of Pharmacy (University of Michigan) Records
- Computer Aided Engineering Network (University of Michigan) Records
- Computing Center (University of Michigan) Records
- East Quad Memory Project (University of Michigan) Collection
- Jennifer Granholm Papers
- John L. Kavanaugh Digital Materials Series
- Law School (University of Michigan) Publications
- Medical School (University of Michigan) Records
- Men's Glee Club (University of Michigan) Records**
- Michigan Public Media (University of Michigan) Records
- Michigan Today (University of Michigan)
- NSENET 20th Anniversary Collection

**Browse by:** Titles Authors Subjects Date

Search within this collection:

**BENTLEY HISTORICAL LIBRARY  
UNIVERSITY OF MICHIGAN**

This Deep Blue collection forms part of a larger Men's Glee Club record group held by the Bentley Historical Library. For more complete index to this record group, please consult the [online finding aid](#). Additional records and information may be found in the [Men's Glee Club Web Archives](#).

**Abstract:**  
University of Michigan choral group founded in 1859; digital materials include correspondence, administrative records, publicity materials, photographs, and audio-visual content.

**History:**  
As early as 1846 University of Michigan students met regularly to sing "for pure enjoyment." In 1859 the university formally acknowledged this activity through the establishment of a Glee Club. Michigan's group was the second such college organization formed in the nation.

- An 'item' is a single entity within a Deep Blue collection and may correspond to:

- An entire series:

[Chronological File, 1886-1969, 2000-2011](#)  
Men's Glee Club (2012)  
 (2.8MB) (296.7MB) (549.4MB) (951.8MB) (153.2MB)

- A subseries, a folder, or individual item (as represented in a finding aid). The following example show 'items' related to the Men's Glee Club 'China Concert Tour' series:

[China Concert Tour - May 12, 2012: Wuhan - Hubei Provincial Museum](#)  
Men's Glee Club, University of Michigan (2012)  
 (105.1MB)

[China Concert Tour - May 12, 2012: Wuhan - Qintai Grand Theater](#)  
Men's Glee Club, University of Michigan (2012)  
 (1.3GB) (66.7MB) (513.4MB) (381.9MB)

[China Concert Tour - May 13, 2012: Shanghai - Oriental Art Center](#)  
Men's Glee Club, University of Michigan (2012)  
 (1.4GB) (66.5MB) (1.3GB) (350.6MB) (1.6GB) (1.8GB) (1.8GB) (1.8GB) (1.8GB) (517.4MB)

[China Concert Tour - May 14, 2012: Shanghai - Yu Yuan Garden](#)  
Men's Glee Club, University of Michigan (2012)  
 (76.2MB)

- As the above examples show, an 'item' may have one or more files or .ZIP packages associated with it.
  - In most cases, there will be a one-to-one relationship between item and file/package.
  - Avoid associating multiple files or packages with a Deep Blue 'item' unless all the materials are closely related (for instance, when a single directory has been split into multiple .ZIP files due to size constraints).

While DSpace was not designed for archival materials, the Division of Digital Curation has established several options (and associated conventions) for describing and depositing materials in Deep Blue. Because Deep Blue has a very flat hierarchy, we cannot create a 'series' in the repository and then nest multiple files or packages within this grouping. This basic limitation has two consequences:

- One or more 'items' may relate to a series, subseries, and/or folder—or parts thereof.
- The intellectual arrangement of a file or .ZIP package will be represented through descriptive metadata and titles, in particular. (See Procedure 4.9 for more information.)

#### 4.7.4.2 Determining How to Package Content

The processing archivist (in consultation with Digital Curation) must determine when it is appropriate to deposit individual files or to package multiple files (or subfolders) into a .ZIP file. This decision will depend upon a number of factors:

- The intellectual arrangement of the material in the finding aid and the actual organization of folders/files in the deposit
- The number of files in the deposit
- The size of files—individually or, if content will be packaged in a .ZIP file, at the directory or subfolder level
- The relative ease with which users will be able to download files. While large audio or video files will be deposited 'as is,' .ZIP packages should not exceed 2 GB, if at all possible
- The level of granularity/specificity with which the digital materials are described and the appropriate level at which to insert links to Deep Blue in the EAD finding aids
- How future additions to the collection (including additional material related to a specific series, subseries, or folder) will be deposited alongside this material.
- The amount of time/labor the packaging will require. **Note:** while we strive for efficiency, the extra time and work required to deposit multiple items (i.e. individual files or smaller .ZIP packages) will give the collection richer descriptive metadata and permit users to have more granular access to materials.

There is no 'one way' to ready materials for deposit into Deep Blue, but the following scenarios will suggest the possible ways in which content may be packaged:

- A deposit with a large number of small Office files related to a single function might be packaged into a single .ZIP file (which might represent an entire series, subseries, or folder).
- If the deposit contains records in separate folders that are related to diverse functions, each of these subfolders may be packaged as a separate .ZIP file. This option will result in multiple packages—perhaps 10, 20, or even more.
- If the deposit contains very large files (uncompressed .TIFF images or video, for example), it may be necessary to divide a single folder into multiple .ZIP file packages so that the materials can be uploaded and downloaded efficiently. In this case, the .ZIP file names should contain the original folder name as well as an indication of the breakdown (alphabetical or chronological):  
*meeting\_minutes-January-April.zip* or *committee\_reports\_A-M.zip*

- Very large files (video or audio, for example) or high-value content that needs to be described or linked to at the item level will be deposited individually.

***\*\*Please note that a single deposit may include any or all of the above packaging options\*\****

***\*\*Discuss appropriate options with Digital Curation\*\****

#### 4.7.4.3 Example Packaging and Deposit Plan

To better understand how content may be deposited, consider the following example from the *Toward A Fair Michigan Records*. The former program director of Toward A Fair Michigan (TAFM) transferred her work files to the library on a single optical disk in 2010. The files were arranged in a single directory with multiple subfolders that reflected different aspects of her work as program director. In processing the collection, the archivist determined that this group of files, taken as a whole, represented a distinct series: **Program Director’s Files**. The subfolders in the main directory were determined to represent a ‘level 2’ hierarchy, as they dealt with specific functions of her role (i.e. Board Materials, Correspondence, Litigation, Press, Programs, etc.).

With these considerations in mind, the TAFM packaging/deposit strategy led to the following representation in EAD:

<u>Container / Location</u>	<u>Title</u>
	<b>Program Director's Files</b> [series]:
	The Program Director's Files series is comprised of born-digital records created and maintained by TAFM Program Director Carol Allan and includes incoming and outgoing correspondence, press releases, organizational bylaws, and logistical information for public events. The series also features audio and video recordings of public forums on the Michigan Civil Rights Initiative. This series contains materials related to the other two series in the collection.
Online	Board, 2006 (ZIP archive file) [ <a href="#">download item</a> ]
Online	Correspondence, 2006 (ZIP archive file) [ <a href="#">download item</a> ]
Online	Leadership Assembly, 2006 (ZIP archive file) [ <a href="#">download item</a> ]
Online	Litigation, 2006 (ZIP archive file) [ <a href="#">download item</a> ]
Online	Organizational Documents, 2005 (ZIP archive file) [ <a href="#">download item</a> ]
Online	Press, 2005-2007 (ZIP archive file) [ <a href="#">download item</a> ]
	Programs
	Audio files
Online	Benton Harbor debate, 2006 (ZIP archive file) [ <a href="#">download item</a> ]
Online	Ellis Cose Colloquium, 2006 (ZIP archive file) [ <a href="#">download item</a> ]
Online	Grosse Pointe Unitarian debate, 2006 (ZIP archive file) [ <a href="#">download item</a> ]
Online	Mackinac Center-Lansing debate, 2005 (ZIP archive file) [ <a href="#">download item</a> ]
Online	Michigan Black Republican Council talk, 2005 (ZIP archive file) [ <a href="#">download item</a> ]
Online	Union Missionary Baptist Church debate, 2006 (ZIP archive file) [ <a href="#">download item</a> ]
Online	Wayne State University debate and summation, 2006 (ZIP archive file) [ <a href="#">download item</a> ]

The large size of the deposit (11.5 GB) required the processing archivist to package the content into multiple zip files. Rather than create arbitrary divisions within the content, each subfolder was packaged in a single .ZIP file, with the exception of the ‘Programs’ folder. This directory included

logistical information about TAFM’s programs as well as some fairly large audio recordings of public events. A decision was made to package each of the audio recordings (i.e., both the original file and preservation copy created by the library) into separate .ZIP files so that users could access a single event without having to download a very large file that contained recordings of all the programs. The materials were placed in Deep Blue in accordance with the intellectual arrangement and deposit plan:

[Program Director's Files - Leadership Assembly](#)

Toward a Fair Michigan (2013)

(2.8MB)

[Program Director's Files - Litigation](#)

Toward a Fair Michigan (2013)

(1.5MB)

[Program Director's Files - Organizational Documents](#)

Toward a Fair Michigan (2013)

(270.9KB)

[Program Director's Files - Press](#)

Toward a Fair Michigan (2013)

(752.2KB)

[Program Director's Files - Programs - Audio Files - Benton Harbor Debate](#)

Toward a Fair Michigan (2013)

(902.8MB)

[Program Director's Files - Programs - Audio Files - Ellis Cose Colloquium](#)

Toward a Fair Michigan; Cose, Ellis (2013)

(763.6MB)

[Program Director's Files - Programs - Audio Files - Grosse Pointe Unitarian Debate](#)

Toward a Fair Michigan (2013)

(735.5MB)

#### 4.7.4.4 Packaging Content with *AutoPro-zipOneFolder.bat*

To zip a single folder and all files and subfolders therein into a single uncompressed .ZIP file, right-click on the folder and select the batch file “AutoPro-zipOneFolder.bat” from the “Send to” context menu.



This operation may be performed on a top-level folder within the processing directory (i.e. \87209\_0001\) or a subfolder thereof. After selecting the batch file, a new CMD.EXE window will open to track the progress of the operation. AutoPro uses 7-Zip to package and verify the .ZIP files; be patient, as large directories will take a longer time. The batch file will package all content within the target folder into a .ZIP file of the same name, located in the same parent directory (see below).

Name	Da
2010	11,
2011	11,
2011.zip	11,

After the .ZIP file has been created and verified, AutoPro will notify you if an error occurred during the packaging process (and give you the option to delete the defective .ZIP file, if one was created).

```

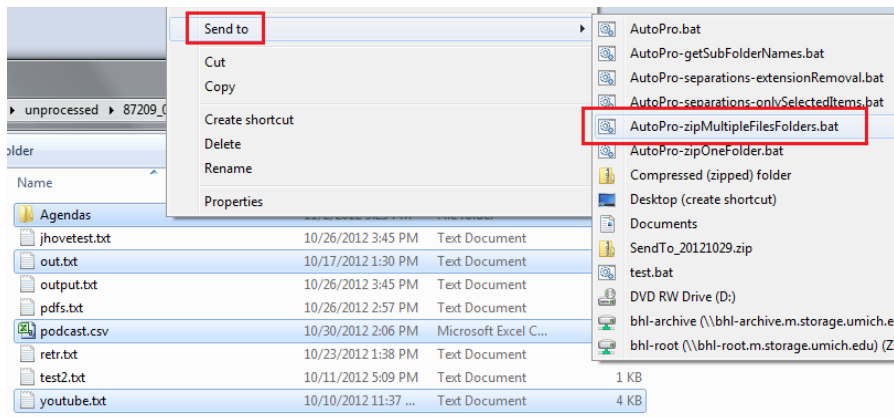
*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.7 (c) 2012
*****
--> PACKAGING FOLDER: C:\BHL\unprocessed\87209_0001\Meeting-Materials\2011
NOTE: Please be patient, as larger files may take longer.
COMPRESSION SUCCEEDED:
--> C:\BHL\unprocessed\87209_0001\Meeting-Materials\2011.zip

```

If the .ZIP file creation was successful, AutoPro will close the window and remove the content that was packaged in the .ZIP file. Please note that in some cases (due to file name length or other issues) AutoPro will be unable to completely clean a directory. If this occurs, you may manually delete the folder/files or consult with Mike.

#### 4.7.4.5 Packaging Content with *AutoPro-zipMultipleFilesFolders.bat*

The other option to package content is to place multiple files and/or folders into a single uncompressed .ZIP file. Please note that all of the target files/folders must be located in the same parent directory and that the resulting .ZIP file will be created in the same directory. To initiate this procedure, select the desired files and/or folders, using the “Shift” or “Ctrl” keys and the left mouse button as needed. Once the appropriate content is selected, right-click on one of the items and choose the batch file “AutoPro-zipMultipleFilesFolders.bat” from the “Send to” context menu.



You will then be asked to enter a filename; use a descriptive, DACS-compliant name for the material, avoiding spaces and non-alphanumeric characters. Please note that AutoPro will supply the .ZIP extension.

```

Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.7 (c) 2012
*****
--> PACKAGING CONTENT:
You have selected multiple folders/files to be added to a .ZIP file stored
in the following directory:
C:\BHL\unprocessed\87209_0001\Meeting-Materials\2010\
At the prompt, enter a filename (without .zip extension) or Q to QUIT.
(Q / filename) SpecialMeetingSessions_2010
Filename: SpecialMeetingSessions_2010.zip
Is filename correctly entered? (Y / N) y

```

After confirming (or changing) the filename, AutoPro will create a .ZIP file of that name in the same directory and then verify the completeness of its contents. If there are any errors in the process, you will be notified by an error message (and given the option to delete the defective .ZIP file, if one was created). If the .ZIP file creation was successful, AutoPro will close the window and remove the content that was packaged in the .ZIP file. Please note that in some cases (due to file name length or other issues), AutoPro will be unable to completely clean a directory. If this occurs, you may manually delete the folder/files or consult with Mike.

#### 4.7.5 Completing the Appraisal and Arrangement Procedures

Once you have completed all the procedures in this workflow step—appraisal and review of content, separations, arrangement, and packaging—select option C from the Appraisal and Arrangement main menu to COMPLETE the procedure.

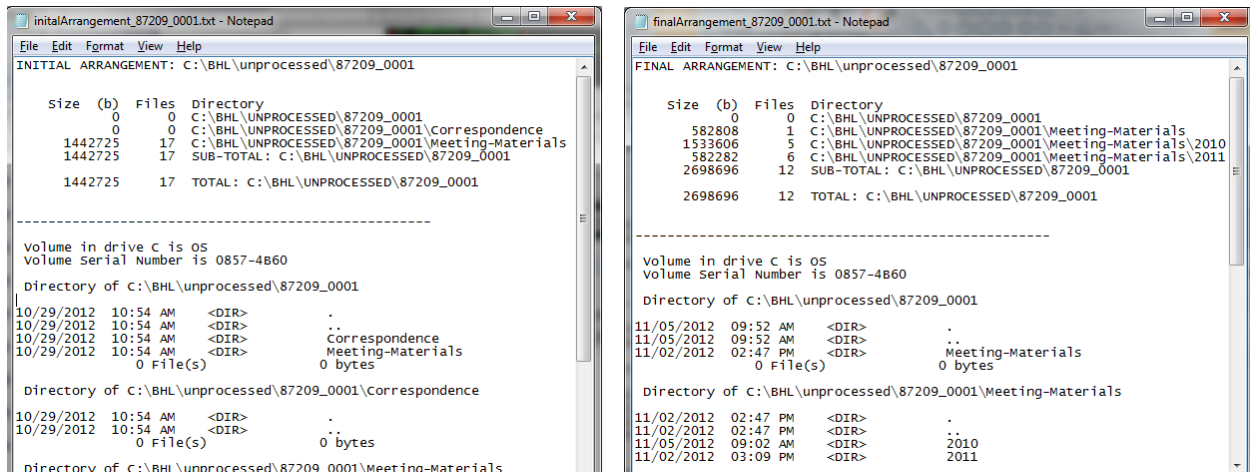
```

*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.6 (c) 2012
*****
APPRAISAL AND ARRANGEMENT
Please select from the following options:
CHARACTERIZE DEPOSIT          REVIEW CONTENT
1: Directory hierarchy        Q: Quick View Plus
2: Relative size of directories U: ULC Media Player (audio/video)
3: File extensions / formats  IRP: IrfanView (raster images)
4: Age of files              INK: Inkscape (vector images)
5: Duplicate content
OTHER
C: COMPLETE procedure
I: view INSTRUCTIONS again
QUIT: QUIT AutoPro and resume appraisal and arrangement at a later time
Enter an appraisal option:

```

It is very important that you wait to complete the step until after all separations, arrangement, and packaging (if necessary) have concluded, as it search for and remove any empty directories in the deposit and create a final record of the deposit’s arrangement.





Taken together, the initial and final arrangement log files document the changes wrought by the Bentley Library on the structure and content of the deposit. Both files include a listing of all the folders in the processing directory (with the number of files and relative size of each), statistics for the deposit as a whole, and then an item-level manifest of the contents of every folder.

After formally completing this step, you will also have an opportunity to view a .CSV spreadsheet of information (number of files and final size in bytes) on the .ZIP files created during the packaging phase.

```

*****
BENTLEY HISTORICAL LIBRARY * * * DIGITAL CURATION DIVISION
AutomatedProcessor v. 0.7 (c) 2012
*****
PACKAGING: ZIP FILE INFO

View information on .ZIP files--number of files and respective size--for
finding aid? This information may be useful in writing the scope and
content note.

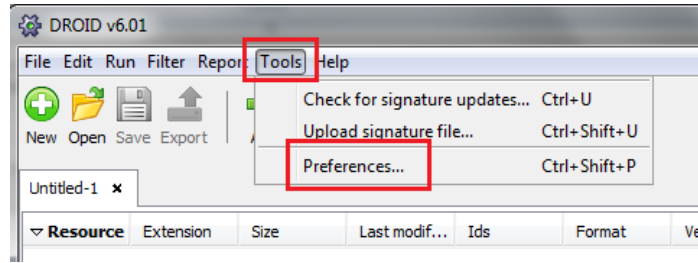
(Y / N)

```

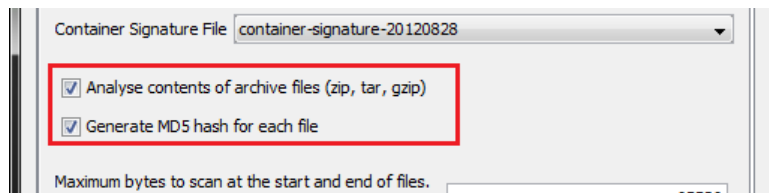
These statistics may be used in the series scope and content notes in the finding aid (see the BHL *Guidelines for Describing Digital Materials in Finding Aids* for more information). If necessary, you may save a copy of the spreadsheet in a local space for later reference.

#### 4.8 Extract Technical Metadata

Upon initiation of this procedure, AutoPro will open DROID so that you may adjust the application's settings. It is essential that you verify these settings, as DROID will provide item-level checksums for all files (including those packaged inside .ZIP files) that may be used at a later date to determine the integrity of content. After approving any updates, click on the *Tools* menu item and select *Preferences*.



Once the Preferences window opens make sure that the boxes next to “Analyse contents of archive files” and “Generate MD5 hash for each file” are checked.



Once you have made the necessary changes, click “OK” and close the DROID window; AutoPro will prompt you to enter an S to verify that the settings are SAVED. The command line version of DROID will now launch and AutoPro will report on the progress as the application creates a profile of the content and then generates a .CSV spreadsheet. **Please be patient!** Deposits with a large number of files or very large individual files (audio, video, email, etc.) may take a long time to be processed—in some cases, hours. Please inform Mike if it appears that DROID is frozen or taking an exceptionally long time to complete its operation.

After DROID has finished, AutoPro will open a spreadsheet of the report for your inspection and then prompt you for a response (see below).

```

*****
          Bentley Historical Library * * * Digital Curation Division
          AutomatedProcessor v. 0.7 (c) 2012
*****

TECHNICAL METADATA EXTRACTION
DROID has generated a CSU spreadsheet for the deposit

Options:
  -C: COMPLETE procedure
  -N: create a NEW DROID report

Be sure to close the spreadsheet before continuing!
<C / N>

```

If there is an error (the most common will be a blank spreadsheet if the process was interrupted), you may run the process again—just be sure to close the Excel window for the report so that a new one may be generated! Otherwise, enter C to COMPLETE the procedure.

#### 4.9 Add Descriptive and Administrative Metadata

This step in the processing workflow requires you to add descriptive and administrative metadata for ‘items’ that will be deposited into Deep Blue. You will be guided by the deposit plan you developed in

Procedure 4.7. The procedure will result in a spreadsheet (to batch upload content to Deep Blue), a Dublin Core XML manifest of materials, and a modified EAD record of administrative metadata for the deposit as a whole.

Please note that you may quit at any point (when given the option) and resume at a later date.

#### ***4.9.1 Make Sure that Intellectual Arrangement and Deposit Plan are Established***

Upon initiation of this step, AutoPro will remind you to finish the intellectual arrangement and description that should have begun with the Appraisal, Arrangement, and Packaging of content (see section 4.7).

```
This procedure will help archivists prepare content for deposit and long-term preservation in the University of Michigan Library's DSpace repository (Deep Blue).
NOTE: Before starting this step, archivists MUST already have:
  1. established the intellectual arrangement of the deposit
  2. described content in a new or updated finding aid
  3. shared the finding aid with a division head or Digital Curation
  4. spoken with Nancy or Mike about how content will be deposited and presented in Deep Blue.
Options:
  -C: CONTINUE; the above requirements have been met!
  -Q: QUIT procedure and return later
<C / Q> _
```

By completing an initial draft of the finding aid or catalog record, you have a thorough understanding of the relationships between different parts of the deposit and, if present, previous accessions to the collection.

As mentioned in Procedure 4.7, DSpace was not designed for archival materials, and so the Division of Digital Curation has established conventions whereby multiple 'items' may relate to a single series, subseries, or folder in the intellectual arrangement and these relationships will be maintained through descriptive metadata (and titles, in particular).

#### ***4.9.2 Add Deposit-Level Metadata***

AutoPro will require you to enter descriptive and administrative metadata about the deposit as a whole, including:

- Creator (entered 'last-name, first-name' if an individual)
- Collection title
- Donor
- Donor ID number
- Year of donation
- Abstract describing the creator and the contents of the deposit (not the collection as a whole)
- Copyright holder (either the Regents of the University of Michigan or a party named by the processor)

```

*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.8 (c) 2013
*****
DEPOSIT METADATA
This step will allow you to enter descriptive and administrative metadata
for the entire deposit.
CREATOR: Smith, Samuel
COLLECTION TITLE: Samuel Smith Papers
DONOR: Susan Smith
DONOR ID: 10234
DONATION DATE: 2012
ABSTRACT: Samuel Smith was a prominent lawyer and political activist in the G
rand Rapids area, born in 1956. Deposit includes digital camera images and O
ffice documents related to political activities with the Michigan state Repub
lican party
Enter 'M' if COPYRIGHT owned by U-M Regents; otherwise enter name: M_

```

After you have supplied each piece of information, AutoPro will give you an opportunity to edit your entries:

```

*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.8 (c) 2013
*****
DEPOSIT METADATA:
CREATOR: Smith, Samuel
COLLECTION TITLE: Samuel Smith Papers
DONOR: Susan Smith
DONOR ID: 10234
DONATION DATE: 2012
COPYRIGHT: Copyright is held by the Regents of the University of Michigan
ABSTRACT: Samuel Smith was a prominent lawyer and political activist in th
e Grand Rapids area, born in 1956. Deposit includes digital camera images an
d Office documents related to political activities with the Michigan state Re
publican party
Do you need to make any changes to the deposit metadata? (Y / N) N

```

If you return to this procedure after having previously started and quit it, the deposit-level administrative and descriptive metadata will display as above and you will have the opportunity to edit it, if necessary.

#### 4.9.3 Check for Pre-Completed Metadata Spreadsheet

If this is the first time you are describing the deposit, AutoPro will check to see if you have a pre-completed descriptive metadata spreadsheet. This spreadsheet **must** be completed in accordance with conventions established for depositing content into Deep Blue. Some metadata ‘massaging’ may be required; consult with Nancy or Mike if you have questions.

```

*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.7 (c) 2012
*****
Do you already have a metadata spreadsheet for the deposit that conforms to
the requirements for Deep Blue? (Consult with Nancy or Mike if necessary.)
(Y / N)

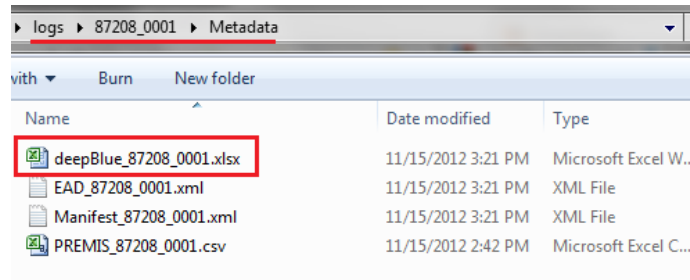
```

If you do not already have a pre-completed metadata spreadsheet, respond to the AutoPro prompt with an *N* and move on to the next section of the User Manual.

If you do have such a spreadsheet (prepared by the donor or created through the reuse of existing metadata), it should be organized with the following columns:

- IDENTIFIER.OTHER: unique BHL digital deposit ID
- DC.TITLE: title of the Deep Blue Item
- DC.DESCRPTION.ABSTRACT: description of the Deep Blue Item
- DC.CONTRIBUTOR.AUTHOR: primary creator of content
- DC.CONTRIBUTOR.OTHER: additional content creators (for people, enter as “Smith, Jane” and separate multiple entries with semi-colons)
- DC.DATE.ISSUED: the year content is deposited in Deep Blue
- DC.DATE.CREATED: the date (expressed in years; may be a range) that content in BHL custody was last modified. Applies to all materials associated with the Item.
- DC.COVERAGE.TEMPORAL: the date (expressed in years; may be a range) that reflects when the information content of the digital material was actually created or used. Applies to all materials associated with the Item.
- DC.TITLE.FILENAME: filename(s) of content associated with the Item; separate multiple entries with semi-colons.
- DC.FORMAT.MIMETYPE: Mime type of files associated with the Item; separate multiple entries with semi-colons. If there are multiple entries, these should be ordered so that they match the order of their respective filenames.
- DC.DESCRPTION.FILENAME: descriptions of the file(s) associated with the Item; separate multiple entries with semi-colons. (NOTE: semi-colons should therefore be avoided in these descriptions.) If there are multiple entries, these should be ordered so that they match the order of their respective filenames.
- DC.DESCRPTION.MD5.CHECKSUM: this field is not required and checksums will be maintained in the DROID report. If included, separate multiple entries with semi-colons and maintain appropriate order of files.
- DC.TYPE: this element indicates the *type(s)* of content in the entire Item; see User Manual section 4.9.11 for more information on content types.
- DC.RIGHTS.ACCESS: this element applies to the entire Item and indicates any access restrictions.
- DC.RIGHTS.COPYRIGHT: this element applies to the entire Item and indicates the content’s copyright holder.
- DC.LANGUAGE.ISO: this indicates the language used for the Item’s metadata entry; in all cases it should be “EN-US”

The spreadsheet should be saved in the deposit Metadata folder (located at `\logs\[deposit-ID]\Metadata`) and must conform to the convention `deepBlue_[deposit-ID].xlsx`. See the example below:



Be sure that the spreadsheet conforms to the above conventions; AutoPro will then prompt you for some general information about the deposit as a whole. You will need to provide date range information about when the files were originally created/used and last modified (since these two dates are often different) as well as what access restrictions are necessary for the deposit. (NOTE: TreeSize Professional can be helpful in determining the range for the deposit's overall last modified dates.)

```

At the prompt, enter the inclusive date range for when the materials in the
deposit were originally created/used. Three example entries are provided:
Examples: a) 2000      b) 2010-2012    c) ca. 1999-2001
2010-2011

At the prompt, enter the inclusive date range for when the files currently
in the Bentley Library's custody were last modified. If the last modified
date is the same as the original creation/use date, hit "Enter".
2012

Summarize the deposit's access restrictions:
  -O: Content is OPEN for research
  -R: Access to content is RESTRICTED; see finding aid for details
<O / R> O

```

If the content of the deposit has already been described in a spreadsheet, this step will conclude the procedure.

#### 4.9.4 Create a New Item

The first time you add descriptive metadata to a deposit, you will be immediately taken to the “Main Item Options” screen:

```

*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.8 (c) 2013
*****

--> ITEM IDENTIFICATION AND DESCRIPTION

In most cases, an "Item" will consist of a single file. Multiple files may
be associated with an Item only if all the files represent a single coherent
intellectual unit.

NOTE: Be sure to discuss the organization of complex deposits with Digital
Curation.

Options:
  -N: create a NEW Item
  -Q: QUIT packaging and resume at a later date
<N / Q>

```

From this point, you may either create a NEW Item or QUIT and resume description at a later date.

After you have added metadata to an Item, you will return to this screen and will have additional options to either REVIEW existing items (which in turn will allow you to add additional content to an item or complete the description thereof) or FINALIZE the deposit's packaging (see section 4.9.13).

#### 4.9.5 Add Item-Level Metadata

The first step in creating a new item involves the entry of descriptive metadata:

```
ITEM METADATA:
This step will allow you to enter descriptive and administrative metadata
for an individual item.

Please remember that the Item title should reflect the hierarchy of the
material's intellectual arrangement.

Enter contributors' names using format "Lastname, Firstname" and separate
multiple names with semi-colons.

If the item contains multiple content types, separate them with semi-colons.
ECHO is off.
See the AutoPro User Manual for more information.

ITEM TITLE: Meeting Materials - State Republican Party Executive Committee
PATH TO ITEM (only if in subfolder; hit "Enter" otherwise): "C:\BHL\unprocess
ed\9994_0001\Meeting-Materials\Republican Party Executive Comm"
DESCRIPTION (no double quote marks or parantheses): Meeting materials collect
ed by Samuel Smith during his service on the Michigan state Republican Party'
s Executive Board
ADDITIONAL CONTRIBUTORS: Davis, Lou; Jones, Sheila
Deep Blue CONTENT TYPE(S): Image; Dataset
NUMBER OF YEARS ITEM MUST BE RESTRICTED (0 if open): 0
```

The elements in this section include:

- Item Title: this will represent the entire intellectual hierarchy of the item and thus should include the series and, if applicable, subseries, folder or even item. Separate each level of hierarchy from the next by a space-hyphen-space. Consider the following intellectual arrangement, in which the audio recording of a debate will be deposited as a single item in Deep Blue:

**Program Director's Files**  
Programs  
Audio Files  
Grosse Pointe Unitarian Debate

The corresponding Item Title should be:

Program Director's Files – Programs – Audio Files – Grosse Pointe Unitarian Debate

It may be necessary to include the date range in the Item Title if there multiple items for a similar resource. Consider the following example from the Office of the Registrar Publications, in which each annual report will be deposited as an individual item:

**Publications**  
Statistics  
Degrees, Awards, and Graduates  
Degree Reports  
2001  
2002

This intellectual arrangement will yield the following Item Titles:

Publications – Statistics – Degrees, Awards, and Graduates – Degree Reports – 2001

Publications – Statistics – Degrees, Awards, and Graduates – Degree Reports – 2002

- **Path to Item:** if the file(s) that will be associated with the Item is/are in the top level of the processing directory, simply hit “Enter.” If, on the other hand, the file(s) is/are within a subfolder in the processing directory, enter the complete path to this folder. Consider the following example, in which the processing directory for the Office of the Registrar deposit is:

Z:\unprocessed\0393\_0001

The Degree Report for 2001 is located in the following directory:

Z:\unprocessed\0393\_0001\degree-reports

The above ‘path’ would therefore be entered as the Path to Item.

- **Description:** provide accurate information about the nature and information in the file(s) associated with the item. In all likelihood, this is the only time this digital material will be described in any detail—this information will aid researchers in their attempt to search for and retrieve digital archival materials from Deep Blue or Google. Be sure to **take your time to include a meaningful description.**
- **Contributors:** include the names of any individuals (entered “last-name, first-name”) or organizations that contributed to the creation of this item. Separate multiple entries with a semi-colon.
- **Deep Blue Content Type:** the administrators of Deep Blue have asked that we identify the “content type” of material we are depositing into the repository. Please note that these ‘types’ do not necessarily correspond to the nature of the Bentley’s digital archives and are used solely for internal purposes in Deep Blue. A text file will open with a list of accepted content types; use semi-colons to separate multiple entries.
- **Number of Years Item Must Be Restricted:** If access to the item is unrestricted you will simply enter a ‘0.’ Otherwise, enter the number of years the content in the item must be restricted from the date of acquisition. For university materials, this will be 20 years for executive



materials, 30 years for personnel materials, 75 years for student records. Please note that attorney-client materials, medical records, and donor instructions may also require temporary or permanent access restrictions.

After you have entered information in all the required fields, AutoPro will then give you an opportunity to make corrections as needed.

```
*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.8 (c) 2013
*****
ITEM METADATA: Meeting Materials - State Republican Party Executive Committee

ITEM TITLE: Meeting Materials - State Republican Party Executive Committee
PATH TO ITEM: C:\BHL\unprocessed\9994_0001\Meeting-Materials\Republican Pa
ty Executive Comm
DESCRIPTION: Meeting materials collected by Samuel Smith during his servic
e on the Michigan state Republican Party's Executive Board
ITEM CONTRIBUTORS: Davis, Lou; Jones, Sheila
CONTENT TYPE: Image; Dataset
ACCESS RESTRICTIONS: Content is open for research

Do you need to make any changes to the Item metadata? <Y / N> N
```

#### 4.9.6 Add Content to the Item

After the Item metadata has been saved, AutoPro will proceed to the “Main Content Options” screen.

```
*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.8 (c) 2013
*****
ITEM: Meeting Materials - State Republican Party Executive Committee
--> IDENTIFICATION AND DESCRIPTION OF CONTENT
Options:
-Enter the full path to an individual file to associate that content
with the Item.
-C: COMPLETE the packaging of this Item
-Q: QUIT packaging and resume at a later date
<PATH / C / Q>
```

You will encounter the above screen each time you have completed the association of content with the Item. From it, you will have the option to ADD new content to the Item, COMPLETE the packaging and description of the Item (in which case you will be returned to the Main Item Options screen described in section 4.9.5), or QUIT and resume at a later date.

#### 4.9.7 Enter Path to File Associated with Item

In the next screen, AutoPro will allow you to provide the full path to a file to be associated with the Item (if needed, you may also complete the Item description or quit and resume the procedure at a later date). To reduce spelling or other errors, it is recommended that you copy and paste the path or use the CMD.EXE AutoComplete functionality.

```

*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.8 (c) 2013
*****

ITEM: Meeting Materials - State Republican Party Executive Committee
--> IDENTIFICATION AND DESCRIPTION OF CONTENT
Options:
-Enter the full path to an individual file to associate that content
  with the Item.
-C: COMPLETE the packaging of this Item
-Q: QUIT packaging and resume at a later date

<PATH / C / Q> C:\BHL\unprocessed\9994_0001\Meeting-Materials\Republican Part
y Executive Comm\january-2011.zip

Are you sure this is the correct file? <Y / N>

```

AutoPro will then prompt you to verify that this is the correct file; if an incorrect path has been entered, you will be directed to re-enter the correct one.

#### 4.9.8 Describing the File

If there will only be one file or .ZIP package associated with the Item, hit “Enter” when prompted to describe the file. If there will be multiple files associated with the item, you will add descriptions to each so that users may accurately search for and retrieve materials. If there are multiple files that are closely related, you may preface the description with “File 1 of 4” or “Part 1 of 4 files” so that users will know there is an interrelated sequence.

```

*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.8 (c) 2013
*****

ITEM: Meeting Materials - State Republican Party Executive Committee
CONTENT: january-2011.zip
--> DESCRIPTION

If there will be more than one file associated with this Item, you should
enter a brief description of the file to differentiate it from the other(s)
If this is the only file associated with the Item, hit "Enter" and continue

Michigan Republican Party executive committee meeting materials from January
2012

Is the above information correct? <Y / N> y

```

AutoPro will give you the opportunity to edit or re-enter the entry; if necessary, you may hit the “up” arrow to bring back the description via the CMD.EXE command history so that you avoid re-typing the entire description).

#### 4.9.9 Original Creation or Use Dates

You will now enter the date (or date range) for when the file was originally created or used. Be sure to use four digit years and, if making an estimate, use the abbreviation “ca.” (with a period). **If the content was scanned or digitized from an original paper, print, or analog resource, include the original creation or use date(s) for the source material** (and not the date on which the digital surrogate was generated).

```
*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.8 (c) 2013
*****

ITEM: Meeting Materials - State Republican Party Executive Committee
CONTENT: June-2011.zip

--> ORIGINAL CREATION OR USE DATES

At the prompt, please enter the year(s) in which the intellectual or
artistic information in the file was originally created or actively used.

  1. Enter 4 digit years; use "ca." if making an estimate (ca. 2000-2005)
  2. If the file is a .ZIP archive, dates should reflect all content therein
  3. If content has been digitized from print or analog materials, use the
    dates when the original items were created or used.

Original creation/use date(s): 2011

Is the above information correct? <Y / N> y_
```

You will then be prompted to verify that the information has been correctly entered. If not, you will have an opportunity to revise the date(s).

#### 4.9.10 Other Information

AutoPro will then record the file's last modified date, retrieve its checksum from the DROID report prepared in Procedure 4.8, and then attempt to identify the file's mime type based on its extension. If the file extension (and mime type) are not recognized, you will be asked to provide a mime type. The following website provides a thorough listing: <http://www.webmaster-toolkit.com/mime-types.shtml>

Consult with Digital Curation if you need assistance.

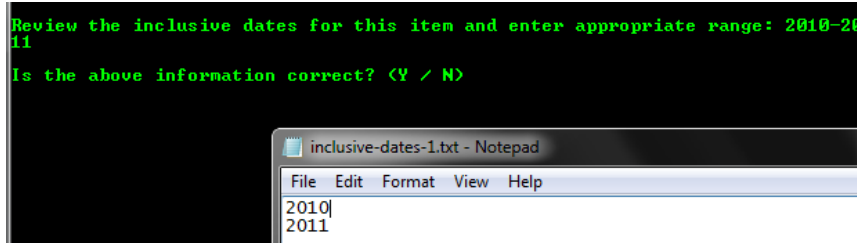
#### 4.9.11 Complete the Packaging and Description of an Item

Having entered all descriptive metadata for the file, you will now asked if you want to ADD new content to the Item, COMPLETE the packaging and description of the Item, or QUIT AutoPro to resume at a later time.

```
ITEM: Meeting Materials - State Republican Party Executive Committee
Options:
  -A: ADD new content to this Item
  -C: COMPLETE packaging content for this Item
  -Q: QUIT packaging and resume at a later date
<A / C / Q> c
```

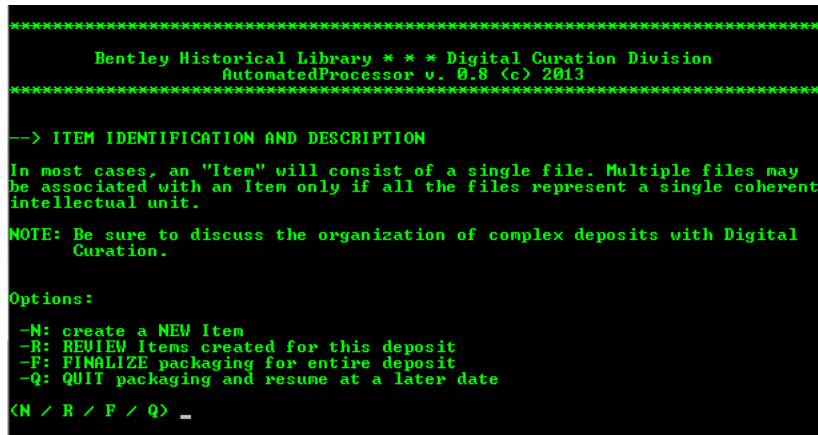
If you ADD new content, you will be returned to the screen described in step 4.9.6, in which you may add the path to another file to be associated with the Item.

When you elect to complete the packaging and description of an Item, AutoPro will record the date range for when the files in the Item were originally created or used. If there are multiple files, you will be asked to enter the date range. AutoPro will open a text file with the dates for each file; review the list and provide an accurate date (or range).



AutoPro will then write out the metadata for the Item and associated files to the Deep Blue spreadsheet and XML manifest.

At this point, you will then be returned to the Main Item Options screen. From here, you may create a new Item, review Item that have already been created, finalize the entire deposit (as described in section 4.9.13), or quit and resume at a later date:



#### 4.9.12 Review an Item

You will access the “Item Review” screen in one of two ways:

1. You have resumed the “Add Descriptive Metadata” procedure after previously quitting it.
2. You select option *R* (“Review Item created for this deposit”) from the Main Item Options screen.

In either case, this screen will display each of the Item you have created for the deposit, listing an Item ID (generated for internal use by AutoPro), title, description, and status (either “completed” or “NOT completed”). Please note that **ALL Items must be formally completed** for AutoPro to correctly generate metadata records.

```

*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.8 (c) 2013
*****

ITEM REVIEW

ITEM ID no.: 1
TITLE: Meeting Materials - State Republican Party Executive Committee
DESCRIPTION: Meeting materials collected by Samuel Smith during his service
on the Michigan state Republican Party's Executive Board
STATUS: Packaging complete

PLEASE NOTE:
1. You may not add files to an Item once it has been finalized.
2. Be sure to "COMPLETE" the packaging of each Item.

Options:

-A: ADD content to an existing Item
-C: COMPLETE packaging of a Item
-U: VIEW contents of Item

-N: create a NEW Item
-F: FINALIZE packaging for entire deposit
-Q: QUIT and resume packaging content at a later date

(A / C / U / F / N / Q)

```

There are two main groups of options on this screen. First, you may perform actions on a particular Item: ADD additional content to an existing Item (if **not** complete), VIEW the contents of an Item, or COMPLETE the packaging and description of an Item. When you select any of these options, AutoPro will ask for the Item ID number:

```

At the prompt, enter the appropriate Item ID number; do not include any
extra spaces.
1_

```

- If you elect to add content, you will be taken to the “Add Content” option described in section 4.9.6 of the User’s Manual.
- If you elect to view content, AutoPro will display the filename and description for all content associated with the Item before returning you to the “Item Review” screen.

```

*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.8 (c) 2013
*****

VIEW FILES ASSOCIATED WITH ITEM
ITEM: Meeting Materials - State Republican Party Executive Committee
FILENAME: january-2011.zip
DESCRIPTION: Michigan Republican Party executive committee meeting materials
from January 2012
FILENAME: June-2011.zip
DESCRIPTION: MEETING MATERIALS FROM jUNE 2011
Press any key to continue . . . _

```

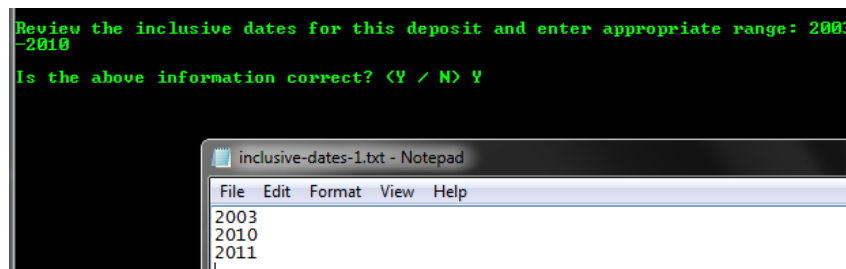
- If you elect to complete packaging and description of the Item, AutoPro will perform the steps described in section 4.9.11 of the User's Manual.

The second group of options available on the "Review Item" screen are ones we have already seen: you may create a NEW Item, FINALIZE the deposit (as described in section 4.9.13), or QUIT and resume at a later date.

#### ***4.9.13 Finalize the Deposit***

When you have created all requisite Items, provided metadata, and associated and described all files with the appropriate Items, you may finalize the deposit.

If there are multiple files and/or items in the deposit, AutoPro will prompt you to review a list of original creation/use dates that you have entered throughout the procedure. Based on this list, enter an appropriate date (or range) for when all the materials in the deposit were originally created and/or used:



AutoPro will then calculate the overall size of the deposit, determine the date range in which files were last modified, and complete the Deep Blue metadata spreadsheet, Dublin Core XML manifest, and the EAD administrative information. You will now be ready to proceed to the next procedure.

### **4.10 Transfer to Long-Term Storage**

This and the following workflow steps will only be completed by Mike, Nancy, or other staff with access to the Deep Blue deposit folder and the BHL Dark Archive. Be sure that you are connected to the appropriate repositories and logged in to the UMROOT domain with your Windows AD password before beginning!

#### ***4.10.1 Transfer to Deep Blue Deposit Folder***

Upon initiation of this step, AutoPro will ask if the materials will be deposited in Deep Blue so that the general public may have access.

```
*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.8 (c) 2013
*****

Initiated: 10. TRANSFER to LONG-TERM STORAGE

*****

This procedure will transfer content to the BHL Dark Archive and, if
necessary, to the \deepblue_deposits\ folder in preparation for deposit to
DSpace.

Will content be deposited in Deep Blue? (Y / N)
```

If the content is restricted in any way, respond with *N* and proceed to the next step.

If the content is unrestricted and permitted to be deposited into Deep Blue, you will be asked to indicate the drive letter on your work station that is mapped to the BHL Interim Repository.

```
*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.8 (c) 2013
*****

Initiated: 10. TRANSFER to LONG-TERM STORAGE

*****

This procedure will transfer content to the BHL Dark Archive and, if
necessary, to the \deepblue_deposits\ folder in preparation for deposit to
DSpace.

Will content be deposited in Deep Blue? (Y / N) y

At the prompt, enter the drive letter where the "deepblue_deposits" folder
is located. Do NOT enter a colon.

Options:
  -QUIT: QUIT and resume transfer at later time
  -Drive letter (i.e. X, Y, or Z, etc.)

<QUIT / Drive letter> Z
```

You will have an opportunity to correct your entry before AutoPro copies the files to the deposit folder. If any errors occur, you will be alerted so that you can retry the operation. Consult with Mike if any issues arise.

#### **4.10.2 Transfer to Dark Archives**

To transfer content to the BHL Dark Archives, you will need to indicate the drive letter to which it is mapped on your work station and AutoPro will request verification.

```
*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.8 (c) 2013
*****

TRANSFER TO DARK ARCHIVE

NOTE: This procedure may only be completed by staff with full access to the
      bhl-root and bhl-archive storage locations.

At the prompt, enter the drive letter to which "bhl-archive" (the BHL Dark
Archive) has been mapped (i.e. Z or Y). Do NOT include a colon.

Options:
  -QUIT: QUIT and resume transfer at another date
  -Drive letter (i.e. X, Y, or Z, etc.)

<QUIT / Drive letter> Y_
```

The material (and metadata) will be packed into “Bags” with the Library of Congress BagIt tool and then copied to the appropriate deposit directory in the BHL Dark Archives. If any errors occur, you will be alerted so that you can retry the operation. Consult with Mike if any issues arise.

#### 4.11 Clean Up

In the final step of the digital processing workflow, AutoPro will delete the processing directory, the backup restore point, and other temporary files. If content has only been transferred to the BHL Dark Archives, do not perform this operation until content has been captured by the “snapshot” functionality of Mainstream storage.

```
*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.8 (c) 2013
*****

Initiated: 11. CLEAN UP

*****

This procedure will clean the processing directory after content has been
successfully transferred to Deep Blue and/or the BHL Dark Archive.

BE SURE TO CLOSE ANY WINDOWS EXPLORER WINDOWS OR APPLICATIONS THAT MAY BE
USING FILES IN THE TARGETED FOLDERS.

The procedure will package log files and delete the following directories:
  - C:\BHL\unprocessed\9994_0001
  - C:\BHL\tmp\9994_0001
  - C:\BHL\backup\9994_0001
  - C:\BHL\logs\9994_0001

Delete Separations directory, too? <Y / N> Y
```

AutoPro will first check to see if the Separations should be deleted as well; check the transfer agreement if you are unsure of the disposition of separated material. After you make your selection, AutoPro will package the log files so that a copy may be deposited in the Bentley’s IFS space for quick reference.



```
*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.7 (c) 2012
*****

TRANSFER LOG FILES
Upload C:\BHL\logs\87206_0001\0001.zip to BHL IFS space:
--> ../bhl/Private/html/processing_log_files/0001/
NOTE: Consult with Mike if problems arise.

Has 0001.zip been successfully uploaded? (Y / Q to QUIT) _
```

Once you have indicated that the .ZIP file of logs has been uploaded, AutoPro will request a final verification that you are ready to delete the working copies of materials.

```
*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.7 (c) 2012
*****

CLEAN UP
- C:\BHL\unprocessed\87206_0001
- C:\BHL\temp\87206_0001
- C:\BHL\backup\87206_0001
- C:\BHL\logs\87206_0001
- C:\BHL\separations\87206_0001

Options:
-D: DELETE folders
-Q: QUIT procedure and perform clean up at a later date
(D / Q) _
```

If you prefer to wait, you may quit the program and return at a later time. Following the deletion of content, the digital processing workflow is concluded; you may press any key and AutoPro will close.

```
*****
Bentley Historical Library * * * Digital Curation Division
AutomatedProcessor v. 0.7 (c) 2012
*****

You have completed processing deposit 87206_0001 with AutoPro, the
Bentley Historical Library's Automated Processor

Good bye!

*****

Press any key to continue . . . _
```

## 5. Version History

<b>Version</b>	<b>Date</b>	<b>Reviewed By</b>	<b>Emendations</b>
02.0	2013-04-09	Michael Shallcross	Major revision of AutoPro and procedures. Removed name normalization step and rewrote all scripts to accommodate irregular file names. Simplified metadata entry and description via command prompt interface.
01.1	2012-11-06	Michael Shallcross	Added extraction of disk images to archive extraction procedures; metadata entry via text files.
01.0	2012-10-29	Michael Shallcross	Final draft used to train graduate students and staff.