

June 19, 2014

Digital Curation Plan Saint Repositorius Council

Piper Chapman test case

Ariadne Rehbein

Table of Contents

Introduction	2
Overview of the files	3
File characteristics.....	4
File formats	6
Metadata	7
Intellectual Property.....	13
Required resources.....	13
Project Management	14
Budget	16
Stakeholders & Responsibilities	16
Institutional Context	17
Appendix	17

Introduction

As the new digital curator at the Saint Repositorius Council (SRC), I am excited to serve as the steward of our growing digital manuscripts collection. Digitized versions of our manuscripts will be preserved and presented online, serving a broad range of visitors. Medieval scholars and the general public alike will have access to the treasures of Saint Repositorius Council, some of which are too fragile to handle in person. With the help of a National Endowment for the Humanities Preservation Grant for Small Institutions, in May 2014, the SRC completed work on its digitization lab. The Saint Repositorious Council is also exploring the option of developing its collection for digital humanities scholarship and applying for another NEH grant.

The head of digitization recently invited Dr. Chapman to a personal digitization tutorial. Dr. Chapman is a local medieval literature scholar at Indiana University and a frequent visitor and supporter of the Saint Repositorius Council. He has published numerous books and articles based on his study of our manuscripts. This scholarly attention was vital in attaining our National Endowment for the Humanities Preservation Grant for Small Institutions. When we begin the digitization process, the manuscripts curator will work closely with Dr. Chapman and other scholars to select manuscripts for digitization. Scholars will add value to the items by contributing metadata, such as descriptions of particular items and their relationship to scholarly developments through bibliographic references.

In this document I lay out an approach to Dr. Chapman's files in a way that demonstrates important curation policies. There is one main difference in the workflow of these files – though they were created using our equipment, they were sent back to us, instead of stored in our repository. Because we had not yet set up our storage, we did not deposit these files, but rather provided them to Dr. Chapman on an external hard drive. Alongside my recommendations for Dr. Chapman's files I explain where the workflow of this test case differs from the standard workflow and demonstrate how digital curation policies for the Saint Repositorious Council are essential to building, maintaining, and providing access to our digitally preserved illuminated manuscript collection.

Overview of the files

Dr. Chapman provided 13 digital files, including 12 images and one metadata file. Among the image files, Dr. Chapman included two other files that do not pertain to Saint Repositorious collection. They were not created using items from our collection. Additionally, since these items are also not described in the metadata, it may be presumed that Dr. Chapman did not intend to send these two files. The issue of receiving extraneous files from scholars will not be encountered in the future since our submission policy will limit submissions to digitized manuscript files that Saint Repositorious staff have saved directly on our institutional network.

The title, object number, and group from Dr. Chapman's metadata file have been noted in the following chart to demonstrate how each item matches a manuscript in our collection.

Chart of basic details for images files sent by Dr. Chapman. The information for files that should not be included in the final collection have been crossed out.

File name	Physical object Type	Title	Alternative Title	Object number
00031501	etching			
00482001	manuscripts	A Battle at Bordeaux		83.MP.150.163v
00543901	manuscripts	A Blackbird	Bestiary (Group title)	83.MR.173.32v
00546701	manuscripts	A Bird on a Dead Wolf	Bestiary (Group title)	83.MR.173.81v
00550501	manuscripts	A Blackbird	De Natura Avium; De Pastoribus et Ovibus; Bestiarium; Mirabilia Mundi; Philosophia Mundi; On the Soul (Group Title)	83.MR.174.40v
00593601	manuscripts	A Battle from the Trojan War	Leaf from Ancient History up to the Reign of Caesar (translated)	83.MP.146.3.recto
06537001	photograph			
30590401	manuscripts	A Beaver	Northumberland Bestiary (Group Title)	2007.16.11v
30592601	manuscripts	A Bear	Northumberland Bestiary	2007.16.27

			(Group Title)	
30596101	manuscripts	A Basilisk	Northumberland Bestiary (Group Title)	2007.16.54v
30596201	manuscripts	A Basilisk; Snakes; Vipers	Northumberland Bestiary (Group Title)	2007.16.55
33394501	manuscripts	Abbey Bible		2011.23

File characteristics

The 10 files relevant to our collection are of excellent quality and may serve as master files. They meet the standards of the Yale University Digitization Shared Practices (2010). For master graphic files, these standards include:

- spatial resolution of 400 – 8000 ppi
- spatial dimensions of 4000 – 8000 pixels across the long dimension (an output size that is 100% of the original material's size, excluding mounts and borders)
- bit depth between 24 – 48 bit color

The head of digitization, Julia Applebloom demonstrated the imaging workflow she has established to Dr. Chapman in creating these files with him. Her scanner calibration and imaging workflow is based on the Technical Guidelines for Digitizing Cultural Heritage Materials: Creation of Raster Image Master Files from the Federal Agencies Digitization Initiative Still Image Working Group (August 2010) (http://www.digitizationguidelines.gov/guidelines/FADGI_Still_Image-Tech_Guidelines_2010-08-24.pdf). As part of the imaging workflow, the digitization head adjusted the scaling, and spatial resolution of the scanner. She corrected the color and adjusted the tone and saturation of the scanner with both the help of aimpoint guidance and through comparing the initial scan to the original item. Lastly, she converted the color profile to Adobe RGB 1998.

JHOVE 2 was used to discover the technical characteristics and format of these objects and therefore determine their acceptability for deposit based on curation policies.

Chart of image file characteristics

File Name	Resolution	Width	Height	Bit depth	Size (bytes)	Compression	Color profile
00482001	800	4572	7047	24	36,915,038	no	Adobe RGB 1998
00543901	800	4244	5892	24	27,202,435	no	Adobe RGB 1998
00546701	800	4438	5951	24	29,682,979	no	Adobe RGB 1998
00550501	800	3093	4520	24	16,909,514	no	Adobe RGB 1998
00593601	800	3273	4271	24	17,323,759	no	Adobe RGB 1998
30590401	800	3219	4340	24	20,651,612	no	Adobe RGB 1998
30592601	800	3219	4340	24	16,168,738	no	Adobe RGB 1998
30596101	800	3219	4340	24	16,225,898	no	Adobe RGB 1998
30596201	800	3219	4340	24	17,190,105	no	Adobe RGB 1998
33394501	800	3129	4468	24	17,201,097	no	Adobe RGB 1998

File formats

All of the image files are currently stored in JPEG format and should be converted into TIFF format to serve as master copies. The original JPEG files will be retained, however, in case a future format migration is required. TIFF is the de-facto standard for master files and is known as sustainable format for several reasons. The most relevant benefits of TIFF over JPEG are lossless compression, the ability to accommodate large file sizes, and self-documenting metadata.

JPEG is also a very widely adopted format, but its major drawback is lossy compression. This means that when a copy or of a JPEG file is saved, some data is lost. According to the Federal Agencies Digitization Initiative Still Image Working Group, JPEG image quality degrades after about three saves. TIFF file derivatives have lossless compression, meaning that copies do not lose data. This feature is important because digital curator will need to refresh copies of the master file, check over file size with checksums, and create derivatives of the master file for access. The encoding of TIFF files is also simple to read and any changes to a file are automatically embedded in the format.

In a typical workflow, image files will be created in TIFF format to serve as master files. Two JPEG files (a small thumbnail file and a file for display in the manuscripts viewer application) will be displayed on our digital collections website. Additionally, a TIFF image file will be available for download. Both TIFF and JPEG files should be available due to the variety of purposes they may serve. Scholars may wish to download them for close analysis and comparison outside the web browser. The general public may wish to use them for a variety of purposes, such as art or personal use. In the case of the Chapman digitization tutorial, JPEG files were created by mistake, but it will not degrade the files in any way to convert them into TIFF as master files, create an access copy TIFF file and create two access copy JPEG files. In accordance with our digital curation policy, we will use ImageMagick within Archivematica to migrate the file format of access copies with efficiency and to reduce human error.

TIFF and JPEG are widely used and well-documented formats¹. They have no particular software or operating system dependencies. The need for format migration is not anticipated during the lifecycle of files in these formats. If it is the case that new image formats become common and TIFF and JPEG are at risk of being obsolete, we will perform migration using the original JPEG files unless this will result in lossy compression. We will make every effort to automate the migration process, select non-proprietary and well-documented file format standards, and document the migration process in the administrative metadata of each file.

¹ TIFF sustainability factors from the Library of Congress:

<http://www.digitalpreservation.gov/formats/fdd/fdd00022.shtml#sustainability>

JPEG sustainability factors from the Library of Congress:

<http://www.digitalpreservation.gov/formats/fdd/fdd00017.shtml#sustainability>

Master copy and access image copies of Chapman files:

Stored in Fedora:

Original copy JPEG, stored in case of future migration
--

Master copy TIFF 800 ppi

Access copies:

JPEG for web viewer and download, 1800 pixels on longest side

TIFF for download 800 ppi

JPEG for thumbnail 190 pixels on longest side
--

Metadata

Metadata is essential for accessing, locating, understanding, and recording interventions on our digitized manuscripts. For his files, Dr. Chapman provided an XML file describing the illuminated manuscripts that were digitized along with an example identifier and two URLs. To meet the submission policies established by the advisory group, this data will be transformed into a widely used standard, Dublin Core.

The group debated the pros and cons of joining the Digital Scriptorium which has its own descriptive standard based on MODS. However, by joining the Digital Scriptorium, we would commit the curation and provision of access to the California Digital Library and UC Berkeley, respectively. Since the Saint Repositorious Council is interested in developing its collection for digital humanities scholarship and applying for another NEH grant we did not want to lock ourselves in to such an agreement. Other projects that have caught the interest of the Saint Repositorious Council are the DM manuscript annotation tool, currently in development at Drew University², and TEI encoding of manuscript descriptions at the Walters Art Museum³. Outside of Digital Scriptorium members, the advisory group could find no institutions with digital illuminated manuscripts using the DS standard, including major institutions like the Bodleian Library, the British Library, and the National Library of the Netherlands (Koninklijke Bibliotheek).

We decided to follow the example of the National Library of the Netherlands, which uses qualified Dublin Core to catalog its objects on a manuscript and illumination level.⁴

² <http://dm.drew.edu/dmproject>

³ <http://www.thedigitalwalters.org/>

⁴ <http://www.kb.nl/en/data-services-apis/medieval-illuminated-manuscripts>

As can be seen in this example record on the NLN Medieval Illuminated Manuscripts website, the NLN has created new field names to suit the unique needs of illuminated manuscripts:

Highlights... Browse... Search... **Medieval Illuminated Manuscripts** Introduction Background Home

Result 1 of 1

Full reference manuscripts [To report an error in the manuscript description or with the images, please click here.](#)
The Hague, KB, 66 B 13

Contents: Valerius Maximus, Des faits et dits mémorables. Translation from the Latin by Simon de Hesdin and Nicholas de Gonesse

Place of origin, date: Loire valley, Master of the Cité de Dieu of Mécon, Master of the Psalter of Jean le Meingre III, Master of the Echevinage de Rouen (illuminators); c. 1470

Material: Vellum, ff. 485, 390x285 (238x167) mm, 42 lines, littera cursiva, Binding: 18th-century brown leather; blind

Decoration: 1 full-page miniature (miniature + margin; 345x255 mm); 6 full-page miniatures (miniature + margin including some lines of text; 345/335x255/240 mm); 8 two-column miniature (220/140x170/165 mm); 7 column miniatures (83/47x74 mm); decorated initials throughout; coloured cadels (ff. 79v, 82v, 174v, 189r, 244r, etc.); penwork initials with pen-flourishes throughout

Provenance: purchased in 1814 in Paris by Tsar Alexander I of Russia and placed in the Imperial Library at St. Petersburg (shelfmark Fr.F.V.IV.2); sold abroad by the Russian government in the 1920s. Purchased by F. Mannheimer of Amsterdam (d. 1939); confiscated during World War II by the German occupying forces and restituted after 1945; since 1953 on permanent loan from the Instituut Collectie Nederland (ICN), Amsterdam (formerly: Dienst voor 's-Rijks Verspreide Kunstvoorwerpen, The Hague, and Rijksdienst voor Beeldende Kunst, The Hague)

Annotation: full-page miniature missing before f. 418 (Opening of Book 9)

Bibliography: Literature about this manuscript

 [Image for this manuscript](#) | [images only](#) | [images and description](#) | [description only](#)

The advisory group analyzed the mapping of Dublin Core terms to NLN terms. The SRC has directly adopted the following terms that NLN uses for manuscripts. Users will see these terms when accessing our digitized collections online (Discovery Information Packages.) These transformations are mapped to Dublin Core, which we use to encode all of our descriptive metadata, stored in Archival Information Packages. Terms are required unless otherwise noted:

NLN/ SRC field name	Dublin Core term
PURL	identifier
Manuscript identifier (The Hague, KB, 66 B13)	identifier
Contents	title
Place of origin	source
Date	date
Material	format
Decoration	format
Provenance (not required)	description
Annotation (not required)	Description
Bibliography (not required)	relation

The SRC descriptive metadata will include the following additional terms to describe manuscripts. Terms are required unless otherwise noted:

SRC field name	Dublin Core term
Language	language
Contributor	contributor
Type	type

The NLN catalogs on an illumination level, often providing several images of illuminations from each folio page. The SRC will catalog on a manuscript and folio level, providing one image of each folio which includes all illuminations on that page.



Example of illumination records on the NLN Medieval Illuminated Manuscripts website:

Highlights... Browse... Search... **Medieval Illuminated Manuscripts** Introduction Background Home

Full reference manuscripts
The Hague, KB, 66 B 13 (1-5 of 77) for (isPartOf all *66 B 13*)



To order a digital scan of these images, click here.

1. The Hague, KB, 66 B 13 fol. 2r



Book 1: Chapter 1
Simon de Hesdin presents his translation of the *Facta et dicta memorabilia* by Valerius Maximus to Charles V, king of France

2. The Hague, KB, 66 B 13 fol. 2r



Book 1: Chapter 1

3. The Hague, KB, 66 B 13 fol. 2r



Book 1: Chapter 1, 2

4. The Hague, KB, 66 B 13 fol. 2r

Book 1: Chapter 1

5. The Hague, KB, 66 B 13 fol. 2r

Book 1: Chapter 1
Numa Pompilius, king of

Fol. 2r: miniature
2r
345X255
iconclass
31A2524
44B113
44B1211
44B1322
44B152
48A12
48C91(-0)
49L72
61B2(CHARLES V [of France])21
61B2(HESDIN, Simon de)21
83(VALERIUS MAXIMUS, *Facta et dicta memorabilia*)

Ten sons of leading Romans are sent to Etruria to be instructed in religious ceremonial; the Etrurians priests are represented as bishops
Fol. 2r: miniature
2r
345X255
iconclass
11P3113
12F219
12F51
25H112
25I146
32B311(ETRURIANS)
32B311(ROMANS)
33A1
49A
83(VALERIUS MAXIMUS, *Facta et dicta memorabilia*, 1.1.1)
98A(-9)

Pontifex Maximus Metellus forbids Consul and Flamen of Mars Postumius, who wanted to leave Rome for a military campaign in Africa, to abandon his religious duties
Fol. 2r: miniature
2r
345X255
iconclass
12F51(FLAMEN)
12F63
44B1211
83(VALERIUS MAXIMUS, *Facta et dicta memorabilia*, 1.1.2)
92B47
98B(METELLUS, L.C.)5
98B(POSTUMIUS ALBINUS, A.)5

The occupation of Rome by the Gauls and C. Fabius Dorsuo's sacrifice to gods: three corpses, a man holding a lamb
Book 1: Chapter 1
2r
345X255
iconclass
12F2121
31E3
32B2(GAULS)
47I2133
61E(ROME)
83(VALERIUS MAXIMUS, *Facta et dicta memorabilia*, 1.1.11)
98B(FABIUS DORSUO, C.)51(-0)

Numa Pompilius, king of Rome, compels the people to observe religious ordinances
Fol. 2r: miniature
2r
345X255
iconclass
31A233
44B113
44B1211
44B1322
83(VALERIUS MAXIMUS, *Facta et dicta memorabilia*, 1.1.12)
98B(NUMA POMPILIUS)5

© Koninklijke Bibliotheek National Library of the Netherlands

SRC will directly adopt the following terms that NLN uses for folios instead. isPartOf can be conveyed through structural data. Using this Dublin Core qualifier prevents OAI-PMH⁵harvesting. Terms are required unless otherwise noted:

NLN/ SRC field name	Dublin Core term
PURL	identifier
Folioidentifier (The Hague, KB, 66 B13 fol 2r)	identifier
Dimensions	format
Text summary (not required)	description
Image subject	subject (using IconoClass classification)

The SRC descriptive metadata will also include the following terms to describe folios. Terms are required unless otherwise noted:

SRC field name	Dublin Core term
Type	type
Title (not required)	title
Date (not required if manuscript has same info or if unknown)	date
Creator (not required if manuscript has same info or if unknown)	creator
Place of origin not required if manuscript has same info or if unknown)	source

Dr. Chapman provided folio-level records for his images. His metadata maps to the following SRC descriptive metadata terms.

Chapman	Example from Chapman records	SRC field name (Dublin Core term)
recordID	info:getty/object/4820	PURL (dc:identifier)
PrimaryTitle	A Battle in Bordeaux	title
MakerName	Master of the Copenhagen Caesar (illuminator)...	creator
Type	Manuscript	n/a – serves same purpose as objectname, but is less specific
Medium	Tempera colors, gold leaf, gold paint, and ink on parchment	material (dc:format)
Place	Place Created: Bruges, Belgium, Europe	place of origin (dc:source)
Date	about 1480 - 1483	date

⁵ <http://www.openarchives.org/OAI/2.0/guidelines-repository.htm#MinimalImplementation-DC>

Source	J. Paul Getty Museum	n/a - captured in administrative metadata
ObjectNumber	83.MP.150.163v	identifier
ObjectName	Manuscript, Folio, detached leaf	type
AlternateTitle	Chroniques (Book Three) (Group Title)	n/a – relationship of a folio to the manuscript that contains it is represented through structural metadata. There will be a link to the appropriate DOI
Department	Manuscripts	n/a - SRC is not divided into departments
Dimensions	Leaf: 48 x 35 cm (18 7/8 x 13 3/4 in.) Justification: 29 x 19.3 cm (11 7/16 x 7 5/8 in.)	Dimensions (dc:format)
Culture	Flemish	n/a - same as source
AlternateNumber	Ms. Ludwig XIII 7, fol. 163v (Manuscript number)	Folio Identifier (dc:identifier)
Topic	How We Live / War / Battles	Image Subject (dc:subject)
imageThumbURI	http://www.getty.edu/art/collections/images/thumb/00482001-T.JPG	n/a, this just an example provided by Chapman of the permanent identifier for the thumbnail. This is administrative metadata that will be linked within METS
recordLink	http://search.getty.edu/museum/records/musobject?objectid=4820	n/a, this just an example provided by Chapman of the proposed URL for the online record.

In addition to descriptive metadata, administrative and structural data will need to be created and associated with the digital files. METS, the metadata encoding and transmission standard was designed to be a container that links digital objects with various types of metadata. METS will contain PREMIS metadata be used to encode administrative metadata such as technical information about the files, preservation actions, and rights information.

By transferring the files into the Archivematica dashboard, the digital curator will be able to extract technical and preservation information about the files and have them automatically encoded into METS. Any file format migrations will be recorded in using PREMIS within the METS document. She will describe each of the items in Dublin Core and Archivematica will automatically nest this descriptive information in a METS document that indicates the file structure. When preparing the files for submission to storage, she will indicate additional administrative information and may describe the SIP as a whole with Dublin Core.

Please see the addendum to view an example metadata record.

PREMIS metadata entry screen in Archivematica:

The screenshot shows a web browser window with two tabs: 'Archivematica Dashboard' and 'ElasticSearch Head'. The address bar shows the URL: `localhost/ingest/b98313ad-cf1d-4ad3-bd1f-1bff13ac479c/rights/add/`. The browser's address bar also shows the following tabs: 'Archivematica', 'ICA-AtoM', and 'Elasticsearch B...'. The Archivematica navigation bar includes the following items: 'Transfer', 'Ingest' (with a red notification badge), 'Archival storage', 'Preservation planning', 'Access', 'Administration', and 'demo'. The breadcrumb trail is 'Ingest / Tutorial Office Docs / Rights / Add'. The main heading is 'Rights' with the subtitle 'Tutorial Office Docs'. The 'Basis' dropdown menu is set to 'Copyright'. The 'Copyright status' field contains 'Under copyright'. The 'Copyright jurisdiction' field contains 'Canada'. The 'Copyright determination date' field contains '2009/09/16' with a note 'Use ISO 8061 (YYYY-MM-DD)'. The 'Copyright start date' field contains '2004/08/17' with a note 'Use ISO 8061 (YYYY-MM-DD)'. The 'Copyright end date' field is empty with a note 'Use ISO 8061 (YYYY-MM-DD)'. The 'Open End Date' checkbox is checked. The 'Copyright documentation identifier' section includes: 'Type' (Donor form), 'Value' (CCA-2009-67), and 'Role' (Copyright holder statement). The 'Copyright note' field contains 'Copyright held by donor'. At the bottom, there are three buttons: 'Save', 'Next', and 'Cancel'.

Intellectual Property

The Saint Repositorious Council retains the copyright to all images and descriptions available through the digital collections website. They are licensed for use under the the Creative Commons Attribution-Share Alike 3.0 Unported License and the GNU Free Documentation License. This means that users are free to download, share, and adapt the images and descriptions on this website as long as they cite the Saint Repositorious Council as their source.

- Sample citation for an image:

W.582, fol 14r, © 2014 Saint Repositorious Council, used under a Creative Commons Attribution-ShareAlike 3.0 license: <http://creativecommons.org/licenses/by-sa/3.0/>

- Sample citation for a description:

Saint Repositorious Council [item name] online description [URL optional] © 2014 Saint Repositorious Council, used under a Creative Commons Attribution-ShareAlike 3.0 license: <http://creativecommons.org/licenses/by-sa/3.0/>

Required resources

Saint Repositorious will use the Archivematica suite of tools to manage its curation workflow. For storage, we would like to join a Private LOCKSS Network and run our node on our current Raid 5 level storage. Archivematica is an open-source suite of curation tools that are combined into a user-friendly dashboard. It facilitates the process of file characterization, format migration, and metadata creation. The digital curator will work with a programming consultant to set up Archivematica and will contact him in case of problems or questions. Through striping and parity across multiple drives, RAID 5 level storage reconstructs a failed physical drive onto a hot spare. Hardware failures are inevitable, so this method prevents the loss of data and user access to data.

Joining a LOCKSS networks would allow us to have redundant copies across a geographically dispersed network with automatic bit stream fixity checks. This is important because catastrophic events, such as natural disasters, may cause RAID level 5 storage to fail. LOCKSS participation would also automate our format migration. It would be a wonderful opportunity to collaborate with other manuscripts repositories working to digitize their collections. We would share the responsibilities of building policies for sustainability and increase the safety of all of our assets as we grow our collections.

Curation of this collection will only involve the work of the digital curator. As is the case with the standard workflow, she will use Archivematica to perform the necessary analyses and migrations of the files and create metadata through the simple user interface of the dashboard. After the completion of assessing the files and translating Dr. Chapman's metadata into Dublin Core (2

hours), creating the metadata records, migrating the file formats, doing another checksum, and uploading the AIP and DIP will take an additional 5 hours.

Project Management

May 19, 2014 – **Receive Submission Information Package**

- copy Dr. Chapman's files onto a staging server for four weeks
- check for viruses, malware, and completeness
- generate checksums

June 16, 2014 – **Assess files**

- transfer into Archivematica in order to generate checksums, identify file forms, and re-check for viruses
- sort files and check for completeness. Remove extraneous files and duplicates
- assess quality and suitability according to collection policies

June 19, 2014 – **Approval of Saint Repositorious administration to include files in online digital collection**

June 23, 2014 – **Produce Archival Information Package**

- verify inventory
- migrate file format to TIFF
- transform existing descriptive metadata into Dublin Core
- input descriptive metadata for each file and for SIP as a whole
- generate technical metadata and preservation action metadata for each file
- input additional administrative metadata, such as rights information
- assign persistent identifiers to files
- Archivematica generates access copies of files and package them into a DIP (Dissemination Information Package that contains all digital files and a METS xml document which contains the exact same metadata as the METS xml document in the AIP)
- generate and compare a new checksum

June 24, 2014 – **Storage and Access**

- send Archival Information Packages to storage
- upload the DIP to the access system (https://www.archivematica.org/wiki/UM_access_1.1)

- contact Dr. Chapman to let him know of ingest completion. Users have access to Dr. Chapman's digitized manuscripts and descriptive metadata via the Saint Repositorious Council online collection

Preservation to ensure digital objects retain integrity and remain authentic, reliable, and usable

Every 5 years:

Refresh files:

- Move data to a new version of the same storage medium or a different storage medium
- Generate checksum of bit stream using JHOVE
- Document the processes applied to refresh data in administrative metadata

As needed:

Migrate format (estimate 5 – 10 years)

- Keep up with the status of software applications needed to access files
- Large scale migration will use ImageMagick or similar software
- Carefully document process of migration to ensure integrity and authenticity of migrated data
- Use JHOVE2 or similar software to identify, characterize, and validate digital objects

As needed:

Update descriptive metadata to link items to new scholarly research and other items within and outside our collection

Budget

Set up of curation environment:

Item	cost
Purchase of IBM Storwize V3700 SFF Dual Control Enclosure	\$8,7999
Private LOCKSS network participation (cost of private box for small baccalaureate college divided by 4)	\$500?
Programmer consultation	\$3,000

Dr. Chapman file curation costs:

Item	cost
Digitization (Digitization Head 1 hour pay)	\$40
Ingest (Digital Curator 7 hours pay)	\$175
Storage of files for 1 year	<\$0.39
User access for 1 year (divided by total size of collections)	\$5
Monitor and refresh/ migrate/ update metadata as needed for 1 year (digital curator salary, 1 collection/ 500 collections total)	\$80

Stakeholders & Responsibilities

- Users of the digital collection, ranging in scholarly expertise and purpose in using the collection
- Digital collections related staff SRC
 - Digitization head – oversee digitization process and create policies
 - Manuscripts curator – select manuscripts for digitization in collaboration with scholars (Current curator has served in this role for 40 years.)
 - Digital curator – oversee and develop policies for stewardship of digitized manuscripts and associated metadata
 - Programming consultant – perform initial set up of archivematica and storage with current SRC servers. Remain on call for bugs that digital curator cannot resolve after troubleshooting
- Administrative staff at the SRC – oversee the SRC as a whole. Digital collections are one new aspect of its mission as an ecumenical group. Many administrators on the Council see the digitized collection aligning with their mission of education and public good, but also see this as an outreach activity that demonstrates the value of the Council and may draw in more funds.

Institutional Context

The Saint Repositorious Council is exploring its options for curation and its role within the larger context of manuscripts repositories and discipline of medieval history. Since the SRC is currently considering applying for an NEH Grant for digital humanities research, it may gain several hundred thousand more dollars to support curation and scholarship, expand the scope of its project and join a LOCKSS network shared with similar repositories. The current suggestions of Archivematica and IBM Storwize Raid level 5 storage match current SRC financial resources to the basic curation needs.

Appendix

Metadata record

For descriptive metadata, Dublin Core mapping “comments” have been provided to show the terms that will appear to users of the digital collections website. This example record does not include a PREMIS record which will store preservation metadata. These records will be automatically generated through Archivematica. The diviprovMD information is a close equivalent. METS encoding examples were copied and modified from the Library of Congress METS Primer and Reference Manual: <http://www.loc.gov/standards/mets/METSPrimerRevised.pdf>

```
<mets:mets xmlns:lc="http://www.loc.gov/mets/profiles/"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:mods="http://www.loc.gov/mods/v3" xmlns:xlink="http://www.w3.org/1999/xlink"
xmlns:mets="http://www.loc.gov/METS/"
xmlns:photo="http://www.loc.gov/mets/profiles/photoObject"
xmlns:premis="info:lc/xmlns/premis-v2" PROFILE="lc:photoObject"
OBJID="loc.natlib.gottlieb.09601" xsi:schemaLocation="http://www.loc.gov/METS/
http://www.loc.gov/standards/mets/mets.xsd">
```

<!--Descriptive metadata -- There may be one dmdSec for the AIP as a whole. Each original file may also have a dmdSec. In this example, I am just describing one folio/object -->

```
<mets:dmdSec ID="DMD1">
  <mets:mdWrap MIMETYPE="text/xml" MDTYPE="Dublincore">
    <mets:xmlData>
      <dc:title>A Basilisk</dc:title>
```

```

        <dc:identifier>305961</dc:identifier><!--srcdc:identifier-->
        <dc:format>Pen-and-ink drawings tinted with body color
and translucent washes on parchment</dc:format><!--srcdc:material-->
        <dc:source>Place Created: England,
Europe</dc:source><!--srcdc: place of origin-->
        <dc:date>about 1250 - 1260</dc:date>
        <dc:type>Folio</dc:type>
        <dc:format>Leaf: 21 x 15.7 cm (8 1/4 x 6 3/16
in.)</dc:format><!--srcdc:dimensions-->
        <Culture>English</Culture>
        <dc:identifier>Ms. 100, fol. 54v (Manuscript number)
&lt;br&gt; L.2007.72.54v (Previous number)</dc:identifier><!--srcdc: folio identifier-->
        <dc:subject>Mythology</dc:subject><!--
srcdc:imageSubject-->
        <dc:subject>Creatures</srdc:subject><!--
srcdc:imageSubject-->
        <dc:subject>Basilisks</dc:subject><!--srcdc:imageSubject-
-></dc:subject>
        <dc:relation>citation</dc:relation><!--srcdc:bibliography-
></dc:subject>
<dc:language>Latin</language>
<dc:contributor>P. Chapman</contributor>
        </mets:xmlData>
    </mets:mdWrap>

</mets:dmdSec>
<amdSec>
    <!-- administrative metadata. There is one amdSec for each object. Each amdSec will
include one techMD, multiple digiprovmDs (preservation metadata), and rightsMDS.-->
    <amdSec>
        <mets:techMD ID="AMD001"> <!--technical information-->
            <mets:mdWrap MIMETYPE="text/xml" MDTYPE="NISOIMG"
LABEL="BasiliskMaster">
                <mets:xmlData>
                    <niso:MIMETYPE>image/tiff</niso:MIMETYPE>
                    <niso:ScanningAgency>Saint Repositorious Council

```

```

        </niso:ScanningAgency>
    </mets:xmlData>
</mets:mdWrap>
</mets:techMD>

<mets:rightsMD ID="ADMRTS1"> <!--rights information-->
    <mets:mdWrap MDTYPE="OTHER"
OTHERMDTYPE="METSRights">
        <mets:xmlData>
            <rts:RightsDeclarationMD
RIGHTSCATEGORY="CCYSA3">
                <rts:Context CONTEXTCLASS="GENERAL
PUBLIC">
                    <rts:Constraints
CONSTRAINTTYPE="RE-USE">
                        <rts:ConstraintDescription>This
volume was published in Great
                                The Saint Repositorious
Council retains the copyright to all images and descriptions available through the digital
collections website. They are licensed for use under the the Creative Commons Attribution-Share
Alike 3.0 Unported License and the GNU Free Documentation License. This means that users
are free to download, share, and adapt the images and descriptions on this website as long as they
cite the Saint Repositorious Council as their source.
                                </rts:ConstraintDescription>
                            </rts:Constraints>
                        </rts:Context>
                    </rts:RightsDeclarationMD>
                </mets:xmlData>
            </mets:mdWrap>
        </mets:rightsMD>

    <!--source information-->
    <sourceMD ID="Ms. 100, fol. 54v (Manuscript number)" >
        <identifier_type>folio
number</identifier_type>

        <source_format_type>folio</source_format_type>
    </sourceMD>

```

```

    <mets:digiprovMD ID="BasiliskMaster">
      <mets:xmlData>
        <mets:mdWrap MDTYPE="OTHER"
OTHERMDTYPE="oclcprov">
          <mets:oclcprov:oclcprov>
            <oclcprov:digiprovMD>
              <oclcprov:fixityAlgorithm>Adler-32
</oclcprov:fixityAlgorithm>

            <oclcprov:fixityCheckResults>202947597</oclcprov:fixityCheckResults>

            <oclcprov:fixityCheckStatus>Success</oclcprov:fixityCheckStatus>
              <oclcprov:fixityCheckDate>2006-
05-15T09:17:49
              </oclcprov:fixityCheckDate>
              <oclcprov:virusSoftware>McAfee
Virus Scan for Linux v.4.40.0
              </oclcprov:virusSoftware>

            <oclcprov:viruCheckStatus>Success</oclcprov:virusCheckStatus>

            <oclcprov:virusCheckDate>2006-05-15T09:17:49</oclcprov:virusCheckDate>

            <oclcprov:mimeType>application/pdf</oclcprov:mimeType>
              </oclcprov:digiprovMD>
            </mets:oclcprov:oclcprov>
          </mets:mdWrap>
        </mets:xmlData>
      </mets:digiprovMD>
    </amdSec>

```

```

<!--File section. By grouping files together, the METS document can be used to manage the files
for display or preservation activities.
-->

```

```

<mets:fileSec>
  <mets:fileGrp USE="Basilisk">
    <mets:file ID="BasiliskMaster" MIMETYPE="image/tiff">
      <mets:FLocat
        xlink:href="http://whhttp://www.saintrepositorriouscouncil.org/tiffMaster/30596101-T.tif"
        LOCTYPE="PURL"/>
      </mets:file>
      <mets:file ID="BasiliskTIFFAccessCopy" MIMETYPE="image/tiff">
        <mets:FLocat
          xlink:href="http://www.http://www.saintrepositorriouscouncil.org/tiffAccess/30596101-
          T.tif" LOCTYPE="PURL"/>
        </mets:file>
        <mets:file ID="BasiliskTrueToSize" MIMETYPE="image/jpeg">
          <mets:FLocat
            xlink:href="http://www.http://www.saintrepositorriouscouncil.org/trueToSize/30596101-
            T.JPG"
              LOCTYPE="URL"/>
          </mets:file>
          <mets:file ID="BasiliskThumbnail" MIMETYPE="image/gif">
            <mets:FLocat
              xlink:href="http://www.saintrepositorriouscouncil.org/thumb/30596101-T.JPG"
                LOCTYPE="PURL"/>
            </mets:file>
          </mets:fileGrp>
        </mets:fileSec>
      </mets:structMap>

<mets:fileSec>
  <mets:fileGrp USE="BasiliskandVipers">
    <mets:file ID="BasiliskAndVipersMaster" MIMETYPE="image/tiff">
      <mets:FLocat

```

```
xlink:href="http://www.getty.edu/art/collections/images/thumb/30596101-T.tif"
LOCTYPE="URL"/>
    </mets:file>
    <mets:file ID="BasiliskandVipersTIFFAccessCopy"
MIMETYPE="image/tiff">
        <mets:FLocat
```

```
xlink:href="http://www.getty.edu/art/collections/images/thumb/30596101-T.tif"
LOCTYPE="URL"/>
    </mets:file>
    <mets:file ID="BasiliskAndVipersTrueToSize"
MIMETYPE="image/jpeg">
        <mets:FLocat
```

```
xlink:href="http://www.getty.edu/art/collections/images/trueToSize/30596101-T.JPG"
    LOCTYPE="URL"/>
    </mets:file>
    <mets:file ID="BasiliskAndVipersThumbnail" MIMETYPE="image/gif">
        <mets:FLocat
```

```
xlink:href="http://www.getty.edu/art/collections/images/thumb/30596101-T.JPG"
    LOCTYPE="URL"/>
    </mets:file>
  </mets:fileGrp>
</mets:fileSec>
</structMap>
<!--Structural map. Indicates the order of the different folio files in the manuscript-->
    <mets:structMap TYPE="physical">
      <mets:div TYPE="manuscript"
LABEL="Northumberland Bestiary ">
        <mets:div TYPE="folio"
LABEL="Basilisk">
          <fptr
FILEID="305961"/></mets:div>
        <mets:div TYPE="folio"
LABEL="BasiliskSnakesAndVipers">
          <fptr
```

FILEID="305962"/></mets:div>

LABEL="ABeaver">

FILEID="305904"/></mets:div>

LABEL="ABear">

FILEID="305926"/></mets:div>

<mets:div TYPE="folio"

<fptr

<mets:div TYPE="folio"

<fptr

<recordID>info:getty/object/</recordID>

</mets:div>