



**THE UNIVERSITY OF MICHIGAN
UNIVERSITY LIBRARY
DIGITAL LIBRARY PRODUCTION SERVICES**

DIGITAL CONVERSION UNIT FOR
THE UNIVERSITY OF MICHIGAN LIBRARIES

**UNIVERSITY DIGITIZATION SPECIFICATIONS
REQUEST FOR PROPOSAL
LAST UPDATE: MAY 1, 2007**



THE UNIVERSITY OF MICHIGAN
UNIVERSITY LIBRARY
DIGITAL LIBRARY PRODUCTION SERVICES

DIGITAL CONVERSION UNIT FOR
THE UNIVERSITY OF MICHIGAN LIBRARIES

UNIVERSITY DIGITIZATION SPECIFICATIONS
REQUEST FOR PROPOSAL
LAST UPDATE: MAY 1, 2007

TABLE OF CONTENTS

<i>I. General Requirements Information.....</i>	<i>1</i>
Purpose of Document.....	1
Communication Guidelines.....	2
Scope of Work.....	2
<i>II. Required Deliverables for Proposals.....</i>	<i>4</i>
Written Proposal Document.....	4
Sample Digital Image Files.....	4
Pricing Matrix for Scanning	5
<i>III. General Requirements for Preferred Vendors.....</i>	<i>5</i>
Subcontracting.....	5
Rights.....	5
Responsibilities.....	5
Digitization Capabilities.....	5
Costs.....	6
Progress Reports.....	6
Invoicing.....	6
<i>IV. Technical Requirements for Digitization Product.....</i>	<i>7</i>
File Directory Structure Requirements.....	7

Image Capture Requirements	8
Image Color Requirements	10
Image Resolution Requirements.....	11
Image Quality Requirements	12
Image Format Requirements.....	12
Image Metadata Requirements.....	14
File Checksum Requirements.....	15
Ancillary Image Files.....	16
<i>V. Evaluation Process.....</i>	<i>18</i>
Vendor Quality Control Expectations.....	18
University Quality Control of Vendor Digital Product.....	18
Digital Quality Control of Files.....	18
Qualitative Evaluation.....	19
<i>VI. Data Delivery Requirements.....</i>	<i>20</i>
Shipping Requirements & Shipping Address.....	20
Delivery Format Options.....	20
<i>VII. Errors in Digitized Product & Rework.....</i>	<i>21</i>
Errors by Vendor.....	21
Errors by University.....	22
<i>VIII. Return of Scanned Materials.....</i>	<i>22</i>
Approval Required for Return of Materials.....	22
Shipping Requirements & Shipping Address.....	22
Timing of the Return of Library Materials.....	23
<i>IX. Revision of Specifications.....</i>	<i>23</i>
Revision of University Digitization Specifications.....	23
Revision of University Technical Specifications.....	23
<i>X. Selection of Preferred Vendors.....</i>	<i>23</i>
Notification of Selection.....	23
Letter of Agreement.....	24
Award of Projects to Preferred Vendors.....	24
<i>Appendix A: Checklists for Digital Product Compliance.....</i>	<i>1</i>
MDP Project Delivered on USB Mass Storage Device.....	1
MDP Project Delivered on Optical Discs.....	2
BibID Project Delivered on USB Mass Storage Device.....	3
BibID Project Delivered on Optical Discs.....	4
File Structure Organization for Hard Disk Drive Data Delivery.....	7

File Structure Organization for Optical Disc Data Delivery.....	8
File Structure Organization for Hard Disk Drive Data Delivery.....	11
File Structure Organization for Optical Disc Data Delivery.....	13
<i>Appendix F: XMP Metadata for JPEG2000 Files.....</i>	<i>18</i>
<i>Appendix G: Sample Scanning Notes Text File.....</i>	<i>22</i>
<i>Appendix H: Sample Bibliographic Target.....</i>	<i>23</i>
<i>Appendix I: Sample Production Note Target.....</i>	<i>24</i>



THE UNIVERSITY OF MICHIGAN
UNIVERSITY LIBRARY
DIGITAL LIBRARY PRODUCTION SERVICES

DIGITAL CONVERSION UNIT FOR
THE UNIVERSITY OF MICHIGAN LIBRARY

UNIVERSITY DIGITIZATION SPECIFICATIONS
REQUEST FOR PROPOSAL
LAST UPDATE: MAY 1, 2007

I. GENERAL REQUIREMENTS INFORMATION

PURPOSE OF DOCUMENT

The University of Michigan Library wishes to establish a list of preferred digitization vendors to provide supplementary services to the university's digitization efforts. This document serves as both as a request for proposal for new vendors and as the general specifications governing any particular project undertaken with a vendor.

This document will establish the parameters and requirements which will guide these digitization projects. This document will establish the general terms for these digitization projects, the financial requirements governing these digitization projects, a general description of the materials selected for these projects, a description of the types of project shipments, the technical requirements for the digital images being produced, the evaluation process which will be applied by the university on the vendor proposals and the digital product of preferred vendors ongoing, and the delivery requirements for providing the digital product to the university.

For purposes of this document, the term "project" shall be defined as a single shipment of University materials which is governed under the terms of this document. Other shipments pending with a vendor under the terms and specification of another agreement shall defer to the requirements of that project.

Vendors have been invited for proposals based on known capabilities for digitization and ongoing business relationships with the University Libraries in general and with the Digital Library Production Service and Digital Conversion Unit in particular. Vendors are welcome to contact the University with questions or concerns at any point during the RFP period and also during the subsequent projects governed by this document.

The University of Michigan reserves the right to revise or amend these requirements and specification at any time. If any substantial revision or amendment to this document occurs (meaning beyond grammatical corrections), the university will communicate the new specifications to all vendors with preferred status as soon as the revision is complete. All

vendors who have preferred status at the time of revision shall have the right to revise their pricing structures in response to the revised specifications. All projects and shipments that are currently underway in a vendor's workflow shall be exempt from any revisions that occur while the digitization is underway.

COMMUNICATION GUIDELINES

Digitization vendors are invited to bring up questions, concerns, technical difficulties, etc. on an ongoing basis through the course of providing digitization services to the University. In order to facilitate efficient exchange of information, the University encourages each vendor to designate a contact person for communication with the University regarding these digitization projects. In order to facilitate answers to vendor questions, the university has designated the following three individuals as primary contact points for vendors participating in this project.

Any administrative or contractual questions may be directed to the following:

Katharina H. Lebowsky
Purchasing Support Specialist
Purchasing Services
Phone: (734) 615-7071
Fax: (734) 615-6235
Email: klebow@umich.edu

Questions about these specifications, data or digitization products, technical matters, or other questions related to the digitization work of this pilot projects may be directed to the following:

Christopher Rennie
Digital Production Services Manager
Digital Conversion Unit
Telephone: 734-615-2836
Facsimile: 734-936-2288
E-mail: ccrennie@umich.edu

Any questions about invoicing, billing or other business-related matters may be directed to the following:

Ann Thomas
Senior Business Services Manager
Library Information Technology Business Services
Telephone: 734-764-6568
Facsimile: 734-763-7885
E-mail: athomas@umich.edu

SCOPE OF WORK

The types of scanning projects covered by this RFP will include but may not be limited to the following:

Michigan Digitization Project

These projects will be the most common type of shipment and materials sent to a preferred vendor. Materials in these projects will be paper-based, primarily textual materials that are unable to be scanned through the ongoing Michigan Digitization Project; the inability to scan these materials will vary but will predominantly include items in poor condition, oversized items, disbound items, items with large foldouts, items with binding issues, and so on.

Projects in this category will be clearly indicated by a packing note in each box which designates the project as a Michigan Digitization Project (MDP) shipment. This designation will affect the file structure and types of page image files required, as detailed later in this document.

Individual volumes in projects in this category will be identified in their digitized data format by the volume barcode located on the inside back cover of the volume; if a vendor finds any volume in projects of this category which does not have an identifiable barcode of this type, the vendor must contact the digital conversion unit immediately for instructions.

MDP projects will be shipped to the vendor with a minimum of preparation work. The shipping containers will include a packing list only. Vendors will be permitted to make their own judgments with regards to the proper color space for individual pages. Criteria for making these judgments are detailed later in this document.

Volumes in these shipments will **not** have a bibliographic target nor a production note target; therefore, the ancillary files associated with these targets are neither expected nor required for volumes in these projects. The AIIM scanning target and notes.txt ancillary files are still required for each volume, however.

Bibliographic Identifier Project

These projects will be the second most common type of shipment and materials sent to a preferred vendor. Materials in these projects will be paper-based, primarily textual materials such as books and bound volumes, pamphlets, manuscripts, archival materials, and so on. These materials may be drawn from the University's general library collections, special collections, archives or museums. Any special handling instructions and requirements will be included with the project documentation and may be discussed with vendor before shipment of the project.

Projects in this category will be clearly indicated by a packing note in each box which designates the project as a Bibliographic Identifier (BibID) shipment. This designation will affect the file structure and types of page images required, as detailed later in this document. Individual volumes or parts of volumes in this category will be identified in their digitized data format by the bibliographic identifier found on the production note target; if a vendor finds any volume in projects of this category which does not have a production note target, the vendor must contact the digital conversion unit immediately for instructions.

BibID projects will be shipped to the vendor after preparation work is performed by the Digital Conversion Unit. The shipping containers will include a packing list, and minimally each volume shall also have a bibliographic target and production note target. The preparation work shall identify specific pages requiring color or grayscale imaging.

Volumes in these shipments will have both a bibliographic target and a production note target; therefore, the ancillary files associated with these targets are expected and required for volumes in these projects as well as the AIIM scanning target and the notes.txt ancillary files.

II. REQUIRED DELIVERABLES FOR PROPOSALS

The following deliverables are expected from all vendors who provide a proposal in response to this RFP. These documents and files will provide the basis for evaluating the vendor's proposal

WRITTEN PROPOSAL DOCUMENT

Vendors invited to submit proposals in response to this RFP must provide a written proposal which details that vendor's ability to meet each of the requirements of this document. The structure of the proposal document is left to the discretion of the vendor, but it must address every requirement detailed in this document.

In addition, vendors submitting proposals should address their ability to meet the general specifications outlined in this RFP, including:

- Overall technical approach; proposed methodology; demonstrated understanding of the scope of work and the requirements.
- Vendor facilities and equipment, including detailed descriptions of the variety of equipment available in their facility (scanning equipment, digital cameras, lighting, book cradles, etc.)
- Post-processing and quality control procedures, including software used for each process
- Previously demonstrated production experience and past performance, including references from other institutions for which the vendor has performed comparable work.
- Capability and experience of key personnel.
- Project management and corporate support capability.

Vendors are invited to submit questions or ask for clarification of this document at any time during the proposal period. Questions should be submitted to the Digital Production Services Manager.

All vendor proposals must be received by the University of Michigan either electronically or in paper by Friday, July 27, 2007.

SAMPLE DIGITAL IMAGE FILES

Vendors who submit proposals in response to this RFP must submit sample digital image files along with their written proposal document. These sample digital image files must minimally include one TIFF formatted image file and one JPEG 2000 formatted file. Both files must be formatted according to the technical requirements detailed in this document, including all metadata fields, though in the sample image files, these metadata fields may be populated with "dummy" or placeholder information.

If the sample images fail to pass evaluation with the University, the University will provide a list of errors it is able to find to the proposing vendor and grant that vendor an opportunity to fix or repair those sample images for resubmission with the original proposal. Resubmitted sample images which do not meet the requirements and specifications of this document may be used to reject that vendor's proposal.

The original submission of sample image files must be received along with the proposal document on Friday, July 27, 2007.

PRICING MATRIX FOR SCANNING

The vendor must also produce a pricing matrix for digitization costs. This matrix must be complete enough to allow the University to accurately and fairly estimate the costs of any individual project for budgetary or recharge purposes. The university strongly prefers a cost per image model of cost matrix, but the vendor may vary if another metric will more accurately details the costs of a project. All cost line items which may appear on a project invoice must be detailed in this pricing matrix, even if the costs cannot be accurately estimated, e.g. shipping costs are highly variable but should be mentioned and addressed in the document.

This pricing matrix must be formulated and submitted to the University along with the proposal document and sample image files by Friday, July 27, 2007.

III. GENERAL REQUIREMENTS FOR PREFERRED VENDORS

SUBCONTRACTING

All digitization will be done on the vendor's premises. No subcontracting of this work will be permitted without the prior written approval of the library.

Vendors whose digitization facilities are outside North America, even if those facilities are part of the vendors' overall business operations, must provide detailed information about these facilities in their proposal. Any costs not included in the vendor's pricing matrix which will be incurred by the university must also be detailed (e.g. shipping costs to international facilities).

RIGHTS

The materials to be scanned, the image files, and the contents of the images will remain the property of the University of Michigan University Libraries. The materials to be scanned, the image files, and/or the contents of the images may not be reproduced or sold without the express written authorization of the University.

RESPONSIBILITIES

The University of Michigan University Libraries Digital Conversion Unit staff are responsible for the projects governed by this agreement and will track the flow of materials through the digitization process from the point of shipment to the vendor and on through to payment for services.

The Digital Production Services Manager is the project manager for the University. His departmental staff will carry out the pre-scanning preparation of the selected material and post-scanning processing of the digital product.

DIGITIZATION CAPABILITIES

Each vendor must supply a detailed description of their own digitization capabilities to the University. This detail should include, but is not limited to, the maximum physical dimensions

of individual scan image at the acceptable resolutions, make and models of scanners, color depth capabilities of scanning equipment, number of scanning stations, and so on.

Costs

Each vendor must supply a pricing matrix to the University which provides enough detail so as to allow the University to accurately estimate project costs within 5% of invoice total. This matrix may be structured according to vendor preferences but must include and detail per image costs according to any document size thresholds, any per image cost differences among bitonal, grayscale, and color capture, and any per image costs for image post-processing.

Vendors must also provide a detailed list of ancillary costs which projects will incur which are unpredictable, such as shipping costs for return of materials, shipping costs for data delivery, media costs for burned optical discs, and so on.

Vendors are also strongly encouraged to provide any additional cost information for rush services, if the vendors provides such service. Details which the University would need to know here would be any targeted or guaranteed turnaround times, additional costs information, and methods for designating a shipment as a rush shipment.

PROGRESS REPORTS

The University of Michigan requires weekly progress reports on the status of projects. These progress reports must include the university shipment number, the date of shipment receipt by the vendor, and an estimated percentage complete. A projected date for completion is highly encouraged.

Progress reports must be provided on the first business day of each week.

INVOICING

The vendor shall provide a detailed invoice for the digitized microfilm within ten (10) business days of shipping the digital product and microfilm to the University. Each vendor's invoice must reference the following:

- The University of Michigan Purchase Order Number
- University shipment number
- Total number of images, itemized according to type (e.g. bitonal, grayscale, and color itemized separately)
- Cost per image rate
- Volume names and item contents for CDs or DVDs delivered
- Any other itemized charges (Cost per optical disc, shipping costs, transportation, etc.)
- Total charge for the shipment

All invoices shall be sent to Ann Thomas, Senior Finance Specialist. Electronic submission of invoices is strongly preferred; however, if paper invoices are the only possibility for a vendor, they must be mailed to the following address within the same required timeframe:

Ann Thomas
Senior Business Services Manager
Library Information Technology Business Services

300 Hatcher Graduate Library North.
920 North University Avenue
Ann Arbor, Michigan 48109-1205

IV. TECHNICAL REQUIREMENTS FOR DIGITIZATION PRODUCT

The following criteria represent the digital imaging requirements for the vendor digitization product of projects governed by this document.

FILE DIRECTORY STRUCTURE REQUIREMENTS

The file directory structure requirements for the digitization product depends significantly on what designation has been given to an individual project. The differences are detailed in the following sections:

Michigan Digitization Projects

Projects of this category and designation require a very flat file structure. The top level of directories consists of directories named for the 14-digit barcode of individual physical volumes. Each physical volume gets its own directory named for the volume barcode and all page images from outside front cover to outside back cover are contained therein with ancillary target image files as applicable.

Depending on the delivery method chosen, these barcode directories may be the very top level directories (when delivering on optical disc media) or may be subdirectories of single directory named for the University shipment number (when delivering on USB mass storage device).

This directory structure and its requirements are further detailed in Appendix C.

Bibliographic Identifier Projects

Projects of this category and designation require a hierarchical file directory structure. The top level of directories consists of directories named for the bibliographic identifier for individual titles. Under these BibID directories will appear one or more subdirectories, named by four digit numbers, for the volume number or year of a multivolume set if applicable. Under these volume directories will appear one or more subdirectories, named by three digit numbers, for the issue number or part of a multivolume set if applicable. The production note target will detail each of these levels in the hierarchy, making the naming and organization of each of these levels of the file structure hierarchy clear. Each physical volume will minimally require its own directory hierarchy and all page images from outside front cover to outside back cover are contained within the third level subdirectory (named with the three digit number); multivolume or multipart bibliographic volumes may share one or more directories in the hierarchy.

Additionally, all continuous tone images (grayscale and color) must be stored in a subdirectory of the three-digit directory name “pages” exactly with letter case sensitivity observed. The contone image files in this “pages” directory must be named with an eight character file name which is identical to its corresponding bitonal image in the directory above. The first zero of the filename of all contone images should also be changed to a lower case letter “p” in the contone image file name. For example, if the 256th page image in a single volume work has a color photograph should have bitonal image named “00000256.tif” in the directory “.../0001/001/” **and** a color contone image named “p0000256.tif” in the directory “.../0001/001/pages/” as well.

Depending on the delivery method chosen, these BibID directories and their subdirectories may be the very top level directories (when delivering on optical disc media) or may be subdirectories of single directory named for the University shipment number (when delivering on USB mass storage device).

This directory structure and its requirements are further detailed in Appendix D.

IMAGE CAPTURE REQUIREMENTS

The following image capture requirements apply to all projects, regardless of their designation as either an MDP project or a BibID project.

The normal image capture requirement for all volumes is to capture them from outside front cover to outside back cover, with all images kept and named in sequential order. File names should start at 00000001 and increment sequentially for each additional image, following the sequence of the original library materials. Page image files must also be named with a file extension appropriate to the image file format, as detailed later in this document. The front cover image should include the spine if possible. All end papers and covers shall be scanned as well unless otherwise instructed by the University.

Michigan Digitization Projects

For projects designated as MDP projects, vendors are required to produce one and only one image for every page within a physical volume. Each image shall be in the required format for the image's content type (bitonal or contone) as detailed later in this document.

Pages containing only black and white text or simple line drawings must be captured as bitonal images. Pages containing half-tone or continuous tone photographs and highly detailed lithographs should be captured as grayscale images. Pages containing any meaningful color elements should be captured as color images.

Bibliographic Identifier Projects

For projects designated as BibID projects, vendors are required to produce one bitonal image for every page within a physical volume. In addition, vendors are required to produce contone images for those pages whose content dictates continuous tone imaging. It is permissible for vendors to capture every page in the color space appropriate for that page and to derive an additional bitonal image file for those pages which require both color and bitonal imaging.

Pages containing only black and white text or simple line drawings must be captured as bitonal images. Pages containing half-tone or continuous tone photographs and highly detailed lithographs should be captured as grayscale images. Pages containing any meaningful color elements should be captured as color images.

The following categories represent special categories which may be encountered in any volume in any project, regardless of designation. Any vendor encountering a special situation which is not addressed in this document should be contact the University for further instruction.

Missing Pages

In all projects, physical volumes may have missing pages. For any bibliographic entity where the pagination is broken by missing pages, missing page targets must be inserted. For projects designated as BibID projects, Digital Conversion Unit staff will attempt to flag each instance where a page is missing from a volume; for projects designated as MDP projects, the minimal preparation of the volume will not be able to flag missing pages. However, should vendor scanning technicians find other pages missing they must insert missing page targets as appropriate.

For example if pages 5 and 6, 21 and 22 are missing from the original text, the scanning operator must replace the missing pages with missing page targets for both the front and the back of each leaf (this example would require 4 missing page targets). Should the scanning operator notice a missing page without a flag, he/she should insert the missing page target. The sequence should follow the original source document sequence and the corresponding files as if the missing page were scanned.

The missing page targets should 8.5 x 11 inch targets regardless of the size of the physical volume. The text of the missing page target must be “Page Missing in Original Volume” in minimally 36 point font.

Foldouts, Centerfold Images, or Two-Page Spreads

These types of pages present special challenges in capture. All foldouts are expected to be scanned by the vendor unless authorization and instruction to the contrary are received from the University. Normally a foldout 11 x 16 can be scanned on most scanners without adjustment to resolution. Foldouts larger than 11 x 16 may be considered oversized and may be a cause to adjust image resolution as necessary.

- If the foldout cannot be safely scanned, the vendor must contact the University for further instruction.
- Foldouts may be removed with prior permission from the University for better image capture. If foldouts are removed, they should be returned to and inserted in their proper order within the volume from which they came.
- Centerfolds (also two-page spreads and uncut plates) will be treated as foldouts in the scanning process. Unless otherwise requested, an extra blank page (8.5 x 11 in) should be inserted after the centerfold image to maintain the correct recto/verso sequence. See Appendix B for sample diagram.
- Any foldout encountered which a vendor cannot capture as one whole image should be reported to the University. University staff will provide the vendor with further instruction.

IMAGE COLOR REQUIREMENTS

The following image color requirements apply to all page images in all projects, regardless of their designation as either an MDP project or a BibID project.

Bitonal Images

Any page image which consists of text only or which consists of line art against the background paper color should be captured as a bitonal image.

Grayscale Images

Any page image which uses halftone or continuous tone photographs, variously shaded gray graphs or diagrams, or variously shaded gray lines to distinguish among multiple chart or illustrative elements should be captured as a grayscale image.

All grayscale images should be captured in an 8-bit color space.

In all projects, vendors are expected to use reasonable judgment to determine whether a grayscale image is most appropriate image capture method; in the case where it is not definitely clear whether to capture in grayscale or color, the vendor can safely assume to use color. Vendors who are unsure about whether to capture any given page image in grayscale are welcome to contact the University for additional instructions.

Color Images

Any page image which uses color photographs, colored bar graphs or diagrams, or colored lines to distinguish among multiple chart or illustrative elements should be captured as a color image.

All color images should be captured in a 24-bit color space, preferably the sRGB color space.

In all projects, vendors are expected to use reasonable judgment to determine whether a color image is most appropriate image capture method; in the case where it is not definitely clear, the vendor can safely assume to use color. Vendors who are unsure about whether to capture any given page image in color are welcome to contact the University for additional instructions.

IMAGE RESOLUTION REQUIREMENTS

The following image resolutions requirements apply to all projects governed by these specifications, regardless of the individual project's designation as an MDP project or a BibID project. Vendors should report their ability to meet these requirements in detail as part of their proposal document.

On all images in all color spaces, X resolution and Y resolution must be identical values.

Bitonal Images

All bitonal page images must have a resolution measured in pixels per inch (ppi) of 600ppi. This resolution should be an uninterpolated optical resolution wherever possible; if the scanning equipment can only achieve this resolution via interpolation, the resolution may be interpolated by the scanning camera and its software as part of the original page scan, not as part of a separate post-processing image adjustment.

If the optical resolution of the scanning equipment is not able to achieve an exact optical resolution of 600ppi but can achieve a resolution within 25ppi up or down from 600ppi, the vendor may resample the image using image correction software either up or down to 600ppi resolution.

If the vendor cannot achieve 600ppi for any page image (e.g. an oversize foldout), the vendor is required to resample the image up to the next "round number" resolution if the difference is less than 5% of the final resolution. For example, if the scanning equipment can achieve a resolution of 575ppi, it is permissible to resample that image up to 600ppi because the additional 25 pixels in resolution is less than 30 pixels (5% of 600ppi). However, if the scanning equipment can only achieve a resolution of 555ppi, the image should then be downsampled to a resolution of 400ppi because the 45 pixel difference in resolution exceed 5% of the final resolution when upsampled to 600ppi. All deviations from the normal 600ppi resolution for bitonal images must be recorded in the notes.txt file for that volume or bibliographic entity.

Bitonal images with a resolution below 400ppi are not acceptable without first contacting the University for instruction on how to proceed.

Continuous Tone (Color & Grayscale) Images

All contone (continuous tone) page images must have a resolution measured in pixels per inch (ppi) of 400ppi. This resolution should be an uninterpolated optical resolution wherever

possible; if the scanning equipment can only achieve this resolution via interpolation, the resolution may be interpolated by the scanning camera and its software as part of the original page scan, not as part of a separate post-processing image adjustment.

If the optical resolution of the scanning equipment is not able to achieve an exact optical resolution of 400ppi but can achieve a resolution within 20ppi up or down from 400ppi, the vendor may resample the image using image correction software either up or down to 400ppi resolution.

Because resolution on many scanning workstations involves physical adjustments to the scanning head, vendors are also permitted to scan contone images at a higher resolution (e.g. 600ppi) and to downsample the final image resolution to 400ppi using postprocessing tools. The University strongly recommends using bicubic interpolation during such downsampling.

If the vendor cannot achieve 400ppi for any page image (e.g. an oversize foldout), the vendor is required to resample the image up to the next “round number” resolution if the difference is less than 5% of the final resolution. For example, if the scanning equipment can achieve a resolution of 381ppi, it is permissible to resample that image up to 400ppi because the additional 20 pixels in resolution is less than 20 pixels (5% of 400ppi). However, if the scanning equipment can only achieve a resolution of 375ppi, the image should then be downsampled to a resolution of 300ppi because the 25 pixel difference in resolution exceed 5% of the final resolution when upsampled to 400ppi. All deviations from the normal 400ppi resolution for contone images must be recorded in the notes.txt file for that volume or bibliographic entity.

Contone images with a resolution below 300ppi are not acceptable without first contacting the University for instruction on how to proceed.

IMAGE QUALITY REQUIREMENTS

All images, being captured from a 2N microfilm print master negative, must have their polarity reversed, so that text and graphics appear as black/gray pixels against a white background.

All 2-up frame images must be split at the gutter (where appropriate) to produce separate images files for each page (recto and verso). Foldouts or other single pages which fill the full microfilm frame must not be split.

All page images must be rotated (where appropriate) so that the page image appears in proper reading position on screen when the image file is opened.

All page images should fill the image frame to the largest extent possible. Preservation microfilming guidelines required that the reduction ratio on single reel of microfilm remain constant throughout the reel; therefore the bibliographic entities may or may not all fill the individual frame on the reel. Vendors should eliminate or minimize the amount of black border which frames the page image either during capture or in post-processing the image.

IMAGE FORMAT REQUIREMENTS

All images files shall be provided in either one of two formats: TIFF or JPEG2000.

TIFF Format Image Requirements

For all TIFF format image files, the internal bitmap must be single strip format. Color TIFF images must all be in the sRGB color space.

Compliance with the TIFF format specification will be verified using the JHOVE object validation software, and all TIFF images must be minimally “well-formed” against the TIFF standard in order to meet acceptability criteria.

Michigan Digitization Projects

For projects designated as MDP projects, all bitonal image files delivered must be in the TIFF format and also must be in CCITT Group 4 compression. Continuous tone images for MDP projects must be delivered in JPEG 2000 format, detailed later in this document.

Bibliographic Identifier Projects

For projects designated as BibID projects, all bitonal image files delivered must be in the TIFF format and also must be in CCITT Group 4 compression. All continuous tone (grayscale or color) image files delivered in the TIFF format must be uncompressed.

JPEG2000 Format Image Requirements

Vendors are required to explain and detail their ability to support this file format in their proposal documents.

The JPEG2000 format for image files is the required format for delivery of the continuous tone (color and grayscale) digitized images produced by projects designated as MDP.

BibID projects may deliver in JPEG 2000 format as an option at present, though still in the same file directory structure. This delivery format will be preferred in the near future by all University digitization projects.

In order to maintain consistency among the JPEG2000 files in the library repositories which are produced from various projects and vendors, the University requires the following parameters for all JPEG2000 files:

- The JPEG 2000 file will conform with JP2 file format as specified in ISO/IEC 15444-1:2000 (i.e., JPEG 2000, Part 1).
- The JPEG 2000 file will be prepared after any image processing or clean-up is performed.
- The JPEG 2000 file's image X origin, image Y origin, tile X origin, and tile Y origin will be 0.
- The JPEG 2000 file will contain only one component.
- The JPEG 2000 file's height and width will be the same as the master image file after transcoding.
- The JPEG 2000 file will be in the sRGB color space.

- The JPEG 2000's progression order will be RLCP (resolution, layer, component, position) or RLPC.
- The JPEG 2000 file will have 2 decomposition levels.
- The JPEG 2000 file will have 8 quality layers.
- The JPEG 2000 transcoding process will use the 9-7 irreversible filter.
- The default slope rate distortion used in transcoding will be 51492.
- No tiling is necessary in the JPEG 2000 file

All image files supplied in the JPEG2000 file format must be compliant with Part 1 of the JPEG2000 core coding system and meet the JP2 basic file format definition. Files must be supplied with the “.jp2” extension.

Many scanning hardware and software packages do not support native image capture to the JPEG2000 format. Transcoding an image file to the JPEG2000 file format is considered acceptable and normative. The file format of the original image file should be a non-distorting format, such as uncompressed TIFF; the original JPEG format should not be used as the original format in order to prevent the image artifacts produced by the compression techniques of that format.

The University of Michigan has adopted the Kakadu open source software package for the transcoding of image files into the JPEG2000 format. Due to variations among the various JPEG2000 software packages available, it is recommended and encouraged to our vendors as an excellent and flexible package for transcoding files to the JPEG2000 format. The University will gladly provide the parameters we use to insure compatible encoding of JPEG2000 image files.

IMAGE METADATA REQUIREMENTS

All image files must have technical and source metadata embedded within the image file itself. The requirements for metadata inclusion for each format are detailed in the following sections.

TIFF Image Metadata Requirements

To help ensure the long-term sustainability of their master images, the University of Michigan Library requires the use of some of the data elements in Draft NISO Z39.87 –2002 Standard, Data Dictionary—Technical Metadata for Digital Still Images to record technical metadata in the TIFF header of each image. In addition to required TIFF tags, some additional tags are utilized to document the provenance of each image.

In addition to the required fields for baseline TIFF images, the following fields of the TIFF header must have the appropriate values inserted:

269	DocumentName	<Barcode>/<image file name> or <BibID>/<image file name> <i>The formats above must be consistent among the files in a single project; the <Barcode> is appropriate for projects designated as MDP projects and the <BibID> format is appropriate for projects designated as BibID projects</i>
306	DateTime	Time and date of image capture in ISO8601 format
315	Artist	Scanning vendor company name

This information is mandatory and will be included in the Library's quality assurance and evaluation processes.

In addition to the required information, the following fields in the TIFF header are recommended for recording additional technical metadata. Presence of this information is not a requirement for satisfactory performance of these specifications.

271	Make	Scanner manufacturer
272	Model	Include model number
305	Software	Include version number

See Appendix E for a more complete explanation with sample values.

JPEG2000 Image Metadata Requirements

All image files submitted in the JPEG2000 format must record technical metadata in an XMP packet written to the UUID box in the JPEG2000 header. This does not supersede or exclude the requirement to write any of the same information (such as the number of pixels for the width or height of the image bitmap) to the required elements of the file header for compliance with the JPEG2000 file format specification.

See Appendix F for a more complete explanation with sample values.

FILE CHECKSUM REQUIREMENTS

Each bibliographic entity subdirectory shall contain a MD5 checksum signature file for all of the image files in that subdirectory in order to verify that the digital image files are complete and intact as the files are moved and migrated from system to system and disk to disk. This signature file is identical to the file supplied in previous University of Michigan digitization specifications.

MD5 checksum files are required for all directories containing images and are required to include all image files regardless of format. This requirement remains in effect even between the different project designations of MDP projects and BibID projects.

This MD5 checksum signature file is required in all directories in which image files are contained and only in such directories. Directories which only contain subdirectories should not contain checksum files, even though those subdirectories may contain image files themselves.

The MD5 signatures for each image file are concatenated into a single file that must be named "checksum.md5" and this file must be stored in the item directory at the same level as the page image files. Checksums are validated throughout the workflow to ensure that files are accurately written from one medium to another. In addition, they provide a check against bit-deterioration and serve to guarantee digital object authenticity.

The MD5 file contains one line for each image in the volume. Each line consists of the 32-character MD5 signature followed by 2 "blank" characters followed by the image file base name. Please notice in the sample lines below that no directory names are present, nor is the directory name present in the MD5 filename.

Here is a sample MD5 file:

786fd62b33b65560d9993d0f5f4235d1	00000001.tif
0fa7abf97a4414aed10b316c59b07249	00000002.tif
f30cc4a3d27f54329b3d9aaa5b2d7bda	00000003.tif
6a621fe605578f95cc66cc27b7ca77b5	00000004.tif
97c664aa9fb998dde78ce2aecbf59d73	00000005.tif
01cb4b01a9de2aa1660da009989f5f13	00000006.tif
1e8cb443ca9d72be25fa545c2a82a541	00000007.tif
745d4e8edfc7492e43eb6053b96cffb2	00000008.tif
f238e2dd728970cb8bb5bc9925c8a409	00000009.tif
a22322fec76543a0df26e08de58d0bfa	00000010.tif
42e9295e8dc45cb082e2759bf3cb4b5c	00000011.tif

It is considered best practice to generate MD5s in the same location as the original scanned images were saved, prior to burning a CD/ DVD or moving the files to another mass storage device or disk subsystem. Vendors are strongly encouraged to verify the MD5 signatures after burning a CD/ DVD or moving the data to another mass storage device or disk subsystem in order to verify that the content was fully transferred to the media.

The checksum algorithm used should be the 128-bit md5 (Message Digest 5) algorithm described in RFC 1321 (<http://www.faqs.org/rfcs/rfc1321.html>).

ANCILLARY IMAGE FILES

The University requires a detailed explanation of the vendor quality control processes, especially as it pertains to verifying the requirements and specification of this document.

AIIM Scanner Test Target

The University requires that that every set of scanned page images, whether a complete physical volume or a subdivision of a physical volume, must include a scan of the Association of Information and Image Management (AIIM) Scanner Test Chart #2, AIIM X441. This target allows the University to evaluate scanner performance, focus, and image clarity for that scanning session.

Each directory should include a scan of the AIIM target which is designated for paper based scanning. The scanned image file for this target must be scanned as a bitonal image in TIFF format with CCITT Group 4 compression; and the file must be named “aiim.tif” exactly with appropriate letter case sensitivity observed. The University shall use the following target evaluation criteria:

- The Bodoni 4 point lower case letters should be clear and distinct
- The diagonal line should be smooth and straight
- Distinct halftone wedges representing the dynamic range present in the source document should be rendered free of moiré patterns at the appropriate screen ruling (normally 133) in either regular or enhanced mode.

If these criteria cannot be met at the settings deemed optimal, the vendor should note this in the notes.txt included with the image files.

This scanning target may be ordered from:

AIIM Headquarters
1100 Wayne Avenue, Suite 1100
Silver Spring, Maryland 20910
Phone: 301.587.8202
Toll free: 800.477.2446
Fax: 301.587.2711
E-mail: aiim@aiim.org
http://www.techstreet.com/cgi-bin/detail?product_id=6845

Scanning Notes Text File

The University requires that that every set of scanned page images, whether a complete physical volume or a subdivision of a physical volume, must include a simple text file which reports any problems scanning, any deviations from normal resolution, or any other information about the scanning process which the vendor judges should be reported to the University.

This file must be a text file named “notes.txt” exactly with appropriate letter case observed.

An example of this file is represented in Appendix G.

Bibliographic Target

The University requires that that every set of scanned page images, whether a complete physical volume or a subdivision of a physical volume, must include a scan of the bibliographic target sheet included for that bibliographic entity.

The scanned image file for this target must be scanned as a bitonal image in TIFF format with CCITT Group 4 compression; and the file must be named “bib.tif” exactly with appropriate letter case sensitivity observed.

These bibliographic target images are only required for projects designated as BibID projects. MDP projects do not produce the bibliographic target, and therefore do not require these target image files. If any bibliographic entity for a BibID project is missing a bibliographic target, the vendor should contact the University immediately for instruction.

An example of this file is represented in Appendix H.

Production Note Target

The University requires that that every set of scanned page images, whether a complete physical volume or a subdivision of a physical volume, must include a scan of the production note target sheet included for that bibliographic entity.

The scanned image file for this target must be scanned as a bitonal image in TIFF format with CCITT Group 4 compression; and the file must be named “prodnote.tif” exactly with appropriate letter case sensitivity observed.

These production note target images are only required for projects designated as BibID projects. MDP projects do not produce the production note target, and therefore do not require these target image files. If any bibliographic entity for a BibID project is missing a production note target, the vendor should contact the University immediately for instruction.

An example of this file is represented in Appendix I.

V. EVALUATION PROCESS

VENDOR QUALITY CONTROL EXPECTATIONS

The University requires a detailed explanation of the vendor quality control processes, especially as it pertains to verifying the requirements and specification of this document.

UNIVERSITY QUALITY CONTROL OF VENDOR DIGITAL PRODUCT

The University shall engage in a formal evaluation both to verify that the digital product of the participating vendors meets the specifications detailed in this document and to evaluate the quality and acceptability of the digital imaging for University digitization efforts. The University reserves the right to reject any digital file or files which fail to meet the specifications and requirements detailed in this document, as determined by the evaluation method which follows. The University also reserves the right to require the rescan of any page images which do not meet the specifications and requirements of this document or to refuse payment, up to and including the whole digital product produced for a shipment of materials. If the digital product of any vendor fails to meet these specifications and requirements after rescan, the university reserves the right to require the return of materials for submission to another vendor.

DIGITAL QUALITY CONTROL OF FILES

The University shall use a variety of means to electronically evaluate the digital product of the vendors. In this electronic evaluation of the digital product of participating vendors, the evaluation will include, but shall not be limited to, the following areas:

Validation of File Structure

The University shall use a variety of automated means as well as manual inspections to insure that the digital file structure meets the specifications in this document. This will include verifying the parent directory is named with the proper identifier according to project designation (MDP vs. BibID) and verifying that all individual bibliographic entities are broken out into appropriate subdirectories that are named according to the specifications of this document.

Validation of File Naming Conventions

The University shall use a variety of automated means as well as manual inspections to insure that the digital file names meet the specifications in this document. This will include verifying the number of characters in each file name as well as verifying that all files in a given subdirectory are named in sequential numeric order, according to the specifications of this document.

Validation of File Metadata

The University shall use a variety of automated means as well as manual inspections to insure that the required metadata elements are present in each digital file according to the specifications in this document.

Validation of Checksum

The University shall use a variety of automated means as well as manual inspections to insure that the required MD5 checksum files are present in each digital file subdirectory and that all digital files are fully valid and verified to a check of the MD5 calculations, according to the specifications in this document.

Validation of File Format

The University shall use a variety of automated means as well as manual inspections to insure that the digital image files are compliant with the image format specification that they purport to be. The primary software tool for accomplishing this will be the JHOVE object validation environment developed by Harvard University. This tool is freely available to the public at the following URL: <http://hul.harvard.edu/jhove/>. The University currently uses version JHOVE 1.1e. Any questions on what constitutes an acceptable file validation output from JHOVE can be directed to the Digital Production Services Manager.

QUALITATIVE EVALUATION

The University shall use a variety of means to qualitatively evaluate the digital product of the vendors. In this qualitative evaluation of the digital product of preferred vendors, the evaluation will include, but shall not be limited to, the following areas:

Evaluation of Image Clarity

All digital images must clearly legible, in proper focus, and provide sharp representation of the original. Individual letters, particularly if they have closed loops or other bounded areas, in the typeface should not display “filled-in” areas and should be clearly identifiable.

For page images which contain illustrations, photographs, or other graphic materials, special attention should be given to eliminate moiré patterns that appear in the scanned image.

All page images should be filtered or processed so as to eliminate or reduce general noise or speckle effects that appear in the digital image.

In all shipments, any frames or images that are judged by the vendor to be impossible to image clearly should be reported to the Digital Conversion Unit and noted in the “notes.txt” file to insure that University evaluation of the vendor’s digital product takes this into account.

Evaluation of Image Skew

All digital images must be deskewed to provide an easily read page image when presented on a display screen. Deskewing techniques and tools should be mindful both of the page edges and the page content.

The University evaluates skew using a level horizontal rule across the image frame compared against lines of text on the page image, which are assumed to be level and square on the page. Deviations greater than 1 degree from the level rule are unacceptable.

Evaluation of Image Crop & Framing

All digital images should fill the frame of the image to the largest extent possible. “Black” borders where the image captures beyond the page itself must be eliminated from the final digital product

VI. DATA DELIVERY REQUIREMENTS

SHIPPING REQUIREMENTS & SHIPPING ADDRESS

The digitization product must be shipped to the following address:

Digital Conversion Unit
The University of Michigan
3212 Buhr Building
837 Greene Street
Ann Arbor, Michigan 48109-3213

DELIVERY FORMAT OPTIONS

The digital images files shall be delivered to the University in one of the following two optional methods. In either case, the delivery medium must be formatted so that they can be accessed from a variety of computer platforms and operating systems.

Portable USB Mass Storage Device

This delivery format is the University’s preferred data delivery method. In order to utilize this delivery method, the participating vendor must contact the University with an adequate amount of time to have a portable USB mass storage device made available for delivery and shipped to the vendor.

The file organization structure specified in the Technical Requirements section must be maintained through the process of storing or moving the data to the mass storage device.

The USB mass storage device will be formatted using the FAT32 file system to maintain the highest possible degree of interoperability among computer systems and operating systems. The mass storage device and the file system may not be altered or reformatted without contacting the University ahead of time.

It is strongly recommended after moving the digital image files onto the mass storage device that the MD5 checksums be verified to insure that all data has successfully migrated onto the mass storage device.

Optical Disc Format

This delivery format is an option available to vendors who are unable to use the USB Mass Storage Device delivery option. The file organization structure in the Technical Requirements section must be maintained through the process of burning the data to optical disc.

The following media choices are both acceptable:

- CD-R (Brand: Mitsui Gold)
- DVD-R (No brand preference at this time)

For shipments of more than 6GB, DVD-Rs are highly preferred. Media may be mixed (e.g. both DVD-Rs and CD-Rs) in a single shipment. Media should be filled as close to maximal capacity as is possible within the file structure guidelines. CD-RW and DVD-RW are not acceptable media.

All media must be written in close compliance to ISO-9660 in order facilitate use on the widest possible set of computing platforms and operating systems. Use of either UDF or Joliet extensions is not acceptable in the media format. All media must be written in a single session format only. Unreadable optical discs will be checked for ISO-9660 compliance with the CD/DVD Diagnostic software package from Infinadyne, Inc.

All optical disc media must be labeled with the electronic label assigned by the University for that project and incrementing sequentially for each disc included in the delivery shipment. The University assign disc label shall consist of a four-character prefix for the project, an underscore character, and three character enumeration beginning with "001" and incrementing one for each disc in the shipment. For example, for a shipment of three DVDs where the University assigned label prefix is "PT08", they would be labeled electronically "PT08_001," "PT08_002," and "PT08_003" respectively. For shipments where contone images are segregated from the bitonal images, separate media must be used for each type of image. Vendors who wish to provide data on optical disc should contact the university for clarification of this labeling process before burning the digitization process to disc.

In addition, all media must be clearly labeled on the physical disc itself with the following information:

- CD/DVD Label assigned by University
- University Shipment Number
- Vendor Name
- Date of Disc Write

VII. ERRORS IN DIGITIZED PRODUCT & REWORK

The staff of the Digital Conversion Unit and DLPS will provide rigorous review, both automated and manual, to insure that the digital data product meets the specifications and requirements of that project. For the purposes of this document, an "error" shall be defined as a digital product of a project which does not meet the requirements of this document, ranging from a single image up to and including an entire shipment.

ERRORS BY VENDOR

All errors made by the vendor which are identified by the University's quality control process must be corrected by the vendor at no additional cost to the library.

The University shall notify a vendor of all errors requiring rescanning or data correction immediately upon completion of quality control for that project. The vendor must then provide a timeframe for the correction and delivery of a new digital product.

Where the vendor cannot meet quality standards or specification due to the condition of the University's original materials, the vendor should immediately contact the Digital Conversion

unit for consultation and advice. This consultation must produce a clear direction for the vendor on how to proceed.

ERRORS BY UNIVERSITY

There are a number of errors for which University shall take responsibility and shall hold the vendor blameless:

- Misdesignation of project type, e.g. projects designated as MDP which should have been designated as BibID or vice versa are the responsibility of the University to correct
- Failure to provide required targets for project type, e.g. projects designated as BibID projects which fail to include a bibliographic target and production note for each volume or bibliographic entity are the responsibility of the University to correct. Vendors encountering this problem must contact the University for advice before shipping the digital product back to the University.
- Failure to provide required identifier for individual volume, e.g. project materials which do not have sufficient information by which to assign a proper identifier to the directory containing the page images. Vendors encountering this problem must contact the University for advice before shipping the digital product back to the University.
- Errors identified after digital product is accepted and physical materials returned to University.

VIII. RETURN OF SCANNED MATERIALS

APPROVAL REQUIRED FOR RETURN OF MATERIALS

The Digital Conversion Unit must approve of the return of all scanned materials in writing via e-mail, fax, or postal mail before materials are shipped. Materials may not be returned before the Library has received and approved the scanned data product. Shipments that are returned before approval is obtained from the Digital Conversion Unit may be subject to financial penalties if page images are unacceptable and require rescan. These penalties may accrue up to the entire scanning cost of the shipment if all images are unacceptable and the vendor has returned the materials before the images were approved.

Shipments of materials must be returned complete and shipped complete on the same day; a shipment may not be broken up into separate shipments upon return of materials. Vendors are strongly encouraged to preserve and reuse the packing boxes and padding materials used to ship the library materials to vendor's facilities.

SHIPPING REQUIREMENTS & SHIPPING ADDRESS

The materials to be returned to the Library must be shipped to the following address:

Digital Conversion Unit
The University of Michigan
3212 Buhr Building
837 Greene Street
Ann Arbor, Michigan 48109-3213

The Library strongly recommends that vendors ship materials to be returned via UPS. If the vendor cannot ship materials to be returned via UPS, the vendor must use a shipper that will load and unload from a ground-level loading area at the Buhr Building. Loading areas at the Buhr Building will **not** accommodate semi-tractor trailer carriers. Palletized loads are not acceptable..

TIMING OF THE RETURN OF LIBRARY MATERIALS

All library materials should be returned to the Digital Conversion Unit promptly upon the approval of the scanning data, but no later than five (5) business days after the scanning data approval is sent to the vendor. If the vendor prefers to only ship on certain days, that request may be approved to deviate from the aforementioned schedule if the Digital Conversion Unit is consulted in advance.

The Digital Conversion Unit should be notified of the date of the return shipment and by which carrier it was shipped; provision of tracking number is also appreciated.

IX. REVISION OF SPECIFICATIONS

REVISION OF UNIVERSITY DIGITIZATION SPECIFICATIONS

These digitization specifications are designed as general reference document which bind all projects which are engaged with a vendor selected as preferred. However, the Library reserves the right to revise any area of these specifications at any time. The Library will report all formal changes immediately to all preferred vendors once edits or clarifications as they are made; and the vendor will have three months to conform to those changes or withdraw from its preferred vendor relationship with the Library from the time of notification.

Input, questions, and comments from the vendors are invited on this specification document on an ongoing basis. In particular, any requirements or specifications that would be impossible for a selected vendor to meet should be addressed as soon as possible. Clarification or revision can be requested by participating vendors on any section; the University will attempt to address any and all concerns in a timely manner.

REVISION OF UNIVERSITY TECHNICAL SPECIFICATIONS

The technical imaging requirements of these digitization project specifications are undergoing constant refinement. Any number of cooperative projects at the Library may require a revision or alteration of these technical specifications as those projects progress. As the digitization needs of the university evolve or as individual projects demand, the University will modify these technical specifications where appropriate and possible for the purposes of this project. Any periodic modifications to the technical requirements after preferred vendors have been selected will be promulgated to all preferred vendors.

X. SELECTION OF PREFERRED VENDORS

NOTIFICATION OF SELECTION

All vendors submitting proposals in response to this RFP will be notified as to the final determination of their proposal within 30 days of that determination.

Any vendor whose proposal is not accepted has the right to resubmit a revised proposal within 30 days of receipt of notification for reconsideration. If the revised proposal is still determined to be unacceptable, then that vendor will be excluded from preferred vendor consideration until the next fiscal year.

Vendors new to the university may be invited to submit a proposal for digitization services at any time; however, if their proposal is accepted, their preferred status will not go into effect until the next fiscal year.

LETTER OF AGREEMENT

If selected as a preferred digitization vendor for the University, the vendor will be expected to sign a letter of agreement stipulating to these specifications as the governing requirements of all work and also stipulating to the accuracy of the cost detail quoted to the university in the vendor's proposal.

AWARD OF PROJECTS TO PREFERRED VENDORS

Selection as a preferred vendor for the university does not guarantee that any percentage or number of projects will be awarded to any vendor selected as preferred. Projects will be awarded based on the University's determination of what vendor best meets the technical requirements of individual projects. Price and prior performance are criteria which influence the award of individual projects to



APPENDIX A: CHECKLISTS FOR DIGITAL PRODUCT COMPLIANCE

MDP PROJECT DELIVERED ON USB MASS STORAGE DEVICE

Are all project directories and files stored in a top-level directory named for the University-assigned shipment number?

Are there an equal number of barcode subdirectories to volumes listed on the packing list?

Are all volumes subdirectories named for the 14-digit barcode of the physical volume?

- Is that barcode on the University supplied packing list?

Is there one and only one image file for each page of each volume?

Are all image files named with proper 8-digit file names?

Do all image files have a proper file extension (either .tif or .jp2)?

Does the sequence of image files for a volume start with 00000001 and increment sequentially for each thereafter?

- Are there appropriate missing page targets in the sequence for volumes with missing pages?

Are all bitonal images in TIFF format?

- Are all bitonal TIFFs in CCITT Group 4 compression?
- Do all bitonal TIFFs have the appropriate TIFF header tags with accurate information?

Are all continuous tone images (color or grayscale) in JPEG 2000 format?

- Do the files meet all of the technical requirements of the specifications for JPEG 2000?
- Is there a UUID box in the JPEG 2000 header? Is there an XMP packet in that UUID box?

Do all images meet the quality and technical parameters required?

Is there a notes.txt file in each barcode subdirectory?

Is there an aiim.tif image file in each barcode subdirectory?

- does the image file have a time and date stamp that indicates it was scanned along with the other images?



MDP PROJECT DELIVERED ON OPTICAL DISCS

Is the optical disc appropriate labeled both physically and electronically?

- Does the physical label have the date the disc was burned and the University-assigned shipment number?
- Does the disc report the same label as appears on the physical label?

Is the optical disc in ISO-9660 format?

Is the optical disc appropriate media for the data shipment size?

Are all top-level directories barcode-named directories?

Are there an equal number of barcode subdirectories to volumes listed on the packing list?

Are all volumes subdirectories named for the 14-digit barcode of the physical volume?

- Is that barcode on the University supplied packing list?

Is there one and only one image file for each page of each volume?

Are all image files named with proper 8-digit file names?

Do all image files have a proper file extension (either .tif or .jp2)?

Does the sequence of image files for a volume start with 00000001 and increment sequentially for each thereafter?

- Are there appropriate missing page targets in the sequence for volumes with missing pages?

Are all bitonal images in TIFF format?

- Are all bitonal TIFFs in CCITT Group 4 compression?
- Do all bitonal TIFFs have the appropriate TIFF header tags with accurate information?

Are all continuous tone images (color or grayscale) in JPEG 2000 format?

- Do the files meet all of the technical requirements of the specifications for JPEG 2000?
- Is there a UUID box in the JPEG 2000 header? Is there an XMP packet in that UUID box?

Do all images meet the quality and technical parameters required?

Is there a notes.txt file in each barcode subdirectory?

Is there an aiim.tif image file in each barcode subdirectory?

- does the image file have a time and date stamp that indicates it was scanned along with the other images?



BIBID PROJECT DELIVERED ON USB MASS STORAGE DEVICE

Are all project directories and files stored in a top-level directory named for the University-assigned shipment number?

Are all subdirectories of the shipment directory named using the BibID indicated on the production note target?

- Do all BibID directories have minimally one volume (four-digit) subdirectory?
- Do all volume (four-digit) subdirectories have minimally one issue (three-digit) subdirectory?

Is there one bitonal image file for each page of each volume?

Are all bitonal image files named with proper 8-digit file names?

Do all bitonal image files have a proper .tif file extension?

Does the sequence of image files for a volume start with 00000001 and increment sequentially for each thereafter?

- Are there appropriate missing page targets in the sequence for volumes with missing pages?

Are all bitonal images in TIFF format?

- Are all bitonal TIFFs in CCITT Group 4 compression?
- Do all bitonal TIFFs have the appropriate TIFF header tags with accurate information?

For volumes with contone images, is there a subdirectory of the issue (three-digit) subdirectory named “pages”?

- Are all contone images in the “pages” subdirectory?
- Are all contone image files named according to standard (with a leading “p” character and the same number as the corresponding bitonal page image)?

Are all continuous tone images (color or grayscale) in TIFF or JPEG 2000 format? (All contone files must be in one format or the other)

- Are all contone TIFFs uncompressed?
- Do all contone TIFFs have the appropriate TIFF header tags with accurate information?

or

- Do the files meet all of the technical requirements of the specifications for JPEG 2000?
- Is there a UUID box in the JPEG 2000 header? Is there an XMP packet in that UUID box?

Do all images meet the quality and technical parameters required?

Is there a notes.txt file in each barcode subdirectory?



Is there a bib.tif target image file in each subdirectory?

Is there a prodnote.tif target image file in each subdirectory?

Is there an aiim.tif image file in each barcode subdirectory?

- does the image file have a time and date stamp that indicates it was scanned along with the other images?

BIBID PROJECT DELIVERED ON OPTICAL DISCS

Is the optical disc appropriate labeled both physically and electronically?

- Does the physical label have the date the disc was burned and the University-assigned shipment number?
- Does the disc report the same label as appears on the physical label?

Is the optical disc in ISO-9660 format?

Is the optical disc appropriate media for the data shipment size?

Are all top-level directories named using the BibID indicated on the production note target?

- Do all BibID directories have minimally one volume (four-digit) subdirectory?
- Do all volume (four-digit) subdirectories have minimally one issue (three-digit) subdirectory?

Is there one bitonal image file for each page of each volume?

Are all bitonal image files named with proper 8-digit file names?

Do all bitonal image files have a proper .tif file extension?

Does the sequence of image files for a volume start with 00000001 and increment sequentially for each thereafter?

- Are there appropriate missing page targets in the sequence for volumes with missing pages?

Are all bitonal images in TIFF format?

- Are all bitonal TIFFs in CCITT Group 4 compression?
- Do all bitonal TIFFs have the appropriate TIFF header tags with accurate information?

For volumes with contone images, is there a subdirectory of the issue (three-digit) subdirectory named “pages”?

- Are all contone images in the “pages” subdirectory?
- Are all contone image files named according to standard (with a leading “p” character and the same number as the corresponding bitonal page image)?



Are all continuous tone images (color or grayscale) in TIFF or JPEG 2000 format? (All contone files must be in one format or the other)

- Are all contone TIFFs uncompressed?
- Do all contone TIFFs have the appropriate TIFF header tags with accurate information?

or

- Do the files meet all of the technical requirements of the specifications for JPEG 2000?
- Is there a UUID box in the JPEG 2000 header? Is there an XMP packet in that UUID box?

Do all images meet the quality and technical parameters required?

Is there a notes.txt file in each barcode subdirectory?

Is there a bib.tif target image file in each subdirectory?

Is there a prodnote.tif target image file in each subdirectory?

Is there an aiim.tif image file in each barcode subdirectory?

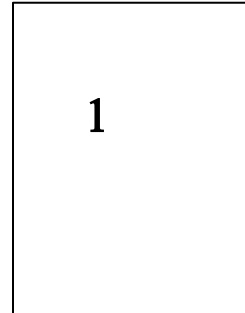
- does the image file have a time and date stamp that indicates it was scanned along with the other images?



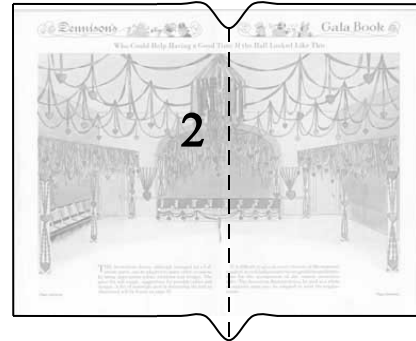
APPENDIX B: FOLDOUTS, CENTERFOLDS, OR DOUBLE PAGE SPREADS

The following appendix diagram demonstrates how to scan centerfold diagram, two-page spread in order to capture the full image and maintain the proper recto/verso order of the volume

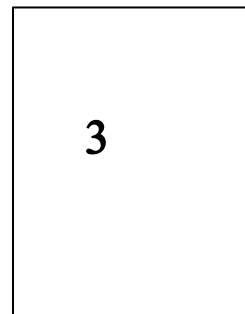
Blank back of left side of centerfold



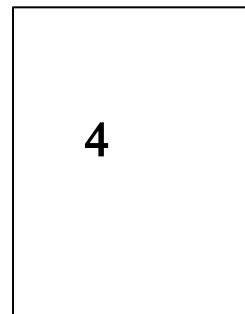
Centerfold in one image (if less than 11 x 17)



Blank back of right side of centerfold



Blank page inserted to maintain recto and verso





APPENDIX C: FILE STRUCTURE ORGANIZATION FOR MICHIGAN DIGITIZATION PROJECT SHIPMENTS

The following section provides a sketch of the file system structure required of the digital product for shipments designated as Michigan Digitization Project (MDP) shipments. This basic file system structure is required regardless of the delivery format chosen (i.e. optical disk or USB hard disk drive); however, the top level directory differs between the hard disk drive delivery method and the optical disc delivery method. This difference is described in detail in the following sections.

FILE STRUCTURE ORGANIZATION FOR HARD DISK DRIVE DATA DELIVERY

For hard disk delivery, this file structure requires that first-level directories to be named according to the University shipment number assigned to that shipment of volumes. The first-level directory must appear on the highest level of the file system hierarchy. If data delivery is accomplished on hard disk drives, multiple shipments may be included on the drive as space on media permits; as a result, multiple shipment number directories may appear on this top level. Shipments placed on the hard disk drive must be complete; if remaining space does not allow a complete shipment to fit on the drive provided, then the Digital Conversion Unit should be contacted immediately for direction.

For hard disk delivery, the files structure requires that subdirectories are created for each volume in the shipment and named according to the 14-digit barcode that appears on the inside of the back cover of the volume. On occasion, a volume may be encountered which has a barcode which appears in another location (e.g. on the back cover) or which has two barcodes. In most cases, the invalid barcode will be “crossed out” or otherwise marred to prevent its use. In all cases, the barcode on the inside back cover should be used as the default barcode to name the volume directory. If the vendor encounters a situation which it cannot clearly discern what barcode is appropriate, then the Digital Conversion Unit should be contacted immediately for direction.

Within those barcode subdirectories, the page image files for the pages associated with that volume should be stored and named according to the pattern specified elsewhere in this document.

The following table sketches out a sample picture of how this file structure should appear for data delivery on hard disk drive:

<Shipment Number>	(Directory)	(Shipment number for all volumes contained within)
<Barcode>	(Directory)	(Barcode of 1 st volume of shipment scanned)
00000001.tif		(TIFF page image file; filename is 8 characters long)
00000002.tif		(TIFF page image file; filename is 8 characters long)
00000003.tif		
00000004.tif		
00000005.jp2		(JPEG 2000 color or grayscale page image; filename is 8 characters long)



00000006.tif	
00000007.tif	
00000008.tif	
checksum.md5	(MD5 checksum signature file for this subdirectory; filename is exact)
aiim.tif	(Scanning target scanned in same session as volume; filename is exact)
notes.txt	(Text identifying scanning issues with this volume; filename is exact)
<Barcode> (Directory)	(Barcode of 2nd volume of shipment scanned)
00000001.tif	(TIFF page image file; filename is 8 characters long)
00000002.tif	
00000003.tif	
00000004.tif	
....	
00000032.tif	
00000033.tif	
00000034.jp2	(JPEG 2000 color or grayscale page image; filename is 8 characters long)
00000035.jp2	(JPEG 2000 color or grayscale page image; filename is 8 characters long)
00000036.tif	
00000037.tif	
checksum.md5	(MD5 checksum signature file for this subdirectory; filename is exact)
aiim.tif	(Scanning target scanned in same session as volume; filename is exact)
notes.txt	(Text identifying scanning issues with this volume; filename is exact)

FILE STRUCTURE ORGANIZATION FOR OPTICAL DISC DATA DELIVERY

For optical disc delivery, this file structure eliminates the first-level directories named according to the University shipment number. Each shipment must have its own set of optical discs to deliver data, if this is the delivery method chosen. Each optical disc must be clearly labeled on its face as to the shipment number, the date the disc was burned, and the optical disc code specified for that shipment by the Digital Conversion Unit.

For optical disc delivery, the files structure requires that the first-level directories are created for each volume in the shipment and named according to the 14-digit barcode that appears on the inside of the back cover of the volume. On occasion, a volume may be encountered which has a barcode which appears in another location (e.g. on the back cover) or which has two barcodes. In most cases, the invalid barcode will be “crossed out” or otherwise marred to prevent its use. In all cases, the barcode on the inside back cover should be used as the default barcode to name the volume directory. If the vendor encounters a situation which it cannot clearly discern what



barcode is appropriate, then the Digital Conversion Unit should be contacted immediately for direction.

For those shipments whose scanned data cannot fit onto a single optical disc, the vendor may break out the shipment volumes onto multiple optical discs. Optical discs must be filled as close to capacity as possible. Each volume's page images burned to a disc must be a complete data set; the vendor may not burn part of a volume's page images onto one disc and the remainder onto another.

Within those barcode directories, the page image files for the pages associated with that volume should be stored and named according to the pattern specified elsewhere in this document.

The following table sketches out a sample picture of how this file structure should appear for data delivery on optical disc:

<Barcode>	(Directory)	(Barcode of 1 st volume of shipment scanned)
00000001.tif		(TIFF page image file; filename is 8 characters long)
00000002.tif		(TIFF page image file; filename is 8 characters long)
00000003.tif		
00000004.tif		
00000005.jp2		(JPEG 2000 color or grayscale page image; filename is 8 characters long)
00000006.tif		
00000007.tif		
00000008.tif		
checksum.md5		(MD5 checksum signature file for this subdirectory; filename is exact)
aiim.tif		(Scanning target scanned in same session as volume; filename is exact)
notes.txt		(Text identifying scanning issues with this volume; filename is exact)
<Barcode>	(Directory)	(Barcode of 2nd volume of shipment scanned)
00000001.tif		(TIFF page image file; filename is 8 characters long)
00000002.tif		
00000003.tif		
00000004.tif		
....		
00000032.tif		
00000033.tif		
00000034.jp2		(JPEG 2000 color or grayscale page image; filename is 8 characters long)
00000035.jp2		(JPEG 2000 color or grayscale page image; filename is 8 characters long)
00000036.tif		
00000037.tif		



- checksum.md5 (MD5 checksum signature file for this subdirectory; filename is exact)
- aiim.tif (Scanning target scanned in same session as volume; filename is exact)
- notes.txt (Text identifying scanning issues with this volume; filename is exact)

In organizing the contents of digital files on CD's or DVD's burned in various digital reformatting projects, the University of Michigan Library has established prescriptive guidelines that must be followed in order for CD's or DVD's to be accepted in fulfillment of contractual guidelines, as well as to guide the work of internal staff who undertake such work. These guidelines conform to ISO-9660 and as such, should be able to be read on a variety of hardware devices. In addition, they allow us to anticipate the structure of directories and files and to develop generic scripts for loading them to our servers.

Name		Notes	Bitonal Examples
CD volume name	Project abbreviation and CD sequence number	Each project is assigned a 3 or 4 character code; the volume name is comprised of this code, an _ (underscore), and the number of the CD in the project series with an appropriate number of leading zeros to make an 8 character volume name. If the label includes a year, it should be represented by the 2nd half of the fiscal year. (ie. Fiscal year 2004-2005 would be pd05_001...)	pz07_001 pi07_038 pd07_001
Directory Level 1	Volume Identifier	Each volume will have its own unique barcode. It can usually be located on the inside of the back cover or in the endsheets.	3901500005479 3901523659417 3901500342256
Individual Image Files	File names	8 places followed by a .tif or .jp2 extension assigned automatically following the sequence of page images scanned.	00000001.tif 00000235.jp2
Additional Files	File names	Files associated with the scanning process, all 8 character filenames	aiim.tif rit.tif (optional) notes.txt checksum.md5 bib.tif (special projects) prodnote.tif (special projects)

Any questions, clarifications, concerns, or problems experienced by the vendors in producing this file structure should be communicated to the Digital Production Services Manager as soon as possible.



APPENDIX D: FILE STRUCTURE ORGANIZATION FOR BIBLIOGRAPHIC IDENTIFIER SHIPMENTS

The following section provides a sketch of the file system structure required of the digital product for shipments designated as Bibliographic Identifier (BibID) shipments. This basic file system structure is required regardless of the delivery format chosen (i.e. optical disk or USB hard disk drive); however, the top level directory differs between the hard disk drive delivery method and the optical disc delivery method. This difference is described in detail in the following sections.

FILE STRUCTURE ORGANIZATION FOR HARD DISK DRIVE DATA DELIVERY

For hard disk delivery, this file structure requires that first-level directories to be named according to the University shipment number assigned to that shipment of volumes. The first-level directory must appear on the highest level of the file system hierarchy. If data delivery is accomplished on hard disk drives, multiple shipments may be included on the drive as space on media permits; as a result, multiple shipment number directories may appear on this top level. Shipments placed on the hard disk drive must be complete; if remaining space does not allow a complete shipment to fit on the drive provided, then the Digital Conversion Unit should be contacted immediately for direction.

For hard disk delivery, the files structure requires that subdirectories are created for each bibliographic entity in the shipment and named in a directory hierarchy according to the designation which appears on the production notes target prepared for that physical volume or for the section of the physical volume by the Digital Conversion Unit.

Within this directory hierarchy, the page image files for the pages associated with that volume should be stored and named according to the pattern specified elsewhere in this document.

The following table sketches out a sample picture of how this file structure should appear for data delivery on hard disk drive:

<Shipment Number> within)	(Directory)	(Shipment number for all volumes contained within)
<Volume #####>	(Directory)	(“Volume” level subdirectory; directory name must be four (4) digit number; Enumeration will be specified in bibliographic target)
<Issue ###>	(Directory)	(“Issue” level subdirectory; directory name must be three (3) digit number; Enumeration will be specified in bibliographic target)
00000001.tif		(TIFF page image file; filename is 8 characters long)
00000002.tif		(TIFF page image file; filename is 8 characters long)
00000003.tif		



00000004.tif	
00000005.jp2	(JPEG 2000 color or grayscale page image; filename is 8 characters long)
00000006.tif	
00000007.tif	
00000008.tif	
checksum.md5	(MD5 checksum signature file for this subdirectory; filename is exact)
aiim.tif	(Scanning target scanned in same session as volume; filename is exact)
bib.tif	(Bibliographic target scanned in same session as volume; filename is exact)
notes.txt	(Text identifying scanning issues with this volume; filename is exact)
<Volume #####>	(Directory) (“Volume” level subdirectory; directory name must be four (4) digit number; Enumeration will be specified in bibliographic target)
<Issue ###>	(Directory) (“Issue” level subdirectory; directory name must be three (3) digit number; Enumeration will be specified in bibliographic target)
long)	
00000001.tif	(TIFF page image file; filename is 8 characters
00000002.tif	
00000003.tif	
00000004.tif	
....	
00000032.tif	
00000033.tif	
00000034.jp2	(JPEG 2000 color or grayscale page image; filename is 8 characters long)
00000035.jp2	(JPEG 2000 color or grayscale page image; filename is 8 characters long)
00000036.tif	
00000037.tif	
checksum.md5	(MD5 checksum signature file for this subdirectory; filename is exact)
aiim.tif	(Scanning target scanned in same session as volume; filename is exact)
bib.tif	(Bibliographic target scanned in same session as volume; filename is exact)
notes.txt	(Text identifying scanning issues with this volume; filename is exact)



FILE STRUCTURE ORGANIZATION FOR OPTICAL DISC DATA DELIVERY

For optical disc delivery, this file structure eliminates the first-level directories named according to the University shipment number. Each shipment must have its own set of optical discs to deliver data, if this is the delivery method chosen. Each optical disc must be clearly labeled on its face as to the shipment number, the date the disc was burned, and the optical disc code specified for that shipment by the Digital Conversion Unit.

For optical disc delivery, the files structure requires that the first-level directories are created for each volume in the shipment and named according to the 14-digit barcode that appears on the inside of the back cover of the volume. On occasion, a volume may be encountered which has a barcode which appears in another location (e.g. on the back cover) or which has two barcodes. In most cases, the invalid barcode will be “crossed out” or otherwise marred to prevent its use. In all cases, the barcode on the inside back cover should be used as the default barcode to name the volume directory. If the vendor encounters a situation which it cannot clearly discern what barcode is appropriate, then the Digital Conversion Unit should be contacted immediately for direction.

For those shipments whose scanned data cannot fit onto a single optical disc, the vendor may break out the shipment volumes onto multiple optical discs. Optical discs must be filled as close to capacity as possible. Each volume’s page images burned to a disc must be a complete data set; the vendor may not burn part of a volume’s page images onto one disc and the remainder onto another.

Within those barcode directories, the page image files for the pages associated with that volume should be stored and named according to the pattern specified elsewhere in this document.

The following table sketches out a sample picture of how this file structure should appear for data delivery on optical disc:

<Barcode>	(Directory)	(Barcode of 1 st volume of shipment scanned)
00000001.tif		(TIFF page image file; filename is 8 characters long)
00000002.tif		(TIFF page image file; filename is 8 characters long)
00000003.tif		
00000004.tif		
00000005.jp2		(JPEG 2000 color or grayscale page image; filename is 8 characters long)
00000006.tif		
00000007.tif		
00000008.tif		
checksum.md5		(MD5 checksum signature file for this subdirectory; filename is exact)
aiim.tif		(Scanning target scanned in same session as volume; filename is exact)
notes.txt		(Text identifying scanning issues with this volume; filename is exact)
<Barcode>	(Directory)	(Barcode of 2nd volume of shipment scanned)



00000001.tif (TIFF page image file; filename is 8 characters long)
 00000002.tif
 00000003.tif
 00000004.tif

 00000032.tif
 00000033.tif
 00000034.jp2 (JPEG 2000 color or grayscale page image; filename is 8 characters long)
 00000035.jp2 (JPEG 2000 color or grayscale page image; filename is 8 characters long)
 00000036.tif
 00000037.tif
 checksum.md5 (MD5 checksum signature file for this subdirectory; filename is exact)
 aaim.tif (Scanning target scanned in same session as volume; filename is exact)
 notes.txt (Text identifying scanning issues with this volume; filename is exact)

In organizing the contents of digital files on CD's or DVD's burned in various digital reformatting projects, the University of Michigan Library has established prescriptive guidelines that must be followed in order for CD's or DVD's to be accepted in fulfillment of contractual guidelines, as well as to guide the work of internal staff who undertake such work. These guidelines conform to ISO-9660 and as such, should be able to be read on a variety of hardware devices. In addition, they allow us to anticipate the structure of directories and files and to develop generic scripts for loading them to our servers.

Name		Notes	Bitonal Examples	Grayscale/Color Explanation
CD volume name	Project abbreviation and CD sequence number	Bitonal: Each project is assigned a 3 or 4 character code; the volume name is comprised of this code, an _ (underscore), and the number of the CD in the project series with an appropriate number of leading zeros to make an 8 character volume name. If the label includes a year, it should be represented by the 2nd half of the fiscal year. (ie. Fiscal year 2004-2005 would be pd05_001...)	nsf_0001 moa_0238 pd05_001	
		Grayscale/Color: shall be burned to a separate CDs than the bitonal scans that they accompany. These CD's should have a letter designation at the end and delete either a 0 or the _ to maintain an 8 character cd label. Follow the same guidelines as specified for bitonal images (3 levels) plus a fourth level containing a pages directory or a figure (figs) directory. These directories are only used if their are files to populate them. Further description is below.	pd05_0038	pd05_038a pd05_038b pd05_038c up to 26 CDs accompanying one bitonal CD



Directory Level 1	Bibliographic Identifier	Each bibliographic entity will have a unique identifier supplied by the library. In most cases it will take the form 0366299 but can have as many as 8 characters should another form be used.	ABN7325 b1407971 0342256	ABN7325 b1407971 0352256
Directory Level 2	Volume Sub-ID Level 1	Four numeric characters; used to designate volumes or years comprising a bibliographic volume. If not necessary, use 0001 as a default	0001 0002 etc. 1921	0001 0002 etc. 1921
Directory Level 3	Volume Sub-ID Level 2	Three numeric characters; used to designate sub-parts or issues that comprise a volume. If not necessary, use 001 as a default.	001 002 etc.	001 002 etc.
Directory Level 4	Container for grayscale and color images	The fourth directory involves only grayscale and color images. These images accompany a set of bitonal page images and are divided into two types of image capture. One called pages where the entire page is captured as a grayscale or color image. The other is called a figure (figs) , where the page is mixed content (text and illustration) and only the illustrations are captured as grayscale or color images. These directories are only necessary if there are files to populate them.	/001/00000023.tif /003/00000189.tif /005/00000012.tif /001/00000156.tif	/001/pages/p0000023.tif /003/pages/p0000189.tif /005/figs/f0000012.tif /001/figs/f0000156.tif
Individual Image Files	File names	8 places followed by a .tif extension assigned automatically following the sequence of page images scanned	00000001.tif ... 00000235.tif	
Individual Image Files for the accompanying grayscale and color images	File names	In Directory level 3, a full set of bitonal images are captured regardless of image content. Where grayscale or color is required to capture the illustration a corresponding grayscale or color file is saved to one of the two Directory level 4 folders. To distinguish between these two types each file begins with either "p" or "f" and saved to the "pages" or "figs" folder.		p0000015.tif p0000185.tif f0000258.tif f0000306.tif
Single volume split on multiple CDs		These instructions apply to volumes, parts, or illustrations where all of the images will not fit on a single CD. Use a "c" appended to the end of Directory Level 3 .	001c	001c
		Examples: pc04_034 contains the first 30 pages of volume 1, part 1 of 0804030. Part 1 is continued on pc04_035.	pc04_034 0804030 0001 001 00000001.tif to 00000030.tif pc04_035 0804030 0001 001c 00000031.tif to 000000xx.tif	pc04_34a 0804030 0001 001 pages p0000001.tif to p0000015.tif pc04_34b 0804030 0001 001c pages p0000016.tif to p0000030.tif



Additional Files	File names	Files associated with the scanning process, all 8 character filenames	aiim.tif; notes.txt checksum.md5 bib.tif prodnote.tif	

Any questions, clarifications, concerns, or problems experienced by the vendors in producing this file structure should be communicated to the Digital Production Services Manager as soon as possible.



APPENDIX E: METADATA TAGS FOR TIFF FILES

To help ensure the long-term sustainability of master images created, the University of Michigan proposes to use some of the data elements in Draft NISO Z39.87 –2002 Standard, *Data Dictionary—Technical Metadata for Digital Still Images* to record technical metadata in the TIFF header of each image. In addition to required TIFF tags, some additional tags are utilized to document the provenance of each image.

The elements identified in the table below should be incorporated into headers for the master TIFF images. These elements include:

- Required fields for Baseline TIFF grayscale images; these unshaded elements are required for a valid TIFF image.
- Required or recommended elements from the Z39.87 Data Dictionary where appropriate TIFF tags exist; these elements are shaded in the following table.
- Elements in italics should be considered recommended; all others are mandatory and will be subjected to quality review.

We continue to review these specifications and may make changes based on experience and/or changes to the Draft Z39.87 Data Dictionary.

TIFF Tag Number	TIFF/EXIF Tag Name	Value	Z39.87 Number	Z39.87 Name	Notes/Sample Values
256	ImageWidth		8.1.5	ImageWidth	
257	ImageLength		8.1.6	ImageLength	
258	BitsPerSample		8.2.1	BitsPerSample	
259	Compression		6.1.3.1	CompressionScheme	
262	Photometric Interpretation		6.1.4.1	ColorSpace	
269	DocumentName	<Reel Number>,<Subdirectory Number>/<Image File name>	6.2.1	ImageIdentifier	ABC1234.0001.001/0000001.TIF
271	<i>Make</i>	<i>Scanner manufacturer</i>	<i>7.6.1.1</i>	<i>ScannerManufacturer</i>	<i>Fujitsu</i>
272	<i>Model</i>	<i>Include model number</i>	<i>7.6.1.2.1-2</i>	<i>ScannerModel</i>	<i>4097D</i>
273	StripOffsets		6.1.5.2	StripOffsets	
274	Orientation		6.2.4	Orientation	
277	SamplesPerPixel		8.2.2	SamplesPerPixel	
278	RowsPerStrip		6.1.5.3	RowsPerStrip	
279	StripByteCounts		6.1.5.4	StripByteCounts	
282	XResolution		8.1.3	XSamplingFrequency	
283	YResolution		8.1.4	YSamplingFrequency	
284	Planar Configuration		6.1.6	PlanarConfiguration	
296	ResolutionUnit		8.1.2	SamplingFrequencyUnit	
305	<i>Software</i>	<i>Include version number</i>	<i>7.6.2.1-2</i>	<i>ScanningSoftware</i>	<i>Fujitsu ScandAll 21 V. 4.2.16</i>
306	DateTime	ISO8601 format YYYY:MM:DD HH:MM:SS	7.9	DateTimeCreated	2004:09:08 00:00:00
315	Artist	[scanning contractor company name]	7.3	ImageProducer	



APPENDIX F: XMP METADATA FOR JPEG2000 FILES

Any participating vendor in this pilot project is invited to submit the digital images produced in the JPEG2000 file format. In order to help facilitate the creation of acceptable JPEG2000 image files with appropriate metadata, the following discussion is offered.

XMP documents must be, by requirement of the XMP specification, to the UUID box of the JPEG2000 file header; it is unacceptable to write this XMP packet to the XML box of the JPEG2000 file header.

The following XML document represents an example of a full XMP packet written to the UUID box of a JPEG2000 image file. The document represents nearly exactly what must appear in the JPEG2000 file headers for files produced in this project, with the only changes being in some of the XML element values from file to file:

```
<?xpacket begin='' id='W5M0MpCehiHzreSzNTczkc9d'?>
<x:xmpmeta xmlns:x='adobe:ns:meta/'>
<rdf:RDF xmlns:rdf='http://www.w3.org/1999/02/22-rdf-
syntax-ns#'>
<rdf:Description rdf:about=''
xmlns:tiff='http://ns.adobe.com/tiff/1.0/'>
<tiff:ImageWidth>2359</tiff:ImageWidth>
<tiff:ImageLength>3229</tiff:ImageLength>
<tiff:BitsPerSample>8</tiff:BitsPerSample>
<tiff:Compression>34712</tiff:Compression>
<tiff:PhotometricInterpretation>1</tiff:PhotometricInt
erpretation>
<tiff:Orientation>1</tiff:Orientation>
<tiff:SamplesPerPixel>1</tiff:SamplesPerPixel>
<tiff:XResolution>300/1</tiff:XResolution>
<tiff:YResolution>300/1</tiff:YResolution>
<tiff:ResolutionUnit>2</tiff:ResolutionUnit>
<tiff:DateTime>2007-01-03
00:00:00+08:00</tiff:DateTime>
<tiff:Artist>University of Michigan - Digital
Conversion Unit</tiff:Artist>
<tiff:Make>Zeutschel</tiff:Make>
<tiff:Model>7000</tiff:Model>
</rdf:Description>
<rdf:Description rdf:about=''
xmlns:dc='http://purl.org/dc/elements/1.1/'>
<dc:source>12345.0002/00000021.jp2</dc:source>
</rdf:Description>
</rdf:RDF>
</x:xmpmeta>
```



<?xpacket end='w'?>

The following represents the same XMP document with annotations from the University of Michigan to explain the importance or relevance of individual elements or values. The XMP document itself is represented by text in the “Courier” typeface and the annotations by the University are represented by text in the “Times New Roman” typeface:

<?xpacket begin='' id='W5M0MpCehiHzreSzNTczkc9d'?>

- This is the XMP packet leader which XMP scanning application search for to identify the presence of XMP metadata in a digital file; it is required by the XMP specification and must be present for the metadata to be validly encoded in the image file

```
<x:xmpmeta xmlns:x='adobe:ns:meta/'>
<rdf:RDF xmlns:rdf='http://www.w3.org/1999/02/22-rdf-
syntax-ns#'>
<rdf:Description rdf:about=''
xmlns:tiff='http://ns.adobe.com/tiff/1.0/'>
```

- These elements define the structure of the XMP packet and the namespaces it is working in; it must be present to validate the XML elements which follow.

```
<tiff:ImageWidth>2359</tiff:ImageWidth>
```

- This element represents the number of pixels in the width of the image bitmap; the value here should be identical with the value of TIFF tag number 256; the value must be numeric; it is expected to vary somewhat from image to image.

```
<tiff:ImageLength>3229</tiff:ImageLength>
```

- This element represents the number of pixels in the height of the image bitmap; the value here should be identical with the value of TIFF tag number 257; the value must be numeric; it is expected to vary somewhat from image to image.

```
<tiff:BitsPerSample>8</tiff:BitsPerSample>
```

- This element represents the number of bits per pixel (sample) in the image bitmap; the value here should be identical with the value of TIFF tag number 258; the value must be numeric and must be “8” for all 8-bit grayscale images.

```
<tiff:Compression>34712</tiff:Compression>
```

- This element represents the code number for the image format; the value must be numeric and must be “34712” for all JPEG2000 images.

```
<tiff:PhotometricInterpretation>1</tiff:PhotometricInt
erpretation>
```

- This element represents the color space of the image bitmap; the value here should be identical with the value of TIFF tag number 262; the value must be numeric and must be “1” for all grayscale images.

```
<tiff:Orientation>1</tiff:Orientation>
```



- This element represents the orientation of the image bitmap in its default presentation; the value must be numeric and must be “1” for all images; the value “1” means that the 0th row represents the visual top of the image, and the 0th column represents the visual left-hand side.

```
<tiff:SamplesPerPixel>1</tiff:SamplesPerPixel>
```

- This element represents the number of color samples for each pixel of the image bitmap; the value here should be identical with the value of TIFF tag number 277; the value must be numeric and must be “1” for all grayscale images and “3” for all color images.

```
<tiff:XResolution>300/1</tiff:XResolution>
```

- This element represents the number of pixels per the number of resolution units in the resolution of the image width; the value here should be identical with the value of TIFF tag number 282; the value must be numeric with a dividing slash followed by another numeric; the requirements for this project state that this value must be “600/1,” “400/1,” or “300/1” whenever that resolution is possible; the X resolution and the Y resolution should always be the same value.

```
<tiff:YResolution>300/1</tiff:YResolution>
```

- This element represents the number of pixels per the number of resolution units in the resolution of the image height; the value here should be identical with the value of TIFF tag number 283; the value must be numeric with a dividing slash followed by another numeric; the requirements for this project state that this value must be “600/1,” “400/1,” or “300/1” whenever that resolution is possible; the X resolution and the Y resolution should always be the same value.

```
<tiff:ResolutionUnit>2</tiff:ResolutionUnit>
```

- This element defines what is the resolution unit; the value must be numeric and must be “2” for all images produced for University shipments; the value here should be identical with the value of TIFF tag number 296; the value “2” means the resolution unit is inches.

```
<tiff:DateTime>2007-01-03  
00:00:00+08:00</tiff:DateTime>
```

- This element records the date and time of image capture; the value must be in ISO8601 format.

```
<tiff:Artist>University of Michigan - Digital  
Conversion Unit</tiff:Artist>
```

- This element records who captured this digital image; the value is textual and should be the name of the vendor or scanning unit.

```
<tiff:Make>Zeutschel</tiff:Make>
```

- This element records what make of scanning equipment captured this digital image; the value here should be identical with the value of TIFF tag number 271; the value is textual and should be the name of the manufacturer of the scanning unit.

```
<tiff:Model>OS7000</tiff:Model>
```




- This element records what model of scanning equipment captured this digital image; the value here should be identical with the value of TIFF tag number 272; the value is textual and should be the name or serial number of the model of the scanning unit.

```
</rdf:Description>  
<rdf:Description rdf:about=' '  
xmlns:dc='http://purl.org/dc/elements/1.1/'>
```

- These elements define the structure of the following element and the namespaces it is working in; it must be present to validate the XML element which follows.

```
<dc:source>12345.0002/00000021.jp2</dc:source>
```

- This element records the provenance of the image; the value here should be identical with the value of TIFF tag number 269; within UM shipments, the value must follow either the format of <Barcode>/<Image File Name> or the format of <BibID>.<Subdirectory #####>.<Subdirectory #####>/<Image File Name>.

```
</rdf:Description>  
</rdf:RDF>  
</x:xmpmeta>  
<?xpacket end='w'?>
```

- These elements close the structure of the XMP packet; it must be present to close the XMP packet within the header.

Vendors seeking greater clarity on the XMP metadata specification should consult the documentation at the following URL:

<http://partners.adobe.com/public/developer/en/xmp/sdk/XMPspecification.pdf>

Vendors seeking greater clarity on the definitions for particular XML elements from the TIFF namespace should consult the TIFF 6.0 documentation at the following URL:

<http://partners.adobe.com/public/developer/en/tiff/TIFF6.pdf>

Any questions, clarifications, concerns, or problems experienced by the vendors in producing this XMP packet in the JPEG2000 image files should be communicated to the Digital Production Services Manager as soon as possible.



APPENDIX G: SAMPLE SCANNING NOTES TEXT FILE

Vendors are expected to record any peculiarities or difficulties in the scanning of materials for all projects in a scanning notes text file. This file must be present in the data delivered back to the University even if there is nothing to report about the scanning of that particular volume or item. Notes that should appear in this file should include missing pages, pages that cannot be imaged at expected resolutions or cannot be imaged clearly due to defacement. All post-processing image cleanup software should also appear in this file.

The file name for this file must be “notes.txt” without the quotation marks.

Sample Scanning Notes Text File Content:

Scanning Notes:

Item ID: <Barcode> or <BibID>

Scanning Problems:

00000057.tif - Unable to achieve 600ppi on foldout; 400ppi resolution used

00000146.tif - Original page is poor quality; unable to obtain clear image

00000181.tif - Missing page; target inserted

00000182.tif - Missing page; target inserted

00000222.tif - Page broken off through text; remainder of page scanned

00000236.tif - Page defaced; interference for clear imaging of printed text.

Post-Processing:

Adobe Photoshop CS2

Custom Deskew Software



APPENDIX H: SAMPLE BIBLIOGRAPHIC TARGET

Below is a screenshot of a sample bibliographic record target. These targets should only appear in projects designated as BibID projects. If a vendor encounter an item in a BibID project without such a target, they should call the University for assistance.

This target must be scanned by the same scanning equipment under the same configuration as the rest of the volume. The scanned target image must be named “bib.tif” and it should be scanned as a bitonal image saved as a TIFF format file in CCITT Group 4 compression

BIBLIOGRAPHIC RECORD TARGET

**Graduate Library
University of Michigan
Digital Conversion Unit**

AAG2862

035/1: : |a (RLIN)MIUG83-S19540
035/2: : |a (CaOTULAS)159943359
040: : |a MiU |c MiU
110:1 : |a Michigan. |b Fisheries Division.
245:10: |a Fisheries research report.
260: : |a [n.p.]
300/1: : |a v.
310: : |a Irregular
362/1:0 : |a no. 1- 1930-
515/1: : |a Some volumes issued in revised editions
550/2: : |a Nos. for 1930- issued by the Institute for Fisheries Research; by the Michigan Fisheries Division.
580/3: : |a Electronic serial mode of access: World Wide Web via the Michigan DNR, Institute for Fisheries Research site.
590/4: : |a Microfiche copy: Microfiche. Washington, D.C. : U.S. Govt. Print. Off.
microfiches : negative ; 11 x 15 cm.
650/1:0: |a Fisheries |z Michigan.
710/1:1 : |a Michigan. |b Institute for Fisheries Research.
740/2:0 : |a Fisheries research report (Online)
856/1:41: |z Selected issues available online: |u
http://www.michigan.gov/dnr/0,1607,7-153-10364_10951_19056-46188-,00.html
998/1: : |c LB |s 9124



APPENDIX I: SAMPLE PRODUCTION NOTE TARGET

Below is a screenshot of a sample production note target. These targets should only appear in projects designated as BibID projects. If a vendor encounter an item in a BibID project without such a target, they should call the University for assistance.

This target provides the vendor valuable information for the file structure. The MirlynID field at the top is the bibliographic ID. The Volume field at the top of the target indicates the enumeration expected in the file directory structure hierarchy. For example, with this sample image, the following file directory hierarchy would be expected:

```
ACD6054 <Directory>  
  -0008 <Directory>  
    -001 <Directory>
```

This target must be scanned by the same scanning equipment under the same configuration as the rest of the volume. The scanned target image must be named “bib.tif” and it should be scanned as a bitonal image saved as a TIFF format file in CCITT Group 4 compression



Production Notes

Sub Project

Mirlyn ID: ACD6054

Volume: 8, Pt. 1

Call number: AP33 S32

Total Pages: 703

Scanning Notes

Item Barcode: _____

Irregular Pagination Photo glued into book on page 189

Missing Pages

Greyscale: 10 or less (give page #'s)

Approximate Number of Greyscale Images: _____

Page 189, added to book

Color: No

Approximate Number of Color Images: _____

Foldouts: No

Pagination

Other Production Notes

Roman numerals I - XIX and 1 blank page before regular pagination