ATTENTION:
This document issued in conjunction with the NEH National Digital Newspaper Program (http://www.neh.gov/grants/guidelines/ndnp.html) award competition. For general program information, see http://www.loc.gov/ndnp/ .

The National Digital Newspaper Program (NDNP)
Technical Guidelines for Applicants

Overview of Technical Approach for 2008-10 NDNP Awards

The National Digital Newspaper Program is a long-term effort and the technical environment will change as the program continues. The National Endowment for the Humanities (NEH) and the Library of Congress (LC) have selected a technical approach to balance long-term objectives and shorter-term constraints. These include:

- convenient accessibility over the World Wide Web for the general public to the entire collection as it grows, through a consistent interface and using proven technology;
- page images of sufficient spatial and tonal resolution to support effective performance of OCR (optical character recognition) software and representation of printed half-tones, given the limitations of microfilm, expecting that future improvements in OCR and image processing will be applied to the same images;
- the use of digital formats with a high probability of sustainability - in particular, using standard formats where possible and proprietary formats only where widely adopted;
- and attention to the cost of digital conversion and maintenance of the resulting assets.

The goal of the initial program phase is to build a Web-accessible NDNP delivery application with sufficient geographic coverage and digital assets to validate the technical approach and to serve as a test bed for future research and development for techniques to enhance the content and access interface, to support effective use by scholars and the general public. This award cycle is a continuation of the initial program development phase.

In succeeding phases of the project, the approach and associated guidelines will be evaluated and revised based on feedback from awardees, experience in providing access to historic newspapers online, and technological advances.

In summary, the current technical approach is based on:

- grayscale images (scanned for maximum resolution possible between 300-400 dpi, relative to the original material) from microfilm
- OCR with word-bounding boxes, uncorrected, with recognition of columns, but without segmentation of pages into articles,
- structural metadata for pages, issues, editions, and titles to support a chronologically-based browsing interface,
- copies of all page images and associated metadata at LC,
- an interface designed specifically for access to historic newspapers in the public domain, mounted at LC (the initial interface will permit full-text searches with retrieval of individual page images, and highlighting of search words on the images), and
- the ability of awardees to re-use any digital assets created for NDNP in other systems or for other purposes.

NEH and LC recognize that other institutions may choose other approaches or formats for their own digital repository and delivery systems and thus either weigh costs and benefits differently or wish for compatibility to existing systems. Applicants may pursue local approaches in parallel with participation in NDNP, with the overall goal of providing effective widespread access to newspapers through scanning and text conversion and evaluating alternative interfaces for navigating and exploring large collections of newspapers. Applicants who use other formats locally must be capable of providing digital assets to the NDNP according to the specifications described below.

The National Digital Newspaper Program supports a consistent technical specification for digital newspaper reproductions and associated metadata in order to maintain parity of services for materials from a variety of institutions and collections and to support the "best practices" of today's understanding of digital preservation needs.

Deliverables

Awardees are expected to deliver the following to the Library of Congress, to allow construction of a permanent archive and a unified interface for searching and browsing the entire NDNP collection. After the cooperative agreements are announced, LC will convene a meeting of awardees to review these technical guidelines, and establish work-plan milestones, and specifications for 2008-10 deliverables.

For each title
- Up-to-date MARC record from the CONSER database, fully conformant to current standards for cataloging U.S. print newspapers,
- Additional title-level metadata related to the title run/s digitized and delivered (see Appendix A: Digital Asset Metadata Elements), and
- Essay – scope and content of each title, history and significance – 500 words.

For each issue/edition
- Structural metadata for issues/editions digitized and organized by date (see Appendix A: Digital Asset Metadata Elements)

For each newspaper page

- Page image in two raster formats
  - Grayscale, scanned for maximum resolution possible between 300-400 dpi, relative to the original material, uncompressed TIFF 6.0 (see Scanning below and Appendix B – File Format Profiles),
  - Same image, compressed as JPEG2000 (see Scanning below and Appendix B – File Format Profiles),
- File with OCR text and associated bounding boxes for words (see OCR details below and Appendix B – File Format Profiles), 1 file per page image,
- PDF Image with Hidden Text, i.e., with text and image correlated (see OCR details below and Appendix B – File Format Profiles),
- Structural metadata to relate pages to title, date, and edition, sequence pages within issue or section; and to identify image and OCR files (see Appendix A: Digital Asset Metadata Elements and Appendix C – XML Metadata Templates), and
- Technical metadata to support the functions of a trusted repository (see Appendix A: Digital Asset Metadata Elements, Appendix B – File Format Profiles and Appendix C – XML Metadata Templates).

Awardees will deliver all digital assets in a METS object structure (Metadata Encoded Transmission Schema), according to an XML Batch template structure. (See Appendix C – XML Metadata Templates.)

For delivery, the awardee shall organize the page images and related files for each newspaper title in a hierarchical directory structure sufficient for identification of the individual digital assets from the metadata provided.

Further options and specifications for delivery will be specified at the initial 2008-10 awardees' meeting, post-award.

For each microfilm reel digitized:

- A second-generation (2N) duplicate silver negative microfilm, made from the camera master, will be barcoded and deposited with the Library of Congress on completion of the award (LC to supply barcodes for all reels), and
- Technical metadata concerning the quality characteristics of the film used for digitization (See Appendix A – Digital Asset Metadata Elements/Reel Information) will be encoded in a METS object with other digital assets (See Appendix C – XML Metadata Templates.).

NOTE: All digital objects must conform and validate to NDNP technical specifications as described in Appendices B and C. See **Technical Validation of Digital Objects** below, for more information.

Technical Details

Selection

The goals of the overall project, the chronological scope (1880-1922), and the intellectual criteria for selecting newspaper titles for this phase are described in the Request for

Proposal guidelines. To ensure the highest quality and most usable digital products and services, the process for selection of a newspaper title for inclusion in the NDNP should also incorporate a technical analysis of the microfilm to be scanned.

For NDNP (and the associated collection of duplicate microfilm negatives (2N)) to be as complete as possible, the following guidelines should be followed:

1. Complete (or majority of) title run should be available on microfilm without restrictions that interfere with the goals of the program;

2. An effort should be made to deliver as complete a title run as possible. Locating and substituting a limited number of scanned images from paper may be necessary to complete the run.

Several technical factors will affect the success of microfilm scanning and optical character recognition (OCR). The following factors should be considered during the selection process. They include:

1. The quality of original text and microfilm capture. Poorly prepared original material, no matter how well microfilmed, yields poor results. Microfilm of bound material may have page curvature, gutter shadows, or out of focus pages that influence digital image quality. Preference in selection should be given to titles on higher quality microfilm.

2. The reduction ratio used when microfilming the original newspaper. This ratio directly influences image quality and OCR results. The lower the reduction ratio (below 20x) the better. (If the reduction ratio is too high to allow scanning at 400 dpi, tests on sample images may be performed to determine if a lower resolution (e.g., 350 dpi) provides acceptable confidence levels in OCR text.)

3. The camera master negative microfilm duplicated for scanning should have resolution test patterns readable at 5.0 or higher. For camera master microfilm without resolution test charts, resolution can be estimated by comparison to film with resolution test charts and original material.

4. Variations in density within images and between exposures. Such variations require adjustment of scanning parameters within a reel. Density readings should follow current standards, but the range should ideally be narrower than the standards allow (e.g. .90-1.20). Best results are obtained from microfilm with variations in density readings of no more than 0.2 within an image and between exposures.

5. Confidence level through OCR testing of sample page images. Searchable text using OCR is a key discovery element of NDNP. For a camera master negative that is questionable with respect to any of the above criteria (resolution, reduction

ratio, densities, etc.), sample digital images may need to be tested for usable OCR confidence levels to determine suitability for selection.

Note: The current guidelines for microfilming newspaper for the USNP are available at http://www.loc.gov/preserv/usnpspecs.html.

Scanning and Master Image Format

Scanning specifications should follow these guidelines:

- scan from a clean second-generation duplicate silver negative microfilm (to be deposited with the Library of Congress at the end of the award period);
- capture specifications are 8-bit grayscale at the maximum resolution possible, between 300 and 400 dpi, relative to the physical dimensions of the original newspaper, rather than the microfilm. For the scanner operator to achieve this, the microfilm reduction ratio must be known or derived by other means;
- a standards-based target film strip should be scanned at the start of each session, to monitor scanning equipment performance. Target test images should be delivered along with the page images;
- provide the master page images, delivered to LC, as uncompressed images in TIFF 6.0 format.

Newspapers microfilmed two sheets per frame should be split into two separate image files (and assigned appropriate metadata). To improve appearance and OCR accuracy, images with more than 3 degrees of skew should be deskewed. Page image files should be cropped to the page edge (not to the text block boundaries), retaining the actual edge and up to ¼ inch beyond. All operations that change the image dimensions, spatial resolution, or orientation (e.g., cropping, deskewing) must be made to the TIFF before OCR, since the OCR output is expected to include bounding-box coordinates to relate words and characters to their position on the page in the search interface. The grayscale master TIFF files delivered to LC must have the same characteristics with respect to cropping and deskewing as the images used for OCR, but no other enhancements (sharpening, contrast enhancement, etc.) used in the OCR-creation process.

To maximize workflow efficiency, existing microfilming target frames may be captured as images and delivered with other digital assets, and described in the reel metadata object (see Appendix C – XML Metadata Templates). Such images will be treated as digital assets for archiving but not normally displayed in the NDNP access interface, as they represent an artifact of the microfilming process rather than intellectual content of the collection.

In addition, a standards-based scanning target film strip should be scanned at the start of each session, to monitor scanning equipment performance. Target test images should be delivered along with the page images, and described in the reel metadata object appropriately (see Appendix C – XML Metadata Templates). Specific test targets and quality analysis tools will be discussed with awardees at the post-award awardees' meeting.

LC and NEH intend to utilize the draft standard NISO Z39.87 Data Dictionary – Technical Metadata for Digital Still Images for master images in the NDNP. To support LC's responsible custodianship of these images, the headers for all image deliverables (TIFF, JPEG2000, and PDF) should incorporate tagged metadata relating to the creation of the images, per Appendix C.

Appendix B lists appropriate data identified by the Library of Congress. Additional tags necessary for rendering of the image (e.g., tile specifications, if used) should be present, as well. Image headers may contain additional compatible tags.

Summary of Scanning Guidelines

1. Digital reproductions should be made from a preservation copy of microfilm, a clean second-generation duplicate silver negative.

2. Technical scanning recommendation: maximum resolution possible between 300-400 dpi, relative to physical dimensions of the original material, 8-bit grayscale, TIFF 6.0 uncompressed.

3. Two-up film should be split so that there is one page image per file.

4. De-skew images with a skew of greater than 3 degrees. (Greater skew leads to less accurate OCR.)

5. Crop to edge of page.

6. Capture microfilm target frames. These image files to be identified in reel metadata; will not be used for display.

7. Capture additional scanning resolution targets (5 images per reel--target will be specified by LC) at the start of each session, to monitor scan quality. These scan target images should be delivered with microfilm target images and page images and identified in reel metadata.

Note: the grayscale images sent to LC must have exactly the same dimensions, spatial resolution, skew, and cropping as the images used for OCR.

OCR and Associated Information

Machine-readable text allows users to search a newspaper or a collection of newspapers for names of people and places, and for phrases, and provides the potential to use more powerful data-mining or natural language analysis techniques to locate relevant articles. The provision of machine-readable text correlated with page images is a tremendous aid to users seeking to navigate the complicated layouts and large, text-intensive pages of newspapers. It permits the examination of the relationships between various articles,

visually and textually. The NDNP access interface is based on a fully automated approach to text conversion without subpage-level segmentation or subpage-level metadata.

OCR creates machine-readable text from scanned page images and permits full-text searching of the contents of newspaper pages. Bounding-box data relates words to their position on the image. Coordinates describe the position and outer dimensions of a box enclosing a character or word, and/or space(s), in the original image. The NDNP application searches uncorrected OCR text at the page-level, using bounding-box coordinates for words to correlate text elements to position on the page, so that search words can be highlighted in the interface.

OCR text must be encoded using the ALTO (Analyzed Layout and Text Object) schema, Version 1-1-041 or greater, with the additional clarifications stated in Appendix B – File Format Profiles.

Important: The page images delivered must correspond in dimensions, orientation, and skew to those used for the OCR. Any text correction must retain the integrity of the ALTO positional coordinates.

Summary of OCR Guidelines

Highlighted elements for OCR files (see Appendix B for full specification):
1. One OCR text file per page image. (Discrete files should be produced for each page, rather than for a multi-page issue or entire title).
2. Each OCR text file name corresponds to the page image it represents.
3. Text in UTF-8 character set.
4. No graphic elements saved with the OCR text.
5. OCR text ordered column-by-column (that is, in a natural reading order).
6. OCR text file with bounding-box coordinate data at the word level.
7. OCR will conform to the ALTO XML schema.

If possible, additional elements for OCR files:
1. Confidence level data at the page, line, character, and/or word level.
2. Point size and font data at the character or word level.

Note: Zones for articles will not be used in the interface. If the OCR process selected by an awardee does generate coordinates for zones, the segmentation data must be removed from the METS/ALTO object prior to delivery to LC.

Other Derivative Files

In addition to the OCR text using the ALTO schema the awardee institution will provide a searchable PDF (Portable Document Format) Image with Hidden Text for each page image and a JPEG2000 compressed image file (.JP2).

PDFs will provide an image of the original page that can be conveniently printed and downloaded, supporting within-page searching for words, external to the NDNP search system. LC will use the separate OCR output file as the basis for search in its access interface. The PDF Image with Hidden Text can be created at the time of processing by the OCR application.

Highlighted elements for PDF files (see Appendix B for full specification)
1. PDF Image with Hidden Text for each page image.
2. Each searchable PDF file name corresponds to the page image it represents.
3. The PDF files should incorporate appropriate XMP metadata per Appendix B – File Format Profiles.
4. The page image will be grayscale, downsampled to 150dpi and encoding using a medium JPEG quality setting.
5. The PDF will not contain any bookmarks, links, named destinations, comments, forms, Javascript actions, external cross references, alternate images, embedded thumbnails, annotations, or private data.

The JPEG2000, Part 1, (or ISO-15444) compressed image files delivered will provide a flexible production master image that can be used to efficiently provide appropriate data to end users. For background information on the origin of the JPEG2000 profile used by NDNP, see the following report, NDNP Historical Newspaper JPEG2000 Profile (http://www.loc.gov/ndnp/pdf/NDNP_JP2HistNewsProfile.pdf).

For the NDNP access interface, LC has developed and employs a zooming capability based on JPEG2000 wavelet compression. This technology not only compresses the newspaper image effectively but also permits the presentation of image segments dynamically, at the user's request.

Highlighted elements of JPEG2000 format:
1. JPEG2000 image for each page image.
2. Each JPEG2000 will incorporate appropriate XMP metadata per Appendix B – File Format Profiles.
3. The JPEG2000 will be 6 decomposition levels, and 25 quality levels.
4. JPEG2000 compression will be 8:1.

Metadata

One aim of the LC/NEH partnership in establishing the National Digital Newspaper Program is to integrate historical newspaper collections digitized by many institutions into a single searchable resource, allowing users to search across multiple titles with a single query. To achieve this while allowing institutions the flexibility to incorporate materials into their own catalog systems and online services, NDNP awardees must deliver to LC copies of descriptive records from CONSER and metadata for various levels of granularity within the digital reproductions.

Each newspaper digitized through NDNP must be supported by coherent metadata, to provide intellectual access and support navigation of the structure of the publication, by date, section, etc. The tables in Appendix A list the elements appropriate at the newspaper title level, the issue/edition level, and the page level. [The tables indicate whether elements are mandatory and whether they are repeatable.] The access interface will permit direct identification and citation at each level through persistent identifiers. The identification of newspapers titles will be based on LCCNs, since not all historical newspapers have ISSNs. These metadata specifications will be discussed at the post-awardees' meeting.

All newspaper titles in NDNP must be described in CONSER records (itself related to holdings information in the USNP UnionList, maintained by OCLC), with a full bibliographic record at the title-level for the original materials. If pre-existing, the CONSER records must be reviewed and updated as necessary by the awardee institution and copies delivered with the project data. The records should be in MARC 21 Communications format.

LCCN is used as the basis for a title identifier instead of ISSN because not all newspapers have ISSNs. All LCCNs provided in metadata must be normalized to MARC21 standard.

Provide issue/edition metadata for all **known** issue/edition occurrences (i.e. if microfilm reel includes information (target or Guide to Contents) indicating an issue/edition was known to be published but is not available as a digital asset at this time, create a record for that issue/edition and use the Issue Present Indicator to indicate the issue/edition the record described is not available.)

Provide page metadata for all **known** page occurrences (i.e. if microfilm reel includes information (target or Guide to Contents) indicating a page was known to be published but is not available as a digital asset at this time, create a record for that page and use the Page Present Indicator to indicate the page the record describes is not available.) Note, however, that a page record should not be created for a page if the issue which the page is part of has been identified as missing.

For issue, the combination of LCCN, Issue Date, and Edition Order can be used as a unique identifier. For page, the combination of LCCN, Issue Date, Edition Order, and Record Sequence Number will be unique.

In addition to Issue and page metadata, also produce reel metadata objects that describe individual scanned reels and filmed targets. This information will be utilized to study the technical characteristics of microfilm used for digitization and their impact on OCR quality, and usability, using a quantifiable approach.

Awardees will deliver all digital assets in METS object structure (Metadata Encoded Transmission Schema), according to an XML Batch template structure. (See Appendix C – XML Metadata Templates.)

Technical Validation of Digital Objects

All NDNP Award digital objects must be *validated* prior to delivery to LC. NDNP utilizes a program-specific software application - distributed to all awardees and updated as needed, - to ensure technical conformance with the digital object profiles and specifications. The software is distributed as the Digital Viewer and Validator (DVV), which allows users to view and validate a batch through a Windows graphic user interface, or to validate from a DOS or Linux command line processor.

NDNP has developed the validation process by using and extending the JHOVE (JSTOR/Harvard Object Validation Environment – see <http://hul.harvard.edu/jhove>) toolkit. JHOVE enables the identification, validation, and characterization of files; each file format, e.g., TIFF, is supported by a separate module. The NDNP Validation Library, "wraps" JHOVE and extends JHOVE's existing TIFF, PDF, and JPEG2000 modules with the NDNP-specific validation rules. In addition, the Validation Library uses a combination of existing XML schemas and Schematron schemas, implementing validation in a custom JHOVE module, and uses JHOVE's format characterization abilities to populate the PREMIS and MIX sections of the METS objects.

For more on the technical approach of digital object validation, see Justin Littman, "A Technical Approach and Distributed Model for Validation of Digital Objects." *D-Lib Magazine*, May 2006. http://www.dlib.org/dlib/may06/littman/05littman.html#18.

Summary of All Digital Asset Deliverables
1. Validated Master digital page image format = TIFF 6.0 uncompressed,
2. Validated OCR text file with bounding-box coordinates = 1 text file per page,
3. Validated PDF Image with Hidden Text = 1 PDF per page, and
4. Validated derivative digital page image format = JPEG2000 (.JP2) using specified compression options,
5. Validated metadata using METS in accordance with guidelines in Appendices A and E (all records should be combined into one dataset).

Note: The four digital files associated directly with a newspaper page (.TIF, .JP2, .PDF, and OCR) are expected to use the same file identifiers with distinct file extensions.

Valid file format examples are available for download at http://www.loc.gov/ndnp/.

Delivery of Digital Assets

Assets should be delivered to LC in the following directory structure. A given delivery device should encompass a single batch. This batch should be stored either at the root of the device or in a batch-specific directory structure. From that directory, successive sub-

directories are based on LCCN, reel number, and issue date. At the root or at the top-level directory an XML Batch file should be created, per the template in Appendix C. The batch file is stored either at the root or within the batch-specific directory. Delivery of digital assets to LC will primarily be via tracked shipment of durable external hard drives (Firewire-enabled). The possibility of delivery via Internet2-enabled server-to-server file transfer will be discussed at the post-award conference (resource planning should be based on use of durable external hard drives). Awardees should plan for adequate temporary storage (approx. 64 Mb per page – including TIFF, JP2, PDF, OCR, XML) during the transfer and verification process at LC. Awardees should plan to deliver data to LC monthly, with an expected response time of 1 month for LC data acceptance and ingestion.

Appendices

(NOTE: Latest versions of these specifications (in use by 2007-09 NDNP Awardees) are available on the Profiles and Specifications page of the LC NDNP Web Site at http://www.loc.gov/ndnp/techspecs.html)

Appendix A: Digital Asset Metadata Elements - Dictionary

Data Description	Data Type	Example	Notes	Repeat-able R= repeatable NR = non-repeatable	Manda-tory M=mandatory MA=mandatory, if available O=optional	Xpath (see XML templates)
GENERAL INFORMATION						
Awardee	string	New York Public Library	If NEH awardee (or vendor for NEH awardee) created the METS record, should be NEH awardee name.	NR	M	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:newspaper:issue"]/metsHdr/mets:agent/mets:name
Source Repository	string	Multiple examples: Library of Congress; Washington, DC or New York Public Library; New York, NY	Owner of source that was digitized; city and state postal abbreviation	NR	M	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:newspaper:issue"]/mets:dmdSec[@ID="pageModsBib1"]/mets:mdWrap/mets:xmlData/mods:mods/mods:relatedItem[@type="original"]/mods:location/mods:physicalLocation/@displayLabel ~or~ mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/mets:dmdSec[@ID="targetModsBib1"]/mets:mdWrap/mets:xmlData/mods:mods/mods:relatedItem/mods:location/mods:phys

						icalLocation/@displayLabel
Source Repository Code	enumeration	DLC	MARC organization code of owner of source.	NR	MA	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:newspaper:issue"]/mets:dmdSec[@ID="pageModsBib1"]/mets:mdWrap/mets:xmlData/mods:mods/mods:relatedItem[@type="original"]/mods:location/mods:physicalLocation ~or~ mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/mets:dmdSec[@ID="targetModsBib1"]/mets:mdWrap/mets:xmlData/mods:mods/mods:relatedItem/mods:location/mods:physicalLocation
Digital Responsible Institution	string	Multiple examples: Library of Congress; Washington, DC or Library of Virginia; Richmond, VA	Awardee; city and state postal abbreviation	NR	M	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:newspaper:issue"]/mets:dmdSec[@ID="pageModsBib1"]/mets:mdWrap/mets:xmlData/mods:mods/mods:note[@type="agencyResponsibleForReproduction"]/@displayLabel ~or~ mets:mets[@TYPE="urn:library-of-

						congress:ndnp:mets:microfilmReel"]/mets:dmdSec[@ID="techTargetModsBib"]/mets:mdWrap/mets:xmlData/mods:mods/mods:note/@displayLabel ~or~ mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/mets:dmdSec[@ID="targetModsBib1"]/mets:mdWrap/mets:xmlData/mods:mods/mods:note/@displayLabel
Digital Responsible Institution Code	enumeration	DLC	MARC organization code of Awardee.	NR	MA	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:newspaper:issue"]/mets:dmdSec[@ID="pageModsBib1"]/mets:mdWrap/mets:xmlData/mods:mods/mods:note[@type="agencyResponsibleForReproduction"] ~or~ mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/mets:dmdSec[@ID="techTargetModsBib"]/mets:mdWrap/mets:xmlData/mods:mods/mods:note ~or~ mets:mets[@TYPE="urn:library

						-of-congress:ndnp:mets:microfilmReel"]/mets:dmdSec[@ID="targetModsBib1"]/mets:mdWrap/mets:xmlData/mods:mods/ mods:note
ISSUE INFORMATION						
LCCN	string	sn83031150	Use canonical form of LCCN for associated title bibliographic description. See http://www.loc.gov/marc/lccn-namespace.html	NR	M	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:newspaper:issue"]/mets:dmdSec[@ID="issueModsBib"]/mets:mdWrap/mets:xmlData/mods:mods/mods:relatedItem/mods:identifier/@type
Volume Number	string	27	Following SICI standard: (1.) All numeric information shall be converted to arabic numbers. (2) Alphabetic data used as enumeration designations shall be transcribed as they appear on the piece, and converted to uppercase.	NR	O	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:newspaper:issue"]/mets:dmdSec[@ID="issueModsBib"]/mets:mdWrap/mets:xmlData/mods:mods/mods:relatedItem/mods:part/mods:detail[@type="volume"]/mods:number
Edition Order	positive integer	1	Default is 1. If more than one edition on this date, number in chronological order.	NR	M	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:newspaper:issue"]/mets:dmdSec[@ID="issueModsBib"]/mets:mdWrap/mets:xmlData/mods:mods/mods:relatedItem/mods:part/mods:detail[@t

						ype="edition"]/mods:number
Edition Label	string	Multiple examples: Late City Final Two Stars (Final Edition) Extra	If present, as printed. If symbol is used to indicate edition label (e.g., two stars), describe the visual symbols and the meaning of those symbols in parentheses (i.e. if 2 stars are used to represent Final Edition, then Field value should be "Two Stars (Final Edition)".)	NR	MA	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:newspaper:issue"]/mets:dmdSec[@ID="issueModsBib"]/mets:mdWrap/mets:xmlData/mods:mods/mods:relatedItem/mods:part/mods:detail[@type="edition"]/mods:caption
Issue Number	string	3	Following SICI standard: (1.) All numeric information shall be converted to arabic numbers. (2) Alphabetic data used as enumeration designations shall be transcribed as they appear on the piece, and converted to uppercase.	NR	O	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:newspaper:issue"]/mets:dmdSec[@ID="issueModsBib"]/mets:xmlData/mods:mods/mods:relatedItem/mods:part/mods:detail[@type="issue"]/mods:number
Issue Date	date	1908-03-21	Actual date issued, corrected if necessary. Use this particular ISO 8601 style: YYYY-MM-DD.	NR	M	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:newspaper:issue"]/mets:dmdSec[@ID="issueModsBib"]/mets:mdWrap/mets:xmlData/mods:mods/mods:originInfo/mods:dateIssued
Issue Present Indicator	string	Not digitized, published	Valid values are: Present; Not digitized, published;	NR	M	mets:mets[@TYPE="urn:library-of-

			Not digitized, not published; Not digitized, publishing unknown. Note: Present (meaning Published and digitized).			congress:ndnp:mets:newspaper:issue"]/mets:dmdSec[@ID="issueModsBib"]/mets:mdWrap/mets:xmlData/mods:mods/mods:note
Issue Present Comment	string	No issue published due to weather.	To record any additional known information indicated in film on missing issues.	NR	O	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:newspaper:issue"]/mets:dmdSec[@ID="issueModsBib"]/mets:mdWrap/mets:xmlData/mods:mods/mods:note/@displayLabel
Issue Date As Labeled	date	1908-03-21	If date printed was in error (not the date issued), this field reflects the incorrect date as printed. Include this field only if date printed was in error. Use this particular ISO 8601 style: YYYY-MM-DD.	NR	MA	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:newspaper:issue"]/mets:dmdSec[@ID="issueModsBib"]/mets:mdWrap/mets:xmlData/mods:mods/mods:originInfo/mods:dateIssued[@qualifier="questionable"]
PAGE INFORMATION						
Section Label	string	B	If present, as printed. Could be blank, "C," "IV," "3," "Business," etc.	NR	MA	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:newspaper:issue"]/mets:dmdSec[@ID="sectionModsBib1"]/mets:mdWrap/mets:xmlData/mods:mods/mods:part/mods:detail[@type="section

						label"]/mods:number
Page Sequence Number	positive integer	13	This orders the records for page records within an issue, regardless of printed page number. See Page Number field below. This field is particularly useful for multi-section titles. "	NR	M	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:newspaper:issue"]/mets:dmdSec[@ID="pageModsBib1"]/mets:mdWrap/mets:xmlData/mods:mods/mods:part/mods:extent[@unit="pages"]/mods:start
Page Number	string	B3	Exactly as printed.	NR	MA	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:newspaper:issue"]/mets:dmdSec[@ID="pageModsBib1"]/mets:mdWrap/mets:xmlData/mods:mods/mods:part/mods:detail[@type="page number"]/mods:number
Page Physical Description	string	microfilm	Valid values: microfilm, microfiche, print.	NR	M	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:newspaper:issue"]/mets:dmdSec[@ID="pageModsBib1"]/mets:mdWrap/mets:xmlData/mods:mods/mods:relatedItem[@type="original"]/mods:physicalDescription/mods:form
Page Present Indicator	string	Present (Published and digitized)	Valid values are: Present; Not digitized, published; Not digitized, not published; Not digitized, publishing unknown.	NR	M	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:newspaper:issue"]/mets:dmdSec[@ID="pageModsBib1"]/mets:mdWrap/mets:xmlData/mods:mods/mods:not

			Note: Present (meaning Published and digitized).			e[@type="noteAboutReproduction"]
Page Present Comment	string	Best copy available	To record any additional known information indicated in film.	NR	O	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:newspaper:issue"]/mets:dmdSec[@ID="pageModsBib1"]/mets:mdWrap/mets:xmlData/mods:mods/mods:note[@type="noteAboutReproduction"]/@displayLabel
REEL INFORMATION						
Reel Number	string	375892205698	Reel number to correspond with LC barcode supplied for all duplicate microfilm reels deposited with LC. --Not mandatory if page is missing (even if page is represented by a target).	NR	MA	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:newspaper:issue"]/mets:dmdSec[@ID="pageModsBib1"]/mets:mdWrap/mets:xmlData/mods:mods/mods:relatedItem[@type="original"]/mods:identifer[@type="reel number"] ~or~ mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/mets:dmdSec[@ID="techTargetModsBib"]/mets:mdWrap/mets:xmlData/mods:mods/mods:relatedItem/mods:identifier[@type="reel number"]
Reel Sequence	positive	1	This orders the records	NR	MA	mets:mets[@TYPE="urn:library

Number	integer		**within a reel**. Indicate position of the image on the microfilm reel. In the case of two-up microfilming, each frame contains 2 images, and each image has its own reel sequence number, in left-right reading order. In the case of orientation-shift within the reel, continue to order the images in left-right reading order.			-of-congress:ndnp:mets:newspaper:issue"]/mets:dmdSec[@ID="pageModsBib1"]/mets:mdWrap/mets:xmlData/mods:mods/mods:relatedItem[@type="original"]/mods:identifer[@type="reel sequence number"] ~or~ mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/mets:dmdSec[@ID="techTargetModsBib"]/mets:mdWrap/mets:xmlData/mods:mods/mods:relatedItem/mods:identifier[@type="reel sequence number"]
Object Format	string	Microfilm	Valid value (currently): Microfilm	NR	M	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/mets:dmdSec[@ID="techTargetModsBib"]/mets:mdWrap/mets:xmlData/mods:mods/mods:relatedItem/mods:physicalDescription
Tech Target Label	string	ISO 16067 Calibrated Target	Describe target scanned (calibration targets scanned at NDNP request or targets already extant in microfilm)	NR		mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/mets:dmdSec[@ID="techTargetModsBib"]/mets:mdWrap/mets:xmlData/mods:mods/mods:tit

						leInfo/mods:title
Titles (on Reel)	string	National Forum Washington Evening Times	MARC 130, if present. Otherwise use MARC 245 $a. List multiple titles in repeating fields.	R	M	mets:mets[@TYPE="urn:library -of- congress:ndnp:mets:microfilmRe el"]/mets:amdSec/techMD[@ID ="reelTechMD"]/mets:mdWrap/ mets:xmlData/ndnp:reelTechMD /ndnp:titles
Start Date	date	1881-11-22	Use this particular ISO 8601 style: YYYY-MM-DD.	NR	M	mets:mets[@TYPE="urn:library -of- congress:ndnp:mets:microfilmRe el"]/mets:amdSec/techMD[@ID ="reelTechMD"]/mets:mdWrap/ mets:xmlData/ndnp:reelTechMD /ndnp:startDate
End Date	date	1881-11-16	Use this particular ISO 8601 style: YYYY-MM-DD.	NR	M	mets:mets[@TYPE="urn:library -of- congress:ndnp:mets:microfilmRe el"]/mets:amdSec/techMD[@ID ="reelTechMD"]/mets:mdWrap/ mets:xmlData/ndnp:reelTechMD /ndnp:endDate
Position	string	2a	1a, 2a, 1b, 2b	NR	M	mets:mets[@TYPE="urn:library -of- congress:ndnp:mets:microfilmRe el"]/mets:amdSec/techMD[@ID ="reelTechMD"]/mets:mdWrap/ mets:xmlData/ndnp:reelTechMD /ndnp:position
Reduction Ratio	string	20x	If stated, transcribe. If not stated, estimate.	NR	M	mets:mets[@TYPE="urn:library -of-

						congress:ndnp:mets:microfilmRe el"]/mets:amdSec/techMD[@ID ="reelTechMD"]/mets:mdWrap/ mets:xmlData/ndnp:reelTechMD /ndnp:reductionRatio
Capture Resolution Original	string	300	Resolution relative to original material, measured in pixels/inch (or mm.).	NR	M	mets:mets[@TYPE="urn:library -of- congress:ndnp:mets:microfilmRe el"]/mets:amdSec/techMD[@ID ="reelTechMD"]/mets:mdWrap/ mets:xmlData/ndnp:reelTechMD /ndnp:captureResolutionOriginal
Capture Resolution Film	string	6000	Resolution relative to microfilm, measured in pixels/inch (or mm.). Capture Resolution Film = Reduction Ratio x Capture Resolution_Original.	NR	M	mets:mets[@TYPE="urn:library -of- congress:ndnp:mets:microfilmRe el"]/mets:amdSec/techMD[@ID ="reelTechMD"]/mets:mdWrap/ mets:xmlData/ndnp:reelTechMD /ndnp:captureResolutionFilm
Guide To Contents Present Flag	boolean	true	Valid values are "true" for present; "false" for missing.	NR	M	mets:mets[@TYPE="urn:library -of- congress:ndnp:mets:microfilmRe el"]/mets:amdSec/techMD[@ID ="reelTechMD"]/mets:mdWrap/ mets:xmlData/ndnp:reelTechMD /ndnp:guideToContentsPresentFl ag
Guide To Contents String	string	Title established April 30…	If present, transcribe text from Guide to Contents, as it appears on film	NR	O	mets:mets[@TYPE="urn:library -of- congress:ndnp:mets:microfilmRe el"]/mets:amdSec/techMD[@ID

						=”reelTechMD”]/mets:mdWrap/ mets:xmlData/ndnp:reelTechMD /ndnp:guideToContentsString
Date Microfilm Created	string	1986	Date of microfilm creation	NR	M	mets:mets[@TYPE=“urn:library -of- congress:ndnp:mets:microfilmRe el”]/mets:amdSec/techMD[@ID =”reelTechMD”]/mets:mdWrap/ mets:xmlData/ndnp:reelTechMD /ndnp:dateMicrofilmCreated
Loose Leaves Flag	boolean	false	For original material; Valid values: true, false	NR	M	mets:mets[@TYPE=“urn:library -of- congress:ndnp:mets:microfilmRe el”]/mets:amdSec/techMD[@ID =”reelTechMD”]/mets:mdWrap/ mets:xmlData/ndnp:reelTechMD /ndnp:looseLeavesFlag
Bound Volume Flag	boolean	true	For original material; Valid values: true, false	NR	M	mets:mets[@TYPE=“urn:library -of- congress:ndnp:mets:microfilmRe el”]/mets:amdSec/techMD[@ID =”reelTechMD”]/mets:mdWrap/ mets:xmlData/ndnp:reelTechMD /ndnp:boundVolumeFlag
Comments	string			NR	O	mets:mets[@TYPE=“urn:library -of- congress:ndnp:mets:microfilmRe el”]/mets:amdSec/techMD[@ID =”reelTechMD”]/mets:mdWrap/ mets:xmlData/ndnp:reelTechMD /ndnp:comments

Dimensions	string	17x23 in.	From original materials, if possible, or from N.W. Ayer & Son. N.W. Ayer & Son's American Newspaper Annual. Philadelphia : N.W. Ayer and Son, 1880-1909.	NR	M	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/mets:amdSec/techMD[@ID="reelTechMD"]/mets:mdWrap/ mets:xmlData/ndnp:reelTechMD /ndnp:dimensions
Pages Per Issue	positive integer	4	Estimate from microfilm, or from N.W. Ayer & Son. N.W. Ayer & Son's American Newspaper Annual. Philadelphia : N.W. Ayer and Son, 1880-1909. See http://www.loc.gov/rr/news/news_research_tools/ayersdirectory.html	NR	M	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/mets:amdSec/techMD[@ID="reelTechMD"]/mets:mdWrap/ mets:xmlData/ndnp:reelTechMD /ndnp:pagesPerIssue
Number Of Resolution Targets	positive integer	5		NR	M	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/mets:amdSec/techMD[@ID="reelTechMD"]/mets:mdWrap/ mets:xmlData/ndnp:reelTechMD /ndnp:numberOfResolutionTargets
Resolution Of Master	floating point number	7.1	Resolution of original camera master microfilm. Record actual number from line-pair resolution test pattern on film, if available.	NR	MA	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/mets:amdSec/techMD[@ID="reelTechMD"]/mets:mdWrap/ mets:xmlData/ndnp:reelTechMD

						/ndnp:resolutionOfMaster
Resolution Comment Master	string			NR	O	mets:mets[@TYPE="urn:library -of- congress:ndnp:mets:microfilmRe el"]/mets:amdSec/techMD[@ID ="reelTechMD"]/mets:mdWrap/ mets:xmlData/ndnp:reelTechMD /ndnp:resolutionCommentMaster
Density Reading Master	floating point number	0.91	Ten readings per reel.	R	M	mets:mets[@TYPE="urn:library -of- congress:ndnp:mets:microfilmRe el"]/mets:amdSec/techMD[@ID ="reelTechMD"]/mets:mdWrap/ mets:xmlData/ndnp:reelTechMD /ndnp:densityReadingMaster
Average Density Master	floating point number	0.95		NR	M	mets:mets[@TYPE="urn:library -of- congress:ndnp:mets:microfilmRe el"]/mets:amdSec/techMD[@ID ="reelTechMD"]/mets:mdWrap/ mets:xmlData/ndnp:reelTechMD /ndnp:averageDensityMaster
Dmin Master	floating point number	0.2		NR	M	mets:mets[@TYPE="urn:library -of- congress:ndnp:mets:microfilmRe el"]/mets:amdSec/techMD[@ID ="reelTechMD"]/mets:mdWrap/ mets:xmlData/ndnp:reelTechMD /ndnp:dminMaster
Resolution Of Duplicate	floating point	6.3	Resolution of the duplicate negative used	NR	MA	mets:mets[@TYPE="urn:library -of-

Negative	number		for digitization. Record actual number from line-pair resolution test pattern on film, if available.			congress:ndnp:mets:microfilmReel"]/mets:amdSec/techMD[@ID ="reelTechMD"]/mets:mdWrap/ mets:xmlData/ndnp:reelTechMD /ndnp:resolutionOfDuplicateNeg ative
Resolution Comment Duplicate Negative	string	Weak 7.1		NR	O	mets:mets[@TYPE="urn:library -of- congress:ndnp:mets:microfilmReel"]/mets:amdSec/techMD[@ID ="reelTechMD"]/mets:mdWrap/ mets:xmlData/ndnp:reelTechMD /ndnp:resolutionCommentDuplic ateNegative
Density Reading Duplicate Negative	floating point number	1.14	Ten readings per reel.	R	M	mets:mets[@TYPE="urn:library -of- congress:ndnp:mets:microfilmReel"]/mets:amdSec/techMD[@ID ="reelTechMD"]/mets:mdWrap/ mets:xmlData/ndnp:reelTechMD /ndnp:densityReadingDuplicate Negative
Average Density Duplicate Negative	floating point number	1.1		NR	M	mets:mets[@TYPE="urn:library -of- congress:ndnp:mets:microfilmReel"]/mets:amdSec/techMD[@ID ="reelTechMD"]/mets:mdWrap/ mets:xmlData/ndnp:reelTechMD /ndnp:averageDensityDuplicate Negative
Dmin Duplicate	floating	0.15		NR	M	mets:mets[@TYPE="urn:library

Negative	point number					-of-congress:ndnp:mets:microfilmReel"]/mets:amdSec/techMD[@ID ="reelTechMD"]/mets:mdWrap/ mets:xmlData/ndnp:reelTechMD /ndnp:dminDuplicateNegative

Appendix B: File Format Profiles and Specifications

NOTE: Latest versions of Profiles and Specifications (in use by 2005-07 NDNP Awardees) are available on the LC NDNP Web Site at
http://www.loc.gov/ndnp/techspecs.html

NDNP TIFF Profile
Version 1.7

CHANGES in 1.7:
1. Changes to resolution limits.

CHANGES in 1.6:
1. Corrections to tags 282, 283, and 42016.

1. The TIFF will conform with the TIFF 6.0 specification, except it is not required that values be word offset[1]. A tag must be provided if it is required by the TIFF 6.0 specification and its value is not the default.
2. The TIFF will be 8-bit grayscale.
3. The TIFF will not be compressed.
4. Image processing will not be applied to the TIFF, except for deskewing. The TIFF will be as close to the original produced by the scanner as possible. Deskewing will be applied if the skew is greater than 3 degrees.
5. TIFF resolution must be maximum possible between 300 and 400 dpi, relative to physical dimensions of the original material. If this is not possible, contact NDNP.
6. The image should be cropped to the page edge (not to the text block boundaries).

In addition to the tags required by the TIFF 6.0 specification, the following tags are required. The tag must be provided if the value is not the default as given in the TIFF 6.0 specification:

TIFF Tag #:	269
TIFF Tag Name:	DocumentName
TIFF Type:	ASCII
Notes:	Use Microfilm reel # (barcode). If not converted from microfilm, use normalized LCCN.

TIFF Tag #:	42016
TIFF Tag Name:	ImageUniqueID
Z39.87 #:	6.2.1
Z39.87 Name:	UniqueImageID

---

[1]This is a common problem with many TIFF producing applications and is readily handled by most TIFF rendering applications.

TIFF Type: ASCII

Notes: Use reel sequence number. Must be unique within reel. If not converted from microfilm, use "[Issue date in CCYY-MM-DD format]_[edition order]_[page sequence number]". For example, "1909-03-20_1_13".

TIFF Tag #:   274
TIFF Tag Name: Orientation
Z39.87 #:   6.2.4
Z39.87 Name: Orientation

TIFF Tag #:   41728
TIFF Tag Name: FileSource
Z39.87 #:   7.1
Z39.87 Name: SourceType
TIFF Type: ASCII
Value: "microfilm" or "microfiche" or "print"
Note: Contact NDNP is none of the permitted values are appropriate.

TIFF Tag #:   315
TIFF Tag Name: Artist
Z39.87 #:   7.3
Z39.87 Name: ImageProducer
Notes: Use the following format: "institution name; scanning contractor".
The semicolon and scanning contractor are omitted if not applicable.

TIFF Tag #:   271
TIFF Tag Name: Make
Z39.87 #:   7.6.1.1
Z39.87 Name: ScannerManufacturer

TIFF Tag #:   272
TIFF Tag Name: Model
Z39.87 #:   7.6.1.2.1
Z39.87 Name: ScannerModelName
Note: Include model number (optional).
Include serial number (required).
Use the following format: "model name, model number, SN#serial number"

TIFF Tag #:   305
TIFF Tag Name: Software
Z39.87 #:   7.6.2.1
Z39.87 Name: ScanningSoftware
Note: Include version.

TIFF Tag #:   306
TIFF Tag Name: DateTime

Z39.87 #:   7.9
Z39.87 Name: DateTimeCreated
Note: See the TIFF 6.0 specification for the proper formatting of this tag.

The following are clarifications of the TIFF 6.0 specification, based on commonly encountered mistakes or additional NDNP requirements:
TIFF Tag #:   256
TIFF Tag Name: ImageWidth
Note: ImageWidth is measured in pixels.

TIFF Tag #:   257
TIFF Tag Name: ImageLength
Note: ImageLength is measured in pixels.

TIFF Tag #:   296
TIFF Tag Name: ResolutionUnit
Value: "2" (inches) or "3" (centimeters)
Note: Specifying a unit of measurement is required. A value of "2" (inches) is encouraged.

TIFF Tag #:   282
TIFF Tag Name: XResolution
Note: Xresolution is a Rational, as defined by the TIFF 6.0 specification. ImageWidth (Tag 256) is the numerator and the length of the source (measured in the units specified in ResolutionUnit (Tag 296)) is the denominator.

TIFF Tag #:   283
TIFF Tag Name: YResolution
Note: Yresolution is a Rational, as defined by the TIFF 6.0 specification. ImageLength (Tag 257) is the numerator and the width of the source (measured in the units specified in ResolutionUnit (Tag 296)) is the denominator.

TIFF Tag #:   277
TIFF Tage Name: SamplesPerPixel
Value: 1

NDNP JPEG 2000 Profile

Version 2.7

CHANGES IN 2.7:
Removed code-block style requirement.

1. The JPEG 2000 file will conform with JP2 file format as specified in ISO/IEC 15444- 1:2000 (i.e., JPEG 2000, Part 1).
2. The JPEG 2000 file will be prepared after any image processing or clean-up is performed. The JPEG 2000 file will correspond with the image that is used for OCR.
3. The JPEG 2000 file's brand will be "jp2 ", version will be "0" and compatibility will be "jp2 ". (Note the space after jp2.)
4. The JPEG 2000 file's image X origin, image Y origin, tile X origin, and tile Y origin will be 0.
5. The JPEG 2000 file will contain only one component.
6. The bit depth of that component will be 8.
7. The JPEG 2000 file's height and width will be the same as the TIFF master file.
8. The JPEG 2000 file's tile header will not contain coding style default, coding style component, quantization default, and quantization component marker segments.
9. The JPEG 2000's progression order will be RLCP (resolution, layer, component, position) or RLPC.
10. The JPEG 2000 will have 6 decomposition levels.
11. The JPEG 2000 will have 25 quality layers. The bits per pixel for each quality layer will be: 1,0.84,0.7,0.6,0.5,0.4,0.35,0.3,0.25,0.21,0.18,0.15,0.125,0.1,0.088,0.07,0.0625,0.005,0.04419,0.03716,0.03125,0.025,0.0221,0.01858,0.015625.
12. The JPEG 2000's code-block size will be 64x64.
13. The JPEG 2000's code-block style will be bypass.
14. The JPEG 2000 will use the 9-7 irreversible filter.
15. The JPEG 2000 will be compressed so that it is about one-eighth of the TIFF or 1 bit per pixel.
16. The JPEG 2000 will use 1024x1024 tiles.
17. The JPEG 2000's color specification must be either the monochrome (grayscale) enumerated color space or the Monochrome Input restricted ICC profile.
18. The JPEG 2000 file will not contain regions of interest or precincts.
19. The JPEG 2000 file will not contain intellectual property rights information.
20. It is recommended that information about the codec used to encode the JPEG 2000 file (e.g., name, version) be included. The preferred method to do this is an an XML Box containing the relevant MIX elements.
21. The JPEG 2000 file will contain an XML Box that conforms with the following:

For newspaper pages:
```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdfsyntax-ns#">
     <rdf:Description rdf:about="urn:library-
ofcongress:ndnp:mets:newspaper:page://#The normalized LCCN/#Date of
publication in CCYY-MM-DD#/#Edition order#/#Page sequence number#"
xmlns:dc="http://purl.org/dc/elements/1.1/">
          <dc:format>image/jp2</dc:format>
          <dc:title>
               <rdf:Alt>
               <rdf:li xml:lang="en">#The title of the
          newspaper#.(#Location of publication#) #Date of publication
          in CCYY-MM-DD# [p #page label#].</rdf:li>
```

```
                              </rdf:Alt>
                    </dc:title>
                    <dc:description>
                              <rdf:Alt>
                              <rdf:li xml:lang="en">Page from #The title of the
                    newspaper# (newspaper). [See LCCN: #The normalized LCCN#
                    for catalog record.]. Prepared on behalf of #responsible
                    organization#.</rdf:li>
                              </rdf:Alt>
                    </dc:description>
                    <dc:date>
                              <rdf:Seq>
                              <rdf:li xml:lang="x-default">#Date of publication in
                    CCYY-MM-DD#</rdf:li>
                              </rdf:Seq>
                    </dc:date>
                    <dc:type>
                              <rdf:Bag>
                              <rdf:li xml:lang="en">text</rdf:li>
                              <rdf:li xml:lang="en">newspaper</rdf:li>
                              </rdf:Bag>
                    </dc:type>
                    <dc:identifier>
                              <rdf:Alt>
                              <rdf:li xml:lang="en">Reel number #The reel number#.
                    Sequence number #The sequence number#.</rdf:li>
                              </rdf:Alt>
                    </dc:identifier>
              </rdf:Description>
</rdf:RDF>
```

Note: dc:identifier may be omitted if the image is not created from microfilm.

For example:
```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdfsyntax-ns#">
      <rdf:Description rdf:about='urn:library-
ofcongress:ndnp:mets:newspaper:page://sn82015056/1910-05-
28/1/1'
xmlns:dc='http://purl.org/dc/elements/1.1/'>
              <dc:format>image/jp2</dc:format>
              <dc:title>
                        <rdf:Alt>
                        <rdf:li xml:lang="en">The National Forum.(Washington,
                    D.C.) 1910-05-28 [p 1].</rdf:li>
                        </rdf:Alt>
              </dc:title>
              <dc:description>
                        <rdf:Alt>
                        <rdf:li xml:lang="en">Page from The National Forum
                    (newspaper). [See LCCN: sn 82015056 for
catalog record.]. Prepared on behalf of Library of Congress.</rdf:li>
                        </rdf:Alt>
              </dc:description>
              <dc:date>
                        <rdf:Seq>
```

```
                    <rdf:li xml:lang="x-default">1910-05-28</rdf:li>
                    </rdf:Seq>
            </dc:date>
            <dc:type>
                    <rdf:Bag>
                    <rdf:li xml:lang="en">text</rdf:li>
                    <rdf:li xml:lang="en">newspaper</rdf:li>
                    </rdf:Bag>
            </dc:type>
            <dc:identifier>
                    <rdf:Alt>
                    <rdf:li xml:lang="en">Reel number 23454234545.
            Sequence number 5.</rdf:li>
                    </rdf:Alt>
            </dc:identifier>
        </rdf:Description>
</rdf:RDF>
```

For targets:
```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdfsyntax-ns#">
        <rdf:Description rdf:about="#urn:library-
ofcongress:ndnp:mets:newspaper:target://#Reel number#/#Reel
sequence number#"
xmlns:dc="http://purl.org/dc/elements/1.1/">
            <dc:format>image/jp2</dc:format>
            <dc:description>
                    <rdf:Alt>
                    <rdf:li xml:lang="en">Target from microfilm reel
            #Reel number#. Prepared on behalf of #responsible
            organization#.</rdf:li>
                    </rdf:Alt>
            </dc:description>
            <dc:identifier>
                    <rdf:Alt>
                    <rdf:li xml:lang="en">Reel number #The reel number#.
            Sequence number #The sequence number#.</rdf:li>
                    </rdf:Alt>
            </dc:identifier>
        </rdf:Description>
</rdf:RDF>
```

To generate a JPEG 2000 with Kakadu, use:
```
kdu_compress -i YOURINPUT.pgm -o YOUROUTPUT.jp2 -rate
1,0.84,0.7,0.6,0.5,0.4,0.35,0.3,0.25,0.21,0.18,0.15,0.125,0.1,0.088,0.0
75,0.0625,0.05,0.04419,0.03716,0.03125,0.025,0.0221,0.01858,0.015625
Clevels=6 Stiles={1024,1024} Corder=RLCP -jp2_box YOURMETADATA.xml
```

Note that the metadata xml file must begin with "xml " followed by a newline.

To generate a JPEG 2000 with Aware, use:
```
j2kdriver -i YOURINPUT.tif -t JP2 -p RLCP -w I97 6 -B 1 -p RLCP --tile-
size 1024 1024 -yB 0 1 -yB 1 0.84 -yB 2 0.7 -yB 3 0.6 -yB 4 0.5 -yB 5
0.4 -yB 6 0.35 -yB 7 0.3 -yB 8 0.25 -yB 9 0.21 -yB 10 0.18 -yB 11 0.15
```

```
-yB 12 0.125 -yB 13 0.1 -yB 14 0.088 -yB 15 0.075 -yB 16 0.0625 -yB 17
0.05 -yB 18 0.04419 -yB 19 0.03716 -yB 20 0.03125 -yB 21 0.025 -yB 22
0.0221 -yB 23 0.01858 -yB 24 0.015625 --set-output-jp2-add-metadata-box
XML YOURMETADATA.xml -o YOUROUTPUT.jp2
```

NDNP PDF Profile

Version 2.1

CHANGES IN 2.1:
1. Removed specific PDF/A requirements.
2. Added general PDF/A recommendation.

1. The PDF will contain only one page, which will contain an image of the newspaper page with text "behind" the image.
2. The image will grayscale, downsampled to 150dpi and encoded using JPEG, using a medium (or 40) quality setting. (If these settings are found to be suboptimal, please contact NDNP.)
3. Only the 14 standard Type1 fonts may be used. These fonts will not be embedded.
4. The page may have a page label. The page label will be the page number as it appears in the image.
5. Text streams will be Flate encoded.
6. Text should not be placed in the same location as other text. (Note: Some OCR applications will place a small amount of text in the same location as other text. However, text should not deliberately be placed in the same location as other text.)
7. The page will not contain any bookmarks, links, named destinations, comments, forms, Javascript actions, external cross references, alternate images, embedded thumbnails, annotations, or private data.
8. The PDF will not be tagged. (Note: PDFs of newspapers tagged with Adobe Acrobat's automated tagging are generally inaccurate and not useful for Read Aloud functionality.)
9. The PDF will open to Fit Page sizing.
10. The PDF will open to single page layout.
11. The PDF will open with neither document outline nor thumbnail images available.
12. The PDF will open with the tool bar, menu bar, and user interface elements visible.
13. The PDF will not open centered in the screen.
14. The PDF will not be encrypted, digitally signed, or have any security.
15. It is recommended that the PDF be linearized (also known as "Fast Web View").
16. The PDF will be compatible with Acrobat 5.0 or later.
17. Except where conflicting with any of the other requirements of this profile, conforming to PDF/A (ISO 19005-1) is recommended.
18. The PDF's document XMP metadata conform with the following:

For newspaper pages:
```
<rdf:Description rdf:about="#The appropriate uuid#"
xmlns:dc="http://purl.org/dc/elements/1.1/">
    <dc:format>application/pdf</dc:format>
    <dc:title>
        <rdf:Alt>
```

```
        <rdf:li xml:lang="en">#The title of the
newspaper#.(#Location of publication#) #Date of publication in
CCYY-MM-DD# [p #page label#].</rdf:li>
        </rdf:Alt>
</dc:title>
<dc:description>
        <rdf:Alt>
        <rdf:li xml:lang="en">Page from #The title of the
newspaper# (newspaper). [See LCCN: #The normalized LCCN# for
catalog record.]. Prepared on behalf of #responsible
organization#.</rdf:li>
        </rdf:Alt>
</dc:description>
<dc:date>
        <rdf:Seq>
        <rdf:li xml:lang="x-default">#Date of publication in CCYY-
MM-DD#</rdf:li>
        </rdf:Seq>
</dc:date>
<dc:type>
        <rdf:Bag>
        <rdf:li xml:lang="en">text</rdf:li>
        <rdf:li xml:lang="en">newspaper</rdf:li>
        </rdf:Bag>
</dc:type>
<dc:identifier>
        <rdf:Alt>
        <rdf:li xml:lang="en">Reel number #The reel number#.
Sequence number #The sequence number#.</rdf:li>
        </rdf:Alt>
</dc:identifier>
</rdf:Description>
```

Note: dc:identifier may be omitted if the image is not created from microfilm.

For example:
```
<rdf:Description rdf:about='uuid:813e0307-eef2-4b52-8fd2-902bfade3ca3'
xmlns:dc='http://purl.org/dc/elements/1.1/'>
        <dc:format>application/pdf</dc:format>
        <dc:title>
        <rdf:Alt>
        <rdf:li xml:lang="en">The National Forum.(Washington, D.C.)
1910-05-28 [p 1].</rdf:li>
        </rdf:Alt>
</dc:title>
<dc:description>
        <rdf:Alt>
        <rdf:li xml:lang="en">Page from The National Forum
(newspaper). [See LCCN: sn 82015056 for catalog record.].
Prepared on behalf of Library of Congress.</rdf:li>
        </rdf:Alt>
</dc:description>
<dc:date>
        <rdf:Seq>
        <rdf:li xml:lang="x-default">1910-05-28</rdf:li>
        </rdf:Seq>
```

```
        </dc:date>
        <dc:type>
              <rdf:Bag>
              <rdf:li xml:lang="en">text</rdf:li>
              <rdf:li xml:lang="en">newspaper</rdf:li>
              </rdf:Bag>
        </dc:type>
        <dc:identifier>
              <rdf:Alt>
              <rdf:li xml:lang="en">Reel number 23454234545. Sequence
      number 5.</rdf:li>
              </rdf:Alt>
        </dc:identifier>
</rdf:Description>
```

For targets:
```
<rdf:Description rdf:about="#The appropriate uuid#"
xmlns:dc="http://purl.org/dc/elements/1.1/">
        <dc:format>application/pdf</dc:format>
        <dc:description>
              <rdf:Alt>
              <rdf:li xml:lang="en">Target from microfilm reel #Reel
      number#. Prepared on behalf of #responsible
      organization#.</rdf:li>
              </rdf:Alt>
        </dc:description>
        <dc:identifier>
              <rdf:Alt>
              <rdf:li xml:lang="en">Reel number #The reel number#.
      Sequence number #The sequence number#.</rdf:li>
              </rdf:Alt>
        </dc:identifier>
</rdf:Description>
```

NDNP OCR Profile

Version 1.10

       CHANGES IN 1.10:

              1. Added clarification of hyphenation.

       CHANGES IN 1.9:

              1. Added prohibition on multiple Strings in same location.

              2. Further clarified natural reading order.

       CHANGES IN 1.8:

              1. ALTO version updated to 1-2.

       CHANGES IN 1.7:

              1. Added clarification about column organization.

1. OCR text will be encoded using the ALTO (Analyzed Layout and Text Object) schema, Version 1-2, with the additional clarifications stated below.
2. The value for MeasurementUnit will be "inch1200," which is 1/1200 of an inch.
3. The use of the SourceImageInformation\fileName element is required. This should include the path if the path contains useful information (e.g., identifying the newspaper title and/or issue).
4. The use of the OCRProcessing element is encouraged.
5. If the OCRProcessing element is used, the use of the ProcessingSoftware element is required. If the software does not have a commercial name, the name of the executable may be used.
6. For all applicable elements, the use of STYLEREFS and language are encouraged.
7. For the PageElement, the use of Height, Width, PRINTED_IMG_NR, QUALITY, POSITION, and PROCESSING are encouraged.
8. For the PageElement, the entire page may be included in the PrintSpace. (Thus, use of TopMargin, LeftMargin, RightMargin, and BottomMargin are not required.)
9. The use of Illustration, GraphicalElement, and ComposedBlock are not required.
10. The use of non-rectangular blocks is not encouraged.
11. The use of SP and HYP are encouraged.
12. For a TextLine, the use of BASELINE is discouraged.
13. For a String, the use of ALTERNATIVE, WC, and CC is encouraged if available.
14. For a String, the use of HEIGHT, WIDTH, HPOS, and VPOS is required.
15. The coordinates of Strings should not overlap. In other words, for each location on a page, there should only be one String. (If alternatives are desired, ALTERNATIVE should be utilized, not multiple Strings.)
16. OCR text must be in natural reading order. Thus, OCR text should reflect columns of the original newspaper and be ordered column-by-column. In addition, the ordering of all elements should reflect the original newspaper. (That is, reading order should be indicated by the ordering of elements, for example, Strings should be in reading order.)

Additional clarifications:

1. For the ProcessingStepSettings, the settings can be specified as the command-line arguments given to the processing software.

2. For a String, the CONTENT should be a word, not a character.

3. If a hyphen splits a word at the end of a line, the OCR file should represent both fragments of the word, the hyphen, and the complete word. See the following example, where the word "experts" was split at the end of a line.

```
<String ID="P5_ST00015" HPOS="5508" VPOS="24344" WIDTH="170"
HEIGHT="61" CONTENT="ex" SUBS_TYPE="HypPart1" SUBS_CONTENT="experts"
WC="0.96" CC="111"/>
<HYP CONTENT="-"/>
</TextLine>
<TextLine ID="P5_TL00003" HPOS="3146" VPOS="24425" WIDTH="2532"
HEIGHT="108">
<String ID="P5_ST00016" HPOS="3146" VPOS="24439" WIDTH="288"
HEIGHT="94" CONTENT="perts" SUBS_TYPE="HypPart2" SUBS_CONTENT="experts"
WC="0.99" CC="00001"/>
```

4. If the hyphenated word occurs in the middle of a line, the hyphen should be left in place. See the following example where the word is "re-examination" occurred in the middle of a line.
```
<String ID="P2_ST03691" HPOS="11428" VPOS="15727" WIDTH="897"
HEIGHT="89" CONTENT="re-examination" WC="0.97" CC="01001011010110"
STYLEREFS="TXT_5"/>
```

Appendix C: XML Metadata Templates

ISSUE XML Metadata Template, v.1.9

```
<?xml version="1.0" encoding="UTF-8"?>
<!--NDNP Newspaper issue profile-->
<!--Justin Littman, OSI-->
<!--Version 1.9-->
<!--Note that an issue may have additional elements/attributes
added/removed as
    it passes from awardee/vendor and through the pre-ingest and ingest
processes.
    Elements/attributes that do not apply to all steps have been
indicated below.
    amdSec (and children):  Should only be added by Validation
Application.
    RECORDSTATUS:  Should only be added by Validation Application.
-->
<!--INSTRUCTIONS:
1.  Omit all comments.  (Comments in all caps may be left in for
clarity.)
2.  Treat all attribute values in brackets as comments that should be
replaced with the appropriate values.
3.  Omit xsi:schemaLocation
4.  Optional fields that are not going to be populated should be
omitted rather than left blank.  For example, if a page does not have a
page number, omit <mods:detail type="page
number"><mods:number></mods:number></mods:detail>.
-->

<!--CHANGES in 1.8:
1.  Added reference to digital signature in metsHdr.
2.  Removed individual file digital signatures.

    CHANGES in 1.9:
1.  Added physicalDescription to page mods to support originals other
than microfilm.
-->
<mets TYPE="urn:library-of-congress:ndnp:mets:newspaper:issue"
PROFILE="urn:library-of-congress:mets:profiles:ndnp:issue:v1.5"
LABEL="[Newspaper title], [issue date][, editionLabel (if present)"
    xmlns:mix="http://www.loc.gov/mix/"
xmlns:ndnp="http://www.loc.gov/ndnp"
xmlns:premis="http://www.loc.gov/standards/premis"
xmlns:mods="http://www.loc.gov/mods/v3"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:xlink="http://www.w3.org/1999/xlink"
xmlns:dsig="http://www.w3.org/2000/09/xmldsig#"
xmlns="http://www.loc.gov/METS/"
xsi:schemaLocation="http://www.loc.gov/METS/ ../schema/ndnpmets.v1-
4.xsd http://www.loc.gov/mods/v3 ../schema/ndnpmods-3-1.xsd
http://www.loc.gov/ndnp ../schema/ndnp.xsd
```

```
http://www.loc.gov/standards/premis ../schema/Object.xsd
http://www.loc.gov/mix/ http://www.loc.gov/standards/mix/mix.xsd
http://www.w3.org/2000/09/xmldsig# ../schema/xmldsig-core-schema.xsd">

     <!--METS HEADER-->
     <metsHdr CREATEDATE="2004-11-12T09:00:00" RECORDSTATUS="Validated"
ADMID="dsig metsPremis"><!--CREATEDATE should be populated with
creation date of the record.  RECORDSTATUS and ADMID should only be set
by Validation Application.-->
        <agent ROLE="CREATOR" TYPE="ORGANIZATION">
           <name><!--Awardee, if the awardee created the METS record.
If LC (or an LOC vendor) created the METS record, should be "Library of
Congress"--></name>
        </agent>
     </metsHdr>

     <!--DESCRIPTIVE METADATA-->
     <!--Descriptive metadata for issue-->
     <dmdSec ID="issueModsBib">
        <mdWrap MDTYPE="MODS" LABEL="Issue metadata">
           <xmlData>
              <mods:mods>
                 <mods:relatedItem type="host">
                    <mods:identifier type="lccn"><!--lccn of
newspaperTitle, normalized according to http://www.loc.gov/marc/lccn-
namespace.html (required)--></mods:identifier>
                    <mods:part><!--required-->
                       <mods:detail type="volume">
                          <mods:number><!--volumeNumber
(optional)--></mods:number>
                       </mods:detail>
                       <mods:detail type="issue">
                          <mods:number><!--issueNumber
(optional)--></mods:number>
                       </mods:detail>
                       <mods:detail type="edition"><!--required-->
                          <mods:number><!--editionOrder
(required)--></mods:number>
                          <mods:caption><!--editionLabel
(optional)--></mods:caption>
                       </mods:detail>
                    </mods:part>
                 </mods:relatedItem>
                 <mods:originInfo><!--required-->
                    <mods:dateIssued encoding="iso8601"><!--
issueDate (required)--></mods:dateIssued>
                    <mods:dateIssued encoding="iso8601"
qualifier="questionable"><!--issueDateAsLabeled (optional)--
></mods:dateIssued>
                 </mods:originInfo>
                 <mods:note type="noteAboutReproduction"
displayLabel="[issuePresentComment, otherwise
omit]">Present</mods:note><!--issuePresentIndicator must be valid
enumeration (required)-->
              </mods:mods>
           </xmlData>
        </mdWrap>
```

```
        </dmdSec>

    <!--Descriptive metadata for section-->
    <dmdSec ID="sectionModsBib1">
         <mdWrap MDTYPE="MODS" LABEL="Section metadata">
             <xmlData>
                 <mods:mods>
                     <mods:part>
                         <mods:detail type="section label">
                             <mods:number><!--sectionLabel (optional)--
></mods:number>
                         </mods:detail>
                     </mods:part>
                 </mods:mods>
             </xmlData>
         </mdWrap>
    </dmdSec>


    <!--Descriptive metadata for page-->
<dmdSec ID="pageModsBib1">
         <mdWrap MDTYPE="MODS" LABEL="Page metadata">
             <xmlData>
                 <mods:mods>
                     <mods:part>
                         <mods:extent unit="pages">
                             <mods:start><!--pageSequenceNumber
(required)--></mods:start>
                         </mods:extent>
                         <mods:detail type="page number">
                             <mods:number><!--pageNumber (optional)--
></mods:number>
                         </mods:detail>
                     </mods:part>
                     <mods:relatedItem type="original">
                         <mods:physicalDescription>
                             <mods:form type="microfilm" /><!--Valid
values are microfilm, microfiche, or print.  See following dmdSec for
an example of a print original.-->
                         </mods:physicalDescription>
                         <mods:identifier type="reel number"><!--
reelNumber (required if converted from microfilm)--></mods:identifier>
                         <mods:identifier type="reel sequence
number"><!--reelSequenceNumber (required if converted from microfilm)--
></mods:identifier>
                         <mods:location>
                             <mods:physicalLocation authority="marcorg"
displayLabel="[sourceRepository]"><!--sourceRepositoryCode (from
http://www.loc.gov/marc/organizations)--></mods:physicalLocation>
                         </mods:location>
                     </mods:relatedItem>
                     <mods:note type="agencyResponsibleForReproduction"
displayLabel="[digitalResponsibleInstitution]"><!--
digitalResponsibleInstitutionCode (from
http://www.loc.gov/marc/organizations)--></mods:note>
                     <mods:note type="noteAboutReproduction"
displayLabel="[pagePresentComment, otherwise
```

```xml
omit]">Present</mods:note><!--pagePresentIndicator must be valid
enumeration-->
                </mods:mods>
            </xmlData>
        </mdWrap>
    </dmdSec>
    <!--Repeat dmdSec for each page-->

    <amdSec>
        <!--TECHNICAL METADATA.-->
        <!--All technical metadata is added by trusted validator-->
            <techMD ID="metsPremis">
            <mdWrap MDTYPE="OTHER" OTHERMDTYPE="PREMIS" LABEL="PREMIS
object metadata">
                <xmlData>
                    <premis:object>
                        <premis:objectCharacteristics>
                            <premis:significantProperties>
                                <ndnp:exemptionSet>
                                    <ndnp:code>TEST1</ndnp:code>
                                    <ndnp:code>TEST2</ndnp:code>
                                </ndnp:exemptionSet>
                            </premis:significantProperties>
                        </premis:objectCharacteristics>
                    </premis:object>
                </xmlData>
            </mdWrap>
        </techMD>

        <!--PREMIS technical metadata for digital master for page 1.
Added by trusted validator.-->
        <techMD ID="masterPremis1">
            <mdWrap MDTYPE="OTHER" OTHERMDTYPE="PREMIS" LABEL="PREMIS
object metadata">
                <xmlData>
                    <premis:object>
                        <premis:objectCharacteristics>
                            <premis:fixity>
                                <premis:messageDigestAlgorithm>SHA-
1</premis:messageDigestAlgorithm>
                                <premis:messageDigest><!--The SHA-1
checksum (Required)--></premis:messageDigest>
                                <premis:messageDigestOriginator>Library
of Congress</premis:messageDigestOriginator>
                            </premis:fixity>
                            <premis:size>1</premis:size><!--Size of the
file in bytes. (Required)-->
                            <premis:format>
                                <premis:formatDesignation>

<premis:formatName>image/tiff</premis:formatName>
                                </premis:formatDesignation>
                            </premis:format>
                            <premis:significantProperties>
                                <ndnp:exemptionSet>
                                    <ndnp:code>TEST1</ndnp:code>
                                    <ndnp:code>TEST2</ndnp:code>
```

```
                                </ndnp:exemptionSet>
                            </premis:significantProperties>
                        </premis:objectCharacteristics>
                        <premis:creatingApplication>
                            <premis:creatingApplicationName><!--
creatingApplicationName (optional)--></premis:creatingApplicationName>
                            <premis:creatingApplicationVersion><!--
creatingApplicationVersion (optional)--
></premis:creatingApplicationVersion>
                            <premis:dateCreatedByApplication>2004-11-
22T00:00:00</premis:dateCreatedByApplication><!--
dateCreatedByApplication (required for master)-->
                        </premis:creatingApplication>
                    </premis:object>
                </xmlData>
            </mdWrap>
        </techMD>
        <!--techMD is repeated for master image for each page-->

        <!--MIX technical metadata for digital master for page 1.
Added by Trusted Validator.-->
        <techMD ID="masterMix1">
            <mdWrap MDTYPE="NISOIMG" LABEL="NISO still image metadata"
>
                <xmlData>
                    <mix:mix>
                        <mix:BasicImageParameters>
                            <mix:Format>
                                <mix:Compression>

<mix:CompressionScheme>1</mix:CompressionScheme><!--Compression scheme
(required)-->
                                </mix:Compression>
                                <mix:PhotometricInterpretation>

<mix:ColorSpace>0</mix:ColorSpace><!--Colorspace (required)-->
                                </mix:PhotometricInterpretation>
                            </mix:Format>
                        </mix:BasicImageParameters>
                        <mix:ImageCreation>

<mix:SourceType>Microfilm</mix:SourceType><!--sourceType (required for
master)-->
                            <mix:ImageProducer><!--awardee; vendor
(required for master; optional for others)--></mix:ImageProducer>
                            <mix:ScanningSystemCapture>
                                <mix:ScanningSystemHardware>
                                    <mix:ScannerManufacturer><!--
scannerManufacturer (required for master)--></mix:ScannerManufacturer>
                                    <mix:ScannerModel><!--
scannerModelName (required for master)--></mix:ScannerModel>
                                </mix:ScanningSystemHardware>
                            </mix:ScanningSystemCapture>
                        </mix:ImageCreation>
                        <mix:ImagingPerformanceAssessment>
                            <mix:SpatialMetrics>
```

```
<mix:SamplingFrequencyUnit>2</mix:SamplingFrequencyUnit><!--
SamplingFrequencyUnit (required)-->

<mix:XSamplingFrequency>400</mix:XSamplingFrequency><!--
XSamplingFrequency (required)-->

<mix:YSamplingFrequency>400</mix:YSamplingFrequency><!--
YSamplingFrequency (required)-->
                              <mix:ImageWidth>1</mix:ImageWidth><!--
ImageWidth (required)-->
                              <mix:ImageLength>1</mix:ImageLength><!-
-ImageLength (required)-->
                         </mix:SpatialMetrics>
                         <mix:Energetics>

<mix:BitsPerSample>8</mix:BitsPerSample><!--Bits per sample (required)-
->
                         </mix:Energetics>
                    </mix:ImagingPerformanceAssessment>
                </mix:mix>
            </xmlData>
        </mdWrap>
    </techMD>
    <!--techMD is repeated for master image for each page, unless a
master image has identical values as a previous master image, in which
case repeating is not necessary-->

    <!--PREMIS technical metadata for primary service image for
page 1  Added by Trusted Validator.-->
    <techMD ID="primaryServicePremis1">
        <mdWrap MDTYPE="OTHER" OTHERMDTYPE="PREMIS" LABEL="PREMIS
object metadata">
            <xmlData>
                <premis:object>
                    <premis:objectCharacteristics>
                        <premis:fixity>
                            <premis:messageDigestAlgorithm>SHA-
1</premis:messageDigestAlgorithm>
                            <premis:messageDigest><!--The SHA-1
checksum (Required)--></premis:messageDigest>
                            <premis:messageDigestOriginator>Library
of Congress</premis:messageDigestOriginator>
                        </premis:fixity>
                        <premis:size>1</premis:size><!--Size of the
file in bytes. (Required)-->
                        <premis:format>
                            <premis:formatDesignation>

<premis:formatName>image/jp2</premis:formatName>
                            </premis:formatDesignation>
                        </premis:format>
                    </premis:objectCharacteristics>
                    <premis:creatingApplication>
                        <premis:creatingApplicationName><!--
creatingApplicationName (optional)--></premis:creatingApplicationName>
```

```
                                    <premis:creatingApplicationVersion><!--
creatingApplicationVersion (optional)--
></premis:creatingApplicationVersion>
                                    <premis:dateCreatedByApplication>2004-11-
22T00:00:00</premis:dateCreatedByApplication><!--
dateCreatedByApplication (optional)-->
                            </premis:creatingApplication>
                    </premis:object>
                </xmlData>
            </mdWrap>
        </techMD>
        <!--techMD is repeated for service image for each page-->

        <!--MIX technical metadata for primary service image for page
1.  Added by Trusted Validator.-->
        <techMD ID="primaryServiceMix1">
            <mdWrap MDTYPE="NISOIMG" LABEL="NISO still image metadata">
                <xmlData>
                    <mix:mix>
                        <mix:BasicImageParameters>
                            <mix:Format>
                                <mix:Compression>

<mix:CompressionScheme>1</mix:CompressionScheme><!--Compression scheme
(required)-->
                                </mix:Compression>
                                <mix:PhotometricInterpretation>

<mix:ColorSpace>0</mix:ColorSpace><!--Colorspace (required)-->
                                </mix:PhotometricInterpretation>
                            </mix:Format>
                        </mix:BasicImageParameters>
                        <mix:ImagingPerformanceAssessment>
                            <mix:SpatialMetrics>

<mix:SamplingFrequencyUnit>2</mix:SamplingFrequencyUnit><!--
SamplingFrequencyUnity (required)-->

<mix:XSamplingFrequency>400</mix:XSamplingFrequency><!--
XSamplingFrequency (required)-->

<mix:YSamplingFrequency>400</mix:YSamplingFrequency><!--
YSamplingFrequency (required)-->
                                <mix:ImageWidth>1</mix:ImageWidth><!--
ImageWidth (required)-->
                                <mix:ImageLength>1</mix:ImageLength><!-
-ImageLength (required)-->
                            </mix:SpatialMetrics>
                            <mix:Energetics>

<mix:BitsPerSample>8</mix:BitsPerSample><!--Bits per sample (required)-
->
                            </mix:Energetics>
                        </mix:ImagingPerformanceAssessment>
                    </mix:mix>
                </xmlData>
            </mdWrap>
```

```
        </techMD>
        <!--techMD is repeated for primary service image for each page,
unless a primary service image has identical values as a previous
primary service image, in which case repeating is not necessary-->

        <!--PREMIS technical metadata for other derivative for page 1.
Added by Trusted Validator.-->
        <techMD ID="otherDerivativePremis1">
            <mdWrap MDTYPE="OTHER" OTHERMDTYPE="PREMIS" LABEL="PREMIS
object metadata">
                <xmlData>
                    <premis:object>
                        <premis:objectCharacteristics>
                            <premis:fixity>
                                <premis:messageDigestAlgorithm>SHA-
1</premis:messageDigestAlgorithm>
                                <premis:messageDigest><!--The SHA-1
checksum (Required)--></premis:messageDigest>
                                <premis:messageDigestOriginator>Library
of Congress</premis:messageDigestOriginator>
                            </premis:fixity>
                            <premis:size>1</premis:size><!--Size of the
file in bytes. (Required)-->
                            <premis:format>
                                <premis:formatDesignation>

<premis:formatName>application/pdf</premis:formatName>
                                </premis:formatDesignation>
                            </premis:format>
                        </premis:objectCharacteristics>
                        <premis:creatingApplication>
                            <premis:creatingApplicationName><!--
creatingApplicationName (optional)--></premis:creatingApplicationName>
                            <premis:creatingApplicationVersion><!--
creatingApplicationVersion (optional)--
></premis:creatingApplicationVersion>
                            <premis:dateCreatedByApplication>2004-11-
22T00:00:00</premis:dateCreatedByApplication><!--
dateCreatedByApplication (optional)-->
                        </premis:creatingApplication>
                    </premis:object>
                </xmlData>
            </mdWrap>
        </techMD>
        <!--techMD is repeated for derivative image for each page-->

        <!--PREMIS technical metadata for ocr text for page 1. Added by
Trusted Validator.-->
        <techMD ID="ocrTextPremis1">
            <mdWrap MDTYPE="OTHER" OTHERMDTYPE="PREMIS" LABEL="PREMIS
object metadata">
                <xmlData>
                    <premis:object>
                        <premis:objectCharacteristics>
                            <premis:fixity>
                                <premis:messageDigestAlgorithm>SHA-
1</premis:messageDigestAlgorithm>
```

```
                                    <premis:messageDigest><!--The SHA-1
checksum (Required)--></premis:messageDigest>
                                    <premis:messageDigestOriginator>Library
of Congress</premis:messageDigestOriginator>
                            </premis:fixity>
                            <premis:size>1</premis:size><!--Size of the
file in bytes. (Required)-->
                            <premis:format>
                                    <premis:formatDesignation>

<premis:formatName>text/xml</premis:formatName>
                                    </premis:formatDesignation>
                            </premis:format>
                    </premis:objectCharacteristics>
                    <premis:creatingApplication>
                            <premis:creatingApplicationName><!--
creatingApplicationName (optional)--></premis:creatingApplicationName>
                            <premis:creatingApplicationVersion><!--
creatingApplicationVersion (optional)--
></premis:creatingApplicationVersion>
                            <premis:dateCreatedByApplication>2004-11-
22T00:00:00</premis:dateCreatedByApplication><!--
dateCreatedByApplication (optional)-->
                    </premis:creatingApplication>
                </premis:object>
            </xmlData>
        </mdWrap>
    </techMD>
    <!--techMD is repeated for ocr text for each page.-->

    <!--XML digital signature for Mets.  Added by Trusted Validator
prior to ingest.-->
    <digiprovMD ID="dsig">
        <mdWrap LABEL="Mets record validation signature"
MDTYPE="OTHER" OTHERMDTYPE="XML-Signature">
            <xmlData>
                <dsig:Signature>
                    <dsig:SignedInfo>
                        <dsig:CanonicalizationMethod
Algorithm="http://www.w3.org/TR/2001/REC-xml-c14n-
20010315#WithComments"/>
                        <dsig:SignatureMethod
Algorithm="http://www.w3.org/2000/09/xmldsig#dsa-SHA-11"/>
                        <dsig:Reference URI="">
                            <dsig:Transforms>
                                <dsig:Transform
Algorithm="http://www.w3.org/2000/09/xmldsig#enveloped-signature" />
                            </dsig:Transforms>
                            <dsig:DigestMethod
Algorithm="http://www.w3.org/2000/09/xmldsig#SHA-11" />
                            <dsig:DigestValue
>a6YyprHlfzy49l30S/sTFDviaE8=</dsig:DigestValue>
                        </dsig:Reference>
                    </dsig:SignedInfo>
                    <dsig:SignatureValue
>FmuNM4sto5vVntkO1+WWi4H0DaiWvzU8Bak7j3K9wQxEWVDvt7EoOg==</dsig:Signatu
reValue>
```

```
                                       <dsig:KeyInfo>
                                            <dsig:KeyValue>
                                                 <dsig:DSAKeyValue>

<dsig:P>/X9TgR11EilS30qcLuzk5/YRt1I870QAwx4/gLZRJmlFXUAiUftZPY1Y+r/F9bo
w9subVWzXgTuA
HTRv8mZgt2uZUKWkn5/oBHsQIsJPu6nX/rfGG/g7V+fGqKYVDwT7g/bTxR7DAjVUE1oWkTL
2dfOu
K2HXKu/yIgMZndFIAcc=</dsig:P><dsig:Q>l2BQjxUjC8yykrmCouuEC/BYHPU=</dsig
:Q><dsig:G>9+GghdabPd7LvKtcNrhXuXmUr7v6OuqC+VdMCz0HgmdRWVeOutRZT+ZxBxCB
gLRJFnEj6EwoFhO3
zwkyjMim4TwWeotUfI0o4KOuHiuzpnWRbqN/C/ohNWLx+2J6ASQ7zKTxvqhRkImog9/hWuW
fBpKL
Zl6Ae1UlZAFMO/7PSSo=</dsig:G><dsig:Y>IQy0sz7MUZsAcnddcGlz8Nc5Cx68xbVNCI
8DjOrahDarTyvFmOymnYNJ+TkmkGJZxGEPt6lGFaRX
mFKwPaq9P6SHOZc567l2XZMIdIIeMrlZ2xOJjzC5H4EvOLxEQgQP3p/9cywaA/fNkRdSfF9
MeLwL
XdJkkjE6zEN7eYgSjkM=</dsig:Y>
                                                 </dsig:DSAKeyValue>
                                            </dsig:KeyValue>
                                            <dsig:KeyName>NDNP1.0</dsig:KeyName>
                                       </dsig:KeyInfo>
                                  </dsig:Signature>
                             </xmlData>
                        </mdWrap>
                   </digiprovMD>
              </amdSec>

              <!--FILE SECTION-->
              <fileSec>
                   <!--File group for page 1-->
                   <fileGrp ID="pageFileGrp1">
                        <file ID="masterFile1" USE="master" ADMID="masterPremis1
masterMix1"><!--ADMID should reference appropriate IDs.-->
                             <FLocat LOCTYPE="OTHER" OTHERLOCTYPE="file"
xlink:href="./page.tif" /><!--xlink:href should be populated with a
file path relative to this file-->
                        </file>
                        <file ID="serviceFile1" USE="service"
ADMID="primaryServicePremis1 primaryServiceMix1"><!--ADMID should
reference appropriate IDs.-->
                             <FLocat LOCTYPE="OTHER" OTHERLOCTYPE="file"
xlink:href="./page.jp2" /><!--xlink:href should be populated with a
file path relative to this file-->
                        </file>
                        <file ID="otherDerivativeFile1" USE="derivative"
ADMID="otherDerivativePremis1"><!--ADMID should reference appropriate
IDs.-->
                             <FLocat LOCTYPE="OTHER" OTHERLOCTYPE="file"
xlink:href="./page.pdf" /><!--xlink:href should be populated with a
file path relative to this file-->
                        </file>
                        <file ID="ocrFile1" USE="ocr" ADMID="ocrTextPremis1"><!--
ADMID should reference appropriate IDs.-->
                             <FLocat LOCTYPE="OTHER" OTHERLOCTYPE="file"
xlink:href="./page.xml" /><!--xlink:href should be populated with a
file path relative to this file-->
```

```
            </file>
        </fileGrp>
        <!--Repeat fileGrp for each page-->
    </fileSec>


    <!--STRUCTURAL MAP-->
    <structMap xmlns:np="urn:library-of-congress:ndnp:mets:newspaper">
        <div TYPE="np:issue" DMDID="issueModsBib"><!--DMDID should
reference appropriate IDs.-->
            <div TYPE="np:section" DMDID="sectionModsBib1"><!--[An
issue may or may not have sections.  If the issue does not have
sections, this div level will be omitted.]-->
                <div TYPE="np:page" DMDID="pageModsBib1"><!--DMDID
should reference appropriate IDs.-->
                    <fptr FILEID="masterFile1" />
                    <fptr FILEID="serviceFile1" />
                    <fptr FILEID="otherDerivativeFile1" />
                    <fptr FILEID="ocrFile1" />
                </div>
                <!--div repeated for each page.-->
            </div>
            <!--div repeated for each section.-->
        </div>
    </structMap>
</mets>
```

REEL XML Metadata Template, v. 1.5

```
<?xml version="1.0" encoding="UTF-8"?>
<!--NDNP Reel profile-->
<!--Justin Littman, OSI-->
<!--Version 1.5-->
<!--Note that a reel may have additional elements/attributes
added/removed as
    it passes from awardee/vendor and through the pre-ingest and ingest
processes.
    Elements/attributes that do not apply to all steps have been
indicated below.
    amdSec (and children):  Should only be added by Validation
Application.
    RECORDSTATUS:  Should only be added by Validation Application.
-->
<!--INSTRUCTIONS:
1.  Omit all comments.  (Comments in all caps may be left in for
clarity.)
2.  Treat all attribute values in brackets as comments that should be
replaced with the appropriate values.
3.  Omit xsi:schemaLocation
4.  Reel sequence numbers should not be assigned to tech targets.  Reel
sequence numbers should begin with the first image on
     the content reel beginning with 1.
-->
<!--CHANGES in 1.5:
1.  Added reference to digital signature in metsHdr.
```

```
2.  Removed individual file digital signatures.
3.  Added physicalDescription to page mods to be consistent with issue.
4.  Added dmdSec for tech targets.
-->


<mets TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"
PROFILE="urn:library-of-congress:mets:profiles:ndnp:microfilmReel:v1.4"
LABEL="[reel number]"
     xmlns:mix="http://www.loc.gov/mix/"
xmlns:ndnp="http://www.loc.gov/ndnp"
xmlns:premis="http://www.loc.gov/standards/premis"
xmlns:mods="http://www.loc.gov/mods/v3"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns:xlink="http://www.w3.org/1999/xlink"
xmlns:dsig="http://www.w3.org/2000/09/xmldsig#"
xmlns="http://www.loc.gov/METS/"
xsi:schemaLocation="http://www.loc.gov/METS/ ../schema/ndnpmets.v1-
4.xsd http://www.loc.gov/mods/v3 ../schema/ndnpmods-3-1.xsd
http://www.loc.gov/ndnp ../schema/ndnp.xsd
http://www.loc.gov/standards/premis ../schema/object.xsd
http://www.loc.gov/mix/ http://www.loc.gov/standards/mix/mix.xsd
http://www.w3.org/2000/09/xmldsig# ../schema/xmldsig-core-schema.xsd">

    <!--METS HEADER-->
    <metsHdr CREATEDATE="2004-11-12T09:00:00" RECORDSTATUS="Validated"
ADMID="digSig"><!--CREATEDATE should be populated with creation date of
the record.  RECORDSTATUS should only be set by Validation
Application.-->
        <agent ROLE="CREATOR" TYPE="ORGANIZATION">
            <name><!--Awardee, if the awardee created the METS record.
If LC created the METS record, should be "Library of Congress"--
></name>
        </agent>
    </metsHdr>

    <!--Descriptive metadata for tech target-->
    <dmdSec ID="techTargetModsBib">
        <mdWrap MDTYPE="MODS" LABEL="Tech Target metadata">
            <xmlData>
                <mods:mods>
                    <mods:titleInfo>
                            <mods:title>ISO 16067 Calibrated
Target</mods:title>
                    </mods:titleInfo>
                    <mods:note type="agencyResponsibleForReproduction"
displayLabel="[digitalResponsibleInstitution]"><!--
digitalResponsibleInstitutionCode (from
http://www.loc.gov/marc/organizations)--></mods:note>
                </mods:mods>
            </xmlData>
        </mdWrap>
    </dmdSec>

    <!--Descriptive metadata for target-->
    <dmdSec ID="targetModsBib1">
        <mdWrap MDTYPE="MODS" LABEL="Target metadata">
            <xmlData>
```

```
                        <mods:mods>
                            <mods:relatedItem type="original">
                                <mods:physicalDescription>
                                    <mods:form type="microfilm" /><!--Microfilm
is only valid value.-->
                                </mods:physicalDescription>
                                <mods:identifier type="reel number"><!--
reelNumber (required)--></mods:identifier>
                                <mods:identifier type="reel sequence
number"><!--reelSequenceNumber (required)--></mods:identifier>
                                <mods:location>
                                    <mods:physicalLocation authority="marcorg"
displayLabel="[sourceRepository]"><!--sourceRepositoryCode (from
http://www.loc.gov/marc/organizations)--></mods:physicalLocation>
                                </mods:location>
                            </mods:relatedItem>
                            <mods:note type="agencyResponsibleForReproduction"
displayLabel="[digitalResponsibleInstitution]"><!--
digitalResponsibleInstitutionCode (from
http://www.loc.gov/marc/organizations)--></mods:note>
                        </mods:mods>
                    </xmlData>
                </mdWrap>
        </dmdSec>
    <!--Repeat dmdSec for each page-->

    <amdSec>
        <!--TECHNICAL METADATA-->
        <!--NDNP technical metadata for reel-->
        <techMD ID="reelTechMD">
            <mdWrap MDTYPE="OTHER" OTHERMDTYPE="NDNP Reel Technical
Metadata" LABEL="NDNP technical metadata for microfilm reel">
                <xmlData>
                    <ndnp:reelTechMD>
                        <ndnp:titles>[Titles (required)]</ndnp:titles>
                        <ndnp:startDate>1919-01-01</ndnp:startDate><!--
Start Date (required)-->
                        <ndnp:endDate>1920-01-01</ndnp:endDate><!--End
Date (required)-->
                        <ndnp:position>1a</ndnp:position><!--Position
(required)-->
                        <ndnp:reductionRatio>[Reduction Ratio
(required)]</ndnp:reductionRatio>
                        <ndnp:captureResolutionOriginal
measurement="pixels/inch">300</ndnp:captureResolutionOriginal><!--
Capture resolution of original (required)-->
                        <ndnp:captureResolutionFilm
measurement="pixels/inch">6000</ndnp:captureResolutionFilm><!--Capture
resolution of film (required)-->

<ndnp:guideToContentsPresentFlag>true</ndnp:guideToContentsPresentFlag>
<!--Guide to contents present (required)-->
                        <ndnp:guideToContentsString>[Guide to Contents
String (optional)]</ndnp:guideToContentsString>
                        <ndnp:dateMicrofilmCreated>[Date microfilm
created (required)]</ndnp:dateMicrofilmCreated>
```

```
<ndnp:looseLeavesFlag>true</ndnp:looseLeavesFlag><!--Loose leaves
(required)-->

<ndnp:boundVolumeFlag>true</ndnp:boundVolumeFlag><!--Loose leaves
(required)-->
                            <ndnp:comments>[Comments
(optional)]</ndnp:comments>
                            <ndnp:dimensions>[Dimensions
(required)]</ndnp:dimensions>
                            <ndnp:pagesPerIssue>4</ndnp:pagesPerIssue><!--
Pages per issue (required)-->

<ndnp:numberOfResolutionTargets>5</ndnp:numberOfResolutionTargets><!--
Number of resolution targets (required)-->

<ndnp:resolutionOfMaster>7.1</ndnp:resolutionOfMaster><!--Resolution of
master (Not required. Provide if available.)-->
                            <ndnp:resolutionCommentMaster>[Resolution
comment master (optional)]</ndnp:resolutionCommentMaster>

<ndnp:densityReadingMaster>0.91</ndnp:densityReadingMaster><!--Density
reading 1 master (required)-->

<ndnp:densityReadingMaster>0.91</ndnp:densityReadingMaster><!--Density
reading 2 master (required)-->

<ndnp:densityReadingMaster>0.91</ndnp:densityReadingMaster><!--Density
reading 3 master (required)-->

<ndnp:densityReadingMaster>0.91</ndnp:densityReadingMaster><!--Density
reading 4 master (required)-->

<ndnp:densityReadingMaster>0.91</ndnp:densityReadingMaster><!--Density
reading 5 master (required)-->

<ndnp:densityReadingMaster>0.91</ndnp:densityReadingMaster><!--Density
reading 6 master (required)-->

<ndnp:densityReadingMaster>0.91</ndnp:densityReadingMaster><!--Density
reading 7 master (required)-->

<ndnp:densityReadingMaster>0.91</ndnp:densityReadingMaster><!--Density
reading 8 master (required)-->

<ndnp:densityReadingMaster>0.91</ndnp:densityReadingMaster><!--Density
reading 9 master (required)-->

<ndnp:densityReadingMaster>0.91</ndnp:densityReadingMaster><!--Density
reading 10 master (required)-->

<ndnp:averageDensityMaster>0.91</ndnp:averageDensityMaster><!--Average
density master (required)-->
                            <ndnp:dminMaster>0.20</ndnp:dminMaster><!--Dmin
master (required)-->

<ndnp:resolutionOfDuplicateNegative>7.1</ndnp:resolutionOfDuplicateNega
```

tive><!--Resolution of DuplicateNegative (Not required. Provide if available.)-->

<ndnp:resolutionCommentDuplicateNegative>[Resolution comment DuplicateNegative (optional)]</ndnp:resolutionCommentDuplicateNegative>

<ndnp:densityReadingDuplicateNegative>0.91</ndnp:densityReadingDuplicateNegative><!--Density reading 1 DuplicateNegative (required)-->

<ndnp:densityReadingDuplicateNegative>0.91</ndnp:densityReadingDuplicateNegative><!--Density reading 2 DuplicateNegative (required)-->

<ndnp:densityReadingDuplicateNegative>0.91</ndnp:densityReadingDuplicateNegative><!--Density reading 3 DuplicateNegative (required)-->

<ndnp:densityReadingDuplicateNegative>0.91</ndnp:densityReadingDuplicateNegative><!--Density reading 4 DuplicateNegative (required)-->

<ndnp:densityReadingDuplicateNegative>0.91</ndnp:densityReadingDuplicateNegative><!--Density reading 5 DuplicateNegative (required)-->

<ndnp:densityReadingDuplicateNegative>0.91</ndnp:densityReadingDuplicateNegative><!--Density reading 6 DuplicateNegative (required)-->

<ndnp:densityReadingDuplicateNegative>0.91</ndnp:densityReadingDuplicateNegative><!--Density reading 7 DuplicateNegative (required)-->

<ndnp:densityReadingDuplicateNegative>0.91</ndnp:densityReadingDuplicateNegative><!--Density reading 8 DuplicateNegative (required)-->

<ndnp:densityReadingDuplicateNegative>0.91</ndnp:densityReadingDuplicateNegative><!--Density reading 9 DuplicateNegative (required)-->

<ndnp:densityReadingDuplicateNegative>0.91</ndnp:densityReadingDuplicateNegative><!--Density reading 10 DuplicateNegative (required)-->

<ndnp:averageDensityDuplicateNegative>0.91</ndnp:averageDensityDuplicateNegative><!--Average density DuplicateNegative (required)-->

<ndnp:dminDuplicateNegative>0.20</ndnp:dminDuplicateNegative><!--Dmin DuplicateNegative (required)-->
                    </ndnp:reelTechMD>
                </xmlData>
            </mdWrap>
        </techMD>

        <!--TECHNICAL METADATA.-->
        <!--All technical metadata is added by trusted validator-->
        <!--PREMIS technical metadata for digital master for page 1.
Added by trusted validator.-->
        <techMD ID="masterPremis1">
            <mdWrap MDTYPE="OTHER" OTHERMDTYPE="PREMIS" LABEL="PREMIS
object metadata">
                <xmlData>
                    <premis:object>
                        <premis:objectCharacteristics>
                            <premis:fixity>

```
                                <premis:messageDigestAlgorithm>SHA-
1</premis:messageDigestAlgorithm>
                                <premis:messageDigest><!--The SHA-1
checksum (Required)--></premis:messageDigest>
                                <premis:messageDigestOriginator>Library
of Congress</premis:messageDigestOriginator>
                            </premis:fixity>
                            <premis:size>1</premis:size><!--Size of the
file in bytes. (Required)-->
                            <premis:format>
                                <premis:formatDesignation>

<premis:formatName>image/tiff</premis:formatName>
                                </premis:formatDesignation>
                            </premis:format>
                        </premis:objectCharacteristics>
                        <premis:creatingApplication>
                            <premis:creatingApplicationName><!--
creatingApplicationName (optional)--></premis:creatingApplicationName>
                            <premis:creatingApplicationVersion><!--
creatingApplicationVersion (optional)--
></premis:creatingApplicationVersion>
                            <premis:dateCreatedByApplication>2004-11-
22T00:00:00</premis:dateCreatedByApplication><!--
dateCreatedByApplication (required for master)-->
                        </premis:creatingApplication>
                    </premis:object>
                </xmlData>
            </mdWrap>
        </techMD>
        <!--techMD is repeated for master image for each page-->

        <!--MIX technical metadata for digital master for page 1.
Added by Trusted Validator.-->
        <techMD ID="masterMix1">
            <mdWrap MDTYPE="NISOIMG" LABEL="NISO still image metadata"
>
                <xmlData>
                    <mix:mix>
                        <mix:BasicImageParameters>
                            <mix:Format>
                                <mix:Compression>

<mix:CompressionScheme>1</mix:CompressionScheme><!--Compression scheme
(required)-->
                                </mix:Compression>
                                <mix:PhotometricInterpretation>

<mix:ColorSpace>0</mix:ColorSpace><!--Colorspace (required)-->
                                </mix:PhotometricInterpretation>
                            </mix:Format>
                        </mix:BasicImageParameters>
                        <mix:ImageCreation>

<mix:SourceType>Microfilm</mix:SourceType><!--sourceType (required for
master)-->
```

```
                                        <mix:ImageProducer><!--awardee; vendor
(required for master; optional for others)--></mix:ImageProducer>
                                        <mix:ScanningSystemCapture>
                                                <mix:ScanningSystemHardware>
                                                        <mix:ScannerManufacturer><!--
scannerManufacturer (required for master)--></mix:ScannerManufacturer>
                                                        <mix:ScannerModel><!--
scannerModelName (required for master)--></mix:ScannerModel>
                                                </mix:ScanningSystemHardware>
                                        </mix:ScanningSystemCapture>
                                </mix:ImageCreation>
                                <mix:ImagingPerformanceAssessment>
                                        <mix:SpatialMetrics>

<mix:SamplingFrequencyUnit>2</mix:SamplingFrequencyUnit><!--
SamplingFrequencyUnit (required)-->

<mix:XSamplingFrequency>400</mix:XSamplingFrequency><!--
XSamplingFrequency (required)-->

<mix:YSamplingFrequency>400</mix:YSamplingFrequency><!--
YSamplingFrequency (required)-->
                                                <mix:ImageWidth>1</mix:ImageWidth><!--
ImageWidth (required)-->
                                                <mix:ImageLength>1</mix:ImageLength><!-
-ImageLength (required)-->
                                        </mix:SpatialMetrics>
                                        <mix:Energetics>

<mix:BitsPerSample>8</mix:BitsPerSample><!--Bits per sample (required)-
->
                                        </mix:Energetics>
                                </mix:ImagingPerformanceAssessment>
                        </mix:mix>
                </xmlData>
            </mdWrap>
        </techMD>
        <!--techMD is repeated for master image for each page, unless a
master image has identical values as a previous master image, in which
case repeating is not necessary-->

        <!--PREMIS technical metadata for primary service image for
page 1  Added by Trusted Validator.-->
        <techMD ID="primaryServicePremis1">
            <mdWrap MDTYPE="OTHER" OTHERMDTYPE="PREMIS" LABEL="PREMIS
object metadata">
                <xmlData>
                    <premis:object>
                        <premis:objectCharacteristics>
                            <premis:fixity>
                                <premis:messageDigestAlgorithm>SHA-
1</premis:messageDigestAlgorithm>
                                <premis:messageDigest><!--The SHA-1
checksum (Required)--></premis:messageDigest>
                                <premis:messageDigestOriginator>Library
of Congress</premis:messageDigestOriginator>
                            </premis:fixity>
```

```xml
                                    <premis:size>1</premis:size><!--Size of the
file in bytes. (Required)-->
                                    <premis:format>
                                        <premis:formatDesignation>

<premis:formatName>image/jp2</premis:formatName>
                                        </premis:formatDesignation>
                                    </premis:format>
                            </premis:objectCharacteristics>
                            <premis:creatingApplication>
                                    <premis:creatingApplicationName><!--
creatingApplicationName (optional)--></premis:creatingApplicationName>
                                    <premis:creatingApplicationVersion><!--
creatingApplicationVersion (optional)--
></premis:creatingApplicationVersion>
                                    <premis:dateCreatedByApplication>2004-11-
22T00:00:00</premis:dateCreatedByApplication><!--
dateCreatedByApplication (optional)-->
                            </premis:creatingApplication>
                    </premis:object>
                </xmlData>
            </mdWrap>
        </techMD>
        <!--techMD is repeated for service image for each page-->

        <!--MIX technical metadata for primary service image for page
1.  Added by Trusted Validator.-->
        <techMD ID="primaryServiceMix1">
            <mdWrap MDTYPE="NISOIMG" LABEL="NISO still image metadata">
                <xmlData>
                    <mix:mix>
                        <mix:BasicImageParameters>
                            <mix:Format>
                                    <mix:Compression>

<mix:CompressionScheme>1</mix:CompressionScheme><!--Compression scheme
(required)-->
                                    </mix:Compression>
                                    <mix:PhotometricInterpretation>

<mix:ColorSpace>0</mix:ColorSpace><!--Colorspace (required)-->
                                    </mix:PhotometricInterpretation>
                            </mix:Format>
                        </mix:BasicImageParameters>
                        <mix:ImagingPerformanceAssessment>
                            <mix:SpatialMetrics>

<mix:SamplingFrequencyUnit>2</mix:SamplingFrequencyUnit><!--
SamplingFrequencyUnity (required)-->

<mix:XSamplingFrequency>400</mix:XSamplingFrequency><!--
XSamplingFrequency (required)-->

<mix:YSamplingFrequency>400</mix:YSamplingFrequency><!--
YSamplingFrequency (required)-->
                                    <mix:ImageWidth>1</mix:ImageWidth><!--
ImageWidth (required)-->
```

```
                                  <mix:ImageLength>1</mix:ImageLength><!-
-ImageLength (required)-->
                              </mix:SpatialMetrics>
                              <mix:Energetics>

<mix:BitsPerSample>8</mix:BitsPerSample><!--Bits per sample (required)-
->
                              </mix:Energetics>
                          </mix:ImagingPerformanceAssessment>
                      </mix:mix>
                  </xmlData>
              </mdWrap>
          </techMD>
          <!--techMD is repeated for primary service image for each page,
unless a primary service image has identical values as a previous
primary service image, in which case repeating is not necessary-->

          <!--PREMIS technical metadata for other derivative for page 1.
Added by Trusted Validator.-->
          <techMD ID="otherDerivativePremis1">
              <mdWrap MDTYPE="OTHER" OTHERMDTYPE="PREMIS" LABEL="PREMIS
object metadata">
                  <xmlData>
                      <premis:object>
                          <premis:objectCharacteristics>
                              <premis:fixity>
                                  <premis:messageDigestAlgorithm>SHA-
1</premis:messageDigestAlgorithm>
                                  <premis:messageDigest><!--The SHA-1
checksum (Required)--></premis:messageDigest>
                                  <premis:messageDigestOriginator>Library
of Congress</premis:messageDigestOriginator>
                              </premis:fixity>
                              <premis:size>1</premis:size><!--Size of the
file in bytes. (Required)-->
                              <premis:format>
                                  <premis:formatDesignation>

<premis:formatName>application/pdf</premis:formatName>
                                  </premis:formatDesignation>
                              </premis:format>
                          </premis:objectCharacteristics>
                          <premis:creatingApplication>
                              <premis:creatingApplicationName><!--
creatingApplicationName (optional)--></premis:creatingApplicationName>
                              <premis:creatingApplicationVersion><!--
creatingApplicationVersion (optional)--
></premis:creatingApplicationVersion>
                              <premis:dateCreatedByApplication>2004-11-
22T00:00:00</premis:dateCreatedByApplication><!--
dateCreatedByApplication (optional)-->
                          </premis:creatingApplication>
                      </premis:object>
                  </xmlData>
              </mdWrap>
          </techMD>
          <!--techMD is repeated for derivative image for each page-->
```

```
        <!--PREMIS technical metadata for ocr text for page 1. Added by
Trusted Validator.-->
        <techMD ID="ocrTextPremis1">
            <mdWrap MDTYPE="OTHER" OTHERMDTYPE="PREMIS" LABEL="PREMIS
object metadata">
                <xmlData>
                    <premis:object>
                        <premis:objectCharacteristics>
                            <premis:fixity>
                                <premis:messageDigestAlgorithm>SHA-
1</premis:messageDigestAlgorithm>
                                <premis:messageDigest><!--The SHA-1
checksum (Required)--></premis:messageDigest>
                                <premis:messageDigestOriginator>Library
of Congress</premis:messageDigestOriginator>
                            </premis:fixity>
                            <premis:size>1</premis:size><!--Size of the
file in bytes. (Required)-->
                            <premis:format>
                                <premis:formatDesignation>

<premis:formatName>text/xml</premis:formatName>
                                </premis:formatDesignation>
                            </premis:format>
                        </premis:objectCharacteristics>
                        <premis:creatingApplication>
                            <premis:creatingApplicationName><!--
creatingApplicationName (optional)--></premis:creatingApplicationName>
                            <premis:creatingApplicationVersion><!--
creatingApplicationVersion (optional)--
></premis:creatingApplicationVersion>
                            <premis:dateCreatedByApplication>2004-11-
22T00:00:00</premis:dateCreatedByApplication><!--
dateCreatedByApplication (optional)-->
                        </premis:creatingApplication>
                    </premis:object>
                </xmlData>
            </mdWrap>
        </techMD>
        <!--techMD is repeated for ocr text for each page.-->

        <!--XML digital signature for Mets.  Added by Trusted Validator
prior to ingest.-->
        <digiprovMD ID="digSig">
            <mdWrap LABEL="Mets record validation signature"
MDTYPE="OTHER" OTHERMDTYPE="XML-Signature">
                <xmlData>
                    <dsig:Signature>
                        <dsig:SignedInfo>
                            <dsig:CanonicalizationMethod
Algorithm="http://www.w3.org/TR/2001/REC-xml-c14n-
20010315#WithComments"/>
                            <dsig:SignatureMethod
Algorithm="http://www.w3.org/2000/09/xmldsig#dsa-SHA-11"/>
                            <dsig:Reference URI="">
                                <dsig:Transforms>
```

```
                                            <dsig:Transform
Algorithm="http://www.w3.org/2000/09/xmldsig#enveloped-signature" />
                                        </dsig:Transforms>
                                        <dsig:DigestMethod
Algorithm="http://www.w3.org/2000/09/xmldsig#SHA-11" />
                                        <dsig:DigestValue
>a6YyprHlfzy49l30S/sTFDviaE8=</dsig:DigestValue>
                                    </dsig:Reference>
                            </dsig:SignedInfo>
                            <dsig:SignatureValue
>FmuNM4sto5vVntkO1+WWi4H0DaiWvzU8Bak7j3K9wQxEWVDvt7EoOg==</dsig:Signatu
reValue>
                            <dsig:KeyInfo>
                                <dsig:KeyValue>
                                    <dsig:DSAKeyValue>

<dsig:P>/X9TgR11EilS30qcLuzk5/YRt1I870QAwx4/gLZRJmlFXUAiUftZPY1Y+r/F9bo
w9subVWzXgTuA
HTRv8mZgt2uZUKWkn5/oBHsQIsJPu6nX/rfGG/g7V+fGqKYVDwT7g/bTxR7DAjVUE1oWkTL
2dfOu
K2HXKu/yIgMZndFIAcc=</dsig:P><dsig:Q>l2BQjxUjC8yykrmCouuEC/BYHPU=</dsig
:Q><dsig:G>9+GghdabPd7LvKtcNrhXuXmUr7v6OuqC+VdMCz0HgmdRWVeOutRZT+ZxBxCB
gLRJFnEj6EwoFhO3
zwkyjMim4TwWeotUfI0o4KOuHiuzpnWRbqN/C/ohNWLx+2J6ASQ7zKTxvqhRkImog9/hWuW
fBpKL
Zl6Ae1UlZAFMO/7PSSo=</dsig:G><dsig:Y>IQy0sz7MUZsAcnddcGlz8Nc5Cx68xbVNCI
8DjOrahDarTyvFmOymnYNJ+TkmkGJZxGEPt6lGFaRX
mFKwPaq9P6SHOZc567l2XZMIdIIeMrlZ2xOJjzC5H4EvOLxEQgQP3p/9cywaA/fNkRdSfF9
MeLwL
XdJkkjE6zEN7eYgSjkM=</dsig:Y>
                                    </dsig:DSAKeyValue>
                                </dsig:KeyValue>
                                <dsig:KeyName>NDNP.1.0</dsig:KeyName>
                            </dsig:KeyInfo>
                        </dsig:Signature>
                    </xmlData>
                </mdWrap>
            </digiprovMD>

    </amdSec>

    <!--FILE SECTION-->
    <fileSec>
        <!--File group for first target-->
        <fileGrp ID="targetFileGrp1">
            <file ID="masterFile1" USE="master" ADMID="masterPremis1
masterMix1"><!--ADMID should reference appropriate IDs.-->
                <FLocat LOCTYPE="OTHER" OTHERLOCTYPE="file"
xlink:href="../page.tif" /><!--xlink:href should be populated with a
file path relative to this file-->
            </file>
            <file ID="serviceFile1" USE="service"
ADMID="primaryServicePremis1 primaryServiceMix1"><!--ADMID should
reference appropriate IDs.-->
                <FLocat LOCTYPE="OTHER" OTHERLOCTYPE="file"
xlink:href="../page.jp2" /><!--xlink:href should be populated with a
file path relative to this file-->
```

```
                </file>
                <file ID="otherDerivativeFile1" USE="derivative"
ADMID="otherDerivativePremis1"><!--ADMID should reference appropriate
IDs.-->
                    <FLocat LOCTYPE="OTHER" OTHERLOCTYPE="file"
xlink:href="../page.pdf" /><!--xlink:href should be populated with a
file path relative to this file-->
                </file>
                <file ID="ocrFile1" USE="ocr" ADMID="ocrTextPremis1"><!--
ADMID should reference appropriate IDs.-->
                    <FLocat LOCTYPE="OTHER" OTHERLOCTYPE="file"
xlink:href="../page.xml" /><!--xlink:href should be populated with a
file path relative to this file-->
                </file>
            </fileGrp>
            <!--Repeat fileGrp for each target-->
        </fileSec>

        <!--STRUCTURAL MAP-->
        <structMap xmlns:np="urn:library-of-congress:ndnp:mets:newspaper">
            <div TYPE="np:reel" ADMID="reelTechMD">
                <!--The techtargetreel are the targets that LC requires be
shot before each reel.  Mandatory.-->
                <div TYPE="np:techtargetreel">
                    <div TYPE="np:target" DMDID="techTargetModsBib">
                        <fptr FILEID="masterFile1" />
                        <fptr FILEID="serviceFile1" />
                        <fptr FILEID="otherDerivativeFile1" />
                        <fptr FILEID="ocrFile1" />
                    </div>
                    <!--div repeated for each tech target.  There must be
5.-->
                </div>
                <div TYPE="np:target" DMDID="targetModsBib1"><!--DMDID
should reference appropriate IDs.-->
                    <!--Note:  To keep this template brief, the same files
are being re-used below.  In a real instance, these should be different
files than above.-->
                    <fptr FILEID="masterFile1" />
                    <fptr FILEID="serviceFile1" />
                    <fptr FILEID="otherDerivativeFile1" />
                    <fptr FILEID="ocrFile1" />
                </div>
                <!--div repeated for each target.-->
            </div>
        </structMap>

</mets>
```

BATCH XML Metadata Template, v. 1.2

```
<!--NDNP Batch profile-->
<!--Justin Littman, OSI-->
<!--Version 1.2-->
```

```
<!--INSTRUCTIONS:
1.  Omit all comments.  (Comments in all caps may be left in for
clarity.)
2.  Treat all attribute values in brackets as comments that should be
replaced with the appropriate values.
3.  Omit xsi:schemaLocation
4.  Omit any element which takes a value if there is no value.  For
example, omit the marcxmlBibRecord element if a MarcXml file is not
being provided.
-->
<!--CHANGES IN 1.2:
1.  Added name attribute.
-->
<batch xsi:schemaLocation="http://www.loc.gov/ndnp
../schema/ndnpBatch.xsd" name="batch_lc_20060106_quebec">
<!--name is the name of the batch.-->
<!--NEWSPAPER TITLE-->
<newspaperTitle lccn="n78890351">./mets.xml</newspaperTitle>
<!--lccn should be lccn, normalized according to
http://www.loc.gov/marc/lccn-namespace.html
-->
<!--Repeat newspaperTitle for each newspaper title-->
<!--
DO NOT DO THIS:
    <newpaperTitle lccn="n78890351" />
-->
<!--ISSUE-->
<issue lccn="n78890351" issueDate="2004-12-03" editionOrder="1">
<!--A relative path of the Mets file.  For example:-->
./1/1/mets.xml
<!--lccn should be a normalized LCCN.-->
</issue>
<issue lccn="n78890351" issueDate="2004-12-03"
editionOrder="1">./1/2/mets.xml</issue>
<!--Repeat newspaperIssue for each newspaper issue-->
<!--REEL-->
<reel reelNumber="[reelNumber]">
<!--A relative path of the Mets file.  For example:-->
./mets.xml
</reel>
</batch>
```