ATTENTION:

This document issued in conjunction with the NEH National Digital Newspaper Program (https://www.neh.gov/grants/preservation/national-digital-newspaper-program) award competition. For general program information, see https://www.loc.gov/ndnp/ .

The National Digital Newspaper Program (NDNP)
Technical Guidelines for Applicants
2020-22 Awards

Document Control

Revision	Date	Summary of Change

UPDATES for 2020-22 NDNP Awards

1. Updated OCR Profile to Version 1.18 which includes option to use *ISO 639-3* language codes when *ISO 639-2* code is not available (Pages 8-10, 42)

Overview of Technical Approach for 2020-22 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 program is to enhance access to newspapers through a Web-accessible NDNP delivery application, open and freely available to the public. This delivery application will provide access to a substantive amount of public domain newspapers selected for their historical value with broad geographic coverage and consistent of sustainable digital assets to support effective use by scholars and the general public. This award cycle is a continuation of the existing program.

As the program progresses, 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) primarily from microfilm
- OCR with word-bounding boxes, 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 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.

Applicants with access to a considerable amount of previously digitized content fitting the intellectual scope of NDNP may propose repurposing some amount of such content for

inclusion in the program. Previously digitized source files eligible for repurposing may include TIFF images scanned as bitonal, grayscale, or color, and reprocessed to technical specifications consistent with existing NDNP digital objects (i.e. all technical deliverables meet NDNP technical requirements except source files). Awardees will be required to submit an initial set of example source files to LC for review. These steps will ensure conformance with NDNP technical standards and to ensure the content will integrate into the substantive amount of NDNP content already available online. More details will be available at the post-award awardees' meeting.

Additionally, while the technical approach is primarily based on scanning from microfilm, some pre-approved amounts of NDNP deliveries may consist of images scanned directly from original newsprint, providing that technical specifications are met.

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 2020-22 deliverables.

For each title

- Up-to-date MARC record from the CONSER OCLC database, fully conformant to current standards for cataloging U.S. **print** newspapers [original format only, not microform],
- Additional title-level metadata related to the title run(s) digitized and delivered (see Appendix A: Digital Asset Metadata Elements), and
- Newspaper History Essay – scope and content (in English) of each title, history, and significance – 500 words.

Note: Conformant MARC records for newspaper titles selected for NDNP should be present or made available by awardees via the OCLC WorldCat database prior to submitting digitized data to LC.

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),
- 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, to sequence pages within issue or section; and to identify associated 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).

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.

For the award

- List of newspaper titles, and associated publishing information, identified as having non-NDNP digitized archives available on the Web, as described in the NEH Program Guidelines (https://www.neh.gov/grants/preservation/national-digital-newspaper-program). This list should include title-level bibliographic information and information on the digital resource available. Further details will be specified at the 2020-22 awardees' meeting, post-award.

Technical Details

Selection

The goals of the overall project, the chronological scope (1690-1963), and the intellectual criteria for selecting newspaper titles for the program are described in NEH Program guidelines (https://www.neh.gov/grants/preservation/national-digital-newspaper-program). Awardee institutions are responsible for determining the copyright status of newspapers published less than 95 years ago. Such titles must satisfy the U.S. Copyright Office's definition of public domain in order to be considered for inclusion in the program (https://www.copyright.gov/help/faq/index.html). More information will be available at the post-award awardees' meeting. 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, within the prescribed date range, 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 or 300 dpi) provides suitable 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 https://www.loc.gov/rr/news/usnp/usnpguidelinesp.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 that contain text blocks exhibiting more than 3 degrees of skew should be de-skewed. 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.

In general, the goal of the NDNP cropping specification is to produce as complete a page image as possible in order to best enable long-term management and access needs into the future. For film created against a white or neutral-colored background, newspaper images may be cropped based on the text block and appropriate padding.

All operations that change the image dimensions, spatial resolution, or orientation (e.g., cropping, de-skewing) 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 the TIFF should have no other enhancements (bitonalization, 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). Capture of these images and creation of derivative files and associated metadata is ***optional***. If included in NDNP deliveries, 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, as specified by Library of Congress, 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. Targets will need to be purchased by individual awardees from a specified source (currently: 35mm Grayscale Preservation Microfilm Target, available from Image Science Associates, http://imagescienceassociates.com, approx. cost $425/strip). (See: http://www.imagescienceassociates.com/mm5/merchant.mvc?Screen=PROD&Store_Code=ISA001&Product_Code=MPT&Categ ory_Code=TARGETS.)

NDNP follows recommendations of the Federal Agencies Digitization Guidelines Initiative (http://www.digitizationguidelines.gov/) and utilizes the standard NISO Z39.87 Data Dictionary – Technical Metadata for Digital Still Images for master images. 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 and rendering of the images (e.g., tile specifications, if used), per Appendix B.

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 requirements: 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 that contain text blocks exhibiting skew of greater than 3 degrees. (Greater skew leads to less accurate OCR.)

5. Crop to include visible edge of page, retaining up to ¼ inch beyond edge**.**

6. Optional: Capture microfilm target frames. These image files will be identified in the reel metadata but will not be used for display.

7. Capture additional scanning resolution targets, i.e., 35mm Grayscale Preservation Microfilm Target, (2 images per reel--target as 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 software 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 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.

Each page digitized for NDNP must be accompanied by OCR text encoded using the ALTO (Analyzed Layout and Text Object) XML schema, Version 2.0, with the additional clarifications stated in Appendix B – File Format Profiles.

Newspapers selected for NDNP digitization must contain text published in languages that have a valid language code according to the international standard *ISO 639-2: Codes for the representation of names of languages: alpha-3 codes* (https://www.loc.gov/standards/iso639-2/). All text in the ALTO XML must be encoded by TextBlock to automate indexing between language sets, using *ISO 639-2*. For languages that are not represented in *ISO 639-2* specification, the *ISO 639-3* (https://iso639-3.sil.org) code may be used if available.

The Library of Congress expects to provide full-text search support in two tiers, based on language. Support will be provided as follows:

Language Analysis Support

The following languages will be supported with (a) exact match techniques and (b) language analysis (stemming, stop words, and/or variants) as applied by the search tools/technologies in use by the Library of Congress' web sites. This list is subject to change based on the supporting software.

Arabic	French	Latvian
Armenian	Galician	Norwegian
Basque	German	Persian
Bengali	Greek	Polish
Brazilian Portuguese	Hebrew	Portuguese
Bulgarian	Hindi	Romanian
Burmese (Myanmar)	Hungarian	Russian
Catalan	Indonesian	Serbian
Chinese	Irish	Spanish
Czech	Italian	Swedish
Danish	Japanese	Thai
Dutch	Khmer	Turkish
English	Korean	Ukrainian
Finnish	Lao	

Exact Match Only Support:

Languages with valid *ISO 639-2* or *639-3* language codes that do not have language analysis support as applied by the search tools/technologies in use by the Library of Congress' web sites will be supported using "exact match" techniques only. Search results will only match exactly what is provided for input, without accounting for stemming, stop words, or variants.

The awardee institution is responsible for providing relevant language expertise to review the quality of the converted content and related metadata. Automated language recognition, if used, should be applied by title, rather than by reel or batch.

Note: When applicable, text printed in Fraktur/black letter fonts must incorporate technical processing that includes Fraktur/black letter specific tools. (This may represent additional cost.) While these fonts were predominately used for German language text, other European languages may be printed in Fraktur/black letter fonts. More information will be available at the annual awardee meeting.

Associated MARC title records should be reviewed and, if necessary, updated in OCLC WorldCat to reflect the presence of non-English text.

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, Version 2.0.
8. All page images must be accompanied by an ALTO XML file containing recognized text.

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 and/or batch, prior to delivery to LC.

For non-English language text:

1. "Language" attribute must be expressed by Textblock, as most appropriate for the content. (Although use of "eng" is recommended for English text, text with unspecified language attribute will default to "eng" (English) for search and discovery purposes).
2. Text should be encoded at the Textblock level, using *ISO 639-2: Codes for the representation of names of languages: alpha-3 codes* (https://www.loc.gov/standards/iso639-2/)**.**
    a. Some languages have both "b" (bibliographic) and "t" (terminology) codes. For NDNP, choose the "b" (bibliographic) code.
    b. For languages that are not represented in *ISO 639-2* specification, the *ISO 639-3* code (https://iso639-3.sil.org) may be used if available.
3. For titles with non-English text, review and, if necessary, update associated MARC record in OCLC WorldCat to reflect non-English language use.
4. When applicable, text printed in Fraktur/black letter fonts must incorporate technical processing that includes Fraktur/black letter specific tools.

Note: Any ALTO text without specific language encoding will be treated as English for text indexing and searching.

Other Derivative Files

In addition to the master TIFF image file and 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 encoded 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 (https://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 ensure LC has access to updated title-level bibliographic records 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 newspaper titles will be based on Library of Congress Control Numbers (LCCNs), since not all historical newspapers have been assigned International Standard Serial Numbers (ISSNs) or another unique identifier. These metadata specifications will be discussed at the awardees' annual meeting.

All newspaper titles selected for digitization under NDNP must be under bibliographic control per U.S. newspaper cataloging guidelines maintained by the Cooperative Online Serials Cataloging (CONSER) program and included in the CONSER database hosted within the OCLC Online Union Catalog (WorldCat). Each title must have a full bibliographic record at the title-level for the original materials (*not microfilm*). If pre-existing, the CONSER records must be reviewed and updated as necessary by the awardee institution. If updated or edited, records should be integrated into OCLC WorldCat before submission of associated digitized pages. LC will harvest the CONSER bibliographic records from OCLC WorldCat at regular intervals for NDNP purposes. If needed, awardees may be required to send .dat export records to LC in MARC 21 Communications format, UTF-8 encoding.

All LCCNs provided in metadata must be *normalized* to MARC21 standard. (See https://www.loc.gov/marc/lccn-namespace.html for more information on *LCCN normalization.*)

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 is used as a unique identifier. For page, the combination of LCCN, Issue Date, Edition Order, and Page 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. Some fields, as indicated in the Metadata Dictionary and XML templates, are ***optional*** and not used within the NDNP system to manage or provide access to data. Awardees should use their own discretion in determining whether capture of this data is useful for their own needs.
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, as well as authenticity (through the incorporation of a digital signature and checksums). The software is distributed as the NDNP Digital Viewer and Validator (DVV), and 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 https://openpreservation.org/technology/products/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, included in the NDNP DVV, "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 preservation and technical metadata sections of Issue and Reel METS XML objects. The NDNP Validation Library also calculates a checksum for each file that is written into the METS XML objects as a digital signature. This signature can be used to verify validity throughout the digital lifecycle.

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**.**

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,
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 C.

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 at https://www.loc.gov/ndnp/guidelines/.

Delivery of Digital Assets

Awardees will deliver all digital assets to LC 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. (See Appendix D – Batch, File and Directory Structure on Delivery Media.) Assets delivered to LC as prescribed in this directory structure are converted by LC to conformance with the "BagIt" specification, a hierarchical package format for transferring digital content (see https://confluence.ucop.edu/display/Curation/BagIt for background information).

A given delivery device should encompass a single batch, with authenticity verified on the delivery media. Awardees will name each batch conforming to NDNP batch naming specifications (see Appendix D). The precise directory structure and batch naming specification will be discussed at the post-award awardee meeting and include successive sub-directories based on LCCN, reel number, and issue date with edition sequence. (Note: initial sample batches follow a different naming scheme than regular batches, see Appendix D for details.) An XML Batch file should be created per the template in Appendix C.

Delivery of digital assets to LC should primarily be via tracked shipment of durable external hard drives without write-protection (USB enabled). The possibility of delivery via Internet2-enabled server-to-server file transfer will be discussed at the annual awardees' conference (resource planning should be based on use of durable external hard drives). Awardees should plan for adequate temporary storage locally (approx. 54 Mb per page – including TIFF, JP2, PDF, OCR, metadata) during the transfer and verification process at LC. Awardees should plan to deliver data batches to LC monthly (no more than 10,000 pages per month), with an expected response time of 6-8 weeks for LC data acceptance and ingestion.

Further options and specifications for delivery will be specified at the 2020-22 awardees' meeting, post-award.

Appendices

Note: Latest versions of these specifications (in use by 2019-21 NDNP Awardees) are available on the Guidelines and Resources page of the LC NDNP Web Site at https://www.loc.gov/ndnp/.

Appendix A: Digital Asset Metadata Elements - Dictionary

NOTES:
- Metadata elements below are described by original object. Elements may appear in more than one digital object per NDNP specifications. Mandatory (and mandatory if available) elements are marked as “M” (or “MA”) and appear in red.

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) and/or Data location
General Information						
Awardee Name	string	New York Public Library	Name of institution that received the NEH award	NR	**M**	mets:mets[@TYPE=“urn:library-of-congress:ndnp:mets:newspaper:issue”]/metsHdr/mets:agent/mets:name mets:mets[@TYPE=“urn:library-of-congress:ndnp:mets:microfilmReel]/metsHdr/mets:agent/mets:name
Award Year	enumeration	2020	Year of NEH award under which the digitization of this content was funded	NR	**M**	Xml:xml[@TYPE=“urn:library-of-congress:ndnp:batch”]/batchHdr/batch:agent/batch:awardYear
Original Source Repository	string	Multiple examples: Library of Congress; Washington, DC or New York Public Library; New York, NY	Owner of original source that was digitized (micro-film or paper) ; 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 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/@displayLabel
Original Source Repository Code	enumeration	dlc	Normalized MARC organization code of owner of	NR	**MA**	mets:mets[@TYPE=“urn:library-of-congress:ndnp:mets:newspaper:issue”]

			source. See https://www.loc.gov/marc/organizations/org-search.php for more information and code list.			/mets:dmdSec[@ID="pageModsBib1"]/mets:mdWrap/mets:xmlData/mods:mods/mods:relatedItem[@type="original"]/mods:location/mods:physicalLocation 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 institution; 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="agency ResponsibleForReproduction"]/@displayLabel mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/mets:dmdSec[@ID="techTargetModsBib"]/mets:mdWrap/mets:xmlData/mods:mods/mods:note/@displayLabel mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/mets:dmdSec[@ID="targetModsBib1"]/mets:mdWrap/mets:xmlData/mods:mods/mods:note/@displayLabel TIFF: Tag 315/institution name PDF: rdf:Description/dc:description/ rdf:Alt/rdf:li JPEG2000: rdf:Description/ dc:description/ rdf:Alt/rdf:li
Digital Responsible	enumeration	Multiple examples: dlc	Normalized MARC organization code of	NR	**MA**	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:newspaper:issue"]

Institution Code		oru vi	Awardee. See https://www.loc.gov/marc/organizations/org-search.php.			/mets:dmdSec[@ID="pageModsBib1"]/mets:mdWrap/mets:xmlData/mods:mods/mods:note[@type="agency ResponsibleForReproduction"] mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/mets:dmdSec[@ID="techTargetModsBib"]/mets:mdWrap/mets:xmlData/mods:mods/mods:note mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/mets:dmdSec[@ID="targetModsBib1"]/mets:mdWrap/mets:xmlData/mods:mods/mods:note Xml:xml[@TYPE="urn:library-of-congress:ndnp:batch"]/batchHdr/batch:agent/batch:awardee
Batch name *Sample* **Batch name** *Production*	string string	batch_dlc_2020sample batch_dlc_apple	For initial sample batch, use this naming structure: batch_[MARC organization code*]_[year of award]sample For production batches, use this naming structure: batch_[MARC organization code*]_[keyword**]. *For MARC Organization code, see https://www.loc.gov/marc/organizations/org-search.php. ** Batch keywords are unique within the deliveries of a given awardee throughout their program participation.	NR	**M**	Xml:xml[@TYPE="urn:library-of-congress:ndnp:batch"]/batchHdr/batch:agent/batch:name

			Keywords (in alpha order for a given grant period) are selected by awardees. NOTE: All batch names must be lower-case. (batch _[orgcode]_[keyword]. Only 2 underscores are permitted per batch name (e.g., batch_dlc_thomasjefferson)			
Scanning Contractor	string	Image Scanning Corporation OR Image Scanning Corporation; Image Subcontractors, Inc.	Contractor that scans to create TIFFs. Additional sub-contractors creating TIFFs may be represented separated by a semicolon and a single space. (Note: Scanning contractor element is combined with Digital Responsible Institution in TIFF Tag 315. If the Digital Responsible Institution ("institution name" element in TIFF tag 315) conducts its own scanning, the "scanning contractor" element can be omitted.)	NR	**MA**	TIFF: Tag 315/scanning contractor (See NDNP TIFF Profile for more information.)
Scanner Manufacturer	string	Scanner Manufacturing Company, Inc.	Manufacturer of scanner used to create TIFFs.	NR	**M**	TIFF: Tag 271
Scanner Model	string	Scanner0001, Model2, SN#12345	Model name, number and serial number of scanner. Use the following format: “model name, model number, SN#serial number” Note: Include model number (optional). Include serial number (required).	NR	**M**	TIFF: Tag 272 (See NDNP TIFF Profile for more information.)

Scanning Software	string	ImageDocSoftware v.10.1	Software used to create original TIFFs or software used to create NDNP deliverable TIFF images. Include version number/name.	NR	**M**	TIFF: Tag 305
Title Information						
LCCN	string	sn83031150 or 2007123234	Use canonical (normalized) form of LCCN for associated title bibliographic description. See https://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 Xml:xml[@TYPE="urn:library-of-congress:ndnp:batch"]/newspapertitle:lccn Xml:xml[@TYPE="urn:library-of-congress:ndnp:batch"]/issue:lccn PDF: rdf:Description/dc:description/ rdf:Alt/rdf:li JPEG2000: rdf:Description JPEG2000: rdf:Description/ dc:description/ rdf:Alt/rdf:li
Title	string	Multiple examples: The national forum. (Washington, D.C.) The Weekly roundabout. (Frankfort, Ky.) The Irish standard. (Minneapolis, Minn. ;	Combine MARC 245$a and (260$a) Or Combine MARC 245$a and (264$a) [for RDA records] Or	NR	**M**	PDF: rdf:Description/dc:title/ rdf:Alt/rdf:li PDF: rdf:Description/dc:description/ rdf:Alt/rdf:li JPEG2000: rdf:Description/ dc:title/ rdf:Alt/rdf:li JPEG2000: rdf:Description/ dc:description/ rdf:Alt/rdf:li

		St. Paul, Minn.) Svět = World. (Cleveland, Ohio)	MARC 130$a (Use of the MARC 130 field is permitted only if the field is present and well-formed.) (If multiple MARC 260$a fields are present, include all within parentheses and separate each with a space.) (If MARC 245$b field is present, optionally include after MARC 245$a separated with a space.)			mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:newspaper:issue"]@LABEL
Volume Number	string	27	Following SICI standard: (1.) All numeric information shall be converted to Arabic numerals. (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	Multiple examples: 1 2 3	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[@type="edition"]/mods:number Xml:xml[@TYPE="urn:library-of-congress:ndnp:batch"]/issue:editionOrder JPEG2000: rdf:Description
Edition Label	string	Multiple examples: Late City Final Two Stars (Final	If present, record as printed. (If not present, do not generate.)	NR	O	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:newspaper:issue"]/mets:dmdSec[@ID="issueModsBib"]/mets:mdWrap/mets:xmlData/mods:mo

		Edition) Extra	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)".)			ds/mods:relatedItem/mods:part/mods:detail[@type="edition"]/mods:caption mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:newspaper:issue"]@LABEL
Issue Number	string	3	Following SICI standard: (1.) All numeric information shall be converted to Arabic numerals. (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 Xml:xml[@TYPE="urn:library-of-congress:ndnp:batch"]/issue:issueDate PDF: rdf:Description/dc:title/ rdf:Alt/rdf:li PDF: rdf:Description/dc:description/ rdf:Alt/rdf:li JPEG2000: rdf:Description JPEG2000: rdf:Description/dc:title/ rdf:Alt/rdf:li JPEG2000: rdf:Description/dc:date/ rdf:Seq/rdf:li

						mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:newspaper:issue"]@LABEL
Issue Present Indicator	string	Not digitized, published	Valid values are: Present; Not digitized, published; Not digitized, not published; Not digitized, publishing unknown. (Note: "Present" means Published and digitized).	NR	**M**	mets:mets[@TYPE="urn:library-of-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.	R	**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, record as printed, to reflect logical Section navigation. (If not present, do not generate.) Could be blank, "C," "IV," "3," "Business," etc. If more than one Section heading per page, record dominant heading or combine headings (limit to approx. 6 words).	NR	O	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 pages 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 JPEG2000: rdf:Description
Page Number	string	B3	Exactly as printed. If not printed, element should be omitted. If not used, element should be omitted.	NR	O	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 PDF: rdf:Description/dc:title/ rdf:Alt/rdf:li JPEG2000: rdf:Description/dc:title/ rdf:Alt/rdf:li
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 TIFF: Tag 41728 (See NDNP TIFF profile for more information on acceptable values.)
Page Present Indicator	string	Present	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:note[@type="noteAboutReproduction"]

			Note: “Present” means Published and digitized.			
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”] 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”] Xml:xml[@TYPE=“urn:library-of-congress:ndnp:batch”]/reel:reelnumber TIFF: Tag 269 PDF: rdf:Description/dc:identifier/ rdf:Alt/rdf:li JPEG2000: rdf:Description/ dc:identifier/ rdf:Alt/rdf:li
Reel Sequence Number	positive integer	1	This orders the records **within a reel**. Indicate position of the image within the microfilm reel sequence.	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 sequence

						number”] 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”] TIFF: Tag 42016 PDF: rdf:Description/dc:identifier/ rdf:Alt/rdf:li JPEG2000: rdf:Description/ dc:identifier/ rdf:Alt/rdf:li
Object Format	string	Microfilm	Valid value (currently): Microfilm	NR	**M**	mets:mets[@TYPE=“urn:library-of-congress:ndnp:mets:microfilmReel”]/mets:dmdSec[@ID=“targetModsBib”]/mets:mdWrap/mets:xmlData/mods:mods/mods:relatedItem/mods:physicalDescription TIFF: Tag 41728 (See NDNP TIFF profile for more information on acceptable values.)
Tech Target Label	string	Preservation Microfilm Scanner Target PMT1	Valid value (currently): Preservation Microfilm Scanner Target PMT1	NR	**M**	mets:mets[@TYPE=“urn:library-of-congress:ndnp:mets:microfilmReel”]/mets:dmdSec[@ID=“techTargetModsBib”]/mets:mdWrap/mets:xmlData/mods:mods/mods:titleInfo/mods:title
Titles (on Reel)	string	Multiple examples: The national forum. (Washington, D.C.) The Weekly roundabout. (Frankfort, Ky.)	Combine MARC 245$a and (260$a) Or Combine MARC 245$a and (264$a) [for RDA records]	NR	O	mets:mets[@TYPE=“urn:library-of-congress:ndnp:mets:microfilmReel”]/mets:amdSec/techMD[@ID=”reelTechMD”]/mets:mdWrap/mets:xmlData/ndnp:reelTechMD/ndnp:titles

		The Irish standard. (Minneapolis, Minn. ; St. Paul, Minn.) Svět = World. (Cleveland, Ohio)	Or MARC 130$a. (Use of the MARC 130 field is permitted only if the field is present and well-formed.) (If multiple MARC 260$a fields are present, include all within parentheses and separate each with a space.) (If MARC 245$b field is present, optionally include after MARC 245$a separated with a space.) Use title that corresponds to the folder in which the reel XML file resides. Make note of multiple titles in "comments" element (below).			
Start Date	date	1881-11-22	Use this particular *ISO 8601* style: YYYY-MM-DD.	NR	O	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/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	O	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/mets:amdSec/techMD[@ID="reelTechMD"]/mets:mdWrap/mets:xmlData/ndnp:reelTechMD/ndnp:endDate
Position	string	2a	1a, 2a, 1b, 2b	NR	O	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/mets:amdSec/techMD[@ID="reelTechMD"]/mets:mdWrap/mets:xmlData/ndnp:reelTechMD/ndnp:position TIFF: Tag 274

Reduction Ratio	string	20x	If stated, transcribe. If not stated, estimate.	NR	O	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/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	O	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/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	O	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/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	O	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/mets:amdSec/techMD[@ID="reelTechMD"]/mets:mdWrap/mets:xmlData/ndnp:reelTechMD/ndnp:guideToContentsPresentFlag
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:microfilmReel"]/mets:amdSec/techMD[@ID="reelTechMD"]/mets:mdWrap/mets:xmlData/ndnp:reelTechMD/ndnp:guideToContentsString
Date Microfilm Created	string	1986	Date of microfilm creation	NR	O	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/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	O	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/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	O	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/mets:amdSec/techMD[@ID="reelTechMD"]/mets:mdWrap/mets:xmlData/ndnp:reelTechMD/ndnp:boundVolumeFlag
Comments	string		Any general comments regarding the reel, duplication process, or content, etc. If listing multiple titles on the reel, include associated LCCNs and date spans of each title.	NR	O	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/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, or other sources.	NR	O	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, or other sources. See https://www.loc.gov/rr/news/news_research_tools/ayersdirectory.html	NR	O	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	2	Refers to microfilming resolution targets.	NR	O	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	O	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/mets:amdSec/techMD[@ID="reelTechMD"]/mets:mdWrap/mets:xmlData/ndnp:reelTechMD/ndnp:resolutionOfMaster

Resolution Of Master: Comments	string		Any comments relating to the resolution of the camera master microfilm	NR	O	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/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	O	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/mets:amdSec/techMD[@ID="reelTechMD"]/mets:mdWrap/mets:xmlData/ndnp:reelTechMD/ndnp:densityReadingMaster
Average Density Master	floating point number	0.95	The average density of the camera master microfilm	NR	O	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/mets:amdSec/techMD[@ID="reelTechMD"]/mets:mdWrap/mets:xmlData/ndnp:reelTechMD/ndnp:averageDensityMaster
Dmin Master	floating point number	0.2	The minimum density point on the camera master microfilm	NR	O	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/mets:amdSec/techMD[@ID="reelTechMD"]/mets:mdWrap/mets:xmlData/ndnp:reelTechMD/ndnp:dminMaster
Resolution Of Duplicate Negative	floating point number	6.3	Resolution of the duplicate negative used for digitization. Record actual number from line-pair resolution test pattern on film, if available.	NR	O	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/mets:amdSec/techMD[@ID="reelTechMD"]/mets:mdWrap/mets:xmlData/ndnp:reelTechMD/ndnp:resolutionOfDuplicateNegative
Resolution Of Duplicate Negative: Comments	string	Weak 7.1	Any comments relating to the resolution of the duplicate negative used for digitization	NR	O	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/mets:amdSec/techMD[@ID="reelTechMD"]/mets:mdWrap/mets:xmlData/ndnp:reelTechMD/ndnp:resolutionCommentDuplicateNegative
Density Reading Duplicate Negative	floating point number	1.14	Ten readings per reel.	R	O	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/mets:amdSec/techMD[@ID="reelTechMD"]/mets:mdWrap/mets:xmlData/ndnp:reelTechMD/ndnp:densityReadingDupli

						cateNegative
Average Density Duplicate Negative	floating point number	1.1	The average density of duplicate master negative used for digitization	NR	O	mets:mets[@TYPE="urn:library-of-congress:ndnp:mets:microfilmReel"]/mets:amdSec/techMD[@ID="reelTechMD"]/mets:mdWrap/mets:xmlData/ndnp:reelTechMD/ndnp:averageDensityDuplicateNegative
Dmin Duplicate Negative	floating point number	0.15	The minimum density point on the duplicate negative used for digitization	NR	O	mets:mets[@TYPE="urn:library-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 are available on the LC NDNP Web Site at https://www.loc.gov/ndnp/guidelines/.

NDNP TIFF Profile

Version 1.9

CHANGES in 1.9:
1. Additional information added to tags 305 and 315.

CHANGES in 1.8:
1. Additional acceptable values added to tag 41728.

CHANGES in 1.7:
1. Changes to resolution limits.

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

NDNP TIFF Profile (Version 1.9)

1. The TIFF will conform to 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
TIFF Type:	ASCII

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

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” or “1” or “2” or “3”
Notes: Contact NDNP if none of the permitted values are appropriate. Permitted values reflect updated values (“1”, “2”, or “3”) and existing NDNP practice values (“microfilm” or “microfiche” or “print”) . [1 = Film Scanner; 2 = Reflection Print Scanner; 3 = Digital Camera]

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 “institution name” is the Digital Responsible Institution. The semicolon and scanning contractor are omitted if not applicable (if Digital Responsible Institution conducted the scanning). Additional institution or contractor/sub-contractor names may be added. Separate each name with a semicolon and a single space.

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 number/name.

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 2.9

CHANGES IN 2.9:

1. Updated punctuation in rdf:Description.

CHANGES IN 2.8:

1. Updated rdf:Description references in XML Box metadata template to match NDNP element names in Digital Asset Metadata Elements – Dictionary.

CHANGES IN 2.7:

1. Removed code-block style requirement.

NDNP JPEG 2000 Profile (Version 2.9)

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:
    1,0.84,0.7,0.6,0.5,0.4,0.35,0.3,0.25,0.21,0.18,0.15,0.125,0.1,0.088,0.07,0.0625,0.05, 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://#LCCN#/#Issue
Date#/#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">#Title#, #Issue Date#, [p #Page Number#].</rdf:li>
                        </rdf:Alt>
                </dc:title>
                <dc:description>
                        <rdf:Alt>
                        <rdf:li xml:lang="en">Page from #Title# (newspaper). [See LCCN: #LCCN# for
                catalog record.]. Prepared on behalf of #Digital Responsible Institution#.</rdf:li>
                        </rdf:Alt>
                </dc:description>
                <dc:date>
                        <rdf:Seq>
                        <rdf:li xml:lang="x-default">#Issue Date#</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 #Reel Number#. Sequence number #Reel
                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-
28/1/2'
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
                2].</rdf:li>
                        </rdf:Alt>
                </dc:title>
                <dc:description>
                        <rdf:Alt>
                        <rdf:li xml:lang="en">Page from The national forum (newspaper). [See LCCN:
                sn82015056 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 #Digital Responsible Institution#.</rdf:li>
                        </rdf:Alt>
                </dc:description>
                <dc:identifier>
                        <rdf:Alt>
                        <rdf:li xml:lang="en">Reel number #Reel number#. Sequence number #Reel
                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
1,0.84,0.7,0.6,0.5,0.4,0.35,0.3,0.25,0.21,0.18,0.15,0.125,0.1,0.088,0.075,0.0625,0.05,0.04419,0.03716,0.03
125,0.025,0.0221,0.01858,0.015625 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 2.4

CHANGES IN 2.4:

1. Updated font requirements.

CHANGES IN 2.3:

1. Updated punctuation in rdf:Description.

CHANGES IN 2.2:

2. Updated rdf:Description references in XMP metadata template to match NDNP element names in Digital Asset Metadata Elements – Dictionary.

CHANGES IN 2.1:

1. Removed specific PDF/A requirements.
2. Added general PDF/A recommendation.

NDNP PDF Profile (Version 2.4)

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. It is recommended that only the 14 standard Type1 fonts be used. These fonts will not be embedded. If it is necessary to use a non-standard font, the appropriate subset of this font should 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">#Title#, #Issue Date#, [p #Page Number#].</rdf:li>
		</rdf:Alt>
	</dc:title>
	<dc:description>
		<rdf:Alt>
		<rdf:li xml:lang="en">Page from #Title# (newspaper). [See LCCN: #LCCN# for catalog
	record.]. Prepared on behalf of #Digital Responsible Institution#.</rdf:li>
		</rdf:Alt>
	</dc:description>
	<dc:date>
		<rdf:Seq>
		<rdf:li xml:lang="x-default">#Issue Date#</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 #Reel Number#. Sequence number #Reel 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 2].</rdf:li>
		</rdf:Alt>
	</dc:title>
	<dc:description>
		<rdf:Alt>
		<rdf:li xml:lang="en">Page from The national forum (newspaper). [See LCCN:
	sn82015056 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
        #Digital Responsible Institution#.</rdf:li>
                </rdf:Alt>
        </dc:description>
        <dc:identifier>
                <rdf:Alt>
                <rdf:li xml:lang="en">Reel number #Reel number#. Sequence number #Reel Sequence
        Number#.</rdf:li>
                </rdf:Alt>
        </dc:identifier>
</rdf:Description>
```

NDNP OCR Profile

Version 1.18

CHANGES in 1.18:
1. Expanded language code specification source to include *ISO 639-3* codes when *ISO 639-2* code is not available.

CHANGES in 1.17:
1. Added clarification for languages that have both "b" (bibliographic) and "t" (terminology) *ISO 639-2* language codes.

CHANGES in 1.16:
1. Expanded language code specification information to include any language with a valid *ISO 639-2* alpha-3 language code.

CHANGES in 1.15:
1. Added language code specification information for Danish, Hungarian, Norwegian, Portuguese, and Swedish text.

CHANGES in 1.14:
1. Added language code specification information for German text.
2. Added clarification about non-English text language codes.

CHANGES in 1.13:
1. Changed version of ALTO to 2.0.
2. Added language code specification for non-English text and acceptable language codes.

CHANGES in 1.12:
1. Required use of HEIGHT and WIDTH for Page element.

CHANGES IN 1.11:
1. Changed version of ALTO to 1-4.

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.

NDNP OCR Profile (Version 1.18)

1. OCR text will be encoded using the ALTO (Analyzed Layout and Text Object) schema, Version 2.0, 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 Page element, the use of PRINTED_IMG_NR, QUALITY, POSITION, and PROCESSING are encouraged.
8. For the Page element, the use of HEIGHT and WIDTH are required.
9. For the Page element, the entire page may be included in the PrintSpace. (Thus, use of TopMargin, LeftMargin, RightMargin, and BottomMargin are not required.)
10. The use of Illustration, GraphicalElement, and ComposedBlock are not required.
11. The use of non-rectangular blocks is not encouraged.
12. The use of SP and HYP are encouraged.
13. For a TextLine, the use of BASELINE is discouraged.
14. For a String, the use of ALTERNATIVE, WC, and CC is encouraged if available.
15. For a String, the use of HEIGHT, WIDTH, HPOS, and VPOS is required.
16. 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.)
17. 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.)
18. Non-English text must be encoded at the TEXTBLOCK, using *ISO 639-2* alpha-3 language codes (https://www.loc.gov/standards/iso639-2/). For languages that have both "b" (bibliographic) and "t" (terminology) codes, use the "b" (bibliographic) code. For languages that are not represented in *ISO 639-2* specification, the *ISO 639-3* (https://iso639-3.sil.org) code may be used if available. Language search support will vary according to the tools/technologies in use by the Library of Congress' web sites. Fraktur/black letter fonts must incorporate OCR technical processing that includes Fraktur/black letter specific tools.

Note: a single ALTO document may have multiple languages encoded within individual TEXTBLOCKs (e.g. bilingual newspaper pages), but a single TEXTBLOCK may only have a single language.

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"/>

5. Because a single reel or batch may contain variations in language and OCR quality, automated language recognition (if used) should be applied to individual titles, rather than across entire reels or batches.

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="[Title], [issueDate][,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
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:SignatureValue>
<dsig:KeyInfo>
<dsig:KeyValue>
<dsig:DSAKeyValue>

<dsig:P>/X9TgR11EilS30qcLuzk5/YRt1I870QAwx4/gLZRJmlFXUAiUftZPY1Y+r/F9bow9subVWzXgTuA
HTRv8mZgt2uZUKWkn5/oBHsQIsJPu6nX/rfGG/g7V+fGqKYVDwT7g/bTxR7DAjVUE1oWkTL2dfOu
K2HXKu/yIgMZndFIAcc=</dsig:P><dsig:Q>l2BQjxUjC8yykrmCouuEC/BYHPU=</dsig:Q><dsig:G>9+Gghda
bPd7LvKtcNrhXuXmUr7v6OuqC+VdMCz0HgmdRWVeOutRZT+ZxBxCBgLRJFnEj6EwoFhO3
zwkyjMim4TwWeotUfI0o4KOuHiuzpnWRbqN/C/ohNWLx+2J6ASQ7zKTxvqhRkImog9/hWuWfBpKL
Zl6Ae1UlZAFMO/7PSSo=</dsig:G><dsig:Y>IQy0sz7MUZsAcnddcGlz8Nc5Cx68xbVNCI8DjOrahDarTyvFm
OymnYNJ+TkmkGJZxGEPt6lGFaRX
mFKwPaq9P6SHOZc567l2XZMIdIIeMrlZ2xOJjzC5H4EvOLxEQgQP3p/9cywaA/fNkRdSfF9MeLwL
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. 1.7

```
<?xml version="1.0" encoding="UTF-8"?>
<!--NDNP Reel profile-->
<!--Justin Littman, OSI-->
<!--Version 1.7-->
<!--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 1.7:
1.  Changed titles, startDate, endDate, position, reductionRatio, dateMicrofilmCreated, captureResolutionOriginal, captureResolutionOriginal, guideToContentsPresentFlag, looseLeavesFlag, boundVolumeFlag, dimensions, pagesPerIssue, numberOfResolutionTargets, resolutionOfMaster, resolutionCommentMaster, densityReadingMaster, averageDensityMaster, dminMaster, resolutionOfDuplicateNegative, resolutionCommentDuplicateNegative, densityReadingDuplicateNegative, averageDensityDuplicateNegative, and dminDuplicateNegative to optional.
2.  Changed reelTechMD to optional.
-->

<!-- CHANGES in 1.6:
1.      Changed the number of required tech targets to between 1 and 5.
 -->

<!--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> Preservation Microfilm Scanner Target PMT1</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 (optional)-->
      <techMD ID="reelTechMD">
         <mdWrap MDTYPE="OTHER" OTHERMDTYPE="NDNP Reel Technical Metadata" LABEL="NDNP
technical metadata for microfilm reel">
            <xmlData>
               <ndnp:reelTechMD>
                  <ndnp:titles>[Titles (optional)]</ndnp:titles>
                  <ndnp:startDate>1919-01-01</ndnp:startDate><!--Start Date (optional)-->
                  <ndnp:endDate>1920-01-01</ndnp:endDate><!--End Date (optional)-->
                  <ndnp:position>1a</ndnp:position><!--Position (optional)-->
                  <ndnp:reductionRatio>[Reduction Ratio (optional)]</ndnp:reductionRatio>
```

<ndnp:captureResolutionOriginal measurement="pixels/inch">300</ndnp:captureResolutionOriginal><!--Capture resolution of original (optional)-->
<ndnp:captureResolutionFilm measurement="pixels/inch">6000</ndnp:captureResolutionFilm><!--Capture resolution of film (optional)-->
<ndnp:guideToContentsPresentFlag>true</ndnp:guideToContentsPresentFlag><!--Guide to contents present (optional)-->
<ndnp:guideToContentsString>[Guide to Contents String (optional)]</ndnp:guideToContentsString>
<ndnp:dateMicrofilmCreated>[Date microfilm created (optional)]</ndnp:dateMicrofilmCreated>
<ndnp:looseLeavesFlag>true</ndnp:looseLeavesFlag><!--Loose leaves (optional)-->
<ndnp:boundVolumeFlag>true</ndnp:boundVolumeFlag><!--Loose leaves (optional)-->
<ndnp:comments>[Comments (optional)]</ndnp:comments>
<ndnp:dimensions>[Dimensions (optional)]</ndnp:dimensions>
<ndnp:pagesPerIssue>4</ndnp:pagesPerIssue><!--Pages per issue (optional)-->
<ndnp:numberOfResolutionTargets>5</ndnp:numberOfResolutionTargets><!--Number of resolution targets (optional)-->
<ndnp:resolutionOfMaster>7.1</ndnp:resolutionOfMaster><!--Resolution of master (optional)-->
<ndnp:resolutionCommentMaster>[Resolution comment master (optional)]</ndnp:resolutionCommentMaster>
<ndnp:densityReadingMaster>0.91</ndnp:densityReadingMaster><!--Density reading 1 master (optional)-->
<ndnp:densityReadingMaster>0.91</ndnp:densityReadingMaster><!--Density reading 2 master (optional)-->
<ndnp:densityReadingMaster>0.91</ndnp:densityReadingMaster><!--Density reading 3 master (optional)-->
<ndnp:densityReadingMaster>0.91</ndnp:densityReadingMaster><!--Density reading 4 master (optional)-->
<ndnp:densityReadingMaster>0.91</ndnp:densityReadingMaster><!--Density reading 5 master (optional)-->
<ndnp:densityReadingMaster>0.91</ndnp:densityReadingMaster><!--Density reading 6 master (optional)-->
<ndnp:densityReadingMaster>0.91</ndnp:densityReadingMaster><!--Density reading 7 master (optional)-->
<ndnp:densityReadingMaster>0.91</ndnp:densityReadingMaster><!--Density reading 8 master (optional)-->
<ndnp:densityReadingMaster>0.91</ndnp:densityReadingMaster><!--Density reading 9 master (optional)-->
<ndnp:densityReadingMaster>0.91</ndnp:densityReadingMaster><!--Density reading 10 master (optional)-->
<ndnp:averageDensityMaster>0.91</ndnp:averageDensityMaster><!--Average density master (optional)-->
<ndnp:dminMaster>0.20</ndnp:dminMaster><!--Dmin master (optional)-->
<ndnp:resolutionOfDuplicateNegative>7.1</ndnp:resolutionOfDuplicateNegative><!--Resolution of DuplicateNegative (optional)-->
<ndnp:resolutionCommentDuplicateNegative>[Resolution comment DuplicateNegative (optional)]</ndnp:resolutionCommentDuplicateNegative>
<ndnp:densityReadingDuplicateNegative>0.91</ndnp:densityReadingDuplicateNegative><!--Density reading 1 DuplicateNegative (optional)-->
<ndnp:densityReadingDuplicateNegative>0.91</ndnp:densityReadingDuplicateNegative><!--Density reading 2 DuplicateNegative (optional)-->
<ndnp:densityReadingDuplicateNegative>0.91</ndnp:densityReadingDuplicateNegative><!--Density reading 3 DuplicateNegative (optional)-->
<ndnp:densityReadingDuplicateNegative>0.91</ndnp:densityReadingDuplicateNegative><!--Density reading 4 DuplicateNegative (optional)-->
<ndnp:densityReadingDuplicateNegative>0.91</ndnp:densityReadingDuplicateNegative><!--Density reading 5 DuplicateNegative (optional)-->
<ndnp:densityReadingDuplicateNegative>0.91</ndnp:densityReadingDuplicateNegative><!--Density reading 6 DuplicateNegative (optional)-->

```
                <ndnp:densityReadingDuplicateNegative>0.91</ndnp:densityReadingDuplicateNegative><!--
Density reading 7 DuplicateNegative (optional)-->
                <ndnp:densityReadingDuplicateNegative>0.91</ndnp:densityReadingDuplicateNegative><!--
Density reading 8 DuplicateNegative (optional)-->
                <ndnp:densityReadingDuplicateNegative>0.91</ndnp:densityReadingDuplicateNegative><!--
Density reading 9 DuplicateNegative (optional)-->
                <ndnp:densityReadingDuplicateNegative>0.91</ndnp:densityReadingDuplicateNegative><!--
Density reading 10 DuplicateNegative (optional)-->
                <ndnp:averageDensityDuplicateNegative>0.91</ndnp:averageDensityDuplicateNegative><!--
Average density DuplicateNegative (optional)-->
                <ndnp:dminDuplicateNegative>0.20</ndnp:dminDuplicateNegative><!--Dmin
DuplicateNegative (optional)-->
            </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>
                  <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:SignatureValue>
                  <dsig:KeyInfo>
                     <dsig:KeyValue>
                        <dsig:DSAKeyValue>

<dsig:P>/X9TgR11EilS30qcLuzk5/YRt1I870QAwx4/gLZRJmlFXUAiUftZPY1Y+r/F9bow9subVWzXgTuA
HTRv8mZgt2uZUKWkn5/oBHsQIsJPu6nX/rfGG/g7V+fGqKYVDwT7g/bTxR7DAjVUE1oWkTL2dfOu
K2HXKu/yIgMZndFIAcc=</dsig:P><dsig:Q>I2BQjxUjC8yykrmCouuEC/BYHPU=</dsig:Q><dsig:G>9+Gghda
bPd7LvKtcNrhXuXmUr7v6OuqC+VdMCz0HgmdRWVeOutRZT+ZxBxCBgLRJFnEj6EwoFhO3
zwkyjMim4TwWeotUfI0o4KOuHiuzpnWRbqN/C/ohNWLx+2J6ASQ7zKTxvqhRkImog9/hWuWfBpKL
Zl6Ae1UlZAFMO/7PSSo=</dsig:G><dsig:Y>IQy0sz7MUZsAcnddcGlz8Nc5Cx68xbVNCI8DjOrahDarTyvFm
OymnYNJ+TkmkGJZxGEPt6lGFaRX
mFKwPaq9P6SHOZc567l2XZMIdIIeMrlZ2xOJjzC5H4EvOLxEQgQP3p/9cywaA/fNkRdSfF9MeLwL
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 between 1 and 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. 1.6

```
<?xml version="1.0" encoding="UTF-8"?>
<!--NDNP Batch profile-->
<!--Version 1.6-->
<!--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 any element which takes a value if there is no value.  For example, omit the encyclopediaEntry
element if encyclopediaEntry is not being provided.
   4. The sequence of elements is required.
-->
<!—

CHANGES IN 1.6:
   1. Added clarification to issue and reel elements.
 CHANGES IN 1.5:
   1. Omitted references to xsi:schemaLocation.
   2. Added clarification about encyclopediaEntry and newspaperTitle elements.
   CHANGES IN 1.4:
   1. Added awardee and awardYear attribute.
   CHANGES IN 1.3:
   1. Added additional clarification about xsi:schemaLocation.
   2. Added additional clarification about sequence.
   3. Added encyclopedia element.
   CHANGES IN 1.2:
   1. Added name attribute.
-->


<batch name="batch_dlc_eagle" awardee="dlc" awardYear="2008" xmlns="http://www.loc.gov/ndnp"><!--
NAME is the name of the batch.  AWARDEE is the awardee MARC Org Code (from
http://www.loc.gov/marc/organizations).  AWARD YEAR is the NEH award phase for which the content was
created, given as a 4 digit year.-->

   <!--ENCYCLOPEDIA-->
   <encyclopediaEntry><!--A relative path of the Mets file.  For example:--
>./encyclopedia/encyclopedia.xml</encyclopediaEntry>
   <!--Repeat encyclopediaEntry for each encyclopedia entry-->
   <!--encyclopediaEntry element used internally by LC. Not required for awardees. -->

   <!--NEWSPAPER TITLE-->
   <newspaperTitle lccn="sn78890351"><!--A relative path of the Mets file.--
>./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:
      <newspaperTitle lccn="sn78890351" />
   -->
   <!--newspaperTitle element used internally by LC. Not required for awardees. -->

   <!--ISSUE-->
   <issue lccn="sn78890351" issueDate="2004-12-03" editionOrder="1"><!--A relative path of the Mets file.--
>./1/1/mets.xml<!--lccn should be a normalized LCCN.--></issue>
   <issue lccn="sn78890351" issueDate="2004-12-03" editionOrder="1">./1/2/mets.xml</issue>
   <!--Repeat issue element for each issue-->

   <!--REEL-->
   <reel reelNumber="[reelNumber]"><!--A relative path of the Mets file. -->./mets.xml</reel>
   <!--Repeat reel element for each reel-->
</batch>
```

Appendix D: Batch, File and Directory Structure on Delivery Media

Scanning from Microfilm

The basic file and directory structure for batches containing content scanned from microfilm is as follows. See below for details related to batches containing content scanned from original newsprint.

See https://www.loc.gov/ndnp/guidelines/examples.html for a link to an example batch scanned from microfilm.

batch_dlc_apple/
- batch.xml
- sn83045433/
  - 0010049324a/
    - 0010049324a.xml
    - 0001.tif
    - 0001.jp2
    - 0001.pdf
    - 0001.xml
  - 1905112901/
    - 1905112901.xml
    - 0004.tif
    - 0004.jp2
    - 0004.pdf
    - 0004.xml

Each batch will have a unique name constructed as “batch” underscore “Awardee MARC Org code” underscore “keyword.” All characters used in batch names shall be lower-cased. All batch names shall include and be limited to 2 underscores. Each keyword will be unique to a given Awardee throughout their participation in the program. Keywords should be alphabetic within a given award.

Example:
batch_dlc_apple

NOTE: For sample batches, associated keyword must be “[AwardYear]sample”. All other batch details are the same.

Example:
batch_dlc_2020sample

In the root directory of the portable hard drive will be the single batch file pointing to each Issue/Edition and reel record on the delivery media.

Example:
batch.xml

Each title will be contained in its own directory, name matching the title’s LCCN.

Example:
/sn83045433/

Within each title directory, subdirectories will be created for each scanned reel, with names matching the barcode on each reel.

Example:
/sn83045433/0010049324a/

Within each reel subdirectory will be one metadata file for the reel, named after its barcode number, the target images for that reel, the images of the currently used standards-based scanning resolution technical targets, and a subdirectory for each newspaper issue/edition, named after its date and edition order (typically 01).

Example reel metadata file:
/sn83045433/0010049324a/0010049324a.xml

Example subdirectory for the only edition from January 24, 1905:
/sn83045433/0010049324a/1905012401/

The metadata for that edition would be:
/sn83045433/0010049324a/1905012401/1905012401.xml

In each issue/edition subdirectory will be the metadata file for that issue, and all TIFF, JP2, PDF and OCR files for the pages in that issue/edition.

If the issue/edition is noted as missing from the guide to contents or other filmed targets, the proper issue/edition subdirectory can be created and contain the metadata file, but **no** TIFF, JP2, PDF or OCR files.

Example for "issue missing November 29, 1905":
/sn83045433/0010049324a/1905112901/1905112901.xml

Image and OCR files shall be named in four digit, one-up manner, according to the order of appearance on the scanned reel, whether the image is a target or page.

For example, a reel might contain, in order, a technical target (for which the scanned TIFF would be named):
/sn83045433/0010049324a/0001.tif
a title target:
/sn83045433/0010049324a/0002.tif
a guide to contents:
/sn83045433/0010049324a/0003.tif
and newspaper pages:
/sn83045433/0010049324a/1905012401/0004.tif
/sn83045433/0010049324a/1905012401/0005.tif
/sn83045433/0010049324a/1905012401/0006.tif
etc.

Each JP2, PDF and OCR file will have the same prefix as the TIFF file from which it was derived, and will be located in the same directory as its matching TIFF.

Example JP2, PDF and OCR files:
/sn83045433/0010049324a/1905012401/0004.jp2
/sn83045433/0010049324a/1905012401/0004.pdf
/sn83045433/0010049324a/1905012401/0004.xml

Scanning from Original Newsprint

For batches containing content derived from newsprint, rather than microfilm negative, substitute a directory called “print” for the reel directory. No targets will be scanned in that directory and no reel level METS XML files will be created. Image and OCR files shall be named in four digit, one-up manner, according to the order of appearance in bound volume, loose issue, or other container. All other file structures remain the same.

Note: Metadata value options (in issue level METS XML and image headers) necessary for describing content derived from print (rather than microfilm) are detailed in the appropriate sections of Appendices A, B, and C. More information will be made available at the post-award awardees’ meeting.

See https://www.loc.gov/ndnp/guidelines/examples.html for a link to an example batch scanned from newsprint.

Example:

```
batch_dlc_apple/
        batch.xml
        sn83045433/
                print/
                        1905112901/
                                1905112901.xml
                                0001.tif
                                0001.jp2
                                0001.pdf
                                0001.xml
```

Appendix E: Associated Online Resources

Program Information

- National Digital Newspaper Program (National Endowment for the Humanities) – https://www.neh.gov/divisions/preservation/national-digital-newspaper-program
- National Digital Newspaper Program (Library of Congress) - https://www.loc.gov/ndnp/
- Chronicling America: Historic American Newspapers - https://chroniclingamerica.loc.gov/

Award Application Information

- National Digital Newspaper Program Application Guidelines (NEH) – https://www.neh.gov/grants/preservation/national-digital-newspaper-program
- National Digital Newspaper Program Technical Guidelines (LC)– https://www.loc.gov/ndnp/guidelines/

Technical Background

- NDNP Guidelines and Resources (LC) - https://www.loc.gov/ndnp/guidelines/
- JSTOR/Harvard Object Validation Environment - https://openpreservation.org/technology/products/jhove/
- Justin Littman, "A Technical Approach and Distributed Model for Validation of Digital Objects." *D-Lib Magazine*, May 2006. https://www.dlib.org/dlib/may06/littman/05littman.html
- BagIt Specification - https://confluence.ucop.edu/display/Curation/BagIt
- Meta|morphosis Tutorials (University of Kentucky), film-to-digital resources - https://www.uky.edu/Libraries/NDNP/metamorphosis/index.html

Technical Information - Metadata

- MARC Standard - https://www.loc.gov/marc/
- METS (Metadata Encoding and Transmission Standard) - https://www.loc.gov/standards/mets/
- CONSER (Cooperative Online Serials) Program - https://www.loc.gov/acq/conser/
- LCCN Namespace - https://www.loc.gov/marc/lccn-namespace.html
- MARC Organizations - https://www.loc.gov/marc/organizations/org-search.php
- ISO 639-2: Codes for the representation of names of languages- Part 2: Alpha-3 codes - https://www.loc.gov/standards/iso639-2/php/code_list.php
- ISO 639-3: Codes for the representation of names of languages- Part 3: Alpha-3 codes for comprehensive coverage of languages - https://iso639-3.sil.org/

Technical Background - Imaging

- Federal Agencies Digitization Guidelines Initiative – http://www.digitizationguidelines.gov/
- Image Science Associates - http://imagescienceassociates.com
- NDNP Historical Newspaper JPEG2000 Profile - https://www.loc.gov/ndnp/guidelines/docs/NDNP_JP2HistNewsProfile.pdf

Additional Background

- United States Newspaper Program (NEH) – https://www.neh.gov/us-newspaper-program
- United States Newspaper Program Technical Guidelines (LC)- https://www.loc.gov/rr/news/usnp/usnpguidelinesp.html
- Newspaper and Current Periodicals Reading Room – Digital Reference Collections
    - Rowell's American Newspaper Directory (1869-1877)
    - Ayer's American Newspaper Annual (1880-1909)
    - Ayer's American Newspaper Annual & Directory (1910-1922)
    - Rowell's and Ayer's All-In-One Search (University of North Texas)
- United States Copyright
    - United States Copyright Office
    - Circular 22 - How to Investigate the Copyright Status of a Work