FDsys Operational Specification for Converted Content (Version 3.3)

(Feb. 2006)

Digitization Specifications and Operating Procedures for Archiving Materials: Creation of Preservation Master Files

For the following content types – Textual, Graphic Illustrations / Artwork, Originals, and Photographs

Specifications and metrics for Converted Content – a functional solution of the Future Digital System (FDsys)

United States Government Printing Office (GPO)

Final
Feb 10, 2006

FINAL

Document Change Control Sheet

Document Title: FDsys Operational Specification for Converted Content

Date	Filename/version #	Author	Revision Description
2/2/2005	*DigitizationSpecs-v.1.doc*	N. Doyle / R. Selvey	First Draft
2/18/2005	*DigitizationSpecs-v1.1.doc*	N. Doyle / R. Selvey	Additions, corrections and input from outside sources (LOC, etc.)
2/18/2005	*DigitizationSpecs-v2.0.doc*	N. Doyle / R. Selvey	Additions, corrections, visuals
3/3/2005	*DigitizationSpecs-v2.1.doc*	N. Doyle / R. Selvey	Revisions, narrowed down Standards list
4/12/2005	*DigitizationSpecs-v2.2.doc*	N. Doyle / R. Selvey	Revisions based on workflow
5/10/2005	*DigitizationSpecs-v2.3.doc*	N. Doyle / R. Selvey	All targets / standards have been established and updated.
5/31/2005	*DigitizationSpecs-v2.4.doc*	N.Doyle / R. Selvey	Sect III.C – Aimpoints have been revised / updated
6/1/2005	*DigitizationSpecs-v2.5.doc*	N.Doyle / R. Selvey	Update submission level Metadata
6/24/2005	*DigitizationSpecs-v3.0.doc*	T. Priebe	Formatted into FDsys template
09/26/05	*DigitizationSpecs-v3.1.doc*	N Doyle	Updates based on digi. suggestions
12/01/05	*DigitizationSpecs-v3.2.doc*	N.Doyle	Changed compression scheme for bitonal to CCITT Group 4
02/10/06	*FDsys Operational Spec for Converted Content - v3.3.doc*	R. Selvey	Change CCP to SIP, update references, crosswalk to other specs, update after green team review, update qc, update ID, workflow, current situation, batch processing, update table of contents
02/16/06	*FDsys Operational Spec for Converted Content - v3.3.doc*	R. Selvey	Updates after PMO review

Table of Contents

FDsys Operational Specification for Converted Content (Version 3.3) 1
Specifications and metrics for Converted Content – a functional solution of the Future Digital System (FDsys) 1
United States Government Printing Office (GPO) 1
DigitizationSpecs-v.1.doc 2
1. Scope 4
1.1 Deliverable 4
1.2 Overview 4
2. Referenced Documents 5
2.1 GPO 5
2.2 Organizational/Standard 5
2.4 Organizational/Standard 5
3. Current Situation 5
3.1 Background and objectives 5
3.2 Conversion 6
3.2.1 Scanning 6
3.2.1.1 Operational situation 6
3.2.1.2 Metrics 7
3.2.2 In3.2.2 Inspection 10
Document Inspection prior to scanning 10
3.2.2.1 Current operational situation 10
Document Characterization 10
Image Capture Classification 11
3.3 Content Management 12
3.3.1 Image Workflow 12
3.4 Storage 12
4. Desired Situation 12
4.1 Background Changes 12
4.1.1 Specific Component 12
4.1.1.1 Objectives 13
4.1.1.2 Metrics 13
4.1.1.3 Priorities among changes 13
5. Benchmarks 13
6. Constraints 16

1. Scope

What is addressed in this document:

- Scanning and format requirements for text, photographs, and graphic materials
- Digitization Environment
- Digitization Standards
- Required hardware/software configurations
- Quality control

Types of scanning projects will include the following:

- Brittle books (serials and monographs)
- Pamphlets and unbound material
- Archival materials
- Bound materials
- Fold-outs, maps, posters, etc.
- Microform (includes microfilm, microfiche, and aperture cards)

This specification does not describe how to create a Submission Information Package (SIP). The SIP will be covered in a separate content package specification.

1.1 Deliverable

The end product of the Conversion Process will be a GPO standard Submission Information Package (SIP).

1.2 Overview

This specification covers all the necessary conversion elements that are required for the creation of a SIP. The components of the conversion solution have been grouped into the following: 1) Conversion Processes; 2) Content Management; 3) Storage.

Converted content is one type of digital content that will be ingested by the Future Digital System. Converted content consists of electronic files created from tangible paper documents, which can be preserved as master files with associated metadata. GPO staff and external service providers “including contractors, library partners, and federal agencies” will provide converted content to the Future Digital System. The end product of conversion is a Submission Information Package (SIP). The SIP must be produced at a level of quality that is adequate to support preservation as well as future iterations of derivative products.

This document is an outline of our scanning specifications and will continue to evolve and improve as technological advancements occur in the digital imaging industry.

2. Referenced Documents

2.1 GPO

- *A Strategic Vision for the 21st Century, December 1, 2004*
- *Concept of Operations for the Future Digital System (ConOps) V2.0*
- *Requirements Document for the Future Digital System (RD) V1.0*
- *FDsys Unique ID specification*
- *FDsys SIP specification*
- *Report from the Meeting of Experts on Digital Preservation*, March 12, 2004

2.2 Organizational/Standard

Colorado Digitization Project - *General Guidelines for Scanning*, CDP Scanning Working Group, Spring 1999. http://www.cdpheritage.org

Digital Library Federation's *Benchmark for Faithful Reproductions of Monographs and Serials* (Version 1, December 2002), http://www.diglib.org/standards/bmarkfin.pdf

Frey, Franziska S., and James M. Reilly. *Digital Imaging for Photographic Collections Foundations for Technical Standards*. Rochester, NY: Image Permanence Institute, Rochester Institute of Technology, 1999.
http://www.rit.edu/~661www1/sub_pages/digibook.pdf.

The Institute for Museum and Library Services' (IMLS) *Framework of Guidance for Building Good Digital Collections* (2001), http://www.niso.org/framework/framework2.pdf

Western States Digital Standards Group: *Digital Imaging Working Group - Digital Imaging Best Practices*, Jan 2003.

2.3 Agency

Puglia, Steven, Reed, Jeffrey, and Rhodes, Erin. *Technical Guidelines for Digitizing Archival Materials for Electronic Access: Creation of Production Master Files-Raster Images*. College Park, MD: U.S. National Archives and Records Administration (NARA), June 2004.
http://www.archives.gov/research_room/arc/arc_info/techguide_raster_june2004.pdf

2.4 Organizational/Standard

The Institute for Museum and Library Services (IMLS) has also published a *Framework of Guidance for Building Good Digital Collections* (2001)

3. Current Situation

3.1 Background and objectives

The present objective internally within the GPO is to establish a prototype conversion activity to develop workflow processes and metrics to create all conversion elements that are required for the creation of a SIP.

The current system was designed to test and validate the viability of various technologies and planned processes. DCS is utilizing a pilot operation during its transition period to analyze, develop, and document reporting requirements for the future system. These requirements can then be incorporated into the evaluation criteria for components of the future system and used to evaluate the cost of implementation.

3.2 Conversion

Scanning is the only element of the conversion solution that has been benchmarked. Other elements, such as audio and video, need definition.

3.2.1 Scanning

A conversion solution does not currently exist within GPO. Digital Conversion Services (DCS) is currently a prototype operation that is producing scanned images only.

3.2.1.1 Operational situation

Current GPO equipment:

Sixteen workstations utilizing flatbed scanners. Scanning capability is 60 pages per hour per workstation/scanner.

Two workstations utilizing Auto-Document Feed (ADF) scanners. Scanning capability is 1000 pages per hour per workstation/scanner.

Equipment Guidelines:

A. Flatbed Scanner

1. *Capabilities*

   a) Allows the operator to place a single sheet or de-bound materials face down on the scan bed.
   b) Suitable for reflective media (e.g. paper, other substrates).
   c) Suitable for transmissive media such as negatives and film.

2. *Limitations*
   a) Size limitations based on scanner bed imaging area.
   b) Productivity dependant on operator performance.
   c) Fragile and brittle looseleaf books

B. Overhead Scanner/Digital Camera

1. Auto-page turning
   a) *Capabilities*
      - Suitable for bound or non-destructible material.
      - Automated features rely less on speed of the operator.
      - Scans pages while unattended or multi-tasking.

   b) *Limitations*
      - Not suitable for fragile or brittle material.
      - Not suitable for looseleaf or de-bound material.

- Size limitations based on camera/scanner imaging area.

2. Manual-page turning
   a) *Capabilities*
      - Suitable for fragile and brittle material.

   b) *Limitations*
      - Productivity dependant on operator performance.
      - Size limitations based on camera/scanner imaging area.

C. Auto-document Feed scanner

1. *Capabilities*
   a) High volume automated processing.
   b) Suitable for de-bound or destructible material.

2. *Limitations*
   a) Scans a limited volume of pages at a time based on the tray size.
   b) Occasionally introduces distortions due to moving or rotation of pages within the feeder.
   c) Size limitations based on scanner imaging area.
   d) Not suitable for rare, valuable, or brittle material.

D. Film Scanner

Used for all types of transmissive media (e.g. *microfiche, microform, negatives, aperture cards, and E-6 slides*).

1. *Capabilities*
   a) Achieves higher resolution necessary for the type/size of media.
   b) Higher quality and dynamic range.

2. *Limitations*
   a) Some film scanners are limited to certain types of media sizes (i.e. 35 mm, medium format, etc), therefore more than one type may be necessary.

3.2.1.2 Metrics

Current GPO Capabilities

Scanning capability for flatbed workflow given existing resources is 60 pages per hour per workstation/scanner.

Environment

A variety of factors will affect the appearance of images, whether displayed or printed on reflective, transmissive or emissive devices or media. Those factors that can be quantified must be controlled to assure proper representation of an image by its environment.

ISO 3664: Viewing Conditions for Graphic Technology & Photography

Monitors (refer to NARA Technical Guidelines – pp. 23)

- The monitor should be set to 24-bits (millions of colors) or greater, and calibrated to a gamma of 1.8 (Mac) or 2.2 (PC).
- Monitor color temperature set to 5000 Kelvin degrees with a desktop background of a neutral gray (avoid images, patterns, and/or strong colors).
- Monitor luminance level must be at least 85 cd/m2 and should be 120 cd/m2 or higher.
- CRT/LCD monitors designed for the graphic arts and multimedia are recommended for a digitization environment.
- Using a target such as the NARA Monitor Adjustment Target or a Kodak Grayscale can be used to adjust the monitor aimpoints of brightness / contrast for calibration *(refer to NARA Technical Guidelines – pp. 24)*

Room

- Ambient room lighting should be kept at or below 5000 Kelvin color temperature and should be dispersed/diffused throughout the room, not directly overhead causing glare problems. *(refer to NARA Technical Guidelines – pp. 23)*
- The room should be relatively dust free by use of a air filter and commitment to keeping all scanning systems free of dust and other particles.

Quantifying Scanner/Digital Camera Performance

Digitization Standards

Tests should be performed on all image capture equipment prior to purchase and throughout the life cycle of the equipment to ensure quality standards and verification of optimal performance. The following standards should be looked at as benchmarking tools to assess all equipment by either requesting test results from the vendor/manufacturer of imaging equipment or performing an evaluation with the use of a test target for performance metrics. These standards can be purchased from ISO at http://www.iso.ch or from IHS Global at http://global.ihs.com or other affiliated standards organizations such as ANSI at http://www.ansi.org/ or AIIM at http://www.aiim.org.

Subject	Document Number
Terminology	
Photography -- Electronic still-picture imaging – Terminology	ISO/FDIS 12231.2. July 2004 or 2005
Data Dictionary - Technical Metadata for Digital Still Images (Draft standard for trial use.)	NISO Z39.87-2002 AIIM 20-2002

Opto-Electronic Conversion Function	
Photography -- Electronic still-picture cameras -- Methods for measuring opto-electronic conversion functions (OECFs)	ISO 14524:1999
Resolution	
Photography -- Electronic still-picture cameras – Resolution measurements.	ISO 12233:2000
Photography -- Electronic scanners for photographic images -- Spatial resolution measurements -- Part 1: Scanners for reflective media	ISO 16067-1:2003
Photography -- Electronic scanners for photographic images -- Spatial resolution measurements -- Part 2: Film scanners	ISO16067-2 Sept. 2004
Photographic & Electronic Imaging (Resolution definition and application for evaluation of photographic and electronic systems.)	ANSI/AIIM TR26-1993
Noise	
Photography -- Electronic still picture imaging – Noise measurements	ISO 15739:2003
Dynamic Range	
Photography -- Electronic scanners for photographic images -- Dynamic range measurements	ISO 21550 Sept. 2004
Viewing Conditions	
Viewing Conditions—Graphic technology and photography	ISO 3664:2000
Viewing Conditions—Graphic Technology – Displays for color proofing	ISO 12646
Color	
Photography and graphic technology – Extended color encodings for digital image storage, manipulation and interchange – Part 1: Architecture and requirements	ISO 22028-1:2004
Graphic technology -- Prepress digital data exchange -- Colour targets for input scanner calibration	ISO 12641:1997
Quality Control	
Recommended Practice for Quality Control of Image Scanners. Provides procedures for ongoing quality control of image scanners, including incorporation of targets.	ANSI/AIIM MS44-1988 (R1993)
Sampling Procedures and Tables for Inspection by Attributes. Includes tightened, normal and reduced plans. (American Society for Quality)	ANSI/ASQ Z1.4-2003
Sampling Procedures and Tables for Inspection by Variables for Percent Nonconforming (American Society for Quality)	ANSI/ASQ Z1.9-2003
Sampling Procedures for Inspection by Attributes of Images in Electronic Image Management (EIM) & Micrographics Systems. Provides guidance in selecting a sampling procedure	ANSI/AIIM TR34-1996

Test Targets

Before the purchasing of new digitization equipment and after the purchase, an initial performance capability evaluation should be conducted with each digitization device. This may involve using test targets to make benchmark assessments in image quality to predict the integrity of such devices and how effective they will be. Tests are also performed to optimize the performance of an image capture device based on operational settings. These test results should be cumulated into a database to track the performance and/or any variability.

Targets used for Benchmark Testing Digital Image Capture Devices

3.2.2 Inspection

In the prototype environment, all scanned images are manually inspected.

Document Inspection prior to scanning

- Determine that all pages are in each publication.
- Determine if there is any damage to publications:
- Torn pages
- Damaged spine
- Stains
- Smudges
- Wrinkles

3.2.2.1 Current operational situation

Eight workstations are dedicated to inspection. Inspection is a manual examination of the page as compared to the image.

Document Characterization

Categories of Material	Handling	Types of Scanners
Type A*:* **Rare, valuable, & brittle** 	• Must be specially handled with white, static-free gloves and treated with care. • Pages turned carefully and book must not be mishandled or dropped. • All areas kept free of extraneous paper dust and dirt through careful measures such as, compressed air or by lightly dusting over the imageable surface. • Some documents may require a translucent protective sleeve prior to digitization.	• Overhead Scanner/Digital Camera – Manual-Page Turning **ONLY** • Flatbed Scanner

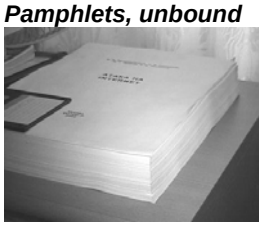
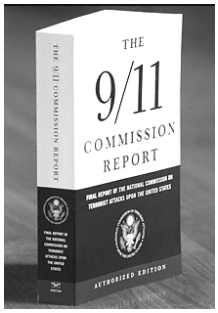
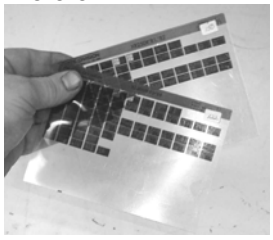
Type B: ***Pamphlets, unbound*** 	• Can be separated and run through an automated feed process. • Can be unfolded and placed flat on an imageable surface. • Some may require removal of binding materials (*ie. staples, stitches, spiral, comb-binding, tape, etc*.)	• Auto-document Feed scanner • Flatbed Scanner
Type C*: Bound* 	• Publications scanned while intact and in its original bound form. • Can be opened and placed flat on an imageable surface.	• Overhead Scanner/Digital Camera – Auto/Manual-Page Turning
Type D: ***Fold-outs, maps, posters*** 	• Can be separated and run through an automated feed process. • Can be unfolded and placed flat on an imageable surface. • Some are larger formats and may require a larger scanner/camera imaging device to capture the whole area.	• Flatbed Scanner • Wide Format Cameras/Scanners
Type E: ***Microform*** 	• Many different formats/sizes that may require specific equipment or handling, therefore more than one type of scanner may be necessary.	• Film scanner (*various types*) • Flatbed Scanner

Image Capture Classification

1. Determine the type of image capture mode performed on each page

   RGB (Color halftones, solid images, photographs, charts, or any type of continuous-tone image)
   Grayscale (Non-color halftones, solid images, photographs, charts, or any other type of continuous-tone image)
   Bitonal (Black and white only – text matter or line-art matter)

3.3 Content Management

3.3.1 Image Workflow

Currently DCS utilizes a manual process for file workflow tracking and management.

The product set selected by DCS will support document/data capture and production/ad-hoc scanning in a single application. The application will also have a strong Application Programming Interface (API) to expand functionality when needed within the functionality of the COTS product selected. Most structured and unstructured documents can be scanned in batches, and the system should have the capability to automatically recognize each document in a batch and process them based on characteristics that have been predefined. The product's workflow should be integrated and manage documents allowing a high level of control over how the diverse types of documents that GPO will manage are processed. The product must provide the capability to define and modify workflows.

The selected product set should combine both document and data capture and allow remote Internet-based capture for future use. Capture stations should be designed with simplistic configuration procedures in place. Capture stations should be located at GPO's HQ site and at possible remote sites—across geographic regions or in the same building—and should be able to synchronize with a central capture site via the Internet. It is important that the product selected have an open architecture that makes it easy to extend the basic application to handle complex, high-volume document processing. The product should also be able to predefine "batch definitions or classes" to allow all classes and types of documents to be captured.

3.4 Storage

Storage of scanned images is on a network server, with standard IT back-up processing in place.

4. Desired Situation

4.1 Background Changes

Create a scanning environment that incorporates automated workflow software, with combinations of scanning equipment and efficient user interfaces to support each area within the workflow.

4.1.1 Specific Component

A Scanning module should be available to create batches, scan and import documents, and edit the contents of batches. A batch is a set of data or jobs to be processed in a single program run or a quantity required for or produced as the result of one operation. In most cases within a digitization environment, a single batch will be an entire publication or group of publications from a single customer/source. After the batches are created, they should be able to be entered into temporary storage in the system, making them available for processing by subsequent modules.

- **Batch creation:** The operator creates the batch by selecting the type of batch to create (the batch class) and then scanning or importing documents and pages. The document images are stored in a temporary folder for further processing by the system.

- **Batch editing:** Once the batch is created, the operator can visually check documents or pages, and edit them as necessary. Editing functions include replacing, reordering, or rejecting documents and pages. Entire documents or individual pages can be rotated and saved in the rotated state.

4.1.1.1 Objectives

To design a system that constructs as many "mini" conversion pipelines that can stand on their own should a failure occur. Each of these mini pipelines or "clusters" contain workflow, scanning, recognition, key-from image, key from paper, QA, storage functionality and the people to staff its stations. All of the clusters are then managed by a site-level workflow manager which normally manages workflow for all of the clusters, provides administrative functions and communicates with sites and services outside of the confines of the current site.

The system will be broken down into as many "independent clusters" as required to help guarantee reliability. Workflow and administrative functions at the site level will also be organized in a way to make sure that backups and administrative tasks are built to make a cluster as independent as possible.

4.1.1.2 Metrics

Metrics of workflow will follow previously mentioned ANSI and ISO standards.

4.1.1.3 Priorities among changes

1) Workflow Software
2) Batch Processing for Digitization of Documents
3) Quality Control Process
4) Process for Metadata Capture

5. Benchmarks

Image Capture Benchmarks for Preservation Masters *(refer to NARA Technical Guidelines – pp. 32-36)*

Scanner Setup (refer to DLF – pp. 3, NARA-pp.52)

Image Types	Bit Depth	Color Mode	Resolution (ppi/spi)	Scale	File Format	Compression
Reflective						
B&W Text Only	1-bit	B&W (bitonal)	600 ppi/spi	100% (1:1)	TIFF	CCITT Group 4
B& W Text with Illustrations (charts, artwork, graphs, photos)	8-bit	Grayscale	400 ppi/spi *		TIFF	None
Color Photos & Illustrations with Text	24-bit	RGB	400 ppi/spi *		TIFF	None

Transmissive						
16mm	36-48 / 16 bit	Color / Grayscale	5000 ppi/spi	1600% (16:1)	TIFF	None
35mm	36-48 / 16 bit	Color / Grayscale	3400 ppi/spi	850% (8.5:1)		
2-1/4"	36-48 / 16 bit	Color / Grayscale	1800 ppi/spi	450% (4.5:1)		
4" x 5"	24-48 / 8-16 bit	Color / Grayscale	800 ppi/spi	200% (2:1)		
8" x 10" +	24-48 / 8-16 bit	Color / Grayscale	400 ppi/spi	100% (1:1)		

* Scanning resolutions for images over 11 x 16" (300 ppi for 8-bit grayscale and 300 ppi for 24-bit RGB color)

1. Originals will be backed with bright white opaque paper for flatbed scanning.

2. **Scan Kodak Grayscale Target (Q-13 or Q-14)**, or an equivalent 14-step or 20-step grayscale, only on publications required to preserve color/grayscale data and to further evaluate of the tonal/dynamic range of the scanning device output.

3. **Choose best defined presets to digitally capture type of publication** – **Based on all these factors**:

a) *Color Mode* – to best define the color of the original publication format.

b) *Scaling* – to best define the digital capturing parameters according to *III.A Scanner Setup specifications.*

c) Size/Crop – assuring that an area of at least 1/4" outside of the parameters of the open page(s) is captured.

d) Resolution – using the correct amount of this is dependant on the type of media as well as the content itself according to *III.A Scanner Setup* specifications. (ie. *transmissive vs. reflective, color vs. grayscale vs. bitonal*)

e) *Descreen* – to remove any printed halftones that cause the obtrusive moiré patterns when digitally capturing from printed material such as newsprint or magazine-type paper.

f) Paper/Print Mode – to determine the optimal settings for the scanner/camera to capture the best rendering of the original (*ie. Some scanner API's have substrate mode [magazine/coated, newsprint, uncoated, photograph] to choose from for the purposes of descreening or other capture features*)

g) For significant embossed seals / images, the flatbed scanner must be set use One Directional Light

h) *Tonal Adjustments* – scanner hardware and software must be equipped and capable of capturing correct highlights/shadows without losing detail. Also, the software should use tools with more controls (Levels and curves) along with numeric feedback.

i) *Color management* could be involved in any settings using proper calibration software for both monitors and image capture devices (Cameras and scanners).

NOTE: Presets will be programmed for each scanner based on these definitions.

Curvature Reduction

If available in the API (Application Programming Interface) of the scanning software, applying an in-process setup to reduce the curvature or rotation of pages during the scanning phase may be necessary.

Aimpoints for Grayscale Target (Tone Compression)

On the preservation master file, the original scan contains a grayscale target. Tone compression is a technique to make the digital reproduction to look like the original in terms of the exact tonal range.

NOTE: This theory should not be applied in all cases, due to each publication's variation in quality attributes due to aging, or the process used in the creation of the publication.

Scanning Aimpoints for Grayscale Target (Q-13) using 24-bit Color Mode

		Neutralized White Point	Neutralized MidPoint	Neutralized BlackPoint
Step or Density	Kodak Q-13/14	A	M	B
	Visual Density	0.05 – 0.10	0.75 – 0.85	1.65 – 1.75
Aimpoint	RGB Level	242-242-242	122-122-122	40-40-40
	% Black	4%	60%	90%
Acceptable Range	RGB Level	236 – 248	116-128	34-46
	% Black	2 – 6%	58 – 62%	88 – 92%

Aimpoint Variability

For the three aimpoint values described above, none should exceed a variability of ± 6 RGB increments per each individual channel: Red, Green, and Blue. You can verify this by using an image sampler in the scanner software tools or an eyedropper tool from image processing software (such as Adobe Photoshop or equivalent) and set to measure an average of either 3 x 3 or 5 x 5 pixels to sample on the grayscale.

Note: never use a point sample or single pixel sample to base your measurement on.

Verification and Save

Results of the scan - All converted images must be inspected to ensure the highest quality possible. Images shall not contain dust representation, digital artifacts, scratches, poor color contrast, poor saturation, incorrect cropping, noise, duplicates, missing images, or any unknown discrepancies not visible on the original tangible piece. Conversion equipment must be configured and maintained to meet the requirements for digitization. This includes but is not limited to profiles, calibration, and cleanliness. All quality discrepancies must be corrected prior to release.

Minimum (submission) level Metadata - Each publication scanned and digitized, must have a minimal level of metadata associated with each TIFF file for preservation purposes. The data elements will consist of bibliographic, technical, and administrative information necessary to track, manage, and preserve the associated files with each title for the future content management system. The TIFF data elements and values (e.g. presented in XML as fields with values associated with file header tags), represent metadata used to render and manage image data.

GPO submission level metadata will capture:

(1) Identity
  (a) Title or caption
  (b) Unique Identifier (persistent locators, filenames, ISNs, etc)

(2) Responsibility
  (a) Author / Creator
  (b) Publisher / Authority
  (c) Rights Owner *

(3) Version / Fixity*
  (a) Version information
  (b) Relationship to other version or manifestations

(4) Representation / Technical / Structure*
  (a) Must incorporate NISO Z 39.87-2002 technical metadata for digital still images
  (b) Structure Information

*If readily available

File Naming Convention –The system identifier requires machine or human indexing for corresponding files that relate to each document. Through a standard naming convention, the process of ingest, storage, search, and retrieval of documents is simplified. Files derived from conversion processes shall be assigned a unique 9 digit alphanumeric identifier conforming to the Code 39 barcoding standard (ANSI: BC1-1995). The first digit is the fixed letter “a” which enables validation for a METS schema later assigned. Digits will include 0-9 and letters A-Z (minus i and o). Scanned Publications at the page level shall be assigned the publication unique ID followed by an underscore and a sequential 5 digit identifier.

Example: A12345678_00001

2. **Submission Information Package (SIP)** – The images may be in RGB, Grayscale, or Bitonal mode and should have a unique identifier and metadata associated with each file. The quality of the files derived from conversion shall conform to the FDsys Operational Specification for Converted Content.

6. Constraints

- Not incorporating automated workflow software will constrain throughput.
- Not upgrading scanning equipment capability will constrain document scanning options.
- Not automating Quality Control process will increase personnel required, and constrain throughput.