

## **Specification for Quality Control (Version 1.1)**

(March, 2006)

Specifications and Operating Procedures for Quality Control:  
Creation of Preservation Master Files

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

**Specifications and Metrics for Quality Control of  
Converted Content** – a functional process within Digital  
Conversion Services (DCS)

***United States Government Printing Office (GPO)***

FINAL



### Table of Contents

- Specification for Quality Control (Version 1.1) ..... 1
- Specifications and Metrics for Quality Control of Converted Content – a functional process within Digital Conversion Services (DCS)..... 1
  - United States Government Printing Office (GPO) ..... 1
- Specification for Quality Control v1.0 ..... 2
- Specification for Quality Control v1.1 ..... 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..... 6
  - 3.1 Background and Objectives for DCS ..... 6
  - 3.2 Quality Control ..... 6
    - 3.2.1 Metrics..... 6
- 4. Image Capture Requirements ..... 8
  - Image Capture Classification* ..... 8
- 5. Quality Control Requirements ..... 9
  - 5.1 Core Requirements ..... 9
  - 5.2 Process Requirements..... 9
  - 5.3 Color Reviewing Requirements ..... 10

# 1. Scope

What is addressed in this document:

- Purpose
- Quality Control requirements for Converted Content
- DCS Operational Procedures
- Environment
- Quality Control Standards
- Required hardware/software configurations
- Suggested software

Types of converted content 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 scan content or create a Submission Information Package (SIP). This information can be obtained from the Operational Specification for Converted Content and FDSys Submission Information Package (SIP) Specification.

## **1.1 Deliverable**

The end product of the Quality Control will be converted content configured to quality requirements and FDSys specifications. Quality Control will ensure that the content is ready for packaging and ingest by FDSys.

## **1.2 Overview**

This specification covers all the necessary requirements and elements that are required for the quality control of converted content.

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 contain content 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 Quality Control 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 21<sup>st</sup> 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](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](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 for DCS

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 Quality Control

The quality control process is done between the scanning and OCR process steps within Digital Conversion Services. Quality Control is expected to deliver a perfect product at a 95% confidence level or above.

A 100% page by page Quality Control should be conducted for material derived from manual single page scanners (flatbed scanners, overhead scanners, digital cameras).

When content is converted from Auto Document Feed (ADF) or Auto Page Turn Scanners, systematic sampling selection shall occur (see ANSI/ASQ Z1.4-2003, ANSI/ASQ Z1.9-2003, ANSI/AIIM TR34-1996).

Operators use an image selection tools (e.g. Thumbs-Plus, etc.) to view and do QA checks for basic quality attributes (resolution, color mode, skew, noise, speckle, out-of-order pages, missed pages, duplicated pages, etc).

Within Digital Conversion Services, operators in each process are challenged to not make mistakes or jeopardize the quality and integrity of each file. Quality Control plays the lead role in providing a product that meets GPO and customer quality requirements. As a result, Quality Control operators are expected to be very detailed oriented and lead by example within the DCS environment. It is the Quality Control operator's responsibility to locate all scanning errors, understand and choose corrective measures, explain corrective measures to scanner operators, and to provide a positive environment when doing such.

#### 3.2.1 Metrics

##### 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/m<sup>2</sup> and should be 120 cd/m<sup>2</sup> 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 environments free of dust and other particles.

**Quantifying Performance**

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
<b>Terminology</b>	
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

## 4. Image Capture Requirements

### 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
<b>Reflective</b>						
<i>B&amp;W Text Only</i>	1-bit	B&W (bitonal)	600 ppi/spi	100% (1:1)	TIFF	CCITT Group 4
<i>B&amp;W Text with Illustrations (charts, artwork, graphs, photos)</i>	8-bit	Grayscale	400 ppi/spi *		TIFF	None
<i>Color Photos &amp; Illustrations with Text</i>	24-bit	RGB	400 ppi/spi *		TIFF	None
<b>Transmissive</b>						
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)

#### **Image Capture Classification**

How to determine the type and settings for each page.

- a) **Color Mode** – to best define the color of the original publication format.
  - **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)
- b) **Size/Crop** – assuring that all original content is captured.
- c) **Resolution** – dependant on the type of media as well as the content itself.

## 5. Quality Control Requirements

### 5.1 Core Requirements

**5.1.1** All publications must be inspected to ensure the highest quality possible.

**5.1.2** To avoid a bias and risk, sample selection shall never be chosen based on judgment or convenience.

**5.1.3** Quality Control shall ensure that all of the requirements are met a 99.5% or above accuracy level.

**5.1.4** In order to acquire a 95% or above confidence level in quality, any sampling shall be conducted following Six Sigma methodologies (see 5.2.2).

**Six Sigma:** A philosophy of managing that focuses on eliminating nonconformance through practices that emphasize understanding, measuring, and improving processes. It's based on the statistical concept of six sigma, measuring a process at only 3.4 defects per million opportunity.

**5.1.5** The following dissatisfiers shall be identified and corrected prior to release:

- incorrect color mode
- Incorrect resolution
- page omissions, duplications, improper sequence
- digital files and folders named not in accordance with specifications
- incorrect cropping
- skew from page placement represented
- dust representation
- digital artifact representation
- scratch representation
- inadequate color contrast
- inadequate brightness
- poor tonal range
- poor saturation
- noise representation

**5.1.6** Quality Control may initiate the configuration and maintenance processes to meet the requirements for digitization. This includes but is not limited to profiles, calibration, and cleanliness.

### 5.2 Process Requirements

**5.2.1** When Flatbed Scanners, Overhead Scanners, Digital Cameras are used for conversion, the following procedures must be followed:

100% page by page Quality Control shall be conducted

**5.2.2** When Auto Document Feed (ADF) Scanners, Auto Page Turn Scanners are used for conversion, the following procedures must be followed:

Systematic sampling selection shall occur.

$$N = (1.96s / \Delta)^2$$

N = minimum sample size  
S = estimate of standard deviation data

1.96 = constant representing 95% confidence level  
 $\Delta$  = the difference you're trying to detect

### 5.3 Color Reviewing Requirements

**5.3.1 When tonal/dynamic range is visibly inadequate for color/grayscale images, the following must be evaluated and corrected.**

**Kodak Grayscale Target (Q-13 or Q-14)**, or an equivalent 14-step or 20-step grayscale, must will be associated to all publications required to preserve color/grayscale data.

***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.

**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  $\pm 6$  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.