

**Library of Congress**  
**Recommended Format Specifications**  
2014-2015

**Introduction**

**Categories of creative content:**

- I. Textual Works and Musical Compositions**
  - i. Textual Works – Print
  - ii. Textual Works – Digital
  - iii. Textual works – Electronic Serials
  - iv. Digital Musical Compositions (score-based representations)
  
- II. Still Image Works**
  - i. Photographs – Print
  - ii. Photographs – Digital
  - iii. Other Graphic Images – Print
  - iv. Other Graphic Images – Digital
  - v. Microforms
  
- III. Audio Works**
  - i. Audio – On Tangible Medium (digital and analog)
  - ii. Audio – Media-independent (digital)
  
- IV. Moving Image Works**
  - i. Motion Pictures – Digital and Physical Media
  - ii. Video – File-Based and Physical Media
  
- V. Software and Electronic Gaming and Learning**
  
- VI. Datasets/Databases**
  - i. Datasets
  - ii. Geospatial Data
  - iii. Databases

## Introduction

Throughout its history, the Library of Congress has been committed to the goal best described in its mission statement “to further the progress of knowledge and creativity for the benefit of the American people”. At its core, the Library's ability to advance the nation's progress has depended upon its collection, which in turn embodies the knowledge and creativity of the many authors, composers, journalists, artists, and scientists whose work is embodied there. The quality of the collection reflects the Library's care in selecting materials and the effort it invests in preserving them and making them accessible to the American people for the long term.

To build such a substantial and wide-ranging collection and to ensure that it will be available for successive generations, the Library relies upon a wealth of expertise. In order to maximize the scope and scale of the content in the collection, the Library calls upon the wealth of expertise in languages, subject matter and trends in publishing and content creation provided by the specialists who identify and acquire material for the Library's collection. But knowledge of the technical characteristics of the production of creative works is required as well. In the past, the lasting power of the collections depended exclusively upon the endurance of such materials as the paper, ink, and binding of a book; the acetate or paper coated with gelatin in a photograph; or the shellac, vinyl, and coated polyester that comprise a sound recording. Although these materials remain in use today, creators and publishers have also begun to employ a wide array of intangible digital formats, as well as continuing to change and adapt the physical formats in which they work. The Library needs to be able to identify the formats which are suitable for large-scale acquisition and preservation for long-term access if it is to continue to build its collection and ensure that it lasts into the future.

The Library also appreciates that the work of identifying the formats of creation best suited to preservation and long-term access will be of benefit to broader communities as well. The Library shares with many other institutions, organizations and corporations the common goal of the creation, dissemination and preservation of knowledge, which benefits the nation socially, economically and politically. In seeking this goal, the Library has a unique position as America's foremost institution dedicated to the preservation and provision of access to the creative output of the nation and the world; and one of its fundamental goals is to provide the benefit of its expertise and knowledge to support and assist those others who could benefit from it in their own missions.

The Library accepts and embraces its responsibility to identify the formats, both analog and digital, which will best meet the needs of all concerned, maximizing the chances for survival and continued accessibility of this creative content well into the future. To meet this need, the Library has decided to develop a set of specifications of formats which it recommends, both internally to its own professionals and externally to creators, vendors and archivists, as the preferred ones to use to ensure the preservation and long-term access of the creative output of the nation and the world. To do so, the Library was able to build upon the experience and knowledge it has accumulated with its partners in the Copyright Office in providing what is known as the “Best Edition Statement”.

The specifications which the Library is now publishing do not replace or supersede the Best Edition Statement, which provides guidance to publishers and creators in fulfilling their obligations with regard to the registration or deposit of their works under the terms of the Copyright Law. Instead, it seeks to complement that work, building upon the knowledge gained from working with the Best Edition Statement and providing a broader set of recommendations, aimed at providing guidance and clarity in

a creative world, which is both rich with potential and rife with pitfalls, and afforded numerous competing options for information format or container.

In 2011, the Library began to examine existing Best Edition Statements and then started to work with experts on a wide variety of formats in which the creators of knowledge produce their work products. Calling upon their knowledge of the creative and publishing landscape and marketplace, and calling upon external experts when necessary, the Library was able to identify six broad categories of creative output: Textual Works and Musical Compositions; Still Image Works; Audio Works; Moving Image Works; Software and Electronic Gaming and Learning; and Datasets/Databases. Within each of these broad categories, there are further subdivisions, down to specific recommendations on various aspects of print books, digital photographs, motion picture films or electronic datasets, for example. This is not to say that the Library is seeking to dictate each and every aspect of the creation of these works. The specifications seek to provide a framework within which creative works should have the flexibility to grow and develop, and also help ensure that these creative works be accessible and authentic into the future.

The Library's commitment to the long-term survival of the creative output of the nation and the world means that this set of specifications is a living document. The creative world by its very nature is a dynamic one and so the framework must live, adapt and grow alongside it. As such, the Library will be revisiting these specifications on an annual basis. It is not expected that this will result in root-and-branch changes in the course of any one of these revisions. It is in fact hoped that, by engaging with the specifications on an annual basis, revisions will be smaller and more manageable as there be less room for the Library's specifications to slip out of sync with developments in the creative world. In the months preceding the annual review, the Library will seek out and request input from stakeholders to ensure that all parties who could use and will benefit from this set of specifications are fully aware and engaged in any and all revisions.

The creation and publication of these recommended format specifications is not intended to serve as an answer to all the questions raised in preserving and providing long-term access to creative content. They do not provide instructions for receiving this material into repositories, managing that content or undertaking the many ongoing tasks which will be necessary to maintain this content so that it may be used well into the future. Tackling each of those aspects is a project in and of itself as each form of content has a unique set of facets and nuances. These specifications provide guidance on identifying sets of formats which are not drawn so narrowly as to discourage creators from working within them, but will instead encourage creators to use them to produce works in formats which will make preserving them and making them accessible simpler. Following these specifications helps make it realistic to build, grow and save creative output for our individual and collective benefit for generations to come.

At a time of such great growth in the production of creative output, not only are the frontiers expanding, but new ones seem to crop up faster than we can grasp them, there is a definite need for some expert guidance, so that this amazing creative content is not lost to us. The Library of Congress appreciates that it is uniquely positioned to provide that guidance and, in fact, that its position has given it that responsibility. The Library is the nation's premier institution instructed to further the progress of knowledge and creativity for the benefit of the American people. In producing and publishing these format specifications, it seeks to meet that charge, and to provide the benefit of its expertise for creators, vendors, and archivists, so that they might succeed in their goals to share and disseminate their creative output and benefit the nation generally.

## I Textual Works and Musical Compositions

### i. Textual Works – Print

(Books, musical compositions, etc.)

#### Preferred:

##### A Paper

- 1 Archival quality paper (ISO 11108: 1996 for Archival Paper)

##### B Printing Process, in descending order of preference:

- 1 Lithography (offset printing press)
- 2 Electrophotography (digital press)
- 3 Inkjet (inkjet printer using stable pigment or dye-based inks)

##### C Binding and Packaging

- 1 Slip-cased, if available
- 2 Binding, in descending order of preference:
  - a Hard cover
    - 1 Library binding (NISO Z39.78-2000)
    - 2 Sewn
    - 3 Glued only
  - b Soft cover
    - 1 Sewn
    - 2 Glued only
    - 3 Spiral- or plastic-bound
    - 4 Stapled
  - c Loose-leaf (including all binders and indexes published as part of the deposit and offered for sale and distribution)

##### D Size

- 1 Larger-sized editions (Note: large-type editions are not preferred over editions with conventional size typefaces)
- 2 For broadsides and musical compositions, the Library prefers items:
  - a In protective folders
  - b Rolled (rather than folded)

##### E Rarity, Special Features, Illustrations

- 1 Limited editions (including those with special binding or special features)
- 2 Editions with the greatest number of unique features (such as pop-ups, overlaps, magnifiers, overlays, tabs, notches, etc.)
- 3 Illustrated editions; original color illustrations preferred over black and white reproductions

##### F Completeness

- 1 Complete work. For items published in a finite number of separate components, all elements published as part of the work and offered for sale or distribution must be submitted.
- 2 All updates, supplements, releases, and supersessions published as part of the work and offered for sale or distribution must be submitted. Insertions (including all binders and indexes) must be received in a regular and timely manner for proper maintenance of the deposit.
- 3 For unaccompanied vocal musical compositions: open score, with each part on separate staff
- 4 For vocal musical compositions with orchestral accompaniment and for instrumental musical compositions:
  - a Full score and up to 13 parts, if applicable; if published only by rental, lease, or lending, submit full score only
  - b Conductor's score and up to 13 parts, if applicable; if published only by rental, lease, or lending, submit conductor's score only

#### Acceptable:

n/a

### ii. Textual works - Digital

(Electronic books, etc. For electronic serials, see sec. I.iii, below; for digital musical compositions, see sec. I.iv, below.)

#### Preferred:

##### A Content

- 1 Formats, in descending order of preference:
  - a XML-based markup formats, with included or accessible DTD/schema, XSD/XSL presentation stylesheet(s), and explicitly stated character encoding
    - 1 BITS-compliant (NLM Book DTD)
    - 2 EPUB-compliant

- 3 Other widely-used book DTDs/schemas (e.g., TEI, DocBook, etc.)
- b Page-layout formats
  - 1 PDF/UA (ISO 14289-1-compliant)
  - 2 PDF/A (ISO 19005-compliant)
  - 3 PDF (highest quality available, with features such as searchable text, embedded fonts, lossless compression, high resolution images, device-independent specification of colorspace, content tagging; includes document formats such as PDF/X)
- 2 Character encoding, in descending order of preference:
  - a UTF-8, UTF-16 (with BOM), US-ASCII
  - b ISO 8859-1
- 3 Rarity and Special Features
  - a Limited editions (including those with special features such as high resolution images)
  - b Editions with the greatest number of unique features (such as additional content, multimedia, interactive elements, etc.)
- 4 Completeness
  - a Complete work. For items published in a finite number of separate components, all elements published as part of the work and offered for sale or distribution must be submitted. Includes all associated external files and fonts considered integral to the publication.
  - b All updates, supplements, releases, and supersessions published as part of the work and offered for sale or distribution must be submitted and received in a regular and timely manner for proper maintenance of the deposit.

#### B Metadata

- 1 Title, creator, creation date, place of publication, publisher/ producer/ distributor, ISBN, contact information
- 2 Include if available: language of work, other relevant identifiers (e.g., DOI, LCCN, etc.), edition, subject descriptors, abstracts

#### C Technological measures

- 1 Files must contain no measures that control access to or use of the digital work (such as digital rights management or encryption)

---

#### Acceptable:

#### A Content

- 1 Formats, in descending order of preference:
  - a Other structured or markup formats
    - 1 XHTML or HTML, with DOCTYPE declaration and presentation stylesheet(s)
    - 2 XML-based document formats (widely-used and publicly-documented), with presentation stylesheet(s) if applicable. Includes ODF (ISO/IEC 26300) and OOXML (ISO/IEC 29500).
    - 3 SGML, with included or accessible DTD
    - 4 Other XML-based non-proprietary formats, with presentation stylesheet(s)
    - 5 XML-based formats that use proprietary DTDs or schemas, with presentation stylesheet(s)
  - b Page-layout formats
    - 1 PDF (web-optimized)
  - c Other formats
    - 1 Rich text format
    - 2 Plain text
    - 3 Widely-used proprietary word-processing formats
    - 4 Other text formats not listed here
- 2 Other character encodings not listed here

#### B Metadata

- 1 Title, creator, creation date, place of publication, publisher/ producer/ distributor, ISBN, contact information
- 2 Include if available: language of work, other relevant identifiers (e.g., DOI, LCCN, etc.), edition, subject descriptors, abstracts

#### C Technological measures

- 1 Files must contain no measures that control access to or use of the digital work (such as digital rights management or encryption)

---

### iii. Textual Works – Electronic serials

#### Preferred:

#### A Content

- 1 Formats, in descending order of preference:
  - a Content compliant with the NISO JATS: Journal Article Tag Suite (NISO Z39.96-201x) with XSD/XSL presentation

- stylesheet(s) and explicitly stated character encoding
- b Page-layout formats
    - 1 PDF/UA (ISO 14289-1-compliant)
    - 2 PDF/A (ISO 19005-compliant)
    - 3 PDF (highest quality available, with features such as searchable text, embedded fonts, lossless compression, high resolution images, device-independent specification of colorspace; content tagging; includes document formats such as PDF/X)
  - 2 Character encoding, in descending order of preference:
    - a UTF-8, UTF-16 (with BOM), US-ASCII
    - b ISO 8859-1
- B Completeness**
- 1 Complete work. All elements considered integral to the publication and offered for sale or distribution must be submitted – e.g., articles, table(s) of contents, front matter, back matter, etc..
  - 2 All updates, supplements, releases, and supersessions published as part of the work and offered for sale or distribution must be submitted and received in a regular and timely manner for proper maintenance of the deposit.
- C Metadata**
- 1 Title-level metadata: serial or journal title, ISSN, publisher, frequency, place of publication
  - 2 Article-level metadata as relevant or applicable: volume(s), number(s), issue date(s), article title(s), article author(s), article identifier (DOI, etc.)
  - 3 Include if available: other descriptive metadata (e.g., subject heading(s), descriptor(s), abstract(s))
- D Technological Measures**
- 1 Files must contain no measures that control access to or use of the digital work (such as digital rights management or encryption)

**Acceptable:**

**A Content**

- 1 Formats, in descending order of preference:
  - a Other structured or markup formats:
    - 1 Widely-used serials or journal non-proprietary XML-based DTDs/schemas with presentation stylesheet(s)

- 2 Proprietary XML-based format for serials or journals (with documentation) with DTD/schema and presentation stylesheet(s)
  - 3 XHTML or HTML, with DOCTYPE declaration and presentation stylesheet(s)
  - 4 XML-based document formats (widely used and publicly documented), with presentation stylesheets, if applicable. Includes ODF (ISO/IEC 26300) and OOXML (ISO/IEC 29500).
- b Page-layout formats
    - 1 PDF (web-optimized with searchable text)
  - c Other formats
    - 1 Rich text format
    - 2 Plain text
    - 3 Widely-used proprietary word processing or page-layout formats
    - 4 Other text formats not listed here
  - 2 Other character encodings not listed here
- B Metadata**
- 1 Title-level metadata: serial or journal title, ISSN, publisher, frequency, place of publication
  - 2 Article-level metadata as relevant or applicable: volume(s), number(s), issue date(s), article title(s), article author(s), article identifier (DOI, etc.)
- C Technological Measures**
- 1 Files must contain no measures that control access to or use of the digital work (such as digital rights management or encryption)

---

**iv. Digital Musical Compositions (score-based representations)**

**Preferred:**

**A Content**

- 1 Formats, in descending order of preference:
  - a XML-based markup music notational format, with included or accessible DTD/schema, XSD/XSL presentation stylesheet(s), and explicitly stated character encoding.
    - 1 MusicXML
    - 2 Music Encoding Initiative (MEI)

- 3 Other widely-used and publicly documented musical notation DTDs/schemas
  - b Page-layout formats
    - 1 PDF/UA (ISO 14289-1-compliant)
    - 2 PDF/A (ISO 19005-compliant)
    - 3 PDF (highest quality available, with features such as searchable text, embedded fonts, lossless compression, high resolution images; includes document formats such as PDF/X)
  - 2 Character Encoding, in descending order of preference:
    - a UTF-8, UTF-16 (with BOM), US-ASCII
    - b ISO 8859-1
  - 3 Rarity and Special Features
    - a Limited editions (including those with special features)
    - b Editions with the greatest number of unique features (such as additional content, multimedia, interactive elements, etc.)
  - 4 Completeness
    - a Complete work. For items published in a finite number of separate components, all elements published as part of the work and offered for sale or distribution must be submitted. Includes all associated external files and fonts considered integral to the publication.
    - b All updates, supplements, releases, and supersessions published as part of the work and offered for sale or distribution must be submitted and received in a regular and timely manner for proper maintenance of the deposit.
    - c For unaccompanied vocal musical compositions: open score, with each part on separate staff
    - d For vocal musical compositions with orchestral accompaniment and for instrumental musical compositions:
      - 1 Full score and all parts, if applicable; if published only by rental, lease, or lending, full score only may be submitted
      - 2 Conductor's score and all parts, if applicable; if published only by rental, lease, or lending, conductor's score only may be submitted
- B Metadata**
- 1 Title, creator, creation date, place of publication, publisher/ producer/ distributor, ISMN, contact information

- 2 Include if available: language of work, other relevant identifiers (e.g., ISBN, DOI, LCCN, etc.), edition, subject descriptors, abstracts
- C Technological Measures**
- 1 Files must contain no measures that control access to or use of the digital work (such as digital rights management or encryption)

**Acceptable:**

**A Content**

- 1 Formats, in descending order of preference:
  - a Other structured or markup formats
    - 1 XHTML or HTML, with DOCTYPE declaration and presentation stylesheet(s)
    - 2 SGML, with included or accessible DTD
    - 3 XML-based musical notation formats that use proprietary DTDs or schemas, with presentation stylesheet(s)
  - b Page-layout formats
    - 1 PDF (web-optimized)
  - c Other formats
    - 1 Widely-used proprietary music notation formats
    - 2 Other music composition formats not listed here
- 2 Other character encodings not listed here

**B Metadata**

- 1 Title, creator, creation date, place of publication, publisher/ producer/ distributor, ISMN, contact information
- 2 Include if available: language of work, other relevant identifiers (e.g., ISBN, DOI, LCCN, etc.), edition, subject descriptors, abstracts

**C Technological Measures**

- 1 Files must contain no measures that control access to or use of the digital work (such as digital rights management or encryption)

## II Still Image Works

### i. Photographs – Print

#### Preferred:

##### A Permanence and appearance

- 1 Unmounted
- 2 Pigmented inks (if digitally printed)
- 3 Fixed, well-washed (if wet chemistry method)

##### B Size

- 1 Largest available (if multiple sizes)
- 2 Min: 8 x 10"
- 3 Max: 28 x 36"

##### C Faithful representation of the work, as published. Equal in quality to the publication version.

##### D Metadata

- 1 Title, creator, creation date, place of publication, publisher/ producer/ distributor, contact information.
- 2 Include if available: language of work, other relevant identifiers (e.g., ISMN, DOI, LCCN, etc.), subject descriptors, abstracts; key or reference to each data field and technical production information (type of paper, how processed, publisher internal tracking numbers).

#### Acceptable:

n/a

### ii. Photographs – Digital

#### Preferred:

##### A Technical characteristics

- 1 Highest resolution available, not rescaled or interpolated
- 2 Highest bit depth available, 16 bits per channel if available
- 3 Specified color space used in published version
- 4 Uncompressed
- 5 Unlayered
- 6 Best available version, full clarity (e.g. highest resolution, highest bit depth, uncompressed).

##### B Formats, in order of preference

- 1 TIFF (uncompressed)
- 2 JPEG2000 (lossless) (\*.jp2)
- 3 PNG (\*.png)
- 4 JPEG/JFIF (\*.jpg)
- 5 Digital Negative DNG (\*.dng)
- 6 JPEG2000 (lossy) (\*.jp2)
- 7 TIFF (compressed)
- 8 BMP (\*.bmp)
- 9 GIF (\*.gif)

##### C Metadata

- 1 If supported by format, include following metadata: title, creator, creation date, place of publication, publisher/ producer/ distributor, contact information. Include if available: language of work, other relevant identifiers (e.g., DOI, LCCN, etc.), subject descriptors, abstracts; key or reference to each data field and technical production information (ex. EXIF metadata from digital camera).
- 2 Otherwise provide separately in external text file.

##### D Technological Measures

- 1 Files must contain no measures that control access to or use of the digital work (such as digital rights management or encryption)

#### Acceptable:

##### A Formats

- 1 TIFF (in Planar format)
- 2 PhotoShop (\*.psd)
- 3 Camera raw (\*.dng, \*.raw, \*.nef, \*.crw)
- 4 JPEG 2000 Part 2 (\*.jpf, \*.jpx)
- 5 FlashPix (\*.fpx)
- 6 Computer Graphic Metafile (CGM, WebCGM) (\*.cgm)
- 7 Encapsulated Postscript (EPS)

##### B Compression

- 1 Lower compression ratios preferred
- 2 Discrete wavelet transform (DWT) preferred to discrete cosine transform (DCT)

##### C Technological Measures

- 1 Files must contain no measures that control access to or use of the digital work (such as digital rights management or encryption)

### iii. Other Graphic Images – Print

#### Preferred:

##### A Permanence and appearance

- 1 Packaging materials equivalent to published form (e.g. binding, box/packaging materials)
- 2 If multiple versions available, provide the most widely distributed edition.
- 3 If limited edition, provide an unnumbered but otherwise identical copy
- 4 For large items, provide rolled, unfolded
- 5 Prefer edition with protective coatings
- 6 Faithful representation of the work, as published. Equal quality to publication version.

##### B Cartographic materials, in order of preference

- 1 Most complete data (all layers, appendices)
- 2 Largest size
- 3 Most widely distributed

##### C Related Materials

- 1 Include indexes, study guides, or other matter if available
- 2 Also include annotations, accompanying tabular or textual matter, or other interpretive aids

##### D Metadata

- 1 Include associated descriptive metadata: title, creator, creation date, place of publication, publisher/ producer/ distributor, contact information. Include if available: language of work, other relevant identifiers (e.g., DOI, LCCN, etc.), subject descriptors, abstracts; key or reference to each data field.

#### Acceptable:

n/a

### iv. Other Graphic Images – Digital

#### Preferred:

##### A Technical characteristics

- 1 Highest resolution available, not rescaled or interpolated
- 2 Highest bit depth available, 16 bits per channel if available
- 3 Specified color space as used in published version
- 4 Uncompressed
- 5 Unlayered
- 6 Largest version available, full clarity (e.g. highest resolution, highest bit depth, uncompressed)

##### B Formats (raster), in order of preference

- 1 TIFF (uncompressed)
- 2 JPEG2000 (lossless) (\*.jp2)
- 3 PNG (\*.png)
- 4 JPEG/JFIF (\*.jpg)
- 5 Digital Negative DNG (\*.dng)
- 6 JPEG2000 (lossy) (\*.jp2)
- 7 TIFF (compressed)
- 8 BMP (\*.bmp)
- 9 GIF (\*.gif)-

##### C Formats (vector), in order of preference

- 1 SVG
- 2 AutoCAD Drawing Interchange Format (DXF)
- 3 Encapsulated Postscript (EPS)
- 4 Shape files

##### D Cartographic materials, in order of preference (for geospatial data, see Section VI: Datasets/Databases)

- 1 Most complete data (all layers, appendices), even if proprietary
- 2 GeoTIFF
- 3 GeoPDF
- 4 GeoJPEG2000
- 5 Shapefile

##### E Related Materials

- 1 Include any associated indexes, study guides, or other matter if available.
- 2 Also include annotations, accompanying tabular or textual matter, or other interpretive aids.

##### F Metadata

- 1 If supported by format, include associated descriptive metadata: title, creator, creation date, place of publication, publisher/ producer/ distributor, contact information. Include if available: language of work, other relevant identifiers (e.g., DOI, LCCN, etc.), subject descriptors, abstracts; key or reference to each data field and technical production information, such as EXIF and other embedded metadata.

- 2 Otherwise provide separately in acceptable text or XML-based format.

#### **G Technological Measures**

- 1 Files must contain no measures that control access to or use of the digital work (such as digital rights management or encryption)

#### **Acceptable:**

##### **A Formats (raster)**

- 1 TIFF (in Planar format)
- 2 FlashPix (\*.fpx)
- 3 PhotoShop (\*.psd)
- 4 Camera raw (\*.raw, \*.nef, \*.crw)
- 5 JPEG 2000 Part 2 (\*.jpf, \*.jpx)
- 6 MrSID (\*.sid)

##### **B Formats (vector)**

- 1 Computer Graphic Metafile (CGM, WebCGM) (\*.cgm)
- 2 Macromedia Flash (\*.swf)

##### **C Compression**

- 1 Lower compression ratios preferred
- 2 Discrete wavelet transform (DWT) preferred to discrete cosine transform (DCT)

##### **D Technological Measures**

- 1 Files must contain no measures that control access to or use of the digital work (such as digital rights management or encryption)

#### **E Related Materials**

- 1 Include indexes, study guides, or other printed matter if available.

#### **Acceptable:**

n/a

---

#### **v. Microforms**

#### **Preferred:**

##### **A Permanence and Appearance**

- 1 Silver halide
- 2 Positive polarity
- 3 Color (when available)
- 4 Polyester film base

##### **B Format (newspapers and newspaper-formatted serials)**

- 1 Roll microfilm

##### **C Format (all other materials), in order of preference**

- 1 Microfiche
- 2 Roll microfilm
- 3 Microfilm cassettes
- 4 Micro-opaque prints

##### **D Size**

- 1 35mm, if roll film.

### III Audio works

#### i. Audio – On Tangible Medium (digital and analog)

##### Preferred:

##### A Sound recordings

- 1 Final production/release version of content rather than pre-production version
- 2 Published Compact Disc (CD audio) rather than Recordable Compact Disc (CD-R, audio format)
- 3 With all jackets, sleeves, enclosures and inserts rather than without
- 4 Stereophonic if originally recorded/released as stereophonic
- 5 Monaural if originally recorded/released as monaural
- 6 Direct Stream Digital (DSD) or other multi-channel (e.g. Surround Sound) version in addition to stereophonic version if released in both
- 7 Vinyl disc (LP) in addition to Compact Disc (CD) if released in both

##### Acceptable:

##### A Sound Recordings

- 1 Recordable Compact Disc (CD-R) rather than vinyl disc
- 2 Vinyl disc rather than audio cassette
- 3 Audio cassette if only released as such

#### ii. Audio – Media-independent (digital)

##### Preferred:

##### A Audio

- 1 Final production /release version of content rather than pre-production version
- 2 Highest native resolution PCM WAVE file of final version produced (44.1 kHz / 16 bit or higher) in addition to Compact Disc (CD) when both are produced
- 3 WAVE file with embedded metadata (Broadcast WAVE) rather than without (LC will specify fields)

- 4 File in native resolution rather than up-sampled resolution
- 5 Very high resolution file formats such as DSD, PCM 176.4khz , 192khz up to 384kh when produced for release in addition to Compact Disc (CD) when both are produced
- 6 DSD in the released version (e.g., surround-sound or stereo)

##### B Streaming/Podcasts

- 1 Uncompressed files rather than compressed.
- 2 Compressed version in a major standard compression scheme rather than non-standard scheme

##### C Accompanying Image/Text Files

- 1 With final version of all accompanying image and text files; higher resolution images rather than lower
  - a TIFF or JPEG formats for images
  - b Text files in PDF

##### D Metadata

- 1 Provide most complete metadata set as delivered to online distributors (e.g. iTunes and Amazon), including but not limited to: song/work title, album title, artist, composer, genre, publisher/label name and issue number, location and date of performance, date of publication, standard identifier (e.g. ISRN, UPC) and any other entity identifiers. Provide data in a standard XML-based format

##### E Technological Measures

- 1 Files must contain no measures that control access to or use of the digital work (such as digital rights management or encryption)

##### Acceptable:

##### A Audio

- 1 Uncompressed file of final release version
- 2 Highest resolution compressed version in a major standard compression scheme
- 3 Lossless compression scheme rather than lossy compression scheme

##### B Accompanying Image/Text Files

- 1 TIFF, JPEG or PDF of text files of final release version
- 2 Other standard images and text formats

##### C Metadata

- 1 Provide most complete metadata set as delivered to online distributors (e.g. iTunes and Amazon), including but not limited to: song/work title, album title, artist, composer, genre, publisher/label name and issue number, location and date of performance, date of publication, standard identifier (e.g. ISRN, UPC) and any other

entity identifiers. Provide data in a standard XML-based format

**D Technological measures**

- 1 Files must contain no measures that control access to or use of the digital work (such as digital rights management or encryption)

OLDER VERSION - DO NOT USE

## IV Moving Image Works

### i Motion Pictures – Digital and Physical Media

#### Preferred:

#### A Theatrically Released Films

- 1 Complete final production/release version of motion picture work in the original production resolution, aspect ratio and frame rate
- 2 Theatrical release version in original gauge (e.g., 70mm, 35mm, 16mm)
- 3 Unencrypted interop Digital Cinema Package (DCP), when theatrical release is not distributed as film

#### B Sound

- 1 Complete final tracks, including any foreign language tracks and descriptive audio, when applicable

#### C Descriptive Metadata

- 1 Release title, Release/Production Date, Production Company and/or Producer, Distributor Name, Country of Origin, Language, Duration, relevant unique identifiers applicable to the work (EIDR, ISAN)

#### D Technological Measures

- 1 Files must contain no measures that control access to or use of the digital work (such as digital rights management or encryption)

#### Acceptable:

#### A Theatrically Released Films

- 1 Blu-ray disc, Blu-ray disc recordable

#### B Audio

- 1 Each language and mix for the final production version shall be in its original channel structure and audio resolution as it was delivered to the content distributor

#### C Descriptive Metadata

- 1 Release title, Release/Production Date, Production Company and/or Producer, Distributor Name, Country of Origin, Language, Duration, relevant unique identifiers applicable to the work (EIDR, ISAN)

#### D Technological Measures

- 1 Files must contain no measures that control access to or use of the digital work (such as digital rights management or encryption)

### ii Video – File-Based and Physical Media

#### Preferred:

#### A Video – File-Based

- 1 Final production version of content rather than pre-production version
- 2 Original production resolution and frame rate (i.e. 1080p24; 720p60, etc.)
- 3 Version and file-based format that was delivered to the content distributor

#### B Audio

- 1 Each language and mix for the final production version shall be in its original channel structure and audio resolution as it was delivered to the content distributor.

#### C Descriptive Metadata

- 1 Title (when applicable: Series Title, Season, Episode Title, Episode No., Production No.), Content Type (e.g., Episode, Documentary, Oral History, News, Music Video, etc.), Release/Broadcast/Production Date, Network, Production Company and/or Producer, Distributor Name, Country of Origin, Language, Duration, relevant unique identifiers applicable to the work (EIDR, ISAN)

#### D Technological Measures

- 1 Files must contain no measures that control access to or use of the digital work (such as digital rights management or encryption)

#### Acceptable:

#### A Video – Physical Media

- 1 Final production version of content rather than pre-production version
- 2 Original production resolution and frame rate (i.e. 1080p24; 720p60, etc.)
- 3 Content contained in standard physical media in the following order of preference:
  - a HD: HDCAM-SR, HDCAM, HD-D5, Blu-ray disc, Blu-ray disc recordable

LOC Recommended Format Specifications  
2014-2015

- b SD: Digital Betacam, Betacam SP, DVD, DVD-r

**B Audio**

- 1 When applicable, each language and mix for the final production version shall be in its original channel structure and audio resolution as it was delivered to the content distributor

**C Descriptive Metadata**

- 1 Title (when applicable: Series Title, Season, Episode Title, Episode No., Production No.), Content Type (e.g., Episode, Documentary, Oral History, News, Music, Video, etc.), Release/Broadcast/Production Date, Network, Production Company and/or Producer, Distributor Name, Country of Origin, Language, Duration, relevant unique identifiers applicable to the work (EIDR, ISAN)

OLDER VERSION - DO NOT USE

## V Software and Electronic Gaming and Learning

This category includes software for Desktop, Mobile/Handheld, Gaming and Learning Systems. **Note:** Gaming and Learning Systems are included in the software category, but often require submission of the proprietary platform as well (e.g. Game Console) or a virtual or emulated version of the proprietary platform that runs on a commonly available operating system.

### Preferred:

#### A Required Content

- 1 With documentation and other accompanying material (e.g. instruction materials, errata, addenda, read me files) A copy of the electronic distribution file if the product was sold to the public as a downloadable file. Note: if the software was a part of a book publication (e.g. a software manual with accompanying discs) then a copy of the book must be submitted with the disc.
- 2 Source Code: A copy of the source code from which the software was developed. This may be on physical media or file-based. However, metadata that specifies which compiler was used to create the final code for commercial release—including the version number and build number of the compiler software—must be included. If the compiler is unique to the project or company (i.e. not commercially available), then a copy of compiler software in the specific version and build used to create this version of the software, along with specifications of the platform the compiler ran on, must be included in the submission. This submission should include the rights clearance for the Library to install and use the compiler to read and use the provided information.
- 3 Operating system: If not a commonly available OS, then a copy of the OS must be provided with the submission.  
**Note:** If there are different versions released at the same time for different operating systems (e.g. for Mac and Windows), a submission will be required for each.

- 4 Platform: If a submission requires a stand-alone or proprietary platform (e.g. a gaming system or child's toy), then the platform or a virtual or emulated version of the proprietary platform is required with submission.

#### B Delivery Method

- 1 Mass storage device: All of the requested information may be included as distinct files or may be grouped together using a method such as BagIt, a tar, or AXF object
  - a Hard drive with USB (universal serial bus) interface
  - b CD-ROM disc
  - c Flash drive with USB interface
  - d DVD disc
- 2 Direct File Submission: These submissions would require grouping in a submission package such as BagIt, tar, or AXF object.

#### C Metadata

- 1 Title, creator, creation date, place of publication, publisher/ producer/ distributor, contact information. Include if available: language of work, other relevant identifiers (e.g., UPCDOI, LCCN, etc.), subject descriptors, abstracts; key or reference to each data field
- 2 Production metadata such as credit, rights, and files which are available at the time of production

#### D Technological Measures

- 1 Files must contain no measures that control access to or use of the digital work (such as digital rights management or encryption)

### Acceptable:

n/a

## VI Datasets/Databases

The Library is aware that, in some cases, the provision of datasets and databases for current research uses (including support for the U.S. Congress) may depend upon native formats and associated software, while preservation and long-term access may depend upon data-migration via transport or export formats, with a concomitant risk of loss of precision and accuracy. Given the focus of this document is preservation and long-term access, the following format preferences favor those outcomes.

### i. Datasets

(For Geospatial Data, see Section VI.ii below)

#### Preferred:

##### A Formats

- 1 Platform-independent, character-based formats are preferred over native or binary formats as long as data is complete, and retains full detail and precision. Preferred formats include well-developed, widely adopted, de facto marketplace standards, e.g.
  - a Self-describing, e.g. JSON, XML-based data formats using well known schemas, XML-based data formats accompanied by schema employed
  - b Line-oriented, e.g. TSV, CSV, fixed-width
  - c Platform-independent native formats, e.g. Excel (.xls, .xlsx)

##### B Supplemental Materials

- 1 Consult the appropriate sections of this document to identify the preferred formats for supplementary material

##### C Delivery Media

- 1 Hard drive; CD-ROM; DVD-ROM, etc.

##### D Metadata

- 1 Deposits should include all applicable metadata, data dictionaries, XML schemas, and technical specifications as appropriate. Discipline-specific metadata standards should be used whenever possible
  - a Include: title, creator, creation date, place of publication, publisher/producer/ distributor, contact information

- b Include if available: language of work, other relevant identifiers (e.g., DOI, LCCN, etc.), subject descriptors, abstracts; key or reference to each data field; checksums.
- c For datasets serving as part of a database: proprietary database package and version

##### E Technological Measures

- 1 Files must contain no measures that control access to or use of the digital work (such as digital rights management or encryption)

#### Acceptable:

##### A Formats

- 1 Transportable exports from widely adopted software packages, e.g. MATLAB, MS Access, R, SAS (xport), SPSS (.sav, .por), Stata
- 2 Non-proprietary, publicly documented formats endorsed as standards by a professional community or government agency, e.g. CDF, HDF

##### B Supplemental Materials

- 1 Consult the appropriate sections of the document to identify the preferred formats for supplementary materials

##### C Delivery Media

- 1 Hard drive; CD-ROM; DVD-ROM, etc

##### D Metadata

- 1 Deposits should include all applicable metadata, data dictionaries, XML schemas, and technical specifications as appropriate. Discipline-specific metadata standards should be used whenever possible
  - a Include: title, creator, creation date, place of publication, publisher/producer/ distributor, contact information
  - b Include if available: language of work, other relevant identifiers (e.g., DOI, LCCN, etc.), subject descriptors, abstracts; key or reference to each data field; checksums.
  - c For datasets serving as part of a database: proprietary database package and version

##### E Technological Measures

- 1 Files must contain no measures that control access to or use of the digital work (such as digital rights management or encryption)

## ii. Geospatial Data

### Preferred:

#### A Formats

- 1 Most complete data (all layers, appendices), even if proprietary
- 2 Formats compatible with widely adopted GIS (e.g. ArcGIS)
- 3 Formats compatible with recommendations and tools from geospatial open source and open data communities
  - a Formats developed or endorsed by the Open Geospatial Consortium (OGC)
  - b Formats supported by well supported open source software libraries such as GDAL, OGR and GeoTools

#### B Supplemental Materials

- 1 Consult the appropriate sections of this document to identify the preferred formats for supplementary material

#### C Delivery Media

- 1 Hard drive; CD-ROM; DVD-ROM, etc.

#### D Metadata

- 1 For metadata information see Federal Geographic Data Committee (FGDC):

<http://www.fgdc.gov/metadata/geospatial-metadata-standards>

#### E Technological Measures

- 1 Files must contain no measures that control access to or use of the digital work (such as digital rights management or encryption)

### Acceptable

n/a

## iii. Databases

### Preferred:

#### A Preservation

- 1 Complete set of the content contained within the database, conforming to preferred specifications in Section VI.i-ii

#### B Access

- 1 Publisher web interface, with
  - a Comprehensive and user-friendly search and discovery
  - b Counter compliant usage statistics
- 2 Delivered preservation content

### Acceptable:

n/a