Library of Congress
Recommended Formats Statement
2021-2022

For online version, see Recommended Formats Statement - 2021-2022 (link)

Introduction to the 2021-2022 revision

The success of the Recommended Formats Statement since it was first launched in 2014 and the way in which it has become such an important tool for the community has encouraged the Library to take the opportunity to look more closely at it for this edition of the RFS. The Library has moved beyond the standard review process which it has undertaken annually in order to the keep the RFS current and relevant. Over the course of the past year, the Library has engaged in a more thorough examination, both of the organizational structure of the RFS and of the processes through which each year's version is revised. This root-and-branch analysis has already resulted in significant changes implemented in 2020 including the new layout of the RFS to include categories for Design and 3D; GIS, Geospatial and Non-GIS Cartographic; and Musical Scores.

Underpinning these changes has been the establishment of a new internal model, which the Library has used in this instance to assess the digital file formats in the RFS. This model is based on the conceptual framework of Levels of Service. The Levels of Service concept helps define the degree to which the Library can manage specific formats throughout the lifecycle by considering both global/community criteria and local/institutional criteria. This allows for a more structured and transparent analysis of the file formats and a clear record of that analysis in a matrix workbook with each content category on a separate worksheet.

The global/community criteria for digital file formats have been based on the seven sustainability factors developed for the Library's Sustainability of Digital Formats website: Disclosure, Adoption, Transparency, Self-documentation, External dependencies, Impact of patents and Technical protection mechanisms. Each of these factors may have different emphasis or importance depending on the community of practice and content type. Some may not be applicable or essential for every format.
The local/institutional factors estimate the level of resources at The Library of Congress available to preserve and manage the digital file formats over time. These include Staff experience and expertise, Software/Hardware/Operating System availability (including appropriate number of licenses), Representation/extent in LC collections/storage and Established workflow/functionality.
The use of this evaluation model has enabled the Library to sharpen and focus its analysis of the digital file formats in the Recommended Formats Statement. In providing a consistent review structure across all content categories, it now serves as a means to document improvements over the years as well as identify gaps that need to be filled.

Overall, the analysis has allowed us to establish clearer definitions of 'Preferred' and 'Acceptable' when categorizing digital file formats in the RFS:

Preferred formats:

- Global/community: Meets or exceeds benchmarks for all relevant sustainability factors
- Local/institutional: The Library of Congress has the skills, experience, workflows, tools and systems to manage and preserve these formats in current systems with confidence.

Acceptable formats:

- Global/community: Meets minimum acceptability across benchmarks or does not meet all relevant sustainability factors.
- Local/institutional: The Library of Congress can manage this format at a basic level of acquisition, management and preservation; and a greater ability for management and preservation is within the Library's capacity with further investment.

The success in using this model in evaluating and assessing the digital file formats in the Recommended Formats Statement opens the possibility of adapting it to apply to those other characteristics of creative works, both physical and digital, which the RFS covers in its remit to address all types of creative works. The Recommended Formats Statement is not intended to serve as an answer to all the questions raised in preserving and providing long-term access to creative content. It does not provide instructions for receiving 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. The RFS provides 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. The Library hopes that the RFS will help make it realistic to build, grow and save creative output for our individual and collective benefit for generations to come.

The Library of Congress, realizing its unique position, is pleased to be able to contribute a resource like the Recommended Formats Statement for the benefit of all involved with creative works. The commitment of time and resources to the ongoing revision and indeed improvement of the RFS reflects the priority the Library places on working collaboratively to ensure that all might succeed in our common goal to share and disseminate creative output and to benefit the nation and the world at large.

.

I. Textual Works

NOTE: See also Musical Scores

i. Textual Works – Print (books, etc.)		
	Preferred	**Acceptable**
A. Paper	1. Archival quality paper (ISO 11108: 1996 for Archival Paper)	
B. Printing Process, in 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 i. Library binding (NISO Z39.78-2000) ii. Sewn iii. Glued only b. Soft cover i. Sewn ii. Glued only iii. Spiral- or plastic-bound iv. 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	

i. Textual Works – Print (books, etc.)		
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.	
G. Metadata	1. As displayed on item: a. Title b. Creator c. Creation Date or Start Date/End Date d. Place of Publication e. Publisher/Producer/Distributor f. ISBN 2. As displayed on item, if available: a. Other relevant identifiers (e.g., DOI, LCCN, etc.) b. Edition c. Subject descriptors d. Abstracts	

ii. Textual Works – Digital		
	Preferred	**Acceptable**
A. Technical Characteristics, in order of preference	Character encoding, in descending order of preference: 1. UTF-8, UTF-16 (with BOM), US-ASCII 2. ISO 8859	Other character encodings not listed in Preferred section
B. Formats, in order of preference	XML-based markup formats, with included or accessible DTD/schema, XSD/XSL presentation stylesheet(s), and explicitly stated character encoding a. EPUB3-compliant. (Other versions of EPUB are also preferred formats but EPUB3 is the most common.) b. BITS-compliant (NLM Book DTD) c. Other widely-used book DTDs/schemas (e.g., TEI, DocBook, etc.) 2. Page-layout formats a. PDF/UA (ISO 14289-1-compliant) b. PDF/A (ISO 19005-compliant) 4. 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)	1. Other structured or markup formats a. XHTML or HTML, with DOCTYPE declaration and presentation stylesheet(s) b. XML-based document formats (widely-used and publicly-documented), with presentation stylesheet(s) if applicable. Includes DOCX/OOXML 2012 (ISO 29500), ODF (ISO/IEC 26300) and OOXML (ISO/IEC 29500). c. SGML, with included or accessible DTD d. Other XML-based non-proprietary formats, with presentation stylesheet(s) e. XML-based formats that use proprietary DTDs or schemas, with presentation stylesheet(s) 2. Page-layout formats a. PDF (web-optimized) 3. Other formats a. Rich text format (RTF) b. Plain text c. Widely-used proprietary word-processing formats
C. Rarity and Special Features	Limited editions (including those with special features such as high resolution images) Editions with the greatest number of unique features (such as additional content, multimedia, interactive elements, etc.)	

ii. Textual Works – Digital		
D. Completeness	• 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. 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.	
E. Metadata	1. As supported by format (e.g., standards-based formats such as ONIX for Books, XMP, MODS, or MARCXML either embedded in or accompanying the digital item): a. Title b. Creator c. Creation Date or Start Date/End Date d. Place of publication e. Publisher/ producer/ distributor f. ISBN g. Contact information 2. Include if available: a. Language of work b. Other relevant identifiers (e.g., DOI, LCCN, original URL, etc.) c. Edition d. Subject descriptors e. Abstracts	
F. Technological Measures	Files must contain no measures (such as digital rights management technologies or encryption) that control access to or prevent use of the digital work.	

iii. Textual Works – Electronic serials		
	Preferred	**Acceptable**
A. Technical Characteristics, in order of preference	1. Character encoding, in descending order of preference: a. UTF-8, UTF-16 (with BOM), US-ASCII b. ISO 8859	1. Other character encodings not listed in Preferred section
B. Formats, in order of preference	1. Content compliant with the NISO JATS: Journal Article Tag Suite (ANSI/NISO Z39.96-2015) with XSD/XSL presentation stylesheet(s) and explicitly stated character encoding 2. Page-layout formats a. PDF/UA (ISO 14289-1-compliant) b. PDF/A (ISO 19005-compliant) c. 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)	1. Other structured or markup formats: a. Widely-used serials or journal non-proprietary XML-based DTDs/schemas with included or accessible DTD/schema, presentation stylesheet(s) and explicitly stated character encoding. b. Proprietary XML-based format for serials or journals (with documentation) with DTD/schema and presentation stylesheet(s) c. XHTML or HTML, with DOCTYPE declaration and presentation stylesheet(s) d. XML-based document formats (widely used and publicly documented). With presentation stylesheets, if applicable. Includes DOCX/OOXML 2012 (ISO 29500), ODF (ISO/IEC 26300) and OOXML (ISO/IEC 29500). 2. Page-layout formats a. PDF (web-optimized with searchable text) 3. Other formats a. Rich text format b. Plain text

iii. Textual Works – Electronic serials		
		c. Widely-used proprietary word processing or page-layout formats d. Other text- or graphic-based formats not listed here that represent textual works
C. Completeness	• 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. Includes all associated external files and fonts considered integral to the publication. • 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.	
D. Metadata	1. Title-level metadata (e.g., standards-based formats such as ONIX for Books, XMP, MODS, or MARCXML either embedded in or accompanying the digital item): a. Serial or journal title b. ISSN and ISSN-L c. Publisher d. Frequency e. Place of publication 2. Article-level metadata as relevant or applicable (e.g., standards-based formats such as ONIX for Books, XMP, MODS, or MARCXML either embedded in or accompanying the digital item): a. Volume(s) b. Number(s) c. Issue date(s) d. Article title(s)	

iii. Textual Works – Electronic serials		
	e. Article author(s) f. Article identifier (DOI, original URL, etc.) 3. Include if available: a. Other descriptive metadata (e.g., subject heading(s), descriptor(s), abstract(s))	
E. Technological Measures	• Files must contain no measures (such as digital rights management technologies or encryption) that control access to or prevent use of the digital work.	

II. Still Image Works

i. Photographs – Print		
	Preferred	**Acceptable**
A. Faithful representation of the work	Equal in quality to the publication version, best edition or master copy	
B. Permanence and appearance	• Unmounted • Pigmented inks (if digitally printed) • Fixed, well-washed (if wet chemistry method)	
C. Size	• Min: 8 x10” • Max: 28 x 36”	Larger sizes may be acceptable if best or only version.
D. Metadata	1. As supported by format: a. Title b. Creator c. Creation Date d. Place of publication e. Publisher/producer/distributor f. Contact information 2. Include if available: a. Language of work b. Other relevant identifiers (e.g., DOI, LCCN, etc.) c. Subject descriptors d. Abstracts e. Key or reference to each data field and technical production information (type of paper, how processed, publisher internal tracking numbers)	

ii. Photographs – Digital		
	Preferred	**Acceptable**
A. Faithful representation of the work	• Equal in quality to the published version, best edition or master copy • In the same format as the master copy	
B. Technical Characteristics	• Highest resolution available, not rescaled or interpolated • Highest bit depth available, 16 bits per channel if available • Embedded color profile or specified color space used in published version • Uncompressed • Unlayered	• Lossless compression or lower compression ratios • Discrete wavelet transform (DWT) preferred to discrete cosine transform (DCT) • Layered, if supported by preferred or acceptable format
C. Formats	• TIFF (*.tif) • JPEG2000 (*.jp2) • PNG (*.png) • JPEG/JFIF (*.jpg) • BMP (*.bmp)	• Photoshop (*.psd, *.psb) • JPEG2000 Part 2 (*.jpf, *.jpx) • Digital Negative DNG (*.dng) • Proprietary Camera Raw formats (*.nef, *.crw) • GIF (*.gif)
D. Metadata	1. As supported by format: a. Title b. Creator c. Creation Date d. Place of publication e. Publisher/producer/distributor f. Contact information 2. Include if available: a. Common embedded schema (e.g., IPTC) b. Language of work c. Other relevant identifiers (e.g., PLUS ID, DOI, LCCN, etc.) d. Subject descriptors e. Abstracts f. Key or reference to each data field and technical production information (e.g. EXIF metadata from digital camera)	Metadata provided separately in external text of XML-based file

ii. Photographs – Digital		
E. Technological Measures	• Files must contain no measures (such as digital rights management technologies or encryption) that control access to or prevent use of the digital work.	

iii. Other Graphic Images – Print (posters, postcards, fine prints)		
NOTE: See also Geospatial Cartographic and Design and 3D		
	Preferred	**Acceptable**
A. Faithful representation of the work	• Equal in quality to the publication version, best edition or master copy	
B. Permanence and appearance	• Packaging materials equivalent to published form (e.g., binding, box/packaging materials) • If multiple versions available, provide the most widely distributed edition. • If limited edition, provide an unnumbered but otherwise identical copy. • For large items, provide rolled, unfolded.	
C. Related Materials	• Includes indexes, study guides or other matter if available • Also includes annotations, accompanying tabular or textual matter or other interpretive aids	
D. Metadata	• As supported by format a. Title b. Creator c. Creation Date d. Place of Publication e. Publisher/producer/distributor f. Contact Information • Include if available: a. Language of work b. Other relevant identifiers (e.g., DOI, LCCN, etc.) c. Subject descriptors d. Abstracts	

iii. Other Graphic Images – Print (posters, postcards, fine prints)		
	e. Key or reference to each data field and technical production information (type of paper, how processed, publisher internal tracking numbers)	

iv. Other Graphic Images – Digital		
NOTE: See also Geospatial Cartographic and Design and 3D		
	Preferred	**Acceptable**
A. Faithful representation of the work	• Equal in quality to the published version, best edition or master copy • In the same format as the master copy	
B. Technical Characteristics	• Highest resolution available, not rescaled or interpolated • Highest bit depth available, 16 bits per channel if available • Specified color space used in published version • Uncompressed • Unlayered	• Lower compression ratios • Discrete wavelet transform (DWT) preferred to discrete cosine transform (DCT) • Layered, if supported by preferred or acceptable format
C. Formats (raster)	• TIFF (*.tif) • JPEG2000 (*.jp2) • PNG (*.png) • JPEG/JFIF (*.jpg) • BMP (*.bmp)	• Photoshop (*.psd, *.psb) • JPEG2000 Part 2 (*.jpf, *.jpx) • MrSID (*.sid) • Encapsulated Postscript (*.eps) • Digital Negative DNG (*.dng) • Proprietary Camera Raw formats • GIF (*.gif)

iv. Other Graphic Images – Digital		
D. Formats (vector)	• Scalable vector graphics (*.svg)	• Computer Graphics Metafile (CGM, WebCGM) • Page-layout formats, e.g. PDF/UA (ISO 14289-1-compliant), PDF/A (ISO 19005-compliant), PDF (highest quality available, with features such as searchable text, embedded fonts, lossless compression, high resolution images; includes document formats such as PDF/X) • Encapsulated Postscript (*.eps)
E. Related Materials	• Includes indexes, study guides or other matter if available • Also includes annotations, accompanying tabular or textual matter or other interpretive aids	
G. Metadata	1. As supported by format: a. Title b. Creator c. Creation Date d. Place of publication e. Publisher/producer/distributor f. Contact information 2. Include if available: a. Common embedded schema b. Language of work c. Other relevant identifiers (e.g., DOI, LCCN, etc.) d. Subject descriptors e. Abstracts f. Key or reference to each data field and technical production information (e.g. EXIF metadata from digital camera)	• Metadata provided separately in external text or XML-based file
H. Technological Measures	• Files must contain no measures (such as digital rights management technologies or encryption) that control access to or prevent use of the digital work.	

v. Microforms		
	Preferred	**Acceptable**
A. Faithful representation of the work	Equal in quality to the publication version, best edition or master copy	
B. Permanence and appearance	• Silver halide • Positive polarity • Color (when available) • Polyester film base	
C. Format (newspapers and newspaper-formatted serials)	Roll microfilm	
D. Format (all other materials), in order of preference	1. Microfiche 2. Roll microfilm 3. Microfilm cassettes 4. Micro-opaque prints	
E. Size	35mm, if roll film	16mm film and other sizes that match the primary production master
F. Related Materials	Include indexes, study guides or other printed matter if available	
G. Metadata	• As supported by format a. Title b. Creator c. Creation Date d. Place of Publication e. Publisher/producer/distributor f. Contact Information • Include if available: a. Language of work b. Other relevant identifiers (e.g., DOI, LCCN, etc.) c. Subject descriptors d. Abstracts e. Key or reference to each data field and technical production information (type of paper, how processed, publisher internal tracking numbers)	

III. Moving Image Works

i. Motion Pictures – Digital and Physical Media		
	Preferred	**Acceptable**
A. Motion Pictures - Digital And Physical Media	• Complete final production/release version of motion picture work in the original production resolution, aspect ratio and frame rate • Theatrical release version in original gauge (e.g., 70mm, 35mm, 16mm) • Unencrypted interop Digital Cinema Package (DCP) with the following characteristics: a. 24- or 48-frame progressive scan b. Minimum projector resolution of 2048 by 1080 pixels c. Image source compression (if used) conforming to ISO/IEC 15444-1 (JPEG2000) d. Image and sound files packaged as either SMPTE or Interop DCPs e. DCP formats (SMPTE ST429-2 and related specifications) Contact archive for guidance regarding master materials (DCDM, DSM, camera original negatives, etc.)	• Commercially pressed DVD or Blu-ray disc
B. Audio	• Complete final tracks, including any foreign language tracks and descriptive audio, 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. Metadata	1. Relevant unique identifiers applicable to the work (EIDR, ISAN) 2. If unique identifier not available, then a) Release title b) Release/Production Date c) Production Company and/or Producer d) Distributor Name e) Country of Origin f) Language g) Duration	

i. Motion Pictures – Digital and Physical Media		
D. Technological Measures	• Files must contain no measures (such as digital rights management technologies or encryption) that control access to or prevent use of the digital work.	

ii. Video – File Based and Physical Media		
	Preferred	**Acceptable**
A. Video – File-based, in order of preference	Final production version with the original production resolution and frame rate (i.e. 1080p24; 720p60, etc.) and file-based format that was delivered to the content distributor. 1. Interoperable Master Format (IMF) consisting of a. Essence files as MXF tracks including video, audio, data and dynamic metadata essences b. Composition playlist c. Packaging data XML files (asset map, packing list, volume index) 2. ProRes a. QuickTime (.mov) container b. 4444 (XQ), 4444 or 422 HQ codecs 3. MPEG-2 a. Compliant with ISO/IEC 13818 4. XDCAM a. MXF b. HD422, SHD422, HD codecs Contact archive for guidance regarding pre-production versions.	FFV1 (version 3) in Matroska (.mkv) container only for content without closed captions and/or timecode information. Viewing proxy such as a) Recordable DVD b) Recordable Blu-ray disc c) MPEG-4 (.mp4)

ii. Video – File Based and Physical Media		
B. Video – Physical Media, in order of preference		1. Complete, final production version with the original production resolution and frame rate (i.e. 1080p24; 720p60, etc.) 2. Content contained in standard physical media in the following order of preference: a. HD: HDCAM-SR, HDCAM, HD-D5, Commercially pressed DVD or Blu-ray disc b. SD: Digital Betacam, Betacam SP
C. Audio		• 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
D. Metadata	1. Relevant unique identifiers applicable to the work (EIDR, ISAN) 2. If unique identifier not available, then: a) Release title b) Release/Production Date c) Production Company and/or Producer d) Distributor Name e) Country of Origin f) Language g) Duration	
E. Technological Measures	• Files must contain no measures (such as digital rights management technologies or encryption) that control access to or prevent use of the digital work.	

IV. Audio Works

i. Audio – On Tangible Media (digital and analog)		
	Preferred	**Acceptable**
1. Sound Recordings, in order of preference	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	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	**Acceptable**
A. Audio, in order of preference	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) 7. Uncompressed files rather than compressed. 8. Compressed version in a major standard compression scheme rather than non-standard scheme	1. Highest resolution compressed version in a major standard compression scheme 2. Lossless compression scheme rather than lossy compression scheme
B. Accompanying Image/Text Files, in order of preference	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	1.
C. Metadata	1. Provide most complete metadata set as delivered to online distributors (e.g. iTunes and Amazon), which may include elements not embedded in a file, including but not limited to: a. Song/work title b. Album title c. Artist d. Composer e. Genre f. Publisher/label name and issue number g. Location and date of performance	

ii. Audio – Media-independent (digital)		
	h. Date of publication i. Standard identifier (e.g. ISRC, UPC) j. Any other entity identifiers 2. Provide data in a standard XML-based format, such as the Electronic Release Notification (ERN-DDEX) 3. RSS feeds are desirable for podcasts	
D. Technological Measures	• Files must contain no measures (such as digital rights management technologies or encryption) that control access to or prevent use of the digital work.	

V. Musical Scores

NOTE: See also Textual Works - Print

i. Musical Scores – Print		
	Preferred	**Acceptable**
A. Paper	• Archival quality paper (ISO 11108: 1996 for Archival Paper)	
B. Printing Process, in order of preference	• Lithography (offset printing press) • Electrophotography (digital press) • Inkjet (inkjet printer using stable pigment or dye-based inks)	
C. Binding and Packaging	• Slip-cased, if available • Binding, in descending order of preference: a. Hard cover i. Library binding (NISO Z39.78-2000) ii. Sewn iii. Glued only b. Soft cover i. Sewn ii. Glued only iii. Spiral- or plastic-bound iv. Stapled • Loose-leaf (including all binders and indexes published as part of the deposit and offered for sale and distribution)	
D. Size	• Larger-sized editions (Note: large-type editions are not preferred over editions with conventional size typefaces) • For broadsides and musical compositions, the Library prefers items: c. In protective folders d. Rolled (rather folded)	
E. Rarity, Special Features, Illustrations	• Limited editions (including those with special binding or special features)	

i. Musical Scores – Print		
	• Editions with the greatest number of unique features (such as pop-ups, overlaps, magnifiers, overlays, tabs, notches, etc.) Illustrated editions; original color illustrations preferred over black and white reproductions	
F. Completeness	• 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. • 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. • For unaccompanied vocal musical compositions: open score, with each part on separate staff • 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	
G. Metadata	1. As displayed on item: e. Title f. Creator g. Creation Date or Start Date/End Date h. Place of Publication i. Publisher/Producer/Distributor j. ISBN 2. As displayed on item, if available: a. Other relevant identifiers (e.g., DOI, LCCN, etc.)	

i. Musical Scores – Print		
	b. Edition c. Subject descriptors d. Abstracts	

ii. Musical Scores – Digital	**Preferred**	**Acceptable**
H. Technical Characteristics, in order of preference	1. Character Encoding, in descending order of preference: a. UTF-8, UTF-16 (with BOM), US-ASCII b. ISO 8859	Other character encodings not listed in Preferred section
I. Formats, in order of preference	1. XML-based markup music notational format, with included or accessible DTD/schema, XSD/XSL presentation stylesheet(s), and explicitly stated character encoding. a. MusicXML b. Music Encoding Initiative (MEI) c. Other widely-used and publicly documented musical notation DTDs/schemas 2. Page-layout formats a. PDF/UA (ISO 14289-1-compliant) b. PDF/A (ISO 19005-compliant) c. PDF (highest quality available, with features such as searchable text, embedded fonts, lossless compression, high resolution images; includes document formats such as PDF/X)	1. Other structured or markup formats a. XHTML or HTML, with DOCTYPE declaration and presentation stylesheet(s) b. SGML, with included or accessible DTD 2. Page-layout formats a. PDF (web-optimized) 3. Other formats a. Widely-used proprietary music notation formats b. Other music composition formats (including graphics-based formats) not listed here
J. Rarity and Special Features	• Limited editions (including those with special features) • Editions with the greatest number of unique features (such as additional content, multimedia, interactive elements, etc.)	
K. Completeness	• 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. • All updates, supplements, releases, and supersessions published as part of the work and offered for sale or	

ii. Musical Scores – Digital		
	distribution must be submitted and received in a regular and timely manner for proper maintenance of the deposit. • For unaccompanied vocal musical compositions: open score, with each part on separate staff • For vocal musical compositions with orchestral accompaniment and for instrumental musical compositions: a. Full score and all parts, if applicable; if published only by rental, lease, or lending, full score only may be submitted b. Conductor's score and all parts, if applicable; if published only by rental, lease, or lending, conductor's score only may be submitted	
L. Metadata	1. As supported by format (e.g., standards-based formats such as ONIX for Books, XMP, MODS, or MARCXML either embedded in or accompanying the digital item): a. Title b. Creator c. Creation Date or Start Date/End Date d. Place of publication e. Publisher/ producer/ distributor f. ISMN g. Instrumentation 2. Include if available: a. Language of work b. Other relevant identifiers (e.g., ISBN, DOI, LCCN, original URL, etc.) c. Edition d. Subject descriptors e. Event f. Abstracts	
M. Technological Measures	Files must contain no measures (such as digital rights management technologies or encryption) that control access to or prevent use of the digital work.	

VI. Datasets

NOTE: See also Geospatial and Cartographic

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		
	Preferred	**Acceptable**
i. Formats	• 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. Formats using well known schemas with public validation tools available b. Line-oriented, e.g. TSV, CSV, fixed-width c. Platform-independent open formats, e.g. .db, .db3 • Any proprietary format that is a de facto standard for a profession or supported by multiple tools (e.g. Excel .xls or .xlsx, Shapefile) • Character Encoding, in descending order of preference: a. UTF-8, UTF-16 (with BOM) b. US-ASCII or ISO 8859-1 c. Other named encoding	In order of preference: 1. Non-proprietary, publicly documented formats endorsed as standards by a professional community or government agency, e.g. CDF, HDF 2. Text-based data formats with available schema
ii. Related Materials	• Consult the appropriate sections of this document to identify the preferred formats for supplementary material	
iii. Delivery Method, in order of preference	1. Public download URLs 2. Automated private download URLs with any necessary API keys or credentials	

i. Datasets		
	3. Hard drive; CD-ROM; DVD-ROM	
iv. 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 2. As supported by format: a. Title b. Creator c. Creation date d. Place of publication e. Publisher/ producer/ distributor f. Contact information g. A list of software used to produce, render or compress the data (if applicable) h. Character encoding 3. Include if available: a. Language of work b. Other relevant identifiers (e.g., DOI, LCCN, canonical URL, etc.) c. Subject descriptors d. Abstracts e. Key or reference to each data field f. Checksums g. Permanent version specifiers h. Information about how the data was collected and any sampling or post-processing which has been applied i. Known copyright terms, especially for datasets which combine data from multiple sources 4. For datasets serving as part of a database: proprietary database package and version	
v. Technological Measures	• Files must contain no measures (such as digital rights management technologies or encryption) that control access to or prevent use of the digital work.	Files in formats which support executable code do not depend on embedded programs for purposes other than display

i. Datasets		
	• Files in formats which support linking or embedding external resources (e.g. XML, JSON, Excel) should be self-contained to remain useful in the event of external service changes. • Files in formats which support executable code (e.g. Excel) do not contain executable code.	(e.g. search, filtering, etc.); the raw data is available without executing code.

ii. Databases		
	Preferred	**Acceptable**
A. Preservation	• Complete set of the content contained within the database, conforming to preferred specifications in sec. VI.i-ii	
B. Access, in order of preference	1. Publisher web interface with a. Comprehensive and user-friendly search and discovery b. Counter compliant usage statistics 2. Delivered preservation content	Documented API

VII. GIS, Geospatial and Non-GIS Cartographic

i. Geographic Information System (GIS) – Vector Data		
	Preferred	**Acceptable**
A. Formats	Most complete data (all layers, appendices), even if proprietary, with a preference for preserving the native format and projection of the data Vector formats compatible with widely adopted GIS including, • Shapefile, which is comprised of at least a SHP, SHX, and DBF file and optionally a PRJ (highly recommended), XML (highly recommended), SBN, and/or SBX. • Esri File Geodatabase • OGC GeoPackage	1. GeoJSON (may have scalability issues) 2. KML 3. GML
B. Delivery Method, in order of preference	• Public download URLs • Automated private download URLs with any necessary API keys or credentials • Hard drives	
C. Metadata	• For metadata information see 191xx ISO standards and Federal Geographic Data Committee (FGDC) • To the extent allowed by the underlying format, include available information about how the data was collected and any post-processing which has been applied	Project and layer files (.mxd, .qgs, .lry) may be acquired with deposits of content data to assist in reviewing materials during deposit process
D. Technological Measures	• Files must contain no measures (such as digital rights management technologies or encryption) that control access to or prevent use of the digital work.	

ii. GIS Vector and Raster Combined		
	Preferred	**Acceptable**
A. Formats	In order of preference:	• TerraGo GeoPDF • Geospatial PDF

ii. GIS Vector and Raster Combined		
	1. Most complete data (all layers, appendices), even if proprietary, with a preference for preserving the native format and projection of the data 2. Vector and raster formats compatible with widely adopted GIS including: a) Esri File Geodatabase b) OGC GeoPackage c) Formats compatible with recommendations and tools from geospatial open source and open data communities, formats developed or endorsed by the Open Geospatial Consortium (OGC), formats supported by well supported open source software libraries such as GDAL, OGR and GeoTools	
B. Delivery Method, in order of preference	1. Public download URLs 2. Automated private download URLs with any necessary API keys or credentials 3. Hard drives	
C. Metadata	• For metadata information see 191xx ISO standards and Federal Geographic Data Committee (FGDC) • To the extent allowed by the underlying format, include available information about how the data was collected and any post-processing which has been applied	
D. Technological Measures	• Files must contain no measures (such as digital rights management technologies or encryption) that control access to or prevent use of the digital work.	

iii. GIS Raster and Georeferenced Images		
	Preferred	**Acceptable**
A. Formats	• Most complete data (all layers, appendices), even if proprietary, with a preference for preserving the native format and projection of the data • Raster formats compatible with widely adopted GIS including GeoTIFF • OGC GeoPackage	• TIFF (.tif) files with accompanying TIFF World File (.tfw and .tifw) • GML in JPEG 2000
B. Delivery Method, in order of preference	1. Public download URLs 2. Automated private download URLs with any necessary API keys or credentials 3. Hard drives	
C. Metadata	• For metadata information see 191xx ISO standards and Federal Geographic Data Committee (FGDC) • To the extent allowed by the underlying format, include available information about how the data was collected and any post-processing which has been applied	
D. Technological Measures	• Files must contain no measures (such as digital rights management technologies or encryption) that control access to or prevent use of the digital work.	

iv. Non-GIS Cartographic		
	Preferred	**Acceptable**
A. Cartographic materials, in order of preference	1. Most complete data (including appendices) with a preference for preserving the native format and projection of the data 2. Largest size 3. Most widely distributed 4. Follows recommended formats in Still Image Works (if material is an image) or Textual Works (if material is primarily textual).	

iv. Non-GIS Cartographic		
B. Faithful representation of the work	Equal in quality to the publication version, best edition or master copy	
C. Permanence and Appearance	• Packaging materials equivalent to published form (e.g., binding, box/packaging materials) • If multiple versions available, provide the most widely distributed edition. • If limited edition, provide an unnumbered but otherwise identical copy. • For large items, provide rolled, unfolded. • Prefer edition with protective coatings. • Faithful representation of the work as published. Equal quality to publication version.	
D. Related Materials	• Includes indexes, study guides or other matter if available • Also includes annotations, accompanying tabular or textual matter or other interpretive aids	
E. Metadata	• As supported by format a. Title b. Creator c. Creation Date d. Place of Publication e. Publisher/producer/distributor f. Contact Information g. • Include if available: a. Language of work b. Other relevant identifiers (e.g., DOI, LCCN, etc.) c. Subject descriptors d. Abstracts e. Key or reference to each data field and technical production information (e.g. EXIF metadata from digital camera	

VIII. Design and 3D

NOTE: See also Still Image Works

i. 2D and 3D Computer Aided Design		
	Preferred	**Acceptable**
A. Formats (raster)	• TIFF (*.tif) • JPEG2000 (*.jp2) • PNG (*.png) • JPEG/JFIF (*.jpg) • Digital Negative DNG (*.dng) • BMP (*.bmp) • GIF (*.gif)	• Photoshop (*.psd, *.psb) • JPEG2000 Part 2 (*.jpf, *.jpx) • Encapsulated Postscript (*.eps)
B. Formats (vector)	• Scalable vector graphics (*.svg) • AutoCAD Drawing Interchange Format (*.dxf) • Shapefile	• Computer Graphics Metafile (CGM, WebCGM) • Extensible 3D (X3D) • Non-proprietary formats endorsed as standards by a professional community or government agency, e.g. IFC, STEP • Page-layout formats, e.g. PDF/UA (ISO 14289-1-compliant), PDF/A (ISO 19005-compliant), PDF (highest quality available, with features such as searchable text, embedded fonts, lossless compression, high resolution images; includes document formats such as PDF/X) • Page-layout formats, e.g. PDF/UA (ISO 14289-1-compliant), PDF/A (ISO 19005-compliant), PDF (highest quality available, with features such as searchable text, embedded fonts, lossless compression, high resolution images, device-independent specification of colorspace, content

i. 2D and 3D Computer Aided Design		
		tagging; includes document formats such as PDF/X) • Encapsulated Postscript (*.eps) • Proprietary vector formats
C. Technical Characteristics	• Highest resolution available, not rescaled or interpolated • Highest bit depth available, 16 bits per channel if available • Embedded color profile or specified color space used in published version • Uncompressed • Unlayered	• Lossless compression or lower compression ratios • Discrete wavelet transform (DWT) preferred to discrete cosine transform (DCT) • Layered, if supported by preferred or acceptable format
D. Related Materials	• Includes indexes, study guides or other matter if available • Also includes annotations, accompanying tabular or textual matter or other interpretive aids	
E. Delivery Method, in order of preference	• Public download URLs • Automated private download URLs with any necessary API keys or credentials • Hard drives	
F. Metadata	3. As supported by format: g. Title h. Creator i. Creation Date j. Place of publication k. Publisher/producer/distributor l. Contact information 4. Include if available: g. Common embedded schema (e.g., IPTC, FGDC, ISO 19115) h. Language of work i. Other relevant identifiers (e.g., PLUS ID, DOI, LCCN, etc.) j. Subject descriptors k. Abstracts l. Key or reference to each data field and technical production information (e.g. EXIF metadata from digital camera)	• Metadata provided separately in external text of XML-based file

i. 2D and 3D Computer Aided Design		
G. Technological Measures	• Files must contain no measures (such as digital rights management technologies or encryption) that control access to or prevent use of the digital work.	

ii. Design (schematics, architectural drawings) - Print		
NOTE: See also Still Image Works		
	Preferred	**Acceptable**
A. Faithful representation of the work	• Equal in quality to the publication version, best edition or master copy	
B. Permanence and appearance	• Packaging materials equivalent to published form (e.g., binding, box/packaging materials) • If multiple versions available, provide the most widely distributed edition. • If limited edition, provide an unnumbered but otherwise identical copy. • For large items, provide rolled, unfolded. • Prefer edition with protective coatings. • Faithful representation of the work as published. Equal quality to publication version.	
C. Related Materials	• Includes indexes, study guides or other matter if available • Also includes annotations, accompanying tabular or textual matter or other interpretive aids	
D. Metadata	• As supported by format a. Title b. Creator c. Creation Date d. Place of Publication e. Publisher/producer/distributor f. Contact Information • Include if available: f. Language of work g. Other relevant identifiers (e.g., DOI, LCCN, etc.) h. Subject descriptors	

ii. Design (schematics, architectural drawings) - Print		
	i. Abstracts j. Key or reference to each data field and technical production information (type of paper, how processed, publisher internal tracking numbers)	

iii. Scanned 3D Objects (output from photogrammetry scanning)		
	Preferred	**Acceptable**
A. Formats		• STL (STereoLithography) • Reflectance Transformation Imaging (RTI) • Polygon File Format (PLY) • Wavefront (OBJ)
B. Related Materials	• Includes indexes, study guides or other matter if available • Also includes annotations, accompanying tabular or textual matter or other interpretive aids	
C. Metadata	• As supported by format a. Title b. Creator c. Creation Date d. Place of Publication e. Publisher/producer/distributor f. Contact Information • Include if available: k. Language of work l. Other relevant identifiers (e.g., DOI, LCCN, etc.) m. Subject descriptors n. Abstracts o. Key or reference to each data field and technical production information (type of paper, how processed, publisher internal tracking numbers)	
D. Technological Measures	• Files must contain no measures (such as digital rights management technologies or encryption) that control access to or prevent use of the digital work.	

IX. Software and Video Games

This category includes software for Desktop, Mobile/Handheld and Video Gaming Systems. Note: Gaming Systems 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.

i. Software and Video Games		
	Preferred	**Acceptable**
A. Content, in order of preference	1. Uncompiled Source Code: A file-based copy of the source code from which the software was developed. 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. 2. Gold master build (specific file types will vary depending on company producing build) 3. Distribution file (e.g. ipa [Mac iOS], apk [Android], exe [Windows]) 4. Disk-based physical release of game/software content	Hard drive/flash drive/writable disk of game/software content
B. Operating system	• Operating system: If not a commonly available OS, then a copy of the OS, submitted as a disk image, must be provided with the submission. • **NOTE**: If there are different versions released at the same time for different operating system (e.g. for Mac, Windows and Xbox One), a submission will be required for each.	
C. Platform	If a submission requires a stand-alone or proprietary platform (e.g. a gaming system or child's toy), then a	If a submission requires a stand-alone or proprietary platform (e.g. a gaming system or child's toy), and a virtual or emulated

i. Software and Video Games		
	virtual or emulated version of the proprietary platform is required with submission.	version of the proprietary platform is not available, a physical platform including controllers and power supply is acceptable.
D. Related materials	• 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.	
E. Delivery method	• Direct File Submission: These submissions would require grouping in a submission package such as BagIt, Tape Archive (tar), or AXF object.	• 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, Tape Archive (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
F. Metadata	1. 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. 2. As supported by format: a. Title b. Creator c. Creation date d. Place of publication e. Publisher/producer/distributor f. Contact information g. Production metadata such as credit, rights and files which are available at the time of production 3. Include if available: a. Language of work	

i. Software and Video Games		
	b. Other relevant identifiers (e.g., UPCDOI, LCCN, etc.) c. Subject descriptors d. Abstracts e. Schema	
G. Technological Measures	• Submissions of uncompiled source code must include the rights clearance for the Library to install and use the compiler to read and use the provided information. • Files must contain no measures (such as digital rights management technologies or encryption) that control access to or prevent use of the digital work.)	• A file containing digital rights management technologies or encryption which requires a stand-alone or proprietary platform is accompanied by a virtual or emulated version of the proprietary platform or a physical version of the platform. However, the file cannot include a connection to a remote server for DRM or encryption services.

X. Web Archives

This format specification covers the Library's preferred format for archived web content, as well as a preferred "format" for presentation of web content for archiving (in other words, best practices for content creators to help in creating preservation-friendly websites). The Library is aware that websites, including blogs, social media and other web content that make up websites, are presented and created in formats for viewing in a web browser, and are often different than the standard format that is recommended for preservation and long-term access. Given that the focus of this document is preservation and long-term access, the following format preferences favor those outcomes, and include recommendations for best practices to better enable preservation of web content.

i. Websites		
	Preferred	**Acceptable**
A. Technical Characteristics	1. Website creators can improve the archivability of web content by following best practices such as: a. Using sitemaps and stable URLs b. Using open formats c. Following accessibility standards, such as: i. Section 508 (https://www.access-board.gov/guidelines-and-standards/communications-and-it/about-the-section-508-standards/guide-to-the-section-508-standards) ii. Web Content Accessibility Guidelines (WCAG) (https://www.w3.org/WAI/intro/wcag) iii. United States Web Design Standards (https://standards.usa.gov/) d. Providing page specific titles and description, publication or update dates, and meaningful web addresses, when possible, to convey the substance of content presented 2. Resources that address this further and may be helpful to content creators can be found on the Library of Congress Guide to Creating Preservable Websites (http://loc.gov/webarchiving/preservable.html)	

i. Websites		
B. Formats	1. The Library, and other organizations involved in web archiving, are preserving web content in the Web ARChive (WARC) format	B. Internet Archive's ARC_IA format, a precursor to the WARC format
C. Delivery Method	1. Capture using tools that produce non-proprietary output, to conform with standard formats and requirements	1. Transmission of WARC or ARC IA files created by web content producers or other archiving organizations
D. Metadata	1. Refer to the WARC ISO-standard specification for mandatory and recommended metadata fields 2. When displaying archived content, the following should be clearly indicated: a. archiving institution, b. dates and time of capture, c. statements about functionality within the archive to distinguish from the live site	1. The ARC IA should be named in a manner that easily identifies the archiving institution (see WARC standard for recommended naming conventions)
E. Technological Measures	1. Websites should not contain measures (such as content behind logins or only accessible through search functions) that control access to or prevent capture of the digital work. 2. Robots.txt restrictions should be set so as not to block crawlers from capturing important content, such as image and style sheets, which allow for replay of the site as it looked at the time of capture.	1. Tools currently available cannot capture all web content, so certain types of web content may not be preservable through web capture at this time. These include: a. Multi-media rich content b. Streaming media c. Deep web content d. Databases