

NARA 2014-04: Appendix A, Revised Format Guidance for the Transfer of Permanent Electronic Records – Tables of File Formats

Quick Links

1. Computer Aided Design	2. Digital Audio	3. Digital Moving Images
3.1 Digital Cinema	3.2 Digital Video	4. Digital Still Images
4.1 Digital Photographs	4.2 Scanned Text	4.3 Digital Posters
5. Geospatial Formats	6. Presentation Formats	7. Textual Data
8. Structured Data Formats	9. Email	10. Web Records

Symbol Key

Preferred Formats 

Acceptable Formats 

Acceptable for Imminent Transfer Formats 

1. Computer Aided Design (CAD)

Computer aided design (CAD) – CAD formats are vector graphics files that rely on mathematical expressions to create multi-dimensional computer graphics intended for use in engineering and manufacturing design. CAD programs can generate representations and animations of two and three-dimensional surface projections of objects.

Preferred Formats

Preferred Formats	Format Specifications
Extensible 3D (X3D)	ISO/IEC 19775-1:2008: (http://www.web3d.org/files/specifications/19775-1/V3.2/index.html)
Standard for the Exchange of Product Model Data (STEP)	ISO 10303-21:2002: (http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=33713) ISO 10303-28:2007: (http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=40646)

Acceptable Formats

Acceptable Formats	Format Specifications
Portable Document Format/Engineering (PDF/E)	ISO 24517-1:2008 Document management -- Engineering document format using PDF -- Part 1: Use of PDF 1.6 (PDF/E-1): (http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=42274)
Universal 3D (U3D)	Universal 3D File Format. Standard ECMA-363. 4th edition (June 2007): (http://www.ecma-international.org/publications/standards/Ecma-363.htm)
Product Representation Compact (PRC)	Acrobat 3D PRC Specification (Version 7094): (http://livedocs.adobe.com/acrobat_sdk/9/Acrobat9_HTMLHelp/API_References/PRCReference/PRC_Format_Specification/)

2. Digital Audio

The Digital audio category encompasses formats used to encode recorded sound as machine readable files by converting acoustic sound waves into digital signals. Digital audio formats are generally composed of both a wrapper format, usually the common name associated with the file extension, and an encoding method or codec.

General requirements for digital audio records:

- Digitize to standards appropriate for the accurate preservation of the original audio, when converting analog material (e.g., audio cassettes, record albums, and reel-to-reel audio tapes). Examples of appropriate methods and formats are available on [NARA's Digitization Services Products and Services](#) page;
- Transfer digital audio at a minimum of 16 bits per sample, but 24 bits per sample is encouraged; and
- Transfer digital audio at a minimum sample rate of at least 44.1 KHz, but sampling at 96 KHz is encouraged.



Preferred Formats

Preferred Formats	Format Version	Codecs	Format Specifications
Broadcast Wave (BWF)	0, 1 & 2	Linear Pulse Code Modulated Audio (LPCM)	European Broadcast Union (EBU). Tech Specification of the Broadcast Wave Format (BWF) – Version 1: (http://web.archive.org/web/20091229093941/http://tech.ebu.ch/docs/tech/tech3285.pdf) Specification of the Broadcast Wave Format (BWF) - Version 2: (https://tech.ebu.ch/docs/tech/tech3285.pdf)
Free Lossless Audio Compression	1.21	FLAC	FLAC Format Specification version 1.21: (http://flac.sourceforge.net/format.html)

(FLAC)			
--------	--	--	--

☆☆ Acceptable Formats

Acceptable Formats	Format Version	Codecs	Format Specifications
Audio Interchange Format (AIFF)	1.3	Linear Pulse Code Modulated Audio (LPCM)	Audio Interchange File Format: "AIFF" A Standard for Sampled Sound Files Version 1.3 Apple Computer, Inc.: (http://www-mmsp.ece.mcgill.ca/Documents/AudioFormats/AIFF/Docs/AIFF-1.3.pdf)
MPEG Audio Layer III (MP3)		MP3enc, Lame	ISO/IEC-11172-3 Information technology – Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s – Part 3: Audio: (http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=22412) ISO/IEC 13818-3:1995 Information technology – Generic coding of moving pictures and associated audio information – Part 3: Audio: (http://www.iso.org/iso/home/store/catalogue_ics/catalogue_detail_ics.htm?csnumber=26797)
Wave Waveform Audio File Format (Wave)		Linear Pulse Code Modulated Audio (LPCM)	Multimedia Programming Interface and Data Specifications 1.0: (http://www-mmsp.ece.mcgill.ca/Documents/AudioFormats/WAVE/Docs/riffmci.pdf)

3. Digital Moving Images

Digital moving images consist of bitmap digital images or “frames” displayed in rapid succession at a constant rate, giving the appearance of movement. This category includes two subcategories: digital cinema which encompasses digitized film; and digital video (including both video digitized from analogue sources and born digital video).

General requirements for digital moving image records:

- Agencies must digitize to standards appropriate for accurate preservation of the original video and audio components, when converting analog material. Examples of appropriate methods and formats are available on [NARA’s Digitization Services Products and Services](#) page; and
- For reformatted video, 8-bit is acceptable but 10-bit is preferred.

3.1 Digital Cinema

☆☆☆ Preferred Formats

Preferred Formats	Format Version	Codecs	Format Specifications
Digital Moving Picture Exchange Bitmap (DPX)	1 & 2	Uncompressed	Society of Motion Picture Television Engineers. SMPTE Standard 268M-1994 (DPX Version 1.0): (http://standards.smpte.org/) Society of Motion Picture Television Engineers. SMPTE Standard 268M-2003 (DPX Version 2.0): (http://standards.smpte.org/)

3.2 Digital Video



Acceptable Formats

Acceptable Formats	Format Versions	Codecs	Format Specifications
Audio Video Interleaved Format (AVI)		Uncompressed 4:2:2	Multimedia Programming Interface and Data Specifications 1.0: (http://www.kk.ij4u.or.jp/~kondo/wave/mpidata.txt)
Quicktime File Format (MOV)		Uncompressed 4:2:2	Apple QuickTime File Format Specification (ISO/IEC 14496-14:2003): (https://developer.apple.com/library/mac/documentation/QuickTime/QTFF/QTFFPreface/qtffPreface.html#//apple_ref/doc/uid/TP40000939)
Windows Media Video 9 File Format (WMV)	9	VC-1	Advanced Systems Format (ASF) Specification Revision 01.20.03 Microsoft Corporation December 2004: (http://msdn.microsoft.com/en-us/library/bb643323.aspx) Windows Media Video 9 encoder: (http://msdn.microsoft.com/en-us/library/windows/desktop/ff819505(v=vs.85).aspx)
MPEG 4		H.264	ISO/IEC 14496-10:2003. Information technology -- Coding of audio-visual objects -- Part 10: Advanced Video Coding (formal name) MPEG-4, Advanced Video Coding: (http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=37729)
MPEG-2 Video (MPEG2)			ISO/IEC 13818-2:2000 Information technology -- Generic coding of moving pictures and associated audio information: Video: (http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=31539)
Material Exchange Format (MXF)		J2K-losslessly-compressed	ST 377-1:2011 Material Exchange Format (MXF) — File Format Specification: (http://standards.smpte.org/content/978-1-61482-517-3/st-377-1-2011/SEC1.abstract?sid=63bac43b-e0e1-40a3-8019-d379a103987e) ISO/IEC 15444-1:2004 Information technology -- JPEG 2000 image

			coding system: Core coding system: (http://www.iso.org/iso/catalogue_detail.htm?csnumber=37674)
--	--	--	---

4. Digital Still Images

Digital still images are files that are sampled and bitmapped as a grid of rectangular dots, picture elements (pixels) or points of color. This category encompasses two subcategories: digital photographs (digitally captured photographs or digital scans of photographic prints or negatives), and scanned text.

4.1 Digital Photographs

Digital photographs include still photographs of natural, real-world scenes or subjects produced by digital cameras, and scanned images of photographic prints, slides, and negatives. The guidance applies to master image files of digital photographs created using medium to high quality resolution settings appropriate for continued preservation.

General requirements for digital photographic records:

- Agencies should use appropriate, professional quality, dedicated photographic equipment when capturing images;
- When converting analog material (photographic prints, glass plate negatives, slides, etc.), agencies must digitize to standards appropriate for the accurate preservation of the original image. Examples of appropriate methods and formats are available on [NARA's Digitization Services Products and Services](#) page;
- Agencies must digitize analog originals at a minimum resolution of 3,000 pixels across the long dimension; and
- NARA prefers images that are uncompressed or which make use of lossless compression.

The requirements for digital photographic records such as aerial photography are described in section 5. "Geospatial formats". Additional special requirements for digital photographs are described in 36 CFR 1237.28.



Preferred Formats

Preferred Formats	Format Versions	Format Specifications
Tagged Image File Format (TIFF)	4, 5, & 6	TIFF Revision 6.0 Final — June 3, 1992 Adobe Systems Incorporated: (http://partners.adobe.com/public/developer/en/tiff/TIFF6.pdf)



Acceptable Formats

Acceptable Formats	Format Versions	Format Specifications

JPEG File Interchange Format (JFIF) with Joint Photographic Experts Group (JPEG) compression	1.02	ISO/IEC 10918-5 Information technology – Digital Compression and coding of continuous-tone still images: JPEG Interchange File Format: (http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=54989) ISO/IEC 10918-1:1994 Information technology – Digital Compression and coding of continuous-tone still images: Requirements and guidelines: (http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=18902)
Digital Negative (DNG)	1.4.0.0	Adobe Digital Negative (DNG) Specification Version 1.4.0.0: (http://www.images.adobe.com/www.adobe.com/content/dam/Adobe/en/products/photoshop/pdfs/dng_spec_1.4.0.0.pdf)
Portable Network Graphics (PNG)	1.2	ISO/IEC 15948:2004 Information technology -- Computer graphics and image processing -- Portable Network Graphics (PNG): Functional specification: (http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=29581)
Jpeg2000 (JP2)	JP2 – Part 1	ISO/IEC 15444-1:2004 Information technology – JPEG 2000 image coding system: Core coding system: (http://www.iso.org/iso/catalogue_detail.htm?csnumber=37674)

4.2 Scanned Text

Scanned text is a photograph of a printed page produced either by a digital camera or scanner.

General requirements for scanned text include the following:

- Agencies must digitize to standards appropriate for the accurate preservation of the information on the printed page. When converting analog or film based material (microfilm, microfiche, slides, etc.), agencies must digitize to standards appropriate for the accurate preservation of the original image. Examples of appropriate methods and formats are available on [NARA's Digitization Services Products and Services](#) page;
- Bitonal (1-bit black and white) images must be scanned at 300-600 ppi. Scanning at 600 ppi is recommended. This is appropriate for documents that consist exclusively of clean printed type possessing high inherent contrast (e.g., laser printed or typeset on a white background);
- Gray scale (8-bit) must be scanned at 300-400 ppi. Scanning at 400 ppi is recommended. This is appropriate for textual documents of poor legibility because of low inherent contrast, staining or fading (e.g., carbon copies, thermofax, documents with handwritten annotations or other markings), or that contain halftone illustrations or photographs; and
- Color (24-bit RGB [Red, Green, Blue]) must be scanned at 300-400 ppi. Scanning at 400 ppi is recommended. Color mode (if technically available) is appropriate for text containing color information important to interpretation or content.



Preferred Formats

Preferred Formats	Format Versions	Format Specifications
Tagged Image File Format (TIFF)	4, 5 & 6	TIFF Revision 6.0 Final — June 3, 1992 Adobe Systems Incorporated: (http://partners.adobe.com/public/developer/en/tiff/TIFF6.pdf)
Jpeg2000 (JP2)	Part 1 (JP2)	ISO/IEC 15444-1:2004 Information technology – JPEG 2000 image coding system: Core coding system: (http://www.iso.org/iso/catalogue_detail.htm?csnumber=37674)
Portable Network Graphics (PNG)	1.2	ISO/IEC 15948:2004 Information technology -- Computer graphics and image processing -- Portable Network Graphics (PNG): Functional specification: (http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=2958)
Portable Document Format/Archival (PDF/A)	PDF/A-1	ISO 19005-1:2005 Electronic document file format for long-term preservation – Part 1: Use of PDF 1.4 (PDF/A-1): (http://www.iso.org/iso/catalogue_detail?csnumber=38920)



Acceptable Formats

Acceptable Formats	Format Versions	Format Specifications
JPEG File Interchange Format (JFIF) with Joint Photographic Experts Group (JPEG) compression	1.02	ISO/IEC 10918-5 Information technology – Digital Compression and coding of continuous-tone still images: JPEG Interchange File Format: (http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=54989) ISO/IEC 10918-1:1994 Information technology – Digital Compression and coding of continuous-tone still images: Requirements and guidelines: (http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=18902)
Graphics Interchange Format (GIF)	87a & 89a	Graphics Interchange Format (sm) Version 89a: (http://www.w3.org/Graphics/GIF/spec-gif89a.txt)
PDF/A-2	PDF/A-2	ISO 19005-2:2011 Document management -- Electronic document file format for long-term preservation -- Part 2: Use of ISO 32000-1 (PDF/A-2): (http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=50655)

4.3 Digital Posters

Digital posters include both posters created digitally and scanned images of analog posters. Posters are generally large in format and usually printed and displayed for advertising and publicizing purposes.

General requirements for digital posters include the following:

- Agencies must digitize to standards appropriate for the accurate preservation of the information of the image. When converting analog or film based material (microfilm, microfiche, slides, etc.),

agencies must digitize to standards appropriate for the accurate preservation of the original image. Examples of appropriate methods and formats are available on [NARA's Digitization Services Products and Services](#) page;

- Bitonal (1-bit black and white) images must be scanned at 300-600 ppi. Scanning at 600 ppi is recommended. This is appropriate for documents that consist exclusively of clean printed type possessing high inherent contrast (e.g., laser printed or typeset on a white background);
- Gray scale (8-bit) must be scanned at 300-400 ppi. Scanning at 400 ppi is recommended. This is appropriate for textual documents of poor legibility because of low inherent contrast, staining or fading (e.g., carbon copies, thermofax, documents with handwritten annotations or other markings), or that contain halftone illustrations or photographs; and
- Color (24-bit RGB [Red, Green, Blue]) must be scanned at 300-400 ppi. Scanning at 400 ppi is recommended. Color mode (if technically available) is appropriate for text containing color information important to interpretation or content.



Preferred Formats

Preferred Formats	Format Versions	Format Specifications
Tagged Image File Format (TIFF)	4, 5 & 6	TIFF Revision 6.0 Final — June 3, 1992 Adobe Systems Incorporated: (http://partners.adobe.com/public/developer/en/tiff/TIFF6.pdf)
Jpeg2000 (JP2)	Part 1 (JP2)	ISO/IEC 15444-1:2004 Information technology – JPEG 2000 image coding system: Core coding system: (http://www.iso.org/iso/catalogue_detail.htm?csnumber=37674)
Portable Network Graphics (PNG)	1.2	ISO/IEC 15948:2004 Information technology -- Computer graphics and image processing -- Portable Network Graphics (PNG): Functional specification: (http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=29581)
Portable Document Format/Archival (PDF/A)	PDF/A-1	ISO 19005-1:2005 Electronic document file format for long-term preservation – Part 1: Use of PDF 1.4 (PDF/A-1): (http://www.iso.org/iso/catalogue_detail?csnumber=38920)



Acceptable Formats

Acceptable Formats	Format Versions	Format Specifications
JPEG File Interchange Format (JFIF) with Joint Photographic Experts Group (JPEG)	1.02	ISO/IEC 10918-5 Information technology – Digital Compression and coding of continuous-tone still images: JPEG Interchange File Format: (http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=54989) ISO/IEC 10918-1:1994 Information technology – Digital Compression and coding of continuous-tone still images: Requirements and guidelines: (http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?)

compression		csnumber=18902)
Graphics Interchange Format (GIF)	87a & 89a	Graphics Interchange Format (sm) Version 89a: (http://www.w3.org/Graphics/GIF/spec-gif89a.txt)

5. Geospatial Formats

Geospatial records include digital cartographic data files and aerial photography that are created and processed in Geographic Information Systems (GIS) or other software applications for spatial analysis.



Preferred Formats

Preferred Formats	Format Versions	Format Specifications
Geospatial Tagged Image File Format	1.8.2	Geo TIFF Format Specification: (http://www.remotesensing.org/geotiff/spec/geotiffhome.html)
Geographic Markup Language	2.0 through 3.2	ISO 19136:2007 & Version 3.2, OGC document 07-036: (http://www.opengeospatial.org/standards/is)
Topologically Integrated Geographic Encoding and Referencing Files	2006 Second Edition	2006 Second Edition TIGER/Line®: (https://www.census.gov/geo/maps-data/data/pdfs/tiger/tiger2006se/tgr06se.pdf)
Keyhole Markup Language	2.2	Open Geospatial Consortium Inc. OGC 07-147r2: (http://www.opengeospatial.org/standards/kml/)



Acceptable Formats

Acceptable Formats	Format Versions	Format Specifications
Vector Product Format		MIL-STD-2407: (http://earth-info.nga.mil/publications/specs/printed/2407/2407_VPF.pdf)
ESRI ARC/INFO Interchange File Format		Reverse engineered specification: (http://avce00.maptools.org/docs/v7_e00_cover.html)
TerraGo Geospatial PDF	GeoPDF Encoding Best Practice Version 2.2	Open Geospatial Consortium Inc. OGC 08-139r2: (http://www.opengeospatial.org/standards/is)
ESRI Shapefile (Compound)	1997 – current version	ESRI Shapefile Technical Description: (http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf)



Acceptable for Imminent Transfer Formats

Acceptable for Imminent Transfer Formats	Format Versions	Format Specifications

Spatial Data Transfer Standard (SDTS)	All versions	ANSI NCITS 320-1998: (http://mcmcweb.er.usgs.gov/sdts/standard.html)
---------------------------------------	--------------	---

6. Presentation Formats

Presentation formats are used to convey graphical information to audiences in the form of a slide show. Presentation formats are not acceptable for use as transfer containers for permanent digital still images.



Preferred Formats

Preferred Formats	Format Versions	Format Specifications
OpenDocument Presentation Format (ODP)	1.0	ISO/IEC 26300:2006 Information technology -- Open Document Format for Office Applications (OpenDocument) v1.0: (http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=43485)
Portable Document Format Archival (PDF/A-1)	PDF/A-1	ISO 19005-1:2005 Electronic document file format for long-term preservation – Part 1: Use of PDF 1.4 (PDF/A-1): (http://www.iso.org/iso/catalogue_detail?csnumber=38920)



Acceptable Formats

Acceptable Formats	Format Versions	Format Specifications
Microsoft Powerpoint 1997-2007 Binary Format (PPT)	8.0	[MS-PPT]: PowerPoint (.ppt) Binary File Format: (http://msdn.microsoft.com/en-us/library/cc313106(v=office.12).aspx)
Microsoft Powerpoint Office Open XML Format (PPTX)		[MS-OI29500]: Office Implementation Information for ISO/IEC 29500 Standards Support: (http://msdn.microsoft.com/en-us/library/ee908652%28v=office.12%29)
PDF/A-2	PDF/A-2	ISO 19005-2:2011 Document management -- Electronic document file format for long-term preservation -- Part 2: Use of ISO 32000-1 (PDF/A-2): (http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=50655)

7. Textual Data

The textual data category refers to two general content types: unformatted (plain text) or formatted. Unformatted plain text (defined in MIME as text/plain) contains basic character information and control or non-printing characters but lacks styling information. Formatted text files include all of the attributes of plain text files but have extended formatting capabilities, for “stylized” or “rich” text features including

italics, bold, colors, hyper-linking, etc.

Agencies must identify the character encoding method used with each text file.



Preferred Formats

Preferred Formats	Format Versions	Format Specifications
ASCII Text	7 bit	ISO/IEC 646:1991 Information technology -- ISO 7-bit coded character set for information interchange: (http://www.iso.org/iso/catalogue_detail.htm?csnumber=4777)
Unicode Text	UTF-8 UTF-16	RTF 3629: UTF-8, A Transformation Format of ISO 10646: (http://tools.ietf.org/html/rfc3629) RFC 2781 UTF-16: An Encoding of ISO 10646: (http://www.ietf.org/rfc/rfc2781.txt)
OpenDocument Text Format (ODF)	OpenDocument 1.0	ISO/IEC 26300:2006 Information technology -- OpenDocument Format for Office Applications (OpenDocument) v1.0: (http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=43485)
PDF/A-1	PDF/A-1	ISO 19005-1:2005 Document management -- Electronic document file format for long-term preservation -- Part 1: Use of PDF 1.4 (PDF/A-1): (http://www.iso.org/iso/catalogue_detail?csnumber=38920)
PDF/A-2	PDF/A-2	ISO 19005-2:2011 Document management -- Electronic document file format for long-term preservation -- Part 2: Use of ISO 32000-1 (PDF/A-2): (http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=50655)



Acceptable Formats

Acceptable Formats	Format Versions	Format Specifications
PDF	PDF 1.7 PDF 1.0-1.6	ISO 32000-1:2008 Document management -- Portable document format -- Part 1: PDF 1.7: (http://www.iso.org/iso/catalogue_detail.htm?csnumber=51502) Adobe® Portable Document Format Version 1.6: http://www.adobe.com/devnet/pdf/pdf_reference_archive.html
Microsoft Word (DOCX) Office Open XML	OOXML Microsoft Word for Windows, version 2007-2010	[MS-OI29500]: Office Implementation Information for ISO/IEC 29500 Standards Support: (http://msdn.microsoft.com/en-us/library/ee908652%28v=office.12%29)
Microsoft Word 97 Binary Document Format (DOC)	8.0	[MS-DOC]: Word (.doc) Binary File Format: (http://msdn.microsoft.com/en-us/library/cc313153%28v=office.12%29.aspx)

8. Structured Data Formats

Structured data comprises the broad category of data that is stored in defined fields. Categories for structured data are as follows:

- Database formats are organized collections of associated data that conform to a logical structure. Database formats are determined by “data models” that describe specific data structures used to model an application and generally include navigational, relational, and hybrid models;
- Spreadsheets are tables made up of columns and rows and which contain cells of data. Relationships between cells can be pre-defined as mathematical formulas;
- Statistical data is the result of quantitative research and analysis. Statistical data formats contain collections of data presented in both tabular and non-tabular form; and
- Scientific data refers to research data collected by instrumentation tools during the scientific process. Scientific data formats are either domain specific within a single field of study, or are multi-domain formats used for transfer of scientific data between domains.

General requirements for structured data include the following:

- Agencies must transfer structured data that is both well-formed according to the syntactical conventions of the format, and valid according to the structural rules defined in any associated schemas or document type definitions (DTDs);
- Value Separated Files, e.g. CSV or comma separated value files, may use a character other than the comma. The pipe or caret are recommended delimiters because they are not commonly found in free text fields. Alternatively, text files encoded with ASCII characters and where each field is a fixed width, is also an acceptable transfer format for use with structured data, even though ASCII is technically a data encoding type. ASCII text files must be accompanied by complete documentation of the record lengths and field widths;
- Data files and databases shall be transferred as flat files or as rectangular tables, that is, as two-dimensional arrays, lists or tables. All records in a database, or rows (tuples) in a relational database, should have the same logical format. Each data element within a record should contain only one data value. A record should not contain nested repeating groups of data items; and
- Structured data must be transferred together with any associated files necessary to verify the validity of the data, e.g., DTDs, schemas, and data dictionaries.



Preferred Formats

Preferred Formats	Format Versions	Format Specifications
Comma Separated Value	N/A	Common Format and MIME Type for Comma-Separated Values (CSV) Files:

(CSV)		(http://tools.ietf.org/html/rfc4180)
OpenDocument Format Spreadsheet (ODS)		ISO/IEC 26300:2006 Information technology -- Open Document Format for Office Applications (OpenDocument) v1.0: (http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=43485)
ASCII Text	7 bit	ISO/IEC 646:1991 Information technology -- ISO 7-bit coded character set for information interchange: (http://www.iso.org/iso/catalogue_detail.htm?csnumber=4777)
JavaScript Object Notation (JSON)		The application/json Media Type for JavaScript Object Notation (JSON): (http://www.ietf.org/rfc/rfc4627.txt?number=4627)
Extensible Markup Language (XML)	1.1	Extensible Markup Language (XML) 1.1 (Second Edition): (http://www.w3.org/TR/2006/REC-xml11-20060816/)

Acceptable Formats

Acceptable Formats	Format Versions	Format Specifications
Microsoft Excel Office Open XML	OOXML Workbook Excel 2007-2010 XLXS Microsoft Excel for Windows, version 2007	[MS-OI29500]: Office Implementation Information for ISO/IEC 29500 Standards Support: (http://msdn.microsoft.com/en-us/library/ee908652%28v=office.12%29)
Microsoft Excel 97 Binary Document Format (XLS)	Version 8.0	[MS-XLS]: Excel Binary File Format (.xls) Structure: (http://msdn.microsoft.com/en-us/library/cc313154(v=office.12).aspx)

Acceptable for Imminent Transfer Formats

Acceptable for Imminent Transfer Formats	Format Versions	Format Specifications
Extended Binary Coded Decimal Interchange Code (EBCDIC)	U.S. EBCDIC	IBM EBCDIC Code Page 0037: (http://www-01.ibm.com/software/globalization/cp/cp00037.html)

9. Email

Email is defined as discrete electronic communications transmitted over the Simple Mail Transfer Protocol (SMTP), between two or more people or entities, in compliance with applicable IETF's Request for Comments (RFC) specifications. Email does not include other functions commonly available via email programs such as calendars, tasks, appointments, newsgroups, or instant messaging. In order for information in a calendar, contact list, address book etc. to be transferred to

NARA, it must be scheduled as a separate item.

Please note that NARA considers email attachments to be a component of the email record and does not require that unseparated email attachments meet the transfer standards specified by the format category under which the attachment alone would fall.

General requirements for email:

- Transfers of email records must consist of an identifiable, organized body of records (not necessarily a traditional series);
- Email messages should include delimiters that indicate the beginning and end of each message and the beginning and end of each attachment, if any. Each attachment must be differentiated from the body of the message, and uniquely identified;
- Email messages transferred as XML files must be accompanied by any associated document type definitions (dtds), schemas, and/or data dictionaries;
- Labels to identify each part of the message (Date, To [all recipients, including cc: and bc: copies], From, Subject, Body, and Attachment) including transmission and receipt information (Time Sent, Time Opened, Message Size, File Name, and similar information, if available). To ensure identification of the sender and addressee(s), agencies that use an email system that identifies users by codes or nicknames, or identifies addressees only by the name of a distribution list should include information with the transfer-level documentation; and
- Email converted to formats not natively used by the email program, and which do not maintain header information (such as RTF or Word documents) are not accepted. Printouts of emails are also not accepted under this Bulletin.



Preferred Formats for Individual Messages

Preferred Formats for Individual Messages	Format Versions	Format Specifications
Internet Message Format (EML)		Internet Message Format: (http://www.ietf.org/rfc/rfc2822.txt) And MIME: (http://tools.ietf.org/html/rfc2045), (http://tools.ietf.org/html/rfc2046), (http://tools.ietf.org/html/rfc2047), (http://tools.ietf.org/html/rfc4288), (http://tools.ietf.org/html/rfc4289), (http://tools.ietf.org/html/rfc2049)
MBOX Email Format (MBOX)		MBOX Email Format: (https://tools.ietf.org/html/rfc4155) And MIME: (http://tools.ietf.org/html/rfc2045), (http://tools.ietf.org/html/rfc2046),

		(http://tools.ietf.org/html/rfc2047), (http://tools.ietf.org/html/rfc4288), (http://tools.ietf.org/html/rfc4289), (http://tools.ietf.org/html/rfc2049)
--	--	---

☆☆ Acceptable Formats for Individual Messages

Acceptable Formats for Individual Messages	Format Versions	Format Specifications
Extensible Markup Language (XML)	1.1	Extensible Markup Language (XML) 1.1 (Second Edition): (http://www.w3.org/TR/2006/REC-xml11-20060816/)
Microsoft Outlook Item Message Format (MSG)		Microsoft Outlook Item Message Format: (http://msdn.microsoft.com/en-us/library/cc463912(v=exchg.80).aspx)

☆☆☆ Preferred Formats for Aggregations of Email

Preferred Formats for Aggregations of Email	Format Versions	Format Specifications
Microsoft Personal Folders Format (PST)		Outlook Personal Folders File Format: (http://msdn.microsoft.com/en-us/library/ff385210%28v=office.12%29.aspx)
MBOX Email Format (MBOX)		MBOX Email Format: (https://tools.ietf.org/html/rfc4155) And MIME: (http://tools.ietf.org/html/rfc2045), (http://tools.ietf.org/html/rfc2046), (http://tools.ietf.org/html/rfc2047), (http://tools.ietf.org/html/rfc4288), (http://tools.ietf.org/html/rfc4289), (http://tools.ietf.org/html/rfc2049)

10. Web Records

Web records consist of web sites and social media sites created and maintained to provide information and services of the United States Government via the World Wide Web. This Bulletin applies to web records managed by an agency that have been appraised and scheduled for permanent retention by NARA. Agencies should harvest websites using a utility that will package component files in a manner that meets the following general requirements.

General requirements for web content records:

- Web records must be accessible via Hypertext Transfer Protocol (HTTP) from a server to a client *browser* when a *URL* has been activated;

- Web content records that share a domain name including content managed under formal agreement and residing on another site must be transferred together;
- All component parts of web content records that have been appraised as permanent including image, audio, video and all other proprietary formats, must be transferred in a manner that maintains all of the original links, functionality and data integrity;
- Dynamic content such as calendars or databases either must be transferred in an acceptable format, or be made accessible as static content;
- All internally referenced URLs must be included with the transfer set; and
- All control information from the harvesting protocol must be maintained.

The following will not be accepted for transfer under this Bulletin:

- Program or administrative records documenting the management of web sites;
- Externally referenced content (e.g., accessed via hyperlink) that resides in a different domain and is not managed for an agency under a formal agreement;
- Static images, (such as screen shots), of web content records, because they do not retain hypertext functionality.

Acceptable Formats

Acceptable Formats	Format Versions	Format Specifications
Web ARChive Format (WARC)	.18	ISO 28500:2009 Information and documentation -- WARC file format: (http://www.iso.org/iso/catalogue_detail.htm?csnumber=44717)
Archive File Format (ARC)	1.0	Arc File Format: (http://archive.org/web/researcher/ArcFileFormat.php)

[Contact Us](#) [Accessibility](#) [Privacy Policy](#) [Freedom of Information Act](#) [No FEAR Act](#) [USA.gov](#)

The U.S. National Archives and Records Administration

PDF files require the [free Adobe Reader](#). 1-86-NARA-NARA or 1-866-272-6272
More information on Adobe Acrobat PDF files is available on our [Accessibility page](#).