28 Aug 2017

Guidelines - File Format Recommendations

James Roger Weaver

Missouri University of Science and Technology, weaverjr@mst.edu

Guidelines - File Format Recommendations

Purpose

The purpose of these guidelines is to provide recommendations and support as to appropriate file types to use when submitting files to Scholars' Mine for long-term preservation and access.

Recommendations

Some file formats are better than others when planning for long-term preservation and deposit to Scholars' Mine. Typically, applications used to open files and their specifications become harder to locate over time. This results in some formats becoming difficult or impossible to read in the future. File formats with the best chance for long-term viability have open specifications, are of high quality and in wide use. Choosing the right file format is key to their long-term preservation.

To ensure the long-term viability of your digital files:

1. Use the recommended file formats,
2. Make certain the format is an uncompressed version, or the highest quality compressed version, or the final production version,
3. Exercise caution when using proprietary formats and digital rights management (DRM) software.

It is not possible to provide a single list of file formats that are appropriate for all use cases, but the list below is a good starting point.

Simplified File Format Recommendations

- **Audio**
    - o Highly recommended: AIFF (.aif, .aiff) or WAV (.wav)
    - o Moderately recommended: MP3 (.mp3), AAC (.mp4, .mp4a, .aac), FLAC (.flac) or ALAC (.m4a)
- **Image**
    - o Highly recommended: Full color images @ 600dpi or higher saved as JPEG2000 Lossless (.jp2) or TIFF 24-bit, uncompressed (.tif, .tiff)
    - o Moderately recommended: lossy compressed formats limited to JPEG (.jpg, .jpeg), JPEG2000 (.jp2), TIFF (.tif, .tiff), or PNG (.png) at highest quality possible.
- **Video**
    - o Highly recommended: Uncompressed Quicktime Movie (.mov); uncompressed AVI (.avi)
    - o Moderately recommended: MPEG-1, MPEG-2 or MPEG-4 encoded video (.avi, .mpg, .mpeg, .mov, .mkv, .mp4)
- **Virtual Reality/3D**
    - o Highly recommended: X3D (*.x3d). Please contact Scholars' Mine to determine best output settings and format
- **Text**
    - o Highly recommended: Open Document Text (.odt), UTF-8 Unicode text (.txt), or PDF/A (.pdf)
    - o Moderately recommended: Rich Text Format (.rtf).
- **Presentation**
    - o Highly recommended: Open Document Presentation (.odp)

- o Moderately recommended: PDF/A (.pdf) for images only.
- **Spreadsheet**
  - o Highly recommended: Open Document Spreadsheet (.ods)
  - o Moderately recommended: Comma separated value CSV (.csv) or Tab-delimited text file (.txt).

Simplified File Format Recommendations

High probability of long-term preservation	Moderate probability of long-term preservation	Low probability of long-term preservation
Audio		
- AIFF (PCM) (.aif, .aiff) - WAV (PCM) (.wav) - Broadcast WAV (PCM) (.wav, .bwav) For music: 24 bit, 96 kHz or better For speech: 16 bit, 44.1 kHz or bette	For music: - AIFF (PCM) (.aif, .aiff) - WAV (PCM) (.wav) - Free Lossless Audio Codec (.flac) - Apple Lossless (.m4a) For speech: - Advance Audio Coding (.mp4, .m4a, *.aac) - MP3 (MPEG-1/2, Layer 3) (.mp3) - MP2 (MPEG-1 Audio Layer II) (.mp2) 128 kbps (mono), 256 kbps (stereo) or better, 44.1 kHz, variable bit rate settings off if possible. AIFF or WAV preferable.	- Advance Audio Coding (.mp4, .m4a, .aac) for music - MP3 (MPEG-1/2, Layer 3) (.mp3) for music - MP2 (MPEG-1 Audio Layer II) (.mp2) for music - Ogg Vorbis (.ogg) AIFC (compressed) (.aifc) - NeXT SND (.snd) - RealNetworks 'Real Audio' (.ra, .rm, *.ram) - Windows Media Audio (.wma) - Protected AAC (.m4p) - WAV (compressed) (.wav)
Audio (MIDI)		
	- Standard General MIDI (.mid, .midi), original or Level 2, Universal SysEx commands only, instruments well documented	- Any other MIDI file not listed here

Image (raster)		
Full color @ 600dpi or higher: - JPEG2000 Lossless (.jp2) - TIFF 24-bit, uncompressed (.tif, .tiff)	- JPEG (.jpg, .jpeg) - JPEG2000 lossy compressed (.jp2) - PNG (.png) - TIFF lossy compressed (.tif, .tiff)	- GIF (.gif) - Macintosh Graphics (.pict) - Photoshop document (.psd) - All other image formats not listed here
Image (vector)		
- Open Document Graphics (.odg)	- Scalable vector graphics (.svg)	- Illustrator document (.ai) - Photoshop document (.psd) - All other image formats not listed here
Video		
- AVI (uncompressed, motion JPEG) (.avi) - QuickTime Movie (uncompressed, motion JPEG) (.mov) Recommendation: Maintain original resolution of video and audio	- MPEG-1, MPEG-2 (.avi, .mpg, .mpeg, .mov) - MPEG-4 ASP (.avi, .mkv, .mp4,) - MPEG-4 H.264 (.avi, .mkv, .mp4, .mov,) - Ogg Theora (.ogg) [requires Xiph.org open-source decoder for Windows or Mac playback] Standard Definition recommendations: 720x486 pixels, 2.5Mbits/second for MPEG-4; 5Mbit/second for MPEG-2. Audio 448kbit/sec, 16 bit, 44.1kHz. High Definition recommendations: 1920x1080 pixels, 8 to 15 Mbit/s. Audio 16 bit, 44.1kHz	- 3GP (.3gp, .mp4) - Flash Video (.flv, .f4v) - MPEG-4 ASP (Divx, Xvid) (.avi) - MPEG-4 H.264 (DRM protected) (.m4v, .mp4) - RealNetworks 'Real Video' (.rv) - Windows Media Video (.wmv)

Virtual Reality/3D		
X3D (.x3d)	VRML (.wrl, .vrml) U3D (Universal 3D file format) STEP (ISO 10303) (.stp .step) IGES (NBSIR 80- 1978/ASME Y14.26M) (.igs, .iges) for basic geometry only	All other virtual reality formats not listed here
Text		
- UTF-8 Unicode text (.txt) - Open Document Text (.odt) - PDF/A	- Word XML (.docx) - Rich Text Format (.rtf) - PDF other than PDF/A (.pdf)	- Word (.doc) - HTML (.htm, .html) - Pages (.pages) - All other text formats not listed here
Presentation		
Interactive - Open Document Presentation (.odp) Static (images only) - PDF/A Animated - See video guidelines	- PDF other than PDF/A (.pdf) - PowerPoint XML (.pptx)	- PowerPoint (.ppt) - Keynote (.key) - HTML (.htm, .html
Spreadsheet		
- Open Document Spreadsheet (.ods)	- Excel XML (.xlsx) - Comma separated value CSV (.csv) - Tab-delimited text file (.txt)	- Excel (.xls) - Numbers (.numbers)

Guideline Review

These guidelines and the actions and activities associated with them will be evaluated regularly to ensure that implemented strategies continue to support the Library's mission and policies, use resources in a cost-effective manner, and adapt appropriately to address evolving law and technologies. This evaluation will be completed at least once every three years.