Recommended Data Formats for Preservation Purposes in the FCLA Digital Archive

This table is intended to help Florida university administrators develop guidelines for preparing and submitting files to the FCLA Digital Archive. It is not intended to suggest that these file formats are allowable formats for ETDs or any other official publication of any Florida university. (See the notes following this table.) Source: FCLA, June 2004.

Media	Preferred	Acceptable	Bit-level Preservation only
Text	- Plain text (encoding: US-ASCII, UTF-8, UTF-16 with BOM) - XML (includes XSD/XSL/XHTML, etc.; with included or accessible schema and character encoding explicitly specified) - Computer program source code (*.c, *.c++, *.java, *.js, *.jsp, *.php, *.pl, etc.)	- Cascading Style Sheets (*.css) - DTD - Plain text (ISO8859-1 encoding) - PDF (*.pdf) (embedded fonts, no encryption) - Rich Text Format (*.rtf) version 1.x - HTML 4.x (include a DOCTYPE declaration) - SGML - OpenOffice (*.sxw)	- PDF (with encryption) - Microsoft Word (*.doc) - WordPerfect (*.wpd) - All other text formats not listed here
Raster Image	- TIFF (uncompressed) - PNG (*.png)	- BMP (*.bmp) - JPEG/JFIF (*.jpg) - JPEG2000 (prefer uncompressed) (*.jp2) - TIFF (CCITT Group 3/4, JPEG, PackBits compression)	- MrSID (*.sid) - TIFF (with LZW compression or in Planar format) - GIF (*.gif) - FlashPix - PhotoShop (*.psd) - All other raster image formats not listed here
Vector Graphics	- SVG	- CGM - WebCGM	- EPS - Macromedia Flash (*.swf) - All other vector image formats not listed here
Audio	- AIFF (uncompressed) (*.aif, *.aiff) - WAV (PCM only) (*.wav)	- SUN Audio (uncompressed) (*.au) - MIDI - Ogg Vorbis (OGG)	- AIFC (*.aifc) - NeXT SND (*.snd) - RealNetworks 'Real Audio' (*.ra, *.rm, *.ram) - Windows Media Audio (*.wma) - MP3 (MPEG-1, Layer 3) (*.mp3) - MP2 (MPEG-1, Layer 2) (*.mp2) - WAV (compressed) (*.wav) - All other audio formats not listed here
Video	- MPEG-1, MPEG-2 (*.mpg)		- AVI ('Video for Windows') (*.avi) - QuickTime Movie (*.mov) - RM (RealNetworks; 'Real Video') (*.rv) - Windows Media Video (*.wmv) - All other video formats not listed here

Spreadsheet/ Database	- Delimited Text - SQL DDL	- DBF (*.dbf) - OpenOffice (*.sxc)	- Excel (*.xls) - All other spreadsheet/ database formats not listed here
Virtual Reality	- X3D	- VRML	- All other virtual reality formats not listed here
Computer Programs	See Preferred Text formats	See Acceptable Text formats	- Compiled / Executable files (EXE, *.class, COM, DLL, BIN, DRV, OVL, SYS, PIF)
Presentation		- OpenOffice (*.sxi)	- PowerPoint (*.ppt) - All other presentation formats not listed here

Notes:

1. UMI ProQuest lists file formats that they will accept for theses and dissertations on their website (http://wwwlib.umi.com/dissertations/about_etds). Some of the formats UMI ProQuest accepts (GIF, AVI) are listed in the 'bit-level preservation only' column in the above table. Because UMI ProQuest currently only provides bit-level preservation, this should not represent a conflict for any universities in deciding on acceptable ETD file formats.

2. No files with viruses will be accepted (refers especially to these file formats: DOC, XLS, MDB, PPT, ZIP, EXE). Please scan your files for viruses with up-to-date virus scanners before transmitting your files to the FCLA Digital Archive.

3. Any files fully or partially encrypted will be accepted for bit-level preservation only.

4. Any files fully or partially compressed with a proprietary compression algorithm (LZW, etc.) will be accepted for bit-level preservation only.

5. It is permissible to archive files in the following formats for transmission to the FCLA Digital Archive: Unencrypted ZIP (*.zip), uncompressed TAR (*.tar), TAR + Gzip compression (*.tar.gz)

6. As a general rule, use platform-independent, vendor-independent, non-proprietary, stable, well-supported formats