

# Preferred File Formats

**Date:** October 2015

**Version:** 1.0

**Contributors:**

**Mark Phillips** Assistant Dean for Digital Libraries

**Ana Krahmer** Supervisor, Digital Newspaper Unit

**Hannah Tarver** Department Head, Digital Projects Unit

**Daniel Alemneh** Supervisor, Digital Curation Unit

**Laura Waugh** Repository Librarian for Scholarly Works



This work is licensed under a Creative Commons Attribution 4.0 International License.

Note: This is an archived version of this document.

An up-to-date version if available at:

<http://www.library.unt.edu/digital-projects-unit/digital-file-formats>

# Preferred File Formats

## Introduction

This document provides a general overview of the file types produced by the activities of the Digital Libraries Division, or accepted for inclusion in the UNT Libraries' Digital Collections.

## Preferred File Formats

These are commonly-accepted file formats for which we intend to maintain bitstream preservation in addition to preserving the functionality and content of the files.

	Mimetype	Common Accepted Extensions
Tagged Image File Format	image/tiff	.tif, .tiff
Joint Photographic Experts Group	image/jpeg	.jpg, .jpeg
Portable Document Format	application/pdf	.pdf
Text file	text/plain	.txt
HTML file	text/html	.htm, .html
Audio Video Interleave file	application/x-troff-msvideo video/avi video/msvideo video/avs-video	.avi
MOV	video/quicktime	.mov
Waveform Audio file	audio/wav audio/x-wav	.wav, .wave
Audio Interchange File Format	audio/aiff audio/x-aiff	.aiff, .aif
ARC	application/octet-stream	.arc, .arc.gz
Web ARChive file	application/warc	.warc, .warc.gz
Portable Network Graphics file	image/png	.png

## Recognized File Formats

These file formats are less preferred. Although we still preserve content and files at the bit level, functionality may be lost over time.

	Mimetype	Common Accepted Extensions
Motion Picture Entertainment Group	video/mpeg audio/mpeg	.mpg, .mpeg
Flash Video Format	video/x-flv	.flv
MPEG Group4	video/mp4	.mp4, .m4v,
Windows Media	video/x-ms-wmv	.wmv
MPEG Group4	audio/mpeg3 audio/x-mpeg-3 video/mpeg video/x-mpeg	.mp3

## Standard Normalization Paths

For some born-digital file types, we create a set of derivatives to display the contents in a more standardized way. Although we retain all file types and derivatives, they may not all be visible or accessible from the public interface.

Original File	Intermediary File(s)	Display/Public Files
Word document (.doc, .docx)	.pdf	.jpg
Portable Document Format (.pdf)	>>	.jpg
Powerpoint (.ppt, .pptx)	.pdf	.jpg
Spreadsheet (.xls, .xlsx)	>>	.csv
OpenDocument file: presentation text document	.pdf, .ppt .pdf, .doc	.jpg