



# Preservation Handbook

## Digital by-products

---

Author	Gareth Knight
Version	1
Date	27 July 2005
Change History	

---



## Definition

The AHDS Collections Preservation Policy outlines its aim to “preserve the significant properties... of data-based digital resources”. This is defined as those aspects of the resource that represent important information content. Logically, this presumes that there are aspects of a resource that are not significant and do not require preservation.

For the purpose of this document, digital by-products refer to digital files created by the operating system (OS) that are not an intended part of the collection. These files are created on the depositor's computer before submission or by the computer used by AHDS staff when preparing the AIP. The purpose of these files will vary: some are designed to avoid corruption of the original file until it is saved; others are intended to store custom settings defined by the user. This information is usually invisible to the user, who is often unaware of its existence. AHDS staff should identify digital by-products and ensure they are not made available in the Archival Information Package.

## Additional Information

- File Extension details for .DS\_Store  
< [http://filext.com/detailist.php?extdetail=DS\\_STORE](http://filext.com/detailist.php?extdetail=DS_STORE) > Last checked 11/11/2004
- DS\_Store Terminator  
< <http://www.macupdate.com/info.php/id/10478> > Last checked 11/11/2004
- Description of how Word creates temporary files  
< <http://support.microsoft.com/kb/211632/EN-US/> > Last checked 11/11/2004
- What is thumbs.db  
< <http://www.fcs.iastate.edu/computer/tips/thumbs/default.htm> > Last checked 20/07/2005



# Technical Environment

## Common Formats

Format	File Extension	Notes
DS Store	.ds_store	An index file created automatically by Mac OS to customise folder appearance, e.g. the shape and size of the window, the position of the window on the desktop and whether file, folder or icon view has been selected. Deleting the .DS_Store the folder would revert to the system default next time. These files should not be considered part of the deposit.
Temporary Files	.tmp, .bak	<p>Software applications often create 'working' copies when editing a file (e.g. editing a MS Word document, WAVE audio file, or video data). These files are usually deleted when the user closes the application, however in some circumstances (e.g. application crash or poorly written software) the files remain. As a result, the size of the deposit will increase and information will be duplicated.</p> <p>File names for temporary files will vary between operating systems. Those created on Mac OS precede the filename with '.' (dot, underscore) and are typically between 40 and 1000 bytes. Those created by Windows are preceded by a '~\$ (tilde, dollar) and can be any size. These should be checked within a Hex viewer or other appropriate software.</p> <p>Both forms of temporary files are stored with hidden attributes, meaning you will display hidden files to become aware of their existence.</p>
Thumbnail files	Thumbs.db	Windows 2000 and upwards automatically create and store cached thumbnails of images stored within a single directory. These are, in most cases, hidden by the host computer and only appear when the data is transferred to another machine. These files are superfluous to the collection and should not be made available in a preservation or dissemination copy.
Log files	WS_FTP.Log	In addition to temporary files, software applications often create log files. These files are useful as a method of confirming that a process was performed (e.g. the transfer of data from a remote FTP server), but unimportant for the deposit. A common example is the WS_FTP.log created by the Windows FTP client. These files should not be stored within the deposit unless it is explicitly mentioned within the documentation.
Shortcuts	*.lnk	SIPs sometimes contain Windows shortcuts that refer to a path on the depositors' system. These files do not contain any data and should not be considered part of the deposit.
Media Promotions		Some media contains text and other files that inform the user of its benefits. The most common example (found within several deposits) are Zip disks that contain a readme.txt file with instructions on 'Getting the Most out of Your Zip Disks' and a Flash animation promoting '50 ways to using your Zip Disk' (50ways.exe). These files are covered by existing copyright agreements and should not be considered part of the deposit.