WISCONSIN HISTORICAL SOCIETY

File Management and Processing Tools

File Management and Processing Tools

Contents

Introduction

This guidance document provides a list of software tools that can assist in electronic file management and processing. This document is intended for records managers at state agencies, or other individuals who find themselves tasked with managing large collections of files.

This is an inexhaustive list of tools compiled by the Wisconsin Historical Society. All tools are able to run on Microsoft Windows operating systems and are to download and use. However, records officers and other state agency staff may need the approval or assistance of their IT departments to install software on their work computers.

The Minnesota Historical Society has also released useful [guides for electronic records management tools](#), some of which are included on this list.

Note: The Wisconsin Historical Society does not provide technical support for any of the tools listed. All tools were examined in January 2018, but the WHS does not continuously monitor these tools.

Batch operations

Batch operations, such as batch file or folder renaming, copying, moving, and deleting, can be useful when handing a large collection of electronic records.

Advanced Renamer
Website: https://www.advancedrenamer.com/

Description: Advanced Renamer is a free program for renaming multiple files and folders at once. Users can set up renaming methods to manipulate file names in various ways.

Extensive user documentation can be found online at https://www.advancedrenamer.com/user_guide/gettingstarted.

FastCopy
Website: https://ipmsg.org/tools/fastcopy.html.en

Description: FastCopy is a fast copy and delete utility for Windows. It can be used to complete bulk copy/move/delete operations on files. It allows you to perform these operations quickly without slowing down other applications on your computer. FastCopy is more powerful than a built-in application like Windows Explorer, because it can perform actions such as file verification, detecting newer versions of files, and using the NSA method for wipe and delete.

Remove Empty Directories (RED)
Website: https://sourceforge.net/projects/rem-empty-dir/

Description: RED searches and deletes empty directories recursively below a given start folder and shows the result in a tree. Users can create custom rules for keeping and deleting folders. Empty files in directories can also be ignored.

TeraCopy
Website: https://www.codesector.com/teracopy

Description: TeraCopy is a file transfer utility for Windows. Like FastCopy, it can be used to complete bulk copy/move/delete operations on files. TeraCopy also includes advanced features like file verification and error detection. It can also be set as the default copy handler on your computer, confirming drag-and-drop file moving operations.

Duplicate file finding and deduplication

File duplicates are a common records management issue. Often, several copies of the same files will exist, occupying valuable file storage and making it difficult to identify records. The tools in this section can be used to identify and remove file duplicates. While some tools can be used for any kind of file, others specialize in duplicate image files.

Auslogics Duplicate File Finder
Website: https://www.auslogics.com/en/software/duplicate-file-finder/

Description: Auslogics Duplicate File Finder is a deduplication application for Windows. Duplicate File Finder can only be used for local files or files on removable media (such as a USB flash drive) – it cannot be used to detect duplicate files on network drives. Duplicate File Finder has an MD5 search engine which allows the program to search for duplicate files by content, regardless of other match criteria.

SimilarImages
Website: https://tn123.org/simimages/

Description: SimilarImages is a utility program to analyze and search large media collections (images/videos) for near duplicates. Near duplicate images may show the same image but be in different file formats or have different compression levels. SimilarImages can also be used to identify images that contain very similar content, such as several photographs taken of a speaker at an event.

SimilarImages first analyzes a file, generating a color/location footprint of a normalized thumbnail image of a file, and then compares these footprints. Analyzation results will be cached and stored on disk, so that subsequent runs become faster.

VisiPics
Website: http://www.visipics.info/index.php?title=Main_Page

Description: VisiPics is a duplicate image finder that incorporates customizable filters in an easy to use interface. It includes easy browsing and an Auto Select feature that allows you to preselect images for deletion based on rules. Like SimilarImages, VisiPics is useful for identifying near duplicate images.

VisiPics supports fewer image formats than SimilarImages, but it has a slightly more user-friendly interface and runs faster.

Disk space analysis

Disk space analyzers are applications that let you understand how folders and files are structured on your disks. These tools can help you quickly identify which folders and files are occupying the most amount of disk space, often using Treemap visualizations. These tools can also often generate customizable reports and file inventories.

SpaceSniffer
Website: http://www.uderzo.it/main_products/space_sniffer/

Description: SpaceSniffer is a free disk space analyzer for Windows. It utilizes Treemap visualizations to display disk space usage. Through SpaceSniffer you can visually identify key folders and useless files. SpaceSniffer has several customization options, export functionality, and allows you to edit files and directories in-application.

TreeSize Free
Website: https://www.jam-software.de/treesize_free/index.shtml

Description: TreeSize Free is a free disk space analyzer for Windows. It utilizes Treemap visualizations to display disk space usage. This tool is the free version of TreeSize Professional, so it does not include features like deduplication or support for Windows Servers. However, it does feature a user-friendly interface and extensive user documentation.

Image viewers

IrfanView
Website: http://www.irfanview.com/

Description: IrfanView is a very fast and compact image viewer and editor for Windows. It includes capabilities for image manipulation and editing, including batch conversion and fast directory browsing. IrfanView can be used to look for images that may have escaped deduplication efforts, images that are poor quality, and irrelevant images.

Integrity checking

Like physical records, digital files are susceptible to a number of risks. Unlike physical records, these risks are not always visible to the naked eye. Storage device aging and failure can result in data degradation and corruption, and human error can sometime result in accidental changes or deletion. These changes can compromise the integrity of a record.

There are several tools that can be used to verify the integrity of files. These tools use checksums, unique "fingerprints" generated by algorithms, to look for data corruption or other changes to files.

ExactFile
Website: http://www.exactfile.com/

Description: ExactFile is a file integrity verification tool. In ExactFile, you can create a "digest", or report, of checksums for all the files in a folder, and you can also include all the files in all the subfolders within that folder. By default, ExactFile produces a report that is stored in that folder alongside the files. This report can be used to verify file integrity in the future and ensure that no files have been corrupted, using either ExactFile or another utility.

Fixity
Website: https://www.weareavp.com/products/fixity/

Description: Fixity is a utility for the documentation and regular review of stored files. Fixity scans a folder or directory, creating a manifest of the files including their filepaths and their checksums, against which a regular comparative analysis can be run. Fixity monitors file integrity through generation and validation of checksums, and file attendance through monitoring and reporting on new, missing, moved and renamed files. Monitoring is automatically performed at regular intervals, scheduled by the user.

Fixity emails a report to the user documenting flagged items along with the reason for a flag, such as that a file has been moved to a new location in the directory, has been edited, or has failed a checksum comparison for other reasons.