

Quick Reference

Using Freely Available Tools to Protect Your Digital Content

Tool Name	Used For	Description	Windows Download	Mac OS Download/Alternative Program
DROID	Checksum (<i>MD5</i>) File Format Identification	Generates MD5 checksums. Performs batch identification of file formats, and links to central registry for technical information (PRONOM). Generates reports (PDF), or saves results as .CSV. Helps to understand your digital assets, assess risk, and plan for action (migration to archival file formats, for example).	z.umn.edu/k68	Mac OS download available
Exact File	Checksum (<i>MD5, SHA1, SHA256, & more</i>)	Helps maintain file integrity: Use to make sure files haven't been changed or damaged over time, or in the process of moving files from one location to another.	www.exactfile.com	DROID captures MD5 checksums. A search for MD5, Sha-1, or Hash in the App Store brings up many free and paid options.
Hash My Files	Checksum (<i>MD5, SHA1, CRC32, SHA-256, & more</i>) Duplication Detection	Helps maintain file integrity: Use to make sure files haven't been changed or damaged over time, or in the process of moving files from one location to another. Saves checksum values as CSV/Tab-Delimited.	z.umn.edu/k69	DROID captures MD5 checksums. A search for MD5, Sha-1, or Hash in the App Store brings up many free and paid options. A search for "file duplicates" also returns many results.
Karen's Directory Printer	Checksum (MD5) File Format Identification	Helps maintain file integrity by generating MD5 checksums. Prints names of every file on a drive, file size, date/time of last modification, and file attributes. Saves information as .txt file.	z.umn.edu/k6a	DROID captures file format information, selected file/folder metadata, and can capture MD5 checksum values.
Metadata Quality Control (MDQC)	Metadata Quality Control	Compare <i>expected</i> technical (and some non-technical) metadata to <i>actual</i> technical metadata. (E.g. run the tool on a group of files returned to you from a vendor).	z.umn.edu/k6b	Mac OS download available
Duplicate File Finder	Duplication Detection	Find and delete duplicate files. Can compare content byte by byte or match on file name.	www.ashisoft.com	Alternative Program: Araxis Find Duplicate Files: z.umn.edu/k6c A search in the App Store for "file duplicates" returns both free and paid options.
Windows Directory Statistics (WinDirStat)	File Visualization	Helps visualize the contents of selected drives and learn more about the properties of the content, and allows for deletion of content.	windirstat.info	Alternative Programs: Disk Inventory X: www.derlien.com GrandPerspective: grandperspectiv.sourceforge.net

Using Freely Available Tools to Protect Your Digital Content Library Technology Conference March 2014



DROID Resource Guide

Tool Name: DROID (Digital Record Object IDentification)

Functionality/Use: The main purpose of DROID is to identify (or confirm) file formats. During this process additional file metadata is captured and recorded in reports. Checksums can also be run and used to find duplicate files. (Windows and MacOS)

To Install:

(Windows)

- Download from The National Archives at:
<http://www.nationalarchives.gov.uk/information-management/our-services/dc-file-profiling-tool.htm> (droid-binary-6.1.3-bin.zip)
- Unzip the folder (if needed).

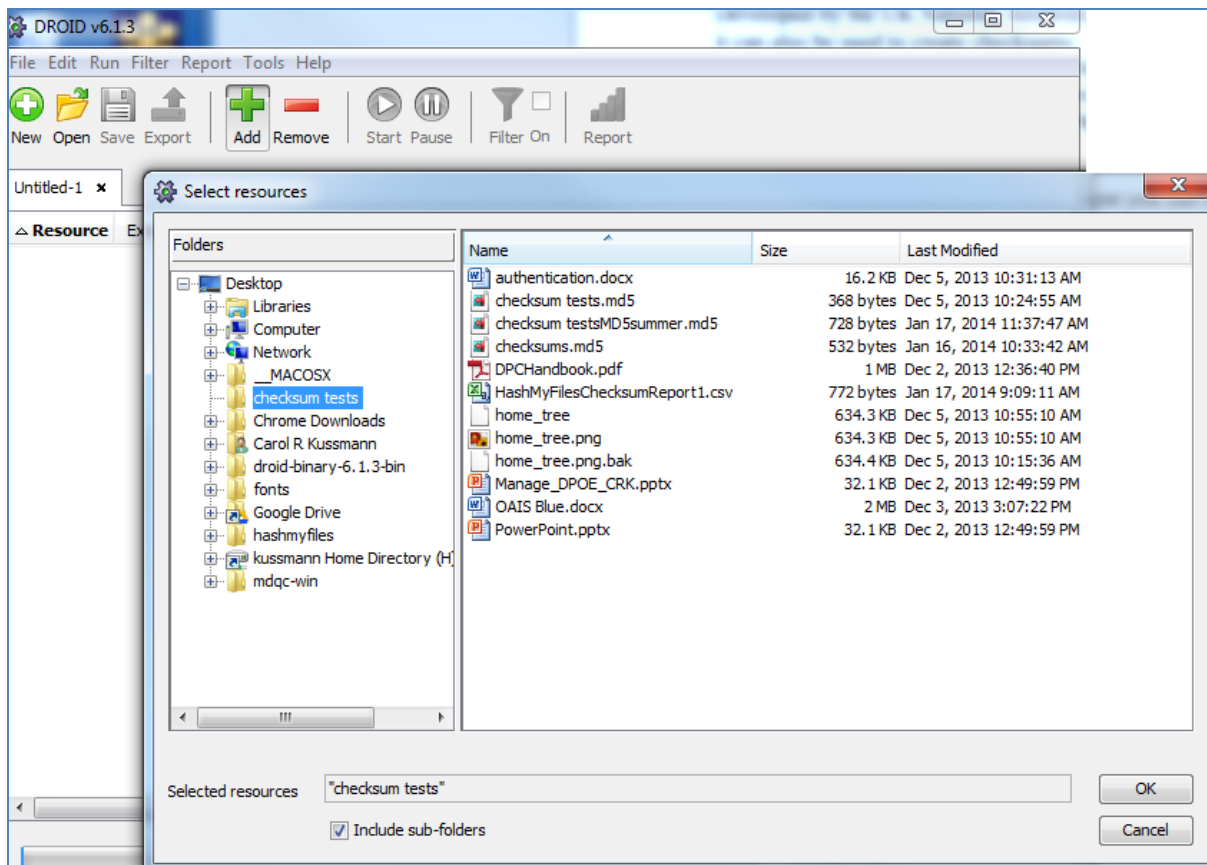
(Mac OS)

- Download from The National Archives at:
<http://www.nationalarchives.gov.uk/information-management/our-services/dc-file-profiling-tool.htm> (droid-binary-6.1.3-bin.zip)
- Unzip the folder (if needed).
- Move the folder to a location in which you can navigate to easily with the “Terminal”.
- Set the Executable Permissions
 - o Open the Running DROID.txt file for instructions/commands.
 - o Open the Terminal
 - o Navigate to the folder that contains the *droid.sh*, *droid-command-line-6.1.3-jar*, and *droid-ui-6.1.3.jar* files.
 - `cd droid-binary-6.1.3-bin` [is all I had to do with the folder directly under my home folder]
 - o Type in:
 - `chmod a+x droid.sh`
 - `chmod a+x droid-ui-6.1.3.jar`
 - `chmod a+x droid-command-line-6.1.3.jar`
 - (If double clicking the jar files does not work with these three commands, try typing in `chmod a+x *`)

To Use:

- Open the droid-binary-6.1.3-bin folder.
- Double click on the *droid.bat* file for Windows and *droid-ui-6.1.3.jar* for Mac OS.
 - o Mac Users: If double clicking the *droid-ui-6.1.3.jar* file does not work and you have also typed in the `chmod a+x *`, you can also start it from the terminal.

- Open the Terminal
 - Navigate to the droid-binary-6.1.3-bin directory
 - Type in “./droid.sh”
- Choose your settings
 - Under the Tools Menu, choose Preferences
 - If you want to calculate MD5 hash values, click the box
 - Note: This will take longer to run.
 - Adjust other settings as desired.
 - Add Folders/Files to Analyze
 - Add the folder you wish to identify by clicking the Add Button (Green Plus)
 - Navigate to the folder
 - Make sure the “Include Subfolder” box is checked
 - Click OK



- Save the Profile
 - Click the Disk/Save button.

- Name the profile
- Save the file (to a location that you can find easily – i.e. desktop)
- Run DROID
 - Click the blue Play/Start button
 - Wait for it to complete (status bar at the bottom)
- Results
 - Results are displayed in table form on the screen

The screenshot shows the DROID v6.1.3 application window. The title bar reads 'DROID v6.1.3'. The menu bar includes 'File', 'Edit', 'Run', 'Filter', 'Report', 'Tools', and 'Help'. The toolbar contains icons for 'New', 'Open', 'Save', 'Export', 'Add', 'Remove', 'Start', 'Pause', 'Filter On', and 'Report'. Below the toolbar, a tab labeled 'DROIDTest1' is active. The main area displays a table with the following columns: Resource, Extension, Size, Last modified, Ids, Format, PUID, Method, Version, Mime type, and Hash. The table lists various files from the desktop, including authentication.docx, checksum tests.md5, DPCHandbook.pdf, HashMyFilesChecksumRe..., home_tree, home_tree.png, home_tree.png.bak, Manage_DPOE_CRK.pptx, and OAIS Blue.docx.

Resource	Extension	Size	Last modified	Ids	Format	PUID	Method	Version	Mime type	Hash
C:\Users\corr0013\Desktop\...			1/17/14 11:39 AM							
authentication.docx	docx	16.2 KB	12/5/13 10:31 AM		Microsoft Word for Windows	fmt/412	Container	2007 onwards	application/v...	
checksum tests.md5	md5	368 bytes	12/5/13 10:24 AM							
checksum testsMD5summ...	md5	728 bytes	1/17/14 11:37 AM							
checksums.md5	md5	532 bytes	1/16/14 10:33 AM							
DPCHandbook.pdf	pdf	1 MB	12/2/13 12:36 PM		Acrobat PDF 1.3 - Portable Docu...	fmt/17	Signature	1.3	application/pdf	
HashMyFilesChecksumRe...	csv	772 bytes	1/17/14 9:09 AM		Comma Separated Values	x-fmt/18	Extension		text/csv	
home_tree		634.3 KB	12/5/13 10:55 AM		Portable Network Graphics	fmt/11	Signature	1.0	image/png	
home_tree.png	png	634.3 KB	12/5/13 10:55 AM		Portable Network Graphics	fmt/11	Signature	1.0	image/png	
home_tree.png.bak	bak	634.4 KB	12/5/13 10:15 AM		Portable Network Graphics	fmt/11	Signature	1.0	image/png	
Manage_DPOE_CRK.pptx	pptx	32.1 KB	12/2/13 12:49 PM		Microsoft Powerpoint for Windows	fmt/215	Container	2007 onwards	application/v...	
OAIS Blue.docx	docx	2 MB	12/3/13 3:07 PM		Microsoft Word for Windows	fmt/412	Container	2007 onwards	application/v...	

- Results can be exported as .csv or xml
 - Export Results as .csv or .xml
 - Click the Export button
 - Select the profile to export
 - Select One Row Per File
 - Click Export Profiles
 - Save with appropriate file name and extension in appropriate location (e.g.: type .csv if you want to export as a comma delimited text file)
 - Click OK
- Results can be used to create reports on certain properties
 - Create Reports
 - Various reports can be created under the Report Icon. For example, to create a report sorted by file extension, follow the steps below.
 - Click the Report Icon
 - Select the profile to run the report on
 - Select the *File Count and Sizes by File Extension Report* from the drop down list
 - Click *Report on Profiles* button
 - A report will be generated

- To save the report, click Export
 - Name the Report, Save to appropriate location
- Re-Save Profile
 - If you would like to reuse the same profile again, remember to save it. (You will be prompted to save it when you close the program.)

Additional Resources for More Information

- UK National Archives Documentation and video:
<http://www.nationalarchives.gov.uk/documents/information-management/droid-how-to-use-it-and-interpret-results.pdf> ; <http://vimeo.com/24718678>
- Practical E-Records Blog Posts: Using DROID for Appraisal: <http://e-records.chrisprom.com/using-droid-for-appraisal/>; <http://e-records.chrisprom.com/more-on-using-droid-for-appraisal/>

Using Freely Available Tools to Protect Your Digital Content Library Technology Conference March 2014



Exact File Resource Guide

Tool Name: Exact File

Functionality/Use: Create and verify checksum values for monitoring files over time.
(Windows only)

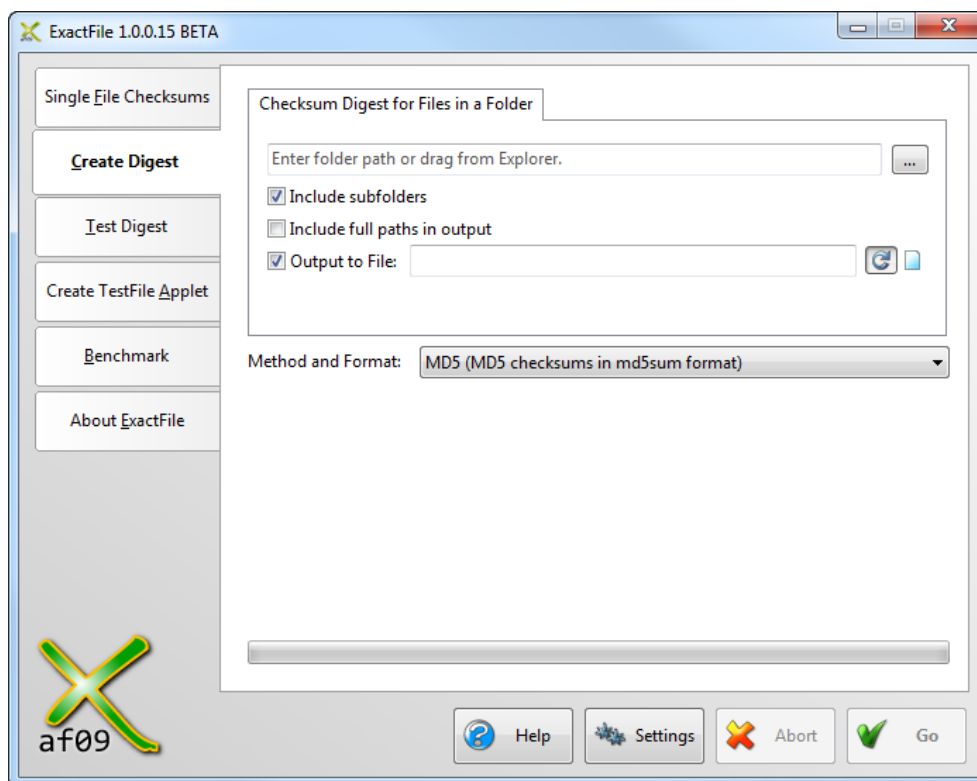
To Install:


- Download at: <http://www.exactfile.com/> (ExactFile-Setup.exe)
- Double click file and select Run
- Type in Admin User Name and Password
- Work through Setup Wizard


To Use:

This program has multiple uses. You are able to run checksums on single files, a group of files, or associate checksums with files copied to and from CDs/DVDs. These instructions explain how to run checksums for a group of files.

- Select the Create Digest Tab.



- Browse to the folder you want to run the checksums on  click OK
- Choose if you will include subfolders, include full paths in output, and the name and location in which to output the checksum file.
 - o Recommendations: Include subfolders; don't include full paths; the output file will automatically be called "checksum" which works well, especially if you keep the file in the folder in which you are running the checksums. (Note: This is the default setting.)
- Select from the dropdown the Method and Format of the checksum
 - o MD5, SHA-1, SHA-256 are often used

MD5 (MD5 checksums in md5sum format) ▼
- Click Go 
- A report will be produced that provides information on when the checksums were run, the checksum value, and the file name. Review if desired, and close it. This file is saved as `checksums.md5` in the location of your choice (folder you ran the checksums on).

```

; Checksums generated by ExactFile 1.0.0.15
; http://www.exactfile.com
; 1/16/2014 10:33:42 AM

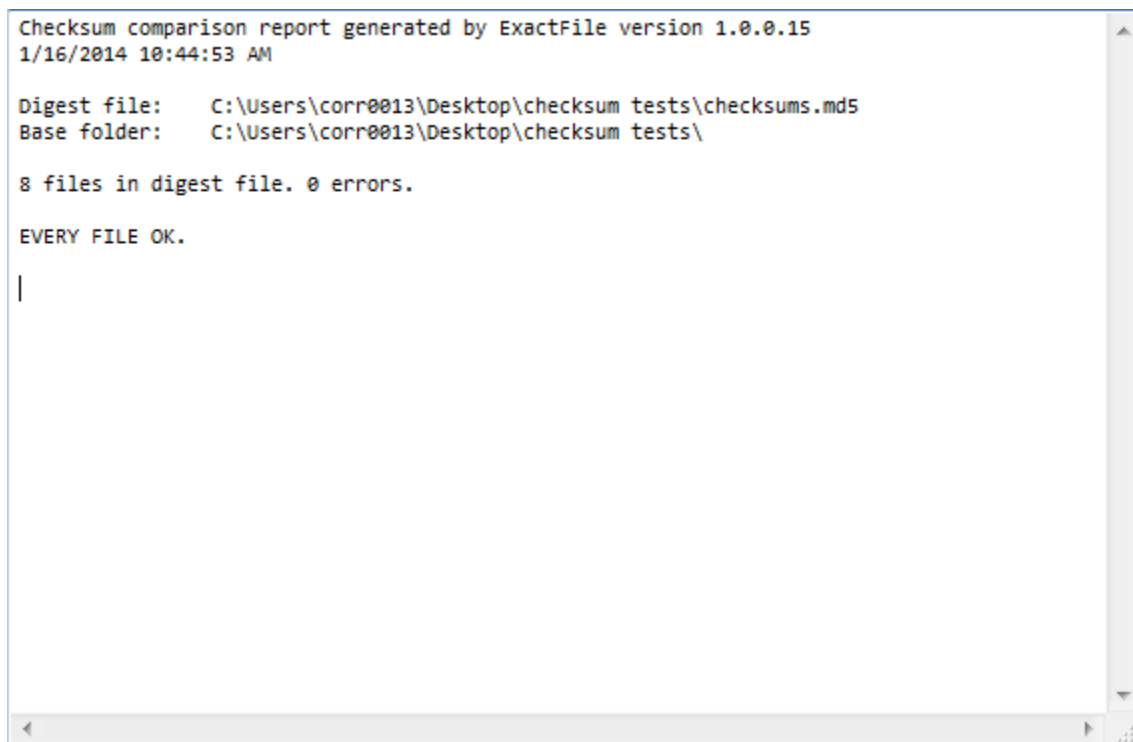
a34424592e896ebe78de333c343605c7 *authentication.docx
e12783f8564468b36ec7004c9dcd0046 *checksum tests.md5
9ba58ef967ded23cf63c9c0ae268db72 *DPCHandbook.pdf
999bd5d23cebbfb9a153ccbd13580b32 *home_tree.png
164f3e97e43c8104eef513662b414f68 *home_tree.png.bak
37f4cbdf03aab3c70632683e927570e4 *Manage_DPOE_CRK.pptx
fe1e8c923461f2177262b850ddf7df51 *OAIS Blue.docx
4d33af964583e2494f83dac66bd1bcc4 *Thumbs.db

; 8 files hashed.

```

To Verify Checksums

- Navigate to the folder you want to verify checksums on
- Double click the “checksum.md5” file
 - o The file should open with ExactFile, but if not you can right click on the file and use “Open With”, and then navigate to Exact File. (If you need to browse to Exact File, it might be in the Program File (x86) folder under Exact File.)
- Once the file opens, it re-runs the checksum to see if any of the files have changed. A report will appear that says “Every File OK” or explains the errors if there are any.



```
Checksum comparison report generated by ExactFile version 1.0.0.15
1/16/2014 10:44:53 AM

Digest file:    C:\Users\corr0013\Desktop\checksum tests\checksums.md5
Base folder:    C:\Users\corr0013\Desktop\checksum tests\

8 files in digest file. 0 errors.

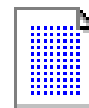
EVERY FILE OK.

|
```

Another guide with additional information can be found here:

http://www.mnhs.org/preserve/records/legislativerecords/carol/docs_pdfs/ExactFileEvaluation.pdf

Using Freely Available Tools to Protect Your Digital Content Library Technology Conference March 2014



Hash My Files Resource Guide

Tool Name: Hash My Files

Functionality/Use: Create checksum value/s for monitoring files over time. The program can calculate multiple checksums at the same time. The program can also identify duplicate files. (Windows only)

To Install:

- Download at NirSoft: http://www.nirsoft.net/utis/hash_my_files.html (the download link can be found 2/3 of the way down the page in light purple)
- Unzip 'hashmyfile.zip'

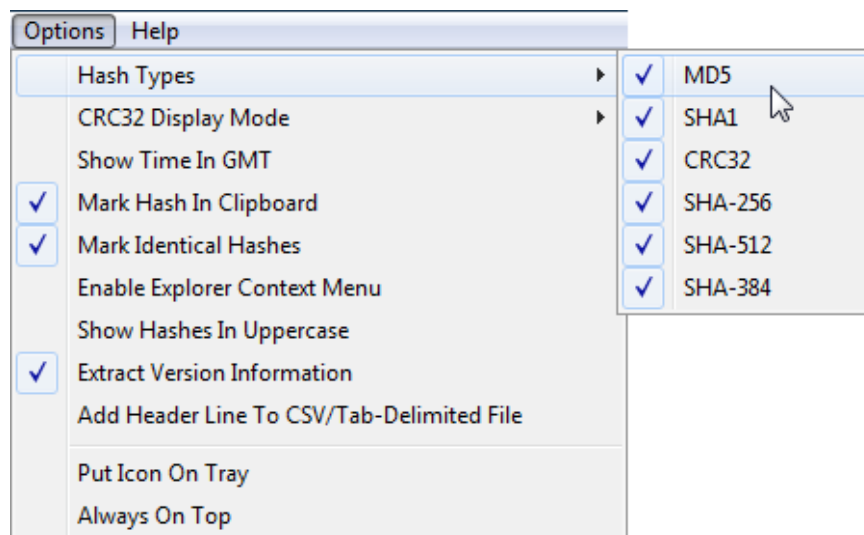
To Use:

- Open the hashmyfiles folder and double click HashMyFile.exe

Selecting Hash Types and Other Options

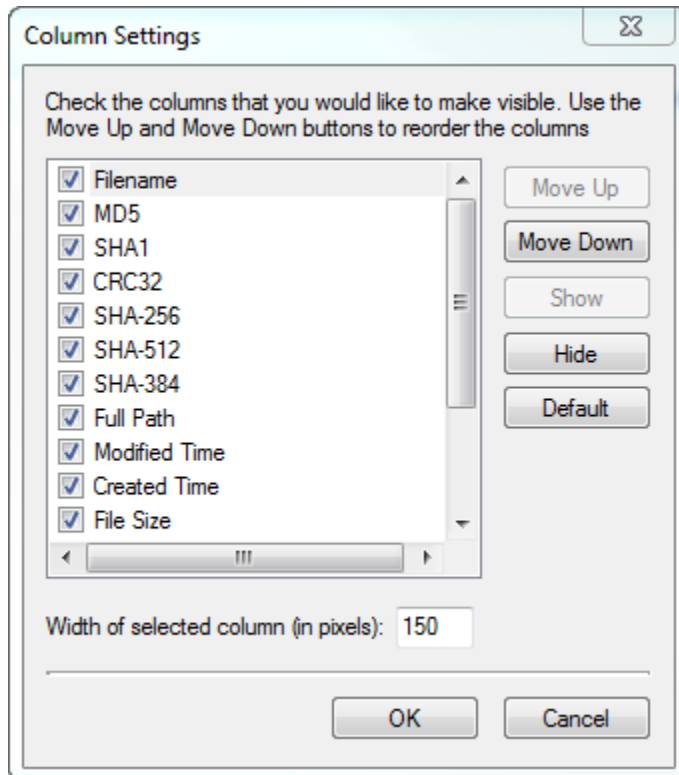
Before you run the program, you may want to take time and choose your settings.

- Under Options, go to Hash Types, and select only the checksum algorithms you want to use. You can only uncheck one box at a time; you will have to navigate to the menu for each one you want to select or remove.
- To take advantage of identifying duplicate files, make sure the "Mark Identical Hashes" is checked.



Choosing Columns to View and for Reports

- Under View, go to Choose Columns, then select the items you want to include on the screen as well as in a CSV report/file. You can also order the columns here by highlighting the item you want to move and click the Move Up or Move Down buttons appropriately.



- The resulting header bar might look like this:

Filename	MD5	Extension	File Size	Identical	File Attributes
----------	-----	-----------	-----------	-----------	-----------------

- After the settings are the way you want them to be you can select files on which to calculate hash values for.

Calculate Hash Values

[There are many ways to do this, only one way is shown below.]

- Go to File, Add Folder
- Browse to the folder you want to add
- Select Add Files in Subfolders to capture any folder hierarchy
- Click OK

The checksums will be calculated and information collected displayed. One added benefit of this program is it visually demonstrates if duplicates are found by color coding and labeling the duplicate files.

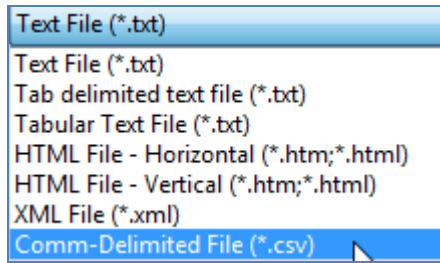
All headers can be used to sort the information.

Filename	MD5	Extension	File Size	File Version	Identical
authentication.docx	a34424592e896ebe78de333c343605c7	docx	16,566		
checksum tests.md5	e12783f8564468b36ec7004c9dcd0046	md5	368		
checksums.md5	d9c7425ac1b253202794ea9648fff76f	md5	532		
DPCHandbook.pdf	9ba58ef967ded23cf63c9c0ae268db72	pdf	1,067,906		
home_tree	999bd5d23cebbfb9a153ccbd13580b32		649,567		1
home_tree.png	999bd5d23cebbfb9a153ccbd13580b32	png	649,567		1
home_tree.png.bak	164f3e97e43c8104eef513662b414f68	bak	649,601		
Manage_DPOE_CRK.pptx	37f4cbdf03aab3c70632683e927570e4	pptx	32,823		2
OAIS Blue.docx	fe1e8c923461f2177262b850ddf7df51	docx	2,079,281		
PowerPoint.pptx	37f4cbdf03aab3c70632683e927570e4	pptx	32,823		2
Thumbs.db	4d33af964583e2494f83dac66bd1bcc4	db	84,480		

11 file(s) NirSoft Freeware. <http://www.nirsoft.net>

Save a Report

- Sort the data and move the column order around if you choose.
- [If you want to include headers in a csv file or tab delimited text file, you will need to click the box indicating your preference under Options.]
- Highlight all of the records you want to include in the report (“Ctrl A” on your keyboard to select all items.)
- To save click the Disk Icon or use the File menu and select Save Selected Items.
- In the Save Window
 - o Select the location you want to save the file in
 - o Name the report
 - o Choose a format



- Click Save

A CSV report looks something like this:

	A	B	C	D	E	F	G
1	Filename	MD5	Extension	File Size	File Version	Identical	
2	authentication.docx	a34424592e896ebe78de333c343605c7	docx	16,566			
3	checksum tests.md5	e12783f8564468b36ec7004c9dcd0046	md5	368			
4	checksums.md5	d9c7425ac1b253202794ea9648fff76f	md5	532			
5	DPCHandbook.pdf	9ba58ef967ded23cf63c9c0ae268db72	pdf	1,067,906			
6	home_tree	999bd5d23cebbfb9a153ccbd13580b32		649,567			1
7	home_tree.png	999bd5d23cebbfb9a153ccbd13580b32	png	649,567			1
8	home_tree.png.bak	164f3e97e43c8104eef513662b414f68	bak	649,601			
9	Manage_DPOE_CRK.pptx	37f4cbdf03aab3c70632683e927570e4	pptx	32,823			2
10	OAIS Blue.docx	fe1e8c923461f2177262b850ddf7df51	docx	2,079,281			
11	PowerPoint.pptx	37f4cbdf03aab3c70632683e927570e4	pptx	32,823			2
12	Thumbs.db	4d33af964583e2494f83dac66bd1bcc4	db	84,480			
13							
14							

Another guide with additional information can be found here:

http://www.mnhs.org/preserve/records/legislativerecords/carol/docs_pdfs/HashMyFilesEvaluation.pdf



Using Freely Available Tools to Protect Your Digital Content Library Technology Conference March 2014

Karen's Directory Printer Resource Guide

Tool Name: Karen's Directory Printer

Functionality/Use: Document selected properties of files and folders within a selected directory. In addition to file/folder level metadata, the program can also be used to run checksums. These properties could be used for documenting provenance, fixity, and capturing other information as necessary. (Windows only)

To Install:

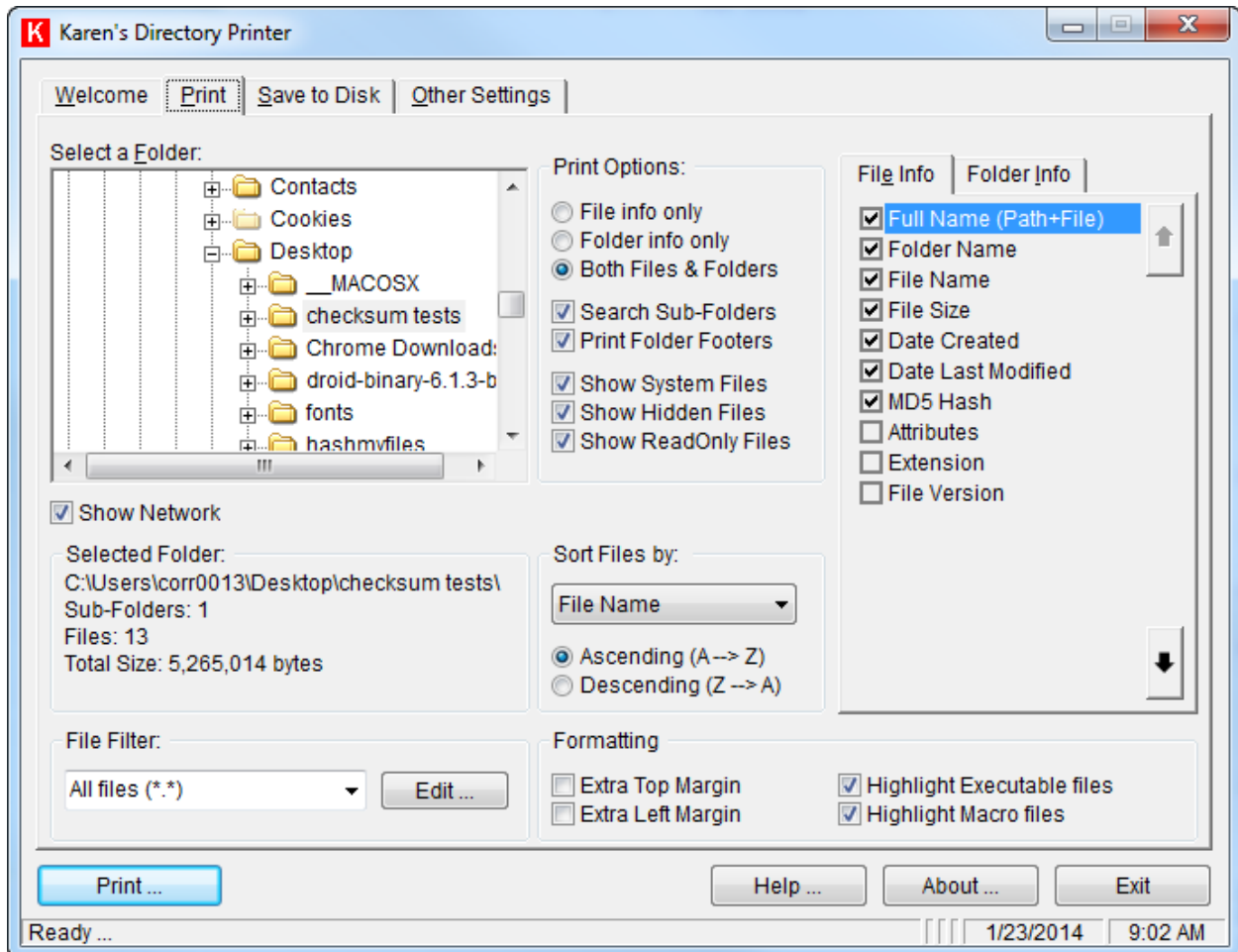
- Download from Karen's Power Tools:
<http://www.karenware.com/powertools/ptdirprn.asp>
- (ptdirprn-setup.exe)
- Double click on the file
- Type in Admin User Name and Password
- Work through the Setup Wizard

To Use:

You can **PRINT** a report or **SAVE** a report. You must choose/set the settings for each option. (You can always print a saved report.)

- Open the program by double clicking on DirPrn.exe
- Set Other Settings
 - o Click the Other Settings tab
 - o Select which File and Folder Information you would like to collect when you print or save a report
 - o Choose your settings (and save them if desired by checking the box)
- To PRINT
 - o Set Search and Print Options
 - Select the folder / location you will run the program on.
 - Navigate to the folder that you would like to start the program in
 - Sort Files if desired (File Filter)
 - Set Print Options
 - Choose Search Sub-Folders for the most complete picture
 - Sort Files if desired.
 - Choose File Info and Folder Info
 - The lists pull options based on the items selected on the Other Settings tab.

- Use the arrows to order the information elements in the way you want to see them.
- Click Print to print a report



Screen Shot of Print Options (very similar to Save to Disk Options)

- To SAVE TO DISK
 - Set Search and Save To Disk Options
 - Select the folder / location you will run the program on.
 - Navigate to the folder that you would like to start the program in
 - Sort Files if desired (File Filter)
 - Set Print Options
 - Choose Search Sub-Folders for the most complete picture
 - Sort Files if desired.
 - Choose File Info and Folder Info

- The lists pull options based on the items selected on the Other Settings tab.
- Use the arrows to order the information elements in the way you want to see them.
- Click Save to Disk
- Save the report
 - Name the File, choose type (.txt, .tsv, .prn, or .dir), click Save.

```

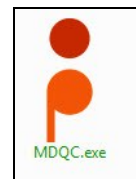
COMMENT Lines beginning with "FILE" contain information about a Single File.
COMMENT Here's the information found in those lines:
COMMENT Attributes <TAB> Date Created <TAB> Date Last Modified <TAB> File Size <TAB> File Name <TAB> Full Name (Path+File) <TAB> Folder Name <TAB> MD5 Hash
COMMENT
COMMENT Lines beginning with "FOLDER" contain information about a Single Folder.
COMMENT Here's the information found in those lines:
COMMENT Folder Name <TAB> Full Name (Path+Folder) <TAB> Folder Size <TAB> Number of Files <TAB> Number of Sub-Folders <TAB> Attributes
COMMENT
COMMENT Lines beginning with "TOTAL" contain information about a Folder, and all its Sub-Folders.
COMMENT Here's the information found in those lines:
COMMENT Folder Name <TAB> Full Name (Path+Folder) <TAB> Folder Size <TAB> Number of Files <TAB> Number of Sub-Folders <TAB> Attributes
COMMENT
COMMENT Computer: ULWIL-517-AB
COMMENT User: kussmann
COMMENT Prepared: 09:20 1/23/2014
COMMENT Folder: C:\Users\corr0013\Desktop\checksum tests\
COMMENT Include Sub-Folders? Yes
COMMENT
FOLDER checksum tests C:\Users\corr0013\Desktop\checksum tests\ 5,265,014 13 1 -----
FILE ---A----- 1/16/2014 10:27 12/5/2013 10:31 16,566 authentication.docx C:\Users\corr0013\Desktop\checksum tests\authentication.docx C:\Users\corr0013\De
FILE ---A----- 1/16/2014 10:27 12/5/2013 10:24 368 checksum tests.md5 C:\Users\corr0013\Desktop\checksum tests\checksum tests.md5 C:\Users\corr0013\De
FILE ---A----- 1/17/2014 11:37 1/17/2014 11:37 728 checksum testsMD5summer.md5 C:\Users\corr0013\Desktop\checksum tests\checksum testsMD5summer.md5 C:\U
FILE ---A----- 1/16/2014 10:33 1/16/2014 10:33 532 checksums.md5 C:\Users\corr0013\Desktop\checksum tests\checksums.md5 C:\Users\corr0013\Desktop\checksum t
FILE ---A----- 1/16/2014 10:27 12/2/2013 12:36 1,067,906 DPCHandbook.pdf C:\Users\corr0013\Desktop\checksum tests\DPCHandbook.pdf C:\Users\corr0013\De
FILE ---A----- 1/17/2014 08:54 1/17/2014 09:09 772 HashMyFilesChecksumReport1.csv C:\Users\corr0013\Desktop\checksum tests\HashMyFilesChecksumReport1.csv C:\U
FILE ---A----- 1/17/2014 08:41 12/5/2013 10:55 649,567 home_tree C:\Users\corr0013\Desktop\checksum tests\home_tree C:\Users\corr0013\Desktop\checksum t
FILE ---A----- 1/16/2014 10:27 12/5/2013 10:55 649,567 home_tree.png C:\Users\corr0013\Desktop\checksum tests\home_tree.png C:\Users\corr0013\Desktop\checksum t
FILE ---A----- 1/16/2014 10:27 12/5/2013 10:15 649,601 home_tree.png.bak C:\Users\corr0013\Desktop\checksum tests\home_tree.png.bak C:\Users\corr0013\De
FILE ---A----- 1/16/2014 10:27 12/2/2013 12:49 32,823 Manage_DPOE_CRK.pptx C:\Users\corr0013\Desktop\checksum tests\Manage_DPOE_CRK.pptx C:\Users\corr0013\De
FILE ---A----- 1/16/2014 10:27 12/3/2013 15:07 2,079,281 OATS Blue.docx C:\Users\corr0013\Desktop\checksum tests\OATS Blue.docx C:\Users\corr0013\Desktop\ch
FILE ---A----- 1/17/2014 08:43 12/2/2013 12:49 32,823 PowerPoint.pptx C:\Users\corr0013\Desktop\checksum tests\PowerPoint.pptx C:\Users\corr0013\Desktop\ch
FILE ---HSA----- 1/16/2014 10:27 1/16/2014 09:42 84,480 Thumbs.db C:\Users\corr0013\Desktop\checksum tests\Thumbs.db C:\Users\corr0013\Desktop\checksum t
FOLDER Extra Files C:\Users\corr0013\Desktop\checksum tests\Extra Files\ 2,367,812 4 1 -----
FILE ---A----- 1/23/2014 09:01 12/2/2013 12:36 1,067,906 DPCHandbook.pdf C:\Users\corr0013\Desktop\checksum tests\Extra Files\DPCHandbook.pdf C:\Users\cor
FILE ---A----- 1/23/2014 09:01 1/17/2014 09:09 772 HashMyFilesChecksumReport1.csv C:\Users\corr0013\Desktop\checksum tests\Extra Files\HashMyFilesChecksumRepr
FILE ---A----- 1/23/2014 09:01 12/5/2013 10:55 649,567 home_tree C:\Users\corr0013\Desktop\checksum tests\Extra Files\home_tree C:\Users\corr0013\Desktop\ch
FILE ---A----- 1/23/2014 09:01 12/5/2013 10:55 649,567 home_tree.png C:\Users\corr0013\Desktop\checksum tests\Extra Files\home_tree.png C:\Users\corr0013\De
FOLDER Duplicates C:\Users\corr0013\Desktop\checksum tests\Extra Files\Duplicates\ 2,794,528 4 0 -----
FILE ---A----- 1/23/2014 09:01 12/5/2013 10:15 649,601 home_tree.png.bak C:\Users\corr0013\Desktop\checksum tests\Extra Files\Duplicates\home_tree.png.bak
FILE ---A----- 1/23/2014 09:01 12/2/2013 12:49 32,823 Manage_DPOE_CRK.pptx C:\Users\corr0013\Desktop\checksum tests\Extra Files\Duplicates\Manage_DPOE_CRK.pptx
FILE ---A----- 1/23/2014 09:01 12/3/2013 15:07 2,079,281 OATS Blue.docx C:\Users\corr0013\Desktop\checksum tests\Extra Files\Duplicates\OATS Blue.docx C:\U
FILE ---A----- 1/23/2014 09:01 12/2/2013 12:49 32,823 PowerPoint.pptx C:\Users\corr0013\Desktop\checksum tests\Extra Files\Duplicates\PowerPoint.pptx C:\Users\cor
TOTAL Duplicates\ C:\Users\corr0013\Desktop\checksum tests\Extra Files\Duplicates\ 2,794,528 4 0 -----
TOTAL Extra Files\ C:\Users\corr0013\Desktop\checksum tests\Extra Files\ 5,162,340 8 1 -----
TOTAL checksum tests\ C:\Users\corr0013\Desktop\checksum tests\ 10,427,354 21 2 -----

```

Screenshot of Save to Disk Report

Additional Resources for More Information

- University of Hull Idiot Guide No. 4: Karen's Directory Printer:
<http://www.hullhistorycentre.org.uk/discover/pdf/Idiots%20Guide%204%20-%20Karens%20Directory%20Printer.pdf>
- Minnesota State Archives User Guide:
http://www.mnhs.org/preserve/records/docs_pdfs/KarensDirectoryPrinter.pdf



Metadata Quality Control Resource Guide

Tool Name: Metadata Quality Control (MDQC)

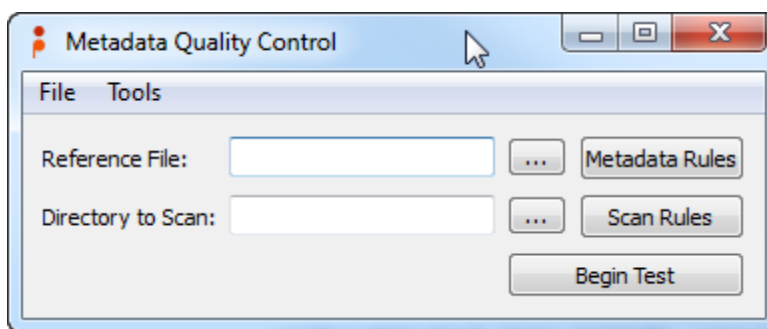
Functionality/Use: Compare metadata fields, for verifying technical and administrative specifications of files. Tool is useful for performing quality control tasks. (Windows and Mac)

To Install:

- Download from AudioVisual Preservation Solutions:
<http://www.avpreserve.com/avpsresources/tools/> (mdqc-win.zip or mdqc-osx.zip)
- Unzip folders (if necessary).

To Use: (Windows based instructions)

- Open the mdqc-win or mdqc-osx folder
- Double click on the MDQC.exe or MDQC.app file



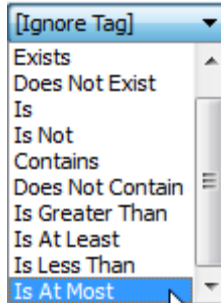
Main Screen

- Point to a Reference File
The metadata fields in the reference file will be used to compare the files in the chosen directory. Make sure this file has the correct information in the fields of concern.
 - o Use the browse button (...) to find the reference file
- Select which tool you would like to use to pull metadata out of the files: ExifTool or MediaInfo (Click on the Tools menu.) MediaInfo only pulls metadata from audio and video files and ExifTool can pull metadata from a variety of other file formats. View a list of ExifTool supported formats:
<http://www.sno.phy.queensu.ca/~phil/exiftool/#supported>

- Set the Metadata Rules

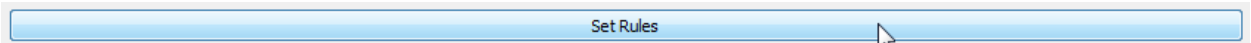
If this does not work on the Mac, see the troubleshooting guide below.

- o Readable metadata fields are pulled out of the reference file and populate a grid.
- o Use the dropdown menu to select how important the field is and what you want the metadata in that field to do:



Drop Down Menu for Metadata Rules

- o When you are done making your selections click the Set Rules button at the top of the screen.



- Point to a Directory to Scan

- o Use the browse button (...) to point to the directory that contains the files you want to compare metadata fields with the reference file's metadata fields.

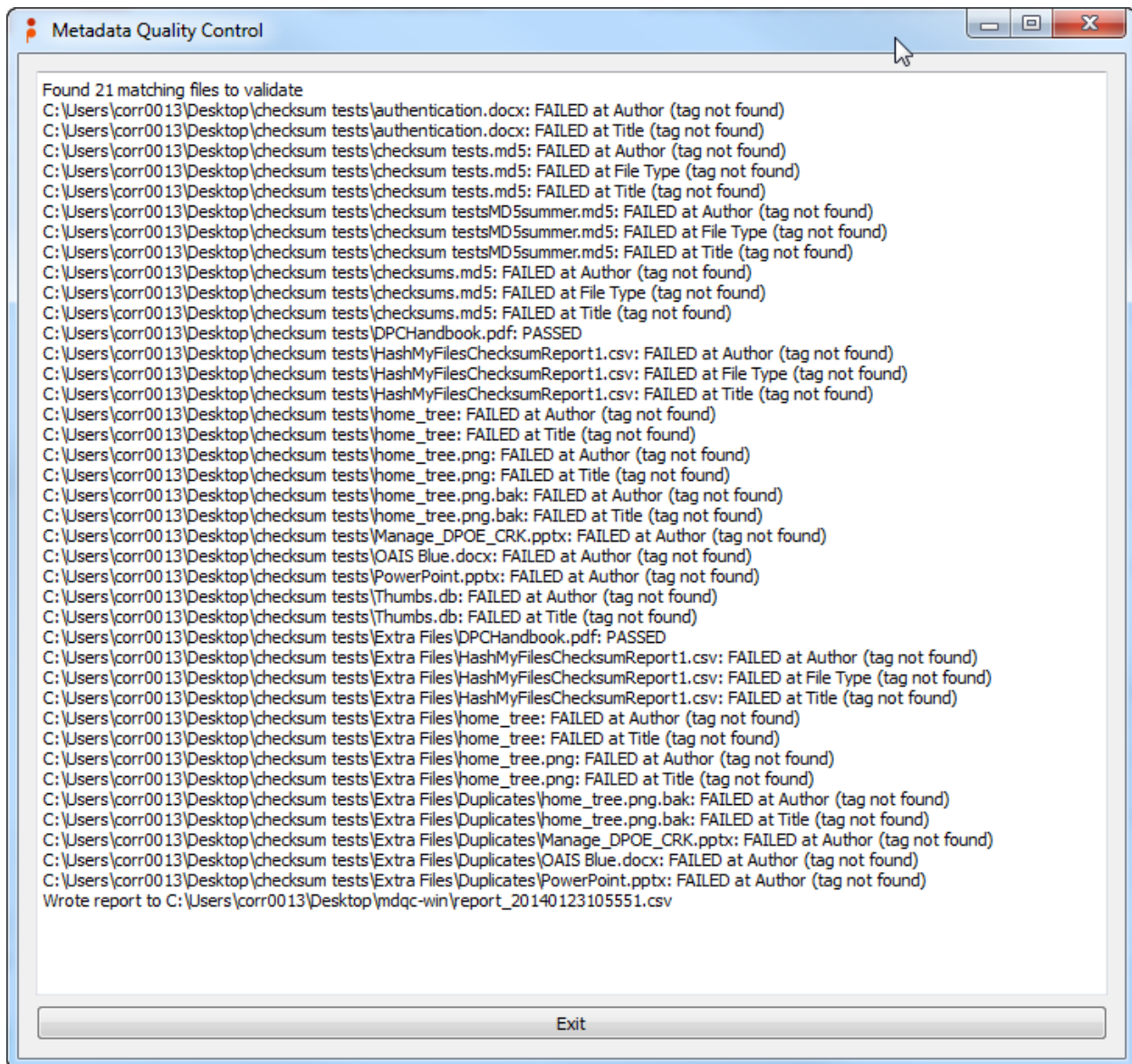
- Set the Scan Rules

- o If you would like to include or exclude files or directories based on what their names contain or do not contain, set these rules.

- Click Begin Test

- A window appears that tells you how many files it found.

- All files are compared to the reference file and an output displays if the file passed or failed and why. (see screenshot below)



On Screen Results

- A report documenting the results is created and saved as a .csv file in the program folder (mdqc-win folder).
- Information on how to use the .csv file is found in the MDQC User Guide below.

Additional Resources for More Information

- Official MDQC User Guide: <http://www.avpreserve.com/wp-content/uploads/2013/10/MDQCUserGuide.pdf>

Trouble Shooting:

Mac:

If you can open MDQC but the Metadata Rules button is not pulling up any information, you might need to install both the ExifTool and MediaInfo.

- ExifTool for OSX: <http://www.sno.phy.queensu.ca/~phil/exiftool/ExifTool-9.47.dmg>
- MediaInfo for OSX:
http://mediaarea.net/download/binary/mediainfo/0.7.66/MediaInfo_CLI_0.7.66_Mac.dmg

Using Freely Available Tools to Protect Your Digital Content Library Technology Conference March 2014



Duplicate File Finder Resource Guide

Tool Name: Duplicate File Finder

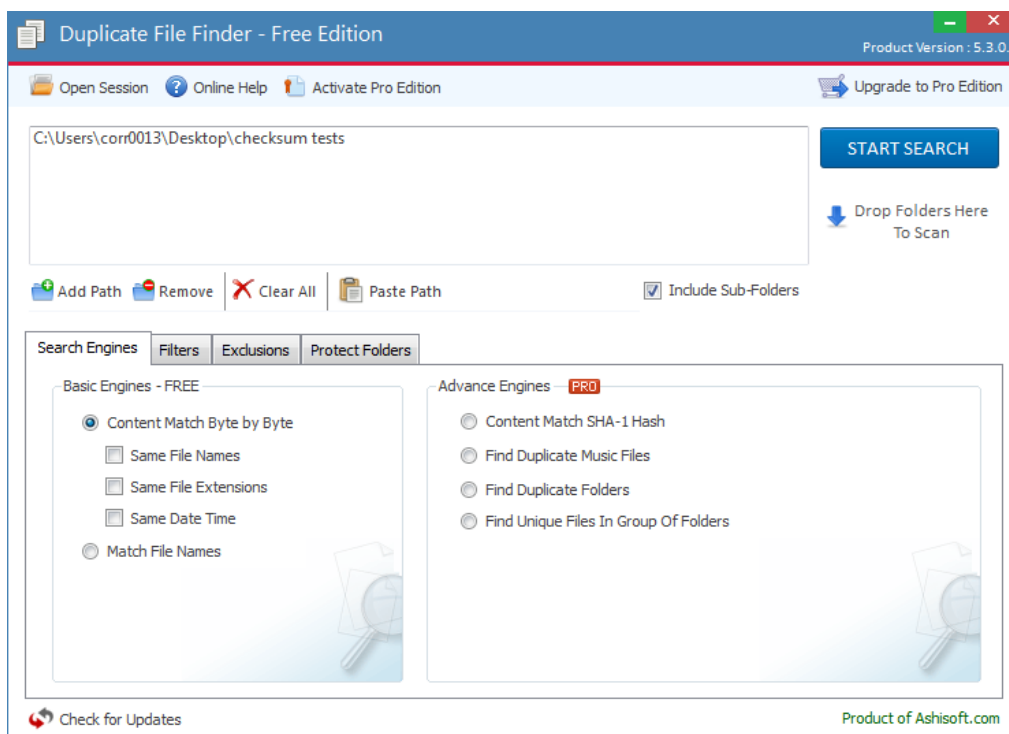
Functionality/Use: Find duplicate files in a selected group of files; you can also choose to immediately delete the files as well. [The free version allows you to find and delete files, but additional features are available in the Pro (paid) Version.] (Windows only)

To Install:

- Download from Ashisoft at: <http://www.ashisoft.com/> (dfsetup.exe)
- Double click file (dfsetup.exe) and select Run
- Type in Admin User Name and Password
- Walk through Setup Wizard

To Use:

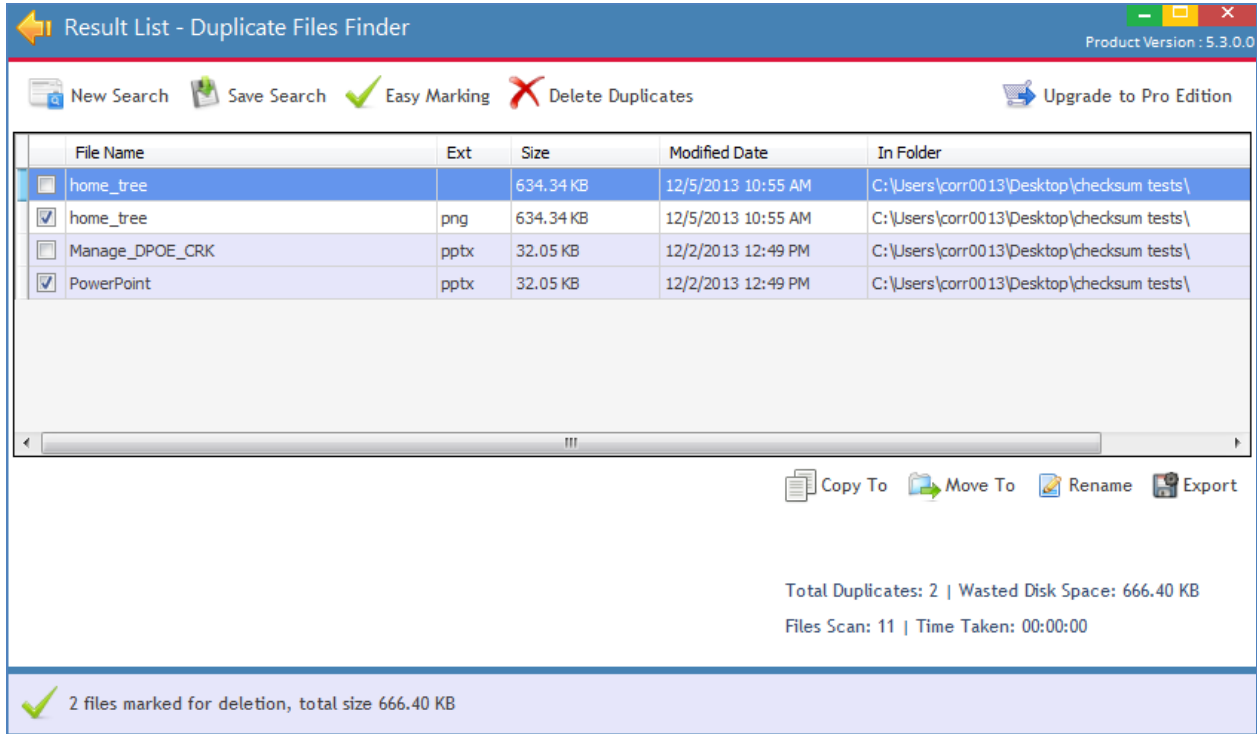
- Double click on program Icon
- Choose the area/s that you want to search for duplicate files (Add Path)
- Choose the method in which to find duplicate files (“Content Match Byte by Byte” is most thorough)
- Click Start Search



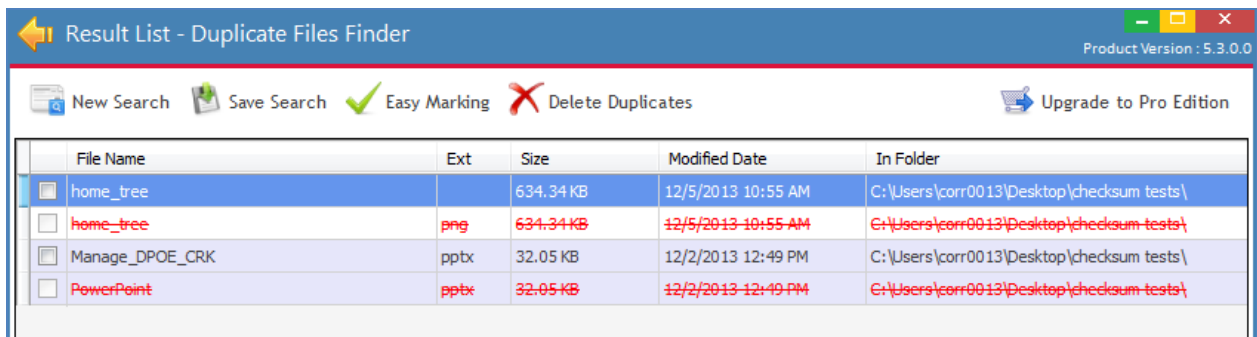
UNIVERSITY OF MINNESOTA

LIBRARIES

- Results will be returned



- Verify files to be deleted (if you want to delete a different file, just check/uncheck the boxes next to the file name). Files with the checkmark will be deleted.
- Delete Duplicates (send to recycle bin or delete immediately)



Additional Resources for More Information

- Online Tutorial from Company: <http://www.ashisoft.com/find-duplicates-step-by-step.htm>
- Guide: http://www.mnhs.org/preserve/records/docs_pdfs/DuplicateFileFinder.pdf

Windows Directory Statistics (WinDirStat) Resource Guide

Tool Name: Windows Directory Statistics

Functionality/Use: “WinDirStat is a disk usage statistics viewer and cleanup tool.” It helps visualize the contents of selected drives, learn more about the properties of selected content, and allows you to delete content. (Windows only. See WinDirStat site for Mac OS program suggestions.)

To Install:

- Download from WinDirStat: <http://windirstat.info/index.html> (windirstat1_1_2_setup.exe)
- Double click on the file
- Type in Admin User Name and Password
- Work through the Setup Wizard

To Use:

- Double click on the program icon
- Choose All Local Drives or Individual Drives
- Click OK

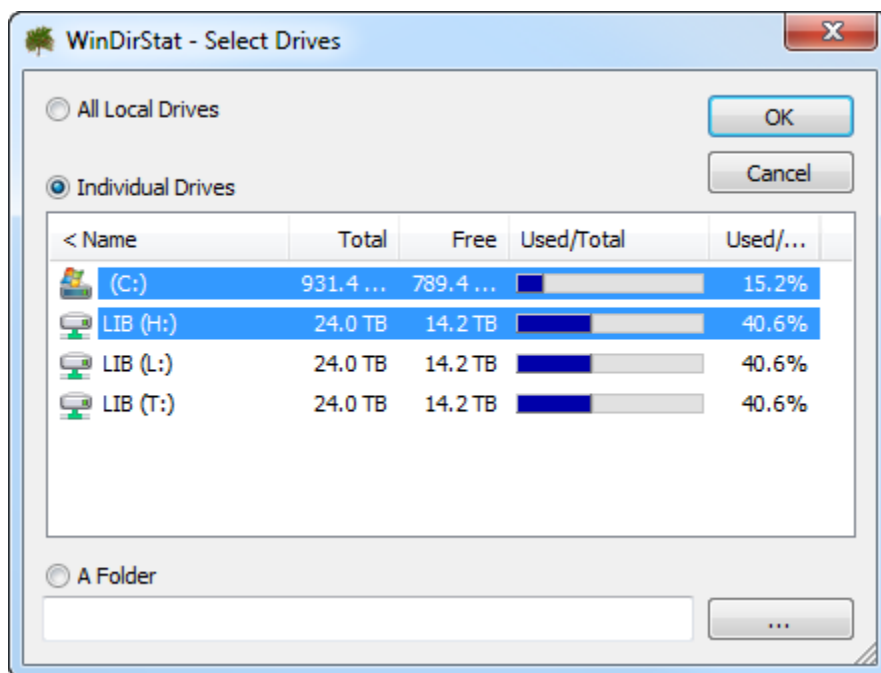
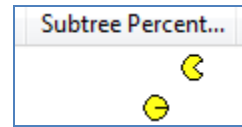


Figure 1: Select Drives



- The program will run. You will see Pacman as the status bar.
- After the drives have been mapped, you will see three areas on the results window.
 - o Directory List (top window similar to windows folder structure, organized by size)
 - o Extension List (right hand window listing file extensions and color legend for the tree map)
 - o Treemap (the colored section at the bottom)

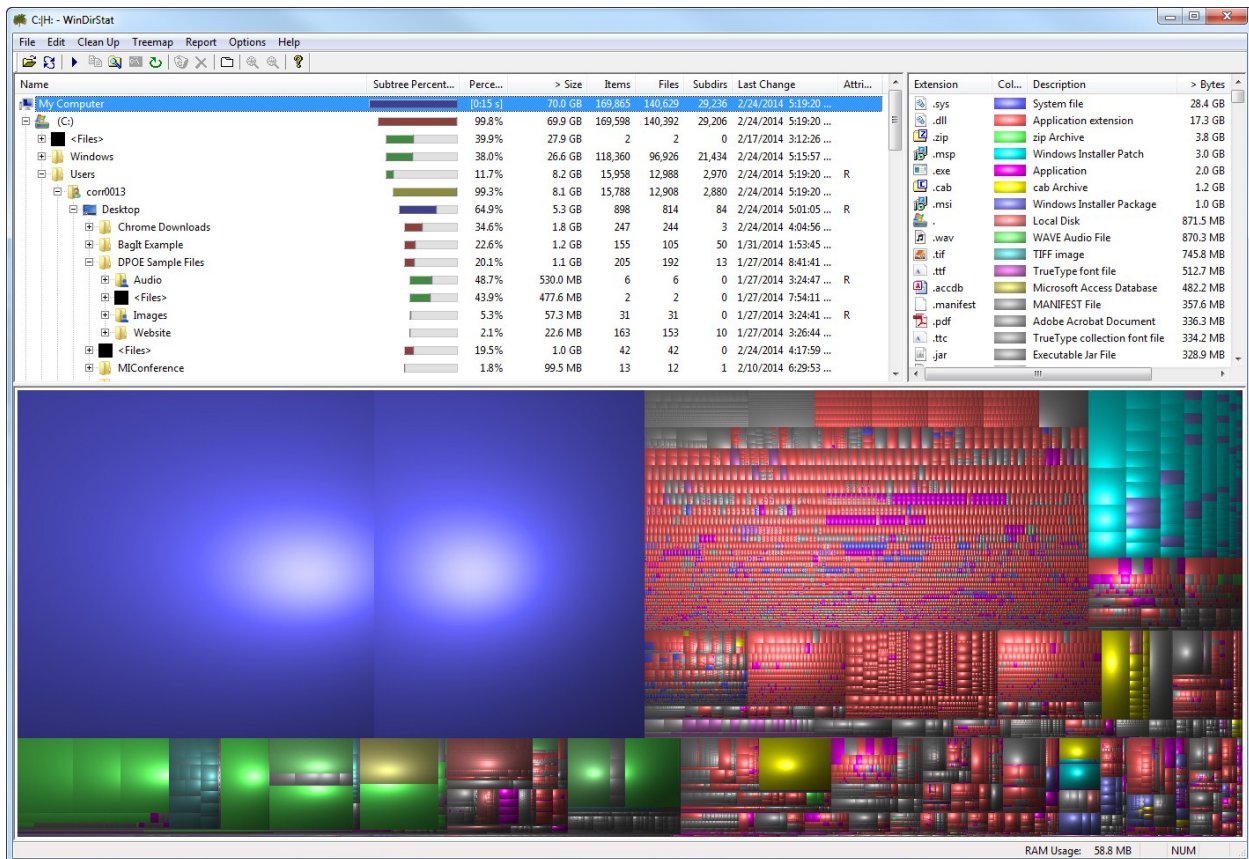


Figure 2: Results Display (Directory List, Extension List, Treemap)

- Use the Directory List to show where a particular set of files can be found and how much space it takes up.

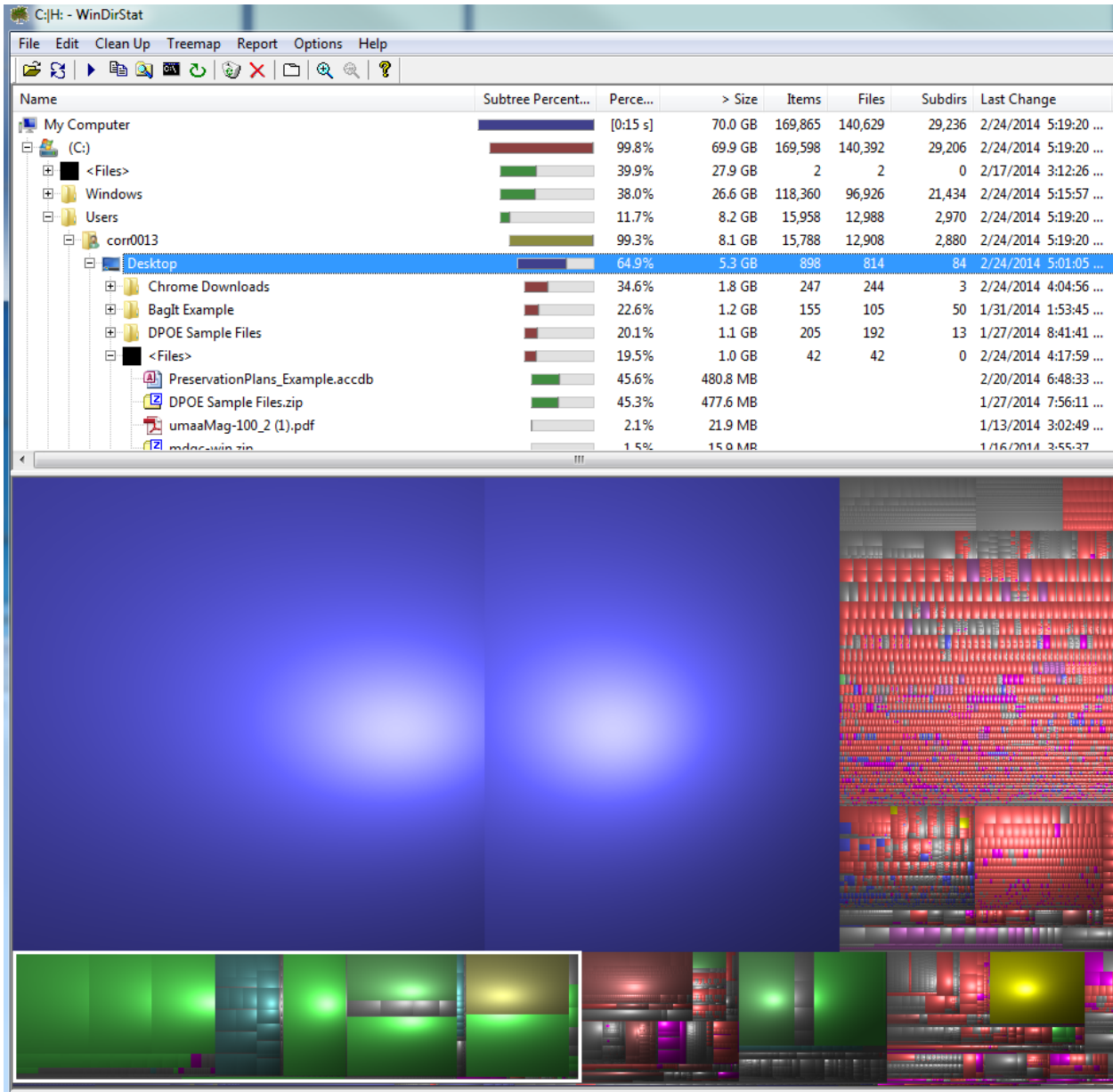


Figure 3: The white highlighted box is my Desktop.
File properties are given in the Directory View.

- Additional file properties can be seen by using the drop down under Clean Up→Properties

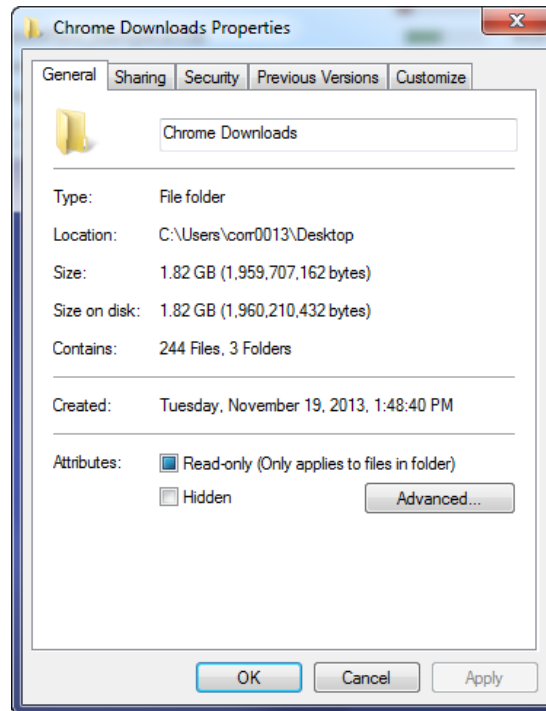


Figure 4: Windows File Property Screen

- Use the Extension List to see the file types you have and visualize them.

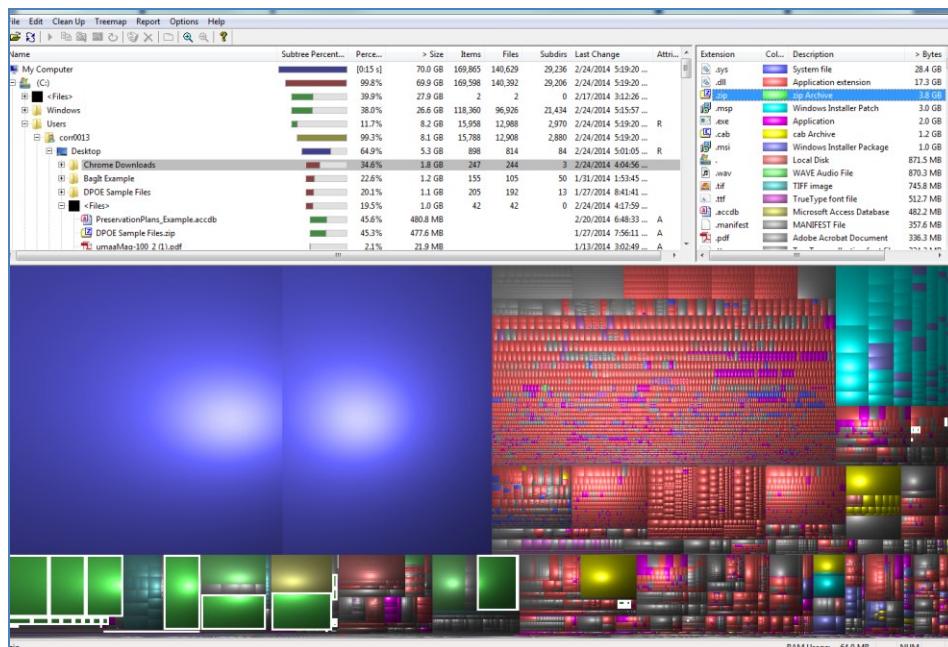


Figure 5: This image shows the .zip files I have (green with the white box around them)

- Use the Treemap to visualize your content.

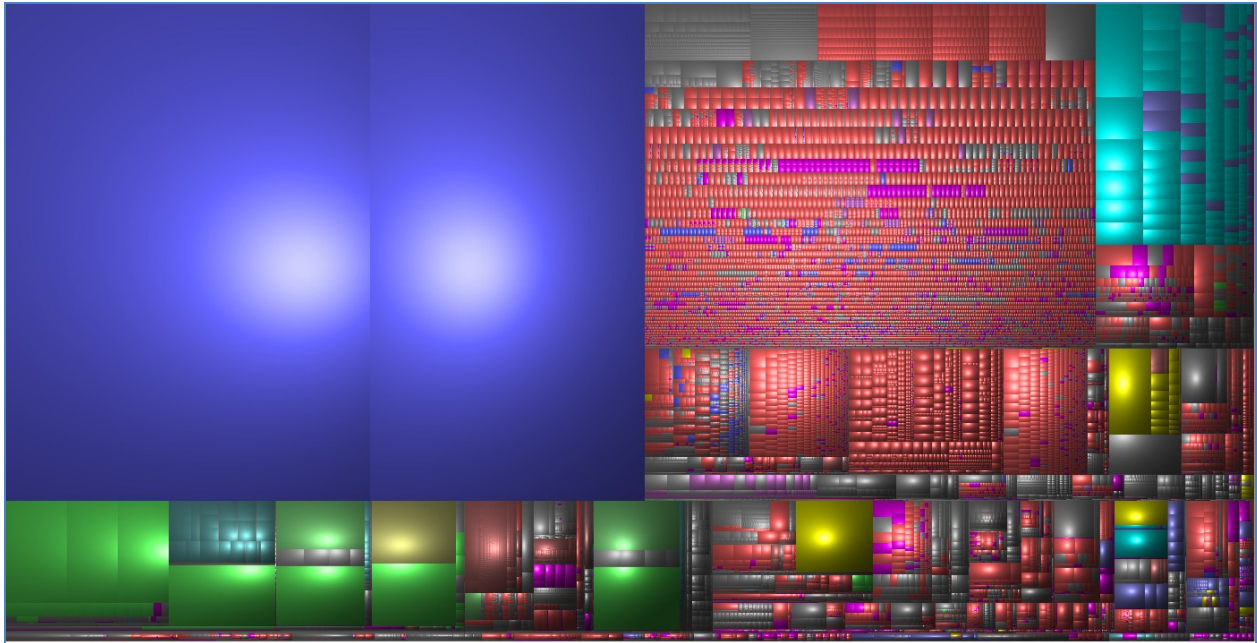


Figure 6: Color coded “Tree Map” that helps you visualize what types of files you have.

- You can also show free space, (Options → Show Free Space)
- You can email reports to users (Using Outlook Profiles?) [Did not try this.]
- You can delete files from the program.
 - o Clean Up → Delete (to Recycle Bin)
 - o Clean Up → Delete (no way to undelete)
- Using other dropdown choices you can:
 - o Open the file location.
 - o Copy the file path.
 - o Open the command line at the desired location.

Additional Resources for More Information

- WinDirStat Website
<http://windirstat.info/index.html>