Case Note

Personal Digital Preservation
Photographs and Video

By Richard Wright

January 2017

'The Digital Preservation Coalition (DPC) Technology Watch Report *Personal Digital Archiving* provides an overall approach and methodology for putting one's "digital house" in order. My purpose in this case note is to add the gory detail about digital photos and video, giving a one-person case study. I say exactly what software and processes I use (and which ones I've considered but rejected, or used and discarded), and try to be frank about successes and failures.'

We have a large, and largely invisible, problem with our digital photographs: there are so many of them! Photos used to cost something like £0.10 per print – and several pounds more for the original unexposed film. We had to be careful how many we took, or had printed. Now, once the camera is bought, the photos are free. Furthermore, with cameras on our mobile phones, laptops, and tablet computers, we have a camera with us almost always – and so we take pictures almost always. We can take hundreds of photos in a day, thousands in one trip, and can accumulate tens of thousands in a decade.

Photo by John E. Payne Re-used with permission from the author. Original image http://www.inpayne.com/family/granpaspix/gppix.html

The Digital Preservation Coalition (DPC) Technology Watch Report *Personal Digital Archiving* provides an overall approach and methodology for putting one's 'digital house' in order. My purpose in this case note is to add the gory detail about digital photos and video, giving a one-person case study. I say exactly what software and processes I use (and which ones I've considered but rejected, or used and discarded), and try to be frank about successes and failures. Note to Apple users: I'm a PC user, so some of the software and processes I describe will not apply to Apple (or to Linux).

In the photo above, we see a collection of 100 glass slides from a century ago, together with the camera used to make the images. These 100 slides are treasured by the great-grandson of the original photographer, each one a precious heirloom.

Unlike an analogue collection, I can't show a picture of my collection of 35,000 (and growing) digital photographs. In this case note I describe my efforts to make my digital collections accessible and intelligible to my own children and grandchildren.
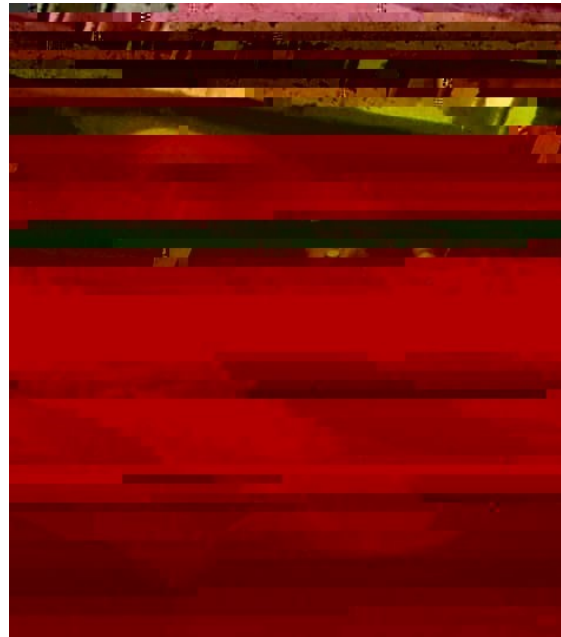
What is the state of all these images I want to preserve for my children and grandchildren? If the files could be rendered into physical form, I imagine the actual condition of the storage and organisation of digital photographs would look like the photo to the right.

The state of digital storage and preservation is largely invisible. It is all too easy for us to continue dumping our photo files in various places. It is all too easy to not worry about – or not even know about – the problems caused by just having a jumble of tens of thousands of photo files.

The condition of the content of the files is also invisible. We do not know whether something is

CC Attribution 2.0 Generic No changes have been made to this image. The author of this image is I <3 happy!! from NY, NY. The image was originally published on Flickr at http://www.flickr.com/photos/65634609@N00/606583152/

rotten in Denmark, because our 'digital Denmark' is invisible and silent. Unless we intervene in some way, we will only discover a problem with a file when we try to view it. Only then – when it may be too late – do we discover that the file is missing, or corrupt, or obsolete.

The images above illustrate just how easily an image in a compressed format (JPG) can be damaged by any sort of an error. On the left is the original, and on the right a version with an error in just one byte (out of 24k). When our photos are just a jumble of files, literally dumped on various kinds of storage, we have no way of knowing which ones remain readable.

Problem One: what have I got?

According to Redwine's Report: 'One of the initial difficulties in gaining control over a body of personal digital files is figuring out how many, and what type, of files exist'.

This is my body of personal digital files:

- 35,000 digital photographs taken over the last 15 years
- my partner's photos: probably another 20,000
- family photos, including those of children, their spouses, and now grandchildren – all taking photos and sending subsets of them to me in various ways
- social media: photos on Facebook and Flickr; some video on YouTube
- The Cloud: mobile devices that automatically upload photos and videos to various services (such as Dropbox, Google Photos, Amazon Cloud, Microsoft OneDrive)
- out there somewhere: photos (and audio, video, and interviews) on various websites and blogs, including websites owned and managed by other people and organisations
- then there are those boxes of pre-2000 real photographs in the form of prints, negatives, and slides (remember slides?).

So the first thing I want to do is to find all these files, then maybe I can get them all in one place. At the very least I need some sort of directory or inventory or catalogue that lists all these 'media assets'. Which brings me immediately to the second problem.

Problem Two: where are my photos (and videos)?

Redwine's Report points us in the right direction: 'Secure storage is one of the most important factors in caring for personal digital archives. Choosing high quality storage media and refreshing it regularly will help avoid hardware obsolescence. In the event that a storage environment is compromised and files are lost, a reliable backup copy will be a vital part of helping you to regain control over your digital archive'.

My problem is that I have so much storage, of so many sorts (as do most people):

- one big computer, three laptops – and a tablet buried somewhere
- also a mobile phone, a camera, and several video cameras
- video captured on digital tape (mini-DV), mini-DVD, flash memory and internal camera memory, and hard drive; most of it is now on files in various places
- my partner: big computer, two laptops, tablet, at least three mobile phones, at least one each still camera and video camera (I lose track; I have enough trouble with my own gadgets)

~

We do not know whether something is rotten in Denmark, because our 'digital Denmark' is invisible and silent.

~

- a storage system attached to an Internet connection (a NAS = Network Attached Storage system)
- media on at least 7 USB hard drives of various sorts, in various drawers
- Amazon's Glacier cloud storage for a backup copy of all my photographs since 2013
- … but photos from my still camera uploaded automatically to Dropbox
- … except in January of this year I started using Amazon Photos for uploading and Amazon Prime for cloud storage – because my partner is paying for Amazon Prime so I *thought* I got free cloud storage
- … but it turns out that only she gets free unlimited cloud storage – not me.
- other various automatic backups used by my partner – to Norton and to Google+ and possibly also to Microsoft's Drive cloud service … every device or application we use tries to link us to one or more cloud services!

To know what I have and where it is, I either need to get everything in one place, or I need some sort of directory or inventory. It will need to point to multiple storage devices, including off line (USB hard drive in a drawer) and cloud storage (Dropbox, Amazon Photos). In my efforts to collect my photo and video files in one place, I've been faced with an additional problem: file names and types.

Problem Three: do I have the right file names and file types?

Redwine again: 'A key moment of neglect – but one which also presents an opportunity for responsible management – occurs at the point when digital files are created or transferred from a camera to a computer or other storage. For example, a folder filled with 2,000 digital photographs sequentially numbered with a camera-generated root (e.g., IMG_0001, IMG_0002) will require a considerable amount of work to manage. File names need to be meaningful in order for an individual to take advantage of digital affordances'.

File names and formats may need to be changed, to have a successful photo and video collection:

- files from cameras and mobile phones come with an automatically created file name. Sometimes it is a date and time (and therefore unique) but sometimes it is just an index number (leading to duplicates)
- software that transfers files from devices to a computer often changes the name – usually changing index number names into a date and time name taken from the embedded metadata
- if data from a device is manually as well as automatically transferred, there is a VERY high probability of having two identical image files with different names and different file-transfer dates. There should be a common 'creation' date in the metadata, but it may be difficult to get a computer system to use that date. So de-duplication of assets is a needed tool and process
- video files very commonly have proprietary formats, often a disguised version of MPEG, therefore conversion to a standard video format is probably also needed
- professional and 'pro-sumer' still cameras may offer a variety of image formats. There should be some sort of RAW (totally uncompressed) format, and this should be kept for archiving – but there are many versions of RAW formats, all proprietary. Images can be converted to uncompressed TIFF, to lossless JPEG or to standard lossy JPEG

~

In my efforts to collect my photo and video files in one place, I've been faced with an additional problem:
file names and types.

~

In my case, as with the example given by Redwine, I had thousands of photo files with index names like IMG_0001. I also had thousands with a file name based on the date of the photo. Sometimes that date was when the photo was taken. Other times that date was when the photo was moved (or even last moved). On some occasions, the date was just wrong because the device in question was set to the wrong date. Digital cameras are notorious for having the wrong date. It was often difficult to use the menu system to set the date and the date could reset whenever the battery died. Modern mobile devices are much better, usually taking a date and time from the mobile phone signal at regular intervals. I will describe later why I no longer worry about the file name but DO worry about having two versions of the same image.

Multiple versions happen (for me) every time I manually empty the SD card from my phone or manually transfer files from my laptop to an external hard drive. I collect all these files into my master filing system. I also collect all the files from cloud storage into the same system. So every file that automatically uploaded to Dropbox or Amazon Cloud is very likely to also come into my master filing system via a manual upload.

Tool One | De-duplication

I use a free programme called AntiTwin. It allows checking of all files in one directory (and sub-directories). It also supports checking files in one storage system against files in another system. It has everything I expect in terms of how to find a true duplicate, and also has multiple options for how to get rid of the duplicates (flag them, move them somewhere else, or delete them entirely).
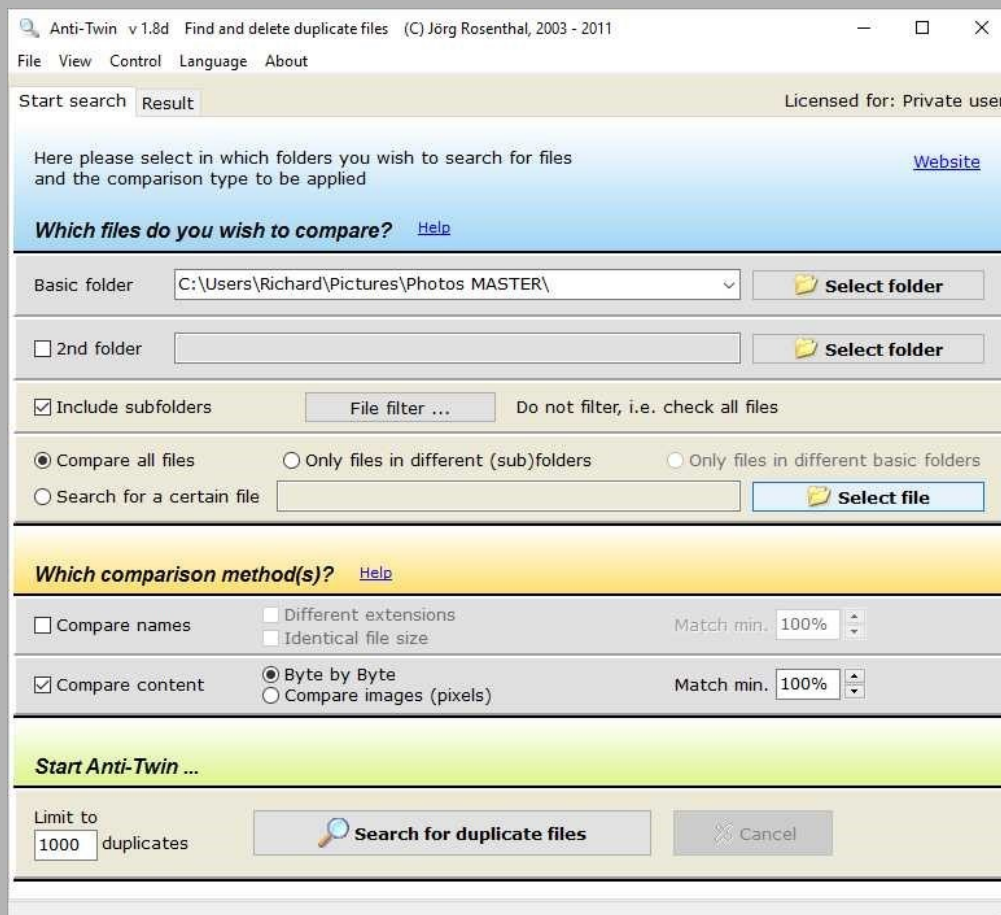
My ambition – keeping my media files

I want to know what I have, where it is, and to preserve these assets for the future. This ambition requires several steps:

- create and maintain an inventory of all media assets. I'm not sure what I'll do with an inventory of 35,000 items, but I want to at least be able to generate such an inventory

- maintain all assets in as few formats as possible, using open standards. This is pretty easy for photographs (except for RAW formats) but not so easy for video. Some of my problems (and solutions) for video will be described below

- maintain backups. Backups provide basic security, however, having files in so many places makes the process difficult. The need for backups convinced me that I also need to try to get all my photos into one place. Backing up digital assets is much easier and more reliable if I only had one 'thing' – one set of files – to be copied to backup storage

- establish confidence that the assets are OK (not lost, altered, or damaged). I could assume that if I have a main copy and a backup copy, then I'm alright. But that's an assumption. Secure long-term preservation is based on having actual evidence that assets are OK

- automate backups. I'm not a librarian and I'm not good at keeping things in order, so for me – the more automation the better (providing the automation does what it should, and no more)

- automate checks that assets are OK. Manual checking of 35,000 assets is impossible; checking this many files requires a tool that checks that a file is correct

My related ambition – to quickly find photographs

Which requires …

- discovery metadata:
    - Time and date
    - Place
    - Subject
    - Identification of people in the photographs

- a metadata database that supports quick search and retrieval; ideally this will be a well-documented database format allowing the data to be exported and used by some future search programme

- and also metadata embedded in every asset file – for preservation (created metadata embedded according to open standards, not in some obscure or proprietary fashion)

~

I now feel that I'm about 80% of the way to the goals listed. Two years ago, I felt less than half-way along, and four years ago I was just facing a mess.

~

Clearly not all in the above can be achieved – at least not by me in the finite amount of time between retirement and the point of eternal departure. One motivation for this note is to describe my personal effort to travel in the general direction of these aspirations. I have used free, open-source and standardised methods in almost all cases. I now feel that I'm about 80% of the way to the goals listed. Two years ago, I felt less than half-way along, and four years ago I was just facing a mess.

I originally thought I could use digital library or asset management software to create an inventory (catalogue), to support search and retrieval, and to support backups (and checking of assets and backups). I quickly found that most of these approaches were unsuitable for various reasons.

How not to build an inventory: taking stock of what you have, creating a list, and saving that list in a digital, searchable format should not be so difficult. Linux people told me that only PC users had a problem in this area, but it is a big problem whether you use PCs or Linux. Probably the worst 'solution' I discovered was advice from a Library of Congress student project: 'We suggested opening a list of the files and taking a screenshot using the Print Screen button on the keyboard'. Clearly that suggestion doesn't scale to 35,000 files, and doesn't produce a machine-readable inventory. So what should I use?

Digital library software: the products I knew about were Greenstone, DSpace and Fedora. Dspace is really for archiving (putting things away), and both Dspace and Fedora take up too much space for a home computer. Fedora could be used as a cloud service, but the first step would be to upload all assets – a situation I faced later when looking at digital asset management systems (see below). I didn't want to manage my 50 GB of images and several hundred GB of video in the cloud. The upload times would be huge and the storage would be costly.

Greenstone, however, was attractive. I've used it since 1998 and it is very well-documented and supported. It has extensions to support the embedded metadata in photographs and it is simple, robust, and easy to use, even on a simple PC. It can be pointed at a collection of files and will create a digital collection – a library – with functions for web access, metadata control and editing, and search and retrieval. There was only one problem: I really wanted to use face recognition on my 35,000 images using software that was incompatible with Greenstone.

Digital asset management = DAM (or media asset management = MAM) systems: DAM systems usually offer media tools beyond the capabilities of digital library systems. Though clearly there is also much overlap. Most DAM/MAM systems are commercial and therefore costly, but there is one open-source product that looked useful: ResourceSpace. Oxfam and many other organisations use this OS software professionally to manage audiovisual (and related) assets.

I would recommend ResourceSpace to an institution (particularly a multi-site or multi-national one) that wants to manage joint assets. However, for personal use, I had a problem pretty much the same as with Fedora. I could have a local version – but only if I had or built a Unix server! Or I could obtain a cloud version, but then I'd have to push all my assets into the cloud. As stated above, that is a very slow process.

~

With the grandchildren using Instagram and other social applications, I doubt if I'd get real support for a 'family digital asset management system'.

~

Cloud upload speeds: I've tried Amazon Glacier, Amazon Cloud Drive and a cloud version of ResourceSpace. On my network connection, the upload time is three to four hours per gigabyte. For my 50 GB photo collection, that's a long time: roughly a week if all goes well with overnight runs. Comparatively, in my actual experience it takes more like two weeks due to failures and glitches.

For a small institution, or for a large but very well-organised family, ResourceSpace is worth considering. It won't run on a home PC but it was surprisingly easy to create a completely private version on a cloud service. There are instructions on ResourceSpace for launching a cloud version on the Google Cloud Platform. That platform gives a 60-day free trial so I 'deployed' my own cloud version of ResourceSpace. I found the process very straightforward and quick. In literally 20 minutes, I had the process all up and running, all my own, on the cloud. The cost was about £3.50 per month, including 7 GB storage – which I didn't pay because of the free trial.

If upload speeds increase sufficiently, I'm still considering ResourceSpace as an option for the future – if the rest of the family would agree to use it for our 'family media resources'. But with the grandchildren using Instagram and other social applications, I doubt if I'd get real support for a 'family digital asset management system'. I still want to mention it, though, because other families – and certainly many small institutions – could now get real professional asset management in the cloud (with full privacy) for a few pounds per month. Although, these services charge for moving media in and out, a cost which could quickly become impractical. However, as a basic 'family archive' – mostly storing media, not much moving media in and out – it could well be cost effective.

Back to an inventory: in searching for inventory and asset management tools, I eventually came across tools that had been created twenty years ago when people had lots of files stored on writeable CDs (CD-ROMs). Software had been written that would make a simple catalogue of the file names and file locations from a collection of such CDs. Some of these programmes continued to be developed and supported and can now be used for internal and external hard drives, USB sticks, SD cards and even cloud storage. Basically, these programmes can create an index (or catalogue) from any storage that can be recognised by a PC and given a 'drive letter' (like the C: or D: drive).

I had been on a two-year search for such capability, and so had other people. There was a blog about a fruitless search for basic asset inventory management, written by David Gewirtz, a self-described 'power user' who couldn't find what he wanted. He searched all the software dedicated to photos (including Picasa) and then went on to search asset management systems as well. In the end, he didn't find anything that would support his professional

requirement to manage 10 GB of image content.

Meanwhile, I found the CD-ROM indexing software and was pleased with it. I went back to David's article to drop a comment only to discover that he'd come to the same conclusion: forget asset management and photo management tools, get a basic index from the old CD-ROM tools.

David is a MAC user and so settled on Neofinder. The same company makes a PC version, abeMeda for Windows.

What do such programmes do? They make an inventory, and can do so for virtually any number of physical storage systems. Point the programme at the storage, get an inventory. There is limited searching of embedded metadata, but the main point is having a database (inventory, catalogue) of all your assets in one place.

These programmes are commercial software, not open-source – AbeMeda costs £29 – so there is a question about long-term support. On the other hand, their survival for 20 years is a good indication of a large user-base and continued interest in the products.

Tool Two | CD-ROM Listing Software

I bought a related product, WinCatalog, because they had a £5 offer if I would also order some business cards (which I needed anyway).
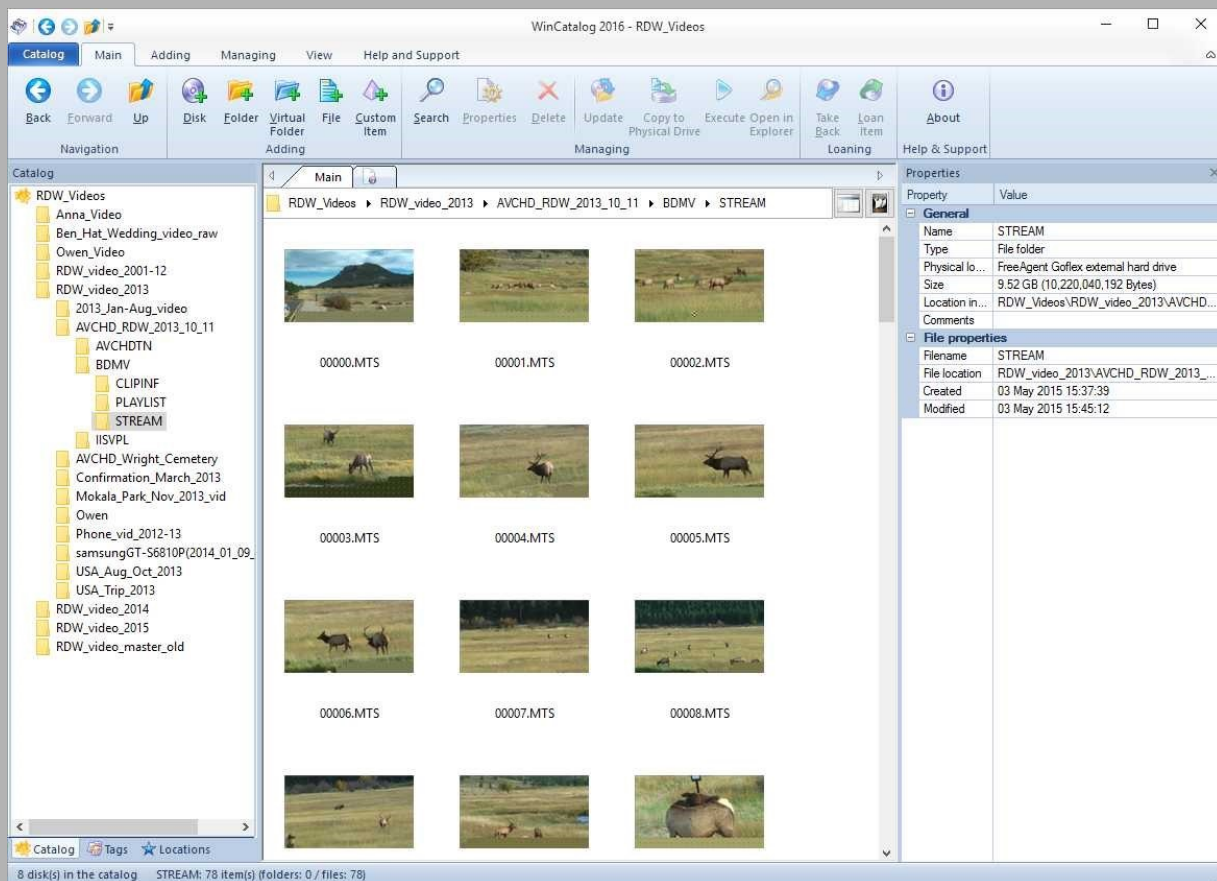
WinCatalog tells me what files I have and where they are. If what I find isn't on my computer, it will tell me which storage system or device it is on and prompt me to load that. The mindset of the tool is still managing files on a collection of CDs – but that mindset is still viable. The basic difference between a 700 MB CD-ROM and a 2 TB USB hard drive is simply size. The hard drive has nearly 3000 times as much storage, but it is still a piece of removable storage with a filing system – and lots and lots of files.

Back to Face Recognition: with WinCatalog I can make a useful database for all my video files, and there is no problem with the fact that I have video on several different hard drives. The software creates thumbnails and allows me to type in short descriptions, and so is a good start for managing my video collections. What I don't have is automatic shot detection, something to break the video into 'shots' and support describing individual shots in order to navigate at a more granular level.

However, my current video camera (which uses a memory card for storage) produces a new file every time the camera starts and stops, so for that camera the media is already divided into shots. Where I have a problem is with my older DV tapes. The ingest for those is a lengthy process and might or might not include shot detection depending on the software. So the solution to shot-level management of my video requires attention to how I ingest the video. I'd use the shot detection of the ingest programme but manually add shot-level metadata in WinCatalog. Ideally, I'd also have a way to move metadata between my 'media manager' (WinCatalog) and my 'media editor' (I use Pinnacle) – but exactly how to do that becomes very application specific.

Video takes up a lot of space, but I don't really have a lot of video files, so I can handle making metadata manually. For photographs, and 35,000 files, I really want automation. Most of my photo files have embedded data for date and time (most of it correct!) and recent files also have embedded geospatial data identifying where the photo was taken. It is a modern irony that my high-spec camera (with good lenses and the capability of producing high-quality photos) is far behind my smart phone in two ways: no geospatial coordinates and no automatic uploads.

Tool Three | Picasa Photo Management

Picasa photo management with face recognition. Now about to be 'retired' by Google! I'll try to keep using it as long as I can. My Picasa3 still runs under Windows 10.

With metadata about 'when' and 'where', the obvious next questions are 'what' and 'who'. The what information will probably never become automatic, but for 'who' face recognition can help. For this I use Google's Picasa -- a free but NOT open-source application that drives many people wild (see the forums and also these forums) but does do a very good job of recognising faces. I now have about 3000 tagged faces in my photo collection representing about 100 people.

The bad news is that Google recently announced it is going to 'retire Picasa'. My copy of the programme runs on my computer, so I can continue to use it locally for as long as I wish. However, pressure to move our content onto the web will inevitably continue to increase in future. Photos already upload automatically from mobile devices to the web. According to Google, the only sensible thing to do with them is leave them there, manage, tag, and apply face recognition to them there. Forget about old fashioned things like knowing where they are stored, or how, or whether they have become lost or corrupted.

I ensure that Picasa writes face recognition metadata into the individual photo files, as well as builds a database of the metadata. Face recognition metadata has a standard and is written as XML. Picasa does write the metadata according to the standard. In principle, I could find another database product (such as Lightroom) that would read all my files and make a new search-and-retrieval system.

With metadata for 'where', 'when', and 'who' I could forget about file names. The important thing is the metadata embedded in each file and in a database for quick search-and-retrieval. The search functionality of Picasa includes the ability to search embedded metadata (although this functionality is not well documented). It will also search faces, locations and manually-added tags. Everything that Picasa lists under the 'properties' of a photo can be searched. This feature allows a primitive version of search by date range. For instance, searching for "2013:08" will find all the August 2013 photos. One can even search for photos that come from a particular camera.

Tool Four | Photo Viewing & Editing

Irfanview is my general photo viewing and editing tool. It has a wide range of photo management functions, many of which can be batch processed.

But I still have worries about security of arbitrary filenames like IMG_0001. I prefer something with meaning, preferably the date the photo was taken. When I eliminate duplicate files, I keep the ones with the date in the file name. If there is no date in the file-name, the name of the file can be changed to use the date from the embedded metadata. Picasa won't do this, but a widely-used open-source photo edit programme called Irfanview will do exactly that. Furthermore, it can do it as a batch process on a whole folder of files (including subfolders as needed).

Currently, to manage my 35,000 photos, I keep a master file (really a whole hierarchy of files) of about 50 GB on my main home computer with a backup on our networked attached storage (NAS) system. The backup automatically updates whenever the master file is changed. For more redundancy, I have a second backup copy on a removable hard drive. Transfer speeds get higher all the time, but still it takes hours to transfer 50 GB. To help ease this time burden, I use a programme called SyncToy that only copies new files or modified files. It's not a toy, and is old, but it's robust. I recommend it for all mass-migrations of files to backup.

Tool Five | Backup Storage Sync

SyncToy manages keeping backup storage in sync with a set of master copies.

Tool Six | Cloud Storage Management

FastGlacier manages Amazon Glacier storage, with costs about $0.01 per GB per month.

If someone breaks in and steals the main computer and the NAS and all the USB hard drives in our house, my photos will be lost. To mitigate this risk, I also keep a copy in Amazon's Glacier storage. This process took a week's work to set up, but now updates only take a day or two, a process which I carry out once a year. The cost is less that £0.35 per month. I use a free programme called FastGlacier to actually move content in and out of the Amazon Glacier storage.

Organising storage to support backup: Originally, for probably ten years, I organised photos in folders by subject. I had folders about trips and folders about specific family members. This system proved awkward when I began to make systematic backups, especially over a slow upload to Amazon Glacier. Any folder that was changed had to be uploaded again, and I was adding new content to most of my 'family member' folders on a regular basis.

To solve this problem, I re-ordered all the photos by date. Every year a new year's worth of photos can then be uploaded to Glacier, but the older year folders are kept sacrosanct. I keep one general folder for putting in photos that should be in older year folders – and someday they will all be re-sorted. At that point, I'll re-upload the whole set to whatever cloud storage system I'm using then.

Integrity checking: in a project on digital preservation that I was involved in, Matthew Addis (now of Arkivum) formulated a simple state-transition diagram showing how to keep content alive. The basic idea is to check files regularly to ascertain if they can still be read and whether they have changed.

As the diagram below shows, once an error has been detected, a repair can be performed simply by replacing the file with the backup copy. The process only goes wrong if the backup is also faulty – in which case there is irrecoverable loss.
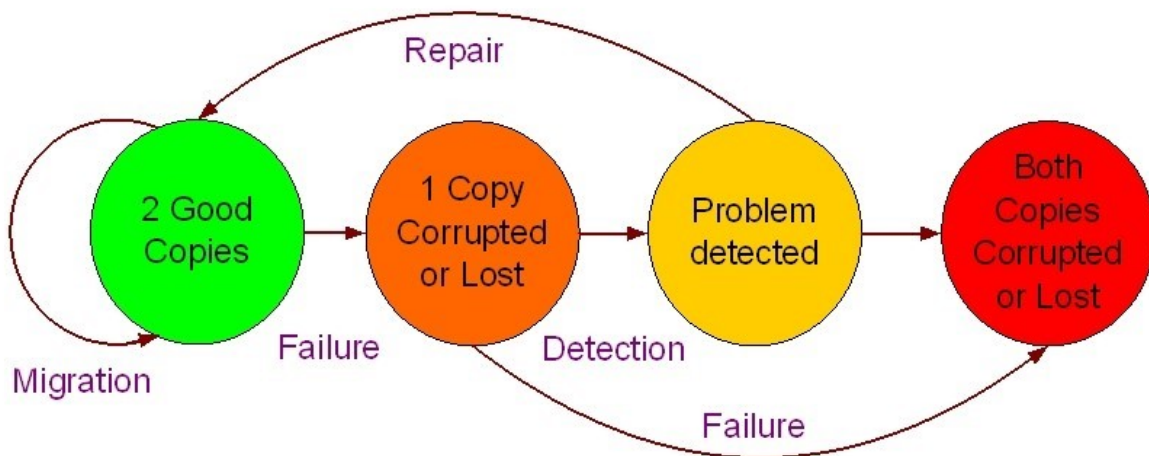
One way to see if a file has changed is to check the metadata: has the date on the file changed? Another method is to check if the size of the file has changed. The problem with both of these approaches is that a real error, such as a corruption introduced when a file is moved from one storage system to another, will not change the date and may well not change the file size – but could still corrupt data within the file.

For a small number of files, manual checking is sufficient: manually see if the files open, and if they 'look right'. This method does not scale beyond something like 100 files, and I have 35,000. The proper way to check files is to have a programme that does the mathematical equivalent of 'looking at them'. A number is computed that represents the state of all the bits in the file. That number, the fixity information, is stored in a database. A file is checked by re-computing the fixity number and comparing it with the original. If they match, there is a very high probability that nothing has changed – the file passes its fixity text.

Tool Seven | Fixity

[Fixity from AVPreserve](#) will automatically and periodically check files, and flag any problems.

I use a free programme from AV Preserve to automatically perform fixity tests once a month. My master copy of all 35,000 photo files sits on my main computer. The fixity check programme is set to run in the background on that machine and to produce a log of any problems. So far I haven't had any and don't really expect any. It is difficult to get statistics on what makes people lose data, but the biggest problems are 1) human error; 2) failure of entire storage systems (a disc crash of memory card corruption). In my view, a principal use of my automated fixity checker is that it forces me to be careful about my approach to storage and backup, thus reducing the chances of human error!

My files get changed all the time. I add metadata using face recognition and do other kinds of manipulation. Fixity is clever enough to recognise an intentional change from an outright error. It is able to detect intentional change because then those files have not only a new fixity number but also different metadata (date of last change) and usually a different size. The file that represents a problem is one that ONLY has a change to the fixity number. These are the files that will cause Fixity to launch the equivalent of a red alert so that the user can take recovery action (e.g. replacing the bad file with a backup copy).

How did I do?

Here is a review of my aspirations:

- create inventory of all media assets. I use WinCatalog

- have all assets in as few formats as possible, using open standards. My photo formats are standard. I have one set of RAW high-quality files, which I'm just keeping as-is. I had a problem with video from an old DV camera that used mini-DVDs for storage. The issue of video formats is really messy, and I did a lot of searching. In the end I used a free programme DVDVob2Mpg http://www.dvdvobtompg.com/ to get a standard file from the DVD-format files. There are still problems with this conversion (metadata gets lost) but the whole area is too intricate to describe here.

- have backups. I have 3 sets of backups for photos, and 1 backup for videos.

- know that the assets are OK (not lost, altered, damaged). I use fixity information.

- have an automatic way to make backups. One backup copy is automatically made, using the software that comes with my NAS system.

- have an automatic way to check that assets are OK. I have monthly automatic fixity checks using the programme Fixity.

What I also wanted was:

- discovery metadata
    - Time and date
    - Place
    - Subject
    - Identification of people in the photographs

- in a database supporting quick search and retrieval; ideally this will be a well-documented database format allowing the data to be exported and used by some future search programme

- and metadata also embedded in every asset file – for preservation
    - created metadata embedded according to open standards, not in some obscure or proprietary fashion

Overall, I have met these aspirations, but discovery metadata comes from a proprietary and now obsolescent application – Google's Picasa software. The database is NOT well-documented and supported, though Google claims it is compatible with LightRoom. Also with cooperation the millions of Picasa users may be able to ensure a way forward with our data.

Perhaps the greatest success I've had in my effort to preserve my photo and video collections for my children and grandchildren is the preservation of all the metadata that I have created (face recognition, description, tags). This metadata is a success because it is not only stored in a searchable database but also embedded in the files themselves. Photographs have the capability for very rich metadata, providing more possibilities than this case note can explore. The important lesson is that metadata in a database supports quick search and retrieval, while embedded metadata supports preservation. Embedding ensures the metadata still exists, even when the database application (as with Picasa) becomes obsolete or otherwise unusable.

About the Author

Before retiring, Richard Wright worked as a Technology Manager at BBC Information & Archives from 1994. He later became a Senior Engineer Archive Researcher. Before taking on the post at BBC, Wright worked in acoustics, speech and signal processing for the US and UK Governments, University College London, RNID and Cirrus Research. Today, Wright runs Preservation Guide, a consultancy on audiovisual preservation. Wright is the author of many reports and publications, including the DPC Technology Watch Report *Preserving Moving Pictures and Sound* (2012).

To contact Wright: preservation.guide@gmail.com

To contact the DPC visit www.dpconline.org