Slide 1

My topic today is Preserving Your Family History Records Digitally. We will be looking at the challenges, solutions and considerations involved in this process.

Slide 2

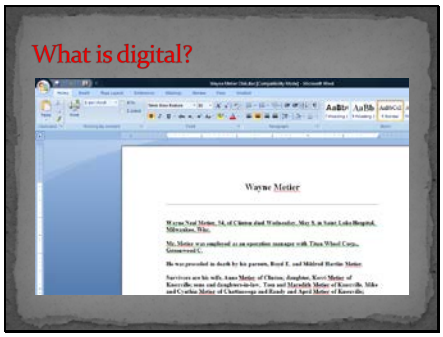
So what is a digital record?

[Speaker holds up a paper death certificate] This is not a digital document—it is obviously paper.

However, if I scan this document onto my computer, this record is now digitized.

Likewise if I download a death certificate from the internet, it is also a digital record.

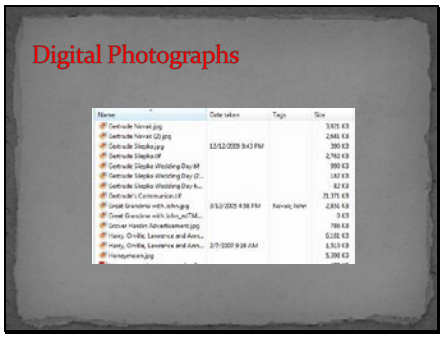
Slide 3



[Speaker holds up a Ancestral Research Report printed on paper] Again, this is not a digital document.

However, if I use Microsoft Word to create the document, this document, as it stored on my computer is in a digital format.

Slide 4



[Speaker holds up a framed photograph] Finally, this is not a digital photograph.

If I scan this photo or take a photo with a digital camera, these photos will be digitized.

So what I will be talking about today are the considerations involved in preserving your **digital** documents and photos for generations to come.

I will **not** be addressing preservation of video and audio files.

Slide 5

I am not an expert in this field. Most of the material presented here was gathered from the online articles indicated.

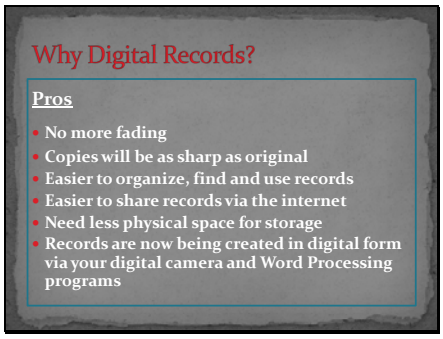
Preserving Your Family History Records Digitally
By Gary T. Wright
Sept. 9, 2011

The Best Way to Archive Anything: L.O.C.K.S.S.
Posted by Dick Eastman
Aug. 17, 2011

8 Blunders People Make When They Scan Photographs
Sally Jacobs, The Practical Archivist

Screen shots of computer instructions shown in this presentation are a computer using Microsoft's Windows Vista Home Premium.

Slide 6



•You don't have to worry about your documents yellowing, fading or losing contrast

•If you make copies of the originals, the copies will be as bright and sharp as the original

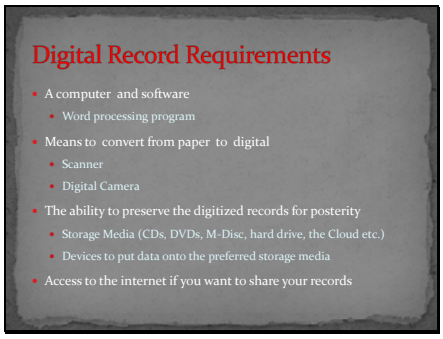
•The computer does the work– it organizes, stores and searches the records for you. You don't have to climb up stairs to an attic or down to a dark basement to retrieve your records. Plus digital records can contain descriptive information which make finding the records so much easier (more on that later)

•Easier to share your records either by email or by using an internet service.

•Don't need bulky boxes and costly containers and the space it takes to store them. You can keep a copy in a safe-deposit box.

•You need to learn to preserve digital records anyway because most photographs and documents are now being created digitally. "Born-digital" is a term coined for records that originate in digital form (most likely your personal history was written using a Word processing program and family photos are now taken using a digital camera).

Slide 7

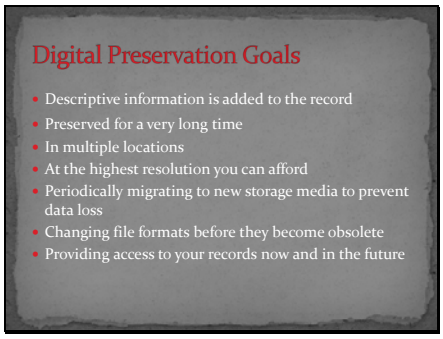


•Obviously, a computer and the necessary software
•You may need a scanner to convert your hard copy vital records, census records, photos, newspaper clippings, maps, etc. into digital format. A digital camera may also be useful for this purpose.
•You'll need storage media on which to store your records and any devices needed to read and display the records
•And of course, you will need access to the internet if you decide you want to share your records with family or other interested parties.

Slide 8

In a nutshell, digital preservation is all about providing long-term protection of your digital records

Slide 9



The goals of the Digital Preservation Process I am going to cover today include:

•Preferably, we want to include descriptive information along with the record. Who is in the photo? When was it taken? Etc.

•Need to ensure that these records will be readable for a very long time

•We need to store our records in multiple locations --Jacques Lowe was John F. Kennedy's **unofficial** White House photographer. Kennedy told him to "stick around and record my administration". Well, his precious negatives were stored in **one** place—the J.P. Morgan Chase bank vault housed in the World Trade Center basement. 40,000 historically significant photo negatives were gone. It would have been better if the negatives were split into groups and housed in multiple bank vaults in multiple cities.

•Preferably, you want your records in the highest resolution you can afford—particularly important for photos

•Periodically you need to refresh the media your records are stored on as digital data recording media may deteriorate more rapidly than say paper. Also the hardware devices needed to read the data may become obsolete.

•Another digital preservation goal is to Change the file formats before they become obsolete

•Finally you want to have access to your records now and in the future

Slide 10

So what are the challenges involved?

First off, on a computer, text, pictures, audio, and nearly any other form of digital information is converted into a string of bits, each of which has a value of 1 or 0. The computer can read these bits and render an image that can be viewed on a computer screen or printed.

Unfortunately, the physical deterioration of whatever you store your records onto—flash drive, CD, external drive, etc. will occur over time causing bit loss which causes data loss.

Slide 11

Kurt Gerecke, a physicist and storage expert at IBM Deutschland GmbH says "that burned CDs have a relatively short life span of between two to five years, depending on the quality of the CD,". "There are a few things you can do to extend the life of a burned CD, like keeping the disc in a cool, dark space, but not a whole lot more."

The problem is material degradation. Optical discs commonly used for burning, such as CD-R and CD-RW, have a recording surface consisting of a layer of dye that can be modified by heat to store data. The degradation process can result in the data "shifting" on the surface and thus becoming unreadable to the laser beam.

"Many of the cheap burnable CDs available at discount stores have a life span of around two years," Gerecke said. "Some of the better-quality discs offer a longer life span, of a maximum of five years."

Cheap CDs fail more often than high quality CDs.

The Gold CDs and DVDs use a Phthalocyanine (thal-o-sy-a-nene) dye which is superior to the typical dye used on CDs.

Store CDs in jewel cases with a central hub. If you put CDs in a sleeve or envelope, every time you remove it, you risk scratches.

Never write on the top of a CD with a regular pen. Use a felt-tip pen to avoid scratching and only write on the round space in the center of the disc. Or use a label on the case to identify the contents.

Gloves are a great idea. The oils on your fingers contain salts and acids that are not good for CDs.

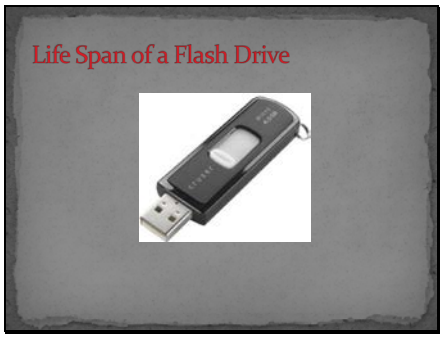
Store CDs away from light, heat and dust.

Slide 12

The average life span of a hard disk drive is hard to forecast. Obviously, a number of factors affect its life span. Heat being a major one. How many hours per day the drive is in use is another. But on average, the life span is said to be 3-5 years.

Also, another consideration when choosing a hard disk over optical discs is when it comes to migrating your files to new media. It is far easier to plug in one hard drive than to insert a lot of CDs or DVDs.

Slide 13



What about the life span of a flash drive?

According to IT Director Rae Williams, “Flash drives are very handy for carrying files from place to place and computer to computer. However, they are relatively volatile storage, so you should never consider them a primary backup for your files. They fail much, much, much more quickly than CDs or hard drives.”

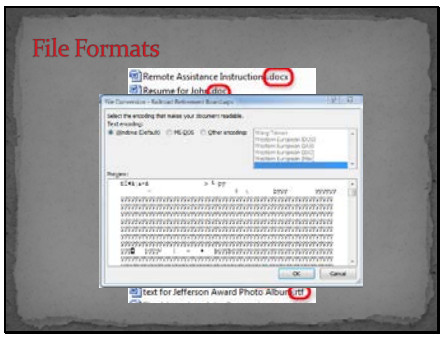
Sally Jacobs, The Practical Archivist, says that Flash Drives “are notoriously unreliable and **not** a viable option for long term digital storage”.

Slide 14

The 2nd Digital Preservation challenge is Obsolescence.

And the format the file is stored in is our 1st consideration.

Slide 15



What do I mean by file format?

In the chart, I show some sample documents with various file formats in which the file format is circled in red. The format of a file is based on the section of the filename following the final period.

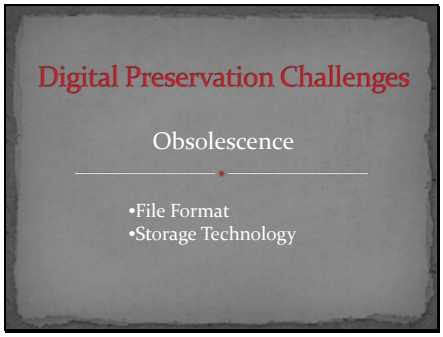
The first red circle indicates a file format of **docx**. This document was created by the program Word 2007 which appends the docx file format to the documents it creates.

The file below it has a **doc** file extension was created by an earlier version of Microsoft Word. In many cases, the file format indicates which program was used to create the document and which program will be needed to read this file.

Why is this important to know?
Let's say a number of years ago, you created a text document detailing your father's memories using Microsoft Works which had a file extension of WPS.

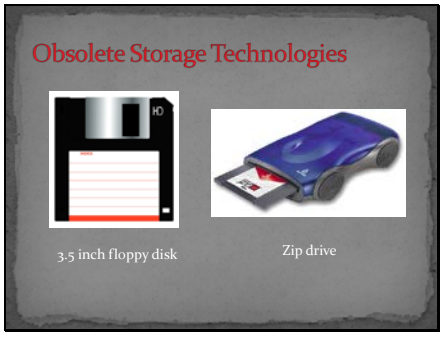
Now in 2011, you might try to open this file using Word 2007 and you may end up with something that looks like this. Same holds true for a file created using the once popular Wordstar.

Slide 16



The second challenge is the obsolescence of storage media and the hardware device necessary to read the files.

Slide 17



Who doesn't remember the 3.5-inch, and 5.25-inch floppy disks developed by IBM, from the mid-1970s to the 2000s.

Now, you may find it hard to find a way to read documents on these devices because the devices used to read them are no longer commonly used.

Recently, while trying to clean out a drawer, I came across this floppy disk that Wolf Camera had put my photos onto. It contained a favorite photo of my husband on it. I no longer have a floppy disk reader in my house so my son had to take it to work where they had a floppy disk reader and copy the photos to my flash drive.

Slide 18

One practical and effective solution to avoid storage media bit loss will soon be available. It is called an M•DISC. M-Disc is a revolutionary optical disc technology developed by Dr. Barry Lunt and Dr. Matt Linford of Brigham Young University and manufactured by Utah-based Millenniata, Inc.

The M-DISC records your records with permanent engraving—like etching stone. Unlike other optical discs that use an organic dye for recording, the M-DISC uses a synthetic stone-like material to record your digital bits. A high powered laser permanently etches the bits into the write layer of this synthetic stone, and they can be read by virtually any DVD or Blu-ray drive for generations to come without any data loss.

In 2009, the U.S. Naval Air Warfare Center Weapons Division at China Lake, California tested four different brands of gold archive-grade DVDs and one regular DVD along with the new M-DISC. The project was an accelerated aging test that evaluated disc stability and readability after being exposed to elevated levels of light, heat, and humidity. A report of the results stated, *"None of the Millenniata media suffered any data degradation at all. Every other brand tested showed large increases in data errors after the stress period. Many of the discs were so damaged that they could not be recognized as DVDs by the disc analyzer."*

Based on this and other internal testing, Millenniata claims that the M-DISC has an archival life measured to last up to 1000 years. And M-DISCs have no special storage requirements to achieve this remarkable archival data life.

Millenniata, in collaboration with LG, has produced a new line of optical disc rewriters capable of reading and writing DVDs, CDs and M-DISCs.

The M-DISCs and M-READY DVD drive were expected to be sold in popular retail stores beginning in Oct. 2011. However, as far as I can tell, they are not yet available in stores but can

be preordered (see hyperlink).

_____

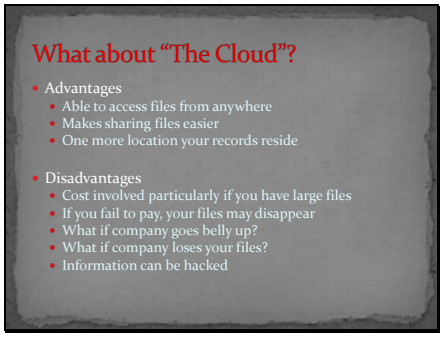
Migration involves copying your family history digital records to a newer storage medium that is about to replace the old technology or to prevent bit loss from occuring. In other words, before the 3-5 year limit on CDs, DVDs or hard drives expire, you must rewrite your records to another storage media to guarantee that you won't suffer data loss.

Similarly, just like the way of the floppy drives, if the storage media your using is losing popular support, you must copy your records to the new technologies before the old devices are obsolete.

Even if they are M-DISCs, they may need to be copied to the M-DISC replacement technology of the future, whatever it turns out to be.

Remember migration work may have to be performed by your posterity or your extended family after you are gone. You need to prepare your descendants for such migrations.

Slide 19



Storing your files and photos online at a service like **Amazon Cloud Service**, **SkyDrive**, **Google WebCity**, etc.

Advantages
You are able to access your files from any computer anywhere in the world.
It makes sharing files easier.
Is one more location your records can reside

Disadvantages
Some of these services are free. Many offer a small amount of storage for free but then require a fee for additional storage.

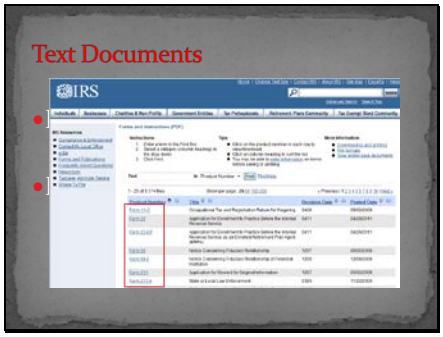
But what happens if you get incapacitated and unable to renew your subscription? Worse yet what happens to your files if you should die? You will lose your records

What if the company goes belly up?

What if the company loses your files? Amazon recently lost some of their customers' data.

Finally, the information runs the risk of being hacked

Slide 20



All of us are probably familiar with **PDF** files. At tax time, you may have to downloaded tax forms in the PDF format. The US government as well as the medical community use this format a lot.

Advantages of PDF file format:

- •Guarantees long-term reproducibility
- •Is efficiently and compactly compressed while the quality is almost always as good or better than other formats
- •Can easily be shared around the world with Mac and Window users alike
- •Possible to search documents archived in the traditional way for specific information
- •Metadata (author, topic, content, keywords, date created, date modified, publisher, etc.) are embedded directly in the PDF file

But over the last 15 years, the PDF file format has added function and overhauled existing ones.

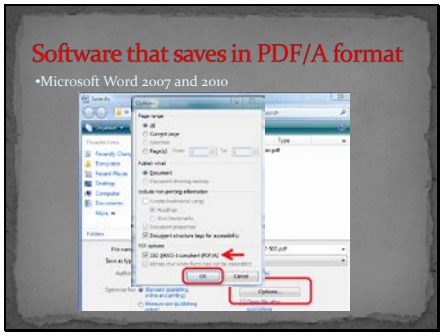
_____

PDF/A

Soin 2005, recognizing the impact of file format obsolescence on digital preservation, the International Organization for Standardization (ISO) defined an "electronic document file format for long term preservation." Based on the Adobe PDF 1.4 format, PDF/A provides a self-contained,

self-describing format that is independent of external sources. For example, it embeds fonts and color information with the content data so that future computer software will be able to render the document exactly as it looked when created. The document will look the same no matter what program and system was used to produce, save or reproduce it.

Slide 21



How can you save documents in PDF/A format?

Microsoft is responding to customer demand by making it possible to create PDF/A documents directly from within the most current Office suites (2007 or higher)

Once you have installed the **Save as PDF** plugin, click on the logo button. This opens a drop-down menu.

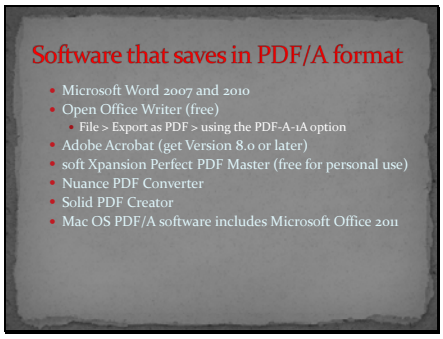
Hover your cursor over **Save As**

Click on the **PDF or XPS option**

Another window opens. Click the **Options…** button

then select **ISO 19005-1 compliant (PDF/A**) under **PDF Options**

Slide 22



Other software that saves in PDF/A format include:

Open Office Writer (free download at http://www.openoffice.org/)

Adobe Acrobat X Pro
Very expensive; however, you can download a free 30-day trial of Acrobat X Pro from the Adobe website and use that to create PDF files on a temporary basis. Also, if you have a student or teacher in the family, you can ask them to purchase the software for you at 80% off lowering the price to $119.

Soft Xpansion Perfect PDF Master
•Free for personal use
•Does not allow for the addition of descriptive information when creating PDF/A files. You must purchase the business version of this product in order to get this capability

Nuance PDF Converter

Solid PDF Creator

MAC OS PDF/A

Slide 23

Let's now turn our attention to photographs.

I need to do some explaining first though.

If I continue to zoom in on a photo enough, I will begin to see these boxes that make up every photo. If it is a color photo you will see colored boxes. These boxes are called pixels. They are the basic component of every image.

Slide 24

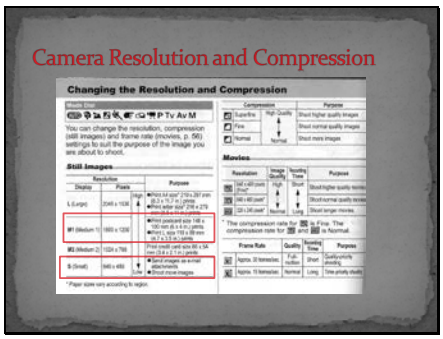
When scanning photos, to get a decent print that doesn't look pixilated (you begin to see the pixels), set your Scanner to a resolution of at least 300 dpi.

This 300 dpi recommendation only applies when you don't need to enlarge the print size. In other words, to print at the same size the photo was scanned as, set your scanner to at least 300 dpi.

But if you are scanning a 3.5 x 5 inch photo and you want to print it out at 5 x 7, you'll need to increase your resolution to 429 dpi.

For a handy resolution calculator, visit the indicated website.

Slide 25

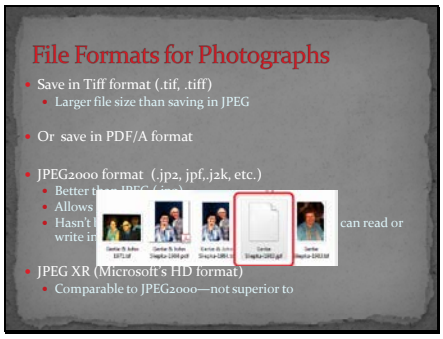


Likewise, set your camera accordingly to enable you to print your photo without pixilation and in the size you want.

Using the chart shown, if I set my camera to take photos using the small resolution setting, I will most likely not be happy with the photo if it is printed. The chart indicates I will need to set my camera to at least the M1 resolution setting in order to print my photos as 4 x 6s. I would need to set my camera to the large setting if I wanted to print my photos in a larger size.

File Size—always preserve records at the highest resolution that is reasonable for your situation. The reason is that a digital record's resolution quality cannot be improved once the digital record is created.

Slide 26



What is the best format to save your archival photos in?

Well, after I scan a photo I'd want to preserve, I could save it in **Tiff** format. Tiff format would result in a larger file size than Jpeg format, because the JPEG format will throw away pixels in order to reduce the file size. This loss of data may result in a lower quality print.

_____

You can also save your special photos in **PDF/A** format**.**
All information necessary to display the document is embedded. Though articles including the one I relied on extensively for this presentation recommend savings photos in PDF/A format, I don't recommend it at this time unless you own the software Adobe Acrobat X PRO and have the knowledge needed to understand the PDF/A Standard technology--which is considerable and specific. I believe the PDF/A format is more suitable for text documents at this time.

_____

Another format you could use is the **JPEG2000** format. **JPEG2000** is a file format that is superior to the earlier 1992 JPEG standard. Same quality as Tiff files but the small size of the old JPG format. It is used strictly for photos. The problem with the new format is that cameras were expected to support the JPEG2000 format. However, that hasn't happened yet nor has there been widespread adoption of the format (web-browsers and email programs) and few programs can read or write in this format.

Also the thumbnails available don't display in my Windows Vista environment which I see as one big drawback. Now perhaps there is a way around this, but at this time I am unaware of it.

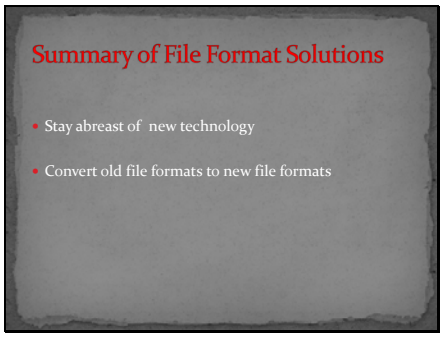
_____

JPEG XR (Microsoft's HD format)
Why go for a new standard? The XR standard appears to be equal but not superior to JPEG 2000 plus less metadata information is available for use.

So until things change, unless you own a copy of Adobe Acrobat X PRO, I recommend you store

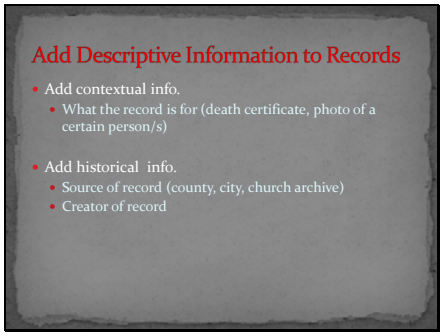
photos in tiff format with lossless compression.

Slide 27



If any file format you use starts to diminish in popular use or a new format is introduced you need to convert your files from the old format to a new format before the old format become obsolete.

Slide 28

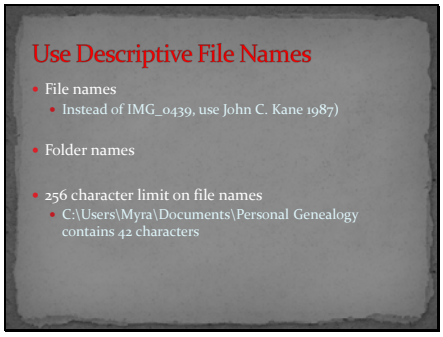


Before starting to preserve your digital records, you will want to setup a **plan** to add descriptive information to your records.

How do you accomplish this?

You must decide on a folder structure before you start scanning. Do you want to organize by a person's name? In other words, do I want to make a folder with an ancestor's name and put all his photos, death certificate, census images in subfolders? Or do I want to create separate folders for all death certificates, census images, etc.?

Slide 29

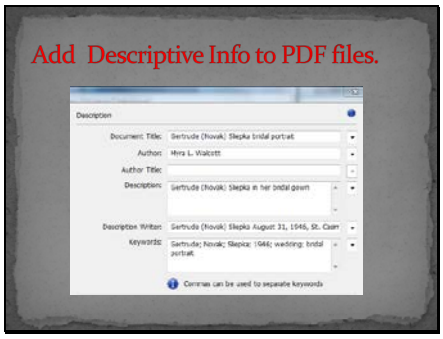


Also you can use folder and file names as part of your descriptive information.

NOTE: Current personal computers have a 256 character limit on file names. This includes the names of all folders that must be opened to get to the file.

So in this example, the path name contains 42 characters. Any subfolders or files added to this path must be 213 characters or less.

Slide 30



How do you add descriptive information to PDF records?

If you are saving records in PDF format, use your pdf software to add this descriptive information before you save in PDF/A format.

Slide 31

Another way to add descriptive information to your records is through the use of tags.

Plus You can make it a whole lot easier to find your pictures and files by applying tags to the files.

Microsoft Office documents and photos stored in Tiff can be tagged.

However, certain file formats cannot be tagged in Microsoft Vista or higher—These include RTF, PNG and TXT.

So let's see how finding files using tags works!

I have applied the tag 1946 to a photo. All I need do in Vista, is click the large logo button and in the search field, type in my tag—1946. Immediately above, Vista will display files having 1946 in the filename or files having a tag of 1946. You'll notice that the indicated file does not have 1946 in its filename.

You can also search for files within a particular folder using tags. Say, I want to find photos in my Pictures folder having a tag of 1946. I would open the folder in Windows Explorer and in the upper right hand corner of the window, there is a search box. I could enter 1946 or I could use the advanced operator **tag:** followed by **1946**. Because I used the advanced operator tag, Vista would only display photos having the 1946 tag.

Slide 32

How do I add tags to a document or photo in Vista or higher?

In a Windows Explorer window, click on the photo or document you want to add a tag to.

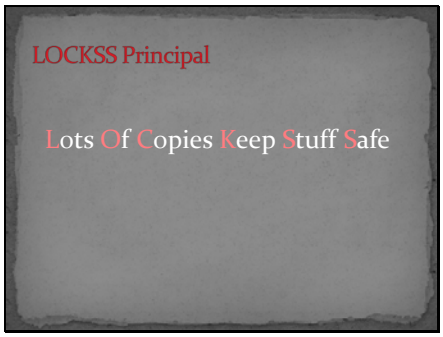
Click in the **Tags: Add a Tag** field located at the bottom of the screen.

Type your tags separated by semicolons.

Click the **Save** button located at the bottom on the right.

In the upper right hand corner of every Explorer window there's a search box. Enter and refine your terms there to find the files you need. Vista's search supports advanced operators, too, like author:, kind: (as in kind:document) and finally, tag:. **Example**: tag:William tag: French

Slide 33



An important preservation principal developed at Stanford University is the ancronym LOCKSS. It stands for Lots Of Copies Keep Stuff Safe. The basic concept is this—the more copies of any type (digital or physical) you archive in different locations, the safer your records will be.

To apply LOCKSS, you should write a minimum of two discs per set of files and store them in two different locations as far apart as practical.

3 discs in 3 different locations is even better.

Any single copy of any digital file is almost guaranteed to be unavailable within a few years.

Moral of this story? Make lots of copies of your records….

Slide 34

To summarize:

Organize

- Setup Folder and file naming conventions using descriptive names
- Decide which file formats you will use

Acquire storage media and equipment you will need to record and view your files

Scan or photograph files with highest resolution

Enrich files with descriptive information either through PDF or tags or descriptive folder and filenames

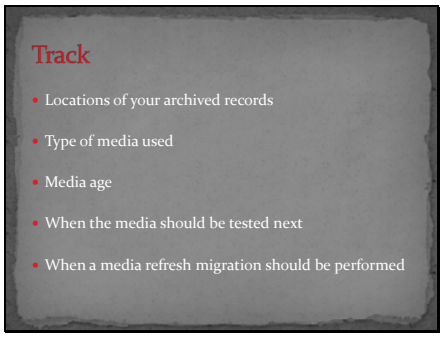
Write the files to your storage media

Verify this copy is readable

Make more copies

Distribute copies

Slide 35



You're not done yet! You need to keep a record of the following information:

Locations of your archived records

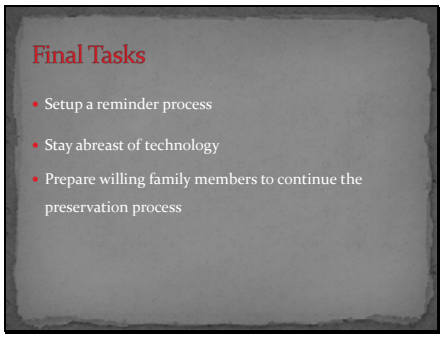
Type of media used

Media age

When the media should be tested next

When a media refresh migration should be performed

Slide 36



Finally, set your calendar to remind you when you need to access the files to make sure they are readable without errors. If some records are beginning to have errors, use an alternate copy to make a fresh copy. Be proactive—don't wait until the errors start to occur.

Again, you must stay abreast of the newest and latest so that you will be prepared for obsolescence of either the file format or storage device you have chosen to store your records.

Enlist a family member to continue your preservation efforts.

Slide 37