Digitization 101 2011-04-11

From Digital Projects Lab Projects Wiki

Contents

Introductions

Mark E. Phillips

Mark Phillips is the Assistant Dean for Digital Libraries at University of North Texas Libraries. He has been involved in the area of digital libraries for the past seven years with projects like the Portal to Texas History and the UNT Library. His areas of interest include open source digital library development and optimizing digitization workflows. <mark.phillips@unt.edu>

Jeremy D. Moore

Jeremy Moore is the Lab Manager of the Digital Projects Unit at the University of North Texas Libraries. He has worked in the digital libraries' arena for three years with projects like the Portal to Texas History and the UNT Libraries' Digital Collections. He is currently completing a MFA in Photography and a MA in Art History at the University of North Texas. <jeremy.moore@unt.edu>

Nathan Hall

Nathan Hall is a Digital Curation Librarian at the University of North Texas Libraries. He has been involved in digitization projects since 2004. A recent highlight of his work is the Environmental Policy Collection. He is currently working on a doctoral dissertation in digital library evaluation at the University of North Texas. <nathan.hall@unt.edu>

Computer Hardware

A top of the line machine is not needed for digitization, but there are a couple of features which are suggested.

Hardware Suggestions

- Desktop computer running Windows 7 or Windows XP
    - Laptops are great, but you want a desktop as they are cheap for how fast they are vs. a laptop, plus they allow the addition of more internal hard drives

(see below).
- Middle of the pack processor
  - You don't need the top of the line processor as the most expensive is always at a premium because it is the fastest. Choose a processor 2-3 steps down from the newest/fastest for the best value.
- 4Gb of RAM (Random-Access Memory)
  - This is the hardware that your programs are loaded into when you are actually using them. Windows XP will only utilize 3Gb so having more than 4Gb is not better with XP.
- 2x internal hard drives
  - This allows you to have a "boot" drive where all of your system files and programs are located and a "data" drive where you can keep all of the actual files you are creating.
  - The internal "boot" drive does not need to be massive, but (currently, Feb, 2009) 500Gb is a good size for your "data" drive.

Desktop Specs at the DPU

- Dell Desktop with 2x 17" monitor (having dual monitors is a boon as you can work with images on one screen and metadata on the other)
- Windows 7
- Intel Core i7 Processor @ 3.07GHz
- 4.0Gb of RAM
- 250Gb Boot Drive
- 1TB Data Drive

Setting up the Computer - Windows XP

There are some nice features built into Windows XP (WinXP from now on) that aid with digitization projects, but you'll have to turn them on yourself.

Adding Details View

When you open a window full of files the default view in WinXP is to show thumbnails, which is helpful when visually looking for an image, but the majority of the time it just clutters up your window and doesn't let you see information pertinent to digitization. Instead of using the thumbnail view, we use details view.

- In any open window find the icon on your toolbar with the green back button that looks like a small window with multicolored thumbnails. It will be to the right after Search and Folders.
- Click the icon, this will cause a drop-down menu to appear, choose "Details".
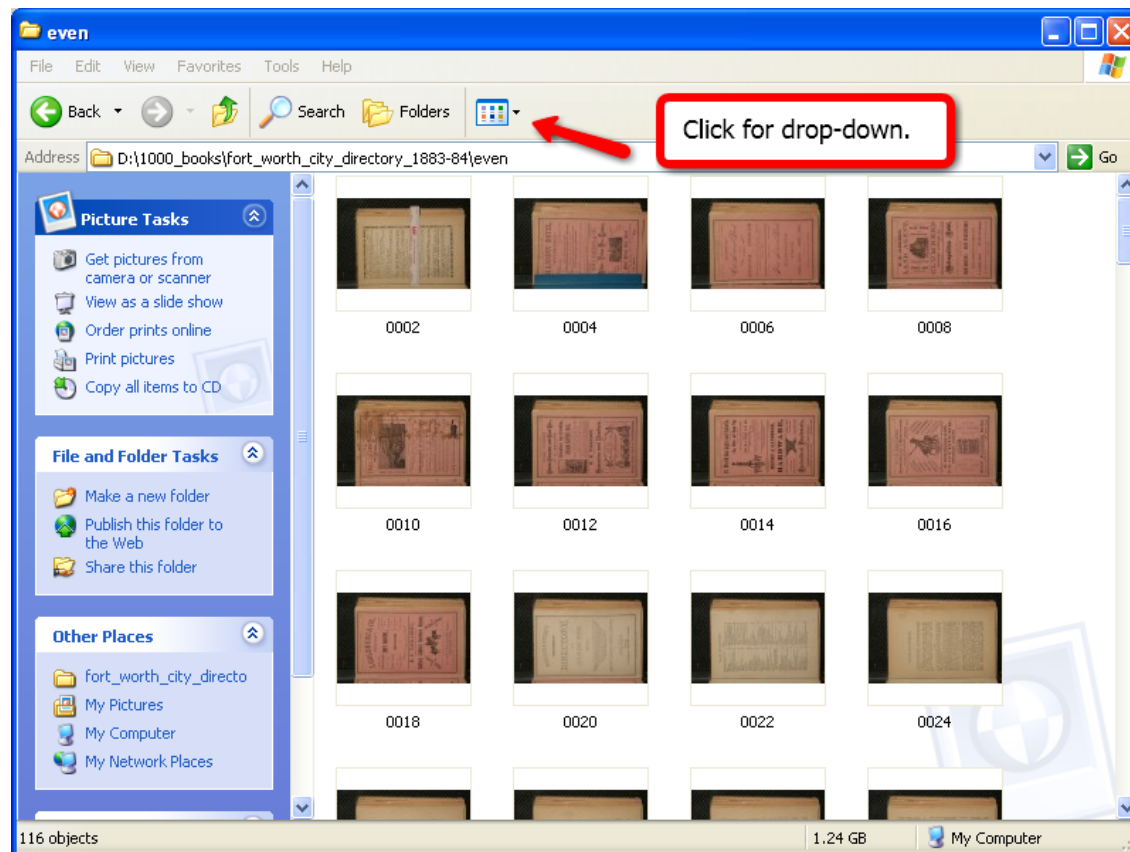
- Now you can add the pixel dimensions of any image files in the folder as a column. To bring up this option right-click on the toolbar right above the files where it says "Name".
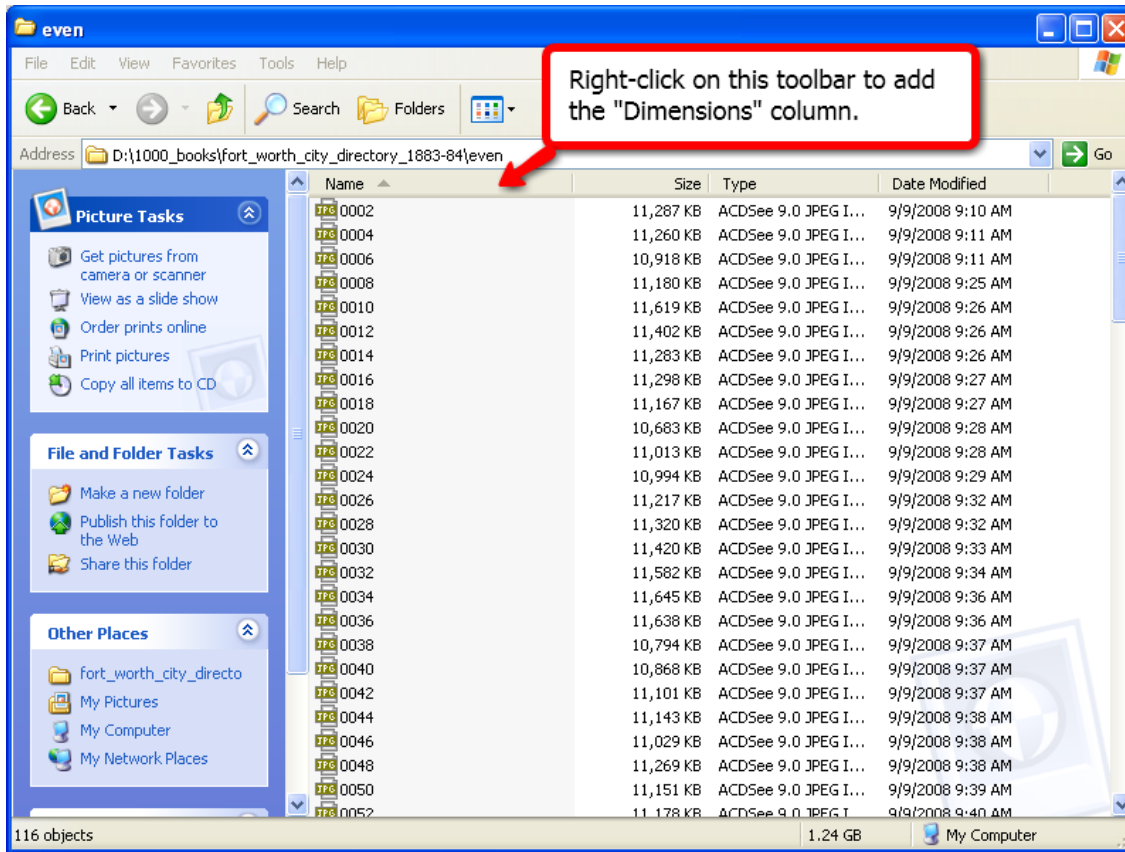
- Choose "More" at the bottom and in the new window that pops up scroll down until you can click the box next to "Dimensions" and then click "OK".
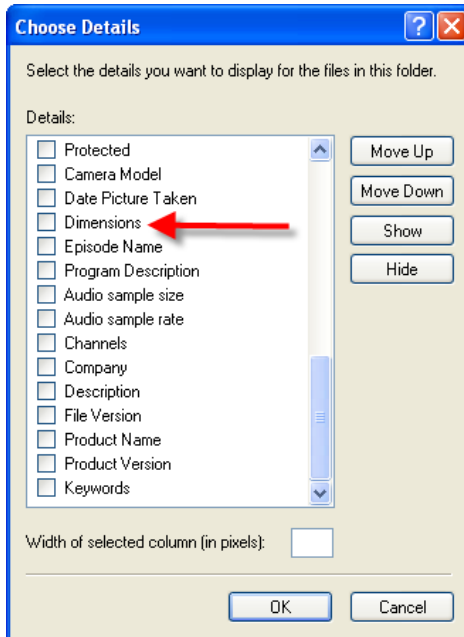
- This adds the dimensions column, but not in an area that is too helpful so you can click on the column name and drag it to the left. I place them in this order: Size, Dimensions, Date Modified, and then Type. I place Type last as I can get the type of file from the extension. At the moment you cannot see the file extension, but we'll add that in the last step when we make our setttings global, i.e. for every window WinXP opens.

Name ▲	Size	Date Modified	Dimensions	Type
00010001	224 KB	2/1/2009 10:34 PM	5084 x 6504	ACDSee 9.0 TIF Im...
00020002	248 KB	2/1/2009 10:34 PM	5084 x 6504	ACDSee 9.0 TIF Im...
00030003	185 KB	2/1/2009 10:34 PM	5084 x 6504	ACDSee 9.0 TIF Im...
00040004	180 KB	2/1/2009 10:34 PM	5084 x 6504	ACDSee 9.0 TIF Im...
00050005	184 KB	2/1/2009 10:34 PM	5084 x 6504	ACDSee 9.0 TIF Im...
00060006	212 KB	2/1/2009 10:34 PM	5084 x 6504	ACDSee 9.0 TIF Im...
00070007	27 KB	2/1/2009 10:34 PM	5084 x 6504	ACDSee 9.0 TIF Im...

Add Status Bar

- Add a status bar to the bottom of your window which will show you the number of images in the folder plus the number you have selected at any time by clicking "View" in the menu bar and clicking "Status Bar" in the drop-down menu.

Making Folder Options Global

- In the menu bar click "Tools" and click "Folder Options..." in the drop-down and a new window pops up.
- Choose the "View" tab.
- In the Advanced settings area make sure "Hide extensions for known file types" is unchecked (clicking adds/removes the checkmark) then click the "Apply to All Folders" button in the top left. A window will pop up asking if you want to "Set all the folders on your computer to match the current folder's view settings (except for toolbars and folder tasks)?" and click "Yes". Click "OK" to close the window.
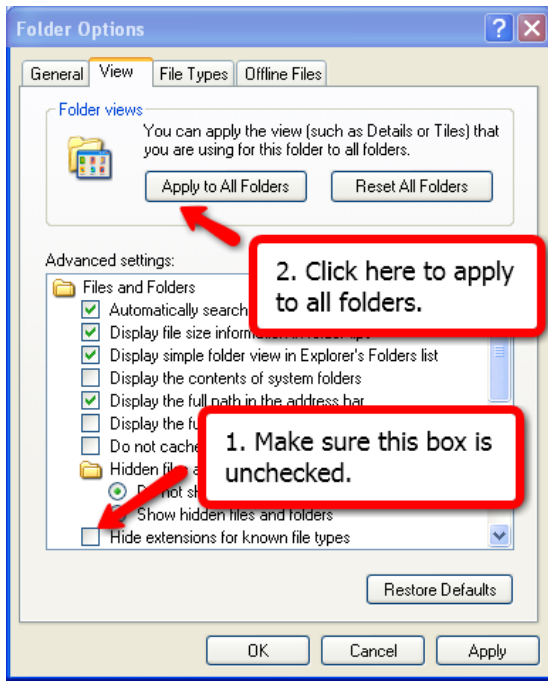

Setting up Windows 7 for Digital Projects

The following steps are taken to customize windows to be more efficient at the tasks commonly performed in the DPU Lab. There are some nice features built into Windows 7 that aid with digitization projects, but you'll have to turn them on yourself.

Setting Up Toolbars and Views

Each section will explain how to add something to the file window in order to make it more efficient for digitalization projects. After each explanation, there is an illustration to show where the buttons are actually located.

In Windows 7, when you view a file window for digital projects, the default window toolbar will show "Organize", "Burn", and "New Folder" (others may be shown, but these are the basic default buttons on the toolbar). See image below.
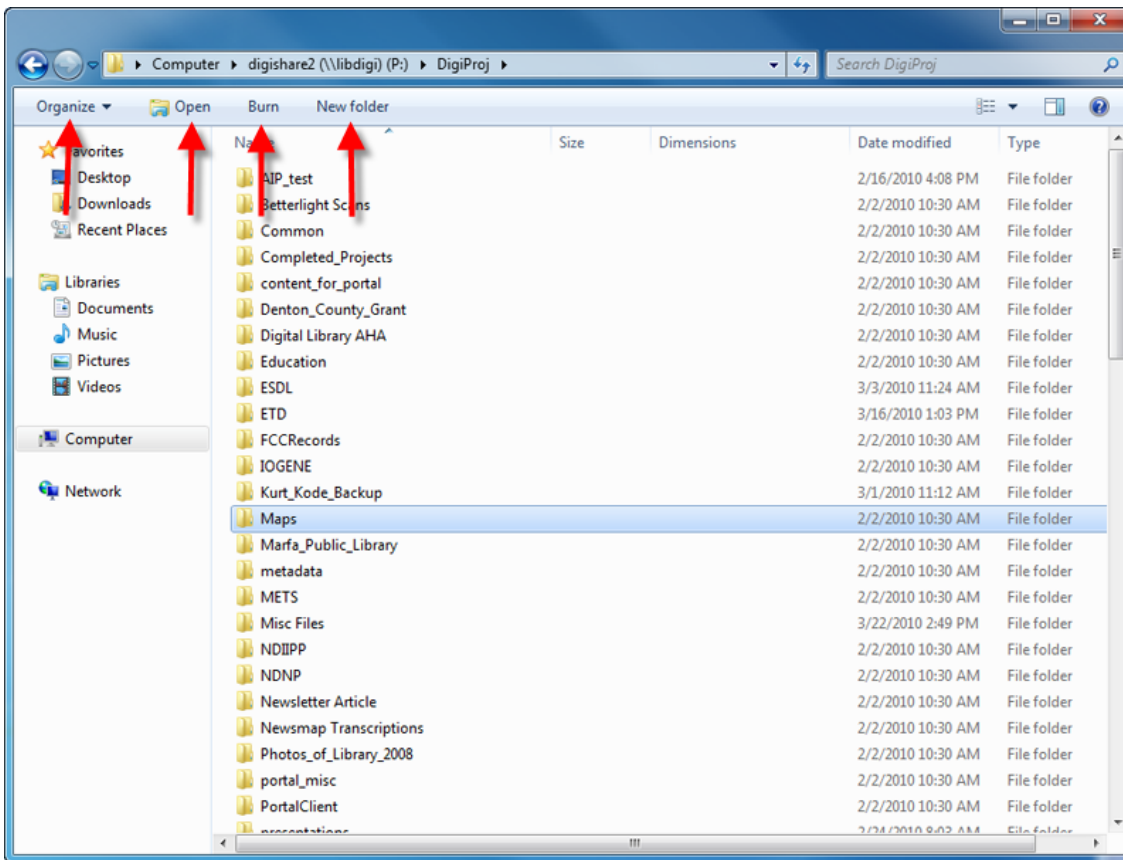
You will need to add the "Menu Toolbar" to the file window.

- Click on "Organize" in the file window and a drop down menu will appear.
- In the drop down menu, move the mouse over "Layout" and another drop down menu appears.
- Be sure that "Menu bar" is checked (click to add/remove check mark).

The "Menu bar" (File, Edit, View, Tools and Help) will appear above the default toolbar at the top of the file window. See image below.
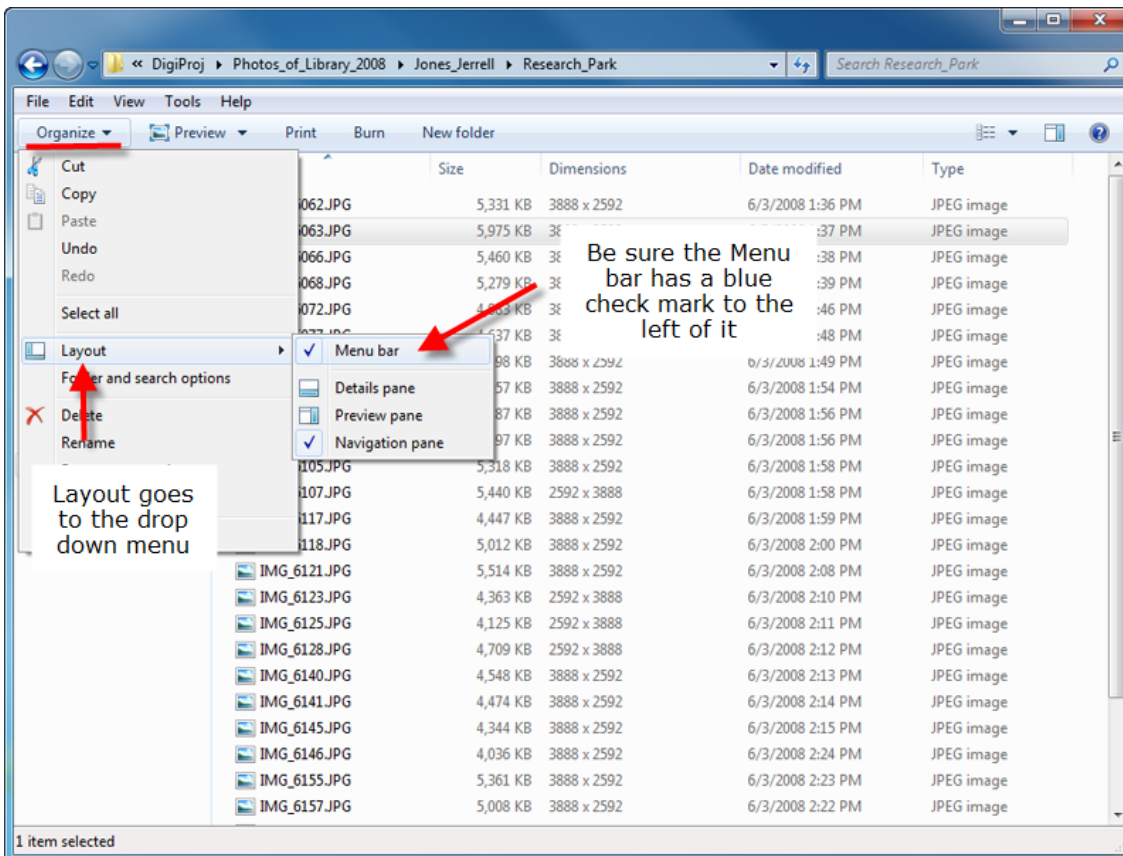
Adding Details View

When you open a window full of files, the default view in Windows 7 is to show Details, which is helpful for most digitalization projects.

- If it is not set to Details, click on "View" in the Main Toolbar. A drop down menu will appear.
- In the middle of the drop down menu, click on "Details" and be sure that a blue dot is to the left of it.

See image below.
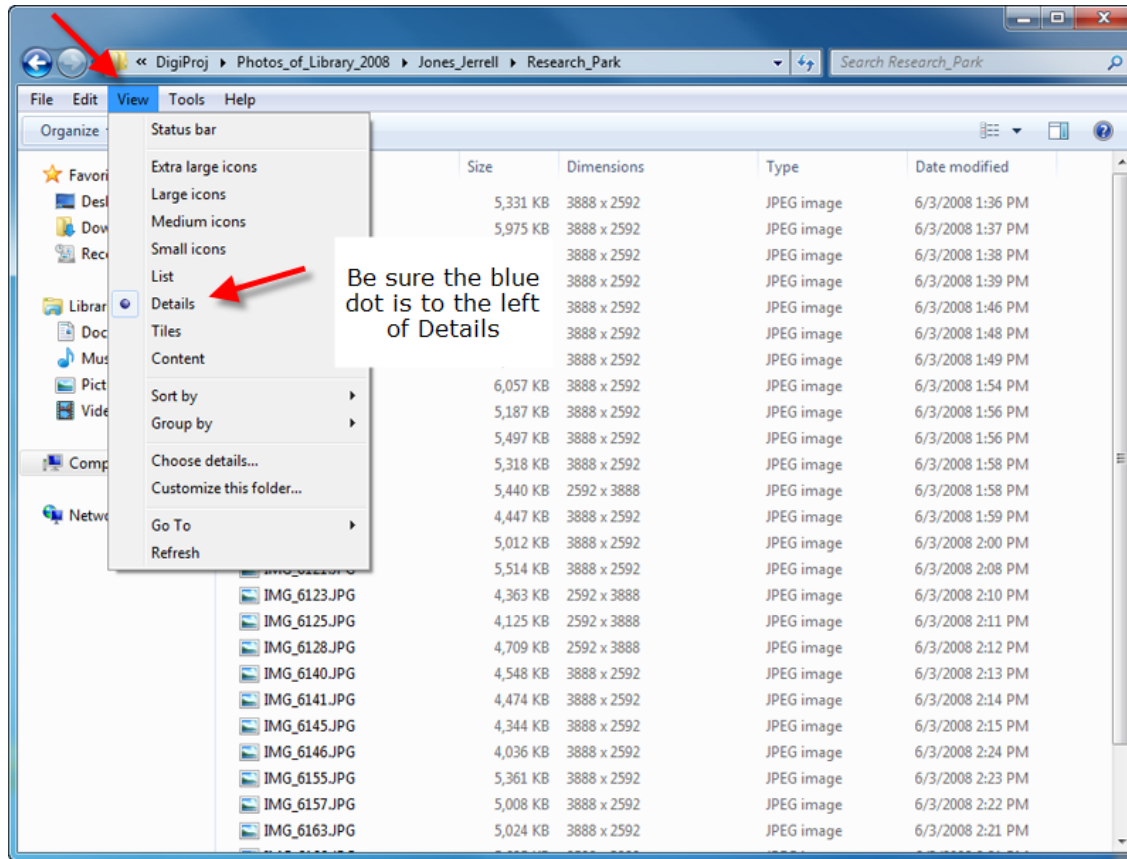
Adding the Dimensions Column

Now you can add the "Dimensions" column in your file window. This will show the pixel dimensions of any image files in the folder.

- Right-click on the toolbar right above the files where it says "Name".

See image below.
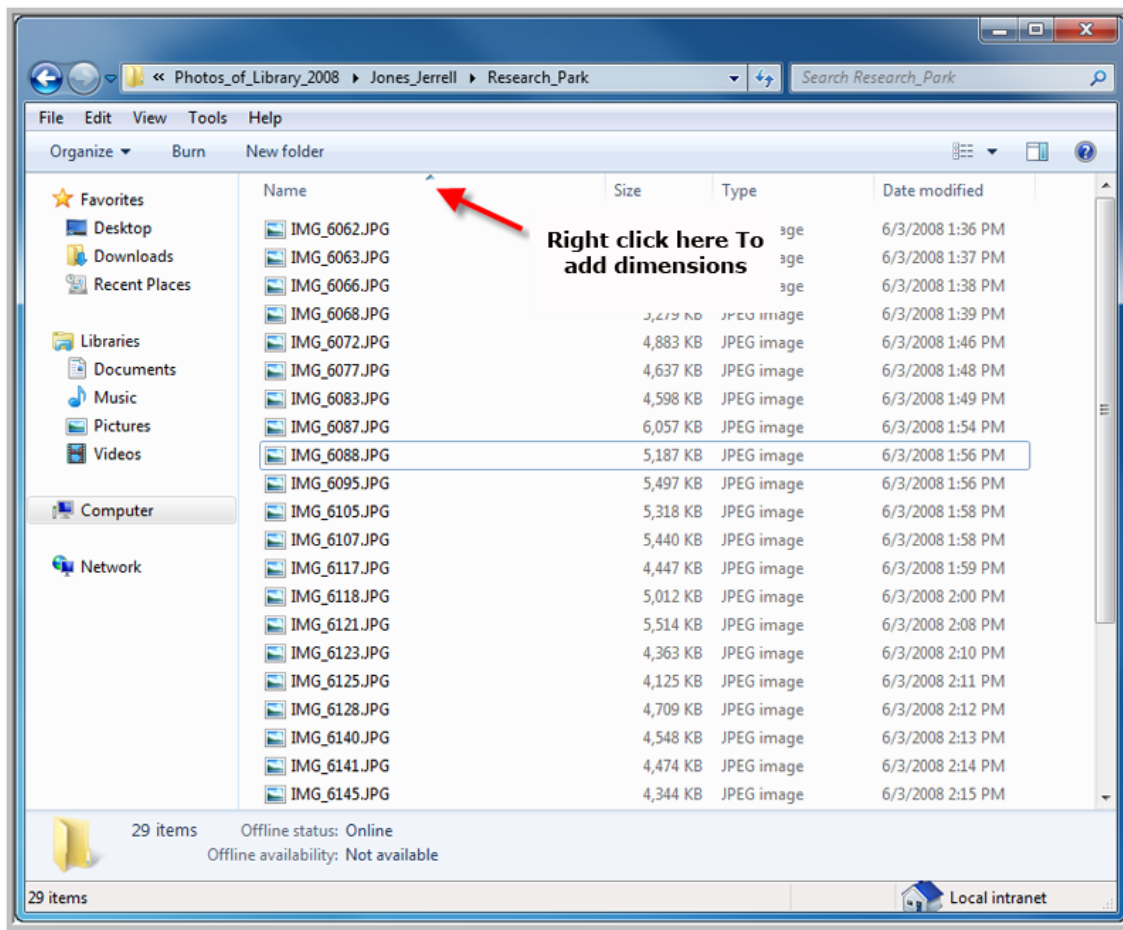
- After you have right clicked on "Name", a new window will pop up.
- At the very bottom of the pop up window, click on "More" and a new window will pop up.
- In the new window, under "Details:", scroll down until you see "Dimensions" and click on it so that it has a blue check mark to the left of it.

See image below.

This adds the Dimensions column to your file window, but not in an area that is too helpful so you need to move it and arrange your columns in the file window.

- Click on the Dimensions column and drag it to the left.
- Click and drag each of the columns to arrange them in this order:

1.Name 2.Size 3.Dimensions 4.Date Modified 5.Type See example image below. This is how your file window should look.

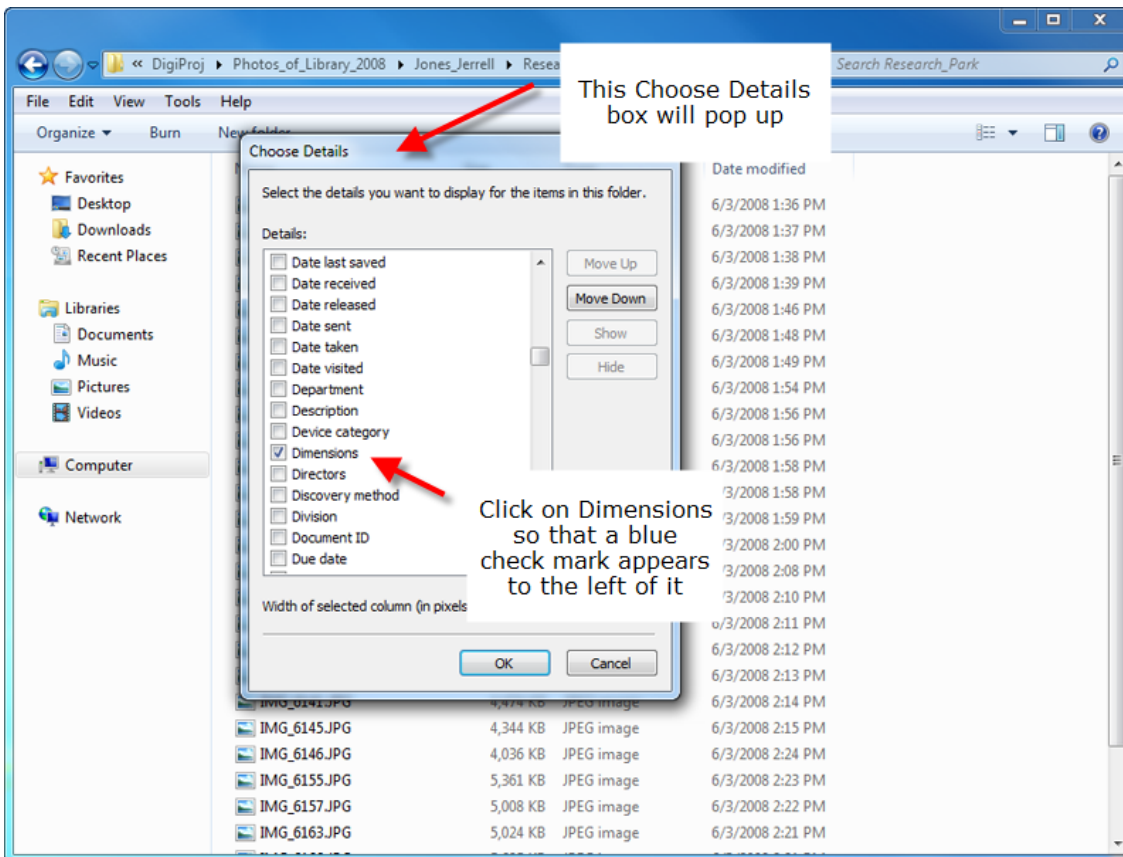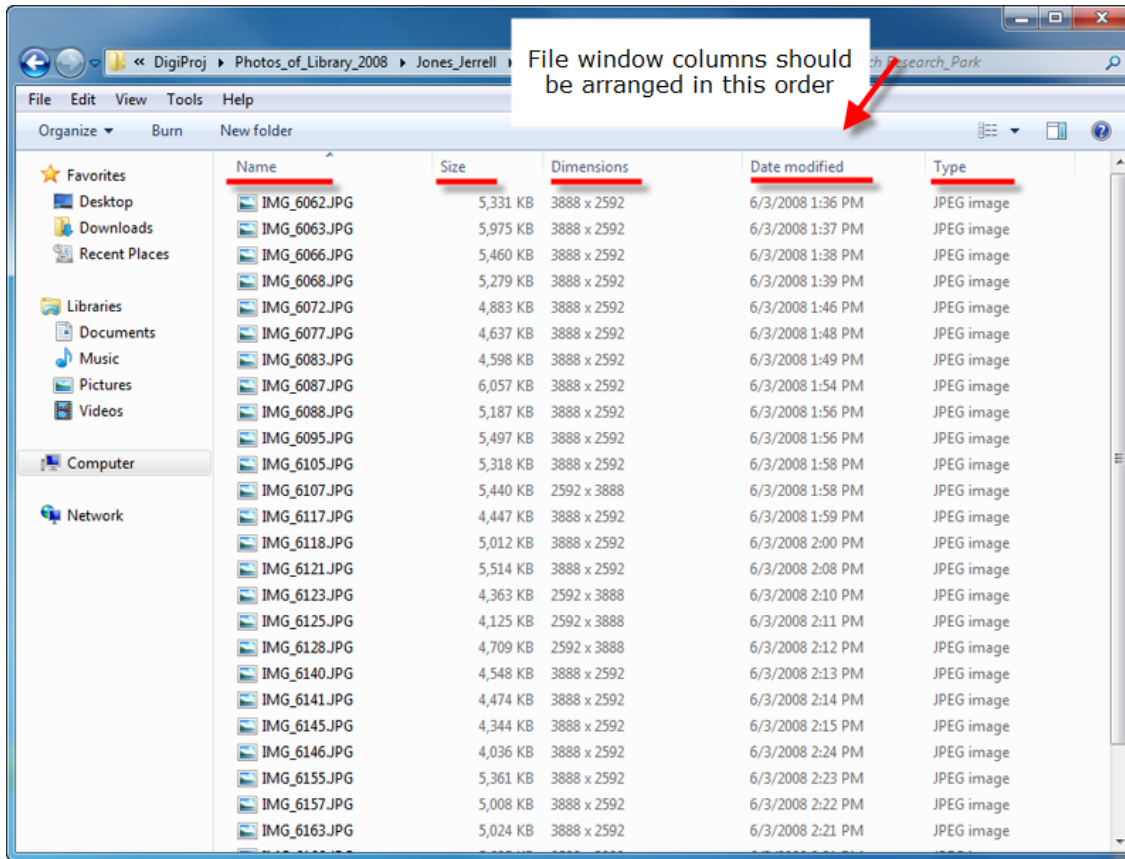The "Type" column is last in the row because you can get the type of file from the extension. At the moment you cannot see the file extension, but we will add that in the last step on this page, when we make our settings global (i.e. for every window that Windows 7 opens).

Add Status Bar

Add a "Status bar" to the bottom of your file window. This will show you the number of images in the folder plus the number you have selected at any time.

- Click on "View" in the Menu bar and a drop down menu will appear.
- The first thing, at the very top of the drop down menu is "Status Bar". Click on it to be sure that a blue arrow is to the left of it (clicking adds/removes the checkmark). Once it has a blue arrow, the status bar has been added to your file window. It will appear at the bottom.

See image below.
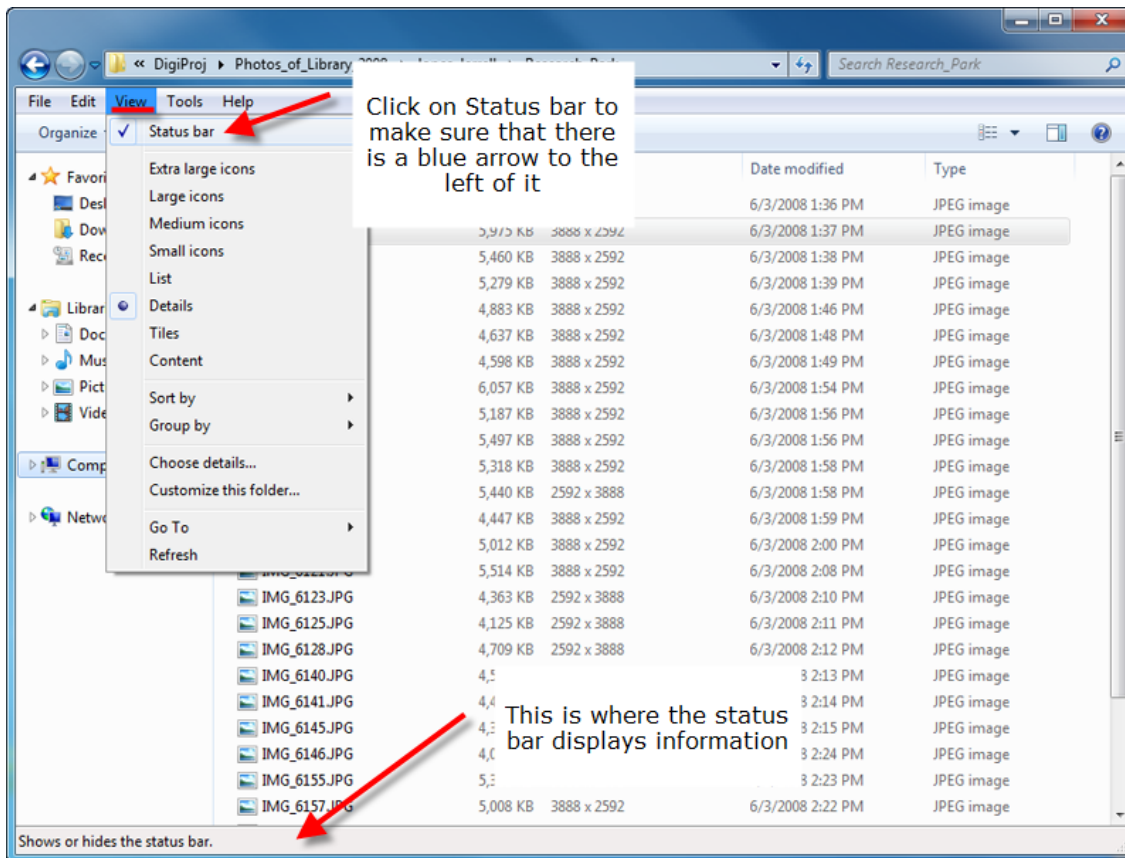
Making Folder Options Global

Now, we need to make the folder options global, meaning they will be used for every window that Windows 7 opens.

- Click on "Organize" in the default toolbar and a drop down menu will appear.
- Click on the "Folder and Search Options" button.

See image below.
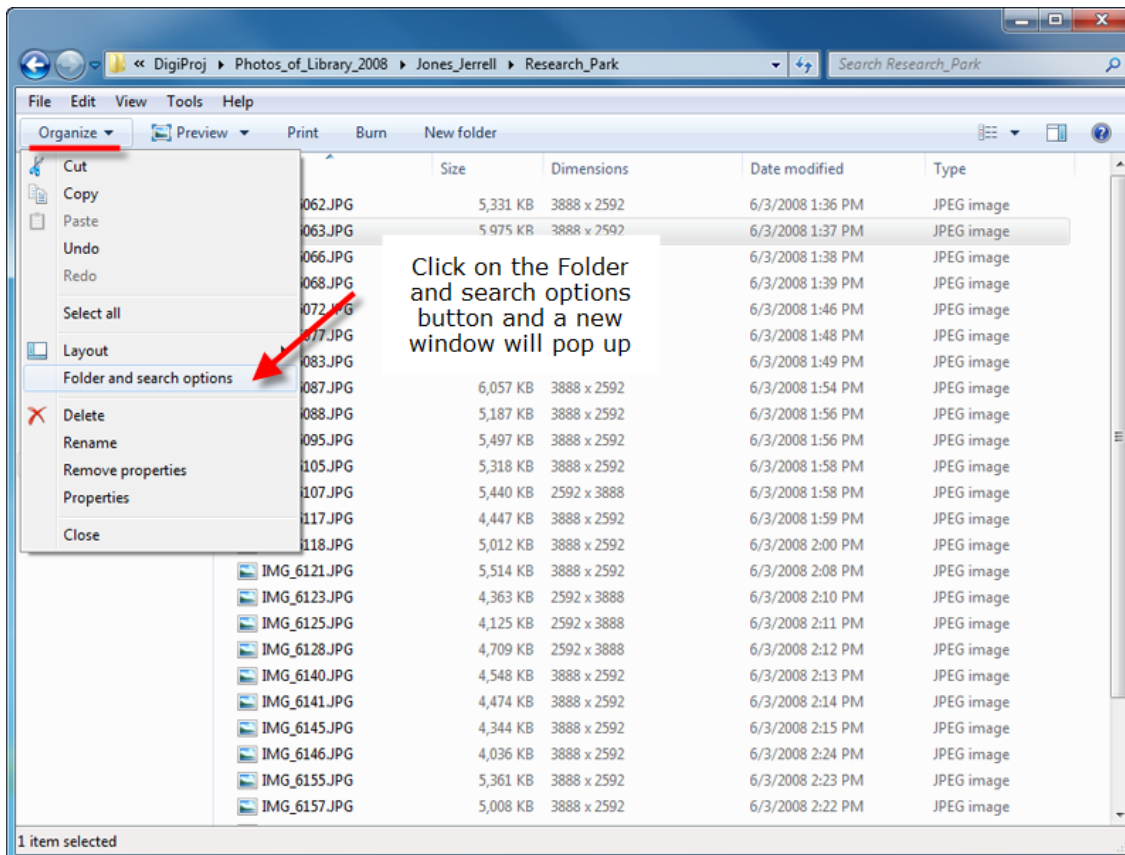
- A new "Folder Options" window will pop up.
- Click on the "View" tab in this window.
- In the "Advanced settings" area, make sure that "Hide extensions for known file types" is unchecked (clicking adds/removes the checkmark).

See image below.
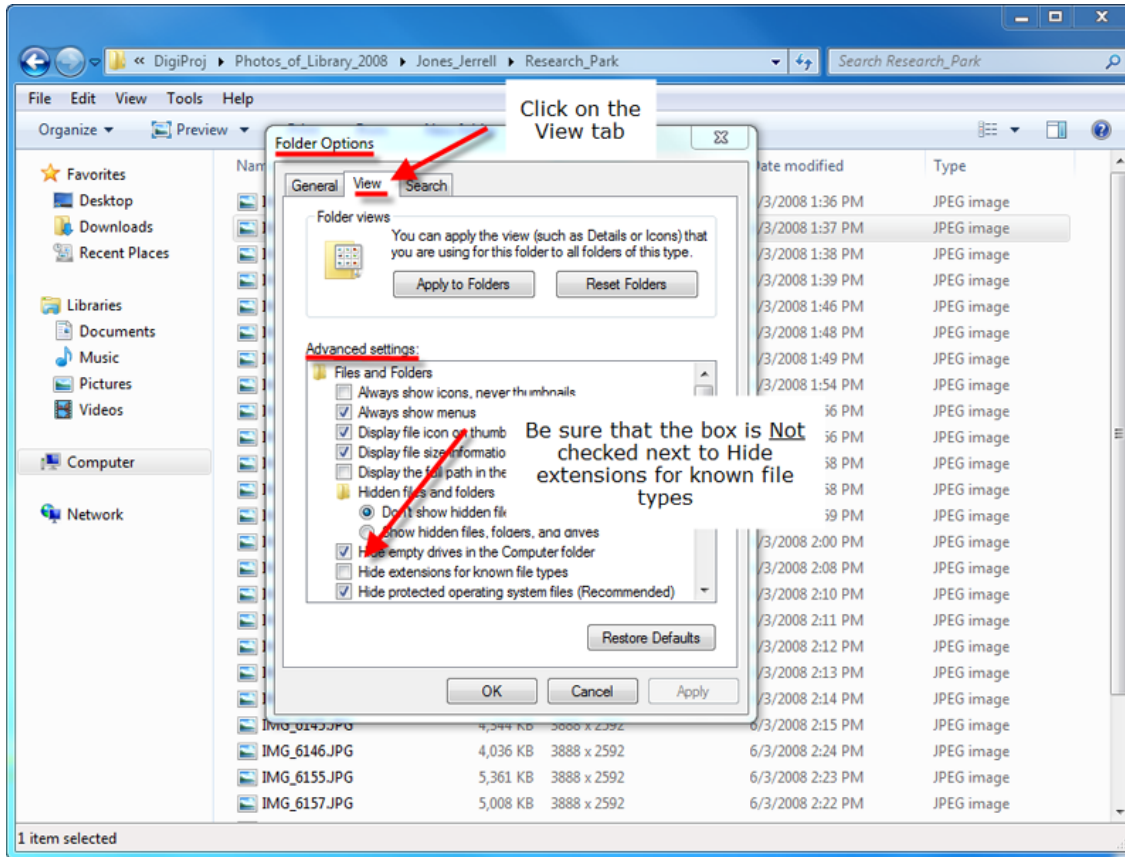
- Click the "Apply to Folders" button in the Top Left under "Folder Views" (this is NOT the "Apply" button at the bottom).
- A window will pop up asking if you want "All folders of this type to match this folder's view settings?" and click "Yes". Click "OK" to close the window and you are finished.

See image below.
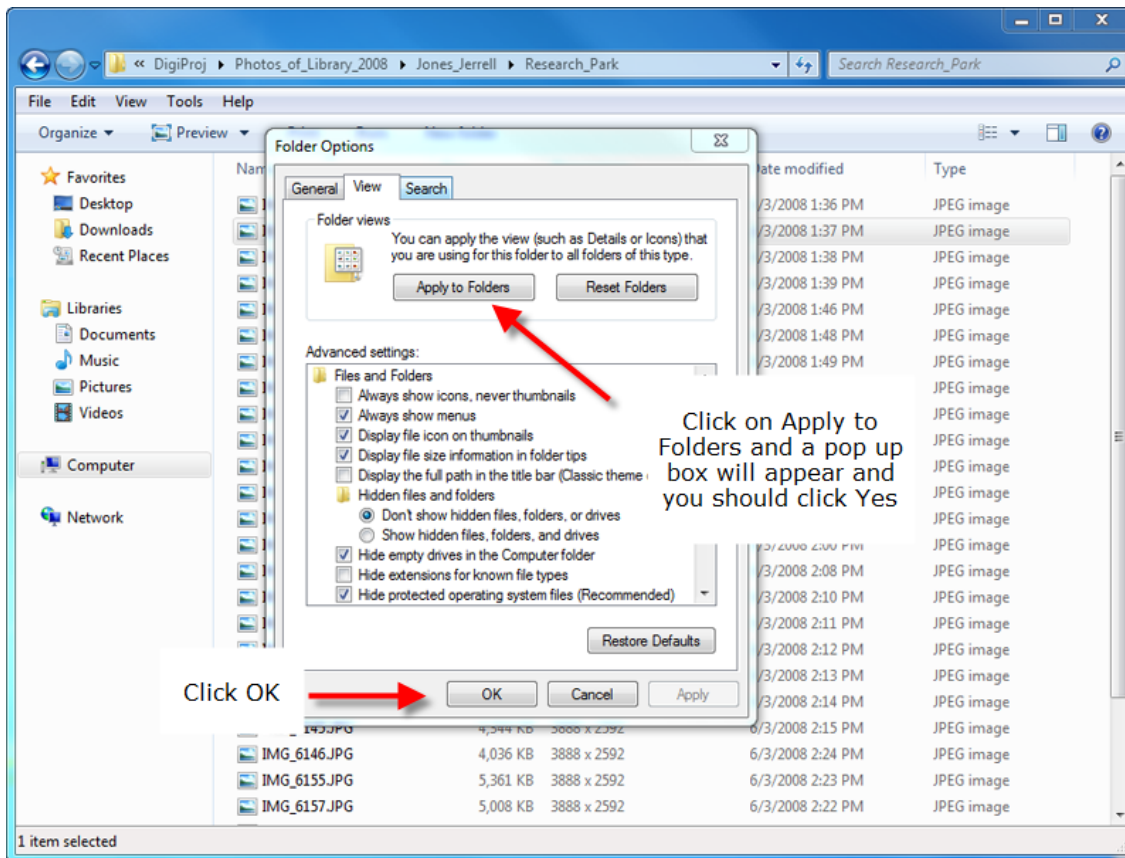
Your file window is successfully organized for digitization projects.

Identifiers

It is important to have identifiers assigned to each object before any digitization is started. A unique and persistent identifier is used throughout the digitization process for filenames, folder structures and linking metadata records with digital objects scanned. Think of having a modern day library without bar codes on the back of books which uniquely identify and link a book on the shelve to a metadata record in the catalog. All too often we see collections in the DPU which have no identifiers and it is challenging and time consuming for our staff to assign meaningful identifiers to other institutions content.

What makes a good identifier?

- Reasonably Unique
- Weather well over time
- Inexpensive to generate

Examples of good identifiers.

- Barcode numbers on books
- OCLC numbers
- LCCN numbers
- Accession Numbers
- Library Catalog Numbers

Examples of less good identifiers.

- Content based identifiers
- Excessively un-Unique

Identifiers in the wild

Most of the identifiers we see in the DPU come from numbers assigned to images within a collection. These are sometimes written in pencil on the back of a photograph or they are added as a sticker to a folder or protective sleeve. These numbers vary based on the institution and how they name collections and items, but some examples are as follows.

- ABB_PEIS-1986-4-15
- DHPS_1998-023-418
- MMLUT_0001
- 1060295

These numbers don't have to be completely unique within an institution, sometimes that's just not practical, what they do need is to be unique within a collection so that it is possible for all of the items in a collection to be identified when the digitization process is being completed.

Folder Management

- One of the biggest issues with digitization is the organization and management of objects you have created for current and past projects. It greatly helps to think this problem through before beginning a project as it can come back and cause problems at a later date.

scanned_for_xxx

- In the DPU we name our main folder "scanned_for_xxx" where "xxx" is the name of the project we're working on. So for example, we would name the main folder "scanned_for_Amon_Carter" for a project funded by the Amon Carter Foundation. If you are only scanning for your own institution you can break it down by project names specific to what you are working on, such as "scans_from_Tomahawk_Family" if you have a collection of photographs from the Tomahawk family.
    - By naming the main folder specific to a project you immediately know the folder contains photographs from the Tomahawk family, and consequently if you're looking for the scans from the Frederick Farm collection you know NOT to look in this folder.
- Inside of the folder scans_from_Tomahawk_family there are then folders delineating each step in the process. The folders are named with a number to keep them in order and also a "To" so there's no confusing what the next step in the process for the items in these folders need to be. I created this example folder just like our production folders we use in the DPU which consists of:
    - 1.ToQC - This is where work just completed goes to live until it can be quality control checked for resolution, skew, cropping, color, and quality (more on this later!).
    - 2.ToDekew - This is for TEXT-based objects that need to be deskewed, i.e. they are skewed and we need to make them straight.
    - 3.ToResize - This is for TEXT-based objects that need to be resized to match the original document.
    - 4.ToOCR - This is for TEXT-based objects that need to have the Optical Character Recognition Engine run on them
    - 5.ToMetadata - This is for any object that is finished and ready to be uploaded, but does not yet have metadata.
    - 6.ToUpload - Folder for files ready to be processed for uploading.
    - 7.Uploaded - Folder for the files that have been uploaded.
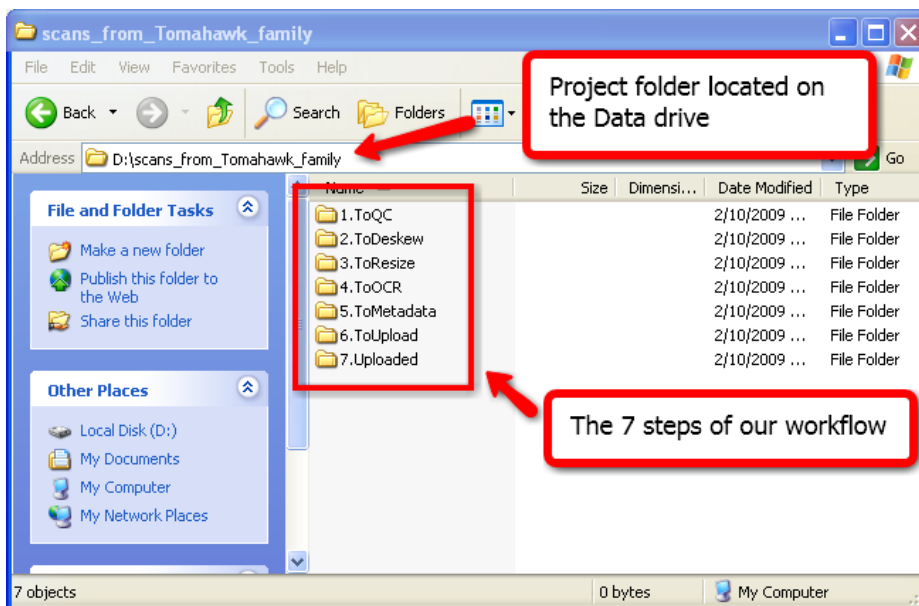
1 Object + 1 Metadata Record = 1 Folder

- Math works a little differently in the DPU when it comes to folder management because, for us, 1 object + 1 metadata record = 1 folder. This means each digital object gets its own folder regardless of how many digital images are associated with it.
    - For example, the front of a photographic print may be scanned as one digital image or you may need to scan the front and back providing two digital images; in both cases you use one folder as you are dealing with one object.
- This method means you ALWAYS have your metadata file with the object it describes and whenever you move the folder it keeps all of the digital images of an object and the metadata together without any extra work.

Naming Object Folders and Files

- Only use A-Z, a-z, 0-9, an underscore ("_") hyphen ("-") period (".") when naming your folders and files.
    - This is for consistency and files cannot be named with non-alphanumeric/special characters
    - The main project folder is the only location where we use the period (".") in our folder names, i.e. "5.ToMetadata".
- The name for the object folder and the name for the included file[s] will match, though if there is more than one digital image then each digital image will need a suffix to separate them. If you have an accession number or internal name/signifier for the object then use that name as it will make it much easier to move back and forth between the digital and the physical object if need be.
    - Example: If the accession number for your object, a photograph, is 1919-Tomahawk-1 then the folder name will also be 1919-Tomahawk-1 and the digital image will be named 1919-Tomahawk-1.tif unless there is more than one image for this object. If there is information on the back of the photograph and you scan both sides then your folder would be 1919-Tomahawk-1 and inside of that folder will be two digital files: 1919-Tomahawk-1_01.tif and 1919-Tomahawk-1_02.tif.
        - Do not add a number suffix if there is only one scan. This way you know there was only one image whereas if you add a _01 suffix to a single image three years later someone may wonder where _02 is....
        - When numbering a series of files always "pad" the number with zeroes in front or they will not show up in order when you open the folder. So if there are less than 10 objects you only need to use two digits (01, 02, 03, 04, etc.), but from 10-99 files you will need to use three digits (i.e. 023, 024, 025, 026, etc.).
- Use the three letter extension for a file: tif, jpg, png, txt.
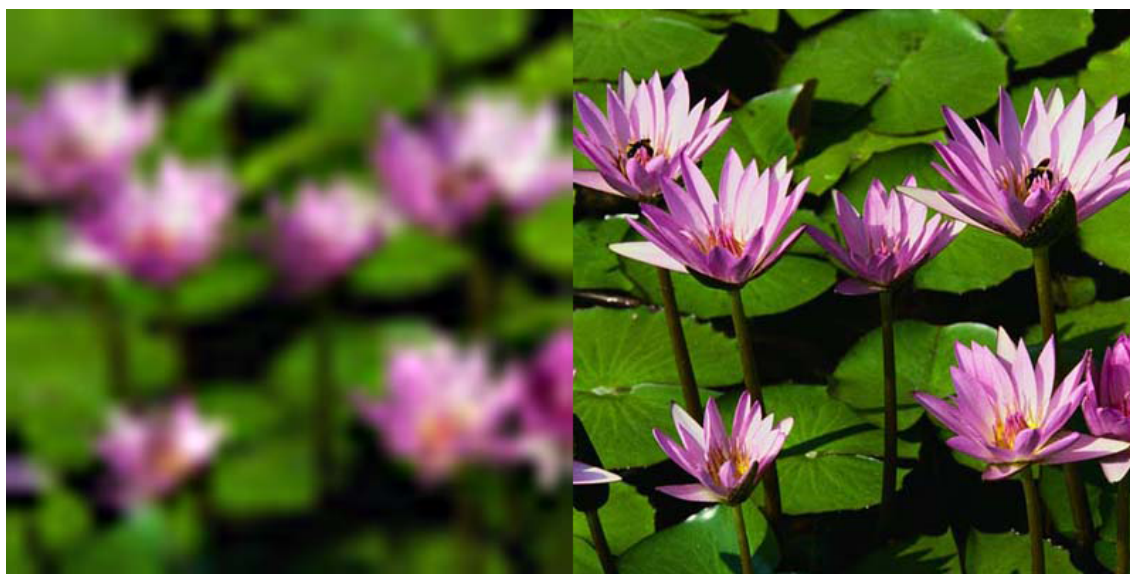
Standards

What is Resolution?

Resolution refers to the density of pixels in an image and is a measurement of height by width at a certain pixel density (ppi). Digital images have no real absolute size or pixel density, only a certain number of pixels in each dimension.

- ppi? dpi?
  - PPI = pixels per inch
  - DPI = dots per inch
    - These terms are incorrectly used interchangeably but refer to two entirely different measurements.
      - PPI is used when the image is still in the computer as you are measuring pixel density.
      - DPI is used when speaking about physical prints as you are talking about an actual number of dots of ink per inch.

The take-home with resolution is you need to have a high enough pixel density (ppi) for the images you scan to meet your needs. You can always res an image down to a lower ppi at the same size, but you can never increase it after the original scan has been made.

One way to think of this is akin to the creation of a mosaic. If you are making a 4' by 4' mosaic tile and you use 1' tiles then your mosaic won't look as much like an image as just 16 big blocks of color. Now if you use 1" tiles you now have 192 blocks to make a picture from and a much better image is created.

The image below exemplifies how this affects the way a digital image looks. The left side is the equivalent of the 4' by 4' mosaic tile using 1' tiles, i.e. 400 pixels x 400 pixels at 8ppi, where a vague idea of the picture is formed, but nothing concrete. The right side is the digital equivalent of using the 1" tiles, i.e. 12x the resolution giving 400 pixels x 400 pixels at 96ppi.

UNT Libraries' Standards

The Digital Projects Unit at the UNT Libraries has developed a simple set of scanning specifications that it uses throughout all of the digital projects carried out in the Lab. The standards were created to provide the highest quality reproductions while keeping in mind the cost of storing the digital copy. Many national and international standards were consulted when creating these standards and we feel that they stand up well as a basic set of scanning standards.

http://www.library.unt.edu/digitalprojects/standards

Any of the suggested scanning resolutions can be taken as a base when doing your scanning, however the cost benefit of scanning at higher resolutions may not result in additional information captured from the original.

Scanning

What Makes a Good Scan

Before diving into the nitty gritty of scanner drivers and Photoshop adjustments it helps to understand what makes a good scan, but how do you define such a moving target as a "good" scan? A scan is a digital representation of a physical object so a good scan would be a faithful reproduction of that object without distractions.

Examples of distractions:

- Skew (or crookedness)
  - If your scan is crooked it not only looks unprofessional it is also very distracting for the viewer, additionally you are increasing the image size of the file when it is crooked compared to the same image which has been deskewed.

- Poor cropping (too tight, too loose)
  - If you crop too tightly then you are not doing justice to the object as information is being stripped away and the image is no longer being faithfully reproduced.
  - If you crop too loosely then the image size of the file is larger than it needs to be and the image will be displayed smaller than it otherwise could be if your system displays fixed width images. If your system displays images on the web at 700 pixels wide and you have a border of 10 pixels all the way around

then there is still 680px of the actual object being displayed, but if your border is 75 pixels then you are losing 150 pixels of width that could display the object, over 20% of the total size!

- Incorrect Image Tones (Too dark/light)
  - If the image is much darker than the original then details in the shadows are now lost and the image is no longer a faithful representation.
  - If the image is much lighter than the original then details in the highlights are now lost and the image is no longer a faithful representation.

Scanning hardware

In the DPU we use a range of Epson flatbed scanners for the multitude of our work as they are fast, high quality, robust, and multi-functional.

The Epson Perfection V700 Photo scanner which we have here with us today is one of Epson's current flatbed scanner and is currently $549.99 direct from Epson and even cheaper through resellers such as Amazon.com where it is $500. This scanner can be use to scan any reflective (prints) or transmissive (negatives, slides) material up to 8.5" x 11.7". It also comes with holders to scan 35mm negatives and slides, medium format negatives and slides, and 4x5 negatives and slides.

For scanning of bound material we typically use a Plustek OpticBook 3600 Scanner which is a "no edge" book scanner meaning that it allows an object to be placed on the scanner in such a way that when scanned the gutter at the binding is not easily seen. This scanner also works well for content that is loose leaf in nature. The scanners have an easy to use interface with built-in software for controlling the scanner.

Scanning Software

In the DPU we use Adobe Photoshop CS4 which is $635 at Amazon.com, but these workflow steps work for any version after Photoshop 7. If you are qualified, Adobe does have educational pricing (check www.adobe.com for details).
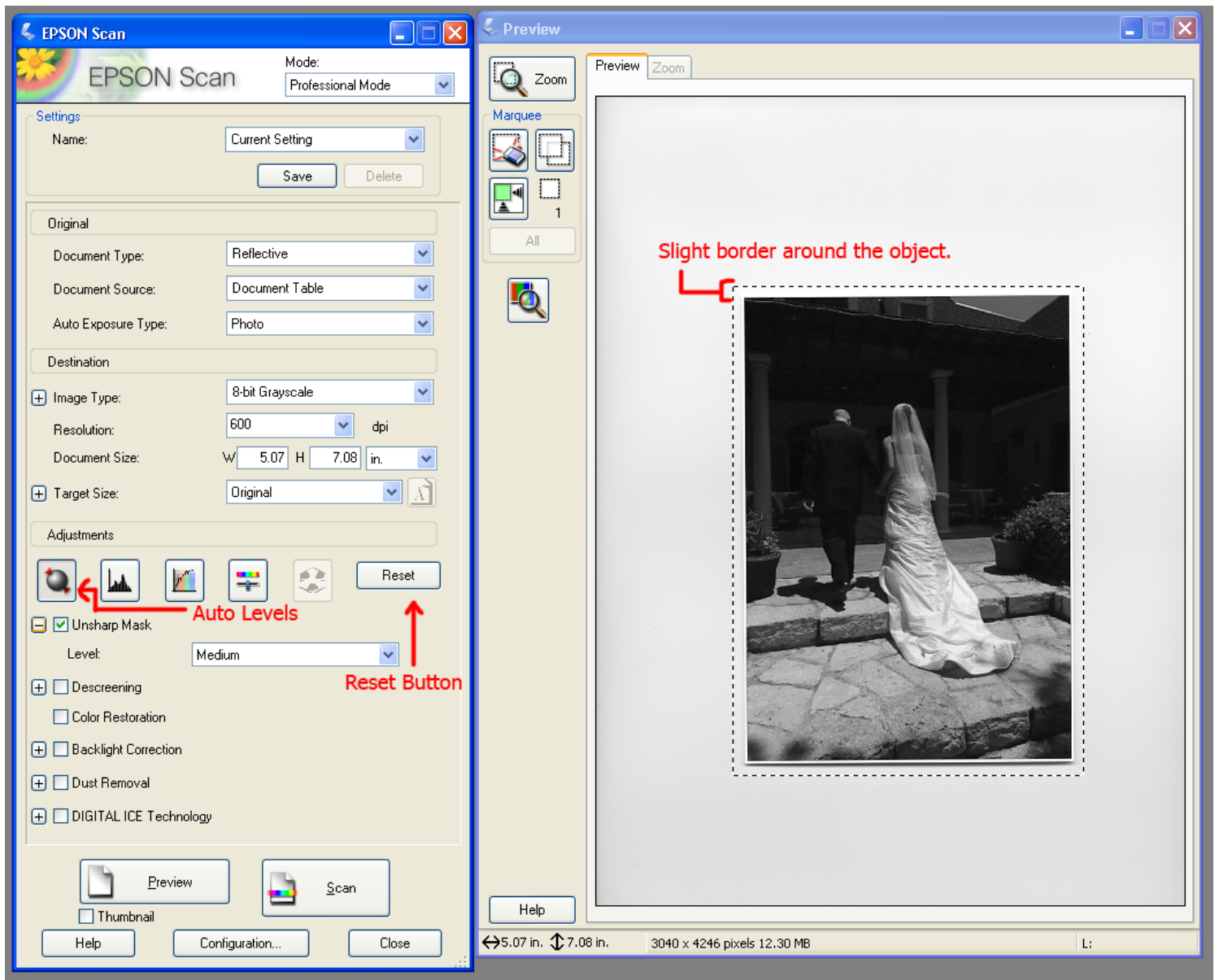
We use Epson's TWAIN driver (fun fact: TWAIN stands for Technology Without An Interesting Name) to run the scanner through Adobe Photoshop CS4 (from now on just referred to as Photoshop) so as soon as the scanner is finished scanning the file is in Photoshop.

"I'll fix it in Photoshop"

Another point that should be expressed is that the mentality of "I'll fix it in Photoshop" is a bad one to have. There are many things that can be tweaked and corrected, but always plan on getting the best scan you can up front. If you're often having to resort to Photoshop to "fix" things then your scanning workflow may need some looking into.

Scanning a photographic print

- Open up Adobe Photoshop CS4. Go to File > Import > Epson Perfection V700/V750.
- A new window pops up that is the Epson scan driver. Make sure the scanning software is set to "Professional Mode".
- Place the object on the scanner (taking care to line it up straight), make sure the reflective mode cover is blocking the transparency adapter (if the underside of the lid is white then it is installed, glass means it is not installed) close the lid, and click Preview in the scanning software.
- Drag a box of "marching ants" around the object in the preview screen. Add a small and even margin around the outside of the object. There is no perfect size for the margin, but you want it to be obvious the entire object has been scanned without going overboard. Consistency is the key here.
- Set the scanner settings:
  - Original
    - Document Type: Reflective
    - Document Source: Document Table
    - Auto Exposure Type: Photo
  - Destination
    - Scan in color or grayscale?
      - 24-bit color: Handwritten document and/or there is color anywhere on the page
      - 8-bit grayscale: Everything else
    - Resolution
      - In accordance with the specs we discussed earlier, we will scan our prints at 600 dpi/ppi (computer software will use "dpi" and "ppi" interchangeably)
  - Adjustments
    - Reset any previous tones settings
      - Click the "Reset" button. It is ¾ of the way down the settings page and is all the way to the right.
    - Auto-Levels of the tones.
      - Click the "Auto-Levels" button. It is the first button on the left on the same row as the "Reset" button.
    - Unsharp Mask
      - Click on (checkmark means it is on) and set to the Medium setting in the drop-down box. Every digital capture needs to be sharpened to take it back to the sharpness of the original.
  - Click Scan
- Close Epson Scan. If you do not close Epson Scan you will not be able to do anything in Photoshop.
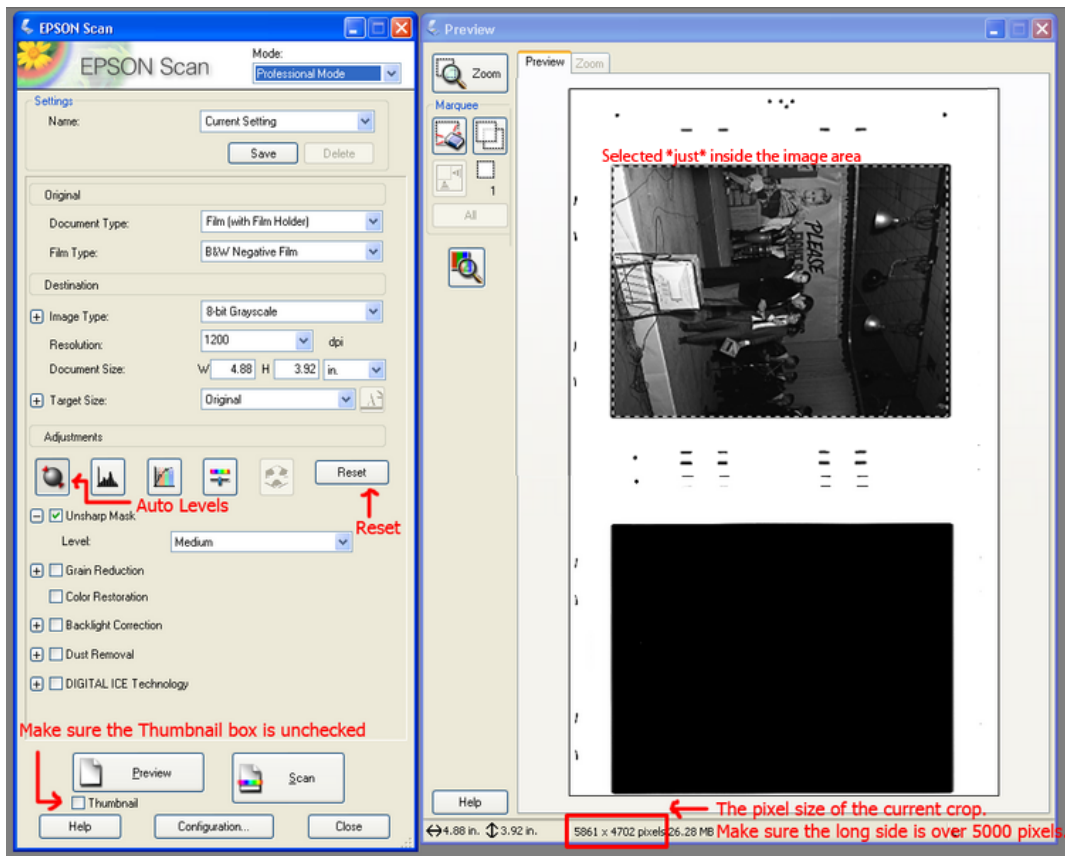
- Rotate the scan
  - If your image is not correctly oriented you can rotate it by going to Image > Rotate Canvas> and then picking the option to make the image correctly oriented.
- Deskew the scan
  - If the page is crooked then use the Ruler to straighten it. The Ruler tool in Photoshop is found in the toolbar with the Eye Dropper. Click and hold on the eye dropper and when a menu opens choose the Ruler tool.
  - Click on one edge of the page and still holding down the mouse button draw a line along the edge of the page.
  - Go to Image > Rotate Canvas > Arbitrary...
  - A window will pop up with the information already filled out, both the direction of rotation and the degrees to rotate. Just click "Ok."
- Set the margin to be white
  - Open Levels by going to Image > Adjustments > Levels... (keyboard shortcut Ctrl + L).
    - Click the White Point dropper which is the eye dropper furthest to the right in the lower right corner of the pop-up window.
    - Click in the margin above the page you scanned.
    - Click "Ok" in the pop-up window.
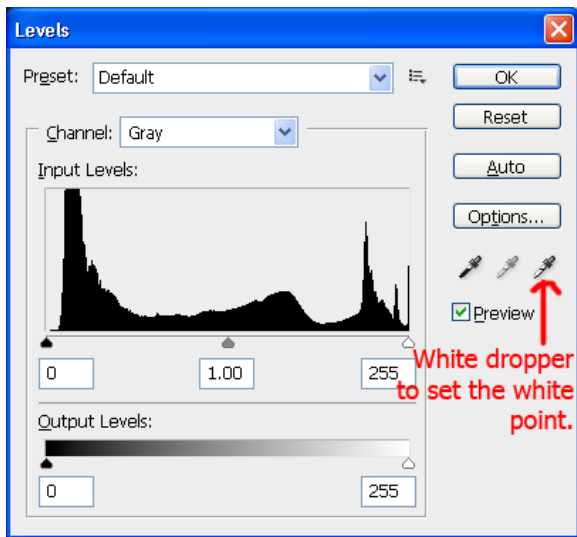
- Screencast of deskewing, re-cropping, and setting the white point: http://screencast.com/t/q98QZ7zgk
- Save the file
  - Go to File > Save As…
  - Enter in the filename and save it as a .tif. The default options for the .tif in Photoshop are what we use.
- Close the file.
- Lather, Rinse, repeat!

Scanning a 4x5 negative

- Open up Adobe Photoshop CS4. Go to File > Import > Epson Perfection V700/V750.
- A new window pops up that is the Epson scan driver. Make sure the scanning software is set to "Professional Mode".
- Place the negative in the 4x5 holder and place the holder on the scanner, make sure the reflective mode cover is NOT blocking the transparency adapter (if the underside of the lid is white then it is installed, glass means it is not installed), close the lid, UNCLICK "Thumbnail" at the bottom of the window below Preview, and click Preview in the scanning software.
- Drag a box of "marching ants" around the object in the preview screen. Add a small and even margin around the outside of the object. There is no perfect size for the margin, but you want it to be obvious the entire object has been scanned without going overboard. Consistency is the key here.
- Set the scanner settings:
  - Original
    - Document Type: Film (with Film Holder)
      - This is the setting you will use with any film that fits into one of Epson's film holders. If you are scanning a negative that does NOT fit in one of the holders (such as a 5x7 or 8x10 negative or glass plate negatives) use the Film (with Film Area Guide) setting.
    - Film Type
      - Choose "Positive Film" for anything that is not a negative, Color Negative Film for any color negative, and last, but not least, B&W Negative Film for B&W negatives.
    - Auto Exposure Type: Photo
  - Destination
    - Image Type:
      - 24-bit color: Color negatives and positives
      - 8-bit grayscale: B&W negatives and positives
    - Resolution
      - In accordance with our specs, we scan negatives and positives to be at least 5000 pixels on the long side. There is a .pdf located at http://www.library.unt.edu/digitalprojects/assets/files/procedures/film_sizes.pdf (there is also a link to it underneath the Photographs (Negatives and Slides) section of our standards page. This .pdf tells you what resolution to scan a particular negative size to meet our 5000 pixel spec.
  - Adjustments
    - Reset any previous tones settings
      - Click the "Reset" button. It is ¾ of the way down the settings page and is all the way to the right.
    - Auto-Levels of the tones.
      - Click the "Auto-Levels" button. It is the first button on the left on the same row as the "Reset" button.
    - Unsharp Mask
      - Click on (checkmark means it is on) and set to the Medium setting in the drop-down box. Every digital capture needs to be sharpened to take it back to the sharpness of the original.
  - Click Scan

- Close Epson Scan. If you do not close Epson Scan you will not be able to do anything in Photoshop.
- Rotate the scan
  - If your image is not correctly oriented you can rotate it by going to Image > Rotate Canvas> and then picking the option to make the image correctly oriented.
- Deskew the scan
  - If the page is crooked then use the Ruler to straighten it. The Ruler tool in Photoshop is found in the toolbar with the Eye Dropper. Click and hold on the eye dropper and when a menu opens choose the Ruler tool.
  - Click on one edge of the page and still holding down the mouse button draw a line along the edge of the page.
  - Go to Image > Rotate Canvas > Arbitrary...
  - A window will pop up with the information already filled out, both the direction of rotation and the degrees to rotate. Just click "Ok."
- Save the file
  - Go to File > Save As…
  - Enter in the filename and save it as a .tif. The default options for the .tif in Photoshop are what we use.
- Close the file.
- Lather, Rinse, repeat!

Digital Preservation

Digital preservation is an active area of research where many major questions remain unanswered, progress however is being made on many fronts and the community as a whole is in a much better place than we were only five years ago. Remember we had centuries to figure out how to preserve paper based documents and well over 150 years to figure out how to take care of photographic images. Some of the technologies and techniques we are looking at today haven't even been around for a decade.

A practical model for digital preservation has been outlined in the makeup of the Federal Depository Library Program and later in a much easier to remember initiative LOCKSS (Lots Of Copies Keep Stuff Safe). The idea of having multiple copies of your master files is a sound, tested and practiced method for making sure you have your saved content in the future.

With todays ever growing volume of digital content, methods suggested as few as five years ago have now been phased out as unscalable and are being surpassed by newer practices. For example most groups don't suggest burning compact disks or even DVDs for your long term storage just because of the management issues, at the DPU we have some projects that would easily span 2,000 DVDs just to make the first copy of the files. The move to spinning disk allows for a greater amount of content to be stored in a smaller amount of space with easier management, however the use of hard disk drives makes it critical to back up all data being stored and maintain at least two copies. With modern prices it is easy to make two and three copies of data on high quality external hard disk drives.

Note: Sales Pitch Beginning

Another option in the area of digital preservation is to work with another institution with a similar mission statement who is interested in helping provide access to and preserving content from various institutions. The Portal to Texas History is one such system which maintains a preservation copy of all digital objects it adds to the system. In the even of a catastrophe it is possible to retrieve these digital objects as a first step in the rebuilding process of valuable collections. I urge you to talk with the Portal staff to get a better idea of how they might help in this area.

Retrieved from "http://digitalprojects.library.unt.edu/projects/index.php/Digitization_101_2011-04-11"

Category: Tutorials