AUSTRALIAN MUSEUM

DIGIVOL -
ARCHIVE HANDLING AND DIGITISING

September 2015

AUSTRALIAN MUSEUM
1 William Street Sydney
NSW 2010 Australia
T 61 2 9320 6000
australianmuseum.net.au

AUSTRALIAN MUSEUM

TABLE OF CONTENTS

1 Introduction 3

1.1 The Australian Museum Archives 3

2 Process and procedures 4

2.1 Health and safety preparation **Error! Bookmark not defined.**

3 Australian Museum Registers, Reports and Magazines 4

3.1 Handling Registers, Reports and Magazines 5

3.2 Assessing Registers, Reports and Magazines for imaging 5

4 Field Diaries and Notebooks 7

4.1 Handling and Assessing Field Diaries and Notebooks for imaging 7

5 Photos and slides 7

5.1 Handling and assessing Photos and Slides for imaging 7

6 Cards and data sheets 8

6.1 Handling and assessing cards and data sheets for imaging 8

6.2 Ethnology Large Cards 19

6.3 Data sheets 24

7 Miscellaneous archival projects 8

8 Workstation setup 8

8.1 Creating a folder 9

8.2 Applications to open on the desktop 9

9 Imaging archival material 14

9.1 Database 18

9.2 Points to consider 26

9.3 Damaged archival material 27

10 Steps to process imaged Archival material 27

11 Image Optimisation 27

11.1 Steps to optimise the images in Photoshop 27

12 Making the PDF 36

12.1 Making the high and low resolution PDFs 36

12.2 Saving documents in PDF/A format 38

13 Converting Transcriptions to Word Documents 40

This document contains information which is provided specifically for training instructions in the DigiVol project at the Australian Museum. This document and content is copyright to the Australian Museum.

Document Author	Version	Release Date	Status	Comments
Leonie Prater	1.0	4 August 2011	Draft	Contributor
Rhiannon Stephens	1.0	4 August 2011	Draft	Contributor
Paul Flemons	1.5	5 August 2011	Final	Review and Approval
Leonie Prater	2.0	August 2015	Final	Update
Rhiannon Stephens	2.0	Sept 2015	Final	Update

Acknowledgements

This document and the processes and procedures it describes have been developed at the Australian Museum (AM) with initial funding assistance provided by the Atlas of Living Australia (ALA) and further funding provided by the AM Foundation and the AM.

This document has been written by the following staff in the AM's Collection Informatics Unit: Paul Flemons (Manager), Leonie Prater (DigiVol Coordinator), Rhiannon Stephens (DigiVol Coordinator) and Michael Elliot (Senior Technical Officer). Special thanks to Vanessa Finney (Archives and Records Manager) for her valuable contribution.

1 Introduction

1.1 The Australian Museum Archives

The Australian Museum Archives holds records created by the Museum which are preserved permanently because of their enduring value for legal, administrative or historical research purposes.

The earliest records include the Minutes of the Committee of Superintendence from June 1836 and Outward Letters books from 1837.

The Archives' collection of historical objects, memorabilia and artefacts complements our paper and photographic records in telling the story of the history and development of the Australian Museum.

The collection comprises material generated within the Museum as well as donations from scientists with a strong connection to the Australian Museum and includes:

- objects used by scientific staff in their research, including scientific instruments and field equipment
- a selection of items from exhibitions that shows how displays have changed over the years
- objects representative of the fabric and infrastructure of the Museum. Most of our heritage furniture collection is still in use today and is part of the NSW S170 Heritage & Conservation Register.
- a selection of items representing how the Museum has been advertised and promoted.

It is important to digitise the AM Archival Records to allow information to be accessed by anyone without further handling or removal of the fragile materials from the Archives.

The archival record images will be available to the AM staff with the aim of some of it being extended to the research communities and the public.

This guide outlines volunteer best practice guidelines and processes for digitising archival material and should be used in conjunction with the archive training video.

2 Health and safety preparation procedures

Before handling archival material, clean hands thoroughly and remove all jewellery. Oils and acids from skin can be transferred to the paper and are potentially harmful to archival material. Determine whether gloves need to be worn when handling the archival material.

Some of the older archival material may have red rot. Red rot is a type of deterioration of the surface of tanned leather occurring when certain tannins degrade to a fine red powder. With some people, the powdery dust may cause mild to moderate skin irritation or minor localised rashes, itching eyes or an allergic reaction similar to hay fever (i.e. sneezing, coughing or runny nose).

In order to minimise any contact with red rot, supplied latex gloves are to be used when handling archival material with red rot present.

It is a requirement to wear covered shoes in the laboratory to avoid personal injury from a falling item or trolley.

3 Australian Museum Registers, Reports and Magazines

Until the 1870s, the specimens in the Museum's collection were individually labelled and most were on public display in the 'cabinets of curiosities'. In 1877, Edward Palmer was employed to compile the first compilation Register of the existing collections. The Palmer Register was followed by 'A' and 'B' Registers, where all acquisitions were collectively catalogued. From 1886, separate specialist registers have been maintained, reflecting the major organism or object groups collected by the Museum.

Registers are important to the Museum's collections as they contain information relating to the specimens held in these collections.

The Museum Archive collection also contains many Museum publications such as the AM annual reports and the AM magazines. The AM magazine was launched in 1921 to satisfy the public's interest in nature and culture with news and feature stories that go behind the scenes.

3.1 Handling Registers, Reports and Magazines

In order to avoid damage when handling a heavy Register, always use two hands to avoid pulling the Register by the spine. Register handling will involve removing the Register from the shelves in archives to the trolley which will be used to transport the registers to the DigiVol laboratory for imaging. At no time, should there be more than two registers placed on top of each other on the trolley.

Australian Museum reports and magazines are historic documents which need to be handled with care and supported with both hands. These Museum archives also need to be transported to the DigiVol laboratory by trolley.

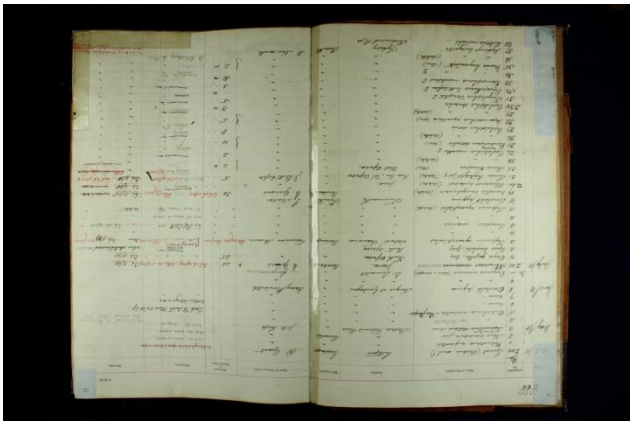
3.2 Assessing Registers, Reports and Magazines for imaging

In the DigiVol Lab, prior to removal of a Register, report or magazine from the archive box, assess if red rot is present and whether gloves are necessary for safe handling. With both hands, remove the Archive item to the cradle and make an assessment with the DigiVol coordinator to determine the following:

- whether the item is fragile and requires extra support to avoid damaging the spine.
- whether the item is numbered.
- whether there are inserts which need to be digitised

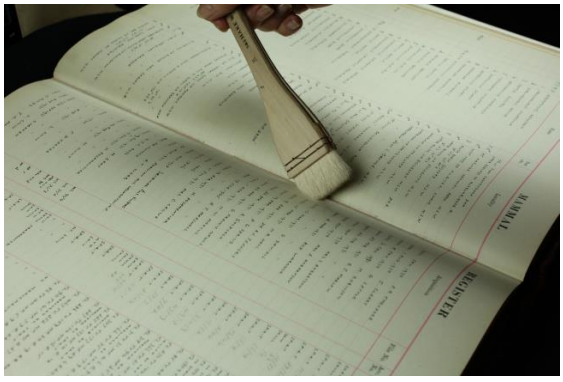
Remove gloves after opening the cover.

3.2.1 Imaging Registers



Place the Register upside down and square on the black cloth which covers the cradle and surrounding workstation area. The image of the black cloth fills the camera frame providing a clearer image by showing a contrast between the edges of a Register and the background. A Register needs to be lying flat in order to take a good image of both pages and avoid shadows which may cause a blurred image. In order to achieve this, assess whether the Register's binding is adequately supported and lying flat on the cradle or whether it will require a cushion to provide extra support to achieve a good image.

When opening the Register or turning a page to be imaged, it is best to support each page by using the flat of your whole hand to turn it from the top left hand corner. As you proceed to turn pages in a Register, the pages may no longer be flat so it will be necessary to place glass weights on the pages to hold them down and in some situations, using finger tips, being careful not to cover any writing.



Use the supplied brush to gently brush any dirt from each page along the inside of the spine joint and away from yourself.

3.2.2 Imaging Reports and Magazines

When imaging reports and magazines, photograph one page at a time. Make sure that the single pages are photographed in the same orientation each time. Place the item square on the black

cloth and try to fill the camera frame with the page with only a small border of black cloth around the edge. The item needs to be lying as flat as possible and make sure that the text on the page is sharp and not blurry. The more effort put in to taking a good image at this stage cuts down the post processing time in the next stage.

4 Field Diaries and Notebooks

The Australian Museum has many field diaries and notebooks of curators, collection managers and research scientists stored safely away in its Archives and Records. DigiVol has made this wealth of information available to many scientists, historians, amateur naturalists and others by digitising and transcribing these valuable documents. Once the diaries have been uploaded onto DigiVol Online website (DigiVol.org) they are transcribed and validated by virtual volunteers. Lists of species observations in these diaries and notebooks are searchable and provide an extraordinary amount of data on fauna and flora which will help us to better understand our biodiversity. Some examples of scientist field diaries and notebooks digitised include M.S.R Sharland, E. Waite, G Holloway and I Thornton.

4.1 Handling and Assessing Field Diaries and Notebooks for imaging

Field diaries and notebooks vary in size, age and fragility. Each diary needs to be assessed to identify any specific handling requirements to avoid any damage to the diary. The pages of the field diaries may need to be numbered, including any additional inserts.

Care needs to be taken to ensure that the pages are adequately supported and lying flat on the cradle to take a good image. An image of each page needs to be taken as a single page and not as a double page.

5 Photos and slides

Apart from photographing scientific specimens, the Museum's early photographers would often capture images of scientists in the field, landscapes, birds, animals and people.

In addition to the Museum's own institutional records, Archives also holds some acquired photographic collections from people who have been associated with the Museum or who have worked in closely related fields.

The Australian Museum's official lantern slide collection began with the opening of our new Halstrom Lecture Theatre in 1910. Through the 1920s to mid 1950s, lantern slides were the Museum's image bank, created, used and re-used for lectures and talks on a huge range of subjects by Museum scientists, educators and visitors, both on site in our lecture theatre and off using portable projectors in community venues in Sydney and regional NSW.

Different Collections throughout the Museum have stored collections of old 35mm and medium format slides. They are an important reference resource for research scientists and staff at the Museum.

5.1 Handling and assessing Photos and Slides for imaging

When taking images of the AM’s very valuable photographic albums, gloves need to be worn when handling each page to avoid damaging the photographs. The photographs may be stored loosely in a box and great care must be taken to ensure that each photographic page is returned to the original order.

As there are special requirements in handling culturally sensitive photographic albums, a volunteer needs to be briefed on these requirements and will be the only person taking images of the culturally sensitive photographs.

Great care must be taken when handling glass lantern slides as they are fragile and very valuable. A damp cloth is used to remove any smudges on the lantern slides prior to taking an image. No cleaning agents are to be used.

Due to the age of some of the 35 mm slides, care must be taken when cleaning the slides to remove dirt, dust and smudges. Special cleaning agents and tools may be used to avoid damaging the slides.

6 Cards and data sheets

Prior to using computers, staff at the Australian Museum used a card index system to store important data about each specimen, cultural object and archival record. This valuable information is being imaged in the DigiVol lab and then imported to the Museum's database to make it easily accessible to Collection staff.

Some Collections have valuable scientific data stored on old style data sheets which are important historic records associated with specimens in the collections.

6.1 Handling and assessing cards and data sheets for imaging

Both index cards and data sheets need to be handled with care. Index cards are to be returned to the tray in the same order. Data sheets may be bound which need special care to ensure that the pages do not tear along the binding.

7 Miscellaneous archival projects

DigiVol volunteers also complete one-off job requests which may vary in content. Some examples have been when volunteers have transcribed data from historic Museum drawing plans into a spreadsheet or transcribed information from Museum Registers to make this information easily accessible to staff. Volunteers are provided with training on the specific requirements of each project.

8 Workstation setup

The workstation is set up with a camera on a copy stand, a cradle and a computer with dual screens.

It is a three step process to capture the image of the Archival item:

1. The Handler prepares the item for imaging.
2. Imaging the item and data entry.
3. Preparing the next page/item for imaging.

At each workstation there will be two volunteers. One of whom will handle and turn the pages of the archival material (the Handler) and the other volunteer will image and database the pages (the Digitiser).

The following steps will guide you in the systematic usage of the equipment (ie computer, camera and database) for digitise archival material.

- Use the generic login to log onto the computer.
- Turn the lights on.
- Turn the camera on and remove the lens cap, place the lens cap in the equipment box.

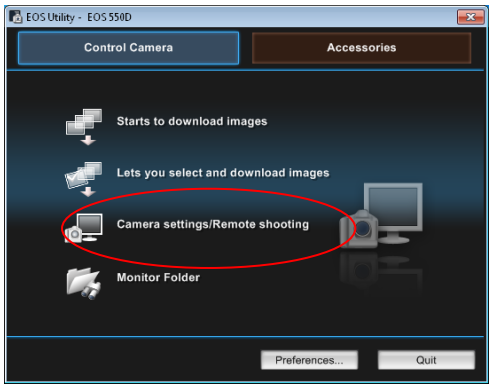
8.1 Creating a folder

Before starting each new archival project, a new folder must be created within the R: under Document Digitising. Consult your supervisor on the naming convention for your project.

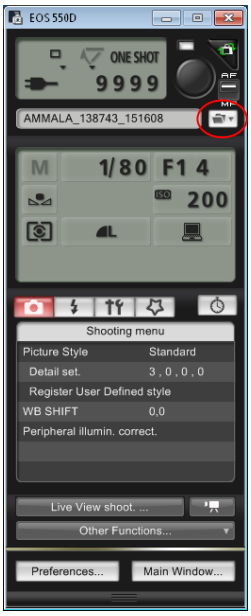
8.2 Applications to open on the desktop

8.2.1 EOS utility

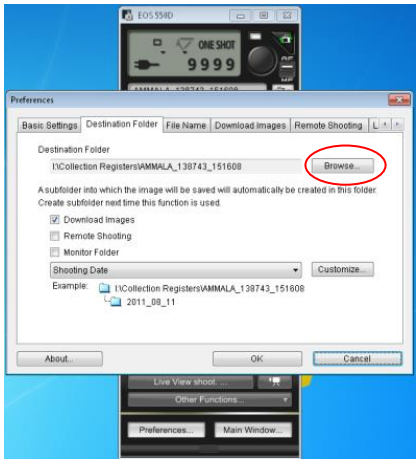
Double click on the EOS Utility icon on the desktop (unless it opened automatically when the camera was turned on), this will open up the below window. Click on Camera settings/Remote shooting.



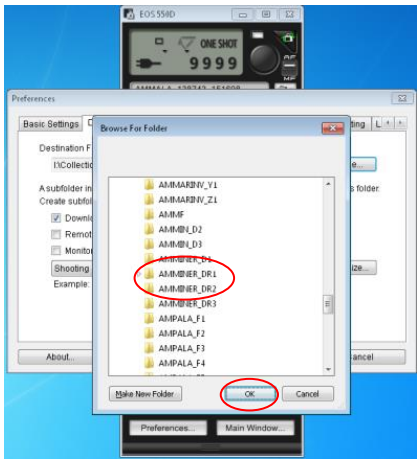
By clicking on Camera settings/Remote shooting, the EOS 550D window will open to the camera's remote shutter control. Click on the folder icon to change the folder that the images are being saved, if needed.



Click on Browse in the Preferences box.

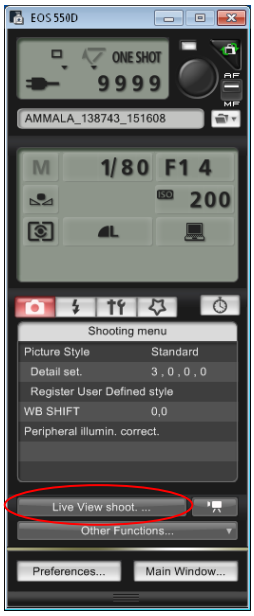


Navigate to and select the folder where the images for the archive project will be saved and press OK

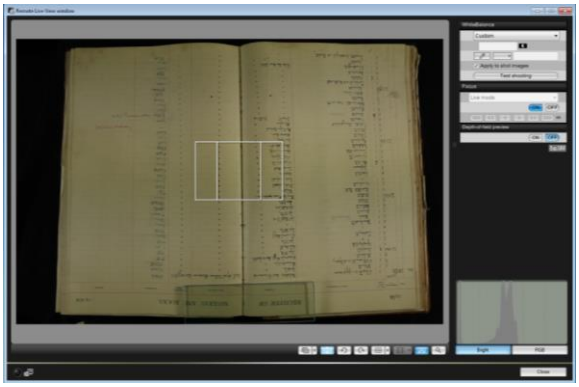


8.2.2 Remote Live View Window

Click on Live View Shoot on EOS utility.



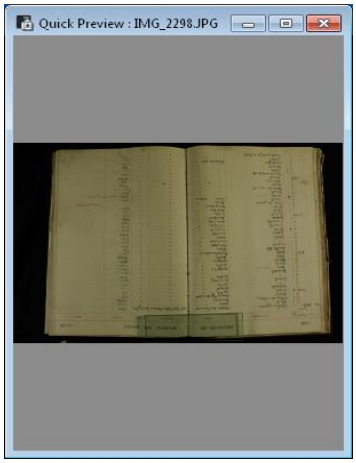
This will open the Remote Live View window.



8.2.3 Quick Preview

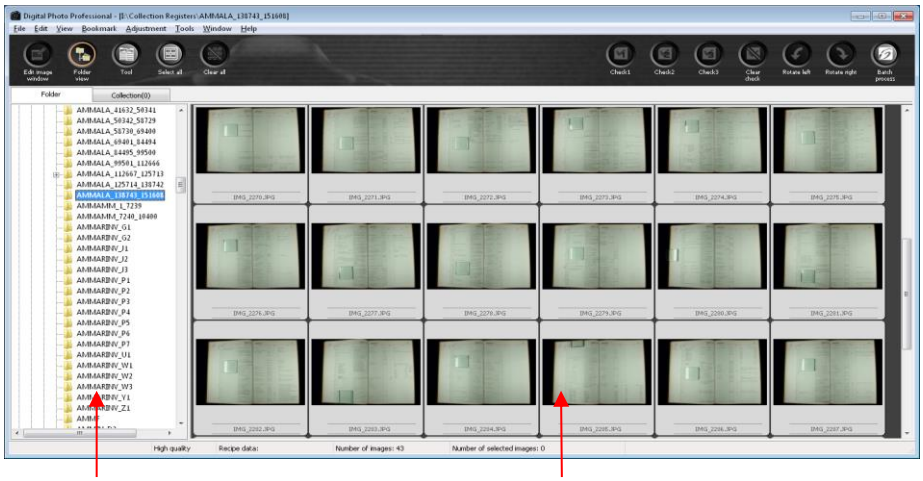
Click on Other Functions on EOS utility and then on Quick Preview.

This will open the Quick Preview window.



8.2.4 Digital Photo Professional

Double click on the Digital Photo Professional icon on the desktop. On the far left hand side navigate to the folder that you will be saving the images to.

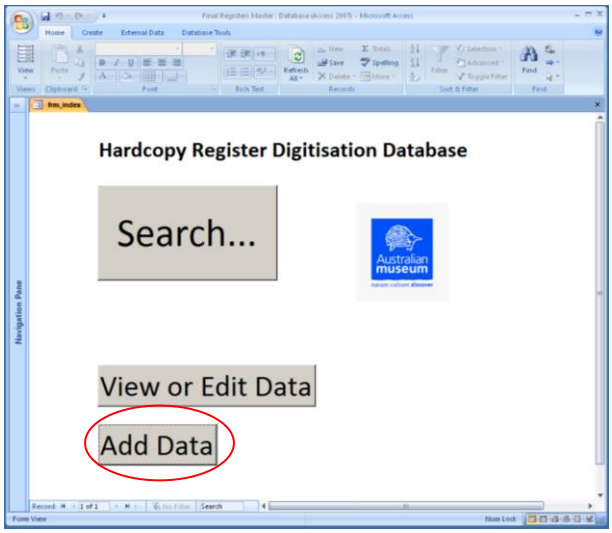


Folder where the image is saved

Image library

8.2.5 Database

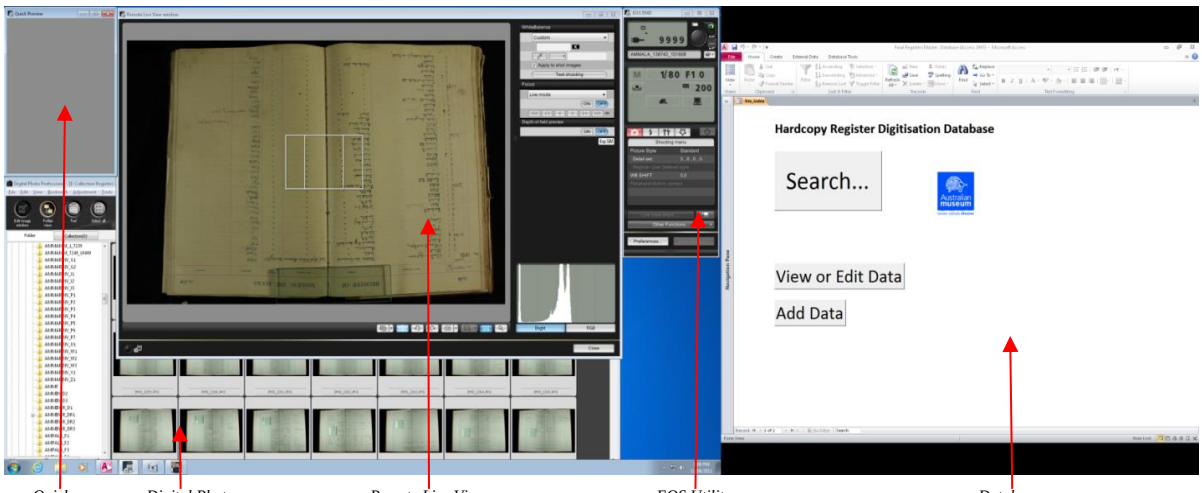
When digitising Registers a small amount of information needs to be recorded about each page of the Register. This information can be recorded in the Register database. Open Final Register Database which can be found on the computer desktop. Click on Add Data.



8.2.6 Position of open windows

The positioning of the open window applications on the dual monitors is a personal preference however guidelines are outlined below.

Monitor 1 (left) Monitor 2 (right)



9 Imaging archival material

The Handler prepares the archival item for imaging making sure that it is centred and flat in the camera view. The first image to be taken will be the item's cover, if it has writing and secondly, an image of the spine's title. The inside pages, preceding the written recorded pages may have writing which needs to be imaged. A microfilmed insert, if included does not need to be imaged. The last page of an item should have the 'Last Page' insert imaged if the item does not have its own recorded last page notation.

Use the camera's remote live view finder, to ensure that the archival material is aligned and in the camera frame and that the black cloth is covering the bookrest and the background.

The camera that we are using in the workstation is a Canon EOS 550D mounted to a copy stand with free standing cold fluorescent lights.

The usual settings for the camera are shutter speed 1/80, aperture F14, ISO200 and custom lighting.

A basic explanation of the camera settings is as follows:

Shutter speed. The time that the shutter is open. These settings are usually measured in seconds or fractions of a second e.g. 1/30, 1/60 and so on. The smaller the fraction the faster the speed (e.g. 1/1000 is much faster than 1/30). The greater the seconds or fraction of a second setting the longer the shutter is open. This can create problems with camera shake due to movement of the copy stand for long exposure times.

Aperture. The aperture is the size of the iris in the lens that also controls the amount of light that the lens lets into the camera and the depth of field (i.e. how much of the image is in focus). The settings for aperture are called f-stops and are usually a standard set of numbers f/2.8, f/5.6, f/8,

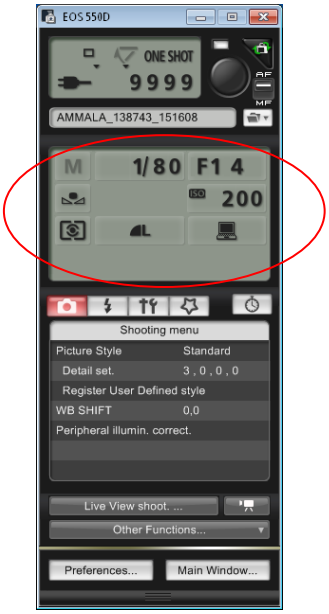
f/11, f/16, f/22, f/32. These numbers are actually fractions (where the number 1 has been replaced by the letter f), but are usually expressed as full numbers (e.g. f 1/16 is usually expressed as “f sixteen”).

Each step up in aperture value halves the amount of light reaching the sensor. Each step down in aperture value doubles it.

Larger apertures (where lots of light gets through) are given smaller f-stop numbers and have shallower depth of field. Smaller apertures (where less light gets through) have larger f-stop numbers and a greater depth of field.

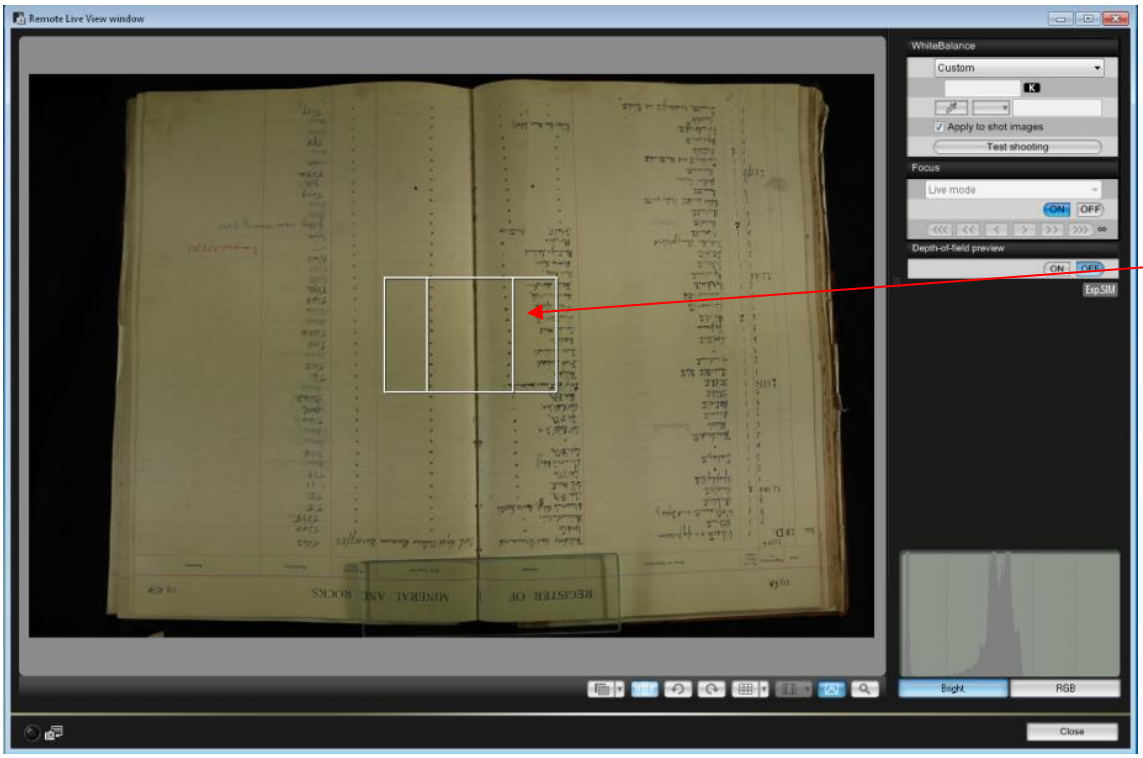
ISO Settings. ISO is the sensitivity to light of the imaging chip in the camera. The lower the number the less sensitive the camera is to light, and the lower the electronic noise.. Higher ISO are generally used in darker situations to allow for a faster shutter speed (like in the Dinosaur exhibition). However, the cost is noisier shots. ISO 200 is sufficient for adequate image quality for specimen data capture.

These parameters can be changed in the EOS 550D window.



In Remote Live View window the point of focus is used to ensure that the Register pages are sharp.

Click on Focus>Live Mode>On to focus the image.



The computer operated shutter control on EOS Utility is used to take the final image.

Shutter control

Take the photo and maximise the Quick Preview window to check the positioning of the item and to make sure the background is covered by the black cloth. Double clicking on the image in Digital Photo Professional can also be used to zoom in on an image to check the focus and clarity.

If the image is not satisfactory, go into Digital Photo Professional and click on the image so that a yellow outline appears, press the delete button and yes to delete the image. Care must be taken that only one image is selected for deletion and if adding information to the database, the image number correlates to that entered in the database.

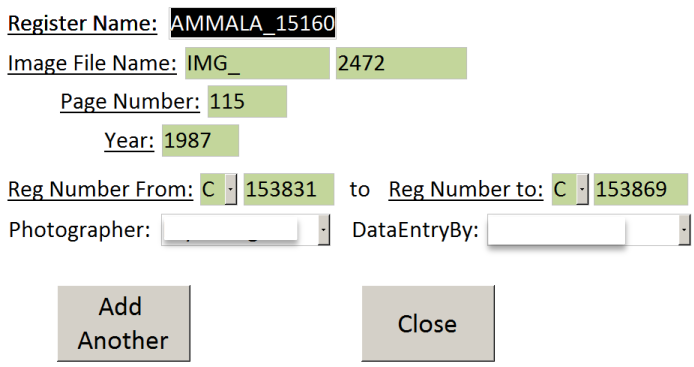
The Digitiser images each page of the Register and checks for a good quality image and writing image clarity.

If a whole archival item has not been imaged during that day, then place the ‘To be Imaged’ insert on the page to be imaged next time. Ensure that all inserts have been removed prior to replacing the item into the archive box and returning to the DigiVol coordinator. Put gloves back on if there is red rot and place on the trolley for transport to archives.

9.1 Database

When digitising Registers the database must be completed for each page. The Digitiser enters the Register page information into the database for each image and the Register Handler provides a quality check on data entered. Check the data in each field on the database matches those on the Register image screen. If you find it difficult to read the information on the Register, you can use the zoom tools in Digital Photo Professional by double clicking on the image. If the information is correct the next page is ready to be imaged by clicking the 'Add Another' button on the database. If any information is incorrect, then overwrite it with the correct information and prepare the next page for imaging. Once the Digitiser is satisfied with the final image the Register Handler can prepare the next page for imaging.

The fields of the database need to be filled to correspond with that on the Register.



Register name: Name of the folder on I: where the images are stored

Image File Name: Camera file name of image i.e. IMG and number in the next field. This can be found below the image in Digital Photo Professional.

Page Number: Page number of the Register (written in pencil at the top of the page).

Year: Register year (if there are multiples on the same page – always use the oldest date e.g. 1977, 1978 – use 1977).

Reg Number From: Prefix in drop down box and 1st register number at the top of page.

Reg Number to: Prefix in drop down box and last Register number.

Photographer: Digitiser's name

DataEntry By: Digitiser's name

Add Another: to save and go to the next record

Close: Save and close at the end of the day.

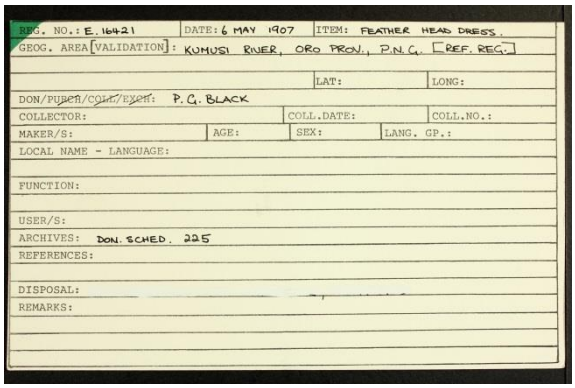
9.2 Imaging Ethnology Large Cards

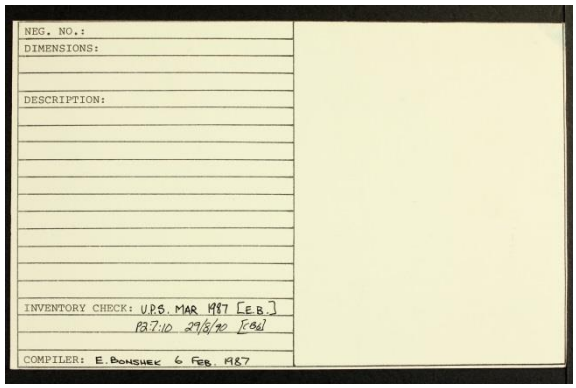
Put the card upside down in the frame on the copy stand.

Use the below settings on the camera with the camera at height 32 cm from the bottom of the camera.

Take the image of the front of the card. Turn the card over and image the back.

You now have 2 images in the **Large** cards folder of both the front and the back of the card.

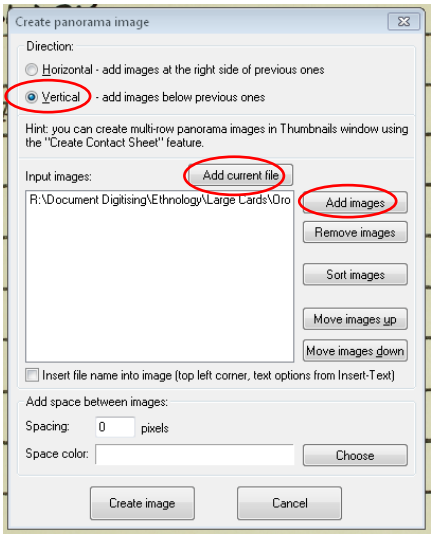




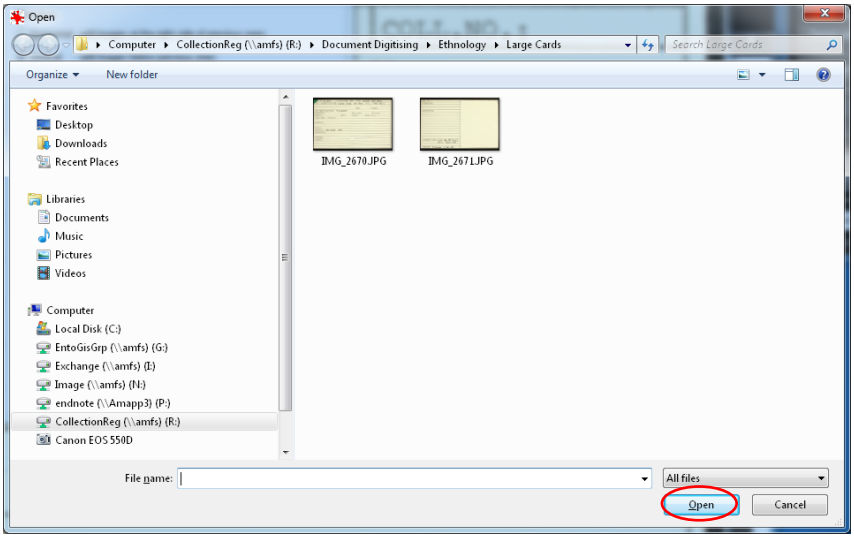
9.2.1 Stitching the images together.

Open up the first image i.e. the front of the card, in Irfanview. You need to open the image from Windows Explorer. It will not work from Digital Photo Professional.

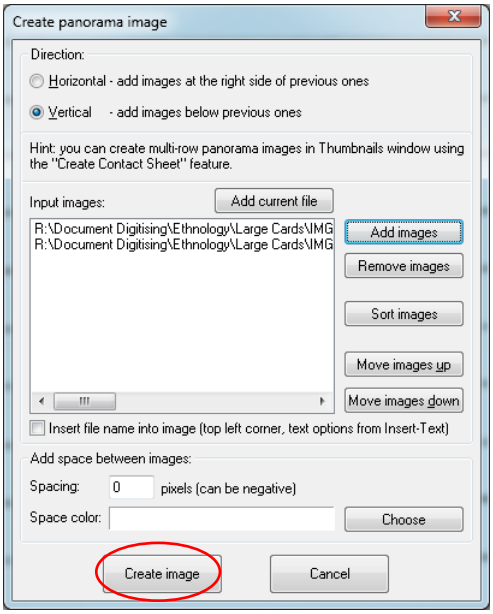
Click on Image>Create Panorama image….>



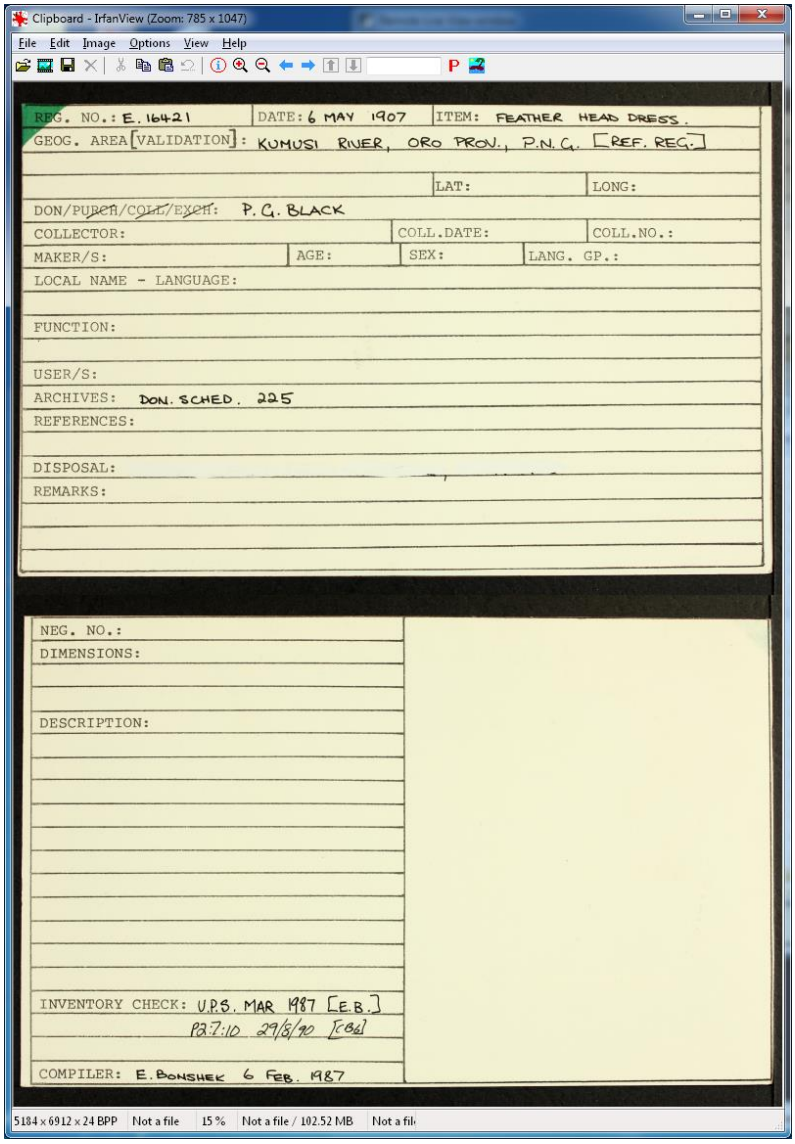
Make sure **Vertical** has been chosen. Then **Add current file**. Then click on **Add images** and find the second photo i.e. the back of the card in the directory (should always be the last photo).



Choose the photo and click on open.



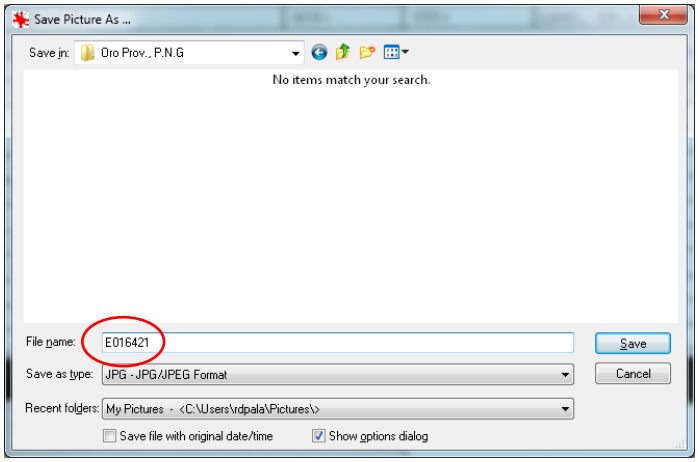
Click on **Create image**



Click on File>Save as

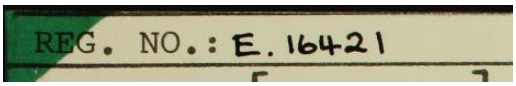
Save the picture in CollectionReg R:>Document Digitising>Ethnology>ETHNO_DRW_1>Large Cards

*Make new folders for each new drawer of cards.



9.2.2 Naming Conventions of Ethnology Cards

Use the Reg. No. on the top left front of the card for the file name. The number after the E must be six digits long. If it is not, add 0s to the front of the number to make it up to the 6 digits. Do not add a full stop after the E. For Example



will become dE016421+01

- If it is a single number with no subsequent cards with the same number it is given a +01 suffix. Thus the image name for a card with E012345 would be E012345+01.jpg
- If the card and subsequent cards have the same number, the suffix advances by 1 for each card. Thus if E012345 is written on 3 individual cards the images for the cards become E012345+01.jpg, E012345+02.jpg and E012345+03.jpg
- If the card number has a dash in it the number after the dash has zeros prefixed to it to become a 3 digit number and the usual suffix is then appended as well. Thus a card with E012345-1 becomes E012345-001+01.jpg and E012345-12 would become E012345-012+01.jpg.
- Rule 2 applies to subsequent cards with the same dash number as they are separate entries in Emu. Thus 2 cards with E012345-1 become E012345-001+01.jpg and E012345-001+02.jpg
- Subsequent cards with advancing dash numbers do not get advancing suffix numbers. Thus E012345-1 and E012345-2 do not become E012345-001+01.jpg and E012345-002+02.jpg. They become E012345-001+01.jpg and E012345-002+01.jpg as they are different reg numbers.

Click on Save.

9.2.3 Record in spreadsheet

Fill in the spreadsheet with the date, folder and file name. Add the photographer.

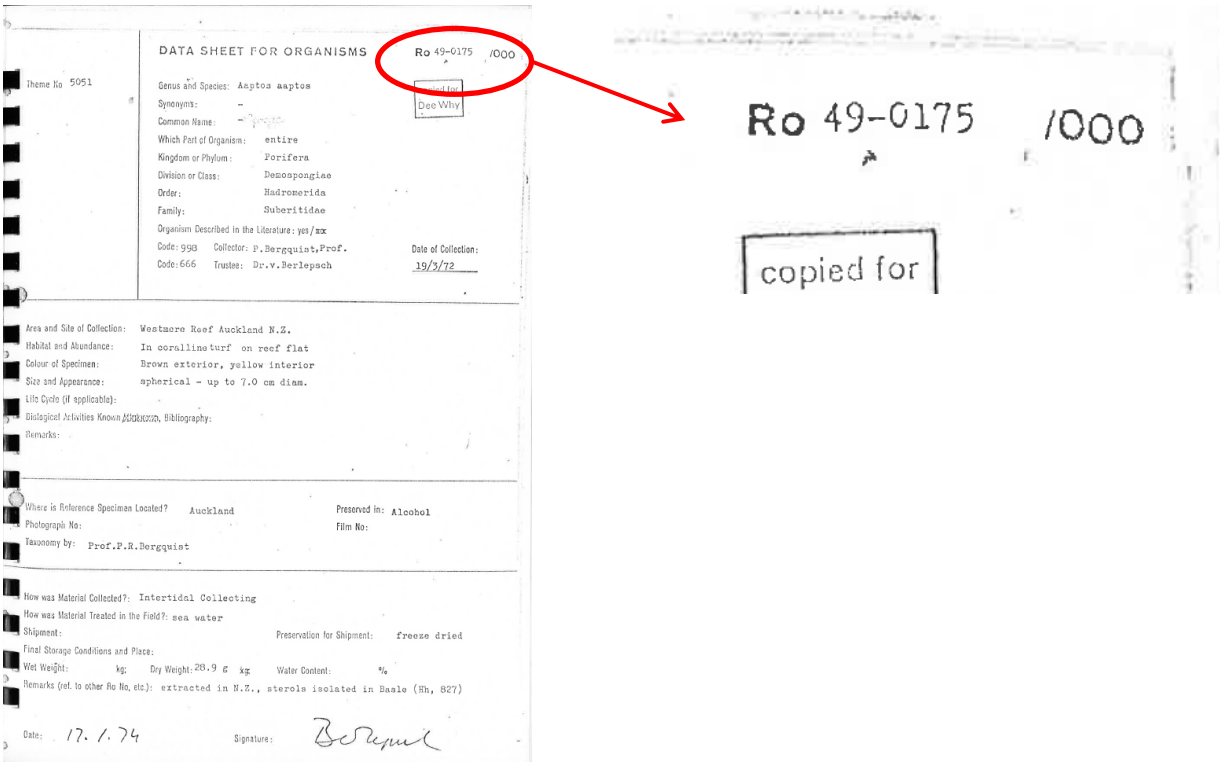
You can then close the image ready for the next card.

Just before the end of the day please check all your images to make sure their filename matches what is on the card. Place a marker after the last card you have finished for the day.

9.3 Data sheets

Each collection will have their own specific naming conventions for entering information into a spreadsheet.

Image the data sheet as a jpeg. Fill up as much of the camera frame as possible, with neat borders and make sure all information is captured. This material will be OCR'd so try and keep as flat as possible.



DATA SHEET FOR ORGANISMS Ro 49-0175 /000

Theme No 5051

Genus and Species: Aaptos aaptos
Synonyms: –
Common Name: –
Which Part of Organism: entire
Kingdom or Phylum: Porifera
Division or Class: Demospongiae
Order: Hadromerida
Family: Suberitidae
Organism Described in the Literature: yes/no
Code: 999 Collector: P.Bergquist,Prof.
Code: 666 Trustee: Dr.v.Berlepsch
Date of Collection: 19/3/72

Area and Site of Collection: Westmere Reef Auckland N.Z.
Habitat and Abundance: In corralline turf on reef flat
Colour of Specimen: Brown exterior, yellow interior
Size and Appearance: spherical - up to 7.0 cm diam.
Life Cycle (if applicable):
Biological Activities Known/Unknown, Bibliography:
Remarks:

Where is Reference Specimen Located? Auckland
Preserved in: Alcohol
Photograph No:
Film No:
Taxonomy by: Prof.P.R.Bergquist

How was Material Collected?: Intertidal Collecting
How was Material Treated in the Field?: sea water
Shipment:
Preservation for Shipment: freeze dried
Final Storage Conditions and Place:
Wet Weight: kg; Dry Weight: 28.9 g kg Water Content: %
Remarks (ref. to other Ro No, etc.): extracted in N.Z., sterols isolated in Basle (Hh, 827)

Date: 17. 1. 74 Signature:

Ro 49-0175 /000

copied for

Rename the image using the RO number at the top right hand of the sheet. In the example attached this would be Ro49-0175. Some of the early sheets use a "/" instead of a "-" but this will cause problems with file paths and the / is to be avoided.

Title	Image	Creator	Description
MulTitle	Multimedia	MulCreator_tab(1)	MulDescriptio
Ro 49-0175	Roche_Data_Sheet_0000_0149	Photographer	Original data sl from Roche Re Institute
			Original data sl from Roche Re Institute
			Original data sl from Roche Re Institute
			Original data sl from Roche Re

Record each file onto the spreadsheet. Fill out the first three columns, title (the Ro number), image folder, and photographer.

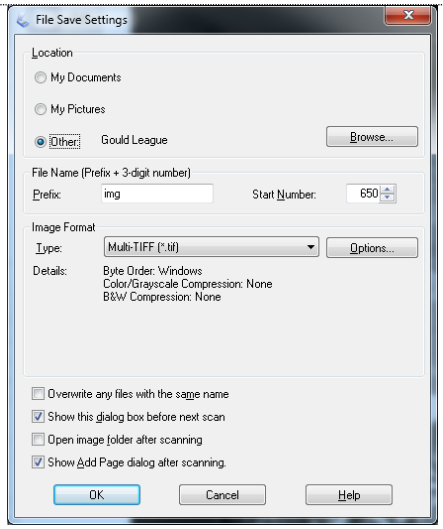
9.4 Slides

The table below outlines the process to scan slides.

1. clean slide with microfiber cloth provided 2. place slide in holder	• If there is writing on the slide, this goes face down • Often there is a paper mount in the slide and the 'glued' side faces up • One side is usually a little; clearer, this side faces down
3. Select **Preview** button at bottom of EPSON Scan	This scans the slide "**Preview scan in progress**" box appears. Wait. When finished, the image will be a small area on screen in the Preview pane
4. Click and drag on the image to define the marquee (the moving dotted line around a rectangle)	If the image of the slide doesn't include the entire slide, try rotating it in the holder or moving it on the platen to get the scanner to guess again about where the side is this time. If you moved the slide, select **Preview** again
5. Select **Zoom** button at top right of Preview pane	This will rescan the slide and zoom in on the smaller area defined by the marquee
6. Adjust the marquee, if required 7. Use the **Rotate** buttons, if required	If you notice the slide is skew to the scan, at this point you can adjust it, select **Preview** and start again
8. Select **Scan** button at bottom of EPSON Scan	

9. Check the details such as **Start Number** in the **File Save Settings** form

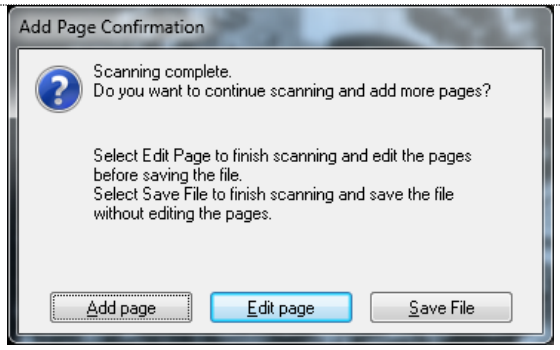
10. OK



The slide is scanned again.

This may take some time

11. select **Save File** at the **Add Page Confirmation** form
12. the image appears in the photo folder
    Initially as a placeholder icon, eventually as a thumbnail of the photo
13. start again with next slide



9.5 Points to consider

When assessing and imaging Registers consider the following:

- Assess the Register to see whether it has been numbered. If not, use the provided pencil to lightly number, in small print, the top right hand pages starting with number 1, 3, 5 etc. Every 15 pages check for numbering accuracy.
- Proceed to take an image of the cover if it has writing on it and take an image of the writing on the spine.
- Take an image of the inside of the front cover if it has writing on it. Fill the database in to be page 0, the 2nd page inside the front cover would be page 0.1 and so on.
- If there is no writing inside the front cover, go to page 1.
- At the end of the Register take a photo of the inside back cover with the insert 'Last page of register' and the register number in the database will be Reg number from 0 to Reg number to 0.

- If there is an extra piece of paper stuck to a page then 2 images need to be taken to capture the information underneath. In the database the actual page will be the page number and the extra information on that page will be the same number plus 0.1 e.g. page 23, page 23.1

At the end of the day, close everything on the computer desktop and shut the computer down. Turn off the camera and turn off the flash. Make sure the workstation is neat.

9.6 Damaged archival material

If any part of the archival material is broken or torn when imaging, immediately consult your DigiVol Coordinator. Any pieces must be preserved and appropriately handled.

10 Steps to process imaged Archival material

After the images of a book/journal etc. have been taken the steps to process the images are as follows:

- Optimise the image by cropping, contrasting, transforming etc.
- Make a high resolution PDF-A version
- Make a low resolution PDF-A version

11 Image Optimisation

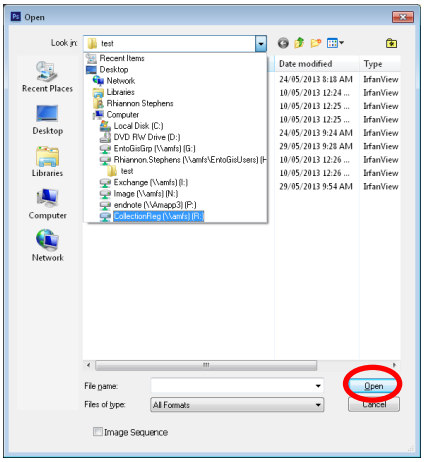
Make a copy of the original jpeg (not RAW) images into a new subfolder to keep the originals available and separate from the 'final' files. Open Adobe Photoshop CS6 whose icon can be found on the desktop.

11.1 Steps to optimise the images in Photoshop

1. Brighten up the image by Auto Contrasting. Do not use Auto Colour to preserve the "antique look" for the images.
2. Crop the images to remove borders and edges/binding of other pages that can be seen. Adobe Acrobat will not apply an OCR layer on a document if the page size is greater than 45 inches.
3. Use the Transform tools (Align, Perspective, and Warp) as needed to attempt to keep pages and text as straight as possible.
4. Use the Clone Tool or Healing Brush to fill in any gaps at the edges
5. Save the image and close it.

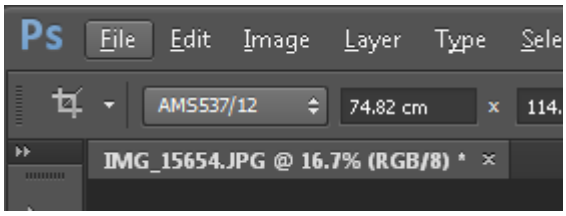
11.1.1 Contrast

Open the image you want to working on in Photoshop by clicking File>Open.



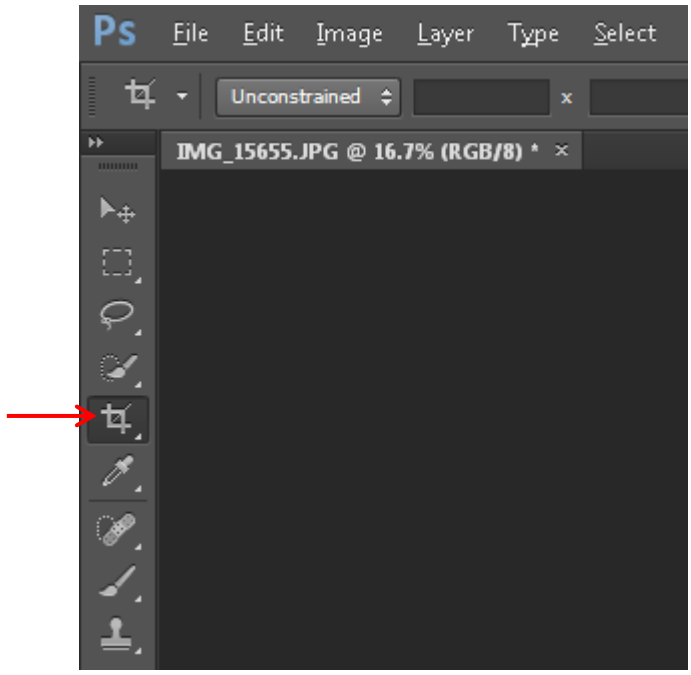
Find the file for the image that you want to work on and click open.

Brighten up the image by Auto Contrasting. Do not use Auto Colour to preserve the “antique look” for the images.



Click on Image>Auto Contrast OR Alt+Shift+Ctrl+L

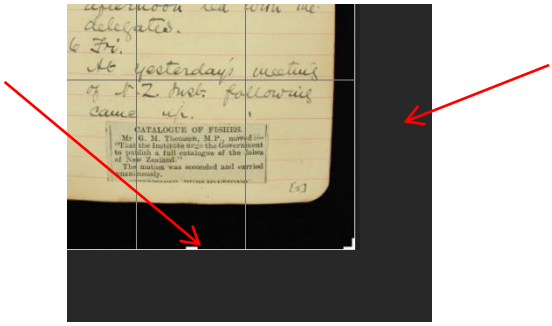
11.1.2 Cropping



Click on the crop tool in the left hand tool bar.

We want to crop all images in the same book to be the same size. At the beginning of each book the ratio of the cropping may need to be changed to comply with that set of images. To do this you need to start by making a crop preset.

Note that in Photoshop CS6 if the image you are working on has a solid border,



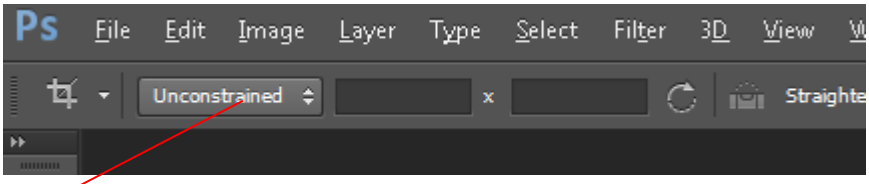
you need to either:

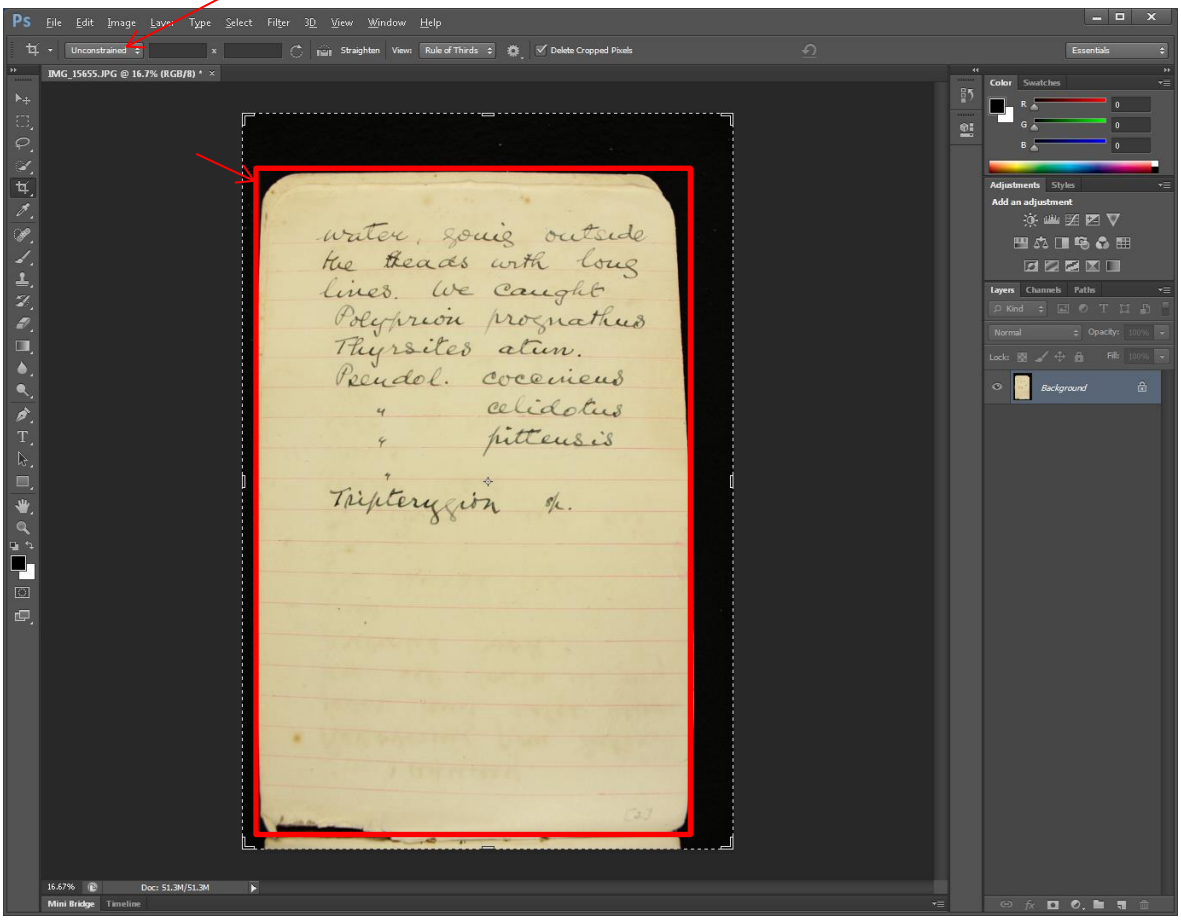
Reset the crop box

Cancel the current crop operation

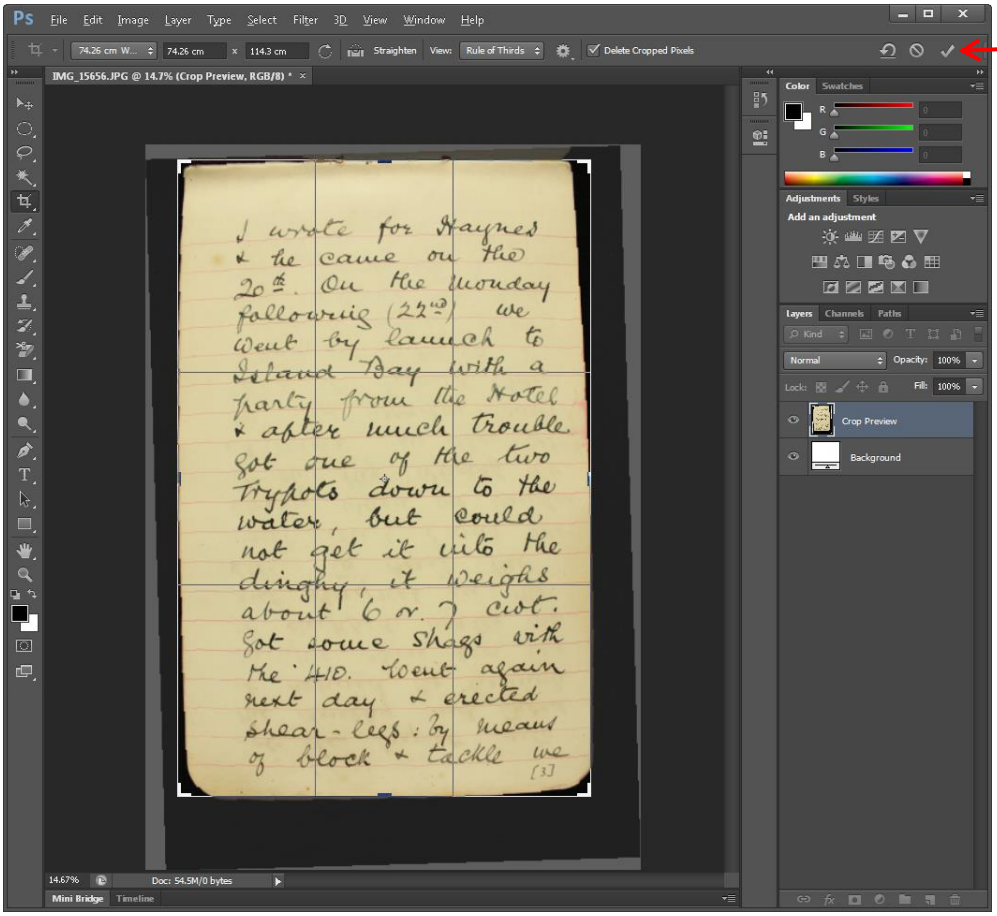
Commit the current crop operation.

Setting a crop preset





To set a crop preset select the crop tool, make sure the top tool bar is displaying ‘Unconstrained’ and drag the cursor around the page. Make sure the page fits neatly into the cropped area and is well centred. Use the arrows to move the page around within the cropped area.

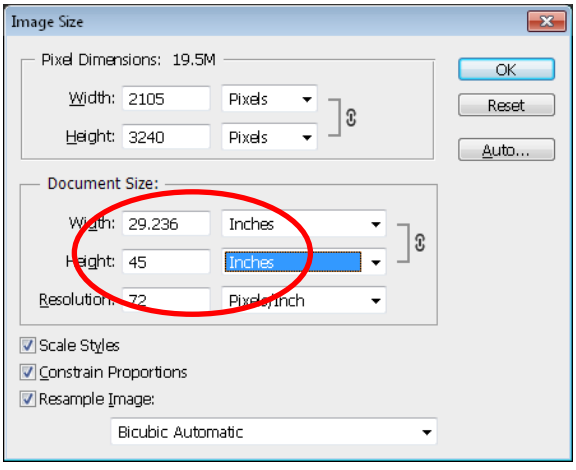


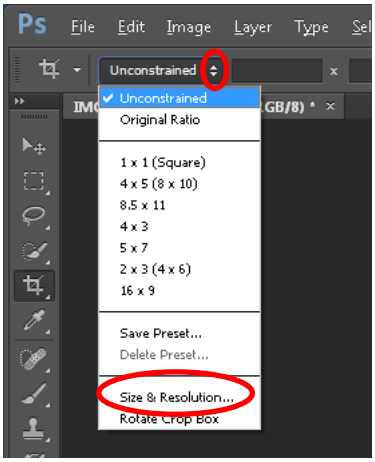
Or use the mouse key outside of the image to tilt and angle the image inside the crop square.

This selected area will be the basis of the cropping page size for the rest of the document.

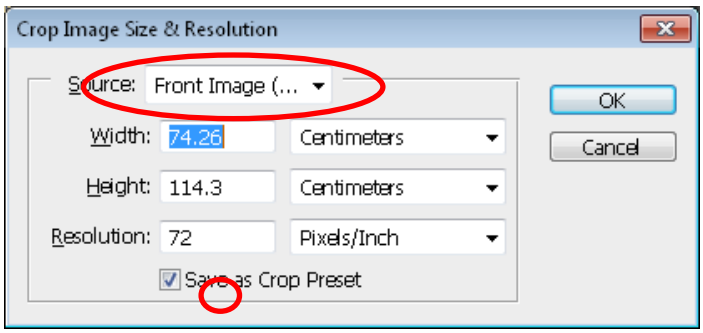
Once you have selected the area click enter or to accept the crop

Go to Image>Image Size and change the largest number between the width and the height to be 45 inches and make sure the Resolution is set at 72 Pixels/Inch and click OK

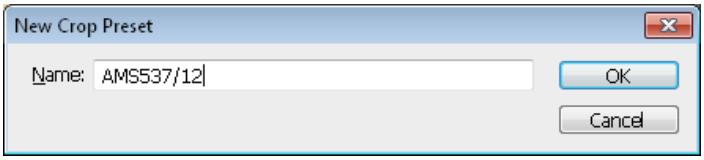




Click on size and resolution within the ratio tool bar



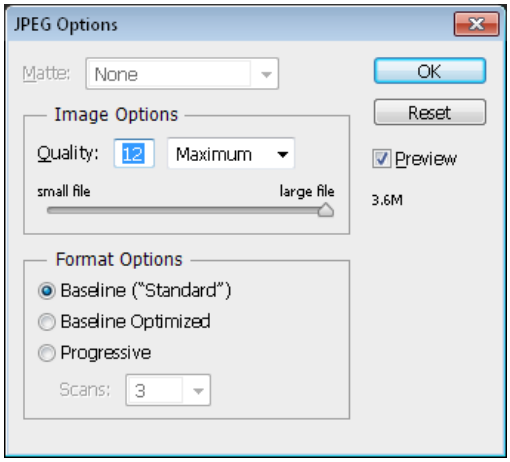
Choose in ‘Source’ the ‘Front Image’ and make sure ‘Save as Crop Preset’ is ticked and click OK.



Change the name of New Crop Preset to reflect the name of the book that you are working on and click OK.

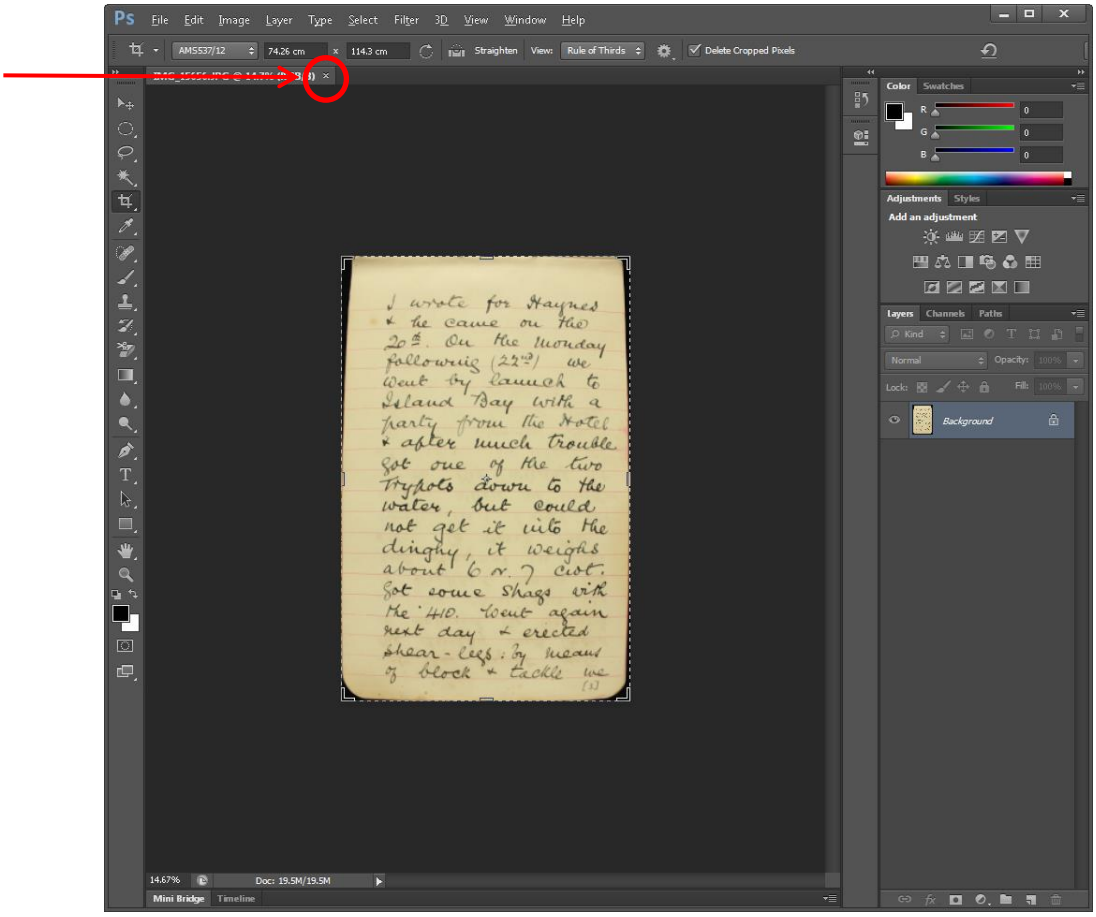
Ctrl S to save the image.

It may ask to Crop the image before saving. Click on ‘Crop’.



Change the 'Image Options' Quality to 12 and click OK.

If the image does not need any other manipulation then this image is now complete.

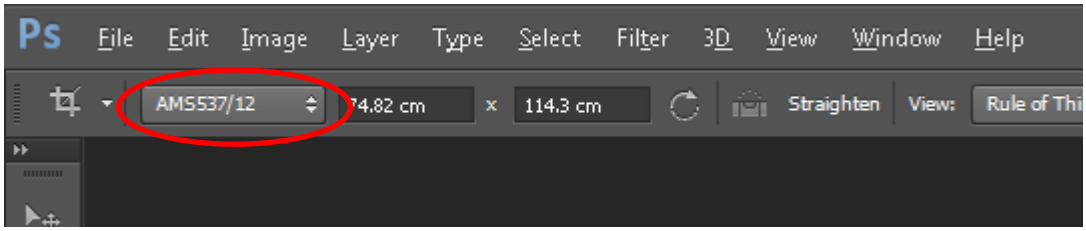


The image can be closed using the 'x' ready for the next image.

Cropping all images

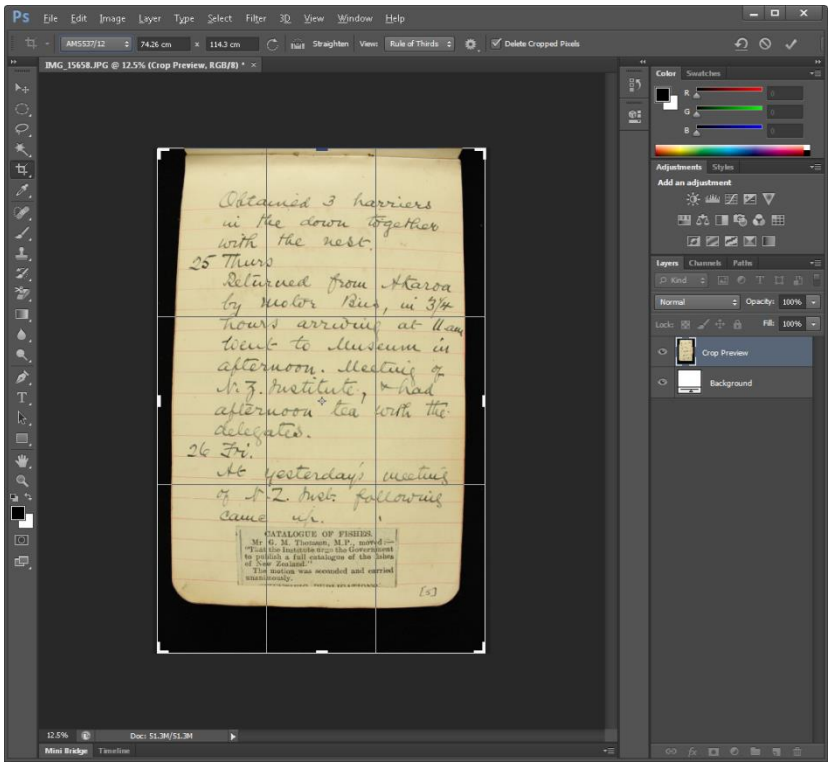
Now that the crop preset has been set it can be used to crop the rest of the images to the same size.

Open the next image (File>Open). Note that several images may be opened at the one time.

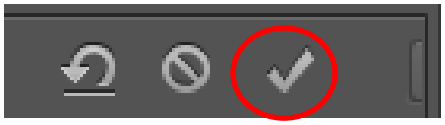


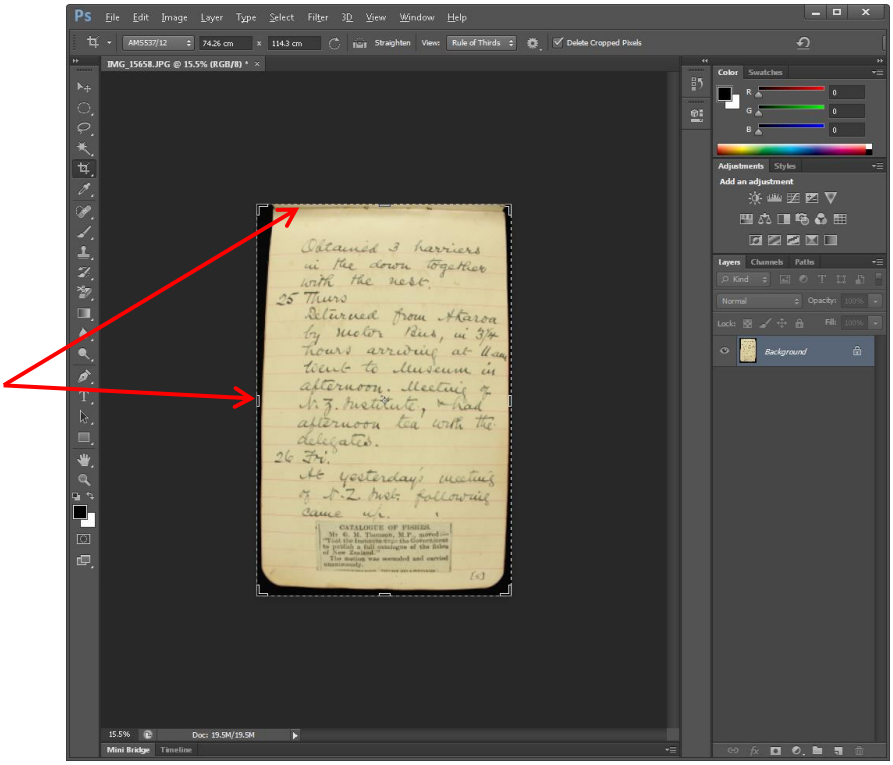
Image>Auto Contrast

Make sure the crop preset that you set previously can be seen at the top.



If the border is a solid border (as above) it means that crop is activated. Click Esc or



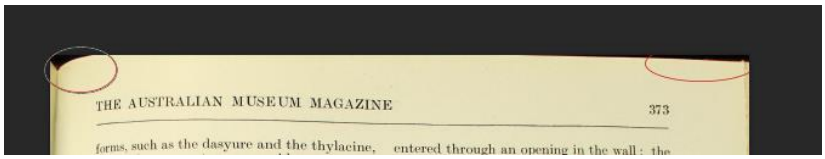


Drag the cursor around the page or move the edges in, align and then click enter to accept and Ctrl S to save.

11.1.3 Transform

Use the Transform tools (Align, Perspective, and Warp) as needed to attempt to keep pages and text as straight as possible. This is especially used for pages that will have OCR applied. It helps to straighten the writing up so OCR works better.

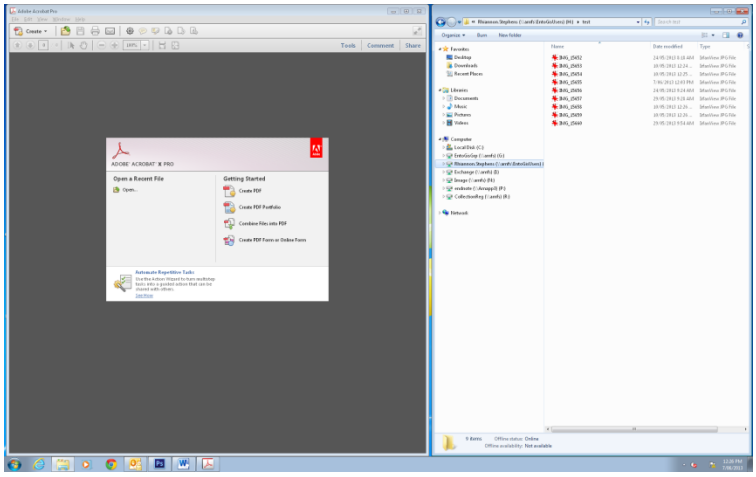
11.1.4 Clone



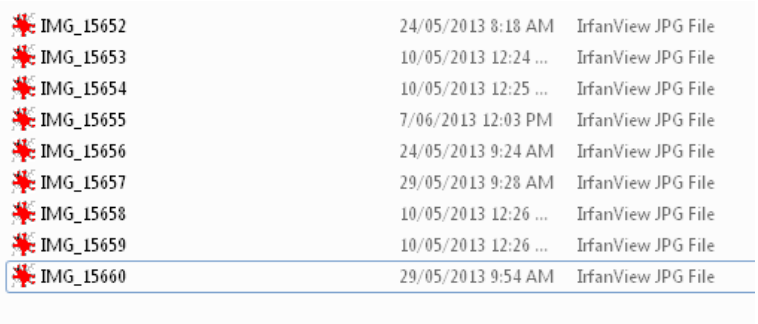
Use the Clone Tool or Healing Brush to fill in any gaps at the edges. Be careful when using these tools that no information is lost or 'rubbed' out.

12 Making the PDF

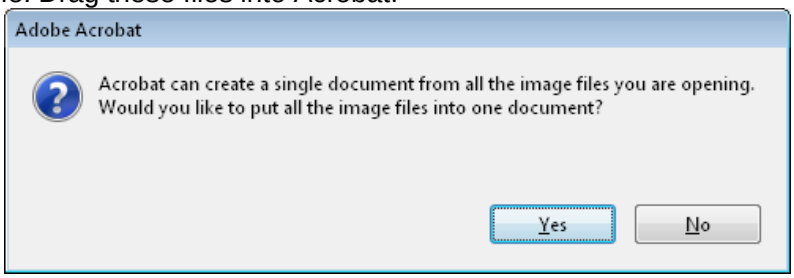
12.1 Making the high and low resolution PDFs



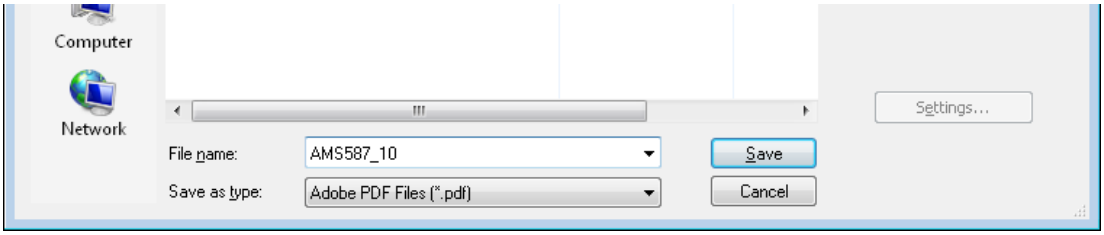
Open Adobe Acrobat X Pro. Keep the window minimised (not full screen). Keep the Windows Explorer folder handy to be able to view both the Acrobat and the folders of images.



Select the images in sequential order by clicking on the first file and holding down shift and clicking the last file. Drag these files into Acrobat.



A message box will appear. Click on Yes.



File>Save As>PDF. The PDF should be called by the AMS number and the volume number. (eg AMS587_10.pdf). Click on Save and save in the Final folder.

If the document has printed text an OCR layer needs to be applied.

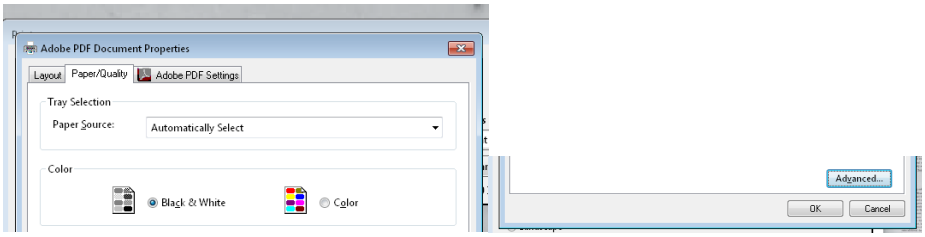
Click on View>Tools>Recognize Text

OCR resolution as 300dpi.

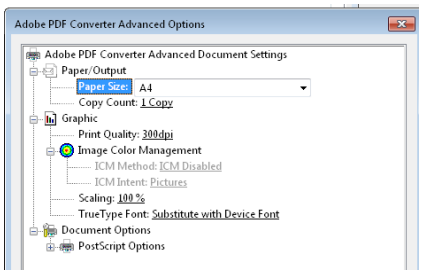
Save the file. This will be the high resolution pdf.

To make the low resolution version – File>Print. Select Adobe Acrobat as the printer and open the Properties.

Paper/Quality = Black & White.



Advanced tab>Graphic = Print Quality = 300dpi.



Click OK.

Click OK to print. Save file as e.g. AMS368_10_lowres.pdf.

Ensure the file is being saved in the correct year's subfolder. Acrobat will default to last used subfolder as the destination path.

Open the "low res" file.

View >Tools >Recognize Text to apply the OCR layer.

OCR resolution as 300dpi.

Save and close to complete the low resolution pdf.

12.2 Saving documents in PDF/A format

12.2.1 Background

This is the process for saving documents to PDF/A-1b preservation standard.

There are two ways to approach the job: in the course of stitching the individual page scans into high-resolution PDFs, the PDF/As could be created as an additional step. Or you could batch-process the job by creating a Preflight Droplet from the Options menu in the Preflight dialogue box and check the specification. Files can then be dragged onto the Droplet.

12.2.2 Saving to PDF/A

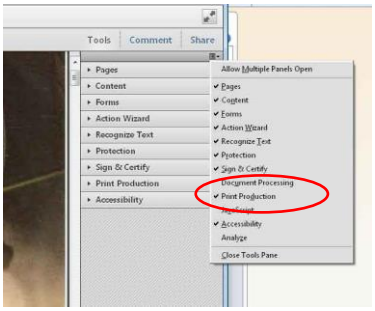
For manually saving to PDF/A 1-b, using Adobe Acrobat Pro X to save a high-res PDF to PDF/A-1b format (with no additional tagging):

1. Open the PDF.

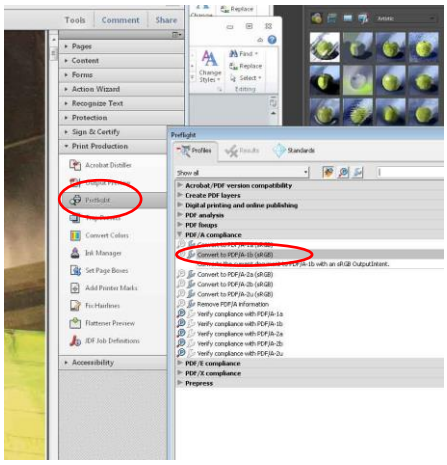
2. If you have been instructed to add metadata: File > Properties > Description (tab)

   Select a language > Advanced > Language dropdown list > select English. Click OK button.

3. To apply PDF/A process: far right of the screen click Tools menu > select Print Production button from list below. NB: If Print Production does not appear, click small icon in corner. A pop-up window will appear. Tick Print Production so that it appears.



4. Click Preflight > Convert to PDF/A1b > Analyze and fix button (bottom right of window)



4. You will be prompted to save to a folder, and then the processing will occur. As a general rule, allow one to two minutes for every 10MB.

5. Re-open the file and check pages in Adobe Reader (to simulate how someone may view the file). If you have correctly saved the file as a PDF/A, you will see a blue bar at the top of Adobe Reader. NB: this will not be the case with other readers.

13 Converting Transcriptions to Word Documents

We need a readable version of the CSV file which will also be uploaded into EMu to be used as a quick reference file.

Once it has been created as a Word document, it needs to be saved as a PDF/A.

The formatting is a quick cleanup of the CSV file for this purpose.

1. Once the CSV file has been downloaded, in Excel select and delete the left hand columns for taskID, externalIdentifier, recordIdx. Also remove the top line of the next column occurrenceRemarks
2. You should now just see one long line for each page. Select all the text, cut and paste into the Word document template <you can create this from the Word example I've attached. I've intentionally made the text style for Normal and Heading 1 text in styles Helvetica as this is a recognised Base-14 font used in archiving>
3. Paste option is Keep text only in Word
4. Check that Paragraph spacing is 1.5; Spacing Before and After is 0pt
5. Find and replace **\n** with a space
6. Find and replace **digit + tab x 3 + para mark** a few times. It looks like this in the Find What field: ^#^t^t^t^p Click tick box for Match case
7. Find and replace an elipses + a colon with a single space. It looks like this in the Find What field: ...:
8. Strange characters which appear with apostrophes or quotation marks need to be removed. They look like this Â ™ € ˜ œ

9. Find and replace double spaces with one space. Do this several times if required.
10. Add the title in the header from the EMu item (assuming it has been created)
11. Save the Word doc as the EMu series number (and add any other sensible identifying)
12. Once the Word doc has been completed, save it as a PDF/A-1b.