

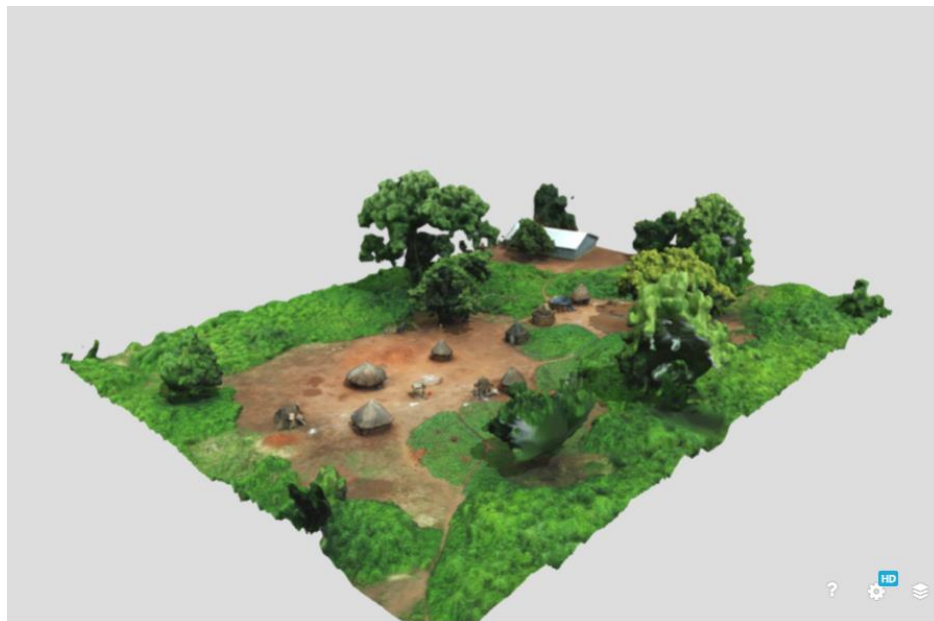
# Other types of Data

Wednesday Morning

MJC-BSANA-Data-Workshop-2024

# Session Outline:

1. File Management with command line
2. Multidimensional Media
3. Digital Images Annotation
4. Digital Video Annotation
5. Photogrammetry



3D model of an Opo homestead showing the open space occupied by the different huts and activity areas that make up the domestic space. EMKP grantee Alfredo Gonzalez Ruibal. Rights owner: Incipit-CSIC License: CC-BY-NC-SA 4.0

# Gathering and managing files

When we decide to extract material from the web for our own research, education etc, we need to make sure that we are collecting and managing the files in the best way possible.

1. Download the multimedia file and store it locally in a repository that you are curating.
2. Always download the file in the maximum resolution possible as you will be able to compress it later.
3. Always record basic metadata about the file, including the place where you obtained it from, the date, the author, title, description and license.
4. Use metadata schemas like Dublin Core to guide you in the selection of metadata terms and labels that can be relevant in your own research.
5. Always remember to generate a filename or file ID number that will allow you to keep track of the file and where it came from.



Coin from the mint of Salacia (MLH A.103.1)(Source: Hesperia, from the P.M.L collection).

Original_File	
File_Name	
Title	
Description	
Measures	
Provenance	
Obverse	
Reverse	

# Commands for File management

Description	Windows Command Prompt	macOS Terminal
Change to the specified directory	cd directory_name	cd directory_name
List files in current directory	dir	ls
List files in the current directory and all subdirectories	dir /s	ls -R
list all files including hidden	dir /a	ls -a

# Commands for File management

Description	Windows Command Prompt	macOS Terminal
Move up one directory level	<code>cd ..</code>	<code>cd ..</code>
Change to the root directory	<code>cd \</code>	<code>cd /</code>
Change the current drive and directory	<code>cd /d path</code>	<code>cd path</code>
Create a new directory	<code>mkdir directory_name</code>	<code>mkdir directory_name</code>
Create a new file	<code>echo.&gt; file_name</code>	<code>touch file_name</code>

# Commands for File management

Description	Windows Command Prompt	macOS Terminal
Delete the specified file	del file_name	rm file_name
Delete all files in the specified folder and its subfolders	del /s folder_name	rm -r folder_name
Delete the specified file	erase file_name	rm file_name
Copy a file to a new location	copy source_file destination	cp source_file destination

# Multidimensional Media

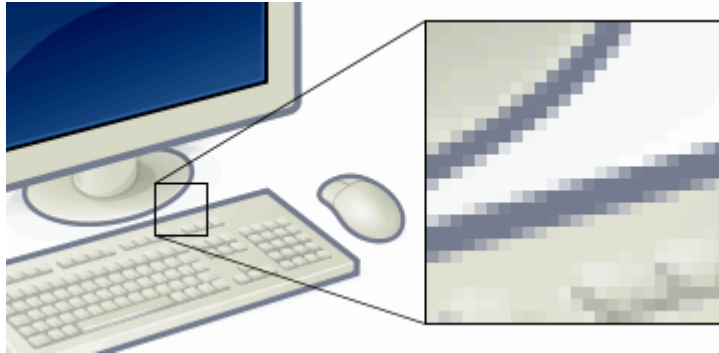
- Covers various types of non-textual data.
- Used to refer to derivative media created through digitisation processes or synthetic images produced via software.
- These processes often result in multi-part, multi-format and, often, large project files which describe the creation process.
- Data experts structure information according to the dimensions used to store the data in the computer system.
- For example a table with two types of values (Name and Surname) uses 2-dimensions as both rows and columns are used to organised the data.

Name	Surname
Paula	Granados
Brandie	Ratliff



# Digital Images

- Data in digital images is organised in a similar way to tables, as *pixels*, which are the smallest element of an image.
- Pixels are organised across 2-dimensions. We usually refer to these dimensions as the  $x$  and  $y$  axis.
- A digital image is a 2D array of pixels.
- Each pixel is characterised by its  $(x, y)$  coordinates and its value.
- Digital images are characterised by matrix size, pixel depth and resolution.
- The matrix size is determined from the number of the columns ( $m$ ) and the number of rows ( $n$ ) of the image matrix ( $m \times n$ ).



Example of pixels. © ed g2s from Wikimedia Commons

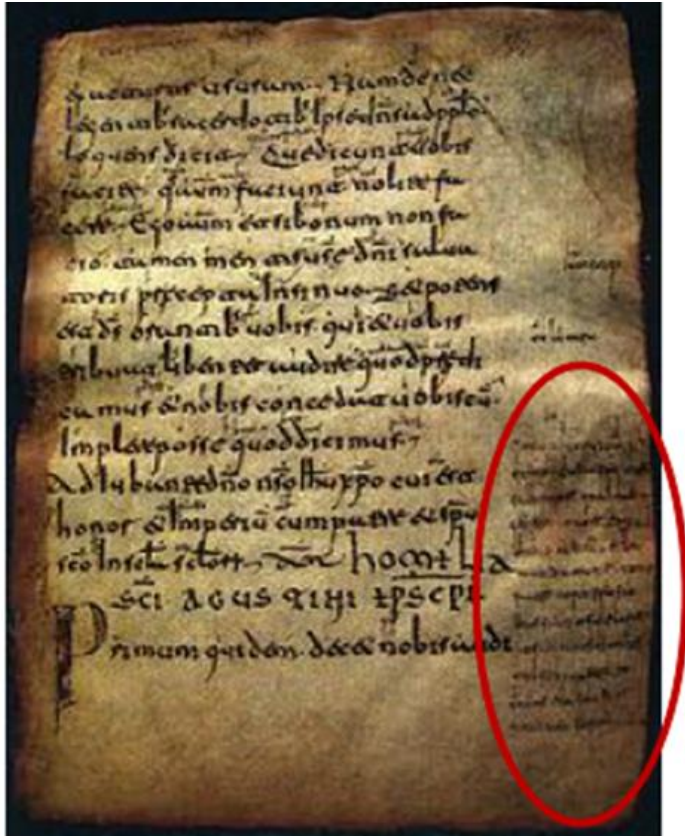
# Digital Images Formats

Most Common Image Formats:

- *TIFF: Tag Image File Format*, stores non-compressed information, making the file size larger.
- *JPEG: Joint Photographic Experts Group*, compresses pixel information making the files size smaller
- *PNG: Portable Network Graphics*,

Digital photography uses formats such as *RAW* to store raw image information. Other software store images in proprietary formats, so the images can only be opened by the respective software.

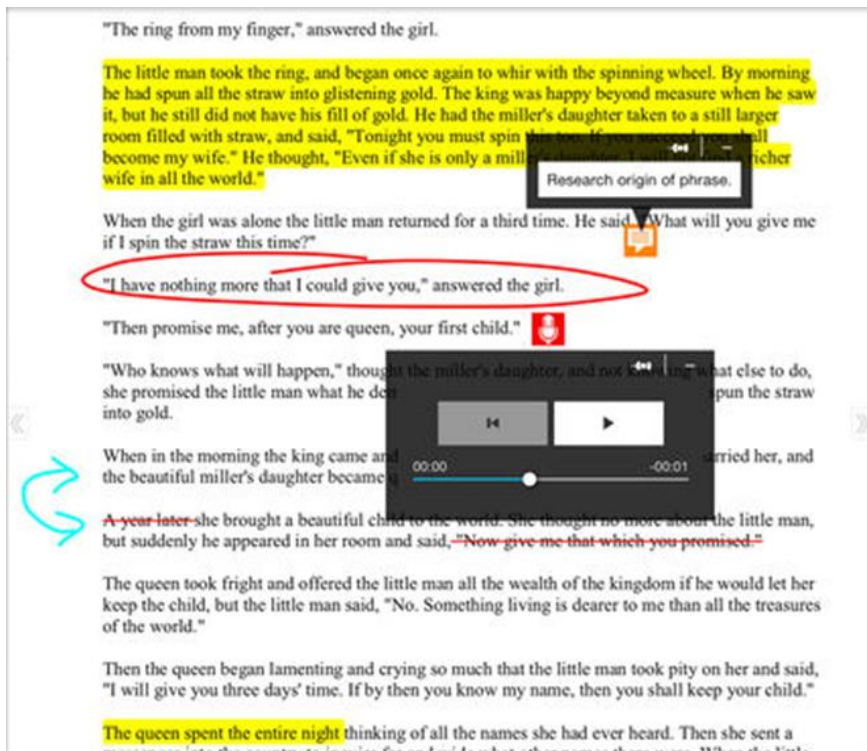
# Annotation



Hand written marginalia to a medieval manuscript.

- Individuals have for long enjoyed the ability to comment or annotate related thoughts during the reading.
- A way to extract information from the text that can then be collected and assimilated in different ways.
- Semantic annotation exists since at least the inclusion of marginalia in medieval manuscripts where it was common to add notes to the margins of the main text to comment on concepts, peoples or places mentioned in the manuscript.

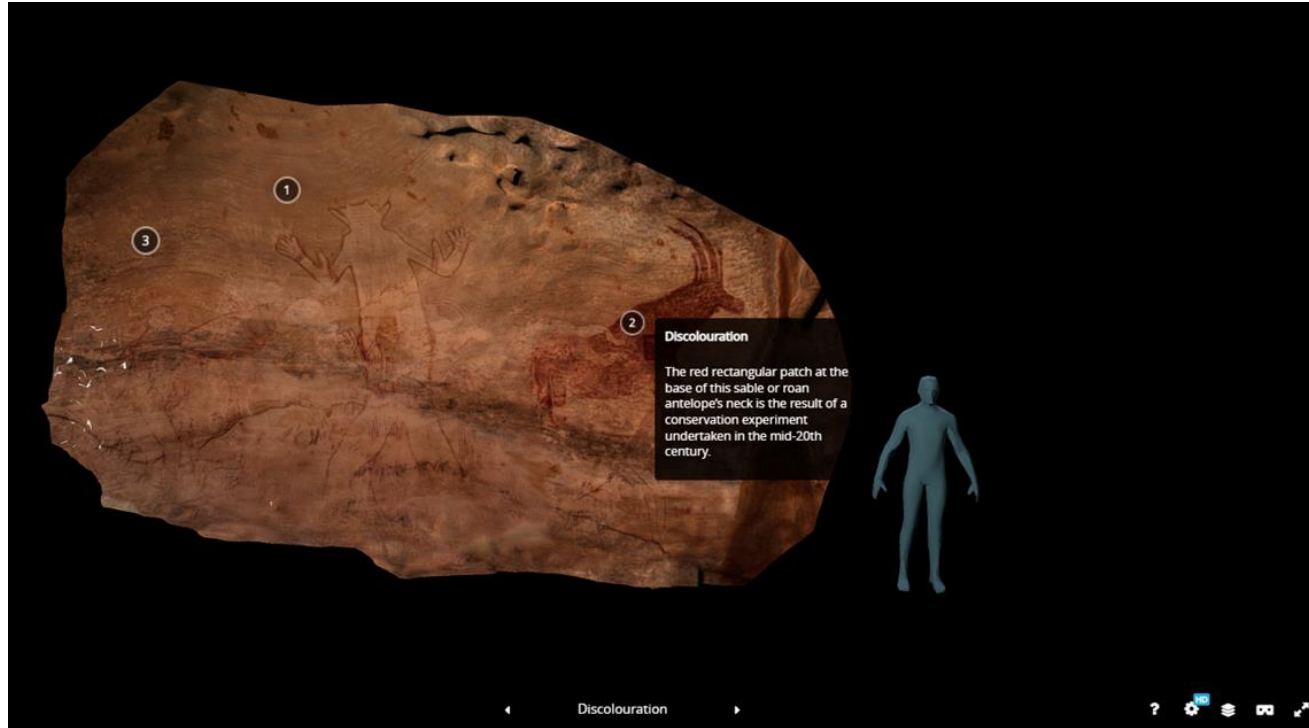
# Annotation



Digital Annotation

- In the digital context, annotation has become more popular in the last decades with the development of technologies.
- Whereas it maintains its original purpose, digital annotation improves potentially the enrichment of the source data, since it allows the user to append additional information to the file.
- Every annotation becomes data in itself.

# Digital Annotation



[Painted panel, Tassili n'Ajjer, Algeria from British Museum Sketchfab Account, African Rock Art project.](#)

# Image Annotation with VIA

Home Project Annotation View Help

Region Shape

Project

Name: via\_project\_15May2024\_11h

All files regular expression

[1] adutta\_swan.jpg  
[2] wikimedia\_death\_of\_socrates.jpg  
[3] Picture1.png


Add Files Add URL Remove

Attributes

Region Attributes File Attributes

attribute name + -

name



	name	type	image_quality
2	not_defined	Unknown (object)	<input type="checkbox"/> Blurred region <input checked="" type="checkbox"/> Good Illumination <input checked="" type="checkbox"/> Object in Frontal View

# Image Annotation with VIA

- Load image
  - Project → add local files
- Draw Regions
  - Select a region shape, press enter
- Create Annotations
  - Click the View → Toggle attributes editor to show attributes editor panel in left sidebar and add the desired file or region attributes (e.g. name).
  - Click View → Toggle annotations editor to show the annotation editor panel in the bottom side. Update the annotations for each region.
- Export Annotations
  - To export the annotations in json or csv format, click Annotation → Export annotations in top menu bar.
- Save Project
  - To save the project, click Project → Save in top menu bar.

# Digital Video

- Digital video comprises a series of digital images rendered sequentially in rapid succession, usually at 24, 30, or 60 frames per second.
- Digital video has many advantages such as easy copying, multicasting, sharing and storage.

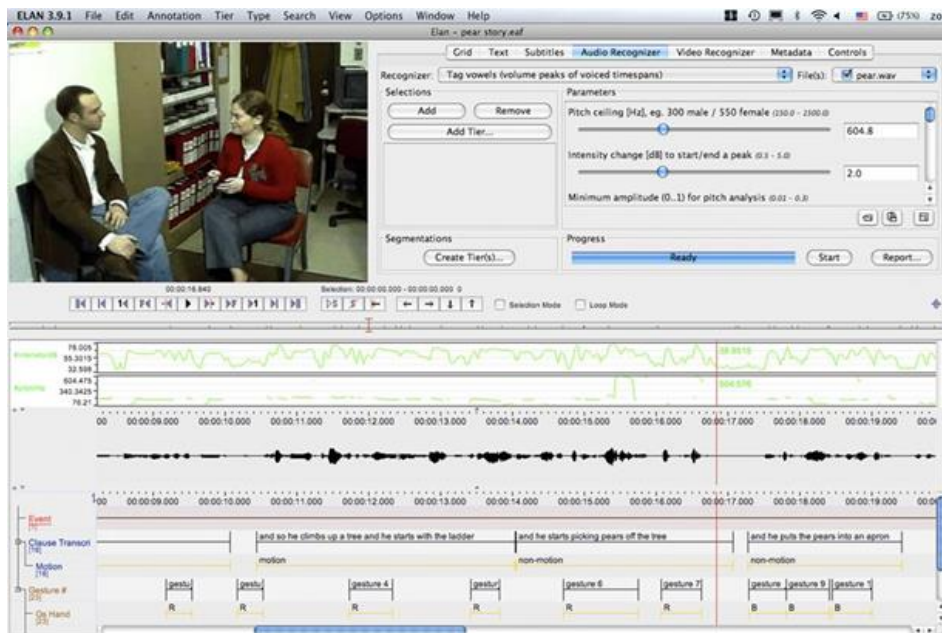
## Video Formats

There are many formats to store video and audio, most common are:

- MP4 (MPEG-4) is a common *container format* which can play in almost all devices and over the web.
- WebM and OGG are open video formats.



# Video Annotation with ELAN



## ELAN annotation software

- Open source and non-proprietary
- Excellent for audio(-visual) interview
- transcriptions and translations
- Time-aligned notes
- Contains all scripts (e.g., Arabic, Thai, etc.)
- One tool can serve multiple purposes
- Different data in different layers
- Lively community and product support
- Data exported in multiple formats

# Video Annotation with ELAN

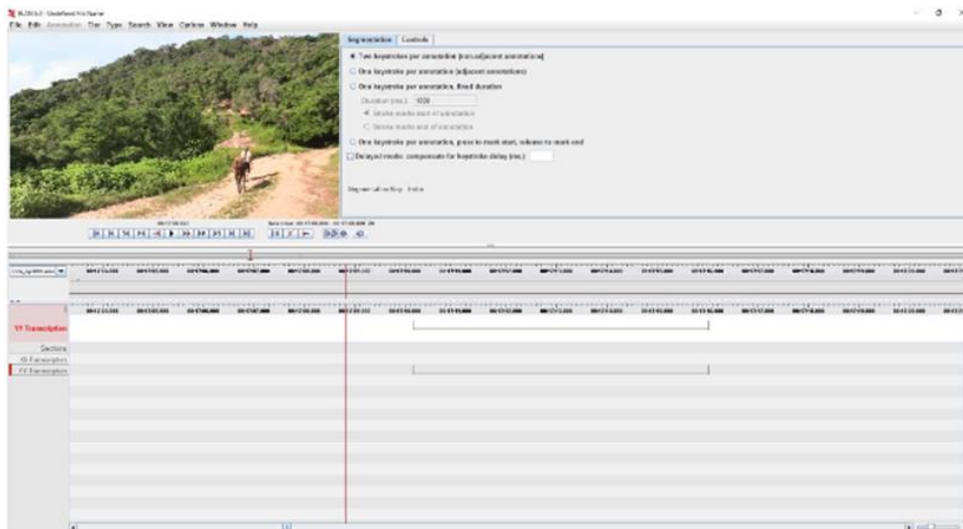
Click the ELAN icon in the folder where you have it stored.

2. Click 'File > New'.
3. In the new tab, navigate to the audio and video files you want to annotate.
4. Select the file and click on ">>" this will move it to the selected files list.
5. Click 'OK'. ELAN should then load the video ready for you to begin annotating.

Dashboard and controls:

1. Control panel for volume and pace
2. Right click on the photo to save current frame as an image.
3. See the photo extracted as particular frame.
4. Zoom in the video to get a closer picture of mouth and interactions.
5. Media streaming panel for video
6. Under media pane video controls, play, go forward, go back
7. Show frame by Frame >F. Also works without a selection.
8. Playing the video frame by frame or second by second will help you see some of the embodied actions of the participants.
9. Make selection on segmentation pane with cursor, play segment, go back or go forwards.

# Video Annotation with ELAN



In 'Segmentation Mode', we can see our video in the top left corner, with settings on the right-hand side. Below is the waveform of our audio file. Below that, is the timeline that shows the timespan of our recording visually, from left to right.

The timeline is divided horizontally into tiers; each tier represents a particular kind of information we want to capture (translation, technique, other commentary).

# Video Annotation with ELAN

1. Create Tiers.
  - a. Type - add type name.
    - i. Translation or Transcription.
  - b. Tier- Create new Tier.
    - i. Add Tier name.
    - ii. Add annotator name.
    - iii. Linguistic type - Transcription.
    - iv. In more options you can change colour, highlight and font.
    - v. If you want to delete default tier, just go to Tier, delete tier and have default selected.
2. Annotation segmentation and transcription.
  - a. Make selection
  - b. Hover over selection on top of Tier section.
  - c. Double click on selection to insert text.
  - d. Move selection if needed.

# Video Annotation with ELAN

Save as word

- a. Go to File > Export As
- b. Export as Interlinear text
- c. Select the different tiers you want to include on the right
- d. Shot time codes - hour millisecond millisecond, time code is underneath
- e. Include Silence duration to show -- SD
- f. Select how the transcript is presented.
- g. Tabs instead of spaces
- h. Save as and give text file a name
- i. Find the new transcript export as a txt file.

Export as subtitles text.

# 3D Images

- Photogrammetry: science and technology of obtaining reliable information about physical objects and the environment through the process of recording, measuring, and interpreting photographic images and patterns of electromagnetic radiant imagery and other phenomena. It has broad applications in various fields such as mapping, architecture, engineering, archaeology, and geology.
- As an accessible alternative to 3D scanning, photogrammetry allows the generation of 3D models from several pictures of the same object taken from different points of view, a method that has proven to be very useful in the documentation of artefacts and material culture in the field.
- 3D models allow the production of high quality visualisations of objects, offering a better understanding of the object's materiality, since the models can be moved and rotated in space, and allowing higher quality zoom-in than other imaging methods.

# 3D Images

The software that we are going to use for this photogrammetry session is called Metashape. Other options of photogrammetry software are 3DF Zephyr and Meshroom.

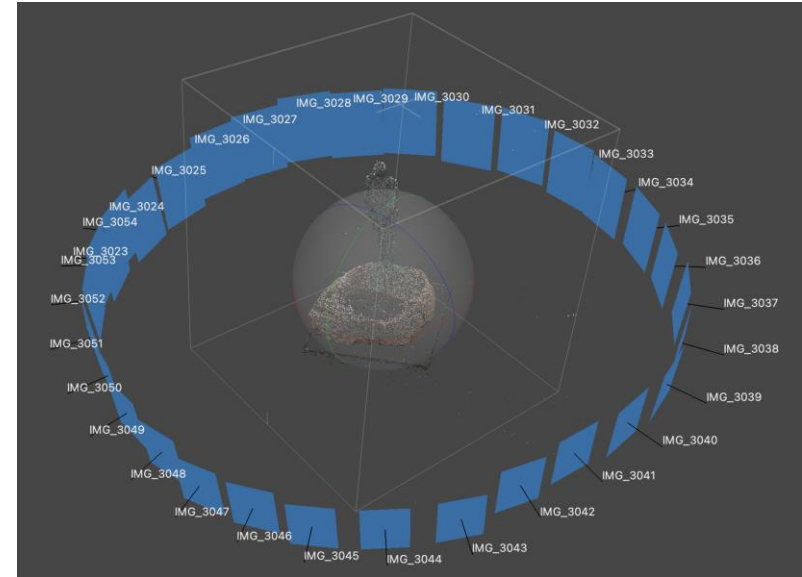
Photogrammetry involves creating 3D models from photographs and requires high-quality image datasets. Here are some sources where you can find free image datasets suitable for photogrammetry:

1. 3Dflow - 3Dflow offers free sample datasets that can be used with their photogrammetry software, 3DF Zephyr. You can download these datasets and use them to practice your photogrammetry skills.
1. Metashape also provides sample datasets to test with the software.
1. Sketchfab - While Sketchfab is primarily a platform for sharing 3D models, it also provides downloadable datasets that include photos used to create these models. Some datasets are available for free, often under Creative Commons licenses.

# Metashape Tutorial

## Step 1: Capturing Images

1. Equipment: Use a digital camera or a smartphone with a good camera. A tripod can help in capturing steady shots.
2. Lighting: Ensure consistent lighting to avoid shadows and overexposure. Overcast days are ideal for outdoor captures.
3. Coverage: Take multiple photos from different angles to cover the entire statue. Overlap each photo by about 60-80% with the adjacent ones.
4. Distance: Maintain a consistent distance from the statue. For a small statue, a distance of 1-2 meters is usually sufficient.
5. Settings: Use manual settings to keep exposure and focus consistent. Set a low ISO to reduce noise and a narrow aperture (high f-number) for greater depth of field.





# Tutorial

## Step 2: Setting Up Metashape

1. Download and Install: Download Agisoft Metashape from the [official website](#) and install it on your computer.
2. Open the Software: Launch Metashape after installation.

## Step 3: Importing Images

1. Create New Project: Start a new project by clicking on "Workflow" and selecting "Add Photos" to import the images you captured.
2. Load Photos: Ensure all images are loaded and visible in the project workspace.

# Tutorial

## Step 4: Initial Processing

1. **Align Photos:** Go to "Workflow" and select "Align Photos". This process detects common points between images and aligns them.
2. **Settings:** Use default settings for accuracy and pair preselection. Higher accuracy settings provide better results but take longer to process.
3. **Review Alignment:** Once the alignment is complete, review the sparse point cloud generated. If it looks correct, proceed. If not, adjust settings or add more images and re-run the alignment.

## Step 5: Dense Point Cloud

1. **Build Dense Cloud:** Go to "Workflow" and select "Build point Cloud". This will create a detailed point cloud of the statue.
2. **Settings:** Choose the quality and depth filtering settings. Medium quality is a good starting point, but higher quality will produce more detail at the cost of processing time.
3. **Start Process:** Click "OK" to start generating the dense point cloud. This may take some time depending on the number of images and your computer's performance.

# Tutorial

## Step 6: Mesh Generation

1. Build Mesh: Once the dense point cloud is generated, go to "Workflow" and select "Build Mesh" to create a mesh from the dense cloud.
2. Settings: Choose the source data (dense cloud), surface type (arbitrary for most cases), and face count. Default settings usually work well.
3. Generate Mesh: Click "OK" to start the mesh generation process.

## Step 7: Texturing

1. Build Texture: After the mesh is created, go to "Workflow" and select "Build Texture".
2. Settings: Choose the texture mapping mode and size. Default settings are a good starting point.
3. Apply Textures: Click "OK" to start the texturing process. This will map the original photos onto the 3D model for a realistic appearance.

# Tutorial

## Step 8: Review and Export

1. Review Model: Once texturing is complete, review the 3D model for any errors or areas that need refinement.
2. Export Model: When satisfied with the model, click on "File" > "Export" and choose the desired format (e.g., OBJ, STL) to export the 3D model.

## Step 9: Post-Processing (Optional)

1. 3D Printing: If you plan to 3D print the model, you may need to use additional software to prepare it for printing, such as MeshLab or Cura.
2. Further Editing: For more advanced editing, you can import the model into software like Blender or Autodesk Meshmixer.