INITIATIVE OF DIGITIZATION IN

NATIONAL ARCHIVES OF MALAYSIA

01

Managing The Digitization Project

ZAIDIL AKMAL BIN ZAINUDDIN
HEAD OF SECTION
DIGITAL ARCHIVES SECTION
NATIONAL ARCHIVES MALAYSIA

Table of Contents

- Defining Things
- Team Building
- Planning and Managing Project
- Digitization in National Archives of Malaysia
- Digitization Standard

Defining Things : Define the stakeholders

- Identify stakeholders
- Make a survey
- Create goal/outcome

Defining Things : Define the Objectives

Why Digitize?

- ❑ To preserve fragile materials
- ❑ Provide better access
- ❑ Added value to resources
- ❑ Fulfill strategic mission and goal of institution

Defining Things : Define the Criteria

- ❑ How much do you digitize?
- ❑ When will the project start?
- ❑ Long-term plans (sustainability)?

Defining Things : Define the Way of Digitization

Outsourcing or in-house? Pros and Cons of Outsourcing :

Pros
1.All in-one-go
2.Latest equipment and trained staff, availability
3.The costs of technology obsolescence, failure, downtime, staff changes, etc., are borne by the service provider

Defining Things : Define the Way of Digitization

Cont'd

Cons
1.Knowledge gap between the vendor and your organisation
2.Moving out fragile materials
3.Less control over digitising process
4.The service provider maybe out of business
5.Security and privacy issue on confidential materials
6.Proprietary data on materials

Defining Things : Define the Way of Digitization

- Permanent staff assigned, equipment purchased, software developed locally
- Temporary staff hired, equipment purchased, software developed locally
- Permanent and temporary staff employed, hardware rented, software "subscription"
- Scanning and metadata creation performed by vendor

Defining Things : Define the Audience of the Project :

For whom?

- Primary audience
- Secondary audience

Defining Things : Define the materials

What is digitized?

- What should/could be digitized?
- Priorities?
- The output of the project?

Team Building : Team Members

- Who are the members in a team? Examples:

Project Manager	IT hardware vendor	System engineer
Data storage vendor	In-house / outsource staff -Data entry -Scanning staff	Web developer
Scanning specialist	Repository / storage staff	Production specialist

Team Building : Committees

- Steering committee
- Working committee
- Technical commitee

Team Building : Stakeholders

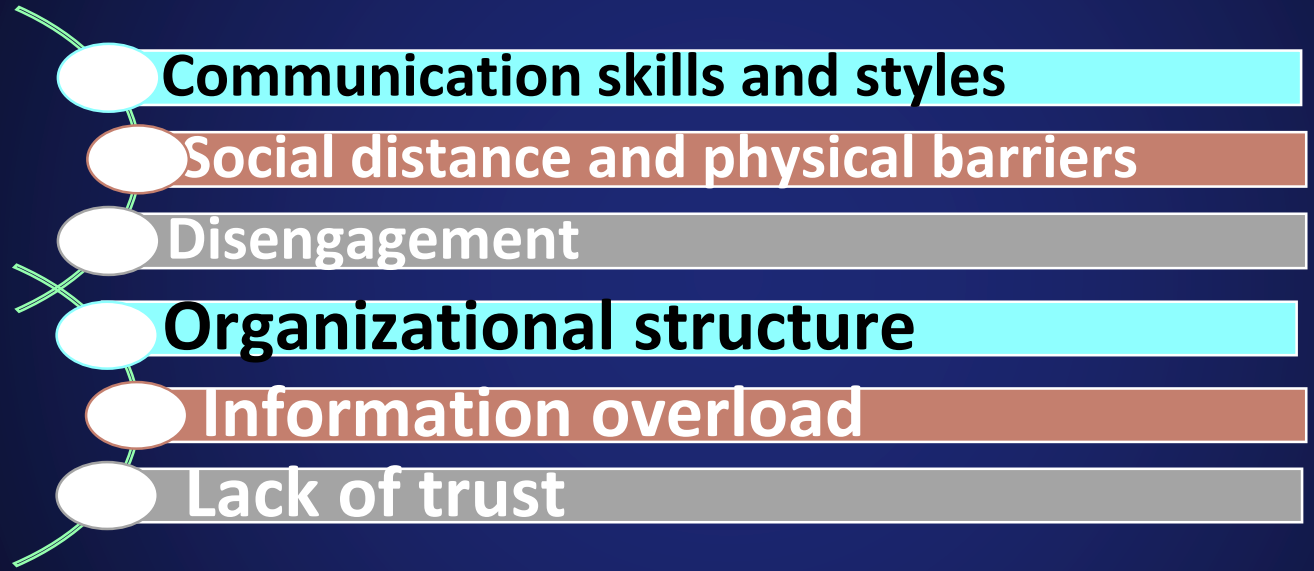
Success of projects depends on developing a core team of stakeholders

Stakeholders may be part of the institution, parent institution, or partners in the project

Team Building : Meetings

- How frequent?
- What is meeting for?

Team Building : Avoid Communication Barriers



Planning and Managing the Project : What is Planning?

- Planning - a primary managerial activity that involves:
  - Defining the organization’s goals
  - Establishing an overall strategy for achieving those goals
  - Developing plans for organizational work activities

Planning and Managing the Project : Purposes of Planning

- Provides direction
- Reduces uncertainty
- Minimizes waste and redundancy
- Sets the standards for controlling

Planning and Managing the Project : Equipment

- The cost is high when the final output quality Is high
- Warranty/maintenance
- Workflow and productivity
- Avoid unproved technologies
- Rent equipment

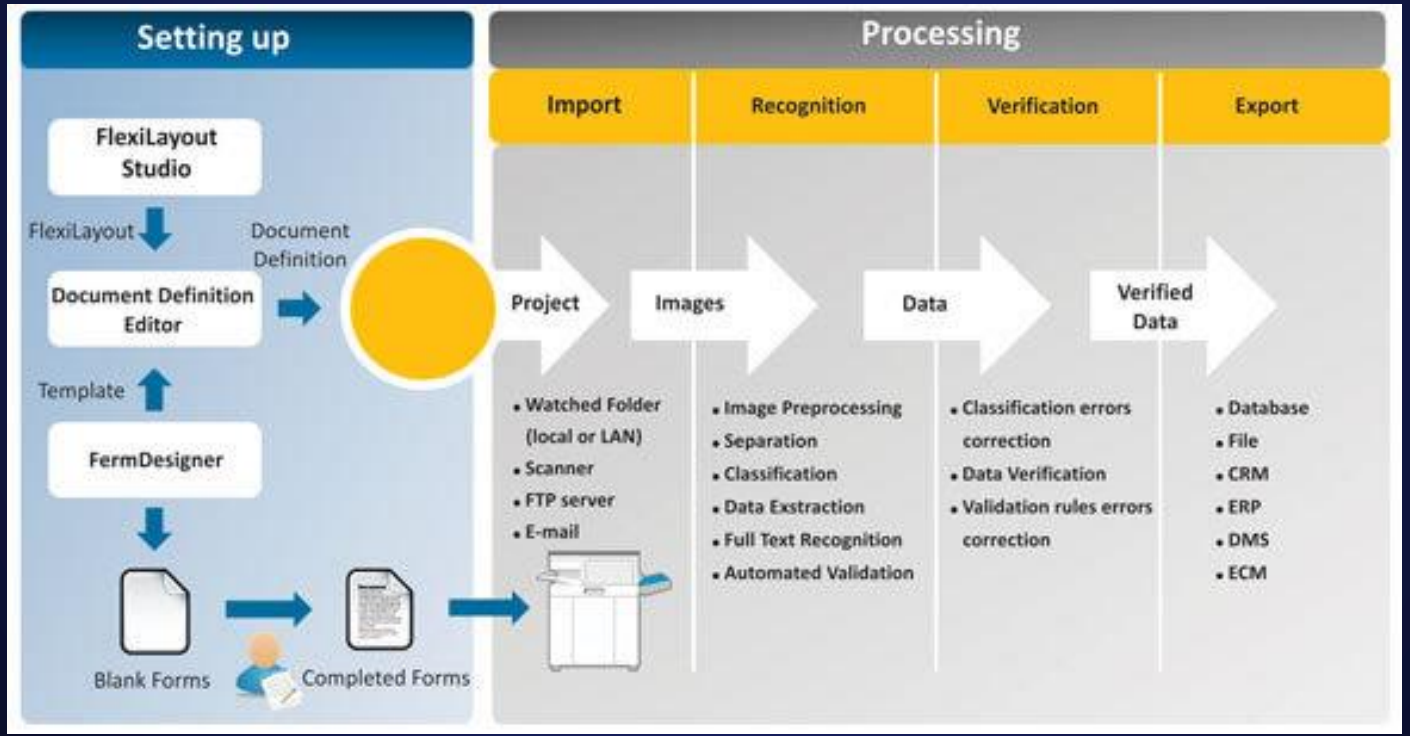
Planning and Managing the Project : Here's the Manager

- Here is the manager/managers
- Where Do Managers Work?
- Manager cannot work alone

Planning and Managing the Project : The Goal of the Project

“When completed at the end of the first year, the project will have digitized a minimum of 5,000 images and related historical records and textual documentation into the AAA Digital Repository.”

Planning and Managing the Project : The Workflow



Planning and Managing the Project : The Timeline

Project Month	Action	Steps Taken	Who is responsible
Jan-Feb 2021	Hire Project staff; buy hardware/software	Interview candidates; training	Project manager, system engineer, storage vendor
March 2021	System and hardware installation, testing and acceptance	Prepare space, monitoring installation, PAT, UAT, FAT	System engineer, storage vendor, scanning / data entry staff,
April 2021	Scanning and metadata creation	Staff scan materials and add data	Scanning and conversion specialists, scanning and data entry staff
May 2021	Publicity, Presentations, Post-Grant activities,	Making promotion	Outreach unit

Planning and Managing the Project : The Collection

- Physical condition
- Copyright
- Value

Planning and Managing the Project : Metadata

- What is metadata?
- Type of metadata
  - Descriptive
  - Administrative
  - Structural

- Managing the chosen metadata

Planning and Managing the Project : Training

The related topics covered during the training/workshop are:

- Introduction on Document Imaging
- Conversion Procedures / Standard Operation Procedures
- Acceptance Procedures
- Quality Assurance / Quality Control
- Uploading

Planning and Managing the Project : Reports

- Progress Report
- Final Report

Digitization in National Archives of Malaysia : Background

- Computerised Archival System and Services (COMPASS)
  - Developed in 1996
  - Started operation on 2 January 1997
  - Archival holdings online database

Digitization in National Archives of Malaysia : Background

- Data conversion from manual list to digital
- Started on 1st March 1998.
- Data of 39,000 files from Selangor State Secretariat keyed in into COMPASS
- In-house

Digitization in National Archives of Malaysia : Data Conversion

- Data and Image Conversion into Digital Format

1. Project DC 1 (2000 – 2001)
   - 63,858 digital data and images.

2. Project DC 2 (2002 – 2003)
   - 117,608 digital data of Selangor State Secretariat

Digitization in National Archives of Malaysia :
Data Conversion

3. Project DC 3 (2004 – 2006)
   - 54,785 digital data and images and 42,284 personal collection materials (SP)

4. Project DC 4 (2007 – 2010)
   - Selangor State Secretariat data conversion
   - 217,432 files
   - 468 artifact materials and 609 personal collection materials

Digitization in National Archives of Malaysia : Digitization by Digital Archives Section

In 2011

1. Pahang State Secretariat
   - Total files : 844
   - Pages : 48,341

2. Publication
   - Total : 1,777
   - Pages : 400,000

Digitization in National Archives of Malaysia : Digitization by Digital Archives Section

Project MyCC (2012-2013)

- Project MyCC
  - Equipment
    - Colour scanner A1 Size (1 Unit)
    - Colour scanner 2A Size(7 Unit)
    - PC/workstation (25 Unit)
  - Digitization
    - 3,000,000 pages of archival materials

Digitization in National Archives of Malaysia :
Digitization by Digital Archives Section

In 2020

- 45,494 images scanned from files

In 2021

- 7 Temporary staff / MySteps – microfilms
- Outsourcing project will start in this month

In 2022

- About 300,000 images from files

Digitization Standard : Empowerment

- National Archives Act 2003 (Act 629)
  - Section 36: Processing and preservation of public archives
    - All public archives shall –
      - (d) where necessary, be transferred to microfilm recordings, photographic copies, or any other forms of storage

- MS 2473:2012
  - Information and documentation – Implementation guidelines for digitization of records

Digitization Standard : MS 2473:2012

- This standard was developed by the Technical Committee on Records Management, as follows:
  - Bank Negara Malaysia
  - Chief Government Security Office, Prime Minister's Department
  - Datarunding Sdn. Bhd.
  - Malaysia Administrative, Modernisation and Management Planning Unit (MAMPU)
  - National Archives of Malaysia
  - PETRONAS, Petroleum Resource Centre
  - Public Service Department of Malaysia
  - SIRIM Berhad (Secretariat)
  - Universiti Kebangsaan Malaysia
  - Universiti Teknologi MARA
  - Department of National Library
  - National Film Development Corporation (FINAS)

Digitization Standard : Clauses

Clause 5 – Preliminary considerations

- 5.1 digitization project viability assessment
  - Clearly outline benefits
  - Budget
  - Resource commitments -

Digitization Standard : Clauses

Clause 5 – Preliminary considerations

- 5.2 Master copies and derivatives
  - What format? TIFF/JPEG/PDF
  - Should destroy or keep hardcopy after digitization
  - Managed according to any records legislation or regulatory requirements

Digitization Standard : Clauses

Clause 6 – Best practice guidelines

- 6.3.3 Non-digital source record preparation
  - Physical preparation for digitization (e.g. staple removal, alignment of single pages)
  - Procedures to enable checking and verification that all target non-digital source records have been included in the digitization process

Digitization Standard : Clauses

Clause 6 – Best practice guidelines

- 6.3.5 Quality control
  - Scanner operation quality control
  - Verification that digital output matches the quantity of non-digital source record input
  - Criteria for checking image quality
  - Processes for re-digitizing

Digitization Standard : Clauses

Clause 6 – Best practice guidelines

- 6.4.2 Back-up procedures
  - Back-up regimes should be documented and back-up copies maintained to a level of security that ensures the authenticity of the records used in recovery situations

Conclusion

Planning and managing are essential to determine needs and cto reate path towards goals in digitization project.

02

WORKFLOW & TECHNICAL SPECIFICATION FOR DIGITIZATION OF ARCHIVAL MATERIALS

NURUL AIQA BINTI MOHAMAD ZAIN
ARCHIVE OFFICER
DIGITAL ARCHIVES SECTION
NATIONAL ARCHIVES MALAYSIA

SCOPE

- INTRODUCTION
- DIGITIZATION ACTIVITIES (BEFORE, DURING & AFTER)
  - ❖ PROCESSES & PROCEDURES OF DIGITIZATION
  - ❖ SPECIFICATIONS OF DIGITIZATION
- CONCLUSION

DEFINITION OF DIGITIZATION

- The process of converting data from analog form (text, image, audio, video) to digital form by using suitable electronics tools (such as scanners) so that the information can be processed, stored, transmitted and accessed through networking, equipments and computers.

THE SELECTION OF ARCHIVAL MATERIALS

3 CRITERIAS:

- Popular materials (regularly referred by researcher)
- Unique, special collection, high value
- Old archive materials

THE NEEDS OF DIGITIZATION

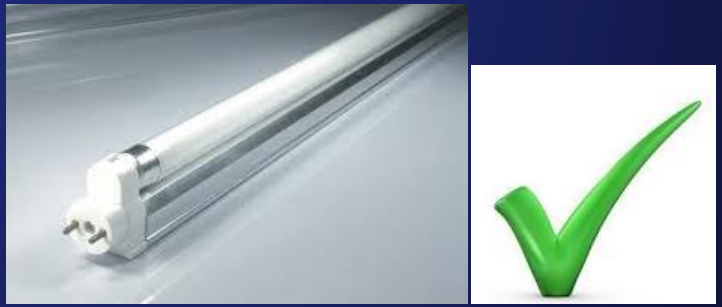
Needs Of Digitization Activities

- Room / Digital Lab
- Records / Archives Materials
- Hardware, Equipment and ICT Software
- Information such as metadata
- Staffs

FEATURES OF DIGITAL LAB

Lighting

- Not too bright and not too exposed to sunlight.
- The use of downlights and standing lights is not permitted. However, the position of the ceiling light can be changed according to the suitability of the position of the scanner and computer monitor.



FEATURES OF DIGITAL LAB

Humidity

- Humidity Workplace - RH 40% (+-5%).

FEATURES OF DIGITAL LAB

Wall Colour

The walls of the digitization room must be painted in matte grey and the percentage of the colour is 60 percent (60%)

Colours other than grey;
and;
colour types glossy ones are not recommended

- as they will cause non -uniform colour reflection and light arrangement on the scanned documents.

FEATURES OF DIGITAL LAB

Scanners Location

- Distance between scanners (*overhead type*) and the wall must be at least 0.5 meter.

- It is to reduce colour reflection from the wall which will affect the output of the digital image produced.

TYPES OF SCANNERS

1. Flatbed/Shedfed Types
2. Overhead Types
3. Camera

TYPES OF SCANNERS

Camera Types

TYPES OF SCANNERS

Flatbed/Shedfed Types

TYPES OF SCANNERS

Overhead Types (High End)

CopyBook Brand

Book Eye Brand

Zeutschell Brand

SCANNERS

Seven (7) A2 size scanners (Zeutschel)
One (1) A1 size scanner (Zeutschel)
Two (2) A3 size scanners (Zeutschel)

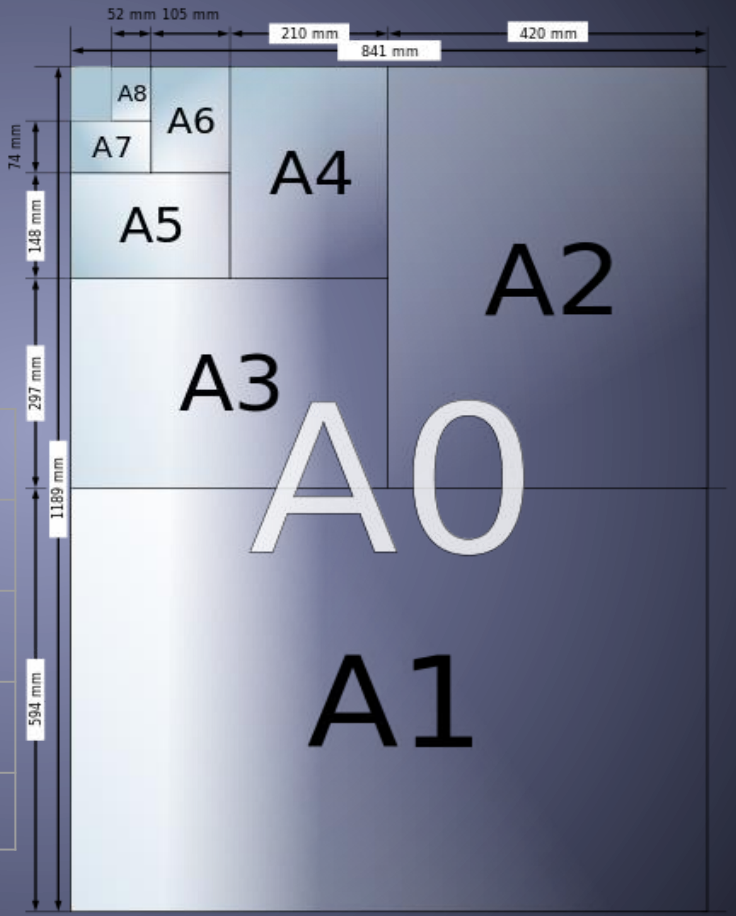
A1 Size

A2 Size

SCANNERS SIZES

The use of scanners depends on the size of the archival materials to be scanned.

MATERIALS SIZES	SCANNERS SIZES	REMARKS
Materials size > A1	A0	2 x A1
Materials size > A2	A1	2 x A2
Materials size > A3	A2	2 x A1
Materials size > A4	A3	2 x A4



Infrastructure & Facilities	Total
Digital Laboratory	1
Description room	1
Scanner A1	1
Scanner A2	7
Scanner A3	2
Microfilm Scanner	1
Workstation (PC)	24
Server Storage	1 PT @ 1000TB

HARDWARE, EQUIPMENTS, ICT SOFTWARE

ICT Equipment

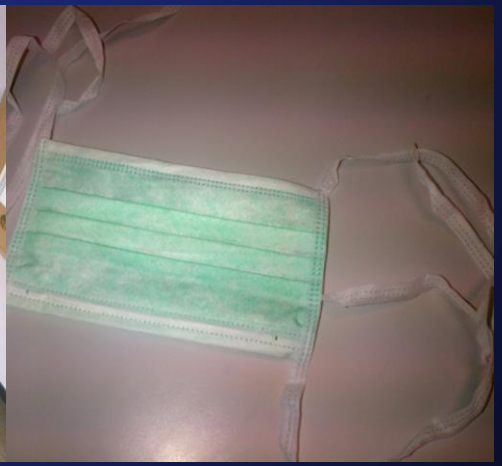
Digital Lab – National Archives of Malaysia

Digital Lab – National Archives of Malaysia

HARDWARE, EQUIPMENTS, SOFTWARE

Equipments

1. Trolley
2. Boxes
3. Mask
4. Gloves



SOFTWARE

Software

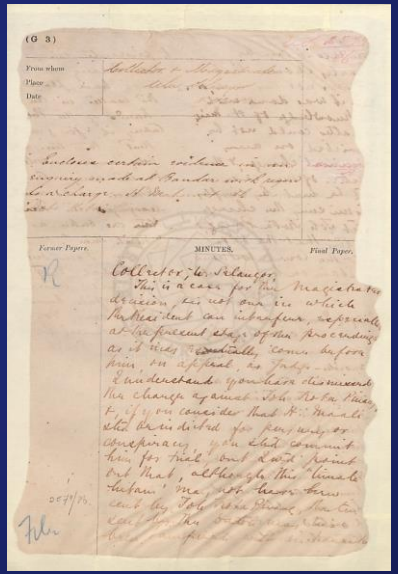
1. Scanners Software
2. Editing Software
3. Optical Character Recognition (OCR) Software
4. Pdf Conversion Software
5. Watermarking Software

SOFTWARE

Image Editing Software

Example :

- Nuance Paperport



SOFTWARE

Watermark Software

Example :

- PDF Watermark

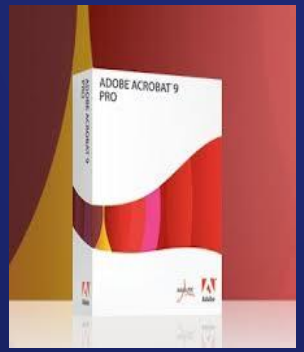
SOFTWARE

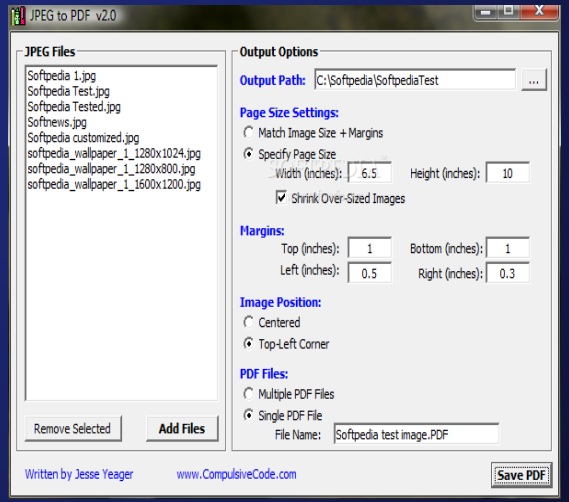
Pdf Conversion Software

- PDF = Portable Document Format
- Combines many JPEG files into one PDF files (merge)
- Rearrange all pages in a PDF document

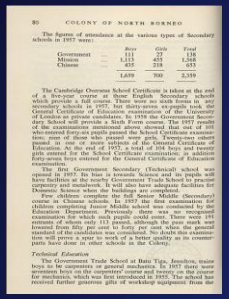
Software example

- Adobe Acrobat
- PDF Converter

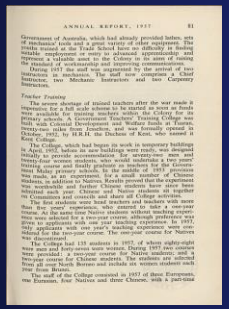




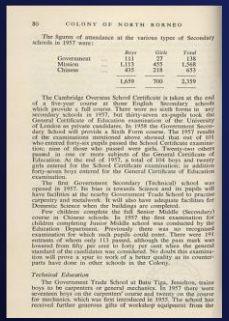
Pdf Conversion Software



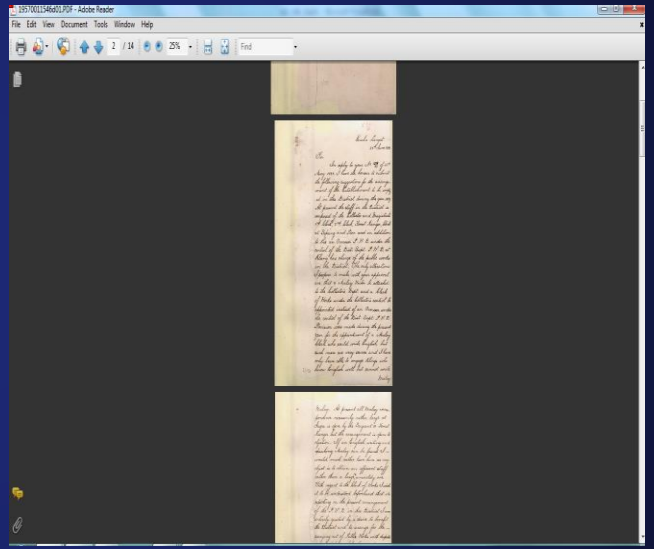
JPEG Image 1



JPEG Image 2



JPEG Image 3



Multipage PDF

Importance of Specification

1. The technical digitization specifications need to be known by each operator before operating the digitization process.

1. A high level of knowledge of digitization specifications allows the digitization operator to solve problems that occur during the process.

1. The use of standard parameters will produce a consistent digital image.

Digital file format

- TIFF – Tagged Image File Format
- PDF – Portable Document Format

DIGITIZATION SPECIFICATION

Materials Types	Format Types	Technical Specification
Document	TIFF Uncompressed, PDF	Colour DPI : 300 dpi Bit Depth : 24 Bit Black and White DPI : 300 dpi Bit Depth : 8 bit Grey-scale DPI : 300 dpi Bit Depth : 16 bit

TIFF vs PDF

TIFF Uncompressed	PDF
Suitable for digital preservation	Suitable for access
Big file size i.e. : 40MB	Small file size i.e. : 9MB
No Compression – 100% data as original	Lossy Compression
BEST QUALITY!	Good quality
Suitable for image/graphic	Suitable for text / image

FILE RESOLUTION - DPI

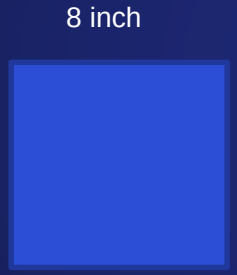
- Image resolution can be defined as how many dots or points there are in one inch of a digital image.

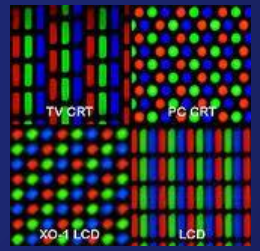
- It is measured in the form of 200, 240, 300, 400 dpi or ppi (*dots per inch or pixel per inch*).

- The total number of pixels in a digital image depends on the resolution and size of the original material.

FILE RESOLUTION - DPI

- For example, an 8 -inch x 10 -inch image material scanned using 100 DPI, will produce an image with a resolution of 800,000 pixels.





Resolusion = (Lenght x dpi) x (Width x dpi) 800,000 pixel = 800 MegaPixel

= (8 inch x 100 DPI) x (10 inch x 100 DPI).

= 800, 000 pixel

FILE RESOLUTION - DPI

- The higher the resolution of a digital image, the larger the file size of the digital image.

- Therefore, if a material is scanned at a high image resolution (for example 1000dpi), then it will produce a high total number of pixels and in turn, it require large storage.

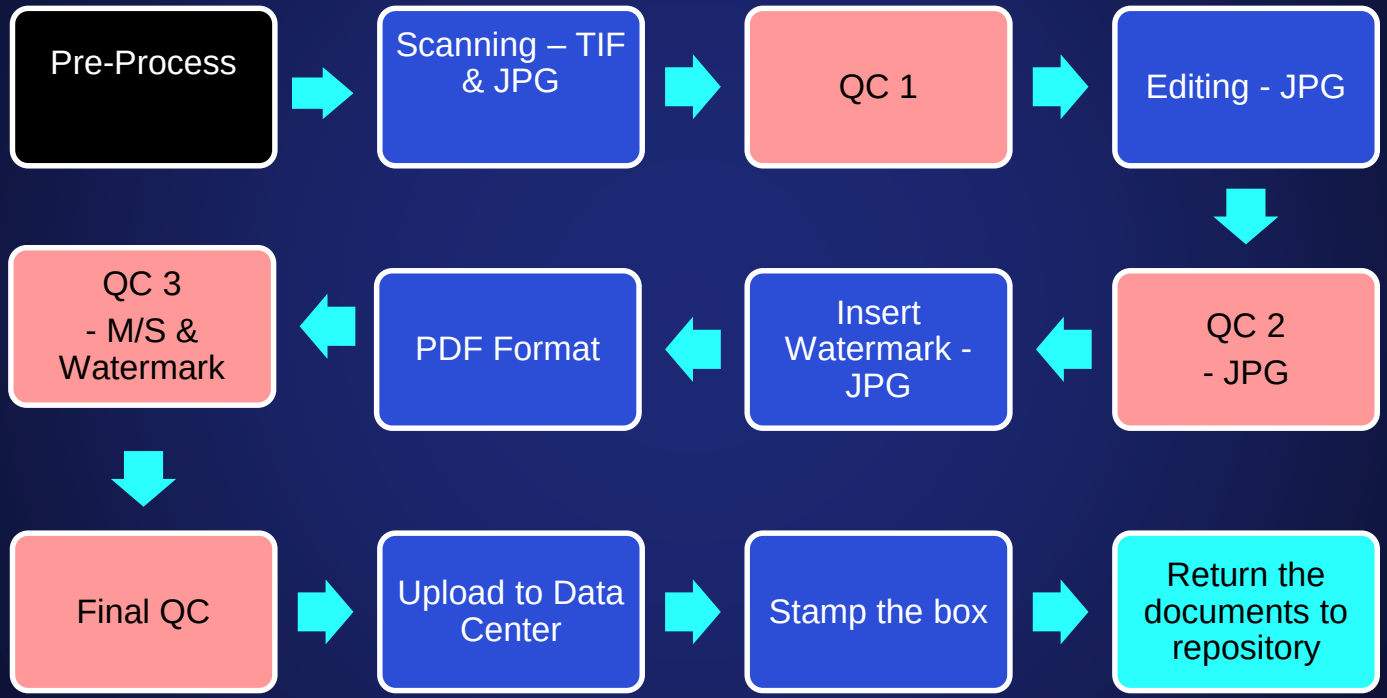
Before Digitization (Phase 1)

1. Identify records to be digitized
2. Select the priority of the record to be digitized
3. Define storage and network for digitized records (system / temporary storage / data center)

DIGITIZATION

(Phase 2)

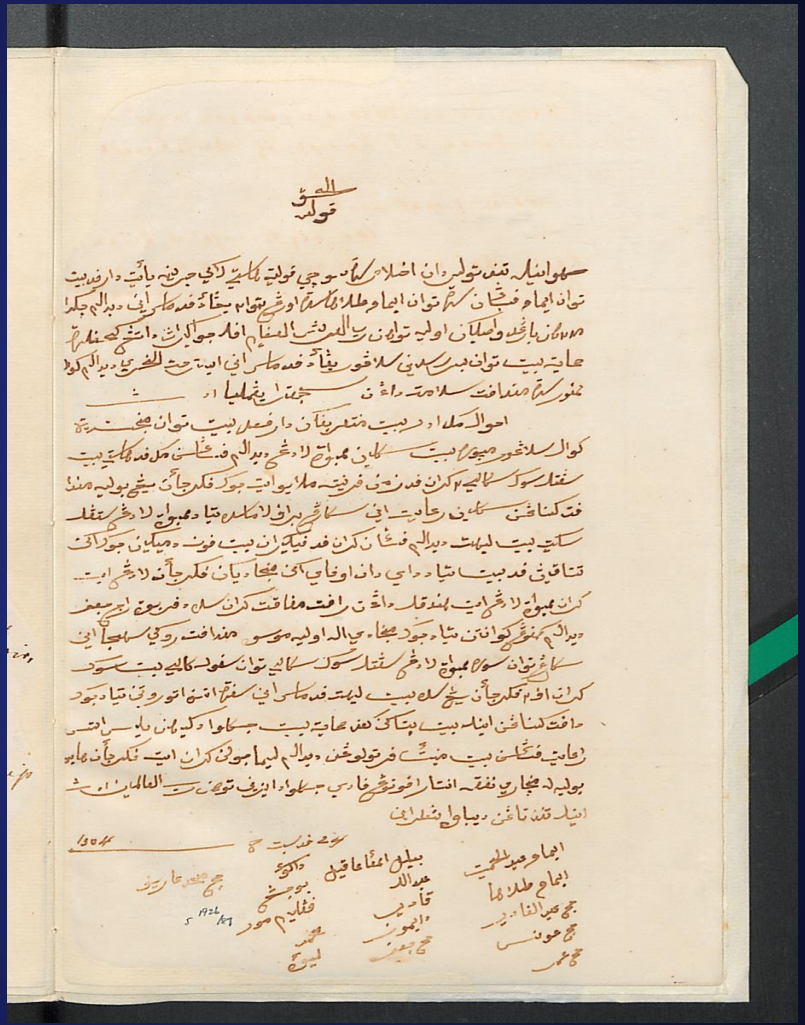
DIGITIZATION WORK FLOW



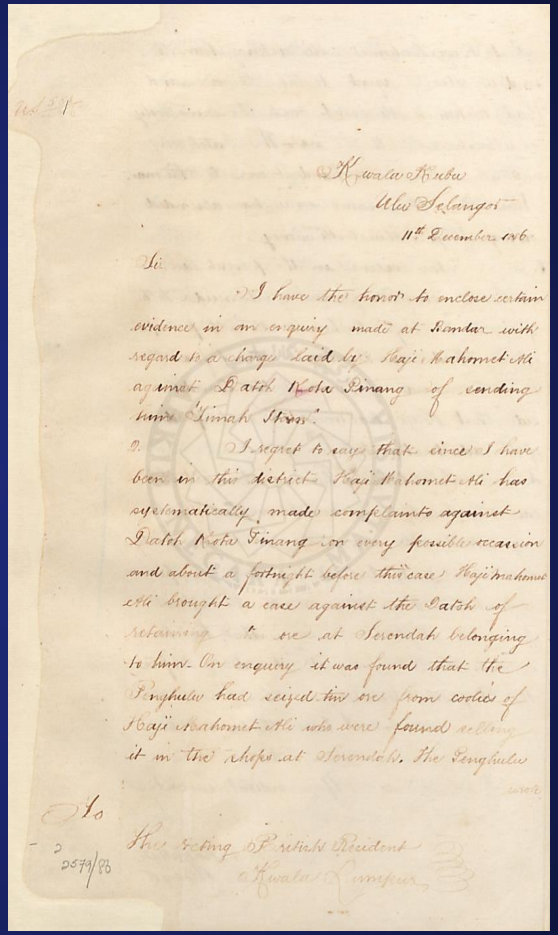
PRE - PROCESSES

Scanning

Example of TIF Image
(preservation)



Example of JPEG image



Kwala Kubu
Ulu Selangor
11th December 1886

Sir

I have the honor to enclose certain evidence in an enquiry made at Rembau with regard to a charge laid by Haji Mahomet Ali against Datoh Kaya Dinang of sending him Limah Stan.

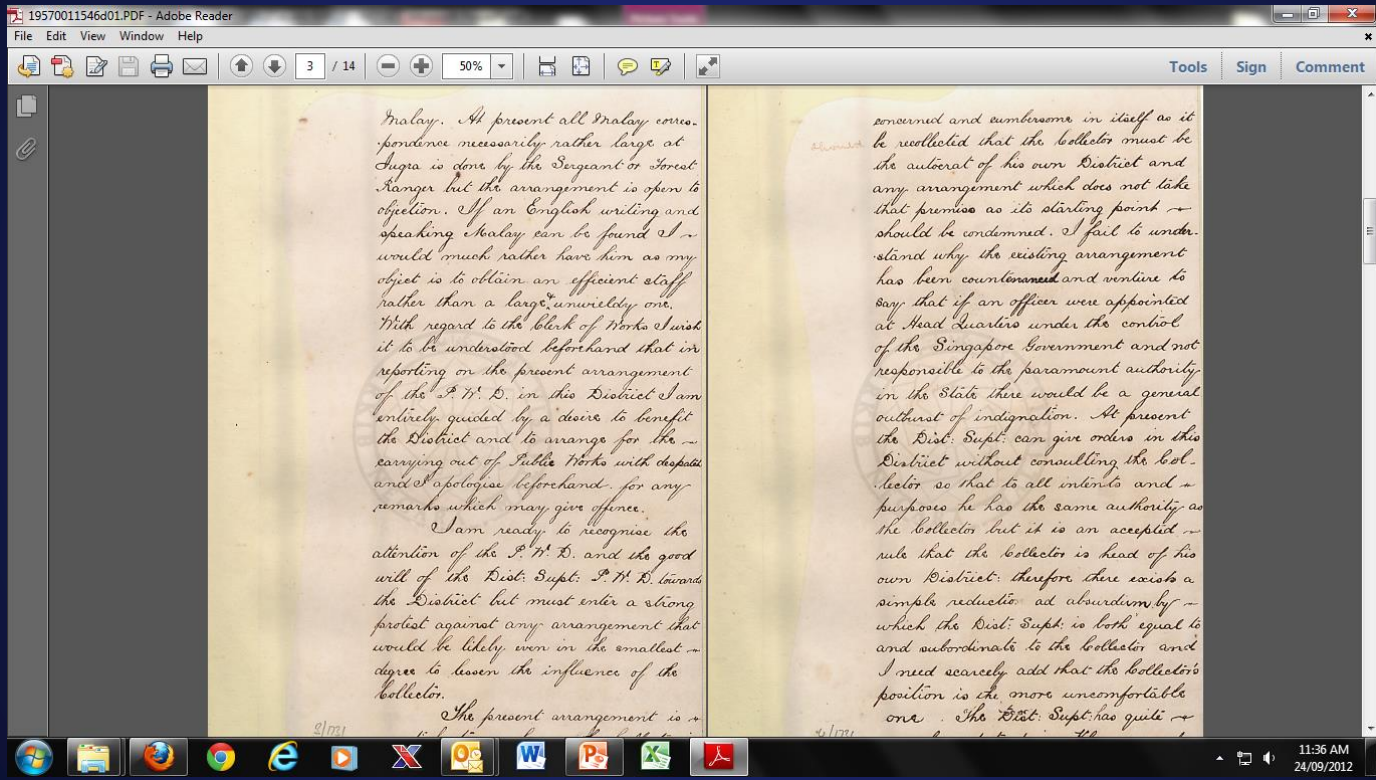
2. I regret to say that since I have been in this district Haji Mahomet Ali has systematically made complaints against Datoh Kaya Dinang on every possible occasion and about a fortnight before this case Haji Mahomet Ali brought a case against the Datoh of returning to me at Serendah belonging to him. On enquiry it was found that the Penghulu had seized this ore from coolies of Haji Mahomet Ali who were found selling it in the shops at Serendah. The Penghulu

To

The Acting Assistant Resident
Kwala Lumpur

2579/86

Example of PDF image



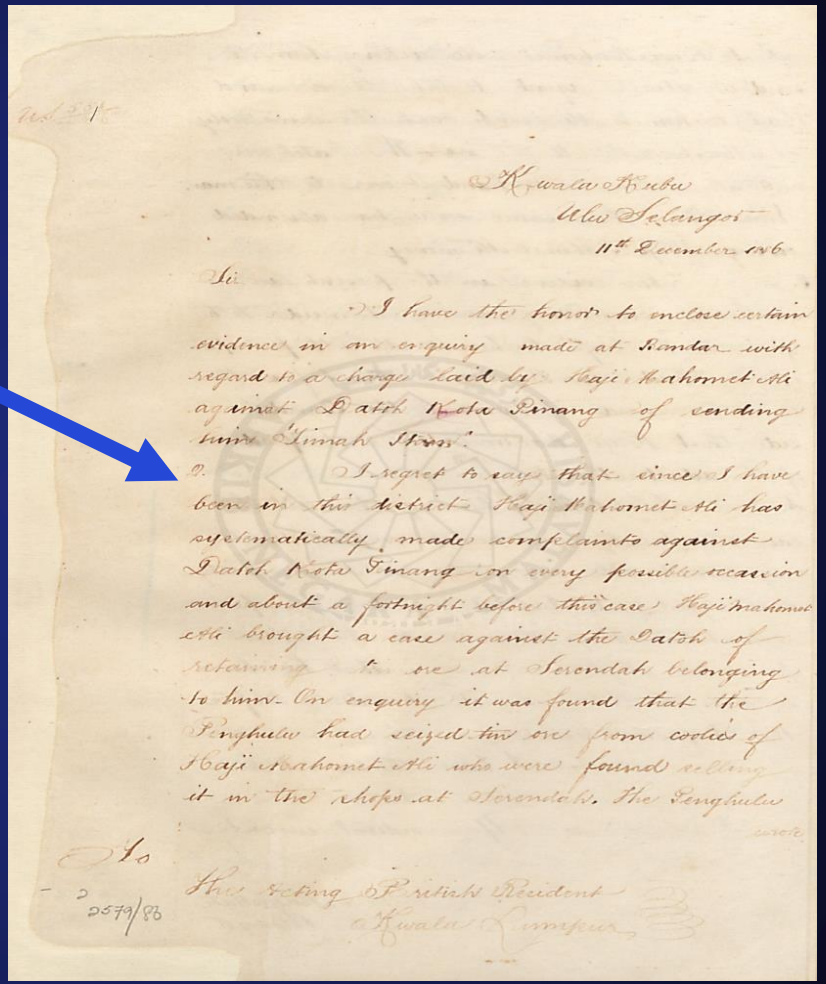
Watermark insertion procedure

The Watermark used is the NAM logo measuring 10.0cm x 10.0cm (standard size)

PDF formats - watermarked.

The watermark should be placed at the center of digital image.

WATERMARK



Kwala Kubu
Ulu Selangor
11th December 1876

Sir

I have the honor to enclose certain evidence in an enquiry made at Bandar with regard to a charge laid by Haji Mahomet Ali against Datoh Kota Dinang of sending him 'Limah Stan'.

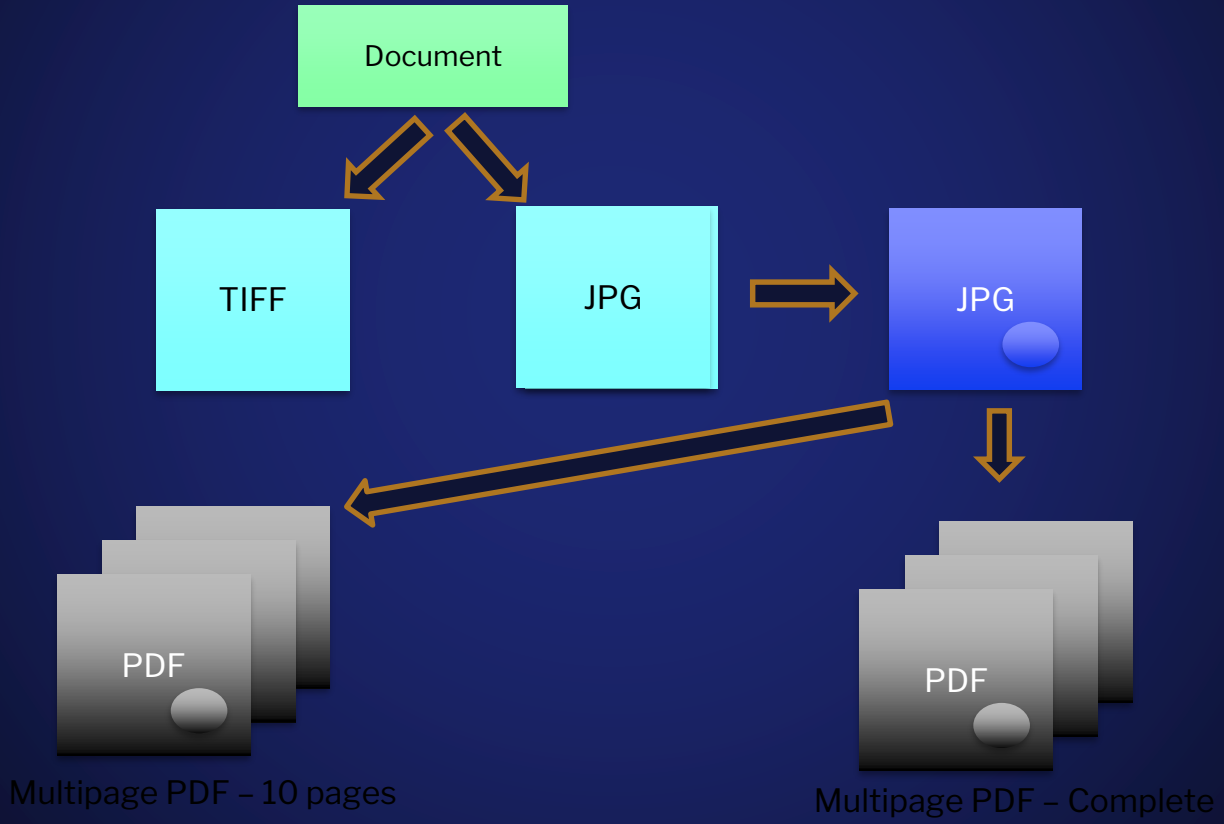
2. I regret to say that since I have been in this district Haji Mahomet Ali has systematically made complaints against Datoh Kota Dinang on every possible occasion and about a fortnight before this case Haji Mahomet Ali brought a case against the Datoh of retaining tin ore at Serendah belonging to him. On enquiry it was found that the Penghulu had seized the ore from coolies of Haji Mahomet Ali who were found selling it in the shops at Serendah. The Penghulu

To

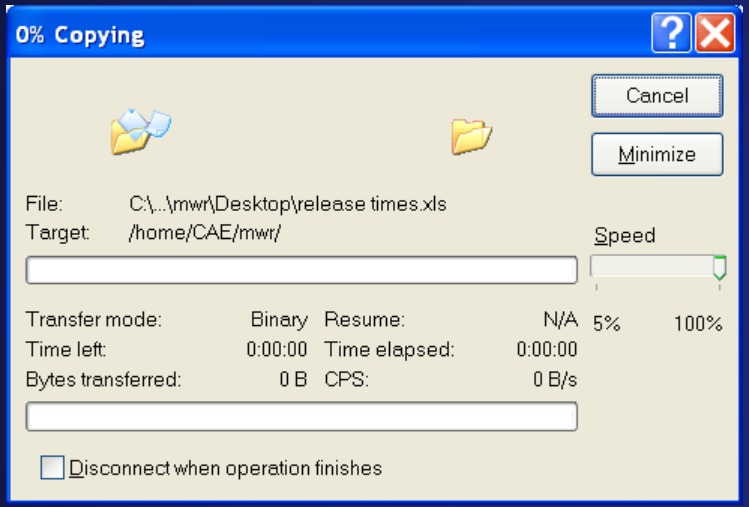
The Acting Assistant Resident
Kwala Lumpur

2579/76

FINAL FORMAT



UPLOADING PROCEDURE



Digital images that have been generated need to be uploaded / saved to storage. Metadata on the system will be linked (link) with the digital image.

After digitization (Phase 3)

- Documentation and analysis / statistics

NO.	TYPES	TOTAL OF ARCHIVES	TOTAL OF IMAGES
1	File	404,386	8,331,258
2	Publication	5,711	1,506,408
3	Microfilm	17,091	1,462,826
JUMLAH		**427,188**	**11,300,492**

Conclusion

- Digitization is the process of converting analogue records to digital form.
- Archival materials digitization requires commitment, knowledge, specifications, hardware, software and procedures
- Archives in digital form are more accessible than physical form

Thank you

DIGITAL ARCHIVES SECTION
NATIONAL ARCHIVES MALAYSIA

Q & A