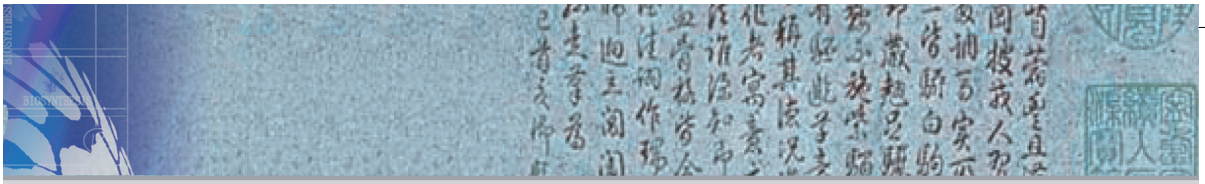
Digitization Procedures Guideline:Integrated Operation Procedures

Ya-Ping Wang, Mei-Chih Chen

Taiwan e-Learning and Digital Archives Program
Taiwan Digital Archives Expansion Project



Publisher's Preface

After the "National Digital Archives Program" was initiated in 2002, members of numerous institutional projects and request-for-proposals projects joined our team to engage in digital work that covered new categories and massive amounts of contents. The first phase of the five year project was successfully completed in 2006. The following years, the "National Digital Archives Program" and "National Science and Technology Program for e-Learning" were integrated into the "Taiwan e-Learning and Digital Archives Program (TELDAP, http://teldap.tw/)", striving to achieve the ultimate goal of "presenting Taiwan's cultural and natural diversity" as it continued to expand digital resources in various fields, and systemically promoted digital achievements in education, research and industries. TELDAP is preparing to actively collaborate with the private sector to drive growth in related industries, not only preserving important cultural assets, but also accelerating the development of a new culture in the digital age of today.

Originally named the "Content Development Division" during the first phase, we were renamed "Taiwan Digital Archives Expansion Project" (http://content.teldap.tw) as a subproject of TELDAP, and took more active measures to expand the sources of digital content, extending our reach to the collections of private institutions and even individuals. We have widely requested proposals for digitization projects related to archives, archeology, philology, geography, ethnicity, art, daily life, animals and plants, and hope to better integrate digital content with different characteristics, to develop them into fun and inspiring digital materials, and to provide them free of charge to the public for education and research; this will also help firms and public or private holding institutions to find cooperation opportunities in value-added applications. Collaboration between the "Taiwan Digital Archives Expansion Project" and other projects under the "Taiwan e-Learning and Digital Archives Program" will help speed up development of educational, research and commercial value-added applications of digital content, which will benefit the presentation of Taiwan's cultural and natural diversity, and allow people everywhere around to understand and appreciate the richness of our history and culture, as well as the beauty of our natural ecology.

While collecting and developing value-added applications of digital contents, whether it may be during the "Content Development Division" or "Taiwan Digital Archives Expansion Project" period, members of this project have continuously followed up on digital workflow related technologies used by public and private institutions and open request-for-proposals projects, and compiled a series of "Digitization Procedures Guideline Books" that introduce various international standards and provide information on digitization technologies and workflows. Since 2005, we have written 21 digitization procedures guidelines on different themes (the full text of all of the 21 books can be downloaded from the "Taiwan Digital Archives Expansion Project" website under "Virtual Library: Digitization Books"), selecting exquisite digital objects, such as ceramics, paintings, calligraphy, and string-bound books, combining the experiences of different institutional projects, and supporting them with domestic and foreign theories and practice results.

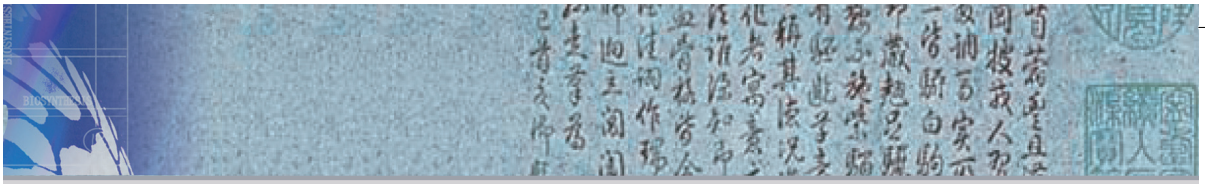
Since 2008, we have continuously revised and expanded our "Digitization Procedures

Guideline" book series, hoping to expand distribution channels so that they may be provided to even more museums, libraries, institutions and individuals for reference. Our preparations are mainly divided into revising existing guidelines for "selected objects" and compiling new guidelines on "common principles". The former refers to revising the existing 21 guidelines with a focus on introducing new digitization technologies and specifications, more practical software and hardware, and digital content protection mechanisms; we expect to revise seven books per year and complete all 21 books within three years. As for compiling guidelines on "common principles," our emphasis will be on the introduction of key concepts, such as the "life cycle" of digital information and quality control, studying multiple types of objects instead of a single type of object, and adopting common principles as the guideline framework. The so called common principles refer to project planning, integrated workflow, audiovisual data, text data, color management, outsourcing management, and digital content protection and authorization. These eight common principles are topics of which we will investigate, study and write guidelines for; we expect to publish eight guidelines in three years.

Guidelines for selected objects and guidelines on common principles in fact complement each another. Guidelines on common principles emphasize on the analysis of important topics in digitization work, guiding readers to thoroughly consider the advantages and disadvantages of digitization. Guidelines on selected objects describe practices and techniques for digitizing specific objects, helping readers to select the most suitable, most effective digitization workflow. By publishing this "Digitization Procedures Guideline" book series, we believe that we are providing institutions and individuals with the intention to engage in digitization work with a series of practical guidelines that provide an overall view, while guiding them step by step through the digital workflow. Here we must stress that the theoretical foundation of this book series is the precious experiences of institutional and request-for-proposal project teams accumulated throughout the years. These experiences allow higher quality digital content to be produced, presented and maintained with less cost, further enriching our digital archives and e-learning content. As we continue to publish our "Digitization Procedures Guideline" book series, we must give special thanks to working partners who were interviewed and colleagues who were a part of writing the guidelines, and are grateful to the scholars and specialists that reviewed and provided their advice on the book series. Finally, we hope that readers will not be reluctant to correct any mistakes or make recommendations that will help us be even better.

Taiwan e-Learning and Digital Archives Program
Taiwan Digital Archives Expansion Project · Digital Archives Sub-project of Project Integration

Project Director
February 10th, 2010



Editor's Preface

The Taiwan e-Learning and Digital Archives Program (TELDAP) is compiling a series of digitization procedures guidelines that can be provided as important reference to institutions and personnel engaged in digitization work, benefiting the enhancement of digital archive quality and consistency in Taiwan. TELDAP's approach is to first gather experiences of institutions that have obtained excellent results in digitizing specific types of objects, including ceramics, paintings and calligraphy, and ancient string-bound books, and then compiled them into individual digitization procedures guidelines; this stage since 2005. Starting in 2008, TELDAP began compiling general guidelines that are non-specific to object type, in which this guideline introduces common methods, technologies and specifications, helping to understand core concepts and methods, such as digital content establishment, data storage and management, application services, and permanent preservation.

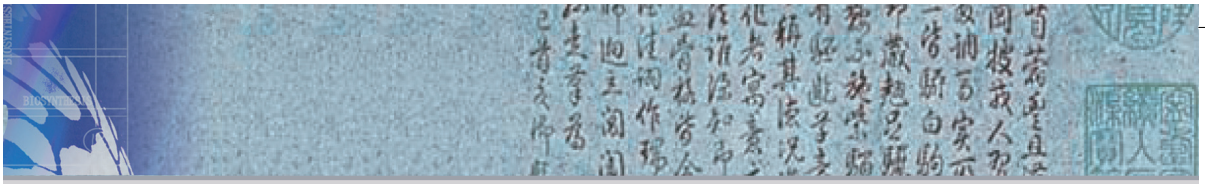
"Digitization Procedures Guideline: Integrated Operation Procedures" provides common operation procedures, methods, technologies and specifications used in system development, and guides the planning of unified and generally suitable operation procedures. The purpose of this guideline is to effectively introduce contents of digitization work in a simple manner, to explain the order of which procedures should be executed, to help understand methods and technologies, to maintain operation quality, to enhance production efficiency, and to elevate development results. Integrated operation procedures refers to methods applicable to any digital archive project, regardless of the type of object the project is digitizing, meaning that such operation procedures are suitable for digitizing maps, architectures, books, specimen, and various cultural or technology objects. This guideline introduces matters involved in the digitization process and long-term management, from preliminary procedures, object digitization, cataloging, file storage, applications, preservation planning, to administrative management, and provides clear and simple rules for planning human resources, timetable and budgets, making it easy to for operators to understand, which ensures system development and application quality, and facilitates the effective establishment and conservation of digital archives. TELDAP divides objects into sixteen categories, while the National Repository of Cultural Heritage divides objects into fifteen categories; each has several categories that the other doesn't. With the appearance of new digital archive projects, e.g. military vessels, movies, and intangible cultural heritages of the Hakka and indigenous people, including any project into an existing category will result in the continuous expansion of guidelines for individual objects. Therefore, it is extremely important to set forth common operation procedures that are non-specific to any object category, as well as become familiar with best practices.

The contents of this guideline possess both width and depth, not only does it pass on experiences that institutions in the National Digital Archives Program accumulated over the years, but also introduces practices of the National Repository of Cultural Heritage, as well as analyzes standards used by international organizations and advanced countries, so that the integrated operation procedures it proposes can be put to the test. The contents of this book were thoroughly planned out, chapter one describes the background and purpose,

hoping to help the public gain a better understanding of digital archive establishment, application and permanent preservation, so that they understand its importance. Chapter two explains the importance of operation procedures, emphasizing on its relationship with quality management; to explain the completeness of integrated operation procedures, chapter two introduces the concept of digital material life cycle, and describes its four main stages: data creation, data management, data preservation, and data provision. Chapter three discusses operation procedures of digital archives, starting from planning operation procedures, operation procedures for different object types, comparison of operation procedures, and main tasks of digital content establishment. Chapter four describes specific contents of integrated operation procedures, including main steps of integrated operation procedures, the relationship between integrated operation procedures and OAIS, comparison of born-digital and digitized materials, selecting digital image and metadata formats, application systems of digital images and metadata, and permanent preservation of digital archives, providing detailed descriptions and clear standards that benefit actual implementation. Finally, in the conclusion and future prospects this guideline once again emphasizes on the necessity of integrated operation procedures, and explains that the procedures described are consistent with OAIS specifications. Definitions and explanations of technical terms are provided in glossary to help build a correct understanding. Hence, this guideline's contents give consideration to both theory and practice, methods and technologies, introduces academic and practical experiences from Taiwan or other countries, and uses them as a basis for establishing integrated operation procedures.

蔡曉慧

February 24th, 2010



Contents

2 Publisher's Preface

4 Editor's Preface

8 ONE. Introduction

I. Foreword 009

II. Contents of this Book 010

12 TWO. The Importance of Operation Procedures

I. Operation Procedures and Quality Management 013

II. Operation Procedures and Digital Information Life Cycle 015

1. Data Creation

2. Data Management

3. Data Preservation

4. Data Provision

21 THREE. Reviews of Operation Procedures in Digital Archiving

I. How to plan operation procedures 022

1. Operation Procedure Considerations

2. Case Considerations

3. Management Considerations

II. Operation procedures for different types of objects 024

1. Sixteen Thematic Objects of Digital Archives

2. Operation Procedures of the Six Themes of Digital Archives

III. Comparison of operation procedures for different objects 028

1. Common operation procedures for different objects

2. Different operation procedures for different objects

IV. Main tasks of digital archive content establishment 034

1. Scanning
2. Photography
3. Audio/Video Recording
4. Metadata Establishment

42 FOUR. Integrated Operation Procedures in Digital Archiving

I. Integrated Operation Procedures in Digital Archiving……043
1. Main steps of integrated operation procedures
2. Relationship between integrated operation procedures and OAIS

II. Comparison of Born-Digital and Digitized Materials……050
1. Born-Digital and Digitized Objects
2. Comparison of Operation Procedures for Born-Digital and Digitized Objects

III. Selecting Formats for Digital Images and Metadata……054
1. Selecting Image Formats
2. Selecting Audiovisual Formats
3. Selecting Metadata Formats

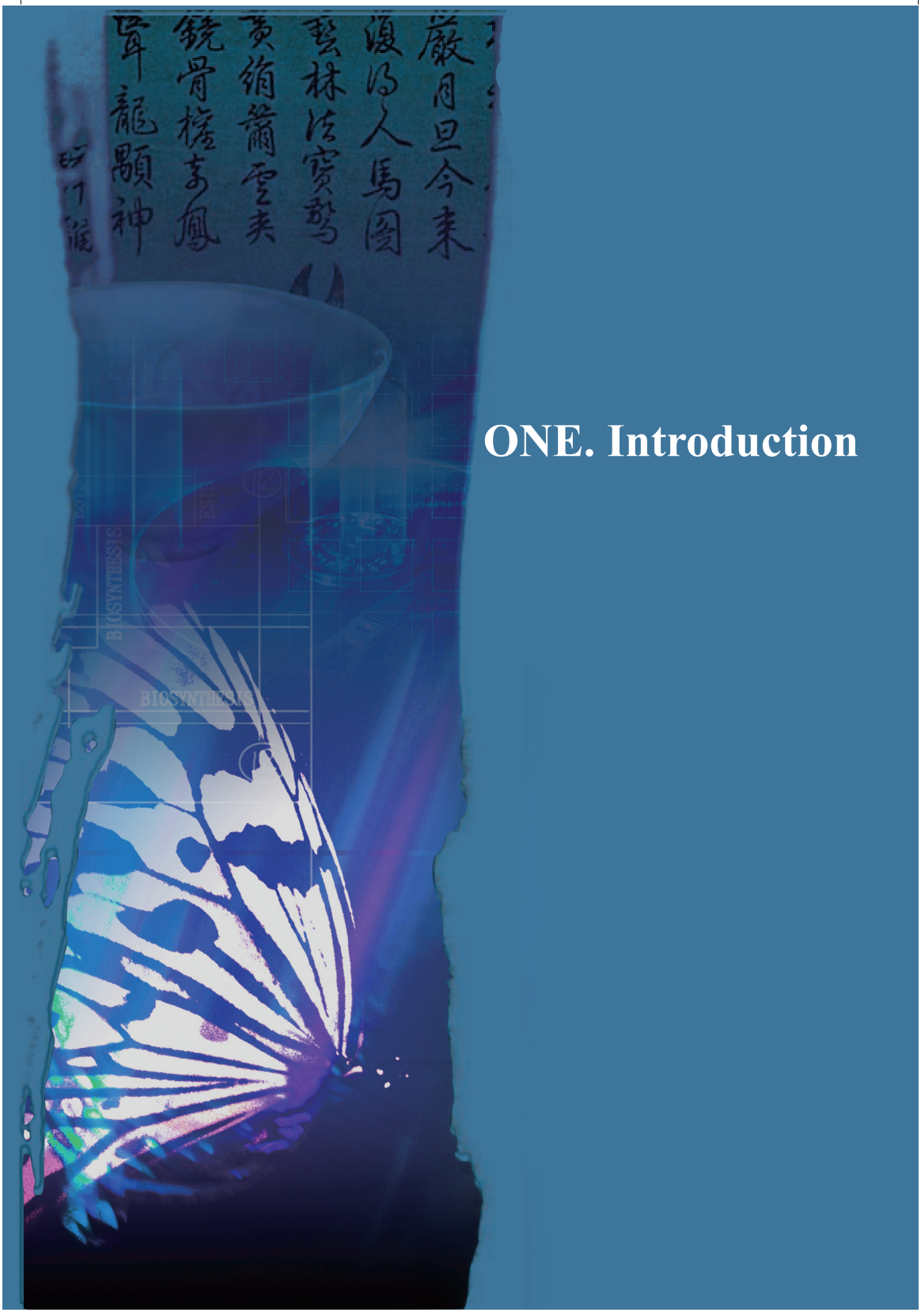
IV. Application Systems for Digital Images and Metadata……068
1. System Establishment Evaluation
2. System Development and Establishment
3. Service, Innovative and Value-added Applications

V. Permanent Preservation of Digital Archives……071
1. Meeting Requirements for Permanent Preservation
2. OAIS and Permanent Preservation

76 FIVE. Conclusions

78 Glossary

83 References



ONE. Introduction

I. Foreword

In recent years, countries around the world have been actively promoting digital archives, using information technology to turn important cultural assets into digital archival resources. In order to enhance the quality of these resources, it is necessary to adopt systematic procedures and establish a set of standards for the complicated tasks that are involved. The establishment of a rich digital repository is like constructing a building, and requires specific procedures and steps, gradually building it one brick at a time. For operation procedures to be complete, personnel should understand standards and methods related to the entire process, reducing human mistakes in system establishment, and achieving the goal of reducing cost and enhancing quality.

The "Taiwan e-Learning and Digital Archives Program" has been implemented since 2002 (the first phase was the "National Digital Archives Program"), fellow colleagues of TELDAP have investigated and recorded digitization procedures and technologies used by public/private organizations and open request-for-proposal projects, and combined them with international standards and digitization technologies to compile a series of "Digitization Procedures Guidelines." Quite a few organizations have established their own set of operation procedures, starting out from an early exploration stage and gradually developing mature operation procedures; thanks to them, we have reference standards for digitizing various types of collections. Operation procedures may vary along with different types of collections due to their significantly different properties. Therefore, the establishment of integrated operation procedures will not only help project units that are implementing digital archive projects for the first time to effectively plan their work, but also serve as reference for project units to use when making decisions that concern digital archives.

The concept of operation procedure originated from the concept of business process – "The process of which participants in a company follow a set of rules, which were defined to achieve an overall or specific goal, and cooperate with other participants in or outside of the company in establishing documents, information or tasks that create value."[1] An operation procedure refers to the entire process of a task from its beginning to end, which may involve multiple departments and coordination of numerous links. In short, an operation procedure is a process that turns an input into output, each step along the way corresponds to specific personnel responsible for specific tasks.

Therefore, integrated operation procedures common operation procedures, methods, technologies and specifications used in system development, and guides the planning of unified and generally suitable operation procedures. Its purpose is to effectively introduce contents of digitization work in a simple manner, to explain the order of which procedures should be executed, to help understand methods and technologies, to

[1] Su Han-Liang, "Design and Construction of Operation Procedure Management Systems", National Taiwan niversity Graduate Institute of Civil Engineering Master Thesis, 2001.

maintain operationquality, to enhance production efficiency, and to elevate development results. Integrated operation procedures refers to methods applicable to any digital archive project, regardless of the type of object the project is digitizing, meaning that such operation procedures are suitable for digitizing maps, architectures, books, specimen, and various cultural or technology objects. This guideline introduces matters involved in the digitization process and long-term management, from preliminary procedures, object digitization, cataloging, file storage, applications, preservation planning, to administrative management, and provides clear and simple rules for planning human resources, timetable and budgets, making it easy to for operators to understand, which ensures system development and application quality, and facilitates the effective establishment and conservation of digital archives. By following operation procedures of digitization, digital archiving work will more effectively and rapidly produce results.

Digitization procedures guidelines compiled by TELDAP in the past focused on specific types of objects in the program's sixteen categories, including painting and calligraphy, ceramics, ancient string-bound books, archives, seal and imprint, mineral specimens, and positives. Each guideline provided detailed digitization procedures for different types of objects so that institutions could use them as reference for digital archiving. Considering that digital archiving encompasses a vast scope of fields and that institutions have accumulated a wealth of experience over the years, effectively introducing the fundamentals of digital archiving to the general public has become a matter of great urgency.

Besides discussing operation procedures of the six object types, this book attempts to breakthrough barriers between each category, and take on a general, integrated perspective, using "integrated" operation procedures to help institutions or individuals implementing digital archive projects to gain a clear understanding of digital archiving from start to end. Furthermore, this book also introduces concepts of "digital information life cycle" and quality management as principles for examining operation procedures, helping project units to be more efficient and elevate work quality.

II. Contents of this Book

Chapter ONE　Introduction:

This chapter explains the purpose for compiling "Digitization Procedures Guideline: Integrated Operation Procedures," and hopes that this book can help institutions or individuals engaged in digital archiving work to grasp core concepts and technologies when facing collections of any type, to effectively carry out operation procedures, and to follow related methods and specifications to establish a system with high quality.

Chapter TWO　The Importance of Operation Procedures:

This chapter describes how it is important that operation procedures and quality management complement each other, and uses the concept of digital information life cycle as a principle for examining operation procedures.

Chapter THREE Reviews of Operation Procedures in Digital Archiving

This chapter describes how to plan operation procedures, and uses the sixteen thematic objects of TELDAP as an example, introducing main digitization procedures of each object, and discusses the differences and similarities between the procedures. This chapter also briefly describes main tasks of establishing digital content.

Chapter FOUR Integrated Operation Procedures in Digital Archiving:

Using the comparison of digitization procedures for different object types in the previous chapter, this chapter draws a flowchart of integrated operation procedures, and uses the flowchart to explain main steps, as well as how to plan and select specifications for digital images, metadata and system establishment.

Chapter FIVE Conclusions

This book collects and shares experience of integrated operation procedures in digital archiving it can be provided as reference to help more institutions engaged in digitization work, and contribute to digital archives in Taiwan.

TWO. The Importance of Operation Procedures

I. Operation Procedures and Quality Management

The concept of operation procedures originated in the business management system, which on one hand stipulated employee's responsibilities, and on the other hand provided operation procedures for employees to follow when executing tasks. Digital archiving is a complicated process, but can be executed in systematic steps and methods once operation procedures are established. Effective control over procedures can improve operation quality; a robust mechanism can reduce cost; simplified operation procedures can enhance productivity.

1. Planning Operation Procedures

Digital archiving turns valuable collections into digital images for the convenience of long-term preservation. Besides truly presenting the appearance of collections, these digital files can also be provided for future value-added applications, which is why quality requirements must increased. To effectively reach standards that satisfy quality requirements, a set of operation procedures and management is necessary to guarantee the quality of digitization results. Whether or not a digital archiving mechanism is mature can be determined by whether or not operation procedures are suitable and smooth, or if there are Standard Operation Procedures (SOP), which maintain consistency of quality and enhance work efficiency. Standard operation procedures are an important link in quality control; besides ensuring stable quality, they also reduce the occurrence of mistakes.

Whether its operation procedures or even more precise standard operation procedures, it is necessary to first verify the target of standardization. Steps, procedures, methods or problems that frequently occur especially need standardization. Operation procedures must go through continuous additions, amendments, abolishment and establishment, check points must be added to detailed descriptions or complicated procedures to maintain quality of the overall operation and boost results and progress. The establishment of standard operation procedures also helps reduce cost on human and material resources.

2. The Importance of Quality Management

"Quality Management" is not only focused on quality requirements, but also means to achieve it. Different components of quality management, including "quality inspection", "quality control" and "quality assurance", all pursue the ultimate objective – work results. The establishment of operation procedures benefits the implementation of quality management; the purpose of quality management is to inspect the accuracy as well as the efficiency of digitization work. Regarding quality in the digitization process, it takes professional equipment and competencies in digital production to complete the series of stages and procedures; pay attention to details in complicated operation

procedures, check points for each step must be established to ensure the integrity of the procedure; continuous communication and planning are required to maintain consistency and stability.

Digital archiving requires a great amount of technologies and human resources, which is why work efficiency and quality is sure to be elevated by providing simple explanations of the entire structure and implementation steps, as well as establishing standard operations procedures. However, operation procedures implemented by domestic institutions are still individually developed and lack a common standard, resulting in varying quality of work results.

3. Digital Archiving Experiences of Taiwan and Other Countries

Learning from the experiences of other countries is also an approach to work planning, the status quo of a few digital archive projects implemented by renowned foreign institutions are provided in the following section as examples: The British Museum developed the COMPASS archiving system with pilot 3D projects to provide a simple method for the general public to access information.[2]

The Metropolitan Museum of Art[3] , Harvard Peabody Museum, and Harvard University Art Museums: Fogg Art Museum and Arthur M. Sackler Museum all use The Museum System (TMS), a museum collection management system developed by Gallery systems. The system involves operation procedures that categorize collections, and the standardization of data input forms. For instance, each object in TMS could have thousands of metadata fields; besides basic cataloging information, these fields are for research departments, archive departments, cataloging departments, exhibition departments and conservation departments to use for management purposes; some fields are for management information if the object is transferred between different departments.[4]

The National Museum of Ethnology in Osaka, Japan constructed an image database with a cross reference index search mechanism as its information system, and proposed a new database management concept using online forums to manage cultural resources.[5]

The Hermitage Museum signed a cooperation agreement with IBM to not only construct the museum's website, but also develop digital imaging, data processing and database functions; their partner relationship was successful and together they achieved

[2] David Jillings (2003), "From Two to Three Dimensions- Building on Experience of 2D Digitization of Prepare For the 3D Digital World", International Conference of Museum Digitization Antiquities, Painting and Calligraphy, Taipei.

[3] Hwai-ling Yeh-Lewis (2003), "Then and Now: Manual Cataloguing system vs. Digitization System", International Conference of Museum Digitization Antiquities, Painting and Calligraphy, Taipei.

[4] Lin Yu-Yun, Lin Kuo-Ping, Chen Shu-Chun, Lin Shih-Chieh, Li Te-Tsai, "Alliance and Collaboration: New Energy of the 'National Digital Archives Program", "Museology Quarterly" 20:3, 2006/07, pages 99-115.

even more goals of digital archiving technology.[6]

The Art Institute of Chicago has a set of illustrated operation procedures to explain steps required for creating high resolution images of artworks, and how to utilize equipment and technologies to produce digital files that are suitable for archive management systems.

C2RMF has used 3D technology and EROS, which is for art preservation and research, for many years.

In Taiwan, the National Palace Museum, Academia Sinica and National Museum of History were among the first of institutions to implement digital archive projects, and each has gradually established a set of operation procedures based on their digitization results accumulated over the years. Academia Sinica compiled a Digital Archive Technology Collection for the National Digital Archives Program, providing numerous related technologies and application methods; the Taiwan Digital Archives Expansion Project compiled a series of digitization procedures guidelines and illustrated manuals; and major museums, including the National Palace Museum, have all planned their own set of operation procedures. Operation procedures of the institutions mentioned above for digital archiving can roughly be divided into preliminary operations, specification selection, object digitization, collection management and exhibition system, and value-added applications; these main procedures can be drawn into a flowchart. Flowcharts are commonly used as a management tool, showing all main tasks in an illustration helps clarify the coherence and sequence of individual tasks, and makes it easy to find the key to problems, which allows work to be more consistent and helps improve the quality of results; improvements can also be effectively made to parts that need to be modified. The more complete a flowchart is, the easier it makes for personnel to understand operation procedures, reducing unnecessary mistakes and improving quality.

Not getting into the differences of operation procedures in different institutions yet, details of each operation procedure should give consideration to digital archiving as a whole; only by defining the key tasks of each step can adjustments be made at any time according to actual conditions. In order to effectively control digital quality, it is necessary to fully understand each step and method, effectively carry out and manage procedures, and provide training in coordination.

II. Operation Procedures and Digital Information Life Cycle

Like all living beings, which go through birth, growth, maturity and death, information, regardless of its type, goes through a similar life cycle – creation, processing, storage and reutilization. Living beings evolve as they progress through

[5] Masatoshi Kubo (2003), "Development of the Image Database at the National Museum of Ethnology: Toward Forum-Type Database Management", International Conference of Museum Digitization Antiquities, Painting and Calligraphy, Taipei.

stages of their life cycle. The creation, storage, utilization, destruction, and dissemination of information achieve the diverse influence of information. Operation procedures in digital archiving also follow the digital information life cycle, from non-existent to existent, progressing one stage at a time. To establish mature operation procedures, steps can be examined using stages of the digital information life cycle.

The digital information life cycle is mainly divided into the following stages: data creation, data management, data preservation and data provision. We can try to match operation procedures in digital archiving to stages of the digital information life cycle to verify the completeness of operation procedures.

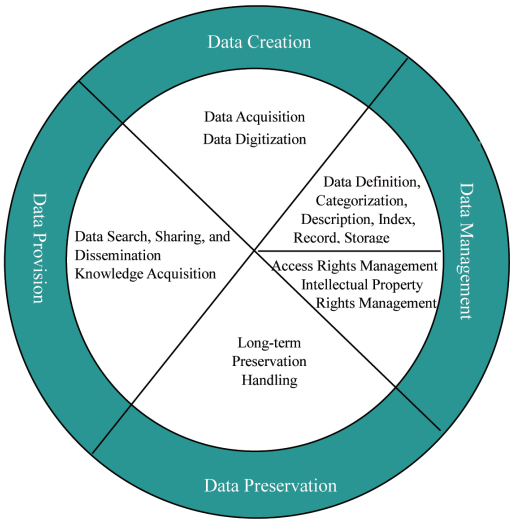


Fig.2-1 Digital Information Life Cycle

1. Data Creation

This stage is when files and information are created, and mainly involves the collection, production, acceptance and acquisition of information on digital objects. Main work contents of this stage include gathering and digitizing data. Data creation is the result of a series of thoughts and decisions, so its structural characteristics and contents directly affect the development and management of data. From the perspective of digitization procedures, the data creation stage consists of preliminary procedures, digital file establishment, quality inspection, and post-processing.

(1) Preliminary Procedures

Before engaging in digitization, establish specifications based on usage purpose and future applicability to satisfy digital archiving requirements. Preliminary procedures include sorting out archival collections, making an inventory of objects, establishing

digitization standards, selecting a digitization method, planning human resources, borrowing objects, and preparing the digitization environment.

(2) Digital File Establishment:

This involves digitization procedures that digitize physical objects. The basic principle of digitization is to not harm the physical object, while truly presenting the object in a digital form. Tasks of digital file establishment include digitization procedures related to photography or scanning.

(3) Quality Inspection:

After digital files are created, they must be carefully examined for unwanted conditions, such as color shift or slanted images. It is imperative that high resolution, high quality images are produced. Quality inspection requires the use of various software and instruments for digital proofing and color calibration. Born-digital content saves time in digital file establishment, but still need to go through quality inspection and might require some modifications.

(4) Post-processing:

Digital files that have been inspected can enter post-processing procedures, including retouching, stitching for large files, file naming, file conversion, image protection, proofing and output.

These procedures all belong to the "Data Creation" stage in the digital information life cycle; their main purpose is to extract and digitize data.

2. Data Management

Data management mainly refers to managing metadata of archival collections; operations include metadata creation, categorization, definition, documents, storage, access rights and intellectual property rights management[7]. The management method adopted should be based on the data's structure, format, compression rate, coding and purpose when it was created. Digitization procedures that belong to the data management stage include:

(1) Metadata Requirements Analysis:

This is the foundation for metadata establishment, and analyzes metadata requirements, as well as how to establish, categorize and define metadata.

(2) Metadata Cataloging:

After cataloging standards are established, metadata recording, cataloging and proofreading, as well as additions and modifications all belong to this procedure, which ensures that each digital object is accurately described.

(3) System Development and Establishment:

[7] Tsai Yung-Cheng, Huang Kuo-Lun, Chiu Chih-Yi et al, "Introduction to Digital Archives Technology", Taipei City: NTU Publishing Center, November 2007, page 82..

In order for data to be effectively managed, the establishment of a database and search system must consider how it will process massive amounts of data. Therefore, the file structure, format and usage purpose should be kept in mind during requirements analysis and design. This process also includes the establishment of an access rights management system to be responsible for maintaining user accounts, as well as providing user rights settings, e.g. system administration rights, data cataloging rights, and data search rights. In addition, records of user operations, e.g. when the archive system was used, which data entry was modified, and what key words were used for searches, are all references for system evaluations, and are a link in the data management system.

(4) Testing for Database and Searching System:

Continuous system tests are required for the maintenance and adjustment of data specifications and formats.

The data management stage builds the foundation for long-term preservation of data, and is also an important process in quality management.

3. Data Preservation

Data "storage" is generally considered to play a passive role in the digital information life cycle, because digital data can become corrupted or not readable after it changes storage media or format. Therefore, data preservation needs to consider storage format, storage environment, data migration, remote backup, and other technical aspects (including storage media and platform)[8]. Digitization procedures that belong to the "Data Preservation" stage include:

(1) Digital File Storage:

This mainly refers to the storage of the source digital file and files that have completed post-processing. Digital materials of this procedure are imported into the database.

(2) Backup Copy and Remote Backup:

"Backup copy" and "remote backup" are both important concepts of data storage, but on different levels. Backup copies are made on a regular basis and carefully preserved. "Remote backup" is when the same data is stored in two different locations, so when one location encounters problems, the other location will operate in its place to reduce loss.

(3) Database:

A database is simply a place where data is stored. It also refers to a system that has

[8] Hsieh Yi-Keng, Tsai Shun-Tzu, "Establishment of Operation Procedures in Digital Archiving from the Perspective of Digital Information Life Cycle", Taipei: Digital Archive Quality Management Seminar, 2007.

a specific format, organization and files stored in it. Therefore, an effective database management system is an important link in data preservation.

Therefore, data "preservation" aims to protect data from being damaged by changes in time, space or information equipment; it strives to achieve long-term preservation of data. Data preservation is a series of strategies and procedures, including:

(1) Refresh: Refreshes, or duplicates, media. However, this strategy is not yet able to guarantee that the data can be searched or processed in the future.

(2) Migration: Converts data formats, structure or standards; converts configuration settings of software/hardware into different settings or a new generation of computer technology.

(3) Emulation: Emulation using software or hardware can predict how the old system will function in a future unknown system, allowing original programs of digital objects to operate.

(4) Standardization: Mainly focuses on the standard format of data, and must determine whether or not existing standards are still applicable and can also be preserved. This is to prevent digital files from being damaged, lost or depreciated because of new formats.

(5) Technology Preservation: Data preservation also involves preservation of the original application, operating system and hardware platform, emphasizing that the original appearance of digital objects can only be reproduced by preserving its original operation environment[9].

In contrast to digital archiving procedures, database maintenance, application system establishment, and even making backup copies (remote backup) all need to consider long-term preservation strategies. Develop a preservation method for the digital archive project based on the management policy.

4. Data Provision

Data provision is the stage where files and information are used and shared; all other stages of the digital information life cycle serve to support this stage. In this stage digitization units attempt to make the data they created easier to use, manage and exchange, so it can be accessed and shared; each entry of data follows a clear context and is a part of the entire management system structure. Operation procedures that belong to this stage include: search system, website services, licensing and innovation services.

(1) Search System:

The most important purpose of digital archiving, aside from data creation and

[9] Chang Huai-Wen, "Metadata and Digital Archive Long-term Preservation and Access – An Introduction to Preservation Metadata", "National Digital Archives Program Newsletter", Volume 3 Issue 1, January 1st, 2004.

acquisition, is to effectively manage these data and allow data to be rapidly found for further application. To satisfy such a requirement, diverse search and access methods should be designed, and data browsing and inquiry functions should be provided; the "Union Catalog of Digital Archives" provides such functions.

(2) Website Services:

Collection analysis and data integration are carried out for promotion and exhibition purposes, strengthening the connection between objects, establishing a knowledge and context structure, and allowing digital content to be widely used, which expands benefits of the data itself; the "Knowledge Web of Taiwan's Diversity" is a website with such properties.

(3) Licensing Services:

Advancements in technology have made it extremely convenient to transfer digital information, so the establishment of a well-rounded licensing mechanism will expand the scope of applications and circulation. However, numerous intellectual property rights issues have been created because digital information is easy to transfer, modify and duplicate, including ownership of intellectual property rights, licensing issues, and establishing content protection mechanisms.

(4) Innovation Services:

Features and concepts of digitized materials can be used as elements in designs, converted, and then incorporated into products that possess market value. These innovative and creative products can be utilized to form a circulating market to create economic value. More and more successful examples can now be found, including the cooperation of National Palace Museum and Franz Corporation has obtained brilliant achievements in Europe and America, as well as the popularity of images licensed for use on LCD monitors in foreign markets.

Data creation and data provision are equally important. Use the most convenient method to extract and create complete digital information, and then plan suitable management mechanisms, information application scenarios, user interfaces, and appropriate applications for provision and dissemination.

THREE. Reviews of Operation Procedures in Digital Archiving

I. How to Plan Operation Procedures

A great variety of operation procedures has been produced in response to the multitudes of object formats in archival collections. This chapter divides considerations when planning operation procedures into three categories for analysis and discussion: operation procedure considerations, case considerations and management considerations

1. Operation Procedure Considerations

(1) Steps: Clearly ordered steps are an important link in operation procedures. Therefore, before planning an operation procedure, one must first understand the order that tasks must be completed, in which is order of digitization procedures can be roughly divided into preliminary procedures, digitization procedures, metadata and database establishment, and value-added applications. Using preliminary procedures as an example, most collections were acquired during early periods, as time passed by, different types of registration and management data were produced for different needs and purposes of the department that accessed it, as a result, the condition of most physical objects is no longer consistent with the source data. Therefore, before engaging in digitization procedures, first examine the condition of the physical collection and make a new list of objects to be digitized, and then arrange the order of which objects will be digitized.

(2) Methods and Technologies: Select suitable digitization methods and equipments based on characteristics of the collection. In the past, the main consideration of selecting digitization methods and technologies was work progress, but overlooked potential damage to the physical collection during the digitization process. Today, more and more emphasis is being put on relic conservation, and following technology advancements, numerous digitization equipments that do not damage physical objects while providing excellent functions have been developed, e.g. cold light and high-end professional multi-purpose scanners. Such digitization equipments significantly reduce damage caused by pressure, light, heat and centrifugal force, while achieving printing and conservation requirements.

2. Case Considerations

Each digitization project is unique in its own way; it is very hard for two projects to be completely identical. In the planning process of each digitization project, besides referring to or borrowing from existing methods, supplements, corrections and improvements are also required so that the project meets historical context and logic requirements. Furthermore, the correctness of function requirements should also be verified. When implementing a digital archive project, give consideration to object selection, purpose, human resources, digitization method, standard selection, management and preservation related issues. When consideration the purpose of digital objects, first consider future applications of digital files, and then decide on the format

of digital files. If digital files will be used in publications, then the resolution should be at least 300dpi for the image to be the same size (1:1) as the original object; therefore, the minimum requirement is RGB 24bit or above (full color), 400~600dpi TIFF (not compressed), and a file size of 100 MB (A4) or above. If the digital image will only be used for online browsing, the smaller the file is the easier it will be for the image to be browsed online; therefore, the recommended specification is RGB 24bit (full color), 150dpi JPEG (compression rate: 50%), and file size of 230kb (A4) or less; the file size can be increased if larger bandwidth is provided. However, TIFF files can be converted into JPEG files, so the most important part of digitization is to create and preserve high quality uncompressed files.

3. Management Considerations

When implementing digital archive projects, numerous factors that affect operation procedure planning might have an influence on work quality, including collection type, personnel, digitization method, work environment, and even policy and funds. To maintain the quality of digitization results, give consideration to the following management related aspects when planning digitization procedures:

(1) Human Resources:

When planning operation procedures in digital archiving, assigning personnel to different tasks based on their competencies, deployment, relationships and problem solving methods can effectively increase the accuracy of cost, time and resource consumption estimates. To achieve even better results, clearly define performance and control standards, assign duties and responsibilities, and make flexible personnel adjustments when necessary. For example, personnel responsible for scanning should be competent in operating professional scanners, and personnel responsible for photographing objects should not only be competent in lighting, but also photography.

(2) Budgets:

Plan the scale of budgets that can be invested in the digital archive project and allocate a budget, estimate the amount of resources and cost required for individual activities of digital archiving, and effectively distribute funds to each link of work. At the same time, verify the amount of resources (human resources, equipment, etc.) required for each task. For example, begin making estimates from the smallest link, including which types of equipment will be used and cost required, to accurately control cost requirements that will achieve the target quality.

(3) Time Schedules:

Digital archiving work is periodical, so after arranging the order of objects for digitization and completing budget planning, consider the progress and time each task will be completed, what parts should be completed before what time to begin with the next phase of work. When planning the timetable, remember to leave some time after

each task for evaluating results and making corrections, so that the next phase of work can begin with proper adjustments already made.

Thorough planning and standardized operation procedures ensure that tasks are completed according to schedule, achieving stability of operation procedures and quality, and enhancing work efficiency and performance.

II. Operation Procedures for Different Types of Objects

1. Sixteen Thematic Objects of Digital Archives

The Taiwan e-Learning and Digital Archives Program organized various thematic groups based on characteristics of project collections during the Program's first phase of implementation. The thematic groups met irregularly, but the conveners of each thematic group met on a regular basis to establish operation procedures and work principles for data establishment, establishing cooperation, division of labor, and coordination mechanisms to help units successfully implement projects. Up to the end of the first phase, there were sixteen thematic groups, including: zoology, botany, geology, anthropology, files and documents, antiquities, painting and calligraphy, maps and remote sensing images, rubbings, rare and ancient books, archeology, news, architecture, audiovisual, language, and Chinese full text.

During project implementation, thematic groups requested that project units recommend scholars or professionals to hold the position of convener. Besides promoting digital archive projects, the convener is also responsible for negotiating data exchange and integration between different projects. The sixteen thematic groups and their corresponding collections are as shown in Table 3-1.

Collections under different themes possess unique characteristics; hence it is hard for

Table 3-1 Thematic Groups and Corresponding Collections

Thematic Group	Collection Content and Form
Zoology	Collections mainly consist of animal specimens found in Taiwan, and includes type specimen, specimen catalogs, labels, specimen photos, habitat photos, basic explanations, distribution data, ecological films, and literature.
Botany	The main digitization target is data on Taiwan's vegetation, including type specimen, labels, images, habitat and pattern photos, basic explanations, distribution data, Flora of Taiwan and related literature.
Geology	Geological collections include minerals, rocks, fossils of vertebrates, invertebrates, micro-organisms, and plants, and drilling core samples.
Anthropology	Mainly consists of field data collected by researchers on daily life, villages and settlements, religious activities, and physique of ethnic groups, including photos, documents, relics (specimen), audio records, video tapes, maps, and references (books).

Thematic Group	Collection Content and Form
Files and Documents	Mainly consists of documents from the Ming and Ching Dynasties, early Minguo, the Nationalist Government Period, Japanese Colonial Rule, and after Taiwan's Restoration, including historical materials, official government documents and private records; in terms of form, collections consist of books, photos, manuscript, and images.
Antiquities	Collections of the antiquities thematic group cross over millenniums, including bronze ware, jade ware, jade, ceramics, lacquer ware, bamboo and wood, seals, sculptures, and locksmith tools.
Painting and Calligraphy	Calligraphy, books of stone rubbings, paintings, and embroideries dated between the Song and Ching Dynasties, as well as modern engravings, Chinese paintings, calligraphy, Western paintings and photography.
Maps and Remote Sensing Images	Historical maps, remote sensing images.
Rubbings	Oracle rubbings, bronze ware rubbings, rubbings of Han Dynasty stone relief, Buddhist rubbings, Liao-Chin-Yuan rubbings, Han wooden slips, picture brick rubbings, rubbings of tomb tablets, and old stone rubbings of Taiwan.
Rare and Ancient Books	Mainly consists of rare and ancient books from the Song to Ching Dynasties and early Minguo, contents include books published during Japanese Colonial Rule, rare books, folk literature, and books of ethnic minorities.
Archeology	Mainly consists of archeological data, collections include archeological relics, field excavation notes and dairies, drawings, and site photos.
News	Mainly consists of electronic newsletters, newspapers, periodicals, and news clips; in terms of form, collections include text, images, photos and audiovisual files of print media and TV media.
Architecture	Mainly consists of "architecture" related structures and building data; in terms of form, collections include text, images, photos and audiovisual data.
Audiovisual	Mainly consists of audiovisual data in the form of a disc, CD, radio broadcast, movie, TV program, or video tape.
Language	Mainly collects and preserves existing corpus, including sign language; collections are mainly in the form of sounds and text.
Chinese full text	Contents include the Twenty-five Histories, Thirteen Classic Works, novels and theater, and Buddhist scriptures; collections are mainly in the form of text.

operation procedures to be entirely the same. In preliminary procedures, some paintings and calligraphy have frames that need to be taken off, geological specimen need to be cleaned and repaired, and documents need to be mounted before following digitization procedures can be carried out.

2. Operation Procedures of the Six Themes of Digital Archives

In the first phase of TELDAP, each project decides its own target, quantity, procedures and timetable, while the thematic group provides recommendations. "Thematic groups" aid cross-agency cooperation, and establish common work specifications. After entering the second phase, the overall objective became "integration and innovation," "deepen value added" and "effective dissemination." Therefore, digital archive projects are beginning to emphasizing on deepening value-added applications of digital content produced in the first phase, as well as planning integrated knowledge structures and sustainable operation guidelines. In this phase "thematic groups" are the trunk for cross-project integration, combining different themes to form a knowledge base that integrates natural history and human history.

To achieve the overall objective, specialists and scholars were invited to organize six new thematic groups on the basis of the sixteen thematic groups in the first phase, the six thematic groups are "Biology and Nature", "Life and Culture", "Languages and Multimedia", "Art and Images", "Maps and Architecture" and "Documents and Files". This six themes are the foundation for holding "Thematic Group Convener Meetings" and "Thematic Group Discussions", in which technical issues in digital archiving are addressed and the most suitable solution is discussed; special projects are implemented to study digital resources outside the scope of existing projects, and to establish a knowledge network, striving to "preserve Taiwan's diversity."

Operation procedures vary with each project and thematic group; a few thematic groups are used as examples and described below:

Table 3-2 Six Themes and Corresponding Thematic Groups

Six Themes	Contents
Biology and Nature	Zoology, Botany, Geology
Life and Culture	Anthropology, Archeology
Languages and Multimedia	Language, Audiovisual, News
Art and Images	Antiquities, Painting and Calligraphy
Maps and Architecture	Maps and Remote Sensing Images, Architecture
Documents and Files	Files, Rare and Ancient Books, Rubbings, Chinese Full Text

For the geology theme, operation procedures consist of three main parts: 1. Preliminary procedures, including cleaning the specimen, repairing fossils, collecting specimen metadata and references, and establishing metadata fields; 2. Objects digitization, including scanning positive and negative images, photographing transparent thin film, 2D photography, 3D photography, and digital image storage and backup; 3. System establishment, including database development and website construction. In this stage text data and metadata are imported into the database, after which images, text, literature and the specimen database are linked together, providing a digital museum for the public to browse through.

For the painting and calligraphy theme, using the National Palace Museum as an example, operation procedures are divided into four parts: 1. Preliminary procedures, including digitization standard establishment, object selection, compiling a list of objects for digitization, recording contents of the original object, examining the negative, photographing objects, re-photographing, scanning negatives and verifying data; 2. Metadata analysis and editing, including coding file names, compiling introductions, webpage descriptions, and creating complete information on artworks; 3. Image digitization, main tasks include checking the quantity and order of negatives, scanning positives and inspecting images; 4. Storage and utilization, main tasks include digital image file protection – adding watermarks, digital data storage and management, and value-added applications.

For the archives theme, using the Institute of History and Philology as an example, operation procedures are divided into three stages. First stage: Organizing physical objects; main tasks include dusting objects, examining objects, mounting, verification and registration, and storing objects in the warehouse. Second stage: Digitization stage; this stage is divided into three parts that are carried out at the same: (1) Image processing, including object scanning, digital photography, digital image verification, revision, and post-processing; (2) Catalog establishment, including directory file cataloging and verification; (3) Value-added analysis, including authority file cataloging and verification. Third stage: Application stage, mainly includes system links and user access.

Operation procedures of the audiovisual theme are roughly divided into three parts: 1. Preliminary procedures, which includes data collection, contacting the creator, signing a license agreement, list compilation, website establishment, and hardware rental and maintenance; 2. Digital file conversion, which includes converting, burning, inspecting, collecting and editing files; 3. System establishment, which includes metadata search, planning and establishment, and uploading WMV files to the website.

For details on operation procedures of each thematic group, please visit the Taiwan Digital Archives Expansion Project website: http://content.ndap.org.tw/index/?p=994.

III. Comparison of Operation Procedures for Different Objects

Collections that belong to the same thematic group can be integrated to a certain extent, e.g. in the "Biology and Nature" theme, collections mainly consist of animal, plant and geological specimen; in the "Languages and Multimedia" theme, collections mainly consist of audiovisual data; in the "Literature and Archives" theme, collections mainly consist of documents and records. However, even the same type of object might involve different operation procedures, as a result of differences in the project's management method or the object's condition. In the following section, we will discuss common and different operations procedures for different objects.

1. Common operation procedures for different objects

After analyzing the six themes of digital archives, main collections can be categorized into the following: specimen, photos, literature, painting and calligraphy, antiquities, architecture, maps, video and audio; Table 3-3 lists operation procedures for different objects, which is then used as a basis for finding similarities and differences.

Table 3-3 Operation Procedures of Different Objects

	Preliminary Procedure	Digitization	Data Preservation	Value-added
Specimen	1. Specimen collection 2. Thorough data check, list compilation 3. Authentication and verification 4. Object selection 5. Specimen cleaning and treatment 6. Data and literature collection 7. Metadata requirement evaluation and analysis	1. Specimen measurement 2. Color calibration and light measurement 3. Digital photography 4. Photographed image inspection 5. Image file conversion 6. Specimen conservation (fish specimen) 7. Data cataloging 8. System development and establishment	1. Backup copy 2. Remote backup 3. Import into database	1. System search 2. Value-added applications

	Preliminary Procedure	Digitization	Data Preservation	Value-added
Photograph	1. Make an inventory of objects and plan tasks 2. Decide the digital file format 3. Select a digitization method 4. Select objects and verify the list 5. Evaluate and analyze metadata requirements	1. Equipment color calibration 2. Digitization operations – scanning 3. Image file naming and adjustments 4. File conversion 5. Data cataloging 6. System development and establishment	1. Proofing 2. Backup copy 3. Remote backup 4. Import into database	1. System management and access
Files and Documents	1. Select a digitization method 2. Work planning 3. Document compilation	1. Establish color management model 2. Scanning operations 3. Post-processing 4. Metadata field recommendations 5. Database establishment 6. Metadata cataloging	1. Digital image backup and management	1. System integration and access

	Preliminary Procedure	Digitization	Data Preservation	Value-added
Photograph	1. Objects sorting 2. Digital specifications selection 3. File naming principles 4. Digitization methods choosing 5. Human resource planning 6. Requirements evaluations and content analysis	1. Dimension measurement and cleaning 2.Preparations before photography 3. Digital photography, color calibration, post-processing 4. Digital proofing 5. System design and testing 6. Data cataloging	1. Save and backup	1. Digital museums 2. e-Learning 3. Creative products
Antiquities	7. Metadata requirements specification 8. Cataloging principles 9. Photography studio arrangements 10. Borrow objects 1. Establish standards 2. Antiquity preparation 3. List preparation 4. Photography system preparation 5. Borrow objects 6. Establish core metadata fields 7. System analysis and establishment	1. Photography and verification 2. Post-processing 3. Image quality inspection 4. Text data collection and data establishment 5. Input complete data of antiquities 6. System testing and modifications 7. Import into database, revision	1. Proofing 2. Save copy and embed watermark into the image	1. Database applications

	Preliminary Procedure	Digitization	Data Preservation	Value-added
Architectures	1. Collect and organize data 2. Object verification and registration 3. Object arrangement and cleaning 4. Metadata design, compilation, planning and establishment	1. Photograph buildings and structural components 2. Film buildings and structural components 3. Photograph and film rotating object 4. Photo retouching 5. Video file conversion, editing and categorization 6. Construct 3D model 7. Interactive platform establishment 8. Database establishment	1. Digital file compilation	1. Interactive platform and database 2. Licensing
Maps	1. Sort out maps 2. Repair maps 3. Make an inventory 4. Fill in data on the preservation condition 5. Plan Metadata fields	1. Scanning 2. Image retouching and inspection 3. File compression and conversion 4. Metadata cataloging 5. System establishment	1. Save and backup scanned image files 2. Backup in various media	1. Establish website for access
Video	1. Collect data 2. Contact the creator 3. Sign a license agreement 4. Check master tapes 5. List compilation	1. Video file conversion 2. Digital file editing 3. Metadata cataloging 4. File conversion in batches 5. Record the file onto a disc and check its contents 6. Import into database 7. Metadata collection, compilation and establishment	1. File storage 2. Backup in various media	1. Establish website for access

	Preliminary Procedure	Digitization	Data Preservation	Value-added
Audio	1. Collect data 2. Prepare digitization materials	1. Recording and file conversion 2. Data analysis, definition and verification 3. Metadata cataloging 4. Lexicon establishment 5. Automatic system establishment	1. File storage 2. Backup in various media	1. Automated system applications

After analyzing and arranging the operation procedures above, we can obtain Table 3-4 common operation procedures for different objects.

Table 3-4 Common Operation Procedures for Different Objects

	Preliminary Procedure	Digitization	Data Preservation	Value-added
Operation Procedure	1. Thorough data check 2. List arrangement and compilation 3. Objects preparation 4. Metadata requirement evaluation and analysis	1. Digitization process (scanning, photography, audio/ video recording, and file conversion) 2. Data verification 3. File conversion 4. Data cataloging 5. System development and establishment	1. Digital file storage (various media) 2. Remote backup 3. Database storage	1. System search 2. Value-added applications

2. Different operation procedures for different objects

After analyzing and arranging operation procedures for different objects, we can obtain Table 3-5 differences of operation procedures for different objects.

Table 3-5 Differences of Operation Procedures for Different Object

	Preliminary Procedure	Digitization	Data Preservation	Value-added
Specimen	1. Specimen collection 2. Authentication and verification 3. Specimen cleaning and treatment 4. Data and literature collection	1. Specimen conservation (fish specimen) 2. Photography		
Photos		1. Scanning	1. Proofing	
Files and Document	1. Document treatment, mounting	1. Scanning		
Painting and Calligraphy	1. Photography studio arrangements	1. Digital proofing 2. Photography		1. Digital museum
Antiquities	1. Photography system preparation	1. Photography	1. Proofing	2. e-Learning
Architecture	1. Sort out and clean objects	1. Construct 3D model 2. Photography, filming		
Maps	1. Repair	1. Scanning		
Video	1. Contact the creator 2. Sign license agreement	1. File conversion		

	Preliminary Procedure	Digitization	Data Preservation	Value-added
Specimen		1. Data analysis, definition and confirmation 2. Lexicon establishment 3. Automatic system establishment 4. Recordings and file conversion		1. Automated system applications

Differences in operation procedures might also result from different digitization methods, digitization projects, or considerations of the execution unit. For example, if scanning or photography is selected as the digitization method, then color calibration must be completed for digitization equipment; documents need to be arranged in order; specimen and antiquities need to be cleaned and treated; documents, painting and calligraphy, and maps involve repair and mounting; audiovisual data need to undergo file conversion and editing; photos, painting and calligraphy, and antiquities need to go through proofing; and buildings and structural components involve photography and filming.

From the perspective of project institutions: preliminary procedures of museum and library related projects require the use of a library system; library related projects have somewhat different coding methods; museum related projects tend to emphasize on digital museum and e-learning applications. If the project unit is a school, then management will vary along with the school's regulations.

IV. Main Tasks of Digital Archive Content Establishment

To select the most suitable digitization method, analyze operation procedures of all projects and possible results, and thoroughly understand characteristics of digital objects. Typically, the establishment of digital content mainly involves scanning, photography, audio recording, video taping and metadata descriptions.

1. Scanning

Objects that are digitized via scanning are mostly flat objects, including: photos, literature and maps.

Commonly used scanners include: drum scanner, desktop platform scanner, desktop scanner with mask, desktop sheet-fed scanner, desktop seamless book scanner, upward facing book scanner, and professional multi-purpose scanner. A brief introduction to their functions and usage is as follows:

(1) Drum scanner: Professional scanner for prints; can scan large objects, but only

if it is a single page; highest resolution can reach 4800dpi, is relatively slow, but centrifugal force causes damage to pyhsical objects; this type of scanner is almost entirely out of production.

(2) Desktop platform scanner: Relatively cheap compared with other scanners, and is further divided into high-end and low-end machines depending on its resolution. When scanning positives, desktop platform scanners are usually required to have a mask. For most platform scanners, the light source is under the object being scanned, so when positives are being scanned, masks allow the positives to be placed below the light source; positives require accurate colors, and therefore need to be placed under the light source. Desktop platform scanners can scan up to A3 size objects, if the object is larger than A3, such as a newspaper or large poster, then it needs to be scanned in separate sections and then combined digitally. Books and periodicals, however, have thickness and areas close to the bindings are hard to scan using this type of scanner.

(3) Desktop sheet-fed scanner: Suitable for single sheet flat objects that can be rapidly fed into the scanner; the objects can be at most A3 size. Sheet-fed scanners face the issue of paper being jammed, so like platform scanners when scanning objects larger than A3 size, the object needs to be scanned in separate sections, and then combined digitally. Thick objects, such as books and periodicals, can not be scanned using a sheet-fed scanner. This type of scanner is mainly used for scanning large quantities of documents; it often can not meet color scale requirements of photos for bright areas or shadows.

(4) Desktop seamless book scanner: Modified desktop scanner that has a slanted side for scanning books and periodicals; the maximum object is A4. However, in order to scan areas close to the binding of thick books, high pressure must be applied to the book during scanning, and might cause pages to drop off.

(5) Upward facing book scanner: Designed based on the concept of re-shooting, upward facing book scanners can scan objects up to A2 or A1 size; the glass lid is used to flatten books so that areas near the bindings can be clearly scanned; scanning with this type of scanner is relatively fast.

(6) Professional multi-purpose scanner: Possesses characteristics of both re-shooting and traditional scanners; can scan photos, books, periodicals, newspaper, maps, painting and calligraphy, textiles, and plant specimen, as well as not curved solid objects, such as glass painting, shadow puppets, and jade of sizes up to A1. The scanning platform's design allows areas near book bindings to be clearly scanned without the use of a glass lid. Scans fast and supports high digital file specifications.

After selecting suitable digitization equipment, make preparations to begin scanning, in which an essential task is equipment color calibration. The main purpose of color

calibration is to reproduce the exact same colors of the physical object when it reaches each piece of equipment. Color calibration procedures include: computer monitor color calibration → digitization equipment (scanner) color calibration → output equipment (printer or printing machine) color calibration.

After completing color calibration, begin scanning objects according to their order. Lay the object flat on the scanner's platform to begin scanning. In addition, objects that use scanning for digitization are easily affected by heat and humidity, which causes different tension on both sides, curving the object as a result. If the platform scanner does not have an upper lid, then other methods must be used to flatten the object, such as adding a transparent acrylic or glass sheet on top, or developing a platform that provides suitable suction to keep the object flat on the platform for digitization.

2. Photography

Objects digitized via photography are mainly solid objects, including: specimen, painting and calligraphy, antiquities, and architecture related. In which painting and calligraphy stand out because they are categorized as large flat objects, most of which are larger than A0 size; therefore, most holding institutions use photography for digitization.

Common photography equipment includes:

(1) Camera, Digital Camera Back:

i. Digital camera: Digital cameras are operated like conventional cameras, but post-processing is done on a computer. Digital cameras use light sensing devices to record images into storage media; the storage media can be reused so less cost is required. Digital cameras produce digital files, which save time on conversion required for conventional cameras, making it the most commonly used camera type for digitization.

ii. Digital camera back: A CCD or CMOS sensor added to the back of a conventional single lens reflex camera; usually used together with a 120 medium format or 4×5 large format conventional camera. Digital camera backs contain light sensing devices and control/storage units. To use a digital camera back, take of the back cover of a conventional camera and replace it with the digital camera back; this will turn a camera that originally used film into a digital camera.

iii. Conventional camera: Images produced by conventional cameras reach a certain quality consideration that make professional photographers still chose to use such single-lens reflex cameras; this is why conventional cameras still have not been entirely replaced by digital cameras. Here we must stress that the positives or film produced by conventional cameras need to be scanned to generate digital files.

Considering future applications of digital images, we recommend using a DSLR camera or medium to large format conventional camera with digital camera back. Besides allowing different lenses to be used, these cameras provide better image quality because more detailed aperture, shutter and view depth adjustments can be made. When selecting a digital camera or camera back, the two main considerations are "pixels" and "CCD size", image quality is greatly affected by these two factors.

In terms of brand selection, we recommend large well-known brands because accessories are easier to purchase and repair is more convenient. If funds and equipment are limited, one shot DSLR cameras in the market can be considered an option for digitization. DSLR cameras currently have up to 20 megapixels, and different brands are all releasing full frame DSLR cameras; these are all good choices under limited funds.

(2) Lens

Besides camera selection, lens selection also affects the quality of images. The light sensing device (CCD or CMOS) of Typical DSLR cameras is smaller than film used by conventional cameras, so when lenses are added, the focal length is increased (generally 1.5-fold to 1.6-fold). When photographing static objects in a photography studio, we recommend using a zoom lens, this way only the lens needs to be adjusted without changing the camera's position. Furthermore, a macro lens can be used for photographing parts of an artwork or details, such as detailed parts of a specimen or an author's inscription.

(3) Lights

Aside from the camera and lens, lights and adjustments are also an important part of photography. For light selection, ancient objects are not as resistant to light, whether it may be natural sunlight or artificial light, and easily become brittle; heat generated by lights also cause different degrees of deterioration or color fading. Therefore, if cultural relic preservation is to be achieved, the light control is not only required during regular collection and exhibitions, but also during digitization; avoid shinning lights that are too bright or for too long. We recommend using standard color temperature cold lights for photographing objects. Although cold lights are continuous light sources, they generate far less heat than tungsten lights, minimizing its damage to objects.

After preparing photography equipment and environment, make preparations for photographing objects. Equipment color calibration must first be completed before photographing objects, equipment that require calibration include the computer monitor, digitization equipment (camera), and output equipment (printer or printing machine).

When photographing objects, pay attention to the following matters:

i. Avoid photographing objects in a hazardous environment.

ii. photography lights (continuous light sources) emit ultraviolet light, so 5000K to 5500K cold lights that do not emit ultraviolet light are used to reduce damage to objects.

However, sometimes the brightness of cold lights is insufficient, so light selection should be based on the current condition of collections.

iii. To satisfy color calibration requirements of printing and proofing, photograph objects along with a basic color card, gray scale card and scale.

iv. After using a light meter to adjust the evenness of light, verify the consistency of apertue and color space.

3. Audio/Video Recording

Objects digitized via video/audio recording are mainly video data and audio data, most of which are data gathering records, interviews, and field investigations.

Common video/audio recording equipment include: traditional betacam recorder, digital video camcorder, radio receiver tape recorder, MD, recorder pen, microphone, wind cap, and various cables.[10]

Basic equipments for audiovisual interviews:

(1) Two or more video cameras to take different angles of the person being interviewed; this is to collect complete information to avoid sudden accidents that might prevent data from being gathered.

(2) Special functions of a directional microphone linked the video camera allow clearer audio data to be gathered.

(3) Mini-microphones pinned on the person being interviewed allow audio data to be clearly and completely recorded.

(4) Set up the video camera on a camera stand to avoid shaking images, while reducing the strain of workers.

(5) Recorder pens make data gathering more convenient, files are easy to read and take less time, benefitting following procedures.

(6) Lighting equipment make video data gathered clearer.

(7) Preparing a camera to take photos allows even more complete data of the person being interviewed to be gathered.

(8) Listen on the microphone and earphone to ensure all sounds are captured.

Operations procedures for interviews:

(1) Prepare at least two video cameras for interviews, one at a fixed position (usually in front of the person being interviewed), and one mobile (this way different angles can be provided to choose from). The position of the person being interviewed can be planned during earlier visits to the interview site.

[10] TELDAP Taiwan Digital Archives Expansion Project, "Operation Procedures of National Taiwan University of Arts Department of Graphic Communication Arts 'Folk Dance Costumes – Taiwan Folk Arts Digital Cultural Village Project", February 2009.

(2) One person interviews while the other records, but the interviewer should not be in the picture.

(3) Any parts that were left out should be recorded in another take.

(4) Notify the person being interviewed of how the audiovisual data will be handled and acquire authorization.

When collecting audiovisual data, pay attention to the following matters:

(1) Be sure to fully understand the person's background.

(2) Check and double check the time and place where data will be gathered, be sure to confirm with the person being interviewed before setting out.

(3) Draft an outline of data gathering; if it is an interview, provide the interview's contents to the person being interviewed in advance.

(4) Visit the site where data will be gathered or the interview will take place in advance.

(5) Make sure the lights are bright enough when gathering data or conducting an interview, and avoid disturbance from other noises.

(6) Double check all equipment before data gathering or an interview.

(7) Try your best to overcome sound reception problems at the site where data is gathered or the interview takes place.

(8) To enhance sound quality, a wireless directional microphone can be pinned on the person being interviewed. Directional microphones have the advantage of reducing noise and enhancing sound quality.[11]

4. Metadata Description

The key to whether or not digital archiving succeeds lies in the establishment of a database with "integrated functions and fields", using efficient knowledge management to achieve rapid resource search and extraction. In which thorough metadata planning is essential to achieving highly efficient data explanation and searching. Therefore, metadata establishment is the most fundamental task in digital archiving; its purpose is to describe the contents and structure of digital archives in a standardized and structural way.[12]

Before developing metadata, there are numerous factors and conditions that must be thoroughly thought out and planned. First is the objective and development direction of the institution implementing the digital archives project, which directly affects the functions of digital objects, and indirectly affects the depth and width of metadata, as

[11] TELDAP Taiwan Digital Archives Expansion Project, operation procedures of Nanya Institute of Technology "Digital Archives of Lyung Gogan Atayal Kgogan Cultural Stories", August 2008.

[12] Li Te-Tsai, "Digital Archives Technology Collection", 2004 edition. Search: January 2010. http://www2.ndap.org.tw/eBook/showContent.php

well as possible developments after being linked to other databases. Furthermore, a clear understanding of collection characteristics and knowledge specific to the respective field is also required. In order to satisfy the two requirements above, early stages of metadata establishment involve gathering scholars of specific fields, system developers, and metadata exerts and discussing basic elements and contents of research, management and maintenance, so that different aspects of knowledge inherent in collections can be represented.[13]

Based on surveys and studies on archival collections, the Metadata Architecture and Application Team (MAAT) established the following eight principles for planning metadata, so that metadata reflect on characteristics of local collections, while providing functions that satisfy academic research requirements, and allow internationalization of online resources:[14]

(1) Meets requirements of international standards, including Metadata format, internet protocol, and searching specification (e.g. Z3 · 0), etc.

(2) Modified from existing standards according to domestic requirements, and not separately developed.

(3) Easy to use, including data establishment, representation, search, interpretation and interaction.

(4) Applicable to multiple languages.

(5) Adopt different metadata formats for different fields, and not try to design a single metadata format that is applicable to all fields.

(6) Adopt different metadata formats, at the same time use XML to construct the structure of metadata, so that it is compatible and integrates different formats, including integration with existing library systems, which allows a single search through all library resources.

(7) Design system management mechanisms and functions in response to different requirements, including access restrictions and intellectual property rights protection.

(8) Possessing expandability, including conversion, storage and representation of different data formats, as well as metadata establishment and correction.

After establishing metadata elements, begin filling in and cataloging metadata of digital files; the person responsible for cataloging must possess profession knowledge

[13] Huang Wan-Yu, Tsai Shu-Yun, "Metadata of the Institute of History and Philology's 'Ethnological Investigation Photos Database", collected in the "Seminar on Digital Archives and Value-added Applications of Anthropological Knowledge", May 2004.

[14] Chen Ya-Ning, Chiang Hui-Ying, Chen Shu-Chun, Liu Chia Hui, "Chinese Metadata Planning and Discussion on its Trial Implementation", Academia Sinica MAAT phase 1, 2 reports. Search: January 2010. http://cdp.sinica.edu.tw/project/04/6_3.htm

of the archival collection, as well as basic word processing ability. It is necessary to check the metadata to ensure that data is correct, especially for text cataloging data that correspond to images. After cataloging metadata, researchers should proofread catalog contents, checking if the right field and information was keyed in. If mistakes were found, make necessary corrections and proofread contents again; any additions should also be completed and proofread in this stage.

In metadata cataloging, we must specially mention the "content description" field. This field was designed to explain contents of the video/audio to the user. In order words, the "content description" field allows users to see an image in a larger social and cultural context, and learn more of the local knowledge behind the image. The "content description" field retextualizes contents that have been detextualized. We must not only carefully handle retextualization so that mistakes or over interpretation doesn't occur, but also understand the complex knowledge of image contents, of the photographed ethnic group, and of the group's social status and living environment. Naturally, being able to hear locals interpret or explain certain matters would be the most ideal, but the explanation of scholars or specialists familiar with the field would also suffice.

The importance of the "content description" field is that it makes up for important image information that can't be provided by other fields using text descriptions. However, descriptions cannot be extended and expanded without limits, and could raise doubts of over interpretation. Furthermore, considering the database's capacity, the more data fields there are the more time it takes the system to match information, reducing search speed and lowers the convenience and time effectiveness of searches. Therefore, we still have to restrict the length of the description so that core descriptions are provided, using the simplest and most thorough way to represent contents of an image.[15]

Digital archive projects cross over multiple fields and disciplines, making its working environment relatively complex. By establishing good operation procedures personnel will be able to clearly understand standards and methods related to the entire digital archiving process. Therefore, the planning of complete operation procedures can enhance digital content quality.

[15] Huang Wan-Yu, Tsai Shu-Yun, "Metadata of the Institute of History and Philology's 'Ethnological Investigation Photos Database", collected in the "Seminar on Digital Archives and Value-added Applications of Anthropological Knowledge", May 2004.

FOUR. Integrated Operation Procedures in Digital Archiving

I. Integrated Operation Procedures in Digital Archiving

In the Taiwan e-Learning and Digital Archives Program (TELDAP), different operation procedures are adopted based on characteristics of the collection being digitized. In the previous section on operation procedures in digital archiving, we tried to find main steps and procedures and draw an integrated flowchart with common operation procedures (Fig 4-2). In the following section we will use this flowchart to explain contents of main steps and procedures in integrated operation procedures.

1. Main steps of integrated operation procedures

From the analysis of the previous chapter we can find that there are four main stages of digitization procedures, as shown in Fig 4-1, including: preliminary procedures, digitization procedures, data preservation and value-added applications.

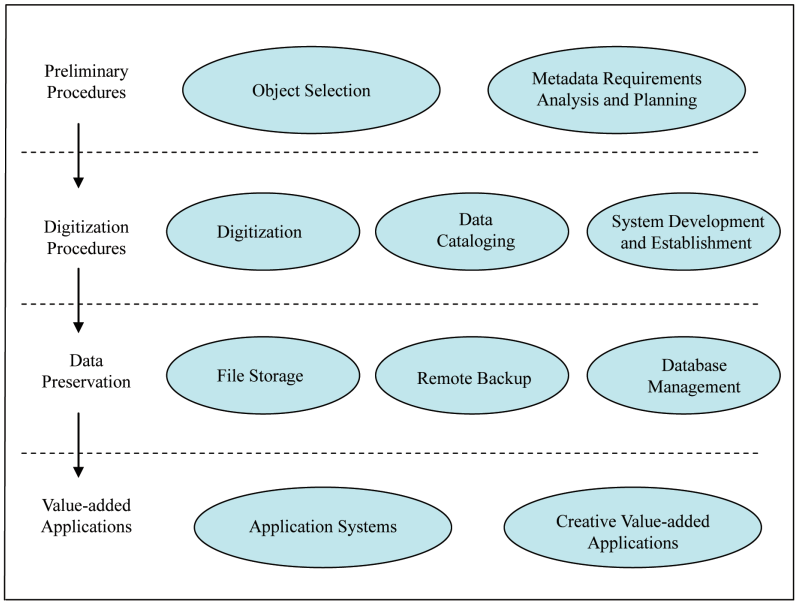


Fig 4-1 Concepts of Digitization Procedures

Source: Taiwan Digital Archives Expansion Project

Following the concepts of digitization procedures above, and combining operation procedures of different objects, we can draw the more detailed “digitization flowchart” shown below (Fig 4-2).

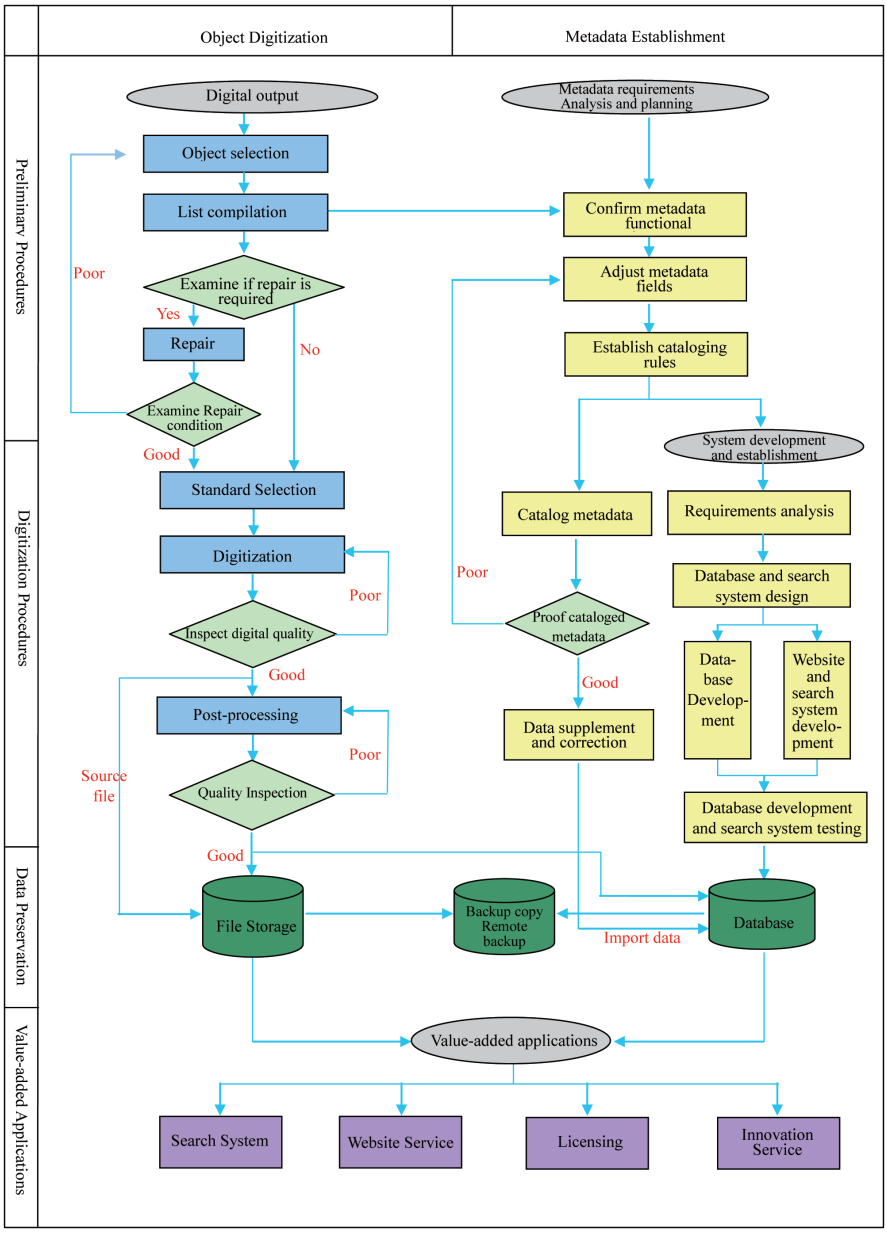


Fig 4-2 Digitization Flowchart

Source: Taiwan Digital Archives Expansion Project

From the flowchart above, we can briefly explain the main stages of integrated operation procedures:

(1) Preliminary Procedures:

Main work items include thoroughly checking data, compiling lists, and selecting objects for digitization (also involves object repair when necessary). At the same time metadata requirements are evaluated and analyzed, drafted, adjusted, and then cataloging principles are established.

(2) Digitization Procedures:

This stage is the key focus of current digital archive projects. Besides object digitization (including scanning, photography, video/audio recording and file conversion), this stage also includes data cataloging and system development. No matter what step you are on, "quality inspection" is an important task, including inspecting the quality of digital objects, quality of post-processing, proofreading metadata, and testing and modifying the database and search system.

(3) Data Preservation:

Data preservation is a long-term strategy of digital archives, and includes digital file storage (storage on different media, storage of the source file, and storage of digitization results), backup copy, remote backup mechanisms, and importing digital files and metadata into the database. The permanent preservation of digitization results will be further discussed in following sections.

(4) Value-added Applications:

Besides sharing digitization results via an online search system, institutions now have to consider value-added applications and sustainable operation of these precious data, so that more people can learn and utilize it. Value-added applications can be categorized into the following: search system, website services, licensing and innovation services. Value-added applications in academics, education, commerce and even social services are all major objectives of this stage.

In these four main stages, there are even more detailed steps; each link complements another, so if any step goes wrong, it should be reviewed and modified to allow the procedures successfully operate. This way the quality of results can be further improved.

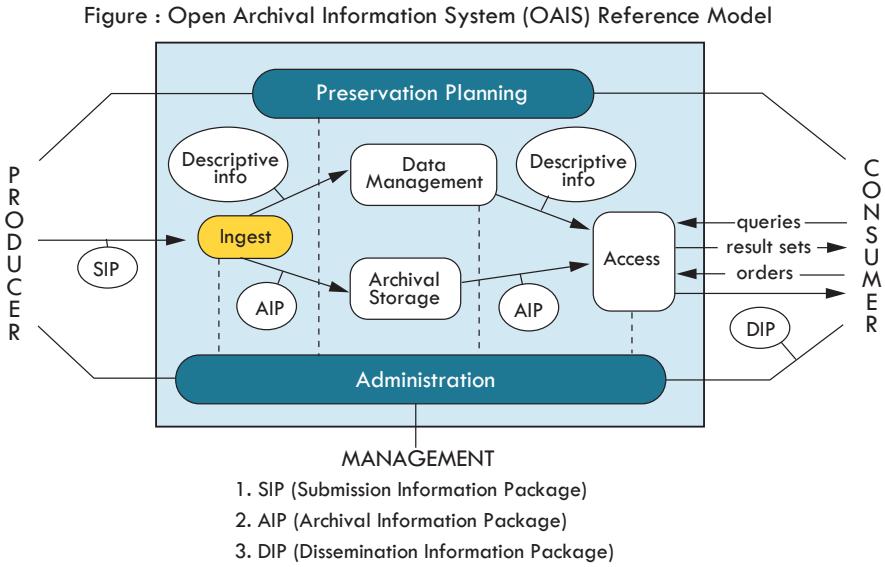
2. Relationship between integrated operation procedures and OAIS

Integrated operation procedures consist of the four stages: preliminary procedures, digitization procedures, data preservation and value-added applications, all of which conform to the concept of digital information life cycle. Comparing with developments of digital archiving operation procedures in foreign countries, we use the ISO Open Archival Information System (OAIS) reference model as an example, and discuss their relationship in digitization procedures.

(1) OAIS (Open Archival Information System)

OAIS (Open Archival Information System) was released in 2003 by the International Organization for Standardization (ISO) with support from the Consultative Committee for Space Data Systems (CCSDS), which is under NASA. OAIS is an archive consisting of an organization of people and systems that recommends long-term preservation. The system model comprises six main functions: Ingesting, Archive Storage, Data Management, Access, Administration and Preservation Planning. The model indicates the relationship between upstream, midstream, downstream, producer and consumer in digital archiving. Depending on the preservation stage, three types of information packages can be found, including: Submission Information Package (SIP), Archival Information Package (AIP), and Dissemination Information Package (DIP). This system was originally designed for storing space data, but has been widely applied to more types of data systems in recent years. The OAIS reference model is as shown in the figure below:

Fig 4-3 OAIS Reference Model[16]



The OAIS reference model is directed towards database functions and requirements, and provides a common architecture for metadata in different systems, provides a conceptual framework, defines the environment in which an archive system exists,defines

[16] Source: Reference Model for an Open Archival Information System (OAIS). CCSDS 650.0-B-1, Blue Book, January 2002

functions and organization of a archive system, supports the basic information structure for processing archives, and achieves the long term preservation and maintenance of digital archives. The six main functions of the reference model are:

i. Ingest

This function is mainly responsible for accepting (Submission Information Packages (SIP) of the producer. In other words, before digital images of collections are established, related standards must be submitted, including data format, type, form, transmission method, physical documents, and internal affairs. Processed SIPs are converted into Archival Information Packages (AIP) that meets the planned archive standards, which are then transferred to data management and archival storage. However, some information cannot be provided by the producer, and some information needs to be refreshed as contents are changed.

ii. Archive Storage

Receives AIP produced by digital images and stores them permanently. Storage must conform to requirements of certain purposes (e.g. permanent preservation, browsing, etc.), storage media is updated at all times, confirmation is made on a regular basis, transfer and error detection are provided, and transfer is completed via AIP.

iii.Data Management

Mainly manages metadata of collections, and provides maintenance and access services, including database management, archive architecture and integrity maintenance, and database refresh. The information system supports operations of the archival institution; descriptive information is output to AIP based on results.

iv. Access

This function helps users to acquire information, turning information requested by users into Dissemination Information Packages (DIP), allowing users to send requests for information, receive responses, and access information.

v. Administration

This stage is mainly responsible for managing operations of the overall archive system, including soliciting producers and signing agreements, specification recommendations and preservation planning, software/hardware architecture maintenance, and receiving operational data from the data management and archival storage functions. The overall function is to establish rules for the digital archive project and manage overall system operations as means for monitoring operation quality.

vi. Preservation Planning

Mainly develops long-term preservation policies to ensure that users can access OAIS information. Therefore, this function provides recommendations for the

long-term storage and search of data, investigates user communities, refreshes information technologies and standards, and makes necessary long-term preservation planning.

OAIS procedures utilize coordination between personnel and systems to complete the long-term preservation and maintenance of digital archives. In which main roles can be divided into producer, administrator and consumer. The producer's role is to provide information to the consumer's system; the administration consists of people who plan the overall policies of OAIS, and assess the policies and operations on a regular basis; consumers use OAIS services to acquire information. Overall usage regulations follow instructions of the administration, preservation strategies and technology developments are established by preservation planning, and then implemented by the administration.

(2) Relationship between integrated operation procedures and OAIS

The completeness of integrated operation procedures can be examined by corresponding them to stages of the digital information life cycle. Main stages of integrated operation procedures include: preliminary procedures, digitization procedures, data preservation and value-added applications, and accord with the digital information life cycle. However, there are some differences in operations procedures of TELDAP with OAIS. As shown in Table 4-1 comparison of OAIS in different countries, similarities of current integrated operation procedure projects with OAIS are they all use international standard formats for collection exchanges. The difference is that the National Digital Archives Program has already been implemented for numerous years, and quite a few digitization manuals have been compiled based on experiences of its participants; these digitization manuals are provided as reference to digital archive projects.

In addition, OAIS plans preservation planning and administration within its operation procedures, which is the same as integrated operation procedures; preservation planning is a necessary process to the enhancement of digital quality. Still, when implementing long-term preservation of digital resources, the OAIS reference model is indeed a pattern that can be followed. The "value-added application" stage of integrated operation procedures is especially worth noting because it only considers applications in search systems, but also considers other value-added applications that increase benefits of digital archives.

[17] Source: Digital Archive e-Learning website, search: January 2010. http://ic.shu.edu.tw/DA/DADL/dam_02_foreign_05.htm

Table 4-1 Comparison of OAIS in different countries[17]

Main Project		Similarities with OAIS	Differences with OAIS
Institution or Project Name	Planning Perspective		
National Digital Archives Program, Taiwan	Metadata requirement design and relative experiences of institutions are used as reference	International standard formats are used for exchange specification	Digitization guidelines are compiled for work experience sharing, and provided to other themes for reference
National Repository of Cultural Heritage, Taiwan	Integrated archival storage; common metadata fields are established for units to fill in	Standards are established for themes of the CCA based on exchange formats of collections	Sharing common specification of digital files and metadata are submitted by the CCA for application a common search platform
Library of Congress, U.S.A.	Digital archiving does not need to accord with all listed procedures; most procedures are executed simultaneously	Emphasizes on administration and preservation planning; provides confirmation of exchangeable files	Gives consideration to value-added applications of digital objects
Canadian Heritage Information Network, Canada	Comprehensive archive planning recommendations, includes short-term and long-term planning	Provides confirmation of exchangeable files	Gives consideration to references on digitization procedures (scanning, photography)
UK Humanities, Arts and Scientific Data Services, England	Plans operation procedures based on the OAIS framework; considers responsibility assignment	Uses OAIS components for planning operation procedures	Stresses on the duties of roles in operation procedures, and their coordination with the information system
Council of Europe Forum, European Union	Focuses on data storage, transfer, and content definition	Gives consideration to overall long-term preservation planning	Only emphasizes on the concepts of archival storage and description

Main Project		Similarities with OAIS	Differences with OAIS
Institution or Project Name	Planning Perspective		
National Library of Australia, Australia	Describes procedures from the perspective of purpose and meaning	Provides confirmation of exchangeable files	Administration and preservation planning evaluations; does not focus on information system development

II. Comparison of Born-digital and Digitized Materials[18]

The first task of digitization procedures is selecting objects to digitize; some objects are born-digital, while others are digitized.

1. Born-Digital and Digitized Objects

Digital objects are generally considered to be the body of digital materials, which may be in the form of digital publications and digital records; both born-digital and digitized are considered as digital objects.[19]

Born-digital objects are data created on digital media, while digitized objects are the digital form of physical collections.

The scope of data preserved in digital archives is extensive, and includes any data that represents a part of human civilization, regardless of the time and form it was produced as long as it possesses value for preservation; culture, science, art, music, animals and plants all fall within the scope of digital archives.

In terms of digital data formats, any digital form (text, audio, video, and 2D/3D objects) of digitization targets (records, antiques, specimen, audio recordings, videos, and architecture) fall within the scope of digital archive data formats. Digital objects serve as substitutes for physical objects, or were born in digital form. When representing digital objects, besides the data itself, other attributes include structure, file format, file size and source.[20]

"Born-digital" data are created in digital form; such data are large and come in numerous forms.[21] In the age of information technology, digital information is characterized by being easy to disseminate, duplicate and not easily distorted, so "born-

[18] Tsai Yung-Cheng, Huang Kuo-Lun, Chui Chih-Yi, "Introduction to Digital Archive Technology", Taipei City: NTU Publishing Center, November 2007, pages 18-19.

[19] Ingeborg Verheul, Networking for Digital Preservation: Current Practice in 15 National Libraries. (Munchen: K.G. Saur, 2006), pp.21.

[20] Chen Chao-Chen, "Long-term Preservation of Electronic Resources", "Information Management for Buddhist Libraries", Issues 25 and 26, June 2001.

[21] Hosted by Chen Chao-Chen, "Status Quo and Trends of Multimedia Digital Archives and Study on Multimedia Information Platform Specifications in Taiwan", commissioned by the Executive Yuan Government Information Office, November 2004.

digital" resources are already a common form of information. Any digital object that was created in digital form, including audiovisual data, electronic documents, and electronic data are all "born-digital" objects.

The opposite of "born-digital" is "digitized", and refers to physical objects within the scope of digital archives (e.g. records, antiques, specimen, audio recordings, videos, and architecture) changed into digital formats via digital archiving technologies.

2. Comparison of Operation Procedures for Born-Digital and Digitized Objects

"Born-digital" objects are digital data when they come into existence, and therefore don't require certain procedures in digital archiving, such as object selection, object repair, and digitization related procedures. Take electronic documents as an example, from the perspective of the four stages of operation procedures – 1. Preliminary procedures: Putting objects in order, compiling lists, and evaluating metadata requirements; 2. Digitization procedures: the documents are already digital files, so digitization procedures aren't required, and the quality of files is directly examined, followed by post-processing, file storage, data cataloging and system development and establishment; 3. Data preservation: Digital file storage (different storage media), remote backup, and import into database; 4. Value-added applications: search system and value-added applications.

For multimedia objects produced by electronic equipment, such as audio recordings of interviews, the four stages of operation procedures would be – 1. Preliminary procedures: Main tasks include planning which parts need to be sorted out, compiling lists, contacting the creator, signing a licensing agreement, and evaluating metadata requirements at the same time; 2. Digitization procedures: Digital quality inspection, post-processing (includes file conversion and editing), digital file storage, data cataloging, and system development and establishment; 3. Data preservation: Includes making backup copies, inspection, filing, and uploading data to the database; 4. Value-added applications: Includes the search system and value-added applications.

Operation procedures of "born-digital" objects are as shown in Table 4-1:

Table 4-2 Operation Procedures of "Born-digital" Objects

	Preliminary Procedure	Digitization	Data Preservation	Value-added
Born-Digital	1. Object selection 2. List compilation 3. Metadata requirements analysis and planning	1. Quality inspection 2. Post-processing 3. File storage 4. Data cataloging 5. System development and establishment	1. Digital file storage (different storage media) 2. Remote backup 3. Import into database	1. Search System 2. Value-added Applications

Source: Taiwan Digital Archives Expansion Project

"Non born-digital" objects were not digital originally, so they involve physical object selection, object repair and cleaning, and object digitization procedures. Using early documents as an example: 1. Preliminary procedures: This is the stage where physical objects are handled, including dusting, inspection, mounting, arrangement, digitization plan selection, work planning, and metadata requirements evaluation and analysis; 2. Digitization procedures: Color management model establishment, photography or scanning, post-processing, metadata recommendations, database establishment, and metadata cataloging; 3. Data preservation: Digital image backup and management; 4. Value-added applications: System integration and access.

Using the "photograph" theme as an example: 1. Preliminary procedures: Includes making an inventory of physical objects and work planning, deciding digital file specifications, selecting the digitization method, object selection and list verification, and metadata requirements evaluation and analysis; 2. Digitization procedures: Includes equipment color calibration, scanning operations, image file naming and adjustments, image file conversion, data cataloging, and system development and establishment; 3. Data preservation: Includes proofing, backup copies, remote backup, and import into database; 4. Value-added applications: System management and access.

Operation procedures of "non born-digital" objects are as shown in Table 4-3:

Table 4-3 Operation Procedures of "Non born-digital" Objects

	Preliminary Procedure	Digitization	Data Preservation	Value-added Applications
Digitized	1. Thorough data check 2. List compilation 3. Borrow objects 4. Metadata requirements evaluation and analysis	1. Object digitization (scanning, photography, video/ audio recording, and file conversion) 2. Data verification 3. File conversion 4. Data cataloging 5. System development and establishment	1. Digital file storage (different storage media) 2. Remote backup 3. Import into database	1. Search System 2. Value-added Applications

Source: Taiwan Digital Archives Expansion Project

The main difference between born-digital and non born-digital objects is in the "data creation" stage. "Born-digital" objects are created as digital data, so they only require quality inspection; if digital quality is poor, then the physical object needs to be re-borrowed and re-digitized; if digital quality reaches standards, then some procedures can be omitted, such as object selection, cleaning and repair, and digitization procedures (scanning and photography).

Differences between operation procedures of born-digital and non born-digital objects are shown in the digitization flowchart below.

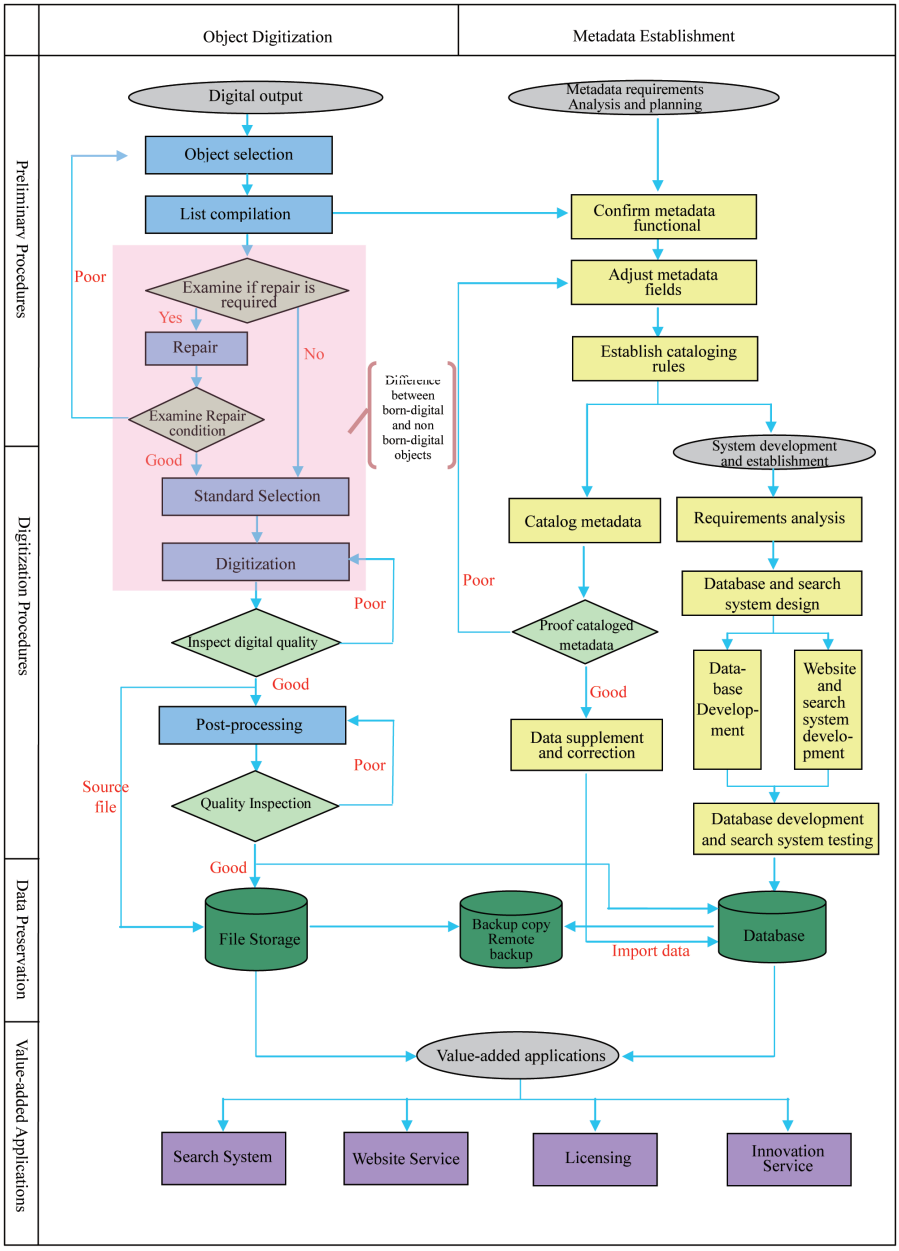


Fig 4-4 The Difference of Digitization Flowchart for Born-digital and Non Born-digital Objects
Source: Taiwan Digital Archives Expansion Project

III. Selecting Formats for Digital Images and Metadata

Besides differences of operation procedures for born-digital and digitized objects, which are in the preliminary procedures stage, the selection of digital images and metadata formats is also extremely important in following procedures. Below we will introduce reference standards commonly used by digital archive projects in Taiwan and overseas. Institutions can select suitable formats based on their requirements.

1. Selecting Image Formats

When planning digital file formats, one must first understand future purposes of digital files. In current digital archive projects, main purposes usually include permanent preservation, commercial value-added applications and online browsing. In which file formats can further be divided into text data and image data:

(1) Text data

i. Source data is an electronic file

If source data is an electronic file typed on a computer, besides saving a copy of the source file, we recommend converting the file into HTML, PDF or RTF.

ii. Source data is a print or manuscript

When the source data is a print or manuscript, in order for full text search to be possible, then the text needs to be keyed in or OCR software should be used to generate a digital file; remaining procedures are the same as the above. If full text search is not required, then only metadata needs to be established; scan the source document, store a permanent copy in an uncompressed format, and then convert the file into other formats according to requirements, such as popular applications and preview formats. Recommendations for these three formats are as follows:

Table 4-4 Text Data Reference Specifications:

Purpose	Permanent Preservation	Popular Applications	Preview
Description	Retains the original appearance of digital archives	For users to view online and print out	For users to preview and choose fields from
File Format	TIFF (not compressed)	JPEG (compressed)	GIF (compressed)
Color Mode	RGB (24bit/pixel) or above	RGB (24bit/pixel) or above	RGB (8bit/pixel) or above
Resolution and Dimensions	Resolution: 300~600dps or higher (Select a suitable resolution based on the quality and importance of the source file; typical prints can use 300dpi)	150dpi~300 dpi or dimensions of 500x400 ~ 1000x700 pixels	72dpi or dimensions of 150x100 ~ 200 x 200 pixels

Source: “National Digital Archives Program Technology Collection” and Taiwan Digital Archives Expansion Project

(2) Image data

If the source data is a photograph, image or map, use a scanner to scan it into an image file, store a permanent copy in an uncompressed format, and then convert the file into other formats according to requirements. Recommendations for the three formats are as follows:

Table 4-5 Image File Reference Specifications:

Purpose	Permanent Preservation	Popular Applications	Preview
Description	Retains the original appearance of digital archives	Provided to users for reproduction, compression or other image processing and exchange purposes	For users to preview
File Format	RAW	TIFF	JPEG
Color Mode	RGB (24bit/pixel) or above	RGB(24bit/pixel) or above	RGB (24bit/piexl)
Resolution and Dimensions	Original size, 300dpi or above	Original size, 300dpi or above (Select a suitable resolution based on the quality and importance of the source file; typical prints can use 300dpi; 600dpi is recommended for replicas; 350dps is recommended for publications)	72dpi, dimensions depend on display requirements (150x100 ~ 200x200 pixels)

Source: “National Digital Archives Program Technology Collection” and Taiwan Digital Archives Expansion Project

"Preservation" and "sharing" are important purposes of digital archives. In the case of paintings and calligraphy, albums and replicas are highly probable applications, so quality standards of digital images must be elevated to ensure the correctness and quality of digital image colors. Digital file formats for painting and calligraphy of two institutions are listed below in Table 4-6 "Image File Specifications of the National Palace Museum" and Table 4-7 "Image File Specifications of Hwa Kang Museum". During actual implementation, institutions should select formats that best suit their purposes.

Table 4-6 Image File Specifications of the National Palace Museu

Purpose / Specifications	Raw: Source File from the Camera	Archive Image: High Resolution	Medium Image: Medium Resolution	Reference Image: Low Resolution	Preview Image: Low Resolution Small Image
Type	RGB (48bits/pixel)	RGB (24bits/pixel)	RGB (24bits/pixel)	RGB (24bits/pixel)	Index colors (8bits/pixel)
File Format	Raw	TIFF	TIFF	JPEG	GIF
Resolution	350dpi	350dpi	350dpi	72dpi	72dpi
Image Size	6668×4992 pixel	6668×4992 pixel	3334×2496 pixel	500×400 ~ 1000×800 pixel	100×100 pixel
File Size	250MB	90MB	20MB~25MB	Less than 200KB	Less than 20KB

Source: Kao Lang-Hsuan, Chen Hsiu-Hua, "Painting and Calligraphy Digitization Procedures Guideline", Taipei City: Taiwan Digital Archives Expansion Project, April 2009

Table 4-7 Image File Specifications of Hwa Kang Museum

	Replica	Permanent reservation	Website	Preview
Format	TIFF	TIFF	JPEG	JPEG
Color Mode	RGB	RGB	RGB	RGB
Image Mode	48bit, 300-600dpi	48bit, 300dpi	24bit, 150dpi	24bit, 72dpi
Application	Not compressed;	Not compressed; high	File size not exceeding	GIF index color
Format	image quality close	quality image	200KB; Dimensions	compression
Purpose	to the original		of 1024x768	
	For prints and publications	Preservation of source data	Internet applications and file conversion	Rapid preview

Source: Kao Lang-Hsuan, Chen Hsiu-Hua, "Painting and Calligraphy Digitization Procedures Guideline", Taipei City: Taiwan Digital Archives Expansion Project, April 2009

2. Selecting Audiovisual Formats

Unlike text data, audiovisual data can chose from a large variety of formats, but following the continuous development of existing formats and new formats, incompatibility between different formats has become a commonly encountered issue. Therefore, the purpose of digitization and future applications should be considered to select suitable audiovisual formats. In addition, audiovisual files require more storage space, which increases cost of storage hardware, so storage format should be a key consideration when funds are limited.

Recommendations for digital file formats of audiovisual data are divided into two categories, audio data and video data.

(1) Audio data

If the source data is a cassette, CD or LP, convert it into a digital file, store a permanent copy in an uncompressed format, and then convert the file into other formats according to requirements. Recommendations are as shown in the two tables below:

Table 4-8 Audio File Reference Specifications-1:

Purpose	Permanent Preservation	Audio Streaming (Online Browsing)
Description	Retains the original appearance of digital archives, and provides the archives to users for reproduction, compression or other image processing and exchange purposes	For users to listen online
Recommended File Format	File Format: WAVE (Microsoft format) Sampling Rate: 44.1 kHz/16~24 bit/ sample Channel: Stereo	File Format: WMA or Mp3 Sampling Rate: 44.1 kHz / 16bit/ sample Channel: Stereo Bandwidth: 64Kbps-128 Kbps File Time: The first 30 to 60 seconds or the complete audio file

Source: National Repository of Cultural Heritage

Table 4-9 Audio File Reference Specifications-2:

Purpose	Description	Recommended File Specifications
Description	Retains the original appearance of digital archives, and provides the archives to users for reproduction, compression or other image processing and exchange purposes	File Format: WAVE (Microsoft format) Sampling Rate: 44.1 kHz/16~24 bit/sample Channel: Stereo
Recommended File Format	Allows broadband users to listen to audio files online	File Format: MP3 Sampling Rate: 44.1KHZ Channel: Stereo Bandwidth: 128kbps
		File Format: RA Coding Type: Two-channel ISDN Stereo Sampling Rate: 16KHZ Bandwidth: 100Kbps

Purpose	Permanent Preservation	Audio Streaming (Online Browsing)
Streaming Narrow Band Format	Allows narrow band users to listen to audio files online	File Format: MP3 Sampling Rate: 22.05KHZ Channel: Stereo Bandwidth: 56kbps
		File Format: RA Coding Type: ISDN Stereo Sampling Rate: 8KHZ Bandwidth: 50Kbps

Source: National Digital Archives Program Technology Collection
http://www2.ndap.org.tw/eBook/showContent.php?PK=157

(2) Video data

If the source data is a video tape, then convert it into a digital file using video extraction software, and store it in the following file formats. Main reference specifications include:

Table 4-10 Video File Reference Specifications-1:

Purpose	Permanent Preservation	Video Streaming (Online Browsing)
Description	Retains the original appearance of digital archives in DVD quality, making it suitable for preserving high resolution objects	For users to view online
Recommended File Format	File Format: MPEG4 or above Video Dimensions: 1920*1080 pixels or above Audio decoded into two-channel stereo Data Rate: 8Mb/sec	File Format: WMV Video Dimensions: 320*240 pixels Data Rate: 150-300kbps File Time: 1-5 minutes or complete (shows main features)

Source: “National Repository of Cultural Heritage” and “Taiwan Digital Archives Expansion Project – Multimedia Thematic Group Digitization Specifications

Table 4-11 Video Format Reference Specifications-2:

Purpose	Description	Recommended File Specification
Permanent Preservation Format (1)	Retains the original appearance of digital archives in VCD quality	File Format: mpg Compression Method: MPEG-1 Dimensions: 352x240 pixels Video Decoding Frame Rate: 29 fps or above Audio Decoded into Two-channel Stereo CD Data Rate: Approx. 1.2 megabits/sec (approx. 150 kilobytes/sec) Jitter Standard Frame Time: 9 mSec or better Average Synchronous Displacement: 1 mSec or better Standard Device Synchronous Displacement: 15 mSec or better Audio Data Streaming (nChannels): 2 or higher Audio Data (nSamplesPerSec): 44100 or higher
Permanent Preservation Format (2)	Retains the original appearance of digital archives in DVD quality; file size is 3~100 times that of format (1), making it suitable for preserving high resolution objects	File Format: mpg2 Compression Method: MPEG-2 Dimensions: 720x480 pixels Video Decoding Frame Rate: 29 fps or above Audio Decoded into Two-channel Stereo AC3 Data Rate: Approx. 6~8 Megabits/sec (approx. 150kilobytes/sec) Jitter Standard Frame Time: 9 mSec or better Average Synchronous Displacement: 1 mSec or better Standard Device Synchronous Displacement: 15 mSec or better Audio Data Streaming (nChannels): 2 or higher Audio Data (nSamplesPerSec): 44100 or higher

Purpose	Description	Recommended File Specification
Streaming Broadband Format	Allows broadband users to view videos online	File Format: mpg4 Compression Method: MPEG-4 Audio Decoded into MP3 Bit-rate: 250Kbps
		File Format: RM Bit-rate: 250Kbps Audio Format: 22 kHz, stereo
		File Format: ASF Bit-rate: 250Kbps Audio Format: 22 kHz, stereo
Streaming Narrow Band Format	Allows narrow band users to view videos online	File Format: mpg4 Compression Method: MPEG-4 Audio Decoded into MP3 Bit-rate: 56Kbps
		File Format: RM Bit-rate: 56Kbps Audio Format: 11 kHz, mono
		File Format: ASF Bit-rate: 56Kbps Audio Format: 11 kHz, mono

Source: National Digital Archives Program Technology Collection
http://www2.ndap.org.tw/eBook/showContent.php

Each institution selects specifications according to their usage purpose and storage method. The table below shows specifications used by the "Taiwan Social and Humanities Video Archive", in which different standards are adopted for different purposes. However, considering the speed of developments in information technology, updating technologies and storage work are issues that should not be overlooked.

Table 4-12 Specifications of the "Taiwan Social and Humanities Video Archive"

	Content	Description
1	Permanent Preservation MPEG-2	File Format: MPEG-2 Video Dimensions: 704 X 480 Video: 15Mbps Audio: 192 Kbps Stereo Music Frame Rate: 30 fps
2	Regional Display RM and WMV	File Format: RM and WMV Video Dimensions: 704 X 480 Video: 1,500 Kbps Audio: 96Kbps Stereo Music Frame Rate: 30 fps

	Content	Description
3	Multiple Bandwidth Display RM and WMV	File Format: RM and WMV Video Dimensions: 352 X 240 Video: 300Kbps, 100Kbps, 34Kbps Audio: 64Kbps Stereo Music, 6Kbps, 20Kbps Mono Music Frame Rate: 15 fps
4	Online Browsing RM and WMV	File Format: RM and WMV Video Dimensions: 352 X 240 Video: 300Kbps, 100Kbps, 34Kbps Audio: 64Kbps Stereo Music, 6Kbps, 20Kbps Mono Music Frame Rate: 15 fps

3. Selecting Metadata Formats

Another fundamental task in preliminary procedures of digital archiving is "metadata requirements evaluation and analysis." Depending on the type of collection being digitized, different metadata standards may be adopted; the selection of metadata standards aims to truly present collection contents to users.

Metadata not only helps users to understand digital resources, but also allows owners and administrators to more effectively manage digital resources, keep track of digital resources, and establish data preservation and data search mechanisms. From another aspect, metadata standards allow collections to be consistently, structurally and comprehensively described, so that digital resources are interoperable, and can be shared across institutions and countries, making space for data to be exchanged and applied.

(1) Metadata Types

Metadata refers to structural standardized background data, which has three types: management, structural and descriptive, and three attributes: semantic, syntax, and lexical. Metadata referred to in literature today are descriptive data, which are descriptions of collections or specific objects, and describe the contents and links of a document or resource so that resources can be indexed, found and identified.

According to the scope of its usage, metadata can be divided into two categories, general and specific. General metadata standards are characterized by being simple, expandable, cross disciplines and cross data types, and represented by the Dublin Core, which emphasizes on resource discovery in digital archives. The Dublin Core has 15 basic elements, has been widely adopted by countries around the world, and is currently the standard used by the TELDAP union catalogs. Specific metadata standards are characterized by being able to provide in-depth description of a specific discipline or field, and emphasizes on common requirements and cataloging standards of a specific field. For example, the Darwin Core is a metadata standard

used by the field of biology, and consists of 50 elements that describe digitized biological materials.

Below we will use commonly used international metadata standards that have been adopted by digital archive projects.

(2) Commonly Used International Metadata Standards

i. DC

The Dublin Core (DC) was a seminar jointly sponsored by the Online Computer Library Center (OCLC) and National Center for Supercomputing Applications (NCSA) in 1995. Its members included scholars and specialists in libraries, computers, the internet and other professional fields. Its purpose was to establish a set of elements to describe electronic documents on the internet, and further help the finding of information. At present, more complicated frameworks developed based on the DC can be found all over the world in libraries, museums, government entities and commercial organizations.

DC main feature is that it establishes a cross-disciplinary element set. The "simple" DC has 15 elements and emphasizes on discovering resources in digital archives. Elements of DC have the following five characteristics: [22]

A. Each element is optional and can be repeatedly cataloged.

B. Elements can appear in any sequence.

C. Cataloging rules are recommended, not required.

D. Cataloging rules are established by the institution using it.

E. Adopting application files that belong to a specific field is encouraged; data contents and data values specified by these application files are followed.

The Union Catalogs established by TELDAP uses a single platform to perform search functions; the platform was developed based on the Dublin Core, in which 15 core elements are as follows:

Title: A name given to the resource.

Creator: An entity primarily responsible for making the resource.

Subject & Keywords: The topic of the resource.

Description: An account of the resource.

Publisher: An entity responsible for making the resource available.

Contributor: An entity responsible for making contributions to the resource.

Date: A point or period of time associated with an event in the lifecycle of

[22] TELDAP WIKI: http://wiki.teldap.tw/index.php, Search: January 2010.

the resource.

Resource Type: The nature or genre of the resource.

Format: The file format, physical medium, or dimensions of the resource.

Resource Identifier: An unambiguous reference to the resource within a given context.

Source: A related resource from which the described resource is derived.

Language: A language of the resource.

Relation: A related resource.

Coverage: The spatial or temporal topic of the resource, the spatial applicability of the resource, or the jurisdiction under which the resource is relevant.

Rights Management: Information about rights held in and over the resource.

Website: http://dublincore.org/

ii. CDWA

Categories for the Description of Works of Art (CDWA) is a standard developed by the Art Information Task Force (AITF), which is subordinate to Getty Trust. CDWA describes the content of art databases by articulating a conceptual framework for describing and accessing information about works of art, architecture, other material culture, groups and collections of works, and related images. CDWA includes 381 categories and subcategories, basic cataloging principles and examples. A small subset of categories are considered core in that they represent the minimum information necessary to identify and describe a work. CDWA Lite is an XML schema to describe core records for works of art and material culture based on CDWA and CCO (Cataloguing Cultural Objects).

Website: http://www.getty.edu/research/conducting_research/standards/cdwa/

iii. EAD

EAD (Encoded Archival Description) is a document type definition. Development of EAD began with the "Berkeley Finding Aid Project (BFAD)" initiated by the University of California, Berkeley, Library in 1993. It became a standard after the Society of American Archivists (SAA) joined in its development in 1998. EAD is currently maintained by the Network Development and MARC Standards Office, which is subordinate to the Library of Congress, and the SAA. EAD is a commonly used metadata standard for archives. Its purpose is to aid the digitization, search and representation of archives and manuscript data; it is a standard machine-readable description format that

facilitates the acquisition of manuscripts and archives online.

Website http://www.loc.gov/ead/

iv. VRA

VRA established by the Data Standards Committee of the Visual Resources Association. The committee is responsible for promoting standards that facilitate the management and exchange of information; members include specialists in the field of digital images, movies, photography, museums and publishing. Elements of VRA aim to facilitate the sharing of information on works of visual culture, and can be considered as the DC of visual resources, which include works and images. VRA adopts the same one to one principle as DC, meaning that one set of metadata only describes one object or work, in which elements can be arranged in any order; the latest version is VRA4.0.

Website: http://www.vraweb.org/projects/vracore4/

v. Species 2000 Data Standard

The Species 2000 project was initiated by the International Union of Biological Sciences in September 1994, and is jointly maintained and developed by technical teams around the world. The motive for establishing the Species 2000 database is to create a validated checklist of all species in the world. This is being achieved by bringing together an array of global species databases covering each of the major groups of organisms, allowing users to rapidly find complete data on a species by entering its name.

Website: http://www.sp2000.org/index.php

vi. DwC

The Darwin Core is a product of the Species Analyst project implemented by the University of Kansas with the purpose to describe living collections and natural history collections. DwC tries to provide common records of specimen under the premise of not considering the content storage mechanism, using a relatively simple method to support data search and the extraction of descriptive information. The latest version is DwC V2.0, which has metadata definitions for 48 elements.

Website: http://wiki.tdwg.org/twiki/bin/view/DarwinCore/WebHome

vii. CSDGM

Developed by the Federal Geographic Data Committee (FGDC), the Content Standards for Digital Geospatial Metadata was released in June 1994 and provides a basic framework for metadata applications. Its purpose is to build a foundation for geospatial data, and to describe contents, quality and access methods of GIS information as a basis for data searches.

Website: http://metadata.teldap.tw/project/project-frame.html

Table 4-13 Commonly Used International Metadata Standards:

Metadata	Example	Introduction
CDWA	Painting and Calligraphy, Rubbings, Slips, Artifacts, Architecture, Relics, and Sites	The purpose of this standard is to develop an information framework for describing works of art (e.g. paintings, carvings, ceramics, architectures, furniture and performing arts) and images.
EAD	Archives, Records	EAD is a finding aid standard that was developed to support the collection and preservation of archives and manuscript repositories, providing a permanent coding standard that allows catalogs, checklists and indexes produced by various institutions (e.g. archives, libraries, museums and manuscript repositories) to be machine readable, which make resources easier to access online.
HISPID	Botany	HISPID is a standard that focuses on data exchange. Information that is exchanged includes categories, technical terms, lists, specimen models, and rare and extinct plants, which covers activities and issues of specimen repositories and botanical gardens. Furthermore, HISPID also attaches importance to the relation with the database structure, reducing obstacles of data exchange operations.
LOM	e-Learning, Learning Resources	LOM is a standard that defines the syntax and semantics of metadata for learning resources. LOM emphasizes on the management, orientation and evaluation of learning resources, providing nine categories and basic elements for applications and development by individual information systems.
Marc	Rare Books	Machine Readable Cataloging (MARC) is more accurately a family of cataloging formats used by libraries instead of a single standard. MARC originated in 1965~6 when the Library of Congress developed a set of standardized data records (LC MARC) to benefit the exchange of catalogs between libraries. It later was adopted by countries around the world and revised according to national conditions, e.g. USMARC (U.S.A.), UK MARC (England), CAN/MARC (Canada), and AUSMARC (Australia).

Metadata	Example	Introduction
SPECIES 2000 VRA	Zoology	Species 2000 uses a set of standard data fields to describe all known species. These data fields were simplified from Global Species Databases (GSD is an array of global species databases covering each of the major groups of organisms), so they are compatible to data fields of numerous different databases. When using Species 2000, users can access information on a specific species by simply entering its name.
SPECIES 2000 VRA	Photographs	This standard focuses on visual resources communities, and provides guidance for describing visual resources, such as artworks, architectures, antiques, and cultural relics. To facilitate information sharing and exchanges between visual resources communities, VRA's design emphasizes on simple elements (17 in total) that are flexible and reusable, and provides recommendations for data values to achieve the consistency and correctness of information.

Source: Digital Archive e-Learning: http://ic.shu.edu.tw/DA/DADL/dam_03_metadata.htm

(3) Metadata Formats of the Six Themes

Metadata standards for each category in the six themes of TELDAP are as shown in the table below:

Table 4-14 Metadata Formats of Six Classification for Different Themes

Theme	Category	Metadata Cataloging Standard	Example
Biology and Nature	Zoology	1. ABCD Schema (Access to Biological Collection Data Schema) 2. Darwin Core (DwC) (Species Analyst 2001) 3. FGDC Biological Metadata Profile 4. UCMP (University of California Museum of Paleontology) Data Model (UCMP1994) 5. Species 2000 Standard Data	Zoological Research of Taiwan: Fish – adopted Species 2000 Standard Data as the reference standard for its metadata architecture
	Botany	1. ABCD Schema (Access to Biological Collection Data Schema) 2. Darwin Core (DwC) (Species Analyst 2001) 3. FGDC Biological Metadata Profile 4. FGDC Vegetation Classification and	

Theme	Category	Metadata Cataloging Standard	Example
		Information Standards 5. Plant Names in Botanical Databases (Bisby 1994) 6. UCMP (University of California Museum of Paleontology) Data Model (UCMP1994)	
	Geology	Paleontology related standards: 1. UCMP (University of California, Museum of Paleontology) 2. The Paleobiology Database Mineral related standards: 1. Mineralogy Database 2. The Mineral Galle Rock related standards: 1. Tom Weiland's Homepage 2. The Stupid Page of Rocks	
Life and Culture	Anthropology	1. CDWA (The Categories for the Description of Works of Art)	
	Archeology	1. CDWA (The Categories for the Description of Works of Art)	
Language and Multimedia	Language	1. OLAC Metadata standard	
	Audiovisual	1. The ECHO Metadata Modeling Report WP3 (2000) (IST-1999-11994)	
	News	1. NewsML (News Markup Language) 2. News Records Metadata Format)	
Art and Images	Antiquities	1. CDWA (The Categories for the Description of Works of Art)	The Hwa Kang Musuem adopted CDWA as the reference standard for its metadata architecture
	Painting and Calligraphy	1. CDWA (The Categories for the Description of Works of Art) 2. VRA (Core Categories for Visual Resources) 3. REACH (for Shared Description of Museum Objects) 4. SPECTRUM (The UK Museum Documentation Standard) 5. CIDOC (The International Committee for Documentation)	
Maps and Architecture	Maps and Remote Sensing Images[23]	1. CSDGM (Content Standards for Digital Geospatial Metadata) 2. ISO TC211 Geographic information and geomatics related standards	

[23] Source: Yang Yi-Ling, Lin Yen-Hung, "Maps and GIS Digitization Procedures Guideline", Taipei: National Digital Archives Program – Content Development Division, 2006.

Theme	Category	Metadata Cataloging Standard	Example
		3. CEN TC287: Geographic Information – Data description – Metadata. Draft V2 - for 2nd informal vote by WG 2, January 1996. 4. CERCO/MEGRIN: GDDD – Geographical Data Description Directory 5. IHO: S57, version 3. 6. ISO / TC 211: ISO 19115:2003 Geographic information – Metadata, 2003-05-08, 140pages. 7. NATO DGIWG/DIGEST: N469 Geographic Information – Data description – Metadata. 8. R.O.C.: NGIS related standards	
	Architecture	1. CDWA (The Categories for the Description of Works of Art)	
Literature and Archives	Archives	1. EAD (Encoding Archival Description) 2. BAC (Business Acceptable Communications) Metadata 3. National Archives of Australia Recordkeeping Metadata Standard 4. SPIRT RKMS, Australian Recordkeeping Metadata Schema 5. ISAD (G) General International Standard Archival Description 6. ISAAR (CF)International Standard Archival Authority Record for Corporate Bodies, Persons and Families 7. MARC 21	Projects that adopted EAD as the reference standard for their metadata: 1. Digital Archives of Famous Personages, Diplomatic Records, and Economic Records in Modern China 2. Digital Archives of Historical and Cultural Relics collected in the Institute of History and Philology – the Grand Secretariat Archives
	Rare Books	1. MARC	
	Metal and Stone Rubbings	1. CDWA (The Categories for the Description of Works of Art)	
	Chinese Full Text	1. TEI markup (Text Encoding Initiative)	

Source: Digital Archive Technology Collection 2006 Version

IV. Application Systems for Digital Images and Metadata

After selecting formats for digital objects, in order for the digital archive system to effectively operate, system development and establishment requires the joint effort of information technicians and content specialists to reap in twice the results. Establishment of an application system that combines digital images with metadata should be planned at the beginning of operation procedures. Key operation procedures of system planning are briefly described below:

1. System Establishment Evaluation

Besides system considerations, metadata planning and design should also be considered when establishing a digital archive system. Therefore, the establishment of a digital archive system includes system planning, metadata analysis, system analysis and design, and system development.[24]

(1) System Planning:

All aspects should be thoroughly considered based on requirements of the project institution. Plan suitable system formats, function descriptions, make a preliminary evaluation, and produce a system specification.

(2) Metadata Requirements Analysis:

Institutions implementing digital archive projects should seek advice from metadata experts to fully understand which metadata standards that are suitable for their collections, draft function requirements, make adjustments whenever necessary, and establish cataloging rules for institutions, metadata experts and system developers to use as a basis for effective communication.

(3) System Analysis and Design:

Before designing the database and search system, system developers should further communicate with institutions to analyze system functional requirements. System function specifications and design specifications are keys to ensure the correctness of operation procedures.

(4) System Development:

System developers develop system functions (e.g. data management, data search, full text search, database query, and browsing) according to the system function specifications, and integrate all digital images and metadata to establish a complete digital archive system.

2. System Development and Establishment

After meticulous planning and analysis, system developers transform institution's system requirements into technical specifications, and design suitable system functions and architecture based on system requirements with consideration to the institution's usage conditions. This consists of two major directions: database and search system design. Institutions need to adjust functions according to their own operations and management. Below are a few commonly used function designs in the system development process:[25]

(1) Warehouse management functions: Integrate management of physical collection with management of digital data.

[24] Tsai Yung-Cheng, Huang Kuo-Lun, Chiu Chih-Yi et al, "Introduction to Digital Archives Technology", Taipei City: NTU Publishing Center, November 2007, page 70.

[25] Same as 24, pages 83-85.

(2) System security mechanisms: Besides system protection and system restoration, also includes data backup and remote backup mechanisms.

(3) Digital rights management mechanisms: Data encryption, rights description, copy detection and tracing.

(4) System expandability and load balance mechanisms: To ensure system service quality, estimate the maximum number of users and data flow within the same time period.

(5) Data exchange functions: Establish data interoperability mechanisms to share research results.

Each establishment process involves continuous tests and corrections in coordination with data additions and corrections, before data can be input to complete the database and search system, and benefit future value-added applications.

3. Service, Innovative and Value-added Applications

Digital archive system not only serve to preserve valuable digital resources, but also provide exhibition, research and education applications, e.g. digital library, public exhibition system and digital museum, depending on the requirements of different user communities.[26]

(1) Digital library: Provides digital data to users for research and education activities. For example, the TELDAP union catalog gathers digital archiving results different institutions, and establishes a single platform for browsing and searching for information.

(2) Public exhibition system: Introduces digital archiving results to the general public; exhibitions aim to be intuitive, fun, simple and interactive. For example, the TELDAP portal repackages professional and profound digital archives, and introduces them to the general public.

(3) Digital museum: Combines features of digital libraries and public exhibition systems, and extends museum functions into the digital world. For example, the National Palace Museum (http://www.npm.gov.tw/) extends its physical museum into the digital world, and provides data search and interactive information to satisfy the general public's research and education requirements.

Besides the three main functions described above, digital archive application systems still have space for value-added applications, including search system, website services, licensing, and innovation services. Search system and website services are generally applied in digital libraries, digital museums, and public exhibitions systems. As for licensing and innovation services, besides academics and education, applications have extended into new fields, such as commerce. Therefore, institutions should also establish a resource management system for digital data preservation and digital resource

[26] Same as 24, pages 350-353.

licensing. The following section will briefly describe important topics of digital data management, as well as permanent preservation of digital resources.

V. Permanent Preservation of Digital Archives

1. Meeting Requirements for Permanent Preservation

There are multitudes of institutions and projects taking part in TELDAP. Years of project implementation have accumulated a considerable amount of digital content, not to mention the great amount of implementation experiences, digital archiving talents that have been cultivated, and network platform technology and applications that have been developed. From the perspective of Taiwan's long-term digital archive policy, these resources are important common assets in our information society, and the key to boosting developments and resolving digital divide. For these reasons, TELDAP has been devoted to the permanent preservation and sustainable operation of digital archiving results, providing a unified interface for long-term preservation of digital data, and developing permanent preservation methods that do not become obsolete as time passes or as technologies, software/hardware equipment, and storage media progress.

Although digital archives have a great advantage in being convenient, provide search functions, and take small storage space, they have also caused quite a few problems. Especially after the continuous production of digitized or born-digital data, information security and preservation issues are entirely different from conventional paper copies. Digital information is easily destroyed, and therefore digital information must be further researched, evaluated and planned.[27]

When evaluating and planning permanent preservation, keep the following considerations in mind:[27]

(1) Storage media: Storage media lack durability and deteriorate rapidly, careless operation can easily do damage.

(2) Storage technology: New technologies appear at ever-growing speeds to replace old ones, but digital data requires both software and hardware to be read, so this also needs to be considered.

(3) Characteristics of digital information: Digital information is easy to disseminate and duplicate, preventing its authenticity from being guaranteed as a result. Digital information is easily damaged and lost during file conversion, refresh or operations.

Kranch believes that long-term preservation methods of digital information include: preservation of original technologies, transfer to new technologies, and transfer to paper or other mediums that can be browsed.[28] Muir's preservation strategies include:

[27] Ou-Yang Chung-Jung, "Discussion on Digital Information Preservation", "Archives Quarterly", 1(2), 2002, pages 36-47.

[28] Kranch, D.A.(1998). Preserving electronic documents, Proceedings of the third ACM conference on Digital Libraries: 295-296

technology preservation, migration and emulation.[29] According to Lawrence et al., migration is the most suitable preservation method;[30] however, after an in-depth exploration of emulation methods, Rothenberg believes that emulation is the most suitable method.[31]

Commonly used or discussed preservation methods include refreshing, migration, emulation, standardization and technology preservation.

(1) Refreshing

Information technology and storage media are developing at ever-growing speeds and it is possible that one day the software and hardware required for reading the storage media will no longer exist. Refreshing is adopted to avoid not being able to read data in storage media. Refreshing refers to storage media refreshing, which is the transfer of digital files from old storage media to new storage media, and is currently the most popular and simple preservation method. However, there are certain limitations to this strategy, e.g. refreshing must be done on a regular basis, or else no one can guarantee that the refreshed storage media will be usable. Furthermore, specific software and hardware are still required for reading data after refreshing, so format limitations, compatibility of software and hardware, and decreasing storage quality of media as time passes by are all factors that need to be considered. The concept of refreshing is easy to understand, and relatively simple in its implementation, but does not guarantee information can be searched and processed in the future.[32]

(2) Migration

Migration is the regular transfer of digital files from old software/hardware to new software/hardware, or from old information technology to new information technology; all files operate under a new system after migration. There are two purposes of migration: (1) To preserve the completeness of digital objects; (2) To allow users to search, display and use digital information and resources.[33] Still, migration faces several issues in planning and implementation: for example, we cannot predict when new technologies will appear, which technologies will become the mainstream, what the storage media of new technologies will be, and whether or not they will be compatible with existing systems; preventing data loss during migration is also another issue that needs to be considered.[34]

---

[29] Muir, A. (2001). Legal deposit of digital publications: A review of research and development activity, Proceedings of the first ACM/IEEE-CS joint conference on digital libraries: 165-173

[30] Lawrence, Gregory W., William R. Kehoe, Oya Y. Rieger, William H. Walters, and Ann R. Kenney. Risk Management of Digital Information: A File Format Investigation, Council on Library and Information Resources, June 2000

[31] Rothenberg, Jeff. "Avoiding Technological Quicksand: Finding a Viable Technical Foundation for Digital Preservation," Council on Library and Information Resources, Jan 1999, pp. 1-35

[32] Chen Hsueh-Hua, Hung Wei-Ping, "Discussion on Long-term Preservation of Digital Information and Resources", Taipei: Student Book Co., published in 2004.

[33] Chen He-Chin, "Discussion on Metadata and Digital Archives", "University Library Journal", 5 : 2, 2001, pages 2-11.

[34] Chou Hsin-Ying, "Maintenance and Preservation of Digital Collections" "Shu Yuen", Issue 51, 2002, pages 72-82.

Migration and refreshing are somewhat different; refreshing cannot guarantee that data is completely copied, and is more prone to compatibility issues, but saves time, money and is easier to implement; migration retains the original appearance of digital resources and provides them for future usage. The main difference between migration and refresh is that the former includes data format, structure or standard conversion, while the later refers to the refreshing of storage media.[35]

(3) Emulation

"Emulation" refers to the process of using software to emulate operations of old systems on new ones. This ensures that data is not lost, maintains consistency of the appearance and behavior of digital files, and allows digital files to operate under the same conditions, even in new equipment and environment. Yet, emulation technology still has its issues, and although it allows digital files to operate under its original conditions, it cannot preserve all documents related.

(4) Standardization

There are a great number of digitization standards, and numerous formats and standards are related to just the creation and application of digital files, including: database, coding standards, metadata, digitization format and search, etc. Standardization focuses on standard formats of data, and therefore considers whether or not formats are still applicable, whether or not they have been preserved, and potential damage, loss, or reduced meaning of digital data due to the adoption of new formats.

(5) Technology Preservation

This is similar to a computer museum in that besides preserving information on hardware for emulation, it also preserves the hardware itself. This method preserves necessary outdated software and hardware, and stores digital files in their most primitive format. All files and equipment operate under their original environment, including applications, operating systems and platform environments.

2. OAIS and Permanent Preservation

OAIS is an archive placed in between a Producer, Consumer and Management, as shown in Fig 4-5.

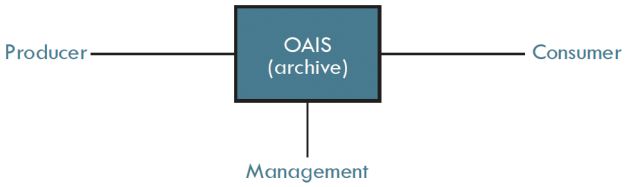


Fig 4-5 OAIS Reference Model[36]

[35] Ou-Yang Chung-Jung, "Discussion on Digital Information Preservation", "Archives Quarterly", 1(2), 2002, pages 36-47.

[36] Source: Sawyer, D., Reich, L., Giaretta, D., Mazal, P., Huc, C., Nonon-Latapie M., et al. (2002). The Open

The OAIS reference model supports both digital objects and physical objects, as well as data objects, which are combinations of the first two kinds of objects.

To achieve data preservation, the OAIS reference model defines the concept and structure of three "information packages":

(1) SIP: Submission Information Package: In OAIS, SIP is converted into AIP for long-term preservation.

(2) AIP: Archival Information Package: Packages produced for preservation requirements; contains complete content information and preservation descriptions.

(3) DIP: Dissemination Information Package: Packages produced by user requests; provides information packages required by users.

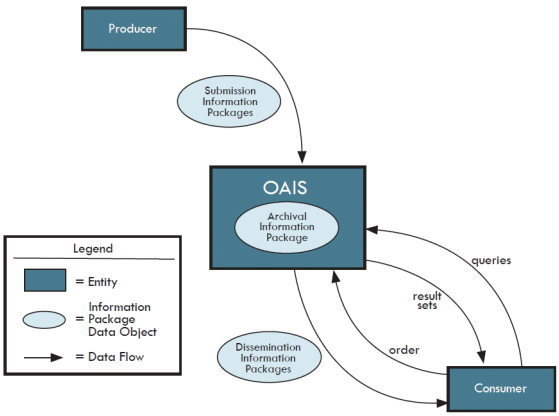


Fig 4-6 Operation procedures of OAIS[37]

OAIS operation procedures: The Producer generates SIP and passes it to Ingest; Ingest converts the SIP into AIP and passes it on to the Archival Storage, while descriptive info of SIP is passed on to Data Management. When Consumers use the correct descriptive information and search tool to find data, the Archival Storage sends AIPs that match the request to Access for conversion into DIP and further on to the Consumer.[38]

Archival Information System (OAIS) Reference Model and its Usage. Retrieved Dec. 3,2008, from http://www.aiaa.org/Spaceops2002Archive/papers/SpaceOps02-P-T5-39.pdf

[37] Source: Sawyer, D., Reich, L., Giaretta, D., Mazal, P., Huc, C., Nonon-Latapie M., et al. (2002). The Open Archival Information System (OAIS) Reference Model and its Usage. Retrieved Dec. 3, 2008, from http://www.aiaa.org/Spaceops2002Archive/papers/SpaceOps02-P-T5-39.pdf0

[38] Source: Sawyer, D., Reich, L., Giaretta, D., Mazal, P., Huc, C., Nonon-Latapie M., et al. (2002). The Open Archival Information System (OAIS) Reference Model and its Usage. Retrieved Dec. 3, 2008, from

Structural descriptions and records developed to manage information of digital objects for the purpose of long-term preservation and access of digital resources is the concept called Preservation Metadata, and benefits data management operations, such as access, control, preservation and migration. Unlike description metadata of MARC and Dublin Core, which are used to find and identify digital objects, Preservation Metadata mainly aids information management and access of digital content.39

OAIS defines functions and requirements of databases, provides a framework for interoperable metadata, and allows different types of metadata to be exchanged and reused. This model ensures that required data for preservation is included. Therefore, many projects adopt Preservation Metadata of the OAIS reference model.[40]

Although digital objects are easy to create, correct and disseminate, its storage media is not as stable as conventional media, e.g. paper. Therefore, if digital archiving results are to continue to be accessed and preserved, information on technologies and conditions that digital objects are created and used must also be preserved. Preservation Metadata is an important method for the preservation of digital object technology information, and supports basic strategies of long-term preservation, such as Refresh, Migration, Emulation, Standardization and Technology Preservation. It can also be considered as a reference standard provided by OAIS for permanent preservation.

Therefore, the definition of an Open Archival Information System (or OAIS) is an archive, consisting of an organization of people and systems, that has accepted the responsibility to preserve information and make it available for a Designated Community (ISO, 2003). In other words, OAIS achieves long-term preservation and maintenance via coordination between people and systems.

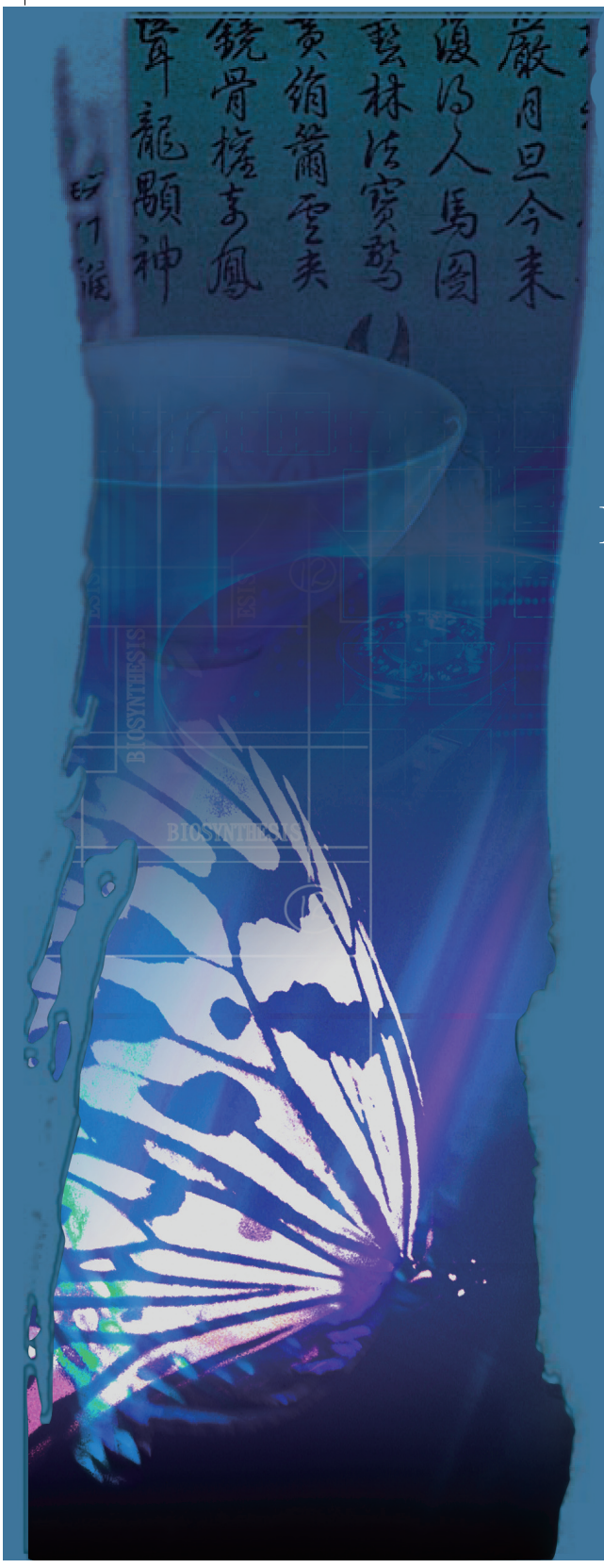
To truly achieve permanent preservation, institutions can develop preservation methods for their digital archive projects based on policies established by the administration, and ensure that administrators and users can access OAIS information. When planning preservation, also factor in the impact of information technology on digital archive storage; make long-term preservation plans.

---

http://www.aiaa.org/Spaceops2002Archive/papers/SpaceOps02-P-T5-39.pdf

[39] National Library of Australia, Preserving Access to Digital Information(PADI) , “Preservation Metadata”, (http://www.nla.gov.au/padi/topics/32.html

[40] Chang Huai-Wen, “Metadata and Long-term Preservation and Access of Digital Archives – Introduction to Preservation Metadata”, Search: January 2010 http://www2.ndap.org.tw/newsletter06/news/read_news.php?nid=352

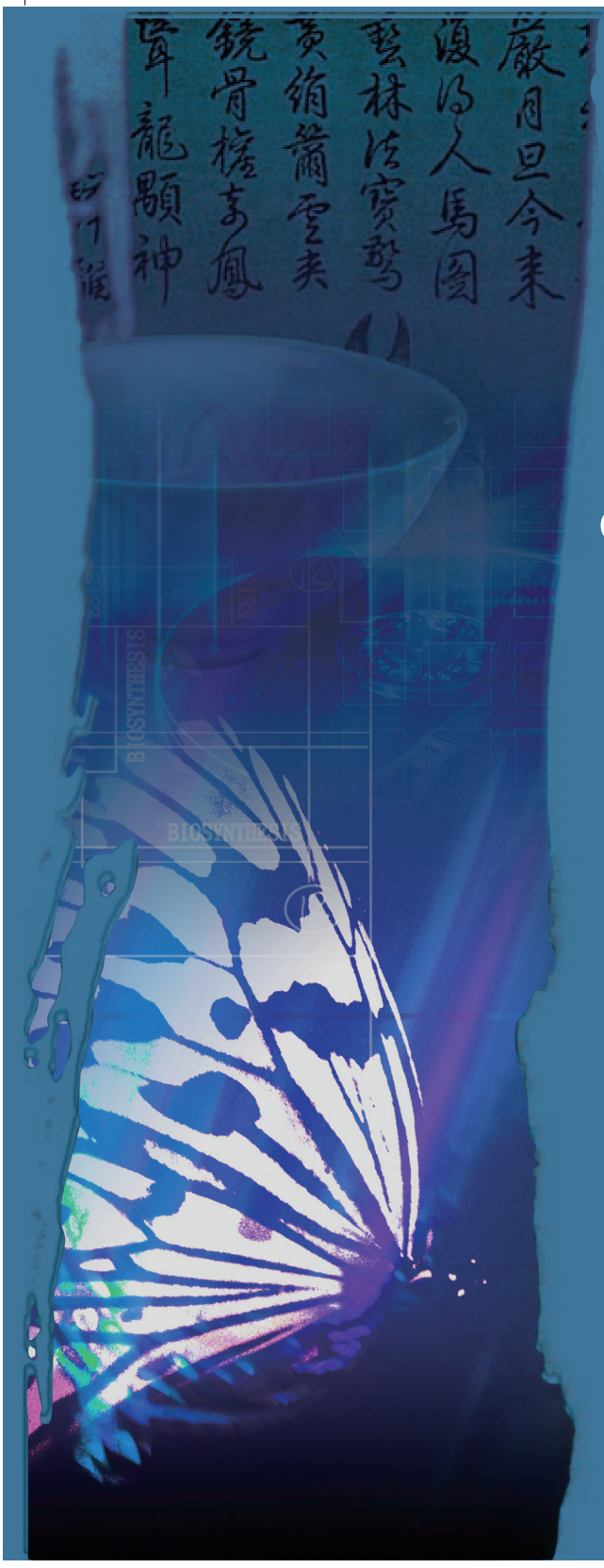


FIVE. Conclusions

Each digital archive project may produce different results due to different personnel, time, technology, digitization objects, and implementation methods. Although different object types have different detailed digital workflows. In fact, we still can find common operation procedures. The purpose of compiling a guideline on integrated operation procedures is to provide system development related operation procedures, methods, technologies and specifications for reference, so that institutions engaged in digital archive work can effectively grasp key points, adopt simple and efficient implementation methods, understand methods and technologies, control operation quality, enhance working efficiency, and elevate digitization results.

The standardization of digitization procedures is an important link to quality and efficiency. Digital workflows handle digital information, and are therefore confined to the life cycle of digital information. Each stage of the digital information life cycle (data creation, data management, data preservation, and data provision) corresponds to a part of digital archiving work, which is why the digital information life cycle can be used to examine the completeness of operation procedures. Using the ISO OAIS reference model as an example, besides conforming to the life cycle of digital information, it also has specific planning strategies for "administrative management" and "preservation plans"; these two major procedures are important issues faced by digital archiving work.

Besides adopt a digital information life cycle approach, this guideline also refers to concepts of the OAIS model and attempts to sequentially describe specifications and implementation methods related to each procedure. We hope that this guideline can provide a core planning blueprint to institutions or individuals who are devoted to digital archiving work, so that they can reduce unnecessary mistakes. At the current stage this guideline can only serve as an example for reference, we will continue work towards establishing standard operation procedures, and hope that readers will offer their opinions or constructive criticism and jointly contribute to digital archiving work.



Glossary

Integrated Operation Procedure	Reference standards for system development related operation procedures, methods, technologies and specifications; guides the planning of unified and common operation procedures. Its purpose is to provide a simple and effective introduction of work contents, to explain the order of which procedures should be executed, to help understand methods and technologies, to maintain operation quality, to enhance production efficiency, and to elevate development results.
Digital Information Life Cycle	Digital materials, like human beings, ecology and enterprises, have life cycles. Studies have generalized stages of digital information life cycle into data creation, data management, data preservation, and data provision.
Data Creation	Data creation is the stage in which files and information are created, and mainly involves the collection, acquisition, production, and acceptance of various digital objects. Main tasks of this stage include data gathering and digitization.
Data Management	Mainly involves the management of metadata, including the establishment and planning of metadata creation, categorization, definition, documents and storage, access rights, and intellectual property rights.
Data Preservation	Data preservation involves the form and environment of which data is stored, including data format, data migration, and whether data is completely or partially stored in a distributed or centralized system. Data preservation emphasizes on protecting data from being damaged by changes in time, space or information equipment, and ensuring that data can continue to be accessed, achieving "long-term" preservation.
Data Provision	Data provision is the stage in which files and information are accessed and shared, all other stages of the digital information life cycle exist to support this function.
Born-Digital Objects	Digital objects are generally viewed as the body of digital materials, its scope encompasses all digital publications and digital records; Born-digital objects refer to objects that were originally created in digital form, e.g. images produced from digital photography are born-digital files.
Digitized Objects	The opposite of "born-digital" is "digitized." Objects are changed into digital form via digitization technology,

	meaning that digital files are produced from physical objects (e.g. documents, artifacts, specimen, audio recordings, videos, architecture, etc.) or files stored in analog form; for example, digital files of film provided by traditional photography are produced via digital scanning.
Color Model	A color model is a typical way of expressing colors as a set of numerical values. For example, the RGB color mode uses three values to represent colors; the first value is the content of Red (R), the second value is the content of Green (G), and the third value is the content of Blue (B).
Color Space	Color space is a mechanism that uses numbers to define the relationship between two color spaces, using a color mapping function to map color signals to different positions in color spaces that have the same color perceptions. For example, Adobe RGB and sRGB are two absolute color spaces based on the RGB color model.
Streaming	Streaming is a technology used for playing files real-time on the internet; data can be transferred section by section so that users don't have to wait for the complete file to be downloaded before they can start viewing it.
Non-Streaming	Non-Streaming refers to downloading audiovisual files and storing them in physical storage space before playing the files, meaning that there will be a file at the user end.
OAIS	OAIS (Open Archival Information System) was released in 2003 by the International Organization for Standardization (ISO) with support from the Consultative Committee for Space Data Systems (CCSDS). OAIS is an archive consisting of an organization of people and systems that recommends long-term preservation. The system model comprises six main functions: Ingesting, Archive Storage, Data Management, Access, Administration and Preservation Planning.
RAW	RAW is the source image file generated by digital cameras. RAW files are not compressed, can render rich image colors and layers, and allow adjustments to white balance and brightness; it is a good storage format if image files will be further applied in the future. The downside is that image processing requires software provided by the camera manufacturer or professional software (Adobe Photoshop,

	Adobe Photoshop Lightroom or Aperture). Still, RAW is gradually being adopted by digitization units as a means for permanent preservation because it holds the original image.
TIFF	TIFF is widely applied in different platforms and software. TIFF uses the LZW algorithm to achieve lossless compression, and is therefore suitable for preserving original images, which can be used for future processing or printing high resolution images. The downside is that TIFF has low compression rates, requiring longer time to open and save files; TIFF files are rarely used for download or upload.
JPEG/JPG	The compression rate is decided when JPEG/JPG (Joint Photographic Experts Group) images are saved; high compression rates will result in low image quality, while low compression rates will result in high image quality close to the original image. JPEG is not suitable for image preservation, which requires high image quality, because it is a lossy compression method. However, its highly effective compression method can create extremely small files that are suitable for browsing online.
GIF	GIF (Graphics Interchange Format) is a bitmap file format established by Compuserve. Its color model uses index colors, it possesses file compression ability, supports website animation functions, and is frequently used in webpage design.
RGB	The most commonly used color model; it is a full color model that comprises of the three primary colors red, green and blue.
sRGB	A monitor color space jointly established by HP and Microsoft; its Gamma value is set at 2.2, and color temperature is set at 6500; s represents Standard, another saying is that s represents the last name of its designer Mike Stokes; sRGB is now widely applied in PCs.
WAV	WAV (Waveform Audio Format) is an audio coding format jointly developed by Microsoft and IBM. It uses uncompressed waveforms of PCM coding so sounds are not distorted, but file size is relatively large among other audio formats; mainly used in Windows.
MP3	MP3 (MPEG-1 Audio Layer 3) is a link in the MPEG standard; it is a digital audio coding and lossy compression

	format. The highly efficient data compression and sound effects of MP3 coupled with convenient players and coding programs, have gradually increased the number of MP3 users and turned into a standard for multimedia files.
WMA	Windows Media Audio (WMA) was developed by Microsoft. Its main feature is that its file size is less than half of MP3, while providing equal sound quality.
MIDI	Musical Instrument Digital Interface (MIDI) was jointly established in 1983 by internationally renowned electronic music producers. MIDI is a type of digital interface, it is a unified format that tells electronic equipment how to make sounds, so software/hardware designed by different manufacturers are interoperable as long as they conform to the MIDI format.
Motion JPEG	Based on JPEG image compression; AVI files use this format.
MPEG	MPEG was proposed in 1988 by the Motion Picture Expert Group, which is under the ISO. MPEG is an audiovisual compression technology that reduces inconvenience brought by the massive amount of data of Motion JPEG. Depending on its purpose, related standards include MPEG-1, MPEG-2, MPEG-4, MPEG-7 and MPEG-21.
AVI	A video file format developed by Microsoft. Due to the popularity of Windows operating systems, AVI has become the most widely used audio/video file format.
QuickTime	A video file format developed by Apple in 1991 that can only be opened by QuickTime player. Its contents are the same as AVI, but QuickTime does not limit the compression format that must be used, it only defines the storage structure of videos.
SIP	Submission Information Package.
AIP	Archival Information Package.
DIP	Dissemination Information Package.

References

Books

王雅萍、林彥宏，《文書檔案數位化工作流程指南》，台北市：數位典藏拓展臺灣數位典藏計畫，2009年04月。

高朗軒，陳秀華 ，《書畫數位化工作流程指南》，台北市：數位典藏拓展臺灣數位典藏計畫，2009年04月。

李佩瑛、王雅萍、高朗軒 ，《數位化工作流程指南:色彩管理》，台北市：數位典藏拓展臺灣數位典藏計畫，2009年04月。

蔡永橙、黃國倫、邱志義等著，《數位典藏技術導論》，台北市：台大出版中心。2007年11月。

楊易玲、林彥宏，《地圖圖資數位化工作流程參考標準》，台北市：數位典藏國家型科技計畫—內容發展分項計畫，2006年。

歐陽崇榮，《數位資訊保存策略》，臺北市：文華圖書館管理。2008年3月初版二刷。

Journal Papers

David Jillings, “From Two to Three Dimensions- Building on Experience of 2D Digitization of Prepare For The 3D Digital World”, International Conference of Museum Digitization Antiquities, Painting and Calligraphy, Taipei, 2003.

Hwai-ling Yeh-Lewis, “Then and Now: Manual Cataloguing system vs. Digitization System”, International Conference of Museum Digitization Antiquities, Painting and Calligraphy, Taipei, 2003.

Masatoshi Kubo, “Development of the Image Database at the National Museum of Ethnology: Toward Forum-Type Database Management”, International Conference of Museum Digitization Antiquities, Painting and Calligraphy, Taipei, 2003.

Elena V. Komarova, “Development of the Digital Collection of the State Hermitage Museum”, International Conference of Museum Digitization Antiquities, Painting and Calligraphy, Taipei, 2003.

Kranch, D.A, “Preserving electronic documents”, Proceedings of the third ACM conference on Digital Libraries, 1998.

Muir, A, “Legal deposit of digital publications: A review of research and development activity”, Proceedings of the first ACM/IEEE-CS joint conference on digital libraries, 2001.

Lawrence, Gregory W., William R. Kehoe, Oya Y. Rieger, William H. Walters, and Ann R. Kenney, “Risk Management of Digital Information: A File Format Investigation”, Council on Library and Information Resources, June 2000.

Rothenberg, Jeff, “Avoiding Technological Quicksand: Finding a Viable Technical Foundation for Digital Preservation”, Council on Library and Information Resources, 1999.

謝易耿、蔡順慈，〈由數位資訊生命週期探討數位典藏工作流程之建立〉，台北：數位典藏品質管理研討會，2007年。

蔡順慈、林昱伍，〈數位典藏工作流程與作業規範之探討〉，《中華傳播學刊》，(10)，2006年12月。

林玉雲、林國平、陳淑君、李士傑、李德財，〈結盟與合作：「數位典藏國家型科技計畫」的新能量〉，《博物館學季刊》20:3，2006年7月。

張懷文，〈詮釋資料與數位典藏長久保存取用－淺談Preservation Metadata〉，《數位典藏國家型科技計畫電子報》，第三卷第一期，2004年1月1日。

林昱伍，《開放式數位典藏資訊系統之建置研究-以華岡博物館爲例》，中國文化大學資訊傳播研究所碩士論文，2004年。

蘇漢良，《工作流程管理系統之設計與建構》，國立台灣大學土木工程研究所碩士論文，2001年。

嚴漢偉，〈中央研究院數位典藏簡介〉，檢索：2009年12月。http://www.ascc.sinica.edu.tw/nl/89/1625/02.txt。

黃如足、梅世杰，〈標準作業程序(SOP)於數位典藏建置之初探〉，檢索：2009年12月，http://datf.iis.sinica.edu.tw/Papers/2003datfpapers/a/A-2.pdf。

張志光，〈器物影像數位化之品質管控作法探討—以故宮器物數位典藏子計畫爲例〉，檢索：2009年12月，http://www.npm.gov.tw/da/ch-htm/pdf/tech/tech-2.pdf。

陳昭珍主持，〈國內外影音資料數位典藏現況與趨勢暨我國影音資訊平台建置相關規範研究〉，行政院新聞局委託，2004 年11月。

黃宛瑜、蔡淑韻，〈史語所「民族學調查照片資料庫」後設資料之擬定〉，《人類學知識的數位典藏與加值應用研討會》，2004年5月。

周欣鶯，〈數位館藏的維護與保存〉，《書苑》，51期，2002年。

陳昭珍，〈電子資源的長久保存〉，《佛教圖書館館訊》，第25/26期，2001年6月。

Online Resources

“Reference Model for an Open Archival Information System (OAIS)”. CCSDS650.0-B-1, Blue Book, January 2002，檢索：2010年1月，http://nssdc.gsfc.nasa.gov/nost/isoas/ref_model.html.

TELDAP WIKI ，檢索：2010年1月，http://wiki.teldap.tw/index.php

數位典藏學習網，檢索：2010年1月，http://ic.shu.edu.tw/DA/DADL/dam_02_foreign_05.htm.
數位典藏國家型科技計畫技術彙編，檢索：2010年1月，http://www2.ndap.org.tw/eBook/showContent.php?PK=157.
文建會國家文化資料，檢索：2010年1月，http://nrch.cca.gov.tw/ccahome/
數位典藏與數位學習國家型科技計畫 拓展臺灣數位典藏計畫網，檢索：2010年1月。http://content.ndap.org.tw/index/?p=994.
Dublin Core: http://dublincore.org/
CDWA: http://www.getty.edu/research/conducting_research/standards/cdwa/
EAD: http://www.loc.gov/ead/
VRA: http://www.vraweb.org/projects/vracore4/
Species 2000 Data Standard: http://www.sp2000.org/index.php
DwC (Darwin Core): http://wiki.tdwg.org/twiki/bin/view/DarwinCore/WebHome
CSDGM: http://metadata.teldap.tw/project/project-frame.html

Digitization Procedures Guideline: Integrated Operation Procedures

Advisory Unit: National Science Council, Executive Yuan, R.O.C.

Issuer: Fu-Shih Lin

Editor-in-Chief: Peng-Sheng Chiu

Executive Editors: Yen-Hung Lin, Ting-Li Lin, Fang-Chih Lin, Lang-Hsuan Kao

Authors: Ya-Ping Wang, Mei-Chih Chen

Reviewer: Associate Professor Shun-Tzu Tsai, Shih Hsin University Department of Information and Communications

Translator: International Collaboration and Promotion of Taiwan e-learning and Digital Archives Project

Publisher: Taiwan e-Learning and Digital Archives Program　Taiwan Digital Archives Expansion Project

Address: No.128, Sec.2, Academia Rd., Nankang District, Taipei City, 115
Institute of History & Philology, Academia Sinica

Tel: +886-2-2782-9555 extn.288

Fax: +886-2-2786-8834

Website: http://content.teldap.tw

Email: content@gate.sinica.edu.tw

Typesetting: Ya-Chen Tasi

Printing: Evergreen International Corporation

Publish Date: 1st Edition March 2010

ISBN: 978-986-02-7219-2

本書譯自拓展台灣數位典藏計畫出版之數位化工作流程指南：整合性工作流程