Vascular Plants Digitization Procedures Guideline

Lang-Hsuan Kao, Mei-Chen Liang

International Collaboration and Promotion of
Taiwan e-Learning and Digital Archives Program

Acknowledgements

This book would not have been possible without the assistance of project directors and staffs of Academia Sinica Biodiversity Research Center, Council of Agriculture Endemic Species Research Institute, Taiwan Forestry Research Institute, National Taiwan University (NTU), National Taiwan Museum and National Museum of Natural Science, who provided details on their actual operations and experiences. Special thanks to Professor Hsieh Chang-Fu of the Institute of Ecology and Evolutionary Biology, NTU for reviewing this Digitization Procedures Guideline, and also to fellow colleagues of this project for providing assistance and suggestions.

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 countless categories and massive amounts of content. The first phase of the five year project was successfully completed in 2006. The following year, 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 content, 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

Contents

Contents

One.
Introduction

I. Research Background

Taiwan possesses extremely rich plant species archives, with plant investigations traced back to as early as gathering records of Englishman Robert Fortune in 1854; for example, the Tetrapanax papyriferus and Lilium formosanum specimens he collected in Tamsui are the earliest formal new records of species discovered. Plant investigations have been conducted in Taiwan for over 150 years and went through several periods; they are the earliest of biological resource investigations to be conducted in Taiwan, and reached their peak during the Japanese Colonial Period (1895-1945). However, specimens gathered during early periods were mostly sent to Japan, Europe and America. It wasn't until Ryozo Kanehira took charge of the Department of Forestry, Central Research Institute, Taiwan Govenor-General Office, did the research center of Taiwan's plants gradually shift to Taipei.

After Taipei Imperial University was established in April 1928, plant investigations were handled by the Office of Plant Systematics and Ecology directed by Yushun Kudo, and the Herbarium of dried preserved specimens (today the Herbarium of the Institute of Ecology and Evolutionary Biology, National Taiwan University) was built in 1929, conducting massive gathering and research. After Taiwan Restoration, economic depression during early periods resulted in lack of resources for academia. Luckily, a great number of middle-generation scholars of plant taxonomy were educated during this period and were devoted to specimen cataloging and researches; the results of well planned plant investigations gradually appeared. Thanks to the efforts and contribution of predecessors, Taiwan's rich plant resource records were transformed one by one into precious literature on Taiwan's plants; the plant specimens that were produced and collected are precious resources of Taiwan's plant history.[1]

Precious plant specimens were accumulated as plant research continued; the immense collections of plant specimens we have were the fruits of painstaking efforts of our predecessors, and have been served as important resources for academic research. However, population growth in recent years and overexploitation of the natural environment have caused great damage to

[1] <Current Research and Status of Taiwan Seed Plants> by Peng Ching-I and Yang Yuen-Po et al., "Special Issue No.11 of the Institute of Plant and Microbial Biology," 1992, pages 55-85.

Taiwan's ecological environment, and also forced many species to the brink of extinction. From another aspect, plant specimens are used repeatedly for research, so fragile specimens will eventually be consumed. Thus, planning a suitable preservation environment is also an important link to the proper preservation of plant specimens. In order to effectively control the categories, characteristics, distribution and quantity of plant resources for the purposes of preservation and sustainable utilization, the most effective path to achieving proper preservation of specimens is to digitize specimen data. Digitization of important cultural assets and collections aims to achieve permanent preservation, establishing a detailed and complete database, and providing data via digital transmission for extensive utilization, thus achieving integration of national resources, facilitating academic exchanges and becoming in line with international standards.

The "Taiwan e-Learning and Digital Archives Program" (TELDAP) was established in 2002 to implement digitization related affairs. The primary objectives of this program are to: digitize important cultural relics and archives, establish national digital archives and databases, and drive cultural, social, industrial and economic development. At present, digitization procedures for various objects are almost complete, and results of the program have gradually appeared; data entries have increased significantly and some are already online for the public to access.

Although TELDAP has already been implemented for several years, considering the rich collections of domestic institutions, insufficient experience with digitization and different methods adopted by different units does not benefit the long-term implementation of digital archives. To address this issue, the first division project of TELDAP – Taiwan Digital Archives Expansion Project assembled six thematic groups, each with several teams for similar topics.[2] The purpose for thematic groups is to periodically convene meetings with institutions within the same theme regarding objects of the same characteristics; they are platforms for communicating digitization procedures, formulating related experiences and exchanging experiences. However, institutions have varying resources and equipment, resulting in significant

[2] Taiwan Digital Archives Expansion Project – Introduction to the six thematic groups, Search: November 2009, http://content.ndap.org.tw/index/?cat=6&team=9.

differences in their digitization procedures. Digitization procedures guidelines were written to integrate digitization procedures and combine experiences of digital archive projects, hoping that they will serve as indicators for digitization. Digitization procedures guidelines not only serve as operation manuals for new personnel of implementation institutions, but also provide reference for institutions planning to implement digital archive projects. On the other hand, these guidelines encourage more units to join in the ranks of digital archives and cultivate professional talents.

II. Current Status Analysis

Taiwan is in a unique geographical location encompassing the tropics and subtropics climate zones; the complex terrain and varying climate have resulted in the complex and special composition of plant communities found in Taiwan. Although Taiwan is small in size, its diverse ecological environment has nurtured rich biological resources, and its high species diversity is unmatched by most countries or regions; Taiwan may be proclaimed an important biological resource pool on earth. According to “Flora of Taiwan” (Huang et al., 2003), Taiwan is home to 4,339 vascular plant species (including 420 varieties and subspecies), and is considered a precious treasure vault on earth (including ferns, gymnosperms, dicotyledons and monocotyledons).

Over the years, six execution units have taken part in the flora team of the Biosphere and Natural Thematic Group, and implemented a total of 10 projects. There archive status of each unit is analyzed in the table below:

[I]	
Unit	Academia Sinica Biodiversity Research Center
Project Name	Database of Native Plants in Taiwan [2002 ~]
Project Website	http://taiwanflora.sinica.edu.tw/
Introduction	Academia Sinica Biodiversity Research Center has a rich collection of data on native plants in Taiwan, literature, precious specimens, and images of wild plants and their habitats, which are provided for academic research, ecological conservation, environmental impact assessment, and development of related industries. This project aims to digitize specimen gathering records of native plants in Taiwan, specimen images, ecology images and literature on plant taxonomy, and establish a database of Taiwan’s plant resources with both text

Introduction	and images, so that users from all sectors will be able to effectively utilize its resources. Also, in coordination with the development of WebGIS technology, this database integrates the longitude and latitude coordinates of where specimens are collected with user requirements to produce “distribution maps of plants in Taiwan,” and links it to contents of the specimen database. The “Database of Native Plants in Taiwan” integrates geographical distribution of plant species, specimen data, literature, and plant images.
[II]	
Unit	Council of Agriculture Endemic Species Research Institute
Project Name	Digital Archives of Taiwan Wild Plant Specimen and Ecological Images [2008, concluded] Digital Archives of Taiwan Wildlife Ecological Images and Plant Specimen [2009 ~]
Project Website	http://plant.tesri.gov.tw/plant/
Introduction	**Digital Archives of Taiwan Wild Plant Specimen and Ecological Images:** The Endemic Species Research Institute was founded in July 1992. Understanding that basic data of plant resources is a basis for academic research and execution of preservation affairs, the Institute has completed biological resource investigations of all counties, cities and offshore islands. Besides making lists of plants discovered and making specimens, the Herbarium of the institute currently has accumulated 23,395 specimens (excluding duplicate specimens), belonging to 224 families and 3,064 species (representing roughly 3/4 of all native plants in Taiwan), as well as tens of thousands of ecology images. For more convenient access by the public and researchers, this project digitizes plant specimens, ecology images and related information, and uploads the data to the Taiwan Wild Plant Database. Digitization of herbarium specimen benefits specimen preservation, reduces specimen consumption, reduces the time researchers spend going back and forth between herbaria, and benefits social education and ecological preservation. **Digital Archives of Taiwan Wildlife Ecological Images and Plant Specimen:** This project aims to digitize the herbarium’s dried preserved plant specimen, positives of vascular plants photographed in the wild, positives of wild animals, bryophytes and microscopic cells, and betacam filmed by the Endemic Species Research Institute during early periods. Related information is uploaded to TaiBif and the Union Catalog. In addition to the abovementioned databases, plant data is also uploaded to Taiwan Wild Plant Database (http://

	plant.tesri.gov.tw/plant/) for the convenience of the public and researchers, serving the purpose of ecological preservation and environmental education.
[III]	
Unit	Taiwan Forestry Research Institute
Project Name	Herbarium Digital Archives Project [2007, concluded] Taipei Botanical Garden Plant Digitization Project [2005 ~ 2007, concluded]
Project Website	http://tpbg.tfri.gov.tw/ http://taif.tfri.gov.tw/
Introduction	**Herbarium Digital Archives Project:** Since the Herbarium of Taiwan Forestry Research Institute was founded in 1904, it has collected over 270,000 specimens, covering over 90% of all vascular plants in Taiwan, including some 1,700 type specimens to which the scientific name of a plant species is formally attached, as well as specimens gathered in Southeast Asia, Japan and pacific islands. Herbarium specimens have long been provided to domestic and foreign scholars as basic data for academic research. Besides general specimens, the Herbarium has completed the digitization of type specimen images, creation of metadata files, and the scanning of the original literature published. Furthermore, the Herbarium completed the digitization of images contained in 543 ancient glass negatives in 2007, and created metadata files for the images. The herbarium's website has an information search webpage for more convenient access by other academic institutions and researchers via the internet, allowing digital resources to be more extensively utilized and reducing the risk of specimens being damaged when borrowed or transported. This herbarium hopes to cooperate with other herbaria in information integration and analysis in the future via data management technologies and high speed network. **Taipei Botanical Garden Plant Digitization Project:** Taipei Botanical Garden was established in 1896, and was originally a nursery garden built by the Japanese when they occupied Taiwan. The nursery garden was formally renamed Taipei Botanical Garden in 1912, and in addition to its original tasks, the garden sent personnel to Europe, America, Australia, Africa and Southeast Asia to gather tree seeds and transport them back for cultivation, making an extremely great contribution to academic research and natural science education. After Taiwan Restoration, Taiwan Forestry Research Institute was devoted to the operation of Taipei Botanical

	Garden, rearranging the garden and actively importing seeds to plant. At present, the garden has over 2,000 plant species and has become the best place for classification research and biology education. This project plans to complete digitization of data over a three year period and include it in the garden's website for academic research, professional and popular education. Contents and scope of the digital database will include images of the origin of plants and a plant database.
[IV]	
Unit	National Taiwan University
Project Name	The Digitization of Plant Specimens and Historical Botanical Literature of Taiwan[3] [2002 ~]
Project Website	http://tai2.ntu.edu.tw/
Introduction	TAI Herbarium has accumulated over 260 thousand specimens since its establishment in 1929, covering over 95% of vascular plants in Taiwan, including 1,000 type specimens and over 60 thousand old specimens with great research value that were gathered in Taiwan during the Japanese Colonial Period. For this reason TAI Herbarium has been the capital of research on plants in Taiwan and Southeast Asia. Thousands of books and journal articles have been published using these samples as research materials, represented by the Flora of Taiwan (1st Edition, Vol.1-6, 1975-1979) and Flora of Taiwan (2nd Edition, Vol. 1-6, 1993-2003). All type specimens, 65 thousand general specimens, origins of specimen, Flora of Taiwan 2nd Edition and related metadata were digitized in the first phase of TELDAP. This project adopts the original digitization framework when digitizing specimens that were not yet digitized, and expands digitization to important historical botanical literature. Besides preserving precious historical literature, this also presents Taiwan's rich biodiversity of the past that has now vanished. Such literature includes important results of plant resource exploration during late Qing Dynasty and the Japanese Colonial Period, most of which are rare publications now out of print. There is great need for systematic collection and digitization in coordination with important specimens of TAI Herbarium; databases established will be included in the website for permanent preservation, academic research, professional and popular education.

[3] Named "Herbarium Digital Archives Project" in the first phase of TELDAP (2002~2006) and "The Digitization of Plant Specimens and Historical Botanical Literature" in the second phase (2007).

[V]	
Unit	National Taiwan Museum
Project Name	Digital Archive Project of National Taiwan Museum – Botany [2004 ~ 2006, concluded]
Project Website	http://irs.ntm.gov.tw/plan_D/plan_Da.htm
Introduction	National Taiwan Museum is the earliest natural history museum established in Taiwan and has stood for nearly a century, serving research, preservation, exhibition and education functions. The botany division is mainly responsible for the investigation, collection, preservation and research of plant data and specimens, and coordinates with exhibitions and educational events. The herbarium of this division is registered in Index Herbarium as the Herbarium of Taiwan Museum (TAIM), and has vascular plant, algae, moss and lichen specimens and samples, including vascular plant specimens gathered before Taiwan Restoration, an abundance of algae specimens gathered in Taiwan's surrounding sea areas, and moss and lichen specimens gathered from five continents. In recent years, the herbarium has actively organized its specimens and related data with the plan to digitize them and create image data, so as to more effectively manage specimens and make access to the specimens more convenient.
[VI]	
Unit	National Museum of Natural Science
Project Name	Taiwan Low Altitude Plants Knowledgebase Subproject[4] [2002 ~] Digital Fungal Hall of Taiwan Subproject[5] [2002 ~] Taiwan Non-vascular Plants and Lichen Subproject [2002 ~ 2006], concluded
Project Website	http://digimuse.nmns.edu.tw/ http://digiku.nmns.edu.tw/fungi_web/
Introduction	**Taiwan Low Altitude Plants Knowledgebase Subproject:** The second phase project utilizes our advantages, using low altitude plants as a basis and further adds species that we do not have. The purpose of this project is to let the public understand the ecological environment and characteristics of low altitude plants in Taiwan, and further enhance their knowledge of ecological and

[4] Named "Taiwan Vascular Plants Digital Archives Subproject" in the first phase of TELDAP (2002~2006) and "Taiwan Low Altitude Plants Knowledgebase Subproject" in the second phase (2007).

[5] Named "Taiwan Fungi Subproject" in the first phase of TELDAP (2002~2006) and "Digital Fungal Hall of Taiwan Subproject" in the second phase (2007).

	environmental protection. Results can also be provided to foreign researchers for them to learn about low altitude plants in Taiwan, as well as education organizations for teaching purposes. This project gradually realizes life-long learning concepts, such as "virtual botanical garden" and "distance education." The database constructed by this project will benefit all future research, teaching, monitoring, development, utilization, preservation and management, so we know that the implementation of this digital archive project is important, valuable and of utmost urgency. **Digital Fungal Hall of Taiwan Subproject:** This project continues to expand a knowledgebase on the diversity of Taiwan's culture and nature, and establishes a digital museum with knowledge as its core. This project constructs a multilayer reusable knowledge framework for its natural and cultural knowledgebase, and provides it to academia for knowledge sharing and application. The knowledge system construction and reutilization model is also shared with other institutions to boost the integration and duplication of development models. The knowledgebase is converted into popular science education regeneration and application content, enriching the quality and quantity of digital content and presenting it in diverse ways, establishing a learning platform for middle and elementary school students and the public.

The data above indicate that there are many different types of plant collections under the flora team, e.g. fungi, non-vascular plants (lichen, moss and algae), and vascular plants (ferns, gymnosperms and angiosperms). In terms of digital archive projects, collections mainly consist of "vascular plant specimens," thus vascular plants are selected first as the target of digitization, and a complete guideline on digitization procedures for the flora team will be gradually established. This guideline integrates experiences accumulated by project implementation units in the flora team, and then applies theories, technical information and management systems gathered from different sources to establish an efficient and complete digitization workflow, which will serve as a basis and principles for controlling work progress and ensuring and enhancing digital quality, thus aiding different institutions in establishing digital archives. In addition, with consideration to budget limitations of different institutions, which might prevent them from purchasing certain digitization equipment, this books provides information on outsourcing, so that users may choose the most suitable digitization method based on project requirements and other considerations.

Two. Digitization Flowchart

The establishment of workflows aims to transform complex tasks into systematic procedures, or standard operating procedures (SOP), which allow digitization personnel to understand specifications and methods in the overall digitization process, thus achieving stable quality and cost down, while enhancing the overall quality of digital archives. To pass down practical experiences, digitization procedures adopted by the flora team over the years is generalized in Figure 2-1, please refer to related chapters (Chapter Three and Chapter Four) for details.

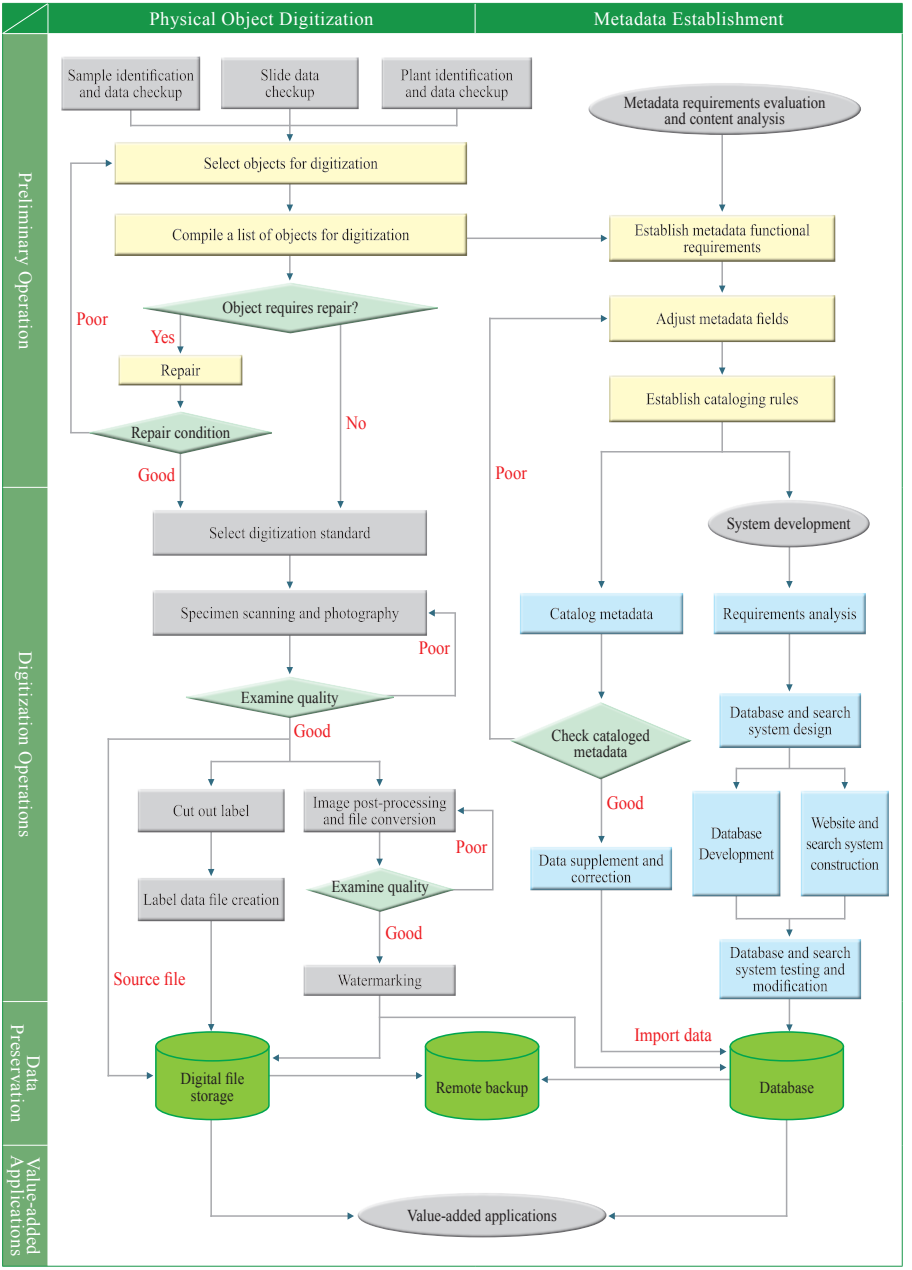


Figure 2-1 Digitization flowchart for plants

Three. Preliminary Operations

"Preliminary operations" play a critical role in the overall digitization workflow, and aim to let digitization procedures be successfully executed in an ideal state. It is recommended for senior researchers to be responsible for overall planning, which comes with the benefit of not only reexamining the entire collection, but also reestablishing management and storage principles if determined necessary. Properly arranging and checking data, from a long-term perspective, is meaningful to institutions concerning the permanent preservation of their collections.

Generally speaking, the digitization workflow can be divided into two parts – "preliminary operations" and "digitization operations." Preliminary operations mainly include thoroughly checking, arranging and examining specimens, and establishing file formats and digitization methods. Digitization operations mainly involve object image creation, metadata analysis, database establishment and webpage search system construction; preservation, sustainable operation and value-added applications of digitization results should also be included in plans. Tasks of each stage must be linked and correspond to one another, at the same time cost, time and human resources should be considered to gain maximum benefits. Work contents that require special attention during preliminary operations are described in the following sections.

I. Work Planning and Standard Establishment

Like the saying "good tools are prerequisite to the successful execution of a job," every item of every link in the digitization workflow must be carefully planned before executing digitization operations in order for digitization work to be smooth and successful.

(I) Establishing standardized forms

Keeping detailed records of specimens that were handled in each procedure is the same as rechecking all specimens in the herbarium, and benefits specimen management and checkup. For institutions with immense collections, this is also an opportunity to update collection management, and key in data of old and new specimens into the computer in response to current computerized management operations. No doubt that this will be more efficient and convenient for both administrators and users.

Furthermore, it is recommended to design tables and checklists for each

link of the workflow to keep accurate records of the actual situation. Using Table 3-1 as an example, each work procedure from preliminary operations to digitization operations is recorded, allowing researchers to more effectively gain control over the actual condition of specimens and manage the digitization process when arranging, checking or searching for specimen data, and make proper adjustments when necessary.

Plant specimens are fragile. Therefore, it is recommended for researchers to complete the entire digitization process once a plant specimen is taken out. This is to reduce the number of times specimens are used and avoid damaging them in the process of moving them. Thus, every aspect should be considered when establishing the digitization workflow, in which the main consideration of preliminary operations should be how to prevent damaging objects while achieving digitization, and thereby lay a solid foundation for future results.

Table 3-1 Checklist for plant specimens [Sample]

Specimen No.	Family No.	English family name	Chinese family name	Cleaned and repaired	Arranged	Checked	Scanned	Storage Location	Cut	Image Processing	Written on CD
059614	132	Leguminosae	豆科	v		v	v		v	August 26th	

Source: The Digitization of Plant Specimens and Historical Botanical Literature of Taiwan, National Taiwan University

(II) Digitization files specifications

File specifications may include file format, color mode, color depth, resolution and size; TELDAP has planned basic file specifications for project units to use as reference (Table 3-2). Basically, it is recommended to set different standards for different requirements, such as permanent preservation, commercial purpose (value-added applications) and public information (online browsing). To design suitable file specifications, institutions can refer the file specifications of members of the flora team

in TELDAP and consider their own requirements; file specifications of The Digitization of Plant Specimens and Historical Botanical Literature of Taiwan, National Taiwan University (Table 3-3) and Academia Sinica Biodiversity Research Center (Table 3-4) are provided below as reference. Finally, digital archives is a one time task, so institutions should look far into the future when establishing file specifications to prevent inadequate file specifications from obstructing value-added applications.

Table 3-2 TELDAP File Specifications

	Permanent Preservation	Commercial Purpose	Public Information (Online Browse)
File format	RAW or TIFF	TIFF	JPEG
Color mode	RGB (24bit/pixel) or above	RGB(24bit/pixel)	RGB(24bit/pixel)
Resolution and size	Original size, 300dpi or higher	Original size, 300dpi	Based on website design requirements, 72 dpi

Table 3-3 NTU File Specifications for Scanning Plant Specimens

	Permanent Preservation	e-Commerce	Public Information
File format	TIFF	BMP	JPEG
Color mode	RGB (24bit/pixel) or above	RGB (24bit/pixel)	RGB (24bit/pixel)
Resolution and size	Type specimen: Original size, 600dpi General specimen: Original size, 300dpi	Original size, 200dpi	100dpi

Table 3-4 Digital File Specifications of Academia Sinica Biodiversity Research Center for Ecological Slides

	Permanent Preservation	Online Browsing	Thumbnail
File format	TIFF	JPEG	GIF
Color mode	RGB (24bit/pixel)	RGB (24bit/pixel)	RGB (256 color)
Resolution and size	2000dpi, uncompressed File size: Roughly 28MB	1200dpi, compressed File size: Roughly 120KB	110dpi, compressed File size: Roughly 120KB

(III) Digitization Method Selection

Different digitization methods will be adopted according to the type, properties and form of objects to be digitized. In other words, the digitization method selected should be based on the object. Plant specimens are precious collections of each institution and possess high academic value. It is inevitable that some early specimens will tear or peel off when being preserved, and various natural disaster and artificial factors might damage the precious specimens. Today, these precious plant specimens can be digitized thanks to modern technology; however, the utmost priority of digitization is to protect the integrity of the object, then digitization procedures are considered.

Considering that the main purposes of plant specimens are for academic research, professional and popular education[6], and that scholars and experts emphasize the characteristics and detailed veins of specimens, therefore digital images must not be distorted, possess identifiable characteristics and present the original appearance of objects; the key of standards establishment is to make images as close to their original appearance as possible. The purpose of digitization is to achieve permanent preservation of collections; various factors need to be avoided in the process of converting physical objects into digital images, in order for the quality of digital images to meet standards.

Based on the above mentioned principles and after factoring in the implementation experiences of institutions in the flora team, this guideline recommends different digitization methods for type specimens and general specimens. Type specimens are precious in that they are irreplaceable, therefore scanning is recommended to acquire digital images with relatively higher resolution; photography using a DSLR camera is recommended for general specimen because of their relatively large quantity to shorten the work schedule of digitization. Although digital images obtained from 10 megapixel DSLR cameras currently in the market are not as good as those obtained from scanners, they are adequate for publication purposes.

In the digitization process, superior digital quality can only be

[6] Flora team/The Digitization of Plant Specimens and Historical Botanical Literature of Taiwan, National Taiwan University et al., "Digitization Procedures Guideline – Flora Team," Taipei: Content Development Division, 2004, page 1.

achieved via excellent equipment, skilled technique and strictly abiding by digitization principles.[7] Therefore, all aspects must first be considered and a suitable digitization method selected to yield twice the result with half the effort.

II. Specimen Cataloging and Repair

Plant specimens that are fragile and hard to preserve should be given priority for digitization. Plant specimens are a sample of a part or whole of a plant, and applied to species certification and classification research, providing researchers with a basis for comparison when identifying plant species, which make them indispensible evidence to international plant research. A valuable plant specimen must possess the following conditions:[8]

1. **Can be accurately identified:** Characteristics of each part (roots, stem, leaves, flower and seeds) must be clearly displayed for convenient observation by users.
2. **Correct and complete specimen label:** The label should list the specimen's category, location and gathering data, such as the names of the gatherers, location gathered, environment and date.
3. **Good preservation status:** Status of the plant original colors, roots, stems, leaves, flowers and seeds are all present and well preserved.

Furthermore, dried specimens are more fragile and easily peel off, especially old specimens, therefore cleaning or repair is a necessary procedure in the workflow. Repair procedures and notices should be set in advance to benefit the permanent preservation of specimens; also, make sure that images are complete and clear. Notices of specimen cleaning and repair are described more in detail below:

(I) Checking and verifying specimen data

Digitization of plant specimens is mainly in the order type specimens → old specimens → general specimens. Type specimens are by far the most precious of plant specimens, followed by old specimens with a long history,

[7] <An Analytical Study of the Evaluation Method of Digital Archive Image Quality> by Hsu Ming-Ching, Wei Yu-Chang et al., "Proceedings of the Digital Archives Conference on Workflows & Quality Management," December 2004, pages 237-238.

[8] <Standards of Plant Specimens>, Herbarium Ngensis, Search: February 2010, http://www.ngensis.com/flora/fn01.htm.

and general specimens, which are prioritized based on their intactness.

Thoroughly checking specimens is the most basic procedure of digitization, and includes specimen selection, categorization and arrangement, specimen verification, and verify and recording specimen label data. The selection of specimens for digitization should carried out by experienced researchers according to plant taxonomy, naming laws and regulations, and books, such as Flora of Taiwan; specimens should be cataloged while being categorized and organized.

Checking and verifying specimens is the most important but also most minute and complicated task, and also follows plant taxonomy, naming laws and regulations, and books, such as Flora of Taiwan. Any mistakes found on labels should be immediate corrected, and should be compared with the latest literature to ensure the validity of data.[9] According to project personnel of TAI Herbarium, for some specimens collected during the Japanese Colonial Period by Japanese or foreign scholars, the plant's scientific names on the labels would be according to the pronunciation of the scholar's native language, resulting in spelling differences; some labels were incomplete or sometimes blurred. These issues rely on careful and patient searches conducted by researchers to be resolved, so that the missing data may be filled in.

Specimen categorization and organization should be carried out while selecting specimens for the systematic planning of archive cabinets and long-term preservation environment. Furthermore, a checklist can be designed according to requirements of each institution for more effective checkups in the future.

(II) Specimen repair

Due to the long period of time plant specimens have been collected, the ratio of damaged specimens has grown as they are frequently moved and used for academic purposes. To protect the completeness of specimens and create image files in good condition, damaged specimens must be repaired once they are discovered during digitization, which will also extend the lifespan of the specimens. Therefore, specimens are all sent for repair after being checked and verified. Tasks in this procedure include the following:

[9] Same as note 8, Search: February 2010.

1. **Specimen cleaning:** When native plants are gathered and made into specimens, they go through a drying process, and will often have fibers peel off or become powder if preserved in a poor environment for long periods of time. This not only interferes with digitization work, but also makes image processing more time consuming, so it is necessary to clean specimens before proceeding with digitization operations. Use a soft brush to clean any specks, then examine the specimens, proceed with digitization if the specimen is in good condition; if not, repair the specimens as necessary.
2. **Specimen repair:** To maintain the original appearance of specimens, researchers should do their best to use branches and leaves of the original specimens that fell off for repair. First use a brush to clean off fiber powder of the specimens on the mounting sheet, and then check for branches and leaves that might have fallen off. Try your best to return the specimen to its original appearance using tape, glue, thread and needle, acid-free paper[10] and wrapper (Figure 3-1) for sowing or sticking (Figure 3-2), fixing the specimens in place and maintaining the completeness of its original appearance. If the mounting sheet is damaged, use a piece of acid-free paper of the same size and paste it on the back of the mounting paper, so as to strengthen the

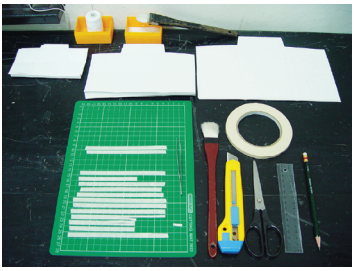


Figure 3-1 Tools required for repairing specimens: tape, needle, thread, tweezers and brush

Figure 3-2 Specimens with fat stem or roots can be fixed with needle and threads

Figure 3-3 Damaged mounting sheet and repair

[10] Acid free paper: Paper or mounting paper for pasting dried preserved specimens. Card paper, white board, coated paper or printing paper is generally used; the paper is cut into 42×29cm (international standard) or 39×27cm.

mounting paper and protect the specimen (Figure3-3).

3. **Notices of specimen repair** (Figure 3-4):
   (1) Tools preparation: It is inevitable for ancient specimens to peel off after being preserved for such a long time and be covered in powder. Repair personnel are required to wear a mask and gloves for hygiene purposes and to reduce damage to specimens caused by direct contact.
   (2) Arranging and fixing specimens: First use a brush to clean any peeled off pieces or powder from the mounting sheet to maintain a complete and clear image. The stems, branches and large hard leaves of a specimen can be fixed in place using threads, ensuring that the specimen is secure on the mounting sheet; tape the more fragile parts to prevent further damage or deformation.
   (3) Specimen preservation: Due to the preciousness of specimens, branches or pieces that peel off and can't be repaired should be wrapped with acid-free paper and pasted on a blank spot of the mounting sheet; in principle they should be pasted at the four corners, so that they don't interfere with the specimens and labels.

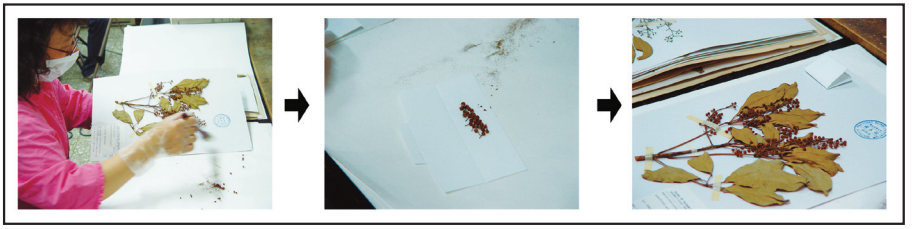


Figure 3-4 Specimen repair workflow

Digitization operations can be carried out once the preliminary operations above are completed.

Four. Object Digitization Procedures

After evaluating the implementation status of institutions in the flora team over the years, this guideline recommends "scanning" as the digitization method for plant specimens because it offers better digital image quality. Scanners convert optical images into electronic signals, and then store them in the computer. The light source for photoelectric imaging can be found within the scanner, signals are received by light sensing devices close to the object, and colors of the object are converted into digital color signals by the scanner's A/D Converter.[11] The purpose of digitization is to present clear and complete images as a replacement of the actual specimens for observation; the main advantage of scanning is that it allows characteristics of specimens to be clearly displayed, duplicating the actual appearance. Therefore, it is recommended to adopt scanning as the digitization method for unique type specimens or old specimens to acquire better digital image quality.

In terms of plant specimens, the main consideration is the clarity of detailed characteristics; in principle specimens should be scanned only once after the image specifications are established, avoiding the possibility of damaging specimens. The first thing to consider when selecting digitization equipment is that it mustn't damage specimens, and that its resolution must reach a certain standard.

In the digitization process of vascular plants, there are two types of images to consider – images of plant specimens and plant ecology images. Digitization operations associated with plant specimen images mainly use a scanner or digital camera; digitization operations associated with plant ecology images mainly use a slide image scanner or digital camera. Each digitization method has its purpose and considerations, which will be analyzed and explained below.

I. Plant Specimens

Scanning intact or repaired specimens is the most critical part of the entire digitization workflow. Scanning allows the specimen's appearance to be clearly displayed and show identifiable characteristics: roots, stems, leaves, flowers and seeds; even data on old labels should be clear and complete.

---

[11] "Digital Photography Technology" by Hsu Ming-Ching, Taipei: Garden City Publishers, October 2001, first edition, pages 80~81.

Some institutions use a DSLR camera for digitization instead of scanner. The main difference is in the resolution of images obtained. It is recommended to use a scanner when digitizing type specimens or other precious specimens to obtain images with higher resolution; for general specimens, which are in relatively large quantities, using a digital camera for digitization comes with the advantages of good image quality and higher digitization efficiency.

(I) Scanning

1. Scanning characteristics:

(1) Image capture: In scanning, light sensing devices receive the complete original image of an object under the scanner's light, clearly capturing every part of the specimen.

(2) Optical precision: Digital colors are generally analyzed by a RGB filter and then synthesized, so the precision that a scanner analyzes a single point affects the sharpness of the scanned image. Thus, the optical precision of a scanner greatly influences image quality. The optical resolution of platform scanners today have reached the level of 600dpi to 5000dpi[12], which has far surpassed the standard required for displaying the original appearance of an object.

2. Scanning principles:

Plant specimens are fragile and easily peel off, which is why the main consideration of scanner selection is to not damage specimens while meeting digital quality requirements. Plant specimens are mounted on A3 size sheets, so scanners should be able to scan A3 size objects. Make sure that specimens and labels are completely in the image during scanning.

3. Scanner selection:

Current scanners can roughly be divided into seven types: pen, hand-held, sheet-fed, platform, slide specific, drum and "flat-bed" scanners used for string bound books.[13] As for scanners that are suitable for scanning plant specimens, this guideline recommends a modified platform scanner that allows plant specimens to be scanned facing upwards or a flat-bed scanner, which reduce the possibility

[12] "Digitization Procedures Guideline: Color Management" by Li Pei-Ying et al., Taipei: Taiwan Digital Archives Expansion Project, April 2009, first edition, page 57.
[13] Same as Note 12, pages 56~57.

of damaging plant specimens when scanning them. Chapter eight offers a more detailed introduction to these two types of scanners and considerations when choosing one of them.

(II) Photography

1. Photography characteristics:

(1) Work efficiency: In the digitization of plant specimens, digital cameras are mainly used for general specimens. Due to the immense collections of herbaria, digitization is completed via photography, unless dealing with type specimens or precious old specimens. This increases digitization efficiency and leaves more time for organizing and verifying labels, thus ensuring the correctness and completeness of data.

(2) Image quality: It is very common for DSLR cameras currently in the market to have 10 megapixels or higher. In addition to the number of pixels, camera functions have been enhanced as well. Experiments have verified that high-end digital cameras are able to produce images of equal quality to traditional cameras.

2. Photography principles:

When photographing specimens, besides properly arranging light sources and the surrounding environment, clean the surface of specimens before photographing them and position the mounting sheet and specimen so that the specimen and label are clear and complete. Furthermore, some plant specimens contain fruits, so pay special attention to the focal distance to prevent images from becoming out of focus. Immediately examine the image once it is taken, re-photograph the specimen if it is not straight, the image has specks, is out of focus or the label is unclear.

3. Camera selection:

To summarize the digitization of plant specimens in different institutions, some use a DSLR camera for photographing specimens, while other use a 4×5 camera with a digital back, in which digital backs are further divided into scanning backs and one shot[14]. Although there are numerous types of cameras, but in principle scanning backs

[14] One Shot refers to the image captured when the camera opens and closes one time.

are still used for digitizing type specimens and old specimens because they offer higher resolution and image quality, but take relatively more time. For digitizing general specimens, institutions can choose DSLR cameras or one shot digital backs, which are more efficient and guarantee acceptable image quality.

The digitization methods above are all capable of recreating the original appearance of specimens. Plants go through a compression and drying process after they are gathered to be made in to specimens, some have lost their original color and texture due to long periods of preservation, and can only be identified via characteristics of their stems, branches and leaves, or even more detailed characteristics, such as leaf veins. This is why digitization methods that are able to recreate the original appearance of specimens are so important, they allow users to clearly see the true appearance of specimens without the need to borrow physical specimens.

II. Ecology Images

The difference between ecology images and specimen images is that ecology images are mainly for education and promotion purposes, while specimen images are for academic research purposes, but both emphasize the appearance and characteristics of plants. During early periods, ecology images were photographed using traditional cameras because digital camera technology was not mature; images were developed into slides to record plant ecology. Following the popularization of digital cameras and the maturing of its technology, digital cameras have completely replaced traditional cameras in photographing plant ecology images. Slides from early periods are now scanned into digital images using slide specific scanners for permanent preservation.

(I) Slide scanning:

1. Scanning principles:

Before scanning, check if the focal distance is clear and lighting is sufficient, and then select suitable slides for scanning. Then, check and verify image data to ensure its correctness.

2. Positive/negative (slide) scanner selection:

Equipment used by plant related project units for scanning slides mainly consists of “slide feeders” and “positive/negative scanners.”

The main purpose of "slide feeders" is to save time changing slides; using the NIKON SF-210 as an example, at most 50 slides can be placed at once for automatic scanning. If a "positive/negative scanner" (e.g. NIKON LS-5000ED) was used alone, then only one slide can be scanned each time.

At present, there are positive/negative scanners (e.g. NIKON LS-9000ED) that have built-in automatic slide feeders and allow scanning of multiple slides; these scanners have good image quality and functions. Considerations of selecting positive/negative scanners are explained more in detail in chapter eight.

(II) Photographing ecology images

1. Photography principles:

The photograph sequence should be based on the phonological phenomena of plants; contents generally include the appearance or habitual behavior of plants, characteristics of branches and leaves, blossoming, and close ups on flower pedals or fruits, which are important characteristics in plant taxonomy. Plants vary greatly in size, so the angle and position of the camera must be adjusted accordingly. Images should be photographed under natural light; adjust the white balance settings of the camera in place of color management.[15]

2. Camera selection

This guideline recommends that institutions use DSLR cameras for digitization. DSLR cameras currently in the market all have decent image quality; however, notice the difference between typical lenses and macro lenses when preparing for close up shots.

III. Digital Image Post-Processing

Object digitization emphasizes the importance of presenting the original appearance of objects. Digital images should show the complete specimen or plant, and should not have any specks on it (e.g. specks on the mounting paper of old specimens). It is recommended to use image processing software,

[15] <Taiwan Forestry Research Institute – Taipei Botanical Garden Plant Digitization Project Digitization Flowchart> by Taipei Botanical Garden Plant Digitization Project, 2008.

such as Adobe Photoshop, to remove any specks in the image or cut the image's borders. Furthermore, specimens usually don't contain rich, bright colors. Therefore, it is not recommended to conduct further color calibration. Basically, scanning is able to recreate the original color and appearance; plant ecology images are photographed under natural light, so adjusting the white balance of the camera should suffice. Tasks and focuses of image post-processing are as follows:

(I) Cutting the image's borders that are asymmetrical

When scanning plant specimens, make sure that the entire specimen is in the image. Notice the completeness of the image when cutting the borders that are imbalanced.

(II) Cutting the specimen label

The label of a plant specimen is generally at the lower part of the mounting sheet, and records the plant's scientific name, gatherer, gathering location and time; this information is provided for metadata establishment. After scanning a plant specimen, important label data can be cut out of the specimen's image and saved in a separate file.

(III) Examining specks and clarity of the image

After completing image digitization, first examine whether or not the image is clear, the specimen is straight, the focal distance is correct, detailed characteristics are displayed, lighting is sufficient and if there are specks or not. Any flaws in the image can be resolved by re-photographing the specimen, otherwise use image processing software to edit the image.

(IV) Image resolution specifications and preservation method

It is recommended to divide image files into three categories – permanent preservation, e-commerce, and public information, depending on their purpose, such as permanent preservation or general applications. Please refer to chapter three for more details on file specifications. Besides file specifications, the preservation method of image files is also a key matter that should not be overlooked. Based on their experiences, members of the flora team recommend three methods for preserving specimen images – saving them on computer, output paper copies, and writing them onto DVD.

Five. Metadata and Database Establishment

The main purpose of digitizing collections is to utilize modern technology to permanently preserve ancient objects with precious meaning, bestowing them with even more profound content and value by transforming them into digital images; databases established for these images provide search functions that make their access more convenient, thus allowing data to extend its reach and bring its potential cultural and academic value into full play. This is the mission of digital archives.[16] Therefore, all aspects must be thoroughly considered and planned before metadata and database establishment.

Taiwan possesses extremely abundant biodiversity resources and a large number of endemic species, thus the primary task at hand is to construct a complete biodiversity database. After integrating resources of digital archive projects, a search system for Taiwan plant resources with both text and images was constructed on the basis of Taiwan's biodiversity resources; the system can be accessed by all users and has become an exchange center for academic resources.

Database establishment mainly considers the interoperability of metadata fields with important plant databases around the world. The TELDAP Metadata Architecture and Application Team (MAAT) and Flora Team continuously adjusted and revised search systems and metadata of different databases, and established an interchangeable and standardized database system to integrate data while maintaining features of each institution's database system; this database system serves as foundation for establishing a single window in Taiwan.[17] To develop a data circulation and sharing mechanism for domestic herbaria that conforms to the structure of international herbarium databases, while meeting domestic requirements, considerations and design principles of metadata are described below.

---

[16] "Practices and Technology of Ancient Chinese Text Digitization" by Hung Shu-Fen, Taipei: TELDAP Training and Promotion Division, February 2004, page 48.

[17] Metadata Architecture and Application Team – Flora Team Work Plan, Search: February 2010, http://metadata.teldap.tw/project/project-frame.html.

I. Definition and Design Principles of Metadata

Metadata is generally defined as "Data about Data[18]." This definition originated in the "Metadata Workshop" co-organized by the Online Computer Library Center (OCLC) and National Center for Supercomputing Applications (NCSA) in March 1995. Scholars and experts in fields associated with library science, computer science and literature coding were invited to the Metadata Workshop, during which the definition "Data about Data" was proposed for metadata[19].

Metadata is a set of structural and standardized background information associated with objects that falls into three categories: descriptive, structural, and administrative, describing the contents and characteristics of each object in terms of semantics, syntax, and lexicology. Metadata allows digital collections to achieve optimal resource discovery performance in a digital environment or system, and effectively provides search, display, management, control and execution functions that facilitates digital resource interoperability and sharing, fulfilling its role as basic information for the permanent preservation of digital collections. In the light of this, digitization projects all consider metadata planning and implementation to be the most fundamental procedure in the digitization workflow.[20]

In addition to domestic considerations, the establishment of metadata should also adopt formats that conform to international standards; the more detailed index functions of the database are, the more information is provided to users, which also benefits data identification. Search functions and how data is displayed depends on metadata fields and the contents of each field. Therefore, when planning metadata fields, institutions should first understand future search functions of the database and how data should be displayed, this

[18] <Metadata>, Wikipedia, Search: February 2010, http://zh.wikipedia.org/zh-tw/Metadata.
[19] Same as Note 18, Search: February 2010.
[20] Metadata Architecture and Application Team – Project Introduction, Search: February: 2010, http://metadata.teldap.tw/introduction/introduction-frame.html.

way necessary metadata fields will not be left out.[21] Data fields of general specimens include scientific name, gathering location, gathering date, gatherer, appraiser, gathering number, origin of the specimens, specimen number and specimen condition; planning of the input format and file structure of specimen data must be completed in advance.

II. Metadata Recommendations for Plants

Metadata fields aim to establish interchangeable and standardized database systems that conform to international standards. The MAAT and flora team jointly established a metadata requirements table that uses the 15 elements of the Dublin Core[22] as a basis, and extends them according to data properties.

In terms of plant databases, "Species Elements" and "Specimen Elements" are important parts of the data structure. Strategy planning and application of international metadata standards, research and development of metadata theories, including metadata methodology, metadata registration center, and knowledge organization system; Table 5-1 "Common metadata elements of the flora team (draft): complied by the MAAT 2003/4/17" can serve as reference for future planning.

[21] "Practices and Technology of Ancient Chinese Text Digitization" by Hung Shu-Fen, Taipei: TELDAP Training and Promotion Division, February 2004, pages 58~59.

[22] The Dublin Core is a set of simple but effective core elements that originated in the first Metadata Workshop co-organized by the OCLC and NCSA, Search: February 2010, http://metadata.teldap.tw/standard/standard-frame.html.

Table 5-1 “Common metadata elements of the flora team (draft)” (v0.2)

類別		中文欄名		英文欄名		欄位定義描述	欄位值
物種名錄		界名	拉丁界名	Kingdom Name	Latin Kingdom Name	拉丁界名	
			中文界名		Chinese Kingdom Name	中文界名	
		門名	拉丁門名	Division Name	Latin Division Name	拉丁門名	
			中文門名		Chinese Division Name	中文門名	
		綱名	拉丁綱名	Class Name	Latin Class Name	拉丁綱名	
			中文綱名		Chinese Class Name	中文綱名	
		目名	拉丁目名	Order Name	Latin Order Name	拉丁目名	
			中文目名		Chinese Order Name	中文目名	
		科名	拉丁科名	Family Name	Latin Family Name	拉丁科名	
			中文科名		Chinese Family Name	中文科名	
		屬名	拉丁屬名	Genus Name	Latin Genus Name	拉丁屬名	
			中文屬名		Chinese Genus Name	中文屬名	
		種小名		Species Epithet		種小名	
		種命名者		Species Author		種小名命名者英文姓名	
		亞種名		Subspecies Epithet		亞種名	
		亞種命名者		Subspecies Author		亞種名命名者英文姓名	
		變種名		Variety Epithet		變種名	
		變種命名者		Variety Author		變種名命名者英文姓名	
		品種名		Form Epithet		品種名	
		品種命名者		Form Author		品種名命名者英文姓名	
		中文名		Chinese Name		台灣植物誌第二版正式採用之中文名	
標本紀錄欄位元素	標本典藏資料	標本館號／標本館館號		Specimen Order Number		標本進館編號／標本館館藏流水號	
		典藏單位	典藏單位	Herbarium/ Institute	Herbarium/Institute	典藏單位名稱	
			典藏單位代碼		Herbarium/Institute Code	典藏單位代碼	
		標本狀況		Specimen Status		標本保存狀況說明	良好／蟲蛀／毀損／遺失／其他
	標本採集資訊	採集者	採集者中文姓名	Collector	Collector Chinese Name	採來者中文姓名(先姓後名)	
			採集者英文姓名		Collector English Name	採集者英文姓名(先姓後名)	
		捎來編號		Collection Number	採集編號		
		採集日期		Collection Date	採集日期<yyy/mm/dd>		

*Country(Chinese) → Conutry (Chinese)

類別		中文欄名		英文欄名		欄位定義描述	欄位值
		模式類型／標本類型		Specimen Category		是否為模式標本模式標本	標本／全模式標本／複模式標本／副模式標本／配模式標本／選定模式標本／複選模式標本／等價選定模式標本／同模式標本／屬模式標本／系列模式標本／新模式標本／群模標本／產地模式標本／
	標本採集地理資訊	國別／地理區域	國別代碼	Country	Country Code	採集國家代碼	
			中文國名		Country (Chinese)	探集國家(中文)	
			英文國名		Country (English)	採集地國家(英文)	
		行政分區	一級行政分區中名	Administrative Area	1 st Administrative Area (Chinese)	一級行政分區(省、府、州)中名	
			一級行政分區英名		1 st Administrative Area (English)	一級行政分區(省、府、州)英名	
			二級行政分區中名		2 st Administrative Area (Chinese)	二級行政分區(縣、市)中名	
			二級行政分區英名		2 st Administrative Area (English)	二級行政分區(縣、市)英名	
			三級行政分區英名		3 st Administrative Area (Chinese)	三級行政分區(鄉、鎮)中名	
			三級行政分區中名		3 st Administrative Area (English)	三級行政分區(鄉、鎮)英名	
		國家公園／保護區	國家公園／保護區中名	National Park/ Reserve	National park/ Reserve (Chinese)	國家公園／保護區或特殊地理區中文名	
			國家公園／保護區英名		National Park/ Reserve (English)	國家公園／保護區或特殊地理區英文名	
		地名	中文地名	Locality	Locality (Chinese)	中文地名	
			英文地名		Locality (English)	英文地名	
		其他描述／環境描述		Additional Descriptions		地點其他描述／包括地形、植群、棲地、立地基質、濕度、光度、物候	
		海拔高度下限		Altitude/Lower Altitude		海拔高度或範圍下限值／單位：公尺	

其他各表同

類別	中文欄名		英文欄名		欄位定義描述	欄位值
標本鑑定資訊	海拔高度上限		Upper Altitude		海拔高度範圍上限值／單位：公尺，本欄不填則代表高度值固定非範圍	
	東／西經		Longitude Direction		E/W	東經(E)／西經(W)
	經度(度／分／秒)		Longitude (Deg/Min/Sec)		經度分為度分秒	
	南／北緯		Latitude/Direction		N/S	北緯(N)／南緯(S)
	緯度(度／分／秒)		Latitude (Deg/Min/Sec)		緯度分為度分秒	
	植物學名索引碼		Code of Plant Scientific Name		標本鑑定學名代碼	
	原始鑑定者	鑑定者中文姓名	Identifier	Identifier/Determined by(Chinese)	鑑定者中文姓名(先姓後名)	
		鑑定者英文姓名		Identifier/Determined by(English)	鑑定者英文姓名(先姓後名)	
	原始鑑定日期		Identification Date		標本鑑定日期＜yyy/mm/dd＞	
標本訂正資訊	館號		Specimen Order Number		標本進館編號	
	訂正紀錄編號		Verification Serial Number		標本訂正紀錄編號	
	訂正後植物學名索引碼		Verified Code of Plant Scientific Name		標本訂正後之學名代碼	
	訂正者	訂正者中文姓名	Verifier	Verifier(Chinese)	標本訂正者中文名(先姓後名)	
		訂正者英文姓名		Verifier(Engish)	標本訂正者英文名(先姓後名)	
	訂正日期		Verification Date		標本訂正日期	
文獻資料	文獻標題	篇名	Title	Title	文獻篇名	
		語文		Language	文獻語文	
	作者	中文姓名	Author	Name	作者中文姓名	
	書刊名	書／刊名	Reference Title	Reference Title	文獻出處英文書刊名	
影像資料欄位元素	媒體種類／影像來源		Origin of Media/Origin of Image		影像種類描述／影像來源的類型	標本／標本照／生態照／手繪圖／其他
	作者	作者(中文)	Author(Chinese)		指攝影或繪圖者	
		作者(英文)	Author(English)		指攝影或繪圖者	
	拍攝日期		Photograph Date		影像攝影日期	

Source: Digital Archives Technology Collection 2007[23]

[23] Digital Archives Technology Collection 2007, Search: February 2010, http://www2.ndap.org.tw/eBook08/showContent.php?PK=25#_附件二_「植物主題小組後設資料共通元素總表初稿（v0.2）」.

The establishment and structure of basic data are extremely important due to complexity of biodiversity data. The collection, management and preservation of basic data can be considered the foundation of biology research. Therefore, whether or not biological information is fully exchanged and easy to access has an immense effect on biodiversity research. The accessibility and extension of internet resources are without doubt the most suitable strategy for providing convenient applications of scientific data on biodiversity. The establishment of a life science database was listed as a key research direction by the National Science Council as early as 1994 when it was planning the discipline and resources of life sciences. With regards to results integration, the GBIF portal of TaiBIF (Taiwan Biodiversity Information Facility)[24] has obtained considerable results in the integration of information on Taiwan's biodiversity, including list of species, list of experts, species descriptions, endemic species, invasive alien species, distribution of species on land and at sea, literature on biodiversity, geographic information, environmental information, and related institutions, organizations, projects, scenic spots, biological databases and publications.

[24] Taiwan Biodiversity Information Facility, http://www.taibif.org.tw/.

Six.
Outsourcing Digitization

When deciding whether or not to outsource digitization (whether to purchase equipment and allocate manpower or let contractors provide both), consider which option has lower cost, shorter time, better quality and higher output. The most important medium of digitization operations is mechanical equipment, e.g. scanners, cameras and digital backs, which convert physical collections into image files that can be browsed online or permanently preserved. The key to good or poor image quality lies in the functionality of equipment; better equipment are more expensive and may even need to be imported from overseas. Luckily, the domestic digital content industry has flourished with great support from government agencies, and as supply and demand are becoming even, existing self-developed equipment or equipment purchased from overseas is sufficient for meeting digitization requirements.

Most domestic institutions are already engaged in digital archive projects; implementation methods include having their own personnel carry out digitization or outsourcing digitization. If digitization involves and requires professional knowledge and skills for identification and decision-making, for example, plant specimens are fragile and precious and their images need to stress certain characteristics, then outsourcing digitization is not recommended. However, if the equipment required for digitization is too expensive, the quantity of objects is too large and requires a long period of time to digitize, or under funding restraints objects are not suitable for being transported long distances, institutions may outsource digitization provided that the contractor carries out digitization in the institution with project personnel to provide assistance, thus reducing the possibility of collections being damaged.

With consideration to the preciousness of collections, the best option is of course for institutions to complete digitization operations themselves. Due to the importance of their collections, some institutions remain skeptical about outsourcing digitization; this might be because they do not understand outsourcing procedures, or had bad experience in the past. Therefore, this chapter further discusses outsourcing, and uses an interview article: “Industry-academia cooperation – Outsourcing Management: Experience sharing of contractors” (See Appendix 4) to provide institutions relatively insufficient financial and material resources an alternative option for digitization.

I. Comparison of Self-Production and Outsourcing

Institutions should first evaluate their equipment and human resources along with the purpose and requirements of their digital archive projects, before determining whether or not to outsource digitization operations. The table below compares advantages and disadvantages of self-production and outsourcing:

Table 6-1 Comparison of Self-Production and Outsourcing

	Advantages	Disadvantages
Outsourcing[25]	1. Saves expenses on personnel, equipment and R&D: Lowers cost 2. Dedicate efforts to core services 3. Reduces management pressure, focuses on allocation of existing human resources 4. More flexible internal organization and functional restructuring 5. Not limited by existing professional knowledge and skills	1. Takes time communicating and coordinating with the contractor 2. Internal operations are easily affected by outsourced operations 3. Concerns over collection management and maintenance
Self Production	1. Proper adjustments can be made in the process 2. More convenient for internal communication and coordination	1. Longer time spent on digitization operations 2. Insufficient expertise, high training cost 3. Additional expenses on equipment required

Source: Compiled by Taiwan Digital Archives Expansion Project

II. Work Items and Procedures of Outsourcing

The flowchart below shows the entire outsourcing process. This flowchart is based on the conclusion of the e-government CIO council meeting on January 9th, 2003 to "establish price standards for outsourcing information affairs and a knowledge management mechanism," which led to the proposal

[25] "Digitization Procedures Guideline: Outsourcing Management" by Kao Chih-Tung, Taipei: Taiwan Digital Archives Expansion Project, April 2009 first edition, pages 15-16.

and implementation of the "government agency information affairs outsourcing promotion project"; the project complied a manual on best practices of outsourcing government information affairs.

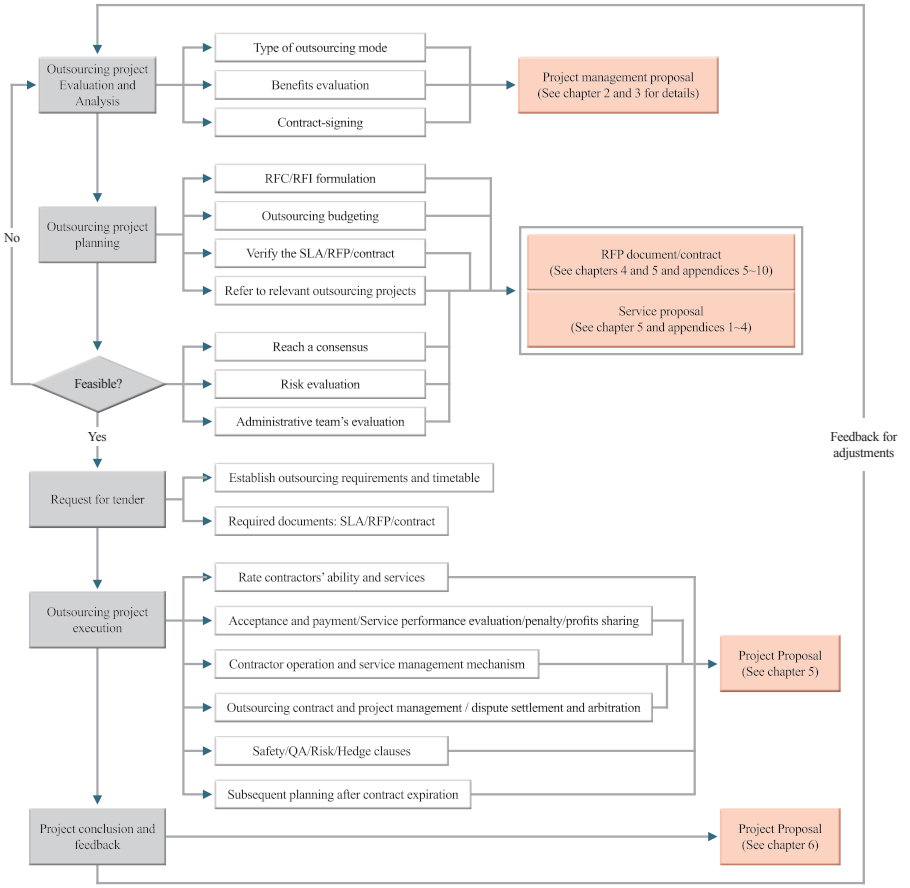


Figure 6-1 Flowchart of outsourcing projects[26]

Important stages of the outsourcing process shown above include:

(I) Outsourcing project evaluation and analysis: Establish specifications and related documents.

[26] Same as Note 25, page 17.

(II) Outsourcing project planning: Budget planning and contract writing

(III) Risk evaluation and verification

(IV) Tender invitation: Set the outsourcing time, goals and quantity of objects to be digitized.

(V) Outsourcing project execution: Establish inspection standards, management mechanisms and penalties if standards are not reached.

Every outsourcing project might have different results due to differences in personnel, time, technology, digitization object, and implementation method, making each project original.[27] Therefore, regardless of if it is an extension project, every link in the outsourcing process should be carefully arranged. Furthermore, communication and coordination with the contractor affects whether or not outsourcing will be successful; this process avoids unnecessary disputes during work and prevents any loss from resulting. For details on procedures and regulations concerning outsourcing, please refer to "Digital Archives Series: Digitization Procedures Guideline – Outsourcing Management."

[27] Same as Note 25, page 129.

Seven. Digital Content Protection

Developments in modern computer technology and the booming internet have not only accelerated the speed at which data is disseminated, but also extended its reach; the public now relies on the internet for work, learning, accessing data and browsing current events. Yet, convenient data dissemination has caused intellectual property rights to be infringed without notice, especially important and valuable electronic data. The prevalence of internet access struck a severe blow to intellectual property rights of creators, making it necessary for digital content protection to cover both creations and their creators, thus ensuring the safety and rights of both parties when using content on the internet.

TELDAP has been implemented for numerous years, and significant achievements have been made under the efforts of project personnel in historical relics, animal and plant specimens, audiovisual data, news and current events, historic architectures and even medicine. In addition to the management and preservation of physical collections, the protection of precious duplications and images on the internet from infringement must rely on the aid of digital content protection technologies.

The main method used for protecting digital content is to limit user access within the licensed scope, or to add copyright descriptions to digital objects, thereby declaring who may use the digital content and how the content may be used. There are currently numerous existing intellectual property rights protection technologies and standards, such as DRM: Digital Rights Management, Digital Watermark and Creative Commons,[28] of which Digital Watermarks are most often used when dealing with plant specimens. The following sections briefly introduce each of the abovementioned protection technologies and related standards.

I. DRM: Digital Rights Management

DRM technology is a protection and management mechanism for digital data that integrates technologies of digital watermarks, cryptology and right expression languages, and allows copyright owners to protect their intellectual property rights. DRM protects digital objects from being modified or damaged during circulation, and its main technology has the following three

[28] Creative Commons Taiwan, http://creativecommons.org.tw/blog/.

characteristics:[29]

Applicable to image (JPEG, TIFF…etc.), audio (MP3, WAV, WMV…etc.) and video (MPEG-1, MPEG-2, MPEG-4…etc.) protection.

The logo and information of digital content copyright can be embedded in images by embedding a watermark, thus protecting the copyright of digital content and ensuring the integrity of the original content (Please see II for an introduction to watermarks).

Encrypting digital content protects it from being illegally accessed by a third party; unnecessary digital asset loss can be avoided by limiting the number of times digital content can be downloaded and used, and by setting access rights.

II. Digital Watermark

Watermark technology is mainly used to embed "copyright information" into digital content, and hides marks (encryption, text or logo) that allow copyright to be identified in the protected data, so that this digital data will retain the marks when downloaded and used. If there is a dispute over copyright in the future, then the mark extracted from the image will show who the legal owner is. Digital watermarks are usually embedded in image, audio and video.[30]

Basically, watermarks consist of two technologies – visible watermark and invisible watermark, each of which is briefly described below.

(I) Visible Watermark

This technology adds an obvious copyright logo to digital data, directly notifying users who the copyright owner is and warns them to respect the copyright when using the data. This advantage, however, is also a disadvantage that many people of complained about because it directly damages image quality and clarity; watermarks that are to obvious often make the file unreadable and reduces the image's original value. To learn more about watermarking technology, visit the "TELDAP Portal," and then click on blog –

[29] <Digital Rights Management Technology>, Industrial Technology Research Institute – Transferrable Communication and Photonic Technologies, Search: December 2009. http://www.itri.org.tw/.

[30] <What are Watermarks?> by Chia Hsin-Chieh, TELDAP e-Newsletter, Search December 2009, http://www2.ndap.org.tw/newsletter06/news/read_news.php?nid=732.

experience watermark technology.[31]

(II) Invisible Watermark

Invisible watermarks cannot be seen from the appearance and acts like hidden signatures or imprints that artists like to put in their artworks, thus protecting the appearance of digital objects. What is different from visible watermarks is that many institutions and research units have higher interest in invisible watermarks, and are dedicating their efforts to its development. When an invisible watermark is embedded in data, it is undetectable by the naked eye and must be extracted using special algorithms by a trustworthy third party for verification, after which it will be able to prove who has copyright ownership, protecting the rights of digital content owners.

After considering actual system operations, requirements that need to be satisfied when using invisible watermark technology are as follows:[32]

1. The medium used after the watermark is added should be as close to the medium used before as possible. In other words, high quality media should be used to contain digital data after watermarks are added, this way it does not lose the value of the original medium.
2. Aside from the legal owner, others should not be able to detect the existence of a digital watermark in the image to ensure safety of the digital watermark.
3. The algorithm used for hiding the watermark must be made public, just like encryption algorithms in cryptology; system security must not be based on the premise that the perpetrator does not know how the system operates.
4. Extraction of the watermark must not require the original medium. This way it won't be necessary to store two copies for extracting watermarks, saving storage space.
5. The watermark extracted must possess robustness, meaning that the watermark should still appear after going through a number of digital signal processing technologies, such as filtering, loosely data compression and crop-and-parse.
6. Digital media usually have immense storage capacity, so the

[31] TELDAP Portal, Search December 2009, http://digitalarchives.tw/.

[32] "Introduction to Digital Watermarks" by Chang Chen-Cheng, National Chi Nan University e-Magazine, Search: December 2009, http://beaver.dlc.ncnu.edu.tw/projects/emag/.

watermark algorithm should have good execution efficiency in order to be practical.

7. Any medium should allow multiple watermarks to co-exist; this way it can protect the rights of multiple owners and reduce the risk of being cracked.

Digital watermarks are the most frequently utilized digital content protection technology for plant related digital archives, and are also the most convenient. However, they still only count as one type of digital content security mechanism and cannot guarantee that digital content is not stolen; more protection technologies are still required to provide digital content with better protection.

III. Creative Commons

Creative Commons is a form of open copyright licensing that is applicable to an extensive range of works, including music, stories, paintings, novels, pictures, website content, movies, sound effects, books and articles. Unlike the conventional copyright licensing contract, Creative Commons allows creators to license their work in advance to anyone who might want to use them. Therefore, any user may directly use the work as long as he/she abides by the licensing terms declared by the creator. No additional license is required from the creator, thus eliminating the possibility of copyright infringement due to inability to contact the creator.[33]

At present, "Creative Commons licenses 3.0 Taiwan" has been formally launched online. The introduction of version 3.0 does not necessarily cause versions 2.5 or 2.0 to lose effect, instead creators can choose the version and type of license that best suits their needs.[34] There are four licensing conditions for all versions of Creative Commons; for example, if the creator were to label his/her work with the license 

, which represents "Attribution-NonCommercial-ShareAlike," it would indicate that the work is protected by the terms of the CC license. When using a work with the abovementioned

[33] <Basic Concept of Creative Commons – What is Creative Commons>, Ministry of Education's website on Creative Commons, Search: December 2009, http://ccnet.moe.edu.tw/faq_cclist01_q1.htm.

[34] <Creative Commons licenses 3.0 Taiwan formally launched online on October 31st, 2009>, Creative Commons Taiwan, Search: December 2009, http://creativecommons.org.tw/blog/.

license, the user would be required to specify the name of the original creator, could freely use it for non-commercial purposes, and would be required to share any derivative works alike with the public. Furthermore, if the license  , which represents “Attribution-NoDerivs,” appears on a shared work, it means that users may not create any derivative works or will be liable for infringement.

Eight. Equipment and Cost Analysis

When digitizing precious data, aside from technical issues, equipment and cost also need to be considered so that proper arrangements can be made. After which institutions should evaluate their capabilities and requirements to select suitable digitization methods. The analysis below is based on equipment and costs described in chapter four, and can be considered when planning digitization operations.

I. Main Equipment and Tools

(I) Scanning equipment

1. "Platform" scanner

Platform scanners mainly use wide-angle lenses and reflex lenses to rapidly capture images to their CCD, and range from 600dpi to 5000dpi. Some scanners have a built-in mask or can add one to scan transparencies, and high-end platform scanners are now able to move on both X and Y axes, ensuring that the lens is always perpendicular to the area being scanned and avoiding image deformation due to the lens; the quality of images produced using such scanners are nearly equal to that of drum scanners, and are able to scan solid objects.[35] When using regular platform scanners to scan plant specimens, the plant specimens need to be turned down facing the glass panel to be scanned, but this damages vulnerable specimens and endangers their integrity. This action not only hinders digital operations, but also causes the glass panel to be scratched, interfering with image clarity and increasing amortization of equipment. In response to characteristics of plant specimens, TAI Herbarium modified the EPSON 10000XL/Pro platform scanner (as shown in Appendix 1) so that specimens can be scanned facing upwards; the modified scanner's functions and results can match that of a flat-bed scanner. The only disadvantage, however, is that the scanner's service life will be reduced, but is still a viable option considering equipment cost.

2. "Flat-bed" scanner

Flat-bed scanners are designed so that data being scanned faces

[35] "Digitization Procedures Guideline: Color Management" by Li Pei-Ying et al., Taipei: Taiwan Digital Archives Expansion Project, April 2009, first edition, page 57.

upwards; light is projected from high above and the image is captured from above the machine. This makes it suitable for scanning plant specimens that shouldn't be turned facing down. When using this type of scanner, all personnel need to do is carefully place specimens on a fixed position, making it suitable for scanning fragile plant specimens or books with brittle paper; this type of scanner can also scan small solid objects, such as puppets and jade.

3. Positive/negative (slide) scanner

Positive/negative scanners are like a small box with a container at its side large enough to place 35mm slides. Furthermore, there are scanners that support multiple formats or 4×5 inch slides. The purpose for designing slide scanners is to scan small areas in high resolution; the resolution of images produced can rival that of middle to high end platform scanners.[36] The most commonly seen equipment in the plant related projects is the NIKON CooLScan series, such as the NIKON SF-210 (slide feeder) and NIKON LS-5000ED (positive/negative scanner) mentioned in chapter four. However, the above mentioned equipment is now out of production, so pay special attention to the after-sales service and rights of products before purchasing them.

In terms of other options, many platform scanners now have a built-in transparency holder, and even if they don't have a transparency adapter, such accessories are easy to purchase. If every image in the strip has the same color and resolution requirements, then multiple transparencies can be scanned at the same time, thus increasing scanning efficiency.[37]

(II) Photography equipment

1. Digital camera

As mentioned in chapter four, most digital cameras currently in the market, whether it is a consumer camera or professional DSLR camera, have over 10 megapixels, and are becoming cheaper and cheaper. Considering future applications of digital images, it is recommended to use a DSLR camera. Besides the advantage of being able to change lenses, more detailed aperture, camera shutter

[36] "Introduction to Imaging" by Howard Besser, translated by Lin Yen-Hung, Taipei: Taiwan Digital Archives Expansion Project, October 2009, first edition, page 50.
[37] Same as Note 36, page 50.

and depth of field adjustments can be made to produce better image quality.[38] In the photography of vascular plants, DSLR cameras are mainly used to photograph plant ecology images, but the tens of thousands of plants in nature come in greatly varying shapes and sizes, so lenses must be fully prepared.

In general, wide-angle lenses can be chosen for photographing whole plants or large plants, so as to satisfy requirements on the composition of the image. To photograph specific parts of a plant or a small cluster of flowers, choose middle to long focal length lenses to highlight the plant in a narrower view. To photograph close-ups on plant details, such as the pistil or veins, choose a macro lens with 1:1 magnifying rate to acquire images with clearer and more delicate lines. Still, equipment used for digitization should be based on the project's requirements and purpose to achieve the best results from fund utilization.

2. Digital back

Digital backs are used together with 120 medium or 4×5 large traditional cameras, and contain light sensing devices and control/storage units. Take off the back of the traditional camera and replace it with the digital back and the traditional camera will turn into a digital camera. Digital backs currently in the market all have at least 20 to 30 megapixels, which is more than adequate for photographing vascular plants. Besides the number of pixels, lens quality, type and size of light sensing devices, and software used to process images are all factors that affect image quality.[39] Although digital backs offer better image quality, they are relatively expensive, so institutions should carefully consider whether outsourcing or purchasing equipment is more economic.

(III) Computer equipment and software

It is recommended to have at least 2 computers, one for scanning or digital photography and the other for data and image processing; large monitors with high resolution and high color quality are recommended.

---

[38] "Ceramics Digitization Procedures Guideline" by Chen Hsiu-Hua, Taipei: Digital Archives Expansion Project, April 2009, first edition, pages 49-50.

[39] "Mastering Digital Printing " by Harald Johnson, Taipei: Flag Publishing Co., 2005, pages 3-22.

Furthermore, basic word processing and image processing software and a DVD recorder are also required; a UPS can prevent data loss and equipment from being damaged by unexpected blackouts.

II. Cost Analysis

(I) Cost elements

Costs associated with digitization include three basic elements: Material cost, Labor cost and Miscellaneous costs:

1. Material cost is the cost of consumables used for digitization work.
2. Labor cost is mainly the salaries of personnel.
3. Miscellaneous costs can be divided into direct costs and indirect costs:
   (1) Direct costs include cost and amortization of information equipment and scanners, and cost of information software.
   (2) Indirect costs include depreciation or rent of the scanning work space, renovation cost, insurance fee, utilities and other.

Costs that are not mentioned, such as time cost of repair or appraisal, are hard to control. For this reason the cost analysis of this guideline only considers labor cost and direct costs when estimating the scanning cost per image; the calculation process shows the relationship between cost elements (human resources, equipment and time) of digitization work.

(II) Cost estimation

1. Calculation method:

There are two ways to calculate cost based on amortization of equipment:

(1) Calculating amortization of equipment based on service life [Labor cost (NTD) + Amortization of equipment (NTD)] / Digital output (Number of images) = Cost per image (NTD/Image)
   A. Labor cost is mainly salaries of personnel
   B. [Equipment cost (NTD) + Software cost (NTD)] / Service life = Amortization of equipment

(2) Calculating amortization of equipment based on digital output [Labor cost (NTD) / Digital output (Number of images)] + {[Equipment cost (NTD) + Software cost (NTD)] / Digital output (Number of images)} = Cost per image (NTD/Image)

2. Calculation example

(1) Basic settings:

An institution is planning to implement a digital archive project and has allocated two personnel, one to be responsible for scanning and one for organizing data and verification. The equipment used includes two computers and one platform scanner.

A. Equipment: Computers: NT$30,000×2 = NT$60,000 Platform scanner: NT$100,000

B. Software: Adobe Photoshop: Roughly NT$20,000

C. Human resources: Monthly salaries of NT$30,000

D. File size: Scanning one A3 size image takes roughly 90 seconds

E. Monthly (20 work days) digital output: Roughly 2,400 images

F. Service life: 4 years (according to the accounting office of each institution)

(2) Calculating amortization of equipment based on service life

A. Labor cost = NT$60,000

B. Amortization of equipment = (60,000 + 100,000 + 20,000) / 4 = NT$45,000 per year or 45,000 / 12 = NT$3,750 per month

Cost per image = (60,000 + 3,750) / 2,400 = NT$26.56

3. Recommendations

The cost estimation of this guideline is only a preliminary evaluation. Institutions implementing digitization projects may use the simple formulas above to roughly estimate their cost, which is mainly based on human resources and equipment:

(1) Human resources: Since scanning operations are fixed, a professional part-time worker can be trained to handle the operations, thus reducing labor cost. Verification requires personnel with image processing expertise, therefore it is recommend to hire full time personnel to ensure image quality.

(2) Equipment cost: Although high-end equipment is expensive, if it benefits digital output, it will reduce labor cost, meaning that the overall cost will not necessarily increase. On the contrary, using cheap equipment with lower digital output will result in higher labor cost, and the overall cost will not necessarily decrease. Therefore, conduct an overall evaluation before purchasing

equipment to decide on the level of equipment to use.

(3) Outsourcing evaluation: If the total cost above is close to the cost of outsourcing scanning operations, then having a contractor work within the institution becomes an option. This way human resources and time can be dedicated to other procedures; working with a contractor will also allow the successful completion of digitization operations.

Nine. Benefits and Prospects

Ever since "TELDAP" was implemented in 2002, institutions have adopted different digitization procedures; the implementation and quality of digitization is affected by the funding, human resources, material resources and other conditions of institutions. In order for new members of TELDAP to rapidly become familiar with operations and begin planning related tasks, we compiled digitization methods of different institutions into this "Vascular Plants Digitization Procedures Guideline." This guideline combines work experiences of different institutions, and provides specific feasible digitization procedures with a theoretical foundation. However, this guideline is still in an orientation stage and has many flaws. Our efforts to better this guideline will be directed in the following directions:

(I) Gradual improvement

This guideline is the first attempt to gather together results and experiences of different projects, so it is hard to avoid parts requiring revision. Institutions implement digitization projects under different circumstances, therefore setting a single standard is inappropriate; institutions still need to adjust their implementation plans according to their own requirements. Furthermore, we hope institutions will provide more feed back regarding digitization methods, procedures, contents or equipment and software, so that we may enrich the contents of this guideline and improve standard procedures, striding towards higher quality and higher standards.

(II) Increase digitization efficiency and digitization quality

For institutions planning to implement digital archive projects in the future, we hope that they will be able rapidly gain a relatively complete concept of digitization work after referring to digitization procedures guidelines for their collections. This way they will not need to spend excessive time exploring other digitization methods and technologies, thus improving digitization efficiency and quality. This will allow more precious collections to be accessible, and achieve resource sharing via the internet.

(III) Keep pace with technological developments and new software and equipment

In an era with new technologies appearing every single day and new equipment and software frequently being introduced, software and hardware recommended by this guideline will one day become obsolete, and cost, human resource and efficiency estimations will change as well. Therefore, institutions may flexibly select equipment according to their actual situation.

Ten.
Conclusion

“Digitization procedures guidelines” are complete sets of operating procedures established to increase work efficiency following the substantial increase in quantity and quality of digital content. Besides preliminary operations, which are established for specific objects, digitization procedures are applicable to objects with similar characteristics. To the future of digital archiving, this digitization procedures guideline is a starting point.

Although there are the above constraints on digitization procedures guidelines, from a long-term perspective, digital archive projects carry the mission to preserve our culture and precious resources, promote popular science education and elevate our international academic status. Digital archiving is a current international trend.

In order to gather and collect even more precious resources, TELDAP has published a series of technical manuals, digitization procedures guidelines and other books. These publications not only describe the spirit, meaning and importance of digital archives, but also present actual contents of digitization workflows, so that the public may better understand the contents and importance of digitization. “Digitization work” is a field that requires both academic and technical expertise to implement. However, we hope that by compiling this guideline more institutions will join in the ranks of digital archives, and thus integrate nationwide resources and use equipment to further build an exchange platform. We also hope that under a mutual cooperation mechanism, we will be able to more effectively display and apply cultural assets in the future.

References

Books

Harald Johnson著，《數位列印輸出聖經》，台北，旗標出版股份有限公司，2005年。

Howard Besser著，《影像製作入門》，林彥宏譯，台北：數位典藏拓展台灣數位典藏計畫，2009年10月初版。

李佩瑛等著，《數位化工作流程指南：色彩管理》，台北：數位典藏拓展台灣數位典藏計畫，2009年4月初版。

洪淑芬著，《文獻典藏數位化的實務與技術》，台北：數位典藏國家型科技計畫 訓練推廣分項計畫，2004年2月。

徐明景著，《數位攝影的技術》，臺北：田園城市，2001年10月，初版。

高芷彤等著，《數位化工作流程指南：委外製作》，台北：數位典藏拓展台灣數位典藏計畫，2009年4月初版。

陳秀華等著，《瓷器數位化工作流程指南》，台北：數位典藏拓展台灣數位典藏計畫，2009年4月初版。

植物主題小組/國立台灣大學植物標本館典藏數位化計畫等著，《數位化工作流程—植物主題小組》，台北：數位典藏內容發展分項計畫，2004年。

Journal Papers

徐明景、魏裕昌等著，〈數位典藏影像品質評量方法之分析研究〉，《數位典藏作業規劃與品質管理研討會論文集》，2004年12月，頁235-247。

彭鏡毅，楊遠波等著，〈台灣種子植物之研究與現況〉，《中央研究院植物研究所專刊第11號》，1992年，頁55-85。

Online Resources

台灣生物多樣性 資訊入口網，http://www.taibif.org.tw/。

〈植物標本的規範〉，延陵科學綜合室，檢索：2010年2月，http://www.ngensis.com/flora/fn01.htm。

張真誠編，《淺談浮水印》，暨大電子雜誌，檢索：2009年12月，http://beaver.dlc.ncnu.edu.tw/projects/emag/。

創用CC—Creative Commons Taiwan，http://creativecommons.org.tw/blog/。

〈創用CC基本概念—何謂創用CC〉，教育部創用CC資訊網，檢索：

2009年12月，http://ccnet.moe.edu.tw/faq_cclist01_q1.htm。

磁軒資訊媒體行銷有限公司簡介，檢索：2010年2月，http://www.cx-media.com/web_main_6.html。

賈馨潔編，〈什麼是浮水印？〉，數位典藏國家型科技計畫電子報，檢索：2009年12月，http://www2.ndap.org.tw/newsletter06/news/read_news.php?nid=732。

〈Metadata〉，維基百科，檢索：2010年2月，http://zh.wikipedia.org/zh-tw/Metadata。

數位典藏技術彙編 2007年版，檢索：2010年2月，http://www2.ndap.org.tw/eBook08/index.html。

數位典藏與數位學習 成果入口網，檢索：2009年12月，http://digitalarchives.tw/。

〈數位版權管理技術〉，工業技術研究院-通訊與光電可轉移技術，檢索：2009年12月，http://www.itri.org.tw/。

Others

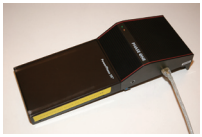
台北植物園植株數位化計畫著，〈林試所—台北植物園植株數位化計畫數位化工作流程圖文說明〉，2008年。

胡嘉穎，中研院生物多樣性中心—台灣本土植物數位典藏計畫數位化情形調查表，2008年10月。

楊綉玉，台灣大學—植物標本與歷史文獻典藏數位化計畫數位化情形調查表，2008年12月。

Appendices

Appendix 1: List of Scanning Equipment

Type	Appearance	Product Features	Source
Large professional multipurpose scanner (Top mounted)		Positive, photograph, newspaper, rare book, thesis, instrument panel, plant specimen, architecture blueprint, satellite image, shadow puppet, glass painting, Chinese painting, Western painting and bottles (Large/small width full color scanner)	http://www.iiri.com
Highly efficient top mounted scanner		For the purpose of scanning objects of various sizes (Luminescence scan does not generate heat)	http://www.zeutschel.com/
Flat-bed High Speed Microfilm System		For scanning high resolution large colored image files (books, newspaper, large documents, maps, engineering drawing and poster)	http://www.zeutschel.com/
Scanning back		High quality, high resolution professional image capturing device	The Project of Historical and Cultural Heritages Developed in IHP: Archeological Data
4×5 Camera		Can be used with scanning backs and large copy stands for scanning objects of various sizes	The Project of Historical and Cultural Heritages Developed in IHP: Archeological Data
Desktop Platform Scanner (Modified)		Resolution up to 2400dpi, and can be modified for scanning plant specimens	http://w3.epson.com.tw/

Appendix 2: Practical Planning of MAAT

Source: Taiwan e-Learning and Digital Archives Program – Metadata Architecture and Application Team (MAAT)

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

Stage 1: Requirements Evaluation and Contents Analysis

Procedure 1: Requirements Interview

The metadata analyzer interviews experts or content providers of a thematic project to understand project properties and metadata requirements. Besides acquiring project background information and coordinating contact methods, the metadata analyzer verifies the project's purpose, objectives and expected benefits. Information gathered at this stage mainly includes:

1. Contact information: Including project members, contact window and contact method.
2. Scheduled progress and timetable of metadata establishment.
3. Scope of metadata: The purpose and properties of metadata, such as who, what, when, why, where, or vocabulary control.
4. Basic information of the current system: Such as text fields (or elements), structure, file quantity, storage format, input method and system, so as to effectively understand advantages and disadvantages of the current system.
5. Metadata application background: Whether the database structure consists of a single metadata system or includes other supporting metadata systems, such as the Geographic Information System (GIS).
6. Role and function of metadata: Functions to describe, search, index or manage resources.

Procedure 2: Project Standards and Case Study

Analyze related metadata standards and cases to evaluate the feasibility of implementing a metadata standard. In terms of standards observation, the "metadata standard selection model" is applied and a suitable metadata type is found after analyzing four aspects of a project – community, data type, discipline and function. For case study, gather information on digital museum projects to understand trends and issues of metadata applications, and use the information as reference for future implementation and development. Results

of this analysis stage will help project members understand differences between their project and other projects with the same or similar characteristics, and thus revise the direction of their project.

Community	Library, Archive, Museum, Herbarium
Data type	Rubbings, Ancient Books, Photos, Audio, Specimens, Documents, Field Data, Textiles, Cultural Relics
Discipline	Literature, Anthropology, Physical Therapy, History, Politics, Zoology, Botany, Art, Buddhism, Minerals, Fossils, Industrial Technology, Linguistics
Function	Refresh, Search, Index, Manage, Description, Exchange, Access, Preservation, Internationalization, Form a complete digital archive system

Metadata Standard Selection Model

Procedure 3: In-depth Analysis of Metadata Requirements

Utilize the work sheets below to more accurately analysis metadata requirements of your project:

1. Metadata requirements form, including element name, definition and cataloging rules.
2. Code table of metadata elements, e.g. control vocabulary.
3. Metadata cataloging examples.
4. Metadata element attributes, including data type, necessary and multiple value.
5. Metadata unique identifier system.
6. Data relationship diagram: expresses the level and relationship of objects in a structural drawing, e.g. fonds, series, volume and document.
7. Association diagram of metadata elements, grouping of metadata elements and linkage relationship.
8. Functional requirements, e.g. input and Chinese character and Japanese character display, default values and link functions.
9. Data query and display requirements, includes different level searches (key words, advanced search) and display (summary, detailed).

10. Metadata management requirements: e.g. file creation process and access rights control.

The work sheets above mainly allow project personnel to:

1. Understand the scope, association, relation and attributes of different metadata categories.
2. Verify the scope of the system or database integrated by metadata, e.g. integration with GIS.
3. Acquire application examples of other projects as a basis for best practices.

Procedure 4: Verify the Metadata Strategy and Interoperability of Standards

Based on previous studies, metadata strategies recommended for digital archive projects include adopting a single metadata standard or multiple metadata standards, or developing a suitable metadata format based on existing metadata standards.

Stage 2: Metadata Functional Requirements

Procedure 5: Compile the Metadata Functional Requirements Document

The purpose of the metadata functional requirements document is to serve as a bridge of communication between the digital archive project, metadata analysis and system development, and build a consensus between different professional fields. Contents of this document include: the document's version and management information, project background, project members, system objectives and scope, standard adopted, metadata elements and structure, attributes of metadata elements (e.g. name, length, data type and primary key), output example, comparison with related standards, association diagram of system scope, system functional requirements (e.g. input and display of Chinese or Japanese characters), control vocabulary or code list, and XML DTD. To summarize the above, main functions of the requirements document include:

1. allowing the digital archive project to verify its metadata requirements.
2. serving as a liaison for the metadata task force and system designers.
3. serving as a basis for revising the metadata system and its functions.
4. comparing the metadata standard with best practices.

Procedure 6: Metadata System Evaluation

Evaluate the feasibility of metadata system development so the project can decide whether to adopt the system of a project with the same or similar characteristics, develop a new system, or jointly develop a system with other organizations (e.g. university or company).

Stage 3: Metadata System

Procedure 7: Best Practices

The purpose of best practices includes: providing application guidance for metadata elements, providing checklists or references for projects to apply standards, or providing a manual for controlling the quality of metadata records. Contents of best practices should include definitions of metadata elements, cataloging principles, system recommendations, cataloging examples and comparison with other metadata standards.

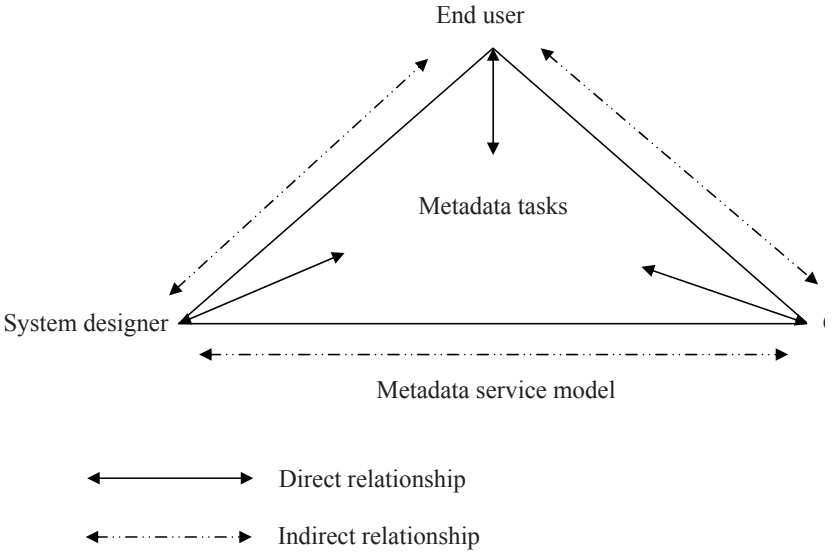
Procedure 8: Metadata System Development

At this stage metadata development in complete and we begin with system development procedures. The mission of system development is to develop a metadata system and tools that conform to the requirements document. During system development, project members, metadata and system developers should continuously exchange and discuss opinions to drive more effective system design. After the system prototype is completed, project personnel and metadata analyzers should feedback system testing and evaluation results for system developers to make adjustments.

Stage 4: Services and Evaluation

Procedure 9: Metadata Services

To control the quality of the metadata mechanism, we planned a “metadata service model” based on the role, relationship and service items of each service mode:



Service Mechanism:

1. Increase the utilization value of project and user related functions in the user interface.
2. Consultation support: Provide consultation to project members regarding metadata and content analysis.
3. Revise metadata system functions: Propose system function revisions based on system testing results.
4. Establish metadata interoperability mechanisms, such as Crosswalk and Meta-search.
5. Provide and harvest metadata records to help metadata become interoperable.
6. Complete a crosswalk of current records, project requirements and elements of the standard adopted for system designers to use as reference.
7. Develop a Chinese version metadata standard.

Role: Members involved include users of the digital archive project (general or content experts), system designers, and the metadata task force.

Relationship: Includes direct and indirect relationships. Direct relationships exist between the metadata task force and general users, and between content experts and system designers; indirect relationships exist between general

users, contents experts and system designers.

Procedure 10: Metadata Operations Evaluation

Examine overall metadata implementation procedures and benefits according to project requirements. Evaluation items include:

1. Quality of metadata records: includes completeness, accuracy, record type, granularity and serviceability. For example, evaluating metadata search quality in terms of decision-making and procedures.
2. Effectiveness of the metadata standard in terms of search function: For example, evaluating the effectiveness and accuracy of searching using name, creator and theme.
3. Usability of metadata production tools: For example, evaluating development benefits of metadata tools and focuses of subsequent designs.
4. Verify implementation benefits of every operating procedure: Verify if it is necessary to implement a certain operating procedure again, such as gaining an even more in-depth understanding of metadata requirements.

Appendix 3: Metadata Requirements and the Dublin Core

Metadata requirements specifications for plants: Metadata requirements specification for native plants in Taiwan

(1) 物種名錄欄位元素

項目名稱		英文名稱	DC
植物學名索引碼		Code of Plant Scientific Name	IDENTIFIER
界名	拉丁界名	Latin Kingdom Name	TITLE
	中文界名	Chinese Kingdom Name	TITLE
門名	拉丁門名	Latin Division Name	TITLE
	中文門名	Chinese Division Name	TITLE
綱名	拉丁綱名	Latin Class Name	TITLE
	中文綱名	Chinese Class Name	TITLE
目名	拉丁目名	Latin Order Name	TITLE
	中文目名	Chinese Order Name	TITLE
科名	拉丁科名	Latin Family Name	TITLE
	中文科名	Chinese Family Name	TITLE
屬名	拉丁屬名	Latin Genus Name	TITLE
	中文屬名	Chinese Genus Name	TITLE
種小名		Species Epither	TITLE
種命名者		Species Author	CREATOR
亞種名		Subspecies Epithet	TITLE
亞種命名者		Subspecies Author	CREATOR
變種名		Variety Epithet	TITLE
變種命名者		Variety Author	CREATOR
品種名		Form Epithet	TITLE
品種命名者		Form Author	CREATOR
栽培變種名		Cultivariety Epithet	TITLE
栽培變種命名者		Cultivariety Author	CREATOR
中文名		Chinese Name	TITLE
中文別名		Other Chinese Name	TITLE
英文名		English Name	TITLE
原始發表年代		Scientific Name Publication Yea	DATE-CREATED
原始發表文獻		Scientific Name Publication Literature	PELATION-REFERENCES
備註		Remarks	DESCRIPTION

(2) 標本紀錄欄位元素

項目名稱		英文名稱	DC
標本典藏資料			
標本館號		Specimen Order Number	IDENTIFIER
標本編號		Specimen Serial Number	IDENTIFIER
典藏單位	典藏單位	Herbarium/Institute	RIGHTS
	典藏單位代碼	Herbarium/Institute Code	IDENTIFIER
標本狀況		Specimen Status	DESCRIPTION
計畫名稱		Project	RIGHTS
建檔日期		Record Creation Date	DATE-CREATED
建檔人員		Record Creation Operator	CONTRIBUTOR
出借狀況	出借紀錄編號	Loan Serial Number	IDENTIFIER
	出借日期	Loan date	DATE
	借出單位	Borrower	DESCRIPTION
	預計歸還日期	Expected Loan Data	DATE
標本交換	交換狀況	Exchange Status	DESCRIPTION
	標本交換單位代碼	Exchange Herbarium/Institute Code	IDENTIFIER
	標本交換日期	Exchange Date	DATE
其他		Remarks	DESCRIPTION
標本採集資訊			
採集者	採集者代號	Collector Code	IDENTIFIER
	採集者中文名	Collector Chinese Name	CREATOR
	採集者英文名	Collector English Name	CREATOR
採集編號		Collection Number	IDENTIFIER
採集日期		Collection Date	DATE
複份標本數量		Duplicate	FORMAT-EXTENT
隨同人員		Companion	CONTRIBUTOR

*__IstAdministrative__ → Ist Administrative
2ndAdministrative → 2nd Administrative
3rdAdministrative → 3rd Administrative
Aditional Description → Aditional Descriptions

項目名稱		英文名稱	DC
證據標示		Voucher	DESCRIPTION
標本類型		Specimen Category	TYPE
溫室栽培		Greenhouse Cultured	DESCRIPTION
壓製與否		Greenhouse Specimen	DESCRIPTION
標本採集地理位置			
國別／地理區域	國別代碼	Country Code	IDENTIFIER
	中文國名	Country(Chinese)	COVERAGE-SPATIAL
	英文國名	Country(English)	COVERAGE-SPATIAL
行政分區	一級行政分區中名	1 stAdministrative Area(Chinese)	COVERAGE-SPATIAL
	一級行政分區英名	1 stAdministrative Area(English)	COVERAGE-SPATIAL
	二級行政分區中名	2ndAdministrative Area(Chinese)	COVERAGE-SPATIAL
	二級行政分區英名	2ndAdministrative Area(English)	COVERAGE-SPATIAL
	三級行政分區中名	3rdADministrative Area(Chinses)	COVERAGE-SPATIA
	三級行政分區英名	3rdAdministrative Area(English)	COVERAGE-SPATIAL
國家公園／保護區	國家公園／保護區中名	National Park/Reserve(Chinese)	COVERAGE-SPATIAL
	國家公園／保護區英名	National Park/Reserve(English)	COVERAGE-SPATIAL
地名	中文地名	Locality(Chinese)	COVERAGE-SPATIAL
	英文地名	Locality(English)	COVERAGE-SPATIAL
其他描述		Additional Description	COVERAGE-SPATIAL
海拔高度／下限		Altitude/ Minimum Altitude	COVERAGE-SPATIAL
海拔高度／上限		Altitude /Maximum Altitude	COVERAGE-SPATIAL
東／西經		Longitude Direction	COVERAGE-SPATIAL
經度(度／分／秒)		Longitude/Deg/Min/Sec	COVERAGE-SPATIAL
南／北緯		Latitude Direction	COVERAGE-SPATIAL
緯度(度／分／秒)		Latitude(Deg/Min/Sec)	COVERAGE-SPATIAL
臺灣二度分帶座標(X)		Taiwan Grid(X)	COVERAGE-SPATIAL

Tropic Mode → Trophic Mode

項目名稱		英文名稱	DC
標本採集地理位置			
臺灣二度分帶座標(Y)		Taiwan Grid(Y)	COVERAGE-SPATIAL
標本棲地資訊			
植群帶	植群帶(中文)	Vegetation Zone(Chinese)	TYPE
	植群帶(英文)	Vegetation Zone(English)	TYPE
植群型	植群型(中文)	Vegetation Type)Chinese)	TYPE
	植群型(英文)	Vegetation Type(English)	TYPE
地形位置		Topography(English)	COVERAGE-SPATIAL
自然度	自然度(中文)	Degree of Naturalness(Chinese)	TYPE
	自然度(英文)	Degree of Naturalness(English)	TYPE
微生育地		Habitat	COVERAGE-SPATIAL
環境光度	環境光度(中文)	Light Intensity(Chinese)	DESCRIPTION
	環境光度(英文)	Light Intensity(English)	DESCRIPTION
環境濕度	環境濕度(中文)	Humidity(Chinese)	DESCRIPTION
	環境濕度(英文)	Humidity(English)	DESCRIPTION
豐富度	豐富度(中文)	Abundance(Chinese)	DESCRIPTION
	豐富度(英文)	Abundance(English)	DESCRIPTION
其他棲地資訊		Habitat Remarks	DESCRIPTION
標本特徵資訊			
生長型		Habit/Life Form	TYPE
植株高度		Plant Height	DESCRIPTION
營養方式(／伴生方式)	營養方式(／伴生方式)(中文)	Tropic Mode (Chinese)	TYPE
	營養方式(／伴生方式)(英文)	Tropic Mode (English)	TYPE
寄主代碼(／伴生植物代碼)		Host Code	IDENTIFIER

Stage(Chinese) → Stage (Chinese)
其他照此修改

項目名稱		英文名稱	DC
物候資料－花期	物候資料－花期(中文)	Phenology-Flowering Stage(Chinese)	DESCRIPTION
	物候資料－花期(英文)	Flowering Stage(English)	DESCRIPTION
物候資料－果期	物候資料－果期(中文)	Phenology-Fruiting Stage(Chinese)	DESCRIPTION
	物候資料－果期(英文)	Phenology-Fruiting Stage(English)	DESCRIPTION
花色		Flower Color	DESCRIPTION
果色		Fruit Color	DESCRIPTION
其他特徵／備註		Additional Character	DESCRIPTION
標本鑑定資訊			
植物學名索引碼		Code of Plant Scientific Name	IDENTIFIER
鑑定者	鑑定者中文姓名	Identifier/Determined by(Chinese)	CREATOR
	鑑定者英文姓名	Identifier/Determined by(English)	CREATOR
鑑定日期		Identification Date	DATE-CREATED
分類評注		Name Comment	DESCRIPTION
標本訂正資訊			
館號		Specimen Order Number	IDENTIFIER
訂正資訊	訂正紀錄編號	Verification Serial Number	IDENTIFIER
	訂正後植物學索引碼	Verified Code of Plant Scientific Name	IDENTIFIER
	訂正者 訂正者中文姓名	Verifier(Chinese)	CREATOR
	訂正者 訂正者英文姓名	Verifier(English)	CREATOR
	訂正日期	Verification Date	DATE-MODIFIED
	建檔日期	Record Creation Date	DATE-CREATED
	分類訂正意見	Taxonomic Identification Opinions	DESCRIPTION

(3) 文獻資料欄位元素

項目名稱		英文名稱	DC
文獻紀錄編號		Code of Literature	IDENTIFIER
植物學名索引碼		Code of Plant Scientific Name	IDENTIFIER
文獻標題	篇名	Title	TITLE
	語文	Language of Title	LANGUAGE
其它標題	篇名	Author	TITLE
	語文	Language of Other Title	LANGUAGE
作者	姓名	Author	CREATOR
	著作方式	Type	CREATOR
	不同語言名稱	Other Name	CREATOR
文獻出處	書刊名縮寫	Reference Abbreviation	SOURCE
	書刊名	Reference Title	SOURCE
	卷	Volume	SOURCE
	期	Number	SOURCE
	頁碼	Pages	SOURCE
出版類型		Type	TYPE
出版者		Publisher	PUBLISHER
出版地點		Publish Location	COVERAGE-SPATIAL
出版年		Publish Year	DATE-CREATED
資料識別代碼	國際標準書號ISBN	ISBN	IDENTIFIER
	國際標準叢刊號ISSN	ISSN	IDENTIFIER
關鍵字		Keywords	SUBIECT
建檔日期		Record Creation Date	DATE-CREATED
註		Notes	DESCRIPTION

(4) 影像資料欄位元素

項目名稱		英文名稱	DC
影像編號		Code of Image	IDENTIFIER
影像來源		Origin of Image	FORMAT-MEDIUM
影像品質		Quality of Image	FORMAT-EXTENT
影像主題		Subject of Image	SUBJECT
圖說		Illustration	DESCRIPTION
植物學名索引碼		Code of Plant Scientific Name	IDENTIFIER
作者		Author	CREATOR
影像提供者		Image Provider	CONTRIBUTOR
拍攝日期		Photograph Date	DATE-CREATED
採集者	採集者(中文)	Collector(Chinese)	CONTRIBUTOR
	採集者(英文)	Collector(English)	CONTRIBUTOR
採集號		Collection Number	IDENTIFIER
採集地點(中文)		Collection Location	COVERAGE-SPATIAL
建檔日期		Record Creation Date	DATE-CREATED
影像檔案名稱		Image File Name	IDENTIFIER
備註		Notes	DESCRIPTION

Source: Digital Archives Technology Collection 2007[40]

[40] Digital Archives Technology Collection 2007, Search: February 2010 http://www2.ndap.org.tw/

Appendix 4: Contractor Interviews

Industry-Academia Cooperation – Outsourcing Management: Experience Sharing of Contractors / Liang Mei-Chen

Introduction

At present, most repositories in Taiwan are already engaged in digitization work. A number of publications can provide insight on digitization operations, including equipment and digitization related books and articles, as well as a series of "digitization procedures guidelines" of the Taiwan Digital Archives Expansion Project compiled for different objects. In which equipment, such as scanners and digital backs, is an important medium for digitization, changing collections into image files that can be browsed online or permanently preserved. Image quality is determined by the functionality of the equipment used, better equipment is more expensive and may even need to be imported, but may not be within the budget of every institution. Furthermore, whether or not the professional competencies of human resources are sufficient is also a factor that must be considered. Therefore, this guideline introduces "outsourcing management" with a focus on important elements of digitization work – equipment and human resources (including professional knowledge).

Digitization work can be implemented two ways: one is for the institution to implement digitization itself (purchase equipment and allocate manpower) the other is to seek assistance from a contractor (outsourcing, the contractor provides equipment and manpower) when the institution can neither purchase high-end equipment nor has sufficient professional knowledge. The main considerations for deciding between the two are low cost, short time period, good quality and high output. With consideration to the preciousness of collections, the best option is of course for institutions to complete digitization operations themselves. However, if there are budget limitations, then institutions might have no choice but to work with a contractor. Due to the importance of their collections, some institutions remain skeptical about outsourcing digitization; this might be because they do not understand outsourcing procedures, or had bad experience in the past. Therefore, we use this interview to discuss outsourcing from the perspective of a contractor,

---

eBook08/showContent.php?PK=100.

and then offer suggestions based on their experience. We hope to explain outsourcing through the experience of a contractor, and provide institutions that intend to implement digitization projects but have relatively insufficient financial and material resources with an additional option.

This article interviews Mr. Li Su, General Manager of CX Media, and asks him to share his cooperation experiences as an "outsourcing contractor" of plant specimen digitization, providing reference for other project units when implementing "outsourcing management."

Outsourced Operations Which is Better?

Generally speaking, outsourced operations can be divided into: stationing personnel and equipment in the repository or moving collections to the contractor for digitization. When precious collections are handed to a contractor for digitization, the institution sustains certain risks: whether or not collections are treated as required, communication problems, quality control and process delay. Sometimes, evaluations are properly conducted, but both parties are still exhausted by the time the case is concluded, and the results are far from ideal. Still, pleasant cooperation experiences with contractors are very common, whether communication and coordination is effective depends on whether if both parties can earnestly make adjustments. As mentioned above, the appropriateness of equipment selected plays an important role in the digitization workflow; most contractors are able to provide appropriate equipment. If project units do not have the funding to purchase required equipment, they may establish outsourcing operations specifications (includes contract) based the quantity and requirements of their collections and make a request for tender.

From the perspective of a contractor, Mr. Li indicates that their company currently decides whether or not to become stationed in an institution based on the size of the collection, and that they will recommend sending collections to their company if the collection is small and requires short digitization time. On the contrary, if institutions require equipment to be stationed, it will depend on whether or not the cost is affordable, other considerations include: work efficiency, preciousness of the collection and output. With consideration to the high cost of equipment, contractors are sometimes forced to stand on principle when deciding whether or not to become stationed in an institution. Mr. Li

recommends fully evaluating the case, including equipment characteristics, object characteristics, and digitization method or difficulty, before deciding whether to take it or not. Notices on bidding are described in the following section.

Whether to Bid or Not Emphasis on Evaluation

Concerning whether to bid or not, according to Mr. Li's experience, he will first evaluate the contents and conditions of the tender, as well as whether or not his equipment has the functions to process special objects. Currently, equipment used is large and expensive and hard to move, so the area is considered if the equipment is to be stationed in the repository. With consideration to the issue of moving equipment and personnel, CX Media currently only stations its equipment and personnel in institutions in the Taipei area. From the company's position, whether or not to bid is based on cost effectiveness and characteristics of the collection.

Although contractors assist with only one link of the digitization workflow, it is the most important link. "Even so" says Mr. Li "Contractors still need to have a certain degree of understanding of the entire digitization workflow to be clear of the institution's requirements. Contractors may even make recommendations based on their expertise and past experiences. To both parties, this is not only a cooperative relationship, but also an opportunity to learn from one another, yielding twice the result with half the effort."

Operation of Mr. Li's company: At the beginning, the company makes recommendations and addresses possible issues based on the object and requirements, or inquires with the institution regarding object characteristics. Then the company designs and formulates a digitization workflow; equipment and collections are always elements in the workflow. After gaining a sufficient understanding, adjustments are made according to special characteristics. Once adjustments are complete, subsequent procedures will be smooth and successful, and work is usually completed as scheduled. This shows the importance of being fully prepared.

The second stage of work is preliminary planning of scanning. As usual Mr. Li formulates a complete set of operation rules and the complete production process, and then adjusts the process or operation method according to the client's requirements. Basically, this part is actually operated, from setting

up equipment, turning devices on, placing objects, setting the focal distance, operating principles and procedures to even image clarity. If project personnel have any suggestions or requirements in the process, then time is spent communicating and making adjustments. The purpose for this process is to formulate a suitable workflow, so that when digitization formally begins, operations will be smooth and successfully with relatively less errors occurring.

Communication before digitization is a critical part of the entire workflow to Mr. Li, he would rather spend time resolving issues and coordinating until a consensus is reached than make countless mistakes during formal operations. As long as preliminary operations are properly carried out, the need for post-processing and less time cost will be required, which is without doubt best for both parties.

Personnel Change Remedial Plan

Personnel flow is common in the workplace, but is a real headache for project units and contractors; project units are especially worried that personnel change of the contractor will affect operating procedures. Mr. Li shares his experience with overcoming this issue: Personnel begin training once they enter the company. They are required to abide by work rules if they are assigned to a specific unit and must understand the digitization workflow before they may be stationed there. During training, personnel memorize all work rules and are careful not to damage specimen in the digitization process, while strictly abiding by specifications of digitization results. Once they clearly understand all concepts, all that's left is for them to accumulate proficiency through experience. If personnel encounter problems that need to be handled using special methods when implementing digitization, for example: inconsistent thickness of a vascular plant specimen (difference in thickness or height between the cone and leaves) that requires additional adjustment of the focal distance to be able to scan a clear image, then they are required to immediately report the problem to their supervisor. This way the problem is rapidly resolved without delaying the digitization schedule. After scanning personnel accumulate a certain amount of experience, they might discover even more efficient operation models and be able to increase work efficiency after confirming with supervisors and coordinating with the project unit. Mr. Li's company effectively resolves issues via communication and supervision,

ensuring that there remain personnel or supervisors clear of the work progress if an operator leaves the job and can immediately fill in the position; fixed workflows and rules are thus all under supervision. With personnel ready to provide support at all times, the possibility of work being suspended is eliminated, allowing digitization operations to be successfully implemented.

Special Situations Adapt to Circumstances

Due to the uniqueness of different objects, different solutions need to be adopted for digitization operations. “Plant specimens” are unique in that they include “vascular plants” and “non-vascular plants.” In Mr. Li’s experience working with project units, he once encountered issues with scanning “moss and lichen”: there are several tufts of moss or lichen on one mounting sheet, each with different height, making it more complicated in terms of focal distance settings compared with general vascular plant specimens; scanning personnel had to spend more time on adjusting and watching focal distance settings to maintain a clear image of all specimens on the mounting sheet. A lot more attention is paid to each specimen on the mounting sheet and the machine needed to be reset for every new mounting sheet. Sometimes it took up to 20 minutes to scan just one mounting sheet; usually efficiency is lost when quality is prioritized. Different scanning procedures are used according to object characteristics, and must be clearly planned in advance. If special situations are discovered only until formal digitization operations, it will result in unimaginable cost.

Basically speaking, plant specimens that are dried and pressed on mounting paper are scanned as solid objects in digitization operations. Besides complex settings required for moss and lichen specimens, plant specimens are characterized by having different heights: such as cones, so in addition to requirements on depth of field, overall clarity (every leaf displayed on the mounting paper) should also be emphasized; for such specimens digitization should not be a pursuit of resolution, but instead clear and solid images obtained from scanning. The most common vascular plant specimens are plants with only dried leaves and branches, for which project personnel are responsible for cleaning any leaf and branch fragments and contractors are responsible for scanning; scanning operations of this type of specimens is much easier.

The majority of plant specimens are vascular plants, cones are only a small part of plant specimens, but both are classified as “plant specimens” to contractors. When considering a request for tender, objects are almost never seen on site, sometimes a few specimens are provided, but most information is purely theoretical. Once difficulties are encountered during actual implementation, different characteristics require different methods and execution time, so the schedule might be delayed without prior planning. Without a clear understanding and consensus between both parties before implementation, it is hard to avoid extreme time cost from appearing.

From Mr. Li’s perspective, project units might neglect such issues, but from his position contractors emphasize the quality of output results, and it is not their concern how much time or manpower is spent. Situations occur because matters are not stated clearly in the specification or contract. For this reason if both parties do not fully communicate and have an understanding, their cooperative relationship will be affected once they encounter an issue that is hard to immediately resolve, and both parties will sustain great loss. Besides the situation described above, Mr. Li believes that other issues contractors will encounter include:

(I) Difference in acceptance standard: The difference refers to inconsistent acceptance standards at different levels of the cooperative unit. Contractors must find a way to overcome various difficulties and maintain flexibility to resolve all issues and accomplish their mission.

(II) Understanding of object characteristics and the digitization workflow: Even though contractors only handle a part of digitization operations, if they do not understand the object and lack the ability to make proper decisions, it will cause them to spend more time and effort to resolve issues encountered during digitization; for example, dry plant specimens such as scanning moss and lichen take longer time. Sometimes, whether the object is easy to digitize is a key issue.

Key to Success　Careful Contractor Selection

If project units need to outsource operations, Mr. Li makes the following suggestions for contractor selection:

(I) Primary conditions to evaluate: First, evaluate the contractor’s technology and equipment; this reduces time cost and achieves high quality output.

Then, consider excellent contractors recommended by other units; contractors with rich experience may have stronger communication and problem solving abilities to reduce operation time. Finally, see if the price is reasonable. Use these three suggestions for the first screening stage to select a proper contractor.

(II) Examine the contractor's performance in past cases: It is recommended to require contractors to provide results of past cases for the project unit to compare; the first consideration is whether or not the contractor's equipment can fully present object characteristics and create clear images. Besides display on a computer monitor, it is recommended to consider future possibilities, such as publication, so use the images for output and digital proofing to avoid not being able to print the images in the future; this prevents selecting contractors with insufficient technology.

(III) Actual operation evaluation: There's no harm in setting a time period (1 or 2 weeks) as a last chance for both parties to decide whether or not to cooperate based on actual results, and further establish contract contents and specifications; this avoids a difference in results presented before and after implementation.

(IV) Importance of communication and coordination: Both parties must reach a consensus on work specifications and acceptance standards before implementing digitization operations, so that digitization work can be successfully implemented without being affected by inconsistent opinions in the process. Mutual trust and sincerity in making adjustments are the key to successfully completing work.

Conclusion

As civilizations around the world have developed until today, computer technology has earned a place of its own, turning digitization into a global trend. Outsourcing boosts industry-academia cooperation, which expands market opportunities for businesses and allows precious relics to be permanently preserved. Industry-academia cooperation stimulates the domestic digital contents industry, and enhances the international competitiveness of Taiwan's academia and digital technology.

Here we would like to thank Mr. Li for accepting our interview, and for sharing his rich experiences and offering precious suggestions. Mr. Li

emphasized the importance of “preliminary operations” and “communication” again and again, and indicated that if these two principles were followed in management or workflows, then it will yield twice the results with half the effort.

Vascular Plants Digitization Procedures Guideline

Advisory Unit: National Science Council, Executive Yuan, R.O.C.
Issuer: Simon C. Lin
Editor-in-Chief: Simon C. Lin
Executive Editors: Meng-Yin Wu, Yu-Ju Lin
Authors: Lang-Hsuan Kao, Mei-Chen Liang
Translator: Tai-Yu Chen
Reviewer: Professor Chang-Fu Hsieh, Institute of Ecology and Evolutionary Biology, National Taiwan University
Publisher: International Collaboration and Promotion of Taiwan e-Learning and Digital Archives Program
Address: No.128, Sec.2, Academia Rd., Nankang District, Taipei City, 115 Institute of Physics, Academia Sinica
Tel: +886-2-2789-8311
Fax: +886-2-2783-5434
Website: http://collab.teldap.tw
Email: teldap-collab@twgrid.org
Typesetting and Printing: EVERGREEN INTERNATIONAL CORP.

Publish Date: 1st Edition July 2012
ISBN:

本書譯自拓展臺灣數位典藏計畫出版之
數位化工作流程指南：維管束植物