

Development of an Archival Management Workflow: the Macau Case

Veng-Ian Chan

Faculty of Science and Technology
University of Macau
Macau
ma76505@umac.mo

Yain-Whar Si

Faculty of Science and Technology
University of Macau
Macau
fstasp@umac.mo

Abstract—Archive is a collection of historical records which have been accumulated over the course of an individual and institution's lifetime. Archive is considered as an invaluable asset which provides us with an insight into the lives and circumstances of past generations. In this paper, we describe the workflow model of an archival management system adopted by Macau Historical Archives. To further analyze the strength and the weaknesses of the current system, the conceptual workflow model from Macau Historical Archives is mapped with OAIS functional model. As the project is still ongoing, the paper mainly studies the modeling requirements, discusses the difficulties have to overcome, and concludes with our future work.

Keywords-archival management; archival process; workflow

I. INTRODUCTION

Archive [1] is a collection of records and documents which need to be kept and conserved as an invaluable asset. Although more and more recent archival management systems are designed for preserving digitized data, a manual archival process [1] by experienced technicians is still needed for handling of traditional paper-based records. In addition, performing such archival management process often requires understanding of historical information, government policies, local custom, and tradition. As a result, archival management processes are usually different from one country to another and these systems are usually tailored for specific needs of a local government.

Operations in an archival process are performed within the archival lifecycle [1]. Archival lifecycle includes tasks such as appraisal, accession, removing contaminants, measuring, initial repair, arrangement and descriptions, digitization, preservation, and utilization by end users. Some of those archival processes involve manual operations and effectiveness of these tasks significantly depends on human decision and implicit knowledge. Archival management processes are best suited for being designed as workflow models [2]. Workflow is a depiction of a sequence of operations, defined as composite tasks that comprise both human and computational subtasks. Workflow management systems are widely adopted for the implementing business processes. Archival workflow [3] is a precise representation of the archival process, reflecting the formal coordination mechanisms between archival activities

and manual processes with respect to the specific requirements of the local governing body.

Integration of workflow modeling concept into archival management problem allows efficient handling of complex and sometime tedious tasks in preservation procedure. However, care must be taken when designing archival workflow since the entire process model is required to be complied with security, legal and standardization requirements. By adopting workflow technology, archival management tasks can be effectively designed to conform to local and international standards. In addition, clear understanding and stringent control of internal archival procedures can also provide organizations with the ability for achieving long term preservation objectives.

Integration of workflow technology also allows better control over retraceability, accountability, and separation of logic (process definition) from archival data. Although such integration proved to be beneficial to archival management community, only a limited number of researches (e.g. [5], [6], [7]) have been carried out in that direction. To the best of authors' knowledge, conceptual as well as implementation issues associated with such integration have not been reported so far in the related literature. These issues include control flow design for semi-automated preservation tasks, resource assignments, security, localization as well as standardization, compatibility, and future extension. Against this background, this paper aims to provide an in-depth analysis on modeling archival workflows based on international standards. Our analysis is demonstrated through the workflow model of an archival management system adopted by Macau Historical Archives.

The paper is structured as follows. A generic archival workflow model is described in section 2. In section 3, we detail the archival management process for Macau. Comparison of archival process with OAIS Model is described in section 4. In section 5 we briefly review related work before summarizing our ideas in section 6.

II. ARCHIVAL MANAGEMENT WORKFLOW MODEL

Modeling archival management process has following distinct characteristics:

1. It is impractical to automate all the processes within archival workflow since some of the tasks are complex and can only be handled by archival specialists. Degree of automation may differ from one area of the workflow to another. For instance, digital records can be archived in more automated way compared to paper-based records.
2. Archival processes are required to follow the international/local archival management standard and cannot be changed arbitrarily.
3. In addition to streamlining workflow processes, other factors from archival management process such as safeness, boundless and reachable should be carefully considered during the design phase. Safeness is used to ensure that the archive documents are securely processed within the archival life cycle whereby taking into account factors such as access right and privacy control (copy rights issues). Instead of limiting the acceptable types of archival document and handling methods, boundless property is used to ensure that existing system can be extended in future to accommodate new requirements. Reachable property is used to ensure that the archived documents can be accessed by public.
4. In any archival workflow, it is crucial to preserve the records' integrity, originality, and authenticity. Archived records are mostly deposited by various government departments and therefore the entire archival process must be accountable for legal reasons.

Based on these underpinnings, we conceptualize a generic archival workflow model in Figure 1. First, records for archiving are evaluated in the appraisal task. Appraisal is the process of assessing whether these records have sufficient value to warrant acquisition by an archival institution. After initial appraisal task, these records are transferred to the archive institute. Next, initial conservation such as basic cleaning is done before they are being grouped in arrangement process. Then the records are forwarded to digitization process and archive description process based on the task assignment schedule. Digitization process transforms the physical archive into the digital archive. After the digitization process is completed, the digitized images are assigned with the corresponding index and description. Archive description process is carried out to analyze, organize and record details of the archive, according to the international description standard.

The next steps in the process involve creating backups and rebinding archives. In backup process, the original files are copied into storage media so that it can be restored if the original data is deleted or damaged. Archive rebinding repackages the archive or the backup media for permanent storage. Finally these records are stored at the permanent storage. Depending of the nature of the archived material, periodic maintenance tasks (i.e. preservation) for the stored record is also scheduled at the end of the process.

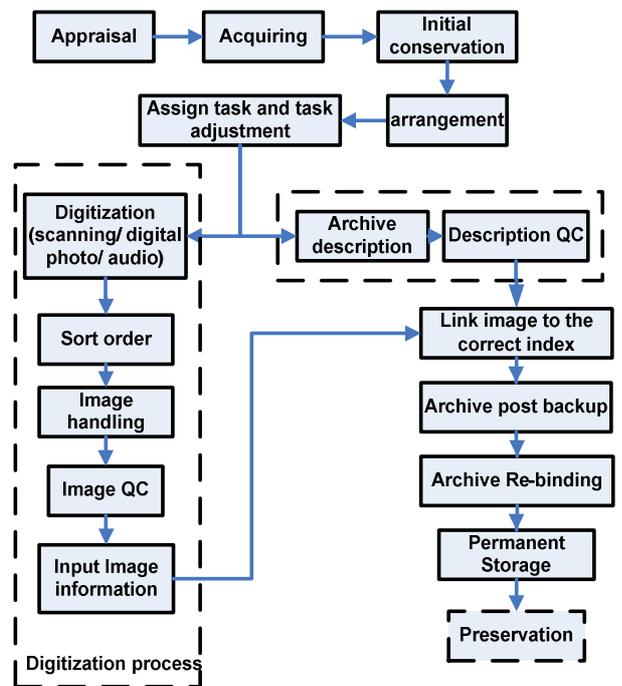


Figure 1. Conceptual archival workflow model

III. CASE STUDY: MACAU HISTORICAL ARCHIVES

Archival management task in Macau was set up in 1952 under the name of “Macau General Archives”. Later, it has changed its name into “Macau Historical Archives” and incorporated into Cultural Affairs Bureau of Macau in 1986. The facilities currently offer safe storage for valuable historical records. Macau Historical Archives plays a key role in the management of recorded information for the Government of Macau Special Administrative Region (Macau SAR). Macau Historical Archives is responsible for preserving invaluable historical records and making them available for the community.

Macau Historical Archives comprises of several working groups including archive, conservation, surrogating and public service. In total, approximately 45000 records are preserved in the archive. These records include articles, books, slides, microfilms, maps, photos and electronic files. Some of the records are acquired from legislatures, other governmental bodies and also from private bodies.

At present, archival management process in Macau is done in semi-automated fashion. A number of issues are observed in Macau archival management process.

Time consuming and limited performance

At present, all archiving tasks are required to be prescheduled on yearly basic. These tasks are organized into batches and assigned to each staff at the beginning of a year. Based on these schedules, a technical staff may process a batch of records/assets and pass it on to another working group after it is completed. Scheduling tasks for an entire year and then processing in batches is inflexible when there are unexpected delays in the process. Moreover, performance of preservation tasks is manually monitored and therefore audit trail data is difficult to be obtained for analysis. As a result, more efficient resource allocation schemes are difficult to be employed.

Security of archival management

Archives are invaluable assets and therefore it is essential to maintain highest level of security for these records. At present, only staff member can assess/process the original archive material. Individual staff members are also carefully selected and assigned for handling archive records based on a predefined schedule. Although the archive documents are strictly controlled by Macau Historical Archives, a number of issues are identified for further improvement. For instance, take-out history is still done manually and there is no tracking mechanism for the archive which has been lent out. In addition, there is not sufficient information for staff to trace the document in case it is missing.

Hard to confirm the consistency and relationship of the archival description

Preparing description for the record being archived is done manually on paper-based media and the responsible staff is required to do a series of processes in order to ensure the description is correct. In addition, he/she may decide how to split or merge the file based on his/her experience. In order to make these decisions, the staff may need to check the content of the original archive. Such checking can only be done by submitting a request to access the original archive and

therefore may cause delays in processing the original archive. Inconsistence in description may also occur due to such access restriction policy.

Official language integration

Chinese and Portuguese are the official languages of Macau. At present, Macau Historical Archives provides archival records description mainly in Portuguese and in Chinese. English language service is not provided at the moment.

Limited Service

At present, users are allowed to access the abstracted information about the archive via the computer terminals located in Macau Historical Archives. For borrowing archived material, a user has to personally contact the front desk and fill in the request form. After validation, archive staff may lend the physical copies to them. Ideally, users should be able to access the records (e.g. digitized archival material) anytime via Internet instead of being restricted by such physical limitations.

In order to alleviate these limitations, Macau Historical Archives is recently undertaking a series of development projects to improve the current archival process based on automated workflows. These workflows are designed in accordance with the existing legal requirements of Macau and provisions are also made for future extension. Macau Historical Archives also recognizes the importance of the international archival standards such as ISAD(G) [3] Descriptive standard and OAIS model [4]. ISAD(G) is a general International Standard Archival Description, defined by International Council on Archives (ICA). ISAD(G) outlines the general rules for preparing archive description. OAIS functional model (see Figure 2) is a reference model for archival process. The reference model consists of producer and consumer of information. In this reference model, emphasis is given on digital information, both as the primary forms of information held and as supporting information for both digitally and physically archived materials.

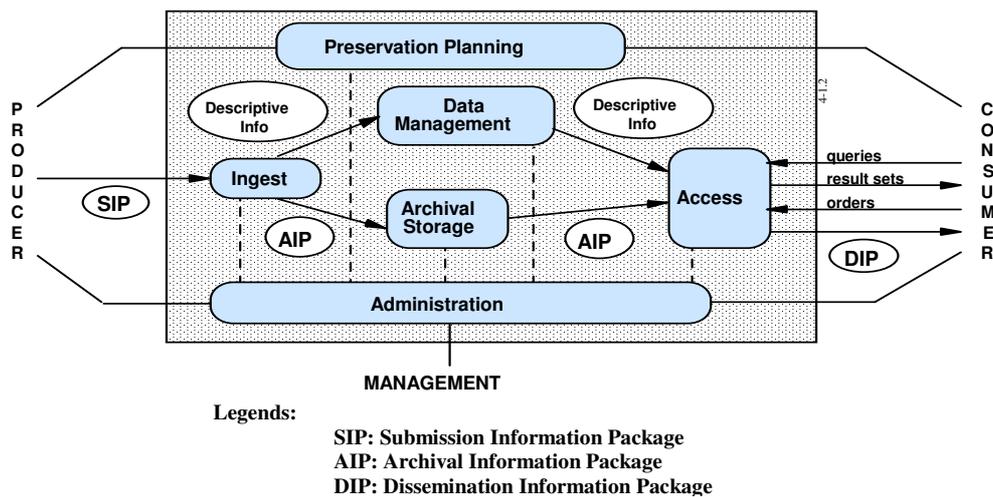


Figure 2. OAIS Function Entities [4]

To further illustrate the Macau archival process, an activity-based conceptual workflow is depicted In Figure 3. Activity-based workflow modeling is a sequence of discrete actions and relationships to be performed to achieve a business objective.

The diagram shows an archival management lifecycle of Macau, starting from an initial appraisal process and terminating at the public access function.

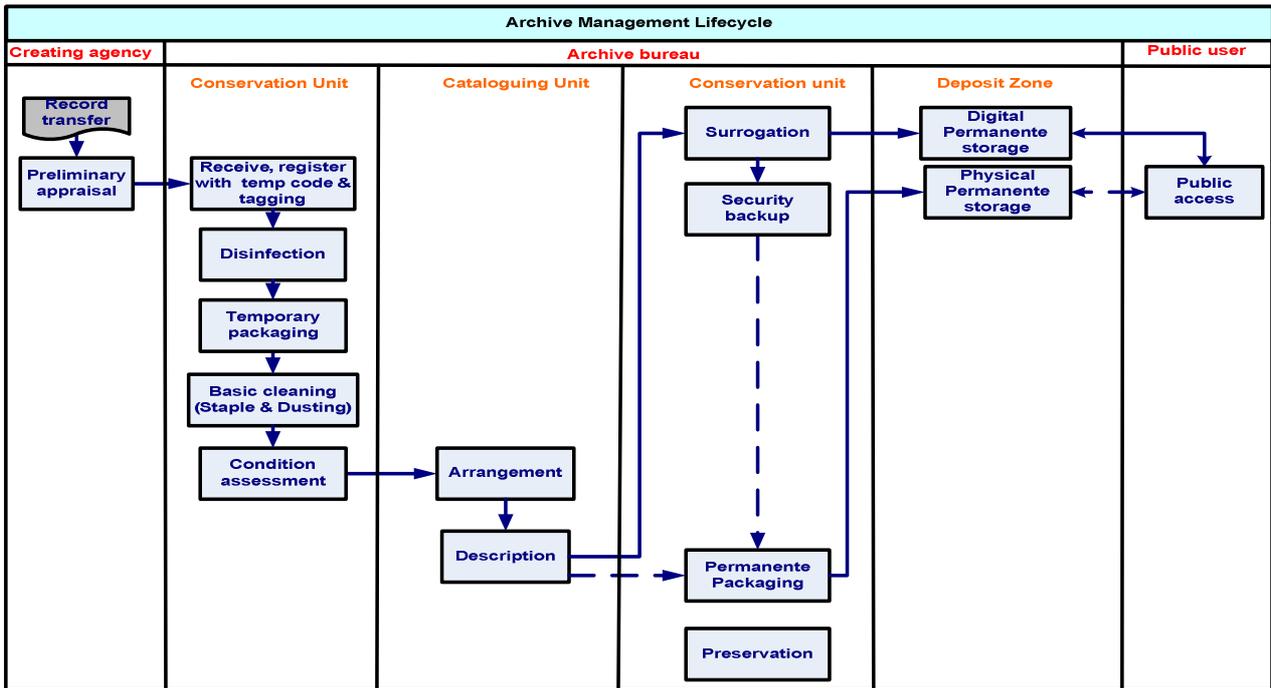


Figure 3. Activity-based conceptual workflow

In addition to main functions outlined in Figure 3, there are a number of smaller tasks within the workflow. These include planning, approving schedule, transferring list received, approving transfer list, locating temporary location, collecting archival documents, tagging temporary code, eliminating archival document, disinfection, basic cleaning, condition assessment, cataloguing, locating permanent location for storage, tagging permanent code, surrogating, image quality check, record linkage, generating microfilm, locating microfilm location, tagging microfilm code, and final record linkage. Thereafter, we will define the specific workflow logic of each process.

Due to the space limitation, we select four archival sub-processes for illustration. The initial conservation process depicted in Figure 4 includes disinfection, basic cleaning and condition assessment tasks. These tasks have to be completed before records are transferred to the description process.

In Figure 4, records designated for archive are processed based on first-in first-out policy. In addition, the condition of the records must be approved by a conservation manager at several stages of the workflow. First, the disinfection is performed for the record. If it satisfies the basic cleaning plan which is made at the beginning of each year, staples and dust from the record are removed. Then, the record is assessed for its condition. Next, the record is placed in an archival folder and tagged with a temporary code. If the record is in poor

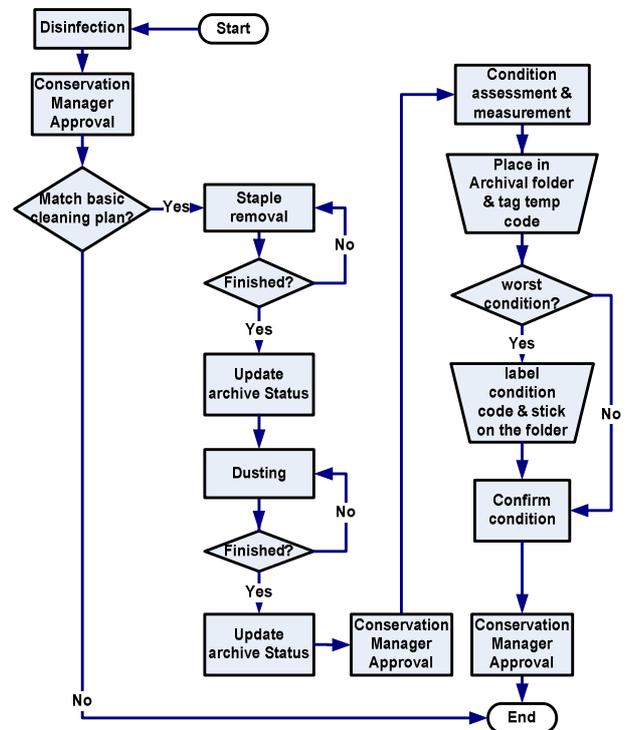


Figure 4. Initial conservation process

condition, it is labeled with a special code for future reference. Finally, an approval is sought from the conservation manager before it is passed on to another process.

After the initial conservation process is completed, the record is passed to the cataloguing process depicted in Figure 5. First, the cataloguing staff will prepare proper description of the record based on ISAD standard. Next, cataloguing manager check the archival description and as well as whether to merge or split the records. If no action is needed or after it has been merged or split, the record(s) is passed to the deposit manager for deciding the location. After that, the record is repackaged for conservation and finally a permanent code for the archive is generated.

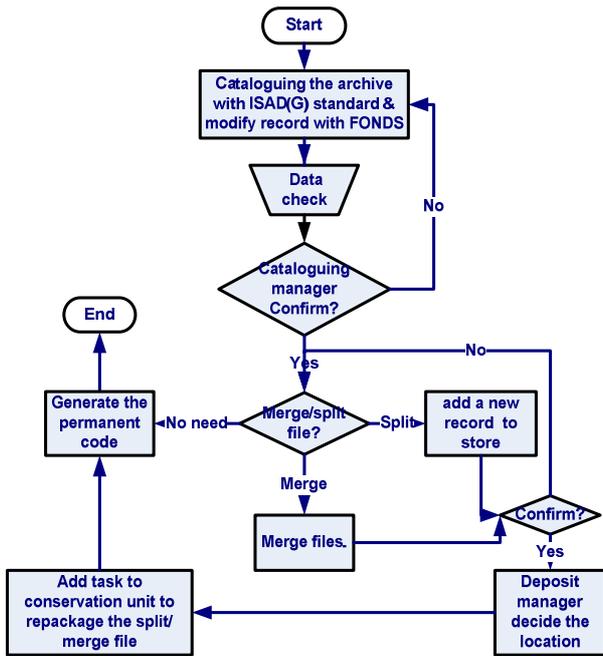


Figure 5. Cataloguing process

The surrogating process is depicted in Figure 6. Surrogating process mainly includes microfilm digitization and paper-based record digitization. These processes transform the microfilms and paper-based records into digital format.

When the surrogating process begins, a conservation manager checks the type of record to be archived. If the record is a microfilm, it will be scanned and a conservation staff will assess the scanning quality. After that, the image will be converted from TIFF format into JPEG and uploaded to the system.

On the other hand, if the material to be archived is a paper-based record, it will be grouped with other records and scanned in batches. Next, a conservation staff will assess the quality of the output. If the quality of the scanned output is acceptable, the image will be written into a microfilm. Two copies of the microfilm are created in the process and tagged with the permanent code. Then a staff member will register and update the microfilm record in archival management

system. After the backup process is completed, the file in TIFF format is deleted from the server.

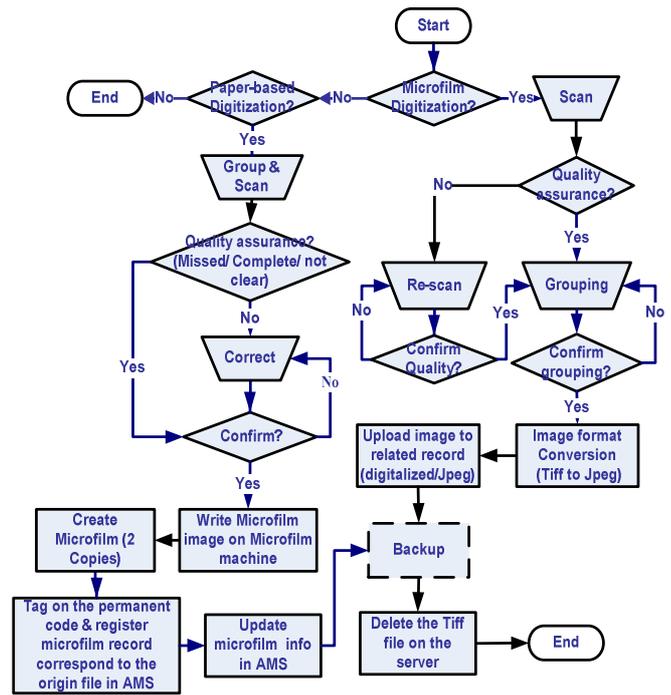


Figure 6. Surrogating process

After the surrogating process is completed, the image file is passed to the backup process depicted in Figure 7. Macau Historical Archives creates 3 backup copies (CD ROM). After the CDs are created and their quality is assessed, the CDs are vacuum-packed and tagged with a permanent code. Finally, a staff member will register and update the data in archival management system and store the backups in a secure place.

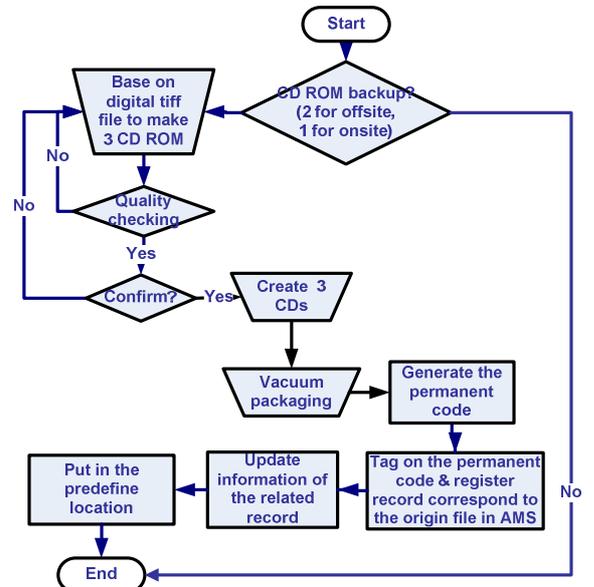


Figure 7. Backup process

Although the workflows described in Figure 3, 4, 5, 6, and 7 are tailored for Macau Historical Archives, these models could be revised and extended for similar situations. For instance, archival management tasks in neighboring regions may involve similar activities. To illustrate such possibility in customization, we compare the proposed archival management workflow with OASIS model in the following section.

IV. COMPARISON OF ARCHIVAL PROCESS WITH OASIS FUNCTIONAL MODEL

The OASIS model is an ISO standard (ISO 14721:2003) [4] and it is recognized as a general guideline for digital preservation and management. This standard provides a form of checklist for identifying limitations as well as for exploring potential extension of existing archival processes. The mapping of the archive functional view (see Figure 2) to OASIS is depicted in Figure 8. The mapping identifies the corresponding entities within the archival management process to the respective OASIS entities. The mapping also highlights the similarities as well as differences between the archival process workflow which is being developed and the OASIS model. From the mapping, we can observe following characteristics:

- Macau Historical Archives negotiates and accepts records from producers or transfer agencies. Transfer agency is a government department which is authorized by law to perform an archival function. A transfer agent needs to register a transfer list of the archival document. Transfer list is a list of files which are being moved from an organization to the Macau Historical Archives. The agent also needs to provide corresponding metadata to Archive bureau for storage. Such activities require prior negotiation between both parties.

- Macau Historical Archives obtains sufficient control of the archival record provided to the level needed to ensure long-term preservation.
- Macau Historical Archives has used ISAD (G) international description standard to prepare description and therefore the community should be able to understand the archival record without the assistance of the experts.
- Macau archival policies and laws are used to ensure that the record is preserved against all reasonable contingencies. In addition, the archival management process includes tasks such as tracking and providing reliable platform and security mechanism for internal and public record transfer.
- Macau Historical Archives is planning to make archival records available to the designated community in the near future.

From the above comparison on OASIS model of mandatory responsibilities, we can see that Macau archival management model covers similar entities and functionalities. In addition, we also find that some of the areas of OASIS model such as preservation planning and monitoring are insufficient. Preservation planning includes generating plans for periodical preservation, migration, and testing. Monitoring includes periodical review of archival records and future migration plans, developing recommendations for archive standards and policies, monitoring changes in the technology development, understanding designated community's service requirements, and establishing a knowledge base for future expansion. Strategies and definition regarding these issues are not clearly stated in OASIS model.

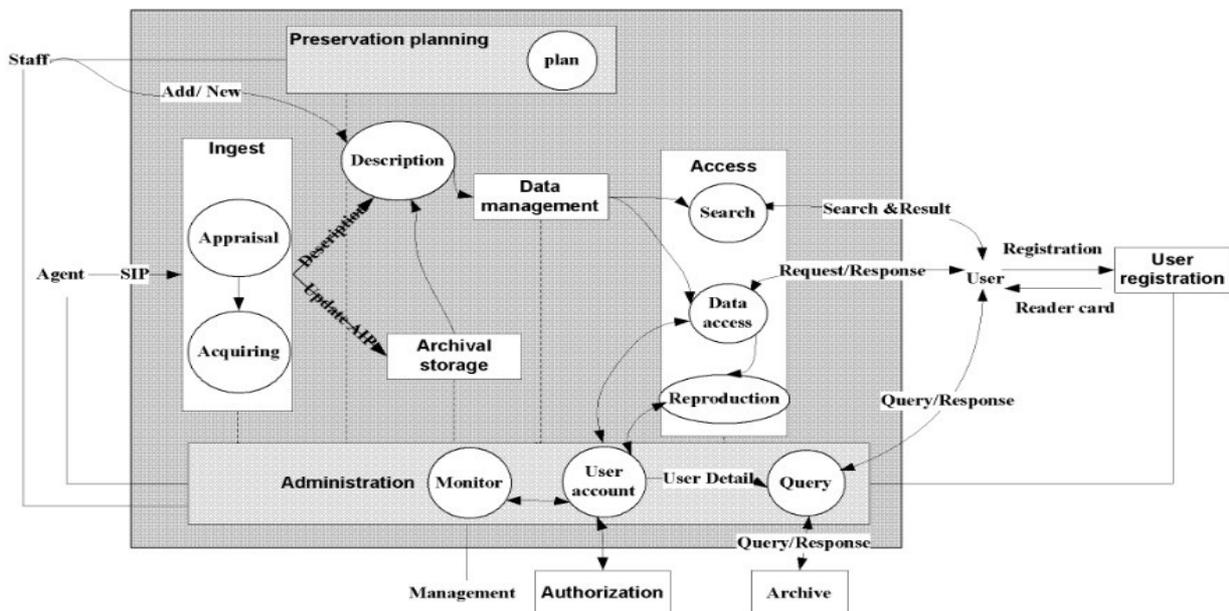


Figure 8. Mapping of archival process into OASIS Functional Model

V. RELATED WORK

Although workflow technology is widely used in managing business processes, less attention has been devoted to integrating workflow modeling with archival management systems. Although several prototype systems have been proposed in recent years, the strength and weakness of such systems has been rarely analyzed in detail. Although OAIS model is accepted as an international reference archive model, it only provides a framework for archival management and long term preservation. OAIS still lacks of the detail clarification on the issues such as monitoring, assignment of roles and resources, and audit trail management.

In [5], Daines et al. describe how process management can be integrated with archival management system. The Integrated Digital Special Collection (IDIN) prototype was developed with the aim of improving management of archival and manuscripts at the Brigham Young University. In their approach, sierra-php [9] is used to develop IDIN system along with workflow management functions. The prototype includes three core IDIN tools (a contact manager, a task manager, and a search application) and one acquisition and accession module.

A conceptual framework for using service-oriented architecture (SOA) for archival management is proposed in [6]. The framework deploys SCA¹-BPEL² service composition approach to build SOA-compliant digital preservation system. In [6], the architecture of the system is specified in SCA model whereas the behavior of each service is defined in BPEL model.

VI. CONCLUSION

The topic on archival workflow has had a relatively little coverage in literature. In this paper, we illustrate the importance of workflow concepts in archival management. Next we propose a conceptual workflow model for archival management process. This model is then extended and further revised for Macau archival management scenario. As a case study, we describe the process models designed for Macau Historical Archives. In addition, to analyze the strengths and the weaknesses of the current system, the conceptual view of archival workflow model from Macau Historical Archives is mapped with OAIS functional model.

As for the future work, Macau Historical Archives is planning to strengthen the archival management system capabilities to overcome some of the limitations described in the previous sections and to enhance monitoring functions for resource allocation. Monitoring task is generally considered as an implicit part of administrative function and often neglected in archival management process design. Monitoring functions can be used to capture audit trail data. Audit trail is the information in records that track a transaction from beginning to end, making it possible to review whether it was done according to relevant policies and standards. The recorded

audit trail data can be used to improve the effectiveness in resource allocation. We are also planning to conduct simulation experiments using Color Petri Nets. The analysis results from the simulation experiment will be compared with the extracted audit trail data to further improve the process model.

Macau Historical Archives is also planning to improve preservation planning process to ensure the longevity of both physical records and the digital information. Preservation is a complex task for protecting materials by minimizing chemical and physical deterioration. Preservation can minimize the loss of information, reduce potential damage caused by the environment, and extend the life of cultural property. Although preservation of physical records is considered to be a common requirement among archival management communities, preservation of digital information is still subject to intense ongoing research. Ensuring the longevity of digital information faces a number of challenging issues such as short media life, obsolete hardware and software, and defunct Web sites [8]. To overcome these issues, Macau Historical Archives is considering extension of current preservation activities while taking into account both physical lifetime and time until obsolete of storage media.

We are also planning to investigate the possibility of providing some of the functionalities via Web services to other organizations. For instance, Web services can make record transfer function (for digital records) accessible to other government departments over standard Internet protocols regardless of platforms and programming languages. One possibility is to wrap around existing modules in archival management workflow to make them network-enabled.

In addition, Macau Historical Archives is planning to extend the current public access services so that archival records are assessable by the users via Internet. Other features such as enhancing security and stricter control of archival data management will be implemented to ensure the integrity and consistency of archival records. As the project is still ongoing, the paper mainly studies the modeling requirements, discusses the difficulties have to overcome, and concludes with our future work.

ACKNOWLEDGMENT

This research is funded by the Research Committee, University of Macau.

REFERENCES

- [1] R. Pearce-Moses, SAA: A Glossary of Archival and Records Terminology, The Society of American Archivists, <http://www.archivists.org/glossary/index.asp>.
- [2] D. Hollingsworth, Workflow Management Coalition, workflow reference model, January 1995.
- [3] ISAD(G): General International Standard Archival Description, Second Edition, INTERNATIONAL COUNCIL ON ARCHIVES, Stockholm, Sweden, 19-22 September 1999, ISBN 0-9696035-5-X.
- [4] Reference Model for an Open Archival Information System (OAIS) CCSDS 650.0-B-, BLUE BOOK, January 2002.
- [5] J. G. Daines and C. L. Nimer, Integrating Process Management with Archival Management Systems: Lessons Learned, Code4lib Journal, Issue 6, March 30, 2009, ISSN 1940-5758.

¹ Service Component Architecture (SCA)

² Business Process Execution Language (BPEL)

- [6] C. Saul and F. Klett, Conceptual Framework for the Use of Service-oriented Architecture Approach in the Digital Preservation, Proceedings of The Fifth International Conference on Preservation of Digital Objects, pp.229-234, iPRES2008, September 29-30, 2008.
- [7] S. S. Chen, Digital Preservation and Workflow Process, in Proceedings of ICADL 2004, Lecture Notes in Computer Science 3334, pp. 61-72, 2004, Springer.
- [8] S. S. Chen, The Paradox of Digital Preservation, Computer, 34(3), pp.24-28, March 2001, IEEE.
- [9] Sierra-php PHP Application Framework,
<http://code.google.com/p/sierra-php>