



# Human Factors in Digitization Selection of Archival Collection

\*Nur Nadiah Athirah Abdul Rajah<sup>1</sup>, Azree Ahmad<sup>2</sup>, AP Azli Bunawan<sup>1</sup>, Nik Nurul Emylana Nik Ramlee(1), Nurul Hanis Kamarudin<sup>1</sup>, Saiful Farik Mat Yatin<sup>1</sup>, Wan Abdul Malek Wan Abdullah<sup>1</sup>

<sup>1</sup>Faculty of Information Management, Universiti Teknologi MARA (UiTM Selangor) Puncak Perdana Campus, Shah Alam, Malaysia.

<sup>2</sup>Faculty of Information Management, Universiti Teknologi MARA (UiTM Kedah), Malaysia.

\*Corresponding author: \*Nur Nadiah Athirah Abdul Rajah: Faculty of Information Management, Universiti Teknologi MARA (UiTM Selangor) Puncak Perdana Campus, Shah Alam, Malaysia.

Email: [nadiahathirah75@gmail.com](mailto:nadiahathirah75@gmail.com)

## Abstract

Some records will be kept for a year as it has enduring value that shows evidence from previous business transaction or evidence from the past event. It will select for permanent preservation called archival collection. In order to preserve the archival collection for future use that might be precious to some organization, digitization is the best solution that can be taken. As to digitize archival collection for public access, some sort of criteria need to be considered. This article discussed on human factors that may affect the selection of archival collection for digitization project.

**Keywords:** Records, Archival collection, Digitization, Human Selection, Public access.

## 1. Introduction

Digitization is a process of converting hard-copy or non-digital records into a digital format in order to provide greater access besides as a part of preservation for the archival collection [1]. In converting records into digital object, commonly it will represented by pixels and processed by technology which then convert into digital representations that enable to be searchable and editable. Archival collection or known as digital surrogates helps communities in defining their collective memory as archival collection provide public access to primary and secondary source materials [2]. In the other hand, archivist becomes the gatekeepers of the collection and material that can be considered as experts on the historical contexts face some problem regarding the digitization of archival collection where archivists need to decide of what to digitize and what to make available for public access [3]. Decision to make archival collection to be digitized must be follow by archivist's skills and rationality where lack of historical expertise may affect digitization selection of archival collection besides effects on culture of the community at large. Moreover, any approach to digitization process should be documented including quality control processes for continuing compliance with the legal environment.

## 2. Problem Statement

In implementing digitization process, there must be a number of risks associated. For example, cost in maintaining the digital images over time, technology, and technical standard used for the images in the future. Based on the previous research, there were

many issues revealed involved in the digitizing archival material such as lack of historical expertise, donor restriction, insufficient funding and more. In some cases, digitizing archival material cannot be done due to lack of time, insufficient general funding, not having necessary technological infrastructure and so on. In the other hand, digitizing archival collection also raise legal issue such as copyright of archival collection through an agreement with the creator of the work.

### *Lack of policy for digitization in organization*

Policy become a best practices as it was written, approved, and researched by responsible administration before been applied within institutions. Research in cultural institution stated that, even though a comprehensive policy in digitization of the archival collection is existed, but it is insufficient to serve as a guideline for cultural institution used [4]. Selection criteria, the principles guidelines, and the most important elements, the standard and protocol which are crucial in digitization project are not included in the policy. In some institution, there is no formal written policy and basically based on habitual practices [2]. But in some cases, institution does have written policy but the institution's administration never reviewed or approved to use it. Regarding the criteria of selecting materials for what to digitize, some information professional simply determine it by themselves without thinking about capabilities of the staff, time-consuming to implement digitization project, technical capabilities, and most important part, about intellectual rights [5].

### *Lack of expertise among archivists*

Stated in a previous research that lack of expertise among archivist in handling digitization process make a reason why human factor is become priority in selecting archival collection for public

access<sup>2</sup>. Besides that, archival collection for public access raised some legal issue such as copyright. Although archivists are an expert to the collection and materials under their purview, they also need to gain some knowledge especially in historical context. This is because where lack of historical expertise may affect digitization selection of archival collection besides effects on culture of the community at large.

#### **Legal issue raised in digitization of archival collection**

Legal issues such as copyrighting arise when archivists have to collect and preserve materials that are copyrighted. Any reproduction through digitization, need to license besides get an authorization from the creator of the work. To make its contents can be search, used, and available online, cultural institutions have to get approval from the owner to make that happen [6]. Furthermore, moral right also become an issue when the digitized of archival collection been modified.

It is because moral right is the subject matter that held by the creator of the original work as it cannot be transferred as long as the length of copyright. Moreover, rights management and protection technologies require licensing to end-user in order to have a little protection for the goods that retain their value over a long period. Watermarks, signature and fingerprints have value to misuse and copyright infringement while encryption can achieve high level of security.

### **3. Research Objectives**

This paper attempts:

- To find out the best practices in digitizing archival material for public access.
- To study the levels of skills needed for archivist in doing digitization.
- To mitigate any legal issues that relates with digitization of archival collection.

### **4. Research Questions**

This research questions will help to address the problems or issues raised in National Archive of Malaysia :

- Is there any best practice in digitizing archival collection for public access?
- What level of skills do archivists need in order to enable him/her in doing digitization?
- What kind of legal issues faced by organization?

### **5. Literature Review**

There are links provides by the culture, between the past, present, and the future [7]. Cultures is also defined as knowledge, belief, art, morals, law, custom and any other capabilities and habits acquired by man as a member of society [8]. It may be forms in everyday lives of the people regardless in the form of published and unpublished texts, images of many types, works of art, artifacts, collectables, historical treasures or similar items kept in the cultural institutions. All of these things enhance understanding between culture, society, and communities besides promote unity. As for human development, cultural and heritage diversity play as an essential aspect that need to be protected. The cultural institution refers to National Archives, National Library, National Art Gallery and Department of Museum that were responsible in holding valuable, unique and rare artifact or resources that represent cultural heritage of member of society. These institutions normally managed by information professionals who are a group of people who involve in handling cultural information within this cultural institution. In this study, human selection becomes an issue in

digitizing archival collection for public access as there is lack in policy in organization.

#### **Policy for Digitization in Organization**

Policy is a principle of action that was adopted or proposed by a government, party, business or individual (<http://en.oxforddictionaries.com/definition/policy>). Policy is important for organization personnel at any level in making decision as it works as guidance to ensure staffs carry out their tasks correctly and responsibly. In order to have a good policy that good for all, there are features to be followed.

For example, matters of principle, action on what to be done and by whom besides authoritative statement for a person with power to do so (<http://sydney.edu.au/legal/policy/what/>). As for Malaysian context, National Library of Malaysia (Department Digitization Policy) in the year 2000 had developed a policy while in the year 2012, standard been produced by Malaysian Standard. Generally, this standard provide the best practice guidelines starting from planning processes, digitization process management, management systems, and records disposition. It discussed more on technical parts such as file formats, resolution, color resolution or bit depth, compression and color management<sup>1</sup>.

#### **Expertise among Archivists**

Digitization project needs information professionals in handling activities related to the project. In this study, archivist was responsible in managing archival collection for digitization project. In Malaysian Standard, recommended that staff skills sets was produced for staff training to support digitization according to their skill area. For example, for management skill area, staffs need to assess the business case for digitization besides negotiating in purchasing, ongoing service, maintenance of equipment and supplies<sup>1</sup>. As for systems analysis skill area, they need to select of scanner hardware, defining storage requirements, testing of configurations and many more. These recommended skills set help archivist in preparing themselves for the right skills that should be equip in order to serve the best for organization through this digitization project. Archivists should know what kind of archival material that brings connection or interaction between the public and collective memory as it is ways where people construct a sense of the past [9]. Other than that, utilizing the availability of technology through the digitization project is dependent on staff expertise [10].

#### **Legal issue raised in digitization of archival collection**

The biggest challenge in digitization of archival collection is the question of copyright. Copyright is where owner or author who has the right to reproduce, publish, perform, communicate and adapt his work besides enable the owner to control the commercial exploitation of his work (<https://www.ipos.gov.sg>). It protects works like novels, paintings, sheet music, plays, and computer programmes. In order to digitize archival collection, it depended on funding and collaboration opportunities [11]. Through the digitization of archival materials, it become advantageous to the archive as there is content that attract their interest and can be access by online. While library is more concern on providing digital substitutes to frail materials besides online dissemination of content minimize risk of third-party misappropriation and misuse of a particular institution's content over the internet. According to Trevar D. Riley-Reid, he said that there is a gap between the rules and laws when they do not match with the needs and demands of users. Besides that, the laws do not fulfill the expectation and requirement of the user when recent updates and amendments created even more question. For example, in sixteenth century Britain, a group of printers in London were allowed to print publications that only approved exclusively by the monarchy. It is because the exclusive rights belonged to the first to print a particular work is to promote learning and prevent monopoly [12]. In this

twentieth century, many digital resources in many institutions where they do not own these assets but license them instead. They still cannot do necessarily on the digital resources even they have permission to digitally copy and electronically distribute the contents.

## 6. Methodology

### Sampling technique

Sampling technique involved population and sampling where both give direction to researcher about the focus group of the research. Population refers to an entire group or elements with common characteristics while sampling is the process whereby a small proportion or subgroup of a population is selected for analysis. Population in this study is information professionals in National Archive of Malaysia that consist of 54 staffs (including states branch). As for sampling, 10 informants will be interviewed as they are under Digital Archive Section. The positions held are Head of Section, Archivists, Archives Assistant, Information Technology Officers, Information Technology Assistant Officer and Office General Assistant.

Table 1: Sampling for this study.

POSITIONS	NO OF PERSONNEL
Head of Section	1
Archivist	2
Archives Assistant	3
Information Technology Officer	1
Information Technology Assistant Officer	2
Office General Assistant	1
<b>TOTAL</b>	<b>10</b>

### Research instrument

Research instrument is a tool for data collection where researcher needs to ensure that the research instrument is valid and reliable. They include interview, observation, reading, and questionnaire [13]. In order to conduct this study, instrument that will be used in getting information from the interview session is by using interview's question. Questions develop according to the topic where the interview can be structured or semi structured. Questions will be divided into 3 parts namely Part (A) Personal Details, Part (B) Digitization Issues - skills, lack in policy or standard, and Part (C) Awareness in Legal Issue such as Copyright. The answers will be analyzed to bring finding towards this study.

### Research design

Research design refers to the strategy and plan to the entire of research process starting from conceptualizing a problem to writing research questions, data collection, data analysis and report writing. After considering the objectives of the study, research questions, and limitations, adopting qualitative method is the best ways in obtained data through interview. Descriptive research is chosen as it describes phenomenon of interest, providing with clear answer of who, what, when, where, why, and how (5W1H) of the research problem and data will be collected through the interview session.

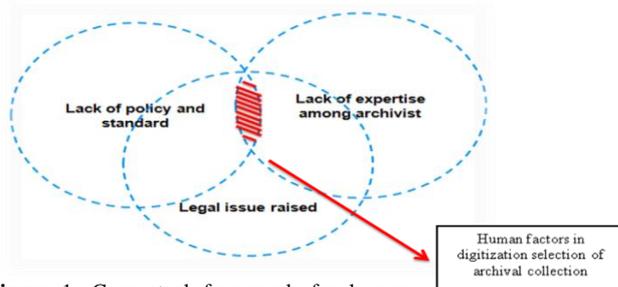


Figure 1: Conceptual framework for human factors in digitization selection of archival collection

### Data collection

From the research instrument, data will be collected through interview and literature review where interview allows researcher to generate ideas and interact with the informants where this session can be structured or semi structured interview. In structured interview, the interviewer follows a set pattern of questions on the interview questionnaire while semi structured interview is a series of open ended questions that allow the researcher to prompt or encourage the informants if the researcher are looking for more information or find what the informants are saying is interesting. This method gives the researcher freedom to discuss some topics in more detail.

### Data analysis technique

Information gathered after the interview session must be analyzed by using a powerful analytical tool, named ATLAS.ti Version 7 that suitable for maximum efficiency and accuracy result in conducting this study and using the thematic analysis for data analysis.

## 7. Conclusion

As a conclusion, this study provides skills and knowledge in digitization of archival collection for public access. Through this study, some aspects should be emphasized, for example policy for digitization in organization. In order to have a better digitization project, policy or standard need to be addressed as it becomes a best practice in this study. Besides that, skills and knowledge among archivist need to be improved from time to time as well as their general knowledge regarding the current issues of digitization. In the other hand, issue related to copyright of archival collection need to give more attention to avoid problem in the future.

## Acknowledgement

This paper was partially funded by:

Conference Support Fund, Institute of Graduate Studies (IPSiS, UiTM).

Academic Development Trust Fund (TAPA), Faculty of Information Management, UiTM.

## References

- [1] MS 2473: 2012 (Malaysian Standard 2473). Information and documentation-Implementation guidelines for digitization of records. <http://www.msonline.gov.my>
- [2] Smith, R. N. (2015). Human selection and digitized archival collections: an exploratory research project about choice of archival materials digitized for online public availability. *Library and Information Science Commons 2015*. <http://dc.umw.edu/etd>
- [3] Conway, P. Modes of seeing: Digitized photographic archives and the experienced user. *American Archivist 2010*, 73(2), 425-62
- [4] Zuraidah Abdul Manaf. Establishing the national digital cultural heritage repository in Malaysia. *Library Review 2008*, 57(7), 537-48
- [5] Levine, M. S. Overview of legal issues for digitization. *Handbook for Digital Projects: A Management Tool for Preservation and Access*; 2000.
- [6] Riley-Reid, T. D. (2015). The hidden cost of digitization –things to consider. *Collection Building 2015*; 34(3), 89-93. <http://dx.doi.org/10.1108/CB-01-2015-0001>
- [7] Liew, C. L. Online cultural heritage exhibitions: a survey of information retrieval features. *Electronic Library and Information Systems 2005*, 39(1), 4-24
- [8] Taylor, E. B. *Primitive Culture*, John Murray: London; 1871.
- [9] Confino, A. Collective memory and cultural history: problems of method. *American Historical Review 1997*; 102(5), 1386-1403
- [10] Allen, N. Collaboration through the Colorado digitization project. [http://firstmonday.org/issues5\\_6/allen/index.html](http://firstmonday.org/issues5_6/allen/index.html)

- [11] Rosati, E. Copyright issues facing early stages of digitization projects .Mobile collection project, 1-16. <http://www.digitalhumanities.cam.ac.uk>
- [12] Warwick, S. *Libraries, Museums, and Archives: Legal Issues and Ethical Challenges in the New Information Era*. Lanham: Scarecrow Press; 2002.
- [13] Annum, G. Research instruments for data collection. [http://campus.educadium.com/newmediart/file.php/137/Thesis\\_Repository/recds/assets/TWs/UgradResearch/ResMethgen/files/notes/](http://campus.educadium.com/newmediart/file.php/137/Thesis_Repository/recds/assets/TWs/UgradResearch/ResMethgen/files/notes/)