

## Digitization of Records and Archives: Quality Control Process

Nurisya Farha Roslan, Saiful Farik Mat Yatin , Nur Aqila  
Azman, Nor Shahira Che Soh

Faculty of Information Management, Puncak Perdana Campus, Universiti Teknologi MARA,  
UiTM Selangor, Malaysia

To Link this Article: <http://dx.doi.org/10.6007/IJARBSS/v8-i9/4586>

DOI:10.6007/IJARBSS/v8-i9/4586

Published Date: 13 October 2018

### Abstract

This paper examines the quality control used in the organization in order to identify, control and maintain the value of the digitized materials. In a growing development, it is significant for the organization to implement the quality control in the business environment and archival institution as well as in the digitization process as quality control ensure the digitized material meet the standard requirement and it is the key element in producing a quality material which consist of accurate and valuable information and record. This paper aims to highlight the purpose of the quality control being implemented in organization, the procedure and requirement needed to establish a quality material with the appropriate equipment.in addition this paper contains the purpose of the quality control in digitization which helps the organization manage their materials appropriately and offers several benefits through its implementation. Furthermore, there are some issues and challenges faced by the organization in sustain the quality control that could lead to the ineffectiveness of the digitization process, it is important for any organization to aware about the issues which it can help in preventing the problems arise within the digitization process, however this paper also highlighted some ways that can be used by the organization in handling and resolve the problems. This paper also compared the implementation of quality control by using machine and manpower which is the most important sources in quality control process. The differences are being discussed is based on the research being done from several articles.

**Keywords:** Quality Control, Visual Checks, Digitization, Records, Archives

### Introduction

**Quality control** is that part of *quality management* focused on fulfilling *quality requirements*. Quality control is a process through which a business seeks to ensure that product quality is maintained or improved and manufacturing errors are reduced or eliminated. Quality control requires the business to create an environment in which both management and employees strive for perfection. This is done by training personnel, creating benchmarks for product quality, and testing products to check for statistically significant variations (Russel, 1920).

The level of quality control (QC) for digitization will vary based on the nature of the project and material being scanned. Typically, images will be processed at the proper resolution, cropped, and correctly oriented during the scanning process. These are all primary QC indicators. During capture, the scanning technician also inspects all images for other primary QC indicators including the correct file name, uniform crop, and color profile. These steps allow the technician to catch any mistakes and save scanning time by resolving these issues early in the process (UCLA Library, 2015). In other word, quality control is an important part of any digitization project. Quality control includes procedures and techniques to verify the quality, accuracy, and consistency of digital files. Quality control should be conducted throughout all phases of the digital conversion process to ensure that the material is only needs to be digitized once, then used and shared many times. Not only that, the quality control also plays a critical role in high volume document digitization projects by making sure that the specified quality standard is reached under cost and time constraints. Quality control is not simply a check on the output of digitization, but a process that should be built into and maintained in the ongoing operation of the digitization work.

### **Issue Of Quality Control**

#### ***Quality of digitized image***

Quality control depends on the quality of the digitized image and is dependent on human analysis, the digitized image of the original paper is a process that requires expertise in producing a material that is more suitable for storage over a long period of time and so that the material can be preserved for future reference purposes. Good material images can retain important content for certain parties. The issues faced by the institution are that where they face the archive materials that suffer from severe image quality issues for instance blurriness, duplicate pages, image order and missing pages, readability issues and lines through scans (Queensland State Archives, 2014). These problem causes the contents of records is difficult to be referred by others in the future and cause failure in digitization process of archival materials. Hence, this issue requires a deep attention from the archives in ensuring the archive materials are in good condition and capable to retained for a long period. The equipment used in the digitization must be service regularly so that the digitized materials can be conducted without any problem and the quality of the materials can be retained for a long period.

#### ***The amount to check***

The essential aspect in quality control is determining the archival materials that through the digitization process need to be checked beforehand to ensure the materials in the permitted conditions for digitization, where the tested materials meets the minimum requirement in quality level in which it can be a time consuming and intensive resource for checking the high amount of materials which consist various types of archival collection such as maps, files and photograph. It can a difficult process as it require the staff to checked all digital images in order to ensure it meets the minimum requirement because there are a bundle of materials in archives. Besides that, every digital materials requires them to be tested and checked to ensure the digitize materials meet the quality guidelines, and every materials checked require appropriate handling and proper care in order to preserve the content of the materials. As the amount of materials increases every day, this issue needed to be address by the institution so that the materials being checked and digitized will not have a problem.

### ***Re-Imaging***

The digitized materials are the approaches that are used to eliminate the space and as one option in ensuring the materials can be preserved over a long period. However, if the digitalized material does not meet the prescribed quality standards document, the public authority needs to recheck the materials to ensure there is no material that does not meet the standards as it may damage the contents of the record itself. If the material obtained does not meet the quality standards, a procedure should be provided where there are instructions on the need to re-examine and recheck the remaining output. Furthermore, the digitalization process needs to be re-implemented to meet the standard aspects which can be a fatigued process and repeated work as they need to ensure that digitized materials that being re-imaged and metadata re-captured are in line with the procedures when the quality standards are not meet.

### ***Staff issue***

There are some of the quality control which relies on the human analysis and judgement, it is usually subjective where the percentage of visual inspection results is different from the point of view of each person, so it is difficult to determine the result of the inspections hence it contribute to the complexity of decision made on the digitized materials. The staff issues that arise in digitization process is there are unskilled and incomprehensible staff about their role in checking the output of digitized materials, and there are responsible staff for the quality control process who are not aware about their work specifications and lack of adequately training and skills. The institution need to ensure that the responsible staffs are adequately trained in order to carry out their duties properly. Training should be provided to the staff responsible for visual inspection, it is important to improve the qualitative communication of effective information between the staff. Besides that, an increase in the inspection of additional quality from supervisory staff should be established to ensure a higher level of consistency.

### **Reasons Of Quality Control**

#### ***Important part of large-scale digital production and preservation***

The quality control is important in order to digitize the books and all the primary sources (Conway, 2011). The reason why quality control needs to be done on books and primary sources is because it will determine the fitness for use, provide incentives for stakeholders who may contribute to digital repository or even donate materials to an archive. The quality control in large scale digitization may require some processes to detect the error. The processes are either using visual checks or automated checks. All the error detected between the two techniques will be calculated and recorded. Without doing the quality control, the errors are still there and the content of the materials will be considered inaccurate to be use by user. Quality control also will help in preservation since the error are all detected during the first process and they do not have to be inspect again to check for the error.

#### ***Produce quality in result***

The materials produce without considering the quality control will affect the user's satisfaction. When materials given to user with a lot of error, it shows that the material has no quality and the information is consider incomplete. Most of the time, quality control is related with scanning process. It is very important to control the results of the scanning business during the digitization process and not only at the end. If the archivist or the

responsible person is not scanning exactly according to the agreements in the quotation and the test, errors in the execution of the technical specifications or in the giving of filenames of the scans or the OCR can still be adjusted in an early stage. This quality control is essential and is in first instance done by the quality manager (Chapman and Leonard, 2013). The people who do the quality control must be particular about the process since it will link to the good result of the material.

#### ***Improve staff techniques, method and skills***

By supplying good equipment to conduct the quality control, it will improve and help in techniques and method for the person who responsible to conduct the quality control process. Since the quality control requires a person to be specific and details, it will be difficult for them to do the quality control by themselves. By organizing training for the staff on how to handle the operational and the equipment, it will improve their skills and increase their quality control knowledge. Then, the good equipment such as automated check that will be able to handle by the staff and with a lot of practice it will improve their skills and techniques. By the skills they have, they can use any kind of technique to produce a good quality material (Chapman and Leonard, 2013)

#### ***Improve preservation strategies***

When doing quality control, the preservation strategies can be improved. This is because the quality control requires proper and double checking about the content of the materials right after the scanning (Chapman and Leonard, 2013). There are many errors detected and from that, the quality control will ensure the errors are detected and should be fix. Through that, the preservation of the materials can be less since the errors has been detected earlier and the staff do not have to fix and preserve the materials regularly since the materials they serve to user is already in good condition and quality.

#### **Requirement Of Quality Control**

Based on the policy stated which encompass the several requirements needed in quality control in order to ensure the flexibility and efficiencies of the digitization process in archival institution. Queensland State Archives (2014) have established an outline of the conditions and requirement of quality control in Digitisation Disposal Policy. This is to ensure that all the archival material that undergoes a digitization process in which they intended to destruct the original copy of the document, must comprehend with the requirements establish and stated in the Principle 3 of the policy by the Queensland State Archives. Within the Principle 3 of Digitisation Disposal Policy, it is necessary for the public entities to determine and organize the appropriate system used to enhance the efficiency and effectiveness of the digitization process. Besides that, it is also important for the public entities to carry out the processes of data capturing as well as managing and controlling the records that have been digitized.

#### ***Timing of Equipment Tests***

The archival institution plays a vital role as there are responsible for ensuring that all equipment used for the purpose of digitization needs to be tested in advance for the purpose of all the digitized processes to be carried out properly and to avoid any problems arising during the process which may impair the archival materials or failure of the digitization.

### *Equipment Calibration*

One of the scales or accuracy made or fixed by one device and measurements made in the same way as a second piece of equipment. Basically, it is the choice made of appropriate equipment that is suitable for digitizing process that depends on the types of the materials that need to be digitized. Additionally, the equipment required in the digitization process should be in accordance with the standards set out in the policy to ensure that the equipment used is appropriate and meet the prescribed conditions for the digitization of the archival material. Otherwise, the calibration checks must be performed at the specified time range based on the manufacturer's specifications.

### *Checking and Approving Output*

The staff responsible for quality control must be adequately trained to adapt to the environment and the situation where it is necessary to overcome all the specification of the work and to deal with all the issues that may arise due to unavoidable events and some factors such as equipment failures used for digitization process. In addition, the staffs need to ensure that subjective visual inspection is consistent and meets the standards set to preserve the material from any damage.

### *Checking the Output*

Requirement of procedure that require the output to be checked such as digital reproductions ratio in which it should be subject to visual inspection and the duration of the original records should be maintained in the archival institution after the implementation of the digitization process to ensure that the quality inspection process can be implemented in the appropriate manner. The aspects that need to be highlighted and addressed for digitized image is the smallest detail legibly captured for instance the decimal point, smallest type size of text, including the clarity of punctuation mark. Besides that, in quality control the details of the record need to be completed so that the retrieval process is easier, and staff must ensure that the content of the digitized materials are same as the original copy which means there are no missing information or any cropped image. Additionally, staff must aware about the accuracy of the colors and information compare with the original copy, to avoid any issues of misinformation of the digitized record.

### *Re-imaging*

This procedure only need to be done when there are some digitized materials that identify to have an incomplete content of the materials or there are materials that did not meet the quality standard which developed by the policy. It is important for any digitized materials in a good condition in order to maintain the integrity and accuracy of the records suitable. Hence, the organization or record office need to ensure that every digitized material preserved in their care must be protected from any risks.

Overall, the requirement developed by the Queensland State Archives must be followed by every record office and archives in order to preserve the integrity and values of the materials so that the content can be retained over a long period of time and can be referred to when it is needed. Robust quality control procedures are important as it is capable of producing a useful and valuable materials that can be used in the future, and if there are any challenge, the public authorities need to show in court that reliable and effective systems and processes

are used in digitizing the materials and working properly according to the guidelines and requirements establishes.

### **Procedure Of Quality Control**

#### ***Stage 1: Automated checks***

The first stage of this process is an automated check of all files digitized in a specified time frame. It means that how frequently an automated check is run varies in digital projects. In conducting outsourcing digitization, the automated checks are run whenever a batch of files from the vendor are received.

Process for digitizing an archival collection:

The archival collection must be neatly organized into a discreet box that would take about a day or two each for scanning. In the scanning process, an automated check will run on all files scanned when a box is complete. If there are more continuous of the collections without natural breaking points, the automated checks will be run weekly or even daily.

Process for other collections:

The process started with checking all images located in a temporary storage directory, then moved to a permanent location after they pass the review. Compared to archival collection, only certain images of other collections within a directory are checked at a given time.

#### ***Stage 2: Visual checks***

While the automated checks described above are extremely effective, they could not obtain all possible quality problems in a digital imaging project. A staff member who supervises the digitization team performs visual review of a sample of images to supplement the automated check. The images are viewed on a monitor in a standard viewing environment at a magnification such that one screen pixel displays one image pixel. The number of images checked varies among digitization projects, but for projects digitized in-house we generally check around 10 percent of the total number of images scanned. Digitizing staff track any anomalies in original items that might contribute to image quality problems. These items are visually inspected in addition to the random sampling of the entire set of images.

#### ***Stage 3: Moving to a centralized system***

This stage started with developing an image validation and processing system with a web-based interface which departments use for delivery of their digitized images to the Digital Library Program. Riley and Whitsel (2005) explains that the image processing system is built in Perl and uses XML configuration files that allow the same application to be used for a number of different image collections. When the image processing system has been extremely successful in codifying, it enforces quality standards for imaging projects. This is not a complete quality control solution if it is not directly controlled by the Digital Library Program. Therefore, it is encouraged for the departments that using the system to enhance its use with a visual review of some percentage of their digitized images, and search for ways to facilitate this within the image processing system.

### **Equipment For Quality Control**

Digitization relies on regularly maintained and correctly calibrated hardware and software to produce high quality images in order to meet the quality baselines. A scanner is the most vital

hardware for digitization, hence it is important to select the best scanner that are suitable such as in term of the size of documents to be scanned, type of feed, functionalities (e.g. page counter), technical support available, image mode, optical resolution, etc. The scanner requires scanning software, therefore it is a must to choose software that can meet the project's needs such as user-friendliness, formats available, functionalities and settings. The scanning hardware system should be tested by the use of scanner test targets or charts in order to establish acceptable levels of quality for digital image capture. Queensland State Archives (2014) discusses that these contain a wide range of material which provides the ability to judge output in carefully measured increments for such aspects as resolution, text, fonts, line widths, color, tonal range, handwriting and halftone.

Scanner are the most significant hardware components since this component is a medium for digitizing collections. Scanners are divided into few types which includes flatbed scanner, overhead scanner, sheet-fed scanner, film scanner and etc. There are various types of scanners because they are used for different purposes. For example, some scanner is used for large scale materials, some for maps, some for photographs, slides, negatives and many more. A scanner for digitization process need to be considered in terms of its interface and specifications. Since digitization requires long hours of using the scanner, it must be fast and reliable. Hence, it is not recommended to use USB and parallel port scanners, SCSI port scanner and IEEE 1394 FireWire. For the scanner specifications, consideration must be done on the scan area, the optical resolution (dots per inch) and its color depth (bits).

Besides that, other hardware components are needed for scanning such as computers with sufficient capacity, high-resolution screens and adequate video card. These criteria are needed and important since a good quality computer as well as the scanner, will help to ensure the quality of the digitized images that will be produced. This is because when digitizing materials, those hardware components such as computers and scanners will determine the quality of the digitized materials where the staff member will check whether all information are completely digitized. The checking process will be conducted properly and precisely in order to avoid any missing information on the digitized material. Therefore, a large size of monitor screens and an advanced scanner are important for producing a high quality digital projects which can save time without having to do repeated digitizing process.



Figure 1: Scanner for oversized materials



Figure 2: Book scanner



Figure 3: Digitization workstation

## Comparison Between Machine And Manpower In Quality Control

### *Time consuming*

**Machine** are more practical and fast since they can detect the error between a minute. The data or material just needed to be entering to the machine then the machine will do the rest. What makes the machine become time consuming is due to the machine can access more data or material at one time to detect the error. So, the machine is considered practical and efficient since it fast and easy to be use and handle.

Compared to **manpower**, it requires human effort, time and knowledge to conduct the control quality. The error found as not as fast as machine since they need to do it one-by-one. It can be fast if there are many manpower who involve in the quality control process. Not only that, manpower also need to take a break and sometimes some of them are on leave. That is why the quality control process using manpower will take a little bit time compared to the machine.

### *Quality of work*

**Machine** can be helpful and sometimes can be harmful. Since the machine are too advance and can process the material in large-scale, but the result can be poor. The errors sometimes are not fully detected. Therefore, some of the machine can only detect for the basics or simple element only such as checksums, file type, file naming, or resolution. They are not details enough to detect the error (Belfiore, 2012).

Compared to **manpower**, they are more details since they are particular about the content of the materials also they do it one-by-one. The technique or process use by manpower is known as visual check. As for example, the manpower might compare digital images to the analog original to check whether the file name assigned to the image is correct, if the colour matches the original, if the digital image is in the correct orientation, that no physical matter is included in the image, and that the pages are in the correct order (Riley and Whitsel, 2005).

### **Costing**

**Machine** requires a good technology infrastructure in order to make it work. For example, it requires a machine, software and electricity. For that, it will need to be purchase and the cost could be expensive. Not only that, the organization also must consider the maintenance cost for the machine if suddenly it breaks down. The bills of electricity also need to be consider since electricity is the main source to ensure the machine are working well.

Compared to **manpower**, it will be less cost since they only use human effort, time and energy so the organization only needs to pay for their salary. There are no maintenance required and the quality control can be done smoothly even it takes some time to detect the error but the most important is the quality of the work is good and can lead to satisfaction.

### **Creativity**

**Machine** is lack of creativity since they only follow the rules and standard of what has been set to them. By that, the result will be consistent and not so outstanding. If they are being set only for checking the maps, then it will only checking the maps with the only way the machine can do.

While for **manpower**, they can be creative in doing the quality control. They can use their skills to ensure they can easily detect the errors and how to make it fast. In addition, manpower also can avoid from being bored so they can change their job task such as not doing the same thing every day but still follow the rules and standard. For example, today the staff is doing the quality control on maps and tomorrow is on papers. By that, they can also learn something new instead of just doing the same thing every day.

### **Conclusion**

To be concluded, quality control (QC) is a process that verifies the quality, accuracy, and consistency of digital images to ensure the products match the specifications outlined in the project plan. The quality control process is particularly important for scanning projects where paper records will be destroyed after scanning an important component of electronic document management in digitization must allow user to create electronic documents that can replace your paper documents. In order to do so, user must implement rules that can allow them to create quality electronic documents while making sure their transfer process is properly documented. This is because even if the user creates quality documents, if the transfer process is not documented, they will not be able to destroy their paper documents. Conversely, if they document their transfer process but they create poor-quality electronic documents, they will not be able to destroy their paper documents. Additionally, quality control is the processes which enable the digitized materials being done properly using the specifications that have been established and by using the appropriate equipment for each type to prevent any mistake that can harm and damage the materials. Furthermore, the process of quality control must be handle by experts so that the quality of materials is sustainable and readable, therefore quality control staff must have expertise in this field or

need to administer for training session to improve their skills as they are responsible in protecting the value of the materials and ensure the quality of digitized materials. Quality control is the most important process that need to be done in order to detect error in materials so the materials being serve to the user is quality and good in physically and content.

### References

- Belfiore, E. (2012). "Defensive instrumentalism" and the legacy of new labour's cultural policies. *Cultural Trends*, 21(2), 103–111. doi:10.1080/09548963.2012.674750
- Chapman, J. & Leonard, S. (2013). Cost and benefit of quality control visual checks in large-scale digitization of archival manuscripts. *Library Hi Tech*, 31(3), 405-418. Retrieved from <http://www.emeraldinsight.com.ezaccess.library.uitm.edu.my/doi/pdfplus/10.1108/LHT-01-2013-0002>
- Digitization quality control workflow.* (n.d). Retrieved from <https://www.investopedia.com/terms/q/quality-control.asp>
- Guide to quality control and quality checklist.* (2016). Retrieved from <http://www.sustainableheritagenetwork.org/system/files/atoms/file/GuidetoQualityControlandQualityChecklists.pdf>
- Riley, J., & Whitsel, K. (2005). Practical quality control procedures for digital imaging projects. *OCLC Systems & Services: International digital library perspectives*, 21(1), 40-48. DOI: 10.1108/10650750510578145