

A STUDY OF THE DIGITIZATION PROCESS TO PRESERVE THE CULTURE AND HERITAGE OF A CIVILIZATION USING NATURAL LANGUAGE PROCESSING AND IT'S IMPACT ON THE SOCIAL, ECONOMIC AND SCIENTIFIC ASPECTS

¹MUKESH MADANAN, ²NORLAILA HUSSAIN, ³ADEEL AHMAD KHALIQ

¹Lecturer, Dhofar University, Department of Computer Science, Salalah, Oman

²Assistant Professor, Dhofar University, Department of Computer Science, Salalah, Oman

³Lecturer, Jazan University, Department of Computer and Network Engineering, Saudi Arabia

E-mail: ¹mukesh@du.edu.om, ²norlaila@du.edu.om, ³akahmad@jazanu.edu.sa

ABSTRACT

Venturing into the arena of digitization, technology has been innovative in establishing digitized manuscripts and e-books. Digitizing and archiving manuscripts and documents blanketing texts and images using digital cameras and scanners are being viewed as a stepping stone to the use of technology. To be a part of Generation Next, focus has to be on up to the minute technology of digitization. This entangles to preserve not only manuscripts but also the culture and heritage depicting it. Preserving the artefacts, manuscripts and paintings through primitive techniques physically in historical societies and libraries has limitations. Highlighted restraints are access to general public and lack of facilities to safeguard valuables and keep in place their vulnerability. Manuscripts and documents when archived through chemical methods do not most of the time depict actual bearings. Digitization of rich culture and heritage documents and manuscripts would recalibrate these constraints and iron out most of these limitations. The paper focuses on the digitization process to preserve the artifacts using natural language processing techniques and indexing and also highlights the future of digitization to enhance the cultural heritage of a nation.

Keywords: *Digitization, Cultural Heritage, Metadata, Indexing, Database Management*

1. INTRODUCTION

Information technologies today are rapidly evolving and have a compelling impact on the day to day routine. Globalization has brought the world within a box of communication where everyone leads a uniform life. This has recently influenced the eradication of culture and heritage from the minds of youth causing loss of their unique identities. The young generation also seems to be very gradually erasing their traditions and incorporating minimum values from their ancestors.

Preservation of culture and heritage has been recognized as the leader in advocating conservation. The legacies and evasive values of physical artifacts have to be passed on to the Generation Next. Manuscripts, documents, archeological sites and monuments at the moment symbolize the cultural heritage of a civilization. Magazines, paintings and artworks also heap up the

civilization history. As a matter, even today's living testimony has to be recorded and passed on to the next progeny.

To inhibit and conserve the culture and heritage, technological aspects such as digitization can lend in hands. Digitization of the culture and heritage will involve documenting, organizing, preserving, exploring and propagating knowledge and information. The cultural diversity of various civilizations has to be digitally catalogued appropriately for preservation. This diversity varies across the demography along with linguistic affiliation and ethnic relations. The culture and heritage of a civilization has been recorded by the ancestors in the form of books, documents, paintings and manuscripts. The architecture and urban divergence influences the culture and heritage of a civilization. Digitization can be seen as transformation of analogue materials into a digital format for use in computer applications.

Data capture or acquisition is done normally using a digital camera or scanners.

With the advent of technology, the preservation techniques also have to be leading-edge. Digital technology has been the magic wand in the techniques recently used to do preservation and conservation. From mainframe computers of the 1940s, which were costly, labour intensive and maintained centrally by large organizations, via the introduction of micro and mini computers in the 1970s and 1980s, to the development of improved, inexpensive processors and memory which influenced the personal computing revolution in 1990s, these changes have dramatically affected the way we live and work [5]. But with the importance of culture and heritage preservation grooming, we have to look forward to the new swing in digitization strategies and techniques. The digitization of the text and literature has to be seen with remarkable vitality. Alongside the preservation of literature and manuscripts, prior importance has to be given to safeguard the manuscripts and art forms. As Cassar M. states “Efforts to preserve and enhance cultural assets not only reduce the risk of permanent loss of diversity among the immovable and moveable heritage, but they can provide important benefits and opportunities for greater social cohesion, reinforcing a common identity and strengthening socio-economic aspirations” [15].

The conventional environment and craftsman’s experience of a civilization is transferred by the cultural heritage. The progress of a civilization is also depicted by the culture. The preservation of the historical monuments and architectures has not been efficiently executed due to the time constraints and technology hindrances. To an extent, natural disasters and wear and tear have also caused the destruction of such valuable piece of history. To render the culture and heritage of the a civilization, the digitization techniques have to be integrated to the preservation methods which in turn would lead to building a repository of preserved civilization. The ancestors have recorded most of the details of history in terms of writings which are preserved using traditional methods such as chemical treatment and preservation techniques in cultural heritage institutions and museums. But such preservation has limitations where access to public and researchers are limited.

The digitization of cultural heritage which has been recorded in the books and manuscripts is a

multidimensional process. Digitization will involve selection, assessment, ranking, project management and tracking during the first phase. The secondary phase will follow with preparation of originals for digitization, metadata collection and creation and digitising. Succeeding the digitising activity, the process of quality management, data acquisition and management, submission of scanned images to systems for storage, and assessment and evaluation of the digitization effort will follow. The methodology or phases will involve planning the entire process followed by digitising the manuscripts and documents and leads to the process of maintenance and error prevention. The phases throughout the digitization project are overlapping to ensure quality and standard. In the research paper by George P. and team it has been mentioned that “The whole process involves digitization, data processing and storage, archival and management, representation and reproduction”[10].The vitality that distinguishes the digital preservation context from the analog one is the sheer scale of it [17].

The selection of the artifacts and manuscripts for digitization will be based on certain criteria such as demand, copy right, funding etc. The digitization will also focus on a framework which spotlights on data capture and creation, access and delivery along with managing the collection. The digitization process can be seen as primary step in preservation which will also involve the data retrieval methods. Sporleder explains” The main reason for most of the large-scale digitization projects, however, is the fact that governments worldwide have come to view cultural heritage as a valuable asset, both ideationally and economically” [2].

1.1 Culture and Cultural Heritage

Culture may be encompassed as the human characteristics and traits depicted by their language, religious belief, habits, entertainment such as music and arts in a group. Historians believe culture virtually represents every aspects of life. The contradiction between culture and cultural heritage is that the cultural heritage is the legacy which the ancestors leave behind in terms of artifacts and monuments for the future generations to learn and follow. Cultural heritage is an important factor and very often it is beneficial in robust varieties. The cultural heritage comprising of buildings, decorative elements, objects and collections, equipments and settlements and agriculture and human traces from history on the land and in water

are recorded in books and ancient manuscripts. Hadžić, O states that “culture can be justified for tourism, for industry and for employment, but it must also be seen as an essential element in preserving and enhancing national pride and spirit” [4]. All the sites and locations represent the vast history of a civilization which has to be preserved and conserved for the future generations. Heritage preservation and interpretation are indispensable elements in the conservation natural and cultural resources [4]. Though modernization and technology has caught hold, the culture and heritage and living tradition have to be upheld. The landscape and architecture is a niche of old world wonders which are very unique and preservation of such heritage should be of prior importance.

Literature and architecture are one of the prehistoric forms of arts and crafts. Though urbanization has taken effect, the rural inhabits are of vital importance. The complexity of literature and architecture has certainly increased with time and become more contemporary. The manuscripts and architecture of the past reveal a lot about the human history. Architectural heritage is the most important facet of cultural heritage. One of the most important parameters in describing cultural heritage is historical space and historical time [16].

The culture of a civilization is very much dependent of the geography and location. The cultural divergence is clearly pictured in most of the writings and paintings. The historic architectural sites and buildings reflect a simpler cultural expression but the private residences reveal the gendered architectural space. Passing a civilization’s exceedingly rich culture and heritage and lifestyle to the generation next is done primarily through manuscripts and literature. The art, culture, folklore, artistry and architecture of a civilization are tremendously viewed as exotic by the tourists. Preservation of the cultural heritage has been of prior importance with focus on retrieval of cultural information which can be achieved through the digitization of the books and documents. A well preserved heritage enables communities to learn about their cultural history truly and chronologically [25].

1.2 Digitization

Due to the introduction of various technical strategies, the preservation of the existing collections and creation of new assortment has been the focus of cultural institutions and governments

all over the world. Latest trends have opened a new doorway to preserve the cultural heritage of nations using the improved techniques and methods of digitization through which digital archives are created and managed. Digitization can be used as a promotional tool as it reduces the distance between the public and the objects which are of historical importance. Moreover the digitised content could be displayed and exhibited through various forms of technology. Digitization breathes new life into materials from the past, and turns it into a formidable asset for the individual user and an important building block of the digital economy [22]. Digitization allows cultural heritage institutes to provide continuous, simultaneous and world-wide access to their collections, including those objects which are usually not on public display due to lack of exhibition space or due to their fragility [2]. The digitization process is non-linear and decisions made about technical standards will affect all future uses of digitised images, whether they are suitable as archival images available for a digital preservation strategy [13].

Digitization refers to art of converting physical analogue originals to digital objects. The digitization strategy has proved to be very much efficient in preservation of the originals in digital format. The digital format organizes the objects into bits of data which are easily addressable and accessible. The process is very loosely explained or defined but it encapsulates a lot of management and technical processes and activities. The selection of artifacts and materials for digitization, processing of the files created, storing and preserving and distribution and retrieval are some of the activities involved in digitization process.

The process is very tedious and requires a well described collection of parameters and instructions on establishing and implementing. Depending on the nature of the collection of digitised objects to be created, the sophisticated technical processes involved varies. The digitization will comprise of document handling, creating and managing metadata, scanning, image processing and indexing. This will also render the disaster recovery features. In cultural heritage documentation, choosing the appropriate technology (sensor, hardware, software), the appropriate procedures, designing the workflow and assuring that the final output is in accordance with the set of technical specifications is always a challenging matter [3]. Post digitization activities include submission to delivery and repository

systems, data collection and management, making digital copies and associated metadata available, assessment and evaluation [11].

1.3 Natural Language Processing

The digitization is only a step towards preserving the cultural heritage. After scanning, the data has to be saved and given unique title or reference numbers which can be termed as metadata. The metadata created gives only information to humans and to facilitate their searching. For machines to understand, the human metadata created will pose a problem because of its language ambiguity and specification. Moreover, it may sometimes lead to substandard results during data retrieval. In certain situations, data retrieval may lead to incorrect results based on the search query. This may be due to the lexical ambiguity in the entered query. In order to access the information saved and stored in the database, the data about the digital content has to be enriched and polished. A way to reduce the lexical ambiguity is to set up a set of metadata and naming conventions. This opens up research topics in natural language processing. Always the language used for the process varies and may not be standardized. The ancient languages may be multi-lingual which requires a lot of understanding and knowledge. In addition to this, there may be issues of the manuscripts being in variety of modes which may possess hindrance.

One way to restrict this language ambiguity is to regulate metadata to certain vocabulary. Controlled vocabularies are often organized hierarchically into a thesaurus or knowledge base which can aid automatic query expansion [2]. This can lead to enforcing a one to one mapping between meaning and form of the object created. The use of metadata from this vocabulary will reduce the errors in data retrieval and provide accurate results. Data cleaning can reduce the errors in the recording of metadata. Normalization can also lend in hand to minimize the errors by segmenting the data.

Natural Language Processing will provide a platform to move from simple text to knowledge base data storage. For accurate data retrieval we need to save and store data in an organized structured form in a knowledge base. As the artifacts are marked and catalogued before the digitization process, each label in the catalogue can

be made a database field in a sequence labeling task. In certain situations, the metadata creation has to be very precise to enable data retrieval of certain artifacts when available in digital format. This can be achieved by automatic metadata creation which is facilitated by most of the scanning softwares during digitization. As the focus is on ancient manuscripts a detail study of the typical spelling and vocabulary may be needed. Data retrieval being querying the databases, we need to be accurate when using modern language to search as the ambiguity of spoken languages may be a hindrance in the process. Therefore, the queries have to be converted to ancient spellings or an alternative may be to convert these ancient manuscripts' index or metadata to new age equivalents.

The digitization project will provide a better understanding of the language processing process including tokenization and parsing. During tokenization, the metadata will be broken down to suitable tokens. Through the process of digitization and saving the digital contents we can also aim at designing a language component which co-relates digital image and textual analysis. The language analysis component can be used to provide accurate data extraction during the information retrieval process. The process will focus on a service-oriented platform keeping in demand the necessity of cultural heritage preservation. Another focal point of the digitization project will be to design and develop tokens, taggers and recognizers along with extractors which can be handy while information searching is carried out. Future enhancements of the digitization project which could be considered would be computational statistics and cognitive science

2. FRACTIONALE AND IMPORTANCE

Literature and cultural heritage provides a connection to the social values and beliefs of the society. The cultural heritage draws people to the ancient traditions and practices which were followed by their ancestors. The study of ancient literature might prove to be beneficial in providing the nationals a sense of unity. Moreover it provides a better understanding of the ancestors' life and their earlier generations. By preserving the literature we indirectly conserve the cultural heritage.

Cultural heritage preservation aims to preserve manuscripts, artifacts, history, cultural relics, monuments and architecture. One of the

benefits which are quite evident from the preservation of cultural is communal support. The cultural heritage comprises a vast amount of large and minute things which the ancestors of the community have written in the manuscripts. The ancestors have recorded their life styles, architectural designs and ancient history in the writings. The agriculture and landscape are also a part of the cultural heritage. A lot of these have been mentioned and recorded in the books, journals and manuscripts.

The purpose of digitization project is to preserve the documents and manuscripts so as to facilitate knowledge to younger generations. The main aspect of the preservation of culture is to archive the history in such a manner which will reduce the physical and chemical deterioration and minimize damage of the original artifact. Primary goal of such a process is to prolong the existence of these manuscripts and ancient documents. The culture of a prior civilization shapes values and aspirations of the existing generations. Digitization is also a means of creating resources that can be re-purposed for unforeseen uses in the future [5].

Cultural integrity is very important which distinguishes people from one another. In addition, the preservation of cultural heritage digitally will enhance the need of assimilation and competence of education. Moreover the preserved cultural heritage in terms of manuscripts and books could be refurbished in case of timely deterioration. Such preserved heritage will attract tourists thereby increasing the economic stability. The temperament of such artifacts by nature is unavoidable but referencing the digital preservation content, they could be restored at any time. The importance of preserving culture, gains importance with time as people relocate from one place to another and become busy to reflect their pedigree. In order to keep the history intact for the days of future and reduce the risk of endangerment it is very essential to preserve the cultural heritage.

Preserving the cultural heritage can be achieved in a number of ways. The latest trend has been to use various ways of technology to speed up and digitise to the maximum the available resources. When thinking of a long term strategy, digitization method ranks at the top of the list. A number of advantages have resulted in this top ranking. One of them is easy accessibility to the general public. Any digitised content can be stored in the digital format and shared through various

means of technology. In addition, when the content is available in digitised form, the number of people accessing the media or object at the same time can be numerous compared to the physical visit. Preserving the cultural heritage digitally will enhance the integration with technology and applications and improve the knowhow. Digital storage will allow information sharing and enhance the popularity of the content in a short span. Going green has been the motto of the Sultanate and preserving contents digitally will certainly reduce paper and save botany. The secondary aim of preservation is reuse. When digitised, the content can be reused and accessed a number of times without causing any damage to the original artifact. The contents are indexed and such a process will enhance data retrieval much easier when compared to the manual process of preservation. In the case of books and manuscripts when preserved digitally can serve as a measure of republishing if the originals are subjected to natural wear and tear. This can in fact serve as a source of disaster recovery and backup. Advantages of the digitization technique does not sum up to the above. It provides a copy right for the content where the accessibility rights are to the original author. The reduction of storage space is another achievement of the digitization technique. The digitization technology will certainly prevent the loss of data caused due to deterioration and mishandling in cultural institutions. The process will also reduce destruction or obsolescence use of the original artifacts. Digitization contributes to the conservation and preservation of heritage and scientific resources; it creates new educational opportunities; it can be used to encourage tourism; and it provides ways of improving access by the citizen to their patrimony [4].

The digitization processes will also open doors for various researches in the direction of natural language processing and image processing. Since the digitised content has to be accessed through various means it has to be enriched to allow smart search and navigation. This literally means linking and connecting related artifacts to each other. Data and information retrieval or access should not be limited to one word search but it should be able to answer complex queries. To enable this sort of searching, it is required to enrich the digitised content in the metadata level. The digitization of the artifacts will certainly increase the accessibility and efficiency but the information embedded in them have to be fully unlocked and made easy to be searched and queried. This will

involve polishing, enriching the information which can be attained by data processing and image processing.

3. STATE OF THE ART

With the emergence of technology and globalization, a need to preserve the history has been observed as an important task. A number of methods have been used recently to capture the history and retain it for the future. A lot of effort has been put into the process of preservation and the momentum is tremendously increasing due to both human and natural factors. The museums, libraries and cultural institutions in the country hold a number of collections which were accumulated over years. The aim of these cultural institutions is to bring the users close to the culture and heritage of the country. Traditional preservation technique still deters the archiving of such cultural heritage manuscripts.

The main advantage of digitization technique is that it allows flexibility depending on the type of material to be digitised. The process is time consuming as well as the expenditure inclusive of manpower and equipment for it is quite high. The process of digitization is very complicated which requires a lot of knowledge in data acquisition and analysis, leading to rendering a modeling of the data or monuments. The process will be inclusive of surveying, drawing and designing, and a lot of post digitization activities such as data management and data storage and retrieval. A lot of policies have to be created and developed by a dedicated team of professionals to start the process of digitization to preserve the culture and heritage. A draft with specified deadlines has to be drawn in order to facilitate easy structure of the project. The process policies will revolve around the cost, time, labor, selection, equipment for digitization and subsequent digitization activities such as data management and data retrieval.



Figure1: Digitisation Process

Digitising the document preserved in cultural heritage institutions and museums only marks the beginning of the process. The digitised content created will be stored in a database management system which should be easily accessible and maintainable by the use of sophisticated devices. Activities such as cleaning, connecting and data refining succeed digitization

and process will also mark an end to the pen and paper based collection management which caused a lot of burdensome to the staff to search for a particular document in chunks of volumes of catalogues. Safe guarding the data loss is main objective of digitization project. The backup process has to organized and well – maintained in order to facilitate future accessibility for various

groups concerned with education, research and innovation.

3.1 Manpower

The main area of digitization is the skill sets of manpower employed to achieve the success and aims and objectives of the project. The initial step of the digitization team will be to collect the ancient manuscripts and documents from the site and bring it to the digitization center which will involve a lot of care and training to handle these artifacts. Prioritizing the artifacts will be a major consideration factor for the people involved. Other responsibilities of the manpower will be cataloging and recording and sorting of the original artifacts to be digitised. In addition to this, the condition and importance of the artifact to be digitised has to be assured keeping in mind the value it will bring to the cultural society. In some special cases to achieve digitization to the maximum, the artifacts and manuscripts might require cleaning which should be only handled by experienced and well qualified staff. The members will be solely responsible for the equipments involved in digitization activity and any faults which may occur should be reverted by proper training before the process starts.

3.2 Selection and Identification of the artifacts

Creation and management of a digital collection will involve data capture, data access and delivery along with the management of the collection. Digitization project is very unique keeping in mind the aim is to preserve the culture and heritage embedded in the documents and manuscripts. The initial step in digitization procedure is the selection or identification of the artifacts and materials. Artifacts are critically assessed to determine their potential values depending on the need and priority. This in turn will also result in the feasibility of the digitization project. A number of factors such as demand, rationale and relevance also guide to the selection criteria. High consideration and concern should be given to the availability of the original artifact for the purpose of digitization. In addition to this, copy rights of the originals have to be attained. Moreover, selection of the available technology needs to be precise so as to accommodate image capturing of such ancient manuscripts. The process of digitization will also focus on the creation of

metadata which has to be feasible to the technical standards.

The digitization project is undertaken to fulfill accessibility need of the clients and users. The type of users who will be accessing the collection is a primary factor along with the delivery of the content to them. The condition and feasibility of the originals have to be carefully examined. Any sort of manuscripts, books, letters, architectural drawings, photographs and prints could be considered and given prior importance. Special care has to be taken when handling such a piece of history as they are always subjected to wear and tear. The manpower involved in handling such historical important artifact should be provided ample training to do so. The physical state and quality of the original artifacts to be digitised has to be considered as this may affect the quality of digital image. One of the central questions for building the digital assets is whether the digital resources will be having sufficient impact in the world [22].

3.3 Equipment and Training

The digitization process will involve the use of lot of softwares and hardwares. Softwares will be required to do the scanning and data capture image of the artifacts. Softwares are also required for formatting the data and tagging of the metadata created. Hardwares such as high precision scanners, hand held laser scanners and high precision digital cameras will be used to scan and capture the digital image of the artifacts. Laser scanners and digital cameras will ensure that the scanning process will be non-contact one and will comply with the damage restriction policies of certain artifacts. Other materials which will come handy are gloves for handling original artifacts and containers for storing the original artifacts before and after digitization.

The staff involved in the process will be accountable for the success of the project. In order to achieve this, proper training has to be provided for using the softwares and hardwares in order to create metadata and explain the digital library. In addition to this, training has to be provided to select and study metadata schemas for explaining each and every digital data created after digitization. Since the value of these historical pieces of artifacts is impeccable, special training should be conducted to educate the staff on how to handle original artifacts.

Preservation of the original artifact is very important as per the culture and heritage norms. Always the physical artifact is a separate entity when compared to the digital artifact created. In our project, we aim to create two digital copies of the artifact, one of which will serve as the master copy and the other will be used for client access.

3.4 Software and Hardware

The digitization process will involve the use four categories of softwares. The first category will be a set of softwares to scan the artifacts and manuscripts and the second category will be used to edit and skew the scanned digital images of the artifacts. The process will also require accurate color reproductions which need monitor calibrating softwares. The photo editing software should facilitate the tagging of images via extended metadata platform. The quality control will be dependent on the color parameters of the digital image created. The scanning process will involve use of scanners equipped with high quality cameras which enables high precision data capture. The scanners selected should be able to produce both grayscale and color images at a wide range of image resolutions. The monitor calibrating software will be used to regenerate the color of the digitised image by quantifying the output of the monitor. In order for the human eye to render these colors, precision between original artifact and digitised content the calibration software is very much essential. The photo editing software will be used for reproducing the color temperature and saturation and hue of the original artifact. The documents and manuscripts will be scanned with or without the use of any auto-correction methods depending on the nature of the artifact. This is done so that the professional working on the calibrating software can modify the color combination accordingly to suite the original document scanned. The format of the digital object should be decided to facilitate the storage and retrieval. We will be crating two versions of the digital artifact-master copy and access copy. The master copy will be saved in the TIFF (Tagged Image File Format) and the access copy will be in JPEG (Joint Photographic Experts Group).The size of the image will be a criteria as this has to be in accordance with the storage and retrieval efficiency. The digitised content will be stored and moved to an integrated data management system (IDMS) which will be the fourth category.

To maintain the quality of the digital content a number of processes might be required. Some of the digital content after scanning will have to be re-oriented to be saved in the digital repository. The digitization professional when scanning and saving the file in the IDMS should not crop the image so as to ensure the entire painting or art or page of book or manuscript has been captured completely. The file naming conventions has to be unique so it can be easy to identify and retrieve information from the database. The digital content after scanning will become a unique identity. Hence a lot of care has to be taken to avoid duplication of the work or the metadata.

Hardware Specifications recommended:

PC Configuration

- Processor: Intel i7
- RAM:8GB
- Hard Disk:2TB
- Network Adapters:2 gigabit Ethernet
- Graphic adapter with Digital output(DVI-D or Display Port),Minimum Resolution:1920 x 1080
- Monitor:24" or larger, Resolution 1929 x 1080 or higher
- Hardware calibrate able Monitor recommended
- Operating system: Windows-[7(64 bit) or above] or Mac OS X [v10.6 or above]

Scanner Specification:

- CCD(Charge Coupled Device) : Trilinear CCD 14.400 pixel
- Lenses: 80mm and 120mm
- Variable Scan Resolution Magnification Range: with 80mm lens: 240-900ppi,with 120mm lens: 400-1800ppi
- Variable Texture Effect: for left mode only
- Variable Reproduction Angle: 0°,10° and 15°
- Original holder: 150 x 250cm
- Maximum size scan: 240ppi: 250 x 150cm,900ppi: 250 x 40cm, 1800ppi: 250x 20cm
- Maximum Scan Height with Light Angle Device: 60cm
- LED lights: with automatic light positioning

3.5 Metadata and Indexing

In the digitization project prior importance is to preserve and retrieve data on client or user demands. For such a purpose, it is important

to store and save data and images in a way which identifies each and every object individually. This information regarding the objects and artefacts which are digitised is termed as metadata. Metadata can be in turn considered to be data about data stored in a database or any other storage sources. Metadata can serve many purposes depending on the person interacting with it. They can be used for technical, administrative, descriptive, preservation, transport and usage purposes.

The registration of the metadata for the digital objects is an important task and has to be done very accurately. A descriptive metadata for each of the digital object created will definitely provide uniqueness. Additional information can be provided in terms of date and time for each metadata. The management of the digital objects is governed by factors such as uncovering of the complete details regarding the digital content, acceptance of the format in which the digital object is saved and stored, extent to which the users can use the digital objects etc. In addition to this, the change of technology which may be used in future should also be considered along with the technical aspects. For longevity and scalability, the best practices used in this project will be documented, covering items such as file naming conventions, metadata entries and structure, and other archival concerns, starting with the digitization documentation [6].

As the goal to create a metadata collection of digital repository is for long term, ensuring the metadata created will facilitate data migration to which ever platform will be used in future is vital. The solution to this is to keep in touch with similar digitization projects which will further motivate knowledge transfer via conferences on digital technology and culture preservation. Re-use and contextualising is crucial for cultural content and if the metadata standards and interoperability rules are followed, the user can create his own virtual collections in minutes, can learn the stories behind the object of his interest, can organise and re-use his personal collections, share with others, print, etc [23]. The metadata will be technically created and will comprise of the following:

Table 1: Metadata Set for Digitization

Technical Metadata	Additional Metadata
Name of File	File Creator
Type of Document or	Contact

manuscript	Information
Application details	City
Date of Creation	Country
Date of Modification	Reference no/Title
Size of File	Description
Bit Depth	Source Provider
Colour Mode and Profile	Copyright Notification
Dimensions	
Resolution of the digital content	

In addition to this, it is recommended to develop and create bibliographic metadata to provide details of the manuscripts

3.6 Quality Assurance and Control

The digitization process is undertaken to preserve the cultural heritage of a civilization. So all measures has to be taken to ensure the quality of the deliverable is above par. The primary step is the quality of equipments used for the process. The scanner used should render the colour quality of the object to be scanned to the maximum. Colour imbalances can be rectified by the use of monitor calibrating softwares and colour processing softwares. These all factors regarding colour aspects need precise training of the staff which is provided in the pre-process. In a digitization process, one of the most important areas is colour management of the digitised content. Monitor calibration is the first and most important step in colour management [1].

Digitization process will result in a data collection of digital objects. The digital objects have to be indexed and given metadata accordingly. In order to this properly, we require a file validating software which will on a timely basis check the virtue of the objects created and stored in the database. Moreover, a frequent back up of the database is required to ensure disaster recovery and back up. The file validating software will be also monitoring the back up to ensure the uniqueness. Long term preservation must be the focus of our project. The policy and staffing has to be made accordingly to facilitate this. The technological quality of the softwares and hardwares used must be carried out periodically by experienced staff in order to avoid any user-oriented issues and breakdown of the system. A Quality Control (QC)

and Quality Assurance (QA) metrics will be designed to keep track of the quality aspects of the artifacts being digitised along with a quality maintenance record for the equipments used. The quality and integrity of born-digital material will have an influence on how rapidly it can move through the production and descriptive process [12].

3.7 Digital Back up and Disaster Recovery

As the aim of the project is preservation, we have to take into account long term preservation and disaster recovery aspects. In order to achieve this, back up of the digital repository of artifacts should be created and stored at on-site location and should be updated on a daily basis depending on the digitization carried out. In addition to this, it is also suggested to keep an additional back up of the digital contents to safeguard the preservation of culture and heritage. This secondary back up should be kept at an off-site location. This back up should also be updated on a regular basis to save the digital collection so as to serve as an emergency option in case of disaster recovery. In addition to this, we can focus on Network Attached Storage (NAS) and Storage Area Network (SAN) to achieve long term for the master copy of the digital content. Similar to the value and security of the original artifact, the security of the digitised content is also important which can be restricted to only authorized person involved in the project [18].

4. DESCRIPTION OF THE PROPOSED DIGITIZATION PROCESS

A draft plan has to be created to facilitate all the stages which will lead to success procedure of digitization project and will ensure proper preservation and data retrieval. A well defined set of policies and benchmarks to maintain quality is also be outlined. A high level DFD is proposed in Figure 2 (Digitization Data Flow) that depicts the proposed methodology of the project. The generic digitization process that will encompass and implement all the key functionality and features for proper execution of this project is defined here.

4.1 Methodology

The entire digitization will comprise of a variety of stages such as consulting the collaborators and planning and defining a digital infrastructure. The digital infrastructure comprise of two key components: the digital preservation network of partners (creators, owners, collectors, users) collaborating to preserve and provide long-

term access to digital content; and the digital preservation architecture (structures, logical components and logical relationships), the technical components that enable digital preservation [17]. Digitization in the modern era has been so far the go to solution for preservation of culture and heritage. The DFD below divides the activities involved in digitization process to mainly two places namely: activities performed at Digitization center and activities performed at the acquisition site.

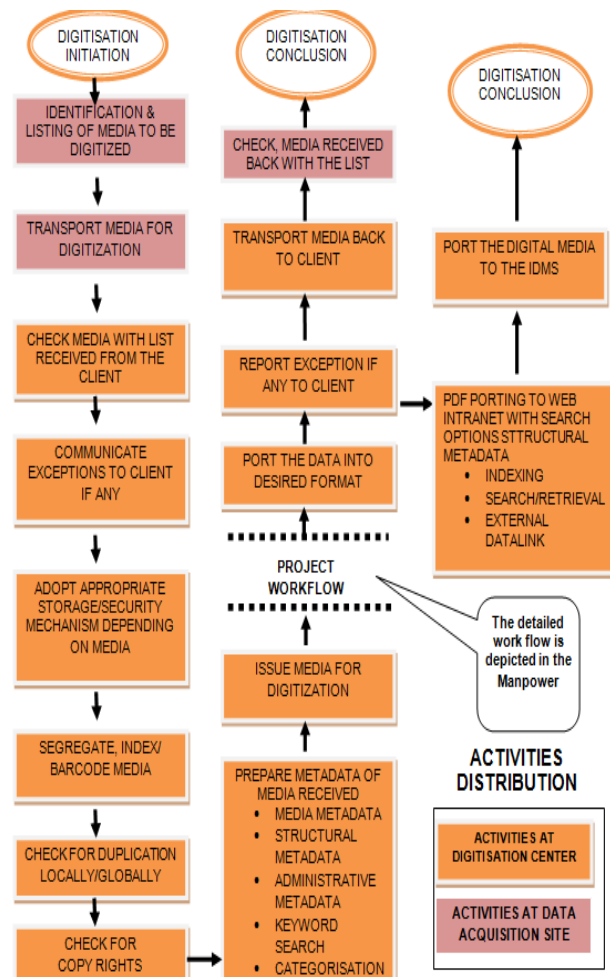


Figure 2: Digitization Data Flow

The process is initiated with the collection of artefacts from the site and cataloguing it to maintain a record. This is followed by the setting up of equipments along with the installation of softwares. After this process, the digitization procedure begins with the copy right valuation followed by indexing and digitising and metadata creation. The digital object is then transported to

integrated database management system for storage after scanning and calibration. For long term archiving, implementation of SAN and NAS is required. This is in addition to the storage of digital content in the hard drives and system storage. The digitization process will add value to the cultural and heritage exhibits and make them more exposed to the scientific world. This will enhance the research and interest of researchers and upcoming generations.

4.2 NAS and SAN Implementation

As the main goal and objective of the digitization process is long term storage and preservation we need to have a database where there is option for advanced storage and information management. Two storage techniques can be considered for this purpose-Storage Area Networks (SAN) and Network Attached Storage (NAS). In addition to this, we will be using traditional data storage model which consists of disk drives attached physically to the database server. The database is connected to the server using SCSI or SSA. But adoption of this strategy is restricted as compliance with scalability, restriction of distances between server and database, multiple storage sharing and performance and speed.



Figure 3: Data Storage Model

Network Attached Storage may be considered a modification of the traditional model where the network sharing of the information is added. In NAS, a number of storage drives are connected to database server which facilitates storage across the network through protocols such as NFS and TCP/IP. Considering the traditional model aspects such as distance between server and database could be eliminated in addition to improvement in the speed and data management. As the model includes a number of networks, the load over the network will be high as the number of requests to be processed will increase.

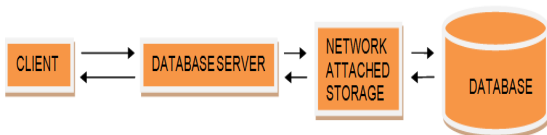


Figure 4: Network Attached Storage

Storage Area Networks reduce the load on the server by providing fiber channel base.SAN uses RAID hardware which provides mirrored caching and dynamic storage allocation. In addition to this, data protection and backups can be easily obtained.SAN also provides support for databases on various platforms. By implementing SAN we could also consider implementing a clustered database. The basic difference between NAS and SAN is that in the SAN the disks in RAID array are not shared at the file system but at the block level.

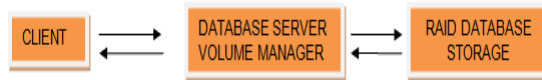


Figure 5: Storage Area Network

In traditional database storage access model, servers accessing storage involves three basic steps-client requests data from the database server over the network, database server does disk I/O and the result is sent to the user or client over the network [24].Compared to the traditional model, the NAS model has five basic steps to access the data[24].

The additional step which is included in the NAS model is that the database server should request the NAS server for the I/O disk check which in turn provides a response to the database server following the reply from the database. In SAN, the client makes data request over the network to the database server, database server performs the I/O over the fiber channel I/O bus to hardware RAID and then database server sends results back to the client over the network [24].

4.3 Tasks and Phases

Following would be the schematic of the various tasks involved in the proposed project.

1. *Requirement gathering Phase:* While the macro level requirements of the project have been understood to get a more granular insight into the requirement of the project, the requirements in the following areas are gathered.
 - a. IDMS customisation
 - b. Document preparation and preservation and handling
 - c. Digitization standards
 - d. QA, QC and metrics
 - e. Metadata requirement

- f. Storage and backup
g. Site inspection for digitization facility
2. *Design Phase*: Based on the inputs from the Requirements gathered the following strategies are designed.
- Document handling strategy
 - Identification of scanners and equipment for image processing
 - Customisation of the automated workflow
 - QC strategy and deliverable structure
 - IDMS customisation and deployment
 - People requirement.
 - Hardware architecture design
 - Storage and backup policy
 - Metadata schema
 - Additional software requirements
 - Site layout and work areas definition
3. *Site, document and equipment preparation*
- Preparation of the site
 - Segmenting the site and layout creation
 - Painting
 - Lighting and backup power
 - Networking
 - Tables chairs and furniture
 - Equipment preparation
 - Power and backup
 - Installation
 - Calibration of scanners
 - Servers, PCs and storage installation
 - Installation of customised IDMS software
 - Installation of workflow-customised with all elements
 - Installation of all accessory software
 - Procurement of document preparation consumables
4. *Implementation phase*
- Kick-off with recap on procedures for various teams
 - Setting Standard operating procedures for
 - Document preparation
 - Scanning
 - QC Protocol
 - Communication
 - Creation of roles, users and tasks and assigning to respective persons.
- d. IDMS links establishment
5. *Operations and production phase*
- Document preparation
 - Metadata Preparation
 - Scanning
 - Segmenting and Structural Metadata preparation
 - Quality control – at various stages of workflow
 - Export of data to the IDMS – Metadata and Images
 - IDMS rendering of data
 - Process Adherence Monitoring
 - Document Accounting
 - Maintenance, preventive and breakdown of the following :
 - Scanners
 - Servers
 - PCs- Peripherals
 - Storage
 - IDMS
 - Infrastructure
- Based on the tasks and phases explained above, the following deliverables are expected:
- Assessment Matrix with Document control sheet
 - Digitised object
 - Data Migration
 - Final Technical report of the project
- 5. MANAGEMENT PLAN AND PROJECT ACTIVITIES**
- The project should be managed by a management committee consisting of the lead researcher and other members. Project technical progress, financial updates, commercialization, dissemination, and technology transfer, as well as project reporting should be discussed on a regular basis. Regular interaction with the industry group will also ensure that the project is meeting the objectives of the interested industry partners.
- Management Plan will be as follows to cart the deliverables:
- 1st Deliverable: The Technical report summarizing the state of the art, requirement specifications and test best deployment experiences.
 - 2nd Deliverables: Digitised Object
 - 3rd Deliverable: Data Migration
 - 4th Deliverable: Final Report

The project is divided into a number of activities. These activities are in turn categorized into different phases. The phases involved in the digitization project are depicted the table.

Table 2: Digitization Activities

S.No	Activity
REQUIREMENTS PHASE	
1	IDMS Customization
2	Document preparation and preservation and handling
3	Digitization standards
4	QA, QC and metrics
5	Metadata requirement
6	Storage and backup
7	Site inspection for digitization facility
DESIGN PHASE	
8	Document handling strategy
9	Identification and scanners and equipment for image processing
10	Customisation of the automated workflow
11	QC strategy and deliverable structure
12	IDMS customisation and deployment
13	People requirement
14	Hardware architecture design
15	Storage and backup policy
16	Metadata schema
17	Additional software requirements
18	Site layout and work areas definition

SITE,DOCUMENT AND EQUIPMENT PREPARATION	
19	A. Preparation of the site i. Segmenting the site and layout creation ii. Painting iii. Lighting and backup power iv. Networking v. Tables chairs and furniture
20	B. Equipment preparation i. Power and backup ii. Installation iii. Calibration of scanners
21	Servers, PCs and storage installation
22	Installation of customised IDMS software
23	Installation of workflow-customised with all elements
24	Installation of all accessory software
25	Procurement of document preparation consumables
IMPLEMENTATION PHASE	
26	Kickoff with recap on procedures for various teams
27	Setting Standard operating procedures for: i.Document preparation ii.Scanning iii.QC Protocol iv.Communication
28	Creation of roles, users and tasks and assigning to respective persons.
29	IDMS links etc established
OPERATIONS AND PRODUCTION PHASE	
30	Document preparation
31	Metadata Preparation
32	Scanning
33	Segmenting and Structural Metadata preparation

34	Quality control – at various stages of workflow
35	Export of data to the IDMS – Metadata and Images
36	IDMS rendering of data
37	Process Adherence Monitoring
38	Document Accounting
MAINTENANCE	
39	i. Scanners ii. Servers iii. PCs- Peripherals iv. Storage v. IDMS vi. Infrastructure
RESEARCH AND DEVELOPMENT	
40	Data Set Preparation
41	Experiments Implementation
42	Results Publication

6. BENEFITS OF DIGITIZATION

Manuscripts and transcripts represented in analogue and images represented in canvas, negatives or glass plates reflect the cultural heritage of a nation. Drawbacks of these representations include bondage to the physical aspect, temperamental of these artefacts and loss of the original quality when duplicated. Digitization eradicates these drawbacks and puts the originality back into the artefacts if duplicated. The digital version procured after the digitization process provides linking to multimedia techniques and independency over the physical barriers is achieved. Storage options are also enhanced with the digitization aspect as it reduces physical space. With the latest trends in computing the digital content can be distributed to wide network of clients. The preservation of cultural heritage has an impact on the economic prosperity of country inclusive of the tourism and business. It also has an effect on the development of environment and education. The advantages of the digitization can be summarized as access, support of preservation

activities, collections development, institutional and strategic benefits, and research and education [5].

Global Accessible:

- The digital content could be accessible to authorized personnel from anywhere in the world.
- The people all over the world can access these digital contents at any time from any part of the world.
- The contents will be available all time 24*7.

Minimize the cost and time of storage:

- The digital objects would reduce the physical storage space compared with storing lot of information as paper documents
- In the commercial area, the space and time required for storage and retrieval of paper based documents cost high. By digitising the objects and artefacts this will be considerably reduced.
- In our digitization process, scanning documents and interpreting them into digital system can reduce the amount of prime storage space and cost required by paper.

Quick Information Retrieval:

- If the documents are scanned it could be accessible from any computer by any authorized people compared with traditional paper based documentation which has to be housed in a file cabinet and request through another person for delivery.
- The digitised content can be retrieved easily and it can be shared and routed to any person who needs it.

Internet access:

- In recent days at most all countries are well equipped and connected through broadband networks; actually people need only internet and gadgets to access these digital documents.
- Moreover these digital documents are available with required format that the person expects.

Lost documents and disaster recovery:

- This digital documentation process would eliminate the loss of documents and information through various elements

- We can recover the data and information immediately after any calamities

Flexible Indexing and Data Retrieval:

- Indexing the images of digital documents can be done several ways, these can help to retrieve files by any word or phrase in the document, this capability is not possible with paper or microfilm based documents which are done manually.
- The digital documentation system can have single or multiple taxonomies and categorization.
- The digital documents distribution is multi-threading and folding, one image can be distributed to multiple terminals simultaneously.

Cultural & Heritage Preservation:

- In this phase of project we would digitise the cultural and heritage related documents and images.
- Essential and significant credentials from important eras in the history will be accessible online.
- The cultural heritage of a civilization will be known to the entire part of the world.
- The project will also make accessible recordings of traditional music, dance, paintings and other fine arts.

Internal relationship:

- The document deliberates on relations with other countries and region over political matters, trades and relationships.
- The comprehensive selection of resources gives detailed reports of the geography and mineral wealth.

Student's benefits:

- These digital documents would be helpful to those studying about culture and civilizations.
- The unique records could be easily accessible by students and researchers.
- The digital documents and photographs could show the region's older days after tremendous infrastructure development like road, flyovers and massive building structures.

The cultural heritage artefacts when available in digital format can tremendously enhance education sector. The educational institutions can directly approach and refer these

historical manuscripts and deliver an educational module based on these collections available. Dynamic learning of history and culture of the country would definitely grasp the interest of students when technology changes the faces of books and manuscripts.

Moreover, the benefits will also be institutional and national and strategic. The project will certainly mark an upgradation digital profile while converting the physical national treasures such as documents, arts and manuscripts to digital format which can be considered very prestigious. The project will in turn lead to the development of collection which is rich in diversity representing the culture and heritage. The rich corpus of data obtained through the project will open gateway for the research community to advance more in the direction of Natural language processing, image processing etc. It will also help the Government to preserve the historical records for life time. The digitised content would help the citizens and international research society to pursue research in this area.

7. ECONOMIC IMPACT OF DIGITIZATION

The digitization can consecutively convert the cultural manuscripts and documents and pictures and arts to ingredients of stabilized economy. All the digital resources of cultural heritage when available through technologies such as internet and intranet can motivate and increase the prospects of tourism and education. It can also trigger the spread of culture to different countries which can in fact influence the architecture and design as well. In addition, this can also provide the civilization profile a better quality of the culture and heritage diversity. As the accessibility of the digital collections increase with digitization, the number of tourists and visitors will get a boost. More over the regional economy will also be influenced due to this factor of tourism. Employment creation can be a prospect of the digitization technology both directly and indirectly influencing the economy of the country. Promotion of the cultural diversity can be triggered along with growth of job creation. In the area of economic development, multilateral development banks, bilateral development agencies, institutions and individuals involved in development programmes have recognized the urgent need to assist developing countries in preserving their cultural resources and assets, and moreover to relate cultural values to development[15].

8. SOCIAL IMPACT OF DIGITIZATION

Digitization would bring the world within a box where exchange of the cultural values could be facilitated. The ancient and antique manuscripts and documents would engage wider publicity and connections. In short it would be a means to improve the information and communication and strengthen the international links. Such a project will increase the awareness of the culture and heritage of the country among the citizens and foreigners. Throughout the phase of digitization, we will have to collaborate with multiple industries and organizations. The participation of the civil structures is highly crucial for the success of cultural heritage research, documentation, preservation and promotion [14]. The growing interest in the links between cultural heritage preservation and mainstream societal concerns such as development, environment, health, education, access to information, construction and the economy is self-evident [15].

9. ACADEMIC, SCIENTIFIC AND INNOVATION SIGNIFICANCE

Digital technologies and applications contribute to the economic strength, societal well-being, and effective governance of a nation [20]. Modernity is the essence of the digital world and digitization can gradually change the face of academia and education. In a way, the technology is modifying the meaning of words and letters. The leap from print to digital format caused to an extent by technology boom will certainly mark an innovation in the publishing and educational institutions. Digitising the collections can be considered to be a technological development stage in the country. Technological development lead to multichannel, multimedia destination management system, serving purposes not only of travel information distribution, planning and fulfillment, but also of travel-related education and entertainment, which is important for sustainable development of tourism [4]. This will also enhance information sharing and knowledge distribution throughout. In long term, digitization can be considered to be building blocks of a digital economy. The project will also spice up accessibility of cultural heritage and knowledge and will inject fuel in areas of advertising and tourism.. Teaching and learning excellence are also important cornerstones of effective digital resources [22]. Moreover, digitization helps researchers and

curators at cultural heritage institute themselves, as information can be found more quickly and curation process can be partly automated [2].

The advantages to academic research and advanced scholarship are equally impressive, and the potential of networked technologies to create a dynamic reading and scholarly environment is driving digitization initiatives at many institutions [5]. The ancient manuscripts, maps, books and documents which were accessible only to a limit, through digitization will re-live and restore its grandeur. The data discovery of the digitised content will also be enhanced. The artefacts which are of historical importance will open up a cross cut relationship between different streams of research such as archaeology and typography. In addition to this, the creation of a digital collection will allow to develop relationships with cultural institutions. The information shared will be on the basis of workflows and knowledge databases and digitization standards. Digitised materials are perfect for training students en masse and preparing them for scholarly research [8]. Besides all these significances, the digitization will ensure perpetuation of the cultural heritage of the nation

10. CHALLENGES IN DIGITIZATION

Digitization is the process by which information about artefacts and specimens held in museums and other collections (e.g. identifications, descriptions, geographical locality information, collecting information, images, etc.) is converted from analog (e.g. ledger books, catalog cards, specimen labels) to digital form and made available to users via the web or technology [21]. During the entire process, the original artefacts which are of historical importance has to handled with much care to avoid damage and deterioration. The conversion of the analog data to digitised format through the process of data capture such as scanning is complicated and requires high precision. The challenges do not engulf to the above but it will only sum up to more issues such as storage of both analog and digital format of the artefacts. The main objective to undertake digitization project being information sharing and knowledge distribution, a lot of attention has to be paid to the method of assessment and mobilization of the digitised content. The hardware and software used in the process has to render high quality and accuracy in order to achieve the objective of the project. Though technology is a hindrance, the main

concern will be the human resource engaged with the project. The staff training has to be done from the initial stage of selection of the artefacts along with the conversion of the manuscripts to digital image and to the storage of this content in an integrated database management system. Digitization projects are rife with copyright problems, and their restrictive legal conditions can often complaint and consternation [7]. Digitization projects face great uncertainty with regard to the level of the copyright license fees due [9].

prompt research aspects but as well be an aspiration to explore more aspects of technology when archives and surrogates created are made available at one's disposal. Through the course of digitization, we not only conserve our past but ensure abduction of years to come as it remodels to our past.

The process of digitization will preserve artefacts and documents and provide access to authorised professional at any location on any time basis.

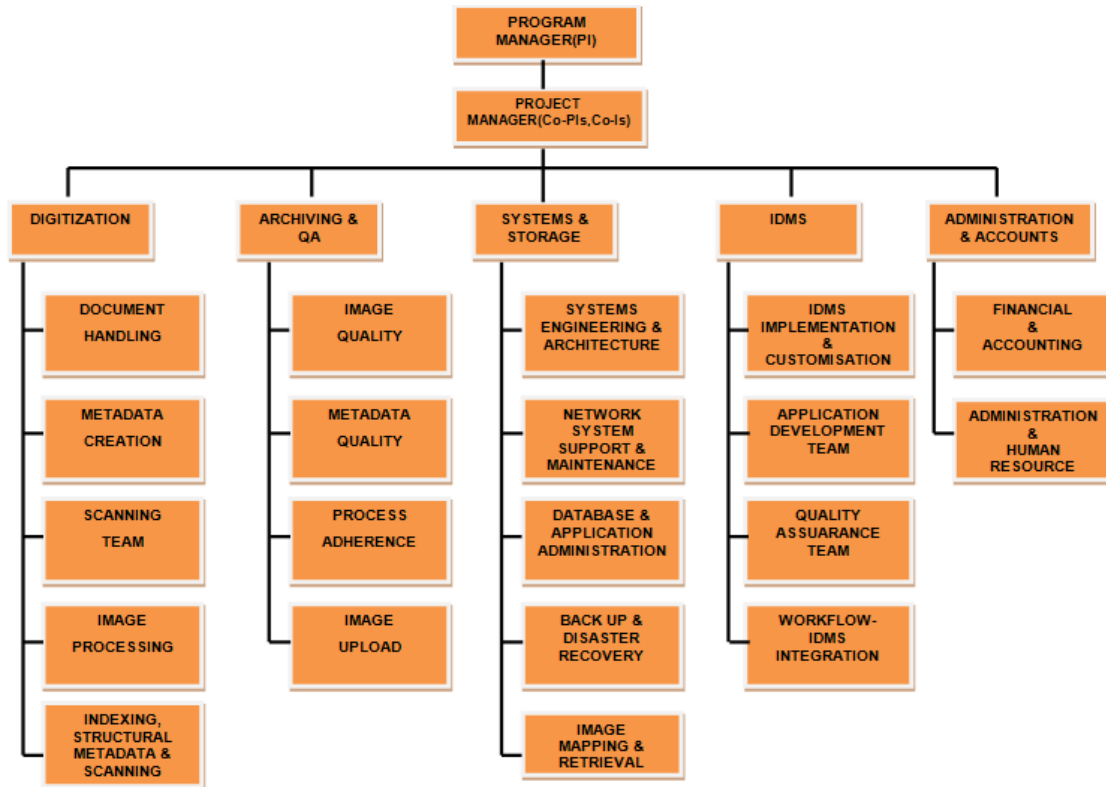


Figure 6: Work breakdown Structure

11. PROJECT EXECUTION BREAKDOWN

The staff involved is to manage and execute the scientific tasks of the project, design the theoretical solutions and supervise and support the other human resources to achieve the outcomes of the projects. They are also responsible for writing technical reports. The work breakdown is show in Figure 6.

12. CONCLUSION

Digitization would literally accomplish archiving heritage and would provide resources to uphold the affluent and prosperous cultural heritage of even the smallest society. This will not only

Data retrieval time will be considerably reduced as well. Data retrieval will open doors to new strategies in natural language processing in terms of creating metadata and indexing the artefacts including lexical analysis and data extraction. Digitization also enables the physical documents to be handled less and hence not detrimental to the documents that are in a fragile condition without compromising on the access.

As the value of the artefacts to be digitised is extremely high and the consequences to damage may be far fetching, scanning will need to

be done on non-contact type overhead scanners which are both gentle on the documents and in most cases may not require the binding of the documents to be removed. This reduces the damage to artefacts some of which are extremely old and precious. Digitization will focus on facilitation of data access and prevention of data loss due to deterioration and obsolescence. Researchers could also approach such collections provided in the repository for analysis and review.

With more sophisticated process and techniques the process of digitization could be more simplified and brought down to a hands-on task. This could provide the way for preserving the culture and heritage through the digital printing techniques which is now available. The museums and cultural preservation centers could employ the process and print digital images of monuments and artifacts and make it available within the vicinity of public. The paper just aims to provide insight to the entire process of digitization through NLP and indexing. The process explained and explored are on the basis of methods followed in the digitization industry. Through the evolution of time, many of the process will be updated depending on the facilities available.

REFERENCES

- [1] Zayed, A. A. (2009). Digitization Workflow and Guidelines - *University of Exeter*.
- [2] Sporleder, C. (2010). Natural Language Processing for Cultural Heritage Domains. *Language and Linguistics Compass*
- [3] Grussenmeyer, P., Landes, T., Voegtle, T., Ringle, K., (2012), Comparison Methods of Terrestrial Laser Scanning, Photogrammetry and Tacheometry Data for Recording of Cultural Heritage Buildings, *International Society for Photogrammetry and Remote Sensing*
- [4] Hadžić ,O.,(2004), Tourism and Digitization of Cultural Heritage, *Mathematical Institute SASA*
- [5] Why digitise? The costs and benefits of digitization, *Facet publishing*, Retrieved from http://www.facetpublishing.co.uk/downloads/file/sample_chapters/digitising%20collections%20chapter%201.pdf
- [6] Okerson,A.,Beaudin,E.,(2005), Iraq ReCollection : A Proposal for Preserving Iraq's Cultural Heritage, *Yale University Library*
- [7] Rikowski,R., (2010), Digitization Perspectives,*Sense Publishers*
- [8] Proffitt,M.,Schaffner,J.,(2009), The Impact of Digitising Special Collections on Teaching and Scholarship ,Reflections on a Symposium about Digitization and the Humanities, *OCLC Programs and Research*
- [9] Korteweg,D.A,Hugenholtz,P.B, (2011), Digitization of Audiovisual Materials by Heritage Institutions: Models for Licenses and Compensations,*SSRN*
- [10] George Pavlidis , Anestis Koutsoudis, Fotis Arnaoutoglou, Vassilios Tsioukas, Christodoulos Chamzas, (2006), Methods for 3D digitization of Cultural Heritage, *Science Direct Journal of Cultural Heritage Volume 8 Issue 1*
- [11] Federal Agencies Digitization Guidelines Initiative, (2009), *Digitization Activities, Project Planning and Management Outline, Version 1.0*
- [12] Poole,N.,(2010), The Cost of Digitising Europe's Cultural Heritage : A Report for the Comité des Sages of the European Commission, *Collections Trust*
- [13] Youngs,K., (2010), Managing the Digitization of Library, Archive and Museum Materials, *Collections Trust*
- [14] Borrissova,V.,(2010), Digitising Cultural Heritage in Bulgaria A Survey of Intellectual Property-related Experiences and Practices
- [15] Cassar,M., (2006), Evaluating the Benefits of Cultural Heritage Preservation: An Overview of International Initiatives, *Centre for Sustainable Heritage*
- [16] Laužikas , R. ,(2005) Digitization of Cultural Heritage: Model of an Integral, Three-Dimensional Spatio-Temporal Thesaurus, *Online Publication System of the University of Tübingen*
- [17] Preserving our Digital Heritage, Plan for the National Digital Information Infrastructure and Preservation Program, (2002), *Digital Preservation-Library of Congress*,Retrieved from www.digitalpreservation.gov/documents/n_diipp_plan.pdf
- [18] Reddy, E.R, (2001)A Proposal for Action Plan for Digitization of University Libraries, *Documentation Research and Training Centre Indian Statistical Institute, Bangalore Centre, India*
- [19] Status of Technology and Digitization in the Nation's Museums and Libraries, (2006),*Institute of Museum and Library Services*
- [20] Sabbagh,K.,El-Darwiche,B.,Friedrich,R.,Singh,M.,(2012),

- Maximizing the Impact of Digitization,*Strategy & Reports, India*
- [21] Prospects for a scientific software innovation institute in biological collections: Digitization,2011,*The Society for the Preservation of Natural History Collections*
- [22] Meyer,E.T, (2011), Splashes and Ripples, Synthesizing the evidence on the impacts of digital resources, *Oxford Internet Institute at the University of Oxford*
- [23] Sotirova,K.,Peneva,J.,Ivanov,S.,Doneva,R.,Dobrev,M.,(2012), Digitization of Cultural Heritage-Standards, Institutions, Initiatives, *Access to Digital Cultural Heritage: Innovative Applications of Automated Metadata Generation*
- [24] Schulz,G.,(2000),Database Configuration: SAN or NAS,<http://www.bytepile.com>
- [25] Günlü,E, Yağcı,K, Pınar,I,(2009),Preserving Cultural Heritage and Possible Impacts on Regional Development: Case of İZMİR, *International Journal of Emerging and Transition Economies*