



Open Access

**Q**LANTIC  
 JOURNAL OF  
 SOCIAL SCIENCES  
 AND HUMANITIES

## Digitization Process of Archives: A Case Study of the Punjab Archives, Lahore

Mukarram Nawaz<sup>1</sup> Muhammad Nadeem Akhtar<sup>2</sup> Dr. Muhammad Tariq<sup>3</sup>  
 Sidra Batool<sup>4</sup> Shahbaz Ali Anjum<sup>5</sup>

**Abstract:** *The digitization process of archives involves the conversion of physical documents, records, and artefacts into digital formats, enabling easier access, preservation, and dissemination of valuable historical information. The present study aimed to explore the digitization and preservation of archives: a case study of the Punjab Archives, Lahore. This study uses a qualitative research methodology in order to achieve the study's goals. An inductive approach was used to collect the data. Data was collected from the staff of the Punjab archives team, the project team and primary sources. Data was converted into written transcripts with the help of NVivo software. The benefits of digitization, such as improved accessibility, enhanced searchability, increased security, efficiency of accessing resources, and the protection of cultural heritage, are discussed, along with potential benefits, challenges, and solutions. The findings of the study revealed that the Punjab archives are the richest and oldest archives in the region. Seven million archives are available in different forms (documents, manuscripts, government records, coins, stamps, lithography, books, pictures, maps, etc.) in the Punjab Archives, and 114,592 documents have been selected for the initial phase of digitization. The purpose of the project was to preserve damaged, forged, and torn records in digital form.*

**Key Words:** Punjab Archives, Digitization Process, Cataloguing, Challenges, Documents

### Introduction

Archives play a crucial role in preserving historical records and cultural heritage. With the advent of technology, many institutions are embracing digitization as a means to safeguard and provide wider access to their collections. This article explores the digitization process of The Punjab Archives in Lahore, highlighting the steps taken, challenges faced, and benefits achieved through this endeavour.

There is a variation in every digitization project in terms of needs, including customized preparation for technology requirements, source material selection and restoration, staffing, budget, and archival storage (Mariner, 2014).

Digitization is a way that includes converting physical records, documents, and other analogue materials into digital formats. This process has gained significant importance in recent years due to its ability to preserve and provide wider access to archival collections. It would be beneficial to begin studying the digitization process of Punjab Archives by establishing a clear definition of the terminologies used for archives. Not all records are archives, but not all archives are records. It is critical to first recognize the differences between their categories.

Xie and Matusiak (2016) pointed out that the core of the digitization process is digital capture, which is the process of sampling analogue data and transforming it into digital signals using equipment like

<sup>1</sup> Digital Imaging Officer, Punjab Information Technology Board, Lahore, Punjab, Pakistan.

<sup>2</sup> Head, Department of Main Library & IT, Ayub Agricultural Research Institute, Faisalabad, Punjab, Pakistan.

<sup>3</sup> Incharge, Library Information Services, COMSATS Institute of Information Technology, Faisalabad, Punjab, Pakistan.

<sup>4</sup> Technical Librarian, Ayub Agricultural Research Institute, Faisalabad, Punjab, Pakistan.

<sup>5</sup> Librarian, Abduds Salam School of Mathematical Sciences, Government College University, Lahore, Punjab, Pakistan.

▪ **Corresponding Author:** Muhammad Nadeem Akhtar ([ch.nadeem@aari.punjab.gov.pk](mailto:ch.nadeem@aari.punjab.gov.pk))

▪ **To Cite:** Nawaz, M., Akhtar, M. N., Tariq, M., Batool, S., & Anjum, S. A. (2024). Digitization Process of Archives: A Case Study of the Punjab Archives, Lahore. *Qlantic Journal of Social Sciences and Humanities*, 5(1), 240-247.

<https://doi.org/10.55737/qjssh.391567320>

scanners, digital cameras, or analogue-to-digital converters. All civilized cultures have long retained records of human activity, including their digitization, preservation, and modernization. Archives are available as proof of social activity and are important information sources for historical inquiry.

The Punjab Archives are the richest and oldest archives in Pakistan. The Punjab Archives, founded in 1923 by Professor Garrett of the Indian Education Service, has worked tirelessly to preserve centuries-old records, including the history of the Tomb of Anarkali, the Mughal Rule in India, the art of miniature painting, the life and times of Ranjit Singh, and miniature paintings and lithographs from the Mughal, Sikh, and British periods. It also has documents from Delhi, Haryana, NWFP, Afghanistan, Nepal, and other countries, putting it in the top tier of South Asian repositories. The Punjab Archives has a large collection of seals, coins, and weapons (Niazi, [2008](#)).

Fatima (2004) states that an archival repository is a location that houses historical records, documents, and other materials. Archives is a term that is often used in the plural to indicate the physical location where permanent documents are arranged, described, conserved, and made available, as well as the department within an organization that is responsible for maintaining the archives.

This was the first of its sort of initiative, in which digital cataloguing criteria were used to locate data throughout the digitization process. After months of research and a thorough examination of Punjab's Rules of Business, digital document codes (DDC) were introduced. These special codes are assigned to departments in addition to the existing file numbers.

## Objectives

- To investigate the digitization process of the Punjab Archives.
- To explore the techniques being used by the Punjab Archives Department for the digitization process of documents.
- To explore the challenges faced by the Punjab Archives during the digitization process.

## Literature Review

Archives are consistently implementing the digitization process for archived assets as part of the organization's mission and vision to facilitate internet access for future generations. The general practice of digitization offers prospects for the growth of internet accessibility because some of these practices use different techniques, such as book scanners.

Institutions or organizations with high workloads often face management problems, especially with archive documents. Good and clean data storage is important to keep data safe, effective and secure and can increase the profitability of the business (Amalia & Panduwinata, [2022](#)).

The rapid development of science and technology in recent years, on the one hand, benefits the efficiency and convenience of many human activities, and on the other hand, these developments also have an impact, especially on archives. , needs an urgent estimate (Wirasaputrakkk, [2022](#)).

Old documents containing valuable historical information and special details are now presented and accessible electronically (Khadjahdkk. [2021](#)).

Every type of organization, large or small, should have a goal to achieve. The purpose of the organization is to ensure efficiency and effectiveness while working for management. Many organizations are not very interested in using technology to do their job. Mail and important documents placed in the room can affect the operations of the institution or institution if not managed properly. In addition, incorrect information management can make it difficult to find information or even lose it (Rozana & Musfikar, [2020](#)).

The concept of global archives has changed from closed buildings to limited open and easily accessible digital archives. By conducting extensive research and extracting the necessary information, the practice of storing information has been replaced by websites, online services and virtual tours. Keeping up with the digital world and opening the doors of archives to the outside world, especially researchers and



historians from universities around the world. Thereupon, Punjab Archives decided to collaborate with Punjab Information Technology Board will digitize the records (Chughtai, 2019).

The John F. Kennedy Presidential Library, according to Ulyatt (2014), is digitizing millions of papers in multiple formats by producing low-resolution, web-ready JPEG files and making them accessible through online locating services. The Archival Material of American Art scanned 76 collections between 2005 and 2009 for a total of 512 linear feet; new technology bought in 2010 could scan a page in colour in about one second. The National Archives of the United Kingdom has digitized approximately sixty million records using a variety of technologies. For the 1911 Survey, five automated scanners operated nonstop to generate over 40,000 images each day. In five days, the images were electronically transported to the Philippines, where they were transcribed and returned with full quality assurance.

Bultmann et al. (2006) conducted a survey of the digitization efforts in the UK research library and archives industry. According to the research, 51 institutions that participated in it were actively involved in digitization. Access to the collections was rated as the main benefit of digitization. The arts, humanities, and social sciences predominated, with sciences, religion, health, and technology following.

The process of digitization allows for the creation of digital substitutes that can shield originals from damage. Participating in the creation of digital material via digitization may significantly increase professional confidence while also being advantageous to users (McMenemy&Poulter, 2005)

According to the research, the "relevance to the aims and objectives of the organization" criterion is the one that is most frequently used when choosing which materials to digitize, followed by user demand, rarity or uniqueness of the material, and the presence of a coherent collection.

Mat and Kolej (2005) mentioned that the methods of storing, preserving, disseminating, and providing efficient access to users in libraries were revolutionized by digitization. The contribution and preservation of heritage and scientific resources were distinguished by digitization. It provides additional educational possibilities, may be utilized to promote tourism, and enhances citizen access to their heritage.

Alhaji (2007) stated that the respondents identified inadequate funds, lack of technical equipment, lack of knowledge and skills, lack of organizational management, and lack of power as major limitations in the digitalisation of tests and documents, as well as in the state. Papers. Thirty Nigerian Universities admitted that the respondents admitted to the survey that inadequate funding, lack of technical equipment, lack of knowledge and skills, lack of policies and digitalization of past examinations were not strong enough to be a significant stop. Thirty Nigerian Universities accepted their national certifications.

According to Rafiq and Ameen (2013), university libraries around the world are launching digitization projects to convert their knowledge resources from analogue to digital formats and provide electronic access. They evaluated the capacities of Pakistani university libraries to launch, complete, and sustain digitization projects, as regarded by university library administration and senior library professionals. They discovered that the study's respondents had a mixed opinion of libraries' digitalization skills.

Jaswal (2016) provided a thorough examination of the digitization process, its issues, challenges, and constraints, highlighting how libraries and documentary heritage in Pakistan are being neglected and threatened with extinction as a result of environmental, economic, and political conditions, as well as human threats such as terrorism. She provided some practical ideas and solutions to overcome the challenges and bottlenecks of digitization work in an environment where archival digitization is not a typical practice.

Numerous studies have been conducted on various elements of archives, not just in Pakistan but all throughout the world, like Ahmad's (2008) investigation into the topic of library and archive records.

The scarcity of information in the country or the government's digitalization efforts is a good study in the literature review. There is no study that can tell us about the current practices regarding the digitization process in Punjabi archives and/or whether they are ready to launch these services to meet the needs of users in the evolving digital age. There is an urgent need for research to better understand the local area and develop local programmes.

## Research Design and Methodology

A cost-effective method was adopted for the collection and analysis of data as part of the data collection strategy. This study uses a qualitative research methodology in order to achieve the study's goals. The purpose of this study was to learn more about the digitization of historical materials as well as identify the process and obstacles faced by personnel during document digitization. The Punjab Archives provides access to all of the resources that are available within the Punjab Archives, including a portion of this population as well. A qualitative approach was deemed to be the most effective method for gathering data and in order to gain a better understanding of the study questions. In order to obtain in-depth qualitative data for this study, an open-ended questionnaire was used along with observational methods to collect data. For the purpose of analysis, the data was codified and made it possible to emerge categories, which, in turn, led to the development of themes as a result of the coding process.

The data were converted into written transcripts, and with the help of NVivo software, the data analysis was performed. During this conversion, the data were also cleaned, and such communication which was not related to the topic or research was not converted into a written transcript. An inductive approach was used to collect the data.

## Data Analysis and Discussion

In the analysis of this study, the respondents know that the digitization process has some requirements, equipment, steps, and challenges during the digitization process.

## Digitization Process of the Punjab Archives

The digitization process is also known as scanning and converting archival records into digital format. The archive records include letters, reports, maps, paintings, etc. The digitization process may also be accepted by taking digital images of the source records, where appropriate.

## Selected Collection for Digitization Process

As part of the three-year collaboration between PITB and the Archives and Libraries Wing of S&GAD, it is advised that the portion of the collection that has undergone some basic cataloguing should be digitized. The reason was the lack of high-level archival expertise to appropriately handle the collection's more vulnerable components because a project of this size has never been attempted in Pakistan. Operating with a low-risk level and a zero-tolerance policy for document damage over the first three years is crucial.

The following collections will be used to choose the 500,000 pages that must be saved, catalogued, and digitalized:

**Table 1**

*Selected collection for digitization*

S. No	Statement	Type of material	Quantity of Documents	Duration	Language
1	Mughal Records	Letters	82	1629-1857	Persian
2	Agencies Record	Letters	22214	1804-1849	English
3	Revenue Department	Letters	19500	1849-1869	English
4	Home Department	Letters	17552	1849-1869	English
5	Lahore Residency	Letters	1900	1846-1848	English
6	Political Department	Letters	18190	1849-1869	English
7	Judicial Record	Letters	16480	1849-1869	English
8	Military record	Letters	2544	1849-1869	English
9	Mutiny Record (Independence War)	Letters	1614	1857-1858	English

## Digitization Process Step by Step

All the data approved by the Research Associates is stored in external drives and transferred to the servers



and (storage area network) SAN storage at the PITB Data Centre. Punjab Archives and Libraries Wing is the custodian of the data, which is stored in external hard drives at their premises. Data back-ups are made on two mediums: tape drives and SAN storage drives, which are kept at the Government of Punjab Data Centre. A diagram of the workflow is appended below:

The digitization process included record keeping, preservation, scanning, quality assurance, bibliography, research communication, physical verification, uploading, cataloguing, and database.

### Record Keeping

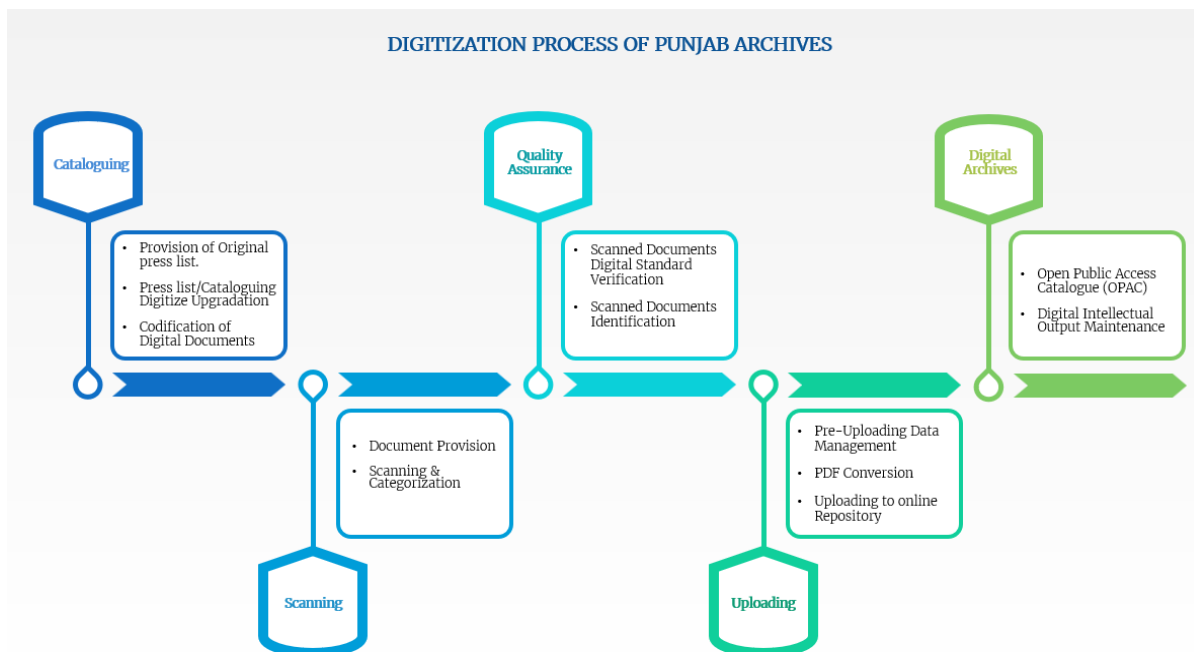
The digitization process starts with manual record-keeping. Some of the basic duties of the records keeper are given below. (1) The record keeper is the first block in the chain and provides material on demand from the preservation. (2) The actual maintenance of documents on a manual basis, from caring for the archival record to protection from damage and loss, is the responsibility of the Record Keeper. (3) Any document that moves from the Preservation Officer, Digital Imaging Officer, or Bibliographer towards the Database is directly returned to the hands of the Record Keeper. After sorting and cleaning the document, the Record Keeper puts that document in the proper place. (4) The Record Keeper is responsible for collecting the record for scanning, etc. and returning it after the completion of the process to the Concerned Officer of the Archives Wing. This is a cyclical thing that goes on in the digitization of archives' records.

### Preservation Process

After the document is handed over, it arrives in the hands of the Preservation Officer. The major responsibility of the Preservation Officer is to decide whether the archival material under consideration for scanning is good enough to be scanned. He/she has to check the archival record with respect to the conservation, preservation and restoration. If records need to be processed or updated, the preservation officer shall create a record of the situation in advance as required or ordered and send it to the Conservation Laboratory for the above work. Likewise, the preservation officer is responsible for training the archive team in the processing of archive records. Therefore, the Preservation Officer must coordinate with other partners when necessary.

### Scanning Process

After the data is recorded and edited from all angles, it reaches the digital image officer. Its main function is to scan files and convert them into TIFF, PDF, JPEG, etc. format.



### Bibliography

The documents gradually reach the bibliographers after scanning. Bibliographers reserve documents by giving them special names. Some of their responsibilities include Creating digital document forms and



accessing all records, numbers, dates, usage dates and numbers, locations, languages and identification codes in the system to store, analyze and create content from important data for the best search results.

### **Cataloguing Process**

The digitization process starts with the cataloguing of the record. Files are obtained from the archive's repository. The project staff fills the data fields in Excel sheets for each file/document. This data set is prepared by keeping into account the software so that all the fields related to online search and web development should be covered.

### **Physical Verification**

In this stage, the record keepers, in collaboration with the cataloguing team representative, checked the issued record by matching it with the complied press lists of the department.

### **Document Identification**

Documents after the PDF conversion are sent for folding. The PDF files are classified according to the archive's press lists and cataloguing data prepared by the cataloguing team. It's the final stage of quality control, where the staff also looks into any discrepancy in the DDC or scanned file. The DDC is checked again to ensure that scanned files and PDF conversions are clear for uploading to the online server.

### **PDF Conversion**

From the combined internal storage, the scanned documents are converted into PDFs. The conversion helps to minimize the size and save files in accordance with the digital document code.

### **Uploading Process**

In this stage, the uploading team, through a special login, access the documents in the internal storage system and uploads the document through the form prepared by the software team. In each uploading, a file is saved by uploading the relevant PDF, and data is obtained through the cataloguing team.

### **Database**

All the data is automatically saved in the assigned database server. The database server is carefully preserved and secured by 24/7 proper air conditioning and a configured entrance at PITB.

### **Techniques Used for Digitization Process**

According to their respondents, the digitization process is used software, staff hiring, scanning resolution, and formats for the digitization process of Punjab Archives,

### **Software used for the Digitization Process**

Punjab Archives employs customized software built in the PHP CodeIgniter framework as a backend and HTML, Bootstrap, and CSS as a front end, with MySQL as the database. This software is used to upload and download digital repositories with online cataloguing. All scanned documents are uploaded, and their cataloguing is updated through a DDC (Digital Document Code).

### **Staff for Digitization Process**

The Punjab Archives hired staff for digitization projects in collaboration with PITB. PITB is managing and leading this team. The project team is controlled by the project director. Project staff works in two teams: the IT team and the technical team. The IT team in PITB reports to the Senior Program Manager, and the technical team in the Punjab Archives Department reports to the Focal Person in the Punjab Archives Department.

The staff of the IT team are software engineers, network administrators, quality assurance engineers, graphic designers, and web developers. The technical staff includes research associates, bibliographers,



research and communication officers, digital imaging officers, preservation officers, record keepers, and data entry operators.

### Resolution used for Scanning in the Digitization Process

Archives followed international standards for scanning resolution, a mixed trend in terms of the resolution used for image-based digitization of documents. The frequency was 300 dpi (dot per inch) used for letters, reports, books, proceedings, and other hand-written and printed documents, and 600 dpi resolutions, used for maps, newspapers, and pictures. The pattern explains the findings that the Punjab archive was digitizing the records at good quality output.

### Formats Used for the Digitization Process

For the process of digitization, TIFF (Tagged Image File Format), JPEG (Joint Photographic Experts Group) and PDF (Portable Document Format) were used as digital file formats and for digital preservation. The TIFF format is used for the digitization of pictures and maps, the JPEG format is used for documents, and the PDF format is used for delivering files.

### Challenges Faced During Digitization Process

The digitization project team faced challenges at different moments during the digitization process. The Punjab Archives were faced with change management issues, security concerns, intellectual property issues, website issues, and time delays (the COVID-19 pandemic) during the digitization process.

### Results and Discussions

Over seven million archives are available in different forms (documents, manuscripts, government records, coins, stamps, lithography, books, pictures, maps, etc.) in Punjab Archives. One hundred fourteen thousand five hundred ninety-two documents have been selected for the initial phase of digitization. Punjab Archives developed custom-based software for digital repositories. The purpose of the software was to catalogue documents and upload them, including all record backups in the PITB. Twenty-four persons were hired for the project and divided into two parts: the technical team and the IT team. Punjab archives followed international standards for scanning and resolution of documents. TIFF, JPEG, and PDF formats were used for digital preservation and uploading in software.

The digitization process worked in the cycle process in different steps: scanning, folding, PDF, cataloguing, and uploading and updating the database. Scanned documents are converted to PDF format against the DDC code. The cataloguing team made a DDC and a PDF and placed them in the appropriate folder. PDF files were imported into the software. All data was automatically saved on the assigned database server. The database server is delicately reserved and secured by 24/7 proper air conditioning and configuring entrance at PITB. The Punjab Archives faced Change management issues, security concerns, intellectual property issues, website issues, and time delay (pandemic COVID-19) problems during the digitization process.

### Conclusion

In conclusion, the digitization process of archives represents a transformative approach to preserve and provide access to our cultural heritage. It offers numerous benefits, including improved accessibility, enhanced searchability, and increased security. However, it requires meticulous planning, technical proficiency, and ongoing preservation efforts to overcome challenges and ensure the long-term sustainability of digitized archives. Collaboration among archivists, information technology specialists, and stakeholders is vital to successfully navigate the complexities of digitization and preserve our invaluable historical records for future generations.

### Recommendations

Based on the case study of the Punjab Archives digitization process, outreach and education efforts should be undertaken to promote the digitized archives and educate users on how to utilize them effectively. In order to attract online traffic from education institutes and research centres around the world, the budget should be allocated for online marketing promotions. Collaboration and partnerships with other

institutions or organizations can bring valuable expertise and resources to the digitization process. Moreover, Continuous evaluation and improvement are necessary to ensure the digitization process remains effective and up-to-date. Regular assessments, user feedback, and keeping abreast of technological advancements will help refine the digitization process and enhance the accessibility of the archives.

## References

- Ahmad, S. (2008). Research study on Library and Archival Records in the Directorate of Sindh Archives. Karachi: Unpublished PhD Thesis.
- Alhaji, U. I. (2007). Digitization of past question papers, dissertations and theses: A case study of 30 Nigerian university libraries. *The International Information & Library Review*, 39(3-4), 228-246. <https://doi.org/10.1016/j.iilr.2007.07.002>
- Amalia, A. T., & Panduwinata, L. F. (2022). Sistem Informasi Manajemen Arsip Elektronik (E-Arsip) Berbasis Microsoft Access Terhadap Efektivitas Penemuan Kembali Arsip Pada SMKN 4 Surabaya. *Jurnal Pendidikan Administrasi Perkantoran (JPAP)*, 10(3), 195-210. <https://doi.org/10.26740/jpap.v10n3.p195-210>
- Archives, P. (2021, September). Punjab Archives and Libraries Wing, S & GAD. <https://archives.punjab.gov.pk/>
- Bültmann, B., Hardy, R., Muir, A., & Wictor, C. (2006). Digitized content in the UK research Library and Archives sector. *Journal of Librarianship and Information Science*, 38(2), 105-122. <https://doi.org/10.1177/0961000606063893>
- Chughtai, A. (2019). The Project of Digitization, Preservation and Modernization of Punjab Archives and Libraries Wing. Lahore: Punjab Information Technology Board.
- Jaswal, B. A. (2016). Digitization of archival collections by libraries in Pakistan: Issues, strategies, challenges and opportunities. *Pakistan Journal of Information Management and Libraries*, 17, 97-108. <https://doi.org/10.47657/201617897>
- Khadjah, U. L., Perdana, F., Kirana Sarasvathi, D. G., & Winoto, Y. (2021). Proses digitalisasi naskah kuno sebagai pelestarian informasi Di Museum bandar Cimanuk Indramayu. *Pustaka Karya : Jurnal Ilmiah Ilmu Perpustakaan dan Informasi*, 9(1), 49. <https://doi.org/10.18592/pk.v9i1.5167>
- Mariner, M. C. (2014). Managing Digital Audiovisual Resources: A Practical Guide for Librarians. Rowman & Littlefield, Lanham, MD. <https://rowman.com/ISBN/9780810891036/Managing-Digital-Audiovisual-Resources-A-Practical-Guide-for-Librarians>
- Mat, S. Y., & Kolej, A. M. (2005). Digitization and sustainability of local collection: an observation of digitization activities among universities libraries. In World Library and Information Congress: 71st IFLA General Conference and Council "Libraries a voyage of discovery". Norway: Oslo. <https://api.semanticscholar.org/CorpusID:107844529>
- McMenemy, D., & Poulter, A. (2005). Delivering digital services: a handbook for public libraries and learning centres. London: Facet. <https://pureportal.strath.ac.uk/en/publications/delivering-digital-services-a-handbook-for-public-libraries-and-1>
- Niazi, L. A. (2008). An introduction to the Punjab Government archives: Tomb of Anarkali. Lahore: Sang-e-Meel Publications. [https://books.google.com.pk/books/about/An\\_Introduction\\_to\\_the\\_Punjab\\_Government.html?id=BeovngAACAAJ&redir\\_esc=y](https://books.google.com.pk/books/about/An_Introduction_to_the_Punjab_Government.html?id=BeovngAACAAJ&redir_esc=y)
- Rafiq, M., & Ameen, K. (2014). Towards a digitization framework: Pakistani perspective. *Pakistan Journal of Information Management and Libraries*, 15, 22-29. <https://doi.org/10.47657/201415757>
- Rozana, L., & Musfikar, R. (2020). Analisis Dan perancangan sistem informasi pengarsipan Surat berbasis web pada Kantor lurah Desa dayah tuha. *Cyberspace: Jurnal Pendidikan Teknologi Informasi*, 4(1), 14. <https://doi.org/10.22373/cj.v4i1.6933>
- Ulyatt, M. (2014). The John F. Kennedy Presidential Library and Museum is a cultural representation of the public memory of the president. *European Journal of American Culture*, 33(2), 117-130. [https://doi.org/10.1386/ejac.33.2.117\\_1](https://doi.org/10.1386/ejac.33.2.117_1)
- Wirasaputra, A., Riduan, F., & Noviana, W. (2022). Dampak Dari Perkembangan Teknologi Informasi Dan Komunikasi. *Jurnal Kreativitas Mahasiswa Informatika*, 3(2), 206-210.