

APPLICATION OF WEB ARCHIVING TECHNOLOGIES IN BNL AND NAB: A PROPOSED MODEL

Md. Rifat Mahmud¹
Raiyan Bin Reza²

Abstract

Web archiving has become a regular activity in many libraries and archival institutions. With the massive spread of internet, preserving the web contents are now being given importance by various countries as the websites contain various important legal, political, educational information. This paper investigates the issues related to web archiving that might be faced by the Bangladesh National Library (BNL) and National Archives of Bangladesh (NAB). The main aim of this paper is to describe current state of web archiving in BNL and NAB. This paper also tried to explore identifying the problems in archiving web contents and proving possible solutions to overcome the problems. Interview method was applied for this study. We interviewed officials from both BNL and NAB and explored relevant literatures to gather information for our work. Web archiving activities are found to be useful in many government libraries and archival institutions around the globe but it is yet to be done in BNL and NAB. The study found that there are many challenges for implementing web archiving in BNL and NAB such as technological difficulties, copyright issues, unskilled manpower, lack of logistical support, etc. should be taken into account while implementing any web archiving programme. Sufficient steps like proper planning, efficient training, logistical support, international cooperation and adequate financial support will help the authorities to establish a successful web archiving programme. Finally, we proposed an intuitive model for NAB and BNL so that it could be considered while taking any web archiving initiative.

Keywords

Web archiving, Bangladesh National Library, National Archives of Bangladesh, Website, Archiving, Bangladesh

¹Postgraduate student, University of Dhaka, Email: rifatmahmud110@gmail.com

²Postgraduate student, University of Dhaka, Email: raiyaandip96@gmail.com

Introduction

Internet has revolutionized the world with billions of active internet users. With more than 4.57 billion internet users worldwide (Statista.com), the internet plays an increasingly important role in people's daily activities. Nowadays, people live with the internet. Websites are integral part of internet. Almost 1.72 billion websites are there (Statista.com). Websites contain tons of information on various issues. Websites are generally very dynamic in nature. They change very quickly whenever any new information is found. It may be necessary for various users to get the previous versions. A user may be in need of both new content and older content (Teevan, 2008). Though web browsers provide features designed to support retrieval, such as bookmarks and history, as do analysis frameworks such as "Things I've Seen" (Dumais et al.,2003), such tools are not designed to manage circumstances where the corpus of information might have changed. The websites containing important information must be captured in real time to ensure that these contents survive for the next generation. Historian Ian Milligan realized the importance of web archiving for preserving important political and cultural events (SSRC.org). The web is changing our way to look at the past events. Historians are looking at archived webpages for their research work more than ever. It is now impossible to write about the events of 1990s without looking at web pages. Websites now contain vast amount of cultural, social and political data (Milligan, 2016). Various web archiving initiatives are going on all over the world to archive various parts of websites (aub.edu.lb).

The importance of archiving web contents is now being realized all over the world. Many researchers of various fields are now in need of archived web pages for their research work. A wide range of institutions are now being involved in archiving web contents. Many national institutions and universities around the world are taking on web archiving programmes for meeting various needs. Even many repositories now have started web archiving. Even some institutions are saving twitter data (Weber, 2019).

Web archiving is the process of gathering up data that has been recorded on the World Wide Web, storing it, ensuring the data is preserved in an archive, and making the collected data available for future research (Niu, 2012). According to International Internet Preservation Consortium (IIPC), web archiving is the process of collecting portions of the World Wide Web, preserving the collections in an archival format, and then serving the archives for access and use. According to Library of Congress, web archiving is the process of collecting documents from the Internet and bringing them under local control for the purpose of preserving the documents in an archive. The act of running web archiving software is often known as "harvesting" or "crawling." Web archivists use various software tools to determine the scope of a distinct web archiving

activity and use web crawlers to pull web resources down to local storage. The terms Web Capture, Web harvesting, Web archiving, and Web collecting are all used synonymously.

In 1996, the internet archive and many national libraries started web archiving activities. The International Web Archiving Workshop (IWA), launched in 2001, has provided a forum for the exchange of experiences and ideas. Then establishment of the International Internet Preservation Consortium (IIPC), in 2003, greatly facilitated international cooperation in the production of standards and open-source resources for web archive construction. There are various integrated systems, third party services and commercial institutions offering web archiving services (Dpconline.com). The journey of internet in Bangladesh was started in the 1990s (Azam, 2007). The scope of internet is expanding day by day in our country. According to Bangladesh Telecommunication Regulatory Commission (BTRC), there are 103.253 million subscribed internet users in Bangladesh. There are numerous government and non-government websites in Bangladesh. These websites may contain information regarding government policies, administration, economics, politics, commerce, science & technology and other necessary fields. Archiving these websites may help the government as well as citizens in the long run. Various national libraries and archives around the world are taking part in various web archiving initiatives to preserve the websites containing various important information regarding their respect (IIPC, 2020). In this paper, we will concentrate on the exploration of web archiving issues and challenges of both Bangladesh National Library and National Archives of Bangladesh.

Literature review

Web archiving is a recent field of study which has emerged at the end of the 20th century with the aim of avoiding a digital black hole in world history. The increasing publication of born digital information through the World Wide Web (WWW) and the international recognition of such publications as part of the heritage have increased the value of archiving web content. A survey was conducted on web archiving initiatives in the developed world. It aims at knowing the state of web archiving initiatives. In these countries, web pages which are known to have importance are archived. They archive a huge amount of content and the whole initiative is managed by a small team. Many initiatives lack effective search mechanisms which often hamper access to archived contents. Various legal barriers make access to archived content difficult. Open access to archival information is still a challenge for many (Gomes et al., 2011). In Asia only four countries are doing national web archiving activities in their respective national libraries on a large scale. These are South Korea, Japan, China and Singapore. Web archiving is now being taken seriously in Sarawak state of Malaysia.

In this state, the state library has a web archiving project named: “Sarawak State Web Archive (SSWA)”. This initiative archives websites published by various state and government institutions. It also arranges many improvement programmes to identify issues and provide suitable solution to the issues. It has issues like legal issues, authorization, preservation, insufficient and unskilled manpower. It also faces technological and storage related problems. But a project has now been taken to tackle the issues and challenges faced by the archiving initiatives. In spite of all the challenges, it has been able to achieve some great milestones. It may not only help in research work but also in preserving web content of various importance (Jamain et al., 2018).

Library and information professionals face various legal and moral issues regarding web archiving. In the USA, archival institutions have legal protections against various lawsuits. But these protections are very limited in many countries. Librarians and information professionals deal with morally complex issues regarding breaking legal barriers for common good. In such conditions, African Studies Association came up with guidelines like:

- a. Not harming other
- b. Disclosing full information regarding objectives, funding, methods and outcome
- c. Reciprocity and equity
- d. Deposition of data and publications (Lor and Britz, 2004)

Another significant problem that may arise in case of web archiving is self- censorship, through which control over intellectual production on the web can be exercised. In some circumstances, web archives are research projects that involve highly sensitive issues such as the creation of dark web archives for terrorism analysis (Cadavid et al., 2013). Copyright and privacy are both inherent issues for web archiving. Many countries have now enacted legislations in favour of archiving web contents. (Schiphof and Beunen, 2006). But the National library of Australia and Netherlands are still grappling with these issues in order to making materials available to public. The legal protection and reduced complexities may speed up the activity of preserving the nation's digital heritage (Glanvile, 2010). Many researchers from social science and humanities background are still unaware of web archives (Costea, 2018). They do not know that web archives may be beneficial to their research projects. But the value of archiving web contents and web pages has been emphasized by both users and nonusers. The issue of legal and technical difficulties is found everywhere. There is an intense need for proper documentation of web contents. It will maintain the validity and trustworthiness of the archived web content. It may be helpful if the users can use

multiple archives at the same time. The availability of different files in the same archives will be good for some researchers (Costea, 2018).

Different web archive collections may function differently. Every web archive collection may have different technological workflow and curation policies. Some key factors may help to understand this phenomenon like mandating and motivating, various technical settings and crawling software. The organizational and technical contexts are necessary for understanding the performance of web archiving process. A framework may help a lot in web archiving for libraries and archiving institutions (Maemura et al., 2018). Web archiving initiatives follow a wide range of activities like selection, acquisition, organization and storage, access etc. All of these activities are influenced largely by various external and internal factors. The external factors may include legal barriers or relationship between content creators and archiving agency. The internal factors depend on the nature of the contents, the nature and capacity of the archival institution. The professionals also need various specialized skills to carry out the activities necessary for web archiving (Niu, 2012).

BNL and NAB-Profile

BNL and NAB were established in 1972. At first, they were functioning under the ministry of education. Now, they work under the ministry of culture. The whole organization consists of 98 staffs. Both the archives and the library are supervised by two separate directors. There are two separate buildings. Both buildings are situated at Sher-E-Bangla Nagar, Dhaka. Both the library and archives contain valuable resources.

BNL: BNL collects all of the country's published materials under the law of the country. It has a rich collection of books. The collection of the Library is the asset of the nation. It covers all major disciplines and branches of human knowledge. In addition to all new local publications each year, the library also acquires foreign books, particularly those on South Asian issues. BNL enhances its collections in different ways, such as by acquiring copyright publications, purchases, institutional / personal donations, complimentary copies, etc.

NAB: The NAB contains important documents like parliament Proceedings and Records, Cabinet Division records, Dhaka Divisional Commissioner's Records, Chittagong Divisional Commissioner's Records, District Records, Collectorate Records, Government Bengal, East Bengal and East Pakistan proceedings/files, Dhaka City Corporation Records, Zilla Parishad (District Board) Records, Narayanganj Municipality Records, newspapers, old maps, press clippings etc. The national library contains books on almost all branches of knowledge, newspapers, periodicals, maps,

microfilms, microfiches etc. NAB launched a five-year (2002-2007) digitization project in 2002 which was called "Digitization of District Records Collected from 1760-1900" (Shuva, 2009). The main objective of the project was to convert traditional district resources into a digital format. The project also aimed to develop a strategy and methodology for the permanent digital archiving of electronic versions of the main resources of the center. However, due to a gap in proper planning and other constraints, the project has not been successful. A further project (2009-2012) named "Collection and scientific preservation of obsolete records" was an effort to provide online access to NAB's digital archives along with the digitization of documents activities.

Web archiving around the globe

In the International Internet Preservation Consortium (IIPC), there are 57 institutions from different countries and continents. Most of the institutions are from Europe. In Asia only few institutions are the members of IIPC. Egypt is the only country in the African continent to be enlisted as IIPC member.

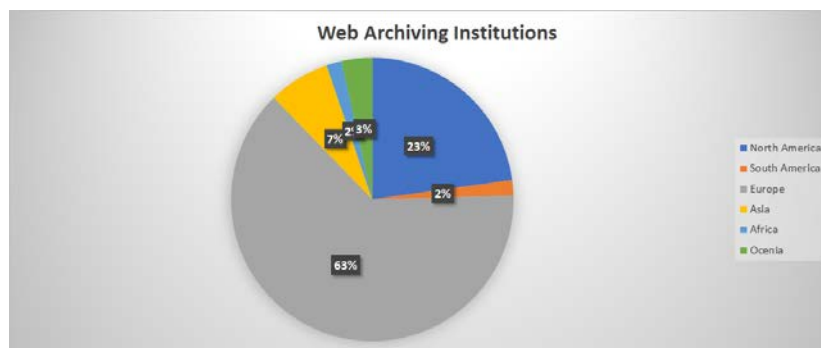


Figure 1: Web archiving around the globe

Source: From the data found at <https://netpreserve.org/about-us/members/>

Objective of the study

Almost all the government departments have websites containing valuable information. There is need for web archiving initiative to preserve such information. The objectives of the study are to:

- a. Explore the potentials to implement the web archiving in BNL and NAB.
- b. Identify the issues hindering web archiving activities.

- c. Provide possible solutions to overcome the problems.
- d. Propose a web archiving model for BNL and NAB.

Research Questions

We have also formulated two important research questions to address the objectives. They are:

RQ1: What are the issues hindering web archiving activities in BNL and NAB?

RQ2: What is the possible web archiving model that can be followed in BNL and NAB?

Methodology

This study was based on interview method. We interviewed 10 high ranked officers of BNL and NAB including deputy directors to gain knowledge regarding the activities related to web archiving. In this study, we also reviewed various scholarly articles and websites regarding web archiving. The websites of BNL and NAB provided us a plenty of information. We also tried to gather various data from the research works related to NAB and BNL. The interview was a semi structured interview. The interview was taken via telephone. We asked some questions regarding the present condition, issues regarding web archiving, providing possible solutions to the problems and future possibilities and plans regarding web archiving. The sample questions are following:

Are there Web archiving activities in BNL and NAB?

Are there Plans regarding web archiving?

Will there be training programs?

Are there necessary personnel?

Do BNL and NAB have technical facilities?

Do BNL and NAB have logistical support?

Do BNL and NAB have legal protection against web archiving related cases?

Do BNL and NAB have integrated system for web archiving?

Do BNL and NAB has financial support?

Data analysis of the study

Unfortunately, from the interview we found the web archiving activities are yet to be introduced to both the BNL and the NAB. The digitization of materials is still in action. The collection of online versions of newspapers and building a strong oral archive are ongoing projects. Still, there is no plan by the directory of archives and libraries to implement web archiving project. Many libraries and archives around the world are facing various issues regarding archiving web contents. Both BNL and NAB will face various issues while implementing a web archiving project. These issues need to be addressed clearly and the authorities should find efficient solutions to all these issues.

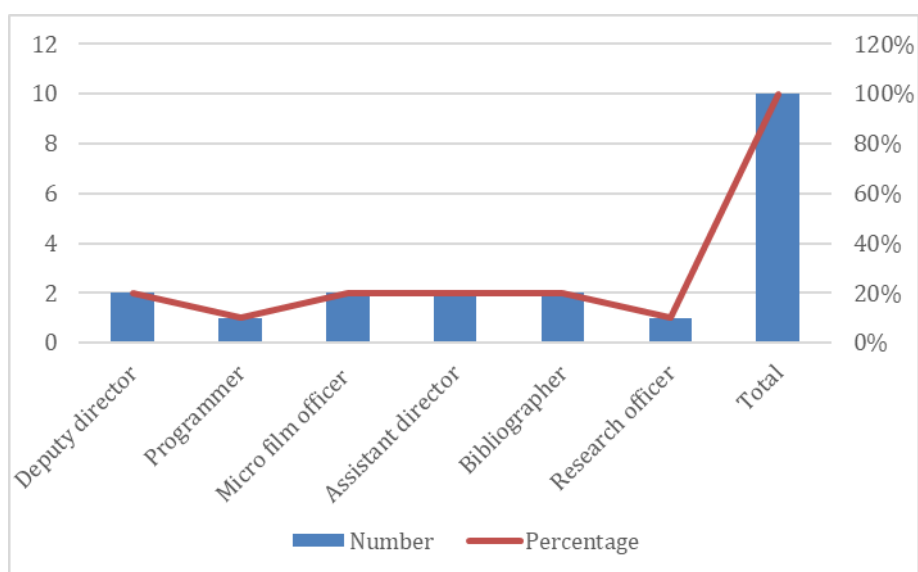


Figure 2: Interviews from BNL and NAB

Web Archiving Issues

There are many issues that need to be fixed whenever the plan is taken by the authorities.

i. Technological difficulties:

Web archiving involves archiving of multimedia products such as, text, sound, graphics, pictures, photographs, video clips etc. which require intensive use of bandwidth which is actually difficult for developing countries like Bangladesh. Moreover, pictures, graphics and video files are really very space-hungry.

Besides, the rapid growth of the new technologies requires urgent action to be taken in the archiving of web content. Electronic material is continuously being updated, deleted or permanently lost and cannot be restored when the website is updated. In these cases, web content will be unavailable. These issues tend to create technological difficulties in implementing web archiving activities. There are also various problems regarding the web crawlers. The web crawlers find it hard to archive deep web materials (Pennock, 2011).

ii. Copyright issues

Many countries protect their libraries and archives legally whenever they face any issue regarding copyright issues. These institutions will remain vulnerable to legal claims. It will be a huge challenge for both BNL and NAB to maintain this issue. A clear framework is always needed to tackle various legal difficulties caused by copyright claims. Some flexibility in the copyright law will allow these institutions to carry out web archiving with ease. In Britain many librarians are supportive of laws to support web archiving activities (Brindley, 2011).

iii. Access issue

NAB & BNL may often deal with the difficulty of giving access to archived web contents. It will not be an easy issue to fix. There will be a need for introducing necessary tools to provide easy access to the web archive. This issue may cause various difficulties if it is not dealt in a proper manner.

iv. Limited and unskilled personnel

According to the interview, there is a smaller number of personnel than necessary in BNL and NAB. Moreover, many of them are not skilled enough. This is also an obstacle to implement the web archiving project in Bangladesh. Unskilled and insufficient number of personnel will hinder any organizational work. Web archiving requires both technological knowledge and expertise. Efficient skill and sufficient number of personnel will be able to perform the tasks related to web archiving in a desired way.

v. Difficulty of integrating various functions

Web archiving requires a wide range of activities like assessing, processing, storing and disseminating. These processes require effective integration. All of these activities need to be planned thoroughly to ensure a good service. Otherwise, it will be difficult to perform the tasks related to web archiving. Finding the right integrated system will support the multifaceted tasks with comfort.

vi. Storage and backup

One of the key problems is the shortage of storage capacity. Archiving of web content will need a huge amount of space. Besides a strong backup and recovery system is needed for web archiving activities. As websites contain various types of information, any storage system should support the activities in a proper manner.

vii. Financial challenges

In any project, financial factors play a crucial role. For web archiving activities we need certain necessary things like training programs, web crawler, storage facilities etc. All these components will require essential financial support to function properly. If the authorities fail to provide necessary funds, the project will face problems.

Benefits of web archiving

Web archiving offers tremendous benefits for everyone. Some of them are outlined in the following:

Preserving important websites

Websites contain valuable information in this age. But they are also very volatile in nature. Archiving the web contents will help to keep track of the changes made in the website. It will also help us to record all the previous information before getting lost.

Cultural heritage

In the past books were regarded a valuable part of the national cultural heritage. Now websites are now providing valuable cultural information. So, it is necessary to keep track of the regular updates of the websites. Web archiving helps a lot in these cases (Pennock, 2016).

Administrative record

Now various governments publish valuable documents and information in online platforms. It is necessary to preserve these information for various purposes. In UK and Ireland, it is obligatory for the archival institutions to archive various government websites (Pennock, 2016).

Academic research

Websites now play a huge role in academic research. Researchers cite various websites for their resources. Archiving the websites will help the researchers.

Historical research

Web archives have greater amount of information for research than modern print media files. They keep diverse types of information which support a wide range of use cases for historical research.

Organizational information

Websites contain certain information about organizations which cannot be found elsewhere. It is necessary to archive these webpages (The national archives, 2011).

Data mining and searching

The automatic extraction of knowledge from huge amounts of data is known as data mining. Data mining functionalities are required for web archives. Web archives offer tools and services to make mining their archived data easier. They provide the service that allows the automatic search and processing of huge number of files. Searching allows users the opportunity to easily access large quantities of knowledge. Most of the web archives offer web address (URL) search as a way for users to access archived materials.

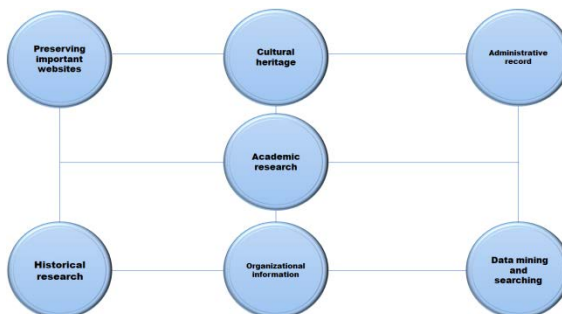


Figure 3: Web archiving benefits

Web archiving process

Any web archiving process must be done with a solid policy. It will help to determine a criterion for archiving web contents. The websites that will be archived must be well identified. The websites will be captured with a help of a web crawler. The web crawler set up process should be done keeping in mind about the time and depth of the whole

procedure (Diana and Amran, 2019). The downloaded items must be verified after crawling. They should be given proper metadata for easy searching. Then they should be stored in the web server.

Proposed model for web archiving

We have proposed a model for web archiving called PSCVNS. We have created this model with the help of other ongoing models around the world (Diana and Amran,2019; Maemura,2018). In our proposed PSCVNS model there are six important functions. They are:

- i. **Policy:** In this phase, a guideline should be formulated for web archiving. It will entail determining the websites for archiving.
- ii. **Selection:** Websites that will be archived should be found and selected timely for the archival process.
- iii. **Capturing:** Web archiving needs a web crawling software. Web crawling software will download the web pages.
- iv. **Verifying:** Verification should be done to check for any mistakes. It will ensure the quality for the overall process.
- v. **Naming:** Providing proper metadata for the archived contents is a must. It should be done carefully.
- vi. **Storing:** The web contents should be stored for preservation. A good storage facility will ensure the safety of the web contents.

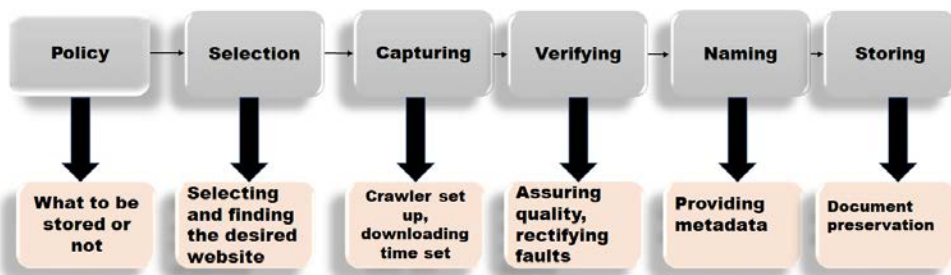


Figure 4: PSCVNS model for web archiving

Recommendations

Proper planning

Any successful project needs a proper plan. The national archives and the national library need a solid plan to match the requirements of web archiving activities. The Bangladesh National Library and National Archives of Bangladesh should adopt a transparent and open framework regarding web archiving, making available their policies. This framework must predict possible issues and should be revised in order to making solutions to the issues.

Training

Many archivists and librarians may lack knowledge regarding web archiving. A sophisticated training programme will provide the skills and knowledge to undertake web archiving activities. The training programme must address the technical issues properly. There might be a need for the continuous training for the personnel for newly emerging issues.

Logistical support

Web archiving initiatives require sufficient tools like proper software and efficient storage facilities. The authorities must take of these facilities to run the project successfully. A good integrated system like PANDORA, Web curator tools can be considered. We may need specialized software for accessing, searching and crawling. There are many open-source software to support these activities (Pennock, 2011).

International cooperation

There are various international institutions working on web archiving like International Internet Preservation Consortium (IIPC). Both the library and the archives must enlist themselves over there. They can also make partnerships with other national libraries and national archives. It will help them to identify common problems and to provide possible solutions to the problems.

Financial support

Adequate financial support should be provided by the government for the implementation and maintenance of web archiving project. Sufficient budget will help a lot in archiving websites.

Sufficient storage facilities

For ensuring that the web archiving operation runs smoothly a server with a huge space capacity must be given. A strong backup and recovery system should also be ensured for the web archiving purpose. The storage system must be built for long periods. Upgrading the storage facilities should be considered because websites are regularly embracing new markup languages and contents (Jackson, 2012).

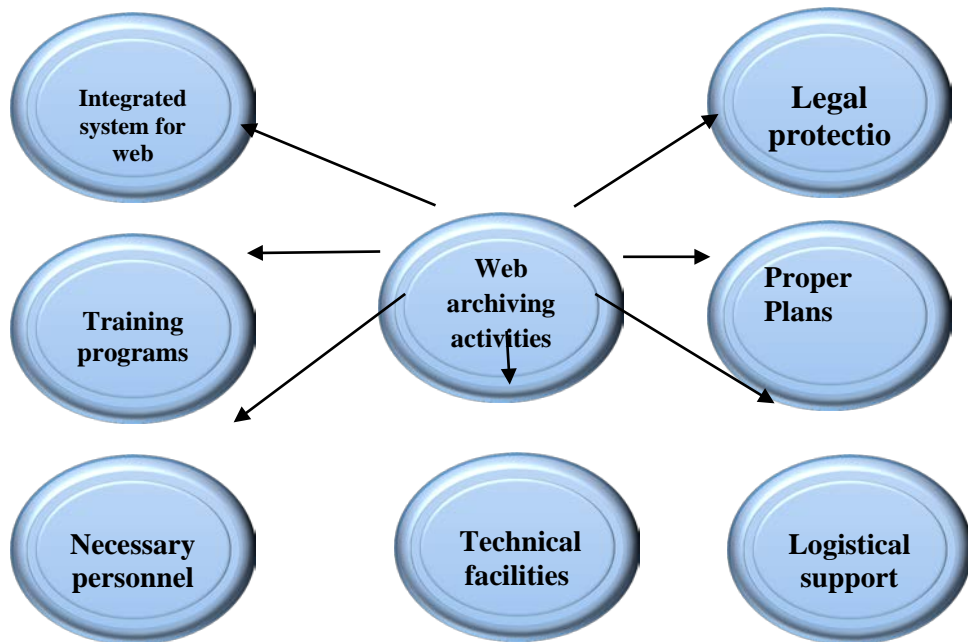


Figure 5: Components needed for web archiving in BNL and NAB

Limitations of the study

In our work, we could not find any previous literature dealing with web archiving in Bangladesh. We also could not find literature dealing with web archiving in South Asian countries. There were also time constraints and various difficulties posed by COVID -19 restrictions.

Conclusion and future research

This paper tried to show the issues faced by the BNL and NAB in implementing web archiving project. Web archiving is getting momentum in libraries and archives all around the world. Web archiving initiatives are mostly seen in the developed countries.

In a developing country like Bangladesh, web archiving initiatives will face multifaceted problems and issues. But web archiving initiatives may be taken in the near future to preserve web contents related to national importance. A smooth plan, solid infrastructure, efficient and skilled manpower will be necessary for any web archiving activities in the national library and the national archives of Bangladesh. The web archiving is relatively a new concept. So, it is understandable that there are still no works related to web archiving in Bangladesh. Further exploration of this field will certainly provide new perspectives. There is also a need for future studies in this topic to provide more in-depth information. Such initiatives will help to provide more insights into the discussions regarding web archiving.

References

- Alvord, T., 2015. Web Archiving. In *Encyclopedia of Information Science and Technology, Third Edition* (pp. 7684-7692). IGI Global.
- Aub.edu.lb.libguides.com. 2020. Libguides: Government Documents: Web Archiving Initiatives. [online] Available at: <<https://aub.edu.lb/libguides.com/c.php?g=276470&p=1843274>> [Accessed 7 September 2020].
- Ayala, B.R. and Caragea, C., 2014. Towards building a collection of web archiving research articles. *Proceedings of the American Society for Information Science and Technology*, 51(1), pp.1-5.
- Azam, S., 2007. Internet adoption and usage in Bangladesh. *Japanese Journal of Administrative Science*, 20(1), pp.43-54.
- Begum, B. and Akter, R., 2016. Open Access Initiatives of the National Archives of Bangladesh (NAB).
- Brindley, L., 2019. British Library International Survey of E-Legal Deposit 2011: Summary of Findings. In CDNL Annual Meeting, Puerto Rico. http://www.cdnl.info/sites/default/files/docs/2011_legaldeposit_survey_Slides_Aug.pdf. Datum pristupa (Vol. 20, No. 10).
- Btrc.gov.bd. 2020. Internet Subscribers in Bangladesh March, 2020. | BTRC. [online] Available at: <<http://www.btrc.gov.bd/content/internet-subscribers-bangladesh-march-2020>> [Accessed 7 September 2020].
- Cadavid, J.A.P., Basha, J.S. and Kaleeswaran, G., 2014. Legal and technical difficulties of web archival in Singapore. *Rev. Prop. Inmaterial*, 18, p.35.
- Costea, M.D., 2018. Report on the scholarly use of web archives. NetLab.
- Dponline., 2020. Web-Archiving - Digital Preservation Handbook. [online] Dponline.org. Available at: <<https://www.dpconline.org/handbook/content-specific-preservation/web-archiving>> [Accessed 7 September 2020].

- Dumais, S., Cutrell, E., Cadiz, J.J., Jancke, G., Sarin, R. and Robbins, D.C., 2003, July. Stuff I've seen: a system for personal information retrieval and re-use. In Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval (pp. 72-79).
- Gomes, D. and Costa, M., 2014. The importance of web archives for humanities. *International Journal of Humanities and Arts Computing*, 8(1), pp.106-123.
- Gradmann, S., Borri, F., Meghini, C. and Schuldt, H., 2011. Research and advanced technology for digital libraries. In Conference proceedings TPDL (p. 177).
- IIPC, A. and members, I., 2020.IIPC Members - IIPC. [online] IIPC. Available at: <<http://netpreserve.org/about-us/members/>> [Accessed 7 September 2020].
- Jackson, A.N., 2012. Formats over time: Exploring UK web history. arXiv preprint arXiv:1210.1714.
- Jamain, J., Yahya, A.L., Muhammad, N. and Abdul Rahman, M.A., 2017. Web archiving issues and challenges in State Government of Sarawak (Malaysia): Do they really need their website to be archived?
- Lor, P. and Britz, J.J., 2004. A moral perspective on South–North web archiving. *Journal of information science*, 30(6), pp.540-549.
- Maemura, E., Worby, N., Milligan, I. and Becker, C., 2018. If these crawls could talk: Studying and documenting web archives provenance. *Journal of the Association for Information Science and Technology*, 69(10), pp.1223-1233.
- Milligan, I., 2016. Lost in the infinite archive: The promise and pitfalls of web archives. *International Journal of Humanities and Arts Computing*, 10(1), pp.78-94.
- Pennock.M 2020.Twittervane: Crowdsourcing Selection - UK Web Archive Blog. [online] Available at: <<https://britishlibrary.typepad.co.uk/webarchive/2011/12/twittervane.html#>> [Accessed 7 September 2020].
- Pennock, M., 2013.Web-archiving.DPC technology watch report, 13, pp.1-45.
- Rahman, M.M., Ahmed, T. and Shirahada, K., 2017. Value Co-Creation in Archival Resources: Exploring the Feature of National Archives of Bangladesh (NAB)'s Open Access Project. *International Journal of Library and Information Services (IJLIS)*, 6(2), pp.37-56.
- Schiphof, T. and Beunen, A., 2006.Legal aspects of web archiving from a Dutch perspective.
- Shuva, N.Z., 2009, January. Creating and managing digital collection at national archives of Bangladesh: a developing country perspective. In Archiving Conference (Vol. 2009, No. 1, pp. 192-196). Society for Imaging Science and Technology.
- Statista. 2020. Digital Users Worldwide 2020 | Statista. [online] Available at: <<https://www.statista.com/statistics/617136/digital-population-worldwide/>> [Accessed 7 September 2020].

- Sumathi, S. and Esakkirajan, S., 2007. Data mining and data warehousing. *Fundamentals of Relational Database Management Systems*, pp.415-475.
- Teevan, J., 2007. "Where'd it go?": How people ask after lost Web information. *Proceedings of the American Society for Information Science and Technology*, 44(1), pp.1-19.
- Weber, M.S., 2020. *The Archived Web: Doing History in the Digital Age*. NielsBrügger. Cambridge, MA: The MIT Press, 2018. 200 pp. \$30.00 (hardcover). (ISBN 9780262039024). *Journal of the Association for Information Science & Technology*, 71(3), pp.374-376.
- Weigle, M., Weigle, M., Sloane, M., Borchert, T., Wambua, M. and Boston, A., 2020. On The Importance of Web Archiving. [online] Available at: <<https://items.ssrc.org/parameters/on-the-importance-of-web-archiving/>> [Accessed 7 September 2020].