

An Exploratory Study of Advantages and Disadvantages of Website Preservation

Rattahpinnusa Haresariu Handisa 

Perpustakaan R.I Ardi Koesoema, Ministry of Environment & Forestry, Indonesia

Paper Type:
Literature Review

Abstract

Background of the study: website preservation should be explored due to several reasons. First, the theme website preservation is classified into a new field so there is a limited number of scientific articles discussing to this topic; Second, website preservation is not an easy task due to its potential challenge. Third, preserving website is an effort to maintain accessibility to beneficial resources for future generations.

Purpose: This paper will explore benefits and drawbacks in preserving websites. It also discuss problem and solution during website preservation process. The result of discussion will benefit for library manager to formulate an accurate policy on how to maintain accessibility of library's website.

Method: Reviewing literature is an analyses method in this paper

Findings: Findings shows that website preservation has a lot of benefits rather than its drawbacks. The website preservation will support humanities researchers to collective memories to special events or importance person in the past; the website preservation assures accessibility to heritage data and it offers opportunity for library to collaborate in developing heritage database.

Conclusion: Website preservation will have a lot of benefits rather than drawbacks. Library will be able to develop collective memory for future generations by preserving website contents. In addition, the website preservation also offers efficiency in collecting and managing metadata due to its advance technology.

Keywords: *exploratory study, website preservation, internet archiving, libraries 4.0, digital archive strategy, accessible website*

Submitted: 19 September 2020

Revised: 15 December 2020

Accepted: 12 May 2021

Online: 29 June 2021

* Correspondence :
Rattahpinnusa Haresariu
Handisa

Email:
rattahpinusa@gmail.com

Introduction

Preserving websites' contents will be a new domain of core business for Gallery, Library and Museum (GLAM) institutions in the future. The website preservation is an integral process of maintaining data, managing recorded data, and ensuring access to preserved data on the internet. That process has a significant contribution to preserve national heritage for future generation (Niu, 2012; Phillips, Alemneh, and Ethiopia, 2015). Furthermore, the importance of preserving websites is considered that websites may contain valuable information. for example, a winning speech of presidential election or an announcement of foreign invasion. That information may become historical manuscripts, and future generation can get benefit from it (Costa, Gomes, and Silva, 2016). By preserving websites, the library could develop collective memory on the community by providing accessible information resources.

Interestingly, a theme of website preservation should be explored due to several reasons. First, the theme website preservation is classified into a new field so there is a limited number of scientific articles discussing to this topic. As cited in wikipedia page, an initial project Website preservation was developed in 2001. The project was called wayback machine whom could view website content (Wikipedia 2021). The project is countinously developed until this present time. Secondly, technology on website has changed gradually over time. An old technology website applied platform web 1.0. Then, technology on website has changed from platform web 2.0 to web 4.0 at this moment. Those change will have impact on how to preserve data on website. Third, website preservationis not an easy task due to its potential challenge. It is predicted that an obsolete technology will be a technical issue on website preservation. In other side, law and regulation also potentially threat on website preservationprocess. Considering those three reasons, library as a part of GLAM institutions should prepare strategies to tackle potential challenges of Website preservation.

This paper will explore benefits and drawbacks in preserving websites. It also discuss problem and solution during website preservation process. The result of discussion will benefit for library manager to formulate an accurate policy on how to maintain accessibility of library's website.

Method

This paper uses a literature review method to explore advantages and disadvatages of website preservation. (Webster, and Watson, 2002) writes a guideline of writing a literature review. That guideline provides seven stages of reviewing literature as follows: Firstly, a reviewer should find relevant articles and the reviewer should focus on a specific topic. Then, the reviewer should cover relevant articles related to that topic. The next stage is structuring the literature review. The structure has two types, namely concept-centric and author-centric. The concept determines the organisation of frameworks or ideas. Meanwhile, the second type of reviewed structure emphasises on the credibility of the author. Selecting appropriate tone and tense is another crucial aspect in the third and fourth stages. The reviewer should select a neutral and informative tone to avoid subjectivity. Similarly, the reviewer should use a present tense to describe a concept because the concept is an ongoing process. Meanwhile, the reviewer should use a past tense when the reviewer cites the authors' views. Developing theory is needed by the reviewer. The reviewer should examine the past research, and the reviewer should predict the future trend of the research. An evaluating literature review is in the sixth stage. Peer feedbacks encourage the reviewer to write a qualified review. The final stage is to discuss findings and to draw a conclusion.

To cite this document:

Rattahpinnusa, H. H. (2021). An Exploratory Study of Advantages and Disadvantages of Website Preservation. *Record and Library Journal*, 7(1). 1-6.

Open access under Creative Commons Attribution-Share A like 4.0 International License

(CC-BY-SA)

Result & Discussions

The Advantage of website preservation

It is argued that accessible library's website has an important role to develop collective memories to future generations. Library as a part of GLAM institutions has responsibility to preserve websites due to values of digital heritages. UNESCO recognises websites as digital heritage because many website contents express the special characteristic of human knowledge. That digital heritage has a significant contribution to the political, educational, and economic development in the future (Cadavid, 2014). Similarly (Gomes, and Costa, 2014) stated that website preservation would supports humanities researchers and historians in term of accessing historical web data. Visi versa, both humanities researchers and historians can be users as well as contributor in providing historical data web. By accessing digital heritage website, future generations will learn historical events in the past so they can develop collective memories about their ancestor.

In addition, website preservation is an effort to maintain accessibility to digital heritage. This preservation could overcome the fragility of obsolete technology. As (Glenville, 2010) mentions that, the fragility of networked digital contents and the rapid change of technological development encourage libraries to archive websites. Every website has a link called Uniform Resource Locator (URL), and that link directs readers finding source materials. In fact, many URL are fragile items. A project named perma.cc showed that broken links were common problems causing inaccessible websites (Potter, 2014). That project provides a clear evidence of the importance of websites preservation.

Furthermore, website preservation offers an opportunity for library to participate in collaborative project. As (Thom, 2012) stated that website preservation will encourage library working collaboratively with other library developing online resources database of archived website contents. Moreover, the website preservation offers efficiency in collecting and managing metadata. According to Lyle (2004 cited in (Niu, 2012)), library can save cost, time, and storage space by developing website repository contents so library's delivery information service could reach broader area coverage and audience. In addition, developing website repository contents does not require a lot of human interference because the process of website preservation is rely on web crawler machine (Crook, 2009). Thus, the library does not necessary spend their budget to hire extra staffs. The next benefit is a wide coverage. According to (Koerbin, 2005), the crawling machine successfully captured various websites with Australian domain.

Considering those argument above, it is clear that website preservation has advantages for library to develop collective memory future generation about historical events in the past. Website in early 2000 is a momentous of digital revolution. At the time, computer and internet are popular items to produce information. The invention of blogspot and html website encourage people to publish their work into internet so other people can access and read digital information. At that time, website become a popular learning resources outside classroom. Although technology of website has weakness point at url link. Experts were trying overcome that weakness point by developing wayback machine. The wayback machine crawled the url website and it stored the website contents in its database. The wayback machine helps people to retrieve information which stored on the obsolete website.

Disadvantage website preservation

In another perspective, it is believed that website preservation has many drawbacks rather than its benefits. In a minor perspective, the website preservation will be an expensive project which can burden time, funds and human resources. According to (Niu, 2012), the website

To cite this document:

Rattahpinnusa, H. H. (2021). An Exploratory Study of Advantages and Disadvantages of Website Preservation. *Record and Library Journal*, 7(1). 1-6.

Open access under Creative Commons Attribution-Share A like 4.0 International License

(CC-BY-SA)

preservation is time consuming tasks because it requires staffs spending their time to select relevant web sites and to evaluate the contents of website. Furthermore, the web preservation is classified into a labour intensive task due to required a great number of employees to evaluate the website contents (Jacobsen, 2008; Koerbin, 2005). The website preservation is an expensive project due to its complexity.

In addition, website preservation is potentially break a legal deposit law in few country. For example is a case study of website preservation project at British Library, England. According to (Tuck, 2008), legal deposit law enables British Library to get permission harvesting website' contents, but that legislation applies for a restricted access on those contents. That finding shows that the British library get only permission to download the website's contents without any right to redistribute those contents. In other case study, Koninklijke Bibliotheek (KB) of Netherland was placed in a similar situation with the British Library. According to (Glenville, 2010), Koninklijke Bibliotheek (KB) finds difficulty to redistribute the personal website's contens due to legal deposit laws. In this case, KB should ask permission to website's owner to harvest it contens and to redistribute it. After get the permission, KB could preserve the website contents. However, without any permission from the website owners to redistribute it, the harvested website's content is illegal to redistribute it. Consequently, archived website can be deleted from the libraries' database the website's owner to requests it (Cadavid, 2014). That finding discourages the British Library and Koninklijke Bibliotheek to disseminate any website's contents.

From discussion above, there are two disadvantages of website preservation. Firstly, the website preservation is a complex process which require skilled staff, availability of supporting funds and having a long period of time. The website preservation process involves selecting and evaluating tasks to identify the relevancy and the importance information on the website contents. Secondly, website preservation will potentially break a legal deposit law if the website's owner does not give permission to library for redistributing website's contents. The identified disadvantages of website preservation may become a consideration for library to conduct website presevation program in the future.

Problem of during website preservation process and its solution

Based on the literature review, this paper recognises that violate copyright law is a problem during website preservation process. An illegal web crawling process can violate copyright law, and copyright infringement has a serious impact on the accessibility. According to (Jacobsen, 2008), copyright law is a potential obstacle for libraries to provide accessible digital materials. Based on that fact, many libraries should prevent breaching copyright law by preparing legal strategies. Firstly, extended legal deposit law is a strategy to mitigate copyright issue. According to Beunen and Sciphof (2006 cited in (Glenville, 2010)), several national libraries in Europe, South Africa, and New Zealand have extended their legal deposit laws to cover digital materials. Revised legal deposit prevents them from copyright infringement.

Moreover, revising legal deposit law strategy can minimise the legal impact of whole domain harvesting. The nature of web crawling is to capture targeted websites without exception. Legal deposit law should be extended to cover digital materials. Thus, the law gives the mandate to preserve digital materials. In the real practice, the State Library of Tasmania' experience, comprehensive legal deposit legislation can minimise copyright issues because that legislation encourages the state library to legally possess duplicates of Tasmanian digital materials (Sokvitne, 1999). By having the legal law, libraries can reduce risks of the copyright infringement during web crawling process

To cite this document:

Rattahpinnusa, H. H. (2021). An Exploratory Study of Advantages and Disadvantages of Website Preservation. *Record and Library Journal*, 7(1). 1-6.

Open access under Creative Commons Attribution-Share A like 4.0 International License

(CC-BY-SA)

Conclusion

In conclusion, website preservation will be compulsory program for libraries in the future. Websites contain special resources of human knowledge, but the fragility of the digital network may harmful for accessibility of the website contents for future generation. Website preservation will have a lot of benefits rather than drawbacks. Library will be able to develop collective memory for future generations by preserving website contents. In addition, the website preservation also offers efficiency in collecting and managing metadata due to its advance technology. Furthermore, Cost efficiency and broad coverage become advantages of website preservation.

Meanwhile, potential problem of violation copyright law during website preservation process can be avoid by revising legal deposit law. This solution can minimise the legal impact of web crawling which capture targeted websites without exception during website preservation process.

Acknowledgement

My appreciation to Australian Award Scholarship whom providing financial support to pursue my post graduate degree in information management.

References

- Cadavid, AP. (2014). Copyright Challenges of Legal Deposit and Website Preservation in the National Library of Singapore. *Alexandria*, 25(1-2), 1–19.
- Costa, M, Gomes, D, and Silva, M. (2016). The Evolution of Website Preservation. *International Journal on Digital Libraries*. <<http://link.springer.com.access.library.unisa.edu.au/article/10.1007/s00799-016-0171-9>>.
- Crook, E. (2009). Website Preservation in a Web 2.0 World. *The Electronic Library*, 27(5), 831–36.
- Glennville. (2010). Web Is Archiving: Ethical and Legal Issues Affecting Programmes in Australia and the Netherlands.” *The Australian Library Journal*, 59(3),128–34.
- Gomes, D, and Costa, M. (2014). Digital Preservation of the Importance of Web Archives for Humanities.” *International Journal of Humanities and Arts Computing*, 8(1), 106–23.
- Jacobsen, G. (2008). Website Preservation: Issues and Problems in Collection Building and Access. *Liber Quarterly*, 18(3-4), 366–376.
- Koerbin, P. (2005). Report on the Crawl and Harvest of the Whole Australian Web Domain Undertaken during June and July 2005.” https://pandora.nla.gov.au/documents/domain_harvest_report_public.pdf.
- Niu, J. (2012). An Overview of Website Preservation. *D-Lib Magazine*. <<http://www.dlib.org/dlib/march12/niu/03niu1.html>>.
- Phillips, ME, Alemneh, DG, and Ethiopia, NA. (2015). *Web-Archiving: Collecting and Preserving Important Web-Based National Resources*. Texas: University of North Texas.
- Potter, A. (2014). *The Why and What of Web Archives*, Library of Congress. <<https://blogs.loc.gov/digitalpreservation/2014/04/the-why-and-what-of-web-archives/>>.
- Sokvitne, L. (1999). Our Digital Island: Web Preservation Issues and Solutions at the State Library of Tasmania. *State Library of Tasmania*.

To cite this document:

Rattahpinnusa, H. H. (2021). An Exploratory Study of Advantages and Disadvantages of Website Preservation. *Record and Library Journal*, 7(1). 1-6.

Open access under Creative Commons Attribution-Share A like 4.0 International License

(CC-BY-SA)

- <<http://ausweb.scu.edu.au/aw99/papers/sokvitne/paper.html>>.
- Thom, C. (2012). Benefits of Digital Preservation - Bishopsgate Library. *wiki opbs lbs*. <http://wiki.opf-labs.org/display/SPR/Benefits+of+Digital+Preservation+-+Bishopsgate+Library>.
- Tuck, J. (2008). Website Preservation in the UK: Cooperation, Legislation and Regulation. *Liber Quarterly*, 18(3-4), 357-365.
- Webster, J, and Watson, RT. (2002). Analysing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly*, 26(2), 13–23.
- Wikipedia. (2021). Website Preservation. < https://en.wikipedia.org/wiki/Web_archiving>.

To cite this document:

Rattahpinnusa, H. H. (2021). An Exploratory Study of Advantages and Disadvantages of Website Preservation. *Record and Library Journal*, 7(1). 1-6.

Open access under Creative Commons Attribution-Share A like 4.0 International License

(CC-BY-SA)