



Research Article

Repercussion of Using Internet Sources: Dilemma for Research Communities

Nilakshi Sharma
Librarian
Government Model College
Deithor, Karbi Anglong, Assam, India
Email: snilakshi4@gmail.com

S. N. Singh
Professor
Department of Library and Information Science
Mizoram University
Aizawl, India
Email: drsinghsn@yahoo.com

Received: 5 Apr. 2022

Accepted: 13 Mar 2023

© 2023 Sharma and Singh. This is an Open Access article distributed under the terms of the Creative Commons-Attribution-Noncommercial-Share Alike License 4.0 International (<http://creativecommons.org/licenses/by-nc-sa/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly attributed, not used for commercial purposes, and, if transformed, the resulting work is redistributed under the same or similar license to this one.

DOI: [10.18438/eblip30144](https://doi.org/10.18438/eblip30144)

Abstract

Objectives – Consultation of internet sources for educational and research purposes is the new normal. As a result of information communication technology, information creation and access are more convenient. The current study was carried out to investigate the proportion of use of internet sources for research purposes by research scholars of three central universities of North East India, namely Tezpur University, Mizoram University, and Rajiv Gandhi University.

Methods – The researchers collected data from 123 respondents through an online questionnaire that was distributed through different social media platforms. The study was conducted among Research Scholars (PhD and M.Phil) of Mizoram University, Rajiv Gandhi University, and Tezpur University.

Results – The research results show that research communities are moving toward digital platforms for searching and consulting their required resources. Most of the respondents consult internet sources for writing their research reports, but they do not format the references properly. Some research scholars do not follow any referencing style for citing web documents, and respondents do not have much awareness about the differences between URLs and DOIs. Research communities also face problems due to the inaccessibility of online documents cited by former researchers.

Conclusion – Since most of the respondents are not familiar with the use of web archives, the current study suggests that higher education institutions should arrange awareness programs on the use of web archives. Research communities should follow the proper referencing formats to acknowledge others' works. Publishers should mandate a citation style for authors and verify the accuracy of the references before publishing articles or other works.

Introduction

The publishing industry has expeditiously adopted e-publishing over print publishing in the last few decades. The internet has introduced new avenues for disseminating information quickly through e-publishing at minimal cost compared to print publishing. Most electronic publishing sources are available over the internet and swiftly accessible, which is indeed a privilege for research communities. Digital repositories and their relationship with the Open Access (OA) movement constitute a new trend in scholarly communication worldwide that will set new standards for information sharing and management (Wani, Gul, & Rah, 2009). In their study, Halloumeh & Jirjees (2016) found that both faculty members and graduate students highly preferred electronic journals over print journals. E-resources occupy a significant portion of the global literature (Baskar, 2017).

Researchers cannot move on the right track with their research without scrutinizing publications of previous research in that field. To build basic understanding of a topic, researchers must consult previous works. Consulting resources is easier since information communication technology made it possible to access almost all sources of information from anywhere at any time without much physical or financial effort. However, while there are advantages of internet sources, there are also some drawbacks. All documents available on the web have an address called a universal resource locator (URL), but these web addresses may vanish over time for reasons such as changing the document location, server issues, change in domain name, etc. This phenomenon is known as link rot, or broken links. Along with several other studies, a study by Bansal (2021) showed that a huge number of URLs cited in reputed journals were missing, posing a problem for researchers.

Unavailability of previously cited resources may create confusion, and finding such a gap may be puzzling to researchers. Vanishing resources also impact author citations. To discover the impact that broken links have on research, the present study was carried out among research scholars from different fields at three central universities in North East India.

Literature Review

With the emergence of information communication technology, publishers have shifted from traditional modes of publishing to web publishing. Web publishing and the use of web information sources go hand

in hand. Studies have claimed that the use of web resources by research communities are increasing. Kumar (2019) stated that advantages like the ability to hyperlink, use of a graphical interface, and remote access are the major reasons for the popularity of the web as a 24/7 information hub. The study revealed that almost 84% of the research scholar respondents used the web daily for writing articles, acquiring information, and other academic purposes.

On the other side, many scientists have shown that the use of web resources is not reliable, since web addresses vanish and documents may become unavailable for readers after a few years. In their study, Bansal and Parmar (2020) found that web resources cited in the *Current Science* journal have disappeared from their original locations with a common error message: *404 File Not Found*. However, according to their research, the average half-life for these missing URLs was 1.76; that means it would take almost 2 years for half of the URL citations to vanish. Lack of constancy in web references implies that persistent availability of online information resources is not sure or guaranteed. Kumar and Sushmitha (2019) found that, out of a total 1105 URLs cited in *Annals of Library and Information Studies* from 2006 to 2015, only 625 URLs remained active, and 480 URLs were missing. They also tested the accessibility of URLs and their recovery through the [Time Travel](#) tool. The Internet Archive had the highest number (56.46%) of missing URLs, followed by Bibliotheca Alexandria Web Archive (15%).

Kumar (2018) studied the prevalence and persistency of URLs cited in articles between 2006 and 2015 of 12 LIS journals from reputed publishers. Results revealed that the LIS researchers were highly dependent on web resources for scholarly information, and as a result, URL citations in LIS scholarly literature have increased in the last few years. The number of URL citations increased from 1,870 in the year 2006 to 5,673 in the year 2015. A total of 93.80% of respondents used URL citations in their scholarly articles; meanwhile 64.02% of respondents mentioned that they choose to use web resources due to easy accessibility.

In another study, Kumar and Kumar (2017) used the [W3C Link Checker](#) tool to find out the availability of URLs and found that 1310 URLs were active, whereas 823 URLs were missing. The researchers recovered a total of 484 URLs with the Internet Archive, but the remaining 339 URLs were not found. Arao, Santos, & Guedes (2015) studied the half-life and obsolescence of the citations included in dissertations and theses from the post-graduate program of science of the Belles-lettres School of the Federal University of Rio de Janeiro. The half-life study of citations was limited to only books and book chapters and showed that some fields have a relatively long half-life period. The results indicated that the half-life of science literature from 2007–2008 is 14 years, and the half-life of science literature from 2011–2012 is 15 years.

But as every problem has a solution, web archiving may be a solution to this serious issue. Gul, Mahajan, & Ali (2014) mentioned that lifelong availability of web resources is not guaranteed, but solutions like web archiving can prevent the decay of web citations. Ben-David & Huurdeman (2014) claimed that the field of web archiving is at a turning point. Previously, in web archiving, the single URL has been the dominant unit for preservation and access. Access tools such as the Internet Archive's Wayback Machine reflect this notion as they allow consultation, or browsing, of one URL at a time. In recent years, however, the single URL approach to accessing web archives is being gradually replaced by search interfaces. Ben-David & Huurdeman address the theoretical and methodological implications of this transition to search; they recommend that web archiving should be appreciated and implemented to preserve web resources, which may help future researchers or lay persons find information.

Aims

This study was carried out to fulfil the following objectives:

1. To find out the use of internet sources among researchers.
2. To know the purpose of using internet sources.
3. To understand the web citing behaviour of researchers.
4. To understand the problems faced by researchers due to broken references.

Scope and Methodology

The study was conducted among PhD and MPhil Research Scholars of Mizoram University, Rajiv Gandhi University, and Tezpur University. Basic information about these three institutions is given in Table 1.

Table 1
Basic Information of Institutions

Name of the Institution	Establishment Year	State	Responding Departments	Total Respondents
Mizoram University (MU)	2001	Mizoram	5	62
Rajiv Gandhi University (RGU)	1984	Arunachal Pradesh	3	22
Tezpur University (TU)	1994	Assam	3	39
Total			11	123

The researchers systematically structured an online questionnaire (available in Appendix) in English. The questions were formulated based on the analysis of studies related to the availability and persistency of web references. It was shared through different social media platforms like WhatsApp, Facebook, and ResearchGate. The distribution of the questionnaire started 20 December 2021, and 123 responses were received by 7 February 2022. Data were statistically analysed and presented in different tabular and graphical forms. Institutional review of our research methods and survey were not required.

Limitations

The questionnaire included questions requesting details such as age. These questions were required so that the data could be analysed according to scholar demographics. However, researchers have been careful to maintain the privacy of respondents.

Results

Demographic Characteristics of Respondents

From five departments at MU, three departments at RGU, and three departments at TU, totals of 62, 22, and 39 respondents respectively were received. As shown in table 2, more female research scholars than male research scholars responded to the questionnaire, and most of the respondents were pursuing a PhD. The age group 26–30 represented 53 respondents (the largest number of respondents from any of the age categories), while only three respondents were in the age group >40.

Table 2
Demographic Structure of Respondents

Demography		Frequency	Cumulative Frequency	Relative Frequency	Cumulative Relative Frequency
Sex	Female	67	67	0.54	0.54
	Male	53	120	0.43	0.98
	Prefer Not to Say	3	123	0.02	1.00
Age	<26	21	21	0.17	0.17
	26–30	53	74	0.43	0.60
	31–35	35	109	0.28	0.89
	36–40	11	120	0.09	0.98
	>40	3	123	0.02	1.00
Degree Pursuing	MPhil	46	46	0.37	0.37
	PhD	77	123	0.63	1.00

Purpose of Using the Internet

All respondents (100%) from all the universities used the internet for research activities (see figure 1). All respondents of Mizoram University used the internet for educational purposes. Majority of respondents also used the internet for social connectivity with others and entertainment purpose. Additionally, Respondents, i.e., 37% of MU, 45% of RGU and 67% of TU use internet for purposes like online marketing, internet banking, shopping etc.

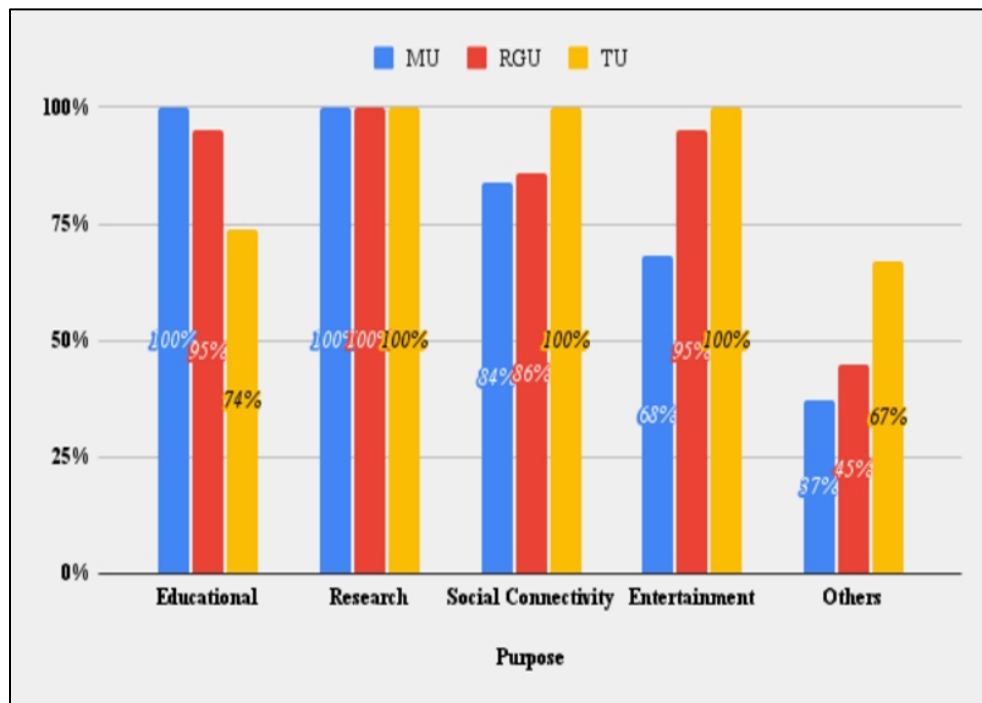


Figure 1
Purpose of using the internet.

Use of Internet Sources in Different Research Activities

Research scholars mostly used various internet sources for the literature review of their research areas (see figure 2). Although literature review is a process that helps in developing research knowledge like understanding methodologies and tools, how to verify results, and so on. Most of the researchers considered literature review as a separate process. They indicated collecting sources related to their research area to be added in the literature review section of their theses. No respondents from RGU were found to use internet sources to understand the methodologies. Meanwhile 17% of respondents from TU prioritized internet sources for using data collection tools, and 30%, 27%, and 10% of respondents from MU, RGU, and TU respectively prioritized use of internet sources for collecting data.

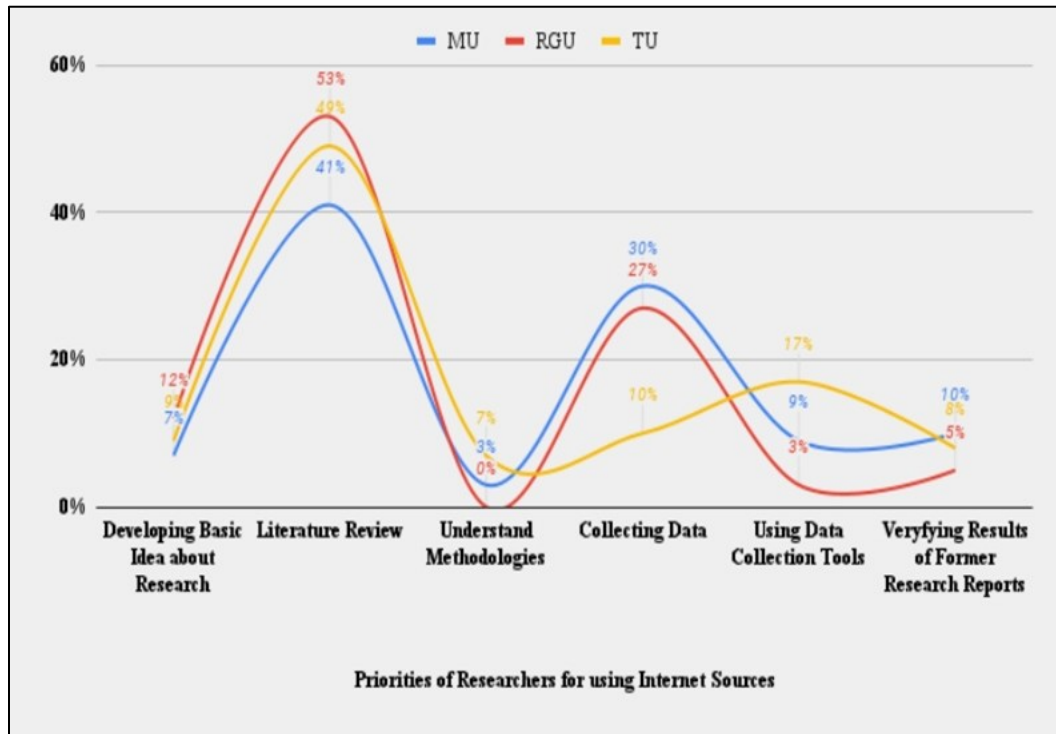


Figure 2
Use of internet sources in different research activities.

Types of Internet Sources Used for Research Purposes

As shown in figure 3, all respondents from all three universities used e-journals and e-books as information sources. Majority of respondents also used electronic theses and dissertations (ETD). All respondents from RGU and TU used websites as information sources; on the other hand, only 79% of MU respondents used websites to collect information. Blogs were an important source of information for 65%, 29%, and 78% of respondents from MU, TU, and RGU respectively. All the sources used by respondents have web addresses that they included in the references of their research reports.

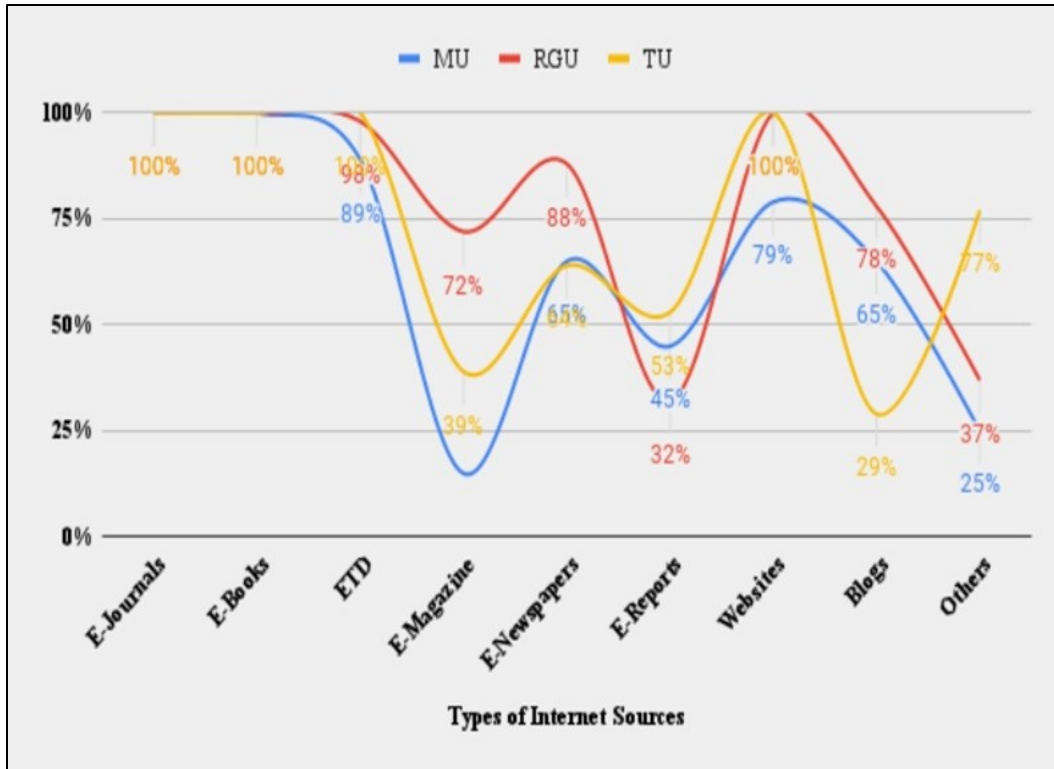


Figure 3
Types of internet sources used.

Document Searching Techniques

Fifty-seven percent, 49%, and 61% of respondents from MU, RGU, and TU respectively searched for a document using the URL of the source (see figure 4). Only 25%, 27%, and 3% of respondents from MU, RGU, and TU respectively used the Digital Object Identifier (DOI) of the document, which may be a reason why researchers could not access many of the documents. DOIs are more stable compared to URLs, but the results show that researchers tend to search web documents through their URLs. Some scholars search documents using keywords or by searching the title of the document in search engines.

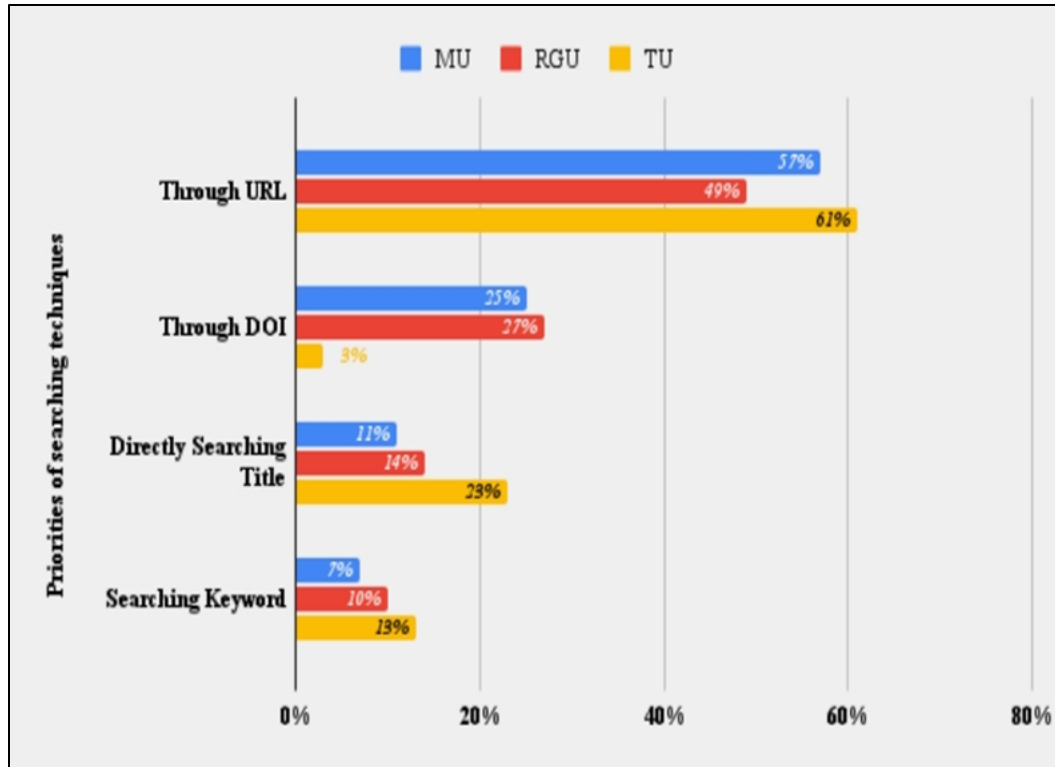


Figure 4
Document searching techniques of respondents.

Citing Behaviour of Research Scholars

Today's research publications are tomorrow's research guides. All the references included in today's research reports may be searched by other researchers for many decades to come. Therefore, it is essential to study the citing behaviour of today's researchers, so that the behaviours can be appreciated, modified, or rectified if needed. The use of reference management software (RMS) was not popular among respondents: very few respondents preferred to use RMS to create references (see figure 5). Most of the respondents added all the available details of the web source, which may not have included the retrieval date. However, 42%, 23%, and 19% of respondents from MU, RGU, and TU respectively created references as per the guidelines given by a particular reference style. Some respondents added only the web address of the source.

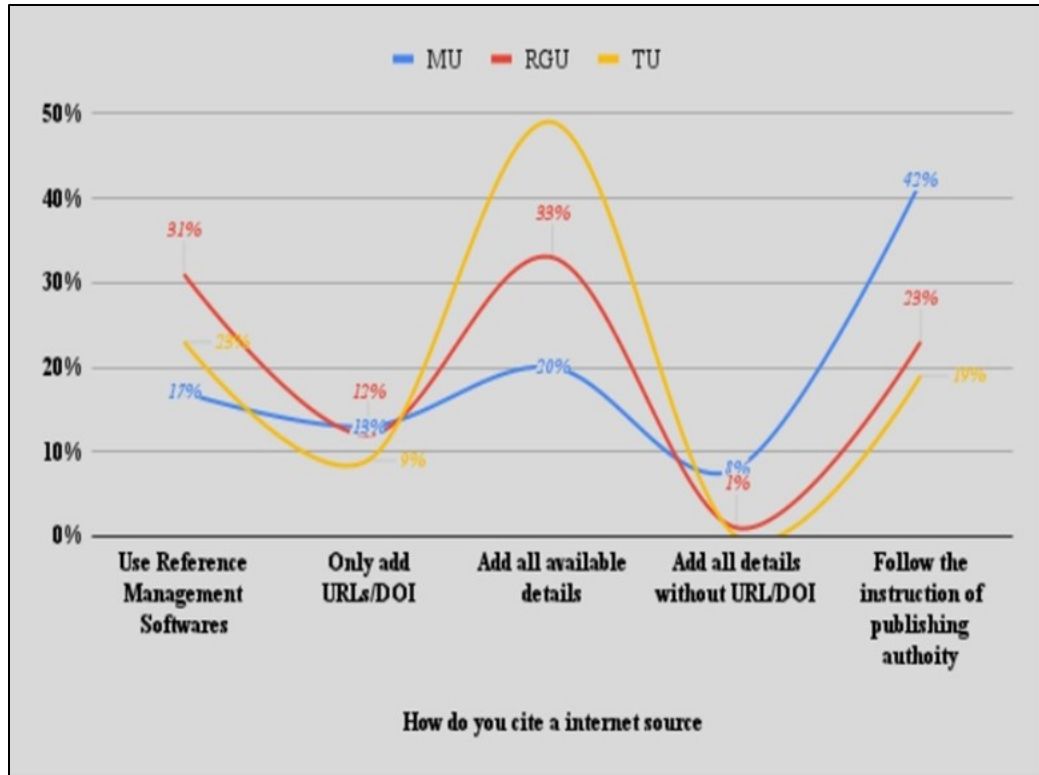


Figure 5
Citing behaviour of research scholars.

Problems Faced Due to Inaccessible URLs

All research scholars responded positively to the question “Have you ever witnessed the HTTP Error during your document search?” For various reasons, web addresses had vanished, and the user had faced problems. A problem faced by research communities is inaccessibility of resources, which may result in confusion among scholars. A few scholars mentioned that their research was derailed when they could not access the important sources. One scholar wrote that “since my research were based on secondary sources, I was reliable on the internet sources to some extent. But while I found that some of the resources are relocated and some have gone missing, it was like disaster to me.” Another researcher wrote that “while I was looking for available literatures to my research area, a lot of literatures were available at that time and I was satisfied with it. But while I started downloading the files, I found some of the most important papers were missing.”

Understanding About Web Archives

To assess scholars’ level of understanding about web archives, the questionnaire included the question “Do you know web archives can help you to access your missing resources?” In the results, one respondent shared their experience of using the Internet Archive (suggested by their research supervisor) for searching missing documents. However, notably, 47% respondents answered the question with “No,” showing a lack of awareness among research communities about using web archives.

Discussion

Along with the increasing rate of using internet sources for research activities, confusion about inaccessibility of important sources is also increasing. Research reports are the intellectual output of an institution; they contain rich data and are a major source of information for present and future generations, but those future generations of researchers will be unable to access this information if web addresses are missing. It is the duty of today's research communities to preserve our intellectual output for future generations. This study interprets that using the internet is the new normal for society. The internet is acting as a bridge to link today's society with knowledge, directly contributing to the process of building a knowledge society.

This study also found that although literature review should be an ongoing process carried out through the research process to aid in developing research knowledge, understanding methodologies and tools, verifying results, etc., many researchers consider the literature review to be an independent part of research. They tend to collect sources that are related to their research topic for the purpose of simply adding a literature review to their research report.

Along with other sources, research scholars use blogs as information sources. Compared to respondents from the other universities, respondents of Tezpur University showed a lower interest in blogs as information sources. This may be due simply to the researchers' personal interests. However, some research scholars may consider blogs as unauthentic, non-reliable information sources since they do not present filtered or processed information.

Focusing on availability of web information sources, DOIs are more stable than URLs, and researchers will have an easier time searching as well as provide a benefit to later researchers by using DOIs instead of URLs. Publishers should also focus on providing DOIs for online articles. Due to the inaccessibility of important resources, research scholars face different issues at different phases of their research, which may impact the number of citations received by an author or a document.

Based on the current study's results, we suggest that researchers should create references properly by using RMS. Researchers and publishers should prioritize permanent web addresses like DOIs or PURLs over URLs. Publishers should verify the accuracy of web references before publishing papers. Awareness of web archives cannot fully resolve the problem of inaccessibility but may reduce it to some extent. Therefore, institutes of higher education should organize awareness programs on the use of web archives, RMS, and similar tools. Content creators and users should use web archives to preserve important information sources. Authors should have a web-archiving system to preserve their own intellectual output for future references. Publishers should consider assigning DOIs to previously published articles. Librarianship is all about managing information with different tools and techniques. Librarians should lead this movement to make the knowledge society aware of this issue and find solutions to it.

Conclusion

Inaccessibility and loss of references may directly impact the research productivity of researchers and institutions. Knowledge creators should emphasize using proper information sources and adding proper references by using RMS. Web references should be properly formatted according to citation guidelines, and the permanence of web addresses should be maintained. Web archives should be used, and for that, awareness must be created among researchers. The current study has helped to explain the dilemma of research communities and to suggest solutions.

Author Contributions

Nilakshi Sharma: Methodology, Data curation, Writing – original draft, Writing – review & editing
S. N. Singh: Conceptualization, Visualization, Investigation

References

- Arao, L. H., da Costa Santos, M. J. V., & da Silveira Guedes, V. L. (2015). The half-life and obsolescence of the literature science area: A contribution to the understanding the chronology of citations in academic activity. *Qualitative and Quantitative Methods in Libraries*, 4, 603–610.
<http://www.qqml.net/index.php/qqml/article/view/271/261>
- Bansal, S. (2021). Decay of URL references cited in DESIDOC Journal of Library & Information Technology. *Library Philosophy and Practice*, Article 5811.
<https://digitalcommons.unl.edu/libphilprac/5811>
- Bansal, S., & Parmar, S. (2020). Decay of URLs citation: A case study of Current Science. *Library Philosophy and Practice*, Article 3582. <https://digitalcommons.unl.edu/libphilprac/3582/>
- Baskar, A. (2017). E-resources and its uses in library. *Journal of Critical Reviews*, 7(08), 3588–3592.
<http://www.jcreview.com/admin/Uploads/Files/61c717c2ea6a74.46602997.pdf>
- Ben-David, A., & Huurdeman, H. (2014). Web archive search as research: Methodological and theoretical implications. *Alexandria*, 25, 93–111. <https://doi.org/10.7227/ALX.0022>
- Gul, S., Mahajan, I., & Ali, A. (2014). The growth and decay of URLs citation: A case of an online library & information science journal. *Malaysian Journal of Library & Information Science*, 19(30), 27–39.
<https://ajba.um.edu.my/index.php/MJLIS/article/view/1781>
- Halloumeh, K. A., & Jirjees, J. M. (2016). Electronic versus print journals in academic libraries in Abu Dhabi: Preference & problems. *Advances in Journalism and Communication*, 4, 113–126.
<https://doi.org/10.4236/ajc.2016.44011>
- Kumar, A., (2019). The use of web and its resources among research scholars of social science at Guru Ghasidas University (A Central University), Bilaspur, Chhattisgarh, India: A study. *Library Philosophy and Practice*, Article 2457. <https://digitalcommons.unl.edu/libphilprac/2457>
- Kumar, D. V. (2018). *URL prevalence and persistence: Measuring the efficiency of Internet Archive and Google in the recovery of vanished URLs* [Doctoral thesis, Tumkur University]. Shodhganga.
<http://hdl.handle.net/10603/227210>
- Kumar, V., & Sampath Kumar, B T. (2017). Finding the unfound: Recovery of missing URLs through Internet Archive. *Annals of Library and Information Studies*, 64(3), 165–171.
<http://op.niscair.res.in/index.php/ALIS/article/view/16709>
- Kumar, D. V., & Sushmitha, M. (2019). Recovery of missing URLs cited in *Annals of Library and Information Studies*: A study of time travel. *Annals of Library and Information Studies*, 66(1), 24–32.
<http://op.niscair.res.in/index.php/ALIS/article/view/22420>

Wani, Z. A., Gul, S., & Rah, J. A. (2009). Open access repositories: A global perspective with an emphasis on Asia. *Chinese Librarianship: An International Electronic Journal*, 27. <http://www.white-clouds.com/iclc/cliej/cl27.htm>

Appendix Questionnaire

Greetings of the Day,

This survey is for research purposes only. We are currently working on an area named 'Repercussion of Using Internet Sources: Dilemma for Research Communities'. We request you to cooperate with us and provide the correct details. We will surely maintain the privacy of the respondents.

Thank you.

** Required*

Basic Details

Name *

(open-ended response)

Name of the Institution *

(open-ended response)

Name of the Department *

(open-ended response)

Date of Birth *

Example: January 7, 2019

Gender *

Mark only one oval.

- Male
- Female
- Prefer not to say
- Other:

Degree Pursuing *

Mark only one oval.

- MPhil
- PhD

Year of joining the current program *

Details Related to Topic

Do you use internet? *

Mark only one oval.

- Yes
- No
- Other:

Please mention purpose of using Internet? *

Check all that apply.

- Educational Purpose
- Research Purpose
- For Social Connectivity
- For Entertainment
- Other:

For which of the research activities do you use Internet Sources? *

Check all that apply.

- To develop basic idea of research
- For literature review
- To explore methodologies
- To collect Data
- To use online data collection tool
- To verify results
- Other:

What types of internet sources do you use? *

Check all that apply.

- E-Book
- E-Journals
- Electronic Theses and Dissertations
- E-Magazine
- E-Newspaper
- E-Reports
- Websites
- Blogs
- Other:

While searching an online resource, what technique do you use? *

Check all that apply.

- I search through the URL
- I search through DOI
- I directly search the title on search engines
- I search through keywords
- Other:

Do you cite all online sources used for the research purpose? *

Mark only one oval.

- Yes
- No
- I cite, but not all sources
- Other:

What method do you use to cite an online document? *

Check all that apply.

- I use Reference Management Software
- I add all available details of the source
- I add all available details except URL/DOI
- I follow the instruction of the journal
- I add only the URL or DOI
- Other:

Have you ever witnessed the HTTP Error during your document search? *

(open-ended response)

Have you ever heard about 'Web Archives'? *

Mark only one oval.

- Yes
- No
- Maybe

Do you know web archives can help you to access your missing resources? *

(open-ended response)

Any other additional information?

(open-ended response)

What do you recommend to overcome the problem of missing web resources?

(open-ended response)