

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Library Philosophy and Practice (e-journal)

Libraries at University of Nebraska-Lincoln

Fall 11-15-2020

Impact and Scope of Open Access Research Data: Perceptions of Library and Information Professionals

Naimat Ullah Shah

PhD Scholar, DLIS, Islamia Univesity Bahawalpur, naimat784@gmail.com

Salman Bin Naeem Dr.

DLIS, Islamia University, Bahawalpur, salmanbaluch@gmail.com

Rubina Bhatti Prof.

DLIS, Islamia University, Bahawalpur, rubytariq@gmail.com

Follow this and additional works at: <https://digitalcommons.unl.edu/libphilprac>



Part of the [Library and Information Science Commons](#)

Shah, Naimat Ullah; Naeem, Salman Bin Dr.; and Bhatti, Rubina Prof., "Impact and Scope of Open Access Research Data: Perceptions of Library and Information Professionals" (2020). *Library Philosophy and Practice (e-journal)*. 4619.

<https://digitalcommons.unl.edu/libphilprac/4619>

Impact and Scope of Open Access Research Data: Perceptions of Library and Information Professionals

By

Naimat Ullah Shah

Ph.D. Scholar,
Department of Library and Information Sciences,
Islamia Univesity Bahawalpur, Pakistan
Corresponding author email: naimat784@gmail.com

Dr. Salman Bin Naeem

Assistant Professor,
Department of Library and Information Sciences,
Islamia Univesity Bahawalpur, Pakistan

Dr. Rubina Bhatti

Professor,
Department of Library and Information Sciences,
Islamia Univesity Bahawalpur, Pakistan

Abstract

This study aims to prove that those scholarly articles belonging to subjects with changing proportions of open-access implementation have an incredible research impact if the research articles are openly accessible online than by paid access. This study also explores the impact of open access research data in Pakistan through the library and information science (LIS) professionals' perception. The foundation of this study includes surveying the research method. The LIS professionals were surveyed to assess their perception towards the possibly better impact of open access research data than non-OA ones. This study indicates that free-access research data have a better research impact and scope than those that have not openly accessible. This study is part of a Ph.D. dissertation. This study's respondents were library professionals working in public sector university libraries of Pakistan, so the library professionals are working in the libraries of colleges, schools, and private sector universities. Data should only be made open and placed online when there is a clear demand for it. Academic institutions should focus on releasing open data. Academic institutions should provide tools that make it easier for the researcher to look inside open datasets.

Keywords: Open Access, Open Data, Research data, Free Access,

INTRODUCTION

Information centers are in a continual development area and modify services to fulfill the quires in less time. The library scientists have unveiled several diversified tools to judge

consumers' attitudes that may act as a road map for bringing progressive modifications in services. One of the usual ways used by the libraries to assess the user approach to information and its utilization is Usability study. Such studies mainly target the computer-human relationships still, the usability studies may not be ignored because of its essential nature.

The open-access (OA) means free of cost available on the internet for the public, restricting none of the users to search, study, save, download, propagate, read, distribute, print, link, grab it for indexing, convert it as data to a software program, or utilize it for any other legal mission, producing no economic, legal or technical hindrances, besides those which are mandatory to be imposed. The only restriction on the re-presentation and distribution may be gifted in the form of copyright to the authors to secure their work's veracity so that their study may not be forwarded without proper citation and acknowledgment (Suber, 2012).

The term Open Access depicts a novel model for the broadcasting of scholarly manuscripts. The process of putting journal's subscriptions on sale while limiting the "way in" just to the payers only, developed in the age of printed journals, will be replaced by the Open Access technique slowly and steadily. A renowned campaigner of Open Access protocol illustrates the OA manuscript as a digital article that is not only available online without any charges but is also mainly free of copyright and licensing precincts (Das, 2008; Suber, 2012).

Model of data sharing through open access

The below figure demonstrates the process of sharing the OA research data that the researcher gathers the data to address the research problem. Then the researcher analyzes the data to draw inferences. The data is then stored in an institutional repository for further use and at last, the data was shared through open access to gain a high-level impact.

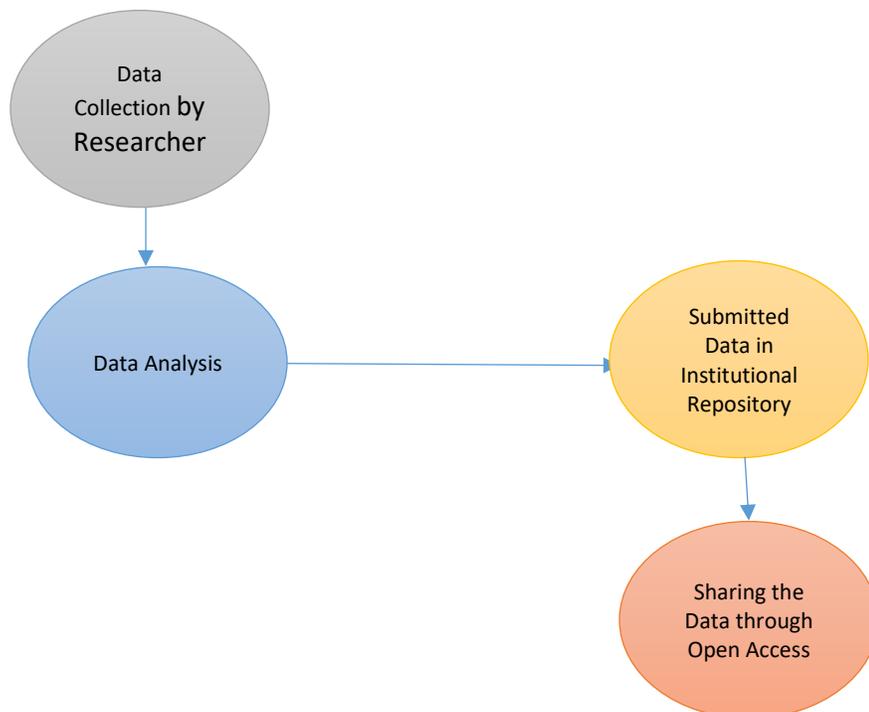


Figure-1 process of sharing research data through open access

Literature review

All the educational materials that reside under the public scope are usually presented with an open license. The literature with open status means a lack of restrictions on copying, utilizing, adapting, and re-representing this data (Salem Jr, 2017). In the modern world, the published manuscripts, as well as citations, are supposed to be the key indicators for the academic skills, research visibility as well as research productivity of a researcher (Brown & Abbas, 2010)

OA research journals may be classified as Green OA journals and Gold OA Journals. Despite these two ways of open access earlier, they were started to be appropriately documented from 2004. the Green OA journal as the publishing of a paper by an author in a non –OA research journal still making it possible to self-archive it in an open-access research archive (Björk, Laakso, Welling, & Paetau, 2014)

Policies and planning developed to access research and data openly are considered to be knotted with the ongoing practice of open access to research manuscripts, e.g., articles in research journals. In the initial hours of electronically publishing practices in the last decade of 20th century, voices raised open access to published literature to deliver the democratic right of free access to information to the public. Open access may be in various ways, including pre-print distribution even before publication, dissemination of post-prints after the manuscripts are being published, and publishing it without any restrictions to read it (Borgman, 2018).

"In the ideal case, the open data along with an open license is accessible, in a prearranged protocol with open and non-proprietary texture involving utilization of URIs as well as the to other data that is usually linked" (Berners-Lee, 2010).

Despite specific problems, the significant propensity is observed while collecting the handsome extent of internal and external data formally called data-lakes. These data-lakes are endeavor data management sites for storing, processing, and investigating data obtained from various sources, e.g. open data sources. Despite immense interest developed in recent years in this sector, the literature and studies reveal that the postulation of 'more use is better' is not a good case (Sadiq & Indulska, 2017).

The data's complexity: Any information transferred from a system's one part to another one will probably be more complicated than the same data in its initial location. The refining of data for making them feasible to be incorporated as a component in a vast compilation of inter-related ideas is done via semantic capabilities or ontologies (Rodríguez-Molina, Martínez, Castillejo, & Rubio, 2017).

Merchant, G., & Pountney, R. (2013) states that "the study has unveiled the IT-related activities and skills as a sample space for the problem that includes computer literacy as well as the partial reusability of open access data concerning schools" (Gruszczynska, Merchant, & Pountney, 2013).

To have a digital future in education processes, continuous efforts should be made to evolve and establish such learning tricks that follow users and target the classrooms and schools. Moreover, the inclusion of digital literacy, Openness of accesses and technological advancements in the pedagogy, and the curriculum are the need of the hour for a better understanding of ongoing practices in the field of digitalization of education processes (Cox, 2016).

The Semantic Web and particularly the Linking Open Data initiative, hearten the organizations to make their data publically available, share it with others and cross-link it via the web. The visibility of information can be interestingly improved via linking it with the sources of other data. This is favorable for both the profiting as well as the non-profiting institution. Libraries will fulfill users' hopes by being a linked data web's component as it will be able to provide incessant information availability in a format that can be easily understood to users who are not experts in library sciences (Hannemann & Kett, 2010).

Libraries have realized their capabilities and therefore, a variety of Institutions are scheduling their information as linked data. However, practically this seems to be an exigent practice. Besides the institutional obstacles, some technical issues may be faced by publishing data for semantic web and grabbing data from it by the cultural heritage institutions e.g. libraries particularly those with meager budgets allocated for IT operations. The paper aims to debate linked data concerning libraries and similar other institutions of cultural heritage, along with the presentation of a tangible report on the experience of the German National Library for establishing such a practice (Ezema & Onyancha, 2017).

It is supposed to be the need of the hour to develop strategic pathways for identifying factors that may affect the efficacy of open data utilities both on theoretical grounds and practical basis. The worth of derived data may be decided. We have identified as much as three corners that are further researched for the progress towards efficacious utilization of open data on the grounds of knowledge. The explored challenges in research reveal teamwork on behalf of researchers from information management, Information technology, data science, pure and social-science, etc., along with open data administrating agencies (Sadiq & Indulska, 2017).

Publishing of a manuscript is usually knotted with the condition to data delivering access to research, no matter if it is going to be published in an open-access forum or not. Releasing data is usually done while submitting the manuscript. Storage of such datasets is done in archives or stores where they get a distinct identification no—linked with the paper. Searching the paper or dataset will lead you to the other partner of the dataset-paper pair (Bourne, 2005).

Practicing within the cloud's rising knowledge basis can assist in dealing with various complex matters that libraries face during the operation of their datasets locally. Disambiguation,

data quality management, detection, individualization, enrichment, and detection are prime instances of those tasks. This is going to explore the tracks for novel services that employ data of multiple institutions. The linked data may be rewarded from the similar struggles done by libraries and other institutions of cultural heritage. The library data is highly qualitative as it is being unruffled, amended and preserved by expert professionals. Therefore, it can turn into the vital factor to be relied upon for the developing semantic web (Hannemann & Kett, 2010).

The trend of publishing scientific manuscripts and books in the format of open access and a facility of providing their prints to the reader upon demand is progressively in practice these days. The open-access leads to the more intense level of broadcasting of manuscripts that results in the fame of the concerned researchers and their affiliated organization and reward them a handsome worth. In open-access format, the task to make possible the access and sustainability are assigned to two varying groups (Standards & Practices, 2013).

The mature research project usually leads to multiple publications and datasets that result in a multifaceted between these two, which reveals their basic concept. Specific and other issues established a relationship between the released data and the published one. Once datasets are stored at a repository or archive, a Digital Object Identifier (DOI) may be assigned to them so that their retrieval may become possible in the same sequence (Standards & Practices, 2013).

Aim of the study

This research's fundamental purpose is to prove that those scholarly articles belonging to subjects with fluctuating free-access implementation data have an incredible research impact if they are openly accessible online than if they have paid access. This research also explores the effects of openly available research data through LIS professionals' perception and the scope of open access JCR listed in Pakistan's research journals.

Objectives

- To know about the ways used to utilize the open-access research data by LIS professionals in Pakistan.
- To know about the perception of LIS professionals towards the impact of open access research data in Pakistan.

Methodology

The data for this study was acquired by using the survey research method. Previous research studies were also consulted to know about the impact and scope of open access research data. Data collection tool was developed with the help of related literature for data collection from the library information science (LIS) professionals to know about the ways used to utilize the open-access data and to see the impact and scope of open data by LIS professionals, which have a significant impact on research productivity in this digital age. The LIS professionals of university libraries in Pakistan were surveyed using an online google form. All of the master's degree holders LIS professionals are the population of this study. A total of 256 questionnaires were distributed using a purposive sampling technique and 207 responses were received back. The researcher used personal phone calls, personal emails, and WhatsApp messages as a follow-

up strategy to enhance the researcher's response rate. Hence the response rate was 83 %. The collected data was first organized and then analyzed using SPSS 22.0 (Statistical Package for the Social Sciences).

Findings and discussion

Data users are seldom found as presenters of a study, which leads to the lack of an established pathway for the cleansing of data. So, results are frequently observed in the form of erroneous practices when dealing with the information (Becker, Duretec, & Rauber, 2017).

Unawareness regarding the data's residual quality is considered the leading menace when we talk about the utilities of open data. Observations reveal that the available open data is rarely used for the same purpose presented as an aim during the data collection protocol. Instead, it is usually utilized for other sorts of utilities. But it is a fact that a single data may be proved to be suitable for a specific field and, at the same time, un-qualitative for others.

Results of Quantitative data and discussion

The results were obtained by analyzing the collected data through a self-administered questionnaire regarding LIS professionals' perceptions of the impact and scope of open access research data. The questionnaire consisted of three parts. The first part dealt with the usage and awareness of LIS professionals towards free access. In contrast, the second part was mainly related to the impact and the third part associated with the scope of open access research data.

The results are as under:-

Gender wise distribution of the data:

The figure-2 demonstrates the gender-wise distribution of the sample of the study. The total response rate was 91%, out of which 167 (80.95%) participants were male, and the remaining 40 (19.05%) were female.

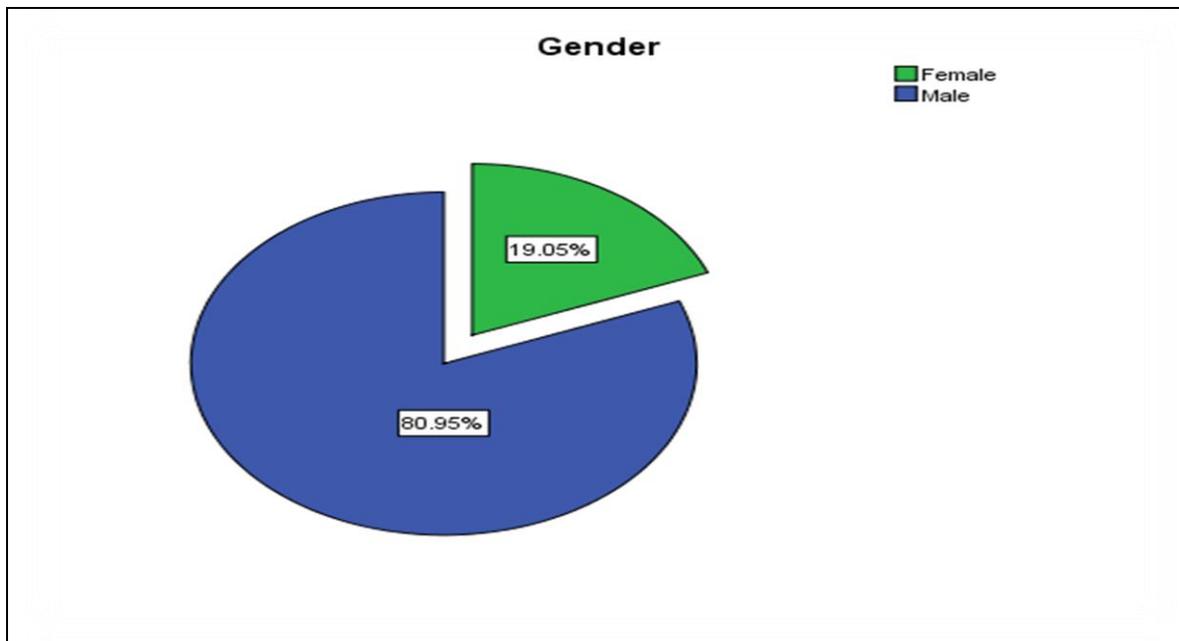


Figure: 2- Gender Wise Distribution (N-207)

Academic Qualification

Figure -3 demonstrates the academic qualification of the respondents of the study. Among the total respondents, 27.5% of the respondents have an M.Phil degree in LIS, and 72.5% have a master's degree in LIS.

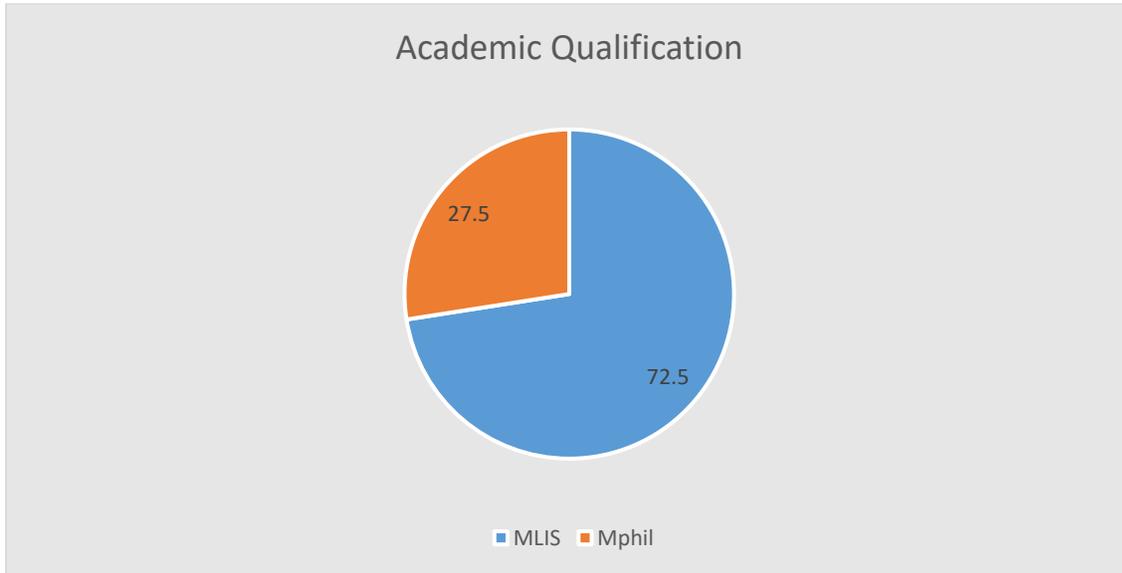


Figure:3- Academic Qualification of the respondents (N-207)

Professional Experience of the Respondents:

The figure-4 shows the professional experience of the librarians. It was found that most of the participants were experienced library professionals, 46% of participants have experience from 9 to 20 years, and thirty percent (28%) of respondents had experience from 4 to 8 years, while 17% of respondents had experience from 1 to 3 years.

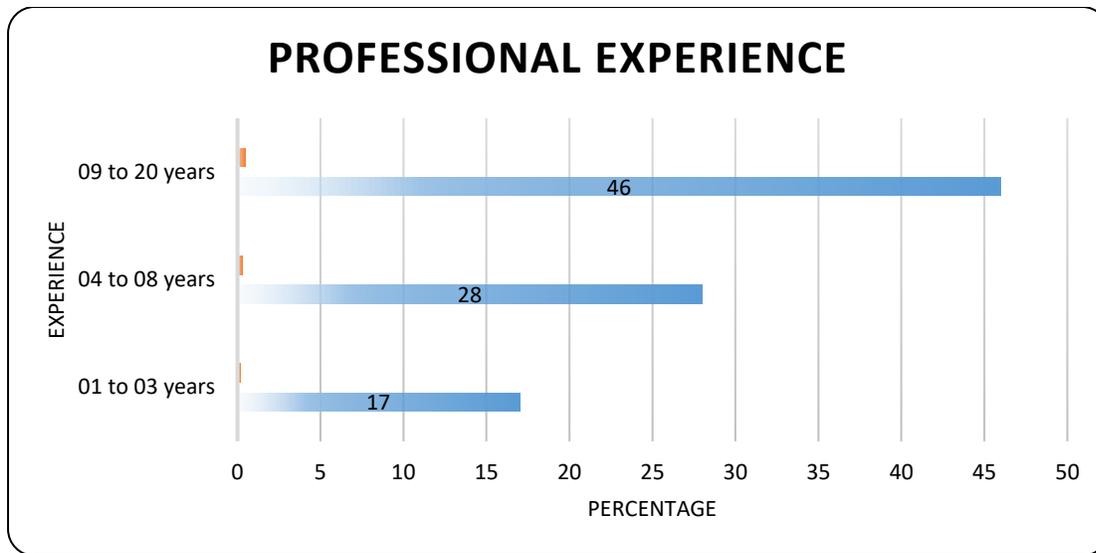


Figure:-4 Professional Experience of the respondents (N-207)

Ways of Open Access Research Data Usage

Open access may be in various ways, including pre-print distribution even before publication, dissemination of post-prints after the manuscripts are being published, and publishing it without any restrictions to read it (Borgman, 2018).

To know the usage of open access research data in many ways, the respondents were asked to rate the statements by choosing the option from a five-point Likert scale (Never-Always). The table results indicate that most of the respondents with the mean score ($\mu=3.53$) are used open access research data for creating reports or article building. The usage of open access research data for exploring the new data is ranked 2nd with a mean ($\mu=3.37$). Most respondents describe with the mean score of ($\mu=3.27$) that open access research data were used for data collection tool building. LIS professionals' concept has website building with the help of open access research data is ranked in the last with a mean score ($\mu=1.74$).

Table:1

Usage of open access data (N-207)

Statement	Mean	Std. Deviation
Report or article Building	3.53	1.250
Research thesis	3.47	1.515
Exploring the Data	3.37	1.208
Data collection tool Building	3.27	1.446
Creating a presentation	3.09	1.481
An application Building (mobile or desktop)	2.95	1.277
Statistical Research	2.71	1.336
Website Building	1.74	.998

Scale: 1-Never. 2-Sometime, 3-Rarely. 4-Oftenly. 5-Always

Impact of Open Access Research Data

To know the impact of open access research data, the respondents were asked to choose an appropriate option from the statements enumerated in the questionnaire using a five-point Likert scale (strongly disagree – strongly agree). The data revealed that (Table-2), most of the respondents are in the opinion that ($\mu=4.24$). The release of open research data is essential because institutions are not very good at building online services. Data should only be made open and placed online when there is a clear demand for it is ranked 2nd with a mean score of 4.14. Academic institutions should provide tools that make it easier for the researcher to look inside open datasets. This is also ranked 2nd with the same mean score of 4.14. The librarians can collaborate with the users regarding their information needs. Also, most of the respondents believe that ($\mu=4.05$) the Open access research data have a significant impact.

Table: 2
Impact of open access research data (N-207)

Statement	Mean	Std. Deviation
The release of open research data is essential because institutions are not very good at building online services.	4.24	.889
Data should only be made open and placed online when there is a clear demand for it	4.14	.793
Academic institutions should provide tools that make it easier for the researcher to look inside open datasets	4.14	.910
Open access research data have a significant impact	4.05	.921
Academic institutions should focus on releasing open data	3.95	1.024
The availability of open access research data is one of the most critical issues for any academic institution.	3.86	1.236
The release of open access research data is essential as a matter of principle in this online digital world	3.81	1.030
Open access can modernize the institution image and research-reputation	3.67	1.111

Scale: 1-Strongly disagree. 2-Disagree, 3-Neutral. 4-Agree. 5-strongly agree

Scope of Open Access Research Data

The perception of LIS professionals regarding the scope of open access research data indicates in the following table. It was asked from the respondents to choose an appropriate option from the five-point Likert scale (Never- Always) to know about the scope of open access research data. The results (table-3) describe that all the participants responded to the question. The results demonstrate that majority of the LIS professionals are used open access research data for digging the data about the subject of law with a mean score of ($\mu=3.10$) as well as the computer science subject is also dug out through open access with a mean score ($\mu=2.90$) which is the second important subject by the respondents. Life sciences with a mean score ($\mu=2.86$), Engineering with a mean score ($\mu=2.67$) and Mathematics with a mean score of ($\mu=2.57$) are ranked in 3rd and 4th

positions, respectively. As revealed from the table, Social Sciences is put at the last rank with a mean score ($\mu=2.00$) in the given categories of contents.

Table: 3-
Scope of open access research data (N-207)

Statement	Mean	Std. Deviation
Law	3.10	1.446
Computer Science	2.90	1.300
Arts and Humanities	2.86	1.389
Life Sciences	2.86	1.424
Engineering	2.67	1.683
Mathematics	2.57	1.287
Politics,	2.48	1.167
Physical Science	2.19	1.167
Social Sciences	2.00	1.140

Scale: 1-Never. 2-Sometime, 3-Rarely. 4-Oftenly. 5-Always

Conclusion:

The open-access data is used for research purposes, articles and reports building, exploring the new data, establishing the data collection tool, thesis building, and preparing the presentations. The release of open access research data is essential because institutions are presently not building online data services and most of the scholars are relying on it. The availability of open access data is one of the most critical issues for any academic institution. It is also discovered that most LIS professionals use open access data about the subjects of law, computer science, Life sciences, Engineering, Mathematics and Social Sciences. The respondents indicate that open access research data significantly impact research institution image than those who have not made their research data openly available.

Recommendations

1. Academic institutions should focus on releasing open access data.
2. Academic institutions should provide tools that make it easier for the researcher to look inside open datasets.
3. The institutions must release open access data, which is important as a matter of principle in this online digital world.
4. Open access can modernize the institution's image and research-reputation.

References

- Becker, C., Duretec, K., & Rauber, A. (2017). The challenge of test data quality in data processing. *Journal of Data and Information Quality (JDIQ)*, 8(2), 7.
- Berners-Lee, T. (2010). Tim berners-lee: The year open data went worldwide: TED.
- Björk, B. C., Laakso, M., Welling, P., & Paetau, P. (2014). Anatomy of green open access. *Journal of the Association for Information Science and Technology*, 65(2), 237-250.

- Borgman, C. L. (2018). Open data, grey data, and stewardship: Universities at the privacy frontier. arXiv pre-print arXiv:1802.02953.
- Bourne, P. (2005). Will a biological database be different from a biological journal? *PLoS Computational Biology*, 1(3), e34.
- Brown, C., & Abbas, J. M. (2010). Institutional digital repositories for science and technology: A view from the laboratory. *Journal of Library Administration*, 50(3), 181-215.
- Cox, J. (2016). Communicating new library roles to enable digital scholarship: A review article. *New Review of Academic Librarianship*, 22(2-3), 132-147.
- Das, A. K. (2008). Open Access to Knowledge and Information: Scholarly Literature and Digital Library Initiatives: The South Asian Scenario: UNESCO House.
- Ezema, I. J., & Onyancha, O. B. (2017). Open Access Publishing in Africa: Advancing Research Outputs to Global Visibility. *African Journal of Library, Archives and Information Science*, 27(2).
- Gruszczynska, A., Merchant, G., & Pountney, R. (2013). " Digital Futures in Teacher Education": Exploring Open Approaches towards Digital Literacy. *Electronic Journal of e-Learning*, 11(3), 193-206.
- Hannemann, J., & Kett, J. (2010). Linked data for libraries. Paper presented at the Proc of the world library and information congress of the Int'l Federation of Library Associations and Institutions (IFLA).
- Rodríguez-Molina, J., Martínez, J.-F., Castillejo, P., & Rubio, G. (2017). Development of Middleware Applied to Microgrids by Means of an Open Source Enterprise Service Bus. *Energies*, 10(2), 172.
- Sadiq, S., & Indulska, M. (2017). Open data: Quality over quantity. *International Journal of Information Management*, 37(3), 150-154.
- Salem Jr, J. A. (2017). Open pathways to student success: Academic library partnerships for open educational resource and affordable course content creation and adoption. *The Journal of Academic Librarianship*, 43(1), 34-38.
- Standards, T. G. o. D. C., & Practices, C.-I. (2013). Out of cite, out of mind: The current state of practice, policy, and technology for the citation of data. *Data Science Journal*, 12(0), CIDCR1-CIDCR75.
- Suber, P. (2012). Ensuring open access for publicly funded research. *British Medical Journal Publishing Group*. Retrieved from <https://www.bmj.com/content/345/bmj.e5184>