

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Library Philosophy and Practice (e-journal)

Libraries at University of Nebraska-Lincoln

2019

Linked Data Applications for Content Management of Serial Publications: A Case Study of Proceedings SZMC

Ihsan Basit

Shaikh Zayed Medical Complex, Lahore, Pakistan, librarian@skzmdc.edu.pk

Hafiz Nazeer Hussain

Lecturer: Department of Library & Information Science, Govt. Postgraduate College, Ravi Road, Shahdara, Lahore, hafiznazeerhussain@gmail.com

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

Basit, Ihsan and Hussain, Hafiz Nazeer, "Linked Data Applications for Content Management of Serial Publications: A Case Study of Proceedings SZMC" (2019). *Library Philosophy and Practice (e-journal)*. 2986.

<https://digitalcommons.unl.edu/libphilprac/2986>

Linked Data Applications for Content Management of Serial Publications: A Case Study of Proceedings SZMC

Ihsan Basit¹
Hafiz Nazeer Hussain²

1. Librarian: Shaikh Zayed Medical Complex, Lahore, Pakistan
2. Lecturer: Department of Library & Information Science, Govt. Postgraduate College, Ravi Road, Shahdara, Lahore, Pakistan

Abstract

Linked Data (LD) is increasingly becoming popular for the content management of serial publications. LD provides a shift in thinking from publishing data in human readable HTML documents to machine readable documents. The objective of LD is to publish the structured data by using specific standards i.e. Uniform Resource Identifiers (URIs), Hypertext Transfer Protocol (HTTP), Resource Description Framework (RDF) SPARQL etc. in order to interlink data and make it more useful through semantic web. This study aims at exploring the current status of LD applications for serial publications. This case study discusses the content management (CM) approach using LD technology for Proceedings Shaikh Zayed Medical Complex (PSZMC), the official medical research journal of Shaikh Zayed Medical Complex, Lahore. WordPress is used to publish contents of PSZMC after digitizing 64 issues (626 articles) with high resolution scanner. This practice-based case study has found that Wordpress is one of the LD applications that is used for the content management of the PSZMC. This is first study which may be helpful for the content management of medical journals in Pakistan.

Keywords: *Linked Data; Content Management of Serials; Proceedings SZMC, WordPress*

Introduction

The idea of Linked Data (LD) emerged in 1989 by Tim Burners-Lee as world-wide web (www). The WWW has been gradually transformed from web of page to semantic web and web of data (Berners-Lee & Fischetti, 2001). Semantic web is a network of data that links data to data, and to people's interests. Organizing content on the Web in this way fosters collaboration between computers and humans. The Semantic Web not only identifies document formats (HTML), but also their content and their meaning, which are represented in a technical language comprehensible to machines (Web Ontology Language (OWL)/ Resource Description Framework (RDF) (Berners-Lee, Hendler, & Lassila, 2001; Bizer, Heath, & Berners-Lee, 2011).

Linked data (LD) technology is gaining great importance in the common vocabulary of information technology, especially in the field of information science and Library management (Wright, 1997). LD is the published data on the web in readable and interpretable format with the sake of enhancement of its value of usability both by machines and humans (Guerrini & Possemato, 2013). Bizer, Heath, and Berners-Lee (2009) conclude LD as tools for publishing data on web to make it a usable part of universe of data space using the specific standards including Universal Resource Identifiers (URIs), Hypertext Transfer Protocol (HTTP), Resource Description Framework (RDF) and SPARQL (a semantic query language for databases which is used for retrieving and manipulating data stored in RDF).

Another most relevant term Linked Open Data (LOD) is referred to as the data which is available free for all and without having copyright issues. LOD collects the open data sets which are available on the web, and the paradigm of its detailed

growth occurring in a very brief period of time which demonstrates the level of interest that linked data has saved in organizations and institutions of different types. The work of this system is based on five steps of LOD, called building blocks i.e. LOD Data Model, Content Rules, Metadata Schema in LOD, Serialization of LOD and Exchanging of LOD (Guerrini & Possemato, 2013).

LD exposes descriptive metadata of serial publications as well as their relationship on the web which allows LD to create links among services of libraries regarding the serial publications and specifically numerous features of Web of data. LD is also capable of solving the challenges of serial publications single record versus compound, recent versus old entries both in print format and online. Library professionals are required to enrich the usage of their collections by adopting LD technologies and their specific standards. Catalogues of libraries provide lot of information regarding published and non-published material. Implementation of LD technologies in cataloging service as well as serial publications may enhance the successful searching ratio for researchers (Kaplan, 2012).

Linked Data continues to expand web of data and is increasing its impact on the web both in its depth and range of data sources. However, the applications of LD technology that can allow ordinary persons to contribute in expanding web of data or developing contents on web are difficult and still lacking. This case study made effort to implement “Wordpress” as Linked Data CMS, an approach allowing existing web content management system (CMS) software to be configured and to display journal contents to medical professionals as well as the research world. Wordpress prototype was implemented as Linked Data CMS approach. This approach provides the tools for semantic web application developers to rapidly develop an entire website based on linked data, while allowing ordinary web users to contribute directly to the web of

data using familiar CMS tools (Taylor, Jekjantuk, Mellish, & Pan, 2013).

Sheikh Zayed Medical Complex (SZMC), Lahore is a multi-disciplined hospital which provides medical education and research from undergraduate to postgraduate levels as well as tertiary healthcare services. SZMC is publishing Proceedings SZMC from 1987 on regular basis. This research publication is quarterly and peer reviewed as well as approved by PMDC. Generally, the present study explores the LD applications and initiatives taken for the digital content management of serial publications and finds the methods to implement LD technological approaches for the content management of SZMC Proceedings in electronic format. This study may help the library professionals to understand how the LD approaches can be applied for managing the institutional publications in digital formats effectively.

Objectives

This study is based on the following objectives:

1. To identify the LD approaches from literature used and implemented for serial publications on web.
2. To explore the methods how to apply LD approaches for content management of serial publications or institutional publications on web.
3. To find the practical solutions for managing the content of SZMC Proceedings on web.

Significance of Study

LD is increasingly becoming popular as content management of serials. Communication within electronic networks has become increasingly content-centric. There is a connection between LD and content management systems (CMS). However, LD applications are gaining importance in libraries to manage knowledge,

digital contents and serial publications as well. The local researchers do not have touched this area of research in Pakistan. Therefore, a study is needed to explore applications of LD for serial publications and to explore the standards of metadata which are used for content management of serial publications. This is the first empirical study on this topic in Pakistan. This is the first study which may help to understand LD tools for content management of medical serials in Pakistan.

Literature Review

There are many ways of digital content management. Thompson and Richard (2013) explored that DRUPAL, a freely available open source framework applied for CM, developed by Dries Buytaert in 2001, could be the best option for content management of serial publication. It was also found that DRUPAL showed some challenges i.e. while adding a vocabulary namespace to the RDF module, they were unable to edit or delete it without a great deal of effort.

Patel, Rathod, and Prajapati (2011) analyzed that three most important open source CMS were used including JOOMLA, DRUPAL and WordPress. WordPress provided a number of add-ons to handle the sites in a better way. It was reported that well renowned sites use this CMS. WordPress was used for blog platform initially but for the last two years, it had been used as a good CMS which facilitated top ranked documentation support.

Kaplan (2012) explored that the dedication of library professionals to confirm a structured export of MARC based data to the LD via Bibliographic Framework Initiative, might direct the mode of serial publications in LD.

Tharani (2015) found that only vendors played their role for developing, storing and sharing the metadata in case of books and research-oriented journals for the libraries. The same metadata collaboration of metadata was required at the

libraries end. Famous search engines like Google are now facilitating with users with automatically indexed and managed resources through LD. Libraries are particularly well positioned to realize this goal with their expertise in search, metadata generation, and ontology development. This study looks forward for taking more initiatives by libraries to become more responsive and to make library resources more appropriate to the process of information creation.

It was found that how latest principles for the serials could be changed to integrate LD models. Serial community would have to work together with other groups of metadata to upgrade and develop standards for serials in an LD environment. This joint work could make better the future of LD. LD models might be able to account effectively for the unique characteristics of serials. Hence the community of serials was gradually implementing LD practices with MARC-based practices, it was also perceived as imperative that these linked data models sufficiently accommodate legacy data (Balster, Rendall, & Shrader, 2018).

Bai and Cho (2009) produced characteristics for serials, description and structure of the bibliographic relationship stages to meet predetermined requirements. The main objectives of this project were:

“(a) to build an organized URI on the basis of FRBR and semantic associations for the serials; (b) to advance the ability of resource describing and organization of the Union Catalog System; (c) to enhance bibliographic and management intelligence by new insights into resources relationships operating at different levels;(d) to optimally streamline workflow of their institute i.e. to support Union Data Processing System to find the connections and match the relevant metadata; and (e) to promote the system of Central Repository and Union Service by higher demonstration of serials in search findings”, (p. 131).

A recent study of Béquet and Oury (2018) focuses on development of new ISSN Portal, which is a part of the new ISSN services. Readers may not be able to find detailed description of different functions of this portal, but only highlights of particular features that support the ISSN International Centre strategy regarding the distribution of information, the inclusion in the LD ecosystem, and the provision of expert services for the serial community at large.

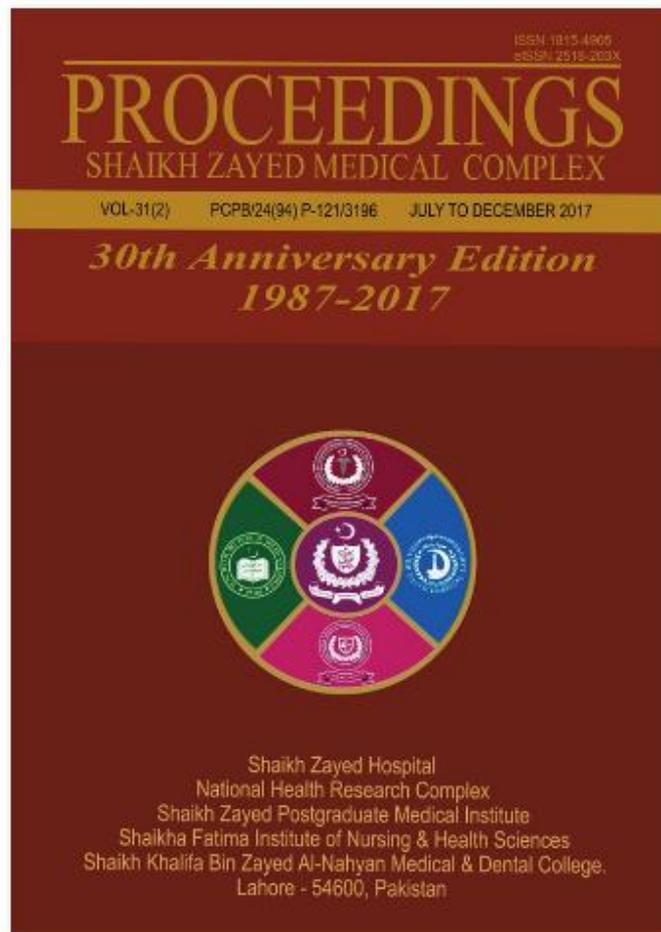
A review of related literature shows that the libraries are rapidly adopting applications of LD technology for information management on web, creating knowledge gateways in terms of Digital Libraries (DL), Online Public Access Catalogue (OPAC) and ILS etc. applying various metadata standards and relating MARC with LD in libraries, especially in content management of serial publications. Publishers use several types of schemes for their content management. However, most of the available studies on LD have been carried out only to explore the historical background and origin of LD. In addition to this, the primary literature could demonstrate comprehensively how different publishers use LD as the tool to manage their contents of serials. A few studies are available which has discussed this issue but not provided comprehensive solution. Those studies could not introduce and implement LD applications for content management of serial publications and could not facilitate to implement LD applications, comprehensively. Libraries are still facing problems regarding content management of serial publications on web. The study may be helpful to fill this gap.

Methodology and Procedures

This paper has addressed the efforts made for the content management (CM) of “Proceedings Shaikh Zayed Medical Complex (PSZMC)” which is an official and peer-reviewed quarterly Journal of Shaikh Zayed Medical Complex,

Lahore, Pakistan. Moreover, this project reports the CM approaches with the help of literature review. The published literature is reviewed to explore the conceptual background of content management of serial publications using LD technologies and to identify CM tools. Wordpress is identified as an LD application for content management on semantic web. Wordpress platform is used to create, manage, edits and publishes contents on the web sites. It meets the requirements of journal contents. It is also used for publishing library websites such as library general information, library services, library resources, various links and web 2.0 tools (Krubu, Idhalama, & Omigie, 2017). This case study demonstrates the process of using Wordpress content management system for the Proceedings of Shaikh Zayed Medical Complex (Figure 1).

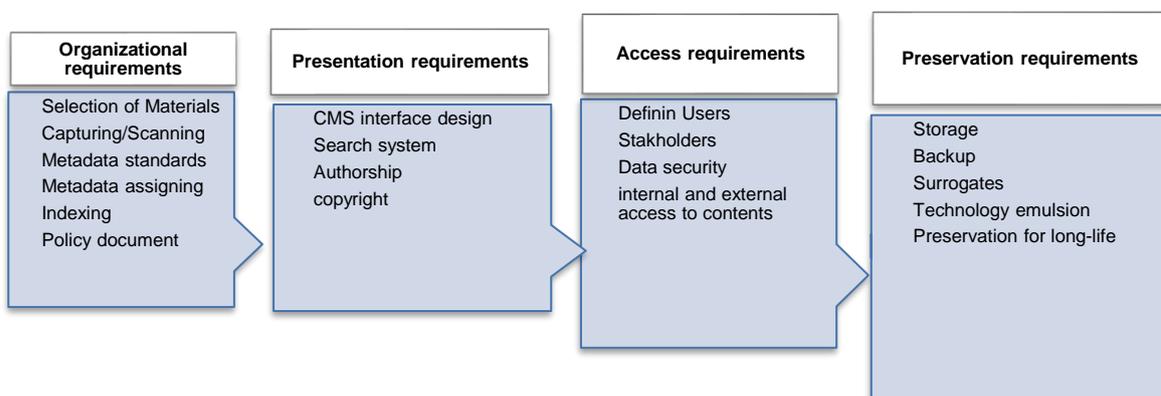
Figure 1: Proceedings of SZMC



Results of Case Study

The retrospective and current published contents of PSZMC were selected for digital CM. The contents of PSZMC comprised of 64 issues (626 articles), starting from 1987- 2015. Research articles from Proceedings were collected after taking approval from hospital higher management and editorial board. Collected data was scanned, analyzed and refined using Linked Data technologies. Refined contents of Proceedings were managed using WordPress (PHP and My SQL based free content management software). BookEye-4 professional scanner was used for scanning of contents of Proceedings and made available for public through the link: www.proceedings-szh.com. The project was initiated in 2017 and completed in 2018. Following steps were taking to manage the contents of PSZMC on semantic web using LD technology. Following steps, approaches and requirements undertaken in this project were identified necessary for the success of the CMS of PSZMC. To understand the areas in which the requirements addressed the functions, activities and procedures were grouped into following four major categories based on their nature as illustrated in Figure 2.

Figure 2: Requirements for CMS of PSZMC



- (1) **Organizational requirements.** Metadata (Dublin core metadata standards), contents, and some other subdivisions, that report about organization of the contents and their related metadata.
- (2) **Presentation requirements.** Used to tackle a system's look-and-feel and statistical requisites.
- (3) **Access requirements.** These requirements may consist of both internal and external security availability. In internal availability, users need to get access to the contents and metadata whereas; external availability grabs other systems' requirements to get access to the contents and metadata.
- (4) **Preservation requirements.** These requirements may include storage, backup and long-life protection of issues for the contents and metadata.

Conclusion

The study concluded that LD content management system is a software system which keeps preserve, organize and disseminate the services of digital collections. By implementing the procedure of CM systems, SZMC considered its requirement of a specific system and used Wordpress as CMS. Various open source LD applications like DRUPAL, JOOMLA, Greenstone, Fedora, DSpace, etc. are available commonly used for CM. WordPress is also one of them to be used for managing serial contents digitally and effectively. The main objective of this case study has been achieved by publishing the contents of PSZMC by using LD technology and making available for everyone at the link www.proceedings-szh.com.

References

- Bai, H., & Cho, Y.-L. (2009). Structuring serials bibliographic relationships through ID linking. Paper presented at the International Conference on Dublin Core and Metadata Applications.
- Balster, K., Rendall, R., & Shrader, T. (2018). Linked Serial Data: Mapping the CONSER Standard Record to BIBFRAME. *Cataloging & Classification Quarterly*, 56(2-3), 251- 261.
- Béquet, G., & Oury, C. (2018). Revisiting the identification of serials: ISSN goes linked. *Insights*, 31(1).
- Berners-Lee, T., & Fischetti, M. (2001). *Weaving the Web: The original design and ultimate destiny of the World Wide Web by its inventor*: DIANE Publishing Company.
- Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The semantic web. *Scientific American*, 284(5), 28-37.
- Bizer, C., Heath, T., & Berners-Lee, T. (2009). Linked data-the story so far. *International journal on semantic web and information systems*, 5(3), 1-22.
- Bizer, C., Heath, T., & Berners-Lee, T. (2011). Linked data: The story so far. In *Semantic services, interoperability and web applications: emerging concepts* (pp. 205-227): IGI Global.
- Guerrini, M., & Possemato, T. (2013). Linked data: a new alphabet for the semantic web. *JLIS. it*, 4(1), 67.
- Kaplan, L. (2012). Laurie Kaplan, "The Future of Serials in a Linked Data World" (2012). *Proceedings of the Charleston LibraryConference*, 368-374.
- Krubu, D. E., Idhalama, O. O. U., & Omigie, C. (2017). Lecturers' perception of students' information literacy skills versus students' actual information literacy levels. *Information Impact: Journal of Information and Knowledge Management*, 8(2), 99-117.
- Patel, S. K., Rathod, V., & Prajapati, J. B. (2011). Performance analysis of content management systems-joomla, drupal and wordpress. *International Journal of Computer Applications*, 21(4), 39-43.
- Tharani, K. (2015). Linked Data in libraries: A case study of harvesting and sharing bibliographic metadata with BIBFRAME. *Information Technology and*

Libraries (Online), 34(1), 5.

Taylor, S., Jekjantuk, N., Mellish, C., & Pan, J. Z. (2013). Reasoning driven configuration of linked data content management systems. Paper presented at the Joint International Semantic Technology Conference.

Thompson, K., & Richard, J. (2013). Moving our data to the Semantic Web: Leveraging a content management system to create the Linked Open Library. *Journal of Library Metadata*, 13(2-3), 290-309.

Wright, R. (1997). The man who invented the web. *Time*, 149(20), 64-68.