7-10-2021

Integration of External Repositories with TemaTres for Easy Access of Related Terms

Sukumar Mandal

Department of Library and Information Science, The University of Burdwan,
sukumar.mandal5@gmail.com

Integration of External Repositories with TemaTres for Easy Access of Related Terms

Dr Sukumar Mandal
Assistant Professor & Teacher-in-Charge, Department of Library and
Information Science, The University of Burdwan
E-mail: sukumar.mandal5@gmail.com

Abstract

Generally, thesaurus construction and visual vocabulary are generated with the help of TemaTres. It shows the relation of different terms in view of a specific facet and sub facet. But this paper has explored extra facilities for users. This facility is a part of a cloud computing system. Now, this research paper has integrated the external repositories and software interface in both offline and online environments. How to integrate these external repositories with TemaTres? What are the metadata sets available for data interoperability and crosswalk? How to access these external repositories from the TemaTres metadata interface? However, this paper has been selected by the popular thesaurus construction and visual vocabulary software TemeTres for easy integration of external repositories regarding these specific questions. The whole process is developed and designed on the basis of configuration of files in TemaTres such as config.tematres.php and image icons. This integrated framework is very helpful to the users and librarians for easy access of thesaurus and visual vocabulary from different external repositories. Finally, this has created the common access interface of metadata for the users.

Keywords: TemaTres, Thesaurus, Visual Vocabulary, External Repositories, and Metadata formats

Introduction

The concept of search interface has changed in the modern repository management system. Digital library is a collection of electronic objects and metadata. Institutional digital repository is also an important aspect in an open access environment. This concept is also known as cloud computing and information mashup systems. There are a lot of search mechanisms available in an Online environment. Metadata search is one of the important tasks for an institutional digital repository. Now, this research paper has explored the integration method of different search repositories with TemaTres for easy access of thesaurus as well as metadata in view of availability from library search repositories. However, integrated twenty number of repositories with the TemaTres user interface such as BASE, CORE, DOAB, DOAJ, Google Busqueda, Google Books, Google images, Google scholar, Google search, OpenDOAR, PQDTOpen, Wikipedia, Shodhganga, DSpace, EPrints, Greenstone, Koha, Omeka, VuFind, and Open Library. Apart from these it also highlights and finds out the formats of different metadata sets such as BS8723-5, DC, MADS, SKOS-Core, VDEX, XTM, Zthes, JSON, and JSON-LD. This is the prototype integrated framework where librarians can easily use these concepts and ideas for the library users. Open access is a part of the online education system. Generally open access resources are accessed with the help of a particular repository, but here shows the new concepts in metadata search mechanism. Different software can be accessed using this integrated framework by applying the site URL and programming language of html, php, and java. Open access digital repositories have been accessed from the TemaTres user interface. It has created the common access window of metadata to search for different keywords which are available in other open

access digital repositories. This is the new tool and technique which helps the library users, how to access the similar metadata from these twenty repositories. A faceted search mechanism is achieved using this system. Modern libraries have been struggling to manage and access the digital resources from cloud computing environments. This problem is to be solved with the help of TemaTres. Generally it can manage the thesaurus and visual vocabulary. In any case, configure the files under TemaTres for easy access of metadata associated with these repositories. So, it can explore and generate the common access interface of different metadata sets and keywords. These twenty external repositories are shown in Figure-1 for integration with TemaTres towards providing better services among the users.

Figure-1: External repositories

Review of Literature

It is possible to implement and enhance the citation linker that enables users to search articles, journals, and books with minimal but sufficient citation metadata (Xu, 2010). It highlights and explores library metadata from usability and accessibility perspectives (Beyene & Godwin, 2018). Digital collections that have been catalogued using high-quality metadata can be retrieved more easily by users of search engines such as Google (Dawson & Hamilton, 2006). It can be really seen as the new "Library of Alexandria" for digital resources from the perspective of teachers, academics and researchers that might want to retrieve and reuse those resources for learning purposes (Gaona-Garcia, Sanchez-Alonso & Fermoso García, 2017). Authors' present efforts towards the development of a digital library environment supporting research at the Medical School of Athens University, Greece (Nikolaidou, Anagnostopoulos & Hatzopoulos, 2005). Metadata for Internet resource discovery, interoperability and digital preservation all point to the continuing need for something like traditional library services to organise, access and preserve networked information (Day, 1999). The Internet has heightened awareness among the information community of the need to provide user-friendly searching and navigation tools that lead to quality information (Heery, 1996). This is the new idea and concept of the functionality of metadata harvesting for regional digital libraries, based on the OAI-PMH protocol (Mazurek...et.al., 2006). Identify these search patterns within a well-curated historical newspaper collection using the existing metadata (Bogaard...et.al, 2019). Explain and describe the wide dissemination of Library of Congress Subject Headings within digital libraries and present some of the advantages and disadvantages of using this controlled vocabulary in digital collections (Walsh, 2011). It discusses the future improvements and potential of OAIster and the OAI-PMH protocol (Hagedorn, 2003). The paper proposed the potential of accessibility metadata in improving knowledge discovery and access in digital library environments, discussed developments in creating accessibility terms for resource description, and attempted to relate those developments to the overall purpose of universal design to finally recommend points for improvement (Beyene, 2017). Overview the genesis of the project, the rationale for architectural design decisions, challenges faced, and our progress to date (Cole & Shreeves, 2004). Emerging trends of metadata practice and standards overview of the environments in which metadata is used, before focusing on metadata for information resources (Dempsey & Heery, 1998). These Discussions represent appropriate technologies and sustainable strategies with the help of open source software that will help Internet portals, digital libraries, virtual libraries and library catalogs to scale better and to anticipate and meet the needs of scholarly and educational users (Mitchell, Mason & Pender, 2004). Understanding and exploring the issues surrounding the cataloguing of maps in archives and libraries. An investigation into appropriate metadata formats, such as MARC21, EAD and Dublin Core with RDF, shows how particular map data can be stored (Beamer, 2009). Transforming metadata schemas in institutional repositories will lead to increased indexing by Google Scholar (Arlitsch & O'Brien, 2012). Identification of criteria for the evaluation and integration of visual search interfaces, proposing guidelines and recommendations to improve information retrieval tasks with emphasis on the education-al context (Gaona-García, Martin-Moncunill & Montenegro-Marin, 2017).

Objectives

Objectives of this paper are as follows:

(i) To explore the integration methods of external repositories with TemaTres.
(ii) To show the thesaurus construction and visual vocabulary interface.
(iii) To highlight and overview the metadata search results from external repositories.

(iv) To display the search icons of different search repositories for easy access of related terms.

Methods of Integration

This is a very easy process to integrate the popular external repositories with TemaTres for easy access of metadata. Now the library professionals are very much benefitted by using these techniques to easily create the thesaurus and visual vocabulary. Apart from these, it also accesses the similar bibliographic metadata from other library repositories mentioned in Figure-1. However, it is required to configure the three important files in TemaTres such as (i) config.tematres.php (ii) config.tematres.php (iii) images. Now, to need the location of these files in TemaTres, in the first case the folder path is (var/www/html/tematres/vocab/config.tematres.php). The second file is located under *var/*www/html/tematres/common/include/config.tematres.php. The third file is located in *var*/www/html/tematres/common/images. Customize and configure the two files under TemaTres in terms of vocab and common are as follows:

config.tematres.php
```
// ID del Tesauro por DEFAULT
$CFG["DFT_TESA"] ='1';
//Config Sites availables for URL search
$CFG["SEARCH_URL_SITES"]
=array("wikipedia","Google exacto","Google
scholar","Google images","Google books","Koha
OPAC","DSpace IDR@BU","EPrints
IDR@BU","Omeka IDR@BU","VuFind
WSD@BU","Google","DOAB","PQDTOpen","D
OAJ","OpenDOAR","BASE","Shodhganga","CO
RE","Open Library");
```

config.tematres.php
```
'DOAB' => Array('favicon' => 'd.png',
                'leyenda' => 'DOAB',    'url' =>
'https://www.doabooks.org/doab?func=search&ui
Language=en&template=&query=STRING_BUS
QUEDA',
                'encode'=>FALSE
        ), 'PQDTOpen' => Array( 'favicon' =>
'p.png', 'leyenda' => 'PQDTOpen',
                'url' =>
'https://pqdtopen.proquest.com/results.html?QryT
xt=STRING_BUSQUEDA',
                'encode'=>FALSE        ),
```

Thesaurus Construction Interface

This is the thesaurus construction interface in TemaTres. It is possible to create a new thesaurus regarding different subjects and concepts. Metadata options are not available in this interface. It is only used for new thesaurus construction. Figure-2 represents the thesaurus construction interface in TemaTres. Relationships of metadata can be displayed in two forms such as candidate term and meta term. Vocabulary configuration is possible under administration. However, it shows other important parameters under administration are the bulk editor, users, export, and import. It is also possible to update the databases on the basis of thesaurus structure. Multilingual thesaurus can be made with the help of a multilingual editor.
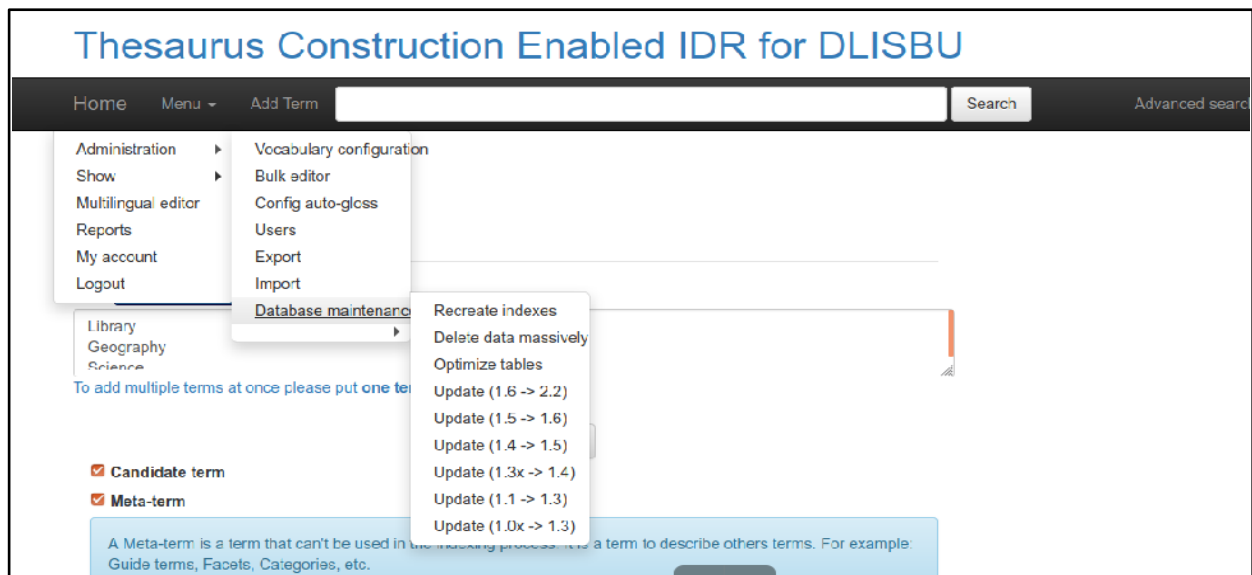
Figure-2: Thesaurus construction interface

Integration Interface of TemaTres

After successfully configure the different files, the metadata access interface will appear where integrated all the repositories such as BASE, CORE, DOAB, DOAJ, Google Busqueda, Google Books, Google Images, Google Scholar, Google search, OpenDOAR, PQDTOpen, Wikipedia, Shodhganga, DSpace, EPrints, Greenstone, Koha, Omeka, VuFind, Open Library. So, all the repositories can easily be accessed from the interface of TemaTres (Figure-3).Users are very much benefitted using this integrated framework.

Figure-3: Integration interface of TemaTres

Metadata Formats

It shows the nine metadata formats in TemaTres for access to different terms related to metadata. These nine metadata formats are BS8723-5, DC, MADS, SKOS-Core, VDEX, XTM, Zthes, JSON, and JSON-LD.

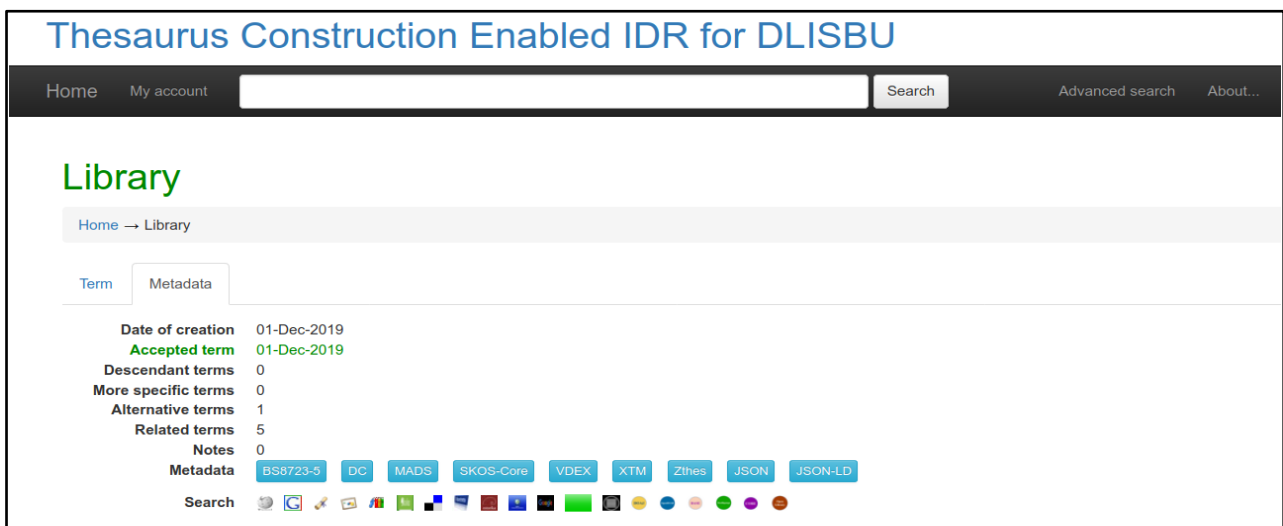| (i) | BS8723-5 | : | Structured Vocabularies for Information Retrieval. It was created specifically to support Part 5 of the standard, known as DD 8723-5, dealing with Exchange formats and protocols for interoperability. |
(ii)	DC	:	Dublin core metadata standard for easy creation of thesaurus and is possible to retrieve among users.
(iii)	MADS	:	Metadata authority description schema used for thesaurus construction and visual vocabulary in view of data interoperability and crosswalk from one system to another.
(iv)	SKOS-Core	:	Simple knowledge organization system core is a metadata standard used for semantic web in different terms.
(v)	VDEX	:	Vocabulary definition exchange is a metadata set used for thesaurus construction.
(vi)	XTM	:	Knowledge maps are created using this XTM format for hierarchy structures.
(vii)	Zthes	:	It provides interoperability for applications that deal with thesauri - semantic hierarchies of terms as described in ISO 2788 and ANSI/NISO Z39. 19. The core specification in the Zthes suite is an abstract model for thesaurus terms.
(viii)	JSON	:	Controlled vocabulary server based on JSON web services to synchronize and notify changes.
(ix)	JSON-LD	:	Linked data facilities have been achieved using the JSON-LD metadata format for creation of controlled vocabularies.

Access Related Terms from TemaTres

Related terms have been accessed in two ways such as software related integration and external repositories integration. It is possible to access local level related terms from Koha, DSpace, Omeka, EPrints, Greenstone, VuFind. This is known as software level integration. On the other hand, external global repositories have been accessed with the help of TemaTres. External repositories are BASE, Google Busqueda, Google Search, Shodhganga, CORE, Google Books, OpenDOAR, DOAB, Google Images, PQDTOpen, DOAJ, Google Scholar, Wikipedia, etc. Apart from these, it is also possible to integrate other repositories with TemaTres by applying the same techniques and concepts. Access metadata from the TemaTres interface by clicking on any icon.

Results of Metadata Formats

Different metadata formats are in view of thesaurus construction and visual vocabulary. Now, the results of metadata formats in 'Library' has been represented as on the basis of metadata formats and their ports in TemeTres thesaurus construction are represented as follows:

(i)	BS8723-5	:	http://localhost/tematres/vocab/xml.php?bs8723Tema=16
(ii)	DC	:	http://localhost/tematres/vocab/xml.php?dcTema=16
(iii)	MADS	:	http://localhost/tematres/vocab/xml.php?madsTema=16
(iv)	SKOS-Core	:	http://localhost/tematres/vocab/xml.php?skosTema=16
(v)	VDEX	:	http://localhost/tematres/vocab/xml.php?vdexTema=16
(vi)	XTM	:	http://localhost/tematres/vocab/xml.php?xtmTema=16
(vii)	Zthes	:	http://localhost/tematres/vocab/xml.php?zthesTema=16
(viii)	JSON	:	http://localhost/tematres/vocab/xml.php?jsonTema=16
(ix)	JSON-LD	:	http://localhost/tematres/vocab/xml.php?jsonldTema=16

These metadata schemas and formats are very important for easy management of big data science

towards present and next level automated and digital library systems. These schemas are very helpful to the library users community as well as library professionals for easy access of similar metadata from other repositories.

Visual Vocabulary Interface

Graph visualization is an attractive interface in a big data environment. It gives the distinctive competence of graph data that is very easily shown in a visual format and provides additional value for data analysts towards library users. It is possible to review the graph which represents the connections, determine areas of interest, narrower terms, broader terms, and related terms which insights the anomalies. It works the same as human eyes and brains. There are many tools and techniques available but this paper has selected the TemaTres visual vocabulary plugin. So, obviously librarians are required to have a lot of knowledge in computers, but in reality it's not possible. This is the extra facilities provided by TemaTres in graphical form. This is a very new and innovative interface for the creation of visual vocabulary. Figure-4 is shown as a visual vocabulary interface for users as well as library professionals.

Figure-4: Visual vocabulary interface

Now, visual vocabulary shows the narrower term, broader term, related term, and adjacent. The Figure-5 represents the broader and narrower term with the help of TemaTres visual vocabulary.
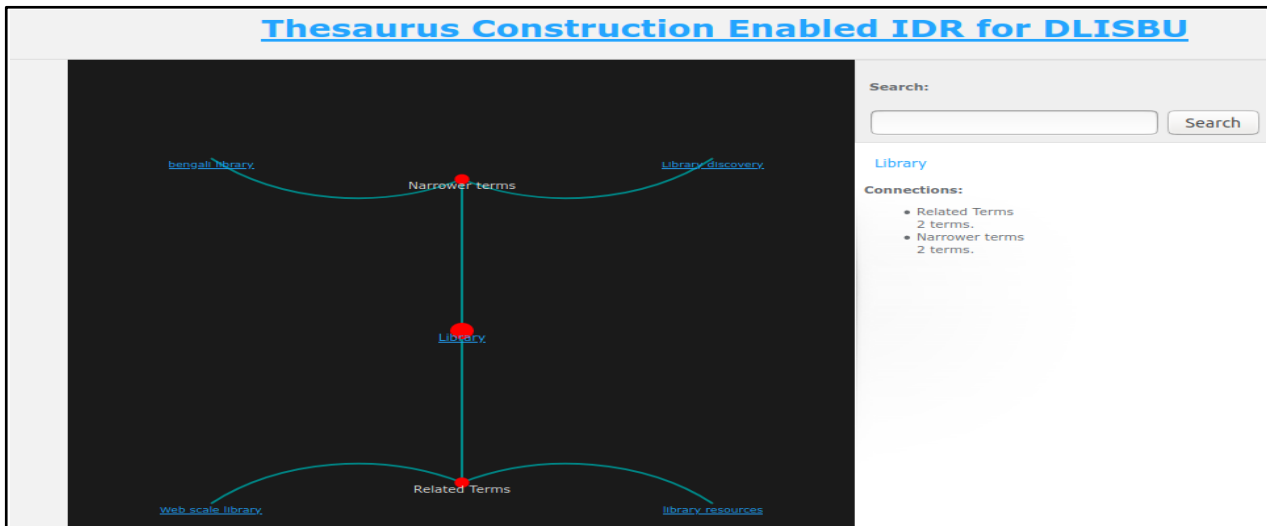
Figure-5: Broader , narrower, related, and adjacent terms

Significance of Integration Interface for Users

Library information technology is an important strand for an open access digital environment. This integrated domain specific framework consists of two sets of parameters such as (i) software integration in an offline environment and (ii) integration of external repositories for easy management of related terms including broader, narrower, adjacent. So, big data science can easily be managed with the help of TemaTres and Visual Vocabulary. It is possible to manage graph visualization both in offline and online environments through availability of related terms in the form of metadata such as METS, MODS, MARC21, and etc. However, library professionals and advanced users are very much attracted to this integrated interface. Bibliographic metadata can easily be accessed from this integrated framework for creation of a common access search window.

Conclusion

This is the common access integrated interface known as the metadata access framework. It is possible to access metadata from external repositories. It is provided in two ways such as software level access and open access educational repositories. Metadata can easily be accessed from different

software such as Koha OPAC, DSpace, EPrints, Greenstone, Omeka with the help of TemaTres metadata search interface. Open access repositories or external repositories are also accessed using this system and technique. This is very helpful to the users in view of the thesaurus construction and visual vocabulary. This is the new innovative technique for easy access of metadata. So, it is clear that TemaTres not only provides the thesaurus regarding the broader term, narrower term, related term, use for, scope note, and etc. Apart from these it also retrieves the metadata against on terms available in TemaTres thesaurus interface. This approach can save the time of the reader and library professionals also. It is possible to display the nine sets of metadata formats. Data interoperability and crosswalk are possible with the help of nine metadata sets mentioned in the section of metadata formats. Users can easily access the metadata from external repositories with the help of TemaTres, just by clicking on an icon in this common integrated framework. However, the total concept is based on programming language but librarians are very benefitted using this prototype system for easy access and retrieving of digital resources among the users.

References

Arlitsch, K. and O'Brien, P.S. (2012). Invisible institutional repositories: Addressing the low indexing ratios of IRs in Google Scholar. *Library Hi Tech*, 30 (1), 60-81. https://doi.org/10.1108/07378831211213210 (Accessed on November 19, 2020)

Beamer, A. (2009). Map metadata: essential elements for search and storage. *Program: electronic library and information systems*, 43 (1), 18-35. https://doi.org/10.1108/00330330910934084 (Accessed on November 5, 2020)

Beyene, W.M. and Godwin, T. (2018). Accessible search and the role of metadata. *Library Hi Tech*, 36 (1), 2-17. https://doi.org/10.1108/LHT-08-2017-0170 (Accessed on November 8, 2020)

Beyene, W.M. (2017). Metadata and universal access in digital library environments. *Library Hi Tech*, 35 (2), 210-221. https://doi.org/10.1108/LHT-06-2016-0074 (Accessed on November 8, 2020)

Bogaard, T., Hollink, L., Wielemaker, J., van Ossenbruggen, J. and Hardman, L. (2019). Metadata categorization for identifying search patterns in a digital library. *Journal of Documentation*, 75 (2), 270-286. https://doi.org/10.1108/JD-06-2018-0087 (Accessed on November 12, 2020)

Cole, T.W. and Shreeves, S.L. (2004). Search and discovery across collections: the IMLS digital collections and content project. *Library Hi Tech*, 22 (3), 307-322. https://doi.org/10.1108/07378830410560107 (Accessed on November 12, 2020)

Dawson, A. and Hamilton, V. (2006). Optimising metadata to make high-value content more accessible to Google users. *Journal of Documentation*, 62 (3), 307-327. https://doi.org/10.1108/00220410610666484 (Accessed on November 17, 2020)

Day, M. (1999). Resource discovery, interoperability and digital preservation: some aspects of current metadata research and development. *VINE*, 29 (4), 35-48. https://doi.org/10.1108/eb040731 (Accessed on October 10, 2020)

Dempsey, L. and Heery, R. (1998). Metadata: a current view of practice and issues. *Journal of Documentation*, 54 (2), 145-172. https://doi.org/10.1108/EUM0000000007164 (Accessed on October 13, 2020)

Gaona-García, P.A., Martin-Moncunill, D. and Montenegro-Marin, C.E. (2017). Trends and challenges of visual search interfaces in digital libraries and repositories. *The Electronic*

Library, 35 (1), 69-98. https://doi.org/10.1108/EL-03-2015-0046 (Accessed on October 16, 2020)

Gaona-Garcia, P.A., Sanchez-Alonso, S. and Fermoso García, A. (2017). Visual analytics of Europeana digital library for reuse in learning environments: A premier systematic study. *Online Information Review*, 41 (6), 840-859. https://doi.org/10.1108/OIR-04-2016-0114 (Accessed on October 24, 2020)

Hagedorn, K. (2003). OAIster: a "no dead ends" OAI service provider. *Library Hi Tech*, 21 (2), 170-181. https://doi.org/10.1108/07378830310479811 (Accessed on October 27, 2020)

Heery, R. (1996). Review of metadata formats. *Program: electronic library and information systems*, 30 (4), 345-373. https://doi.org/10.1108/eb047236 (Accessed on October 21, 2020)

Mazurek, C., Stroinski, M., Werla, M. and Węglarz, J. (2006). Metadata harvesting in regional digital libraries in the PIONIER network. *Campus-Wide Information Systems*, 23 (4), 241-253. https://doi.org/10.1108/10650740610704117 (Accessed on November 16, 2020)

Mitchell, S., Mason, J. and Pender, L. (2004). Enabling technologies and service designs for collaborative Internet collection building. *Library Hi Tech*, 22 (3), 295-306. https://doi.org/10.1108/07378830410560099 (Accessed on November 18, 2020)

Nikolaidou, M., Anagnostopoulos, D. and Hatzopoulos, M. (2005). Development of a medical digital library managing multiple collections. *The Electronic Library*, 23 (2), 221-236. https://doi.org/10.1108/02640470510592933 (Accessed on November 12, 2020)

Xu, F. (2010). The SFX citation linker and its enhancements. *Library Hi Tech*, 28 (3), 377-390. https://doi.org/10.1108/07378831011076639 (Accessed on November 7, 2020)

Walsh, J. (2011). The use of in digital collections. *Library Review*, 60 (4), 328-343. https://doi.org/10.1108/00242531111127875 (Accessed on November 14, 2020)