

**DESIGNING AND DEVELOPING DIGITAL CONTENT
MANAGEMENT SYSTEM THROUGH OPEN SOURCE
SOFTWARE AND STANDARDS**

Dr. Sukumar Mandal

Assistant Professor, Department of Library and Information Science

The University of Burdwan, Burdwan – 713 104

Email – sukumar.mandal5@gmail.com

Abstract

Conventional cataloguing focuses mainly on print materials for organizing and disseminating information. In recent decades the digital contents are appearing in an accelerating rate and thus information scientists are forefronting the problems relating to their bibliographical description and management. It provides metadata related information like creator, title, language and other parameters based on open source standards. While typically the cataloguing process results in the production of library catalogues, and thus the need arises towards making varieties of discovery tools for resource management. The paper shows how digital content management and image cataloguing can be done by using open source standards and open source software. Single window based interface is one of the important aspects in the automated and digital library system. In this practical experimentation the open source software Omeka and the open standards like Visual Resource Association (VRA) core and Dublin Core have been used. Apart from these, advanced level parameters for designing and developing the image cataloguing and digital resource management in different libraries have been used. Thus the paper has tried to solve the problems of image cataloguing in museums, libraries and other such organizations. The whole system and relevant standards have to be generated by using Ubuntu operating system because of its high security and reliability qualities.

Keywords: VRA Core, Dublin Core, Image Cataloguing, Omeka, Reference Management

1 Introduction

Cataloguing is one of the crucial activities in any type or size of a library. Most of the libraries are doing manual cataloguing, but some libraries are using machine generated cataloguing systems (Aspeli, 2007). However, image cataloguing still remains one of the important problems in all types of libraries. This paper attempts to solve this lingering problem. This image cataloguing process, is equally helpful for digital resources as well as its management which recently a librarian faces a great challenge to organize this digital resources as well as to maintain and management the user interfaces (Hsu, Ke, Yang, 2006). Majority of the

parameters are extensively useful for managing both the digital documents and images such as neatline map, open archive initiative for protocol metadata harvesting, subject list, photos, references, map, collection tree (Arthur, 2013). The different computing parameters such as Google Books, Amazon Book cover images, archive grid, digital public library of America, Google Scholar, Hathi trust, JSTOR, Library of Congress and WorldCat are the means for managing the user interfaces and also the cloud computing can be used equally for managing user interfaces (Seadle, 2011). In this regard open source formats such as atom, dc-rdf, dcmes-xml, json, omeka-json, omeka-xml and vra can be used the vocabularies and other concepts for visual formats (White, 2002). Browsing the collections in user interfaces of different items and their bibliographical description can be regularly managed through this image cataloguing technique. In this process metadata can also be managed because of the fact that the software that would be used for image cataloguing is prepared in PHP programming language (Goans, Leach, & Vogel, 2006). For management of metadata and fulltext document both the add and delete function can be operated in this windows. The Omeka open source software can be applied to execute the total process and this software is competent enough for image cataloguing, searching bibliographical data, browsing as well as metadata collection trees with fulltext materials of various file formats such as pdf, html, jpeg, audio, videos and other formats and browse including tags, search items, subject and references (The Content Management Handbook, 2005). User will get the benefit to narrow the search field such as Dublin Core and Visual Resource Association (VRA) Core for development of digital collections in search item fields. The Omeka open source image cataloguing software which is devised by George Mason University can equally extend the facility of single window based search system.

2 Objectives

The objectives of this research paper are:

- a) To explore the possibilities of using open source software Omeka for the management of digital resources.
- b) To explore the usability of open source standards like Visual Resource Association and Dublin Core metadata element sets in cataloguing.
- c) To find out the advanced level parameters in Omeka for easy management and retrieving of documentary resources – both regional and multilingual.
- d) To explore the possibility of incorporating all the advanced level parameters in a single window of Omeka on Ubuntu operating system for image cataloguing and digital content management.

3 Methodology

The methodology used in this study is very simple, systematic and practical. When Omeka website is opened, two sets of files are found – Windows files and Ubuntu files. Among them, Ubuntu zip file format was selected for easy installation and configuration of Omeka and the Omeka folder was kept under the web server in Ubuntu (e.g. /var/www/html). Then the Image Magic software was installed as it is one of the Omeka supported software. It may be mentioned that Omeka is based on LAMP architecture (L for linux operating system; A for Apache web server; M for MySQL database structure and P for PERL i.e. programming environment report language or practical extraction report language). These software can be easily installed from the Tasksel menu in Ubuntu operating system. For smooth running of Omeka some changes in the configuration for the two files such as apache2 and sites available under the Apache2 web server can be changed as per requirement which supports the best processing of image cataloguing in the firefox web browser. In the next step a permission is to be made so that Omeka folder under the web server can work for apache linking and SQL database connectivity. Then terminal to be open and creation of new user and database with the help of SQL based commands as well as new configuration is to be made i.e db.ini file for user name, database name, and database password. In this final step Omeka installation option can be run from the Mozilla Firefox web browser and this to be written <http://localhost/omeka/install/install.php> which is illustrated through the Figure – 1 and in this stage admin login and password to be used for managing and updating the new documents or digital resources in Ubuntu operating system.



Figure – 1: Omeka installation interface

4 OSS for Image Cataloguing and Digital Content Management

Image cataloguing and digital content management is possible by using the open source standards and open source software available in the field of automated and digital library system. In this paper two standards such as Visual Resource Association (VRA) Core and Dublin Core have been used for image cataloguing and digital content management in Omeka open source software. Obviously Omeka can solve the problem of image cataloguing and cataloguing their full-text versions. But it is difficult manage different metadata fields and their sub-fields using VRA. But in the present experimentation this problem has been successfully solved. Apart from this, it also managed metadata standards like Dublin core very known to us. But here image cataloguing lot of fields are included in Dublin Core schema for easy and efficient management of regional and multilingual resources both in admin and user interfaces. These two standards are very user-friendly in Omeka open source software.

4.1 Visual Resource Association (VRA) Core

VRA i.e. Visual Resources Association was developed during 1982 in North America. It helps in easy creating, managing and retrieving the metadata of different digital images and other media. The purpose of this association is easy and smooth management of images in the field of education, commerce and cultural heritage. Publish the variety of resources and also offers the educational programs for the users in different subject fields (Molen, 2011). It can perform different tasks including preservation, analog resources of visual culture and access of different digital resources with regional languages. It can easily manage different types of library operations like cataloguing, classification and integration of technology including citation styles of different subject fields. Researchers and advanced level users can be benefited by using this standard and techniques. Management of broader and narrower information which support the different file formats of visual information in library related resources and understanding the experience of cultural heritage. Different types of standard are available in online environment, but here only VRA core methods and techniques have been used because these help to easily manage images and their cataloguing. All the VRA core parameters and their fields and subfields have been included in this system. Table – 1 represents the VRA core elements of image cataloguing and uploading of full-text documents in both regional and other languages because it is a Unicode based standard.

Table – 1: Parameters of VRA core in Omeka

ID	VRA Work	State Edition	Display Attributes Notes State Edition Attributes and Sub-elements
Title	Display Attributes Notes Title Elements	Style Period	Display Attributes Notes Style Period elements
Agent	Display Attributes Notes Agent Attributes and Sub-elements	Subject	Display Attributes Notes Subject Attributes and Sub-elements
Cultural Context	Display Attributes Notes Cultural Context Elements	Technique	Display Attributes Notes Technique Elements
Date	Display Attributes Notes Date Attributes and Sub-elements [YYYY-MM-DD]	Textref	Display Attributes Notes Textref Attributes and Sub-elements
Description	Display Attributes Notes Description Elements	Worktype	Display Attributes Notes Worktype Elements
Inscription	Display Attributes Notes Inscription Attributes and Sub-elements	Relation	Display Attributes Notes Relation Elements
Location	Display Attributes Notes Location Attributes and Sub-elements	Rights	Display Attributes Notes Rights Attributes and Sub-elements
Material	Display Attributes Notes Material Elements	Source	Display Attributes Notes Source Attributes and Sub-elements
Measurements	Display Attributes Notes Measurements Elements		

4.2 Dublin Core

Dublin core standard is also an important standard for image cataloguing and full-text document cataloguing. Dublin core was developed during 1995 in USA at the metadata workshop hosted by the Online Computer Library Center (OCLC) library consortium and the National Center for Supercomputing Applications (NCSA) (International Conference on Dublin Core and Metadata

Applications., & Dublin Core Metadata Initiative, 1995). It is a small set of vocabulary terms and their relationships which can be used to represent the web related resources including videos, images, web pages, photos, file formats and also the physical resources like CDs, DVDs, books, and other objects, like artworks, and other link identifier. The details of metadata sets can be found on Dublin Core Metadata Initiative website and the original fifteen sets of metadata elements known as the Dublin Core Metadata Element Set have been included in different standards like IETF RFC 5013, ISO Standard 15836-2009 and NISO Standard Z39.85. It can be used for multiple purposes like simple resource description to combining the metadata items vocabularies; to providing the interoperability and crosswalk of vocabularies in the Linked Open Data (LOD) and Semantic Web applications in cloud computing environment. The Omeka open source software use the DCMI metadata terms which is the current set of Dublin Core vocabulary consisting of fifteen terms plus the qualifier terms (Table - 2) and each term has a unique URI that can manage the RDF properties also.

Table – 2: Parameters of Dublin Core qualified metadata elements in Omeka

Title	Is Part of
Subject	Is Referenced By
Description	Is Replaced By
Creator	Is Required By
Source	Is Version Of
Publisher	References
Date	Replaces
Contributor	Requires
Rights	Extent
Relation	Medium
Format	Bibliographic Citation
Language	Spatial Coverage
Type	Temporal Coverage
Identifier	Accrual Method
Coverage	Accrual Periodicity
Alternative Title	Accrual Policy
Abstract	Audience
Table of Contents	Audience Education Level
Date Available	Mediator

Date Created	Instructional Method
Date Accepted	Provenance
Date Copyrighted	Rights Holder
Date Submitted	License
Date Issued	Conforms To
Date Modified	Has Format
Date Valid	Has Part
Access Rights	Has Version
	Is Format of

5 Advanced Level Parameters

The Omeka consists of many advanced level parameters that can help the students, teachers, researchers and also advanced users. The important, necessary and most relevant parameters can be installed successfully in this system. Actually these parameters are related with plugin in Omeka for the management of digital resources and their different functions. These parameters are:

- (i) **Zotero** : This parameter manages the references and their citation both in admin and user interfaces (Puckett, 2011). This open source software is a reference management software and this can be incorporated in Omeka interfaces.
- (ii) **Tags** : Tag means tagging the item which can help to create the alternative term against in one term under the collection.
- (iii) **Fedora Connector** : This plugin makes it possible to connect items in Omeka collection with objects in Fedora Commons repositories and automatically generate the Dublin Core records for the items.
- (iv) **Map** : After integrating the geo-location plugin in Omeka, the map field or option appears in admin interface. It displays the country map in respect of different items in Omeka collections.

- (v) **Catalog Search** : It searches the records in other catalogues which are available on different websites like Archive Grid, the DPLA, Google Books, Google Scholar, the Hathi Trust, JSTOR, the Library of Congress, and WorldCat.
- (vi) **Admin Images** : Omeka 2.x allows uploading of images on different sample pages and their cataloguing in respective fields with sub-fields. It can easily manage the images for different item types.
- (vii) **Bulk Metadata Editor** : It allows to search, add search and replace the curators to update the different metadata fields and sub-fields of item records easily and quickly in Omeka admin interface and sometimes user interface also.
- (viii) **Collection Tree** : This is one of the important parameters in Omeka because it can manage the different terms like group, sub-group, section and sub-sections, division, sub-division also. Users can easily see their necessary resources from the Omeka user interface.
- (ix) **CSV Import** : Imports the CSV file from the Omeka admin interface in different metadata element sets because it fully supports the CSV format.
- (x) **Derivative Images** : Create or recreate the derivative images like thumbnails or other formats in Omeka. This facility is available in Omeka admin interface.
- (xi) **Flickr Import** : Imports the Flickr photosets and galleries into Omeka admin interface of different collections and their item types for preserving the metadata.
- (xii) **OAI-PMH Harvester** : Harvests the metadata elements of different items from the OAI-PMH data providers in Omeka admin interface (Cole & Foulonneau, 2007). This

- is the great plugin which can manage the harvested documents available in online environment.
- (xiii) **Ngram** : It can generate the ngram related graphs from the existing Omeka collection in admin interface and user interface also.
- (xiv) **Reports** : Reports plugin allows the users to create reports from admin interface in Omeka collections of different digital items. It can also generate the barcode for identifying the HTML and PDF formats in Omeka.
- (xv) **Simple Vocab** : This helps in creating the controlled vocabularies of different items and collections in Omeka.
- (xvi) **Solr Search** : It is possible in Omeka to just connect to solr port for accessing the solr documents by configuring the server host, server port and URL.
- (xvii) **Neatline** : It is a visual resource of lightweight, extensible framework for creating the different items – paintings, maps, images, photographs and other materials.

6 Single Window Based Interface

Advanced level parameters in Omeka for the designing and developing the image cataloguing in content management system and digital resource management on Ubuntu operating system has been discussed in section 1.4. Figure – 2 represents the single window based interface in Omeka which incorporates all the advanced level parameters. It may be mentioned that Omeka uses a of parameters that can manage the references, citations, solr search, harvesting from the other server, subject list, browsing, csv import, zotero integration and many other activities.

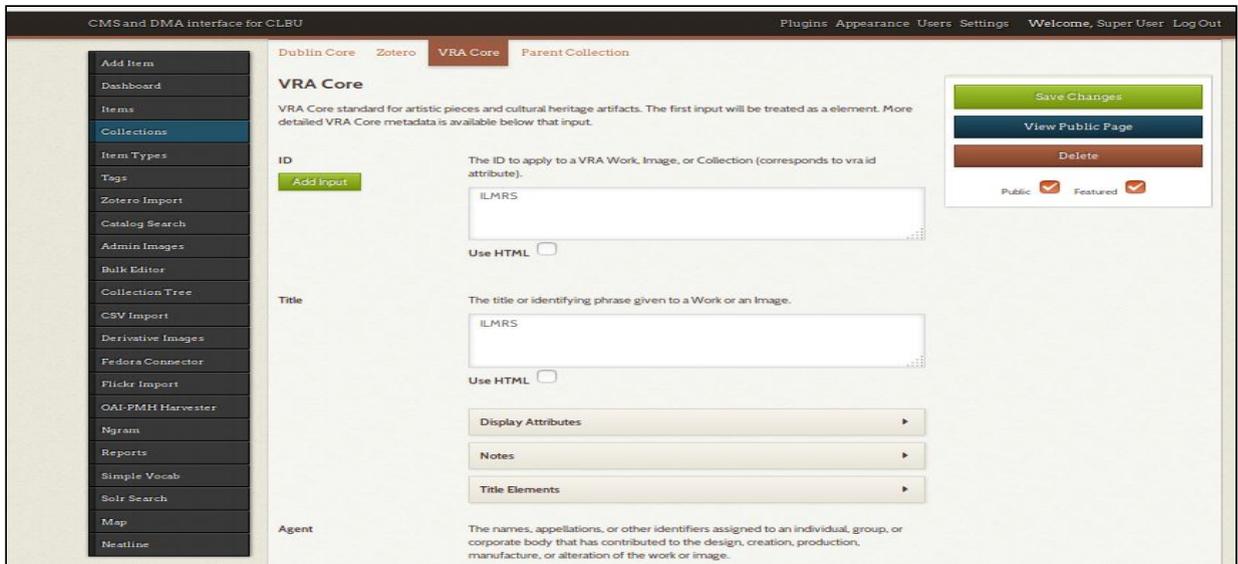


Figure – 2: Single Window based interface in Omeka

7 User Interface for Accessing Image and Digital Resources

This is the innovative user interface (Figure - 3) in Omeka which can access images with description of metadata and full-text resources. Browse the subjects in alphabetical and here advanced search facilities is also available in different item types with collections.



Figure – 3: User interface in Omeka

7.1 Collection Tree

Figure – 4 represents the collection tree interface in Omeka through which users can access the collection accordance with the collection against in a specific item types. Most of the users in libraries can access their necessary documents and display the metadata both in Dublin Core format and Visual Resource Association (VRA) Core format.

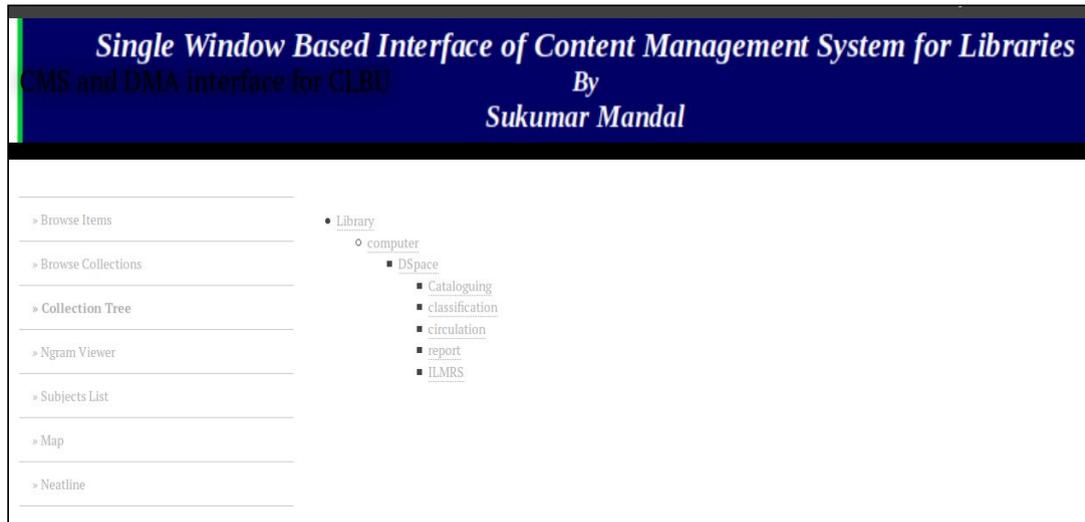


Figure – 4: Collection tree interface in Omeka

7.2 Subject List

The alphabetical arrangement of subject list can be done very efficiently in Omeka user interface. Figure – 5 represents the subject list interface in Omeka and also found the metadata and different file formats. Library of Congress (LoC) gives different subject headings lists in library OPAC. This study also explores how to managed the subject list in user interfaces as same as LoC (Library of Congress, 2009). WorldCat subject searching is also possible against a specific term because it fully supports the cloud computing based services for the different users like teachers, students, researchers and etc.

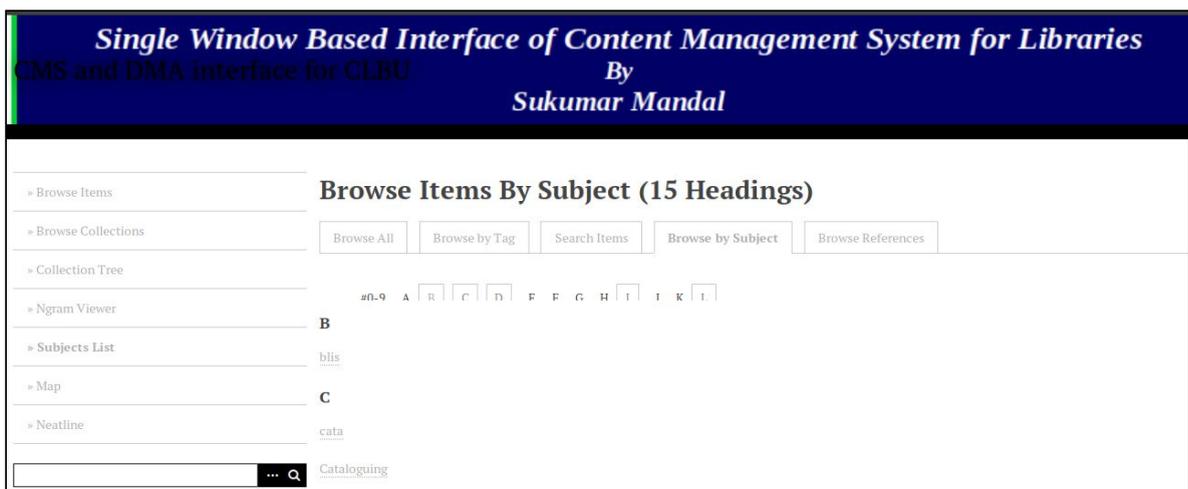


Figure – 5: Subject list interface in Omeka

7.3 View Map

Geolocation can be found from the MAP view interface in Omeka. Each and every items is to identify the correct location against in every collection. The Map interface in Omeka is represented in Figure – 6. Searching can be done country wise and satellite wise for geolocation of a particular resource. This is the most comprehensive tool for users as well as libraries because its source code is fully open and anybody can download, install and configure it in accordance with the installation instruction.

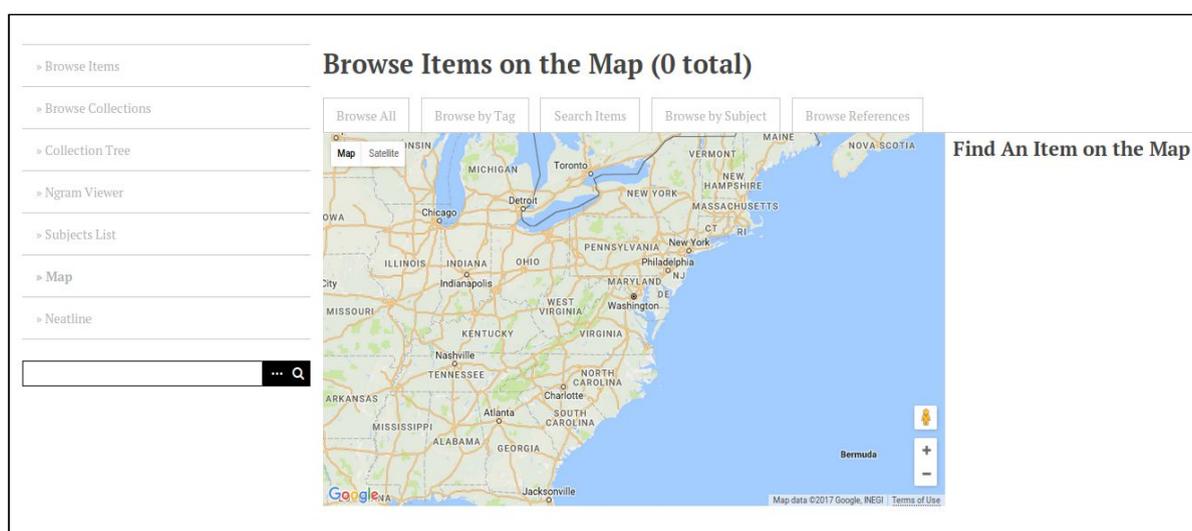


Figure – 6: Map interface in Omeka

8 Findings

The important findings of the study are stated as follows:

- (i) It is possible to easily manage the references in user interface for any collection or any digital item of different variety.
- (ii) It is possible to integrate all the advanced level parameters in a single window based interface.
- (iii) It is possible to retrieve the information alphabetically.
- (iv) It is possible to harvest the records from other servers.
- (v) Solr search can also be performed from the single window interfaces in Omeka.

- (vi) Subject tagging is also possible from the user interface and admin interface in Omeka.
- (vii) Graph display is also possible through the ngram in Omeka.
- (viii) BibTex import is possible through Zotero in Omeka.
- (ix) CSV import can be made in Omeka admin as well as user interface.
- (x) Fedora connect is possible in Omeka through connector from the server port.
- (xi) Report generation and barcode generation can be made from Omeka interface.
- (xii) Generate the simple vocabulary in Omeka admin interface and retrieving through user interface.
- (xiii) Display the country map in collection and item types of Omeka.
- (xiv) Neatline display can be made from the Omeka user interface.

9 Conclusion

Today web publishing is one of the important activities of any institution or library. Image publishing is a part of such publishing that can enter data and these data can be searched through user interfaces. It gives the programmers opportunity to generate and customize plugins in Omeka, which provides multiple features. From the above discussion, it is clear that the web publishing is possible through the Omeka user interface. It has been shown above how using two open source standards, VRA Core and qualified Dublin Core, a single window based interface can be designed and developed where all the advanced level parameters can be integrated. Researchers can easily access their citation against any item types of Omeka user interfaces. The whole system can be developed in Ubuntu operating system because of its high security and reliability qualities. This integrated framework can be used in any institution for the better management of image cataloguing with metadata for each item. Cloud computing and information mashup is also possible through user interface because it fully supports cloud based services. It is very user friendly use as it consists of a lot of readymade plugins, which just require customization and configuration. Multilingual documents can also be managed in different museums and other institutions by using this open source software and the open standards.

References

1. Arthur, M. H. (2013). *Expanding a Digital Content Management System: For the Growing Digital Media Enterprise*. Hoboken: Taylor and Francis.
2. Aspeli, M. (2007). *Professional Plone development: Building robust, content-centric web applications with Plone 3, an open source content management system*. Birmingham: Packt.
3. Cole, T. W., & Foulonneau, M. (2007). *Using the Open Archives Initiative protocol for metadata harvesting*. Westport, Conn: Libraries Unlimited.
4. Goans, D., Leach, G., & Vogel, T. M. (2006). Beyond HTML: Developing and re-imagining library web guides in a content management system. *Library Hi Tech*, 24(1), 29–53.
<https://doi.org/10.1108/07378830610652095>
5. Hsu, T.-Y., Ke, H.-R., & Yang, W.-P. (2006). Unified knowledge-based content management for digital archives in museums. *The Electronic Library*, 24(1), 38–50.
<https://doi.org/10.1108/02640470610649236>
6. International Conference on Dublin Core and Metadata Applications., & Dublin Core Metadata Initiative. (1995). *DC-: Proceedings of the International Conference on Dublin Core and Metadata Applications*. Madrid: Universidad Carlos III.
7. Library of Congress. (2009). *Library of Congress subject headings: Free-floating subdivisions, genre/form headings, children's subject headings*. Washington, D.C: Library of Congress, Cataloging Distribution Service.
8. Molen, F. (2011). *Get ready for cloud computing: A comprehensive guide to virtualization and cloud computing*. Zaltbommel: van Haren Publishing.
9. Seadle, M. (2006). Content management systems. *Library Hi Tech*, 24(1), 5–7.
<https://doi.org/10.1108/07378830610652068>
10. *The Content Management Handbook*. (2005). *New Library World*, 106(9/10), 474–475.
<https://doi.org/10.1108/03074800510623173>
11. White, M. (2002). Selecting a content management system. *VINE*, 32(2), 34–39.
<https://doi.org/10.1108/03055720210473830>