

## **Greenstone Open Source Digital Library Software in the Context of Arabic Content**

Dr. Ramadan Elaïess

Dept. of Library and Information Science

University of Benghazi, Libya

[ramelaïess@gmail.com](mailto:ramelaïess@gmail.com)

### **ABSTRACT**

The aim of this paper is to explore Greenstone open source digital library software which considered to be one of most essential open-source Content Management Systems (CMSs) that are available for use in creating, organizing, and managing Arabic content on the Internet. It focuses on the appropriateness of the system for Arabic content from different perspectives, such as its ability to support the Arabic language and to sustain and maintain different file formats. It also aims to examine Greenstone software and discuss its weaknesses and strengths in terms of managing Arabic digital content. Greenstone search facilities were tested and evaluated in this study using a quasi-experimental approach. The

National Oil Cooperation standalone CDS/ISIS database in Libya (NOC) was converted into web interface using Greenstone software to explore the system's functionality and to test its search and browse facilities. This database consists of 3400 Arabic bibliographic records. As previously stated, this paper aims to assess Greenstone retrieval facilities in the context of Arabic digital content to provide a full picture of the system's suitability for the Arabic language and to highlight its limitations and strong points in the overall process of information retrieval.

**KEYWORDS:** Greenstone software, Open source software, Digital libraries, Content management systems

## 1. INTRODUCTION

The importance of today's technology and its capability for expanding Arabic digital content, in addition to the recommendations of the World Summit on the Information Society (WSIS) in Geneva, 2003 and Tunis, 2005 regarding the exploitation of modern ICT to foster the productivity and expansion of research activity, has contributed to a great extent in encouraging researchers and developers in the Arab world to begin taking advantage of Content Management Systems (CMSs) to organize, manage, and disseminate Arabic digital content on the Internet.

Since the emergence of CMSs, both commercial and open-source software, Arabic content has been in a state of continuous transformation from paper-based format to digital format thanks to recent advanced technologies. However, this shift requires effective CMSs to manage and organize such content and make it available online in order to facilitate users' access to diverse and comprehensive resources on the web. There has also been a growing tendency among developers in the Arab world towards the adaptation and implementation of CMSs which do not support Arabic by default due to the absence of systems that support this language to organize and disseminate Arabic content on the Internet. In fact, CMSs are made for designated communities and local environments.

To expand research activities in the Arab world and provide information resources to scholars and researchers, CMSs would appear to be the appropriate choice. Therefore, there is a need to explore currently available CMSs to assess their quality and suitability for Arabic digital content. Exploiting such advanced technologies could assist in enhancing Arabic

content on the Internet and strengthen its presence. This in turn can assist the largest possible audience to access diverse and comprehensive resources on the web and have a positive impact on Arabic content on the Internet. Furthermore, the promotion of CMSs in a non-English language environment, e.g. in Arabic, will also help to support international and inter-cultural understanding and awareness, provide resources to researchers, expand non-English and non-Western content on the Internet, and contribute to scholarly research.

This paper is structured as follows. In the next part of this paper, background and related work is presented, followed by the concept of Content Management Systems (CMSs) and open-source software (OSS). The purpose of this section is to highlight the importance of today's technology for expanding research activity in the Arab world by creating and facilitating access to digital content on the web. The evaluation of Greenstone software is then presented, followed by description of the way used for testing Greenstone system. We conclude with opportunities for future work and a general discussion of the evaluation process.

## 2. BACKGROUND AND RELATED WORK

The world is witnessing a considerable transformation from print based-formats to digital formats thanks to advanced computing technology, which has had a profound impact on organizing, managing, and disseminating nearly all previous formats of publications in digital formats on computer networks. Due to the progress in CMSs, currently almost all types of text, still and moving images, soundtracks, music, and almost all known formats can be managed and organized on the internet

and can be stored and retrieved on computer magnetic disks (Ashoor, 2000). The improvement and expansion of Content Management Systems (CMSs) have assisted developers worldwide in designing highly practical institutional repositories and sophisticated digital libraries to expand and facilitate access to digital content on the Internet. An institutional repository and digital library is a new concept for collecting, managing, disseminating, and preserving research works created in digital form by researchers in individual organizations and institutions. These tools can assist in narrowing the digital divide or the digital gap (through knowledge sharing) between Third World countries and developed countries, which continues to widen between the information rich and the information poor. This in turn can have positive impact by acting as an instrument for sustainable economic development. Digital libraries can provide cost-effective information retrieval system, offer materials to the largest possible audience, and can simplify the process of information search, material lookup, and typical library functions (Malkawi, 2007).

As it is always important to define terms, it would be meaningful to begin with the definition of the term 'digital library'. There is no doubt that there are many different views in the literature as to the actual nature of digital libraries. This paper does not intend to provide a comprehensive collection of definitions of the digital library, but rather a number of representative definitions. A variety of terms, such as electronic library, hybrid library, library without walls, cyber library, virtual library and so forth, are still used interchangeably. Arms (2000) views a digital library as "managed collection of information with

associated services, where the information is stored in digital formats and accessible over a network". Witten and Bainbridge (2003) define the digital library as "a focused collection of digital objects, including text, video, and audio along with methods for access and retrieval, and for selection, organisation and maintenance of the collection". The digital library federation (DLF) defines digital libraries as "organizations that provide the resources, including the specialised staff, to select, structure, offer intellectual access to, interpret, distribute, preserve the integrity of, and ensure the persistence over time of collections of digital works so that they are readily and economically available for use by a defined community or set of communities." According to Jeng (2006), digital library has the following characteristics:

- It is an organized and managed collection of digital objects
- It is accessible over the Internet or server
- It is a global information infrastructure
- It should offer a service

The last point indicates that there is a difference between a digital collection and a digital library in that a digital library should offer a service to end-users. Therefore, a digital library is considered to be a collection of information objects and a collection of services that should be provided by the digital library. "The definition of a digital library that came up in the March 1994 Digital Library Workshop emphasised that a full service digital library must accomplish all the essential services of traditional libraries and also exploit the well-known advantages of digital searching, storage and communication" (Chowdhury, 2002).

In this respect, Leiner (2009) reports that “there are a large and varied set of services, including services to support management of collections, services to provide replicated and reliable storage, services to aid in query formulation and execution, services to assist in name resolution and location, etc.” at the World Summit on the Information Society in Geneva in 2003 and the Tunis Commitment in 2005 the desire to build a people-centred and development-oriented information society, where everyone can create, access, utilise, and share information and knowledge was expressed (World Summit on the Information Society, 2005). An information society is defined as “a type of society in which information and information access plays a central role, economically, socially and individually” (Merriam Webster, 2011). The information society can be identified by exploring citizens’ access to and use of the internet, e-government, e-learning, e-health, buying and selling on-line and e-business readiness. It may be concluded from the former definitions that everyone in an information society participates in the process of handling of information as a producer or as a consumer and that access to information is also available to everyone. Digital libraries can assist in narrowing the digital divide which is still widening between developed countries and Third World countries. The term ‘digital divide’ was coined in the 1990s to describe the growing gap between those who have access to, and the skills to use, ICT and those who have limited access or none. In other words, it is the gap between the information rich and the information poor. Digital libraries can also assist in the process of transition to an information society in many ways. Establishment of digital content and long-term preservation strategies of the

information created, purchased, and harvested in digital form along with an ICT policy framework to highlight the priorities of providing information for health and education as well as for remote specialists and researchers, is considered a major step in the advancement of developing countries. Meyyappan (2000) reports that “reviews of digital libraries show the diversified collection of information resources available to users ranging from full-texts of journal and conference papers to CD-ROM databases, theses and dissertations, e-journals, e-books, examination papers, images and photographs, maps, audio, video and multimedia resources, manuscripts and so on”. This means that the world is witnessing a considerable transformation from print formats to digital formats and that almost all known formats can be stored and retrieved from digital libraries which can be accessed in a variety of ways, for example through:

- Institutions’ library WebPages
- Web pages of specific digital libraries such as NDLTD, NCSTRL, NZDL, ACM digital library, and Pub Med
- Subject gateways, such as SOSIG, Biz/ed and OMNI
- Search service providers such as Dialog, Ovid online and Proquest
- Web search tools: search engines including meta search engines and specialty search engines such as Google, Ask Jeeves, Kartoo, Vivisimo, etc., and directories such as Yahoo (Chowdhury, 2000).

Digital libraries assist in the dissemination of and access to information in many ways. Pandian (2001) indicates the following advantages:

- Digital libraries bring the library to the user
- Searching and browsing

- Support of full-text searching
- Information can be shared more easily
- Current Information can be updated easily
- Information is always available (not limited to space and time)
- New forms of information become possible (digital representation of objects)
- Allows networking and exchange of ideas
- Preservation and archiving
- Faster scholarly communication and research
- Multilingual capability

Other advantages of digital libraries as mentioned by Haddouti (2007) are:

- Users can access information anywhere
- Reduction of bureaucracy by access to information
- Information is not necessarily located in same place
- Understanding the catalogue structure is not necessary
- Cross-references to other documents speed up the work of users
- Full-text search
- Wide exploration and exploitation of the information

In the last four years, the digital library paradigm received considerable interest from scientists from various areas (universities, arts, libraries, industries, and so forth.). The term 'digital library' is the result of information proliferation and technological advances. Witten (2006) states that "digital libraries can assist human development by providing a mechanism for distributing information on priorities and materials that addresses specific community problem besides benefit from methods of information distribution".

Digital libraries are spreading all over the world. All types of libraries, e.g. academic, national, public and specialist libraries, have started building digital libraries. The number of digital libraries on the Internet is increasing dramatically. Managers of libraries and information centers across the world have realized the importance of digital libraries, their potential contribution to the advancement of the services provided to end-users and their positive effect on the performance of the parent body. A considerable number of specialist libraries in institutions, organizations, enterprises, and industries have built their digital collections and have initiated the provision of digital library services using various software programmes. Building a low cost digital library is becoming less complex thanks to the availability of open source CMS software required for administrating digital content on the Internet.

### **3. CONTENT MANAGEMENT SYSTEMS (CMSs)**

Scholars in all scientific fields are taking advantage of the wealth of online information and services to ask new questions, create new kinds of scholarly products, and reach new audiences (Borgman, 2007). The Internet lies at the core of an advanced scholarly information infrastructure to facilitate distributed, data- and information-intensive collaborative research. The advent of the Internet has changed almost every aspect of our lives and a new generation has grown up surrounded by digital technologies and media. Weber (2007) states that "[I]n a growing list of countries, the use of one form of digital technology or another has permeated almost every aspect of contemporary social life,

changing the ways in which we work and study, influencing culture, media, and social rituals.”

The spread of Internet access led to the development of CMSs, which have always interested of information technology specialists seeking to facilitate digital content on the Internet. Therefore, the number of commercial and open source CMS software programmes is increasing. Feller (2005) states that “the widespread expansion of Internet access in the early 1990s led to a dramatic acceleration of open-source activity. The volume of contributions and diversity of contributors expanded sharply, and numerous new open-source projects emerged, most notably Linux”. Interactions between commercial companies and the open-source community also became commonplace in the 1990s.

A content management system offers a way to manage large amounts of web-based information that avoids the trouble of coding all of the information into each page in HTML by hand. As is well-known, there are two main categories of CMS software, i.e. commercial CMS software and open-source CMS software.

Open-source CMS is increasingly considered as an alternative to commercial CMS due to costs of production and maintenance which have risen considerably. Consequently, the free access and high level of functionality of open-source CMS software, are two of the reasons for increased usage and interest.

Open-source software (OSS) is software for which the source code is freely available for anyone to see and manipulate. The terms “free software” and “open-source software” refer to software products distributed under terms that allow users to use, modify, and redistribute the software

Free and open-source software however, is more than a set of terms of distribution. F/OSS is primarily a collection of tools and processes with which people create, exchange, and exploit software and knowledge in a manner which has been often called revolutionary (Feller, 2005).

A considerable number of digital libraries and institutional repositories are currently accessible over the Internet thanks to open-source CMSs systems. The use of open-source CMS software for managing and organizing Arabic digital content on the Internet in the Arab world has recently increased as open-source CMSs can be modified and redistributed. Source Forge is one of the best-known sites for downloading a variety of software required for managing and disseminating digital content. The following are some of the best-known software programmes in this domain, according to the Registry of Open Access Repositories (ROAR)

- Greenstone
- Dspace
- Fedora
- Bepress
- Eprints
- Keystone
- CDS ware
- DigiTool

#### **4. GREENSTONE SYSTEM: AN OVERVIEW**

As regards the following features:

- System requirements and system support which include:
  - Flexibility: support for various formats
  - Installation support
  - Frequency of version updates

Greenstone software supports various formats and can run on different operating systems such as Windows, Linux, UNIX, and Mac. This is

important, as libraries in the Arab world use different operating systems. Therefore, there is a need to adopt software that could support various operating systems. The system also provides detailed documentation for each of the systems it supports. The details can be found on the Greenstone wiki. There are also constant updates for the software, which can be found on the Source Forge website (Witten, 2008).

- Design and usability
  - Simplicity and use of the software
  - How configurable is the software?
  - How well are the support/help features and services implemented?
  - Are the extra features sufficient: handles multilingual interface, high-speed search, multiple output formats?
  - Accessibility of the software

As regards accessibility, Greenstone is open-source software. The software is highly customisable. Documentation about how to customise the Greenstone interface can be found on the Greenstone wiki. As regards simplicity, retrieved documents appear on the first screen in a simple and customisable format. Greenstone is one of the leading applications in this domain and can handle documents in multiple languages. Greenstone is also capable of displaying the user interface in multiple languages and handling collections consisting of text, pictures, audio, and video. Additionally, the Greenstone system support various file formats such as MS-Word, PDF, HTML, PostScript, JPEG, and GIF (Witten, 2008).

- Search and retrieval features
  - Search simplicity
  - Advanced search features
  - Ease of use: Powerful, easy-to-use search engine

- Information organisation
- Browsing features
- Access control and privacy

Greenstone is customisable for simple search and advanced search, and the software is powerful. A large number of digital libraries in the world have been designed using Greenstone software because the software is easy to use and has a powerful search engine. Indexing is offered for the text documents and specific metadata fields. Searching capabilities are provided for defined sections within a document (e.g., title, chapter, or paragraph) or in a whole document. Stemming and case-sensitive searching is also available. With regard to access control and privacy, Greenstone software has a built-in access control mechanism which allows collections, and even individual documents, to be restricted to authorised users using a password protection scheme. In order to create tools that effectively search, categorize, and give context to data, highly enriched annotated metadata are required to describe attribution, contents, formats, usage conditions, rights, and so forth. Greenstone is capable of implementing various metadata standards such as Dublin Core, Metadata Encoding and Transmission Standard METS, and so forth. As regards information organisation, Greenstone is also capable of organising collections and has powerful browsing facilities.

- Interoperability
  - Greenstone is highly interoperable software. It incorporates a server that can serve any collection over the Open Archives Protocol for Metadata Harvesting (OAI-PMH), and Greenstone can harvest documents over OAI-PMH and include them in a collection. Z39.50 protocol is also supported by Greenstone software.

## **5. TESTING GREENSTONE SYSTEM**

The objective of this test was to ensure that each element of the application met the functional requirements required for the test process. The first phase of testing was carried out to ensure the system's installation and functioning, prior to its implementation. The second phase of testing was conducted to verify that all components of the system interfaced with each other correctly and that there were no gaps in the data flow. The third phase of testing was carried out to ensure that the system provided acceptable response times for retrieval of information. In this test, some bibliographic records were retrieved from the system and the response time proved encouraging (1 to 2 seconds). Greenstone has two separate interfaces, the librarian interface and the user interface. Both interfaces were tested to make sure that the system runs as it should.

## **6. CONVERTING THE NOC CDS/ISIS DATABASE TO WEB INTERFACE**

The National Oil Cooperation (NOC) CDS/ISIS database, which is a standalone database, was converted to web interface using the Greenstone software so that users can search and browse the central library catalogue the banner. The user can browse the publications by title, author, editor, keyword, type of information source, and by location of the source in the central library. Figure 1 displays the

from anywhere, provided they have Internet access and basic ITC infrastructure.

Since CDS/ISIS records are limited to 32,000 characters (in version 1.5), CDS/ISIS databases generally do not contain the full text of documents; this first type of conversion is thus normally used to provide easier web access to a bibliographic database (through indexing or browsing on any of the CDS/ISIS fields). If any of the CDS/ISIS fields contain hyperlinks to external resources, they will be active in the resultant Greenstone library.

## **7. EVALUATING GREENSTONE RETRIEVAL FACILITIES**

The following screenshot shows the CDS/ISIS collection. This collection consists of bibliographic details of about 3400 records taken from the NOC CDS/ISIS database. One can browse the publications by title, author, editor, keyword, type of source and location. Advanced searches can be performed using title, creator, keyword, and year of publication, Dewey class, and terms occurring in the text and the search terms can be combined using Boolean operators. By default, search is case sensitive, although the preferences can be changed. Browse facilities can be clearly seen under English interface in advanced search mode and Figure 2 demonstrates the Arabic interface in advanced search mode.

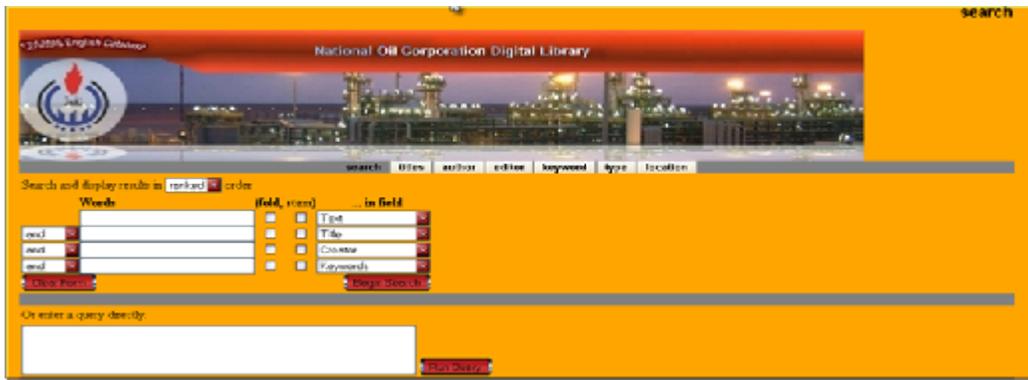


Figure 1: English interface in advanced search mode



Figure 2: Arabic interface in advanced search mode

The following screenshot displays titles from the English CDS/ISIS collection in browse mode. It can be seen that titles are arranged alphabetically from A – Z and then

from 1 – 9. Users can browse the titles under particular letters. Users could also browse the database using author, editor, and keyword in the same way.

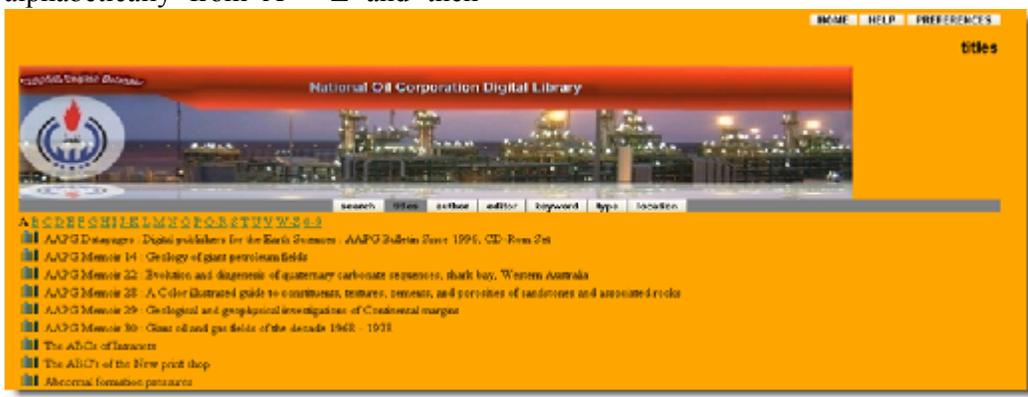


Figure 3: Greenstone browses facility

In order to test Greenstone search facilities, various techniques were applied as follows:

- Single word search: such as the word “communication”.
- Exact term: The search term can be a word or a phrase. One can use a search term, e.g. “advertising” or a phrase “information marketing”.

- Two words or phrases search: This type of search is usually used in a query to retrieve documents that have two words or phrases that appear close to each other, such as “cost accounting”. It retrieves records where the words ‘cost’ and ‘accounting’ are within two words of each other.
- Boolean search: ‘AND’, ‘OR’, ‘NOT’ are used for Boolean combinations.
- Stemming and folding

It is important to mention in this context that search facilities in Greenstone provide users with different options to modify retrieval process. In view of this, they can set and change their preferences according to their needs. The preferences button also enables users to change the search preferences in regard to query mode, query style, case differences, word ending, and search history.

The aforementioned options include the following:

- Query mode: simple query mode **or** advanced query mode (allows Boolean searching using! &, |, and parentheses)
- Query style: normal with single line (query box) **or** fielded with 2-8 fields.
- Word endings: ignore word endings **or** whole word must match.
- Return up to: 50-200 hits with 20-50 hits per page.

### 7.1 Single Word Search

To test simple query mode in Greenstone search facilities, the word “physics” was entered directly into the space provided to run query. The system retrieved 5 hits from which 3 hits were relevant and matched the query. Figure 4 shows the first hit. As can be seen the word “physics” appears under the title and statement of responsibility.



Figure 4: Simple query mode in Greenstone

Figure 5 demonstrates search hits for the same word (“physics”). The system retrieved the target word incorrectly as the word appears under the metadata

field instead of the title and statement of responsibility field. Thus, the result did not match the query and was considered incorrect.



Figure 5: Search hits

### 7.2 Exact Term Search

The term “classification” was selected in the simple query mode to examine the system’s ability to retrieve a particular scientific term. The system retrieved 14 hits, of which eight were relevant. Six hits did not match the

query and considered to be inaccurately as shown in Figure 6. In fact, the system retrieved words occurring in the metadata fields which did not correspond with the search logic.



Figure 6: Exact term search

### 7.3 Two Words or Phrases Search

This type of search is usually used in a query to retrieve documents that have two words or phrases that appear close to each other. To perform such a type of search, the phrase “remote sensing” was entered directly into the space provided to run the query. The system retrieved over 50 records that matched the query. Word count: remote: 12,

sensing: 0, following: 150, as shown in Figure 7.

From the 150 search hits, no more than 9 hits matched the search query accurately. Furthermore, the system appeared to be unable to distinguish between the word “remoteness” and the word “following” or “after”. As a consequence, many hits were in relation to the word “following” instead of the word “remoteness”.

To compare between the Arabic and English languages in terms of Greenstone's search facilities, the same phrase was entered directly into the

space provided to run a query in the English database. Figure 8 shows the results

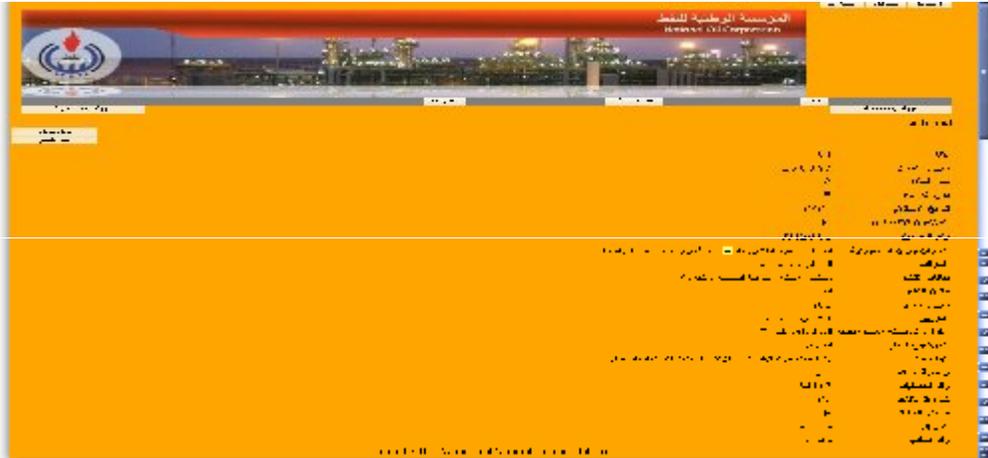


Figure 7: Two Arabic words or phrases search

As can be seen in Figure 8, the Greenstone system retrieved 4 hits for the phrase "remote sensing", all of which were relevant, accurate, and exactly matched the search query. This clearly indicates that Greenstone retrieval facilities suit the English language much better than the Arabic language and that a great deal of work must be done to improve the system's

performance in managing Arabic content. In fact, the Arabic language has not received adequate attention from CMSs developers; hence, there is a need to draw their attention to the weaknesses of their products that have a negative influence on the dissemination of Arabic content on the web.



Figure 8: Two English words or phrases search

#### 7.4 Boolean Search

To test Boolean combinations, the following search was performed. The term “oil” was combined with the term “petroleum” in the search field as structured by the Greenstone system. The search terms were combined using Boolean operators. The two terms were entered directly into the space provided to run the query in the Text field. The system should retrieve records that contain the term “oil **AND NOT**

petroleum”). It is important to mention in this context that searching these terms in the text field would mean that Greenstone should search for the target terms in any text in the database because all the texts ought to be indexed by the system. The results, as demonstrated in Figure 9 indicate that the system retrieved 94 hits for the word “oil” and 99 hits for the word “petroleum”, of which 37 documents matched the query.



Figure 9: Query for “oil” AND NOT “petroleum” combination

The system retrieved 94 hits for the word “oil” and 99 hits for the word “petroleum” from which 50 documents matched the query (“oil **OR** petroleum”), as demonstrated in Figure

10. For the combination of “oil **AND** petroleum”), 193 documents matched the query. This indicates that the Boolean search in Greenstone works as it should.



Figure 10: Query for (Oil OR Petroleum) combination

### 7.5 Case-folding and Stemming Features

According to Witten (2006), case-folding replaces all upper-case characters in the query with their lower-case equivalents, treating upper-case versions of words as equivalent to lowercase words. Case-folding treats for example, *Digital* and *DIGITAL* and *Library* and *LIBRARY* as equivalent to *digital* and *library*). “Stemming reduces words by stripping off suffixes, converting them to neutral stems for example, *libraries* is deemed equivalent to *library*” (Witten,2006). To test stemming feature in the Greenstone system, the word “libraries” in Arabic was entered in the field search, then the stem option was checked and stemming feature does not work efficiently for the Arabic language.

checked. Greenstone retrieved 69 hits for this query, all of which were related to the word libraries. Figures 11 and 12 illustrate hits which were retrieved for both words by the system. Generally speaking, by checking the stem box, the system is supposed to convert the word “libraries” to a neutral stem and to retrieve a few hits for the word “library” too. This did not occur, which indicates that the stemming and case-folding feature does not work well for the Arabic language. The word “library” was entered into the field search and the stem option was checked to test this feature again. The system retrieved 7 hits relevant to the word library. This means that the case-folding



Figure 11: Case-folding and Stemming Features

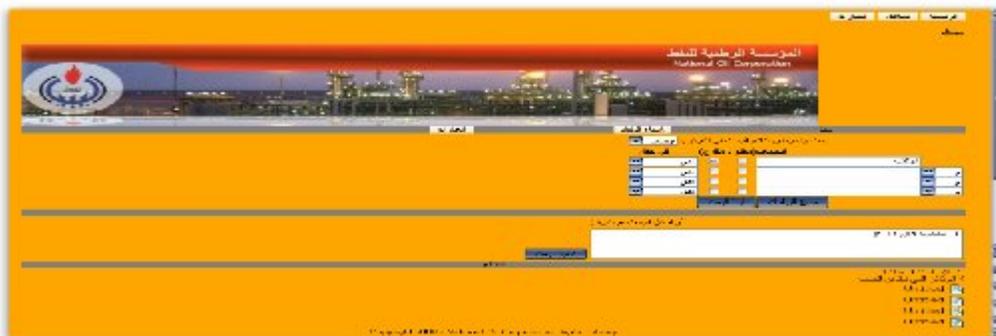


Figure 12: Results of stemming features

## 8. CONCLUSION

This paper has highlighted and reviewed a number of open-source digital library software programmes that are known as Content Management Systems by information specialists. Utilizing such software can assist in expansion of Arabic digital content on the Internet as they can help in creating, organizing, and disseminating information to the widest possible audience. From the literature review, we have come to the conclusion that current open-source digital library software still lacks certain functionalities perceived to be important. However, it must be noted that each software package has individual strengths and weaknesses. Greenstone software has proved to be suitable software for managing Arabic digital content as it supports the Arabic language as well as different file formats. The outcomes of evaluating Greenstone functionality and its search facility in this study have proven encouraging, as the software has been shown to be capable of organizing, retrieving, and managing Arabic content. As a result, Greenstone can be relied upon to a great extent in the future for disseminating Arabic digital content on the web.

The digital divide is a global phenomenon. A widening digital divide can only widen social divisions and tensions. The problem of the relation

between the access to and the availability of ICT and the participation in the development of the information society is widely recognised. Access to ICT in the Arabic-speaking world can have long-lasting benefits for quality of life, as individuals can use ICT to develop personal interests, further their education, receive job training and so on. Therefore, governments in the Arab world should take more targeted policy actions aimed at bridging the digital divide. Content Management Systems have become a necessity for establishing institutional repositories, digital libraries and for the expansion of Arabic content on the Internet, which can assist in narrowing the digital divide through knowledge sharing. These tools can have positive impacts as instruments for sustainable economic development in the Arab world. Setting up the necessary networked infrastructure and providing the requisite hardware and software in a time of global change has become more important than ever to overcome the dearth of Arabic content on the Internet. Our hope is that this paper contributes to the scholarly work and makes a difference to educators, scholars, and policy makers in the Arab world.

## REFERENCES

1. Arms, W., 2000. *Digital libraries*. Cambridge: MIT Press.
2. Ashoor, S., 2000. Planning the electronic library: suggested guidelines for the Arabian Gulf region. *The Electronic Library*, 1(2), pp. 29-39.
3. Chowdhury, G., 2002. Digital libraries and reference services: present and future. *Journal of Documentation*, 3(58), pp. 33-39.
4. Goh, Dion Hoe-Lian, 2006. A checklist for evaluating open source digital library software. *Online Information Review*, 30(4), pp. 360-379.
5. Feller, J. [editor], 2005. *Perspectives on free and open source software*. Cambridge: MIT Press.
6. Goutam, B., 2010. An evaluative study on the open source digital library software for institutional repository: Special reference to Dspace and Greenstone digital library. *International Journal of Library and Information Science*, 2(1). pp. 001-010.
7. Haddouti, H., 2005. *Digital Library Initiatives*. [online] available at: <<http://www.unesco.org>> [Accessed 18 June 2011].
8. Jeng, H., 2006. *Usability of the digital library: an evaluation module*. Ph.D. New Jersey: State University of New Jersey.
9. Leiner, M. and Richard, S., 2009. *A working definition of digital library*. [online] Available at: <<http://www.diglib.org/about/dldefiniton.htm>> [Accessed 16 July 2011].
10. Malkawi, M., 2007. *Digital library services in special libraries*. [online] Available at: <[http://www.unesco.org/webworld/public\\_domain/](http://www.unesco.org/webworld/public_domain/)> [Accessed 18 June 2011].
11. Merriam Webster Online. Available at: <[www.merriamwebster.com/dictionary/experimental](http://www.merriamwebster.com/dictionary/experimental)> [Accessed 24 May 2011].
12. Meyyappan, N. and Chowdhury, G., 2000. A review of twenty digital libraries. *Journal of Information Science*, 5(26), pp. 24-35.
13. Registry of Open Access Repositories (ROAR). *Eprints and other repositories publications*. [online] Available at: <<http://roar.eprints.org/?action=home&q>> [Accessed 15 July 2011].
14. Rick, W., Jeph, R. and Tom, K. Building a local CMS at Kent State. *Library Hi Tech*, 24 (1) 2006 pp. 69-101
15. UNESCO. *World Summit on the Information Society (WSIS)*. [online] Available at: <<http://www.itu.int/wsis/index.html>> [Accessed: 21 July 2011].
16. Witten, I., 2006. Digital libraries for the developing world. *Interactions*, 2(4), pp. 47-49.
17. Witten, I. Bainbridge, D. and Nichols, D., 2003. *How to build a digital library*. San Francisco: Morgan Kaufmann.
18. Yin, Leng Theng. 2009. *Handbook of research on digital libraries: design, development, and impact*. Hershey. New York.