



# Multilingual Digital Libraries: A review of issues in system-centered and user-centered studies, information retrieval and user behavior

Evgenia Vassilakaki<sup>a</sup>, Emmanouel Garoufallou<sup>b,\*</sup>

<sup>a</sup> Department of Library Science and Information Systems, Technological Educational Institute of Athens, Aigaleo, Athens, Greece

<sup>b</sup> Department of Library Science and Information Systems, Alexander Technological Educational Institute of Thessaloniki, Sindos, Thessaloniki, Greece

Available online 2 September 2013

## KEYWORDS

Multilingual digital libraries;  
Multilingual information retrieval;  
Multilingual information access;  
System-centered studies;  
User-centered studies;  
System design;  
User behavior;  
Literature review

**Abstract** *Purpose:* This study aims to identify, collect and critical review the research literature on Multilingual Digital Libraries in English language from 1997 to 2012.

*Design/methodology/approach:* The present literature review has followed the rules of systematic review. In particular, the identified relevant papers were categorized based on their expressed aim on two core themes, that of system-centered and user-centered studies. The assigned papers were further analyzed and six sub-themes emerged for the system-centered studies and four for the user-centered studies. Additional categorization was also provided according to type of publication.

*Findings:* The literature concerning Multilingual Digital Libraries is vast and mainly focuses on two aspects the "System" and the "Users". The majority of papers tried to meet the challenges raised for enabling multilingual information retrieval in Digital Libraries. Unfortunately, these efforts undertaken by a small number of researchers or research groups apparently working in isolation and therefore resulting in the development of numerous different tools and techniques. Relatively few studies have focused on the user and aimed to explore users' behavior and expectations when interacting with Multilingual Digital Libraries. As a result, further research is needed to reach to some tangible and usable findings.

*Originality/value:* This literature review captures the diversity of the research conducted regarding multilingual information access and retrieval in Digital Libraries. It organizes the vast literature in comprehensive themes and sub-themes enabling easy access to specific information.

*Limitations:* This study reviews only papers in English due to language restrictions from 1997 to 2012.

© 2013 Elsevier Ltd. All rights reserved.

\* Corresponding author.

E-mail addresses: [evevasilak@gmail.com](mailto:evevasilak@gmail.com) (E. Vassilakaki), [mgarou@libd.teithe.gr](mailto:mgarou@libd.teithe.gr), [garoufallou@gmail.com](mailto:garoufallou@gmail.com) (E. Garoufallou).

## Introduction

Electronic information has always been produced in a variety of different languages but for many years its collection, management and dissemination has been treated separately, for each language. A Digital Library was, and to an extent remains, the main body used to organize and access electronic information. A Digital Library is defined as “an extension and enhancement of information storage and retrieval system that manipulates digital data in any medium (texts, images, sounds; static or dynamic images) and exists in distributed networks” (Borgman, 2000, p. 42). However, until recently Digital Libraries enabled retrieval of information in one language at a time.

The Multilingual Digital Library (MDL) is a recent development and it is defined as “a Digital Library that has all functions implemented simultaneously in as many languages as desired and whose search and retrieval functions are language independent” (Pavani, 2001, p. 73). Multilingual Digital Libraries aim to support Multilingual Information Access and Retrieval. In this context, a user can search for relevant information using search criteria in a chosen language (preferably in his/her mother tongue) and retrieve all the documents which match those criteria in a unified list, regardless of the language of the documents or its indexing (Chen & Gey, 2004; Jorna & Davies, 2001).

The growing interest in providing access to multilingual information in any format led to the extension of the traditional information retrieval techniques and development of systems to support multilingual information retrieval and access (Yang, Lee, & Chen, 2009). Research in this field can be traced back in 1970's with the work of Salton (1969, 1973) but has become a major research focus during the last decade (Jones, 2002). A significant number of papers were published summarizing efforts in various sub-fields of multilingual information retrieval and Digital Libraries such as: a) *natural language processing* (Steinberger, 2011), b) *log analysis* (Agosti, Crivellari, & Di Nunzio, 2011; Di Nunzio, Leveling, & Mandl, 2011), c) *interoperability* (Fox & Marchionini, 1998), d) *multimedia retrieval and information extraction* (Peters, 2012), e) *interface design* (Peters & Picchi, 1997), f) *knowledge management concepts and technologies* (Chen, 2004b) to name just a few.

This paper reviews the research literature on Multilingual Digital Libraries in English from 1997 to 2012. Previous studies have reviewed some of the relevant literature. In particular, Agosti, Di Nunzio, Ferro, and Peters (2007) provided an overview of the evolution of the Cross Language Evaluation Forum from 2000 to 2007; Mizera-Pietraszko (2007) reviewed the cross-language information retrieval systems of some Digital Libraries whereas Erich, Coman, and Tîrziman (2010) focused on specific Digital Library initiatives in Romania. Finally, Tripathi and Tripathi (2007) provided an overview of current status of Digital Libraries in India. However, these studies were either not focused on MDLs (see Agosti, Di Nunzio et al., 2007), or reviewed studies in specific regions such as Romania, India (see Mizera-Pietraszko, 2007, Erich et al., 2010; Tripathi & Tripathi, 2007).

A few studies reviewed the way multilingualism was treated in Digital Libraries. In particular, Borgman (1997) focused on recording the multi-lingual issues involved in designing Digital Libraries whereas Haddouti (1999) reviewed the state of the art in multilingual text retrieval and access. In addition, Chen (2004a, 2004c) provided an overview of multilingual information access in Digital Libraries whereas Liew (2009) documented the research conducted for Digital Libraries from 1997 to 2007. He (2012) reviewed the outcomes of a practical session regarding achievements concerning services, users, the collection development, and the supporting technology for enabling multilingual information retrieval in Digital Libraries. One of the most recent literature reviews was that of Diekema (2012) but she focused on the core set of literature referring to Multilingual Digital Libraries.

This paper is structured as follows: Firstly, the methodological approach undertaken for conducting the systematic review is outlined. The identified themes regarding Multilingual Digital Libraries literature are described with a thorough and critical discussion of papers' objectives. Finally, conclusions are presented accompanied by recommendations about the directions research takes in terms of multilingual information retrieval and access in Digital Libraries.

## Methodology

The present literature review has followed the rules of systematic review (Hemingway & Brereton, 2009). Search terms such as “Multilingual Digital Librar\*”, “multilingual information retrieval”, “multilingual information access”, “cross language information retrieval”, “cross language information access”, “cross lingual information retrieval”, “cross lingual information access”, “translingual information retrieval” and “translingual information access” both alone and in conjunction were run on ACM Digital Library, Library, Information Science & Technology Abstracts (LISTA), Library & Information Science Abstracts (LISA), Citeseer, Google Scholar, e-prints in Library & Information Science (e-LiS), Digital Library of Information Science & Technology (DLIST). It should be noted that when a relevant paper was identified, its references were examined to identify any additional relevant papers. The searches were conducted between August to October 2012.

The identified literature was further analyzed and two core themes emerged based on each paper's expressed aim and objectives. These were system-centered studies and user-centered studies. Once the papers were assigned to these two core themes, a second analysis was conducted to further categorize each theme. In this context, six sub-themes emerged for the system-centered studies whereas four for the user-centered studies theme again based on each paper's expressed aim and objectives. Each paper was then assigned to the relevant theme creating a table (see Annex A, Table 4). It is worth mentioning that some papers had more than one aim and/or objectives and as such fell in more than one theme. In these cases, the paper was assigned to all the relevant themes.

Additional categorization was also provided according to type of publication (see Annex A, Table 3). In this context, a

Download English Version:

<https://daneshyari.com/en/article/355758>

Download Persian Version:

<https://daneshyari.com/article/355758>

[Daneshyari.com](https://daneshyari.com)