This chapter appears in the book, *Design and Usability of Digital Libraries: Case Studies in the Asia Pacific*
edited by Yin-Leng Then and Schubert Foo © 2005, Idea Group Inc.

Chapter IX

Cross-Lingual Information Retrieval:
The Challenge in Multilingual Digital Libraries

Christopher Yang
The Chinese University of Hong Kong, China

Kar Wing Li
The Chinese University of Hong Kong, China

Abstract

Structural and semantic interoperability have been the focus of digital library research in the early 1990s. Many research works have been done on searching and retrieving objects across variations in protocols, formats, and disciplines. As the World Wide Web has become more popular in the last ten years, information is available in multiple languages in global digital libraries. Users are searching across the language boundary to identify the relevant information that may not be available in their own language. Cross-lingual semantic interoperability has become one of the focuses in digital library research in the late 1990s. In particular, research in cross-lingual information retrieval (CLIR) has been very active in recent conferences on information retrieval, digital libraries, knowledge

management, and information systems. The major problem in CLIR is how to build the bridge between the representations of user queries and documents if they are of different languages.

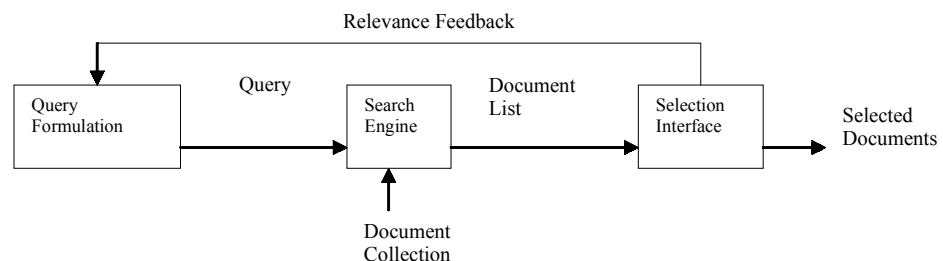# Cross-Lingual Information Retrieval

Information retrieval is a process by which users seek to locate documents that contain information about the subject of their query. The information retrieval process is distinguished from the conventional database access paradigm by the user's desire to find documents about a subject rather than data that directly answers the query (Oard & Dorr, 1996). Lancaster (1968) defined that an information retrieval system does not inform (i.e., change the knowledge of) the user on the subject of his inquiry but it merely informs on the existence (or non-existence) and whereabouts of documents relating to his request. *Figure 1* depicts a typical information retrieval system.

With the advance of technologies in World Wide Web and digital libraries, the traditional discipline of information retrieval has been extended to multimedia including image, graphics, video, and speech, and multi-languages, including European, Asian and other languages. Users expect to submit a textual query to the digital library in one language and obtain all the relevant information in all media and languages. In this chapter, our focus is the challenge of cross-lingual information retrieval in digital libraries.

The followings are some scenarios of the application of CLIR in digital libraries:

a.   The digital library has a collection of monolingual documents but it supports users who speak different languages. In this case, the queries may be of

Figure 1. A typical information retrieval system

Related Content

Metadata Usage in Serbia: Describing Paper and Digital Holdings
Adam Sofronijevic, Natasa Dakicand Jelena Andonovski (2012). *International Journal of Digital Library Systems (pp. 22-32).*
www.irma-international.org/article/metadata-usage-serbia/68815

A Survey of Digital Forensic Techniques for Digital Libraries
Yue Li (2011). *International Journal of Digital Library Systems (pp. 49-66).*
www.irma-international.org/article/survey-digital-forensic-techniques-digital/59888

The Concept of Metadata for Digital Information Resources with Special Reference to Dublin Core (DC)
Sunil Tyagi (2013). *Design, Development, and Management of Resources for Digital Library Services (pp. 160-170).*
www.irma-international.org/chapter/concept-metadata-digital-information-resources/72455

Issues for Library Management When Implementing Large-Scale Programmatic Change
Kathy Arsenault, Ardis Hansonand Joan Pelland (2003). *Building a Virtual Library (pp. 165-179).*
www.irma-international.org/chapter/issues-library-management-when-implementing/5959

QiVMDL - Towards a Socially Constructed Virtual Museum and Digital Library for the Preservation of Cultural Heritage: A Case of the Chinese "Qipao"
Yin-Leng Theng, Yanling Luoand Gladys Theng Sau-Mei (2010). *International Journal of Digital Library Systems (pp. 43-60).*
www.irma-international.org/article/qivmdl-towards-socially-constructed-virtual/48202