

Access to Scholarly Information across Disciplines, Languages, and Alphabets

YooJin Ha

Clarion University of Pennsylvania, USA

EXECUTIVE SUMMARY

The global conglomerate enterprise called libraries has spent over a century providing access to specialized trans-language and trans-alphabet information by converting non-Latin alphabets to Latin through transliteration. This endeavor has recently been challenged by efficiencies of scale provided by machine translation. A collective case study of United States national library practices shows that transliteration is still a priority for access to monographic materials at a bibliographic level, although the intended end-users are often confused by such practices and rely more on access through translation. It appears that well-established systems can co-exist in isolation from preferred systems even when both are needed. The result of such a case study shows that separate silos exist in the world of bibliographic access systems for monographic materials with the possibility that the future morphing of electronic materials including books and serials may clarify and possibly resolve this core access issue.

ORGANIZATION BACKGROUND

By its very purpose, scholarly information is meant to be shared based on the premise that knowledge advances as individuals work toward the common goal of developing theories and models. This sharing of knowledge is assumed to transcend country borders, cultures, and languages. Nonetheless, in the current Information Retrieval (IR) environment, there exist impediments and obstacles confronting users who seek to access and use online databases when seeking information in different languages.

These problems are exacerbated when users seek information written in languages not known to them. Difficulties are encountered at various stages in the search process: constructing query formulations, interpreting the core bibliographic record, and evaluating the document itself (Dolamic & Savoy, 2010; Ha, 2008; Vanopstal, Stichele, Laureys, & Buysschaert, 2012). There are certain differences in pursuing information by users with different language backgrounds. Although, currently, most systems are trying to support users seeking information across languages, it is recognized that machine translation functions are not yet stable and might be considered an experiment in the making, ripe for future user research.

Bibliographic information, defined here in its broadest sense as information about information, can be used to create access pointers to more primary information. If the access pointers, such as index terms and name of publisher, are restricted by language, then those seeking information will be denied a key opportunity to locate source material which might satisfy their information query. This is especially perplexing if the original intent of producing the information was done to be shared and to expand knowledge.

Currently, translation can be used to access information in other languages and the avenues to do this are limited but expanding. Translation, however, only partly addresses the complexities of going from one alphabet to another since it does not guarantee that similar items will remain together. Transliteration, the isomorphic linking of one alphabetic sound symbol to a symbol in another alphabet attempts to address this collocation by offering consistency when providing bridges for users to transfer from one language to another (Ha, 2010). Transliteration, however, the substitution of characters from one alphabet to another, succeeds when one human can use the new text as a transparent replacement for the original text written in a different alphabet. Extensive transliteration efforts have been undertaken for decades at the national libraries of the world. Using manually produced efforts, humans have transliterated millions of bibliographic records with the assumption that end users could traverse from a non-Roman script to a Roman script.

The purpose of this chapter is to explore how scholarly information in different languages is accessed at the secondary or bibliographic level. The focus here is on access to digital materials shared across several disciplines for works written

17 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/access-scholarly-information-across-disciplines/82652

Related Content

Best Practices in Data Warehousing

Les Pang (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition* (pp. 146-152).

www.irma-international.org/chapter/best-practices-data-warehousing/10812

Pattern Preserving Clustering

Hui Xiong, Michael Steinbach, Pang-Ning Tan, Vipin Kumar and Wenjun Zhou (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition* (pp. 1505-1510).

www.irma-international.org/chapter/pattern-preserving-clustering/11019

Program Comprehension through Data Mining

Ioannis N. Kouris (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition* (pp. 1603-1609).

www.irma-international.org/chapter/program-comprehension-through-data-mining/11033

Quality of Association Rules by Chi-Squared Test

Wen-Chi Hou (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition* (pp. 1639-1645).

www.irma-international.org/chapter/quality-association-rules-chi-squared/11038

A Genetic Algorithm for Selecting Horizontal Fragments

Ladjel Bellatreche (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition* (pp. 920-925).

www.irma-international.org/chapter/genetic-algorithm-selecting-horizontal-fragments/10930