Developing a Digital Library of Historical Records in Traditional Mongolian Script

Bilgsaikhan Batjargal, Ritsumeikan University, Japan
Garmaabazar Khaltarkhuu, Mongolia-Japan Center, Mongolia
Fuminori Kimura, Ritsumeikan University, Japan
Akira Maeda, Ritsumeikan University, Japan

ABSTRACT

This article discusses the authors’ approach in creating a digital library which can be used to access ancient historical documents written in traditional Mongolian script using a query in modern Mongolian (Cyrillic). They introduce a Traditional Mongolian Script Digital Library (TMSDL), which aims to preserve over 800 years of historical records written in traditional Mongolian for future use and to make them available for public viewing. The authors present their approaches to solve two major problems for realizing such a digital library: 1) accessing ancient documents using a query in modern Mongolian, and 2) rendering traditional Mongolian script properly. They implement an “ancient-to-modern information retrieval” method, which takes into account historical differences in the writing systems of the ancient and modern Mongolian languages. The authors’ results are explored further in the article.

Keywords: Digital Library, Historical Records, Information Retrieval, Information Systems, Library Science, Traditional Mongolian Script

INTRODUCTION

In recent years the role and importance of digital cultural heritage preservation has been increasing in the Asia-Pacific region, as well as worldwide. There has been much advancement in digital library research and development, and the significance of digital library systems that preserve cultural heritage has also increased. Internationalization and digital cultural heritage preservation worldwide require digital library systems to support various collections in many different languages from different cultures. There is no doubt that if any culture or language is omitted, we as a society would lose parts of our history, language and culture. The Mongols established the largest contiguous land empire in world history, and at that time the ancient Mongol Empire kept its records in traditional script. Considering the importance of long-term accessibility of these linguistic and cultural
records, this paper discusses the creation of a digital library of Mongolian historical records written in traditional Mongolian script.

There are over 50,000 catalogued manuscripts and historical records written in traditional Mongolian script stored in the National Library of Mongolia. About 21,100 of them are handwritten documents and over 9,400 of those are related to the history of Mongolia. In addition, many other manuscripts and books are kept in other countries. Despite the importance of storing and preserving original historical materials, the Mongolian climate and storage facilities are not satisfactory to keep historical records for long periods of time (Tungalag, 2005). An efficient and effective way to preserve and protect materials of historical importance while making them publicly available is to digitize them and to create a digital library. Thus we introduce the TMSDL, which aims to preserve ancient Mongolian historical records.

MONGOLIAN LANGUAGE

Mongolian is spoken by most of the Mongolian population as well as Inner Mongolians and other groups of people who live in several provinces of China and the Russian Federation (Poppe, 1954). One of the biggest challenges with the Mongolian language is that Mongolians have used numerous writing systems, with traditional Mongolian script being the most popular and longest-surviving script for over 800 years. The Mongolian language has also been written phonetically with Chinese characters, Square or Phags-pa script, Todo or Clear script, Soyombo script, Horizontal square script, Latin, and Cyrillic script (Shagdarsuren, 2001). The traditional Mongolian script is written vertically, from top to bottom, in columns advancing from left to right.

The sounds of words changed as the Mongolian language evolved, but the spelling remained unchanged. Therefore, there are sometimes differences between written Mongolian and spoken Mongolian. In 1946, the Cyrillic script was adapted to Mongolian with the addition of two additional characters, and Cyrillic became the official Mongolian language script. At that time, the spelling of modern Mongolian in the Cyrillic alphabet was based on the pronunciations in the Khalkha dialect, the largest Mongol ethnic group (Sečenbagatur, Tuyag-a, & Ying, 2005; Svantesson, Tsendina, Karlsson, & Franzén, 2005). This was a radical change and it separated the Mongolian people from their culture and historical archives written in traditional Mongolian script. Traditional Mongolian script preserves the ancient language and reflects the Mongolian language spoken in the ancient period, while modern Mongolian reflects the unique pronunciations in modern dialects.

By the end of the 20th century the traditional Mongolian script had made an official, government-decreed return. As a result, modern Mongolian language has two distinct writing systems: Cyrillic and traditional Mongolian. Traditional Mongolian is a distinct dialect with grammar different from that of modern Mongolian. Knowing traditional Mongolian and reading documents in traditional Mongolian script is becoming equally important consideration for Mongolians as modern Mongolian in Cyrillic. However using literacy in modern Mongolian to retrieving information from traditional Mongolian documents is not a simple task because the Mongolian language has changed substantially over time. Having an information retrieval system that deals with cross-period and cross-script historical documents has become an essential demand. We believe the proposed TMSDL will help users to access materials written in traditional Mongolian easily by allowing modern Mongolian query input.

RELATED WORK

Much research has been conducted in the last decade on cross-language information retrieval (CLIR)—a technique to retrieve documents in one language through a query in another language. There has been little research on information retrieval techniques for historical
The Concept of Metadata for Digital Information Resources with Special Reference to Dublin Core (DC)
www.igi-global.com/chapter/concept-metadata-digital-information-resources/72455?camid=4v1a