

Siam Society Digital Archive

Royal Chitradon, High Performance Computing Center, National Electronics and Computer Technology Center, Thailand.

Project Concept

This project is a pilot project of Digital Archive that is planned to build up these following know-how and Technologies for collecting our Culture, Linguistic and History in Digital Format. This will be started with Rare Book Collection at Siam Society Library.

These old rare book collection cannot be translated in-time, before these will be damaged by age, temperature and humidity. The best way to keep all of these valuable knowledge in history, culture and linguistic is putting these digital images in the archive system.

The knowledge which are kept in these rare books cannot be translated and studied because of these following reasons:

Lack of Expertise in these old script translation,

Translation has to be done at the Library, and

Lack of Dictionary and Methodology, most of the experts, they translate by their own experience and their own written books.

Solution

This project plans to provide service on computer network, these are contents and services:

Digital Archive

- Collection of rare book image collection, and
- Collection of those translated rare books.

Knowledge Base

- Dictionary of these of scripts and current Thai, and
- Grammar (These will be started in the next phase of this project.).

Engine

- String Search, and

- Semantic Search (These will be started in the next phase of this project.).

This will let most of the expert in Thailand and other place to study and tranlate these rare book collection.

System Architecture

The system will be grouped into three main parts: Digital Archive, Knowledge Base and Engines.

Digital Archive

Lossless Image Collection		
Lossy Image Collection	Metadata	Translated Document
Lossy Image for Browsing		

There will be some Image Processing, Indexing and Datacompression routines which have to be developed.

Knowledge Base

Dictionary Grammar	Dictionary of those old Scripts and Current Thai Grammar of those old in comparison with current Thai
-----------------------	---

Engine

Script search engine	Script search engine on metadata and translated documents.
Semantic search engine	Semantic search engine, this will utilize Dictionary.

Some Technologies

Most of Technogies which we need to run this project is under developmet in another project, Thailand Integrated Water Resource Management Project, there are some technologies which we need to explore:

- How to record these rare book ? (Flat-Bed Scanner, Digital Camera, Positive Color Film)
- Digital Format for Lostless and Lossy Image?

- Storage Media, Backup System and Portability?
- Indexing, Auto-Indexing, Archiving and Retrieval?
- How to buildup Knowledge Base on the top of this Digital Archive?
- Script and Font?

References

1. Atlas of World Culture, URL <http://www.wsu.edu:8000/~dee/ATLASES.HTM>.
2. Berkeley Buddhist Center, URL <http://ishi.lib.berkeley.edu/buddhist/bbrc/>
3. Ciolek, T. Matthew, The Electronic Cultural Atlas Initiative (ECAI), URL <http://www.ciolek.com/ECAI/ecai-home.html>.
4. Indra's Net Home, URL <http://indra.indranet.net/s1/ebti/HTM/Ebti.htm>
5. Preserving Digital Information, Report of the Task Force on Archiving of Digital Information, commissioned by The commission on Preservation and Access and the Research Libraries Group, May 1, 1996.
6. The John C. and Susan L. Huntington Archive of Buddhist and Related Art, URL <http://kaladarshan.arts.ohio-state.edu/Default.html>.