Ontology-based Metadata of Thai Culture: Heet Sib Song (Twelve Months Festival)

Wirapong Chansanam and Kulthida Tuamsuk

iSchool@KKU-Thailand
Information Science Department, Humanities and Social Sciences Faculty, Khon Kaen University, Thailand
Outline

• Background
• Problem statements
• Research methodology
• Preliminary result
• Conclusion and future work
Background

• Local Wisdom in Thailand
• Variety data sources
• Different academies
• Semantic search
Problem statements

• Information access system
• The metadata criteria in a research setting for Thai culture
• Mistakes cause contradictions and incompleteness
Integrating Various Data Sources

• The main challenge in integrating data from various sources is in resolving schema and data conflicts

• Approaches to this problem include using a federated database architecture, or providing a multi-database interface. These approaches are geared more towards providing query access to the data sources than towards supporting analysis

• Types of data integration:
  • Physical integration – convert records from various data sources into a common format (e.g. ‘XML’)
  • Logical integration – relate all data to a common process model
  • Semantic integration – allow cross-reference and possibly inferencing of data with regards to a common metadata standard or ontology
Research methodology

• In this study, we attempted to develop an extensible and adaptable architecture
to perform the integration of various data sources into a data repository
environment using an ontology-based data mediator approach

• The components of this architecture include:
  • Knowledge-based – stores the ontology; consists of:
    • The abstraction model – domain level concepts
    • The database description model – metadata record of data sources
    • The mappings model – how data elements relate to attributes in the abstraction model
    • The transformations model – metadata of available methods to transform data elements
      from one data source to another
  • Data mediators – provides each data source an interface to the repository
    and resolving data conflicts between any different representations;
    necessary classes generated from the Thai custom ontology
  • Data repository – provides access to integrated data for analysis and
    decision-making
A Thai Culture Prototype Architecture

- Data Source DB I (RDBMS, e.g. DHRG Database)
- Data Source DB II (ORDBMS, e.g. KKU Library Data warehouse)

*possible use of JDBC metadata to obtain DB descriptions
*alternatively, a common metadata exchange standard such as XML could be used
*XML data binding could be used to generate APIs for data validation or transformation

*Thai custom ontology can be created and modified via Protégé-5.5.0; underlying format is RDF
*abstraction model in the ontology is extensible to any domain
 major target: develop the knowledge-based as a component, use an open source

*possible use of Extensible Stylesheet Language Transformations to execute data transformations

Knowledge-based
- Abstractions
- Data Descriptions
- Data Mappings
- Transformation Descriptions

Goal DB

Repository Mediator

(Data Repository environment, e.g. H2)
Challenges of various data integration

• Increasingly large volumes of data are being made available
• Data sources are often developed by different people with differing requirements for differing purposes
• Data sources may therefore be various in terms of their:
  • Data model
  • Query interfaces
  • Query processing capabilities
  • Database schema or data exchange format
  • Data types used
  • Terminology adopted
• Integrating data sources to meet the needs of new users or new applications requires reconciliation of such heterogeneities
Thai Custom Ontology-based access to an Integrated Virtual Database Resources

User query ➔ Result

Query Processor

Thai Custom Ontology

Integrated with Thai Culture Databases schema

Mappings

Source Database schema

Mappings

Source Database schema

Source Database schema

Thai Custom Metadata
Preliminary result

• Thai Custom Metadata

• Thai Custom Ontology
  • Classes and subclasses of Thai Custom Ontology
  • Object properties of Thai Custom Ontology
  • Data properties of Thai Custom Ontology
  • Instances (Individuals) of Thai Custom Ontology
<table>
<thead>
<tr>
<th>No.</th>
<th>Elements</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Tradition name</td>
<td>The name used for the tradition</td>
</tr>
<tr>
<td>2</td>
<td>Local tradition name</td>
<td>Other names used to refer to that custom</td>
</tr>
<tr>
<td>3</td>
<td>Organized month</td>
<td>International calendar month names for tradition organizing</td>
</tr>
<tr>
<td>4</td>
<td>Time according to the lunar calendar</td>
<td>The time period specified according to the format in the Thai lunar calendar</td>
</tr>
<tr>
<td>5</td>
<td>Objective</td>
<td>What the organizers want to appear according to expectations in organizing traditions</td>
</tr>
<tr>
<td>6</td>
<td>Activity</td>
<td>The model of the method of performing expression the skits</td>
</tr>
<tr>
<td>7</td>
<td>Ritual</td>
<td>Model of practice patterns with objects things, words, and chants or text used in the performing</td>
</tr>
<tr>
<td>8</td>
<td>Literature</td>
<td>Stories, fairy tale, fables, jataka, allegory, books, prose works, poetry or writing</td>
</tr>
<tr>
<td>9</td>
<td>Belief</td>
<td>A strong faith in supernatural powers that control human destiny, comments that are consistent with what is involved in any subject</td>
</tr>
<tr>
<td>10</td>
<td>Place</td>
<td>An area or area used for holding a tradition</td>
</tr>
<tr>
<td>11</td>
<td>Ceremony</td>
<td>A person who performs the primary duty of a leader in the ceremony</td>
</tr>
<tr>
<td>12</td>
<td>Participant</td>
<td>People who take part in the ceremony</td>
</tr>
<tr>
<td>13</td>
<td>Equipment</td>
<td>Objects or things used to organize a tradition</td>
</tr>
<tr>
<td>14</td>
<td>Buildings</td>
<td>Something built for the tradition or that occurred after the tradition was held</td>
</tr>
</tbody>
</table>
Classes and subclasses of Thai Custom Ontology
Object properties of Thai Custom Ontology
Data properties of Thai Custom Ontology
Instances (Individuals) of Thai Custom Ontology
Conclusion: Metadata model for databases integration

- The metadata model here contains the information needed for the data integration process.

- The database description model contains language independent class definitions that closely mirror the physical layout of a source database. In our prototype model, the database description is simply a class containing a set of database entries. A model is provided for two distinct entry-types: *field-entries* (from flat-file data sources) and *column-entries* (from relational data sources). Entries are essentially instances of the *attribute* class.

- Modeling the database metadata as an ontology provides flexibility when trying to describe heterogeneous data sources. For instance, the model can be easily extended to describe Native XML databases.

- How the models are used in data integration:
  - The source *database* attributes are mapped to the appropriate *abstraction* characteristic through *mappings*. When an abstraction defines multiple representations for the same characteristic attribute, *transformation* functions are defined to convert between them.
Future work: Thai Culture Information Retrieval on the Semantic Web

- Query User Interface
- Indexing
- Document Retrieval
- Document Annotation
- Thai Custom Ontology
- Thai Culture Databases
- Thai Culture Physical Books Databases

Semantic Analysis of query using Thai Culture knowledge domain

Comparison of query with annotated documents

Mappings
Thank you for your attention

Question & Answer