Model Driven Data Integration with MDA Approach

MCMS project built TVU’s data warehouse to integrate distributed institutional data into a united data model and repository with model driven approach.

Model driven engineering increases productivity, improves communication, and reduces the amount of handcrafted coding and rework for data integration by defining an abstracted integration model and transforming it into program codes.

Data Integration
Data integration involves combining data residing in different sources and providing users with a unified view of these data.

The common processes of data integration includes data extraction, cleansing and merging.

A Model Transformation Framework
Model Driven Architecture (MDA) uses different levels of models which are PIM and PSM, and supports automatic model transformation between these models including the transformation between PSM and real code. Once PIM is defined in abstracted level, real source codes to execute model extraction, integration and customization are generated from this model transformation framework.

However, no existing data modelling tool supports PIM design for data integration and automatic transformation between PIM and PSM. That’s why we proposed a data merging model as extended UML and its transformation into CWM which is an industrial standard for data warehouse.
Model Driven Data Integration with MDA Approach

Data Sources of MCMS
- Historical data of TVU's 9 different data sources have been integrated.

An Example of Data Merging Model and its Transformation
- This example shows a merging of faculty data from two different data sources. Once data merging PIM is defined using the proposed data model, it is transformed into data merging PSM in oracle warehouse builder, then PL/SQL code is generated from the PSM.

Proposed Data Model
- A data merging meta model has been proposed to define relationship between model elements from different models in conceptual level. We used UML to express merging types and relationship of merging entities.

Model Transformation
- A defined merging model has been transformed into data merging model in oracle warehouse builder manually by the transformation rule below during the integration work. (The automatic transformation engine is currently under development.)

<table>
<thead>
<tr>
<th>DM Model</th>
<th>Transformation Rule</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMElement</td>
<td>IF DMElement is connected with DMSource, then generate a reference to an existing table. IF DMElement is connected with DMTarget, create new table schema including primary key and foreign key constraints. IF a DataElement has a primitive type, change table constraint on foreign key to reference a proper element.</td>
</tr>
<tr>
<td>DMSource</td>
<td>Create data mappinng as matrix, if the source and target elements do not have the same type. Insert data type change function before the mapping data.</td>
</tr>
</tbody>
</table>

References
- A Case Study on Model Driven Data Integration for Data Centric Software Development. Hyeonsook, K., Ying, Z., Samia, O. and Tony, C., DSMM’09.
- Applying MDA to the development of Data Warehouses, Mazon, J., Trujillo, J., DOLAP’05.

Home page http://samsa.tvu.ac.uk/mcms
Email for further information samia.oussena@tvu.ac.uk