

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.

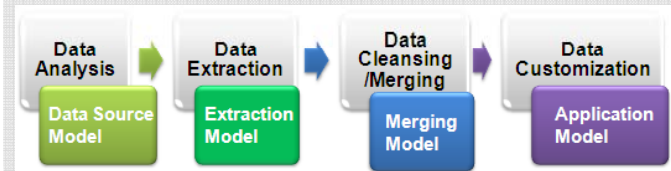
The common processes of data integration includes data extraction, cleansing and merging.



Model Driven Data Integration

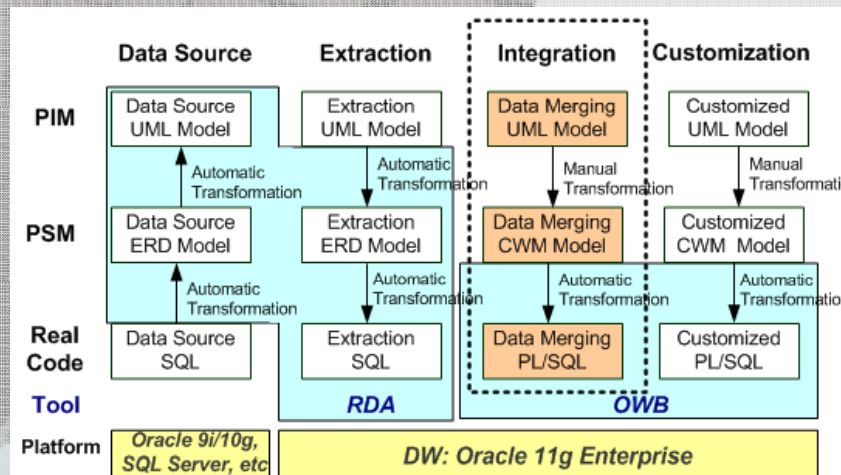
Model driven data integration is to incorporate and utilizes metadata proactively across the data integration process. It reduces the complexity of data integration by decoupling data and metadata.

In this approach different data models (meta data) such as extraction model and merging model are required for each data integration process.



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.



- PIM: Platform Independent Model
- PSM: Platform Specific Model
- UML: Unified Modelling Language
- ERD: Entity-Relationship Diagram
- CWM: Common Warehouse Model
- RDA: Rational Data Architect
- OWB: Oracle Warehouse Builder

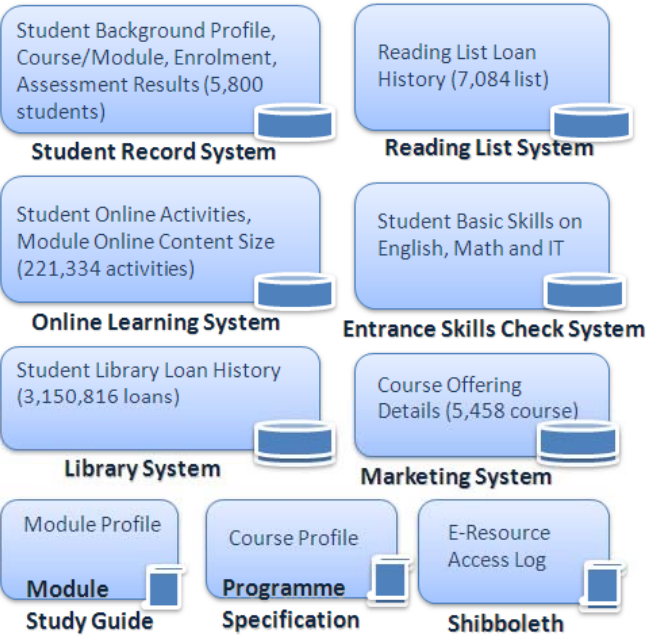
However, no existing data modelling tool supports PIM design for data integration and automatic transformation between PIM and PSM. That’s why we proposed a data merging model as extended UML and its transformation into CWM which is an industrial standard for data warehouse.



Model Driven Data Integration with MDA Approach

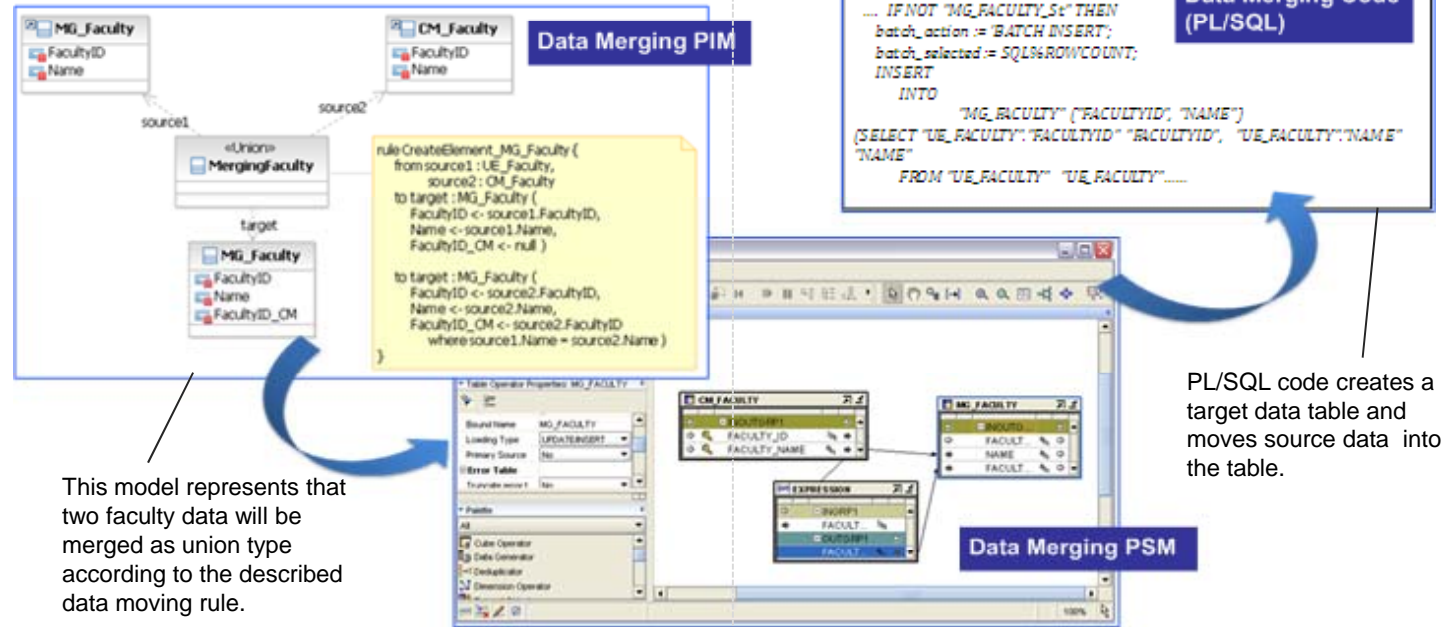
Data Sources of MCMS

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



An Example of Data Merging Model and its Transformation

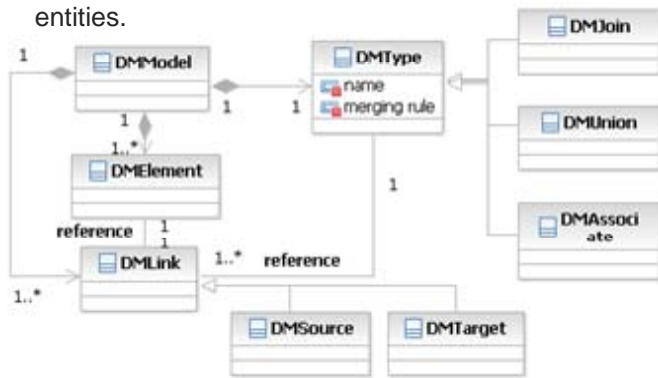
- This example shows a merging of faculty data from two different data sources. Once data merging PIM is defined using the proposed data model, it is transformed into data merging PSM in oracle warehouse builder, then PL/SQL code is generated from the PSM.



This model represents that two faculty data will be merged as union type according to the described data moving rule.

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.)

DM Model	Transformation Rule
<i>DMElement</i>	<ul style="list-style-type: none"> -If <i>DMElement</i> is connected with <i>DMSource</i> link, generate a reference to an existing table. -If <i>DMElement</i> is connected with <i>DMTarget</i> link: create new table schema including primary key and foreign key constraints. -If an attribute of <i>DMElement</i> is not a primitive type, change table constraints on foreign key to reference a proper element.
<i>DMUnion</i>	<ul style="list-style-type: none"> -Create data mappings as much as the number of <i>DMSource</i> links. -According to the mapping order in rule script, each data mapping from a source to a target is transformed into each attribute connection between source and target elements in turn. -If attributes of source and target are not same type, insert data type change function before mapping data.
.....

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.
- The Data Warehouse Toolkit, second edition, John Wiley & Sons. Kimball, R. and Ros, M. 2002.
- MDA Explained. The Model Driven Architecture: Practice and Promise. Kleppe, A., Warmer, J. and Bast, W. 2003. Addison-Wesley.
- Common Warehouse Metamodel (CWM), Specification 1.1, Object Management Group (OMG). 2003.
- 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.a.c.uk