A Mapping of Common Information Model: A Case Study of Higher Education Institution

Abdullah Fajar, Setiadi Yazid, Mame S. Sutoko

Faculty of Engineering, Widyatama University, Indonesia
E-mail : {abdullah.fajar, setiadi, mame}@widyatama.ac.id

ABSTRACT

Most of higher education institution in Indonesia has heterogeneous application software for managing their institution. Maintaining all application software run smoothly and able to collaborate one to other is very difficult. Each application has own schema and usually not prepare for collaboration. This paper will map a case study of a higher education institution. This paper describes a mapping of common information model which captured from all application software that has been implemented and review model conceptually. Common information model should reduce number of interfaces and provide a basis of integrating applications in a way that reduces the coupling among of them. Minimal functional impact should be achieved by allowing them to be upgraded or replaced.

Keywords: *Common Information Model, Heterogeneous Application, Higher Education*

1. INTRODUCTION

Most of higher education institution in Indonesia has heterogeneous application software for managing their institution. Maintaining all application software run smoothly and able to collaborate one to other is very difficult. Each application has own schema and usually not prepare for collaboration. There are two alternatives to solve that problem:

(1) Develop new application to replace old application in order all application has integrated schema;
(2) Integrated all application using middleware in order all application able to collaborate one to another.

These alternatives has advantages and disadvantages, but this paper will not explore one of alternatives but will focus how prepare mediator (will be implemented as integration broker) for heterogeneous applications. Mediator that can be captured will known as common information

Jardine [1] has introduced a concept that provide common information model under the name conceptual schema. Traditional integration practices required point-to-point (two-schema architecture) interfaces between each application and the data sources it depended on. The alternatives practices propose three-schema architecture that provides a mediator as integration broker between them which this papers refers as common information model (CIM).

There is three applications that can be reviewed as a mapping of common information model in this case study, there are: payroll, human resources academic application. This review will answer several questions to achieve the main goal of common information model. The question is whether a mapping fulfil one or more feature of CIM. The main goal of CIM is to reduce number of interfaces and provide a basis of integrating application in a way that reduces the coupling among of them. The outcome of CIM should reduce function impact minimally.

The rest of paper will describe conceptual background on the next section. After that a mapping of the model will be described and followed by discussion. Conclusion in last section describes summary, future direction and critical issues regarding common information model.

2. CONCEPTUAL BACKGROUND

A common information model (CIM) defines information that is available for sharing among multiple business processes and the applications that support them. These common definitions are neutral with respect to the processes

that produce and use that information, the applications that access the data that express that information, and the technologies in which those applications are implemented [2].

The key thesis of the ANSI/SPARC Committee was that traditional integration practices required *point-to-point* interfaces between each application and the data sources it depended on [1], as illustrated in Figure 1. This required the development of a mapping between the definition of the data available from the source (the *internal schema*) and the definition of the data required by the application (the *external schema*). The committee referred to such an approach as a *two-schema architecture*. The interfaces between the components had to implement this mapping in order to transform the physical representation, the syntax, and the semantics of the data from its source to its destination.

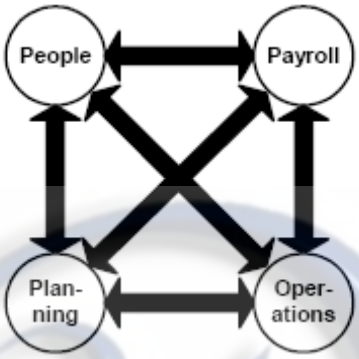

Figure 1 Two Schema Architecture

The ANSI/SPARC Committee recommended an alternative architecture, as depicted in Figure 2, in which interaction among applications is mediated through an integration broker that implements what they called a *conceptual schema*, and what this paper refers to as a *common information model* or *CIM* (also known as a *Common Business Information Model*, *Canonical Business Information Model*, *Normalized Information Model*, *Common Business Object Model*).

Figure 2 Three Schema Architecture

In this architecture, individual applications map their schemas only to the conceptual schema and interface only to the component that implements the conceptual schema (herein referred to as the *integration broker*), which is responsible for translating data from the source application first into the neutral form of the CIM and from that into the form required by the target application. Since each exchange passes from the internal schema of the source application through the conceptual schema of the integration broker to the external schema of the destination, the committee referred to this approach as a *three-schema architecture*.

The comparison of the schemas [2] is described in table below:

Table 1: Comparison between Two and Three Schemas Architecture

Two Schemas Architecture	Three Schemas Architecture
Number of Interfaces: As these systems mature, the number of interfaces to be built and maintained could grow with the square of the number of applications.	**Number of Interfaces:** The number of interfaces to build and maintain is substantially less, growing linearly with the number of applications, rather than with the square of that number.
Redundancy: The transformations required to implement communications are implemented redundantly by multiple interfaces. Each application must design its own approach to merging data from multiple sources.	**Redundancy:** The broker manages the common tasks of transforming and merging data from multiple sources, a task that would have to be done redundantly by each application in the two-schema architecture.
Impact Assessment: When an application changes in a way that affects its internal or external schema, every application that maps to that schema must be examined, reverified, and possibly revised.	**Impact Assessment:** When an application changes in a way that affects its mapping of its schema to the CIM, the only other mappings that must be examined, reverified, and possibly revised are those that contain data whose definition in the CIM has changed.
Scope of Knowledge: When an application is being upgraded to support new requirements, or when a new application is added to the architecture, architects have to examine every other application to determine what interfaces are required.	**Scope of Knowledge:** The effect of the three-schema architecture is to hide the sources and targets of data. An application's integration architects need only know about the common information model; they need not know the sources and targets of its contents. Hence, those sources and targets can, in most cases, be changed without critical impact on the application.

The common information model has re-emerged as a viable architectural approach, not as an architecture for physical databases but as an approach to achieving application integration. Variations on this approach to achieving the goals of the CIM usually offer one or more of three features:

1. Ontology
2. Standard Exchange Format
3. Integration Framework

2.1 Ontology

An *ontology* is a specification of the information appropriate to a particular business process or domain. It defines the entities that can be the subjects of information, the properties of those entities, the relationships among them, and in some cases the fundamental operations that can be performed among them. The CIM is essentially an integrated ontology that embraces the ontology of all the communities that need to share information. An ontology typically leaves undefined the mechanisms by which information based on the ontology is exchanged. That is left to those who implement the ontology. However, some ontologies are embedded in an architectural framework that provides standard mechanisms for exchanging information.

2.2 Standard Exchange Format

Some industry organizations are developing standards for how to define and exchange information based on agreements about common information. They provide standard languages and mechanisms, and leave it as a separate exercise, which they often facilitate, to specify the information content to be expressed in those languages.

1. United Nations Directories for Electronic Data Interchange for Administration, Commerce and Transport (UN/EDIFACT)
2. Open Applications Group (OAG)
3. Organization for the Advancement of Structured Information Standards (OASIS)
4. World Wide Web Consortium (W3C)

2.3 Integration Framework

The standards mentioned above try to achieve the objectives of the CIM by standardizing the content and/or format of the exchange package, and leaving the question of how a component uses or produces that package entirely to its implementation. An alternative approach attempts to optimize the integrated system by exploiting specific technologies.

3. MAPPING OF COMMON INFORMATION MODEL FROM INFORMATION SYSTEM

The key role of the CIM in all of the higher education institution activities is to reduce the coupling between components of an information system so that those components can be revised or replaced without a major redesign of other components in order to accommodate the change. Common information model have four dimensions [2] to implement, there are:

1. Neutrality : this property must respect to the application and technologies and have level or degree for its neutrality. Level of neutrality refers to John Zachman framework for an Information Architecture [3-4].
2. Scope : CIM defines the information to be comprehended, that is, its ontology, this the kind of information that can be exchanged through it.
3. Process : CIM as the basis for large scale integration is a practice for the process around it to be well founded.
4. Normalization : The value of normalization to the development of the CIM is that it assures that any given type of data is to be found in only one place in the collection. Normalization simplifies the mapping with application schemas by assuring non-redundancy in the CIM.

The steps of mapping will involve the cycle that drawn below:

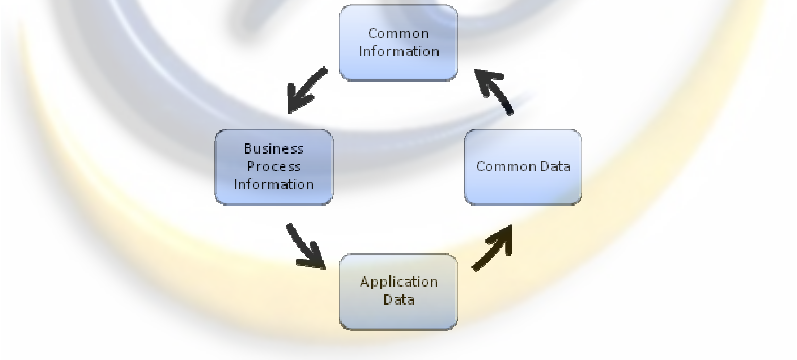

Figure 3 Cycle of CIM Mapping

The figure above can be explained as below:

1. Common data must be mapped to the common information that provides its meaning (its semantics).
2. Business process information must be mapped to the common information to provide a basis for information flow across processes and to validate definitions of common information against the rules of business processes.
3. Application data must be mapped to business process information to provide meaning to it in the context of supporting the process.
4. Application data must be mapped to common data to support the exchange of data through CIM-based middleware.

4. CASE STUDY: ACADEMIC, PAYROLL AND HUMAN RESOURCES

Academic is main business process in higher education institution, human resource and payroll is most common business process. Many application of academic and payroll is developed separately and in the worst case implemented

in different platform and environment. Common information model will be mapped from these two applications to prepare integration broker. Here below, table of mapping from academic and payroll application:

Table 2 Mapping of Academic, Payroll and Human Resources for CIM

Academic	Payroll	Human Resources
Common Data: 1. Academic Member 2. Academic Activity	**Common Data:** 1. Employee Bank Account	**Common Data:** 1. Employee
Common Information: 1. Academic member consist of: Lecturer, Professor, Assistant or tutor. 2. There is daily activity in teaching learning that involve academic member. 3. Each member academic records their activity in teaching according schedule.	**Common Information:** 1. Every month each employee receive salary via bank account 2. Each employee receives additional allowance from academic activity, over time or other else.	**Common Information:** 1. There is a record regarding absence or present of each employee 2. There's a record regarding employee activity
Business Process Information: 1. Teaching Activity Procedure 2. Midterm and Final Exam Procedure 3. Propose a research Procedure 4. Article publishing procedure 5. Conference Submission Procedure	**Business Process Information:** 1. Employee Attendant procedure 2. Salary generation procedure 3. Allowance administration procedure	**Business Information:** 1. Employee Registration/Administration Procedure 2. Employee Reporting Administration
Application Data: 1. Lecturer entity 2. Course entity 3. Course Schedule entity 4. Research and Publishing entity	**Application Data:** 1. Bank Account Entity 2. Salary Entity 3. Allowance Entity 4. Employee Attendant Entity.	**Application Data:** 1. Employee Entity 2. Portofolio Entity

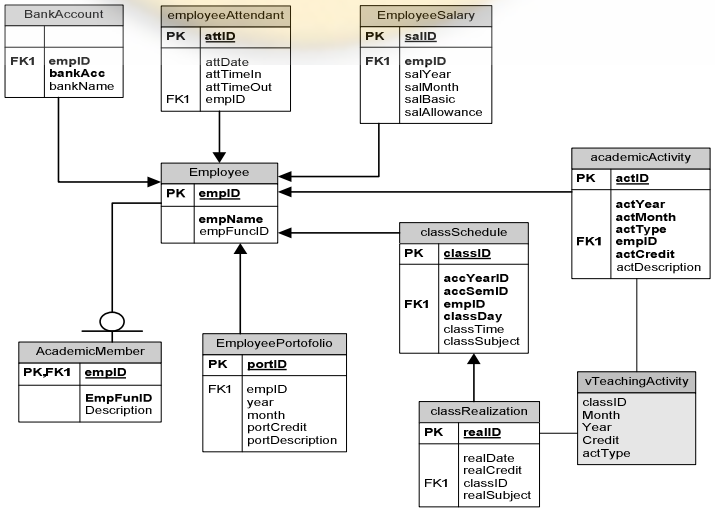

Figure 4 Entity Relationship Diagram

Based on table above, entities that will be involved in CIM such as (1) Employee; (2) Employee Attendant; (3) Employee Portofolio, (4) Employee Salary, (5) Academic Member, (6) Academic Activity, (7) Teaching Activity, (8) Class Schedule, (9) Class Realization

Based on figure 4 above, it can be capture several interface that will be included in CIM, there are:

1. Academic ← {Academic Member, Employee, Teaching Activity, Employee Attendant, Employee Salary} → Payroll
2. Academic ← { Academic Member, Employee, Academic Activity, Employee Portofolio} →Human Resources
3. Human Resources ← {Employee, Employee Attendant, Employee Salary} → Payroll

Common Information Model that can be generated from description above (drawn in figure 5):

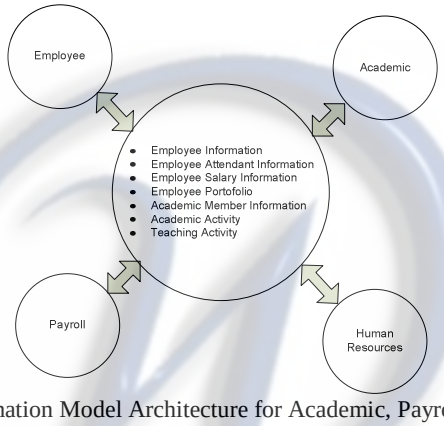

Figure 5 Common Information Model Architecture for Academic, Payroll and Human Resources

5. DISCUSSION

Case study from previous section gives a map what information should be shared between them and ready to implement using industry standard. The most popular standard for integration is XML using web service. Dimension that must be considered in this model is applied such as neutrality, scope, process and normalization. Neutrality of this model as mention by [3-4] is adequate, scope of case study has already limited to focus on shared object/entity. Process and normalization has been applied to get entity in normal form as mention by Chen [5].

6. CONCLUSION

This paper already success describing a sample case to illustrate how to prepare mediator for middleware using common information model. The key of role in modelling to minimize impact to function must be taken carefully while implementation process. This paper has introduced preparation to middleware development. The future works that can be identified such as how to detail this model in XML format.

REFERENCES

[1] D. Jardine "ANSI/SPARC Database Model (Three Schema Architecture)" in *Proceeding of the 1976 IBM Share Working Conference on Data Base Management System*, 1976.
[2] J.A Fulton, "Common Information Model" in *Encyclopedia of Database Technology and Management*. Hershey PA, Idea Group Reference, 2006, pp. 78–86.
[3] J. Zachman, "A Framework for Information Systems Architecture," *IBM System Journal.*, vol. 26, no. 3, pp. 276.
[4] J. Zachman, and J Sowa "Extending and Formalizing the Framework for Information System Architecture" *IBM System Journal.*, vol. 31, no. 3, pp. 590.
[5] P.Chen, "The Entity Relation Model: Towards a Unified View of Data," *ACM Transaction on Database System.* vol. 1, no. 1, pp. 9 - 36.