Interactive Mapping Tool for HL7 RIM-to-Relational Database Using Knowledge Game

Mariyum Gul, Muhammad Afzal, Maqbool Hussain, Mehtab Alam, Khalid Latif, Hafiz Farooq Ahmad
NUST School of Electrical Engineering and Computer Science (SEECS), Pakistan
{08msitmaryamgul, muhammad.afzal, maqbool.hussain, Mehtab.alam, khalid.latif, farooq.ahmed}@seecs.edu.pk

Abstract-Understanding the meaning of messages exchanged between healthcare organizations, has long been realized as one of the key problem, known as interoperability issue. Taking heterogeneity into account, semantic mappings between data elements of databases (hospital, laboratory or any healthcare organization database) and a standard data/information model such as HL7 RIM is the need of the day. Finding such mapping manually is very difficult and error-prone task. In this paper we propose an interactive and unique technique in perspective of user, that is find mappings through question-answer game.

I.INTRODUCTION

Healthcare organizations are growing with rapid speed, which increase complexity in controlling the flow of data between their various departments. To deal with this complexity most healthcare organizations define their own rules, vocabulary and knowledge base. Instead of reducing the complexity, all these become source of adding more complication due to heterogeneity in data sources [1].

So the intricate issue that the healthcare organizations are facing today to integrate their data is **Semantic interoperability.** Standards play a vital role in development of the cost effective and interoperable systems. In the healthcare domain, Health Level 7 (HL7) provides standards relating to exchange of information among different healthcare organizations. HL7 Reference Information Model (RIM) is a static model that is the source for the data contents of all HL7 V3 messages. In healthcare organization, relational databases are the source for all information. In order to integrate the HL7 standard in the healthcare organization, there is a need of mapping between HL7 RIM more specifically the RMIM (Refined Message Information Model) and database schema. It is very complicated task due to heterogeneity of data models, schema structures and the query language they support [2].

II. BACKGROUND: REPRESENTING HEALTHCARE DATA IN THE HL7 RIM AND MAPPING COMPLEXITIES

HL7 RIM is the root of all information models and works as the vital component for V3 development process. It is a comprehensive, non-discipline-specific, object-oriented information model of patient care and of the providers, institutions, and activities involved [3].

HL7 RIM is divided into Domain Message Information Model (D-MIM), is a specialization of RIM. Refined Message Information Model (R-MIM) further specializes the D-MIM. It contains the message specific content and all its content is drawn from particular D-MIM in which it is used. For example, Laboratory DMIM has two RMIMs: Result Event and Result Query.

Common Message Element Types (CMETs) are normally used in RMIMs classes. These are work product for communicating a common, valuable and reusable concept [4].

In the healthcare systems, HL7 RIM provides the bases to achieve interoperability. In order to make other healthcare organizations HL7 RIM compliant, there is need of mapping between database schema's of healthcare organization and RIM. Many mapping techniques exist for this purpose but don't provide mapping at good rates.

In paper [2], authors worked on developing an 'Automation Tool for HL7 RIM-to-Relational Database Mapping'. The basic methodology followed is to analyze various clinical databases and then identify the matching fields in corresponding to HL7 V3 Lab Domain. In this tool the repository is introduced that contains the matching fields and tables/classes. It works as the knowledge base. Using this knowledge base, mappings are performed. By introducing knowledge repository it overcomes many weakness of simple data mapping, for example, it introduces some sort of semantic in the form of heuristic algorithms, once analysis is done then mapping is performed automatically.

Although in [2] the proposed architecture has solved the mapping problem to some extent in order to lessen the gap between HL7 V3 standard and local health databases yet mapping is also not fully automated**.** Most of the mapping is required to perform manually because HL7 RIM is complex to understand due to following reasons;

- Complex Data types

- Correct mappings to missing HL7 RIM attributes, for example: There is no corresponding attribute in HL7 RIM for "age".
- Complexity of applying the RIM. For example, data stored in one local "*Patient"* table requires mapping data to at least five classes in the HL7 RIM: *Role, Entity, Living Subject, Person,* and Organization.
- End user might come across a confusing situation because of ambiguity in the data. For example, the "id" attribute of patient table on database side have more than one mapping options whether to map with the "id" attribute of person or patient class on RIM side.

To resolve above mentioned complex situations, healthcare operators (e.g. laboratory administrator) are involved. There are two basic issues if the user is allowed to map such mapping directly without providing some sort of help or guidance.

- It is highly possible that he/she might wrongly map the field just because of not knowing the context.
- Knowing about the context for healthcare operator before mapping requires deep understanding of each and every concept of HL7 RIM, which is a highly cumbersome phenomenon.

To free the user from such intricacies, we are proposing in this paper an interactive and user friendly technique known as mapping via question/answer game. Details of this technique are provided in section next.

III. PROPOSED TECHNIQUE

Due to the complexity of RIM, it is a tedious task for user to do mappings manually. We introduce the question answer game in which the attributes from a database are mapped with HL7 RIM. These attributes were previously mapped manually by domain experts [2].The underlying approach depends on a question bank covering different aspects of R-MIM. As a case study we have only taken the R-MIMs of laboratory domain. End-user will be asked a number of questions; according to the answer of the user, the required mapping is performed.

In our approach we need two ontologies; ontology of R_MIM and Questions ontology. The R-MIM ontology is specialization of HL7 RIM ontology developed elsewhere [1]. Question ontology, on the other hand, doesn't share its base with other ontology. It defines the ambiguity patterns and question sentence. It sets the bases for questions. Questions are designed in order to dig out the ambiguity in mapping.

A. Methodology

- Analyze the R-MIM thoroughly and design a number of questions.
- Make the ontology of these Questions. QuestionBody class in question ontology contains the question text. The Question Type defines the type of the question True/False, Multiple Choice etc [5].
- Also make the ontology of R-MIM and link with question ontology by using the property *has question* whose domain is entry class in R-MIM ontology and range is QuestionBody class in question ontology as shown in Fig.1. Entry class is the entry point for each R_MIM i.e. entry point for Laboratory R-MIM of 'Result Event' is *Observation Event Choice*.

B. Process Flow

Fig.2 shows the process flows of whole architecture.Step.1 to step.3 are related to [2], which provides the base for our work. Step.4 to step.8 shows process flow of proposed architecture.

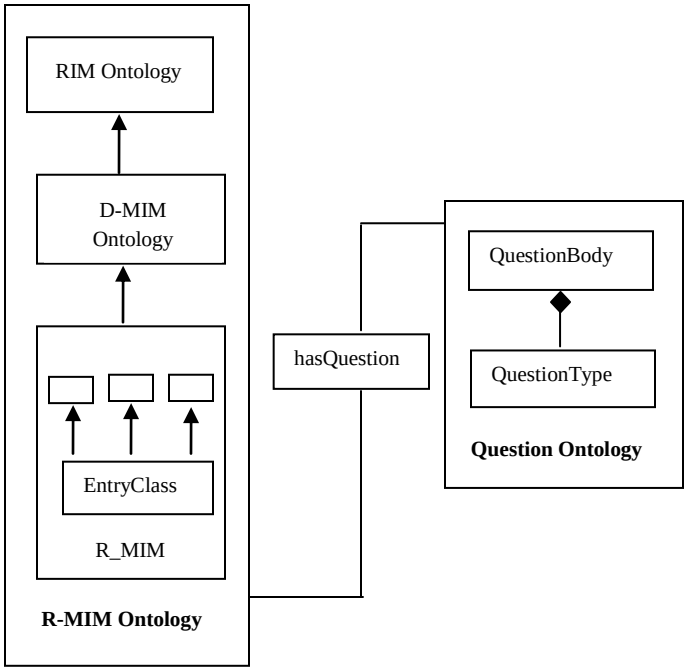


Fig.1.Preview of R-MIM and Question Ontology

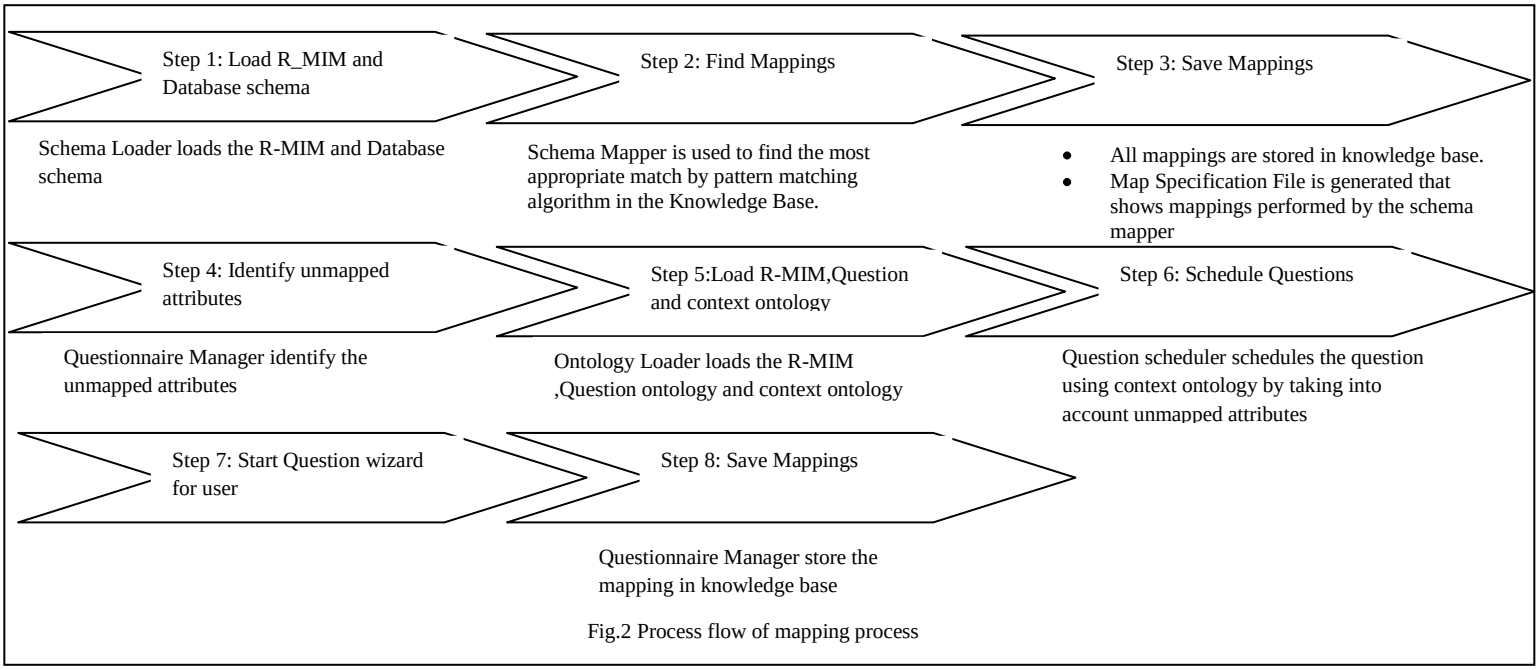


Fig.2 Process flow of mapping process

C. *Architecture Components*

The major components of the architecture are shown in Fig.3. This is an extension of our previous work reported in [2].

1) *Question Manager* controls all the components of the question mapping process. It extracts those attributes that are not mapped automatically. When mapping is performed by question mapping process, it stores these mapping in *knowledge base* again.

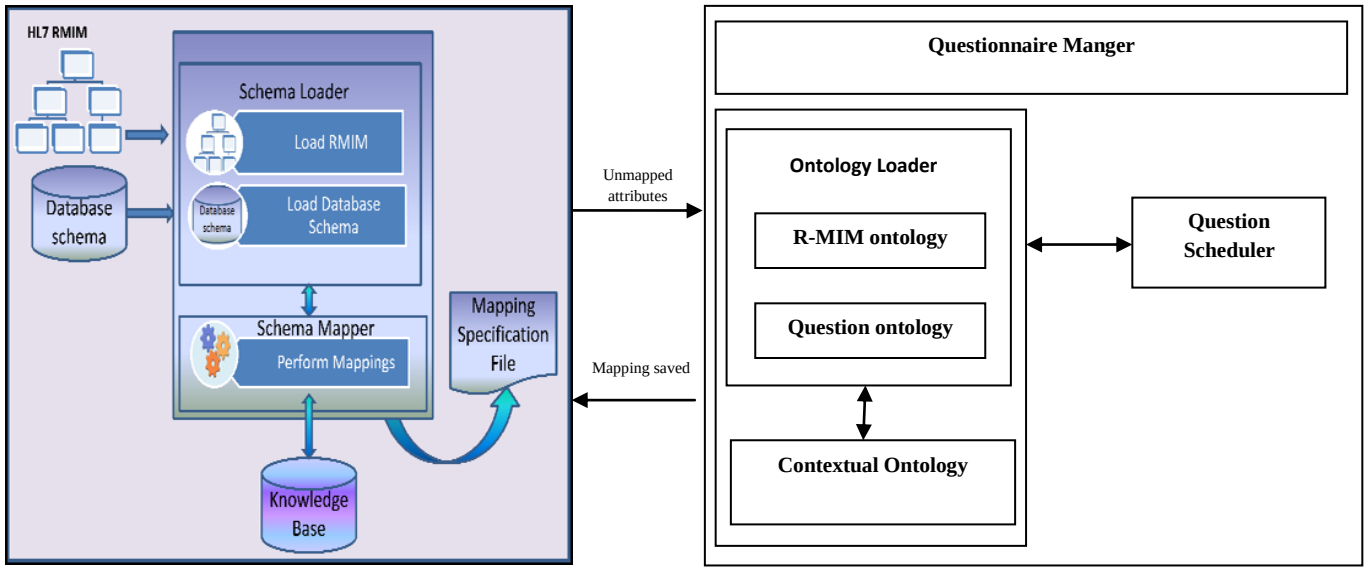
2) *Ontology Loader:* loads the R-MIM ontology and the corresponding question ontology and connect them with each other as shown in Fig.1

3) *Contextual Ontology:* provides the context for example; when the question will start, where to stop, what are the sequences of the questions etc.

4) *Question Scheduler:* the basic purpose of this component is to schedule the questions in question ontology according to the attributes that are taken by Question manager from Fig.3 (a).

D. *TEST SCENARIO*

In many cases the mapping from database to R-MIM is one-to-many, means one table of database is mapped to



(a) Schema Mapping Process

(b) Question Mapping Process

Fig.3 Databases to R-MIM Mapping

multiple tables of R-MIM.

E. TEST SCENARIO

In many cases the mapping from database to R-MIM is one-to-many, means one table of database is mapped to multiple tables of R-MIM.

Here we take *Result Event* RMIM of HL7 V3 Laboratory domain. The *id* attribute in the patient table of database can be mapped to the person and patient table of the R-MIM as shown in the Fig.4. Confusion arises as the system tries to map it to R-MIM. Many other situations occur where mapping algorithms find it to finalize the appropriate mapping. These mappings can be found by asking appropriate question from the domain expert in order to reduce ambiguity. In the case of "id" for example, we assume that no mapping is found by the tool existing. The content of question will be like *whether this "id" is used for person identification in hospital*. If user replies with 'Yes' then this "id" is mapped with the Patient class of R-MIM, in the case of 'No" it will be mapped with the "id" attribute of the Person Class.

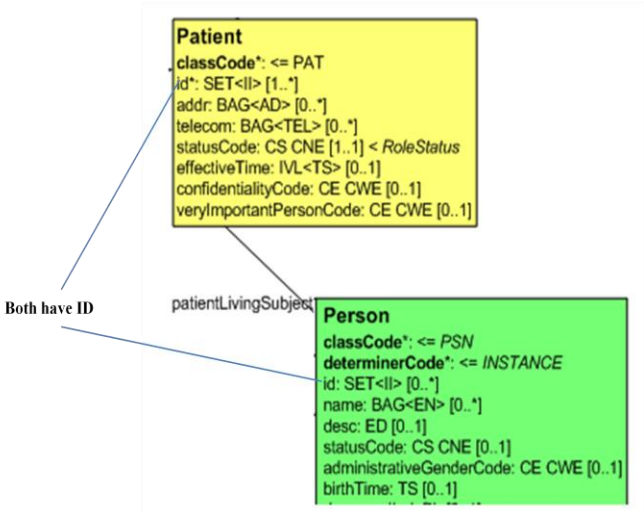


Fig 4: Person and patient class of R_MIM having attribute Id [4]

IV. DISCUSSION

The use of the RIM as a persistence model is known as **RIM Based Application Architecture (RIMBAA). RIMBAA** provides framework for those who are interested in using the RIM for application and database design. This frame work helps to create databases according to the HL7_RIM, so the databases are HL7_RIM complaint from their initial stage. The architecture presented in this paper, provides the mechanism for HL7_RIM complaint to those databases which are already created [6].

V. CONCLUSION

We have proposed a truly dynamic and automated technique for mapping the healthcare schema to HL7-RIM in order to create the proper motivation for end users to get involved. This technique required two ontologies R-MIM ontology and question ontology. The proposed architecture provides the mechanism to map the unmapped attributes by using user-friendly questions answer game. The questions are designed in such a way that builds the interest of user in order to find out accurate mapping. In this way users will be succeeded to achieve mappings for the non-trivial situations where he/she gets confused to decide the correct mapping.

VI. REFERENCES

[1] B. Orguna, J.Vub, "HL7 ontology and mobile agents for interoperability in heterogeneous medical information systems," Science Direct, vol.36, Issue 7, Pages 817-836, July-August 2006

[2] Shagufta Umer, Muhammad Afzal, Maqbool Hussain, Hafiz Farooq Ahmad, Khalid Latif, "Design and Implementation of an Automation Tool for HL7 RIM-to-Relational Database Mapping," 10th International HL7 Interoperability Conference, 2009

[3] William T F Goossen, Judy G Ozbolt, Amy Coenen, "Development of a Provisional Domain Model Health Level 7 Reference Information Model for the Nursing Process for Use within the Health Level 7 Reference Information Model," American Medical Informatics Association, vol. 11, no. 3, May / Jun 2004.

[4] "Graham Grieve, et al, Transmission Infrastructure, ANSI/HL7 V3 IM, R1-2004, HL7 Normative 2006"

[5] Radu Balog-Crisan, Ioan Roxin, Ioan Szilagyi, "Ontologies for a Semantic Quiz Architecture", Proceedings of Ninth IEEE International Conference on Advanced Learning Technologies, 2009

[6] RIMBAA, http://www.hl7.org/Special/committees/java/overview.cfm