DESIGN AND IMPLEMENTATION OF AN AUTOMATION TOOL FOR HL7 RIM-TO-RELATIONAL DATABASE MAPPING

Shagufta Umer, **Muhammad Afzal**, Maqbool Hussain, Hafiz Farooq Ahmad, Khalid Latif

Presenter

NUST SCHOOL OF ELECTRICAL ENGINEERING & COMPUTER SCIENCES, PAKISTAN.

ABSTRACT

Abstract

Exploring HL7 interoperability with the local clinical databases.

Heterogeneity in data models - Difficult to bring local clinical schemas in compliance with HL7 messaging.

HL7 version 3 messages can be parsed to any relational database composite of tables attributes and associated identifiers.

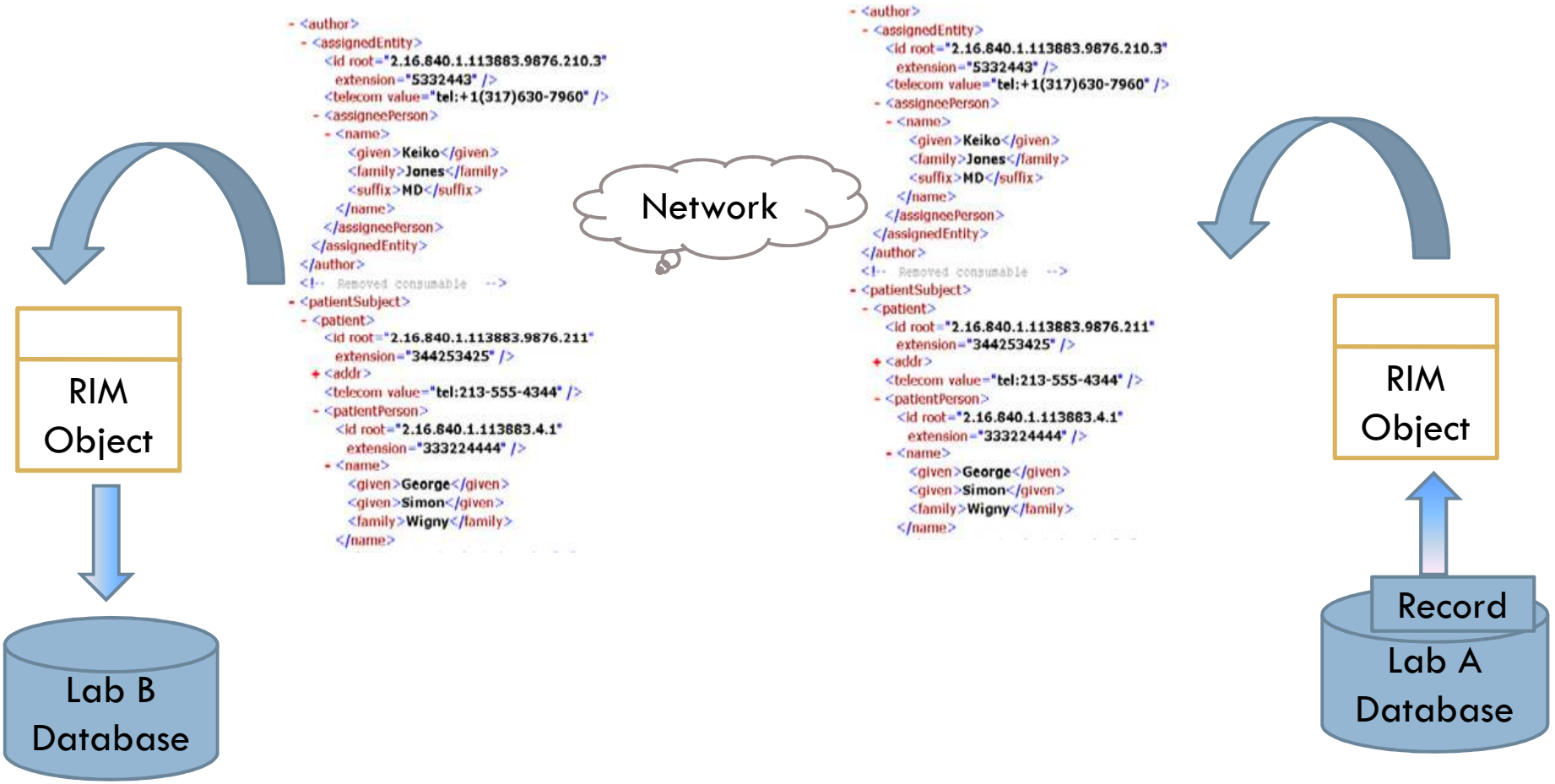
In this paper, we explored the RIM to clinical schema mappings and proposed a scheme for dynamically mapping clinical schemas to RIM.

INTRODUCTION

Introduction

- Healthcare organizations store their data in traditional relational databases.
- Information exchange between two healthcare systems requires *common syntax, shared vocabularies*.
- HL7 RIM is a comprehensive UML model representing healthcare concepts.
- There exist either one-to-one or one-to-many correspondences of fields *mapping between clinical data model and RIM model*.

HL7 Messaging Process



RIM to Schema Mappings Challenges (1/2)

- Heterogeneity issues because of
  - Varying data models (ER, EAV, UML)
  - Differences in schemas
  - Query languages they support
  - Domain terminologies they recognize

- Same concept, different name.
  - Medicine: Drug, pills, Medicine etc
- Same name, different concepts.
  - Doctor: Author, Performer, Verifier etc

RIM to Schema Mappings Challenges (2/2)

- Finding RIM to schema mappings is the bottleneck!
  - mostly done by hand
  - labor intensive & error prone

- RIM covers the whole healthcare arena, but complex to understand.
  - Complex Data types
  - Correct mappings to RIM attributes is complex problem.
    - Age : No exact corresponding attribute in HL7 RIM
    - NIC: No representation in HL7 RIM etc.
  - Codes and vocabularies. Local schemas don't observe codes.
  - IDs and sequence numbers are used for interoperability purpose, no such practical usage in clinical systems.

Problems in Existing Approaches

No automation

Time consuming task

Error prone user intervention in identifying the appropriate mappings.

Proposed Methodology

Proposed Methodology

- Analyze some of the clinical databases, more specifically laboratory databases.
- Patterns in the clinical databases are collected.
- Example;
  - patient name in clinical databases mostly mentioned as:
    - *pname*
    - *Name*
    - *PatientName*
    - *FirstName/LastName*
    - *pFirstName/pLastName*

- Maintaining a ***Mapping Knowledge Repository*** using the patterns seen in the clinical database terminologies.

Proposed Architecture

Schema Loader:

Loads the data model of the target clinical database and the corresponding RMIM Object Model.

Schema Mapper:

This component is the core engine of the process. For every column in the database, it will search the most appropriate match in the Knowledge Base and bind it to the RMIM accordingly.

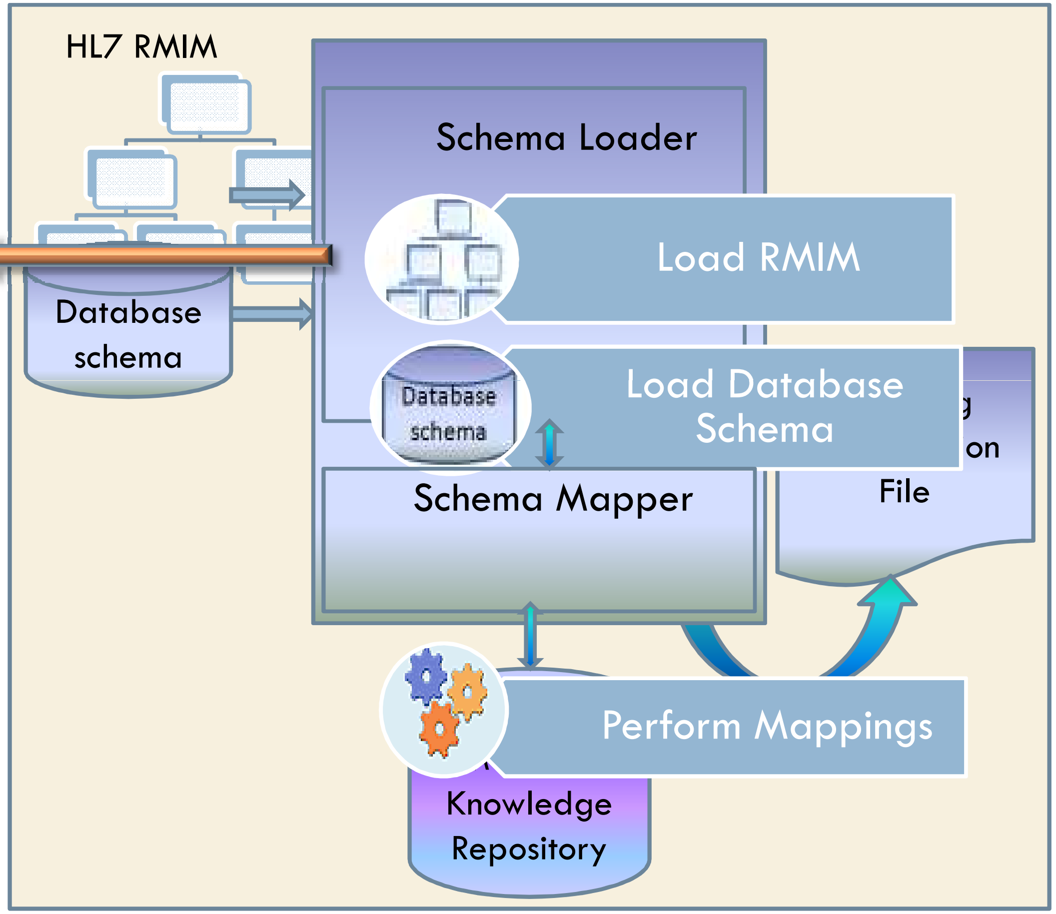


Fig. Proposed Architecture

Implementation & Evaluation Strategy

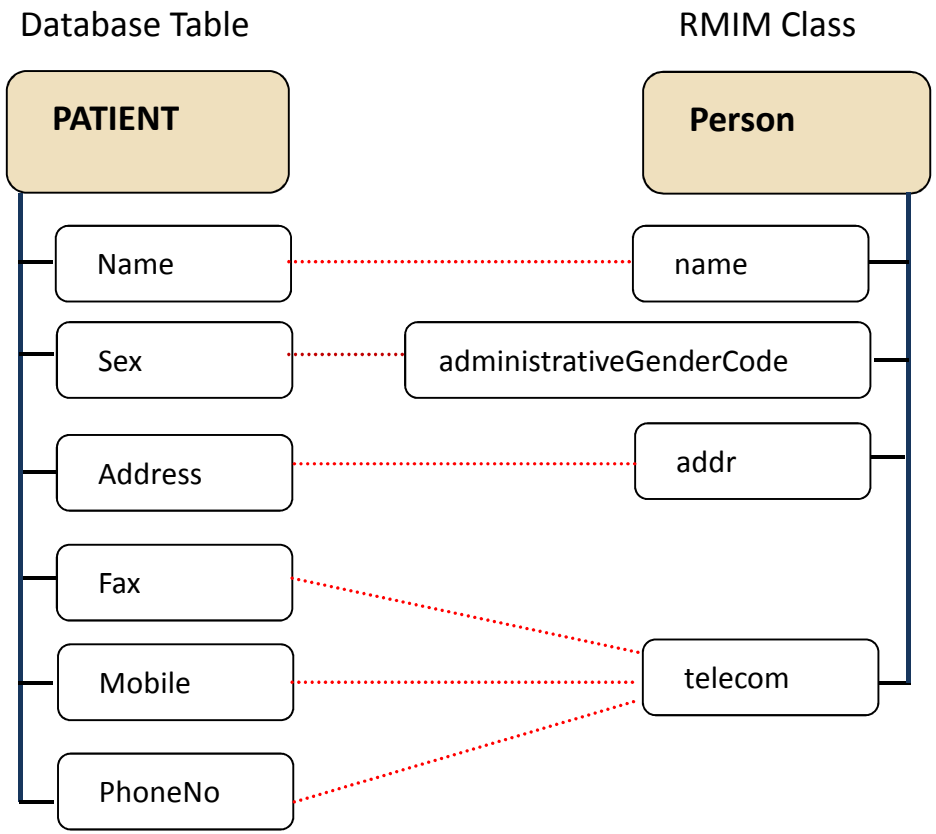
Implementation & Evaluation Strategy….

- Implemented as a case study to support CITILab, Pakistan.

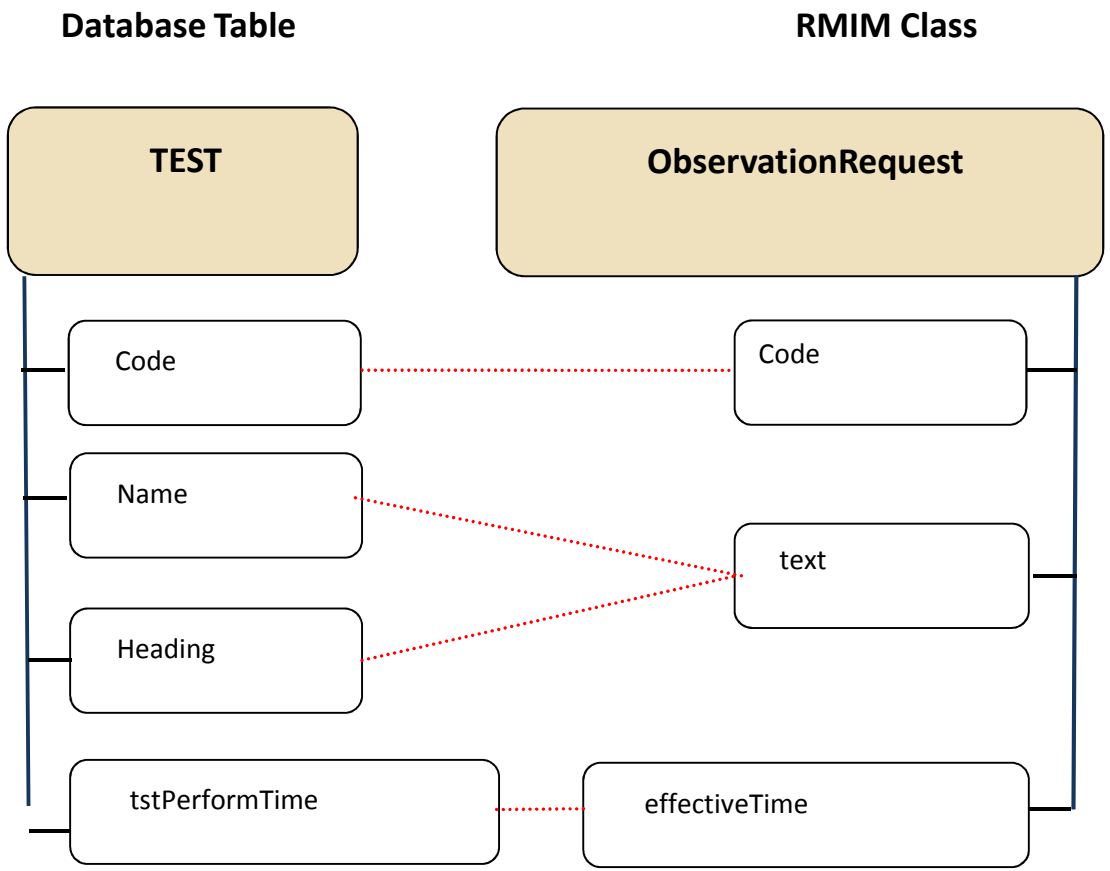
- The CITILab database currently runs in
  - Windows environment and
  - uses the Microsoft SQL relational database management system (ver. 2000).

- Information centered on the concept of a “test order” and “test results”.

Mapping CITILab to RIM (1/2)



Mapping CITILab to RIM (2/2)



Evolution of the Tool

- The Knowledge Repository will be helpful in mapping multiple clinical databases
- The *repeating patterns* are helpful in refining the knowledge repository.
- So the tool will evolve by adding new mappings in knowledge repository with minimal efforts.
- The tool being an open source, the community may contribute in its evolution process.

Conclusion & Future Work (1/2)

- Mapping provides ease in sharing and exchange of the information.
- Understanding and mapping RIM with clinical schemas require a thorough and deep understanding of each and every concept.
- Clinical schemas are developed having no single structure and representation.

Conclusion and Future Work (2/2)

- Adding Intelligence through ontology based mapping.
- Extract any database schema and build automated ontology of database.
  - Overcomes the mapping knowledge repository.
- Later on, this ontology will be used for the dynamic RIM to clinical schema mapping.

References

References

[1] Health Level 7, http://www.hl7.org , November 2008.

[2] HL7 Reference Information Model, ANSI/HL7 V3 RIM, R1-2003, 12- 3- 2003.

[3] Walter Sujansky, "Heterogeneous Database Integration in biomedicine", Journal of Biomedical Informatics, 285-298, 2001.

[4] Rene Spronk, The RIMBA Technolgoy Matrix, http://www.ringholm.de/docs/03100_en.htm

[5] HL7 Version 3 Guide, http://www.hl7.org/v3ballot2008JAN/html/help/V3guide/v3guide.htm

[6] "Laboratory Domain", HL7 V3 Lab, R1, Last Ballot: January 2008.

[7] "Laboratory Domain", HL7 UK V3 Lab, R1, Affiliate Ballot: May 2008.

[8] Lyman, J. Pelletier, S. Scully, K. Boyd, J. Dalton, J. Tropello, S. Egyhazy, C., Dept. of Health Evaluation Sci., Virginia Univ., Charlottesville, VA, USA, "Applying the HL7 reference information model to a clinical data warehouse", Oct. 2003, Volume: 5, On page(s): 4249- 4255 vol.5

[9] caAdapter, https://cabig.nci.nih.gov/tools/caAdapter , December 2008.

[10] caAdapter Model Mapping Service Hands-On Training, November, 2007.

[11] http://www.connectingforhealth.org/, February 2009.

[12] Ringholm, http://www.ringholm.de , January 2009.

[13] NeoTool, http://www.neotool.com December, 2008.

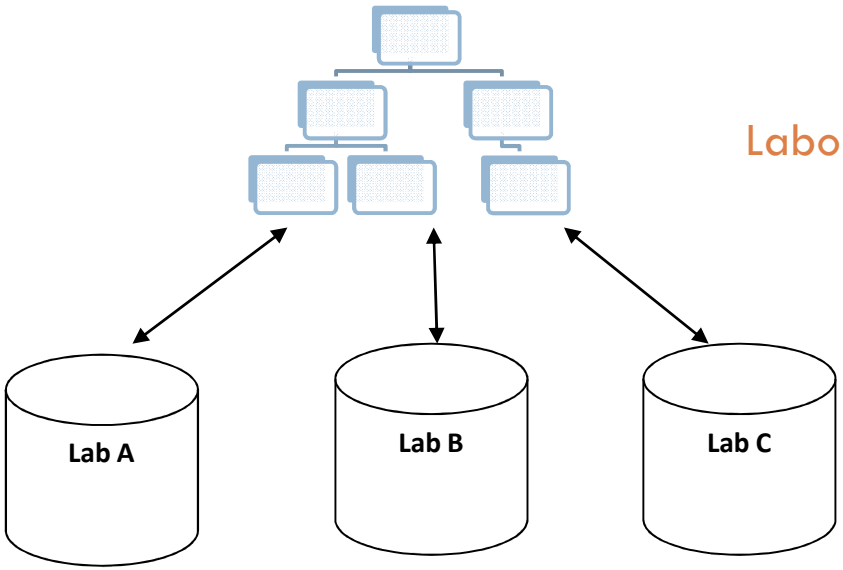
Questions & Answers

?

Training the TOOL

Q: How many clinical database will be analyzed to serve the mapping process?

- Manual Mappings for 3 Laboratory databases first and then build a mapping knowledge repository.



Laboratory RMIM

Laboratory Databases

Mapping Strategy

Q: What is the mapping strategy?

- **Representation matching:** Matching the Table Name or the Column Name.
- **Data types Matching:** Identifying some mechanism to handle the complex RIM data types with respect to the Clinical schema data types.
- Column with respect to its placement in a table can also be used to handle the mappings.

Constraint Handler

Q: How mapping conflicts are handled?

- If two or more mappings are identified for a single concept then ask the user to decide which mapping is most appropriate.

- If any mapping is wrongly identified than the user can request to remap the particular table/field.

Interface for the Mapping Knowledge Repository

Q: How the mapping knowledge repository would evolve ?

- Knowledge repository would evolve with the passage of time through community interaction.
- GUI is provided to update the **mapping knowledge repository**.
- **Categorize the RIM classes**, e.g. Classes involved in Patient Information, Test Information, Specimen Information, Equipment Information etc. This will help the user in finding the appropriate mappings.
- **Provide user-level description**: Besides RIM complexity, the description provided to the user should be simple and clear.
- Display to the user the appropriate choices for the mappings.
- Allow the user to select appropriate mappings.

Backup Slides

Related Work (1/2)

Applying the HL7 Reference Information Model to a Clinical Data Warehouse

Lyman, J. Pelletier, S. Scully, K. Boyd, J. Dalton, J. Tropello, S. Egyhazy, C.

This paper focused on:

Mappings from a local clinical data (CDR) warehouse to RIM based classes.

The CDR is a unique information resource at University of Virginia Health System.

Problems:

Proposes static mapping , only RIM to CDR.

No rules for extending the concept to other databases.

Does not focus on automating the RIM to schema mappings process.

caAdapter

A framework National Cancer Institute US

- caAdapter offers the capability to map object models to data models.
- The Mapping Tool component has a GUI-based front-end application with drag-and-drop capability for mapping.
- caAdapter mapping service
  - requires human intervention for manually tracing out all attributes in RIM.
  - Requires laborious and time consuming efforts.
- No mechanism for collecting RIM context for target database schema.