 ISBN 978-83-60810-22-4

Semantic Interoperability for Infectious Diseases Reporting System

Murugavell Pandiyan
Research Student,
Kalasalingam University,
Krishnan koil, India.
Email:
murugavell.pandiyan@tecom.ae

Osama El Hassan
Sr. Technical Consultant
TECOM Investments,
Dubai, UAE.
Email:
osama.elhassan@tecom.ae

Zakaria Maamar
Professor,
Zayed University,
Dubai, UAE.
Email:
zakaria.maamar@zu.ac.ae

Pallikonda Rajasekaran,
Associate Professor,
Kalasalingam University,
Krishnan koil,India.
Email:
mpraja@klu.ac.in

***Abstract*—The healthcare setting is multifaceted, comprised of many different components including private, governmental, and regulatory agencies. There is always a necessity of timely and reliable information exchange among these agencies especially on "Infectious Disease" information due to their criticality. The heterogeneity of the systems used by these agencies has led us into designing and developing an interoperable solution to exchange data effectively among several independent yet collaborating health authorities at both state and national levels. This research work articulates the efforts put into achieving an interoperable "Infectious Diseases Reporting System" that incorporates ontology-based semantic rules to align different infectious disease coding standards and to deploy Web services for collecting data from remote sources. This effort is a first step towards achieving a policy-based interoperable Infectious disease monitoring system which can be used across different yet collaborating regulatory agencies.**

***Index Terms*—Interoperability, Semantics, Healthcare system.**

I. Introduction

INFECTIOUS diseases outbreaks demand a timely and proportional response to mitigate effects on public health. Management of these outbreaks is becoming a growing concern in public health, as it requires extreme actions and coordination between governing authorities at both state and national levels. Tracking and identifying emerging infectious diseases and epidemic outbreaks in particular poses critical challenges on public health researchers and practitioners. Dealing with large numbers of incoming reports and alerts requires an automated system which performs real time analysis on a centralized repository of collected clinical information. As such, intrinsically intensive information exchange should be exercised between healthcare facilities and this repository, which is normally operated by a regulator. Regulators use these collected information to identify, manage, and investigate infectious diseases outbreaks. Such data, when visually and adequately represented, support the education of healthcare providers within participating facilities and improves the outcome of disease outbreaks management.

A. Challenges

Gathering information from various sources (i.e., healthcare facilities) in real-time is a challenge because of the diverse and heterogeneous nature that labels interfacing healthcare applications. Moreover, the capability of aligning and integrating healthcare data standards varies drastically from one coding system to another. There is a necessity here to design an advanced interoperable system that is empowered with a semantic layer to retrieve and map the information onto a unified standard to support alerting and surveillance. The design of such an interoperable system is complex and challenging due to the following factors:

1. Each healthcare facility has its own software to manage patient's data e.g., Physician Practice Management System (PPMS) and Electronic Medical Records (EMR). Moreover, even standardized EMR systems might be interfaced differently and sub-systems such as Lab systems are isolated in terms of their used standards (i.e., proprietary standards) and thus raise extra integration challenges.
2. Each facility has its own policies and procedures for reporting patient data.
3. The network infrastructure and real time reporting may affect the data availability.
4. Last but not least, from a regulatory body point of view, since the infectious disease is a growing concern in public health, it is necessary to collaborate with other health authorities to exchange and manage the related information and alerts.

To address the aforementioned factors, a platform is needed to enable interoperable information exchange between independent healthcare systems.

B. Proposed Solution

This research work presents the design and development of semantic information integration for Infectious Diseases Reporting system using Web services as a core component for information exchange with ontology-based rules. Ontology has been predominantly used to represent well-categorized concepts and to support mappings between models. The central ontology has been designed for the system and expressed in OWL [1][2] so as to semantically assign diseases represented in various coding standards to a single unified entity.

The rest of the paper is organized as follows. Section 2 provides some background information and literature survey. Section 3 presents the architecture and implementation details. Section 4 discusses some design decisions and their impact on the implementation. Section 5 summarizes the research outcomes and outlines future work.

II. Literature Survey

In [3] Iqbal et al. propose an ontology based model for an Electronic Medical Record that targets Chronic Disease Management with focus on providing a coherent information structure to support other acute diseases and co-morbidities. However, this model is patient-centric as it comprises longitudinal information of the patient, but it lacks the capability of collecting data from heterogeneous system.

In [4] Sampalli et al. propose an ontology model for patient profiles. In this model, frequently occurring procedure/diagnosis terms are extracted from patient charts, converted into SNOMED CT(Systematized Nomenclature of Medicine-Clinical Terms) and then categorized using the ontology concepts. The categories include medical, physical, psychosocial, rehabilitation, and nutrition. However, since the rules associated with the ontology are applied on patient charts for recurrent medical terminologies, the degree of accuracy between the doctors' diagnosed term and the converted concept identifiers cannot be easily measured. Such a study can be helpful for statistical analysis but not on patient treatment.

In [5] Ngamnij Arch-int et al. describe a semantic bridge ontology that provides Web services with the same service information but different parameters – according to the proprietary database structures – to integrate or exchange data and obtain common semantic meanings.

Iqbal et al. and Sampalli et al. demonstrate the use of ontology based rules in conceptual categorization and the value-added of ontology to enhance the expressiveness of concept. In our work, we have used "criticality" in "Infectious Disease" as the key indicator of expressiveness based on the diagnosis code. Regarding Arch-int et al. the ontology rules are not used to create a conceptual model for the relationships of the data from the disparate systems.

III. Our Infectious Diseases Reporting System

A. *Design*

We have divided our system into the following modules: Interoperable Web Service, Ontology-based Infectious Disease, Semantic Bridge to communicate with other health authorities, and Business Intelligence and Reporting.

Our infection diseases reporting system is built upon a set of Web services that interface with existing systems to collect infectious diseases. The content of these systems is structured differently with different standards (e.g., ICD -10 [6] or ICD -9 [7]). It is worth nothing that some local healthcare facilities send their disease codes as "descriptive text" to Web services. Thus we need an ontology bridge to map counterpart meanings of single concept (e.g., disease name "malaria") that are defined by different standard codes into a uniform entity.

The regulatory body defines "criticality" of each infectious disease. Each infectious disease has several characteristics properties but we focused only on "Criticality. We utilized OWL to build the relationship between ICD-10 codes and ICD -9 [7] class as "union-of" relationship to streamline a unified infectious disease classification range: Highly Critical, Critical, etc. to establish an "infectious disease centric" ontology model on patient data. The collection of these ontology-based entities reflects our conceptual model. The role of "Ontology-based Infectious Disease" module is to build such a model in order to represent and reason about the "criticality" of infectious diseases.

For a single regulatory body, there is a necessity to communicate its captured infectious diseases incidents to other health authorities at state and national level. Therefore we provide a web service based mechanism to exchange collected incidents data. Each health authority defines its own web service WSDL (Web Services Description Language) description. Since the WSDL parameters for web service methods will be different for each health authority, although the information are same, it is necessary to build ontology based rule to get the semantic meaning of the appropriate parameters. The role of "Semantic Bridge" module is to bridge the internal artifact properties with the WSDL parameter names of external health authority's web services.

The "Business Intelligence and Reporting" module is built to provide the regulatory body a statistical research and education program.

The architecture of our infection diseases reporting system is built upon 5 layers (Fig. 1): Resource, Mediation, Application Layer, Business Intelligence and Reporting, and Semantic Bridge.

The Resource Layer hosts a Centralized Data Repository, which has Patient Profile information that is stored in an Object Relational Database.

The mediation layer hosts three components:

- *Interoperable Web service*

It is implemented in SOAP and facilitates the Infectious Diseases data retrieval that will be triggered by the facilities when a patient encounter of infectious disease occurs. The Web service converts the retrieved patient data into internal artifacts and passes the information through the Ontology Engine.

- *Ontology Engine*

It facilitates building the conceptual model of the "Criticality" of the disease based on the data retrieved in Web services by issuing SOQL query to the ontology rules on "Infectious Disease". Fig. 2 shows the "Architecture of Ontology engine for Semantic Bridge". The "Cholera" disease is denoted in OWL as follows.

```
InfectiousDisease rdf:ID="Cholera">
 <CriticalityCondition rdf:resource="&xsd;string">
HiglyCritical
</CriticalityCondition>
 <DiseaseDescription rdf:resource="&xsd;string">
Cholera is an infectious Disease
</DiseaseDescription>
<owl:unionOf rdf:parseType="Collection">
 <owl:Thing rdf:about="#CholeraICD10" />
 <owl:Thing rdf:about="#CholeraICD9" />
</owl:unionOf>
```

- *Alert system*

It is implemented as an e-mail notification system that informs the regulatory body authority when a patient has some infectious diseases.

The Application Layer consists of Web interface screens for viewing Patient profile data and also for assigning proper coding to the reported Infectious Disease if it is submitted as a descriptive text without standard coding terminology. If the "descriptive text" contains medical terminology, there is an automated search of the ICD -10 with that medical term. If this search manages to fetch the appropriate code, it will stamp that internal artifact disease coding as "Automated text conversion". We have built web portal on top of the ICD -10 to search for codes or description, which will help the medical coding experts to assign "descriptive text" proper coding and convert them to conceptual model.

The Business Intelligence and Reporting Layer retrieves the data stored in Centralized Data Repository and transforms them into Dimensional data model and stored in "Warehouse" Object Relational database. We have built "Reports".

The Semantic Bridge Layer allows to address the problem of semantic discrepancies of WSDL parameter names. We have built some common ontology rules on the parameter names between the internal artifact property names and WSDL parameter names. For instance, let us take an internal artifact property called "patient name" that is codified as "PatName" using a WSDL parameter and "PatientName" as another internal artifact property. These two parameters are semantically the same and hence, are described in OWL model as "same-as" relationship. This mapping is illustrated for other parameters in Table I.

TABLE I.
WSDL PARAMETER NAMES WEB SERVICES ARE MAPPED TO INTERNAL ARTIFACT PROPERTY

Internal artifact property name	Wsdl parameter name for the web service 1	Wsdl parameter name for the webservice 2
PatientName	Name	PatName
MRNNo	MedicalRecordNumber	MRNNumber
Healthcard	HealthcardNo	HealthcardNumber
DateOfBirth	DOB	BirthDate
Gender	Sex	Gender
DiagnosisCode	ICD10Code	ICD9Code
DiagnosisStatus	Status	Status
PhysicianName	ClinicianName	Providername

The following "Patient Name" property in OWL

```
<Parameter rdf:Id="PatientName">
 <parameterFrom>InternalArtifact</parameterFrom>
 <parameterID>PatientName</parameterID >

<owl:sameAs rdf:resource="#Name">
<owl:sameAs rdf:resource="# PatName">

</ Parameter>
<Parameter rdf:ID="Name">
 <parameterFrom>Webservice1</parameterFrom>
 <parameterID>Name</parameterID>

</ Parameter>

<Parameter rdf:ID="PatName">
 <parameterFrom>Webservice2</parameterFrom>
 <parameterID>PatName</parameterID>

</ Parameter>
```

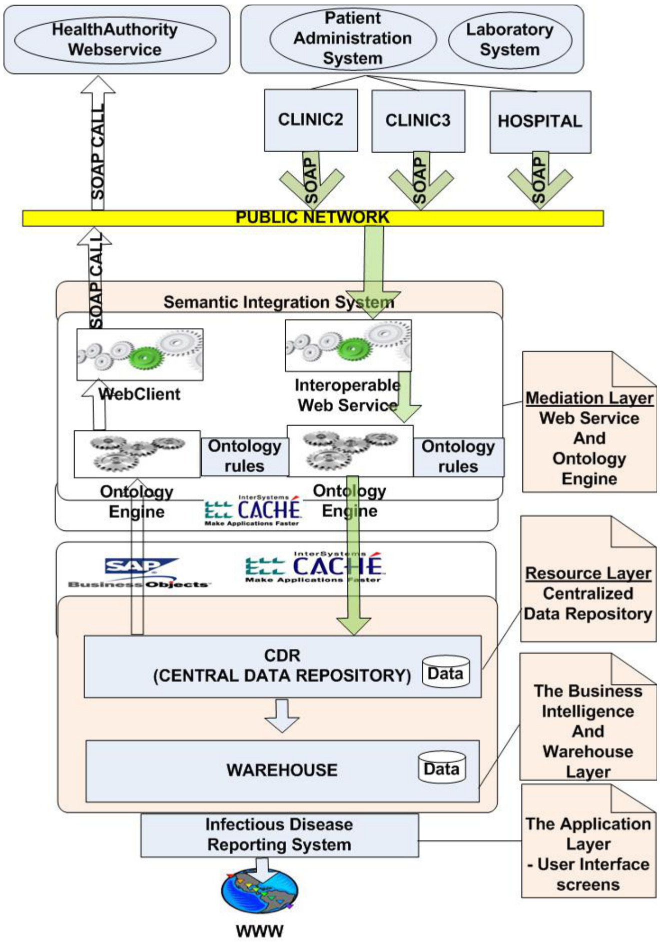

Fig 1 Architecture of Infectious Disease Reporting System

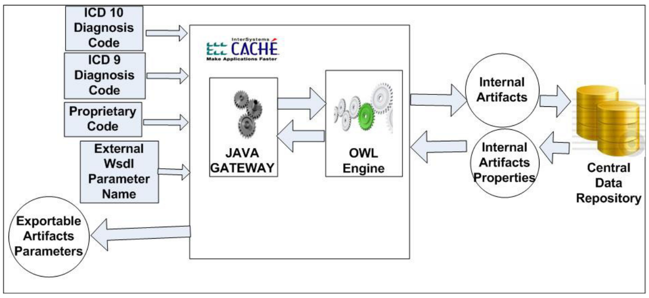

Fig 2 Architecture of Ontology Engine for Semantic Bridge

B. *Dataflow*

When a patient is diagnosed with an infectious disease, the clinical system launches a Web service that will send the data to our system. These data go will be managed by some ontology rules before they get stored in the Central Data Repository (CDR) as "Patient Profile". Moreover these data will be subject to some mining as follows.

1. The Healthcare facilities send the data of infectious disease to the Web service.
2. The Web service delivers the data to the semantic engine. It is implemented on Intersystem's cache having a "Java Gateway" and "OWL" engine. This engine converts the data into a conceptual model of internal artifacts.
3. The internal artifacts finally are stored in the "Central Data Repository".
4. The data are mined in the "warehouse repository" for statistical reports and charts generation.
5. To report the infectious disease to an external health authority Web service, the ontology relationships are established between the internal artifacts parameters and the external callable WSDL parameters. This will match the semantic meaning of the WSDL parameters and makes web service client as semantically integrated with the external WSDL parameters.

C. *Technologies Used and Delivered Components.*

We have built the web service on top of Intersystem's Cache implementation. The infectious disease data collected is transformed into data mining for the statistics research, which is built on Business Objects. The web pages are built on "Cache Server Page". For the transformation of the data to the conceptual model with OWL for both Infectious Diseases and the Semantic wsdl parameters, the jena engine is implemented. Table II shows the technology components and their vendors list. Table III shows the components designed and delivered.

TABLE II.
SUMMARY OF IMPLEMENTATION TECHNOLOGIES

Technologies	Brand
Database & Language	Intersystem's Cache [8]
Reports & Graphs	Business Objects [9]
OWL	Jena[10]

TABLE III.
MAJOR COMPONENTS DESIGNED AND DELIVERED

Component	Description of the Component
Web service	This service exposes web method "ReportDisease" and collects for the data submitted from the external request from the Healthcare facility medical systems
User Interface screens	Web page for searching ICD 10 disease codes and description. Used by medical coding experts to assign proper coding
Semantic Bridge	To achieve semantic interoperability, we have used the ontology based rule on the internal artifact property name with the wsdl parameter names of the external web services
E-mail alerting system.	Automated e-mail system was built to notify the regulatory body for the infectious disease reporting.
Reporting system	We have built data warehouse for statistical research with ad-hoc reporting tool on Business Objects
Security	Security Layer on the Web pages with the authentication on Active Directory

IV. DISCUSSION

The architecture presented in this paper lays down the basis for the development of an Interoperable medical application that targets multi-regulators "Infectious Disease Reporting System". The ontology developed in this study adds another layer of expressiveness whereas Web services secured the reach ability of the existing systems across different platforms. To achieve semantic integration, we have built a set of ontology rules on WSDL parameters with the internal artifacts parameters. The key limitation in this work is the inability to capture medical tests and results, which form together the fundamental support for confirming diagnosis of infectious disease. To achieve this goal, we have to expand the array of data to accommodate medical tests and their results by providing ontology vocabulary on the LOINC codes to convert them into conceptual model.

V. CONCLUSION AND FUTURE WORK

This research work is driven by the immense needs and posed challenges for integrating different healthcare systems to monitor and manage public health indicators such as infectious diseases. As such, this work can be extended in many aspects such as putting forward a policy-based framework to allow injecting regulators rules. This direction should follow the approach specified by El-Hassan et al. [11] which allows specifying rules for accessing resources (e.g.. patients data) in both normal and emergency situations. Additionally, advances in web services that advocate collaboration between independent agents have to be incorporated. In this perspective, web services research presented by Khosravifar et al. [12] demonstrate a viable candidate for community-based web services that manage service-based interactions between equal parties.

APPENDIX

A. Section of WSDL Description of Web service

```
<?xml version='1.0' encoding='UTF-8' ?>
<definitions xmlns:http='http://schemas.xmlsoap.org/wsdl/http/' xmlns:SOAP-ENC='http://schemas.xmlsoap.org/soap/encoding/' xmlns:mime='http://schemas.xmlsoap.org/wsdl/mime/' targetNamespace = 'http://HIRAS.ae' xmlns='http://schemas.xmlsoap.org/wsdl/'>
<types>
<s:schema elementFormDefault='qualified' targetNamespace = 'http://HIRAS.ae'>
<s:element name="ReportDisease">
<s:complexType>
  <s:sequence>
    <s:element name="FacilityID" type="s:string" minOccurs="0" />
    <s:element name="PatientName" type="s:string" minOccurs="0" />
    <s:element name="MRNNo" type="s:string" minOccurs="0" />
    <s:element name="Healthcard" type="s:string" minOccurs="0" />
</s:schema>
</types>
<message name="ReportDiseaseSoapIn">
  <part name="parameters" element="s0:ReportDisease" />
</message>
<message name="ReportDiseaseSoapOut">
  <part name="parameters" element="s0:ReportDiseaseResponse" />
</message>
<portType name='InfectiouDiseaseReportingSoap'>
  <operation name='ReportDisease'>
    <input message='s0:ReportDiseaseSoapIn' />
    <output message='s0:ReportDiseaseSoapOut' />
  </operation>
</portType>
<binding name='InfectiouDiseaseReportingSoap' type='s0:InfectiouDiseaseReportingSoap' >
    <soap:binding transport='http://schemas.xmlsoap.org/soap/http' style='document' />
  <operation name='ReportDisease' >
    <soap:operation soapAction='http://HIRAS.ae/CDR.InfectiousDisease.ReportDisease' style='document' />
    <input>
      <soap:body use='literal' />
    </input>
    <output>
      <soap:body use='literal' />
    </output>
  </operation>
</binding>
<service name='InfectiouDiseaseReporting' >
  <port name='InfectiouDiseaseReportingSoap' binding='s0:InfectiouDiseaseReportingSoap' >
    <soap:address location='http://HIRAS.ae/CDR.InfectiousDisease.cls' />
  </port>
</service>
</definitions>
```

B. OWL representation of Infectious Disease

```
<owl:Class rdf:ID="InfectiousDisease">
<rdfs:label xml:lang="en">InfectiousDisease </rdfs:label>
</owl:Class>
<owl:ObjectProperty rdf:ID="CodeScheme">
<rdfs:domain rdf:resource="#InfectiousDisease"/>
<rdfs:range rdf:resource="#CodingScheme"/>
</owl: ObjectProperty >
<owl:DatatypeProperty rdf:ID="DiseaseDescription">
<rdfs:domain rdf:resource="#InfectiousDisease "/>
<rdfs:range rdf:resource="&xsd;string"/>
</owl:DatatypeProperty>
<owl:DatatypeProperty rdf:ID="CriticalityCondition">
<rdfs:domain rdf:resource="# InfectiousDisease"/>
<rdfs:range rdf:resource="&xsd;string"/>
</owl:DatatypeProperty>
```

ACKNOWLEDGEMENT

This research work would have not been completed without the enormous support of Dubai Healthcare City's top executives who granted us access to their "Arab Health Award Winning Product, namely,"HIRAS- Healthcare Information Reporting and Analysis System". We particularly thank the CEO, Dr. Ayesha Abdulla and Dr. Abdulkareem Al-Olama. Additionally thanks go to Imad Choucair and Naima Shaikh from the CIO office in TECOM Investment for the continuous encouragement and support.

REFERENCES

[1] M. K. Smith, C.Welty, and D. McGuinness. OWLWeb Ontology Language Guide. W3C Recommendation, *http://www.w3.org/TR/owl-guide/* ,May 14, 2011.

[2] W3C. Owl web ontology language-reference. LSDIS Lab, University of Georgia,2004. http://www.w3.org/TR/owl-ref/.

[3] Iqbal, A.M.; Shepherd, M.; Abidi, S.S.R., "An Ontology-Based Electronic Medical Record for Chronic Disease Management" in *Jan. 2011 44th Hawaii International Conference* pp. 4–6.

[4] Sampalli, T. Shepherd, M. Duffy, J.., "A Patient Profile Ontology in the Heterogeneous Domain of Complex and Chronic Health Conditions" in Jan. 2011 System Sciences (HICSS), *2011 44th Hawaii International Conference*.pp 4-6.

[5] Arch-int, N.; Arch-int, S "Semantic information integration for electronic patient records using ontology and web services model" in *April. 2011 Information Science and Applications (ICISA), International Conference*.pp 3-5

[6] International Classification of Diseases. *http://www.cdc.gov/nchs/icd/icd10.htm,* May 14, 2011.

[7] International Classification of Diseases, Ninth Revision (ICD-9) *http://www.cdc.gov/nchs/icd/icd9.htm*, May 24, 2011.

[8] Intersystems Cache. *http://www.intersystems.com/cache/*, May 24, 2011.

[9] Business Objects *http://www.sap.com/solutions/sapbusinessobjects/index.epx*, May 24, 2011

[10] Jena A Semantic Web Framework for Java. *http://jena.sourceforge .net/* , May 14, 2011.

[11] El-Hassan, Osama and Fiadeiro, Jos'e Luiz and Heckel, Reiko "Managing socio-technical interactions in healthcare systems" in 2008 *Proceedings of the 2007 international conference on Business process management*.

[12] Khosravifar, B.; Bentahar, J.; Moazin, A.; Maamar, Z.; Thiran, P. "Analyzing Communities vs. Single Agent-Based Web Services: Trust Perspectives", in *July 2010 Services Computing (SCC), 2010 IEEE International Conference*.