ABSTRACT
The integration of Electronic Patient Record (EPR) systems is at the centre of many of the new regional and national initiatives to integrate clinical processes across department, region, and national levels. Web Service technologies offer significant solutions to provide an interoperable communication infrastructure but are unable to support precise definitions for healthcare messages, functionality, and standards, required for making meaningful integration. The lack of interoperability within healthcare standards adds complexity to the initiatives. This heterogeneity exists within two versions of same standard (e.g. HL7), and also between standards (e.g. HL7, openEHR, CEN TC/251 13606).

We therefore introduce an integration platform PPEPR (Plug and Play Electronic Patient Records), which is based on the principles of a semantic Service-Oriented Architecture (sSOA). PPEPR solves the problem of interoperability at the semantic level. A key focus of PPEPR is that once a patient information is captured, should be available for use across all potential care processes.

Categories and Subject Descriptors
D.2.12 [Software]: Software Architectures; Interoperability [Domain-specific architectures]

Keywords
eHealth, SOA, Web services, Semantic Interoperability

1. INTRODUCTION
Healthcare enterprise applications (e.g. Electronic Patient Record Systems (EPRs)) have been created within hospitals to address specific end user requirements. The EPRs, however, exist as islands of information with little or no connectivity between them. What connectivity has been implemented has been a largely manual effort with significant resources spent in routing and mapping healthcare messages between those systems [8]. In some cases special interfaces have had to be developed to allow EPR systems to communicate. Consider the following set of problems:

• Most hospitals still have obsolete standards or protocols running as their critical applications.
• Healthcare standards are under constant development and improvements.
• IT or healthcare professionals may diverge from the use of healthcare standards (e.g. HL7v2.x, v3), CEN TC/251 13606, openEHR etc.) intended meanings, thus defeating the purpose of those standards.
• Healthcare standards formatted in XML solve the integration or interoperability problem at a syntactic level, but domain specific solutions are required to achieve meaningful integration.

In this paper we introduce PPEPR, an integration platform for heterogeneous EPRs. PPEPR is based on the design principles of a semantic SOA Reference Architecture and built (It’s first prototype) around semantic Web service technology. The structure of this paper is as follows: first we describe the relation between healthcare, SOA, Web services, and the emergent need of semantics. Then we introduce the PPEPR approach and a simplified version of the PPEPR architecture. Next, we introduce our approach for ontology and an adapter framework development. Finally we explain how PPEPR addresses the EPRs integration requirements.

2. EPRS INTEGRATION: SOA, WEB SERVICES, & SEMANTICS
Gradually, organisations are adopting SOA as their fundamental architecture for systems development [4]. For healthcare systems there are two conceptual viewpoints (both are valid):

• Implementing a general SOA framework (common infrastructure, tools and approaches), “HL7 is just another content type”

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

SAC’08 March 16-20, 2008, Fortaleza, Ceará, Brazil
Copyright 2008 ACM 978-1-59593-753-7/08/0003 ...$5.00.
Implementing an HL7-based messaging architecture that can use different messaging and transports, including Web services.

The first tends to lead to the conclusion that HL7 should just define content and the second suggests HL7 should define the whole stack. Various initiatives focus on each of the viewpoints; creating an overall enterprise SOA where HL7 is just one content type amongst others (e.g. X.12), and the need for HL7 to fully define an operating model for those using less complete frameworks such as basic HL7 LLP messaging. These divergent viewpoints map to issues on the use of the various wrapper layers around HL7 content, the level of overlap between HL7 functionality, and advanced general frameworks such as Web Services. The majority of existing EPRs do not employ Web services, creating significant challenges for any integration system based on advanced research/technology (e.g. semantic Web services), without imposing constraints on existing systems (e.g. requiring the upgrade of existing EPRs to use Web services).

This transition to the use of Web services imposes significant costs to the implementing enterprise. To counter this difficulty PPEPR provides three types of integration between non-Web services and Web services enabled EPRs.

Web services provide the technology foundation for implementing and delivering SOA platforms. However, SOA itself is not a complete solution for the integration of healthcare information systems since the two core challenges of conventional computing - search and integration - (also known as semantic gap of SOA) are not addressed. Therefore, SOA significantly and fundamentally depends on solutions to fill the semantic gap so that its full potential can be achieved.[1] The PPEPR architecture employs a semantic Service-Oriented Architecture (sSOA) and thus acts as such a solution. The architecture will be realized by developing a cross-functional, semantically enabled integration platform with input from Galway University Hospital (GUH).[2]

3. PPEPR APPROACH

The development approach is divided into two steps: First, Modelling the healthcare domain knowledge at a conceptual level (during design time) by, (a) annotating healthcare functionalities (Functional ontology) (b) annotating healthcare messages (Message ontology) and, (c) mapping ontologies at the schema level. Second, Integrate EPRs (during run time) using the modelled concepts by, (a) mediating ontologies at instance level (b) mediating healthcare messages and functionalities at instance level. The above approach helps to delineate scope and responsibilities as well as organise, maintain and access the healthcare data, information and knowledge. The key outcomes of PPEPR are Ontologies (Message, Functional) and the Adapter Framework to connect to traditional healthcare software.

4. PPEPR ARCHITECTURE

Figure 1 shows the simplified architecture of PPEPR. The main components are:

- **EPRs**: In the first prototype, PPEPR incorporates HL7 (2.x, v3, and CDA) compliant EPRs. In advanced versions

Figure 1: Simplified PPEPR architecture

of PPEPR, openEHR and CEN TC/251 13606 will be incorporated within the integration framework.

**Adapter Framework**: The adapter framework is a common platform for enabling communication between systems that use different data formats and it is an application specific software component that transforms data from its source to target format. The framework provides an adapter skeleton and allows developers to register (plug-in) their adapters and invokes the target application through one of the registered adapters. Direct invocation is also possible if the input message is described in the target-required format. The adapter’s main role is at run-time. Grounding rules executed by adapter are specified at design-time. Along with message transformations, the adapter will also create a request (goal) describing services to be invoked at the receiving ends. These goals will later be resolved by the semantically-enabled middleware during process mediation. The main components within the Adapter Framework are:

- The Communication Manager is a gateway to the adapter framework which can be accessed by MessageListener via a Web-service interface or TCP/IP port (for non-Web Service EPRs). The Message Listener waits for any new message incoming from the network. Once requested by a back end application, the Communication Manager interprets the request and invokes either the Adapter Manager or the Update Manager. If the invocation is successful, it returns a response message which contains either a 'success' message or the 'response' returned from the target application.

- The Adapter Manager is the core component of the Adapter Framework which schedules job between adapters and the data format Validator. It updates and refreshes the list of the adapters as they are registered and unregistered.

- The Validator is responsible for validating the data format as claimed by the sending application. An XML message, for example, sent by back end application, should pass the XML data format validation process.

- The Adapter Pool is a “repository or container” of all the registered adapters.

- The XML2WSML adapter is a specialized adapter which transforms (lifts) the XML message (syntactic) to Web ser-
service modeling language (WSML) (semantic).

The WSML2XML adapter is a specialized adapter which transforms (lowers) the WSML (semantic) to XML message (syntactic).

The Database Manager manages database connection, exceptions, and utilities.

**Semantics-Enabled Middleware (WSMX)**: A detailed description of WSMX, WSML, and Web service modeling tool (WSM+), is outside the scope of this paper, therefore we explain WSMX and WSML from the PPEPR perspective, mainly focusing on data and process mediation. Healthcare standards define a set of messages and functionalities which may not match across standards or systems. These functionalities are exposed as Web services and the first step towards semantic Service-Oriented Architecture (sSOA) is to create semantic descriptions of Web services (during design-time), listing 1 shows the snippet of WSML Web service description of HL7 v3 order filler Web service.

Listing 1: WSML Web service description of the order filler Web service

In listing 1(top to bottom), the namespace section describes the identifiers (IRIs, sQNames) for message ontology (MO) and functional ontology (FO). The usesMediator keyword identifies the mediator (e.g. ooMediator) which connects different Goals, Web Services and Ontologies, and enable inter-operation by reconciling differences in representation formats, encoding styles, business protocols, etc. In capability section, preconditions describe the condition over information space state before the execution of a Web Service. For example, non-functional properties describes the type of message accepted as input and corresponding HL7 v2 OrderFiller Web Service (e.g. HL7 v2 and HL7 v3 complaint EPRs interactions). Interface describe how to interact with a service from the requester point of view (choreography) and how the service interacts with other services and goals it needs to fulfill in order to fulfill its capability (orchestration), which is the provider point of view. The OrderFillerInterface describe transition rules as a part of choreography, for example, request and corresponding Acknowledgment (response) messages. In orchestration section, the input and output messages and respective WSDL grounding is defined, for example, wsdlinterfaceMessageReference (ServiceInterface/ Operation/input or output message).

The listing 1 describes only two messages (OrderActivateFulfillmentRequest, OrderActivateFulfillmentRequest_Ack), mediator used, choreography, and orchestration associated with these messages. The complete message exchange pattern is shown in Figure 2 and dotted lines within listing 1 indicates the description of all the messages exchanged by orderfiller Web service. The WSML web service description are stored in semantically-enabled repository (e.g. WSMO4RDF repository within ORDI SG framework), WSMX picks-up these service descriptions automatically during discovery, composition, choreography, and orchestration of Web services. The WSML Web service descriptions are created manually during each setup of existing and/or new EPRs, which is a major integration task for HL7 compliant EPRs within PPEPR integration platform.

The WSMX mediation (data and process) components have three main functions; one used during the design time and the other two used during runtime. First, a set of mappings is created by domain experts (e.g. healthcare professionals), using a mapping tool - WSMT. Second, the "Mapping Rule Generator" within mediation components is called internally during runtime to transform the mappings into mapping rules. Mappings express the similarities between the two ontologies, while the mapping rules describe how these similarities are used in order to transform instances of the first ontology in instances of the second ontology. Mapping rules express the mappings in an executable way. The third and final step of the mediation process is the execution of the mapping rules. These mapping rules are received from the "mapping rules generator" and executed inside the WSMX rule execution environment. 

---


---

5. Ontology Development

Ontology development for PPEPR includes message ontologies to represent the information sent between the components and functional ontologies to represent the ordering of relations in the execution of the message exchange patterns.

5.1 Message ontology

The ontology of a healthcare standard is developed from the specification provided in the XML Schema format. This is done by processing grounding rules defined by the developers that transform the HL7 XML Schema to an ontological representation (Bi-directional). Heterogeneities at the ontological level are identified and mapping definitions between ontologies are created using an ontology modelling tool WSMT. Both Grounding and Mapping definitions at the conceptual level are stored and used by the adapter and semantic enabled middleware at run-time.

Listing 2: XML Message (Simplified version): HL7 v3 laboratory observation order activates fulfillment request

Listing 2 shows a simplified XML message from a HL7 v3 compliant EPR and Listing 3 shows the equivalent WSML instance. A detailed explanation of the HL7 message structure and the XML to WSML transformation process are outside the scope of this paper. Therefore we will briefly explain the syntactic (XML) and semantic (WSML) relationship of both the messages. The message POLB_IN002120 (HL7 v3) is a “laboratory observation order activate fulfillment request” message. This interaction message is created when a “fulfillment request” is communicated between the “order placer” and the “fulfiller” (e.g. Figure 2). Grounding Rules (e.g. XSLT\textsuperscript{15}) are applied to transform the POLB_IN002120 xml message to the WSML instance. For example, the “controlActProcess” tag in the XML message is equivalent to the “controlActProcess” relation within the WSML instance. Similarly, the “given” tag of the XML message is equal to the “given” relation instance. Thus, the “Dr. Rise” value in the “given” tag transforms into the second parameter of the “given” relation instance. Instances of WSML concepts are used in relation instance parameters to link the different relation instances. This is similar to the way complex types in XML Schema maintain a nested form within an XML message.

\textsuperscript{15}http://www.w3.org/TR/xslt

Listing 3: WSML Instance (Simplified version): HL7 v3 laboratory observation order activates fulfillment request.

WSML uses similar XPath\textsuperscript{16} patterns to represent “instance” as well as “relationInstance” of concept and the relation, therefore it looses the structure of original XML message. To resolve this issue during the lowering process, we use XPath’s context node feature to uniquely identify instance, relation instance and attribute values. For example, the “given” relationInstance (see second last line listing 3) and its first parameter “en” are pointed to by the contexts “//wsml:relationInstance [wsml:memberOf = ’given’]” and “//wsml:parameterValue [wsml:parameter= ’parameter1’]/wsml:value” respectively (see listing 4).

Listing 4: XSLT (Simplified version): Lowering WSML Instance (Listing 2) to XML message (Listing 1)

5.2 Functional ontology

HL7 categorises healthcare events that are used to annotate healthcare functionalities. Similar to standards in other domains, such as RosettaNet\textsuperscript{17} for Supply Chain Management in ICT, HL7 not only defines the message content, but also the business logic to achieve certain functionality in the health care domain. Figure 2 shows the choreography

\textsuperscript{16}http://www.w3.org/TR/xpath
\textsuperscript{17}http://www.rosettanet.org/
(message exchange patterns) between the order placer and the order fulfiller whereas the activity diagram in figure 3 shows the process model of the order placer to achieve the actual healthcare process. It is sufficient if both parties, the process placer and the process fulfiller model and execute a process according to the message exchange patterns defined in HL7 and shown in the left and right part of figure 3.

Figure 2: Interaction Diagram in HL7 Fulfillment Request

```
Figure 3: Business Process of Order Placer in a HL7 Fulfillment Request
```

To model and execute such message exchange patterns, it is necessary to employ a process modelling and execution standard which is able to reference ontological elements and allow their mapping within the model. BPEL4SWS introduces a new mechanism to reference ontological elements within a business process description. BPEL4SWS enables defining a complex message exchange between two partners without dependency on WSDL which allows us to include goals in the execution process. BPEL4SWS introduces a new element, the `<b4s:conversation>` element, independent from WSDL. This element enables grouping of interaction activities and thus enables defining a complex message exchange between two partners.

Interaction activities in BPEL4SWS link messages to goals. The interaction activities are grouped in a conversation which allows us to link to a Web Service Modeling Ontology (WSMO) Goal[5]. The WSMO Goal describes what the process expects from a service that plays the partner role in this particular interaction by means of the requested capabilities.

Listing 5: Snippet of BPEL4SWS according to HL7 Fulfillment Request

```
Listing 5: Snippet of BPEL4SWS according to HL7 Fulfillment Request

6. EPR INTEGRATION: PPEPR

Healthcare is a complex domain, comprising vendors, standards, legacy systems, and information systems which inherently differ from one another. PPEPR provides a unique approach to interoperability. The core solution lies in enabling semantic interoperability between existing and new EPR systems. The PPEPR architecture considers three types of integration between EPRs based on their web service capabilities (or lack thereof).

**Type 1: Existing EPR [non-Web services]** `<==>` Existing EPR [non-Web services]

This type of interaction is focussed on existing EPRs, which are mostly HL7 v2.x based. At this level EPRs will be able to exchange messages in EDIFACT and XML format. The majority of current HL7 v2.x based systems do not use Web services, and their immediate requirement is to exchange messages. The PPEPR adapter can interact bi-directionally with these EPRs at a specified TCP/IP port.

Limitations: The goal of PPEPR is to achieve the maximum level of interoperability between existing and/or new systems, without upgrading existing or introducing new EPR systems. Even though HL7 v2.x has categorized the events in healthcare by considering service functionality which reflects the business logic in this domain, implementations are not service-oriented. These service functionalities could be used to deploy Web services by upgrading existing EPR systems to Web service enabled EPR. This transition, although

18http://lists.w3.org/Archives/Public/public-ws- semann/2007Jun/0000
19http://www.wsmo.org/
20http://en.wikipedia.org/wiki/UN/EDIFACT
outside the scope of PPEPR, would enable those systems to benefit from advanced interoperability.

**Type 2:** Existing EPR [non-Web services] $\iff$ Web-Service enabled EPR

This type of integration is the most complex (e.g. HL7 2.x $\iff$ HL7 v3), since EPRs (non-Web services) are required to communicate with the other EPRs (Web-services) via SOAP.

Limitations: As, one side of the EPR integration process does not employ Web services the Web Service Enabled EPR cannot invoke the services directly. PPEPR’s Adapter can receive the message at a TCP/IP port, converting the message to XML if the received message is via EDI, creating a SOAP message, then invoking the EPR at the other end, after mediation. This transformation and routing mechanism is defined at design time. Therefore, two or more EPRs at this level should be aware of each other, before they interact. Also, this type of integration cannot enjoy the interoperability advantages (e.g. discovery, composition, orchestration, etc.) of semantic Web services at the service and/or process level but achieves the major goal of sharing healthcare messages.

**Type 3:** Web-Service enabled EPR (1) $\iff$ Web-Service enabled EPR (2)

This type of integration in PPEPR architecture offers the best interoperability solution by achieving syntactic as well as semantic interoperability.

7. RELATED WORK

PPEPR approach is based on semantic Service-Oriented Architecture (sSOA) and other projects which employs the similar principal are:

- COCOON is a 6th Framework EU integrated project aimed at setting up a set of regional semantics-based healthcare information infrastructure with the goal of reducing medical errors. In order to enable a seamless integration of eHealth services, Semantic Web Services technology is applied.
- ARTEMIS is a STREP project supported in the 6th Framework by the European Commission. ARTEMIS aims to develop a semantic Web Services based Interoperability framework for the healthcare domain. Artemis has a peer-to-peer architecture in order to facilitate the discovery and consumption of healthcare web services.
- RID and SemanticHEALTH projects are EU road map projects with Special Emphasis on Semantic Interoperability. The road map will be based on consensus of the research community, and validated by stakeholders, industry and Member State health authorities.

The major differences between eHealth projects described above and our approach are:

- Other projects impose constraints on the EPRs (e.g. transition from traditional to Web services) but our aim is to integrate them without imposing any constraint on existing or proposed EPRs.
- None of the other projects provide an integration solution between non-service-oriented and service-oriented (Web services).

- PPEPR applies a new mechanism to describe the communication between two partners without a dependency on WSDL. As described in section 5.2, BPEL4 SWS introduces a new element, $\langle$b4s:conversation$\rangle$, which is not dependant on a $\langle$partnerLinkType$\rangle$ and as such is not dependant on WSDL. This element enables the grouping of interaction activities and thus enables defining a complex message exchange between two partners. For example, Artemis introduced business process template “BP template” to model business process at design time. BP template is dependent on WSDL (e.g. $\langle$partnerLinkType$\rangle$) to describe a contract between two partners in terms of roles and corresponding WSDL portTypes. Also, for interaction activities “BP template” mainly relies on partnerLink and a WSDL operation.

8. CONCLUSIONS

Current healthcare enterprise applications need greater flexibility and scalability to meet the challenges of heterogeneity of healthcare systems at all levels - data, process, services, etc. The architecture of any integration system holds the key to offer a dynamic, flexible and scalable solution. As we have discussed above, healthcare is a complex domain. Actors include vendors, standards, legacy systems, and information systems all of which must interoperate to provide healthcare services. PPEPR provides an interoperability solution without imposing any constraint on existing or proposed EPRs. The major advantage of PPEPR approach is the “Plug and Play” feature, which requires minimal configuration during setup. Within scope of PPEPR, EPRs are categorized based on international healthcare standards: (1) HL7 v2.x, (2) HL7 v3, and (3) CEN TC251 13606/openEHR. These three categories of standards are quite different in their overall approach towards an EPR, but share many semantic similarities, especially the Reference Information Models (RIM) of HL7 v3, and CEN TC251 13606/openEHR. These similarities are the linchpin of ontology development and mapping.

We have presented two major aspects of PPEPR: ontology specification and adapter framework development. Each ontology specified as part of this project is further categorised into a message and a functional ontology. A functional ontology describes the semantics required for interaction between EPR systems, based on the interaction events within healthcare standards to provide domain based workflows between interactive EPR systems. A message ontology explicitly defines the semantics associated with healthcare messages required to be exchanged for data interoperability between EPR systems. Differences in these specifications are resolved by semantic mapping techniques.

The adapter framework receives the grounding rules developed at design-time and plays a crucial role in integrating EPRs at run-time. The adapter framework follows a similar design pattern to that of the Java Connector Architecture (JCA). Similarity in features include component based development and deployment, flexibility in deployed adapters, and automated data format validation.

The first prototype of PPEPR has successfully integrated HL7 v3 and HL7 v2.x within an experimental environment; 25

\[\text{http://www.semantichealth.org/}\]

\[\text{http://java.sun.com/j2ee/connector/}\]
Future developments include mapping and mediating between HL7, CEN TC251 13606, and openEHR standards. PPEPR will be validated in conjunction with end user clinicians in leading Irish Hospitals. In the first prototype of PPEPR we have not yet covered all integration issues (e.g. security) during adapter development. However, since we employ component based development, future components can be easily integrated. Also, this paper focuses on introducing PPEPR architecture, data and process mediation. Other service-oriented features such as publishing, discovery, etc. are outside the scope of this paper.

9. ACKNOWLEDGMENTS
We would like to thank Brahmananda Sapkota, Armin Haller, and James Cooley, for their comments and input to this document. This material is based upon works supported by the Science Foundation Ireland project Lion under Grant No. (SFI /02/CE1/I131 ) and by Enterprise Ireland under Project PPEPR (CFTD 2005 INF 224).

10. REFERENCES