DEVELOPING INTEROPERABLE SEMANTIC E-HEALTH TOOLS FOR SOCIAL NETWORKS

Juha Puustjärvi
Helsinki University of Technology, Innopoli 2, Tekniikantie 14, Espoo, Finland

Leena Puustjärvi
The Pharmacy of Kaivopuisto, Nëitsytpolku 10, Helsinki, Finland

Keywords: Social networks, Web 2.0, e-Health, Medical applications, Semantic web, Ontologies, Knowledge management.

Abstract: Many studies have indicated that most patients are not satisfied with the medical treatment information on the Web though many e-health tools provide links to materials or other websites that have information about patient’s health conditions or medications. In addition, many studies have demonstrated that patients should have easy access to their own health information as well as to any information they need in order to make decisions about their own health care. However, while there are a variety of tools for managing and sharing medical information, no integrated tool for health information management and sharing has been developed. Satisfying this challenge requires a means to capture and interconnect information from various sources which are relevant to one patient and create personal health space containing links to the health information that are related to the customer or of which the customer is interested in. In this paper we describe our work on developing a personal health assistant, which integrates the tools supporting personal health records, information therapy and health oriented blogs. Technically the personal health assistant is based on knowledge management technologies, and it is easily extensible to capture additional e-health tools.

1 INTRODUCTION

Although the term social networking has started to be used fairly recently, online social networks existed before the Web in the form of email discussion lists and bulletin boards, and which are still commonly used. Nowadays social networking has encouraged new ways for communicating and sharing information (Childs, 2007). Social networking Web sites are regularly used by millions of people.

Advances in social networking and the widespread use of Internet are also changing the way health care is provided (Lewiset al., 2005). In particular, health care provision is moving from a disease oriented model, where the treatment decisions are made almost exclusive by physicians based on their clinical experience, to a patient oriented model, where patients are active participants in the decision making process about their own health (Tuil et al., 2006, Trevana et al., 2006). The term e-health is commonly used to describe this evolution in health care.

Many studies have indicated that most patients are not satisfied with the medical treatment information on the Web though many e-health tools provide links to materials or other websites that have information about patient’s health conditions or medications (Butcher, 2007; Kemper, 2008). In particular, they have regarded many sites to be overly commercial, or they could not determine the source of the information (Puustjärvi and Puustjärvi, 2008).

E-health covers many fields including electronic health records (Vesely, 2006, Ghani et al., 2008; EHR, 2009), personal health records (Kaelber et al., 2008; Raisinghani, 2008), evidence based medicine (Metler and Kemper, 2004), information therapy (Kemper, 2008) and disease management (Stalidis et al., 2001). There are also various kinds of e-health tools focusing different fields of health care, each...
having their own user interfaces (Puustjärvi and Puustjärvi, 2006).

A problem here is that these e-health tools do not interoperate. This lack complicates the usability of these systems as well as restricts the services that the e-health tools can provide.

In order to illustrate the potential gain of interoperation (or integration) let us consider the tools supporting personal health records (PHRs), Information therapy (Ix) and health oriented blogs. A PHR is a record of a consumer that includes data gathered from different sources such as from health care providers, pharmacies, insures, and the consumer (Angst et al., 2008). Its main goal is to provide a complete and accurate summary of the health and medical history of a consumer (Agarwal and Angst, 2006) while the goal behind information therapy is to prescribe the right information to right people at right time (Mettler and Kemper, 2004). Hence, by integrating PHR system and a system providing Ix we could automate Ix based on the content of the PHR as it describes patient health and medication history and the present state as well.

Further, by integrating e-health oriented blogs to PHR tool we could automatically deliver the blog items to patient, which are related to patient’s medication or illnesses. So from patient’s point of view the integrated health tool would be like an electronic newspaper that is personalized according to patient’s dynamic health and medication profile.

This kind of automation in delivering personal health information would be useful as many studies have demonstrated that the provision of information therapy can increase compliance with treatment regiments, satisfaction with the health care provider and medical facility, and improve the ultimate health outcome for the individual (Butcher, 2007). It is also turned out that patients who do not understand their treatment instructions, disease management, or prescription requirements are more likely to mishandle their health, be hospitalized more frequently, and have much higher medical costs than their more involved counterparts (Kaelber et al., 2008).

In this paper, we describe our work on the integration of the tools supporting PHRs, Ix and health oriented blogs. We call such integrated tools as personal health assistants or PHAs for short. In this paper we do not consider PHAs from security point of view though it is important aspect as it specifies how well the tool can provide confidentiality and nonrepudiation by authenticating the parties involved and maintaining access control. Instead the focus of this paper is to consider a PHA from interoperation and data management points of view.

The main advantage of a PHA is that patients do not need to navigate on the Web to find evidence-based medical information or relevant blog items. Instead the relevant information or its links are automatically delivered to patients’ PHA. Which information is delivered depends on the content of the patient’s personal health records, and thus the delivery can be targeted automatically.

The corner stone of the system is the PHA-ontology, which captures the ontologies used in PHR, Ix and in the blog. Technically PHA is based on knowledge management technologies and is easily extensible to capture additional e-health tools.

The rest of the paper is organized as follows. First, in Section 2, we consider the application integration/interoperation strategies used in developing the PHA, and the role of Knowledge base in storing the health information constituting the PHA-ontology. Then, in Section 3, we present the components of the PHA, and introduce the notions of Semantic blog, Semantic PHR and Semantic Ix.

After this, in Section 4, we consider the PHA-ontology, which is shared by all the interoperable e-health tools. We represent parts of the ontology in a graphical form as well as in OWL (Web Ontology Language OWL) (OWL, 2006). We also illustrate how ontology instances can be presented in RDF (Resource Description Language) (RDF, 2004). Finally Section 5 concludes the paper by discussing the advantages and disadvantages of our represented solutions.

2 THE ROLE OF KNOWLEDGE MANAGEMENT IN PHA

2.1 Application Integration and Interoperation Strategies

Basically the term integration refers to the idea of putting diverse concepts together to create an integrated whole (Singh and Huhns, 2005). Instead interoperation refers to making applications work together by sharing the appropriate messages but without any single conceptual integration.

Even though the approaches for the interoperation of various applications vary considerable, the principal distinction between Information-oriented, Process-oriented and Service-
oriented and Portal-oriented application interoperability can be done (Lithicum, 2004).

- In **Information-oriented** approaches applications interoperate through a database or knowledge base.
- In **Process-oriented** (also called workflow-oriented) approach the interoperability is controlled through a process model that binds processes and information within many systems.
- In **Service-oriented** application interoperability is achieved by providing the infrastructure for such method sharing.
- In **Portal-oriented** application integration a multitude of systems can be viewed through a single user interface, i.e., the interfaces of a multitude of systems are captured in a portal that user access by their browsers.

From user’s point of view our used application integration strategy follows the portal oriented approach as the multitude of e-health tools can be viewed through a single user interface. On the other hand, we use the Information-oriented approach in achieving the interoperability between the e-health tools. That is, the tools interoperate through sharing a knowledge base, and the ontology is developed by integrating the ontologies of the interoperable e-health tools.

### 2.2 Knowledge Management and PHA

Knowledge management (Daconta et al., 2003) concerns with acquiring, accessing and maintaining knowledge within an organization. Knowledge management system refers to a computer based system for managing knowledge in organizations. A knowledge base is a special kind of database for knowledge management. It provides the means for the computerized collection, organization, and retrieval of knowledge for various applications.

Today an ever expanding set of knowledge management systems are using the technologies of the Semantic Web. That is, knowledge is organized according to ontologies, and automated tools are used in accessing and maintaining knowledge.

In particular, knowledge management is considered to be important because organizations view internal knowledge as an intellectual asset from which they can draw greater productivity, create new value, and increase their know-how (Daconta et al., 2003).

We argue that this is also true with respect to e-health. That is, by acquiring, accessing and maintaining health-oriented knowledge we can develop e-health tools that create new value and increase the productivity of health care. Therefore both knowledge management and the development of e-health tools should be developed iteratively, so as to provide mutual feedback.

However, for the present, the deployment of the knowledge management technologies in e-health is quite limited. The main obstacle is that the developed schemas, such as those based on HL7 RIM (Dolin et al., 2001), are too weak with respect to their semantics.

In developing knowledge oriented systems the key idea is to revolve all applications around the shared ontology. In our case, it means the integration of the PHR-ontology, the Blog-ontology and the Ix-ontology and then revolving the Semantic Blog, Semantic PHR and the semantic Ix around this integrated ontology as illustrated in Figure 1. The integrated ontology is called PHA-ontology. So the components of the PHA interoperate by accessing the shared PHA-ontology.

![Figure 1: The Architecture of a PHA.](image)

User interacts with the PHA by a browser, and so all the documents for user are presented in HTML. However, all the content in the knowledge base are in OWL, i.e., represented by XML-documents.

The required transformation between these representation formats can be automatically done.
This requires that a specific style sheet is specified for the translation for each document type. A language associated with style sheets is XSLT (Extensible Stylesheet Language) (Harold and Scott Means, 2002). It is a markup language that uses template rules to specify how a style sheet processor transforms an XML document.

3 THE COMPONENTS OF THE PHA

We now consider blogs, PHRs and Ix. We first give a short overview of these terms, and then we consider their semantic variations, i.e., how they can be implemented by exploiting a knowledge base.

3.1 Semantic Blogs

Blog represents a technology of Web 2.0, which is a controversial term in that various definitions have given for the term Web 2.0. A commonly used definition is that Web 2.0 refers to a second generation of services available on the Internet that let people collaborate and share information online. It is also regarded as synonymous with the term Internet based social networking. The term social network usually refers to a social structure made of individuals or organizations called "nodes," which are connected by one or more specific types of interdependency, such as friendship.

Blogs provide a way for representing content in social networks. Typically blogs are web pages that contain a series of frequently generated entries by a person or a group. A personal blog is an ongoing diary or commentary by an individual. The entries in a blog (text, links, figures, video or audio files) are presented in chronological order with the latest entry listed first.

Blogs are typically used on specific subjects, which can be specified by tags such as keywords or small phrases. Each item in a blog can be associated with one or more tags. Usually sites which provide tagging functionalities also combine it with sharing capabilities, i.e., allow someone to share his or her blog items with other people.

Many blogs use RSS feeds (i.e., RSS documents) for allowing user subscriptions and thus leverages the creation of blog networks. RSS is most commonly translated as “Really Simple Syndication”. Syndication refers to making web feeds available from a site in order to provide other people with the summary of the website’s recently added content. In general, RSS is a family of web feed formats used to publish frequently updated works such as blog entries in a standardized format. Web feeds includes full or summarized text and metadata such as publishing dates and authorship.

By the term semantic blogs we refer to blogs which data is organized according to an ontology, called blog ontology. The purpose of the blog ontology is to describe the concepts of the domain in which the blog takes place. Hence, blog ontology describes the concepts, as well as their relationships, such as blog-entry, predecessor and subject.

In our adopted approach semantic blogs are maintained by a knowledge management system, which provides data management functionalities such as queries, updates and insertions. Such functionalities can also be used in producing summaries. Therefore in using semantic blogs RSS feeds that are used with traditional blogs are not needed. Neither tagging mechanism is needed as the blog ontology captures the classification of the blog entries, i.e., the relationships between blog entries and classification system (e.g., a taxonomy) can be specified in the blog ontology.

3.2 Semantic PHRs

A PHR is a record of a consumer that includes data gathered from different sources. It includes information about medications, allergies, vaccinations, illnesses, laboratory and other test results, and surgeries and other procedures.

A PHR should provide a complete and accurate summary of the health and medical history of a consumer. It is accessible to the consumer and to those authorized by the consumer. It is not the same as electronic health record (EHR), which is designed for use by health care providers.

PHRs can be classified according to the platform by which they are delivered. In internet-based PHRs health information is stored at a remote server, and so the information can be shared with health care providers. Some PHRs also have the capacity to import data from other information sources such as a hospital laboratory and physician office. However, importing data to PHRs from other sources requires the standardization of PHR-formats.

Various standardization efforts on PHRs have been done. In particular, the use of the Continuity of Care Record (CCR standard) of ASTM (CCR, 2009) and HL7’s Continuity of Care Document (CCD standard) (CCD, 2009) has been proposed. From technology point of view CCR and CCD-standards represent two different XML schemas designed to
store patient clinical summaries. However, both schemas are identical in their scope in the sense that they contain the same data elements.

A problem with XML-based PHRs is that their data is document-centric-data, i.e., they are collections of documents such as documents including lab tests, prescribed medications and illnesses. Instead the effective usage of PHRs is mainly data centric, meaning that data should be extracted from various documents and then integrated according to specific criteria.

For example, a consumer may be interested to know the average blood pressure and/or blood sugar concentration during the time periods he or she was using a drug for blood pressure, or the consumer may be interested to know the cholesterol values when he or she was on a diet. Unfortunately the computation required by such queries is not provided by the query languages such as XPath (XPath, 2008) and XQuery (XQuery, 2008) which are designed to address XML documents.

In order to allow high expression power in accessing PHRs we have developed an ontology for the data stored in PHRs. It describes the concepts of the domain in which PHRs take place. Hence, a PHR-ontology describes the concepts, as well as their relationships, such as demographics, insurance information, immunizations, allergies, diagnoses, procedures and medication. In developing the PHR-ontology, we have exploited the XML-schema of the CCR file, which is originally developed for storing patient clinical summaries.

The PHRs which data is organized according to an ontology we call semantic PHRs. That is, semantic PHRs exploit knowledge management systems, and thus they provide high expression power in accessing the PHR-ontology. The PHA-ontology, which is presented in Section 4, includes the PHR-ontology.

### 3.3 Semantic Information Therapy

Information therapy is a type of healthcare information service that has emerged in the past decade. The goal behind information therapy is to prescribe the right information to right people at right time (Kemper, 2008). Information therapy is also described as “the prescription of specific evidence based medical information to specific patients at just the right time to help them make specific health decisions or behavior changes” (Mettler and Kemper, 2004).

Information therapy applies to a wide range of situations and context. For example, information therapy may be a physician-written prescription telling a patient what to read, or it may use to help a patient to make treatment decision such as whether to continue medication.

Information therapy can be compared to similar concepts in medicine such as drug therapy, physiotherapy or bibliotherapy. However, information therapy differs from these in the sense that by exploiting information technology information therapy aims at providing personalization, targeting and documentation.

There is a variety of paper-based mediums for delivering information therapy such as handing out information pamphlets or sending them through the post. There are also many electronic infrastructures (such as electronic medical record systems, personal digital assistants, order entry systems and personal health records) that have been proposed for delivering information therapy.

By semantic information therapy we refer to the ontology based management of the information entities such that prescribing can be automated by exploiting the IX-ontology. The idea behind this is that the information entities can be modelled in an IX-ontology in the way that their relationship to other relevant health care concepts (e.g., diseases and medication) can be specified.

In our used IX-ontology the class InformationEntity (IE) is further divided into subclasses such as ProductIE, DiseaseIE, ColesteroTestIE, and BloodPressureTestIE. Their relationships to other relevant classes in the PHA-ontology are considered in the next section.

### 4 PHA-ONTOLOGY

Originally ontology is the philosophical study of the nature of being, existence or reality in general, as well as of the basic categories of being and their relations (Antoniou and Harmelen, 2004). Traditionally listed as a part of the major branch of philosophy known as metaphysics, ontology deals with questions concerning what entities exist or can be said to exist, and how such entities can be grouped, related within a hierarchy, and subdivided according to similarities and differences.

In computer science, an ontology is a general vocabulary of a certain domain, and it can be defined as “an explicit specification of a conceptualization” (Gruber, 1993). Essentially the used ontology must be shared and consensual terminology as it is used for information sharing and exchange.
Essentially ontology tries to capture the meaning of a particular subject domain that corresponds to what a human being knows about that domain. It tries to characterize that meaning in terms of concepts and their relationships. It is typically represented as classes, properties, attributes and values. Depending on the generality level of conceptualization, different types of ontologies are needed (Puustjärvi and Puustjärvi, 2009). Each type of ontology has a specific role in information sharing and exchange.

The purpose of our used PHA-ontology is to describe the concepts of the domain in which blogs, PHRs and Ix take place. In order to illustrate its content consider Figure 2, which represents a subset of the PHA-ontology. In this graphical representation ellipses represent data and object properties, rectangles represent data and object representation ellipses represent classes and subclasses, and rectangles represent data and object properties. Classes, subclasses, data properties and object properties are modelling primitives in OWL.

In the graphical ontology:
- the classes Patient and Medication are connected by the object property Uses,
- the classes BlogItem and Medication connected by object property Relates, and
- the classes informationEntity and Medication are connected by the object property Associates.

Note that based on the above classes and object properties it is possible to process the queries presented in Section 1 (Introduction). That is, based on the medication the patient uses, the PHA can automatically deliver to the patient the blog items and information entities that are related to medication the patient uses.

The OWL Web Ontology Language is designed for use by applications that need to process the content of information instead of just presenting information to humans. By an ontology language it is possible to write explicit, formal conceptualizations of domains. So OWL facilitates greater machine interpretability of Web content than that supported by XML, RDF, and RDF Schema by providing additional vocabulary along with a formal semantics.

A part of the graphical ontology of Figure 2 is presented in OWL in Figure 3.

![Figure 2: A PHA-ontology in a graphical form.](image)

![Figure 3: Representing the PHA-ontology in OWL.](image)
statements: the first states that its type is ProductName in the PHA-ontology, and the second states that its BrandName in the PHA-ontology is Diclofenac.

```xml
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#
  xmlns:po=http://www.lut.fi/ontologies/PHA-ontology#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns:do="http://www.doctorontology.com/ontology/#"
  xmlns:med="http://www.medontology.com/ontology/#"
  xmlns:swt="http://www.swtontology.com/ontology/#"
  xmlns:act="http://www.actontology.com/ontology/#">
  <po:Medication rdf:about="211708-8">
    <po:Contains>Voltaren</po:Contains>
    <po:BrandName>Diclofenac</po:BrandName>
    <po:StrengthUnit>Tabs</po:StrengthUnit>
    <po:StrengthValue rdf:datatype="&xsd;integer">30</po:StrengthValue>
  </po:Medication>
</rdf:RDF>
```

Figure 4: Representing ontology instances in RDF.

5 CONCLUSIONS

Internet has changed the way people work, bank and shop, but a similar change in health care has been small-scale. However, recent interest in social networking and the evolution towards patient-centric healthcare is speeding this change. At the same time the use of patient-centric e-health tools is rapidly increasing. These tools cover many fields including electronic health records, personal health records, telemedicine, evidence based medicine, information therapy and disease management.

A problem is that the e-health tools each have their own interfaces. By integrating the e-health tools we can achieve two gains: simplify user interaction and provide new more advanced services. The situation is analogous with many organizations having heterogeneous legacy systems each having own user interfaces. Hence also the solutions developed for the integration and interoperability of organizational and business applications can be adopted for the e-health case as well.

We have designed an e-health tool (PHA) which captures the functions of a personal health record, information therapy and health oriented blog. From enterprise application integration (EAI) point of view it represents information oriented integration approach, and from technology point of view it represents knowledge management technologies such as RDF and OWL.

Moving from XML-archives to semantic PHA requires the introduction of the PHA-ontology, and transforming the XML-based medical and health information in the format that is compliant with the PHA-ontology. This transformation can be executed automatically. However, a specific stylesheet must be specified for each transformed document type.

The PHA-ontology stores the urls of the information entities and blog items, and so the information entities as well as the blog items may be stored in any server. However, the content of the PHRs should be stored in centralized way in the knowledge base as otherwise making expressive queries on the PHRs is not possible.

REFERENCES


XQuery, 2008. XQuery 1.0: An XML Query Language. Available at: http://www.w3.org/TR/xquery/.