Chapter 2.7
An SOA-Based Architecture to Share Medical Data with Privacy Preservation

Mahmoud Barhamgi
Claude Bernard Lyon 1 University, France

Djamal Benslimane
Claude Bernard Lyon 1 University, France

Chirine Ghedira
Claude Bernard Lyon 1 University, France

Brahim Medjahed
University of Michigan - Dearborn, USA

ABSTRACT

Recent years have witnessed a growing interest in using Web services as a reliable means for medical data sharing inside and across healthcare organizations. In such service-based data sharing environments, Web service composition emerged as a viable approach to query data scattered across independent locations. Patient data privacy preservation is an important aspect that must be considered when composing medical Web services. In this paper, the authors show how data privacy can be preserved when composing and executing Web services. Privacy constraints are expressed in the form of RDF queries over a mediated ontology. Query rewriting algorithms are defined to process those queries while preserving users' privacy.

1. INTRODUCTION

Recently, Web services have started to be a popular medium for data publishing and sharing on the Web (Carey, 2007; Gilpin, Yuhanna, Leganza, Heffner, Hoppermann, & Smillie, 2007). Modern enterprises are moving towards a service-oriented architecture for data sharing on the Web by putting their databases behind Web services, thereby providing a well-documented, interoperable method of interacting with their data. We call this type of Web services as *DaaS*

DOI: 10.4018/978-1-61350-323-2.ch2.7

Web services (*Data-as-a-Service Web services*). DaaS services are becoming increasingly popular in the eHealth industry as a viable solution to access and manipulate the Electronic Health Record (EHR). The European Commission, in fulfillment of its action plan (European Commission, 2004) to promote interoperability among European eHealth systems, has supported many projects that address the interoperability problem by adopting the DaaS Web service technology as an interoperability platform among healthcare facilities, medical research centers and health institutions in Europe. One of the most prominent projects is ARTEMIS (Dogac et al., 2006), where DaaS services are used to access and manipulate the different components of the medical records (EHRs) that are held by proprietary data sources/ information systems in healthcare facilities.

While individual DaaS Web services may provide interesting medical information alone, in most cases, users' queries require the invocation of several services. For instance, let us consider the following query: "*what are the tests performed in ABC Lab by patients who have been administered Glucophage in XWZ hospital?*" Let us assume that ABC Lab and XWZ hospital provide two DaaS services S_{ABC} and S_{XYZ}, respectively: S_{ABC} returns the tests performed by a given patient in ABC Lab and S_{XWZ} returns the list of patients that have been administered a given drug in XWZ hospital. The execution of the above mentioned query involves the *composition* of S_{ABC} and S_{XYZ} services. Web service composition is a powerful solution for building value-added services on top of existing ones (Singh, 2001). One can for example, reconstitute the entire healthcare record by compositing the DaaS services which provide its primitive data elements like, *allergies*, *Medications*, *Operations*, etc.

Patient data privacy preservation is one of the most challenging problems in the medical DaaS service Web composition. Privacy is the right of individuals to determine for themselves when, how and to what extent information about them is communicated to others (Westin, 1967). Users are reluctant to use online services for fear that their private data may be disseminated to untrusted parties or used for unintended purposes (LeFevre, 2007).

Data privacy preservation has received a considerable attention in recent years. Earlier work focused on preserving privacy in centralized settings through anonymization (LeFevre, 2007). However, users' private information is often scattered across independent distributed data sources. Applying anonymization techniques on each data source in isolation is not suitable. For example, a patient *Sarah* may have her personal information (national identifier or SSN, age, sex, address, etc) stored at provider *A*'s data source and her lung cancer screenings tests stored at provider *B*'s. If data sources at *A* and *B* were anonymized in isolation, then the anonymized records are not joinable with other datasets. Hence, queries such as "*return the female patients who have developed lung cancer in XWZ city"* (which requires joining A's and B's relations) cannot be answered even if *Sarah* agreed to release her address, sex, and medical tests.

In this paper, we present a privacy preserving DaaS service composition approach that allows healthcare givers (e.g., physicians, nurses, healthcare planers, medical researchers, etc) to answer their queries while preserving the privacy of patients and their medical data. Our proposed approach assumes the existence of a medical ontology to capture consensual and shared medical knowledge. We model each DaaS Web service as *RDF View* over the medical ontology to explicitly represent its semantics. We define a query rewriting approach to answer queries over DaaS Web services while preserving patients' privacy preferences. The idea behind our approach is the following: given a query over the medical ontology and a set of RDF views modeling DaaS Web services, reformulate the query into an expression that refers only to the RDF views and provides the answer to the query while preserving privacy.