Chapter 29
Semantic Reconciliation of Electronic Health Records Using Semantic Web Technologies

Karim El Guemhioui
University of Quebec in Outaouais, Canada

Steven A. Demurjian
University of Connecticut, USA

ABSTRACT

In this paper, the authors present an approach to reconcile the semantics of distinct medical terms found in personal health records (PHRs - that store data controlled by patients) and electronic medical records (EMRs - that store data controlled by providers) that are utilized to describe the same concept in different systems. The authors present a solution for semantic reconciliation based on RDF and related semantic web technologies. As part of the solution, the authors utilize a centralized repository of ontologies to: uniformly interrogate the medical coding systems in which those terms are defined, extract all of their published synonyms, and save the results as RDF triples. The final step in the process is to employ a reasoner to infer non-explicit synonymy among those terms, hence evidencing the underlying semantics to the PHR and EMR systems for possible further processing.

1. INTRODUCTION

In this age where smartphones are ubiquitous, and computer apps geared towards personal health issues proliferate, people are becoming more and more self-conscious of their health and incline to resort to these apps to keep track of their medical condition.[1] As a result, many of them end up assembling in apps health and fitness data from different sources that include several personal health records (PHRs)[2] that maintain data controlled by patients and electronic medical records (EMRs) that maintain data managed by medical providers. Patients utilize the mobile health (mHealth)[3] apps that seem to deal with distinct and unique health issues while in fact they are referring to the same health condition. While a medical

DOI: 10.4018/978-1-7998-1204-3.ch029

provider is easily able to identify related different terms in PHRs and/or EMRs referring to the same condition (i.e., problem), this exercise is less obvious for most individuals (whose medical knowledge is usually superficial), and outright impossible for the machine hosting these electronic records and the computer applications we rely on to make some sense out of this data.

The multiplication and affordability of electronic mobile devices (smartphones, phablets, tablets, "Fitbits", wearable healthcare devices, etc.), and their embracing of the technologies by an increasing number of diverse stakeholders (e.g., patients, family members, home health care aids, therapists, etc.) are paving the way for a shift in the gathering and integration of personal healthcare data from PHRs and EMRs coupled with its readiness for the various healthcare and medical service providers. Therefore, the integration of PHRs and EMRs will play a pivotal role in this move to a safer, centralized, more efficient, consumer-driven healthcare system. Just to give a simple example: mHealth applications that provide access to health and fitness data, gathered from multiple PHRs and EMRs, will allow us to have our medical history available in case of emergency, when traveling, seeking a second opinion, or switching health insurance companies (HealthIt, 2014).

To accomplish this goal, we are starting to see official efforts focused on encouraging provider organizations to give consumers access to their own data; e.g., the Blue button initiative (HealthIT, 2016). More and more sophisticated PHR apps are also starting to support access to and/or import of electronic medical records (EMRs) or electronic health records (EHRs)[4]. A new term has even been coined: a tethered (or connected) PHR to designate a PHR linked to a specific healthcare organization's electronic health record (EHR) system or to a health plan's information system (Daglish & Archer, 2009). With a tethered PHR, patients can access their own records through a secure portal and see, for example, the trend of their lab results over the last year, their immunization history, or due dates for screenings.

While there is no agreed upon standard for a PHR, a significant amount of standardization efforts has been done regarding the annotation, classification, and tagging of health and fitness data in EMRs and EHRs. Among the most widespread standards, at least in North America, are: HL7 v2.x (1989), CDA (2005), LOINC[5], SNOMED[6], UMLS[7], and FHIR (2014). PHRs accessing/downloading these remote electronic data are therefore exposed to their underlying format. This paper focuses on the data in the PHR obtained/constructed from remote EMRs/EHRs sources, and leverages existing standardization efforts to propose an approach to the semantic reconciliation of some critical information appearing in PHRs (and especially tethered PHRs). As for the medical terms used in these EHRs to describe a health condition, a treatment, a medication, etc., we are assuming that they comply with agreed upon terminologies defined in specialized thesauri, controlled vocabularies, reference books, dedicated ontologies, etc., which we will call generically "medical coding systems" in the remainder of this paper. Also, for the sake of clarity, we are limiting ourselves in this work to clinical terms describing a medical condition (i.e., an observed problem).

This paper presents a solution for semantic reconciliation that utilizes a centralized repository of ontologies to identify the relevant terms, extract all published synonyms, and save the results as RDF triples. To facilitate the discussion, we present a scenario that illustrates a rather common issue in PHRs, namely, that the same medical condition is being designated by different medical terms in different EMRs/EHRs, creating therefore a semantic hiatus and preventing appropriate machine processing. This scenario is utilized to demonstrate our approach to semantically reconcile these related terms, whether they have been chosen from the same or from distinct medical coding systems. The proposed solution is built around the Resource Description Framework (RDF) and its related standards RDFS, OWL, SPARQL, etc. (Allemang and Hendler, 2011). Note that the presented approach straightforwardly extends to terms dealing with others issues than medical condition (e.g., medication, treatment, lab results).