e-Health – For Continuity of Care
C. Lovis et al. (Eds.)

doi:10.3233/978-1-61499-432-9-740

A framework for privacy-preserving access to next-generation EHRs

Vassiliki KOUFI[a,1], Flora MALAMATENIOU[a], Aggeliki TSOHOU[b] and George VASSILACOPOULOS[a]
[a] *Department of Digital Systems, University of Piraeus, Greece*
[b] *Department of Computer Science & Information Systems, University of Jyväskylä, Finland*

Abstract. Although personalized medicine is optimizing the discovery, development and application of therapeutic advances, its full impact on patient and population healthcare management has yet to be realized. Electronic health Records (EHRs), integrated with data from other sources, such as social care data, Personal Healthcare Record (PHR) data and genetic information, are envisaged as having a pivotal role in realizing this individualized approach to healthcare. Thus, a new generation of EHRs will emerge which, in addition to supporting healthcare professionals in making well-informed clinical decisions, shows potential for novel discovery of associations between disease and genetic, environmental or process measures. However, a broad range of ethical, legal and technical reasons may hinder the realization of future EHRs due to potential security and privacy breaches. This paper presents a HIPAA-compliant framework that enables privacy-preserving access to next-generation EHRs.

Keywords. Next-generation Electronic Health Records, HIPAA, Privacy, Attribute-based Access Control

Introduction

Personalized medicine has emerged as an essential strategy of modern practice to drive the most effective clinical decisions for optimal patient and population health [1]. A personalized medicine approach to providing health services may involve integrating genomics technologies and advances with clinical and family histories in order to more coherently tailor therapeutics to individual patients [2]. Thus, a key supporting component to effecting personalized medicine is the Electronic Health Record (EHR) providing the non-genetic factors.

In addition to the traditional clinical narrative, EHRs are currently capturing structured data relating to all aspects of care, including diagnosis, medication, laboratory test results and radiological imaging data containing integrated patient data. However, much potential can be realized if a broader set of information were incorporated in EHRs such as: (a) patient information contained in Personal Health Records (PHRs), (b) health information from medical devices connected to patient such as from assistive telecare systems, (c) social care information retrieved on request from social care organizations, (d) health information extracted from various healthcare

[1] Dr V. Koufi, Dept. of Digital Systems, University of Piraeus, 185 34, Greece; Email: vassok@unipi.gr.

systems such as primary and hospital care electronic medical records - EMRs; and (e) genomics information such as genotype and sequence data extracted from biobanks and genetic databanks. This can introduce a new era of personalized medicine where this new generation of EHRs will constitute a computable collection of fine-grained longitudinal phenotypic profiles, facilitating cohort-wide investigations and knowledge discovery on an unprecedented scale, thus leading to more effective prevention, diagnosis and treatment of diseases.

Despite the foreseen potential of next-generation EHRs, their multi-owner and multi-user nature raises severe privacy concerns. A number of approaches for addressing similar concerns have been proposed solely for EHR or PHR systems [3]. To address these issues in next-generation EHRs, as defined above, an innovative framework is required under which data will be pulled together from all the data sources in a privacy preserving manner. This paper presents such a framework which enables EHR data to be processed fairly, for specific purposes (clinical practice and research) according to the privacy policies set by different owners, and, where needed, based on consent of the persons concerned. Thus, clinical practice can be supported (a) by providing healthcare professionals with a readily available rich picture of patient data at the bedside, and (b) by enabling researchers to perform population-wide research in order to obtain EHR-derived knowledge that will bridge the translational gap between bench and bedside and will contribute to the a realization of personalized and stratified medicine. The proposed framework is deployed on a trusted cloud environment and draws upon Attribute-based Access Control (ABAC) paradigm for providing fine-grained access control to users. For research purposes, the proposed framework provides unrestricted access to data hosted in entities (health and social care organizations, etc) covered by Health Insurance Portability and Accountability Act (HIPAA) which are previously de-identified according to the HIPAA Privacy Rule. Hence, the likelihood of breaching patient privacy is reduced.

1. Methods

Figure 1 illustrates a high-level architectural view of the proposed framework which is described by a three-tier model, comprising the data layer, the client layer and a trusted cloud layer.

The client layer is a web portal, whereby medical professionals and other researchers request access to next-generation patient EHRs.

The data layer comprises the remote data resources providing the various chunks of data comprising a next-generation EHR. These resources are heterogeneous and reside at different settings. In particular, for the purposes of this paper, it is assumed that PHR data are stored in a cloud-based PHR system and, while EMR and social data are stored in data repositories of geographically distributed and organizationally disparate health and social care providers, respectively. In addition, information from medical devices attached to patients are uploaded to the PHR through a special-purpose mobile application. Finally, genomic information is assumed to be hosted in several research centers. The various parts of a next-generation EHR are accessible by relevant web services and are owned by different entities according to the specific security policy (e.g. owner of PHR and genomic data is the patient, while owner of the clinical/social data is the health/social provider where the patient has received care). Hence, each part of the EHR may be governed by different institutional or personal

policies and practices with respect to confidentiality, security and release of information. However, as health and social care organizations constitute covered entities under the HIPAA Privacy Rule, they must comply with the Rule's requirements for safeguarding the privacy of protected health information.

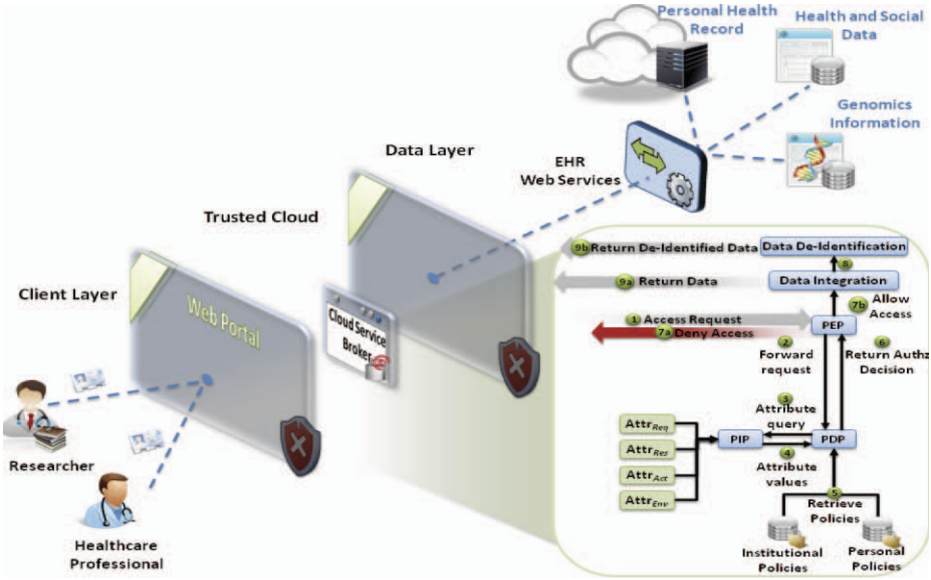

Figure 1. System Architecture

The middle tier is a cloud service broker which serves as a mediation gateway that handles interactions between users and data resource providers. It is deployed on a trusted cloud and provides access to integrated patient EHR data while ensuring compliance to the relevant HIPAA and patient-defined policies by properly authorizing users of the system, i.e. medical doctors and researchers. In doing so, the service broker draws upon the Attribute-based Access Control (ABAC) paradigm. In essence, the Service Broker mediates between requestors (healthcare professionals, researchers) and resources (EHR web services) to decide whether access of a given requestor to given resource should be permitted or denied by taking into account the attributes of the requestor, the resource, the action and the context holding at the time of the attempted access (operational, technical, or situational).

Essentially, access control is enforced through attribute-based access control policies which are defined as 4-tuples (*Req*, *Act*, *Res*, { p-Attr$_{Req}$, p-Attr$_{Res}$, p-Attr$_{Act}$, p-Attr$_{Env}$}) stating that a requestor *Req* is allowed to perform action *Act* to an asset *Res* subject to constraints {p-Attr$_{Req}$, p-Attr$_{Res}$, p-Attr$_{Act}$, p-Attr$_{Env}$} imposed by the attributes of *Req*, *Act*, *Res* and *Env*. The attributes of each entity participating in an access control decision are:

- Requestor attributes: User ID, Department of hospital and/or university where he belongs, user roles, group memberships etc.
- Resource attributes: Uniform Resource Identifier (URI) of the web service to be invoked, specific method of the web service to be executed.
- Action attributes: type of request (read, write, update EHR), purpose of request (clinical practice, research).

- Environment attributes: Time and location of attempted access, type of communication channel (e.g. protocol), client type (e.g. smart phone, etc).

Each request for gaining access to a resource is received by the Policy Enforcement Point (PEP) protecting it. The PEP makes an authorization call to the Policy Decision Point (PDP), which in turn queries for additional subject, resource, action and environment attributes from the appropriate Policy Information Point (PIP). After acquiring the requested attributes, the PDP retrieves the relevant policies from the policy repository and the policies created by the PHR owner, evaluates the request against these policies (function illustrated in Figure 2) and returns a response (and applicable obligations) to the PEP in the form of an authorization decision to grant or deny access. An obligation is information returned with the decision upon which the PEP may or may not act (e.g. an obligation may contain additional information concerning a decision to deny). If access is permitted, the PEP grants the requester access to the resource; otherwise, access is denied.

```
1  ISSUE access_request(Req, Act, Res)
2  BEGIN attribute_collection
3     FOR each Entity in {Req, Act, Res, Env}
4        FOR each Attr_Entity,I in Attr_Entity
5           Acquire attribute value
6           Add attribute to the Entity attribute set
7        END FOR
8     END FOR
9  END attribute_collection
10 BEGIN make AC decision based on attribute vales and user-defined policies
11    PDPdecision = PDP_make_AC_decision(Req request to Act on Res under Env)
12       PHRSysPolicy= retrieve user-defined policies related to Req
13             IF (PDPdecision = 'permit' and PHRSysPolicy = 'permit') THEN
14                   Forward the request to the relevant Web Service
15             ELSE
16                   Deny the request
17             END IF
18 END make AC decision based on attribute vales & user-defined policies
19 FUNCTION PDP_make_AC_decision(Req request to Act on Res under Env)
20    retrieve policies(Req, Act, Res)
21    evaluate policies (Req, Act, Res, Env)
22    return decision
23 END FUNCTION
```

Figure 2. Function for access control policy evaluation

In case that the user is a researcher, s/he can obtain unrestricted access to the sets of data required for his specific research purpose. However, in order to prevent access to unauthorized patient data, EHR data are de-identified in accordance with the HIPAA Privacy Rule once it is retrieved from the various data resources. In particular, the "Safe Harbor" method is used according to which 18 specific identifiers are removed (names, geographic subdivisions smaller than a state, all elements of dates, biometric identifiers etc) [6]. Although genomic data can uniquely identify a person, in the proposed framework, these data are not totally excluded from the data set returned to the researcher. In particular, during the de-identification process, specific parts of these data which can be used for identifying a person are removed, while the remaining information is made available for research. Due to the large volume of data usually required for research purposes, Quality of Service (QoS) needs to be ensured. In the proposed mechanism, this is ensured by detecting and rejecting bad queries, and/or throttling datasets, for example by enforcing limitations on data query size.

2. Results

To illustrate the functionality of the proposed framework, a prototype system was implemented that draws on the ePrescribing healthcare process depicted in Figure 1. Prototype system implementation was based on the Oracle 11g SOA and, in particular, on (i) Integrated Service Environment (ISE) for EHR web services development and (ii) an enterprise portal for healthcare professionals, collaborating healthcare organizations and researchers to access content. The system was deployed on a laboratory cloud computing infrastructure. The platform used for the generation of sample patient PHRs is Tolven ePHR where health information from medical devices (i.e. weight, blood pressure and blood glucose measurements) are uploaded automatically via a relevant Android application [5][6]. Semantic interoperability between all pieces of information comprising a next-generation EHR is achieved by means of ontologies. The ABAC policies have been defined using eXtensible Access Control Markup Language (XACML) [7]. Due to lack of space, a comprehensive description of the prototype implementation will be provided elsewhere.

3. Discussion

Next-generation EHRs comprising clinical, non-clinical and genomic data, hold great promise for improving patient safety as they can support well-informed decision making by providing readily access to richer patient information at the bedside. Moreover, they also constitute a vast repository of data that could be mined in the context of population-wide research aiming at bridging the translational gap between bench and bedside and moving towards a realization of personalized and stratified medicine. This paper presents a framework that sets the ground to realize these goals by facilitating privacy-preserving access to next-generation PHRs. System evaluation is a task to be undertaken in the near future aiming at determining the system usability. Thus, its potential weaknesses may be revealed suggesting alterations in the system design and directions for future work.

References

[1] Waldman SA, Terzic A. Patient-centric clinical pharmacology advances the path to personalized medicine, Biomarkers Med. 2011 Dec;5(6):697–700.
[2] Issa AM. Personalized Medicine and the Practice of Medicine in the 21st Century, Mcgill J Med. 2007 Jan;10(1):53–57.
[3] Heinze O, Birkle M, Köster L, Bergh B. Architecture of a consent management suite and integration into IHE-based regional health information networks. BMC Med Inform Decis Mak. 2011 Oct; 11:58.
[4] Oracle 11g SOA suite [Internet]. [cited 2014 Jan 15]. Available from: http://www.oracle.com/us/products/middleware/soa/suite/overview/index.html
[5] Tolven Healthcare Innovations[Internet]. The electronic Personal Health Record. [cited 2014 Jan 15]. Available from: http://www.tolven.org/ephr.html
[6] Koufi V, Malamateniou F, Vassilacopoulos G. An Android-Enabled Mobile Framework for Ensuring Quality of Life through Patient-Centric Care. Paper presented at MIE; 2012, Aug 26 – 29; Pizza: Italy.
[7] OASIS [Internet]. eXtensible Access Control Markup Language (XACML) [cited 2014 Jan 14]. Available from: https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=xacml