A Model of Privacy and Security for Electronic Health Records

Pulkit Mehndiratta, Shelly Sachdeva, and Sudhanshu Kulshrestha

Jaypee Institute of Information Technology
{pulkit.mehndiratta,shelly.sachdeva,sudhanshu.kulshrestha}@jiit.ac.in

Abstract. Information and communication technology has created excellent development in over the past few years in the field of medicine and healthcare. Healthcare is constantly undergoing changes, with new medical technologies, business models and research findings. The requirements for security and privacy are also very critical and very difficult to satisfy in case of Electronic Health Records (EHRs) data especially as compared to any other data. This is due to the conflicting needs of clinicians (who demand open and easy access to databases) and the patients (who prefer closed and private access to information stored in databases). The potential and capabilities of IT and its influence on the Indian healthcare is of utmost importance. Thus, this study examines the current status of security and privacy of various healthcare services/solutions implemented for electronic health records in India. This topic has not been sufficiently addressed by the existing healthcare solutions based on standards. The authors aim to bridge this gap by proposing a model to protect the security and privacy for Standardized Electronic Health Records EHRs database systems. A simulative analysis for the implementation of the proposed model has been presented. This will help in large scale deployment of secured Electronic Health Record systems that will benefit hospitals and their users.

Keywords: Security and Privacy, Electronic Health Record, Developing Country, India.

1 Introduction

1.1 Electronic Health Records

The medical domain is vast and complex. Electronic Health Records (EHRs) are the paperless solution to a disconnected healthcare world that runs on a chain of paper files. They provides new opportunities, improves productivity, reduces the administrative burdens, reduce cost and medical errors. These become cavillous in the case of an emergency where the patient may be unable to communicate this information. These provide doctors with more timely access to potentially life-saving information at the point of care while diminishing the paper trail.

In general, an EHRs database includes lifelong history of medical documents for any person which includes clinical statements such as observations, laboratory

A. Madaan, S. Kikuchi, and S. Bhalla (Eds.): DNIS 2014, LNCS 8381, pp. 202–213, 2014.

tests, diagnostic imaging reports, treatments, therapies, drugs administered, and allergies. Thus, over a period of time the size of the database becomes very large and very fast leading to various management and security issues.

"As more of our medical records are stored electronically, the threats to our security and privacy increase"[1]. Electronic health records form an integral part of the healthcare system and it is imparitive that EHRs are safe because there is evidence that breaches in security have an impact on patients health care. In a survey conducted in 2006 [11], 62% of the public said "The use of electronic medical records makes it more difficult to ensure patients' privacy". However similar proportions recognized the potential for EHRs in cost and error reductions and increased patient safety. Thus, unless privacy and security problems are resolved, EHRs will not be widely adopted.

1.2 Privacy and Security

The definition of privacy emphasizes the control over the Personally Identifiable Information that should always rest with the data subject. Taking control over this information/data from the subjects takes away his/her privacy. Whereas, security is defined as the extent to which this personally identifiable information can be stored and shared in such a manner that access to the information is limited to authorized parties.

On one hand, best solution to protect the privacy is that data subject should volunteer their own information as they may want to delegate only some (or all) the controls to others. On the other hand, security to the EHRs systems can be provided by the means of physical security of the system, using access control mechanism, or by the use of firewalls and encyrption techniques.
Four major threats (to privacy of data) identified by US National Research Council [23] on medical privacy refer primarily to insider attacks. Many systems like Microsoft Health Vault and Google Health comply with data protection acts by letting the patients decide on the usage and disclosure of their data. But these fail in satisfying essential requirements to privacy and security.

In United States, much of the sensitive data such as insurance information, sensitive patient communications, and personally identifiable information are protected under the U.S. law (the Health Information Technology for Economic and Clinical Health (HITECH) Act). The legislation that regulates release of health-care information is Health Insurance Portability and Accountability Act (HIPAA)[3]. Where on one hand EHRs have to follows strict guidelines of the standards they are based upon, there are legal aspects also when it comes to personal identifiable information. According to Spanish law [12, 13], an article states that *"no personal data will be stored in files that do not meet the requirements of integrity and security"*.

Electronic Health Record data become critical in the case of an emergency where the patient may be unable to communicate this information. These provide doctors with more timely access to potentially life-saving information at the point of care while diminishing the paper trail.

In developing countries like India, the conventional system of medication is still restricted to paper and pen. With a population of over 1 billion, EHRs databases can be a boon to the existing system. EHRs represent lifelong documentation of medical history for any patient and the size of the database increases exponentially. So, an efficient protocol and model is required similar to what is proposed in [9,10]. Thus, it is of utmost importance to provide doctors and patients with modern facilities like computer and mobile based medical solution. This will ease the work of practitioners and make it more effective and productive. At the same time security and privacy of the data has to be maintained in the system. Few of the breaches that occurred in past six to eight months around the globe [4] are due to lack of security and privacy measures and it has affected the lives of patients. ISO/TS 18308 standard gives the definitions of security and privacy issue for EHRs [2].

This paper contributes to the current status of EHRs in India and what are the various security and privacy issues. Section 2 throws light on whether various EHRs implemented in India follow any standards, along with what are the guidelines to be taken into consideration while creating any EHRs solution. The security is required for all the components/layers encountered between the front-end and back-end. Section 3 details the privacy and security measure required before deploying any EHRs database system. A simulation of various techniques has also been discussed in this section providing detailed interaction between various layers on EHRs database systems. Section 4 present the discussion. Section 5 illustrate the conclusions.

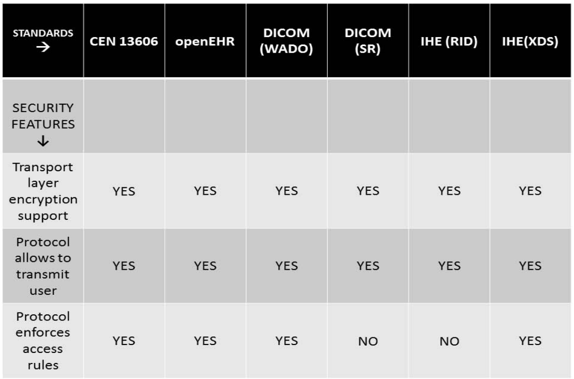
2 Privacy and Security Concerns in Healthcare

The main difference for healthcare data as compared with any other industry is mainly the sensitivity of healthcare data. Patients want their private data to be not disclosed without their consent. Clinicians and researchers want access to patient data in order to come up with concrete solutions in their work. Information is therefore required to be provided in such a way that the personal identity information cannot be disclosed.

Some of the patients records were left exposed to public at University of Michigan Medical center on the internet because the center thought that they were on a server protected by a password [25]. Various other breaches have been reported in [4]. Thus, unless privacy and security problems are resolved, EHRs will not be widely adopted.

There are several standards for EHRs interoperability, such as the Health Level 7 (HL7) Clinical Document Architecture (CDA), CEN EN 13606 EHRcom, openEHR[15], Digital Imaging and Communications in Medicine Structured Reporting (DICOM SR), Web Access to DICOM Persistent Objects (ISO WADO), integrating the Healthcare Enterprise (IHE), Retrieve Information for Display (RID) and IHE Cross-Enterprise Document Sharing (XDS) [14].

ISO/TS 18308 standard gives the definitions of security and privacy issue for EHRs. The Working Group 4 of International Medical Informatics Association (IMIA) was set up to investigate the issues of data protection and security within the healthcare environment. Its work to date has mainly concentrated on security in EHRs networked systems and common security solutions for communicating patient data. The European AIM/SEISMED (Advanced Informatics in Medicine/Secure Environment for Information Systems in MEDicine) project is initiated to address a wide spectrum of security issues within healthcare and provides practical guidelines for secure healthcare establishment. The general technical standard on information security is under study by ISO/SC 27, while the medical-specialized technical standard and the guidelines for using such a standard are under development by ISO/TC 215 WG4, CEN/TC 251 WG3, and HL7 WG13. Figure 1 presents the security features of some existing standards in EHRs. It is prepared by studying enhancement (including two standards) of security features [14].



STANDARDS →	CEN 13606	openEHR	DICOM (WADO)	DICOM (SR)	IHE (RID)	IHE(XDS)
SECURITY FEATURES ↓						
Transport layer encryption support	YES	YES	YES	YES	YES	YES
Protocol allows to transmit user	YES	YES	YES	YES	YES	YES
Protocol enforces access rules	YES	YES	YES	NO	NO	YES

Fig. 1. Security Features of Standardized EHRs

Although, many architectures have been proposed in order to provide better security and privacy to the user but still almost all the systems are lacking in providing the same. The disparity between patients needs and desires for security & privacy and what is provided by some of the electronic health record systems, is illustrated by the results of a study commissioned by HSS [21], which found that the privacy policies of Personal Health Record (PHR) vendors, a type of health controlled by the patients, "lacked the standard components of privacy notices". Thus, one has to come with a system which is patient oriented, so that patient can use and interact with the system in easier manner and at the same time provides security and privacy [22].

2.1 Case Study

Healthcare is moving to the world of computing and the use of new technologies. The ability to quickly and pro-actively react to changes in the healthcare industry will likely become a necessity to stay ahead of the increasing demands of reducing costs, compliance mandates, security concerns and improving the quality of patient care. A standardized format and content of a patient's clinical record helps promote the integration and continuity of care among the various providers of care to the patient. If the data perceived conforms to healthcare standard(openEHR/ HL7), a vendor may conclude that data credibility is high. Keeping this in mind, the research's motive is to explore the privacy and security of standard based EHRs application for its use in India.

EHRs Services in India: In India, very few agencies are working towards area of Health Informatics and Electronic Health Records (EHRs). C-DAC (Center for Development of Advanced Computing) has developed various solutions such as *E-Sushrut [5], DIGHT [6], Mercury, E-Sanjeevni, Tejhas, Ayusoft* etc. Most of these solutions are indigenously developed and managed by C-DAC only. We have done an extensive study of the architecture of all the products and solutions developed and tried to evaluate the security and privacy component in it.

Among these, *E-Sushrut* [5] is the most comprehensive and widely deployed Health Information System. This system incorporates an integrated computerized clinical information system for improved hospital administration and patient health care. HISP India [18] is a training, analysis and evaluation organization creating HIS (Health Information Systems) for South East Asia and is headed by University of Oslo, Norway. HISP has developed few products which cater the EHRs solution requirement. For example, its flagship product DHIS2 which has been recently deployed in a block of Punjab in India. The real time version streamlines the flow of patients and simultaneously empowers workflow to perform to their peak ability, but the security and privacy of the patients data is only limited to the user-level access control mechanism. No attention has been paid to the data encryption and anonymity which could lead to inferences from the available information. The system also lacks in various measures to protect it from network attacks. Thus, very critical and highly confidential information can easily be compromised due to lack of proper measures.

Project DIGHT (Distributed Infrastructure for Global EHRs Technology) [6] proposed to have a separate module for security and privacy which will provide secure storage and access of EHRs, along with privacy to the user. But, till date no such module has been developed/implemented for India to suffice the purpose.

EHRs Standards in India: A report on EHRs standards for India approved by Ministry of Health and Family Welfare (MHFW) has been released in August 2013 [19]. It stated the following factors as shown in (figure 2) of data privacy and security to be considered while creating EHRs solution.

1. **Access Control:** A mechanism for uniquely identifying a user and enabling the user with control for authorized access of EHRs. Although, an exception could always be there for critical circumstances.

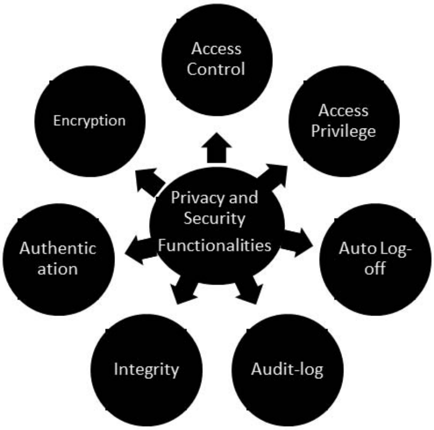


Fig. 2. Privacy and Security Functionalities for Creating EHRs solutions

2. **Access Privilege:** The access to an EHRs of a user should be dependent upon the varying privilege of the care provider(s), as specified by the EHRs administrator.
3. **Auto Log-off:** This is a feature which encapsulates the session management as well as state management. A session time-out should be defined and the states of the record thereof should be saved as intact.
4. **Audit-log:** A log should be maintained which would store each and every action performed over any record. It should include timestamp, user id & the action performed.
5. **Integrity:** The standards should not be compromised in the data transit and also, the deletion of any record as well as audit-log should be prohibited.
6. **Authentication:** User authentication mechanism should be deployed to check the authenticity of the user locally or on a network.
7. **Encryption:** A suitable mechanism should be used for same with substantial key length. Also, in cases of alteration and intrusion, the verification technique should be there e.g. hashing. Also, the data transaction should contain the substantial id information that the receiver can perform security audit trails along with access control decisions.

All the product designed and developed by C-DAC are lacking in security and privacy component. Thus, we have proposed a model to provide security and privacy to the end-user.

3 Proposed Model

There is no present case for India to maintain privacy and security of standardized EHRs database systems. However, MHFW has presented set of factors (please see section 2) to be included as a requirement for secure EHRs systems in the future. Considering various privacy and security concerns and their manifestations, the authors propose a model for EHRs database systems as shown in figure 3. We have tried to included the security at each reference layer of the standardized EHRs. Also, a simulation of various techniques that can be applied to the each layer has been discussed in the sections below.

3.1 Proposal for Secure Model

Our proposal as shown in figure 3,aims to provides security and privacy to the user (data subjects). This shows the function wise reference layer model of the EHRs system. The goal is, how we can include security and privacy techniques on each layer of this reference model of electronic health record database systems to give maximum security as well as state of the art privacy to the data subjects.

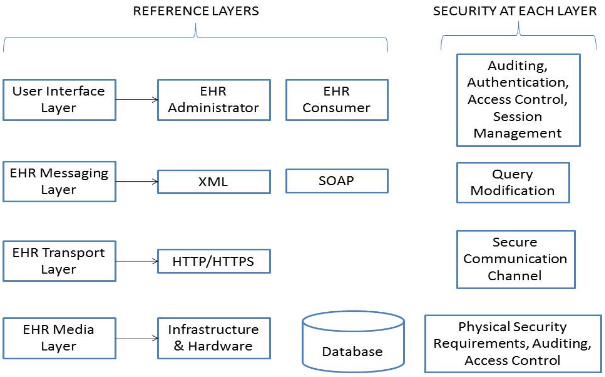


Fig. 3. Reference model for the Standardized Electronic Health Records Database systems with privacy and security measures at each layer

1. **The User Interface Layer:** It is the top most layer where the end user (doctors, nurses and pateints) is interacting with the system.In healthcare domain, the users vary in their background skills and have variable needs. Moreover, they interact with systems in various contexts. Techniques like access control and audit logging can be very helpful to provide adequate security to this layer. For example, if a patient is suffering from fever, or any minor disease, he may share his information, But, a patient will hesitate if he/she is suffering from any sexual or acute disease such as, tuberculosis. The consent of the patient should be taken and they should be familiarize with all the implications once the data is shared. Authentication mechanism also plays an important role when a person has to interact with the system. This is because unwanted entities can be stopped for accessing the system.
2. **The Messaging Layer:** This layer consists of the various messaging formats of the EHRs system which are mainly in XML (Extensible Markup

Language) or use SOAP (Simple Object Access Protocol) for exchanging the structured information through a web service. We need to specify various rules and regulations for the query evaluation purposes. in order to provide security to this later Iif any query is violating the rules or the guidelines then the modification technique should be used to protect the unwanted sharing of data. For example, controlled query evaluation enforces security policies for confidentiality in information systems [27, 28]. Also, the system must provide means where user queries can be modified. This is important in healthcare environment when the user wants to query data from another hospital.

3. **The Transport Layer:** This is most important layer for any EHRs service as all the request are forwarded via this layer itself. Using a secure communication channel i.e. HTTPS can make data-flow more secure as this protocol encrypts the data to be transferred over the web. Further, this protects the service from various attacks like packet sniffing.
4. **The Media Layer:** EHRs databases are complex, vast and temporal in nature. Media layer is holding the most important asset i.e., the actual database of the system. It also comprises of all the hardware and the infrastructure too. Privacy can be increased by storing data as anonymous as possible. Database separation can be done based upon the demographic information and the healthcare data. Access Control and audit logging play a critical role in securing the data from insider attacks. Moreover, the physical security of the database is also necessary.

Above mentioned reference model with security features can be applied to any standardized EHRs solution.

3.2 Simulation of Proposed Model Using Various Techniques

In the last section, we have discussed various security and privacy measures that can be included at each layer of any standardized electronic health record system thus, to make it more secure and usable for end user. This section presents simulation of privacy and security at various layers proposed in reference model through various techniques. Figure 4 shows how the various layers will interact with each other internally once the system is deployed for use. The tag cloud consists of various techniques (numbered from 1 t0 8 as shown in figure 4). It throws light on how the various techniques that are part of the tag cloud, are being used to suffice the goal of providing security and privacy to the EHRs systems. The figure detailswhich technique should be used in which layer for the smooth handling and storage of sensitive healthcare data. It also shows the interaction between different layers through secure communication channel, sending query and its retrieval with secure message being exchanged.

1. **The Demographic and Electronic Health Record Separation:** The database can be further separated into set of two databases. One database will store healthcare information and the other database will be having the

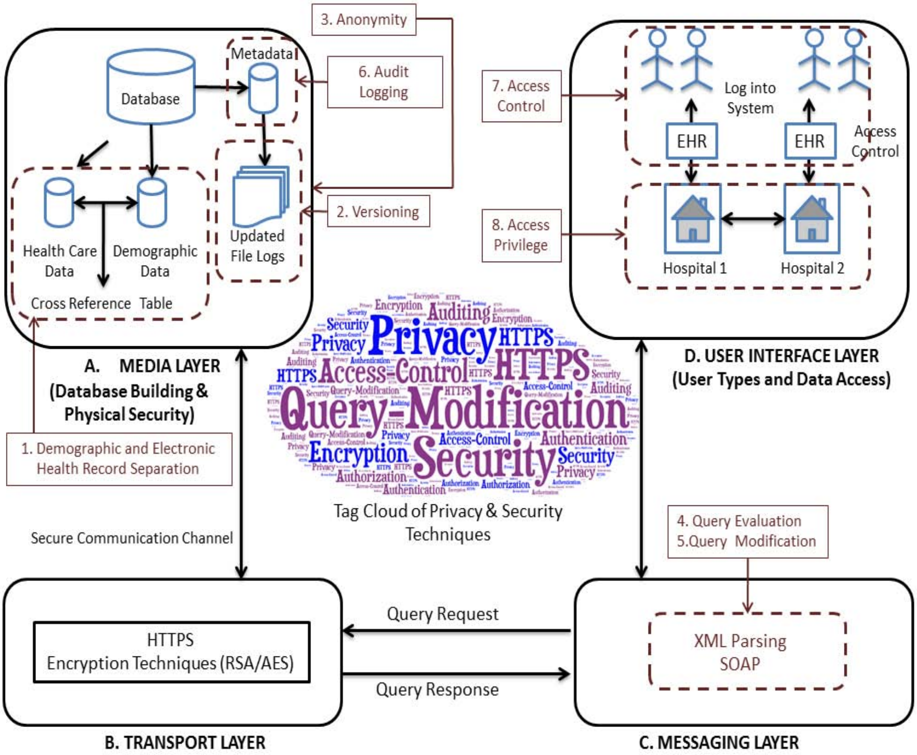


Fig. 4. Simulation for Proposed Model using Various Techniques

demographic details of the patients. A cross-reference table will be maintained to match the record from both the databases.

2. **Versioning:**Along with demographic separation, a meta data log should be maintained that contains the time stamp information when any changes are made to any of the records. At the same time, the new record should be declared as active while the previous record becomes inactive.
3. **Anonymity:** The use of anonymity techniques like k-anonymity and L-diversity will make databases more generalized. Thus, in case of crisis the loss of sensitive information may be minimized. Since, the L-diversity prevents the homogeneity attack and the background knowledge attack posed by k-anonymity, hence choosing L-diversity for sensitive health data would be better option for standardized EHRs systems.
4. **Query Evaluation:** User run queries should be evaluated based on the privacy policy and guidelines to make sure they do not violate it. In case of violation, appropriate measures like evaluation or modification of the query should be done. For example, controlled query evaluation enforces security policies for confidentiality in information systems [17],
5. **Query Modification:** The system must provide means where user queries can be modified [16]. This is important in healthcare environment where

querying medical information may vary depending on the role of end-user and the way information is protected. The e-Diamond project [24] overcomes the query modification issues (multiple views and removing order, primary keys and joins) for medical databases. Measures (taken in discussed method) will reject the query if it does not fulfill the criterion of privacy policy. It is a strict method but considering security concerns, it may suffice for the task. Lets take an example of runtime query enrichment.

Runtime Query Enrichment: The search of sensitive data can be constrained within predefined range of values and thereby it is restrained from direct querying by the end user. The user query is enriched automatically by the system before fetching the sensitive data within the predefined range of values. For example, Blood Pressure (BP) is a sensitive data. Assume that BP level from 90 to 120 is a range of a healthy person and 120 to 140 is a range of a non-healthy person. Therefore, the person falling in the range of BP level 120 to 140 may not like to disclose their BP levels to all other people and thereby, they would like to have a user access control on their BP level readings. In this case, the database column that holds the BP level readings is kept hidden from the users. And the user query is enriched at the runtime to invoke the data of the persons who fall in the range of BP level of 90 to 120. Consider the user query is "select * from person". Then this query is enriched through query modification as "select * from person where BPLevel >= 90 and BPLevel <=120".

6. **Audit Logging:** A log should be maintained that which person with what rights has accessed the system. Along with that, a log should include queries that a user executed on the system. This log combined with new queries can be used to evaluate queries in order to prevent privacy policy violations.
7. **Access Control:** Patients should have control over their own sensitive healthcare data, a technique has been proposed in [10]. Its should be decided by them whether they want to share the information with hospitals for research purposes or not.
8. **Access Privilege**: Hospitals should be having the privilege that once the data has been kept with hospital for any patient they can share among other hospitals for the research and development purposes.

4 Discussions

The literature shows that some solutions in India have mentioned to take security and user privacy into consideration [5, 6]. But, their proposals are not in compliance with the standards proposed by MHFW or any other international act or standardized policy set like HIPAA or HITECH Act. Thus, their is a need for imposing very stringent security policies and procedures. Security issues such as authentication, availability, confidentiality, integrity, access control, data ownership, data protection policies, user profiles and standard model need to be taken into consideration for EHRs. Techniques like *k-anonymity* [7] and *L-diversity* [8] should be used to make data more private and anonymous to disable the inferences from the databases. Incorporating security measures and privacy preserving techniques, organizations will benefit from increased user confidence, convenience, and speed of access to information.

5 Conclusions

We surveyed the problem of security and privacy for various EHRs already implemented and under development in India. In India, there is no specific Care Delivery Organization(CDO) [20] who monitors and follow a particular generic service delivery model. CDOs are legal entities whose primary mission is the delivery of healthcare related products and services. Our findings also implicate that most of the current systems lack for proper security and privacy measures being taken for the system and the user information. It outlines some of the security and privacy measures (figure 2), approved by the MHFW for EHRs solutions in India. But, still no solution has been developed on the basis of those guidelines. However, an insistence on demonstrated factors including implementation become part of future mandatory privacy and security concerns for EHRs systems.

It has been proposed what to look for in a secure EHRs system for complying with the security and privacy concerns at each layer (figure 3). A simulation for the implementation of the proposed model has been presented through various techniques like demographic and electronic heath records separation, query evaluation and modification, audit logging, access control and access privilege.

A very high level of security and privacy is required for the front-end user application and the back-end database. Thus, in future we aim to conduct actual implementation of the proposed model for an EHRs system based on semantic interoperability standard (such as openEHR or HL7).

References

1. State of the Union. Address of William J. Clinton USA (January 19, 1999)
2. ISO/TS 13606, `http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=50121`
3. HIPAA Health Privacy Rule Act, `http://www.hhs.gov/ocr/privacy/`
4. Top 10 Data Security Breaches in 2012, `http://www.healthcarefinancenews.com/news/top-10-data-security-breaches-in-2012`
5. E-Sushrut, `http://www.cdacnoida.in/healthcare.asp`
6. DIGHT: Distributed Infrastructure for Global eHr Technology, `http://dight.sics.se/?q=node/3`
7. Sweeney, L.: k-Anonymity: A model for protecting privacy. International Journal on Uncertainty,Fuzziness and Knowledge Based Systems (2002)
8. Machanavajjhala, A., Gehrke, J., Kifer, D.: L-diversity: Privacy beyond k-anonymity. In: Proceedings of the 22nd International Conference on Data Engineering, Atlanta, GA, USA, April 3-8 (2006)

9. Addas, R., Zhang, N.: Support Access to Distributed EHR's with Three levels of Identity Privacy Preservation. In: Proceedings of Sixth International Conference on Availability, Relaibility and Security, Vienna, Austria, August 22-26 (2011)
10. Huda, M.N., Yamada, S., Sonehara, N.: Privacy-aware access to patient-controlled Personal Health Records in emergency situations. In: Proceedings of Third International Conference on Pervaisve Health, London, UK, April 1-3 (2009)
11. Donelan, K., Miralles, P.D.: supra note 17, at 66 (2006)
12. Law 41/2002 of November 14, basic regulator of the patient autonomy and rights and obligations of clinical information and documentation matters. BOE 274, sec. 1, pp. 40126-40132 (November 14, 2002)
13. Law 15/1999 of December 13, of the Protection of Personal Data BOE 298, sec. 1, pp. 43088-43099 (December 13, 1999)
14. Eichelberg, M., Aden, T., Riesmeier, J., Dogac, A., Laleci, G.: A survey and analysis of Electronic Healthcare Record standards. ACM Comput. Surv. 37(4), 277–315 (2005)
15. The openEHR Foundation, http://www.openehr.org
16. Wong, E., Stonebraker, M.: Access control in a relational data base management system by query modification. ACM SIGMOD (1975)
17. Biskup, J., Bonatti, P.A.: Controlled Query Evaluation for Known Policies by Combining Lying and Refusal. Annals of Mathematics and Artificial Intelligence 40(1-2), 37–62 (2004)
18. Health Information Systems Programmme, http://hispindia.org/
19. Electronic Health Record Standards For India, http://blog.digmed.in/2013/09/22/e-h-r-standards-for-india-goi-report/
20. Adams, J., Bakalar, R., Boroch, M., Knecht, K., Mounib, E.L., Stuart, N.: Healthcare 2015 and Care Delivery", IBM (white paper) (2013), http://www-03.ibm.com/industries/ca/en/healthcare/files/hc2015_full_report_ver2.pdf
21. Personal Health Records Need a Comprehensive and consistent Privacy and Security Framework, CTR. FOR DEMOCRACY AND TECHNOLOGY (June 9, 2009), http://www.cdt.org/policy/personal-health-records-need-comprehensive-and-consistent-privacy-and-security-framework
22. Tejero, A.: Advances and current state of the security and privacy in Electronic Health Records: Survey from a social prospective. Journal of Medical Systems 36, 3019–3027 (2012)
23. For the Record: Protecting Electronic Health Information, Committee on on Maintaining Privacy and Security in Health Care Applications of the National Information Infrastructures, National Research Council (1997)
24. Power, D., Slaymaker, M., Politou, E., Simpson, A.: Protecting sensitive patient data via query modification. In: SAC 2005. ACM (March 2005)
25. Carter, M.: Intergarted electronic health records and patients privacy: possible benefits and real dangers. Medical Journal of Australia 172, 28–30 (2000)