Data Protection Issues of Integrated Electronic Health Records (EHR)

Eirini C. Schiza[1,3], Georgios J. Fakas[2], Constantinos S. Pattichis[1], Nicolai Petkov[3], and Christos N. Schizas[1]

[1] eHealth Lab, Department of Computer Science, University of Cyprus, Cyprus
ischiza@cs.ucy.ac.cy
[2] Department of Computer Science and Engineering, Hong Kong University of Science and Technology, Hong Kong
[3] Intelligent Systems Group, Johann Bernoulli Institute of Mathematics and Computer Science, University of Groningen, The Netherlands

Abstract— **An Electronic Health Record (EHR) of a patient in perspective maintains the medical history of the citizen electronically in medical databanks serviced locally or is cloud based. The ownership and the access control should belong to the citizen and this should be done under the supervision of his personal doctor. Audit trails and security measures must be implemented for making sure that EHR systems properly collect, store, retain and use the patient health information for the better service of the citizen when in need of medical treatment. EU and other countries are determined to find solutions, impose policies and standards as to implement EHR at national level and international levels. In this article the main security issues are presented, the EU directives and legislations in data protection and privacy from the use of EHR are considered, and proposed solutions are analyzed. Finally, it is explained why EHR can and should remain a safe tool.**

Keywords— **eHealth, security, Electronic Health Record (EHR), interoperability, data protection.**

I. Introduction

In the last few years, Electronic Health Record systems have received a great attention in the literature, as well as in the industry and health policy makers. Although they are not widely used yet they are expected to contribute a lot to healthcare savings, increase health care quality and reduce medical errors. The ideal EHR is defined as a collection of continuously updated health-related facts and medical data associated with a patient. The EHR can be a dynamic electronic record that chronologically stores a citizen's medical data from approximately nine months before birth to their death. EHR management systems will enable storage and retrieval of patient data, facilitating physicians to provide safer and effective care through embedded clinical decision support and intelligent diagnostic systems, and can provide useful information through the collection of data for medical research purposes [17]. A significant progress is made in the quest for EHR, which will improve the quality and safety of patient care and achieve real efficiencies in the healthcare system. Many benefits can be attributed to an integrated structured EHR environment such as better management of resources, improved care coordination, chronic disease management, nation and world wide access of medical data, elimination of medical errors and delays, reduced operational cost and patient involvement in their therapy. The approach that Europe has taken over the last few years and more recently by the announcement under the Horizon 2020 ambitious Work Programme of the challenge titled *Health, Demographic Change and Wellbeing*, shows the determination of Europe to find solutions, impose policies and standards, to support the eHealth and patient centered practices mainly through the implementation of national EHR systems [18].

Privacy and confidentiality, personal data, and data protection issues are highly relevant when discussing EHR in its local and pan-European legal and regulatory context. It is thus critical to amend existing health related legislations to create ground for accommodating EHR systems [18].

II. Electronic Health Record

Based on the epSOS (European program, Smart Open Services for European patients) patient summary [20] we formed a structure how the EHR would be and what it should contain.

An EHR must contain general information about the patient, a medical summary consisting of the most important clinical patient data, a list of the current medication including all proscribed medicines that the patient is currently taking and information about when this record was generated and updated by whom.

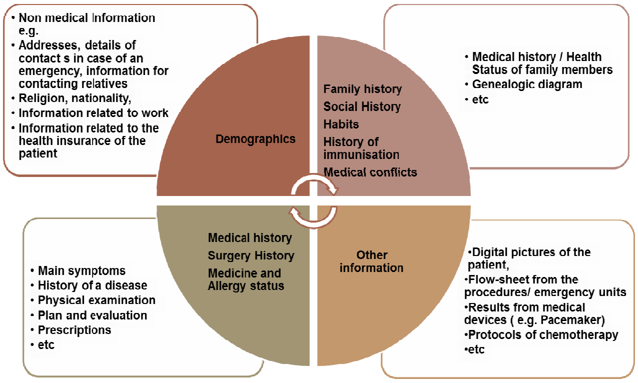
• Non medical Information e.g.
• Addresses, details of contact s in case of an emergency, information for contacting relatives
• Religion, nationality,
• Information related to work
• Information related to the health insurance of the patient
Demographics
Family history
Social History
Habits
History of immunisation
Medical conflicts
• Medical history / Health Status of family members
• Genealogic diagram
• etc
• Main symptoms
• History of a disease
• Physical examination
• Plan and evaluation
• Prescriptions
• etc
Medical history
Surgery History
Medicine and Allergy status
Other information
•Digital pictures of the patient,
•Flow-sheet from the procedures/ emergency units
•Results from medical devices (e.g. Pacemaker)
•Protocols of chemotherapy
•etc

Fig. 1 Electronic Health Record structure

The original version of this chapter was inadvertently published with an incorrect chapter pagination 781–784 and DOI 10.1007/978-3-319-32703-7_152. The page range and the DOI has been re-assigned. The correct page range is 787–790 and the DOI is 10.1007/978-3-319-32703-7_153. The erratum to this chapter is available at DOI: 10.1007/978-3-319-32703-7_260

E. Kyriacou et al. (eds.), *XIV Mediterranean Conference on Medical and Biological Engineering and Computing 2016*,
IFMBE Proceedings 57,
DOI: 10.1007/978-3-319-32703-7_153

As shown in Figure 1, the EHR can be divided into 4 sections. The first section is Demographics, and contains all non-medical information about the patient. The second section is medical, surgery history with medication and allergy status. This section includes all information for each procedure the patient has been or is going through. Another section is the family history which includes also social history, habits, any immunizations and other general information. The last section is for other information such as digital pictures of the patient's situation and results and protocols from any procedures. The epSOS patient summary is the base for building/creating a complete EHR. To use this sort of EHR you need to form a system to support it and each person participating in the implementation, development and use of this system has an ethical obligation to respect the patient's confidentiality and take the necessary actions in order to protect the patient data. What are the factors, however, that can guarantee the data protection in an EHR system, a system that may involve a variety of entities including the hardware, software, people, network, policies and procedures? A model of a security scheme around the patient data can be built by considering a variety of entries; for example, security can be defined according to the implementation architecture, the software at each location, and the organizational policies.

Leak and exposure of sensitive data, such as the patient data, can harm the owner in many ways. This concern has been accompanied by the development of different standards and frameworks to meet EHR challenges. Unauthorized people, for example, may change the original data, or be informed on a sensitive issue. Such an unauthorized access may take place not only in the case of the attacker who uses the vulnerabilities of the system to obtain access to the system, but also by system users who can use their position to be informed on a sensitive matter without having the consent of the owner. Furthermore, manufacturers and product sellers, and pharmaceutical companies can find very useful data in an EHR database, which will make them very competitive in an unlawful way. With the EHR systems, the fact that data may need to be transferred over the network and the physical owners may be more than one, introduces more complex security issues which affect the confident of patients in using such systems. So we need to use the highest levels of data infrastructure, virus prevention, spam filtering, and encryption measures.

The EHR includes some of the most sensitive data of an individual, and therefore deserves the highest degree of protection against all kinds of abuse making the sensitive nature of the processing and thus requiring larger and special handling to protect them.

The EHR, although it is an evolving concept oriented to provide improved healthcare quality and the patient must not deny the right of each patient to ensure the confidentiality and protection of sensitive personal data, but to safeguard. For this reason, it is necessary to establish an appropriate legislative framework so that all the necessary procedures and actions to be taken must comply with the EU legislation.

The benefits of accumulated medical records are self-evident, but we need to spare a thought, too, for the amount of risk they create. Privacy and protection of personal data is of the utmost importance both to the European Union and to each of its member states - and maintaining the privacy of sensitive health data is becoming an ever greater challenge, particularly in the area of cloud computing. Is it possible to preserve privacy at all nowadays? This paper will suggest a practical framework for accomplishing maximum avoidance of data leakage based on the EU directives.

Considering an EHR of a patient where all his medical record and history is maintained electronically, under his ownership and responsibility, and under the supervision of his doctors (e.g. a cloud service provided by a third party). However the case the following three categories of security issues need to be addressed:

(1) Privacy-preserving data publishing **(PPDP*);*** hospitals and governmental services may have to share such data due to research purposes or regulations. Data publishing of personal data must preserve privacy.

(2) Access Control of EHR; the medical record consists of a long medical history of the patient; various specialists may have different privilege accesses on parts of this data. This must be discussed further in the context of the doctor's rights.

(3) Cryptography; Cryptography plays three major roles in the implementation of secure systems: (1) Secrecy and integrity, (2) Authentication, and (3) Digital signatures.

All these facilities are necessary since data will be maintained in third parties and their remote access will be required. In the sequel we discuss in detail these issues [21].

III. Privacy-Preserving Data Publishing (PPDP)

Hospitals, medical centers, etc. may have to share their data either for research purposes or due to regulations [1, 6, 7, 8, and 19]. For example, licensed hospitals in California are required to submit specific demographic data on every patient discharged from their facility [1]. Thus, detailed person-specific data in its original form often contains sensitive information about individuals, and publishing such data violates individual privacy. The best choice to achieve better security is to develop methods and tools for publishing data in a more hostile environment, so that the published data remains practically useful while individual privacy is preserved. This undertaking is called *privacy-preserving data publishing* (PPDP). In the past few years, research communities have responded to this challenge and proposed many approaches [9].

Another example is when a research center requests from a hospital to publish patients records for a research and the external table, e.g. an open election catalogue, is then accessible to the attacker too because these external catalogues are open due to open data access regulations and directives. Then we can assume that every person with hospital record has a record in the open table. In Cyprus, for example, such electoral catalogues are not yet legally open. Note that each of these attributes does not uniquely identify a record owner, but their combination, called the *quasi-identifier* [4, 14] (denoted as *qid* in this report), often singles out a unique or a small number of record owners. For instance, in this case the set of attributes *Job*, *Sex*, and *Age* can form a *qid.* Thus, joining the two tables on the common attributes *Job*, *Sex*, and *Age* may link the identity of a person to his/her *Disease*. For example, *Doug*, a male lawyer who is 38 years old, is identified as an *HIV* patient by *qid* = {*Lawyer, Male,* 38} after the join. To prevent record linkage through *QID*, Samarati and Sweeney [7] proposed the notion of *k-anonymity*: if one record in the table has some value *qid*, at least $k-1$ other records also have the value *qid*. In a *k*-anonymous table, each record is indistinguishable from at least $k-1$ other records with respect to *QID*. Other PPDP approaches include *l*-Diversity [10, 11, 12], *t-Closeness* [10], *Personalized Privacy* [16], *FF-Anonymity* [15], *δ-Presence* [13], *ε-Differential privacy* [5], etc. [10].

IV. Access Control of EHR

The issue raised is whether the various specialists should have the same accesses on the various parts of the EHR. For instance, should the personal doctor (GP) of a patient have full access on someone's EHR and for how long? Can a GP provide access to other doctors? For instance, the GP may provide a complete access to a hematologist but a very limited access to a podiatrist with the consent of the patient. Similarly, a specialist before admitting a patient should have full EHR access and give also, as needed, access to other medical staff under his supervision on certain parts of it. In limited cases a patient may not be permitted to have accesses to his own EHR due to some mental or psychological reason. Such cases should be regulated. It should be emphasized once more that the patient has the complete ownership of his EHR (the practicalities of this are being studied further in the context of our research). There are different implementations of access control policies for determining the access that subjects may have on objects; i.e. ***discretionary***, ***mandatory***, and ***role-based*** [2]. The role-based policy is elaborated further as it is more dynamic and appropriate model for EHR.

Role-Based Access Control (RBAC). In Role-Based Access Control, subjects are assigned to roles that line up with roles that users hold in real life. A role could represent a set of actions and responsibilities that a subject has in their job; for example, the patient, the personal doctor, a specialist, a family, etc. The patient may have full read access to his record but will not have the privilege to modify certain parts of his EHR such as medical details. The GP will be given full read access by the owner, and also full privileges for allowing other colleagues or collaborators for accessing parts of the EHR. In an EHR system, various roles would be created that represent different levels of access. Then, users in the system would be placed in their appropriate role and receive authorization accordingly.

V. Cryptography

Cryptography plays three major roles in the implementation of secure systems. We explain how it can be useful in the case of EHR which need to be stored in a cloud, communicated digitally or need to be accessible remotely [2].

A. *Secrecy and Integrity*

Cryptography is used to maintain the secrecy and integrity of information whenever it is exposed to potential attacks; for example, during the storage on a cloud or during transmission across networks that are vulnerable to eavesdropping and message tampering. It exploits the fact that a document or a message that is encrypted with a particular encryption key can only be decrypted by the owner or the recipient who knows the corresponding decryption key. Encryption also maintains the integrity of the encrypted information, provided that some redundant information such as a checksum is included and checked. For instance, in the EHR context, the EHR is encrypted and the cloud provider cannot decrypt it. Only users having the decryption key can access the record and only once decrypted it makes sense, or better, if it includes some value agreed (such as a checksum of the message) then users can read the record also verify its integrity, i.e. it has not been altered .

B. *Authentication*

Cryptography is used in support of mechanisms for authenticating communication between pairs of security principals (any entity that can be authenticated by the system, such as a user account, a computer account, or a thread or process that runs in the security context of a user or computer account). A principal who decrypts a message successfully using a particular key can assume that the message is authentic if it contains a correct checksum or some other expected value. They can infer that the sender of the message possessed the corresponding encryption key and hence deduce the identity of the sender if the key is known only to two parties. Thus, if keys are held in private, a successful decryption authenticates the decrypted message as coming from a particular sender. For instance, a doctor wishes to access files held by a cloud server. Another third party (e.g. *www.verisign.com*) is an authentication server that is securely managed and issues users with passwords and holds current secret keys for all of the principals in the system it serves (generated by applying some transformation to the user's password). For example, it knows the doctor's and the cloud server's keys.

C. Digital Signatures

Cryptography is used to implement a mechanism known as a *digital signature*. This emulates the role of a conventional signature, verifying to a third party that a message or a document is an unaltered copy of one produced by the signer. Digital signature techniques are based upon an irreversible binding to the message or document of a secret known only to the signer. This can be achieved by encrypting the message – or better, a compressed form of the message called a *digest* – using a key that is known only to the signer. Public-key cryptography is generally used for this: the originator generates a signature with their private key, and the signature can be decrypted by any recipient using the corresponding public key. For instance, a specialist wants to sign a medical report of a patient so that any subsequent recipient can verify that he is the originator of it. When this is done, other medical staff can verify that the specific specialist is the originator of the document when they access it irrespective of the route taken to reach them.

VI. CONCLUSIONS

In this paper, we explained the necessity of EHR and the benefits which an eHealth environment can gain from a properly developed EHR databank. An interoperable, patient centered, and remotely access EHR provides the best way of collecting, storing, retaining and using patient health information. The implementation of EHR is potentially vulnerable to many security challenges; in this paper these concerns were identified and solutions are proposed to be considered during the implementation. Another concern of the paper was to show that in spite of security concerns EHR can be a safe tool for all partners involved.

REFERENCES

1. D. M. Carlisle, M. L. Rodrian, C.L. Diamond. California inpatient data reporting manual, medical information reporting for California (5th Ed), Tech. rep., Office of Statewide Health Planning and Development. 2007.
2. G. Coulouris, J. Dollimore, T. Kindberg, G. Blair. Distributed Systems-Concept and Design (Fifth Edition), Addison-Wesley, May 2011.
3. A. Crowell. A Survey of Access Control Policies, University of Maryland
4. T. Dalenius. Finding a needle in a haystack - or identifying anonymous census record. *J. Official Statistics 2*, 3, 329–336. 1986.
5. C. Dwork. Differential privacy. In *Proceedings of the 33rd International Colloquium on Automata, Languages and Programming (ICALP)*. 1–12. 2006.
6. G. J. Fakas, A novel keyword search paradigm in relational databases: Object summaries, DKE, 2011.
7. G. J. Fakas, B. Cawley, and Z. Cai, Automated Generation of Personal Data Reports from Relational Databases, JIKM, 2011.
8. G. J. Fakas, Z. Cai, N. Mamoulis, Diverse and Proportional Size-*l* Object Summaries for keyword Search, SIGMOD, 2015.
9. B. Fung, K. Wand, R. Chen, P. Yu. Privacy-preserving data publishing: A survey of recent developments. ACM Computing Surveys (CSUR) 42, no. 4 (2010): 14.
10. N. Li, T. Li, M.Venkitasibramaniam. *t*-closeness: Privacy beyond *k*-anonymity and *l*-diversity. In *Proceedings of the 21st IEEE International Conference on Data Engineering (ICDE)*. 2007.
11. A. Machanavajjhala, D Kifer, J Gehrke, M.Venkitasibramaniam. *l*-diversity: Privacy beyond k-anonymity. *ACM Trans. Knowl. Discov. Data 1*, 1. 2007.
12. A. Machanavajjhala, J. Gehrke, D. Kifer, M.Venkitasibramaniam. *l*-diversity: Privacy beyond k-anonymity. In *Proceedings of the 22nd IEEE International Conference on Data Engineering (ICDE)*. 2006.
13. M. E. Nergiz, M. Atzori, C. Clifton. Hiding the presence of individuals from shared databases. In *Proceedings of ACM SIGMOD Conference*. ACM, New York, 665–676. 2007.
14. L. Sweeney. Achieving k-anonymity privacy protection using generalization and suppression. *Int. J. Uncertainty, Fuzziness, Knowl.-Based Syst. 10*, 5, 571–588. 2002.
15. K. Wang, Y. Xu, A. Fu, R Wong. *ff*-anonymity:When quasi-identifiers are missing. In *Proceedings of the 25th IEEE International Conference on Data Engineering (ICDE)*. 2009.
16. X. Xiao, Y. Tao. Personalized privacy preservation. In *Proceedings of the ACM SIGMOD Conference*. ACM, New York. 2006.
17. HealthIT, *What is an electronic health record (EHR)?,* Available at: http://www.healthit.gov/providers-professionals/faqs/what-electronic-health-record-ehr
18. K. C. Neokleous, E. C. Schiza, C. S. Pattichis, C. N. Schizas, *A Patient Centered Electronic Health System: An Example for Cyprus,* ICIMTH 2014 – International Conference on Informatics, Management, and Technology in Healthcare, July 10-13, 2014, Athens, Greece. http://www.icimth.com/icimth-2014 vol.202, pp.111-114, 2014.
19. Data Protection Act, Subject Access Requests, 1998. http://en.wikipedia.org/wiki/Data_Protection_Act_1998.
20. epSOS, *Patient Summary,* Available at: http://www.epsos.eu/epsos-services/patient-summary.html
21. C. N. Schizas, eHealth week 2015, *eHealth for the Citizen for Better Privacy and Data Protection,* Available at: https://www.eiseverywhere.com/file_uploads/c8815f88eff6764218aeeeb4e06cd22a_SCHIZAS_eHealthfortheCitizenforBetterPrivacyandDataProtection.pdf