

IT infrastructure for National Electronic Health Records in Luxembourg – Acceptance occurs when benefits outweigh disadvantages

S. Benzschawel, H. Zimmermann, M. Da Silveira,
U. Roth, A. Jahnen

Public Research Center Henri Tudor

(stefan.benzschawel, heiko.zimmermann, marcos.dasilveira, uwe.roth,
andreas.jahnen)@tudor.lu

29, avenue John F. Kennedy, Luxembourg, 1855, LU

Abstract: Electronic Health Records (EHR) systems manage the most intimate and private information. The acceptance of EHR systems is proportional to their positive balance of benefits weighted against the risk of insufficient data protection. An objective, better treatment process for patients on the one side and a highly secured system on the other side are the most important preconditions. The aim of this paper is to present the requirements to design architectures to national eHealth platforms and to describe the architecture proposed by SAN-TEC team.

Introduction

Compilations of medical records are the basis for treatment decision support systems. They support practitioners to diagnosis diseases and select therapies or medical pathways by taking into account the patient's current symptoms and his complete medical history. With the advances of the Internet, medical data can be more easily shared between healthcare professionals. But this context also has some critical risks and the confidentiality and reliability of medical data need to be assured. The contribution of this work is on the data protection level within national EHR architectures. In this paper, we present an architectural solution to manage 3 kinds of data:

- (a) Small and structured documents, where personal data can be separated from medical data. Example: CDA laboratory results.
- (b) Small documents where personal and medical data cannot be separated. Example: PDF discharge letters with patient identity.
- (c) Large documents where personal and medical data cannot be separated. Example: DICOM images with patient identity.

System Architecture & Security

The proposed system deals with these kinds of data: (a) Laboratory results which are pseudo-anonymized and stored in a central medical data repository. (b) PDF documents are encrypted and stored in a centrally encrypted DMS (Data Management System), (c) X-ray images are stored in encrypted decentralized data repositories (for example on servers in the hospital's DMZ (demilitarized zone)). Also national PACS backup systems are good alternatives and may act in a second role as decentralized repositories.

To avoid attacks from external or from internal users the system splits person-identifying data and medical-result data into different systems, operated by different institutions. One system (called TTP – trusted third party) keeps the person-identifying data and maps it to pseudonyms. A second system (CMReg – central medical registry) keeps pseudonyms together with their corresponding medical data (see Fig. 1).

Dealing with pseudo-anonymized data

Primary systems (that produces medical data) are charged to split the information into two messages: one with the person-identifying data; and one with the medical data. The messages are sent to TTP and CMReg, respectively. Both messages have the same link-ID which will be used by CMReg to request to TTP a pseudonym. CMReg will never know the patient's identity and TTP will never know where the medical data is stored.

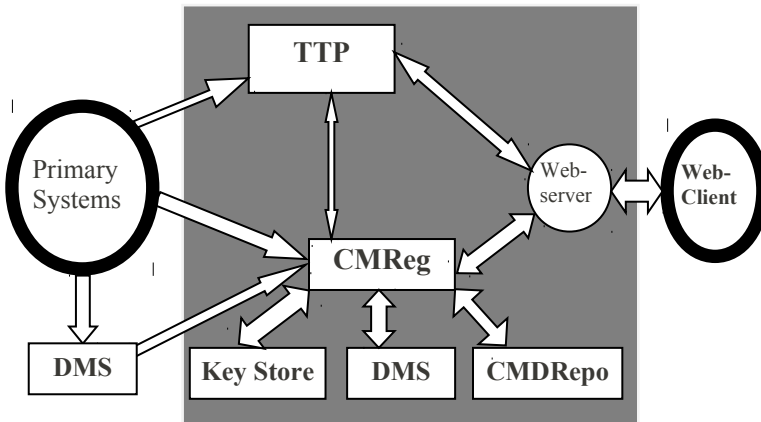


Fig. 1: Simplified Architecture representation

The medical data document is stored in the CMDRepo (central medical data repository) and the CMReg keeps meta-data (location, pseudonym, etc.) about the document.

When a Web-client queries data from the system, the query is splitted into two messages. The TTP receives person-identifying data (of the patient) and a token; the CMReg receives the query and the same token. The CMReg asks to TTP for the pseudonym(s) corresponding to the token. The TTP replies to the CMReg with a list of pseudonyms that match to person-identifying data sent by the Web-server (with the same token). These pseudonyms are used to perform the query in the medical data storages. The query results are encrypted with the public key of the client and sent back.

Dealing with non-anonymized data

As long as the system can assure that medical data does not contain person-identifying data, the pseudo-anonymization works well. For documents that pseudo-anonymization cannot be guaranteed, a new procedure must be implemented.

For small documents like patient discharge letters in CDA or their PDF versions this problem can be solved by: (a) charging the primary system to generate a symmetric key; (b) encrypt the medical data with the symmetric key; (c) encrypt the symmetric key asymmetrically with the public-key of the *KeyStore* server; (d) send the encrypted key and the encrypted medical data to CMReg; (e) CMReg sends the medical data to DMS and the encrypted key to KeyStore; (f) CMReg creates meta data (locations, pseudonyms, etc.) about the documents.

For large and unstructured documents like x-Ray images, where a separation of data is impossible, the transmission time can become a bottleneck to the system. In that case, data may reside on the producers' side instead of getting copied to the central encrypted DMS. Therefore we considered to have Encrypted Decentralized Data Repositories. The procedure to register these documents in the EHR system is quite similar to small documents (e.g., PDF files): (a) the documents are symmetrically encrypted (as explained before), the key is sent to the *KeyStore* server and the medical data location is sent to CMReg (to complete the meta data). The only difference is that documents are stored on the producer side.

The retrieving process is the same for small or huge documents. When the document is requested by a client, the symmetric key is decrypted by the

KeyStore server and re-encrypted asymmetrically with the public key of the client. Both, document and encrypted key, are sent to the client.

Statistic evaluations are possible only on non encrypted medical results which are stored in the CMDRepo. If further statistical research has to be done and personalized data like age and sex are necessary, special exceptional queries must be allowed to query the TTP.

Conclusion

This paper describes the architectural approach of SANTEC to improve confidentiality and reliability of medical data within a national eHealth platform. The proposed architecture was designed to satisfy the following requirements:

- Medical data must be separated from person-identifying data.
- The same organization must not have complete access to both data stores. A TTP is charged to map real identities and pseudonyms while a CMReg is charged to map pseudonyms and medical data.
- Data can be sent as structured or unstructured file formats.
- If possible, medical data must be pseudo-anonymized.
- Non pseudo-anonymized data must be encrypted.
- Data must be accessible via secured connections over the Internet.
- Encrypted data and the encryption key cannot be available to the same person or organization (e.g. server manager).
- The results of a query can only be readable by the requester.
- Users must be identified.
- The access' rights of users must be predefined. The services are available according to the users' rights.
- Pseudonyms of patients are never sent outside of the platform. They are not visible for clients and primary systems.
- Statistic evaluations are possible only with non encrypted medical data.
- The system must foresee a solution to verify patients' consents. Consents are associated to pseudonyms.

Acknowledgements

The authors thank their partners from the Health Ministry of Luxembourg for their very helpful advises and for providing insights into organizational and legal aspects of the eHealth platform.