Integrating Electronic Health Records in Clinical Decision Support Systems

Eider Sanchez, Carlos Toro and Manuel Graña

Abstract Electronic Health Records (EHR) are systematic collections of digital health information about individual patients or populations. They provide readily access to the complete medical history of the patient, which is useful for decision-making activities. In this paper we focus on a secondary benefit of EHR: the reuse of the implicit knowledge embedded in it to improve the knowledge on the mechanisms of a disease and/or the effectiveness of the treatments. In fact, all such patient data registries stored in EHR reflect implicitly different clinical decisions made by the clinical professionals that participated in the assistance of patients (e.g. criteria followed during decision making, patient parameters taken into account, effect of the treatments prescribed). This work proposes a methodology that allows the management of EHR not only as data containers and information repositories, but also as clinical knowledge repositories. Moreover, we propose an architecture for the extraction of the knowledge from EHR. Such knowledge can be fed into a Clinical Decision Support System (CDSS), in a way that could render benefits for the development of innovations from clinicians, health managers and medical researchers.

Keywords Electronic health record · Clinical decision support system · Knowledge extraction · Semantic model

---

E. Sanchez (✉) · C. Toro
Vicomtech-IK4 Research Centre, Mikeletegi Pasealekua 57, 20009 San Sebastian, Spain
e-mail: esanchez@vicomtech.org

E. Sanchez
Biodonostia Health Research Institute, P. Doctor Begiristain S/N, 20014 San Sebastian, Spain

E. Sanchez · M. Graña
University of the Basque Country UPV/EHU, Computational Intelligence Group, Computer Science Faculty, P. Manuel Lardizabal 1, 20018 San Sebastian, Spain

M. Graña
ENGINE Centre, Wrocław University of Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland

Y.-W. Chen et al. (eds.), *Innovation in Medicine and Healthcare 2015*,
Smart Innovation, Systems and Technologies 45,
DOI 10.1007/978-3-319-23024-5_37

1 Introduction

Electronic Health Records (EHR) are systematic collections of digital health in-formation about individual patients or populations. They include demographical data, medical history, medication and allergies records, immunization status, laboratory test results, radiological images, vital signs, personal statistics, such as age and weight, and even billing information. The main objective of EHR is to make patient clinical information as well as all their medical history available for future clinicians that will be treating this patient. In such way, the diagnosis could be improved and the prescribed treatments could be better suited to each patient.

During the last years the medical community has identified EHR as valuable as-sets and hard work has been done in order to improve and integrate them in clinical and hospital environments [1, 2–4]. In particular, interoperability has been broadly addressed, so that EHR of a patient visiting different Health Systems can travel safely with the patient integrating the information along the way. Different EHR standards have been defined, such as CEN/ISO 13606, HL7 (RIM, CDA) and OpenEHR [1, 2, 4, 5]. Such standards determine the structural characteristics of EHR, as well as the ones needed for communication purposes with other EHR. Results obtained at research and technological levels have been successful, so that interoperability of EHR management systems is a reality nowadays, even though they are still not implemented in most clinical or hospital environments.

Nevertheless, such vision does not take advantage of the complete potential offered by EHR [6]. In particular, all patient data registries stored in EHR implicitly reflect different clinical decisions made by the clinical professionals that participated in the assistance of patients. More in detail:

- Which patient parameters have been used in each decision (e.g. which medical tests have been performed, which treatments were prescribed, which interventions carried out, etc.).
- Which criteria have been followed during such decisions (whether such criteria follows clinical guidelines and protocols).
- Which has been the result of the decisions made on the patient (e.g. the effect of the prescribed treatments, the success versus failure of such a treatment).

Thus, the exploitation and reuse of the implicit knowledge in EHR could also be used to improve the knowledge on the mechanisms of a disease or the effectiveness of the treatments. At a clinical level such necessity is clear: nowadays the mechanisms of quite a lot of relevant diseases as well as the treatments applied to fight them are still unknown. For instance, oncology and neurology could be two domains in which such knowledge is still needed.

This paper proposes a methodology that allows the management of EHR not only as data containers and information repositories, but also as clinical knowledge repositories. The main objectives are: (i) to improve and enlarge the knowledge of the mechanisms of a disease, (ii) to evaluate the effectiveness of the applied therapies and interventional procedures, (iii) to evaluate whether the clinical practice developed follows clinical guidelines and protocols, (iv) to identify similar patients for facilitating enrolment of patients in clinical studies in general, (v) to identify groups of similar patients for the stratification of the patient population, (vi) to measure the quality of the clinical practice developed by a clinical team, and (vii) to generate preliminary evidence of a certain hypothesis (e.g. a certain treatment is not effective or valid for patients of a certain type) that could lead to start clinical studies.

In this paper we propose an architecture for the extraction of the knowledge from EHR. Additionally, we also propose feeding such knowledge into a Clinical Decision Support System (CDSS), in order to handle the knowledge in such a way that could be useful for clinicians, health managers and medical researchers alike.

This paper is structured as follows: Sect. 2 introduces EHR and CDSS. Section 3 proposes the methodology for the semantic modelling of EHR and the knowledge extraction. Section 4 proposed a methodology for reusing such knowledge to provide clinical decision support. Finally, Sect. 5 discusses some relevant aspects of our approach.

2 Related Concepts

EHR systems, as well as CDSS are introduced in this Section, in order to provide the relevant concepts regarding the aspects covered in this paper.

2.1 Electronic Health Record

According to the definition provided by the Institute of Medicine (IOM) Electronic Health Records (EHR) are defined as "a longitudinal collection of electronic information about the health status of patients, introduced or generated by members of the medical team in a health organization or hospital". The main characteristics of EHR are the following: (i) persistence of the information during every step of the clinical assistance, (ii) unambiguous patient identification, by means of a universal identifier for each patient, (iii) interoperability with other systems, (iv) standardization of information storage, (v) representation of the contents in an understandable manner by other healthcare professionals, (vi) usability, easy to use by all healthcare professionals, (vii) legal value of every document contained, signed by the corresponding responsible, and (viii) security and privacy of the data.

EHR are intended to provide a benefit in different domains, such as (i) healthcare: EHR stores all patient data; (ii) teaching: the information contained is useful for the learning of clinical cases; (iii) research: patient data can be reused for performing clinical studies; (iv) management: the costs of clinical procedures, patient billing, and clinical and economic indicators can be calculated based on patient data, and (v) legal: the assistance provided to a patient can be certified and the legal responsible in cases of failure or are identified.

Different EHR standards have been developed. Health Level 7 (HL7), for in-stance, is a Standards Developing Organization (SDO) oriented to heath information and interoperability. HL7 develops different standards, covering areas such as messaging and data interchange, rules, syntax, visual integration, context, clinical document architecture, functional model, and labelling. For EHR, the HL7 Version 3 is currently the most applied. It is based on a Reference Information Model (RIM) [7] and covers specially messaging aspects. Clinical Document Architecture (CDA) complements the RIM, focusing on the data structure of the EHR.

OpenEHR is an open standard detailed for the development of a complete and interoperable computational platform for clinical information [7]. The technical specifications of information, service and clinical models are detailed. A dual model is implemented, formed by a Reference Model (information) and an Archetype Model (formal definition of clinical concepts). An archetype is formed by three different parts: (i) descriptive information (identifier, code of the clinical concept described and metadata); (ii) restriction rules (regarding cardinality, structure and content), and (iii) ontological definitions (vocabulary).

ISO EN 13606 is formed by 5 parts: (i) Reference Model: a general model of information to communicate with the clinical history of a patient; (ii) Specification of the archetype model: a generic model of information and a representation language for the individual instances of archetypes (archetype description language, ADL [8]; (iii) Reference archetypes and term lists: a normative section with the list of codified terms to be used in the attributes of the reference model and an informative section in which examples archetypes are presented to show how to map using ISO 13606 structures the clinical information codified in HL7 v3 or OpenEHR, (iv) Security characteristics that individual instances must comply, and (v) Inter-change model: contains a set of models for communication purposes based messages and services. ISO 13606 has adopted the Dual Model of OpenEHR for the representation of every health data introduced in the EHR, which has 2 parts. (i) The first one is the reference model: a generic class model that represents the generic properties of the information in the EHR [9], following the different perspectives of the company, the information, computational, engineering and technology. The second is the archetype model: the definition of the clinical content of the data. Archetypes are metadata used to represent the specific characteristics of clinical data. They are a formal definition of a clinical concept based on the reference model.

The relationship between HL7, OpenEHR and ISO EN 13606, reported by Schloeffel [7], is depicted in Fig. 1.

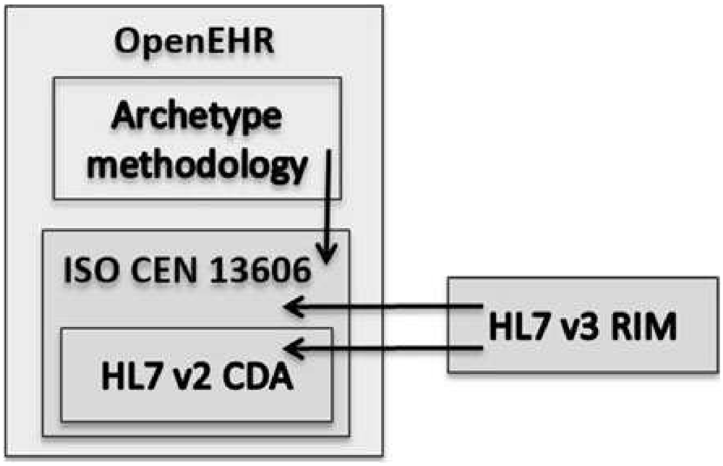

Fig. 1 Relationship of standards HL7, OpenEHR and ISO EN 13606 by Schloeffel (Schloeffel, 2006)

2.2 Clinical Decision Support Systems

We adhere to the definition of CDSS given in [10] stating that CDSS are active intelligent systems that use patient clinical data to generate case specific advice. According to [11], the main task of CDSS consists of the retrieval of relevant knowledge and patient data (coming from medical devices, evidence provided by the medical community, and clinical guidelines and protocols) and their analysis to perform some action, often the generation of recommendations. The target user can be a physician or any other medical professional, a medical organization, a patient or patient's caregivers or relatives. The goals of CDSS are: (i) to facilitate assessment of patient data, (ii) to foster optimal decision making, problem solving and acting, in different contexts and tasks (such as diagnosis and treatment), ensuring that decision makers have all the necessary knowledge to make a correct decision, and (iii) to reduce medical errors [12, 13]. A wide variety of tools can be included in CDSS, some examples are: (i) computerized alerts and reminders, (ii) clinical guidelines, (iii) order sets, (iv) patient data reports and dashboards, (v) documentation templates, (vi) diagnostic support, and (vii) clinical workflow tools [14]. The technologies in which such tools and interventions are based are sparse (e.g. data mining techniques, communication protocols, knowledge acquisition techniques, semantic representation and reasoning, etc.).

We will focus on Knowledge-based CDSS, which benefit from a symbolic representation of knowledge about a particular domain, and the ability for reasoning about solutions of problems within that domain, [15]. The general model of Knowledge-based CDSS proposed by Berner et al. [12] consist of 4 elements: (i) an input, (ii) an output, (iii) a Knowledge Base and (iv) a reasoning engine.

Knowledge-based CDSS in our approach are focused in two main functionalities: generation of recommendations [16, 17] and the management of the underlying knowledge that will drive such recommendations [18].

3 Proposed Methodology for Knowledge Extraction from EHR

We propose a methodology for the extraction of the implicit knowledge in EHR. Two different layers are proposed: an integration layer and a semantization layer. Figure 2 depicts the proposed architecture.

3.1 EHR Integration Layer

This layer is intended to provide the corresponding modules capable of integrating the information contained in different EHR (i.e. EHR of different types, structures, standards, origins). Each EHR will develop a different integration module.

The main idea underlying this approach is the requirement of no information loss from the current existing EHR (i.e. interoperable EHR) to the proposed semantic EHR. In our approach we intend to build novel EHR that could be easily integrated to existing systems currently running in hospitals and clinical organizations.

Taking into account legal issues of data protection and privacy, the system will be provided of an anonimization layer that will ensure that all data processed is stripped of personal data. This process is followed by data extraction and clustering of data fields. Clustering processes will be qualitative due to the heterogeneous nature of the data in the HER, conventional quantitative clustering do not deal well with categorical or descriptive data. To this end, a tool to be considered is the

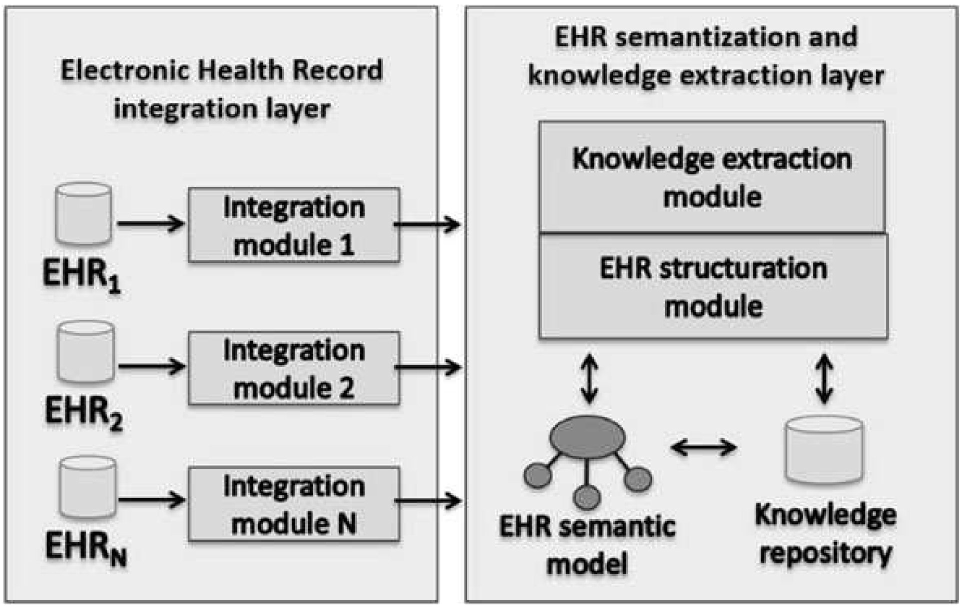

Fig. 2 Architecture for the knowledge extraction from EHR

Formal Concept Analysis [19], based on a lattice approach to the discovery of semantic classes and ontology creation. This approach only requires the existence of some kind of ordering, and is able to work on very heterogenous data. Also the recent advances include some kinds of adaptive processes which allow the semantic ontology to evolve in time with the new data.

3.2 EHR Semantization and Knowledge Extraction Layer

The main objective of this layer is to extract the knowledge from the EHR and store it in a knowledge repository in which it is represented in a way that will allow future reuse and reasoning over it.

3.2.1 EHR Structuration Module

We propose a new semantic model for the representation of the contents of the EHR. Our model extends the Dual Model (of ISO EN 13606 and OpenEHR) with a triple approach in which not only the patient, the disease and the performed medical tests are represented (dual approach), but also a decision model is represented. The decision model contains both, the decision made and the context of such decision: (i) patient data (socio-demographic data, data from anamnesis and data from the medical tests performed to the patient), (ii) decision criteria considered during decision-making, (iii) the objective of the decision (e.g. fast recovery, survival without surgery, avoid blood transfusion), and (iv) the result of the decision (level of success achieved with regard to the objective searched).

The works reported in [18, 17] are aimed to mine the experience present in the history of interactions between doctors and patients. This approach is greatly benefitted by the formalization of the EHR, so that actual experiential events (decisions) can be readily extracted and traced by the entries in the EHR. Therefore, we propose the usage of Decisional DNA and SOEKS [20] to model decisions contained in the EHR.

3.2.2 Knowledge Extraction Module

The mining of the decision model will generate new knowledge in the system, such as the assessment of the effectiveness of the treatments prescribed. We propose the development of different algorithms for the knowledge extraction from EHR. The development of natural language processing algorithms will be a key success factor for such mining. Some approaches have been previously developed, specially regarding codification of clinical terms into different terminologies (i.e. SNOMED CT, CIE-9, CIE-10, UMLS, etc.). However, error rates are still high.

4 Integration of EHR into CDSS

We intend to benefit from the architecture proposed in [18, 17] for Semantically enhanced CDSS (S-CDSS), based on:

- The user layer
- Data, knowledge and experience repositories.
- A multiagent architecture consisting of 9 distinct agents: majordomo, data handling, data translation, knowledge and decision, experience handling, reasoning, application, user characterization, and standards and interoperability.

A new agent in the architecture, the EHR integration agent, will implement the architecture proposed in the previous section (Fig. 3).

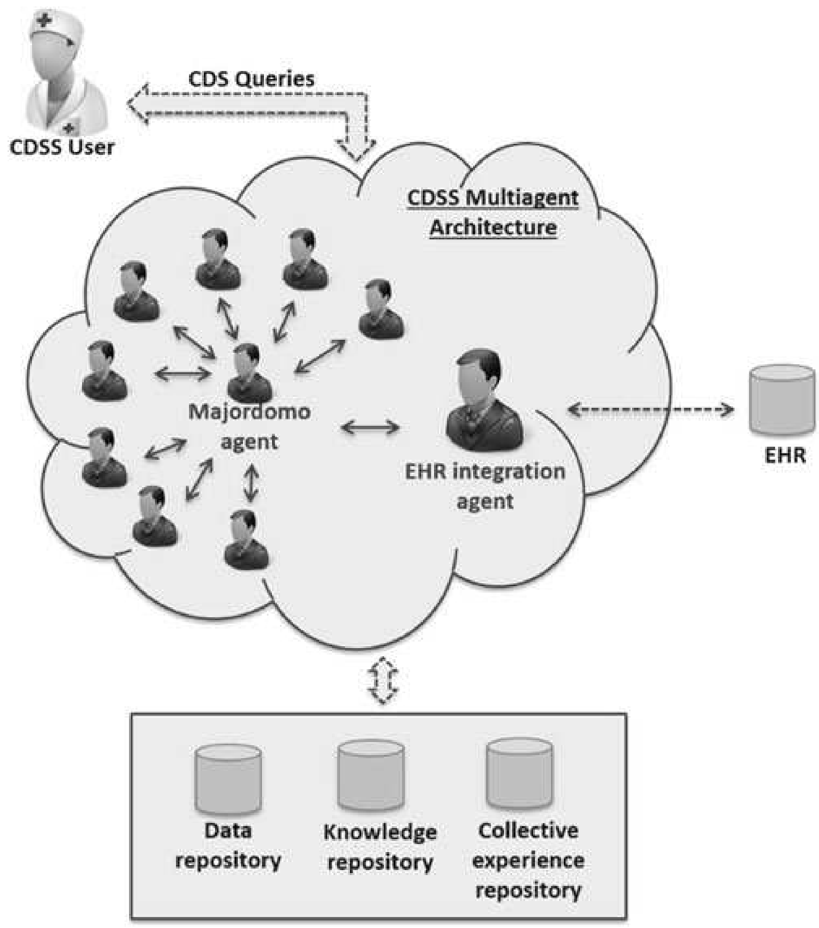

Fig. 3 Integration of EHR into CDSS

5 Conclusions

In this paper we have proposed an architecture for the semantization of EHR, that allows knowledge extraction and reuse from EHR. Additionally, we have also presented a methodology for integrating such Semantic EHR into CDSS. We are actually planning the validation of this approach on real data provided by private companies, or made available for research at the Internet, such as the ones provided by OpenMRS (http://openmrs.org/).

Acknowledgements M Graña was supported by EC under FP7, Coordination and Support Action, Grant Agreement Number 316097, ENGINE European Research Centre of Network Intelligence for Innovation Enhancement

References

1. Sumita, Y., Takata, M., Ishitsuka, K., Tominaga, Y., Ohe, K.: Building a reference functional model for EHR systems. Int. J. Med. Inform. **76**(9), 688–700 (2007)
2. Blobel, B.: Advanced and secure architectural EHR approaches. Int. J. Med. Inform. **75**(3–4), 185–190 (2006)
3. Martínez Costa, C., Menárguez-Tortosa, M., Fernández-Breis, J.T.: Clinical data interoperability based on archetype transformation. J. Biomed. Inform. **44**(5), 869–880 (2011)
4. Martínez-Costa, C., Menárguez-Tortosa, M., Fernández-Breis, J.T.: An approach for the se-mantic interoperability of ISO EN 13606 and OpenEHR archetypes. J. Biomed. Inform. **43**(5), 736–746 (2010)
5. Garde, S., Hovenga, E., Buck, J., Knaup, P.: Expressing clinical data sets with openEHR archetypes: a solid basis for ubiquitous computing. Int. J. Med. Inform. **76**(3), S334–S341 (2007)
6. Bakker, A.: Access to EHR and access control at a moment in the past: a discussion of the need and an exploration of the consequences. Int. J. Med. Inform. **73**(3), 267–270 (2004)
7. Schloeffel, P., Beale, T., Hayworth, G., Heard, S., Leslie, H.: The relationship between CEN 13606, HL7, and openEHR. In: Proceedings of HIC 2006 and HINZ 2006, p. 24 (2006)
8. Eichelberg, M., Aden, T., Riesmeier, J., Dogac, A., Laleci, G.B.: A survey and analysis of electronic healthcare record standards. ACM Comput. Surv. **37**(4), 277–315 (2005)
9. Kalra, D., Ingram, D.: Electronic health records. In: Zieliaski, K., Duplaga, M., Ingram, D., (eds.) Information Technology Solutions for Healthcare, Health Informatics, pp 135–181. Springer, London (2006)
10. Liu, J., Wyatt, J.C., Altman, D.G.: Decision tools in health care: focus on the problem, not the solution. BMC Med. Inform. Decis. Mak. **6**(4) (2006)
11. Greenes, R.A: Clinical Decision Support: The Road Ahead. Elsevier Science, Boston (2011)
12. Berner, E.S., La Lande, T.J.: Overview of Clinical Decision Support Systems. In: Berner, E.S. (ed.) Clinical Decision Support Systems, Theory and Practice, 1, 2nd edn. Springer, New York (2007)
13. Peleg, M., Tu, S.W.: Decision support, knowledge representation and management in medicine. In: 2006 IMIA Yearbook of Medical Informatics: Assessing Information—Technologies for Health, pp 72–80. Schattauer Verlagsgesellschaft GmbH, Stuttgart (2006)
14. Osheroff, J.A., Teich, J.M., Middleton, B.F., Steen, E.B., Wright, A., Detmer, D.E.: A roadmap for national action on clinical decision support. JAMIA **14**(2), 141–145 (2007)

15. Carson, E.R., Collison, P.O., Kalogeropoulos, D.A.: Towards knowledge-based systems in clinical practice: development of an integrated clinical information and knowledge management support system. Comput. Methods Programs Biomed. **72**, 65–80 (2003)
16. Sanchez, E., Toro, C., Artetxe, A., Graña, M., Sanin, C., Szczerbicki, E., Carrasco, E., Guijarro, F.: Bridging challenges of clinical decision support systems with a semantic approach. A case study on breast cancer. Pattern Recogn. Lett. **34**(14), 1758–1768 (2013)
17. Sanchez, E., Toro, C., Graña, M., Sanin, C., Szczerbicki, E.: Extended Reflexive Ontologies for the generation of clinical recommendations. Cyberntetics Syst. **46**(1–2), 4–18 (2015)
18. Sanchez, E., Peng, W., Toro, C., Sanin, C., Graña, M., Szczerbicki, E., Carrasco, E., Guijarro, F., Brualla, L.: Decisional DNA for modeling and reuse of experiential clinical assessments in breast cancer diagnosis and treatment. Neurocomputing **146**(25), 308–318 (2014)
19. Poelmans, J., Ignatov, D.I., Kuznetsov, S.O., Dedene, G.: Formal concept analysis in knowledge processing: a survey on applications. Expert Syst. Appl. **40**(16), 6538–6560 (2013)
20. Sanin, C., Szczerbicki, E.: Decisional DNA and the Smart Knowledge Management System: A process of transforming information into knowledge. In: Gunasekaran, A. (ed.) Techniques and Tools for the Design and Implementation of Enterprise Information Systems, pp. 149–175. IGI, New York (2008)