Volume 6, Issue 9, September 2016 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Enhanced and Secure Electrical Health Record Retrieval Protocol

C. Senthil Selvi
Ph.D Scholar, Department of Computer Science,
Bharathiar University, Coimbatore,
Tamilnadu, India

Dr. N. Vetrivelan
Professor, Department of Computer Science, Principal in
Srinivasan College of Arts and Science, Perambalur,
Tamilnadu, India

Abstract— As the use of Electronic Medical Records (EMRs) becomes more widespread, so does the need to look and give viable information disclosure on them. Information disclosure strategies will permit experts and other healthcare stakeholders to find significant pieces of information in the growing corpus of accessible EMRs. The victory of Web look motors has appeared that keyword questions are a valuable device for locating significant information in an instinctive and viable manner. However, questions emerge of the form: What are the semantics of keyword questions on EMRs? What is a vital result? What is the role of medical and clinical ontologies and Lexicons like SNOMED (Systematized Classification of Human and Veterinary Medicine) in answering such queries? In this position paper we introduce the issue of keyword-based information disclosure on EMRs and enumerate the salient challenges that must be addressed to encourage quality information discovery. The objective is to make interest in new medical information administration relook initiatives, and conceivably make new paradigms for utilizing medical data. The primary center of the paper is the newest XML-based EMR standard created by the Health Level Seven (HL7) group, the Clinical Archive Engineering (CDA) Release 2.0, although the same issues emerge for any other standard hierarchical format.

Keywords— Electronic Health Records, Health Information Technology, Information Integrity, EMR Technology.

I. INTRODUCTION

The National Health Information Network (NHIN) and its data-sharing building blocks, RHIOs (Regional Health Information Organizations), are encouraging the widespread reception of Electronic Medical Records (EMR) for all hospitals inside the next five years. To date, there has been little or no effort to characterize strategies or approaches to rapidly look such archives and return vital results. One of the most promising models for EMR The definition and reception of this standard presents new challenges to related computer science disciplines like information management, information mining and information retrieval. In this position paper we study the issue of facilitating information disclosure on a corpus of CDA documents, i.e., given a question (query) and a set of CDA EMRs, find the substances (regularly subtrees) that are “good” for the query, and rank them according to their “goodness” with respect to the query. The victory of Web look motors has appeared that keyword questions are a valuable and instinctive information disclosure approach. Therefore, we mainly center on keyword questions in this paper, although some issues going beyond plain keyword questions are moreover examined.

As an example, consider the usual situation where a specialist wants to check conceivable conflicts between two drugs. Keyword question “drug-A drug-B death” could be submitted to discover cases where a patient who took both drugs died. Note that the word “death” can be indicated in many diverse parts of a CDA document, and moreover synonyms or related terms like “mortality” can be used instead. The latter can be tackled by leveraging fitting medical ontologies like SNOMED Clinical Terminology (SNOMED CT) as discussed below.

The key positioning criteria found in current systems as well as the bibliography are (a) relevance, (b) quality (authority) and (c) specificity. It is testing to characterize the information disclosure semantics for CDA archives such that the three aforementioned key positioning criteria are considered, given the hierarchical structure and particular semantics of CDA, and the fundamental references to outside substances like dictionaries, ontologies, separate text, or multimedia patient data. Medical Lexicons and ontologies regularly used in CDA are SNOMED CT and LOINC. We moreover study how past work on information disclosure on XML information (Area 2.2) can be leveraged, and what limitations might exist in this unique domain. We note that our study does not discuss the vital privacy issues involved in accessing patient information, as required by HIPAA.

The broadened variant of this work describes more challenges and discusses more related work. The rest of this paper is organized as follows: Area 2 presents a background exposition of current clinical information models and a brief survey on information disclosure on XML data. Area 3 addresses the challenges that we have recognized to execute information disclosure on a corpus of EMR documents. Our concluding remarks are displayed in Area 4.

II. BACKGROUND

In this Area we audit key models used to represent clinical information and EMRs and present past work on information disclosure on general XML documents.

2.1 Clinical Information Model and Ontologies

Reference Information Model (RIM): HL7 is a language, and every dialect has a grammar. The HL7 RIM indicates the punctuation of HL7 messages and the fundamental building squares of the dialect and their permitted relationships. For more details see.

Systematized Classification of Medication (SNOMED): SNOMED has grown up into a comprehensive set of over 150,000 records in twelve diverse chapters or axes. SNOMED Clinical Terms (SNOMED CT) is a universal health care terminology and infrastructure.

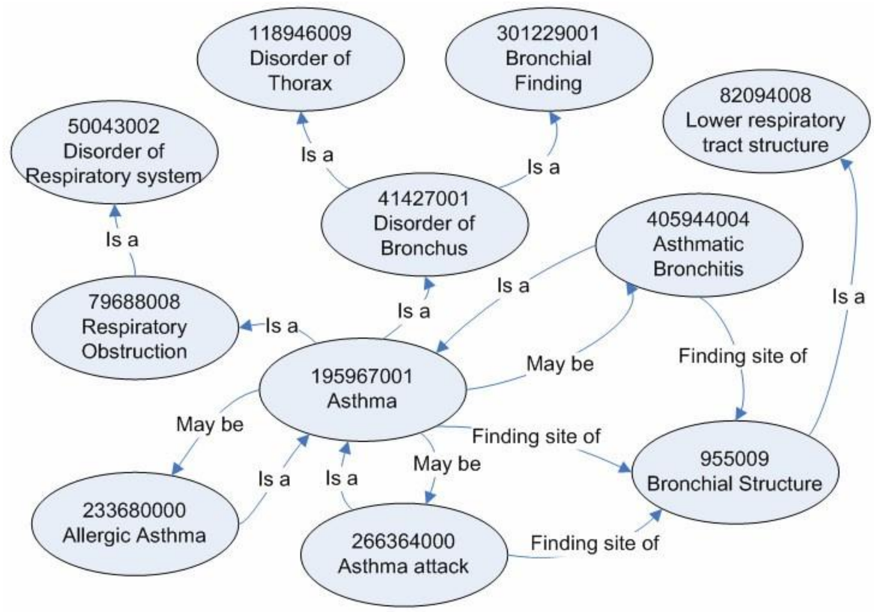

Figure 1 shows a sub-diagram of the SNOMED ontology graph

Clinical Archive Architecture: The Clinical Archive Engineering (CDA) is an XML-based archive markup standard that indicates the structure and semantics of clinical documents, such as discharge summaries and progress notes, for the purpose of exchange. It is an American National Models (ANSI) approved HL7 standard, intended to become the de facto electronic medical record. Figure 2 depicts a sample CDA archive D1, which is wrapped by the "ClinicalDocument" element, as it appears in Line 2 of this figure.

2.2. Looking XML Archives

In this Area we present an overview of past work on looking XML documents. This corpus of work will be seen as the starting point to present the challenges of information disclosure on CDA XML archives in Area 3. XRANK positions the XML parts by generalizing the Page-Rank algorithm, combining the positioning of parts with keyword proximity. XLook positions the results taking into thought both the degrees of the semantic relationship and the significance of Figure 1: Partial SNOMED ontology for the term "Asthma" the keyword. Cohen et al. present an broadened framework to indicate the semantic relationship of XML elements. XIRQL utilizes a diverse strategy to compute its ranking, defining index units, particular entity sorts that can be indexed and used for tf-idf computation.

III. CHALLENGES OF INFORMATION DISCLOSURE ON CDA DOCUMENTS

In this Area we present a arrangement of challenges that have to be addressed to effectively perform information disclosure on a corpus of CDA documents. For simplicity we center on plain keyword queries, although the same challenges are valid for semi-organized questions as well a semi-organized question is a question where partial information about the structure of the results is provided. Detailed talk and examples for each of the challenges are displayed in the broadened version.

We discuss why the general work on looking on XML archives (Area 2.2) is not adequate to give quality information disclosure on CDA XML documents. The key reasons are the complex and domain-particular semantics and the frequent references to outside information sources like Lexicons and ontologies. We use Archive D1 depicted in Figure 2 as our running example.

3.1. Structure and Scope of Results

In contrast to customary Web look where entirety HTML archives are returned as question results, in the case of XML archives and especially CDA documents, we need to characterize what a vital question result is. Past work has studied diverse approaches to characterize the structure of results.


```
<LevelThree>
<header>
 <!-- same as LevelOne header-->
</header>
<body>
 <procedure>
  <paragraph>
   <healthcare.code identifier="P5-20100"
    name.of.coding.system="SNM3"
    local.coding.system="N">Chest X-Ray
   </healthcare.code>
  </paragraph>
 </procedure>
 <findings>
  <paragraph>RLL nodule</paragraph>
 </findings>
 <impressions>
  <Clinical_observation>
   <observation_value_txt code="M-03010"
    source="SNM3"/>Nodule
    <Observation_relationship>
     <relationship_type_cd code="GC006"
      source="SNM3"/>in the
      <Clinical_observation>
       <observation_value_txt code="T-28400"
        source="SNM3"/>RLL,
      </Clinical_observation>
    </Observation_relationship>
    <Observation_relationship>
     <relationship_type_cd code="G-C022"
      source="SNM3"/>suggestive of
      <Clinical_observation>
       <observation_value_txt code="M-80001"
        source="SNM3"/>malignancy.
      </Clinical_observation>
    </Observation_relationship>
  </Clinical_observation>
 </impressions>
 <recommendations>
  <paragraph>I notified the ordering physician of this
   finding.</paragraph>
 </recommendations>
</body>
</LevelThree>
```

Figure 2: HL7 CDA Sample Document.

A corpus of works consider a entirety subtree as result, that is, a result is unambiguously defined by the lowest fundamental ancestor (LCA) hub of the keyword nodes. We allude to this approach as subtree- as-result. For example, XRANK favors deeply nested elements, returning the deepest hub containing the keywords as the most particular one, having more content information. In contrast, a way as the result is proposed by; where a insignificant way of XML hubs is returned that collectively contain all the question keywords. Note that we use the term “path” loosely to differentiate it from the subtree-as-result approach, since it can be a collection of meeting ways (a tree) for more than two question keywords. We allude to this approach as path-as-result. It is unclear whether the subtree-as-result or the path-as-result is a better fit for looking CDA documents. The talk on insignificant information unit beneath sheds more light to this aspect. Another issue is the scope of a result, in particular, whether results spanning over EMRs should be produced. Finally, specialists would like to be able to indicate the results’ pattern in some cases, which in turn limits the sorts of parts searched for the question keywords.

3.2. Insignificant Information Unit (IIU)

It is testing to characterize the granularity of a piece of information in a way that it is selfcontained and meaningful, yet at the same time specific. For example, in Archive D1 returning the “value” component of Line 45 without the preceding “code” component is not vital for the user. Hence, the “value” component is not an fitting IIU, whereas the enclosing “Observation” component could be.

Furthermore, for some questions it is required to incorporate into the result some parts that do not contribute in interfacing the question keywords or are part of the MIU of such a interfacing node. For instance, the “patientPatient” component should be included in the result of question “Asthma Theophylline” if a practitioner submits the query, yet not if a specialist does. Such personalization issues are further discussed in Area 3.11.

3.3. Semantics of Hub and Edge Sort

It is testing to incorporate the rich semantic information accessible for the clinical domain, and especially for the parts of a CDA document, in the results' positioning process. At the most basic, a area expert statically assigns a weight to each hub and edge type, as in BANKS. In expansion to that, we can relegate a significance to entirety ways on the pattern as explained below.

Furthermore, it is desirable that the degrees of semantic affiliation are adjusted dynamically exploiting significance feedback and learning techniques.

3.4. Access to Lexicons and Ontologies

CDA archives routinely contain references to outside dictionary and ontology sources through numeric codes. As an example, archive D1 incorporates references to LOINC and SNOMED CT in Lines 34 and 38 respectively. Hence, it is no longer enough to answer a question considering the CDA archive in isolation, as is done by all past work on information disclosure on XML archives (Area 2.3). In this setting, the question keywords may allude to content in the CDA archive or an ontology that is connected to the CDA archive through a code reference. For example, the question keyword "appendicitis" may not be present in the archive yet its code might be present, so we need to go to the ontology and look for the question keyword there.

3.5. Access to Lexicons and Ontologies

We need to relegate an fitting esteem to each of the relations present in the ontologies. SNOMED CT, for example, has four diverse sorts of relationships: (1) Defining characteristics, (2) Qualifying characteristics, (3) Historical connections and (4) Other relationships. Figure 1 incorporates relations such as "May be", "Finding site of" and "Has finding site" in expansion to the most fundamental "Is a" relationship. Stricter and stronger relations in the ontology should intuitively have a higher weight. Furthermore, we need to take into thought the bearing of the edges. A conceivable approach to measure the degree of affiliation between hubs of an ontology diagram is to execute ObjectRank on the ontology graph, as described by Hwang et al.

3.6. Arbitrary Levels of Nesting

We can find an arbitrary number of levels of nesting and recursion in the definition of parts and sections, as appeared in the way component section. Component. Area in Lines 58-63 of Figure 2. Taking into thought the semantics of the document, the interconnection relationship rule of XSEarch, where the same tag may not appear twice in internal hubs of a result path, cannot be applied since the same tag can appear twice in a vertical way (top-to-bottom).

3.7. Free Content Embedded in CDA Archives

In some cases, plain content descriptions are added to certain areas to enrich the information about the record or to express a real life property not codified in Lexicons or ontologies. As a first measure, customary text-based Information Recovery systems should be included in the Engineering to support such cases. Another technique to address the coexistence of semi-organized and unstructured information is displayed in, where IR and closeness rankings are combined.

3.8. Time and Area Attributes

After discussing with medical researchers and practitioners, we found that time and area are basic attributes in most queries. For instance, for the question "drug-A drug-B" the specialist is probably looking for any conflict between these drugs, and henceforth the time separation between the prescriptions of these drugs for a patient is a basic piece of information. Area is moreover vital since two patients located in nearby beds in the hospital should be seen as associated since infections tend to transmit to neighbouring beds. Clearly, it is testing to standardize the representation of such area information inside an EMR.

Furthermore, time and area can lead to the definition of metrics comparative to the inverse archive frequency (idf) in Information Retrieval. For instance, asthma is more fundamental in summer; henceforth a patient who has asthma in winter should be positioned higher for the question "asthma". Similarly, a patient who has the influenza in a town where no one else has it should be positioned higher for the question "flu". These associations are too complex since time can be used to characterize time, distance, or periodicity. Similarly, area connections can be indicated either inside a hospital or over towns.

3.9. EMR Document-as-Question

An alternative question sort to the plain keyword question is utilizing a entirety (or part of) EMR (CDA) archive as the query. This approach can be used to find comparative CDA documents, that is, CDA archives of patients with comparative history, demographic information, treatments, and so on. The client should be able to customize and personalize such an information disclosure device to fit her needs. For instance, a specialist may not consider the physician's (author of CDA document) name when matching CDA documents, and could indicate that a generic solution should be seen as identical to the non-generic equivalent. Past work on archive content similitude and XML archive structural similitude can be used to solve this problem.

Furthermore, such document-as-question questions can be used to find medical writing significant to the current patient. In this scenario, the EMR application could have a button named "significant literature" that invokes an information disclosure calculation on PubMed or other medical sources. Price et al. present a first attempt towards this direction, where they extract all MeSH terms (MeSH refers to the U.S. National Library of Medicine's controlled vocabulary used for indexing articles for MEDLINE/PubMed) from an EMR (not particular to CDA) and then question MEDLINE utilizing these terms. The structured position of CDA archives can conceivably permit more elaborate looking algorithms where multiple terms that are structurally correlated can construct a single and more focused question on medical writing sources.

3.10. Handle Negative Explanations

A substantial fraction of the clinical observations entered into patient records are expressed by means of negation. Elkin et al. found SNOMED-CT to give coverage for 14,792 ideas in 41 health records from Johns Hopkins University, of which 1,823 (12.3%) were recognized as negative by human review. Today, one has to examine the terms preceding a diagnosis to determine if this diagnosis was barred or not. Ceusters and Smith propose new ontological connections to express “negative findings”. It is testing to handle such negative explanations for an information disclosure question in a way that the client can indicate whether negated ideas should be barred or not from the look process.

3.11. Personalization

The information disclosure engine should give personalized results depending on the preferences of each individual user. For example, for diverse doctors, diverse substances and connections in the CDA parts are more important. For some healthcare providers, the solution may be more significant than the observation, or the solution may be more significant than the specialist name. Moreover the connections in ontologies may be seen differently. Furthermore, depending on whether a client is a nurse, a pharmacist, a technician or a physician, the system could automatically relegate diverse weights on edges and hubs of the CDA Object Model to encourage the information needs of the users.

IV. CONCLUDING REMARKS

We have introduced the issue of Information Disclosure on Electronic Medical Records (EMR), enumerating a arrangement of challenges that must be addressed to give quality information disclosure services on EMRs, specifically on HL7 CDA documents. The key challenges are related to the semantics of the architecture, the XML structure of CDA documents, and the interconnection of EMR archives with ontologies and dictionaries. Additional challenges incorporate the incorporation of time and area semantics, as well as handling negative statements. We hope that this work will spawn new relook directions to address these challenges. The successful realization of information disclosure on EMRs is expected to have a great impact on the quality of healthcare.

REFERENCES

[1] Denise L. Anthony; Celeste Campos-Castillo, “Do Health Care Users Think Electronic Health Records are Important for Themselves and Their Providers? Exploring Group Differences in a National Survey”, Healthcare Informatics (ICHI), 2013 IEEE International Conference on Year: 2013 Pages: 141 – 146.

[2] Weiwei Jiang; Haishun Wang; Xiaomei Xu; Chun Peng, “Individual Self-Service Electronic Health Records: Architecture, Key Technologies and Prototype System”, Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC), 2011 International Conference on Year: 2011 Pages: 574 – 579.

[3] Pieterjan De Potter; Pedro Debevere; Erik Mannens; Rik Van de Walle, “Next generation assisting clinical applications by using semantic-aware electronic health records”, Computer-Based Medical Systems, 2009. CBMS 2009. 22nd IEEE International Symposium on Year: 2009 Pages: 1 – 5.

[4] Usman Mushtaq; Kevin Hall, “Power and agency in health information technology: towards a more meaningful participatory design for sustainable development”, Science and Technology for Humanity (TIC-STH), 2009 IEEE Toronto International Conference Year: 2009 Pages: 1036 – 1041.

[5] Leslie Mertz, “Electronic Health Records Usher in a New Era in Health Care: Making the Transition to Health Information Technology”, IEEE Pulse Year: 2012, Volume: 3, Issue: 6 Pages: 43 – 50.

[6] Anju Mammen; Richard Weeks, “Electronic Medical Record (EMR) technology acceptance by healthcare professionals in South Africa”, Proceedings of PICMET '14 Conference: Portland International Center for Management of Engineering and Technology; Infrastructure and Service Integration Year: 2014 Pages: 3539 – 3548.

[7] Vichita Vathanophas; Tullawat Pacharapha, “Information Technology Acceptance in healthcare service: The study of Electronic Medical Record (EMR) in Thailand”, PICMET 2010 TECHNOLOGY MANAGEMENT FOR GLOBAL ECONOMIC GROWTH Year: 2010 Pages: 1 – 5.

[8] Yabin Xu; Yunmei Shi; Xuhong Liu, “The integration technology of multi-source heterogeneous EMR Based on semantic grid”, Information Theory and Information Security (ICITIS), 2010 IEEE International Conference on Year: 2010 Pages: 642 – 645.

