An Ontology-Based Electronic Medical Record for Chronic Disease Management

by

Ashraf Mohammed Iqbal

Submitted in partial fulfillment of the requirements
for the degree of Master of Computer Science

at

Dalhousie University
Halifax, Nova Scotia
February, 2011

DALHOUSIE UNIVERSITY

Faculty of Computer Science

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled "An Ontology-Based Electronic Medical Record for Chronic Disease Management" by Ashraf Mohammed Iqbal in partial fulfillment of the requirements for the degree of Master of Computer Science.

Dated: 16, February, 2011

Supervisor: _________________________________

Co-Supervisor: _________________________________

Reader: _________________________________

DALHOUSIE UNIVERSITY

DATE: 16, February, 2011

AUTHOR: Ashraf Mohammed Iqbal

TITLE: An Ontology-Based Electronic Medical Record for Chronic Disease Management

DEPARTMENT OR SCHOOL: Computer Science

DEGREE: MCS CONVOCATION: May YEAR: 2011

____
Signature of Author

DEDICATION PAGE

To My Niece and Nephew

Labiba and Shahfin

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Effective chronic disease management ensures better treatment and reduces medical costs. Representing knowledge through building an ontology for Electronic Medical Records (EMRs) is important to achieve semantic interoperability among healthcare information systems and to better execute decision support systems. In this thesis, an ontology-based EMR focusing on Chronic Disease Management is proposed. The W3C Computer-based Patient Record ontology [32] is customized and augmented with concepts and attributes from the Western Health Infostructure Canada chronic disease management model [27] and the American Society for Testing and Materials International EHR. The result is an EMR ontology capable of representing knowledge about chronic disease. All of the clinical actions of the proposed ontology were found to map to HL7 RIM classes. Such an EMR ontology for chronic disease management can support reasoning for clinical decision support systems as well as act as a switching language from one EMR standard to another for chronic disease knowledge.

LIST OF ABBREVIATIONS USED

ADL	Acrhetype Definition Language
ASTM	American Society for Testing and Materials
CDA	Clinical Document Architecture
CDM	Chronic Disease Management
CIR	Clinical Investigator Recording
CPR	Computer Based Patient Record
EHR	Electronic Health Record
EHCR	Electronic Health Care Record
EMR	Electronic Medical Record
EPR	Electronic Patient Record
GEHR	Good Electronic Health Record
HIPAA	Health Insurance Portability and Accountability Act
HL7	Health Level Seven
LOINC	Logical Observation Identifiers Names and Codes
OCR	Optical Character Recognition
OWL	Web Ontology Language
POMR	Problem Oriented Medical Record
PROMIS	Problem-Oriented Medical Information System
RIM	Reference Information Model
SNOMED-CT	Systematized Nomenclature of Medicine--Clinical Terms
SOAP	Subjective Objective Assessment Plan
SWRL	Semantic Web Rule Language
W3C	World Wide Web Consortium
WHIC	Western Health Infostrcuture Canada

ACKNOWLEDGEMENTS

I would like to thank Dr. Michael Shepherd for supervising me towards my Masters Thesis. It would not be possible for me to complete my work without his proper guidance, support, encouragements and patience.

I would also like to thank Dr. Syed Sibte Raza Abidi for guiding me throughout the entire time of my thesis. His insightful thoughts have been very much helpful in solving different problems of this thesis.

Many thanks to Dr. Grace Paterson for being a reader of this thesis. Her comments have been really useful not only to make necessary corrections of my thesis but also at different occasions throughout the entire period the thesis was done.

I would also like to acknowledge the great support from Green Shield Canada Foundation for funding this research.

I feel lucky to get the chance of working in an environment of WIFL lab at Dalhousie University. All of the fellow members of both WIFL and NICHE labs have been really helpful.

My parents, brother, sister, and wife have always given me moral support to work hard. I would like to give my special thanks to Barbara Borden at Faculty of Computer Science, Dalhousie University for providing me moral support and guidance.

CHAPTER 1 INTRODUCTION

Chronic diseases are long-term, rarely cured, and have a slow progression. They have been identified as the major cause of deaths the world over [1]. According to the Director General of the World Health Organization [2], chronic diseases have surpassed acute diseases as the major cause of deaths, resulting in 60% of deaths, and are projected to increase a further 17% in the next ten years. In Canada, about two-thirds of total deaths were caused by chronic diseases in 2009 [3]. Chronic diseases are associated with massive direct medical care costs and indirect social costs. In 2004 in Canada alone, such costs exceeded $93 billion [3].

Effective chronic disease management is therefore necessary for cost reduction and quality care. The inter-related elements of such effective chronic disease management include [4]: health system (creating a culture, organization and mechanisms for high quality care), self-management support (educating and supporting patients to take a proactive role in self-care), delivery system design (ensuring the delivery of effective, efficient clinical care and self-management support), decision support (providing better clinical care based on scientific evidence and patient preferences), clinical information system (organizing patient data for timely access and better care delivery), and the community (encouraging patients to participate in effective community programs). Clinical information systems are key components in such a management plan for readily accessing data of individuals and populations, tracking and monitoring the progression of diseases, taking effective care plans, and providing data to better execute the decision support systems.

An Electronic Medical Record (EMR) is the key component of a clinical information system to capture the longitudinal medical records of patients. An EMR can reduce errors in data entry, ensure timely accessibility of information by simultaneous multiple users, can be used to support reasoning in decision support systems, and hence can reduce care costs. The knowledge representation of EMR is imperative to ensure semantic interoperability and to facilitate reasoning by decision support systems.

1.1 RESEARCH STATEMENT

Some EMR standards have already been proposed in the literature [34-37, 43], while HL7 provides the most widely used messaging standard among healthcare information systems. There also have been some efforts in representing the knowledge of EMRs through building an ontology [31, 32, 44-48]. Unfortunately, very few of these efforts have yet fulfilled the complete terminological representations and information elements needed for successful chronic disease management. While there have been some successful implementations of complete chronic disease management models [27-30, 74, 75], these consider only the information needs for chronic disease management. However, an EMR provides such information needs not only for chronic disease management, but also for other acute and infectious diseases. Moreover, very few successful harmonizations [29] between the HL7 messaging standard and these EMR standards have yet been evolved.

We argue that a coherent ontology of an EMR is necessary to address chronic disease management while providing support for acute diseases. Such an EMR ontology cannot only serve the purposes an EMR is primarily expected to do, but can also play an important role in classifying taxonomies and providing conceptualization of medical records. Such an EMR ontology can further used by clinical decision support systems for reasoning purposes. This EMR ontology should also fully capture the clinical message elements of HL7.

These imply two important research questions we intend to answer in this thesis:

- Can we represent knowledge through building an ontology for an Electronic Medical Record that focuses primarily on Chronic Disease Management while capturing data for other acute diseases?

- Can we map the HL7 Reference Information Model with the EMR ontology to ensure that the clinical messages written in HL7 are fully captured by the EMR ontology?

1.2 Research Challenges

The above mentioned research objectives are associated with a number of research challenges, such as:

a) EMR Ontology

The EMR ontology should be capable of representing all concepts necessary for appropriate classification of terminologies in a medical record. The concepts should have the necessary properties to hold longitudinal patient data. The semantic relationships between these concepts should be correctly defined to ensure proper reasoning and execution by decision support systems.

b) EMR ontology and CDM Model

Successful chronic disease management requires long-term interventions and follow-ups. Planned clinical actions, in addition to the actual actions, with associated monitoring of outcomes are crucial particularly for such chronic disease management. Thus, it is important to ensure that the concepts, properties and semantic relationships necessary for a CDM model are well supported by the resultant EMR ontology. At the same time, the EMR ontology should be capable of supporting other acute diseases while maintaining a coherent ontological representation.

c) EMR ontology and HL7 Reference Information Model

To ensure that the clinical messages written in HL7 are fully captured by the resultant EMR ontology, a mapping is necessary between HL7 RIM (the object-oriented model of HL7 version 3) and the EMR ontology. Since RIM provides a more robust clinical information model, a refined model of it should be devised.

1.3 Solution Approach

Our main objective of this thesis is to develop an ontology-based EMR that will hold data for chronic disease management while serving a general purpose. Since, there are some differences between the information model of an EMR and the information model required for chronic disease management, the W3C proposed Computer-Based Patient Record (CPR) ontology [32] is customized and mapped onto the WHIC proposed chronic disease management model [27] as shown in Figure 1.1.

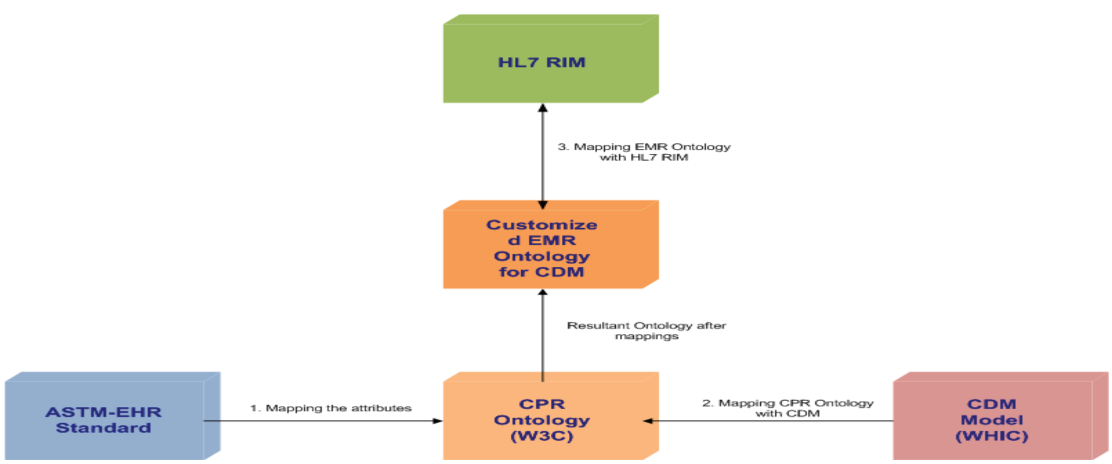


Figure 1.1 Solution approach of Ontology-based EMR for Chronic Disease Management.

We imported the attributes necessary for concepts of this ontology from the standard data elements proposed by ASTM International for an EHR [43]. Our proposed ontology ensures a structured means for data entry by integrating the controlled vocabulary from SNOMED-CT. Furthermore, the resultant ontology is mapped to HL7 RIM to capture the clinical messages written in HL7.

1.4 Background

1.4.1 Electronic Medical Record

An Electronic Medical Record (EMR) is a digital, comprehensive, longitudinal, integrated set of medical records ensuring on-time, multiple access of present and past medical information, with built-in restrictions so that only the right people can access the requisite information. Although there are several definitions of EMR, no single one is yet internationally accepted.

According to HINA [5], an EMR is "an electronic longitudinal collection of personal health information usually based on the individual, entered or accepted by health care providers, which can be distributed over a number of sites or aggregated at a particular source. The information is organized primarily to support continuing, efficient and quality health care. The record is under the control of the consumer and is stored and secured securely."

Similarly, OHIH [6] defines an EMR as "a longitudinal collection of personal health information of a single individual, entered or accepted by health care providers, and stored electronically. The record may be made available at any time to providers, who have been authorized by the individual, as a tool in the provision of health care services.

The individual has access to the record and can request changes to its contents. The transmission and storage of the record is under strict security."

According to CPRI [7], an EMR is "a virtual compilation of non-redundant health data about a person across a lifetime, including facts, observations, interpretations, plans, actions, and outcomes. Health data include information on allergies, history of illness and injury, functional status, diagnostic studies, assessments, orders, consultation reports, treatment records, etc. Health data also include wellness data such as immunization history, behavioral data, environmental information, demographics, health insurance, administrative data for care delivery processes, and legal data such as consents."

Electronic Medical Record (EMR) is also sometimes referred to as *Electronic Heath Record (EHR)*, *Electronic Patient Record (EPR)*, *Electronic Health Care Record (EHCR)*, *Computerized Patient Record (CPR)*, etc., in different countries and at different healthcare organizations. D. Garets et al. [8] distinguished EMR from those, stating that EHR is a subset of each care delivery organization's EMR and is owned by a patient or stakeholder whereas EMR is owned by a care delivery organization.

The idea of computerized patient information instead of using paper was first introduced in the 1960s by Lawrence L. Weed, a physician [9]. Based on Weed's idea, information scientists and physicians at the University of Vermont started developing an EMR in 1967 called *Problem Oriented Medical Record (POMR)*, which was first used in 1970. The EMR plays a central role in today's computerized health information system, upon which other clinical systems (e.g., decision support systems, order processing systems, etc.) are dependant.

1.4.2 Advantages of Electronic Medical Record

An EMR can improve the quality of healthcare in a number of ways. It reduces the physical space required for storing the files and folders in healthcare organizations. Multiple users (e.g., general practitioners, patients, medical staff, insurance companies, etc.) can access the same electronic record at the same time, which is not possible in paper-based systems. Furthermore, it is no longer necessary for patients to bring their medical documents to practitioners, which precludes their losing the documents. Moreover, handwritten records may result in poor legibility [10], which in turn might result in medical errors. For this reason, *Optical Character Recognition (OCR)*, which scans handwritten records into electronic format, would not be successful. Lærum Hallvard et al. [13] conducted a survey on 70 physicians from 6 hospitals to compare the usages of OCR and EMR and found that where the majority of the physicians found EMR more easy, 22-25% of the physicians found retrieval of patient data in EMR more difficult.

One of the major advantages of an EMR is interoperability, which is loosely defined as the ability of different information technology systems and software applications to communicate and to exchange data accurately, effectively, and consistently, and to use the information that has been exchanged [11]. This ensures on-time, secured data exchange between different units of an organization and also between different healthcare organizations, which in turn improves the quality of care. It also reduces the chance of providing duplicate medical services (e.g., laboratory tests, diagnoses, etc.). As stated in [12], such integration can reduce administrative overhead costs, data errors, and observed outcomes can be more rapidly identified.

In addition, EMR can make billing more accurate and fast by being integrated into the billing system. It preserves historical medical records and thereby ensures instant accessibility and retrieval of old medical records when needed. Also, analyzing these data can be useful for researchers as well as for the management of healthcare organizations when making decisions about resource management and quality improvement. These data can also be useful as inputs to clinical decision systems.

1.4.3 Ontology

One way of representing knowledge is through building an ontology. The term ‘ontology’ originated in metaphysics, a branch of philosophy, and is a study of ‘being’ or ‘existence’. It determines what entities exist or are said to exist, grouped together or subdivided according to their similarities and differences. Recently, ontology is being used in semantic web design to imply a technical meaning. Tom Gruber [18] suggests that ontology is a “formal, explicit specification of a shared conceptualization”, whereas Fredrik Arvidsson and Annika Flycht-Eriksson [19] define is as such: “An ontology provides a shared vocabulary, which can be used to model a domain – that is, the type of objects and/or concepts that exist, and their properties and relations.”

A more technical definition of ontology is provided by J. F.Sowa [21] as follows: “The subject of ontology is the study of the categories of things that exist or may exist in some domain. The product of such a study… is a catalog of the types of things that are assumed to exist in a domain of interest *D* from the perspective of a person who uses a language *L* for the purpose of talking about *D*.”

Another technical definition of ontology in terms of its necessary components is given by Further, Uschold et al. [22] as follows: “An ontology may take a variety of forms, but necessarily it will include a vocabulary of terms, and some specification of their meaning. This includes definitions and an indication of how concepts are inter-related, which collectively imposes a structure on the domain and constrains the possible interpretations of terms”.

According to Grigoris Antoniou and Frank van Harmelen [20], an ontology typically consists of a finite set of terms (classes of objects of the domain), relationships among such terms, properties of relationships, value restrictions, disjointness between terms, and specification of logical relationships between objects.

Although there are wide range of definitions for an ontology available from a mathematics point-of-view, we have adopted the one used by Fahim Imam at [93]:

An Ontology, O can be defined as the 4-tuple $O = (C, \Re, I, A)$, where, C is the set of concepts, $\Re$ is the set of binary relations, I is the set of instances, and A is the set of axioms. Thus, according to this definition, each ontology should primarily have four sets of components, which are described below.

Concepts: Concepts (also called Classes) of an ontology are abstract object (i.e., whose existences are independent of time and location) categories or types in real world. Concepts can generalize (i.e., contain), or can specialize (i.e., subsume) other concepts. For example, the concept 'Person' in a clinical domain can further be specialized by other concepts like 'Patient', 'Physician', 'Nurse' etc.

There are two important categories of concepts in an ontology:

- **Primitive Concepts**: The concepts which only have necessary conditions in an ontology are called primitive concepts. Necessary conditions of a statement are such conditions those must be satisfied for the statement to be true. For example, 'Chronic Kidney Disease' is a 'Chronic Condition' – this is a necessary condition, making 'Chronic Kidney Disease' a primitive concept.

- **Named Concepts**: The concepts which both have necessary and sufficient conditions in an ontology are called named concepts. Sufficient conditions of a statement are such conditions that if satisfied, then the statement must be true. For example, a 'Patient' is a 'Person' is a necessary condition, and, a 'Patient' has a 'Medical-Problem' is a sufficient condition, which makes the 'Patient' a named concept.

Relations: Relations (also known as properties) in an ontology are binary predicates which relate between two concepts, or two relations. This ontology component can primarily be divided into two kinds:

- **Taxonomic Relations**: These relations (known as 'is-a', 'is-kind-of', 'is-part-of') define the hierarchy of concepts in an ontology, i.e., the generalizations/specializations relations among different concepts. For example, we can specify 'Patient' as a sub-concept of 'Person' using the taxonomic relation: Patient is-a Person.

- **Associative Relations**: These (popularly known as 'has' or 'is' relations) relate between two concepts in an ontology. The source concept in an associate relation is called Domain, and the target concept is called Range. For example, if we want to specify that a person has an age, we can use the associative relation, 'hasAge' between 'Person' and 'Age' concepts. Inspired by descriptive logic, these relations can further have some special properties:

    - **Reflexive Relations**: These (also known as equivalence relations) relate between two equivalent concepts. For example, if we say that the concepts 'Physician' and 'General Practitioner' are related by a reflexive relation, that essentially means that these two concepts are equivalent.

    - **Functional Relations**: Each Domain value can be associated with at most one unique Range value, i.e., if $x\Re_F y$ and $x\Re_F z$, then, $y = z$, where $\Re_F$ denotes a functional relation. For example, each person can have only one date of birth – this condition can be satisfied by using a functional relation.

- **Transitive Relations**: If $x\Re_T y$ and $y\Re_T z$, then $x\Re_T z$, where $\Re_T$ denotes a transitive relation. For example, John is an ancestor of Rob, and Rob is again an ancestor of David implies to the fact that John is an ancestor of David.

- **Symmetric Relations**: If $x\Re_S y$, then $y\Re_S x$, where $\Re_S$ denotes a symmetric relation. For example, John is married to Marry implies that Marry is also married to John.

Instances: The Instances are the basic 'ground level' objects for concepts in an ontology [93]. For example, 'Diabetes Type 2' can be an instance of the concept 'Medical-Problem' in a medical ontology.

Axioms: Axioms in an ontology are formulas (i.e., propositions in mathematics) to specify the interdependencies of concepts or relations on other components (i.e., on other concepts, relations, instances) of that ontology. For example, if 'husband' and 'married' are two relations in an ontology, we can specify a very simple logic axiom: $\forall_{x,y} husband(x,y) \rightarrow married(x,y)$

1.5 Thesis Outline

This thesis is organized into the following chapters.

Chapter 2 discusses various EMR standards currently available in the literature, HL7 messaging standards, chronic disease management models, as well as proposed ontologies on EMR and various ontology mapping techniques.

Chapter 3 briefly describes the proposed ontology-based electronic medical record for chronic disease management.

Chapter 4 describes the results of the ontology evaluation. Our proposed EMR ontology is evaluated against standard ontology design principles [23, 24]. Two sample medical records represented in HL7 [25, 26] are instantiated using the proposed ontology.

Chapter 5 summarizes the findings, highlights the limitations of our work, and recommends some future directions of research in related areas.

CHAPTER 2 LITERATURE REVIEW

2.1 Introduction

The objective of this thesis is to propose an ontology on Electronic Medical Record which will focus mainly on chronic disease management while serving the purposes of other acute and infectious diseases. In this chapter, we will first discuss some of the Electronic Medical Record standards from the literature and critically compare them. We will then briefly describe the object-oriented Reference Information Model of the most widely used message standard in today's healthcare information settings, Health Level Seven (HL7).

Although there have been some ongoing efforts in developing information models for chronic disease management models in different geographical locations (e.g., Norway, France, Canada etc.), we found the one proposed by Western Health Infostructure Canada [27] to be the most complete in terms of implementation. Thus, we will describe this model along with its vocabulary domains and mapping results with HL7.

W3C proposed a Problem-Oriented Medical Record (POMR)-based ontology which they call Computer-based Patient Record (CPR) [32]. We found this OWL-DL based ontology as the most robust and comprehensive one on an EMR that incorporated some other top-level ontologies (e.g., BFO 1.1, BIOTOP, FMA etc.) to ensure a sound and coherent means of necessary terminological representations required by an EMR. Details of this ontology are also described in this chapter.

An agreed standard methodology for ontology development is important to ensure re-usability of ontologies and interoperability among different applications. A few such methodologies proposed in the literature are discussed, and critically analyzed in the next section. The basic technical details of the most prominent ontology language, OWL, are discussed next.

Mapping among different distributed, heterogeneous ontologies has been an interesting topic in the literature. Different techniques of such mapping are particularly important in our research since we need to map between different ontologies (e.g., mapping between the Chronic Disease Model and EMR ontology). We will describe various state-of-the-art techniques, methods and tools of ontology mapping and will also depict some comparisons among these in the following section.

Finally, we will summarize our findings of the literature review by offering some insightful thoughts of our own.

2.2 Electronic Medical Record Standards

Extensive research is being carried out to develop comprehensive standard models for EMRs. These standard models provide a logical structure of information content. They also specify the relationship of this content to clinical concepts (architectural standards) and specify the syntax and representation of EMR information to be interchanged. Open-EHR, CEN-EN 13606, and the Problem-Oriented Medical Record (POMR) are examples of standards for EMR logical structures, while HL7 provides the most widely used messaging standard among healthcare information systems.

2.2.1 Open-EHR

Open-EHR [35], adopted from the Australian Good Electronic Health Record (GEHR), provides specifications for shared EMR, which is more “technology-based” than “standards-based” [33]. It consists of a two-level model, one level of which is a simple reference model (RM) and the other, a formal constraints model, which is called archetype. The RM describes the basic structure of clinical information whereas the archetype models provide architectural standards for EMR information and share common clinical definitions specified in the shared Open-EHR archetype repository. Each archetype can be considered a model containing clinical contents and can be expressed in a constraint formalism form. The information model is proposed as a separate model from the demographic information model, and the Extract package is archetyped containing both of them. The most important package of this RM model is 'ENTRY', which is based on the Clinical Investigator Recording (CIR) ontology [37]. It is claimed to fill the greatest portion of the POMR [37]. The CIR ontology is shown in Figure 2.1.

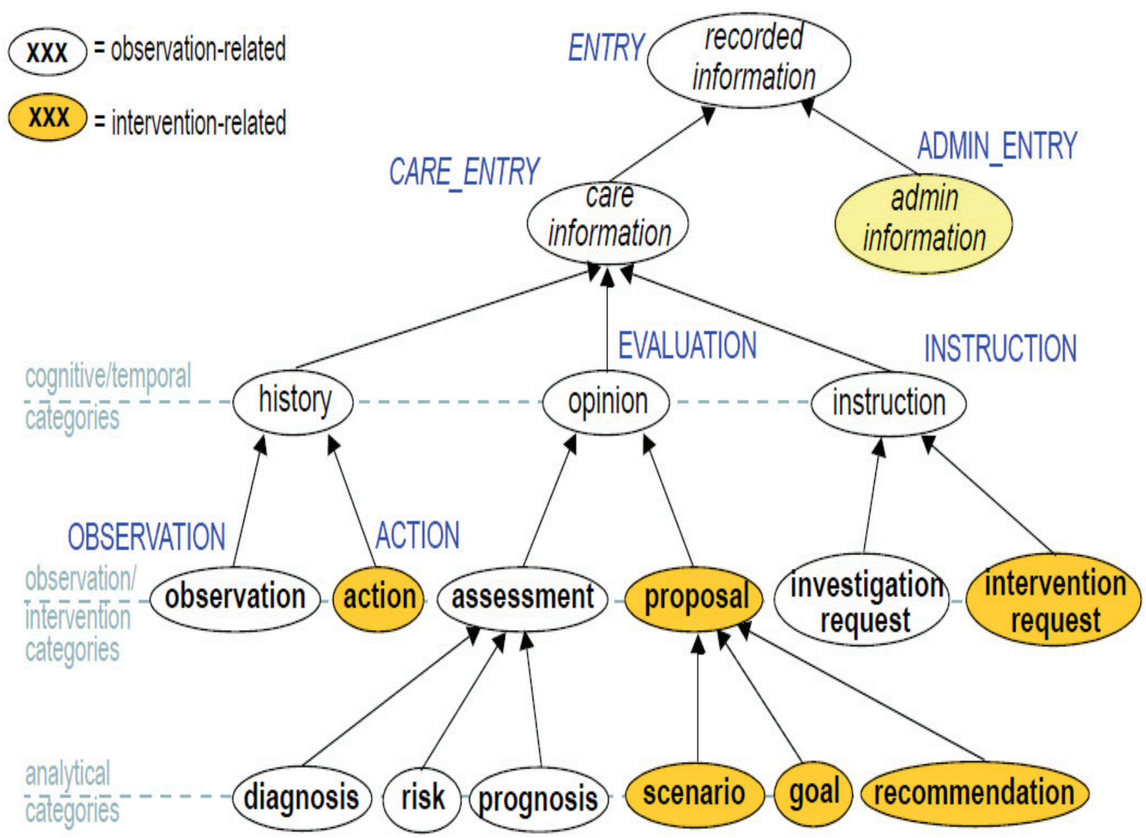


Figure 2.1 Clinical Investigator Recording (CIR) ontology, figure taken from [37].

As shown in Figure 2.1, the entities have been primarily classified into two broader categories: observation-related (entities involved to investigate and identify possible diagnosis) and intervention-related (entities involved in the intervention for any diagnosed medical issue). The ENTRY package of open-EHR has been devised from this CIR ontology with the following specializations.

ADMIN_ENTRY: open-EHR ENTRY package distinguishes the administrative entries from clinical entries. This class records all non-clinical information including admission information, appointments, discharge, billing, insurance information, etc.

CARE_ENTRY: This is the main concept to record all the longitudinal clinical actions of a patient. The specializations of this concept are discussed below.

OBSERVATION: This class is used to record the phenomenon or state of interest to do with the patient, e.g., pathology results, blood pressure readings, family history and social circumstances, physical examination, etc. [37]. This is similar to the 'Observation' class of HL7 RIM. This class relates to both 'Subjective' and 'Objective' concepts of SOAP structure used in POMR.

ACTION: This class is used to record what is actually done to the patients. It is different from OBSERVATION in the sense that OBSERVATION is used only to record information about the situation of a patient, not the actual action done to them.

EVALUATION: This class is used to record different assessments (e.g., diagnosis, risk, etc.) undertaken, as well as suggested care approach, planned target, etc.

INSTRUCTION: INSTRUCTION in *open*-EHR defines clinical actions and observations planned to be performed in the future. During the encounters, physicians generally keep notes on their planned actions which are satisfied by this class. This is also very crucial for long-term treatment plans. This class relates to the 'Plan' concept of SOAP structure used in POMR.

The Open-EHR model was developed to address the lack of representing semantics in healthcare information systems [35]. The designers of the model [35] realized the importance of having a knowledge-oriented system that included ontologies, terminologies and a semantically-enabled computing platform with patient-centric EHR to solve this problem. Various reusable archetypes of the Open-EHR model employ semantics of data from standardized terminologies, while the reference model ensures interoperability with standardized data representation. Moreover, Open-EHR reused the CIR ontology in its ENTRY package. Another important advantage of this model is its detailed implementation specifications. It proposed an open-source guideline of Java implementation (not yet complete) [35], a standardized query language called Archetype Definition Language (ADL) to retrieve information from archetypes, and a standardized specification for connecting with different terminologies.

Despite its advantages, the Open-EHR model has a number of limitations. The reference model defines the top-level generalized structure of medical information and is fully dependent in its controlled archetype repository. It is an ongoing research and development process to incorporate more archetypes and some practical scenarios which cannot be logically adjusted into current archetypes [33] (e.g., the episodic tracking of treatment). Moreover, the archetypes are designed by human experts and hence are error-prone. Although some efforts have been made to build OWL ontologies on Open-EHR archetypes with SWRL rules [71], the fullest benefits would not be achievable unless it involves a model-independent archetype design specification.

2.2.2 CEN EN 13606

CEN EN 13606, the European standard EMR, is based on the open-EHR archetype model. It is composed of five parts [33]: i) the RM defines the EHR information to be communicated, ii) the archetype interchange specification provides the generic model of information of archetype instances, iii) reference archetypes and term lists maintain the rules and associated data objects for EHR interaction, iv) security requirements and distribution rules specify the requirements and mechanisms of access rights of EHR components, and v) exchange models describe a set of models for service based or message based communication. The archetypes can be represented in a standard format called Archetype Definition Language (ADL), which is compatible with HL7 RMIMs and CMETs [33].

CEN EN 13606 is the result of a 15-year research and development project carried out by European scientists to develop a suitable EHR standard for European countries. This model is very similar to Open-EHR with two-level models. The biggest difference between these two models is that, unlike Open-EHR, CEN EN 13606 does not provide specifications for full EHR but only for EHR Extract. CEN EN 13606 also has archetype-related problems like Open-EHR (discussed in Section 2.1.1).

2.2.3 Problem Oriented Medical Record (POMR)

The POMR was proposed by Weed in 1969 [36] as a means of storing medical data in a structured way to ensure its ready accessibility. This was a problem-centric theoretical model supported by another structure called SOAP used to take progress notes. There are four main components involved in constructing such EMR structures: the problem list, database, initial plans, and progress notes. The problem list contains the titled and numbered list of problem headings, the status of the problem, and the date of the first entry of the problem. It may also include a short description about each problem with information such as symptoms, laboratory investigations, etc. Some socio-medico factors such as social problems, risk factors and psychiatric problems may also go under this list. The updates of problems go under the problem heading with the observation date. The database mainly contains information about previous clinical history. Although there are ongoing debates about the definition of the database, Weed suggested forming the database with the routine information that clinicians usually ask patients [38]. The initial plans reflect the initial goals in the practitioners' minds after observing the patient and incorporate diagnostic lists, information to be monitored, probable therapy and patient's education [38]. The progress notes (also known as follow-up notes) are captured in four sub-sections: Subjective (symptoms or absence of expected change), Objective (results of investigation), Assessment (notes based on the previous two sub-sections), and Plans (plans for further investigation/ medication) [39].

There have been some successful implementations of POMR [40, 41]. Weed initiated computerized POMR at the University of Vermont in 1969, and in 1976 developed a hypertext EMR system called Problem-Oriented Medical Information System (PROMIS) [40]. The system was implemented using a touch screen for data entry by clinicians. Although it was observed that PROMIS was less time consuming for clinicians than using a standard paper format, it was not widely accepted mainly due to its non-conventional method of data entry. PKC [41] proposed the clinical sections for a patient: Screening, Health Maintenance, Medical Problems – Active, Medical Problems – Inactive and Assets with possible sub-sections under each of these (e.g., the sub-sections for active medical problems were goal, basis, status, disability, follow courses, etc.). CPOMR supported both free text and coded data elements in data entry. In practice, the data elements in one hierarchical list may need to interact with those in another. This was not supported by the proposed CPOMR. The complete information model of their proposed CPOMR was not published by the authors.

Despite its simplicity, there are some shortcomings with the POMR model. The major limitation is that it considers only medical problems, and there is no way to link between different problems. It was also found that the SOAP structure for progress notes might be unnecessarily complex for simple problems.

2.2.4 ASTM-EHR

The American Society for Testing and Materials (ASTM) International provided an EHR standard [43] mainly to define the attributes necessary for the successful implementation of an EHR. They adopted the traditional POMR approach and classified the clinical data into eight main categories: Patient, Problem, Encounter (contains encounter and referral information), Practitioner, Order/Plan (i.e., request for a procedure/observation and care plan), Service Instance (e.g., medications, immunizations and procedures), Observation (e.g., screening information, lab results, physical examination) and Service Master

(master tables for ensuring controlled vocabulary of attributes). They proposed 119 essential data elements under these clinical entities.

Unlike Open-EHR and CEN EN-13606, ASTM is not just a top-level model for EHR but provides a detailed information model for an EHR with necessary attributes. This POMR-based standard is particularly popular for its comprehensive list of attributes for an EHR. Nevertheless, there are some important attributes missing in its specification (e.g., attributes for Goals). Also, it does not provide any specification about how this EHR can be used with standard medical vocabularies (e.g., LOINC, SNOMED-CT, etc.).

2.2.5 Summary

Currently available EMR standards are based on various structural perspectives while constructing the information models from different aspects. The information model of HL7 is the most widely used messaging standard. Among the EMR standards, HL7 is act-centric, Open-EHR and CEN EN 13606 information models are based on elements specified in archetypes, and POMR is problem-centric. Open-EHR and CEN EN 13606 have limitations in that they are fully dependent in their controlled archetype repository. It is an ongoing research and development process to incorporate more archetypes and some practical scenarios which cannot be logically adjusted into current archetypes [33] (e.g., the episodic tracking of treatment). Traditional POMR also has some limitations such as the lack of a suitable way to maintain the narrative notes of healthcare professionals, and linking and relating different problems. Moreover, the SOAP structure for progress notes might be unnecessarily complex for simple problems. Despite these and other drawbacks, a POMR-based information model can still adequately capture clinical data while ensuring the problem-centric orientation of clinical information that reflects the same procedures that clinicians usually follow in practice.

2.3 HEALTH LEVEL SEVEN (HL7)

HL7 (Health Level-7) [34] is an ANSI-accredited standard providing organization. In version 3, they incorporated the Reference Information Model (RIM), an object model with attributes, codes and vocabularies for representing the logical relationships among different entities involved in a clinical information domain and for specifying the complete life cycles of events carried by shared messages. As well, HL7 provides an XML-based messaging standard called Clinical Document Architecture (CDA) to specify the controlled architecture of contents in shared clinical documents. These CDA documents are both human and machine readable.

2.3.1 HL7 Reference Information Model (RIM)

In version 3, HL7 defined the logical relationships among healthcare entities in Reference Information Model (RIM). The complete lifecycle of healthcare events carried by HL7 messages are also defined by this object-oriented model. The latest version of RIM was published in September 2010 [50]. The core classes of this act-centric model, with associations among them, are shown in Figure 2.2.

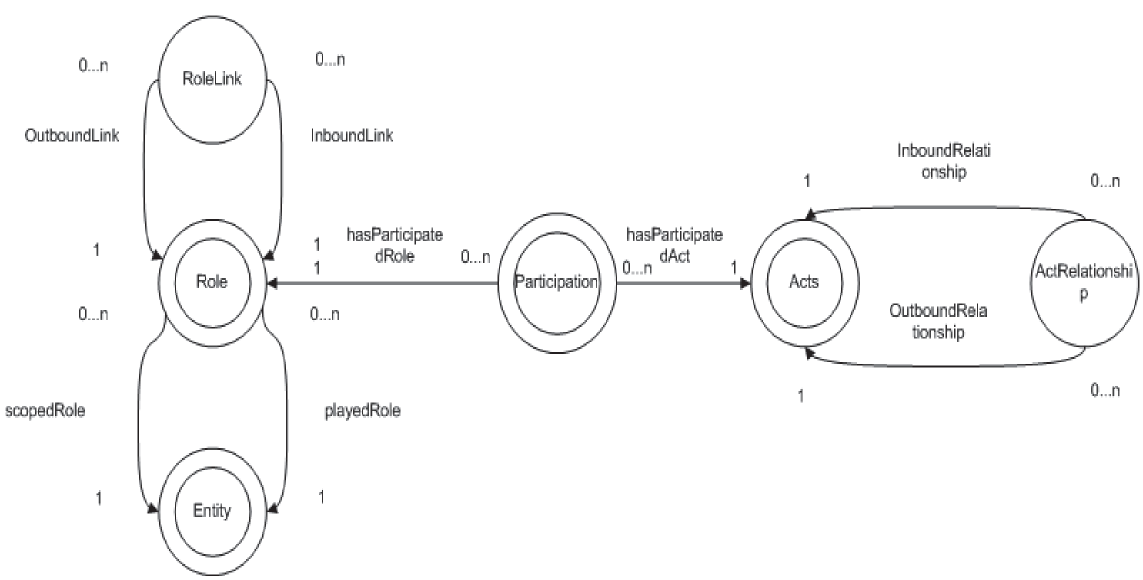


Figure 2.2 Core classes of HL7 Reference Information Model with associations.

The core class of this model is ‘Acts’ which specifies all actions and events of healthcare services. These ‘Acts’ can be related to other ‘Acts’ through ‘ActRelationship’ which defines the necessary information of such relation (e.g., type of relation, priority and sequence of these ‘Acts’, etc.). An ‘Entity’ is a particular physical thing (e.g., a person), a group of physical things (e.g., birds), or an organization. These entities play a certain role to perform an act. For example, a person plays the role of a general practitioner during the encounter with a patient. However, the same person plays the role of a patient if she undergoes surgery. ‘RoleLink’ is used in this model to represent the connection (e.g., dependency) between two roles. The association between an act and a role is defined by using the class ‘Participation’ (the kind of such association is represented by the attribute of ‘Participation’, typeCode).

HL7 RIM [50] has defined ‘Act’ as a record of an action that has already been performed, is to be performed, or is being performed. It has some specializations including classes for financial and insurance information (e.g., Account, FinancialContract, FinancialTransaction, InvoiceElement etc.), clinical actions (e.g., Observation, Procedure, PatientEncounter, SubstanceAdministration), or specifications of clinical documents (e.g., ContextStructure). Since, in this thesis, we are primarily interested in clinical activities, we are intentionally avoiding the financial classes in this discussion. ‘Observation’ is an important class in RIM that records any new information about a subject [n13]. Such information may include the diagnostic test results, physical examinations, history screening of a patient as well as the observation results for device or other entities (e.g., device temperature). ‘Procedure’ is a clinical act which is expected to bring an alteration of the physical condition of a subject [50]. This class represents traditional clinical procedures such as surgery, acupuncture, etc., as well as environmental alterations such as straightening rivers, etc. A subclass ‘Procedure’ is ‘SubstanceAdministration’, which defines the actions of medications and vaccinations. Interactions between the patient and physicians/care providers are recorded in the class ‘PatientEncounter’.

The attributes of HL7 RIM are defined for use with the data-types specified by HL7 itself, and also are bound to take values from the HL7 vocabulary domains. HL7 has its own data type standard as defined by the Normative Ballot version 3 of their data type on May 2009 [50]. HL7 showed two main reasons for defining its own standard and not adopting any existing standard [50]: namely, there are varieties of existing standards with differences in implementation technologies; these standards are not powerful enough to express many essential concepts (uncertainty, precision, physical quantities, etc.) of the healthcare domain.

HL7 data types are categorized into five broad groups [50]:

- Foundation: Provides core structural data types, e.g., boolean, collection of items (both ordered and unordered), and history items.
- Basic: Basic building blocks, e.g., text, coded concepts, identifiers, names and addresses.
- Quantities: Both simple numeric numbers (e.g., integers) and more complex quantities and ratios.
- Quantity Collections: Complex expressional facilities for sophisticated concepts, (e.g., timing specifications).
- Uncertainties: Extend the other four types of data to express uncertain values and outcomes.

HL7 defined its own vocabulary domain while using some external vocabulary standards (e.g., LOINC, SNOMED-CT etc.). The internal vocabularies are useful to bind coded values to some attributes. Three most important such attributes being used by Act, Role and Entity classes are classCode, moodCode or determinerCode and Code. classCode represents the exact class represented as a class in the RIM hierarchy [50], moodCode (in Act) or determinerCode (in Entity) differentiates between an instance and a kind of Act or Entity and also explains whether the instance is an occurrence or intent (in Act), Code is a specialization of classCode to further classify the exact class.

Researchers have been pointing out different logical shortcomings of HL7 RIM from ontological perspectives that sometimes prevent ensuring a comprehensive and sound knowledge representation of its concepts. A sound ontology should be capable of tracking the distinction between occurrents and continuants to ensure appropriate classification. Unfortunately, HL7 RIM cannot clearly distinguish between the two, which leads to logical inconsistencies and results in confusing representations of its ontological concepts. Both Vizenor [76] and Vizenor, Smith, Ceusters [77] reported this problem using particular examples. In HL7 RIM, a document (collection of information) is defined as an Act (a subclass of Acts [50]), not as an Entity, with different properties [76] such as persistence (the clinical document which remains unaltered for a time period) and stewardship (the document is maintained by a legitimate person or organization). Clearly, the persistence property represents continuants and the documents are not unfolded at different phases. Again, Acts are occurrents, not continuants, but is generalization of documents in RIM. Thus, the HL7 RIM could not differentiate between continuants and occurrents.

A comprehensive ontology should differentiate between universals and particulars. The Role class in HL7 RIM represents the roles that individual entities may play. One of the specializations of this is LicensedEntity, which is a licensed or qualified entity with the capability of performing specific functions and is scoped by an organization [50]. For example, the authority gives permission to a healthcare provider to provide healthcare services. Now, the instantiation of such roles are clearly examples of dependent particular continuants. Again, there are some roles which are universals (e.g., have generalizations) [77]. Thus, HL7 RIM failed to distinguish between universals and dependent particular continuants.

Vizenor, Smith, Ceusters [77] claimed that RIM gave more emphasis to the primary act than to the secondary act and failed to adequately differentiate between the two in all its parts. Both of these two acts are attributed to some physical entities and can differ even if both represent the same real-world scenario. For example, Mr. Smith documented that Dr. Philip observed the systolic value of blood pressure of patient Mr. Carlos was

125 mmHg. Here, the act of documentation carried out by Mr. Smith is the primary act and the observation by Dr. Philip is secondary act. In HL7, the documentation (primary act) is of more interest than what actually was observed (secondary act) [77]. Now, if another instance describing the same real-world scenario (observation) but different primary act is initiated, RIM would consider it as another act instance, which would cause confusion for the recipient [77].

Smith and Ceusters [78] pointed out another shortcoming of HL7 RIM from an ontological perspective. According to HL7's definition [7], an Act is a "record of something that is being done, has been done, can be done, or is intended or requested to be done." However, the examples of Act provided by them are: "(1) clinical observations, (2) assessments of health condition (such as problems and diagnoses), (3) healthcare goals, (4) treatment services (such as medication, surgery, physical and psychological therapy), (5) acts of assisting, monitoring or attending, (6) training and education services to patients and their next of kin, (7) notary services (such as advanced directives or living will), (8) editing and maintaining documents, and many others" [50]. This causes confusion as to whether an Act is an action itself or a record of the action.

2.3.2 Summary

Despite the logical inconsistencies of the HL7 Reference Information Model, it is the most widely used and the best choice for message communication in today's clinical information systems. Hence, we have chosen medical records written in HL7 to use in our thesis.

2.4 Chronic Disease Management Models

There have been some efforts to develop an information model for chronic disease management at different geographical locations. McGuire [74] proposed such an information model considering the business processes required and involved in chronic conditions. However, this model is more business-oriented than technology-based and thus lacks some important implementation details, such as standard coded values for data elements. There have also been initiatives taken at different geographical locations [75] (e.g., Germany, the Netherlands, Switzerland, Australia, British Columbia, etc.) to emphasize the information needs for that particular region. Western Health Infostructure Canada (WHIC) proposed a model for Canada [27] which is possibly the best choice, in that it provides implementation details with a complete information model (with entities, attributes, relationships, constraints), code tables and mappings with HL7. We will discuss this model in the next section.

2.4.1 The WHIC Chronic Disease Management Model

The Western Health Infostructure Canada (WHIC) proposed the Chronic Disease Management (CDM) Model [27] with data standards for chronic disease management and mapping of HL7 messaging standard with this data model. They also provided the implementation details of this infostructure within the participating jurisdictions of Canada (i.e., British Columbia, Alberta, Saskatchewan, and Manitoba) [27]. Although they [27] provided a general data model for chronic disease management, the clinical data elements were chosen to focus mainly on three chronic diseases (diabetes, hypertension, chronic kidney disease) chosen by the participating jurisdictions.

The WHIC-proposed detailed data model of CDM is shown in Figure 2.3. A person-focused, "Problem-Oriented Medical Record" approach was adopted for defining this model [50]. This provided necessary options to keep track of the problem-centric planned actions and clinical goals and also to follow up on executions of actions, a crucial aspect for successful management of chronic conditions.

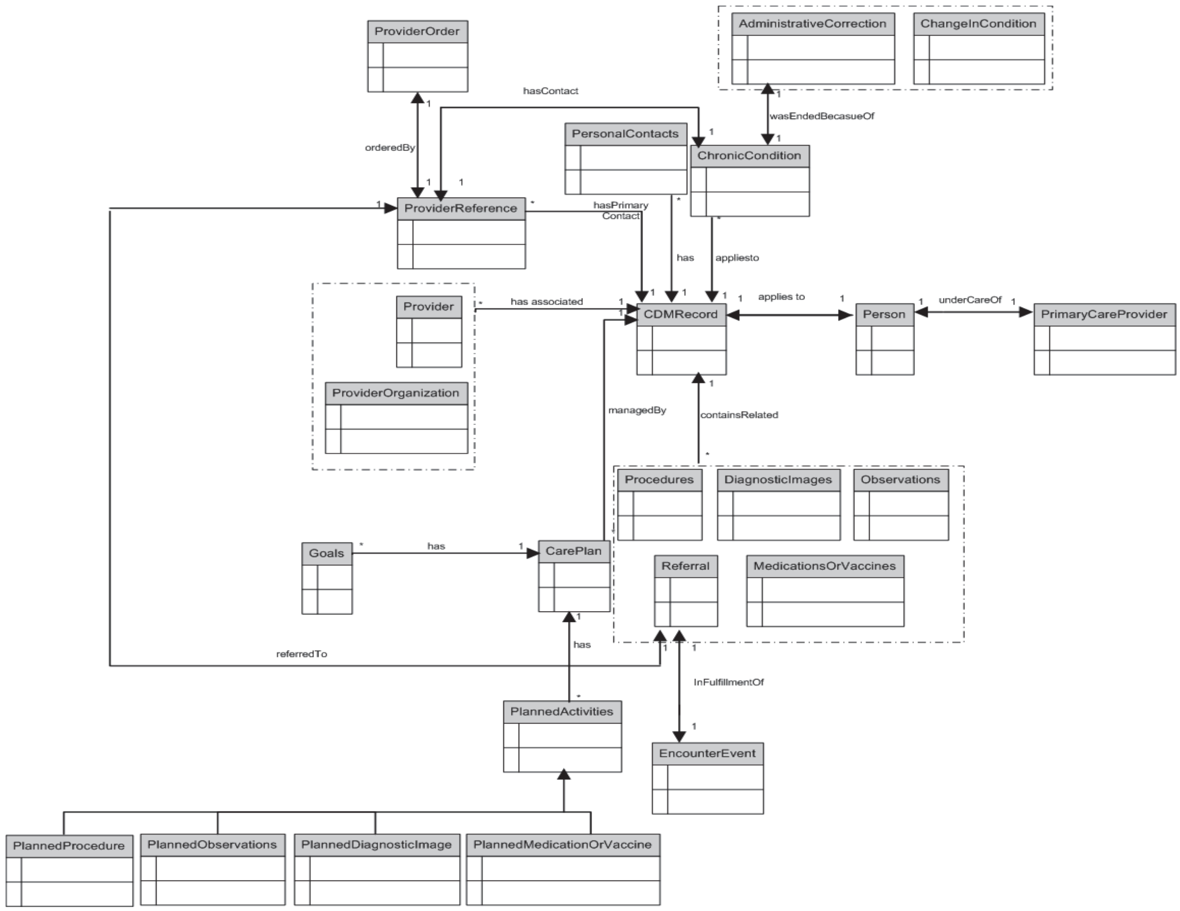


Figure 2.3 Information model of Chronic Disease Management Model.

The entities of this model are discussed below.

CDM Record: The CDM Record is used to contain information about all the chronic conditions of a person. It incorporates both person-centric information (e.g., Personal Contacts, Providers, etc.) and condition-centric information (e.g., Observations, Procedures, etc.).

Person: The Patient with at least one chronic condition.

Providers: The Providers can be either clinically qualified individual (e.g., physician, nurse, etc.) or other individual (e.g., parent) who participates in the delivery of care to the patient [50]. The 'Provider Type' attribute is used to distinguish between different types of providers participating in the chronic disease management.

Provider Organizations: A group of providers who acts as a single entity to the delivery of care to the patient.

Primary Care Provider: The provider who is clinically qualified and who is primarily responsible for delivering care to the patient.

Provider Reference: This is used to keep track of the accountability and responsibility of a clinically qualified person involved previously in the delivery of care to the patient.

Provider's Order: The request of the provider to perform particular clinical act (e.g., observations, procedures, referral, etc.).

Chronic Conditions: A diagnosis of the patient which is chronic and hence is the focus of the record.

Procedures: Procedures are such types of medical interventions whose outcomes are expected to make some desired changes in the patient's condition. Examples include education, skill development of the patient, surgery, etc.

Medications/Vaccines: The medications or vaccines a patient is receiving or has received [27].

Observations: This is a very important concept of the proposed model which provides the information about the findings related to a patient's condition. It incorporates all the subjective and objective findings, allergies, lab tests, diagnoses, medical conditions, assessments, etc. [27]

Diagnostic Images: Contains the results in the form of an image of a particular diagnostic procedure (e.g., ultrasound image).

Referral: The clinical intervention of referring a patient to a specialized clinically qualified person.

Encounter Event: This concept is used to keep track of encounters, particularly for referrals.

Care Plan: Describes the plans for the management of chronic conditions of a patient. Since chronic conditions are long-term by nature, this concept is particularly important to follow up the patient in future encounters.

Goals: A clinical target for the patient to be achieved by a target date (e.g., BMI<30 within the next two months).

Planned Activities: A care plan contains one or more planned activities. The planned activities are classified into four categories: planned procedures, planned observations, planned diagnostic images and planned medications/vaccines.

Planned Procedures: Procedures planned to be undertaken to improve the patient's conditions.

Planned Observations: Observations planned to be performed to assess the status of the patient's conditions.

Planned Diagnostic Images: The particular diagnostic imaging procedures which are planned to be performed as parts of the care plan.

Planned Medications/Vaccines: The medications/vaccines planned to be taken by the patient.

WHIC proposed its own vocabulary for chronic disease management while using six external vocabulary standards: HL7 (for demographical codes), LOINC (for diagnostic tests, physical exams, medical problems, procedures, medications, immunizations, etc.), ICD-10 (for some medical problems and physical exams), DSM-IV (for other medical problems), CCI (for some procedures), ATC/DDD (for some medications and immunizations). They also provided detailed implementations of mapping with HL7 and designed a RMIM for their proposed CDM.

2.4.2 Summary

The WHIC-proposed Chronic Disease Management Model [27] has been successfully implemented in some Canadian jurisdictions. These jurisdictions provided a detailed documented implementation guideline that incorporates the object-oriented entity-relationship model, their proposed vocabulary code tables for chronic diseases, and mapping of their model with HL7 RIM. Nevertheless, this POMR-based model takes into account only information pertinent to chronic disease management and does not consider data elements necessary for an EMR implementation. We believe that some entities essential for successful chronic disease management are missing in this model (e.g., Encounter-related information). We also argue that using six different vocabulary standards for this model is unnecessary, as the most robust vocabulary standard SNOMED-CT has almost all of the codes proposed by WHIC.

2.5 Ontologies on Electronic Medical Record

Bayegan et al. [44] proposed a process ontology for EMR which incorporated family-care workflow processes, clinical activities, different participants, and interactions of participants with a patient-record system. Their thesis (part of the NHS Clinical Headings project) was particularly interesting because it defined the minimal number of clinical headings necessary in a clinical setting, and also because their model was compatible with HL7. A similar effort was carried out by Scheuermann et al. [46], where they mainly focused on the disease and diagnosis manifestations. There are also some top-level ontologies in the literature (e.g., BFO [45], BIOTOP, etc.) which can be further customized and expanded. Inspired by all these initiatives, W3C proposed an OWL-DL Computer-based Patient Record (CPR) ontology [32], which is briefly described in the following section.

2.5.1 The Computer-based Patient Record (CPR) Ontology

W3C first started to develop a Problem-Oriented Medical Record Ontology in 2006. The goal was to define a minimal set of healthcare information terms while ontologically grounding HL7 RIM as a process model and using the criteria outlined in the traditional POMR structure [31]. This led to the Web Ontology Language (OWL)-based ontology in November 2009, called the Computer-based Patient Record (CPR) ontology [32]. Some parts of this ontology were taken from other top-level ontologies (e.g., BFO 1.1, BIOTOP, FMA, etc.) to ensure a sound and coherent means of necessary terminological representations required by an EMR. The core concepts of this ontology are shown in Figure 2.4.

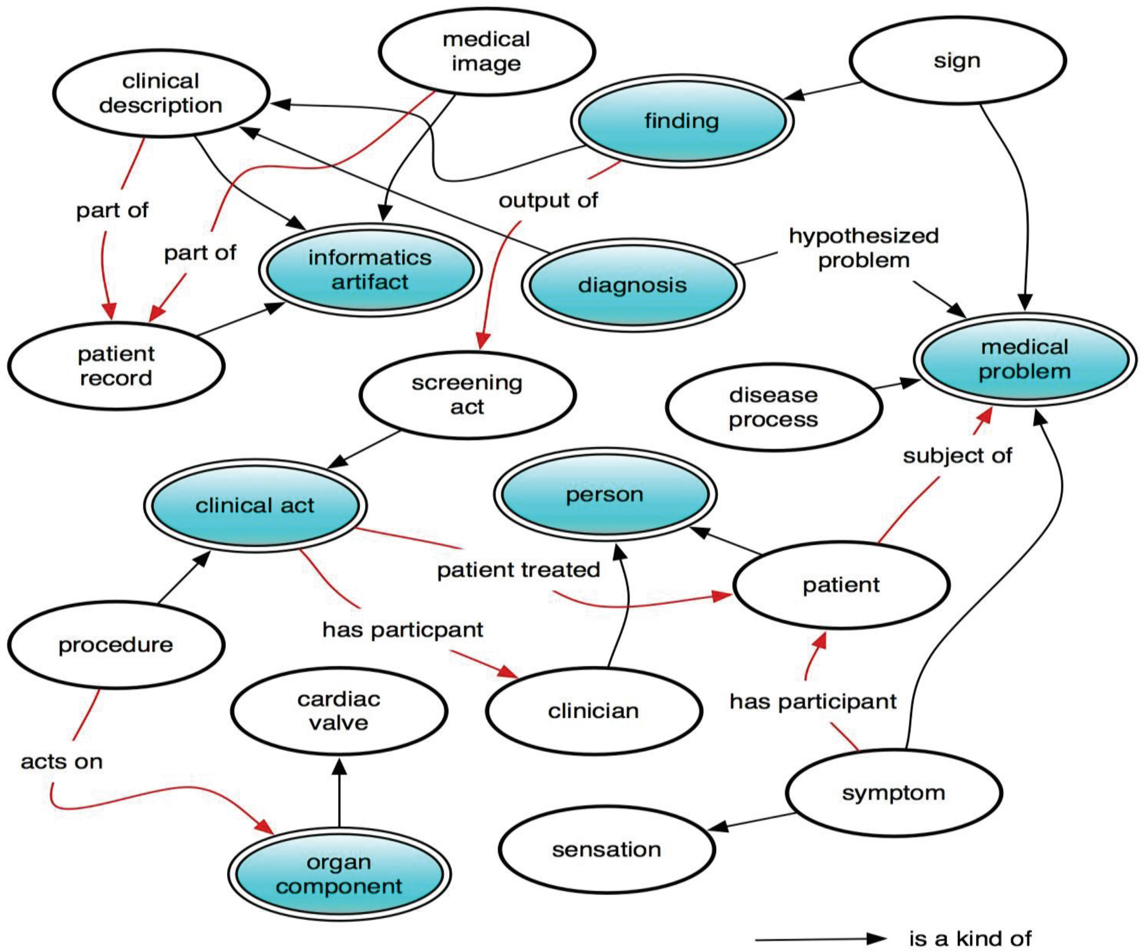


Figure 2.4 Main concepts of CPR ontology taken from [32].

The top-level concepts of the CPR archetypes are shaded and shown with double circles in Figure 4. These are described below:

Clinical Acts: The most important concept of CPR ontology is Clinical Acts, which is used to model various clinical tasks and activities and the information flow in these activities. This ontology used the process ontology of defining clinical processes as a workflow model proposed by Bayegan et al. [44] for defining the minimum clinical headings that are important for clinical communication and documentation. These clinical headings were put under the 'span:Process' class of BFO Ontology [45] to ensure proper classification of occurent and continuants data.

There are four specializations of Clinical Acts: Clinical Administration Act, Clinical Investigation Act, Procedure, and Therapeutic Act. A Clinical Administration Act is defined as any administrative act which is not itself investigatory or therapeutic and is done for either the assessment or treatment (e.g., patient appointment). A Clinical Investigation Act is used to discover the status, causes and mechanisms of a patient's health condition and is further classified into four classes: Clinical Analysis Act (used to generate the clinical hypothesis based on the condition of disease, physical examination, lab results, etc.), Diagnostic Procedure (the process of assessing the diagnosis; includes both laboratory or radiologic procedures), Laboratory Tests (the process of quantitative or qualitative test of a substance in laboratory), Screening Act (collecting data from different aspects (e.g., clinical examination, medical history, social history, family history, etc.) to identify problems). A Procedure is a type of act which is taken to improve the patient’s condition. This concept is used in this ontology to incorporate both diagnostic and therapeutic procedures and is aligned with the definition of Procedure in HL7 RIM. Therapeutic acts are activities which are taken to improve or maintain the physical condition of a patient. This incorporates medical therapy (e.g., surgery), physical therapy (e.g., exercise), and psychological therapy (e.g., request to read an article that will improve the patient’s psychological status).

Medical Problems: In this ontology, medical problems are defined as entities which incorporate the signs, symptoms and confirmed diseases of a patient. Signs are abnormalities interpreted by clinicians during physical examinations whereas symptoms are particular sensations reported by the patient themselves. The disease process has been defined as either pathological disease or etological agents while re-using the ontological framework for disease and diagnosis proposed by Scheuermann et al. [46].

Findings: Findings are clinical examinations done by a clinical expert during an encounter to assess the condition of patient’s body parts.

Diagnosis: Diagnosis is not confirmed but hypothesized medical problem recorded during clinical analysis acts.

Informatics Artifacts: Informatics artifacts represent the pertinent information stored in an EMR. It includes all the clinical artifacts encountered in a patient, digital entities (e.g., diagnostic images), and other longitudinal information (e.g., clinical findings, symptoms). This concept is used to distinguish between the records of an action and the actual action itself.

Person: A person can be either the patient him- or herself or the clinically qualified person (e.g., nurse, general practitioner, etc.).

Organ Components: Organ components are the anatomical and pathological entities which take part in different clinical procedures and screening acts.

The CPR ontology is engineered in Protégé using OWL-DL language. Although it has all the necessary concepts an EMR should have, it lacks the properties of these concepts and the implementation of vocabulary binding in this ontology.

2.5.2 Summary

The W3C proposed OWL-based Computer-based Patient Record (CPR) ontology [32] is a successful implementation of the concepts necessary for an EMR. It adopted the POMR approach to define the clinical headings of such an EMR. However, there appears to be an ongoing effort by its developers to define the properties of the concepts of this ontology. Also, the vocabulary should be bound to this ontology so that the EMR concepts can use coded values where necessary.

2.6 ONTOLOGY DEVELOPMENT METHODOLOGIES

An agreed standard methodology for ontology development is important to ensure re-usability of ontologies and interoperability among different applications. Moreover, proper guidelines on methods and techniques can ease such ontology development and can ensure successful implementations, engineering and maintenance of ontologies. These methodologies should clearly suggest all the activities required throughout a complete life cycle of ontology development, and each of these activities should be well supported by appropriate techniques or tools. These techniques and tools should be complied with the modern ontology languages (e.g., OWL) and environments (e.g., Protégé). There should also be possible guidelines to avoid common errors and to properly maintain built ontologies. These methodologies should be transparent and descriptive enough and no prior knowledge of the developer on ontology development should be assumed.

A number of standard methodologies for ontology development [79-90] developed since the early 80's. Most of these are based on experiences gained during developing particular ontologies. Some of these concentrate on building ontologies from scratch [84, 87] where others facilitate re-using existing ontologies. Interestingly, some efforts [87, 90] have been observed to build methodologies based on the standard software development processes. We will describe some of these ontology development methodologies in the following sections.

2.6.1 Cyc

The Cyc research project was initiated in 1984 with the aim to build an information system capable of reasoning general common-sense just like human beings [79]. The Cyc methodology was derived from the experiences gained during development of the Cyc Knowledge Base (KB), a formal representation of human knowledge used in everyday life. This KB is considered as an ontology as it can be used as a common base to build different intelligent systems and also as a common platform for their communications. A formal first-order predicate language, CycL, has been proposed to codify the Cyc KB.

The Cyc methodology of ontology development can be described by the three phases carried out during development of the Cyc KB:

Phase 1 – Manual Extraction of Knowledge: Explicit and implicit common-sense knowledge is manually coded in CycL and stored into the Cyc KB during the first phase of the Cyc methodology. The authors [79] claimed that manual extraction of knowledge is necessary because the current natural language and machine learning systems are not capable enough to extract such common-sense knowledge.

Phase 2 – Computer Aided Extraction of Knowledge: Since common-sense knowledge has already been stored into the Cyc KB in the first phase, the second phase proposes to codify that knowledge aided by some tools.

Phase 3 – Computer Managed Extraction of Knowledge: The third phase aims to extract common-sense knowledge automatically by using tools. A human effort is required only for the difficult parts of the text where such automatic extraction of knowledge is not possible.

The first two phases of this methodology have already been successfully implemented. Currently, the Cyc KB contains nearly five hundred thousand terms, about fifteen thousand types of relations, and about five million facts (assertions) relating these terms [80]. Cyc has also developed a number of tools for the computer-aided knowledge extraction [80], such as: WordNet Browser (allowing users to relate WordNet dictionary with the Cyc KB), HTML Browser (a hypertext way of browsing the Cyc KB), Lexicon Editor (for editing the Cyc lexicon) etc.

There also have been some efforts [80, 81] to implement the third phase, i.e., automatic extraction of common-sense knowledge. Taylor et. al. [81] proposed two classical machine learning approaches, the Naïve Bayes algorithm and the Support Vector Machine, for this purpose and obtained 98% precision and recall for a sample Cyc dataset. Recently, natural language systems have been applied to correctly parse many different sentence types, including ambiguous and syntactically complex inputs and for handling negation, modals, and nested quantifiers [80].

The Cyc methodology is possibly one of the oldest methodologies for ontology development, based on real-world experiences. The release of OpenCyc, an open source version of the Cyc technology, and of ResearchCyc, a version of Cyc aimed at the research community, by the Cycorp [80] is expected to fast the completion of its third phase. However, its ambition to code the encyclopedic knowledge in hand during the first phase is complex and error-prone. Moreover, the lack of documentation about the methodology is problematic from an ontological engineering point-of-view. The definition of a complete life cycle has been proposed by the Cycorp [80], which is believed to overcome this limitation. It is also expected that the life cycle will include the error avoidance and correction mechanisms, pre-development and post-development processes, which are absent in the current specifications.

2.6.2 Enterprise Ontology Methodology

Like the Cyc methodology, the Enterprise Ontology Methodology was also proposed from the real-world experiences gained during developing an ontology. Uschold and King [83] felt the necessity of such a methodology in 1995 while developing the Enterprise Ontology, an enterprise modeling process ontology, at Edinburgh. There are four main steps involved into this top-down approach:

Identify Purpose: Uschold and King [83] started their ontology development methodology with identifying the high level purposes of the ontology and the intended users of that ontology. They proposed a number of questions to be answered by the knowledge engineer and other development members before they actually start building the ontology [83], such as: whether the ontology will be re-used by a small group or by a larger community, whether the ontology will be used as means to structure a knowledge base or as part of a knowledge base, whether the ontology will specify a vocabulary or a meta-level specification of a logical theory etc. They suggested that the ontology developers should conduct a survey about these competency questions before the actual development.

Building the Ontology: This is the main step of ontology development – from knowledge capture to the codification of such knowledge while re-using existing ontologies. This step was broken down into three sub-steps:

- **Ontology Capture**: The ontology capture involved into identifying the possible concepts in the particular application domain, and the relationships among them. It was proposed to be done at the knowledge level, and thus was independent of any particular coding language. The particular tasks suggested to perform at this step were [83]: to identify the key concepts and relationships in the domain of interest, to define precisely and unambiguously these concepts and relationships, and to identify terms for each of these concepts and relationships.
- **Coding:** Once the knowledge is acquired, the next step was to represent such knowledge explicitly using a formal specification language (e.g., Prolog).
- **Integrating Existing Ontologies**: During either or both previous steps in building the ontology, existing ontologies should be reused. However, Uschold and King [83] indicated this as a complicated problem and failed to suggest some specific ways of how that could be done.

Evaluation: Uschold and King [83] defined this step as the definition of ontology evaluation proposed by Gomez-Perez [67], i.e., the technical judgment of the ontology along with its software environment and documentation against some formal requirements, competency questions, or real world scenarios.

Documentation: This step suggests detail documentation of the ontology development and management, including all assumptions made, the definitions of both concepts and primitives etc.

The successful project developed using this methodology was the Enterprise Ontology Project in 2000. Although this methodology provided a standard ontology development terminology, with a clear life cycle and structure, it fails to provide details about some important open-ended questions, such as: is it practical to specify the competency questions for the scope determination at the very beginning?, how to re-use the existing ontologies? etc. Also, this approach lacks the use of tools during the ontology development. However, considering the time the methodology was initiated, it was a great contribution, and continuous efforts could have been resulted into some better methodologies to serve the purpose of today's ontology engineering.

2.6.3 TOVE Methodology

Another ontology development methodology experienced while developing the TOronto Virtual Enterprise (TOVE) Project ontology was proposed by Gruniger and Fox [84] in 1995. This methodology concentrates on developing an ontology from the scratch and follows the informal specifications to the formal first-order logic representation of knowledge. This methodology consists of six steps:

Capture of Motivating Scenarios: These motivating scenarios were described as story problems or examples which are not adequately addressed by the existing ontologies. These scenarios could also incorporate the possible informal solutions to those scenario problems in terms of possible semantics for the objects and relations in that particular domain.

Formulation of Informal Competency Questions: During this step, the requirements of what the ontology should be are constructed in the form of competency questions. These questions are usually formed from the motivating scenarios and later are used to evaluate the ontology. The ontology, once developed, should be able to answer these questions using its terminology, axioms and definitions. This step is quite analogous to the first step in the Enterprise Ontology Methodology described previously.

Specification of the Terminology within a Formal Language: Once the informal competency questions are constructed, the informal set of terminologies can be easily derived from the terms used in those questions. A formal language (e.g., KIF) can then be used to specify these terminologies in formal way. These formal terminologies should be able to answer the informal competency questions.

Formulation of Formal Competency Questions using the Terminology: The formal set of terminologies is then used to derive the formal competency questions. The same language used to specify the formal terminologies can be used for this purpose.

Specification of Axioms and Terms within the Formal Language: The axioms (i.e., the definition of terms and constraints in their interpretation) are expressed as first-order sentences. Axioms are also considered to provide the semantics of these terms. If the set of axioms are not able to answer all of the formal competency questions, then additional axioms are suggested to be included until all those questions can be fully answered.

Establish Conditions for Characterizing the Completeness of the Ontology: The formal competency questions, axioms, and instances are used to define the conditions under which the solutions to those questions can be considered complete.

The TOVE methodology uses its own terminology, with a clear structure of ontology development. This methodology aims to build an ontology from the scratch and is highly dependent on the initial set of informal competency questions, which implies a possible bottleneck of it. Moreover, it does neither provide any guideline of how existing ontologies can be re-used during the development phases, nor provide detail documentation of the ontology development life cycle. Although Gruniger and Fox [84] indicated their efforts to develop tools for facilitating such methodology, they have not yet proposed any such tool.

2.6.4 KACTUS

The KACTUS methodology [85], resulted from the ESPRIT 8145 KACTUS project, aimed to investigate how knowledge can be re-used in complex technical systems and how ontologies can be used to support that. This methodology was proposed by Schreibur et. al. [85] and was conditioned by application development while re-using the existing application ontologies. They provided a browse-able library of available application ontologies. Every time an application is built, the corresponding ontology for that application in the library is refined. This methodology consists of three steps:

Specification of the Application: The context of the application is defined in this step. This includes the definition of terms, tasks to be performed by that application, and other components the application will try to model.

Preliminary Design: Based on the information defined in the previous step, this step searches for the existing ontologies in the library which can be re-used and extended. These ontologies are generally used for other applications. If no such ontologies are found in the library, a new ontology should be developed from the scratch.

Ontology Refinement and Structuring: The ontologies found in the previous step are customized, extended, refined and structured to meet the specifications defined in the first step. Alternatively an ontology is built from the scratch, if no match was found in the previous step.

This methodology is well supported by its own terminology and by the KACTUS toolkit. However, the steps involved into it are very abstract and not described properly from a technical point-of-view. Moreover, the existing KACTUS library contains frame-based ontologies where the expressiveness of ontologies are at the frame level. Thus, further works are necessary to take the advantages that a description-logic ontology has over a frame-based one.

2.6.5 SENSUS

The SENSUS methodology was proposed by Swartout et. al. [86] in 1996 with the aim to derive domain-specific ontologies from a large-scale ontology called SENSUS. The SENSUS ontology, a natural language based ontology, was resulted into after extracting and merging with different electronic knowledge sources such as: the PENMAN upper model, WordNet, ONTOS etc. These sources were further merged with English, Spanish and Japanese dictionaries using some merging tools. Currently, the SENSUS ontology contains [82] more than 50,000 concepts organized into a hierarchical level of abstraction. However, since the main objective of this ontology is to be used as a top-level ontology, it does not cover more detail terms specific to a particular domain. The SENSUS methodology [86] consists of five steps:

Identifying Seed Terms: The domain experts identify the seed terms – a series of terms necessary for the particular domain.

Linking Seed Terms to SENSUS: The domain experts then manually link these seed terms to the SENSUS ontology.

Incorporating Root Concepts: All the concepts that are in the path from the new terms to the root of the SENSUS ontology are then included into the resulting domain ontology.

Adding New Terms: The terms those are necessary but could not yet be covered were then manually added.

Adding Some Complete Sub-Trees: The sub-trees of the SENSUS ontology having many nodes already present in the resulting ontology are then added. This step is done manually based on the assumptions that if most of the nodes of a sub-tree are relevant to the domain, then the rest are also relevant. This step also incorporates adding additional domain terms necessary for such nodes.

Clearly, the SENSUS methodology employs a top-down tree-based approach of ontology development. This approach is interesting in the sense that it targets to have a common ontology to be customized for different domains, and thus can better facilitate ontology merging or alignment between two ontologies developed following the same methodology. However, the main limitation is the lack of automated tools and the total dependence on the manual integration by the domain expert. Moreover, like most of the methodologies discussed in the previous sections, SENSUS has not provided any documentations of its life cycle.

2.6.6 Methontology

Methontology, proposed by Fernandez et. al. [87] in 1997, is a well-structured ontology development methodology from scratch. This methodology was developed in the Laboratory of Artificial Intelligence at the Polytechnic University of Madrid while building an ontology in the domain of chemicals. It combines the ontology development activities based on both software development process and knowledge engineering methodologies, and facilitates the ontology development at the knowledge level. This approach includes the set of activities necessary for the ontology development, the particular techniques to perform each of these activities, a complete life cycle based on the prototypes. This methodology is well supported by the ODE tool (for single user), and its web version, WEB-ODE (for multiple users). The life cycle of this methodology has the following seven steps [87]:

Specification: The particular aim of this step is to produce either an informal, semi-formal or formal specification of ontology requirements using natural language, a set of intermediate representations, or a set of competency questions [87]. The specification should clearly specify the purpose (i.e., intended uses, scenarios), scope (i.e., set of terms with characteristics and granularity) and level of formality (e.g., semi-formal, rigorously formal etc.) of the ontology. A middle-out approach was proposed to construct the informal list of terms.

Knowledge Acquisition: The requirements can be gathered gradually (i.e., from more general to more detail level) and can be considered independent activity in the ontology development process. A few particular techniques have been proposed for this purpose [87]: Non structured interviews with experts (to construct a preliminary draft), Informal text analysis (to study the main concepts written in books), Formal text analysis (to identify the structures and specific components of an ontology), and Structured interviews with experts (to get specific knowledge about the ontology and its evaluation criterion).

Conceptualization: The domain knowledge is structured in a conceptual model during this step. To do so, a complete glossary of terms (i.e., concepts, instances, verbs and properties) is constructed first. The requirement document is used to construct the initial glossary of terms, and new terms can be added during any step of the ontology development life cycle. Once the glossary is guilt, a concept classification tree and a verb diagram are built. The concept classification trees are further classified into more specific sets [87]: Data Dictionaries (all the useful and potentially usable domain concepts, their meanings, attributes, instances, etc.), Tables of Instance Attributes (attributes and their values at the instance), Tables of Class Attributes (describe only the concepts), Tables of

Constants (specify the information related to the domain of knowledge that always take the same value), Tables of Instances (describe only the instances), and Attributes Classification Trees (a graphical representation of the attributes and constants, as well as the inference rules for them). Similarly, the verbs diagrams are also further divided into two sub-sets [87]: Verbs Dictionary (the declarative expression of meaning of verbs) and Table of Conditions (the set of conditions to be considered before or after performing an action). Finally, these intermediate representations are used to construct a table of formulas and a table of rules.

Integration: Re-using existing ontologies is investigated at this step by i) inspecting meta-ontologies which can be re-used for the intended domain, and ii) finding out the libraries which provide semantics of terms coherent with the terms of identified in the previous step.

Implementation: A suitable language (e.g., LOOM, Prolog etc.) is used to implement the final form of the ontology.

Evaluation: The evaluation of the final ontology is proposed to be carried out through verification (i.e., to judge the correctness of the ontology), and validation (i.e., the final ontology is capable to answer the competency questions).

Documentation: Finally the ontology development processes and activities are documented along with the source codes used to build the ontology.

Methontology [87] aims to reduce the gap between the 'ontology art' and 'ontology engineering' by providing a well-structured methodology, a detail and complete life cycle associated with the techniques to carry out different activities. Although it primarily concentrates on building an ontology from scratch, it also provides some basic technical details about how existing ontologies can be re-used. However, further technical details are necessary in re-using ontologies (e.g., how to plug-in the meta-ontologies with the glossary of terms), conceptualization (e.g., how the concept classification trees can be constructed—whether in UML or other in other standards). Furthermore, up-to-date tools are necessary to facilitate ontology development using current technologies (e.g., a Protégé plug-in).

2.6.7 Ontology Development 101

Noy and McGuiness from Stanford University [88] proposed this methodology in 2001 targeting the beginners in the ontology development field. They used a seven step approach of ontology development and showed an example of the Wine Ontology built in Protégé. These steps are discussed below.

Determine Domain and Scope: During this step, the scope of the ontology is defined, the intended users and maintainers of the ontology are considered, and the competency questions are articulated.

Consider Re-use: If any of the existing ontologies can be re-used, that should be considered for the ease of implementation and for obtaining better interoperability among applications.

Enumerate Important Terms: The possible terms and their properties are identified at this step.

Define Classes and Class Hierarchy: The terms identified in the previous step with independent existence are considered as classes of the ontology. The class hierarchy is defined according to either top-down, bottom-up, or middle-out approach, whichever best suited for the particular domain.

Define Properties: The remaining terms are used as properties. Noy and McGuiness [88] divided properties into four parts: intrinsic properties (i.e., essential or inherent property), extrinsic property (i.e., neither essential nor inherent), part-of property (both physical and abstract parts of an object), and relationships (relationships between individual members of a class and other items).

Define Restrictions: The restrictions on properties are defined in this step. The restrictions include the value type, allowed values, cardinality of a property, domain and range of a property etc.

Create Instances: The ontology is populated with instances in the final step.

This methodology aims to the beginners of ontology development, and thus is well supported by a complete life cycle with guidelines to avoid common errors. It also provides detail specifications on encoding ontology in OWL language using Protégé. However, this methodology concentrates only on the actual development of an ontology, and does not provide any details about testing (other than checking against the competency questions) and maintenance plans. Moreover, none of the real-world examples show evidence of success of applying this methodology.

2.6.8 On-To-Knowledge

Sure [89] proposed On-To-Knowledge (OTK) methodology in 2003 with the aim to develop and maintain ontology-based knowledge management applications into enterprises. Their methodology was supported by a plug-in, OntoKick, they developed for the OntoEdit. This plug-in was capable of extracting concepts, relations, and instances based on the requirements specified in the ontology requirements specification document (ORSD) format. The OTK methodology consists of five steps:

Feasibility Study: At first, both technical and economical feasibilities are studied, the domain and scope of the ontology are determined, and the involved people are identified.

Kickoff: The requirements are captured in an ontology requirements specification document (ORSD), and a semi-formal description is created.

Refinement: A prototype ontology is formalized based on the semi-formal requirements created in the previous step.

Evaluation: The usefulness of the generated ontology is evaluated against several methods, such as: the ontology engineer checks the fulfillment of the ORSD requirements and the competency questions, the applicability of the ontology in the application environment is tested based on the feedbacks from beta users.

Maintenance: The ontology is maintained and changes are made based on the feedbacks from users.

The most important aspect of the OTK methodology is the OntoKick plug-in, which semi-automatically extracts ontology components based on the requirements. However, currently the ontologies built according to this methodology are encoded into OIL language, and is supported by the OntoEdit tool. This results into a serious bottleneck from the interoperability point-of-view. Although OTK has provided a detail specification about their life-cycle, this methodology lacks examples of real-world applications.

2.6.9 UPON

Inspired by the software development life cycle approach, De Nicola et. al. [90] presented a use-case driven, incremental and iterative ontology development methodology in 2005 called UPON (Unified Process for ONtology building). They argued that although there are some differences between the software development and ontology development processes, the basic phases are the same. This methodology was based on the Unified Software Development Process where UML diagrams were used to design and evaluate the ontologies. UPON aimed to derive ontologies those can be used both by human beings and by automated systems.

After each cycle of the ontology development, the methodology results into a new version of the ontology. Each cycle consists of four phases:

Inception: During this phase, the requirements are gathered and analyzed conceptually.
Elaboration: A set of fundamental concepts are identified and loosely structured.
Construction: This phase concentrates on the actual implementation of the ontology.
Transition: The ontology is tested and released.

Each of these phases is further divided into iterations and each iteration consists of five workflows:

Requirements: This workflow determines the scope of the ontology and its purpose by using a storyboard and an application lexicon. A set of competency questions are constructed to evaluate the ontology in later steps, and use-cases are identified and prioritized.

Analysis: The existing ontologies are investigated to be re-used, the UML diagram for the application scenario is modeled and domain concepts are defined.

Design: This involves the taxonomic relations among concepts.

Implementation: The previously created UML diagram is used to formalize the ontology in an ontology language (e.g., OWL).

Test: The resultant ontology is evaluated against both the competency questions set during the Requirements workflow, and against the use-cases to test for the completeness of an ontology.

The UPON methodology is interesting in the sense that it employs the standard software development approach, and is thus expected to be adopted by the knowledge engineers and domain experts with previous experience of software development. Furthermore, this methodology has provided detail specifications about their life cycle with the implementation specifications for recent ontology development language, OWL. The major limitation is the lack of real-world applications based on this methodology. Although UML tools can partially fulfill the automation of this methodology, further tool supports are necessary (e.g., for the Test workflow).

2.6.10 Summary

We have reviewed a few ontology development methodologies in this section. The survey results indicate that despite a number of efforts being carried out, this is still comparatively a pre-mature field of research. There is no standard metric or evaluation criteria to measure the effectiveness of these methodologies. Ferenandez-Lopez and Gomez-Perez [82] investigated the maturity of some of these methodologies against the IEEE standard of software development and suggested that Methontology is the best possible methodology. However, this survey was done in 2002 and some recent methodologies (e.g., OTK, UPON etc.) have been evolved after it was conducted. A comparatively recent survey was conducted by Dahem and Hahn [91] in 2009, where they evaluated methodologies against their own defined criteria and found that the Ontology Development Methodology 101 fulfilled the most number of criteria (seven out of thirteen). However, both of these surveys concluded that further research is required to develop a more matured methodology with tools integrated into all phases of ontology development.

In our work, we need a methodology that facilitates re-using existing ontologies. And for this reason, we cannot use the methodologies which primarily focus on developing ontologies from scratch (e.g., TOVE and Methontology). KACTUS is inappropriate since it concentrates only on frame-based ontologies. Similarly, Enterprise Ontology Methodology, SENSUS, and OTK are not suitable as they lack tool supports (e.g., OTK has tool supports only for OIL language). We aim to derive a light-weight EMR ontology, and Cyc should not be considered for its unnecessary complexity. Although both UPON and Ontology Development 101 methodologies avoid these limitations, they lack evidences of successful real-world applications based on them. However, UPON provides the most adequate life cycle with some supporting tools among the existing methodologies.

2.7 Ontology Implementation Languages

A standard language for ontology implementation is mandatory to achieve semantic interoperability among information systems, and to ensure re-usability of ontologies. This language should model an ontology with sufficient expressiveness, should support users with well-defined syntax and semantics, and should possibly be reasoned automatically. Moreover, such a language should be both easily readable by humans and processable by machines.

This section does not pretend to present a state-of-the-art review on ontology implementation languages, rather primarily intends to illustrate the technical details of the Web Ontology Language (OWL) – the most powerful language proposed for this purpose [92]. Some W3C recommended previous standards (e.g., RDF, RDFS, OIL, DAML+OIL etc.) are also discussed here, because these highly influenced the design of OWL. Other classic ontology languages (e.g., ONTOLINGUA, KIF, OCML etc.) are intentionally avoided as these are mostly out-dated.

2.7.1 Resource Description Framework (RDF)

The Resource Description Framework (RDF) is a framework that represents web information by using the XML serialization syntax. It has a simple graph data model independent of any specific serialization syntax, a formal semantics with provable inference based on rules defined as RDF data, URI-based extensible vocabulary, data-types represented into XML schemas, and assertion capabilities of expressions. The vocabulary in RDF is specified by its vocabulary description language called RDF Schema (RDFS). RDFS is a semantic extension of RDF for providing formal description about groups of related resources and the relationships among them.

The major advantage of RDF is that its XML-based syntax is very user friendly. However, the major limitation identified in RDF and RDFS is their limited expressivity. For example, RDF uses only binary ground predicates, and RDFS is limited to a subclass hierarchy and a property hierarchy, with domain and range definitions of these properties [92]. There are some more technical limitations in RDF for knowledge representations, such as: disjointness, union, intersection of classes are not allowed, cardinality restrictions on properties are not permitted, special properties (e.g., transitive, inverse etc.) cannot be specified etc.

2.7.2 DAML+OIL

DAML+OIL is a joint effort to define a richer ontology implementation language resulted from the American proposal of DAML-ONT and the European language OIL. It is a description logic-based semantic mark-up language which extended RDF and RDFS with more powerful primitives. It represents concepts, taxonomies, binary relations, functions and instances, and avoids most of the technical limitations of RDF. Moreover, it allows automatic reasoning although the suitability of its reasoners is questionable.

2.7.3 Web Ontology Language (OWL)

OWL extends the XML syntax of RDF to provide more readability by humans and better accessibility of semantics by machines. RDF descriptions are re-used to define the instances, and primitives from RDF schemas are also re-used in OWL. OWL uses very similar primitives to DAML+OIL with more powerful expressiveness. Some of its important elements are discussed below.

Class Elements: Classes are defined as a owl:Class element with unique 'ID' attribute of RDF. Each class is considered as a sub-class of a predefined class, owl:Thing, and as a super-class of another predefined class, owl:Nothing. Two special elements, owl:disjointWith and owl:equivalentClass, have been defined to specify the disjointness and equivalence of classes.

Property Elements: Each property uses an object as domain and either another object or data-type value as range. There are two kinds of properties in OWL:

- **Data-Type Properties**: These properties relate objects to data-type values. There are a number of primitive data-types (e.g., String, Integer, Date etc.) already defined in OWL. User-defined data-types can also be provided through XML schema.
- **Object Properties**: These relate an object to another one.

owl:inverseOf and owl:equivalentProperty are used to specify inverse and equivalent properties respectively.

Property Restrictions: Different restrictions can be applied to the properties in OWL ontology, including their cardinalities, the values they can take etc. In general, the owl:Restriction element is used to indicate any of such restrictions and can contain any of the followings:

- owl: minCardinality and owl:maxCardinality: To specify the minimum and maximum cardinalities a property can have.
- owl:someValuesFrom: This specifies the existential quantifier of predicate logic, i.e., a property should take at least one value from the data range specified.
- owl:allValuesFrom: This specifies the universal quantifier of predicate logic, i.e., a property should take all values from the data range specified.
- Owl:hasValue: This restricts the property to take a specific value.

Special Properties: A number of special properties are proposed in OWL to fulfill corresponding binary relations:

- owl:TransitiveProperty: To define the transitive property (e.g., isGreaterThan).
- owl:SymmetricProperty: To define the symmetric property (e.g., isSiblingOf).
- owl:FunctionalProperty: Restricts a property to take at most one unique value (e.g., dateOfBirth).
- owl:InverseFunctionalProperty: Restricts a property where two different objects can not have the same value (e.g., isPatientIDOf).

Enumerations: An enumeration is a collection of values and is specified in OWL by using the element owl:oneOf.

Instances: RDF syntax is used to declare the instances in OWL.

Three Species of OWL: Three species of OWL language have been proposed with different levels of expressiveness, complexity and reasoning facilities:

- **OWL-Full**: The full OWL language with all the primitives is called OWL-Full. While this sub-language provides the highest level of expressiveness among its species, no automatic reasoning is possible because of its complexity. Moreover, it does not correspond to the Description Logic.
- **OWL-DL**: The OWL-DL has a number of restrictions on use of constructors, such as: a property cannot have some values from a data-type and other values from a class (vocabulary partitioning), the partitioning must be stated explicitly, a data-type property cannot be functional (property separation), transitive properties cannot be restricted on cardinality, anonymous classes can only be used as domain and range of equivalent and disjoint properties. Although this sub-language is less expressive than OWL-Full, it reduces the complexity and can be automatically reasoned.
- **OWL-Lite**: The OWL-Lite further restricts the OWL-DL, such as: the cardinality of a property can take maximum value of 1, a few elements cannot be used (e.g., owl:disjointWith, owl:one of, owl:hasValue etc.), anonymous classes can no longer be used by equivalent properties. Although this OWL sub-language further reduces the complexity and easier to grasp, it looses some important expressiveness capabilities.

2.7.4 Summary

Among all of the languages discussed in this section, OWL provides the most powerful expressivity with a robust syntax [92]. Moreover, its NLP based reasoners are able to perform more powerful reasoning. It has also provided three options to use based on the expressiveness and reasoning requirements. For all these reasons, OWL is the best option in our work. However, further works are necessary in OWL to handle exceptions.

2.8 Ontology Mapping Techniques

Mapping among different distributed heterogeneous ontologies is important not only to provide a common layer for accessing information from or exchanging information between different ontologies of related application domains (i.e., ontology alignment), but also to construct a resultant ontology from two ontologies (i.e., ontology merging). Research on ontology mapping is being carried out to develop automated, semi-automated and manual systems by using a number of different fields ranging from machine learning (e.g., CAIMAN, GLUE, etc.), concept lattices (e.g., FCA-Merge), linguistics (e.g., SMART, PROMPT, PROMPTDIFF, ONION, etc.), heuristics (e.g., PROMTDIFF, ConceTool, etc.). Below, we will discuss and compare various prominent frameworks, tools and methods of ontology mapping.

2.8.1 CAIMAN, GLUE

CAIMAN [58] is an ontology mapping method that uses machine learning techniques for text classification. It employs the probability measurement to determine which concept of the target ontology matches with a concept of source ontology.

GLUE [53], a semi-automated system of ontology mapping, uses multiple machine learning techniques for similarity measurement, each of which focuses on a different type of information (e.g., taxonomic structure, data instances, etc.). It uses three learners: a Content Learner that employs the Naïve Bayes' Theorem for text classification based on the content of instances, a Name Learner that is similar to the previous one but uses the full name of an instance rather than its content, and a Meta Learner that combines the prediction results of the previous two and gives weights to each learner based on how much it trusts the prediction. GLUE also uses another technique called relaxation labelling to label the nodes of a graph (that is, the output of this system) based on the features of their neighbourhoods, given a set of constraints.

CAIMAN [58] was proposed specifically to map concepts in different web documents using similarity measurement techniques of text mining. This technique was designed and evaluated for text documents mapping in information retrieval, with the authors claiming they were the first to do so. The similarity measurement technique of their ontology mapping methodology depended on linguistic similarities of concepts (i.e., terms in documents). Hence, this technique is not suitable for other ontology mapping tasks where linguistic similarities between concepts are less important (e.g., mapping between two EMR ontologies).

GLUE [53] has the key features in that it employs multiple similarity measurement techniques while incorporating multiple learning strategies for ontology mapping. Nevertheless, this mapping method considered only concept-level mapping for two given ontologies and not at other levels (e.g., attributes, relationships, constraints, etc.). Moreover, we believe that it could be further extended to incorporate logics with machine learning techniques for better mapping results.

2.8.2 MAFRA

MAFRA (Ontology Mapping FRAmework for distributed ontologies in the semantic web) [54] is a distributed ontology mapping framework to automatically detect the similarity between two ontologies based on a technique called Semantic Bridge. This framework consists of five horizontal and four vertical modules. The horizontal modules are as follows:

i) Lift & Normalization: First, the two input ontologies are normalized to a uniform representation (e.g., RDF) to avoid the possible syntax differences and make semantic differences more clear. This is done by a tool called LIFT that converts the DTDs, XML schemas and relational databases to the structural level of an ontology.

ii) Similarity Discovery: At this module, the similarities between source and target ontologies are discovered.

iii) Semantic Bridging: The mapping between the source and target ontologies is defined at this module, which consists of two steps: concept bridge (translates source instances into target ones) and property bridge (translates the source properties into target ones).

iv) Execution: Once all the mapping rules are defined, it populates an ontology of mapping constructs called Semantic Bridge Ontology in DAML+OIL format.

v) Post-processing: After the execution is done, the results are checked and possibly can improve the quality of ontology mapping.

The four vertical modules are as follows:

i) Evolution: Ensures the synchronization of the semantic bridges with the changes made at either the source or target ontology.

ii) Cooperative Consensus Building: Harmonizes between the two parties in the mapping process.

iii) Domain Constraints and Background Knowledge: Uses WordNet or another domain specific thesaurus to improve similarity measures between two ontologies.

iv) Graphical User Interface (GUI): The GUI is used to show the similarity measures and transformation results.

The semantic bridge is the core component of ontology mapping in MAFRA. However, it considers only the taxonomies of two ontologies for such mapping, not any procedural mechanism. Moreover, this semantic bridge can only be represented in the DAML+OIL language and not in any other languages (e.g., OWL).

2.8.3 SMART, PROMPT, PROMPTDIFF

Noy and Musen [55-57] developed a series of semi-automated ontology alignment and merging tools based on linguistic similarity matches between concepts. They started with SMART in 1999, which not only matches two ontologies based on class names but also investigates linguistic similarities in common suffixes, common prefixes, shared substrings, slot names, slot value types, structures of merged concepts, etc. It also suggests possible similarities to users, identifies conflicts, and suggests possible solutions to resolve those conflicts.

PROMPT [56], another semi-automated ontology alignment and merging tool, goes beyond what SMART provides in that, while providing linguistic similarities, it generates a list of suggestions for the engineer based on linguistic and structural knowledge. Wherever an automated decision is not possible, it asks for the input of the user, points out conflicts, and provides suggestions to resolve them.

Noy and Musen [57] developed their latest tool, PROMPTDIFF, to compare structures of various versions of an ontology using different heuristic matchers. It applies these heuristic matchers chronologically, i.e., the output of one matcher as an input of another until it produces some form of change.

SMART [55] facilitates both ontology merging and alignment, iteratively takes inputs from the human expert, and automatically provides suggestions for possible matchings based on those inputs. The biggest problem of this method is that it provides the initial list of possible matches solely based on linguistic similarities of concepts. Moreover, the SMART tool works only for frame-based ontologies and not for description logic ones. PROMPT [56] incorporates structural knowledge in addition to linguistic knowledge to provide the intial suggestions. PROMTDIFF [57] further enhanced such similarity measurements by incorporating heuristic matchers. However, both of these tools expect only frame-based ontologies as inputs and do not work description logic ontologies.

2.8.4 FCA-Merge

FCA-Merge [59], another semi-automated ontology merging method, is based on formal concept analysis, lattice exploration, instances of source, and target ontologies. The overall method works in three steps:

i) Extraction of instances and generation of the formal context for both source and target ontologies. A lattice of concepts is derived by using natural language techniques.

ii) The next step is to compute the pruned concept lattice by using an algorithm called TITANIC.

iii) The final step is a semi-automated one, asking assistance from the domain engineer for resolving possible conflicts while providing an automatic support in a Q & A manner.

FCA-Merge [59] was proposed only for merging two ontologies, not for ontology alignment. It incorporated the NLP techniques and employed a bottom-up approach of such ontology merging. However, this technique is particularly suitable for merging text documents, since it expects common instances in the two given ontologies and the pruned concept lattice is constructed on the basis of the concept frequencies.

2.8.5 IF-Map

Kalfoglou and Schorlemmer [60] proposed an automated method, IF-Map, for mapping multiple ontologies. This technique consists of four steps:

i) Ontology Harvesting: Acquisition of existing ontologies from different sources.

ii) Translation: Automatically translating source ontologies into horn logics.

iii) Informorphism Generation: This is the main step in the mapping procedure where the program automatically detects any logic informorphism between the two ontologies.

iv) Display of Results: Displays the mapping results in RDF format.

IF-Map [60] is particularly interesting in that it was targeted to map ontologies automatically while employing Horn Logic in its ontology mapping procedure. Also, since it produces the final mapping results in RDF format, it can easily be used to map description logic ontologies. Nevertheless, it requires a reference ontology for such mapping to reduce the interoperability problem, which does not always exist for different domain ontologies.

2.8.6 ONION

Mitra and Wiederhold [62] proposed a semi-automated tool called ONtology compositION system (ONION). The objective of their proposed tool was to resolve heterogeneity between different ontologies using articulation rules for mappings. These articulation rules were constructed manually. The linguistic matcher was then used to find out all the possible pairs of concepts in two ontologies and to assign a similarity score to each matching pair. The pairs having similarity scores higher than a certain threshold value (set by a domain expert) were used to construct the articulation rules. Further, an inference-based matcher indentified matches based on these rules and any custom rules set by the domain expert. A GUI was used to show the matches to the domain expert where (s)he could add, delete, or modify any suggested match. The authors [62] claimed that full automation of this system was not possible due to lack of adequate natural language processing technology.

ONION [62] matches concepts between ontologies based on both linguistic similarities of concepts and heuristic methods (i.e., matching rules). This method works fine for specialized ontologies with controlled vocabulary (e.g., two medical ontologies with specific disease names) but fails to produce acceptable results for ontologies with more generalized vocabulary.

2.8.7 ConcepTool

ConcepTool is another semi-automated tool developed by Compatangelo and Miesel [63]. They [63] used a description logic reasoner and represented the two ontologies as two enhanced entity-relationship models. Linguistic and heuristic inferences were used to map the attributes of these two models. The analyst is prompted to resolve conflicts (if any occurred).

ConcepTool [63] is similar to MAFRA [54] in that both use a semantic bridge for ontology mapping. However, ConcepTool uses both linguistic and heuristic inferences with a description logic reasoner and hence can be used for description logic ontologies. It was reported that this method does not work well for complex constraints in any given ontology.

2.8.8 Comparison among various ontology mapping techniques

There is not yet any standard metric to measure the effectiveness and suitability of an ontology mapping method or tool [52]. However, we compared them by using the metrics proposed by Choi et al. [52] with some additional metrics of our own (e.g., level of automation). Table 1 summarizes the results of this comparison. Some of these tools have been proposed with the aim of ontology merging (e.g., SMART, PROMPT, FCA-Merge, etc.) whereas others aim to map between two ontologies (e.g., MAFRA, ConcepTool, etc.). While most of these tools take two ontologies as input, some (e.g., ONION) take terms of two ontologies as input. As shown in Table 2.1, there is a wide range of technologies being used by these tools. For instance, CAIMAN, GLUE use machine-learning techniques, FCA-Merge use concept lattices, SMART, PROMPT, PROMPTDIFF, ONION use linguistics similarities, and PROMTDIFF, ConceTool use heuristics. All of these tools (except IF-Map) require assistance from a human expert in generating the final mapping and merging results. These tools need one or more structured, instance-based, lexical and domain knowledge data.

Table 2.1 Comparison among different ontology mapping tools (parts of this have been taken from [51]).

Ontology Mapping Tool	Input	Output	User interaction	Mapping strategy or algorithm	Level of autom ation	Str uct ure d kno wle dge	Ins tan ce-bas ed Kno wle dge	Lex ical kno wle dge	Do mai n kno wle dge
CAIMAN, GLUE	Two taxonomies with their data instances in ontologies	A set of pairs of similar concepts	User-defined mappings for training data , similarity measure, setting up the learner weight, and analyzing system's match suggestion	Multi-strategy learning approach (machine learning technique)	Semi-automatic	No	Yes	Yes	Yes
MAFRA	Two ontologies	Mappings of two ontologies by the Semantic bridge ontology	The domain expert interface with the similarity and semantic bridging modules and it has graphical user interface	Semantic Bridge	Semi-automatic	Yes	Yes	Yes	Yes
SMART, PROMPT, PROMPTDI FF	Two input ontolog ies	A merged ontology	The user accepts, rejects, or adjusts system's suggestions.	Heuristic-based analyzer	Semi-automatic	Yes	No	Yes	No
FCA-Merge	Two input ontologies and a set of documents of concepts in ontologies	A merged ontology	Generating a merged ontology requires human Interaction of the domain expert with background knowledge	Linguistic analysis & TITANIC algorithm for Computation for pruned concept lattice	Semi-automatic	Yes	Yes	Yes	Yes
IF-Map	Two ontologies	A merged ontology	Results are shown in RDF format. A Java front-end is also provided that can be accessed from the Web	Identifies logic informorphis m between two ontologies	Automati c	Yes	No	Yes	No
ONION	Terms in two ontologies	Sets of Articulation rules between two ontologies	A human expert chooses or deletes or modifies suggested matches using a GUI tools	Linguistic matcher, inference-based Heuristics	Semi-automatic	Yes	No	Yes	Yes

Ontology Mapping Tool	Input	Output	User interaction	Mapping strategy or algorithm	Level of automation	Structured knowledge	Instance-based Knowledge	Lexical knowledge	Domain knowledge
ConcepTool	Two ontologies	Alignment of these two ontologies	An enhanced entity-relationship model	Description logic reasoned, linguistic and heuristic inferences	Semi-automatic	Yes	No	Yes	Yes

2.8.9 Summary

The selection of a suitable ontology mapping tool or method should depend on the purpose of such mapping and the domain of the ontologies to be mapped. None of these tools can universally serve the purpose of ontology mapping. The major problem with most of the methods using natural language processing or machine learning techniques is that these techniques rely on linguistics rather than semantics (e.g., the same concept can be used in two ontologies with the name of 'Patient' and 'Person').

In this thesis, we are particularly interested in mapping between different medical domain ontologies (e.g., CDM and POMR ontologies). Moreover, we need a tool that allows such mappings for ontologies in OWL-DL languages. We cannot use CAIMAN for this purpose, as it depends on linguistic similarities of two ontologies. Likewise, SMART cannot be used for the same reason, and MAFRA is not suitable for our purpose as it considers only the taxonomies of two ontologies and accepts ontologies only in DAML+OIL form. In addition, SMART, PROMT, PROMTDIFF work only for frame-based ontologies and hence are not suitable, nor is FCA-Merge, since it expects common instances in the two given ontologies and is suitable only for text documents. IF-Map can also not be used for our purpose since our target is to map different application ontologies, and we do not have a canonical ontology for the two ontologies to be mapped.

Finally, since we are dealing with mapping less specific specialized ontologies, ONION cannot be used for our purpose.

For the concept-level mapping of such ontologies (e.g., mapping between ASTM-EHR and CPR ontology concepts), we can use GLUE and ConcepTool. However, we need to provide accurate matching rules for such mapping.

2.9 SUMMARY

In this chapter, we first discussed some of the Electronic Medical Record standards, namely, Open-EHR, CEN EN 13606, ASTM-EHR and Problem-Oriented Medical Record. We have critically examined each of these while keeping in mind the purpose for which we need an EMR standard. We found that although all of the afore-mentioned standards have some shortcomings, POMR would be the best choice for our methodology. We have also discussed the HL7 Reference Information Model and the logical inconsistencies reported about it in the literature.

Next, we discussed the information models available for chronic disease management while giving special emphasis to the WHIC-proposed CDM model. We found an ongoing effort to develop coherent ontology for EMR. An OWL-DL ontology for the W3C-proposed CPR was found to be of particular interest for our purpose and thus was briefly described.

Different methodologies for ontology development were described and analyzed for the usefulness in our work. A technical overview of the ontology language OWL has been presented. Various ontology mapping techniques, tools and methods were discussed and critically compared. We have also tried to draw a conclusion regarding which of these could be used in our methodology.

CHAPTER 3 METHODOLOGY

3.1 Our Approach

Our purpose is to build an ontology-based EMR which focuses on chronic disease management while providing a coherent information structure to support other acute diseases and co-morbidities. This EMR should be patient-centric by nature and should hold the longitudinal information of patients. It should also facilitate coded data entry by using standard clinical vocabulary and be mapped with HL7 RIM to ensure that the clinical messages in HL7 can be fully captured by this ontology. The proposed methodology is shown in Figure 3.1.

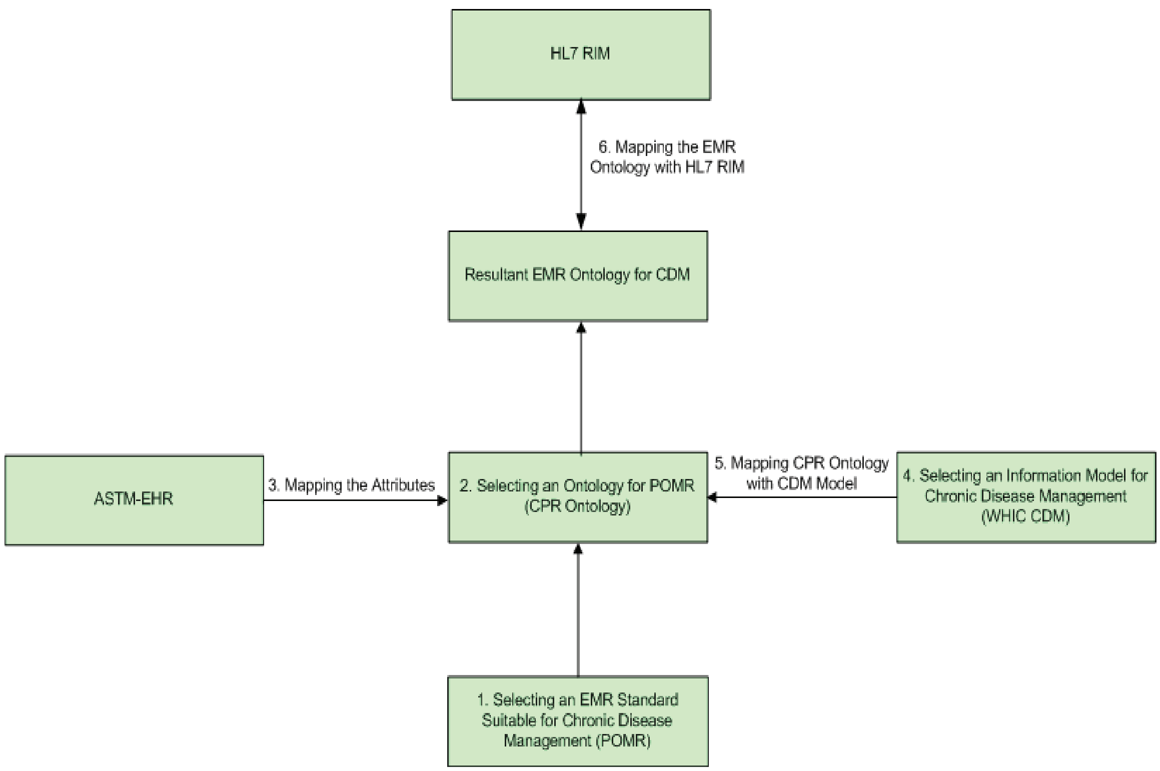


Figure 3.1 Proposed methodology of Ontology-based EMR for Chronic Disease Management.

The first step of our methodology was to choose an appropriate EMR standard for chronic disease management. We have decided not to use Open-EHR and CEN EN-13606, as these reference models define generalized structures of medical information and are fully dependent in controlled archetype repository. Instead, POMR was chosen as the EMR standard to be used. The rationale of choosing POMR is that it captures and stores the clinical information in a problem-oriented way, which is best suited for chronic disease management.

The next step was to build an ontology on POMR. We reused the CPR ontology [32] discussed in Chapter 2. The CPR ontology is based on a POMR standard and thus could be easily used. However, as discussed earlier, this ontology lacks the necessary data-type properties for holding data by its concepts. We created these properties in the CPR ontology by incorporating the attributes from equivalent concepts of the ASTM-EHR model [43]. We mapped the concepts between the CPR ontology and ASTM-EHR for this purpose.

A suitable chronic disease management information model was required to ensure that the information elements necessary particularly for successful chronic disease managenet were well supported by the resultant EMR ontology. We found the WHIC-proposed Chronic Disease Management Model [27] appropriate for this purpose. This POMR-based model provides a detailed information model with concepts, attributes, relations among concepts, constraints (i.e., property restrictions) and coded data elements pertinent to successful chronic disease management. In addition, implementation details for mapping HL7 messages into their proposed CDM model were also provided.

Next, we mapped the WHIC proposed Chronic Disease Management Model [27] and the CPR ontology to ensure that the concepts and properties necessary for chronic disease management are well supported by the resultant ontology, which we call the EMR ontology. We converted the vocabulary proposed in the CDM model into SNOMED-CT since it provides a robust and powerful vocabulary in the clinical domain.

Finally, we mapped the resultant EMR ontology with the HL7 RIM to ensure that the clinical documents in HL7 can be completely captured by the EMR ontology.

It is worth mentioning that the underlying model of the ASTM-EHR model, the CPR ontology and the CDM model used in these mapping procedures, is POMR.

3.2 Mapping between the ASTM-EHR and the CPR Ontology

3.2.1 Problem Specification

As discussed in the previous section, the CPR ontology does not have the data-type properties for concepts in it to hold the medical information. This results into a serious bottleneck of the ontology in proper classifications and reasoning. To overcome this limitation, we have decided to map the concepts of the ASTM-EHR model [43] to the equivalent concepts of the CPR ontology. This mapping problem can be specified more formally as:

Given two sets of concepts, $C_1 = \{c_{11}, c_{12}, c_{13}, \ldots\ldots\ldots, c_{1m}\}$ and $C_2 = \{ c_{21}, c_{22}, c_{23}, \ldots\ldots., c_{2n}\}$, on the same application domain, we have to derive a set of semantic relations between these two such that, $c_{1i} \Re c_2 \middle| \{\{c_{1i} \in C_1\} \wedge \{c_2 \subset C_2\}\}$, where, C_1 defines the concepts (i.e., domains for the relations) in the ASTM-EHR model and C_2 defines those (i.e., co-domains for the relations) in the CPR ontology.

Since these relations are equivalent, clearly, these are reflexive, i.e., should map among semantically equivalent concepts from two sets. This implies to $c_{1i} \Re c_2 \equiv \{c_{1i} = c_{2j} \middle| \{\{c_{2j} \in c_2\} \wedge \{c_2 \subset C_2\}\}\}$. Since the relations are equivalent, these are also symmetric and transitive.

Once the mappings among the concepts are done, the next step will be to fit the attributes from the ASTM-EHR model to the corresponding data-type properties in equivalent concept(s) of the CPR ontology. This procedure is based on the assumption defined below (the definition of formal context is taken from [59]).

Assumption: Given two formal contexts $\mathrm{K} := (C, A)$ and $\mathrm{K}' := (C', A')$, where, C and C' are sets of concepts, A and A' are sets of attributes, and I and I' are binary relations between C and A, and C' and A' respectively. If two concepts, c and c' from the two sets, are semantically equivalent and has two sets of attributes, a_1 and a_1' respectively, then, there exists a finite set of attributes, a common both in c and c'.

For example, if there are two semantically equivalent concepts ‘Person’ and ‘Patient’ in two formal contexts, then, there exists a finite common set of attributes (e.g., PersonID, Name, Address etc.).

This assumption leads to a simple task of attribute fitting as data-type properties into the concepts of the CPR ontology from the equivalent concepts of the ASTM-HER model, once the concept-level mapping is done.

3.2.2 Mapping Procedure between the ASTM-EHR and the CPR Ontology

Before we discuss the actual procedure taken to do the concept-level mapping between these two information models, we are going to discuss the similarities and the mismatches between them. This will help us identifying the research challenges and thus better designing the methodology for such mapping.

Similarities between the ASTM-EHR and the CPR Ontology: Both ASTM-EHR and CPR ontology provide information models for an EHR. Another important similarity from a technical viewpoint is that both use the POMR model as their underlying model, and thus both share semantically equivalent concepts.

Mismatches between the ASTM-EHR and the CPR Ontology: The biggest difference between these two is that the CPR ontology provides an ontological representation for the EHR whereas the ASTM-EHR aims to provide the model for the structure and content of an EHR. Moreover, the ASTM-EHR provides a complete model for an EHR, but, the CPR ontology lacks the data-type properties in its concepts (as discussed before).

ASTM International defines 119 core attributes organized under fourteen concepts (these are defined as 'entity segments' in the ASTM-EHR specification [43]), shown in the following sub-section, which are necessary for any EHR information model [43]. These fourteen concepts are grouped under the eight categories discussed in section 2.2.4 (e.g., the category 'Service Instance' contains concepts 'Medications', 'Immunizations' and 'Procedures'). In order to ensure proper granularity of information, we are going to consider the fourteen concepts of the ASTM-EHR, rather than its eight main categories.

Since the ASTM-EHR is not an ontology, we cannot directly apply any of the ontology mapping techniques/tools discussed in Chapter 2 as all of those take ontologies as inputs for mapping. We could build an ontology for the ASTM-EHR model and then could apply one or more of those techniques, which would have been more time consuming and impractical for these limited number of concepts. Rather, we have decided to manually derive the relations between the concepts of these two models. For each and every concept in the ASTM-EHR model, we chose the semantically equivalent concepts in the CPR ontology. We believe that applying the domain knowledge in such a way would also result more accurate and effective results.

Most of the ontology mapping methods in the literature assign a probabilistic weight for the similarities between a pair of concepts, and usually set a threshold value for showing the possible matches (e.g., a threshold value was set by the domain expert in ONION [62]). In our case, we can see such similarity weights to be booleans, i.e., given two concepts c_i and c_j from the two models; the similarity between these two is defined by

$$P(c_i, c_j) = \begin{cases} 1, & if \quad c_i = c_j \\ 0, & Otherwise \end{cases}$$

Here, the two concepts equal implies that the two concepts are semantically equivalent.

3.2.3 Mapping Results

We were able to map the ASTM-EHR concepts into corresponding concepts of the CPR ontology. The mapping results are shown in Figure 3.2. Since for some domains, we found multiple co-domain values, clearly, the mapping results show a relational mapping (not a functional mapping) between the two sets of concepts. Direct mapping was possible in most of the cases. Some ASTM-EHR concepts were mapped into more than one CPR ontology concept (e.g., 'Therapy/Procedures' were mapped to two different concepts – 'Therapeutic Act' and 'Procedure'). In such cases, we divided the attributes of the ASTM-EHR concept and mapped these into the appropriate CPR ontology concepts.

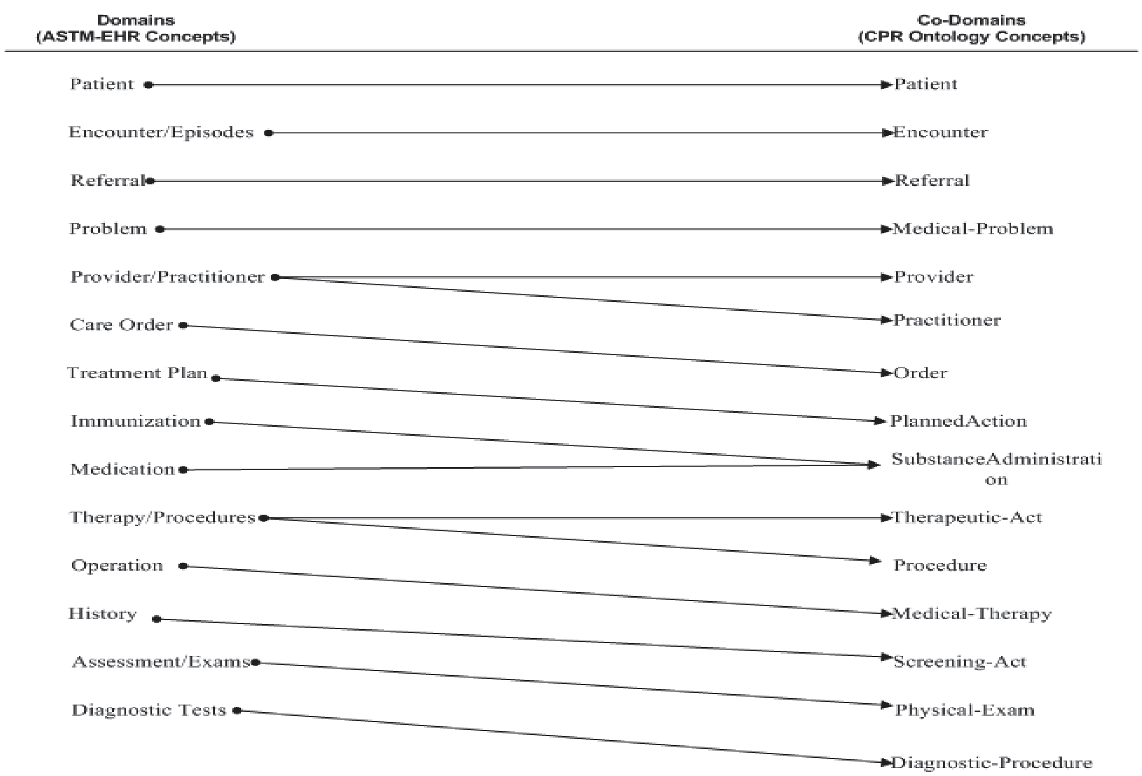


Figure 3.2 Concept-level mapping results between the ASTM-EHR and the CPR ontology.

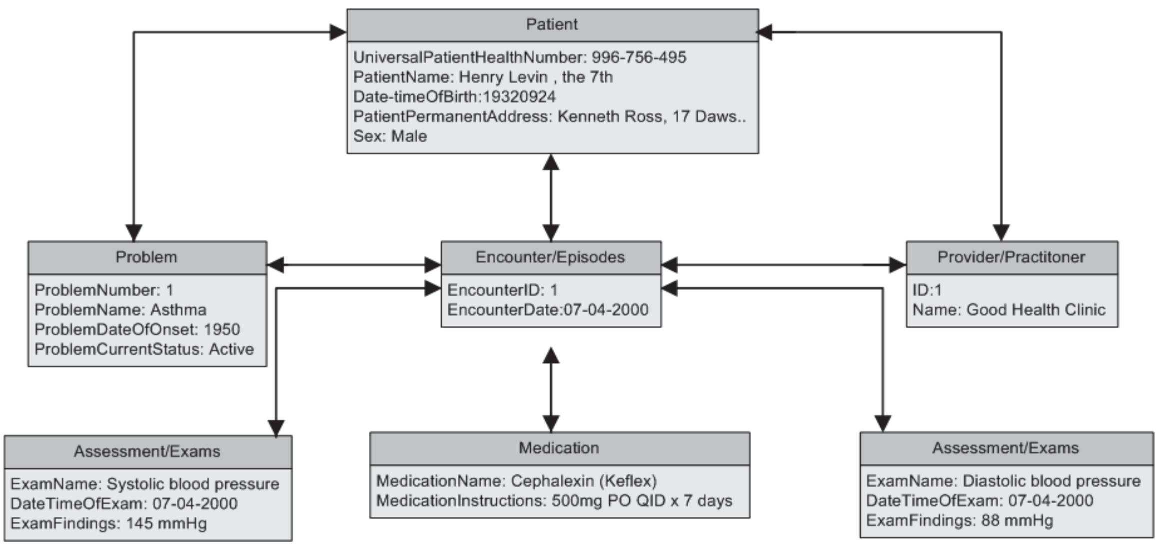
We also found the same co-domain value for multiple domain values such as 'SubstanceAdministration' both for 'Immunization' and 'Medication'. This is acceptable since both of these concepts share most of the properties in them and a single property, 'MedicationOrVaccinationType' would be used to successfully distinguish between them.

Once the concept-level mapping was done and we fitted the appropriate attributes from the ASTM-EHR model, the CPR ontology concepts had all the data-type properties necessary for an EHR. This can be checked by investing the mapping results in the reverse direction (i.e., from the CPR ontology to the ASTM-EHR model). We found that most of the concepts of the CPR ontology had the necessary properties for holding instances, with some exceptions (e.g., 'Clinical Administration Act', 'Clinical Analysis Act' etc.). However, since these concepts specialize some other concepts (e.g., 'Clinical-Act'), they hold some basic properties (e.g., action-date, note-text etc. for 'Clinical Analysis Act') inherited from their parents.

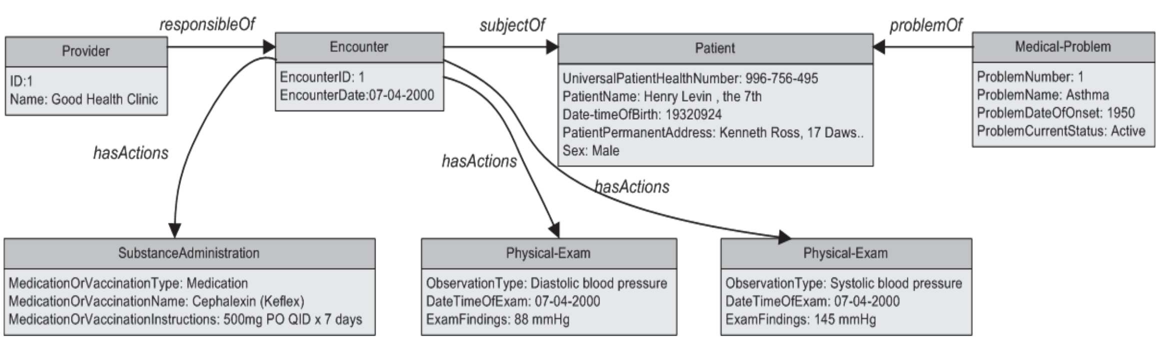
An example can be used to show how this mapping works. Here, we are using a small medical record (adopted from [70]) to instantiate that in the ASTM-EHR model, and then in the CPR ontology based on these mapping results. The medical record can be stated as follows:

"Henry Levin , the 7^{th} is the patient with DOB: 24-09-1932; Permanent Address: Kenneth Ross, 17 Daws Rd. Blue Bell, MA, 02368, Tel: (888) 555-1212; Sex: Male. He has been suffering from Asthma since 1950, went to see his healthcare provider, Good Health Clinic, on April 07, 2000. The provider found his blood pressure was 145:88 and suggested to take Cephalexin (Keflex) 500mg daily for next seven days."

This simple medical record incorporates the demographical information both for patient and healthcare provider, and two types of clinical actions: physical examination, and substance administration. This can be easily instantiated into the ASTM-EHR model which is shown in Figure 3.3 (a). As shown in the figure, the concepts and the attributes of the ASTM-EHR model are capable to represent this example record in the existing structure of this model. We also instantiated the record into the CPR ontology which we found after the mapping was done with the ASTM-EHR model, and thus incorporating the necessary data-type properties into it. The instantiation results are shown in Figure 3.3 (b). As shown in the figure, it represents very similar results to those with the ASTM-EHR model. However, as discussed before, we have only single concept, 'SubstanceAdministration', both for 'Medication' and 'Immunization' into the ontology, and a single property 'MedicationOrVaccinationType' is used to distinguish between them.



(a)



(b)

Figure 3.3 Instantiation results for the sample medical record, both for (a) the ASTM-EHR model, and, (b) the CPR ontology.

This example indicates the validity of our mapping results for a simple medical record. Moreover, instantiation results for some complete medical records in the CPR ontology will be shown in Chapter 4.

3.3 MAPPING BETWEEN THE CPR ONTOLOGY AND THE CDM MODEL

Since there are some differences between the information model of an EMR and the information model required for chronic disease management, a mapping scheme between these two is crucial. We mapped the CDM model onto the CPR ontology to make sure that our resultant EMR ontology contains all the necessary information elements required for successful chronic disease management. This problem can be more formally specified as:

Given two models $M_1 = (C_1, R_1, A_1)$ and $M_2 = (C_2, R_2, A_2)$, we have to derive a resultant model,

$$M' = \{(C', R', A') \mid C' \in (C_1 \cup (C_2 - C_1)) \wedge R' \in (R_1 \cup (R_2 - R_1)) \wedge A' \in (A_1 \cup (A_2 - A_1))\},$$

where,

C_1 is the set of concepts in the CPR ontology,

C_2 is the set of concepts in the CDM model,

R_1 is the set of binary relations (both taxonomic and associative) in the CPR ontology,

R_2 is the set of binary relations (both taxonomic and associative) in the CDM model,

A_1 is the set of data-type properties in the CPR ontology,

A_2 is the set of attributes in the CDM model.

Clearly, the problem applies to the mappings between these two standards at the concept level, the attribute level and at the relation level. We are going to discuss each of these mappings problems, mapping procedures and results in the following sections.

3.3.1 Concept-level Mapping between the CPR ontology and the CDM model

Problem Specification: The concept-level mapping can further be broken down into smaller parts:

- Identifying the semantically equivalent concepts in the CPR ontology and the CDM model, i.e., the set of concepts, $C = \{x \cap y | x \in C_1 \wedge y \in C_2\}$, where, C_1 and C_2 are the set of concepts in the CPR ontology and the CDM model respectively.

- Identifying the set of concepts in the CDM model for which no semantically equivalent concepts are found in the CPR ontology, i.e., the set of concepts, $C = \{y - x | x \in C_1 \wedge y \in C_2\}$, where, C_1 and C_2 are the set of concepts in the CPR ontology and the CDM model respectively.

- Identifying the appropriate hierarchical position in the CPR ontology for the concept set found in the previous step.

Mapping Procedure: For the first step discussed in the problem specification, we have to find out a set of semantically equivalent relations between the two models. The second step will result into the set of concepts for which no such equivalent relations were found, and the third step would require implementations of ontology engineering in the CPR ontology.

==

The Concept-Level Mapping Procedure between the CPR Ontology and the CDM Model (Pseudo Code)

==

Inputs: Two sets of concepts, $C_1=\{c_{11}, c_{12}, c_{13},\ldots\ldots\ldots, c_{1n}\}$ and $C_2=\{c_{21}, c_{22}, c_{23},\ldots\ldots\ldots, c_{2m}\}$, for the CPR ontology and for the CDM model respectively, and the taxonomic relations (is-a) among the concepts in a set.

Output: A resultant set of concepts $C' = \{x \mid x \in (C_1 \cup (C_2 - C_1)\}$, such that,

- $c_{1i}\Re_T c_{1j} \mid (c_{1i}, c_{1j}) \in C_1 \rightarrow c'_i\Re_T c'_j \mid (c'_i, c'_j) \in C'$, and similarly,
  $c_{2i}\Re_T c_{2j} \mid (c_{2i}, c_{2j}) \in C_2 \rightarrow c'_i\Re_T c'_j \mid (c'_i, c'_j) \in C'$, where,
  $\Re_T$ denotes the taxonomic relations.

==

```
For each concept in the CDM model do
Begin
        Is it possible to find the semantically equivalent concept in the CPR ontology?
        If yes then
                Continue;
        Else
                Does a generalized concept for that exist in the CPR ontology?
                If yes then
                        Create a specialized class in the CPR ontology for this concept of the CDM
                        model.
                Else
                        Create a top-level concept for this concept of the CDM model.
                End If;
        End If;
End
```

Figure 3.4 Pseudo Code for the Concept-level mapping between the the CPR ontology and the CDM model.

To solve this mapping problem, we followed the steps shown in the pseudo code in Figure 3.4. For each concept in the CDM model, we investigated whether a semantically equivalent concept is already present in the CPR ontology. For example, the concepts ‘Person’ in the CDM model and ‘Patient’ in the CPR ontology are semantically equivalent. If no such direct mapping was possible, we checked whether any generalized concept for that was present in the CPR ontology. For example, we did not find any direct semantic match for the concept ‘PlannedAct’. However, we found a generalized concept for it, ‘clinical-act’, was present in the CPR ontology, and thus we created a new concept for ‘PlannedAct’ as a child of ‘clinical-act’. If no generalized concept was found in the CPR ontology, we created a top-level concept for this. For example, we created a top-level concept for ‘Goal’.

The semantic matching between different concepts was done manually, i.e., by applying domain knowledge of these two models. We could apply some ontology mapping techniques (e.g., ConcepTool, GLUE, etc.) for these mappings. However, the CDM model was not implemented as an ontology and we needed to develop an ontology on that before we could apply one or more of those ontology mapping techniques. Although it was more time-consuming to perform the mappings manually, we argue that it was necessary to achieve more accurate results.

Mapping Results: It was observed that the CPR ontology lacked some very crucial concepts for chronic disease management (e.g, 'Planned Actions', 'Goals', 'Referrals' etc.). We created these concepts under the appropriate hierarchy. Some concepts were found in both the CDM model and the CPR ontology (e.g., 'Patient', 'Physician', 'Procedure', 'Diagnostic Procedure' etc.). Out of 26 concepts in the CDM model, we had to create 8 new in the CPR ontology. The detail results are shown in Appendix A.

The example shown in section 3.2.3 can be used to show how our mapping works. In the CDM model, the concept 'Person' is used to encode the patient information, 'Provider' for the healthcare provider, 'Chronic Condition' for medical problem, 'MedicationOrVaccination' for the medication, and 'Observation' for the physical examinations. There is no concept corresponding to 'Encounter' in the CDM model. The instantiation results for the CPR ontology is shown in Figure 3.3, which indicates that the concept-level mapping results between these two models are valid.

3.3.2 Relation-level Mapping between the CPR ontology and the CDM model

Problem Specification: The relation-level mapping between the CPR ontology and the CDM model can be formally specified as:

Two sets of binary relations, $\Re_1 = (\Re_{11}, \Re_{12}, \Re_{13}, ... \Re_{1n})$ and $\Re_2 = (\Re_{21}, \Re_{22}, \Re_{23}, ... \Re_{2m})$, for the CPR ontology and for the CDM model respectively with the constraints associated with these relations are given. We have to derive a resultant set of relations $\Re' = \{x \mid x \in (\Re_1 \cup (\Re_2 - \Re_1)\}$, such that,

- $c_{1i}\Re_1 c_{1j} \mid (c_{1i}, c_{1j}) \in C_1 \rightarrow c_i'\Re' c_j' \mid (c_i', c_j') \in C'$, and similarly,
- $c_{2i}\Re_2 c_{2j} \mid (c_{2i}, c_{2j}) \in C_2 \rightarrow c_i'\Re' c_j' \mid (c_i', c_j') \in C'$, where,

  (c_{1i}, c_{1j}) , (c_{2i}, c_{2j}), (c_i', c_j') denote the concepts in the CPR ontology, the CDM model and the resultant model respectively.

Thus, we have to satisfy that the relations between concepts in the CDM model are also present between the equivalent concepts in the CPR ontology. Here, the binary relations imply the associative relations and do not include the taxonomic relations as we already considered those while doing the concept-level mapping between these two. For example, there exists an associative relation, 'has', between the domain 'Care Plans' and the range 'Goals' in the CDM model. The resultant model should include a similar relation between semantically equivalent concepts of those two in it.

Mapping Procedure: Since we already mapped the concepts of the two models, the relation-level mapping would be easier. To solve this mapping problem, we followed the steps shown in the pseudo code in Figure 3.5. For each relation present in the CDM model, we investigated whether a similar relation exists between the semantically equivalent concepts in the CPR ontology. If that holds true, we investigated for the constraints (i.e. property restrictions) associated with that relation. This is important to satisfy the conditions necessary for an information model on chronic disease management. Otherwise, we created a relation in the CPR ontology between the semantically equivalent concepts.

===

The Relation-Level Mapping Procedure between the CPR Ontology and the CDM Model (Pseudo Code)

===

Inputs: Two sets of binary relations, $\Re_1 = (\Re_{11}, \Re_{12}, \Re_{13}, \ldots\ldots \Re_{1n})$ and $\Re_2 = (\Re_{21}, \Re_{22}, \Re_{23}, \ldots\ldots \Re_{2m})$, for the CPR ontology and for the CDM model respectively with the constraints associated with these relations. Here, the binary relations do not include the taxonomic relations.

Output: A resultant set of relations $\Re' = \{x \mid x \in (\Re_1 \cup (\Re_2 - \Re_1))\}$, such that,

- $c_{1i}\Re_1 c_{1j} \mid (c_{1i}, c_{1j}) \in C_1 \rightarrow c_i'\Re' c_j' \mid (c_i', c_j') \in C'$, and similarly,
- $c_{2i}\Re_2 c_{2j} \mid (c_{2i}, c_{2j}) \in C_2 \rightarrow c_i'\Re' c_j' \mid (c_i', c_j') \in C'$, where,

  (c_{1i}, c_{1j}) , (c_{2i}, c_{2j}), (c_i', c_j') denote the concepts in the CPR ontology, the CDM model and the resultant model respectively.

===

```
For each relation, c_2i ℜ_2 c_2j in the CDM model do
Begin
    Does there already exist a similar relation in the CPR ontology between the semantically
    equivalent concepts of c_2i and c_2j ?
    If yes then Begin
        Does that relation in the CPR ontology fulfill the constraint requirements specified in
        ℜ_2 ?
        If Yes then Continue;
        Else
            Apply that constraint for the relation in the CPR ontology.
        End If;
    Else
        Create a relation in the CPR ontology between the semantically equivalent concepts of
        c_2i and c_2j .

    End If;
End
```

Figure 3.5 Pseudo Code for the Relation-level mapping between the the CPR ontology and the CDM model.

The constraints on relations in the CDM model were modeled as object-type properties in the Protégé ontology while applying different conditions. For example, the property was made functional to satisfy the constraint 'functional'. The 'mandatory' constraints on relations in the CDM model have been modeled by setting the minimum cardinalities of corresponding object-type properties to one. Similarly, necessary conditions were applied on those properties to satisfy the constraint ‘mandatory’.

The semantic matching between different relations was done manually, i.e., by applying domain knowledge of these two models. We could not directly apply any of the ontology mapping techniques for the same reason discussed in section 3.3.1.

Mapping Results: We found that no similar relations exist in the CPR ontology for most of the relations in the CDM model. This can be explained as the differences between these two model – the CDM model aims to satisfy the information needs only for chronic disease management where the CPR ontology defines a more general information structure for an EHR. Out of the 17 relations in the CDM model, we had to create 13 new in the CPR ontology. The detail results are shown in Appendix B.

The example shown in section 3.2.3 can be used to show how our relation-level mapping results. The instantiation results for that sample record for the CDM model is shown in Figure 3.7. Since we created an additional concept, ‘Encounter’ in the CPR ontology, the relation ‘*hasActions*’ connecting this with clinical actions corresponds to the relation ‘*containsRelated*’ in the CDM model. Similarly, other equivalent matches will be found in Figure 3.3 for the instantiations of the same record in the CPR ontology.

3.3.3 Attribute-level Mapping between the CPR ontology and the CDM model

Problem Specification: The attribute-level mapping between the CPR ontology and the CDM model can be formally specified as:

Given a set of data-type properties, $D_1 = (D_{11}, D_{12}, D_{13}, ... D_{1n} \mid (D_{1i}, D_{1j}, D_{1k}, ..) \in c_{1l})$ in the CPR ontology and a set of attributes

$A_2 = (A_{21}, A_{22}, A_{23}, ... A_{2m} \mid (A_{2i}, A_{2j}, A_{2k}, ..) \in c_{2l}))$ in the CDM model, where, $(D_{1i}, D_{1j}, D_{1k}, ..)$ is a set of data-type properties in the CPR ontology which belong to a concept c_{1l} in it, and similarly, $(A_{2i}, A_{2j}, A_{2k}, ..)$ is a set of attributes which belong to a concept c_{2l} in the CDM model. We have to derive a resultant set of data-type properties for each concept c_{1l} in the CPR ontology such that,

$$((D_{1i}, D_{1j}, D_{1k}, ..) \cup (A_{2i}, A_{2j}, A_{2k}, ..)) \in c_{1l} \mid (D_{1i}, D_{1j}, D_{1k}, ..) \in c_{1l} \wedge (A_{2i}, A_{2j}, A_{2k}, ..) \in c_{2l} \wedge (c_{1l} = c_{2l})$$

Here, $(c_{1l} = c_{2l})$ implies that these two concepts are semantically equivalent.

Here, we have to ensure that there exist corresponding data-type properties in the CPR ontology concept for each and every attribute present in the semantically equivalent concept in the CDM model. This will ensure proper classification of concepts particularly important for chronic disease management in the CPR ontology.

Mapping Procedure: Since we already mapped the concepts of the two models, the attribute-level mapping would be easier. To solve this mapping problem, we followed the steps shown in the pseudo code in Figure 3.6. For each concept present and for each attribute of it in the CDM model, we investigated whether a similar data-type property exists in the semantically equivalent concepts in the CPR ontology. If no such direct mapping was found for a particular attribute, we investigated whether a similar data-type property is available in the generalized concept of it in the CPR ontology. This will still satisfy our purpose as the properties can be inherited to specialized concepts in an ontology. For example, the property 'ObservationType' was inherited to the concept 'Physical-Exam' from its generalized concept, 'clinical-investigation-act'. If that does not hold, we created a data-type property in the concepts of the CPR ontology.

==

The Attribute-Level Mapping Procedure between the CPR Ontology and the CDM Model (Pseudo Code)

==

Inputs: A set of data-type properties, $D_1 = (D_{11}, D_{12}, D_{13}, \ldots\ldots D_{1n} \mid (D_{1i}, D_{1j}, D_{1k}, \ldots) \in c_{1l})$ in the CPR ontology and a set of attributes $A_2 = (A_{21}, A_{22}, A_{23}, \ldots\ldots A_{2m} \mid (A_{2i}, A_{2j}, A_{2k}, \ldots) \in c_{2l}))$ in the CDM model, where, $(D_{1i}, D_{1j}, D_{1k}, \ldots)$ is a set of data-type properties in the CPR ontology which belong to a concept c_{1l} in it. And similarly, $(A_{2i}, A_{2j}, A_{2k}, \ldots)$ is a set of attributes which belong to a concept c_{2l} in the CDM model.

Output: A resultant set of data-type properties for each concept c_{1l} in the CPR ontology such that,

$((D_{1i}, D_{1j}, D_{1k}, \ldots) \cup (A_{2i}, A_{2j}, A_{2k}, \ldots)) \in c_{1l} \mid (D_{1i}, D_{1j}, D_{1k}, \ldots) \in c_{1l} \wedge (A_{2i}, A_{2j}, A_{2k}, \ldots) \in c_{2l} \wedge (c_{1l} = c_{2l})$

Here, $(c_{1l} = c_{2l})$ implies that these two concepts are semantically equivalent.

==

```
For each concept, c_2l in the CDM model do
Begin
    For each attribute, (A_2i, A_2j, A_2k, ...) ∈ c_2l do
    Begin
        Does there exist a similar data-type property in c_1l of the CPR ontology, where c_2l and
        c_1l are semantically equivalent?
        If Yes then continue;
        Else
            Does there exist a similar data-type property in a generalized concept of c_1l in
            the CPR ontology that will be inherited to it?
            If Yes then continue;
            Else
                Create a data-type property in c_1l
            End If;
        End If;
End
```

Figure 3.6 Pseudo Code for the Attribute-level mapping between the the CPR ontology and the CDM model.

We found the semantic matching between different attributes manually, i.e., by applying domain knowledge of these two models. We could not directly apply any of the ontology mapping techniques for the same reason discussed in section 3.3.1.

Mapping Results: We found that out of 82 attributes present in the CDM model, there were no similar data-type properties for 37 in the CPR ontology. These were mostly related to the planned clinical actions. This is justifiable as we had to create concepts for those in the CPR ontology. The detail results are shown in Appendix C.

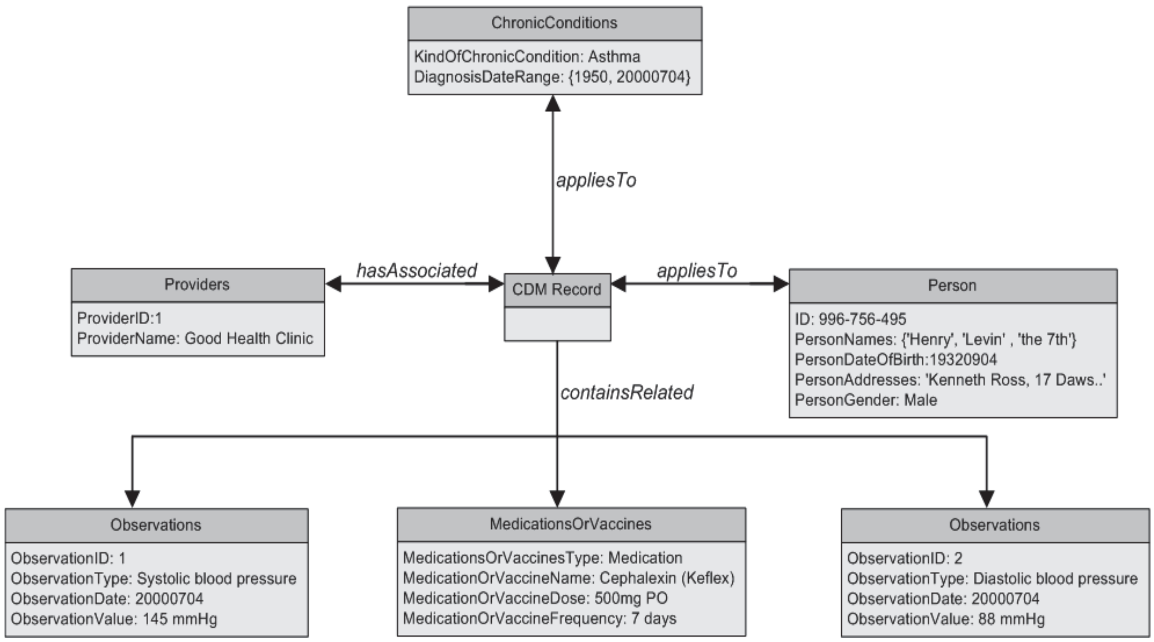


Figure 3.7 Instantiation results for the sample medical record for the CDM model.

The example shown in section 3.2.3 can be used to show how our attribute-level mapping results. The instantiation results for that sample record for the CDM model is shown in Figure 3.7. As shown the figure, the CDM model lacks some of the attributes (e.g., medical problem status, problem number etc.) necessary for proper classification of concepts. This is because the CDM model only provides information pertinent for chronic disease management, not for a complete EHR. However, the examples show the validity of our mapping results for a simple medical record. In chapter 4, we will also show instantiation results in the resultant EMR ontology for a complete medical record used by the CDM model specification [69]. This will be useful to show that a record on the CDM model can successfully be used to instantiate our resultant ontology.

3.4 THE RESULTANT EMR ONTOLOGY

After all of these mappings of the CPR ontology with the ASTM-EHR model and with the CDM model, we found the resultant EMR ontology capable to represent clinical knowledge. We are going to describe the components of this ontology here.

3.4.1 Concepts

The resultant EMR ontology contains a few concepts which have been found necessary for appropriate classification of clinical information and for holding necessary data elements for successful chronic disease management. The core concepts with relations among them are shown in Figure 3.8 and are discussed briefly below.

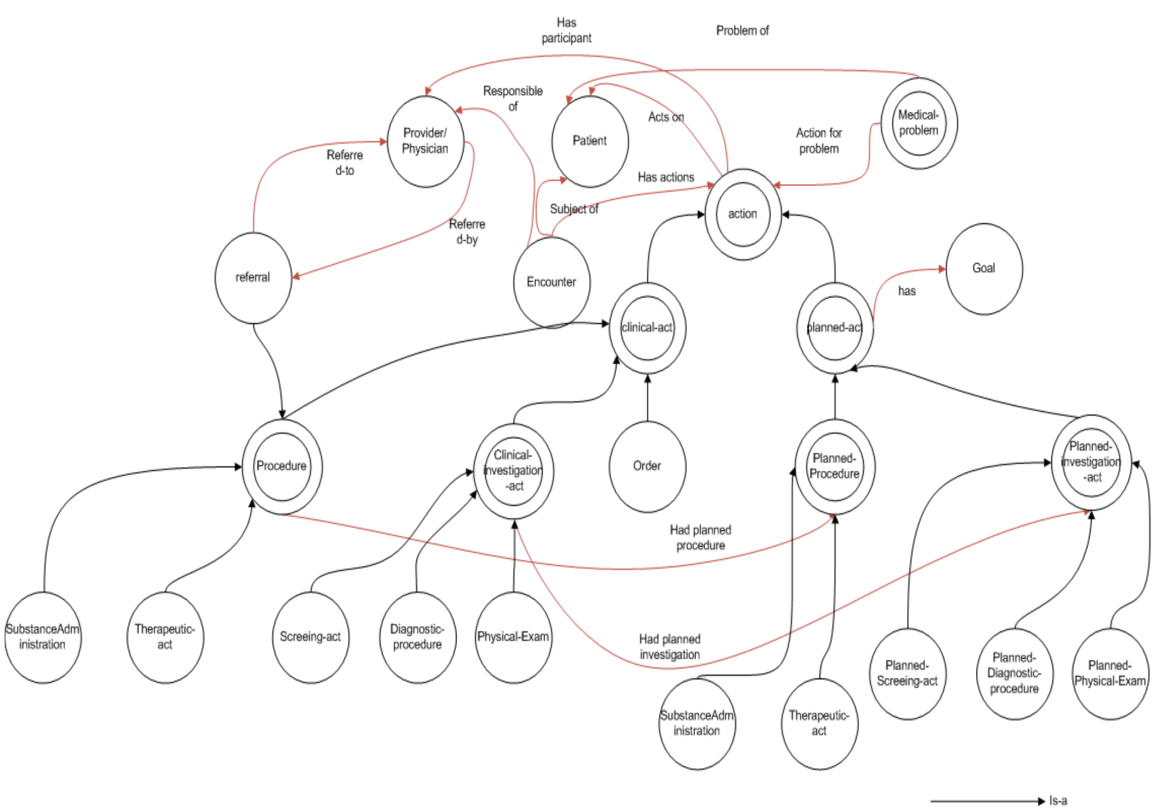


Figure 3.8 The core concepts and relations among them of the resultant EMR ontology.

Actions: The most important concept in the ontology are ‘Action’, which describe all the clinical actions already taken or intended to be taken in future. This implies the primary classification of this concept into two specializations: ‘clinical-act’ (i.e., the actions already taken), and ‘planned-act’ (i.e., the intended actions). There have been relations between the specializations of ‘clinical-act’ with corresponding specializations of ‘planned-act’. This ensures proper tracking and monitoring of a planned action, i.e., whether it has actually been carried out or not.

The ‘clinical-act’ is further classified into specialized concepts:

- Procedure: As discussed in section 2.5, ‘Procedure’ is such type of clinical action whose outcome is expected to improve the patient’s condition. These include patient education (particularly important for chronic disease management), skill development of the patient, therapies, referral etc. This concept has three further specializations in our ontology: ‘Therapeutic-Act’ (for representing different medical, psychological therapies including surgery), ‘Referral’ (e.g., the procedure of referring a patient to a specialized), and ‘Substance-Administration’ (which represents clinical actions related to both medication and vaccination).

- Clinical-Investigation-Act: This act is taken to investigate for possible diagnosis of a patient, and is further specialized into three concepts: ‘Screening-Act’, ‘Physical-Exam’, and ‘Diagnostic-Procedure’. ‘Screening-Act’ is used to take different medical, social, family histories of a patient, ‘Physical-Exam’ is done usually by a clinician to find out possible signs of a patient, and ‘Diagnostic-Procedure’ is the process of assessing the diagnosis; includes both laboratory or radiologic procedures.

- Order: This relates to clinical orders instructed to be performed by physician/healthcare providers to perform a particular task, e.g., order for 100 new cabins in a clinic for in-patients.

This conforms to the concepts defined in the CPR ontology, and, the planned actions (with goals) from the CDM model fill the gaps in it.

Medical Problems: Medical problem is the central key concept in any POMR-based EMR structure, and thus in the resultant EMR ontology. All the clinical actions are organized under medical problems for a patient. Some coded properties for medical problems (e.g., kind of problem, recall indicator, etc.) have been imported from the CDM model into the resultant ontology.

Person: We have mainly two kinds of persons involved in the resultant ontology: 'Patient', and 'Physician'. However, we made a separate class for 'Provider' and made it different from 'Physician' in the resultant ontology. We also created a specialization of this, 'PrimaryCareProvider', which is primarily responsible for providing care to a particular patient.

Encounter: We introduced a new concept, 'Encounter', into this ontology which, we believe, is very important to keep track of follow-ups and long-term care plans.

Other Concepts: Some other concepts (e.g., 'AdministrativeCorrection') needed to be created in the CPR under the appropriate hierarchy. We also kept the additional concepts (e.g., 'MaterialEntity') present in the CPR ontology.

We made the concepts to be disjoint wherever appropriate. For example, the concepts 'Procedure' and 'Clinical-Investigation-Act' are clearly disjoint. A number of concepts in the Protégé ontology have been made primitive classes by applying necessary and necessary and sufficient conditions. For example, it is necessary for a 'Referral' to be referred by a 'Physician'. However, it is necessary and sufficient for a 'Referral' to be referred by a 'Physician' and also referred to another 'Physician'. These conditions have been modeled as $\text{Physician} \quad \mathit{referredBy} \quad \text{Referral} \rightarrow \text{Referral}$, and, (Physician *referredBy* Referral) $\wedge$ (Referral *referredTo* Physician) $\leftrightarrow$ Referral respectively.

3.4.2 Relations

The binary relations in an ontology are used to relate between two concepts, or two relations in it. We are going to describe the taxonomic and associative relations used in the resultant EMR ontology below.

Taxonomic Relations: These represent the taxonomic hierarchical structure among concepts in an ontology and are represented as '*is-a*', '*is-kind-of*' or '*is-part-of*' relations. Such a hierarchical structure depends on definitions of the concepts. As shown in Figure 3.8, we ensured proper taxonomic relations between concepts. For example, clearly the definition of 'Diagnostic-Procedure' indicates that these clinical actions are conducted to investigate possible diagnoses of a patient, and thus should better specialize the concept 'Clinical-Investigation-Act', and not the concept 'Procedure'.

Associative Relations: These are used to relate between two concepts in an ontology and examples of such relations include '*has-a*', '*has-property*' etc. As shown in Figure 3.8, we have defined the necessary such relations between concepts in the EMR ontology. As discussed before, these associative relations resulted into after mapping the CPR ontology with the CDM model. However, since we created a new concept 'Encounter' in the resultant ontology, it is associated with clinical actions by such an associative relations, '*hasActions*', which ensures that the clinical actions corresponding to an encounter will be grouped together. Further, this concept also relates with the concept 'Patient', and 'Physician', and there exists a relation between 'Patient' and 'Medical-Problem' to group the medical problems for a patient together. These relations have been modeled in the Protégé ontology as object-type properties while satisfying the required cardinality constraints.

3.4.3 Vocabulary

We ensured a structured way of data entry by using the codes described in the CDM model [32]. We converted all the codes of the CDM model [32] into corresponding SNOMED-CT codes and integrated these into our EMR ontology. We chose SNOMED-CT since it provides the most robust vocabulary domains. Also, using a single vocabulary standard ensures ease of implementation and integration with other clinical information systems.

For each of the code values described in the CDM model, we found the appropriate equivalent code value in the SNOMED-CT vocabulary. We found all of those code values were successfully mapped into SNOMED-CT. An example of such conversion for the coded values of ‘ObservationType’ from LOINC code values to corresponding SNOMED-CT codes are shown in Table 3.1.

Table 3.1 Examples of code value conversion for data-type property ‘ObservationType’.

LOINC Concept (Code)	**Corresponding SNOMED-CT Concept (Code)**
Adverse Drug Reaction (44939-7)	Adverse reaction to drug (62014003)
Medication Concerns (28174-1)	Medication response (405177001)
Person History (35090-0)	Demographic history detail (302147001)
Family History (10157-6)	Family health status (405205002)
Smoker (11367-0)	Smoker (77176002)
Cigarette packs/day (8663-7)	Cigarette consumption (230056004)
Cigarette exposure in pack years (8664-5)	Cigarette pack-years (401201003)
Alcohol drinks/week (44940-5)	Alcohol intake (160573003)
Height (3137-7)	Body height measure (50373000)
Weight (8350-1)	Body weight measure (27113001)
Waist circumference (8280-0)	Waist circumference (276361009)
Heart-rate (8893-0)	Heart rate (364075005)
Heart rhythm (8884-9)	Pulse rhythm (364095004)
Systolic blood pressure (8459-0)	Systolic blood pressure (271649006)
Diastolic blood pressure (8453-3)	Diastolic blood pressure (271650006)
A1C (4548-4)	Hemoglobin A1C (269823000)
Total Cholesterol (14647-2)	Dietary cholesterol intake (289183008)
HDL Cholesterol (14646-4)	High density lipoprotein (443835004)

LOINC Concept (Code)	Corresponding SNOMED-CT Concept (Code)
LDL Cholesterol (22748-8)	Low density lipoprotein (226117007)
TC:HDL cholesterol ratio (32309-7)	High density lipoprotein (HDL)/total cholesterol ratio measurement (104583003)
Triglycerides (14927-8)	Finding of triglyceride level (365795001)
Serum-Creatinine (14682-9)	Finding of serum creatinine level (365757006)
Albumin-creatinine ratio (14959-1)	Albumin/creatinine ratio measurement (250745003)
Creatinine clearance – calculated (33914-3)	Creatinine renal clearance, function (102811001)
24 hour urinary protein (21482-5)	24 hour urine output (395060000)
Serum Potassium (2823-3)	Total body potassium (251838003)
Serum Sodium (2951-2)	Serum appearance (314037008)
Hemoglobin (35183-3)	Hemoglobin finding (250220000)
Transferrin Saturation (2505-6)	Total iron binding capacity (117173006)
Serum calcium (2000-8)	Calcium volume (416320009)
Serum phosphorus (14879-1)	Phosphate (102822002)
Intact Parathyroid Hormone (14866-8)	Measurement of parathyrin antibody (117924003)
Serum albumin (1751-7)	Finding of albumin level (365801005)
Alanine aminotrasferase (1742-6)	Alanine measurement (104479001)
Creatine Kinase (2157-6)	Finding of creatine kinase level (398137007)

In order to bind these vocabulary values into the resultant EMR ontology, we constructed a top-level concept, 'Vocabulary' which contains all the code table values with two properties: concept name (i.e., the concept name in SNOMED-CT), and code (i.e., the exact code in SNOMED-CT for this concept). These were then bound to the appropriate data-type properties in the Protégé EMR ontology. For example, to bind the concept 'VocObservationType' (which contains the coded values described in Table 3.1) into the data-type property 'ObservationType' of concept 'Physical-Exam', we used universal restrictions on this property such that, $\forall PhysicalExam.ObservationType \in VocObservationType$

3.5 Mapping between the EMR ontology and the HL7 RIM

3.5.1 Problem Specification

Since in version 3 of HL7, the structured documents (e.g., CDA) are based on its object-oriented model RIM, we are interested to map our EMR ontology to the RIM. This is important as we expect to get a refined subset of the RIM that is sufficient for holding necessary clinical information of our interest.

Since WHIC already mapped their proposed CDM model with the RIM and found a refined model of it [29], we had to map the additional concepts and properties used in the EMR ontology with it. The problem can be specified as:

Given two sets of formal contexts $H = (C_H, \Re_H, A_H)$ and $E = (C_E, \Re_E)$, where, C_H and C_E are the sets of concepts in the HL7-RIM and the EMR ontology respectively, A_H and $\Re_E$ are the set of attributes and data-type properties in the HL7-RIM and the EMR ontology respectively, and $\Re_H$ is the set of relations in the HL7-RIM, we have to derive a subset of H, $H' = (C'_H, \Re'_H, A'_H)$ which will be semantically closest to those in another set E

As specified in the problem, we are not going to change the existing structure of the RIM, rather intend to devise a subset of it, i.e., we are keeping the set of relations among the concepts in the resultant subset those are present in the set $\Re_H$.

These mapping results can be used to extract the portion in any CDA document useful to instantiate our EMR ontology, and thus can further be reasoned. However, in our thesis we are just going to have a refined subset of RIM, and not actually implementing how a particular portion of CDA can be extracted.

3.5.2 Mapping Procedure

The mapping procedure was done manually by choosing the closest possible concept (or property) of HL7 RIM for each concept (or property) of the EMR ontology. The pseudo-code of the mapping procedure is shown in Figure 3.9.

As shown in the pseudo-code, first we are finding the closest possible match for each of the concepts in the EMR ontology. If no semantically close match can be found, we are marking that concept for further investigation. Otherwise, we are adding the relations found for that particular concept with already added concepts. Next, we are similarly checking for possible semantic matches for each and every data-type properties of that concept in the EMR ontology.

```
========================================================================================
The Mapping Procedure between the HL7-RIM and the EMR Ontology (Pseudo Code)
========================================================================================
Inputs: Two sets of formal contexts H = (C_H, ℜ_H, A_H) and E = (C_E, ℜ_E), where C_H and C_E are
the sets of concepts in the HL7-RIM and the EMR ontology respectively, A_H and ℜ_E are the set of
attributes and data-type properties in the HL7-RIM and the EMR ontology respectively, and ℜ_H is the
set of relations in the HL7-RIM.

Output: A subset of H, H' = (C'_H, ℜ'_H, A'_H) which will be semantically closest to those in another set
E
========================================================================================
Initialize: H' := null
For each concept in the EMR ontology do
Begin
        Is it possible to find the semantically closest concept in the HL7-RIM?
        If yes then
        Begin
                Add that concept of HL7-RIM into C'_H
                Does there any relation exist in HL7-RIM for this concept with any other concept already

                added into C'_H ?
                If yes then add that relation into ℜ'_H ;
                For each data-type properties of this concept in the EMR ontology do
                Begin
                        Is it possible to find the semantically closest attribute in the HL7-RIM for this?
                        If yes then add that relation into A'_H ;
                        Else
                                        Mark that data-type property and continue;
                        End If;
                End
        Else
                Mark that concept and continue;
        End If;
End
```

Figure 3.9 Pseudo Code for the Concept-level mapping between the EMR ontology and the HL7-RIM.

The main mapping challenge we found was that the two models we are trying to map have a number of logical differences, such as, the RIM does not differentiate between occurants and continuants where the EMR ontology does that. However since RIM has a very robust vocabulary domain, we found that vocabularies can be used to make such differences. For example, the 'actCode' and 'moodCode' attributes of 'Act' can be used to differentiate between medical problems (continuant) and diagnostic procedures (occurant).

The closest semantic matches between the two models were found by applying the domain knowledge. Alternatively, we could apply some ontology mapping techniques (e.g., ConcepTool, GLUE, etc.) for these mappings. However, in doing this, we needed an ontology on the latest release of the RIM standard which is absent. Moreover, we believe that careful observations inherent in manual mapping, together with appropriate domain knowledge, helped us achieve more accurate results.

3.5.3 Mapping Results

We found that all of the concepts of the EMR ontology were successfully mapped into corresponding concepts of the HL7 RIM. Out of 80 data-type properties of the EMR ontology which were not previously mapped by WHIC, 8 were partially mapped (e.g., 'PatientPermanentAddress') and 10 could not be mapped (e.g., 'Physician.UniversalIDNumber') to the RIM. The properties for which we could not find any map, or found partial maps were mostly demographic attributes, which we believe would not have much impact on the reasoning procedures by a decision support system. The details of the mapping results are shown in Appendix D.

Since HL7 RIM provides a very robust information model to capture the clinical data from almost every aspect, we devised a smaller refined model of it, based on the mapping results. We found such a subset of HL7 RIM, which is shown in Figure 3.10.

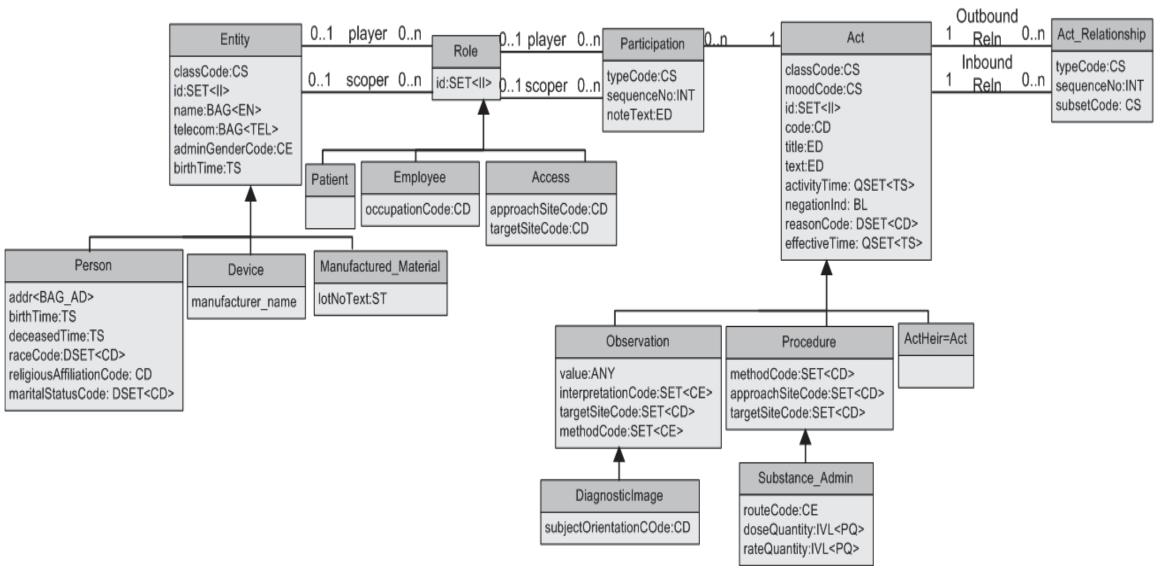


Figure 3.10 Subset of HL7 RIM for the proposed EMR ontology.

All the clinical actions of the proposed ontology were found to be mapped into the RIM classes, 'Observation' and 'Procedure'. 'ActHeir' class has been chosen to satisfy the condition that an 'Act' can have another 'Act' as part of it. The act codes have been used to represent the concepts such as planned actions (with moodcode='Goal'), medical problems (with code='Condition') etc. of this ontology. The 'Person' class of the RIM maps with both 'Patient' (with role 'Patient') and 'Practitioner' (with role 'Employee') of the proposed Ontology. 'Device' class has been chosen to map with the material entities whereas 'ManufacturedMaterial' represents 'Medication' and 'Vaccination' of the proposed ontology with role 'Access'. 'ActRelationship' and 'Participation' are core classes of the RIM being used to link between different Acts, and to define the particular actions an 'Entity' is playing within a particular 'Role'.

We are going to use the same example which we used in section 3.2.3 to show how the resultant subset of the RIM can be used to capture the clinical information. The instantiation results for that sample medical record in the RIM subset is shown in Figure 3.11.

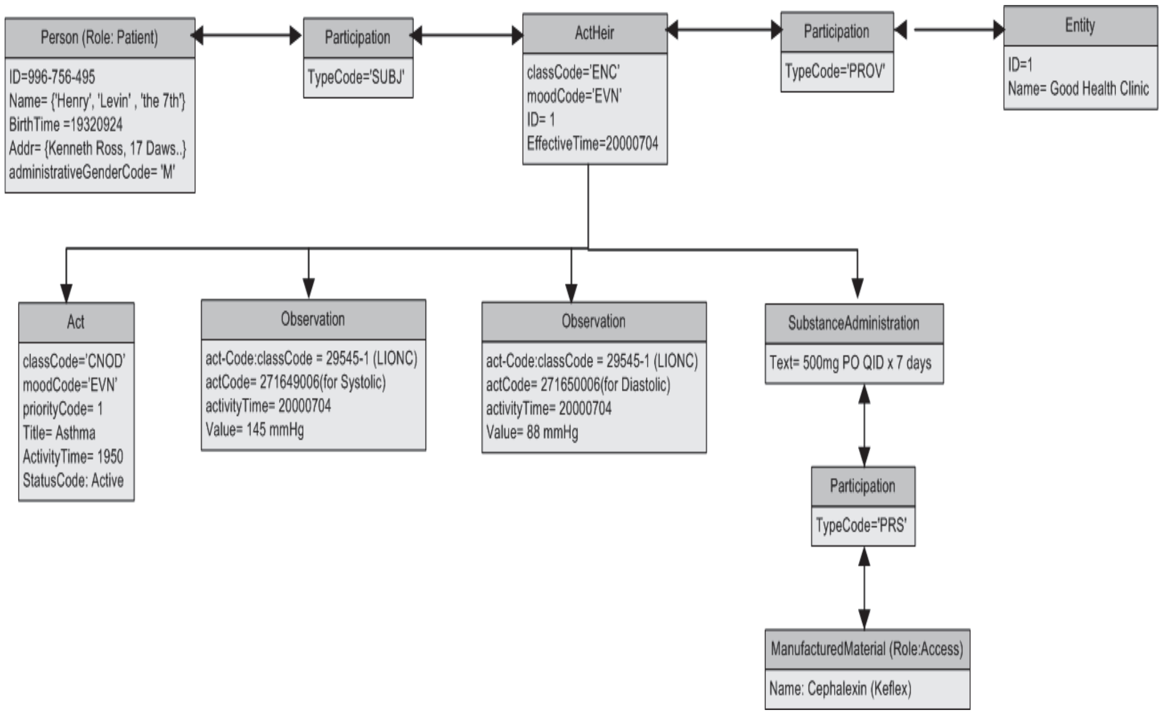


Figure 3.11 Instantiation results for the sample medical record for the subset of the RIM.

As shown in the figure, the sample medical record could be successfully instantiated in our resultant RIM subset which indicates validity of our mapping results. The use of the concept 'ActHeir' to represent 'Encounter' ensured the reflexive closure of generalization relationships (i.e., an 'Act' can contain another 'Act'). Some coded attributes of the RIM (e.g., actCode, classCode, typeCode, moodCode etc.) are used to differentiate between different concepts in the example. For example, although 'Act', 'ActHeir' and 'Observation' are under the same base class in the RIM (i.e., 'Acts'), these coded attributes could successfully differentiate between the concepts, 'Medical-Problems', 'Encounters' and 'Physical-Exam' respectively.

It is important to note here that the mapping procedure of this section has resulted into a subset of the RIM (i..e., a data record), which is sufficient to be used for carrying CDA documents or HL7 messages with the clinical information pertinent for the resultant EMR ontology. Thus, the EMR ontology can be instantiated only with this subset of RIM information.

3.6 Summary

We have proposed an EMR ontology primarily focusing on chronic disease management. This ontology is also capable of representing knowledge of other acute and infectious diseases. Since we have customized and used the WHIC-proposed CDM model [27] in our methodology, some of the data-type properties of this proposed ontology are currently bound with vocabularies for three chronic diseases (Hypertension, Chronic Kidney Disease and Diabetes). However, the structure of the ontology is general and can easily capture vocabularies necessary for other diseases.

All the mappings discussed in this chapter were done by using the domain knowledge. There are numerous ontology mapping techniques available in the literature (discussed in Chapter 2). However, these required some extra effort to build new ontologies (e.g., for the CDM model and for the latest version of the HL7 RIM). As well, most of these ontology mapping techniques need input from a human expert. Our mapping ontologies are heterogeneous, and thus we have decided to make the mapping rules manually.

The mapping results show promising results, and the resultant EMR ontology is robust while ensuring structured data entry. We believe that the proposed methodology is important in the sense that it does not only propose an EMR ontology for chronic disease management, but also be used as switching language among various EMR standards. Moreover, the mapping of this ontology with the HL7 RIM standard ensures a step forward to capture clinical records in HL7 standard, and instantiate the EMR ontology with that HL7 record, and thus, can further be reasoned by decision support systems. We will see some instantiation results of the EMR ontology for medical records in HL7 in chapter 4.

CHAPTER 4 RESULTS

4.1 Evaluation Criteria

We evaluated our proposed EMR ontology against standard ontology design principles proposed in the literature [67, 68]. We checked the proper classification of taxonomy and logical consistencies of the proposed ontology using reasoners. We also instantiated two sample medical records written in HL7 [69, 70] using the proposed ontology. Both of these two medical records represent the general characteristics necessary to successfully manage chronic diseases.

4.2 Evaluation for compliance with standard design principles of ontology

We manually checked the compliance of our proposed EMR ontology against two sets of standard design principles: Gomez-Perez's ontology design principles [67] and Bodenreider's design principles [68]. While the design principles proposed by Gomez-Perez [67] are a bit more theory-oriented and abstract, the ones proposed by Bodenreider [68] are more development-oriented. We found that our proposed EMR ontology satisfies the basic principles of Gomez-Perez [67]. Among the design principles proposed by Bodenreider [68], some were already enforced by Protégé. We examined the rest and found that the criteria 'Non-leaf classes must have at least two children' was partially satisfied by our ontology with some exceptions (e.g., the 'Provider' class has only one child, 'PrimaryCareProvider').

4.3 Checking Logical Consistencies using Reasoners

We checked the logical consistencies and proper classification of taxonomies in the proposed EMR ontology automatically using two reasoners – Pelett 2.2.2 [72] and

RacerPro 2.0 [73]. Pellet is an OWL-DL reasoner with a plugin for it in Protégé. RacerPro is another description logic ontology reasoner used to identify possible logicac inconsistencies and improper classification of taxonomy. Both of these reasoners found no logical inconsistencies (e.g., loops) in our EMR ontology.

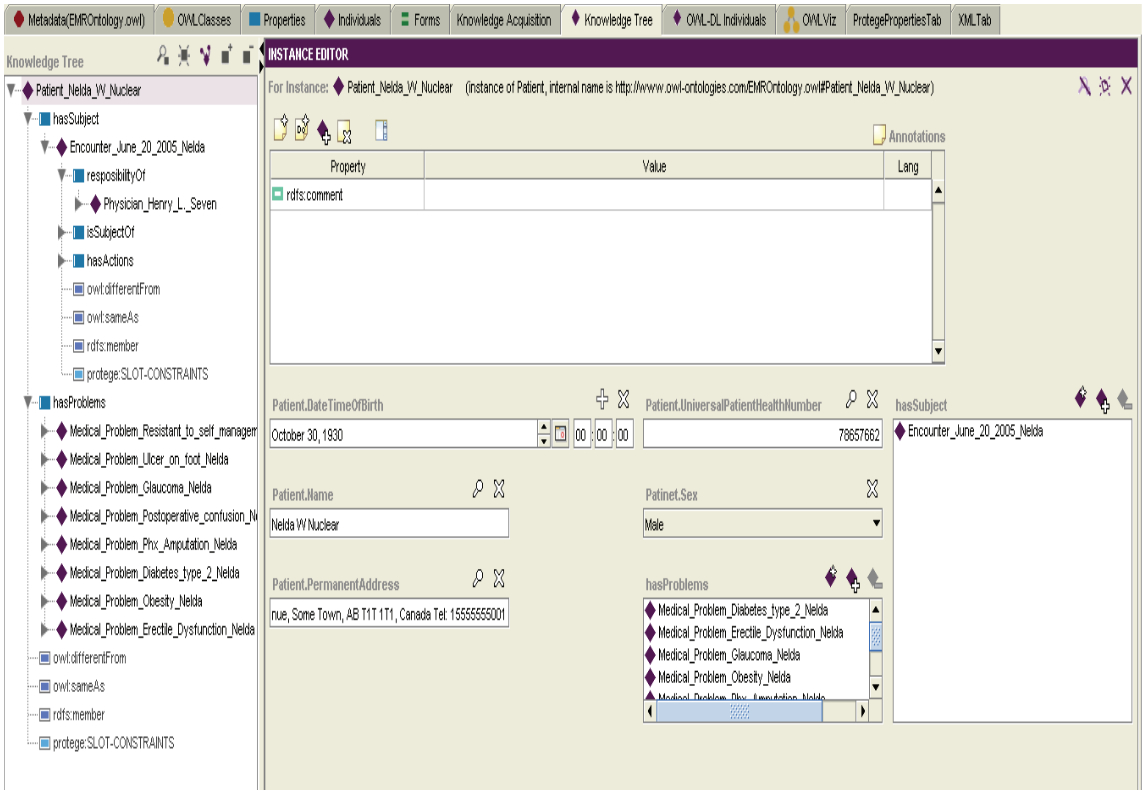
4.4 Instantiations of the proposed EMR ontology

4.4.1 Instantiation for the first medical record

The first medical record [69] we used was proposed by WHIC [27] to test their proposed CDM Model. This medical record is written in HL7 version 3, incorporates necessary data elements for chronic disease management, and is organized according to the structure of the CDM Model [27].

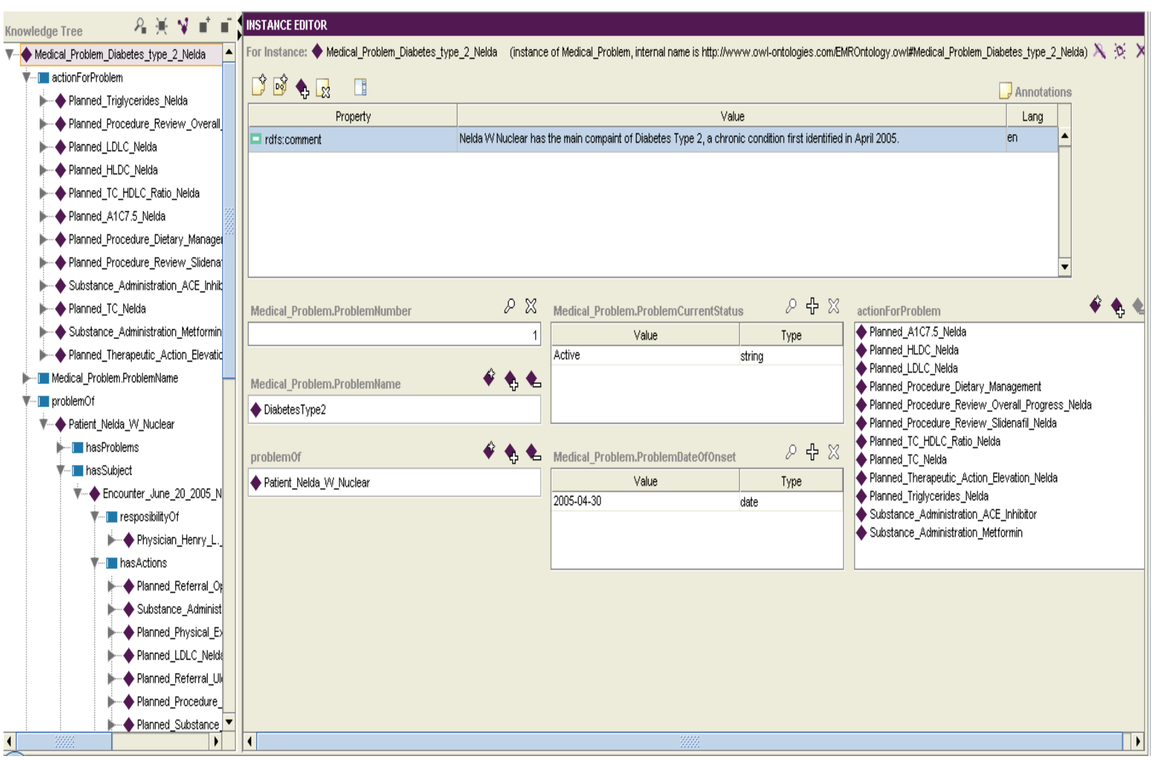
This medical record is about a 74 years (at the time the medical record was written) male patient, Nelda W Nuclear, from Some Town of British Columbia, Canada. The patient had a chronic condition, Diabetes Type II and seven other associated medical problems, namely, Resistance of Self Management, Ulcer on Foot, Glucoma, Post-Operative Confusion, Phx Amputation, Mild Obesity and Erectyle Disfunction. Nelda had previously undergone a transsexual surgery. He went to see his physician, Henry L. Seven, on June 20, 2005, when the physician planned for a number of procedures (e.g., daily dressings), referrals (e.g., ulcer clinic), physical examinations (e.g., planned blood pressure), and substance administrations (e.g., Rx aspirin and Zincaps). Each of these planned actions had an associated particular goal to achieve within a particular date (e.g., the planned weight of 85 kg by 20-09-2005). Henry also suggested Nelda to take some medications (e.g., ACE Inhibitor) to improve his conditions.

Since we already know that the CDM Model is based on POMR structure, each of the clinical actions was organized under particular medical problems (e.g., a procedure is planned to improve one or more problems). Some snapshots of instantiation results for this medical record into our Protégé EMR ontology are shown in Figure 4.1.



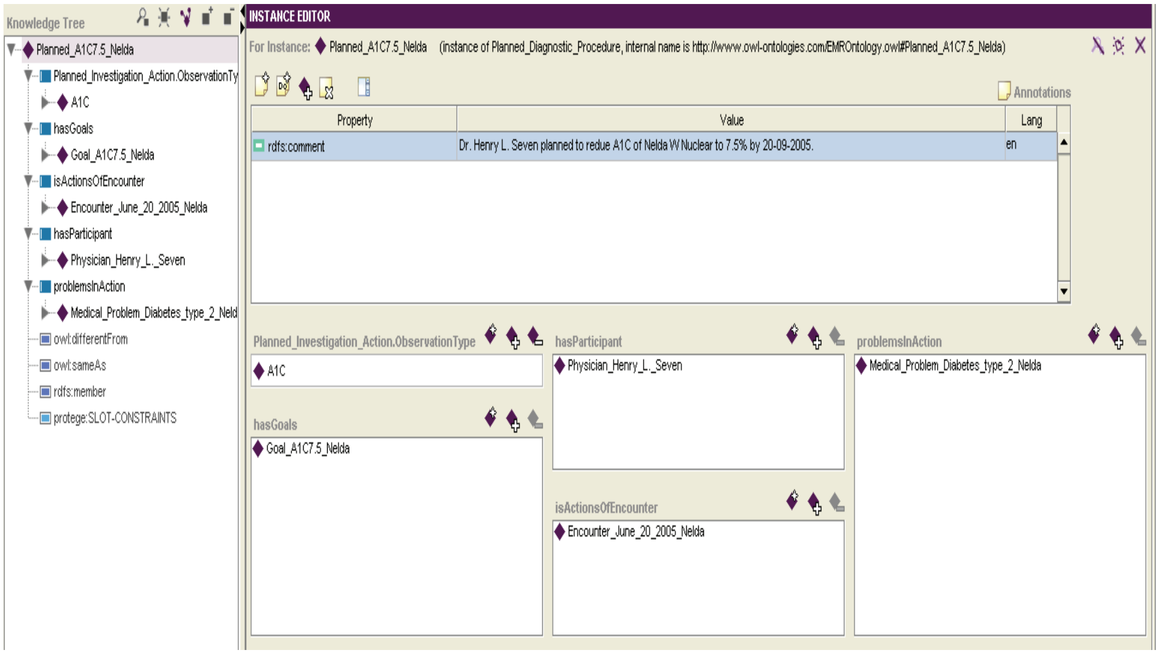
Metadata(EMROntology.owl)
OWLClasses
Properties
Individuals
Forms
Knowledge Acquisition
Knowledge Tree
OWL-DL Individuals
OWLViz
ProtegePropertiesTab
XMLTab
Knowledge Tree
Patient_Nelda_W_Nuclear
hasSubject
Encounter_June_20_2005_Nelda
resposibilityOf
Physician_Henry_L._Seven
isSubjectOf
hasActions
owl:differentFrom
owl:sameAs
rdfs:member
protege:SLOT-CONSTRAINTS
hasProblems
Medical_Problem_Ulcer_on_foot_Nelda
Medical_Problem_Glaucoma_Nelda
Medical_Problem_Phx_Amputation_Nelda
Medical_Problem_Diabetes_type_2_Nelda
Medical_Problem_Obesity_Nelda
Medical_Problem_Erectile_Dysfunction_Nelda
INSTANCE EDITOR
For Instance: Patient_Nelda_W_Nuclear (instance of Patient, internal name is http://www.owl-ontologies.com/EMROntology.owl#Patient_Nelda_W_Nuclear)
Annotations
Property
Value
Lang
rdfs:comment
Patient.DateTimeOfBirth
October 30, 1930
Patient.UniversalPatientHealthNumber
78657662
hasSubject
Encounter_June_20_2005_Nelda
Patient.Name
Nelda W Nuclear
Patinet.Sex
Male
Patient.PermanentAddress
hasProblems
Medical_Problem_Diabetes_type_2_Nelda
Medical_Problem_Erectile_Dysfunction_Nelda
Medical_Problem_Glaucoma_Nelda
Medical_Problem_Obesity_Nelda

(a)

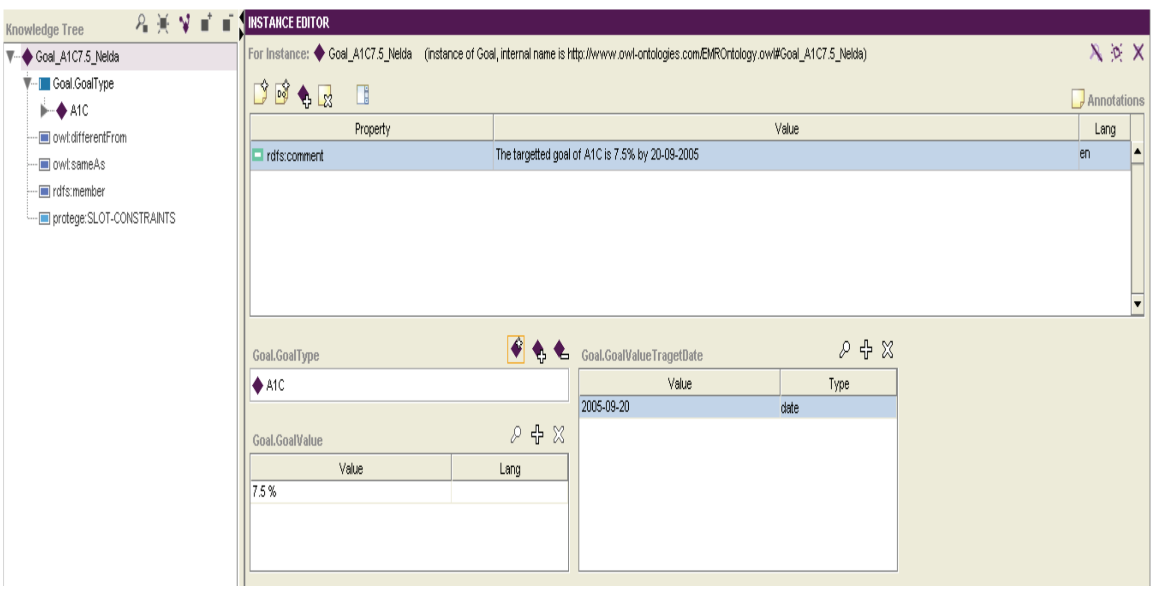


Knowledge Tree
Medical_Problem_Diabetes_type_2_Nelda
actionForProblem
Planned_Triglycerides_Nelda
Planned_LDLC_Nelda
Planned_HLDC_Nelda
Planned_TC_HDLC_Ratio_Nelda
Planned_A1C7.5_Nelda
Planned_TC_Nelda
Medical_Problem.ProblemName
problemOf
Patient_Nelda_W_Nuclear
hasProblems
hasSubject
resposibilityOf
hasActions
INSTANCE EDITOR
For Instance: Medical_Problem_Diabetes_type_2_Nelda (instance of Medical_Problem, internal name is http://www.owl-ontologies.com/EMROntology.owl#Medical_Problem_Diabetes_type_2_Nelda)
Annotations
Property
Value
Lang
rdfs:comment
Nelda W Nuclear has the main compaint of Diabetes Type 2, a chronic condition first identified in April 2005.
en
Medical_Problem.ProblemNumber
1
Medical_Problem.ProblemCurrentStatus
Value
Type
Active
string
actionForProblem
Planned_A1C7.5_Nelda
Planned_HLDC_Nelda
Planned_LDLC_Nelda
Planned_Procedure_Dietary_Management
Planned_Procedure_Review_Overall_Progress_Nelda
Planned_Procedure_Review_Sildenafil_Nelda
Planned_TC_HDLC_Ratio_Nelda
Planned_TC_Nelda
Planned_Therapeutic_Action_Elevation_Nelda
Planned_Triglycerides_Nelda
Substance_Administration_ACE_Inhibitor
Substance_Administration_Metformin
Medical_Problem.ProblemName
DiabetesType2
problemOf
Patient_Nelda_W_Nuclear
Medical_Problem.ProblemDateOfOnset
2005-04-30
date

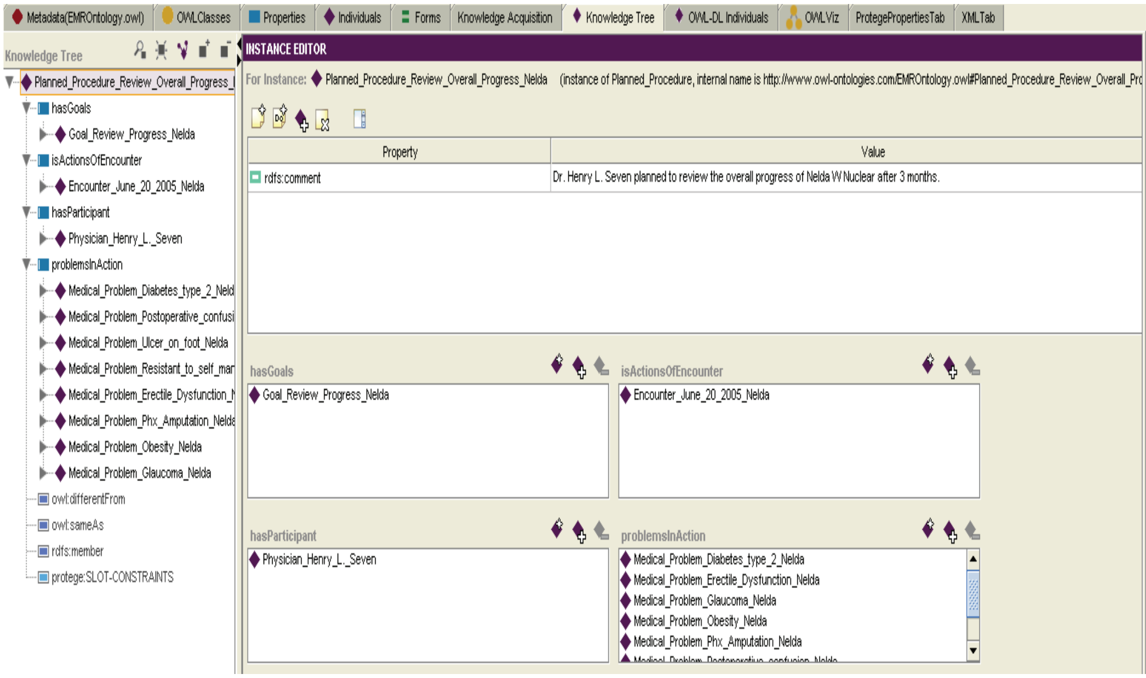
(b)


Knowledge Tree
INSTANCE EDITOR
For Instance: Planned_A1C7.5_Nelda (instance of Planned_Diagnostic_Procedure, internal name is http://www.owl-ontologies.com/EMROntology.owl#Planned_A1C7.5_Nelda)
Planned_A1C7.5_Nelda
Planned_Investigation_Action.ObservationTy
A1C
hasGoals
Goal_A1C7.5_Nelda
isActionsOfEncounter
Encounter_June_20_2005_Nelda
hasParticipant
Physician_Henry_L._Seven
problemsInAction
owl:differentFrom
owl:sameAs
rdfs:member
protege:SLOT-CONSTRAINTS
Annotations
Property
Value
Lang
rdfs:comment
Dr. Henry L. Seven planned to redue A1C of Nelda W Nuclear to 7.5% by 20-09-2005.
en
Planned_Investigation_Action.ObservationType
Medical_Problem_Diabetes_type_2_Nelda

(c)


Knowledge Tree
INSTANCE EDITOR
For Instance: Goal_A1C7.5_Nelda (instance of Goal, internal name is http://www.owl-ontologies.com/EMROntology.owl#Goal_A1C7.5_Nelda)
Goal_A1C7.5_Nelda
Goal.GoalType
A1C
owl:differentFrom
owl:sameAs
rdfs:member
protege:SLOT-CONSTRAINTS
Annotations
Property
Value
Lang
rdfs:comment
The targetted goal of A1C is 7.5% by 20-09-2005
en
Goal.GoalValueTragetDate
Type
2005-09-20
date
Goal.GoalValue
7.5 %

(d)



(e)

Figure 4.1 Examples of instantiation results for the medical record of [69] in the Protégé EMR ontology: (a) Patient, Nelda W Nuclear, (b) A medical problem, Diabetes Type 2, of patient Nelda, (c) Planned diagnostic procedure, A1C for Nelda, (d) Particular goal of A1C associated with planned diagnostic procedure, and (e) Planned overall review of Nelda after 3 months.

Figure 4.1 shows a few examples of instantiations for some of the important concepts in the EMR ontology. It starts with the instance of patient, Nelda W Nuclear in Figure 4.1 (a), shows its demographic information, as well as her medical problems and encounter with Dr. Henry L. Seven. Among 8 medical problems Nelda has, we are showing here the instantiation results only for Diabetes Type 2 in Figure 4.1 (b), which relates to the clinical actions carried out during current encounter particularly for this problem. Next, in Figure 4.1 (c), we have gone into details about one of these particular actions, i.e., planned diagnostic procedure, A1C. Figure 4.1 (d) shows the particular goal planned to be achieved for A1C, with date and value. Figure 4.1 (e) illustrates the instantiation for the planned review of overall progress of Nelda W Nuclear after 3 months, which not only is related to Diabetes Type 2, but to all of the problems.

Here, we have intended to show a few instantiation examples, from top-level concept (e.g., medical-problem) to in depth (e.g., a particular goal related to a planned action). The detail instantiation results for [69] are shown in Table 4.1.

Table 4.1 Detail instantiation result of the proposed EMR ontology for medical record of [69].

CLASS 'Individual'	***Property***	**Value / CLASS 'Individual'**
Patient "Nelda W Nuclear"	*PatientName*	Nelda W Nuclear
	Date-time of birth	10-1930
	Patient permanent address	Village Avenue, Some Town, AB T1T 1T1, Canada Tel: 15555555001
	Sex	Female
	UniversalPatientHealthNumber	7-865766-2
Practitioner "Henry L. Seven"	*ID*	1
	Name	Henry L. Seven
	Address	Home Health Care Clinic 6666 Home Street Some Town AB T1T 1T1 Canada
	responsibleOf	Encounter "June 20, 2005"
Encounter "June 20, 2005"	*EncounterDate*	20-06-2005
	EncounterPurpose	Checkup
	EncounterID	1
	subjectOf	Patient "Nelda W Nuclear"
	has Actions	Medical-History-Screening "Transexual Surgery" Planned-procedure "Re-emphasize need for self management", Planned-procedure "Re-emphasize need for self management", Planned-procedure "Medications in a "blister pack", Planned-referral "Ophthalmologist", Planned-procedure "Daily dressings", Planned-procedure "Proper foot care", Planned-procedure "Elevation", Planned-substance-administration "Rx aspirin and Zincaps", Planned-referral "ulcer clinic", Planned-substance-administration "Sildenafil", Planned-referral "dietitian", Planned-procedure "dietary management", Planned-clinical-exam "Weight", Planned-clinical-exam "Waist circumference", Planned-clinical-exam "BP", Planned-clinical-exam "sexual dysfunction", Planned-diagnostic-procedure "A1C", Planned-diagnostic-procedure "TC", Planned-diagnostic-procedure "HDL C", Planned-diagnostic-procedure "LDL C", Planned-diagnostic-procedure "TC ; HDL cholesterol ratio", Planned-diagnostic-procedure "Triglycerides", Planned-procedure "review overall progress", Planned-clinical-exam "BP", Planned-procedure "review slidenafil", Substance - administration "ACE Inhibitor", Substance - administration "Metformin", Substance - administration "Influenza vaccine"
medical-problem "Diabetes type 2"	*medical-problem.ProblemNumber*	1

CLASS 'Individual'	***Property***	**Value / CLASS 'Individual'**
	medical-problem.ProblemName	Diabetes mellitus type 2 (VocMedicalProblem = 'Diabetes mellitus type 2' where VocMedicalProblem.Code=' 44054006' and VocMedicalProblem.ConcetName= ' Diabetes mellitus type 2 (disorder)') [From SNOMED-CT]
	medical-problem.ProblemDateOfOnset	04-2005
	medical-problem.ProblemCurrentStatus	Active
	problemOf	Patient "Nelda W Nuclear"
medical-problem "Confusion"	*medical-problem.ProblemNumber*	2
	medical-problem.ProblemName	Postoperative confusion (VocMedicalProblem = 'Postoperative confusion' where VocMedicalProblem.Code=' 404906000' and VocMedicalProblem.ConcetName= ' Postoperative confusion (disorder)') [From SNOMED-CT]
	medical-problem.ProblemCurrentStatus	Active
	problemOf	Patient "Nelda W Nuclear"
medical-problem "Glaucoma"	*medical-problem.ProblemNumber*	3
	medical-problem.ProblemName	Glaucoma (VocMedicalProblem = 'Glaucoma' where VocMedicalProblem.Code=' 23986001' and VocMedicalProblem.ConcetName= ' Glaucoma (disorder)') [From SNOMED-CT]
	medical-problem.ProblemCurrentStatus	Active
	problemOf	Patient "Nelda W Nuclear"
medical-problem "Ulcer on foot"	*medical-problem.ProblemNumber*	4
	medical-problem.ProblemName	Ulcer on foot (VocMedicalProblem = 'Ulcer on foot' where VocMedicalProblem.Code=' 95345008' and VocMedicalProblem.ConcetName= ' Ulcer on foot (disorder)') [From SNOMED-CT]
	medical-problem.ProblemCurrentStatus	Active
	problemOf	Patient "Nelda W Nuclear"
medical-problem "Erectile dysfunction"	*medical-problem.ProblemNumber*	4
	medical-problem.ProblemName	Erectile dysfunction associated with type 2 diabetes mellitus (VocMedicalProblem = 'Erectile dysfunction associated with type 2 diabetes mellitus' where VocMedicalProblem.Code=' 428007007' and VocMedicalProblem.ConcetName= ' Erectile dysfunction associated with type 2 diabetes mellitus (disorder)') [From SNOMED-CT]
	medical-problem.ProblemCurrentStatus	Active
	problemOf	Patient "Nelda W Nuclear"
medical-problem "Phx Amputation"	*medical-problem.ProblemNumber*	5
	medical-problem.ProblemName	Traumatic amputation (VocMedicalProblem = 'Traumatic amputation' where VocMedicalProblem.Code=' 262595009' and VocMedicalProblem.ConcetName= ' Traumatic amputation (disorder)') [From SNOMED-CT]
	medical-problem.ProblemCurrentStatus	Active
	problemOf	Patient "Nelda W Nuclear"
medical-problem "Mild obesity – BMI =	*medical-problem.ProblemNumber*	6

CLASS 'Individual'	*Property*	Value / CLASS 'Individual'
30"	*medical-problem.ProblemName*	Obesity (VocMedicalProblem = 'Obesity' where VocMedicalProblem.Code=' 414916001' and VocMedicalProblem.ConcetName= ' Obesity (disorder)') [From SNOMED-CT]
	medical-problem.ProblemCurrentStatus	Active
	problemOf	Patient "Nelda W Nuclear"
medical-problem "Resistant to self management ideas"	*medical-problem.ProblemNumber*	7
	medical-problem.ProblemName	Health management deficit (VocMedicalProblem = 'Health management deficit' where VocMedicalProblem.Code=' 38334009' and VocMedicalProblem.ConcetName= ' Health management deficit (finding)') [From SNOMED-CT]
	medical-problem.ProblemCurrentStatus	Active
	problemOf	Patient "Nelda W Nuclear"
Medical-History-Screening "Transsexual Surgery"	*observationType*	Transsexual Surgery (VocObservationType = 'Surgically transgendered transsexual, female-to-male (finding)' where VocObservationType.Code=' 407379008' and VocObservationType.ConcetName= ' Surgically transgendered transsexual, female-to-male (finding)') [From SNOMED-CT]
	HistoryTakingEventDate	20-06-2005
	actsOn	Patient "Nelda W Nuclear"
Planned-procedure "Re-emphasize need for self management"	*procedureType*	Respiratory health self management education (VocObservationType = 'Respiratory health self management education' where VocObservationType.Code=' 427625003' and VocObservationType.ConcetName= ' Respiratory health self management education (procedure) (procedure)') [From SNOMED-CT]
	procedureDetails	education and skills development
	actionForProblem	medical-problem "Resistant to self management ideas"
	actsOn	
Planned-procedure "Re-emphasize need for self management"	*procedureType*	Respiratory health self management education (VocObservationType = 'Respiratory health self management education' where VocObservationType.Code=' 427625003' and VocObservationType.ConcetName= ' Respiratory health self management education (procedure)') [From SNOMED-CT]
	procedureDetails	education and skills development
	actionForProblem	medical-problem "Resistant to self management ideas"
	actsOn	Patient "Nelda W Nuclear"
Planned-procedure "Medications in a "blister pack"	*procedureType*	Medication education (VocObservationType = 'Medication education' where VocObservationType.Code=' 967006' and VocObservationType.ConcetName= ' Medication education (procedure)') [From SNOMED-CT]
	procedureDetails	Medications to be packaged in a "blister pack" – discuss with pharmacist
	actionForProblem	medical-problem "Resistant to self management ideas"
	actsOn	Patient "Nelda W Nuclear"
Planned-referral "Ophthalmologist"	*ReferralReason*	requires specialist management
	ReferredTo	Ophthalmologist
	actionForProblem	medical-problem "Glaucoma"
	actsOn	Patient "Nelda W Nuclear"
Planned-procedure "Daily dressings"	*procedureType*	Dressing of ulcer (VocObservationType = 'Dressing of ulcer' where VocObservationType.Code=' 182532000' and VocObservationType.ConcetName= ' Dressing of ulcer (procedure)') [From SNOMED-CT]

CLASS 'Individual'	***Property***	**Value / CLASS 'Individual'**
	procedureDetails	Daily dressings by visiting nurse
	actionForProblem	medical-problem "Ulcer on foot"
	actsOn	Patient "Nelda W Nuclear"
Planned-procedure "Proper foot care"	*procedureType*	Foot care (VocObservationType = 'Foot care' where VocObservationType.Code=' 385955003' and VocObservationType.ConcetName= ' Foot care (regime/therapy)') [From SNOMED-CT]
	procedureDetails	Proper foot care
	actionForProblem	medical-problem "Ulcer on foot"
	actsOn	Patient "Nelda W Nuclear"
Planned-procedure "Elevation"	*procedureType*	Elevation (VocObservationType = 'Elevation' where VocObservationType.Code=' 103720008' and VocObservationType.ConcetName= ' Elevation (Procedure)') [From SNOMED-CT]
	procedureDetails	Elevation
	actionForProblem	medical-problem "Ulcer on foot"
	actsOn	Patient "Nelda W Nuclear"
Planned-substance-administration "Rx aspirin and Zincaps"	*MedicationOrVaccinationType*	Medication
	MedicationOrVaccinationName	Aspirin and Zincaps
	MedicationOrVaccinationInstruction	Rx aspirin 150 mg / day and Zincaps one daily to medication regimen
	actionForProblem	medical-problem "Ulcer on foot"
	actsOn	Patient "Nelda W Nuclear"
Planned-referral "ulcer clinic"	*ReferralReason*	If no improvement found
	ReferredTo	Ulcer clinic
	actionForProblem	medical-problem "Ulcer on foot"
	actsOn	Patient "Nelda W Nuclear"
	ReferralReason	requires specialist management
Planned-substance-administration "Sildenafil"	*MedicationOrVaccinationType*	Medication
	MedicationOrVaccinationName	Sildenafil
	actionForProblem	medical-problem "Erectile dysfunction"
Planned-referral "dietitian"	*ReferralReason*	Obesity
	ReferredTo	Dietitian
	actionForProblem	medical-problem "Obesity"
	actsOn	Patient "Nelda W Nuclear"
Planned-procedure "dietary management"	*procedureType*	Dietary regime management (VocObservationType = 'Dietary regime management' where VocObservationType.Code=' 410175003' and VocObservationType.ConcetName= ' Dietary regime management (Procedure)') [From SNOMED-CT]
	procedureDetails	Reduction and consistency in total carbohydrate intake, Increase intakes of grains and complex CHO, Reduce fat intake, Maintain a diary of food intake, Attempt to increase physical activity
	actionForProblem	medical-problem "Obesity"
	actsOn	Patient "Nelda W Nuclear"
Goal "Weight: 85 kg"	*GoalType*	Weight (VocObservationType = 'Weight' where VocObservationType.Code=' 27113001' and VocObservationType.ConcetName= ' Body weight (observable entity)') [From SNOMED-CT]
	GoalValue	85 kg

CLASS 'Individual'	***Property***	**Value / CLASS 'Individual'**
	GoalValueTargetDate	20-09-2005
Planned-clinical-exam "Weight"	*hasGoal*	Goal "Weight: 85 kg"
	actsOn	Patient "Henry Levin , the 7th"
Goal " Waist circumference 38cm"	*GoalType*	Waist circumference (VocObservationType = Waist circumference' where VocObservationType.Code=' 276361009' and VocObservationType.ConcetName= ' Waist circumference (observable entity)') [From SNOMED-CT]
	GoalValue	38 cm
	GoalValueTargetDate	20-09-2005
Planned-clinical-exam "Waist circumference"	*hasGoal*	Goal " Waist circumference 38cm"
	actsOn	Patient "Henry Levin , the 7th"
Goal " BP 140 /990mmHg"	*GoalType*	Blood Pressure (VocObservationType = 'Blood Pressure' where VocObservationType.Code=' 75367002' and VocObservationType.ConcetName= ' Blood Pressure (observable entity)') [From SNOMED-CT]
	GoalValue	140 /990mmHg
	GoalValueTargetDate	20-09-2005
Planned-clinical-exam "BP"	*hasGoal*	Goal " BP 140 /990mmHg"
	actsOn	Patient "Henry Levin , the 7th"
Goal " improvement of Sexual dysfunction"	*GoalType*	Sexual function (VocObservationType = 'Sexual function' where VocObservationType.Code=' 76859005' and VocObservationType.ConcetName=' Sexual function (observable entity)') [From SNOMED-CT]
	GoalValue	improvement by slidenafil
	GoalValueTargetDate	20-09-2005
Planned-clinical-exam "sexual dysfunction"	*hasGoal*	Goal " improvement of Sexual dysfunction"
	actsOn	Patient "Henry Levin , the 7th"
Goal " A1C: 7.5%"	*GoalType*	A1C (VocObservationType = 'A1C' where VocObservationType.Code=' 408591000' and VocObservationType.ConcetName= 'Hemoglobin A1c (HBA1c) target (observable entity)') [From SNOMED-CT]
	GoalValue	7.5%
	GoalValueTargetDate	20-09-2005
Planned-diagnostic-procedure "A1C"	*hasGoal*	Goal " A1C: 7.5%"
	actsOn	Patient "Henry Levin , the 7th"
Goal " TC 6.0 mmol /L"	*GoalType*	TC (VocObservationType = 'TC' where VocObservationType.Code=' 390896004' and VocObservationType.ConcetName= 'Target cholesterol level (observable entity) (observable entity)') [From SNOMED-CT]
	GoalValue	6.0 mmol /L
	GoalValueTargetDate	20-09-2005
Planned-diagnostic-procedure "TC"	*hasGoal*	Goal " TC 6.0 mmol /L"
	actsOn	Patient "Henry Levin , the 7th"
Goal " HDL C 1.5 mmol / L"	*GoalType*	HDLC (VocObservationType = 'HDLC' where VocObservationType.Code=' 390896004' and VocObservationType.ConcetName= 'Target cholesterol level (observable entity) (observable entity)') [From SNOMED-CT]
	GoalValue	1.5 mmol / L
	GoalValueTargetDate	20-09-2005
Planned-diagnostic-procedure "HDL C"	*hasGoal*	Goal " HDL C 1.5 mmol / L"
	actsOn	Patient "Henry Levin , the 7th"
Goal " LDL C <2.6 mmol /L"	*GoalType*	TC (VocObservationType = 'TC' where VocObservationType.Code=' 390896004' and VocObservationType.ConcetName= 'Target cholesterol level (observable entity) (observable entity)') [From SNOMED-CT]
	GoalValue	<2.6 mmol /L

CLASS 'Individual'	***Property***	**Value / CLASS 'Individual'**
	GoalValueTargetDate	20-09-2005
Planned-diagnostic-procedure "LDL C"	*hasGoal*	Goal " LDL C <2.6 mmol /L"
	actsOn	Patient "Henry Levin , the 7th"
Goal " TC ; HDL cholesterol ratio 5"	*GoalType*	TC:HDL cholesterol ratio (VocObservationType = 'TC:HDL cholesterol ratio' where VocObservationType.Code=' 390896004' and VocObservationType.ConcetName= 'Target cholesterol level (observable entity) (observable entity)') [From SNOMED-CT]
	GoalValue	5
	GoalValueTargetDate	20-09-2005
Planned-diagnostic-procedure "TC ; HDL cholesterol ratio"	*hasGoal*	Goal " TC ; HDL cholesterol ratio 5"
	actsOn	Patient "Henry Levin , the 7th"
Goal " Triglycerides <1.9mmol / l"	*GoalType*	Triglycerides (VocObservationType = 'Triglycerides' where VocObservationType.Code=' 14740000' and VocObservationType.ConcetName= 'Triglycerides measurement (procedure)') [From SNOMED-CT]
	GoalValue	<1.9mmol / l
	GoalValueTargetDate	20-09-2005
Planned-diagnostic-procedure "Triglycerides"	*hasGoal*	Goal " Triglycerides <1.9mmol / l"
	actsOn	Patient "Henry Levin , the 7th"
Planned-procedure "review overall progress"	*procedureType*	Review of care plan (VocObservationType = 'Review of care plan' where VocObservationType.Code=' 425268008' and VocObservationType.ConcetName= ' Review of care plan (Procedure)') [From SNOMED-CT]
	procedureDetails	review overall progress after 3 months
	actsOn	Patient "Nelda W Nuclear"
Planned-clinical-exam "BP"	*observationType*	Blood Pressure (VocObservationType = 'Blood Pressure' where VocObservationType.Code=' 75367002' and VocObservationType.ConcetName= ' Blood Pressure (observable entity)') [From SNOMED-CT]
	examSummary	Weekly
	actsOn	Patient "Nelda W Nuclear"
Planned-procedure "review slidenafil"	*procedureType*	Review of medication (VocObservationType = 'Review of medication' where VocObservationType.Code=' 182836005' and VocObservationType.ConcetName= ' Review of medication (Procedure)') [From SNOMED-CT]
	procedureDetails	review slidenafil after 3 months
	actsOn	Patient "Nelda W Nuclear"
Substance - administration "ACE Inhibitor"	*MedicationOrVaccination Type*	Medication
	MedicationOrVaccinationName	ACE Inhibitor
	actsOn	Patient "Nelda W Nuclear"
Substance - administration "Metformin"	*MedicationOrVaccination Type*	Medication
	MedicationOrVaccinationName	Metformin
	actsOn	Patient "Nelda W Nuclear"
Substance - administration "Influenza vaccine"	*MedicationOrVaccination Type*	Vaccination
	MedicationOrVaccinationName	Influenza vaccine
	actsOn	Patient "Nelda W Nuclear"

4.4.2 Instantiation for the second medical record

The second medical record [70] was proposed by the HL7 Continuity of Care Record work group. This medical record is also written in HL7 version 3. The interesting thing is that by instantiating this medical record into our EMR ontology, we can test whether the ontology works for medical records in some forms other than the CDM Model. Since, the example medical record in [70] aims to represent Continuty of Care Record, and basically represents a patient having active problems with Asthma, this best suit with our work.

The medical record represents a 67 years male patient, Henry Leven the Seventh, from MA, who has active medical problems of Asthma, Pneumonia, and Myocardial Infarction. He went to his health care provider, Good Health Clinic, on 4th April, 2000. During that encounter with Henry, the responsible clinician took information about his next to kin, screened family history (e.g., father died because of Myocardial Infarction), medical history (e.g., Henry smoked 1 pack cigarette per day during 1947-1972). The clinician also performed some physical examinations (e.g., height, weight, blood pressure etc.), ordered some diagnostic procedures (e.g., WBC, K, CL etc.), and suggested some medications (e.g., Albuterol inhalant), and vaccinations (e.g., Influenza virus). The doctor ordered for total left hip replacement, and planned to perform the Pulmonary function test to better understand the current situation of Asthma. The detail results of the instantiation are shown in Table 4.2.

Table 4.2 Detail instantiation result of the proposed EMR ontology for medical record of [70].

CLASS 'Individual'	***Property***	**Value / CLASS 'Individual'**
Patient "Henry Levin , the 7th"	PatientName	Henry Levin , the 7th
	Date-time of birth	24-09-1932
	Date of earliest held entry	07-04-2000
	Patient permanent address	Kenneth Ross 17 Daws Rd. Blue Bell, MA, 02368 Tel: (888) 555-1212
	Sex	Male
	UniversalPatientHealthNumber	996-756-495
	hasRelative	Relative "Henrietta Levin"
Relative "Henrietta Levin"	RelativeName	Henrietta Levin
	RelativeAddress	Tel:(999)555-1212
	RelationshipToPatient	Next to Kin

CLASS 'Individual'	*Property*	Value / CLASS 'Individual'
	hasRelationhipWith	Patient "Henry Levin , the 7th"
Provider "Good Health Clinic"	*ID*	1
	Name	Good Health Clinic
	responsibleOf	Encounter "April 7, 2000"
Encounter "April 7, 2000"	*EncounterDate*	07-04-2000
	EncounterPurpose	Transfer of care
	EncounterID	1
	subjectOf	Patient "Henry Levin , the 7th"
	has Actions	Family History "Father"→Screening-act, Family History "Mother"→Screening-act, Medical-History-Screening "Cigarette Smoking", Medical-History-Screening "Alcohol consumption", Medical-History-Screening "Adverse Reaction--Penicillin", Medical-History-Screening "Adverse Reaction--Aspirin", Medical-History-Screening "Adverse Reaction-- Codein", Substance - administration "Albuterol inhalant", Substance - administration "Clopidogrel (Plavix)", Substance - administration "Metoprolol", Substance - administration "Prednisone", Substance - administration "Cephalexin (Keflex)", Substance - administration "Influenza virus vaccine", Substance - administration "Influenza virus vaccine", Substance - administration "Pneumococcal polysaccharide vaccine", Substance - administration "Tetanus and diphtheria toxoids", Clinical-Exam "Height: 14 Nov, 1999", Clinical-Exam "Height:7 April, 2000", Clinical-Exam "Weight: 14 Nov, 1999", Clinical-Exam "Weight: 7 Apr, 2000", Clinical-Exam "Systolic blood pressure: 14 Nov, 1999", Clinical-Exam "Diastolic blood pressure: 14 Nov, 1999", Clinical-Exam "Systolic blood pressure: 7 Apr, 2000", Clinical-Exam "Diastolic blood pressure: 7 Apr, 2000", Diagnostic-procedure "HGB", Diagnostic-procedure "WBC", Diagnostic-procedure "PLT", Diagnostic-procedure "NA", Diagnostic-procedure "K", Diagnostic-procedure "CL", Procedure "Total hip replacement, left", Planned-diagnostic-procedure "Pulmonary function test"
medical-problem "Asthma"	*medical-problem.ProblemNumber*	1
	medical-problem.ProblemName	Asthma (VocMedicalProblem = 'Asthma' where VocMedicalProblem.Code=' 195967001' and VocMedicalProblem.ConcetName= ' Asthma (disorder)') [From SNOMED-CT]
	medical-problem.ProblemDateOfOnset	1950
	medical-problem.ProblemCurrentStatus	Active
	problemOf	Patient "Henry Levin , the 7th"
medical-problem	*medical-problem.ProblemNumber*	2

CLASS 'Individual'	***Property***	**Value / CLASS 'Individual'**
"Pneumonia"	*medical-problem.ProblemName*	Pneumonia (VocMedicalProblem = 'Pneumonia' where VocMedicalProblem.Code=' 233604007' and VocMedicalProblem.ConcetName= ' Pneumonia (disorder)') [From SNOMED-CT]
	medical-problem.ProblemDateOfOnset	Jan 1997
	medical-problem.ProblemCurrentStatus	Resolved
	problemOf	Patient "Henry Levin , the 7th"
medical-problem "Myocardial Infarction"	*medical-problem.ProblemNumber*	3
	medical-problem.ProblemName	Myocardial Infarction (VocMedicalProblem = 'Myocardial Infarction' where VocMedicalProblem.Code=' 22298006' and VocMedicalProblem.ConcetName= ' Myocardial Infarction (disorder)') [From SNOMED-CT]
	medical-problem.ProblemDateOfOnset	Jan 1997
	medical-problem.ProblemCurrentStatus	Resolved
	problemOf	Patient "Henry Levin , the 7th"
Family History "Father"→Screening-act	*observationType*	Family History (VocObservationType = 'Family History where VocObservationType.Code=' 405205002' and VocObservationType.ConcetName= ' Family health status (observable entity)') [From SNOMED-CT]
	HistoryRelationshipSourceToPatient	Father
	Value	Myocardial Infarction (cause of death), age at onset:57; Hypertension, age at onset:40
	actsOn	Patient "Henry Levin , the 7th"
Family History "Mother"→Screening-act	*observationType*	Family History (VocObservationType = 'Family History where VocObservationType.Code=' 405205002' and VocObservationType.ConcetName= ' Family health status (observable entity)') [From SNOMED-CT]
	HistoryRelationshipSourceToPatient	Mother
	Value	Alive, Asthma, age at onset:30
	actsOn	Patient "Henry Levin , the 7th"
Medical-History-Screening "Cigarette Smoking"	*observationType*	Cigarette consumption (VocObservationType = 'Cigarette consumption' where VocObservationType.Code=' 230056004' and VocObservationType.ConcetName= ' Cigarette consumption (observable entity)') [From SNOMED-CT]
	HistoryTakingEventDate	07-04-2000
	Value	1 pack per day from 1947-1972
	actsOn	Patient "Henry Levin , the 7th"
Medical-History-Screening "Alcohol consumption"	*observationType*	Alcohol consumption (VocObservationType = 'Alcohol consumption' where VocObservationType.Code=' 160580001' and VocObservationType.ConcetName= ' Alcohol consumption unknown (finding)') [From SNOMED-CT]
	HistoryTakingEventDate	07-04-2000
	Value	From 1973
	actsOn	Patient "Henry Levin , the 7th"
Medical-History-Screening "Adverse Reaction--Penicillin"	*observationType*	Adverse drug reaction (VocObservationType = 'Adverse reaction' where VocObservationType.Code=' 281647001' and VocObservationType.ConcetName= ' Adverse reaction (disorder)') [From SNOMED-CT]
	HistoryTakingEventDate	07-04-2000
	Value	Hives on Penicillin, status:active

CLASS 'Individual'	***Property***	**Value / CLASS 'Individual'**
	actsOn	Patient "Henry Levin , the 7th"
Medical-History-Screening "Adverse Reaction--Aspirin"	*observationType*	Adverse drug reaction (VocObservationType = 'Adverse reaction' where VocObservationType.Code=' 281647001' and VocObservationType.ConcetName= ' Adverse reaction (disorder)') [From SNOMED-CT]
	HistoryTakingEventDate	07-04-2000
	Value	Wheezing on Aspirin, status:active
	actsOn	Patient "Henry Levin , the 7th"
Medical-History-Screening "Adverse Reaction-- Codein"	*observationType*	Adverse drug reaction (VocObservationType = 'Adverse reaction' where VocObservationType.Code=' 281647001' and VocObservationType.ConcetName= ' Adverse reaction (disorder)') [From SNOMED-CT]
	HistoryTakingEventDate	07-04-2000
	Value	Nausea on Codein, status:active
	actsOn	Patient "Henry Levin , the 7th"
Substance - administration "Albuterol inhalant"	*MedicationOrVaccination Type*	Medication
	MedicationOrVaccinationName	Albuterol inhalant
	MedicationOrVaccinationInstructions	2 puffs QID PRN wheezing
	Status	Not Found
	actsOn	Patient "Henry Levin , the 7th"
Substance - administration "Clopidogrel (Plavix)"	*MedicationOrVaccination Type*	Medication
	MedicationOrVaccinationName	Clopidogrel (Plavix)
	MedicationOrVaccinationInstructions	75mg PO daily
	Status	Not Found
	actsOn	Patient "Henry Levin , the 7th"
Substance - administration "Metoprolol"	*MedicationOrVaccination Type*	Medication
	MedicationOrVaccinationName	Metoprolol
	MedicationOrVaccinationInstructions	25mg PO BID
	Status	Not Found
	actsOn	Patient "Henry Levin , the 7th"
Substance - administration "Prednisone	*MedicationOrVaccination Type*	Medication
	MedicationOrVaccinationName	Prednisone
	MedicatioInstructions	20mg PO daily
	MedicationOrVaccinationDate	28-03-2000
	Status	Not Found
	actsOn	Patient "Henry Levin , the 7th"
Substance - administration "Cephalexin (Keflex)"	*MedicationOrVaccination Type*	Medication
	MedicationOrVaccinationName	Cephalexin (Keflex)
	MedicationOrVaccinationInstructions	7 500mg PO QID x 7 days (for bronchitis)
	MedicationOrVaccinationDate	28-03-2000
	Status	Not Found
	actsOn	Patient "Henry Levin , the 7th"
MaterialEntity "Automatic implantable cardioverter/defibrillator"	*MaterialEntityName*	Automatic implantable cardioverter/defibrillator
	DateSupplied	Not Found
MaterialEntity "Total hip replacement prosthesis"	*MaterialEntityName*	Total hip replacement prosthesis
	DateSupplied	Not Found
MaterialEntity "Wheelchair"	*MaterialEntityName*	Wheelchair
	DateSupplied	Not Found
Substance - administration "Influenza virus vaccine"	*MedicationOrVaccination Type*	Vaccination
	MedicationOrVaccinationName	Influenza virus vaccine
	MedicationOrVaccinationDate	11-1999
	Status	Not Found
	Source of Information	Not Found
	actsOn	Patient "Henry Levin , the 7th"
Substance - administration "Influenza virus vaccine"	*MedicationOrVaccination Type*	Vaccination
	MedicationOrVaccinationName	Influenza virus vaccine

CLASS 'Individual'	*Property*	Value / CLASS 'Individual'
	MedicationOrVaccinationDate	12-1998
	Status	Not Found
	Source of Information	Not Found
	actsOn	Patient "Henry Levin , the 7th"
Substance - administration "Pneumococcal polysaccharide vaccine"	*MedicationOrVaccination Type*	Vaccination
	MedicationOrVaccinationName	Pneumococcal polysaccharide vaccine
	MedicationOrVaccinationDate	12-1998
	Status	Not Found
	Source of Information	Not Found
	actsOn	Patient "Henry Levin , the 7th"
Substance - administration "Tetanus and diphtheria toxoids"	*MedicationOrVaccination Type*	Vaccination
	MedicationOrVaccinationName	Tetanus and diphtheria toxoids
	MedicationOrVaccinationDate	1997
	Status	Not Found
	Source of Information	Not Found
	actsOn	Patient "Henry Levin , the 7th"
Clinical-Exam "Height: 14 Nov, 1999"	*observationType*	Height (VocObservationType = 'Height' where VocObservationType.Code=' 271603002' and VocObservationType.ConcetName= ' Height / growth measure (observable entity)') [From SNOMED-CT]
	DateTimeOfExam	14-11-1999
	ExamFindings	177 cm
	actsOn	Patient "Henry Levin , the 7th"
Clinical-Exam "Height:7 April, 2000"	*observationType*	Height (VocObservationType = 'Height' where VocObservationType.Code=' 271603002' and VocObservationType.ConcetName= ' Height / growth measure (observable entity)') [From SNOMED-CT]
	DateTimeOfExam	07-04-2000
	ExamFindings	177 cm
	actsOn	Patient "Henry Levin , the 7th"
Clinical-Exam "Weight: 14 Nov, 1999"	*observationType*	Weight (VocObservationType = 'Weight' where VocObservationType.Code=' 27113001' and VocObservationType.ConcetName= ' Body weight (observable entity)') [From SNOMED-CT]
	DateTimeOfExam	14-11-1999
	ExamFindings	86 kg
	actsOn	Patient "Henry Levin , the 7th"
Clinical-Exam "Weight: 7 Apr, 2000"	*observationType*	Weight (VocObservationType = 'Weight' where VocObservationType.Code=' 27113001' and VocObservationType.ConcetName= ' Body weight (observable entity)') [From SNOMED-CT]
	DateTimeOfExam	07-04-2000
	ExamFindings	88 kg
	actsOn	Patient "Henry Levin , the 7th"
Clinical-Exam "Systolic blood pressure: 14 Nov, 1999"	*observationType*	Systolic blood pressure (VocObservationType = 'Systolic blood pressure' where VocObservationType.Code=' 271649006' and VocObservationType.ConcetName= ' Systolic blood pressure (observable entity)') [From SNOMED-CT]
	DateTimeOfExam	14-11-1999
	ExamFindings	132 mmHg
	actsOn	Patient "Henry Levin , the 7th"
Clinical-Exam "Diastolic blood pressure: 14 Nov, 1999"	*observationType*	Diastolic blood pressure (VocObservationType = 'Diastolic blood pressure' where VocObservationType.Code=' 271650006' and VocObservationType.ConcetName= ' Diastolic blood pressure (observable entity)') [From SNOMED-CT]
	DateTimeOfExam	14-11-1999
	ExamFindings	86 mmHg

CLASS 'Individual'	***Property***	**Value / CLASS 'Individual'**
	actsOn	Patient "Henry Levin , the 7th"
Clinical-Exam "Systolic blood pressure: 7 Apr, 2000"	*observationType*	Systolic blood pressure (VocObservationType = 'Systolic blood pressure' where VocObservationType.Code=' 271649006' and VocObservationType.ConcetName= ' Systolic blood pressure (observable entity)') [From SNOMED-CT]
	DateTimeOfExam	07-04-2000
	ExamFindings	145 mmHg
	actsOn	Patient "Henry Levin , the 7th"
Clinical-Exam "Diastolic blood pressure: 7 Apr, 2000"	*observationType*	Diastolic blood pressure (VocObservationType = 'Diastolic blood pressure' where VocObservationType.Code=' 271650006' and VocObservationType.ConcetName= ' Diastolic blood pressure (observable entity)') [From SNOMED-CT]
	DateTimeOfExam	07-04-2000
	ExamFindings	88 mmHg
	actsOn	Patient "Henry Levin , the 7th"
Diagnostic-procedure "HGB"	*Test Requested*	HGB
	TestDateTime	14-11-1999
	NumericMeasurementOrAnalyteValue	13.2
	NumericMeasurementOrAnalyteName	g/dl
	actsOn	Patient "Henry Levin , the 7th"
Diagnostic-procedure "WBC"	*Test Requested*	WBC
	TestDateTime	14-11-1999
	NumericMeasurementOrAnalyteValue	6.7
	NumericMeasurementOrAnalyteName	ul
	actsOn	Patient "Henry Levin , the 7th"
Diagnostic-procedure "PLT"	*Test Requested*	PLT
	TestDateTime	14-11-1999
	NumericMeasurementOrAnalyteValue	123
	NumericMeasurementOrAnalyteName	meq/l
	actsOn	Patient "Henry Levin , the 7th"
Diagnostic-procedure "NA"	*Test Requested*	NA
	TestDateTime	07-04-2000
	NumericMeasurementOrAnalyteValue	140
	NumericMeasurementOrAnalyteName	meq/l
	actsOn	Patient "Henry Levin , the 7th"
Diagnostic-procedure "K"	*Test Requested*	K
	TestDateTime	07-04-2000
	NumericMeasurementOrAnalyteValue	4
	TestComments	meq/l
	actsOn	Patient "Henry Levin , the 7th"
Diagnostic-procedure "CL"	*Test Requested*	CL
	TestDateTime	07-04-2000
	NumericMeasurementOrAnalyteValue	102
	TestComments	Meq/l
	actsOn	Patient "Henry Levin , the 7th"
Procedure "Total hip replacement, left"	*ProcedureName*	Total hip replacement, left
	ProcedureStartDate	1998
	actsOn	Patient "Henry Levin , the 7th"
	actionForProblem	medical-problem "Asthma"
Planned-diagnostic-procedure "Pulmonary function test"	*Test Requested*	Pulmonary function test
	PlannedReassessmentDate	21-04-2000
	actsOn	Patient "Henry Levin , the 7th"

CLASS 'Individual'	*Property*	Value / CLASS 'Individual'
	actionForProblem	medical-problem "Asthma"

4.4 SUMMARY

The WHIC-proposed sample HL7 message was successfully instantiated into our EMR ontology. This is justified, since we are using the proposed CDM model as an underlying mapping model in our proposed methodology. However, the instantiation of the example medical record for the HL7 Continuity of Care Record [69] was particularly interesting because of the need to evaluate whether the proposed ontology can capture the medical records of some other formats. The instantiation results for this record are summarized in Table 4.2.

As shown in Table 4.2, we observed that most of the concepts of the example were successfully instantiated into our proposed ontology. While we encountered that the medication or immunization status (e.g., Active) could not be captured into our ontology, we could encode most of the clinical terms, with a few exceptions (e.g., the diagnostic procedure, HCO3). However, we believe that this is because we did not incorporate those clinical terms as vocabularies into our ontology, and there is scope (e.g., ObservationType) to add those, if needed.

CHAPTER 5 CONCLUSION

5.1 Thesis Contributions

This thesis makes the following contributions towards the EMR ontology for chronic disease management.

a) EMR ontology for CDM

We provided a POMR-based EMR ontology for chronic disease management. This ontology is capable of representing all the terminologies expected from an EMR in a clinical setting and needed for successful chronic disease management. Furthermore, we ensured structured data entry by integrating controlled vocabulary from SNOMED-CT.

b) Mapping results of EMR ontology with HL7 RIM

We believe that a harmoinization between clinical messaging standard (e.g., HL7) and an EMR is important. This thesis provides some interesting mapping results of such a harmonization between RIM and our EMR ontology. Since our EMR ontology is based on the POMR standard, as do some other EMR standards (e.g., open-EHR, CEN-EN 13606, ASTM EHR), these results are not limited only for this EMR structure, but rather can be used for similar mappings of other EMR standards with HL7.

c) A uniform EMR structure for various standards

Our proposed EMR ontology can be used as a switching language among various EMR standards. We mapped ASTM-EHR with our proposed EMR ontology in this thesis. Similarly, some other POMR-based EMR standards (e.g., open-EHR, CEN-EN 13606) can be easily mapped with it.

5.2 Limitations And Future Directions

This thesis has a number of limitations. First, the proposed EMR ontology holds data elements necessary only for three chronic diseases (i.e., Diabetes, Hypertension and Chronic Kidney Disease). It would be interesting to incorporate coded data elements for other chronic disease in future research.

The mapping procedures have been conducted fully by applying the domain knowledge on the mapping ontologies. It would be interesting to investigate the possibilities of applying one or more existing semi-automated ontology mapping tools for this purpose.

We evaluated our proposed ontology by instantiating with two medical records, both written in HL7. Both of these two medical records are organized in a Problem-Oriented way. Additional medical records, possibly from other standards, should also be instantiated to see whether the proposed ontology can capture data from them.

While mapping with HL7 RIM, we have observed some differences between it and our proposed ontology from an ontological point of view. For example, the proposed ontology differentiates between continuants and occurants, whereas HL7 RIM does not. We realize that some further technical implementations are necessary to reason over HL7 messages automatically and to fit RIM entities into appropriate concepts (e.g., Observation of HL7 RIM into clinical-examination, diagnostic-procedure, etc.).

5.3 Conclusion

We have prototyped a patient-centric, longitudinal Electronic Medical Record ontology focusing mainly on chronic disease management. Since the ontology is based on the POMR information model, it can also be used for the treatment of other acute diseases. This ontology ensures structured data entry by using SNOMED-CT-controlled vocabulary codes. We have successfully mapped the proposed ontology onto HL7 RIM to ensure that clinical messages would be successfully captured by this ontology. As the ontology is implemented in OWL-DL, decision-support systems can be implemented through reasoning over the Description Logic (DL) representation. The evaluation results show that our proposed ontology can capture the elements of clinical records and has the capability of representing the knowledge on a patient's medical records.

Although the ontology is currently capable of holding data elements for three special chronic diseases (i.e., Diabetes, Hypertension and Chronic Kidney Disease), we believe its structure is sufficiently general and flexible to incorporate elements for other chronic diseases.

The research reported in this thesis is part of a larger project to create clinical decision-support systems for chronic diseases, instantiated by EMRs. Our EMR ontology will not only support this project but can also be used as a switching language among various EMR standards.

BIBLIOGRAPHY

[1] World Health Organization, “Chronic Diseases”, http://www.who.int/topics/chronic_diseases/en/, retrieved on November 23, 2010.

[2] World Health Organization, “Preventing CHRONIC DISEASES – a vital investment”, http://www.who.int/chp/chronic_disease_report/contents/foreword.pdf, retrieved on November 23, 2010.

[3] Public Health Agency of Canada, “Chronic Diseases in Canada”, http://www.phac-aspc.gc.ca/publicat/cdic-mcc/29-4/ar_06-eng.php, retrieved on November 23, 2010.

[4] Improving Chronic Illness Care, “The Chronic Care Model”, http://www.improvingchroniccare.org/index.php, retrieved on November 23, 2010.

[5] A Health Information Network for Australia, Commonwealth of Australia. National Electronic Health Record Taskforce. ISBN 0 642 44668 7. July 2000.

[6] OHIH:2001, Office of Health and the Information Highway, Health Canada. “Tactical Plan for a pan-Canadian Health Inforstructure”, 2001 Update.

[7] CPRI:1995, Description of the Computer-based Patient Record (CPR) and Computer-based Patient Record System. Computer-based Patient Record Institute, May 1995.

[8] Garets, D. and Davis, M., Electronic Medical Records vs. Elec-tronic Health Records: Yes, There is a Difference; White Paper, August 26, 2005, HIMSS Analytics LLC.

[9] http://ezinearticles.com/?History-Of-Electronic-Medical-Records&id=254240, retrieved on November 23, 2010.

[10] Walker J., Pan E., Johnston D., Adler-Milstein J., Bates D., Middleton B. The Value Of Health Care Information Exchange And Interoperability. Health Affairs. Web Exclusive, January 19, 2005.

[11] Adapted from the IEEE definition of interoperability, and legal definitions used by the FCC (47 CFR 51.3), in statutes regarding copyright protection (17 USC 1201), and e-government services (44 USC 3601).

[12] "Confidentiality, Data Security, and Cancer Research: Perspectives from the National Cancer Institute", http://www3.cancer.gov/confidentiality.html, retrieved on November 23, 2010.

[13] Hallvard Lærum, MD, Tom H. Karlsen, MD, and Arild Faxvaag, MD, PhD (2003). "Effects of Scanning and Eliminating Paper-based Medical Records on Hospital Physicians' Clinical Work Practice". *Journal of the American Medical Informatics Association* **10**: 588–595.

[14] Johnston, Doughlas, et al. "The Value of Computerize Provider Order Entry in Ambulatory Settings: Executive Preview." Wellesley, MA: Center for Information Technology Leadership, 2003.

[15] Hillestad, Richard et al.: "Can Electronic Medical Record Systems Transform Health Care? Potential Health Benefits, Savings, and Costs", Health Affairs, 2005.

[16] Health Insurance Portability and Accountability Act (HIPAA) http://www.hipaa.org/ , retrieved on November 23, 2010.

[17] Clayton L. Reynolds MD, FACP, FACPE (March 2006): Paper on Concept Processing Retrieved July 27, 2006.

[18] Tom Gruber (1993). "A translation approach to portable ontology specifications". In: *Knowledge Acquisition*. 5, pp: 199-199.

[19] Fredrik Arvidsson and Annika Flycht-Eriksson, http://www.ida.liu.se/~janma/SemWeb/Slides/ontologies1.pdf, retrieved on November 23, 2010.

[20] Antoniou, Grigoris and Harmelen, Frank van, "A Semantic Web Premier" (2004), The MIT Press, Cambridge, Massachusetts, London, England, ISBN: 0-262-01210-3.

[21] Sowa, J., F., "Building, Sharing, and Merging Ontologies", http://www.jfsowa.com/ontology/ontoshar.htm, retrieved on November 23, 2010.

[22] Uschold, M. , King.,M, Moralee, S., and Zorgios., Y., "The Enterprise Ontology", The Knowledge Engineering Review, Special Issue on Putting Ontologies to Use, pp 31-89, 1998.

[23] Health Level Seven, http://www.hl7.org/index.cfm, retrieved on November 23, 2010.

[24] Health Level Seven Reference Information Model, http://www.hl7.org/library/data-model/RIM/C30202/rim.htm, retrieved on November 23, 2010.

[25] Hsiao CJ, Burt CW, Rechtsteiner E, Hing E, Woodwell DA, Sisk JE. (2008) "Preliminary estimates of electronic medical records use by office-based physicians: United States, 2008". Health E-Stat. National Center for Health Statistics. 2008. http://www.cdc.gov/nchs/products/pubs/pubd/hestats/hestats.htm, retrieved on November 23, 2010.

[26] Davis Karen, Doty Michelle M., Shea Katherine, Stremikis Kristof (2008). "Health Information Technology and Physician Perceptions of Quality of Care and Satisfaction". In Health Policy, published online Nov. 25, 2008. http://www.commonwealthfund.org/Content/Publications/In-the-

Literature/2009/Jan/Health-Information-Technology-and-Physician-Perceptions-of-Quality-of-Care-and-Satisfaction.aspx, retrieved on November 23, 2010.

[27] Western Health Infostructure Canada, "Chronic Disease Management Model – Introduction", 2005.

[28] Western Health Infostructure Canada, "CDM Record Message Design Document", 2005.

[29] Western Health Infostructure Canada, "CDM Data Standard – HL7 Mapping", 2005.

[30] Western Health Infostructure Canada, "Chronic Disease Management Model – Appendix E - Code Tables", 2005.

[31] W3C, Problem-Oriented Medical Record Ontology, http://esw.w3.org/HCLS/POMROntology#A_Problem-Oriented_Medical_Record_Ontology, retr. May 25, 2010.

[32] W3C, Compter-Based Patient Record Ontology, http://code.google.com/p/cpr-ontology/, last retrieved on November 23, 2010.

[33] NEHTA, "Review of Shared Electronic Health Record Standards (Version 1.0 – 20/02/2006)", http://www.nehta.gov.au/component/option,com_docman/task,doc_download/gid,68/Itemid,139/, retr. May 25, 2010.

[34] Health Level 7, http://www.hl7.org/index.cfm, last retrieved on May 25, 2010.

[35] Open-EHR, http://www.openehr.org/home.html, last retrieved on May 25, 2010.

[36] Weed, L.L., "Medical Records, Medical Education and Patient Care", The Problem-Oriented Record as a Basic Too, Cleveland, Case Western Reserve Univ. Press, 1969.

[37] Beale, T. “The openEHR Reference Model - EHR Information Model.”, Revision: 5.1.1, http://www.openehr.org/releases/1.0.2/architecture/rm/ehr_im.pdf, last retrieved on May 25, 2010.

[38] Petrie, J. C., McIntyre Neil. “The Problem-Oriented Medical record – Its Use in Hospitals, General practice and Medical Education”. Churchill Livingstone, ISBN: 0 443 01 405 1, 1979.

[39] Kalra, Dipak. “Advancing the Realization of EHR Interoperability.” International Symposium on Biomedical Informatics in Europe, http://www.infobiomed.net/symposium/pdf/session_e/Dipak_Kalra.pdf, last retrieved on May 25, 2010.

[40] Schultz JR. “A History of the PROMIS Technology:An Effective Human Interface”, In Goldberg A (ed) A History of Personal Workstations. Ad Wesley, 1988.

[41] PKC, “A Problem Oriented Approach to the Computerized Patient Record”, http://www.pkc.com/papers/pomr.pdf retr. May 25, 2010.

[42] Begoyan, A. "An Overview of ARDS for Electronic Health Rexords." Integrated Design and Process Technology, IDPT-2007, 2007.

[43] ASTM Standard Practice for Content and Structure of the Electronic Health Record (EHR), http://www.astm.org/Standards/E1384.htm, last retrieved on January 25, 2011.

[44] Bayegan, E., “Knowledge Representation for Relevance Ranking of Patient-Record Contents in Primary-Care Situations”, PHD Thesis, Norwegian University of Science and Technology, 2002, http://www.idi.ntnu.no/grupper/su/publ/phd/bayegan-thesis.pdf, last retrieved on May 25, 2010.

[45] Barry Smith, Basic Formal Ontology, http://www.ifomis.org/bfo, last retrieved on May 25, 2010.

[46] Richard H. Scheuermann, Werner Ceusters, and Barry Smith, "Toward an Ontological Treatment of Disease and Diagnosis", Proceedings of the 2009 AMIA Summit on Translational Bioinformatics, 2009, 116-120.

[47] A L Rector PD Johnson S Tu C Wroe and J Rogers, "Interface of inference models with concept and medical record models", in Proc Artificial Intelligence in Medicine Europe (AIME-2001) Springer, 2001, 314-323.

[48] Gomez-Perez A, Fernandez-Lopez M, Corcho O. Ontological engineering with examples from the areas of knowledge management, e-commerce and the Semantic Web. London ; New York: Springer, 2004.

[49] Bodenreider . Investigating subsumption in SNOMED CT: An exploration into large description logic-based biomedical terminologies. Artificial intelligence in medicine 2007;39:183.

[50] HL7 Reference Information Model, http://www.hl7.org/v3ballot/html/welcome/environment/index.htm, last retrieved on May 25, 2010.

[51] Kalfoglou Y., Schorlemmer M.,Ontology mapping: the state of the art, The Knowledge Engineering Review 18(1):1-43, January 2003.

[52] Choi, N., Song, I.Y., Han, H.: A survey on ontology mapping. ACM SIGMOD Record 35(3), 34–41 (2006).

[53] A. Doan, J. Madhavan, P. Domingos, and A. Halevy. Learning to map between ontologies on the semantic web. In *Proceedings of the 11th International World Wide Web Conference (WWW 2002), Hawaii, USA*, May 2002.

[54] Nuno Silva, Joao Rocha, *"MAFRA – An Ontology Mapping FRAmework for the Semantic Web"*, Proceedings of the 6th International Conference on Business information Systems; UCCS, Colorado Springs, CO, May 2003.

[55] Natalya Fridman Noy and Mark A. Musen, *"Smart: Automated Support for Oontology Merging and Alignment",* Proceedings of the Twelfth Banff Workshop on Knowledge Acquistion, Modeling, and Management, Banff Algeberta, 1999.

[56] N. Noy and M. Musen, "*PROMPT: Algorithm and Tool for Automated Ontology Merging and Alignment.*" Proceedings of the National Conference on Artificial Intelligence (AAAI), 2000, *Austin, TX, USA*.

[57] N.F. Noy and M. Musen. PROMPTDIFF: A Fixed-Point Algorithm for Comparing Ontology Versions. In *Proceedings of the 18th National Conference on Artificial Intelligence (AAAI'02), Edmonton, Alberta, Canada*, August 2002.

[58] M. Lacher and G. Groh. Facilitating the exchange of explicit knowledge through ontology mappings. In *Proceedings of the 14th International FLAIRS conference, Key West, FL, USA*, May 2001.

[59] Gerd Stumme, Alexander Maedche, *"FCA-Merge: Bottom-Up Merging of Ontologies"*, In proceeding of the International Joint Conference on Artificial Intelligence IJCA101, Seattle, USA, 2001.

[60] Y. Kalfoglou and M. Schorlemmer. Information-flow-based ontology mapping. In *On the Move to Meaningful Internet Systems 2002: CoopIS, DOA, and ODBASE.* Lecture Notes in Computer Science 2519, pages 1132–1151. Springer, 2002.

[61] S. Prasad, Y. Peng, and T. Finin. Using explicit information to map between two ontologies. In Proceedings of the AAMAS 2002 Wokshop on Ontologies in Agent Systems (OAS'02), Bologna, Italy, July 2002.

[62] P. Mitra and G. Wiederhold. Resolving terminological heterogeneity in ontologies. In *Proceedings of the ECAI'02 workshop on Ontologies and Semantic Interoperability, Lyon, France*, July 2002.

[63] E. Compatangelo and H. Meisel. Intelligent support to knowledge sharing through the articulation of class schemas. In *Proceedings of the 6th International Conference on Knowledge-Based Intelligent Information & Engineering Systems (KES'02), Crema, Italy*, September 2002.

[64] Domenico Beneventano, Sonia Bergamaschi**,** Francesco Guerra, Maurizio, "*Synthesizing an Integrated Ontology",* IEEE Internet Computing, September • October 2003.

[65] AnHai Doan, Pedro Domingos, Alon Halevy, "*Learning to Match the Schemas of Data Sources: A Multistrategy Approach",* Machine Learning, 50 (3): 279-301, March 2003.

[66] John Li, *"LOM: A Lexicon-based Ontology Mapping Tool",* Proceedings of the Performance Metrics for Intelligent Systems (PerMIS. '04), 2004.

[67] Gomez-Perez A, Fernandez-Lopez M, Corcho O. Ontological engineering with examples from the areas of knowledge management, e-commerce and the Semantic Web. London ; New York: Springer, 2004.

[68] Bodenreider . Investigating subsumption in SNOMED CT: An exploration into large description logic-based biomedical terminologies. Artificial intelligence in medicine 2007;39:183.

[69] Western Health Infostructure Canada, "CDM Schemas and Sample Messages", 2005.

[70] HL7 Implementation Guide: CDA Release 2 – Continuity of Care Document (CCD), https://modeling-mdt.projects.openhealthtools.org/ds/getDSMessageAttachment.do/CCD-Model-Report.pdf?dsForumId=159&dsMessageId=1521&dsAttachmentId=1840&dsAttachmentMime=application/pdf, last retrieved on May 25, 2010.

[71] Leonardo Lezcano, Miguel-Angel Sicilia, Pablo Serrano-Balazote. Combining OpenEHR Archetype Definitions with SWRL Rules – A Translation Approach. EMERGING TECHNOLOGIES AND INFORMATION SYSTEMS FOR THE KNOWLEDGE SOCIETY, Lecture Notes in Computer Science, 2008, Volume 5288/2008, 79-87.

[72] Pellet: The Open Source OWL 2 Reasoner, http://clarkparsia.com/pellet/, last retrieved on November 23, 2010.

[73] RacerPro, http://www.racer-systems.com/products/racerpro/, last retrieved on November 23, 2010.

[74] McGuire, Michael R., Steps Toward a Universal Patient Medical Record: A Project Plan to Develop One, Universal Publishers, Boca Raton, FL, 2004.

[75] Chronic Care Model Gallery, The Chronic Care Model, http://www.improvingchroniccare.org/index.php?p=CCM_Gallery&s=149, last retrieved on November 23, 2010.

[76] Vizenor L. Actions in health care organizations: an ontological analysis. Medinfo 2004;11(Pt 2): 1403-7.

[77] Vizenor L, Smith B, Ceusters W (2004) Foundation for the electronic health record: An ontological analysis of the HL7's reference information model. http://ontology.buffalo.edu/medo/HL7_2004.pdf . last retrieved on November 23, 2010.

[78] Smith B, Ceusters W (2006) HL7 RIM: an incoherent standard. Studies in health technology and informatics 2006;124, pp:133-8.

[79] Lenat, D. and Guha, R. (1990) Building Large Knowledge-Based Systems: Representation and Inference in the Cyc Project, Addison-Wesley Publishing Company Inc.

[80] The Cyc Foundation, http://www.cyc.com, last retrieved on November 23, 2010.

[81] Taylor, M.E., Matuszek, C., Klimt, B., Witbrock, M.: Autonomous classification of knowledge into an ontology. In: Proceedings of the Twentieth International FLAIRS Conference (2007), pp: 140-145.

[82] FERNANDEZ-LOPEZ and A GOMEZ-PEREZ. Overview and analysis of methodologies for building ontologies. The knowledge engineering review 2003;17:129.

[83] Uschold M, King M. Towards a Methodology for Building Ontologies.: Artificial Intelligence Applications, Institute,Universify of Edinburgh; 1995.

[84] Gruniger, M. and Fox, M., Methodology for the design and evaluation of ontologies. Proceedings of the IJCAI; 1995.

[85] Schreibur, A., Wielinga, B. and Jansweiger, W. (1995) The KACTUS view on the 'O' word, In Proceedings of the 7th Dutch National Conference on Artificial Intelligence, The Netharlands, 159-168.

[86] Swartout, W., Patil, R., Knight, K. and Russ, T. (1996) Toward Distributed Use of Large-Scale Ontologies, In Proceedings of the Tenth Workshop on Knowledge Acquisition for Knowledge-Based Systems.

[87] Fernandez, M., Gomez-Perez, A. and Juriston, N. (1997) METHONTOLOGY: from Ontological Art towards Ontological Engineering, In Proceedings of the AAAI97 Spring Symposium Series on Ontological Engineering, March, Stanford, USA, 33-40.

[88] Noy, N. and McGuiness, D. (2001) Ontology Development 101: A Guide to Creating Your First Ontology, Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880.

[89] Sure, Y. (2003) A Tool-supported Methodology for Ontology-based Knowledge Management, York Sure, In the Ontology and Modeling of Real Estate Transactions, Ashgate.

[90] De Nicola, A., Missikof, M. and Navigli, R. (2005) A Proposal for a Unified Process for Ontology Building: UPON, Database and Expert Systems Applications, 655-664.

[91] Dhalem, N., Hahn, A. (2009) User-Friendly Ontology Creation Methodologies – A Survey, In Proceedings of the Fifteenth Americas Conference on Information Systems, San Fransisco, CA, 1-9.

[92] Antoniou, G. and Harmelen, F. (2004) Web Ontology Language: OWL, In Handbook on ontologies. Berlin; New York: Springer, p. 67-92.

[93] Imam, F. T. (2007) An Inconsistency Tolerant Approach to Ontology Merging, Master's Thesis, Computer Science, Saint Francis Xavier University, Canada.

APPENDIX A Mapping Results between the CDM Model and the CPR Ontology Concepts

CDM Concept	CPR Ontology Concept	Matching	Mapping Comments
Person	Patient	F	In CDM, 'Person' is the person whose record is represented, i.e., the patient with chronic condition.
Primary Care Provider	PrimaryCareProvider	F	We created a specialization, cpr:PrimaryCareProvider of cpr:Provider with Role 'PrimaryCareProviderRole'
EncounterEvent	Encounter-->Referral	F	
Procedures	therapeuticAct U procedure	F	
Diagnostic Images	diagnosticImage	F	
Observations	clinical-investigation-act U clinical-finding	F	
Referral	Referral	F	
MedicationsOrVaccines	SubstanceAdministration	F	
Goals	PlannedGoals	F	A new class has been created
CarePlan	PlannedAct	F	
Planned Activities	PlannedAct	F	
PlannedProcedures	PlannedAct-->clinical-act(therapeuticAct U procedure)	F	
PlannedDiagnostic Images	PlannedAct-->clinical-act(diagnostic-imaging procedure)	F	A new concept has been added under diagnostic procedure
PlannedObservations	PlannedAct-->clinical-act(clinical-investigation-act U clinical-finding)	F	
PlannedReferral	PlannedAct-->clinical-act(Referral)	F	
PlannedMedicationsOr Vaccines	PlannedAct-->clinical-act(SubstanceAdministration)	F	
Chornic Condition	medical-problem	F	
Provider	Provider	F	
ProviderOrganization	ProviderOrganization	F	A new top-level concept, 'ProviderOrganization' has been added for this purpose
ProviderReference	medical-problem.ProblemResponsiblePractitionerID	F	
ProviderOrder	Order	F	A new concept has been added under clinical-act for representing Orders of clinicians
PersonalContact	Patient.FamilyMemberName,	P	No scope of incorporating neighbors or other contacts
	Patient.FamilyMemberRelationship		
AdministrativeCorrection	AdministrativeCorrection	F	A new concept has been added under 'clinical-administration-act'
ChangeInCondition	medical-problem.CurrentStatus	P	If coded values are used for CurrentStatus of medical problems, this can be used to track whether the patient is still under treatment for a particular chronic condition.
CDM Record	patient record	F	
Supporting Elements	clinical-act U clinical-finding	F	

APPENDIX B Mapping Results between the CDM Model and the CPR Ontology Relations

CDM					CPR Ontology				Comments
Relationship	Domain	Range	Constraint		Relationship	Domain	Range	Constraint	
appliesTo	CDMRecord	Person	functional, mandatory		appliesTo Patient	patient-record	Patient	functional, mandatory	a new object property is created
hasAssociated	CDMRecord	Providers or Provider Organizations	Max 50		hasAssociated	patient-record	Provider U ProviderOrganization		a new object property is created
appliesTo	CDMRecord	Chronic Conditions	Max 25		appliesTo Conditoin	patient-record	medical-problem		a new object property is created
has primary contact	CDMRecord	Provider Reference	Required, 1		has primary contact	patient-record	Provider U ProviderOrganization		a new object property is created
has	CDMRecord	PersonalContacts	Max 2						Since patient-record is already connected to Patient, its personal contacts will be also associated with patient-record
contains related	CDMRecord	Procedures/ Observations/ Diagnosic Images /Referral/ Medications or Vaccines	required, Max 1000000		contains	patient-record	cpr:clinical-act ⊔ cpr:representational-artifact ⊔ cpr:image		
managed by	CDMRecord	Care Plans	reuired, Max 1000		managed by	patient-record	PlannedAct		a new object property is created
appliesTo	CDMRecord	Person	functional, mandatory		appliesTo Patient	patient-record	Patient	functional, mandatory	a new object property is created
hasAssociated	CDMRecord	Providers or Provider Organizations	Max 50		hasAssociated	patient-record	Provider U ProviderOrganization		a new object property is created
appliesTo	CDMRecord	Chronic Conditions	Max 25		appliesTo Conditoin	patient-record	medical-problem		a new object property is created

CDM					CPR Ontology				Comments
Relationship	Domain	Range	Constraint		Relationship	Domain	Range	Constraint	
has primary contact	CDMRecord	Provider Reference	Required, 1		has primary contact	patient-record	Provider U ProviderOrganization		a new object property is created
has	CDMRecord	PersonalContacts	Max 2						Since patient-record is already connected to Patient, its personal contacts will be also associated with patient-record
contains related	CDMRecord	Procedures/ Observations/ Diagnosic Images /Referral/ Medications or Vaccines	required, Max 1000000		contains	patient-record	cpr:clinical-act ⊔ cpr:representational-artifact ⊔ cpr:image		
managed by	CDMRecord	Care Plans	reuired, Max 1000		managed by	patient-record	PlannedAct		a new object property is created
under care of	Person	Primary Care Provider	required, 1		under care of	Patient	Primary Care provider	functional	a new object property is created
has contact	Chronic Conditions	Provider Reference	required, 1		hasContact	medical-problem	Provider U ProviderOrganization	functional	a new object property is created
was ended because of	Chronic Conditions	Administrative correction/ Change in Condition	required, 1					functional	a new data type property of medical-problem has been created 'EndReason'
InFulfillmentOf1	Referral	Encounter Event	required, 1						a new object property is already created from Encounter to clinical-act (Referral is a specialization of clinical-act)
referred to	Referral	Provider Reference	required, 1					functional	a new object property is already created from Referral to Provider

CDM					CPR Ontology				Comments
Relationship	Domain	Range	Constraint		Relationship	Domain	Range	Constraint	
ordered by	Provider's Order	Provider Reference	required, 1		orderedBy	Order	Provider U ProviderOrganization	functional	a new object property is created
has	Care Plans	Planned Activities	required, Max 1000						a new object property is created from PlannedAct to clinical-act
has	Care Plans	Goals	required, Max 100		contains	PlannedAct	PlannedGoals		
Provider's order	Supporting Elements	Provider's Order	required, 1		orderOf	Order	clinical-act	functional	a new object property is created
entered / recorded by	Supporting Elements	Provider Reference	populated, 1		composedBy	Action U representational-artifact	Person	functional	
appliesTo	CDMRecord	Person	functional, mandatory		appliesToPatient	patient-record	Patient	functional, mandatory	a new object property is created
hasAssociated	CDMRecord	Providers or Provider Organizations	Max 50		hasAssociated	patient-record	Provider U ProviderOrganization		a new object property is created
appliesTo	CDMRecord	Chronic Conditions	Max 25		appliesToConditoin	patient-record	medical-problem		a new object property is created
has	CDMRecord	PersonalContacts	Max 2						Since patient-record is already connected to Patient, its personal contacts will be also associated with patient-record
contains related	CDMRecord	Procedures/ Observations/ Diagnosic Images /Referral/ Medications or Vaccines	required, Max 1000000		contains	patient-record	cpr:clinical-act ⊔ cpr:representational-artifact ⊔ cpr:image		
managed by	CDMRecord	Care Plans	reuired, Max 1000		managed by	patient-record	PlannedAct		a new object property is created
under care of	Person	Primary Care Provider	required, 1		under care of	Patient	Primary Care provider	functional	a new object property is created

CDM					CPR Ontology				Comments
Relationship	Domain	Range	Constraint		Relationship	Domain	Range	Constraint	
has contact	Chronic Conditions	Provider Reference	required, 1		hasContact	medical-problem	Provider U ProviderOrganization	functional	a new object property is created
was ended because of	Chronic Conditions	Administrative correction/ Change in Condition	required, 1					functional	a new data type property of medical-problem has been created 'EndReason'
InFulfillmentOf1	Referral	Encounter Event	required, 1						a new object property is already created from Encounter to clinical-act (Referral is a specialization of clinical-act)
referred to	Referral	Provider Reference	required, 1					functional	a new object property is already created from Referral to Provider
ordered by	Provider's Order	Provider Reference	required, 1		orderedBy	Order	Provider U ProviderOrganization	functional	a new object property is created
has	Care Plans	Planned Activities	required, Max 1000						a new object property is created from PlannedAct to clinical-act
has	Care Plans	Goals	required, Max 100		contains	PlannedAct	PlannedGoals		
Provider's order	Supporting Elements	Provider's Order	required, 1		orderOf	Order	clinical-act	functional	a new object property is created
entered / recorded by	Supporting Elements	Provider Reference	populated, 1		composedBy	Action U representational-artifact	Person	functional	
appliesTo	CDMRecord	Person	functional, mandatory		appliesToPatient	patient-record	Patient	functional, mandatory	a new object property is created
hasAssociated	CDMRecord	Providers or Provider Organizations	Max 50		hasAssociated	patient-record	Provider U ProviderOrganization		a new object property is created
appliesTo	CDMRecord	Chronic Conditions	Max 25		appliesToConditoin	patient-record	medical-problem		a new object property is created

APPENDIX C Mapping Results between the CDM Model Attributes and the CPR Ontology Data-Type Properties

CDM Attribute	CPR Ontology Property	Matching	Mapping Comments
Person	**Patient**	F	
Person identifiers	UniversalPatientHealthNumber	F	These two match in the sense that both are unique numbers for each patient in a clinical setting.
Person names	PatientName	F	
Person gender	Sex	F	
Person date of birth	DateTimeOfBirth	F	
Person date of death	DateTimeOfDeath	F	Newly added
Person phone numbers and emails	PhoneNumber and Email	F	Newly added
Person addresses	PatientPermanentAddress	P	
Person language types	LanguageType	F	Newly added
Primary Care Provider	**Primary Care Provider**		
Primary care provider identifier	Provider.ID	F	
Primary care provider name	Provider.Name		
Primary care provider phone numbers and emails	Provider.PhoneNumber and Provider.Email	F	Newly added
Primary care provider location name	Provider.Location	F	Newly added
Provider Organization	**Provider Organization**		
Provider facility id	ID	F	Newly added
Provider facility name	Name	F	Newly added

CDM Attribute	CPR Ontology Property	Matching	Mapping Comments
Provider facility phone numbers and emails	PhoneNumber and Email	F	Newly added
Provider	**Provider**		
Provider type	Type	F	Newly added
Provider specialty	Specialty	F	Newly added
Provider identifier	ID	F	Newly added
Provider name	Name	F	Newly added
Provider location name	Location	F	Newly added
Provider phone numbers and emails	PhoneNumber and Email	F	
Chronic Condition	**medical-problem**		
Kind of chronic condition	problemKind	F	Newly added
Diagnosis of condition date range	conditionStartDate and conditionEndDate	F	Newly added
Recall indicator	recallIndicator	F	Newly added
Care plan cross references	PlannedAct-->clinical-act -->medical-problem	F	
Provider Reference		F	
Provider identifier	Provider.ID	F	
Date entered, recorded	representational-artifact.compositionDateTime	F	
Procedures	**therapeutic-act/ procedure**		
Procedure type	TherapeuticAct.therapyType and Procedure.procedureType	F	Newly added
Procedure identifier	clinical-act.ID	F	Newly added
Procedure date	TherapeuticAct.therapyStartDateTime, TherapeuticAct.therapyFinishDateTime and procedure.DateTime	F	
Procedure occurred indicator	ClinicalAct.occuredIndicator	F	
Procedure non performance reason	ClinicalAct.NonPerformanceReasonCode	F	
Procedure details	therapeutic-act.details and procedure.details	F	Newly added
Medications/Vaccines	**substance-administration**		
Medication / Vaccine flag	MedicationOrImmunization	F	Newly added
Medication / Vaccine type	MedicationType	F	Newly added
Medication / Vaccine name	MedicationName	F	
Medication / Vaccine prescription or dispense identifier	Order.OrderID-->substance-admin	F	
Medication / Vaccine dose	Medicationdose	F	
Medication / Vaccine frequency	Medicationfreq	F	
Medication / Vaccine route	MedicationRoute	F	
Medication / Vaccine occurred indicator	ClinicalAct.occuredIndicator	F	
Observations	**clinical-investigation-act U clinical-finding**		
Observation type	act-type and finding-type	F	Newly added
Observation identifier	clinical-act.ID	F	

CDM Attribute	CPR Ontology Property	Matching	Mapping Comments
Observation value	diagnostic-procedure.NumericMeasurementOr AnalyteInterpretation, clinical-examination.ExamFindings, medical-history-screening-act, clinical-finding.value	F	
Observation interpretation	diagnostic-procedure.NumericMeasurementOr AnalyteInterpretation	F	
Observation method code	clinical-act.Method	F	
Observation date	clinical-finding.Date, diagnostic-procedure. TestDateTime, diagnostic-procedure.TestDateTimeResultReported, clinical-examination.DateTimeOfExam, screeningDateTime	F	
Observation occurred indicator	clinical-act.occuredIndicator	F	
Observation non performance reason	clinical-act.nonPerformance ReasonCode	F	
Observation normal range value	diagnostic-procedure.normalRangeValue	F	Newly added
Observation details	diagnostic-procedure.TestComments, clinical-examination.ExamSummary, medical-history-screening-act, clinical-finding.Details	F	
Diagnostic Images	**diagnostic-image, diagnostic-procedure**		
Diagnostic image type	diagnostic-image.Type	F	Newly added
Diagnostic image identifier	clinical-act.ID	F	
Diagnostic image method code	clinical-act.Method	F	
Diagnostic image date	diagnostic-procedure.	F	
Diagnostic image occurred indicator	TestDateTime, diagnostic-procedure. TestDateTimeResultReported	F	
Diagnostic image non performance reason	clinical-act.nonPerformance ReasonCode	F	
Diagnostic image details	diagnostic-procedure.TestComments	F	
Referral	**Referral**		
Referral identifier	clinical-act.ID	F	Newly added
Referral reason code	reasonCode	F	Newly added
Referral details	referralDetails	F	Newly added
Referral occurred indicator	clinical-act.occuredIndicator	F	Newly added
Care Plans	**PlannedAct**		
Care plan identifier	ID	F	Newly added
Care plan attachment	Attachment	F	Newly added
Planned Activities	**PlannedAct**		
Reassessment date	reassesmentDate	F	Newly added
Goals	**PlannedGoals**		
Goal type	goalType	F	Newly added
Goal value	goalValue	F	Newly added
Goal value target date	goalTargetDate	F	Newly added

CDM Attribute	CPR Ontology Property	Matching	Mapping Comments
Planned Procedures	**PlannedAct-->clinical-act(therapeuticAct U procedure)**		
Planned procedure type	Procedure.procedureType	F	PlannedAct-->clinical-act(therapeuticAct U procedure)
Planned Medications/Vaccines	**PlannedAct-->clinical-act(SubstanceAdministration)**		
Medication / Vaccine flag	MedicationOrImmunization	F	PlannedAct-->clinical-act(SubstanceAdministration)
Medication / Vaccine type	MedicationType	F	
Medication / Vaccine name	MedicationName	F	
Planned Observations	**PlannedAct-->clinical-act(clinical-investigation-act U clinical-finding)**		
Planned observation type	act-type and finding-type	F	PlannedAct-->clinical-act(clinical-investigation-act U clinical-finding)
Planned observation diagnosis value	diagnostic-procedure.NumericMeasurementOr AnalyteValue	F	
Planned Diagnostic Image	**PlannedAct-->clinical-act(diagnostic procedure)**		
Planned diagnostic image type	diagnostic-image.Type		
Encounter Event	**Encounter-->Referral**		
Referral encounter date	Encounter.DateTime	F	
Referral encounter non performance reason	clinical-act.nonPerformance ReasonCode	F	
Referral encounter occurred indicator	clinical-act.occuredIndicator	F	
Provider's Order	**Order**		
Medication / Vaccine dispense date	OrderDateTime-->substance-admin	F	
AdministrativeCorrection	**AdministrativeCorrection**		
End reason	EndReason	F	Newly added
ChangeInCondition			
	medical-problem.CurrentStatus		
End reason	medical-problem.EndReason	F	Newly added
CDM Record	**patient record**		
End reason	EndReason	F	Newly added

APPENDIX D Mapping Results between the EMR Ontology and the HL7 RIM

EMR Ontology Concept/ Property	HL7-RIM Concept/Attribute	Matching	Comments
Patient	**Person with Role 'Patient'**	**F**	
Patient.Birthplace	Place	F	Place is a specialization of Entity in RIM which can be linked with the Person for Patient for BirthPlace
Patient.DateOfEarliestHeldEntry		N	No such match found in RIM
Patient.DateOfLatestHeldEntry	Act.activityTime	P	Assuming that the document containing the current record is the latest entry at that time.
Patient.DateTimeOfBirth	LivingSubject.birthTime	F	Inherited to Person class from its parent class LivingSubject
Patient.EducationLevel	Person.educationLevelCode	F	
Patient.EthnicGroup	Person.ethnicGroupCode	F	
Patient.FamilyMemberName	Person class for Patient linked with another Person instance for relative	P	
Patient.FamilyMemberRelationship	Person class for Patient linked with another Person instance for relative	P	
Patient.MaritalStatus	Person.maritalStatusCode	F	
Patient.Occupation	ObservationOccupation	F	
Patient.PatientName	Entity.name	F	RIM has more facilities to break down name into first name, lastname, title etc.
Patient.PatientPermanentAddress	Person.addr	P	No scope to define whether the address is permanent
Patient.Race	Person.raceCode	F	
Patient.RecordHoldingLocationID	Entity→Place	P	There is a way in RIM to define the location where the document was created.
Patient.Religion	Person.religiousAffiliationCode	F	
Patient.Sex	LivingSubject.administrativeGenderCode	F	
Patient.UniversalPatientHealthNumber	Entity.id	F	
Physician	**Person with Role Physician**	F	
Physician.Address	Entity.addr	F	
Physician.id	Entity.id	F	
Physician.Name	Entity.name	F	
Physician.Profession	Employee.occupationCode	F	
Physician.UniversalID Number		N	No such match for UniversalID. However, the Entity ID is unique for each physician
Clinical-Exam	**Observation with actCode = 29545-1 from LOINC**	F	
Clinical-Exam.DateTimeOfExam	Act.activityTime	F	
Clinical-Exam.ExamFindingComments	Act.Text	F	
Clinical-Exam.ExamFindings	Observation.Value	F	
ClinicalAct	**Act**	F	
ClinicalAct.NonPerformanceReasonCode	Act.reasonCode	F	
ClinicalAct.occuredIndicator	Act.negationInd	F	
diagnostic-procedure	**Observation with classCode='DG' and moodCode='EVN'**	F	Alternatively, RIM can also use code from LOINC to denote this concept
diagnostic-procedure.MicrorgAttribut	Entity.desc	P	A list of attributes for a microbiological organism.

EMR Ontology Concept/ Property	HL7-RIM Concept/Attribute	Matching	Comments
diagnostic-procedure.NumericMeasurementOrAnalyteInterpretation	Observation.InterpretationCode	F	
diagnostic-procedure.NumericMeasurementOrAnalyteName	Act.Title	F	
diagnostic-procedure.NumericMeasurementOrAnalyteValue	Observation.Value	F	
diagnostic-procedure.TestComments	Act.Text	F	
diagnostic-procedure.TestDateTime	Act.activityTime	F	
diagnostic-procedure.TestDateTimeResultReported	Act.effectiveTime	F	
diagnostic-procedure.TestReportText	Act.Text	F	
diagnostic-procedure.TestRequested	Act.Title	F	
diagnostic-procedure.TestRequestOrderingTreatmentFacility	Observation→Participation (type='ORG')→Entity	F	'ORG' here stands for the origin of location where the diagnostic procedure was originated
diagnostic-procedure.TestRequestPerformingFacility	Observation→Participation (type='DST')→Entity	F	'ORG' here stands for the origin of location where the diagnostic procedure was happened
Encounter	**Act with classCode='ENC' and moodCode='EVN'**	F	
Encounter.DateTime	Act.effectiveTime	P	
Encounter.EncounterID	Act.id	F	
medical-history-screening-act	**Observation with classCode='SCR' and moodCode='EVN'**	F	
medical-history-screening-act.CurrentHabits	Observation.Value where Act.code = 'HBT'	F	
medical-history-screening-act.HistoryOfPreviousIllnesses	Observation.Value	F	
medical-history-screening-act.HistoryRelationshipSourceToPatient	Observation→Participation (typeCode='Source')→Person	F	
medical-history-screening-act.HistoryTakingEventDate	Act.activityDate	F	
Medical-problem	**Act with classCode='CNOD' and moodCode='EVN'**	F	
medical-problem.ProblemCurrentStatus	Act.StatusCode	F	
medical-problem.ProblemDateOfOnset	Act.activityTime	F	
medical-problem.ProblemName	Act.Title	F	
medical-problem.ProblemNumber	Act.priorityCode	P	There is no direct match. Hoever, priorityCode is possibly the closest match for that.
PlannedAct	**Act with moodCode='INT'**	F	INT moodCode stands for Intention, i.e., planned acts
Goal	**Observation with moodCode='GOL'**	F	GOL moodCode stands for Goal
Goal.GoalDate	Act.activityDate	F	
Goal.GoalType	Observation.code	F	
Goal.GoalValue	Observation.value	F	
Goal.reassessmentDate	Act.effectiveTime	F	
Referral	**Act with classCode='PCPR' and moodCode='RQO'**	F	

EMR Ontology Concept/ Property	HL7-RIM Concept/Attribute	Matching	Comments
Referral.reasonCode	Act.code	F	Act codes can take values from standard vocabularies like SNOMED-CT, LOINC etc.
Referral.referralDetails	Act.Text	F	
Referral.ReferralID	Act.ID	F	
Referral.referralNonPerformanceReason	Act.reasonCode	F	
Referral.referralOccuredIndicator	Act.negationInd	F	
substance-administration	**ManfactureMaterial with Act as SubstanceAdministration**	F	
substance-administration.MedicationDateOfLastRefill		N	No match found for last refill date
substance-administration.MedicationDose	SubstanceAdmin.doseQuantity	F	
substance-administration.MedicationFreq	SubstanceAdmin.rateQuantity	F	
substance-administration.MedicationInstructions	Act.text	F	Inherited to SubstanceAdministration from its parent class Act
substance-administration.MedicationName	Entity.name	F	
substance-administration.MedicationNotes	Act.text	F	Inherited to SubstanceAdministration from its parent class Act
substance-administration.MedicationPrescriber	SubstanceAdministration→Participation (with typeCode='Prescription')→ Person instance for physician	F	
substance-administration.MedicationPresTime	Act.activityTime	F	
substance-administration.MedicationRoute	SubstanceAdmin.routeCode	F	
substance-administration.MedicationVehicle	Act.text	P	Not directly matches, however could be included into the detail text of SubstanceAdministration Act
therapeutic-act	**Procedure**	F	
therapeutic-act.NameOfTherapy	Act.Title	F	
therapeutic-act.TherapistsRecommendations	Act.Text	F	
therapeutic-act.TherapistsResponseAssessment	Act.Text	P	This is used in the EMR ontology for the therapists to specify her assessment of patients response, there is no straight forward mapping found in RIM, however, this could be defined in the Text field under a subheading.
therapeutic-act.TherapyFinishDateTime	Act.ActivityTime	F	The ActivityTime of class Act in RIM defines a duration, thus, is used to specify both start and end times of a procedure.
therapeutic-act.TherapyStartDateTime	Act.ActivityTime	F	