Ontology-driven Approach to Health Data Management for Remote Healthcare Delivery

Poly Sil Sen
Department of Information Technology,
Techno India, SaltLake,
Kolkata, India
polysilsen14@gmail.com

Shabnam Banerjee
Department of Palliative Care and Psychooncology,
Tata Medical Center,
Kolkata, India
shabnambanerjee.mck@gmail.com

Nandini Mukherjee
Department of Computer Science and Engineering,
Jadavpur University,
Kolkata, India
nmukherjee@cse.jdvu.ac.in

ABSTRACT

In the current scenario, it is possible to provide remote healthcare facilities to people living in rural and remote areas using mobile, health sensor and cloud technologies. However, in order to support a remote health framework and to store and access data generated through this system, it is necessary to develop a standard ontology. Till date there is no standard format to hold health data generated from a remote health framework which uses sensors to measure the clinical parameters and store the data in the cloud. In this paper we propose an ontology for health data for a remote health framework, particularly focusing on primary healthcare delivery. This ontology handles health data and care for common causes of illness and death in child upto five years in developing countries. However, it can be extended to support primary healthcare in general.

KEYWORDS

Remote Healthcare, Health Sensors, Ontology

1 INTRODUCTION

Currently available Electronic Health Record (EHR) formats are flat and not suitable for proper representation of complex data including streaming data generated from e-health sensors or social media data. On the other hand the recent usage scenarios of e-health applications and the new opportunities for delivering remote healthcare services based on new technologies give rise of the necessity of new hardware, software and storage types for storage and management of such data, which are nowadays categorised as big data.

Health data produced through the use of healthcare applications are so large, complex and diverse that they are difficult or impossible to manage with traditional hardware, software and data management tools and methods [11]. In order to represent such huge and diverse data, an appropriate representation is sought. Ontology represents entities and their interrelationships among each other. An ontology is needed to integrate databases and data stores to share information and for use by artificial agents for knowledge acquisition.

There is no standard, available ontology to represent health data that can be used for general purposes. In particular, no ontology has yet been proposed for the use in a remote health framework that uses sensors. Our specific objective is to deliver primary healthcare services on remote basis using a sensor-cloud framework [12] in rural areas of India. Therefore, an ontology is required for representation and management of data generated within this framework. In this work, we propose an ontology to cover this domain.

The ontology proposed in this paper accommodates healthcare data for young infants and sick children upto five years covering common causes of illness and death following IMNCI, Integrated Management of Neonatal and Childhood Illness [6]. More than ten million children die during their first five years of life in developing countries. Seventy percent of these deaths are because of measles, malaria, pneumonia, diarrhoea and malnutrition. Our ontology contains clinical signs and symptoms and assessment, so that preliminary diagnosis can be made that whether a patient has malaria, measles, pneumonia, diarrhoea, ear infection, malnutrition and can be treated. Domain knowledge related to health data of the patient is represented in this ontology in such a way that it can be used to decide whether the child requires immediate referral to hospital, specific care at outpatient health facility, care at home, monitoring or observation [6].

The rest of the paper is organised as follows: in Section II used IMNCI process of child care is explained, in Section III state of the art of ontology in health care is discussed, in Section IV the proposed ontology is discussed, in Section V verification and validation of the ontology is discussed. Finally Section VI concludes the paper with a direction for future work.

2 USED PROCESSES FOR CHILD CARE

The process of treating and caring a child is considered following IMNCI- Integrated management of Neonatal and Childhood Illness [6] document. Young infants upto 2 months are treated in one order and sick children upto 5 years are treated in another manner. This process does not cover chronic illness or emergency due to accidents etc.

In case of young infants the following steps are adopted according to [6]:

1. Checking of possible bacterial infection or jaundice
2. Symptom assessment for diarrhoea
3. Checking of feeding problem, malnutrition and immunization
4. Checking of other problems
5. Classification of degree of severity and treatment action identification
6. Selection among urgent referral to hospital, treatment at outpatient health facility and care at home is made
3 ONTOLOGY IN HEALTH CARE : STATE OF THE ART

Ontology provides a formal specification of the domain knowledge, including the structure of information, relations, properties etc. Pete lanance points out that medical information systems need to be able to communicate complex and detailed medical data securely and efficiently [7]. He also mentions that such a task requires thorough analysis of the structure and the concepts of medical terminologies and can be achieved by constructing medical domain ontologies for representing medical terminology systems. One important benefit of such ontologies is that they provide support for integration of knowledge and data. Several proposals have been made for developing ontologies in medical domain.

In [8], an ontology has been proposed to take care of chronically ill patients. The ontology provides support for decision making and covers 19 diseases, 2 syndromes and five social issues. In [9], an ontological framework along with medical work flow has been proposed for care processes, resources, organizations, medical policies in order to minimize the gap between healthcare people and IT. It has provision for task scheduling and resource utilization. In [17], an ontology based context model for monitoring health of elderly patients at home to provide chronic care is proposed. The ontology here is appropriate for the particular case mentioned. It does not work with any health data standard. In [10], it is observed that clinical data may have uncertainty. Treatment is highly affected by this uncertainty. Here, an ontology is proposed to handle uncertainty in clinical data by preparing a QoD aware telemedicine system.

In [16], authors have developed ontology for traffic sensor network. The architecture used has four layers including the bottom sensor layer, and data layer above it. Raw sensed data is stored in this data layer. The sensor data absorption into data layer is not explained or discussed in detail in terms of ontology. In [18], an approach to generate a semantic ontology from a raw set of sensor data is proposed. Continuous data is transformed into discrete data points by dividing the entire period into a number of time slots.

However, none of the above ontologies are generic enough to hold data for primary healthcare delivery. Furthermore, there is no scope of storing streaming sensor data for continuous monitoring from remote places or diverse type of data as may be necessary for remote healthcare delivery services. This paper proposes an ontology for child healthcare with consideration of the particular requirements of health data.

4 PROPOSED ONTOLOGY

In this section we describe the proposed ontology for storing health data of a remote health framework that uses sensor and cloud. The ontology covers common causes of illness and death in children up to five years.

In this process we use the available ontologies for electronic health records [19], sensors [26] and other available ontologies [20], [21], [22], [23], [24], [25], [27] which are relevant in this context. Thus, the developed ontology conforms to available related ontologies as much as possible. The ontology follows the Minimum Data Set described in Indian Standard for Electronic Health Records [5]. The relationships among different entities are also included in [1] with required modification.

It is to be noted that most of the terms used in the proposed ontology are taken from or inspired from either [5] or [6] or the different ontologies reused in this work and mentioned in the previous paragraph.

The ontology presented here has been developed using the concepts and guidelines of IMNCI document [6]. Although, the ontology is developed for the care of infants and children, the ontology with slight modification can also be used for remote delivery of primary healthcare in general. In Figure 1, the entities of the ontology are shown.

Figures 2, 3 show different sections of the ontology. Types of entities are shown using straight lines and relationships among different entities are shown using circles and dotted lines with arrows. Parents are pointed using arrows (solid) direction. Directions of the arrows for relationships indicate direction of relationships.

In the remaining part of this section we give an overview of the proposed ontology.

A Person (person) can be a Patient, a Doctor, an EmergencyContactPerson or a CarePersonAtHome to take care of a Patient. A Person can be all or some or one of them.

Name of a Person is stored in DemographicInfo. DemographicInfo contains entities BasicClinicalParameters and ContactDetails. BasicClinicalParameters contain details like dateOfBirth, gender, bloodGroup, foodHabit, socioEconomicStatus etc. ContactDetails entity contains Address, contact numbers and email addresses. DemographicInfo contains information of a Person needed to identify him/her. These information must not be shared always or kept separate.
Symptoms are kept in ClinicalSymptoms. The different types of ClinicalSymptoms are EmergencySymptoms, PrioritySymptoms, GeneralDangerSymptoms. ClinicalSymptoms entity is a collection of all such ClinicalSymptoms. In other words ClinicalSymptoms entity contains EmergencySymptoms, PrioritySymptoms, GeneralDangerSymptoms entities.

Doctor observed signs during a Visit are kept in ClinicalSignsAll. It is called ClinicalSignsAll as there is an entity that contains ClinicalSigns. The different types of ClinicalSigns are EmergencySigns, PrioritySigns, GeneralDangerSigns. ClinicalSignsAll entity is a collection of all such ClinicalSigns. In other words ClinicalSignsAll entity contains EmergencySigns, PrioritySigns, GeneralDangerSigns entities.

Doctor asks questions to the Patient or CarePersonAtHome on the basis of the Complaint made by the Patient. The answers to these questions form the ClinicalAssessment entity during a Visit. ClinicalAssessment entity has two types ClinicalAssessmentChild, ClinicalAssessmentInfant (not shown in the diagram). The types are maintained as the sets of questions asked for an infant and a child are quite dissimilar though not disjoint.

Doctor checks the attributes of the BasicClinicalParameters like blood group, food habit, socio-economic status and also HistoryAll. Patient contains HistoryAll. Using the same logic we have used before, HistoryAll contains ImmunizationHistory, AllergyHistory, FamilyHistory, PastHistory and PresentHistory all of which are different types of History entity. For a grown up female HistoryAll also contains MenstrualObstericHistory (not shown in any figure) which is again a type of History entity. We use the available History entity [19] as it has a specific structure that can be reused by ImmunizationHistory, AllergyHistory, FamilyHistory, PastHistory and PresentHistory and MenstrualObstericHistory.

Thus on the basis of BasicClinicalParameter, MeasuredClinicalParameter, ClinicalSymptomsAll, ClinicalSignsAll, ClinicalAssessment, HistoryAll, initial DiagnosisInVisit is made and Investigation is prescribed. During later visits, refinement is made to DiagnosisInVisit after going through Investigation results. DiagnosisInVisit stores Disease that has been diagnosed so far. Disease entity is kept separately as every Disease is treated in a particular way. In our work we have considered only few diseases which occur more frequently among the children. These are Malaria, Measles, Pneumonia, Malnutrition and Jaundice, Diarrhoea etc.

Investigation can be of two types ContinuousMonitoring and DiscreteMonitoring. DiscreteMonitoring is done for an instance or for a small duration of time. ContinuousMonitoring is done over a time period. DiscreteMonitoring has two types OneTimeInvestigation and PeriodicInvestigation. For example Doctor may prescribe to perform a certain pathological test. This may be advised to be done once or thrice keeping a gap of say one week.

Both, OneTimeInvestigation and PeriodicInvestigation may be DirectlyFed or ManuallyFed. All DirectlyFed data comes from AbstractedSensorData. ContinuousMonitoring is also obtained from AbstractedSensorData. SensorOutput is processed to generate AbstractedSensorData. SensorOutput comes from Sensor. AbstractedSensorData uses SensorOutput data that comes from virtual sensors deployed in our Sensor-Cloud environment.

In our work, sensors are used to record all directly fed investigations whether continuous or discrete. If it is continuous then timestamp of start of observation and end of observation is noted. For each observation, the value to be measured. Each sensor has a specific unit of measurement.

Figure 1: Hierarchy of different information objects in the ontology for health data of remote health framework

A Patient has a Complaint and for the Complaint, Doctor initiates a TreatmentEpisode. A TreatmentEpisode requires one or more than one Visit(s) to the Doctor.

Healthworker is an entity of type Person. She collects information from CarePersonAtHome. She observes a Patient. She works for a HealthCenter which is the place where TreatmentEpisode takes place or HealthCenter is the place of initiation of observation of a Complaint or situation. TreatmentEpisode may also start in a Hospital or a Doctor’s chamber depending on the available facilities of the place of stay of the Patient.

In a Visit, height, weight, body temperature, systolic blood pressure, diastolic blood pressure etc. are measured. These measured values are kept in MeasuredClinicalParameter. Patients’ words about the Complaint are stored in ClinicalSymptomsAll during the Visit. It is named ClinicalSymptomsAll as there is another entity called ClinicalSymptoms. ClinicalSymptoms has types MainSymptoms, SpecificSymptoms and GeneralSymptoms. The ClinicalSymptomsAll entity is a collection of all symptoms mentioned in MainSymptoms, SpecificSymptoms and GeneralSymptoms. ClinicalSymptomsAll on the other hand has two types ClinicalSymptomsChild and ClinicalSymptomsInfant. In other words ClinicalSymptomsAll entity contains MainSymptoms, SpecificSymptoms and GeneralSymptoms entities.
On the basis of DiagnosisInVisit and Disease diagnosed, TreatmentPlan is suggested. Doctor advises CarePersonAtHome to take care of Patient.

In this ontology we have reused the entities like Sensor, SensorOutput [26], History (written as History in some places), Information object, Person, ClinicalContext [19] of the ontologies we have used.

SuggestedCare of TreatmentPlan can have four sub types UrgentRefToHospital, SpecificCare (SpecificCareAtOutPatient), CareAtHome, FollowUpCare. UrgentRefToHospital contains ReferralNote.

EmergencySigns help a health care person to decide whether a Patient needs UrgentRefToHospital. UrgentRefToHospital contains a ReferralNote containing details like treatment given, immunizations needed etc. If a Patient has to be treated at an Out Patient Health Facility then SuggestedCare becomes SpecificCare. CareAtHome is suggested if Patient can be taken cared and treated at home. Once a Patient is cured or almost cured he/she may be requiring FollowUpCare.

Another observation is that some medical problems or diseases like jaundice, diarrhoea, malnutrition are treated as members of ClinicalSigns.

5 VERIFICATION AND VALIDATION OF THE ONTOLOGY

In [14] different numerical metrics are used to evaluate an ontology like relationship, attribute, inheritance richness, knowledge base, class richness, class connectivity and class importance, cohesion. It is felt that health domain has some compulsion to not to satisfy such requirements always properly. This is because health domain, depends on requirements of medical science, human anatomy, diseases, clinical signs, symptoms, assessment facts. This can not or must not be measured in numerical metrics. Thus, health ontology must not be evaluated in terms of different numerical parameters on relationships, attributes, inheritances, class connectivity, class usage, cohesion etc.

This ontology has been verified and validated using the concepts stated in [13]. In our case, realizing that we are working in the domain of health data, we have reused existing ontologies in order to improve its credibility.

This ontology is for storage of health and other related data for a remote health framework using different sensors and cloud. It stores health data mentioned in [5] and [6]. It has been realized that the ontology that we are developing will require additions and alterations to accommodate wider scope of usage in health domain as well as for usage of more sensors and new facilities of cloud. The domain is health domain and for wider usage, it may require future extensions. For example it does not cover gynecological problems or cancer. These topics may be added later. Thus the ontology may go through modification for minor or major changes like addition of new diseases.

The ontology is developed using Protege 5.0 [30]. In this work, we have tried to work with currently available versions of different softwares.

Verification - In this work, our frame of reference was supposed to be “requirements specification, competency questions, real world” [13]. Our aim has been to develop the ontology so that it satisfies the requirements of IMNCI and the basic requirements specified in Minimum Data Set of Indian Standard for Electronic Health Records [5]. HermiT 1.3.8 Reasoner available with Protege was used to verify if there is any error at every step of development.

We have verified the ontology by adding individuals from real world and by discussing and verifying it by a domain expert.

Validation - Here we ensure that the ontology models the real world it is supposed to model. We have evaluated this ontology by both methods that are used to evaluate an ontology dependent and independent of the tool used to develop the ontology.

We have used ontology validator tool Manchester [28] to validate the ontology as specified by ontology validator tool in w3C [29]. This tool is independent of the tool used to develop the ontology. Methodologies suggested in [14] have also been studied for checking the ontology. Different types of lexical checks on the names given and checks on measuring the complexity of an ontology are available along with other checks. Whether an appropriate natural language explanation is given for a class can also be checked.

It is felt after going through [13], [14] and [15] that there must be some usage of artificial intelligence to study and find out the domain knowledge expressed in any ontology. We have not come across any such ontology validation and verification method. OntoCheck of [15] is dependent on the tool used to develop the ontology.

The ontology has been checked manually for the following errors as mentioned in [13] and the results are as follows: Inconsistency - It does not contain circulatory error, partition error or semantic error [13].

Incompleteness - It does not seem to miss any important concept as of now keeping in mind [6] and [5]. All the required parts are used and implemented. But we do not claim that the ontology will not change or develop further when there is change in requirements in the domain in future. All disjoint classes are defined: for example manually fed investigation type vs. directly fed investigation type, ContinuousMonitoring vs. DiscreteMonitoring.

Grammatical errors are not present, identical formal definition of classes and instances are not found. Our ontology is an application ontology and it has been validated by domain expert.

It has been checked manually that it is consistent as it does not contain contradictions. It is concise as it does not contain any redundancy. It seems to be expandable according to the domain. It is not sensitive as “small changes in definition doesn’t alter a set of well defined concepts” [13].

6 CONCLUSION

In this paper, an ontology about general terms for health data of a remote health framework using sensor, is proposed. It contains health and care data of infants and children upto five years following IMNCI [6] guidelines. This ontology uses existing ontologies on Electronic Health Record, Sensor and other related available ontologies. The purpose of developing this ontology is not to extract, learn, verify such health data available in web page automatically but to represent health data of remote health framework based on sensors and cloud. But it may be used for this purpose with improvement. It is verified and validated as mentioned before.

Ontologies are domain specific, application specific, level of detail specific. Available ontologies may be reused with or without
It is found that ontology development process is not an interconnected process. There is no communication among the different parts of the world regarding this matter. Few ontologies are shared by making them available to public.

Health sector is such a sector that standard ontology in this domain with different levels of detail and different scopes are needed. Health ontology requires generic coverage as well as specific coverage depending on the scope of the implementation and its requirement. The level of detail of different attributes and entities are application specific and depends on the scope and requirement of the application.

An ontology validation and verification method must incorporate artificial intelligence based approach to find and check correctness of domain knowledge represented in an ontology.

Ontology may be used to convert a generic data model to a data model of Big data solution.

REFERENCES


Figure 3: Procedure to treat a complaint

REFERENCES


[27] http://ontologydesignpatterns.org/ont/dul/DUL.owl
[29] https://www.w3.org/2001/sw/wiki/owl_validator