Child cancer follow-up ontology and information system

J. A. El-Subaihi[1], F. Anton[2], D. Mioc[2], S. V. Andersen[3]
& B. Lausen[4]
[1]Department of Applied Mathematics and Computer Science, Technical University of Denmark, Denmark
[2]Division of Geodesy, National Space Institute, Technical University of Denmark, Denmark
[3]Pallas Informatik A/S, Denmark
[4]Department of Pediatrics and Adolescent Medicine, Rigshospitalet, University of Copenhagen, Denmark

Abstract

An increase in chronic diseases in Danish healthcare can be explained by the corresponding increase of population longevity. Health professionals will not be able to keep up with treating those diseases, due to the many existing and new cases of chronic diseases. This results in mistakes in treatment processes, compensations to patients due to medical negligence and duplication of work and effort. In order to address a solution for healthcare practitioners, a small sub-group of patients and diseases is chosen from all chronic diseases. Namely, children diagnosed with cancer.

This research brings the methodology for child cancer treatment plan that produces an ontology to create a conceptual model and a database model. To construct the ontology, the "methontology" method is used as a structured approach for the ontology process. The method guides the ontology developer from scratch to building a complete model. The ontology is developed in two phases. In the first phase, research from other countries and process models are reviewed and the generic model is built from this research. The generic model is adapted to the ontology for the Danish hospitals including the NOPHO-ALL 2008 protocol. To develop the ontology, a data dictionary is first proposed. Then, the relationships between concepts are identified and verified: the oriented graph, where nodes are concepts and oriented edges are dependence relationships,

where the definition of the concept at the origin of the edge depends on the concept at the destination of the edge, must be a directed acyclic graph. Finally, the ontology resulting from the previous steps is implemented in Protégé-OWL. The conceptual model follows directly and univocally from the ontology: an entity-relationship diagram in UML notation.

Keywords: child cancer follow-up, ontology, information system, ERD.

1 Introduction

In Denmark, there are 53 hospitals with a capacity of 18.303 patients. The distribution of the staff is: 14 % doctors, 22% nurses, 23 % healthcare personal and the last 40 % are diverse professionals, like psychologist, cleaning, IT-department etc. [1]. The hospitals have experienced an increase of 25% in treated patients from 2007 to 2010 [1]. One of the biggest issues in the healthcare system is the increasing number of chronic ill patients. Generally, the hospitals and healthcare use a lot of time and resources on treatments plans for chronic ill patients, both before and after diagnosis. This is done with handwritten individualized plans. Diagnosis is a time consuming process and a lot of mistakes can occur [2–5]. Therefore, research for developing and handling treatments, guidelines and protocols, by computer based systems is necessary.

It is documented that there is a big saving on resources with the right treatment plan after the chronic ill patient gets diagnosed [6, 7]. In 2011, the healthcare system reported 35.000 events in hospitalization that have resulted in an injury of the patient [8]. The major causes reported are medication errors and treatment mistakes. They have resulted in a payout of 814 million DKK from the regions [9] – this is expected to increase, due to the investigations of the remaining cases. But it also means that those patients, who got the wrong treatment, are readmitted because of malpractice and the hospitals are obligated to spend enormous resources to treat the patients again [9]. Because of the lack of resources, the Danish healthcare system had made an effort on clearly describing the process of treatment of different chronic diseases [10]. The intent was to have uniformity in the disease handling internationally, but still it is in paper format and some treatments decisions are made ad-hoc. This clearly shows the lack of usage of information technology in the treatment of the patient and uniformity is not met, since the guidelines usually exist in paper format. By adopting information technology in health care, patient care and knowledge will improve, and treatment time and cost will decrease. An information system has to be developed to handle the increasing number of patients [3]. The advantages of using ontology as a representation model for treatment plans are that they have the necessary expressiveness and reasoning capabilities to use the knowledge ontology model to recommend task specific guidelines. The ontology has also enough information to support expressiveness and reasoning capabilities, to develop the decision system [11, 12].

By using computerized treatment plans, the benefit is clear [11, 12]:

1. Savings of recruitment of healthcare professionals, which leads to saving of payroll.

2. Saving of time, this leads to redistribution of resources.
3. Reduction of recurring errors in treatment, this will have the effect of saving human lives, patient complication due to wrong treatment and savings of compensation given to patients because of wrong treatment.
4. Uniformity in treatment, which implies even though the patients move from one region to another region and are getting treatment in other hospitals, the healthcare practitioners can continue with the prescribed treatment plan.

2   Methodology

The "Methontology" methodology was constructed in 1997, and was based on an IEEE standard for software development [13]. This methodology can be used to build completely new ontologies or can be used to further develop existing ontologies [13–15]. An ontology is typically built in a lifecycle way i.e., specification, conceptualization, formalization, implementation and maintenance. While developing the ontology, the methodology guides to specific techniques performing each activity [13]. The methodology uses a middle-out strategy, and the main phase is the conceptualization phase, which defines the organization of the specific knowledge, that is independent of the environment [13–15].

Domain: The domain of the child cancer treatment plan is child cancer management. The ontologies purpose is to define the different task involved with hospitalized children with cancer.

Purpose: The purpose is to develop a child cancer treatment plan for management of child cancer in the hospital. The intended development is to propose a model to create an automated system for the hospital to be used in child cancer treatment applications.

Knowledge sources: The knowledge sources of child cancer treatment plans exist (from discovery to treatment) in research papers, meetings with hospital doctors and in the World Wide Web.

Conceptualization: Conceptualization is done by middle out approach; the main idea is to develop a generic model, by following the three stated points:

- Identify the key concepts
- Align the similar concepts
- Relationships between the key concepts are identified

Evaluation: Should be made by different treatment plans and by using a reasoning engine [16].

2.1.1 Web ontology language (OWL)

The OWL language is used for creating, managing and sharing data among ontologies [17], which are recommended by the World Wide Web consortium (W3C) [18].

OWL is an expressive ontology description of complex relations between concepts. In other words, it is a mark-up language for publishing and sharing data using ontologies on the Web and was designed for being read by computer applications [17]. It is derived from web ontology languages. The syntax is

normally RDF/XML tags and schema. This is the most used language to represent ontologies [17, 19–21].

RDF has a known model description that is known in AI, the framework metadata description model {object/value, model, attribute} [17, 19–21]. RDF models use XML-Based serialization, which looks like the XML mark-up language.

2.1.2 Protégé

Protégé makes it possible to make knowledge-based applications and construction of domain model with ontologies [19–21]. It supports creation, visualization and manipulation of ontologies. This is also the most common used editor to build ontologies. It is written in Java, and it is free. Protégé supports a lot of different database management systems that have JDBC drivers.

2.2 Current practices for child oncology treatment plans in Denmark – Rigshospitalet

The oncology department for children at Rigshospitalet, Denmark treats children who are diagnosed with cancer. It has almost 50 different international treatments protocols [22], in which they treat children with cancer. The department makes their own treatment plans guidelines from the protocol using Microsoft Word and they are printed as a PDF file, so no changes can be added. The treatment plans are based on the treatments protocols [22].

There is a guiding line of who is making what. The doctors decide the treatment and the nurses follow the guidelines from the doctors and the treatments plans for the patient [22]. The diagnose and the treatment of children are made more difficult because there are many different types of child cancer, treatments plans and chemotherapy treatments that are different from adult types of cancer [22]. There can be delays in the treatments, due to other side effects or bad responses to the specific treatment [22].

The treatment is decomposed into treatment blocks or repeated treatments with specific treatments days. The first day of a patient's treatment is noted as treatment day 1. A treatment cure could last from 1 to 30 days, the most obvious treatment consists of several cures, and the cures could last 3-5 days. A treatment could be composed of more treatment regimes, and the estimated length of the treatments is long, sometimes between a half year to almost two years and half [22]. The treatments beginning are already timed-stamped from the protocol with specific cures days or blocks. Whether the treatment can start or be postponed depends often on blood count values. If the blood count values are fulfilled to the standards, then the treatment can be started. If not, then the treatment has to be postponed until the blood counts values are good again, for starting the treatment. All the chemotherapy treatments have to be given in a very intensive way, which means the treatments within a block cannot be delayed or postponed, because of a bad treatment response, due to the survival [22].

2.2.1 The ALL-2008 Treatment Protocol short description

The Nordic countries started to cooperate in 1981 with a goal to treat diagnosed children with ALL (Acute Lymphoblastic Leukemia, all this section has been written after consulting [23]). That cooperation has resulted in an increase of the healed children rate from almost 50% to 75%. The protocol to the ALL treatment keeps changing and due to the experience and continued cooperation between the countries, it aims for better survival of ALL patients. The latest protocol to treat ALL is NOPHO-ALL 2008. The protocol used to treat ALL is made by the Nordic Society of Paediatric Haematology and Oncology (NOPHO), which is a research group with participants of Denmark, Finland, Norway, Sweden, Iceland, Estonia, Lithuania and Latvia. When the patients are diagnosed with leukemia, the chemotherapy treatments are applied according to the NOPHO ALL 2008 protocol. The first treatment phases in the protocol are the "Initial treatment", which lasts 4 weeks. In the second and fourth week in the initial state, the bone marrow is checked. The effects of the treatment are checked by bone marrow examination after 2 and 4 weeks of treatment.   Depending on the bone marrow results, the patient is divided into 3 different groups:

Standard risk (SR): Are for those children who have leukemia with fewer than 100 billion leucocytes / L at the beginning of the disease and responds well to treatment. The treatment duration is 2 ½ years. This group is about half of all children with ALL.

High risk (HR): Are for those children who have unfavorable disease factors – WBC over 100 billion. / L, T-cell leukemia, certain chromosomal abnormalities in leukemia cells – and has shown poor efficacy of starting treatment. The duration of treatment is 2½ years. This group is about 10% of children and adolescents with ALL.

Intermediate risk (IR): The remaining 35% of children and adolescents with ALL. The duration of treatment is 2½ years. Re-induction as the SR, but which is slightly more intense. It is two times more during the process, and to reduce the amount of residual disease.

     The Generic model is designed by using the information found on treatment plans for children with cancer, by using a step-by-step approach, which is:

1. Identifying the concepts, shown in Table 1.
2. Align the similar concepts, shown in Table 2.
3. Identifying new concepts from the process plans.
     The final Generic model after several iterations is shown in Figure 1.

3   Data dictionary

The Danish healthcare system has only data dictionary definitions for central health specialist concepts [24]. The data dictionary does not include concepts for the treatment and roles in hospitals. A proposal for a data dictionary based on information derived from the protocol and the proposed generic ontology model

Table 1:   Some the concepts and sub-concepts identified from treatment plans.

Concept	Sub concept
Hospital diagnose	Physical examination
Hospital treatment	Surgery
Hospital monitoring	Investigations
	Re-consulting
Hospital roles	
Hospital discharge	
Follow up	
Patient home observation	

Table 2:   The concepts identified and aligned from the identification.

Concept	Sub concept	
Hospital Diagnose	Blood test	
	Samples	Biopsy
		Bone marrow
	Scans	MRI
		CT
		Bone
		PET
		X-ray
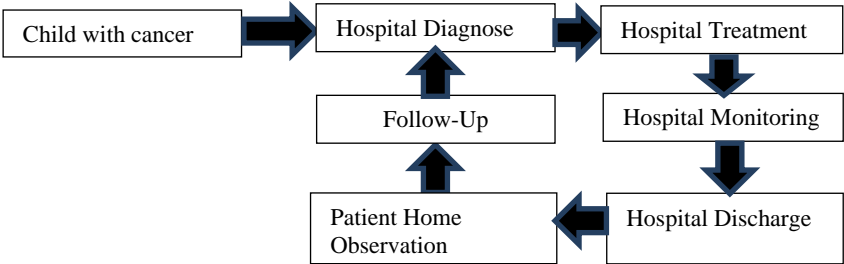
Figure 1:   The final generic model.

is made. The following shows some of the concepts (showed in bold) and the definitions of the concepts.

Hospital Department: A specific institution for care and treatment of child cancer [25]. It is the children's department with highly specialized functions in oncology [26]. The diagnoses of the children are made in the hospital department and the follow-ups for the children who were receiving treatment for cancer are also done in the hospital department. The oncologist and the nurse, who are

trained to care for children with cancer, work in the hospital department. The hospital department has rooms with room numbers, where beds for the admitted children are provided if they are admitted, when going through diagnose and treatment process.

Hospital department admission: The hospital department acceptance of providing health care facilities to the child, who is provided with a room and nursing [26]. The child with cancer resides at least overnight during the treatment process [27], which means that a room and a bed are provided for the child.

Hospital department outpatient admission: The hospital department acceptance of providing health care facilities to the child, who is receiving treatment and leaving after the treatment [28]. The child does not require a bed. This is decided by the multidisciplinary team and by the protocol treatment guidelines.

The final identification and descriptions of the concepts resulted in an identification of 65 different concepts.

4   Ontology development

Before implementing the data dictionary in Protégé-OWL, the explanation of the relationships between concepts are further shown in Table 4.Entity # 1 is the concept which is relating to the other concept entity # 2. The approach for defining each relationship between concepts is simply modus tollens from logic, i.e., if A → B and B is false, then A is false. That led to 101 different relationships identified between the concepts (see Figure 2 and Table 3).

Table 3:   The definition relationships that the diagnose concept induces.

| Entity # 1	Entity # 2	Description
Diagnose	Hospital Department	Diagnose is made in the hospital department specialized in oncology, if the hospital department specialized in oncology does not exist, then diagnose cannot be made.
Diagnose	Physical examination	A part of diagnose is the physical examination, if the physical examination is not done, then diagnose is incomplete.
Diagnose	Scans for children over the age of 6-8 years	A part of diagnose for children over 6-8 years is scans, if the scans are not made, then diagnose is incomplete.
Diagnose	Scans for children under the age of 6-8 years	A part of diagnose for children under 6-8 years is scans is scans, if the scans is not made, then diagnose is incomplete.
Diagnose	Samples	A part of diagnose is samples, if the samples are not made, then diagnose is incomplete.
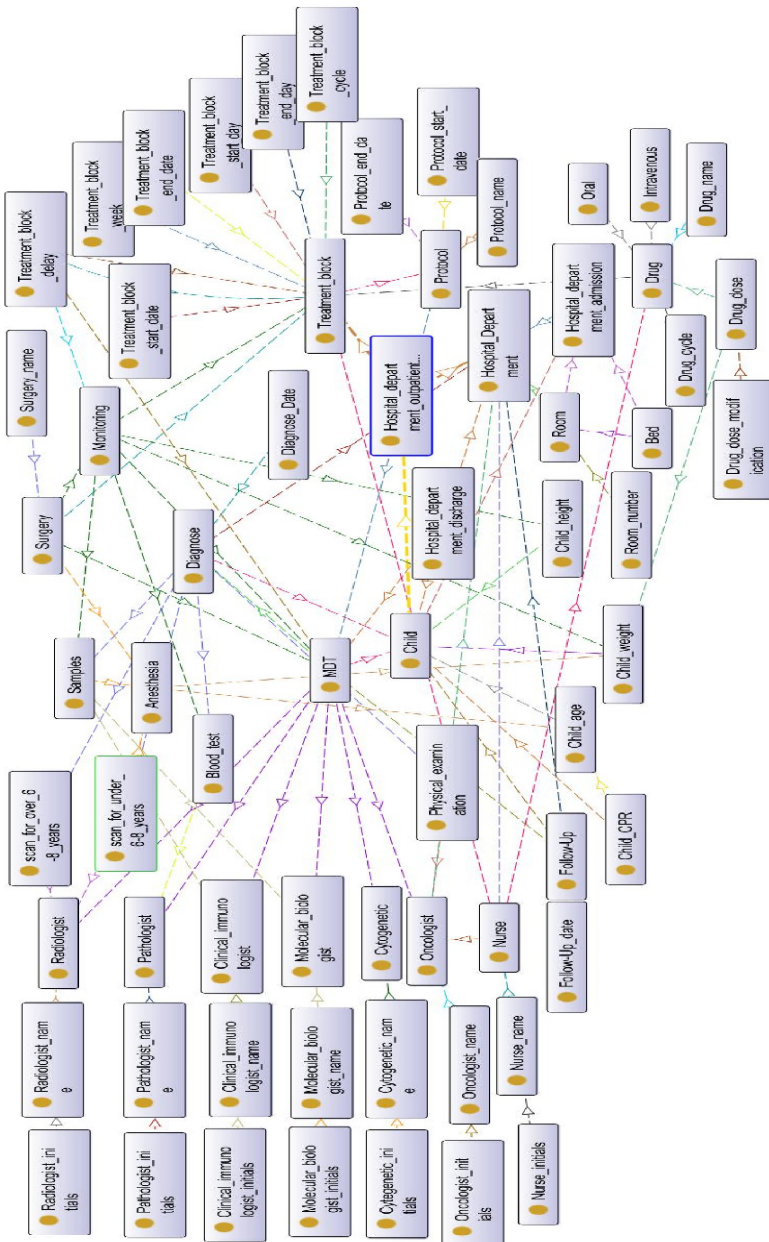
Figure 2:   The relationships between the concepts. Represented as a DAG.

Example: If Anesthesia implies Surgery, then it must be true that if surgery does not exist, then anesthesia does not exist either. This is clearly false, since anesthesia is also used for the scans in some cases and samples. So the right sentence is surgery implies anesthesia. Then it must be true that if anesthesia does not exist, then the surgery does not exist, since surgery cannot be done without anesthesia.

4.1 Logical model for the data dictionary

The ontology model in Figure 2 is a directed acyclic graph (DAG), which shows the relations between the concepts. The directed edges shown in the ontology model represent which concepts define other concepts, as explained in Table 3.

The relationships between the concepts identified during the ontology development imply also the cardinality of the corresponding relationships between the corresponding entities in the database modeling. The next steps towards the logical model are made from the acyclic graph. That leads to defining the format/length of the entities corresponding to the concepts identified in the ontology model, where the domain is the concept name, which is a class and the object properties are the format and range definition depending on the type of the format.

4.2 Entity-relationship diagram (ERD)

The definition for the logical model in a database system is represented by the concepts assembled in one schema with the relationships between them, the attributes for the entities and the primary keys are specified [29]. The ontology presented in Figure 2 is transformed into an ERD (in UML), shown in Figure 3.

5   Conclusions

This paper has described how an ontology has been built by using the Protégé development tool, with the OWL language. Secondly, it showed that using the methontology method for developing the ontology for child cancer treatment and follow-up has resulted in a clear structured method for the process of the treatment planning of child cancer.

This leads to the development of the data dictionary: identification of key concepts and aligning key concepts through several refinements. The concepts and the relationship between these concepts were established, which meant that the ontology could be implemented and no inconsistencies were found.

After developing the ontology of the generic model, the ontology was modified to the NOPHO ALL-2008 treatment protocol, with the hospital processes identified. The identification of concepts and the description of those were made. From the data dictionary, the ontology model for the concepts was made and implemented as classes in Protégé-OWL. The ontology development resulted in an acyclic graph that shows the causal relationships between definitions of concepts. The ontology model was further used to develop the logical model. The logical model was derived by defining the format and length

Figure 3:   The entity-relationship diagram.

of each concept's data type. The ERD was obtained from the ontology by mapping concepts to either entities or their attributes and relationships between concepts to relationships between the corresponding entities with their cardinalities.

Creating the physical database model in the respective departments within the hospitals will have a potentially big impact on the treatment processes. This leads to a reduction of medical errors, more available time for each patient, and savings on compensations to the patient and better resource allocation. Furthermore, every time a treatment process plan or treatment plan is changed due to the patient's response on the treatment, it is not necessary to start over by reading the patient journal or protocol, but simply change or delay the planned treatment. This will result in savings in time, reducing administrative tasks, reducing the errors and providing more time for patient care.

References

[1] Henning, L. (2012). Fakta om sundhedsvæsenet - sundhedsvæsenet i tal. http://www.regioner.dk/Aktuelt/Temaer/Fakta+om+regionernes+effektivite t+og+%C3%B8konomi/Kopi+af+Fakta+om+sundhedsv%C3%A6senet.asp x

[2] Bjørnsson, K. (03-10-1997). Samarbejdet sikrer mod fejlbehandling, Online.   http://www.dsr.dk/Sygeplejersken/Sider/SY-1997-40-32-2-Diverse.aspx

[3] Lesh, A.K. and Laguerre, R. (4. 1. 2007). United States Patent Application Publication, Pub. No.: US 2007/0005621 A1 Information systems using healthcare ontology

[4] Chalortham, N. Buranarach, M. and Supnithi, T. (2009). Ontology Development for Type II Diabetes Mellitus Clinical Support System, Online. www.text.hlt.nectec.or.th/marut/papers/dcare_kicss2009.pdf

[5] Sesen, M.B., Banares-Alcantara, R., Fox, J., Kadir, T. and Brady, M. J., OWL-ED (2012). Lung Cancer Assistant: An Ontology-Driven, Online Decision Support Prototype for Lung Cancer Treatment Selection. www.webont.org/owled/2012/papers/paper_22.pdf

[6] Beck, L. (2. 4. 2012), Medwatch. Millioner at spare på kroniske syge patienter. www.medwatch.dk/Sygdom___Sundhed/article4279774.ece

[7] N, F. and McGuinness, L. D. (2001). Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880, March 2001. *Ontology Development 101: A Guide to Creating Your First Ontology*

[8] Flemming, P. S; Berlingske (2012). Patienter må være på vagt for at undgå fejlbehandling.Online.   www.b.dk/nationalt/patienter-maa-vaere-paa-vagt-for-at-undgaa-fejlbehandling

[9] Ritzau; Børsen (2012). Fejlbehandlinger koster 814 mio.kr. Online. www.borsen.dk/nyheder/politik/artikel/1/224691/fejlbehandlinger_koster_8 14_mio_kr.html

[10] Sundhedsstyrelsen 200; (11. 04. 2008). Forløbsprogrammer for kronisk sygdom – Generisk model og forløbsprogram for diabetes. www.sst.dk/publ/publ2008/Plan/Kroniks/kronisk_forløbsprogrammer2008. pdf

[11] The K4CARE. Knowledge Based Homecare eServices for an Ageing Europe. www.k4care.net/

[12] Ilaria, C. (2007). The University of Leeds School of Computing. Ontology-based Representation and Reasoning about the History of Science.

[13] Abela, C., Department of Intelligent Computer Systems, Faculty of ICT at the University of Malta, Lecture notes (n.d). Methodologies and Methods for Building Ontologies www.Staff.um.edu.mt/cabe2/lectures/ webscience/docs/ontology_engineering.pdf

[14] Fernández-López, M. and Gómez-Pérez, A. (2002). Research project funded by the IST Programme of the Commission of the European Communities as project number IST-2000-29243. OntoWeb Deliverable 1.4: A survey on methodologies for developing, maintaining, evaluating and reengineering ontologies

[15] Sharif, S., Dalhousie University (2012). Characterization of longitudinal Care plan Model For managing chronic diseases: A care plan ontology to computerize paper-based care plans

[16] Reasoner Definition; www.en.wikipedia.org/wiki/Semantic_reasoner

[17] W3C OWL Web Ontology Language Reference. www.w3.org/TR/owl-ref/

[18] W3Ci, www.en.wikipedia.org/wiki/World_Wide_Web_Consortium

[19] *Ceusters*, W., Santos, D.C.M., and Fielding, M., Language and Computing, Making computers understand natural language. Online. www.ifomis.org/Downloads/Ceusters.ppt

[20] Smith, M.K., Welty, C. and McGuinness, L. D. (2009). OWL Web Ontology Language Guide, W3C Recommendation.Online. www.w3.org/TR/owl-guide/

[21] Protégé, www.protege.stanford.edu/overview/protege-owl.html

[22] Person, B. L., Personal communication, 07 March 2012, Department of Pediatrics and Adolescent Medicine, Rigshospitalet, University of Copenhagen, Denmark

[23] NOPHO ALL Protocol-2008. Online. www.nopho.org

[24] Sundhedsstyrrelsen Begrebsbasen, Online. www.begrebsbasen.sst.dk/Document/4a7da506-6210-4b76-856b-048e2e380dd4.htm

[25] National cancer institute at the national institute of health. NCI Dictionary of Cancer Terms, Online. www.cancer.gov/dictionary

[26] Sundhedsstyrelsen (2011). Pakkeforløb for kræft hos børn, Online. www.sst.dk/publ/Publ2012/SYB/KPforloeb/KraeftBoern.pdf

[27] The Free Dictionary by Farlex, Online. www.medical-dictionary.thefreedictionary.com/hospital+admission

[28] National Cancer institute, www.cancer.gov/dictionary?CdrID=44520

[29] Elmasri, R. and Navathe, S. B. sixth edition Pearson; Database Systems Models, Languages, Design, And Application Programming