Bachelor Thesis

"Data Modeling of Health Information Systems with openEHR"

Author:	Christina Pahl
Subject:	Biomedical Engineering
Matriculation number:	42212
Date and place of release:	14.06.2011, Ilmenau

Responsable Professor:

Prof. Dr. Vesselin Detschew
TU Ilmenau, Fakultät für Informatik und Automatisierung, Institut BMTI, PF 100565, 98684 Ilmenau, DE, vesko.detschew@tu-ilmenau.de

Advisor:

Dr. Daniel Weingärtner
Universidade Federal do Paraná, Departamento de Informática, Po.Box 19081, 81531-990 - Curitiba, Paraná, BR, danielw@inf.ufpr.br

Abstract

The increasing demand for more efficiency in healthcare results in new approaches of electronic healthcare record systems. Hence, the openEHR foundation developed an approach defining specifications for electronic health record (EHR) systems. The key innovation relating to a multi-level approach is the separation of clinical data from record keeping concerns. This approach leads to the coveted interoperability between different healthcare information systems.

The approach consists of two levels. The first level is built by a stable reference model (RM). The second level is alterable and represents restrictions of the first level which are called archetypes. Advantages arise from the use of one RM which causes interoperability and the reuse of archetypes which leads to rapid development.

This thesis investigates in how far the openEHR approach is usable for the modeling of healthcare domain related data provided by the national health system (SUS) of Brazil. SUS offers universal free access to health services. Using real data, the author modeled a template consisting of archetypes which were reused or edited. In contrast to the already available scientific works, this thesis analyzes that topic from the viewpoint of a biomedical engineer combining computational and medical skills accordingly.

The result is that the major part of prenatal and demographic data is represented by already existing unchanged archetypes. The other part required merely few modifications of archetypes. Altogether, the representation of prenatal data recorded in the local health unit (HU) of the city of Colombo in Brazil required no creation of archetypes. All data was almost entirely represented within archetypes provided on the clinical knowledge manager (CKM).

The conclusion is that the complexity built by the vast specifications requires a team to build new archetypes properly. However, pregnancy related clinical data demanded from the Brazilian Health Ministry and recorded in the HU was nearly completely represented by already existing archetypes.

This work offers a detailed look at the openEHR theory, its practical implementation and finally an evaluation of the process.

Affidavit

I hereby declare that the following bachelor thesis "Data Modeling of Health Information Systems with openEHR" has been written independently and on my own without using any other sources and aids as I stated. Where I used other sources I clearly marked them as not my own. This thesis has not been received by any examination board, neither in this nor in a similar form.

Ilmenau, 14.06.2011

_____

Christina Pahl

Note of thanks

I thank my responsible Professor Dr. Vesselin Detschew from the Ilmenau University of Technology (TUI), with whose assistance this thesis could only be developed. I am also indebted to my advisor Dr. Daniel Weingärtner from the Federal State University of Paraná (UFPR), who enabled this work as a part of a semester abroad and provided competent support on all my questions. I would also like to express my gratitude to Sergio Freire from the Federal State University of Rio de Janeiro (UFRJ), who supported my work through his experience in archetype modeling. I also take this opportunity to thank Dr. Luciana Schleder Gonçalves and Marcelli Sampaio de Almeida from the UFPR, whose explanations of clinical terms used in Brazilian Healthcare System have been incorporated in this work. Further contributions given by Heather Leslie, director of clinical modeling, and Marco Aurélio Borges are also acknowledged hereby. Finally, I would like to thank the graduate students from the UFPR to whom I have had the pleasure to work out the demands of the HU, friends and family. Opinions expressed in these notes are mine and do neither necessarily reflect the policy of UFPR nor of TUI.

Table of content

1. Introduction

This thesis tries to investigate if openEHR is applicable to the representation of real data based on records done in Brazilian HUs and demanded by the Health Ministry of Brazil.

Clinical software based on the same openEHR RM can communicate with any other system using this standard and complies thus the needs towards interoperability. Unlike the ordinary way to build medical knowledge and information on a same single-layer, the openEHR approach separates information from knowledge with a multi-level strategy. Here, information can be considered as "statements about specific entities", whereas knowledge as "statements which apply to all entities of a class".[1] In other words, knowledge relates to individual facts like clinical data and information to stable facts like Systematized Nomenclature of Human and Veterinary Medicine (SNOMED) for example. This separation also results in better maintainability. Research has revealed that the division of the generic information into a small model and the domain knowledge into a large model improves interoperability.[2] OpenEHR uses exactly this type of separation, although it was also shown that this separation induces less flexibility in development.

Any modification in the healthcare domain only affects the knowledge of the second layer, where the Archetype Model (AM) is located. The actual information layer remains the same. Standardized data representation on the RM enables unified interfaces and interoperability. This results in an easier and cheaper maintenance and durability. Any archetype-based Hospital Information System requires a constant involvement of healthcare providers to be updated regarding the latest clinical data[3]. Healthcare is facing different problems nowadays. Among these are the size of the clinical domain and its rate of change. Here, openEHR tries to provide future-proof systems. Due to multi-contact health, a promising approach towards improved interoperability of information is offered. Furthermore, decision support and

1 Beale: Archetypes and the EHR, p.17, 2002
2 Michelson et al.: Comparing different approaches to two-level modeling of electronic health records, 2005
3 Păun et al.: Local EHR Management Based on openEHR and EN13606,p. 5, 2009

population medicine can be provided through queryable longitudinal data[4]. Therefore, observation of standardized data over time and space is provided[5].

The theoretical part of this thesis deals with the openEHR architecture and its components. Furthermore, the practical part examines the use of this approach. Here, forms just like they are used to record patient related data form clinical content for the application. Tools provided are used for electronic representation of demanded data. The development environment will be presented and evaluated.

In December 2010, the Brazilian government has already started giving consideration as to how this approach may be implemented into its healthcare system. At the same time the department for informatics at the UFPR in the city of Curitiba held a symposium on this topic. This work emerges from controversial discussions on this potential approach.

1.1 Fields of application

The exchange of information between different healthcare suppliers is of increasing importance. Since patients become mobile and studies demanding comparability are carried out, interoperability is an essential requirement for health record systems. Many organizations, from local institutions to powerful international healthcare organizations, are interested in solutions guaranteeing interoperability.

Experiences meeting the demands of interoperability by the use of archetypes in EHR systems are still rarely reported. The point is that there are no studies about large clinical environments with equipment for more than 600 in-patients or similar using this approach.

Swedish researchers investigated in 2009 the convertibility of EHR models. The research group used a national EHR product called COSMIC[6] as a template model and the specifications of the openEHR RM and AM. By the mapping between the local template model and the RM/AM a conversion preserving nearly all structural and semantic definitions of the original models has been achieved. Regional templates have been converted into the

4 Beale: Archetypes and the EHR, p.8, 2002
5 http://weber.ucsd.edu/~tkousser/Beck%20Notes/longitude20041short.pdf
6 Rong et al.: Archetype-based conversion of EHR content models: pilot experience with a regional EHR system, 2009

openEHR archetype format. Hereby, a method has been formulated enabling the sharing of different model formats.

Denmark was investigating if openEHR could be used in a national context[7]. Knut Bernstein, a national pioneer in research on the implementation of the openEHR RM, looked for methods in order to implement archetypes and templates on national level. Therefore, he tried out some elements of the openEHR methodology in cooperation with local hospitals. Benefits during the use of archetypes were reported. The ease of operation, stability and security were considered as positive. The national generic IM called BEHR[8] (Basic Structure for Electronic Health Records) was compliant with the EHR system. The only problem that occurred, was the unacceptability of the standards which were elaborated by the National Board of Health of Denmark by healthcare professionals.

In order to facilitate data storage in clinical trials, the University of Heidelberg in Germany explores how clinical data sets can be expressed with openEHR archetypes. A report[9] published in 2010 describes electronic patient record containing archetyped data. The benefits using archetypes were described as: integration of existing data sets, multiple use of mappings and meta analysis.

The University of Moratuwa in Sri Lanka is also doing research on an implementation of a national EHR based on the openEHR architecture. In relation to the publication of the year 2005, the Sri Lankan scientists intended to use openEHR in demographical matters to handle the patient profiles[10].

Brazil lists with four institutions one of the highest concentration of scientific institutions researching on the field of openEHR. In 2009 publications[11] documented archetypes built in order to fit the interface between epidemiological and clinical concepts. Furthermore investigations on the field of semantic interoperability were done using the RM and the AM.

7 Bernsteinet et al.: Can openEHR Archetypes Be Used in a National Context? The Danish Archetype Proof-of-Concept Project, 2009
8 Bruun-Rasmussen: National report of EHR implementation in Denmark ,p. 23, 2009
9 http://www.openehr.org/wiki/download/attachments/5996988/SecondaryUse_Kohl.pdf?version=1&modificationDate=1290696080000
10 http://www.openehr.org/shared-resources/usage/academic.pdf
11 http://www.openehr.org/shared-resources/usage/academic.html

A Hospital in São Paulo called, “Hospital das Clinicas”, belonging to its Federal University applied archetypes for clinical observations and radiological reports.

1.2 Aim

This paper shall be looked upon as an introduction to the openEHR approach. Furthermore, it shall give a description of the development process of archetypes with the most widespread operating System, Windows. Valuations arise from the viewpoint of a biomedical engineer. Furthermore, this work is consisting of the following points of interest:

1. **Familiarizing** with the openEHR approach
2. **Establishing** a development environment
    - 2.1 Medical forms
    - 2.2 Archetype editor
    - 2.3 Template designer
3. **Implementation**
4. **Evaluation**

The idea is to provide an overview over openEHR as well as the approach provided and subsequently implement this knowledge. The main aim is to use the environment offered, to do an electronic representation of real data. This includes archetypes and templates. Finally, management of data modeling within the openEHR environment will be examined.

1.3 Material and methods

Focusing on data distributed by the openEHR online community, scientific works dealing with the openEHR approach and data published by the Health Ministry of Brazil, theoretical basics are provided.

The information about openEHR, which is mainly provided by the foundation itself, comprises several specifications, explanations and clinical models. The homepage provides different access points, so that user with different backgrounds can get started with customized information. The author used technical as well as clinical information distributed on the same homepage. Moreover, the community supplies user with tools, which support authoring of archetypes and templates.

Furthermore, this work relates to various papers of which the majority is freely accessible online. Usually these papers describe the openEHR approach from the viewpoint of scientists applying it in different contexts. The criticism on the openEHR expressed therein flows for the sake of objectivity into this thesis. Moreover, various perspectives on openEHR lead to a better understanding of this approach. Data was taken from the HU in order to present it in a digital way. Further specific information was gained through the work with people dealing with information processing. Further data was taken from the official website of the Health Ministry of Brazil.

Through the change of location during writing process, email exchange became an important way to communicate. The author moved from a locality dealing with openEHR to a place, where no specialists exist working on the topic of openEHR. The process became independent from others. The application of researched tools was performed in a self-study. Provided support in form of manuals and videos was used to a certain extent. Some content was not compatible with latest versions. Due to poor documentation, the use of programs was based on experimental practical tests.

1.4 Type and location of research

According to the type of research[12] this work combines various types. First of all, this work takes part in an applied and active research as the intention is to represent data in form of a focused action and not target an acquisition of knowledge for the sake of itself. Further on, this research relates to development conducted to measure the usability of openEHR approach and therefore this research also relates to an evaluation research.

This work represents the concluding part of the Bachelor´s course in Biomedical Engineering at the Technical University of Ilmenau in Germany. Furthermore, it is based on the research at the department for Informatics at the Federal University of Paraná in the city of Curitiba in Brazil.

1.5 Structure of the thesis

The thesis is structured in six main chapters. Following this introduction, the second chapter is devoted to the theory of openEHR and illustrates the components and its relations to each

12 http://www.disability.wa.gov.au/Research/Definitions/ResearchTypes.htm

other. Afterwards, in chapter three, the problem is described more in depth. In the following fourth chapter the development process of the demographic and the clinical data is described in detail. Based on this practical part, in chapter five application results will be presented. At the end, in chapter six, conclusions will be drawn. All contributions will be shown and proposals for future research will be subsequently given.

2. Healthcare Record

The ancient Greek physician Hippokrates of Kos (ca. 460-370 BC) was probably the first person recognizing the importance of a health record. He taught to note medical findings and their treatments in a clear manner, so that they can be passed to prospective physicians[13]. Today handwritten data gets digitalized and current information is recorded to a large extent on computers.

2.1 EHR

In general, electronic health record implies the collection of digital health information about persons. This information encompasses past medical history, progress notes, vital signs and several other reports. Furthermore it integrates the healthcare data of the systems which follow: Administration, Clinical, Lab, Nursing, Pharmacy and Radiology. In conclusion, an EHR has the ability to support care-related activities. It can deliver evidence-based support and contribute to the clinical quality management.

2.1.1 EHR in HIS

Because of the complexity of HUs, it is necessary to find a solution facilitating the clinical daily routine. A hospital information system (HIS) describes a service consisting of several components supporting the management of information in a clinical context. The main advantage of the use of a HIS can be considered as the guarantee of quality in between the clinical environment. Besides other, it facilitates the process of decision making in clinical context by offering pathways, which are suggesting evidence based medicine.

Moreover, it offers a structured overview of patient data which can be connected via internet to a pharmacy calling attention to allergies and reactions towards administered drugs. In addition to the vantages for patient and medical staff, the finance administration is reaping the benefits of an information system in a decisive way. A system enables precise control of spending as well as automatic cost evaluation based on data, which is stored by the system and contributes to a well-organized finance budget.

13 Margotta: The Story of Medicine, p.68, 1968

Nowadays it is required that HISs ensure interoperability without losing security of patient information. Therefore, the discussion about the use of the openEHR approach, enabling EHR within a local clinical context is of interest.

2.2 OpenEHR

The openEHR community leaps on the EHR-idea and provides means for creating healthcare records consisting of demographic and clinical information.

Enabling health records by standard based, open source implementations, it dedicates itself to "support the needs of patients and clinicians everywhere"[14]. With regard to the self given definition, the online community claims to be non-profit and independent. There is no independent confirmation or refutation of these attributes.

As a proceeding of a non-profit organization, the use of the tools provided by the openEHR community is free of charge. The applications are compatible with the standard operating system Windows and available on the openEHR homepage[15]. Furthermore, joining the development group enables access to the source code of its applications. Actually, all components needed to model a clinical information system are available on the web.

OpenEHR is originally based on the European Union´s GEHR-Project. At this point, GEHR is an acronym for "Good European Health Record". This program was started in 1991 with the objective to create a common architecture for computerized health records and ended in 1995. Involving seven European countries the ventures intention was to develop an architecture meeting common acceptation and application across Europeans clinical computer systems[16]. In addition to that, the early GEHR ambition covering clinical demands is nowadays translated into action by the use of openEHR complying with the global needs.

At the present time, the implementation of EHR in healthcare systems is supported in several countries. A key reason using openEHR lies in a multi-level approach meeting the requirements of data exchange between different care providers. The layering follows a separation into the domain knowledge model and the generic information model. This separation leads to a division of functions.

---

14 Ingram: The origins of openEHR. openEHR Foundation, London, 2002, see: www.openehr.org
15 http://www.openehr.org/download/software.html
16 http://www.chime.ucl.ac.uk/work-areas/ehrs/GEHR/

While clinicians manage the clinical content, technicians manage the applications[17]. Thus, the development process gets more flexible[18]. The exchange of archetypes enables a joint cross-border development. International research teams are pursuing the goal of creating a cross-national communication standard for healthcare.

Flexibility is furthermore reached by the use of universal generic structures and types of data[19]. Data management gets facilitated by the use of those general structures and information. Especially the management of frequently changing models, as is common in healthcare, is simplified. Clinical content can be revised by clinicians modifying archetypes without impacting the supporting technical infrastructure.

2.2.1 Clinical aims

OpenEHR was designed as a record-keeping computer service with clinical care specified requirements enabling a facilitated interoperability.

The patient-centric openEHR offers solving and decision support for diagnostic purposes. This clinical decision support improves the patient data safety. Through repeated medical investigations the support has an economically highly important effect: reduction of costs. Another aim of openEHR is to integrate patient views with knowledge recourses. The vast body of knowledge includes not only guide lines but also terminologies as well as computerized libraries.

The advantages of clinical computerization can be seen in the decrease of errors as like as the improvement of communication besides others. Communicating data leads for example to the prevention of repeated investigations, more effective search for information and a better analysis of predictive risk factors. Regarding the above mentioned aims the patient attendance becomes significantly improved.

In the end, data of clinical relevance shall be exchanged satisfying the demands of security, comparability and therefore unambiguousness. One may limit the access to sensitive data so that only authorized persons have an insight. In order to understand person related data at all

17 Leslie et al.: Engaging Clinicians in Clinical Content: Herding Cats or Piece of Cake?
18 Demski: Kommunikation in ByMedConnect – Umsetzung des CCR mit EN13606, Deutschland, S.10, 2010
19 Brass et al.: Health Data Management with an Archetype Driven EHR System in Low Resource Environments, p.4, NS

it is of major importance that this information underlies unambiguousness. For this reason, it is necessary that the data relates to the same model of reference. Furthermore, the comparison of information for studies leads to assuming interoperability a special position.

2.2.2 Interoperability

The primary purpose of openEHR is the interoperability of data. Different healthcare providers shall be able to exchange information so that data can be interpreted properly. Here, interoperability is subdivided into:

- **Syntactic** interoperability
- **Semantic** interoperability

Syntactic interoperability is relating to the word syntax which means "orderly or systematic arrangement of parts or elements" in the English Oxford Dictionary. As an idea of a structured composition since Ancient Greek[20] correct syntax is a prerequisite between at least two systems in order to execute actions in a syntactic interoperable way[21]. An example for such an action would be the exchange of clinical information between three hospitals. Otherwise, the exchange of unstructured information would lead to obstacles in understanding and thus misunderstandings.

Syntactic interoperability is the cause of semantic interoperability. It is the "ability of two or more systems or components to exchange information and to use the information that has been exchanged"[22]. Everyone taking part in this action has to understand the exchanged information in the same way. Therefore, it is indispensable for the participating systems to refer to any kind of reference for purposes, a RM for example. Through the exchange of data with its semantic model (also known as archetypes), openEHR enables semantic interoperability.

20 http://csmt.uchicago.edu/glossary2004/syntax.htm
21 Eichelberg et al.: A Survey and Analysis of Electronic Healthcare Record Standards, p.2, 2005
22 IEEE Standard Computer Dictionary: A Compilation of IEEE Standard Computer Glossaries, New York, 1990

This openEHR purpose of health data enables personal health information "to be unambiguously recorded in one system, transported to another system and utilized as though the data had been authored natively within the second system"[23].

Integrating security into interoperable systems is a decisive task to do. A closer look on the methods keeping data safe will be given in the next paragraph.

2.3 Security

Maintaining the confidentiality in terms of a concealment of disclosed data as well as the privacy itself, in order to limit the insight in personal data, are necessarily considered demands.

Actually, the above mentioned requirements are standing in conflict to the claims of clinical professionals who demand for quick access to the personal data. There are two widely known approaches to reduce this problem.

One strategy provides a differentiation of the access. By separating sensitive data referring to sexual and mental health items from the rest of the health record, the insight in data can be demarked in different accession facilities. Considering the interrelatedness of health information, the deficiency of this solution is noticeable. Furthermore, possible conclusions out of a medication list make the protection of diagnosis pointless.

Another idea is to specify access control by defining roles. However, role-based access will always be problematic due to temporal changes. The in theory well executable approach does not keep up with the fluctuating reality of the clinical environment. Altogether, it is possible decreasing risk with above mentioned techniques to an acceptable restraint. The only problem here is usability.

2.3.1 OpenEHR security approach

The foundation supplies just the most necessary policy principles to avoid attacks. It rather passes the security responsibility to system developers.

23 Leslie et al.: Engaging Clinicians in Clinical Content: Herding Cats or Piece of Cake?

The general part of the guidelines[24] embodied in openEHR includes indelibility, audit trailing and anonymity. Hereafter a brief overview of the principles:

Indelibility depicts the impossibility of erasing already recorded information. Moreover, the foundation foresees audit trailing in order to save all the changes done to the EHR. Depending on the development configuration, anonymity is offered by separating data on different physical computers or servers. Hereby, an expropriation of only one information source impedes the attribution of the information to the personal identities or other way around.

Due to the subsidiary information about the principles considering the access control the paragraph forgoes its elaboration. Significant is that the access control is role-based and that defaults for the EHR can be set by patients.

2.4 EHR Model

The following UML diagram is relating to the IM overview[25] and shows how the healthcare consumer data becomes part of the EHR.

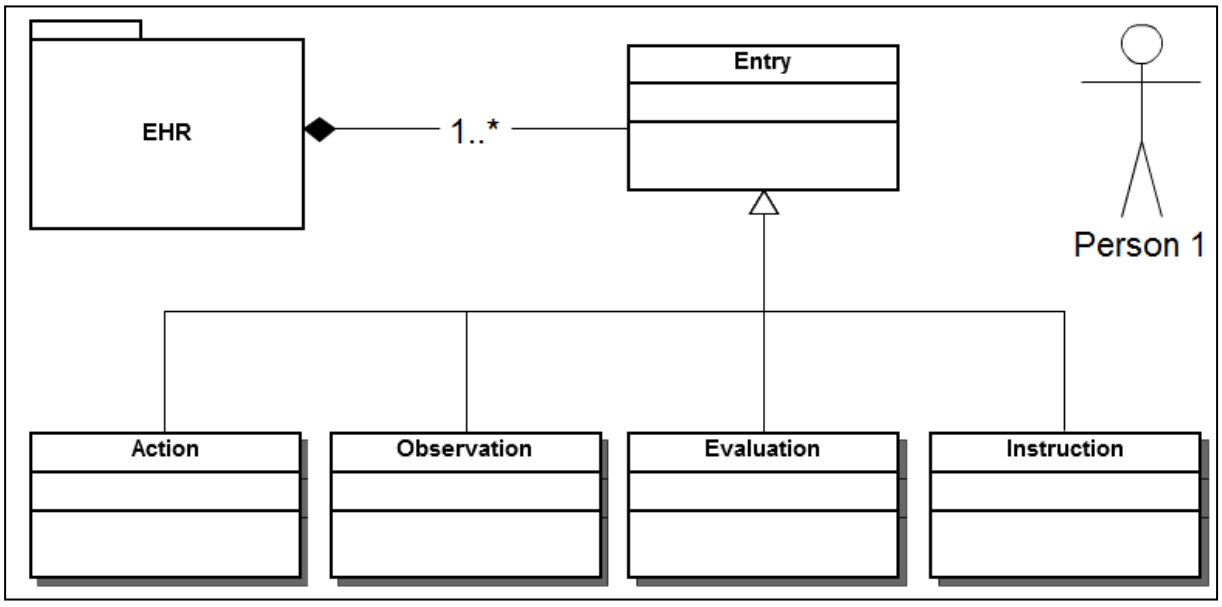


Figure 1: data processing, Source: The author

Every person receiving healthcare and being therefore recorded within the EHR is identified with at least one unique number. The identification number (ID) is an example for an

24 Security Policy: see www.openEHR.org
25 Beale et al.: EHR Information Model, p.19, 2008

identifier in order to distinguish one person from others. The ID builds a bridge between the information divided into knowledge and information to be able to use the recorded data in a person-related way.

An inheritance relationship is drawn for four child classes (Action, Observation, Evaluation, Instruction) and the super class called Entry.[26] The Entry is the main semantic class of the RM. The subclasses inherit functions from this main class. Furthermore, Entry has a compositional relationship with the EHR which is actually a set of Entries. An EHR package instance will always have at least one Entry class instance.

Moreover, archetypes are based on those subclasses. Consequently, every archetype is an Entry which is part of the EHR.

2.5 Structure

The multi-level approach, or to be more precise, the two-level approach consists of a stable and a variable layer.

Even if the development of RM involves a great deal of effort, it is worth the work because of an easy implementation, a more simple maintenance and benefits in terms of safety comparing to one level approaches. Figure 2 illustrates this approach. The information on the RM is stored separately from the knowledge on the AM. Whereas the RM constitutes the stable part, AM is variable in its content[27].

Making up the first level, the stable RM is implemented in software and by that independent from mutable content. Coded terms are an example for invariable data. This stable part is based on technical concerns. Basically, the RM is a standardized data representation which describes the health record itself.

The RM consists of various Information Models (IMs)[28]:

- **EHR** IM: defines context semantics
- **Demographics** IM: specification of demographic service

26 Reference Model: EHR IM, p. 19, 2008
27 4.1 Ontological Separation: http://www.openehr.org/releases/1.0.2/
28 http://www.openehr.org/releases/1.0.2/

- **Common** IM: refers to concepts and design patterns used in higher level openEHR models
- **Data Structures** IM: provides generic data structure
- **Data Types** IM: defines a set of data types
- **Support** IM: defines types, identifiers and specification used in the rest of the specifications
- **Integration** IM: supports expression of legacy data

Building the second level, the AM includes archetypes and templates which act as a gateway to terminologies and guidelines. Thereby, the AM offers those models which are important to describe the semantics of archetypes and templates. The Archetype Definition Language (ADL) is included here.[29]

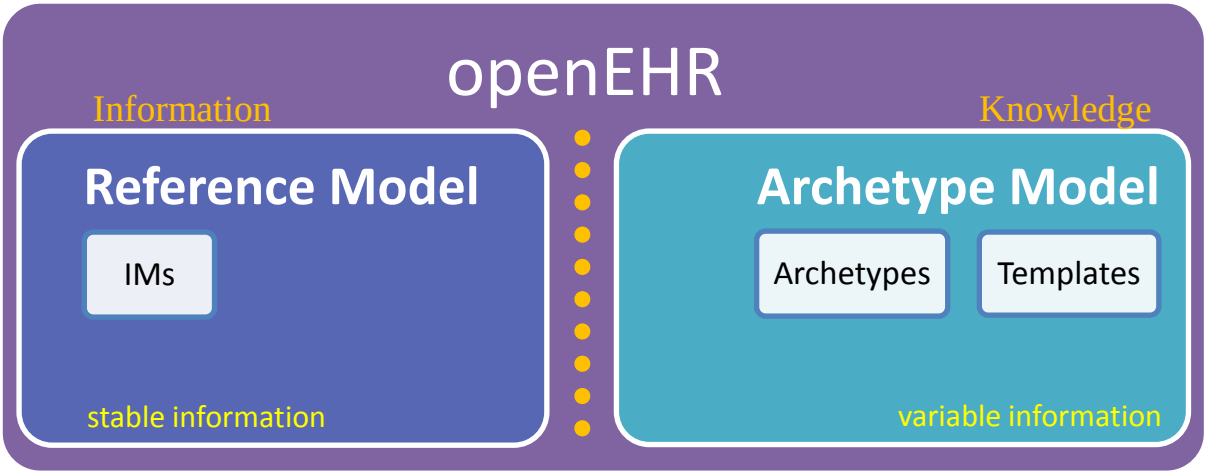


Figure 2: openEHR approach, Source: The author

The AM is very different to the RM in the quality of its data so that the maintenance and development should be handled separately. Doing that enables the reusability of the same RM in different hospitals based on an EHR system.

2.5.1 Reference Model

Ensuring data interoperability, the RM is the basis of the openEHR approach. The RM is the most generic data model that can be developed in the openEHR framework. It contains highly generic knowledge and business rules. The most basic terms required by other packages like

29 openEHR Primer: http://www.openehr.org/sharedresources/ openehr_primer.html

the IMs listed below are described by it. Its optional kept data can be set as mandatory on the second level when archetypes are created based on the restrictions given by the RM.

The RM used by the openEHR Foundation is embodied with special information consisting of a number of coded attributes. This embodiment leads to a very generic architecture satisfying many requirements mainly resulting in functional interoperability. Due to that, the same reference architecture can be used for several care systems. Those ones can strongly differ in the type of semantic data they record. Hence, the RM can be used for public as much as for clinical information without any modification of its architecture. That leads to a facilitated implementation and to the required flexibility. The RM cannot be overridden by archetypes.

2.6 Archetypes

Archetypes are part of the second layer, which is known as the knowledge building AM. Archetypes are flexible specifications of healthcare information parts. They can be expressed in a computable way using ADL. The manner of EHR communication based on archetype methodology has been developed by openEHR and the European and international standard defining organizations CEN/ISO. Furthermore, archetypes are instances of the AM itself. The RM imposes restrictions upon the AM. By that, every archetype is written with respect to the RM. Here, subsets of instances being conform to its own subject are defined e.g. “Pregnancy”. On this second level the optional kept data can be set as mandatory. By the definition of attributes and cardinality the knowledge gets specialized.

To sum it up, the second level allows an individual requirements meeting specialization. Those parts of the RM which are not a subject to change will be taken over entirely by the system anyhow. The instances of the AM are principally based on a semantic class of the RM. These instances are: Observation, Evaluation, Instruction and Action. An Observation class can be a record of things which are measurable or observable[30]. The Evaluation class that follows can be an assessment, an opinion, a goal or something similar. The Instruction class can be based on evaluation and is at least used to record things which have to be done until a time in the future. The last instance, Action, records activities which were performed in clinical context.

30 Heard et al.: openEHR Archetypes & Templates 101, p.3, 2007

Every archetype is expressed in the same formalism. Significant for an archetype is its reusability. The recycle ability speeds up the work process significantly. If an already modeled archetype is too general or lacks detail, it can be specialized. Any specialization will always be matched by queries based on the parent archetype and in relation the specifications given in the RM. This relationship between archetypes follows the principle of subsumption, which is nothing more than a classified summarizing of those. The other relationship possible between archetypes is composition. Composition means the chaining of archetypes.

An archetype connects information structures to formal terminologies. Archetypes are language neutral. It is possible to express their information in any language so that international demands towards the variety of communication can be satisfied. The formal structure of an archetype includes various concepts that one can be plugged into the ADL. Among these main concepts is "archetype", wherein the name of archetype is stated. Furthermore, "concept" is an abstract idea of the archetype and "language" states all languages the archetype is written with. Moreover, languages are defined in standards like ISO for example. Furthermore, "definition" builds the main part of the archetype through enabling clusters, elements and items with the syntax cADL. Besides this, "ontology" is a section wherein data items defined in "definition" are listed and shortly described.

A closer look on the syntaxes dADL and cADL is given in paragraph "2.8 ADL". Moreover, it is possible to create a new version of an archetype. This is only necessary if the data of the present archetype is incompatible with the previous version[31]. This is the case, when a mandatory item of the previous version is regarded as improper. The method applied retains the original archetype unchanged and creates a new version containing the changes. To avoid the creation of a new archetype instead of a new version it is important to keep the identifier the same:

Pregnancy.v1

Pregnancy.v2

In addition, the new version should relate to the previous one in content matters. Versioning not only helps to keep the EHR current but also future-proof.

31 Beale et al.: The openEHR Archetype System, p.15 , 2007

2.6.1 Design principles

The foundation presents 14 principles to design an archetype. Contributing to a better comprehension the following strive to what the reader should know at this point.

- An archetype defines a model of content.
- The constraints defined by it are based on a RM.
- The complexity of archetypes correspondents equally to the complexity of the concepts given in the information model.
- Archetypes correspond to the same ontology level business concepts belong to
- They can be composed and specialized (see **2.7** Templates)
- They are internally hierarchical structured
- They are completely translatable

An archetype can be shared, reused, revised and specialized. Moreover, it can be versioned. All of these actions can be taken without conflict with the RM[32].

2.6.2 Purpose

Archetypes are general-purpose. The main intention is to enable experts modeling concepts. Referring to the healthcare domain, archetypes map clinical knowledge. Every concept of clinical matters is represented by each archetype by a constraint of the RM. It is expected that, clinical experts develop this constraints with their knowledge.

Furthermore archetypes are used for specialized searching in the sense of comparing the data to specialized archetypes. Similarly, studies can be undertaken in a facilitated way by the comparison of data included in archetypes. Hence, archetypes support human interaction they are also developed for computational purposes like the efficient querying of data based on the structure of archetypes[33]. Although they can be used directly the archetypes are generally encapsulated by templates for computational purposes. It means in effect the archetypes are applied to data by using templates.

32 Heard et al.: openEHR Archetypes & Templates 101, p.3, 2007
33 http://www.openehr.org/releases/1.0/architecture/am/archetype_principles.pdf

Archetypes, which are usually authored and transmitted in the ADL format, can be parsed to other formats, for example XML.

2.6.3 Clinical Knowledge Manager

Clinical Knowledge Manager (CKM) is an online platform[34] where clinical data can be shared in order to publish it after a collaborative development. This process consists of reviews by heterogeneous teams including specialists with different backgrounds. Besides archetypes, further resources like templates and metadata relating to clinical models are included here. There are two related platforms available: openEHR CKM[35] and NEHTA CKM[36]. Both repositories are powered by Ocean Informatics. OpenEHR CKM represents the first version of such a collaborative environment. Thus, it contains data developed many years ago. Latest archetypes are starting to come through NEHTA CKM. This collaboration makes an effort to engage national Australian healthcare. It concentrates on latest design but offers only 30% of archetype resources accessible on openEHR CKM.

Both CKM´s provide a noticeable similar user interface. A search tool facilitates screening of resources. Content can be displayed in different ways. A mind map, tabbed view, ADL and XML are among these. Moreover, reports are provided on both CKM´s stating actual statistics.

2.7 Templates

A template can be considered as a composition of archetypes. Their function is to narrow the choices of the archetypes for specific purposes so that only the needed information is displayed on the screen.

A template constrains archetypes down to make them more practical and therefore usable for the clinicians. That includes the ability to make optional data points mandatory and to bind data points to terminology subsets appropriate for a given setting. Moreover, templates cannot overrule archetypes.

34 http://www.openehr.org/knowledge/
35 www.openEHR.org/knowledge
36 www.dcm.nehta.org.au/ckm

2.7.1 Design principles

The following three design guidelines describe the idea of a template more precisely[37]. The first principle characterizes a template as an aggregation of archetypes. Secondly, templates can only inherit constraints given in an archetype. Consequently, templates do not introduce new semantics to archetypes and therefore they do not change their structure. The only changes that can be made on templates are "assigning default values, addition of mandatory items and specifying subsets for usage"[38]. In practice, archetypes are composed on templates to screen forms.

At least, due to the fact that templates inherit attributes from the archetypes, the record of template identifiers is not necessary. Furthermore, templates define the user language which can be Portuguese or English for example.

2.7.2 Purpose

Templates aim at chaining archetypes. The terms preset by the RM and therefore also in archetypes can only be modified in a limited way. These changes apply to the reduction and restriction in optionality or erasing of intended structures defined in the referenced archetypes.

Likewise, templates constitute a basis for constraining data to local needs just as validating data from external sources relevant to the particular use of archetypes.

2.8 ADL

The Archetype Definition Language is a formalism used to express the semantics of archetypes.

It can be applied to write archetypes for any domain where very generic information models are used to describe data in a system[39]. Archetypes are used at this point to constrain the structures of those generic items to represent the particular concepts. By this way future-proved information systems can be also guaranteed.

ADL uses mainly three Syntaxes:

37 Beale et al.: Archetype Definitions and Principles, p.9, 2007
38 Heard et al.: openEHR Archetypes & Templates 101, p.5, 2002
39 Beale et al.: Archetype Definition Language, p.13, 2008

1. data definition form of ADL (dADL): used for widely-accepted primitive[40] syntax types
2. constraint form of ADL (cADL): describes other more sophisticated[41] constraint syntax types
3. First Order Predicate Logic (FOPL) which is a formal logical system also used in mathematics as well as linguistics and is used to describe constraints on data which are instances of some IMs.

Here, dADL is used as a formalism to express descriptive meta-data. The dADL formalism models the data. Its syntax is developed as a human readable computer processing language. In order to be preferred over common formalisms like XML, it is claimed to be more comprehensive. Regarding the specifications[42] of openEHR this attribute is fulfilled by adhering to object-orientated semantics besides others. Finally, the dADL syntax was created to facilitate the exchange of data. Moreover, the cADL syntax "enables constraints on data defined by object-orientated information models to be expressed in archetypes or other knowledge definition formalisms"[43]. The cADL formalism models constraints upon data, which was modeled with dADL before. Thereby, data represented on the RM becomes restricted.

After all, ADL is parsable so it can be analyzed or separated into more easy processing components. With parsing tools it is also possible to convert ADL into various other formalisms like XML for example.

40 Beale et al.: Archetype Definition Language, p.49, 2008
41 Beale et al.: Archetype Definition Language, p.49, 2008
42 Beale et al.: Archetype Definition Language, p.23-24, 2008
43 Shinji Kikushi et al.: Databases in Networked Information Systems, Japan, p.164,2010

3. Brazil and openEHR

In 2004 the Brazilian Federal Government introduced an agreement to decrease the maternal and newborn mortality. The approach was to improve the pre-natal attendance. Two years later the Ministry of Health published a manual for the local districts. The instructions recommended to the healthcare professionals considered a pioneering treatment of pregnant women. For the first time, a protocol of care was foreseen to document stages of pregnancy.

These protocols are forms relating to gestation and are imposed for any HU dealing with that subject as an obligation. Displayed data shall be represented within an EHR. Dating up the manual way of attendance, openEHR became the focus of attention.

3.1 Healthcare in Colombo

The HU in Colombo is actually called "Jardim das Graças*"* and one of many small governmental run HUs situated this city. The staff is consisting of physicians, nurses, community health agents and several technicians who attend a local area divided into further micro areas. Every patient living in one of these areas is recorded. The services provided are mainly general but focused on family health matters. Actually there are two options of a Brazilian HU to attend their patients:

1. With health agents known as "Agentes comunitários de saúde"
2. Without those agents.

"Jardim das Graças" attends their local patients with those health agents. The main function of these agents is to build a bridge between the HU and the patient. This job does neither qualify for medical practice nor nursing care. On the part of the municipality the agent registers notable diseases and births, monitors pregnant women, promotes breastfeeding and ensures compliance of vaccination schedule besides others.

3.1.1 Data used in the health unit

Before a pregnant woman gets attended by medical staff, it is necessary to register her. Therefore, there exists a sheet provided by the Brazilian Ministry of Health, which is in the following referred to as Form 1 and includes general information about a pregnant patient. She will be stated as Healthcare Consumer in the diagram. The HU itself, which assumes the

role of the healthcare provider, will be called Service Provider. The registration form demands information from the service provider and the healthcare consumer separately on the same form. The following Figure 3 and 4 describe requested data relating to Form 1 attached in the appendix under chapter 10.2.

Service Provider	Explanation
Name	Name of the establishment, where the registration took place.
Code Number (CNES)	Every HU has an own code number provided by the government called CNES (Cadastro Nacional de Estabelecimento). This Portuguese abbreviation can be translated as "national establishment registration"
Name of the city	e.g.: Colombo
Code of the city (IBGE)	The Brazilian institution for geography and statistics (IBGE) provides a special code for municipalities.
Initials of the Federative Unit	Every of the 26 Brazilian States has its own initials.
Code of the Federative Unit (IBGE)	State or Federal District provided by (IBGE)

Figure 3: Healthcare Provider Form 1, Source: The author

Healthcare Consumer	Explanation
SISPRENATAL-ID	Every gestation gets an Identification Number provided by the governmental project.
PSF area and micro area	Apart gestation, patients get attended by the family healthcare program, which is offering services in the particular area, where the consumer herself lives. The ID of the area she belongs to and the initials of the micro area are demanded.
Name of the pregnant women	Full name
Birthday	dd/mm/yy
Name of the mother	It is important to state the name of the mother of the consumer as a reference person
Color of skin	There exist 5 options, which can be chosen to classify the ethnicity of the patient: white, black, yellow, brown and indigenous
Address	Complete address of the pregnant woman

	consisting of: number, street, neighborhood, city, zip code and telephone number.
SUS number	Rarely asked. The Brazilian National Public Healthcare System provides an identification number for every registered person.
CPF number	An identification number provided by the Brazilian government to every Brazilian citizen.
Certificate	Details about the certificate quoting the name of the person (birth or marriage certificate).
ID	Identification number, general registration number of every person living in Brazil.
Employment Record Card number	Every legal worker has a card recording the occupation, length of service, pay and any accidents at work.
Date of first consultation	Important for the units finance
Date of last menstruation	Due to impossibility of a correct reconstruction of conception date, this date is used to give evidence about the stage of pregnancy
CBO	Code classifying the profession. A doctor has for example another code than a nurse.
Signature and stamp	The signature of one of seven qualified professionals and the stamp of the unit.

Figure 4: Healthcare Consumer Form 1, Source: The author

In this case data written in grey was not directly associated to demographic data. “Date of first consultation” and ”CBO” are actually demographic data in a clinical context. Thus, it will not be presented in that context. “Menstruation” seemed to be of strict clinical character. The last information in grey serves bureaucratic regulation.

In comparison to other countries, the public health system in Brazil demands more information from the patient. The health consumer has to know the code of the area she is living in, her CPF number, SUS number, ID as well as her employment card number.

The employment card number informs the healthcare provider about the type, time and payment of the patient’s occupation. The ID is nothing new at this point. Otherwise, the SUS number combines health record to each health consumer, so that any provider is always informed about the specific healthcare history of a patient. Moreover, the CPF is a unique number provided to at least every tax payer in Brazil by the government of the Federative Republic. Finally, the PSF helps the medical staff to localize the health consumer and associate her to the responsible HU.

All these data together are recorded in Form 1 and contributes mainly to the processing for demographic information. After an identification of the pregnant women follows a gradual attendance classified by the number of consultations. This classification results out of the mode of payment, which is taking frequency of attendance into consideration. Here, the government is considering the number of consultation. These consultations are based on clinical data pointed out below.

Another blank, Form 2, is also distributed by the Ministry of Health and contains further demographic information but mainly clinical data. It follows a division into three main parts. Comprised demographic information relates to age, educational background and civil state besides others. The following parts deal with clinical information. One section refers to personal and family related anamnesis. Moreover, details about prior pregnancies are asked in the same part. Finally an examination follows. This medical examination is divided into two parts. One demands one-time tests and the other consists of further tests, which are repeated frequently.

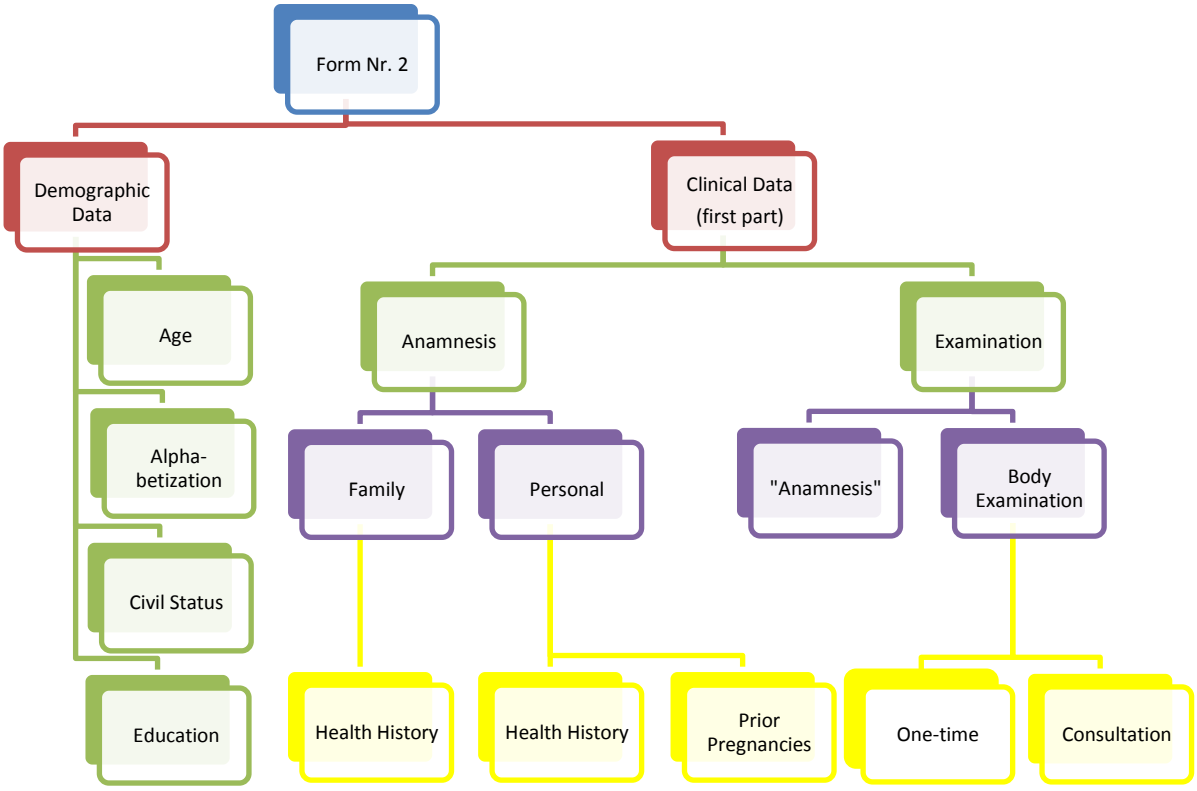


Figure 5: Constraint of Form 2, Source: The author

The second part of Form 2 relates to the actual birth and possible complications. Finally, part three demands data about the newborn and the post natal situation of its mother. These last two sections will not be included in the practical work of this thesis. It is not demanded to represent post-natal data now. Therefore, only pre-natal information referring to the first part of Form 2 is explained in detail. In Figure 6 data is filtered according to the aspects mentioned beneath.

First of all, it was necessary to distinguish demographic data from clinical information. In the section, wherein demographic information is requested, the focus is on four different matters. Among them are age defined in years, the question if the patient is literate or not and the civil status as an extended indicator whether the health consumer is married. This marital status offers here four options: married, widowed, single or divorced. Furthermore, the health consumer has to state whether she has a fundamental, intermediate, higher education level or none of those levels mentioned here.

With respect to clinical information a division is done into anamnesis and examination. While anamnesis is data gained during questioning the health consumer, examination is a way to get data by the means of physical investigation. One anamnesis refers to the family history in order to calculate risks. The other anamnesis relates to the health consumer herself. Her personal health history is asked as well as facts about prior pregnancies.

Clinical data also demands information about examinations. Here, data about examinations done in the past and examinations performed in situ is demanded. “Anamnesis” refers to examinations like previous laboratory tests and hospitalizations during pregnancy. “Body examination” can be divided into one-time examinations, which are done only once, and examinations performed periodically. It is necessary to quote number and date of consultation.

3.2.2 Integration of demanded data

The two forms together are an important part of the Brazilian HU attendance concerning pregnant patients in HUs. Figure 6 refers to the original flow chart published in a final work at the UFPR. Based on information given there by Sampaio de Almeida [44] the following diagram

44 Sampaio de Almeida: Sistema de Informação para a atenção pre-natal no municipio de Colombo, Brazil, 2010

was created and shows the already described data to a lucid arrangement. Moreover, it shows how decisions are made and in which way attendance is structured.

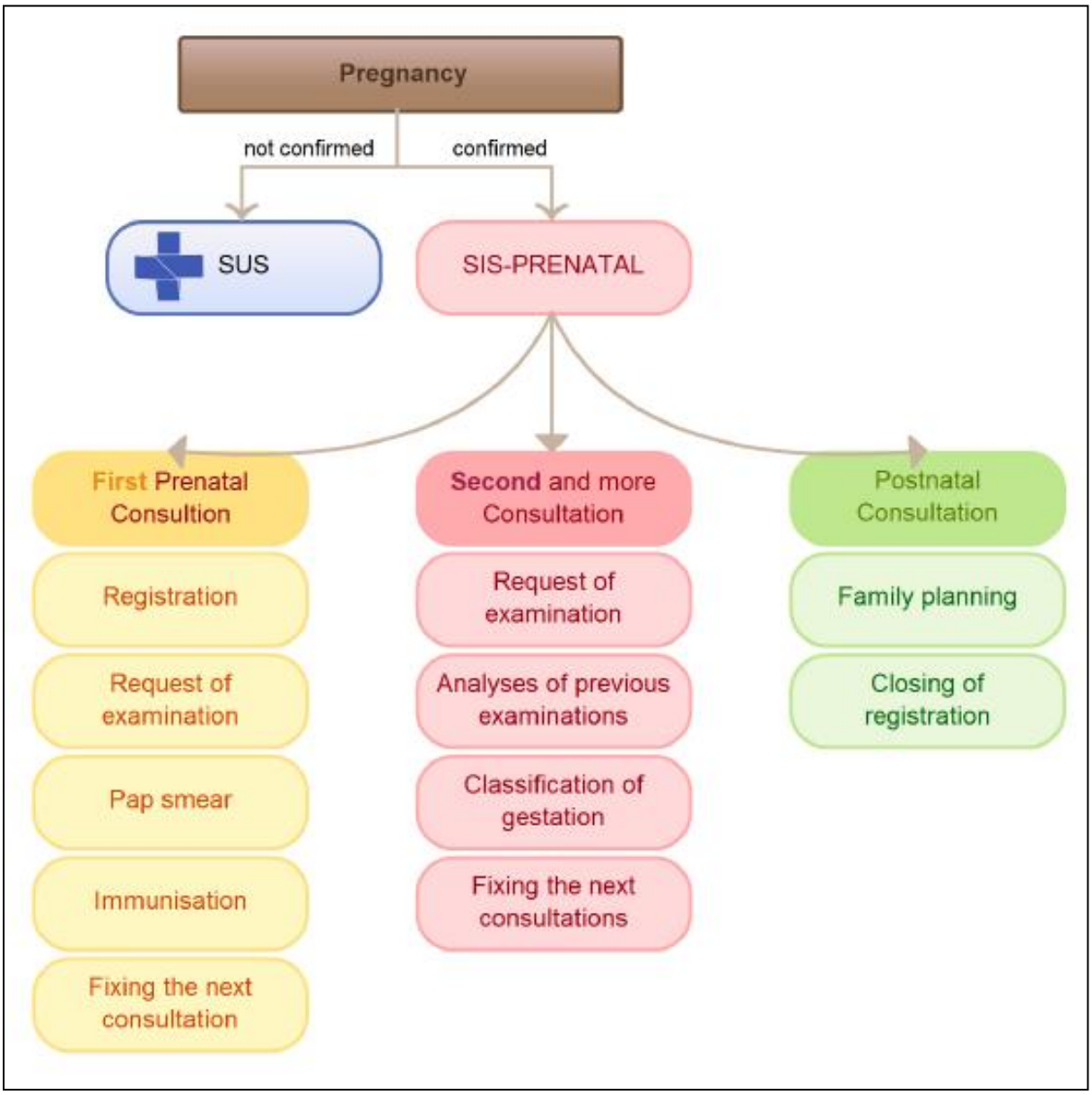


Figure 6: Attendance Tree, Source: The author (restriction of the flow chart of Sampaio de Almeida)

In order to attend the patient with Form1 and 2, it is necessary to confirm the gestation first. If the health consumer is not pregnant she will be attended as usual with the services provided through the SUS program, which is open to every Brazilian habitant. In the case of pregnancy the expectant mother will get special care through the SIS-PRENATAL program. The idea is to provide special support by informing about the correct behavior during the different phases of pregnancy, preparing birth as well as possible and enclosing the care with consolations after birth.

The attendance can be classified into three distinct stages which can be in turn subdivided into sub classes. These classes represent an excerpt of actions done in relation to the progress of pregnancy. The first column includes the registration explained in detail before. Moreover, previous examinations are requested and others are done. The first consultation is the basis for the next ones.

Besides the request of examination within following consultations, the previous examinations will be analyzed and the gestation classified. The classification is based on weight, blood pressure and uterus size besides others. After these steps the next consultation will be fixed once again. The following consultations will have the same procedure till the child is born.

After the birth, the clinicians talk about the post-natal phase. The consultation in between this phase, being the last one, foresees the construction of a family plan. This plan is considering the actual family constellation as well as eventual plans about having more children in the future. Finally the consultation will be formally closed.

Furthermore, the handwritten form containing the health consumer's information is sent to the Municipal Secretary of Health, which transmits it to the State Secretary for Health up to the Ministry for Health. This treatment consisting of several phases is actually foreseen to reduce infant mortality in Brazil.

3.3 From data collection towards representation

The idea is to have a frequently used system implemented with openEHR as a proof of concept. It is not implemented yet for any HU. Therefore, the solution to that problem is the application of openEHR with the use of HU related data. Since this approach shows widely acceptance among technicians and even governments like Denmark, it is a reasonable task to investigate the ability of representation by applying this approach.

Therefore, HU related data are studied, relevant information is taken out and divided into two different types of data sets. By categorizing the data according to their content the first steps are done. Furthermore, regarding the design principles of archetypes, which describe the reuse of already existing archetypes also like versioning, these data representations will be done with respect to the first level of the approach, the RM. Doing this, ensures conformity with the openEHR proceeding. In order to create a system adapted to the needs of openEHR, the used data has to be pre-processed. This pre-procession merges into demographic archetypes

representing the information necessary to register a person and into clinical archetypes relating to medical history and examinations.

Moreover, the restricted data will be visualized in order to work with it in a clinical environment. Thus, the composition of the partly edited and reused archetypes into a template will be done. All these steps will be analyzed and evaluated hereafter.

4. Data representation

Before starting to create an own archetype a structure should be considered, which is showing data in a clear manner. This pre-structure helps to gain an overview and to be conform to the design principles proposed by openEHR.

Because information on forms is often subject to an order which is not primary showing the relations between these, a reorganization of data is required. This includes the extraction of required data from the documents, a subsequent classification and rearrangement. Doing this also improves the recognition of data already represented in archetypes accessible on the CKM. Many of these archetypes relate to general data demanded on documents and thus can be reused within the context.

The strategy of reuse helps the archetype to be aligned with the principles of openEHR so that already existing ones can be used wherever possible. The development of a new archetype should only be considered if the other two options are not possible. If necessary, the archetype will be modified.

The CKM provides archetypes relating mainly to three types of status: draft, team review and published. Here, "status" refers to the state of progress. A "draft" represents an archetype newly created. Furthermore, "team review" refers to an archetype, which was already reviewed by a team. If the teams, consisting of editors and reviewer, are of the view that the archetype satisfies the needs, the status is set as published. Furthermore, there are two more status: obsolete and rejected. Once an archetype is published and does not fit to the needs anymore it can be set as obsolete. Moreover, a not published archetype can be rejected when its content is represented better by another archetype for example.[45]

However, in order to create an archetype it is advisable to download a visual tool also known as an archetype editor. The first archetypes should be generated with such an editing tool because of the assistance they offer to build archetypes with user-friendly interfaces, which generate automatically the ADL code. An experienced developer can use a simple editor to create an archetype. Notepad++ is a free source text editor, which is used too.

[45] CKM functional overview: http://www.openehr.org/wiki/display/healthmod/CKM+functional+overview

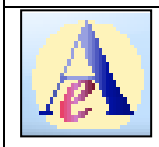
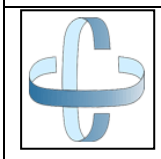
Logo/Name/Version/Release/ Operating System	General Information	Positive Aspects (+)	Negative Aspects (-)
 Ocean Informatics Archetype Editor 2.2.775 Beta 18.04.2011 Windows	This editor was created by Sam Heard and financially supported by the Government of Australia.	+User friendly interface through symbols +direct access to CKM content without opening homepage	-Help documentation does not refer to actual version -only openEHR RM
 OpenEHR Archetype Editor 2.2.775 Beta 18.04.2011 Windows	This editor is the same as the Ocean Informatics editor with the exception that it has another logo.	See above.	See above
 LinkEHR-Ed Archetype Editor RC3 (Windows) 05.04.2011 Windows, Linux	Main component of a project supported by the Spanish Education Ministry, that ended in 2007 Applicable for more than one RM	+Helpful documentations available +offers more than one RM	-Deficits in usability
 Linköping Archetype Editor LiU v0.5.2 20.03.2007 Windows	Initially developed as part of a Master´s thesis at the Department of Biomedical Engineering at the Linköpings University in Sweden		-poor documentation -Links on Homepage are inoperable

Figure 7: Archetype Editor Chart, Source: The author

The author used Microsoft Windows as the operating system to use the openEHR tools in order to create archetypes and also to compose them to templates. It is important to know that some of the tools offered are not compatible with other systems than Windows. Furthermore, the installation of the Java Development Kit (JDK) was required on the operating system, because at least free editors are implemented in Java. Out of the few archetype editors available for free, all four of them are presented briefly below:

At the date of writing this thesis, the openEHR foundation provided the Archetype Editor 2.2.775[46] which was released on the 18. April 2011. It is free, open source and easy to use. The tool was developed by volunteers of the Ocean Informatics project team under the leader ship of Sam Heard. It shows benefits in the use because of its interface using symbols. Furthermore, the technicians are helpful in case of difficulties using the editor. The tool is updated frequently. The Archetype Editor is available for Windows only. The Ocean Informatics and openEHR Archetype editors are the same. The only thing that differentiates them is the logo. Both were financially supported by the Commonwealth Government of Australia.

Furthermore, the LinkEHR-Ed Archetype Editor[47]is another open source and free of charge tool which is available in the World Wide Web. The editor was created by the IBIME Group. This Biomedical Informatics Group is supported by the Spanish government. There are manuals available helping the user to create his first archetypes. Practical instructions can be found, inter alia on the video-sharing website YouTube[48]. The user benefits from the good email based support through the developers and helpful manuals facilitating the independent working process. It enables a slightly different processing of archetypes than the openEHR editor. There are no symbols to drag-and-drop but subunits to select. The user can switch between a hierarchical tree view and the ADL workbench. The demographic archetype was created with LinkEHR-Ed. Moreover, it is possible to edit archetypes based on different RM´s. Depending on the organization, one can choose between five RM´s. It is not possible to load archetypes from the CKM automatically like the openEHR tool it offers. LinkEHR-Ed supports the openEHR RM also like the CEN EN13606 RM which is offered on the interface and defined by the European Committee for Standardization as a European EHR Standard.

46 download: http://www.openehr.org/download/software.html
47 download: http://pangea.upv.es/linkehr/?page_id=10
48 http://www.youtube.com/watch?v=P3BWI4LYEMo

ADL files can be parsed to the machine-readable XML formalism. By that the archetype can be represented in the internet and shared on an interoperable level and even visualized in a different tool. This editor is available for Windows as well as for Linux. It gets released yearly and the latest version is accessible since in April 2011.

The Linköping Archetype editor[49] LiU v0.5.2 is a Java based-tool. It was initially a master´s thesis work at the Linköping University in Sweden, but later performed in the framework of the EU-funded Network of Excellence. Its user interface consists of box-elements and is in its navigation less intuitive in comparison to those tools mentioned above. Besides ADL the user can write in XML. A website[50] with the download package offers links to several files like a list of known bugs. The problem is that especially the link to this list is not working. This tool supports only the openEHR RM. There are no releases accessible besides the one from 2007. Altogether, it is not as powerful as the other tools available for free.

Actually, ADL gets easily translated into XML by each tool. All in all, LiU Archetype Editor is due to its poor documentation and its interface sorted out first. The Ocean Informatics or openEHR Archetype editor respectively is in comparison with the LinkEHR-Ed editor because of its power of great interest. The author decided to use the LinkEHR-Ed editor first in order to develop a demographic archetype because it is documented well and therefore easy to start with. The Ocean Informatics editor was chosen for the creation of clinical archetypes because it seemed to be adequate for the development of clinical data due to its archetype repository including several types of archetypes needed to represent various requested HU specific information.

Some hints drawn up by the clinician and co-author of the openEHR specifications, Sam Heard, to develop an archetype were used. Heard contributed to the openEHR society to a great extent. The available experience and advices of this person were used to create archetypes. The following instructions[51] were applied here:

1. Reuse of archetypes wherever possible
2. Representing a single concept with a single archetype
3. Keeping archetypes small

49 Download:http://www.imt.liu.se/mi/ehr
50 http://www.imt.liu.se/mi/ehr/tools/
51 Heard et al.: openEHR Archetypes & Templates 101, p.8-9, 2007

The reuse of archetypes insofar an important point as the approach of the openEHR Foundation foresees reusing archetypes to facilitate the interoperability by keeping the archetypes to a manageable number avoiding congruency. The second point intends to keep an archetype concrete by using just one concept for each archetype. The last part intends to keep the archetypes easy to handle at a manageable level. This is a way to limit the complexity of archetypes.

During the process of this work frequent changes were noticed. Versions were up dated, new tools were accessible like the template designer for example and data sets were renewed. In order to be "up to date" this thesis refers to the latest state of the art. Therefore, the development process was subject to several revisions in order to satisfy the current needs.

4.1 Representing demographic data

Demographic archetypes are based on the openEHR Demographic Information Model. Figure 8 represents the structure of the demographic RM. This RM consists of various classes and their attributes. A demographic archetype can either describe a person, group, organization or agent. These sub classes inherit from actor, which is a class and can have one or more roles. A role can be in this case either related to the behavior of the healthcare provider or the consumer. In order to describe the demographic information of a person it is not necessary to define a role and its capabilities.

Moreover, actor inherits from party. The party class can have a relation to another class called contacts which has at least one address. Furthermore, party can have one or more party identities. Such a party identity is a classification that enables a differentiation between the utilization of names for one of the sub classes. Finally, a party can have relationships, which describe the relation to a second party or more. The data types here are named with the class prefix "DV_", and inherit from the Data_Value class.

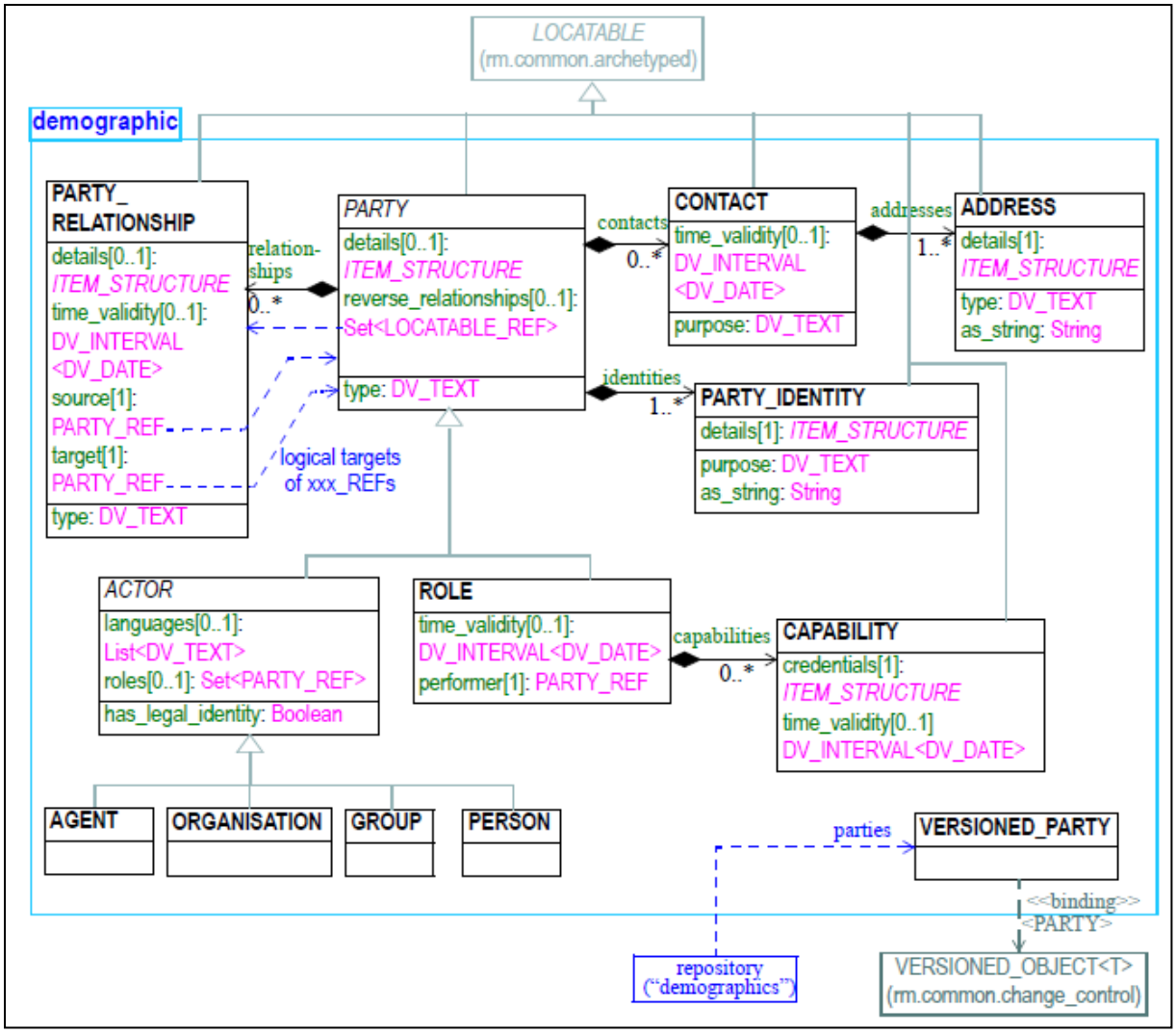


Figure 8: Demographic Package, Source: openEHR

Person´s demographic data are representable with this package. First of all, the forms used in the HU for the record of demographic data were studied. The demographic information was gathered separately from the clinical. Demographic archetypes available on CKM were taken into consideration. Archetypes available on this platform contain contacts of the people who created them, so that contacting the developers is possible in all cases. In cooperation with researchers from UFRJ, a demographic archetype was modified.

Due to redundancy the author will not use age and alphabetization from Form 2. It is assumed that age is information directly relating to birthday and alphabetization can be inferred from education. Thus, only civil status and education background are demographic data taken from Form 2. Besides education background, the other data is already represented in already existing archetypes.

Adhering to the guiding principle of reusing archetypes, the author scanned on the CKM existing ones to find contents fulfilling the specific demands. Many options exist to represent required data. Demographic archetypes were filtered by the subject "organization". In this process the author selected an archetype, which had the status of a draft and the ID openEHR-DEMOGRAPHIC-ORGANISATION.**organisation**.v1. This archetype is able to represent required demographic data about the healthcare provider. Below its mind map view is shown:

Figure 9: Organization archetype, Source: openEHR

Figure 9 shows clearly arranged how data is organized in the archetype structure. Details, contacts and identities refer to slots. The details section for example refers to a further archetype used to identify the provider. This archetype referred to is openEHR-DEMOGRAPHIC-CLUSTER.**provider_identifier**.v1 and includes besides others an identifier designation wherein a number or code assigned to the provider can be stated in order to identify that provider. Changing the occurrences from 1…1 (only one occurrence) to 1…* (one to infinite occurrences) allows storing more than one identity codes. Here, the CNES code can be represented.

The contact node refers to address and includes two further archetypes representing the physical and electronic address of the organization. The slot is actually not correctly specified on the CKM. Instead of relating to two electronic addresses, one should be relation to a physical address. Exactly in openEHR-DEMOGRAPHIC-ADDRESS.**address-provider**.v1 data about the city name should be stored. Furthermore, the IBGE codes are represented here. "Identities" as a further node is used to identify the organization by its names. That node refers to the archetype openEHR-DEMOGRAPHIC-PARTY_IDENTITY.**organisation_name**.v1. Here the name of the unit can be stored. Changing the occurrences from 1…1 to 1…* allows to store more than one name and further name related data like the initials of the HU demanded in Form 1. The relationship node is representing relations to other organizations and departments. "Description" refers only to a descriptive text about the archetype.

So far, this archetype, which is relating to other archetypes, represents the demanded data. The provider´s name is represented within the identities section as well as initials of the federative unit. The name of the city will be shown in the contacts node. Furthermore, numbers like the ID, which are used to identify HUs on a national level, are represented within the details node.

Up to this point, it is shown that data about the organization can be stored using one archetype. But when it comes to the healthcare provider, a second archetype is necessary. Therefore it is essential to be able to state the function of the organization, which is described before. Here, the archetype openEHR-DEMOGRAPHIC-ROLE.**healthcare_provider_organisation**.v1 stores the PSF area and micro area demanded from the healthcare consumer.

Figure 10: healthcare-provider-organization archetype, Source: openEHR

Due to the redundant content of this archetype, the city name is represented in the details node, where it refers to the archetype openEHR-DEMOGRAPHIC-PARTY_IDENTITY.**organisation_name**.v1, which is containing a slot with further archetypes. One of those identifier archetypes included, is called openEHR-DEMOGRAPHIC-CLUSTER.**identifier_other_details**.v1 and contains an element referring to the name of the city. The representation of data demanded from the provider requires the following archetypes:

- openEHR-DEMOGRAPHIC-PARTY_IDENTITY.**organisation_name**.v1
- openEHR-DEMOGRAPHIC-ADDRESS.**address-provider**.v1
- openEHR-DEMOGRAPHIC-CLUSTER.**provider_identifier**.v1

An alternative to the method above would be the creation of a specialization using the openEHR-DEMOGRAPHIC-ORGANISATION.**organisation**.v1. To do this, downloading the archetype is necessary first. Afterwards the specialization can be done by choosing “specialization” within the editor and naming it. By calling the specialization “healthcare_provider” the name of the specialized archetype occurs then like: openEHR-

DEMOGRAPHIC-ORGANISATION.**healthcare_provider-organisation**.v1. The original content will be taken over and can be edited.

The patient´s demographic data was split into four main parts: names, numbers, backgrounds and address. “Names” refers to the own name and the mother´s name. “Numbers” represents an amount. “PSF” stands here for the code of the area and its micro area the pregnant woman is classified through the family health program called as PSF “Backgrounds” relates to marital status, ethnic group and education background. Moreover, the address is further demographic data demanded. The author describes in the following how a demographic archetype is created referring to other archetypes also including an already existing archetype which was specialized before.

Figure 11 presents the interface of the LinkEHR-Ed tool wherein the demographic archetype is created. This tool was selected because it enables the creation of archetypes based on the demographic RM. First of all the organization has to be chosen. It can be selected between CEN and openEHR. The author chose openEHR because due to the task, this work refers to the international openEHR and not to the locally restricted CEN organization.

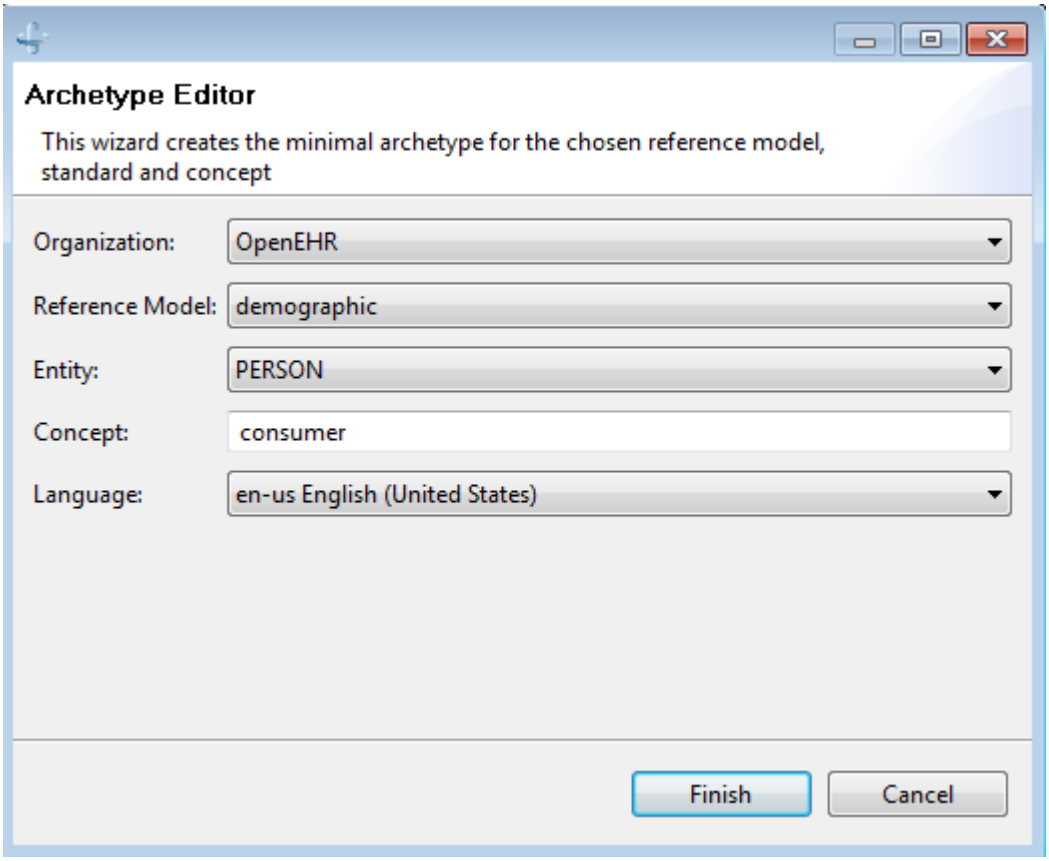


Figure 11: LinkEHR Ed , Source: The author

The second step requires a choice between two RMs. Besides the demographic RM the EHR RM can be selected. It is not possible to create with the EHR RM a demographic archetype so that the demographic RM needs to be chosen. There are different archetype classes available according to their supposed use: Address, Agent, Capability, Cluster, Contact, Element, Group, Item List, Item Single, Item Table, Item Tree, Organization, Party Identity, Party Relationship, Person and Role. They define the matter of data represented in the archetype .Because of representing demographic data of pregnant women, the entity "Person" was chosen. The concept was chosen according to the healthcare consumer. At least one idiom has to be selected. Afterwards a translation can be done, so that the archetype can be read in various languages. The author chose English due to its status as a commonly used language for electronic communication.

The developed demographic archetype OpenEHR-demographic-PERSON.**consumer**.v1 shows Figure 12 in the composition mode:

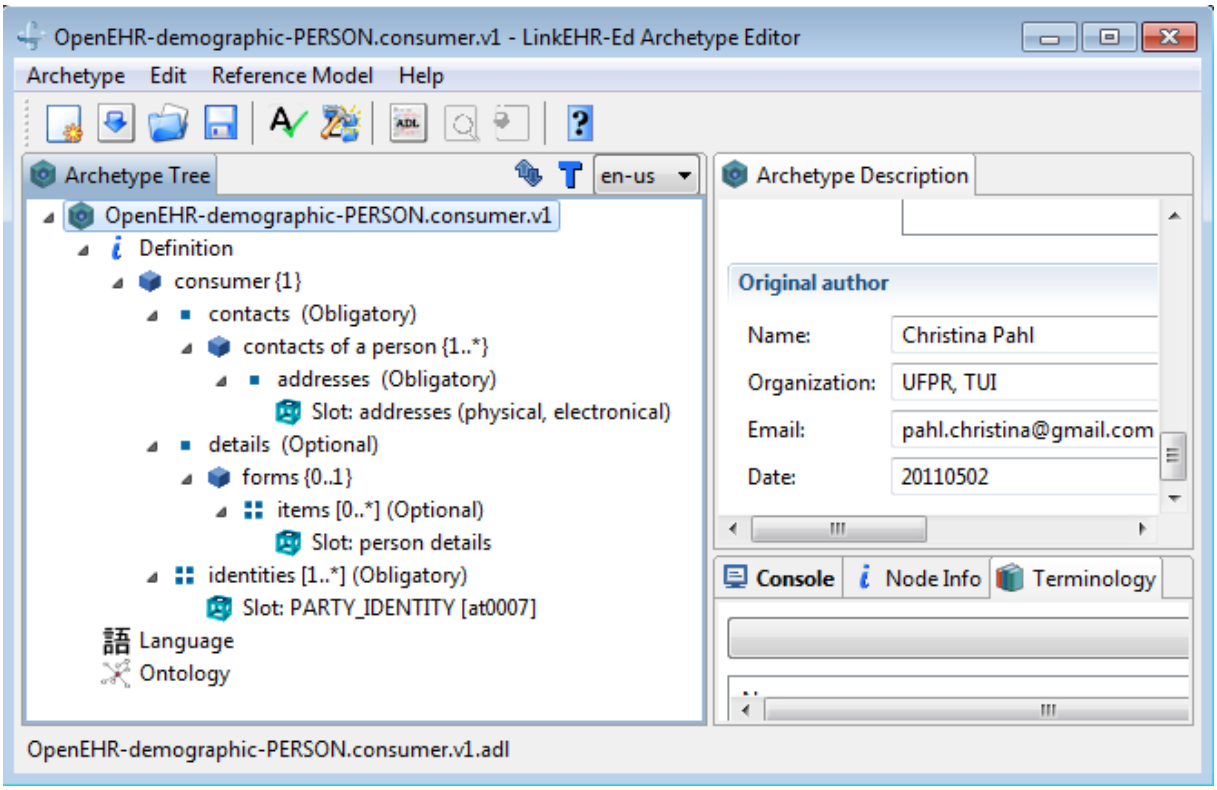


Figure 12: Archetype development with LinkEHR Ed, Source: The author

The patient demographic data is splitted into three main parts: contacts, details and identities. Moreover, these parts are subdivided into slots. A slot is a chaining point of at least one archetype. Slots are special nodes where archetypes can be added.

“Contact” for example is subdivided into a slot wherein two archetypes are located. These archetypes describe addresses. The author decided to use electronic and physical addresses as patient contacts. By doing so, not only the demanded physical address is represented but also an alternative like e-mail address or the required telephone number besides others.

The openEHR-DEMOGRAPHIC-ADDRESS.**electronic_communication**.v1 archetype refers to a telephone, mobile, fax, pager, e-mail, URL or other medium of communication. Furthermore, the patient has to add the type of use of the contact. It is necessary to give the information if the contact is either for personal or for business use. Every medium is also combined with operating hours. The patient is able to decide if she likes to attend in between special time period like in the morning, during working time or only at night for example.

The second archetype called openEHR-DEMOGRAPHIC-ADDRESS.**address**.v1 includes the required information of a typical Brazilian home address. First of all the address can be set as unknown in the case it is inadequately described. Otherwise, the following information about the address will be directly assigned to the patient: street suffix code in order to qualify the street name in directional references, street name, street number and street type code in order to identify the type of public thoroughfare. Furthermore, the country identifier representing the country component of an address, the town, the postal code, a delivery point identifier which is a unique number assigned to a postal address as designated by the postal service, the census area known as an area defined by the Brazilian Geographic and Statistics Bureau according to the special distribution of the population, and the district. The attributes of the address archetype represent a geographic location which can be used as postal address, residential address or in other means. Address related demands like the IBGE codes are satisfied here, because each detail about a Brazilian address can be stored.

In order to store information about the area the person is attended through the PSF program, the archetypes had to be specialized to openEHR-EHR-CLUSTER.**Health_Unit_Area-address**.v1. Herein, two elements were added representing data about area and micro area. This information could not be stored on the original archetype.

The second main section “Details” contains all documentations except for names and identities. This part includes a reference to the openEHR-DEMOGRAPHIC-ITEM_TREE.**person_details**.v1 archetype. Because of its content which is relating perfectly to the only element missing, this archetype was extended by one data more, the education background. Therefore, the above-mentioned archetype was downloaded and specialized as

openEHR-DEMOGRAPHIC-ITEM_TREE.**education-person_details**.v1. The education background was added within this specialization and then put into the slot relating to "Details".

In order to specialize the archetype, it is necessary to localize a section wherein the data fits best. Each of the four different clusters representing demographic data within the archetype was taken into consideration. Here, a cluster refers to a " grouping class within trees that allow subtrees to be specified"[52] The first cluster contains elements referring to the birth of the person. The items are inter alia about the date and place of birth. The second cluster deals with the death of a person. The third one refers to the person additional data and the last cluster to biometric information. Out of the four clusters included in the archetype, the author chose the third one.

"Education Status" is additional and supplements the already included data like marital status. The DV_Coded_Text class is used to state the different levels of education. Terms directly referring to that rubric are subdivided into codes relating to the same mother element. Here, each code stands for a different level.

The archetype itself is originally written in Portuguese but translated into English. Therefore, the specialization was edited in both languages, too. Doing so preserves consistency. Furthermore, only editing the element in the definition section is not a sufficient change. Figure 13 shows the composition view of the specialized archetype as well as the ADL code referring to the parts in front of the arrows. The edited element has to be explained in the ontology part of the archetype, too.

52 file:///C:/Program%20Files%20%28x86%29/Ocean%20Informatics/Template%20Designer/Help/TemplateDesignerRepository.html

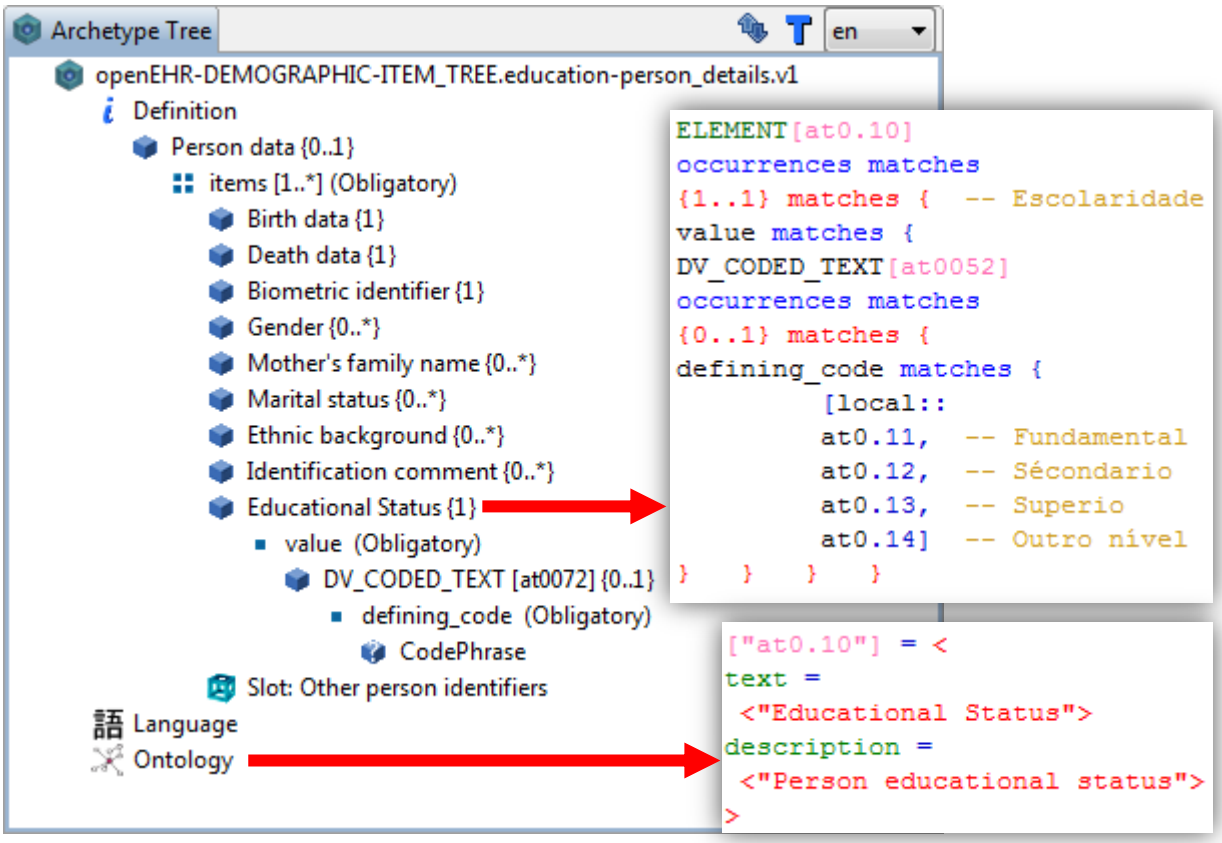


Figure 13: Person details specialization, Source: The author

In order to preserve the form, information in the ontology was added in English as well as Portuguese. Afterwards, this specified archetype is added into the slot of the OpenEHR-demographic-PERSON.**consumer**.v1 archetype. This archetype refers to civil status, mother´s name, ethnic background, education background and date of birth. Furthermore, this archetype contains a slot referring to an archetype called openEHR-DEMOGRAPHIC-CLUSTER.**person_identifier**.v1. This archetype includes an identifier element named "identifier main data" wherein all identity codes and numbers can be stored. To these numbers belong CPF, ID, SIS-PRENATAL, SUS and the employment card number.

The identities section of the OpenEHR-demographic-PERSON.**consumer**.v1 includes the name of the healthcare consumer. A slot was created automatically named PARTY_IDENTITY and the archetype openEHR-EHR-CLUSTER.**person_name**.v1 was added as reference. The author has specialized this archetype. This specialized archetype openEHR-DEMOGRAPHIC-PARTY_IDENTITY.**Certificate-person_name**.v1 represents within two further elements of data about a certificate quoting the name of the person. Regarding the RM PARTY_IDENTITY can only include names. In order to prevent misunderstandings the author quotes at this point the RM wherein the intention of this class is described in detail.

"Instances of PARTY_IDENTITY, linked to PARTY by the attribute identities are intended to express the names of people, organisations, and other actors - that is names which are "owned" by the party, e.g. self-declared (in the case of institutions and companies) or by virtue of social relations (names given by parents, tribes etc). Identifiers of Parties given by other organisations, or the state are not represented in this way, and should be recorded in the PARTY.details structure instead"[53]

From this quote, it is obvious that PARTY_IDENTITY has nothing in common with further particulars establishing the identity of a person. In this case these are particulars such as the identification number of a person or a fingerprint as a biometrical parameter besides others. The intention of this class is merely the definition of a name. To sum it up, the representation of healthcare consumer data requires the following archetypes:

- openEHR-DEMOGRAPHIC-ADDRESS.**electronic_communication**.v1
- openEHR-EHR-CLUSTER.**Health_Unit_Area-address**.v1
- openEHR-DEMOGRAPHIC-ADDRESS.**address**.v1
- openEHR-DEMOGRAPHIC-ITEM_TREE.**education-person_details**.v1
- openEHR-DEMOGRAPHIC-CLUSTER.**person_identifier**.v1
- openEHR-DEMOGRAPHIC-PARTY_IDENTITY.**Certificate-person_name**.v1

All these archetypes are included in the OpenEHR-demographic-PERSON.**consumer**.v1 which composes them to one archetype. In the following Figure 14 and 15 resumes all information demanded and location of its representation within the created archetype:

Demographic Provider Data	Archetype	Details
Name	openEHR-DEMOGRAPHIC-PARTY_IDENTITY.**organisation_name**.v1	ELEMENT[at0010] Name
Code Number	openEHR-DEMOGRAPHIC-CLUSTER.**provider_identifier**.v1	ELEMENT[at0001] Identifier designation
Name of the city	openEHR-DEMOGRAPHIC-ADDRESS.**address-provider**.v1	ELEMENT[at0007] Suburb/town/locality
Code of the city (IBGE)	openEHR-DEMOGRAPHIC-ADDRESS.**address-provider**.v1	ELEMENT[at0004] Census area
Initials of the	openEHR-DEMOGRAPHIC-	ELEMENT[at0024]

53 Reference Model: Demographic Information Model, p.10

Federative Unit	PARTY_IDENTITY.**organisation_name**.v1	Abbreviated name
Code of the Federative Unit (IBGE)	openEHR-DEMOGRAPHIC-ADDRESS.**address-provider**.v1	ELEMENT[at0008] State/territory/province identifier

Figure 14: Demographic Provider Data, Source: The author

Demographic Consumer Data	Archetype	Details
SISPRENATAL-ID	openEHR-DEMOGRAPHIC-CLUSTER.**person_identifier**.v1	ELEMENT[at0001] Identifier main data
PSF Area and micro area	openEHR-EHR-CLUSTER.**Health_Unit_Area-address**.v1	ELEMENT[at0.2] PSF Area ELEMENT[at0.6] PSF Micro Area
Name of the pregnant women	openEHR-DEMOGRAPHIC-PARTY_IDENTITY.**Certificate-person_name**.v1	CLUSTER[at0002] Structured name
Birthday	openEHR-DEMOGRAPHIC-ITEM_TREE.**education-person_details**.v1	ELEMENT[at0010]Birth data
Name of the mother	openEHR-DEMOGRAPHIC-ITEM_TREE.**education-person_details**.v1	ELEMENT[at0032]Mother's family name
Color of skin	openEHR-DEMOGRAPHIC-ITEM_TREE.**education-person_details**.v1	ELEMENT[at0034] Ethnic background
Education Status	openEHR-DEMOGRAPHIC-ITEM_TREE.**education-person_details**.v1	ELEMENT[at0.10] Education Status ELEMENT[at0.11] Fundamental ELEMENT[at0.12] Intermediate ELEMENT[at0.13] Higher ELEMENT[at0.14] Other Level
Physical Address	openEHR-DEMOGRAPHIC-ADDRESS.**address**.v1	ELEMENT[at0028] Street name ELEMENT[at0029] Street number ELEMENT[at0003] District, ELEMENT[at0006] Postal code [at0007] Suburb/town/ locality
Electronic Address	openEHR-DEMOGRAPHIC-ADDRESS.**electronic_communication**.v1	ELEMENTS [at0020] Telephone [at0007] Details
SUS number	openEHR-DEMOGRAPHIC-CLUSTER.**person_identifier**.v1	ELEMENT[at0001] Identifier main data
CPF number	openEHR-DEMOGRAPHIC-CLUSTER.**person_identifier**.v1	ELEMENT[at0001] Identifier main data
Certificate	openEHR-DEMOGRAPHIC-PARTY_IDENTITY.**Certificate-**	ELEMENTS [at0.2] Certificate name [at0.4] Certificate details

	person_name.v1	
ID	openEHR-DEMOGRAPHIC-CLUSTER.**person_identifier**.v1	ELEMENT[at0001] Identifier main data
Employment Record Card number	openEHR-DEMOGRAPHIC-CLUSTER.**person_identifier**.v1	ELEMENT[at0001] Identifier main data

Figure 15: Demographic Consumer Data, Source: The author

4.2 Clinical archetypes

Clinical data is represented different than demographic information. Due to the complexity of information, a preparatory work consisting of separation, structuring and classification of data was necessary.

In order to edit, specialize or build clinical archetypes, the Ocean Archetype Editor was used. This modeling tool is only referring to the openEHR-EHR RM and thus supports only creation of clinical content. Moreover, the editor provides direct access to the CKM. This enables downloading archetypes directly from this platform without using a web browser like it is the case with the LinkEHR-Ed editor.

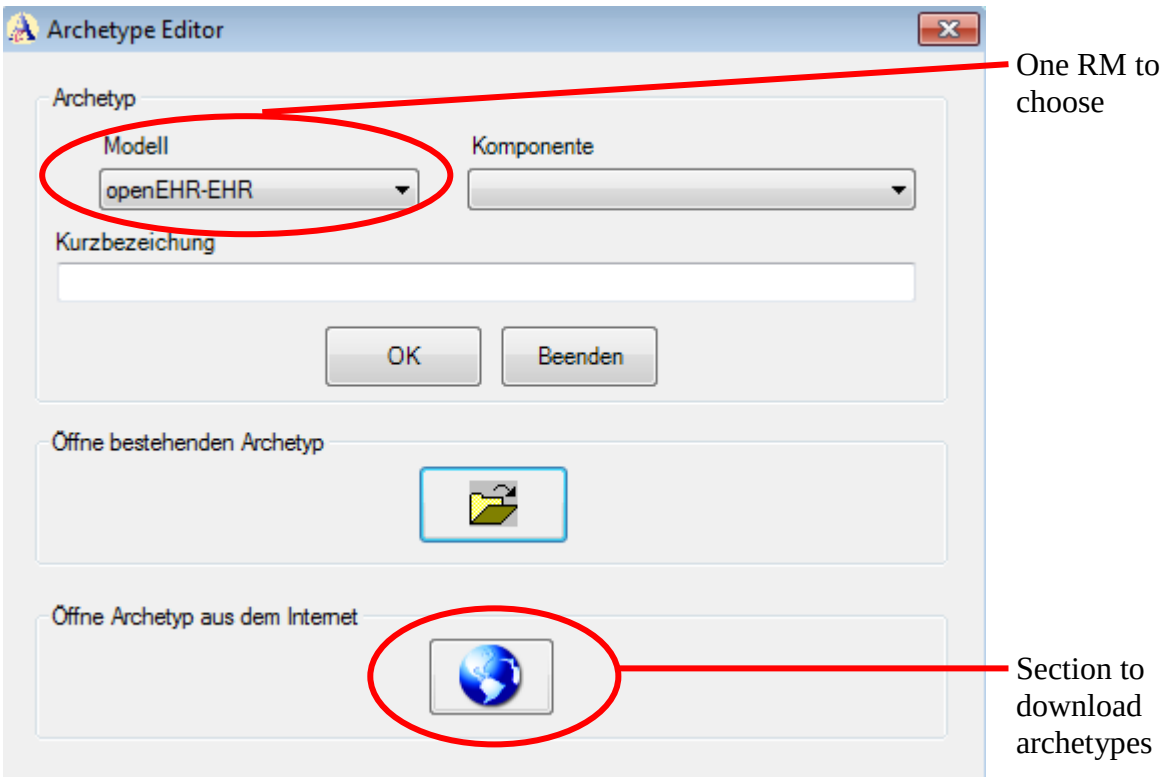


Figure 16: Ocean Informatics, Source: The author

The editor was used on a computer system with settings in German language. Thus, the figure's components are in German. The red ellipses show the sections above-mentioned wherein the RM is set as the only option and another section wherein any archetype can be automatically downloaded from the CKM platform if an internet connection is available.

Before working with the editor each data item was studied. The information´s meaning in the specifically context was worked out as well as connections of all data among themselves. A chart was created referring to the characteristics of information with focus on data type and already existing archetypes they refer to. Afterwards, this selection of archetypes was analyzed more precisely. Archetypes satisfying the demands completely were used. Only a few information demanded was unrepresented on the CKM.

First, an archetype openEHR-EHR-SECTION.**health_unit**.v1 was created using the Ocean Archetypes Editor presented above. "Section" is a type of archetype and is based on the openEHR Section class. This class enables data within a composition to be organized. Hence, the archetype can be seen as a pre-structure. Using a further tool, template editor, archetypes can be put into the slots displayed in red here. These slots all relate to a generic term, here in green, displayed on the later template. The creation of a slot is associated with the definition of the archetypes class. This class can be defined as Observation, Evaluation, Instruction or Action besides others.

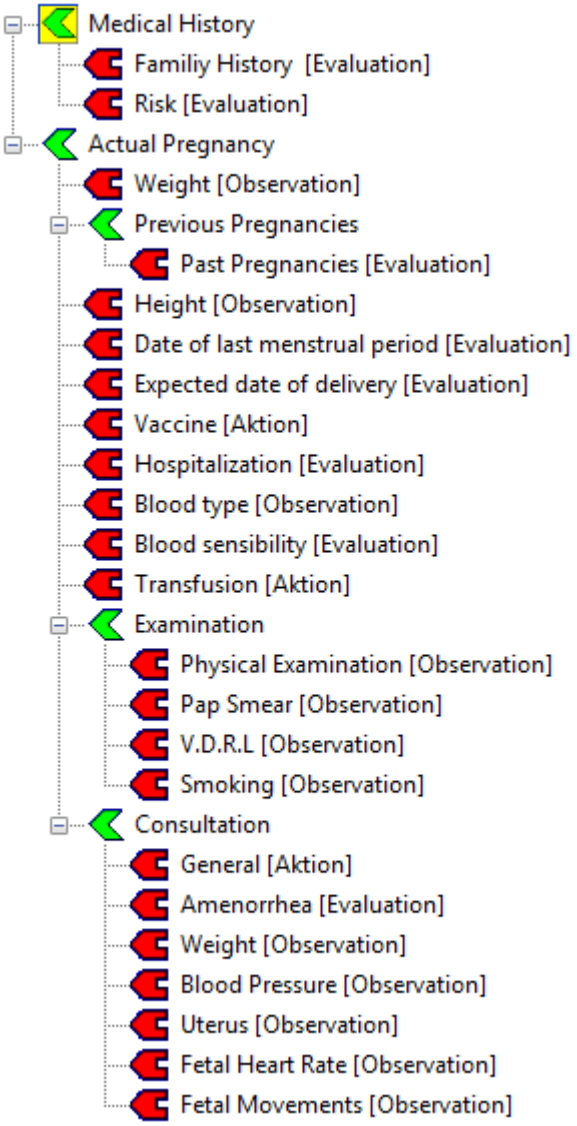


Figure 17: Section archetype, Source, The author

The classification represented in Figure 17 is based on the preparatory work. Data was extracted from the contexts preset by the forms. This data, dealing with prenatal information, were selected and redundancies

summed up. The last step considered a rearrangement of these data displayed above.

On the basis of this structure, the author compared archetypes to the concepts presented above relating to further details on the forms. A complete conformance of demanded data and already created archetypes was the case for a few. Conformity is given for the following archetypes:

- openEHR-EHR-OBSERVATION.**blood_match**.v1
- openEHR-EHR-OBSERVATION.**body_weight**.v1
- openEHR-EHR-OBSERVATION.**fetal_movement**.v1
- openEHR-EHR-OBSERVATION.**heart_rate**.v1
- openEHR-EHR-OBSERVATION.**height**.v1
- openEHR-EHR-CLUSTER.**exam-generic**.v1.
- openEHR-EHR-CLUSTER.**exam-uterine_cervix**.v1
- openEHR-EHR-CLUSTER.**exam-mouth**.v1
- openEHR-EHR-CLUSTER.**exam-uterus**.v1
- openEHR-EHR-OBSERVATION.**blood_pressure**.v1
- openEHR-EHR-CLUSTER.**physical_properties**.v1
- openEHR-EHR-CLUSTER.**individual_professional**.v1
- openEHR-EHR-CLUSTER.**person_name**.v1
- openEHR-EHR-OBSERVATION.**lab_test**.v1
- openEHR-EHR-OBSERVATION.**substance_use-tobacco**.v1
- openEHR-EHR-ACTION.**medication**.v1
- openEHR-EHR-CLUSTER.**individual_professional**.v1

The archetypes displayed in gray are under construction and therefore not confirmed for public use yet. The ones in green are already published and ready for public use. Blood match represents the blood group and the Rhesus factor in the way it is requested on Form 2 under “actual pregnancy”. The body weight of the pregnant patient is asked within two coherences. Once it is recorded as normal body weight before pregnancy, the weight gets frequently recorded in the context of the growth of gestation during consultations. Height is demanded in the same section and fully represented in an archetype already published.

Furthermore, data relating to actual pregnancy demands information about the unborn child. Within the frequently done consultations, fetal heart rate and movement are recorded. Fetal movement is relating to occurrences of movements of the fetus during pregnancy and represented in the archetype openEHR-EHR-OBSERVATION.**fetal_movement**.v1. The rate of the fetal heart is information which is not occurring in an archetype which is directly relating to the fetus. The archetype openEHR-EHR-OBSERVATION.**heart_rate**.v1 relates to

any rate in the context of a heart. Data recorded about fetal heart rate can be stored in this archetype.

The demanded examination results are represented within the observation archetype openEHR-EHR-OBSERVATION.**exam.**v1 including a slot for further clusters. These clusters were chosen in such a way from the ones offered by the CKM, so that they are meeting the demanded type of examination on Form 2 as closely as possible. In order to allow more than one archetype in this slot the amount of occurrences had to be changed in the original observation archetype from minimal 0 to 1 occurrence to maximum infinite occurrences.

The clusters, representing the needs are: openEHR-EHR-CLUSTER.**exam-generic**.v1, openEHR-EHR-CLUSTER.**exam-uterine_cervix**.v1, openEHR-EHR-CLUSTER.**exam-mouth**.v1 and openEHR-EHR-CLUSTER.**exam-uterus**.v1. The first three archetypes are used in the examination part and the last one, referring directly to uterus, is used in the consultation part. The archetype openEHR-EHR-CLUSTER.**physical_properties.**v1 is an archetype which is not directly used but within a slot of openEHR-EHR-CLUSTER.**exam-uterus**.v1 in order to store data about the size of the uterus. In order to state the size a paper[54] was used indicating common parameters. This paper suggests diameter, length and width as possible physical quantities.

Furthermore, within the frequently repeated consultation the blood pressure of the pregnant women has to be recorded. Therefore, it was possible to use the already published archetype openEHR-EHR-OBSERVATION.**blood_pressure**.v1. Herein, in which maximum (systolic) as well as minimum (diastolic) blood pressure can be stored.

Moreover, within a specialized archetype used to store general data about the consultation, the author relates by an added slot to the name of the professional doing the examination. Using the archetype openEHR-EHR-CLUSTER.individual_professional.v1 and openEHR-EHR-CLUSTER.**person_name**.v1 in the context of the openEHR EHR RM and not the demographic RM, the author enabled a further use of two unchanged archetypes. Within the first archetype "Brazilian Classification of Occupations" can be stored. This data is demanded on Form 1. Furthermore, it offers a slot to represent the name of the professional.

54 Esmaelzadeh et al.: Normal uterine size in women of reproductive age in northern Islamic Republic of Iran, 2004

Another archetype openEHR-EHR-OBSERVATION.**lab_test**.v1 was used as a basis for two different data: Papanicolaou test and V.D.R.L. The first test is used in gynecology to detect cancerous processes in the cervix. Moreover, V.D.R.L is an acronym for Venereal Disease Research Laboratory test and stands for a syphilis test. Form 2 demands data about the result of V.D.R.L and date, the result was issued. Both data can be stored without any changes. Furthermore, the test status of the Papanicolaou test is demanded. The coded element "Test status" offers at this point various test statuses, among these "final" and "never performed". Two different tests can be represented using this archetype in its original form.

In order to represent the information whether the pregnant woman is a smoker or not and if so how many cigarettes she is smoking per day, the archetype openEHR-EHR-OBSERVATION.**substance_use-tobacco**.v1 was used. This archetype already represents a specialization and contains all information requested. Here, the information if and how often the patient is smoking tobacco products can be stored using the element "Frequency". Following the demands of Form 2 the items "daily use" and "no use" coded under this element are of interest. Moreover, using the quantity element "Number smoked" data about the amount of cigarettes smoked per time unit can be represented.

Representing vaccination related data required a bottom-up approach. Starting with the very special archetype openEHR-EHR-ITEM_TREE.**medication-vaccine**.v1, of which specialization is described later, the author looked for a referring super-archetype. In the later template, only Cluster and Element archetypes can be inserted into slots within Entry or Cluster archetypes to support the particular use-case[55]. Furthermore, the Item_Tree class can only be used within Instruction and Action classes. The Instruction archetype offered on the CKM lacks compliance with the subject. Here, transfusion is not intended to be an instruction but a record of a tetanus vaccination done in the past of a patient. Thus, an Action archetype matching with this requirement was chosen. The archetype openEHR-EHR-ACTION.medication.v1 was used here. This archetype enables the recording of actions taken with regard to medication. The slot for an Item_Tree archetype included there was specified at design time. It is not possible to create a slot for this class of archetype after the design process. In order to use the specialized archetype openEHR-EHR-ITEM_TREE.**medication-vaccine-Tetanus**.v1 within the Action archetype, it was necessary to embed the specialization first. Using drag-and-drop, this specialization was put into the repository of archetypes the

55 Heather Leslie, 2011

super-archetype can relate to. Figure 18 shows archetypes, which were not modified in content.

Clinical Consumer Data	Archetype	Details
Blood group, Rhesus factor	openEHR-EHR-OBSERVATION.blood_match.v1	ELEMENT[at0004] AB0 ELEMENT[at0005] Rhesus
Body weight	openEHR-EHR-OBSERVATION.body_weight.v1	ELEMENT[at0004] Weight
Occurrence of fetal movements	openEHR-EHR-OBSERVATION.fetal_movement.v1	ELEMENT[at0005] Occurring now
Fetal heart rate	openEHR-EHR-OBSERVATION.heart_rate.v1	ELEMENT[at0004] Rate
Height of the patient	openEHR-EHR-OBSERVATION.height.v1	ELEMENT[at0004] Height/Length
Examination	open EHR-EHR-OBSERVATION.exam.v1	CLUSTER[at0005] allow_archetype
Name of a named body part and its condition	openEHR-EHR-CLUSTER.exam-generic.v1	ELEMENT[at0.9] Name of object ELEMENT[at0005] Normal statement
Condition of cervix	openEHR-EHR-CLUSTER.exam-uterine_cervix.v1	ELEMENT[at0005] Normal statement
Condition of teeth and related structures	openEHR-EHR-CLUSTER.exam-mouth.v1	ELEMENT[at0005] Normal statement
Appearance of uterus	openEHR-EHR-CLUSTER.exam-uterus.v1	ELEMENT[at0002] Clinical description
Maximum and minimum blood pressure	openEHR-EHR-OBSERVATION.blood_pressure.v1	ELEMENT[at0004] Systolic ELEMENT[at0005] Diastolic
Size of uterus	openEHR-EHR-CLUSTER.physical_properties.v1	ELEMENT[at0034] Diameter ELEMENT[at0029] Length ELEMENT[at0031] Breadth
CBO, name	openEHR-EHR-CLUSTER.individual_professional.v1	ELEMENT[at0011] Professional Identifier CLUSTER[at0015] allow_archetype
Name of professional	openEHR-EHR-CLUSTER.person_name.v1	ELEMENT[at0001] Unstructured name
Papanicolaou test and V.D.R.L	openEHR-EHR-OBSERVATION.lab_test.v1	ELEMENT[at0005] Test name ELEMENT[at0078] Result ELEMENT[at0075] Datetime result issued
Smoking	openEHR-EHR-OBSERVATION.substance_use-tobacco.v1	ELEMENT[at0005] Frequency ELEMENT[at0012.1] Number smoked
Vaccination	openEHR-EHR-ACTION.medication.v1	allow_archetype ITEM_TREE matches

Figure 18: Reused archetypes, source: The author

The first column refers to the matter being represented in each archetype presented in the adjacent column. Wherein exactly in the archetype data is represented can be taken out from the last column. This one displays the assignment of information within each archetype so that it can be easily reconstructed and understood relating to the mentioned archetype. All these archetypes were not modified in its content.

The following archetypes are modified. Specialization was used to make a generic parent archetype more specific for two particular purposes: transfusion and general information about the consultation. This is needed in the case that an internationally published archetype needs to be adapted to local variations, what helps to keep the core content consistent with the published archetype. If the archetype is not meeting actual needs anymore and further requirements are identified, modifications have to be made. At this point versioning the particular archetype is the solution. "Within the CKM governance processes, backwardly compatible changes will be released as a revision v1.x, whereas as a non-backwardly compatible change to the published archetype will be released as a new version v2 or v3." (Heather Leslie, **2011**) Those archetypes changed are shown hereafter:

- openEHR-EHR-EVALUATION.**risk-SUS**.v1
- openEHR-EHR-EVALUATION.**risk-family_history-SUS.**v1
- openEHR-EHR-EVALUATION.**pregnancy-SUS**.v1
- openEHR-EHR-ACTION.**procedure-transfusion**.v1
- openEHR-EHR-ACTION.**procedure-SUS**.v1
- openEHR-EHR-ITEM_TREE.**medication-vaccine-Tetanus**.v1

Data referring to medical history is already available on the CKM. Archetypes representing personal medical history and family history only required specialization. It is interesting, that the family history archetype already represents a specialization of the personal risk archetype. Meeting the needs of the Brazilian healthcare program SUS, both archetypes had to be adapted to its particular demands. Due to SUS related modifications, the name openEHR-EHR-EVALUATION.**risk-SUS**.v1 is extended by a freely chosen naming, here SUS, joined by an hyphen. This style is always used to characterize a specialization. That specialization includes changes in form of a differentiation of the text element condition, on which the risk assessment is focused, into a coded text. This DV_Coded_Text refers to the demanded abnormalities: urinary infection, infertility, diabetes, arterial hypertension, pelvis or uterus surgery and others. Moreover, the author set the occurrences of the coded text items as unbounded because more than one abnormality could have occurred. Furthermore, a Boolean element stating the presence of abnormality was added. Similar changes were done in order to

store data in the specialization openEHR-EHR-EVALUATION.**risk-family_history-SUS.**v1. Here, single abnormalities were listed within a DV_Coded_Text. Among these are: diabetes, arterial hypertension, multiple pregnancies and others. Moreover the occurrences were also set as unbounded for the same reason. Here, a Boolean data type is already enabling the representation of the presence of an abnormality.

The specialization openEHR-EHR-EVALUATION.**pregnancy-SUS**.v1 was created. This archetype includes various additions. It is divided into three main clusters: conception, previous pregnancies and current pregnancy. The date of the last menstrual period and a statement about the correctness of this statement is required. Therefore, it was necessary to add a Boolean element to the already existing date element in order to indicate doubts referring to the stated date. This was done in the conception cluster, wherein date of last menstrual period was already included.

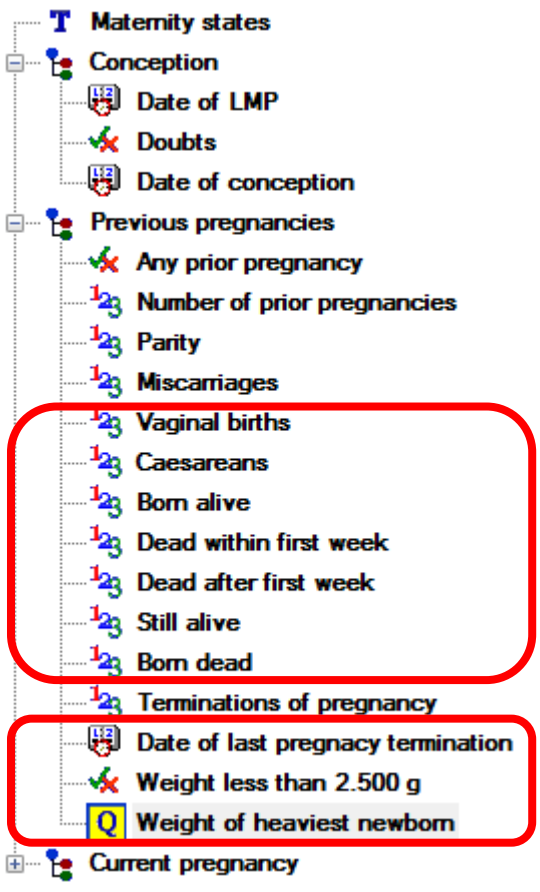


Figure 19: Previous pregnancies, Source: The author

The previous pregnancies cluster of the specialized archetype version contains information high lightened in red boxes, which are supplemented to the original version. The information whether the patient has born more than three times or never, can be taken from the information about the number of prior pregnancies. Thus, the author did not edit these data.

Form 2 also requests the numbers of vaginal births, caesareans, children who were alive during birth as well as newborns, who died during the first week and after the first week. It is demanded to state how many children are still alive and how many were dead during birth. In addition to that, it was necessary to add the date of last termination of pregnancy. This information is demanded on the form and complements the already existing data about the number of terminations.

This cluster was also supplemented in the matter of weight relating to the newborn. In the case that a birth occurred with a child with less than 2.500 g, it should be marked. Moreover, it is also necessary to state the weight of the heaviest newborn. Therefore, a quantity element was created wherein the weight in kg can be entered. The rest of that cluster was kept in the way it is occurring in the first version of the same archetype. This added content is marked with red boxes.

During consultations it is necessary to state the actual time of menstrual period absence in weeks. Amenorrhea can be calculated from the last menstrual period (LMP), which is already represented in the conception Cluster. The system can calculate that information, thus it should not be stored.

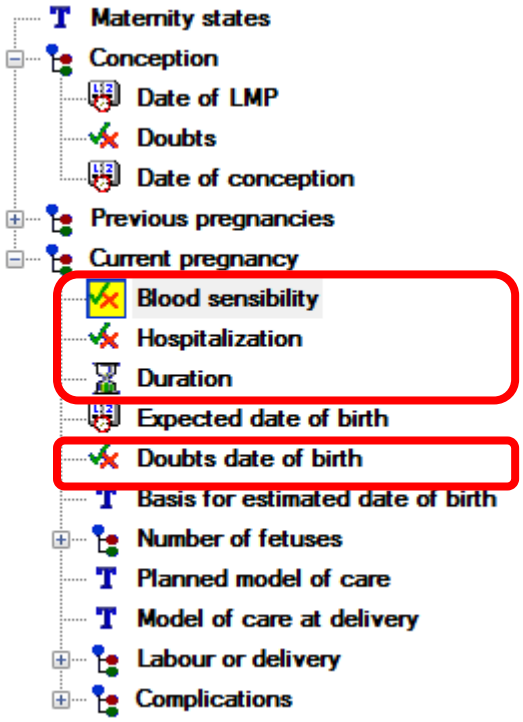


Figure 20: Current pregnancy, Source: The author

Besides this information, details about blood match and Rhesus factor are demanded. Exactly this information is already covered by the archetype openEHR-EHR-OBSERVATION.**blood_match**.v1 above-mentioned. The only blood related information not included, is sensibility. The author assumes at this point, that this data fits better into the context of a pregnancy archetype. It is known that different rhesus factors between a pregnant woman and the fetus can cause sensibility of the mother against the antigens of the unborn child. This sensibility can occur in the case of placenta related micro traumas and cause complications. Here, the specialized pregnancy archetype version includes this sensibility data. Regarding the document, this data needs to be stored as Boolean data. Furthermore, information about any hospitalization during gestation is requested. Thus, an element was added to this cluster consisting of a Boolean data type storing either yes or no statements towards hospital stay. In addition to this, it is necessary to add duration of stay in hospital while pregnancy. Therefore, a data type representing duration in days was complemented in order to complete hospitalization related data in the way it is demanded. This cluster, dealing with the actual pregnancy, contains originally an element, wherein the date of expected birth can be stored. Half of information relating to date of delivery is already

covered by this data type. The other part had to be added. This relates to an element, wherein doubts about the stated date can be indicated. By the use of a Boolean data type, the possibility to state a doubt about this statement was given. Up till now, all changes to the pregnancy specialization were presented.

Transfusion is a subject not occurring on the knowledge manager in the sense of an observation. There are two transfusion archetypes on the openEHR CKM: an instruction and an Action archetype. The first mentioned relates to instructions for giving a transfusion and the last archetype is used to record the actions taken during transfusion. Both archetypes do not meet the demanding requirements of a record done during pregnancy in the past of a person. Thus, storing data about a blood transfusion done during pregnancy required the specialization of the archetype openEHR-EHR-ACTION.**procedure**.v1. This archetype describes a clinical activity that has been carried out for therapeutic or diagnostic purposes. It relates to the seldom used class Action. Herein, the name of the activity "transfusion" was added and two items inside of the protocol part. These items are a DV_Boolean stating whether the action was transformed or not and finally a DV_Text wherein the name of locality where the action was completed can be stored. Moreover, archetypes with an Action class provide a pathway[56] to store activities. Recording transfusion against the "completed" pathway step will allow the date to be captured due to the RM. Here the date does not need to be modeled explicitly. The archetype openEHR-EHR-ACTION.**procedure-transfusion**.v1 represents all data requested about a transfusion done in the past by specifying an Action archetype.

In order to represent general data recorded during consultation for administration issues, the archetype openEHR-EHR-ACTION.**procedure**.v1 was specialized here again. The specialization openEHR-EHR-ACTION.**procedure-SUS**.v1 contains a few changes. The name of process was chosen as "consultation" because it will be used during all consultations which follow. Data items added refer to name of the professional executing the physical examination and to the number of consultation. The name serves to the allocation of response of results recorded during examination. Here, the archetype openEHR-EHR-CLUSTER.**person_name**.v1 was used to store the name of the professional. Moreover, it was not necessary to use any item to track the frequency of examinations received by the pregnant patient before birth. This information can be taken from the instances of the archetype

56 Beale et al.: EHR Information Model, p.62, 2008

consultations, which are stored. The knowledge about that number is necessary in order to calculate the payment for the HU. The path of that specialization was set as "in progress" because the state refers to an action being performed as an entire row of consultation and not only one. It is not necessary to add date. Recording "consultation" against the "in progress" pathway step will allow the date to be captured due to the RM. Thus, the date does not need to be modeled explicitly. This is similar to the previous specialization using the Action class.

With reference to the super-archetype openEHR-EHR-ACTION.**medication**.v1, the specialized archetype openEHR-EHR-ITEM_TREE.**medication-vaccine**.v1 was further specialized. An item, stating the tetanus vaccine, was added within a vaccine element. Moreover, a DV_BOOLEAN is also put into this archetype in order to state whether vaccination did already start. The number of vaccination is covered by the element "Sequence number". Figure 21 shows data demanded on the forms and the modified archetypes they are relating to:

Clinical Consumer Data	**Archetype**	**Details**
Family history	openEHR-EHR-EVALUATION.**risk-family_history-SUS**.v1	ELEMENT[at0.0.34] Diabetes ELEMENT[at0.0.35] Arterial Hypertension ELEMENT[at0.0.36] Multiple pregnancies ELEMENT[at0.0.37] Others ELEMENT[at0.31] Presence
Medical History	openEHR-EHR-EVALUATION.**risk-SUS**.v1	ELEMENT[at0014] Urinary infection ELEMENT[at0015] Infertility ELEMENT[at0016] Diabetes ELEMENT[at0017] Arterial hypertension ELEMENT[at0018] Pelvis/uterus surgery ELEMENT[at0019] Others ELEMENT[at0020] Presence
Pregnancy	openEHR-EHR-EVALUATION.**pregnancy-SUS**.v1	ELEMENT[at0008] Date of LMP ELEMENT[at0.109] Doubts ELEMENT[at0021] Any prior pregnancy ELEMENT[at0.110] More than three pregnancies ELEMENT[at0033] Number of prior pregnancies ELEMENT[at0017] Parity ELEMENT[at0031] Miscarriages ELEMENT[at0.115] Vaginal births ELEMENT[at0.116] Caesareans ELEMENT[at0.117] Born alive ELEMENT[at0.118] Dead within first week ELEMENT[at0.119] Dead after first week ELEMENT[at0.120] Still alive ELEMENT[at0.121] Born dead ELEMENT[at0.122] Date of last pregnancy termination

		ELEMENT[at0.112] Weight less than 2.500 g ELEMENT[at0.123] Weight of heaviest newborn ELEMENT[at0.125] Blood sensibility ELEMENT[at0.126] Hospitalization ELEMENT[at0.127] Duration ELEMENT[at0013] Expected date of birth ELEMENT[at0.128] Doubts date of birth
Transfusion	openEHR-EHR-ACTION.**procedure-transfusion**.v1	ELEMENT[at0002.1] Transfusion ELEMENT[at0.6] Done during gestation ELEMENT[at0.7] Local name
Consultation	openEHR-EHR-ACTION.**procedure-SUS**.v1	ELEMENT[at0002.1] Consultation ELEMENT[at0047] In progress ELEMENT[at0.63] Professional name
Vaccination	openEHR-EHR-ITEM_TREE.**medication-vaccine-Tetanus**.v1	ELEMENT[at0.0.4] Tetanus ELEMENT[at0.0.5] Started ELEMENT[at0022] Sequence number

Figure 21: Specialized archetypes, Source: The author

All information cannot be represented using already existing archetypes in their original form. Thus, six archetypes required a specialization.

4.2.2 Composing Archetypes

In order to use archetypes in a composed way and meaningful structured a further tool is needed. Here, the author used Template Designer version 2.6.1214Beta in order to compose single archetypes to an assorted template. This tool is provided by Ocean Informatics and allows visualizing templates as HTML. This tool enables data entry and thus builds the basis for the electronic data form representing information demanded on the paper forms.

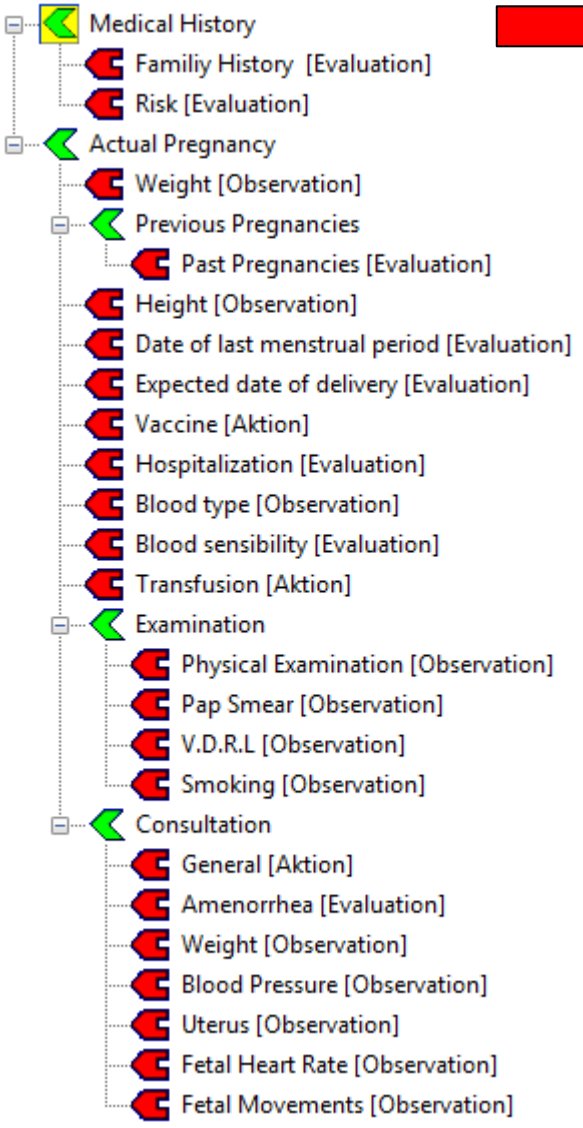
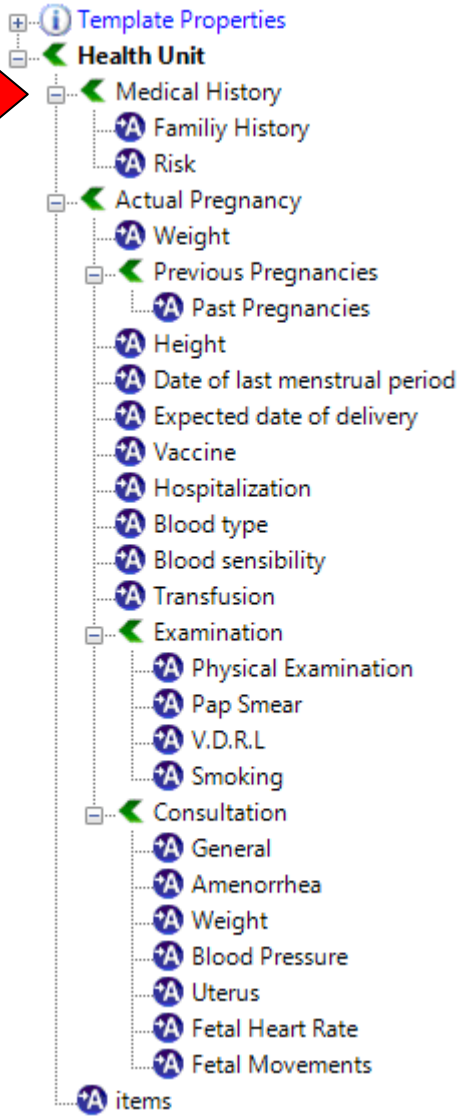
Archetype Editor	Template Designer
	

Figure 22: Archetype editor vs. archetype designer, Source: The author

Here, this archetype, whose development is described in the previous chapter, is shown on the left side of Figure 22. Opening this archetype with the Template Designer offers the view displayed on the right side of the same. Within this template two properties, template name and its purpose, are characterizing each template in order to differentiate it from others. The blue circles with the capital letter “A” indicate slots. These slots can be specialized in their

capacity. Thus, one slot can contain a defined number of archetypes. The next step consists of an allocation of archetypes out of a set. This repository needs to be filled with archetypes which are foreseen to be put into these previously created slots. Through drag and drop of archetypes into associated slots their content is listed.

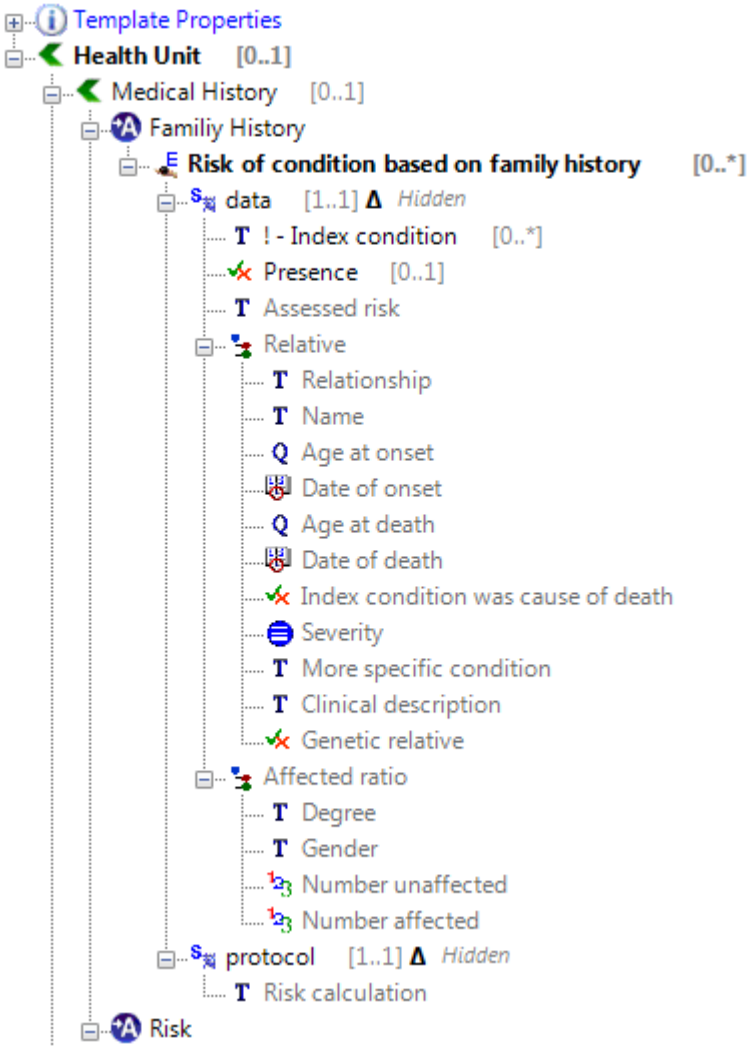


Figure 23: Health Unit-template, Source: The author

Figure 23 shows an example of the first slot filled with the archetype openEHR-EHR-EVALUATION.**risk-family_history**.v1. It is possible to view each element of this archetype and to put attributes on single clusters, elements or even further slots included in this archetype. These adjuncts apply to occurrences besides others. Furthermore, it is possible to set items mandatory or to hide them on the later form by choosing the corresponding setting. The ones with the numbers in the square brackets will appear on the form. Moreover, they give details about the number of occurrences. Besides setting of default values, it is possible to use only features of archetypes which are needed. Choices can get limited as required for the particular use in the HU in the city Colombo.

The symbols in front of each item show to which type of data the element refers to. Here, “T“ means DV_TEXT, tick and cross represent a DV_BOOLEAN, “Q” stands for quantity, and so on. All these data types cannot be changed on the level of template designing. This is only possible by the use of any template editor.

The author filled each slot with at least one archetype presented in the previous chapter. Some slots required more than one archetype due to further slots included within the same. The occurrence of items was selected in order to hide unimportant data. Thereby, a reduced view of data focused only on the ones demanded, appears on the form and other data is faded out. Although information is hidden once, it can get always be integrated on the electronic form again by marking the attribute of occurrence. Here, the view gets restricted to the first part of the whole template which is dealing with medical history. The intention is to show the template looks like when it gets automatically translated into HTML by the Template Designer. The HTML view is one of a few views provided by the designer.

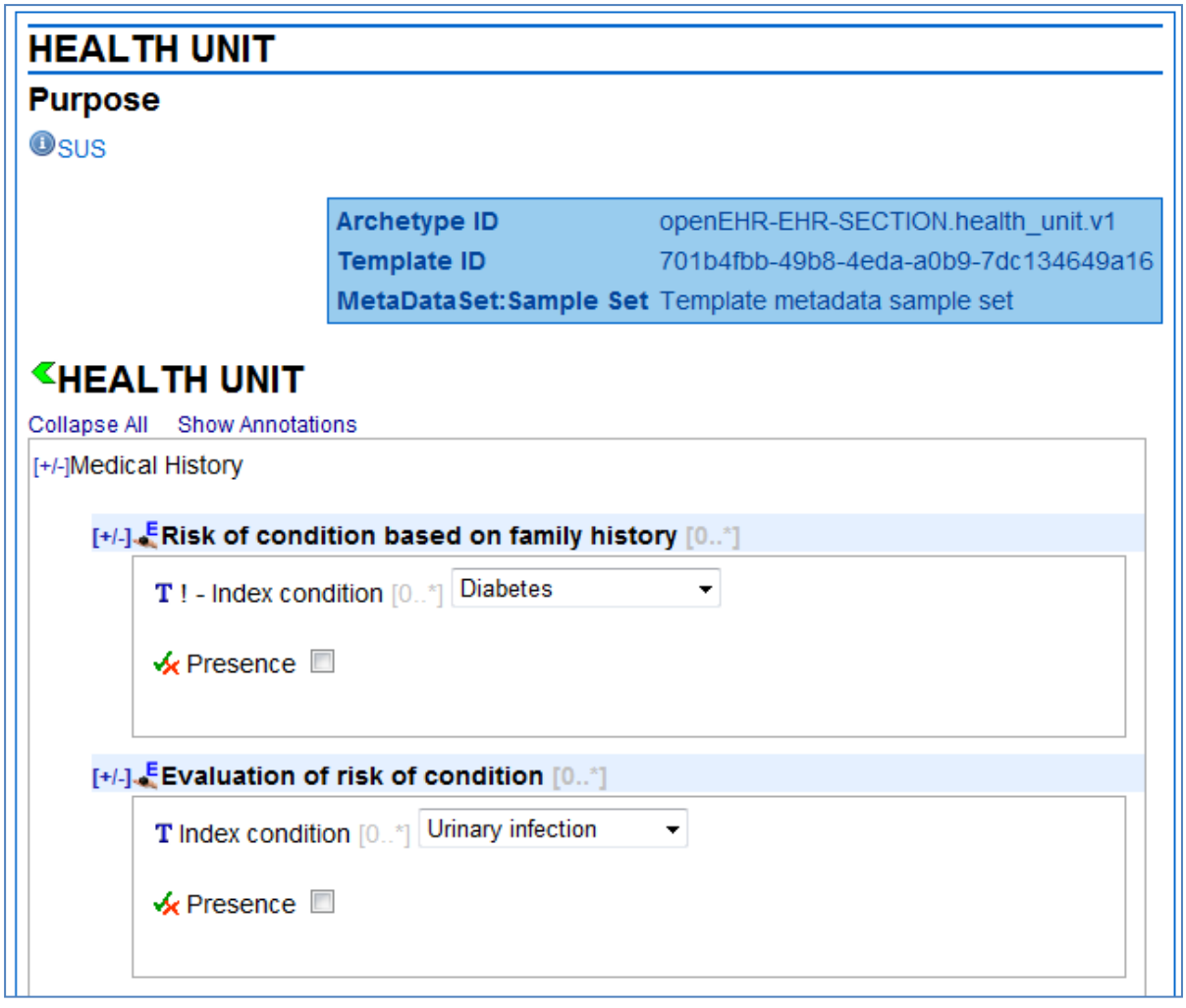


Figure 24: Health Unit-URL, Source: The author

As one can see, only those items marked as “occurring” on the template structure (Figure 24) are shown on the HTML view. An HTML view of the entire developed template is given in the appendix.

4.3 Problems during application

During the process of representing demanded data several problems occurred. Regarding ADL, its dADL syntax is not accepted by all openEHR users and thus rarely used. Many specialists even prefer XML as a format to define archetypes instead of ADL. In contrast to the primary intention of dADL to facilitate data exchange, this use is not popular among experts.

Getting started with the videos provided and other material about LinkEHR-Ed made processing easy. Switching to Ocean Informatics archetype editor with the experience gained before, facilitated the use. The lack of a useful manual explaining the potentialities of Ocean´s editor was a problem. The guide provided is not conform to the latest editor version and also poorly documented. Furthermore, this tool is not able to display demographic archetypes. Here, LinkEHR-Ed is preferable to Ocean´s editor.

First of all, the development process required a detailed knowledge about the single definitions. In order to use the openEHR environment in a correct way, it was necessary to get a deep insight into single specifications. It took a great deal of attention and further consultations with specialists in order to interpret the content in a right way. Because there does not exist a model explaining the details of a clinical archetype similar to the model provided for demographic content, it is not an easy task to create clinical archetypes corresponding to the RM in an adequate way. Therefore, clinical archetypes were created reduced to needed data only and following the example of similar archetypes provided on CKM. By doing so, mistakes can be prevented. Furthermore, only a few people deal with clinical archetypes. There is no domain specialist available to discuss the models, which is one big difficulty.

Regarding the usability of the archetype editor LinkEHR-Ed used for the representation of demographic archetypes, a few difficulties were noticed. The user interface is compared to the Ocean Informatics Archetype editor not nearly as easy to use. LinkEHR-Ed offers a structure of archetypes in form of a tree view, which can be added by the use of elements shown by mouse click on the particular item. As a user does not always comprehend which element needs to be defined by which sub elements offered. The author preferred a direct way to add elements within an archetype through viewing ADL code and adding missing elements in the way similar data was already represented. Some archetypes were overloaded with translations,

so that keeping an overview became challenging besides the fact that added items need to be translated into all languages provided.

Moreover, the LinkEHR-Ed software used was not updated to the version 1.0.2 of the openEHR specifications at this point. Additionally, the tool inserted some archetype nodes in wrong places. Moreover, some domain types like code phrases and quantities could not be represented. In order to fix that problem, it was necessary to work with the ADL workbench and put nodes to the right places or to edit relevant items to particular places within the workbench. By that, information which was not accurate placed could be placed manually. In summary, the Link-EHR-Ed did not contribute sufficiently to a satisfactory level of addition of elements within demographic archetypes.

Furthermore, it is necessary to use a free mind mapping application in order to enable a graphic view out of the tree structure provided by the LinkEHR-Ed editor. There is no renderer implemented in LinkEHR-Ed. A solution therefore, is the free of charge open source tool FreeMind[57] besides others.

The Ocean Informatics archetype editor is in comparison to LinkEHR-Ed much easier in use. Only knowledge about data types was needed in order to create or edit archetypes by drag-and-drop from object column to workbench. Out of the tree view provided the code becomes translated automatically into several different formats like HTML, XML and ADL besides others. Drag-and-drop is actually also possible within LinkEHR-Ed but relates to another operating principle. The selected item can get moved on the nested structure in order to reorder the objects of a multiple attribute for example.

Besides the editor related ease of operation, another issue occurred relating to the state of archetypes. Those not published are subject to changes. Thus, a certain degree of flexibility is required working with archetypes provided on CKM. An example for this case is the person´s details related demographic archetype, which was changed during the process of this work besides other data sets. The fact that the template designer was made accessible during the practical phase of this work by Ocean Informatics, put in the meantime a further point of interest into this work and required flexibility, too. Before, no tool offering similar services

57 Download: http://freemind.sourceforge.net/wiki/index.php/Main_Page

was accessible for free. Working with archetypes, means dealing with data under processing conditions at this point.

Ocean Informatics template designer showed problems in the matter of terminology. It was not possible to change the interface into English after installing the software on a computer with different language settings. The template designer showed difficulties in loading changes. Because there is no real time refresh on the render above the template implemented, the possibility to see the final result automatically does not exist.

Further difficulties occur due to the governance of the used CKM repository which shows slow changes on the platform. Statistics published on the CKM report that from the total number of archetypes, stated at the time of writing as 292, only 9 archetypes are published yet. That means that since the pilot version of CKM dated on 2008 not even 4% of archetypes are elaborated sufficiently enough to be used correctly in EHR contexts. That is among other things caused by bad management of repository and results in a delay of the use of archetypes in EHR´s.

One of those problems caused by human interaction and difficult to manage is the fact, that people with different backgrounds interpret data in different ways. Thus, misinterpretations occur due to designations dependent on people´s background. Uniqueness should be considered of high relevance during contribution to archetype processing.

5. OpenEHR application result

In this scale of analysis, it was shown that the CKM provides enough archetypes to represent an excerpt of particular data demanded by the Health Ministry. Hence, it was not necessary to create new archetypes. The entire process of representing demanded data in form of archetypes using the tools chosen by the author leads to an encouraging achievement.

The aim of this thesis has been to investigate if openEHR is a possible approach in modeling HISs. An incidental aim has been to find out how much effort it requires to represent data with the approach. In order to study this, the author applied that specification with the use of selected free of charge and open source tools.

First of all, an analysis of openEHR theory was necessary. The vast literature on this specification published by openEHR[58] comprises documents with around 900 pages. Half of this amount is derived from specifications on the RM and the rest refers to the AM as well as requirements besides others. This work mainly referred to the following specifications as a preparation for the use of tools:

Reference Model	Archetype Model
EHR IM	Archetype Principles
Demographic IM	Archetype Definition Language ADL (1.4)
Data Structures IM	
Data Types IM	

Figure 25: Specifications, Source: The author

The author related to the EHR IM in matters of a detailed logical EHR architecture. In order to create archetypes with correct attributes and relationships of demographic entities, the specification about the Demographic IM was used. Moreover, development of entities, no matter of demographic or clinical content, requires correct semantics. Due to that, the author referred to Data Structures IM and Data Types IM illustrating their correct use. The utilization of the language specially designed to the needs of archetypes required a comprehension of its syntax. Information about that was taken out of the ADL specification offering a precise explanation. From applying this knowledge, archetypes were developed. A detailed

58 Specifications: http://www.openehr.org/svn/specification/TRUNK/publishing/roadmap.html

description was used, given in the Archetype principles. Further information referred to, had been taken out of other references listed in the text. Among these specifications “Data Types IM” was most often used by the author. This specification explains in detail for which use each data type is supposed to.

The forms, provided by the Brazilian Ministry of Health, contain demographic as well as clinical information. The data is not clearly separated. Therefore, all included data was segmented first and then extracted. The extraction and subsequent rearrangement of relevant information resulted in two new structures representing either data of clinical or demographic content.

Further contemplation of both structures, lead to a simplified classification of items in archetype related categories. Single archetypes were able to cover various data. Thus, demographic information was classified referring to existing archetypes. Due to complexity, clinical data was inserted into an excel spreadsheet, translated into English and associated with archetype elements. In the case data was already represented, the element´s identification number was added. Otherwise information became marked as not existing within the archetype set provided on the CKM.

Out of this aid structures, data was assigned to existing archetypes or related ones in order to edit those. Changes were kept to a minimum at the utmost additions. No data required a complete archetype creation process. Due to the release of a free of charge template designer, Ocean Informatics template designer, during writing process, it made part of that work. In order to test Ocean Informatics Template Designer, a set of archetypes covering demanded data of clinical content only was used. The author assumed this set as sufficient to gain useful results. Doing so, offers the demonstration of a complete process of data representation based on openEHR. This step was supposed to be the enclosing practical part.

Based on a Section archetype created especially for the template designer with the editor of the same company like the template tool, archetypes could be composed, restricted and viewed as a HTML form reduced to demanded data. The author worked subsequent to

proposals published by Sam Heard in order to represent data of interest as conform as possible to the guidelines provided.[59]

All in all, 23 different archetypes were necessary to cover data demanded by the Unified National Health System (SUS), on which policies and plans of the Brazilian Ministry of Health are actually based. The pie diagram (Figure 26) demonstrates the distribution of archetypes in two categories: edited ones and reused archetypes.

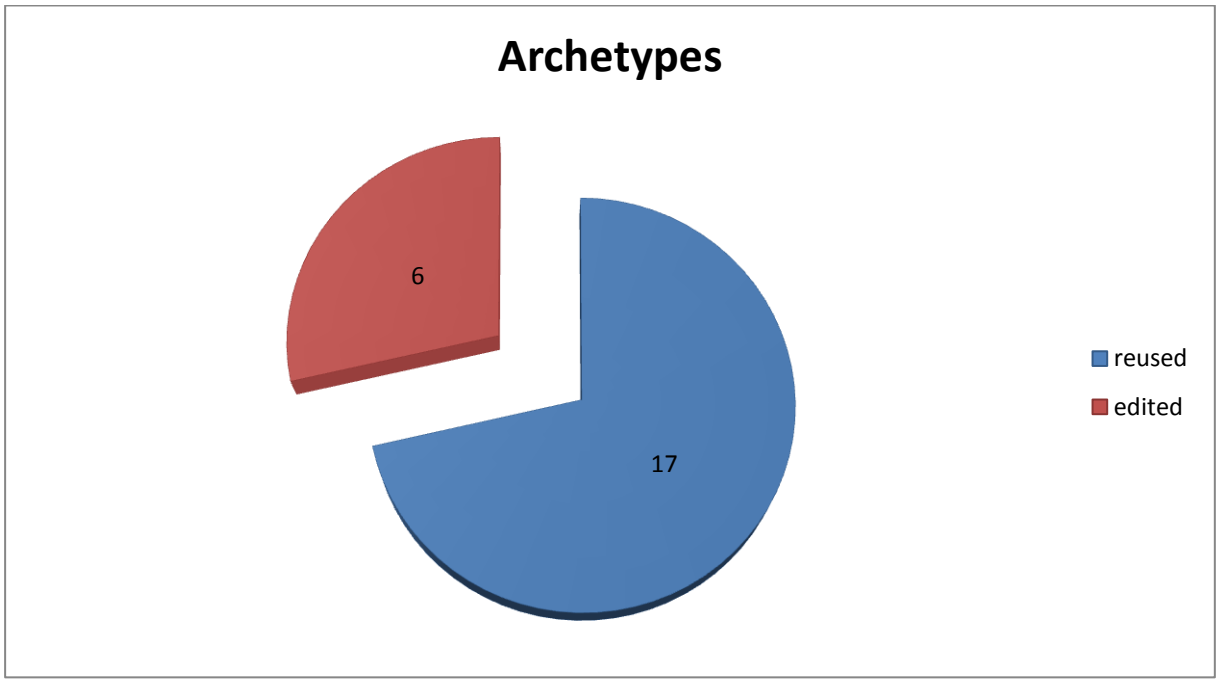


Figure 26: Distribution of used Archetypes, Source: The author

The major part of archetypes is reused and the rest modified conform to the first of the three guiding principles set by Sam Heard. Circa 26% of the archetypes were modified and 74% only reused. This result can be considered as a real achievement and will further underpin confidence that the openEHR approach is absolutely usable towards representation of clinical data. All in all, regarding the peculiarity of data this result is more than satisfying.

5.1 Usability of openEHR

In terms of applicability, the author noticed the conceptional defiance occurring through the vast amount of specifications provided. For a single person it is difficult to keep an overview over the openEHR approach theory. Thus, it proved to be useful to get a feeling for

59 Heard et al.: openEHR Archetypes & Templates 101, p.7, 2007

archetypes by studying their structure and refer to specifications in case of doubts. This synchronous working method enables concrete processing and parallel an application related understanding of more or less abstract specifications.

Creating archetypes by oneself is possible but not recommend when the subject is of complex character. Hence, archetypes should be created within heterogeneous teams in order to take critics immediately into consideration and to speed up publication process of archetypes and thereby to bring these into EHR related circulation. The online available tools do not differ a lot in the range of options. The several archetype editors offer to a greater or lesser extent intuitive use. The two archetype editors applied were the well documented LinkEHR-Ed and the user friendly Ocean Informatics editor. The author prefers the last one due to an intuitive arrangement of items and resulting facilitated use.

The Ocean Informatics template designer is the only archetype composing tool found, which is free of charge. Thus, there is no further template creation tool in order to compare this one. After setting the repository paths in a correct way and regarding a previous creation of a section archetype consisting of sections the later archetypes will be put in, the following steps are very simple. The only things to be considered are a section archetype, correct adjusted paths and a subsequent selection of items which are of interest in order to hide the insignificant rest. The last step includes the illustration of the template. This is so simple, that no explanation is necessary. All in all, the usability of openEHR and its environment including tools provided is feasible.

6. Conclusion and summary

The approach of openEHR is applicable mutatis mutandis to specific demands. The CKM offering a wide range of clinical and demographic resources builds an inevitable reference in matters relating to archetype work. Caution is advised in matters of designation. Some terms describe the same subject, what should be considered when archetypes are quarried for particular content. All possibilities of naming should be tried out. Hence, content should be reconsidered and archetypes with a more general content taken into consideration. In the case that a direct search was not successful, representing data as an adequate specification is an alternative.

It should be regarded that many archetypes were built when CKM was published. It was put to use several years ago. Although they have good content, their design does not meet the latest requirements. Thus, NEHTA tries to improve the structure. The use of four main classes results in ambiguity. Instruction is easy to differentiate from other classes. The same is with Action, which is as an execution of Instructions. But when it comes to allocation of data to Observation or Evaluation it is not clear which class fits best. Observation is used as evidence for something observed or measured. Whereas, Evaluation describes an interpretation of the status observed or measured before. Here, things need to be observed objectively and will be evaluated subjectively. It cannot be excluded that everything observed, will be of objective character. An examination is always subject to a valuing eye when recorded. These two classes can cause problems.

Furthermore, archetype developers suspect that the Item_Tree may not get a lot of use in the future. Reviewers will find it hard to discern why there is apparently inappropriate content in one or the other. They will not understand that it is to support the complementary archetype's function. The fact that, referring to Item_Tree within an archetype is only possible during archetype design time, complicates the use of the approach. This class is not appropriate to in future archetype related work.

It is not possible to keep an overview over all specifications in short time. Thus, a team of specialists should be consulted in case of doubts. Teams composed of people with different backgrounds improve the quality of modeled clinical knowledge. When creating archetypes it is obligatory to take this into account. Due to the fact that the author did not have any contact person to consult in vicinity, necessary conversations were based on the web.

Regarding the effort to represent healthcare related information, it was a manageable work leading to this result, which was based on openEHR and related tools provided. Due to a culture of online available archetypes, user can benefit by reusing most of them. Although it should be regarded that the majority is still subject to changes. Two main repositories can be taken into consideration. The openEHR CKM[60] provides most of archetypes. About this repository it is known that it takes a lot time to get review archetypes. Whereas, NEHTA CKM[61] offers less archetypes but improved in design. Here, processing time is unknown.

It was not studied whether the multilevel approach, which is separating stable information from variable content, enables interoperability. Although this was not in the scope of that thesis, but would be an interesting completion. Interoperability improves data exchange and facilitates focusing on clinical content. That is actually one of the main reasons for the use of openEHR as an approach. In summary, openEHR is besides some negative aspects thoroughly usable in case of representation of clinical data.

6.1 Summary of contributions

First of all, it was shown that openEHR is an adequate approach representing data of clinical content. A partial representation of real data used by the HU was done. Its purpose was to verify whether the approach satisfies the requirement of particular data. Based on these results, the HU can use suggested archetypes listed above, which were changed to a certain extent. This facilitates the later effort on the subject of data representation using the same approach. Trough regarding recommendations towards tools and methods by the author, more effective work can be achieved.

Furthermore, content related change of openEHR-DEMOGRAPHIC-ITEM_TREE.**person_details**.v1 referring to "Education status" called the archetype authors' attention and will be regarded within coming review processes. Many systems will need this kind of information. Thus, it is required to wait for the CKM discussion so that also further changes in the archetype are made after a discussion by the community. Moreover, a mistake was found on the CKM within the demographic archetype openEHR-DEMOGRAPHIC-

60 openEHR CKM: http://openehr.org/knowledge/
61 NEHTA CKM: http://dcm.nehta.org.au/ckm/

ORGANISATION.**organisation**.v1 relating to a slot. Here, a wrong allocation of archetypes was provided. This problem was resolved immediately.

The results of this thesis are of relevance for the Brazilian healthcare system. At this stage, the openEHR approach has the status of a draft by the means of a regulation proposal for EHR format. Thus, the author decided to make this content available. During writing process, provisions were made for the publication of an article in "Cadernos de Saúde Pública" (Reports in Public Health). This journal publishes articles of scientific merit that contribute to the study of public health in general as well as in related disciplines. Altogether, that thesis not only contributes to improving health record in a local but also in a national context.

6.2 Future Research

From the authors point of view some things should be considered towards future research on openEHR. Perceived need for improvement refers to instructions for the use of tools like editors and the template designer. Offering workshops which explain the use of tools in order to create archetypes is a good thing to do. But the international character of users requires an alternative to that, which enables instructions online. The first step towards distance education is given by a few videos explaining the use of LinkEHR-Ed.

In order to improve the governance of archetypes on the openEHR CKM, so that archetypes can get published quicker and used in real environments, a public repository should be considered. This step would offer an improvement of governance and thus faster processing. Furthermore, the coexistence of Observation and Evaluation should be reconsidered. One archetype class satisfying the needs of both would facilitate decisions based on these classes. The same is with Item_Tree. Due to its dubious functionality causing irritations among viewers, this class should be avoided.

Based on these results, further effort can lead to a complete representation using the same approach by representing person related data in the way it can be used for electronic management of Brazilian HUs. And finally, a more intelligent search tool would improve the work with clinical knowledge on the CKM. Redundancies could be prevented and performance strengthened.

7. List of abbreviations

ADL	Archetype Definition Language
AM	Archetype Model
C3SL	Centro de Computação Científica e Software Livre; center for scientific science and open source software
cADL	Constraint Archetype Definition Language
CBO	Brazilian Classification of Occupations
CEN	European Standardization Committee
CPF	Cadastro de Pessoa Fisica; individual registration number for taxpayer
dADL	Data Archetype Definition Language
EHR	Electronic Health Record
GEHR	Good European Health Record
IBGE	Instituto Brasileiro de Geografia e Estatística; Brazilian Institute of Geography and Statistics
IM	Information Model
HIS	Hospital Information System
HU	Health Unit
ISO	International Organization for Standardization
JDK	Java Development Kit
NEHTA	National E-Health Transition Authority
NS	No Statement
openEHR	open Electronic Health Record
RM	Reference Model
SISPRENATAL	Sistema de Informação em Saúde (SIS) no Pré-Natal (prenatal project provided by the Brazilian government)
SM	Service Model
SNOMED CT	Systematized Nomenclature of Medicine Clinical Terms
SUS	Sistema Único de Saúde; Brazilian National Public Health System
TM	Template model
UFRJ	Universidade Federal Do Rio De Janeiro
UFPR	Universidade Federal Do Paraná
UML	Unified Modeling Language
V.D.R.L.	Venereal Disease Research Laboratory
XML	Extensible Markup Language

8. Table of figures

9. References

Allan, Jennifer et al.: *Patient-Centered Documentation: An Effective and Efficient Use of Clinical Information Systems*, Journal of Nursing Administration 30(2):90-95, **2000** http://www.jonajournal.com/pt/re/jona/abstract.00005110-200002000-00006.htm;jsessionid=G1kM04PmQSzv792TpfVkzG1JVblm6pqYdb2QQn4fGMfmplZGcv2q!-1725731959!-949856144!8091!-1.

Beale et al.: *Demographic Information Model*, **2008** http://www.openehr.org/releases/1.0.2/architecture/rm/demographic_im.pdf

Beale et al.: *Data Structures Information Model*, **2008** http://www.openehr.org/releases/1.0.2/architecture/rm/data_structures_im.pdf

Beale et al.: *Data Types Information Model* , **2008** http://www.openehr.org/releases/1.0.2/architecture/rm/data_types_im.pdf

Beale et al.: *EHR Information Model*, p.19, **2008** http://www.openehr.org/releases/1.0.2/architecture/rm/ehr_im.pdf

Beale et al.: *Archetype Definition Language*, p.49, **2008** http://www.openehr.org/releases/1.0.2/architecture/am/adl.pdf

Beale et al.: *Archetype Definitions and Principles*, p.9, **2007** http://www.openehr.org/releases/1.0.2/architecture/am/archetype_principles.pdf

Beale, Thomas et al.: *Data Information Type Model, Ocean Informatics*, Centre for Health Informatics and Multi-professional Education-University College London, p. 19-20, **2008**

Beale, Thomas: *Archetypes and the EHR*, Ocean Informatics, Australia, p.8, **2002**

Bernsteinet, Knut et al.: *Can openEHR Archetypes Be Used in a National Context?* The Danish ArchetypeProof-of-Concept Project, **2009** http://person.hst.aau.dk/ska/MIE2009/papers/MIE2009p0147.pdf

Biomedical Informatics Group (IBIME): *LinkEHR-Ed Archetype Editor v0.8 User's Manual*, ITACA Institute, Universidad Politécnica de Valencia, Spain, **2008**

Brass et al.: *Health Data Management with an Archetype Driven EHR System in Low Ressource Environments*, p.4, NS http://www.helmholtz-muenchen.de/fileadmin/IBMI/PDF/Med-E-Tel_EHRflex.pdf

Bruun-Rasmussen, Morten: *National report of EHR implementation in Denmark* ,p. 23, Denmark, **2009**

Chen, Rong et al.: *Archetype-based conversion of EHR content models: pilot experience with a regional EHR system*, Department of Biomedical Engineering, Linköping University, Linköping, Sweden, **2009**, http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2715396/?tool=pubmed

Christian Kohl et al.: *Facilitating Secondary Use of Medical Data by Using openEHR Archetypes*, University Hospital Heidelberg, Germany, **2010** http://www.openehr.org/wiki/download/attachments/5996988/SecondaryUse_Kohl.pdf?version=1&modificationDate=1290696080000

Demski, Hans: *Kommunikation in ByMedConnect – Umsetzung des CCR mit EN13606*, Helmholtz Zentrum München, Arbeitsgruppe MEDIS, Institut für Biologische und Medizinische Bildgebung, S.10, Deutschland, **2010**http://www.helmholtz-muenchen.de/fileadmin/BYMEDCONNECT/IMG/ByMedConnect_GMDS_WS_Mannheim__Demski_.pdf

Esmaelzadeh et al.: *Normal uterine size in women of reproductive age in northern Islamic Republic of Iran***, 2004** http://www.emro.who.int/publications/emhj/1003/normal.htm

Garde et al.: *openEHR archetypes in electronic health records: The path to semantic interoperability?*, Proceedings of the 50th Annual Conference of the German Society for Medical Informatics, Biometry and Epidemiology, **2005** http://www.egms.de/en/meetings/gmds2005/05gmds484.shtml.

Heard, Sam et al.: *open*EHR Archetypes & Templates 101, p.3, Slovakia, **2007** http://data.nczisk.sk/old/nczisk/ZI/stvrtok/3_SamHeardArchetypesTemplates.pdf

IEEE Standard Computer Dictionary: *A Compilation of IEEE Standard Computer Glossaries*, New York, **1990**

Ingram, D.: *The origins of openEHR*. openEHR Foundation, London, **2002**

Leslie et al.: *Engaging Clinicians in Clinical Content: Herding Cats or Piece of Cake?* Medical Informatics in a United and Healthy Europe, **2009**

Margotta, Roberto: *The Story of Medicine*, New York: Golden Press, **1968**

Michelson, L. et al.: *Comparing different approaches to two-level modeling of electronic health records*, Department of Health Science and Technology Aalborg University, Denmark, 2005 http://www.ncbi.nlm.nih.gov/pubmed/16160245

National Institutes of Health:, *Electronic Health Records Overview*, National Center for Research Resources, **2006**, http://www.ncrr.nih.gov/publications/informatics/ehr.pdf

Păun et al.: Local EHR Management Based on openEHR and EN13606,p. 5,Romania, **2009** http://www.springerlink.com/content/54368wkq5540424r/

Sampaio de Almeida, Marcelli: *Sistema de Informação para a atenção pre-natal no municipio de Colombo*, Departamento de Enfermagem, Universidade Federal do Paraná, Brasil, **2010**

Shinji Kikushi et al.: *Databases in Networked Information Systems*, 6th International Workshop, DNIS 2010, Aizu-Wakamatsu, Japan, p.164,**201**

10. Appendix

10.1 OpenEHR Packages

The graphic given below shows the relation between models building the three main packages of openEHR. A breakdown of the major packages by smaller packages is listed respectively besides. This figure illustrates the granularity of the openEHR structure. The concept makes the reuse of bottom-most levels in higher levels possible. By that packages appear omnipresent. The exchange between packages of different hierarchical classification gets facilitated.

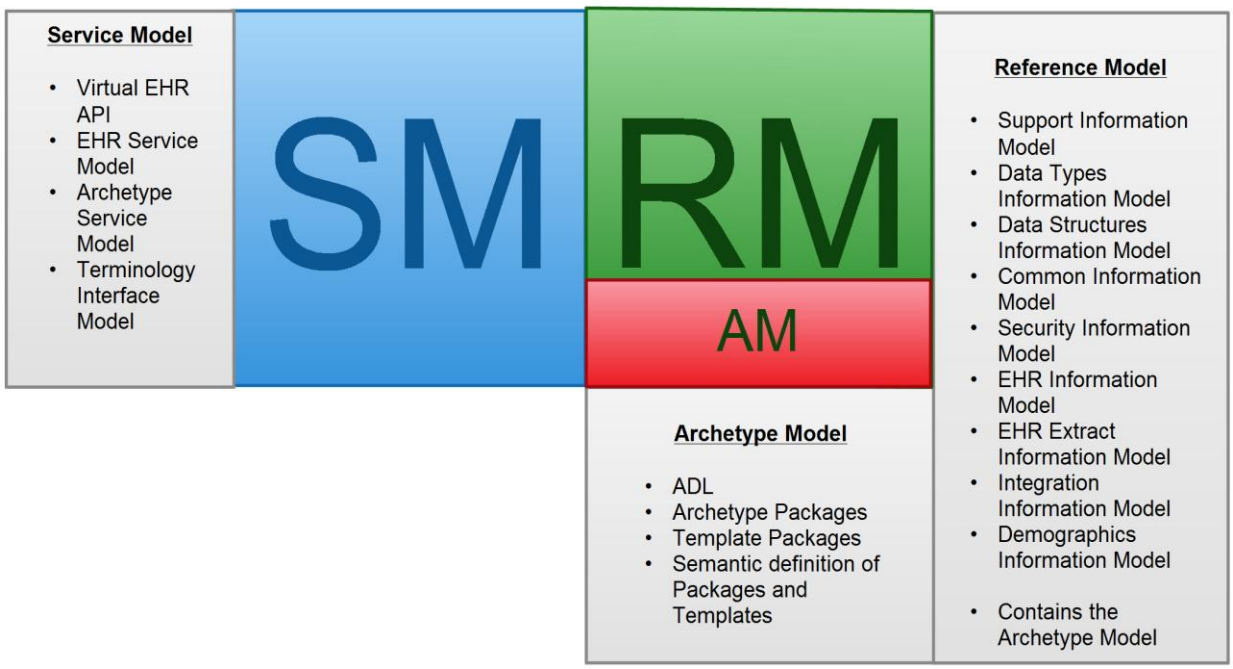


Figure 27: Relation between the openEHR Packages, Source: The author

10.2 Form 1

MINISTÉRIO DA SAÚDE

PROGRAMA DE HUMANIZAÇÃO NO PRÉ-NATAL E NASCIMENTO
COMPONENTE I - INCENTIVO À ASSISTÊNCIA PRÉ-NATAL

FICHA DE CADASTRAMENTO DA GESTANTE

1. Nome do Estabelecimento Assistencial de Saúde	**2.** Código do Estabelec. no CNES
3. Nome do Município	**4.** Código do Município no IBGE
5. Sigla da UF	**6.** Código da UF no IBGE

IDENTIFICAÇÃO DA GESTANTE

7. Número da Gestação no SISPRENATAL	**8.** Gestante acompanhada pelo PSF	
	Código da Área	Microárea
/ - . .		

9. Nome da Gestante	**10.** Data de nascimento
	/ /
11. Nome da Mãe da Gestante	**12.** Raça/Cor

13. Endereço Residencial			
Número	Complemento	Bairro	
Município		CEP	Telefone

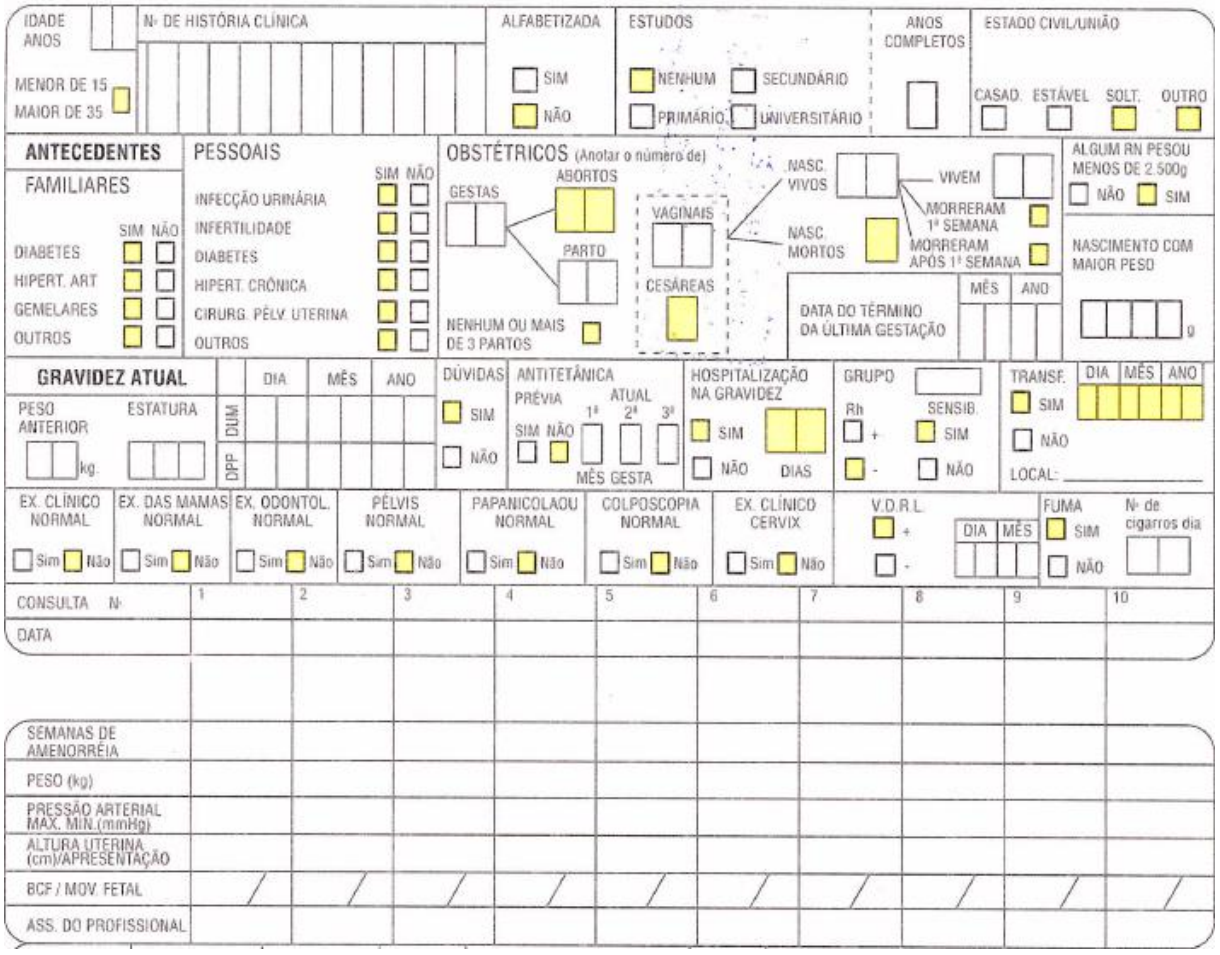
Preencher com apenas um dos seguintes documentos

14. Nº do Cartão SUS	**15.** Nº do CPF	
16. Certidão de Nascimento ou Casamento		
Nome do Cartório	Livro	Folha
17. Identidade		
Número	Órgão Emissor	
18. Carteira de Trabalho		
Número	Série	UF
19. Data da 1ª Consulta de Pré-natal	**20.** Data da Última Menstruação	
/ /	/ /	
21. Assinatura e Carimbo do responsável pela primeira consulta de Pré-natal	**22.** Código Brasileiro Ocupacional (CBO) do responsável pela primeira consulta de Pré-natal	

- Preencher no campo **22** o CBO do profissional do responsável pela primeira consulta, conforme tabela:
  223505 - Enfermeira; **223116** - Médico do PSF/Comunitário; **223545** - Enfermeira Obstetra;
  223115 - Clínica Geral; **2235C1** - Enfermeira do PSF;
  223132 - Ginecologia/Obstetrícia; **2235C2** - Enfermeira do PACS;

- No campo **12** a raça/cor, conforme tabela:
  1 - Branca; **4** - Parda;
  2 - Preta; **5** - Indígena.
  3 - Amarela;

10.3 Form 2


IDADE ANOS
MENOR DE 15
MAIOR DE 35
Nº DE HISTÓRIA CLÍNICA
ALFABETIZADA
SIM
NÃO
ESTUDOS
NENHUM
SECUNDÁRIO
PRIMÁRIO
UNIVERSITÁRIO
ANOS COMPLETOS
ESTADO CIVIL/UNIÃO
CASAD.
ESTÁVEL
SOLT.
OUTRO
ANTECEDENTES FAMILIARES
SIM NÃO
DIABETES
HIPERT. ART
GEMELARES
OUTROS
PESSOAIS
SIM NÃO
INFECÇÃO URINÁRIA
INFERTILIDADE
DIABETES
HIPERT. CRÔNICA
CIRURG. PÉLV. UTERINA
OUTROS
OBSTÉTRICOS (Anotar o número de)
GESTAS
ABORTOS
PARTO
NENHUM OU MAIS DE 3 PARTOS
VAGINAIS
CESÁREAS
NASC. VIVOS
NASC. MORTOS
VIVEM
MORRERAM 1ª SEMANA
MORRERAM APÓS 1ª SEMANA
DATA DO TÉRMINO DA ÚLTIMA GESTAÇÃO
MÊS
ANO
ALGUM RN PESOU MENOS DE 2.500g
NÃO
SIM
NASCIMENTO COM MAIOR PESO
g
GRAVIDEZ ATUAL
PESO ANTERIOR
kg.
ESTATURA
DUM
DPP
DIA
MÊS
ANO
DÚVIDAS
SIM
NÃO
ANTITETÂNICA
PRÉVIA
SIM NÃO
ATUAL
1ª 2ª 3ª
MÊS GESTA
HOSPITALIZAÇÃO NA GRAVIDEZ
SIM
NÃO
DIAS
GRUPO
Rh
+
-
SENSIB.
SIM
NÃO
TRANSF.
SIM
NÃO
LOCAL:
DIA MÊS ANO
EX. CLÍNICO NORMAL
EX. DAS MAMAS NORMAL
EX. ODONTOL. NORMAL
PÉLVIS NORMAL
PAPANICOLAOU NORMAL
COLPOSCOPIA NORMAL
EX. CLÍNICO CERVIX
Sim Não
V.D.R.L.
+
-
DIA MÊS
FUMA
SIM
NÃO
Nº de cigarros dia
CONSULTA Nº
1 2 3 4 5 6 7 8 9 10
DATA
SEMANAS DE AMENORRÉIA
PESO (kg)
PRESSÃO ARTERIAL MAX. MIN. (mmHg)
ALTURA UTERINA (cm)/APRESENTAÇÃO
BCF / MOV. FETAL
ASS. DO PROFISSIONAL

MINISTÉRIO DA SAÚDE — GOVERNO FEDERAL

PARTO

HOSPITAL:

IDADE GESTA — MENOR DE 37 OU MAIOR DE 42

TAMANHO FETAL CORRESPONDE — SIM / NÃO

INÍCIO — IND. / ESP.

MEMBRANAS — INT. / ROT. — DATA: HORA, DIA, MÊS — RUTURA — CEF. / PELV. / TRAN.

TERMINAÇÃO — EXP. / CES. / FORC. / OUT. — HORA, MINUTO, DIA, MÊS, ANO

NÍVEL DE ATENÇÃO — 3º / 2º / 1º / DOMIC. / OUTRO

ATENDEU:

EPSIOTOMIA — SIM / NÃO

LACERAÇÃO — SIM / NÃO

DEQUIT. EXP. — SIM / NÃO

PLACENTA COMPL. — SIM / NÃO

MORTE FETAL — NÃO / SIM

MOMENTO — GRAV. / PARTO / IGNO.

MÉDIC. ENF/PAR. AUX. EMPIR. OUTROS

PARTO

NEONATO

MEDICAÇÃO NO PARTO — ANESTESIA LOCAL / ANESTESIA REGIO. / ANESTESIA GERAL

ANALGÉSICO / TRANQUILIZANTE / OCITOC. / ANTIBIÓTICO / OUTROS / NENHUM

PATOLOGIA NA GESTAÇÃO/PARTO/PUERPÉRIO

- GRAVIDEZ MÚLTIPLA
- HIPERTENSÃO PRÉVIA
- PRÉ-ECLÂMPSIA
- ECLÂMPSIA
- CARDIOPATIA
- DIABETES
- INFECÇÃO CRÔNICA
- OUTRAS INFECÇÕES
- PARASITOSES
- DESPROP. CEF. PÉLV.
- HEMORRAG. 1º TRIMES.
- HEMORRAG. 2º TRIMES.
- HEMORRAG. 3º TRIMES.
- ANEMIA CRÔNICA
- RUTURA PREMAT. MEM.
- INFECÇÃO PUERPER.
- HEMORRAG. PUERPE.
- OUTRA
- AMEAÇA DE PARTO PREMAT. / NENHUMA

RECÉM NASCIDO

SEXO — F / M

V.D.R.L. — – / +

APGAR MINUTO — 1º / 5º — 6 OU MENOS

REANIMAÇÃO — SIM / NÃO

PESO AO NASCER — g. — MENOS DE 2.500g

IDADE POR EXAME FÍSICO — Sem. — MENOR DE 37 SEMANAS

PESO/I.G. — ADEQ. / PEQ. / GRAN.

EX. FÍSICO IMEDIATO — NORMAL / ANORMAL

ESTATURA — cm

PER. CEF. — cm

EX. FÍSICO PRÉ-ALTA — NORMAL / ANORMAL

EX. NEURO. — NORMAL / ANORMAL / DUVIDOSO

PATOLOGIAS

- M. HIALINA
- S. ASPIRAT.
- OUTROS SDR
- APNÉIAS
- HEMORRA.
- HIPERBILI.
- INFECÇÃO
- NEUROL.
- A. CONG.
- OUTRA
- NENHUMA

RN ALOJ./CONJ. — SIM / NÃO

ALTA DO RN — SADIO / TRANSF. / C/PATOL. / ÓBITO

IDADE NA ALTA/ TRANSFERÊNCIA — DIAS / HORAS

IDADE AO FALECER — DIAS / HORAS

ALIMENTAÇÃO — PEITO / MIXTO / ARTIFIC.

PUERPÉRIO

HORA OU DIAS PÓS-PARTO OU ABORTO			
TEMPERATURA			
PULSO (BATIMENTO/ MINUTO)			
PRESSÃO ARTERIAL MÁX./MIN. (mm Hg)			
INVÓL. UTERINA			
CARACTERÍSTIC. DOS LÓQUIOS			

ALTA MATERNA — SADIA / TRANSFERIDA / C/PATOLOGIA

MORTE MATERNA — GRAVIDEZ / PARTO / PUERPÉRIO

ORIENT./CONTRACEPÇÃO — CONDON / D.I.U. / ORAL / OUTRA / LIG. DE TRO. / RITMO / NENHUM

MINISTÉRIO DA SAÚDE — GOVERNO FEDERAL