

# **POLITECNICO DI TORINO**

## Master of science in Biomedical Engineering



### **Clinical Modeling of openEHR Archetypes and Templates: a case of study**

**Supervisor**

Prof.ssa Monica Visintin

**Candidate**

Alessia Posterino

**27 March 2020**



# Index

<b>Summary</b> .....	1
<b>Chapter 1</b> .....	3
<b>Information technology in the healthcare world</b> .....	3
1.1 Healthcare IT state of art.....	4
1.2 Interoperability importance and issue .....	6
1.3 Overview of health standards .....	8
1.4 Computerised healthcare in Italy and Europe .....	10
1.4.1 A case of study.....	12
1.5 Future steps.....	15
<b>Chapter 2</b> .....	16
<b>openEHR</b> .....	16
2.1 Clinical objectives .....	20
2.2 Multi-level modeling and archetypes .....	22
2.3 Separation of responsibilities .....	25
2.4 Archetypes and Templates .....	25
2.4.1 Class of Archetypes .....	28
2.5 EHR Server.....	31
2.5.1 Identification of the EHR.....	34
<b>Chapter 3</b> .....	35
<b>CardioFilo and openEHR</b> .....	35
3.1 Clinical Modelling Suite .....	36
3.2 Clinical Knowledge Governance.....	37
3.3 Creating clinical data models .....	38
3.3.1 Archetype and template on the previous chemotherapy.....	38
3.3.2 Archetype on the previous radiotherapy .....	50

3.3.3 Archetype on diabetes.....	56
3.4 Problems found in the development of clinical data .....	63
<b>Chapter 4 .....</b>	<b>64</b>
<b>GDPR/Legal Issue .....</b>	<b>64</b>
4.1 GDPR key principles.....	66
4.2 Possible reference model.....	67
4.2.1 Data Inventory .....	68
4.2.2 Risk and Impact Analysis .....	69
4.2.3 Organisational measures .....	71
4.2.4 Technical safety measures .....	72
4.2.5 Technical Measures on Applications.....	73
4.3 Scope of the GDPR .....	73
4.4 Sanctions .....	74
4.5 The GDPR on health data.....	75
<b>Conclusions .....</b>	<b>78</b>
<b>Acknowledgements.....</b>	<b>79</b>
<b>Bibliography and Sitography .....</b>	<b>80</b>

# Index of Figures

Figure 1.1 [Types of interoperability].....	7
Figure 2.1 [Holistic integration of health information, patient-based] .....	18
Figure 2.3 [Archetype Meta-architecture] .....	23
Figure 2.4 [Multilevel Modeling and Software Engineering].....	24
Figure 2.5 [structure of an Archetype Repository] .....	30
Figure 2.6 [Blood Pressure Archetype].....	31
Figure 2.7 [minimum EHR in openEHR system] .....	32
Figure 2.8 [tipi di informazione in relazione al processo clinico] .....	33
Figure 3.9 [Date Class containing basic archetype information].....	53
Figure 3.10 [Class State, containing data object information].....	54
Figure 3.11 [Protocol Class, contains information on how the information was collected or measured].....	55
Figure 3.12 [Class History].....	56
Figure 3.13[ date class of diabetes archetype] .....	59
Figure 3.14 [details of the date class of the diabetes archetype].....	60
Figure 3.15 [details of the date class of the diabetes archetype].....	61
Figure 3.16 [protocol class, contains information on how the information was collected or measured].....	62
Figure 4.1 [governance and management model of the action envisaged by the GDPR in Healthcare].....	68
Figure 4.2 [Non-exhaustive example of a list of typical treatments of a Healthcare Provider].....	69
Figure 4.3 [Non-exhaustive example of a list of typical treatments of a Healthcare Provider].....	70

# Summary

The CardioFilo project was born with the aim of becoming a health care model for patients suffering from heart diseases, such as atrial fibrillation (a fairly common heart disorder in which the atrium beats faster and more uncontrollably), subjects undergoing coronary angioplasty (a technique used to treat coronary heart disease and based on the dilation of the artery tract that is occluded, by means of a balloon catheter) or stroke (the process of cell death, necrosis, which affects the myocardium, the heart muscle, following the occlusion of a coronary artery or one of its branches). After a stroke or other acute event and after surgery, cardiology patient care should not stop with returning home. Everyone, in fact, should be directed to a path of secondary prevention and functional recovery. This project, through the creation of a web application for the doctor and a mobile one for the patient and for the doctor as well, in order to have an overview of the situation of the individual patient, wants to be a tool for secondary prevention of heart disease, with the aim of optimizing the patient's therapy, achieve clinical stabilization, provide support on risk factors related to a wrong lifestyle, periodically evaluate the achievement of objectives and also ensure the patient follows the therapies.

The thesis work was carried out at Abinsula srl, a company operating in technology and information field. The application requirements were suggested by the cardiologists of the San Giovanni Molinette hospital and of A.O.U. San Luigi Gonzaga, who requested the implementation of the application. In a previous thesis work, a version of the web and mobile application was developed.

This thesis work mainly focuses on the modeling of clinical data, through archetypes, using openEHR is an open specification in healthcare information technology that describes the management, storage, retrieval and exchange of healthcare data in electronic health records (EHR). The work has been carried out in collaboration with the doctors of San Giovanni Molinette Hospital and of A.O.U. San Luigi Gonzaga; initially several meetings took place in order to better define the contents and the clinical information, useful for the subsequent creation of the archetypes, and try to understand how to best meet their requests and needs. Then we moved on to the actual creation of archetypes/templates, where each clinical concept corresponds to a risk factor reported in the CardioFilo application. The modeling was done through dedicated softwares called "Archetype Editor" and "Template Designer". After having developed them, they were presented to the openEHR community, and a process of verification and validation and possible publication of the archetypes/templates by the community supervisors was started.

With the creation of these archetypes and subsequent templates, an attempt has been made to provide a means to express the definitions and content that clinicians and patients need to record, clinical information has been expressed in a way that is reusable and in some cases universal. The use of archetypes is supposed to ensure a high degree of interoperability between the various EHRs adhering to the openEHR approach.

A multidisciplinary exchange of information is needed to ensure good patient health management so that morbidity and mortality can be reduced. It was decided to use the openEHR standard as it allows a "common language", which enables all those involved to interact effectively.

# Chapter 1

## Information technology in the healthcare world

The medical field is one of the most important areas where information technology can bring a significant contribution to make hospital processes more efficient and less expensive, providing tools to simplify, support, manage and control the work of healthcare professionals. Within hospital facilities there are many different information systems for the management of clinical, diagnostic, administrative and other data flows. These systems are developed by various companies using different technologies, so that it is necessary to implement interoperable solutions to ensure, by information exchange, the interaction among the systems. [1]

The term "e-health" usually refers to all possible ICT applications in the medical field. A more accurate definition states: e-health is a discipline that lies between medical information technology, public health and business, so it deals with the provision of healthcare services through information technology.

Health Informatics is an evolving scientific discipline that deals with the collection, rescue, recovery, communication and use of medical data. This discipline applies computer methods and technologies to problem solving, decision making and health care improvement in all areas of biomedical sciences. [2]

## 1.1 Healthcare IT state of art

The first definition of medical record was born with Hippocrates, introduced together with the concepts of diagnosis and prognosis, with the aim of improving patient observation. Nowadays, thanks to information technology and electronics, electronic health records (EHR) have been developed. The first legislation in Italy to regulate the diary of a patient's state of health dates back to 1938, with the Petragani law.

The idea behind EHR is that any contact between a person and the operating system, such as medical examination, hospital admissions, laboratory tests, generates a detailed digital description that can be collected in an electronic file and consulted by the patient and authorized persons.

The development of information technology in the medical field began in the 1950s, when computers started to be used in this field. However, until the '70s (when there was a rapid development of technologies and hardware) everything remained bound to simple experiments, in fact the term *health informatics* was not yet coined. The term *medical informatics* and *nursing informatics* first appeared in publications only in the 1970s. [1]

The main objective currently concerns the definition of standards that, respecting patient privacy, allow the efficient exchange of information: there is no need to have an electronic health record if it is not possible to compare information from different laboratories or hospitals.

There are two international protocols for the treatment of health data:

1. DICOM, dedicated to the transfer of digitized images between different systems;
2. Health Level Seven (HL7), dedicated to clinical and administrative information in electronic format.

There are currently three profiles used:

1. European: EHRCOM;
2. Australian: HISA;
3. North American: IHE.

One of the biggest challenges is to be able to share clinical information between health professionals in different locations and patients. This includes the openEHR system, a specific open system in health information technology that describes the management, storage, retrieval and exchange of health data in electronic health records (EHR).

In an openEHR system the data are fully interoperable regardless of the programming language and database technology used internally, all data relating to a person's health is stored in a person-centred, vendor-independent EHR.

The content of documents based on the openEHR standard is defined by clinical practitioners and can quickly adapt to the continuous evolution of modern medicine. This standard is used by a number of international organizations like England, Brazil and Japan; openEHR helps to define the standard of medical information technology for the benefit of patients and healthcare professionals. [3]

Medical record management systems have traditionally been developed by software companies which have typically designed a database and a user interface on it to support the work processes of healthcare professionals. Over time, changes have been made on data model and user interfaces. In particular, these models were not owned by healthcare organizations but by software companies, despite the data controllers are the healthcare organizations themselves. Thinking about of a project in which to exchange data between different systems, there is a high probability that integration is not completely possible because two different data models that are under the control of two different software companies are needed. Such situations of data "hostage" have been the basis for healthcare IT business models for the benefit of service and software providers, but they do not maintain a fully sustainable behaviour when considering the interoperability between hospital facilities.

Fascicolo Sanitario Elettronico (FSE), in Italy, is the tool through which citizens can trace and consult the whole history of their health life, sharing it with health professionals and ensuring a more effective and efficient service.

To date there are

- 26,887,409 digitized reports
- 17 regions, 11 of which are members of interoperability
- 11,708,944 FSE active

Fascicolo Sanitario Elettronico is the first manifestation of e-Health culture in Italy by which an architecture at the complete service of interaction between health professionals, family doctor or paediatrician (PG, general practitioner), specialist doctor, citizen has been designed. The effective implementation of the FSE at national level and its subsequent dissemination will also be able to generate significant savings of paper, but also enables a phase of complete clinical and administrative processes review as well as the entire public health organization. Therefore, it differs from the Health

Care Dossier in which health documents and information accessible through this tool have been generated by several structures and not by a single health care structure.

## 1.2 Interoperability importance and issue

*Interoperability* is the ability of a product or system, whose interface is fully declared, i.e. without any hidden parts of code, to interact with other products or systems, existing or still in development, without any restrictions on access or implementations [4].

The aim is the exchange of information between these systems in order to achieve certain objectives. Interoperability is a key aspect if the use of information technology in the medical field has to be possible on a large scale. Just considering that there are many different vendors on the market offering their systems and machinery for hospitals, these one must be able to be integrated with each other and able to communicate. Doctors should be able to be connected:

- with each other when it comes to patient transfers;
- with pharmacists to avoid mistakes that are harmful to patients' health;
- to the patient's medical records.

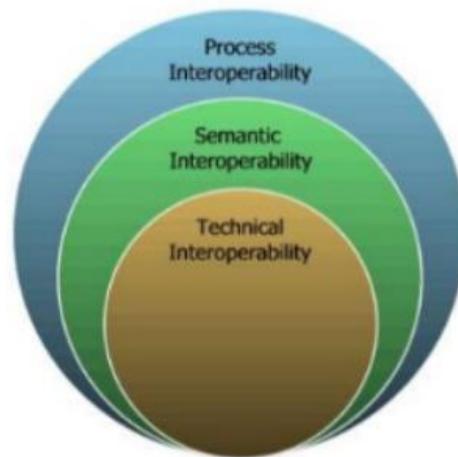
In addition, hospitals must be linked to each other to support the transfer of information between them, and laboratories must be connected to the patient's medical data, i.e. have access to it.

In addition to the medical and patient data aspects, the economic issue should be considered. The population increases and ages, equipment and treatment costs, moreover the paper version of the medical record leads to errors and is therefore inefficient. So for both aspects, economic and paper model, it is necessary and useful to invest in the transition from paper model to electronic model.

In order to ensure interoperability, it is very important to work on standards for data exchange and sharing, so that different systems and organisations can communicate. Three main aspects can be defined in this area of interoperability (see Figure 1.1):

1. *Technical aspects*: ability to move information from one system to another;
2. *Semantic aspects*: systems must interpret data in the same way, so a shared vocabulary must be adopted;
3. *Syntactic aspects*: to ensure that the business processes of the organizations, to which the different systems belong, can work together. Refers to the file format to be used or, better still,

the Application Programming Interface (API). The achievement of this objective is facilitated by the presence of appropriate international standards. [1]



*Figure 1.1 [Types of interoperability]*

When it comes to interoperability in healthcare, the main problem lies primarily in the way data is written and stored. Universal interoperability could be achieved if all doctors recorded data in the same structure and using the same data model. Health data are becoming increasingly digitised and electronic health records containing patient clinical data must be available, identifiable and understandable.

To support automated clinical decision and other machine-based processing, the data must be structured and standardized.

The problem is to agree and find the integration formulas that work. E-Health is not uniformly adopted in European regions. The Italian health system, is gradually lagging behind other European countries in terms of telemedicine: now we are 21st in the Euro Health Consumer Index census [5]. Every day in hospitals, clinics, laboratories or through dedicated apps, a large amount of data is collected that is potentially useful to improve the health of all of us, but poorly connected (and connectable) to each other. We are faced with a fragmented ecosystem due to multiple independent services, the different vendors involved and their proprietary software and data formats, as well as the bureaucratic/legal constraints that may arise in this area [6].

In summary, the advantages of interoperability are:

- Better communication of health data: sometimes simply making the right information available at the right time can improve the treatment process;
- Patient safety: The lack of important clinical data concerning patients can lead to errors with even serious consequences.



Figure 1.2 [interoperability concept]

The objective, therefore, is to have detailed information available in order to aim at personalised medicine. The guidelines must be "data intensive" and high specialisation of health information systems is required. It will be important to achieve interoperability of the solutions as soon as possible, with the highest value being attributed to the data: what is missing today is a health data model. In short, the goal is to create the data model that enables and implements interoperability [5].

### 1.3 Overview of health standards

The management of patient information through non-aligned information systems is a very complex process due to:

- technological differences;

- different representations of patient data;
- the distribution of individual patient data in several healthcare facilities.

It is essential, therefore, for a company to have applications allowing the integration of old systems with new ones while maintaining stable what is already present in the database, i.e. makes the various components of an information system communicate with each other safeguarding the existing information assets.

Topics concerning health standards appeared for the first time in the 1960s. One of the first attempts to build health standards is made by ASTM (American Society for Testing and Materials), regarding standards for message exchange, medical data storage and health system security. Subsequently, several proposals for standards are developed. Listed below are some of the most relevant standards:

- *Health Level 7 - HL7*: standard dedicated to the construction of a framework for the exchange, integration, sharing and retrieval of health data, oriented both to clinical issues (e.g. patient data) and organizational issues (e.g. hospitalizations) of health services. It is an attempt to develop a common language for the exchange of messages in the medical field, in progress since 1987. It defines an explicit and shared conceptual model for the creation of health message structures.
- *Digital Imaging and Communication in Medicine - DICOM*: open standard for the management, storage, display and transmission of medical information in order to make medical imaging diagnostics. Computers are not only used to display images, but also to produce 3D images and models from data collected by devices using techniques such as PET, CT, MRI, SPECT. Practically everything that deal with images. It was necessary to develop a standard for the connection and exchange of images from medical devices, as there are many different manufacturers of devices. The DICOM standard is composed of many parts and is quite wide because it has to cover several branches. Each of them has a subgroup that is responsible for developing that part of the standard. This standard is mainly used in radiology, but also in cardiology, mammography and other areas.
- *OpenEHR*: open specification that mainly concerns electronic health records and the standardisation of the data contained in them.

In order to be able to make different systems communicate with each other and to integrate their functionalities, it is necessary to use a common language for the exchange of information.

In this thesis work, the openEHR standard, which forms the basis of the project, is discussed in depth.

## **1.4 Computerised healthcare in Italy and Europe**

Faced with a mature awareness of the role that digitization can play in healthcare, both in the provision of new models of care and in favour of the sustainability of the Health Service, Italy still shows a picture of backwardness. It appears to be insufficient investment in digital health by the public and private sector. Moreover not so digital is the management of the chronic patient, and more generally of the citizens.

Against a total health expenditures, between the public system and the direct disbursement of citizens, which has stabilized in the last five years at around 145-150 billion euros, the estimated need for 2025 is around 210 billion. To this figure should also be added the fact that the Italian population over 65 is growing strongly, already representing 21.8% of the total and projected to almost 35% (one in three citizens) by 2051. [7]

In the perspective of the Europe 2020 Strategy, the digitization process of Italian healthcare still appears to be lagging behind the majority of EU countries on the basis of available indicators. The insufficient performance reflects Italy's low level of eHealth spending, which in 2015 amounted to 1.2% of public health spending, compared to the EU average of between 2 and 3%, with peaks close to 4%. If the National Health Service were to become more computerized with a more advanced sharing of information, a more advanced interaction between patients, doctors, health care operators and facilities, it would allow a gain in efficiency, an optimization in service delivery, a reduction in medical error, an increase in patient safety, an improvement in the management of chronic diseases. [8]

The elements which are hampering the development of e-health in Europe, and which therefore need to be overcome, are:

- first of all, the lack of habit, and consequently of security, in the relationship with digital by users;

- the impossibility of interoperability between IT systems;
- the absence of clear reference standards;
- too often insufficient training among healthcare professionals. [9]

Because of these obstacles, systems must be devised that aim to continue to promote health, prevent disease and provide patient-centred care that meets the needs of citizens. Health and care systems need reforms and innovative solutions to become more resilient, accessible and effective in providing quality care to European citizens. Europe's health and care systems face major challenges such as ageing, multimorbidity, shortages of health workers and the growing problem of preventable non-communicable diseases caused by risk factors such as tobacco, alcohol and obesity and other diseases, including neurodegenerative and rare diseases. A further growing threat is posed by infectious diseases due to increased resistance to antibiotics and new or re-emerging pathogens.

Public health and long-term care costs are rising increasingly in the EU Member States and are expected to continue to do so. Data is a key element in making a digital transformation possible. Health data may be available in various forms and are not managed in the same way in all EU Member States or within national health systems. They are often not even available to the patients themselves or to public authorities, medical staff or researchers to assist them in developing or providing a better diagnosis, treatment or personalised care. Even when they are available, health data are often tied to technologies that are not interoperable, which is an obstacle to their widespread use. Market fragmentation and lack of interoperability between health systems is an obstacle to an integrated approach to disease prevention and to treatment and care better adapted to citizens' needs.

To date, the deployment of digital health and care solutions is slow and varies widely between Member States and regions. Further action at EU level is crucial to accelerate the proper use of digital solutions in public health and healthcare in Europe. Healthcare computerization presents problems specifically related to electronic data exchange, in particular the risk of privacy breaches, cyber security risks and data quality and reliability. As regards the scope for future EU action, respondents gave priority to:

- The development of EU-wide standards for data quality, reliability and cyber security;
- EU-wide standardisation of electronic health records;
- Better interoperability through open exchange formats.

The development of specifications for a European electronic health record exchange format should be based on open standards and appropriate technical expertise, taking into account the potential use of data for research and other purposes. The Commission will also monitor the cross-border interoperability of electronic health record systems and, once in use, the adoption of the European electronic health record exchange format in the EU.

### 1.4.1 A case of study

According to the study conducted by CENSIS on 26 July 2016, it's noticeable that the progress of European countries in terms of digital health is calculated on the basis of four indicators, for each of them we will compare Italy with the rest of the EU Member States:

1. *Search for online information on health issues by citizens:* Italy is below the EU average in 2015, ranking 27th out of 28. The percentage of internet users searching for health information online was 46% compared to an average of 56% in EU countries (see Figure 1.3).

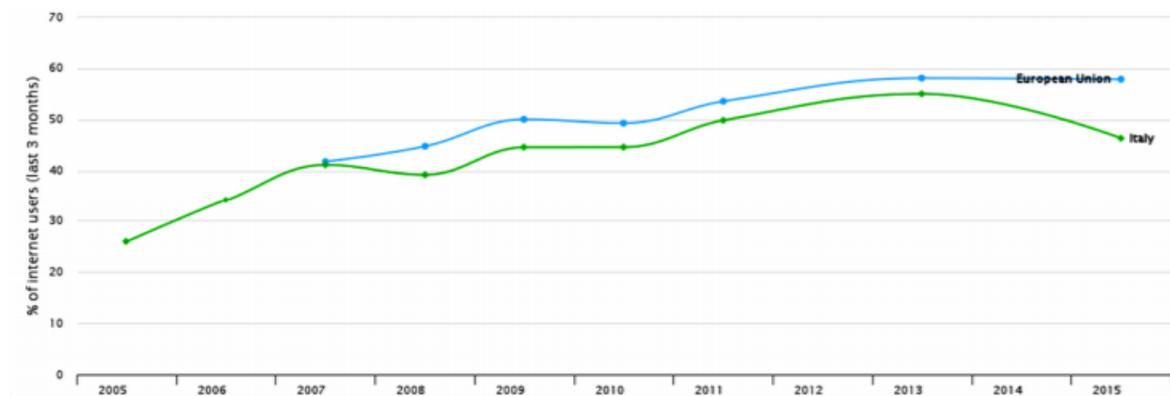


Figure 1.3 [% of internet users seeking online health information]

2. *Booking of medical visits via Web by patients:* Italy was 12th among the 28 European countries in 2014; the indicator is 10% compared to the EU average of 12.5%. The most advanced countries are Spain with 36%, Finland with 35% and Denmark with 34% (see Figure 1.4).

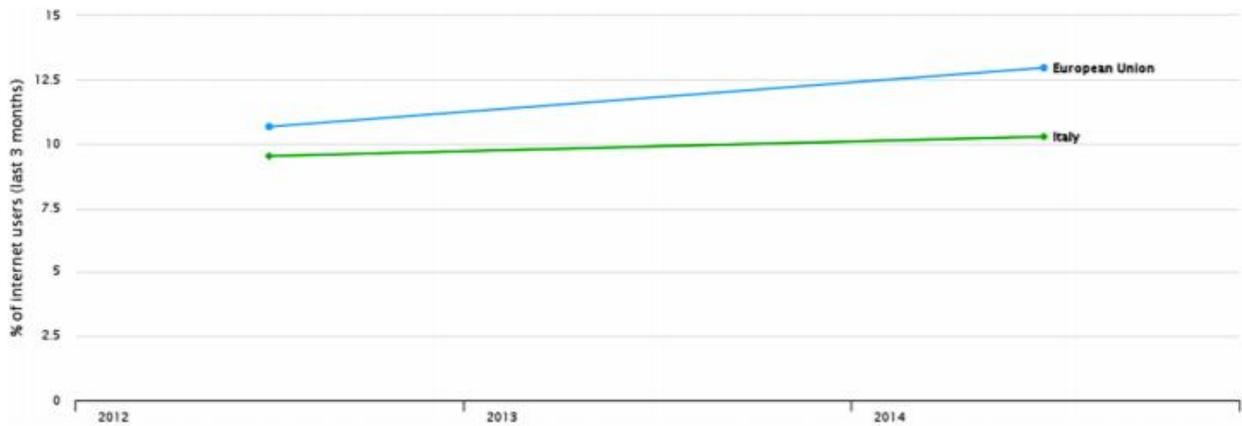


Figure 1.4 [% of internet users booking web visits].

3. *General practitioners sending prescriptions electronically to pharmacists:* Italy occupies the 17th position with 9%, far behind the leading countries: Estonia (100%), Denmark (100%), Croatia (99%), Sweden (97%), Iceland (96%), the Netherlands (94%) (see Figure 1.5).

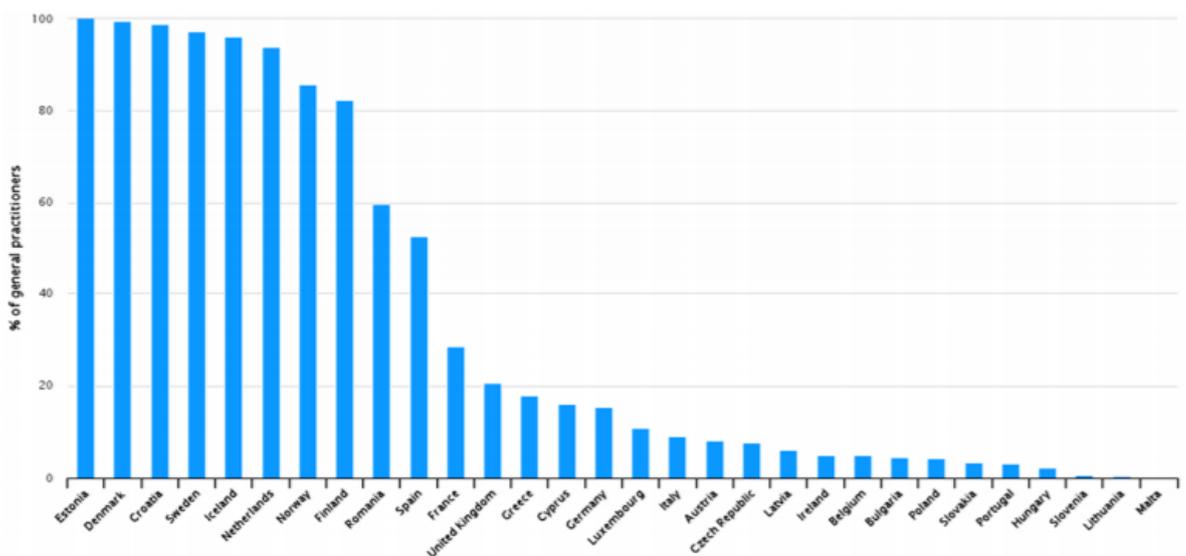


Figure 1.5 [% of general practitioners electronically sending prescriptions to pharmacists].

4. *General practitioners sharing patients' medical data with other healthcare professionals*: also for this field the only data available dates back to 2013, Italy ranks 14th with 31% of general practitioners sharing patients' medical data with other healthcare professionals. The most advanced country is Denmark with 92% (see Figure 1.6).

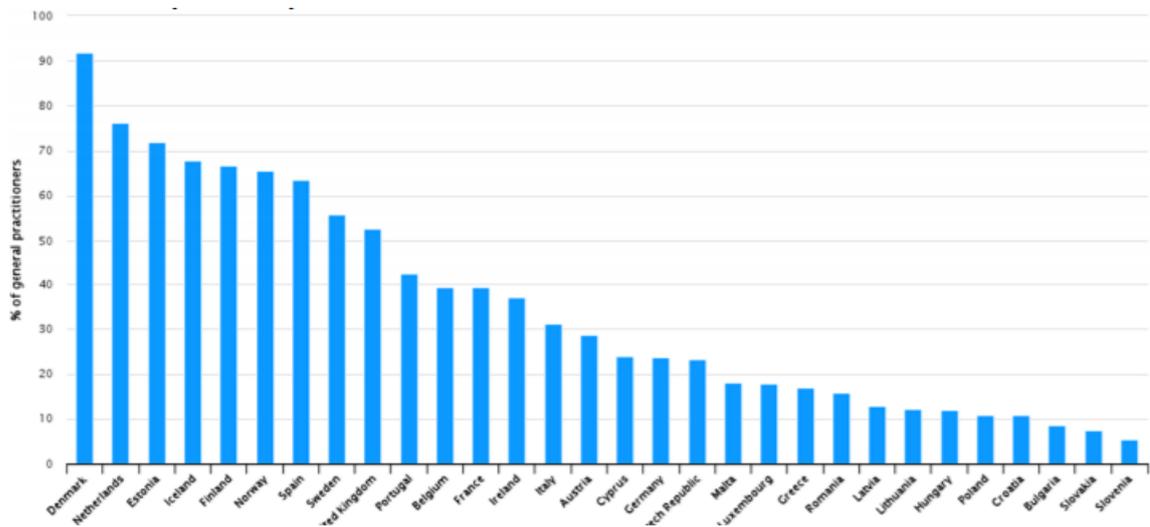


Figure 1.6 [% of general practitioners share patient medical data with other health care professionals and practitioners]

Finally, we are going to mention a general indicator that includes all the others listed above. In this indicator Italy obtains a score of 0.26 (out of a maximum of 1). The European countries with the best score are:

- Denmark with 0.87;
- Finland with 0.84;
- Spain with 0.72;
- Holland with 0.71;
- Sweden 0.67.

It is concluded that Italy compared with the rest of the European Union countries remains far from the best and with poor results. The poor performance is considered to be caused by low spending on digital health. [8]

## 1.5 Future steps

Innovative digital solutions can promote people's health and quality of life and enable better organisation and delivery of health and care services. For this to happen, these solutions must be designed to meet the needs of people and health systems and be implemented so that they are adapted to local contexts.

Digital technologies should be seen as an integral part of health and care and geared towards achieving the wider objectives of health systems. The rapid use of innovative digital solutions is possible:

- Cooperating at European level;
- Sharing the related experiences;
- By measuring the effects;
- By transferring innovation between Member States and regions.

The active engagement of all parties is essential to create a winning strategy that benefits people, health systems and the market. [10]

## **Chapter 2**

### **openEHR**

OpenEHR is an open specification in health information technology that describes the management, storage, retrieval and exchange of health data in electronic health records (EHR). Medical terminology is constantly evolving and therefore standardization is necessary to develop systems that are able to interface with others.

A traditional EHR usually refers to a specific application or system of the provider and thus risks data being blocked by the provider. In an openEHR system, on the other hand, data is fully interoperable regardless of the programming language and database technology used internally, all data related to a person's health is stored in a vendor-independent, person-centric EHR.

openEHR specifications are the result of 15 years of research and development in EHR and new paradigms (including what has become known as the archetype modeling methodology for content specifications, e.g. blood pressure archetype modeling) and are managed by the OpenEHR Foundation, a non-profit foundation that supports research, development and implementation. OpenEHR specifications include an EHR Extract specification, which is a document defining and describing openEHR data types and their architecture. It includes use cases that include EHR communication, other clinical content messages and EHR synchronization.

These specifications do not primarily concern the exchange of data between EHR systems since this is the core of other standards such as EN 13606 (specifies the communication of part or all of a patient's EHR). ISO 13606 is used to support direct care that is provided to people who are identifiable

or to support population monitoring systems such as disease registries and public health surveillance. The use of EHRs that have different purposes and require anonymisation or merging of individual records, such as teaching, administration and reporting, service management, research and epidemiology, are not at the core of ISO 13606. [12]) and HL7.

OpneEHR is currently used by many research organisations, governmental and non-governmental organisations, charities and non-profit organisations, as one of the key points of openEHR is the common clinical information models, the sharing of EHRs. Content models developed at international, national or regional level represent a solid basis for ensuring technical and semantic interoperability.

Promoting the interoperability of electronic health records through the use of standards is a high objective, but at the same time it is indispensable to plan and, probably, to achieve. To this end, it is necessary that all those involved in the processes of diagnosis and treatment converge towards increasingly uniform practices, defining which clinical vocabularies or ontologies to adopt, the techniques to use for the exchange of health documents, the graphical interfaces to be used, how to assign patient identifiers, the rules to follow for the management of privacy and security and so on. These results, moreover, have an important consequence: to design a functional digital tool to accompany the patient throughout the entire care process, both in the same care structure and, in a transversal way, in several care contexts belonging to a given territorial reality or, in other circumstances, to different realities, national or supranational. It should be stressed that, following a holistic approach, the real beneficiary of an efficient system is only the citizen, around whom the fragmentation of information concerning his "clinical history" can find unity ( see Figure 2.1).

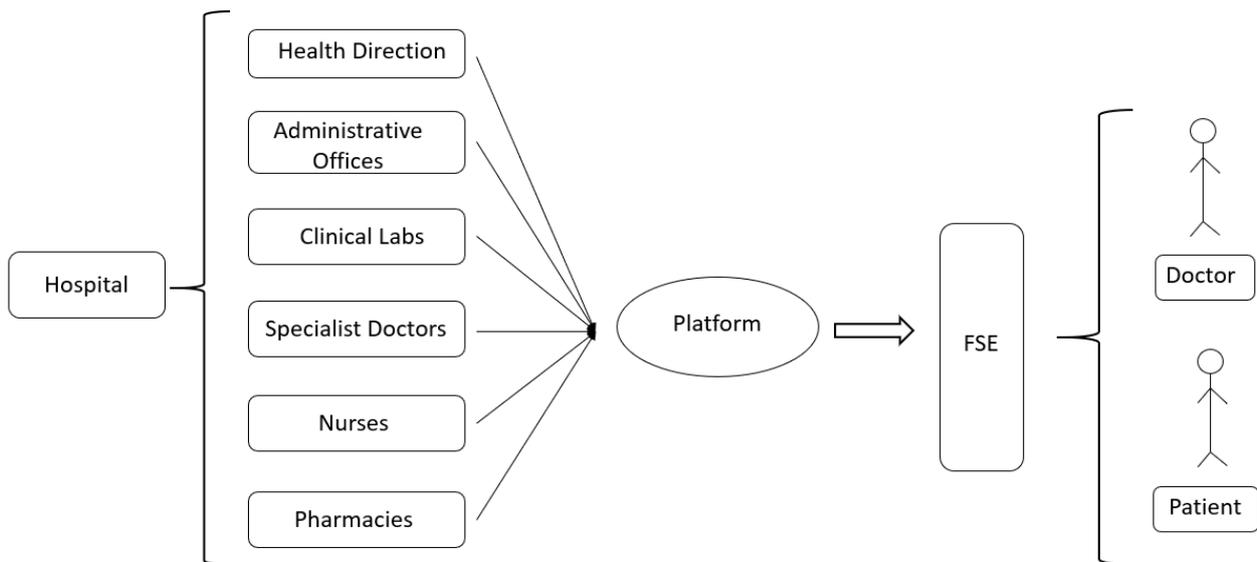


Figure 2.1 [Holistic integration of health information, patient-based]

Some of the issues to be addressed are the following:

- How do the various softwares know the semantics of data? (Semantic Interoperability)
- How to build an electronic patient record shared by all the various devices/organisations that may need access to it?
- How to create systems that can survive the changes?

A two-tier approach is used to address these issues in openEHR:

- *Reference model*: it contains the basic information models that define how the data is represented in the medical record. In this way, those who will maintain this level will not always have to change the model if there are changes. For example, the model contains the definition of data types.
- *Archetype Object Model*: it contains models that can be developed with tools that define clinical knowledge.

In summary, the Reference Model specifies the basic components, which, mixed together, can compose different archetypes. Archetypes can be created to model different concepts, such as blood pressure, weight, medication and so on. The archetype is created with a language called ADL (Archetype Definition Language). There are tools that allow to create and modify archetypes and templates, they generally have graphical interfaces so the user does not need to deal with the code.

This is an indication that everyone can use openHER, as it is a cross-sector application, in the sense that it involves several professionals interacting with each other. Archetypes are used by Templates, which are the result of aggregations of certain archetypes. Templates are basically aggregations of several archetypes, where each one contains a clinical concept.

Thanks to the "Template Designer" and "Archetype Editor" tools of Ocean Informatics [11], it is possible to create templates and archetypes, compile them, and eventually serialize them in other languages such as XML and JSON, and then proceed, for example, to the creation of GUIs (Graphic User Interface) related to the template. Finally, one last aspect should be mentioned: there is an ad hoc language that is built to perform queries based on AQL archetypes - Archetype Querying Language. This allows to be independent from the database technology used, leaving the system developer to create a mapping between the AQL statement and the native querying mechanism.

The openEHR clinical modeling approach involves several artifacts:

- Archetypes: these are the constituent elements of clinical concept (medical observation based on direct examination of the patient and non-surgical treatment of the various pathologies); typically one archetype per clinical concept;
- Templates: which aggregate, combine and bind archetypes to create context-specific clinical data sets and documents such as clinical notes, discharge summary documents or messages that will be used in EHR systems;
- Terminology: a set of values that provide the allowed coding for specific data fields, based on underlying terminologies such as "Systematized Nomenclature Of Medicine Clinical Terms", SNOMED CT, provides the main general terminology to be used in the Electronic Health File; it contains, in fact, more than 311.000 active concepts, continuously increasing, with unique meanings; ICD-10 "International Classification of Diseases", ICD-10, currently in its tenth revision, collects an international classification of diseases and other related problems, particularly useful for statistical and epidemiological studies, as well as for the management of questions concerning public health and hygiene and many others. [13]
- Collaboration processes: for physicians and other domain experts to ensure that archetypes and models are validated, agreed upon and clinically approved;
- A governance process: to ensure that a library of archetypes, models and terminology subsets are developed, managed and disseminated according to transparent business rules.

## 2.1 Clinical objectives

The requirements to be followed during the development of openEHR, as far as assistance is concerned, are set out below:

- The need for a life-long, patient-centered, electronic health record that involves a holistic view of patient needs versus niche problem-solving and decision support techniques for limited diagnostic purposes;
- Integration of different patient views (general practitioner, emergency and acute care, pathology, radiology, computerised patient order entry, etc.) with the vast body of knowledge resources available (terminology, clinical guidelines and computerised libraries);
- Support for clinical decisions to improve patient safety and reduce costs through repeated medical investigations;
- Access to standards-based computer applications.

EHR Integrated Care has a great promise: to generalize and make widely available the benefits of computerization that have been demonstrated individually and in isolated contexts. These can be summarized as:

- Reducing adverse events resulting from therapeutic errors such as interactions, duplication or inappropriate treatments and the associated flow costs;
- Improve timely access to critical information and reduce the physician's time searching for information;
- Reduce the incidence of neglected patients in the healthcare system due to non-communication of information;
- Reduce the duplication of investigations and other tests and procedures due to unavailability of results in the local IT environment;
- Improved prevention and early detection based on predictive risk factor analysis, which is possible with quality EHR data;
- Improved decision making through decision support tools with access to the entire patient record;
- Improve access and calculation of evidence-based guidelines;
- Increase targeted health initiatives known to be effective, based on patient criteria;
- Reduced admissions and readmissions.

A comprehensive statement of EHR requirements covering many of these is ISO 18308, for which an openEHR profile has been created. The requirements summarized above are described in more detail in the openEHR information model specifications.

The openEHR architecture is designed to support the construction of many types of systems. One of the most important is the integrated patient record for shared care, and it is illustrated in the figure 2.2.

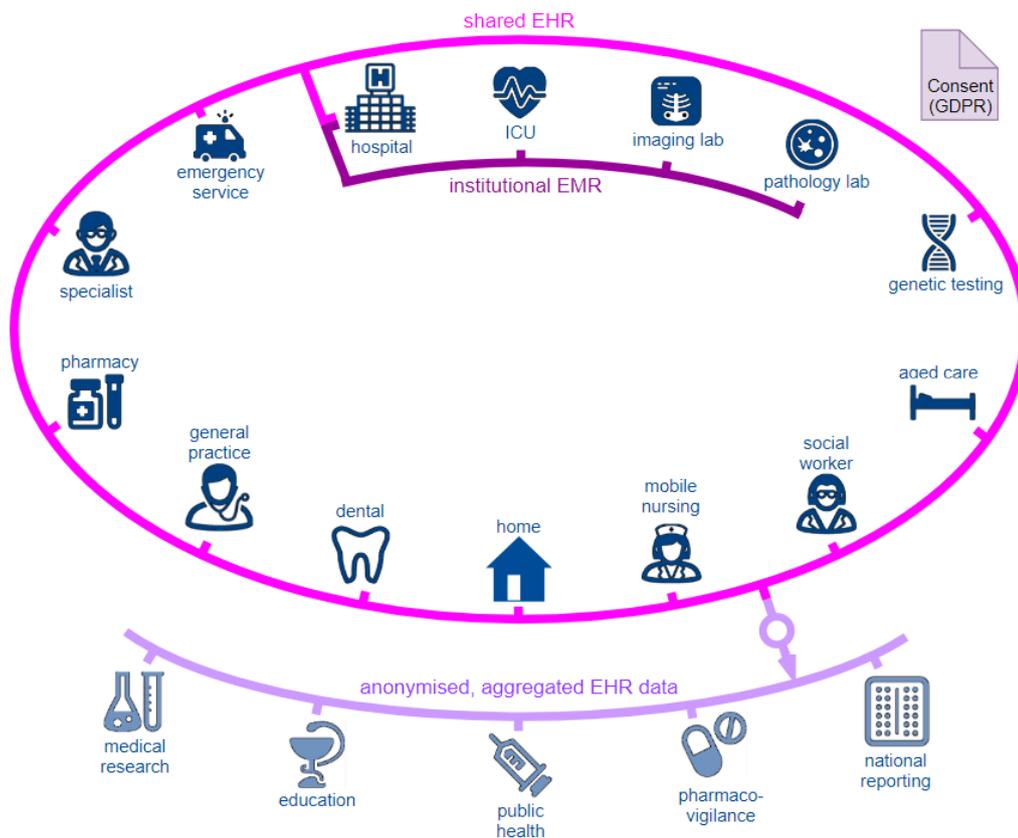


Figure 2.2 [Community Shared-care Context]

In this form, openEHR services are added to the existing IT infrastructure to provide a shared and secure medical record for patients who are seen by any number of health professionals in their community context. Overall, a number of important categories of systems can be implemented using openEHR, including:

- EHR of the shared care community or regional health services;
- summary EHRs at national, state, provincial or similar levels;

- small GP (General Practitioner) desktop systems;
- hospital EMRs (Electromagnetic Radiation);
- consolidated and summary EHRs in federative environments;
- legacy data purification and validation gateways;
- secure web-based EHR systems for mobile patients.

Systems containing anonymous or pseudonymized patient records can also be implemented, as the openEHR architecture defines an electronic patient record where demographic links (e.g. to the national registry or via national health care number) are optional. Where such links are used in the institutional EMR or in the shared EHR context, they can be easily removed in a process of anonymisation.

## 2.2 Multi-level modeling and archetypes

As previously mentioned, openEHR is based on an approach known as "multilevel modelling" or rather "two level modelling", i.e. clinical information is separated from technical knowledge. This allows for self-adaptation and more maintainable systems. There are three levels of models that are required for a system:

1. *Reference model (RM)*: Basic information models that define how data is represented in the medical record;
2. *Re-usable content element definitions*: Archetype definition, i.e. clinical content is defined;
3. *Context-specific data set definitions*: Template definition, i.e. the set of various data extrapolated from different archetypes.

Multi-level modeling significantly changes the dynamics of the system development process. Requirements are collected through various meetings and discussions with specialized personnel, depending on the case of interest. Projects and templates are followed by testing and implementation.

The openEHR Reference model is a basic information model that defines how the data must be represented in the medical record. For example, the openEHR RM model contains the definition of data type or the medico-legal requirements, whereas the Archetype and Template model are groups of data

representing all the clinical knowledge. The same archetype can be used in several different cases giving the possibility to model data only once and to do it according to the standard. The international archetype library is the Clinical Knowledge Manager (CKM) and contains about 500 archetypes. An archetype defines a complete, precise and clinically meaningful concept. This provides the basis for the semantic interoperability of clinical knowledge.

Archetypes and templates also act as a well-defined semantic gateway to terminologies, classifications and computerised clinical guidelines.

The use of archotyping in openEHR engenders new relationships between information and models, as shown in figure 2.3.

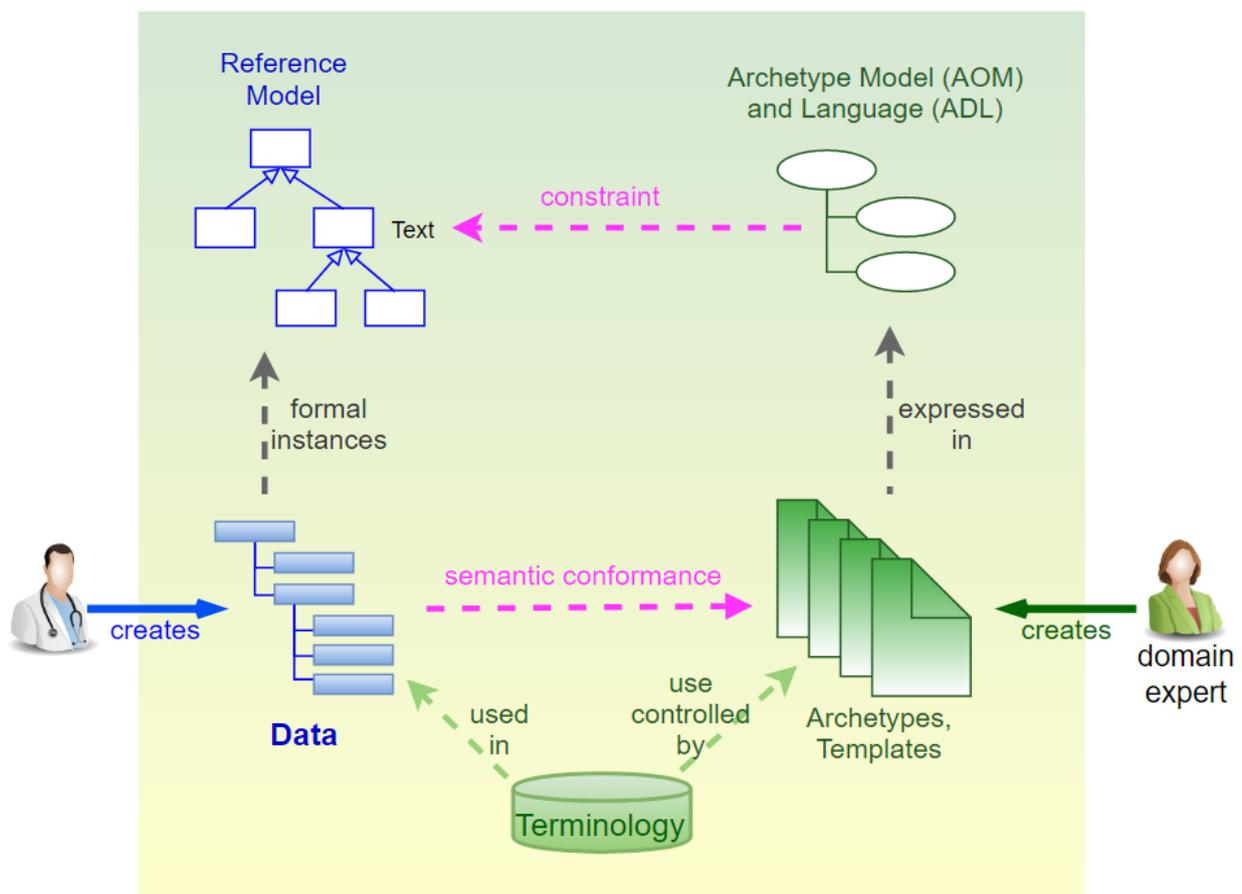


Figure 2.3 [Archetype Meta-architecture]

In this figure, 'data' (shown on the bottom left) conforms in the usual way to an object model (top left). Systems engineered in the 'classic' way (i.e. all semantics are encoded in the software or in the database) are limited to this type of architecture. With the use of multi-level modelling, runtime data now conform

semantically to archetypes as well as concretely to the reference model. All archetypes are expressed in a generic Archetype Definition Language (ADL), which is the basis of ISO standard 13606-2.

When using openEHR, the softwares create models with the help of archetypes, then these models can be reproduced through a user interface and represent the basic structure. Anything that enters the end user's system is validated by the openEHR engine. If the validation is unsuccessful, an error will be generated, if it is successful the data will be stored. Every openEHR compliant system will be able to receive and understand all data. Ergo, interoperability technique and semantics in action!

Therefore, using openEHR, data models are not blocked by software companies, but are publicly available to any vendor wishing to design software based on openEHR (see Figure. 2.4).

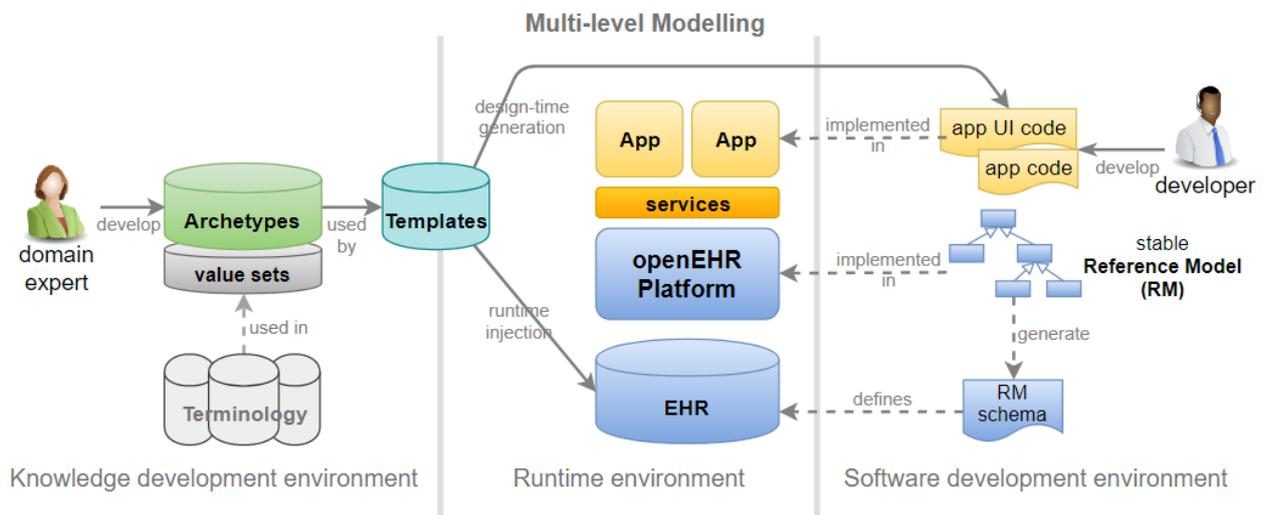


Figure 2.4 [Multilevel Modeling and Software Engineering]

The advantage of this multi-level approach is that the archetypes are independent of technology and semantics.

## 2.3 Separation of responsibilities

In openEHR, IT responsibilities are separated, so by implementing this division, even more complex domains can be dealt with simply by dividing the various functions into different areas, i.e. creating a "system of systems".

If responsibilities are divided, there is also a separation of views, so it will be necessary to define:

- The information everyone processes;
- How they will communicate.

In this separation of responsibilities approach, the various functionality areas are implemented and modeled as an autonomous service with its own well defined interface. For example, a health care environment contains services at three levels of implementation: provider organization (hospital, clinic, etc.); service network (e.g. regional health service) and national network. These levels are intended to be related to three perspectives of care:

- health care delivery: what happens in a provider company, such as a clinic or hospital;
- continuity of care: the passage of the patient through multiple clinics and meetings to achieve a care process designed to achieve a goal;
- health system: the perspective of a national health system, including public health, planning, quality reporting, etc.

## 2.4 Archetypes and Templates

In the multi-level modelling approach, information is formally modelled at two levels. A first level is the reference model, an object model from which software and data can be built. The concepts of the openEHR reference model include elements such as Composition, Section, Observation and various types of data such as Quantity and Coded Text. The second level defines the Archetype, i.e. the various clinical contents, and the Templates, i.e. the set of various data extrapolated from different archetypes. The concepts defined at this last level include for example "Blood Pressure Measurement", "HbA1c Result".

All information conforming to the openEHR reference model (RM) and the collection of information models (IM) is "archetypal", which means that the creation and modification of the content, and subsequent querying of the data, is controllable by archetypes. The archetypes are in turn separated from the data and are stored in their own archive.

Archetypes are the constituent elements of clinical concept: typically you have one archetype per clinical concept.

The openEHR Foundation defines the archetype in the following way:

*"An archetype is a computable expression of a domain content model in the form of structured constraint instructions based on a reference model. OpenEHR archetypes are based on the openEHR reference model. Archetypes are all expressed in the same formalism. In general, they are defined for broad reuse, however, they may be specialized to include local particularities. They can accommodate any number of natural languages and terminologies".*

Archetypes are therefore widely described clinical concepts. Each concept is described by many possible specialists in order to obtain a maximum data set for the concept. They are data sets that represent all clinical knowledge. The same archetype can be used in several different cases giving the possibility to model data only once and to do it according to the standard. Archetypes are designed by clinical or other domain experts and very often require a significant study of a disciplinary area. The development process can be national or international and requires peer review and testing in real systems. This fits in with the semantic value of archetypes, i.e. as reusable content models.

The international archetype library is the Clinical Knowledge Manager (CKM) and contains about 500 archetypes. An archetype defines a complete, precise and clinically significant concept. This provides the basis for the semantic interoperability of clinical knowledge.

While archetypes are a shared realisation of an activity and a clinical model, models are a more local matter, and can therefore be created by the various experts according to the specific case of interest, combining several archetypes. A template will typically be created on three elements:

- what you want to be in a screen form or in a report;
- what archetypes are already available;
- the local use of terminology.

Templates will generally be created locally by tools that conform to the openEHR Template Object Model. [13]

Therefore, ultimately, openEHR templates are more detailed specifications representing aggregations of specific elements of one or more archetypes; for example, up to 60 archetypes may be required to represent all the clinical data elements needed for a summary of a given clinical report.

Each archetype in the template is bound for the proposed clinical scenario, hiding non-essential elements and revealing only those necessary. *Templates* are extremely important as they allow the expression of clinical diversity, even when identical archetypes are used. For example, 10 same archetypes could be used to represent the results of the cardiac examination, but the different constraints will allow appropriate levels of detail to be expressed for use by a general practitioner or cardiologist.

## 2.4.1 Class of Archetypes

Archetypes are the turning point of openEHR's structure, they enable the involvement of clinicians and experts in the design and collaboration of standardized clinical content specifications for electronic health records.

The specification is expressed in Archetype Definition Language (ADL), which is an ISO standard, but able to be viewed and reviewed in 'clinician-friendly' formats, as structured definitions and mind maps. Archetyped data will have the same meaning no matter what context it is used within the EHR and, similarly, no matter which EHR system is used or what language is used.

When archetypes are created, it is important to consider that they must be, as far as possible, interoperable, stable and long-lived. For this reason it is not easy to create archetypes for every case of use, although archetypes can be reviewed, it is preferable to minimize their review process. For this reason, before stating that an archetype is ready for publication and therefore for subsequent use, many experts, clinicians and organisations are involved.

Archetypes can be divided into classes according to their content (see Figure 2.5). Each of these has specific attributes that support the recording and reuse of clinical information.

1. *Composition Class*: can be seen as a container, where the information related to an EHR is stored inside. In particular they represent real documents containing clinical measurements, for example "Encounter" (use as a generic document-level container for recording details of a single interaction, contact or care event between a subject of care and healthcare provider. The contact may be face to face, via telephone or another electronic medium). This class may contain a Section that in turn contains all the clinical details in the archetypes belonging to the Entry class.
2. *Class Section*: represents the header of the document. Usually it does not contain any semantic meaning, but it is used to insert the archetypes related to the Entry class and the Cluster class, which are the most representative of the clinical content. In turn, the Section class is located within a composition class.
3. *Cluster Class*: represents a group of elements put together, they are used in many archetypes and clinical scenarios (e.g. "resignation"). This class can be added to other archetypes and do not have any protocol, timing or event related to it.
4. *Class Structure*: A structure archetype models a structure such as a tree or a list of items and allows reuse in other archetypes;

5. *Class Element*: An element archetype represents a single item, that can be reused.
6. *Entry Class*: This class contains information that will have the same meaning regardless of where it is used. Entry can be entered within the Composition and Section classes, but can also be evaluated independently of them. The Entry class in turn is divided into four sub-classes:
  - i. *Observation*: The Observation class is an archetypal class dedicated to the modeling of data derived from the observation of any phenomenon or state of interest of the patient, such as the results of analysis of a given pathology, blood pressure readings, but also the family history and social circumstances told by the patient to the doctor, patients' answers to the doctor's questions during a physical examination and answers to a psychological evaluation questionnaire. Observations differ from actions in that actions are interventions, whereas observations record only information about the patient's situation, not what is done about it. The significant information of an observation is expressed in terms of:
    - a. *Data*: the current data that is recorded, expressed as a history of events, each of which can be a complex data structure such as a list, table, single (value) or tree, in its own right.
    - b. *Status* : any particular information on the status of the subject of the entry necessary to correctly interpret the data, which is not already known in the medical record (for example if the patient is a pregnant woman, or currently undergoing chemotherapy). For example, level of experience (rest, post-marathon), position (lying, standing). The form of the state attribute is the same as the given attribute: an EVENT HISTORY of ITEM\_STRUCTURE.
    - c. *Protocol*: details of how the observation was conducted, which may include a particular clinical protocol (e.g. Bruce protocol for treadmill training ECG) and/or information on observation tools and methods. This information can always be safely omitted from the user interface, i.e. it has no relation to data interpretation.
    - d. *History*: Contains information about observation times and the "length" of the information.

- ii. *Evaluation*: This subclass is used when, for example, results are to be recorded and acquired. More precisely, they are called "meta-observations", i.e. they represent the views of the physician. Evaluation, unlike Observation, does not contain a temporal model.
- iii. *Instruction and Action*: the former represents assertions about what should happen in the future from the point of view of successful clinic-related processes, such as "the order of drugs". On the other hand, the Actions are assertions about what has been done during clinical activities, such as "administration of the drug". Action is the subclass that completes the Instruction and records its status which could be considered in three different ways: "complete", "programmed" and "cancelled". [14]

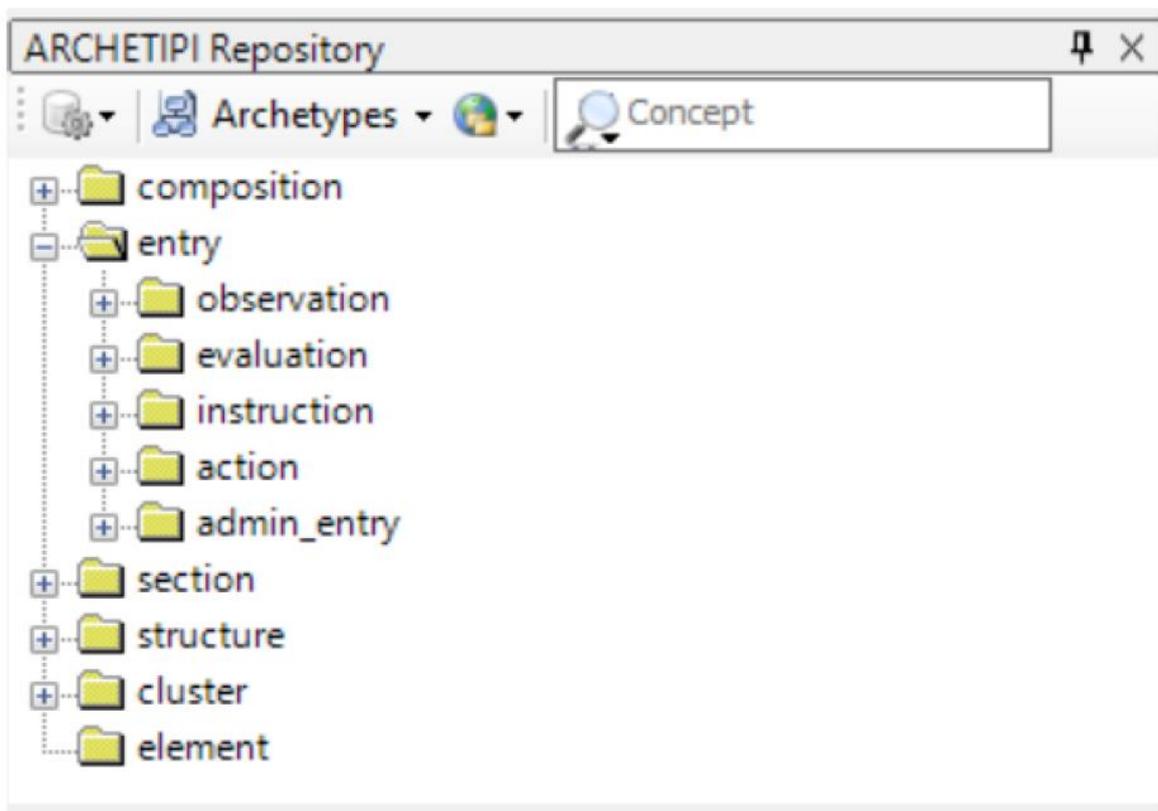


Figure 2.5 [structure of an Archetype Repository]

An example of the observation archetype Blood Pressure is shown in Figure 2.6:

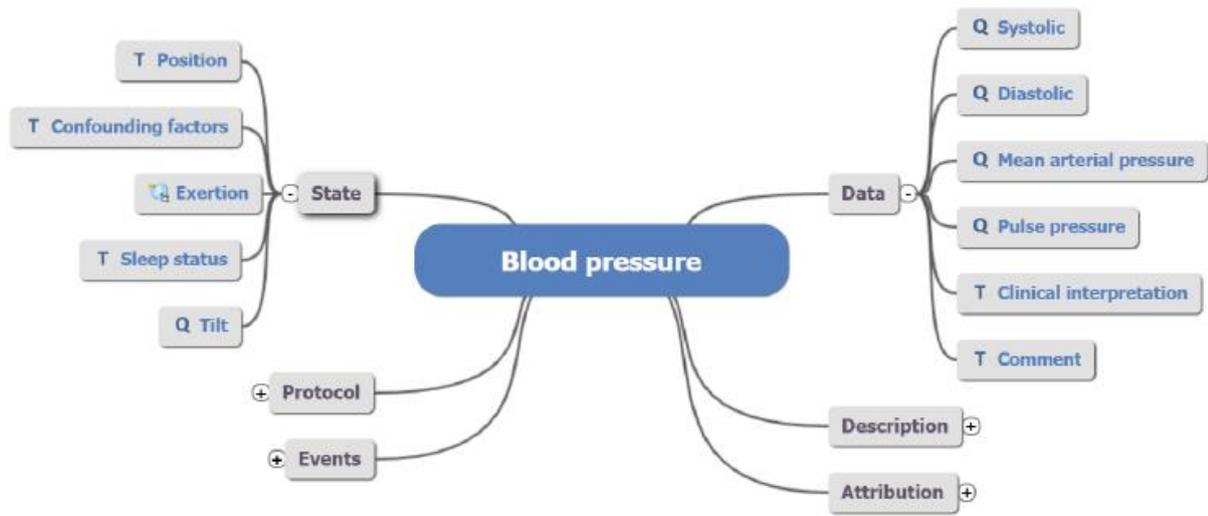


Figure 2.6 [Blood Pressure Archetype]

## 2.5 EHR Server

A system based on openEHR consists of a logical and distinct repository, corresponding to an organizational entity legally responsible for the management and governance of the health data contained within. This could be a regional health service that serves multiple provider companies or a single provider company such as a larger hospital.

In information terms, a minimum EHR system based on openEHR consists of an EHR repository, an archetype repository, terminology (if available) and demographic/identity information, as shown in figure 2.7.

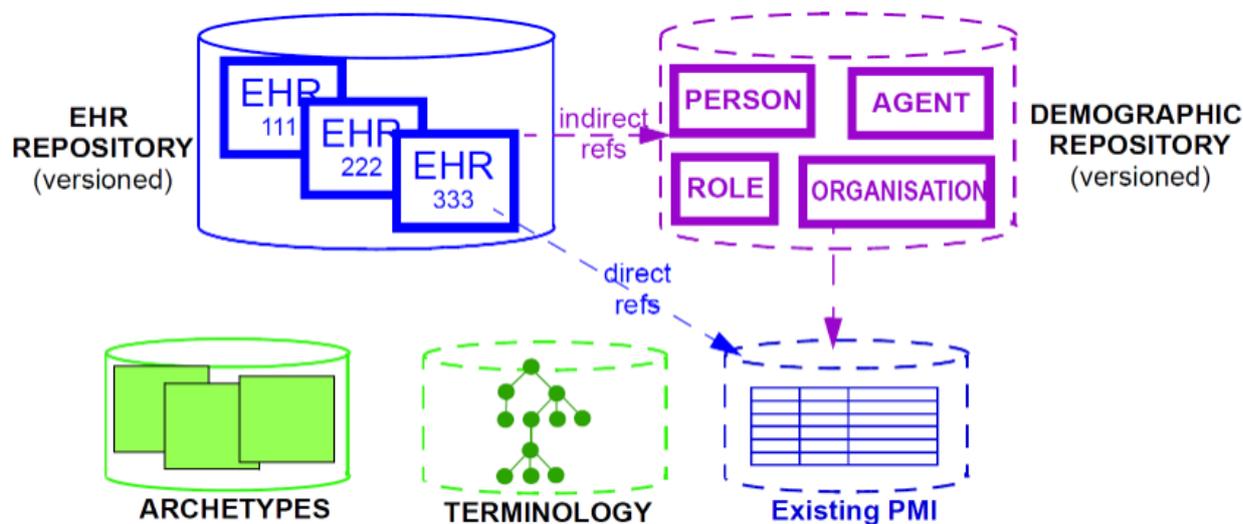


Figure 2.7 [minimum EHR in openEHR system]

The latter may be in the form of an existing PMI (main patient index) or other directory, or it may be in the form of an openEHR demographic repository. In the EHR it is possible to disregard demographic data and/or patient identification data. In fact, the basis of openEHR is precisely the separation of demographic information from the EHR, in this way it is possible to consider a single EHR anonymously without having within it data relating to the identity of the patient.

Once the clinical information has been created in openEHR, it is expressed in "Articles". An item refers to clinical information that may be a single piece of information or may contain a number of pieces of information (e.g. a laboratory test result). All this information will become archetypes for EHR, the choice of the type of archetype will depend on the clinical process considered (see Figure 2.8).

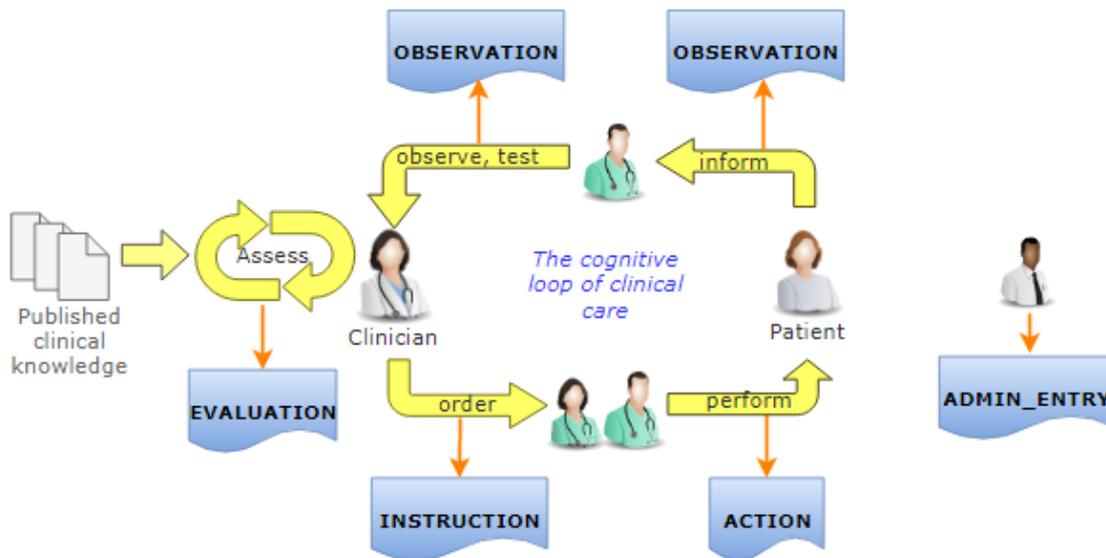


Figure 2.8 [tipi di informazione in relazione al processo clinico]

The figure shows the iterative cycle of information creation related to a specific clinical case. The reported system has been divided into a part related to the patient and a part related to the clinician. The clinician can include both the patient, if he is under observation or undergoing treatment, and health care professionals such as nurses or laboratory technicians, who are entrusted with the task of understanding the patient's clinical status and assigning any treatment. A problem is solved by forming observations, opinions (hypotheses) and prescribing actions (instructions) for the next steps, which can be further investigated in order to solve the problem, and finally the instructions are carried out (actions).

An important part of the investigation process, and of health care in general, is the intervention. For an information system, it is very difficult to manage something that will happen in the future because it can be difficult to track this event within the systems (e.g. a surgical intervention).

The Entry class of the openEHR approach helps in this regard, in particular the Instruction subclass to declare and specify future events and the Action subclass to record what really happened. This model contains some features:

- a single, flexible way of modelling all interventions, whether they be single drug medication orders or complex hospital-based therapies;

- a way of knowing the state of any intervention, in terms of the states in a standard state machine; this allows a patient's EHR to be queried in a standard way so as to return "all active medications", "all suspended interventions" etc.;
- a way of mapping particular care process flow steps to the standard state machine states, enabling health professionals to define and view interventions in terms they understand;
- support for automated workflow, without requiring it.

## **2.5.1 Identification of the EHR**

Each EHR within the openEHR system is identified by an EHR ID, which represents an overall unique indicator. If the system you use is well distributed it is rare for a single patient to identify more than one EHR ID for several EHRs, this can happen if the system is not well distributed and therefore for each subject you have an EHR with its own different ID in each institution. Therefore, the presence of more than one EHR ID per patient in a distributed context is evidence of a lack of systematic connectivity or identification service.

To try to ensure data interoperability, for each patient there is a unique correspondence with only one EHR in the whole system, so no global and unique system should contain two EHRs for the same patient, if this does happen, it means that the system is not well distributed and connected, as previously stated.

When the subject comes to a new institution, he or she may make a request to the environment identification service to ensure that there is already one EHR for him/her in the system. If the result of the search is positive, patients may request to clone their existing EHR, in whole or in part, or create a new one but using the same ID in both cases. In this way the subject can start, for example, a new course of treatment in different places because his/her EHR ID would be the same. [14]

## Chapter 3

### CardioFilo and openEHR

The Cardiofilo project aims to become a health care model for patients suffering from heart disease. In this context it becomes necessary to understand what are "cardiovascular diseases"; we mean a group of pathologies among which the most relevant are ischemic heart disease, arterial hypertension, circulatory diseases of the brain, arrhythmias, diseases of the vessels. These diseases share the main risk factors and, in many cases, etiopathogenesis on an ischemic basis (mainly atherosclerotic disease of the heart and/or coronary, cerebral or peripheral arteries) and are therefore grouped together both for epidemiological analysis and preventive actions. The severity and number of their main outcomes, acute myocardial infarction, congestive heart failure, sudden death, stroke, gangrene, place cardiovascular disease in first place as a cause of death in industrialized countries and are responsible for a significant proportion of permanent disability.

The main risk factors identified include advanced age, male sex, high LDL cholesterol levels, low HDL cholesterol levels, high blood pressure (hypertension), smoking, diabetes, familiarity with cardiovascular disease, previous chemotherapy and/or radiotherapy, HIV in treatment, numerous kidney failure, obesity and sedentary lifestyle. [15]

In this thesis work, the attention was turned to the modeling of clinical data, through archetypes and templates, related to risk factors, not present in the CKM and useful for the CardioFilo project, in particular:

- previous chemotherapy;
- previous radiotherapy;
- diabetes;

For the modelling of clinical data according to the openEHR standard, the following tools were required:

- *Clinical Knowledge Management (CKM)*: The international Archetypes and Templates library, developed by the openEHR community;
- *Archetype Editor (AE) from Ocean Health Systems*: used for the creation of archetypes not existing in the CKM and based on the Language Definition Archetype (ADL) language;
- *Ocean Health Systems' Template Designer (TD)*: used to merge multiple archetypes to create custom templates.

### 3.1 Clinical Modelling Suite

Current openEHR tools for clinical modelling consist of:

- *Archetype Editor*: is an open source tool with a "drag and drop" user interface that allows you to create and modify archetypes locally. The technical aspects of the openEHR reference model and the archetype definition language (ADL), which lie at the basis of the archetypes, are largely hidden from clinical authors. The software allows the design of all classes of archetypes: Compositions; Sections; all types of items, such as Observation, Evaluation, Education, Action and Admin Entry; and Clusters and Elements.
- *Template Designer*: an open source tool with a 'drag and drop' user interface that allows local authoring and editing of templates. The Template Designer is used in machining contexts to create downstream software artifacts specific to the model, such as Template Data Objects (TDO), which are programming interfaces, and also Data Schema Template (TDS), are XML schemas that determine an XML payload that can be used in a document, message, or other communication. All of these artifacts allow trained non-openEHR developers to immediately use openEHR semantic models to build and deploy the software.

## 3.2 Clinical Knowledge Governance

As openEHR archetypes and models were designed by the international community, it quickly became clear that these clinical models required formal support for publication and governance in order to achieve the goal of semantic interoperability. Development of an online knowledge repository commenced testing in late 2008, and is known as the Clinical Knowledge Manager (CKM). It preserves and administers models, archetypes and subsets of terminology as primary resources, as well as related documentation. CKM was developed to enable individuals to share their knowledge, know-how and experience to achieve greater personal and collective benefit than they could achieve if they were alone. By combining a crowd sourcing approach with the collaborative power of the internet CKM supports a shared online community to collaborate together to define high quality, well reviewed clinical archetypes and templates for use in patient care. There are no limits to participation, all interested parties can voluntarily participate in revisions of archetypes or propose an archetype to the community. [16]

The third key aspect of CKM is about rigorous but transparent governance guaranteeing appropriate processes for the management, maintenance and dissemination of the clinical resources. Clinical Knowledge Administrators are responsible for the operations of the CKM, the management of the participating community and the management of a homogeneous and high quality library of clinical content resources, minimizing any gaps between the concepts of archetypes and avoiding potential duplication of archetypes. Coordinators coordinate reviews of the archetype and model community, gradually refining each resources until they are fit to be published as stable artefacts ready for implementation. Translators are able to provide translations and terminologists insert terminological constraints once the archetypes are stable and published. Grassroots clinicians require minimal training in order to participate in archetype or template reviews, they contribute where they have domain expertise, ensuring the clinical content is appropriate, while computer scientists and engineers can contribute their knowledge and skills to ensure that the archetype is technically valid and suitable for implementation [17]. The openEHR Clinical Knowledge Manager (CKM), under the guidance of the openEHR Foundation, is adopted as a source of excellence internationally. As of July 2014 there are over 1100 experts registered on the openEHR CKM from 83 countries, with many archetypes translated into multiple languages to enable cross country sharing of health data. Australia, Norway and Slovenia have established CKM applications for their national eHealth programmes. Negotiations are ongoing with Brazil for a national instance. The UK instance is currently serving the Scottish National Health Service and intends to use it for the UK National Health Service. [18]

### **3.3 Creating clinical data models**

For the definition of the archetypes related to the risk factors mentioned above, several meetings were held with the doctors of San Giovanni Molinette Hospital and of A.O.U. San Luigi Gonzaga, who requested the implementation of the application, in order to define the requirements for clinical data. After fixing with them all the medical contents for the risk factors, we then moved on to the actual implementation of the clinical models through the "Archetype Editor" and "Template Designer" software.

#### **3.3.1 Archetype and template on the previous chemotherapy**

The progress made in health care in recent years has led to an improvement in the survival of cancer patients, but mortality and morbidity due to side effects from cancer treatment have also increased. Among the most important and worrying side effects are cardiovascular disease (CVD), leading even to premature death in patients who have recovered from cancer. Despite progress in the cardio-oncological field, many aspects of cardiovascular disease induced by cancer drugs and radiation remain to be clarified.

In principle, the complications of CVD resulting from cancer therapy can be divided into:

- myocardial dysfunction and heart failure (HF);
- pulmonary hypertension;
- hypertension;
- valvular disease;
- arrhythmias, especially those induced by drugs that prolong QT;
- thromboembolic disease;
- peripheral vascular disease and stroke;
- coronary artery disease (CAD);
- pericardial complications. [19]

With regard to the development of the Chemotherapy risk factor archetype, it was decided to proceed as follows (see Figures 3.1; 3.2; 3.3 and 3.4):

1. The "Archetype Editor" software was used;
2. It was decided, among the many existing classes of archetypes, to create an archetype within the class "ENTRY". The information an ENTRY will mean the same thing regardless of where it is used: this is a fundamental design feature of openEHR architecture (examples may be: Blood pressure - systemic arterial blood pressure; Intravascular pressure, pressure at a point within the vascular system; Weight - the weight of the whole body).
3. It was decided to create an Observation archetype;
4. The archetype has been divided into the following areas:
  - Class Data, containing the basic information:
    - *Tumor type*: The type of tumor being staged.
    - *Anatomical site*: The anatomical site where the assessed cancer is situated.
    - *Tumor staging*:
      - ✓ Tumor size : according to the TNM classification, referring to the letter T, the range goes from T1 tumour smaller to T4 tumour larger.
      - ✓ Involved lymph node: indicated by the letter N, in the range from 0 no involved lymph node to N3 many involved lymph nodes
      - ✓ metastases: Indicated with the letter M, in the range M0 metastases present, M1 absent.
    - *Date of the last chemotherapy cycle*;
    - *Number of chemotherapy cycles*: total number of chemotherapy cycles performed.
    - *Start date of treatment*: the day the chemotherapy started.
    - *End date of treatment*: the day of the end of the chemotherapy.
    - *Drug class*: Classes of drugs with clear potential of cardiotoxicity
      - ✓ *Checkpoint inhibitors*: Atezolizumab, Avelumab, Cemiplimab, Durvalumab, Ipilimumab, Nivolumab, Pembrolizumab, Altri.
      - ✓ *Miscellaneous*: Everolimus, Temsirolimus, Altri.
      - ✓ *Proteasome inhibitors*: Bortezomib, Carfilzomib, Altri.
      - ✓ *Small molecule tyrosine kinase inhibitors*: Afatinib, Axitinib, Dasatinib, Imatinib mesylate, Lapatinib, Neratinib, Nilotinib, Pazopanib, Sorafenib, Sunitinib, Altri.
      - ✓ *Monoclonal antibodies*: Bevacizumab, Pertuzumab, Ramucirumab, Trastuzumab, Altri.

- ✓ *Antimicrotubule agents*: Docetaxel, Paclitaxel, Altri.
  - ✓ *Antimetabolites*: Capecitabine, Clofarabine, Fluorouracil, Gemcitabine, Altri.
  - ✓ *Alkylating agents*: Cisplatin, Cyclophosphamide, Ifosfamide, Altri.
  - ✓ *Anthracyclines*: Doxorubicina, Epirubicin, Idarubicin, Liposomal anthracyclines, Mitoxanthone, Altri.
- Class State, that contains information on the subject of the data, in particular for this case the “*Place where chemotherapy is carried out*”.:
    - *Domicile*: Chemotherapy is performed at the patient's home if it is an oral therapy using tablets or capsules
    - *Hospitalization*: Chemotherapy is performed during the patient's hospitalization if it is a therapy other than oral therapy using tablets or capsules
    - *Other*.
    - *Confounding factors* : Narrative description of any issue or factors that may impact on the chemotherapy.
  - Class Protocol, that contains information on how the information was collected or measured, in particular for this "*type of chemotherapy*".:
    - Endovenous via CVP / agocannula;
    - via orally or capsules;
    - via intramuscular;
    - via subcutaneously;
    - via arterial;
    - via intrathecal;
    - via intracavitary.
  - Class History, which contains "duration of chemotherapy", the maximum period being the last 5 years.

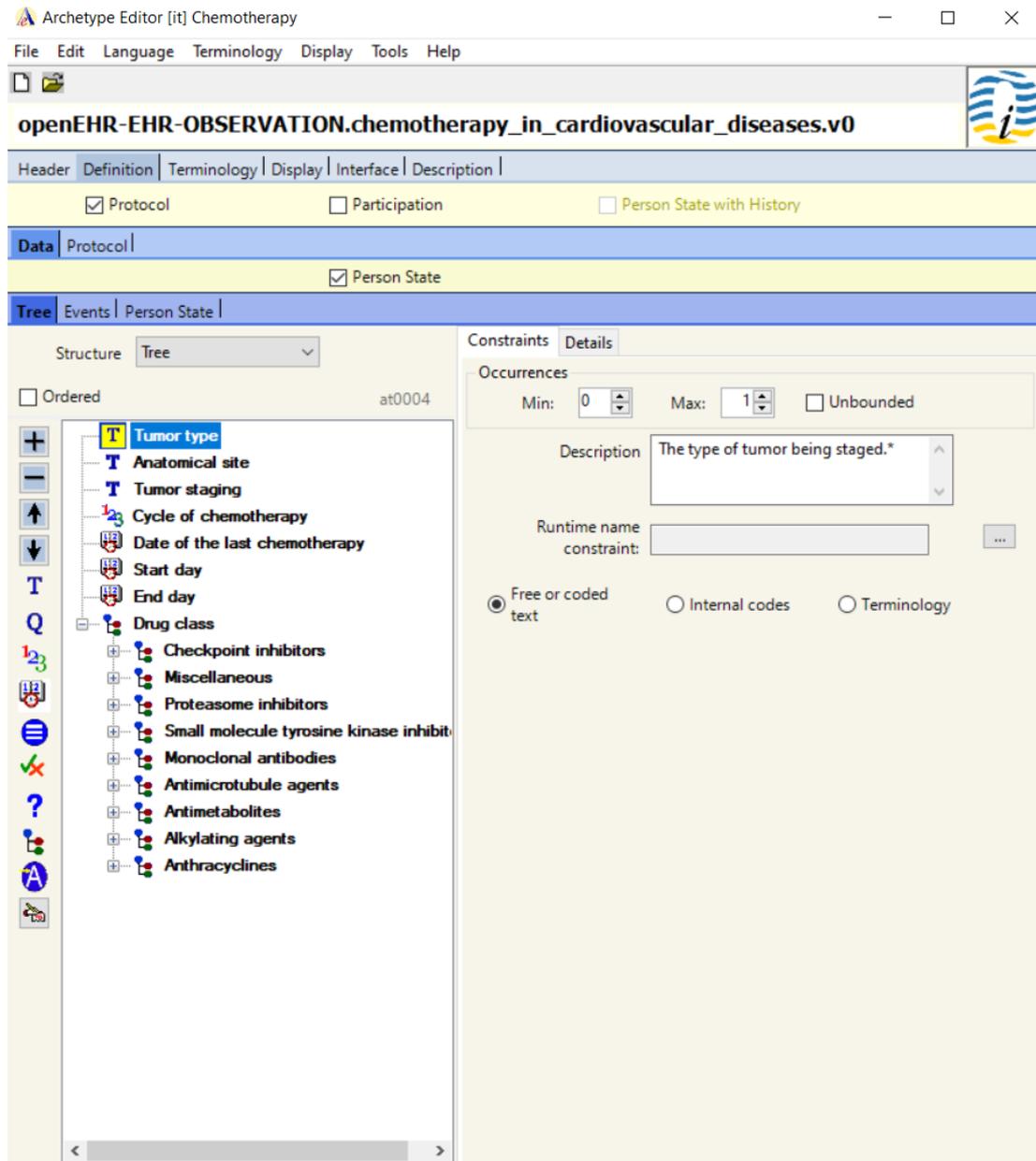


Figure 3.1 [Date Class containing basic archetype information]

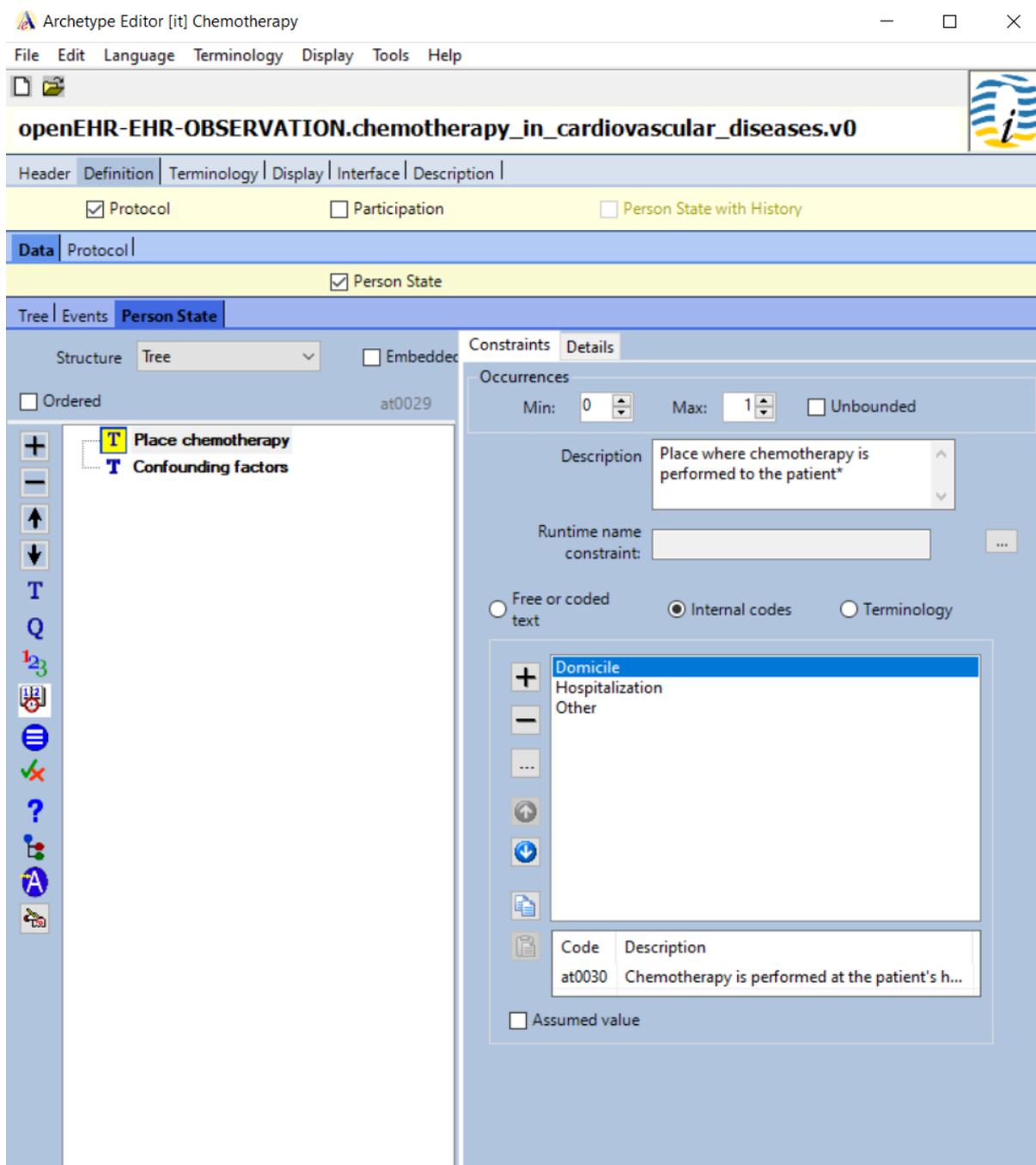


Figure 3.2 [Class State, containing data object information]

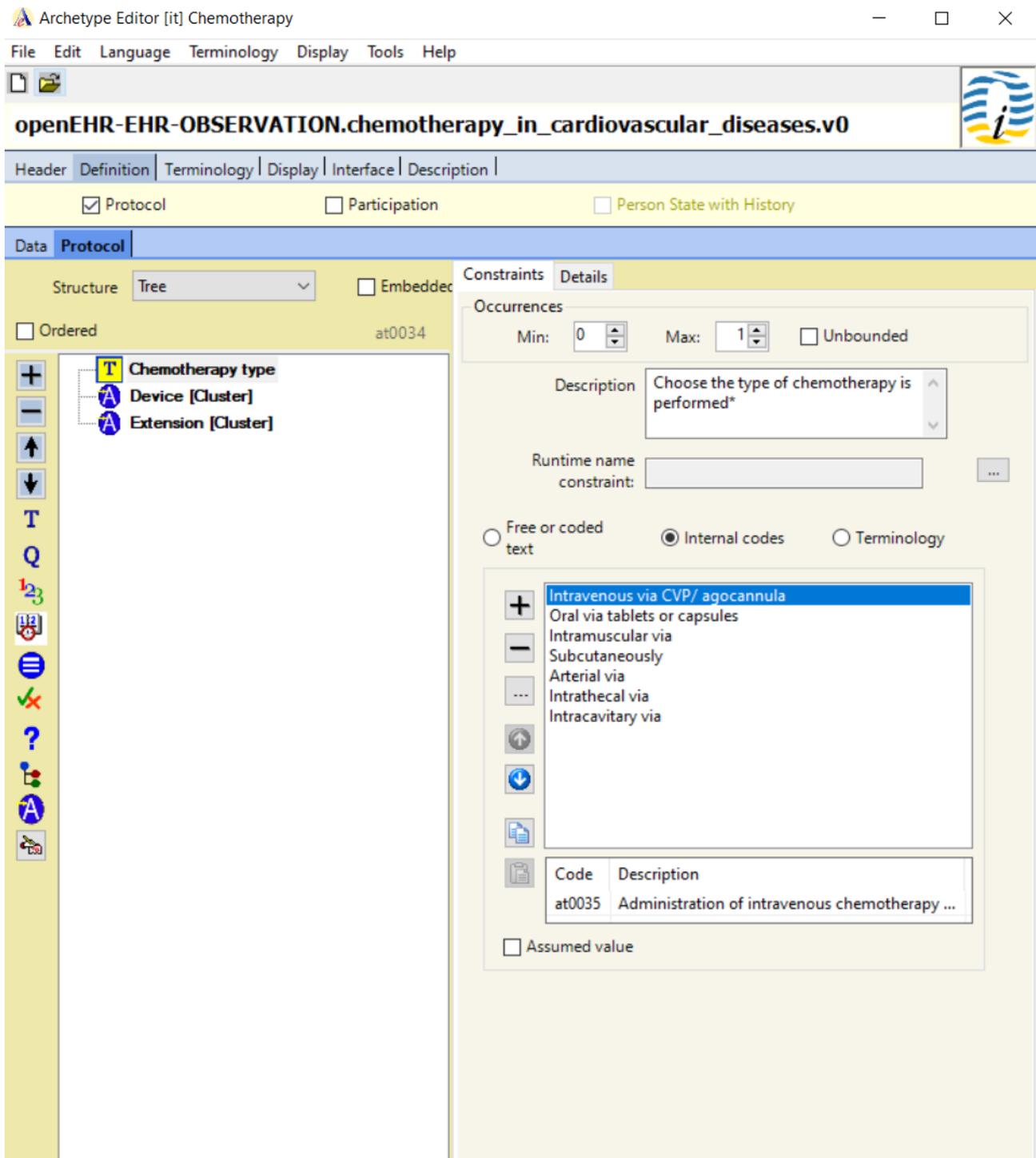


Figure 3.3 [Protocol Class, contains information on how the information was collected or measured]

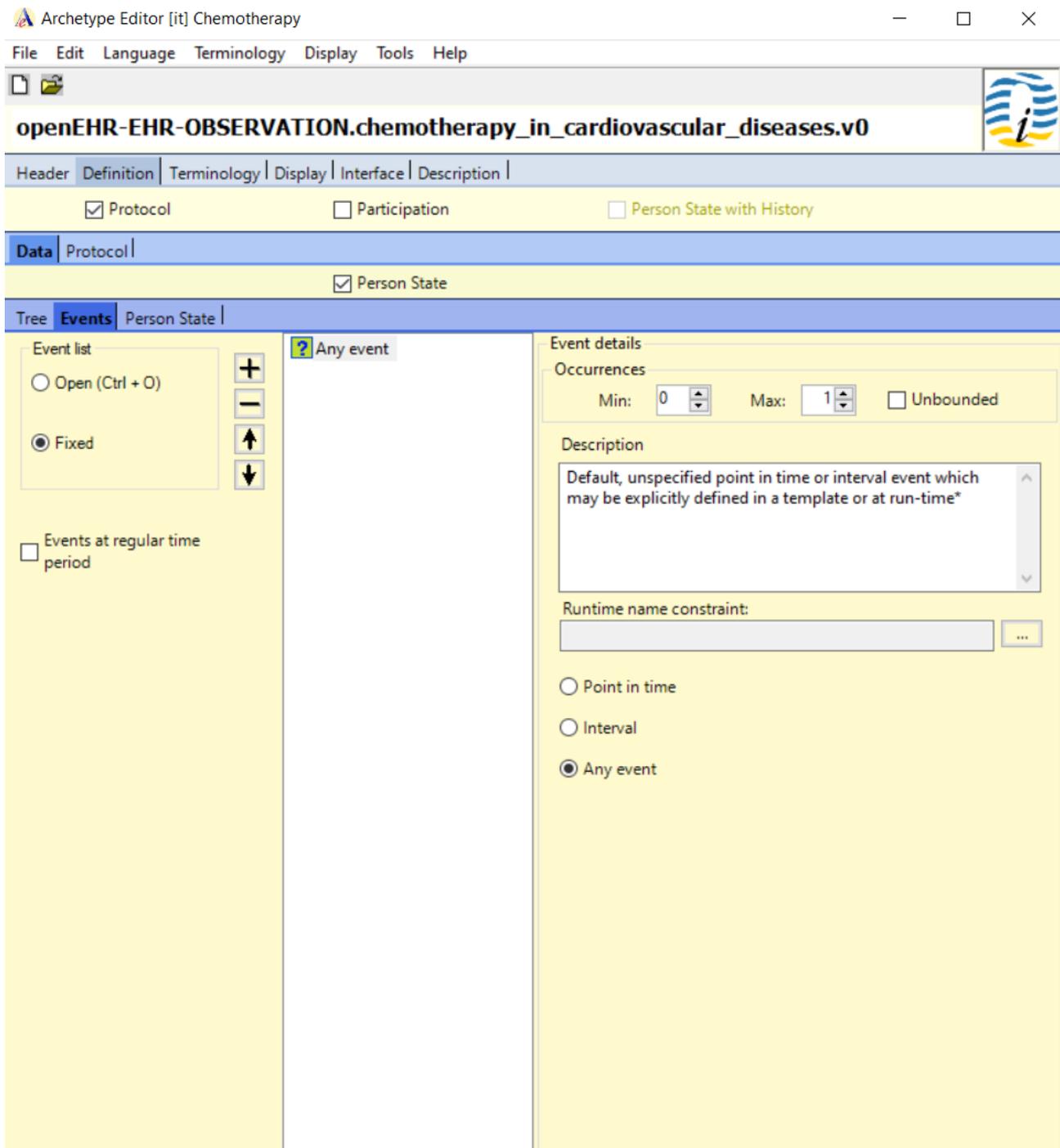


Figure 3.4 [History class, contains duration of chemotherapy]

After completing the development of the chemotherapy archetype, it was uploaded to the CKM site so that it could begin the process of review, approval and then publication by the foundation. Immediately after the archetype was uploaded, a collaboration began with the CKM's founders/resources to understand what changes needed to be made to improve the archetype in order to make it as general as possible and to make it reusable by anyone, in an effort to ensure

interoperability. For this reason it was decided to move to the development of a template related to "chemotherapy for cardiovascular diseases".

Specifically, the Template was created using some parts of the archetype on chemotherapy and reusing some already existing archetypes, of all these only the parts useful for the clinical case of interest have been taken into consideration.

In detail, the following archetypes were considered for the creation of the Template (see Figures 3.5; 3.6; 3.7 and 3.8):

- Encounter: use as a generic document-level container for recording details of a single interaction, contact or care event between a subject of care and healthcare provider. The contact may be face to face, via telephone or another electronic medium).
- Reason for encounter: the following elements have been taken into account:
  - For the class Data:
    - Contact type;
    - Presenting problem.
- Problem/Diagnosis: the following elements have been taken into account:
  - For the class Data:
    - Problem/Diagnosis name;
    - Clinical description;
    - Body site;
    - Cluster “Structured body site”;
    - Date/time clinically recognised;
    - Severity;
    - Cluster “Specific details”;
    - Course description;
    - Date/time of resolution;
    - Cluster “Status”;
    - Diagnostic certainty;
    - Comment.
  - For the class Protocol:
    - Last updated;
    - Cluster “Extension”;

- Medication summary: the following elements have been taken into account:
  - For the class Data:
    - Medication name;
    - Current use;
    - Clinical description;
    - Onset of use;
    - Cumulative dose;
    - Cessation of use.
  - For the class Protocol:
    - Last updated;
    - Cluster “Extension”;
- Chemotherapy in cardiovascular diseases: the following elements have been taken into account:
  - For the class History:
    - Any event;
  - For the class Data:
    - Cycle of chemotherapy;
    - Date of the last chemotherapy;
    - Drug class.
  - For the class State:
    - Place chemotherapy;
    - Confounding factors.
  - For the class Protocol:
    - Chemotherapy type;
    - Cluster “Device”;
    - Cluster “Extension”.
- TNM stage-clinical: the following elements have been taken into account:
  - For the class History:
    - Any event;
  - For the class Data:
    - Cancer type;
    - Anatomical site;

- Primary tumor (cT);
- Regional lymph nodes (cN);
- Distant metastasis (cM);
- Residual tumor (R);
- Other descriptor;
- Stage;
- Retreatment (r);
- Autopsy (a);
- For the class Protocol:
  - TNM Edition;
  - Cluster “Extension”.

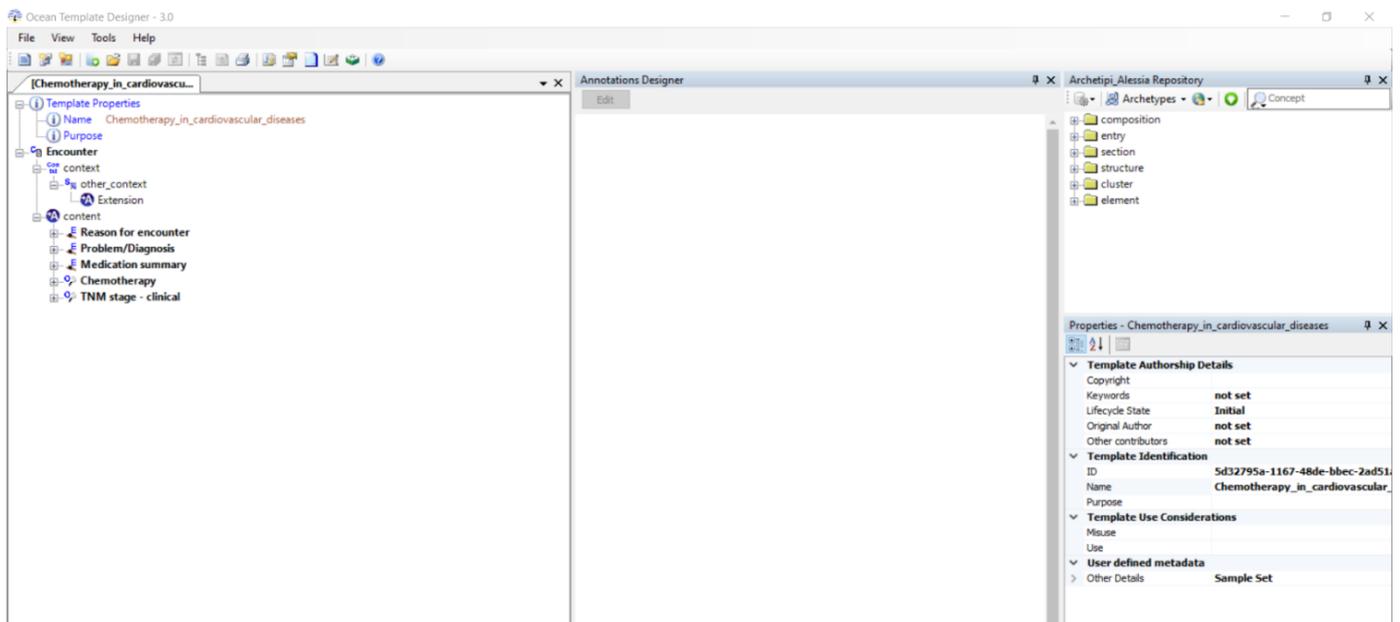


Figure 3.5 [template: chemotherapy\_in\_cardiovascular\_diseases]

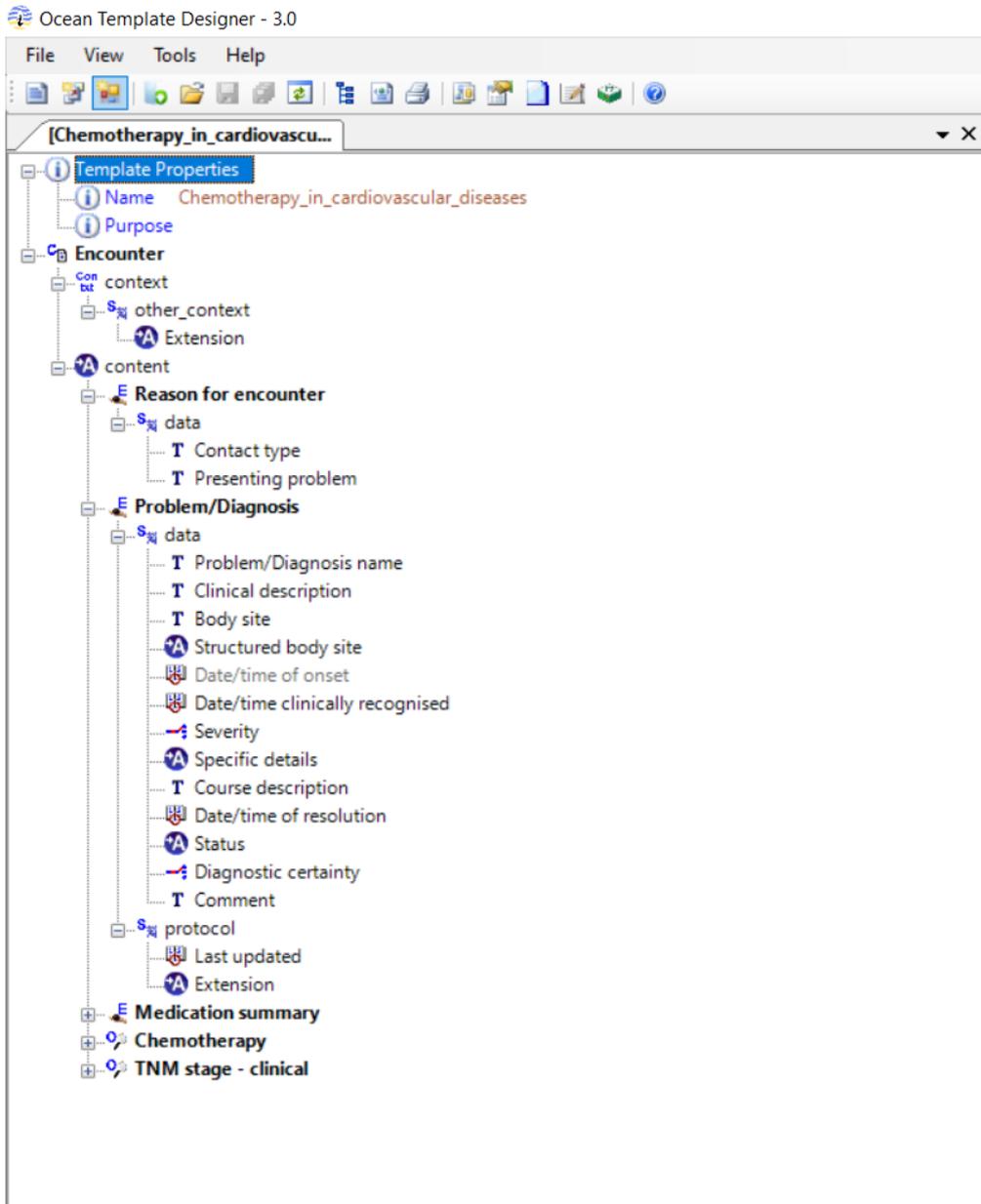


Figure 3.6 [Template details: chemotherapy\_in cardiovascular\_diseases]

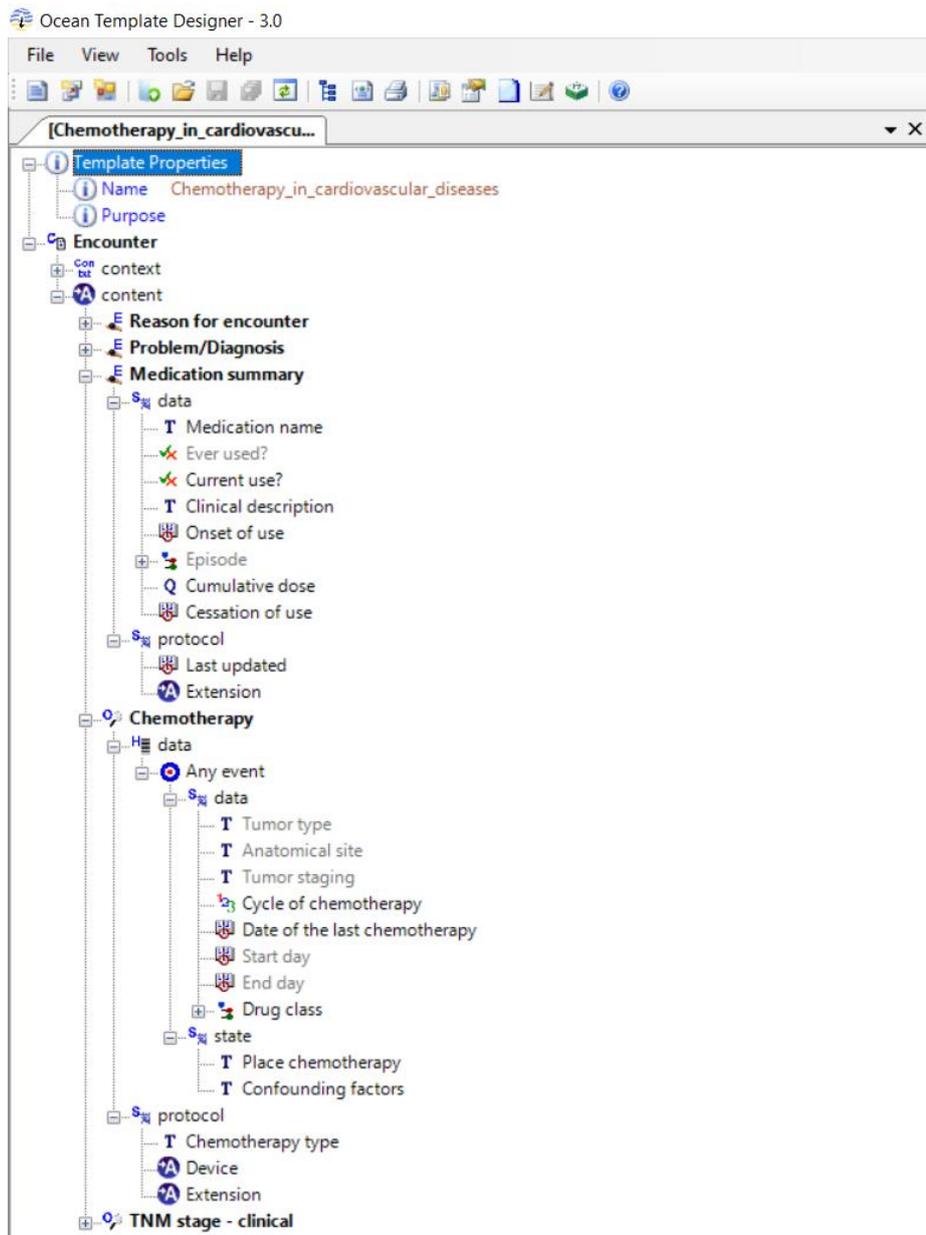


Figure 3.7 [Template details: chemotherapy\_in cardiovascular\_diseases]

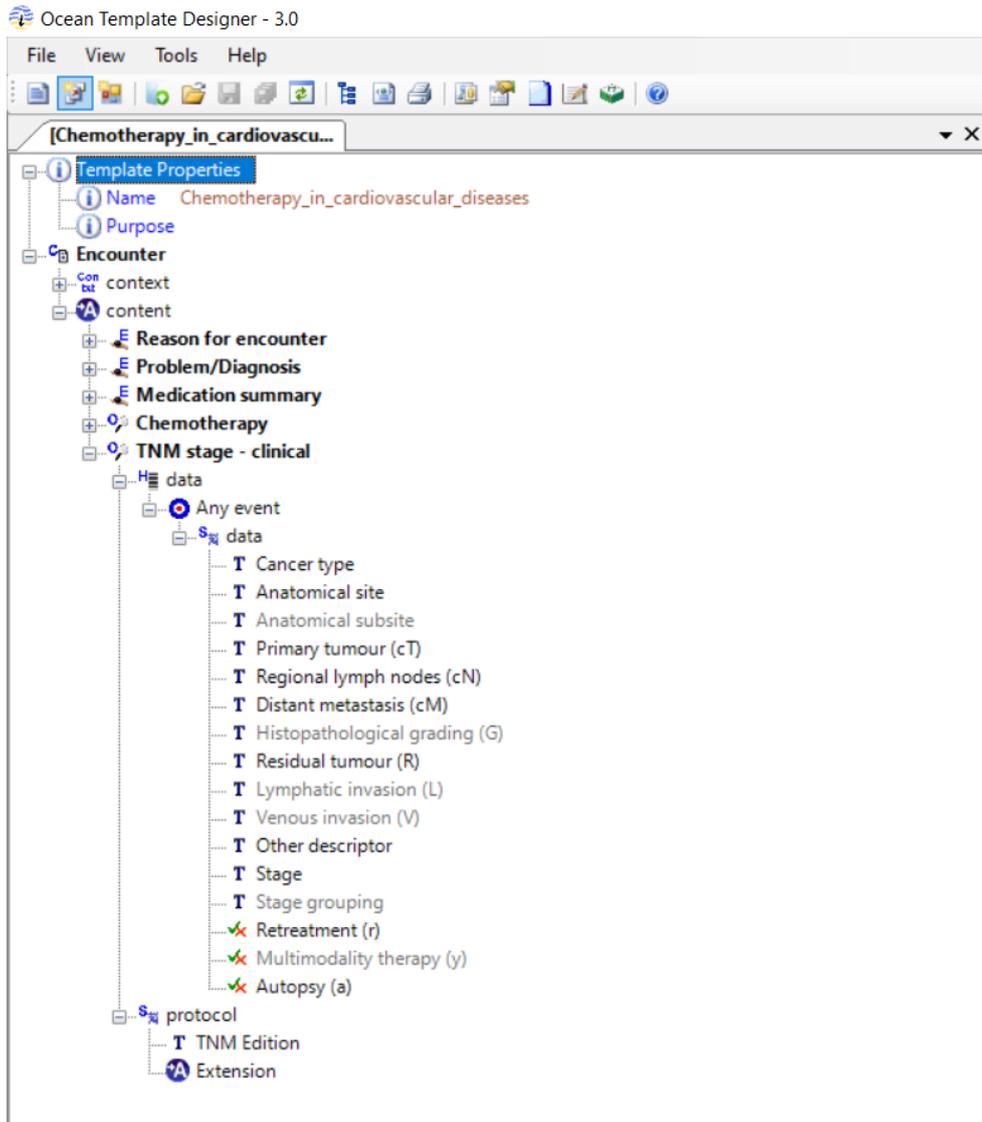


Figure 3.8 [Template details: chemotherapy\_in cardiovascular\_diseases]

### 3.3.2 Archetype on the previous radiotherapy

Radiation therapy is a type of physical therapy that uses high-energy radiation to contain the tumor and kill the cancer cells so that they can no longer proliferate. The radiation that can be used for radiation therapy against cancer are:

- X-rays;
- Gamma rays;
- Rays emitted by certain types of radioactive particles.

Radiation therapy can be administered in two ways:

1. External radiotherapy: so called because the source of the rays is located outside the body;
2. Internal radiotherapy or Brachytherapy: so called because the radioactive substance enters the patient's body.

Radiotherapy can also have a purpose:

- Curative: when the radiation aims to cure the disease by eradicating cancer cells;
- Palliative: aimed at relieving symptoms and pain in cases where the disease cannot be cured;
- Prophylactic: is performed after surgery to eradicate any residual cancer cells.

The appearance of side effects is very variable from one patient to another. Some show only mild effects, others more bothersome, and this depends both on the general health condition and on the location of the tumour and the type of treatment. Among the most important and worrying side effects are cardiovascular disease (CVD). Despite progress in the cardio-oncological field, many aspects of cardiovascular disease induced by cancer drugs and radiation remain to be clarified.

For the creation of a new archetype on the risk factors of the previous radiotherapy, the following steps have been performed (see Figures 3.9, 3.10, 3.11 and 3.12):

- The Archetype Editor software was used;
- It was decided to create an archetype observation;
- The archetype has been divided into the following classes:
  - Class Data, contains basic information:
    - Anatomical site;
    - Mediastinal radiotherapy;
    - Total radiation dosage;
    - Start day;
    - End day;
  - Class State: which contains information on the object of the data, in this case consists of:
    - Radiotherapy place:
      - where the patient is admitted, if the patient's condition is precarious or if there are simultaneous cycles of chemotherapy.
      - clinic
    - Confounding factors: Narrative description of any issue or factors that may impact on the chemotherapy

- Class Protocol, which contains information on how the information was collected or measured, in this particular case we refer to the "*type of radiotherapy*":
  - Internal radiotherapy or brachytherapy divided into:
    - Interstitial Brachytherapy;
    - Endocavitary Brachytherapy;
    - Episcleral brachytherapy;
    - Dosage of brachytherapy that can be minimal or high.
  - External radiotherapy:
    - conformational radiotherapy;
    - modulated intensity radiotherapy;
    - image-guided radiotherapy;
    - 4-dimensional radiotherapy;
    - stereotactic radiotherapy (SABR) and radiosurgical therapy (SRS);
    - adaptive radiotherapy.
  - Systemic radiotherapy:
    - via drip;
    - orally.
- Class Storia, contains:
  - Any Event;

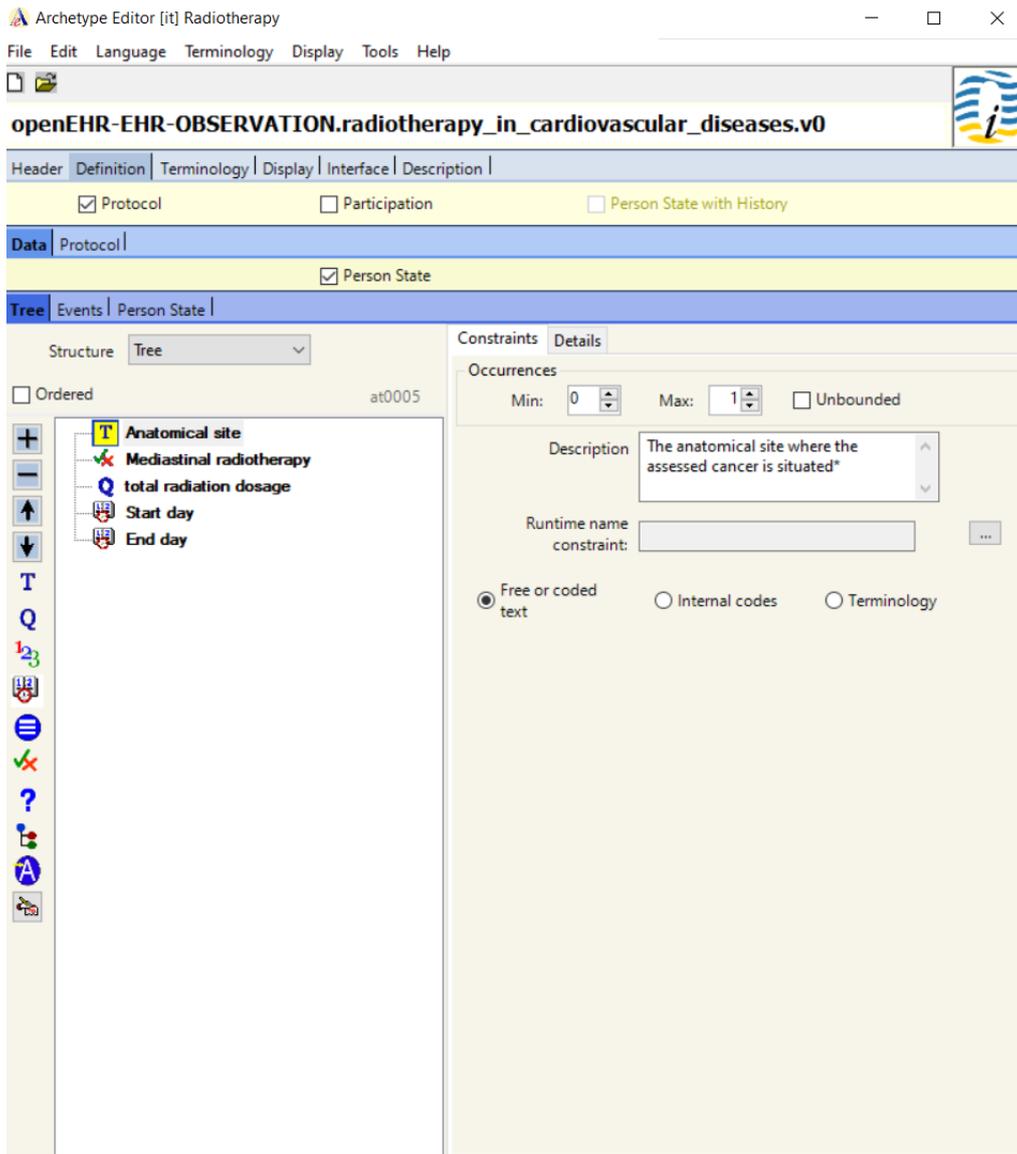


Figure 3.9 [Date Class containing basic archetype information]

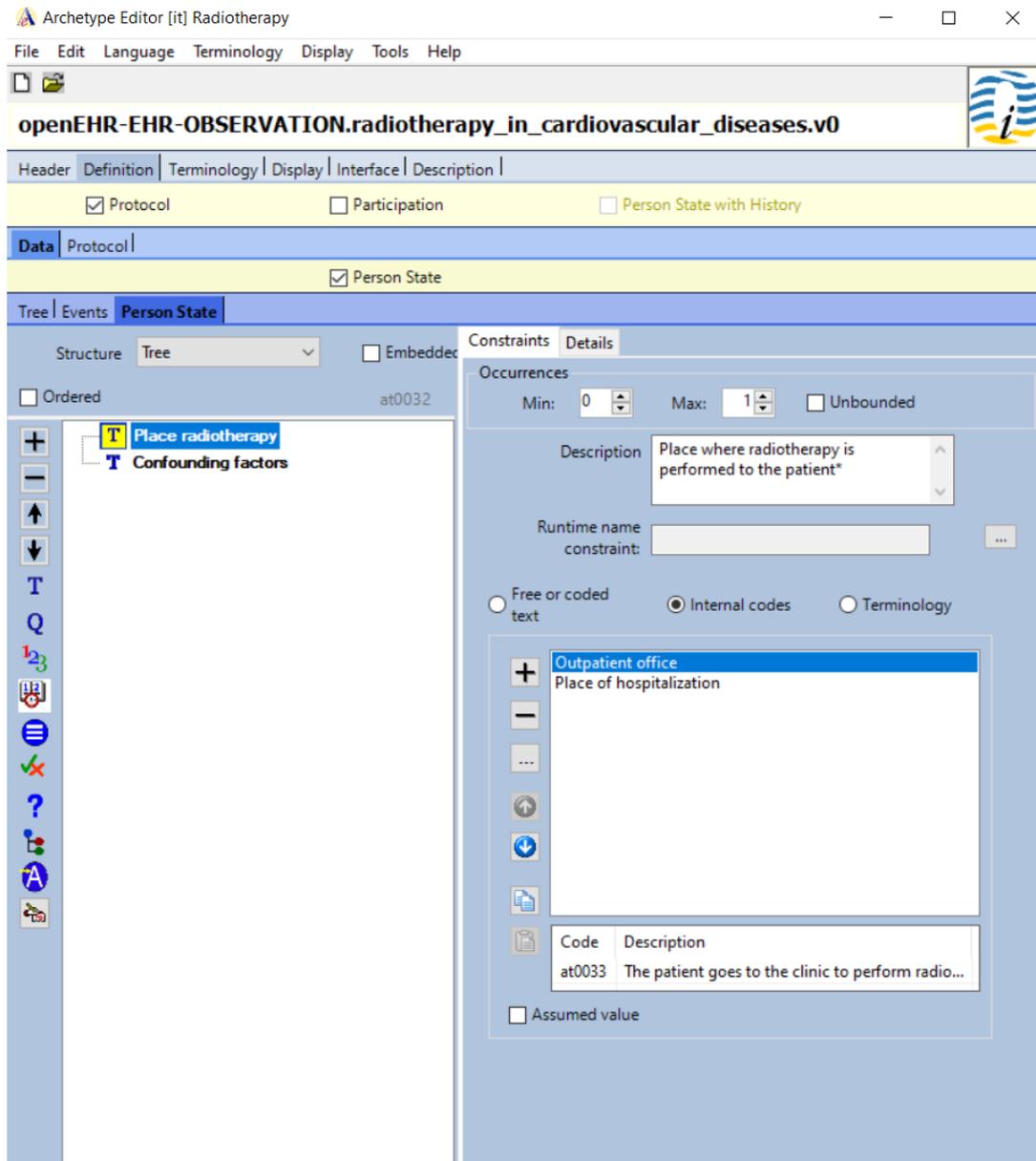


Figure 3.10 [Class State, containing data object information]

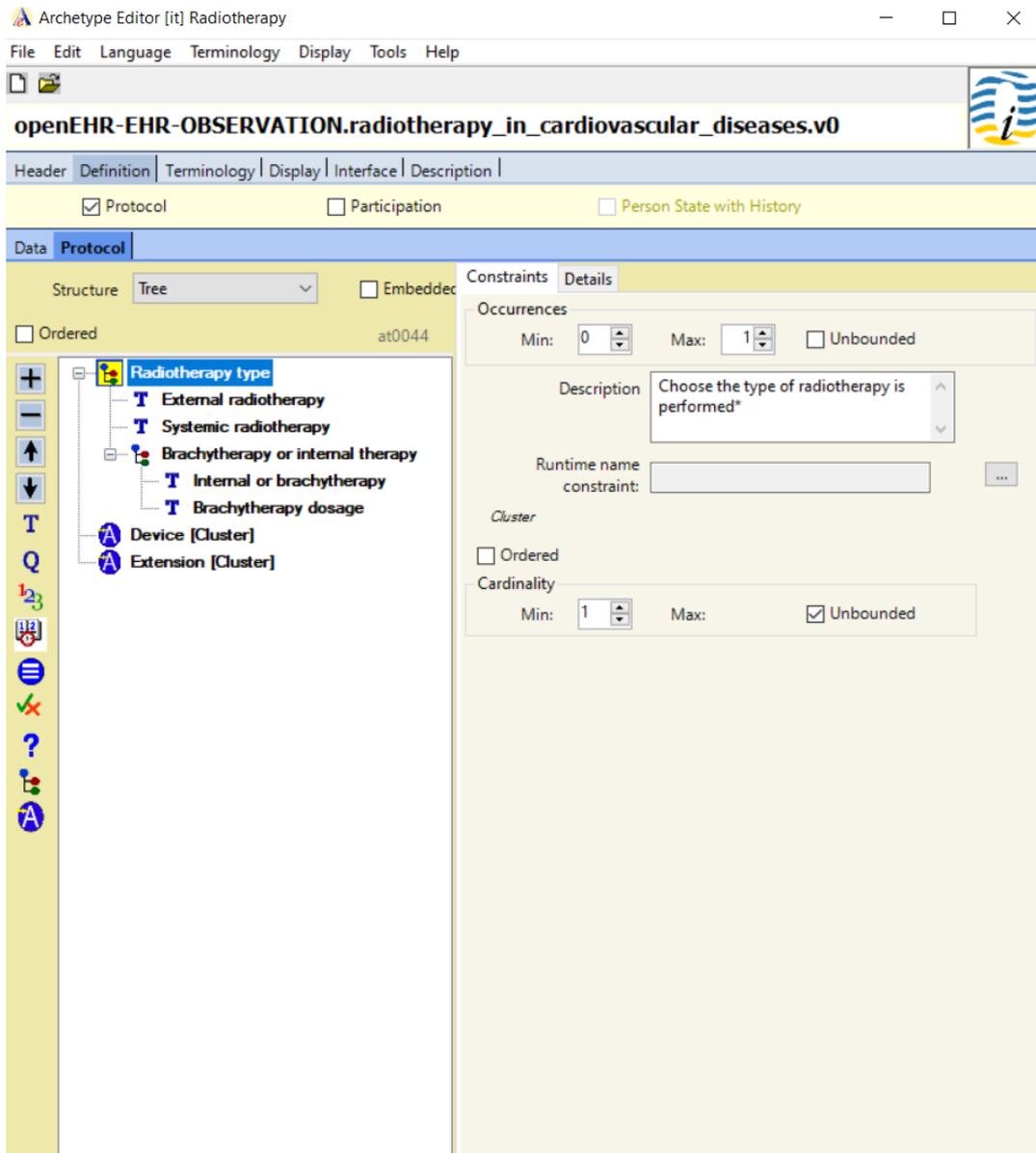


Figure 3.11 [Protocol Class, contains information on how the information was collected or measured]

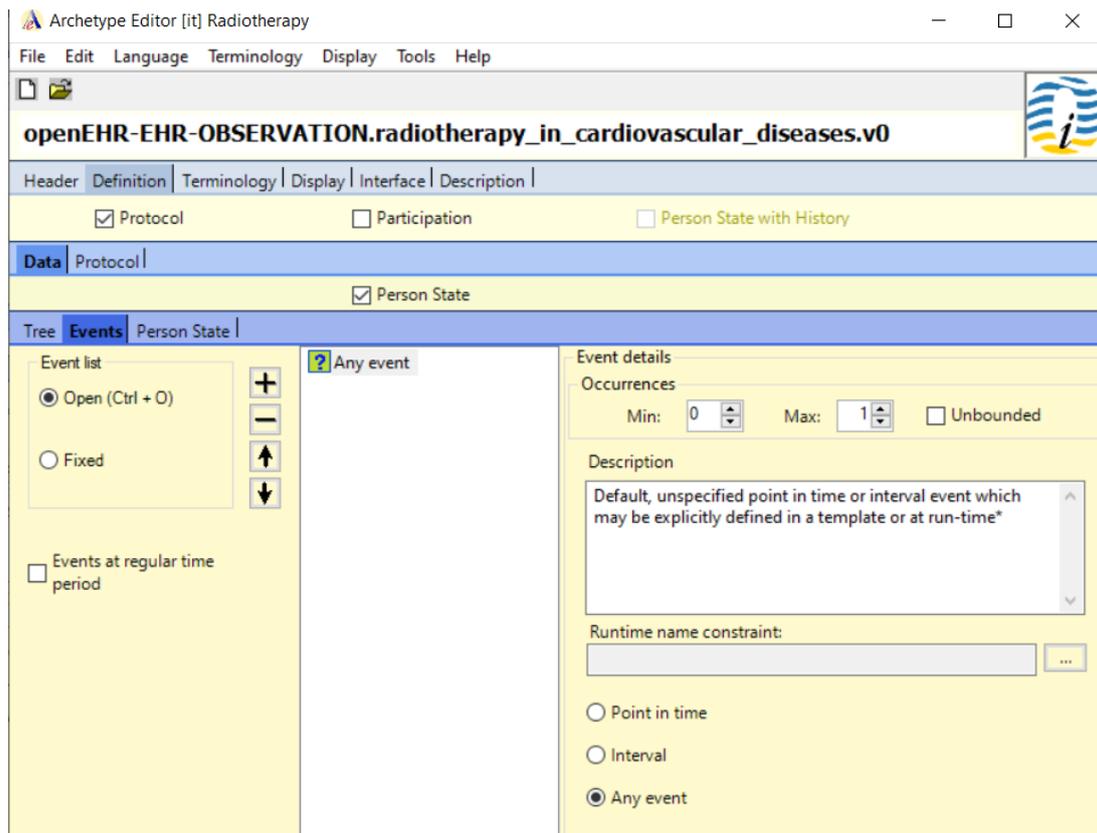


Figure 3.12 [Class History]

After completing the development of the radiotherapy archetype, it was uploaded to the CKM site so that it could begin the process of review, approval and then publication by the foundation. Immediately after the archetype was uploaded, a collaboration began with the CKM's founders/resourcers to understand what changes needed to be made to improve the archetype in order to make it as general as possible and to make it reusable by anyone, in an effort to ensure interoperability.

### 3.3.3 Archetype on diabetes

Diabetes includes a group of disorders caused by hyperglycemia (according to the definition currently used in the United States, fasting glycemia must be greater than or equal to 126 mg/dl or greater than or equal to 200 mg/dl 2 hours after an oral loading of 75 g of glucose, on 2 or more occasions). The risk factors, mentioned in the introduction, interact widely in causal or concausal chains. For example,

body weight control (through nutrition and physical activity), as well as being fundamental for the reduction of cardiovascular risk, is the cornerstone of primary prevention of diabetes.

The effective interventions for the prevention of cardiovascular disease and diabetes complications to be implemented have been summarized and are:

- Physical activity: studies have shown that moderate or intense physical activity reduces coronary heart disease and stroke;
  - Diet: moderate consumption of fruit and vegetables reduces ischemic cardiovascular disease;
  - Tobacco smoking: studies have found an association between smoking and global mortality and ischemic cardiovascular disease;
  - Interventions aimed at reducing blood pressure: with physical activity and an appropriate lifestyle, it has been shown that blood pressure can be reduced. However, there is no evidence that these interventions can reduce mortality and morbidity;
  - Interventions aimed at reducing cholesterol levels: lowering cholesterol levels in asymptomatic subjects reduces the risk of cardiovascular events;
  - Interventions aimed at reducing blood sugar levels in diabetics: controlling blood sugar levels with insulin and/or oral hypoglycemic agents can reduce the risk of cardiovascular disease.
- [20]

The following steps were taken to create a new archetype of risk factors for diabetes (see Figures 3.13, 3.14, 3.15 and 3.16):

- The Archetype Editor software has been used;
- It was decided to create an archetype Evaluation;
- The archetype was divided into the following classes:
  - Class Data, contains basic information:
    - Diabetes type: type 1, type 2, gestational, impaired fasting glucose, impaired glucose tolerance;
    - Start date;
    - Glicemic control:
      - Fasting glycemia [mg/dl];
      - HbA1c [%];
      - HbA1c date;
      - Random plasma glucose [mg/dl];
      - 2 hours plasma glucose on OGTT [mg/dl];

- Glycemic target achieved.
- Microvascular complications monitoring:
  - Nephropathy:
    - Nephropathy (True/False);
    - Stage;
    - eGFR [ml/min/1.73m<sup>2</sup>];
    - microalbuminuria [mg/24h];
    - stage;
    - other.
  - Retinopathy:
    - Retinopathy (True/False);
    - Fundus oculi finding;
    - Other.
  - Neuropathy:
    - Neuropathy (True/False);
    - History and physical examination;
    - Nerve conduction and autonomic;
    - Other.
- Microvascular complications monitoring:
  - Blood pressure [mmHg];
  - Carotid/femoral ultrasonography;
  - ECG finding;
  - LDL-Cholesterol [mg/dl];
  - Other.
- Therapy:
  - Insulin (True/False).
- Class Protocol, which contains information on how the information was collected or measured, in this particular case we refer to the "*type of instruments*":
  - glucometers;
  - other.

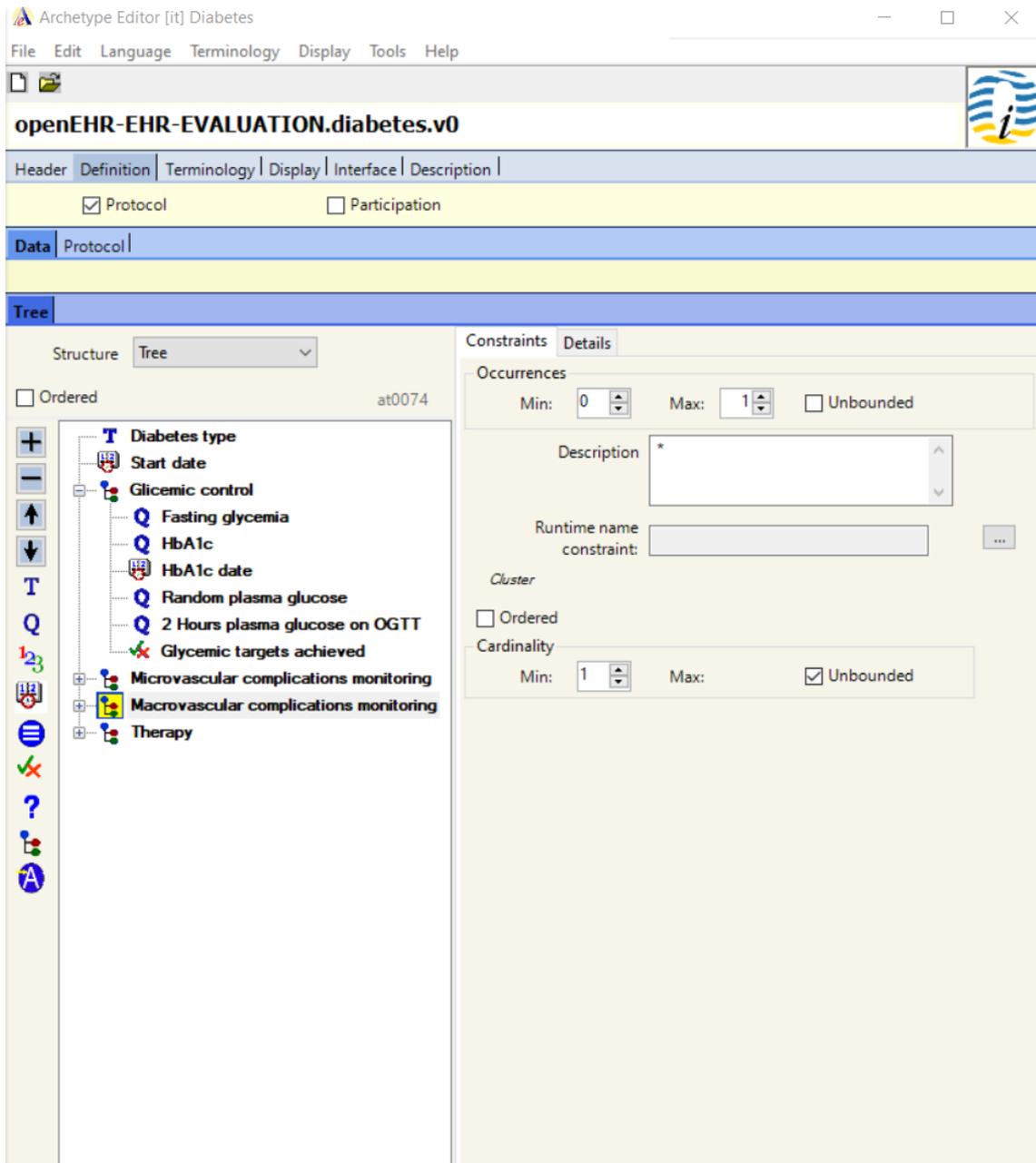


Figure 3.13[ date class of diabetes archetype]

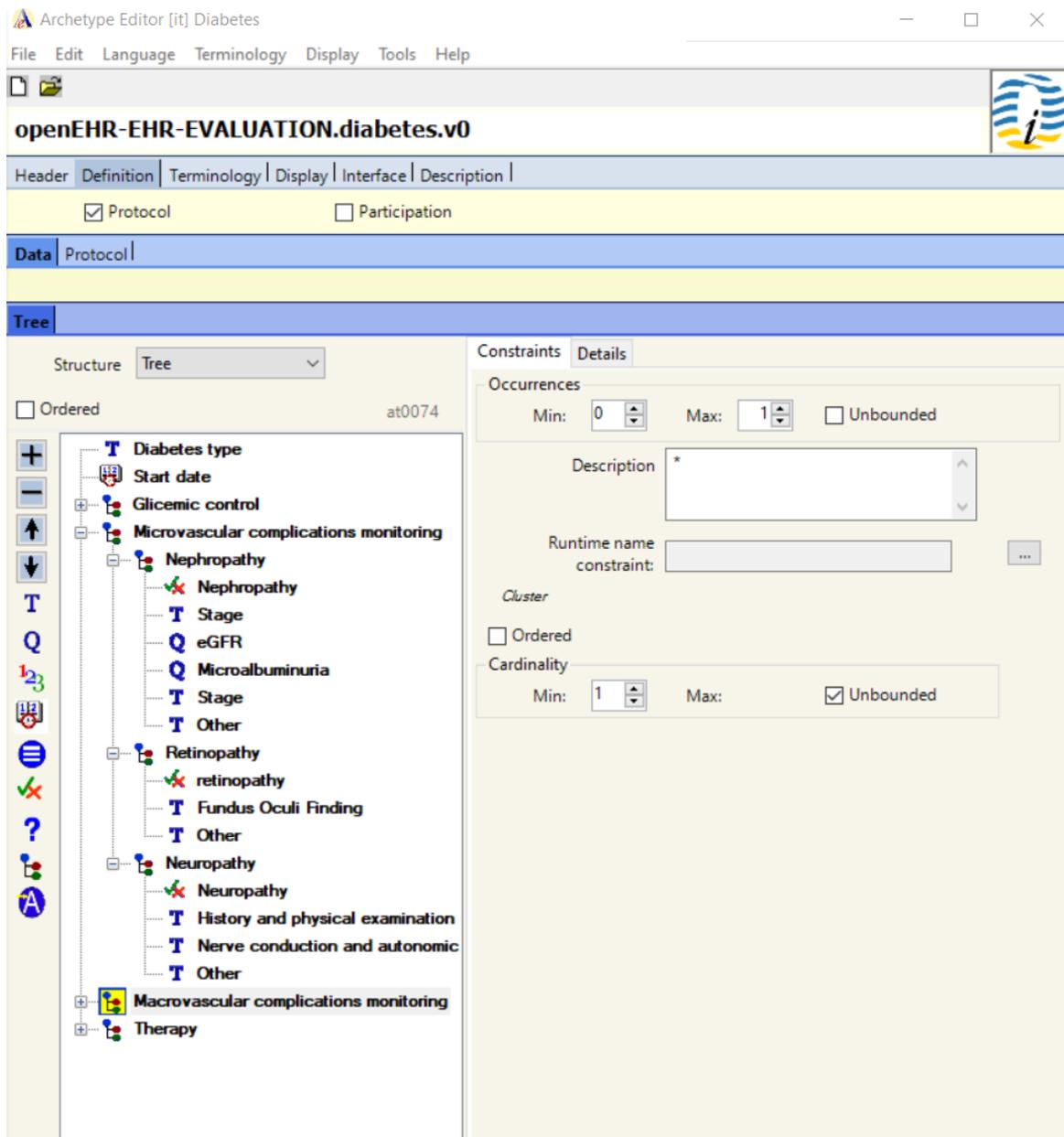


Figure 3.14 [details of the date class of the diabetes archetype]

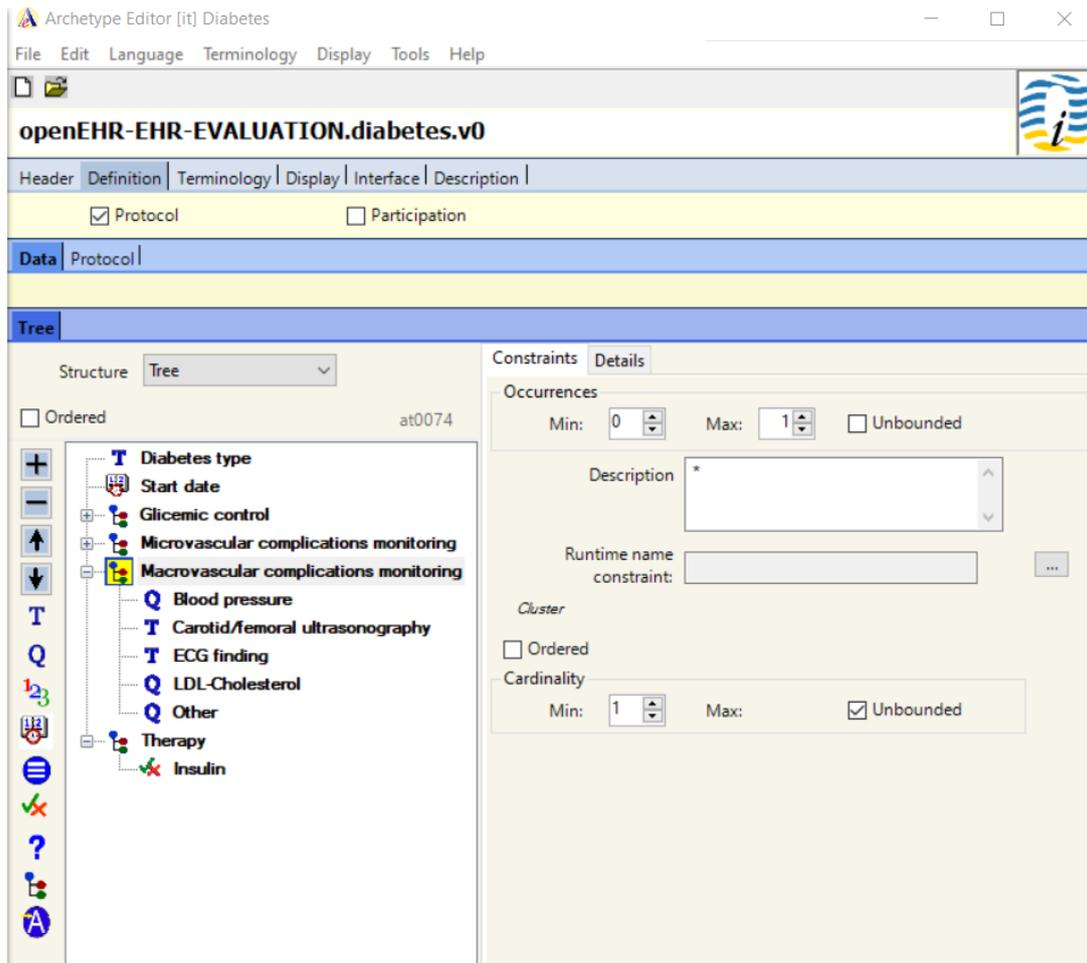


Figure 3.15 [details of the date class of the diabetes archetype]

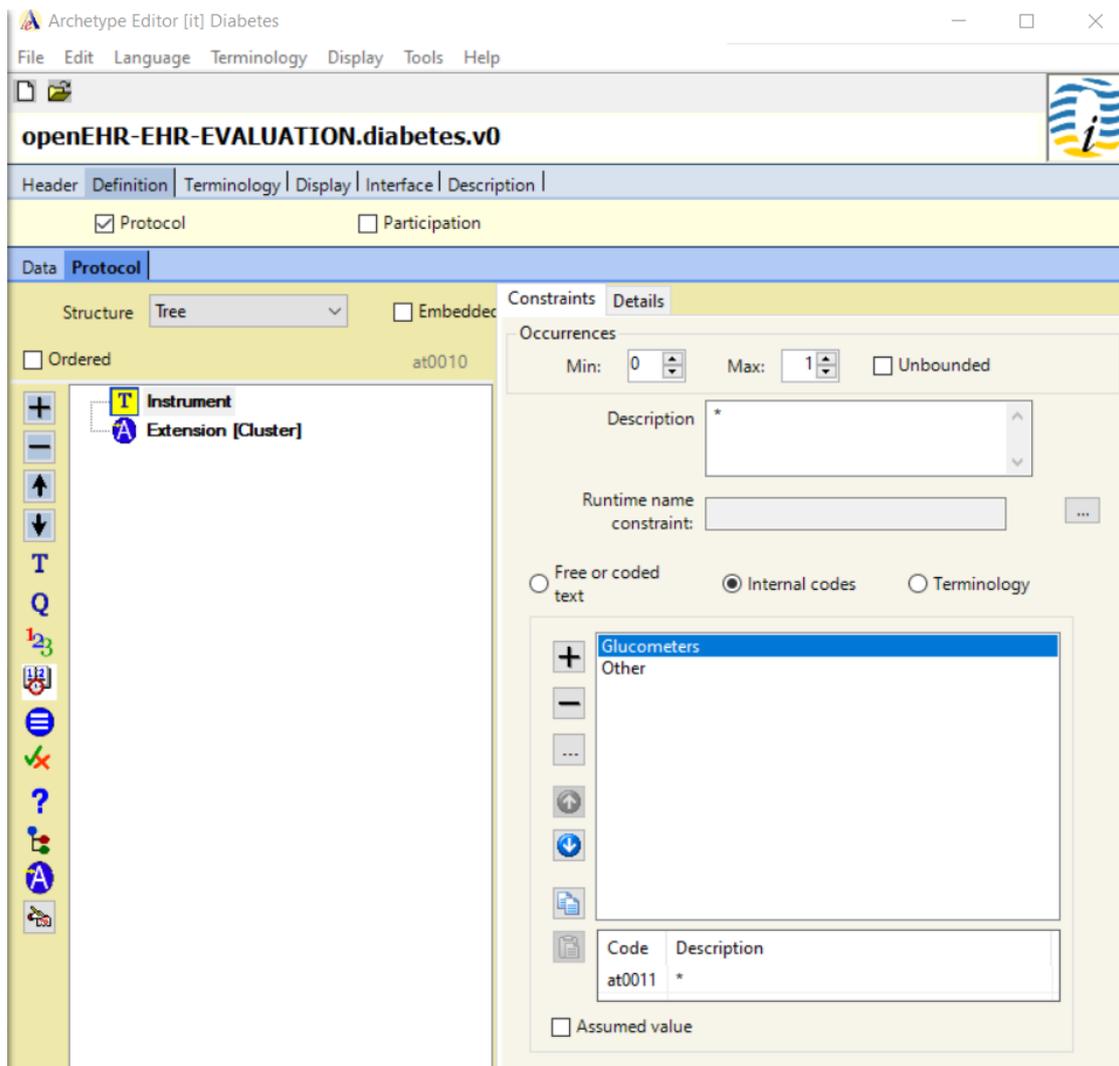


Figure 3.16 [protocol class, contains information on how the information was collected or measured]

After the development of the archetype the goal is always to propose it to the openEHR community. This, however, will be done later, as with the CKM managers at the moment the focus has been on the first two created archetypes (previous chemotherapy and radiotherapy, especially on the first of these) so as to be able to focus on one archetype at a time, to define in detail and improve the data models so that they can be published and reused by more realities.

In addition, in this thesis work, in collaboration with the doctors of the San Giovanni Molinette hospital and of A.O.U. San Luigi Gonzaga, work also started on two more archetypes, related to the risk factors reported in the CardioFilo application, in particular the archetypes on hypertension and dyslipidemia are being developed and these too will later be proposed to the openEHR community in order to continue the collaboration for a possible publication of the data models that have been designed.

### **3.4 Problems found in the development of clinical data**

In the development of clinical data model, it was not always possible to find the archetypes of interest for the CardioFilo project from CKM, so it was necessary to move to the creation of new data models that meet the requirements for risk factors related to cardiovascular disease.

The design of archetypes, according to the openEHR standard, is a complex process that needs to involve many stakeholders. It is precisely in this context that the figure of the biomedical engineer is placed, who stands between the technical professional figure, fundamental for the generation of the system architecture, and the healthcare personnel who provide the clinical knowledge useful to develop the archetypes to be included in the medical records.

The latter being a complete set of data, they should be reusable as much as possible in order to ensure the inclusion of all possible cases of use.

The work carried out has presented some difficulties both from a technical and clinical point of view. In particular with regard to the first aspect, the main criticalities are emerged in the use of the two softwares necessary for the creation of archetypes and templates: they are open source tools with a 'drag and drop' user interface so that the response of the tools was not always immediate and sometimes it was necessary to restart before proceeding to the desired operations. Another criticality, which however concerns only the "Archetype Editor" software, concerns the units of measurement that were often difficult to identify within the software itself.

Finally, the greatest difficulty was to try to develop data models that were suitable for the clinical case of interest, but that at the same time respected the openEHR standard in order to guarantee integration and interoperability. To do this, an attempt was made to create a "clinical-technical" team that could interact in an appropriate way in order to design archetypes functional and reusable. [21]

## Chapter 4

### GDPR/Legal Issue

In the field of e-Health, careful research carried out on the priorities set by the Data Protection Authority (DPA) in this area, reveals that of ensuring the maximum protection of patients' health data while encouraging the development of new technologies in personal care. This means that all health care facilities that hold personal data processing will have to place themselves in the same guarantee perspective focused on the rights of the data subject, continuously balancing the benefits related to the use of technologies with the risks potentially detrimental to the rights and freedoms of patients.

In this regard, it is suggested, first of all, to refer to the rules and prescriptions dictated over the years by the Data Protection Authority within specific sectors or issues placed under the magnifying glass of the Guarantor Authority as they have a potentially high level of risk.

In particular, reference is made to the following areas:

- electronic health file;
- electronic health dossier;
- online reporting;
- remote monitoring of cardiac defibrillator patients;
- booking of specialist visits (e.g. to the unified reservation center and pharmacies);
- interconnection of databases;
- IoT.

The primary objective of the Data Protection Authority's prescriptions, in perfect accordance with the provisions of the Regulation, is to ensure that the healthcare facilities that collect, use and store data do so by having the entity of the risk under control for each processing activity carried out in an e-

Health context, which is preventively assessed in order to provide for the activity in safety from the point of view of data protection.

The overview that emerges from the 2017 Security Report of the Clusit tends to confirm a scenario of constant and daily "red alert" with regard to cybersecurity issues in the health sector, which assumes a general tendency to further worsen if the phenomenon is not countered with great determination.

This is unavoidable, due to the constant increase in the overall attack surface exposed by our digital society: not only for the increasingly widespread "smart working", realized through a mix of mobile, cloud and social tools, often used in a promiscuous and insecure way (i.e. mixing personal digital life with work) but also for the impetuous diffusion of IoT devices, typically lacking the most elementary security measures, not only in the consumer sector but also in production contexts or for critical applications (for example in e-health or smart-city).

These phenomena correspond to the growing aggressiveness of the attackers, who, taking advantage of the numerous vulnerabilities of the system (of a cultural, organizational and technological nature), achieve very high illicit profits against risks which, unfortunately, are still practically non-existent, given the great objective difficulty in pursuing these conducts with the regulatory instruments and resources available, despite the excellent efforts of the authorities in charge.

The Clusit 2017 Report shows that Healthcare, at international level, represents the sector that has suffered the greatest increase in cyber attacks during 2016 compared to 2015 (+102%).

The Cyber Security Report 2014 of La Sapienza University of Rome<sup>2</sup> showed how, within the PA (Public Administration), Healthcare is a sector that is poorly prepared to deal with the issue of security, both in its aspects of awareness and in the planning of organizational and technological solutions. Moreover, Healthcare tends to be one of the PA sectors in which, being complex systems, it is assumed that not only data breach problems are on the agenda, but also problems related to operational continuity, backup, unauthorized use of data. Finally, it is reported that two thirds of the sample of Local Health Authorities and Hospital Organizations surveyed do not carry out Risk Management activities. [22]

The adaptation to the GDPR (General Data Protection Regulation) EU Regulation 2016/679 and to the AgID (Agenzia per l'Italia Digitale) guidelines on minimum ICT security measures for the PA (Circular n°1 of 17.3.2017 published in OJ of 4.4.2017) is therefore a concrete opportunity to improve the quality and safety of ICT services in Healthcare, to protect both citizens who use social and health

services; and professionals who provide them using ICT as a prerequisite for the ordinary activities that are carried out in Healthcare Agencies and Hospitals, as well as ICT professionals who operate in Healthcare, whether they are operators of Healthcare Agencies or Technology Solution Providers.

## 4.1 GDPR key principles

EU Regulation 679/2016 on the protection of personal data ( in the following referred to as GDPR) entered into effect on 24 May 2016, and will become directly applicable from 25 May 2018 (deadline for adaptation), repealing Directive 95/46/EC.

The GDPR completely overturns the perspective of privacy regulation by establishing a regulatory framework focused on the duties and accountability of the Data Controller. The new discipline requires this entity to ensure compliance with the principles it contains, but also to be able to prove it, by adopting a series of tools that the GDPR itself indicates, starting from a careful assessment of risks and impacts, with a planning from the outset of a series of activities that may involve cultural, organizational and technological changes, as well as significant investments of an economic nature.

The concept of " responsibility" translates into the fact that the owner is called to demonstrate that the treatments are consistent with the provisions of the GDPR, to plan and implement technical and organizational measures to prove their adequacy, and to activate a monitoring model of the technical and organizational measures implemented.

In this logic two key assumptions of the GDPR system are introduced:

- the "Privacy by design", therefore the need to design the Security and Privacy measures already in the design phase of the information systems;
- the "Privacy by default", i.e. the ability to design Security and Privacy measures by default, as a prerequisite for the normal functioning of company information systems. (art. 25)

In addition, the principles (art.5) of lawfulness of the processing are reiterated, which can only be possible if the data subject has expressed explicit consent (which the Owner must demonstrate to have collected, art.7), of adequacy, relevance and not excessive data with respect to the purposes for which they are processed.

Following the statement of the key principles, special attention is paid to the Rights of the data subject regulated in a specific Chapter of the GDPR (Chapter III):

- *Information on data processing* (art.12): it must be made in a concise, transparent, intelligible and easily understandable form and where attention is paid to the need to provide precise indications (art.13) on the purpose of processing, any recipients / users of data, the period of retention of data, how to request correction or deletion of the same;
- *Access to the data* by the interested party (art.15) which provides in paragraph 3 the possibility for the interested part to receive a copy of the processed data;
- *Rectification and deletion of data*: the right to rectification (art.16), the right to be forgotten (art.17) and the right to limitation of data processing (art.18) with the obligation to notify the data subject in case of rectification, cancellation or limitation (art.19);
- *Data portability*: the interested party has the right to receive in a structured format, in common use and readable by automatic device, the personal data concerning him/her and has the right to pass this data to another Data Controller (art.20);
- *Right of opposition*: the data subject has the right to object to the processing of data concerning him/her at any time (art. 21) and the right not to be subject to an automated decision-making process, including profiling (art.22). [22]

## 4.2 Possible reference model

The complexity of the interventions to be carried out in the area of digital innovation in Healthcare requires the evaluation, during the planning phase (privacy by design) of the company information system, of two basic considerations:

- The concept of "absolute security" does not exist: any system is vulnerable. Securing a system means planning a set of procedures and tools to reduce risks as much as possible or to acceptable levels (Pfleeger, 2004, Cinotti, 2006);
- The planned interventions are of a cultural, organisational, technological and economic nature. These interventions in the field of digital innovation, security and privacy do not constitute "limitations" to an extended and pervasive use of ICT in Healthcare, but represent actions of qualification and improvement of the company information system.

A possible governance and management model of the actions envisaged by the GDPR in Healthcare is represented in Figure 4.1, which tends to identify five areas of activity that will be dealt with in greater detail in the following paragraphs.

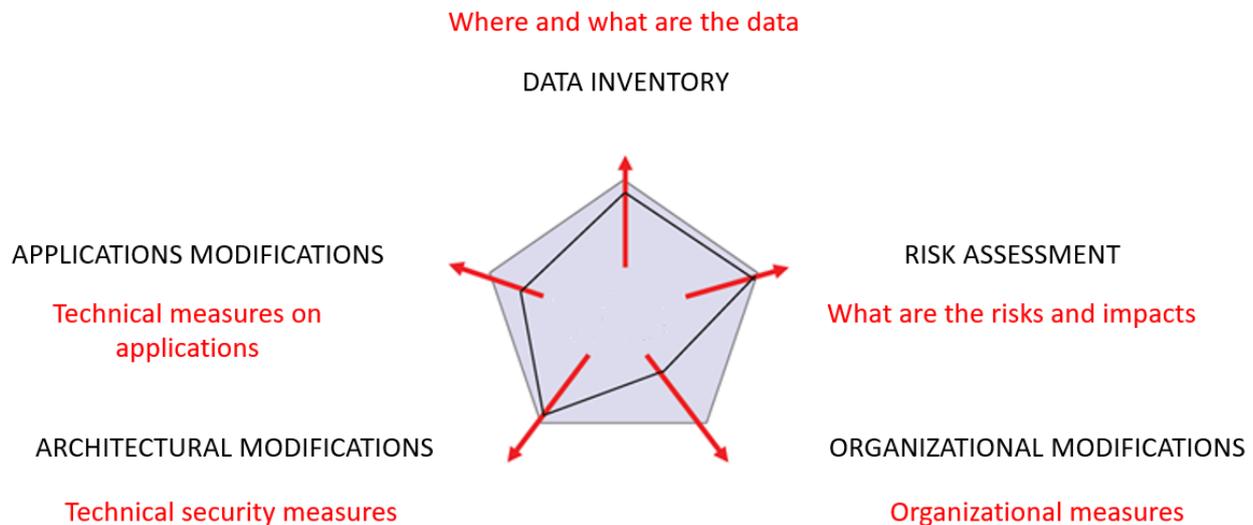


Figure 4.1 [governance and management model of the actions envisaged by the GDPR in Healthcare]

## 4.2.1 Data Inventory

The objective of the first area of activity is aimed at the knowledge of the data processed. It is to be aware of the active treatments and therefore to know for each type of treatment/data treated at least the following information:

- Purpose of the treatment;
- Description of the use data;
- Type of processed data;
- The categories of recipients (to whom the data will be communicated);
- Where is the database (location in the server farm/in the cloud etc);
- Limits for deletion/loss of data;
- Description of security measures (organizational and technological);

- Description of backup and restore measures.

For this purpose, GDPR provides for the creation of a "processing register" (art.30) which can also be written in electronic format (i.e. digital or paper).

Considering that it is difficult to find an application map or a list of all the applications in Healthcare Companies, in the logic of the first application, it is sufficient to draw up a list (xls file) containing the above mentioned information for each company macro-process supported by ICT (see Figure 4.2). [22]

Treatment register					
Functional macroprocesses	Workflow description	Treatment purpose	Type and description of data treated	Possible users	Descriptions security measures
Registry	Management of basic coding and Master Patient Index.				
Customer workflow management	Experimenting with a CWM system that allows you to trace the patient's path in PS by providing indications even the chaperones. Needed analysis and revision organization of the reception area which appears fragmented.				
Clinic Chemical Laboratory	Laboratory standard flow with order entry phase and reception of reports and structured departmental data. To activate microbiology automation. For external users, direct acceptance with unified booking system and paper reports preparation				
Operating Block	Standard operating block management flow with recording of operating times and production of the hall report surgical				
Document Management	I activate a document management system				

Figure 4.2 [Non-exhaustive example of a list of typical treatments of a Healthcare Provider]

## 4.2.2 Risk and Impact Analysis

Having taken note of the data processed in the various company macro-processes, in the second area of activity to be implemented, the problem arises of assessing the risks associated with the processing and an assessment of the impacts related to data protection.

Also in this case, in the logic of the first application (in the following of the document more articulated approaches are defined) it is suggested the construction of a risk map that allows to estimate, for each type of processing, a generic risk index.

In this context it is suggested the construction of a risk matrix that could be articulated in:

- external risks related to the provider of the technological solution;

- internal risks related to the data processing processes in the specific process;
- technological risks linked to the security of the technological infrastructure;
- management risks linked to the organization of the company ICT team.

For each type of risk, verification items are identified and assigned a risk score. This allows to define a generic risk for each operational process supported by ICT.

Subsequently, an impact assessment can be carried out for each operational process supported by ICT. In this regard, AgID has produced a map with three types of impacts: organizational impacts, service impacts, technological impacts (see Figure 4.3). Also in this case, verification items are defined by type of impact to which a score is attributed. This allows to define an impact index for each operational process supported by ICT. In this specific case, however, particular consideration must be given to impacts on people's rights and freedoms, which are not simply attributable to service impacts.

Structure Impacts								
low 1, medium 2, high 3, very high 4, critical 5	Macroprocess	Registry	Customer Workflow	Reservation and billing	Imaging	Operating block	Passive active cycle	Document Management
Type Impact								
Service Impacts	Importance of the service for business purposes	4	3	4	3	4	3	2
	the interruption results in an immediate impact/disagio to users	4	4	4	4	4	2	2
	Type and volume of users involved	4	3	4	4	3	2	2
	It's possible recover ex post the data not acquired	4	1	4	2	2	2	1
	alternative procedures are possible	4	1	4	2	2	2	1
	level of damage to the Company	4	3	4	3	3	2	2
	Total	24	15	24	18	18	13	10
Organizational Impacts	Number of locations involved	3	2	3	4	4	1	2
	Number of staff involved	3	2	3	4	4	2	2
	Interruption causes the process to stop	4	1	4	2	3	3	2
	Interruption leads to impacts on others interdependent processes/systems	4	2	4	3	3	3	2
	Total	14	7	14	13	14	9	8

Figure 4.3 [Non-exhaustive example of a list of typical treatments of a Healthcare Provider]

The sum of the risk index and the impact index determine a Severity Index for each operating process supported by ICT.

The Severity index multiplied by a probability of occurrence of the event (related to the measures already adopted) and a vulnerability index (typical of the complexity of the specific process) determines a real Risk index for each operational process supported by ICT, assessed overall in its aspects of generic risk, probability and vulnerability impacts.

This type of approach also allows to use the results obtained for the Business Continuity Plan in line with AgID indications.[22]

### 4.2.3 Organisational measures

The objective of the activities in this area is mainly to verify:

- Definition of the Privacy Organigram;
- Appointment of Data Protection Officer;
- Definition Ownership/Co-ownership (Pdta);
- Appointment of the Authorized to the treatment (already appointed pursuant to Legislative Decree 196/2003) in the light of the Recommendation expressed by the Guarantor Authority;
- Documentary management of Security and Privacy interventions (Organisational and Technical documentation);
- Notification of Data Breach.

There are no particular complications regarding the definition of the privacy organization chart, which tends to overlap with that of Legislative Decree 196/03, and consequently the appointment of the Owner, the Managers and, although not expressly provided for by GDPR, the Authorized to process personal data on behalf of the Owner or Managers. In the latter case, the Data Protection Authority has expressed the opportunity for data controllers and data processors to maintain the organizational structure and the procedures for appointing data processors as outlined over the years, including through the interventions of the Guarantor itself. In this context, it is also worth mentioning the opportunity, present in art.28 paragraph 6, to use "standard contracts" for the appointments of personnel authorized to the treatment if they carry out homogeneous activities (e.g. service technician, nursing staff, medical staff, etc..) simplifying the documental management of privacy.

The issue of the data controller (art.26) is a particularly interesting issue in Healthcare because it simplifies the regulatory framework with respect to treatments that can be carried out by several data controllers on a specific treatment process. It mainly refers to the theme of the use of new organizational models, of a transversal type (hospital-territory-domicile), called PDTA, for the treatment of chronicity. Frequently these new organizational models involve treatment carried out by

several professionals in different places and times but also belonging to different legal entities. The GDPR provides for the possibility of contractualising co-ownership, defining mutual responsibilities and providing clear and transparent information to citizens (as set out in articles 13 and 14).

A second interesting topic concerns the appointment and role of the DPO (Art. 37, 38, 39). The DPO is a new figure who must have specific knowledge of data protection legislation (on privacy and security) and practice. He or she may be an employee or an external consultant/company, who must be involved in all data protection activities. He or she carries out consultancy tasks for the Data Controller and the Data Processors, supervises compliance with the activities carried out to comply with GDPR, gives advice to the Supervisory Authority, representing the point of contact between the latter and the Company of reference.

A third issue of interest is the mandatory notification of the data breach. The notification, in case of violation of personal data, is mandatory within 72 hours to the Guarantor Authority (art. 33) but also mandatory (art. 34) against the person concerned.

The last issue of a certain interest also for Healthcare is the creation (or use if existing) of a document management system for all the documentation produced on data protection for exhibition purposes to third parties aimed at demonstrating in an objective and transparent way the activities implemented for compliance with the GDPR, in line with the principle of accountability.[22]

#### **4.2.4 Technical safety measures**

This involves implementing and describing the technological measures put in place in order to ensure an adequate level of risk with regard to the rights and freedoms of individuals deriving from the processing operations carried out, including those relating to confidentiality, integrity and availability of data (art. 32) for each operational process supported by ICT. In this regard, it is suggested to adopt a global (multidimensional) and systemic approach considering that it refers in particular to four areas of activity that impact on business processes and require Cultural, Organizational, Technological, Economic interventions.

In particular, reference is made to the unified management of the identity and access profiles of data users (IAM) and the tracking of their activities, the ability to ensure on a permanent basis the confidentiality, integrity, availability and resilience of processing systems and services; the ability to detect events and incidents related to data breach management, the ability to promptly restore the availability and access of personal data in the event of a physical or technical incident; the implementation of a procedure/system to regularly test, verify and evaluate the effectiveness of technical and organisational measures to ensure the security of the processing. [22]

#### **4.2.5 Technical Measures on Applications**

Among the application measures to be put in place, particular attention must be focused on:

- the integration of the application systems with the identity and access profile management system;
- pseudonymisation and encryption of personal data (Art. 32);
- the unified management of information, consent, correction and deletion (art. 7 to 10 and 15 to 19). With regard to the latter point, the analysis of the application map must make it possible to specify which procedures acquire the citizen's consent/restrictions and how they make this attribute transparent to all the applications involved in the individual care process in order to ensure that personal data are consulted only by the staff actually involved in a specific care process and only for the duration of the care process. [23]

### **4.3 Scope of the GDPR**

EU Regulation 2016/679 has as its object the protection of individuals with regard to the processing of personal data and the free movement of data within the European Union which cannot be restricted or prohibited.

Article 4(1) of the Regulation defines "personal data" as "any information concerning an identified or identifiable natural person", including personal data subject to pseudonymisation techniques.

Anonymous data, on the other hand, is excluded in its entirety from the scope of application of the Regulation.

With regard to "processing", it is defined in art. 4, par. 2 of the Regulation as "any operation or set of operations, performed with or without the aid of automated processes and applied to personal data or sets of personal data, such as collection, recording, organisation, structuring, storage, adaptation or modification, retrieval, consultation, use, communication by transmission, dissemination or any other form of making available, comparison or interconnection, restriction, erasure or destruction".

According to art. 3, the Regulation applies to Data Controllers and Data Processors established in a Member State of the European Union, or in a place subject to the law of a Member State, regardless of whether the processing is carried out in the Union or not, if in any case it falls within the scope of the activities of an establishment of a Data Controller or Data Processor in the Union. The Regulation also applies to Data Controllers or Data Processors not established in the European Union if they process personal data of data subjects located in the European Union, when the processing activities specifically concern them:

- the offer of goods or the provision of services to such data subjects in the Union, regardless of whether payment by the data subject is mandatory;
- the monitoring of their behaviour insofar as such behaviour takes place within the Union.

## **4.4 Sanctions**

Just a few lines to remind that the unlawful processing of data or loss of data provides both criminal liability for failure to take minimum security measures, and civil liability because the failure to take appropriate measures determines an obligation to pay compensation under Article 2050 of the Civil Code and under Article 15.del D.Lgs.196/03 and administrative liability. With regard to the latter, the new GDPR (art.83) provides for administrative fines of up to 10 million euros and for companies equal to 2% of the worldwide group turnover in case of violation of the obligations of the Data Controller (art. 8.11 from 25 to 39, 42, 43) and up to 20 million euros and for companies equal to 4% of the worldwide group turnover in case of violation of the conditions of consent (art. 5,6,7,9) to respect the rights of the data subject (art. 12 to 22) and data transfers from one Data Controller to another of third countries (art. 44 to 49). [22]

## 4.5 The GDPR on health data

Article 4 n.15 GDPR provides that health data: *“are personal data relating to the physical or mental health of a natural person, including the provision of health care services, which reveal information concerning his or her state of health”*. More specifically, Recital 35 GDPR provides that *“Personal data relating to health should include all data concerning the health status of the data subject which reveal information related to the data subject's past, present or future physical or mental health status. This includes [...] any information concerning, for example, a disease, disability, risk of disease, medical history, clinical treatment or physiological or biomedical status of the data subject”*.

Well, the European legislature has paid specific attention to health-related data and has included them among the 'special categories of personal data' whose processing is in principle prohibited under Article 9(1) GDPR. In par. 2, the same Regulation provides for a series of exceptions to the prohibition of the treatment of the data relative to health that, in the health field, the Data Protection Authority has generally led back to the cases that:

*“g) the processing is necessary for reasons of public interest relevant on the basis of Union or Member State law”* (art. 9, par. 2, letter g) GDPR) identified by art. 2-sexies of the Privacy Code;

*“i) treatment is necessary for reasons of public interest in the field of public health, such as protection against serious cross-border threats to health or ensuring high standards of quality and safety of healthcare and medicinal products and medical devices, based on Union or Member State law providing for appropriate and specific measures to protect the rights and freedoms of the data subject, in particular professional secrecy”*. This is the case, for example, of health emergencies resulting from earthquakes and food safety;

*“h) treatment is necessary for preventive medicine or occupational medicine purposes, assessment of the employee's ability to work, diagnosis, health or social care or treatment or management of health or social systems and services [hereinafter "treatment purpose"] on the basis of Union or Member State law or in accordance with a contract with a health professional”*. In particular, it should be specified that for the purposes of treatment, health data may only be processed by a professional subject to professional secrecy, or under his responsibility, or by persons in any case subject to the obligation of secrecy (art. 9, par. 3, GDPR and Cons. 53; art. 75 of the Privacy Code).

On this point, it is very important to point out that the Guarantor has specified that the health professional, subject to professional secrecy, must no longer require the patient's consent for the treatment necessary for the healthcare service, regardless of whether he or she works as a freelance or as an employee, or whether he or she works within a public or private healthcare facility. This regulatory evolution follows a new concept of privacy well represented by art. 1, par. 1, GDPR which, in addition to protection, provides for the free movement of personal data.

In this perspective, the Guarantor has specified that where the processing of data relating to health is not strictly necessary for the purpose of treatment referred to in letter h) of Article 9, paragraph 2, even if carried out by health professionals, the owner will have to identify another legal basis or possibly the consent of the person concerned (Articles 6 and 9, paragraph 2, GDPR). In this regard, the Guarantor has deemed it appropriate to identify some hypotheses that do not fall under processing based on necessary purposes and which consequently require the consent of the data subject, including:

- Treatments related to the use of medical apps, through which data (including medical data) of the person concerned are collected for purposes other than telemedicine or when, regardless of the need or otherwise of the purpose of the treatment, the data may be accessed by persons other than health professionals or other subjects bound to professional secrecy (For example, it requires the explicit consent of the person concerned the treatment carried out by pharmacies for point accumulation programs. This an additional purpose, not necessary, compared to the pharmaceutical assistance traditionally carried out within the National Health Service).
- The same applies to treatments carried out in the health field by private legal entities for promotional or commercial purposes (e.g.: promotions on screening programs), or for treatments carried out by health professionals for commercial or electoral purposes.

Special consideration should be given to treatments carried out through the Electronic Health Record (d.l. 18 October 2012, n. 179, art. 12, paragraph 5). The provisions of the sector prior to the application of GDPR, compliance with which is now expressly provided for by art. 75 of the Privacy Code, provide that in such cases the condition of lawfulness of the treatment is the acquisition of consent.

Another fundamental element of the processing of personal data (art. 5, par. 1, letter a), GDPR) is related to transparency. In this context, in order to make those concerned aware, it is necessary to offer them information in a concise, transparent, intelligible and easily accessible form, with simple and clear language. The GDPR makes the data controller responsible for transparency and it is therefore up to the data controller to choose the most appropriate way to provide the information.

With particular reference to health documentation, the Guarantor points out that the legal system provides for different references for conservation times, which have not been modified by the GDPR and therefore remain fully in force. By way of example, the measure refers to the following cases:

- the doctor who carries out examinations for the issue of the certificate of fitness for competitive sports must keep the documentation regarding the examinations carried out for at least five years (art. 5, Ministerial Decree 18/02/1982);
- the medical records, together with the relative reports, must be kept for an unlimited time (Cir. Min. Sanità of 19 December 1986 n. 900 2/AG454/260);
- the radiological iconographic documentation must be kept for a period of not less than ten years (art. 4, d.m. 14 February 1997);

If, on the other hand, the retention periods for specific health documents are not laid down by a legal provision, pursuant to Article 5(1)(1)(b) of the directive, the Commission may, in the event that the retention period for specific health documents is not set out in a legal provision. e) of the GDPR, the data controller must identify this period so that the data are *"kept in a form that allows the identification of data subjects for a period of time not exceeding the achievement of the purposes for which they are processed; [...](<<limitation of storage>>)"* or, if this is not possible, indicate the criteria used to determine this period (Articles 13 and 14, also in paragraph 2, letter a, of the GDPR).

[24]

## Conclusions

The main objective of this work, at present, concerns the definition of standards that, while respecting privacy, allow the efficient exchange of information. Standardisation of clinical content could provide a long-term sustainable approach. This can lead to a national strategy where doctors cooperate with e-Health software in their preparation. This should ensure that healthcare is more efficient, timely, safe and patient-centred.

As the goal is to achieve health care standardization, it wants to ensure that an open standard like openEHR, based on data and interoperability, should be adopted as already done abroad. Indeed, this standard is an optimized platform focusing on data modeling and using over 500 complex archetypes designed to provide a maximum set of data elements. This breadth and depth inevitably leads to a fairly high level of complexity, with the possibility of extending the archetypes where required.

Whether the future is openEHR or any other standard, the fundamental requirement is that they should be open standards, so that data can be shared, protected and appropriately regulated by patient consent. The use of proprietary standards is currently the biggest problem facing the Health Service in terms of interoperability. Although the trend is changing, there are still too many 'closed' providers that are making interoperability too difficult and stifling technological innovation around the provision of digital medical solutions that can change and improve the present status.

In conclusion, it can be said that for the development of archetypes the use of a formal methodology makes it easier to define and adopt interoperable archetypes, improves quality and allows greater re-use between different information systems and research and development projects.

# Acknowledgements

Vorrei dedicare queste ultime pagine del mio elaborato alle persone che mi hanno sostenuta e che hanno creduto in me. Vorrei dire grazie:

Alla Professoressa Monica Visintin, per avermi permesso di intraprendere questo innovativo percorso di tesi, per avermi supportato e consigliato in ogni scelta e per essere stata sempre disponibile, gentile e per avermi guidato nello svolgimento del mio lavoro.

A Paolo Doz, per avermi consentito di entrare a far parte di un ambiente a me sconosciuto, dandomi così l'opportunità di relazionarmi con il mondo del lavoro. Grazie anche a Gabriele e Sara che hanno saputo ascoltare ed interpretare le mie esigenze e mi hanno supportato, sostenuto e motivato, con loro spero di poter continuare questa collaborazione. Ringrazio anche tutto il resto della famiglia Abinsula perché con la loro presenza giornaliera hanno reso questi mesi più leggeri e sorridenti.

A Carloalberto, che mi ha aiutato a portare a termine il mio lavoro di tesi e a raggiungere così questo traguardo. Grazie a Matteo per i suoi consigli e la sua supervisione sul lavoro svolto.

Ai miei genitori, che anche da lontano sono riusciti ad essere sempre presenti. Grazie per aver fiducia in me e nelle mie capacità e grazie per il continuo sostegno che mi dimostrate.

A mia sorella per l'affetto e la pazienza e per essere il mio punto di riferimento.

A Sasindolo, il cui pensiero costante rallegra le mie giornate e illumina ogni momento buio.

A Nino, per aver condiviso le mie ansie e le mie tensioni in tutti questi anni. Grazie per la forza che hai saputo e sai infondermi nella vita ogni giorno.

Ai miei amici, che ogni giorno hanno condiviso con me gioie, sacrifici e successi, senza voltarmi mai le spalle. L'affetto e il sostegno che mi hanno dimostrato rendono questo traguardo ancora più prezioso.

# Bibliography and Sitography

- [1] Hovenga, E.J.S., Health Informatics: An Overview, IOS Press, 2010.
- [2] Eysenbach G, What is e-health? J Med Internet Res, 2001
- [3] L'unione sarda, cultura – Pag. 44, 15 ottobre 2008
- [4] <http://interoperability-definition.info/it/>
- [5] <https://www.01health.it/mercato/interoperabilita-missione-software-sanita/>
- [6] <https://www.aboutpharma.com/blog/2019/10/25/interoperabilita-una-sfida-da-vincere-per-una-vera-sanita-digitale/>
- [7] <https://www.fondazioneveronesi.it/magazine/articoli/cardiologia/sanita-digitale-italia-indietro-rispetto-al-resto-deuropa>
- [8] CENSIS – ImpresaLavoro, “*Le condizioni per lo sviluppo della Sanit`a Digitale: scenari Italia-UE a confronto*”, Roma luglio 2016.
- [9] <https://www.psbprivacyesicurezza.it/sanita-digitale-europea-innovazioni-progresso/>
- [10] COMMISSIONE EUROPEA, “*comunicazione della commissione al Parlamento Europeo, al consiglio, al comitato economico e sociale Europeo e al comitato delle regiono, relativa alla trasformazione digitale della sanità e dell’assistenza nel mercato unico digitale, alla responsabilizzazione dei cittadini e alla creazione di una società più sana*”. Bruxelles, 25.4.2018 COM (2018) 233 final.
- [11] <https://oceanhealthsystems.com/ocean-tools>
- [12] <https://www.iso.org/standard/40784.html>
- [13] [https://specifications.openehr.org/releases/BASE/latest/architecture\\_overview.html](https://specifications.openehr.org/releases/BASE/latest/architecture_overview.html)
- [14] <https://openehr.atlassian.net/wiki/spaces/healthmod/pages/2949191/Introduction+to+Archetypes+and+Archetype+classes>
- [15] [https://www.epicentro.iss.it/focus/piano\\_prevenzione/veneto/Veneto-cardiovasc%20e%20diabete.pdf](https://www.epicentro.iss.it/focus/piano_prevenzione/veneto/Veneto-cardiovasc%20e%20diabete.pdf)
- [16] <https://omowizard.wordpress.com/2014/07/19/clinical-modelling-openehr-style/>
- [17] [https://www.researchgate.net/publication/277667443\\_The\\_openEHR\\_approach](https://www.researchgate.net/publication/277667443_The_openEHR_approach)
- [18] Clinical Programme Lead, openEHR Foundation, Director of Clinical Modelling, Ocean Informatics “*The openEHR Approach to Clinical Modelling*”, Heather Leslie, luglio 2014.
- [19] “*2016 ESC Position Paper on cancer treatments and cardiovascular toxicity developed under the auspices of the ESC Committee for Practice Guidelines*”. European Heart Journal (2016) 37, 2768-2801. DOI: 10.1093/eurheartj/ehw211. 27 November 2019

[20] <https://www.epicentro.iss.it/focus/piano-prevenzione/veneto/Veneto-cardiovasc%20e%20diabete.pdf>

[21] G. E. Bente Christensen, “Evaluating model-driven development for large-scale ehrs through the penehr approach,” *International Journal of Medical Informatics*, pp. 43–54, 2 2016.

[22] Italian Cyber Security Report “*Consapevolezza della minaccia e capacità difensiva della pubblica Amministrazione*”. Research Center of Cyber Intelligence and Information Security “Sapienza” Università di Roma, Dicembre 2014.

[23] ASSOCIAZIONE ITALIANA SISTEMI INFORMATIVI IN SANITÀ, privacy e sicurezza a supporto

dell’innovazione digitale in sanità: il nuovo GDPR

[24]<https://www.agendadigitale.eu/sicurezza/privacy/protezione-dei-dati-in-sanita-tutti-i-paletti-del-garante-privacy/>