**Istituto di Ricerche sulla Popolazione
e le Politiche Sociali - CNR**

Towards a clinical trial ontology using a Concern-Oriented-Approach

Crenguta Bogdan, Daniela Luzi,
Fabrizio L. Ricci, Luca D. Serbanati

What is IRPPS?

IRPPS is an Interdisciplinary Research Institute that conducts studies on demographic and migration issues, welfare systems and social policies, on policies regarding science, technology and higher education, on the relations between science and society, as well as on the creation of, access to and dissemination of knowledge and information technology.

www.irpps.cnr.it

Towards a clinical trial ontology using a Concern-Oriented-Approach

Crenguta Bogdan, Daniela Luzi, Fabrizio L. Ricci, Luca D. Serbanati

Abstract

To reduce costs and enhance the research quality in clinical trials (CT) a more systematic approach is needed for CT automatization as well as for enhancing interoperability at various levels of the research process. To this aim a conceptual model of CTs should be developed. At the base of any modelling approach there are partitioning criteria which enable us to dominate the complexity of the modelled universe. In this report we introduce an original analysis method based on stakeholders' concerns to partition the CT conceptual domain in stakeholder-oriented sub-domains. Mental representations of stakeholders related to each concern are identified as clusters of concepts related to each other. We consider such a cluster as a semantic rationale for the associated concern. The concepts found in the semantic rationales populate the universe of discourse specific to each stakeholder and compose the stakeholder's vocabulary. Some concepts are shared with other stakeholders while others are specific to one stakeholder, some concepts are CT specific while others are medical or general concepts. In this way a concern-oriented ontology of the CT can be created. The method is illustrated for the subject selection criteria, a component of a CT project, but it can be used for any other component of the CT protocol. The taxonomy of the vocabulary of CT concepts and the relative semantic rationale net give us a valuable structure for software development especially in solutions based on a service-oriented architecture.

Key-word: Clinical trial, Selection criteria, Ontology, Concerns

Riassunto

Per ridurre i costi e migliorare la qualita' della ricerca nei trial clinici (CT) e' necessario un approccio piu' sistematico all'automazione dei CT per rinforzare l'interoperabilita' a vari livelli del processo di ricerca. Per questo scopo e' stato sviluppato un modello concettuale di CT. Alla base di ogni approccio di modellizzazione ci sono criteri di partizione che ci permettono di dominare la complessita' dell'universo da modellare.

In questo rapporto noi introduciamo un metodo originale di analisi basato sui concern degli stakeholder per partizionare il domino concettuale dei CT in sotto-domini orientati agli stakeholder. Le rappresentazioni mentali degli stakeholder relative a ciascun concern sono identificati come cluster di concetti collegati ad altri concetti. Noi consideriamo ciascun cluster come una base razionale per il relativo concern. I concetti trovati nelle basi razionali popolano l'universo del discorso specifico per ogni stakeholder e compongono il vocabolario degli stakeholder. Alcuni concetti sono condivisi con altri stakeholder, mentre altri sono specifici di uno stakehoder; alcuni concetti sono specifici dei CT, mentre altri sono concetti medici o generali. In questo

modo un'ontologia orientata ai concern per i CT puo' essere creata. Il metodo e' illustrato utilizzando i criteri di selezione dei soggetti, una componente di un progetto di CT, ma puo' essere usato per ogni altra componente del protocollo del CT. La tassonomia del vocabolario dei concetti dei CT e la rete delle relative basi razionali ci fornisce una struttura possibile per lo sviluppo del software specialmente se si adotta una soluzione basata su architetture orientate ai servizi.

Parole chiave: Trial clinici, Criteri di selezione, Ontologia, Concern

Crenguta Bogdan è ricercatore presso Ovidius University, Constanta, Romania (e-mail: cbogdan@univ-ovidius.ro).

Daniela Luzi è ricercatrice presso l'Istituto Superiore di Sanità (e-mail: daniela.luzi@irpps.cnr.it).

Fabrizio L. Ricci è ricercatore presso l'Istituto di Tecnologie Biomediche (e-mail: fabrizio.ricci@itb.cnr.it).

Luca D. Serbanati è dirigente di ricerca presso il Consiglio Nazionale di Ricerca (email: luca@serbanati.com).

Index

1. Introduction

A clinical trial (CT) is a "controlled study performed in human subjects and intended to discover, and/or evaluate, and/or verify safety, effectiveness, clinical and pharmaceutical effects, and adverse reactions of new drugs, devices, treatments, preventive measures, or other medical interventions in treating, preventing or diagnosing a specific disease or condition" [NCI]. The whole process involves high investments and several participating organizations, uses many resources, and is composed of numerous sub-processes, each of them carried out by agents playing specialized roles for the production of services and artefacts needed to the final evaluation of the research [CFLRSV].

To reduce costs and enhance the research quality new approaches to the research project design and management as well as more IT support are needed. The current trend is to expand the IT support and automatize as many CT activities as possible. Nowadays, software products are available on the market for CRF (Case Report Form) gathering and processing, Clinical Trial Management, etc., however other CT activities as protocol authoring remain largely manual because they are highly creative and intensely knowledge-based. Another bottleneck in the CT process is the lack of standardization and, as a consequence, of seamless interoperability between organizations, sub-processes, or agents at various levels of the research [FCLRSV].

The aim of our research is to provide CT researchers with new methods and techniques able to supply the CT design, setup, and conduction with more effective IT support and enhance interoperability between the CT components [LRS]. We consider that a comprehensive ontology of clinical research can contribute to solve these issues achieving interoperability between software applications and information systems of CT organizations, extending automatization in new areas of the CT process, as the protocol authoring, and facilitating the CT standardization. Building such an ontology is a challenge for analysts and new methods to obtain it are welcomed.

The complexity of clinical research requires to partition the domain of the real world under study in loosely coupled parts and for each part a successive partitioning to be carried out until the resulting parts can be easily described. Finding suitable partitioning criteria is a challenge for any system analyst. This is why many partitioning criteria are proposed by various system development methodologies. Our claim is that that the partitioning criteria are tightly related to the CT stakeholders' interests or concerns in developing or operating the system.

In a clinical trial, stakeholders are people who directly or indirectly are affected by or influence the system development and/or use. First of all, stakeholder is anyone whose jobs will be altered by the CT project, who supplies or gains information from the CT, or whose power or influence within the organization may increase or decrease according the success or failure of the CT. A clinical trial is the result of collaboration but also mitigation of possible divergences or disputes between various stakeholders. This is why analysis of stakeholders interests, preoccupations and desires may guide us towards a comprehensive model of CTs.

In [BS] we defined the concern in engineering as a problem-related care of one or more stakeholders involved in construction or operation of an artefact in its natural environment. The care can derive from an interests, desires or preoccupations of stakeholders for their environment's evolution. The concerns we identify in a CT design and evolution partition the knowledge on the CT universe in concern rationales which can be separately identified and described. Then we can merge them in a comprehensive ontology.

In this report a method of identification of concepts in the clinical trial domain is presented. The report presents a concern-oriented approach to obtain the ontology of clinical trials by conceptualizing the CT domain from the point of view of stakeholders.

2. Ontologies

An ontology is a formal specification of the intent of concepts and the intensional relationships that can exist between concepts. Using logical axioms it is a declarative model of a domain. As a formal system, an ontology allows us to build proofs of assertions viewed as theorems or conclusions. In respect to other models, ontologies , allow accurate expression of meaning of models.

Ontologies were introduced in the computer science domain over ten years ago and since then they gained an important role in Artificial Intelligence, Computational Linguistics, Database Theory, and Web Semantic.

The ontology notion was borrowed from philosophy, where "Ontology" term is "a systematic account of Existence" - aiming to account for all forms and modes of being. This broad definition was interpreted in different ways, depending on the sense that one takes with respect to what "existence" is. In Artificial Intelligence, where the ontology term was most prevalent used, what "exists" is exactly what can be represented.

According to Guarino's definition "An ontology is a logical theory accounting for the intended meaning of a formal vocabulary, i.e. its ontological commitment to a particular conceptualization of the world "[Guarino]. A conceptualization is a set of conceptual (intensional) relations defined on a domain space [Guarino].

2.1 Medical-domain ontology

Much was done for medical terminology and development of specialized languages in health care in order to enhance communication into the community of healthcare providers and consumers. However this languages and lexicons are informal and need much implicit knowledge. Medical information systems need to be able to communicate complex and detailed medical concepts (possibly expressed in different languages) unambiguously. To this aim medical-domain ontologies should be defined for medical terminology. Usually such domain ontologies are extensions of "foundational ontologies" which represent domain-independent concepts like physical or logical objects, events, processes, quality etc.

Ontologies allow us to integrate knowledge and data, so their use can help the development of more powerful and more interoperable information systems in healthcare, especially based on service-oriented architectures. Such systems can reuse and share patient electronic records and integrate legacy systems with new ones in an enterprise integrated system.

But creation of medical-domain ontologies is difficult task and requires a deep analysis of the semantics of concepts belonging to medical terminologies. There are some valuable attempts, but there is much to do, too.

2.2 Clinical trial-domain ontology

An ontology for clinical trials can help CT standardization enabling agreement on the meaning of CT vocabulary and facilitate interoperability and integration across applications. It can be represented in a specialized language enabling automatic querying and reasoning and in this way extending automatization of CTs. And the last but not the least, ontologies allow to integrate knowledge and data, so their use can help the development of more powerful and more interoperable information systems for CTs, especially based on service-oriented architectures. Such systems can reuse and share patient electronic records and integrate legacy systems with new ones in an enterprise integrated system [LRS].

Creation of CT-domain ontologies requires a deep analysis of the semantics of concepts belonging to medical terminologies. First, the CT concepts are identified and a concept subsuming hierarchy (taxonomy), which forms the ontology's backbone, is created. Then associations which define relations between concepts are identified and instantiated accordingly. Finally, a formal specification of the ontology is written.

In our research we use DOLCE (Descriptive Ontology for Linguistic and Cognitive Engineering) [MGBGO] with its extension D&S (Descriptions and Situations) [GCLB], [GM], [MVBCFGG] as a "foundational"[5] ontology where we "plug in" our CT ontology as well as OntoClean, a valuable methodology for analysis and validation of the ontological adequacy of taxonomic relationships. A brief introduction to OntoClean, DOLCE, and D&S follows.

---

[5] A foundational ontology characterizes general terms (e.g. entity, event, process, spatial and temporal location) and basic relations (e.g. part-of, quality-of, participation, dependence).

2.3 OntoClean methodology

A concept is a cognitive and semantic entity that allows organizing and interpreting the elements supplied us by either the perception of the actual reality, applying previous experiences to actual situations or our mental representations. Every concept can be informally described by three elements: name, intention and extension.

The name of a concept is the word which we use when we express the concept in natural language. The intent of a concept describes its semantic meaning in natural language. The extension of a concept is the set of all individuals from the world that verify the intention of the concept. In this case, we say that individual is an instance of that concept in the world, in the sense that it verifies the intention of the concept. Also, in this case we can say that the individual does the property "to be (an instance of) that concept" to be true. For instance, my computer verifies the property "to be a computer", or in other words, it is an extension of the concept "computer".

For ontological analysis of the concepts existing in the semantic rationales we have used OntoClean [Guarino, Welty]. The OntoClean methodology applies the notions used for ontological analysis in philosophy to analyzing conceptual modeling in information systems. In this methodology, the properties are defined as the semantics (or intents) of linguistics expressions like "being a concept _name", the extension of a concept is called class.

OntoClean [GW, 00], [GW, 02], [GW, 04] uses notions as *essence*, *identity*, and *unity*, to specify aspects of the meaning of the properties and relations that compose an ontology. The following aspects were determinant to choose OntoClean:
- provides formal notions (meta-properties) necessary for ontological analysis;
- contains clear steps and rules of correct constructing of a taxonomy and ontology of concepts;
- is the meta-language of DOLCE, supplying an ontology for it.

In this methodology, the properties are defined as semantics of linguistics expressions (intension) and for each concept there is a maximal state of affairs (or a possible world) that contains all the instances of the concept (extension).

The ontological analysis in OntoClean consists in the following steps for each property:
1. assign to each property of a meta properties set;
2. arrange the properties in a backbone taxonomy;
3. verify if the constraints imposed by the meta properties are violated in taxonomy,
4. analyze the non-rigid properties like phased sortals, roles, mixins and attributions.

These semantic aspects considered are represented by four metaproperties: rigidity, identity, unity, and dependence, which impose several constraints on the taxonomic structure of an ontology. The analysis of these constraints helps in evaluating and validating the choices made.

A property of an entity in a given world corresponds to a unary predicate in a first-order logic which induces a class in that world, that is a set of entities (the property instances) that exhibit that property. A property has a syntactic aspect, the predicate, and an extensional one, the class. OntoClean defines subsumption as relation between properties: a property p subsumes q if and only if, for every possible state of affairs, all instances of q are also instances of p. Given an entity and one of its properties, the property is essential to that entity if it must be true of it in every possible world, i.e. if it necessarily holds for that entity.

From an ontological point of view, a property is characterized by OntoClean meta-properties which are presented in Table 1. The relevant properties in an ontology should be verified if the meta-properties (noted R, I, U, O, D) hold or not for them. The result is for any property a code combination like this: +O-R~U+D, where +, -, and ~ indicate the presence, absence and, respectively, contrary of a metaproperty.

Symbol	Name	Description	Choices
R	rigidity	A property essential to all its possible instances	+R,-R,$\neg$R[6],~R
I	identity	It refers to the problem of being able to recognize individual entities in the world as being the same	+I,-I
U	unity	It refers to being able to recognize all the parts that form an individual entity	+U,-U, ~U
O	own	It marks those properties that supply their "own" identity criteria and that are not inherited from the subsuming properties	+O,-O
D	dependency	A property p is specific dependent on an other property if and only if, for all its instances, there exists instances of q which do not overlap with instances of p, that is q may have parts that are not shared with p	+D,-D

Table 1. Meta-properties in OntoClean

---

[6] $\neg$R indicates a semi-rigid property, that is a non-rigid property (-R), but not anti-rigid (~R).

According to the metaproperties combination, OntoClean ontologically and formally classifies the properties in eight categories: types, quasi-types, categories, material and formal roles, phased sortals, mixins and attributions [GW, 00].

OntoClean also imposes constraints on the subsuming relation and hypothesizes on identity.

2.4   DOLCE Ontology

In our research we use DOLCE with its extension D&S as a "foundational"[7] ontology where we "plug in" our CT ontology. DOLCE is an ontology of particulars, in the sense that its domain of discourse is restricted to particulars[8]. We adhered to the DOLCE +D&S ontology because:
-    DOLCE captures the ontological categories belonging to natural languages and human commonsense;
-    being a formal ontology [Guarino], DOLCE and D&S precisely define its notions and relations used;
-    DOLCE and D&S contain all categories and relations we need for writing our CT ontology.

DOLCE is a formal ontology making use of concepts from linguistics, philosophy, and mathematical logic. It provides clear semantics for the natural language, clear motivations for the adopted distinctions and strict rules on how to specify terms and relations.

Furthermore, DOLCE is an ontology of particulars, in the sense that its domain of discourse is restricted to them. The fundamental ontological distinction between universals and particulars can be informally understood by taking the relation of instantiation as a primitive one: particulars are entities that cannot have instances that are individuals; universals represent ideas or abstract entities which can have instances. Universals are described by means of properties and conceptual relationships used for particulars' study.

Since their domains of discourse are disjoint, we take the ontology of universals as formally separated from that of particulars. Of course, universals appear in an ontology of particulars, until now they are used to organize and characterize them.

2.4.1 The Top Categories

DOLCE deals with four top categories of particulars: endurants, perdurants, qualities and abstracts [MGBGO]. The taxonomy of these categories with their basic sub-categories appears at the top of the diagram in Fig. 1. All categories are assumed to be mutually disjoint and to cover the whole domain of particulars.
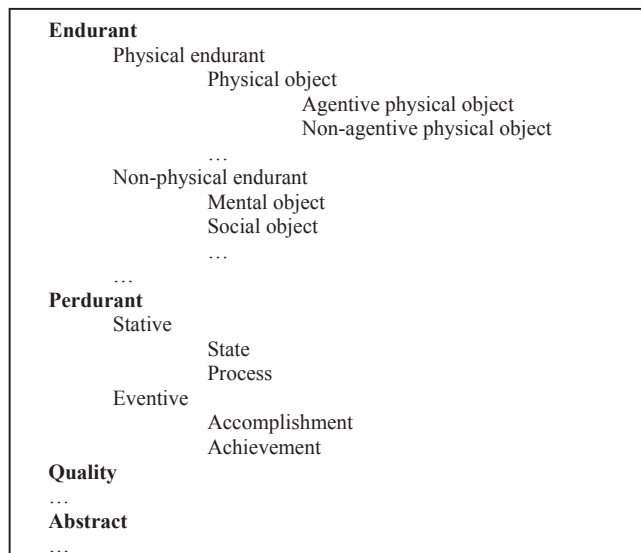
```
Endurant
     Physical endurant
               Physical object
                         Agentive physical object
                         Non-agentive physical object
               …
     Non-physical endurant
               Mental object
               Social object
               …
     …
Perdurant
     Stative
               State
               Process
     Eventive
               Accomplishment
               Achievement
Quality
     …
Abstract
     …
```

Fig. 1 DOLCE Taxonomy

---

[7] A foundational ontology characterizes general terms (e.g. entity, event, process, spatial and temporal location) and basic relations (e.g. part-of, quality-of, participation, dependence).

[8] Particulars are entities that cannot have instances that are individuals, whereas universals represent ideas or abstract entities which can have instances.

By definition, endurants are particulars in space, which participate at least in one perdurant (e.g. substances, objects, social entities, concepts). Perdurants are particulars in time, which have at least one participant (e.g. events, states, processes, phenomena). Therefore, the main relation between endurants and perdurants is that of participation. For example, a person (i.e., an endurant) can participate in a discussion (that is a perdurant) at some time. A person's life is also a perdurant, where a person participates throughout all its duration. In such cases, we say the endurant x (person) participates in the perdurant y (discussion or life) during the "t" time (shortly, PC(x, y, t)).

Another way of characterizing endurants and perdurants has been proposed by K. Hawley (cited in [MGBGO]): something is an endurant if and only if: (1) it exists at more than one moment and (2) statements about what parts of it must be relative to some time or other. In other words, the distinction is based on the different nature of the parthood relation when applied to the two categories: endurants need a time-indexed parthood, but perdurants do not. Indeed, the statement "this pencil is part of my pencil box" is incomplete unless you specify a particular time, but "my youth is part of my life" does not require such specification.

The second difference between endurants and perdurants is related to their behavior in time. Endurants are always wholly present (i.e., all their proper parts are present) at any time of their presence. Perdurants, however, just extend in time by accumulating different temporal parts[9], so that at any time they are present, they are only partially present, in the sense that some of their proper parts (for example, their previous or future phases) might not be present. For example, the piece of paper you are reading now is wholly present, but some temporal parts of your reading are not present any more. Philosophers say that endurants are entities that are in time but lack temporal parts (put in other words, all their parts flow with them in time). Perdurants, however, are entities that happen in time and have temporal parts (all their parts are fixed in time).

Qualities can be seen as the basic entities we can perceive or measure: colors, weights, speeds, sizes, and so on. Qualities are dependent particulars, "inherent" in either endurants or perdurants. Qualities inhere also to entities: each entity (including qualities themselves) comes with its own qualities, which exist as long as the entity exists. Thus, we distinguish between the quality and its value. The value is called quale and describes the position of an individual quality within a certain conceptual space (called by DOLCE quality space).

In this ontology, space and time are considered quality types, such as color and weight. The spatial (temporal) individual quality of an entity is called *spatial* (*temporal*) *location* and its quale is called *spatial* (*temporal*) *region*. According to their temporal and spatial locations, perdurants and endurants have a symmetric behavior: perdurants have a well-defined temporal location, but their spatial location seems to come indirectly from the spatial location of their participants; similarly, most endurants have a clear spatial location, but their temporal location comes indirectly from that of the perdurants they participate in. We will use the notation PRE(x, t) for the statement "x being present at t time" whenever a particular (endurant or perdurant) x exists, i.e. if its temporal region (quale) exists in the quality space.

Abstracts are particulars neither in space nor in time (e.g. sets, regions, metric spaces, etc.).

2.4.2 Formal ontological relations

Apart of identity relation DOLCE offers the "parthood" relation that corresponds to the unity relation. It is given in two versions: temporal parthood (for perdurants) and time-indexed parthood (for endurants).

Constitution is a relation between particulars and it is different from identity and parthood. This relation states that an entity (perdurant or endurant) constitutes another entity on some period. It uses the notation K(x,y,t) for "x constitutes y during t time".

Participation is a time-indexed relation between endurants and perdurants, expressing the fact that an endurant participates in an occurrence (perdurant) during some time interval. Time indexing is provided by the temporal location of the perdurant at a time interval, while the respective spatial location at a space region is provided by the participating endurant.

Quality is a relation that holds between a quality and what it inheres to: another quality, endurant or perdurant.

Let us now informally introduce some useful definitions based on the notion of dependence. DOLCE focuses on *ontological dependence* holding primarily between particulars. A particular *x* is *specifically constantly dependent* (SD) on another particular *y* if and only if, at any time *t*, *x* can't be present at *t* unless *y* is also present at *t*. For example, a person might be specifically and constantly dependent on his/her brain. A particular *x* is *one-sided specifically constantly dependent* (OSD) on another particular *y* if and only if *x* is specifically and constantly dependent on *y* and *y* doesn't constantly depend on *x*. A particular *x* is *generically constantly dependent* (GD) on a property *ϕ* if and only if, at any time *t*, *x* can't be present at *t*, unless a certain instance *y* of

---

[9] *x* is a temporal part of *y*, if and only if *x* is a part of *y* such that all parts of *y* temporally co-located with *x* are parts of *x*.

ϕ is also present at t. For example, a person might be generically and constantly dependent on the fact of having a heart.

2.5 D&S Ontology

The Descriptions and Situations ontology [GCLB]. [GM], [MVBCFGG], shortly D&S, defines a theory aimed to support a first-order manipulation of theories and models. It is used as an ontology design pattern for structuring application ontologies that require contextualization or when an ontology needs to be extended with non-physical objects, such as organizations, plans, narratives, rules, schedules, parameters, diagnoses, etc., as often it is the case in the CT ontology.

It is based on a formal definition of the description and situation concepts.

In D&S:
- a situation is an unitarian state-of-affairs satisfying a description (its unity criterion). It may represent a clinical data set, a set of CRF data, a diagnostic, etc.;
- a description is a non-physical object which represents an interpretation of a state-of-affairs in a non-physical context. It is conceived by an agent (human, artificial, collective or social). It may represent a diagnosis, a clinical research, etc.
- a concept is a social object which is defined by a description. Once defined, a concept can be used in other descriptions. The relation "classifies" relates concepts to particulars and even concepts to concepts.
- three kinds of concepts are reified: course, role and parameter.

As the name indicates, D&S is based on a formal definition of the *description* and *situation* concepts. A *description* is a non-physical object which represents an interpretation of a state-of-affairs in a non-physical context; hence it is generically dependent on some agent and communicable.

In D&S a state-of-affairs satisfying a description is a *situation*. In other words, it is a particular which represents a state of affairs, or a relation, tuple, or a fact, under the assumption that its components 'carve up' a view (a *settin*g) on the domain of an ontology by virtue of a description.

A concept is a social object, which is defined by a description. Once defined, a concept can be used in other descriptions. The relation classifies relates concepts to particulars (and possibly even concepts to concepts) at some time. In D&S three kinds of concepts are reified: course, role and parameter.

The role property is anti-rigid and dependent property. Any role admits a description and D&S ontology defines the role intension as its description that has the property that it classifies the endurants. We agree with this definition, because during the protocol writing, the writing committee are interested and define the role description of the subjects.

3. Concerns

In software engineering the separation of concerns is a decomposing method of a system into smaller, more manageable and comprehensible parts, each of which deals with a care of a particular area of interest or concern. The concern separation allows analysts to structure and developers to better manage complex systems by mapping the whole system in a multi-concern space and focusing one problem at a time. Further refinement of the system model can be followed for a while along one-concern dimension, independently of other dimensions. In this dimension, sub-concerns of the main concern can be used to partition once again the system.

The idea to separate concerns isn't new, it was stated in 1987 by D. L. Parnas in the paper [Parnas] where he proposed to encapsulate features in separate modules in order to localize changes to them and deal with one important issue at a time.

Twelve years later, P. Tarr et al. considered again this principle in [TOHS] and describe what they called the "tyranny of the dominant decomposition", that is decomposition of the artefacts of a software system into modules according to a dominant dimension. For example, the object-oriented methodology decomposes the software artefacts according the dimension of objects. Instead, the structured methodologies does not impose a dominant dimension and although both data and function dimensions may be used, the developer cannot simultaneous decompose the software artefacts by using the two dimensions, so he/she ultimately chooses a dominant dimension.

Since then the concern separation issues spread out in the software engineering world and even if the concern concept was not precisely and clearly defined, a number of concern oriented approaches have been proposed. These range from models for requirements engineering like AORE [RSMA] and Cosmos [SR] to the multidimensional approaches, like Hyperspace [OT]. Work has also been carried at the implementation level through extensions of the object-oriented programming like aspect-oriented programming [EFA] or subject-oriented programming [HO]. These works showed that the concern separation in software engineering has an effect on several quality attributes of the software systems like understandability, extensibility, maintainability, reusability and adaptability.

3.1   Intentional states of mind

In order to identify the concerns and their relations, our method firstly recommends to analyse the stakeholders' preoccupations, interests, needs, desires and beliefs, and identify how they generate concerns, in other words how the stakeholders reason. For this, we resort to contemporary philosophy that gives us many theories of mind.

The most known and accepted theory is the Functionalism which models the states of mind (beliefs, concerns, desires, being in pain, etc.) by considering solely their functional role: transformers of sensory inputs in behavioral outputs, in causal relations with other states of mind [Block].

Functionalism introduces the states of mind in the cause-effect relationship model where the states of mind have a causal role in the body's behavior [Block]. Abstracting states of mind with their functional role allows their multiply realization, that is they are able to be manifested even in computers, so long as computers performs the appropriate functions. In playing its role, a state of mind can be characterized by:

- causes (as they result, for instance, from perceptions),
- effects (as the organism behaves), and
- causal interactions with the other states of mind.

Many states of mind as concerns, beliefs, interests, needs, desires, and preoccupations are intentional[10], that is they are directed at or about or of some entity in the world: an object, phenomenon or state of affairs [Block]. Any state of mind, which has something it is about, is intentional. The intentionality is also applied when we expressed other states of mind, desires, for instance. Beliefs and perceptions have a world-to-mind direction of fit, while interests and desires have a mind-to-world direction of fit.

States of mind are in close relation with knowledge, that is information evaluated, processed, and structured by the human mind in conclusions or explanations so that it can be used according to current needs. They are based on mental representations that contain knowledge structures and are related by psychological mechanisms to other mental representations in order to create the meaning of something perceived or of interest. The mental representations can be expressed in natural language.

---

[10] Intentionality is a relationship between mental acts and the external world.

In our method, the beliefs take a special place because they are closely related to concerns. They represent convictions a stakeholder holds for true in a given situation, independently of the nature of the source of the conviction: either from perception, or an inference from previous knowledge, or a verbally transmitted knowledge. Of course, a belief may be true for one or more stakeholders (thus a piece of knowledge for them), but false for others. When it is accepted as true by all of us (or in all possible worlds) it is a piece of general knowledge.

3.2 Stakeholders and their concerns

Individual stakeholders in CTs can be classified depending on their role. A *role* is a unit of defined responsibility that may be assumed by one or more individuals (human being, team, or organization) in a collaborative environment aimed to deliver an artifact or service. The role is determinant for the concerns the stakeholder exhibits during the environment evolution. Our claim is that a concern is an intentional state of mind arising from preoccupation, interest, need or desire in a problem which the stakeholder has identified in the real world.

There are many roles of stakeholders in clinical research: sponsor, investigator, subject, data manager, writing committee, ethical committee, Institutional Review Board and so on. Sometimes it is useful to consider collective stakeholders as the writing committee formed by individual stakeholders, such as study coordinator, coordinating physician, data manager, statistician, etc.

In [BS] we defined the (stakeholder's) concern as a problem-related care of one or more stakeholders involved in the construction or operation of an artefact in its natural environment. The care can derive from stakeholders' interests, desires, needs or preoccupations for the environment's evolution. In our case, the main artefact is an informational or software system, but it may be the case of any artefact (documents, program products, etc.)

We identified the problem resulting from a concern as a pair of situations, called *initial state* and *final state*, a kind of snapshots of the limited world of interest. The problem's initial state is the current situation in this world as the stakeholder perceives it. The concern occurs when the stakeholder feels the need to change the initial state because he/she has found something missing inside or in divergence with his/her expectations. For this he/she intends to do something in order to achieve the final state, that is another (for the moment, imaginary) situation which matches his/her expectations or goals.

These two elements are respectively considered as hypothesis and conclusion of the match-to-expectations tentative plan the stakeholder will have to fulfil. If grouped with the role of the stakeholder who manifests interest or preoccupation about the problem, this pair becomes the *high level specification* of a concern that the stakeholder will try to solve. We consider a problem's initial state containing all data, information, and knowledge necessary to obtain the final state of the problem. For example, design and management of the CT are based on the following two concerns which reflect the two main preoccupations of the CT managers: a) find out if the new drug or procedure is effective and guarantee the safety of the subjects enrolled in the CT and b) test the effectiveness of the new therapy and/or procedure.

C1	*Name*: Care not to worsen the general clinical state of the subjects during the clinical trial and follow-up	
	Problem	*Hypothesis*: Objectives were defined in order to start a clinical research on a new drug and/or procedure.
		Conclusion: The drug or procedure risks were minimized as much as possible and in the case of appearance of adverse events, their gravity was minimized, the safety of the subjects was preserved during the study and follow-up.
	Stakeholders: Sponsor, Scientific Coordinator	
C2	*Name*: Care to attain the objectives of the clinical trial	
	Problem	*Hypothesis*: Objectives were defined in order to start a clinical research on a new drug or procedure.
		Conclusion: The clinical trial objectives were attained.
	Stakeholders: Sponsor, Scientific Coordinator	

During the process of solving the concerns C1 and C2 many other concerns will come out from various stakeholders. Moreover, a stakeholder may have several concerns related to the same problem. Each of them derives from one of the aspects with which the problem challenges the stakeholder's ability to solve it. The

aspect - it could be quality of service, data security, risk government, etc. - originates from the stakeholder's beliefs or knowledge.

While a stakeholder acts for solving a concern, situations that trigger new concerns may occur. Between these concerns and the original one a causal relation exists. For instance, the concerns beginning with C4-C7 derive from the stakeholder's preoccupation to solve C3 (see bellow). But it is also possible that the natural evolution of the environment where the stakeholder operates to solve a concern generates new situations of concern without an apparent causal relation with another concern. The new concerns could be predictable or unexpected ones. They should be considered and solved by the stakeholder or delegated to other stakeholders responsible for solving them.

4. Towards an Ontology of the Clinical Trial

In [BS] a concern-oriented method for information system analysis is proposed. The method has five steps:

1. identification of stakeholders;
2. identification of concerns;
3. concern classification;
4. identification of relations between concerns; and
5. concern prioritization.

The method also includes guidelines for designing an ontology of the domain under analysis. For this, seven steps should be added to the basic method:

6. identification of semantic rationales in concern specifications;
7. identification of concepts used in the semantic rationales and their classification in three categories: those belonging to the medical concepts, those belonging to some general, CT-independent ontology, and those strictly belonging to the CT vocabulary. Of course, the separation criterion belongs to the semantics of each concept;
8. ontological analysis of the intension of the concepts;
9. choosing a foundational ontology to be extended by our ontology;
10. classification of the concepts conforming the foundational ontology;
11. identification of the concept mapping (conceptual relations) in the semantic rationale with the classification of relations conforming to the foundational ontology; and
12. definition of the ontology using a formal logic language.

Such an ontology is structured according to stakeholders' responsibilities and concerns through the identification of shared and specific vocabularies.

In the followings we present this method for the sample case of an ontology of the selection criteria of CT subjects. In this case the concerns we consider are mainly related to the writing committee stakeholder.

4.1 Concerns identification

A CT is based on the study protocol that is a complex planning document which describes in detail the activities to carry out during the research in order to achieve the CT objectives. The protocol is derived from a previous document, the CT outline [FLMRTV]. Both documents include the characteristics for selection of the study population - a homogeneous and reliable number of subjects identified on the basis of a statistical analysis on a specific pathology - as a primary condition to obtain trustworthy results. One of the first tasks in the CT process is to define the person characteristics needed to select the study population [CFLRSV]. A complete, consistent, and non-redundant set of characteristics is the primary condition to obtain trustworthy results. These characteristics are used to describe inclusion criteria, that is criteria that prospective subjects must meet to be eligible to participate to the study, and exclusion criteria, that is, criteria that may exclude a potential subject from participation in the study, even if this subject satisfies the inclusion criteria. One compulsory exclusion criterion is the not signing of the informed consent. The subject selection criteria are an aggregation of inclusion and exclusion criteria.

In this paper we focus our analysis on the subject selection criteria and their main stakeholder, the writing committee, which has the following concern regarding the identification of the subject selection criteria:

C3	*Name*: Care to obtain the characteristics of the study population	
	Problem	*Hypothesis*: Following information is available: 1. Trial objectives, trial design, statistical considerations 2. Previous CT projects and their scientific results
		Conclusion: Characteristics of the CT subjects are identified and specified as inclusion/exclusion criteria.
	Stakeholders: Writing committee	

As already stated to solve the concern C3, the characteristics of the study population should be identified and gradually specified as inclusion and exclusion criteria. As a consequence other concerns, C4-C8 are derived representing subtasks of the comprehensive task to solve the problem associated to C3. For deriving the new concerns the members of the writing committee make use of tacit knowledge and their work practice (for example, design methodologies, design guidelines, operative manuals, etc).

In the table 2 four samples of knowledge are specified through an identification code and a description in natural language of their corresponding mental representations.

Code	Mental representation description in natural language
K1	The subjects who have the disease in the stage considered by the study are eligible.
K2	The subjects who have other medical condition that could bias the study compliance and follow-up are not eligible.
K3	The therapeutic procedure administration will affect the clinical manifestation of the disease.
K4	Results of previous phases of the CT or experimental laboratory results are available and a list of ascertained side effects and/or adverse events is known.

Table 2. Samples of CT knowledge

According to K1, the set of possible CT participating subjects considers those who have the disease in the stage under study. This is why, in order to solve the concern C3, a first care is to find how the clinical manifestation of the disease is related to the objectives attainment and the selection criteria identification, that is the concern C4:

C4	*Name*: Care to find how the clinical manifestation of the disease is related to the objectives attainment and the selection criteria identification	
	Problem	*Hypothesis*: Following information is available: 1. Trial objectives, trial design, statistical considerations 2. Previous CT projects and their scientific results
		Conclusion: How the disease in the stage treated in the study will be tested in order to get the diagnostic?
		What basic characteristics (gender, race, life expectancy, performance status) have the study population affected by the disease?
		How the clinical manifestation of the disease might affect the safety of the subjects and the attainment of the objectives?
		What particularities of the disease are important for the identifying the selection criteria?
		How the pathogenesis of the disease might affect the general clinical state of the subjects and the inferring of the selection criteria?
	Stakeholders: Writing committee	

As we observed from the high level specifications of the concerns C1-C3, the description of the final state is usually very general, so it may be indicated also helping with the questions which the stakeholder has to answer in order to obtain the final state. Examples of such concerns specifications are C4-C8.

From the CT general objective as well as from the concerns C1-C4 and from the knowledge pieces K1-K4 originate the following concerns:

C5	*Name*: Care to determine how the components of the study therapeutic procedure and its administration might affect the general state of the subjects and/or influence the attainment of the objectives	
	Problem	*Hypothesis*: Following information is available: 1. Trial objectives, trial design, statistical considerations 2. Previous CT projects and their scientific results
		Conclusion: How the kind of therapeutic procedure and its components affect the general clinical state of the subjects and the attainment of the objectives?
		The therapeutic procedure and its components have been tested on the subjects with the age less than 18 years? If so, what effects they had? In this study, might also take part the subjects with the age less than 18 years?
		The therapeutic procedure can have effects on the developing of the fetus?
	Stakeholders: Writing committee	

C6	*Name*: Care to determine the way in which previously adverse events of the study therapeutic procedure administration might affect the general clinical state of the subjects and/or influence the attainment of the objectives	
	Problem	*Hypothesis*: Following information is available: 1. Trial objectives, trial design, statistical considerations 2. Previous CT projects and their scientific results *Conclusion*: The therapeutic procedure contains teratogenic agents or an abortifacient preparation? How other side effects and/or adverse events might affect the general clinical state of the subjects and/or influence the attainment of the objectives?
	Stakeholders:Writing committee	
C7	*Name*: Care to determine the way in which other treatments or drugs might affect the general clinical state of the subjects and/or the attainment of the objectives	
	Problem	*Hypothesis*: Following information is available: 1. Trial objectives, trial design, statistical considerations 2. Previous CT projects and their scientific results *Conclusion*: How other (prior or concomitant) treatments/drugs might affect the general clinical state of the subjects and/or the attainment of the objectives? How other therapeutic (or other kind) procedures might affect the general clinical state of the subjects and/or the attainment of the objectives? What other drugs might be forbidden because they provoke responses (e.g. diseases, adverse events, etc.) that are monitorized in the study? What other drugs might be administrated during the study to attenuate the risk of adverse events appearance, on condition that the attainment of the objectives is not affected?
	Stakeholders: Writing committee	
C8	*Name*: Care to determine if other diseases or medical conditions might affect the general clinical state of the subjects and/or the attainment of the objectives	
	Problem	*Hypothesis*: Following information is available: 1. Trial objectives, trial design, statistical considerations 2. Previous CT projects and their scientific results *Conclusion*: How other diseases that might provoke the same adverse events that might appear during the study affect the general clinical state of the subjects and/or the attainment of the objectives? How other diseases might affect the general clinical state of the subjects and/or the attainment of the objectives?
	Stakeholders: Writing committee	

Other concerns derive from the analysis of particular views of the selection criteria, for example the ethic aspects originated in C1. For instance the next concern, C9, comes from the responsibility of the writing committee to specify selection criteria of quality:

C9	*Name*: Care to specify unambiguous, consistent, and not redundant selection criteria as a sum of inclusion and exclusion criteria	
	Problem	*Hypothesis*: Following information is available: 1. Trial objectives, trial design, statistical considerations 2. Previous CT projects and their scientific results *Conclusion*: The selection criteria of the CT subjects have the following quality attributes: non ambiguity, consistency, and non redundancy
	Stakeholders: Writing committee	

4.2 Beliefs Related by Concerns

Our method firstly proposes an analysis of the stakeholders' beliefs (but also their preoccupations or interests) for identifying the causal relations between states of mind. Mental representations of the beliefs are expressed in natural language as we did for pieces of explicit knowledge. Table 3 collects the results of the analysis of the concerns C1-C4. In the table "Bx" is used for coding mental representation which corresponds to a belief where B stands for a belief and x is a number.

State of mind code	Mental representation description in natural language
B1	There is the risk that the general clinical state of the subjects may become worse during the clinical trial and follow-up
B2	The safety of the subjects is preserved during the clinical trial and follow-up
B3	The objectives of the clinical trial are fulfilled
B4	The study population is formed by subjects.
B5	The study population has characteristics
B6	The characteristics of the study population are defined by the inclusion and exclusion criteria.
B7	The inclusion and exclusion criteria form the selection criteria.
B8	The clinical manifestation of the disease has relation with the attainment of the objectives and the selection criteria identification
B9	The disease in its stage has a diagnostic as it resulted from tests
B10	The existence of the disease and its stage affect the life expectancy of a person
B11	The life expectancy is a characteristic of study population
B12	The study population has other basic characteristics: age, gender, race and performance status
B13	The existence of the disease and its stage affect the performance status of a person
B14	The clinical manifestation of the disease might have some particularities that are important to identify the selection criteria.
B15	Any disease has a pathogenesis.
B16	The pathogenesis of the disease might affect the general clinical state of the subjects and the inference of the selection criteria.

Table 3. Beliefs which originate concerns C1-C4

4.3 Dependency in the Concern Space

We represent the causal relations between intentional states of mind (beliefs, preoccupations, etc.) as dependencies.

There are two kinds of dependence between states of mind: direct and bound.

Direct dependence. It indicates a causal relation describing that the origin of one component of the relation resides in the other component. The relation may relate: a) a concern and its relative belief, preoccupation, interest; and b) two concerns, beliefs, preoccupations, interests, and knowledge.

The first case can be exemplified with the concern C1 which is derived from a preoccupation P1: "the general clinical state of the subject does not be worsen during the clinical trial and follow-up". In its turn P1 causally depends on the belief B1. These dependencies between states of mind are described with dashed line as in the diagram in Fig. 2.
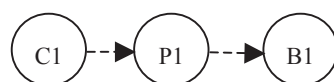

Fig. 2. Direct dependence

The second case comes out when there are beliefs which cause preoccupations or interests on the same problem considered from two different points of view. For example, the knowledge K4 says that adverse events

of the study therapeutic procedure administration have been ascertained by previous tests and/or research studies cause the preoccupation to minimize them during the study. This preoccupation expresses the preoccupation P1, the same we used above, but from the point of view of the adverse events occurrence risk. The risk could be mitigated if, for instance, we require a normal function of major organs (e.g. cardiac function at least Grade 2). These requirements contribute to building of the inclusion criteria for subjects' enrolment.

Another situation is when a belief or a piece of knowledge is the consequence of another belief or knowledge. For instance, the belief B9 may be the consequence of the next knowledge: "the tests have been done and their corresponding value ranges of results were delivered".

The direct dependence relation between two beliefs or knowledge also comes from the ontological relation between the intentions of at least two concepts from the mental representations of the beliefs. The relation is one of the kinds: subsumption or part_of or dependency.

To explain this case, let us consider the beliefs that appear during the concern C4 and their relations (Fig. 3). Because C4 depends (historically) on the belief B8, we can suppose that the reasoning process begins with the state B8. This state gradually originates mental representations of the contained concepts: clinical manifestation of a disease, CT objective, selection criterion, etc. In the context of these representations, the disease clinical manifestation concept is the most important. It is related to the representation of the disease concept, due to the link of dependency type between disease and clinical manifestation.
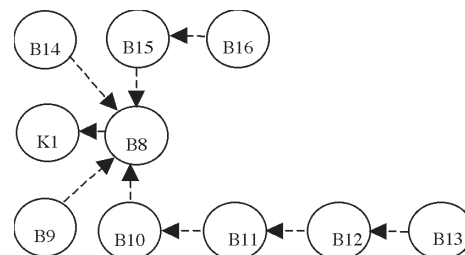

Fig. 3. Direct dependence relations between the beliefs related to C4

Using the same reasoning process, we can infer that the mental representation of the disease concept is bound to representations of other concepts (generated on knowledge), like diagnostic, life expectancy of a person, and pathogenesis.

The link between disease and diagnostic is expressed in the belief B9 which expresses the fact that at enrollment time the disease of the potential subject is diagnosed and a diagnostic was delivered after some tests have had been carried out.

If we re-do the reasoning process, we can conclude that the belief B8 (the starting state) caused the state B9, otherwise the B9 state would not exist if the B8 state did not exist. This is in fact the dependence relation between B9 and B8.
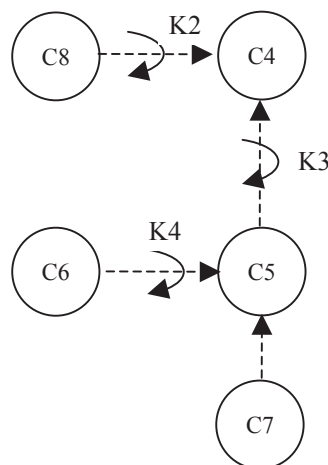

Fig. 4 Direct dependence relations between C3-C9

Similarly, if we apply the last two steps in the case disease-life expectancy of a person and disease-pathogenesis, we obtain dependency relations between B10 and B8, respectively B15 and B8. These relations are showed in Fig. 34. In the figure there are also represented pieces of knowledge which are at the basis of the dependency relation.

The direct dependence relation between two concerns is inferred by their beliefs and pieces of knowledge. It appears when there is a dependency relation between two beliefs from two different concerns and this usually exists due to another piece of knowledge. For instance, the concern C5 depends on C4, because there is a dependency relation between a belief from C5 and B8. The latter relation exists because of the piece of knowledge K3, as we showed in the Fig. 34.

     Bound dependence. This relation is between two states of mind (interests, needs or desires, and beliefs), and expresses the fact that one state of mind is based on the other. For instance, the concern C2 causally depends on an interest I1 "the objectives of the clinical trial to be fulfilled". The mental representation expressed by this sentence is the same as the mental representation of the sentence "the objectives of the clinical trial are fulfilled" which is related by the belief B3. At the ending of the clinical trial, this sentence is true or not, depending on whether the CT was successfully closed. In this case, we use the bound dependence to express the fact that the mental representation of an interest is the same or has the same semantic content as the mental representation of a belief. This dependence between two stats of mind is shown in the Fig. 35.
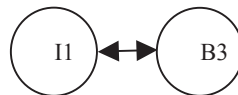

Fig. 5 Bound dependence

4.4 Concern rationales

We define the concern's *rationale* a cluster of concepts and their conceptual relations used for the concern solving. In other words, the concern's *rationale* describes the semantic domain, which the stakeholder should manage in order to solve the concern.

The concepts of such a rationale come from informal descriptions of the mental representations of the beliefs of the stakeholder (-s) during the concern's life cycle. The beliefs to be considered are those associated to the concern's hypothesis and conclusion, but also those which represent milestones in the concern solving process, mainly the hypothesizes and conclusions of concerns the concern depends of.

Using the beliefs from the table 3 we present in the Table 5 the composition of each rationale (noted with R) together with the associated concern code.

Semantic rationale code	Mental representations codes	Concern code
R1	B1, B2	C1
R2	B3	C2
R3	B3, B4-B7	C3
R4	B8-B16	C4

Table 5. The composition of rationales

4.5 Ontological analysis of concern rationales

In this step we gather the concepts identified in the concept rationales. Each concept is described as an OntoClean property. We assign OntoClean meta-properties to all concepts and to any other property which comes out from the analysis. In the Table 6 we present the list of concepts found in the rationales of the concerns previously analysed. For each concept its combination of meta-properties and the formal ontological category according to OntoClean are associated.

ConceptName	Concept intent	Meta-properties	OntoClean Category
Objective	In CT, we conceptualize an objective as a description of an interest of the stakeholders in the execution of an evaluation activity during a whole CT or follow-up. In addition, we impose the conditions that the evaluation activity drives any statistical planning for the trial and any objective is a part of a CT protocol.	+O+R+I+U+D	Type
Selection	The selection criteria are considered in CT as medical	+O+R+I+U+D	Type

ConceptName	Concept intent	Meta-properties	OntoClean Category
Criteria	description elements that are necessary to allow an individual to participate and become a subject in a specific clinical trial. This property is a rigid one for the CT protocol.		
Disease	We define the intention of the disease as any abnormal process of the person that causes discomfort, dysfunction, or distress to the person affected or those in contact with the person. Sometimes the term is used broadly to include injuries, disabilities, syndromes, symptoms, deviant behaviors, and atypical variations of structure and function.	+O+R+I+U+D	Type
Clinical Manifestation of Disease	We refer at the clinical manifestation of a disease as its internal and external manifestation in the body of the ill person.	-O+R+I+U+D	Quasi-Type
Disease Description	The description of a disease, including its clinical manifestation	-O+R+I+U	Quasi-Type
Stage of Disease	The stage of a disease is a clinical attribute used to specify and describe disease appearance or processes.	-O+R+I~U+D	Quasi-Type
Test	According to SNOMED-CT, a test is a medical procedure that involves testing a sample of blood, urine, or other substance from the body. Tests can help determine a diagnosis, plan treatment, check to see if treatment is working, or monitor the disease over time.	+O+R+I-U+D	Type
Diagnostic	Description of the result of the diagnosis. In CT, all the subjects have to get a diagnostic, and that is why, we consider that diagnostic is a rigid property.	+O+R+I-U+D	Type
Diagnosis	We agree with definition of NCI of diagnosis as the act or process of identifying or determining the nature and cause of a disease or injury through evaluation of patient history, examination, and review of laboratory data.		
Study Population	According to CDISC, the study population is any finite or infinite collection of subjects from which a sample is drawn for a study to obtain estimates for values that would be obtained if the entire population were sampled.	-O+R+I~U+D	Quasi-Type
Subject	Participant of CT to who is administrated the intervention or control.	-O~R+I+U+D	Material Role
Participant	Participant is a person or entity with a role in healthcare or a clinical study (CDISC).	-O~R-I+U+D	Formal Role
Person	Living being with the rigid property of having DNA.	+O+R+U-D	Type
Clinical Trial	Type of research study that tests how well new medical treatments or other interventions work in people. Such studies test new methods of screening, prevention, diagnosis, or treatment of a disease.	+O+R+U+D	Type
Gender	The gender is a subject's quality that can take one of the values: male, female.	-O+R+I~U+D	Quasi-Type
Race	In CT, the race is a quality of a subject that can take the values: American Indian or Alaska; Native, Asian, Black or African American, Native Hawaian or Other Pacific Islander, White.	-O+R –I ~U+D	Quasi-Type
Age	The age is a subject's quality	-O+R –I ~U+D	Quasi-Type
Life Expectancy	An assessment finding of mortality of a population that suffer by a certain disease.	-O+R -I ~U+D	Quasi-Type
Performance	According to NCI Thesaurus, the performance status	+O+R+I ~U+D	Type

ConceptName	Concept intent	Meta-properties	OntoClean Category
Status	is a measure of how well a patient is able to perform ordinary tasks and carry out daily activities.		
Pathogenesis of Disease	We use the definition of the pathogenesis as the development of a disease and the chain of events leading to that disease.	+O+R+I+U+D	Type
Description of Pathogenesis of Disease	Description of the pathogenesis of the disease	-O+R+I+U-D	Quasi-type
General Clinical State of Subject	In CT, the general clinical state of subject is defined as the description of the physical and mental integrity of the subjects who participate in clinical trial.	-O+R-I–U+D	Quasi-type
Characteristic of Study Population	Description of the distinguishing qualities or prominent aspects of the study population.	-O+R+I+U+D	Quasi-Type

Table 6. Classification of the concepts of R1-R4 according to OntoClean

According to OntoClean the rigid properties (i.e. categories, type and quasi-types) form the backbone taxonomy (Fig. 6). In the backbone taxonomy corresponding to the semantic rationales of SR1-SR4, we observe that we do not explicitly define the concepts of entity, description, state and process. We only make the common sense assumption that the first one subsumes the last three ones.

As we see in the table 5 we have two roles: participant and subject. The former is a formal role, meaning that is anti-rigid, dependent and does not carry identity, i.e. anything (person or organization) that participates in a clinical trial is a participant. The subject concept is a material role i.e. it is also an also anti-rigid (due to the subsumption relation between subject and participant) and subsumes the person concept, from which the subject inherits the identity criteria. But the last subsumption relation (person and subject) can not appear in the taxonomy because otherwise it violates the OntoClean constraint that a rigid property can not subsume a semi-rigid or anti-rigid property.

We also enrich the backbone taxonomy with the concepts identified as roles in the table 5 and we obtain an ontology according to OntoClean and corresponding of the rationales R1-R4 which is presented in Fig. 6.

4.6 Sample of Clinical Trial Ontology Definition

In Table 4 concepts derived from the concerns C1-C3 and regarding the selection criteria definition are defined. Some of them are CT specific (selection criteria, inclusion criteria, exclusion criteria), while others (person, variable, value of variable, maximal domain) belong to some related ontologies.

Concept	Formal semantic	Description
person	$Person(p) =_{df} APO(p) \wedge \exists c (DNA(c) \wedge OSD(p,c))$	Agentive physical object that specifically and one-side depends on DNA.
selection criteria	$SelectionCriteria(c) =_{df} Description(c) \wedge \exists w,t(WritingCommittee(w) \wedge Conceives(w,c,t)) \wedge \exists i1,i2,e,t (InclusionCriterion(i1) \wedge InclusionCriterion(i2) \wedge ExclusionCriterion(e) \wedge TimeInterval(t) \wedge K(i1,c,t) \wedge K(i2,c,t) \wedge K(e,c,t))$	Description conceived by the writing committee and constituted by the inclusion and exclusion criteria. The description has two parts: inclusion and exclusion criteria. Moreover, the persons have to have signed and dated an informed consent, K1 and K2.
variable	(A) $Variable(v) \rightarrow Description(v)$ (A) $Variable(v) \wedge \forall p1,s1 (Particular(p1) \wedge Situation(s1) \wedge SAT(s1,v)) \rightarrow \exists p2,s2, t (Particular(p2) \wedge TimeInterval(t) \wedge (P(p1,p2) \vee PC(p2,p1,t) \vee qt(p1,p2)) \wedge Situation(s2) \wedge P(s1,s2) \wedge SettingFor(s2,p2) \rightarrow \exists d (Description(d)) \wedge SAT(s2,d) \wedge PP(d,v))$	Generic description of various situations of a particular (namely endurant, perdurant or quality) used in the description of another particular. Such variable exists due to a kind of relation between the two particulars. The relation can be: parthood, participation, or quality.
value of variable	(D) $ValueOf(s, v) =_{df} Situation(s) \wedge Variable(v) \wedge SAT(s,v)$ (A) $Variable(v) \rightarrow \exists s1,s2 (Situation(s1) \wedge Situation(s2) \wedge ValueOf(s1,v) \wedge ValueOf(s2, v) \wedge \forall x1,x2(Particular(x1) \wedge$	A situation that satisfies a variable description. A variable can take different values.

Concept	Formal semantic	Description
	SettingFor(s1,x1)∧Particular(x2)∧SettingFor(s2,x2)→ x1≠x2))	
maximal domain	MaximalDomain(v)=$_{df}$ Variable(v)∧∀s (Situation(s)∧SAT(s,v) → valueOf(s,v))	Domain of all values of a variable.
characteristic of persons	(A) CharacteristicOfPersons(p,v) → Person(p)∧Variable(v) (A) CharacteristicOfPersons(p,v)→ ∃v1(Variable(v1)∧ CharacteristicOfPersons(p,v1)∧PP(v1,v)) (A) ¬∀p, v(Person(p)∧Variable(v)∧ CharacteristicOfPersons(p,v)→ ∃v1(Variable(v1)∧ CharacteristicOfPersons(p,v1)∧PP(v1,v)))	Variable that applies on a person. In addition, the characteristics of a person can be parts of other characteristics.
assessment domain	(A) AssessmentDomain(v) → ∃d(Variable(d)∧ MaximalDomain(d)∧ PP(v,d)) (A) AV(s,v) → ValueOf(s,v) (A) ∃s,v(Situation(s)∧Variable(v)∧ ValueOf(s,v) → ¬AV(s,v)	Sub-domain of the maximal domain which contains the values that are considered in the study.
inclusion criteria	(D) InclusionCriteria(d, p)=$_{df}$ Description(d)∧Person(p)∧ ∃v(Variable(v)∧CharacteristicOfPersons(p,v)∧PP(v,d)) (A) ∀d,p(Description(d)∧Person(p)∧ InclusionCriteria(d, p)→ ∃s(Situation(s)∧AV(s,d))	Description conceived by the writing committee of a non-empty group of assessment domain for a selection list of characteristics of persons. It is used to select (eligible) persons for the CT.
exclusion criteria	(D) ExclusionCriteria(d, p)=$_{df}$ Description(c)∧ Person(p)∧ ∃v(Variable(v)∧CharacteristicOfPersons(p,v)∧PP(v,d)) (A) ∀d,p(Description(d)∧Person(p)∧ ExclusionCriteria(d, p)→ ∃s(Situation(s)∧¬AV(s,d))	Description conceived by the writing committee of a non-empty group of assessment domain for a selection list of characteristics of persons. It is used to eliminate persons from the eligible ones.

Table 4 Excerpt from the ontology of the subject selection criteria

From Table 3 and Table 4 the concepts belonging to the concerns R1-R4 are gathered and their taxonomy is represented according to DOLCE+D&S in
Fig. 8.
From the general rationales, we only can identify few relations of three kinds:
1. the dependency relations between qualities and the types or quasi-type which they inhere to;
2. the parthood relations between concepts; and
3. the constitution relations between concepts.
We present some examples of these three relation kinds.

The relation between a quality and its type or quasi-type is ontologically treated in DOLCE by the quality_of formal relation. So, for the attributions from the R4 we have the following quality_of relations:

      qt(Stage of Disease, Disease)
      qt(Race, Person)
      qt(Life Expectancy, Person)
      qt(Sex, Person)
      qt(Performance Status, Person)

From the rationale R3, we also identified the following parthood relations between concepts:
      P(Inclusion Criterion, Selection Criteria)
      P(Exclusion Criterion, Selection Criteria)

From R3, we identified the following constitution relation:
      K(Subject, Study Population, t)

5. Conclusions

In this paper, a concern-oriented method for designing an ontology is presented. The method is based on stakeholders' concerns to partition the CT conceptual domain in stakeholder-oriented sub-domains. The stakeholders' concerns come from their interests, desires, and preoccupations and depend on their beliefs and knowledge.

Mental representations of stakeholders related to each concern are identified as clusters of concepts related to each other. We consider such a cluster as the rationale of the associated concern. The concepts found in the rationales populate the universe of discourse specific to each stakeholder and compose the stakeholder's vocabulary. Some concepts are shared with other stakeholders while others are specific to one stakeholder; some concepts are domain-specific while others are medical or general concepts.

The method is illustrated to derive the ontology of a component of the clinical trial protocol, the subject selection criteria but it can be used for any other component of the CT. Our current research is focused on the conceptualization of the CT protocol, the planning document where all core concepts of the clinical trial are found, in order to encourage standardization and improve interoperability in clinical research.

6. References

[Block]   Block, N., ed.: Readings in Philosophy of Psychology, 2 vols. Vol. 1, Cambridge, Harvard (1980)

[BS]   Bogdan, C., Serbanati, L. D.: Toward a Concern-Oriented Analysis Method for Enterprise Information Systems, IEEE International Multi-Conference on Computing in the Global Information Technology (ICCGI 2006), Bucharest, Romania (2006)

[CFLRSV]   Collada, A., Fazi, P., Luzi, D., Ricci, F. L., Serbanati, L. D., Vignetti, M.: Toward a Model of Clinical Trials. Biological and Medical Data Analysis. 5th International Symposium, ISBMDA, Barcellona, Spain, Lecture Notes in Computer Science, Springer-Verlag, Berlin (2004)

[EFA]   Elrad T., Filman R. E., Bader A., Aspect-Oriented Programming, Communications of the ACM, 44(10), 2001

[FCLRSV]   Fazi, P., Collada, A. L., Luzi, D, Ricci, F. L., Serbanati, L.D., Vignetti, M.: A Proposed Clinical Trial Model: Analysing the CT Process. *Applied Clinical Trials*, Vol. 15, 1, (2006)

[FLMRTV]   Fazi P., Luzi D., Manco M., Ricci F.L., Toffoli G., Vignetti M.: WITH: a System to Write Clinical Trials using XML and RDBMS. . Proc. AMIA Symp. 2002.

[GCLB]   Gangemi, A., Catenacci, C., Lehmann, J., Borgo, S.: Task Taxonomies for Knowledge Content, EU 6FP METOKIS Project D07 (2005)

[GM]   Gangemi, A., Mika, P.: Understanding the Semantic Web through Descriptions and Situations", International Conference ODBASE03, Italy, Springer (2003)

[Guarino]   Guarino N., "Formal Ontology and Information Systems", *Formal Ontology in Information Systems*, *Proceedings of FOIS'98, Trento, Italy, 6-8 June 1998.* Amsterdam, IOS Press

[GW, 00]   Guarino N., Welty C. A., "A Formal Ontology of Properties", *Proceedings of 12th Int. conf. on Knowledge Engineering and Knowledge Management,* LNCS Springer Verlag, 2000

[GW, 02]   Guarino N., Welty C. A., Evaluating Ontological Decisions with OntoClean, Communications of the ACM, 45(2):61--65, February 2002

[GW, 04]   Guarino N., Welty C. A., An Overview of OntoClean, in S. Staab and R. Studer, editors, Handbook on Ontologies in Inf. Sys., Springer, 2004

[HO]   Harrison W., Ossher H., Subject-oriented Programming (a Critique of Pure Objects), Conference on Object Oriented Programming: Systems, Languages, and Applications, 1993

[LRS]   Luzi, D., Ricci, F.L., Serbanati, L.D.: E-Clinical trials supported by a service-oriented architecture. To be published in MedNet 2006: The 11th World Congress on Internet in Medicine, Toronto (2006)

[MBGGO]   Masolo, C., Borgo, S., Gangemi, A., Guarino, N., Oltramari, A.: WonderWeb Deliverable D18. Ontology Library. IST Project 2001-33052 WonderWeb: Ontology Infrastructure for the Semantic Web (2003)

[MVBCFGG]   Masolo C., Vieu L., Bottazzi E., Catenacci C., Ferrario R., Gangemi A., Guarino N., Social Roles and their Descriptions, Proceedings of the Ninth International Conference on the Principles of Knowledge Representation and Reasoning. AAAI Press, 2004

[NCI]   NCI Thesaurus: http://nciterms.nci.nih.govT

[OT]   Ossher H., Tarr P., Multi-dimensional of concerns and the Hyperspace approach, M. Aksit editor, Software Architectures and Component Technology, Kluwer Academic Publishers, 2001

[Parnas]   Parnas D. L., On the criteria to be used in decomposing systems into modules, Communications of the ACM, 5(12), 1987.

[RSMA]   Rashid A., Sawyer P., Moreira A., Araújo J., Early Aspects: a Model for Aspect-Oriented Requirements Engineering, Proceedings of the IEEE Joint International Conference on Requirements Engineering, 2002

[SR]   Sutton S. M. Jr., Rouvellou I., Modeling of software concerns in Cosmos, Proceedings of the 1st International Conference on Aspect-Oriented Software Development, Enschede, The Netherlands, 2002

[TOHS]   Tarr P., Ossher H., Harrison W., Sutton S., N Degrees of Separation: Multi-Dimensional Separation of Concerns, Proceedings of International Conference on Software Engineering, 1999
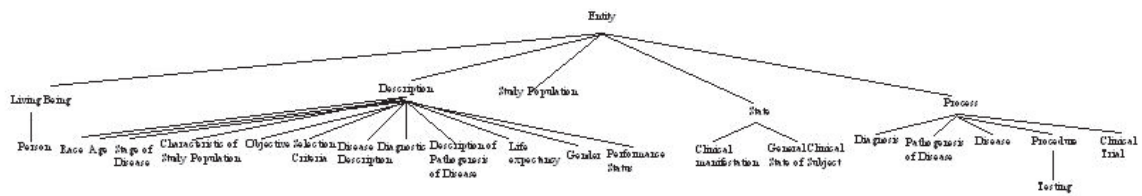
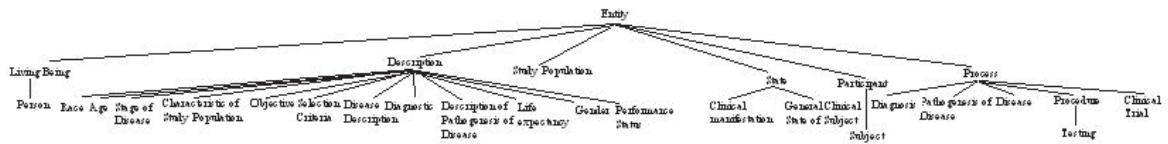**Fig. 6 The backbone taxonomy of the properties of the rationales R1-R4 according to OntoClean**

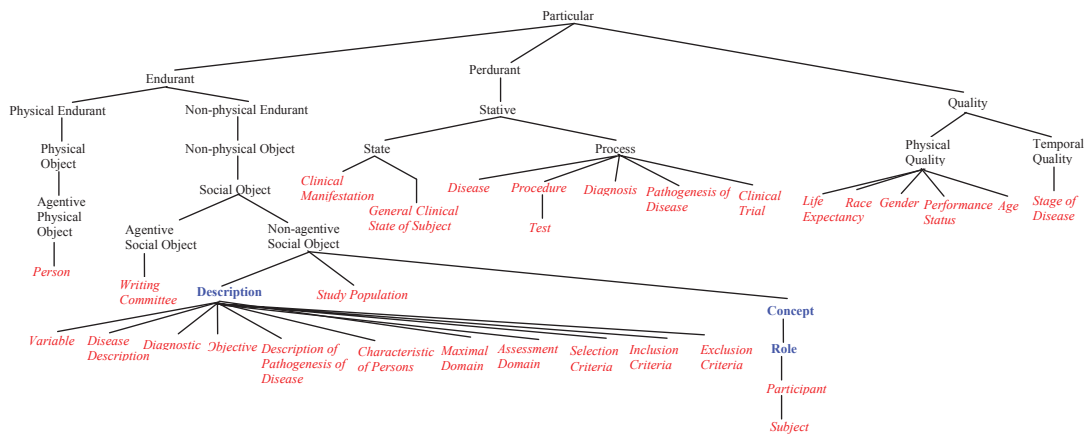**Fig. 7 Classification of properties of the rationales R1-R4 according to OntoClean**

Fig. 8 Classification according to DOLCE + D&S of the concepts that belong to the rationales R1-R4