

preprint version of paper published in *Computing, Philosophy, and Cognitive Science*, G. D. Crnkovic and S. Stuart (eds.),
Cambridge: Cambridge Scholars Press, 2007, 104-122

Ontology as the Core Discipline of Biomedical Informatics

Legacies of the Past and Recommendations for the Future Direction of Research

Barry Smith ^a and Werner Ceusters ^b

^aIFOMIS (Institute for Formal Ontology and Medical Information Science), Saarland University, Germany and
Department of Philosophy, University at Buffalo, NY, USA

^bECOR (European Centre for Ontological Research), Saarland University, Germany

1. Introduction

The automatic integration of rapidly expanding information resources in the life sciences is one of the most challenging goals facing biomedical research today. Controlled vocabularies, terminologies, and coding systems play an important role in realizing this goal, by making it possible to draw together information from heterogeneous sources – for example pertaining to genes and proteins, drugs and diseases – secure in the knowledge that the same terms will also represent the same entities on all occasions of use. In the naming of genes, proteins, and other molecular structures, considerable efforts are under way to reduce the effects of the different naming conventions which have been spawned by different groups of researchers. Electronic patient records, too, increasingly involve the use of standardized terminologies, and tremendous efforts are currently being devoted to the creation of terminology resources that can meet the needs of a future era of personalized medicine, in which genomic and clinical data can be aligned in such a way that the corresponding information systems become interoperable.

Unfortunately, however, these efforts are hampered by a constellation of social, psychological legal and other forces, whose countervailing effects are magnified by constant increases in available data and computing power. Patients, hospitals and governments are reluctant to share data; physicians are reluctant to use computerized forms in preparing patient reports; nurses, physicians and medical researchers in different specialities each insist on using their own terminologies, addressing needs which are rarely consistent with the needs of information integration.

Here, however, we are concerned with obstacles of another type, which have to do with certain problematic design choices made thus far in the development of the data and information infrastructure of biomedicine. The standardization of biomedical terminologies has for some years been proceeding apace. Standardized terminologies in biomedicine now exist in many flavours, and they are becoming increasingly important in a variety of domains as a result of the increasing importance of computers and of the need by computers for regimented ways of referring to objects and processes of different kinds. The Unified Medical Language System (UMLS), designed to “*facilitate the development of computer systems that behave as if they ‘understand’ the meaning of the language of biomedicine and health*” (NLM

2004), contains over 100 such systems in its MetaThesaurus (NLM 2004a), which comprehends some 3 million medical and biological terminological units. Yet very many of these systems are, as we shall see, constructed in such a way as to hamper the progress of biomedical informatics.

2. International Standard Bad Philosophy

Interestingly, and fatefully, many of the core features which serve as obstacles to the information alignment that we seek can be traced back to the influence of a single man, Eugen Wüster (1898-1977), a Viennese saw-manufacturer, professor of woodworking machinery, and devotee of Esperanto, whose singular importance turns on the fact that it was he who, in the middle of the last century, founded the technical committee devoted to terminology standardization of the International Organization for Standardization (ISO). Wüster was almost single-handedly responsible for all of the seminal documents put forth by this committee, and his ideas have served as the basis for almost all work in terminology standardization ever since.

ISO is a quasi-legal institution, in which earlier standards play a normative role in the formulation of standards which come later. The influence of Wüster's ideas has thus been exerted in ever wider circles into the present day, and it continues to make itself felt in many of the standards being promulgated by ISO not only in the realm of terminology but also in fields such as healthcare and computing.

Unfortunately these ideas, which have apparently never been subjected to criticism by those involved in ISO's work, can only be described as a kind of International Standard Bad Philosophy. Surveying these ideas will thus provide us with some important insights into a hitherto unnoticed practical role played by considerations normally confined to the domain of academic philosophy, and will suggest ways in which a good philosophy of language can help us develop and nurture better scientific terminologies in the future.

We surmise further that Wüster's ideas, or very similar ideas which arose independently, could be embraced by so many in the fields of artificial intelligence, knowledge modelling, and nowadays in Semantic Web computing, because the simplification in our understanding of the nexus of mind, language and reality which they represent answers deep needs on the side of computer and information scientists. In subjecting Wüster's ideas to critical analysis, therefore, we shall also be making a contribution to a much larger project of exploring possibilities for improvement in the ways in which computers are used in our lives.

3. Terminologies and Concept Orientation

The thinking of ISO Technical Committee (TC) 37 is that of the so-called Vienna School of Terminology, of which Wüster (1991) and Felber (1984) are principal movement texts (for a survey see Temmerman (2000, chapter 1)). Terminology, for Wüster and Felber, starts out from what are called *concepts*. The document ISO CD 704.2 N 133 95 EN, which bears the stamp of Wüster's thinking, explains what concepts are in psychological terms. When we experience reality, we confront two kinds of objects: the concrete, such as a tree or a machine, and the abstract, such as society, complex facts, or processes:

As soon as we are confronted with several objects similar to each other (all the planets in the solar system, all the bridges or societies in the world), certain essential properties common to these objects can be identified as characteristics of the general concept. These characteristics are used to delimit concepts. On the communicative level, these concepts are described by definitions and represented by terms, graphic symbols, etc. (ISO CD 704.2 N 133 95 EN)

A concept itself, we read in the same text, is "a unity of thought made up of characteristics that are derived by categorizing objects having a number of identical properties." To

understand this and the many similar sentences in 150 documents, we need to understand what is meant by ‘characteristic’. On the one hand, again in the same ISO text, we are told that a characteristic is a property that we identify as common to a set of objects. In other texts of ISO, however (for example in ISO 1087-1), we are told that a characteristic is a “mental representation” of such a property. This uneasy straddling of the boundary between world and mind, between property and its mental representation, is a feature of all of ISO’s work on terminology, as it was a feature of Wüster’s own thinking. Terminology work is seen as providing clear delineations of concepts in terms of characteristics as thus (confusingly) defined. When such delineations have been achieved, then terms can be assigned to the corresponding concepts. Wüster talks in this connection of a ‘realm’ (*Reich*) of concepts and of a ‘realm’ of terms (Wüster 1991, p. 1), the goal being that each term in a terminology should be associated with one single concept through “permanent assignment” (Felber 1984, p. 182).

4. Problems with the Concept-Based View of Terminologies

The above should seem alien to those familiar with the domain of medicine, however, because there we often have to deal with classes of entities for which we are unable to identify characteristics which all their members share in common. Terms are often introduced for such classes of entities long before we have any clear delineation of some corresponding concept.

The reason for this miscalibration between the ISO view of terminology and the ways terms in medicine are actually used turns on the fact that the notion of concept which underlies the terminology standards of ISO TC 37 and its successors *has nothing to do with medicine at all*. As Temmerman points out (2000, p. 11), Wüster was ‘an engineer and a businessman ... active in the field of standardisation’ and was concerned primarily with the standardisation of *products*, entities of the sort which truly are such as to manifest characteristics identifiable in encounters of similars *because they have been manufactured as such*. Vocabulary itself is treated by Wüster and his TC 37 followers ‘as if it could be standardised in the same way as types of paint and varnish’ (Temmerman, p. 12).

In those areas – like manufacturing or trade – which were uppermost in the mind of Wüster and of TC37 in its early incarnations, the primary purpose of standardization is precisely to bring about a situation in which entities in reality (such as machine parts) are *required* to conform to certain agreed-upon standards. Such a requirement is of course quite alien to the world of medicine, where it is in every case the entities in reality which must serve as our guide and benchmark. However, even in medicine – for reasons which increasingly have to do not only with ISO edicts but also with the expectations of those involved in the development of software applications – terminologists have been encouraged to focus not on entities in reality but rather on the concepts putatively associated therewith. The latter, it is held, enjoy the signal advantage that they can be conveyed as input to computers. At the same time they can be identified as units of knowledge and thus serve as the basis for what is called ‘knowledge modelling’, a term which itself embodies what we believe is a fateful confusion of *knowledge* with the true and false *beliefs* to which, in a domain like medicine, many of the concepts in common use correspond.

Some critical remarks about certain conceptions in ISO TC 37 documents have been recently advanced (Areblad and Fogelberg 2003), and the proposed alternative certainly represents an advance on Wüster in its treatment of individual objects. As concerns what is general, however, this new work still runs together objects and concepts, identifying specific kinds or types of phenomena in the world with the general concepts created by human beings. In this way, like Wüster, it leaves itself with no benchmark in relation to which given concepts or concept-systems could be established as correct or incorrect. Moreover, it leaves no way of

doing justice to the fact that bacteria would still have properties different from those of trees even if there were no humans able to form the corresponding concepts.

The Kantian Confusion

We can get at the roots of the problem of Wüsterian thinking if we examine what ISO CD 704.2 N 133 95 EN has to say about individual particulars and the proper names associated with them:

If we discover or create a singular phenomenon (individual object, e.g., the planet Saturn, The Golden Gate Bridge, the society of a certain country), we form an individual concept in order to think about that object. For communication purposes we assign *names* to individual concepts so that we can talk or write about them.

When parents assign names to their children, according to this view, and when they use such names for purposes of communication with others *they are not talking about their children at all*. Rather, they are talking about certain individual concepts which they have formed in their minds. This confusion of objects and concepts is well known in the history of philosophy. It is called “Kantianism”.

Wüster and Felber and (sadly) very many of the proponents of concept-based terminology work who have followed in their wake, as also very many of those working in the field of what is called ‘knowledge representation’, are subject to this same Kantian confusion. One implication of the fact that one is unsure about whether one is dealing with objects or with concepts is that one writes unclearly. This, for example, is how Felber in his semi-official text on terminology (presenting ideas incorporated in relevant ISO terminology standards) defines what he calls a ‘part-whole definition’:

The description of the collocation of individual objects revealing their partitive relationships corresponds to the definition of concepts. Such a description may concern the composite. In this case the parts, of the composite are enumerated. It may, however, also concern a part. In this case the relationship to an individual object subordinate to the composite and the adjoining parts are indicated. (Felber, *op. cit.*, cited exactly as printed)

The Realist Alternative

The alternative to Kantianism in the history of philosophy is called realism, and we have argued in a series of papers that the improvement of biomedical terminologies and coding systems must rest on the use of a *realist ontology* as basis (Smith 2004, Fielding *et al* 2004, Simon *et al.* in press). Realist ontology is not merely able to help in detecting errors and in ensuring intuitive principles for the creation and maintenance of coding systems of a sort that can help to prevent errors in the future. More importantly still, it can help to ensure that the coding systems and terminologies developed for different purposes can be provided with a clear documentation (thus helping to avoid many types of errors), and that they can be made compatible with each other (thus supporting information integration). Note that we say ‘*realist ontology*’ (or alternatively, with Rosse and Mejino (2003), ‘*reference ontology*’) in order to distinguish ontology on our understanding from the various related things which go by this name in contexts such as knowledge representation and conceptual modelling.

Ontology, as conceived from the realist perspective, is not a software implementation or a controlled vocabulary. Rather, it is a theory of reality, a ‘*science of what is, of the kinds and structures of objects, properties, events, processes and relations in every area of reality*’ (Smith 2003). It is for our purposes here a theory of those higher-level categories which structure the biomedical domain, the representation of which needs to be both unified and coherent if it is to serve as the basis for terminologies and coding systems that have the requisite degree and type of interoperability.

Ontology in this realist sense is already being used as a means of finding inconsistencies in terminologies and clinical knowledge representations such as SNOMED (Ceusters W, Smith B. 2003; Ceusters *et al.* 2004; Bodenreider *et al.* 2005), the Gene Ontology (Smith, Köhler and Kumar 2004), or the National Cancer Institute Thesaurus (Ceusters, Smith and Goldberg, *in press*). The method has also proved useful in drawing attention to certain problematic features of the HL7 RIM, more precisely on its confused running together of acts, statements about acts, and the reports in which such statements are registered (Vizenor 2004). This makes the HL7 RIM inadequate as a model for electronic patient records (so that it is to be regretted that experiments in this direction are already taking place). On the positive side, it has been embraced by the Foundational Model of Anatomy and by the Open Biomedical Ontologies Consortium as a means whereby precise formal definitions can be provided for the top-level categories and relations used in terminologies, in a way that will both support automatic reasoning and be intelligible to those with no expertise in formal methods (Smith *et al.*, 2005).

5. Formal methods for coding systems

Biomedical terminologies or coding systems can be integrated together into larger systems, or used effectively within an EHR system (which means: without loss or corruption of information), only on the basis of a shared common framework of top-level ontological categories. Often one talks in this connection merely of the sort of regimentation that can be ensured through the use of languages such as XML, or through technologies such as RDF(S) (W3C 2004) or OWL (W3C 2004a) – ontology languages that currently enjoy wide support through their association with the Semantic Web project.

On closer inspection, however, one discovers that the ‘semantics’ that comes with languages like RDF(S) and OWL is restricted to that sort of specification of meaning that can be effected using the formal technique of mathematical *model theory*. This means that meanings are specified by associating with the terms and sentences of a language certain abstract set-theoretic structures in line with the understanding of semantics that has followed in the wake of Alfred Tarski’s ‘semantic’ definition of truth for artificial languages (Hodges n.d.). Model theory allows us to describe the minimal conditions that a world must satisfy in order for a ‘meaning’ (or ‘interpretation’ in the model-theoretic sense) to be assignable to every expression in an artificial language with certain formal properties. Unfortunately, however, entities in reality are hereby substituted by abstract mathematical constructs embodying only the properties shared in common by all such interpretations. A formal semantic theory makes as few assumptions as possible about the actual nature or intrinsic structure of the entities in an interpretation, in order to retain as much generality as possible. In consequence, however, the chief utility of such a theory is not to provide any deep analysis (or indeed any analysis at all) of the nature of the entities – for example of the biomedical kinds and instances – described by the language. Rather, the power of formal semantics resides at the logical level, above all in providing a technical way to determine which inferences are valid (Guha and Hayes 2002).

In our view, in contrast, the job of ‘semantics’ as this term is used in phrases such as ‘semantic interoperability’ is identical to that of ontology as traditionally understood. Thus it does not consist in the construction of simplified models for testing the validity of inferences. Rather, its task is to support the alignment of the different perspectives on reality embodied in different types of coding and classification systems; to this end it must provide us with a *common reference framework* which mirrors the structures of those entities in reality to which these different perspectives relate.

6. Basic Formal Ontology

One such reference framework, which has been developed by the Institute of Formal Ontology and Medical Information Science in Saarbrücken, is Basic Formal Ontology (BFO) (Grenon and Smith 2004, Grenon *et al.* 2004), one of several closely related ontological theories proposed in the recent literature of realist ontology (for a survey and comparison see Masolo *et al.* 2004). BFO rests on the idea that it is necessary to develop an ontology that remains as close as possible to widely shared intuitions about objects and processes in reality.

It consists in a number of sub-ontologies, the most important of which are:

- SNAP, ontologies indexed by time instants and analogous to instantaneous snapshots of what exists at a given instant
- SPAN, ontologies indexed by time intervals and analogous to videoscopic representations of the processes unfolding across a given interval

– corresponding to the fundamental division between *continuants* (entities, such as organisms or blood corpuscles, which endure self-identically through time), and *occurrents* (processes, such as heart bypass surgeries or increases in temperature, which can be divided along the temporal axis into successive phases). Each SNAP ontology is a partition of the totality of objects and their continuant qualities, roles, functions, etc., existing in a given domain of reality at a given time. Each SPAN ontology is a partition of the totality of processes unfolding themselves in a given domain across a given temporal interval. SNAP and SPAN are complementary in the sense that, while continuants alone are visible in the SNAP view and the occurrents in which they are involved are visible only in the SPAN view, continuants and occurrents themselves exist only in mutual dependence on each other.

SNAP and SPAN serve as the basis for a series of sub-ontologies at different levels of granularity reflecting the fact that the same portion of reality can be apprehended in an ontology at a plurality of different levels of coarser or finer grain from whole organisms to single molecules. What appears as a single object at one level may appear as a complex aggregate of smaller objects at another level. What is a tumour at one level may appear as an aggregate of cells or molecules at another level. What counts as a unitary process at one level may be part of a process-continuum at another level. Since no single ontology can comprehend the whole of reality at all levels of granularity, each of the ontologies here indicated is thus partial only (Kumar *et al.* 2004).

Dependent entities, both within the SNAP and within the SPAN ontologies, are entities which require some other entity or entities which serve as their bearers. Dependent entities can be divided further into relational (for entities – such as processes of infection – dependent on a plurality of bearers) and non-relational (for entities – such as a rise in temperature – dependent on a single bearer).

Processes are examples of dependent entities on the side of occurrents: they exist always only as processes *of* or *in* some one or more independent continuants which are their bearers. Qualities, roles, functions, shapes, dispositions, and powers are examples of dependent entities on the side of continuants: they exist always only as the qualities (etc.) of specific independent continuants as their bearers: a smile smiles only in a human face; the function of your heart exists only when your heart exists.

Universals and particulars: Entities in all categories in the BFO ontology exist both as universals and particulars. You are a particular human being, and you instantiate the universal *human being*; you have a particular temperature, which instantiates the universal *temperature*;

you are currently engaging in a particular reading act, which instantiates the universal *reading act*. In each case we have a certain universal and an associated plurality of instances, where the term ‘instance’ should be understood here in a non-technical way, to refer simply to those objects, events and other entities which we find around us in the realm of space and time (and thus not, for example, to entries, records or values in databases). ‘Universal’, too, connotes something very simple, namely the general kinds or patterns which such particular entities have in common. Thus to talk of the universal *red* is just to talk of that which this tomato and that pool of ink share in common; to talk of the universal *aspirin* is to talk of that which these aspirin pills and those portions of aspirin powder share in common. That universals in this sense exist should be uncontroversial: it is universals which are investigated by science. It is in virtue of the existence of universals that medical diagnoses are able to be formulated by using general terms, and that corresponding standardized therapies can be tested in application to pluralities of different cases (instances) existing at different times and locations (Swoyer 1999).

Again, in part because of the influence of Wüsterian thinking, both universals and particulars have been poorly treated in biomedical terminologies and in electronic health records thus far. While biomedical terminologies ought properly to be constructed as inventories of the universals in the corresponding domains of reality (Smith *et al.* 2005), they have been conceived instead as representations of the concepts in peoples’ heads. While electronic health records ought properly to be constructed as inventories of the instances salient to the health care of each given patient (including particular disorders, lesions, treatments, etc.), they have in fact been put together in such a way that in practice only human beings (patients, physicians, family members) are represented on the level of instances, information about all other particular entities being entered in the form of general codes – in ways which cause the problems outlined in (Ceusters and Smith 2005). Instances have also been inadequately treated in the various logical tools used in the fields of terminology and EHR. (The Tarskian approach referred to above encourages, again, the logical treatment, not of actual particular entities in corporeal reality, but rather of those abstract mathematical surrogates for such entities which are created *ad hoc* for the logician’s technical purposes.)

Ontology and epistemology: The BFO framework distinguishes, further, between ontology and epistemology. The former is concerned with reality itself, the latter with our ways of gaining knowledge of reality. These ways of gaining knowledge can themselves be subjected to ontological treatment: they are *processes* of a certain sort, with cognitive agents as their continuant bearers. This fact, however, should not lead us to confuse epistemological issues (pertaining to *what and how we can know*) with ontological issues (pertaining to *how the world is*). Thus ‘finding’ is a term which belongs properly not to ontology but rather to epistemology, and so also do UMLS terms such as ‘experimental model of disease’. It is the failure to distinguish clearly between ontology and epistemology – a failure that is comparable in its magnitude to the failure to distinguish, say, between physics and its history or between eating and the description of food) – which is at the root of the confusions in Wüster/ISO thinking, and in almost all contemporary work on terminologies and knowledge representation and which leads for example to the identification of *blood pressure* with *result of laboratory measurement* or of *individual allele* with *observation of individual alleles*.

Already a very superficial analysis of a coding system like the ICD (for: International Classification of Diseases: World Health Organization, n.d.) reveals that this system is not in fact a classification of *diseases* as entities in reality (Bodenreider *et al.* 2004). Rather it is a classification of *statements* on the part of a physician *about disease phenomena which the physician might attribute to a patient*. As an example, the ICD-10 class *B83.9: Helminthiasis*,

unspecified does not refer (for example) to a disease caused by a worm belonging to the species *unspecified* (some special and hitherto uninvestigated sub-species of *Acanthocephalia* or *Metastrongylia*). Rather, it refers to a statement (perhaps appearing in some patient record) made by a physician who for whatever reason did not specify the actual type of *Helminth* which caused the disease the patient was suffering from. Neither OWL nor reasoners using models expressed in OWL would complain about making the class *B83.9: Helminthiasis, unspecified* a subclass of *B83: Other helminthiasis*; from the point of view of a coherent ontology, however, such a view is nonsense: it rests, again, on the confusion between ontology and epistemology.

A similar confusion can be found in EHR architectures, model specifications, message specifications or data types for EHR systems. References to a patient's gender/sex are a typical example (Milton 2004). Some specifications, such as the Belgian KMEHR system for electronic healthcare records (Kmehrer-Bis, n.d.) include a classification of what is called "administrative sex" (we leave it to the reader to determine what this term might actually mean). The possible specifications of *administrative sex* are then *female*, *male*, *unknown*, or *changed*. *Unknown*, here, does not refer to a new and special type of gender (reflecting some novel scientific discovery); rather it refers merely (but of course confusingly) to the fact that the actual gender is not documented in the record.

7. An Ontological Basis for Coding Systems and the Electronic Health Record

Applying BFO to coding systems and EHR architectures means, in the first place, applying it to the salient entities in reality – to actual patients, diseases, therapies – with the goal of making coding systems more coherent, both internally and in their relation to the EHRs which they were designed to support. But it is essential to this endeavour that we establish also the proper place in reality of coding systems and EHRs themselves, and that we understand their nature and their purposes in light of a coherent ontological theory. Coding systems are in fact as real as the words we speak or write and as the patterns in our brains, and we can use the resources of a framework like BFO in order to analyze how both coding systems and EHRs relate a single reality in a way which is compatible with what is known informally by the patients, physicians, nurses, etc. toward whom they are directed.

Referent tracking is a new paradigm for achieving the faithful registration of patient data in electronic health records, focusing on *what is happening on the side of the patient* (Ceusters W., Smith B. 2005a) rather than on *statements made by clinicians* (Rector *et al.* 1991). The goal of referent tracking is to create an ever-growing pool of data relating to concrete entities in reality. In the context of Electronic Healthcare Records (EHRs) the relevant concrete entities, i.e. particulars as described above, are not only particular patients but also their body parts, diseases, therapies, lesions, and so forth, insofar as these are relevant to their diagnosis and treatment. Within a referent tracking system (RTS), all such entities are referred to explicitly, something which cannot be achieved when familiar concept-based systems are used in what is called "clinical coding" (Ceusters W., Smith B. 2005b).

By fostering the accumulation of prodigious amounts of instance-level data along these lines, including also considerable quantities of redundant information (since the same information about given instances will often be entered independently by different physicians), which can be used for cross-checking, the paradigm allows for a better use of coding and classification systems in patient records by minimizing the negative impact that mistakes in these systems have on the interpretation of the data.

The users who enter information in a RTS will be required to use IUIs (Instance Unique Identifiers) in order to assure explicit reference to the particulars about which the

information is provided. Thus the information that is currently captured in the EHR by means of sentences such as: “this patient has a left elbow fracture”, would in the future be conveyed by means of descriptions such as “#IUI-5089 is located in #IUI-7120”, together with associated information for example to the effect that “IUI-7120” refers to the patient under scrutiny or that “IUI-5089” refers to a particular fracture in patient #IUI-7120 (and not to some similar left elbow fracture from which he suffered earlier). The RTS must correspondingly contain information relating particulars to universals, such as “#IUI-5089 is a fracture” (where ‘fracture’ might be replaced by a unique identifier pointing to the representation of the universal *fracture* in an ontology). Of course, EHR systems that endorse the referent tracking paradigm should have mechanisms to capture such information in an easy and intuitive way, including mechanisms to translate generic statements into the intended concrete form, which may itself operate primarily behind the scenes, so that the IUIs themselves remain invisible to the human user. One could indeed imagine that natural language processing software will one day be in a position to replace in a reliable fashion the generic terms in a sentence with corresponding IUIs for the particulars at issue, with the need for manual support flagged only in problematic cases. This is what users already expect from EHR systems in which data are entered by resorting to general codes or terms from coding systems.

If the paradigm of referent tracking is to be brought into existence, at least the following requirements have to be addressed:

- a mechanism for generating IUIs that are guaranteed to be unique strings;
- a procedure for deciding what particulars should receive IUIs;
- protocols for determining whether or not a particular has already been assigned a IUI (except for some exceptional configurations that are beyond the scope of this paper, each particular should receive maximally one IUI);
- practices governing the use of IUIs in the EHR (issues concerning the syntax and semantics of statements containing IUIs);
- methods for determining the truth values of propositions that are expressed through descriptions in which IUIs are used;
- methods for correcting errors in the assignment of IUIs, and for investigating the results of assigning alternative IUIs to problematic cases;
- methods for taking account of changes in the reality to which IUIs get assigned, for example when particulars merge or split.

An RTS can be set up in isolation, for instance within a single general practitioner’s surgery or within the context of a hospital. The referent tracking paradigm will however serve its purpose optimally only when it is used in a distributed, collaborative environment. One and the same patient is often cared for by a variety of healthcare providers, many of them working in different settings, and each of these settings uses its own information system. These systems contain different data, but the majority of these data provide information about the same particulars. It is currently very hard, if not impossible, to query these data in such a way that, for a given particular, all information available can be retrieved. With the right sort of distributed RTS, such retrieval becomes a trivial matter.

This, in turn, will have a positive impact on the future of biomedicine in a number of different ways. Errors will be more easily eliminated or prevented via reminders or alerts issued by software agents responding to changes in the referent tracking database. It will also become possible to coordinate patient care between multiple care organisations in more efficient ways. An RTS will also do a much better job in fulfilling the goals of the ICD and its precursors, namely to enable information integration for public health. It can help specifically in the

domain of disease surveillance, an area of vital concern on a global scale that has the potential not only to improve the quality of care but also to provide a means for controlling costs, in particular by promoting effective cooperation among healthcare professionals for continuity of care.

8. Toward the Future

European and international efforts towards standardization of biomedical terminology and electronic healthcare records have been focused over the last 15 years primarily on syntax. Semantic standardization has been restricted to issues pertaining to knowledge representation (and resting primarily on the application of set-theoretic model theory, along the lines described in section 5. above). Moves in these directions are indeed required, and the results obtained thus far are of value both for the advance of science and for some concrete uses of healthcare informatics applications. But we can safely say that the syntactical issues are now in essence resolved. The semantic problems relating to biomedical terminology (polysemy, synonymy, cross-mapping of terminologies, and so forth), too, are well understood – at least in the community of specialized researchers. Now, however, it is time to *solve* these problems by using the theories and tools that have been developed so far, and that have been tested under laboratory conditions (Simon *et al.* 2004). This means using the right sort of ontology, i.e. an ontology that is able explicitly and unambiguously to relate coding systems, biomedical terminologies and electronic health care records (including their architecture) to the corresponding instances in reality.

To do this properly will require a huge effort, since the relevant standards need to be reviewed and overhauled by experts who are familiar with the appropriate sorts of ontological thinking (which will require some corresponding effort in training and education). Even before that stage is reached, however, there is the problem of making all constituent parties – including patients (or at least the organizations that represent them), healthcare providers, system developers and decision makers – aware of how deep-seated the existing problems are. Having been overwhelmed by the exaggerated claims on behalf of XLM and similar silver bullets of recent years, they must be informed that XML, or Descriptive Logic, or OWL, or even the entire Semantic Web, can take us only so far. And of course we must also be careful to avoid associating similarly exaggerated expectations with realist ontology itself. It, too, can take us only so far.

The message of realist ontology is that, while there are various different views of the world, this world itself is one, and that this one world, because of its immense complexity, is accessible to us only by a corresponding variety of different sorts of views. It is our belief that it is only through reference to this world that the various different views can be compared and made compatible (and not by reference to ethereal entities in some ‘realm of concepts’). To allow clinical data registered in electronic patient records by means of coding (and/or classification) systems to be used for further automated processing, it should be crystal clear whether entities in the coding system refer to diseases or to statements made about diseases, to acts on the part of physicians or to documents in which such acts are recorded, to procedures and observations or to statements about procedures or observations. As such, the coding systems used in electronic healthcare records should be associated with a precise and formally rigorous ontology that is coherent with the ontology of the healthcare record as well as with those dimensions of the real world that are described therein. And they should be consistent, also, not with information models concocted by database designers from afar, but rather with the common-sense intuitions about the objects and processes in reality which are shared by patients and healthcare providers.

9. Recommendations

Concrete recommendations for further progress thus include the following:

1. Given that most existing international standards in terminology and related fields were created at a time when the requirements for good ontologies and good controlled vocabularies were not yet clear, efforts should be made to inform people of the urgent need for more up-to-date and more coherent standards.
2. The work of ISO TC 37 (on terminologies) and of the technical committees which have fallen under its sway (CEN/TC251, ISO/TC215, etc.) should be subjected to a radical evaluation from the point of view of coherence of method, intelligibility of documentation, consistency of views expressed, usability of proposed standards, methods for testing, and quality assurance.
3. Through collaboration between users and developers, objective measures should be developed for the quality of ontologies.
4. By applying these quality measures, a publicly available top-level ontology should be developed on the basis of consensus among the major groups involved in biomedical ontology development, almost all of whom are present within the EU; this top-level ontology should be complemented with extensions for biomedicine and bio-informatics.
5. Objective measures should be developed for ascertaining the quality of tools designed for the support of information integration in such a way that, when resources are invested in the development of ontologies and associated software in the future, clear thresholds of success can be formulated and corresponding standards of accountability imposed.
6. Existing terminologies and ontologies should be assessed for their compatibility with the major top-level ontologies, and efforts should be devoted to ensuring such compatibility in the future.
7. Principles should be established setting forth the appropriate use of ontologies in EHR systems, including investigations of the merits of systems which, in addition to general terms from coding systems, also incorporate reference to particulars in a systematic way.
8. The ontological mistakes in the HL7 RIM should be thoroughly documented and modifications should be proposed to make the HL7 approach consistent with a faithful treatment of the different kinds of entities that exist in the domain of healthcare and are relevant for patient data collection and for the communication of information content between healthcare institutions.
9. A Europe-wide institution should be developed for the coordination of ontology research and knowledge transfer in order to promote high-quality work and to avoid redundancy in investment of ontology-building efforts. Open competitions should be developed which are designed to find the best methodologies for harvesting healthcare data, with real gold standards and real measures of success governing applications of the results to clinical care and public health, integration with genomics-based data to develop personalized care, integration with the data gathered by third parties, e.g. by drug companies.

10. Conclusion

We have argued that what is needed if we are to support the kind of information integration to which we all aspire is not more or better *information models* but rather a *theory of the reality* to which both coding systems and electronic health records are directed.

Applying a sound realist ontology to coding systems and to EHR architectures means in the first place ensuring that the latter are calibrated not to the denizens of Wüster's 'realm of concepts' but rather to those entities in reality – such as particular patients, diseases, therapies, surgical acts, and the universals which they instantiate – which form the subject matter of healthcare. In this way we can make coding systems more coherent, both internally and in their relation to the EHRs which they are designed to support, and externally in relation to the patients, physicians, nurses, etc. toward whom they are directed.

Acknowledgments: Work on this paper was carried out under the auspices of the Alexander von Humboldt Foundation, the EU Network of Excellence in Medical Informatics and Semantic Data Mining, and the Project "Forms of Life" sponsored by the Volkswagen Foundation.

References

- Areblad M, Fogelberg M. 2003 "Comments to ISO TC 37 in the revision of ISO 704 and ISO 1087." CEN/TC251 WGII/N03-17 2003-08-27.
- Barry Smith, Werner Ceusters, Bert Klagges, Jacob Köhler, Anand Kumar, Jane Lomax, Chris Mungall, Fabian Neuhaus, Alan Rector, Cornelius Rosse 2005 "Relations in Biomedical Ontologies", *Genome Biology*, 2005, 6 (5), R46.
- Bodenreider O, Smith B, Kumar A, Burgun A. 2005. (Forthcoming) "Investigating subsumption in DL-based terminologies: A case study in SNOMED CT." in *Artificial Intelligence in Medicine*, forthcoming.
- Bodenreider, Olivier, Smith, Barry, Burgun, Anita 2004 "The Ontology-Epistemology Divide: A Case Study in Medical Terminology", *Third International Conference on Formal Ontology (FOIS) 2004*, 185-195.
- Ceusters W, Smith B, Kumar A, Dhaen C. 2004 Mistakes in Medical Ontologies: Where Do They Come From and How Can They Be Detected? in Pisanelli DM (ed.) *Ontologies in Medicine. Proceedings of the Workshop on Medical Ontologies, Rome October 2003*, Amsterdam: IOS Press, Studies in Health Technology and Informatics, vol 102, 145–64.
- Ceusters W, Smith B. 2003 "Ontology and Medical Terminology: Why Descriptions Logics are not enough", *Proceedings of the Conference Towards an Electronic Patient Record (TEPR 2003)*, San Antonio, 10-14 May 2003 (electronic publication).
- Ceusters W, Smith B. 2005a "Referent Tracking in Electronic Healthcare Records". Accepted for MIE 2005, Geneva, 28-31 Augustus 2005.
- Ceusters W, Smith B. 2005b. Strategies for Referent Tracking in Electronic Health Records. (Download draft). Proceedings of IMIA WG6 Conference on "Ontology and Biomedical Informatics". Rome, Italy, 29 April - 2 May 2005. (in press).
- Felber, H. 1984 *Terminology Manual*. Unesco: International Information Centre for Terminology (Infoterm), Paris.
- Fielding, James M., Simon, Jonathan, Ceusters, Werner and Smith, Barry 2004 "Ontological Theory for Ontological Engineering: Biomedical Systems Information Integration", *Proceedings of the Ninth*

- International Conference on the Principles of Knowledge Representation and Reasoning (KR2004)*, Whistler, BC, 2-5 June 2004, 114-120.
- Grenon, Pierre and Smith, Barry 2004 “SNAP and SPAN: Towards Dynamic Spatial Ontology”, *Spatial Cognition and Computation*, 4: 1, 69–103.
- Grenon, Pierre and Smith, Barry and Goldberg, Louis 2004 “Biodynamic Ontology: Applying BFO in the Biomedical Domain”, in D. M. Pisanelli (ed.), *Ontologies in Medicine: Proceedings of the Workshop on Medical Ontologies, Rome October 2003*, Amsterdam: IOS Press, 20–38.
- Guha RV, Hayes P. 2002 “LBase: Semantics for Languages of the Semantic Web. NOT-A-Note” 02 Aug 2002 (<http://www.coginst.uwf.edu/~phayes/LBase-from-W3C.html>)
- Hodges, Wilfrid (n.d.) “Model Theory”, *Stanford Encyclopedia of Philosophy* (<http://plato.stanford.edu/entries/model-theory/>).
- ISO 2002 ISO 18308: *Health Informatics – Requirements for an Electronic Health Record Architecture*. (<http://www.iso.ch/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=33397>).
- Jonathan Simon, James Fielding, Mariana Dos Santos, Barry Smith, “Reference Ontologies for Biomedical Ontology Integration and Natural Language Processing”, *International Journal of Medical Informatics*, in press.
- Kmehr-Bis n.d. “Kind Messages for Electronic Healthcare Record, Belgian Implementation Standard” (<http://www.chu-charleroi.be/kmehr/htm/kmehr.htm>).
- Kumar, Anand, Smith, Barry and Novotny, Daniel 2004 “Biomedical Informatics and Granularity”, *Comparative and Functional Genomics*, 5, 501–508.
- Masolo C., Borgo S., Gangemi A., Guarino N., Oltramari A. 2004 *WonderWeb Deliverable D18: Ontology Library* (<http://wonderweb.semanticweb.org/deliverables/documents/D18.pdf>).
- Milton, Simon K. 2004 “Top-Level Ontology: The Problem with Naturalism”, in Achille Varzi and Laure Vieu (eds.), *Formal Ontology and Information Systems. Proceedings of the Third International Conference (FOIS 2004)*, Amsterdam: IOS Press, 2004, 85–94.
- NLM (National Library of Medicine) 2004 UMLS fact sheet, updated 7 May 2004 (<http://www.nlm.nih.gov/pubs/factsheets/umls.html>).
- NLM (National Library of Medicine) 2004a UMLS MetaThesaurus fact sheet, updated 7 May 2004. (<http://www.nlm.nih.gov/pubs/factsheets/umlsmeta.html>).
- Rector AL, Nolan WA, and Kay S. Foundations for an Electronic Medical Record. *Methods of Information in Medicine* 30: 179-86, 1991.
- Rosse, C. and Mejino, J. L. V. Jr. 2003 “A Reference Ontology for Bioinformatics: The Foundational Model of Anatomy”, *J Biomed Informatics*, 36:478–500.
- Simon, Jonathan, Fielding, James M. and Smith, Barry 2004 “Using Philosophy to Improve the Coherence and Interoperability of Applications Ontologies: A Field Report on the Collaboration of IFOMIS and L&C”, in Gregor Büchel, Bertin Klein and Thomas Roth-Berghofer (eds.), *Proceedings of the First Workshop on Philosophy and Informatics. Deutsches Forschungszentrum für künstliche Intelligenz*, Cologne, 65–72.
- Smith, Barry , Köhler, Jacob, Kumar, Anand 2004 “On the Application of Formal Principles to Life Science Data: a Case Study in the Gene Ontology,” In: Erhard Rahm (Ed.): *Data Integration in the Life Sciences, First International Workshop, DILS 2004*, Leipzig, Germany, March 25-26, 2004, (Lecture Notes in Computer Science 2994), Springer, 79–94.
- Smith, Barry 2003 “Ontology”, in Luciano Floridi (ed.), *Blackwell Guide to the Philosophy of Computing and Information*, Oxford: Blackwell, 155–166.
- Swoyer, Chris 1999 “How Ontology Might be Possible: Explanation and Inference in Metaphysics,” *Midwest Studies in Philosophy*, 23, 1999; 100–131.
- Temmerman R. 2000 *Towards New Ways of Terminology Description*. Amsterdam: John Benjamins.
- Vizenor, Lowell 2004 “Actions in Health Care Organizations: An Ontological Analysis”, *Proceedings of MedInfo 2004*, San Francisco, 1403–10.

W3C 2004 “RDF Semantics”. W3C Recommendation 10 February 2004 (<http://www.w3.org/TR/rdf-nt/>).

W3C 2004a “OWL Web Ontology Language Semantics and Abstract Syntax”. W3C Recommendation 10 February 2004 (<http://www.w3.org/TR/owl-semantics/>).

Werner Ceusters and Barry Smith 2005 “Tracking Referents in Electronic Health Records”, *Medical Informatics Europe* (MIE 2005), Geneva.

Werner Ceusters, Barry Smith, Louis Goldberg (in press) “A Terminological and Ontological Analysis of the NCI Thesaurus”, *Methods of Information in Medicine*.

World Health Organisation n.d. ICD-10 - The International Statistical Classification of Diseases and Related Health Problems, tenth revision (<http://www.who.int/whosis/icd10/>).

Wüster, E. 1991 *Einführung in die allgemeine Terminologielehre und terminologische Lexikographie*. Bonn: Romanistischer Verlag, Germany.