

Referent Tracking for Corporate Memories

¹Werner CEUSTERS, ^{1,2} Barry SMITH

¹ *Ontology Research Group, NY State Center of Excellence in Bioinformatics and Life Sciences,
701 Ellicott Street, Buffalo NY, 14203.*

² Department of Philosophy, University at Buffalo,
ceusters@buffalo.edu, phismith@buffalo.edu

1 Introduction

Corporate memories (CM) are information systems designed to keep track of the history and evolution of an enterprise with the goal of using past experience to enhance performance in the future. Well designed CMs should contain data about both the enterprise and the environment in which it operates. The former, traditionally embodied in what is referred to as an ‘enterprise model’, consists of data about the organisational structure and operating procedures of the enterprise, its mission and strategic objectives, its staff, their skills and competences, the products and services the company is able to deliver, and, most importantly, data about projects or business transactions brought to a successful (or unsuccessful) end. A CM’s environment model includes data about prospects and clients, competitors and partners, applicable laws and regulations, and techniques and methodologies proposed by scholars in the field to complement the results of research carried out by the company itself.

For understandable reasons, CM technology is standardly approached from a backward-looking perspective, employing passive knowledge management techniques with the prime goal of making electronic documents more easily accessible. To this end, documents are manually or semi-automatically annotated with tags that reformulate words or relevant phrases in a document in a more structured and standardised manner (e.g. occurrences of the words *car*, *van*, *bus*, etc. are all tagged with the compound *motor vehicle*), or with meta-tags that add additional context to phrases or paragraphs (e.g. *important*, *motivation*, *ignore*, etc.). Tags and meta-tags are designed to enable retrieval of documents or document sections on the basis of queries issued by users with specific information needs. Ontologies, including ontologies developed within the framework of the Semantic Web, have been proposed in this context as a means of organising and standardizing these meta-tags in such a way as to create more powerful CM applications which would work over corporate networks in such a way as to link together various heterogeneous systems. When these meta-tags are organised in a structure that reflects more or less the way the enterprise is structured, they form what is referred to as ‘enterprise ontologies’.

CM applications can also, however, be used for the development of more pro-active systems, in which data that reflect changes in either the organisation or its environment are able to trigger warnings indicating business opportunities for or, at the opposite extreme, imminent hazards to the proper functioning of the enterprise. We describe below the technology which can support such CM pro-activity, and which is designed to allow CM technology to exploit automatically the lessons learned from past experiences for the future benefit of the enterprise. The same technology can also be applied to other purposes. Thus it can be used to audit CM systems for quality assurance purposes and also to compare the CMs of separate enterprises (or heterogeneous parts of a single enterprise) in ways which may be useful for purposes of due diligence for example in assessing the feasibility of potential collaborations or company mergers.

To achieve these goals, CM applications must be freed from the barriers placed around them by traditional knowledge management paradigms, which means that they must be required to mirror more faithfully those portions of reality which are salient to the workings of the enterprise, including the changes that occur as time passes by. The purpose of this chapter is to demonstrate how a number of foundational technologies that are currently proposed to build proactive electronic healthcare systems can be used to achieve analogous objectives in the area of enterprise ontologies and corporate memories.

2 Background

2.1 Corporate Memories

The word ‘corporate memory’, including its quasi-synonym ‘organisational memory’, is interchangeably used to denote distinct though related entities. Originally, the term referred to a specific type of ‘collective memory’ found in organisations and groups, primarily commercial enterprises, and which, according to social and behavioural scientists descending from Durkheim, is something supra-individual which cannot be reduced to the memories in the minds of single individuals (Wexler, 2002). Collective memory so conceived typically covers various kinds of information about (1) external contacts, (2) internal know-how, (3) the types of authority and influence exerted not only by company owners but also employee associations, (4) the behavior of customers, (5) operational rule sets and routines, and (6) operations implementation strategies that determine how the information about how all of these things interact with the company’s primary business (Beckett, 2000).

With the advance of computer technology, however, this original meaning has been superseded by notions of collective memory that refer to physical embodiments such as a *representation* of the knowledge and information in an organisation (Van Heijst, Van der Spek, & Kruizinga, 1996), or to the *repository* in which such a representation is stored (Prasad & Plaza, 1996). Other authors put more emphasis on dynamic aspects of corporate memory, conceiving the latter as the computer system which embodies a company’s entire stock of knowledge assets, including accumulated know-how (skills), and makes them available to enhance the efficiency and effectiveness of knowledge-intensive work processes (Kühn & Abecker, 1997). While enhanced efficiency is of course the objective, some authors prefer to include failure to achieve this objective in their definition of corporate memory and therefore characterise it as: ‘*the means by which knowledge from the past which is stored in a repository is brought to bear on present activities, thereby resulting in higher or lower levels of organisational effectiveness*’ (Stein, 1995). Indeed, although the development of a corporate memory can prevent ‘*memory losses*’ (Benkard, 2000), its use can lead either to increased performance (Park & Bunn, 2003) or to a competitive disadvantage. The latter may result from putting too much emphasis on the organisational memory so that companies no longer notice changes in the environment and become trapped in habitual decision making (Berthon, Pitt, & Ewing, 2001) or from a failure to adapt existing roles and organisational structures to new situations (Walsh & Ungson, 1991).

How to build corporate memory systems is a research topic in its own right, since any such system has to be able to communicate with the majority of computer systems already installed in the company and to re-use the information they contain. Since this involves issues of semantic interoperability, it is no surprise that ontologies have become essential components of corporate memory systems, contributing to a wide variety of tasks, including determining content relevant for inclusion (Kalfoglou, Alani, O’Hara, & Shadbolt, 2002), describing intended use models (Vasconcelos, Kimble, Gouveia, & Kudenko, 2001), cross-mapping the data- and knowledge bases dispersed over an intranet that need to contribute to the corporate memory (Vasconcelos, Kimble, & Gouveia, 2000), assisting information retrieval (Kung-Jiuan & Yuh-Min, 2006), guiding information extraction (Baumann, Malburg, Hofe, & Wenzel, 1997) or steering software agents (Gandon, Poggi, Rimassa, & Turci, 2002). Most prominent, however, are the ontologies that describe organisational aspects of the enterprise, and are therefore called *enterprise ontologies*.

2.2 Enterprise Ontologies

Where corporate memories capture primarily what is *specific* for an enterprise, such as information about its employees, projects, business rules, contracts, and so forth, enterprise ontologies capture primarily what is *generic*. The first ontologies of this sort were developed in the course of The Enterprise Project in the United Kingdom (Stader, 1996) and the TOVE project in Toronto, Canada (Fox, 1992). Both initiatives formed a bridge between the older *enterprise modelling* paradigm (Schrieber, 1970) (Gershefski, 1970) and the newer paradigms of *enterprise engineering* and *enterprise architecture* as

interdisciplinary fields bringing together the engineering and strategic management disciplines and treating entire enterprises as products to be designed, deployed, and maintained (Liles & Presley, 1996).

One of the outcomes of the Enterprise project was *The Enterprise Ontology*, which is described by its authors as ‘a collection of terms and definitions relevant to business enterprises’ (Uschold, King, Moralee, & Zorgios, 1998). Approximately 90 terms are defined, grouped in 5 clusters labelled ‘activity’, ‘organisation’, ‘strategy’, ‘marketing’ and ‘time’.

The TOVE project comprehends several enterprise-related ontologies, including a *Resource Ontology*, aimed at describing the resources involved in production planning and scheduling (Fadel, Fox, & Gruninger, 1994), an *Activity Ontology* (Gruninger & Fox, 1994), a *Cost Ontology* (Tham, Fox, & Gruninger, 1994) reflecting the ABC (Activity Based Costing) paradigm (Tham, 1999), a *Quality Ontology* (Kim, Fox, & Gruninger, 1995) (Kim, Fox, & Gruninger, 1999), and an ontology for performing measurements (Kim, Sengupta, Fox, & Dalkilic, 2006).

Of more recent date is the *REA Enterprise Ontology* (Geerts & McCarthy, 2002), whose acronym derives from the framework’s original primary object components: economic Resources, economic Events, and economic Agents, and which is based on the REA accounting model (McCarthy, 1982).

2.3 Why are such systems not in use?

Despite the massive interest and research activities in corporate memories and enterprise ontologies, reports on success stories are limited to unverifiable marketing claims or mere speculations. This is witnessed by statements such as ‘*Many initiatives, governmental and commercial, have pursued the grand vision of “transparent access” – making all data available to all consumers (users and applications), in a way the consumer can interpret, anywhere and at any time. Among large-scale enterprises, success stories in achieving such visions seem rare or nonexistent*’ (Rosenthal, Manola, & Seligman, 2001) or papers such as (Hill, 2006), of which the title alone: ‘*Service Taxonomy and Service Ontologies Deliver Success to Enterprise SOA*’ merely implies the existence of actual success without however providing any evidence for it.

There are several reasons for this. In (Partridge & Stefanova, 2001), for instance, it is argued that neither the TOVE and Enterprise Ontology meet the criteria of clear characterisation and domain coverage, and that the latter cannot be compensated for by merging them because they do not share a common view of what an organisation is. This is then a motivation to embark on the development of a new ontology: the Core Enterprise Ontology (Bertolazzi, Krusich, & Missikoff, 2001). When this system was analysed by other scholars, however, then it too was found not to meet certain crucial requirements, which again led to the creation of a new artifact (Osterwalder, Lagha, & Pigneur, 2002). And so it goes on.

Our research on the (lack of) quality of ontologies developed in areas such as natural language understanding (Ceusters, Smith, & Fielding, 2004), biomedical informatics (Smith & Ceusters, 2003), healthcare information technology (Smith & Ceusters, 2006) and digital rights management (Ceusters & Smith, 2007) has demonstrated that the main reason for failure is the adoption of a modelling approach rooted in traditional expert-system-based approaches to knowledge representation and centered around the representation of ‘*concepts*’ (Smith, 2004) or ‘*conceptualisations*’ (Smith, Ceusters, & Temmerman, 2005). These approaches do not take reality seriously enough into account. This is because concepts in the minds of experts are always in one way or another simplifications of the reality to which they are intended to correspond. Representations of concepts in computer systems add a further level of simplification (and thus a further removal from reality) by imposing restrictions of expressivity needed to guarantee computational tractability of the systems which result. Indeed, when knowledge engineers and information analysts proceed by first defining concepts and relationships and then connecting them to some area of investigation, then they have things precisely the wrong way round. What they should be doing is finding a way to refer directly to the concrete real-world entities to which their systems must relate at the very start of their analysis and of making these entities serve as anchors for their analysis. Ignoring reality and focusing instead on concepts has given rise to several so-called ontologies in which the ontologies themselves, rather than reality, have become the objects of study, so that the quality of one ontology is gauged by the degree to which it conforms to a second ontology (Goossenaerts & Pelletier, 2003).

Focusing on reality, in contrast, can provide an independent benchmark for the correctness of an ontology, and facilities to reflect how changes introduced in successive versions of an ontology relate to changes in reality.

This, on realist grounds, unjustifiable focus on conceptualisations which deviate in substantial ways from the structures found in reality, seems also to hold for ontologies developed in the context of enterprise engineering. (Huhns & Stephens, 2002) for instance, describes a methodology under which a multiplicity of ontology fragments, representing the semantics employed by independent third parties, can be related together automatically without the use of a global ontology. On closer inspection of the examples provided, however, it becomes clear that the resulting ontology contains many erroneous associations. Interestingly, Huhns & Stephens do not consider this to be problematic. Indeed they assert that a '*consensus ontology is perhaps the most useful for information retrieval by humans, because it represents the way most people view the world and its information. For example, if most people wrongly believe that crocodiles are a kind of mammal, then most people would find it easier to locate information about crocodiles if it were located in a mammals grouping, rather than where it factually belonged*' (Huhns & Stephens, 2002, pp 89).

We argue that for ontologies ever to become useful in mission-critical domains like business or medicine, they must be built on the basis of an approach that is resolutely grounded in reality. Ontologies which are intended to be used more specifically in the context of enterprise engineering and corporate memory systems must be able to reflect not only how our perceptions and beliefs about reality change in the course of time but also how reality itself changes, and how the former are related to the latter. If for the purposes of a given ontology application it is judged relevant that many people believe that crocodiles are mammals, then this fact should indeed be represented in the ontology; but it should be represented as a *false belief*, rather than the subject of that belief as a natural fact.

In the following sections, we describe how to achieve these ends in such a way as to achieve a level of sophistication in ontology development that is able to draw a clear distinction between reality and our (managers', employers', customers') conceptualizations thereof.

3 Ontologies and faithfulness to reality

3.1 Basic Formal Ontology

The core of our proposal is Basic Formal Ontology (BFO), a framework that is designed to serve as basis for the creation of high-quality shared ontologies especially in the domain of natural science, and that embraces a methodology which is realist, fallibilist, perspectivalist, and adequatist (Grenon, Smith, & Goldberg, 2004). BFO holds, accordingly, (1) that reality and its constituents exist independently of our (linguistic, conceptual, theoretical, cultural) representations thereof, (2) that our theories and classifications can be subject to revision, (3) that there exists a plurality of alternative, equally legitimate perspectives on reality, and (4) that these alternative views are not reducible to any single basic view.

BFO subdivides reality into a number of basic categories. First, it distinguishes *particulars* from *universals*; the former being entities such as Microsoft Inc. or the specific contract the company signed with the 15 universities in Ohio in 1999, the latter being entities, such as *company* and *contract*, which have the former as their instances.

Second, it distinguishes, within the realm of particulars, between *continuants* and *occurrents*. Continuants are those entities, such as Microsoft and its current CEO, that endure continuously through a period of time while undergoing changes of various sorts. Occurrents, in contrast, *are* such changes; they are entities which unfold over a certain time span through their successive temporal parts or phases – thus they are the entities otherwise called 'processes,' 'actions', 'events.' However, not all entities are segmentable in this way because there are beginnings and endings and other boundaries in the realm of occurrents, entities which are instantaneous: the latter are analogous to the edges and surfaces of objects in the realm of continuants. Just as such spatial boundary-entities can exist only as the boundaries of three-dimensional spatially extended objects, so temporal boundary-entities can exist only as the boundaries of temporally extended processes. Typically, the beginning and ending of an occurrent, as well as everything

that takes place between these two points, are parts of the occurrent itself. The beginning and ceasing to exist of a continuant, in contrast, are not parts of the continuant itself, but rather parts of that occurrent which is its life or history.

Third, there is the distinction between *dependent* and *independent* entities, where each dependent entity is defined as being such that it cannot exist without some independent entity which plays the role of bearer. A contract, for example, cannot exist without contracting organisations or persons.

Fourth, there is the distinction between *fiat* and *bona fide* entities, which is based on the opposition between bona fide (or physical) and fiat boundaries, the latter being exemplified especially by boundaries – such as the boundary of Utah, or of the 20th century – introduced via human demarcation (Smith & Varzi, 1997).

BFO also distinguishes three major families of relations between the entities just sketched: (1) $\langle p, p \rangle$ -relations, obtaining between particular and particular (for example: Steve Ballmer being the CEO of Microsoft); (2) $\langle p, u \rangle$ -relations, obtaining between particular and universal (for example: Steve Ballmer's being an instance of the universal *person*); and (3) $\langle u, u \rangle$ -relations, obtaining between universal and universal (for example: *software company* being a subkind of *company*) (Smith, Ceusters, Klagges et al., 2005). The importance of this distinction is exemplified by the fact that relationships such as parthood have distinct properties at the particular and at the universal levels and that ignoring these distinctions has led to a number of erroneous representations of relations crucially important in the domain of medical care (Donnelly, Bittner, & Rosse, 2006).

3.2 Granular Partition Theory

The second element of the proposal is Granular Partition Theory, a highly general framework for understanding the ways in which, when cataloguing, classifying, mapping or inventorising a certain portion of reality (POR), human beings and other cognitive agents divide up or partition this reality at one or more levels of granularity (Bittner & Smith, 2003). The resultant partitions are composed of *partition units* (analogous to the cells in a grid) and the theory provides a formal account of the different ways in which such units can correspond, or fail to correspond, to the entities in reality towards which they are directed. The theory takes account for example of the degree to which a partition represents the mereological structure of the domain onto which it is projected, and also of the degree of completeness with which a partition represents this domain.

Drawing on this framework, we have proposed a calculus for use in quality assurance of complex representations created for clinical or research purposes in the context of both ontology evolution (Ceusters & Smith, 2006a) and ontology mapping (Ceusters, 2006). The calculus is based on a distinction between three levels (Smith, Kusnierczyk, Schober, & Ceusters, 2006):

- 1) the level of reality (for example on the side of specific enterprises);
- 2) the cognitive representations of this reality (for example as embodied in observations and interpretations on the part of business analysts);
- 3) the publicly accessible concretisations of these representations in artefacts of various sorts, of which ontologies and corporate memories are specific examples.

Representations are composed in modular fashion of sub-representations built out of *representational units* that are assumed to correspond to PORs, whereby:

- 1) each such unit is assumed to be veridical, i.e. to conform to some relevant POR on the basis of our best current understanding (which may, of course, be based on errors);
- 2) several units may correspond to the same POR by presenting different though still veridical views or perspectives of this reality, for instance one and the same event may be described both as an event of buying and as an event of selling;
- 3) what is to be represented by the units in a representation depends on the purposes which the representation is designed to serve.

In ontologies and terminologies the representational units are terms from some natural or formal language and are assumed to refer to universals or defined classes; in corporate memories the representational units refer to enterprise-specific entities either at the level of universals or at the level of instances.

3.3 The Referent Tracking Paradigm

Referent tracking, as the third and final step, is a new paradigm for representing instances in reality that has been introduced to support the entry and retrieval of data in the Electronic Health Record (EHR) (Ceusters & Smith, 2006b). Its purpose is to avoid the ambiguity that arises when statements in an EHR refer to disorders, lesions and other entities on the side of the patient exclusively by means of generic terms from a terminology or ontology. In (Parsons & Wand, 2000) it is argued that problems in schema integration, schema evolution, and interoperability are precisely the consequence of ambiguities of this sort, which are deeply rooted in the erroneous assumption adhered to in database design circles according to which entities can be referred to only as instances of pre-specified classes. The authors term this *the assumption of inherent classification* and make the case that this assumption violates philosophical and cognitive guidelines on classification.

Referent tracking avoids such ambiguities by introducing unique identifiers, called IUIs or Instance Unique Identifiers, for each numerically distinct entity that exists in reality and that is referred to in statements in a record. The referent tracking paradigm thereby expands the entities uniquely identified for EHR purposes far beyond the current range, which is restricted to entities such as patients, care providers, buildings, machines and so forth, to include fractures, polyps, seizures and a vast variety of other clinically salient real-world instances. In the context of corporate memories, analogously, IUIs would be assigned not merely to the various organizations and persons relevant to the enterprise (companies, employees, customers, and so forth) but also to contracts, applicable laws, meetings, sales transactions, accidents in manufacturing facilities, deliveries, and so forth. For many of the latter entities unique identifiers will exist already in the various information systems of a large corporation. Our proposal is that these identifiers should be assembled in a single corporate memory store, where they will constitute an evolving dynamic map of the corporation and of all events and processes with which the corporation is involved.

At least the following requirements have to be addressed if the paradigm of referent tracking is to be given concrete form in a Referent Tracking System (RTS) destined to serve the needs of an enterprise: (1) a mechanism for generating IUIs that are guaranteed to be unique strings; (2) a procedure for deciding which particulars should receive IUIs; (3) protocols for determining whether or not a particular has already been assigned a IUI (each particular should receive maximally one IUI); (4) rules governing the processing of IUIs in information systems in general, including rules concerning the syntax and semantics of statements containing IUIs; (5) methods for determining the truth values of propositions that are expressed through descriptions in which IUIs are employed; (6) methods for correcting errors in the assignment of IUIs, and for investigating the results of assigning alternative IUIs to problematic cases; and (7) 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 department of a large company. Clearly, however, the referent tracking paradigm will most optimally serve its purpose when it is used in a distributed, collaborative environment such as a large company with several offices dispersed over a wide area. One and the same customer is often served by a variety of departments within a single large company, many of them working in different settings, and each of these settings may use its own separate information system. These systems contain different data, but these data often provide information about the same particulars. Under the current state of affairs, it is 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 in very many cases a trivial matter and this even on a meta-company level. It could for instance give considerably added value to services of the kind delivered by Factiva which uses a four step automated and manual process to ensure that everything is correctly categorized. Customers can receive the data either as an XML feed or a Web service for integration into their corporate intranets, or their CRM or competitive-intelligence systems (Drew, 2006).

3.4 Services of a Referent Tracking System

An RTS should offer at least three services: (1) to generate unique identifiers to be used as IUIs, (2) to keep track of the IUIs generated, and (3) to provide access to the IUIs stored.

As to the first, several schemes for generating strings that are guaranteed to be unique are already in use, the most prominent of which is Microsoft's Globally Unique Identifier paradigm (GUID), which implements UUIDs (Universally Unique IDs) as defined by the Open Software Foundation in the specification of its Distributed Computing Environment (Williams & Kindel, 1994). The advantage of the GUID paradigm is that unique identifiers can be generated easily on any machine with a network card without the need to resort to a central authority to guarantee uniqueness.

UUIDs have recently been standardized through ISO/IEC 9834-8:2004, which specifies format and generation rules that enable users to produce every 100 nanoseconds 128-bit identifiers which are either guaranteed to be, or have a high probability of being, globally unique (International-Standards-Organisation, 2004). The standard also specifies the procedures for the operation of a Web-based Registration Authority for UUIDs.

If RTS services would be offered by a player external to a specific organisation; it might be beneficial that this player not only registers IUIs, but also certifies the uniqueness of the strings to be used within a given IUI-repository and guarantees that the assignments claimed to have been made by given authors were indeed made by those authors. This can be compared to the services offered by trusted third parties in private key management for asymmetrical encryption purposes (Bellare & Rogaway, 1996).

The second service involves what we shall refer to as the *IUI-repository*, whose purpose is to keep track of the identifiers assigned to already existing entities, or reserved for entities that are expected to come into existence in the future. It will do this in such a way that (i) each IUI represents exactly one particular, and (ii) no particular is referred to by more than one IUI. These two requirements are not easy to fulfil, since both depend on the ability and willingness of users to provide accurate information. This, however, introduces no problems different in principle from those already faced by the users of existing systems when called upon to provide information of a non-trivial and occasionally sensitive sort.

The third service, here called the referent-tracking database (RTDB), should provide access to the information entered into a given corporate memory about the particulars referred to in the IUI-repository. Where the IUI repository is an inventory of concrete entities that have been acknowledged to exist, and, consequently, of what IDs to use if one wants to refer to them, the RTDB is an inventory of descriptions concerning the features and interrelations of these entities and the ways in which they change in the course of time. The RTDB, too, does not need to be set up as a single central database but can rely on any paradigm for distributed storage.

The primary role of the RTDB is to keep track of the features of given particulars and of their relationships to other particulars as they change through time, and of the assertions that have been made about such particulars, including those assertions which have been shown to be false (stored, for example, for the purposes of providing an audit trail). It has an important role also in helping users to determine whether particulars they encounter for the first time have been registered already in the IUI-repository, or whether a new IUI must be created for use in new descriptions. To be sure, this places some additional burden on the person who has to enter the information; but, given that cases such as this are likely to be of high salience, the time perceived as being lost at this stage will likely be recovered when searching for information thereafter.

4 Applying BFO and Referent Tracking to corporate memories and enterprise ontologies

For the remainder of this paper, we will give a few examples of how the theories and paradigms described above can be used to detect and solve a number of problems and inconsistencies that we (and others) encountered in studying the literature on enterprise engineering and corporate memories.

Quite common is the inclusion of representational units in an ontology that do not have a counterpart in reality. This happens at the level both of relationships and of the entities which serve as their relata. Consider the difference between the "Sale" and "Have-Capability" relationships as defined in the

Enterprise Ontology (Uschold et al., 1998). A ‘Sale’ is (acceptably) defined as ‘*a relationship constituting an agreement between two Legal Entities to exchange a Product for a Sale Price*’, in keeping with the Enterprise Ontology’s treatment of relationships as entities in their own right that can thus be instantiated. Two instances of Legal Entity thus enter into a single instance of the Sale relationship. The “Have-Capability” relationship, on the other hand, is defined as ‘*a relationship between a Person and an Activity denoting that the Person is able to perform the Activity*’. The first problem here is the confusion of use and mention: relationships themselves do not ‘denote’; this is the task of the corresponding denoting expressions. But more importantly: being able to engage in an activity does not require that any instance of such an activity exists. Under BFO, properties of this sort would be represented correspondingly as falling within the realm of abilities (powers, functions, dispositions), in such a way that the existence of a capability does not imply the existence of any realization of this capability.

The use-mention confusion (which is common not only among enterprise ontology developers) confuses the level of reality with the level of our representations thereof. It is found, for example, in (Goossenaerts & Pelletier, 2003): the latter correctly argues that the Enterprise and TOVE ontologies do not emphasize the distinction between things and their changes on the one hand and conceptual entities on the other hand, drawing their analysis from the work of Bunge (Bunge, 1977) and specifically from its application in the Bunge-Wand-Weber model in the domain of information systems (Wand, Storey, & Weber, 1999). This analysis led them to develop the PSIM Ontology (*Participative Simulation environment for Integral Manufacturing renewal*), which was inspired also by earlier work conducted in the European Research Project CIMOSA (AMICE-Consortium, 1989) and from Peircean Semiotics (Hoopes, 1991). The result, however, is not without its own dramatic misinterpretations. Thus we read that the PSIM ontology distinguishes three main categories: Activity, Object and Information (element), an Information element being defined as: ‘*a characteristic of either an object or activity or information, which is used to constrain directly or indirectly the involvement of an object in an activity*’ (Goossenaerts & Pelletier, 2003, pp 45). Whereas we can to some extent sympathise with their proposal to classify ‘*the time needed to perform an activity*’ or ‘*how an activity has to be performed*’ as information elements, it is surely a violation of sound ontological principles to consider ‘*how the enterprise is organised*’, ‘*the way the responsibilities are distributed among the enterprise*’, and ‘*the weight of a piece of material*’ to be information. Weight, for BFO, is a dependent continuant that depends on the material object of which it is the weight, and this independent of whether or not a cognitive being has any sort of information about the matter.

Confusions of this sort, as we argued elsewhere (Smith, Ceusters, & Temmerman, 2005), are a direct result of the concept orientation in ontology. The latter leads quite often also to a blurring of the distinction between instantiation and subtyping. Whereas in BFO instantiation is a relationship between a particular and a universal, subtyping is a relationship between universals. Nothing which is an instance can itself have instances, while something that is a subtype, can itself have other subtypes. As is correctly recognised in (Uschold et al., 1998), the distinction between a type of entity, and a particular entity of a certain type, i.e. an instance, is not consistently made when using natural language. This does not, however, mean that it is acceptable when the authors of the Enterprise Ontology ‘*intentionally blurred this distinction*’ in the informal description of their ontology (Uschold et al., 1998, pp 35). And when the methodological work underlying the Core Enterprise Ontology allows John Doe to be an instance of “consumer”, and “consumer” to be an instance of “entity” (Bertolazzi et al., 2001), the result is a mistake that is impermissible in any serious ontology work.

Note that it is not just natural language that blurs the mentioned distinction: traditional database design paradigms exhibit the same type of confusion, as well as some ontology authoring environments such as Cyc (Foxvog, 2005) : a table about cars may contain ‘instances’ such as ‘Volkswagen’ or ‘Audi’. But again, under a realist paradigm, such a representation can only be the result of a sloppy analysis in which a car brand is mistaken for a car.

The rigorous identification schemes proposed by the Referent Tracking paradigm are an important first step in doing away with such confusions, and schemes of the given sort are indeed beginning to acquire acceptance in the enterprise engineering environment, as witnessed by a recent case study exploring the possible complementarity of the Demo Engineering Methodology for Organizations (DEMO) and the

Object Role Modelling (ORM) paradigm (Dietz & Halpin, 2004). DEMO enables the business processes of organizations to be modelled independently of how these processes are implemented, thereby focusing on the communication acts that take place between human actors in the organization. ORM enables business information to be modelled in terms of *fact types* as well as the business rules that constrain how the fact types may be populated for any given state of the information system and how derived facts may be inferred from other facts. One important feature of ORM is its requirement for the inclusion of at least one identification scheme for each entity type, which functions as an identity criterion for instances of that type. Because of this requirement, data use cases, i.e. samples of information, can be used to seed an initial model. However, if ORM is to be used for the purposes of building an ontology rather than a database schema, then developers should pay attention to the fact that several records in a database may refer to the same entity in reality. This is certainly the case for example when databases maintained in originally distinct organisations are merged because of a company take-over.

5 Conclusion

For a company to anticipate and manage change for the future, to design appropriate strategies that will create business value for customers, and to improve profitability in current and new markets, it must have insight into its present business environment as a complex dynamic whole comprehending the activities, resources, markets, customers, products, services, regulations and costs associated with the enterprise. This insight, the key to strategic intelligence, comes from attempts to improve the capabilities of the company's managers and workers to learn about changes in the business or industry environment, for example as summarized in (Marchand, Davenport, & Dixon, 2000). Corporate memories are crucial to building and sustaining such strategic intelligence, and ontologies *combined with referent tracking* are key to building corporate memories. Although ontology is in essence a philosophical discipline that follows the principles of high quality scientific inquiry in order to discover what holds in reality, the development of ontologies as artifacts for use in computer systems is too often conducted in a way that ignores reality. We have presented a suite of theories and paradigms that are intended to bring reality back into business, and to save business from the conceptual models of their IT personnel.

6 References

- AMICE-Consortium. (1989). *Open System Architecture for CIM, Research Reports of ESPRIT Project 688* (Vol. 1). Berlin Springer Verlag.
- Baumann, S., Malburg, M., Hofe, H. M. a. m., & Wenzel, C. (1997). *From Paper to a Corporate Memory - A First Step*. Paper presented at the KI-97 Workshop on Knowledge-Based Systems for Knowledge Management in Enterprises.
- Beckett, R. C. (2000). A characterisation of corporate memory as a knowledge system. *Journal of Knowledge Management*, 4(4), 311-319.
- Bellare, M., & Rogaway, P. (1996). The exact security of digital signatures – How to sign with RSA and Rabin. In *Lecture Notes in Computer Science* (Vol. 1070, pp. 399-416): Springer.
- Benkard, C. L. (2000). Learning and Forgetting: The Dynamics of Aircraft Production. *The American Economic Review*, 90(4), 1034-1054.
- Berthon, P., Pitt, L. F., & Ewing, M. T. (2001). Corollaries of the Collective: The Influence of Organizational Culture and Memory Development on Perceived Decision-Making Context. *Journal of the Academy of Marketing Science*, 29(2), 135-150.
- Bertolazzi, P., Krusich, C., & Missikoff, M. (2001). *An Approach to the Definition of a Core Enterprise Ontology: CEO*. Paper presented at the OES-SEO 2001, International Workshop on Open Enterprise Solutions: Systems, Experiences, and Organizations.
- Bittner, T., & Smith, B. (2003). A Theory of Granular Partitions. In M. Duckham, M. F. Goodchild & M. F. Worboy (Eds.), *Foundations of Geographic Information Science* (pp. 117-151). London: Taylor & Francis Books.

- Bunge, M. (1977). *Treatise on Basic Philosophy, Ontology I: The Furniture of the World* (Vol. 3). Boston: Reidel.
- Ceusters, W. (2006). Towards A Realism-Based Metric for Quality Assurance in Ontology Matching. In B. Bennett & C. Fellbaum (Eds.), *Formal Ontology in Information Systems* (pp. 321-332). Amsterdam: IOS Press.
- Ceusters, W., & Smith, B. (2006a). A Realism-Based Approach to the Evolution of Biomedical Ontologies. In *Proceedings of AMIA 2006* (pp. 121-125).
- Ceusters, W., & Smith, B. (2006b). Strategies for Referent Tracking in Electronic Health Records. *Journal of Biomedical Informatics*, 39(3), 362-378.
- Ceusters, W., & Smith, B. (2007). Referent Tracking for Digital Rights Management. *Forthcoming in International Journal of Metadata, Semantics and Ontologies*.
- Ceusters, W., Smith, B., & Fielding, J. M. (2004). LinkSuite™: Formally Robust Ontology-Based Data and Information Integration. In E. Rahm (Ed.), *Lecture Notes in Computer Science* (Vol. 2994, pp. 124-139): Springer.
- Dietz, J. L. G., & Halpin, T. A. (2004). Using DEMO and ORM in Concert: A Case Study. *Advanced Topics in Database Research*, 3, 218-236.
- Donnelly, M., Bittner, T., & Rosse, C. (2006). A formal theory for spatial representation and reasoning in biomedical ontologies. *Artificial Intelligence in Medicine*, 36(1), 1-27.
- Drew, R. (2006, March 20, 2006). In Google's Shadow. *Computerworld*.
- Fadel, F. G., Fox, M. S., & Gruninger, M. (1994). A Resource Ontology for Enterprise Modelling. In I. o. I. Engineers (Ed.), *Proceedings of the Third Industrial Engineering Research Conference* (pp. 455-460).
- Fox, M. S. (1992). *The TOVE Project: Towards A Common-sense Model of the Enterprise* (Technical Report): Enterprise Integration Laboratory.
- Foxvog, D. (2005). Instances of Instances Modeled via Higher-Order Classes. In *Proceedings of the Workshop on Foundational Aspects of Ontologies (FOnt 2005)* (pp. 46-54).
- Gandon, F., Poggi, A., Rimassa, G., & Turci, P. (2002). Multi-Agent Corporate Memory Management System. *Journal of Applied Artificial Intelligence*, 16(9-10), 699-720.
- Geerts, G., & McCarthy, W. E. (2002). An Ontological Analysis of the Primitives of the Extended-REA Enterprise Information Architecture *The International Journal of Accounting Information Systems*, 3, 1-16.
- Gershefski, G. W. (1970). Corporate Models--The State of the Art. *Management Science*, 16(6), 303-312.
- Goossenaerts, J., & Pelletier, C. (2003). Ontology and Enterprise Modeling [Electronic Version]. Retrieved November 16, 2006 from <http://is.tm.tue.nl/staff/jgoossenaerts/4PublicPdf/PSIM%20book%20ch%205%20Ontol&EM.pdf>.
- Grenon, P., Smith, B., & Goldberg, L. (2004). Biodynamic Ontology: Applying BFO in the Biomedical Domain. In D. M. Pisanelli (Ed.), *Ontologies in Medicine* (pp. 20-38). Amsterdam: IOS Press.
- Gruninger, M., & Fox, M. S. (1994). An Activity Ontology for Enterprise Modelling [Electronic Version]. Retrieved November 22, 2006 from <http://www.eil.utoronto.ca/enterprise-modelling/papers/gruninger-wetice94-act.pdf>.
- Hill, M. (2006). Service Taxonomy and Service Ontologies Deliver Success to Enterprise SOA [Electronic Version]. *SOA Webservices journal*, 6. Retrieved November 23, 2006 from <http://webservices.sys-con.com/read/175385.htm>.
- Hoopes, J. (1991). *Peirce ON SIGNS. Writings on Semiotic by Charles Sanders Peirce*. Chapel Hill and London: The University of North Carolina Press
- Huhns, M. N., & Stephens, L. M. (2002). Semantic Bridging of Independent Enterprise Ontologies. In K. Kosanke (Ed.), *Enterprise Inter- and Intra-Organizational Integration: Building International Consensus* (pp. 83 - 90). Boston, MA: Kluwer Academic Publishers.
- International-Standards-Organisation. (2004). ISO/IEC FDIS 9834-8:2004. Information technology – Open Systems Interconnection – Procedures for the operation of OSI Registration Authorities:

- Generation and registration of Universally Unique Identifiers (UUIDs) and their use as ASN.1 Object Identifier components.
- Kalfoglou, Y., Alani, H., O'Hara, K., & Shadbolt, N. (2002). Initiating Organizational Memories Using Ontology Network Analysis. In *Knowledge Management and Organizational Memories Workshop, 15th European Conf. Artificial Intelligence (ECAI)* (pp. 79-89): IOS Press, Amsterdam.
- Kim, H. M., Fox, M. S., & Gruninger, M. (1995). An Ontology of Quality for Enterprise Modelling. In *Proceedings of the Fourth Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises* (pp. 105-116): IEEE Computer Society Press.
- Kim, H. M., Fox, M. S., & Gruninger, M. (1999). An ontology for quality management - enabling quality problem identification and tracing. *BT Technology Journal*, 17(4), 131-140.
- Kim, H. M., Sengupta, A., Fox, M. S., & Dalkilic, M. (2006). A Measurement Ontology Generalizable for Emerging Domain Applications on the Semantic Web. *Journal of Database Management*, 18(1), 20-42.
- Kühn, O., & Abecker, A. (1997). Corporate memories for Knowledge Management in Industrial Practice: Prospects and Challenges. *Journal of Universal Computer Science*, 3(8), 929-954.
- Kung-Juan, Y., & Yuh-Min, C. (2006). Ontology-Based Knowledge Retrieval in Organizational Memory. In *First International Conference on Innovative Computing, Information and Control (ICICIC'06)* (pp. 566-569).
- Liles, D. H., & Presley, A. R. (1996). *Enterprise modeling within an enterprise engineering framework*. Paper presented at the 96 Winter Simulation Conference.
- Marchand, D. A., Davenport, T. H., & Dixon, T. (2000). *Financial Times—Mastering Information Management, Complete MBA Companion in Information Management*. London: FT Prentice Hall.
- McCarthy, W. E. (1982). The REA Accounting Model: A Generalized Framework for Accounting Systems in a Shared Data Environment. *The Accounting Review*, LVII(3), 554-578.
- Osterwalder, A., Lagha, S. B., & Pigneur, Y. (2002). *An ontology for developing e-business models*. Paper presented at the IFIP DSIAge'2002.
- Park, J. E., & Bunn, M. D. (2003). Organizational memory: A new perspective on the organizational buying process. *The Journal of Business & Industrial Marketing*, 18(2/3), 21.
- Parsons, J., & Wand, Y. (2000). Emancipating Instances from the Tyranny of Classes in Information Modeling. *ACM Transactions on Database Systems*, 25(2), 228-268.
- Partridge, C., & Stefanova, M. (2001). A Synthesis of State of the Art Enterprise Ontologies - Work in Progress [Electronic Version]. Retrieved November 22, 2006 from <http://citeseer.ist.psu.edu/632089.html>.
- Prasad, M. V. N., & Plaza, E. (1996). Corporate Memories as Distributed Case Libraries. In *Tenth Knowledge Acquisition for Knowledge-Based Systems Workshop* (pp. 40-41 40-19). Banff, Canada.
- Rosenthal, A., Manola, F., & Seligman, L. (2001). Getting Data to Applications—Why We Fail—Part 1: Common Fallacies. *The Mitre Information Technology Advisor*, 1(10), 1-2.
- Schrieber, A. N. (1970). Corporate Simulation Models. *Management Science*, 17(1), 120-121.
- Smith, B. (2004). Beyond concepts: ontology as reality representation. In *Proceedings of the third international conference on formal ontology in information systems (FOIS 2004)* (pp. 73-84). Amsterdam: IOS Press.
- Smith, B., & Ceusters, W. (2003). Towards Industrial-Strength Philosophy; How Analytical Ontology Can Help Medical Informatics. *Interdisciplinary Science Reviews*, 28(2), 106-111.
- Smith, B., & Ceusters, W. (2006). HL7 RIM: An Incoherent Standard. In A. Hasman, R. Haux, J. v. d. Lei, E. D. Clercq & F. Roger-France (Eds.), *Studies in Health Technology and Informatics. Ubiquity: Technologies for Better Health in Aging Societies - Proceedings of MIE2006* (Vol. 124, pp. 133-138). Amsterdam: IOS Press.
- Smith, B., Ceusters, W., Klagges, B., Köhler, J., Kumar, A., Lomax, J., et al. (2005). Relations in biomedical ontologies. *Genome Biology*, 6(5), R46.

- Smith, B., Ceusters, W., & Temmerman, R. (2005). Wüsteria. In R. Engelbrecht, A. Geissbuhler, C. Lovis & G. Mihalas (Eds.), *Studies in Health Technology and Informatics* (Vol. 116, pp. 647-652). Amsterdam: IOS Press.
- Smith, B., Kusnierczyk, W., Schober, D., & Ceusters, W. (2006). *Towards a Reference Terminology for Ontology Research and Development in the Biomedical Domain*. Paper presented at the KR-MED 2006, Biomedical Ontology in Action.
- Smith, B., & Varzi, A. C. (1997). Fiat and Bona Fide Boundaries: Towards an Ontology of Spatially Extended Objects In *Lecture Notes In Computer Science* (Vol. 1329, pp. 103 - 119). London, UK: Springer Verlag.
- Stader, J. (1996). *Results of the Enterprise Project*. Paper presented at the 16th Annual Conference of the British Computer Society Specialist Group on Expert Systems
- Stein, E. W. (1995). Organizational Memory: Review of Concepts and Recommendations for Management. *International Journal of Information Management*, 15(2), 17-32.
- Tham, K. D. (1999). *Representation and Reasoning About Costs Using Enterprise Models and ABC*. Unpublished Dissertation, University of Toronto, Toronto, Canada.
- Tham, K. D., Fox, M. S., & Gruninger, M. (1994). A Cost Ontology for Enterprise Modelling. In *Proceedings of the Third Workshop on Enabling Technologies - Infrastructures for Collaborative Enterprises*. West Virginia University.
- Uschold, M., King, M., Moralee, S., & Zorgios, Y. (1998). The Enterprise Ontology. *The Knowledge Engineering Review*, 13(1), 31-89.
- Van Heijst, G., Van der Spek, R., & Kruizinga, E. (1996). *Organizing Corporate Memories*. Paper presented at the Tenth Knowledge Acquisition for Knowledge-Based Systems Workshop.
- Vasconcelos, J., Kimble, C., & Gouveia, F. (2000). A Design for a Group Memory System using Ontologies. In *Proceedings of the 5th UKAIS (United Kingdom Academy for Information Systems) Conference* (pp. 246-255). Cardiff: McGraw-Hill.
- Vasconcelos, J., Kimble, C., Gouveia, F. R., & Kudenko, D. (2001). *Reasoning in Corporate Memory Systems: A Case Study of Group Competencies* Paper presented at the 8th International Symposium on the Management of Industrial and Corporate Knowledge.
- Walsh, J. P., & Ungson, G. R. (1991). Organizational Memory. *The Academy of Management Review*, 16(1), 57-91.
- Wand, Y., Storey, V., & Weber, R. (1999). An Ontological Analysis of the relationship Construct in Conceptual Modeling. *ACM Transactions on Database Systems*, 24(4), 494-528.
- Wexler, M. N. (2002). Organizational Memory and Intellectual Capital. *Journal of Intellectual Capital*, 3(4), 393-415.
- Williams, S., & Kindel, C. (1994). The component object model: A technical overview. [Electronic Version]. Retrieved November 14, 2006 from http://msdn.microsoft.com/library/default.asp?url=/library/en-us/dncomg/html/msdn_comppr.asp.