An Ontology for Virtual Organization Breeding Environments

Joël Plisson, Peter Ljubič, Igor Mozetič, and Nada Lavrač

Abstract—Companies and individuals connect into networks to share their resources with the purpose of achieving a common goal. The field of collaborative network organizations covers various types of organizational structures. Knowledge, which is stored in such networks, can be separated into two different levels. First, there is a common knowledge about the organizational structure itself that can be used and reused in any of such networks. The second level represents the domain-specific knowledge, which such networks cover and use to function. In this paper, we address both levels, first, by proposing an ontology representing the common vocabulary and identifying the actors and relationships in a specific type of network, namely virtual organization breeding environment (VBE), and second, by proposing a methodology for extracting network-specific knowledge related to competencies. The instantiation of the proposed VBE ontology and the developed approach to semiautomated construction of competencies have been applied to real problem scenarios of Virtuelle Fabrik, a Swiss–German cluster of companies in mechanical engineering.

Index Terms—Collaborative network organization, ontology, ontology construction, virtual organization.

I. INTRODUCTION

THIS PAPER focuses on collaborative networked organizations (CNOs), which are constituted by a variety of entities (e.g., organizations and people) that are largely autonomous, geographically distributed, and heterogeneous in terms of their operating environment, culture, social capital, and goals [1]. Here, the term network refers to a type of organizational structure that takes its coordination beyond the company boundaries [2]. The heterogeneity of networks appears within several dimensions. First, in terms of time scale one can distinguish between long-term associations, like strategic alliances of companies, and short-term associations that are simply trying to complete a certain task. Furthermore, these associations can be profit or nonprofit oriented. Heterogeneity can also emerge from the type of participants involved in the collaboration: they can be individuals or organizations. This diversity has led to the identification of specific types of collaborative networks [1].

Several types of collaborative networks exist, namely virtual enterprises, professional virtual communities, collaborative virtual laboratories, etc. In this paper, we focus on recently emerged virtual organization breeding environments (VBEs) [3]. A VBE is an association (also known as a cluster) of organizations and their related supporting institutions, which have both the potential and the will to cooperate with each other through the establishment of a base long-term cooperation agreement and interoperable infrastructure [1]. VBEs respond to business opportunities by forming virtual organizations (VOs). While a VBE represents a long-term association, a VO is a short-term association with a specific goal of acquiring and fulfilling a business opportunity.

It is agreed that networking, as a new way of collaboration, brings benefits to its members, but collaboration implies sharing of knowledge and communication between participants. As the participants might come from different fields or follow a different philosophy, it is necessary to introduce a mechanism to share common understanding of the information, and to agree on a controlled vocabulary used to communicate. An ontology provides a representation of knowledge, which can be used and reused in order to facilitate the comprehension of concepts and relationships in a given domain, and the communication between different domain actors by making the domain assumptions explicit. These actors can be software agents or people that need to access or share a piece of information [4].

Ontologies have proven to be a unambiguous and compact way of knowledge representation for mutual understanding as they provide a basis for sharing information. In order to share the same terminology, the participants of the VBE need to agree on the terms that they intend to use for collaboration. If several organizations joining a VBE share the same underlying concepts (or terminology) (for example on the Web or on their intranet), then software agents are able to extract and aggregate information and use it to gather the data and to answer some queries. Such agents can also support a process of VO creation by proposing more or less optimal VOs based on their competencies.

The goal of this paper is to outline existing techniques and to propose new methods and techniques appropriate for modeling VBEs. The main contributions of the paper are the following. First, the paper proposes an ontology representing the common vocabulary and identifying the actors and relationships in a specific type of network, namely VBE. In this way, the paper contributes to the formalization of the informal notions of VBE and VO in a formal ontology language. This formalization is not merely a codification in an ontology language because it requires a detailed elaboration of all the concepts, consideration

of mutual dependencies between concepts, and the overall consistency of the developed ontology. Additionally, we have made the CNO ontology, built in Protégé, available at a public web site,[1] so that it can be redistributed and/or modified.[2] The site also includes an extensive glossary of terms used/introduced in this paper.

Second, the paper proposes a methodology for semiautomated ontology construction for the needs of VBEs, enabling the extraction of network specific knowledge related to competencies. Finally, the instantiation of the proposed VBE ontology and the developed approach to semiautomated structuring of competencies have been applied to and validated on two real-problem scenarios of Virtuelle Fabrik, a Swiss–German cluster of companies in mechanical engineering.

Section II gives an outline and definitions of ontologies. Different types of ontologies are presented in Section II-A. Their encoding and structure is presented in Section II-B. Ontology construction methodologies are presented in Section II-C.

Section III is dedicated to ontologies related to the business domain. It presents their content and how they were implemented, as well as how they are intended to be used.

In Section IV, an ontology for VBEs, developed in Protégé, is presented. Its purpose is to establish the vocabulary used in the domain of VBE and to identify the main actors and their roles in a VBE. The ontology can be used and reused in different VBEs, and refined with the specificities of each network. One of the most important characteristics that changes from one network to another is a set of partners' competencies, representing the ability of the network (or its subset) to participate in particular types of projects and to perform specific tasks. This cannot be generalized to all VBEs, and thus, the information must be gathered specifically for each VBE. While, the construction of the ontology is made manually, it is possible to refine it in a semiautomated fashion, as shown in Section V.

Section V describes the semiautomated refinement of the manually constructed ontology. First, it presents a methodology to extract and (hierarchically) structure company competencies from their text descriptions, which complements the initial ontology (that is universal for all VBEs) with more specific information such as competencies. This section also includes a case study implemented for competency structuring of the Virtuelle Fabrik industrial cluster.

In Section VI, we validate the general ontology on two cases from the Virtuelle Fabrik cluster. From their recent business history, we analyzed the formation of two actual VOs, which were created to carry out two different tasks and formalized them in the proposed VBE ontology.

II. ONTOLOGIES AND THEIR CONSTRUCTION

An ontology provides a representation of knowledge, which can be used and reused, in order to facilitate the comprehension of concepts and relationships in a given domain and the communication between different domain actors. The most basic type

of ontology is a set of terms representing a controlled vocabulary (e.g., a glossary), which are the terms that people agree to use when dealing with a common domain. By providing definitions, an ontology helps people and machines to use the same terms for expressions leading to better mutual understanding. The role of an ontology is not limited to providing a common vocabulary; complex ontologies can also constrain the usage of knowledge by giving axioms or microtheories and define the relations between the different components.

A. Types of Ontologies

The content of an ontology depends both on the amount of information and on the degrees of formality that is used to express it. Generally, we distinguish two main types of ontologies: lightweight and heavyweight [5]. A lightweight ontology is a structured representation of knowledge, which ranges from a simple enumeration of terms to a graph or taxonomy, where the concepts are arranged in a hierarchy with a simple (specialization, is–a) relationship between them. A heavyweight ontology adds more meaning to this structure by providing axioms and broader descriptions of knowledge. As a word can have several senses, knowledge can also be interpreted in different ways, which creates ambiguity in the knowledge base. Axioms and constraints tend to reduce the ambiguity by restricting and constraining the usage of information, for instance by specifying what is possible to do with it and what is not.

The encoding of an ontology varies from informal to a highly formal representation. A lightweight ontology is usually informal and sufficient to define concepts and basic relationships between them. A formal ontology contains axioms and definitions usually stated in logic. It is also called heavyweight because it can support more complex queries and deliver comprehensive answers.

The degrees of complexity of knowledge expressed in an ontology can vary from one ontology to another. This is also true for the spectrum of knowledge. An ontology may cover one or several domains or even focus on a specific aspect. In all cases, the construction of an ontology involves the choice of appropriate concepts that will best describe the knowledge represented in the ontology. These choices are called "ontological commitments" and are described by the ontology. "We say that an agent commits to an ontology if its observable actions are consistent with the definitions in the ontology" [4].

The number of concepts chosen and their specificity make the ontology fall into one of the following five categories [5]–[7], ordered by degrees of specialization.

1) Upper (top level, generic, foundational, etc.) ontology: A foundational ontology contains very general concepts that can define the most abstract entities (object, event, physical, abstract, etc). All other concepts specialize this top level representation.
2) Core ontology: A core ontology comprises knowledge about a field or area of expertise that may include several different disciplines such as law, computer science, etc. Only the most representative (core) concepts and

relations of each discipline are kept in order to represent their union. These core concepts usually contain the root of each domain-specific ontology that represents a discipline.

3) Domain ontology: Knowledge encoded in a domain ontology is more specific. It presents a more specialized view of the concepts defined in a core ontology, where only the most important ones are kept. A domain ontology tries to cover all the aspects of one domain (e.g., medicine, law, etc.) and the interactions between them.

4) Task (application) ontology: Knowledge can even be limited to the minimum required to fulfill the needs for one task. An application ontology can be reduced to a part of one domain ontology or even a mix of two or more different domain ontologies. Only the concepts that are suitable to the task and to the comprehension of the process are kept.

5) General (common) ontology: In a general ontology, the knowledge is reusable in different domains. The knowledge can represent different kinds of concepts, such as the units of measure (time, space, etc.) or even general relations and axioms that are applicable in most of the domains, such as the is–a (hierarchical) relation or the part-of (meronymical) relations.

B. Ontology Representation and Encoding

From the representation point of view, an ontology can be presented in several ways. Generally, concepts in an ontology are first grouped into several taxonomies with the is–a and subclass relations. Then, these taxonomies are linked together with other relations such as meronymy (part-of) or any other predefined relation. In order to be shared across the web, an ontology is encoded in a format that facilitates the interchange. The most basic language to use is XML [8], but it has many disadvantages, such as the lack of description power and the lack of commitment concerning the modeling primitives. Namely, the concepts are nested without telling what kind of relations are binding them together. These deficiencies require extensions of XML.

RDF [8] is the first layer on top of XML, which adds semantic information to the data. RDF allows for representation of ground binary relations in the form of triples <*subject, predicate, object*>. The encoding in RDF is usually guided by an RDF schema (RDF(S) [9]) which specifies the classes and properties that are intended to be used during the encoding process. RDF(S) allows to add more meaning within the definition of classes, properties, and other resources. For instance, it allows to give the range and domain of the properties that are defined. RDF(S) solves some semantic problems and is appropriate for encoding lightweight ontologies. But for heavyweight ontologies it still lacks expressive power.

A step toward heavyweight ontologies encoding was made with web ontology language (OWL). OWL [10] is based on the description logics formalism and is divided into three sublanguages: OWL Lite, OWL DL, and OWL Full.[3] The main

advantage of OWL is that description logics has been an established research field for many years, and thus, it benefits from all the reasoning algorithms already developed and optimized. The knowledge expressed in such a language is formally defined and contains axioms that restrain its usage to a certain context, thus removing ambiguities during the reasoning process.

The choice of one of these languages has to be made according to the requirements assessed when building an ontology.

C. Methods for Ontology Construction

Most existing ontologies were generated manually. Such a process is time intensive, error prone, and exhibits problems in maintaining and updating ontologies. For this reason, researchers are looking for alternatives to enable generating ontologies in a more efficient and effective way. Ontology learning has emerged as a subarea of ontology engineering, due to the rapid increase of web documents and advanced techniques shared by the information retrieval, machine learning, natural language processing, and artificial intelligence communities. This section provides an overview of the field of manual and semiautomated ontology constructions, and highlights some contributions to ontology generation and evolution.

1) Manual Construction of Ontologies: It is important to emphasize some fundamental rules in ontology design, which help making decisions during an ontology construction. First, there is no single correct way to model a domain as there are always alternative views. The best solution almost always depends on the application and the extensions to be made. Second, the construction of an ontology is an iterative process. And finally, concepts in the ontology should reflect the objects and relationships in the domain of interest. The following methodology and recommendations taken from [11] rely on developing an ontology using Protégé [12], which we used for the construction of the ontology, described in Section IV.

The development of an ontology starts with the definition of its domain and scope. Important aspects are also users, who and in what way will use the ontology, and which kind of questions the ontology should provide answers to. One of the ways to determine the scope of the ontology is to sketch a list of questions that the knowledge based on the ontology should be able to answer, also referred to as competency questions.

The next step is to consider reusing existing ontologies. Ontologies are coded in a strict manner, which resembles object-oriented programming. Therefore, it is worth considering whether somebody else already covered the domain, at least partially. It is worth reusing and refining such ontologies, in order to reduce construction time and not to "reinvent the wheel." There are many libraries of reusable ontologies in the literature and on the web (e.g., www.daml.org, http://protege.stanford.edu/plugins/owl/owl-library/).

After reusing existing ontologies is considered, enumeration of all the terms we would like to explain to the user takes place. The list of important terms and their properties will also help us define the scope of the ontology.

Based on the list of terms, the definition of classes and the class hierarchy is done. There are several possible approaches

---

[3]See http://www.w3.org/TR/owl-features/ for details.

to developing a class hierarchy. A top-down development starts with the definition of the most general concepts in the domain and subsequent specialization of the concepts. On the other hand, a bottom-up development begins with the definition of the most specific classes, which are the leaves of the hierarchy, following the grouping of these classes into more general concepts. Usually, a combination of both approaches takes place. One might start with a few top-level concepts and a few specific concepts, which can then relate to the middle-level concepts. Another alternative is to start at the middle-level concepts and then specialize and generalize them to obtain a full hierarchy.

Defined classes need the properties to provide enough information to answer the competency questions from the first step. This can also be seen as internal structure of the concepts. Usually, the properties relate to those terms from the list of all terms that were not defined as classes. For each property in the list, it is necessary to determine which class it describes. These properties become slots attached to classes. In general, there are several types of object properties that can become slots in an ontology. These can be intrinsic properties that belong to a thing by its very nature (e.g., flavor of wine), extrinsic properties (e.g., name, position etc.), parts if an object is structured, or relationship to other individuals. It is important to keep in mind that properties of general classes are inherited by specialized classes.

Properties are presented as slots, and slots can have different facets that define some constraints on properties like the value type, permitted values, cardinality, and other features of the value, the slot can take. For example, name is usually a slot with value type string, while age takes an integer value. If it is the age of an employee, some allowed values can be specified, for instance, it cannot be negative or higher than 100. The types that appear in ontology construction software are string, number, Boolean, enumerated, and instance type. The last one allows definitions of relationships between individuals.

The last step in an ontology construction is creating individual instances of classes in the hierarchy by filling in the slot values.

2) Learning Ontologies From Documents: There are several approaches to ontology learning from texts. They are based on the use of text corpora. A corpus is a set of texts that should be a representative of the domain, should be complete, and cover all the aspects of a certain domain, and in addition it should be accepted by the domain experts. All methods, which learn ontologies from texts, try to exploit certain linguistic features of terms and surrounding terms to build concepts. Therefore, we can identify two levels:

1) linguistic level, where knowledge is described through linguistic terms;
2) conceptual level, where knowledge is described using concepts and relations between concepts.

The different techniques of ontology learning from texts are based on how linguistic-level structures are projected or mapped to the conceptual level. To achieve this, methods use a combination of natural language processing and statistics. Khan and Luo [13] presented a method for building a domain ontology using clustering techniques and wordnet [14]. Pattern-based extraction [15] recognizes the relations at the conceptual level

from sequences of words in the text that follow a given pattern. Using association rules [16], relationships between concepts (perhaps already in hierarchy) can be established. Conceptual clustering [17] takes as an input a set of concepts, which are then grouped according to a semantic distance. The distance is calculated using syntactic role (e.g., *train* and *car* can appear with the same syntactic role in sentences "John travels by *car*" and "Ann travels by *train*"); therefore they are considered semantically close.

3) Other Techniques for Ontology Learning: Other techniques for ontology learning take advantage of the structure that is represented in the data. Learning from structured data takes one of the following forms.

1) Ontology learning from a dictionary relies on the use of a machine readable dictionary to extract relevant concepts and relations among them.
2) Ontology learning from semistructured data is trying to create an ontology from sources, which have any kind of predefined structure, such as XML schemes.
3) Ontology learning from relation schemes tries to learn an ontology by extracting relevant concepts and relations from knowledge in databases.

III. BUSINESS-RELATED ONTOLOGIES

In the last decade, many projects aimed at creating ontologies for different purposes (wordnet [14], cyc [18], sumo [19], etc.). This section focuses on the most important ontologies concerning the domain of business and enterprise modeling.

A. AIAI Enterprise Ontology

An enterprise ontology is a collection of terms and definitions used in organizations. The AIAI (Artificial Intelligence Applications Institute)enterprise ontology [20] was developed in the scope of the enterprise project, whose goal was to provide a set of tools for enterprise modeling. The available enterprise tool set contains a procedure builder, for capturing process models, an agent toolkit for supporting agent development, and a task manager for integration and visualization. The ontology was used in order to ensure a consistent communication between agents, either human or software. The enterprise ontology built within the enterprise project is not meant to be a complete ontology describing the enterprise domain. It only presents the most frequent terms used in this field. Thus, the ontology has to be enriched for each specific business case.

The enterprise ontology is divided into five top-level concepts: activities and processes, organization, strategy, marketing, and time. The organization part contains the terms representing the actors, which play a role in an enterprise. They can have legal responsibilities or not, be a human or a machine. These terms are then used to model activities and processes. The activity part includes the concept of resources and skills, which are needed and the effects of the activity. In other words, it contains the concept of input–output. The central concept of the strategy part is purpose. Purpose captures the idea either of something, which a "plan" can "help achieve" or that an organization unit can be responsible for. Finally, the marketing part

describes sales. Sale is an agreement between two legal entities for the exchange of a product for a sale–price.

B. Toronto Virtual Enterprise Ontology (TOVE)

The TOVE enterprise ontology was developed in the scope of the TOVE project [21]. The TOVE ontology is a formal representation of the enterprise domain. As the enterprise ontology, it is divided into several top-level concepts to segment the enterprise into general categories: activity, states, causality, time, resources, and organizational structure.

In this paper, we focus primarily on the resources and organization parts of the ontology. In TOVE, the resource ontology comprises two sets of terms or assertions. First, the resources are defined in terms of knowledge, role, mobility, and division of the resource. The role of the resource represents its nature, for instance, whether it is a product, a tool or a work area. The mobility specifies the possibility of moving the resource from one place to another or not. The divisibility of the resource specifies if the resource can be divided into several resources, without affecting its role in an activity. Each division must be able to be consumed by an activity. Once these basic terms are defined, more complex ones are introduced such as the nature of the resource or its capacity. The nature of a resource means that a resource can be continuous or discrete. The capacity of a resource represents its availability at a certain point of time.

In the organization ontology, an organization entity can be an organization–individual or an organization–group denoting several people (e.g., board of directors, teams, etc.). Each organization has properties, such as organization role, skills, constraints, etc. The organization role specifies the goal that the organization has to achieve. Each role has attached skills, processes, policies, etc., which are necessary to complete the goal.

The concepts encoded in the ontology are also enriched by a set of axioms that define and constrain the interpretation of these concepts. The ontology is formalized using first-order logic, allowing answering questions by using the TOVE reasoning engine.

C. Business Management Ontology

The primary goal of the business process management ontology (BPMO) [22] is to provide a stable platform for the semantically rich definition of business processes, in order to better align information technology (IT) with business. The BPMO allows to define private and public processes, business entities, business objects, and services that implement process activities. It follows the UN/CEFACT modeling methodology (UMM) for business process and information modeling.

The BPMO uses the concept of business entities and business objects for process modeling. Their definitions rely on the UN/CEFACT glossary, which defines a business entity as "something that is accessed, inspected, manipulated, produced, and so on in the business." Business entities are defined for all the terms that are intended to be used in business, and thus, form a kind of glossary. Once these entities have been defined, they are generalized under new concepts called business objects.

For instance, the business entities customer and supplier may be represented by a business object named party, which is a generalization of customer and supplier [23].

The BPMO also introduces the notion of process task concept type. It describes, which role performs a task, the business entities and business documents it is related with, and the resources it consumes.

Every task represents a defined context, which includes the following items [23].
1) Role: A logical abstraction of one or more physical actors, usually in terms of common responsibility or position. An actor may be a member of one or more roles. Example: mortgage clerk.
2) Business document: The set of information components that are interchanged as part of a task. A business document may participate in a message flow. Example: private mortgage loan application form.
3) Durable information entity: An information entity that a task needs to perform its function, which must be represented in a persistent storage mechanism, and whose state must exist beyond the lifetime of the service (application) that implements the task. It may be composed of multiple business objects. Example: private mortgage loan application information.
4) Resource: A real object that can be identified. Example: flatbed scanner.

Currently, the BPMO comprises approximately 650 classes. The ontology is available in the OWL format.

IV. ONTOLOGY FOR VIRTUAL ORGANIZATION BREEDING ENVIRONMENTS

There are many new concepts and terms in the field of collaborative networks (CNs) and CNO [1]. A structured way to achieve a common understanding of the newly introduced terminology is by using an ontology for this domain. Other reasons for ontology building include the ability to reuse the present knowledge, making domain assumptions explicit, splitting the declarative domain knowledge from the operational knowledge, and possible analysis of domain knowledge.

There are also more specific reasons due to the nature of working in CNOs. Such work is dynamic in the sense that organizations and individuals may join and leave CNOs. This section focuses on VBE[4] [3], whose main purpose is to enable fast VO creation.[5] An introductory and learning phase for organizations joining a VBE, should therefore, be as short as possible, and ontologies have proven to be an unambiguous and compact way of knowledge representation. In addition, ontologies provide a basis for sharing information not only among people but also among software agents. If several organizations joining a CNO share the same underlying concepts (e.g., on the web or on their

---

[4]Existing VBEs are usually called "clusters" or "industrial clusters," while some other authors use the term virtual web instead of VBE [24], [25].

[5]Since Mowshowitz [26] used the term virtual organization for the first time, the variety of different terms and definitions were created to describe this new form of network organizations, including virtual company [27] and virtual enterprise [28]. See [25] for a detailed explanation of various terms and concepts.

intranet), then software agents are able to extract and aggregate information and use it to gather the data and to answer queries. Such agents can also support a process of VO creation by proposing VOs based on their competencies for achieving a business goal.

For every ontology, its domain and scope should be determined first. This includes the specification of the domain and the potential use of the ontology. A set of questions an ontology should answer can also help determining the level of detail. First, there are some general questions concerning the newly introduced terms, and later some questions concerning the real cases (instances) of those and related terms. The questions one can ask are the following.

1) What is a CNO?
2) What is a VBE?
3) What is a VO?
4) What is the difference between VBE and VO?
5) What form can the following actors take (e.g., a Person, a company etc.) and what are the tasks of:
   a) VBE administrator;
   b) VBE adviser;
   c) Support institution;
   d) VBE partner;
   e) VO broker;
   f) VO planner;
   g) VO coordinator.
6) What is a business opportunity?
7) How is a business opportunity handled when it arises?
8) Does a VBE administrator have to be a member of the VBE?
9) Who was a broker of VO1?
10) How many times was X a VO broker?
11) Can a VO broker be also a VO coordinator?
12) Is it necessary for a VO support provider to be a VO partner?
13) Is it necessary for a VO partner to be a VBE member?

These questions show what level of detail goes into the operational phase of VBEs. From the questions we can also see that different parts of the ontology emerge:

1) a general part, which defines the structure and function of a VBE;
2) a part describing roles of participants of VBE and VO;
3) a part of the ontology describing organizations' competencies, resources, and their availability.

The ontology of this paper is based on the concepts developed in the ECOLEAD project (www.ecolead.org), as well as on the entity relationships and concepts for virtual organizations developed in the SolEuNet project [29]. The proposed CNO ontology is implemented within the Protégé framework following the methodology described in Section II-C1, and is available as GPL-licensed software through a public web site,[6] so that it can be redistributed and/or modified. The site includes also an extensive glossary of terms used in this paper. Moreover, the OWL plugin was chosen in order to formalize the concepts and allow reasoning in the future.

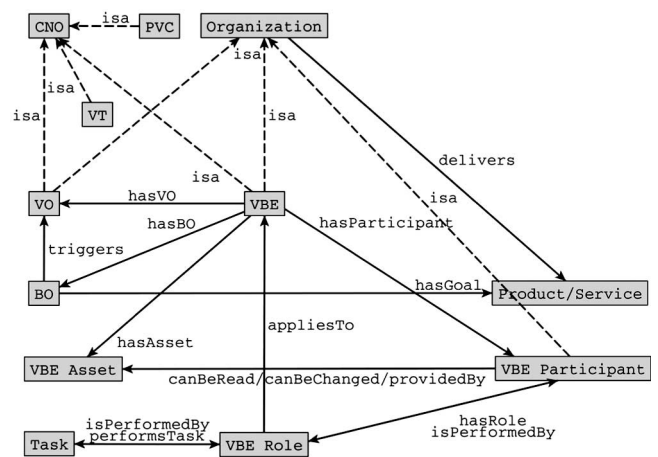[6]http://kt.ijs.si/software/CNOntology/

Fig. 1. Top-level concepts of the ontology. Specializations are shown using a dashed line, while other relations (along with their names) are represented with a solid line. Professional virtual communities (PVCs) and virtual teams (VTs) as forms of CNO are out of scope of this paper.

A. Top-Level of the CNO Ontology (Domain Ontology)

The top-level concepts linking individual parts of the VBE ontology are shown in Fig. 1. In further subsections other concepts of the ontology, including their instances for better understanding are introduced. The two top most concepts are (CNO) and organization. A CNO is a special type of CN comprising only organized collaborations while, in general, collaborative networks include both organized and nonorganized collaborations. More generally, a collaborative network [1] is constituted of a variety of entities (e.g., organizations and people) that are largely autonomous, geographically distributed, and heterogeneous in terms of their operating environment, culture, social capital, and goals. These entities collaborate to better achieve common or compatible goals, and their interactions are supported by a computer network.

Organization is a company, corporation, firm, enterprise or institution, or part thereof (whether incorporated or not, public or private) that has its own function(s) and administration, which supplies products or services to other organizations [30]. All organizations have their profiles, competencies, can perform one or more processes, and deliver some products and/or services. Special types of organizations are VO, VBE, and VBE participant. VO and VBE are also special types of CNOs, since they represent alliances of companies and individuals. VBE participant represents an entity collaborating with other entities in the VBE and VO. A VBE is "an association (also known as cluster) or pool of organizations and their related supporting institutions that have both the potential and the will to cooperate with each other through the establishment of a "base" long-term cooperation agreement and interoperable infrastructure" [1]. A VBE responds to business opportunities by forming VOs. As an organization, a VBE has also competencies, which are not limited to the union of the competencies of its participants. The VBE competencies are the result of combining two or more participants' competencies to realize more complex projects (e.g., building a highway, a bridge, etc.). Each participant can take one or more
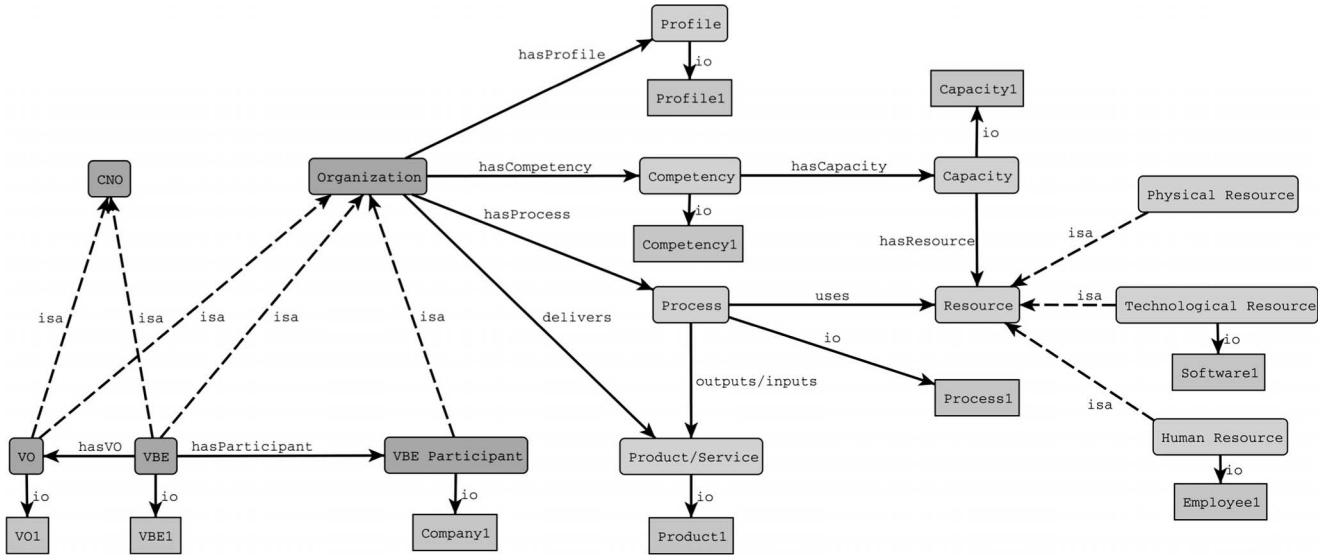
Fig. 2.   Part of the competencies ontology of an organization. Instances are linked to concepts with io (instance of) relations.

different roles that are further defined in subsequent subsections. On the other hand, a VO is a short-term association with a specific goal of being active in fulfilling a Business Opportunity (BO). A BO is a time or occasion with favorable combination of circumstances that is suitable to start a business. VO represents a temporary alliance of diverse organizations that form a collaborative network, sharing knowledge, skills, and resources in order to respond to a specific BO. The partners of the VO are selected from the VBE participants, according to their competencies and availability to deliver products or services required to fulfill the BO. To facilitate the processes within the VBE, the participants have access to several assets, such as business rules, software tools, specifically developed for the VBE. These assets are called VBE assets and are kept in the so-called VBE bag of assets; they are further described in Section IV-C.

B. Organization-Related Ontology

The organization part of the ontology introduces concepts necessary to describe organizations' ability to perform certain operations at a certain time. Concepts and their relations are shown in Fig. 2. The concept profile is a set of structured information describing the organization in general such as name, contact information, description, etc. Each organization covers one or more competencies, which define its capability to perform processes. A process is a structured, managed, and controlled set of interrelated activities that uses resources to transform inputs into specified outputs. The final aim of an organization is to offer certain products or services to the customers at a certain time. Therefore, competency is related to capacity, describing its availability in terms of start time, duration, and availability. The attribute availability of capacity is expressed as percentage, and specifies, which amount of the resource can be used (for instance, a software engineer capable of writing 1000 lines of code per week available for one month at 50% means that he will be able to write 1000*4*0.50 = 2000 lines of code). Resource represents an element consumed in a process, which performs
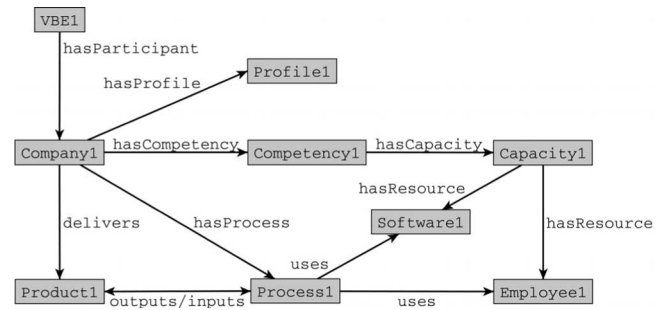

Fig. 3.   Example of instantiation of organization's ontology.

a number of operations that can transform inputs into outputs. There are three specializations of resources, namely physical, technological, and human resource. Organization's physical resource refers to buildings, machines, equipment, etc., within the organization, technological resource refers to hardware and software within the organization, while human resource refers to the individuals within the organization. Resources are split into three categories, due to different attributes used to describe them.

Fig. 3 shows instances and their relations for better understanding of the ontology. As an example there is one instance of VBE (VBE1), which has a participant named company1. This company has a certain profile and only one competency. This competency is available at a certain time (capacity1) with two resources, namely software1 and employee1. Company1 delivers its product (product1) through process1, that uses previously mentioned resources.

1) VBE Role-Related Ontology: A participant is any organization within a VBE and can participate in a different way. Each participant in the VBE can have one or more specific roles, which are classified into three main categories:

1) roles that apply to members of the VBE;
2) roles that apply to support institutions;
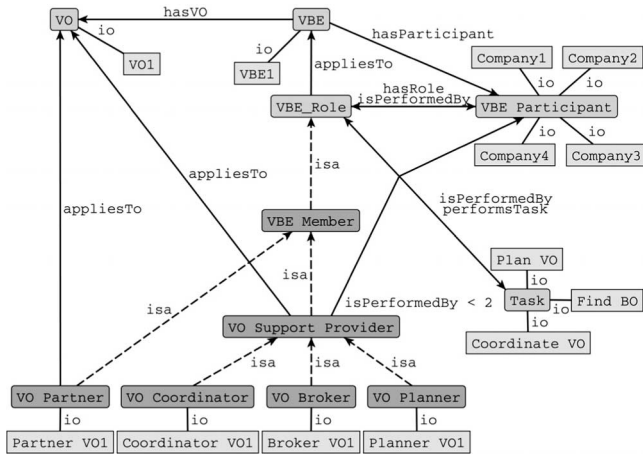3) roles that apply to public entities.

Fig. 4.   Roles and their structuring within a VBE.

Fig. 5.   Example of VBE, VO, its participants, and their roles.

All these roles have in common a relation with the concept task, which specifies all the necessary tasks to be performed within the VBE and VO. For each role, certain tasks must be performed. Public entity is the role taken by a participant, which is not registered in the VBE. We do not cover it in any detail here. Member and support institution roles are defined in the following subsections since they are more complex and relevant for the ontology.

2) VBE Member Role: VBE member is the basic role (shown in Fig. 4) of those organizations registered in the VBE and are willing to participate in the VBE's activities. The principal activity carried out by members of the VBE is to collaborate in VOs. When involved in one or more VOs, they can take different roles such as simple VO partner or VO support provider. The support provider role is further divided into three subroles: VO coordinator, VO broker, and VO planner. These roles can be performed by only one VBE member, whereas the role of simple VO partner can be taken by more than one member that constitute the VO. VO partner is the role of a VBE participant in a VO. Opportunity broker is the role of a participant, which identifies and acquires new collaboration opportunities BO, by marketing VBE competencies and assets and negotiating with potential customers. The VO planner is responsible for identifying the necessary competencies and capacities, select the appropriate partners accordingly and structure the new VO. In many cases, the roles of broker and planner are performed by the same actor. The VO coordinator coordinates the VO during its life cycle in order to fulfill its goals. Fig. 5 shows an example using instances of the concepts previously defined. The figure illustrates the case of a VBE (VBE1) with four participants (companies 1–4). Each participant is assigned a role, VO partner, broker, planner, and coordinator, respectively, for companies 1–4. Each role has assigned a certain number of tasks that the owner of the role has to perform. In addition, each role is attached to a specific VBE and VO, since a member can have several roles in different VBEs and VOs.

3) Support-Institution-Related Ontology: Many industry sectors (such as real estate, construction, banking, education, maintenance, etc.) collaborate with several clusters and serve
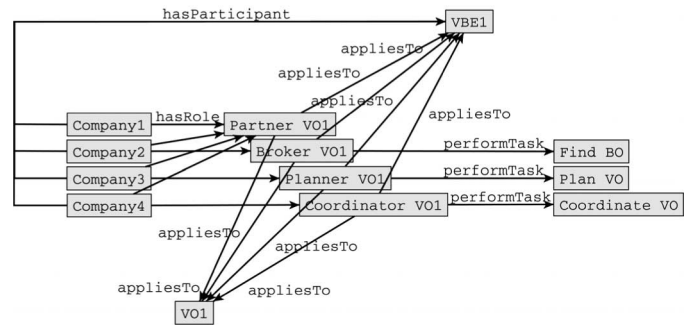
as support institutions that provide services. The same is true within a VBE: the VBE has its own competencies to solve specific problems, but may require the help of support institutions for other tasks where it does not have enough expertise or does not want to get involved. Support institutions can provide a broad range of services, such as training, research, consulting, information services, legal and contractual services, etc. Moreover, support institutions always try to keep up with new trends and technologies in their field of expertise, and thus, tend to propose a solution that is up to date and optimized for the client. Usually, employees inside the support institution are also trained in order to better understand the needs of the client and try to continuously improve their methods.

Support institutions can be used in a VBE for different purposes:
1) to promote entrepreneurship;
2) to promote VBE capabilities;
3) to help VBE members to achieve competitiveness;
4) to provide an appropriate ICT infrastructure;
5) to make easy the installation of enterprises that add value to the VBE;
6) to make easy the application of modern theories in industry;
7) to provide the best coordinated support;
8) to direct funds for industry development in a VBE;
9) to open markets via interregional and international trade agreements;
10) to encourage SME's growth;
11) to provide appropriate infrastructures;
12) to collaborate with brokers in commercial missions;
13) to collaborate between diverse governmental levels (e.g., local, regional, and national), and/or
14) to support research and development activities.

For instance, building companies make use of legal support institutions to have insurance for each of their contracts. Support institutions can even be used in collaborative networked organizations to train new partners, perform administrative tasks, or even to make and keep the network running. Some clusters used in the case study in Section V use a support institution to create the business development of the cluster and to implement new innovative concepts. In this case, the support institution provides services in the form of consulting or even workshops, where tools are designed specifically for the needs of the cluster.
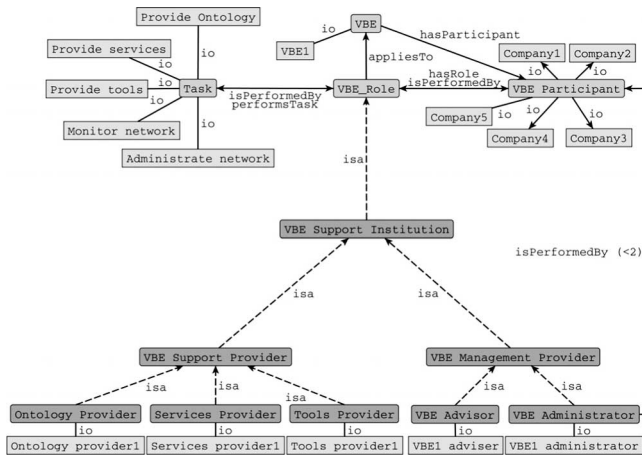
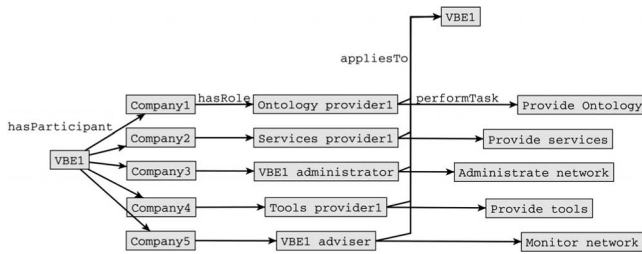Fig. 6. Support institution roles and their instantiations.

Fig. 7. Example of VBE with five participants.

A support institution also helps to develop marketing strategies and Public Relations (PR) materials such as leaflets, web sites, etc., to promote new products. If the support institution invests sufficient effort and involvement, its help may increase the productivity of the client, and thus, lead to the increase of the revenues.

In the ontology of Fig. 6, the role VBE support institution is divided into two main subroles:

1) VBE support provider, which includes ontology provider, services provider, and tools provider roles;
2) VBE management provider which includes VBE adviser and VBE administrator roles.

The roles under the support provider concept deliver services useful for the VBE, such as an ontology (ontology provider), software tools (tools provider), and legal services (services provider). The roles under the management provider concept deliver management services. The VBE administrator is the role of a VBE participant responsible for the VBE operation and evolution, promotion, and cooperation among the VBE members. He/she is also filling the skills/competencies gaps in the VBE by searching and inviting new organizations in the VBE and manages the VBE general processes.

The VBE Adviser is responsible to monitor the network and suggest recommendations to the administrator.

Fig. 7 shows an example of a VBE with five participants (companies 1–5), which have support institution roles. Each of them is attached to a specific task with respect to the role. Each

role is attached to a specific VBE, since a participant can have several roles in different VBEs.

C. VBE Bag of Assets-Related Ontology

In business and accounting an asset is anything owned, which can produce future economic benefit. It can take a form of possession or right to take possession. It can be owned by a person or a group acting together, e.g., a company. Its value can be expressed in monetary terms. Assets can be classified according to the generally accepted accounting principles as follows.

1) Current assets: Cash and other assets expected to be converted to cash.
2) Long-term investments: These are to be held for many years and are not intended to be disposed in the near future.
3) Fixed assets: Purchased for continued and long-term use in earning profit in a business. This group includes land, buildings, machinery, furniture, tools, wasting resources (assets which decline in value over time, e.g., gas, oil, etc.).
4) Intangible assets: These lack physical substance and usually are hard to evaluate. They include patents, copyrights, franchises, goodwill, trademarks, trade names, etc.
5) Other assets: This type includes a high variety of assets, most commonly long-term prepaid expenses, long-term receivables, property held for sale, etc.

All these types of assets can be found inside a single company. Some of them can also be a property of a VBE. The main purpose of the VBE related assets is to speed up and improve the process of a VO creation (which is the main task of a VBE).

In a VBE, the following potential assets have been identified.

1) General policies in the form of documents, books, leaflets to help (new and old) members to easily follow the guidelines of a VBE.
2) Sample contracts to speed up the contracting phase.
3) General legal issues related to the sector.
4) Information of interest, specific to the sector.
5) Links to other sources of information.
6) "Lessons learned." This is a database system designed to collect and make available lessons learned in the business. It enables the knowledge gained from past experience to be applied to current and future projects. Its intention is to avoid the repetition of past failures and mishaps, as well as the ability to share observations and best practices. Through this resource members seek to facilitate the early incorporation of quality into the design of their products and services.
7) FAQs.

The VBE asset concept comprises the name and the description of an asset. Each asset belongs to a VBE encoded as the property has asset of the concept VBE. Different rights to access an asset can be specified for reading and changing using properties can read and can change, respectively. The owner of an asset is identified using a property provided by. The users

and the owner of an asset must be of the type VBE participant, as presented in Fig. 1.

V. EXTENDING THE VBE ONTOLOGY WITH COMPANY-SPECIFIC INFORMATION

The competency part of the VBE ontology is meant to keep information about companies' competencies and their availability. This information is specific to each VBE and must be identified and structured at the creation phase of the VBE. This information is updated every time a partner joins or leaves the network, or when an existing partner acquires new competencies. The number of domains and their diversity makes it unmanageable to create a universal competency ontology, which can be used in all VBEs. Therefore, we propose a methodology that speeds up the process of gathering and structuring companies' competencies, using machine learning techniques. Identification of a hierarchy of competencies (see Section IV-B) and linking the organizations with competencies (with the "has competency" relation) is performed using clustering on company profiles.

A. Methodology for Semiautomated Ontology Construction

The proposed methodology for structuring the competencies consists of the following steps.

1) Data gathering (yields textual data).
   a) Data can be gathered manually through questionnaires filled in by companies.
   b) Alternatively, data are also available on the web, including company home pages and public registers. In this case, a data-gathering method employed can be focused web crawling [31].
2) Preprocessing [of textual data into bag-of-words (BOW)]. Raw textual data is processed in the following steps.
   a) Markup tags and stop-words elimination.
   b) Stemming and/or lemmatization. Words are presented in 'normal' form by elimination of suffixes and prefixes [32].
   c) Transformation into the BOW representation, where a document is encoded as a feature vector with word frequencies as elements. Elements of vectors are weighted with weights inverse document frequency (*IDF*) [33]. All the ith elements are multiplied with $IDF_i = \log(N/df_i)$, where N is the total number of documents and df_i is document frequency of the ith word (the number of documents in which the ith word appears). Such vectors are called *TFIDF* vectors.[7]
3) Structuring (of BOW into clusters): Structuring of the BOW representations is performed by document cluster-

ing [34]. We applied document clustering to automatically build a hierarchy of companies, based on their descriptions, with a subset of relationships between the groups of companies. In our experiments, we used a k-means hierarchical clustering system gCLUTO [35]. The result of clustering is a simple ontology—a taxonomy, which is a tree structure with classes, subclasses, and instances.

4) Visualization (of the taxonomy): Many methods were developed for visualization of text documents, or high-dimensional data in general. Some examples are themeview, themeriver, topic islands [36], and self-organizing maps [37]. In this work, we have applied mountain visualization [35].
5) Ontology evaluation means a comparison to existing ontologies and/or manual evaluation by domain experts. Subsequently, the developed hierarchy can be manually refined and elaborated to improve the deficiencies revealed by the evaluation.

B. Competency Structuring of Virtuelle Fabrik

The proposed methodology for structuring the competencies was applied to a real-life VBE, the Virtuelle Fabrik industrial cluster of mechanical engineering companies.[8] The steps taken were as follows.

1) VF company descriptions were collected manually through questionnaires and translated from German to English.
2) Descriptions were transformed into BOW representation.
3) Structuring was performed by the application of gCLUTO.
4) Mountain visualization of obtained clusters was applied.
5) Results were verified by a human expert (the VO broker).

Company profiles of 50 partners of the Virtuelle Fabrik industrial cluster were made available for the experiment. Each company was described by its name, number of employees, products, services, and their core competencies.

C. Data Gathering and Preprocessing

The procedure of translation of German company descriptions into English was as follows.

1) Each company was assigned a unique numeric identifier. The company identifier and company name were not used as input information for text clustering.
2) An initial document, written in German, was taken. It included descriptions of companies, their competencies, and products.
3) From the document, the data on competencies and products of the companies was extracted, resulting in about 1300 words.
4) All stop words were removed (i.e., und, mit, zu, verbs, adjectives).
5) The translation of the remaining set of German words into English was performed by a web-based translation engine.

---

[7]TFIDF (term frequency—inverse document frequency) is a weight often used in information retrieval and text mining. This weight is a statistical measure used to evaluate how important a word is to a document. The importance increases proportionally to the number of times a word appears in the document, but is offset by how common the word is in all of the documents in the collection or corpus.
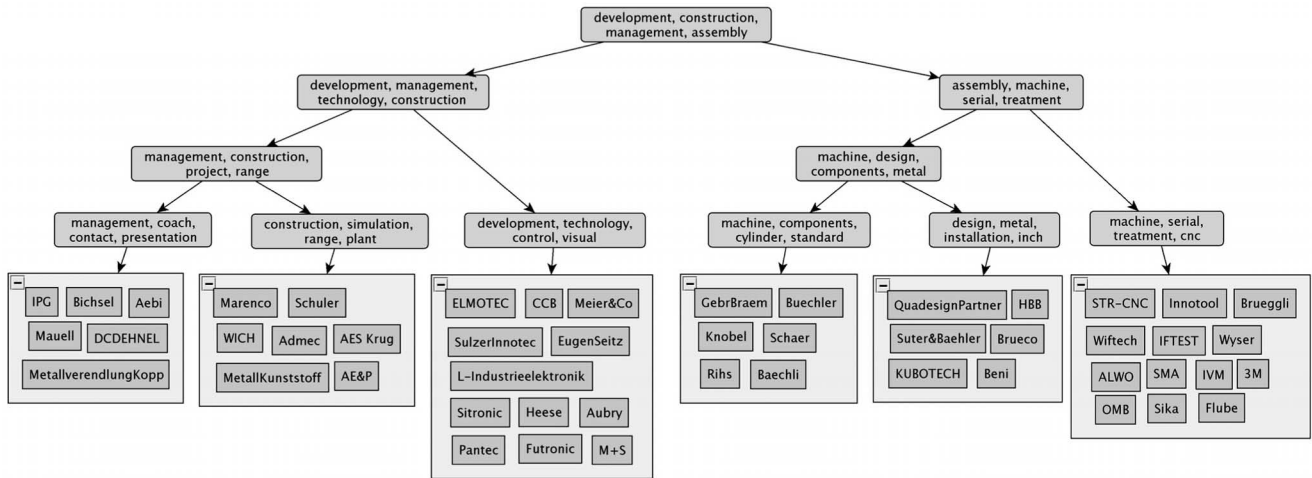
[8]http://www.vfeb.ch/

Fig. 8. Hierarchy of 50 Virtuelle Fabrik company descriptions obtained by agglomerating bottom level clusters into six higher-level clusters.

6) Manual translation of nontranslated words was performed, since the system was unable to translate approximately 5% of the words.

7) Finally, the unification of the words was performed by word lemmatization (e.g., "system" and "systems" were transformed into "system").

8) Text documents were transformed into word lists, representing simplified descriptions of the original text describing the companies.

D. Clustering

*g*CLUTO is a publicly available interactive clustering, visualization, and analysis system [35]. *g*CLUTO performs stop-words removal and stemming in text preprocessing, followed by *k*-means clustering, using a predefined number of clusters of leaf-level nodes as the stopping criterion.

The initial dendrogram induced from 50 company profiles was obtained by the *k*-means hierarchical clustering method available as part of *g*CLUTO. Each node in the dendrogram is automatically accompanied by a list of most representative words from the document/cluster. We have decided, due to space restrictions, to limit the company descriptions to only four most representative keywords. In *k*-means clustering, we selected $k = 2$, in order to get simple hierarchical splits. We decided to limit the description of concepts at higher levels of the concept hierarchy also to four keywords only, in order to preserve the readability of the hierarchical structure of concepts.

The hierarchical structure of Fig. 8 was produced from the initial dendrogram, by cutting the dendrogram at the level, where differences between successive cluster levels are maximal. This resulted in six competency class clusters, described by automatically extracted keywords. Company names are added to the leaves of the hierarchy, in order to simplify the interpretation of the obtained structure.

E. Mountain Visualization of Profile Categories

The *g*CLUTO system [35] offers advanced cluster visualization tools, which we have used to visualize the results of VF
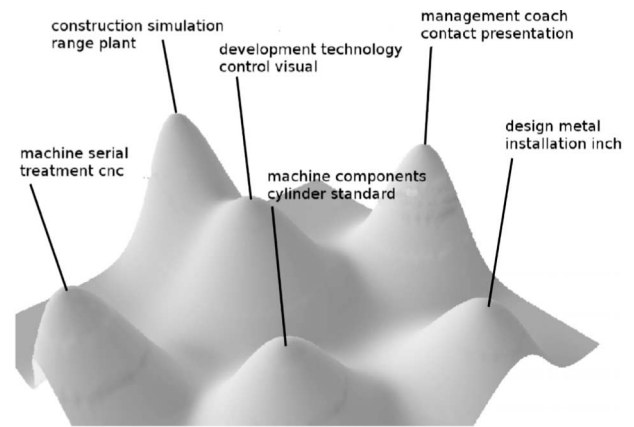

Fig. 9. Mountain visualization of six clusters, described with the most descriptive words, for 50 Virtuelle Fabrik companies.

clustering. The results in Fig. 9 provide an overview of the clusters of competencies, their strength and their homogeneity. Peaks in Fig. 9 represent individual clusters. The shape of each peak is a Gaussian curve, used as a rough estimate of the distribution of the data within each cluster. The height of each peak is proportional to the clusters internal similarity. The volume of a peak is proportional to the number of elements contained within the cluster. The resulting Gaussian curves are added together to form the terrain of the mountain visualization of *g*CLUTO.

VI. VALIDATION

The proposed formalization of the VBE ontology and the result of automated VBE competency structuring were validated in two ways. First, the ontology was instantiated to two existing VOs, and second, we collected responses from VBE experts regarding the appropriateness of the formalized concepts and relations in the ontology.
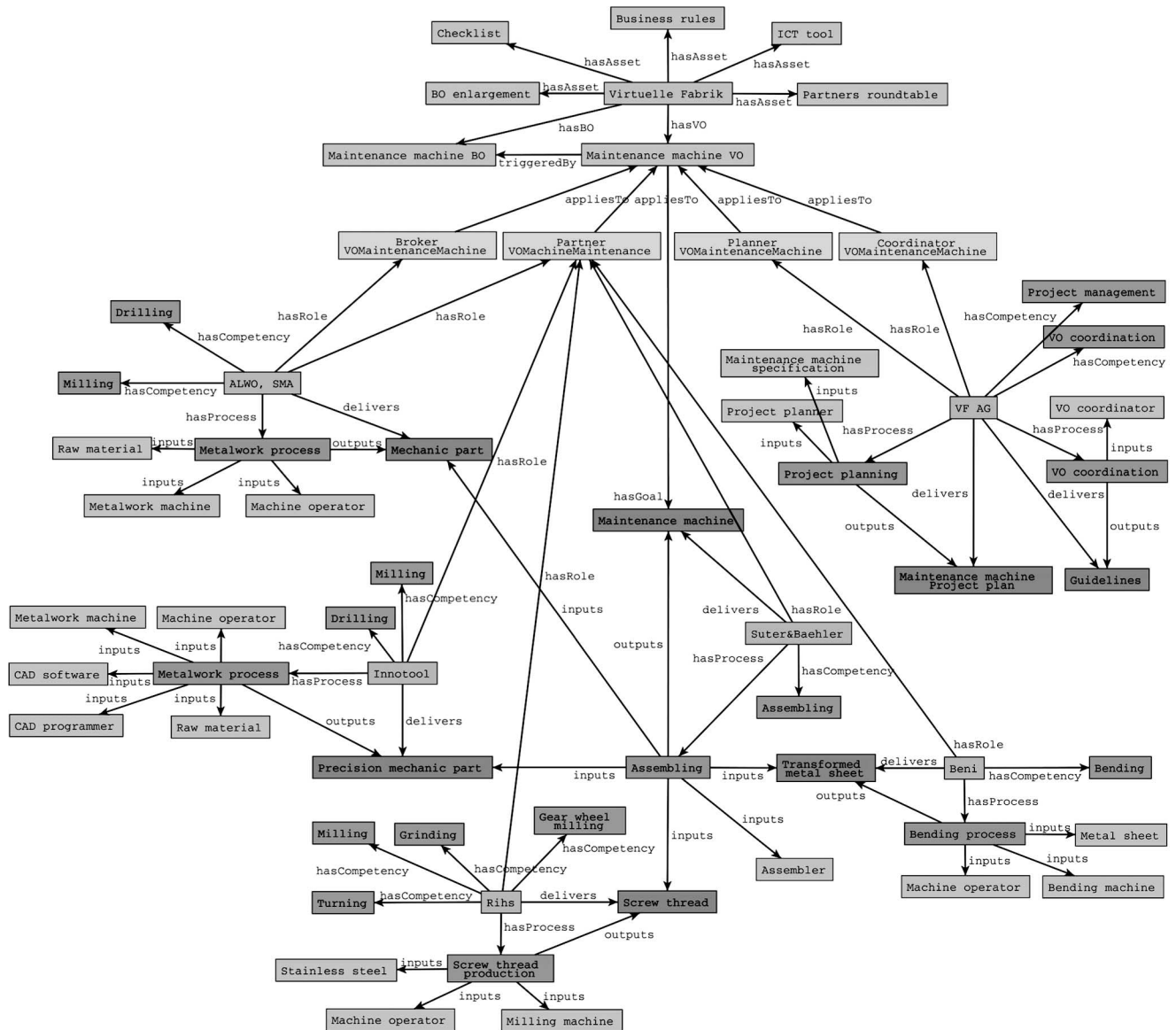
Fig. 10. VO created (within VF) for the task of construction of a maintenance machine for a nuclear power plant.

A. Instantiation of the VBE Ontology on Two Cases of VO

We have instantiated the VBE ontology on two actual VOs formed by the Virtuelle Fabrik industrial cluster. The first VO was created for the task of construction of a maintenance machine for a nuclear power plant. The second VO was formed for the task of a gearbox construction. The two instances of the VO ontology are shown in Figs. 10 and 11, respectively.
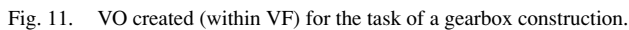
In accordance with the VBE ontology, each VO consists of partners that have one of the following roles: VO partner, or VO support provider (VO coordinator, VO broker, or VO planner). In this validation experiment, we were particularly interested in the VO partners and their competencies.

The VO created for the task of construction of a maintenance machine for a nuclear power plant, shown as a node "maintenance machine VO" in Fig. 10, consisted of five VO partners: ALWO, SMA, Innotool, Rihs, Suter and Baehler, and Beni, shown as nodes in the figure. Each partner is linked to its

competencies (needed for the particular business opportunity for which the VO has been created) by the "has competency" relation.

Notice that these five partners all appear on the right-hand side node (super-cluster labeled by words "assembly, machine, serial, treatment") in Fig. 8. This indicates that they have stronger similarity in terms of their descriptions as compared to the companies on the left-hand side node. However, they are also diverse, as they belong to three different subclusters of the right-hand side supercluster, meaning that they have some complementary competencies.

Similarly, the VO created for the task of a gearbox construction, shown in Fig. 11, consists of five companies: Beni, Knobbel, ALWO, SMA, and Innotool, which also fall in the same supercluster of Fig. 8. Again, these companies are sufficiently diverse, as they belong to the three different subclusters of the right-hand side supercluster. The five partners were not

Fig. 11. VO created (within VF) for the task of a gearbox construction.

sufficient for completing the task, therefore external partners (with the role "public entity") were contracted in order to cover the missing competencies: Stebler and Brunner.

In addition to the above strengths—the usefulness of the VBE ontology for formalizing the particular VOs—we can identify also some weaknesses.

1) First, the labeling of the cluster nodes (the 50-company hierarchy of Fig. 8) is not satisfactory, since the labels were chosen automatically by the algorithm. One of the reasons for such labels is the fact that companies were described with relatively short descriptions. Instead, for text and web mining, rich and wordy descriptions would yield better results, since the techniques can extract more essential and discriminating terms from large and redundant texts. Note that, in this experiment we have performed only steps 1–4 of the proposed methodology for semiautomated ontology

construction. In step 5, the human expert should actually revise the ontology and name the nodes appropriately by more meaningful cluster names.

2) Second, the competency structure is limited to a tree structure (a taxonomy) in contrast to a thesaurus. A thesaurus allows to link individual leaves to several nodes in the cluster, which is often more desirable.

Note that the above validation is qualitative only, as no quantitative measure of quality can be applied. We have, however shown the appropriateness of the semiautomated ontology construction approach in another study [38], in which we applied the described methodology to over 7000 company descriptions from the Yahoo! business directory. In this study, we have been able to quantitatively evaluate the approach by comparing the automatically constructed clusters with the original Yahoo!'s structure, created manually.

B. Feedback From VBE Experts

Further validation of the ontology involved asking the representatives of two different VBEs (Virtuelle Fabrik and IECOS) involved in the ECOLEAD project, whether the proposed VBE ontology accurately describes their activities and can help them in their work. Below are the comments that we received.

1) Virtuelle Fabrik: "The paper describes our organization in a very accurate way. Especially Fig. 11 could help us to illustrate activities within the scope of a running VO. The ontological description of competencies will also help us in further software projects and ECOLEAD activities to improve the accuracy of assumptions."

2) IECOS: "For IECOS, the ontology you propose in your paper is accurate and appropriate for describing the VBE. We think it is complete as it is, and is also simple and easy to be understood by anyone, but could be improved with some of our feedback comments below." The IECOS feedback comments are summarized in the following items.

    a) First, there should be a link between "process" and "competency" (see Figs. 2 and 3 in the paper) since "a competency allows to perform a process(es)."

    b) Another observation that could be considered is to add the class "practice" to the ontology and relate it to "process" because "a process is supported by a practice(s)."

    c) Finally, a "competency" is related to a "process" through the element "capability" because "a competency has a processing capability related to a process."

VII. Conclusion

The paper presents an ontology for a large class of CNOs, namely VBEs, which deal with long-term alliances of COs. The practical emphasis of the paper is on structuring the competencies of VBEs, and on the ontology instantiation on two case studies. Competencies play a major role in identifying business opportunities, acquiring projects, and forming VOs that run and complete the projects. The ontology has been validated through its instantiation on the two case studies and by the feedback obtained from the representatives of two VBEs involved in the ECOLEAD project.

In the future, we will extend the ontology to cover collaborating individuals in the form of Professional Virtual Communities (PVCs) and Virtual Teams (VTs). The generic ontology will be also incorporated into the VBE management software, which will support the complete VBE/VO life cycle: VBE creation and instantiation, search and identification of collaboration opportunities, as well as VO creation, management, operation, and dissolution.

References

[1] L. M. Camarinha-Matos and H. Afsarmanesh, "Collaborative networks: A new scientific discipline," *Intell. Manuf.*, vol. 16, pp. 439–452, 2005.

[2] Keywords and definitions for io [Online]. Available: http://ccs.mit.edu/21c/iokey.html.

[3] L. M. Camarinha-Matos and H. Afsarmanesh, "Elements of a base VE infrastructure," *Comput. Ind.*, vol. 51, pp. 139–163, 2003.

[4] T. R. Gruber, "A translation approach to portable ontologies," *Knowl. Acquis.*, vol. 5, no. 2, pp. 199–220, 1993.

[5] A. Gomez-Perez, M. Fernandez-Lopez, and O. Corcho, *Ontological Engineering*. Berlin, Germany: Springer-Verlag, 2004.

[6] J. Breuker, A. Valente, and R. G. F. Winkels, "Legal ontologies: A functional view," in *Proc. 1st Int. Workshop Leg. Ontologies*, 1997, pp. 23–36.

[7] G. van Heijst, A. T. Schreiber, and B. J. Wielinga, "Using explicit ontologies in KBS development," *Int. J. Hum. Comput.*, 1996, pp. 183–292.

[8] D. Beckett. (2004). RDF/XML syntax specification. W3C recommendation published [Online]. Available: http://www.w3.org/TR/rdf-syntax-grammar/.

[9] D. Brickley. (2004). RDF vocabulary description language 1.0: RDF Schema. W3C recommendation published online [Online]. Available: http://www.w3.org/TR/rdf-schema/.

[10] D. L. McGuinness and F. van Harmelen. (2004). OWL Web ontology language overview. W3C recommendation published [Online]. Available: http://www.w3.org/TR/owl-features/.

[11] N. F. Noy and D. L. McGuinness, "Ontology development 101: A guide to creating your first ontology," Stanford Knowl. Sys. Lab., Stanford, CA, Tech. Rep. KSL-01-05, 2001.

[12] "The Protégé project." (2000) [Online]. Available: http://protege.stanford.edu.

[13] L. Khan and F. Luo, "Ontology construction for information selection," in *Proc. 14th IEEE Int. Conf. Tools Artif. Intell.*, Washington, DC, 2002, pp. 122–127.

[14] C. Fellbaum, *WordNet, An Electronic Lexical Database*. Cambridge, MA: MIT Press, 1998.

[15] M. A. Hearst, "Automatic acquisition of hyponyms from large text corpora," in *Proc. 15th Int. Conf. Comput. Linguist.*, Nantes, France, 1992, pp. 539–545.

[16] R. Agrawal, T. Imielinski, and A. N. Swami, "Mining association rules between sets of items in large databases," in *Proc. 1993 ACM SIGMOD Int. Conf. Manage. Data*, pp. 207–216.

[17] R. Michalski, "Knowledge acquisition through conceptual clustering: A theoretical framework and algorithm for partitioning data into conjunctive concepts," *Int. J. Policy Anal. Inf. Syst.*, vol. 4, pp. 219–243, 1980.

[18] D. Lenat and R. Guha, *Building Large Knowledge Based Systems: Representation and Inference in the CYC Project*. Reading, MA: Addison-Wesley, 1990.

[19] I. Niles and A. Pease. (2001). "Towards a standard upper ontology" [Online]. Available: http://citeseer.ist.psu.edu/niles01towards.html.

[20] M. Uschold, M. King, S. Moralee, and Y. Zorgios, "The enterprise ontology," *Knowl. Eng. Rev.*, vol. 13, Nov. 1998.

[21] M. S. Fox, "The TOVE project: A common-sense model of the enterprise, industrial and engineering applications of artificial intelligence and expert systems," in *Industrial and Engineering Applications of Artificial Intelligence and Expert Systems*. Lecture Notes in Artificial Intelligence, vol. 604, Berlin, Germany: Springer-Verlag, 1992, pp. 25–34.

[22] The open source business management ontology. (BMO) [Online]. Available: http://www.bpiresearch.com/Resources/RE_OSSOnt/re_ossont.htm.

[23] BPMO tutorial defining a private business process in a knowledge base. (2003). [Online]. Available: http://www.bpiresearch.com/BPMO_Tutorial.pdf.

[24] S. Goldman, R. Nagel, and K. Preiss, *Agile Competitors and Virtual Organizations: Strategies for Enriching the Customer*. New York: Reinhold, 1995.

[25] U. Franke, "The knowledge-based view (KBV) of the virtual web, the virtual corporation, and the net-broker," in *Knowledge Management in Virtual Organizations*, Hershey, PA: Idea Group, 2000, pp. 20–41.

[26] A. Moshowitz, "Social dimensions of office automation," *Adv. Comput.*, vol. 25, pp. 335–404, 1986.

[27] S. Goldman and R. Nagel, "Management, technology and agility: The emergence of the new era in manufacturing," *Int. J. Technol. Manag.*, vol. 8, no. 1/2, pp. 18–38, 1993.

[28] M. Hardwick, D. Spooner, T. Rando, and K. Morris, "Sharing manufacturing information in virtual enterprises," *Commun. ACM*, vol. 39, no. 2, pp. 46–54, Feb. 1996.

[29] D. Mladenić, N. Lavrač, and S. Moyle, *Data Mining and Decision Support: Integration and Collaboration*. Boston, MA: Kulwer, 2003.

[30] ISO definitions [Online]. Available: http://www.bizmanualz.com/ISO9000-2000/ISO_definitions.html.

[31] M. Ester, M. Gross, and H.-P. Kriegel, "Focused web crawling: A generic framework for specifying the user interest and for adaptive crawling strategies," presented at the 27th Int. Conf. Very Large Databases, Rome, Italy, 2001.

[32] M. Porter, "An algorithm for suffix stripping," *Program*, vol. 14, no. 3, pp. 130–137, Jul. 1980.

[33] S. C. Deerwester, S. T. Dumais, T. K. Landauer, G. W. Furnas, and R. A. Harshman, "Indexing by latent semantic analysis," *J. Amer. Soci. Inf. Sci.*, vol. 41, no. 6, pp. 391–407, 1990.

[34] M. Steinbach, G. Karypis, and V. Kumar, "A comparison of document clustering techniques," in *Proc. KDD Workshop Text Min. 6th ACM SIGKDD Int. Conf. Knowl. Discovery Data Min.*, Aug. 2000, pp. 109–110.

[35] M. Rasmussen and G. Karypis, "gCLUTO—An interactive clustering, visualization, and analysis system," Dep. Comput. Sci. Eng., Univ. Minnesota, Tech. Rep. TR 04-021, 2004.

[36] N. E. Miller, P. C. Wong, M. Brewster, and H. Foote, "TOPIC ISLANDS— A wavelet-based text visualization system," in *Proc. IEEE Vis. '98*, pp. 189–196.

[37] T. Kohonen, *Self-Organization and Associative Memory*. 3rd ed. Berlin, Germany: Springer-Verlag, 1989.

[38] P. Ljubič, N. Lavrač, D. Mladenić, J. Plisson, and I. Mozetič, "Automated structuring of company profiles," *Metodološki zvezki*, vol. 3, no. 2, 2006.

Peter Ljubič received the B.Sc. degree in computer science from the University of Ljubljana, Ljubljana, Slovenia, in 2002.

He is currently a Researcher in the Department of Knowledge Technologies, Jožef Stefan Institute, Ljubljana. His current research interests include machine learning, data-mining, and natural language processing.

Igor Mozetič received the Ph.D. degree in computer science from the University of Ljubljana, in 1988.

He is a Senior Associate in the Department of Knowledge Technologies, Jožef Stefan Institute, Ljubljana, Slovenia. He was a Fulbright Scholar at the University of Illinois at Urbana-Champaign, Urbana-Champaign, IL. He was a Visiting Professor at George Mason University and at Technical University of Vienna, Vienna, Austria. He was also a Senior Research Fellow at the Austrian Research Institute for Artificial Intelligence. His current research interests include machine learning, logic-based knowledge representation, automated reasoning, and semantic web.

Nada Lavrač received the Ph.D. degree in technical sciences from the University of Maribor, Maribor, Slovenia, in 1990.

She is currently the head of the Department of Knowledge Technologies, Jožef Stefan Institute, Ljubljana, Slovenia. From 1993 to 1996, she was the Scientific Coordinator of the European Scientific Network in Inductive Logic Programming (ILPNET) and from 2000 to 2003, a Cocoordinator of the 5FP EU project Data Mining and Decision Support for Business Competitiveness: A European Virtual Enterprise (SolEuNet). She is the author and editor of numerous books and conference proceedings, including *Inductive Logic Programming: Techniques and Applications* (Kluwer 1997), *Relational Data Mining* (Springer 2002), and *Data Mining and Decision Support: Integration and Collaboration* (Kluwer 2003). Her current research interests include machine learning, relational data mining, knowledge management, and applications of intelligent data analysis in virtual organizations, medicine, and bioinformatics.

Joël Plisson received the B.Sc. degree in computer science from Epitech School, Paris, France, and is currently working toward the Ph.D. degree at the International Postgraduate School Jožef Stefan, Ljubljana, Slovenia.

He is a Researcher in the Department of Knowledge Technologies, Jožef Stefan Institute. His current research interests include natural language processing, text mining, and machine learning.