ISSN:0975-9646

Interoperability between Clouds: an Exploratory Study on Virtual Organizations and Linked Data

Liviu Gabriel Cretu
Business Information Systems Department, Alexandru Ioan Cuza University of Iasi
Bd. Carol I, no 11, Iasi, Romania

***Abstract*— Inter-operability between services provided by different clouds has become one important research topic now that the new information silos have moved to the clouds. Since the end user has virtually no tool at hand to address inter-operability issues between closed cloud environments, different integration and orchestration techniques are needed to deploy integration agents external to the service execution platform. In this article we discuss the opportunity of adopting virtual organization (VO) models for services orchestration using Semantic Web (SW) technologies and Linked Data (LD). Following an exploratory study on VOs and LD, the article introduces the idea of an open model for virtual organizations having no central repository or broker. In the end we discuss a number of research challenges towards the technical implementation of the model..**

***Keywords*— Cloud computing, virtual organizations, Linked Data, Semantic Web.**

I. Introduction

According to the research made by the Open Group [1], integration issues are second, after security, biggest concern of companies who planned to use the clouds. Moreover, the US Federal Cloud Computing Strategy suggests [2] that, at least in the first stage of migration, enterprises will only partially move their processes in the public clouds (mostly, non-core business processes) and they will have to carefully select the providers for different service types. There are two critical issues that appear when one company is planning to adopt cloud solutions: 1) lack of tools to integrate on-premises business information systems with the cloud (at least for data, service and security levels); 2) even if it has the internal pool of resources to create such tools, most clouds are closed to public access to their core components and data structures. Moreover, once services migrated into clouds, the main question is how to integrate resources from different clouds within various business processes.

One way to assembly services from different clouds in one single process would be to use virtual organization (VO) models. The virtual organization is an entity created as a result of ad-hoc selection and configuration of autonomous services, assembled and coordinated by a broker in order to serve a particular request [3]. The configuration might be the same or can change at different moments in time, depending on contextual needs. The main difficulty in deploying any VO solution is to build common vocabularies needed to identify partners and create inter-operable networks of services. To overcome such limitations, there are a number of VO frameworks that emphasize the benefits of using ontologies as a source of shared knowledge and Semantic Web (SW) technologies to process this knowledge base. Within this area, Semantic Web (SW) provides promising solutions nowadays with the proliferation of the semantic technologies based on Resource Description Framework (RDF) [4]. RDF provides an infrastructure to uniquely identify and merge both distributed data and metadata. RDF Schema (RDFS) and Web Ontology Language (OWL) are W3C standards for representing semantic models. While RDFS offers a simple vocabulary for describing schemas or metadata, OWL provides a richer vocabulary (on top of RDFS) with a set of pre-built formalisms for expressing logical definitions and constraints [5]. Using RDF, ontologies and controlled vocabularies have been increasingly applied in many domains within the last years, such as in Medicine, Biology, eGovernment, Web Services, Blogs, Social Web etc. This trend is becoming even more prominent as more vocabularies (RDFS vocabularies or OWL ontologies) are defined for and used by datasets in the Linked Open Data Cloud [6].

Following the advancement of SW technologies and VO frameworks, we have been motivated to explore their applicability to find a solution for the problem of cross-clouds inter-operability. The ontologies built using RDFS and OWL enable the integration of distributed data and services without assuming a single, monolithic, centrally controlled knowledge base. These also enable semantic inter-operability, data integration and meaningful search as well as progressive capturing of new insights, shared understanding and new formal structures.

In this paper we present the preliminary results of our studies on applying open market mechanisms to service discovery for engaging in collaborative scenarios. The rest of the paper is organized as follows. Section two gives details about our motivation. The next two sections present a study of the state of the art today regarding VO frameworks and semantic web technologies. Section five introduces the basic idea of an open virtual organization model and the next section discusses the research challenges in the realization of this model.

II. Motivation

Interoperability between clouds, as well as between clouds and on-premises services, is needed to serve real-world business processes of virtual enterprises. Figure 1 depicts such a process where each task requires resources/services available in different clouds. The main issue in this case refers not necessarily to the heterogeneity of services but more important to the lack of tools the end-user may easily use to assembly services available in clouds within its own business process. Similar to enterprise information silos before, the new information silos have moved in the clouds.

The complexity of the situation comes from the fact that in almost any business scenario one cloud is not enough. Figure 1 shows a business process that uses Salesforce cloud to execute core CRM processes, Google Drive to integrate documents repositories, Amazon services for inventory management, DHL services to trace the delivery of orders and finally the Authorize cloud to process payments. This makes already five clouds plus the business process management (BPM) system that have to be integrated only for one business process. Aside inter-operability issues, complex problems arise in case of the need to change one of the used clouds, i.e. the Salesforce cloud. In this case, the processes in the BPM cloud have to be rewritten – not only to change the CRM interfaces for specific activities, but potentially the flow of the process as well. The cost of switching the clouds is the main reason for the new type of vendor lock-in and consequently leads to less flexibility for the company.

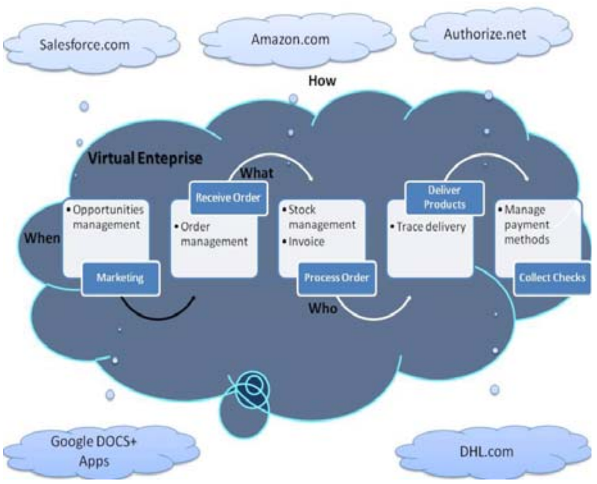

Fig. 1 Workflow example using resources and services from different clouds

The main motivation of the study presented in this paper can be summarized by the following research question: would it be possible today to create a customized configuration of cloud services in the form of a virtual organization designed to support the execution of a business process? A couple of specific questions may be derived from here: 1) what kind of architecture would be needed to create the environment where a business process uses an optimal configuration of cloud services that freely "compete" in offerings for various business process activities; 2) how to design a collaborative business process in such way to not specify the exact services to be used but rather only the semantics to constraint the services selection at run-time

We propose the flowing hypothesis: the state-of-the art in information technology today should allow the externalization of the semantics mediation between heterogeneous cloud services to an open market place where services may be discovered and organized into optimal business solutions by means of declarative queries.

To validate the hypothesis we conduct an exploratory study that takes into account a review of relevant literature, technologies and European projects on virtual organizations and semantic web technologies for open data. Following this study a new VO model is introduced and future research challenges to fulfil the proposed model are discussed.

III. Virtual Organizations

The virtual organization term has its roots back in the early 90's when Raymond Miles and Charles Snow [7] first described the agent-broker network organization (dynamic network). Later on, the idea has been transformed into a new organization design paradigm by popular works of Davidow, Hammer and Cunningham [8,9]. The authors share the same vision of an organization system with the following distinctive characteristics: vertical disaggregation, internal and external brokering, full-disclosure information systems, and market substitutes for administrative mechanisms. However, there is an important difference between VOs and networks. Networks tend to develop stable relationships based on unique regional culture and competences or business needs: a strong-coupled system of partners providing their skills based on long-term contracts as well as a common cultural background. Networks can be often seen as breeding environments for VOs [10] where the term VO refers to ad-hoc designs based on the switching principle [11]. The switching principle refers to the ability of the organization to dynamically re-allocate resources in virtual activities. A common example of applying the switching principle is the order payment process when the user or the system (by means of a set of pre-defined business rules created by the broker) may select one type of payment, from a list of available methods, based on some contextual variables. The switching principle offers an interesting perspective on virtual organizations as systems created on the basis of resource selection from the (electronic) market.

A study made by Katzy [12] shows three types of VOs: a) supply-chain VO in manufacturing industries; b) star (main contractor) topology used in construction industries; and c) peer-to-peer VO in creative and knowledge industries. In a supply-chain topology, it is the business process that is designed and governs the partners' interaction. In a star topology, partners interact with one central hub or strategic centre, while partners in peer-to-peer topology have multiple relationships between all nodes without hierarchy. Another classification shows two main types of VOs: 1) agent-based systems (individuals, agents, goals, individual and group behaviour, rules) [13], [14] and 2) service-based systems (systemic approach on relationships between objectives, events, entities, nodes, services, and the required coordination and management frameworks) [15], [16], [17].

A. *Agent-oriented virtual organization frameworks*

The research efforts regarding agent-based VO frameworks head to ensuring cooperative behaviour in scenarios populated with heterogeneous agents and led by their own interests. Castelfranchi summarizes relevant literature and identifies two main areas of research [18]: 1) imposing restrictive facilities for the actions of agents, and thus being impossible for them to deviate from the desired behavior (the approach severely limits the autonomy of agents); 2) restricting the environment, in which agents interact, through the use of business rules and leaving the freedom for the agents to follow or violate them. Usually, the first case deals with the relationship between tools for workflow management and agent-oriented systems while in the second case the concept of electronic institutions is introduced as a virtual replica of the institutions that govern the real world.

McGinnis published a framework for designing virtual organizations seen as a result of inter-connections that take place in a society of agents [19]. The authors formally describe the agents society as a function of establishing relationships between components such as agents, services, roles, workflow and contracts. The work defines the rules governing the combination of the five components that underlie the formation and implementation of virtual organizations within a service oriented architecture (SOA). They offer a systemic approach, with a high degree of abstraction, over the virtual organization seen as a system based on agents. A series of similar previous works may be used to add more details to the big picture: a) a voting protocol for the agents that make up the VO [20], b) the formal representation of contracts [21].

Taking a similar approach, a series of papers analyze the norms that may be applied to the behaviour of agents through the so-called electronic institutions (EI). An electronic institution is considered a key component in the supervision of agent-based virtual organizations. Rules are declared by the EI to govern the public behavior of agents. In this regard, Sierra [22] proposes a framework for defining and applying such rules. The authors aim to combine the Islander (a pragmatic modeling language for electronic institutions) with a methodology for the development of intelligent agents (Prometheus). Oliveira and Lopes have also developed a framework [23] that uses rules engines to apply a set of rules (the normative system) in a context called "institutional reality" (body of facts that exist into the engine's working memory at a certain moment). The agents will then always act within this kind of context. The authors identify three types of rules: constitutive rules, institutional and operational. Similar approaches that propose the use of rules as a restrictive environment for agents' behaviour can be found in [24].

B. *Service-oriented virtual organization frameworks*

There are a number of works showing that the creation of virtual organization can only occur through the integration of ontologies and semantic technologies in service-oriented systems. Thus, in [25] and [26] one can find an ontology-oriented service-based VO modeling framework addressing the inherent inter-operability problems that can arise in heterogeneous service-oriented environments. The authors present a simplified architecture that facilitates the dynamic reconfiguration of services based on requests expressed by customers. A request is sent to the system and is served by an ad-hoc organization of heterogeneous services that have been previously registered in a semantically-harmonized environment based on Semantic Web technologies.

Much of the research that focus on a top-down approach in the design of virtual organizations assume the existence of the VO breeding environment mentioned earlier in this study. Camarinha-Matos and Afsarmanesh [27] design a framework for VOs in such environments: (1) characterization of the opportunity for collaboration; (2) creation of the VO draft plan; (3) search and selection of partners; (4) negotiations; (5) detailed plan of the VO; (6) contracting; (7) launching.

IV. Semantic Web technologies to support Inter-operablity between services in different clouds

Semantic Web technologies seem to be capable today of providing the needed functionality to support the switching principle of VOs. The Internet of Services has emerged as the fulfilment of the true SOA vision: assembling distributed and heterogeneous Web Services into business processes. The first driver was a new language, WS-BPEL, which came in as a way to specify and coordinate business processes using web services. WS-BPEL initiative has its grounds in the conventional WS specifications and was superseded by a more compelling one that tries to formalize aspects like data-depended behavior, exceptional conditions and long-running interactions. The initiative tried to address the problems related to sustaining conversations between different online partners, but the conversation language, even with the machine-independent XML support to interchange data and with WS-BPEL to discover and coordinate services as business operations, is still in its infancy being too much dependent to the business actors' way of understanding, formulize and interpret the business entities and processes. To overcome such limitations, open vocabularies might be used together with standard RDF and OWL technologies. Thus, in the following sub-sections we study the sate of the art in the areas of open data, service discovery as well as contextualization and matching techniques

A. *Web of Data*

The Web of Data is about distributed data over the Internet based on Linked Data (LD) principles [28] and is the Semantic Web made the right way, according to Tim Berners Lee. Data describing things is marked with a semantic annotation at the source site level and then the description URI is used by search engines to indicate the meaning of the concept when used in other web pages. The Linked Data principles essentially dictate that every piece of data on the web should be given an HTTP URI which, when looked up, should offer useful information using standards like RDF and SPARQL. To create public ontologies on the web XML, a W3C industry standard, was widely accepted and used as a convenient information representation and exchange format. XML itself don't carry semantics, but it serves as the base syntax for the leading ontology languages. RDF (Resource Description Framework) [29] is a standard way for simple descriptions to be made. What XML is for syntax, RDF is for semantics - a clear set of rules for providing simple descriptive information. RDF Schema then provides a way for those descriptions to be combined into a single vocabulary. RDF enforces a strict notation for the representation of information, based on resources and relations between them. As referred to in its name, RDF strength is in its descriptive capabilities, but it still lacks some important features required in an ontology language which are addressed by the next layer in the Semantic Web architecture: Web

Ontology Language (OWL) [30]. OWL is a language for Web ontologies definition and instantiation. OWL enhances RDF vocabulary for describing properties and classes: relations between classes (e.g. subclasses), cardinality (e.g. "exactly one"), equality, richer typing of properties, characteristics of properties (e.g. symmetry) and instances. OWL is the W3C recommendation for ontology definition, but other standards also support similar characteristics (DAML+OIL), for instance). OWL 2 is an extension and revision of the first version of OWL (referred to hereafter as OWL 1.0). Some of the new features of OWL 2 are just syntactic (e.g. disjoint union of classes) while others offer new expressivity including keys, property chains, richer data types, data ranges, qualified cardinality restrictions, asymmetric, reflexive and disjoint properties, and enhanced annotation capabilities. In order to satisfy the computational limitations required in real world applications, the latest specification of OWL 2 introduces three new tractable profiles. These profiles are sub-languages (syntactic subsets) of OWL 2 that offer important advantages in particular application scenarios. OWL 2 EL enables polynomial time algorithms for all the standard reasoning tasks. It is particularly suitable for applications where very large ontologies are needed, and where expressive power can be traded for performance guarantees. OWL 2 QL enables conjunctive queries to be answered in LogSpace using standard relational database technology. It is particularly suitable for applications where relatively lightweight ontologies are used to organize large numbers of individuals and where it is useful or necessary to access the data directly via relational queries (e.g., SQL). OWL 2 RL enables the implementation of polynomial time reasoning algorithms using rule-extended database technologies operating directly on RDF triples. It is particularly suitable for applications where relatively lightweight ontologies are used to organize large numbers of individuals and where it is useful or necessary to operate directly on data using rule-based techniques. Rule Interchange Format [31] or SWRL [32] may be considered as alternative solutions to express rules on RDF triples.

B. Ontologies and controlled vocabularies in Linked Data clouds

Ontologies seem to explode in many domains within the last years, such as in medicine, biology, e-Government, academic publications and so on. This trend is becoming even more prominent as more vocabularies are defined for and used by datasets in the Web of Data. Linking Open Data (LOD) and Linked Open Vocabularies (LOV) (http://lov.okfn.org/dataset/lov/) are two well known projects striving to organize these vocabularies in order to make some sense to humans as well, not only to machines. Most vocabularies are usually focused on the specification of the structure of some set of objects, e.g., their characteristics and properties, or on the functionality these objects can provide, such as when semantically describing services. For instance, many of those ontologies and vocabularies are being used by datasets in the Linked Data Cloud for publishing structured data. Examples range from general purpose vocabularies, such as the Dublin Core standard (http://purl.org/dc/terms/) for publishing metadata descriptions, to more domain specific vocabularies, such as FOAF (http://xmlns.com/foaf/0.1/) for describing people, their activities and their relations to other people, or the Product Types Ontology (http://www.productontology.org/) which provides definitions for types of product or services (from Wikipedia), extending GoodRelations (http://purl.org/goodrelations/v1 - vocabulary for publishing details of products and services) and schema.org (http://schema.org/docs/full.html - specification of a set of object types and associated properties), to cross domain ontologies, such as the DBPedia ontology (http://wiki.dbpedia.org/Ontology). Similarly, some well-known ontologies, such as OWL-S (http://www.ai.sri.com/daml/services/owl-s/1.2/), WSDL-S (http://www.w3.org/Submission/WSDL-S/) and WSMO (http://www.wsmo.org/TR/d2/v1.3/), are being used to semantically annotate services on the Web (e.g., http://www.ai.sri.com/daml/services/owl-s/examples.html and http://iserve.kmi.open.ac.uk/browser.html) with descriptions as instances of that model, and which expose, share, and connect these descriptions with other existing relevant data, information, and knowledge in the LOD
(e.g, dbPedia - http://dbpedia.org/,
Freebase - http://freebase.com/ ,
Products Types Ontology
dataset - http://www.productontology.org,
BioCatalogue - http://www.biocatalogue.org/,
iServe - http://iserve.kmi.open.ac.uk/,
Geonames - http://www.geonames.org/ontology/,
Pachube - https://pachube.com/, etc.)

Under such circumstances characterized by continuous addition of new ontologies to the global Web Database, integration and relationships discovery between data items within different Linked Data source is one of the hot topics of the moment. Within this area, VoID (http://www.w3.org/TR/void/) is an RDF Schema vocabulary for expressing metadata about RDF datasets that has been adopted as a W3C standard. It is intended as a bridge between the publishers and users of RDF data and it has been designed with a special purpose: to express, among others, links between datasets (figure 1).

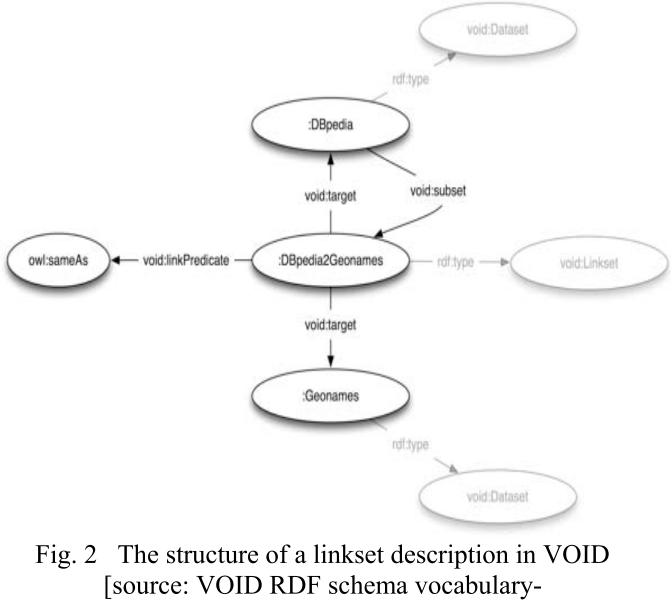

Fig. 2 The structure of a linkset description in VOID [source: VOID RDF schema vocabulary- http://www.w3.org/TR/void/]

C. Semantic annotations for service discovery

The Semantic Web Services (SWS) [33]– developed by the OASIS-Semantic Execution Environment (SEE) initiative – is a set of standards to bring semantics to SOA via web ontologies, and the most important artifacts are WSMO (Web Services Modeling Ontology), WSML (Web Services Modeling Language) and WSMX (Web Services Execution Environment). WSML defines four fundamental entities: Ontologies, Goals, Web Services and Mediators, and provide ontological specification for these entities and aims to integrate Semantic Web and Web Services technologies. WSML is a language with a specific formal syntax to define semantics of WSMO. WSMX represent an implementation for discovering, selecting, mediating and invoking Semantic Web Services. WSMX is designed as a computational software environment capable of interpreting and acting on the semantic descriptions of WSMO.

While billions of services will be available, many with similar functionality (similar semantics on function and parameter levels) [40] yet different syntax, the main challenge in the Internet of Services will be: how one can find, select/filter and use such plethora of services to run a real business process in the most optimal way. Recently, a paradigm shift has made obsolete the field of global service repositories like UDDI (Universal Description Discovery and Integration), moving research to the state of the art of today: various technologies based on RDF, ontologies, Web of Data and semantic annotations (SA) technique. Regarding the semantic annotation of resources, research efforts are heading to semantically describe services on the web. Ontologies, like OWL-S8, WSDL-S9 and WSMO10, are being used to semantically annotate Web Services with descriptions as instances of that model, and which expose, share, and connect these descriptions with other existing relevant data, information, and knowledge in the Web of Data. Another example is the hRESTS microformat (HTML for RESTful Services) [34] designed to obtain machine-readable descriptions of Web APIs described in the form of HTML service documentation for developers.

All these techniques use the SA mechanism to enhance the XML (or XHTML) description of services. However, there is one important limitation with SA: the first question before creating an annotation would be what aspect of the content must be represented. So far, the majority of WS annotation models can represent only one aspect. The authors of hREST present the same idea. They show that in order to obtain some level of automation one needs to capture four aspects of service semantics: information model (a domain ontology) represents data, especially in input and output messages; functional semantics specifies what the service does, by means of functionality classification or through preconditions and effects; behavioral semantics defines the sequencing of operation invocations when invoking the service; and non-functional descriptions represent service policies or other details specific to the implementation or running environment of a service. The authors propose an extension to hREST using WSMO-Lite [35] that defines a lightweight ontology for the four kinds of semantics, and uses SAWSDL (Semantic Annotations for WSDL and XML Schema) [36] to annotate WSDL documents with instances of that ontology. This makes WSDL-based Web services amenable to SWS automation.

D. Contextualization and matching

From the early development stages of the RDF data model there was a requirement for qualifying knowledge in order to limit its validity and credibility. As Linked Data is considered implicitly valid, with vocabularies and reasoners aimed to produce new statements and knowledge connectivity rather than to impose what is allowed and what not, reification was proposed as a mean to express belief, subjectivity, thus providing a way of making statements about statements (http://www.w3.org/TR/rdf-mt/#ReifAndCont). However, verbosity and ambiguous semantics made this design pattern unpopular, to the benefit of an alternative proposal – the N-quads, a solution that breaks the RDF model, but is supported by RDF queries. The "named graphs" supported by SPARQL can be treated as "contextual graphs", using the graph identifier as a resource with its own semantics, acting as an encompassing context for the entire graph content and its semantics, even if an ambiguous one (the context-content relationship is not explicitly named) (http://www.w3.org/TR/rdf-sparql-query/#namedGraphs). It has multiple advantages over reification and it's already supported by alternative syntaxes (N-quads (http://sw.deri.org/2008/07/n-quads/), TriG (http://www4.wiwiss.fu-berlin.de/bizer/trig/), TriX (http://www.hpl.hp.com/techreports/2004/HPL-2004-56.html).

The semantic matching problem has been a research challenge for a long time now. Several scenarios have been identified by Euzenat and Shvaiko [36] where a key requirement is a matching mediator that provides aligned views of one party to the others or the common query interface. Thus heterogeneity is conserved, but overcome. A number of matching frameworks exist both in literature as described in [37], and at tools level such as the Prompt extension to Protégé for matching and versioning various knowledge representation formalisms (http://protege.stanford.edu/plugins/prompt/prompt.html). However, one needs to establish the scope and coverage of this mediator and how communication protocols should access it. For VOs, common ontology alignment is not desirable; the preference goes to peer-to-peer matching strategies and matching-service level agreements.

E. European Projects on Cloud Inter-operability

Table 1 summarizes the main results of recent European Union (EU) funded projects dealing with inter-operability issues between different clouds. The study shows active research at the EU level in the area of semantic mark-ups and ontologies for inter-operable services and proposing business process definition annotations for service discovery. It also suggests a certain trend of EU research today: the usage of ontologies and semantic annotations to support workflow-oriented inter-operability solutions based on virtual organizations.

Table 1 EU funded ICT projects in cloud inter-operability

Project	Main result
SUPER	• Business Process Ontology Framework • (limited) Set of domain ontologies for e-business (e.g. Organizational terms) • BPEL4SWS – BPEL extension for using SWS • Mark-ups for major BPM languages to use SWS (annotations) • Tools to work with BPEL
SPIKE	• Semantically enriched service bus that enables both static and dynamic binding of concrete executable services to the representations of tasks in a predefined abstract process model. • User-centric portal that enables the definition, setup, and management of virtual organizations (designed VOs).
MOSAIC	• Cloud ontology. • Language standardization. • Agents acting on semantic descriptions. • Workflow editors.
SOA4ALL	• Semantic Web technologies to enhance service descriptions - WSMO-lite and USDL • Web based tools and editors for business processes
InteliGrid	• The semantic grid platform. • The product data grid. • The ontology services (gridspace, VO, services, resources, business process objects) • Grid enabled applications • Ontology specification for virtual organizations
VISP	• Software Platform for exchange and combine internet services using workflows. • Ontology for telecom services.
BREIN CONTRAIL	• SLAs in VOs • Dynamic SLAs
CROSSWORK	• Mechanisms for automated workflow formation and enactment in forming Networks of Automotive Excellence (NoAE) • Domain ontology developed
SHAPE	• Semantically-enabled heterogeneous service architecture (extension of SOA to include SWS, P2P, Grid, Agents) • Contribution to SoaML and Model Driven Engineering to generate components based on the proposed architecture
LOD2	• Tools to facilitate the usage of LD in e-business scenarios.

V. Towards an Open Model for Virtual Organizations

A number of conclusions can be drawn from the study presented in the previous sections. Firstly, VOs are based on the idea of dynamic allocation of resources to better serve a common goal or a new one, but the switching principle, which is the key ingredient, does not have a common technical solution yet. The latest trend in VO frameworks is heading towards the usage of ontologies and semantically harmonized environments for agents to communicate based on contracts (suggesting the need of a common language) and to follow the rules imposed by electronic governance institutions. On the other hand, EU research projects focus more on workflows than on independent agents. To design semantics based inter-operable solutions, there are other concerns such as trust, motivation, context, SLA and so on equally important to data and functional integration.

Secondly, Semantic Web has evolved to Linked Data that provides the foundation for building the Web of Data as globally distributed resources with no common repository. The Web of Data already contains vocabularies to describe primitive data as well as terms from many domains. On top of these, declarative query languages such as SPARQL have already reached certain level of maturity, now moving to the mass adoption phase. On the other hand, the new version of OWL 2 offers new features to operate directly on RDF triples using rule-based techniques.

Thirdly, the latest European research initiatives seem to agree on the idea that ontologies should be included in daily business activities and should penetrate at the very level of enterprise business processes. However, most of these projects create yet another universal language (ontology) either to define VOs or to annotate inter-operable services. The main issue with semantic annotations is that they have to be designed specifically for each service description language such as WSDL, WADL or IDL. The main issue is that there are multiple annotation technologies for one service description language as well as multiple vocabularies for semantic annotations to describe various views on the same service (functional and non-functional requirements). The service provider has to have the willingness to modify the description of the offered services to include (multiple) semantic descriptions. Even so, the issue of semantic similarities between vocabularies used in such descriptions remains to be solved.

These conclusions validate our initial hypothesis: using VOs and LOV it should be possible to externalize the semantics mediation between heterogeneous cloud services to a public virtual market place where services may be organized into optimal solutions in order to fit the specific needs of each process. Thus, we propose an open model for VO formation that should allow Linked Data-based discovery and matching of published semantic descriptions of individual workflow tasks, either required or offered (figure 3). The open virtual organization (Open VO) is the result of a query gathering an assembly of resources described in the Web of Data to fulfill the processing needs of a business process. The Open VO is a designed VOs since each VO definition is an organization of linked resources created as a result of published requirements and needed to serve workflows.

At the core of the Open VO one may find an ask/bid market model. An Ask is an RDF description of a workflow task published by the client to declare the functional as well as non-functional needs (requirements). Bids are RDF descriptions of tasks published by service providers to promote the capabilities (functional, non-functional) of their services. The semantic descriptions of activities use datasets from LOD and, as such they are based on Linked Open Vocabularies (LOV) - e.g. Product (Products Types Ontology); Owner (FOAF); Business Unit (GoodRelations); Location (dbPedia); Time (OWL-Time). This way, workflow tasks become resources in the LOD that may be queried by means of RDF-based declarative query languages such as SPARQL and matched using OWL-based technologies. The organization of task descriptions should take into account a multi-view approach (also a recommendation of ISO/IEC 42010 standard [39]). This approach allows users to specify multiple concerns to be taken into account in cross-clouds collaborative workflows (e.g Data, Function, Location, Time, Objective, Service Level Agreements).

The novelty of the approach refers to the implementation of the switching principle in the absence of a central broker or common vocabulary. Since the Open VO actually is the result of a semantic query (e.g. SPARQL), every execution of this query may result in a different configuration of Bids for the same configuration of Asks. Using LOVs, the client may freely specify all kinds of functional and non-functional requirements as long as a vocabulary exists. When the need of a vocabulary extension is identified, the extension should simply be defined and published. Interested service providers or third party service aggregators will use the new extension to publish new data about Bids.

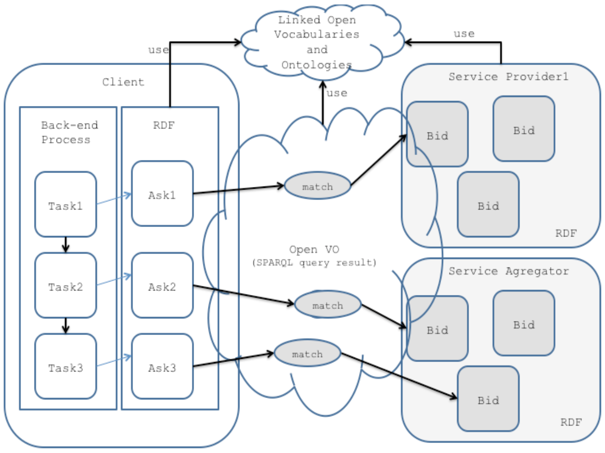

Fig. 3 The Open VO model

The task description should be a set of rules (using some formalism like previously mentioned RIF, OWL2RL or SWRL) as the most used technique today to easily express business semantics. Rules build on facts, and facts build on concepts as expressed by terms [38]. The terms used in these rules are triples published in the Web of Data. The Asks and Bids are distributed on participant's web sites, without any central repository. Queries and OWL technologies may then be used to match Asks and Bids. When participants are using different LOVs to represent the same meaning, formal representations of equivalencies are needed to bridge the semantics.

The Open VO model is built on four principles:

1. No universal business vocabulary - since the Asks and the Bids may use different vocabularies for similar terms, the market matching mechanism should use semantic matching techniques based on OWL or integration vocabularies such as VoID to accommodate vocabularies mismatches.
2. No reference architecture – the Open VO is a semantic model obtained as a result of a query on Linked Data.
3. No semantic annotation of service descriptions - instead of semantic annotations within the service description, the model proposes the externalization of the semantics. This principle induces a number of advantages: 1) semantic descriptions may be stored anywhere on the Internet, other than the home site of the service owner, creating a distributed network of wrappers each one potentially using a different knowledge attachment techniques or knowledge representation formalism; 2) the responsibility of creating the semantic description may be transferred to the service consumer, the one that actually needs the description; 3) since one single service may be reused in multiple activities using different vocabularies, the number of potential collaboration scenarios as well as the chances to enter collaboration networks increase dramatically.
4. Open market – the service clients and providers should be free to publish their requirements and offers without the need to register in any central repository.

VI. RESEARCH CHALANGES

Analysing the results of our previous study and the open VO model described earlier we identify a number of research challenges in the fields of VO frameworks, semantic techniques for services discovery and integration and markets of services.

We have seen earlier there are a number of VO frameworks that have been developed on the same building blocks as Open VO: rules and ontologies. However these mainly refer to agent-based emergent VOs where the main the use of ontologies is to address the need of a common language while rules have to constraint the behaviour of agents. Our model promotes the usage of rules and ontologies to describe the needs for services and the counterpart offerings on the open market. The open market is not a central repository but the LOD as a whole.

By removing the constraints of a fixed data model, the Open VO needs a framework for seamless construction of business interlinguas that may take into account as many specific concerns (views) as needed (e.g Data, Function, Security, Objective, Service Level Agreement, Location etc). Semantic matching components are also needed: 1) to exploit the potential of declarative query languages such as SPARQL while taking into account reification and contextualization; 3) to identify the matching Ask-Bids by matching the individual views in various LD clouds.

Regardless of the large number of research papers, prototypes and languages available, we can still note little progress in the direction of semantically enriched service technologies mass adoption. We can identify at least two types of major limitations that are still preventing the wide spread of services usage and open the gates for future research. Firstly, we consider semantic annotations of services at the level of service description language as not the most flexible technique as long as the standardization of semantic description languages is hard to achieve. The vastness of the business world simply cannot be captured by one single standard ontology and/or language. There will always be a need for alternative vocabularies and dialects specific to each industry. At the same time, the centralized management of common vocabularies needs strong commitment and computational power from users in order to evolve and to serve a large business community if not the whole world. Secondly, there are very few tools today for efficient service discovery, selection and composition in heterogeneous semantic environments. Moreover, the traditional approach regarding the semantic annotation of the service descriptions still needs the willingness and the availability of the service provider to include the needed additional information. There is also a certain level of redundancy induced by the need to describe the same semantics using different semantic technologies.

Such limitations mainly impose restrictions that can be briefly summarized as follows: 1) limited number of business community members; 2) barrier to entry for new comers; 3) information silos and duplication of data generated by world wide isolated groups of companies. Only open collaboration models may overcome these. Common understanding is not a fundamental requirement for entities operating on the same market of semantics – rather, tools for knowledge matching, service level agreement and decision support are necessary to establish a temporary common semantic communication and negotiation channel (while the partners keep holding on to their own subjective views, or even acting based on different views in different scenarios)

The Open VO proposes to decouple the semantic description from service description schema and thus allowing the separation between multiple views over the same service. This means to use specific LOVs and to further define their relationships in order to cover specific concerns needed when publishing Asks/Bids. For example, common views such as Data, Function, Security, Location, SLA etc. may be required to specify the needs/offering of one single task and these may be realized using more than one ontology. Suppose one Ask task description in our Ask-Bid model uses dbPedia while a Bid description uses GoodRelations ontology. The main question is what vocabulary will be used to create a bridge ontology such that the two descriptions to match. In this area, the VoID vocabulary seems to offer great potential, as it offers support for building ad-hoc of relationships between vocabularies. Exploiting or extending this vocabulary integration schema might be a good idea in order to add a sense of orientation for the end-user when it comes to knowledge (re-)organization by different views.

Regarding the ubiquitous usage of semantic annotations we have to admit they still have to make use of some vocabulary. In this case two services may use different vocabularies yet having the same meaning (e.g. in SAWSDL the modelReference property may point out to any model. Thus, the target model still has to be discovered by the tool since there is no indication of its type). Even if the tool will understand the language used by the target model, how would it be possible for that tool to make inferences regarding the similarities between any vocabulary using the same or any other language? Say the service specifies some QoS like price and availability and another service offers the same information only using different currency and time vocabularies. In this case an (semi-)automated mechanism is needed for the machine to understand that both services are similar in nature.

On the fields of contextualization and matching, using the named graphs supported by SPARQL the context-resource can further be described in order to explicit its semantics, then the context description itself can be further qualified by meta-contexts (context of a context) and so on, allowing for expressing subjective (or alternative) views on reality, and even further (on the meta-scale) subjective views on subjectivity provided by other sources. This approach opens still untapped potential for the development of trust and reputation models assigned to linked data sets (and even to other trust and reputation providers, allowing a collaborative approach able to express reputation of reputation providers, trust of trust providers and so on). It also generates challenges in the areas of: 1) criteria for knowledge filtering (for RDF browsers, agents, reasoners), knowledge mapping, automated service level agreement negotiations; 2) a way of relaxing ontologies by allowing the coexistence of alternative, even contradictory views on the same concepts; 3) protocol-level recommender systems based on semantic negotiations might reuse existing techniques such as the so called HTTP content negotiation, which uses header fields such as See Also to redirect between resource and knowledge repositories.

VII. CONCLUSIONS

In this paper we have explored new ways to combine sate of the art virtual organization frameworks and semantic web technologies in the pursuit of inter-operable cloud services. In this study we have taken into account relevant literature, SW technologies for open data and EU funded research initiatives having objectives related to our research. The findings have been used to formulate the idea of an open model for virtual organizations (Open-VO) as a novel approach to use the open market mechanisms and semantic web technologies using linked open vocabularies. In the end we have discussed a number of research opportunities to technically fulfil the Open VO model.

The state of the art today in the area of virtual marketplaces for pieces of software is lead by private implementations of mobile applications markets such as Android Market (the software store developed by Google and accessible from any mobile device with Android OS) and Apple App Store (the software store developed by

Apple and accessible from Apple devices). While it is clear that such examples are based on proprietary infrastructures and closed protocols and APIs the question is what kind of technology one should create to support the proper description of services such that any kind of services to be discoverable and usable in business collaborations.

The idea of linking service descriptions to datasets in the Web of Data, and removing the need of a central repository and universal business vocabularies, generates new opportunities for SPARQL based inference engines to identify equivalent services that may be used within a specific context (e.g. a certain business process instance). Moreover, there will be no need for UDDI-like repositories since the shared semantics will be published in the Web of Data just as any other RDF dataset. As a consequence, new methods for the discovery and orchestration of services are needed to use external semantic descriptions of services published as Linked Data. Decoupling semantic description from service description language also generates services reusability research opportunities by transferring the responsibility of providing an adapted description for specific needs from the service provider to any third party including the client.

REFERENCES

[1] Harding, C. Cloud Computing for Business: The Open Group Guide. Van Haren Publishing. 2011

[2] Kundra, V. Federal cloud computing strategy. cdn3.ctovision.com

[3] Cretu, L. (2006). Virtual Organizations: Beyound Dynamic Networks. Available at SSRN 920405. 2011

[4] Lassila, O., & Swick, R. R. Resource description framework (RDF) model and syntax specification. 1999

[5] McGuinness, D. L., & Van Harmelen, F. OWL web ontology language overview. W3C recommendation, 10(2004-03), 10. 2004

[6] Bizer, C., Heath, T., & Berners-Lee, T. Linked data-the story so far.International Journal on Semantic Web and Information Systems (IJSWIS),5(3), 1-22. 2009

[7] Miles, R. E., & Snow, C. C. Organizations: new concepts for new forms. Firms, Organizations and Contracts, Oxford, 429-441. 1996

[8] Davidow, W. H., & Malone, M. S. The virtual corporation: Structuring and revitalizing the corporation for the 21st century. HarperCollinsPublishers. 1992

[9] Hammer, M., & Champy, J. Reengineering the Corporation: Manifesto for Business Revolution, A. HarperCollins. 2009

[10] Camarinha-Matos, L. M., & Afsarmanesh, H. Creation of virtual organizations in a breeding environment. In Proceedings of INCOM (Vol. 6, pp. 583-592). 2006

[11] Mowshowitz, A. The switching principle in virtual organization.Electronic Journal of Organizational Virtualness, 1(1), 6-18. 1999

[12] Katzy, B., Zhang, C., & Löh, H. Reference models for virtual organisations. In Virtual Organizations (pp. 45-58). Springer US. 2005.

[13] Patel, J., Teacy, W. L., Jennings, N. R., Luck, M., Chalmers, S., Oren, N., ... & Thompson, S. Agent-based virtual organisations for the grid. Multiagent and Grid Systems, 1(4), 237-249. 2005

[14] Frąckowiak, G., Ganzha, M., Paprzycki, M., Szymczak, M., Han, Y. S., & Park, M. W. Adaptability in an Agent-Based Virtual Organization–Towards Implementation. In Web Information Systems and Technologies (pp. 27-39). Springer Berlin Heidelberg. 2009.

[15] Cretu, L. G. Cloud-based Virtual Organization Engineering. Informatica Economică, 16(1), 98-109. 2012

[16] Stoica, M., & Ghilic-Micu, B. Virtual organization-cybernetic economic system. Modeling partner selection process. Economic Computation and Economic Cybernetics Studies and Research, 43(2), 73-83. 2009.

[17] Grefen, P., Eshuis, R., Mehandjiev, N., Kouvas, G., & Weichhart, G. Internet-based support for process-oriented instant virtual enterprises. Internet Computing, IEEE, 13(6), 65-73. 2009

[18] Castelfranchi, C. Engineering social order. In Engineering societies in the agents world (pp. 1-18). Springer Berlin Heidelberg. 2000

[19] McGinnis, J., Stathis, K., & Toni, F. A formal framework of virtual organisations as agent societies. arXiv preprint arXiv:1001.4405. 2010

[20] Pitt, J., Kamara, L., Sergot, M., & Artikis, A. Formalization of a voting protocol for virtual organizations. In Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems (pp. 373-380). ACM. 2006

[21] Udupi, Y. B., & Singh, M. P. Contract enactment in virtual organizations: A commitment-based approach. In AAAI (Vol. 6, pp. 722-727). 2006

[22] Sierra, C., Thangarajah, J., Padgham, L., & Winikoff, M. Designing institutional multi-agent systems. In Agent-Oriented Software Engineering VII (pp. 84-103). Springer Berlin Heidelberg. 2007

[23] Cardoso, H. L., & Oliveira, E. Electronic institutions for B2B: dynamic normative environments. Artificial Intelligence and Law, 16(1), 107-128. 2008

[24] Vázquez-Salceda, J., Aldewereld, H., & Dignum, F. Implementing norms in multiagent systems. In Multiagent system technologies (pp. 313-327). Springer Berlin Heidelberg. 2004.

[25] Val, E. D., Criado, N., Carrascosa, C., Julián, V., Rebollo, M., Argente, E., & Botti, V. THOMAS: a service-oriented framework for virtual organizations. In Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1-Volume 1 (pp. 1631-1632). International Foundation for Autonomous Agents and Multiagent Systems. 2010.

[26] Zhou, L., Chen, H., & Mao, Y. A semantic-based framework for virtual organization management. In ChinaGrid Annual Conference, 2008. ChinaGrid'08. The Third (pp. 294-299). IEEE. 2008.

[27] Camarinha-Matos, L. M., & Afsarmanesh, H. Creation of virtual organizations in a breeding environment. In Proceedings of INCOM (Vol. 6, pp. 583-592). 2006

[28] Bizer, C., Heath, T., & Berners-Lee, T. Linked data-the story so far. International Journal on Semantic Web and Information Systems (IJSWIS), 5(3), 1-22. 2009

[29] Lassila, O., & Swick, R. R. Resource Description Framework (RDF) model and syntax specification. 1998

[30] Beckett, D., & McBride, B. RDF/XML syntax specification (revised). W3C recommendation, 10. 2004.

[31] Kifer, M. Rule interchange format: The framework. In Web reasoning and rule systems (pp. 1-11). Springer Berlin Heidelberg. 2008

[32] Horrocks, I., Patel-Schneider, P. F., Boley, H., Tabet, S., Grosof, B., & Dean, M. SWRL: A semantic web rule language combining OWL and RuleML. W3C Member submission, 21, 79. 2004

[33] McIlraith, S. A., Son, T. C., & Zeng, H. Semantic web services. Intelligent Systems, IEEE, 16(2), 46-53. 2001.

[34] Kopecky, J., Gomadam, K., & Vitvar, T. hrests: An html microformat for describing restful web services. In Web Intelligence and Intelligent Agent Technology, 2008. WI-IAT'08. IEEE/WIC/ACM International Conference on (Vol. 1, pp. 619-625). IEEE. 2008

[35] Booth, D., Haas, H., McCabe, F., Newcomer, E., Champion, M., Ferris, C., & Orchard, D. Web Services Architecture, W3C Working Group Note, 11 February 2004. World Wide Web Consortium, article available from: http://www. w3. org/TR/ws-arch. 2004.

[36] Euzenat, J.: Eight questions about Semantic Web annotations. In Intelligent Systems, vol 17, issue 2, IEEE Computer Society, Washington, DC. 2005

[37] Jiang, B., & Luo, Z. A New Algorithm for Semantic Web Service Matching. Journal of Software, 8(2), 351-356. 2013

[38] Ross, R. Principles of the Business Rules Approach. Addison-Wesley information technology series. ADDISON WESLEY Publishing Company Incorporated. 2003

[39] ISO. ISO/IEC 42010:2007 Systems and software engineering - Recommended practice for architectural description of software-intensive systems, http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=45991 2007

[40] Strîmbei, Cătălin. "Smart Data Web Services." Informatica Economică 16.4 (2012).