LOTED2: an Ontology of European Public Procurement Notices

Isabella Distinto [a,b,*], Mathieu d'Aquin [c], Enrico Motta [c]

[a] *Engineering Tributi S.p.A., Via Giovanni Battista Trener 8, 38121, Trento, Italy*
[b] *ISTC-CNR Laboratory for Applied Ontology, Via alla Cascata 56/C, Povo, 38123, Trento, Italy*
E-mail: isabella.distinto@loa.istc.cnr.it
[c] *Knowledge Media Institute, The Open University, Walton Hall, Milton Keynes, MK7 6AA, United Kingdom*
E-mail: {mathieu.daquin,enrico.motta}@open.ac.uk

Abstract. This paper describes the construction of the LOTED2 ontology for the representation of European public procurement notices. LOTED2 follows initiatives around the creation of linked data-compliant representations of information regarding tender notices in Europe, but focusing on placing such representations within their legal context. It is therefore considered a legal ontology, as it supports the identification of legal concepts and more generally, legal reasoning. Unlike many other legal ontologies however, LOTED2 is designed to support the creation of Semantic Web applications. The methodology applied for building LOTED2 therefore seeks to find a compromise between the accurate representation of legal concepts and the usability of the ontology as a knowledge model for Semantic Web applications, while creating connections to other relevant ontologies in the domain.

Keywords: Legal ontologies, European public procurements, Linked Open Data, Open Government Data, Tenders

1. Introduction and motivation

Over the last twenty years, AI & Law research has widely investigated the field of legal ontology. Up to now, efforts have focused on the development of legal ontologies, aimed at the representation of fundamental basic concepts of legal knowledge and at the description of specific legal domains, to support legal case-based reasoning, decision support systems, legal compliance checking, and more generally to map the complexity of legal knowledge to formal languages such as OWL. Specific methodologies and approaches used for building these ontologies show that legal ontology engineering has its own peculiarities [11], requiring particular conceptual structures for the representation of legal concepts, including notions of legal rules, preconditions, legal consequences, etc. [38,41].

At the same time, the Linked Data trend and the emergence of second generation Semantic Web applications have highlighted the significance of intelligence arising from the integration of dispersed and heterogeneous data from many sources, rather than from closed knowledge based systems [14]. Thus for this kind of applications, there is a need for flexible and modular ontologies, which can be easily integrated to discover non trivial connections between data.

Since law impacts and affects the everyday life of all individuals, representing legal knowledge in the Semantic Web scenario is both a timely need and a challenge. To see this, one has only to consider the huge amount of Open Government Data released at present. Indeed, the e-Government sector is one of the major

---

drivers in the emergence of Open Linked Data and governments are making accessible a large number of datasets, about a wide range of topics, such as spending reports, administrative staff organizations, public healthcare, etc. However, without the right interpretation, data are only raw data, which by themselves may not provide useful information.

Open Government Data are in many cases related to the legal domain and, as a result, legal ontologies may play a key role uncovering the semantics of these data and driving the integration of this information with other datasets. Thus, it would be possible to build semantic e-Government applications, which may provide a significant contribution in bridging the gap between citizens and institutions [25].

Nevertheless, as recently emphasized by some scholars, the shift towards the Linked Data framework and the new paradigm of Semantic Web applications implies a reconsideration in the legal ontologies' role and utility, "questioning the need for a highly axiomatised and unified knowledge representation" and conceiving instead "a new way of designing legal ontologies and of embedding them into architectures for legal information systems and other web services" [10].

As far as we know, at present time no explicit attempts have been made to figure out alternative methodologies for building legal ontologies aimed at supporting web services and, in particular, Semantic Web applications. It is still not clear how we may or should rethink the approach to legal ontologies so as to meet current and growing demand for better access to the legal information by disclosing the semantics of the increasing amount of legal (and legal-related) data. Moreover, we do not know to what extent it is possible for the legal ontology engineer to come to terms with the new challenges, meeting the need of Linked Data and Semantic Web advanced scenarios, while at the same time, preserving the richness of contents, the constraints and the other peculiar features of the legal domain. Furthermore, we are still not able to predict what a change of mind in legal ontologies' role, design, context and use may entail in terms of both benefits and risks.

All things considered, in an attempt to clarify at least part of these uncertainties, in this paper we explore a novel approach to the design of legal ontologies, having in mind their use in the Linked Data framework and specifically in Semantic Web applications as means for delivering better access to the legal information enclosed in Open Government Data. The approach we advocate here represents a sort of compromise between an accurate representation of the legal knowledge and the requirements demanded by the Web of Data, since we try to represent legal concepts in such a way as to drive the construction of Linked Data applications (or lightweight Semantic Web applications). Besides, and against the backdrop of the challenges arising from the Legal Semantic Web applications' design, the work inspires and motivates a discussion on the peculiarities of legal ontologies: their relationship with the Semantic Web and their ability to be integrated with (even) non-legal ontologies in the Linked Data framework.

The application scenario for the investigation of this approach is the European public contracts domain. The public procurement domain is a complex and very technical legal field but also a strategic one, since it accounts for 17% of the EU Gross Domestic Product (GDP). As such, it represents one of the major public finance levers: a significant boost to the European countries economies depends precisely on this public expenditure. Because of its significance, public procurement concerns many stakeholders including: *communities of citizens*, who may be interested in monitoring the management of the *res publica* (i.e. public affair) and of the operating expenses used on behalf of the public interest; *traders*, for whom it represents interesting opportunities, although sometimes not easy to grasp, especially for small and medium enterprises (SMEs); and, of course, *public administrations*, which need to use procurement to deliver infrastructures, public services and goods to citizens.

Data on public contract notices are open data by their nature; by law they must be accessible, because the whole public procurement process should be based on transparency and the advertising of contract notices is an essential step of the public procurement transparent procedure. Indeed, public administrations must aim to get 'value for money' (public money, taxpayers' money) by choosing, through transparent procedures, the best offer among the largest possible number of candidates. Therefore, any contract notice is issued with the intention to reach the largest number of potential candidates, because the more economic operators take part in the tender, the more competition increases. Naturally, as the competition increases, the chance to get better products and services increases as well.

Today, public institutions at all levels (regional, national or supranational) publish their procurement notices on the Web. In the vast majority of cases, the advertisement on the institutional web-sites is an essential condition, required by law, for the conduct of tenders. In the European Union, tender notices for the

award of public contracts 'above the EU threshold'[1] are published on the TED (*Tenders Electronic Daily*) website[2], which is the online version of the 'Supplement to the Official Journal of the European Union' dedicated to European public procurement. The TED system, as such, represents the main source of information concerning European procurement notices for every business entity across the European Union: it is continuously updated with procurement notices issued by several public institutions, authorities, public bodies, etc. from all the European and the EEA countries (about 1500 procurement notices per week); it is maintained by an European Institution (the European Commission) which guarantees the reliability of information, ensuring also the availability of 'data' 24/24h. At the same time, the procurement notices published at the European level provide, indirectly, also much information concerning the public institutions, authorities, etc. that have issued the contract notice and even more (e.g. where contracting authorities are located, in which sector they carry out their main activity, how many EU public contracts they award in a year, how much money a regional authority, a council, or even a country, spend on education, health, environment, etc.). In other words, the TED portal represents one of the most impressive sources of *Open Government Data*, which may wield their potential in the Linked Data scenario and, at the same time, can be used to test new design conception of legal ontologies to drive Legal Semantic Web applications.

In the recognition of these potentialities, we have developed an OWL 2 DL ontology of European public procurement notices, whose main purpose is to serve as a means for ensuring a better access to procurement information. Though information about procurement notices is nowadays largely available on the Web (and specifically on the TED web site with regards to the European procurements), we believe that in many cases it is not fully understandable (and therefore usable) by non-experts or public procurement newbies, notably by small and medium enterprises. Accordingly, the ontology we have built has been conceived to offer – paraphrasing Dietz's words – a new understanding of procurements "such that one would be able to look through the distracting and confusing appearance [of them] right into their deep kernel" [19]. At the same time, we decided not to hide the legal complex-

ity of the domain, by reducing the ontological commitments. Rather we tried, accepting a sort of compromise, to preserve the legal concepts' representation behind the "little semantics" necessary to engage a larger audience than legal experts. Such an expressive modeling of the domain allows the discovery of connections with other domains, e.g., business domains, and the integration with other relevant ontologies, specifically with Good Relations [30].

We call this new ontology LOTED2, since it can be considered as an evolution of LOTED[3] 'Linked Open Tenders Electronic Daily' [45], a project that pioneered the use of Linked Data to enrich the data about public procurement notices contained in RSS feeds of the TED system. Compared to the original LOTED, LOTED2 does not only provide an RDF vocabulary for representing the data exposed through the TED website, but specifically focuses on modeling, in an ontology, the legal context of these data, supporting a more complete, more accurate and better interpretable view over these data for applications to exploit.

The rest of this paper is structured as follows: Section 2 outlines related work on reusing Linked Open Data associated with the European public procurement notices and on other existing ontologies on public contracts. In Section 3 we describe the methodology we used to build the LOTED2 ontology, discussing its aims and its design. In Section 4 we describe the ontology in detail, its modularity structure and the modules by which it is composed; in Section 5 we discuss the integration of LOTED2 with the most widely used ontology for describing e-commerce scenarios, namely Good Relations, while in Section 6 we provide some examples of instantiation of the ontology with TED data. Finally, in Section 7 we reiterate the main results from our work on LOTED2, emphasizing the lessons learned and highlighting further research questions.

2. Related work

2.1. LOTED ontology

The LOTED ontology[4] was developed inside the eponymous project, in order to introduce an additional level of structure on top of the data extracted from the RSS feeds of the TED system. It has been conceived for the need of the platform and structured to

---

[1]For the current values of the thresholds, see http://www.ojec.com/Threshholds.aspx.

[2]http://ted.europa.eu/

[3]http://loted.eu

[4]http://loted.eu/ontology

enrich procurement notices data with automatically discovered links to Geo-Names and DBpedia. So, it is a lightweight ontology, realized to match the semi-structured representation (namely the tabular summary of data) of tender notices as published by the TED website.

Although the LOTED ontology satisfies the requirement of usability and is very suited to the purpose for which it has been designed, it does not actually represent knowledge about the domain, as it merely defines the structure of data objects in the domain. However, authors were aware of this aspect and a more expressive representation of the domain of procurement was planned as part of future work on the LOTED project. The LOTED2 ontology has been developed to achieve this goal.

2.2. The Public Contracts Ontology (PCO) developed under the LOD2 project[5]

LOD2[6] is a large-scale Integrated Project co-funded by the European Commission within the FP7 Information and Communication Technologies Work Program, with the overall aim to creating knowledge out of interlinked data and to develop tools and methodologies for exposing and managing very large amounts of structured information on the Data Web and to test and bootstrap a network of high-quality domains, also based on multi-lingual ontologies, from sources such as Wikipedia and OpenStreetMap. A work package of this project, the WP9A ("LOD2 for a Distributed Marketplace for Public Sector Contracts") is dedicated to exploring and demonstrating how the application of linked data principles for procuring contracts in the public sector may help to bridge the gap between advanced countries and countries with low online participation of enterprises in public tenders. So, the main purpose of this work package is to build a linked data

infrastructure in order to produce a "business impact and achieve an effective resource allocation through emulating the market process of meeting supply and demand"[7].

At the heart of this infrastructure there is the Public Contracts Ontology (PCO). The authors state they are not interested in modeling every aspect related to a contract, but only "information which is available in existing systems on the Web" and "which will be usable for matching public contracts with potential suppliers" [18]. In other words, the goal of this ontology is to model a public contract as a whole, but without going into details of the domain.

PCO is more articulated than LOTED: it is not built to model the data structures of a particular system (TED), but rather tries to represent a variety of aspects of the domain, taking into account the integration with other ontologies (Good Relations, VCard[8], Payments Ontology[9], Call for Anything[10] and also LOTED). It therefore provides a broader vision of the domain compared to LOTED: some relevant aspects of the domain, such as lots, are represented in this ontology.

Nevertheless, PCO shows some weaknesses that cannot be ignored, even while keeping in mind that it is not a legal ontology. In particular, the ontology aims to provide a conceptual description of both the tendering phase and the phase of the execution of the contract. However, there is not a clear conceptual distinction between these different scenarios in the PCO. Indeed, in the ontology the different concepts of call (call for tenders, such as contract notice) and contract (public contract) are often confused.

First of all, the class `pc:Contract` is modeled as subclass of the "Call for Anything" class `c4n:Call`. In some sense, a contract notice (not a contract) can be considered as a 'call', a call through which it is announced a competitive bidding for the award of a public procurement contract. In other words, a contract notice is a call-for-tenders, which may be submitted for the award of a public procurement contract. Thus, the contract is 'the stakes' of the competitive bidding an-

---

[5]Note that the version of the Public Contract Ontology we are referring to was available at the time of the first submission of this article (3/09/2012) at http://purl.org/procurement/public-contracts#. The Deliverable 9a.1.1 ("Framework for Creating Linked Data in the Domain of Public Sector Contracts" – available at http://static.lod2.eu/Deliverables/deliverable-9a.1.1.pdf) describes in details this earlier version of the ontology. However, a new version of PCO has been released recently (see http://opendata.cz/public-contracts-ontology), which substantially differs from the previous one. In particular, the authors of the ontology seem to have addressed many of the points we raise in this section.

[6]http://lod2.eu

[7]http://lod2.eu/WorkPackage/wp9a.html

[8]http://www.w3.org/TR/vcard-rdf/

[9]http://data.gov.uk/resources/payments

[10]Call for Anything, which is available at http://vocab.deri.ie/c4n is a vocabulary to describe calls in general, including call for tenders. As a generic vocabulary, it is intended to be extended by more domain-specific vocabularies. Briefly, the model relates a `c4n:Call` issued by a `foaf:Agent` to a `c4n:Event`. This event must meet certain conditions, such as a deadline and a spatial scope.

nounced through a notice (i.e. a call), not the call itself. Secondly, the Contract so modeled in PCO has both an estimated and actual price, at the same time. When the notice is published the price is only an estimated one. The final price will be set only at the end of the competitive bidding, on the basis of the award criterion (for an example the lowest price) and of the tender bids submitted. So, it would be appropriate to keep strictly separate the different phases of contract notice publication from the tendering phase, from the award phase of the contract and then from its execution.

Conceptual confusion increases by defining the Class `pc:Contract` as equivalent to the Class `loted:Tender`[11]. The declared aim of the ontology is to match awarding authorities' demand and traders' offering; in our vision, this aim cannot be achieved simply by declaring that a (proposed) contract (demand) is equivalent to the tender bid (offer). It is not correct, from a conceptual point of view, as well as potentially harmful. In PCO, there is not a specific definition of the class of awarding entities; they are just identified as business entities. From the point of view of the market, this is true and also useful in order to achieve integration with the Good Relations ontology. However, forgetting the ontological definition of this important aspect of the domain entails also forgetting that the procurement domain concerns the 'Public Agencies to Business' scenario and not the 'Business to Business' one.

Summarizing, this ontology tackles complex subjects through a too simplistic approach.

2.3. The 10ders Information Services Project

'10ders Information Services'[12] is a Project co-financed by the Spanish Ministry of Industry, Commerce and Tourism and by the European Regional Development Found. The aim of this project is to exploit information about public procurement notices using Semantic Web technologies and Linked Open Data approach, in order to build many services, especially targeted to SMEs. The set of services produced by the research group is currently available on a proprietary platform, Euroalert.net[13] [35], a brand owned by Gateway SCS.

Basically, the commercial services offered to small and medium enterprises span from tenders alerting systems on the basis of the subscriber profile model, to the offer of reports on the major public buyers for their products and services. It also provides a customized data mining analysis of public procurement tailored to the interest of each client.

Aside from the commendable purpose to provide an unique access point ('a pan-European platform') both for EU relevant tender notices published on TED, and for notices below EU threshold published on a wide range of buyer profiles of national, regional and local levels, it should be considered that there are many services providing mail alerts about tenders and other mechanisms of this kind. However, the novel approach of this project is that the platform system is built using structured open data instead of screen-scraping techniques.

From the same research academic group (WESO) comes also the MOLDEAS ('Methods on Linked Data for E-procurement Applying Semantics') work [2]. This project can be considered as a broader framework, which includes also the experience gained in developing Euroalert.net. Indeed, MOLDEAS aims to apply the Semantic Web and Linked Open Data approaches to public procurement notices, defining a set of goals. The first goal is to transform government controlled vocabularies such as CPV[14], CPC[15] and Eurovoc[16] (now available in SKOS[17]) into RDF, SKOS or OWL. The second one is to enrich and model information inside public procurement notices with these controlled vocabularies but also with geographical information available in the Linked Data cloud. Then, procurements information is published in a SPARQL endpoint providing a node for the linked data cloud and enhanced services (search and sort, matchmaking, geo-reasoning, statistics, etc.) on data.

The main advantages claimed by authors of Euroalert and MOLDEAS are essentially the decrease of information dispersion (arising from the different sources where contract notice above and under EU threshold are published); the unification of data models and formats and the support to multilingual issues (through EUROVOC resources used for the en-

---

[11]Within PCO this term denotes the tender bid submitted by the economic operator, that in the same ontology is called "Supplier".

[12]http://rd.10ders.net

[13]http://euroalert.net

[14]http://simap.europa.eu/codes-and-nomenclatures/codes-cpv/codescpv_en.htm

[15]http://simap.europa.eu/codes-and-nomenclatures/cpc/index_en.htm

[16]http://eurovoc.europa.eu/

[17]http://www.w3.org/2004/02/skos

richment of public procurement notices). An ontology about public contracts is still in development, as part of WESO group's future work.

2.4. On Legal Ontologies and Linked Data Models

Looking at existing work on ontologies of public procurement domain, there is no mistaking the fact that these artifacts are suitable to be employed in Semantic Web applications. However, as mentioned above, these ontologies do not really take into account the legal aspect of the domain. This is not a fault in itself, provided the domain is described in accordance with the conceptualization dictated by the legal sources.

It is also interesting to note that none of the considered work shows a reference to legal ontologies. Given that public procurement is a legal domain, it might be worth asking why legal ontologies and the Linked Data world are so distant. A brief survey on legal ontologies and an analysis of the peculiarities related to the legal knowledge representation may help to clarify the background.

The larger group of legal ontologies developed up to now, can be classified as two diverse categories.

The first is represented by core legal ontologies, namely ontologies that draw largely on legal theory in order to extract general patterns for describing the main concepts of legal knowledge: agents, roles, norms, normative effects (such as obligations, permissions and prohibitions) and temporal properties (e.g., about normative rules validity, efficacy). Examples of this type of ontologies are: the LRI-Core ontology[8], DOLCE+CLO (Core Legal Ontology) [23], the Ontology of Fundamental Concepts [40] and LKIF-Core [31].

The second group of legal ontologies is represented by specific legal domain ontologies, i.e., conceptualizations of a particular field of legal knowledge drawn from authoritative sources such as national or European laws (e.g., concerning privacy, intellectual property, copyright, etc.). Examples are the IPRONTO ontology [17], the Copyright Ontology [24], the Customer Complaints Ontology [34], the Consumer Protection Ontology [44].

The aim of core legal ontologies is to provide an interoperability framework from which other legal ontologies could inherit general categories, and in that way trying to overcome also comparative and multilingualism issues, among diverse legal systems. On the other hand, legal domain ontologies are built with a specific application-perspective in mind, usually grounded in closed systems. As a consequence, they are focused on the reasoning potentialities more than on the integration with other ontological resources in the open, heterogeneous and large scale Semantic Web environment.

The experience gained so far shows that to deal with the complexity of legal knowledge through ontologies and within the limitations of the Semantic Web languages, requires an intensive design process. Legal ontologies are usually made of a significant amount of axioms and restrictions upon classes, which represent legal concepts, even because of rigid conceptual constraints featuring the legal domain. In fact, the conceptualisation of the legal domain depends on legal sources, which dictate the semantics of legal categories through - more or less precise and explicit - definitions and descriptions, in that way constraining the interpretations (or, at least, part of them).

In contrast, the domains described by the ontologies widely used by the Linked Data community, like for example FOAF[18] or Good Relations, are known in advance. As such, the models representing those domains do not require a strong ontological commitment aimed at explaining and, rather, at restricting the terms' semantics, by avoiding non-intended (or *non-legally-intended*) meanings. So, they are typically based on propositional logic, whereas legal ontologies need at least first order logic, even though some authors have invoked the use of propositional logic for representing legal texts, such as [1], in order to avoid syntactical ambiguities. There are at least two reasons why powerful representation languages are needed for legal ontologies.

First of all, legal knowledge cannot be bridled into propositional structures (atomic propositions and propositional connectives): the truth preservation in law does not make sense, because categories of true or false cannot be applied to normative propositions. Secondly only a small part of legal knowledge may be expressed through propositional logic, which is not adequate to support legal reasoning [39]. The reason is that the first step of legal reasoning is to identify a legal concept, and the utility of ontological legal concept representation is precisely to provide a way to classify individuals.

Given this background, we have the impression that there is 'no room' for legal ontologies in the Linked Data scenario, where lightweight ontologies

---

[18]http://xmlns.com/foaf/spec/

like FOAF or Good Relations seem to rule the scene. Because of their characteristics, legal ontologies look set to dedicated and closed systems, to perform complex tasks, but not of integrating themselves into the Web of Data. Nevertheless, some features of legal ontologies are not necessarily incompatible with Linked Data requirements.

On the one hand, the representation of legal concepts and relations among them is useful to verify and fix the correctness of Linked Data statements on legal content, through the use of reasoners such as Hermit [42] or Pellet [43]. On the other hand, the new triples generated, i.e. inferred or materialized, can be written back into the RDF model, adding a new level of granularity in the data. This type of information is just what is useful to drive the integration of ontologies with other related ontologies and datasets.

There are still grounds, however, for asking whether the structural diversity among legal ontologies and the common Linked Data models represents an obstacle to the realization of legal Semantic Web applications and, when needed, to the integration of legal ontologies with other kinds of semantic resources. If so, it is not trivial to figure out how the shift towards a new way of conceiving legal ontologies may be attained. The work we are going to describe here, namely the LOTED2 ontology, can indirectly be considered as an 'experiment' carried out having in mind these questions, and thus, trying to give them an answer.

3. Intentions, aims and design of LOTED2

As said in the introduction, the public procurement domain is a peculiar one: it is a legal domain, i.e. a domain regulated by norms, but it is also a public finance domain, since it represents the means by which public agencies seek and acquire goods and services on the market, and then, one of the way in which public money is spent. In fact, when we speak about public procurement, we are also referring to the commerce transactions between public agencies and businesses (*PA2B*). As such, public procurement is an ideal field to test not only the construction of Legal Semantic Web applications aimed at supporting the meet between public demand and businesses offering, but also the intersection between legal ontologies and other related ontologies. One of these related ontologies is certainly represented by Good Relations, a lightweight ontology for annotating offerings as well as the searching for goods and services on the Web. Because of its

spread and popularity, Good Relations can be also considered as an exemplar ontology of the Linked Data trend.

Therefore, the original intention that has inspired the LOTED2 ontology's design has been that of reusing both the most widely used core legal ontology (at the state of the art of legal ontological engineering), namely LKIF-Core, and the most widely used ontology for the e-commerce (at the state of the art of Semantic Web applications and the Linked Data experience), namely Good Relations. This with the purpose of producing an ontology of public procurement notices consistent with the upper legal categories – inherited by a core legal ontology – and with the e-commerce Linked Data vocabulary.

Nonetheless, to the best of our ability and effort, a full integration of the two ontologies was not possible. Therefore, the attempt to capitalize on the most representative domain-related ontological resources, has led us to set the design and the content of LOTED2, by agreeing some kind of compromise between the ontology in development and the two conceptual models that we have taken into account as our points of reference.

Although LOTED2 does not inherit, as legal domain ontology, its upper classes from LKIF-Core ontology, it explicitly refers[19] to this core ontology as a source of patterns for modeling the legal foundations of the procurement domain, especially regarding legal sources and the classification of (legal) agents. On the other side and contrary to our reference legal ontology, LOTED2 describes extensively the domain through the RDF model, while limiting the use of OWL axiomatizations for the definition of the public procurement's most relevant legal categories. Such a design has been employed with the aim of facilitating the LOTED2 integration with Good Relations and other non-legal ontologies, and its usage as part of a Semantic Web application. In that way, we tried to reach a trade-off between a full legal 'conceptual coverage' and a 'pragmatic sustainability' [13].

To do this, the ontological modeling activity has been conducted by the domain expert and relies heavily on a comprehensive analysis and interpretation of the legal sources governing the domain. That is why we call LOTED2 a legal ontology.

In the following, we outline in details the aims of the ontology along with the methodology followed for its construction.

---

[19]See annotations to the ontology.

3.1. Aims of LOTED2

LOTED2 has been designed for the following specific purposes:

– to express the (main) legal concepts of the domain of public contract notices as defined in legal sources (European Directives on public contracts);
– to support rich semantic annotation, indexing, search and retrieval of tenders documents, such as contract notices;
– to make possible the reuse of semi-structured data extracted from the TED system;
– to enable the integration with other ontologies and vocabularies about related domains.

The ultimate goal of the ontology is to permit the construction of Legal Semantic Web applications that support public procurement by matching demand and supply.

3.2. Methodology used for building LOTED2: a backward path in re-constructing the procurement data semantics

Two main Directives cover the European public contracts domain: the Directive 2004/18/EC[20] and the Directive 2004/17/EC[21]. The first regulates the coordination of procedures for the award of public works contracts, public supply contracts and public service contracts by contracting authorities (i.e. authorities operating in the so-called 'ordinary sectors'); the second one regulates the procurement procedures of entities operating in the water, energy, transport and postal services sectors (i.e. 'utilities sectors').

These two legal sources represent the reference point to derive the exact meaning of terms used to describe the procurement domain, and to extract the information needed to build any logical theory which would formalize the domain knowledge. Besides this fundamental consideration (albeit an obvious one), another aspect concerning the European tender notices published on the TED system must be taken into account. According to the Directives, the tender notices sent by contracting authorities/entities to the European Commission shall be formulated in accordance to standard forms[22]. Since the Directives prescribe which information must be included in each kind of notice, the standard forms approved by the EU Commission are the result of a standardization process of all the procurement notices, according to the requirements imposed by the Directives. The TED system provides a full version of each tender document in the original language, and also a compact view in the language selected by the user. From the notices available in these formats, semi-structured data can be extracted in the form of a tabular summary, as shown in Figure 1.

TI	Title	UK-Milton Keynes: cleaning services
ND	Document number	382532-2011
PD	Publication date	08/12/2011
OJ	OJ S	236
TW	Place	MILTON KEYNES
AU	Authority name	The Open University
OL	Original language	EN
HD	Heading	Member states - Service contract - Contract notice - Restricted procedure
CY	Country	UK
AA	Type of authority	6 - Body governed by public law
DS	Document sent	05/12/2011
DD	Deadline for the request of documents	16/01/2012
DT	Deadline	16/01/2012
NC	Contract	4 - Service contract
PR	Procedure	2 - Restricted procedure
TD	Document	3 - Contract notice
RP	Regulation	5 - European Union, with participation of GPA countries
TY	Type of bid	1 - Global tender
AC	Award criteria	2 - The most economic tender
PC	CPV code	90910000 - Cleaning services 90911000 - Accommodation, building and window cleaning services 90919200 - Office cleaning services
OC	Original CPV code	90910000 - Cleaning services 90911000 - Accommodation, building and window cleaning services 90919200 - Office cleaning services
RC	NUTS code	UKJ12
IA	Internet address (URL)	www.open.ac.uk
DI	Directive	Classical Directive (2004/18/EC)

Fig. 1. Tenders Electronic Daily tabular summary of Contract notice n. 382532-2011

When there are standard forms for drafting a certain type of document, these can be used as starting point for the analysis of each individual document drafted according to the established format. In a certain sense, model forms are descriptions of classes of documents and, each new document drafted, is an instance of one of these classes. The analysis of these model forms may tell us information about the terms identi-

---

[20]http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX: 32004L0018:en:HTML
[21]http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX: 32004L0017:en:HTML

[22]These forms, required by Commission Directive 2001/78/EC of 13 September 2001 on the use of standard forms in the publication of public contract notices, are available at http://simap.europa.eu/buyer/forms-standard/index_en.htm

I.2) Type of the contracting authority

○ Ministry or any other national or federal authority, including their regional or local sub-divisions	○ Body governed by public law
○ National or federal agency/office	○ European institution/agency or international organisation
○ Regional or local authority	○ Other: *(please specify)*
○ Regional or local agency/office	

Fig. 2. Excerpt of a standard form for contract notices

fying parts of the document and the interrelationships among each other. In many cases those terms identify ontological classes and the structural relationships between terms can be considered as object properties or data-type properties. Thus, filling out a new document means the creation of new instances for each class.

On the other side, we can derive the exact meaning of terms contained in standard forms only through the analysis of the authoritative sources: the sources of law, indeed, contain the definitions of the relevant terms pertaining to the domain they cover (explicit knowledge), and at the same time, they provide us with information about the nature of legal concepts: what are their features (or properties) and their relations with other concepts, in the whole context of the domain (implicit knowledge). The first is a plain literal analysis, the latter, instead, is the result of a deeper domain-expert's interpretation.

As extensively discussed (among others) in [11], various approaches and methodologies have been adopted in legal ontology engineering: in particular, ontologies which rely on the interpretation of the sources of law usually are built following a top-down approach; constrastingly, ontologies which rely on textual analysis of legal texts are built following a bottom-up approach.

It is well known, even by the legal ontological engineering community [20], that both approaches show points of strengths, as well as some weaknesses. Legal ontologies built following a top-down approach, may be reused across different application scenarios, because they provide conceptual frameworks generic enough for being reused or specialized by other ontologies. However, legal ontologies built following such an approach may be not adequately linked to textual sources or to real data structure, since they usually contain many theoretical (or legal foundational) definitions.

Instead, following a bottom-up approach (e.g. by extracting a vocabulary from standard forms and analyz-

ing only the syntactic relations among terms), it is possible to discover useful terminological information at a larger scale and in an easier way. However, most of the times, following a bottom-up approach, the obtained result is too detailed, typically described by means of propositional logic. As a consequence, it becomes extremely difficult to discover fine-grained abstractions concerning broader legal concepts and, at the same time, the risk of inconsistency increases. Moreover, the bottom-up approach does not help in discovering and, then, representing more complex legal conceptual structures, which are necessary to emulate the legal reasoning.

Taking into account all these premises, the design approach of LOTED2 is based both on a top-down approach (extraction of legal concepts from legal sources) and on a bottom-up one (analysis of standard forms). We believe that, given the existence of different kinds of textual sources (standard-forms and legal sources) in the domain of interest, in this way it is possible to achieve a better match between language and conceptualization, and then, a good level of correspondence between terms, which identify data, and ontological classes.

- Firstly: the raw data structure, as shown in the tabular summaries, is *re-placed* in its context. More precisely, data are matched with their corresponding entries in the (standard-forms) notices, from which they are extracted.
- Secondly: the entries contained in the standardized notices are interpreted according to the articles of the Directives, where they are described in detail (explicit knowledge), and also, through a comprehensive analysis of the domain, as it emerges from the interpretation of the whole legal sources covering the domain (implicit knowledge).
- Thirdly: the resulting re-composed "puzzle" is formalized as much as possible through the language used for coding the ontology.

Consider for example, the excerpt of the tabular summary related to the contract notice n. 382532-2011 (see Figure 1) the data 'AA', labeled 'Type of Authority' (which in this specific case is a 'body governed by public law'), is extracted from the entries of the standard form as shown in Figure 2.

Then, an analysis of the Directives is conducted in order to search definitions or other kind of information pertaining to the "Type of Authority" record (or ontological class), especially by analyzing *why* this kind of information is relevant and why it must be included in contract notices. In this specific case, a definition of types of authorities empowered to issue contract notices is contained in the the Article 1 no. 9 of the Directive 2004/18/CE. This specification is relevant because these types of authorities, by issuing a contract notice, play the role of 'contracting authorities'. If one relies only on the tabular summary or on the standard form analysis, it is not clear that 'type of authority' and 'contracting authority' are different types of ontological categories, the latter being a role.

In a certain sense, such a mixed approach facilitates the *re-construction* of the procurement data semantics, by following a sort of *backward path* in discovering, through the help of their normative reference texts, what they really mean.

This path (analysis of data ⇒ analysis of standard-forms ⇒ analysis of legal sources) is documented inside the LOTED2-core ontology, through the annotation properties. In fact, a comment (`rdfs:comment`) is attached to the majority of classes, object and data properties, in order to indicate the legislative reference which has been taken into account for the purposes of the ontological analysis and, specifically, for the definition of that class, object property and data property.

However, the semi-structured data rendered by the TED system represent only the essential part of all the information contained in these documents. Other relevant information is lost. This is the case, for example, when a proposal for the acquisition of similar or related supplies/works/services may result in contracts being awarded at the same time in the form of separate lots. Since in some cases, economic operators may submit an offer for a single lot only (in those cases in which it is admitted a partial type of bid, for example) this type of information may be very helpful for participation in tenders of SMEs. An XML standard for legal documents, which allows also RDFa assertions in order to link the structural part of the text with ontological classes, such as AkomaNtoso [3,46] or CEN Met-

alex [6], may be used in order to tag the full content of tender documents.

4. LOTED2-core ontology and its modules

In this Section, we describe the legal ontology of European Public Procurement notices we have produced, namely LOTED2-core. The integration of LOTED2-core with the Good Relations ontology is called LOTED2-extended and is described in Section 5.

In order to facilitate maintenance, LOTED2 has been designed with a modular approach: specifically, ten modules (Figure 3) compose the LOTED2-core ontology. Modules have been conceived to be 'self-contained', 'independent' and 'reusable' [15].

An extended version of LOTED2 ontology is composed by LOTED2-core ontology, GoodRelations4Tenders and VCard.

GoodRelations4Tenders is a version of the GoodRelations ontology compliant with the public procurement domain, while, as well known, VCard describes a mapping of the VCard specification to RDF/OWL. The goal of the integration of LOTED2-core ontology with VCard is to promote the use of VCard for the description the description of 'points of contacts' and addresses of awarding authorities. GoodRelations4Tenders, instead, represents the *key link* between demand and supply side of public procurement.

LOTED2-core is a framework module (as well as LOTED2-extended module). Its function is to hold together the modules by which is composed LOTED2-core ontology. Both the LOTED2-core ontology and its extended version (called LOTED2-extended) are available   at http://loted.eu/ontology2   and   at https://code.google.com/p/loted2/source/browse/.

In the following, we provide an overview of the modules composing LOTED2-core and a description of the main inferences that the ontology supports.

4.1. Procurements Subjective Scope

The Procurements Subjective Scope module describes the classes of legal persons who are empowered to issue a tender notice (generally called 'call for tenders') and to award a public procurement contract, i.e. to play the role of awarding legal entities. An enumeration of these entities is contained in art. 1 (9) Directive 2004/18/EC and in art. 2 of Directive 2004/17/EC.
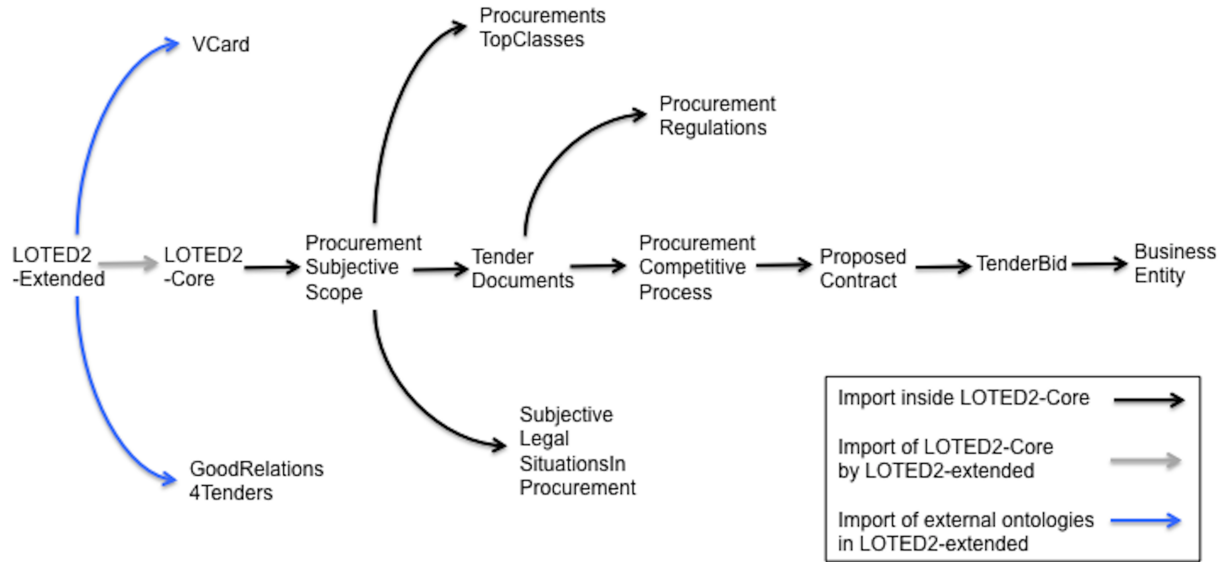
Fig. 3. Dependencies between LOTED2-core modules and external ontologies (GoodRelations4Tenders and VCard) in LOTED2-extended

The first type of entities may play the role of contracting authorities; the second type of entities may play the role of contracting entities. Although both the Directives refer to 'contracting authorities' and 'contracting entities' as two different types of entities, these terms denote, from an ontological point of view, roles rather than types of entities.

Consider for example the case of a body governed by public law, that, on the one hand, issues a notice for the award of a public contract and, one the other, may submit a tender bid in a tender announced by another entity[23]. Furthermore, any public authority or public body cannot be considered a contracting authority or entity per se. It assumes this feature only when it issues a contract notice or another type of call for tenders. In other words, this is an anti-rigid property [27], since every instance of a public authority, body governed by public law, ministry, etc. is not essentially a contracting authority or entity.

However, for the purposes of the European legislator, the distinction between types of entities, which may play the role of contracting authority, and entities, which may play the role of contracting entities, is very relevant. From this distinction, in fact, depends the ap-

plication of Directive 2004/18/EC (on ordinary sectors) or of Directive 2004/17/EC (on utilities sectors). Hence, LOTED2 includes two classes that refer to intermediate legal concepts, namely 'entity operating in ordinary sectors' and 'entity operating in utilities sectors'. If an entity falls in the class of the first type of entity, then it may play the role of contracting authority. Instead, if an entity falls in the second type, then it may play the role of contracting entity. The definition of these classes is based on the main activity carried out by an entity. A detailed list of ordinary sector activities and utilities activities is obtained by standard forms, but should not be considered as a *numerus clausus* (i.e. not limited in their number and content). By means of an annotation property (`Loted2:tedLabel`) each type of these activities and each type of legal person who issues a notice is annotated with the corresponding TED data's label. This annotation property is provided with the purpose to drive the user in the implementation of the ontology by using the TED data as instances.

4.2. Tender Documents

The Procurement Subjective Scope module is connected to the Tender Document module through the object property `Loted2:issues`. This module describes the majority of tender documents available on the TED system, issued by the entities defined in the Directives. The aim of this module is to provide a full description of tender documents, which represent no-

---

[23]The forth whereas of Directive 2004/18/EC and the eleventh of Directive 2004/17/CE, in this regard, state that Member States should ensure that the participation of a body governed by public law as a tenderer in a procedure for the award of a public contract does not cause any distortion of competition in relation to private tenderers.
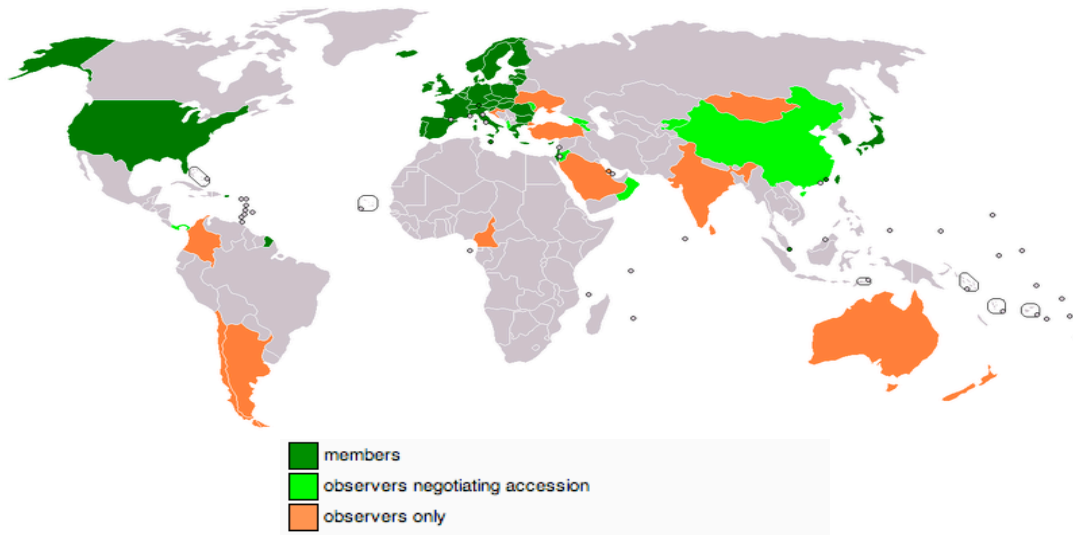
Fig. 4. WTO Members and observers in the Government Procurement Agreement (source: Wikipedia)

tices. Other types of tender documents such as 'specification' or 'descriptive tender documents' are described because they are strictly related to tender notices. These documents are attached to a contract notice in order to describe in details the type of service, good or works and the manner in which the competitive bidding is conducted.

Data contained in the tabular summaries of TED are basically related to tender notices. So, this module, more than others, has been built emphasizing the bottom-up approach. As in the Procurement Subjective Scope module, also in the Tender Document module the annotation property `Loted2:tedLabel` is used. Another annotation property, `Loted2:tedDataID` is used for providing a reference to the ID of the data to which a data-type property refers.

4.3. Procurement Regulation

This module describes the legislative sources that regulate the public procurement domain. Apart from the Directives, many other legislative sources regulate the European procurements domain. The most important is the Government Procurement Agreement (GPA), a pluryilateral treaty signed by a number of WTO (World Trade Organization) parties (Figure 4), with the purpose to open up as much as possible public procurement business to international competition [7,16]. The scope and coverage of GPA is based on the type of procurement, the type of entity and the monetary threshold defined in Appendix I of the Agreement.

It is worthwhile to emphasize that signatories may negotiate the coverage of GPA with other parties, on the basis of reciprocity. For example, the utilities sector is not covered by GPA with respect to Canada and there are many limitations with respect to USA and Japan too, just to name the most relevant. This means that we cannot consider the GPA as a unique treaty for each signatory, since there are rather many bilateral agreements (Canada-EU, USA-EU, et.) negotiated by parties under the GPA framework. Thus, simply saying that a tender is within the coverage of the GPA is not enough to clarify the geographical scope of the application of each single bilateral agreement to the single contract notice. These significant divergences in the application of GPA entail a considerable complexity in defining which specific regulation covers a tender document.

Unfortunately TED system data do not provide an effective help in this direction since they are articulated in an incoherent manner. In fact the field of 'Regulation' data is referred once to the political geographical area of the country in which the entity that issued the tender notice is based (European Union, European Economic Area); once to the type of authority that issued the notice (European Investment Bank, European Bank for Reconstruction and Development, European Monetary Institute, European Institution/Agency or International Organization); once to the reason for which the notice is issued (External aid and European Development Fund) and finally once to the actual regulation, although not always identified with the exact

wording (Agreement between the European Community and the Swiss Confederation, GPA, etc.).

The Procurement Regulation module of LOTED2-core ontology aims to provide the right interpretation of TED 'Regulation data', i.e. which legal source covers the single tender document. This type of information is necessary to define the jurisdiction of the regulation, namely the geo-political reference within the norm is applied and its effects are binding [9,26]. An additional module covering this aspect will be part of our future work.

4.4. Procurement Competitive Process

A notice is issued by an entity acting as contracting authority or contracting entity in order to announce a competition. There are many types of competitions, based on the type of notice. The most relevant type of competition is the tender (i.e. the competitive bidding) that is announced through a contract notice (but not only). Instead, a design contest is announced through a design contest notice. A qualification system is announced through a notice on the existence of this type of system, that we can consider as a competitive procurement process. This system, in fact, is used by only entity operating in utilities, for seeking qualified economic operators, which meet predefined qualification criteria that must be satisfied by potential providers of specific types of works, services and supplies. An applicant of this type of notice, which satisfies these criteria, is registered in the system as potential candidate or contractor for the particular type of contract. A notice on the existence of a qualification system with call for competition is a notice through which is announced both a qualification system and a tender, in which may participate only operators recognized as qualified according to the system.

A procurement competitive process takes place with an established administrative procedure, the award procedure that can be of different types (open, restricted, negotiated, etc.). LOTED2-core ontology does not describe award procedures.

4.5. Subjective Legal Situations

This module describes roles played by agents in procurements competitive processes and in organizations. As highlighted in the description of the Procurements Subjective Scope, terms such as Contracting Authority or Contracting Entity denote roles rather than types of entities. In particular they denote roles than only cer-
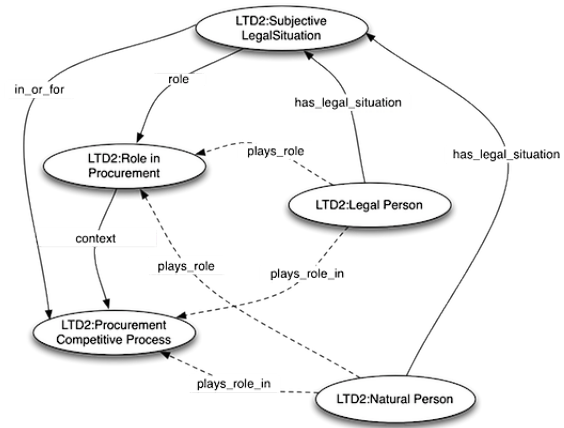

Fig. 5. Subjective Legal Situations and Roles in LOTED2-core

tain entities may play: only entities operating in ordinary sectors may play the role of contracting authority while only entities operating in utilities sectors may play the role of contracting entities. The two terms indicate basically the same concepts, namely a property that an entity assumes when awards a public contract and when carries out all the set of actions required for the awarding process of a contract. The first action is to issue a notice.[24]

So, by issuing a contract notice certainly an entity starts to play the role of awarding legal entity. Apart from the role of awarding legal entity, this module of LOTED2 enables us to describe also roles played by 'business entities' in the competitive processes of procurements, and in organizations. For example, a natural or legal person (operating on the market as 'economic operator') who has submitted a tender bid for the award of a proposed public contract is a natural or legal person who plays the role of 'tenderer'. And if this agent has submitted the best tender bid, then assumes the role of 'successful tenderer'. Another important aspect in procurements' role modeling deserves to be emphasized. Just as a body governed by public law may play the role of both awarding legal entity and tenderer, so business entities may play many roles in different procurements competitive processes.

---

[24]Please note that in some specific cases public authorities may award a contract without issuing a notice: this is the case in which a tender takes place with a negotiated procedure without contract notice. Of course, this case is not examined in this paper, since LOTED2 is an ontology for describing tenders notices published on TED.

Consider that the eligible customers of many business entities are public authorities and so their core activity is precisely the participation in tenders, even more than one at the same time. Then, this matter is not trivial. In order to represent also these cases, the described module represents (legal) situations more than simply roles. An agent may have more than one subjective legal situation that is related to a role played in a context. In the case of procurement the context is the procurement competitive bidding. These concepts and relationships have been modeled reusing and adapting the 'Social Reality pattern' [32,33], one of the proposed content patterns available from the catalogue of Ontology Design Patterns initiative [21,22].

4.6. Proposed Contract

Since LOTED2 is an ontology of public procurement notices, it does not represent public contracts per se. In fact, LOTED2 aims to describe the semantics of notices concerning the award of public contracts and not that of public contracts in themselves. A contract notice is the means whereby a competitive bidding for the award of a public contract is announced. So, the commitment of the ontology is to capture the information of public contract to be awarded (or proposed contract) not of the public contract awarded or in its execution. For this reason, the ontology specifies the class of `Loted2:ProposedContract` rather than of Contract.

Consider the case in which the tender has been declared unsuccessful: can we speak about a contract or not? Of course not, because the contract has not been awarded and then has not been signed by parties. So, in the stage of notice publication there is not a contract, but a contract to be awarded. This is also clear if we consider that a contract notice is also known in legal doctrine as *invitatio ad offerendum*, namely an invitation to make an offer for a proposal of contract. And only the successful bidder will be party of the contract. The connection between the contract notice (the *invitatio ad offerendum*) and the proposed public procurement contract module is via a property chain:

```
Loted2:throughWhichIsAnnounced o
Loted2:forAwardOf ⊑
Loted2:throughWhichIsAnnouncedTenderForAward.
```

Every contract or proposed contract has an object, namely the subject matter of the contract. A law-full object is an essential of a contract or proposed contract. According to legal doctrine, object of contract can be intended either as the commitment that parties agree to assume (and the consequently transfer or creation of rights and/or modification or settlement of existing bonds) or as the description of the real object (good or service) to which relates the contract (i.e. the substantial content of the contract) [5].

The analysis of standard forms for contract notices and of the other tender documents shows that the object of contract is intended in the second sense, i.e. as a description of the type of good, service or work that the entity issuing the notice seeks. A rather significant aspect of procurement domain is the division of proposed public contract in lots. In some cases and under certain conditions, entities issuing a contract notice for the acquisition of similar or related supplies/works/services may decide to split the proposal into separate single proposed contracts to be awarded with the same tender. These single proposals that are also parts of a general proposal are called lots.

Since in some cases, economic operators may submit an offer for a single lot only (namely in all the cases in which is admitted a partial type of bid) this type of information may be very helpful for participation in tenders of SMEs. In fact, usually a small enterprise is specialized in one particular sector with one main offering and so it might be interested in bidding for the single lot rather than for the global proposal. Therefore, one of the commitments of LOTED2 ontology is to accurately model this particular aspect of the domain, namely lots, even by ensuring that through ontology were made possible certain inferences.

Lots are proposals of contract as well as proposal of contracts not divided into lots, but they are different in the sense that they depend - in the sense of existential dependence - by the general proposal which include them. At the same time, the general proposal is also a particular one, since it is ontologically dependent on its parts (lots, indeed); otherwise it would be empty with respect to the object of the contract. In other words, there is an ontological mutual dependence between them. Furthermore, the lot and the general proposal share the same nature: they are both proposal of contract. On the other hand, lots are also different and independent from each other, because they have different subject matter and they can be awarded to different tenderers. Therefore, it is not easy to understand if their being part of another proposal should be understood in the sense of *composition* or in the sense of *constitution*. So, being aware of this subtle distinction, we have decided to represent it into the ontology by introducing different values (named individuals) of 'type of proposal' (see in the ontology the class `ProposedContractType`),

namely `Lot`, `ProposedContractNotDividedIntoLots` and `ProposedContractDividedIntoLots`. Only the first one, namely a proposed contract of type lot, can be part of the latter (in the ontology see the definition of domain and range to the object property `isDividedInto` and its inverse `partOf`).

When a proposed contract is divided into lots through the contract notice is announced a tender for award also single lots. Through LOTED2 it is possible to infer that through the same contract notice is announced a tender for award each single lot. This result has been achieved through a set of General Class Axioms combined with property chains.

Two object properties are inferred in order to show which lots will be awarded through the same tender
(`Loted2: forAwardOf o Loted2:isDividedInto`
`⊑ Loted2:forAwardLot`)
and which lots to be awarded are announced through the same contract notice
(`Loted2:throughWhichIsAnnouncedTenderForAward`
`o Loted2:isDividedInto ⊑`
`Loted2:throughWhichIsAnnouncedTenderForAwardLot`).
An example in the published version of LOTED2 ontology shows the inference so described (see individual `Loted2:Notice1`).

4.7. Tender Bid

This module describes the tender bid, namely the offer that may be submitted by the economic operator in the competitive bidding for awarding a public contract. Note that, in english, the same word 'tender' denotes the 'race', i.e. the 'competition' for the public contract, and also the meaning of the offer for a public contract. It is an ambiguous term and, arguably, this ambiguity has led the PCO developers to confuse the 'race' with the 'offer'. For this reason, we have decided to call the offer *tender bid*, because it is a bid made in the context of a tender. So, the tender bid class aims to describe the offers submitted by economic operators in a tender for the award of public contracts.

An important aspect is that tender bids may be of three different types: a tender bid, indeed, may be either a *partial tender bid*, a *global or partial tender bid*, or a global tender bid. In those cases, respectively, an offer may be submitted for exactly one lot, for one or more than one lot, or necessarily for all lots. This feature denotes a sort of quality that a tender bid has in a single tender. Furthermore, a tender bid has only one of these quality in each tender. The type of tender bid is, then, strictly related to the lots' issue, because both

a partial tender bid type and a global or partial type imply the existence of a proposed contract divided into lots; instead the global tender bid type may be admitted both in tenders through which are awarded contracts not divided into lots and in tenders through which are awarded contracts divided into lots.

A tender bid is evaluated on the basis of an award criterion: the lowest price or the most economically advantageous offer. The first is based only on the criterion of price; the second one is based on a set of combined criteria defined by the authority issuing the contract notice. Furthermore a tender bid will be opened in a certain place (defined into the contract notice) and at a certain date-time; and it may be drawn up in a specific EU language or not.

4.8. Business Entity

This module describes the class of the entities playing the role of economic operators (to simplify hereinafter called 'economic operators'). Economic operators are the subjects to whom the invitation to submit an offer for a proposed public contract is addressed. In other words, an economic operator is the potential counterpart of the awarded contract.

However, not every economic operator can sign a public contract. There are several eligibility requirements, based on certain criteria that must be fulfilled by an economic operator in order to participate in a competitive bidding. Another module describing these requirements will be part of our future works. This module describes private legal persons who are business entities, since they assume the legal form of 'incorporates', 'society', 'cooperative', etc. These classes are modeled following the taxonomy of LKIF-core ontology.

4.9. Top

This is an upper module in which are contained abstract classes, even useful to match the LOTED2 ontology with core ontologies and in particular with core legal ontologies, in order to foster interoperability. In particular, many classes of the 'Procurement Top Classes' module are modeled following the LKIF-core ontology schema. This relationship between LOTED2-core ontology and LKIF-core is what we call *a compromise* accepted in designing LOTED2-core ontology. Indeed, as mentioned above, the initial aim of LOTED2 project was to build an ontology of European public procurement notices integrated with both

GoodRelations and LKIF-core ontology. However, during the development of the ontology we had to take some decisions about how to integrate the two ontologies, which use not only different patterns of knowledge representation, but also different modalities of language's use. This, for instance, is the case of the the time representation.

Given the importance of time factor in legal domain, LKIF-core represents time (such as 'date of publication') as classes, while Good Relations represents time through data type properties such as `xsd:dateTime`. As far as we know, for instance, through OWL constructs there is no way to infer that the date of publication of the contract notice, which according to the LKIF ontology is a Class (namely `PublicationDate`), coincides with the beginning of the validity of the `gr:Offering`, which, according to the Good Relations model, is represented as a dataType property (`validFrom`). Therefore, we had to make a choice between the two options, and taking into account the main goal of the ontology, namely the construction of a Semantic Web application aimed at improving the demand and offer in public procurement, we thought that the approach of Good Relations was more suitable for our purposes. Nevertheless, this is just our choice and nothing prevents to take the opposite decision, but in principle it would be not advisable to choose both options at the same time, otherwise the result would be a useless duplication.

Furthermore, we believe that the LKIF-core ontology has been mostly conceived for representing legislative documents and not also administrative documents, such as tender notices. In fact, every legal document is also a legal source and there is not a clear distinction between legislative sources (which are sources of law in the proper sense) and legal sources, such as a contract (or a proposal of contract), which are sources of law only *inter partes*, i.e. only among the contractual parties. We agree on the fact that even a contract, a proposed contract, a contract notice and in general an administrative document (i.e. a legal document) are, in a broad sense, sources of law, but this representation does not help us in distinguishing clearly between tender documents and the legal sources that regulate them. Given the need in our domain to make explicit the differences between tender documents and norms about EU public procurement, we have introduced the class of Legislative Document, which comprises EU Directives, and the class of Administrative Document, which comprises tender documents.

On the other hand, in GoodRelations, a business entity is a "*legal agent* making a particular offering" and it can be "*a legal body* or a person". This natural language statement is translated in the GoodRelations ontology by representing Organization and Person as SubClasses of BusinessEntity. This ontological representation is obviously instrumental to the needs of e-commerce Semantic Web applications. However, it implies also that *all* the persons and *all* the organizations are also intrinsically *business entities*, and it is obvious that, from the point of view of the legal domain, this ontological representation is not acceptable, neither for LKIF-core, nor for LOTED2-core, because a legal body in the legal domain is not essentially a business entity. Certainly, a legal body can act also as a business entity, but there is something more to it than its buying activity. First: a business entity in the legal domain is only that particular type of organization which is recognized by the law as a business entity, according to well defined criteria. Second: it is hard to accept the idea that, for example, a Parliament, a Ministry or an European Institution are primarily business entities. It is true that they buy products and services in the market, but their buying activity is not the main activity they carry out; it is only an instrumental activity, functional to the achievement of their institutional goals. Third: in general and also from the legal knowledge point of view, natural persons are not primarily business entities.

For all these reasons, LOTED2-extended maintains a sort of 'conceptual independence' from both LKIF-core and Good Relations.

5. LOTED2-extended: the integration of LOTED2-core with GoodRelations4Tenders

Works carried out until now [2,18] bring up the integration of ontologies about public procurements with Good Relations ontology. In particular, as we have seen in Section 2, LOD2 ontology reuses some classes of Good Relations (Offering, Business Entity, etc.) and also WESO Research group pays attention to an integration with Good Relations, as part of its future works. Even the creator of Good Relations, Martin Hepp, often makes reference to public procurement as an interesting application domain for his ontology [29]. It is generally agreed that the object property `gr:seeks` could play a key role in order to promote the opening up of procurements domain to large scale Se-

mantic Web applications. However, the matter, in our opinion, deserves more attention for many reasons.

Good Relations is, currently, the best candidate ontology to complement an ontology of public procurement notices because it represents a strategic domain, such as the e-commerce in a non trivial way, addressing a complex domain and covering "the many intricate situation that this domain requires", "is widely used currently in the e-commerce and linked data communities", "it is easily applicable, actually applied and recommended by the stakeholders from the targeted domain" [13].

This is certainly true and in addition to that, we think that is also an ontology that inspires the recognition of some isomorphic patterns between legal and economic concepts, namely between the market and its legal superstructure. Nevertheless, it should be remembered that Good Relations is an ontology built with the purpose to meet the needs of B2C (Business to Consumer) or B2B (Business to Business) scenarios. In other words, Good Relations has been developed bearing in mind only the private sector. Therefore, it is not fully adaptable to the PA2B (Public Agencies to Business) scenario, because the public procurement domain has its own peculiarities that cannot be ignored, if one wants to build a Semantic Web that does not distort the principles of EC law.

Consider the principle of equality of treatment on which the Art. 23 of Directive 2004/18/EC is inspired. This article states that, unless justified by the subject-matter of the contract, technical specifications shall not refer to a specific make or source, or a particular process, or to trade marks, patents, types or a specific origin or production with the effect of favoring or eliminating certain undertakings or certain products. Such reference shall be permitted only on an exceptional basis, where a sufficiently precise and intelligible description of the object of the contract is not possible. In these specific cases, such reference shall be accompanied by the words "or equivalent". Since Good Relations, according to its purposes, describes brands, types of products (such as models), and obviously also the origin of products, the integration of an ontology about public procurements like LOTED2 with this ontology could result in potentially law distorting effects.

Indeed, these aspects come overtly into conflict with EC procurement principle of equality of treatment. Moreover, there are also other aspects whose incompatibility with the principles of law is less clear, even though equally potentially harmful.

For example, Good Relations defines the object property 'image=depiction', through which it is possible to link a product to its image available on the web. This is a very useful class in the private market scenario. In the PA2B context, instead, the use of this class may create problems. Indeed, with an extensive interpretation of the Article 23 one may argue that a link to a certain image could be used in order to indicate exactly one specific type of product. This can be a sort of trap that may lead into a conflict with the principles of Procurement Regulations.

Given these remarks, we have amended the Good Relations ontology, removing all parts non compliant with EC procurement principles. We have called this version 'Good Relations 4 Tenders'. Apart from EC Directives-non compliant classes like `gr:Brand`, `gr:ProductAndServiceModel`, etc. and related properties like `gr:hasBrand`, `gr:hasMakeAndModel`, `gr:hasManufacturer`, etc., in GoodRelations4Tenders ontology we have removed also other classes and properties not required by the procurements domain such as `gr:OpeningHoursSpecification`, `gr:acceptedPayment Methods`, etc..

Changes in this release compared to the original Good Relations ontology consist of just the elimination of certain classes and properties. The original structure of Good Relations has been preserved.

Another point on which is worth dwelling is about how an integration between an ontology of public contracts and Good Relations should be understood. For example, Public Contracts Ontology (PCO) developed by LOD2 group adheres completely to the conceptual model of Good Relations, detrimental to the particular (and different) domain of public contracts. Instead the aim of LOTED2 ontology is to represent as closely as possible legal concepts pertaining procurements domain. Such a conceptual analysis allows the discovering of the possible connections with concepts of other domains.

So, by explicitly specifying concepts of proposed contract and of *invitatio ad offerendum* (namely call for tenders or contract notice, etc.), the connection between Good Relations ontology and an ontology of procurement becomes clear; so clear that this connection can be inferred. Consider the use of the property `gr:seeks` in the public procurement domain. When an entity issues a contract notice through which announces a tender for the award of a proposed contract, actually is seeking the object of the contract. So the object of the contract is the Offering that this entity invites to submit. In LOTED2-extended this aspect is
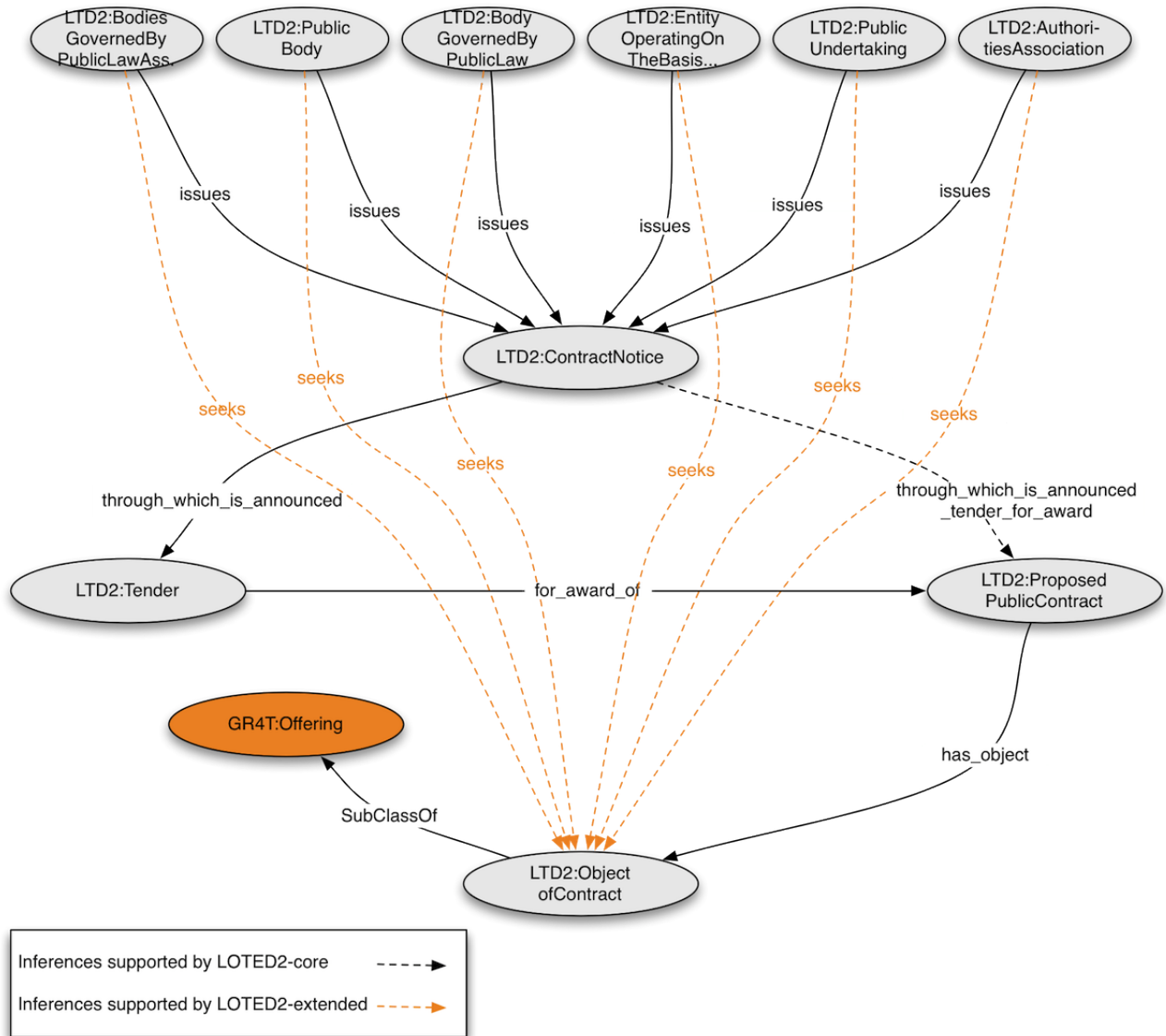
Fig. 6. Inferences supported through the integration between LOTED2-core and GoodRelations4Tenders in LOTED2-extended

automatically inferred using the reasoner, through a property chain (as shown in Figure 6).

This information is an useful one for matching public demand side with the offerings side of market. Recently, an initiative has started to make accessible information about corporations as Open Data: the 'Open Database of the Corporate World' (*OpenCorporates*)[25]. This database contains data about more than 44 millions of companies around the world. Many use-

ful information are associated to each registered company, like the jurisdiction, the kind of legal form of the company, the current status (active, inactive, in-liquidation, etc.), and also the type of product or service that a company offers. Exactly this type of information can be very useful for our purposes, since we could use it to instantiate the LOTED2-extended ontology, in that way helping the business entities to be informed about the EU public procurement opportunities available in their field of activity. In other words, linking TED data with the OpenCorporates data should al-
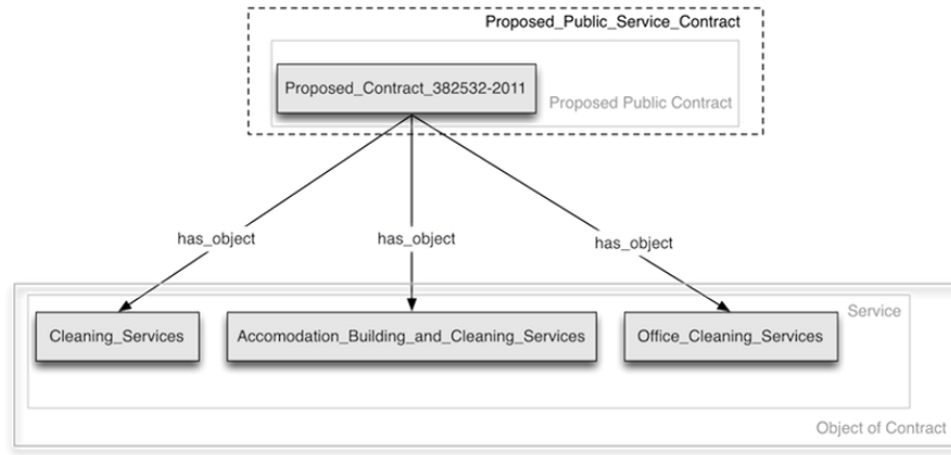
---

[25]http://OpenCorporates.com/

Fig. 7. Inferred classification of the proposed public contract n. 382532-2011 under the category of proposed public service contract

low the matching between demand and supply side, by alerting every company, which offers the same type of product or service sought by an entity issuing a public procurement contract notice.

We assume, indeed, that there are many economic operators (especially small and medium enterprises), which, looking at the world of public procurement, are skeptical or discouraged. They simply do not understand from where and how they can enter into this market, because, for instance, they are not aware of the existence of the TED system or because they do not have confidence with it. Therefore, they are not informed enough about the business opportunities coming from the public sector. To be informed in a simple way or just encouraged in participating in the European tenders, they need to know which public agency is searching for the specific product or service they sell.

In our opinion, this is the *deep kernel* of public procurement notices.

Nevertheless, the LOTED2 ontology helps in accessing also other information, as we show in the following by instantiating the model with some examples of TED data.

6. An instantiation of tender notices with LOTED2

In this section, we give some elements of the instantiation of TED data with respect to the ontology that we have just presented, highlighting in particular the inferences that LOTED2 supports. The instances used as examples are extracted from the semi-structured data contained in the tabular summaries and also from

the full text of two real contract notices published (according to the standard forms) on the TED system. As said in the section 3.2, the information contained in the full text of the tender documents can be marked-up by using a Legal XML standard, or just extracted by adopting scraping techniques.
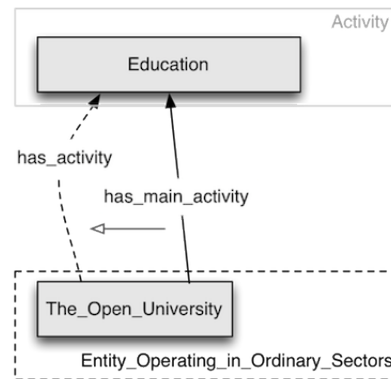

Fig. 8. Inferred classification of The Open University under the category of Entity Operating in Ordinary Sectors

As said, entities are classified on the basis of the main activity they carry out in the two diverse categories of entities operating in the utilities and entities operating in the ordinary sectors. The LOTED2 core ontology lists a set of instances of activities, which are included in the class of ordinary sectors activities and in the utilities. According to the contract notice n. 382532-2011, the main activity of The Open University is Education. Since Education is an Ordinary Activity, the ontology correctly infers that The Open Uni-

versity is an Entity Operating in Ordinary Sectors. The inferences of this case are represented in the Figure 8.

The dashed rectangle represents the inferred classification of the instance, which is indeed an Entity Operating in Ordinary Sectors, while the arrow with the empty tip represents the subProperty relation between the two represented object properties. The dashed arrow represents the inference on the property 'has activity'. Indeed an entity may carry out more than one activity, but only one is the main activity.

The classification of the entity on the basis of its main activity is particularly useful to verify that 'contract notices-public (or ordinary) sectors' are issued by entities operating in the ordinary sectors, while 'contract notices-utilities' are issued by entities operating in utilities. There are indeed different standard forms for these two categories of contract notices. This aspect is addressed through a DL restriction on the two classes, namely:

(*a*)   `ContractNotice-PublicSectors` ⊑
`ContractNotice` ⊓
`(∃issuedBy. EntityOperatingInOrdinarySectors)`
(*b*)   `ContractNotice-Utilities` ⊑
`ContractNotice` ⊓
`(∃ issuedBy. EntityOperatingInUtilities)`

Moreover, as explained in the previous sections, the application of the EU Directives 2004/18 or 2004/17[26] depends by this distinction.

Another inference allows to classify a proposed public contract as 'proposed public service contract', 'proposed public works contract' and 'proposed public supply contract' on the basis of the declared object of the contract. The Figure 7 shows an example of this type of inference. It must be emphasized that the contract notices do not indicate explicitly whether a proposed contract is a mixed one (i.e. a proposed contract with both services or supplies, or works and services, or works and supplies as object) because the awarding authority qualifies the type of contract (for instance as services public contract) before the publication of the contract notice, on the basis of some detailed rules. For this reason, a public service contract is expressed in the ontology as a public contract that has at least one Service as Object of Contract, i.e. not excluding, for instance, that a proposed public service contract may have works as object of contract.

Now, consider the contract notice n. 382532-2011, whose tabular summary is shown in Figure 1. The con-

tract notice is issued by The Open University and it announces the tender 382532-2011 for the award of a proposed contract, whose object is cleaning services and related. Figure 9 shows the matching between the data coming from this contract notice and the Open-Corporates database. The information coming from OpenCorporates can be filtered by searching for all the active (i.e. non in-liquidation and non inactive) corporations registered in the UK, whose sector of activity, according to the Standard Industrial Classification (SIC)[27] includes "cleaning services". Of course, there are a lot of corporates corresponding to these features and we have included only a few of them in Figure 9. As for the other figures, the gray rectangles represent the instances; the bounding rectangles represent the classes of the instances and the dashed black arrow represents the inference supported by LOTED2-core, while the dashed orange arrows represent the inferences supported by the LOTE2-extended ontology.

As explained in section 4.6, in some cases, the proposal of public contract is split into (more than one) smaller single proposals. These proposals, called lots, are basically the result of the subdivision of a single bigger proposal into many. This aspect is particularly relevant because the division in lots enables SMEs (small and medium enterprises) to access also EU tenders that otherwise would be accessed only by large turnover's enterprises. In fact, the EU Commission considers the sub-division of contracts in lots as a good practice in tenders' management and encourages its use[28].

Therefore, we have paid close attention to this aspect in LOTED2. Figure 10 shows an instantiation of

---

[26] See the two General Class Axioms in the ontology

[27] Note that the SIC system of classification can be matched with the Common Procurement Vocabulary (CPV) that is a system of classification for the description of the object of the contract to be awarded. The Common Procurement Vocabulary identifies more than 9400 products with a code composed of 9 numbers. The aim of the CPV is to standardize, by means of a single classification system for public procurement, the terms used by contracting authorities and entities, describing the object of contracts through a uniform nomenclature. The CPV is translated into 22 official languages of the European Community. It simplifies the task of drafting notices, since it describes the subject matter of contracts, and helps also the drafting of statistics on public procurement, because it is compatible with trade monitors used throughout the world (especially those used by the United Nations). The CPV code is also included in the tabular summaries and in LOTED2-core ontology is associated to the object of contract through a dataType property.

[28] See, for instance: http://ec.europa.eu/internal_market/publicprocurement/docs/sme_code_of_best_practices_en.pdf
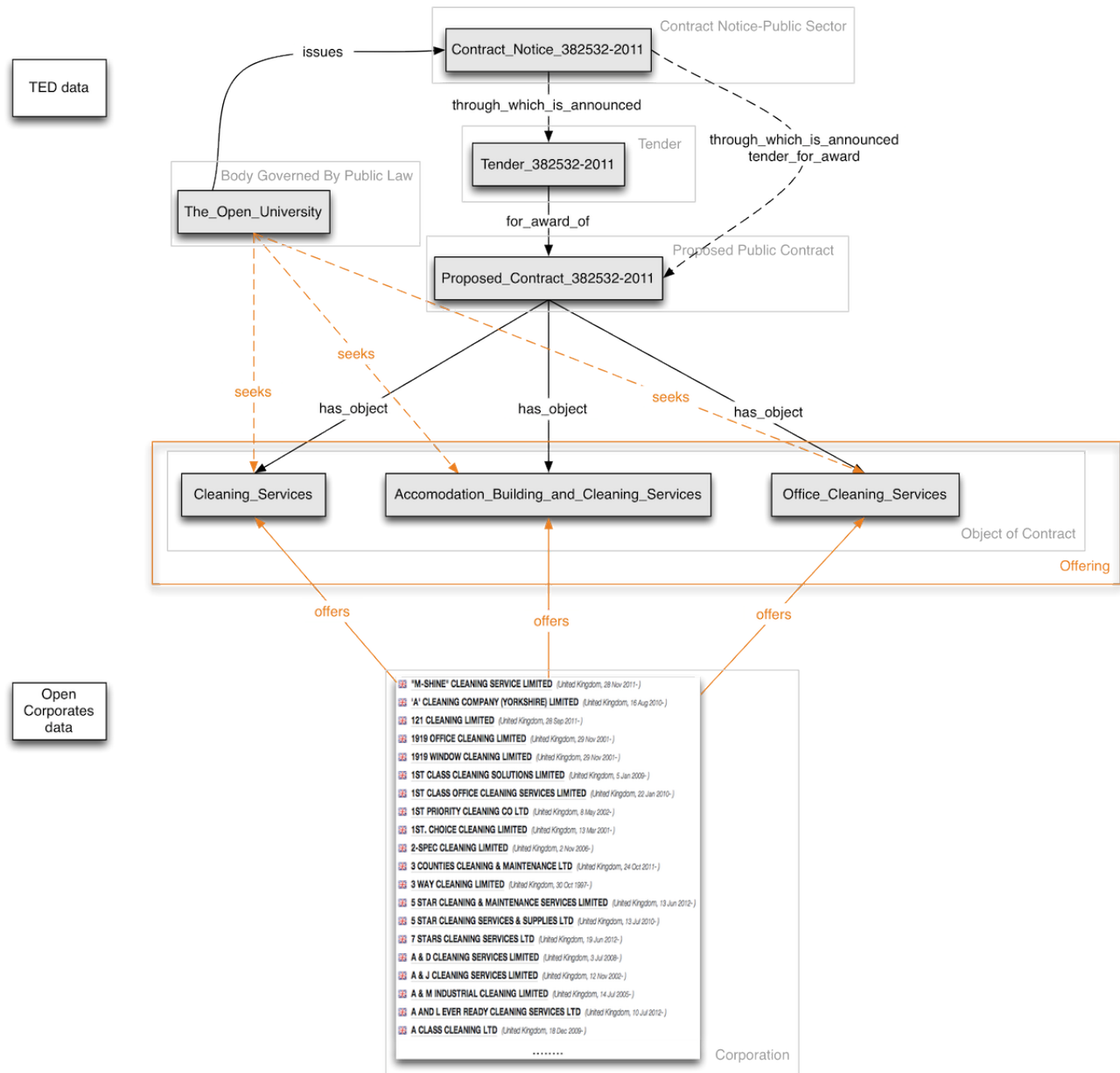
Fig. 9. An example of instantiation of LOTED2-extended ontology with TED and OpenCorporates data
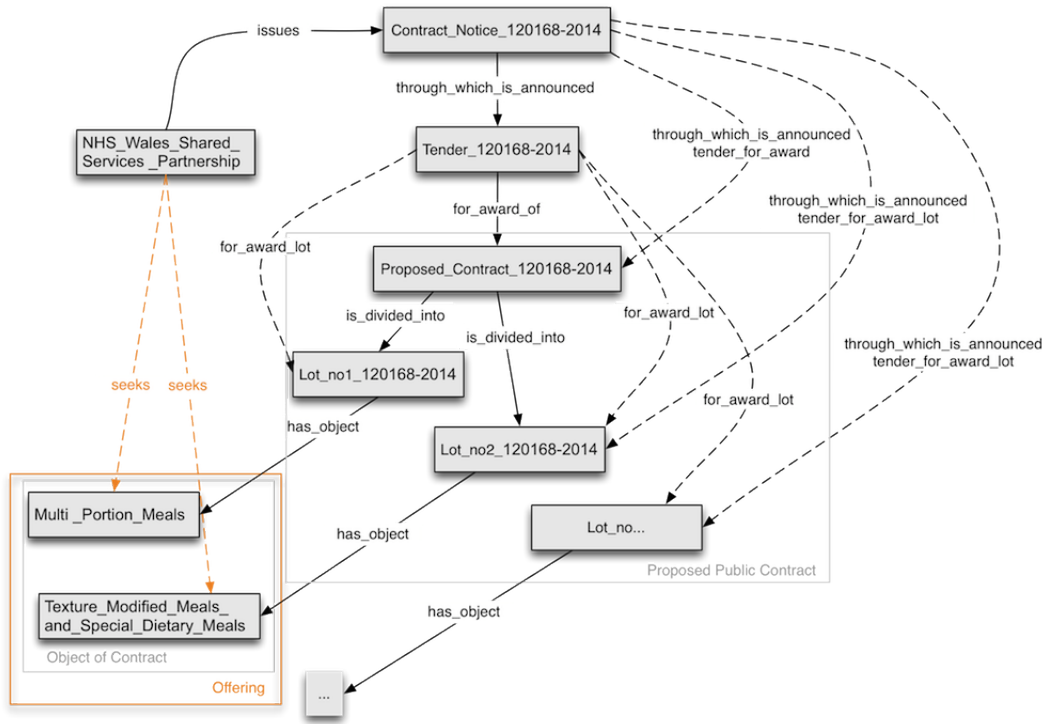
Fig. 10. An example of inferences concerning lots with an instantiation of contract notice n. 120168-2014

the contract notice n. 120168-2014[29] with respect to this particular. Thanks to the property chains described in sec. 4.6, LOTED2 infers that in case of proposed contracts divided into lots, the awarding authority is seeking all the objects of all the lots.

7. Questions, lessons learned and future work

Currently, more and more Open Government Data are available on the web. Since, in many cases, these data cover subjects strongly related to the legal domain, there are grounds for asking whether and how legal ontologies might support the construction of Legal Semantic Web applications that could be used for better delivering public sector information to the citizens. To that end, namely trying to give an answer to these questions, we have carried out an 'experiment' in the European public procurement field, building an ontology called LOTED2, whose main purpose is to support the construction of a Legal Semantic Web application

for enhancing the meeting between the public demand and the market offering.

LOTED2 is a legal ontology inasmuch it tries to represent legal concepts as they result from the interpretation of the legal sources. However, unlike many other legal ontologies, LOTED2 tries to find connections with other Linked Data models relevant to the domain and, just in order to achieve this purpose, it has been built seeking for a balance between the accurate representation of legal concepts and the usability of the ontology as a knowledge model for Semantic Web applications.

Figure 11 shows the sketch of a Semantic Web application that aggregates and reconciles information about EU public procurement notices and business entities, drawing upon the ontology we have developed. The right side of the figure represents real world entities, as public institutions and economic operators. When a public institution issues a contract notice, the data pertaining to this contract notice are instantiated into the LOTED2 ontology; data from OpenCorporates become instances of LOTED2 as well (as shown in Figure 9). Then, thanks to the inferences that LOTED2 supports and by means of a reasoner, the offering and demand side are matched.

---

[29]The contract notice is available at http://ted.europa.eu/udl?uriTED:NOTICE:120168-2014:TEXT:EN:HTML
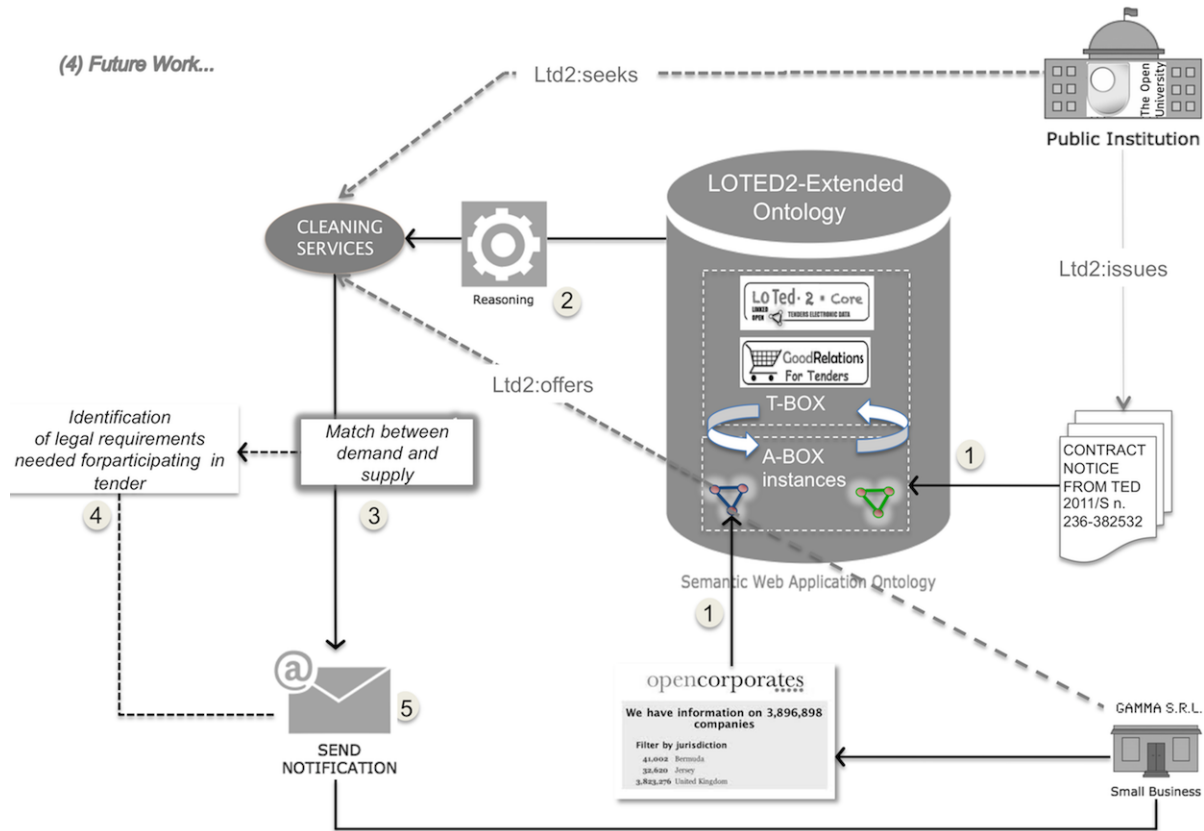
Fig. 11. Sketch of a Semantic Web application employing the LOTED2 ontology

In principle, thus, there is a place for legal ontologies in the Linked Data and Semantic Web applications scenario, but, in practice, there are still other broader issues to take into consideration.

It is worth noting that OpenCorporates makes available also data about non-European business entities and, as we have highlighted in the previous sections, under certain circumstances even the non-European traders may participate in EU tenders. The information about jurisdictions and regulations, therefore, is essential for the construction of a Semantic Web application (like the one drawn in Figure 11) which aims at including also non-European business entities' data. Otherwise, by linking data of companies from states not involved within the scope of application of the regulation that covers the contract notice, a sort of *semantic noise* can be generated.

Consider, for example, the contract notice 2011/S n. 236-382532 (Figure 1) issued by 'The Open University' (that is a body governed by public law) for

the award of a public contract concerning cleaning services. There are, indeed, two key questions related to this contract notice, with not easy solutions for traders who are interested in the application for this notice.

First: the Sigma Corporation is registered in Canada. Can the Sigma Corporation apply for this notice issued by The Open University (UK)?

Second: the Gamma S.r.l. is registered in France. What types of documents should the Gamma S.r.l. submit to The Open University for participating in the tender n. 236-382532?

The answer to the first question depends on the type of 'Regulation' that covers the contract notice. The contract notice states that the Government Procurement Agreement covers this particular type of contract, but this is not enough to give an answer to the first question. It is necessary to consider all the exceptions agreed by single parties to the application scope of the GPA (in this case between Canada and EU). The answer to the second question, instead, depends on the

criteria that must be met by business entities that want to participate in EU tenders.

Thus, in order to avoid alerting economic operators which are not admitted to the participation of this specific tender (e.g. because they have their registered headquarter in countries which are not covered by the GPA) an additional part of reasoning (which currently is not addressed by this LOTED2 version) is needed (step no. 4 in Figure 11). Furthermore, it would be useful to provide also information concerning the requirements needed for participating in the specific tender, customized on the basis of the economic operator's nationality.

It is worth mentioning that many initiatives have been started by the European Commission in order to improve the access of traders (in particular of SMEs) to the public procurement market in EU. This kind of vision has inspired, for example, the Peppol EU Project[30], which aims to make possible electronic communication between any company in the European Union with any governmental institution for all procurement processes. Another initiative is *e-certis*[31], namely a system that helps authorities and economic operators in identifying the different certificates and attestations (that are evidences for requirements) frequently requested in procurement procedures across the 27 Member States, two Candidate Countries (Turkey and Croatia) and the three EEA countries (Iceland, Liechtenstein and Norway). These two initiatives provide us with many useful resources for modeling the set of requirements that must be fulfilled by economic operators (eligible economic operators) for participating in European tenders (such as, for example, the absence of conviction by final judgment for participation in a criminal organization, fraud, money laundering, etc.).

On the other side, even if the use of a Semantic Web application like the one shown in Figure 11 is limited within the European borders, there are other issues to take into careful account.

It is not by chance that we decided to show in Figure 9 just the OpenCorporates data filtered with respect to the UK jurisdiction. In fact, in accordance with the "Open Company Data Index"[32] – that is monitored and calculated by OpenCorporates itself according to some metrics like the free availability and searching of company data, the type of license associated to the

data, the publicity of information regarding company directors and so on – the UK is the country that obtains the best score (90/100) in terms of accessibility and 'openness' of company registration information. In contrast, data about companies based in many other European countries like, for instance, Germany, France or Italy are not *Open Data* and therefore, these are not included in the OpenCorporates database. As shown by the Open Company Data Index, in fact, in the case of Germany and France (score 20/100), it is only possible to search for basic company data, but the dataset can not be downloaded or accessed over an API and it is not released under an open license; instead, in the case of Italy, there is only a limited possibility to search for data about companies (score 10/100).

There are important legal questions related to this issue. Is there an *Open*, or rather, a *Closed Data Divide*? If, in a future not too far away, Legal Semantic Web applications like the one drawn in Figure 11 were really adopted, could this *Data Divide* entail an advantage for certain countries with respect to others? And, therefore, could the inequalities in terms of access to data jeopardize also the equality of treatment in European tenders participation? If this is the case, do we still need to find just a compromise between legal ontologies and Linked Data models? Should we not rather wonder about the impact, in general, of Semantic technologies on the legal domain?

In our opinion, this could actually be the case, and the work we have done so far should be considered just the first step towards this direction. For the time being, there is a number of lessons that we have learnt form our work.

7.1. The Semantic Web and the legal domain.

In open systems the heterogeneity, the scale, the data quality, the reliability of information raise significant problems of control of the legal effects of heterogeneous linked data, or 'smart' data. Moreover, the inequalities in data accessibility affect the right of equal access to core information. Because of this, building Legal Semantic Web applications will require the ability to foresee and to cope with all kinds of risks that may emerge, in order to avoid that the challenge will turn into a danger.

7.2. Semantic interoperability among legal ontologies and Linked Data models.

The integration of legal ontologies with ontological resources related to different domains represents a

---

[30]http://www.peppol.eu

[31]http://ec.europa.eu/markt/ecertis/login.do

[32]http://registries.opencorporates.com/

fairly unexplored field. A sort of *rigidity*, which derives from the fact that authoritative sources drive the conceptual model, characterizes ontologies about legal contents. This rigidity contrasts with the heterogeneity of the Semantic Web, which instead is characterized by a non-unified and non-univocal knowledge representation. Because of these aspects, it is no accident that legal ontologies are mostly conceived in closed systems rather than open ones.

7.3. The value of "little semantics" in the legal domain.

The attempt in integrating LOTED2-core and Good Relations shows also other things. On the one hand, through the properties "seeks" and the mirror inverted "offers", Good Relations represents in an intuitive way the intents of parties when they issue an invitation to treat or a proposal for the conclusion of a contract (an 'offer' in legal terms). In a certain sense, Good Relations shows the final part of the legal superstructure behind the terms "seeks" and "offers". On the other hand, the integration of the LOTED2 ontology with Good Relations shows that also in the legal domain "a little semantics goes a long way" [28]. Nevertheless, it must be emphasized that in the legal domain this little semantics is just the superficial layer of the whole of legal knowledge upstream. Many other relevant legal issues (like that of requirements for the participation in tenders) require an intensive design to be addressed within the limitations of the Semantic Web languages. That is because, in the legal domain, question answering is not only information retrieval. Information retrieval is not enough, since "question requires some deduction or inference before an appropriate answer can be given" and "regulations may contain many different articles about the same topic and one can only assess whether something is permitted or not by understanding the full documentation". "A rather detailed understanding is required, in particular, because regulations generally contain complex structures of exceptions" [4]. In other words, question answering in the legal domain is not a trivial matter.

7.4. Compliance by design in legal ontologies modeling.

The variety of related work about procurements (up to now there are three ontologies of public contracts, including LOTED2, and another is in development) reveals the strong interest of the Semantic Web community in representing legal knowledge. However, the Semantic Web community should consider the matter more carefully and the Legal Informatics community should try to take up the challenge. In this respect, it is worthwhile to emphasize that an exciting research field is emerging on the use of ontologies for checking compliance of legal documents or processes with norms regulating them [12,34], even by combining ontologies with rules formalized through interchange standards specific for the legal domain [37]. Nevertheless, at the same time, the aspect of compliance by design should be more stressed in legal ontology engineering. Ontologies for compliance and ontologies in compliance might be considered as two sides of the same coin.

Nowadays, technologies in general, and even Semantic technologies are causing so many fast and unexpected changes that the Law is under constant and severe pressure. We can not pretend that these major changes do not also require legal solutions and a way to re-think our approach to the legal knowledge. Surely, re-thinking the legal knowledge in the Semantic Web is a challenging task, not at all simple. However, it is worth to face the challenge, starting from the lesson we have learnt so far.

References

[1] L. E. Allen. Language, Law and Logic: Plain Legal Drafting for the Electronic Age. *Computer Science and the Law*, pages 75-100, Cambridge University Press, New York, 1980.

[2] J. M. Álvarez, J. E. Labra, A. Marín, and J. L. Marín. Semantic Methods for Reusing Linking Open Data of the European Public Procurement Notices. In *Poster Session of the Extended Semantic Web Conference 2011 PhD Symposium*, Crete, 2011.

[3] G. Barabucci, L. Cervone, A. Di Iorio, M. Palmirani, S. Peroni and F. Vitali. Managing semantics in XML vocabularies: an

experience in the legal and legislative domain. In *Proceedings of Balisage: The Markup Conference* 2010. Balisage Series on Markup Technologies, vol. 5 (2010)

[4] V. R. Benjamins, P. Casanovas, J. Breuker and A. Gangemi. Law and the Semantic Web, an Introduction. In V. R. Benjamins, P. Casanovas, J. Breuker and A. Gangemi, editors, *Law and the Semantic Web: Legal Ontologies, Methodologies, Legal Information Retrieval and Applications, Lecture Notes in Computer Science*, pages 1-17, Springer Berlin Heidelberg, 2005.

[5] C. M. Bianca. Diritto Civile: Il Contratto, pages 1-447, Giuffrè, Milano, 2000 (in Italian).

[6] A. Boer, R. Winkels, F. Vitali. Metalex XML and the legal knowledge interchange format. In P. Casanovas, G. Sartor, N. Casellas and R. Rubino, editors, *Computable models of the law, Lecture Notes in Computer Science*, pages 21-41, Springer Berlin Heidelberg, 2008.

[7] C. H. Bovis. EU Public Procurement Law. Edward Elgar Publishing, 2008.

[8] J. Breuker and R. Hoekstra. Epistemology and ontology in core ontologies: FOLaw and LRI-Core, two core ontologies for law. In *Proceedings of the EKAW04 Workshop on Core Ontologies in Ontology Engineering*, pages 15-27, Northamptonshire, UK, 2004.

[9] P. Capps, M. D. Evans, S. V. Konstadinidis, editors. Asserting Jurisdiction: International and European Legal Perspectives, Hart publishing, USA, 2003.

[10] P. Casanovas, G. Sartor, M. A. Biasiotti, M. Fernandez-Barrera. Introduction: Theory and Methodology in Legal Ontology Engineering: Experiences and Future Directions. In G. Sartor, P. Casanovas, M. A. Biasiotti, M. Fernandez-Barrera, editors, *Approaches to legal ontologies, Law, Governance and Technology Series*, pages 1-14, Springer Netherlands, 2011.

[11] N. Casellas. Legal Ontology Engineering, *Law, Governance and Technology Series*, pages 1-297, Springer Netherlands, 2011.

[12] N. Casellas, J. Nieto, A. Meroño, A. Roig, S. Torralba, M. Reyes, P. Casanovas, *Ontological Semantics for Data Privacy Compliance: The NEURONA Project*. In AAAI Spring Symposium: Intelligent Information Privacy Management, 2010.

[13] M. d'Aquin and A. Gangemi. Is there beauty in ontologies? *Applied Ontology*, vol. 6, n.3, pages 165-175, IOS Press, 2011.

[14] M. d'Aquin, E. Motta, M. Sabou, S. Angeletou, L. Gridinoc, V. Lopez and D. Guidi. Toward a New Generation of Semantic Web Applications. *IEEE Intelligent Systems*, 23(3):20-28, 2008.

[15] M. d'Aquin, A. Schlicht, H. Stuckenschmidt, M. Sabou. Ontology modularization for knowledge selection: experiments and evaluations. In *Proceedings of the 18th International Conference on Database and Expert Systems Applications DEXA*, 2007, pp. 874-883.

[16] C. De Koninck and T. Ronse. European public procurement law: the European public procurement directives and 25 years of jurisprudence by the Court of Justice of the European Communities: texts and analysis. Kluwer Law international, BV, The Netherlands, 2008.

[17] J. Delgado, I. Gallego, S. Llorente and R. García, IPROnto: An ontology for digital rights management. In *16th Annual Conference on Legal Knowledge and Information Systems, JURIX 2003*, volume 106 of *Frontiers in Artificial Intelligence and Applications*, IOS Press, 2003.

[18] J. Klímek, T. Knap, J. Mynarz, M. Necăskỳ and V. Svátek. LOD2 deliverable 9a. 1.1: Framework for creating linked data

in the domain of public sector contracts. Also available from WWW: http://static.lod2.eu/Deliverables/deliverable-9a, 2012.

[19] J. L. Dietz, What is Enterprise Ontology? *Enterprise ontology*, pages 7-13, Springer Berlin Heidelberg, 2006.

[20] E. Francesconi, S. Montemagni, W. Peters and D. Tiscornia. Integrating a bottom-up and top-down methodology for building semantic resources for the multilingual legal domain. In *Semantic Processing of Legal Texts*, pages 95-121, Springer Berlin Heidelberg, 2010.

[21] A. Gangemi. Ontology design patterns for semantic web content. In *The Semantic Web-ISWC 2005*, pages 262-276, Springer Berlin Heidelberg, 2005.

[22] A. Gangemi and V. Presutti. Ontology design patterns. In *Handbook on Ontologies*, pages 221-243, Springer Berlin Heidelberg, 2009.

[23] A. Gangemi, M. T. Sagri, D. Tiscornia. A Constructive Framework for Legal Ontologies. In R. Benjamins, P. Casanovas, J. Breuker and A. Gangemi, editors, *Law and the Semantic Web*, pages 97-124, Springer Berlin Heidelberg, 2005.

[24] R. García, A Semantic Web Approach to Digital Rights Management, PhD thesis, Technologies Department, Universitat Pompeu Fabra. Also available from WWW: http://rhizomik.net/ roberto/thesis 2006.

[25] A. Gómez-Pérez, F. Ortiz-Rodriguez, B. Villazón-Terrazas. Legal Ontologies for the Spanish e-Government. In R. Marín, E. Onaindía, A. Bugarín, and J. Santos, editors, *Current Topics in Artificial Intelligence*, pages 301-310, Springer Berlin Heidelberg, 2006.

[26] T. F. Gordon, G. Governatori, A. Rotolo. Rules and Norms: Requirements for Rule Interchange Languages in the Legal Domain. In G. Governatori, J. Hall and A. Paschke, editors, *Rule Interchange and Applications, Lecture Notes in Computer Science*, pages 282-296, Springer Berlin Heidelberg, 2009.

[27] N. Guarino, and C. Welty. Evaluating Ontological Decisions with OntoClean. In *Communications of the ACM*. 45(2): 61-65, ACM Press, New York, 2002.

[28] J. Hendler. On beyond ontology. Keynote talk, *International Semantic Web Conference*, Sanibel Island, Florida, USA, 2003.

[29] M. Hepp. Ontology Engineering for Linked Data: What Makes For A Good Ontology? Keynote talk. In *EKAW 2010 Workshop W5*, page 1, Lisbon, Portugal, 2010.

[30] M. Hepp. GoodRelations: An Ontology for Describing Products and Services Offers on the Web. In A. Gangemi and J. Euzenat, editors, *Knowledge Engineering: Practice and Patterns, Lecture Notes in Computer Science*, pages 329-346, Springer Berlin Heidelberg, 2008.

[31] R. Hoekstra, J. Breuker, M. Di Bello and A. Boer. LKIF Core: Principled Ontology Development for the Legal Domain. In J. Breuker, P. Casanovas, M. C. A. Klein, E. Francesconi *Law, Ontologies and the Semantic Web. Channelling the Legal Information Flood, Frontiers in Artificial Intelligence and Applications*, vol. 188, pages 21-52, IOS Press, Amsterdam, The Netherlands, 2009.

[32] R. Hoekstra. Ontology Representation. Design Patterns and Ontologies that Make Sense. *Frontiers of Artificial Intelligence and Applications*, vol. 197, IOS Press, Amsterdam, The Netherlands, 2009.

[33] R. Hoekstra. Representing social reality in OWL 2. In E. Sirin and K. Clark, editors, *Proceedings of the 7th International Workshop on OWL: Experiences and Directions (OWLED 2010)*, CEUR Workshop Proceedings, 2010.

[34] M. Jarrar. Towards Methodological Principles for Ontology Engineering. Phd thesis, Vrije Universiteit Brussel, 2005.

[35] J. L. Marín, Á. Marín, M. Rodríguez, R. Calmeau, J. M. A. Rodríguez and J. Labra. Euroalert.net: building a pan-European platform to aggregate public procurement data and deliver commercial services for SMEs powered by open data. In *Workshop Share-PSI.eu*, 2011.

[36] M. Palmirani, G. Contissa, R. Rubino. Fill the Gap in the Legal Knowledge Modelling. In G. Governatori, J. Hall and A. Paschke, editors, *Rule Interchange and Applications, Lecture Notes in Computer Science*, pages 305-314, Springer Berlin Heidelberg, 2009.

[37] M. Palmirani, G. Governatori, A. Rotolo, S. Tabet, H. Boley, A. Paschke. LegalRuleML: XML-Based Rules and Norms. In F. Olken, M. Palmirani and D. Sottara, editors, *Rule - Based Modeling and Computing on the Semantic Web, Lecture Notes in Computer Science*, pages 298-312, Springer Berlin Heidelberg, 2011.

[38] A. Ross. Tû-tû. *Harvard Law Review*, pages 812-825, 1957.

[39] G. Sartor. Legal Reasoning A cognitive approach to the law. In E. Pattaro, ed. in chief, *Treatise of Legal Philosophy and General Jurisprudence*, Vol. 5, Dordrecht, The Netherlands, Springer, 2005.

[40] G. Sartor. Fundamental legal concepts: a formal and teleological characterisation. *Artificial Intelligence and Law*, Vol.14 n.1, pages 101-142, Kluwer Academic Publishers, 2006.

[41] G. Sartor. The Nature of Legal Concepts: Inferential Nodes or Ontological Categories? *Law Working Papers*, EUI, Florence, 2007.

[42] R. Shearer, B. Motik and I. Horrocks. HermiT: A highly-efficient OWL reasoner. In A. Ruttenberg, U. Sattler and C. Dolbear, editors, *Proceedings of the 5th International Workshop on OWL: Experiences and Directions, OWLED 2008*, pages 26-27, Karlsruhe, Germany, 2008.

[43] E. Sirin, B. Parsia, B. C. Grau, A. Kalyanpur and Y. Katz. Pellet: A practical owl-dl reasoner. *Web Semantics: science, services and agents on the World Wide Web*, pages 51-53, 5 (2), 2007.

[44] D. Tiscornia, E. Francesconi, Building Semantic Resources for Legislative Drafting: The DALOS Project. In P. Casanovas, G. Sartor, N. Casellas, R. Rubino, editors, *Computable Models of the Law, Lecture Notes in Artificial Intelligence*, pages 56-70, Springer Berlin Heidelberg, 2008.

[45] F. Valle, M. d'Aquin, T. Noia and E. Motta. LOTED: Exploiting Linked Data in Analyzing European Procurement Notices. In *1st Workshop on Knowledge Injection into and Extraction from Linked Data* collocated with EKAW 2010, Madrid, Spain, 2010.

[46] F. Vitali, F. Zeni, Towards a country-independent data format: the Akoma Ntoso experience. In C. Biagioli, E. Francesconi and G. Sartor, editors, *Proceedings of V Legislative XML Workshop*, pages 239-252. European Press Academic Publishing, 2007.