

PPROC, an Ontology for Transparency in Public Procurement

Editor(s): Pompeu Casanovas, Universitat Autònoma de Barcelona, Spain, and Victoria University, Australia; Monica Palmirani, University of Bologna - CIRSFID, Italy; Silvio Peroni, University of Bologna, Italy, and National Research Council, Italy; Tom van Engers, University of Amsterdam, The Netherlands; Fabio Vitali, University of Bologna, Italy

Solicited review(s): Aldo Gangemi, LIPN Université Paris13-CNRS-Sorbonne Cité, France and ISTC-CNR Roma, Italy; Axel Polleres, Vienna University of Economics and Business (WU Wien), Austria; Angelo Di Iorio, University of Bologna, Italy; one anonymous reviewer

José Félix Muñoz^{a, c, *}, Guillermo Esteban^a, Oscar Corcho^b and Francisco Serón^c

^a *Aragonese Foundation for Research & Development (ARAID), Spain*

^b *Ontology Engineering Group, Universidad Politécnica de Madrid, Spain*

^c *Legal-Business Laboratory, Universidad de Zaragoza, Edifice I+D, Walqa Technological Park, 22197 Huesca, Spain. * jfm@unizar.es*

Abstract. Public procurement or tendering refers to the process followed by public authorities for the procurement of goods and services. In most developed countries, the law requires public authorities to provide online information to ensure competitive tendering as far as possible, for which the adequate announcement of tenders is an essential requirement. In addition, transparency laws being proposed in such countries are making the monitoring of public contracts by citizens a fundamental right. This paper describes the PPROC ontology, which has been developed to give support to both processes, publication and accountability, by semantically describing public procurement processes and contracts. The PPROC ontology is extensive, since it covers not only the usual data about the tender, its objectives, deadlines and awardees, but also details of the whole process, from the initial contract publication to its termination. This makes it possible to use the ontology for both open data publication purposes and for the overall management of the public contract procurement process.

Keywords: ontology, public procurement, open government data, legal institution

1. Introduction

In the context of public procurement, advertising has always been an essential part of the process, as it fulfils a dual purpose: on the one hand, it is a resource for improving competitive tendering, and on the other, it constitutes an instrument for transparency and for the monitoring of the behaviour of the contracting authorities [1]. This second purpose is becoming increasingly important because one of the third-generation human rights is free access to public sector information, which is now included in the laws of the majority of developed countries [2].

With the progress of electronic government, the publication of information regarding contracting

procedures increasingly began to be performed using electronic means. For instance, some European directives from 2004 (Directives 2004/17/EC and 2004/18/EC)¹ created a specific mechanism called the “buyer profile”. All public sector entities must publish on it the notices about the contracts that they put out to tender. The most important of these announcements is the contract notice, where the characteristics of the contract are explained and businesses are invited to compete. Other advertisements report the progress of the contract,

¹ Today replaced by the directives of the European Parliament and of the Council: 2014/24/EU, of 26 February 2014 on public procurement and repealing Directive 2004/18/EC; and 2014/25/EU, of 26 February 2014, on procurement by entities operating in the water, energy, transport and postal service sectors and repealing Directive 2004/17/EC.

such as the award or the formalization notices. In consequence, buyer profiles have become the central information hub for companies and citizens when it comes to public procurement.

However, their use has been severely limited by the major functional and technical differences between the different profiles of different public authorities and the lack of interoperability among them, which makes the integrated processing of the information published on them very hard (for instance, aggregating the total income being awarded to a specific company across all the public authorities in a group of countries). One of the currently adopted solutions for this heterogeneity problem consists in forcing all public authorities to publish on a single website the announcements regarding tender procedures. In Spain, this site is the Public Sector Contracting Platform (PSCP).²

This solution may be sufficient in order to comply with the needs of competitive tendering, because it is enough to publish a limited set of announcements. However, transparency requires giving citizens much more information and, in addition, transparency practices can vary greatly depending on the policies followed by each authority. Therefore, from the perspective of transparency, the solution involves preparing standards that may be used by public administrations to publish all the information that they consider appropriate. This is the main reason why we have created the PPROC ontology.³

The rest of the paper is structured as follows. We start with a comment about the data models used in public e-Procurement applications (section 2). In section 3 we analyse existing ontologies in this domain (namely PCO, LOTED2 and MOLDEAS). Then we describe the ontology development method used (section 4) and the structure and main components of PPROC, as well as its relationships with other ontologies (section 5). In section 6 there is a description of its usage, including the experience of two Spanish public administrations that have been early adopters of this ontology and are now using it in their production environments to publish structured information about public contracts in their buyer profiles. In the conclusion, we describe future works and make some observations about the influence of the adoption of the ontology on the information system of these public administrations.

² <https://contrataciondelestado.es>

³ <http://contsem.unizar.es/def/sector-publico/pproc>

2. Data Models and Knowledge Representation in Public e-Procurement Systems: Process, Document and Institution Oriented Alternatives

The Zaragoza City Council (666,000 inhabitants) and the Provincial Government of Huesca – which assists 202 small municipalities in the province – publish nowadays the information about their contracts using PPROC. Prior to this, one of the tasks was the analysis of the information used in their contract management applications. In this analysis, we observed that the structure of the databases of these management systems was closely related to the temporal succession in which the information was being generated or received by the management bodies. This leads to an organization of data that we may call “process-oriented”. Even though the relational databases used by public administrators can be very diverse since they have been created by different software providers, we believe that this type of organization appears very often, because is closely related with the way of working of civil servants.

The electronic communication between the agents involved in public procurement is every day more important. There are various initiatives whose purpose is to create standards for this communications, as CEN BII,⁴ developed within the scope of the EU. In these standards, XML formats are defined to make it possible to structure the messages exchanged by the various agents involved in electronic procurement. In addition, as we have already mentioned, both the EU and national governments have created websites whose function is to centrally publish information about public tenders. The announcements that are published on these websites are among the first exchanges of information performed electronically using structured messages in the context of electronic government. However, to do this, the administrations chose to create “de facto” standards, such as the one established for the TED eSenders⁵ or CODICE,⁶ defined for the Spanish PSCP. The objective of these standards is to achieve interoperability, addressing communication between systems. Therefore, their domain is limited to the

⁴ CEN Workshop on Business Interoperability Interfaces for public procurement in Europe (<http://www.cenbii.eu/>) is an initiative of the European Committee for Standardization (<https://www.cen.eu>).

⁵ The Tenders Electronic Daily (<http://ted.europa.eu>) is the online version of the Supplement to the Official Journal of the EU, dedicated to public procurement.

⁶ <https://contrataciondelestado.es/wps/portal/codice>

information that, at any time, is transmitted between the various organizations that are involved in the process. Moreover, the structure of the information is provided by the content of the documents that are exchanged. Consequently, we can call these XML standards “document-oriented”.

Unlike the above situations, the development of ontologies should not be based on elements associated exclusively with the procedure – such as the sequence in which information or communication actions are created – or with the exchange of messages, but with a general, all-encompassing view of the reality to be represented. In our case, we have not aimed at the creation of models of general legal concepts, as happens in some core legal ontologies [3], but at the modelling of a specific social mechanism, used to connect the contracting process of public sector entities. From the perspective of the philosophy of law, this objective is related to the concept of a “legal institution” through which the physical, social and legal elements that comprise a given “social mechanism” are identified and described – such as marriage or contracts, for example – focused on the attainment of a defined objective.

According to an approach based on the “theory of the institution”, the central focus of the model would be public procurement, considered a legal institution whose purpose is the attainment of a “product”: a public contract [4]. We have considered this approach appropriate because considering public procurement (i.e. the domain to be represented) as a legal institution consisting of a group of resources focused on the attainment of an objective is closely associated with the functional aspects of organizations, as is the case with computer applications or tools. For all these reasons, we can state that the PPROC ontology is “institution-oriented”.

In PPROC the semantic relationships of the model are organized according to the role that each concept plays within the “institution” that is represented. The science of the law has been used in order to identify and define these semantic relationships, which is devoted to studying and organizing the legal elements that comprise institutions and the relationships between them. Therefore, the institution-oriented perspective is appropriate for developing a model whose structure will be in accordance with that defined by legal doctrine. In this way, using dogmatic categories, we have differentiated in PPROC, on the one hand, the objective elements, including the purpose of the contract, from the subjective elements (the parties),

and, on the other hand, the material elements, which are the elements related to the merits of the matter, from the formal elements, which are the ones related to the proceeding. In accordance with these classification criteria, the ontology is divided into four blocks, as described in section 5.

Furthermore, the concepts defined by the legal doctrine in each country are fairly uniform. In consequence, one of the benefits derived from this strategy can be to avoid the very different levels of granularity and the lack of consistency between the data models of different public administrations that we have observed in our work.

3. Related Work

In [5] an exhaustive study about the numerous initiatives implemented for the use of semantic technologies in e-Procurement was presented. Among these, some have focused on the announcement of information regarding public contracts. In the European context, the first experience was LOTED2 [6], which expanded the LOTED ontology [7] with the goals of (a) expressing the main legal concepts of the public contract announcements defined in legal sources, (b) supporting rich semantic annotation, indexing, search and retrieval of tender documents, (c) making it possible to reuse semi-structured data extracted from the TED system and (d) enabling the integration with other ontologies and vocabularies about related domains. The ontology bases most of its content on the two directives (2004/17/EC and 2004/18/EC), which at the time of development regulated public contracts in Europe, and is the result of a thorough study of legal documents. This means that the legal content of European procurement is heavily present and rigorously represented in LOTED2.

Another initiative focused on public procurement is MOLDEAS (Methods On Linked Data for E-procurement Applying Semantics) [8], an ontology focused on the representation of information contained in the announcements about public tenders. The objective of this ontology was to provide a pan-European standard about public procurement data, enriching it with the classifications of already-existing products and publishing it by following established open data guidelines.

A third initiative in the EU is the Public Contracts Ontology (PCO),⁷ implemented within the framework of the LOD2 Project [9]. One of its objectives was to demonstrate the application of Linked Data for the publication of information about contracts in the public sector. With this purpose in mind, PCO models the main aspects of public contracts, although not in great depth. The ontology considers “only the information that is publicly available in existing systems on the Web [...], mainly produced during the tendering phase”. Hence, the result is a lightweight ontology that reuses widely accepted ontologies and vocabularies such as VCard, Payments Ontology, schema.org, Call for Anything vocabulary and GoodRelations.

PCO and MOLDEAS use diverse sources of information, trying to identify an information core of the domain of public contracts, describing the main concepts of public procurement without delving much into detail. On the contrary, LOTED2 is focused on defining a complete legal ontology and its main sources of information are European directives. As a result, these ontologies differ also in their complexity. PCO defines most of the information that a public contract may need, but some specific relations, roles or behaviours are not strictly represented (e.g. the contracting body or the distinction between objective and subjective award criteria). For its part, LOTED2 represents almost every aspect of public procurement, including the properties needed to label information from the TED website, with the result that this model is closely related to the text of the 2004 directives.

After studying these ontologies, the decision to develop the PPROC ontology was taken: on the one hand, because two of them (PCO and MOLDEAS) did not have the degree of detail required for our purposes, and on the other, because the LOTED2 model was considered too complex and excessively centred in legal texts. Nevertheless, the main reason for undertaking a new development was that the main objective we chased with the use of the ontology was to improve the transparency of public contracting processes. From this perspective, none of the three ontologies studied were satisfactory, because transparency was not their primary goal. In consequence, they did not detail many of the public procurement concepts that are necessary for transparency purposes and they were not designed to facilitate the understanding by citizens of the information provided in the buyer profile.

4. Ontology Development

For the design of PPROC we applied techniques that are actually supported by most existing methods (e.g. the one proposed by Noy [10]) and the followed strategy was basically bottom-up. Firstly, we defined the domain – which was given by the goals, especially by transparency – and the territorial scope: the European Union. Afterwards we specified the final users and identified the ontological and non-ontological resources. The enumeration of the relevant terms was performed by legal experts from universities and public administrations, and by members of a software company. The latest task was a review, which was carried out when PPROC started being used to label the data of two buyer profiles.

4.1. Requirements

The first step was to define a clear scope of the size and complexity suitable for dealing with the development of the ontology. In a first approximation, the knowledge required to express the information actually published in the buyer profile was identified. Moreover, we include some information needed for transparency and not covered by other ontologies. This information relates to issues such as the reasons for choosing a specific procedure, members of the committees, the changes in the contract that are allowed, and the legal remedies and their resolution. As a result, it is possible to inform citizens about the entire life of the contract and not just about the procurement process.

We also defined the territorial scope, which was the European Union. However, as this is a fairly detailed model, there are terms (approximately 20%) that refer to particular features of Spanish law and even regional particular features within this. This does not represent an obstacle to the use of the ontology in other EU countries, as we basically describe the model established through directives, and users can ignore or adapt the particular features when there are none or when they are applied differently in their country. Similarly, the ontology can be used in Latin American countries, as there is a fairly close affinity between Spanish legislation and their legislation, which also includes an announcement mechanism that is equivalent to the buyer profile [11]. With regard to the possible use of PPROC in other countries, the international regulatory text with the greatest scope is the Agreement on Government Procurement

⁷ <https://github.com/opedatacz/public-contracts-ontology>

of the World Trade Organization,⁸ which does not establish a detailed regulation of procurement procedures.

In the specification of final users, we considered the following: contract authorities, suppliers, citizens, control authorities, and researchers and agencies interested in the study and analysis of public procurement. Below, we identify several of the competency questions [12] that the ontology should solve, provided by several stakeholders who wanted to make use of the data that was going to be published for different purposes.

- First 50 contracts with most budget
- List of the latest contracts awarded
- List of contracts by type
- List of contracts by procedure
- List of contracts grouped by managing department (i.e. water and sewer, gardening)
- List of suppliers that worked with public authorities in the year 2014
- List of steps taken by a contract
- Number of formalized contracts between 11/11/2011 and the current date
- Actual price of all the contracts started, awarded or formalized in 2011, 2012 and 2013
- Total price of the formalized contracts with the supplier MULTITEC
- Identifier, subtype and date of the formalized contracts with the supplier URBANCO
- URI and name of the managing department with the largest amount of contracts

Finally, we identified the ontological and non-ontological resources to be used. The non-ontological resources included the Common Procurement Vocabulary (CPV).⁹ It is exclusively used in the public procurement of EU countries and consists of a main vocabulary for defining the object of a contract without entering into great detail, and a supplementary vocabulary for adding further qualitative information. There are also some implementations of this vocabulary in RDF, such as the one available at <http://cpv.data.ac.uk> or the one described by Alvarez-Rodríguez and colleagues [13]. On the other hand, the reused ontologies are: Public Contracts Ontology (PCO), Organization Ontology,¹⁰ Friend Of A Friend (FOAF)¹¹ SKOS,¹² schema.org,¹³

Dublin Core Metadata Terms (dcterms)¹⁴ and Good Relations.¹⁵

We have made available the Ontology Requirement Specification Document (ORSO), which we initially built following our ontology engineering methodology, in figshare.¹⁶ For the purposes of completeness and better understandability, this ORSO is also linked from the ontology (both in the HTML documentation and as a value of the `rdfs:seeAlso` property associated with the ontology).

4.2. Development

The enumeration of the relevant terms for the model was performed with two sets of stakeholders. On the one hand, the company iASoft, which has developed the buyer profiles of numerous Spanish administrations, compiled the fields included in the various documents published on buyer profiles. On the other hand, several legal experts analysed the annexes of the European directives and Spanish legislation that specify the announcement models for their publication. Then, the ontology development team consolidated this information and used it as a basis to prepare an initial list of entities, including cardinality, domain and range for properties. This list had 111 first-level entities, and the team divided some of them into several second-level entities, of which there were 42 in total. This approach was completed and validated by the contracting managers of three public administrations: the Zaragoza City Council, the Provincial Government of Huesca and the Regional Government of Aragón.

At the same time, the classes and properties present in other ontologies that may be used to describe entities at a higher level were identified. Later, in order to link this set of entities to each other, in a joint task between ontology developers and legal experts, classes and properties were defined to organize the contents according to their nature and function. Therefore, although the strategy was basically bottom-up, as it started from the most detailed elements present in the buyer profiles and in the annexes to prepare an initial approximation of the ontology, an effort was also made to make them

⁸ http://www.wto.org/english/docs_e/legal_e/gpr-94_02_e.htm

⁹ http://simap.europa.eu/codes-and-nomenclatures/codes-cpv/codes-cpv_en.htm

¹⁰ <http://www.w3.org/TR/vocab-org/>

¹¹ <http://www.foaf-project.org/>

¹² <http://www.w3.org/TR/skos-reference/>

¹³ <http://schema.org/>

¹⁴ <http://dublincore.org/documents/dcmi-terms/>

¹⁵ <http://www.heppnetz.de/projects/goodrelations/>

¹⁶ <http://dx.doi.org/10.6084/m9.figshare.1314591>

consistent with the highest-level concepts defined in other ontologies. Finally, the ontology was implemented in OWL.

Another question to be determined during ontology development was which classes should be declared as being of mandatory use and which of recommended use. The sources of knowledge used did not resolve the question, as laws do not establish that there are fields that must necessarily be published and the practice of the different administrations varies considerably. In the absence of these sources, the CODICE standard was taken as a reference for Spain. Here, the properties that relate to the following ontology classes are declared as mandatory: `dcterms:title`, `pc:contractingAuthority`, `pproc:awardDate`, `pc:tender`, `pc:supplier` and `dcterms:identifier`.

These tasks were carried out during 2013 and the beginning of 2014. In April 2014, two of the aforementioned public administrations (Zaragoza City Council and the Provincial Government of Huesca) started labelling their buyer profiles according to the PPROC ontology, producing instances of the different classes and properties of the ontology. This activity served as a basis for a review of the ontology, which was carried out jointly by the legal experts, the public administrations and the ontology development team. The objective of the review was to fulfil the labelling expectations of the public administrations whilst maintaining the legibility of the model. Finally, at the end of 2014, the PPROC 1.0.0 version was published.

5. Ontology Description

The PPROC ontology is composed of 78 classes and 129 properties that make it possible to represent the contract and the procedure for its preparation. The class `pproc:Contract` is the main class for the definition of a contract. For the representation below, the other classes are divided into four blocks, each one of which includes the classes that are directly related to the following points of the contract: 1) the object, which is the supply that the contract covers; 2) the parts, which are the agents that participate in the procurement process and, when appropriate, in the contract (the contracting authority, tender, awarded tender, etc.); 3) the procedure, composed of the steps taken until the execution of the contract; and 4) the fulfilment, which includes actions that must be taken after the execution. Below we show the main

classes that comprise each block, some of the decisions taken in the design and, when appropriate, the relationship with other ontologies.

5.1. The `pproc:Contract` Class

The class `pproc:Contract` contains the basic information about the contract and serves as an entry point to link to the other classes. We define `pproc:Contract` as a subclass of `pc:Contract`. We have taken PCO as a starting point for the preparation of PPROC and the `pc:Contract` class is used as the domain of data properties such as `dcterms:title`, `pc:tenderDeadline` and `pc:actualEndDate`, among others. Many of these properties are being reused in the PPROC ontology and thus do not require changes. However, many other properties are rewritten, as PPROC has specific classes to describe these properties. The reason is that we provide a separate class for each kind of element, which helps ontology users to better define the scope of searches. Some PPROC classes are aligned towards PCO, and we define them as `rdfs:subClassOf` of their corresponding PCO classes.

Public contracts may belong to many different categories. Initially, we considered two different alternatives for categorizing these types of public contracts: SKOS classification schemes (e.g. the PCO ontology uses two of them for this purpose: `pc:ProcedureTypeScheme` and `pc:KindScheme`, to which `pc:Contract` is linked through the properties `pc:procedureType` and `pc:kind` respectively) and several class taxonomies for the different types of contracts. We have opted for the latter option (creating several class taxonomies to replace the use of `pc:KindScheme`), because (a) we wanted to specify more clearly the different types of contracts that we may have to deal with and (b) we added some class restrictions relevant for some of these classes (e.g. `pproc:ContractWithLots owl:subClassOf pproc:lot some pproc:Lot`).

Furthermore, contracts are not only classified according to their administrative type. Contracts can also be extendable, harmonized (i.e. reach the threshold of regulation of European procurement directives), private or multiannual, etc. (see Fig. 1). In addition, contracts may be subdivided into lots. Each lot is a contract in itself, with a defined object, which can be awarded separately but that forms part

of a main contract. PCO instantiates contracts, with or without lots, and lots as `pc:Contract`; that is, this class is the domain and range of the property used to relate contracts to their lots (`pc:lot`). In PPROC (as in LOTED2) specific classes for each of the cases are created.

As an example of the use of the ontology, we include the representation in JSON-LD of a supply contract for the acquisition of training equipment for the fire brigade of the Zaragoza City Council. The representation begins with some properties linked directly to the `pproc:Contract` class (see Fig. 2).

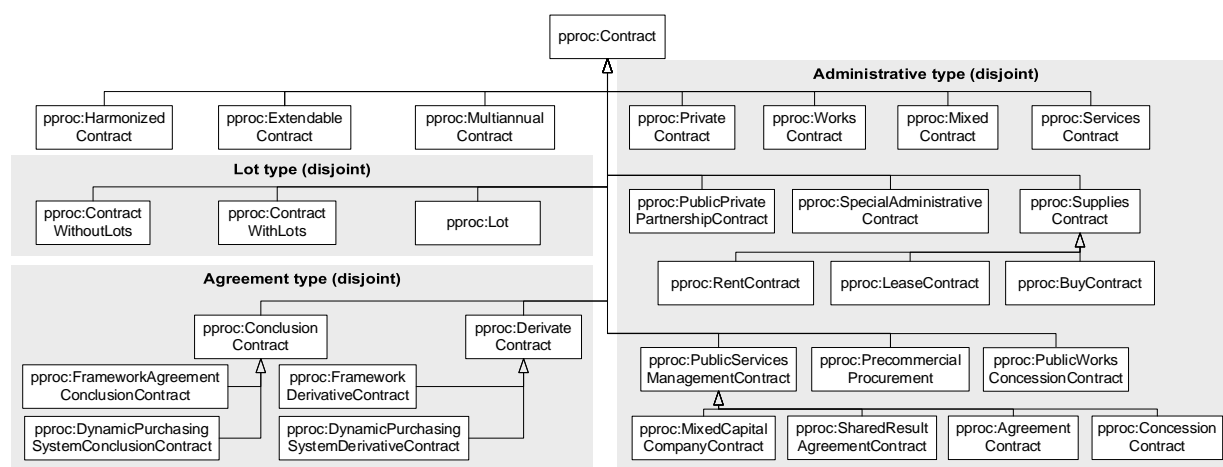


Figure 1. Classifications of the `pproc:Contract` class.

```

{
  "@context" : "http://contsem.unizar.es/docs/context.jsonld",
  "@id" : "contzar:0308848-14",
  "@type" : [
    "pproc:Contract",
    "pproc:SuppliesContract",
    "pproc:HarmonizedContract",
    "pproc:MultiannualContract",
    "pproc:ContractWithoutLots"
  ],
  "dcterms:title" : "Fire training system",
  "dcterms:identifier" : "0308848/14"
}

```

} Defines the object as a supplies contract, harmonized multiannual and without lots

} Defines the name and identifier of the contract

Figure 2. Representation in JSON-LD of some properties linked directly to the `pproc:Contract` class of a supply contract.

5.2. Contract Object

Two different (non-exclusive) approaches can be used to define the product or service that forms the contract object. The first one consists in using the `pproc:object` or `pc:mainObject` properties, which are especially appropriate for using the Common Procurement Vocabulary (CPV). This op-

tion must be included in the ontology, as the use of the CPV is mandatory in contracts subject to harmonized regulation and the control bodies recommend its use in all contracts. However, the scope of the CPV is restricted to public procurement, meaning that it is advisable to have a second way of describing the products and services for extended use in public Procurement. With this objective in mind, the `gr:Offering` class of GoodRelations is used [14].

To describe prices and payment options we use the `gr:PriceSpecification` class. However, when defining the price of a contract, a bundle of objects or a tender we find that `gr:PriceSpecification` and its known subclasses are not enough to describe some prices. Usually a `gr:Offering` contains a set of items the contracting authority is going to acquire. Its price can be set using the `gr:UnitPriceSpecification` class, which defines a price of a single item of the offering (e.g. an offering contains 100 printers with the `gr:UnitPriceSpecification` specifying the price of a single printer). However, usually it is also necessary to define the prices of a set of objects as a whole, treating them as a single package, and to

this end, PPROC includes the `pproc:BundlePriceSpecification` class. In addition, it must be remembered that, although in the majority of contracts the contracting authority receives the service and pays a price for it, there are also contracts in which it is the authority that offers a service (the provision of a public service, for example) and receives a financial consideration.

Finally, the classes that describe the criteria that will be used to award the tender also form part of this block. These are divided into objective criteria (such as the price and the delivery period), which can be quantified and applied through automated systems or by holding auctions, and subjective criteria, which are valued by experts.

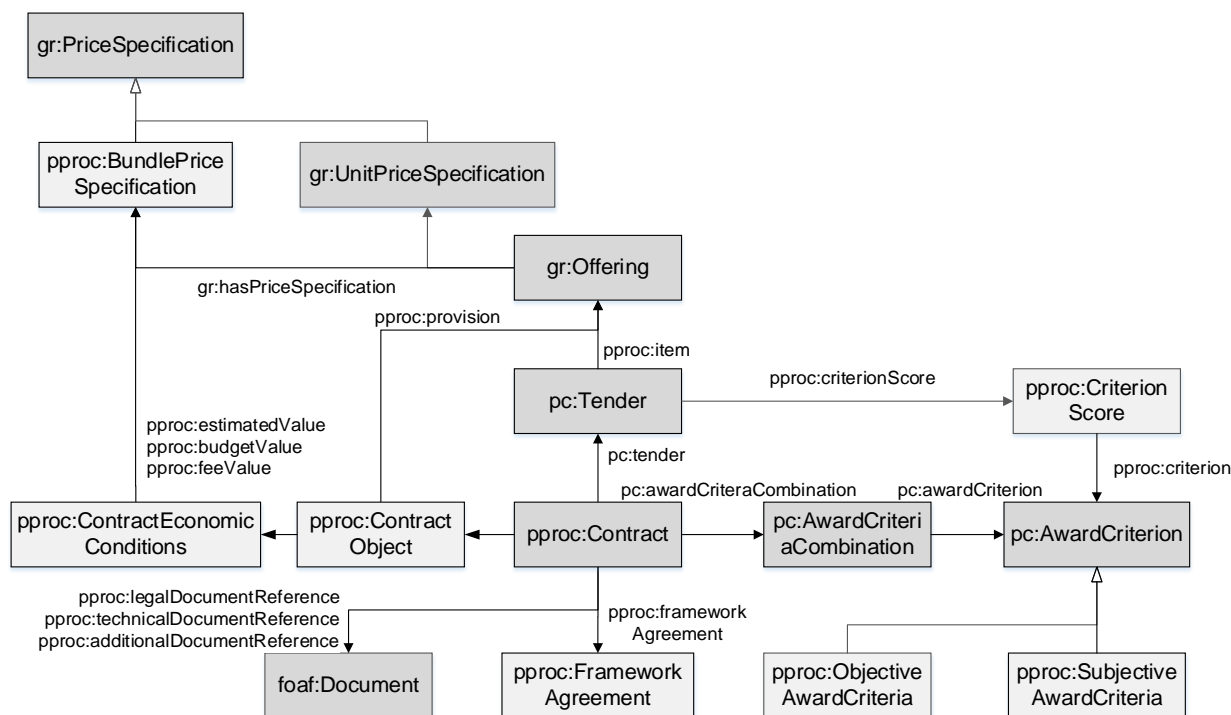


Figure 3. Contract object-related classes.


```

{
  ...
  "pproc:ContractObject" : {
    "@id" : "contzar:0308848-14/ContractObject",
    "@type" : "pproc:ContractObject",
    "pproc:mainObject" : "cpv:code-44212320",
    "pproc:provision" : {
      "@id" : "contzar:0308848-14/Object",
      "@type" : "gr:Offering",
      "dcterms:title" : "Fire training system using
                        industrial platform to Zaragoza
                        City Fire Brigade"
    },
    "pproc:contractEconomicConditions" : {
      "@id" : "contzar:0308848-14/ContractEconomicConditions",
      "@type" : "pproc:ContractEconomicConditions",
      "pproc:budgetPrice" : {
        "@id" : "contzar:0308848-14/priceWithVAT",
        "@type" : "pproc:BundlePriceSpecification",
        "gr:hasCurrencyValue" : "1005656.00.00",
        "gr:valueAddedTaxIncluded" : "true",
        "gr:hasCurrency" : "EUR"
      },
      "pproc:priceReviewAllowable" : "false"
    },
    "pc:awardCriteriaCombination" : {
      "@id" : "contzar:0308848-14/AwardCriteriaCombination",
      "@type" : [
        "pc:AwardCriteriaCombination",
        "pc:AwardCriteriaLowestPrice"
      ],
      "pc:awardCriterion" : {
        "@id" : "contzar:LowestPriceCriterion",
        "@type" : "pc:AwardCriteria",
        "pc:criterionName" : "Price",
        "pproc:evaluation" : "Minor price better score",
        "pc:criterionWeight" : "100"
      }
    }
  }
}

```

Linked data CPV code of the object of contract and the name of the contract object

Budget price of the contract, specifying tax is included and not reviewable

Award criteria combination, consisting in only one criterion of the lowest price

Figure 4. Representation in JSON-LD of the contract object of a supply contract.

5.3. Contract Parts

To describe the parts involved in a public procurement procedure we use the Organization Ontology, which includes the classes and properties needed to describe organizational structures and their hierarchy, through the `org:Organization` class and `org:subOrganizationOf` property. The role that an organization plays in a given procedure or contract – contracting authority, delegating authority, the organization on whose behalf a

contract is being made, the contracting body, managing department and the specific supplier of a tender (tenderer) – is established by the property used to link it with the contract (see Fig. 5). Note that in most cases we were able to use the `org:subOrganizationOf` property to describe the relation between the contracting body and the managing department belonging to a contracting authority.

To describe the persons grouped together to perform a task of the procedure, we use the `pproc:Committee` class. These committees are

known as contract bodies and may have different functions in the procedure. To differentiate them according to these functions, PPROC includes three subclasses of `pproc:Committee`. Members belonging to a concrete organization or committee are stated by using membership properties such as `s:member` or `org:memberOf`. Here we propose using properties belonging to other ontologies and vocabularies, such as the Organization Ontology itself, `schema.org`, Friend Of A Friend or `SKOS`. There are several contents where the location or a specific place should be known (e.g. the office of the contracting authority or a tenderer, the location where the goods should be left or the place of a meeting), and to define them we propose using the `s:Place` class and properties of `schema.org`.

Next, the class `pc:Tender` is reused to describe proposals made by suppliers. PCO uses two properties to link tenders to their related contract: `pc:tender` and `pc:awardedTender`. Once again, we believe that the properties defining the nature of an object can be replaced by defining the `rdf:type` of such an object. Therefore, although we still use the `pc:tender` property, we created subclasses to further define a `pc:Tender` (`pproc:AwardedTender`, `pproc:AcceptedTender`, `pproc:ExcludedTender` and `pproc:FormalizedTender`). Finally, tenderers are also defined through the Organization Ontology and linked using the `pc:supplier` property from a `pc:Tender`.

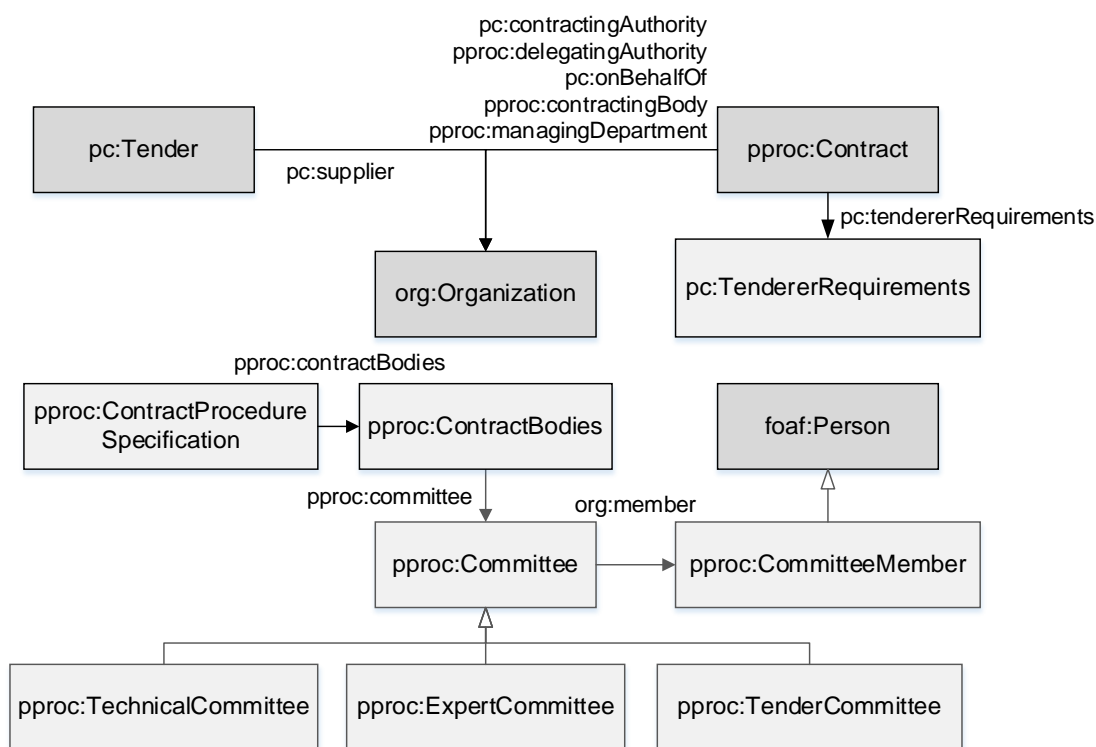


Figure 5. Parts-related classes.

```

{
  ...
  "pproc:managingDepartment" : {
    "@id" : "orgzar:fireServiceRescueAndCivilProtection",
    "@type" : "org:Organization",
    "dcterms:title" : "Fire service rescue and civil protection"
  },
  "pc:contractingAuthority" : {
    "@id" : "orgzar:1",
    "@type" : "org:Organization",
    "dcterms:title" : "Zaragoza City Hall"
  },
  "pproc:tenderersRequirements" : {
    "@type" : "pproc:TenderersRequirements",
    "pproc:requiredClassification" : "35B",
  },
  "pc:tender" : [
    {
      "@id" : "contzar:0308848-14/Tender1",
      "@type" : [
        "pc:Tender",
        "pproc:AwardedTender",
      ],
      "pproc:supplier" : {
        "@type" : "org:Organization",
        "s:name" : "Drager Safety Hispania S.A.",
        "org:identifier" : "A83140012"
      },
      "pproc:awardDate" : "2014-08-20",
      "pproc:formalizedDate" : "2014-09-18"
    },
    {
      "@id" : "contzar:0308848-14/Tender2",
      "@type" : "pc:Tender",
      "pproc:supplier" : {
        "@type" : "org:Organization",
        "s:name" : "Safety Fire Prevention Applied S.L.",
        "org:identifier" : "B74024472"
      }
    }
  ]
}

```

Definition of the managing department and contracting authority

Required classification of tenderers

Information about first (awarded) tender and the supplier of needs

Information about second tender

Figure 6. Representation in JSON-LD of the contract parts of a supply contract.

5.4. Procedure and Fulfilment

Another block of information refers to the procedure followed during the procurement process. The first aspect to represent is the kind of procedure and, in contrast to the decision made while classifying contracts and tenders, we use SKOS to define the kind of procedure and its urgency, using two concept schemes (`pproc:ProcedureTypeScheme` and `pproc:UrgencyTypeScheme`). The use of

SKOS is justified in this case because of its simplicity for this type of information where we do not have specific properties and property restrictions applicable to the different types of contracts and tenders. Both properties determining the procedure and urgency have `pproc:ContractProcedureSpecification` as their domain.

We included in PPROC all the information about it that could be useful to suppliers, such as tender requirements and briefing meetings. However, the information about the procedure is very important for

the control of contracting and, therefore, the ontology also includes classes to describe other points, such as the people that participate in the procedure or possible resources and their result (see Fig. 7). It is also necessary to know whether the type of procedure used is the one related to the contract, and the ontology has specific properties to do this, such as `pproc:Assumption` protecting the chosen procedure type.

Also, the term of the contract does not end with the execution, which is the time when the contracting procedure is considered to be finished. Contracts are often modified at a later time through specific procedures, which often change points such as the price or the term for completion. These modifications can be used to breach the principles of the contracting and, therefore, a fourth block is dedicated to this phase, which we call fulfilment. This contains classes that

make it possible to represent the conditions and limitations that possible modifications to the contract are subject to. If these occur, they can also be represented through the `pproc:ContractModification` class. Finally, it also includes classes to include the final result of the contract.

Regulation regarding these classes varies in the different territorial scopes. For example, the legislation for the Aragon region requires the publication of modifications to contracts, which is not the case with the Spanish legislation (although it will probably require it in the next reform of the law), or with European legislation. However, in any case, the block was considered necessary because the publication of modifications should be considered as good practice in terms of transparency.

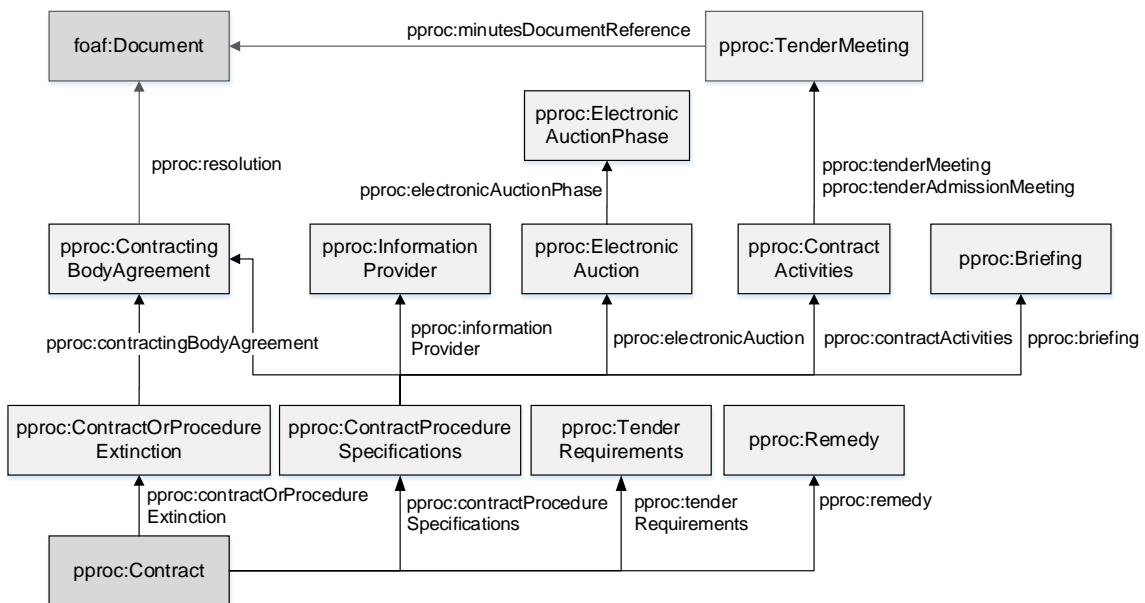


Figure 7. Procedure-related classes.

```

{
  "pproc:contractProcedureSpecifications" : {
    "@id" : "contzar: 0308848-14/ContractProcedureSpecifications",
    "@type" : "pproc:ContractProcedureSpecifications",
    "pproc:urgencyType" : "pproc:Regular",
    "pproc:procedureType" : "pproc:RegularOpen",
    "pproc:notice" : {
      "@id" : "contzar:0308848-14/ContractNotice/anuncioDOUE",
      "@type" : "pproc:ContractNotice",
      "pproc:noticeDate" : "2014-05-14",
      "pproc:noticeSite" : "DOUE",
      "pproc:noticeWeb" : "http://ted.europa.eu/udl?uri=TED:
                           NOTICE:161407-2014:TEXT:ES:HTML&src=0"
    },
    "pproc:tenderDossierStartDate" : "2014-05-07",
    "pproc:tenderDeadline" : "2014-06-16T00:00:00",
  },
  "pproc:contractTemporalConditions" : {
    "@type" : "pproc:ContractTemporalConditions",
    "pproc:estimatedEndDate" : "2014-12-31"
  }
  "pproc:contractExecutionConditions" : {
    "@type" : "pproc:ContractExecutionConditions",
    "pproc:maxSubcontracting" : "50.0",
  }
}

```

Specification of the procedure as regular and open

Information about the contract notice

Specification of some execution and temporal conditions

Figure 8. Representation in JSON-LD of the contract procedure and fulfilment of a supply contract.

6. Usage and Evaluation

PPROC is already being used by two public authorities of different size and scope (Zaragoza's City Council and the Provincial Government of Huesca), which are publishing open data about their public contracts so that they can be used not only by potential tenderers but also by citizens for transparency purposes. In both cases the corresponding PPROC-based RDF data is stored in a SPARQL endpoint,¹⁷ which can be queried for complex information needs. Additionally, in the case of Zaragoza, contracts are offered through their city API, providing data in JSON-LD (as described in the previous examples), as well as in other common formats such as JSON, CSV or XML. Finally, the whole buyer profile of the Zaragoza City Council is now completely based on the usage of PPROC-based data.¹⁸ At the time of writing, the Regional Government of Aragon has also begun work in order to publish its buyer profile by using the ontology.

With regard to its future use by local administrations, it is important to mention that PPROC is recommended as the ontology to be used by smart cities offering their public contract data according to the proposed technical norm of the Spanish Association for Standardization and Certification (AENOR), UNE 178301 on Open Data for Smart Cities.¹⁹ This technical norm also proposes the use of another nine vocabularies for the publication of data on transport, tourism, air quality, businesses, etc., and has been jointly developed by a combination of private companies, cities and other regional and national public administrations in Spain.

The ontology is also used to publish information about public sector contracts in Spain as a whole. To do this, a mapping with the CODICE XML standard has been performed, and a continuous transformation process is being carried out. These are published in a SPARQL end point,²⁰ which as of February 2015 contains 12 million RDF triples, including information about 199,611 public sector contracts.

¹⁷ <http://www.zaragoza.es/ciudad/risp/sparql.html>,
<http://www.dphuesca.es/sparql>

¹⁸ <http://zaragoza.es/ciudad/gestionmunicipal/contratos/>

¹⁹ <http://www.aenor.es>

²⁰ <http://pproc.unizar.es:8890/sparql/>

Table 1. Results of the query “Count of Contracts by Type” executed with the SPARQL endpoint that contains the public sector contracts in Spain

contractType	num
pproc:ServicesContract	85406
pproc:WorksContract	19189
pproc:PublicWorksConcessionContract	135
pproc:RentContract	15680
pproc:PrivateContract	3601
pproc:BuyContract	46275
pproc:DynamicPurchasingSystemDerivativeContract	268
pproc:PublicPrivatePartnershipContract	25
pproc:SpecialAdministrativeContract	3856
pproc:FrameworkDerivativeContract	2300
pproc:Contract	199611
pproc:DynamicPurchasingSystemConclusionContract	27
pproc:SuppliesContract	47012
pproc:FrameworkConclusionContract	1453
pproc:PublicServicesManagementContract	1671

All of the competency questions about the contracting procedures defined in the ORSD have been transformed into SPARQL and can be resolved using the ontology.²¹ For example, the enquiry “Count of Contracts by Type”, which in SPARQL format has the following syntax:

```
PREFIX pproc: <http://pproc.unizar.es/def/sector-publico/pproc#>
SELECT ?contractType count(distinct ?contract) as ?num
WHERE {
  ?contract a pproc:Contract
  . ?contract a ?contractType
}
GROUP BY (?contractType)
```

This query executed with the SPARQL endpoint that contains the public sector contracts in Spain, provides the result shown in the table 1.

However, as we have already mentioned, the information available in the different administrations varies greatly. Thus, while all these competency questions can be solved in the Zaragoza City Council’s SPARQL end point, in the case of the Provincial Government of Huesca, only the first four questions can be answered.

The Spanish Law 19/2013, regarding transparency, access to public information and good governance establishes a set of indicators and data that public entities have to publish on their websites. Some administrations also have their own rules about

transparency, such as the Ordinance of Transparency and Free Access to Information of Zaragoza City Council,²² which adds a considerable amount of data and indicators to the minimum established by law. We have confirmed as part of our evaluation that the PPROC ontology makes it possible to create SPARQL queries that literally reflect the text of the freedom of information act and the ordinance, such as the following:

- Budget volume in percentage of contracts awarded by each of the procedures provided for in the legislation.
- A list of all contracts awarded by the City, classified by type and amount, indicating the object, the amount of the bid, the award and the final cost, the procedure for the award, the instruments through which they have been published where appropriate, the number of participating tenders, the awarded tenders, the duration or timing of planned and actual implementation, modifications, and any other information of special interest to the public.

This is all covered now at the buyer profile of Zaragoza City Council (see footnote 18).

²¹ <https://github.com/pproc/pproc-sparql>

²² http://www.zaragoza.es/ciudadania/gobierno-abierto/transparencia/detalle_Normativa?id=3983

7. Conclusion and Future Work

In this paper we have described PPROC, an ontology for the description of public procurement that we have created with the aim of publishing, in a structured and standardized manner, public procurement information on the buyer profiles of public authorities. This ontology has the potential to improve efficiency (for example, since it enables computerized consultations of the profiles of the various administrations) and to increase transparency (including, for example, information about the committee members or the contract modifications).

PPROC has been developed following standard practices in ontology development, identifying competency questions with different stakeholders (public authorities, companies already working for them and legal experts), and published according to well-established recommendations for Linked Data vocabulary publishing. In this regard, it has been included at lov.okfn.org.²³ We have also provided examples of its usage, especially focusing on the generated JSON-LD context and a set of SPARQL queries that provide answers to the proposed competency questions.

With regard to future actions to be taken, one is to make the mapping between PPROC and the Open Contracting Standard (OCDS) that is being developed by the World Wide Web Foundation.²⁴ Version 1.0 of this standard was launched in November 2014 and a final 1.0 version will be published in summer 2015. The second one is the adaptation of the ontology to the reform of public procurement that will take place with the transposition of directives 2014/24/UE and 2014/25/UE to the different countries. In Spain there is a draft of the new act at an advanced stage of development, but which has not been made public yet. In any case, we expect that the impact of the modifications on the ontology will not be significant.

Other initiatives focus on the use of the ontology as a resource for the integration of information about contracting procedures. In this regard, work is being carried out in partnership with the Observatory of Public Procurement,²⁵ and a new-faceted search engine has been developed that acts on all of the infor-

mation labelled using PPROC, which already includes over 200,000 contracts. The structure of the data used in the TED system has also been analysed to carry out the mapping using PPROC and the possibility of creating a SPARQL end point with this information is being studied.

The appraisal made by the final users of the ontology has been positive. Its adoption has resulted in a major improvement, not only regarding information provided to suppliers and to citizens, but also to the information system of the organizations. Specifically, it has facilitated the integration of information about contracts, which used to be managed in various services and through different applications, and has made it possible to give information a structure that is closely linked to the knowledge and terminology used by experts in procurement. PPROC has been developed with a perspective oriented to the legal institution, and that makes interdisciplinary work between engineers and legal experts easier during the development of the ontology and, once it is finished, more understandable for legally trained users and possibly for everyone else too.

Acknowledgements. PPROC was developed through the project “Optimization of public procurement using semantic technologies (ContSem)”, led by the company iASoft (Oesia) and funded by the Spanish Ministry of Industry, Commerce and Tourism (TSI-020606-2012-4). The authors would like to thank Miguel Angel Bernal (co-author of the ontology), and Carlos Bobed and Carlos Becana (contributors) for their cooperation in the development of PPROC, and also those responsible for procurement and the buyer profile and the technicians of the participating administrations in the project: the Zaragoza City Council (María Jesús Fernández, Ana Budría, Laura Fernando, Victor Morlán and Rubén Notivol), the Provincial Government of Huesca (Cristina de la Hera, Montserrat Rodríguez and Javier Casado) and the Regional Government of Aragón (Ricardo Cantabrana, José María Subero and Eva Sanz).

References

- [1] Miroslav, M., Miloš, M., Velimir, Š., Božo, D., Đorđe, L.: Semantic technologies on the mission: Preventing corruption in public procurement. *Computers in Industry*, 65 (5), 878–890 (2014). doi:10.1016/j.compind.2014.02.003
- [2] Ackerman, J.M., Sandoval-Ballesteros, I.E.: The global explosion of freedom of information laws. *Administrative Law Review*, 58 (1), 85–130 (2006)

²³ <http://lov.okfn.org/dataset/lov/vocabs/pproc>

²⁴ <http://standard.open-contracting.org/>

²⁵ <http://www.obcp.es/>

- [3] Agnoloni, T., Fernández-Barrera, M., Sagri, M.T., Tiscorni, D., Venturi, G.: When a FrameNet-Style Knowledge Description Meets an Ontological Characterization of Fundamental Legal Concepts. In: Casanovas, P., Pagallo, U., Sartor, G., Ajani, G. (eds.) *AI Approaches to the Complexity of Legal Systems*. 93–112. Springer, Berlin (2010). LNCS 6237. doi: 10.1007/978-3-642-16524-5_7
- [4] MacCormick, N.: Norms, institutions, and institutional facts. *Law and Philosophy*, 17, 301–345 (1998). doi: 10.1023/A:1006034203352
- [5] Alvarez-Rodríguez, J.M., Labra, J. E., Ordoñez de Pablos, P.: New trends on e-procurement applying semantic technologies. Current status and future challenges. *Computers in Industry*, 65 (5), 800–820 (2014). doi:10.1016/j.compind.2014.04.005.
- [6] Distinto, I., d’Aquin, M., Motta, E.: LOTED2: An ontology of European public procurement notices. *Semantic Web journal*, Accepted (2014)
- [7] Valle, F., d’Aquin, M., Noia, T. D., Motta, E.: LOTED: Exploiting Linked Data in Analysing European Procurement Notices. In: Proceedings of the 1st EKAW Workshop on Knowledge Injection into and Extraction from Linked Data (2010)
- [8] Alvarez-Rodríguez, J.M., Labra, J. E., Cifuentes, F., Alor-Hernandez, G., Sánchez, C., Luna, J.A.G.: Towards a pan-European e-procurement platform to aggregate, publish and search public procurement notices powered by linked open data: the MOLDEAS approach. *International Journal of Software Engineering and Knowledge Engineering*, 22 (03), 365–383 (2012). doi: 10.1142/S0218194012400086
- [9] Nečaský, M., Klímek, J., Mynarz, J., Knap, T., Svátek, V., Stárka, J.: Linked data support for filing public contracts. *Computers in Industry*, 65 (5), 862–877 (2014). doi: 10.1016/j.compind.2013.12.006
- [10] Noy, N. F., McGuinness, D. L.: Ontology development 101: A guide to creating your first ontology Natalya. *Stanford Medical Informatics Technical Report*, SMI-2001-0880, (2001)
- [11] Moreno Molina, J.A., Panorama Comparado de la Contratación Pública en América Latina, *Observatorio de la Contratación Pública*, www.obcp.es (2015)
- [12] Grüninger, M., Fox, M.S., Methodology for the design and evaluation of ontologies. In: Skuce D. (ed.), *Workshop on Basic Ontological Issues in Knowledge Sharing at IJCAI’95*, Montreal (1995)
- [13] Alvarez-Rodríguez, J.M., Labra-Gayo J.E., Rodríguez-González, A., Ordoñez de Pablos, P.: Empowering the access to public procurement opportunities by means of linking controlled vocabularies. A case study of product scheme classifications in the European e-procurement sector. *Computers in Human Behavior*, 30, 674–688 (2014). doi:10.1016/j.chb.2013.07.046
- [14] Svátek, V., Kompuš P., Dudáš, M., Nečaský, M., Klímek, J., Procurement notice enrichment using product ontologies. In: *Proceedings of the 11th International Conference on Semantic Systems, SEMANTICS ’15* (2015), 200–203. doi: 10.1145/2814864.2814891