

# A Linked Term Bank of Copyright-Related Terms

Víctor RODRIGUEZ-DONCEL<sup>1,a</sup>, Cristiana SANTOS<sup>b</sup>, Pompeu CASANOVAS<sup>b,c</sup>,  
Asunción GÓMEZ-PÉREZ<sup>a</sup>

<sup>a</sup>*Ontology Engineering Group, Universidad Politécnica de Madrid*

<sup>b</sup>*Autonomous University of Barcelona-Institute of Law and Technology*

<sup>c</sup>*Royal Melbourne Institute of Technology-Centre for Applied Social Sciences*

**Abstract.** A multi-lingual term bank of copyright-related terms has been published connecting WIPO definitions, IATE terms and definitions from Creative Commons licenses. These terms have been hierarchically arranged, spanning multiple languages and targeting different jurisdictions. The term bank has been published as a TBX dump file and is publicly accessible as linked data. Models for the RDF data structure are based on Lemon and W3C Recommendations. The term bank has been used to annotate common licenses in the RDFLicense dataset.

**Keywords.** Term bank, linked data, copyright, legal localization, multilingualism

## Introduction

Term banks (also known as *terminologies*), controlled vocabularies and content-related thesauri help keeping translations consistent and help choosing the most adequate term when precision is required. Further, the *localization* of legal texts require of specialized terminologies where the exact concept in a legal system must be invoked.

The work presented in this paper describes a terminology created in a half-automated process, where terms and their definitions have been extracted and integrated from different lexical sources and mapped in a supervised process.

The resulting terminology has been published<sup>2</sup> in the TBX format –ISO 30042[1]– which is the standard for the exchange of terminologies; and also published in RDF, according to the schema described by Cimiano et al. [4]. The RDF version is especially suitable for establishing links with other terminologies, and IATE has been taken as external reference of some of the extracted terms. In addition, a full license has been annotated with the terms herein contained.

The use of a terminology of legal terms found in licenses is not exhausted with the mere translation or localization. Once in a digital format, it can alleviate the task of identifying the key elements in new licenses as in [5] or can help the study of comparative law.

The paper is organized as follows. Section 1 describes the motivation for having a term bank of copyright-related terms published as linked data. Details on the followed

---

<sup>1</sup> Corresponding author: vrodriguez@delicias.fi.upm.es

<sup>2</sup>The copyright-related terms term bank is online at: <http://lider2.dia.fi.upm.es:8080/tbx2rdf/resource/>

methodology and publication are given in Section 2; Section 3 provides the related work and finally Section 4 contains the conclusions.

## 1. Background knowledge: legal term banks as linked data

The representation of copyright and related rights constitutes a part of legal knowledge currently at the limelight of European policy. Progress has been made in delivering copyright-related actions identified in the Digital Agenda<sup>3</sup>, the Intellectual Property Strategy<sup>4</sup> and in the "Licences for Europe"<sup>5</sup>. Moreover, it is worth to consider the Draft Report of the European Parliament<sup>6,7</sup> towards an harmonization of copyright across Europe and updating it to meet current challenges, which safeguards fundamental rights and enables the offer of innovative online services in the EU.

The complexity of the regulatory system in this field, together with the variety of the corpus of copyright (patchwork of international and European sources, such as the Berne Convention for the Protection of Literary and Artistic Works, the WIPO Copyright Treaty, the Directive 2001/29/EC<sup>8</sup> (Copyright Directive), amongst other correlated sources<sup>9</sup>), poses difficulties to search, retrieve and understand the legal information in this domain. Moreover, in a pluralistic legal order [14] the "[EU] legislation is drafted in several languages and [...] the different language versions are all equally authentic. An interpretation of a provision of [EU] law thus involves a comparison of the different language versions"<sup>10</sup>, in accordance with the principle of

---

<sup>3</sup> Communication on content in Digital Single Market (COM(2012) 789 final).

<sup>4</sup> In order to modernise the EU copyright legislative framework, "A Single Market for Intellectual Property Rights" (COM(2011) 287 final) was announced, which proposed series of measures to promote an efficient copyright framework for the Digital Single Market that include short and long-term key policy actions in various areas: patents, trademarks, geographical indications, multi-territorial copyright licensing, digital libraries, IPR violations, and IPR enforcement by customs.

<sup>5</sup> As a premise for a cultural policy and from a structured stakeholder dialogue, industry-led solutions were put forward by stakeholders as a contribution to improve the availability of copyright-protected content online in the EU. Available at <http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52012DC0789>

<sup>6</sup> Draft Report on the implementation of Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society (2014/2256(INI)), 14/01/2015. A legislative proposal on copyright reform is expected for September this year 2015.

<sup>7</sup> In particular, the Report calls for the harmonization of copyright terms and exceptions across Europe, new exceptions for emerging use cases like audio-visual quotation, e-lending and text and data mining, as well as the adoption of an open norm to "allow for the adaptation to unanticipated new forms of cultural expression".

<sup>8</sup> The purpose of Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society (Copyright Directive 83), is to implement the WIPO Copyright Treaty and to harmonise aspects of copyright law across Europe, such as copyright exceptions.

<sup>9</sup> Connected legal instruments: the Directive 2009/24/EC of the European Parliament and of the Council of 23 April 2009 on the legal protection of computer programs, the Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases, the WTO's Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS).

<sup>10</sup> Case 283/81 *CILFIT e.a.* [1982] ECR 3415, paragraph 18.

linguistic equality<sup>11</sup>, which entails a "full multilingualism" [15]. Settled case-law refers that *'the need for a uniform interpretation of [EU] regulations makes it impossible for the text of a provision to be considered in isolation but requires, on the contrary, that it should be interpreted and applied in the light of the versions existing in the other official languages [...] [A]ll the language versions must, (...) be recognised as having the same weight'*.<sup>12</sup>

Identifiable hindrances prevent cross-border access to legal information:

- Disclosure of open data makes it difficult to retrieve relevant and useful information due to its overload and oversupply (large assortments of data);
- Legal documents are published as plain text without hyperlinks to the official legal resources, averting navigation and reasoning among documents; national and EU websites are sometimes poorly interconnected or they use different identification systems;
- Data is not always published in computational formats like RDF or XML for Linked Open Data, but in heterogeneous formats;
- Ambiguity and polysemy of legal terms [6]: the terminological misalignment and the conceptual misalignment [9] between the terminology used at the EU level from that of the national level, even when implementing EU directives [7];
- Context-specificity of legal terms: the meaning of terms is related to the context of the legislation defining it (several context-specific definitions of legal terms with a common thread)<sup>13</sup>;
- Cultural-specificity of legal terms: the meaning of terms is related to the context of the legal and political culture to which these terms belong to (think not only of the classical distinction between Common and Civil law countries, but of cultures with ideographic languages such as Chinese and Japanese) [8].

The need for cross-border multilingual access to legislation is required for legal practitioners, such as judges, lawyers, translators, legal drafters and scholars, but also to other decision-makers, amongst enterprises, public administrations and citizens, subject to regulatory compliance (even outside their own area of expertise and also jurisdiction), in order to: i) exploit legal (open) data and therefore produce new innovative services for the legal information provision market; ii) to predict the impact of implementing the EU legislature in each MS by enriching [7] structurally the documents (with navigable references along legal texts) and semantically (with concepts from ontologies and annotations); iii) information retrieval, automatic translation, automated reasoning; iv) ensure the principle of legal certainty; v) may serve to strengthen the textual (or literal interpretation), and teleological interpretation upon which the ECJ's reasoning primarily rests.

Our work integrates the Linked Open Legal Data [16] momentum that illustrates *"the accessibility and semantic interoperability of legal sources"*<sup>14</sup>. Some of the

---

<sup>11</sup> See EEC Council: Regulation No 1 determining the languages to be used by the European Economic Community, [1958] OJ L 17/385.

<sup>12</sup> See Case C-257/00 *Givane and Others* [2003] ECR I-345, para. 36 and C-152/01 *Kyocera* [2003] ECR I-13833, para. 32.

<sup>13</sup> This point is illustrated by the ruling of the ECJ Case 283/81 *CILFIT e.a.* [1982] ECR 3415, paragraph 19.

<sup>14</sup> European Council, Draft Strategy on European e-justice 2014–2018, 2013 (2013/C 376/06).

advantages for rendering multi-lingual, multi-jurisdictional legal term banks published as linked data are:

- Clear separation and identification of concepts and terms, as data fits a formalized model and every resource is identifiable in a permanent manner.
- Easy browsing from a term in one language to an equivalent term in another language, although this makes only full sense when a preferred term is specified<sup>15</sup>.
- Easy browsing among general terms and the jurisdiction-specific terms, as concepts can be hierarchically organized. This clarity helps towards the harmonization of copyright terms in the EU, an explicit goal in the EU copyright roadmap<sup>16</sup>.
- Easy comparative analysis, as multiple sources are provided;
- Improve the findability and facilitating the unequivocal identification of concepts and corresponding terms at both European and national levels;
- For organization of conceptual domain knowledge and its availability of interrelated data sets on the Web in standard formats.

## 2. Linked resources and methodology

### 2.1. Publication format

In order to build the present linked term bank, several resources have been considered:

a) **WIPO documents** (World Intellectual Property Organization) publishes documents which include glossaries on copyright-related terms<sup>17</sup>. WIPO is an especially authoritative source as the custodian of the treaties on copyright signed by almost every country. Given their almost-universal validity, the definitions provided by WIPO are attributed thus to the most general concepts.

b) **IATE** (Inter-Active Terminology for Europe) is the current EU's inter-institutional terminology resource database, created from several preexisting databases like EUROCAUTOM (Commission), TIS (Council) and EUTERPE (Parliament), among others. IATE is managed by representatives from different institutions including the authoritative entities like the European Court of Justice or the Translation Centre for the Bodies of the European Union. IATE contains more than 8 million terms in all official 24 EU languages. It has been recently published as a linked data resource [4].

c) **Creative commons licenses** are text documents published along with the referred work, and usually symbolized by icons, summaries or hyperlink references. Creative Commons licenses, massively adopted by the internet culture, have been published in versions tweaked for up to 60 different jurisdictions and different languages. These licenses commence with the definition of the key terms, which typically address the

---

<sup>15</sup> For example, IATE defines *preferred term* as: "a term which should be used instead of any other (equally correct) synonym(s) present, for harmonisation purposes"

<sup>16</sup> For a explicit mention, see the "Draft Report on the implementation of Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society (2014/2256(INI))"

<sup>17</sup> [http://www.wipo.int/edocs/pubdocs/en/copyright/891/wipo\\_pub\\_891.pdf](http://www.wipo.int/edocs/pubdocs/en/copyright/891/wipo_pub_891.pdf)

ones used in the target jurisdiction. From version 4.0<sup>18</sup>, Creative Commons aimed at a neutral text, capable of fitting every legal system. Consequently, these definitions have been added to the general concepts and not to the jurisdiction-specific concepts.

**d) Other resources.** Finally, the term bank can be linked to other linked resources to make it a highly connected linked data resource. In particular, the term bank of copyright-related resources has been linked to dbpedia and lexvo. The use of these resources was possible as they had been published under open licensing modalities.

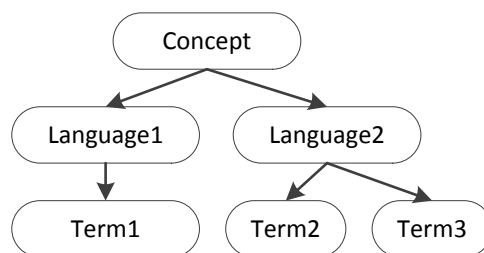
The methodology followed to create the term bank has been the following:

1. **Collection of top concepts.** Key copyright-related concepts have been extracted along with a general definition from the WIPO glossaries.
2. **Mapping to IATE.** The linked data version of IATE [4] version was systematically queried in search of direct matches. From the different sources of IATE, the legal one was preferred over others when more than one term matched. The task was verified and completed manually from the official IATE place<sup>19</sup>.
3. **Addition Creative Commons terms.** Over 100 creative commons terms have been defined, including the different versions, different jurisdiction ports and different languages. These resources are well classified in the RDFLicense dataset<sup>20</sup> [3], which also provides the links between license identifiers and legal texts. Creative Commons issued versions of the same license adapted to different jurisdictions before their version 4.0. Definitions from version 4.0 were added to the general concepts.

The publication of the dataset was made according to the linked data publication guidelines<sup>21</sup> and those specific for term bases [4].

## 2.2. The copyright term bank

A ‘terminology database’ or term bank is a database of abstract concepts together with the precise terms that refer to them, usually in different languages, as shown in Figure 1.



**Figure 1.** Concepts and terms in terminology databases.

<sup>18</sup> “version 4.0 international license [...] is the most up-to-date version of our licenses, drafted after broad consultation with our global network of affiliates, and it has been written to be internationally valid.”, <http://creativecommons.org/version4>

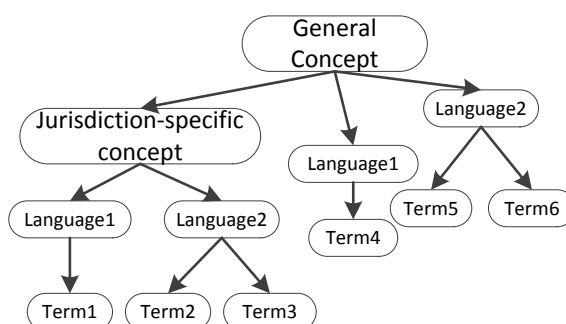
<sup>19</sup> <http://iate.europa.eu/>

<sup>20</sup> <http://rdflicense.appspot.com/>

<sup>21</sup> <http://www.w3.org/TR/ld-bp/>

Concepts are typically accompanied by definitions, whereas terms are sometimes provided with additional information like the source, reliability, domain, additional notes, comments and the context of use.

In order to build our term bank, the structure has been extended to tackle the multi-jurisdiction information that is provided, and jurisdiction-specific concepts have been arranged as subspecies of general concepts (Figure 2).



**Figure 2.** Concepts in a legal terminology database.

### 2.3. Publication format

TermBaseExchange (TBX) is the industry XML standard language used to represent terminology data, sometimes used as native format, sometimes as interchange format. It is published by ISO as standard ISO30042 and by the Localization Industry Standards Association (LISA). The following excerpt shows an example of the published terminology in TBX:

```

<termEntry id="Derivativework (ES)">
<langSetxml:lang="es">
<tig>
<term>obra derivada</term>
<termNote type="termType">fullForm</termNote>
<descrip type="reliabilityCode">3</descrip>
</tig></langSet>
<langSetxml:lang="ca">
<tig>
<term>obra derivada</term>
<termNote type="termType">fullForm</termNote>
<descrip type="reliabilityCode">3</descrip>
</tig></langSet>
</termEntry>
  
```

For an advanced format where linking to other sources is made more straightforward, the RDF data structure as in [4] has been chosen. This format assumes the Lemon model [2], whose representative schema is shown in Figure 3. In any case, the meta-operational relationship between legal reference and coreference has to be worked out.

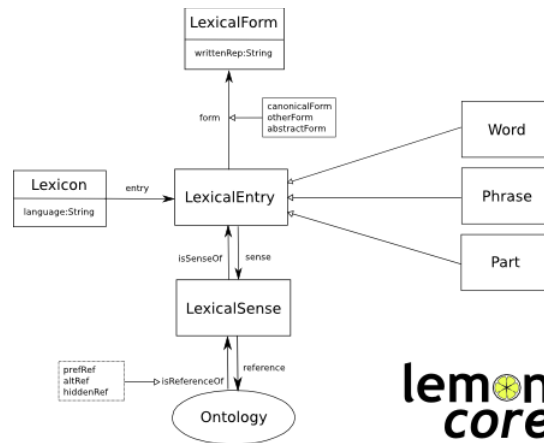


Figure 3. The Lemon Core Model (figure taken from <http://lemon-model.net>).

The example in the excerpt that follows shows two concepts: the concept of “derivative work” (lines 8-17) and the concept of “derivative work” in particular in the Spanish jurisdiction (lines 19-28). “Derivative work” is a general concept (`skos:Concept`) that can be linked to the corresponding IATE concept (74645) and even dbpedia resource (“Derivative\_work”). “Derivative work (ES)” is an abstract concept enshrined in 5 terms in 5 languages (Spanish, Catalan, Galician, Basque, Aranese) for which the creative commons licenses have a translation of the Spanish port. One of these terms is shown in lines 30-43, “obra derivada” in Galician language.

```

01 @prefix rdfs:<http://www.w3.org/2000/01/rdf-schema#> .
02 @prefix skos:<http://www.w3.org/2004/02/skos/core#> .
03 @prefix lemon:<http://lemon-model.net/lemon#> .
04 @prefix tbx:<http://tbx2rdf.lider-project.eu/tbx#> .
05 @prefix ontolex:<http://www.w3.org/ns/lemon/ontolex#> .
06 @prefix dct:<http://purl.org/dc/terms/> .
08 <http://tbx2rdf.lider-project.eu/converter/resource/cc/derivative%20work>
  a skos:Concept;
  rdfs:label"derivative work";
  lemon:definition [
    rdfs:label"a new work that translates or transforms one or more original
    copyrighted pre-existing works"@en;
    dct:source"WIPO" .
  ];
  lemon:reference <http://dbpedia.org/resource/Derivative_work> ;
  skos:closeMatch <http://tbx2rdf.lider-project.eu/data/iate/IATE-74645> .

19 <http://tbx2rdf.lider-project.eu/converter/resource/cc/derivative%20work%20%
28ES%29>
  a skos:Concept;
  rdfs:label"derivative work (ES)";

  cc:jurisdiction <http://dbpedia.org/resource/Spain> ;
  skos:narrower <http://tbx2rdf.lider-project.eu/converter/resource/cc/derivat
ive%20work> ;
  ontolex:isSenseOf <http://tbx2rdf.lider-project.eu/converter/resource/cc/obr
a%20derivada_gl> , <http://tbx2rdf.lider-project.eu/converter/resource/cc/lan%2
0eratorri_eu> , <http://tbx2rdf.lider-project.eu/converter/resource/cc/%C3%B2br
a%20derivada_oci> , <http://tbx2rdf.lider-project.eu/converter/resource/cc/obra

```

```
s%20derivadas_es> , <http://tbx2rdf.lider-project.eu/converter/resource/cc/obra%20derivada_ca> .

30 <http://tbx2rdf.lider-project.eu/converter/resource/cc/obra%20derivada_gl>
  a ontolex:LexicalEntry;
  rdfs:label "obra derivada_gl";
  lemon:language "gl";
  dct:source "http://creativecommons.org/licenses/by/3.0/es/legalcode.gl";
  tbx:reliabilityCode "3";
  skos:definition "e. Consideraranse obras derivadas aquelas obras creadas a partir da licenciada, como por exemplo: as traducións e adaptacións; as revisións, actualizacións e anotacións; os compendios, resumos e extractos; os arranxos musicais e, en xeral, calquera transformación dunha obra literaria, artística ou científica. Para evitar a dúbida, se a obra consiste nunha composición musical ou gravación de sons, a sincronización temporal da obra cunha imaxe en movemento (synching) será considerada como unha obra derivada para os efectos desta licenza.";
  ontolex:sense <http://tbx2rdf.lider-project.eu/converter/resource/cc/derivative%20work%20%28ES%29> .
```

### 3. Related Work

In the literature, different methods exist for approaching the multilingual complexity of European law, for example controlled vocabularies, implemented in terminology database (such as IATE run by all the main EU Institutions that we have resort to in our work), thesauri (as EUROVOC), semantic lexicons or lightweight ontologies (as WordNet, EuroWordNet and, in the legal domain, JurWordNet) that we evoke here. EuroVoc Thesaurus<sup>22</sup> is the most important multilingual, multidisciplinary standardized thesaurus created by the EU, covering the activities of the EU. EuroVoc is managed by the Publications Office, which moved forward to ontology-based thesaurus management and semantic web technologies conformant to W3C recommendations as well as latest trends in thesaurus standards. However, EuroVoc represents a wide-coverage and faceted thesaurus built specifically for processing the documentary information of the EU institutions: the legal terminology is quite poor and limited to the legal fields belonging to the competence of EU.

The CELLAR repository provides semantic indexing, advanced search and data retrieval for multilingual resources to the information system of the Publications Office of the European Union information system. Resources and their Functional Requirements for Bibliographic Records (FRBR) embrace both the web of data perspective and the library or “bibliographic” data perspective [22]. Its new ontology development assumes that “the FRBR classes are collectors of resource metadata at their specific taxonomy level”, thus, allowing a direct constant access to the FRBR levels [22, p. 35]. This represents certainly an improvement over the existing model, as it enhances the accessibility of the OP multilingual documents. However, its scope is also limited to the vocabulary of EU documents.

The Legal Taxonomy Syllabus [6] is a multilevel, multilingual ontology that takes a comparative law perspective to the modeling of legal terms and concepts from EU Directives, helping to increase European terminological consistency. Syllabus is an

<sup>22</sup><http://eurovoc.europa.eu/drupal/>



open-access database linking European terms with national transposition law and also linking terms horizontally (i.e., between national legal orders).

LexALP [20] uses a technique defined for general lexical databases to achieve cross language interoperability between languages of the Alpine Convention. This multilingual legal information system combines three main components, i) a terminology data base, ii) a multilingual corpus, and iii) the relative bibliographic database. In this way the manually revised, elaborated and validated (harmonised) quadrilingual information on the legal terminology (i.e. complete terminological entries) will be closely interacting with a facility to dynamically search for additional contexts in a relevant set of legal texts in all languages and for all main legal systems involved.

The multilingual lexical database version of WordNet, EuroWordNet [19], compounds wordnets expressing lexica of 8 European languages. The wordnets are structured in terms of synsets (sets of synonymous words). Each synset in the monolingual WordNets is linked to the others by cross-lingual equivalence relations to the English synsets recorded by the Inter-Lingual-Index (ILI). The database can be used for monolingual and cross-lingual information retrieval. The LOIS [19] database is compatible with the EuroWordNet architecture, and forms an extension of the EWN semantic coverage into the legal domain. Within this framework, LOIS contributes to the creation of a European Legal WordNet.

#### 4. Conclusions

We have perceived a particular European policy deference towards rendering the copyright and related rights domain more accurate. We have framed some of the advantages for yielding a multi-lingual, multi-jurisdictional legal term bank published as linked data in this domain. Therefore our work presents an effort to achieve a technical and semantic interoperability among linguistic domain concepts.

However, creating a term bank of legal terms is a time-consuming task where expertise in the law of different countries is needed and even domain-specific terminologies require a considerable effort. Legal terminologies, legal concepts and legal knowledge are not synonymous.

Several problems might be raised: (i) ISO standards secure the exchange of terminologies but do not manage the *legal value* of such terminologies, (ii) as said, Version 4.0 of Creative Commons aims at a neutral text, capable of fitting every legal system, but nothing prevents legal operators (e.g. judges) to offer different interpretations of general concepts at the jurisdiction-specific level; (iii) the term “bank” related to other linked resources such as *dbpedia* or *lexvo* entails a more careful examination of this kind of relationship, as the *valence* (i.e. the number of edges incident to the vertex) of resources might not be equivalent, and in fact they are not.

Defining `owl:sameAs` relationships require a more careful examination of the functional entrenchment of legal sources [21], and a more extended comparative work of both legal and cultural systems [22]. This work will be completed in the next future with the annotations of a complete dataset of existing licenses.

**Acknowledgements.** This work is supported by the EU FP7 LIDER project (FP7 – 610782), by DER2012-39492-C02-01 CROWDSOURCING, and by the fellowship 520250-1-2011-1-IT-ERASMUNDUS EMJD.

## References

- [1] ISO 30042:2008 (2008). Systems to manage terminology, knowledge and content – TermBase eXchange.
- [2] McCrae, J., Spohr, D., and Cimiano, P. (2011). Linking lexical resources and ontologies on the semantic web with lemon. In *The Semantic Web: Research and Applications - 8th Extended Semantic Web*
- [3] Rodriguez-Doncel, V., Villata, S., and Gomez-Perez, A. (2014). A dataset of RDF licenses. In Hoekstra, R., editor, *Proceedings of the 27th International Conference on Legal Knowledge and Information System*, pp.–189.
- [4] Cimiano, P., McCrae, J., Rodriguez-Doncel, V., Gornostay, A. Gomez-Perez and B. Simoneit. *Linked Terminology: Applying Linked Data Principles to Terminological Resources*, in *Proc. of the 4th Biennial Conference on Electronic Lexicography* (to appear 2015)
- [5] Cabrio, E., Aprosio, A. P., and Villata, S. (2014). These are your rights: A natural language processing approach to automated RDF licenses generation. In *The Semantic Web: Trends and Challenges*, pp. 255–269. Springer
- [6] G. Ajani, G. Boella, L. Lesmo, A. Mazzei, and P. Rossi. Terminological and ontological analysis of european directives: multilinguism in law. In *11th Int. Conf. on Artificial Intelligence and Law (ICAIL)*, pp. 43–48, 2007.
- [7] G. Ajani; G. Boella; M. Martin; A. Mazzei; D. P. Radicioni; P. Rossi, *Legal Taxonomy Syllabus 2.0*, in "3rd Workshop on Legal Ontologies and Artificial Intelligence Techniques joint with 2nd W. on Semantic Processing of Legal Texts", June 8th, 2009, Barcelona, Spain.
- [8] Nakamura, Makoto, Yasuhiro Ogawa, and Katsuhiko Toyama. "Extraction of Legal Definitions and Their Explanations with Accessible Citations." P. Casanovas, U. Pagallo, M. Palmirani, G. Sartor, *AI Approaches to the Complexity of Legal Systems, AICOL-V*, . Springer Berlin Heidelberg, 2014. 157-171.
- [9] L. Lesmo, G. Boella, A. Mazzei (2006); *Multilingual Conceptual Dictionaries Based on Ontologies: Analytical Tools and Case Studies*. In *Proc. Of V Legislative XML Workshop*, pp. 1-14, EPAP
- [10] J. McCrae, D. Spohr, P. Cimiano (2011). Linking lexical resources and ontologies on the semantic web with lemon. In *Proc. of the 8th extended semantic web conference*, pp. 245-259, G. Antoniou et al. (Eds.), Vol. Part I. Springer-Verlag
- [11] V. Lyding, E. Chiocchetti, G. Sérasset, and F. Brunet-Manquat(2006). The LexALP information system: term bank and corpus for multilingual legal terminology consolidated. In *Proc. of the Workshop on Multilingual Language Resources and Interoperability (MLRI '06)*. Association for Computational Linguistics, Stroudsburg, PA, USA, 25-3
- [12] Kea European Affairs (2010) Study concerning Multi-territory licensing for the online distribution of audiovisual works in the European Union. Final Report Prepared for the European Commission, DG Information Society and Media.
- [13] Francesconi, Enrico; Spinosa, Pierluigi; Tiscornia, Daniela. A linguistic-ontological support for multilingual legislative drafting: the DALOS Project, P. Casanovas et al. (Ed.) *LOAIT June 4th, 2007*, Stanford University, Stanford, CA, USA. *CEUR-WS 321, 2008*: 103-11.
- [14] Poiares Pessoa Maduro, Luis Miguel, *Interpreting European Law: Judicial Adjudication in a Context of Constitutional Pluralism*, *European Journal of Legal Studies*, 2007, 1, 2.
- [15] D. Hanf and E. Muir, *Le droit de l'Union européenne et le multilinguisme*, in D. Hanf, E. Muir et K. Malacek (eds), *Langue et construction européenne*(Cahiers du Collège d'Europe, Bruxelles, 2010) at 23.
- [16] Casellas, Núria, *Linked Legal Data: A SKOS Vocabulary for the Code of Federal Regulations*, available at: [http://www.semantic-web-journal.net/system/files/swj311\\_2.pdf](http://www.semantic-web-journal.net/system/files/swj311_2.pdf)
- [17] Sartor, Giovanni, *Legislative Information and the Web, Legislative XML for the Semantic Web, Principles, Models, Standards for Document Management*, Sartor, G., Palmirani, M., Francesconi, E., Biasiotti, M.A. (Eds.), *Law, Governance and Technology Series*, Springer, 4, 2011.
- [18] Kemp, Richard, *Legal aspects of managing Big Data*, *Computer Law & Security Review*, Volume 30, Issue 5, Pages 463-614 (October 2014).
- [19] Peters W., Sagri M.T., Tiscornia D., (2007), *The structuring of legal knowledge in LOIS*, *Artificial Intelligence and Law*, 15:117-135.
- [20] Lyding, Verena, et al., *The LexALP Information System: Term Bank and Corpus for Multilingual Legal Terminology Consolidated*, *Proceedings of the Workshop on Multilingual Language Resources and Interoperability*, pages 25–31, Sydney, July 2006.
- [21] Casanovas, Pompeu, Núria Casellas, and Joan-Josep Vallbé. "Empirically Grounded Developments of Legal Ontologies: A Socio-Legal Perspective." P. Casanovas, G. Sartor, M.A. Biasiotti, M. Fernández-Barrera (Eds) *Approaches to Legal Ontologies*. Springer Netherlands, 2011. 49-67.
- [22] Francesconi, Enrico, Marc W. Küster, Patrick Gratz, and Sebastian Thelen. "The Ontology-Based Approach of the Publications Office of the EU for Document Accessibility and Open Data Services." In *Electronic Government and the Information Systems Perspective*, pp. 29-39. Springer International Publishing, 2015.