

# Multilingual Conceptual Dictionaries Based on Ontologies: Analytical Tools and Case Studies

Leonardo Lesmo<sup>o</sup>, Guido Boella<sup>o</sup>, Alessandro Mazzei<sup>o</sup>, Piercarlo Rossi\*

<sup>o</sup>*Dipartimento di Informatica, Università di Torino - Italy - {lesmo,guido,mazzei}@di.unito.it*

\**Dipartimento di Scienze Giuridiche, Università di Torino - Italy - p.rouge@inwind.it*

**Abstract.** This paper describes the philosophy behind our tool called “Legal Taxonomy Syllabus”, the analytical instruments it provides and some case studies. The Legal Taxonomy Syllabus is an ontology based tool designed to annotate and recover multi-lingua legal information and build conceptual dictionaries. The Legal Taxonomy Syllabus allows to build legal dictionaries in a bottom up fashion starting from the annotation of legal terms by lawyers and to let ontologist refine the resulting taxonomies of concepts. The Legal Taxonomy Syllabus and its analytical tools provide help to lawyers to study the peculiarities of European Union Directives concerning the polysemy of legal terms, and the terminological and conceptual misalignment. By means of two case studies we show how the Legal Taxonomy Syllabus can help the process of drafting and translating the Directives.

**Keywords:** Multilingual Legal Ontologies, dictionaries, terminologies

## 1. Introduction

The European union each year produces a large number of Union Directives (EUD), which are translated in each of the community languages. The EUD are sets of norms that have to be implemented by the national legislations. The implementation of a EUD however can not correspond to the straight transposition in a national law. An EUD is subject to further interpretation, and this process can lead to unexpected results. Comparative Law has studied in details the problematics concerning EUD and their complexities.

Based on this research, in this paper, we describe the tool for building multilingual conceptual dictionaries we developed for representing and analysing the terminology and concepts used in EUD.

The main assumptions of our methodology, motivated by studies in comparative law, are the following ones:

- Terms and concepts must be distinguished; for this purpose, we use lightweight ontologies.

- Each national legislation may refer to a distinct legal ontology. Furthermore, we distinguish the ontology which is implicitly defined by EUD, the EU level. We do not assume that the transposition of a EUD introduces automatically in a national ontology the same concepts present at the EU level.
- Corresponding concepts at the EU level and at the national level can be described by different terms in the same national language.

The main features of our tool for building multilingual conceptual dictionaries are:

- The system has been designed to be web based, and, thus, it is distributed; in this way, different teams can work on different languages at the same time from different places.
- The system provides different interfaces to different class of users. We assume that people with different expertise contribute to the construction of the dictionary. The first class is composed by legal experts which annotate the terms and relate them to concepts by using basic semantic relations. The second class refines the work of legal experts by adding more complex relations among the concepts
- The system allows to visualize and navigate the ontologies underlying the terms at both the European and national levels, and the relations among them.
- The system allows to mine the ontological data in order to find peculiar situations and possible inconsistencies in the use of terms between the EU level and the national ones.

In this paper, we show how the tool is used to build a dictionary of consumer law.

The structure of this paper is the following one. In Section 2 we stress two main problems which comparative law has raised concerning EUD and their transpositions. In Section 3 we describe how the methodology of the Legal Taxonomy Syllabus allows to cope with these problems In Section 4 we describe how the conceptual dictionary is built in a bottom-up fashion, by distinguishing the role of legal experts from the one of ontology engineers. Finally, in Section 5, we present which analyses can be performed on the Legal Taxonomy Syllabus and two cases studies concerning translation and drafting on EUD.

## 2. Terminological and conceptual misalignment

Comparative law has identified two key points in dealing with EUD, which makes more difficult dealing with the polysemy of legal terms: the *terminological* and *conceptual misalignment*.

In the case of EUD (usually adopted for harmonising the laws of the Member States), the terminological matter is complicated by their necessity to be implemented by the national legislations. In order to have a precise transposition in a national law, a Directive may be subject to further interpretation. Thus, a same *legal concept* can be expressed in different ways in a Directive and in its implementing national law. The same legal concept in some language can be expressed in a different way in a EUD and in the national law implementing it. As a consequence we have a terminological misalignment. For example, the concept corresponding to the word *reasonably* in English, is translated into Italian as *ragionevolmente* in the EUD, and as *con ordinaria diligenza* in the transposition law.

In the EUD transposition laws a further problem arises from the different national *legal doctrines*. A legal concept expressed in an EUD may not be present in a national legal system. In this case we can talk about a conceptual misalignment. To make sense for the national lawyers' expectancies, the European legal terms have not only to be translated in a sound national terminology, but they need to be correctly detected when their meanings are to refer to EU legal concepts or when their meanings are similar to concepts which are known in the Member states. Consequently, the transposition of European law in the parochial legal framework of each Member state can lead to a set of distinct national legal doctrines, that are all different from the European one. In case of consumer contracts (like those concluded by the means of distance communication as in Directive 97/7/EC, Art. 4.2), the notion to provide in a *clear and comprehensible manner* some elements of the contract by the professionals to the consumers represents a specification of the information duties which are a pivotal principle of EU law. Despite the pairs of translation in the language versions of EU Directives (i.e., *klar und verständlich* in German - *clear and comprehensible* in English - *chiaro e comprensibile* in Italian), each legal term, when transposed in the national legal orders, is influenced by the conceptual filters of the lawyers' domestic legal thinking. So, *klar und verständlich* in the German systems is considered by the German commentators referring to three different legal concepts: 1) the print or the writing of the information must be clear and legible (*gestaltung der information*), 2) the information must be intelligible by the consumer (*formulierung der information*), 3) the language of the information

must be the national of consumer (*sprache der information*). In Italy, the judiciary tend to control more the formal features of the concepts 1 and 3, and less concept 2, while in England the main role has been played by the second concept, though considered as plain style of language (not legal technical jargon) thanks to the historical influences of plain English movement in that country.

### 3. The methodology of the Legal Taxonomy Syllabus

To manage properly terminological and conceptual misalignment we distinguish in the Legal Taxonomy Syllabus project the notion of legal term from the notion of legal concept and build a systematic classification based on this distinction. The basic idea is that the basic conceptual backbone consists in a taxonomy of concepts (ontology) to which the terms can refer to express their meaning. One of the main points to keep in mind is that we do not assume the existence of a single taxonomy covering all languages. In fact, it has been convincingly argued that the different national systems may organize the concepts in different ways. For instance, the term *contract* corresponds to different concepts in common law and civil law, where it has the meaning of *bargain* and *agreement*, respectively (Sacco [1999], Pozzo [2003]). In most complex instances, there are no homologous between terms-concepts such as *frutto civile* (legal fruit) and *income*, but respectively civil law and common law systems can achieve functionally same operational rules thanks to the functioning of the entire taxonomy of national legal concepts (Graziadei [2004]). Consequently, the Legal Taxonomy Syllabus includes different ontologies, one for each involved national language plus one for the language of EU documents. Each language-specific ontology is related via a set of *association* links to the EU concepts, as shown in Fig. 1.

Although this picture is conform to intuition, in Legal Taxonomy Syllabus it had to be enhanced in two directions. First, it must be observed that the various national ontologies have a reference language. This is not the case for the EU ontology. For instance, a given term in German could refer either to a concept in the UK ontology or to a concept in the EU ontology. In the first case, the term is used for referring to a concept in the national UK legal system, whilst in the second one, it is used to refer to a concept used in the European directives. This is one of the main advantages of Legal Taxonomy Syllabus. For example *klar und verständlich* could refer both to concept **Ger-379** (a concept in the German Ontology) and to concept **EU-882** (a concept in the European ontology). This is the Legal Taxonomy Syllabus solution for facing the

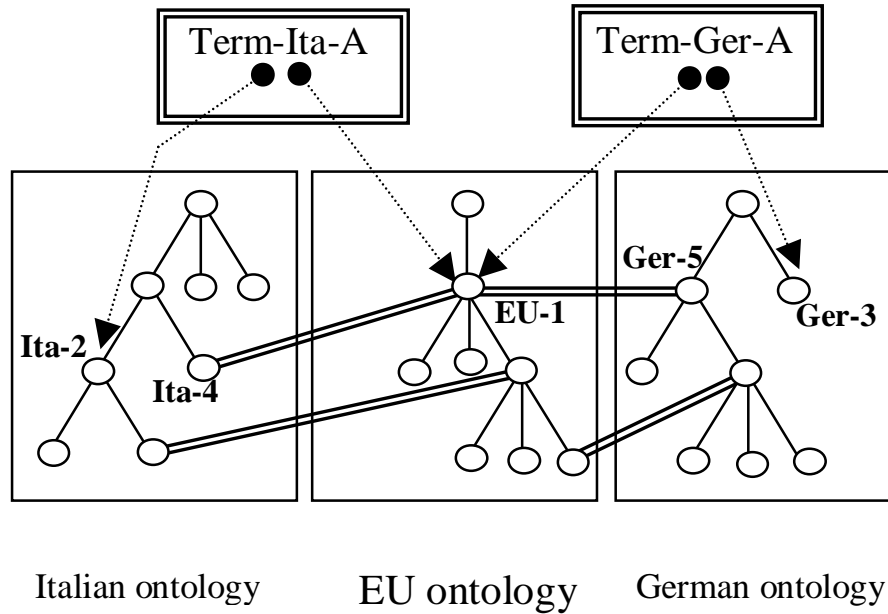


Figure 1. Relationship between ontologies and terms. The thick arcs represent the inter-ontology “association” link.

possibility of a correspondence only partial between the meaning a term has in the national system and the meaning of the same term in the translation of a EU directive. This feature enables Legal Taxonomy Syllabus to be more precise about what “translation” means. It puts at disposal a way for asserting that two terms are the translation of each other, but just in case those terms have been used in the translation of an EU directive: within Legal Taxonomy Syllabus, we can talk about direct EU-translations of terms, but only about indirect national-system translations of terms. The situation enforced in Legal Taxonomy Syllabus is depicted in Fig. 1, where it is represented that: The Italian term *Term-Ita-A* and the German term *Term-Ger-A* have been used as corresponding terms in the translation of an EU directive, as shown by the fact that both of them refer to the same EU-concept EU-1. In the Italian legal system, *Term-Ita-A* has the meaning Ita-2. In the German legal system, *Term-Ger-A* has the meaning Ger-3. The EU translations of the directive is correct insofar no terms exist in Italian and German that characterize precisely the concept EU-1 in the two languages(i.e.,

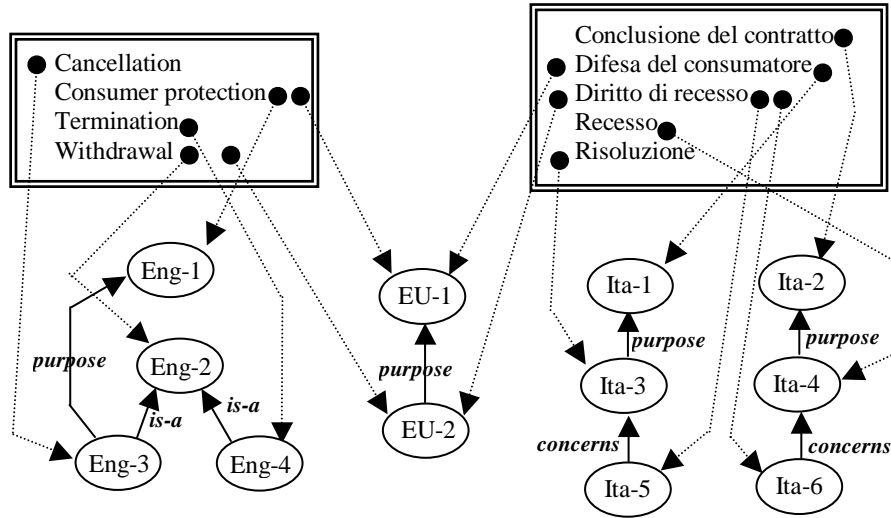


Figure 2. An example of interconnections among terms.

the “associated” concepts *Ita-4* and *Ger-5* have no corresponding legal terms). A practical example of such a situation is reported in Fig. 2, where we can see that the ontologies include different types of arcs. Beyond the usual *is-a* (linking a category to its supercategory), there are also a *purpose* arc, which is self-explanatory, and *concerns*, which refers to a general relatedness. The dotted arcs represent the reference from terms to concepts. Some terms have links both to a National ontology and to the EU Ontology (In particular, *withdrawal* vs. *diritto di recesso* and *difesa del consumatore* vs. *consumer protection*).

The last item above is especially relevant: note that this configuration of arcs specifies that: 1) *withdrawal* and *diritto di recesso* have been used as equivalent terms (concept EU-2) in some European Directives (e.g., Directive 90/314/EEC). 2) In that context, the term involved an act having as purpose the some kind of protection of the consumer. 3) The terms used for referring to the latter are *consumer protection* in English and *difesa del consumatore* in Italian. 4) In the British legal system, however, not all *withdrawals* have this goal, but only a subtype of them, to which the code refers to as *cancellation* (concept Eng-3). 5) In the Italian legal system, the term *diritto di recesso* is ambiguous, since it can be used with reference either to something concerning the *risoluzione* (concept Ita-3), or to something concerning the *recesso* proper (concept Ita-4).

#### 4. Building the Legal Taxonomy Syllabus knowledge base

The Legal Taxonomy Syllabus aims at building an ontology based representation of legal knowledge, but it differs from other proposals in the ontology field. The underlying philosophy of the project is to affirm the difference between knowledge in the law field and in other disciplines. First of all, legal knowledge is specific of national entities or institutional entities like the EU. So from the start, it is open to differences. Moreover, it is based on different philosophical traditions, and open to a continuous reinterpretation process.

Finally, we accept the idea of several scholars (Sacco [1987], Vanderlinden [1995], De Groot [1999]) that there is more in legal knowledge than the knowledge contained in statutory rules, since also case law, jurisprudence must be taken into account.

For these reasons, it is difficult to adopt a traditional top-down approach to the development of legal ontologies (see e.g. Visser and Bench-Capon [1998] for a survey of ontologies for the law).

The top-down approach works well for the topmost level, where the basic conceptual primitives are precisely defined (concept, relation, role, qualia, processes, etc.), and the representation instruments are put at disposal of those who build the ontology. Moreover, this level can be based on insights coming from cognitive science, a fact which assures the generality of the proposed primitives (e.g., the DOLCE approach of Gangemi et al. [2002]).

However, when the core ontology level is considered, it becomes more difficult to proceed in a top down fashion. The need to deal with a long tradition of studies which proceed in a parallel way in different countries and following different philosophical traditions makes it unlikely to try to achieve from the beginning an agreement on a common core ontology of the basic legal concepts (see the example of contract in Section 3). This is instead possible in other disciplines, e.g., biological data or specific domains related to computer science.

To avoid the risk that the knowledge engineers do not take into account the interpretation process of the legal specialists on the real multilingual data, we use a different bottom up approach. Many top down ontologies aim at modelling the legal code but not the legal *doctrine*, that is the work of interpretation and re-elaboration of the legal code which is fundamental for transposing EUD into national laws. In the development of the ontologies described in the previous section, we used a two-step procedure.

As a first step, terms are collected in a database together with the legal sources where they appear, and the underlying concepts are identified. In this phase, polysemy is identified, different meanings are

separated. Moreover, terms at the European level are associated to their transpositions in the national languages. As a second step, for each different ontology (i.e., each specific language ontology and the general EU ontology), the set of concepts is organized in an ontology which can be different for different national traditions.

At the end of these two steps, the result is a light-weight ontology rather than an axiomatic one. Only relations among concepts are identified without introducing restrictions and axioms. The function of these ontologies is to compare the taxonomic structure in the different legislations, to provide a form of intelligent indexing and to draw new legal conclusions.

On the basis of this two step procedure, we introduce different interfaces for the two kinds of users for the Legal Taxonomy Syllabus. The first kind of users (“lawyer”) is allowed to collect the legal terms, identifying the concepts. Moreover these users can only employ basic ontological relations (e.g. “IS-A” that allows to build a taxonomy, “near synonym”) to connect the various concepts. We stress that this process regards each distinct national ontology corresponding to the various languages, as well as the European ontology that is common to the various languages.

The second kind of users (“ontologist”) is allowed by the interface to rearrange the concepts inserted into the system with the aim to take into account the national legal doctrine. These users can define and employ complex ontological relations in order to clarify some peculiar relation among legal concept deriving from national legal doctrine.

## **5. Can the Legal Taxonomy Syllabus improve the EUD production? Analytical tools and case studies**

The Legal Taxonomy Syllabus allows several ways for studying the data and identify critical issues concerning the use of EUD terms in the European and national context. We present in this section the analytical tools of the Legal Taxonomy Syllabus and two case studies.

First of all, the ontology of the European level is shared between the national languages. Hence, for each concept in the ontology, it should be possible to find (at least) a term describing it in each national language. This is, however, not guaranteed. See the case study in Section 5.1 below where this specific issue is used for the translation of EUD.

Mind that the lack of a term in a certain language for a certain concept does not necessarily mean a problem, but only that the work does not proceed in all languages at the same time.

Second, terms at the European level are transposed in the national level when a Directive is implemented. The European term, however, is not transposed directly at the national level: first of all, an equivalent concept can already exist at the national level but described by a different term, or the European term can be already used for other concepts at the national level, etc., as discussed in Section 2. Thus, the Legal Taxonomy Syllabus offers a method for listing the European terms which are transposed in a different way at the national level, European terms which already exist at the national level with other meanings, or even terms which are transposed at the national level into (different) terms which exists also at the European level with a different meaning.

Third, the Legal Taxonomy Syllabus can show ontological discrepancies between the European and national level. Given a pair of concepts at the European level and the pair of corresponding transpositions, it is possible that they are not related by the same ontological relation (IS-A, etc.) in the European and national level (see the example of Figure 2).

Finally, it is possible to use the Legal Taxonomy Syllabus to translate terms in different national systems via the concepts which they are transposition of at the European level. For instance suppose that we want to translate the legal term *credito al consumo* from Italian to German. In the Legal Taxonomy Syllabus *credito al consumo* is associated to the national meaning *ita-175*. We find that *ita-175* is the transposition of the European meaning *EU-26* (*contratto di credito*). *EU-26* is associated to the German legal term *Kreditvertrag* at European level. Again, we find that the national German transposition of *EU-26* corresponds to the national meaning *ger-32* that is associated with the national legal term *Darlehensvertrag*. Then, by using the European ontology, we can translate the Italian legal term *credito al consumo* into the German legal term *Darlehensvertrag*.

Now we describe how the Legal Taxonomy Syllabus can improve the EUD production. The EUD production has two basic processes, e.g. drafting, translation. We believe that the Legal Taxonomy Syllabus can help in working on these two process, and in the implementation of directives as well. Below we present two case studies which explain how the Legal Taxonomy Syllabus can help in translation and drafting.

### 5.1. TRANSLATION

By searching in the European ontology of the Legal Taxonomy Syllabus, we find that the legal concept *EU-74* correspond to the legal words *comerciante*, *commerçant*, *trader*, *commerciant* in Spanish, French,

English and Italian respectively (Fig. 3). In other words, in four language versions several EC provisions on consumer law currently employ the hyperonym *trader*. Moreover we find that there is no equivalent word for German. This consideration can help translators that can use a new term to denote the legal concept EU-74.

From the point of view, at national level, both the Italian and the German legislator employ only one term in the general definitions instead of the different ones provided by the EC Directives (Professionista, in Decreto Legislativo 6 settembre 2005, n. 206, Art. 3.1.c., Unternehmer, in BGB 14). This is the result of consolidating consumer law in accordance with the national private law.

However, the effectiveness of such solution remains questioned. In German legislation one general definition today covers all the cases of trader, seller, final seller, and so forth, but the rules are not clearly unified. In Italian legislation the term *professionista* coexists with other specific consumer provisions relating to *venditore* (seller) without disambiguating the different uses. Moreover, the term *professionista* also refers to one diverse concept, arising out from the rules on *professioni liberali* (Art. 2229 ff.).

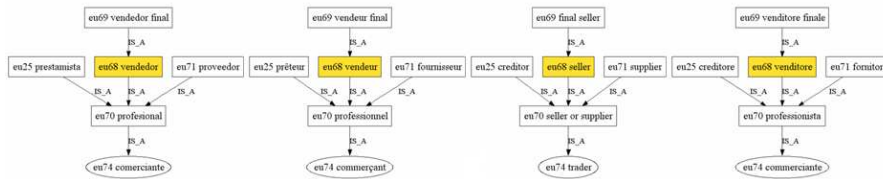


Figure 3. A fragment of the EU level ontology in different languages

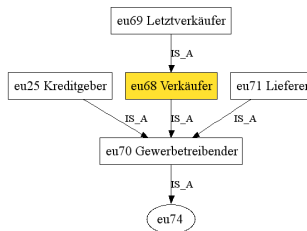


Figure 4. The same fragment in German is not completely translated

## 5.2. DRAFTING

The EC secondary law on contract law is assumed to be partially incoherent: in addition to the well known reasons regarding substantive law (competing rules and policies), the terminological-conceptual matters should not be underestimated. Starting from this alleged situation, the European Commission has promoted the recasting of consumer law provisions. The Legal Taxonomy Syllabus could be useful to highlight some terminological inconsistencies of EC law in order to (re)draft EC consumer law.

Any attempt to recast the consumer law provisions should consider carefully that only an ontological mapping of all the relations among concepts might avoid inconsistent definition.

One example may be drawn out from Directive 84/450/EEC on misleading advertising (and Directive 97/55/EC concerning comparative advertising, amending that Directive). In all language versions *advertising* is the hyperonym of *misleading advertising* and *comparative advertising* (see Fig. 5) Anyway the EUD drafters have to consider carefully the ontological status of these two terms for two distinct reasons:

1. *Advertising* is not only the hyperonym of *misleading advertising*, but also another concept, although correlated, concerning the information duties to be provided to consumers (i.e. consumer credit, sale of goods).
2. *Misleading advertising* is strictly connected to the *misleading commercial practices* that represents a broader concept including *comparative advertising* in some circumstances (as shown in Directive 2005/29/EC Art. 6.2 lit. a)

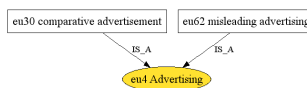


Figure 5. The EU level ontology on advertising

## 6. Conclusions

In this paper we discuss some aspects of the Legal Taxonomy Syllabus, a tool for building multilingual conceptual dictionaries for the EU law. The tool is based on lightweight ontologies to allow a distinction of concepts from terms. Distinct ontologies are built at the EU level and for each national language, to deal with polysemy and terminological and conceptual misalignment. The Legal Taxonomy Syllabus has been extended by adding analytical tools which help the users to identify complex situations and misleading translations of terms when EUD are implemented in the national legislations. Thus, the Legal Taxonomy Syllabus can be fruitfully used in the process of drafting, translating and implementing EUD.

Future work is to study how the Legal Taxonomy Syllabus can be used as a thesaurus for the EUD. Even if the current model the domain is limited to consumer law, the Legal Taxonomy Syllabus provides a general purpose tool for dealing with EU law. As discussed by Rossi [2005], the current instruments provided by EurLex have several limitations, in particular, the Eurovoc thesaurus is based on the EU Treatise structure rather than on a taxonomy of law. The Legal Taxonomy Syllabus could provide the basis for a thesaurus based on ontologies built by the lawyers while populating the legal dictionaries.

## References

- G.R. De Groot. Das übersetzen juristischer terminologie. In De Groot and R. Schulze, editors, *Recht und bersetzen*, page 11 ff, Baden-Baden, 1999. Nomos.
- A. Gangemi, N. Guarino, C. Masolo, A. Oltramari, and L. Schneider. Sweetening ontologies with dolce. In *Proc. EKAW 2002*, Siguenza (SP), 2002.
- M. Graziadei. Tuttifrutti. In P. Birks and A. Pretto, editors, *Themes in Comparative Law*, pages -. Oxford University Press, 2004.
- B. Pozzo. Harmonisation of european contract law and the need of creating a common terminology. *European Review of Private Law*, 6:754–767, 2003.
- Lesmo Graziadei Boella Mazzei Rossi. The next EurLex: what should be done for the needs of lawyers belonging to different national legal systems? In *Procs. of Jurix EU-Workshop*, 2005.

- R. Sacco. La traduction juridique : un point de vue italien. *Les Cahiers du Droit*, 28(4):845 ff, 1987.
- R. Sacco. Contract. *European Review of Private Law*, 2:237–240, 1999.
- J. Vanderlinden. *Das Übersetzen juristischer Terminologie*. Story-Scientia, Bruxelles, 1995.
- P.R.S. Visser and T.J.M. Bench-Capon. A comparison of four ontologies for the design of legal knowledge systems. *Artificial Intelligence and Law*, 6:27–57, 1998.

