European Legal Taxonomy Syllabus: A multi-lingual, multi-level ontology framework to untangle the web of European legal terminology *

Gianmaria Ajani [a], Guido Boella [a,**], Luigi di Caro [a], Livio Robaldo [c], Llio Humphreys [a], Sabrina Praduroux [a], Piercarlo Rossi [b] and Andrea Violato [d],

[a] *Università degli Studi di Torino, Torino, ITALY*

E-mail: gianmaria.ajani@unito.it, {guido,dicaro,humphreys,praduroux}@di.unito.it

[b] *Università del Piemonte Orientale, Novara, ITALY E-mail: piercarlo.rossi@unipmn.it*

[c] *University of Luxembourg, Interdisciplinary Centre for Security, Reliability and Trust (SnT),*

Luxembourg, E-mail:livio.robaldo@uni.lu

[d] *Nomotika srl, Torino, ITALY E-mail: andrea.violato@nomotika.it*

Abstract

The final publication is available at IOS Press through `http://dx.doi.org/10.3233/AO-170174`.

This paper describes a new concept of legal ontology together with an ontology development tool, called European Legal Taxonomy Syllabus (ELTS). The tool is used to model the legal terminology created by the Uniform Terminology project on EU consumer protection law as an ontology.

ELTS is not a formal ontology in the standard sense, i.e., an axiomatic ontology formalized, for instance, in description logic. Rather, it is a *lightweight* ontology, i.e. a knowledge base storing low-level legal concepts, connected via low-level semantic relations, and related to linguistic patterns that denote legal concepts in several languages spoken in the European Union (EU). In other words, ELTS is a multi-lingual and multi-jurisdictional terminological vocabulary enriched with concepts denoted by vocabulary entries, with semantic relations between different concepts. The choice of such an architecture is based on past studies in comparative law and is motivated by the need to reveal the differences between national systems within the EU. Past literature in comparative law highlights that axiomatic ontologies freeze legal knowledge in an unreal steadiness, i.e., they render it disconnected from legal practice. Much more flexibility is needed to make the knowledge base acceptable to legal practitioners.

ELTS was developed together with legal practitioners on the basis of the comparative view of European law. The ontology framework is designed to help professionals study the meaning of national and European legal terms and how they inter-relate in the transposition of European Directives into national laws. The structure and user interface of ELTS is suitable for building multi-lingual, multi-jurisdictional legal ontologies in a bottom-up and collaborative manner, starting from the description of legal terms by legal experts. It also takes into account the interpretation of norms, the dynamic character of norms and the contextual character of legal concepts in that they are linked to their legal sources (legislation, case law and doctrine).

Keywords: ontology, multi-lingual, legal ontology, EU law

1. Introduction

Ontology as a branch of philosophy is described by Smith (2008, p.155) as "the science of what is, of the kinds of objects, properties, events, processes and relations in every area of reality". Ontologies that explicitly describe reality, or provide "an explicit specification of a conceptualization" (Gruber, 1993, p.1), are also an important area of computer science. The objective in computer science is to provide people, or more typically artificial agents, with structured and navigable knowledge about entities and their inter-relations. Ontologies can help people share and reference knowledge about concepts in general or specialist areas, in one or more languages. Ontologies can also be used for information technology tasks such as semantic searches, interoperability between systems, or to facilitate reasoning and problem solving in artificial intelligence.

Ontologies can work very well in certain scenarios. For instance, ontologies have proven to be useful in these scientific and medical domains for overcoming terminological differences. For instance, the Gene Ontology[1] was developed to address a common problem:

> [I]f you were searching for new targets for antibiotics, you might want to find all the gene products that are involved in bacterial protein synthesis, and that have significantly different sequences or structures from those in humans. If one database describes these molecules as being involved in 'translation', whereas another uses the phrase 'protein synthesis', it will be difficult for you - and even harder for a computer - to find functionally equivalent terms.

In our view, the accuracy and usefulness of the Gene Ontology is also due to the fact that the medical/biological domain contains clear, explicit and unambiguous conceptualisations that are shared by most scientists working in the field. Not all domains are like this, and the legal domain is an extreme case where the precise meaning of concepts vary in different contexts and are difficult to pin down. For instance, the concept of "offense" in the legal domain is highly debatable, and many incompatible (and informal) definitions of that concept have been proposed by jurists.

The peculiarities of the legal domain are many, and we will examine them in detail in Section 2. For now, suffice to say that laws are written in "legalese" - a domain-specific sublanguage that inherits all the expressivity and ambiguity of natural language with additional terms of its own that are often obscure, subject to changes over time, contextually defined, ill-defined (i.e., some 'obvious' elements are not made explicit), subject to interpretation to deal with their vagueness, defined in incompatible ways in different legal sources, or difficult to translate into different languages. As Peller (1985, p.1182) stated: "legal discourse can never escape its own textuality".

For this reason, building axiomatic ontologies for the legal domain would appear to be a difficult, if not impossible, task, especially where the ontology creators are computer scientists with little expertise in law. Most likely, the axioms defining the ontological concepts would not be shared by a substantial subset of the members of the scientific community. Moreover, axiomatic ontologies would freeze legal knowledge in an unreal steadiness, making them difficult to be accepted by legal practitioners. In our view, the observations of Wahlgren (1994, p.80-81) remain true today:

> [S]everal contributions in the field of AI and law appear to be incompatible with respect to the understanding of the law and legal work ... It is no secret that very few systems have been accepted by the legal community. The problems, however, are not primarily of a technical nature. With little doubt, the difficulties are more closely related to a too shallow understanding of the requirements of the domain taken as whole.

Most legal ontologies of necessity describe a socially constructed reality (Searle, 1995), rather than natural reality or artifacts. In the legal domain there is no shared "objective" reality available for the creators of the ontology to conceptualize. Laws are (usually) postulated in vague terms in order to be applicable to a large variety of (specific) real contexts and situations, and their legal interpretation is

---

[1] http://www.geneontology.org

later established on a case-by-case basis by courts and legal doctrine, contributing to settle a kind of "de facto" legal culture around norms. Understanding legal terminology requires a deep understanding of legal culture and social values, a situation made more complex by the existence of several legal jurisdictions, each having their own set of concepts, as well as their own social, historical, and cultural characteristics, and where, due to globalisation and subsidiarity, concepts from one jurisdiction are sometimes adopted and adapted to other jurisdictions with different languages, creating multi-lingual terminological complexities. This is particulary evident in the context of the European Union, where 28 jurisdictions and 24 languages coexist, and where the legal language used in EU legislation is not always the same as the one used in the corresponding national state.

Nevertheless, legal ontologies are an active field of research because they could be useful in a range of different scenarios. Legal ontologies could help legal practitioners and scholars keep up to date with continuous changes in the law and understand legal sub-languages outside their own areas of expertise or jurisdiction. They could help legislators draft legislation with clarity and consistency. Moreoover, a multi-jurisdictional legal ontology could help show the inter-relationship between, for instance, national and European Union terms and foster harmonization, a need recognized, e.g., by the Mandelkern Group Report (Mandelkern, 2001), for the EU Commission. The group stressed the need for internal coherence and consistency in the use of EU legal terms, as well as external coherence - consistency in the transposition of legal concepts into national law. A multi-jurisdictional legal ontology could potentially help translators of EU legislation become aware of how their choice of terms will be interpreted in different jurisdictions and help drafters of national legislation ensure greater consistency in their use of terminology when transposing European legislation. Such an ontology could also serve as a useful tool in legal search, information retrieval, automatic translation, automated reasoning and regulatory compliance verification. It could help users find similar or related legislation, or compare the transposition of laws in national jurisdiction - useful for legal scholars in comparative law or lawyers who deal with cross-border issues and international financial institutions.

This paper outlines a different approach to legal ontologies. We don't build a legal ontology starting from the point of view of computer scientists; rather we look at how to extend ontology models starting from the point of view of legal practitioners, seeking to address their needs first. To achieve this objective, legal experts must be kept in the loop from the beginning, so that they can properly assess the different dimensions involved in the conceptualisation of legal terms in each different jurisdiction in the EU.

Several legal ontologies proposed in the literature had involved legal practitioners (cf. (Sartor et al., 2013)). An example is the Ontology of Professional Judicial Knowledge (OPJK) (Casanovas et al., 2007), which was modeled on the basis of a repository of professional judicial knowledge containing nearly 800 questions regarding daily practice.

However, the role of legal experts in these past works has been mostly devoted to the *content* of the ontology, while, in our approach, they mostly collaborate in the *modelling* of the ontology, i.e., the definition of an architecture that can both take into account all needed components and that can model these components at a level of expressivity acceptable from a practical point of view.

As such, the research question of this paper is how to build a multi-lingual, multi-level legal ontology that can help legal professionals, multi-national organisations and professionals involved in trans-national activities to manage the deep, complex and interconnected knowledge required for understanding laws in different jurisdictions within the EU. The sub-questions are the following:

– Given that a 'deep' understanding of legal systems is necessary to bring out explicit conceptualisations, which legal theoretic foundation can we give to our ontology framework, to enable legal experts to be involved in the design, population and updating of ontologies, so that the ontologies properly reflect the law in the EU context?
– Given that the predominant ontology framework of a static world where terms have clear unequivocal meanings is unsuitable for the legal domain, how can we design an ontology framework that enables legal experts to express the range of different conceptualisations of terminology and their relation to one another? In particular:

∗ legal terms are defined in many different pieces of legislation, and are further articulated in case law and legal scholarship. How can a legal ontology manage all these definitions, so that a specialist user can select the most appropriate definition for specific situations?
∗ how to deal with the complexity of the EU legal system, which is not only multi-lingual but also multi-level?
∗ laws change all the time, and so do legal definitions. How should legal ontologies deal with normative change?
∗ if there are several context-specific definitions of legal terms with a common thread, how can we create definitions in the ontology resulting from interpretation without discarding context-specific ones?

The methodology we used to address our research question was developed around four main pillars, each of which we considered to be essential to achieve our objectives:

– We used as a benchmark the knowledge contained in a detailed European-wide legal terminological analysis developed in the European project Uniform Terminology[2]. This project was developed and evaluated to properly handle the specificities of legal terminology in the domain of consumer law, considering phenomena that previous terminological vocabularies, ontologies, etc. cannot model, and putting legal interpretation and evolution of norms over time at the center of legal reasoning.
– We designed a (lightweight) legal ontology schema satisfying the requirements posed by the analysis of comparative law carried out in the context of the Uniform Terminology project. The lightweight ontology approach aims to simplify ontologies management in order to gain the acceptability of legal practitioners, and ensure their active participation in the building/updating of the knowledge base.
– We developed a web-based legal ontology tool conforming to the ontology schema. The tool is usable by legal practitioners who are not expert in existing ontology tools. Integrity constraints are enforced by the tool's interface and its database rather than made explicit in a declarative form.
– We created a multi-lingual legal ontology of European consumer law constructed with the above tool using data from the European project Uniform Terminology, in order to prove the viability of our novel approach and ontology schema.

We will use the general term European Legal Taxonomy Syllabus (ELTS) to refer to either the schema, the tool or the ontology when it is clear from the context which one we are referring to; otherwise, we will specify explicitly which dimension we are referring to.

The ELTS tool has two parts: a repository of legal concepts and a legal terminological database storing linguistic patterns in each of the supported languages. ELTS focusses on the concepts that are associated with the terminological entries via many-to-many relations. The temporal evolution of concepts is represented explicitly, searches are indexed by time, and comparisons can be made between linked concepts in the different languages. Most importantly, ELTS represents explicitly the conceptual misalignments between the different national systems or between the EU and the national levels.

European Legal Taxonomy Syllabus is a key component of the Eunomos system (Boella et al., 2016) which has been further developed as the Menslegis[3] regulatory compliance commercial service. Eunomos was also used in the EUCases project[4] to help find case law in several EU countries (see subsection 6). Furthermore, new research activities involving European Legal Taxonomy Syllabus and Eunomos are currently under development in the context of the ongoing projects "MIREL: MIning and REasoning with Legal texts"[5] and "ProLeMAS: PROcessing LEgal language in normative Multi-Agent Systems"[6].

The decision to create a *lightweight* (non-axiomatic) legal ontology was motivated by the need for high flexibility to take into account actual inconsistencies and conflicts in the law. As stated above, legal

---

[2]http://www.uniformterminology.unito.it
[3]http://www.nomotika.it/IT/MensLegis/Flyer, distributed by Nomotika srl.
[4]http://www.eucases.eu
[5]http://www.mirelproject.eu
[6]http://www.liviorobaldo.com/Prolemas.htm

terms are often vague and ill-defined, and subject to interpretations and changes over time. Consistency is not always considered to be a desirable property in the legal domain, which means that an effective formalization of conflicting resolution rules is still an open research problem[7]. On the other hand, such formalizations cannot be used by legal practitioners, who are the target users of our ontology.

This paper focuses only on modeling legal concepts related to legal terminology. An extension to model norms as specific concepts of their own is beyond the scope of this paper, but is described by Boella et al. (2012). Moreover, we do not discuss here the problem of connecting a legal ontology with top-level ontologies, although we acknowledge the desirability of such mappings for certain applications.

Section 2 discusses philosophical and linguistic issues that have impacted on the design of ELTS. Section 3 and section 4 describe the ontology schema and tool with examples from the Uniform Terminology project. Section 5 describe the interfaces, the technologies and also outlines technologies to be implemented into the system to help address the 'resource bottleneck' (Hepp, 2007) of creating and updating ontologies. Section 6 describes the recent developments of ELTS. Section 7 describes other legal ontologies of note and the similarities and differences in comparison with our approach. Conclusions end the article.

2. Vision

The way of the world is to make laws, but follow custom.
Charles de Montesquieu. Book 1, Essais, 1595

2.1. Comparative Law

To understand the difficulty of defining concepts from the legal domain in an ontology, it is useful to look at the legal process and how legal concepts interact with the social and political environment. Our perspective on laws and legal concepts is largely informed by insights from comparative law. Comparative law has emerged as a thorough approach to investigating and describing laws and legal systems, and the analysis that emanates from the investigation can help explain the meaning of legal terms as used by practitioners. One of the most influential comparative lawyers is Rodolfo Sacco. Starting from the premise that norms are not "legal flowers without stem or root" (Sacco, 1991, p.27), he identified factors that influence how laws are interpreted in different jurisdictions.

First, all legal systems have several *legal formants*, otherwise known as sources of law - codes and statutes, judicial decisions, legal scholarship and political ideologies:

> The civil lawyer may say that this rule comes, in principle, from the code; the common lawyer may say it comes from a particular statute or from judicial decisions; and yet they both will learn their law initially from the books of legal scholars (Sacco, 1991, p.22).

The importance of these legal formants vary considerably in different jurisdictions and different areas of law - case law is more important in France than in Italy, some areas of English law are subject to more statutes than others - although all these legal formants have some influence, whatever the official model of the law might say. The existence of multiple legal formants creates uncertainty, since they are rarely in complete harmony on a point of law. And yet, this does not usually stop the law from functioning.

Legal formants are not the only factors that influence legal interpretation. There are also invisible factors such as the beliefs or mentality of the interpreters, their social and cultural background. Sacco calls these factors *cryptotypes*. Such factors rarely need to be articulated in a mono-culture. Comparative law helps reveal hidden cryptotypes when a seemingly equivalent rule is interpreted in different ways in different legal jurisdictions, or when an implicit rule is made explicit in another legal system. Sacco cites as an

---

[7]In the context of the projects ProLeMAS and MIREL, and drawing from previous approaches to Natural Language Semantics, we are currently studying the development of flexible fine-grained logical representations that are able to handle conflict resolution and, more generally, the non-monotonic defeasible nature of legal text. Cf. section 6 below.

example the issue of whether an heir can transfer property before possessing it. Belgian interpreters of the Civil Code have deemed such transfers invalid, but the French upheld them. The discrepancy is explained by the fact that while the Code itself does not support such transfers, the old Roman law did, and the custom carried over into French law.

Sacco noted that it is quite common to find that not all legal rules are fully articulated. A *synecdoche* occurs when only part of a phenomenon is indicated when referring to the whole. He gives an example that the legal definition of contract in French law refers to the will of the parties without mentioning the need for the will to be declared or that there needs to be a good reason for the parties to declare their will and for the law to respect it. Filling in the gaps requires knowledge of the legal culture and custom. Unwritten rules are passed on from one generation of jurists to another.

Sacco claims that identification of legal formants, cryptotypes, and synecdoche were found "almost as a by-product" (Sacco, 1991, p.388) in comparing different legal systems, but has led to a more profound understanding of how the law functions than the legal philosophies of Kelsen, Hohfeld and others. The reason is that limiting the study of law to a single legal system leads one to ignore features that appear to be too 'obvious' to mention. Such features are not necessarily 'obvious' or common to all legal systems, and their discovery reveals the unwritten rules and values that underpin the law in different legal systems. The comparative approach can go beyond the letter of the law to find its true meaning (Sacco, 1991, p.16-17):

> An abstract idea finds concrete expression in a given legal language much as, in biology, a genotype or distinctive set of genes is expressed in the phenotype or outward form of a plant or animal. The jurist of an individual country studies the phenotype. The comparativist must study the genotype of which it is the expression.

Another important contribution of comparative law is the exploration of the interplay between legal formants, and the awareness that while the law constantly evolves, legal formants rarely move together in sync, so that conflicting valid interpretations are inevitable (Sacco, 1991, p.394). Sacco (1991, p.24) states:

> Comparison recognizes that the "legal formants" within a system are not always uniform and therefore contradiction is possible. The principle of non-contradiction, the fetish of municipal lawyers, loses all value in an historical perspective, and the comparative perspective is historical *par excellence*.

This is in contrast with Kelsen's conclusion that contradictory norms within the same legal system cannot exist, in line with a universalist view of law (Kelsen, 1992, p.112):

> If legal cognition encounters legal norms that contradict one another in content, it seeks, by interpreting their meaning, to resolve the contradiction as a mere pseudo-contradiction. If this effort fails, legal cognition disposes of the material to be interpreted, disposes of it as lacking in meaning altogether and therefore as non-existent in the legal sphere *qua* realm of meaning.

Sacco has had an immense influence on legal scholarship, most especially in his bottom-up approach to the analysis of legal concepts. For example, Graziadei (2004) analyses the gradual acceptance of interest as a return on capital in European countries achieved through different paths. There is no synonymy relation between terms-concepts such as *"frutto civile"* (legal fruit) and *"income"* from civil law and common law respectively, but these systems can achieve functionally similar operational rules thanks to the functioning of the entire taxonomy of national legal concepts. The study shows the influence of changes in religion, economy and philosophy on the evolution of these concepts in different civil and common law jurisdictions as well as the practical implications of conceptual differences.

In summary, comparative law suggests that there is more to norms than what is in legislation, that analysing different national systems gives a deeper understanding of law, and that differences between national systems can be revealed with a bottom-up approach. This view is an alternative to conventional top-down approaches to legal ontologies, which bear a close resemblance to Kelsen's doctrine of the unity of law:

> "As it is the task of natural science to describe its object - reality - in one system of natural laws, so it is the task of jurisprudence to comprehend all human law in one system of norms" (Kelsen, 1941).

The following two subsections develop further the Vision of comparative law that is the basis of our research, while focussing on the two aspects addressed in ELTS: (1) Legal interpretation and the evolution of norms, and (2) Multi-linguality and Multi-jurisdictionality.

2.2. Legal interpretation and the evolution of norms

As stated above, laws may be defined by statute, but are often re-elaborated several times via scholarly or judicial legal interpretation. Legal interpretation is a necessary part of the legal process, and established interpretation methodologies can modify or extend legal rules. Rawls (1955) stated that legal rules are of necessity more general than the myriad present and future scenarios to which such rules must be applied. Liebwald (2012) states that civil law countries often refer to Savigny's canons of interpretation: grammatical (a literal reading of the norm itself and nothing else), systematic (taking into account the domain or legal system in general), historical (based on the purpose of the norm as revealed in the preamble or preliminary discussions) and teleological (based on the 'independent will of the norm' or the will of the interpreter). Similar canons of interpretation exist in the Common Law tradition - the 'plain meaning rule' (which corresponds to the grammatical rule above), the 'mischief rule' (corresponding to the historical rule), and the 'golden rule' where a word's usual meaning can be disregarded to avoid an absurd result.

There are no firm rules as to when to use such canons of interpretation, which means that in practice, they are used at will to best serve the demands of justice or the values and prejudices of the interpreter.

Ambiguity and vagueness in language has been studied extensively by philosophers (Russel, 1923), linguists (Zhang, 1998) and cognitive scientists (Margolis and Laurence, 1999). All these issues have a concrete importance in the legal context, since the interpretation of language has implications for the people involved. Many researchers have faced these problems with specific regard to the legal domain, such as Endicott (2008), Poscher (2011), Soames (2012), among others.

Liebwald (2012) argues that although natural language is inherently imprecise in general, there is even more vagueness in legislative drafting and judgments, and that this vagueness is often intentional - to cover unexpected cases, to leave space for more specific rules in the future, or due to a reluctance to spell out explicit rules where there is political or social disagreement. While the law does contain determinate concepts such as 'age of consent' or speed limits, it is also full of indeterminate and malleable concepts such as 'good faith' or 'reasonable discretion'.

Liebwald (2012) described the evolution of the meaning of 'essence of marriage' as a case in point. The concept derives from The Marriage Act of 1938 in Germany, used to assess whether divorce or annulment is morally justified. During the Nazi era, marriages between Aryan and non-Aryan spouses, or between wives older than their husbands, were considered contrary to the 'essence of marriage'. The Act survived in West Germany and East Germany, but were interpreted quite differently. In West Germany, marriage was interpreted as an absolute, predetermined moral order, and therefore indissoluble. In East Germany, however, marriages were evaluated in terms of the spouses' ability to serve social ideals, particularly the ability to work. The Act has also survived in Austria, where the 'essence of marriage' is defined as the possibility of parenthood, whether or not the spouses want to, or are able to, have children. According to this definition therefore, same-sex couples cannot get married.

The courts are known to stretch the meaning of legal terms as much as possible in order to achieve the desired result. Each party involved in a trial argues for the most convenient and amenable interpretation to support their legal position. In their dialectic legal reasoning, the litigants will try to argue and convince the judge to accept their preferred interpretation. This means that legal drafting and, ultimately, judicial decision-making requires meticulous and attentive analysis of all possible readings of relevant norms.

The legal domain is constantly subject to evolving conceptualisations and neologisms. Legal sublanguages, despite their conservative style, show remarkable flexibility and dynamism in redefining new words and coining new ones when the need arises. Tiersma (2006) states that "the language of lawyers can sometimes be surprisingly creative and innovative. Lawyers are quick to coin a new word when their existing vocabulary is insufficient. Consider recent additions to the legal vocabulary, such as a *shrinkwrap* license (where a software user agrees to terms written in the software itself, or in a user's manual, merely by

opening the box), or a *clickwrap* license (where the user clicks on a box or an icon to indicate acceptance of the terms of the license)."

Our work does not extend to further analysis on this subject. Our contribution is rather to propose a computational system that helps experts find insights and relevant information so that they have the right keys to (possibly) make the "right" interpretation. For this reason, our work lies in the domain of *supporting tools* for the legal domain.

The practical conclusions we draw from the above discussion is that legal terms are inextricably linked to their historical, social and legal context, and that the meaning of legal terms depends strictly on other concepts, such as the purpose of the legislation in which the term was defined. All these dimensions influence the interpretation of the meaning of a legal term. A literal reading of laws is not enough to understand them, and generalised terminologies should be treated with caution. Moreover, norms, and thus the meaning of legal terms, is subject to continuous change, due to explicit or implicit modification of norms by different actors in the legal system.

2.3. Multi-linguality and Multi-jurisdictionality

Achieving shared conceptualisations of law is difficult in any legal system. The problem is confounded in Europe, which is increasingly governed by multiple jurisdictions - European, national and sometimes regional as well - and with many official languages. The last amendment to Regulation No. 1 of 15 April 1958 recognises twenty-four languages as having the status of official and working languages in European institutions. This poses a significant challenge, since each of the twenty-eight Member States that make up the European Community have their own cultural baggage that no one, let alone the Community legislature, is able to escape. The Sapir-Whorf hypothesis (Hoijer, 1954) maintains that it is impossible for a concept in one language to be imported wholesale into another language due to linguistic relativity. Nevertheless, the European Union steadfastly seeks to achieve 'harmonisation' of laws in whole sectors of diverse legal disciplines.

Harmonisation of EU law is a complicated matter. For Regulations, the implementation is automatically 'binding in its entirety and directly applicable in all Member States', while Directives are 'binding, as to the result to be achieved, upon each Member State to which it is addressed, but shall leave to the national authorities the choice of form and methods.'[8]

The procedure of creating European Union laws in a multi-lingual environment has its own complications. Although debates on legislation in the European Parliament and Council can take place in all official languages with the assistance of interpreters, the working draft legislation under discussion is usually only available in one language - English or French or occasionally German. At the end of the drafting phase, a team of specialist legal translators translate the text into the other official community languages, subject to consistency checks by the EU's General Translation Team.

A research group at the European Commission, aiming to consolidate existing EU law, worked on the 'Principles of the Existing EC Private Law' or 'Acquis Principles' (ACQP), which would provide a common terminology as well as common principles to guide uniform implementation and interpretation of European law. The Acquis Principles were sketched by scholars in European Private Law from the so-called *Acquis Communautaire*, a collection of the existing body of EU primary and secondary legislation as well as European Court of Justice decisions. This glossary aims to minimise conceptual differences and semantic ambiguity in the EU legal process, and is of great importance in the activities of translators as well as legal professionals.

Whether the Acquis Principles can truly address the considerable challenges it seeks to address is a moot question (Legrand, 2006). For instance, it does not solve the problem of conceptual and terminological misalignment altogether, since Directives need to be transposed into national law using terms that make sense within the national legal system. In fact, it is precisely this second level of translation that causes most problems. Transposing a rule often means having to use and adapt a different, and sometimes

---

[8] Article 249, Treaty of Rome 1957.

conflicting, lexical baggage to the traditional national one. The incoherent mix of different cultures and traditions results in translations that are often inaccurate or insufficiently precise from a legal point of view.

There are several possibilities in the transposition process, as depicted in Figure 1, where the lines connecting concepts from the EU level to national ones represent the *"implement"* relation.
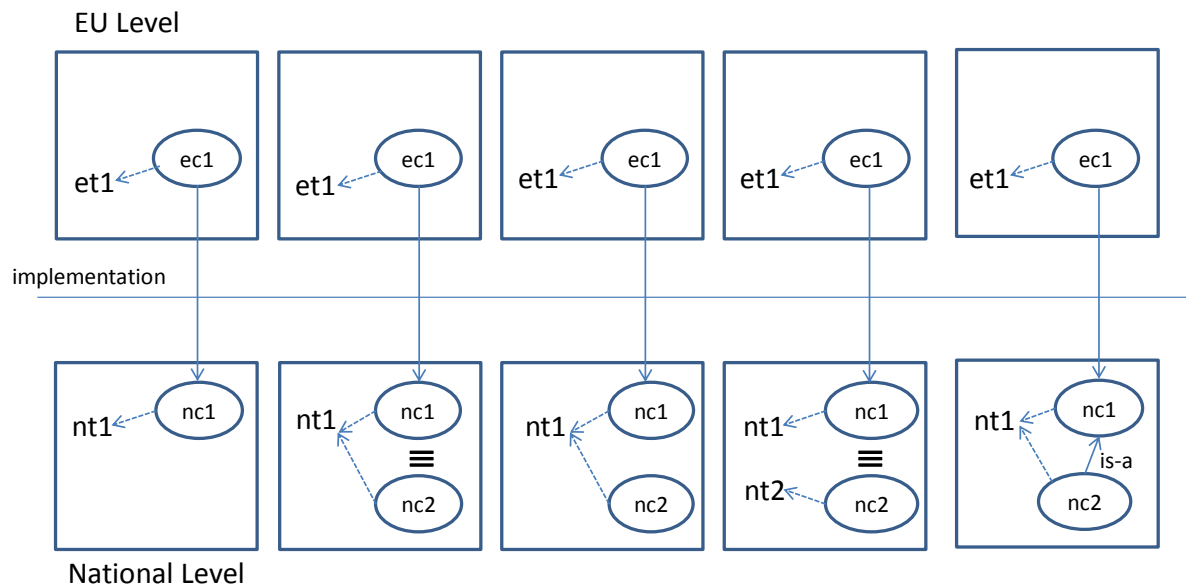

Figure 1. The possible misalignments (1)-(5) between the European and national levels.

The general scenario is composed of a concept ec_1 from an European Directive, which is expressed in the translation of the Directive into a national language with the term et_1. When this Directive is transposed into national law, the concept ec_1 will correspond (hopefully but not necessarily coherently) with a national concept nc_1 expressed by the term nt_1. Different situations arise depending on whether nc_1 and nt_1 already exist at the national level and whether they are already associated with each other or with other terms and concepts. Only in the first two cases are there no misalignment problems:

1. The concept nc_1 is new at the national level and expressed in the transposition with the same term used at the European level $nt_1 = et_1$, not yet used at the national level. *Example*: Undertakings for collective investment in transferable securities (UCITS), regulated by Directive 2009/65/EC of the European Parliament and of the Council of 13 July 2009 on the coordination of laws, regulations and administrative provisions relating to undertakings for collective investment in transferable securities (UCITS).

2. The concept nc_1 is already present at the national level as the equivalent concept nc_2 and expressed by the same term used at the European level $nt_1 = et_1$. The difference compared to the previous case is that, in case 1, there is a single concept expressed with the same description (for example, the EU directive borrowed the definition from that national system), while in case 2 the concepts are defined differently but then nc1 and nc2 are equivalent in their legal effects. *Example*: Good faith as used by Article 3 of Council Directive 93/13/EEC of 5 April 1993 on unfair terms in consumer contracts, i.e. "A contractual term which has not been individually negotiated shall be regarded as unfair if, contrary to the requirement of good faith, it causes a significant imbalance in the parties' rights and obligations arising under the contract, to the detriment of the consumer". The notion of good faith to control unfair clauses is also used within national legal jurisdictions. German law can be cited as an example where (lack of) good faith is one of the criteria used to qualify a clause unfair. See, in particular Article 307 BGB, which provides that: "Provisions in standard business terms are ineffective if, contrary to the requirement of good faith, they unreasonably disadvantage the other

party to the contract with the user. An unreasonable disadvantage may also arise from the provision not being clear and comprehensible".

3. The concept nc_1 is new at the national level and expressed in the transposition with the same term used at the European level $nt_1 = et_1$, which was already used at the national level to mean concept nc_2. Polysemy is introduced at the national level. Example: as argued in (Falkner et al., 2004), the term "parent" in Council Directive 2010/18/EU may be considered as polysemous and can mean "person whose children were born after Directive 2010/18/EU came into force" or "person whose children where born after the *transposition* of Directive 2010/18/EU came into force".

4. The concept nc_1 is already present at the national level but expressed by a different term (nt_2) to the one used at the European level: $nt_2 \neq et_1$. Synonymity is introduced at the national level. *Example*: Legislative Decree No. 170 of 24 May 2004 (Attuazione della direttiva 2002/47/CE, in materia di contratti di garanzia finanziaria), whose Article 8 (1) states that: "the realization requirements and the valuation criteria of financial activities and of the relevant financial obligations must be commercially reasonable". Here the term "reasonable" (ragionevoli) means "correct", in the sense of "compliant to fair dealing standards".

5. The concept nc_1 is already present at the national level and expressed by the same term used at the European level $nt_1 = et_1$, but the meaning of nc_1 is more general or more specific[9] than the corresponding European concept ec_1. *Example*: The term 'consumer' is present in the legislation of the EU member states but in some member states it assumes a broader meaning to that adopted for that term by EU legislation.

Literal or inexact translations are chosen ad hoc in Directives and transposition laws, and as Table 1 and 2 show, the terminology can change over time at both the European and national levels.

Language	Directive 1997/7/EC	Directive 2002/65/EC
DE	Lauterkeit	Treu und Glauben
EN	good faith	good faith
ES	buena fe	buena fe
FR	loyautè	bonne foi
IT	lealtà	buona fede

Table 1

Terms for *"good faith"* Art. 4(2) Directive 1997/7/EC and Art. 3(2) Directive 2002/65/EC

Directive 1997/7/EC	D.Lgs. 206/2005	Directive 2002/65/EC	D.Lgs. 190/2005
lealtà	buona fede e lealtà	buona fede	correttezza e buona fede

Table 2

Terms for *"good faith"* in Italian European Directives and their national implementation

While terminological misalignment can lead to confusion, far more insidious is the effect of conceptual misalignment. For instance, in Directive 97/7/EC, Art. 4.2, the legal terms *"klar und verständlich"* in German, *"clear and comprehensible"* in English, and *"chiaro e comprensibile"* in Italian, appear to have the same meaning until read with the filters of lawyers' domestic legal thinking. *"Klar und verständlich"* in the German system is considered by German commentators as referring to three different legal concepts:

1. The font must be clear and legible (*"gestaltung der information"*),
2. The information must be intelligible to the consumer (*"formulierung der information"*),
3. The language of the information must be in the national language of the consumer (*"sprache der information"*).

---

[9]Note that the fact that two concepts are connected via the *"implement"* relation does not entail they are the same concept; in the national legislation, the implemented concept could have a broader or narrower meaning.

In Italy, the judiciary tend to check more for concepts 1 and 3, and less concept 2, while in England and Wales, more attention has been paid to concept 2, interpreted as requiring a plain style of language (as opposed to technical legal jargon) due to the historical influences of the Plain English Campaign[10] in the United Kingdom.

These problems are confounded by the fact that in every jurisdiction, terms have different meanings in different contexts and over time. This can make it difficult to interpret national legislation that transposes a set of European Directives all at once.

Another issue to bear in mind is that definitions of terms in European Directives and national legislation are sector-specific, and different definitions will apply in other sectors. Consider, for instance, the definition of 'consumer' in the English version of European Directive 1999/44/EC, Art. 1.2 (on sale of consumer goods):

> Consumer: shall mean any natural person who, in the contracts covered by this Directive, is acting for purposes which are not related to his trade, business or profession

which has a different meaning to the definition of consumer given in Council Directive 90/314/EEC, Art. 2.4 (on package holidays):

> "Consumer" means the person who takes or agrees to take the package ('the principal contractor'), or any person on whose behalf the principal contractor agrees to purchase the package ('the other beneficiaries') or any person to whom the principal contractor or any of the other beneficiaries transfers the package ('the transferee').

The European Commission has acknowledged these problems. In its technical report for the European Commission, the Mandelkern Group (Mandelkern, 2001) admitted "lack of simplicity, clarity and accessibility of European provisions —such as unclear, confusing terminology, incomplete or inconsistent regulations or use of vague terms". Moreover, they stressed the need for internal coherence and consistency in the use of EU legal terms, as well as external coherence - consistency in the transposition of legal concepts into national law. In (Commission of the European Communities, 2003), the Commission stated that certain market inefficiencies can be attributed to different interpretations of European contract law due to vague terms and rules, and to the fact that EU contract law consists of several highly sector-specific Directives rather than a coherent approach to contracts in general. Their solution was a Common Frame of Reference (CFR) which would provide common principles, terminology, and legal rules to address gaps, conflicts, and ambiguities emerging from the application of European law.

In summary, an ontology framework for EU law must be not only multi-lingual and multi-jurisdictional but also multi-level. It must be able to model different possible terminological and conceptual misalignments even within the same language. Moreover, it must be able to model both the contextual nature of legal meanings and the interpretations that result from generalizing and merging the meaning of terms across different contexts. Finally, it must be able to include also the conceptualization deriving from draft legislation or scholarly proposals for revision of the law.

3. European Legal Taxonomy Syllabus - The ELTS Ontology Schema

As stated in the Introduction, ELTS is composed of a legal ontology schema, a web-based legal ontology tool conforming to the ontology schema, and a multi-lingual legal ontology on European consumer law constructed with the tool. In this section, we present the legal ontology schema by describing the motivations behind it resulting from the legal analysis in Section 2. In this section, we will present the online tool to construct and browse the ontology, showing for each feature of the tool relevant examples from the consumer law ontology derived from the terminology of the Uniform Terminology project we use as a benchmark to show the feasibility and expressivity of our approach.

---

[10]http://www.plainenglish.co.uk

The European Legal Taxonomy Syllabus is an ontology framework designed to address the issues raised in Section 2 above. The most important insight from lawyers, which informed our design, was that the meaning of a legal term depends on its context (jurisdiction, domain, legislation, timeframe, interpretation). We designed an ontology schema that aims to make these considerations explicit. Our system attempts to model interpretations beyond the letter of the law as well as temporal evolution of concepts in an intuitive way, allowing users to traverse different definitions and determine which definitions are most relevant to their query.

From a pragmatic need to model European law and national transpositions, the ontology framework must be both multi-lingual, multi-jurisdictional and multi-level. This allows links to be made between different national ontologies, so that users can find similar terms in other languages and other jurisdictions, and compare their meaning.

The schema has been designed to support the definition of concepts on the basis of a comparative law methodology. Following Sacco (1991), we chose to adopt a bottom-up approach to ontology creation, i.e., to compare low-level concepts among different legal systems. In our view, a comparative law methodology ensures a non-superficial understanding of legal terms. Starting from low-level elements rather than abstract or composite concepts fosters evidence-based conceptualisations and generally gives rise to less disagreement among ontology contributors. Section 7 discusses how our approach differs from previous ones in this regard. We also adopt the view of comparative law that legal concepts are influenced by formants other than legislation, and ensure that the ontology should provide space for annotation and citations of case law and doctrine.

The main purpose of the European Legal Taxonomy Syllabus tool is to support the work of legal practitioners, scholars and translators in multi-lingual and multi-jurisdictional contexts such as the European Union, to help share technical knowledge and analyse the law in all its complexity. As a secondary aim, the system can be used to build automated tools, e.g., for information retrieval and translation.

Since the ontology framework is primarily designed for human reference, it supports lightweight rather than axiomatic ontologies. In the classification of Giunchiglia and Zaihrayeu (2009) we use "(informal) lightweight ontologies":

> Note that lightweight ontologies are much easier to be understood and constructed by ordinary users. ... designing a full-fledged ontology (expressed, for example, in OWL-DL) is a difficult and error-prone task even for experienced users ...

The choice of building a lightweight ontology was motivated by the need to develop a more user-friendly system, thereby enlarging the possible audience of contributors and users, and at the same time, reducing the costs of building an ontology. It was also driven by the consideration that many peculiarities of law, such as interpretation, penumbra, interaction with social values, metaphors and dynamics are far from having commonly accepted solutions in logic (as discussed in Section 2).

Our analysis of the legal domain led us to identify the following features in the ontology schema to allow the representation of relevant information in the legal domain. Such features are not always straightforward to represent using standard approaches to ontology design.

Terms and concepts: the varying and highly contextualised meaning of legal terms means that there needs to be a structured way to allow more than one meaning for terms in a legal ontology. ELTS separates terms and concepts, allowing terms to be mapped to different concepts and to have concepts mapped to more than one term (in the same language, or in different languages in the case of multi-lingual nations and of EU law). Terms can be either single words or multi-words (cf. examples below). Therefore we have many-to-many relations between terms and concepts, thereby allowing both synonymity and polysemy. Since we are in a multi-lingual context, a term in our system is structured as the term itself together with the jurisdiction identifier and the relevant language, in order to account for multiple languages in the same jurisdiction. The idea of neatly separating the lexicon from the conceptual level is of course not new (cf. (Buitelaar, 2010)) and is the foundation of several models, including the well- known Lemon lexicon model (McCrae et al., 2011). Lemon

was designed to model lexicon and machine-readable dictionaries relative to ontologies in OWL standard. It has been successfully used to publish many dictionaries/ontologies as Linked Data e.g., (Villegas and Bel, 2015), (Ehrmann et al., 2014).

Sources: each concept is linked to its legal source(s), possibly more than one, since a concept can arise from multiple parts of legislation and also from the interaction of legislation, case law and doctrine. However, listing the sources is not enough. For the sake of clarity, concepts are associated with a description in natural language. Nevertheless, it is important to identify the legal sources, since they contain important information about scope and purpose.

Domains: in addition to the contextual information of legal sources, the concepts are classed in domains, traversal with respect to the jurisdictions and levels, in order to organize knowledge and improve search and browsing. Each concept can be associated with more than one domain.

Multi-lingual and multi-jurisdictional: the multi-jurisdictional nature of the EU requires not only a multi-lingual ontological framework associating concepts with terms in different languages, but also a multi-jurisdictional one. The ELTS schema involves separate ontologies for each jurisdiction whose concepts are in turn mapped to terms in relevant languages. Specific relations connect concepts from the ontologies of different jurisdictions, which are separate from the relations within the same ontologies. In particular, the relation *"implement"*, described in more detail below, connects concepts in the EU ontology with concepts in the national ontologies.

Multi-level: besides being multi-lingual and multi-jurisdictional, the ELTS schema distinguishes between the EU level and national levels: these constitute separate ontologies. Note that the EU level contains a single ontology, where all concepts have associated terms in the different languages of the Member States considered. This, however, is a simplification, since it is possible that there are unwanted divergences among different languages even at the EU level. Concepts in the EU level ontology can be associated with terms in all the languages of the Member states. Concepts in the different ontologies at the national level can be associated only with the terms of the languages of the nation they belong to.

Ontological relations: due to the holistic nature of the law, legal concepts are better understood in relation to others. Therefore within each ontology, the concepts are linked via ontological relations such as *"is-a"*, *"part-of"* and by more specific legal relations such as legal *"purpose"*, expressing the legal goal (e.g., *"consumer protection"*) that the legal system aims to achieve with that concept. Relations among concepts may change over time (as stated below).

Implementation relation: Concepts at the EU level can be connected to national level concepts by an *implementation* relation, representing how the concept has been transposed into one national legal system. Given the separation of terms and concepts, the term associated with an EU level concept is not necessarily the same term used to express the same concept in the implementing legislation at national level. The relation is many-to-many, since a national level concept may be the fusion of more than one EU level concept and/or a Member State may express an EU level concept in multiple ways.

Dynamic nature of meaning: the ontology schema must account for the fact that almost every legislation brings new definitions of terms that effectively replace prior conceptualisations. Therefore the ontology must represent the current legal situation, and yet researchers may still need to refer to deprecated conceptualisations for historical purposes or to trace the evolution of terms. This also raises the problem of what happens to the ontological relations associated with the replaced concept, a sort of frame problem. Since we are dealing with a semi-automated context, the proposed solution is that the new concept should inherit all the relations of the replaced one and it is the responsibility of the knowledge engineer to remove the outdated ones and possibly introduce new relations in accordance with authoritative interpretations.

Interpretation: there is a tension in the law between the highly contextual character of meaning, which leads to multiple meanings for one term, each one associated with specific sources, and the need to systematize legal knowledge. The legislator can introduce in a new legislation a new meaning for a term, whose utility can go beyond the context of that legislation. Due to interpretation, the meaning

of the term can be extended also to other concepts denoted by the term in other contexts. The frequent merging of meanings assigned to legal terms that takes place in legal reasoning or simultaneous transposition of multiple directives means that more complex concepts can emerge which do not necessarily replace contextual meanings in all situations. This situation cannot be simply modelled by *"is-a"* relations, since the concepts resulting from the interpretation are neither necessarily more general in meaning, nor necessarily the simple intersection of the more contextualized meanings. Rather, what is generalized is the context of use of the concept. Moreover, the original contextual meaning of a term must always be available to the user and not only the merged one deriving from interpretation.

Conceptual drafts: the ontology must be able to accomodate the conceptualization also of draft legislation, to compare the resulting "draft" ontology. Glossaries created to achieve consistency in legal terminology, such as the CFR and ACQP above, may contain conceptualisations that are yet to be accepted officially. ELTS allows the creation of temporary legal ontologies whose concepts are linked to current legal ontologies until such time as the old concepts are replaced when the draft legislation becomes law.

Given the non-formal character of ELTS, which cannot be defined in a standard ontology language such as OWL, the ELTS ontology schema is provided here as a UML semi-formal representation (Figure 2). The ELTS database schema is of course very close to the ontology schema in Figure 2.
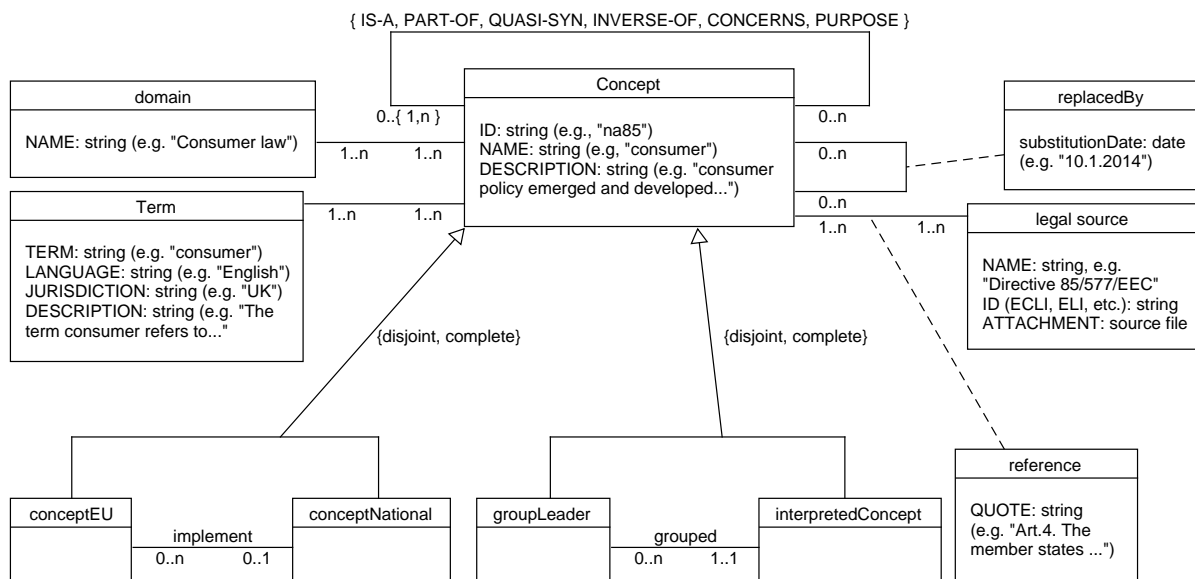
Figure 2. The UML representation of the ELTS Ontology Schema.

The details of the UML schema in Figure 2 are explained, step-by-step, in the next section.

Note that the main class of the ontology schema is "Concept", which is related to all the other classes. The workflow of ontology building with ELTS is to start from the terms (class "Term" in Figure 2), i.e. from the linguistic patterns occurring in legal texts. Nevertheless, ELTS is not intended as a terminological vocabulary but rather a (lightweight) ontology, focussed on the definition of the legal concepts (to which the terms are connected via many-to-many relations).

4. European Legal Taxonomy Syllabus - The ELTS Ontology Tool and the ELTS Ontology

The following sections describe the European Legal Taxonomy Syllabus Tool developed in accordance with the above schema. For each feature, we describe examples taken from the Uniform Terminology project[11] to validate the ability of the tool to model complex legal knowledge. The data collected by the project is on the subject of consumer law. Terms were initially extracted from a *corpus* of 24 EC directives, and 2 EC regulations. Occurrences of such entries were detected from national transposition laws from English, French, Spanish, Italian and German jurisdictions. The actual number of annotated terms and concepts are provided in Tables 3 and 4, respectively.

Language	National	European
French	8	47
Italian	28	52
English	71	75
Spanish	41	60
German	66	98
total	214	332

Table 3

Number of current terms in ELTS

Language	National	European
French	7	43
Italian	24	45
English	54	71
Spanish	34	56
German	52	75
total	171	290

Table 4

Number of current concepts in ELTS

4.1. Terms and Concepts

Law is a highly polysemous domain, and the meaning of terms is highly context-sensitive. Legal terms are often defined differently in different legislation and other legal formants, and each definition can be regarded as a different concept. At the same time, some concepts can be expressed in different terms, particularly when transpositions from European law sometimes import literal translations rather than using the terms that have evolved organically in the national jurisdiction. All this means that an accurate ontology of law, particularly in a multi-jurisdictional environment, has to deal with possible relations between terms and concepts.

ELTS is structurally different from ontologies that permit only one meaning for every term. In ELTS, terms (lexical entries) are separated from concepts. Figure 3 shows the sub-schema corresponding to the concepts, the terms and the legal sources from which they are taken.
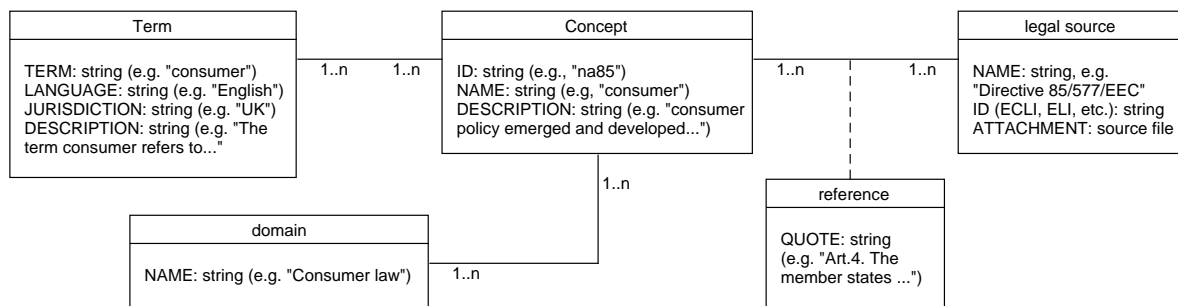
Figure 3. UML sub-schema of Figure 2: Concepts, Terms, and legal Sources.

Concepts - associated with descriptions, source, and domains - are identified by unique identifiers. A concept is mapped to one or more terms. A term, which could be a single word such as *"contract"* or a multi-word term such as *"clear and comprehensible"*, is mapped to one or more concepts.

---

[11] http://www.uniformterminology.unito.it

Moreover, the concepts are associated with quotations from the sources contributing to their definition. In Figure 3, "reference" is an association class that adds an attribute to the many-to-many associations between concepts and legal sources. That attribute allows the exact portions of text to be specified, e.g. an article of a European Directive in which the concept was defined.

Figure 4. An ELTS interface showing legislation and a list of the concepts it contains with anchors and highlighted terms. The Figure refers to the concept "adecuado". That concept is associated with the term "necesario", which belongs to the jurisdiction of Spain and occurs in (the Spanish version of) Directive 2000/31/E.

The text of the sources is maintained in a database of documents, so that it is possible to establish a link between the concept and the whole text of the source in which it is defined. In this way, it is possible to show the user the text of legislation and case law enriched with links to occurrences of concepts and highlighting terms related to those concepts (see Figure 4). Section 5 describes in detail the interface for creating references to sources starting from the legal text.

In Figure 5, we illustrate how a concept associated with a term is shown to the user. The concept associated with the term *"consumer"* at the European level is shown in a tabular format. For ambiguous terms, more than one table is shown for each different meaning. The language, the alternative linguistic realizations associated with the concept, and the domain ("Consumer law") all precede the description of the term in natural language.

As stated above, ELTS allows the insertion in the text of paragraphs where concepts are defined. However, to gain better understanding of legal concepts, it is often necessary to consider a broader fragment. For example, in the case of *"consumer"*, the definition is not enough, and it is necessary to collect multiple paragraphs where consumer protection norms are presented and discussed. References to legislation, notes and attached documents are also optional fields. The implementation and national associations fields are discussed below. Figure 6 shows the concept at the national level.

Example. The system helps the user understand documents by showing where concepts (and their terms) are highlighted. The occurrences of the concepts are listed at the beginning of the documents and are based on the references to the text within the concept entities. To understand how it works, compare a simplified printout showing the text of a directive in Spanish language in Figure 4, with the printout of the concept in Figure 7. The references associated with one meaning of the term *"adecuado"* are used to identify

Figure 5. The European concept associated with the English term *"consumer"*.

the paragraphs of the legislation text where the term occurs (e.g., the paragraph *"Articulo 3 Mercado Interior"*). Note that in the text the reference is not to the term *"adecuado"* but to a specific concept to which the term is linked, with id=163. The concept shown in Figure 7 is among several associated with that term (see part of the list of concepts in Figure 8). This is achieved with pattern-matching techniques, since the text in ELTS (unlike the Eunomos system (Boella et al., 2016) derived from the Syllabus, see Section 6) is not structured as legislative XML.

4.2. Ontological Relations

To help users understand the inter-relationship between concepts, ELTS contains a number of ontological relations between concepts (see Table 5). Notwithstanding classic relations such as the "*is_a*" relation linking a category to its supercategory, it also includes the "*purpose*" relation, which links a concept to the legal principle behind it, and "*concerns*", which refers to a general relatedness between concepts and is often used to link complex concepts to the basic concepts that contribute to their definitions. Legal practitioners have also deemed it useful to have a similarity relation between concepts, and this similarity relation is not the same as strict synonymity between terms: "*quasi_synonym*".
In this way, European Legal Taxonomy Syllabus enables clear modeling of the similarities and differences in conceptual inter-relations in different jurisdictions. It is important to note that these structures

Figure 6. The national concept associated with the English term *"consumer"*.

Relation	Meaning	Property
IS_A	Something is a specific kind of another thing	Transitive
PART_OF	Something is part of another thing	Transitive
QUASI_SYN	Two concepts are almost, but not quite, identical	Non-transitive
INVERSE_OF	Something has the opposite meaning of another thing	Non-transitive
CONCERNS	A concept object of another concept	Non-transitive
PURPOSE	Links a concept to the legal principle behind it	Non-transitive

Table 5

Ontological relations in ELTS

can also change. With the modification of concepts in new legislation at the national or European level, new purposes can be introduced or old ones rendered obsolete. Note that the above ontological relations connect only concepts within a single ontology (the EU ontology or the one of a Member state) and are kept distinct from relations across different ontologies such as *"implementation* or *"translation*. Figure 9 shows the interface for displaying the ontology involving a concept (in this case, the Spanish concept "prestamista").

Example. Figure 10 shows how the *"purpose"* relation can provide clues to the applicability of a concept. The configuration of arcs shows that *"withdrawal"* and *"recesso"* have been used as equivalent terms in some European Directives (e.g., Directive 90/314/EEC), where the terms involve an act having as its purpose *"consumer protection"* (in Italian, *"difesa del consumatore"*). In the legal system of England and Wales, however, not all *"withdrawals"* have this purpose, only a subtype, which are referred to as

Livello europeo

Azioni	
Lingua giuridica	Spanish
Termine	• adecuado • suficiente • necesario
Domini	Consumer Law
Descrizione	⊟ El término "adequate" aparece en la Directiva, y es traducido en la versión española como "suficiente", "adecuado" y "necesario" (art. 3, 17 y 19). Se trata de artículos prevalentemente dirigidos a los Estados, pero se pueden encontrar artículos de la Ley Ley 34/2002, de 11 de julio, de servicios de la sociedad de la información y de comercio electrónico, que aplican lo dispuesto en la Directiva. En el art. 8, dedicado a las "Restricciones a la prestación de servicios y procedimiento de cooperación intracomunitario", aparece el término "insuficiente", en relación a los deberes del órgano competente para requerir al Estado miembro en que esté establecido el prestador afectado para que adopte las medidas oportunas, en el caso de que no las adopte o resulten inadecuadas.
Riferimenti	⊟ 2000/31/CE (Originale) (Commenta) [...] Artículo 3 Mercado interior [ ] (Originale) (Commenta) [...] 4. Los Estados miembros podrán tomar medidas que constituyen excepciones al apartado 2 respecto de un determinado servicio de la sociedad de la información si se cumplen las condiciones siguientes: a) Las medidas deberán ser: i) necesarias por uno de los motivos siguientes: - orden público, en particular la prevención, investigación, descubrimiento y procesamiento del delito, incluidas la protección de menores y la lucha contra la instigación al odio por motivos de raza, sexo, religión o nacionalidad, así como las violaciones de la dignidad humana de personas individuales, - protección de la salud pública, - seguridad [...] (Originale) (Commenta) [...] 5. en caso de urgencia, los Estados miembros podrán establecer excepciones a las condiciones estipuladas en la letra b) del apartado 4. Cuando así ocurra, las medidas se notificarán con la mayor brevedad a la Comisión y al Estado miembro a que hace referencia el apartado 1, indicando las razones de la urgencia según el Estado miembro. [...]
Note	...
Allegati	
Implementazioni	• suficiente

Figure 7. The European concept associated with the Spanish term *"adecuado"*, with its synonyms. Note that the interface language here is Italian.

"cancellation" (concept Eng-3). On the other hand, the concept *"diritto di recesso"* (*right of withdrawal*) in the Italian implementation of the Directives (concept Ita-5) is related to *"recesso"* (concept Ita-3) via a *"concerns"* relation. *"Diritto di recesso"* is a complex concept that includes *"recesso"* in its definition. In other words, ELTS relates the right of enforcing the specific act of withdrawal to the act itself.

4.3. Interlinked Ontologies

It is well known that different legal systems have different conceptualisations. For instance, the term *"contract"* corresponds to different concepts in the common law and civil law traditions, where it has the meaning of *"bargain"* and *"agreement"* respectively (Sacco, 1999), (Pozzo, 2003). In the European Union the situation is complicated further by the presence of another multi-national regulation, the *Acquis commu-*

Termine europeo "adecuado"

Azioni	Mostra ontologia
ID	
Lingua giuridica	Spanish
Termine	• adecuado • suficiente • necesario
Domini	
Descrizione	El art. 10 de la Directiva està dirigido a los Estados; por esta razon no ha sido transpuesto ni en la Ley 9/1999, de 12 de abril, ... ⊞
Riferimenti	⊞ 97/5/EC
Note	...
Allegati	
Implementazioni	• suficiente

Azioni	Mostra ontologia
ID	
Lingua giuridica	Spanish
Termine	• adecuado
Domini	
Descrizione	El término aparece en el art. 11 de la Directiva, que está dirigido a los Estados. Por tanto, no es objeto de transposición. En materia ... ⊞
Riferimenti	⊞ 97/7/EC
Note	...
Allegati	
Implementazioni	

Azioni	Mostra ontologia
ID	
Lingua giuridica	Spanish
Termine	• adecuado
Domini	

Figure 8. A table listing different concepts associated with the term *"adecuado"* (note that the interface here is in Italian).

nautaire, which has been translated into all the languages of the Member States. It cannot be assumed that the transposition of an European Directive (EUD) necessarily introduces in a national ontology the same concepts that are present at the EU level, since national legislators may transpose European legislation in a way that accords with the national legal tradition. Terminological misalignment can occur even in the same language, as a legal concept can be expressed in different ways in a Directive and in its implementing national law.

As such, the European Legal Taxonomy Syllabus was designed to be a framework of separate and inter-linked ontologies, one for each relevant national jurisdiction plus one for the European level. Each national ontology is related via a set of implementation links from the EU ontology, as shown in Figure 11. The set of concepts may be partitioned into the set of concepts belonging to the EU ontology and those belonging to the national ontologies. The latter are related to the former via an *"implement"* ontological relation that, as stated above, is a many-to-many relation, since a national level concept may be the fusion of more than one EU level concept and/or a Member State may express an EU level concept in multiple ways. The separation of the ontologies at different levels, combined with the implementation links and the separation

Ontology

Ontology graph

⊟ "comerciante"
    ⊟ IS_A "profesional"
        ■ **IS_A "prestamista"**
    ⊟ IS_A "vendedor"
        ■ IS_A "vendedor final"
    ■ IS_A "proveedor"

European level

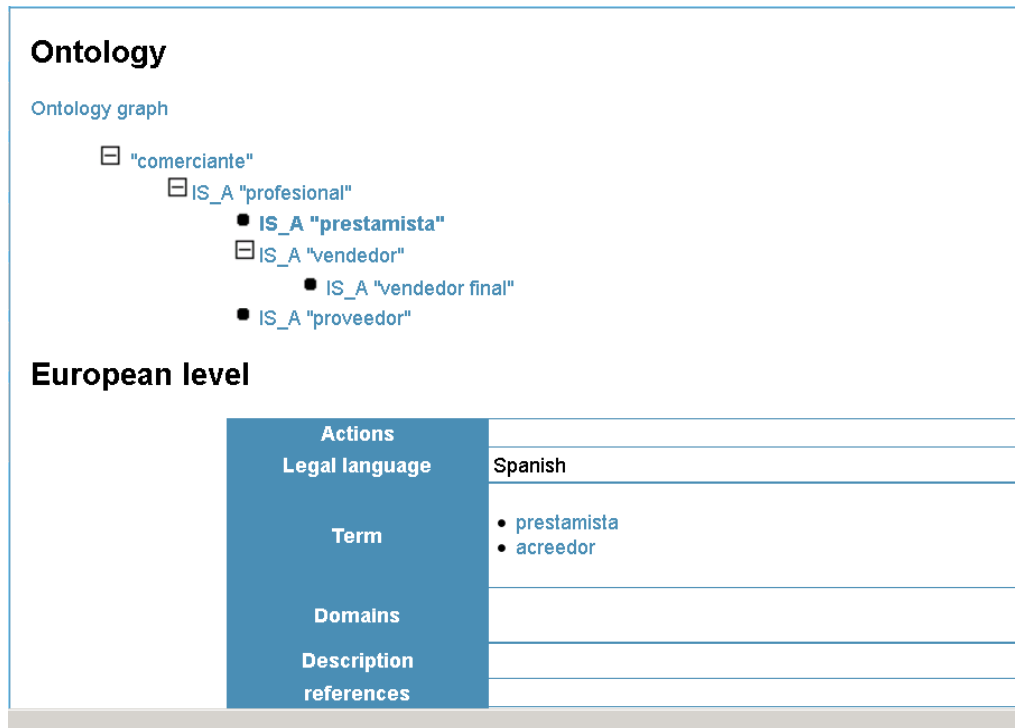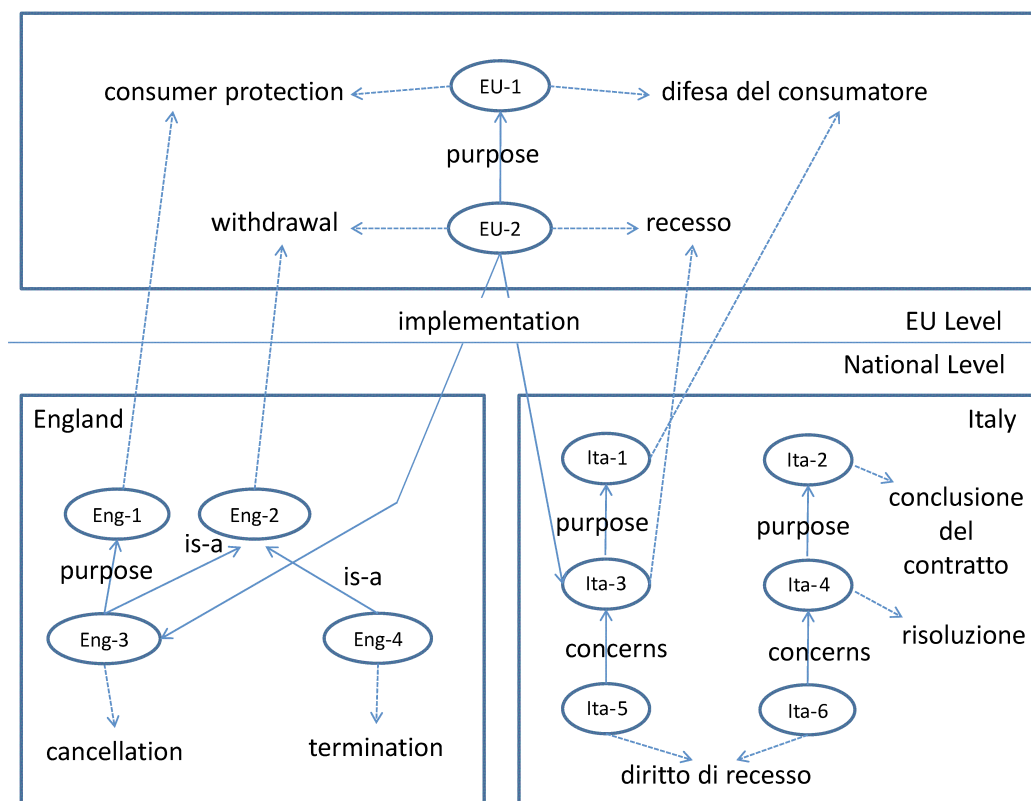| Actions	
Legal language	Spanish
Term	• prestamista • acreedor
Domains	
Description	
references	

Figure 9. The ontology relations of the Spanish EU level concept *"prestamista"*.

Figure 10. The mapping of European legal terms to national legal terms in Italy and England & Wales

of concepts and terms at each level, allows ELTS to model all the possible types of misalignment described in Section 2.3.
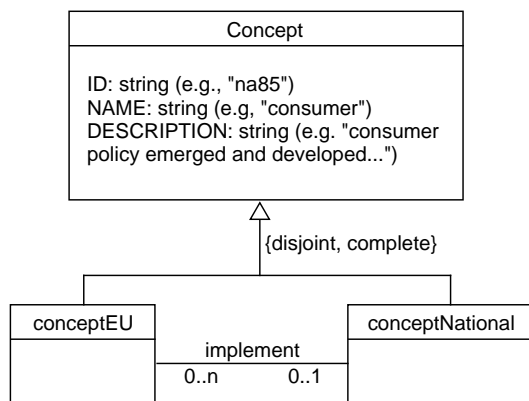


```
                    ┌─────────────────────────────────────┐
                    │              Concept                 │
                    ├─────────────────────────────────────┤
                    │ ID: string (e.g., "na85")            │
                    │ NAME: string (e.g, "consumer")       │
                    │ DESCRIPTION: string (e.g. "consumer  │
                    │ policy emerged and developed...")    │
                    └─────────────────────────────────────┘
                                     △
                              {disjoint, complete}

       ┌─────────────────┐    implement    ┌─────────────────┐
       │   conceptEU     │                 │ conceptNational │
       │                 │   0..n   0..1   │                 │
       └─────────────────┘                 └─────────────────┘
```

Figure 11. An UML sub-schema of Figure 2: the *"implement"* ontological relation.

Example. Figure 10 above shows relations between concepts and terms in three different ontologies for the legal systems of Italy, England & Wales, and the European Union respectively. The dotted arcs represent references from concepts to terms. Terms can be mapped to a national ontology and to the EU ontology. As the figure shows, in the Italian legal system, the term *"diritto di recesso"* is ambiguous as it can be used to mean that one party can unilaterally rescind a contract (the objective being to protect the consumer) or it can be used to mean an agreement by two parties to rescind a contract - one particular way to bring a contract to an end. Linking the term to two distinct concepts bring clarity to the end-user. Note that a complication arises also in England & Wales, in that the term used at the European level, *"withdrawal"*, is ambiguous in the national context with two possible meanings, denoted by *"cancellation"* and *"termination"*, and only the first concept has the same legal purpose as the European concept.

Using ELTS, we can derive direct EU-to-national translations of terms using the implementation links, as well as implicit national-to-national translations of terms.

Direct links can be explicitly created by legal experts to specify a relation between EU concepts and those at the national level. This is the meaning of the field *"national associations"* in Figure 6 above. This results in indirect links: if we start from a concept at a given national level which is the implementation of an EU level concept, by following the inverse of the implementation link we reach the concept at European level. Since the EU level concept holds for all Member states and should be implemented in each state, we can see how that concept is mapped onto further concepts at the various national levels and which terms are associated with them.

Suppose that we wish to find the German national legal term for the Italian national term *"credito al consumo"*. ELTS can reveal that *"credito al consumo"* is linked to national concept Ita-175, which is the transposition of European concept eu26, which is linked to the term *"contratto di credito"*. Eu26 is also linked to the German legal term *"Kreditvertrag"* at the European level. We then find that the national German transposition of eu26 corresponds to national concept Ger-32 that is associated with the national legal term *"Darlehensvertrag"*. This enables us to translate the Italian legal term *"credito al consumo"* into the German legal term *"Darlehensvertrag"*. Thus, it is possible to use ELTS to translate terms from one national system into another national legal system via the transposed concepts at the European level, i.e., by using the implicit associations.

Of course, since different jurisdictions do not always have the same concepts, translation is not always that straightforward. Figure 12 depicts the legal concept eu74 corresponding to the legal words *"comerciante", "commerçant", "trader", "commerciante"* in Spanish, French, English and Italian respectively. Several EU provisions on consumer law employ the hypernym *"trader"* in English and other languages. But there is no equivalent word in German (Figure 13). The clear structure of ELTS can help expose terminological gaps to be resolved by translators and terminology managers.
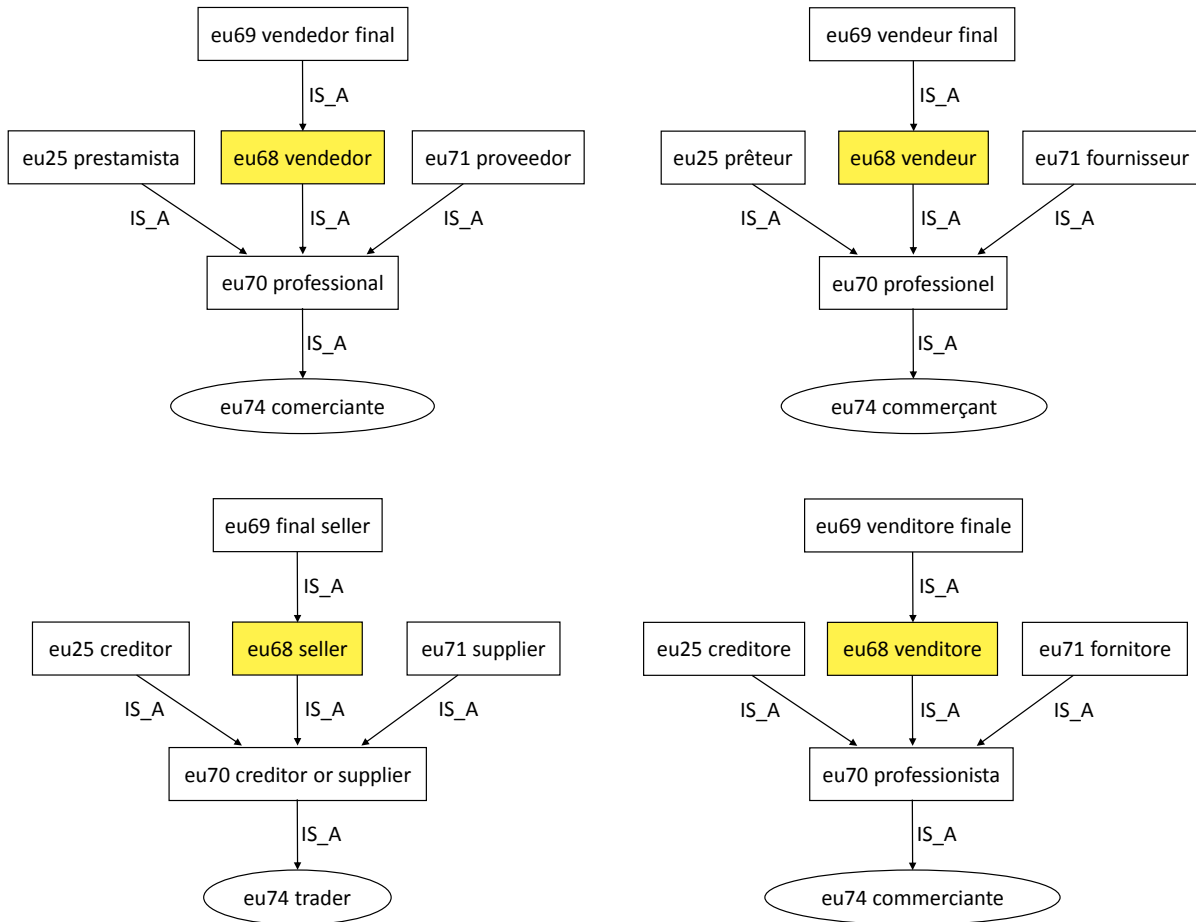
Figure 12. A fragment of the EU level ontology in (four) different languages
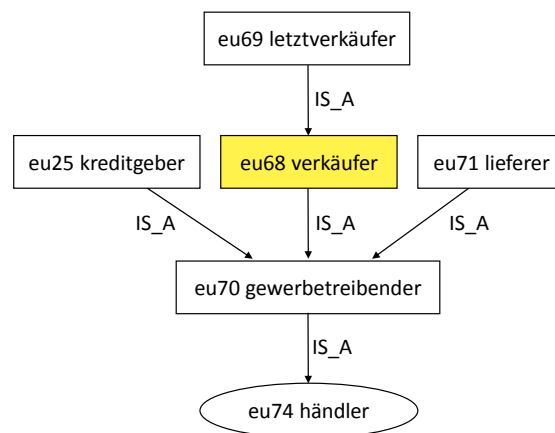
Figure 13. The German counterpart of the same fragment of the ontology as shown in Figure 12.

4.4. Temporal Dimension

Legal concepts constantly evolve. When new legislation is approved and enacted, it can introduce a number of new definitions that can change the meaning of terms defined by old laws, rendering the old conceptualisations obsolete. Normative change (Palmirani and Brighi, 2006) is an open issue in building tools for formal models of legal frameworks. There are two types of normative change: explicit change and implicit change. In the first case, the new norm explicitly states the abrogation of a specific paragraph of an old law (see, e.g., (Spinosa et al., 2009)). Alternatively, the new law can define a term in a way that contradicts definitions in previous laws without mentioning these laws explicitly. The same can take place if a judgement, or even an authoritative scholarly work, changes the meaning of a term. In these cases, the new definition implicitly renders the old definition obsolete.

The evolution of concepts is often ignored in the design of ontologies, since for most domains, users are usually only interested in the current state of affairs. A standard way of representing the temporal dimension of concepts is to associate them with a specific lifetime. This is the approach undertaken, for instance, in the W3C PROV ontology[12], where the lifetime is defined as the period between a generation time and an invalidation time, which are two time units, e.g., two dates. In W3C PROV, the evolution of concepts is modeled by associating entities with *alternates* via special relations such as '*specializationOf*' and '*alternateOf*'. Those entities are different from the original entities and may have different features, e.g., different lifetimes[13].

We believe that it is essential for legal ontologies to provide mechanisms for keeping track of conceptual changes, since deprecated definitions may still be useful, e.g., when looking at cases occurring before the introduction of new legislation, or in order to understand the evolution of terms. Similar concerns lie behind the work of Palmirani and Brighi (2006) in modelling time-specific multi-version coordinated legal texts.

We are aware of the difficulties of modeling time in artificial intelligence and in formal ontology. On the other hand, the expressivity provided by W3C PROV and similar solutions is beyond what is needed for the legal domain, where it is sufficient to represent the legal concepts that stand at the current time, along with any (deprecated) ones that have been replaced.

As such, European Legal Taxonomy Syllabus introduces a simpler solution inspired by the way time is dealt with in the databases of data warehouses - we add a new ontological relation called *"replaced_by"*. In the UML ontology schema shown in Figure 2, the *"replaced_by"* relation corresponds to an association class (see Figure 14).
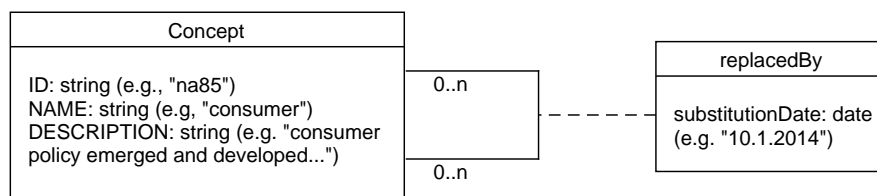

Figure 14. UML sub-schema of Figure 2: the *"replaced_by"* ontological relation.

The *"replaced_by"* relation allows new concepts to explicitly replace old ones while old concepts are still retained in the system. The semantics of this relation is procedural, in that it is embedded in the way the database deals with this relation when it is added. The new ontological relation is dealt with by the system with some particular characteristics:

– A *"replaced_by"* relation brings with it a new data field not present in the other relations: the substitution date.

---

[12] https://www.w3.org/TR/prov-o
[13] Cf. https://www.w3.org/TR/prov-constraints/#entities-activities-and-agents

- When the user searches in the concepts database, replaced concepts will not be shown unless the user specifies a certain date in the past. This enables the user to obtain a snapshot of the legal ontology for any particular moment.
- When a new concept is an update of an old one, all the ontological relations linked to the old concept are automatically copied to the new concept. If some of these relations are no longer valid for the new concept, the interface allows manual intervention from the user.

Example. In Figure 15, EUD 2002/65/EC introduces a new conceptualisation of consumer which effectively replaces the old conceptualisation. This is modelled in the European Legal Taxonomy Syllabus with an explicit assignment of a *"replaced_by"* relation. As the concept of consumer is replaced, the system gathers all relations pertaining to the old concept and assigns them to the new concept. For instance, the *"end consumer"* concept , which is modelled as a sub-type of consumer in the ontology, is automatically linked to the new concept (and the link to the old concept is maintained to preserve a snapshot of the past).
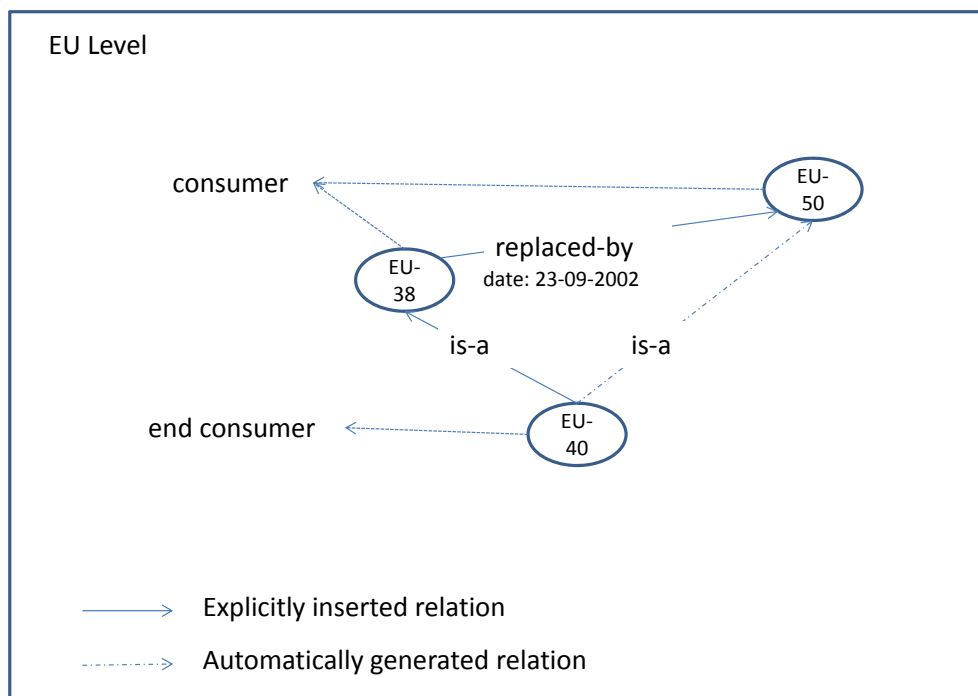

Figure 15. Relating an expired concept to a concept that replaces it.

4.5. Concept Interpretation

While legal professionals need access to legislative definitions to have a precise model of European law, we accept the idea of several scholars ((Sacco, 1987), (Vanderlinden, 1995) and (De Groot, 1999)) that there is more to legal knowledge than the knowledge contained in statutory rules, since legal professionals must also take into account case law and jurisprudence.

Legal reasoning can involve conceptualisations that are legislation context-specific but may also involve considering a broader fragment. Lawyers and legislators often combine sector-specific concepts into more abstract concepts with broader contexts of application. This happens, for example, in order to describe complex entities, such as consumer protection, in all their aspects, even if the single EU Directives concern specific contexts only, e.g., online commerce. This is particularly important when considering European Union terms, since different definitions may not all make their way into national legislation. For example, the Italian legislature has transposed sets of European Directives into unified Italian legislation, rather than transposing each one individually, thus implicitly merging the contexts of applications. The legal concepts

are defined as the merging of all the concepts provided by the individual Directives in a specific sector, as a result of doctrinal interpretation of the Directives.

The European Legal Taxonomy Syllabus introduces the 'interpreted' concept to deal with this situation. As more legislation-specific definitions about terms are collected, it becomes possible to merge them and apply them to a wider context. Interpreted concepts are not necessarily linked to single pieces of legislation, but are conveniently recognized as a generalisation of legislation-specific concepts, to which links must be maintained to enable tracing back to the specific sources. ELTS also introduces a new kind of ontological relation called *"grouped"*: it is a non-transitive relation where the more general conceptualisation, called the *group leader*, represents an abstracted concept that groups a number of more context-specific conceptualisations for the same term directly from legislation. Of course, also the group leader is related to this term.

As shown in Figure 16, the set of concepts can be partitioned into two subsets: abstracted concepts and interpreted concepts. All abstracted concepts are group leaders. Some of them are not associated with any interpreted concept, i.e. with any context-specific interpretation. The others are connected via the relation *"grouped"* to their interpreted concepts. The interpreted concepts are in turn associated with the legal sources (e.g., case law) where the corresponding specific interpretation was used.
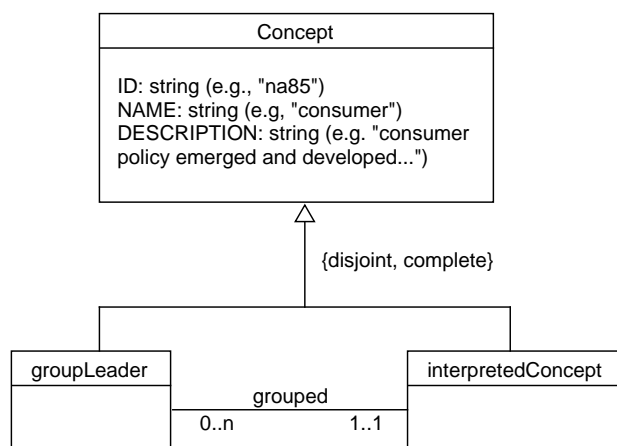

Figure 16. An UML sub-schema of Figure 2: partition into group leaders and interpreted concepts.

This solution is in keeping with a bottom-up approach to ontology building. Users can navigate the ontology at different levels of detail, depending on their goals. When users search ontologies for particular terms, only group leader conceptualisations are shown in the results by default, and more specific conceptualisations are excluded from the results unless explicitly requested.

Example. Figure 17 shows two similar conceptualisations of *"consumer"* grouped into an abstract concept. In the Italian version of the EUD 93/13/EEC, consumer is defined as follows:

"consumatore": qualsiasi persona fisica che, nei contratti oggetto della presente direttiva, agisce per fini che non rientrano nel quadro della sua attività professionale;

which translates literally into English as:

"consumer": means any natural person who, in contracts covered by this Directive, is acting for purposes which are outside his professional activity;

The definition of consumer in EUD 2002/65/EC is broader:

"consumatore": qualsiasi persona fisica che, nei contratti oggetto della presente direttiva, agisce per fini che non rientrano nel quadro della sua attività commerciale o professionale;

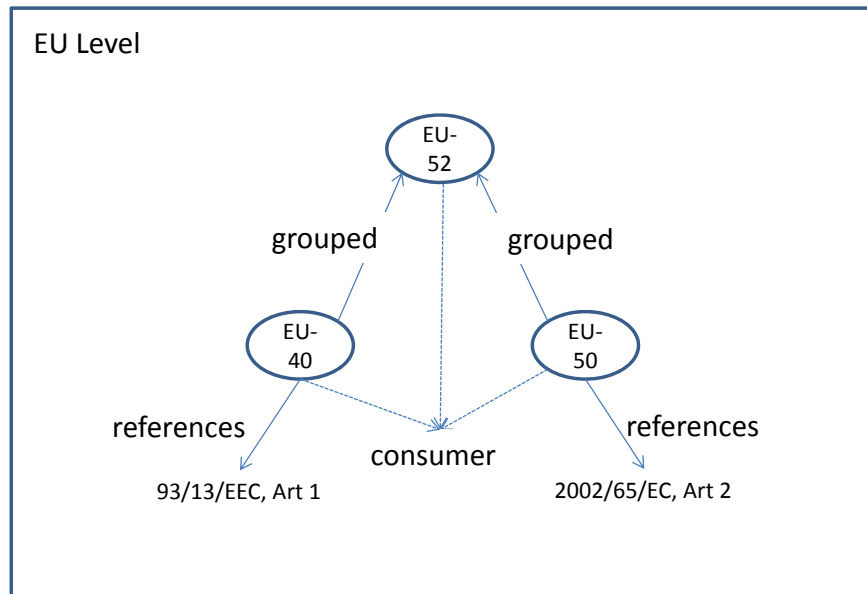which translates literally into English as:

Figure 17. Grouping similar concepts into a concept applied to a more general context.

"consumer": means any natural person who, in contracts covered by this Directive, is acting for purposes which are outside his business or professional activity;

Please note that, in contrast with English, in Italian the second definition of consumatore is broader than the first one, since the term 'professionale' (professional) does not include 'commerciale' (business).

Such a divergence of term definition can often occur without any real difference in the reading of the original legislation, since EUDs usually target a specific sector. EUDs covering different sectors can provide different definitions, and as many views on the same concept. Lawyers and legislators started to combine highly sector-specific concepts into concepts with broader meaning and context, in order to describe (complex) entities, such as'consumatore' in all of its aspects. This is useful because in recent years, implementing legislation in Italy often transpose groups of EUDs rather than single EUDs. The juridical concepts are thus defined as the union of all the sectorial concepts provided by the individual EUDs, as a result of the doctrinal interpretation of directives.

We remark that these problems are common to all European languages. Consider, for instance the definition of consumer, in the English version of EUD 1999/44/EC, Art. 1.2: [...] (a) consumer: shall mean any natural person who, in the contracts covered by this Directive, is acting for purposes which are not related to his trade, business or profession; [...] that has a different meaning with respect to the definition of consumer given in Council Directive 90/314/EEC, Art. 2.4: [...] 'consumer' means the person who takes or agrees to take the package ('the principal contractor'), or any person on whose behalf the principal contractor agrees to purchase the package ('the other beneficiaries') or any person to whom the principal contractor or any of the other beneficiaries transfers the package ('the transferee') [...]". ELTS is able to represent both the more specific dimension related to the definitions in EUDs and the dimension that results from doctrinal interpretation of European law.

Note that the classes groupLeader and interpretedConcept could be better seen as *roles* of the class Concept, rather than sub-classes. In the current version of ELTS, we adopted this simplified model in that the treatment of roles in juridical ontologies is still an open problem (see (Boella et al., 2007)).

4.6. Conceptual Drafts

Provisional legal systems can arise during the course of protracted legal reform or review of legal terminology, which, due to the role of doctrinal elaboration in law, influence legal concepts and legal reasoning, even before official approval. Such situations present the challenge of how to extend the ontological

framework to include a new provisional ontology representing possible reforms, and how to then represent the relation between current and provisional ontologies. A case in point is the *Acquis Communitaire* Principles glossary, intended to replace existing European terminology, and which influenced legal reasoning on European law when issued in draft form.

One major feature of the ELTS approach is that it is possible to classify legal information as belonging to different levels. The ELTS approach was devised to be general enough to account for heterogeneous legal sources (like, e.g., EUDs and *Decreti Legislativi* for European and Italian national levels respectively), and flexible enough to be extended by adding further levels, arising from drafts of legislations. To add a new level into the system, we connect a new legal ontology to an existing one. The new level is linked via explicit associations across ontologies connecting a concept belonging to the new ontology to a concept belonging to the existing ontology.
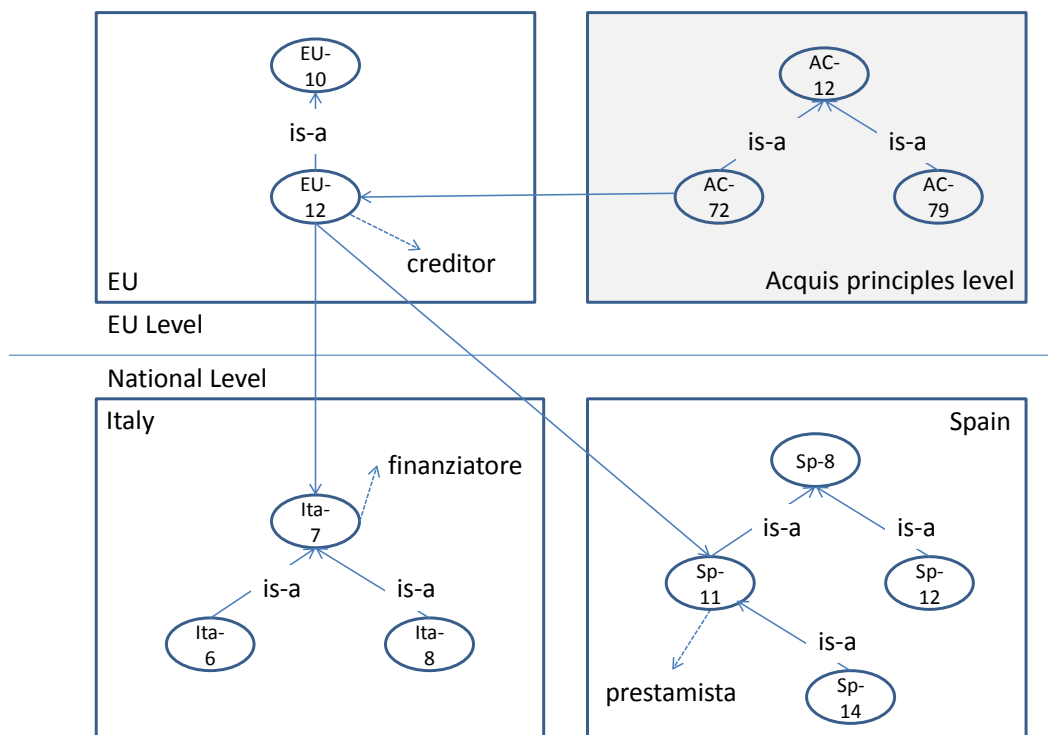

Figure 18. Mapping Acquis Principles concepts to existing European level concepts, which are in turn mapped to national level concepts.

Example. This approach was employed to introduce the *Acquis Communitaire* Principles ontology into ELTS. In Figure 18, the then current EU concept eu12 (linked to the English legal term *"creditor"*) is explicitly linked to the Italian national concept Ita-7 expressed as *"creditore"* and the Spanish national concept Sp-11 expressed as *"prestamista"*. Inserting an explicit link between the *Acquis* concept AC-72 and the EU concept eu12 results in concept AC-72 being implicitly associated with concepts Ita-7 and Sp-11. Allowing paths through more than two ontologies greatly facilitates conceptual and terminological analysis in a dynamic and interconnected environment, offering new insights to scholars in comparative law.

5. ELTS Implementation

Having described the key functionalities of the system, we now discuss key points regarding system implementation. Designing a multi-level ontology system based on contributions from legal experts from various jurisdictions imposes technical requirements which prohibits the use of standard ontology tools:

– The system must be able to separate concepts and terms;
– The system must have a web-based platform in order to obtain contributions from different organisations in different countries; and
– The system must have an user-friendly interface suitable for legal, rather than ontology, experts.

5.1. Ontology Suite

The ELTS ontology framework is based on software from the open-source (and thus modifiable) Gene Ontology project[14], from which it inherits the logical and graphical representation, and the ability to import ontologies and export them to representation languages such as OWL. The Gene Ontology consortium (often referred to as GO) was founded with the aim of creating a consistent terminology from different databases. The resulting tools and databases were then made freely available. The Gene Ontology consortium provides several tools: GO database, whose schema is public; Amigo, a web application for consulting ontologies; and DAG-Edit, a tool for browsing, searching, and modifying ontologies.

In the Gene Ontology, GO terms are viewed as graph nodes and relations as arches with labels. The nodes and arcs are implemented as term and term-to-term relations, while the types of relationships are also stored as terms in a special ontology for cataloguing relation types. European Legal Taxonomy Syllabus manages ontologies in a similar manner, but with two key differences:

1. Relations are not between terms, but between terms and their meanings, and
2. The types of relations are considered not as terms but as separate concepts stored in a special database table.

The decision at the outset to use the code of Gene Ontology project, rather than general-purpose tools for ontologies like Protégé, was based on its relative ease, flexibility and efficiency. Constraints on the ontology are enforced by the interface and by the database rather than by the formalism, which makes it easier to adapt to user needs.

The DB schema of ELTS, which is very close to the ontology schema illustrated above, facilitates reasoning on the database via SQL queries. In Description Logic formalisms such as OWL-DL, this would be harder to model since the reasoning would need to traverse different ontologies (EU and different national ones). Complex queries that require the use of a cycle of several SQL queries on the target ontology are made more efficient by ensuring that the path from each node to all its ancestors is precomputed by reflexive transitive closure.

5.2. Web-based Platform

Since the creation of multi-level ontologies is a coordinated effort between several groups, the system should ideally be easily accessible from different computers without the constraints of any particular operating system or the need to install specialist software. As such, a new web-based interface was designed to manage ontologies via a database server offering the scalability, concurrency, and integrity checks required. The ELTS web application is implemented in PHP.

The design of the system follows the classic three-tier architecture: a database layer, a logic layer and a presentation layer. In the web application, the level of logic is divided into two modules, User Interaction Handler and Business Logic Layer, both of which are divided into several modules designed to perform

---

[14]http://www.geneontology.org

Figure 19. The search interface of ELTS.

specific functions. The classes Smarty[15] and GraphViz[16] are two external projects used to connect the level of User Interaction Handler at the level of the user interface itself. The Smarty class is a vehicle for the interpretation of HTML templates. It comes with a very consistent user interface without having to affect the structure of the main application. Figure 19 shows the search interface of ELTS.

5.3. Security and Integrity

While users interact with ELTS through a web browser, the system has to support concurrent access and modification of data to maintain its integrity. As such, authentication is required to access the editing part of the system, and any changes are recorded in a log file to track what changes were made and who made them. Different levels of authorization can be associated with different users:

1. Site navigation;
2. Comments;
3. Concept management;
4. Relation type management;
5. Full text management;
6. Administration.

---

[15] http://www.smarty.net
[16] http://www.graphviz.org

The system's state is always checked after operations that could potentially lead to an inconsistent state, and there are important constraints on the system. Each term must always be defined by at least one concept, and each concept must always be associated with at least one term. When the user enters a new term, the system checks that there is at least one concept associated with it; likewise, if the user enters a new concept, the system checks that there is at least one term associated with it. If all is well, the system proceeds with the insertion, otherwise the operation is cancelled and an error message is returned.

For the insertion and deletion of ontological relations and associations between national and European concepts, the system checks that the two concepts the user wishes to connect exist and are in the same language.

To enter any ontological relations between different concepts, the following constraints must be met:

- the user must be authenticated and have the necessary permissions to enter a concept.
- the concepts to be linked need to exist in the database.
- the type of ontological relationship specified must exist in the database.
- The two concepts to be linked must be in the same ontology and have the same legal language.

5.4. User Interface

Unlike Gene Ontology's Amigo, the ELTS web interface not only displays dictionaries and ontologies, but also provides a mechanism for creating and deleting terms, definitions and ontologies. ELTS is designed for modeling legal concepts by comparative law methodology. Key aspects of the management of the European Legal Taxonomy Syllabus ontologies are intended to be carried out by legal experts in multiple jurisdictions who do not have specialist IT knowledge. Certain user interface requirements arise from the nature of the domain and user profile. In particular:

- contributors and end-users need user-friendly interfaces to be able to visualize and navigate ontologies at both European and national levels, and the relations between different ontologies. They also need links between ontologies and legislation in order to find patterns of usage and inconsistencies.
- different types of contributors need specific interfaces for their work. Legal experts need an intuitive interface for annotating terms and relating them to concepts with basic semantic relations. Ontology experts can then use a different interface to add more complex relations among the concepts.
- ELTS should allow registered users, who are not authorised to add concepts and terms, to comment in a web 2.0 fashion on the ontology (see the bottom of Figure 6), in order to gather feedback and build knowledge in a shared way.

The development of the content of ELTS ontologies is based on a two-phase procedure. In the first phase, 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 and different conceptualisations are represented separately. Also, terms at the European level are associated with their transpositions in the national languages.

This first phase of defining legal concepts in the context of European directives and national legislation requires contributors to search and read a large number of documents. In order to facilitate this process, the ELTS system includes a database that contains full versions of the relevant European directives and national laws in text format.

After manual inspection of legal documents in the ELTS database, the legal expert can specify correspondences among the terms found, and save the result of his/her analysis in an ontology structure. The workflow is as follows:

- The user creates a new concept named with the term (s)he intends to define;
- The user selects relevant legal texts: the term attached to the concept is used as the default query to retrieve documents;
- After choosing one of the documents in the search results, the full text of the legal document is displayed, with each occurrence of the searched term highlighted;
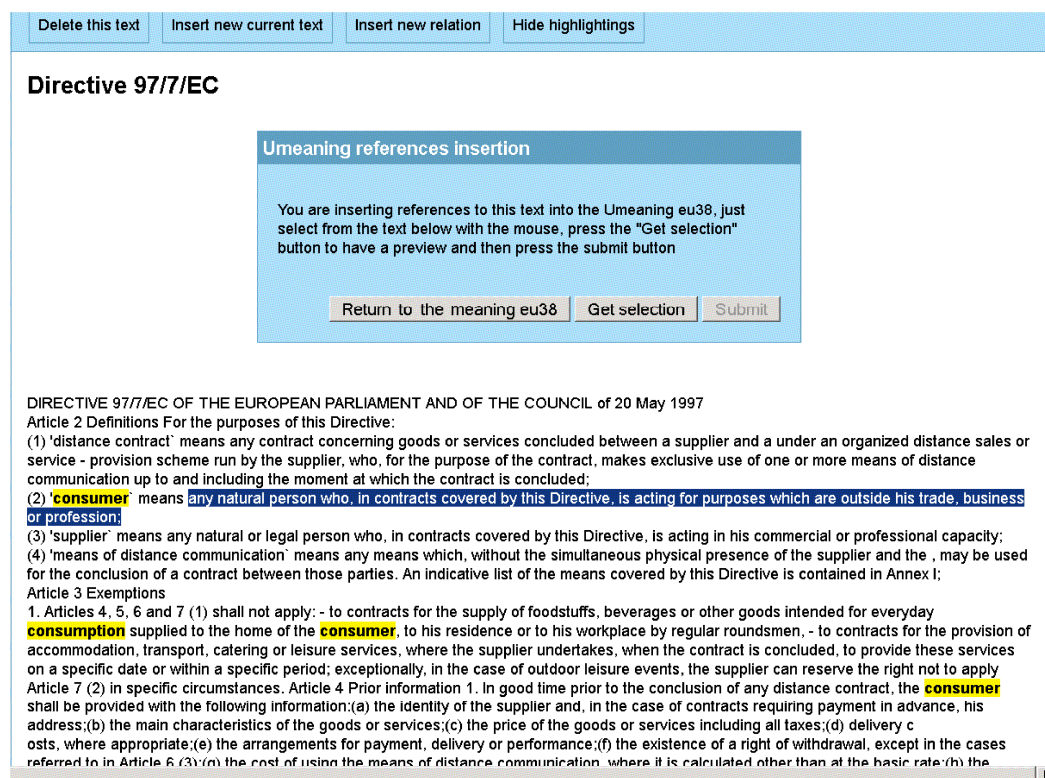
Figure 20. The interface for adding references while editing a concept.

– Finally, the user selects which part of the text should go in the citation and confirms the insertion in the references database.

The links between the concepts, terms and original definitions in legislative sources remain thereafter. This means that should users wish to review the origin of concepts, they can access the documents that provided the definition for a given legal term. The last step of this workflow is illustrated in Figure 20.

An alternative workflow starts from the text of the legislation, where the user can select a part of the text and create a concept from it (see Figure 21). Level (EU or national), language, and references are automatically derived from the metadata of the legislation.

In the second phase, for each different ontology, the set of concepts are organized in a way that is suitable for the relevant jurisdiction. The result is a principled lightweight, but not axiomatic, ontology.

Figure 22 shows the editing interface from the point of view of a user who is allowed to insert and modify concepts. The operations available listed above the concept are: inserting a new reference to legislation (which passes to the legislation search interface and then to a page to select a paragraph in the text of the legislation, as discussed above), inserting domains[17] to which the concept belongs, deleting terms associated with the concept, inserting new terms associated with the concept (thus creating synonyms), creating the description of the concept in another language (only for EU=level concepts), and adding further documents as attachments. The remaining actions allow users to link the current concept to other concepts: linking to concepts implementing the EU level concept in a Member State ontology, inserting a child concept using the ontological relations allowed by the system, inserting parents and children using the *"is-a"* relationship, adding concepts that are similar but not identical to that under consideration (*"quasi-synonym"*), creating concepts that group other concepts due to interpretation, and adding new concepts replacing existing ones.

---

[17]At the moment, ELTS only contains concepts belonging to the consumer law domain; however, the tool allows for the definition of new domains.

Figure 21. The interface for creating a concept from selected text while reading legislation.

Since the tool was designed to be used in multi-lingual settings, the language of the interface can be changed. However, descriptions of concepts in several languages are possible only at the European level. In that case, the system automatically displays the description of the EU concept in the language selected for the interface (see the menu on the right in Figure 5).

The construction of multi-level, multi-language ontologies involves handling large numbers of terms and concepts and attention to preserving consistency. As such, users are provided with tools that display statistics about the ontologies and their terms (see Figure 23).

Similar tools are used to study phenomena concerning legal languages. For example, the system can show polysemy (17% of terms are associated with more than one concept, see Figure 24), synonymity (18% of concepts are associated with more than one term), or how many concepts are transposed into a national system with a term that is different from the one used in the Directive in the same language (above 13%).

6. Current Developments of ELTS

ELTS is the basis for the Eunomos legal document and knowledge management system (Boella et al., 2016), which was originally developed to support regulatory compliance in the context of the ICT4Law[18] project and has been further extended in subsequent years in the context of the EUCases project[19]. It may be used in the context of the ongoing projects ProLeMAS (PROcessing LEgal language in normative Multi-Agent Systems)[20] and MIREL (MIning and REasoning with Legal text)[21].

---

[18] http://www.ict4law.org
[19] http://www.eucases.eu
[20] http://www.liviorobaldo.com/ProLeMAS.htm
[21] http://www.mirelproject.eu.

Figure 22. The concept editor interface of ELTS.

Eunomos, even if it is not multi-lingual and multi-level, has made two advancements to ELTS: firstly, it considers as concepts not only definitions but also (structured) norms, still in a lightweight ontology. Second, the text of the legislation is managed via legislative XML, in particular, the Normeinrete standard (Francesconi, 2006).

In order to help users understand prescriptive as well as constitutive norms, Eunomos includes a complex concept type called 'prescription', with links to a number of different simple concepts. A prescription is defined as a concept which is necessarily connected by relations with other concepts: 1) deontic clause: the type of the prescription: i.e., obligation, prohibition, permission, exception; 2) active role: the addressee of the norm, a concept subsumed by the concept role (e.g., citizen, director); 3) passive role: the beneficiary of the norm; 4) crime: the concept in the ontology of crimes resulting from violation of the prescription (if it is an obligation or prohibition), a concept which is often defined in other legislation, such as the Penal Code.

The constraints between prescriptions and related concepts are not modeled in the ontology, as it is lightweight. Such constraints are rather implemented, as in ELTS, at the insertion interface of a prescrip-

Search results 584 entries		
ID	Legal language	Term
eu3		
	de_DE	Ersatzlieferung
	en_GB	replacement
	es_ES	sustitución
	fr_FR	remplacement
	it_IT	sostituzione
eu4		
	de_DE	Werbung
	en_GB	Advertising
eu6		
	en_GB	adequate
	it_IT	adeguato
eu9		
	de_DE	ausreichend

Figure 23. Statistics about the translation of terms associated with European level concepts in the different languages of Member States.

tion, and maintained by the database. Eunomos is used in the context of regulatory compliance within the Menslegis[22] commercial service for compliance distributed by Nomotika s.r.l.[23], a spinoff of the University of Torino, Italy, in which five of the authors are partners.

Concerning the text of legislation, Eunomos automatically downloads them from Italian legislative portals, transforming them into legislative XML format, in order to structure the text (articles, paragraphs, etc.), to identify references via rule-based techniques drawn from Robaldo et al. (2011a) and Robaldo et al. (2011b), and to classify them via statistical approaches such as developed by Boella et al. (2013b) and Boella et al. (2014). Legislative XML also allows the set of documents to be transformed into hypertext, thereby allowing users to access specific parts of the text (e.g., an article) with a Normeinrete identifier.

ELTS has being further developed within the EUCases project. The main outcome of the EUCases project is a unique pan-European law and case law Linking Platform transforming multilingual legal open data into linked open data after semantic and structural analysis. The web-based EUCases Linking Platform (Boella et al., 2015) provides services linking EU law and case law to legislative acts and court decisions of six EU member states: Austria, Bulgaria, France, Germany, Italy and United Kingdom.

Within EUCases, ELTS is used to retrieve legal documents about consumer law in the EU. Specifically, a multilingual access module allows users to enter full text queries, each of which is then automatically

---

[22]http://www.nomotika.it/IT/MensLegis/Flyer
[23]http://www.nomotika.it

Figure 24. Statistics about polysemic terms.

associated with a list of ELTS concepts. The set of concepts is then expanded by following the ontological relations defined in ELTS. It is possible to set parameters that specify the maximum number of steps for crossing ontological relations as well as the weights to be attributed to each relation; these weights are then used to rank the returned documents for the query. For each of the six supported languages, the legal terminological entries associated with each concept are then searched in the documents in order to identify and rank the terms to be returned to the user.

Moreover, as in Eunomos, but in a multi-lingual and multi-level fashion, the legal documents - legislation, case law and articles from open access journals - are downloaded and converted into a legal XML standard. In this case the standard is Akoma Ntoso[24], the emerging international OASIS (Organization for the Advancement of Structured Information Standards) standard in the field.

The multi-lingual access module developed in EUCases could be extended into a multi-lingual module suitable for legal drafting to support translating from one EU language to another, and this was advocated during the demo presentation of ELTS at the JIAMCATT Joint Inter-Agency Meeting on Computer-Assisted Translation and Terminology conference[25] (Violato, 2011).

---

[24] http://www.akomantoso.org
[25] http://jiamcatt.org

Finally, further research activities on ELTS are currently under study in the context of the ongoing projects ProLeMAS and MIREL. ProLeMAS is a Marie Sklodowska-Curie Individual Fellowship (IF) research project aimed at filling the gap between current logical frameworks designed to represent legal knowledge and the richness of Natural Language Semantics. The project proposes to merge Jerry R. Hobbs's logic for Natural Language Semantics (Hobbs, 1998), which is grounded on the notion of reification (Davidson, 1967) and flat logical structures (Robaldo, 2011), with Input/Output Logic (Makinson and van der Torre, 2000), a well-known framework in deontic logic and normative reasoning. The result is a new logical framework called "Reified Input/Output Logic" (Robaldo et al., 2016).

Furthermore, ProLeMAS aims to implement a concrete NLP pipeline for semi-automatically building formulae in Reified Input/Output Logic from existing legal documents. The pipeline will be possibly integrated with Eunomos. Reification is a flexible formal mechanism to handle ambiguity and vagueness occurring in natural language. Drawing from Robaldo (2010a), Robaldo (2010b), Reified Input/Output Logic is based on flat representations of Natural Language meanings, which facilitates handling of a large variety of linguistic phenomena such as named entities, defeasibility, anaphora, etc. In light of this, we will possibly study potential links between formal items used in the formulae and legal terms included in the European Legal Taxonomy Syllabus.

MIREL is a Marie Sklodowska-Curie Research and Innovation Staff Exchange (RISE) research project, i.e. it funds short-term exchanges for staff belonging to project partners in order to promote networking, knowledge sharing and the skills development of staff members. MIREL involves sixteen academic and industrial partners, with at least one from each continent, including the University of Turin and the spin-off Nomotika s.r.l.. The project will create an international and inter-sectorial network for defining formal frameworks and for developing tools for mining and reasoning with legal texts, with the aim of translating these legal texts into formal representations that can be used for querying norms, compliance checking, and decision support. MIREL addresses conceptual challenges, such as the role of legal interpretation in mining and reasoning, as well as computational challenges, such as the handling of big legal data, and the complexity of regulatory compliance. It bridges the gap between the community working on legal ontologies and NLP and the community working on reasoning methods and formal logic. Moreover, it involves industrial partners in the future development of innovative products and services in legal reasoning and their deployment in the marketplace. Eunomos and the European Legal Taxonomy Syllabus will be involved in both the theoretical and applicative research activities of MIREL.

In the context of the mentioned research projects and user feedback we are collecting from Menslegis customers, the system is currently under revision. We are extending it with new data and we are adding new functionalities to the interface. In particular, while Menslegis has been initially designed to deal with Italian legislation only, as discussed above, in the future we are going to extend its scope at the EU level. Of course, that will possibly require a redesign of the system, as well as the ELTS, which in turn would lead to new validations, user feedback analyses, and overall performance evaluation.

7. Related Work

The most well-known and influential legal ontologies are top-down abstract explorations of legal concepts and their relation to real-world objects, designed for machine reasoning, and informed by legal philosophy and cognitive science. Notable examples are the Frame-Based Ontology of van Kralingen (1995), the Core Legal Ontology (Gangemi et al., 2005), LKIF Core (Hoekstra et al., 2007) and the Functional Ontology of Law of Valente (1995) and Engers et al. (2008). The latter (LRI Core) is informed by Hart, Kelsen and Hohfeld, and Figure 25 shows its top-level categories (cf. also (Breuker, 2004)). These ontology projects focus on abstract high-level concepts first. The rationale for deriving specifications from legal theory, as opposed to legal practice, is that "they have the advantage of having been discussed and debugged in the course of years" (Valente and Breuker, 1994, p.144).

Such an approach to building ontologies is generally known as 'top-down' and is the opposite of the 'bottom-up' approach used in ELTS, where ontologies are initially populated with atomic concepts. Top-
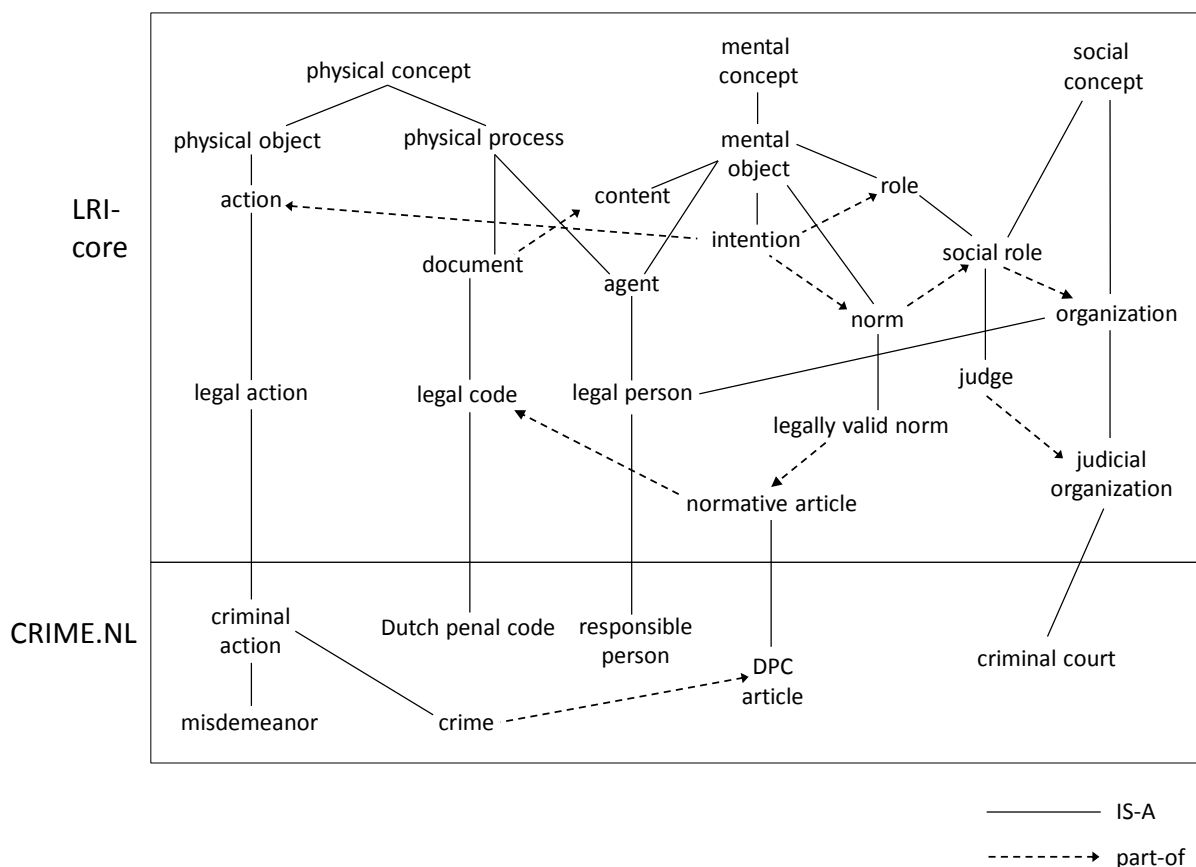
Figure 25. Diagram of a fragment of LRI-core, a top ontology for law from (Engers et al., 2008)

down legal ontologies have not been created for human understanding or for comparing terminologies in European and national jurisdictions. On the other hand, ELTS was specifically designed to be used by legal practitioners. The bottom-up approach is our preference for this scenario, especially because the fact that certain philosophical conceptualisations are accepted in legal theory does not necessarily mean that they are accepted in legal practice.

The bottom-up approach comes at some cost. As Uschold and Grüninger (1996, p.13) explains, "A bottom-up approach results in a very high *level of detail*. This in turn (1) increases overall *effort*, (2) makes it difficult to spot *commonality* between related concepts and (3) increases risks of *inconsistencies* which lead in turn to (4) *re-work* and yet more effort.". Moreoever, ontologies are a costly, resource-intensive endeavour, and the law is a particularly difficult and untidy domain to model. Furthermore, 'abstract' concepts are not precise enough to allow accurate machine reasoning on legal concepts. However, as our examples in Section 2 show, the reality is that legal concepts and their inter-relationships are by no means universal among different jurisdictions.

In the literature, most ontologies used in machine reasoning are universalist legal ontologies. For instance, the LKIF Core Ontology (Hoekstra et al., 2007) set out to model basic concepts of law identified by citizens, legal professionals and legal scholars, as well as top-level concepts (location, time etc) and intentional concepts (beliefs, roles). The system struggles to reason on EU Directive 2006/126 on driving licences, a relatively straightforward regulation, due to the well-known problem of 'resource bottleneck', (Hepp, 2007). One of the biggest challenges for creating ontologies for machine reasoning, as opposed to merely for human understanding, is the level of basic knowledge and inter-connections a machine needs to be provided with. The second challenge is machine reasoning in a social domain. If it's just about possible to model precise terms such as 'at least 16 years old' or 'maximum design speed of not more than 45 km/h' (Hoekstra et al., 2007), open-textured norms are fraught with problems (Rissland and Skalak,

1991), making it almost impossible to determine what is prohibited or permitted by deduction, even from one piece of legislation, let alone consider the effects of other legal formants. As Hart famously pointed out (Hart, 1958, p.607), difficulties can arise even for seemingly straightforward constitutive norms: 'A legal rule forbids you to take a vehicle into a public park. Plainly this forbids an automobile, but what about bicycles, roller skates, toy automobiles? What about airplanes?...We may call the problems which arise outside the hard core of penumbral instances "problems of the penumbra"...If a penumbra of uncertainty must surround all legal rules, then their application to specific cases in the penumbral area cannot be a matter of logical deduction, and so deductive reasoning, which for generations has been cherished as the very perfection of human reasoning, cannot serve as a model for what judges, or indeed what anyone, should do'.

Why do so many legal ontologies attempt to define legal concepts in formal logic? Perhaps it is simply a matter of adhering to the predominant trend of ontology development in general. Hepp (2007, p.5) states that "(f)or a large fraction of ontology researchers, formal logic as a means (i.e., modality) for expressing the semantic account is a constituting characteristic of an ontology" before adding, "[i]n our opinion, it is highly arguable whether formal logic is the only or even the most appropriate modality for specifying the semantics of a conceptual element in an ontology." This is especially true for law, our domain of interest. Legal conceptualisations are social objects (Searle, 1995) and are jurisdiction-, language-, domain- and often statute-specific. The differences between such conceptualisations can only really be understood in the rich expressiveness of natural language since "like language generally, legal discourse can never escape its own textuality" (Peller, 1985, p.1182).

ELTS - we stress again - has been designed from a different perspective and to serve different goals. It is intended to help human users - legal and business professionals, translators and legislators - understand the meaning of legal terms used in European and national legislation, particularly for comparing national laws and the interpretation of European laws in different countries in the European Union.

There are some instruments provided by the European Union which are intended to serve these goals, but they have several limitations. The Eurovoc thesaurus[26] used in European Union institutions is structured on the basis of the EU Treatise rather than on a principled taxonomy of law. That structure denotes the origin of the European Union as an international organisation, but has no connection with the classifications implied in the legal discourse within the Member States and in the national enforcement of EU law. Moreoever, Eurovoc is a multi-lingual terminology of one dimension - the European Union - which ignores the different terminology used in national laws. This drawback is also shared by IATE [27]. It is clear that a more systematic way of organising legal terminology is needed, and the most advanced technology for this is ontologies.

The European Court of Justice shares many of our aims in dealing with EU level and national level terminology with its project *"Vocabulaire juridique multilingue comparé: contribution de la Direction générale de la traduction de la Cour de justice de l'Union europénne"*[28], but it does not use ontologies.

Another project at the EU level concerning terminology is LISE (Legal Language Interoperability Services) FP7 EU project[29], which addresses the urgent need for consolidated administrative nomenclatures and legal terminologies as tools to enhance interoperability and cross-border collaboration. Despite similarities with ELTS, ontologies are also not used in this case.

Similarly, Legivoc[30] (Vibert et al., 2013) is an Internet-based platform dedicated to the promotion, editing and alignment of legal vocabularies. Funded in part by the EU Commission and administered by the French Ministère de la Justice, it provides tools to help align terms belonging to different vocabularies. It includes the vocabularies of 15 countries, which have been translated into more than 20 languages. However, as for LISE, Legivoc is not an ontology.

---

[26] http://eurovoc.europa.eu. See (Boella et al., 2013a) for an example of application using Eurovoc.
[27] http://iate.europa.eu/
[28] http://www.legalaccess.eu/?Vocabulaire-juridique-multilingue-156
[29] http://www.lise-termservices.eu
[30] legivoc.org

Recent years has witnessed legal ontologies designed primarily for information retrieval which happen to also provide textual explanations of terms. One could say they have the same goals as the European Legal Taxonomy Syllabus in reverse priority. They have been developed using a bottom-up approach, since they first and foremost need to include terms that actually appear in the documents to be retrieved.

OntoPrivacy (Cappelli et al., 2007) is an ontology based on a glossary of keywords extracted from the Italian Personal Data Protection Code and includes legal, public administration and generic words. The terms are organised into a lightweight ontology containing textual definitions while also identifying synonymy, hypernymy and hyponymy relations. The trees are then linked to a top-level ontology, – inspired by LRI Core, the Core Legal Ontology and Legal-RDF Vocabularies – with five main classes: events, scenes (where and when an event occurs), physical objects, non-physical objects and roles (played by individuals or groups). The work is interesting in that it bridges the gap between legal glossaries and legal ontologies. It is also similar to our approach in that the data was collated by legal experts. However, it is not clear which legal formants apart from legislation were used to define the contents. They do not mention the European dimension, which has had a significant impact on this area of law in Italy. It is also not clear whether, and if so how they deal with polysemy, terminological variation, evolution in the meaning of legal terms or the issue of alternative legal interpretations of legal terms.

Some proposals, e.g., JurWordNet (Gangemi et al., 2003) and the LOIS ontology (Tiscornia, 2006), among others, are aimed at extending and adapting Wordnet (Miller, 1995) for the legal domain. WordNet is an extensive general glossary that has been used as an ontology for information extraction due to its extensive coverage. Concepts and terms are separated, with terms organised into sets of synonyms (known as synsets) mapping to different contextual concepts (known as senses). Of course, Wordnet cannot be used for information extraction on legal texts because the synsets refer to the generic lexicon used in natural language, not to the specific lexicon used in "legalese".

On the other hand, the Wordnet schema has some limitations in expressivity that need to be overcome in order to properly model legal concepts. Several semantic relations in Wordnet denote broader meanings than the ones required for the legal domain; even the most basic 'IS_A' relation is used both for instances (e.g., Beethoven 'IS_A' composer), hypernyms (e.g., songwriter 'IS_A' composer), and roles (e.g., dog 'IS_A' worker_animal). Another limitation imposed by WordNet is that relations are lexical rather than semantic, which means that relations such as 'purpose', 'grouped' or 'replaced_by' are not allowed.

Some of the limitations of Wordnet have been addressed in the work of Gangemi et al. (2002), where the well-known DOLCE ontology was introduced for the first time, and where it has been shown that by using DOLCE, together with the OntoClean methodology (Guarino and Welty, 2002), it is possible to create a more 'principled' ontology of WordNet, correcting the use of the same relation names for different types of relations.

JurWordNet (Gangemi et al., 2003), a legal ontology of Italian legal terms, is organised in accordance with more precise relations, identified via DOLCE & OntoClean principles. JurWordNet was also developed primarily for information extraction. Other stated goals are semantic tagging of legislative texts, question answering, automatic tagging, norm comparison and knowledge sharing. It also carefully connects legal concepts to layman concepts, since legal concepts are often specialisations of generic layman's terms. The ontology was specifically constructed as an extention of ItalWordNet (Roventini et al., 2000) and EuroWordNet (Vossen, 1998), which are in turn extensions of WordNet (Miller, 1995).

The LOIS ontology (Tiscornia, 2006) is a multi-lingual version of JurWordNet. It contains monolingual WordNets linked by equivalence, near-equivalence or subsumption relations to an English synset in the Inter-Lingual-Index (ILI). Synsets in other languages refer to glosses in the English ILI to get their meaning.

The LOIS ontology is the most similar ontology to ELTS in its multi-linguality, domain, separation of concepts and terms, and handling of polysemy. Both ontology frameworks contain terms from European and national jurisdictions. The main difference is that ELTS has a separate ontological structure for every jurisdiction, with terms in different languages mapping to jurisdiction-specific definitions, whereas the LOIS ontology has one conceptual structure to which all terms in the different languages are mapped.

This is an important limitation inherited from EurWordNet which can impede accurate modeling of differences in conceptions across different jurisdictions. The LOIS ontology works around this limitation by connecting to DOLCE2.1-LitePlus + CLO (Gangemi et al., 2005) where it duplicates LOIS concepts and connects them to semantic relations. Some lexical relations in WordNet are practically the same as DOLCE and CLO semantic relations (e.g., hyponym and sub-class-of), which introduces the risk of mis-alignment. Of course, ontologies "are dependent on the purposes they are made for" (Chandrasekaran et al., 1998), and the strengths of the LOIS ontology may well outweigh its weakness in the context of its main purpose - monolingual and cross-lingual information retrieval.

The main structural weakness of the LOIS ontology is addressed in later work (although that work has a different purpose). The Consumer Protection Ontology within the DALOS Project (Francesconi et al., 2007), which builds on the LOIS ontology, acknowledges that "in a legal domain one cannot transfer the conceptual structure from one legal system to another" (ibid, page 104) and, like ELTS, creates one conceptual layer for EU legal concepts which is mapped to lexical layers in different languages. The intended aim of the ontology is to support multi-lingual EU legislative drafting.

However, the system described does not include a conceptual layer for national jurisdictions. We believe that the multi-jurisdictional framework of ontologies in ELTS offers a more comprehensive system for human understanding of legal terms in European and national jurisdictions. Legal experts can explore differences in different jurisdictions, and translators can gain greater awareness of different conceptions at European and national levels to help them understand the impact of their translation choices.

The EU Employee Legal Ontology by Desprès and Szulman (2007) is also related to this section. It is a precise domain ontology about employment law in European Directives developed with a bottom-up approach. Terms were selected from two European directives about employees, and structured into two different micro-ontologies. The ontologies contain terminological concepts and non-terminological concepts. Terminological concepts are the terms themselves, with legal definitions in natural language, and a list of synonyms. Non-terminological concepts are labelled relation between terminological concepts, either structural (classic relations such as hypernyms or meronyms) or functional (domain-specific relations, e.g., 'is protected by national employment law'). The two micro-ontologies were then merged, and connected by subsumption to higher-level concepts defined formally in the LRI-Core Legal Ontology (Breuker, 2004), and the DOLCE ontology (Borgo and Masolo, 2009). The use of textual and formal elements in this work is indicative of a dual purpose. The work was initiated by French jurists who wanted to compare the use of the concept 'employee' within the Directives before their transposition into national law. However, its stated aim is to support the development of formal models of legislation to be used in legal knowledge-based systems. The way in which legal knowledge-based systems might use the ontology is not described, neither is the issue of open-textured norms. More interestingly in the context of the focus of the present paper, the relationship between European and national terminology is not explored, even though the Directives stipulate that employees are subject to employment law in national jurisdictions. No language other than French is considered, even though the meaning of European terms can be influenced by terminology in other languages and jurisdictions. Neither are other legal formants (jurisprudence, legal scholarship) consulted in exploring the meaning of concepts. Definitions for terms are taken only from the Directives, which are regarded as 'self-contained and cohesive' (Desprès and Szulman, 2007, p.191), without any input from legal experts save for final evaluation. Moreoover, while the ontology has the expected classical ontological relations, and is also rich in domain-specific relations, relations between different conceptions of the same terms are not described, e.g., after merging definitions of terms, is it still possible to access the more context-specific definitions in the original micro-ontologies? When terms are redefined due to amendments to Directives, is it still possible to access earlier conceptualisations?

We conclude this section with a brief mention of the Ontology of Professional Judicial Knowledge (Casanovas et al., 2007). Although it explores a different legal concern (the judicial view of legal procedures in Spain), it has important similarities to ELTS in its emphasis on the collaborative involvement of legal experts in populating the ontology. OPJK explores differences in interpretation (in this case, via a wiki) before reaching generalised definitions, which can be compared to the concept abstraction mechanism in ELTS (see Section 4.5 above). Discarded conceptualisations are archived, and reasons for their

archival recorded. ELTS also has a mechanism for marking conceptualisations obsolete but available to view, via the 'replaced_by' mechanism described in Section 4.4 above.

8. Conclusions

This paper outlines a methodological approach, schema and tool for building a multi-lingual, multi-level legal ontology to help legal professionals involved in transnational activities, or multinational organisations, manage the deep, complex and interconnected knowledge required for understanding laws in different jurisdictions.

The first contribution of this paper is an analysis of the legal domain, the observation that 'deep' understanding of legal systems is necessary to bring out explicit conceptualisations, and the promotion of a comparative law methodology for building ontologies in the context of EU law. Very few legal ontology research is based on such a thorough and pragmatic methodology (though the suggestion is not new - see (Schäfer, 1998)). Indeed, many legal ontologies are built under universalist assumptions that, in our view, render them disconnected from legal practice. Of course, the universalist view of law has a long and distinguished history in legal philosophy. Natural law theory, whose proponents include Aristotle, Thomas Aquinas and Grotius, held that human law in all jurisdictions derive from one root. However, comparative law has a different goal: to bring out different interpretations within the same, as well as different, jurisdictions, in order to achieve deeper understanding of legal concepts and legal reasoning in practice.

Our analysis adopts a bottom-up approach to the construction of legal ontologies, rather than a top-down approach. The top-down approach represents a viable solution for the implementation of formal ontologies, and was indeed used in other proposals, but, given the multi-lingual character of ELTS, it would seem that the bottom-up approach represents a better choice for our purposes. As Winter (1989, p.1134) states, "more abstract concepts are likely to show the greatest cultural variation because they are indirect, imaginative extensions of the physical and social experiences that ground them."

The second contribution is the design of a multi-level ontology schema based on insights from comparative law, with innovative extensions that transcend the limitations of standard ontologies in the domain of law:

- The varying and highly contextualised meaning of legal terms means that there needs to be a structured way to allow more than one meaning in a legal ontology: the ELTS solution is to separate terms and meanings, allowing terms to be mapped to different meanings, and with each meaning linked to its legal sources - legislation, case law and/or doctrine - and related to a particular subdomain.
- The multi-jurisdictional nature of the EU requires not only a multi-lingual ontological framework, but also a multi-level one: the ELTS schema involves ontologies for each national jurisdiction and for the EU jurisdiction, and concepts are mapped to terms in relevant languages and connected by ontological relations. Implementation and translation relations connect concepts across ontologies.
- Some legislation brings new definitions of terms and interpretation can change their meaning, so that new concepts effectively replace prior conceptualisations; yet researchers may need to refer to deprecated conceptualisations for historical purposes or to trace the evolution of terms: ELTS does not remove deprecated conceptualisations, they are simply linked to new conceptualisations with a time-stamped *"replaced_by"* relation.
- The frequent merging of meanings assigned to legal terms that takes place in legal reasoning or the simultaneous transposition of multiple Directives means that concepts with a wider scope of application can emerge that do not necessarily replace more specific ones defined by legislation in all situations: ELTS handles this with a specific relation called *"grouped"* connecting context-generalised concepts to more context-specific concepts.
- Glossaries created to achieve consistency in legal terminology may contain conceptualisations that are yet to be accepted officially: ELTS allows the creation of temporary legal ontologies whose concepts are linked to current legal ontologies until such time as the old concepts are replaced.

The third contribution is the implementation of the above schema in the development of a user-friendly ontology tool. We chose to use a lightweight ontology, with no formal semantics, to make it suitable for obtaining contributions from legal experts with little or no specialist knowledge of ontologies, and with a web-based platform to enable collaboration from various institutions.

Finally, ELTS was used to build knowledge about consumer law in five national jurisdictions and the European Union, and this paper includes examples from this data which validates our methodological approach. We agree with the statement that '[a] jurist ... is the only person competent to decide whether two ideas taken from different legal systems correspond to each other and whether a difference in rules is tantamount to a difference in concepts' (Sacco, 1991, p.14). ELTS's web-based platform and user-friendly interface renders it suitable for the involvement of legal experts on a continuous basis. We used as a benchmark the ability of our framework to model a detailed European-wide legal terminology analysis developed in the European project Uniform Terminology[31] in the domain of consumer law. Comparative lawyers from the University of Turin were involved in the design, population and updating of ELTS. Comparative lawyers from several leading European universities from the Uniform Terminology project produced the term descriptions for different European jurisdictions, with annotations about influential jurisprudence and scholarly doctrine.

European Legal Taxonomy Syllabus has been used in further developments. It is at the basis of the Eunomos system (Boella et al., 2016) on which the commercial regulatory compliance service Menslegis is built, and is used in the EUCases project to help search for case law in several EU countries.

References

Boella, G., Di Caro, L., Graziadei, M., Cupi, L., Salaroglio, C. E., Humphreys, L., Konstantinov, H., Marko, K., Robaldo, L., Ruffini, C., Simov, K., Violato, A., and Stroetmann, V. (2015). Linking legal open data: Breaking the accessibility and language barrier in european legislation and case law. In *Proceedings of the 15th International Conference on Artificial Intelligence and Law*, ICAIL '15, pages 171–175, New York, NY, USA. ACM.

Boella, G., Di Caro, L., Humphreys, L., Robaldo, L., Rossi, R., and van der Torre, L. (2016). Eunomos, a legal document and knowledge management system for the web to provide relevant, reliable and up-to-date information on the law. *Artificial Intelligence and Law*, to appear.

Boella, G., Di Caro, L., Rispoli, D., and Robaldo, L. (2013a). A system for classifying multi-label text into EuroVoc. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Law*, ICAIL '13, pages 239–240, New York, NY, USA. ACM.

Boella, G., Di Caro, L., and Robaldo, L. (2013b). *Semantic Relation Extraction from Legislative Text Using Generalized Syntactic Dependencies and Support Vector Machines*, pages 218–225. Springer Berlin Heidelberg, Berlin, Heidelberg.

Boella, G., Di Caro, L., Ruggeri, A., and Robaldo, L. (2014). Learning from syntax generalizations for automatic semantic annotation. *Journal of Intelligent Information Systems*, 43(2):231–246.

Boella, G., Humphreys, L., Martin, M., Rossi, P., and van der Torre, L. (2012). Eunomos, a legal document and knowledge management system to build legal services. In *AI Approaches to the Complexity of Legal Systems. Models and Ethical Challenges for Legal Systems, Legal Language and Legal Ontologies, Argumentation and Software Agents*, pages 131–146. Springer, Berlin.

Boella, G., van der Torre, L., and Verhagen, H. (2007). Roles, an interdisciplinary perspective. *Applied Ontology*, 2(2):81–88.

Borgo, S. and Masolo, C. (2009). Ontological foundations of dolce. In Staab, S. and Studer, R., editors, *Handbook on Ontologies (Second Edition)*, pages 361–382. Springer, Berlin.

Breuker, J. (2004). Constructing a legal core ontology: LRI-core. In *Proceedings of WONTO 2004*.

Buitelaar, P. (2010). Ontology-based semantic lexicons: mapping between terms and object descriptions. In ren Huang, C., Calzolari, N., Gangemi, A., Lenci, A., Oltramari, A., and Prevot, L., editors, *Ontology and the Lexicon*, pages 212–223. Cambridge University Press. Cambridge Books Online.

Cappelli, A., Lenzi, V. B., Sprugnoli, R., and Biagioli, C. (2007). Modelization of domain concepts extracted from the Italian privacy legislation. In *Proceedings of the 7th International Workshop on Computational Semantics, Tilburg, the Netherlands 2007*.

Casanovas, P., Casellas, N., Tempich, C., Vrandečić, D., and Benjamins, R. (2007). OPJK and DILIGENT: ontology modeling in a distributed environment. *Artificial Intelligence and Law*, 15(2):171–186.

Chandrasekaran, B., Josephson, J., and Benjamins, R. (1998). The ontology of tasks and methods. In *Proceedings of the Banff Knowledge Acquisition Workshop*.

Commission of the European Communities (2003). A more coherent European contract law an action plan. *COM(2003) 68 final*.

---

[31]http://www.uniformterminology.unito.it

Davidson, D. (1967). The logical form of action sentences. In Rescher, N., editor, *The Logic of Decision and Action*. University of Pittsburgh Press.

De Groot, G. (1999). Das Ubersetzen juristischer Terminologie. In Groot, D. and Schulze, R., editors, *Recht und Ubersetzen*, page 11 ff, Baden-Baden. Nomos.

Desprès, S. and Szulman, S. (2007). Merging of legal micro-ontologies from European directives. *Artificial Intelligence and Law*, 15(2):187–200.

Ehrmann, M., Cecconi, F., Vannella, D., McCrae, J. P., Cimiano, P., and Navigli, R. (2014). A multilingual semantic network as linked data: lemon-BabelNet. In *Proceedings of the 3rd Workshop on Linked Data in Linguistics*.

Endicott, T. (2008). *The Value of Vagueness*.

Engers, T., Boer, A., Breuker, J., Valente, A., and Winkels, R. (2008). Ontologies in the legal domain. In *Digital Government*, volume 17 of *Integrated Series In Information Systems*, pages 233–261. Springer, Berlin.

Falkner, G., Hartlapp, M., Leiber, S., and Treib, O. (2004). Non-compliance with EU directives in the member states: Opposition through the backdoor? *West European Politics*, 27(3).

Francesconi, E. (2006). The Norme in Rete project: Standards and tools for Italian legislation. *International Journal of Legal Information*, 34(2):358–376.

Francesconi, E., Spinosa, P., and Tiscornia, D. (2007). A linguistic-ontological support for multilingual legislative drafting: the DALOS project. In *Proceedings of the Workshop on Legal Ontologies and Artificial Intelligence Techniques - LOAIT*, pages 103–111.

Gangemi, A., Guarino, N., Masolo, C., Oltramari, A., and Schneider, L. (2002). Sweetening ontologies with DOLCE. In *Knowledge engineering and knowledge management: Ontologies and the semantic Web*, pages 166–181. Springer, Berlin.

Gangemi, A., Sagri, M., and Tiscornia, D. (2005). A constructive framework for legal ontologies. In *Law and the Semantic Web*, volume 3369 of *Lecture Notes in Computer Science*, pages 97–124. Springer, Berlin.

Gangemi, A., Sagri, M. T., and Tiscornia, D. (2003). Jur-Wordnet, a source of metadata for content description in legal information. In *Proceedings of the workshop on legal ontologies & web based legal information management, part of The international conference of artificial intelligence and law (ICAIL 2003), Edinburgh*.

Giunchiglia, F. and Zaihrayeu, I. (2009). Lightweight ontologies. In *Encyclopedia of Database Systems*, pages 1613–1619. Springer, Berlin.

Graziadei, M. (2004). Tuttifrutti. In Birks, P. and Pretto, A., editors, *Themes in Comparative Law*. Oxford University Press, Oxford.

Gruber, T. (1993). A translation approach to portable ontology specifications. *Knowledge acquisition*, 5(2):199–220.

Guarino, N. and Welty, C. (2002). Evaluating ontological decisions with OntoClean. *Communications of the ACM*, 45(2).

Hart, H. L. A. (1958). Positivism and the separation of law and morals. *Harvard Law Review Vol. 71 No. 4*.

Hepp, M. (2007). Possible ontologies. how reality constrains the development of relevant ontologies? *Internet Computing, IEEE, Volume 1, Issue 1*, pages 90–96.

Hobbs, J. (1998). The logical notation: Ontological promiscuity. In *Chapter 2 of Discourse and Inference*. Available at http://www.isi.edu/~hobbs/disinf-tc.html.

Hoekstra, R., Breuker, J., Di Bello, M., and Boer, E. (2007). The LKIF Core ontology of basic legal concepts. In *Proceedings of the Workshop on Legal Ontologies and Artificial Intelligence Techniques (LOAIT 2007)*.

Hoijer, H. (1954). The Sapir-Whorf hypothesis. *Language in culture*, pages 92–105.

Kelsen, H. (1941). The pure theory of law and analytical jurisprudence. *Harvard Law Review*, 55.

Kelsen, H. (1992). *An Introduction to the Problems of Legal Theory*. Oxford University Press, Oxford.

Legrand, P. (2006). Antivonbar. *Journal of Comparative Law*, 1(1):13–40.

Liebwald, D. (2012). Law's capacity for vagueness. *International Journal for the Semiotics of Law Vol. 25 No. 4*, pages 29–50.

Makinson, D. and van der Torre, L. W. N. (2000). Input/output logics. *Journal of Philosophical Logic*, 29(4):383–408.

Mandelkern, D. (2001). Mandelkern group on better regulation final report. *Paris: EU*.

Margolis, E. and Laurence, S. (1999). *Concepts: Core Readings*. Bradford books. BRADFORD BOOK.

McCrae, J., Spohr, D., and Cimiano, P. (2011). Linking lexical resources and ontologies on the semantic web with lemon. In *Proceedings of the 8th Extended Semantic Web Conference on The Semantic Web: Research and Applications - Volume Part I*, ESWC'11, pages 245–259, Berlin, Heidelberg. Springer-Verlag.

Miller, G. (1995). WordNet: a lexical database for English. *Communications of the ACM*, 38(11):39–41.

Palmirani, M. and Brighi, R. (2006). Time model for managing the dynamic of normative system. In *Electronic Government*, pages 207–218. Springer, Berlin.

Peller, G. (1985). The metaphysics of American law. *California Law Review Vol. 73*, pages 1151–1290.

Poscher, R. (2011). *Ambiguity and Vagueness in Legal Interpretation*. Oxford Handbook on Language and Law. Oxford University Press.

Pozzo, B. (2003). Harmonisation of European contract law and the need of creating a common terminology. *European Review of Private Law*, 6:754–767.

Rawls, J. (1955). Two concepts of rules. *The Philosophical Review*, 64:3–32.

Rissland, E. and Skalak, D. (1991). Cabaret: Rule interpretation in a hybrid architecture. *International Journal of Man-Machine Studies*, 34(6):839–887.

Robaldo, L. (2010a). Independent set readings and generalized quantifiers. *The Journal of Philosophical Logic*, 39(1):23–58.

Robaldo, L. (2010b). Interpretation and inference with maximal referential terms. *The Journal of Computer and System Sciences*, 76(5):373–388.

Robaldo, L. (2011). Distributivity, collectivity, and cumulativity in terms of (in)dependence and maximality. *The Journal of Logic, Language, and Information*, 20(2):233–271.

Robaldo, L., Caselli, T., and Grella, M. (2011a). Rule-based creation of timeml documents from dependency trees. In *Proceedings of the 12th International Conference on Artificial Intelligence Around Man and Beyond*, AI*IA'11, pages 389–394, Berlin, Heidelberg. Springer-Verlag.

Robaldo, L., Caselli, T., Russo, I., and Grella, M. (2011b). From italian text to timeml document via dependency parsing. In *Computational Linguistics and Intelligent Text Processing - 12th International Conference, CICLing 2011, Tokyo, Japan, February 20-26, 2011. Proceedings, Part II*, pages 177–187.

Robaldo, L., Humphreys, L., Sun, L., Cupi, L., Santos, C., and Muthuri, R. (2016). Combining input/output logic and reification for representing real-world obligations. In *Post-proceedings of 9th International Workshop on Juris-informatics (JURISIN 2015). Lecture Notes in Artificial Intelligence*.

Roventini, A., Alonge, A., Calzolari, N., Magnini, B., and Bertagna, F. (2000). ItalWordnet: a large semantic database for Italian. In *Proc. of 2nd International Conference on Language Resources and Evaluation, LREC*.

Russel, B. (1923). Vagueness. *Australasian Journal of Philosophy and Psychology*, 1.

Sacco, R. (1987). La traduction juridique : un point de vue italien. *Les Cahiers du Droit, 28(4):845 ff*.

Sacco, R. (1991). Legal formants: A dynamic approach to comparative law installments I and II. *American Journal of Comparative Law*, 39.

Sacco, R. (1999). Contract. *European Review of Private Law*, 2:237–240.

Sartor, G., Casanovas, P., Biasiotti, M., and Fernndez-Barrera, M. (2013). *Approaches to Legal Ontologies: Theories, Domains, Methodologies*. Law, Governance and Technology Series, Vol. 1. Springer Verlag, Dordrecht, Heidelberg, London, New York, 2013.

Schäfer, B. (1998). Inheritance principles and the community of heirs. *Formal Ontology in Information Systems, Amsterdam, IOS, Amsterdam*, pages 108–121.

Searle, J. R. (1995). *The Construction of Social Reality*. Free Press, New York.

Smith, B. (p.155, 2008). Ontology. In *The Blackwell guide to the philosophy of computing and information*. Wiley.

Soames, S. (2012). Vagueness in the law. In Andrei, M., editor, *The Routledge Companion to Philosophy of Law*, page 95. Routledge.

Spinosa, P., Giardiello, G., Cherubini, M., Marchi, S., Venturi, G., and Montemagni, S. (2009). NLP-based metadata extraction for legal text consolidation. In *Proceedings of the 12th International Conference on Artificial Intelligence and Law ICAIL*, pages 40–49. ACM.

Tiersma, P. M. (2006). Some myths about legal language. *Law, Culture and the Humanities*, 2(1):pp. 29–50.

Tiscornia, D. (2006). The LOIS project: Lexical ontologies for legal information sharing. In *Proceedings of the V Legislative XML Workshop*, pages 189–204.

Uschold, M. and Grüninger, M. (1996). Ontologies: Principles, methods and applications. *Knowledge Engineering Review*, 2(2).

Valente, A. (1995). *Legal knowledge engineering: A modelling approach*. IOS Press, Amsterdam.

Valente, A. and Breuker, J. (1994). Ontologies: the missing link between legal theory and AI & Law. In *Legal knowledge based systems JURIX 94: The Foundation for Legal Knowledge Systems*, pages 138–150. Lelystad: Koninklijke Vermande.

van Kralingen, R. W. (1995). *Frame-based Conceptual Models of Statute Law*. Kluwer Law International.

Vanderlinden, J. (1995). *Comparer les droits*. Story-Scientia, Bruxelles.

Vibert, H.-J., Jouvelot, P., and Pin, B. (2013). Legivoc - connectings laws in a changing world. *Journal of Open Access to Law*, 1(1).

Villegas, M. and Bel, N. (2015). PAROLE/SIMPLE 'lemon' ontology and lexicons. *Semantic Web*, 6(4):363–369.

Violato, A. (2011). Legal taxonomy syllabus. In *Demo at JIAMCATT Conference (the Joint Inter-Agency Meeting on Computer-Assisted Translation and Terminology at the United Nations Office)*.

Vossen, P. (1998). *EuroWordNet: a multilingual database with lexical semantic networks*. Kluwer Academic, Boston.

Wahlgren, P. (1994). A general theory of Artificial Intelligence and Law. In *Legal knowledge based systems JURIX 94: The Foundation for Legal Knowledge Systems*, pages 79–83.

Winter, S. L. (1989). Transcendental nonsense, metaphoric reasoning, and the cognitive stakes for law. *University of Pennsylvania Law Review*, 137(4):1105–1237.

Zhang, Q. (1998). Fuzziness-vagueness-generality-ambiguity. *Journal of pragmatics*, 1.