Scientific
Research

Legal Concepts and Multilingual Contexts in Digital Information

Daniela Tiscornia*, Maria Teresa Sagri

Institute of Legal Information Theory and Techniques (ITTIG-CNR), Italian National Research Council, Florence, Italy.
Email: *tiscornia@ittig.cnr.it, sagri@ittig.cnr.it

ABSTRACT

Analysis of legal documents cannot be reduced to their linguistic expression. When dealing with legal texts a distinction should be made between the linguistic form (normative statement) and the content (norm). Norms are conceived as the interpreted meaning of linguistic contexts (partitions in a legal text, fragments of judicial decisions, etc). In the interpretative process, legal practitioners make reference to an external system of concepts that can be assumed, even if not universal, at least shared by a large social and cultural community. As a consequence, legal interpretation and legal conceptualization are strongly influenced by cultural, social and political factors and changes in meaning of legal concepts occur within a synchronic localization in different cultures and within a diachronic evolution process of the environment in which they are created. This dynamic nature of legal knowledge poses serious problems in communicating legal information, and even more in the realm of digital communication, where, in order to manage knowledge across national borders, there is a strong demand of shared vocabularies embedding a shared understanding of legal concept. This paper aims at showing how lexical/terminological resources and conceptual structures, such as ontologies, can be constructed by means of the innovative tools provided by ICT, like the Semantic Web languages, the computational linguistic and the ontology engineering techniques. Such semantic resources may have a strategic role in representing meaning evolution and in investigating social and cultural influences in linguistic uses and, as a further goal, solutions offered by ICT for bridging the gap between form and content can also provide innovative visions of theoretical issues.

Keywords: Legal Translation; Legal Concepts; Cross-Lingua Retrieval; Multilingual Semantic Web

1. Introduction

The semantic of law has always been a "hot" topic in the debate among legal scholars, involving critical issues like the role of interpretation, the textuality of law and the existence of a *core* set of fundamental constituents with a stable meaning. This last point has been denied by modern legal theory, by arguing that basic notions of law as, for instance "ownership" or "responsibility" are questionable because they are subject to historical changes and evolutions, just like everything else is. The idea of questioning what has been considered for many years "given" in a certain context, might be a useful way of finding new solutions to new problems and questions. Since nothing is isolated but almost everything is connected with global reality, changes in the world will affect even old beliefs and well-established ideas and, in addition, will produce different outcomes depending on the previously existing context. A recently emerging approach within legal theory proposes to apply evolutionary science methods to law.[1] Whether we agree or not

with this proposal, it demonstrates the need for applying an evolving perspective to legal issues. This innovative position affects the study of language as well, because the dynamic nature of legal concepts has its counterpart in the evolution of meaning in legal terminologies. The traditional "univocity" ideal that assumes, in technical domains, a strict correspondence between term and concept are inadequate to fit the complexity of law, while terminological initiatives based on a socio-cognitive approach,[2] that takes into account the interaction between language, the mind and the social world enable the dynamic aspects to be managed.

This paper aims at showing how lexical/terminological resources and conceptual structures, such as ontologies, can be constructed by means of the innovative tools provided by ICT, like the Semantic Web languages, the computational linguistic and the ontology engineering techniques. Such semantic resources may have a strategic

---

*Corresponding author.

[1]See, among others [1].
[2]"Traditional terminology believes that the concept system is to be seen as independent from the term system, and that consequently, unlike words, terms are context independent: the meaning of a term is a concept" [2].

role in representing meaning evolution and in investigating social and cultural influences in linguistic uses and, as a further goal, solutions offered by ICT for bridging the gap between form and content can also provide innovative visions of theoretical issues.

The article is structured as follows: The first part (Sections 1-3) outlines the *problems*: the systematic complexity of the regulatory world, the gaps between common sense meanings in ordinary language and technical meanings in legal language, the challenge of multilingualism and the need for conceptual coherence at transnational level. The second part (Sections 4 and 5) explains how semantic technologies can offer and propose *solutions* to many of the problems that have been outlined and, how formal data model set out by ontology engineering can give relevant contributions to the clarifications of theoretical questions in the domain of language analysis and legal theory.

2. Legal Concepts and Legal Language

Legal knowledge strictly depends on its linguistic expression: the law has to be communicated and social and legal rules are mainly transmitted through their oral and written expression. Even in customary law there is almost always a phase of verbalization that enables legal rules to be identified or recognized; even if the law cannot be reduced to the language that expresses it, nonetheless, it cannot escape its textual nature.

Despite the strict connection, law and language are two autonomous but structurally similar systems: both are endowed with rules that underlie the construction of the system itself, guide its evolution and guarantee its consistency. Both are conditioned by the social dimension in which they are placed, whereby they dynamically define and fix their object in relation to a continually evolving social context.

Let us consider the creative power of the legislator in creating new legal entities: legislative definitions have a *constitutive force*, so we can assume that, for any new definition, a new concept is added in the legal system. Definitions affect lexical meaning of terms in different ways: delimiting, extending and narrowing [3]. Delimiting definitions refer to lexical meaning, but they remove some of the vagueness, for instance in Directive 2008/104, Art.3 "worker: any person who, in the Member State concerned, is protected as an employee under national employment law and in accordance with national practice". Extending definitions expand the scope of a term with respect to its common meaning. The typical case is that of deeming provisions: "Directive 1985/374, Art. 2: Definition of Producer: The manufacturer of a finished product, the producer of any raw material or the manufacturer of a component part and any person, who, by

putting his name, trade mark or other distinguishing feature on the product presents himself as its producer. Without prejudice to the liability of the producer, any person who imports into the Community a product for sale, hire, leasing or any form of distribution in the course of his business shall be deemed to be a producer within the meaning of this Directive and shall be responsible as a producer." Narrowing definitions limit the scope of a term with respect to its common meaning. For instance, Directive 1990/270, Art.1: "worker: any worker as defined in Article 3 (a) of Directive 89/391 who habitually uses display screen equipment as a significant part of his normal work".

Beside the explicit definition of a new concept, implicit changes in meaning occur within a diachronic process in relation to the cultural, political and social evolutions of the environment in which they are created. To be able to grasp the mutation in the social context, many of the legal concepts are *open textured*, and can only be defined extensionally through legal instances. It is mainly through the work of the judiciary that the meaning of terms, like "public interest", "public morals", can be dynamically modified and registered. From a strictly semantic point of view, we cannot expect to find any direct "referent" in reality, contrary to what happens for concepts in natural sciences, but, instead, examples of factual situations denoted by such kind of concepts or conceptual reformulations built by legal interpretation.

Interpretation is the main concern in the work of legal practitioners, that begins with the identification of the legal sources; attorneys look for the "backing" [4] on which to build their arguments in defence of the goal they wish to reach, judges evaluate the arguments in the light of the normative framework on which to found the decision, public administrators are expected to deliver services to citizens based on a listing of rights and duties extracted from the norms. Even the legislator introduces changes to the regulation of a social environment based on an conceptual reconstruction of the existing normative context; this picture enables the prediction of the innovation impact of in terms of social, economic, and cultural feed back.

All these parties carry out an operation of conceptual and systematic reformulation[3] that goes beyond the identification of the single norm (or set of norms) relevant for resolving the individual case. Only in few cases, a norm can be conceived of as the interpreted meaning of written regulations that correspond to a partition in a legal text, like articles, subsections, etc.; in legal practise, the normative context is the product of a process of reformula-

---

[3]"Legal systems have a complex structure whose internal consistency is guaranteed by meta rules of recognition [5] that regulate the dynamic evolution of the normative corpus (implicit and explicit rules of abrogation), its hierarchical organisation and relations with other legal systems.

tion based on the interpretation of a set of logically entailed linguistic expressions. Literal interpretation of selected legislative statements combined with judicial opinions, are furthermore integrated by teleological interpretation and by extra-judicial arguments, such as value judgements, socio-cultural factors, subjective intuitions. The reasoning process of legal experts can be viewed as a path requiring multiple steps: crossing recognition, reconstruction, organization, literal interpretation, conceptual modelling.

Therefore, can we say that legal knowledge coincides with access to the primary sources or, in other words, can we consider to have full knowledge of the norms by merely accessing the legislation? As cited in [6], one of the barriers to render the law accessible is that: "To a worryingly large extent, statutory law is not practically accessible today, even to the Courts whose constitutional duty it is to interpret and enforce it. There are four principal reasons. … First, the majority of legislation is secondary legislation. … Secondly, the volume of legislation has increased very greatly over the last 40 years … Thirdly, on many subjects the legislation cannot be found in a single place, but in a patchwork of primary and secondary legislation. … Fourthly, there is no comprehensive statute law database with hyperlinks which would enable an intelligent person, by using a search engine, to find out all the legislation on a particular topic".

3. Multilingualism

If we move from a monolingual (and national) dimension to a multilingual (and trans-national) dimension, a further complexity arises: legal terminologies used in both European and non-European legal systems express not only the legal concepts which operate in the different countries, but also reflect the deep differences existing between the various systems and the varying interpretations given by lawyers in each system. Given the structural domain specificity of legal language, we cannot talk about "translating the law" to ascertain correspondences between the legal terminology in various languages, since the translational correspondence of two terms satisfies neither the semantic correspondence of the concepts they denote, nor the requirements of the different legal systems.

In the European context, multilingualism affects the comprehensibility of legal documents from a dual point of view: on the one hand, as already mentioned, the difficulty of establishing meaning correspondences between concepts that reflect different legal systems (and social/cultural contexts), in order to find consistent horizontal equivalences [7]. On the other, the need to guarantee vertical consistency between the legal language of the national system and that of the European Union,

since European Directives provide measures that should be implemented in national legislation. For this purpose, any Directive contains a series of norms defining the semantic meaning of terms contained in the acts. Member States can either choose to implement these definitions literally, or they can opt for a different definition, for multiple definitions, or no definition at all, obviously respecting the preconditions set by the European law. In order to produce conceptually equivalent legislative texts and an harmonised implementation of its rules, European Union is inevitably forced to favour generic terminological choices in its texts.

The influence of European terminology is particularly evident in the English context. In the early period of European legislative production the predominance of French left a significant mark on legislative drafting, which is more evident in common law countries.[4] To give some examples, stylistic uses in European legislative language, like, for instance, the prevalence of abstract terms instead of verbs ("opportuneness"[5]) and formal expressions, like "permit", instead of "allow" and "consent" instead of "agreement" are unusual in common English. By implementing European rules within the UK legal system, the impact of the new linguistic uses produces alterations or duplications in meanings: "regulation" is a kind of secondary legislation in UK legal system; it also denotes a legal act enacted by the European legislative bodies that has a direct binding force in Member States. Further examples can be found in commercial law, like in the 2002 Sale and Supply of Goods Act to Consumers Regulations, where a terminological (and conceptual) distinction is made between "warranties", offered by the seller unless the parties have commonly agreed otherwise; and "guarantees", which are mandatory and binding on the seller towards the consumer when the transaction concern movables (UK implementation of Directive 99/44/EC).

A further phenomenon is the importation of foreign terminology from other legal orders, a process that often brings with it a modification in the original meaning, due to the process of adaptation to the normative contexts in which the new terms are placed. Some interesting cases concern the introduction of English terms as "copyright" and "deregulation" into the Italian legal system. The former substitutes the term "diritto d'autore" in legal practice. In effect, on a normative level, the concept is meant only to refer to the economic aspects of intellectual

---

[4]A possible explanation could be that in civil law countries, like in Italy the influence of French codifications had already introduced a relevant number of legal terms from French.

[5]"Further reflection on the opportuneness of non-sector specific measures such as an optional instrument in the area of European contract law", Communication by the Commission to the European Parliament and the Council on "A More Coherent Contract Law—An Action Plan" GUCE C 63-15.03.03, Par. 4.3., p. 15.

property and, therefore, to a more restrictive meaning than the original one in common law systems. "Deregulation" has entered into use in Italy by substituting the term "delegificazione", meaning "delegation of the law-making power to non legislative bodies"; it has, therefore, completely changed the original meaning ("definition of common rules of behaviour between private parties").

A similar phenomenon related to the Asian law concerns the transposition of concept expressed in Chinese characters from Japanese legal texts to Korean and Taiwan legal systems [8]; for example the term: 親権 in the official English translation is "parental authority" ("a person who exercises parental Authority holds the rights and bears the duty to care for and educate the child", Japanese Civil Code,); the transplant of the concept in the Korean Civil Code, translated in 친권, maintains the same meaning, as expressed by the English translation ("parents shall have parental authority of their minor child", Korean Civil Act, 909 (1)). Whilst in the Taiwan legislation the Chinese expression has assumed the meaning of "parental right to be exercised for the best interest of the child", probably influenced by the UN Convention on the rights of the child. All those cases are a clear demonstration of the necessity of locating linguistic translations within conceptual and social contexts.

4. Accessing and Communicating Legal Information in the Digital Society

Linguistic and conceptual complexity is, in the digital information systems, combined with technical barriers. While the amount of public sector information made available by governments for free access and re-use is continuously increasing and among it legal information has reached an unprecedented coverage, the availability of satisfactory, complete and reliable on-line information services for legal experts and non-experts has still to come.

Documents and information in both structured and non-structured form and in different formats are stored in local and often inaccessible databases; despite several initiatives for legal documents standardisation,[6] the level of interoperability is still low, not to mention the very poor level of semantic information attached to documents that prevent conceptual interconnection and sharing of information. In searching information, most of the effort is left to the user to query, collect and integrate the in-

formation in a time consuming search activity, to overcome the technical barriers that keep the available legal digital collections in a universe of separated worlds. Especially in case law, legal professionals are fully aware of the difficulties in searching for relevant precedents which they can use to argue their side in a case.[7]

The lack of a complete and cross-national legal information system has been recognized by The European Parliament, who on the 2008 adopted a non-legislative Resolution on the role of the national judge in the European judicial system[8]: "... noting that complete and up-to-date information on Community law is not available in a systematic and proper manner to many national judges, Parliament calls on the Member States to renew efforts in this area as a true European judicial area in which effective judicial cooperation can take place requires not only knowledge of European law, but also mutual general knowledge of the legal systems of the other Member States. It welcomes the Commission's intention to support the improved availability of national databases on national court rulings concerning Community law and is of the opinion that all national judges should have access to databases containing pending references for preliminary rulings from all Member States".

Several examples of the crucial difficulties in managing the multilingual panorama of European Community can be provided, demonstrating how, in several social contexts, the terminological complexity reflects the problems of finding a methodology for bridging diversities and harmonising legal rules. In the area of private law, the two most famous projects, the PECL (Principles of European Contract Law) and the PETL (Principles of European Tort Law) include in the design of a shared conceptual area also a proposal for the use of a standardized terminology.

In criminal law, to support documents exchange in transnational criminal proceedings, a codified multilingual vocabulary for Criminal records has been defined within the institutions of the European Criminal Records Information System (ECRIS)[9].

5. The Semantic Web Approach

The context that arises from the above considerations is characterized by the textual dependence of legal knowl-

---

[6]Initiatives on adoption of XML standards for the representation of legislative document structures and metadata have been brought on both at national and international level in different countries in recent years. To cite the most successful, XML.gov in the U.S., Crown XML Schema in the U.K. provide the most rich and complete datasets made available by governments in open XML. Other initiatives in European countries, like NIR (NormeInRete) standard in Italy or Metalex in the Netherlands have also lead to further development for a pan-african standard (AkomaNtoso) and to the international initiative of Metalex/CEN global interchange standard of legal sources.

[7]On the theorical difficulties in classifing legal precedents according to cognitive and conceptual parameters, see [9], p. 26: "The principal contending methods of legal classification are formal classification of legal doctrine based on logical relations among legal rules; function-based classification based on the social roles of legal rules, and reason-based classification based on common rationales underlying legal rules and decisions".
[8]Eurlex document: A6-2008-0224.
[9]Annex A of the Council Decision 2009/316/JHA of 6 April 2009 on the establishment of the European Criminal Records Information System (ECRIS) in application of Article 11 of Framework Decision 2009/315/JHA.

---

edge, by the heterogeneity of sources and by the lack of a conceptual shared model. To face such complex scenario, ICT solutions are addressed at improving technical interoperability but also at achieving a consistent semantic interoperability, by establishing a upper "semantic layer" where to express the conceptual interconnections among linguistically different entities. Without semantic, a reliable sharing of information will remain a dream. To give an example, in legal information systems one crucial task is cross lingual retrieval of legal documents. As above highlighted, cross lingual and trans-national search requires that difference between lexical meanings and legal meanings has to be made explicit also taking into account that legal meanings are relative to each legal system. Such kind of conceptual linking is missing in traditional search strategies, based on the *syntactical mapping* of string of characters in the query to string of characters in documents or web pages.

The challenge of multilinguality is just one of the goals that the Semantic Web paradigm intends to achieve. The *semantic web* approach aims at capturing and at making machine understandable the conceptual content of digital data, by means of sets of semantic information attached to each document expressed in machine processable formats.

In the following we outline some of computational challenges, limiting the description of technical details to the key aspects necessary to understand what follows, since discussion of computational aspects are outside the scope of this paper. Our main aim is to describe some models for the construction of such upper conceptual layer, outlining the different levels of expressivity and semantic complexity according to the tasks addressed.

Under the name of "semantic technologies" an integrated environment of languages, architectures and methodologies is made available by ICT in order to perform these challenge: "The use of semantic technologies such as RDF, ontologies, topic maps, etc. is not very popular yet. This is remarkable as the most important aspect of the data that is being exchanged is in fact its linguistic meaning. Without meaning, data does not become information, it is just data. Semantics gives meaning to data and that is very useful".[10]

Methodological choices for moving from syntactic to semantic level concern:

- the way semantic information are attached to documents (manually by legal experts or documentalists; or by automatic classification and indexing)
- the type of conceptual sources: open key-words automatically selected through statistical ranking; or by reference to an external system of concepts (*controlled vocabularies*, *classification systems*, *thesauri*,

semantic lexicon, *ontologies*)

- the level of granularity and formalisation of such external system of concepts: this choice mainly depends on the task the computational applications are expected to perform: from the practical need to provide efficient legal information systems, to support in natural language processing, up to more complex tasks, like legal knowledge representation and reasoning.

Of course, the more tools for semantic processing of legal data are able to capture all the aspects of the legal knowledge, the more semantic modelling is accurate: from the identification of legal sources (connecting unique identifiers[11]), to the description of the document structural elements[12] to the automatic analysis of narrative structure (detection of argumentative patterns, semantic classification of legislative statements [10]; and finally, tools for concept extraction and ontology learning [11]. These last tools provide or integrate external vocabularies, like the one described in the next section.

5.1. Legal Taxonomies, Thesauri and Ontologies

While in medical domain several terminological and conceptual resources are available and collected in a open inventory[13] in the domain of law such kind of general repository doesn't exist yet, not even local or specialized lexicons, especially not multilingual.

Terminological resources can be composed by a simple list of terms, to which definitions can be added (glossary); when organized in taxonomic trees (classifycation/headings lists) and enriched by semantic relations, they are called thesauri. The most important and used thesaurus in Europe is Eurovoc[14] the multilingual thesaurus created by the European Union for indexing all documents produced by European Institutions; localized in 27 European languages, Eurovoc is going to be adopted for indexing also national legal documents; moreover, several projects are currently going for mapping Eurovoc to the main European Thesauri [12]. Despite the wide coverage and the large diffusion, legal terminology in Eurovoc is quite poor and limited to the legal fields belonging to the competence of European Union. Thesauri are based on a shallow semantic structure that only express taxonomic hierarchy and generic conceptual relation (for instance, in Eurovoc "ownership" subsumes heterogeneous specific concepts as "private ownership" and "acquisition of property"); as they are expected to drive acceptable classification of information and reasonable translation of terms in cross lingual retrieval, the organization of trees is more oriented to fit the user expected information needs more than to express conceptual spe-

---

[10]SPOCS, Deliverable D1.1 & D1.2, Survey of Syndication Solutions & Multilingualism, 2010, p. 29.

[11]http://tools.ietf.org/html/draft-spinosa-urn-lex-01.
[12]See footnote 6.
[13]www.geneontology.org/.
[14]www.eurovoc.eu.

cialization.

When semantic relations between terms are interpreted as formal properties of conceptual entities, those computational artifacts are called ontologies [13]: they intend to express language independent system of concepts used in a certain domain. Depending on the type of description they provide, ontologies can be top-level formal or lightweight. Formal ontologies provide a language-independent and axiomatised description of concepts, while lightweight ontologies are poorly axiomatised and describe mainly the lexicalised form of concepts. In the legal domain, initial efforts lead to the top-down development of highly axiomatised formal ontologies, known as *legal core ontologies*, containing few concepts which are considered the least common denominator of legal knowledge.[15] Typical classes are: deontic positions (duties, right, permission, power), artificial person (organization, company, etc..), legal capacity, liability, delegation, etc.

5.2. Semantic Models: The WordNet Model

The development of top-down ontologies was found to be costly and slow, and alternative strategies were explored for the bottom-up development of *lightweight* legal ontologies. The resulting ontologies contain far more legal concepts, anchored in terms appearing in the texts and based on a weak abstraction model, since the elements (classes, properties, and individuals) of the ontology depend primarily on the existing contexts. In the last years the functionalities of automatic text mining has widely improved and ontology learning tools are now the most used means for ontology building.

These mainly consist in mining big legal textual corpora in order to extract representative concepts that are further organized in semantic models. The first and mostly used model for building lightweight ontologies is WordNet [14]. WordNet, initially designed as a lexical resource, not as a formal ontology, is organized around the notion of a synset. A *synset* is a set of one or more uninflected word forms (called *wordsense*) with the same part-of-speech that can be interchanged in a certain context. For example, {case, law suit} form a noun synset because they can be used to refer to the same concept. More precisely, each synset is a set of word-senses, since polysemous terms are distinct in different word-senses, e.g., {property_1, attribute, dimension} and {property_2, belongings, holding}. Each word sense belongs to exactly one synset and each *word sense* has exactly one *word* that represents it lexically. A synset is often further described by a gloss and is related to other synsets by semantic relations, of which the most important are hy-

pernymy/hyponymy (between specific and more general concepts), meronymy (between parts and wholes), and antonymy (between semantically opposite concepts).

The multilingual version of WordNet, EuroWordNet, [19] is composed by monolingual autonomous lexicons, expressing lexica of six European Languages; each synset in the monolingual WordNets is linked to the others by cross-lingual equivalence relations to the English synsets recorded by the Inter-Lingual-Index (ILI). Language-specific synsets from different languages that are linked to the same ILI-record are considered conceptually equivalent. Thus the network of equivalence relations (complete equivalence, near equivalence, or equivalence as a hyponym or hypernym) determines the interconnectivity of the indigenous WordNets.

A specialization of the Eurowordnet model for legal terminology (LOIS[16]) was built by automatically selecting relevant terminology from parallel corpora extracted from the European Directives in six EU languages and integrated by terms selected from national legislation. The emphasis was on the detection of legal conceptual terminology, *i.e.* terminology that is specific to the legal domain, as opposed to the common sense concepts of generic WordNets.

To face the peculiarities of legal language, the WordNet model has been specialized by adding domain specific relations. S*tructural relations* reflect actual systemic connections between legal concepts: the *defined as* relation (stating that a term has a certain set of necessary and sufficient conditions, turning it into a concept with a specific meaning), the *consists of* relation, anchoring the elements of the definition into a document or multiple documents, and the *implemented as* relation, providing a reference to the elements of a European definition from the definitions in national legislation. This is especially interesting for checking differences that may occur between concepts in Directives and concepts in national legislation. Further drelations have been introduced, for instance, the legal notion of *functional similarity* among legal institutions, (Camera dei Deputati, Assemblée nationale, Congreso de los Diputados).

The need of a stronger semantic characterization of lexical ontologies has been emerging in recent years, probably due to the increasing prevalence of ontology learning vs. abstract conceptualization. Several proposals for anchoring lexicons to foundational ontologies [21] have proved to be effective for general purpose resources, but too generic for specific domains like the legal one. Recent proposals focuses on the design of modular architecture where lexical and domain semantic can be expressed at distinct levels, enabling multiple links to

---

[15]Some examples include Functional Ontology of Law [15], LRI-Core [16], Core Legal Ontology [17], LKIF-Core [18].

[16]The LOIS data base, composed by about 35,000 concepts in six European languages (English, German, Portuguese, Czech, and Italian, linked by English) [20].

external domain ontologies, for instance ontological distinctions such as between "contract" as a document and "contract" as a legal transaction can help in sense disambiguation and in setting cross lingual correspondences. A terminology may refer to different domains (for instance, civil law and crime law), while a specific domain may have different terminological specialization (in nation legal systems). This issue will be further discussed in the next section.

6. A Modular Architecture

The methodologies for conceptualization exclusively based on the bottom-up extraction of linguistic knowledge make two strong assumptions: that textual corpora contain all the information necessary to understand the meaning of term and that the reverse process is possible, namely, "that the conceptual structure matches the lexical structure and the former can be deduced from the latter" [22]. On the other hand, top-down ontological models assume that a common cultural background exists, on which the main meaning commitments can be exploited and shared. We argue in favour of a multi layered model where language and concepts are integrated, but kept separated, to allow one (among the many) possible perspectives of a domain description to be expressed in an ontology, to which multiple contextualisations correspond in the lexicon.

Recent Semantic Web frameworks clearly distinguish the two layers and how they interact. This is of crucial relevance in order to enable re-use, localization and mapping among existing ontologies. The lexical layer comprises [23] terms, definitions and contexts of the domain entities that populate the ontology; terms are "concept labels" rather than concepts, so labels can be associated with more than one concept and a concept may be expressed by more than one label (including synonyms or terms in different languages). All information about the domain, for instance proper sub-class relation, are expressed in the ontology.

This distinction would enable the ontology builder to separate a legal concept from its lexical representation (by simple term or compound expression) within a linguistic system and in different linguistic systems. It enables to distinguish among synonyms (preferred label, alternative label) and to manage the fact that, even in technical language like the legal one, terms can be polysemous, and should therefore be assigned to more than one concepts. Moreover, separating terms from concepts will allow for a further distinction to be made with respect to legal terms (describing domain-specific legal notions) or lexical items, which are general language items used in legal discourse, and provide their real-world embedding which is not specific to the legal

domain.

Changes in one of the two (lexical and the conceptual) layers will produce reciprocal impacts [23], the introducetion of a new legal concept will propagate changes in its lexicalizations, not necessarily extended to all the languages and legal systems, but only in those where the new concept is introduced. For instance the introduction of a new definition at European level will produce effects in the European systems (and terminologies) only, the transplant of a concept from one legal system to another will affect the target terminology only. The addition of new term in the lexical layer, (for instance the Chinese expression of "parental authority" in the Taywan context) will impact the conceptual layer, by duplicating the concept in order to express both *parental authority* and *parental right* in the child interest.

7. Conclusions

The evolution of legal systems and the rising complexity of reality, globalisation and specialisation in each scientific field, made it clear that legal analysis cannot be closed within itself but it must be a multi-faceted investigation aimed at explaining law and society. The use of tools taken from other disciplines, like the use of cognitive science, economics and sociology can enrich legal research. Even Information and Communications Technologies can play a role, as the aim of computational applications is to provide communications among users from different countries, languages and cultures; to reach this goal, computational solutions cannot avoid to oversimplify the theoretical challenges, since addressed to reach practical goals; but, on the other side, they offer rigour and clarity in theoretic feed-back and concrete testing.

In the legal domain, the construction of ontologies encounters additional theoretical barriers. If we rely on the analytical view that legal knowledge lies entirely in the interpretation of written sources[17], the contradiction between the heterogeneity of legal systems and the need to communicate the law is re-proposed without practical solutions. Cultural dimension and social changes influence not only linguistic uses but also the process of contextual interpretation of written sources. On the extreme opposite side, the realistic[18] approach considers that only the fundamental basic concepts of law [24] ("right",

---

[17]The formalistic approach historically belongs to the civil law tradition as for instance in the Italian legal system; Civil Code, Art.12 General Provisions. Legal Interpretation—"In applying the law, no other sense but that which is made apparent by the meaning of the words according to the links among them and by the intention of the legislator can be attributed to it; if a dispute cannot be solved with a single provision, other provisions that regulate similar cases or analogous material shall be referred to; if there is still doubt, the decision shall be based on general provisions in the national legal order."
[18]American realism reflects the importance of the binding force of case precedents in common law systems.

"power", "obligation") can be considered universal, whilst the majority of legal concepts are no more than containers of meaning, or "formants" [25] that have their referents in social rather than physical reality.

Without doubt, the conceptualisation of law uses extra-linguistic knowledge that emerges from a theoretical debate, from a stratification of uses, from a collective or wide agreement about the meaning. Apart from being effective means of communication, what computer tools can offer in term of theoretical feed back is a tight framework, consistent collection of multiple contexts of use, and a clear framework, by which to represent the complex and dynamic nature of the phenomenon. They can also offer effective guidelines for legal drafters in order to produce more coherent and semantically consistent texts.

REFERENCES

[1] B. Bottalico, S. Azzini and A. Santosuosso, "Universal v. National in Bioethics. The Informal Transnational Flow of Legal Standards and a Theoretical-Practical Proposal," 2009. http://www.pulib.sk

[2] R. Tennerman, "Questioning the Univocity Ideal. The Difference between Socio-Cognitive Terminology and Traditional Terminology," *Hermes*: *Journal of Linguistics*, Vol. 18, 1997, pp. 51-91.

[3] L. Mommers, W. Voermans, W. Koelewijn and H. Kielman, "Understanding the Law: Improving Legal Knowledge Dissemination by Translating the Contents of Formal Sources of Law," *Artificial Intelligence and Law*, Vol. 17, No. 1, 2009, pp. 51-78. doi:10.1007/s10506-008-9073-5

[4] S. Toulmin, "The Uses of Argument Cambridge," Cambridge University Press, London, 1958.

[5] H. L. Hart, "The Concept of Law," Clarendon Press, Oxford, 1958.

[6] N. Holmes, "Accessible Law," 2011. http://blog.law.cornell.edu/voxpop/2011/02/15/accessible-law/

[7] G. Ajani, "Terminology and Search Functions. 25 Years of European Law Online," Publications Office of the EU, Luxembourg, 2007, pp. 129-136.

[8] L. S. A. Huey and Y. Matsuura, "Law Pack—Koala's Gift for Comparative Law in Context," *Proceedings of Law via the Internet Conference*, Hong Kong, 11 June 2011.

[9] E. Sherwin, "Legal Taxonomy," *Legal Theory*, Vol. 15, No. 2, 2009, pp. 5-54.

[10] E. De Maat, K. Krabben and R. Winkels, "Machine Learning versus Knowledge Based Classification of Legal Texts," *Proceedings of Legal Knowledge and Information Systems—JURIX* 2010: *The* 23*rd Annual Conference*, Amsterdam, 16-17 December 2010, pp. 87-96.

[11] E. Francesconi, S. Montemagni, W. Peters and D. Tiscor-nia, "Semantic Processing of Legal Texts. Where the Language of Law Meets the Law of Language," Springer, Berlin, 2010.

[12] G. Bartoloni and E. Francesconi, "Sharing Knowledge by Conceptual Mapping: The Case of EU Thesaural Interoperability," *Proceedings of the JURIX Conference*, IOS Press, Amsterdam, 2010, pp. 17-26.

[13] T. R. Gruber, "A Translation Approach to Portable Ontology Specification," *Journal of Knowledge Acquisition —Special Issue*: *Current Issues in Knowledge Modelling*, Vol. 5, No. 2, 1993, pp. 199-220.

[14] C. Fellbaum, "WordNet: An Electronic Lexical Database," MIT Press, Cambridge, 1998.

[15] A. Valente, J. Breuker and B. Brouwer, "Legal Modelling and Automated Reasoning with ON-LINE," *International Journal of Human-Computer Studies*, Vol. 51, No. 6, 1999, pp. 1079-1125. doi:10.1006/ijhc.1999.0298

[16] J. Breuker, B. Brouwer and A. Valente, "Use and Reuse of Legal Ontologies in Knowledge Engineering and Knowledge Management," In: *Law and the Semantic Web*: *Legal Ontologies*, *Methodologies*, *Legal Information Retrieval*, *and Applications* (*Lecture Notes in Computer Science*/ *Lecture Notes in Artificial Intelligence*), Springer Verlag, London, 2005, pp. 36-64.

[17] A. Gangemi, M.-T. Sagri and D. Tiscornia, "A Constructive Framework for Legal Ontologies," In: B. Selic and A. Gangemi, Eds., *Law and the Semantic Web—Legal Ontologies*, *Methodologies*, *Legal Information Retrieval*, *and Applications*, *Lecture Notes in Computer Science*, *Subseries*: *Lecture Notes in Artificial Intelligence*, Springer, Amsterdam, 2005, pp. 97-124.

[18] R. Hoekstra, J. Breuker, M. Di Bello and A. Boer, "The LKIF Core Ontology of Basic Legal Concepts," The Leibniz Center for Law, Amsterdam, 2008, pp. 43-63.

[19] P. Vossen, "EuroWordNet: A Multilingual Database with Lexical Semantic Networks," Springer, Berlin, 1998.

[20] W. Peters, D. Tiscornia and M. T. Sagri, "The Structuring of Legal Knowledge in LOIS," *Artificial Intelligence and Law*, Vol. 15, No. 2, 2007, pp. 117-135.

[21] C.-R. Huang, N. Calzolari, A. Gangemi, A. Lenci, A. Oltramari and L. Prevot, "Ontology and the Lexicon," Cambridge University Press, Cambridge, 2010. doi:10.1017/CBO9780511676536

[22] C. Roche, "Saying Is Not Modelling," *Proceedings of NLPCS* 2007 (*Natural Language Processing and Cognitive Science*), Funchal, June 2007, pp. 47-56.

[23] P. Cimiano, E. Montiel-Ponsoda, P. Buitelaar, M. Espinoza and A. Gomez-Pérez, "A Note on Ontology Localization," *Applied Ontology*, Vol. 5, No. 2, 2010, pp. 127-137.

[24] W. N. Hohfeld, "Fundamental Legal Conceptions as Applied in Judicial Reasoning," *The Yale Law Journal*, Vol. 26, No. 8, 1917, pp. 710-770. doi:10.2307/786270

[25] R. Sacco, "Prospettive della Scienza Civilistica Italiana All'Inizio del Nuovo Secolo," *Rivista di Diritto Civile*, Vol. 51, No. 4, 2005, pp. 417-441.