FrameNet model of the suspension of norms

Conference Paper · June 2011

DOI: 10.1145/2018358.2018385 · Source: DBLP

4 authors:

Monica Palmirani
University of Bologna
97 PUBLICATIONS **844** CITATIONS
SEE PROFILE

Marcello Ceci
University of Luxembourg
18 PUBLICATIONS **64** CITATIONS
SEE PROFILE

Daniele P. Radicioni
Università degli Studi di Torino
73 PUBLICATIONS **418** CITATIONS
SEE PROFILE

Alessandro Mazzei
Università degli Studi di Torino
67 PUBLICATIONS **359** CITATIONS
SEE PROFILE

FrameNet Model of the Suspension of Norms

Monica Palmirani
CIRSIFD
University of Bologna
monica.palmirani@unibo.it

Marcello Ceci
CIRSIFD
University of Bologna
m.ceci@unibo.it

Daniele Radicioni
Department of Computer Science
University of Turin
radicion@di.unito.it

Alessandro Mazzei
Department of Computer Science
University of Turin
mazzei@di.unito.it

ABSTRACT
One open problem in the AI & Law community is how to provide computers with a basic understanding of legal concepts, and their relationship with legal texts and with the legal lexicon. We propose to add a layer to connect the linguistic description of the provisions to syntactic patterns using FramNet that can be exploited thought NLP tools. A deep-parsing and shallow-semantics approach has been devised to interpret and retrieve the characterizing components of legal modificatory provisions. In this paper we single out the case of *efficacy suspension* and show how FrameNet approach can provide profit especially to isolate temporal parameters and their interpretation.

Keywords
FrameNet, legal knowledge modelling, NLP, semantic interpretation.

1. INTRODUCTION
One of the main goals in the research conducted in the last ten years on digitalization in the legal domain has been to provide techniques for detecting the linguistic legal content from the text for favouring consolidation ([15],[10],[3]), for helping the legal drafting activity ([7],[17]) or for extracting arguments for supporting logic rules and metadata ([1],[4]).

In earlier works in this area ([12],[14]) we detected some regularity in the linguistic structure of modificatory provisions, and we showed how this regularity, coupled with a prestructured XML markup, could help NLP tools correctly qualify a modificatory provision [11] using pattern matching approach. We now propose a more formalised tool devised for filling the gap between the legal lexicon and its semantic pattern, for facing more complex linguistic detection. Several studies in literature are looking to FrameNet ([2]) with interest, and the Italian FrameNet Project has been recently started ([3]). It is also started the conversion in OWL[1]. We present here a specialization of FrameNet designed to model the modification of norms suspended efficacy for the Italian environment and lexicon.

ICAIL'11 June 6-10, 2011, Pittsburgh, PA

A *suspension* may be defined as an action by which a textual provision interrupts the efficacy of a legal text (or fragment thereof) for a given period. We focus on the suspension of efficacy for some reasons: in fact, it witnesses of relevant drafting needs (more about this later on) and it is by far more complex on a linguistic perspective, thereby requiring further modelling efforts with respect to integration, substitution and repeal modifications [12].

The challenge here is to detect and interpret such actions (when a suspension starts and ends, when it is interrupted or extended, etc.), since the meaning of a modification depends on how the language is used, for example, on whether verbs are used in the passive or in the active form. Suspension is often confused with other modifications (such as derogation and disapplication[2]), and textual substitutions sometimes modify a suspension arguments (such as duration). On the other hand, suspending a norm's efficacy makes norms significantly more dynamic over time, thus also affecting the judge's application of such norms.

We thus take into account the suspension of efficacy (§ 2) definition and anatomy; we identify, on the base of a number of relevant documents analysed, different classes of suspension (§ 3). And these classes we then modelled using FrameNet (§ 4). Finally we added such novel layer to the NLP system architecture to be able to deal with further kinds of modificatory provisions (§ 5).

2. SUSPENSION OF EFFICACY ANALYSIS
In light of the foregoing considerations, we aim to model the suspension of efficacy with FrameNet so to facilitate automatic detection of arguments in the text. We chose this kind of modification (suspension) because it is more complex and rich with arguments than other temporal modifications. Furthermore suspension is a relevant modification, in that it is often used as a legislative drafting technique for introducing a temporary law.[3] This need stems from for two main reasons: when the topic is so complex but urgent that it is necessary to have a temporary solution (e.g. Genetic Law); when some time is needed to fully apply the new dispositions (e.g. Euro Law in 1999). *Suspension* may be defined as the action by which a textual provision interrupts the efficacy of a legal text (or fragment thereof) for a given period [8]. Suspension is based on the rationale that some norms so strongly affect their addressees (citizens, businesses,

[1] http://www.loa-cnr.it/codeps/owl/ofntb.owl

[2] *Disapplication*: the secondary law suspends the primary law.

[3] See the Arizona Legislative Bill Drafting Manual 2009, p. 6; Texas Legislative Council Drafting Manual, p. 58, Main Legislative Drafting Manual, p. 18; Alaska Manual of Legislative Drafting, p. 23.

social actors) that an adequate period is needed for them to tune into the process. Our goal is to track this rationale over time even if suspension may come by a variety of different modifications. In other words, we aim at capturing not only suspensions arguments but also its process and evolution over the time, to put the suspension into relation with its underlying normative rationale.

The suspension can be either *explicit* or *implicit*, depending on the language of the provision in question. And, temporally, it can be either *defined* or *undefined*. A suspension is defined when the period during which a norm efficacy is interrupted is explicitly stated in the text. Provision clearly indicates a beginning and an end (or an initial and a final event). By contrast, a suspension is undefined when the interval during which a norm's efficacy is interrupted is not explicitly set out in any part of the suspending provision. This class of suspensions includes at least three subclasses as follows: (i) *sine die* suspension (without an ending date); (ii) suspension conditioned by an external event (e.g., "Article 5 is suspended for a six-month period starting from entry into force of the Treaty"); and (iii) suspension intervals described with a set of other parameters such as the duration (e.g., "Article 5 is suspended for four months starting from 31 December 2010). The text often needs to be interpreted to detect the correct value.

Another important case we consider is that where a suspension provision is modified. A suspension is usually reflexive, with the law introducing the suspension being the same as that affected by it (this is a role usually devoted in Italy to a law's closing articles). However, it is not unusual to see a later provision modifying the suspension for the same reasons that led to its introduction.

3. LANGUAGE REGULARITY OF SUSPENSION PROVISION

In order to model, and consequently extract, semantics from laws introducing or modifying a suspension of efficacy, we have surveyed a large body of norms that High Court of Cassation legal practitioners have semantically annotated in semi-automatic way using Norma-Editor tool [16], in the last 5 years. The data collection includes about 29,000 documents[4] (divided in 46,483 sections/articles) dating from 2005 to 2009, all of them coming from the Italy's *Official National Gazette* and converted into NiR[5] XML schema definition DTDv2.0 format using Norma-Editor structural parsers. On this body of documents we did a linguistic analysis to extract patterns for each type of suspension provision. The total modifier sections/articles are 19,203 out of 46,483, representing 41.31% of the total sections/articles. The suspending documents (act) are in total 90 (0.31%), for a total of 104 modifier articles (0.54%). The language of suspension exhibits a certain regularity making it possible to fill the gap between the legal lexicon and the rules of suspension. The logical structure of the suspension norm is, following the annotation of FrameNet:

[*PassiveNorm* Article 14] [*Copula* is] **suspended** [*TermporalArguments* for 6 months]

Ten lexical units (see § 4) directly point to a suspension of efficacy in our document sample. There are: three verbs: suspend, disapply, and apply; five nouns: validity, efficacy, application, effect and force and two adjectives: valid and efficacious.

Most of them create no ambiguity in identifying a suspension of efficacy: only the terms force, validity, and valid need an explicit disambiguation, as they usually refer to a norm's coming into force and to its validity, but they are seldom improperly used by the Italian legislator, who actually uses them to mean "efficacy."

Two frame elements (see § 4) accompany a suspension-*evoking* element. They are as follows.

(1) PASSIVENORM. This element represents a norm whose efficacy will be suspended for a certain period. Such norm typically occurs as the subject of a sentence or clause containing a relevant term, except when the subject of the clause is *efficacy*.

(2) TEMPORALARGUMENTS. These time expressions (or time markers) define the time at which a modification of efficacy is to take effect: they do so by specifying a beginning or an end or a period. The lexical forms are quite regular: beginnings are typically signalled by use of function words and phrases such as *from* and *starting from*, *as from*; endings are typically signalled by *until*; and periods by *for*, *during*, *in the year*, and suchlike. No definite connections could be found between specific suspension-related terms and time expressions, even though verbs expressing "process events" can give us some clue about how to interpret or classify a TemporalArgument; thus, *cease* can only introduce a suspension starting time, and *concern* only a period. TemporalArguments can be implicitly derived from the time of a suspension coming into force, or they can be explicitly bound to the entry into force of a law other than the PassiveNorm: this other norm will occur as the object of a preposition following the term *force*, which forms part of a time expression ("PassiveNorm is suspended starting from the entry into force of law Y").

4. MODELLING SUSPENSIONS USING FRAMENET

FrameNet —a lexicon-building project developed at Berkeley University—that produces a set of *frames* able to describe the concepts expressed through a specific language, filling in such matter the gap between ontology and lexicon. The frames are composed by *frame elements (FEs)*: the main parts of the concept. The words (nouns, verbs, adjectives, etc.) that are regularly used for evoking the *frame elements* in a particular context are called *lexical units* (*LU*). The roles of the words inside of the sentence under the grammatical point of view (*GF*) (subject, object, verbs, etc.) defines a second layer of qualification of the frame. The structural type of the phrase defines phrase type (*PT*) (noun phrase, propositional verb, etc.). So the frame elements, by their lexical units, describe the conceptual structures of sentences. This makes it possible to map the main parts of speech (verbs, nouns, etc.) and to couple them with the frames evoked by the words. Our goal is to model the frames for the Suspension using all the FrameNet instruments: FEs and LU, GF and PT. Secondary the frames are modelled in two levels and related each others with a relationship.

[4] The database is owned by the Italian High Court of Cassation and it is not released with Open Source licence. CIRSFID cooperated to the XML mark-up project, providing technical and legal assistance.

[5] NiR – Norme In Rete Italian XML standard for legislative documents.

Two frames are created on the first layer, namely, the ***Efficacy_Inclusion*** and ***Efficacy_Exclusion*** frames, which on the second layer will be merged into the ***Main_Suspension*** frame expressing its meaning in terms of lack of efficacy.

FRAMENAME (Efficacy_Inclusion) FES ([*PassiveNorm*], [*Period_Start*], [*Period_End*]) [*PassiveNorm*] has **efficacy** [*Period_Start*] to [*Period_End*]
FRAMENAME (Efficacy_Exclusion) FES ([*PassiveNorm*], [*Period_Start*], [*Period_End*]) [*PassiveNorm*] [*Copula* is] **suspended from** [*Period_Start*] to [*Period_End*] [*PassiveNorm*] has **not efficacy from** [*Period_Start*] to [*Period_End*] [*PassiveNorm*] has **efficacy until** [*Period_End*]
FRAMENAME (Main_Suspension) FES ([*PassiveNorm*], [*Suspension_Start*], [*Suspension _End*]) [*PassiveNorm*] [*Copula* is] **suspended from** [*Suspension_Start*] to [*Suspension_End*] [*PassiveNorm*] has **efficacy from** [*Suspension_Start*] to [*Suspension_End*]

On the first layer, the relevant LUs evoke either the *Efficacy_Inclusion* or the *Efficacy_Exclusion* frame.

Efficacy_Inclusion LUs (*effectiveness.n, efficacy.adj, apply.v, valid.adj, validity.n, effect.n, application.n, force.n*)
Efficacy_Exclusion LUs (*suspend.v, disapply.v, cease.v+efficacy.n, not.adv+*Efficacy_Inclusion)

On the second layer, the ***Main_Suspension*** frame will be modelled by inheriting the *Process* frame. Suspension is therefore treated as a process, with a "target" represented by the *Passive_Norm* (carried over unchanged from the first layer) and whose state is affected by one or more events: it starts with the event *Suspension_Start* event and/or ends with the *Suspension_End* event.

Finally, the start of the process can be advanced or postponed by another norm, and the same can happen to its end. These events will be represented in four specific frames, subclasses of the ***Suspension_Modification*** frame whose modelling will be presented in a future work for space problem.

4.1 Clearance of the Suspension Model

A problem in the detection of the suspending provisions is to distinguish the suspension action from the exception. The exception is a modification of the norm where the rules are restricted respect the original scope. The matter limitation can involve three main possible aspects:

- the agents addressed (e.g. touristic services);
- the geospatial parameters (e.g. Abruzzo region);
- the temporal time when the fact should be considered (e.g. the earthquake of the 2009).

The main problem, also considering the intellectual activity of the law interpretation, is to decide if the modification affects only the scope or definitely the temporal effectiveness of the norms. We find the same linguistic elements (verbs, arguments) of the suspension, but the result is completely different, more oriented to limit the range of the law application rather than delimitate the temporal suspension of the norms. In the matter of fact we find structural and linguistic common elements that could create misleading in the human, as well as in the machine, approach. Both modifications often include the description of an event (e.g. date). The linguistic expression "*non si applicano*" (it will not be applied) occurs in the suspension as well as in the exception.

The exception, as well as suspension, could also be integrated by conditionals that make more complex the text analysis (e.g. the article 2 shall not be applied in case of accession of the European Union). Finally the exception often delegates to a normative citation the specification of the derogation matter (e.g. the article 4 shall not be applied in the countries defined in the article 45). In this case the derogation matter is really complex to identify without access the referred document. For the abovementioned reasons it is difficult, especially using regular patterns only, to clearly discriminate the two change models, especially in the borderlines examples.

5. NLP SYSTEM: SEMANTIC INTERPRETATION OF SUSPENSIONS

The annotation of modificatory provisions –including suspensions– is a three steps process. Although these steps have been illustrated in previous work (full details are provided in [15]), we briefly recall them in order to make the paper more complete and readable. We then show how the FrameNet formalization is used in the semantic interpretation process, pointing out the benefits due to encoding the knowledge about modifications in declarative form.

5.1 System Architecture

In the first step we look for the possible location of a modificatory provision within the document, and we simplify the input sentences, so to prune text fragments that do not convey relevant pieces of information. In the second step we perform the syntactic analysis of the retrieved sentences; in the third step we semantically annotate the retrieved provisions through a tree matching approach. We briefly illustrate the first two steps and then focus on the annotation phase and on the semantic interpreter design.

The input to the system is encoded in the NiR XML standard format for Italian Legislative Text. The NiR format encodes the structural elements used to mark up the main partitions of legal texts, as well as its atomic parts (such as articles, paragraphs, subparagraphs, and lettered and numbered items) and any non-structured text fragment. Additionally, the NiR standard includes in its DTD a part describing modifications, to implement this model in XML. Based on the XML structure, we retain the text excerpts contained between some meaningful tags (e.g.,<corpo>, which is the Italian word for *body*, where the modifications may be found). The text tagged by <rif> (Italian abbreviation for reference) and <virgolette> (Italian word for *quotes*) is then rewritten with the IDs of the corresponding tags. For example, given the XML encoding of a sentence such as "L'efficacia del decreto ministeriale 17 novembre 2006 è sospesa fino alla data del 30 aprile 2007." (The efficacy of the Ministerial Decree is suspended until the date of April 30th, 2007), we rewrite the sentence like "L'efficacia del RIF12 è sospesa fino alla data del 30 aprile 2007". This sentence, which is much simpler to analyze with no loss of information, is then given in input to the parser.

The parser adopted is a broad coverage rule based parser for Italian. It relies on a morphological dictionary of Italian (about 25,000 lemmata) and on a rule-based grammar that describes dependency structures. The output of the parser is a dependency tree that makes explicit the structural syntactic relationships occurring between the words of the sentence. Each word in the sentence is associated with a node of the tree, as depicted in Figure 1. The nodes are linked via labeled arcs that specify the role of the dependents with respect to their governor (the parent). In the considered example, "efficacia" (efficacy) is the subject of the verb "(è) sospesa" ((is) suspended), while "è" (is) is the auxiliar, marked with *aux*. A special node "trace" is framed by a dashed line and labeled *t*: it specifies that the deep subject of the suspension (the agent, in terms of roles) is not expressed. Finally, the temporal argument is in a dependent that is labeled as a *modifier*.

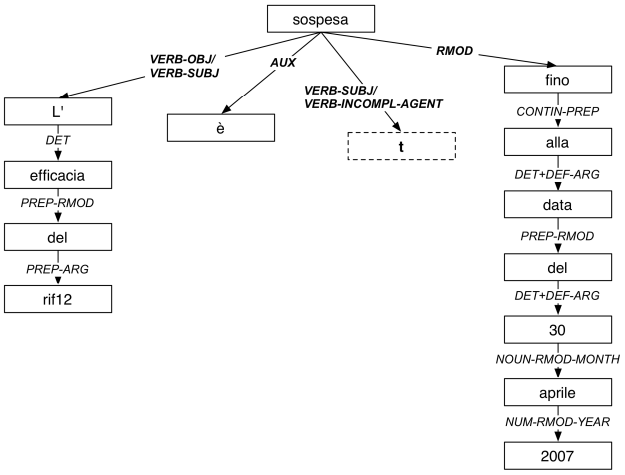

Figure 1 – The (simplified) dependency tree structure for sentence "The efficacy of the rif12 is suspended until the date of April 30th, 2007".

5.2 The Interpretation Process

Modifications are represented by means of semantic frames, composed by slots [6]. Retrieving a modificatory provision amounts to choosing the frame describing that modification, and to filling its slots with the correct arguments. The task of the semantic interpreter is twofold. First it consists in inspecting the dependents of the verb on the one hand, and in inspecting the frames and the available syntactic and semantic information on the other hand. Then the semantic interpreter is charged to find the frame that best fits to current setting. Secondly, once the appropriate frame has been individuated, the related set of rules is applied to retrieve the fillers for the frame slots. The information stored in the FrameNet formalization is thereby fundamental, since it provides a necessary interface between the syntactic and the semantic levels. Additionally, it allows formalizing syntactic and semantic knowledge about modificatory provisions in a *declarative* (as opposed to *procedural*) manner. That is, the FrameNet formalization allows illustrating the rationale underlying and governing the application of rules, since it puts together both the information about the modification, and their grammatical and syntactical possible realizations.

The semantic interpreter is charged to test whether the root node of the syntactic tree is a verb, and if it belongs to the modificatory provisions taxonomy ([13]). For example, given the root verb, we take the verb lemma sospendere (suspend), we search for it in the knowledge base, and find that it is a possible instantiation of the legalCategory *suspension*, together with the verbs disapplicare (to cease to apply), applicarsi (enforce), etc.. In this case we have a fundamental cue that the sentence being analyzed contains a modificatory provision. We earlier mentioned that both *efficacy inclusion* and *efficacy exclusion* frames involve the following slots: *PassiveNorm*, [*Position*], *Start* and *End*. The FrameNet encoding illustrates how such information is linguistically realized, so that the semantic interpreter can exploit it to access the rules base.

Once discovered that the modification is a suspension, the appropriate set of rules is executed so to exploit the information grasped through the FrameNet formalization to retrieve the correct slot fillers from the parse tree. Filling a modification frame corresponds to finding an appropriate mapping between (tree) dependents and (frame) slots. To carry on with the sentence under consideration, we have seen (§ 4) that a typical (syntactic) construction for the Efficacy_Exclusion frame is:

PassiveNorm is suspended from Start to End. (1)

which could be rewritten into:

P is suspended from S to E. (2)

In practical cases it may happen that either the Start or the End argument is lacking, therefore determining an open time span, where one of the two temporal arguments may be absent. Among many possible variants of the sentence in (2), a slightly different linguistic construction (§ 4) can be

The efficacy of P is suspended from S to E. (3)

Once the semantic interpreter recognizes the appropriate frame, further relevant information can be made available and exploited, that is related to the syntactic structure:

[The efficacy of P]$_{subj}$ is suspended [from S]$_{rmod}$ [to E]$_{rmod}$.(4)

Furthermore, the FrameNet formalization provides a compact description of (some of) the possible syntactic realizations of the modificatory provisions. That is, the locution "The efficacy of P" is expected to occur in a branch of the parse tree rooted under the main (suspend: governor) verb. Namely, in the branch containing the subject of the sentence, and accordingly labeled *verb-subj*. The processing of such tree branch will allow extracting the reference to the passive norm. Similarly, extracting both the start time and the end time will imply traversing the tree branches labeled *rmod*. As suggested in the description of the frame, the presence of words/locutions such as "a partire da" (starting from), "a decorrere da" (starting day will be) or "fino a", "sino a" (until) will provide precious cues about the starting and ending times of the suspension time span. Based on such information, the set of rules related to each modification are executed to test the content of the verb arguments and the verb modifiers to fill the slots of current frame. The rules are charged to discover whether in the syntactic arguments like subject, object or in any modifier are present any meaningful locutions or constants, such as RIF. In this way we can conveniently map the syntactic pattern described in the FrameNet formalization onto the set of semantic slots.

The whole interpretation process is illustrated in Figure 2.

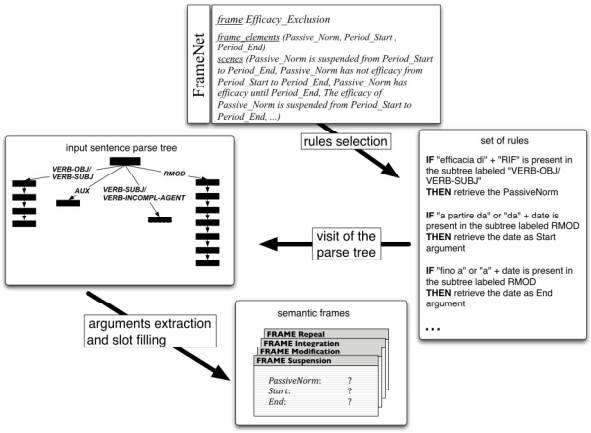

Figure 2 – The semantic interpretation process.

6. CONCLUSIONS

This paper, that is a continuation of a previous research work [12], describes a robust legal and linguistic methodology for approaching legal texts analysis, with special emphasis to temporal modifications. The adoption of the FrameNet approach allows to integrate into the NLP [9] implemented system a wealth of information about legal language phenomena, that span over different layers, such as the legal one, the grammatical one and the syntactic one. The adoption of a FrameNet-based approach yields as benefit that we obtain a declarative description of modifications. In turn, such decoupling between declarative knowledge and procedural parts of the system helps separating legal knowledge from its use, which is not only more convenient on a software engineering perspective, but it is also helpful in extending the system coverage. The results of the first experiments of the system seem to corroborate the approach undertaken; however an extensive experimentation will be necessary to fully access the goodness of the approach. Future works will involve investigating the related modification of exceptions in its connections to suspensions, in order to yield a broader coverage of the modifications handled and a deeper comprehension of legal and linguistic phenomena.

7. ACKNOWLEDGMENTS

The authors thank the ICT4Law project (http://www.ict4law.org/), coordinated by Ajani G., financed by Piedmont Region, for supporting this research.

8. REFERENCES

[1] Biagioli C., Francesconi E., Passerini A., Montemagni S., and Soria C., "Automatic Semantics Extraction in Law Documents." In Proceedings of International Conference on Artificial Intelligence and Law, 2005, pp. 133–39, 2005.

[2] Collin F. B. , Fillmore C. J., and Lowe J. B., "The Berkeley FrameNet Project.", Proceedings of the 17th International Conference on Computational Linguistics, August 10–14, 1998, Montreal and Quebec, Canada, 1998.

[3] De Maat E., van de Ven S., Winkels R., van Engers Tom M., "Automated Handling of Amending Documents and Resulting Consolidations.", JURIX 2009, pp. 116-125, 2009.

[4] De Maat E., Winkels R., "Automated Classification of Norms.", in Sources of Law. Semantic Processing of Legal Texts 2010, pp. 170-191, 2010.

[5] Dini L., and Bosca A., "Dependency Based Valence Induction for an Italian FrameNet.", In Proceedings of GL2009, pp. 27–35, 2009.

[6] Fillmore C. J., "Scenes-and-frames semantics.", In Antonio Zampolli, editor, Linguistic Structures Processing, pp. 55–79. North Holland, Amsterdam, 1977.

[7] Francesconi E., and Passerini A., "Automatic Classification of Provisions in Legislative Texts. In ICAIL 2007, vol. 15, pp. 1–17, 2007.

[8] Guastini R., *Teoria e dogmatica delle fonti*. Milan: Giuffrè, 1998.

[9] Lesmo L., "The Rule-Based Parser of the NLP Group of the University of Torino.", Intelligenza Artificiale, 2(4), pp. 46–47, June 2007, 2007.

[10] Lesmo L., Mazzei A., and Radicioni D. P., "Extracting Semantic Annotations from Legal Texts." Hypertext 2009, pp. 167–72, 2009.

[11] Lesmo L., Mazzei A., and Radicioni D. P., "Semantic Annotation of Legal Modificatory Provisions.", In AI*IA 2009: 10th Congress of the Italian Association for Artificial Intelligence, ed. R. Serra, LNAI, Berlin: Springer-Verlag, 2009.

[12] Mazzei A., Radicioni D. P., and Brighi R., "NLP-Based Extraction of Modificatory Provisions Semantics." ICAIL 2009, pp. 50–57, New York: ACM Press, 2009.

[13] Palmirani M., "Time Model in Normative Information Systems." In proceedings of the workshop The Role of Legal Knowledge in eGovernment, ICAIL2005, 2005.

[14] Palmirani M., and Brighi R., "Model Regularity of Legal Language in Active Modifications.", AICOL Workshop 2009, pp. 54–73 Springer, 2010.

[15] Palmirani M., Brighi R., and Massini M., "Automated Extraction of Normative References in Legal Texts.", In Proceedings of ICAIL 2003, pp. 105–6, New York: ACM Press, 2003.

[16] Palmirani. M., and Brighi R., "An XML Editor for Legal Information Management." In proceedings of DEXA 2003, Workshop on E-Government, Prague, September 1–5, pp. 421–29, Berlin and Heidelberg: Springer-Verlag, 2003.

[17] Soria C., Bartolini R., Lenci A., Montemagni S., and Pirrelli V., "Automatic Extraction of Semantics in Law Documents.", In Proceedings of the 5th Legislative XML Workshop, pp. 253–66, ed. C. Biagioli, E. Francesconi, and G. Sartor. European Press Academic Publishing, 2007.