An Approach to Legal Rules Modelling and Automatic Learning

Enrico FRANCESCONI
ITTIG-CNR, via de' Barucci 20, Florence (Italy)

Abstract. In this paper an approach to legal rules modelling based on a semantic model for legislation and oriented to knowledge reusability and sharing is presented. An automatic methodology able to support rules learning is proposed as well: it is based on techniques of knowledge extraction from legislative texts. This combined approach is aimed at giving a contribution to bridge the gap between consensus and authoritativeness in legal knowledge representation.

Keywords. Legal Knowledge Modelling, Knowledge Extraction, Machine Learning, NLP techniques.

1. Introduction

Knowledge modelling represents a structural pre-condition for implementing the Semantic Web concept as well as intelligent systems dealing with legal information [1]. One of the main problems in this field, addressed in literature, is the existing trade-off between *consensus* and *authoritativeness* in legal knowledge representation.

Consensus is an issue faced in knowledge representation in general [2], since ontological conceptualization has to be shared between stakeholders [3]. Several approaches have been undertaken to reach consensus in legal knowledge representation: for example the *common-sense terms* approach [4] based on common sense understanding of the terminology identifying concepts, as well as the *folksonomy* approach[1] based on social and collaborative activities of concepts selection and categorization [5]. Knowledge representation in the legal domain, however, shows peculiarities due to the importance of having authoritative systems based on legal rules for legal assessment and reasoning [6], or advanced search engines able to retrieve not just documents but also the contained norms [7]. Both common-sense terms and folksonomy approaches are well suited to reach *consensus* on domain concepts, however, when applied to the description of *legal rules*, the gap between consensus and authoritativeness is usually emphasized. For example, by the common-sense terms approach, social and communicative words typical of the legal domain can be provided [8]: in this approach experts may provide de-

[1] Folksonomies (or social tagging mechanisms) have been widely implemented in knowledge sharing environments; the idea was first adopted by the social bookmarking site del.icio.us (2004) http://delicious.com

scription of rules on entities as well as translating them into technical terminology [4], but this activity might reduce *consensus*. Similarly, in the folksonomy approach stakeholders may provide description of rules regulating entities, which might reduce *authoritativeness*.

Nowadays a very active research area is represented by knowledge acquisition from texts [9], since electronic texts still represent the most widely used communication medium on the Web. This approach can play an important role in legal knowledge acquisition, since written text is the most widely used way of communicating legal matters [10]. Knowledge acquisition techniques can be used for implementing taxonomies or suggesting concepts for upper level ontologies, mainly hand-crafted by domain experts, as well as for identifying *legal rules* [10,35].

In this paper an approach to support the acquisition of legal rules contained in legislative documents is presented: it is based on a semantic model for legislation and implemented by using knowledge extraction techniques over legislative texts. This methodology is targeted to provide a contribution to bridge the gap between consensus and authoritativeness in legal rules representation, because it contributes to reach consensus by limiting human intervention in legal rules description, which are extracted from authoritative texts as the legislative ones.

This paper is organised as follows: in Section 2 an approach to legal rules modelling and acquisition is presented, in Section 3 a semantic model for legislative texts is introduced, in Section 4 a knowledge acquisition methodology is shown and tested, finally in Section 5 some conclusions, discussing the benefits of the described learning approach, are reported.

2. An approach to legal rules modelling and acquisition

The proposed approach to legal knowledge acquisition is based on learning techniques targeted to extract *legal rules* from text corpora. Legal rules are essentially "speech acts" [11] expressed in legislative texts regulating *entities* of a domain: their nature therefore justifies an approach aimed at the analysis of such texts.

Therefore, the proposed knowledge acquisition framework is based on a twofold approach:

1. Knowledge modelling: definition of a semantic model for legislative texts able to describe legal rules;
2. Knowledge acquisition: instantiation of legal rules through the analysis of legislative texts, being driven by the defined semantic model.

This approach traces a framework which combines top-down and bottom-up strategies: a top-down strategy provides a model for legal rules, while a bottom-up strategy identifies rules instances from legal texts. The bottom-up knowledge acquisition strategy in particular can be carried out manually or automatically. The manual bottom-up strategy consists, basically, in an analytic effort in which all the possible semantic distinctions among the textual components of a legislative text are identified. On the other hand the automatic (or semi-automatic) bottom-up strategy consists in carrying out the previous activities being supported by tools able to classify Rules, according to the defined model, and to identify the involved Entities. In this paper the automatic bottom-up strategy is presented.

3. Knowledge modelling

The proposed approach is based on knowledge modelling oriented to interoperability and reusability, and it is based on the separation betweeen types of knowledge to be represented by Semantic Web standards. The need of identifying and separating different types of knowledge has been widely addressed in literature [12]. For example [13] criticised a common tendency to indiscriminately mix domain knowledge and knowledge on the process for which it is used, speaking of *epistemological promiscuity*. Similarly [14] and [15] pointed out that usually knowledge representation is affected by the nature of the problem and by the applied inference strategy; this key-point is also referred by [14] as *interaction problem*: it is related to a discussion regarding whether knowledge about the domain and knowledge about reasoning on the domain should be represented independently. In this respect [16] pointed out that the separation of both types of knowledge is a desirable feature, since it paves the way to knowledge sharing and reuse.

The knowledge model proposed in this work reflects these orientations and it is organized into the following two components:

1. Domain Independent Legal Knowledge (DILK)
2. Domain Knowledge (DK)

DILK is a semantic model of Rules expressed in legislative texts, while DK is any terminological or conceptual knowledge base (thesaurus, ontology, semantic network) able to provide information and relationships among the Entities of a regulated domain. The combination of DILK with one or more DKs is able to provide a formal characterization of Rules instances. For this reason we call the proposed methodology to legal knowledge modelling the *DILK-DK* approach.

3.1. DILK

DILK is conceived as a model for legal Rules, independently from the domain they apply to. In literature several models (classification) of legal rules have been proposed, from the traditional Hohfeldian theory of legal concepts [17] until more recent legal philosophy theories due to Rawls [18], Hart [19], Ross [20], Bentham [21], Kelsen [22].

In this respect a particular attention is worth to be given to the work of Biagioli [23]. Combining the work of legal philosophers on rules classification with the Searlian theory of rules preceived as "speech acts", as well as the Raz's lesson [24] to perceive laws and regulations as a set of *provisions* carried by speech acts, Biagioli underlined two views or *profiles* according to which a legislative text can be perceived: a) a structural or *formal profile*, representing the traditional legislator habit of organizing legal texts in chapters, articles, paragraphs, etc.; b) a semantic or *functional profile*, considering legislative texts as composed by *provisions*, namely fragments of regulation [23] expressed by speech acts. Therefore a specific classification of legislative provisions was carried out by analysing legislative texts from a semantic point of view, and grouping provisions into two main families: *Rules* (introducing and defining entities or expressing deontic concepts) and *Rules on Rules* (different kinds of amendments). Rules are provisions which

aim at regulating the reality considered by the including act. Adopting a typical law theory distinction, well expressed by Rawls, they consist in:

- *constitutive rules*: they introduce or assign a juridical profiles to entities of a regulated reality;
- *regulative rules*: they discipline actions ("rules on actions") or the substantial and procedural defaults ("remedies").

On the other hand, Rules on Rules can be distinguished into:

- *content amendments*: they modify literally the content of a norm, or their meaning without literal changes;
- *temporal amendments*: they modify the times of a norm (come-into-force and efficacy time);
- *extension amendments*: they extend or reduce the cases on which the norm operates.

In Biagioli's model each provision type has specific arguments describing the roles of the entities which a provision type applies to (for example the *Bearer* is argument of a *Duty* provision). *Provision types* and related *Arguments* represent a semantic model for legislative texts [23]. They can be considered as a sort of metadata scheme able to analytically describe fragments of legislative texts. For example, the following fragment of the Italian privacy law:

> *"A controller intending to process personal data falling within the scope of application of this act shall have to notify the "Garante" thereof, . . . "*

besides being considered as a part of the physical structure of a legislative text (a *paragraph*), can also be viewed as a component of the logical structure of it (a *provision*) and qualified as a *provision* of type *Duty*, whose arguments are:

Bearer:	"Controller";	*Object*:	"Process personal data"
Action:	"Notification"	*Counterpart*:	"Garante"

The specific textual anchorage of the Biagioli's model represents, in our point of view, its main strength. Since the DILK-DK approach aims at representing Rules instances as expressed in legislative texts, we consider the Biagioli's model, limited to the group of Rules, as a possible implementation of DILK. "Rules on rules" affect indirectly the way how the reality is regulated, since they amend Rules in different respects (literally, temporarily, extensionally): therefore such provision types are not part of DILK model. On the other hand their effects on Rules has to be taken into account for knowledge acquisition purposes.

3.2. DK

In legislative texts *Entities* regulated by provisions are expressed by lexical units, however no additional information on such entities are provided. This information can be provided by a *Domain Knowledge* (DK) providing conceptualization of entities expressed by language-dependent lexical units[2]. Information on such

[2] "Typically regulations are not given in an empty environment; instead they make use of terminology and concepts which are relevant to the organisation and/or the aspect they seek

entities at language-independent level, as well as their lexical manifestations in different languages have to be described by a DK. A possible architecture for describing a DK has been proposed within the DALOS project[3]; it is organized in two layers of abstraction:

- *Ontological layer*: conceptual modelling at language-independent level;
- *Lexical layer*: language-dependent lexical manifestations of the concepts at the Ontological layer.

More details on the DALOS DK architecture, as well as a possible implementation of it for the domain of consumer protection, can be found in [26].

4. Knowledge acquisition

Knowledge acquisition within the DILK-DK framework consists of two main steps: 1) DILK instantiation, 2) DK construction.

4.1. DILK instantiation

The DILK instantiation phase is a bottom-up strategy for legislative text paragraphs classification into *provision types*, as well as specific lexical units identification, assigning them roles in terms of *provision arguments*. The automatic bottom-up strategy, here proposed, consists in using tools able to support the human activity of classifying provisions, as well as to extract their arguments. Three main steps can be foreseen:

- Collection of legislative texts and conversion into an XML format [27]
- Automatic classification of legislative text paragraphs into provisions [28]
- Automatic argument extraction [29]

Legislative documents are firstly collected and transformed into a jurisdiction-dependent XML standard (NormeInRete in Italy, Metalex in the Netherlands, etc.). For the Italian legislation a module called xmLegesMarker, of the xmLeges[4] software family, has been developed [27]: it is able to transform legacy contents into XML so to identify the formal structure of a legislative document.

4.1.1. Automatic classification of provisions

For the automatic classification of legislative text paragraphs into provison types, a tool called xmLegesClassifier of the xmLeges family has been developed. xmLegesClassifier has been implemented using a Multiclass Support Vector Machine (MSVM) approach, as the one reporting the best results in preliminary experiments with respect to other machine learning approaches [28]. With respect to

to regulate. Thus, to be able to capture the meaning of regulations, one needs to encode not only the regulations themselves, but also the underlying ontological knowledge. This knowledge usually includes the terminology used, its basic structure, and integrity constraints that need to be satisfied." [25]

[3] http://www.dalosproject.eu

[4] http://www.xmleges.org

[28], in this work MSVM is tested on the Rules provision family, as first step of DILK instantiation. Documents are represented by vectors of weighted terms and some preprocessing operations are performed on pure words to increase their statistical qualities:

- Stemming on words in order to reduce them to their morphological root
- Stopwords elimination
- Digits and non alphanumeric characters represented by a unique character

Moreover feature selection techniques are applied to reduce the number of terms to be considered, thus actually restricting the vocabulary to be employed (see e.g. [30]). We tried two simple methods:

- An unsupervised *min frequency* threshold over the number of term occurences in the training set, so to eliminate terms with poor statistics.
- A supervised threshold over the Information Gain [31] of terms, which measures how much a term discriminates between documents belonging to different classes. The Information Gain of term w is computed as:

$$ig(w) = H(D) - \frac{|D_w|}{|D|}H(D_w) - \frac{|D_{\bar{w}}|}{|D|}H(D_{\bar{w}})$$

where H is a function computing the entropy of a labelled set ($H(D) = \sum_{i=1}^{|C|} -p_i \; log_2(p_i)$, being p_i the portion of D belonging to provision type i), D_w is the set of training documents containing the term w, and $D_{\bar{w}}$ is the set of training documents not containing w. This method basically allows to select terms with the highest discriminatory power among a set of provision types.

Once basic terms have been defined, a vocabulary of terms $\mathcal{T}$ can be created from the set of training documents $\mathcal{D}$, containing all the terms which occur at least once in the set. A single document d is represented as a vector of weights $w_1, \ldots, w_{|\mathcal{T}|}$, where the weight w_i represents the amount of information which the i^{th} term of the vocabulary carries out with respect to the semantics of d. We tried different types of weights, with increasing degree of complexity:

- a *binary* weight $\delta(w, d)$: presence/absence of the term within a document;
- a *term-frequency* weight $tf(w, d)$: number of times a term occurs within the document (measure of its representativeness of a document content);
- a combination of *information gain* and *term-frequency* ($ig(w, d) * tf(w, d)$);
- a *tf-idf* [32] weight: term specificity degree with respect to a document.

A wide range of experiments was conducted over a dataset made of 258 Rules instances, collected by legal experts, distributed among 6 provision classes (Tab. 1). After terms preprocessing, we tried a number of combinations of the document representation and feature selection strategies previously described. We employed a *leave-one-out* (loo) procedure for measuring performances of the different strategies and algorithms. For a dataset of n documents $D = \{d_1, \ldots, dn\}$, it consists of performing n runs of the learning algorithm, where for each run i the algorithm

Class labels	Provision Types	Number of documents
c_0	Definition	10
c_1	Liability	39
c_2	Prohibition	13
c_3	Duty	59
c_4	Permission	15
c_5	Penalty	122

Table 1. Dataset of provision types

is trained on $D \setminus d_i$ and tested on the single left out document d_i. The loo accuracy is computed as the fraction of correct tests over the entire number of tests. Table 2 reports loo and train accuracy for the different feature selection and document representation strategies (binary (δ), term frequency (tf), infogain*term frequency ($ig * tf$), term frequency-inverse document frequency (tf-idf)).

#	repl. digit	repl. alnum	use stem	weight scheme	min freq sel.	max IG sel.	loo acc (%)	train acc (%)
0	no	no	no	δ	2	500	89.53	100
1	yes	no	no	δ	2	500	88.76	100
2	yes	yes	no	δ	2	500	88.76	100
3	yes	yes	yes	tf	2	500	91.09	100
4	yes	yes	yes	tf-idf	2	500	89.15	100
5	yes	yes	yes	ig	2	500	89.15	100
6	yes	yes	yes	ig*tf	2	500	89.15	100
7	yes	yes	yes	δ	2	250	89.92	100
8	yes	yes	yes	δ	2	100	82.55	100
9	yes	yes	yes	δ	2	50	82.17	96.12
10	yes	yes	yes	δ	2	1000	90.31	100
11	yes	yes	yes	δ	0	500	92.24	100
12	yes	yes	yes	δ	2	500	92.64	100
13	yes	yes	yes	δ	5	500	92.24	100
14	yes	yes	yes	δ	10	500	89.92	100

Table 2. Detailed results of MSVM algorithm for different document representation and feature selection strategies.

While replacing digits or non alphanumeric characters does not improve performances, the use of stemming actually helps clustering terms with common semantics. The simpler binary weight scheme appears to work better than term frequency, while significant improvements can be obtained by performing feature selection with Information Gain. The binary weight scheme appears to be the best one, probably for the small size, in terms of number of words, of the provisions in our training set; this fact makes statistics on the number of occurences of a term less reliable. Finally, Table 3 shows the confusion matrix for the best classifier,

Classes	c_0	c_1	c_2	c_3	c_4	c_5
c_0	122	0	0	0	0	0
c_1	1	9	4	0	1	0
c_2	0	3	55	0	1	0
c_3	2	0	1	6	1	0
c_4	1	1	3	0	8	0
c_5	0	0	0	0	0	39

Table 3. Confusion matrix for the best MSVM classifier.

the MSVM indexed 12, reporting prediction details for individual classes. Rows indicate true classes, while columns indicate predicted ones. Note that most errors are obtained in classes with fewer documents, for which poorer statistics could be learned.

4.1.2. Automatic provision arguments extraction

A tool called xmLegesExtractor[5] [29] of the xmLeges family has been implemented for the automatic detection of provision arguments. xmLegesExtractor is realized as a suite of NLP tools for the automatic analysis of Italian texts (see [33]), specialized to cope with the specific stylistic conventions of the legal parlance. A first prototype takes in input legislative raw text paragraphs, coupled with the categorization provided by the xmLegesClassifier, and identifies text fragments (lexical units) corresponding to specific semantic roles, relevant for the different types of provisions (Fig. 1). The approach follows a two–stage strategy. The first stage con-

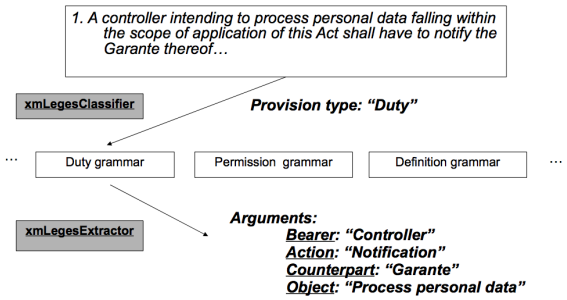

Figure 1. xmLegesClassifier combined with the grammar approach used by xmLegesExtractor.

sists in a syntactic pre–processing which takes in input a text paragraph, which is tokenized and normalized for dates, abbreviations and multi–word expressions; the normalized text is then morphologically analyzed and lemmatized, using an Italian lexicon specialized for the analysis of legal language; finally, the text is POS-tagged and shallow parsed into non–recursive constituents called "chunks". The second stage consists in the identification of all the lexical units acting as arguments relevant to a specific provision type. It takes in input a chunked representation of legal text paragraphs, locating relevant patterns of chunks which represent entities with specific semantic roles within a provision type instance, by using a specific provision type oriented grammar (Fig. 1).

Some experiments testing the reliability of xmLegesExtractor have been carried out on a subset of 209 provisions. For each class of provisions in the dataset

Class labels	Provision type	Dataset	Precision	Recall
c_2	Prohibition	13	85.71%	92.30%
c_3	Duty	59	69.23%	30.50%
c_4	Permission	15	78.95%	100.00%
c_5	Penalty	122	85.83%	89.34%
	Total	209	82.80%	73.68%

Table 4. xmLegesExtractor experiments

the total number of semantic roles to be identified are collected in a gold standard dataset; this value was then compared with the number of semantic roles correctly identified by the system and the total number of answers given by the system. Some results are reported in Tab. 4.

[5]xmLegesExtractor has been developed in collaboration with the Institute of Computational Linguistics (ILC-CNR) in Pisa (Italy)

4.2. DK construction

Lexical units identified by xmLegesExtractor represent language-dependent lexicalizations of provision arguments. More information on related entities, as well as their relations within a specific domain, can be obtained by mapping lexical units to concepts in existing Domain Knowledges (DKs), if any. On the other hand the extracted information can be consider as a ground to construct DKs (in terms of thesauri or domain ontologies). Actually the construction of them is not a specific task of *legal* ontologists, but of ontologists *tout court*, since a DK has to contain information on entities of a domain independently from a legal perspective. This aspect is important in order to conceive a legal knowledge architecture whose components can be reused. A DILK-DK learning approach only suggests language-dependent lexical units for DKs, which can be implemented by projecting lexical units on a large text corpora of a specific domain, inferring conceptualizations by term clustering, as well as using statistics on recurrent patterns for discovering term relationships. This issue is out of the paper scope; a vast literature exists on this topic, therefore the interested reader can refer to [34].

5. Conclusions

A knowledge modelling approach for the legal domain, called DILK-DK, has been presented. It aims to keep distinct domain knowledge from its legal perspective. Moreover an automatic approach based on machine learning and NLP techniques to support a bottom-up knowledge acquisition from legislative texts within the DILK-DK framework has been shown. The proposed learning approach for legal knowledge acquisition can provide the following benefits: a) it contributes to implement taxonomies or suggest concepts for hand-crafted ontologies [35]; b) it contributes to bridge the gap between authoritativeness and consensus for legal rules representation, since it is able to extract rules directly form legislative texts, which are authoritative sources (by definition), nevertheless promoting consensus, since rules are automatically extracted from legal sources, limiting human interaction.

References

[1] J. Breuker, P. Casanovas, M. Klein, and E. Francesconi, eds., *Law, Ontologies and the Semantic Web. Channelling the Legal Information Flood*, IOS Press, 2009.

[2] A. Gangemi, N. Guarino, C. Masolo, A. Oltramari, and L. Schneider, "Sweetening ontologies with dolce," in *Proc. of the 13th EKAW Conference*, 2002.

[3] R. Studer, V. R. Benjamins, and D. Fensel, "Knowledge engineering: Principle and methods," *Data Knowledge Engineering*, vol. 25, no. 1-2, pp. 161-197, 1998.

[4] R. Hoekstra, J. Breuker, M. Bello, and A. Boer, "LKIF core: Principled ontology development for the legal domain," in *Legal Ontologies and the Semantic Web*, 2009.

[5] T. Gruber, "Where the social web meets the semantic web (keynote abstract)," in *The Semantic Web – Proc. of the 5th International Semantic Web Conference*, p. 994, 2006.

[6] J. Breuker, S. van de Ven, A. El Ali, M. Bron, S. Klarman, U. Milosevic, L. Wortel, and A. Forhecz, "Developing harness," ESTRELLA Deliverable 4.6/3b, 2008.

[7] C. Biagioli and F. Turchi, "Model and ontology based conceptual searching in legislative xml collections," in *Proc. of LOAIT*, pp. 83-89, 2005.
[8] J. Breuker and R. Hoekstra, "Core concepts of law: taking common-sense seriously," in *Proc. of Formal Ontologies in Information Systems*, 2004.
[9] P. Buitelaar, P. Cimiano, and B. Magnini, "Ontology learning from text: an overview," in *Ontology Learning from Text: Methods, Evaluation and Applications*, pp. 3-12, 2005.
[10] G. Lame, "Using nlp techniques to identify legal ontology components: concepts and relations," *Lecture Notes in Computer Science*, vol. 3369, pp. 169-184, 2005.
[11] J. Searle, *Speech Acts: An Essay in the Philosophy of Language.* CUP, 1969.
[12] N. Casellas, *Modelling Legal Knowledge through Ontologies. OPJK: the Ontology of Professional Judicial Knowledge.* PhD thesis, Autonomous University of Barcelona, 2008.
[13] J. Breuker and R. Hoekstra, "Epistemology and ontology in core ontologies: FOLaw and LRI-core, two core ontologies for law," in *Proc. of EKAW WS on core ontologies*, 2004.
[14] T. Bylander and B. Chandrasekaran, "Generic tasks for knowledge-based reasoning: the "right" level of abstraction for knowledge acquisition," *International Journal on Man-Mach. Stud.*, vol. 26, no. 2, pp. 231-243, 1987.
[15] B. Chandrasekaran, "Generic tasks in knowledge-based reasoning: high-level building blocks for expert system design," *IEEE Expert*, vol. 1, no. 3, pp. 23-30, 1986.
[16] W. Clancey, "The epistemology of a rule-based expert system: a framework for explanation," TechRep. STAN-CS-81-896, Stanford University, Dep. of Computer Science, 1981.
[17] W. N. Hohfeld, "Some fundamental legal conceptions", Greenwood Press (1978)
[18] J. Rawls, "Two concepts of rule," *Philosophical Review*, vol. 64, pp. 3–31, 1955.
[19] H. Hart, *The Concept of Law.* Clarendon Law Series. Oxford University Press, 1961.
[20] A. Ross, *Directives and Norms.* London: Routledge, 1968.
[21] J. Bentham and H. L. A. Hart, *Of Laws in General.* London: Athlone, 1970 (1st ed. 1872).
[22] H. Kelsen, *General Theory of Norms.* Clarendon Press, Oxford, 1991.
[23] C. Biagioli, "Towards a legal rules functional micro-ontology," in *Proc. of LEGONT*, 1997.
[24] J. Raz, *The Concept of a Legal System*, 2^{nd} edition, Clarendon Press, 1980.
[25] G. Antoniou, D. Billington, G. Governatori, and M. Maher, "On the modeling and analysis of regulations," in *Proc. of ACIS*, pp. 20–29, 1999.
[26] T. Agnoloni, L. Bacci, E. Francesconi, W. Peters, S. Montemagni, and G. Venturi, "A two-level knowledge approach to support multilingual legislative drafting," in *Law, Ontologies and the Semantic Web*, pp. 177-198, IOS Press, 2009.
[27] L. Bacci, P. Spinosa, C. Marchetti, and R. Battistoni, "Automatic mark-up of legislative documents and its application to parallel text generation," in *Proc. of LOAIT Workshop*, pp. 45-54, 2009.
[28] E. Francesconi and A. Passerini, "Automatic classification of provisions in legislative texts," *Int. Journal on Artificial Intelligence and Law*, vol. 15, no. 1, pp. 1-17, 2007.
[29] C. Biagioli, E. Francesconi, A. Passerini, S. Montemagni, and C. Soria, "Automatic semantics extraction in law documents," in *Proc. of ICAIL*, pp. 133-139, 2005.
[30] Y. Yang and J. Pedersen, "A comparative study on feature selection in text categorization," in *Proc. of the* 14^{th} *Int. Conference on Machine Learning*, pp. 412-420, 1997.
[31] J. Quinlan, "Inductive learning of decision trees," *Mach. Learning*, vol.1, pp. 81-106, 1986.
[32] C. Buckley and G. Salton, "Term-weighting approaches in automatic text retrieval," *Information Processing and Management*, vol. 24, no. 5, pp. 513-523, 1988.
[33] R. Bartolini, A. Lenci, S. Montemagni, V. Pirrelli, and C. Soria, "Automatic classification and analysis of provisions in italian legal texts: a case study," in *Proc. of WORM*, 2004.
[34] P. Buitelaar and P. Cimiano, eds., *Ontology Learning and Population: Bridging the Gap between Text and Knowledge*, IOS Press, 2008.
[35] S. Walter and M. Pinkal, "Definitions in court decisions – automatic extraction and ontology acquisition," in *Law, Ontologies and the Semantic Web*, pp. 95-113, 2009.