2014 International Conference on Intelligent Networking and Collaborative Systems

# A Knowledge Based System for Guidance and Training on Legal Concepts

Nicola Capuano, Saverio Salerno
Dep. of Information Engineering, Electric Engineering and Applied Mathematics – University of Salerno, Italy
{ncapuano, salerno}@unisa.it

Carmen De Maio
Department of Computer Science
University of Salerno, Italy
cdemaio@unisa.it

***Abstract* — The paper describes a research aimed at defining theoretical and technological components enabling a citizen to obtain guidance and training on legal concepts starting from a textual description of a case. The defined system is able to detect relevant legal concepts from the textual description also relying on a ontology and on the enrichment of the case text with common-sense knowledge. Detected concepts are used to generate a training path aimed at providing the citizen with the basis for understanding legal issues the case deals with. The training path is then enriched with legal information like relevant laws and jurisprudence retrieved on an external legal repository. This research is part of an initiative aimed at defining innovative technologies to support on-line mediation.**

***Keywords* — *adaptive learning systems; semantic search; online dispute resolution; text analysis; knowledge representation.***

## I. Introduction

Since 1993, when first modern Italian mediation statutes were issued, mediation practice was virtually non-existent despite the high success rates of mediated cases. To improve mediation usage and to streamline the work of Italian courts, the Italian Government introduced, in 2011, mandatory pre-trial mediation of civil and commercial cases.

The new Italian mediation model is demonstrating to sensibly speed-up the settlement of disputes and there has also been significant positive attention paid to it at European level. On the other end, it needs efficient and effective tools to support the explosion of mediation cases. Citizens need also to be sensitized to the benefits of mediation and must be trained on how mediation works and how to access it.

eJRM[1] "*electronic Justice Relationship Management*" is a project, supported by the Italian Ministry of University and Research, that brings together University researchers and practitioners in the fields of law and computer science with the aim of defining, implementing and experimenting innovative technologies for on-line mediation.

Among the other things, the project provides innovative features like the possibility, for a citizen, to formalize a case in natural language or through a structured interview [1] and let the system provide guidance and training about involved legal concepts by generating a customized *Knowledge Path* (KP) integrating engaging training modules based on *story-telling* [2] with additional information resources.

The KP is purposed to let the user get a first idea about his case and to help him decide whether to sue the other party, to initiate a mediation, or to give up. The purpose of this paper is to describe the knowledge-based process of KP generation, the ontological models and techniques behind it as well as the first results obtained in their implementation.

After a brief excursus of related work in section 2, we describe in section 3 the ontological structure we defined to support the KP generation process, that is then detailed in section 4. Section 5 describes the developed prototype while concluding remarks and plans for future work follow.

## II. Related Work

Earlier attempts to use knowledge representation in legal systems date back to time when semantic technologies were still in their embryo form. Among those, we cite the *Frame-based Ontology of Law* (FBO) [3] that considers a legal system made of norms, acts and concepts, in contrast with the *FOLaw* ontology [4] adopting a functional perspective that relies on a structure made up of several different kinds of knowledge.

A more recent initiative is the project *Estrella* [5] aimed at defining a *Legal Knowledge Interchange Format* (LKIF), based on RDF and OWL, for the representation of legal concepts. LKIF is composed of about 200 concepts that can be used as starting point to define a legal system. LKIF is at the basis of *HARNESS* [6], a system able to assess if a case complies with specific legal rules.

Another project purposed to the introduction of semantic technologies to the legal field is *ICT4Law*. In this project, a *Legal Taxonomy Syllabus* [7] has been defined to represent legal information at different levels distinguishing between terms and their interlingua meanings. *Ontomedia* [8] defined instead a specific *Mediation Core Ontology* (MCO), based on OWL, for the semantic representation of legal documents acquired during the mediation steps. The developed system, basing on MCO, provides advanced document retrieval features as well as knowledge-based tools for the formal definition of new mediation cases.

*LOIS* [9] is a multilingual lexical resource in the legal field based on *WordNet* including about 35.000 concepts in five European languages. The related *DALOS* project [10] have built an ontology-linguistic resource, based on LOIS, to be used in EU legislative drafting process as well as in the national transpositions of EU Directives.

[1] http://www.ejrm.it

978-1-4799-6387-4/14 $31.00 © 2014 IEEE
DOI 10.1109/INCoS.2014.32

IEEE computer society

The analysed existing systems are mainly targeted to support legal specialists in juridical tasks like the formal definition and the legal assessment of a case, the semantic retrieval and the alignment of legal documents, etc. Few attempts have been made so far to use such technologies to support citizens with few or no juridical background in obtaining guidance and training on legal topics.

In fact common citizens can't be forced to use complex semantic tools as well as formal languages to describe the case on which they would be supported. They should use, for this purpose, user-friendly approaches like natural language or structured interviews. While adopting natural language, the "language gap" problem must be also taken into account. Common citizens do not use juridical language and, often, rely on non-appropriate terms.

Moreover, to provide guidance to the common citizen doesn't simply mean to find relevant resources with respect to the expressed case (like in traditional or semantic-based information retrieval). It also means to arrange a feasible training path between found resources to transfer, in the most effective way, the relevant knowledge about concepts connected to the input case.

The system described in this paper goes exactly in this direction, trying to provide innovative features directly to citizens mixing semantic, ontology-based approaches with models and techniques coming from adaptive learning systems [11]. It also adopts a technique based on common-sense knowledge to overcome the "language gap" issue.

## III. THE LEGAL ONTOLOGY

At the basis of the defined algorithms, there is a *Legal Ontology* able to formally describe legal concepts and link them to related training and informative resources (e.g. legal principles, precepts, sentences, learning objects, etc.).

The *Legal Ontology* is a structure composed of concepts and relations between concepts. More formally it can be defined as a graph $O(C, R_1,\dots R_n)$ where $C$ is a set of nodes that represent the concepts and each $R_i$ is a set of edges which correspond to a specific type of relation. Two sets of relation types are defined, the first (informative) is purposed to define a structured dictionary of legal terms, the second (educational) is aimed at introducing useful properties for training.

The *informative set* includes the following relations compliant to the SKOS[2] specification.

- *NT* (*a*, *b*) means that the concept *a* is a *narrower term* with respect to the concept *b* i.e. *a* has a more specific meaning that *b*;
- *BT* (*a*, *b*) means that the concept *a* is a *broader term* with respect to the concept *b* i.e. *a* has a more general meaning that *b*;
- *RT* (*a*, *b*) means that the concept *a* is generally *related to* the concept *b*.

The *educational set* includes the following relations in accordance with [12]:

- *HP* (*a*, *b*) means that the concept *b is part of a* i.e. *a* is understood if and only if every *b* so that *b* is part of *a* is understood;
- *IRB* (*a*, *b*) means that the concept a *is required by* the concept *b* i.e. a necessary condition to study *b* is to have understood *a* before;
- *SO* (*a*, *b*) means that the *suggested order* between the two concepts is that *a* precedes *b* i.e. to favour learning, it is desirable to study *a* before *b*.

Moreover, to each node $c \in C$ several information is connected: a name label $N(c)$, a textual description $D(c)$ and a weighted set of terms $T(c)$ characterizing the concept. This latter set, generated by the *enrichment* process described in III.B, is purposed to add common-sense meaning to each concept to limit the "language gap" issue described so far.

### A. *Ontology concepts definition*

To fill the defined ontological structure with concepts and relations, we started from existing legal ontologies as described in section II. Unfortunately existing models are, in some cases, very application specific, while, in other cases too abstract and purposed to describe top-level concepts like process, legal-action, role, norm, rule, etc.

Our legal ontology is instead mainly intended to classify legal topics rather then to define legal abstractions. In order to populate it, we started from the *EuroVoc*[3]: a multilingual thesaurus defined by the EC and built with SKOS for the classification of directives, laws and treaties. The current version of *EuroVoc* includes 6.883 concepts available in 23 official EU languages. To each concept a textual description is attached as well as a list of aliases.

*EuroVoc* includes hierarchical relations among concepts like *BT* (broader term) and *NT* (narrower term) as well as associative relations like *RT* (related term). The thesaurus covers all the activity fields of EU institutions. The adoption of *EuroVoc* ensures interoperability with legal databases and repositories given that it is adopted by all EC institutions and by European national parliaments. On the other end, the provided classification level is too high and it is not enough to support effective legal topic detection from text.

To improve the model, especially with respect to Italian law, we analysed other existing legal taxonomies adopted by legal portals and repositories. Among the other, we selected the *ItalGiure*[4] repository from the Italian Court of Cassation that includes Italian laws and main jurisprudence from any law sector classified on the basis of an internal taxonomy.

Given that the mediation is mainly adopted for civil law cases, we restricted our analysis in the civil section of the repository that adopts a classification based on a taxonomy of 12.701 terms. Repository resources are also classified on *EuroVoc* but the two classification are not directly linked.

[2] http://www.w3.org/2004/02/skos/

[3] http://eurovoc.europa.eu/

[4] http://www.italgiure.giustizia.it/

We imported all *ItalGiure* concepts in our model as well as hierarchical relations that have been replicated with *NT* and *BT* relations. Unfortunately *ItalGiure* concepts do not support associative relations (e.g. the *RT* relation).

To link imported *ItalGiure* and *EuroVoc* concepts inside our model, we analysed indexed legal resources. Let be $c_I$ an *ItalGiure* concept, $c_E$ an *EuroVoc* concept and *Res*(*c*) the set including all resources classified under the concept *c* in the repository, we calculate the *relatedness* between any couple of concepts $c_I$ and $c_E$ in this way:

$$Rel(c_I, c_E) = \frac{|Res(c_I) \cap Res(c_E)|}{|Res(c_I) \cup Res(c_E)|}$$

For any couple of concepts so that $Rel(c_I, c_E)$ is greater than a given threshold (heuristically set to 0,3), we added a relation $RT(c_I, c_E)$ in our model. The figure 1 shows an excerpt of the model centred on the concept of mediation. Concepts in italics come from *EuroVoc*, other concepts come from *ItalGiure*, dashed lines are added relations.

Educational information has been manually added to the model for a small subset of concepts for which training modules were available. The figure 2 shows a subset of concepts covering the theme of civil mediation from the educational point of view. Concepts in bold have been added to the model for educational purposes. Relations from the educational set have been also added.

Available legal resources are linked with ontology concepts the resource deals with. Such resources may be of different kinds: *ItalGiure* provides many legal principles, precepts and sentences to be used for informative purposes but additional resources have been provided for educational purposes and included in an external *educational repository*. Both kind of resources are used to compose a KP.

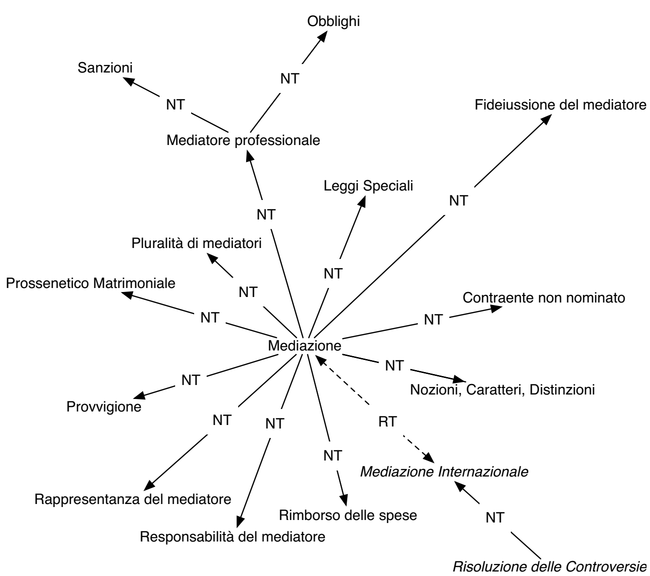


Figure 1. Excerpt of the legal ontology focused on mediation.

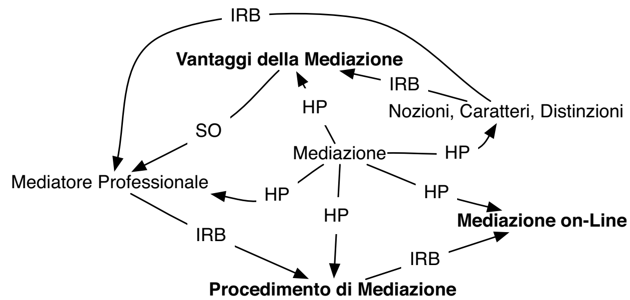


Figure 2. Excerpt of the legal ontology focused on educational relations.

## B. Ontology enrichment

This step is aimed at calculating the set of weighted terms *T*(*c*) characterizing each ontology concept exploiting common-sense knowledge held in *Wikipedia*. In other words, for each ontology concept *c*, this step calculates:

$$T(c) = \{(t_1^c, w_1^c), (t_2^c, w_2^c), \dots, (t_n^c, w_n^c)\}$$

where each $t_i$ is a link to a *Wikipedia* topic (i.e. an article page) while each $w_i$ measures the relevance of the topic with respect to the concept *c*.

To do that, the description *D*(*c*) of each ontology concept is analysed and the most relevant *n-grams* (sequences of *n* words) are selected basing on the *keyphraseness* defined as the probability for an *n-gram* to be a *Wikipedia* link [13]:

$$Key(n\text{-}gram) = \frac{Link(n\text{-}gram)}{Count(n\text{-}gram)}$$

where *Link*(*n-gram*) is the number of topics in which the *n-gram* appears as a link in *Wikipedia* while *Count*(*n-gram*) is the number of topics in which the *n-gram* appears.

All *n-grams* with a *keyphraseness* over a threshold are *candidate topic referrals* for the concept *c*. Unfortunately, a *n-gram* may refer to different *Wikipedia* topics. In order to select the right topic for each concept, a disambiguation step is needed. This is done by relying on two different measures: the *relatedness* and the *commonness*.

For each pair of topics $t_x$ and $t_y$, the sets *X* and *Y* of all hyperlinks that appear in the text of the topics are identified and their overlap $X \cap Y$ is calculated. Let be *N* the total number of *Wikipedia* topics, the *relatedness* between $t_x$ and $t_y$ is defined as follows:

$$Rel(t_x, t_y) = 1 - \frac{max(log|X|, log|Y|) - log|X \cap Y|}{N - min(log|X|, log|Y|)}$$

Moreover, the *commonness* of a topic $t_x$ with respect to a given *n-gram* is defined as:

$$Com(t_x | n\text{-}gram) = \frac{Link(n\text{-}gram | t_x)}{Link(n\text{-}gram)}$$

where *Link*(*n-gram*|$t_x$) is the number of topics in which the *n-gram* appears as a link to the topic $t_x$.

Let be *Ctx* the set of *context topics* (i.e. topics associated with candidate referrals that do not require disambiguation), we can associate a score to each topic *t* associated with an ambiguous candidate referral *n-gram* in this way:

$$Score(t|n\text{-}gram) = \frac{\sum_{t' \epsilon Ctx} Rel(t, t')}{|Ctx|} \times Com(t|n\text{-}gram)$$

For each candidate topic referral, a new pair ($t^c$,$w^c$) is so included in *T*(*c*) where $t^c$ is the referred Wikipedia topic with the maximum score according to the previous equation while $w^c$ corresponds to the value of *Key*(*n-gram*).

## IV. Knowledge Path Generation

A *Knowledge Path* (KP) is composed of a training path and a set of additional information resources. Basing on the legal ontology described in section III, a KP is generated from the description of a legal case. The KP generation process can be summarized in three steps described below.

- *Extraction of legal concepts*. In this step, the system tries to detect relevant legal concepts connected to the current legal case from the legal ontology. A weight is associated to each detected concept and expresses its relevance in the current case.
- *Training path generation*. Starting from extracted legal concepts, from the legal ontology and available learning objects, the system generates a *training path* aimed at providing the basis for understanding legal issues dealt with the input case.
- *Information resources addition*. Starting from extracted legal concepts, from the legal ontology and information resources included in *ItalGiure*, the system enriches the *training path* with relevant legal information like legal principles, precepts and sentences.

The Next sub-sections provide additional details about all these steps.

### A. Extraction of legal concepts

This step is aimed at the detection of legal concepts from a case text *s*. The output is a set *C*(*s*) of pairs ($c_i$, $\mu_i$) where $c_i$ is a concept from *C* and $\mu_i$ specifies the relevance of $c_i$ within the case text. The first operation is the extraction of a list of weighted terms from *s*:

$$T(s) = \{(t_1^s, w_1^s), (t_2^s, w_2^s), \dots, (t_n^s, w_n^s)\}$$

where, as seen is section III.B, each $t_i$ is a link to a *Wikipedia* topic while each $w_i$ measures the relevance of the topic with respect to the case text. The process to calculate *T*(*s*) is the same seen to calculate *T*(*c*) but using the case text rather then the concept description as input.

According to [14], to detect concepts in *s*, for each *c* ∈ *C*, a matching is performed between *T*(*s*) and *T*(*c*) through a combination of the standard measures of *precision* (*P*) and *recall* (*R*) [15] with the following equations:

$$P(s,c) = \frac{\sum_{t_j^c = t_k^s} \left(w_j^c * w_k^s\right)}{\sum_{t_j^c = t_k^s} \left(w_j^c\right)} \quad R(s,c) = \frac{\sum_{t_j^c = t_k^s} \left(w_j^c * w_k^s\right)}{\sum_{t_j^c = t_k^s} \left(w_j^s\right)}$$

The relevance score $\mu$ between *c* and *s* is then calculated with the following equation:

$$\mu = \frac{1}{|T(s)|} R(s,c) + \left(1 - \frac{1}{|T(s)|}\right) P(s,c)$$

Concepts from the legal ontology that have a relevance score over a given threshold (heuristically set to 0,5) are added to *C*(*s*) together with the relevance score itself.

### B. Training path generation

The generation of the *training path* is done starting from the set of concepts *C*(*s*) extracted from the case text. The process, according to [16], is made in three subsequent steps.

The *first step* is aimed at building, from the ontology *O*, the simplified graph *O'*(*C*, *HP'*, *IRB'*, *SO'*) where *HP'* is the inverse relation of *HP*, *IRB'* and *SO'* are initially set to *IRB* and *SO* but they are modified by applying the following rule: each arc *ab* ∈ *IRB'* ∪ *SO'* is substituted with arcs *ac* for all *c* ∈ *C* such that there exist a path from *c* to *b* on the arcs from *HP'*. The figure 3a shows the graph *O'* obtained from the ontology reported in figure 2.

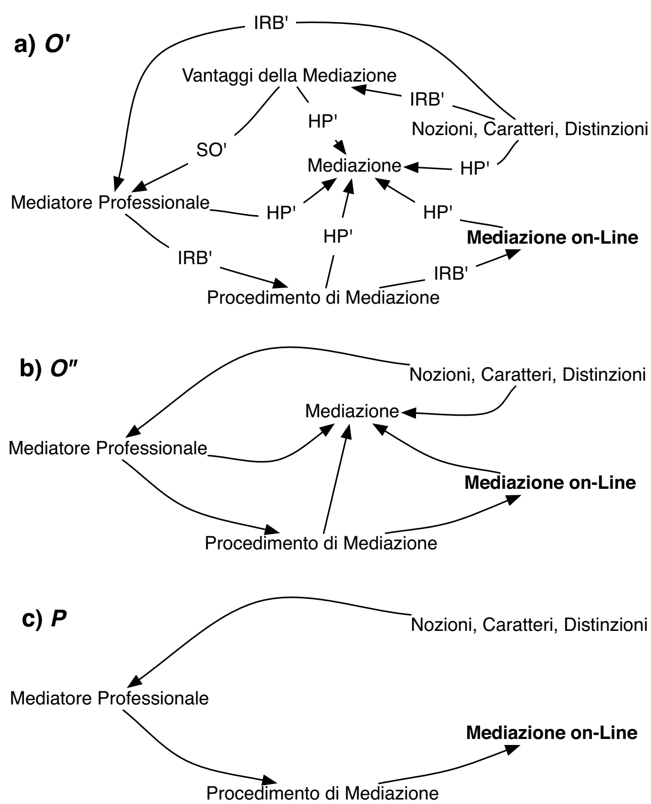


Figure 3. Example of the training path generation process.

The *second step* is aimed at building the graph $O''(C', R)$ where $C'$ is the subset of $C$ including all concept that must be taught according to $C(s)$ i.e. $C'$ is composed by all nodes of $O'$ from which there is a ordered path in $HP' \cup IRB'$ to concepts in $C(s)$. $R$ is initially set to $HP' \cup IRB' \cup SO'$ but all arcs referring to concepts external to $C'$ are removed. The figure 3b shows the graph $O''$ obtained from $O'$ starting the target concept "*Mediazione on-Line*".

The *third step* finds a linear order between nodes of $O''$ by using depth-first search so by visiting the graph nodes along a path $P$ as deep as possible. Then it deletes from the obtained path all non-atomic concepts i.e. all concept $a$ so that $ab \in$ HP for some $b$. This ensures that only leaf concepts (with respect to the *HP* relation) will be part of path $P$. The figure 3c shows the path $P$ obtained from the graph $O''$.

The system then finds feasible learning resources linked to concepts in $P$. The obtained sequence is the *training path*.

### C. Information resources addition

In this step, once the *training path* has been generated, additional informative resources are added to it in order to obtain the KP. To do that, first of all, the set $D$ of legal resources connected to each concept $c_i$ of $C(s)$ is retrieved from *ItalGiure*.

Let be $\{c_1^d, c_2^d, \ldots, c_m^d\}$ the list of all concepts connected with a resource $d \in D$, a matching rate $rank_d$ is calculated for each retrieved resource basing on the relevance $\mu_i$ associated with each concept $c_i$ of $C(s)$ through the following equation:

$$rank_d = \sum_{c_i^d = c_j} \mu_j$$

Legal resources with a rank value greater then a threshold (heuristically set to 0,5) are added to the *training path*. The obtained structure is the resulting KP.

## V. Prototype and Evaluation

The figure 4 shows a screenshot of the developed system. The user can write the legal case text in the uppermost box. In the box below, concepts extracted from text are displayed together with the relevance score $\mu$. In this case, a legal case about a car accident is provided as input and the concepts of *insurance*, *provisional driving licence* and *circulation of vehicles* are detected by the system with the higher score.

By pressing on each concept, the legal ontology on the right box opens and displays the position of each concept within the *NT* hierarchy. Moreover, in the lowermost box, connected legal resources are displayed. By pressing on every resource it is possible to display it from the *ItalGiure* repository.

By pressing on the *training path* button, the user can access another page where he can follow a beginner course generated on the fly to cover principles connected with extracted concepts (*on-line mediation* and *civil liability of the motorist* in the specific case). Within this page (see figure 5) he can chose any learning object from the sequence on the left and read the selected resource in the main panel.


Figure 4. Snapshot of the prototype system.

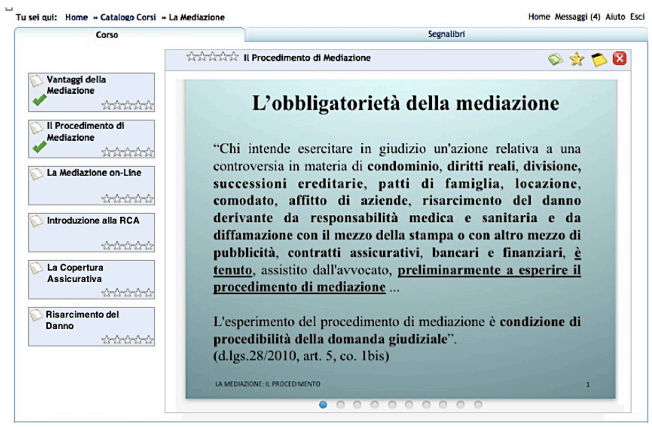


Figure 5. Snapshot of the training path delivery page.

A first experimentation has been performed to evaluate the effectiveness of the search process connected with the *information resources addition* phase described in IV.C. To evaluate algorithm performances, we analysed 50 queries generates from different input text.

The performance of the approach is assessed in terms of precision and recall measures, considering the analysis through micro-average of the individual precision-recall curves [15]. Let be $Q = \{Q_1, Q_2, \ldots, Q_n\}$ a set of queries, $D$ all the relevant resources for the given set of queries $Q$. For each query $Q_i$, we consider $\lambda = 20$ steps up to its maximum recall value and measure the number of relevant documents retrieved at each step $\lambda$.

The micro-averaging of recall and precision (at the generic step $\lambda$), is defined as follows:

$$Rec_\lambda = \sum_{Q_i} \frac{|R_{Q_i} \cap B_{\lambda,Q_i}|}{|R|} \quad Prec_\lambda = \sum_{Q_i} \frac{|R_{Q_i} \cap B_{\lambda,Q_i}|}{|B_\lambda|}$$

where $R_{Qi}$ is the set of relevant resources for a given query $Q_i$, $B_\lambda$ the set of retrieved resources at the step $\lambda$ and $B_{\lambda,Qi}$ is the set of all relevant resources, retrieved at the step $\lambda$, for the query $Q_i$.

The figure 6 shows the tendency of the micro-average of recall/precision curve evaluated on the collection set, and compares our approach with a well known keyword-based search engine called *Lucene*[5]. Let observe that the system performance in term of data retrieval is near to *Lucene*. In particular, the precision is quite constant, it doesn't decrease meaningfully when the recall grows.

It is important to stress out that the effectiveness and accuracy of the semantic search is strongly dependent on the quality of the underlying legal ontology, along with the enrichment mechanism. The latter, as testified by the results, improves its effectiveness as the length of the input text query increases, since the more information is available, the better the tool is able to correctly understand its context and come up with meaningful concepts.

## VI. CONCLUSION

We have presented in this paper original research results on the automatic construction of a knowledge path to provide guidance and training on legal concepts,. The approach here proposed starts with the detection of legal concepts (from a legal ontology) on a case expressed in natural language, continues with the creation of a training path covering such concepts with feasible learning resources and concludes with the addition of informative legal resources retrieved on an external legal repository.

In order to fill the *language gap* between legal language and common citizen language, an enrichment process has been proposed to add common-sense meaning taken from *Wikipedia* to each ontology concept.

We have also described a first prototype implementation of defined models and techniques. It is worth noting that, despite the current version of the prototype implements all theoretical components, the user interface is still in a embryo form and it is currently being improved. The integration with other components for on-line mediation of the eJRM project is also in progress.

Preliminary experimentation results of the sematic search features behind the *information resources addition* step have been presented in V while additional results on the legal concept extraction phase are presented in [14]. Additional experimentation steps, also with real users as well as in conjunction with other eJRM modules, are planned for the next months and will be reported in a future work.

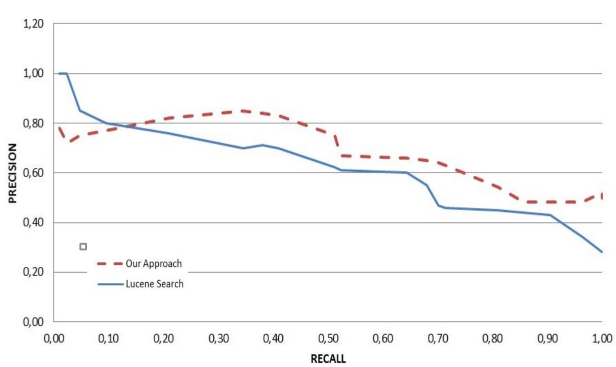


Figure 6. Micro-averaging precision/recall

[5] http://lucene.apache.org


## ACKNOWLEDGMENT

This work has been partially supported by the eJRM (electronic Justice Relationship Management) project, co-funded by the Italian Ministry of Instruction and Research (ref. PON 01-01286).


## REFERENCES


[1] G. Arosio, G. Bagnara, N. Capuano, E. Fersini, D. Toti, "Ontology-driven Data Acquisition: Intelligent Support to Legal ODR Systems", in Proc. Int. Conf. on Legal Knowledge and Information Systems (JURIX 2013), IOS Press, pp. 25-28, 2013.

[2] N. Capuano, C. De Maio, A. Gaeta, G.R. Mangione, S. Salerno, E. Fratesi, "A Storytelling Learning Model for Legal Education", in Proc. Int. Conf. on e-Learning (e-Learning 2014), ACM, 2014.

[3] R.W. van Kralingen, "A conceptual frame-based ontology for the law". Proc. Int. Workshop on Legal Ontologies (LEGONT 97), pp.15-22, 1997.

[4] A. Valente, J. Breuker, "A functional ontology of law", in Towards a Global Expert System in Law, Cedam Publishers, 1994.

[5] R. Rubino, A. Rotolo, G. Sartor "An OWL ontology of fundamental legal concepts" in Proc. Int. Conf. on Legal Knowledge and Information Systems (JURIX 2006), IOS Press, pp. 101-110, 2006.

[6] S. van de Ven, R. Hoekstra, J. Breuker, L. Wortel, A. El-Ali, "Judging Amy: Automated legal assessment using OWL 2", in Proc. of OWL: Experiences and Directions (OWLED 2008), 2008.

[7] G. Ajani, G. Boella, L. Lesmo, M. Martin, A. Mazzei, D.P. Radicioni, P. Rossi, "Legal Taxonomy Syllabus version 2.0", Proc. 3rd ICAIL Workshop on Legal Ontologies and Articial Intelligence Techniques, CEUR, 2009.

[8] M. Poblet, P. Casanovas, J.M.L. Cobo, "Online Dispute Resolution for the Next Web Decade: The Ontomedia Approach", in Proc. 10th Int. Conf. on Knowledge Management and Knowledge Technologies, 2010.

[9] M.A. Biasiotti, M.T. Sagri, D. Tiscornia "LOIS: Building a Multilingual Wordnet for the Legal Domain", in Proc. 4th Int. EGOV Conference 2005, Aug. 22-26, 2005.

[10] E. Francesconi, D. Tiscornia, "Building Semantic Resources for Legislative Drafting: The DALOS Project", in Computable Models of the Law, LNCS, Springer, vol. 4884, pp 56-70, 2008.

[11] N. Capuano, M. Gaeta, P. Ritrovato, S. Salerno, "How to Integrate Technology Enhanced Learning with Business Process Management", J. of Knowledge Management, Emerald, vol. 12 n. 6, pp. 56-71, 2008.

[12] N. Capuano, G.R. Mangione, A. Pierri, S. Salerno, "Personalization and Contextualization of Learning Experiences basing on Semantics", Int. J. of Emerging Technologies in Learning, vol. 9, n. 7, pp. 5-14, 2014.

[13] D. Milne, I. Witten, "An open-source toolkit for mining Wikipedia" in Proc. New Zealand Computer Science Research Student Conf., NZCSRSC, 2011.

[14] N. Capuano, C. De Maio, S. Salerno, D. Toti, "A Methodology based on Commonsense Knowledge and Ontologies for the Automatic Classification of Legal Cases", Proc. Int. Conf. on Web Intelligence, Mining and Semantics (WIMS 2014), 2014.

[15] C.J. Van Rijsbergen "Information retrieval" (2nd ed.). University of Glasgow: Dept. of Computer Science, 1979.

[16] N. Capuano, M. Gaeta, A. Marengo, S. Miranda, F. Orciuoli, P. Ritrovato, "LIA: an Intelligent Advisor for e-Learning", Interactive Learning Environments, T&F, vol. 17, n. 3, pp. 221-239, 2009.