D5.2

Preliminary exploitation plan

Grant Agreement nº:	690974
Project Acronym:	MIREL
Project Title:	MIning and REasoning with Legal texts
Website:	http://www.mirelproject.eu/
Contractual delivery date:	30/04/2017
Actual delivery date:	02/05/2017
Contributing WP	5
Dissemination level:	Confidential
Deliverable leader:	APIS
Contributors:	APIS, UL

Document History

Version	Date	Author	Partner	Description
0.1	10/04/2017	Hristo Konstantinov	APIS	Initial draft
0.2	26/04/2017	Livio Robaldo	UL	Final draft
1.0	02/5/2017	Hristo Konstantinov	APIS	Final Version

Contributors

Partner	Name	Role	Contribution
APIS	Hristo Konstantinov	Coordinator of the deliverable	Coordinating the writing of the deliverable; Interacting with the other partners to collect data.
UL	Livio Robaldo	Editor	Editing many sections of the deliverable; Interacting with the other partners to collect data.

Table of Contents

Executive Summary

This deliverable describes preliminary exploitation solutions and deployed scenarios of the first sixteen months of activity of the project MIREL. The secondments in MIREL were devoted to exchange knowledge and expertise between the partners fit to (a) develop joint theoretical research, (b) devising computational solutions which can be used in the legal informatics market by the three industrial partners of the consortium, (c) defining, preparing, and submitting new joint research project proposals.

1 Introduction

The secondments in MIREL during the first year of the project were devoted to exchange knowledge and expertise between the partners fit to (a) develop joint theoretical research, (b) devising computational solutions which can be used in the legal informatics market by the three industrial partners of the consortium, (c) defining, preparing, and submitting new joint research project proposals. The deliverable is structured according to (a)-(b)-(c).

In the first section, titled "Exploitation of research results", we list the main research achievements for which it could be identified some kind of exploitation, intended as further research activities to be carried out beyond the network of MIREL.

Since most MIREL partners are universities or research centers, the first kind of exploitation we consider in MIREL is of course the definition of new research projects, jointly defined between MIREL partners and possibly other non-MIREL partners, which has been involved thanks to the MIREL network and/or the events, organized in the context of the project (conferences, tutorials, etc.). Joint project proposals, most of which are built on top of the research results listed in section 2, are listed in section 4 "Joint research project proposals".

On the other hand, MIREL also involves three industrial partners: APIS EOOD, Nomotika SRL, and DLVSystem SRL. MIREL partners consider industrial exploitation of the research results crucial for the advancement in legal informatics, therefore lot of efforts have been devoted to investigate how the research achievements of the academic partners of MIREL can be used to enhance the performance of the products of the three companies and to devise new projects involving the partners. The industrial exploitation is described in section 3 "Industrial exploitation".

2 Exploitation of research results

2.1 LegalRuleML

On the 14th of March 2017, the OASIS LegalRuleML XML format has become a standard, after a long standardization process carried out by the MIREL partners UNIBO and Data61.

LegalRuleML (cf. [Athan et al, 2015]) defines a rule interchange language for the legal domain. The work enables modeling and reasoning that allows implementers to structure, evaluate, and compare legal arguments constructed using the rule representation tools provided. LegalRuleML extends the standard RuleML[1] with formal features specific to legal norms, guidelines, policies and reasoning; The standard, expressed with XML-schema, is able to represent the particularities of the legal normative rules with a rich, articulated, and meaningful markup language, such as: (1) defeasibility of rules and defeasible logic, (2) deontic operators (e.g., obligations, permissions, etc.), (3) temporal management of rules and temporality in rules, (4) classification of norms (i.e., constitutive, prescriptive), (5) isomorphism between rules and natural language normative provisions, (6) identification of parts of the norms (e.g., bearer, conditions), and (7) authorial tracking of rules.

> Together with the legal XML standard Akoma Ntoso[2], LegalRuleML will be used in several future research activities and projects; for instance, it will be used in the projects DAPRECO and euLinkedDataHub described below in subsection 4.1 and 4.2 respectively as a interchange format to link the legal documents tagged in Akoma Ntoso with legal ontologies and with formulae representing the meaning of the norms.

2.2 Reified Input/Output logic

Reified Input/Output logic [Robaldo and Sun, 2017] is a new formalism for representing natural language norms in a deontic settings. This incorporates main insights of Input/Output logic [Makinson and van der Torre, 2000], a well-known formalisms in deontic logic, with techniques coming from the literature in Natural Language Semantics, in particular Reification [Davidson, 1967]. The use of reification for the proper representation of norms have been investigated in WP1 and WP2 of MIREL.

Specifically, reified Input/Output logic combines Input/Output logic with the reification-based approach of Jerry R. Hobbs [Hobbs, 2008]. The latter is a wide-coverage logic for Natural Language Semantics able to handle a large set of linguistic phenomena into a simple logical formalism. Reified Input/Output logic represents the first step of a long-term research

---

[1] http://wiki.ruleml.org/index.php/RuleML_Home

[2] www.akomantoso.org

aiming at filling the gap between Input/Output logic and the richness of Natural Language Semantics.

> We plan in our future work to exploit reified Input/Output logic as the underlying formalism for applications in legal informatics to process and reason on existing legal texts, which are available in natural language only. We will start pursuing this line of research in the context of the project DAPRECO, described below in subsection 4.1.1.

2.3 Text similarity for identification of National Implementing Measures

The partner UNITO has applied his long-standing expertise in semantic text similarity techniques to automatically detect the transposition of European Union (EU) directives into the national law (National Implementing Measures, NIMs henceforth). The latter has been spotted by the MIREL partner APIS as an important pivotal service for current legal document management systems such as EuroCases[3] (cf. subsection 3.1); the company has offered access to UNITO to its data in order to train and evaluate the system developed.

A PhD student working at UNITO, Rohan Nanda, has developed a text similarity approach for automated transposition detection of EU directives [Nanda et al., 2016], which is now the topic of his PhD thesis. The student belongs to the Joint International Doctoral Degree in Law, Science and Technology (LAST-JD)[4], whose principal investigator is UNIBO and where the MIREL partners UNITO, UL, INRIA, Nomotika, and Data61 are involved.

The effective transposition of EU law into Member States is important in legal informatics in order to achieve the policy goals defined in the Treaties and secondary legislation. NIMs are the legal texts officially adopted by the Member States to transpose the provisions of an EU directive into national law. The European Commission is responsible to monitor the NIMs to ensure the correct implementation of the directive. However, the measures undertaken by the Commission to monitor NIMs are time consuming and expensive, as they resort to manual conformity checking studies and legal analysis. In the system developed at UNITO uses semantic similarity techniques to automatically identify the transposition of EU directives into the national law at a fine-grained provision level. The system is drawn from and employs techniques coming from the NLP literature, such as [Dumais, 2004].

> The desired goal of the research is to act as a legal support tool for legal professionals and EC officials carrying out cross-border legal research. The tool would identify specific provisions of NIMs which transpose a particular article of the directive for different

---

[3] http://eurocases.eu/

[4] http://www.last-jd.eu/

Member States. The industrial partner APIS is interested in further development and exploitation of the legal support tools, and the extension of its functionality towards identification of similar judgments to a particular court decision and of similar provisions of national law. In connection with this, the partners UNITO, APIS, and UL submitted an H2020 project on Big Data integrating the tool for NIMs identification proposed in [Nanda et al., 2016]. The project proposal is described below in subsection 4.2.

2.4 Named Entity Recognition

The partners UNITO, UL, and UNIBO have further developed the NLP procedures for identifying named entities in legal texts proposed in [Robaldo et al, 2012]. The procedures exploit the dependency parser TULE[5], developed at UNITO in the past decades, and extend its knowledge base fit to recognize specific entities such as proper names of persons, organizations and institutions, locations (cities, regions, etc.), quantities, and so on, which are of paramount importance for effectively processing legal text. TULE does not include an effective NER module. It simply tags proper names according to certain syntactic heuristics mostly based on the occurrence of capital letters. The rule-based NER module of TULE has been developed by Livio Robaldo (UL) within the project ProLeMAS[6].

The new version of the system originally proposed in [Robaldo et al, 2012] is able to output file in Akoma Ntoso, an XML legal standard developed at UNIBO. As pilot case, it has been used within the project OpenSentenze[7] of the partner UNIBO, in order to semi-automatically anonymize case law from the Italian Court of Cassation before being published in the website. It will be used and further extended in the project euDataHub, described below in subsection 4.2.

Also the partner APIS worked on Named Entity Recognition, and specifically on Task 2.2 (develop NLP systems for mining named entities and concepts, in order to populate the ontology), together with the MIREL partner UNITO. During the first year of the project, APIS completed all its secondments on MIREL, by seconding five developers of the company at UNITO. During the secondments, it was decided that the work of APIS' developers in Task 2.2 will be focused on mining legal texts from the EuroCases database (see subsection 3.1), identifying Named Entities and EuroVoc[8] terms in them (including detecting legal citations to other legal documents) and linking identified legal terms to EuroVoc ontology. This will contribute to the semantic annotation of texts which is a

---

[5] http://www.tule.di.unito.it

[6] http://www.liviorobaldo.com/prolemas.html

[7] http://www.opensentenze.it

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

precondition for further mining tasks, such as identification of similar norms, similar documents (similar legislative instruments or court decisions) or conflicting provisions.

Furthermore, APIS' programmers further improved and developed the existing NLP tools of APIS for Named Entity Recognition (NER) of legal citations in EU legal instruments, in particular in their Italian language version, as well as the tools for recognizing concepts and terms from existing ontologies and thesauri such as the European Legal Taxonomy Syllabus [Ajani et al., 2017]. A web-based service (REST API) for recognition of legal citations to EU legal instruments in the EU legal documents has been implemented; that service which will be used by UNITO for further NLP tasks. The work on the recognition of legal citations in the Italian language version of EU legal instruments was coordinated with dr Sabrina Praduroux, a research fellow at UNITO. To this end, she provided a set of guidelines for the syntax of legal citations to EU law in Italian language. More details are provided below in subsection 3.1.

> We plan in the future to further extend the NER modules developed and use them in our future research project and activities. The first project where we will use it will be euLinkedDataHub, described below in subsection 4.2, where UL and UNITO will extend and tune the NER module fit to recognize company data for regulatory compliance and risk assessment.

2.5 Ontology Learning

The ontology learning part of the MIREL project is the core of WP2, thus representing an important task for the whole pipeline. Moreover, this task is also strictly connected with the use cases of WP4, in particular those about technical documents and multilingual corpora of norms.

The partner UNITO has been working on Ontology Learning techniques since years, proposing solutions such as [Di Caro et al., 2008] and [Boella et al, 2016b] and [Boella et al, 2016c] in the legal domain. In the MIREL project, UNITO extended some of its approaches for keyphrases extraction. In particular, these works have been exploited in further research such as [Adebayo et al., 2016] for automatically extracting keyphrases from legal text.

This work is also in line with those in WP3, with the aim of producing ontological information (automatically or semi-automatically extracted from legal corpora and existing ontologies) which can be used for the reasoning part (compliance, error-checking, inference, and so forth). Serena Villata, working at the MIREL partner INRIA and leader of WP4, has also extensively worked on these aspects in collaboration with Laura Alonso Alemany from University of Cordoba. The resulting work, jointly written from INRIA and the University of Cordoba [Cardellino et al, 2017, 2017a, 2017b], uses alignments of open-domain and legal-domain ontologies (Wikipedia and LKIF) proposes an ontology population approach

for legal ontologies. They exploit Wikipedia as a source of manually annotated examples of legal entities, and they align YAGO, a Wikipedia-based ontology, and LKIF, an ontology specifically designed for the legal domain. Through this alignment, the LKIF ontology is populated with the aim to obtain examples to train a Named Entity Recognizer and Classifier to be used for finding and classifying entities in legal texts. Since examples of annotated data in the legal domain are very few, they apply a machine learning strategy called curriculum learning aimed to overcome problems of overfitting by learning increasingly more complex concepts. In addition to Wikipedia, cases from the corpus of the European Court of Human Rights have been annotated as well. Results are satisfying and foster further research in this direction.

These results will be exploited as follows: (i) assessment of the alignment YAGO-LKIF to be improved with the help of the legal experts in UNIBO and UNLP, (ii) the produced resources are the basis for several tasks requiring named entity recognition and linking, (iii) definition of the guidelines for an annotation methodology for legal entities in legal text, and (iv) implementation of a pipeline that provides the output of this system to an ontology-based reasoned as those resulting from WP3 so that this NLP component is the input of the reasoning one.

These technologies will be used to improve the legal information extraction and visualization modules in the upcoming version of the MenslegiS system, described in Section 3.2.

Finally, the ontology learning and population mechanisms developed in MIREL will be crucial for extending existing research activities of the WP2 and WP3 partners, such as [Cardellino et al, 2015], [Di Caro and Boella, 2016], and [Di Caro and Boella, 2015].

2.6 Ontology alignment and restructuring

The MIREL project has carried out some extension of preliminary works of UNITO in the topics of aligning and restructuring ontologies. In particular, UNITO focused on one of the most relevant ontologies in the legal domain, i.e., EuroVoc.

EuroVoc is a 23-languages thesaurus that covers the activity of the European Union. It contains more than 6000 descriptor terms that are organized in a 8-levels hierarchy. The main advantage of EuroVoc consists in the assignment of a set of descriptors to documents, which can be used to perform documentary search, document management and cross-language alignment. Unfortunately, descriptor terms are not in mutual exclusion with each others, sharing contextual similarities that may lead to skewed data, especially in document classification. UNITO presented a novel method of corpus analysis using Latent Dirichlet Allocation (LDA) in order to study whether and how EuroVoc descriptors capture actual documents content [Di Caro and Siragusa 2016].

In 2017, Giovanni Siracura will be seconded at Stanford University in order to further research on the techniques presented in [Di Caro and Siragusa 2016], possibly together with researchers at Stanford University, in particular those belonging to the Stanford NLP group[9].

2.7 Rule Extraction from Legal Documents

The collaboration between the MIREL partners INRIA and Data61, together with a joint collaboration with the University of Trento, led to the design and implementation of a framework that, by combining different Natural Language Processing techniques, is able to extract of rules from legal documents, i.e. to identify in legal texts conditions in natural language describing what is permitted, forbidden or mandatory in the context they regulate, and formalize them in a set of machine-readable conditions.

Specifically, the approach combine the linguistic information provided by WordNet[10] together with a syntax-based extraction of rules from legal texts, and a logic-based extraction of dependencies between chunks of such texts. Such a combined approach leads to a powerful solution towards the extraction of machine-readable rules from legal documents. The NLP tools used for processing the linguistic information are the Stanford parser[11] and the Combinatory Categorial Grammar (CCG) parser tool including the Boxer framework [Curran et. al, 2007], which have been extensively described in MIREL deliverable "2.1 - Collection of state-of-the-art NLP tools for processing of legal text".

The approach has been evaluated over the Australian "Telecommunications consumer protections code", at disposal of the partner Data61, and it will be used within the project NORSEW, described below in subsection 4.7.

In the future, INRIA and Data61 plan to extract and formalize the rules of the "Telecommunications consumer protections code" in LegalRuleML, the XML legal standard described above in subsection 2.1.

2.8 Conceptual and formal frameworks for legal informatics

University of La Plata (UNLP) carried out theoretical research with the MIREL partner UNIBO, in the context of the secondments from UNLP done in the first year of the project and currently ongoing. The secondments from UNLP to UNIBO, all planned in the first

---

[9] https://nlp.stanford.edu/
[10] https://wordnet.princeton.edu/
[11] http://nlp.stanford.edu/software/lex-parser.shtml

biennium of the project, have been rescheduled so that other researchers from University of La Plata could be seconded in UNIBO, specifically prof. Menni and prof. Mainero.

Prof. Menni is a Computer Scientist dedicated to research in mathematics and logic, having a long-standing collaboration with prof. Smith in research on modal logic; Prof. Mainero is a linguist with 5-years work expertise in administrative norms applications and design. Their secondments in UNIBO will be devoted to combine computer science background, mathematical expertise in logic and solid experience in high-level government administration, sharing expertise towards the objectives of WP1.

> Prof. Menni and Prof. Smith will continue the research on the relation between classical and multi-relational semantics. Specifically, the idea is to prove that the latter derives from the former via standard categorical (in the sense of Category Theory) constructions. In more detail, one very basic categorical construction is that of SLICING. The category of Kripke frames embeds in the topos E of non flexive graphs. The conjecture is that the category of multi relational frames (with relations indexed by a set I) embeds into the slice category E/I' where I' is an object in E. Interestingly, the object I' needed to make this precise seems not to be a Kripke frame. This research will bring benefits to both UNIBO, whose researchers could devise new future works for the research they conducted on WP1 also together the MIREL partner Data61, e.g. [Governatori et al., 2016], [Rotolo and Calardo, 2016], and to UNLP, that will improve the courses on Logic Programming and on Category Theory at UNLP in light of the results obtained in this research.
>
> Joint publications and research between prof. Menni (UNLP), prof. Smith (UNLP), prof. Rotolo (UNIBO) and prof. Calardo (UNIBO) on these topics are currently ongoing and will continue after the MIREL project.

2.9 Legal market studies with Nuance-SF

During the first year of MIREL, a researcher from UNITO (S. Pradoroux, secondment #58) has been seconded to Stanford University. During that secondment, a new contact was made with Nuance San Francisco, a world top company in the NLU sector, which was crucial for a novel interaction of the partners (and goals) of the project with Nuance San Francisco.

> This collaboration led to a publication [De Paiva et al., 2016], and future plans within the scope of the project. [De Paiva et al., 2016] survey the legal tech market, particularly the ones in Silicon Valley, classifying and analyzing a number of legal start-ups operating there.
>
> The connection between UNITO and Nuance Communications (specifically, the Nuance office in Turin) began before the MIREL project in a completely different context, i.e., that of automatic identification of correction rules in the automotive domain. In this sense, MIREL strengthened this important and fruitful academy-company asset, which will also continue even out of the reciprocal goals in legal-informatics.

2.10 Large-scale normative reasoning

During the first year of the project, the MIREL partner HUD, leader of WP3, hosted two researchers from UNS and one researcher from Data61. In turn, a researcher from HUD has been seconded to UCT. These secondments were devoted to investigate various research topics, among which:

- Analysis of the structure of normative rules and the corresponding representation using OWL-2 [12], while addressing issues such as types of rules (e.g. obligation, prohibition, permission), priority of rules in legal domain (e.g., lex specialis and lex posterior), representation of rules and automatic detection of exceptions and priorities.

- Complexity of reasoning over normative knowledge, including reasoning under uncertainty and probabilities in legal domain, which is high even in simple cases, and possible approaches to handle it (e.g. Answer Set Programming, defeasible logic, argumentation, etc.)

- Modelling of obligations and their compliance/violation, in terms of modalities in Defeasible Logic Programming (DeLP)[13] and representation of complex scenarios that may arise in the legal domain via an Abstract Argumentation System [Vreeswijk, 1997] with recursive attack and supports relations (the ASAF formalism), with a special investigation on the advantages offered by ASAFs with respect to other formalisms such as defeasible logic.

- Parallelization techniques for inconsistency-tolerant query answering on ontological databases, expressed in the language of Datalog+/-[14]. As managing inconsistency is a key issue when reasoning with information from the legal domain, it has been planned to validate the inconsistency-tolerant query answering algorithms on ontologies built from data extracted from legal texts.

- Investigating the relationships between CSIRO defeasible logic reasoner, SPINdle[15], and the distributed reasoning engine developed by HUD. It has been studied the algorithms to be used, current limitations/constraints of both implementations, and the feasibility of merging the two reasoners to support the reasoning of legal norms as specified by the package. Besides, a use case (which is reasonable large) has also been identified and will be used as one of the test case in the future.

HUD and the other MIREL partners involved in these research activities intend to carry out research work on modelling regulations and reasoning with legal knowledge. A particular focus of our work will be around "large-scale normative reasoning" (normative

---

[12] https://www.w3.org/TR/owl2-overview
[13] http://lidia.cs.uns.edu.ar/delp_client/
[14] https://en.wikipedia.org/wiki/Datalog
[15] http://spindle.data61.csiro.au

reasoning with big data), a term we intend to coin and develop. This poses major research challenges as normative reasoning approaches thus far work in centralized main memory, which cannot work in the face of big data. HUD has specific plans with Data61 on large-scale normative reasoning, with a focus on the telecommunication and financial domains. Joint exploitation through combining our areas of expertise is under discussion also with UNITO and UL, on normative reasoning and compliance in conjunction with geodata, while, with UNIBO, on large-scale normative reasoning in the context of cloud computing.

HUD has a firm intention to seek national research funding in this area from EPSRC[16]. Furthermore, HUD get in touch with Thomson Reuters[17] who have expressed interest in the MIREL project. Future discussions will focus around identifying concrete areas where collaboration between them, HUD, and possibly other MIREL partners, can be established.

3 Industrial exploitation

3.1 APIS

The participation of APIS within MIREL has been mostly committed to the extension of the EuroCases[18] system. EuroCases is a legal informational service providing access to case law related to the application of European Union law. It is multi-lingual, in that it can presents information in several available languages, and it is accessible via Web. EuroCases has been registered as trade mark and distributed by APIS, and it is built upon the achievements of the company within the FP7 EUCases[19] project to which it participated together with the MIREL partners UNITO, Nomotika, and UNIBO. EuroCases is continuously updated by the company and used in the context of the research and industrial projects where it is involved, e.g. MIREL.

Specifically, as explained above in subsection 2.4, in 2016 APIS completed all its secondments in MIREL (all hosted by the MIREL partner UNITO), and integrated the results obtained and the techniques learnt during these secondments within the EuroCases system.

The EuroCases system has been enriched with the following functionalities:

---

[16] https://www.epsrc.ac.uk/
[17] https://www.thomsonreuters.com/
[18] http://eurocases.eu/
[19] http://eucases.eu

- Improvement and further development of the existing NLP tools used in EuroCases for Named Entity Recognition (NER) of legal citations in EU legal instruments, in particular in their Italian language version. The extension to legal citations in the Italian language was carried out via a collaboration with dr Sabrina Praduroux, a legal expert from UNITO who provided guidelines for legal citations to EU law in Italian language.

- Development of a web-based service (REST API) for recognition of legal citations to EU legal instruments in the EU legal documents. Although this web-service will be used within EuroCases, free access to it has been granted also to UNITO and Nomotika for further NLP tasks, in the context of new industrial and research plans the three MIREL partners will possibly carried out together (see subsection 3.2).

- Development of a tool for automated extraction and conversion in an XML-document with simplified structure of EU legislative acts to be integrated in EuroCases. The tool allows for a facilitated and quicker performance of the algorithms for detection of similarities between legal texts.

- Improvement and further development of a module for mining legal texts from the EuroCases database and identify EuroVoc terms therein and linking identified legal terms to EuroVoc ontology. This will contribute to the semantic annotation of texts which is a precondition for further mining tasks, such as identification of similar norms, similar documents (similar legislative instruments or court decisions) or conflicting provisions. With the view of improving the work of the future similarity algorithms to be developed, new experiments on IATE[20], the EU's inter-institutional terminology database, have been started. This database could especially useful because of its rich linguistic and terminological resources, and especially because of the wealth of synonyms it offers.

- Improvement of the already developed EuroCases algorithms for detection of similarities between legal texts. To this end, a POS-tagger and NER in the module has been implemented. The recognition of legal citations as an important part of NER contributed for identification of similar provisions in cases where the legal citation in the referring provision must be replaced by the text of the referred provision in order to reconstruct the real will of the legislator and to retrieve the actual meaning of the norm. Furthermore, it was decided to experiment with improving similarity algorithms by extraction and compilation in a thesaurus of the terms defined in EU legislative acts (the so called "legal definitions" or "legal concepts"). This will allow, for instance, to compare and find similarities between a provision containing a legal concept and a provision containing terms used within the definition of this concept. Another experimental solution emerged from the above-mentioned discussion was connected with the development of an algorithm for topic extraction from legal texts at the level of a provision.

Besides the work above, the secondments of APIS started new research activities (not originally planned in MIREL) and triggered discussions about potential future collaborations.

---

[20] http://iate.europa.eu

In particular, a new PhD student, Rohan Nanda, worked with APIS in the following tasks:

- Development of a tool for automated calculation of: a) precision and recall used for measuring the effectiveness of the algorithms developed by UNITO for finding similar provisions of legislative acts, and b) the optimum where precision and recall have reached maximum values [Nanda et al., 2016].

- Development of a web-based interface for visualisation in a table of comparison of similar provisions of normative acts, in particular for visualisation in a transposition tables of provisions of EU directives and the corresponding provisions of national implementing measures.

The activity and the results obtained in this research has been integrated in the H2020 project "euLinkedDataHub" that UNITO, APIS, and UL are submitting in April 2017 together with other partners (see below subsection 4.2) and they will constitute the core of the PhD thesis of Rohan Nanda, whose defense is expected by January 2019. The research activity will be continued and extended at least until then.

Finally, the programmers of APIS seconded at UNITO visited the company Nomotika, which is also based in Turin, and discussed with them potential collaboration between UNITO as a research institution, Nomotika as a spin-off of the Turin University specialised in compliance management in the legal field and APIS as an experienced legal publisher of national and EU legal contents. See next subsection for more details about potential future collaborations between the three MIREL partners.

3.2 Nomotika

Nomotika s.r.l is a spin-off of UNITO researching and developing cutting-edge ICT solutions for the daily work of legal practitioners. The main product of the company is the MenslegiS system[21], which is the commercial version of Eunomos [Boella et al, 2016], an advanced legal document management system, developed in the past years via a collaboration between UNITO and UL, that integrates NLP procedures and XML legal standards for the management of legal texts for regulatory compliance.

Serena Villata (INRIA) was seconded at Nomotika for two months in 2016 (secondment #121). In the past, INRIA worked on the topic of ontology-based information extraction from licensing information applying machine-learning techniques, together with the MIREL partners Data61, UNIBO, and University of Cordoba [Cabrio et al., 2014], [Cardellino et al.,

---

[21] http://www.nomotika.it/EN/MensLegis/Flyer

2014], [Cardellino et al., 2015]. The results of her work have been exploited to define the two tools called NLL2RDF[22] and Licentia[23].

During secondment #121, Serena Villata studied together with the Nomotika personnel how to generalize the approach proposed in NLL2RDF and Licentia in such a way that this kind of processing is applicable to legal texts in general, and not only to licenses, i.e. whether it can be integrated within the MenslegiS system as well.

> In light of the interaction and the discussion stemmed from this study, Nomotika decided to reshaped the MenslegiS system, i.e. to reimplement it via the Semantic Web technologies mastered by Serena Villata, in order to favour the integration in the system of legal ontologies and modern machine learning techniques to populate them. Also, the documents currently included in MenslegiS database are annotated in the NormeInRete standard[24], an obsolete XML markup language, which is suitable to handle Italian legislation only. In the new shape of MenslegiS, the documents will be converted into the Akoma Ntoso XML standard, possibly via a commercial collaboration with the MIREL partner APIS, whose EuroCases system (cf. subsection 3.1.1) includes a database of EU case law and legislation tagged in Akoma Ntoso and built in the context of the past EU project EUCases[25], to which the MIREL partners APIS, UNITO, Nomotika, and UNIBO participated.
>
> Specifically, the new Semantic Web technology that will be used in the reshaping of MenslegiS is Semantic Media Wiki (SMW)[26], a free, open-source extension of software that powers Wikipedia. SMW allows to store and query data within the wiki's pages. All data created within SMW can easily be published via the Semantic Web, thus facilitating the integration with other systems. Furthermore, many spin-off extensions are available for SMW, which can allow the building a powerful and flexible knowledge management system.

In light of this, Serena Villata moved 2 months of secondment #119 from Data61 to Nomotika, in order to give a tutorial about SMW at Nomotika and follow the activities connected with the reimplementation of MenslegiS. The major modification has been approved by the PO and the tutorial will take place the 15th and 16th of May 2016. Several researchers working at UNITO and UL, as well as programmers and legal experts of Nomotika, have been invited to the tutorial.

---

[22] http://www.airpedia.org/nll2rdf/

[23] http://licentia.inria.fr

[24] http://www3.cirsfid.unibo.it/didattica/upload/46_NormeInRete.pdf

[25] http://www.eucases.eu/start.html

[26] https://www.semantic-mediawiki.org

3.3 DLVSystem

DLVSystem s.r.l. is a spin-off company of University of Calabria. The main product of the company is the DLV system, which is still acknowledged as one of the most powerful Answer Set Programming Systems and the most widely used in commercial applications. DLVSystem protect intellectual property of our solutions and prototypes by the means of scientific publications, and reserve the right of licensing all its implementations and commercial solutions.

Nowadays, DLVSystem is closely working with the MIREL partner HUD in particular, with whom it is seeking to develop big data capabilities for Answer Set Programming (ASP) on legal texts, and to study more elaborate mechanisms for preference handling. These capabilities will be integrated in the DLV system. During the first year of the project, DLVSystem hosted two researchers from HUD while, in summer 2017, researchers and developers from DLVSystem will be seconded in HUD. Other secondments, also with respect to other MIREL partners, are planned for the forthcoming years.

DLVSystem and researchers from HUD worked on mass parallelization of ASP and on mapping normative reasoning to ASP.

Concerning mass parallelization of ASP, HUD and DLVSystem have exchanged essential related publications that will enable further developments, and have agreed on an ongoing collaboration beyond the MIREL project. A major achievement coming from these secondments was gaining the insights, coming from experts on ASP, which allowed for a better understanding of the topic and the trends within the ASP community. Future work involves the development of an ASP grounder, which will be based on MapReduce in order to allow the instantiation of larger ASP programs.

Concerning mapping normative reasoning to ASP, the joint work focused on the exploitability of ASP as a tool for performing normative reasoning. It has been investigated to what extent ASP provides a suitable language for encoding the knowledge required for performing normative reasoning. At this stage, i.e. after the performed preliminary analysis on mapping, ASP is believed to be able to encode such knowledge.

> Within the MIREL project, DLVSystem aims at widening the application of ASP. In particular, it will support research activities related to reasoning about legal domains, providing specific technologies based on a long-time expertise in the design, development and engineering of advanced logic-based Artificial Intelligence and knowledge representation and reasoning solutions. The idea is to increase the knowledge in the specific domain, in order to take advantage from new commercial opportunities and intensify profitability; indeed, the Knowledge Management market is currently enjoying an increasing trend. Both private companies and public administrations are interested and in need for using applications and technologies based on the research activities carried out within the MIREL project. Therefore, scientific and technological advances auspicated in

this project allow development of new powerful, well-tailored knowledge management tools, and innovative applications on top of them. The twofold objective of DLV is:

1) Increasing the sales of commercial licenses;

2) Increasing of the sales of professional services and consultancies.

In the past, DLVSystem did not establish a specific kind of collaboration with other MIREL partners. Nevertheless, DLVSystem found that several subjects among the MIREL partners are very well suited for starting a profitable cooperation in order to reach the goals already expressed above. Indeed, DLVSystem believes that the project is allowing to start things up, and that the goals are likely to be fulfilled beyond the duration of the EU funding – DLVSystem is therefore interested in continuing the collaborations afterwards.

During visits by HUD researchers, DLVSystem got new knowledge on reasoning with big data. They intend to use this new knowledge for grant applications to local and national bodies, seeking to apply ASP to big data, possibly in collaboration with other MIREL partners, specifically UL, UNITO, Data61 and UNIBO (cf. subsection 2.7).

More generally, DLVSystem aims to cooperate with MIREL's partners in a long-term perspective in order to empower the know-how needed for effectively implementing a number of "market-ready" solution in the specific domain of the project. Currently, collaborations have been already established with industrial stakeholder to reserve the possibility to assess any future prospective development.

4 Joint research project proposals

4.1 DAPRECO

"DAPRECO: DAta Protection REgulation COmpliance", is a research project that has been retained for funding under the FNR/CORE schema[27] in October 2016.

DAPRECO aims at building a knowledge base representing in a machine-readable format the provisions in the upcoming General Data Protection Regulation[28] (GDPR) and some ISO standard concerning data protection. The idea is to correlate the provisions in the standard with those in the GDPR in order to measure the overlapping between the two. ISO standard are not part of the law but they can be an argument in a court in order to avoid liability. Correlations are however defeasible, in that they are subject to legal interpretation. The DAPRECO knowledge base will be able to keep track of the several interpretations via defeasible mechanisms.

---

[27] https://www.fnr.lu/funding-instruments/core/

[28] http://ec.europa.eu/justice/data-protection/reform/files/regulation_oj_en.pdf

The underlying formalism that has been chosen for the DAPRECO knowledge base is reified Input/Ouput logic [Robaldo and Sun, 2017], briefly described in subsection 2.1.2. And, during DAPRECO formal legal ontologies will be developed to represent the concepts at stake and connected with the items used in the logic, in order to provide both a high-level and a low-level conceptual organization.

As described in subsection 2.1.2, reified Input/Output integrates recent research in Deontic Logic with recent research in Natural Language Semantics. Specifically, it integrates in Input/Output logic, a well-known deontic formalism developed at the University of Luxembourg, with reification, a well-known technique in natural language semantics, which have been investigated in WP1 and WP2 of MIREL.

Furthermore, the partner UNIBO has been involved in DAPRECO as external collaborator: it will contribute with its long-standing expertise in legal ontology building in order to build the ontology for the GDRP. For this reason, in the summer 2017, researchers from UL and UNIBO will be seconded in Stanford University in order to evaluate the GDPR ontology together with the researchers Codex[29], the Stanford center for legal informatics.

Finally, it is worth noticing that DAPRECO has been resubmitted in 2016, in that in 2015 it has been rejected. One of the main critics received from FNR reviewers in 2015 was that the previous underlying logical formalism was deemed to not be appropriate enough for representing legal knowledge. By substituting it with reified Input/Output logic and by involving UNIBO as external partner, the project has been evaluated more positively and it has been retained for funding from FNR in 2016.

4.2 euLinkedDataHub

The project euLinkedDataHub is a new project proposal involving the MIREL partners APIS, UNITO, and UL. It has been submitted to the H2020 ICT-14 Big Data PPP call: "Big Data PPP: cross-sectorial and cross-lingual data integration and experimentation"[30].

The project will develop a cloud-based open data platform for regulatory compliance and risk management, to serve as a Big Data hub for integration and cross-linking of legal, administrative, company, public procurement, geographical, mass and social media data from a variety of public and private sources. Company data as well as legislation and case law, news, Twitter comments, etc. will be analysed and cross-linked in order to find dependencies and building services for the market. The module for National Implementation Measures developed at UNITO (cf. subsection 2.1.2) as well as the module for Named Entity Recognition in legal texts (cf. subsection 2.1.3) will be integrated in the platform.

The project has been proposed by APIS and UNITO. It involves several partners across Europe, both academic and industrial. UL will be a partner of the project; Specifically, a new

---

[29] https://law.stanford.edu/codex-the-stanford-center-for-legal-informatics
[30] http://ec.europa.eu/research/participants/portal/desktop/en/opportunities/h2020/topics/ict-14-2016-2017.html

professor on Big Data hired at the University of Luxembourg, Martin Theobald[31], has been involved in the project thanks to the network of MIREL. He will provide the Big Data infrastructure to (1) build an integrated and multilingual platform for the large-scale storage, retrieval and analysis of European legal-regulations and compliances texts, including structured and unstructured data with company information, public procurement notices, news and other sources relevant to risk management, (2) link with static data collections (e.g., legal texts, related publications and news articles) and semantic representations (Linked-Open-Data: Yago, DBpedia, Gazetteers/Geonames, etc.), (3) perform real-time linking and integration with news feeds (e.g., RSS) and other streaming data sources (e.g., Twitter).

4.3 Media-EveryWhere

INRIA submitted the project Media-EveryWhere to the Call ICT-20 – Tools for smart digital content in the creative industries[32]. The project aims to develop technologies to decentralise the production, publication and description of content in the creative industry, in order to redistribute power back to all stakeholders along the value chain.

Media-Everywhere relies on the generation of a vast amount of metadata about creative works that is decentralised and distributed among stakeholders across the value chain, both in a raw form or inside a Distributed Ledger. A critical aspect for earning the trust of creative workers to use the digital platform and the applications developed on top of it, is to ensure that data is contributed/edited by the rightful owners or licensers and that access is granted only to the agents (both human or machine) designated by the creator or licenser of the work being described.

Decentralization and distribution enables more resilient, secure, and trustworthy applications, but it also poses several technical challenges on the management of digital rights. The project offers models and algorithms for identification, authentication, rights verification, and security & trust management in smart contracts. It takes also into account distributed data management features, together with the human component and interfaces for privacy, accountability, and informed consent. It also defines the core vocabularies needed to define the theoretical and technical concepts.

Authorization, authentication, and semantic-based access control are part of the proposal, and the related WP is lead by INRIA. The link between Semantic Linked Data and the blockchain is another important part of the project where INRIA will contribute. In this project, INRIA builds upon the knowledge acquired thanks to the MIREL project, in particular regarding the collaboration with the MIREL partner Data61.

---

[31] http://wwwen.uni.lu/recherche/fstc/computer_science_and_communications_research_unit/members/martin_theobald
[32] https://ec.europa.eu/research/participants/portal/desktop/en/opportunities/h2020/topics/ict-20-2017.html

4.4 ADaPT

ADaPT is a project submitted to the DG JUSTICE CALL "Support training activities on the data protection reform" [33]. The project is led by UNIBO and it involves UL and other universities that has been found through the MIREL network, such as the university of Wien and the university of Hannover, which have also been involved in the LAST-JD RiOE proposal (cf. subsection 4.1.8 below).

The project ADaPT will provide training activities and a hand-on set of tools addressing (a) general transversal issues as well as (b) different areas which will be subject to changes after the adoption of the General Data Protection Reform. The general transversal issues will be the following: (1) data protection authorities' role and responsibilities, (2) administrative and judicial remedies available in case of violation of data protection rules, (3) the impact of the CJEU and the ECtHR case law on the rights to privacy and data protection, and their interplay with other rights (for instance the connection between the right to be forgotten to freedom of expression and the right to freedom of economic initiative), (4) the regulation of data flows (including the new EU-US Privacy Shield Agreement) and the issues pertaining to the extra or ultra-territorial application of data-protection law, (5) issues pertaining to the extra-territorial application of the EU GDPR (in particular, the application of EU law to Extra European commercial operators offering goods and services in the EU), (6) the jurisdiction and liability of Internet intermediaries (including search engines), in particular, the interface between Data Protection Law and E-Commerce Law. (7) the connection between data protection on the one hand and security and law enforcement on the other hand (the use of Big Data for prevention of crimes and offences; the implication of the data protection in the context of the raising phenomena of cyber-bullying, cyber-harassment, identity theft, etc.).

The selected areas will include medical data, data related to migrants and asylum seekers (following the reform of the Dublin III regulation, data shared on social networks (with particular attention to right to be forgotten) and intermediaries liabilities, big data and algorithmic decision-making, protection of children data, and data protection in the domain of robotics and artificial intelligence (from self-driving cars to health-care robots).

4.5 LegaMaRR

LegaMaRR (Machine Reading and Reasoning over Legal Texts) is a project submitted to the ANR/FNR call[34] which involve UL and Université Paris Nord (prof. Aldo Gangemi). The project has been conceived during the MIREL "Third Workshop on Legal Knowledge and the Semantic Web" organized by UL and UNIBO. Prof. Aldo Gangemi was the invited speaker of the workshop. The project involves technologies developed at Université Paris

---

[33] http://ec.europa.eu/research/participants/portal/desktop/en/opportunities/rec/topics/rec-rdat-trai-ag-2016.html

[34] https://www.fnr.lu/international-cooperation/anr/

Nord (specifically, the FRED system[35] for Natural Language Processing) and reified I/O logic, illustrated in subsection 2.1.2. The main idea of the project has been conceived from C. Santos, who was a PhD student from Joint International Doctoral (Ph.D.) Degree in Law, Science and Technology (LAST-JD), whose principal investigator is UNIBO and that involves the MIREL partners UL, UNITO, Data61, and Nomotika.

The project proposes to develop an automated text analytics platform (Legal Machine Reading and Reasoning, LegaMaRR) to identify potential privacy and data protection pitfalls in the privacy policies declared by companies on the Web. In particular, policies of companies are analyzed via FRED and transformed in first-order logic representations, which are then converted into reified I/O logic with respect to a legal ontology. Reified I/O logic enables reasoning with respect to the ontology, which allows to identify pitfalls in privacy policies similar to the ones stored in the ontology.

The final objective of the project is to deploy the result obtained from FRED and reified I/O logic in a monitoring functionality to detect front-runner documents from the concerned stakeholders. Pitfalls in privacy policies create an asymmetry of knowledge between the parties (companies and data subjects). Mainly in relation to the digital economy, data subjects are inadequately informed about and thus unaware of infringements of personal data protection rules, which stem from hidden and vague terms.

4.6 PRODA

PRODA is a research project that has been submitted for funding under the FNR/CORE schema[36] in April 2017. The project has been proposed by UL, and it has been written in strict collaboration with the MIREL partners Stanford and UNITO.
The aim of the PRODA (PROtection of DAta) project is to design an automated reasoner for legislative texts. The chosen case study is the EU General Data Protection Regulation (GDPR). PRODA built on the project DAPRECO, described in subsection 4.1.

Innovatively, the PRODA project turns a given knowledge base representation of the GDPR into an automated reasoning system, to be paired with existing Legal technology (Legaltech) applications. The reasoner supports norm compliance checking and consistency checking.

This project has two parts. The first part is devoted to the design of a reasoner, using reified Input/Output logic, described above in section 2.2. The second part is devoted to its automation via an embedding into Higher-Order Logic (HOL) logic and Benzmuller's LEO-III theorem-prover for HOL[37]. Prof. Christoph Benzmuller visited UL at the beginning of 2017 and he has been involved in MIREL activities, e.g., he was an invited speaker of the workshop "International Workshop on Logic-Based Formalisms for Legal Reasoning"[38],

---

[35] http://wit.istc.cnr.it/stlab-tools/fred

[36] https://www.fnr.lu/funding-instruments/core/

[37] http://page.mi.fu-berlin.de/lex/leo3/

[38] https://sites.google.com/view/lbflr2017/home

organized by the partner DLVSystem in the context of MIREL, in order to interconnect WP2 (mining) with WP3 (reasoning) of the project.

Within PRODA it is planned to exploit some of the advantages offered by a new modal translation of Input/Output logic, whose discovery grew out of a MIREL visit to Stanford University and to Nuances Commmunication, San Francisco (secondment #59).

4.7 NoRSeW

"NORSEW: NOrmative Reasoning in the SEmantic Web" is a research project that has been submitted for funding under the FNR/CORE schema[39] in April 2017. The project has been proposed by UL, and involves also the MIREL partners INRIA and UCT.

One of the main open problems in Legal Informatics is the automatic (or semi-automatic) management of the normative information contained in the legal texts: the extraction and formal organization of normative rules, combined with tools for reasoning about them in order to perform specific tasks, as compliance checking.

The project NORSEW aims at investigating such an aspect of Legal Informatics, focusing primarily on ontology modelling and normative reasoning: through the appropriate modification of instruments that are accepted standards in the Semantic Web area, the OWL framework with its editors and reasoners, the goal of the project will be the development of software tools aimed at representing norms and reasoning about them. We will pair such instruments with NLP tools designed to extract deontic rules from legal texts, specifically the NLP tools for extracting rules developed by the MIREL partners INRIA and Data61, together with a joint research collaboration with the University of Trento, described above in subsection 2.7.

In such a way the project will delineate a general approach that, starting from norm-mining, passes through the formalization of the normative content of a legal text in dedicated ontologies, and ends up with reasoning tools that can query such ontologies.

4.8 LAST-JD RIoE

The MIREL partners UNIBO, UNITO, UL, APIS, Nomotika, Data61 and others found through the MIREL network prepared a Joint Doctocal proposal "LAST-JD: Law, Science and Technology Joint Doctorate - Rights of the Internet of Everything (RIoE)". The proposal has been submitted to the call H2020-MSCA-ITN-2017 (Innovative Training Network)[40].

---

[39] https://www.fnr.lu/funding-instruments/core/

[40] https://ec.europa.eu/research/participants/portal/desktop/en/opportunities/h2020/calls/h2020-msca-itn-2017.html#c,topics=callIdentifier/t/H2020-MSCA-ITN-2017/1/1/1/default-group&callStatus/t/Forthcoming/1/1/0/default-group&callStatus/t/Open/1/1/0/default-group&callStatus/t/Closed/1/1/0/default-group&+identifier/desc

LAST-JD (RIoE) will define an international joint doctorate educational training (15 fellowships) to create a common platform of knowledge and language for early stage researchers working in Law, Science and Technology and specializing in Internet of Everything (IoE) and to establish structures for long-term cooperation, strengthening relationships among the leading universities, research institutes and enterprises and to continuously develop the research training platform that European industries rely on, thus avoiding the gap in Europe with respect to similar programs in Stanford, Harvard, etc.

The 15 PhD students will be trained in a joint academic/industrial program with a common training plan, cutting-edge training-by-research, high quality supervision, complementary and transferable skills training, summer schools, etc., in order to pursue an innovative research project that will tackle a timely and important scientific problem with an interdisciplinary approach (Computer Science, Law and Ethics) and to transfer expertise/know-how among the partners of the Consortium and with external groups via industrial internships, networking activities, inter-sectorial exposure, secondments, workshops, sharing of learning material, public engagement and outreach activities.

Doctorate projects will be jointly supervised within a joint governance structure in which the beneficiaries as well as the partner organizations will be actively involved. The doctorate program will expose PhDs to different sectors and they will acquire a comprehensive set of transferable skills. The projects are designed to enhance collaboration and interaction across these disciplines, to integrate non-academic partner organizations, and to develop interdisciplinary methodologies for building, studying and regulating IoE. It will create a generation of young researchers trained in the above topics. They will end up with the knowledge and entrepreneurial capacity to create innovative companies or work in research, industry or public administration.

5 Conclusions

This deliverable has presented the activity of MIREL during the first sixteen months of the project with respect to the exploitation of the results as well as the future plans for exploitation.

In the first section, titled "Exploitation of research results", the main research achievements for which it could be identified some kind of exploitation have been listed. Since most partners in the MIREL are universities or research centers, many of these results has been used to design new research projects, listed in section 4 "Joint research project proposals".

One of the most important research results listed in section 2 is the XML legal standard LegalRuleML, which has become an OASIS standard in March 2017 after a long joint work of the MIREL partners UNIBO and Data61. LegalRuleML will become a worldwide used XML standard for tagging rules on legal documents, i.e. to associate norms in natural language with machine-readable representations referring to their meaning, that may be used

in applications for legal informatics, in order to perform reasoning on legal text. In other words, LegalRuleML represents an important interchange representation format for linking the two main component in MIREL: "mining" and "reasoning", thus it represents a basic component to achieve the main objective of the project.

For this reason, LegalRuleML will be used, together with the other XML legal standard Akoma Ntoso, aiming at tagging the structure of legal documents and the relevant information occurring therein, in most future project in legal informatics, e.g. DAPRECO (described in subsection 4.1) and euDataHub (described in subsection 4.2).

Further advancements have been achieved also in Natural Language Semantics applied to the legal domain (reified I/O logic, described in subsection 2.2), Natural Language Processing applied to the legal domain (specifically: Text Similarity, described in subsection 2.3, Named Entity Recognition, described in subsection 2.4, Ontology learning, alignment, and restructuring, described in subsections 2.5 and 2.6, Rule Extraction, described in subsection 2.7), Conceptual and theoretical issues addressed in WP1 (cf. subsection 2.8), and Reasoning technique applied to the legal domain (WP3), specifically the research on large-scale normative reasoning described in subsection 2.10. On parallel, a legal market study analysis has been conducted (cf. subsection 2.9), important for WP4 and for having a better understanding of the market scenario where the industrial partners of MIREL are moving.

In section 3 of this deliverable, we illustrated industrial exploitation carried out by the three MIREL partners APIS, Nomotika, and DLVSystem.

APIS already completed all its secondments in MIREL; five programmers of the company have been seconded to UNITO to learn more about cutting-edge techniques for Named Entity Recognition and Ontology population (EuroVoc). They integrated the results of this activity within the EuroCases system, patented and distributed by APIS. Furthermore, they discussed with UNITO and Nomotika possible future projects and research/industrial activities to be carried out together, possibly involving other MIREL partners.

Nomotika SRL, during the past secondment of Serena Villata (secondment #121), took the decision to re-implement the basic architecture of the MenslegiS system, the main product of the company. The technology at the basis of MenslegiS has been deemed to be obsolete. The company decided to rebuilt it according to the Semantic Media Wiki technology, which is mastered by Serena Villata, and transfer all current knowledge base of the MenslegiS system in Semantic Media Wiki. Serena Villata will carry out a tutorial about Semantic Media Wiki in May 2017 in the context of secondment #119.

Finally, DLVSystem extended the main product of the company, the DLV system, with the reasoning techniques learnt from the other MIREL partners, especially HUD, during the secondments hosted at the company. For the future, the company plans to extend the DLV system further towards the handling of Big Data, a research topic that MIREL partners want to seriously take into account for the future research activities. The future collaborations will possibly involve other MIREL partners such as UNIBO and UNITO and will be devoted to the definition of new project proposal submissions.

In section 4 of the deliverables, the joint project proposals by MIREL partners have been illustrated. One of them (DAPRECO, subsection 4.1) has been already retained for funding, while all the others are currently under evaluation.

The projects euDataHub (subsection 4.2), Media-EveryWhere (subsection 4.3), and LAST-JD RIoE (subsection 4.8) are H2020 proposals involving several other partners besides those in MIREL, many of which has been found through the network of MIREL. ADaPT (subsection 4.4) and LegaMaRR (subsection 4.5) also involve links created by the network of MIREL, the first from a link created from UNIBO to professor Pascal Bouvry in UL, the latter between UL and prof. Gangemi, working at Paris-3, who was an invited speaker of the "Third Workshop on Legal Knowledge and the Semantic Web", organized as an event of MIREL. Finally, PRODA (subsection 4.6) and NorSeW (subsection 4.7) have been submitted to the FNR/CORE call, the same call of the project DAPRECO. As DAPRECO, they involve MIREL partners as external collaborators and use technologies developed in the project via a collaboration between the proponents and other MIREL partners.

References

[1]   [Adebayo et al, 2016] Adebayo, J., Di Caro, L., Boella, G. A Supervised KeyPhrase Extraction System. Proceedings of the 12th International Conference on Semantic Systems. 57-62, ACM.

[2]   [Ajani et al, 2017] G. Ajani, G. Boella, L. Di Caro, L. Robaldo, L. Humphreys, S. Praduroux, P. Rossi, A. Violato: The European legal taxonomy syllabus: A multi-lingual, multi-level ontology framework to untangle the web of European legal terminology , Applied Ontology, to appear.

[3]   [Athan et al, 2015] Tara Athan, Guido Governatori, Monica Palmirani, Adrian Paschke, Adam Z. Wyner: LegalRuleML: Design Principles and Foundations. Reasoning Web 2015.

[4]   [Boella et al, 2016] G. Boella, L. Di Caro, L. Humphreys, L. Robaldo, P. Rossi, L. van der Torre: Eunomos, a legal document and knowledge management system for the Web to provide relevant, reliable and up-to-date information on the law, Artificial Intelligence and Law, Vol. 24, Issue 3.

[5]   [Boella et al, 2016b] G Boella, L Di Caro, A Ruggeri, L Robaldo. Learning from syntax generalizations for automatic semantic annotation. Journal of Intelligent Information Systems 43 (2), 231-246.

[6]   [Boella et al, 2016c] Boella, G., Di Caro, L., Robaldo, L. Semantic relation extraction from legislative text using generalized syntactic dependencies and support vector machines. International Workshop on Rules and Rule Markup Languages for the Semantic Web.

[7]   [Cabrio et al., 2014] Cabrio E., Palmero A., Villata S. (2014). These Are Your Rights - A Natural Language Processing Approach to Automated RDF Licenses Generation. In Proceedings of the 11th European Conference on Semantic Web (ESWC 2014), Lecture

Notes in Computer Science, vol. 8465, p. 255-269. Springer.

[8] [Cardellino et al., 2014] Cardellino C., Villata S., Gandon F., Governatori G., Lam H., Rotolo N. (2014). Licentia: a Tool for Supporting Users in Data Licensing on the Web of Data. In Proceedings of the 13th International Semantic Web Conference, vol. 1272, p. 277-280.

[9] [Cardellino et al., 2015] Cardellino C., Villata S., Alonso Alemany L., Cabrio E. (2015). Information extraction with Active Learning: A Case Study in Legal Text. In Proceedings of the 16th International Conference On Computational Linguistics and Intelligent Text Processing, Lecture Notes in Computer Science, vol. 9042, p. 483-494. Springer.

[10] [Cardellino et al, 2017] Cardellino, C., Teruel, M., Alonso Alemany, L., Villata, S. Learning Slowly To Learn Better: Curriculum Learning for Legal Ontology Population. Proc. of the 30th Florida Artificial Intelligence Research Society (FLAIRS), 2017.

[11] [Cardellino et al, 2017a] Cardellino, C., Teruel, M., Alonso Alemany, L., Villata, S. Legal NERC with ontologies,Wikipedia and curriculum learning, In Proc. of European Chapter of the Association for Computational Linguistics (EACL), 2017.

[12] [Cardellino et al, 2017b] Cardellino, C., Teruel, M., Alonso Alemany, L., Villata, S. A low-cost, high-coverage Legal Named Entity Recognizer, Classifier and Linker. In Proc. of the 16th International Conference on Artificial Intelligence and Law (ICAIL), 2017.

[13] [Curran et al. 2007] Curran, J.R., Clark, S., Bos, J.: Linguistically motivated large-scale NLP with c&c and boxer. In Carroll, J.A., van den Bosch, A., Zaenen, A., eds.: ACL 2007, Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics, June 23-30, 2007, Prague, The Association for Computational Linguistics (2007)

[14] [Davidson, 1967] Davidson, D. 1967. The logical form of action sentences. In Nicholas Rescher, editor, The Logic of Decision and Action. University of Pittsburgh Press.

[15] [De Paiva et al., 2016] Legal Tech Start-ups: State of the Art and Trends. In Proceedings of the Workshop on 'MIning and REasoning with Legal texts' collocated at the 29th International Conference on Legal Knowledge and Information Systems., 2016.

[16] [Di Caro and Boella, 2015] Di Caro, L., Boella, G. Semantic Similarity Reasoning. FETLT 2015: 127-138.

[17] [Di Caro and Boella, 2016] Di Caro, L., Boella, G. Automatic Enrichment of WordNet with Common-Sense Knowledge. LREC 2016.

[18] [Di Caro and Siragusa 2016] Laying bare EuroVoc by means of Latent Dirichlet Allocation. In Proceedings of the Workshop on 'MIning and REasoning with Legal texts' collocated at the 29th International Conference on Legal Knowledge and Information Systems., 2016.

[19] [Di Caro et al., 2008] Di Caro, L., Candan, S., Sapino, M.L. Using tagflake for condensing navigable tag hierarchies from tag clouds. Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining. 1069-1072, ACM.

[20] [Dumais, 2004] Dumais, S. Latent semantic analysis. Annual review of information science and technology, 38(1):188–230, 2004.

[21] [Governatori et al., 2016] Guido Governatori, Francesco Olivieri, Erica Calardo, Antonino Rotolo, Matteo Cristani), Sequence Semantics for Normative Agents. In Principles and

Practice of Multi-Agent Systems conference agent-based modeling programming, 2016.

[22] [Hobbs, 2008] Hobbs, J. R. 2008. Deep lexical semantics. In Proc. of the 9th International Conference on Intelligent Text Processing and Computational Linguistics (CICLing-2008), Haifa, Israel.

[23] [Makinson and van der Torre, 2000] Makinson, David and Leendert W. N. van der Torre. 2000. Input/output logics. Journal of Philosophical Logic, 29(4): 383-408.

[24] [Nanda et al., 2016] Nanda, R. and Di Caro, L. and Boella, G. A Text Similarity Approach for Automated Transposition Detection of European Union Directives, In Proc. of the 29th International Conference on Legal Knowledge and Information Systems (JURIX2016), 2016.

[25] [Robaldo et al, 2012] L. Robaldo, L. Lesmo, D. Radicioni: Compiling Regular Expressions to Extract Legal Modifications, In proc. of 25th International Conference on Legal Knowledge and Information Systems (JURIX2012). Amsterdam, 2012.

[26] [Robaldo and Sun, 2017] L. Robaldo and X., Sun: Reified Input/Output logic: Combining Input/Output logic and Reification to represent norms coming from existing legislation, The Journal of Logic and Computation, to appear.

[27] [Rotolo and Calardo, 2016] Antonino Rotolo Erica Calardo, Quantification in Some Non-normal Modal Logics, In Journal of Philosophical Logic, online first, 2016.

[28] [Teruel et al, 2017] Teruel, M., Cardellino, C., Alonso Alemany, L., Villata, S. Learning Slowly To Learn Better: Curriculum Learning for Legal Ontology Population. Proc. of the 30th Florida Artificial Intelligence Research Society (FLAIRS), 2017.

[29] [Vreeswijk, 1997] Vreeswijk, G. Abstract argumentation systems, Artificial Intelligence, 90 (1), 225 - 279, 1997.