

LREC 2012 Workshop on

Semantic Processing of Legal Texts (SPLeT-2012)

CALL FOR PAPERS

27 May 2012, Istanbul

Workshop description

The legal domain represents a primary candidate for web-based information distribution, exchange and management, as testified by the numerous e-government, e-justice and e-democracy initiatives worldwide. The last few years have seen a growing body of research and practice in the field of Artificial Intelligence and Law which addresses a range of topics: automated legal reasoning and argumentation, semantic and cross-language legal information retrieval, document classification, legal drafting, legal knowledge discovery and extraction, as well as the construction of legal ontologies and their application to the law domain. In this context, it is of paramount importance to use Natural Language Processing techniques and tools that automate and facilitate the process of knowledge extraction from legal texts.

Since 2008, the SPLeT workshops have been a venue where researchers from the Computational Linguistics and Artificial Intelligence and Law communities meet, exchange information, compare perspectives, and share experiences and concerns on the topic of legal knowledge extraction and management, with particular emphasis on the semantic processing of legal texts. Within the Artificial Intelligence and Law community, there have also been a number of dedicated workshops and tutorials specifically focussing on different aspects of semantic processing of legal texts at conferences such as JURIX-2008, ICAIL-2009, ICAIL-2011, as well as in the International Summer School “Managing Legal Resources in the Semantic Web” (2007, 2008, 2009, 2010, 2011).

To continue this momentum and to advance research, a 4th Workshop on “Semantic Processing of Legal Texts” is being organized at the LREC-2012 conference to bring to the attention of the broader LR/HLT (Language Resources/Human Language Technology) community the specific technical challenges posed by the semantic processing of legal texts and also share with the community the motivations and objectives which make it of interest to researchers in legal informatics. The outcome of these interactions are expected to advance research and applications and foster interdisciplinary collaboration within the legal domain.

New to this edition of the workshop are two sub-events to provide common and consistent task definitions, datasets, and evaluation for legal-IE systems along with a forum for the presentation of varying but focused efforts on their development.

The first sub-event will be a shared task specifically focusing on dependency parsing of legal texts: although this is not a domain-specific task, it is a task which creates the prerequisites for advanced IE applications operating on legal texts, which can benefit from reliable preprocessing tools. For this year our aim is to create the prerequisites for more advanced domain-specific tasks (e.g. event extraction) to be organized in future SPLeT editions. We strongly believe that this could be a way to

attract the attention of the LR/HLT community to the specific challenges posed by the analysis of this type of texts and to have a clearer idea of the current state of the art. The languages dealt with will be Italian and English. A specific Call for Participation for the shared task is available in a [dedicated page](#).

The second sub-event will be an online, manual, collaborative, semantic annotation exercise, the results of which will be presented and discussed at the workshop. The goals of the exercise are: (1) to gain insight on and work towards the creation of a gold standard corpus of legal documents in a cohesive domain; and (2) to test the feasibility of the exercise and to get feedback on its annotation structure and workflow. The corpus to be annotated will be a selection of documents drawn from EU and US legislation, regulation, and case law in a particular domain (e.g. consumer or environmental protection). For this exercise, the language will be English. A specific Call for Participation for this annotation exercise is available in a [dedicated page](#).

The main goals of the workshop and associated events are to provide an overview of the state-of-the-art in legal knowledge extraction and management, to explore new research and development directions and emerging trends, and to exchange information regarding legal language resources and human language technologies and their applications.

Areas of Interest

The workshop will focus on the topics of the automatic extraction of information from legal texts and the structural organisation of the extracted knowledge. Particular emphasis will be given to the crucial role of language resources and human language technologies.

Papers are invited on, but not limited to, the following topics:

- Construction, extension, merging, customization of legal language resources: terminologies, ontologies
- Information retrieval and extraction from legal texts
- Semantic annotation of legal textual corpora
- Legal text processing
- Multilingual aspects of legal text semantic processing
- Legal thesauri mapping
- Automatic Classification of legal documents
- Logical analysis of legal language
- Automated parsing and translation of natural language arguments into a logical formalism
- Linguistically-oriented XML mark up of legal arguments
- Dialogue protocols for argumentation
- Legal argument ontology
- Computational theories of argumentation that are suitable to natural language
- Controlled language systems for law.

Submissions

Submissions are solicited from researchers working on all aspects of semantic processing of legal texts. Authors are invited to submit papers describing original completed work, work in progress, interesting problems, case studies or research trends related to one or more of the topics of interest listed above. The final version of the accepted papers will be published in the Workshop Proceedings.

Short or full papers can be submitted. Short papers are expected to present new ideas or new visions that may influence the direction of future research, yet they may be less mature than full papers. While an exhaustive evaluation of the proposed ideas is not necessary, insight and in-depth understanding of the issues is expected. Full papers should be more well developed and evaluated. Short papers will be reviewed the same way as full papers by the Program Committee and will be published in the Workshop Proceedings.

Full paper submissions should not exceed 10 pages, short papers 6 pages; both should be typeset using a font size of 11 points. Style files will be made available by LREC for the camera-ready versions of accepted papers. Papers should be submitted electronically, no later than February 10, 2012. The only accepted format for submitted papers is Adobe PDF. Submission will be electronic using START paper submission software available at <https://www.softconf.com/lrec2012/SPLeT2012/>.

Note that when submitting a paper through the START page, authors will be asked to provide essential information about resources (in a broad sense, i.e. also technologies, standards, evaluation kits, etc.) that have been used for the work described in the paper or are a new result of your research. For further information on this new initiative, please refer to <http://www.lrec-conf.org/lrec2012/?LRE-Map-2012>.

Selected contributions to a Special Issue of AI&Law Journal

After the Workshop a number of selected, revised, peer-reviewed articles will be published in a Special Issue on Semantic Processing of Legal Texts of the *AI and Law Journal* (Springer).

Important Dates

Paper submission deadline: 10 February 2012

Acceptance notification sent: 5 March 2012

Final version deadline: 23 March 2012

Workshop date: 27 May 2012

Workshop Chairs

- Enrico Francesconi (Istituto di Teoria e Tecniche dell'Informazione Giuridica of CNR, Florence, Italy)
- Simonetta Montemagni (Istituto di Linguistica Computazionale of CNR, Pisa, Italy)
- Wim Peters (Natural Language Processing Research Group, University of Sheffield, UK)
- Adam Wyner (Department of Computer Science, University of Liverpool, UK)

Address any queries regarding the workshop to: lrec_legalWS@ilc.cnr.it

Program Committee (tbc)

Kevin Ashley (Univ of Pittsburgh)

Johan Bos (University of Rome, Italy)

Danièle Bourcier (Humboldt Universität, Berlin, Germany)

Thomas R. Bruce (Cornell Law School, Ithaca, NY, USA)

Pompeu Casanovas (Institut de Dret i Tecnologia, UAB, Barcelona, Spain)
Jack Conrad (Thomson-Reuters)
Matthias Grabmair (University of Pittsburgh, USA)
Carole Hafner (Northeastern Univ.)
Antonio Lazari (Scuola Superiore S.Anna, Pisa, Italy)
Leonardo Lesmo (Dipartimento di Informatica, Università di Torino, Torino, Italy)
Carl Malamud (Public.Resource.Org)
Marie-Francine Moens (Katholieke Universiteit Leuven, Belgium)
Thorne McCarty (Rutgers Univ.)
Raquel Mochales Palau (Catholic University of Leuven, Belgium)
Paulo Quaresma (Universidade de Évora, Portugal)
Robert Richards (Legal Informatics blog)
Tony Russell-Rose (Endeca)
Erich Schweighofer (Universität Wien, Rechtswissenschaftliche Fakultät, Wien, Austria)
Rolf Schwitter (Macquarie Univ)
Manfred Stede (University of Potsdam, Germany)
Daniela Tiscornia (Istituto di Teoria e Tecniche dell'Informazione Giuridica of CNR, Florence, Italy)
Tom van Engers (Leibniz Center for Law, University of Amsterdam, Netherlands)
Giulia Venturi (Scuola Superiore S.Anna, Pisa, Italy)
Vern R. Walker (Hofstra University School of Law, Hofstra University, USA)
Radboud Winkels (Leibniz Center for Law, University of Amsterdam, Netherlands)