
Legal NLP Introduction

Adeline Nazarenko* — **Adam Wyner****

* *LIPN, Université Paris 13 – Sorbonne Paris Cité & CNRS*

** *Swansea University, School of Law and Department of Computer Science*

ABSTRACT. Language and law have always had a close relationship, the latter being primarily a “discourse”. Currently, the automatic processing of the legal language is a major issue because of the growing footprint of the law on the web and its complexity in contemporary globalised societies. In addition, through the prism of a specialised language, here legal language, we can measure the progress of natural language processing (NLP). The goal is to integrate different processes into operational applications that meet specific needs. This goal is particularly challenging and important in the case of legal language due to the intertwined levels of linguistic analysis, from the analysis of character strings (for the identification of citations, for example) to argumentation. This issue of TAL aims to draw attention to the issues and challenges of legal NLP, to present recent research in this field, and, more broadly, to show how different methods of analysis are organised for this specialised language.

RÉSUMÉ. La langue et le droit entretiennent depuis toujours des liens étroits, ce dernier étant d’abord un « discours », mais le traitement automatique de la langue juridique représente aujourd’hui un enjeu majeur du fait de l’empreinte croissante du droit sur le web, de son ouverture et de sa complexification dans les sociétés contemporaines mondialisées. Par ailleurs, le prisme d’une langue de spécialité, ici la langue juridique, permet de mesurer les progrès du traitement automatique des langues. Il s’agit d’intégrer différents traitements dans des applications opérationnelles qui répondent à des besoins spécifiques, et ce défi de l’intégration est particulièrement important dans le cas de la langue juridique du fait de l’intrication des paliers d’analyse linguistique, depuis l’analyse des chaînes de caractères (pour le repérage des citations, par exemple) jusqu’à celle de l’argumentation. Ce numéro de TAL vise à attirer l’attention sur les défis et les enjeux du traitement automatique de la langue juridique, à montrer l’intérêt des recherches récentes dans ce domaine, mais aussi, plus largement, à montrer comment différentes méthodes d’analyse s’organisent pour une langue de spécialité.

KEYWORDS: Law, LegalTech applications, processing of languages for special purposes.

MOTS-CLÉS : droit, applications juridiques, traitement des langues de spécialité.

1. Introduction

The aim of this issue of the *TAL journal*, entitled “Automatic Processing of Legal Language”, is to contribute to the analysis and exploitation of legal information using Natural Language Processing (NLP), considering theoretical problems and specific linguistic phenomena as well as how the analyses are integrated or applied to processing legal sources. It highlights work currently being done in the field, stimulates new research strands, and shows more broadly how different methods of analysis can be organised for processing special language.

Language and Law have always been closely linked, the latter being fundamentally a “discourse”. Legal language is to be understood here in a broad sense: written and oral language, legal texts and judgments, as well as regulatory texts such as decrees, regulations, contracts or requirements. As Language and Law are inseparable, it is essential to develop methods of Legal NLP in order to understand legal language and discourse, to develop tools supporting the exploitation of legal sources for law enforcement, as well as to enable transparent, international, and interoperable legal systems over the web.

This introduction presents a selection of challenges that legal language processing faces today; as we are not offering a complete review of the field, there are naturally other research challenges and approaches which could be discussed¹. Nonetheless, our brief survey shows that NLP is a key issue in the development of LegalTech (Section 2), that legal language raises specific difficulties in terms of NLP (Section 3), and that integration remains a major problem when NLP technologies are to be used in domain-specific but also in real-world applications (Section 4). Section 5 also introduces the papers that compose this special issue and illustrate various facets of the challenges that current research in NLP is addressing.

2. LegalTech – context and opportunities

Broadly, LegalTech refers to technologies from Computer Science that are applied to a range of areas related to legal practice and materials. In order to set the context and opportunities, we scope the discussion. LegalTech has a wide range of application areas to help law firms and organisations with daily activities related to document support (creation, revision, storage and retrieval), legal proceedings (providing electronic documents in the course of litigation and government investigations, legal research of (non-)legal sources to support decision-making) and more generally all aspects related to the dematerialisation of legal services from text and paper to digital form. Some of these areas are document and language centered (document storage and retrieval, electronic discovery, legal research, and document automation/assembly) and highly relevant to Legal NLP.

1. See recent editions of the JURIX and ICAIL conferences as well as issues of the *Journal of Artificial Intelligence and Law*.

Another well-developed area of research and development has been Forensic Linguistics, which applies linguistic analysis to legal materials in proceedings (Osslon, 2009; Coulthard and Johnson, 2010). Amongst the many topics in Forensic Linguistics, we can find linguistic analyses of discourses by participants in police investigations and interviews, courtroom exchanges, authorship identification along with associated opinions, security analysis, and translation of documents in multilingual legal contexts.² To a greater or lesser extent, research in Forensic Linguistics has applied NLP technologies, though this is not essential to the endeavours, which can often be carried out by manually annotating or marking up text.

Language processing has, then, been central to address tasks and purposes in areas of LegalTech and Forensic Linguistics. To the extent that these tasks and purposes have solutions and are already commercialised, we can say they are applications of existing technologies to well defined and scoped textual issues. We find large, well-established legal information service providers such as Thomson Reuters and Oracle, law firms such as Pinset Masons and Riverview Law, as well as a host of startups touching on a full spectrum of issues.^{3, 4} However, for the NLP research community, the aim is to take up opportunities in identifying and addressing challenging textual issues. We mention some of them.

NLP technologies were first used for assisting the drafting of legal documents. One common approach to automated support relies on a decision-tree model of drafting, where a document template (e.g. a contract) is automatically refined and instantiated according to the drafter's local decisions (Sprowl, 1980; Gordon, 1989). Such approaches use quite basic NLP technology to provide contracts. However, NLP now has significant opportunities in the analysis of legal documents, which have become available in very large scale, for example, enabling the mining of contractual relationships across a corpus of documents, e.g. global oil and gas concessions.

Traditionally, law schools, legal offices, and legislative counsels have produced guidelines to explain what should be the internal structure and constituents of legal documents (e.g. resolutions, executive orders, contracts, regulations), how to express rules and decisions unambiguously using precise legal terminology, and how to handle cross-referencing between sources, etc. In this context, NLP is naturally used to support drafting by controlling the structure of the documents, the length of the sentences, and the use of recommended terms (Höfler, 2012).

A more complex issue is related to the control of legality and consistency of legal sources. As anyone knows, legal documents can be very long, they are frequently updated, and they are part of large legal systems which are subject to interpretation and evolve with various social and political factors. Legal actors must ensure a consistent and up-to-date use of terminology, they must control the compatibility and the consistent evolution of rules that come from different jurisdictions, and they must check

2. See <http://www.iafl.org/> for relevant conferences and journals.

3. <https://angel.co/legal-tech-1>

4. <https://www.legalgeek.co/startup-map/>

the legality of the sources and decisions with respect to constitutional norm. Beyond surface analysis, these controls involve a deep understanding of the legal texts and logical reasoning. The development of standards for encoding the structure of documents (e.g. *LegalDocML*⁵) and the semantic content of rules [e.g. *LegalRuleML* (Athan *et al.*, 2015)] is a prerequisite for the construction of tools to facilitate control.

Retrieving information is also a major challenge, considering the huge bulk of laws and decisions that are produced over the years in modern societies. They cover economic, social and political issues on the local scale as well as worldwide. For example, citizens want to know which rules apply to their district when they want to restore their houses; employers must know the applicable labor legislation; and trade is governed by international treaties and agreements. One must be able to retrieve the relevant texts and to extract the specific legal and statutory rules that are relevant to a given case. Most countries have an official website for the legislation⁶, regulations, and legal information. However, ensuring the publication, interoperability, and accessibility of these resources calls for advanced semantic and search technologies. There are needs for richer metadata (e.g. date, jurisdiction, legal matter, keywords, etc.) but also fine-grained search or navigation tools. For instance, one should be able to directly find the decisions related to a given topic that derive from the transposition of a given European directive (Mimouni *et al.*, 2014).

Related to information retrieval is linking legal resources. In their daily work, legal professionals, such as barristers, judges, prosecutors, legal advisers, analyse the law and the cases, search for precedent cases, identify relevant legislative or regulatory documents, and consider jurisprudence. In other words, a range of resources must be leveraged to gain a “wholistic” view of the matter at hand. Thus, there must be a *legal semantic web*, where sources are accessible, interconnected, annotated such that legal professionals can query, explore, and possibly enrich corpora with new cases and documents with new interpretations (Casanovas *et al.*, 2016).

Information retrieval and linking serve to support decision-making, reasoning, and compliance. Taking into consideration contracts, legislation, and regulations, legal professionals ought to be able to determine whether a given action is legally compliant or secure, what legal determination follows from given input information, or what one’s liabilities are. Without such capabilities, business, government, and individual activities can be unclear or problematic. For example, one cannot manage multimedia content without mastering the various rights attached to the elements, where the rights are often encoded in distinct contracts. One would want to automatically compare contracts. To ensure legal certainty, one needs tools to analyse the legal documents, extract and formalise the rules attached to each type of content, reason over the set of rules that apply to a specific business process, and test that the action abides by compliance or security protocols.

5. See https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=legaldocml, which subsumes *Akoma Ntoso* <http://www.akomantoso.org/>

6. e.g. <https://www.legifrance.gouv.fr/> or <http://www.legislation.gov.uk/>

In addition to reasoning with respect to a given body of legal documents, disputes about the interpretation or application of the law arise. Argumentation is often at the heart of legal reasoning and judgment: one wants to validate the arguments and claims put forward as well as to suggest new pro and con arguments. Supporting argumentation calls for NLP technologies to mine, collect, and interrelate arguments in legal sources, to populate knowledge representation of argumentation models, and to reason over the models (Wyner *et al.*, 2010). The ultimate goal is to support legal professionals in designing, understanding and controlling arguments and more generally to assist legal reasoning.

The various topic areas above (drafting, managing, retrieving, linking, reasoning, compliance, and so on) apply as much across jurisdictions as within. Thus, there are also needs for tools to support comparative law, which is crucial for working with the law in international and global settings. Here again, we need to understand the texts in depth, to extract and model the rules they contain, then to reason on those rules to check their consistency and identify flaws or redundancies. Moreover, these operations must be applied on a large scale and with respect to different legal systems, where terminology and concepts must be correlated or aligned between jurisdictions.

In sum, drafting, publishing, querying, linking, and reasoning over legal sources in ways that can be exploited by legal practitioners as well as by citizens and governments raise significant challenges that relate to NLP in connection with document engineering, information retrieval, knowledge representation and reasoning, as well as more generally decision support.

3. A challenging field for NLP

NLP is itself a mature sub-area of Artificial Intelligence with a range of well-known rule-based or machine-learning techniques along with tools designed to split sentences, tokenise texts, lemmatise words, tag words with part-of-speech, parse sentences, enrich text with semantic roles, recognise named entities, extract relations between entities, identify discourse markers, perform anaphoric reference, classify texts or textual passages, and so on. For certain classes of text, e.g. newspaper and narrative text, such techniques and tools are often highly successful.

However, such techniques and tools are usually developed on and for specific textual types. Yet, what works for newspaper or narrative texts does not succeed for corpora of legal language. The challenges are not so much the volume of texts, but 1) the technicalities needed to account for particular, complex structures and patterns and 2) the specificity of performance requirements. We discuss these below.

3.1. *Complexity of legal sources*

Legal sources are complex in many ways. We highlight some of them.

The language itself is complex. To start with, the lexicon is rich. It includes uniquely legal expressions (e.g. *charged* or *defendant*) some of which may have internal structure (e.g. *without prejudice to any claim*), which are mostly standardised or codified in dictionaries of legal language. In addition, there are legal sub-areas, e.g. *family*, *criminal*, and *contract* law, each of which codify different terminologies. Both the legally-specific lexical items within and across legal sub-areas are important for the processing of legal documents. Moreover, in at least the common law context, criminal law, tort law, civil responsibilities, and to some extent legislation and regulations embed ordinary language or descriptions, which must also be analysed. As legal concepts bear on everyday occurrences, such as parking restrictions or temporal relations in a murder, legal and non-legal expressions must correctly be tied together. Considering the variation in time, jurisdictions, and languages, the lexical complexities proliferate.

In terms of syntax, sentences are generally well constructed (as opposed to “noisy” texts published on social media, for instance) but often long and highly structured, with embedded or intertwined clauses and sometimes ambiguities (e.g. prepositional clause attachment, scope of an adverb). The presentational format may break a long sentence into parts that are distributed in lists, e.g. “A British citizen is a person: (a) born in the UK; (b) born to a British citizen; ...”, which themselves may contain sublists. Specific NLP modules must be developed for analysing the various form of lists that can be found in corpora. In practice, it can be observed that the current parsers, e.g. the Stanford Parser, which have been trained on news corpora, can fail to parse long sentences and to resolve ambiguities of legal texts.

This complexity is also found at the semantic level. Anaphora resolution is problematic given the length and complexity of sentences. Logical translations, while largely feasible for short sentences of non-legal language, are infeasible with legal texts: it is difficult to determine the scope of the logical operators (negation, conjunction, disjunction, and conditional) with respect to noun phrases and verb phrases; similarly, the scope of quantifiers can be problematic. These issues are not specific to legal language, but their resolution is of the utmost importance in the law field, because they are prominent in legal sources and because legal reasoning is based on logical analysis and the interpretation of those sources.

Stylistics also matters. There are few standards that hold across international jurisdictions for how legislation or regulations are expressed other than those standards specifically set by international organisations. With respect to common law contexts, case reports are largely up to the presentational preferences and styles of the judges and clerks who write up the report. NLP tools must confront the absence of widespread, consistent, and homogeneous linguistic expression.

Beyond the language of individual texts, the corpora themselves are complex because they are very structured and closely interconnected. An article of law often cannot be interpreted in isolation. It must be considered in the light of the law of which it is a part and of all the texts which are attached to it, through a term definition, by references and citations, or by semantic relations. For instance, when the enforcement of a given law t_2 , which transposes the European directive t_1 , is suspended upon the adoption of a decree t_3 that specifies its conditions of application, none of the three texts can be taken independently of the others. In case law, cases are semantically linked in relations of *upholds*, *overrules*, and similar concepts (so-called Shephardisation). Finally, legal texts are not immutable, but can be wholly or partially updated, leading to portions of the text referring to updated or overwritten law. Several versions of the same law article usually coexist, each one being valid for a specific time frame.

In the various ways outlined above, legal texts are complex, rich, and diverse, making them challenges to current NLP tools.

3.2. Performance requirements for legal applications

Another challenge for processing legal texts is related to the performance requirements of the legal system. Such requirements may vary from one application or use-case to another, and they may be especially high in legal NLP.

In legal case-based reasoning, for instance, it is important to carry out research to identify relevant precedent cases or applicable legislation. Here, legal information service providers have large teams of legal experts to analyse and index legal texts or to provide summaries of the contents of legal texts (also known as headnotes), providing meta-data that can be used in searches. Nonetheless, legal professionals widely accept the results of querying such resources, which can then be subjected to further manual filtering. Were open NLP techniques to be applied to the source texts, the results ought to at least mirror those provided by human experts. The requirements on getting the correct textual resources may be very high, given that a legal argument is only as well supported and defended as the volume and accuracy of the material used. In principle, failing to adequately defend against even a single precedent case or to take a piece of regulation into account could be fatal.

In the domain of the translation of legal texts, given that the language of the law is paramount, the quality of the translation ought to be not only very high, but ought to be validated in some systematic and transparent manner. After all, an obscure phrase in either the source or translated text could have legal ramifications should the phrase be the lynchpin of the dispute.

Even where precision and recall were hypothetically perfect, the results may not satisfy the requirements of the legal problem. Suppose we have a very large corpus of death penalty decisions split between verdicts of innocence or guilt. Suppose a machine learning classifier has outstanding performance in classifying the cases. From this, one might suppose that given all the information required as input to a particular

novel case (on a par with that provided to cases in the case base), the system classifies the verdict, *e.g.* as guilty. Should the defendant then be accepted as guilty? It is highly unlikely that the legal system much less the defendant would accept this as a definitive result. The law has long functioned by providing a full explanation to justify the decision. In part, this is to ensure that the application of the law is based on precedent and existing legislation in a clear, transparent, and systematic fashion; and in part, such decisions can be reused in a variety of ways, *e.g.* linking cases, setting precedent, and as the basis for legal appeals. Without a fine-grained representation of the internal structure of the decision, explanation to the requisite degree would be unavailable. Given the current technology, machine learning does not provide such *explanations* or structured information. This highlights that the NLP techniques ought to serve the purposes of the law and legal setting as they are.

It is not just about finding documents, making a decision, and explaining it, but often users need contextual elements, syntheses, and methodological guides. Consider a citizen accessing some online legal advisory facility using question-answering. She may want not only a document relevant to addressing a legal issue, but auxiliary information as well as guidance on how to work through towards resolution of the matter. Given the complexities of legal sources noted above, it would be unrealistic to simply link the citizen to the source material. Nor can it reasonably be expected that, from the textual sources, the citizen can understand how to navigate the process. The requirements are that the citizen needs some exactly relevant (*i.e.* high precision and recall) and clear digest of the legal materials (*e.g.* a summarisation) as well as a structured path through the legal procedures.

In legal applications, users expect very high performance in terms of results, explanations of the results, guidance on the law relevant to the results, all of which ought to be provided by easy to use tools.

4. The challenge of integration

Focusing on a special language, we can measure the progress of natural language processing. Here we focus on the language of the law, though similar issues arise for any language for specific purposes. In addition, there must be a strong focus on the integration across linguistic analyses, from the character string level (to identify for instance the citations and the list structures), to the variety of standard NLP tasks (parsing, anaphora, ambiguity resolution, etc.), to the multilingual, and to knowledge levels that support argumentation and reasoning over various legal systems. In the most advanced applications, all these various levels of analysis must be addressed and combined.

When one is interested in processing a specialised language, one necessarily has a transversal vision, which covers all levels of analysis. This is particularly true in the case of legal language, which raises great questions of understanding and reasoning:

– document engineering is a requirement. Processing of legal sources presupposes document normalisation and a standard for encoding those documents. This issue has been identified for a long time. Important standard proposals have been made for encoding the structure of legal documents, *e.g.* Akoma Ntoso⁷ (Casanovas *et al.*, 2016) or LegalDocML (Athanasopoulos *et al.*, 2015);

– syntactic analysis is a challenge due to the complexity and precision of legal language but also to the presence of ambiguities. Both statistical and logic-based approaches have been tested with mixed results (Wyner and Peters, 2011; Dragoni *et al.*, 2016);

– terminology is an issue in any specialised language but legal terms and idioms (*e.g.* *établissement d'utilité publique*), which are often complex and difficult to understand for the lay man, have a strong semantics; they are often the keystone of legal reasoning; In legal texts, legal terminology is often mixed with the vocabulary of the field (domain terminology) covered by the law, which can itself be complex. For instance, a legal text about cyberlaw will discuss how legal concepts, *e.g.* obligations or rights, apply to specific domain terminology for aspects of computers and communication technologies, *e.g.* Wi-Fi or passwords, that are neither legally defined, nor legal concepts. This is an additional difficulty for processing legal sources (Bonin *et al.*, 2010);

– stylistics is important, if we consider that drafting guidelines aims to reduce ambiguities and ease the reading of legal documents;

– semantic analysis – be it shallow or deep – is at the heart of legal content management, from information retrieval to consistency checking, and of legal reasoning. Various subtasks, such as semantic annotation (Francesconi, 2016) or rule extraction (Dragoni *et al.*, 2016), have been addressed, but developing a generic legal semantic parser for legal sources remains an open research issue⁸;

– discourse analysis helps to organise and contextualise legal contents. In particular, the analysis of document networks has attracted a lot of attention (Winkels and de Ruyter, 2011; Boulet *et al.*, 2011; Christensen *et al.*, 2016);

– argumentation being the basis of legal reasoning, it is essential to rely on NLP to help produce, check and mine arguments. The analysis of legal arguments has been a topic for some time (Moens *et al.*, 2007);

– beyond NLP, knowledge engineering is also required to design semantic resources or ontologies (Sartor *et al.*, 2013) which can be used to ground semantic analysis, to formalise and apply the legal rules that cannot be exploited in their natural language form, and to enable temporal reasoning in a field where various document timelines are intertwined, related, for instance, to publication, promulgation, in force.

To integrate different analyses into operational applications that meet specific needs, trade-offs must be found, perhaps even dynamically, between the depth, re-

7. <http://www.akomantoso.org/>

8. <https://www.matthes.in.tum.de/pages/74hcgqw5dmwj/Semantic-Analysis-of-Legal-Texts-SERIT>.

liability, coverage of analysis, and volume of text to process. Unfortunately, integration is considered as a technical rather than as a scientific issue. Various options have been proposed but the constraints are seldom explicit, so the trade offs are rarely documented. Thus it is not known if the proposed solutions are optimal.

Another integration issue is related to the management of ambiguities and errors. The various levels of document and linguistic analysis are interdependent: on the one hand, the ambiguities that appear at one level can be solved at another level, but on the other hand, errors can also be propagated from one level to another, thus impacting the quality of the overall analysis. Therefore, the interactions between the various levels of analysis must be controlled.

Various architectures and workflows have been proposed to tackle some of the NLP integration issues⁹. While, they have not been designed for legal source processing *per se* but legal source analysis is an interesting playground to test and compare their relative strengths and weaknesses, due to its complexity and broad analysis spectrum.

However, integration issues go well beyond text analysis. Applications involve domain knowledge regarding the entities at stake (ontologies) and the reasoning rules. In real applications, NLP is only one of the technologies to be implemented together with semantic technologies, logic and knowledge engineering, data and decision sciences. Developing applications for the various actors involved in legal businesses (legal professionals and their clients, citizens, etc.) is a huge project ahead of us.

5. Presentation of this TAL issue

The present issue has two papers that presents two different aspects of research related to legal NLP.

The paper by Jaromir Savelka, Vern R. Walker, Matthias Grabmair and Kevin D. Ashley, entitled “Sentence Boundary Detection in Adjudicatory Decisions in the United States” addresses the specific problem of segmenting legal texts into sentences. It proposes an in-depth analysis of a specific and supposedly simple NLP task, sentence boundary detection, that is a prelude to many more complex ones. This paper illustrates the complexity of legal language and its impact on the quality of the analyses that can be done and on the applications that rely on them.

The second paper focuses on an application. It shows how beneficial a question/answering system on maritime regulations can be for the commander of a boat who needs to know all the regulations that are relevant to a specific type of ship, at a particular time, and in a given space, or for the supervisory authorities in charge of identifying infringements and risk situations on the part of boat commanders. The paper entitled “Un système de question/réponse automatique dans le domaine légal :

9. See, for instance, the General Architecture for Text Engineering (GATE, <http://gate.ac.uk/>), the Natural Language Toolkit (NLTK, <http://www.nltk.org/>) or the Apache Unstructured Information Management Architecture (UIMA, <http://uima.apache.org/>).

le cas des réglementations maritimes”, by Yannis Haralambous and Cheikh Kacfeh Emani, also shows that NLP, knowledge engineering and semantic web technologies must be combined to develop such an question/answering system.

6. References

- Athan T., Governatori G., Palmirani M., Paschke A., Wyner A. Z., “LegalRuleML: Design Principles and Foundations”, in W. Faber, A. Paschke (eds), *Reasoning Web*, Springer, p. 151-188, 2015.
- Bonin F., Dell’Orletta F., Venturi G., Montemagni S., “Singling out Legal Knowledge from World Knowledge. An NLP-based approach”, *Proceedings of LOAIT 2010.*, Fietole, Italy, 2010.
- Boulet R., Mazzega P., Bourcier D., “A Network Approach to the French System of Legal codes- Part I: Analysis of a Dense Network”, *Journal of Artificial Intelligence and Law*, vol. 19, p. 333-355, 2011.
- Casanovas P., Palmirani M., Peroni S., van Engers T. M., Vitali F., “Semantic Web for the Legal Domain: The next step”, *Semantic Web*, vol. 7, n° 3, p. 213-227, 2016.
- Christensen M. L., Olsen H. P., Tarissan F., “Identification of Case Content with Quantitative Network Analysis: An Example from the ECtHR”, in F. Bex, S. Villata (eds), *Legal Knowledge and Information Systems – JURIX 2016: The 29th Annual Conference*, vol. 294 of *Frontiers in Artificial Intelligence and Applications*, IOS Press, p. 53-62, 2016.
- Coulthard M., Johnson A. (eds), *The Routledge handbook of forensic linguistics*, Routledge, 2010.
- Dragoni M., Villata S., Rizzi W., Governatori G., “Combining NLP Approaches for Rule Extraction from Legal Documents”, *Proceedings of the Workshop on ‘Mining and REasoning with Legal texts’ collocated at the 29th International Conference on Legal Knowledge and Information Systems*, 2016.
- Francesconi E., “Semantic Model for Legal Resources: Annotation and Reasoning over Normative Provisions”, *Semantic Web Journal*, vol. 7, n° 3, p. 255-265, 2016.
- Gordon T. F., “A Theory Construction Approach To Legal Document Assembly”, In *Pre-Proceedings of the 3rd International Conference on Logic, Informatics, and Law*, p. 485-498, 1989.
- Höfler S., “Legislative Drafting Guidelines: How Different Are They from Controlled Language Rules for Technical Writing?”, in T. Kuhn, N. E. Fuchs (eds), *Proceedings of Controlled Natural Language CNL2012*, Springer, p. 138-151, 2012.
- Mimouni N., Nazarenko A., Salotti S., “An Approach for Searching and Browsing a Network of Legal Documents”, in S. F. Radbud Winkels, Nicola Lettieri (ed.), *Network Analysis in Law*, Edizioni Scientifiche Italiane, p. 183-207, 2014.
- Moens M.-F., Boiy E., Palau R. M., Reed C., “Automatic Detection of Arguments in Legal Texts”, *Proceedings of the 11th International Conference on Artificial Intelligence and Law*, ICAIL ’07, ACM, New York, NY, USA, p. 225-230, 2007.
- Osslon J., “Forensic linguistics”, in V. Muhvic-Dimanovski, L. Socanac (eds), *LINGUISTICS – Volume I*, Encyclopedia of life support systems, EOLSS Publishers/UNESCO, p. 378-393, 2009.

- Sartor G., Casanovas P., Biasiotti M., Fernandez-Barrera M., *Approaches to Legal Ontologies: Theories, Domains, Methodologies*, Springer Publishing Company, Incorporated, 2013.
- Sprowl J., “Automated Assembly of Legal Documents”, in B. Nesbitt (ed.), *Computer Science and Law – an advanced course*, Cambridge University Press, p. 195-203, 1980.
- Winkels R., de Ruyter J., “Survival of the Fittest: Network Analysis of Dutch Supreme Court Cases”, in M. Palmirani, U. Pagallo, P. Casanovas, G. Sartor (eds), *AI Approaches to the Complexity of Legal Systems (AICOL-III)*, Springer, p. 106-115, Aug., 2011.
- Wyner A., Mochales-Palau R., Moens M.-F., Milward D., “Approaches to Text Mining Arguments from Legal Cases”, in E. Francesconi, S. Montemagni, W. Peters, D. Tiscornia (eds), *Semantic Processing of Legal Texts*, Springer-Verlag, Berlin, Heidelberg, p. 60-79, 2010.
- Wyner A., Peters W., “On Rule Extraction from Regulations”, in K. Atkinson (ed.), *Legal Knowledge and Information Systems – JURIX 2011: The 24th Annual Conference*, IOS Press, p. 113-122, 2011.

Thanks to the members of the specific reading committee of this special issue:

- Kevin Ashley (University of Pittsburgh School of Law, USA)
- Danièle Bourcier (CERSA, CNRS & Université Panthéon-Sorbonne, France)
- Pompeu Casanova (UAB Law School, Universitat Autònoma de Barcelona, Spain)
- Jean-Pierre Chevallet (LIG, Université Grenoble Alpes & CNRS, France)
- Catherine Faron Zucker (I3S, Université de Nice & CNRS, France)
- Aldo Gangemi (LIPN, Université Paris 13 – Sorbonne Paris Cité & CNRS, France)
- Meritxell Fernández-Barrera (ELDA/ELRA, France)
- Thomas Lebarbé (Litt&Arts, Université Grenoble Alpes & CNRS, France)
- François Lévy (LIPN, Université Paris 13 – Sorbonne Paris Cité & CNRS, France)
- Aurélien Max (LIMSI, Université Paris Sud & CNRS, France)
- Marie-Francine Moens (Department of Computer Science, Katholieke Universiteit Leuven, Belgium)
- Isabelle Moulinier (Capital One, Sr. Director, Data Science, USA)
- Monica Palmirani (Computer Science and Law, Bologna University, Italy)
- Wim Peters (Sheffield Natural Language Processing group, UK)
- Patrick Saint-Dizier (IRIT, CNRS & Université Paul Sabatier, France)
- Serena Villata (I3S, CNRS & Nice University, France)
- Giulia Venturi (Institute of Computational Linguistics, CNR, Italy)
- Rabdoub Winkels (Computer Science and Law, University of Amsterdam, Netherlands)

and to the members of the editorial board of *TAL Journal* (see <http://www.atala.org/content/comité-de-rédaction-0>)