

# Automated Recommendation of Templates for Legal Requirements

Amin Sleimi\*, Marcello Ceci\*, Mehrdad Sabetzadeh<sup>†</sup>\*, Lionel C. Briand<sup>\*†</sup>, John Dann<sup>§</sup>

\*SnT Centre for Security, Reliability and Trust, University of Luxembourg, Luxembourg

<sup>†</sup>School of EECS, University of Ottawa, Canada

<sup>§</sup>Central Legislative Service (SCL), Ministry of State, Luxembourg

Email: {amin.sleimi, marcello.ceci, lionel.briand}@uni.lu, m.sabetzadeh@uottawa.ca, john.dann@scl.etat.lu

**Abstract**—[Context] In legal requirements elicitation, requirements analysts need to extract obligations from legal texts. However, legal texts often express obligations only indirectly, for example, by attributing a right to the counterpart. This phenomenon has already been described in the Requirements Engineering (RE) literature [1]. [Objectives] We investigate the use of requirements templates for the systematic elicitation of legal requirements. Our work is motivated by two observations: (1) The existing literature does not provide a harmonized view on the requirements templates that are useful for legal RE; (2) Despite the promising recent advancements in natural language processing (NLP), automated support for legal RE through the suggestion of requirements templates has not been achieved yet. Our objective is to take steps toward addressing these limitations. [Methods] We review and reconcile the legal requirement templates proposed in RE. Subsequently, we conduct a qualitative study to define NLP rules for template recommendation. [Results and Conclusions] Our contributions consist of (a) a harmonized list of requirements templates pertinent to legal RE, and (b) rules for the automatic recommendation of such templates. We evaluate our rules through a case study on 400 statements from two legal domains. The results indicate a recall and precision of 82,3% and 79,8%, respectively. We show that introducing some limited interaction with the analyst considerably improves accuracy. Specifically, our human-feedback strategy increases recall by 12% and precision by 10,8%, thus yielding an overall recall of 94,3% and overall precision of 90,6%.

**Index Terms**—Legal Requirements, Requirements Templates, AI-assisted RE, Natural Language Processing

## I. INTRODUCTION

The elicitation of requirements for IT systems in regulated domains such as labor and healthcare necessarily includes (a) the identification of the laws and regulations that are applicable to the domain in question, and (b) the extraction, by means of legal interpretation, of the legal requirements entailed by the applicable laws and regulations. Since requirements analysts typically do not have the legal expertise to handle these activities, they usually rely on the advice of legal professionals. This type of collaboration, if done without any automated assistance, is costly and time-consuming. Besides, the communication gap that exists between requirements analysts and legal professionals may result in missed legal requirements or legal requirements that are inaccurate or impossible to implement in IT systems. Providing automated support for directly extracting legal requirements from legal texts is an important step toward addressing these challenges.

Legal provisions often state criteria and rules that lead to legal requirements; however, an individual legal statement may

express more than one rule or criterion. In many cases, legal statements affect more than one stakeholder (addressee) and in different ways: attributing to a person an obligation that benefits a counterpart has the automatic effect of attributing to that counterpart a right. So, for example, an obligation for a bank in terms of confidentiality of financial information (bank’s viewpoint) entails a corresponding right for the customer of the bank for secure authentication (customer’s viewpoint).

Our investigation of multiple legal texts suggests that around one out of every six legal statements expresses multiple legal requirements, one for each applicable stakeholder’s viewpoint. However, from a linguistic point of view, a legal provision is normally drafted taking into account one and only one of such viewpoints, in order to avoid redundancy. For example, the obligation of a stakeholder (e.g., a bank as in the previous example) is often expressed only by attributing a right to their counterpart (e.g., the bank’s customers). The presence of multiple angles to a legal statement introduces a *viewpoint issue* [2] for legal requirements extraction, namely, the issue of detecting obligations (and legal requirements) even though they are only indirectly expressed by a legal statement.

Requirements analysts are interested in writing *good requirements*, i.e., requirements that are unambiguous, testable, clear, correct, understandable, feasible, independent, atomic, necessary, and implementation-free [3], [4], [5]. In good requirements, the required action is expressed in the active voice, and from the viewpoint of the addressee (i.e., the IT system or one of its stakeholders). For IT systems that operate in regulated domains, solving the viewpoint issue is necessary in order to write good legal requirements. This need has been corroborated by Breux et al. [1], who deem it necessary to “increase requirements coverage, since obligations derived from rights [...] may be operationalized as requirements.”

To help requirements analysts with this viewpoint issue, it is necessary to (a) identify the presence of multiple viewpoints in a legal statement, and (b) suggest a different legal requirement for each of these viewpoints. The best way to represent a plurality of requirements is by using templates [6], [7]. Requirements templates assist the requirements analyst in writing requirements that follow best practices in the RE community.

In a previous paper [8], we devised an approach for extracting statement- and phrase-level legal metadata at a linguistic level. In this paper, we utilize the extracted metadata for automatically recommending legal requirements templates,

thus assisting requirements analysts with legal requirements elicitation. We rely on the phrase-level metadata types *action*, *target* and *violation*, and the statement-level metadata types *obligation*, *prohibition*, *permission*, and *penalty*.

This paper is motivated by two observed limitations in the literature on legal requirements elicitation:

**1) Lack of a harmonized view of templates for legal requirements.** While the RE community acknowledges the importance of requirements templates and systematic legal requirements elicitation, there is no consensus on the templates for legal requirements. Different strands of work propose different templates, but none provide sufficiently complete coverage of legal requirements templates.

**2) Lack of automated support for the recommendation of templates for legal requirements.** As our previous research suggests [8], NLP techniques have considerably improved in recent years. This raises the prospect that modern NLP techniques may be accurate enough for automated requirements extraction from legal texts. However, to the best of our knowledge, a fully fledged application of NLP has not yet been attempted in legal requirements elicitation.

**Research questions.** Throughout the paper, we investigate three Research Questions (RQs). RQ1 tackles the first limitation above, while RQ2 and RQ3 tackle the second.

**RQ1: What are the adequate and sufficient templates for legal requirements?** RQ1 aims at developing a harmonized set of templates for legal requirements with a sufficient level of expressiveness. To this end, we review and reconcile several existing proposals of legal requirements templates. Our answer to RQ1 is the first contribution of the paper: a set of legal requirements templates.

**RQ2: Can one define template-recommendation rules over linguistic cues from legal texts?** RQ2 investigates the possibility to define rules for template recommendation that rely on linguistic cues from legal texts. We designate as linguistic cues the output of NLP technologies such as constituency parsing, dependency parsing and verb lexicons (e.g., VerbNet), as well as the semantic metadata extracted following our existing approach [8]. In order to define template recommendation rules, we conduct a qualitative study over 1000 randomly selected legal statements from the labor and health domains. Specifically, we annotate the statements with the appropriate templates from the ones identified in RQ1. We use the results of this study for defining rules for automatic template recommendation. The answer to RQ2 is the second contribution of the paper: a set of NLP-based rules for the recommendation of legal requirements templates identified in RQ1.

**RQ3: How accurate is our approach at recommending legal requirements templates?** RQ3 aims at evaluating the accuracy of our approach for template recommendation. Our evaluation is based on 400 legal statements randomly selected from both health and labor laws. Our empirical results suggest that our approach has a recall of 82,3% and precision of 79,8%. A follow-on analysis of the recommendation errors reveals that most of the errors can be easily avoided with

limited interactive guidance from the analyst. We show that by incorporating into our approach a lightweight human-feedback component, recall and precision increase by 12% and 10,8%, respectively, thus resulting in an overall recall of 94,3% and an overall precision of 90,6%.

**Overview and structure.** Section II reviews the related work. Section III answers RQ1 by describing our harmonization of existing legal requirements templates. Section IV presents our qualitative study and the recommendation rules resulting from it, thus answering RQ2. Section V answers RQ3 through a case study that evaluates the accuracy of our approach. Section VI discusses threats to validity. Section VII concludes the paper.

## II. BACKGROUND AND RELATED WORK

In this section, we review the relevant literature from RE, specially concerning requirements elicitation, and from AI and Law, specially concerning legal knowledge representation.

**Foundations from legal theory.** A systematic account of the relationship between legal positions was first investigated by J. Bentham [9] and further formalized by W. N. Hohfeld. The Hohfeldian system [10] distinguishes eight terms for legal positions: right, privilege, power, immunity, duty, no-right, liability, and disability. Each term in the Hohfeldian system is paired with one opposite and one correlative term. In this work we are interested in correlatives, i.e., legal positions that entail each other. For example, the right of a party entails a correlative duty for the counterparty: an employee has the right to obtain a copy of the payslip, which entails a correlative duty for the employer to provide the employee with such payslip.

**Balancing rights and obligations in RE.** The RE community has already highlighted the viewpoint issue in legal texts. Darke & Shanks [2] provide a conceptual framework to “increase requirements coverage by integrating viewpoints representing particular perspectives or set of perceptions of the problem domain”. Breaux et al. [1] “identify implied rights and obligations [...] to ensure requirements coverage and consider multiple viewpoints”. The authors show three ways to balance rights and obligations, dealing with delegations, direct provisions, and indirect provisions. Kiyavitskaya et al. [11] highlight how EU Directives may contain “two-level provisions that impose an obligation on member states and at the same time guarantee a right for an individual person”.

**Requirements Templates.** Considerable work has been devoted to structuring requirements through template suggestion. Palomares et al. [12] report on the use of patterns in RE in a comprehensive survey. First attempts include Robertson’s study [13] on “how event/use case modelling can be used to identify, define and access requirements patterns” and Dwyer et al.’s set of templates [14] for the specification of verifiable requirements through state machines. The latter involves manual mapping and transformation of requirements into logical expressions. More recently, Mavin et al. [15] present the Easy Approach to Requirements Syntax (EARS). EARS templates have a high-level perspective, and do not adequately account for actors and stakeholders other than the IT system.

Other contributions are specifically aimed at capturing legal requirements. Breaux & Gordon [7] present a list of generic templates to highlight information within legal provisions in the Legal Requirement Specification Language (LRSL). LRSL is aimed at encoding legal provisions for developers and policy makers. It accounts for conditions, actions, the syntactic structure of the legal provision, and the different stakeholders of the IT system. In the previously mentioned work, Breaux et al. [1] present a methodology for extracting rights and obligations from regulations using a semantic model. They define a list of patterns for such rights and obligations. Young & Anton [6] present a list of templates for translating provisions into legal requirements. These templates have IT systems as their main focus and take into account the different stakeholders' viewpoints. Yoshida et al. [16] update the templates proposed by Young & Anton by adding templates for definitions and processing data objects as first-class components. Although they present a method for automatically suggesting templates, the implementation has important limitations, the most notable being its exclusive focus on functional requirements, thus not accounting for non-functional and quality requirements.

Contributions from AI and Law focus on representing legal requirements with logical rules rather than templates. LegalRuleML [17] is a rule language that classifies statements into facts and norms, further specialized into constitutive, prescriptive, and penalty statements. LegalRuleML provides a solution to accurately express complex legal rules, but it is not supported by automatic extraction of concepts.

**NLP technologies.** As noted before, the potential of NLP technologies has increased with recent advancements. *Semantic Role Labeling* [18], [19] is the activity of assigning semantic roles to each of the predicate's arguments in a sentence. These roles usually capture the semantic commonality between instantiations of actors or artifacts across the language. The most notable contribution in the field is FrameNet [20], rooted in the theory of frame semantics. *Deep language analysis* [21] consists of using knowledge of linguistics to extract knowledge from text. It is a type of analysis that takes into account the nuances and complexities of linguistic constructs such as negation and conditionality. A *verb lexicon* is a lexical database of the different variations of syntactic representations of verbs in a sentence. VerbNet [22] is a verb lexicon that incorporates both semantic and syntactic information about verb types following Levin's classification of verbs [23].

**Use of Semantic Legal Metadata.** In a previous paper [8], we proposed a conceptual model of semantic legal metadata for RE. The proposed metadata types provide information about the statements and phrases contained in legal provisions. We further developed an approach to automatically extract their proposed metadata types using NLP techniques. We rely on [8] for automatically extracting from legal texts the metadata that form the cues for our recommendation rules. We do not elaborate further on our harmonized conceptual model and instead refer the reader to [8], where we also discuss the state of the art on legal requirements extraction [24].

### III. LEGAL REQUIREMENTS TEMPLATES

In this section, we present a synthesis of the different approaches to requirements templates outlined in the previous section, in order to devise a harmonized set of templates to express legal requirements from multiple viewpoints.

**Required features.** We begin by presenting the features that we need for our legal requirements templates:

- In order to represent multiple viewpoints [1], our templates will express pairs of corresponding statements, each formulated from the viewpoint of a different stakeholder. The first required feature for our templates is therefore **to be able to handle different stakeholders as subject**.
- Legal drafting practices often implicitly refer to a stakeholder by referring to the data objects they are related to. For example, the obligation "A person must write a report that contains [...]" is often expressed in the form "Report must contain [...]" This raises the need for our templates to **handle different data objects as first-class components**.
- Finally, we want to **present the templates in a textual form**, due to the ubiquitous and universal use of natural language in RE, especially in the elicitation and specification of legal requirements which typically involve different stakeholders with different expertise [25], [26].

Table I compares the approaches presented in the previous section against our required features. We note that LRSL supports different stakeholders, and so does Breaux et al.'s patterns. The latter also balances rights and obligations, which is paramount for handling multiple viewpoints. However, LRSL has a graphical representation and Breaux et al.'s patterns have an itemized representation. This does not fit with our required feature of templates being in textual format. Requirements that are not in a textual representation bring with them the need for additional training; for legal requirements, this would include training legal professional who may not be keen to use formal languages.

EARS proposes widely known and used textual templates for requirements. However, it is not suited for our objectives as it does not handle different stakeholders and data objects, depending on viewpoints. To cover these aspects, we adopt Young & Anton's and Yoshida et al.'s templates as a starting point and enhance them with multiple viewpoints.

**Statement types.** Having compared the approaches from the literature and identified the features of our templates, we proceed to define the type of rules that we want to represent. The reference model covers the Hohfeldian concepts that are relevant to RE, namely duty and right [27]. To do this, we focus on four statement types in our conceptual model [8] that are sources of legal requirements: obligation, prohibition, permission, and penalty.

Although *obligation* and *prohibition* are presented as distinct statement types in our reference conceptual model, we note that a prohibition is just a linguistic construct to express a negative obligation. Simply put, a prohibition requires that a specified action does not take place in the system. Since we are focusing on the semantic content of legal statements, we

TABLE I  
MAPPING OF APPROACHES TO REQUIREMENTS TEMPLATES

Related Work	Support for Different Stakeholders	Support for Different Data Objects	Textual Template	Balancing Viewpoints
Easy Approach To Requirements Syntax [15]	✗	✗	✓	✗
Yoshida et al.'s Functional Requirements Templates [16]	✗	✓	✓	✗
Young and Anton's Templates for Legal Requirements [6]	✗	✓	✓	✗
Legal Requirement Specification Language [7]	✓	✗	✗	✗
Breaux et al.'s patterns [1]	✓	✓	✗	✓

group those two statement types into a single one: **duty**. Duties are the main source of legal requirements, and the easiest to transform into requirements when expressed directly by a legal statement. Detecting indirectly expressed duties in statements with multiple viewpoints is the main focus of our study.

**Permissions** can express two types of legal rules, either a right or a power. *Rights* are a secondary source of legal requirements. From rights are derived obligations, that can subsequently be transformed into legal requirements. We note, however, that not all rights entail obligations. *Powers* attribute to one or more public servants a legal competence or duty. From the point of view of requirements elicitation, they have the same effect as rights, in that they often (but not always) entail an obligation for the liable stakeholders.

With regard to **penalties**, we note that the requirement engineer should extract duties from instances of the phrase-level metadata type *violation*. For example, in the sentence “Anyone inciting acts of hatred against a person is punished by an imprisonment of eight days to two years” the violation item is “inciting acts of hatred against a person” and the corresponding duty (prohibition) is “Individuals are forbidden from inciting acts of hatred against a person”.

The remaining statement types in the reference conceptual model, i.e. *facts* and *definitions*, are outside of the scope of this study. These statement types have constitutive effects and do not prescribe behaviors, as explained by Ceci et al. [28]. They can, however, interact with rules that express requirements and therefore affect those requirements. An example is the statement “Article 13 applies to public health workers”. Rules with such interactions are called *metarules* in formalizations such as LegalRuleML [29]. The present research does not deal with metarules, since they pose challenges that are far from our main focus here, i.e., detecting multiple viewpoints. Also, our reference conceptual model does not cover metarules extraction. This limitation implies that our requirements might be missing additional stakeholders and conditions that are introduced by the metarules. A possible approach to handle metarules is to follow Breaux [30], which uses state-event tables and transition tables to link “events generated from rights and obligations [...] to pre-conditions of other rights

and obligations”. Until a solution for the automatic handling of metarules is achieved, it is possible to circumvent this limitation by asking a legal expert to analyze the metarules and manually amend the affected requirements accordingly. Automatic identification of metarules involves detecting and analyzing cross-references [31].

In addition to the three statement types described above (duty, permission, penalty), we classify our templates into two categories depending on whether the action supports a *target*: **intransitive requirements** are those where the action does not support a target, and have the structure “Actor <modality> Action”; **transitive requirements** are those where the action supports a target, and have the structure “Actor1 <modality> Action to Actor2”. The concept of target of a legal requirement is defined by Young & Anton [6] as “the intended recipient of the actor’s action” and is a phrase-level concept in our reference conceptual model.

**Legal requirements templates.** Based on the above classification, we derive six legal requirement templates. We classify these templates into three categories:

(1) **Legal statements with no counterpart.** These statements express only one legal requirement, i.e., they carry a single viewpoint. The action of the requirement is directly expressed by the main verb, which does not have a beneficiary – hence the denomination of “legal statements with no counterpart”. The templates in this category translate into *intransitive requirements*. This category includes two templates, depending on the classification of the legal statement itself:

- a - *Duty with no counterpart*, e.g., “The bank must undergo a standardized accounting exercise each end of year.” The legal requirement is “Bank shall undergo a standardized accounting exercise each end of year.”
- b - *Permission with no counterpart*, e.g., “Personal property that was deposited at the time of bankruptcy may be claimed.” The legal requirement is “Depositor shall be capable of claiming personal property if it was deposited at the time of bankruptcy.”

(2) **Legal statements with correlative statements.** These statements express two correlative legal requirements. Of these two legal requirements, the main one is directly expressed by the verb, and the other is indirectly expressed as the correlative of the main one. An example is the sentence “The user shall obtain a copy of his personal data from the website.” This statement reads as a right for the user to obtain a copy of his personal data from the website, and as an obligation for the website to provide a copy of the personal data to the user. The general template for this category corresponds to *direct provision* in Breaux et al.’s work [1], and translates into two instances of *transitive requirement*.

This category includes two templates, depending on the classification of the main legal statement:

- c - *Duty with correlative permission*, e.g., “The creditors of the bankrupt are required to file at the district court the declaration of their claims.” The corresponding legal requirements are “Creditor shall file the declaration of

TABLE II  
EXCERPT OF LEGAL REQUIREMENTS TEMPLATES

Template Category	Proposed Template	Example
Statement with no counterpart	Actor <modality> Action	A covered entity shall document a restriction in accordance with §160.530(j) of this subchapter.
	The system <modality> Action	The system shall document a restriction in accordance with §160.530(j) of this subchapter.
Statement with correlative statement	Actor_1 <modality_1> Action_1 to Actor_2	The data subject shall be able to obtain confirmation as to whether or not personal data concerning him or her are being processed from the controller.
	Actor_2 <modality_2> Action_2 to Actor_1	The controller shall provide confirmation as to whether or not personal data concerning him or her are being processed to the data subject.
Statement with implied statement	Actor_1 <modality> Delegation_Action Actor_2	A consumer shall have the right to request to a business that they disclose to that consumer the sale of personal information.
	Actor_2 <modality_2> Action_2 to Actor_1 if R1	A business shall disclose the sale of personal information to the consumer if requested by the consumer.

claims at the district court” and “The district court shall be able to obtain the declaration of claims.”

- d - *Permission with correlative duty*, e.g., “The user shall obtain a copy of her personal data from the website.” The corresponding legal requirements are “The user shall be able to obtain a copy of her personal data from the website” and “The website shall provide a copy of the personal data to the user.”

(3) **Legal statements with implied statements.** These statements express two requirements: a legal requirement directly expressed by the sentence, and another implied legal requirement. An example is the sentence “The Minister delegates to the Police Administration the notification to the driver.” This statement reads as a power statement for the authoritative entity (i.e., the Minister) to delegate the notification; this means that each exercise of the power by the Minister implies an obligation for the Police Administration to perform the notification. The general template for this category corresponds to the template for *delegation* in Breaux et al.’s work [1], and translates into two templates: a *transitive requirement* for the requirement directly expressed by the text, and another that is either *transitive* or *intransitive* depending on the implied action. It is also important to notice that the implied requirement is pre-conditioned on the invocation of the original delegation: in the words of Breaux [30], “a stakeholder must first be delegated a right before they can invoke that right.” Considering that, in legal requirements elicitation, we focus on obligations, we can rephrase that into “a stakeholder must first be delegated an obligation before they are subject to that obligation.”

In this third category, we have two templates depending on the classification of the main legal statement:

- e - *Permission with implied duty*, e.g., “The bankrupt may have the circumstances reported by the district court.” The corresponding legal requirements are “The bankrupt shall be able to request that the district court report the circumstances” and “The district court shall report the circumstances if requested by the bankrupt.”
- f - *Penalty with implied duty*, e.g., “Anyone who incites acts of hatred against a person is punished by an imprisonment

of eight days to two years.” The corresponding legal requirements are “The court shall punish with an imprisonment of eight days to two years anyone who incites acts of hatred against a person” and “Individuals are forbidden from inciting acts of hatred against a person.”

An excerpt of our set of legal requirements templates is presented in Table II, and the complete set is available in an online appendix<sup>1</sup>. The set of templates presented in this section provides an answer to RQ1. Using these templates, it is possible to capture legal requirements expressed both directly and indirectly in legal statements, with the exclusion of metarules for which further research is necessary.

#### IV. RECOMMENDING TEMPLATES FOR LEGAL REQUIREMENTS

In this section, we report on a qualitative study aimed at deriving rules for the automatic recommendation of the legal requirements templates presented in the previous section.

##### A. Study context and data selection

Our qualitative study is based on 1000 statements randomly selected from the labor and health laws of Luxembourg (500 statements from each law).

The choice of the labor and health laws was motivated by three factors. First, due to these domains being widely known, legal experts found them to be good showcases for automated legal requirements recommendation. Second, the provisions in the labor and health laws are interesting from an RE perspective, due to their broad implications for the IT systems used by employers, courts and public offices such as the tax department, healthcare institutions and insurance companies. Third, our preliminary study on 200 statements from five different legal domains highlighted labor and health laws as the domains where the viewpoint issue is more common (about 20% of legal statements in the labor domain and 15% in health domain carry multiple viewpoints).

As it is the case with most legal texts, the source texts in our study contain statements with enumerations and lists embedded in them. To treat these statements properly, we took the common legal text pre-processing measure of merging the

<sup>1</sup><http://shorturl.at/hxzKL>



TABLE III  
RULES FOR REQUIREMENTS TEMPLATE RECOMMENDATION

Extraction Rule	Example
IF ("corresponding verb" == False AND statement type = type1 ) Then type1 with no counterpart	A covered entity shall <b>document</b> a restriction in accordance with §160.530(j) of this subchapter.
IF ("corresponding verb" == TRUE AND "implied trigger" == False AND statement type = type2 ) Then type2 with correlative	The data subject shall be able to <b>obtain</b> confirmation as to whether or not personal data concerning him or her are being processed from the controller.
IF ("corresponding verb" == TRUE AND "implied trigger" == TRUE AND statement type = type2 ) Then type2 with implied	A consumer shall have the right to <b>request</b> to a business that they disclose to that consumer the sale of personal information.

beginning of a statement with its individual list items to form complete, independent sentences.

A legal expert (second author of the paper) annotated the 1000 statements with the applicable template category from the ones presented in Section III. We note that the legal expert decided on matters of legal interpretation within the scope of the annotation guidelines detailed above. We discuss the foreseen validity threats in Section VI.

### B. Rules for Legal Requirements Templates Recommendation

Table III presents the template recommendation rules that we derived by analyzing the 1000 statements in our study. To maximize accuracy over these templates, we did five iterations over the 1000 statements, progressively refining the rules. For the first iteration, we built rules for a batch of 200 statements. From the second iteration, we evaluated against a new batch of statements and refined the rules until saturation of the evaluation metrics over all the batches.

The element highlighted in blue in each rule of Table III is the marker, i.e., the target of that rule.

The first step in our approach for recommending legal requirements templates is to use the statement-level semantic metadata we developed in Sleimi et al. [8] to classify the legal provision as expressed in the legal text. As noted in Section III, our conceptual model has six different statement types: fact, definition, penalty, permission, obligation, prohibition. As noted before, we discard statements classified as *fact* or *definition*, as they do not express requirements. Each template recommendation has two parameters: a statement type and a template category classification. The statement type assigned to the legal statement by the metadata extraction module is used to restrict the choice of templates. For example, if the legal statement as expressed by the text is classified as *obligation*, the possible templates will be restricted to *duty with no counterpart* and *duty with correlative right*.

The second step consists of extracting and processing the main verb. We extract the main verb using the following Tregex patterns:<sup>2</sup>

- SENT <(VN=mark)
- SENT <(VPinf <(VN <(VPP=mark)))
- SENT <(PP <(VPinf <(VN <(VPP=mark))))
- SENT <(VPinf <(VN <(VINf=mark)))
- SENT <(PP <(VPinf <(VN <(VINf=mark))))

Note that the keyword *mark* in the rules leads to extracting the verb that forms the linguistic root of the action. We

<sup>2</sup>VPinf means an infinitive clause. VPP means a nonfinite clause. VN means a verbal nucleus. Vinf means an infinitive verb. PP means a prepositional phrase. For details about Tregex, we refer the reader to [32].

illustrate some of our rules in Table III, in order to facilitate understanding and to discuss some important technicalities of the rules. Having identified the main verb, the next step consists of fetching the possible frames for that verb, using VerbNet. Our qualitative study indicates in fact that the main verb is the most reliable element from which we can infer the presence of multiple viewpoints applicable to the statement (also roles can signal such viewpoints, but they are often left implicit in legal statements). More specifically, the presence of multiple viewpoints is signaled by the main verb of the sentence being a *corresponding verb*. Corresponding verbs are two verbs that express two viewpoints on the same situation. They share the same roles, but with their orders switched. Let us consider, for example, the corresponding verbs *to send* and *to receive*: the subject of *send* is the indirect object of *receive*, and vice versa. Our definition of corresponding verbs is in line with the definition provided by Breau et al. [1] for *direct provisions*, i.e., “provisions which have a binary opposite where the subject of the activity assumes the value of the co-requisite attribute”. We classify the main verb as a *corresponding verb* if VerbNet includes, within the possible frames for said verb, one of the following roles: recipient, patient, experiencer, or theme (as an animate object). The design heuristics for this choice are described below. If a corresponding verb is detected, we know that the statement is not a *statement with no counterpart* (the first group of templates presented in Section III). We then proceed to verify whether the main verb, already marked as corresponding verb, is also an *implied trigger*. We define an *implied trigger* as a verb expressing an interaction where an agent assigns a right or obligation to another person, in line with the definition of *delegation* provided by Breau et al. [1]. In order to detect the presence of an implied trigger, we look for the verb in our curated list of delegation verbs<sup>3</sup>.

Fig. 1 presents a summary of the approach. To illustrate the rules, we now describe how they apply to the example statements in the second column of the table.

*Statement 1* reads as: “A covered entity shall **document** a restriction in accordance with §160.530(j) of this subchapter.” The statement is classified as obligation, and the main verb (the *mark* keyword in the tregex pattern previously described) is “to document”, that is not a correlative verb. As a consequence, the template *duty with no counterpart* is recommended.

*Statement 2* reads as: “The data subject shall have the right to **obtain** confirmation as to whether or not personal data concerning him or her are being processed from the controller.”

<sup>3</sup><http://shorturl.at/wEI57>

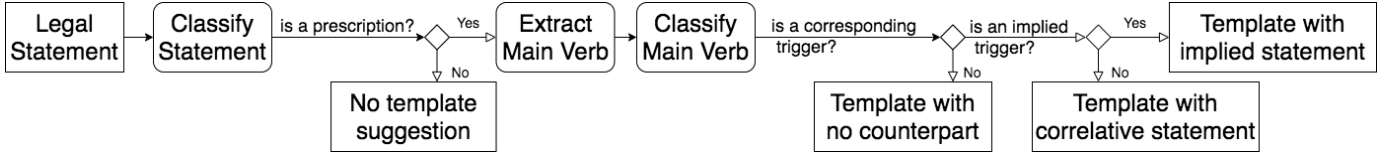


Fig. 1. Overview of Our Approach for Requirements Template Recommendation

The statement is classified as permission, and the main verb, “to obtain”, is a correlative verb but not an implied trigger. The template *permission with correlative duty* is recommended.

*Statement 3* reads as: “A consumer shall have the right to **request** that a business disclose to that consumer the sale of personal information.” The statement is classified as permission, and the main verb, “to request”, is both a correlative verb and an implied trigger. Therefore the template *permission with implied duty* is recommended. The recommended template for the implied statement is that of *transitive requirement*.

**Design heuristics.** During our study, we reviewed the list of corresponding verbs for our dataset in light of legislative drafting practices. We note that the roles *experiencer* and *patient* on their own do not constitute cues of a corresponding verb, especially in the absence of the role *agent*. In other cases, although the frame includes the role *theme as an animate object*, the verb does not express corresponding statements. As a result, we excluded 83 verbs and two phrasal verbs which met the aforementioned criteria. In addition, we implemented two heuristics resolving two issues related to legal drafting practices that we encountered during the qualitative study:

- Several statements in the qualitative study have main verbs that on their own do not have a correlative (e.g., “to keep”). However, in these statements, these verbs are part of compound expressions that prompt a correlative statement (e.g., “to keep confidential”). We also made the same observation with nominalizations; our first heuristic is related to the presence of these compound expressions or nominalizations: After we extract the main verb, we validate whether one such expression is present in the statement. If that is the case, we consider the complete compound expression instead of the main verb only. This allows us to correctly recommend a correlative template based on the expression that contains the verb, rather than on the verb alone.
- In some cases, the statements have a main verb that would under normal circumstances prompt a correlative statement. Take, for example, the verb “to take” which has a correlative “to give”. However, this verb can be part of a phrasal verb that would indicate a statement with no counterpart, for example, “to take effect”. Here, we cannot elicit any legal requirement about a correlative action. The second heuristic is related to the presence of correlative verbs that can be part of phrasal verbs. After we extract these correlative verbs, we do not immediately recommend a template “with correlative statement”. Instead, we consider the presence of these phrasal verbs to prevent incorrect recommendations.

The rules and design heuristics presented in this section enable the automatic recommendation of legal requirement templates, thus providing an answer to RQ2.

## V. EMPIRICAL EVALUATION

In this section, we describe our implementation and measure the accuracy of our approach through a case study.

### A. Implementation

Our template recommendation rules are implemented using Tregex and Java. The rules utilize the outputs of the classic NLP pipeline for syntactic analysis. We also use our framework [8] for semantic metadata extraction.

### B. Accuracy of the Template Recommendation

**Case study description.** The objective of our case study is to measure the accuracy of the template recommendation rules of Table III against a ground truth. To build the ground truth, a legal expert manually classified 400 randomly selected legal statements from the labor and health laws, in addition to the 1000 statements of our qualitative study (see Section IV). The construction of the ground truth took place after the conclusion of our qualitative study. As explained in Section III, we excluded legal statements that express metarules. In order to have a cohesive dataset for the ground truth, we also excluded statements that have contractual effects, i.e., statements for which the interpretation of contract law is necessary in order to be able to identify correlative statements. An example is the statement “The employer may terminate the contract after thirty days”. Here, the action “to terminate” by itself would have no consequence for the employer. However, depending on the contract that is terminated, this could lead for example to (a) a severance package (i.e., a correlative right for the employee) or (b) a temporary prohibition for the former employee to approach the clients of the employer if there is a non-compete clause.

Our analysis of the results did not lead to new template recommendation rules.

**Analysis procedure.** Each template recommendation has two parameters: a statement type and a template-category classification. The first is assigned following the statement-level metadata in our conceptual model [8]; because it is not a contribution of this paper, we do not evaluate it. The second parameter is assigned by the approach described in Section IV. We evaluate the second parameter, i.e., the automated template recommendation rules, using the following notions:

- A recommended template is a *match* if it has the same template category as the ground-truth classification.
- A recommended template is *misclassified* if it has a different template category than the ground-truth classification.
- A ground-truth statement for which the methodology did not provide a match is considered as *missed*.

Our evaluation results are presented in columns 4 through 8 of Table IV. For each category of templates, we provide the

TABLE IV  
STATISTICS FOR TEMPLATE RECOMMENDATIONS

Corpus	Template Category	Ground Truth Results	Correct (TP)	Misclassified (FP)	Missed (FN)	Precision	Recall
Labor	Statement with no counterpart	96	80	11	16	87,9%	83,3%
	Statement with correlative statement	47	41	15	6	73,2%	87,2%
	Statement with implied statement	13	9	1	4	90,0%	69,2%
	<b>Subtotal</b>	<b>156</b>	<b>130</b>	<b>27</b>	<b>26</b>	<b>82,8%</b>	<b>83,3%</b>
Health	Statement with no counterpart	107	84	14	23	85,7%	78,5%
	Statement with correlative statement	37	32	22	5	59,3%	86,5%
	Statement with implied statement	16	14	3	2	82,4%	87,5%
	<b>Subtotal</b>	<b>160</b>	<b>130</b>	<b>39</b>	<b>30</b>	<b>76,9%</b>	<b>81,3%</b>
<b>Total</b>		<b>316</b>	<b>260</b>	<b>66</b>	<b>56</b>	<b>79,8%</b>	<b>82,3%</b>

TABLE V  
ERROR ANALYSIS FOR TEMPLATE RECOMMENDATIONS

Template Category	Result Type	Nominalization	Phrasal Verb	Legal Terminology	Error in Metadata	Design Heuristics
Statement with no counterpart	Misclassified	11	8	0	1	5
	Missed	7	11	8	8	5
Statement with correlative statement	Misclassified	5	15	6	7	4
	Missed	6	2	0	0	3
Statement with implied statement	Misclassified	1	0	2	0	1
	Missed	1	0	4	1	0

number of correct matches, misclassified and missed template-rule recommendations, and scores for precision and recall.

Each match counts as a true positive (TP). Each misclassified recommendation counts as a false positive (FP), and each missed recommendation counts as a false negative (FN).

Precision is computed as  $|TP|/(|TP| + |FP|)$  and recall as  $|TP|/(|TP| + |FN|)$ . The final row in the table shows the overall results. Note that the overall precision and recall scores are computed over all the recommended templates across both domains, and are not the averages of the precision and recall scores for the individual template categories.

**Results.** We first discuss the results for the recommended templates for *statements with no counterpart*. Second, we present the results for the *statements with correlative statement* and the *statements with implied statement*. Third, we discuss the discrepancies between the results in the two legal domains considered in our case study. Finally, we perform an error analysis on the misclassifications and missed template recommendations.

**Results for statements with no counterpart.** Out of 189 recommended templates annotated as *with no counterpart*, 164 were correct matches and 25 were misclassifications. 39 statements with no counterpart in the ground truth were missed. The error analysis is presented in Table V. Our error analysis (summary) results are presented in columns 3 through 7 of Table V. For each category of templates, we provide the number of inaccuracies leading to errors, for each situation. These situations are formally introduced and discussed at the end of this subsection. We obtain an overall precision of 86,7% and an overall recall of 80,7%.

**Results for statements with correlative statement.** Out of 110 recommended templates annotated as *with correlative statement*, 73 were correct matches and 37 were misclassifications. 11 statements with correlative statement in the ground truth were missed. The error analysis is presented in Table V.

We obtain an overall precision of 66,3% and an overall

recall of 86,9%.

**Results for statements with implied statement.** Out of 27 recommended templates annotated as *with implied statement*, 23 were correct matches and 4 were misclassifications. 6 statements with implied statement in the ground truth were missed. The error analysis is presented in Table V.

We obtain an overall precision of 85,1% and an overall recall of 79,3%. We note that we did not have enough statements from this template category in the dataset (7,25%) to draw meaningful conclusions on the accuracy of our rules.

**Legal Domains.** Regarding the two legal domains, our recommendation rules performed well in both cases, but slightly better over the labor law: we obtained a precision of 82,8% and a recall of 83,3%, while for the health law the precision was 76,9% and the recall 81,3%. The difference is due to the fact that, while for the labor law we could easily classify and exclude contractual obligations, for the health law there was no clear way of excluding rules that require external knowledge. The error analysis, however, confirmed that the types of errors are equally distributed across the two domains of our study.

**Answering RQ3.** We can now provide an initial answer to RQ3 based on our quantitative results: our recommendation rules achieve good accuracy with a recall of 82,3% and a precision of 79,8% over the two examined domains in our case study. Despite being good, these accuracy results are still far from perfect. Therefore, based on the quantitative results alone, our study suggests that analysts will need to carefully validate the recommended templates and discard the incorrect ones. Nonetheless, as we are going to argue next, these quantitative results per se are not reflective of the true usefulness of our approach, as they can be considerably improved by introducing a human-feedback component.

**Error analysis.** To identify the root causes for automation inaccuracies, we analyzed the misclassified and missed template recommendations. As indicated in Table V, the inaccuracies stem from five different situations: (a) Nominalizations, (b) Phrasal verbs, (c) Legal terminology, (d) Errors in automated metadata extraction and (e) Design decisions and heuristics.

In *nominalizations* (deverbal nouns [33]), as noted by Breaux et al. [1], the main action (e.g., to investigate) is nominalized in a form (e.g., investigation) that we did not encounter during our design heuristics process (see Section IV).

In the case of *phrasal verbs* such as “to take effect”, we



populated a list to be used by the heuristics; however, this list turned out to be incomplete during our case study, though we would expect it to become increasingly more complete as we cover more domains. Besides, in some cases, we could not devise heuristics because the errors stemming from the new design decision outweighed the correct recommendations. A possible approach to solve the issue would be to paraphrase these constructs, but automating such paraphrasing is outside the scope of this paper.

In the case of verbs that are part of *legal terminology*, the issue is that the semantics of the verb when used in a legal statement can be different from the semantics of the verb in its general lexicon. For example, the verb “to suspend” supports a corresponding statement in its general meaning (“the employer cannot suspend the payment of the salary in ordinary circumstances”). However, when used in its legal meaning (“This regulation is suspended until December 31st, 2010”) it does not express any legal requirement. We made an effort to build a list for this terminology during our design heuristics process, but it turned out to be incomplete.

Regarding the *errors in automated metadata extraction*, which originate from our metadata extraction framework [8], a validation of the necessary metadata by a human annotator might reveal these errors. Nevertheless, and until a more accurate solution for automated metadata extraction is available, we consider this semi-automatic approach the best trade-off between the human effort required and the accuracy of the results. Finally, we note that the errors stemming from our own *design decisions and heuristics* are outweighed by the correct recommendations, which was the reason for our adoption of these heuristics in the first place during the qualitative study.

### C. Observations and Lessons Learned

In this section, we present the observations and lessons learned from our case study.

**Integrating Human feedback.** Based on the error analysis presented above, our approach can be modified to enable smart and minimal interactions with the analyst to prevent most of the errors in template recommendations. Such a semi-automated process would prompt the intervention of the analyst for tasks that are relatively simple, such as reformulating a nominalized verb or disambiguating between the general and legal use of a verb. In an interactive mode of use, the analyst’s intervention would occur as a pre-processing step in order to assist our approach (Fig. 1) with correctly extracting actions from statements. Take for example the statement “The investigation may be performed”. Here, the template recommendation approach would prompt the analyst to answer the following simple yes/no question: “The nominalization *investigation* is the main action of the statement. Is this correct?” With this human-in-the-loop component, we can identify the majority of the legal requirements (38 out of 56) currently missed by our fully automated approach.

The observations presented above clearly show how the presented semi-automated approach would provide much higher recall and precision. Specifically, adding this human-feedback

component would increase recall by 12% and precision by 10,8%, thus yielding an overall recall of 94,3% and an overall precision of 90,6%. To conclude, our quantitative and qualitative analyses clearly suggest that a semi-automated, interactive approach is a better option for legal template recommendation.

**Action in legal statements.** We highlighted in Section IV how identifying the main action expressed by a legal statement is key to detecting the presence of multiple viewpoints. The main action is in fact more important than the statement type for identifying the presence of multiple viewpoints. We also note that detecting the type of action from the main verb contained in the legal statement is especially difficult in four situations: (a) nominalizations, (b) phrasal verbs, (c) legal terminology, and (d) implicit roles (subjects or counterparties).

The first three issues were described in the previous subsection. The issues of nominalizations and phrasal verbs are not specific to the legal domain. A possible solution could be (a) the identification of these linguistic constructs, and (b) their resolution through lemmatization (for nominalizations) and the use of a locution thesaurus<sup>4</sup> (for phrasal verbs). Regarding verbs that employ legal terminology, semantic-role labeling should in theory be able to tackle different linguistic forms to express actions. Unfortunately, the most robust tools such as FrameNet are not trained for the legal language, and thus not very effective in the context of our work.

Roles that are left implicit in the text are either referenced by anaphora or totally omitted. A common drafting technique is the conjugation of the main verb in the passive form, omitting the agent. For this reason, during the elaboration of our rules for template recommendation (see Section IV), we could not rely on the presence of roles within the legal sentence. For the same reason, we cannot rely on the presence of roles for deriving heuristics that address the first three issues.

Solving the issue of implicit roles therefore seems to be a priority, because of the potential that roles carry for creating new recommendation rules. Solving this issue requires anaphora resolution for implicit roles and a domain model for omitted roles.

**The importance of a domain model.** As noted in the previous subsection, some misclassifications were caused by the fact that the actions contained in the legal text only expressed multiple perspectives when seen in the light of the applicable legal framework. For this reason, we had to exclude contractual obligations from our study. In order to overcome this limitation, it is necessary to extract knowledge from additional sources, e.g., contract law for contractual obligations. As noted also in the previous observation on omitted roles, this could be achieved by relying on domain models.

**The importance of corresponding statements in extracting requirements from legislation.** Our preliminary study on five different legal domains highlighted that the relevance of the issue is domain dependent: while about one out of

<sup>4</sup>A locution is a sequence of words (a phrase) in the sentence that has the same grammatical (semantic) value of a single word. A locution thesaurus is therefore a resource that groups locutions with their corresponding words according to semantic similarity.

four statements in health and labor laws express multiple viewpoints, other domains have a much lower ratio. This is only marginally due to drafting techniques, and rather depends on the type of legal relations that are predominant in the domain. For example, health and labor are domains that are interested by many constitutional guarantees and therefore the laws in these domains often attribute rights to subjects. On the hand, the commerce and environment domains are more focused on ensuring due diligence by the operators and the requirements are therefore often expressed directly as duties.

We notice nevertheless that the criteria to detect multiple viewpoints, that we embedded in our recommendation rules, work well across domains.

## VI. THREATS TO VALIDITY

**Internal validity.** A potential threat to internal validity is related to the subjectivity of legal analysis and how it affects the elicitation of legal requirements. Oftentimes, complex requirements specifications are organized hierarchically with the requirements expressed at multiple levels of granularity. The same principle applies to legal requirements, but we are not aware of any systematic means for defining granularity levels for legal requirements. To mitigate subjectivity about granularity levels, we restrict our work to the granularity level at which the underlying legal text(s) have been articulated. Furthermore, we note that the coding in both the qualitative study of Section IV and the ground truth of Section V was done by the second author. To mitigate against potential subjectivity caused by the involvement of an author in coding, we set clear, upfront criteria for the analysis of the legal statements, and therefore for the ground-truth construction. First, annotations were limited to legal statements that explicitly led to legal requirements for IT systems and/or their stakeholders. Second, we completed the coding component of our qualitative study before defining any recommendation rules. Third, we left out statements that involved the interpretation of external sources, e.g., metarules and contractual obligations. Finally, we did not apply our implementation to the legal statements in the ground truth until the coding was completed.

**External validity.** The first threat to external validity is related to the generalizability of our results; for this, we refer the reader to our observations on legal domains in Section V.

The second threat to external validity is related to the differences between languages. Our corpus of legal texts is in French. Our current tool support would thus not readily work for other languages and needs to be adapted in terms of both linguistic cues and heuristics. In the particular case of the English language, one has access to highly developed NLP frameworks that have been trained on very large corpora; this is likely to increase the accuracy of requirements template recommendations, but verifying this claim requires separate empirical investigations.

The third threat to external validity is related to the other ways in which a legal statement can express more than one legal rule. Our goal in this paper was to extract from a given legal statement all the legal requirements that are the result of

multiple viewpoints. However, this is not the only case of one-to-many relations between legal statements and legal rules.

There can in fact be multiple possible interpretations for a legal statement. A multitude of interpretations is different from a multitude of viewpoints in that interpretations are alternatives, i.e., they cannot be valid in the same legal context at the same time. Detecting, extracting, and comparing alternative legal interpretations are topics for research in AI and Law and outside the scope of our current work.

It is also possible that multiple legal statements (e.g., a duty and a definition) are contained within a single legal sentence. An example is “The controller is forbidden from storing sensitive data, which means data that holds sensitive information”. These statements are expressed in different parts of the sentence, and the sentence could be split into two different sentences, one per rule, without altering its meaning and without redundancy except for the phrase “sensitive data”. This process is, however, not always straightforward, and it can be argued that establishing the granularity of rules in a legal provision is a matter of legal interpretation. Extracting multiple rules merged into a single legal statement remains outside the scope of the present research.

## VII. CONCLUSION

In this paper, we presented an approach to automatically recommend templates for legal requirements based on legal statements, thus assisting requirements analysts with legal requirements elicitation. To do so, we first defined a set of templates that account for multiple viewpoints. These templates are grouped into three categories: statements with no counterpart, statements with a correlative statement, and statements with an implied statement. We then devised, using Natural Language Processing, automated rules for recommending suitable requirements templates. We evaluated our approach on 400 statements from labor and health laws in Luxembourg. Our results show good accuracy with a recall of 82,3% and a precision of 79,8%.

We further collected and synthesized knowledge about the verb constructs that were the cause of incorrect recommendations. We outlined how such knowledge can be leveraged for developing a semi-automated, human-in-the-loop approach that can much more accurately identify suitable requirements templates based on minimal input from legal experts.

For future work, we intend to include a domain modeling element in our approach. This would support the elicitation of legal requirements from inter-connected legal statements, thus resolving the challenge of metarules and cross-references. We would further like to expand our approach so that it not only recommends suitable templates, but also fills (populates) the templates by pulling in relevant information from the underlying legal statements.

**Acknowledgments.** This paper was supported by SCL, Luxembourg’s National Research Fund (FNR) under grant PUBLIC2-17/IS/11801776, and NSERC of Canada under the Discovery, Discovery Accelerator and CRC programs. We thank Nicolas Sannier for helpful discussions.

## REFERENCES

- [1] T. D. Breaux, M. W. Vail, and A. I. Antón, "Towards regulatory compliance: Extracting rights and obligations to align requirements with regulations," in *Proceedings of the 14th IEEE International Requirements Engineering Conference (RE'06)*, 2006, pp. 46–55.
- [2] P. Darke and G. Shanks, "Stakeholder viewpoints in requirements definition: A framework for understanding viewpoint development approaches," *Requirements Engineering*, vol. 1, no. 2, pp. 88–105, 1996.
- [3] E. Hull, K. Jackson, and J. Dick, *Requirements Engineering*. Springer London, 2010.
- [4] D. Leffingwell and D. Widrig, *Managing Software Requirements: A Use Case Approach*. Addison-Wesley, 2003.
- [5] R. Young, *Effective Requirements Practices*. Addison-Wesley, 2001.
- [6] J. Young and A. Antón, "A method for identifying software requirements based on policy commitments," in *Proceedings of the 18th IEEE International Requirements Engineering Conference (RE'10)*, 2010, pp. 47–56.
- [7] T. D. Breaux and D. G. Gordon, "Regulatory requirements traceability and analysis using semi-formal specifications," in *Requirements Engineering: Foundation for Software Quality*. Springer, 2013, pp. 141–157.
- [8] A. Sleimi, N. Sannier, M. Sabetzadeh, L. C. Briand, and J. Dann, "Automated extraction of semantic legal metadata using natural language processing," in *Proceedings of the 26th IEEE International Requirements Engineering Conference (RE'18)*, 2018, pp. 302–311.
- [9] J. Bentham and H. Hart, *Of laws in general*. University of London, Athlone Press, 1945.
- [10] W. N. Hohfeld, "Fundamental legal conceptions as applied in judicial reasoning," *The Yale Law Journal*, vol. 26, no. 8, pp. 710–770, 1917.
- [11] N. Kiyavitskaya, A. Krausov, and N. Zannone, "Why eliciting and managing legal requirements is hard," in *2008 Requirements Engineering and Law*, 2008, pp. 26–30.
- [12] C. Palomares, C. Quer, and X. Franch, "Requirements reuse and requirement patterns: a state of the practice survey," *Empirical Software Engineering*, vol. 22, no. 6, pp. 2719–2762, 2017.
- [13] S. Robertson, "Requirements patterns via events/use cases," in *Proceedings of the Third Conference on the Pattern Languages of Programs (PLoP'96)*, 1996.
- [14] M. B. Dwyer, G. S. Avrunin, and J. C. Corbett, "Patterns in property specifications for finite-state verification," in *Proceedings of the 21st International Conference on Software Engineering (ICSE'99)*, 1999, pp. 411–420.
- [15] A. Mavin, P. Wilkinson, A. Harwood, and M. Novak, "Easy approach to requirements syntax (ears)," in *Proceedings of the 17th IEEE International Requirements Engineering Conference (RE'09)*, 2009, pp. 317–322.
- [16] Y. Yoshida, K. Honda, Y. Sei, H. Nakagawa, Y. Tahara, and A. Ohsuga, "Towards semi-automatic identification of functional requirements in legal texts for public administration," in *Proceedings of the 26th Annual Conference on Legal Knowledge and Information Systems (JURIX'13)*, 2013, pp. 175–184.
- [17] M. Palmirani, G. Governatori, A. Rotolo, S. Tabet, H. Boley, and A. Paschke, "Legalruleml: Xml-based rules and norms," in *Proceedings of the 5th International Symposium on Rule-Based Modeling and Computing on the Semantic Web (RuleML'11)*, 2011, pp. 298–312.
- [18] L. Márquez, X. Carreras, K. C. Litkowski, and S. Stevenson, "Semantic role labeling: An introduction to the special issue," *Computational Linguistics*, vol. 34, no. 2, pp. 145–159, 2008.
- [19] D. Gildea and D. Jurafsky, "Automatic labeling of semantic roles," *Computational Linguistics*, vol. 28, no. 3, pp. 245–288, 2002.
- [20] C. F. Baker, C. J. Fillmore, and J. B. Lowe, "The berkeley framenet project," in *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics (COLING-ACL'98)*, 1998, pp. 86–90.
- [21] Y. Miyao and J. Tsujii, "Deep linguistic analysis for the accurate identification of predicate-argument relations," in *Proceedings of the 20th International Conference on Computational Linguistics (COLING'04)*, 2004, pp. 1392–1398.
- [22] K. Kipper, A. Korhonen, N. Ryant, and M. Palmer, "A large-scale classification of english verbs," *Language Resources and Evaluation*, vol. 42, no. 1, pp. 21–40, 2008.
- [23] B. Levin, *English Verb Classes and Alternations: A Preliminary Investigation*. University of Chicago Press, 1993.
- [24] N. Zeni, L. Mich, and J. Mylopoulos, "Annotating legal documents with gaiust 2.0," *Int. J. Metadata Semant. Ontologies*, vol. 12, no. 1, pp. 47–58, 2017.
- [25] A. Ottensooser, A. Fekete, H. A. Reijers, J. Mendling, and C. Menic-tas, "Making sense of business process descriptions: An experimental comparison of graphical and textual notations," *Journal of Systems and Software*, vol. 85, no. 3, pp. 596–606, 2012.
- [26] Z. Sharafi, A. Marchetto, A. Susi, G. Antoniol, and Y. Guéhéneuc, "An empirical study on the efficiency of graphical vs. textual representations in requirements comprehension," in *Proceedings of the 21st IEEE International Conference on Program Comprehension (ICPC'13)*, 2013, pp. 33–42.
- [27] S. Ghanavati, D. Amyot, and A. Rifaut, "Legal goal-oriented requirement language (legal GRL) for modeling regulations," in *6th International Workshop on Modeling in Software Engineering, MiSE*. ACM, 2014, pp. 1–6.
- [28] M. Ceci, T. Butler, L. O'Brien, and F. Al Khalil, "Legal patterns for different constitutive rules," in *AI Approaches to the Complexity of Legal Systems*. Springer International Publishing, 2018, pp. 105–123.
- [29] T. Athan, H. Boley, G. Governatori, M. Palmirani, A. Paschke, and A. Z. Wyner, "OASIS LegalRuleML," in *Proceedings of the International Conference on Artificial Intelligence and Law (ICAIL'13)*, 2013, pp. 3–12.
- [30] T. D. Breaux, "A method to acquire compliance monitors from regulations," in *Proceedings of the 3rd International Workshop on Requirements Engineering and Law (RELAW'10)*, 2010, pp. 17–26.
- [31] N. Sannier, M. Adedjouma, M. Sabetzadeh, and L. C. Briand, "An automated framework for detection and resolution of cross references in legal texts," *Requirements Engineering*, vol. 22, no. 2, pp. 215–237, 2017.
- [32] R. Levy and G. Andrew, "Tregex and tsurgeon: tools for querying and manipulating tree data structures," in *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC'06)*, 2006, pp. 2231–2234.
- [33] O. Gurevich, R. Crouch, T. King, and V. De Paiva, "Deverbal nouns in knowledge representation," *Journal of Logic and Computation*, vol. 18, 01 2006.