LegalLanguage: A Domain-Specific Language for Legal Contexts

3 authors:

Ambrósio Soares
Universidade do Algarve
3 PUBLICATIONS **2** CITATIONS

Paula Martins
Universidade do Algarve
48 PUBLICATIONS **234** CITATIONS

Alberto Rodrigues Da Silva
University of Lisbon
256 PUBLICATIONS **1,982** CITATIONS

Some of the authors of this publication are also working on these related projects:

Project   SilabMDD View project

Project   BOA - Learning Objects Repository View project

Metadata of the chapter that will be visualized in SpringerLink

Book Title	Advances in Enterprise Engineering XIII	
Series Title		
Chapter Title	LegalLanguage: A Domain-Specific Language for Legal Contexts	
Copyright Year	2020	
Copyright HolderName	Springer Nature Switzerland AG	

Corresponding Author	Family Name	**Soares** Text
	Particle	
	Given Name	**Ambrósio Alves**
	Prefix	
	Suffix	
	Role	
	Division	
	Organization	Faculdade de Ciências e Tecnologia da Universidade do Algarve
	Address	Faro, Portugal
	Email	a20982@ualg.pt
Author	Family Name	**Martins**
	Particle	
	Given Name	**Paula Ventura**
	Prefix	
	Suffix	
	Role	
	Division	
	Organization	Faculdade de Ciências e Tecnologia da Universidade do Algarve
	Address	Faro, Portugal
	Email	
	ORCID	http://orcid.org/0000-0002-3199-4517
Author	Family Name	**Silva**
	Particle	**da**
	Given Name	**Alberto Rodrigues**
	Prefix	
	Suffix	
	Role	
	Division	INESC-ID, Instituto Superior Técnico
	Organization	Universidade de Lisboa
	Address	Lisbon, Portugal
	Email	
	ORCID	http://orcid.org/0000-0002-7900-9846

Abstract	Nowadays legal ontologies have been used in the legal domain, however, being poorly explored in legislative and production processes. This paper analyses the adoption of legal ontologies as a tool to support these processes, in particular, related to activities span from the submission of bills and their subsequent authoring and ratification. This paper introduces the state of the art of legal (or normative)

ontologies; and also discusses some application examples. The analysis of this state of the art allows us to identify some problems, namely regarding the activities involving the authoring and validation of laws that tend to be very human-intensive and error-prone. As a consequence of this analysis, we introduce the LegalLanguage, a language particularly suitable for the authoring and specification of law(s) in a more rigorous and systematic way, that would allow to keep track different types of intra and inter-laws relationships (e.g., structural, order or temporal relationships between articles or even between laws). Finally, a simple illustrative example is used and shows the importance of a language like LegalLanguage in the production of normative documents.

Keywords	Legal tech - Legal ontologies - Domain Specific Language

LegalLanguage: A Domain-Specific Language for Legal Contexts

Ambrósio Alves Soares[1]([envelope]), Paula Ventura Martins[1] [ID],
and Alberto Rodrigues da Silva[2] [ID]

[1] Faculdade de Ciências e Tecnologia da Universidade do Algarve,
Faro, Portugal
a20982@ualg.pt
[2] INESC-ID, Instituto Superior Técnico, Universidade de Lisboa,
Lisbon, Portugal

Abstract. Nowadays legal ontologies have been used in the legal domain, however, being poorly explored in legislative and production processes. This paper analyses the adoption of legal ontologies as a tool to support these processes, in particular, related to activities span from the submission of bills and their subsequent authoring and ratification. This paper introduces the state of the art of legal (or normative) ontologies; and also discusses some application examples. The analysis of this state of the art allows us to identify some problems, namely regarding the activities involving the authoring and validation of laws that tend to be very human-intensive and error-prone. As a consequence of this analysis, we introduce the LegalLanguage, a language particularly suitable for the authoring and specification of law(s) in a more rigorous and systematic way, that would allow to keep track different types of intra and inter-laws relationships (e.g., structural, order or temporal relationships between articles or even between laws). Finally, a simple illustrative example is used and shows the importance of a language like LegalLanguage in the production of normative documents.

Keywords: Legal tech · Legal ontologies · Domain Specific Language

1 Introduction

The need to allow the adequate improvement of practices during the writing phase of normative texts (e.g., laws, regulations) of legal-based organization (e.g., parliamentary, lawyers society), led to the study of legal ontologies as a convenient tool to better support the activities of the production process (e.g., the legislative process) in the period of preparation of proposals of laws that includes a set of phases and acts such as the creation and authoring, but also validation and publishing of such normative texts. The empirical studies found in the literature on legal ontology, like [1–3], present aspects for the understanding and use of the subject in the perspective of legal application (judicial process) and not in the process of legislative production (legislative process). There are few studies [4] on the use of legal ontologies as supporting parliamentary activities, of the legislative process.

Figure 1 shows the representation of the two top-level legal processes: the process of producing laws (legislative process); and the process of applying or using of such laws (judicial process).
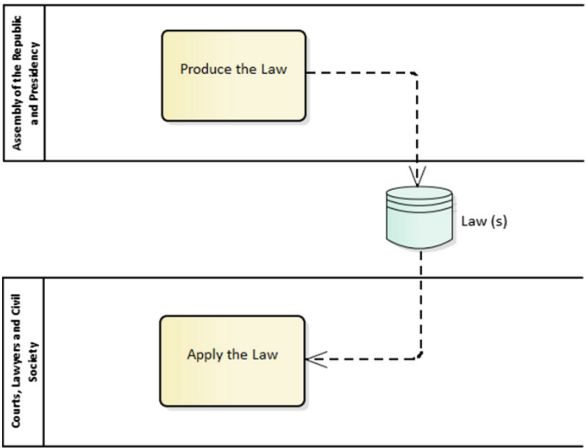

Fig. 1. The two top-level legal processes (in BPMN notation)

The legislative process (production laws) involves the authoring, validation and publication of Laws; while the judicial process (application the law) initiates and receives the final version of the Law (defined from the previous process), giving continuity to the activities related to the use and application of laws. The focus of our research is mainly related to the support of the first process (i.e., the production of laws and other normative texts), which can be considered at the parliamentary scope but also on other levels e.g., at the level of lawyer societies.

Normative legal documents written directly in natural languages (e.g., English, Spanish or Portuguese) present usually coherence failures, as for example (i) at the morphological level, where one deals with the composition of the words and their nature; (ii) at the level of the lexicon, which interprets the individual meaning of words; (iii) at the syntactic level, which focuses on the analysis of the sentence's composition; or (iv) at the semantic level where the meaning of the words or sentences are dealt with.

In the legal domain (in particular parliamentary scope) the responsible for producing laws tend to use common Word Processors tools (like Microsoft Word, OpenOffice or Google Docs) for drafting legal normative acts. However, these software tools do not provide specific support to facilitate such processes. The absence of adequate and specific tools does not help to avoid or mitigate the problems related to the legislative production process, such as: deficiencies in the analysis, navigation, search, traceability and creation of laws; absence of easy navigation between laws and their relationships; absence of traceability mechanisms between laws (e.g., with repealed or in force semantics); lexical and structural ambiguity in words, phrases and expressions; lack of clarity in the text of the body of laws already enacted; or difficulty to automatically check and normalize such laws.

To mitigate some of these problems we propose a legal-specific language, named as "LegalLanguage", that would allow to specify normative texts in a more rigorous and systematic way when compared with the traditional human-intensive and error-prone approaches. This research comes from our previous work on designing languages and tools to improve the rigor of text specifications in disparate domains like privacy policies [5–7] and requirements engineering [8–10]. In addition, recently we have also researched legal ontologies and concluded that the existent proposals are still limited or incomplete [11].

This paper is structured as follows. Section 2 introduces the state of the art on legal ontologies particularly applied in the legislative domain; Sect. 3 introduces and overviews the key aspects of the proposed LegalLanguage; Sect. 4 shows an illustrative example of applying the LegalLanguage; finally, Sect. 5 presents conclusions and overviews future work.

2 Legal Ontologies

The definition of the concept *"ontology"* differs in the context of computer science, and information of reasoning in the philosophical context [12]. According to Gruber, an ontology is a description of concepts and relationships that exist for an individual or a community [13]. The term is borrowed from philosophy, where an ontology is a systematic description of existence. In 2009, Gruber updated the definition of the concept in the context of computer science and information, as a set of representation primitives *(classes, attributes or properties, relations)* with which to model a knowledge domain [14].

Depending on the domain, legal ontologies have several approaches. Figure 2 distinguishes three different approaches, namely: semantic, epistemological and ontological approach. The semantic approach focuses on the meaning of a representation of elements and relationships in a given domain; the epistemological approach is relative
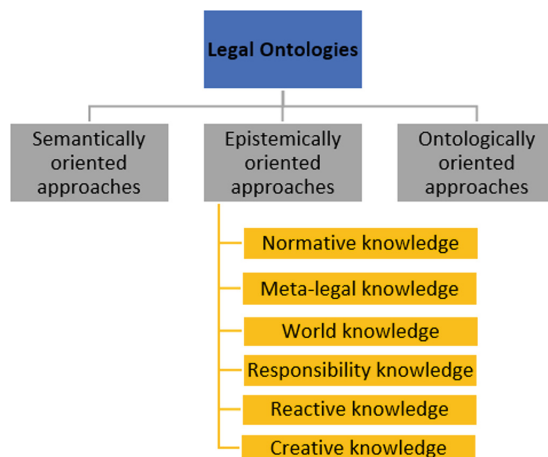

Fig. 2. Different approaches of legal ontologies (source: [12])

to the cognitive knowledge of a domain; the ontological approach gives greater prominence to the entities and relationships that constitute a domain.

The epistemological approach presented in Fig. 2 distinguishes six basic types of knowledge, from an example of this type of approach called functional ontology, that is: normative knowledge, meta-legal knowledge, knowledge of the world, knowledge of responsibility, reactive knowledge, and knowledge creative [12]. Normative knowledge, from the definition of the term "normative" which refers to rules, consists of elements of the legal field which prescribe the behavior of persons in the society in which they are inserted according to certain accepted rules.

Meta-legal knowledge consists of legal rules that govern relationships between different people in society and not directly behavior. World Knowledge does not necessarily mean geographical position, it consists of knowledge of the elements of the world, and how these elements determine how the world looks. Responsibility Knowledge consists in the linking of normative knowledge to reactive knowledge. In the normative case where there are presuppositions to exist of norms, the responsibility is the duty of an agent causing an event to bear the consequences of the own behavior. Reactive knowledge consists of the sanctions imposed if an agent violates a standard and is held accountable for it. Finally, the Creative knowledge of the "creative" concept, which highlights the imagination and ability to create something new, original. It consists of information about created institutions and other entities that arise from law enforcement.

The problems, presented in Sect. 1, related to the improvement in the normative treatment of legal texts, as well as the lack of ontologies related to the structure of legal documents, justify the accomplishment of this work on the normative knowledge type of the epistemological approach, focusing on the normative domain and emphasizing the structure of basic concepts of a law.

The semantic and ontological approaches to the particular domain, presented in Fig. 2, apply in the representation of structural elements of a law and relations, and not as a support for the writing of normative texts, which is the focus of our work.

The following is a brief description of some of the most important legal ontologies. At the conclusion of the chapter, the comparative study of these ontologies is presented based on several criteria, namely the categories of classification presented in Fig. 2, concepts of each ontology and capacity to respond to identified problems.

2.1 LKIF

The Legal Knowledge Interchange Format (LKIF) ontology is based on "basic concepts of law". Terminology issues become the main object of ontology.

LKIF has two main objectives: to allow translation between legal knowledge bases written in different formats and representation formalisms and, secondly, as a knowledge representation formalism that is part of a larger architecture for the development of knowledge systems tool.

Based on the categories presented in Fig. 2, LKIF is identified on the type semantic approach, since it deals with the translation and representation of knowledge, which fits within the principles of the semantic approach.
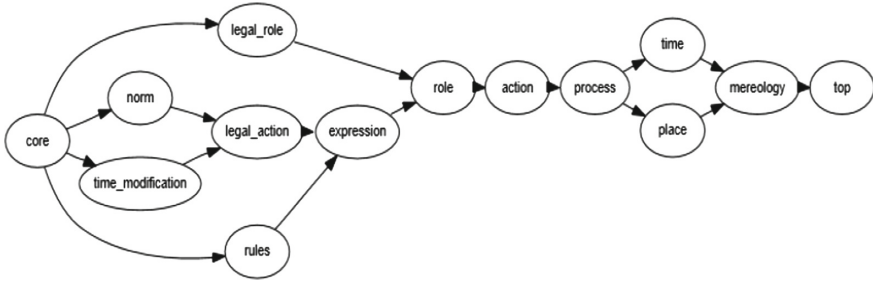
Fig. 3. Dependencies between the main modules of the LKIF ontology (source: [1]).

This ontology defines in five closely related modules the most abstract concepts: top, place, mereology, time, and space-time.

In the top module, LKIF's top ontology is largely based on the upper level of LRI-Core, but has less ontological commitment in that it imposes fewer constraints on subclasses of the major categories. The place module partially implements the theory of relative places in Languages to define and instantiate Web Ontologies (OWL) and based on Description Logics (DL). The mereology module defines theory or logical-mathematical study concepts of the relations between the parts and the whole, and of the relations between the parts within a whole (affiliation, etc.). The time module provides an implementation of the OWL DL of the theory of time [1].

The space-time module consists of basic level concepts that are distributed in four modules: process, role, action and expression.

The process module extends the LKIF top ontology module with a definition of changes, processes (being causal changes) and physical objects. This module introduces a limited set of properties to describe the participants' roles in the processes (see Fig. 3). The role module defines a typology of roles (epistemic roles, roles, personal roles, organizational roles) and the role-linking property. The action module describes the vocabulary to represent actions in general, and does not commit to a particular theory on thematic roles. Actions are processes performed by agents (actor of action). The expression module describes a vocabulary for reporting propositions and propositional attitudes (belief, intention), qualifications, statements, and the media. In addition, it extends the module papers with a number or epistemic roles, and is the basis for the definition of norms.

The remaining basic concepts are extended through three modules that form the legal ontology: legal-action, legal-role and norm. The legal-action module extends the action module with a set of legal concepts related to the action and the agent, such as public acts, public agencies, legal person, natural person, etc. The legal-role module extends the role module with a small number of legal concepts related to roles, legal professions, etc. The norm module is an extension mainly on the expression module, where norms are defined as qualifications. It also defines a number of legal sources, such as legal documents, customary law, etc., and a typology of rights and powers [15]. Further details on the explanation of the ontology can be found in the bibliography in [1].

2.2   CLO

The CLO (Core Legal Ontologies) approach aims to support the construction of legal domain ontologies [16], being classified as an ontological approach (Fig. 2), where the essential concepts of this ontology will be described in Fig. 4.
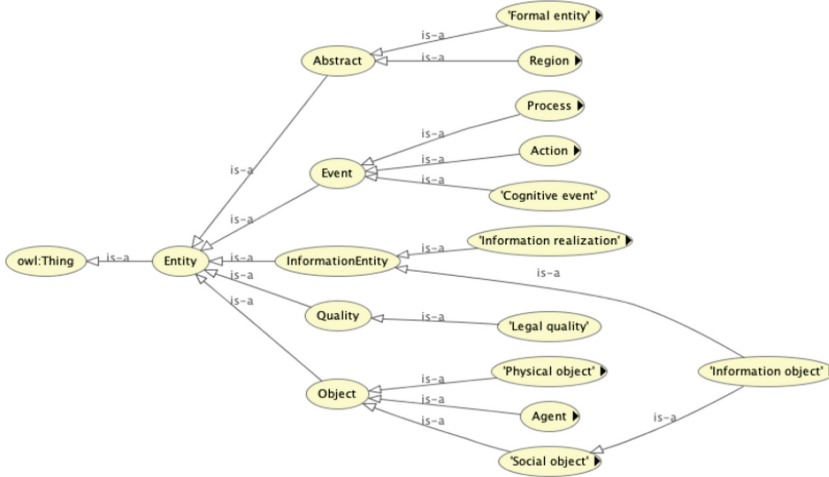

Fig. 4.   Context and stated situations (source: [2])

The CLO provides types and relationships for the heterogeneous entities from the legal domain, be it the physical, cognitive, social or legal realm. According to stratification, entities from different layers may be spatially temporally co-located, being completely different and (mutually or unidirectionally) dependent [16].

The role of the CLO is to separate entities/concepts that belong to the general law theory from concepts specific to national legal systems or a specific legal domain. The CLO also aims to fill the gap between domain-specific concepts and abstract categories of formal top-level or foundational ontologies such as DOLCE [2].

2.3   ELTS

The ELTS (European Legal Taxonomy Syllabus) approach describes a tool that has been used to construct multilingual concept dictionaries, allowing the distinction between terms and concepts for the European Union (EU) legislation [3]. The European Union Guidelines (EUDs) constitute a set of legal standards that must be implemented by national legislation and translated into the language of each Member State. The problem of multilingualism in European legislation and the management of the EUDs are complex, since the implementation of a EUD does not correspond to a direct transposition of a law from a member country to another member country.

The ELTS includes different ontologies, one for each national language involved, for example Italian (Ita) and German (Ger), plus one for the European Union

(EU) document language. Each language-specific ontology is related by means of a set of links (association) with the concepts of the UE, as shown in Fig. 5.
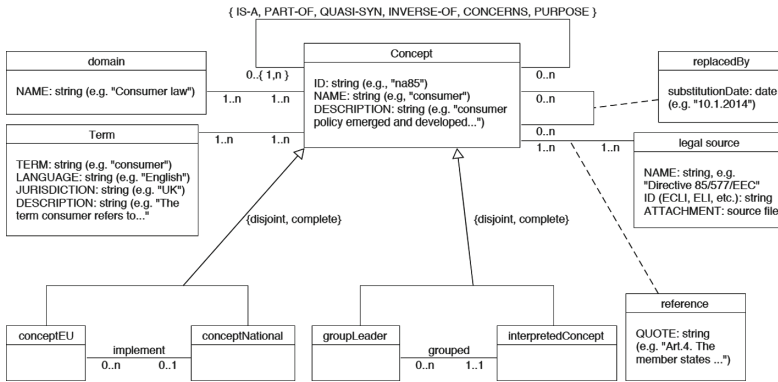

Fig. 5. ELTS ontology schema (UML language, source: [3])

This kind of ontology fits into the type of semantic approach since it deals with the meaning of EU terms coupled with the ontological approach in which it compares two domain-specific ontologies with EU semantic concepts.

2.4 Ontology Reference Model for Normative Acts

The ontological reference model for Normative Acts (NAs) is a modeling approach based on the *OntoUML* ontology [4]. The main objective of this reference model is to represent only the structural elements of NAs.
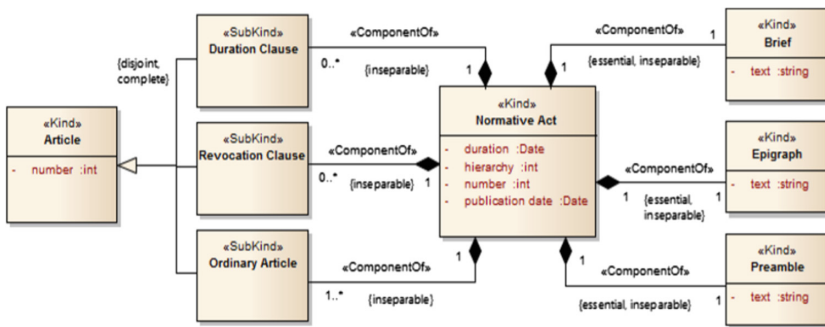

Fig. 6. Compositions of Normative Acts (source: [4])

Figure 6 shows the essential elements of this reference model. Articles may be *Ordinary Articles, Revocation Clauses* or *Duration Clauses*. *Ordinary articles* also known as regular articles are articles that affirm a new communication. *Revocation*

Clauses consists of articles that revoke other articles. The *Duration Clauses* are articles that affirm a validation time. Every NA must be composed of at least one *Ordinary Article*. All NA has preliminary mandatory elements *Preamble*, *Epigraph* and *Summary*. The *Preamble* is a term that means *introduction*, *initiation*, or *initial statement* of NA. It is a short text that predates the first chapter and sets out a brief explanation of the content discussed in NA. The *Epigraph* is a title or phrase that serves as a theme or subject introduction. The *Brief* is a collection of information, data collection most relevant to the development of an NA.

This type of approach is classified in the type of normative knowledge since the emphasis is the representation of elements and relations of a law.

2.5   Generic Model of Relationship (GMR)

The GMR (Generic Model of Relationship) approach aims to assist in organizing information with an emphasis on relationships between concepts and information units [17]. GMR consists of three main entities: (i) concept, (ii) information unit and (iii) relationship. Figure 7 illustrates the elementary class hierarchy of the GMR model.   AQ3
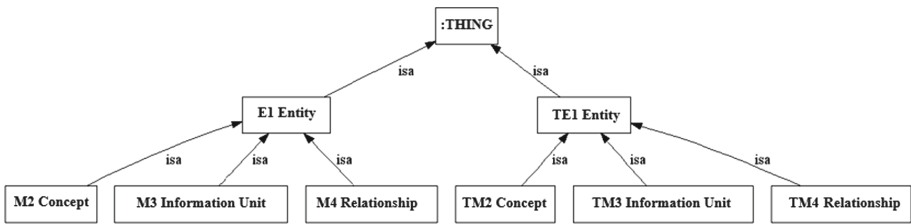

Fig. 7. GMR entities (source:)

The extension of GMR ontology applied in the field of Legal and Juridical Information (GMR-LJI) works on classes M3 Information Unit and M4 Relationship (Fig. 7). The extension applied in class M3 information Unit results in the creation of subclasses: (i) Periods; (ii) Events that start or end these Periods; (iii) Activities and their Agents (Roles) and Participants. Regarding the class M4 Relationship, it gives rise to the subclass Relationships between domain classes. Figure 8 presents the ontology details related to the Activity class (subclass iii of M3). In this model, we highlight the greatest contribution of the author including the legislative process. The model describes the dynamic component through temporal relationships between agent and their activities. However, for legal documents, the model limits to describe the temporal evolution of norms and dispositions, without existing structural or content aspects of legal documents.
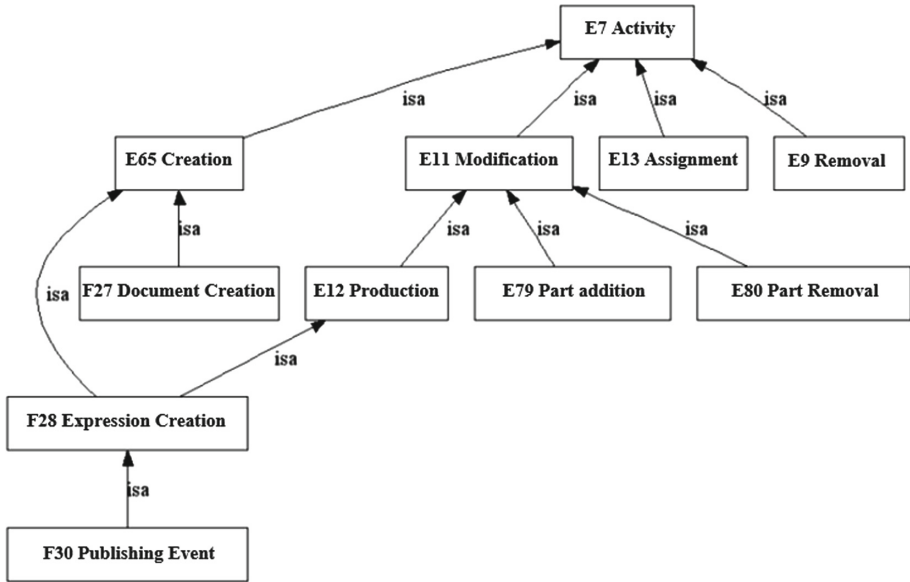
Fig. 8. GMR activities (source:)

2.6   New Developments on Law Content Aspects

The legal domain involves large amounts of concepts, terms and documents. Legal documents are rapidly changing or evolving. New developments in this area must take into consideration syntactic and semantic knowledge [18]. In the legal domain, different ontologies can be defined depending mostly on the task for which they are built.

Several projects involving ontologies dealt with topics such as legal decision support systems, translation and interpretation of legal texts. Sections 2.1 to 2.4 presented different ontologies manually developed from scratch (top-down approach). These approaches were created with a specific domain goal, without paying to legal ontology learning (bottom-up approach). Ontology learning aims to identify terms, concepts, relations or axioms to support ontology building [18].

In the last two decades, different proposals emerged based on Artificial Intelligent techniques to extract concepts and relations among them from unstructured legal texts. In 2005, Lame [19] applied Natural Language Processing (NLP) techniques to create an ontology of French Law dedicated to information retrieval. Saia and Quaresma [20] proposed a methodology to automatically create legal ontologies that allow enriching legal documents. Mezghanni and Gargouri [21] created an approach for ontology learning from Tunisian Legal texts designed for legal information retrieval. Ghosh et al. [22] presented a semiautomatic ontology construction technique (1) reusing existing ontologies to extract similar and complementary information; and (2) capturing relevant legal concepts and relations from textual sources using NLP techniques. Hwang et al. [23] presented a technique for an automatic ontology construction from a structured text (databases). This approach involves NLP and data mining techniques for concept and relationship extraction.

The extraction of rules or conditions from legal texts is a difficult task, so the adoption of a single Natural Language Processing (NLP) approach would not lead to satisfiable results. Considering the limited adoption of current frameworks, Dragoni et al. [24] adopted and combined a set of NLP techniques towards the extraction of rules from legal documents.

In 2019, Fawei et al. [25] presented a methodology that leads to the creation of a legal ontology and a corresponding set of rules. This was the first fine-grained methodology for constructing legal OWL ontologies with Semantic Web Rule Language (SWRL) rules.

Design Engineering and Modelling for Organizations (DEMO) is a theory and method to model interactions between individuals and/or organizations(set of actors) based on a communication-centric approach (social interactions) [26]. The Performance in Social Interactions (PSI) is a component of DEMO that declares how the coordination-acts performed by actors are represented in patterns called transactions. Gouveia and Aveiro [27] present a design of two sources of law based on the transaction axiom of the PSI theory. The authors also analyze the assumptions mismatch between law and DEMO/PSI. In the studied cases, similarities were observed between the main concepts of the previously mentioned ontologies and the essential concepts of DEMO/PSI.

2.7   Discussion

Table 1 summarizes the comparison of the ontologies introduced above. The compared features allow characterizing these ontologies. In addition, it also allows identifying some limitations of these proposals and justifies the motivation for the design of the proposed LegalLanguage.

The LegalLanguage results from the analysis of these ontologies but have a distinct focus: while some concepts are similar, there is a different interest and application domain. Although, the ontological reference model for normative acts proposed by Pedro et al. [4] is also related to the structure of legal documents, it is less flexible since the model details the hierarchical and rigid structure of a NA. This conceptual model was constructed for communication and learning purposes without considering mechanisms to support the writing of normative texts of laws and intra and inter-laws relationships.

In an initial stage of this research work, based on the fact that several ontologies focus semantics of the law, we decide to focus in the structure of a law document. The addition of semantic description will be considered in future work.

3   LegalLanguage Overview

This section overviews the proposed metamodel and the mechanism involved in the construction of law documents. For clarity purposes, Sect. 3.1 presents the LegalLanguage metamodel and provides textual definitions of the elements. Section 3.2 describes the DLS implementation with Xtext [28].

3.1   LegalLanguage Meta-model

As seen in Sect. 2 a legal ontology is an explicit way to represent Laws as rigorous models, which capture and represent common concepts of a generic Law. Figures 9 and 10 present the meta-model (i.e., the abstract syntax as UML class diagrams) of the proposed LegalLanguage. Compared with the ontological reference model for normative acts, the proposed meta-model is more flexible since includes enumerations that only affect the semantics of the respective meta-classes.

Table 1. Comparative analysis of ontologies.

Ontologies	LKIF	CLO	ELTS	ORM for NAs	MGR
Features					
Ontological reference models	N	Y	Y	Y (UML class diag)	Y
Ontological domain of NAs	Europe	Europe	EU	Brazil	Brazil
Concept of					
Law	Y(docs)	Y	Y	Y (NAs)	Y
Legal terms	N	Y	Y	N	N
Legal concepts	Y	Y	Y	N	N
Legal norms	N	Y	Y	NE	Y
Legal texts	Y	Y	Y	N	N
Legal facts	NE	Y	N	N	N
Normative context	N	Y	N	Y	N
Legal situation	N	Y	N	N	Y
Process	Y	Y	N	N	N
Action	Y	NE	NE	N	N
Legal-action	Y	NE	NE	N	N
Role	Y	NE	Y	N	N
Legal-role	Y	NE	Y	N	N
Rules	Y	NE	Y	N	Y
Agent	Y	Y	Y	N	Y
Person	Y	Y	Y	N	Y
Legal-source	Y	Y	Y	N	N
Change	Y	Y	N	N	N
Qualified	NE	Y	NE	N	Y
Approach					N
Semantically oriented	Y	N	Y	N	N
Epistemically oriented					
Normative knowledge	N	N	N	Y	N
Meta-legal knowledge	N	N	N	N	N
World knowledge	N	N	N	N	N
Responsibility knowledge	N	N	N	N	N
Reactive knowledge	N	N	N	N	N
Creative knowledge	N	N	N	N	N
Ontologically oriented	N	Y	N	N	Y

Subtitle: Y (Yes); N (No); NE (Not explicit)

According to this metamodel, the **Law** is composed of several Articles ordered sequentially, which can be further structured in Divisions. The **Law** is the general concept to represent "normative texts" and is classified by a *type* (e.g., Constitution, International Law, Ordinary Law, Public Regulation, Private Regulation), shall be identified by a *unique id* (number) and defined by relevant *dates* (e.g., publication, activation, disactivation). Optionally a Law can be further classified by a *sub*type, for example: Constitution: original or revision; International law: treatise, regulation, directive; Ordinary law: Law, Decree-law, national or regional; Regulation: regulation, contract, tort. It is also possible to define the application domain of the law (e.g., agricultural, aviation, banking, public, civil), and also the possibility to define the current state in which the law can be found (e.g., in edition, submitted to approve, approved, active, suspended, revoked).
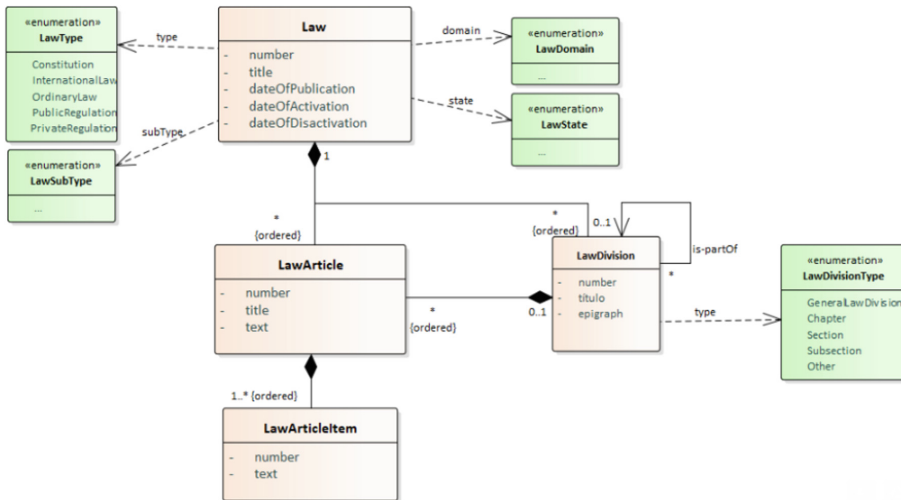

Fig. 9. The meta-model of structural concepts of Law

The **Article** is the basic element of a law, shall be univocally identified by a sequential *number* (e.g., 1, 2, 3), and shall have an optional *title*, a *text* and also a set of *items*.

The **Law** can be structured by a set of **Divisions**, whose properties are *epigraph*, *number*, *type* of division (e.g., Chapter, Section, Sub-section) and *title*. The **Division** may be composed of other divisions, defining hierarchies if relevant. Furthermore, an Article can be assigned to a Division, and on the other hand, a Division may aggregate several Articles.

One may also consider relationships between laws and hierarchies (e.g., order of importance) for the law of laws that assume its subordination to a higher origin. For

example, at a country level, laws are hierarchically inferior to the Constitution law. The laws (e.g., ordinary, delegated and complementary) and the regulatory decree, which aims to regulate the provision of the law, follow the same guidelines. Another common relationship between laws (or specific articles of laws) is the revocation: when a new law appears, it may revoke a previous old law, and then the former becomes active and the latter revoked.

Figure 10 shows the LegalLanguage support for the definition of **relations between laws** (inter-law relations) as well as **relations between articles** of the same law (articles intra-law relations) or even from distinct laws (articles inter-law relations). In addition, these relations are typified, which means we may add some semantic to them (e.g., revokes, revoked, depends, specializes).
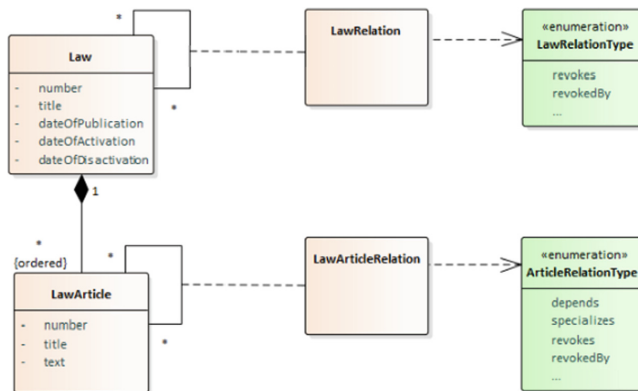

Fig. 10. The meta-model with relations between laws and between articles (same or different laws)

3.2 LegalLanguage Implementation Aspects

The development of a Domain-specific Language (DSL) requires to be aware of two aspects: the characteristics of a DSL and details of the domain to which the language will be contextualized. The particular domain considered is the legal domain, in particular, the scope of parliamentary with activities of the legislative process of drafting laws. To implement the proposed LegalLanguage, we considered the meta-model presented in Figs. 9 and 10.

The Xtext [28], developed as part of the Eclipse Modeling Framework, generates a parser, a serializer, a meta-model and a DSL editor from the grammar of the language.

The code snippet Spec. 1 shows the definition of the LegaLanguage's Law grammar:

```
Law:      'Law' name=ID  (nameAlias=STRING)? ':' type=LawType (':' subType=LawSubType)? ('['
                    ('domain' subject=LawDomainType)?
                    ('state' state=LawStateType)?
                    (relationType+= LawRelationType relations+= RefLaw)*
                    ('number' number=INT '/' year=INT)?
                    ('title' title=STRING)?
                    ('dateOfPublication' dateOfPublication=Date)?
                    ('dateOfActivation' dateOfActivation=Date)?
                    ('dateOfDisactivation' dateOfDisactivation=Date)?
                    ('description' description=STRING)?
                    (lawArticles+=LawArticle*)
                    (lawDivisions+=LawDivision*)
           ']')? ;
enum LawType: Constitution | InternationalLaw | OrdinaryLaw |  PublicRegulation | PrivateRegulation |…;
enum LawRelationType: revokes | revokedBy;
```

Spec. 1. LegaLanguage's Law grammar

The code snippets Specs. 2 and 3 show the definition in Xtext of the LegaLan-guage's Division and Article grammars:

```
LawDivision:
           'Division' name=ID  (nameAlias=STRING)? ':' type=LawDivisionType  ('['
                    ('partOf' partOf=[LawDivision | QualifiedName] )?
                    ('epigraph' epigraph=STRING)?
                    ('number' number=INT)?
                    ('title' title=STRING)?
                    (divisionArticles+=LawArticle*)
           ']')?;
enum LawDivisionType: Chapter | Section | Subsection | Other;
```

Spec. 2. LegaLanguage's Division grammar

```
LawArticle:
           'Article' name=ID  (nameAlias=STRING)?
           ('['
                    (relationType+= ArticleRelationType relations+= RefArticle)*
                    ('number' number=INT)
                    ('title' title=STRING)?
                    ('text' text=STRING)?
                    (itens+=LawArticleItem*)
           ']');
...
enum ArticleRelationType:       depends | specializes | revokes | revokedBy;
```

Spec. 3. LegaLanguage's Article grammar

The Eclipse IDE also includes the LegalLanguage's plug-ins. An editor of laws will be automatically available to assist in the writing of normative texts like laws, and intra or inter-laws relationships. Currently, it is possible to create documents similar to the example presented in the following session.

4 Illustrative Example

The legislative process is crucial in the drafting of a law, it includes a set of phases and acts duly ordered and executed, whose content, form and sequence follow a series of rules specific to each parliament.

```
Package EU_Laws
Law Directive_EC_46_95 "Directive n.º 95/46": InternationalLaw: InternationalLaw_Directive [
          domain IT
          state Revoked
          revokedBy EU_Laws.Regulation_EU_679_2016
          number 95/46
          dateOfPublication 24-Oct-1995
description "DIRECTIVE 95/46/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of
24 October 1995 on the protection of individuals with regard to the processing of personal data and on the
free movement of such data. THE EUROPEAN PARLIAMENT AND THE COUNCIL OF THE
EUROPEAN UNION, ..."
```

Spec. 4. Illustrative example. Law header specification in LegaLanguage

Legislative drafting is a complex process, due to the number of copies that are analyzed by the plenary in each ordinary session, and that undergoes constant changes during the phases of the legislative process until the final drafting of the Law. It involves a huge volume of documents in physical format (paper) and many stakeholders.

Spec. 4 illustrates the LegaLanguage specification to instantiate a practical example of the "Directive 95/46/EC" related to the protection of individuals with regard to the processing of personal data and on the free movement of such data.

Spec. 5 illustrates the specification of the article-3 of the Directive 95/46/EC in LegalLanguage. In particular, it is relevant to stress that this article specializes the article-6 (defined in Sect. 1), and also is revoked by article-15 defined in the former EU Regulation 679/2016. The "revokedBy" is an example of an articles inter-law relation, while the "specializes" is an example of an articles intra-law relation.

```
Article Article_3 "Article 3" [
          revokedBy EU_Laws.Regulation_EU_679_2016.SECTION_II.Article_15
          specializes SECTION_I.Article_6
          number 3
          title "Scope"
          ArticleItem Name "..." [ number 1  text "...." ]
]
```

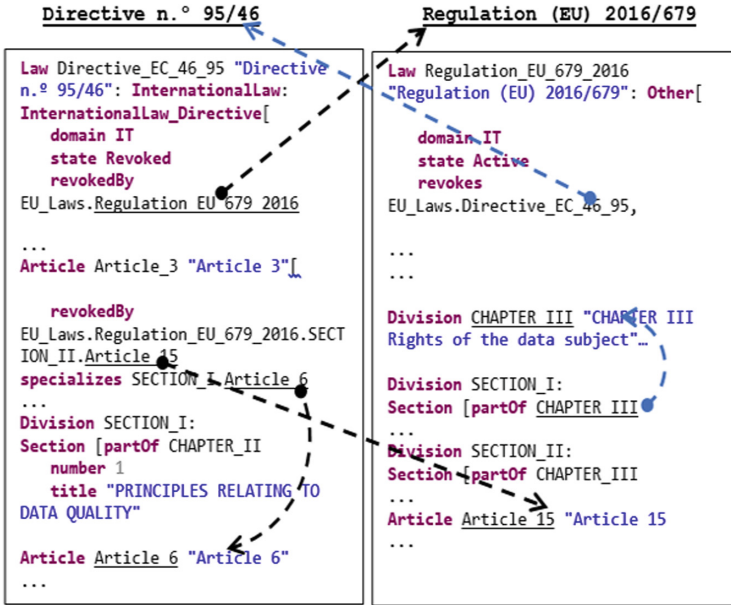**Spec. 5.** Illustrative example. Article specification in LegaLanguage

Fig. 11. Relationships between two types of laws

As suggested in Fig. 11, the authors also defined relations between several types of laws, in which each law relates to other(s), also allowing to link articles. Figure 11 illustrates an example where two types of the European Union legislation have relations, namely: Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 and Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995. The Regulation (EU) 2016/679 revokes the Directive 95/46/EC (General Data Protection Regulation). On the opposite site, Directive 95/46/EC is revoked by the Regulation (EU) 2016/679. In this example, the relationship between articles is also represented.

5   Conclusion

This paper proposes a DSL for the rigorous specification of normative documents, in particular laws produced in the parliamentary scope. This DSL, named as LegalLanguage, shall support the activities related to the elaboration of laws.

The comparative study of different legal ontologies made it possible to note that some problems, mentioned in Sect. 1, in particular, the activities of laws' elaboration continue prone to several error types. The proposed language allows to represent laws in a more rigorous and explicit form. This language allows the definition of structural patterns, like chapters, articles, sections and subsections, typically found in legal documents.

Although there are some tools to assist in the authoring process of legal ontologies, the selection of a framework for the development of DSLs was the most feasible at the time the problems in this specific-domain were identified. The choice of the development tool required a search for support infrastructures that facilitated an agile and iterative evolution of the domain-specific language, as well as essential resources for its construction. Using the LegalLanguage editor, the normative drafting activities can be improved and less error-prone compared to the actual manual process.

The illustrative examples allow to show the applicability of LegalLanguage in the current drafting of laws. Regarding the main objective of this paper, the LegalLanguage application in the legislative production process was satisfactory, in the sense that it helps in laws elaboration activities.

In the example presented, the LegalLanguage allowed to identify opportunities for future works, such as extend the language to address the remaining identified problems. For example, the realization of empirical studies that will make possible to include improvements on the LegalLanguage. Nowadays, several gaps persist between language and technology, our vision is to establishing a simplified way of working among the various actors in the legislative process. LegalLanguage will act as a facilitator during the writing phase of normative texts (e.g., laws, regulations) of legal-based organization.

Future evaluations will also deliver further input for improving LegalLanguage and eventually to create a supporting tool. Future research will also have to consider applying LegalLanguage in the validation of legal documents in different formats, such has words documents.

References

1. Hoekstra, R., Breuker, J., Di Bello, M., Boer, A.: The LKIF core ontology of basic legal concepts. In: Proceedings of LO A IT 07: II Workshop on Legal Ontologies and Artificial Intelligence Techniques, pp. 43–63 (2007)
2. Francesconi, E., Tiscornia, D.: Building semantic resources for legislative drafting: the DALOS project. In: Casanovas, P., Sartor, G., Casellas, N., Rubino, R. (eds.) Computable Models of the Law. LNCS (LNAI), vol. 4884, pp. 56–70. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-85569-9_4
3. Ajani, G., et al.: European legal taxonomy syllabus: a multi-lingual, multi-level ontology framework to untangle the web of European legal terminology, vol. 3. IOS Press (2006)
4. Pedro, B.P.F., Guizzardi, R.S.S., Garcia, A.S.: An Ontology Reference Model for Normative Acts, Brazil (2013)
5. Caramujo, J., et al.: RSL-IL4Privacy: a domain-specific language for the specification of privacy-aware requirements. Requirements Eng. **24**(1), 1–26 (2019)
6. Ribeiro, A., Silva, A.R.: RSLingo4Privacy studio: a tool to improve the specification and analysis of privacy policie. In: Proceedings of ICEIS 2017 (2017)
7. da Silva, A.R., Caramujo, J., Monfared, S., Calado, P., Breaux, T.: Improving the specification and analysis of privacy policies: the RSLingo4Privacy approach. In: Proceedings of ICEIS 2016 (2016)

8. da Silva, A.R.: Linguistic patterns and linguistic styles for requirements specification (i): an application case with the rigorous RSL/business-level language. In: Proceedings of EuroPLOP 2017 (2017)

9. da Silva, A.R., Paiva, A.C.R., da Silva, V.E.R.: A test specification language for information systems based on data entities, use cases and state machines. In: Hammoudi, S., Pires, L.F., Selic, B. (eds.) MODELSWARD 2018. CCIS, vol. 991, pp. 455–474. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-11030-7_20

10. Gonçalves, L.P., da Silva, A.R.: Towards a catalogue of reusable security requirements, risks and vulnerabilities. In: Proceedings of ISD 2018 (2018)

11. Soares, A.A., Martins, P.V., da Silva, A.R.: A systematic literature review of legal ontologies. In: Proceedings of CAPSI 2018 (2018)

12. Mommers, L.: Ontologies in the legal domain. In: Poli, R., Seibt, J. (eds.) Theory and Applications of Ontology: Philosophical Perspectives, pp. 265–276. Springer, Heidelberg (2010). https://doi.org/10.1007/978-90-481-8845-1_12

13. Gruber, T.: A translation approach to portable ontologies. Knowledge Acquisition (1993)

14. Gruber, T.: Toward principles for the design of ontologies used for knowledge sharing. Int. J. Hum.-Comput. Stud. **43**, 907–928 (1995)

15. Rubino, R., Rotolo, A., Sartor, G.: An OWL ontology of fundamental legal. In: Legal Knowledge and Information Systems. JURIX 2006: The Nineteenth Annual Conference. Frontiers, vol. 152 (2006)

16. Gangemi, A., Sagri, M.-T., Tiscornia, D.: A constructive framework for legal ontologies. In: Benjamins, V.R., Casanovas, P., Breuker, J., Gangemi, A. (eds.) Law and the Semantic Web. LNCS (LNAI), vol. 3369, pp. 97–124. Springer, Heidelberg (2005). https://doi.org/10.1007/978-3-540-32253-5_7

17. Lima, J.: Modelo Genérico de Relacionamentos na Organização da Informação Legislativa e Jurídica, Brasilia (2008)

18. Mezghanni, I.B., Gargouri, F.: Towards an Arabic legal ontology based on documents properties extraction. In: 12th International Conference of Computer Systems and Applications (AICCSA), Marrakech, Morocco (2015)

19. Lame, G.: Using NLP techniques to identify legal ontology components: concepts and relations. In: Benjamins, V.Richard, Casanovas, P., Breuker, J., Gangemi, A. (eds.) Law and the Semantic Web. LNCS (LNAI), vol. 3369, pp. 169–184. Springer, Heidelberg (2005). https://doi.org/10.1007/978-3-540-32253-5_11

20. Saias, J., Quaresma, P.: A methodology to create legal ontologies in a logic programming information retrieval system. In: Benjamins, V.R., Casanovas, P., Breuker, J., Gangemi, A. (eds.) Law and the Semantic Web. LNCS (LNAI), vol. 3369, pp. 185–200. Springer, Heidelberg (2005). https://doi.org/10.1007/978-3-540-32253-5_12

21. Mezghanni, I.B., Gargouri, F.: Learning of legal ontology supporting the user queries satisfaction. In: 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT), Warsaw, Poland (2014)

22. El Ghosh, M., Naja, H., Abdulrab, H., Khalil, M.: A ontology learning process as a bottom-up strategy for building domain-specific ontology from legal texts. In: 9th International Conference on Agents and Artificial Intelligence (ICAART 2017), New York (2017)

23. Hwang, R.-H., Hsueh, Y.-L., Chang, Y.-T.: Building a Taiwan law ontology based on automatic legal definition extraction. Appl. Syst. Innov. **1**(3), 22 (2018)

24. Dragoni, M., Villata, S., Rizzi, W., Governatori, G.: Combining natural language processing approaches for rule extraction from legal documents. In: Pagallo, U., Palmirani, M., Casanovas, P., Sartor, G., Villata, S. (eds.) AICOL 2015-2017. LNCS (LNAI), vol. 10791, pp. 287–300. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00178-0_19

AQ4

25. Fawei, B., Pan, J.Z., Kollingbaum, M., Wyner, A.Z.: A Semi–automated ontology construction for legal question answering. New Gen. Comput. **37**, 453–478 (2019)
26. Dietz, J.L.G.: Enterprise Ontology: Theory and Methodology. Springer, Berlin (2006). https://doi.org/10.1007/3-540-33149-2
27. Gouveia, D., Aveiro, D.: DEMO/PSI theory and the law of the land. In: Aveiro, D., Pergl, R., Guizzardi, G., Almeida, J.P., Magalhães, R., Lekkerkerk, H. (eds.) EEWC 2017. LNBIP, vol. 284, pp. 50–65. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-57955-9_4
28. Bettini, L.: Implementing Domain-Specific Languages with Xtext and Xtend, Second edn. Packt Publishing Ltd., Birmingham (2016)

Author Query Form

Book ID : **481324_1_En**

Chapter No :   **3**

Please ensure you fill out your response to the queries raised below and return this form along with your corrections.

Dear Author,

During the process of typesetting your chapter, the following queries have arisen. Please check your typeset proof carefully against the queries listed below and mark the necessary changes either directly on the proof/online grid or in the 'Author's response' area provided below

Query Refs.	Details Required	Author's Response
AQ1	This is to inform you that corresponding author has been identified as per the information available in the Copyright form.	
AQ2	Per Springer style, both city and country names must be present in the affiliations. Accordingly, we have inserted the city names in affiliations. Please check and confirm if the inserted city names are correct. If not, please provide us with the correct city names.	
AQ3	Kindly note that captions text of Figs. 7 and 8 seems to be incomplete. Please check and confirm.	
AQ4	Please check and confirm the updated Ref. [24] is okay.	