A new Usage for Semantic Technologies for eGovernment: Checking Official Documents' Consistency

Fred Freitas[1,2], Zacharias Candeias Jr[3] and Heiner Stuckenschmidt[2]
[1]Federal Universidade of Pernambuco, Cidade Universitária, Brazil
[2]University of Mannheim, Germany
[3]Information Technology Agency for the State of Pernambuco, Brazil
fred@cin.ufpe.br
zacharias.candeias@ati.pe.gov.br
heiner@informatik.uni-mannheim.de

Abstract. Semantic technologies, and particularly the ones related to the Semantic Web and its ontologies, have proven useful for many government related applications and prototypes, such as service configuration, automatic service connection among many others. This is possible because the Semantic Web is based on ontologies, which, in practical words, stands for a detailed conceptualization of a domain and its concepts, relations, constraints and axioms, defined in an unambiguous manner using formal logic. On the other hand, official documents, and particularly legal ones like law codes, often contain semantic deficiencies that are not realized by their authors. The most common among them are ambiguities, inconsistencies and under specifications. These deficiencies are certainly a source of systems' and databases' integration problems and confusion during their usage, when the definitions' intended meanings can differ depending upon the stakeholder. During the ontology development of a domain as simple as vehicles, we have witnessed such phenomena. The necessity of defining the different vehicle types in detail for classification and checking purposes shed light on some of these deficiencies present in two Brazilian legal codes. In this work, we present the building process of the ontology, the resulting ontology and show how these deficiencies were evidenced during its construction. This fact actually opens up new possibilities in the usage of semantic technologies, as guides to check whether official documents are ontologically and logically correct, by not containing ambiguities, under specification or inconsistencies.

Keywords: ontology-based analysis of texts, semantic deficiencies, vehicles, eGovernment, law, ontology engineering, law consistency, official documents' consistency

1. Introduction

It is quite common, even routine, that people have different accounts about a text they had read or even a dialogue they had had. The ambiguities, inconsistencies and other semantic problems arise in readings or talks due to a number of reasons, which are studied in depth by many different branches of knowledge, ranging from human sciences like Philosophy and Linguistics to exact sciences like Mathematical Logics and Artificial Intelligence (Jurafsky and Martin, 2000). The main reason of these misinterpretations on human communication, however, seems to lie on the very nature of natural languages: natural languages emerged as a result of a long – indeed, never lasting - constructive communication process. The process *per se* is rather a solution than a problem, but the different interpretations of terms, phrases, intentions and meanings conveyed in written or spoken communication do cause misunderstandings. in natural languages two interlocutors expect to share a same interpretation about single phrases and entire texts, without relying on any sort of formal semantics, understood by both of them. Their expectations, unfortunately, are not always fulfilled.

To avoid the same problem to happen in a higher scale in computer communication, semantic technologies, and particularly the ones related to the Semantic Web and its ontologies (Berners-Lee et al, 2001), have proven useful for many government related applications and prototypes, such as service configuration, automatic service connection among many others. This is possible because the Semantic Web is based on ontologies, which, in practical words, stands for a detailed conceptualization of a domain and its concepts, relations, constraints and axioms. These specifications must be defined in an unambiguous manner using formal logic. On the other hand, official documents, and particularly legal ones like law codes, often contain semantic deficiencies that are not realized by their authors. The most common among them are ambiguities and inconsistencies ought to linguistic problems like polysemy (i.e., one word with multiple senses), as well as under specifications and inconsistencies. These deficiencies are certainly a source of integration problems and confusion during their usage, when the intended meanings can differ depending upon the stakeholder. During the ontology development of a domain as simple as vehicles, we have witnessed

Reference this paper as: Freitas, F, Candeias Jr, Z and Stuckenschmidt, H. "A new Usage for Semantic Technologies for eGovernment: Checking Official Documents' Consistency" *Electronic Journal of e-Government* Volume 8 Issue 2 2010, (pp121-134), available online at www.ejeg.com

such phenomena. The necessity of defining the different vehicle types in detail for classification and checking purposes shed light on some of these deficiencies present in two Brazilian legal codes.

The law codes are responsible for defining vehicles' categories in an unambiguous manner for many purposes, e.g. tax calculations. In this work, we show these deficiencies and show how they were identified during ontology construction, describing this process and its resulting ontology in detail. This fact actually opens up new possibilities in the usage of semantic technologies, as guides to check whether official documents are ontologically and logically correct, by not containing ambiguities, under specifications or inconsistencies.

The article is organized as follows: section 2 describes semantic technologies and the Semantic Web techniques very briefly; section 3 presents the process of ontology construction; section 4 discusses the domain of vehicles and the two Brazilian official documents (legal codes) used to construct the ontology; section 5 displays some of the deficiencies found out throughout the process; section 6 discusses usages of legal ontologies; section 7 envisions what types of checking could be tested against official documents as well how these could be carried out in methodical fashion, together with other envisioned applications; section 8 brings related work; and section 9, conclusions.

2. Semantic Web and eGovernment

According to the words of Tim Berners-Lee (2001), one of the responsible for the establishment of the World Wide Web, ontologies are the central components and motivation of the Semantic Web, a Web in which the software "understands" and processes data from Web pages, according to the context surrounding these pages, the meaning of the jargon used on them, etc.

How can ontologies provide context for Web pages? In practical words, ontologies comprise definitions of concepts, properties, relations, constraints, axioms, processes and events about a certain domain or universe of discourse. If a Web page instantiate a body of definitions about the domain that this page refers to, the software or agent which handles the page can make use of a precise, clear, formal semantics that defined *a priori* knowledge related to the pieces of information present in the page. This trend of information processing is already provoking changing over eGovernment processes over the Internet. Due to the lack of ambiguity, for example, semantic technologies can be employed in information integration and system interoperability of government services which can be misinterpretations free, particularly if the description logic OWL (Ontology Web Language) (Patel-Schneider et al, 2004), which is recommended by the W3C as one of the standards for the Semantic Web, is used for defining the terminology employed in the specific domain. A practical example should make things clearer and is presented below.

Comte and Léclére (2006) report on a semantic reasoning framework for the French social welfare eGovernment system. The problem for the French was providing a unified vocabulary and its corresponding meanings to which the welfare systems should comply to, as a first step towards shared interoperability among eGovernment systems. Thus, an ontology of social welfare services for family household service professions was one of the resources defined for this goal. This ontology contains a definition of what is a house hold professional, in the following form:

$$\begin{aligned}\text{HouseHoldProfessional} \equiv\ & \text{Person} \sqcap \text{hasValue(lives,?x)} \sqcap \\ & \exists\ \text{workFor.}\ (\ \text{Person} \sqcap \text{hasValue(lives,?x)}\ ) \sqcap \\ & \forall\ \text{workFor.}\ (\ \text{Person} \sqcap \text{hasValue(lives,?x)}\ )\end{aligned} \quad (1)$$

This definition means that such a professional is:

- A person (i.e., can only be filled with instances of the class Person), and (meant by the symbol $\sqcap$)
- Who lives somewhere, in a certain address x - this is defined in the predicate hasValue(lives,?x) -, and
- Who works for *some* (meant by the symbol $\exists$) person who lives in the same address x and
- Who works *only* (meant by the symbol $\forall$) for persons who live in the same address as she (?x).

Note that this definition is unambiguous, as it does not give rise to any false interpretation of household in this domain, and is treated by systems as so, if they are endowed with a description logic reasoner such as RACER (Haarslev and Moller, 2001). If an object or database instance fulfills this definition, it is automatically classified by the reasoner as an instance of the class HouseHoldProfessional. This will enable indexing electronic resources and representing the content of document-based resources in a semantic annotation base, according to the semantic web jargon. Hence, a query to this knowledge base will allow finding relevant information from different media, and a way to access it.

But this is not the only facet where ontologies promise to change drastically the current reality of information handling for eGovernment. Other areas such as knowledge management, tutoring and geographic systems, among many other application areas for governments, are also starting to take advantage of ontologies to improve their systems' abilities.

In the next section, we present the construction process for an ontology of Vehicles, which was designed according to various Brazilian law codes, for the purposes of document exchange, tax calculations, vehicle identification and classificattion, among others.

3. Building a vehicles' ontology

In this section, we will first enumerate the knowledge sources used for the initial sub-process of knowledge acquisition and, in the second sub-section, the construction process itself, encompassing all the steps of the Methontolgy methodology (Gómez-Pérez et al, 2004).

3.1 Knowledge sources

Although the State Government of Pernambuco is one of the states that offers most advanced eGovernment online services in Brazil through its portal (Governo de Pernambuco, 2010), such as touristic information, crimes' and vehicles' registry, only now Semantic Web technologies are being used, with a twofold purpose: First, to ensure that the interoperability among database schemata and the document exchange between systems will take place without risk of misinterpretations or mistakes. The second reason lies on the reuse possibilities: most laws in Brazil are federal; therefore, the definitions generated in this effort are likely to become standard and can be reused in many other states of the country. It can even be reused in other contexts, such as private organizations that build applications compliant to the current laws. It was written in OWL, since we will need to classify instances of vehicles for taxes assignment, for example. The construction of this artefact was built in conformance with the following sources:

- Brazilian Traffic (law) Code (CTB) (Ministério das Cidades, 2008) that states the Brazilian traffic laws in general terms, so that specific details have to be defined elsewhere;
- Resolutions and deliberations from Contran (Conselho Nacional de Trânsito, 2010, the National Traffic Code) which make for the details missing in CTB;
- Sites of the State and Federal Traffic Departments – in Brazil, Traffic Departments are the public institutions empowered to enforce traffic laws,
- Many existing database schemas and systems' manuals, from the State Traffic Department of Pernambuco (Governo de Pernambuco. DETRAN-PE, 2010),
- Interviews with analysts responsible for the systems and databases.

Next subsection describes the process of ontology construction in detail.

3.2 Construction process

The Methontology methodology (Gómez-Pérez et al, 2004) was followed very closely during the Vehicles' ontology construction, such that all documents required by it were produced throughout the ontology construction process, as can be seen in the following. The Methontology methodology is composed by 12 steps, which can be seen in figure 1 (Blazquez et al 98).

First, we built a *Glossary of Terms* that included all the terms (concepts, instances, attributes, verbs, etc.) of the domain and their description, as partly shown in Table 1.

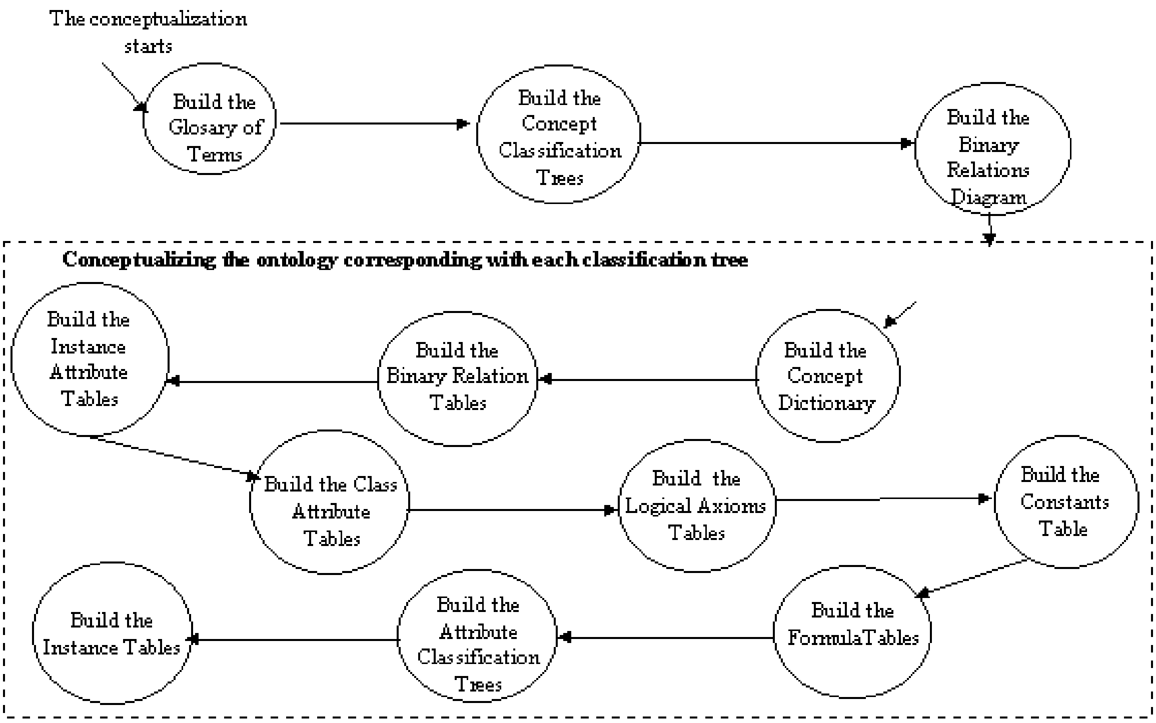

Figure 1: The knowledge structuring process of the Methontology methodology (Blazquez et al 98)

Table 1: A small fraction of the *Glossary of Terms* for the vehicles' ontology

Name	Synonyms	Acronyms	Description	Type
Year			Year in which the vehicle was produced	Attribute
Uses Fuel			Type of fuel used by the vehicle	Relation
Automobile			Automotor vehicle with more than three wheels (usually four) that transport up to nine people	Concept
Engine			Componente de Veículo que gera a impulsão em veículos automotores	Concept
Passenger Position			The way how the passenger is transported in the vehicle	Concept
Seated			A way how the passenger is transported	Instance (or Individual) of the class Passenger Position
Internal Combustion Engine	Explosion engine[1]		Engine fed with inflammable fuel	Concept

Departing from this glossary, we built up the five *Concept Classification Trees* that divided the 102 classes of the ontology: Vehicle, Person, Vehicle Function, Vehicle Component and Propulsion Type. These classes and trees were necessary and sufficient for the comprehensive definition of vehicles, according to the Brazilian law. For this domain, the relations displayed in these trees are only exhaustive, since new vehicle categories have to be included in the codes. There are, therefore, only two possible classification types: partitions – in which the compound classes can share individuals - and mutually disjoint partitions. An example of the former is portrayed in Figure 2.

[1] This sense is only used in Portuguese (motor a explosão)

In the next step, the definition of the *Binary Relations Diagram,* the most relevant relationships between the main class Vehicle and its main features (such as propulsion type, driver position[2], driving control device, function, owner type, etc) and components, were defined, as displayed in Figure 3. The specialization of these features and specific constraints on vehicles involving other classes accurately defines the different vehicle types that constitute the ontology backbone.

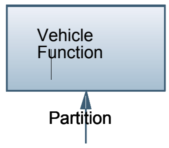

Figure 2: An example of a concept classification tree from the vehicle function class of the vehicles' ontology

After this initial phase, we started the conceptualization itself, which is composed of *'intermediate representations'*, according to the Methontology jargon (Blasquez et al, 98):

- The *Concept Dictionary*, with each concept and its attributes (i.e., relations with a basic type, such as a number or string) and relations with other classes (see a small part of it at Table 2). There are, in the ontology, 39 primitive classes (i.e., which are not defined by axioms) and 63 defined classes (in terms of the axioms).
- The *table of binary relations*, which details each relation with its domain and range concepts, as well as its mathematical properties and inverse relation, if any (see a small part of the binary relations is at Table 3).
- *Class Attribute Tables,* which describe each attribute domain, type, measure unit, precision, possible values and cardinality (i.e., how many possible instances are possible for each attribute relation). Examples of class attributes are shown in Table 4.
- *Instance attribute table.* We had only one instance attribute, which is the minimum age for driving (18 years).
- A *Logical Axioms Table,* which defines the terminology in description logic - see example on Table 5. Note that, although tricky, in this example most of the conditions stated are meant to differentiate cars from other vehicles: the position of passengers is seated (so tricycles are not cars, even if the latter have only three wheels), it can carry up to 9 passengers (this is in the law code, intended not to confuse with buses or vans), and containing neither rear cargo areas (like pickup trucks) nor fifth wheels (like tractor trucks).
- An *Instance Table*, which contains the instances necessary for defining the concepts and axioms properly (Table 6). Note that we had to define categories of speed and displacements, instead of working with continuous values, once the description logic reasoners are not able to compute with real numbers. It can just count instances, like the two or three wheels; therefore, we had to create rather artificial instances like UpTo50cc or MoreThan3500kg.

In the next section we present the resulting Vehicles' ontology.

[2] This information piece is valuable, for instance, to distinguish between motorcycles, in which drivers go mounted, and motor scooters, in which drivers go seated.

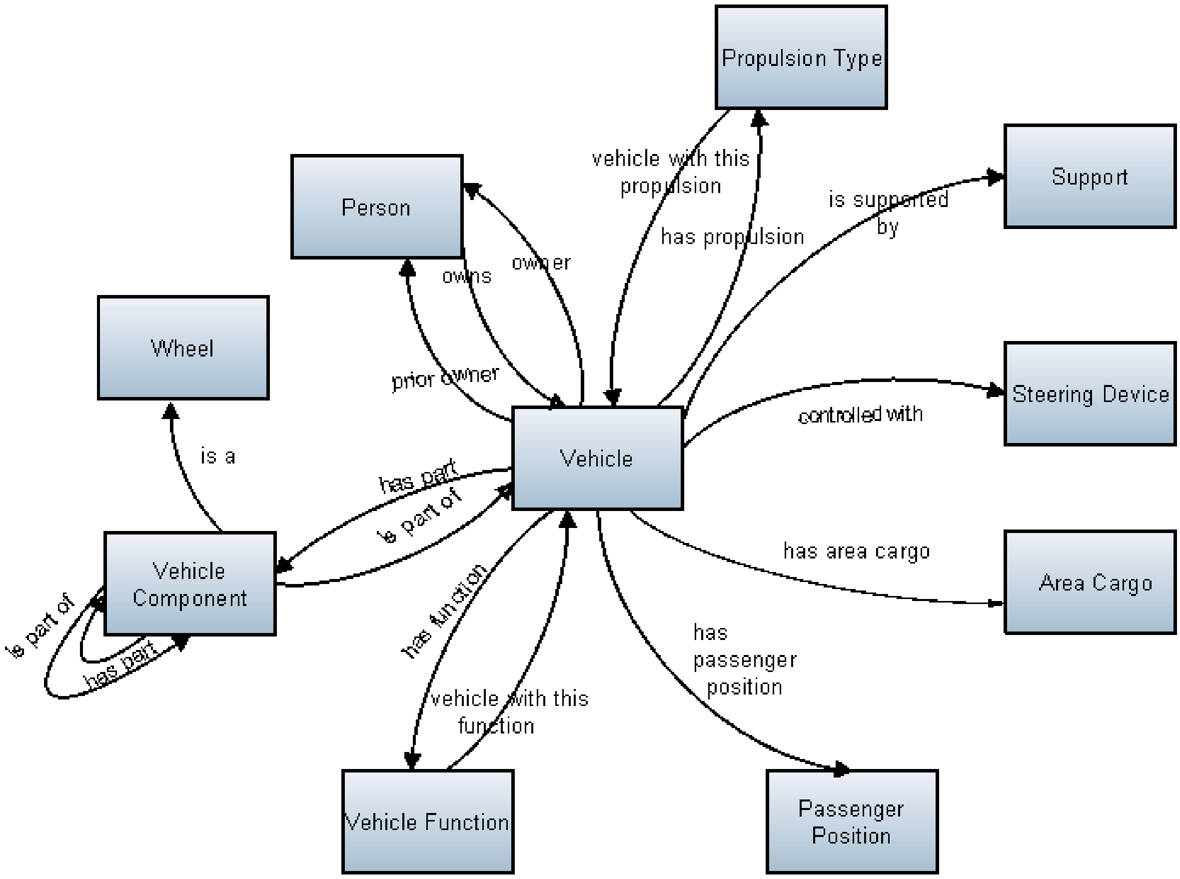

Figure 3: The *Binary Relations Diagram* for the vehicle ontology

Table 2: A small part of the concept dictionary for the vehicles' ontology

Concept	Attributes	Relations
Owner	Name	owns
Person	tax number Identity card	
Firm	Code number	
Vehicle	year colour acquisition date used for Collective Transport own manufacturing used for paid transport make model invoice market value	has rear area cargo controlled with passenger position owner former owner has Function has Part has Propulsion maximum number of passengers

Table 3: A small part of the *table of binary relations* for the vehicles' ontology

Relation	Domain (Concept)	Cardinality (maximum)	Range (Concept)	Matemathical Properties	Inverse Relation
passenger position	Vehicle	1	Passenger position	functional	
owns	Person	N	Vehicle		owner
owner	Vehicle	N	Person		owns

Table 4: A small part of the *Instance Attribute* for the vehicles' ontology

Attribute Name	Domain (Concept)	Datatype	Unit	Precision	Allowed Values	Cardinality
year	Vehicle	Integer	1800..2010	1		(1,1)
model year	Vehicle	Integer	1800..2010	1		(0,1)
capacity	Automotor Vehicle	Integer	kg or number of passengers	1	1..thousands	(1,1)
cargo capacity	Automotor Vehicle	Integer	kg	1	thousands	(1,1)
passengers capacity	Automotor Vehicle	Integer	number of passengers	1	1..100	(1,1)
id card	Person	String				(1,1)

Table 5: An entry in the logical axioms table for the Vehicles' ontology

Axiom Name	Automobile definition
Description	Vehicle with motorized traction, having 3 to 6 wheels, transporting up to 9 people, without a rear cargo or fifth wheel, and whose function is passenger transport,
Expression	Automobile ≡ Vehicle ⊓ ∃hasPropulsion.MotorizedTraction ⊓ >= 3 hasPart.Wheel ⊓ <=6 hasPart.Wheel ⊓ hasValue (passengerPosition, Seated) ⊓ hasValue(MaximumNumberOfPassengers, Until9Passengers) ⊓ ¬ (∃ hasPart.(RearCargoArea ⊔ FifthWheel))
Referred Concepts	Vehicle, MotorizedTraction, Wheel, RearCargoArea, FifthWheel
Referred Attributes	hasPropulsion, hasPart, passengerPosition, MaximumNumberOfPassengers

Table 6: The *Instance Table* for the vehicles' ontology

Instance	Concept	Meaning
UpTo50cc	Cc	≤ 50 cc
HigherThan50cc	Cc	> 50 cc
LightVehicleWeight	VehicleWeight	≤ 3,500 kg
HeavyVehicleWeight	VehicleWeight	> 3,500 kg
Walking	PassengerPosition	
Mounted	PassengerPosition	
Seated	PassengerPosition	
I nTheVehicle	Support	Indicates that the support that connects to a external part is located in the vehicle
InTheExternalPart	Support	Indicates that the support that connects to a external part is located in the external part

4. The vehicles' domain and resulting ontology

As a guiding principle, we decide to stick to the law categories and their definitions during the vehicles' ontology design. So, according to the laws, vehicles are classified in the dimensions of User category, Traction, Function and Type. Although the classes under these dimensions are mentioned in the codes, most of their definitions is not present, but only some relevant features. Therefore, a hard task was to appropriately portrait these classes into axioms in a way that they can correctly infer to which class a given described vehicle belongs. We also tried to endow the ontology with the maximum reusability, by sticking to the criteria recommended by Gruber (1995).

Just to have a flavour of how complex these axioms turned out to be, these are the definitions, in Portuguese and English, that corresponds to a Moped (*Ciclomotor*, in Portuguese), according to the Annex I of CTB, which contains "Concepts and Definitions" related to vehicles' categories:

> *"Ciclomotor – veículo de duas ou três rodas, provido de um motor de combustão interna, cuja cilindrada não exceda cinqüenta centímetros cúbicos [...] e cuja velocidade máxima [...] não exceda a cinqüenta quilômetros por hora."*

> *"Moped – vehicle with two or three wheels, endowed with in an internal combustion engine restricted with a maximum engine displacement of fifty cubic centimeters […] and whose maximum speed does not exceed fifty kilometers per hour."*

And this is its translation into Description Logic:

$$\text{Moped} \equiv \text{Vehicle} \sqcap (= 2\ \text{hasPart.Wheel} \sqcup = 3\ \text{hasPart.Wheel}) \sqcap$$
$$=1\ \text{controledWith} \sqcap \forall \text{controledWith.Handlebar} \sqcap$$
$$\text{hasValue}\ (\text{cc}, \text{UpTo50cc}) \sqcap \text{hasValue}\ (\text{maxSpeed}, \text{UpTo50KmH}) \qquad (2)$$

We have the following remarks about the definition above:

- In order to support these axioms, we had also to deploy support definitions in the form of classes like vehicle components, accessories, functions and user types, as well as 27 data and 18 object properties.
- The axiom above also states that Mopeds are vehicles driven by handlebars, although this was not stated in any law code. This is, however, the common understanding of the law by people, lawyers and judges *by default.*
- On the other hand, we dropped the reference about the internal combustion engine as a necessary and sufficient condition for Mopeds, since engines with the property *cc* can only be of internal combustion.
- It was necessary to introduce wheels as individuals, since there can be different types of wheels even in a same vehicle, like tractors,
- And also to define categories of speed and displacements, instead of working with continuous values, once OWL reasoners are not able to compute with numbers. It can just count instances, like the two or three wheels.

In the next section, we look into the semantic deficiencies that were evidenced throughout ontology design process.

5. Semantic deficiencies regarding vehicles found in the law codes throughout the process of ontology construction

Though acceptable and, in most of the times, manageable on daily life, small misunderstandings in texts are potentially dangerous not only in the exact sciences, but also plagues the field of Law. It is so commonly accepted in many countries that laws are ambiguous or form an inconsistent set of regulations, in which some laws contradict others, that a deluge of trials depend upon judges personal law interpretation. Indeed, our claim is that a new field of ontology application could be to clarify, disambiguate and correct semantic deficiencies contained in law cases. We have experienced such difficulties and provided solutions for that case study that we will describe in the following, describing ambiguities, ill-defined modelling cases and inconsistencies found in the laws during the definitions.

5.1 Ambiguities

For instance, the Portuguese term *Carroçaria* (or *Carroceria*) consists of a Polysemy (a word with more than two senses) in Brazil. A first denotation of this word stands for rear area cargo. This sense is employed in the aforementioned Resolution 291 in its Fifth Article:

> *"Art. 5° Em caso de [...] caminhão, com carroçaria aberta ou fechada, […]"*
>
> *"Art. 5. In case of […] a [...] truck, with open or closed rear area cargo, […]"*

In another word sense, *Carroçaria* stands for vehicle body. Contran Deliberation number 64 (Conselho Nacional de Trânsito - CONTRAN, Deliberação no. 64, 2008), another document which disciplines weights and capacities registers in vehicles, uses *both* senses in its only five pages. In its item 3.1.3, it uses the first sense (area cargo):

> *"3.1.3 Veículo [...] que recebeu carroçaria ou implemento [...]"*
>
> *"3.1.3 Vehicles [...] that received a rear area cargo or new implement [...]"*

While in item 4.2.3, it refers to the second (vehicle body):

> *"4.2.3 – [...], a indicação [do peso] deverá ser afixada na parte externa da carroçaria na lateral dianteira."*

> *"4.2.3 – [...], the [weight] indication must be shown on the right side of the vehicle body external part."*

Polysemy can indeed lead to ambiguities. In this case, for instance, both classes can fit to the given description, since the item could be easily (and correctly) interpreted as:

> *"4.2.3 – [...], the [weight] must be shown on the right side of the rear area cargo external part."*

5.2 Ill-defined modelling

This consists by far the most frequent semantic deficiency. For instance, motorcycles can be classified into two disjoint categories: passengers and cargo, but no further details are conveyed in the law codes. So, how should we make clear distinction between the two? Description Logics fits quite well to the task of correctly determining the exact perimeter of each class, as well as the inheritance relations among them. One solution could be based on the fact that cargo motorcycles are employed for cargo transport. Defining cargo motorcycles on this way may bear little usefulness, though; see the following definition:

$$\text{CargoMotorcycle} \equiv \text{Motorcycle} \sqcap \exists\ \text{hasFunction.CargoTransport} \quad (3)$$

This axiom states that for cargo motorcycles to be recognized as such, in systems that classify vehicles, iff there is an instance c of CargoTransport. a motorcycle instance m and an instance of object property hasFunction (m, c). Nevertheless, this modeling lacks ontological engagement: if any motorcycle is used for light cargo transport (e.g. transporting a single document), it would be classified as a cargo one! The intended meaning of cargo motorcycles for the traffic institutions is based more on vehicle's components instead. So a definition coherent with this stance is displayed below:

$$\text{CargoMotorcycle} \equiv \text{Motorcycle} \sqcap \exists\ \text{hasPart.CargoArea} \sqcap \exists\ \text{hasFunction.CargoTransport} \quad (4)$$

5.3 Inconsistencies

Inconsistencies need at least two axioms. Therefore, they are sometimes hard to locate, since they can hide in distinct law codes. Nonetheless, we have even found inconsistencies in a same official document. For instance, the Law number 9,602 of CTB, brings the following statement:

> *"Art. 96. Os veículos classificam-se [...] quanto à tração: a) automotor; b) elétrico; c) de propulsão humana; d) de tração animal"*
>
> *"Art. 96. Vehicles are classified according to their traction into [...]: a) automotor; b) electrical; c) human propulsion [like bikes]; d) animal traction"*

Although not explicitly stated, this law's default interpretation considers these partitions (automotor, etc) disjoint, for instance, electrical and automotor vehicles:

$$\text{ElectricalVehicle} \sqsubseteq \neg\text{AutomotorVehicle} \quad (5)$$

Electrical vehicles were not defined in the codes - they are vehicles moved by their electrical engines. Engines in Brazil that are driven on roads are divided into two distinct classes, electrical and internal combustion engines. The set of definitions below exhibits this state of affairs, assuring via a closure axiom that an electrical vehicle is and can only be moved by an electrical engine (6):

$$\text{ElectricalVehicle} \sqsubseteq \exists\text{isMovedBy. ElectricalEngine} \sqcap \forall\text{isMovedBy. ElectricalEngine} \quad (6)$$

$$\text{Engine} \equiv \text{ElectricalEngine} \sqcup \text{InternalCombustionEngine} \quad (7)$$

Electrical and internal combustion engines are disjoint too:

$$\text{ElectricalEngine} \sqsubseteq \neg\ \text{InternalCombustionEngine} \quad (8)$$

On the other hand, the Annex I of CTB brings additional details about automotor vehicles:

> *“Veículo Automotor - todo veículo a motor de propulsão que circule por seus próprios meios [...]. O termo compreende os veículos conectados a uma linha elétrica e que não circulam sobre trilhos (ônibus elétrico).”*
>
> *“Automotor Vehicle - every vehicle with a propulsion engine that can move by its own means [...]. The term encompasses the vehicles connected to an electrical wire and which do not move over tracks ([like] electrical buses).”*

Note that, on purpose, this definition does not accept electrical trains. However, this statement entails *lots* of inconsistencies. Vehicles that move by their own have to possess an engine; so

$$AutomotorVehicle \equiv Vehicle \sqcap \exists\ isMovedBy.\ Engine \quad (9)$$

Unfortunately, this definition clashes with the one of electrical vehicle, in any interpretation. The inconsistency was caused by a conflict between the axioms that declares automotor and electrical vehicles disjoint as well as their engines (axioms 5 and 8 respectively) and the actual concept of automotor (9). They cannot simultaneously hold. The most reasonable interpretation, which we included in our ontology, is to drop the disjointness axiom 5. After this change, the system automatically subsumes electrical vehicles as a subclass of automotor.

The same problem applies to all the subclasses of electrical vehicles, like electrical bus and electrical trains. They were as a consequence all inconsistent. Moreover, another flaw in the law was explicitly portrayed in Article 96, which considers electrical trains as not being automotor vehicles, despite of moving due to electrical engines.

6. Discussion

If in a domain as simple as vehicles such semantic flaws arise, they certainly occur even more often in more complex law branches, like commercial disputes, tax regulations, and family law, to cite but a few. In these law domains, opposite law interpretations are acceptable in many cases, since contradictions among different law codes abound. Portraying a correct modeling of these branches in thorough details is certainly to be conducted with all the care and requires a clear identification of the correct or intended law interpretations in a way that prevent contradictions among the codes that shall appear.

Our claim with the current work is that ontologies can be used in the field of law to improve the quality of law creation. Support tools that benefit from such ontologies could be used in parliaments, courts and attorneys’ offices, provoking a deep behavior change in the way of handling, changing and working with laws. Another good consequence and usage of law ontologies defined in this way is bridging the gap between law and systems, particularly for eGovernment. By using a common representation, such systems have less risk of not following closely laws and regulations. Indeed, even public services can be available via Semantic Web Services to take on the job. Therefore, typical tasks accomplished by law ontology-based tools could comprise:

- Checking whether new laws conflict with older ones;
- Identifying law possible interpretations (as logic models);
- Providing support for choosing the correct interpretations to be included in a particular law article;
- Clearly showing the chain of laws to be changed in case a change is required;
- Simulating situations in order to check defense or accusation possibilities;
- And many other tasks of the same flavor.

7. Envisaged application

We devise a concrete application of the present work: the development of a semantic law checker prototype, as mentioned in the previous section. Such semantic law checker should make for more defying modeling cases, like stating the laws of traffic ontologically. In these scenarios, false agreements, as pointed by Guarino (1998), and ill-defined modeling - see many examples at (Guizzardi, 2006) -, that are not being dealt here, will probably take place often, as laws from the same or from different codes are bound to interweave and conflict. Therefore, a reputed semantic law checker needs to employ top ontologies (like DOLCE (Gangemi et al 2003), for instance) to assign relevant features of the ontologies' classes, and meta-property checking, using the OntoClean methodology (Guarino & Welty 2002).

Figure 4 displays the architecture of a semantic law checker. As the user (attorneys, judges, law authorities) describes laws, the system builds the models and warns of inconsistencies or alternative models. If the law ontologies are already built, it can also be used to check whether a legal action (process or law suit, for instance) is applicable.

Such type of systems could clearly show the different law interpretations (or, in more formal, logical words, models) and point out when they conflict, or when a valid interpretation cannot come out from a set of laws. Engers et allii (2001) have already implemented a system like this, by using UML modeling and production rules (Giarrantano & Riley 94). The problem with rules is that, as any non-monotonic reasoning mechanism, they cannot take into account logical models, what would hamper checking classes' consistency and the checking of the meta-property compatibility of inheritance relations among classes in the flavor of OntoClean, unless the meta-properties compatibilities are explicitly stated as rules.

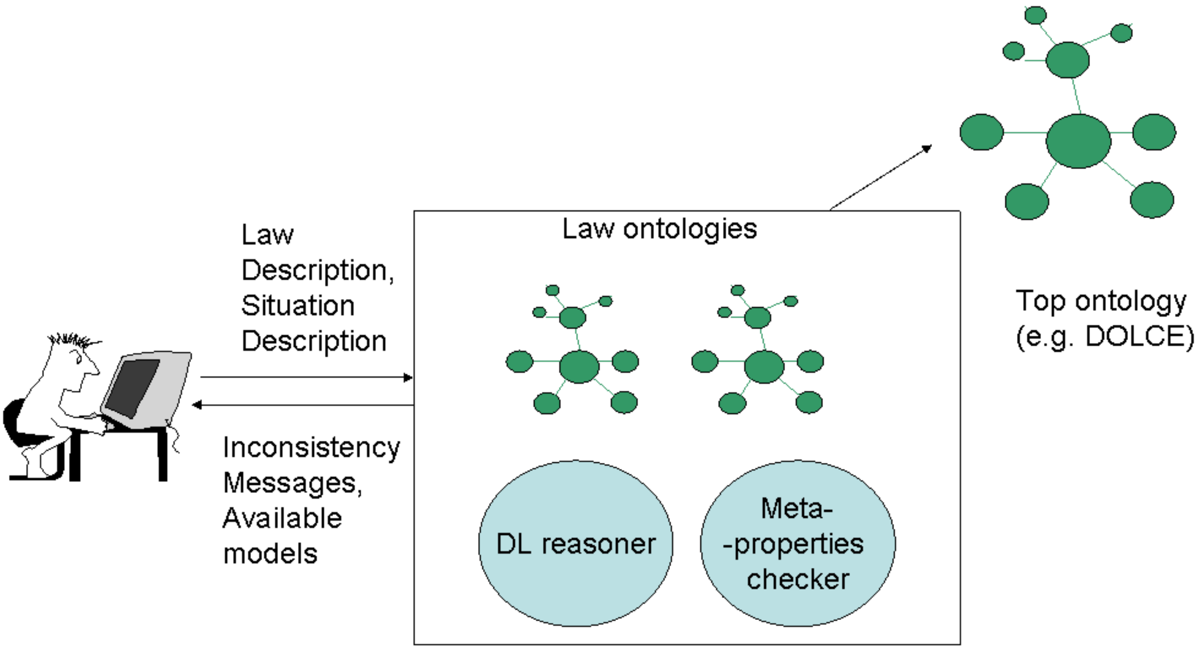

Figure 4: Architecture of a semantic law checker

8. Related work

The field of Law has attracted strong interest from ontology researchers since mid-90s, when ontologies started occupying its relevant place in the Artificial Intelligence mainstream. Ever since, many scientific events and journals that publish papers linking ontologies and law are being held, including specific workshops, like the ones on regulatory ontologies (WORM – Jarrar and Gangemi 2004), legal ontologies (LOAIT – van Engers et al 2001) and Legislative XML, this latter discussing also Semantic Web applications (LXML - Winkels et al 2008).

One of the first works to address the complexity of law modeling, which requires deep meta-knowledge, was the Core Ontology of Law by Valente and Breuker (1996). They grounded their work under the premise that any branch of law must make use of a common set of concepts like responsibility, normative knowledge, agents (in the law sense), regulations, etc. Breuker (2004) has also shown the role that top ontologies can play as a basis for defining knowledge for a specific law subfield (like family law, for instance). Many applications and prototypes take advantage of legal ontologies nowadays, e.g., systems that facilitate citizens to check whether disputes are feasible, by checking the law possibilities (Steenbergen, 2005).

As for the usage of ontologies to validate laws, Engers et allii (2001) have produced a system that employs legal ontologies to provide flexibility to systems, including checking for semantic anomalies among law articles and regulations. It relies on a framework similar to the one suggested here. Their system validates laws using UML/OCL specifications. The risk lies on the fact that these languages have no formal semantics. Although a tool to test whether our laws are correct is future work, we are using OWL, which is based on the well-defined Description Logic formalism. Good features of their tool also include the determination of possible decision paths for juridical decisions and the easy maintainability of the systems against law changes. Since these changes are rather frequent, representing knowledge declaratively is the best solution to cope with the task.

Another good example of ontology-based law system was produced by Hoekstra et allii (2009). Their system employs legal (mostly lightweight) ontologies that are used to search for contextual legal information in texts. Information pieces collected from these tests are used to select case frames. If a case frame is selected, then the legal action can take place. The system seems to be flexible mainly because it is not formal, once it is, based on case frames and retrieved texts.

The formal system HARNESS (van de Ven et al, 2008) makes good usage of core law and law ontologies (e.g. library laws) stated in OWL, where these latter's classes contain many subclasses of the former's classes. The system is capable of recognizing when a law is not being followed in a deep and principle way, once, by doing that practice (inheriting and qualifying law types, laws have elegant hierarchies.

As for ontology on vehicles, although this theme is present in many description logic examples and exercises, and there are many of them, but not with the same focus and perspectives. For instance, Lukibanov (2005) shows ontologies that support design activities at DaimlerChrysler, whose focus are on car parts, connections, compatibility among the parts, etc. Another ontology (Schlenoff, 2005) focus on the activities of controlling vehicles and war equipment automatically via a multi-agent system, containing classes such as obstacles, missions, tasks, etc.

Nevertheless, we have found out only one ontology focused on our subject, vehicles' categories, although in a different context, the domain of car sales (Straccia, 2010). Thus, in order to assess the (quantitative) quality of the resulting ontology, we have carried out a comparison of ours with this ontology. Numbers are shown in Figure 5.

These numbers tell us that the Car ontology resembles more a database than an ontology, once it encompasses a high number of datatype properties while just few object properties. The datatype properties are intended to characterize the different types of vehicles under negotiation. Beyond this deficiency, the cars' ontology does not contain a single axiom, so the classes are artificially organized in a hierarchy. This organization is naturally not supported by the classes' definitions as we defined in our ontology, so it is at least questionable. Moreover, for not presenting axioms, a reasoning classifier cannot take advantage of it.

9. Conclusions

With this work, we tried to make the case for applying Semantic Web technologies for checking the consistency of official documents, particularly law codes. We have presented a case study on the field traffic law, in which some inconsistencies were found out in Brazilian law codes during the design of an ontology to describe Brazilian vehicle types according to definitions conveyed in these codes. Our hypothesis for future work include a larger and more complex case study with more convoluted law codes, in which plenty of ontological mistakes and inaccuracies are supposed to arise, as well as a good (and challenging) implementation.

In the context of this work, we have outlined a framework of possible semantic deficiencies on texts that is not on the scope of this article. The current work together with this framework enables the construction of an automatic semantic law checker, that can serve to lawyers, judgers and law makers in order to analyze the consistency of law codes.

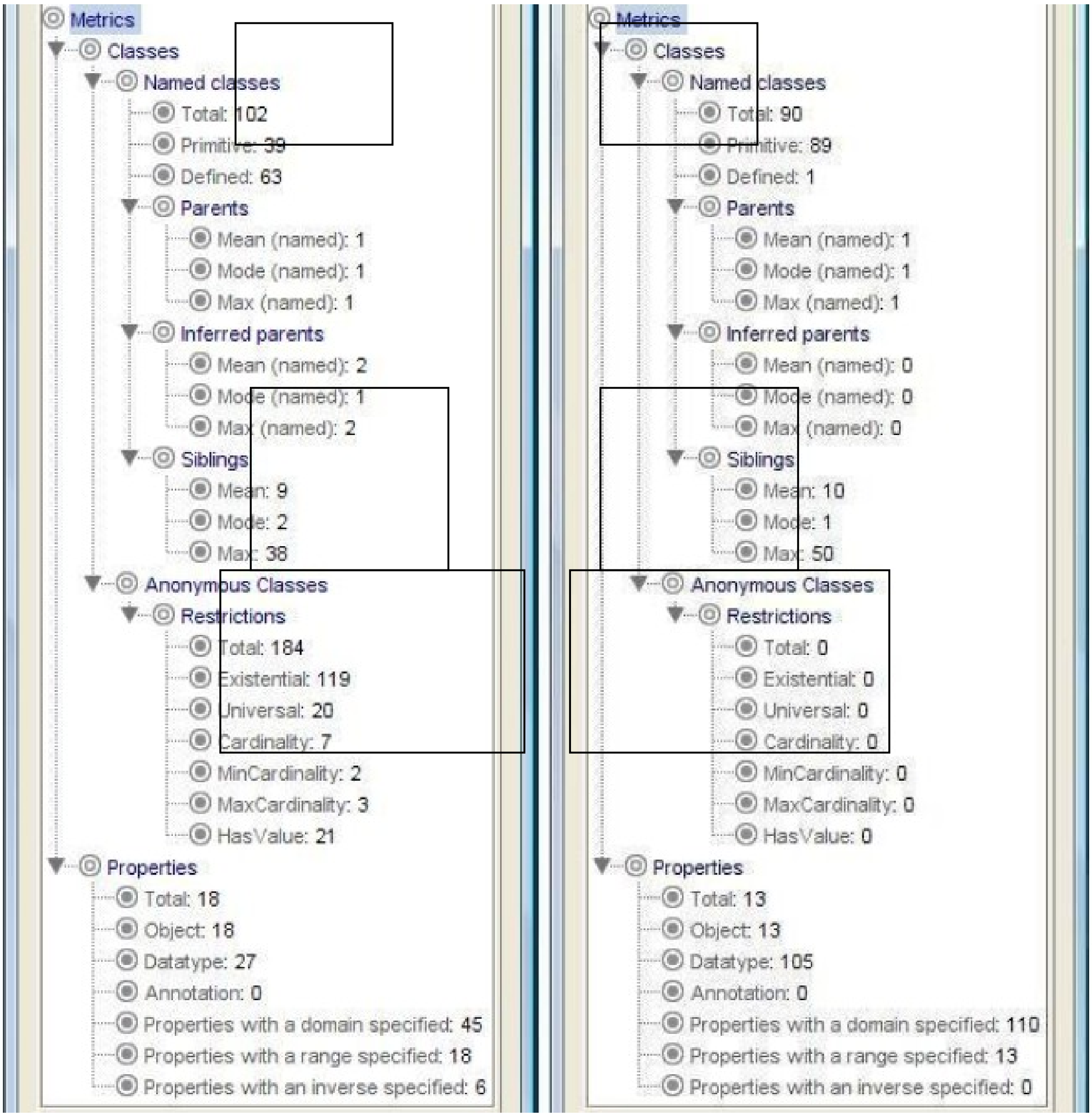

Figure 5: Comparison between the ontology developed here and the Car ontology (Straccia, 2010), the latter, by not containing axioms and having few relations, resembles more a database, with its many attributes

10. References

Berners-Lee, Tim; Hendler, James and Lassila, Ora. "The Semantic Web". Scientific American Magazine, May 2001.

Blázquez, M, Fernández-López, M, García-Pinar, JM and Gómez-Pérez, A, "Building ontologies at the knowledge level using the ontology design environment", Knowledge Acquisition of Knowledge-Based Systems Workshop (KAW) (http://ksi.cpsc.ucalgary.ca/KAW/KAW98/ blazquez/), 1998. [Accessed 10 Oct 2010].

Breuker, J. Constructing a Legal Ontology: LRI-Core. In Freitas, F. Stuckenschmidt, H., Volz, R. Proceedings of WONTO-2004, Workshop on Ontologies and Their Applications. LivroRápido, Recife, Brazil. 2004.

Comte, Frédéric and Leclére, Michel..A Semantical Reasoning Framework for eGovernment of the French Social Welfare System, The Semantic Web meets eGovernment, 2006 AAAI Spring Symposium Serie, Stanford University, California, USA, March 27-29, 2006.

Conselho Nacional de Trânsito - CONTRAN, Deliberação Nº 64, de 30.05. 2008 www.denatran.gov.br/download/Deliberacoes/DELIBERACAO_CONTRAN_64_08.pdf [Accessed 10 Jan 2010].

Conselho Nacional de Trânsito - CONTRAN, Resolução Nº 291 de 30.05. 2008 www.denatran.gov.br/download/Resolucoes/RESOLUCAO_CONTRAN_291.pdf , 2008. [Accessed 10 Jan 2010].

Gangemi, A., Guarion, N., Masolo, C. and Oltramari, A. "Sweetening WORDNET with DOLCE", AI Mag., Vol. 24, No. 3., pp. 13-24, 2003
Giarrantano, J. Riley, G. Expert Systems – Principles and Programing, 2nd. Ed. PWS Plubishing Company, 1994.
Goméz-Pérez, A.; Fernández-López, M.; Corcho, O. Ontological Engineering. Berlin Heidelberg New York:Springer-Verlag, 2004.
Governo de Pernambuco. DETRAN-PE. Departamento Estadual de Trânsito de Pernambuco. www.detran.pe.gov.br , 2009. [Accessed 10 Jan 2010].
Governo de Pernambuco. Portal Pernambuco. www2.pe.gov.br/web , 2010. [Accessed 10 Jan 2010].
Gruber, T. Toward principles for the design of ontologies used for knowledge sharing. International Journal of Human-Computer Studies, Volume 43 , Issue 5-6 Nov./Dec. 1995.
Guarino, N.Formal Ontology in Information Systems.In:Guarino,N(ed) 1st Int. Conf. on Formal Ontology in Information Systems(FOIS'98). IOS Press, Amsterdam, 3–15, 1998.
Guarino, N., Welty, C. Evaluating ontological decisions with OntoClean. Communications of ACM. Feb 2002.
Guizzardi, G. Ontological Foundations for Structural Conceptual Models. PhD thesis, University of Twente, Netherlands, 2005.
Haarslev, Volker and Moller, Ralf. RACER system description. (2001) In Proc. of the Int. Joint Conf. on Automated Reasoning (IJCAR), June 18-23Siena, Italy, 2001.
Hoekstra. R. BestPortal: Lessons learned in lightweight semantic access to court proceedings. In Proc. of the 22nd Int. Conf. on Legal Knowledge and Information Systems (JURIX 2009), IOS Press. 2009.
Jarrar, M., Gangemi, A. 2nd Workshop on Regulatory Ontologies (WORM'04), www.jarrar.info/WORM_2004.htm , 2004. [Accessed 10 Oct 2010].
Jurafsky, D., Martin J. Speech and Language Processing. Prentice-Hall, 2000.
Lukibanov, O. Use of Ontologies to Support Design Activities at DaimlerChrysler. Presentation slides. 8th Intl. Protégé Conference - Madrid, Spain, July 18-21, 2005 - http://protege.stanford.edu/conference/2005/slides/2.1_Lukibanov.pdf , 2005. [Accessed 10 Oct 2010].
Ministério das Cidades, Código de Trânsito Brasileiro e Legislação Complementar. www.senado.gov.br/web/codigos/transito/httoc.htm , 2008. [Accessed 10 Jan 2010].
Ministério das Cidades, Conselho Nacional de Trânsito, Departamento Nacional de Trânsito. www.denatran.gov.br/contran.htm , 2009. [Accessed 10 Jan 2010].
Patel-Schneider, P., Hayes, P., Horrocks, I.: OWL - Web Ontology Language Semantics and Abstract Syntax W3C Recommendation. www.w3.org/tr/2004/rec-owl-semantics-20040210/. , 2004. [Accessed 10 Jan 2010].
Schlenoff, C., Washington,R., Barbera, T. An Intelligent Ground Vehicle Ontology for Multi-Agent System Integration. International Conference on Integration of Knowledge Intensive Multi-Agent Systems, (KIMAS 2005), Waltham, USA, 2005.
Steenbergen, W. Rationalizing Dispute Resolution: Form best alternative to the most likely one. Proceedings of the 3rd ODR Workshop, Brussels, Dec.8-10, 2005.
Straccia, U.. Umberto Straccia site. http://gaia.isti.cnr.it/~straccia , 2009. [Accessed 10 Oct 2010].
Valente, A., Breuker, J. Towards Principled Core Ontologies. In B.R. Gaines and M. Mussen, editors, Proceedings of the KAW-96, Banff, Canada, 1996.
van de Ven, S., Hoekstra, R., Breuker, J., Wortel, L., El-Ali. A. Judging Amy: Automated legal assessment using OWL 2. In Proc. of OWL: Experiences and Directions (OWLED 2008 EU), October 2008.
van Engers, T.,Glassée, E. Facilitating the Legislation Process Using a Shared Conceptual Model. IEEE Intelligent Systems, 16(1):50-58, Jan/Feb. 2001.
van Engers, T., Santor, G., Ganguemi, A. LOAIT Workshop – Legal Ontologies and Artificial Intelligence Techniques www.itig.cnr.it/loait/loait07.htm , 2004. [Accessed 10 Oct 2010].
Winkels, R., Francesconi, E., Boer, A., Hoekstra, R., Palmirani, M., Sartor, G. Workshop on Legislative XML www.itig.cnr.it/Jurix08/images/ProgramLXML08.pdf , 2008. [Accessed 10 Oct 2010].