

Piotr Stolarski¹, Tadeusz Tomaszewski¹

¹ Poznan University of Economics, Department of Information Systems, al.
Niepodleglosci 10, 60-967, Poznań, POLAND
{P.Stolarski, T.Tomaszewski}@kie.ae.poznan.pl

In this paper we describe a legal framework for simple tax status detection of transaction parties as a part of a hybrid eCommerce information system. We also focus on one, in our view more interesting point encountered during the model setup which is the case of legal knowledge modeling in such a way that it is immune to fast-pace changing of jurisdiction. This kind of quality is especially important in the tax and commerce law domain as regulations connected to those two branches are characterized by a high volatility.

1 INTRODUCTION

Recent years witnessed an intense development of eCommerce and its technologies. Although the eBusiness systems market seems to mature yet with still emerging technologies but it is certain that we are still merely on the brink of the true revolution promised by web semantics or web services. So far the undisturbed expansion of electronic economy was in fact due to projected potential and common policy yet this progression may not for long be without any threats. As long as the whole industry was at its startup stage it was left mainly unnoticed by institutions. However as eMarkets are becoming a major source of profits and centers of added value in the global economy landscape, they (and their participants) have been recognized as vital objects of interest of stakeholders from different countries and different legal systems. Additionally it has to be taken into consideration that while institutions of national authority work on a local basis, the eCommerce performance is worldwide. These should generate a growing need for the eCommerce world combined with legal knowledge systems capable of interpreting the eCommerce matters with regard to and in context of the applicable legal recognition.

Constructing the general framework has also another advantage over creating a hardcoded built-in rules-based system as the model approach enable the use of similar specific solutions in different jurisdictions and by larger group of entities including on-line stores, eMarkets themselves, on-line auctions systems and individual contractors as well as different tax institutions. In the paper we delimit ourselves to address the issues of sale of finished product domain only.

The paper is divided into five main sections which are as follow: 2. Legal Models, eBusiness & Taxes – is the general introduction to the legal knowledge models, together with an ontology definition, additionally we give a general overview of the knowledge sources subject to analysis during conceptual work on the tax ontology; in 3. Domain-specific Cases section we demonstrate selected

examples of problems specific to the legal and tax domain description. 4. Putting that all Together – binds the gathered experiences into a comprehensive model proposition whereas 5. is the sum of Lessons Learned.

2 LEGAL MODELS, eBUSINESS & TAXES

2.1 Legal Ontologies

Legal knowledge modeling is a vigorous research discipline. Currently two main directions may be pointed out, namely: Automatic argumentation and ontological representation. While the first branch is more focused on legal discourse and active use of the model, the latter one create an inevitable background providing legal-specific notion of understanding the domain terms which has to be somehow represented in any legal-aware system.

The legal ontologies are specific ones on account of the difficulty of proper representation of often very subtle nuances among terms, highly-specialized corpuses used to incorporate knowledge from – being mostly formal documents with language of particular features. Moreover, there are also some extended expectations formulated towards such knowledge models connected to their use. A concise description of works related to legal ontologies may be found for instance in [1].

In our works we utilize the definition of ontology similar to the one taken from [2]. This definition is compatible with most ontology description languages (OWL, WSDL) [3, 4].

According to Gangemi [5] such knowledge models may be useful in the legal domain in a number of contexts: 1) For knowledge validation; 2) Classification of instances or facts; 3) As helpers for information extraction; 4) For automation of planning; 5) For formalizing case abstractions within a general framework; 6) Advance reasoning in rule-based systems.

As it will be showed later at least the first, second, third as well as the last one are within the scope of creation of framework for tax analysis.

EBusiness is the conduct of automated business transactions by means of electronic communications networks [6]. In comparison to the above definition the eCommerce is being defined as a process of focusing on buying and selling products and services via Internet. It is therefore regarded to be the narrower term to eBusiness. As mentioned earlier, the efforts of creation of legal ontologies are intensified for last decade and there exist a rather large resource of those containing and dealing with a general legal vocabulary [7]. We perceive a lack of oriented models describing more precisely matters of a concrete branch of law, a statute or even only some specific regulations. It is certain that with new opportunities of use arising and with demand form the software systems for facilities enabling easy to (re)use dedicated expert knowledge which additionally will be change-proof the specific ontologies will be of great value. It is however a very hard task to develop a knowledge model which will embrace all the mentioned needs and other constraints (i.e. security concerns) thus the intense ongoing research effort in this track is essential. For the same reason also in the tax law not much work has been done in recent years although the requirement for a concise tax ontology together with a suggested outline has been addressed by Melz and Valente in [8] since their work no essential progress was reported. It is also important that their proposition was based

on the U.S. law. A vocabulary concerning taxes is also enclosed in large projects such as LegalRDF [9]. It is however a general approach focusing on taxonomical representation.

The interesting fact is that the necessity of supplying the comprehensive tax knowledge model seems commonly recognized in practice by the many authors of the ontologies from eCommerce domain¹.

2.2 Tax Regulations - EU Context

The EU policy states that the tax regulations are the issue of national sovereignty and therefore the harmonization efforts are limited only to ensuring that some very general rules are applied. The regulations then aim to provide compatibility between systems of the Member States without immediate integration. Thus the regulations in the legislation of EU may be found mainly in high-level, general documents². Yet as early as in 2000 the EU authorities regarded eCommerce as the international issue rather than a local one [10].

2.3 Polish Taxation System

Polish tax law is an independent branch of law it utilizes many institutions from the civil law yet because of its character it is much more related to the administrative part of legal system. From the point of view of the sources of regulations the three parts may be distinguished, namely:

- The Tax Ordination statute [11] – introducing all the common and important general norms.
- A group of tax-specific statutes – bringing institutions specific to the given type of tax.
- Regulations delegated from the statutes – lower level acts.

The Tax Ordination is a large about 45000 words document divided into 10 Divisions that actually instantiate the taxation system. On the other hand the statute does not establish any tax itself. According to the Polish Constitution [12] “[...] imposition of taxes, as well as other public imposts, the specification of those subject to the tax and the rates of taxation, as well as the principles [...] shall be by means of statute” which means actually - by the mentioned tax-specific bills.

Another important aspect of tax law especially in Poland is the level of turbulence and incorporated amendments. For instance the bill of Civil Deeds Tax was changed 17 times, the Tax Ordination 53 and the Personal Income Tax bill 148 times³. Therefore even this demonstrates the great need to take into account altering regulations in the modeling effort.

¹ Semantic content search engine Swoogle returns about 125 result documents with the keyword „tax”; an essential part of them being the eCommerce related ontologies.

² See for instance Articles 90 and next of the Treaty Establishing the European Community

³ Basic statistics available on Polish Parliament's legal acts database: <http://www.sejm.gov.pl/prawo/prawo.html>

3 DOMAIN-SPECIFIC CASES

Let us consider the following case. We assume that we are able to deliver the legal knowledge representation of the actually valid part of the tax law connected to matters of eCommerce which may be done using the developed ontology stub. The actual validity means that the model reflects the state of the law for – let say – 31 January 2008. The very fragment of such representation is visible on Fig. 1.

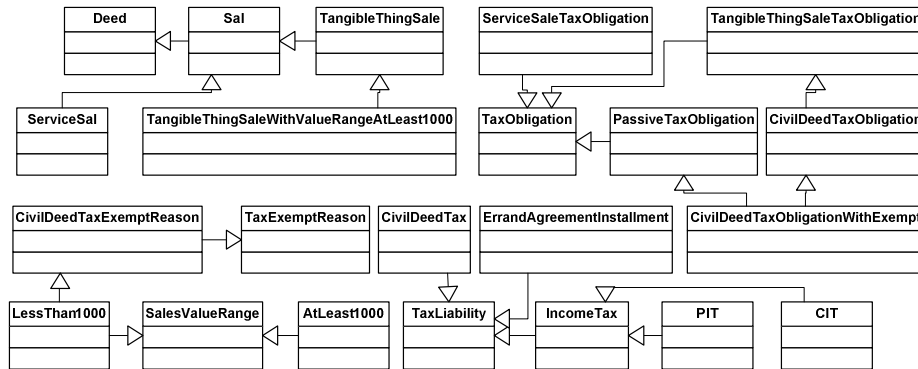


Fig. 1 Fundamental concepts of the tax ontology.

It has to be clarified that the above model brings in a major simplification – it assesses the instantiation of tax obligation independently for any subject.

Let us consider the consequences of it if in our system we commence to analyze some specific transactions dated 18 December 2006. It is important to bear in mind that in case of the bill of Civil Deeds Tax the last amendment came into force with the beginning of year 2007. As a result it means that we have to cope with the common legislation problem meaning addressing changes in regulations which in turn yields or may yield particular modification of the real-life scene (or at least statuses of their actors). We will take a closer look at this problem now.

ECommerce transactions are commonly stored, for the sake of retrospective examination, in some kinds of databases or logs. The unofficially emerging standard may be the one proposed by Google – the ELF/ELF2⁴. Such sets of data are the source of indeed valuable information also when it comes to tax investigation and categorization. This simple format contains information about each separate transaction and any item being the object of the transaction. In our modeling case it may look like this:

```
!36530 123.123.123.123 [18/Dec/2006:11:31:45 -0800] - -
1000.99 - - Warsaw - 23452 PL "Mozilla" ""

36530 123.123.123.123 [18/Dec/2006:11:31:45 -0800] "LG LCD
Display 17in" - 1000.99 1 - "Mozilla" ""
```

The first two lines describe the whole transaction whereas the next two represent the sole item of the transaction. We should also note, that it seems – that that particular format does not support any information about the currency of transaction.

If we extract interesting facts from the log we may represent it in the form of a

⁴ https://secure.urchin.com/helpwiki/ELF_%26_ELF2_Log_Formats

semantic net like the one presented on Fig. 2.

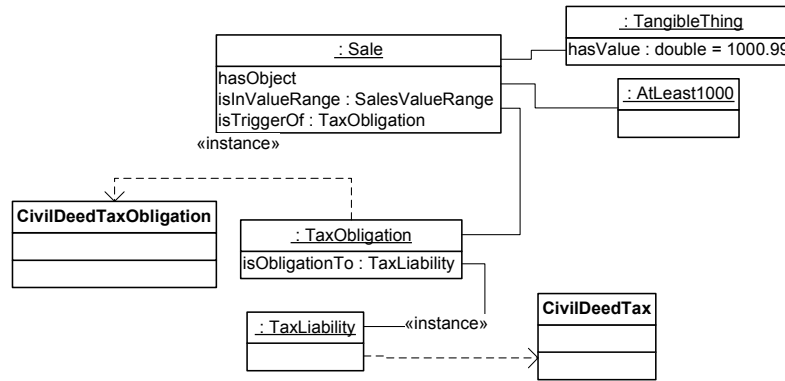


Fig. 2 Simple representation of the presented example.

The same information may be transformed into OWL and integrated with our tax ontology. Yet, as our tax legal knowledge model reflects so far the legislation state dated the end of January 2008, after applying a reasoner we get the result that on the side of the seller there exist the TaxObligation and it is a plain CivilDeedTaxObligation which would be correct supposing that the analyzed transaction had taken place under the ruling of the amended bill of Civil deed Tax. Unfortunately it did not.

This example excuses enough for introducing some kind of ontology evolution or versioning system. There is a number of works on the topic of ontology version tracking, for example [13]. The one method most close to our needs is proposed in [14]. Unluckily the application of model of Eder and Koncilia has disadvantages (mentioned in the referred paper). The authors, however do claim that apart from presented “OWL standard employing” approach there should be other approaches possible to attain similar (or better) results – amongst them the “Meta-ontology” one – which we developed in our model.

The authors of the model assert that “main advantage of this approach is that any ontology description language [...] can be used to define the different ontology versions” which suggest that they mean to use some external resources to describe the ontologies which is not a necessity. It seems that there is a simpler possibility – to employ the OWL Full features in order to extend meta-classes, meaning owl:Class and owl:ObjectRelation. Below we present a short OWL Full code which shed a light on how it may be achieved – however a detailed description is out of scope of this paper.

```

<owl:Class rdf:ID=" VersionedClass">
  <rdfs:subClassOf rdf:resource="&owl;Class"/>
  <rdfs:subClassOf><owl:Restriction>
    <owl:onProperty
rdf:resource="#hasPreviousVersion"/>
    <owl:someValuesFrom
rdf:resource="#VersionedClass"/>
  </owl:Restriction></rdfs:subClassOf>
</rdfs:subClassOf>
</owl:Class>

```

```

<owl:Class rdf:ID="VersionedObjectProperty">
  <rdfs:subClassOf rdf:resource="&owl;ObjectProperty"/>
  ...

```

In this solution the main advantage we perceive is the theoretical⁵ possibility of use of any standard reasoner to make inference on both meta-ontology and ontology levels. This means that it should be possible to check the consistence of the ontology also against its versioning description layer. Another improvement relies on the fact that we allow to identify classes not only by their validity dates but also by explicit enumeration of their predecessors and ancestors which resembles more the evolution description in comparison to “weak” versioning.

After discussing the improvements we will return to our initial case. If we sum up the initial knowledge model with the meta-ontology versioning solution we will get a result similar to Fig. 3.

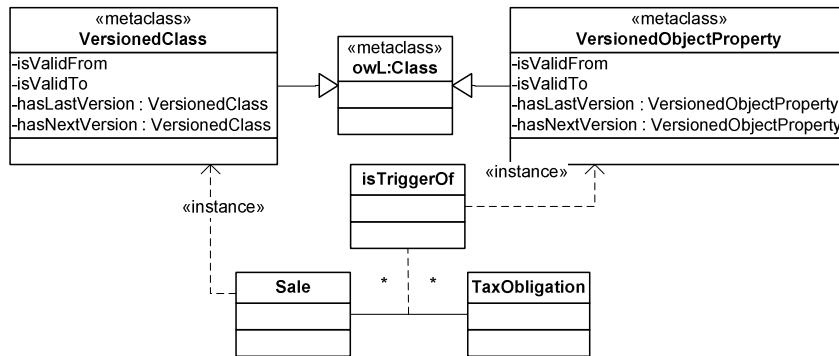


Fig. 3 Introduction of meta-level into tax legal ontology.

Now in contrast to the earlier analysis we are able to obtain a proper result - on the side of the seller there is no TaxObligation – at least not of type CivilDeedTaxObligation (which does not exclude that other side of transaction will be the subject of such obligation or that other type of taxes: VAT and PIT will become applicable). There is still also an issue of incorporating knowledge about conditions of utilizing different branch versions (parts) of the time-dimension cast of the knowledge model.

4 PUTTING THAT ALL TOGETHER

The example from the previous section represents a single important but not the only issue connected with the tax law knowledge representation. In fact we introduced only a very simple case with a number of simplifications. More to that we actually dealt with merely one type of tax (only mentioning others) – featuring relatively small number of norms and concise regulation. However the studying of a larger resource of legal acts of the domain and additional other source texts (professional press materials and legal advises) we came to formulate a number of assertion about

⁵ Unfortunately, there is no reasoner that fully support all OWL Full features – for the list of most common OWL reasoners, see <http://www.w3.org/2001/sw/WebOnt/impls#Implementations>

prospect system of tax legal status analysis for the eBusiness. On the basis of these assertions we are able to view an outline of model representing the structure of such a system. The model is shown on Fig. 4.

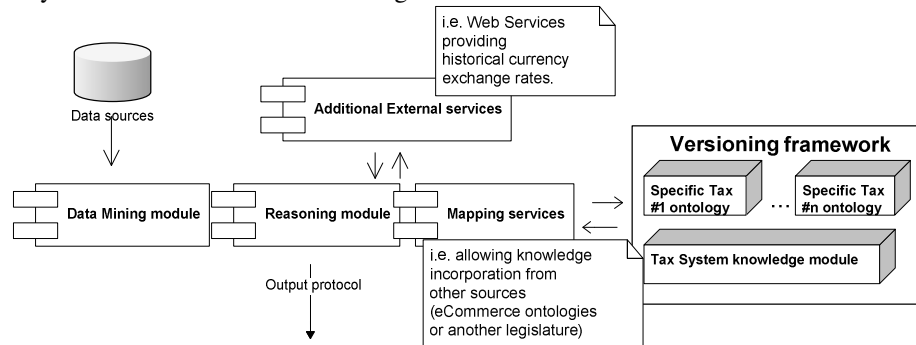


Fig. 4 Outline of tax analysis system model.

Though it is a sketch rather, we are certain that the model may well serve as a reference in most of the solution for a wide range of tax-issues tackling.

5 LESSONS LEARNED

In the paper besides short introduction to the field of legal knowledge modelling putting stress to legal ontologies and tax knowledge representation we portrayed the general assumption of Polish tax law. We also added an international context as we recognize its importance in the eCommerce and eBusiness world. On the top of that we presented a specific case pointing out difficulties that should be addressed by the system of eCommerce tax analysis. Then we described an outline of model of such system in order to solve more general problems than those given in the example. On the other hand we referred to the model of ontology evolution showing its new realization (in comparison to the authors' original one) via the extension on the meta level.

As the consequence we came to formulate the following postulation as the future recommendations: 1) Turbulent legal environments produce a requirement for versioning in real-life legal knowledge models; 2) We acknowledge the earlier formulated need for creating comprehensive highly specialized tax domain ontologies; 3) In order to come up with effective solutions a number of obstacles have to be overcome; 4) The suggested solution of utilizing the meta-classes to describe ontological elements' evolution is another interesting and particularly hopeful approach; 5) This approach may be successfully used in eBusiness tax analysis infrastructure.

REFERENCES

1. Benjamins R., Casanovas P., Breuker J., Gangemi A., (Eds.) Law and the Semantic Web.

- Legal Ontologies. Methodologies, Legal Information Retrieval, and Applications. Springer Verlag, LNAI 3369, Berlin, 2005.
2. Maedche, A. & Staab, S., Measuring Similarity between Ontologies. In: Proc. of the European Conference on Knowledge Acquisition and Management - EKAW-2002, Madrid, Spain, LNCS/LNAI 2473, Springer, 2002.
 3. Knublauch H., Musen M. A. & Rector A. L.. Editing Description Logic Ontologies with the Protégé OWL Plugin. International Workshop on Description Logics (DL2004), Whistler, Canada, 2004.
 4. www.wsmo.org/TR/d37/v0.1/20080125/d37v0.1_20080125.pdf (read on 17-02-2008).
 5. Gangemi A., Design patterns for legal ontology construction, Proceedings of LOAIT '07 II Workshop on Legal Ontologies and Artificial Intelligence Techniques, 2007.
 6. Papazoglou M. P. & Ribbers P., e-Business. Organizational and technical Foundations, Wiley & Sons, 2006.
 7. Breuker J., Boer A., Hoekstra R & van den Breg K., Developing content for LKIF: Ontologies and frameworks for legal reasoning, in T.M. van Engers (editor) Legal Knowledge and Information Systems. Jurix 2006, IOS Press 2006.
 8. Melz E., & Valente A., Modeling the Tax Code in R. Meersman (editor) On the Move to Meaningful Internet Systems 2004: OTM 2004 Workshops, Springer Verlag, LNCS 3292, Berlin 2004.
 9. <http://www.hypergrove.com/legalrdf.org/inventory.html> (read on 17-02-2008).
 10. http://ec.europa.eu/publications/booklets/move/17/txt_en.pdf (read on 17-02-2008).
 11. Ustawa z dnia 29 sierpnia 1997 r. - Ordynacja podatkowa. Dz.U. 1997 nr 137 poz. 926.
 12. <http://www.sejm.gov.pl/prawo/konst/angielski/kon1.htm> (read on 17-02-2008).
 13. Flouris G., On the Evolution of Ontological Signatures, Proceedings of BIS 2007 Workshops on Ontology Evolution (OnE'07), Poznan 2007, ISBN-13: 978-83-916842-4-5.
 14. Eder J. & Koncilia Ch., Modelling Changes in Ontologies, in R. Meersman (editor) On the Move to Meaningful Internet Systems 2004: OTM 2004 Workshops, Springer Verlag, LNCS 3292, Berlin 2004.