

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/280836104>

Enhanced Legal Model of Hungarian Labour Law

Conference Paper · September 2013

CITATIONS

0

READS

63

3 authors, including:



Akos Szoke

Budapest University of Technology and Economics

16 PUBLICATIONS 53 CITATIONS

[SEE PROFILE](#)



András Föhrécz

Budapest University of Technology and Economics

11 PUBLICATIONS 22 CITATIONS

[SEE PROFILE](#)

Enhanced Legal Model of Hungarian Labour Law

Abstract

National governments and large organizations have published laws on the web or intranets in order to provide information for their citizens, businesses and their employees. Although these web sites are great legal sources and these approaches can provide information how to interpret a legal text, precise understanding cannot be achieved with their apparatus.

We have developed the Emerald framework and methodology that provides new approaches for legal modelling than common rule-based systems by utilizing recently emerged Semantic Web technologies, algorithms focused on domain-specific problems and state-of-the-art presentation of complex models enhancing transparency. The framework provides tools to handle formal models and source documents jointly by linking different representations of the same information.

To obtain a proof-of-concept we modelled the Hungarian Act I. of 2012 on Labour Code.

Keywords

legal modelling, Metalex, ontology, RDF, SWRL

1 Introduction

National governments and large organizations have published laws on the web or intranets in order to provide information for their citizens, businesses and their employees (e.g. on government portal of the Netherlands¹, UK² and Hungary³). Although these web sites are great legal sources and these approaches can provide information how to interpret a legal text, precise understanding cannot be achieved with their apparatus.

¹ Dutch government websites: Laws and regulations, <http://wetten.overheid.nl/>

² The official home of UK legislation, <http://www.legislation.gov.uk/>

³ NJT: Hungarian national repository of laws and regulations, <http://njt.hu/>

We have developed the Emerald⁴ integrated legal modelling framework and methodology, which supports building legal knowledge models. It provides new approaches for legal modelling by utilizing recently emerged Semantic Web⁵ technologies, algorithms focused on domain-specific problems and state-of-the-art presentation of complex models enhancing transparency.

In our approach, the expressiveness and semantic richness of a legal text are increased at three levels. The framework provides tools to handle formal models and source documents jointly by linking different representations of the same information.

- The **first level** focuses on the standard representation of legal documents (structure, metadata, annotations and references). At this level semantic search and versioning are supported, the standard representation makes document translating and interchanging easier. Documents and document text fragments can be referenced and linked to other objects.
- The **second level** realizes conceptual modelling. We describe the strict terminology without defining inference axioms regarding to concepts. This level of modelling supports visualization and interpretation services. Direct and hidden relations between concepts can be discovered, visualized and explained. The concepts defined at this level can be linked to the document text fragments of the first level.
- The **third level** extends the formal terminology of the conceptual model with logical model, i.e. with logical rules that describe the normative knowledge of legal text sources as well as with application-specific data. At this level we can support inference services, e.g. legal advisor applications, consistency checking, test case evaluation applications.

2 Emerald Functional Overview

Emerald is an integrated framework that builds on Semantic Web standards and supports the whole lifecycle of legal modelling: we produce standardized XML documents from legal source texts [1, 2], extend them with metadata, annotations and resource links, build formal conceptual and logical models and develop inference-based applications on top of these models [3].

⁴ The work is supported in part by KMOP-2009-1.1.1. grant

⁵ Semantic Web on Wikipedia, http://en.wikipedia.org/wiki/Semantic_Web

Emerald consists of three main modules, a content visualizer, an editor and dialogue service.

2.1 Web Visualizer

This module can be used from a web browser for managing and visualizing legal source documents and conceptual models. It supports semantic search, temporal and language versioning and integration with the conceptual and logical model.

2.2 Desktop Editor

This module supports building formal models from legal text sources and to design reasoning applications on top of these models. It supports consistency and completeness check of legal models and user can create dialogue-based or batch applications for testing purpose or for building end-user services.

2.3 Web Dialogue

The module is a dialogue-based legal advisor application. An application built from legal model appears for users as a dialogue. The general knowledge of the legal model is completed with actual data given by the user until the system collects enough information to reach a conclusion or to calculate a numerical value. The inference mechanism can explain the reason of questions and results.

3 Emerald Technical Background

3.1 Document Representation

Legal source documents are transformed into the jurisdiction independent CEN Metalex⁶ format and managed in an XML-based document server to support semantic searching and indexing [4, 5, 6, 7].

3.2 Knowledge Representation

We comply with the current Semantic Web standards to enable legal information serving via the.

⁶ CEN MetaLex: Interchange Format for Legal and Legislative Resources, <http://www.metalex.eu/>

The Web Ontology Language (OWL) [10] is a standardised set of knowledge representation languages for authoring ontologies. From now on, we are referring OWL 2 (DL) as OWL, and interpret these ontologies using the Direct Semantics. The Direct Semantics is compatible with the SROIQ description logic, a fragment of first order logic with useful computational properties [13]. In our approach, the OWL allows explicit and formal representation of meaning of terms and relations between them. Therefore, it enables construction of machine-accessible models of ontologies underlying particular domains of knowledge. Important characteristics of OWL2 ontologies that they can be processed by a description logic (DL) reasoner. Major services offered by these reasoners are *subsumption testing* (whether or not a concept is a specialization of another concept) that can provide an inferred ontology class hierarchy from an asserted one, and *consistency checking* (whether or not all concepts refer to valid, non-empty categories).

Due to the nature of OWL, the more complex domain knowledge cannot be represented easily or cannot be represented in OWL at all: especially, the normative legal statements. The SWRL (Semantic Web Rule Language) is the most straightforward rule extension to OWL [12]. SWRL extends the set of OWL axioms to include Horn-like rules. Rules are generalization of axioms that allows overcoming certain limitations of OWL e.g. implementing cyclic conditions. However the model-theoretic semantics of SWRL extends OWL beyond the point of decidability and practical implementations, the standard trade-off for expressivity. Therefore a restrictive semantics called the DL-safe interpretation have been introduced [14], where variables in rules bind only to explicitly named individuals. The restriction is implicit in many rule systems and means that rules do not interfere with terminological knowledge directly. Please note, the deontic notions of normative knowledge (e.g. obligations, permissions and prohibitions) are not easily translated into this formalism, as the defeasible nature of deontic operators is not supported. A possible solution is to use a hybrid reasoner as in [15] where a new logical layer is introduced for deontic operators. In this paper we assume that the domain tolerates the usage of a monotonic formalism by reducing deontic operators into descriptions of permitted states of the world.

Conceptual models are represented as SKOS thesauri [9] or OWL 2 DL ontologies [10, 11] (we simply refer to OWL ontology) and reasoning is supported with Pellet⁷ reasoner. Emerald has its own rule language (based on SWRL [12]) with a concrete syntax derived from OWL Manchester syntax and the semantic is an extension of SWRL DL-safe language semantics.

⁷ Pellet OWL 2 Reasoner for Java Homepage, <http://clarkparsia.com/pellet>

3.3 Software Implementation

The Emerald Desktop Editor is built on the open source Java and Eclipse technology. The Web Visualizer and Web Dialogue modules are implemented as browser based applications using the GXT toolkit. In the Emerald backend system documents are stored in an XML database; knowledge elements are stored in an RDF database.

4 Emerald Modelling Methodology

A modelling methodology has been developed to increase the expressiveness and semantic richness of legal texts. Depending of the required services, there can be three levels of vertical modelling. Emerald offers tools to manage and visualize information from different levels in parallel, linking different representations of the same information.

4.1 Level 1 - Document standardization

In some cases it is enough to obtain legal documents with standard structure and metadata, with semantic search, version handling and cross referencing capabilities.

- **Step 1.1:** CEN Metalex conform XML structure
- **Step 1.2:** Document extended with metadata
- **Step 1.3:** Document extended with annotations
- **Step 1.3:** Document extended with resource links

4.2 Level 2 - Conceptual modelling

The next level of modelling identifies concepts (i.e. SKOS or OWL) and their relations in the legal document and links them to text fragments.

- **Step 2.1:** Taxonomy
- **Step 2.2:** SKOS thesaurus
- **Step 2.3:** OWL ontology

4.3 Level 3 - Logical modelling

The final level of modelling is creating rules and defining application-specific data (e.g. goals) based on these rules. Our reasoning method uses a hybrid – ontology and rule based – inference mechanism, rules operates on the elements of the ontology.

- **Step 3.1:** Structured rules
- **Step 3.2:** Formal rules
- **Step 3.3:** Application specific data – goals and questions

5 Modelling of the Hungarian Labour Code

To demonstrate legal modelling with Emerald we modelled the Hungarian Act I. of 2012 on Labour Code. The document-oriented and semi-automatic steps of modelling are applied to the whole act, while formal modelling steps refer to selected parts of the source text.

5.1 Level 1 - Document standardization

Step 1.1: Structured XML document

A standardized document structure is produced preserving the original content. Legal sources can be imported in HTML and are transformed to MetaLex conform XML format. In our implementation this is a semi-automatic process.

The original source text of Fig. 1 is transformed to structured format of Fig. 2. Structural elements are identified and so can be referenced.

Section 79

- (1) The right of termination without notice may be exercised, without giving reasons:
- a) by either party during the probationary period;
 - b) by the employer in connection with fixed-term employment relationships.
- (2) In the case of termination under Paragraph b) of Subsection (1), the employee shall be entitled to absentee pay due for twelve months, or if the time remaining from the fixed period is less than one year, for the remaining time period.

Fig. 1 Source text: Section 79 from the Hungarian Act I. of 2012 on Labour Code

```
<Article metalex-ext:lexid="art79">
  <Category><TextVersion xml:lang="en">§</TextVersion></Category>
  <Index><TextVersion xml:lang="en">79.</TextVersion></Index>
  <SubPart metalex-ext:lexid="art79par1">
    <Index><TextVersion xml:lang="en">(1)</TextVersion></Index>
    <List>
      <SentenceFragment>
        <TextVersion xml:lang="en">The right of termination without notice may be
exercised, without giving reasons:</TextVersion>
      </SentenceFragment>
      <SentenceFragmentSubPart metalex-ext:lexid="art79par1claa">
        <Index><TextVersion xml:lang="en">a)</TextVersion></Index>
        <SentenceFragment>
          <TextVersion xml:lang="en">by either party during the probationary
period;</TextVersion>
        </SentenceFragment>
      </SentenceFragmentSubPart>
      ...
    </List>
  </SubPart>
  ...
</Article>
```

Fig. 2 Structured text: Section 79 from the Hungarian Act I. of 2012 on Labour Code

Step 1.2: Document extension with metadata

Standard metadata are added to the source document. Metadata standards come from public vocabularies, like MetaLex for legal-specific terms (e.g. Publication date) and Dublin Core for general terms (e.g. Title). Fig. 3 demonstrates this extension of the example document:

```
URI: http://www.multilogic.hu/emerald/biblio/hu/parliament/act/2012;1/@/2012-07-01/hun
Title: Act I of 2012
Label: Labour Code
Regulation type: act
Issue date: 2012
Issue ID: 1
Country: Hungary
Language: en
Efficacy start date: 2012-07-01
Publication date: 2012-01-06
```

Fig. 3 Document metadata

Step 1.3: Document extended with annotations

Annotations are added to the source document. Emerald makes possible to interpret certain fragments – or the whole body – of the legal source text. The interpretation can be executed either based on the glossary or without it.

5.2 Level 2 - Conceptual modelling

Step 2.1: Taxonomy

The first step of conceptual modelling is preparing of a classification system from the concepts found in the source text.

Step 2.2: SKOS thesaurus

The high level SKOS thesaurus can be used for conceptual modelling either in itself or as a first step before extending the model to the full OWL capabilities. We can build the thesaurus either from scratch or based on existing public thesauri, e.g. EuroVoc⁸. We can build our own vocabulary reusing these concept descriptions. As an example we show the description of “probationary period”. The RDF fragment in Turtle syntax (see Fig. 4) tells us, that the concept has an alternative name “trial period”, “personnel administration” is a concept with a broader meaning, and “traineeship” is a concept that is somewhat related.

```
probationary_period rdf:type skos:Concept ;
    skos:prefLabel "probationary period"@en ;
    skos:altLabel "trial period"@en ;
    skos:broader personnel_administration ;
    skos:related traineeship .
```

Fig. 4 SKOS thesaurus

Step 2.3: OWL 2 Ontology

The OWL ontology can be used directly or after modelling with SKOS and extending that high level abstraction. OWL enables formal definition of concepts using logical axioms expressed in description logic. Fig. 5 demonstrates a part of the OWL 2 ontology:

⁸ EuroVoc, Multilingual Thesaurus of the European Union, <http://eurovoc.europa.eu/>

```

Class: TerminationWithoutNotice
  Annotations: rdfs:label "Termination without notice"@en
  SubClassOf: Termination

Class: TerminationWithNotice
  Annotations: rdfs:label "Termination with notice"@en
  SubClassOf: Termination

DisjointClasses: TerminationWithoutNotice, TerminationWithNotice.

```

Fig. 5 OWL 2 ontology

5.3 Level 3 - Logical modelling

Step 3.1: Structured Rules

In this step we create the so-called *structured rules* from normative legal text sources to provide easier reading for humans. A structured rule has a <consequence> if <precondition> form similar to formal logical rules, but consists of free text fragments. This form can be easily transformed to formal Emerald rule syntax. Fig. 6 demonstrates the structured rule form of Section 79 (1) of the source text:

```

the right of termination without notice may be exercised without giving reasons
IF      the initiator of termination without notice is employer or employee AND
        employee is in probationary status

the right of termination without notice may be exercised without giving reasons
IF      the initiator of termination without notice is employer AND
        employee's employment relationship is fixed-term

```

Fig. 6 Structured rules

Step 3.2: Formal Rules

The second structured rule is translated here in Fig. 7. Identifiers prefixed with a question mark are variables, which bind to a specific value during rule execution. In the example, “termination” and “initiator” are OWL object properties (relations), whereas “employer”, “fixed term employment” and “termination without giving reason” are OWL classes (categories).

```

Dialogue: D_vat_payment_obligation
Annotations:
  em:label "Has the employer the right to terminate my employment relationship?"@en
Goal: Gla
Annotations:
  em:template "The employer has the right to terminate my employment
relationship."@en
  has_the_right_to_terminate_employment_relationship(employer)
Goal: Glb
Annotations:
  em:template " The employer doesn't have the right to terminate my employment
relationship."@en
  {not has_the_right_to_terminate_employment_relationship}(employer) .

```

Fig. 8 Application specific data

6 Evaluation

Our legal modelling approach is based on three main elements: semantically enriched document, conceptual and logical formal model, as well as linking these different representations of the same information.

An Emerald knowledge base is a rule base where each and every entity is backed by semantic description in a formal ontology. Rules describe normative knowledge, which reflects the regulatory nature of law. So our approach is recommended for processing normative legal documents.

We have prepared so far 22 corporate regulations and 10 legislative resources in 2 language and about 4 time versions per documents. The time requirement of processing an average size document (10-25 pages) was 0,5-1 day for document standardization, 3-5 days for conceptual modelling, and 5-10 days for logical modelling.

Our approach aims at decreasing the ambiguousness of legal texts, increasing the probability of finding the relevant legal materials, and utilizing the application of legal reasoners. It is implemented both as a service for citizens and businesses and as a modelling environment for legal drafters.

References

1. A. Boer, R. Hoekstra, R. Winkels, T. van Engers, and F. Willaert. METAlex: Legislation in XML. In T. Bench-Capon, Aspasia Daskalopulu, and R.G.F. Winkels, editors, Legal Knowledge and Information Systems. Jurix 2002: The Fifteenth Annual Conference,

- Frontiers in Artificial Intelligence and Applications, pages 1–10, Amsterdam, 2002. IOS Press.
2. Enrico Francescon Maria Angela Biasiotti Giovanni Sartor, Monica Palmirani, editor. Legislative XML For The Semantic Web: Principles, Models, Standards For Document Management. Springer, 2011.
 3. András Föhrhéc and György Strausz. An ontology-based rule chaining algorithm for legal expert systems. In Computational Intelligence and Informatics (CINTI), 2011 IEEE 12th International Symposium on, pages 443–447, nov. 2011.
 4. Rinke Hoekstra. The MetaLex Document Server - Legal Documents as Versioned Linked Data. In Harith Alani and Jamie Taylor, editors, Proceedings of the 10th International Semantic Web Conference ISWC 2011, page 16. Springer, 2011.
 5. E. Francesconi. Technologies for European Integration. Standards-based Interoperability of Legal Information Systems. European Press Academic Publishing, 2007.
 6. IFLA Study Group on the Functional Requirements for Bibliographic Records. Functional requirements for bibliographic records : final report. 2009
 7. M. Duerst and M. Suignard. Internationalized Resource Identifiers (IRIs). RFC 3987 (Proposed Standard), January 2005
 8. W3C. Rdf primer, 2004.
 9. Sean Bechhofer and Alistair Miles. SKOS Simple Knowledge Organization System Reference. W3C recommendation, W3C, August 2009.
<http://www.w3.org/TR/2009/REC-skos-reference-20090818/>.
 10. W3C OWL Working Group. OWL 2 web ontology language document overview. Technical report, W3C, October 2009. <http://www.w3.org/TR/2009/REC-owl2-overview-20091027/>.
 11. Peter F. Patel-Schneider, Boris Motik, and Bernardo Cuenca Grau. OWL 2 web ontology language direct semantics. W3C recommendation, W3C, October 2009.
<http://www.w3.org/TR/2009/RECowl2-direct-semantics-20091027/>.
 12. Ian Horrocks, Peter F. Patel-Schneider, Harold Boley, Said Tabet, Benjamin Grosf, and Mike Dean. SWRL: A semantic web rule language combining OWL and RuleML. Technical report, World Wide Web Consortium, 5 2004.
 13. Ian Horrocks, Oliver Kutz, and Ulrike Sattler. The Even More Irresistible SROIQ. In Proc. of the 10th Int. Conf. on Principles of Knowledge Representation and Reasoning (KR2006), pages 57–67. AAAI Press, June 2006.

14. Peter F Patel-schneider. Safe rules for OWL 1.1. Knowledge Creation Diffusion Utilization, pages 1–4, 2008.
15. Joost Breuker, Saskia van der Ven, Abdallah El Ali, Marc Bron, Rinke Hoekstra, Szymon Klarman, Urosh Milosovics, Lars Wortel, and András Förhécz. Developing HARNESS. Estrella deliverable 4.6, University of Amsterdam, <http://relay.leibnizcenter.org/>, 2008.