Legal Knowledge and Information Systems
K.D. Ashley (Ed.)
IOS Press, 2013


doi:10.3233/978-1-61499-359-9-25

# Ontology-driven Data Acquisition: Intelligent Support to Legal ODR Systems

G. AROSIO [a], G. BAGNARA [b], N. CAPUANO [c], E. FERSINI [b] and D. TOTI [c]
[a] *Consorzio Milano Ricerche*
[b] *University of Milano-Bicocca*
[c] *University of Salerno*

**Abstract.** We describe a system for computer-assisted writing of legal documents via a question-based mechanism. This system relies upon an underlying ontological structure meant to represent the data flow from the user's input, and a corresponding resolution algorithm, implemented within a local engine based on a Last-State Next-State model, for navigating the structure and providing the user with meaningful domain-specific support and insight. This system has been successfully applied to the scenario of civil liability for motor vehicles and is part of a larger framework for self-litigation and legal support.

**Keywords.** computer-assisted writing, ontology, data flow, civil liability, algorithm

## 1. Introduction

According to the 2011 report disclosed by the European Parliament, the increasing trend in the use of Alternative Dispute Resolution (ADR) counts about 410.000 cases in 2006, 473.000 in 2007 and more than 500.000 in 2008. More recent and impressive statistics are related to the Italian context, with a particular focus on mediation (one of the available schemas for ADR). According to the last statistics provided by the Italian Ministry of Justice, about 215.000 cases have been addressed through ADR between March 2011 and December 2012. Italy's Central Bank has estimated a 16 billion euro loss in terms of Gross Domestic Product caused by the slowness of civil justice, highlighting the necessities of encouraging alternative resolutions of disputes both from a citizen's and "justice systems" point of view. These numbers have envisaged ICT to be the key action in this area, encouraging therefore to shift from Alternative Dispute Resolution to Online Dispute Resolution (ODR). ODR, originated from the synergy between ADR and ICT, is a type of dispute resolution involving technology and the Internet to facilitate and speed up the resolution of out-of-court disputes. Several initiatives have been investigated for supporting ODR [6,1,5]. One of the main limitations of these systems relates to the collection of information for enabling any decision: claims and requirements are collected by a fixed-structure template to be filled in by parties. eJRM, acronym of electronic Justice Relationship Management [3], represents an ongoing Italian initiative aimed at dealing with semantic representation and machine-learning reasoning mechanisms for improving the awareness of citizens to personally evaluate the outcome of a potential litigation, to be guided to a non-conflict settlement and to be assisted in selecting the potential legal support. In this context, a fundamental role is played by those mechanisms for acquiring information from citizens in order to enable either "artificial" or "human" reasoning mechanisms. For this purpose, in this paper we propose an Ontology-driven Data Acquisition system (ODA), which is part of the eJRM initiative and is basically a context-sensitive adaptive questionnaire: it is meant to mimic the exploratory behavior exhibited by mediator practitioners in order to acquire relevant information from citizens, thus allowing a radical improvement of two major processes: (1) Online Mediation, online management of activities related to the mediation process, and (2) Self-Litigation, capability of a citizen to autonomously classify, formalize and understand the potential outcome of a dispute. The proposed system has found a first application in litigation related to civil liability of motor vehicles.

## 2. Ontology-driven Data Acquisition system: ODA

A computer-based, self-administered, interactive questionnaire has been designed in order to collect useful information for enabling either Online Mediation or Self Litigation processes, where the system selects pertinent questions depending on a citizen's individual responses. The considered case study relates to civil liability regarding motor vehicles, a mediation discipline that is getting more and more relevant in Italy[1]. ODA is made up of two main components: (1) an ontological structure aimed at modeling the juridical knowledge related to a specific application domain and (2) a logical engine targeted at exploring the ontological structure in order to provide questions and collect responses to/from the citizen. These components are detailed in the following subsections.

### *2.1. Ontological structure*

The ontological structure proposed in ODA is based on the definition of a certain number of ontological subsets, or "sub-ontologies", each aimed at modeling a corresponding subset of questions for the adaptive questionnaire, have been defined, by drawing upon the methodological approach for defining ontological models exploited in the ARISTOTELE Project [4]. Its building blocks are listed below.

- Schema concepts. This category includes the following concepts: *Yes/No question*; *Multiple-choice question*; *Answer to multiple-choice question*; *Norm*; *Superclass concept*. These represent all the potential types a specific instance might belong to, and have been implemented as `owl:Class` in the OWL formalism[2].
- Instance concepts. Instance concepts represent instances of the aforementioned schema concepts, which can be acquired via questions presented to the user.
- Relationships/Predicates. The relationships modeled in ODA are the following: *assume*, whose object needs to be verified in the flow in order for the system to show the question concerning its subject; *assumeAND*, where all the object concepts sharing the same subject must be verified in order to show the question related to their subject; *assumeOR*, for modeling multiple-choice questions, by link-

[1] In 2011, 297.636 cases were raised at Justices of the Peace, who are judges competent for the determination of the compensation of damages arising from motor vehicles and craft accidents involving sums up to €20,000

[2] http://www.w3.org/TR/owl-features/

ing an instance of “Multiple-choice question” with mutually-exclusive instances of “Answer to multiple-choice question”; *violatedWhen* and *verifiedWhen*, linking a norm with a concept that could lead to its violation or compliance.

In order to prevent potential decidability problems within the semantic engine when dealing with constructs like `intersectionOf`, `unionOf` etc. coming from the OWL DL/OWL Full languages (semi-decidable and undecidable, respectively), RDF-plain constructs have been used to implement them. More specifically, *assume*, *assumeAND* and *assumeOR* are modeled with the OWL construct `owl:objectProperty`. Negation within a triple consists of a concept (subject), a relationship (predicate) and another concept (object) such that its object can have an associated boolean value (true or false): the boolean values have been modeled via the `rdf:Statement` construct and a specific `booleanValueForObject` property.

The modeling criteria are indeed generic and can be in principle applied (provided they are fine-tuned accordingly) to any type of questionnaire, just like their corresponding navigation logics that take advantage of them. Besides, the use of only simple RDF and OWL Lite constructs ensures a complete decidability of the inference processes carried out upon such an ontology.

The modeled concepts and predicates, necessarily of a normative nature, allow the questions to be presented to the user. Concerning our case study related to civil liability of motor vehicles, the basic concepts to be acquired refer to the applicability of the motor vehicle civil liability condition to the case raised by the user, as well as the existence of elements justifying the responsibility of the driver and/or owner of the motor vehicle involved. With the aim of verifying the actual responsibility of the driver, as well as contributory negligence of others involved in the accident (passengers, cyclists, pedestrians), we have introduced several questions tending to assess the behavior of all such subjects, as well as their violation of any norm or prohibition as set out by the Italian road traffic regulations. Lastly, a set of concepts has been modeled for the determination of the damages to be paid in accordance with the general criteria of the Italian Civil Code.

### 2.2. *Logical engine*

A logical engine able to provide a context-sensitive adaptive questionnaire has been defined with two main goals: (1) to explore the ontology earlier defined in order to gather concepts to be characterized by the user, and (2) to show the user the question related to the given concept and acquire his/her response. Each concept is directly connected with the corresponding question to be presented to the user. In order to provide a mechanism able to firstly explore the ontological structure and consequently manage questions/answers, the logical engine has been based on the “Last State-Next State” model (LSNS) [2]. According to this model, a given concept to be acquired (what question is currently processed) could lead to several potential subsequent concepts to be explored (where the system might go from here given the user’s input). The proposed logical engine implements a short-term memory approach based on predicate priorities. This allows us to provide questions related to a given concept only if prerequisite concepts have been previously verified (according to their ontological properties). The logical engine exploits some predicate priorities to explore the ontology and manage questions/answers.

Given a concept to be verified (*currentConcept*), the engine first explores prerequisite concepts related to it with a primary predicate (*assumeAND*) and subsequently provides the user with the corresponding questions (if not provided yet). This means that if more than one concept shares the same prerequisite (*childConcept*) that has already been acquired, the question related to the prerequisite will not be presented again. If the question type related to a *childConcept* is a multiple-choice, then the concept corresponding to the answer selected by the user is exploited to call recursively the logical engine. Alternatively, if the question type is boolean and the user response does not match the boolean value expected in the ontology, then the question flow is terminated in the current recursion cycle. This allows us to limit the number of questions shown to the user: when a *currentConcept* presumes several *childConcepts* that must be verified and one of these *childConcepts* does not match the user's response, then the questions related to the other *childConcepts* will not be presented. Once the whole first branch of *childConcept* has been acquired, the engine explores those concepts related to the *currentConcept* with a secondary predicate (*assume*) providing the user with the corresponding questions. While the multiple-choice questions are managed analogously to the *assumeAND* branch, the boolean ones do not need to be handled because all the *childConcepts* related to the *currentConcept* must be acquired. Once the *currentConcept* has been acquired, the question flow continues by exploring *parentConcepts* in order to assess whether any law violation has occurred.

## 3. Conclusion

In this paper we have presented ODA, an Ontology-driven data acquisition system, founded upon an underlying ontological structure and a Last-State Next-State local engine, which is able to provide a context-sensitive, adaptive questionnaire. ODA is meant to enhance Online Mediation and Self-Litigation processes, all the while helping users dramatically save time and effort in dealing with legal matters.

## Acknowledgments

This work has been supported by the eJRM Project (ref.: PON01_01286) and by the grant "Dote Ricercatori" - FSE, Regione Lombardia.

## References

[1] Bellucci, E. and Zeleznikow, J.: Family winner: A computerised negotiation support system which advises upon australian family law. In *ISDSS2001*, 2001, pages 74-85.

[2] Bouamrane, M.M., Rector, A., Hurrell, M.: Ontology-driven adaptive medical information collection system. In *Foundations of Intelligent Systems. Springer Berlin Heidelberg*, 2008, pages 574-584.

[3] Fersini, E., Archetti, F., Messina., E.: Towards a smooth e-justice: semantic models and machine learning. In *Proc. of the Conference on Integrated System Design and Technology, Mallorca*, 2012.

[4] Paolozzi, P., Del Nostro, P., Orciuoli, F., Ritrovato, P., Toti, D.: A Semantic-Based Architecture for Managing Knowledge-Intensive Organizations: The ARISTOTELE Platform. In *proceedings of the IEEE International Conference on Web Information System Engineering (WISE), ESE Workshop*, 2012.

[5] Uijttenbroek, E.M., Lodder, A.R., Klein, M.C.A., Wildeboer, G.R., van Steenbergen, W.R., Sie, L.L., Huygen, P.E.M. and van Harmelen, F.: Retrieval of Case Law to Provide Layman with Information about Liability: Preliminary Results of the BEST-project. In *Computable Models of the Law*, 2008.

[6] Zeleznikow, J., Meersman, R., Hunter, D. and van Helvoort, E.: Computer tools for aiding legal negotiation. In *proc. of the 6th Australasian Conference on Information Systems*, 1995.