Legal Compliance Support with an Ontology-based Information System

Albert Meroño-Peñuela
Institute of Law & Technology
U. Autònoma de Barcelona
Faculty of Law, Campus UAB
Bellaterra (08193), Spain
albert.merono@uab.cat

Núria Casellas
Institute of Law & Technology
U. Autònoma de Barcelona
Faculty of Law, Campus UAB
Bellaterra (08193), Spain
nuria.casellas@uab.cat

Sergi Torralba
Institute of Law & Technology
U. Autònoma de Barcelona
Faculty of Law, Campus UAB
Bellaterra (08193), Spain
sergi.torralba@uab.cat

Mario Reyes
S21sec. C/ Alcalde Barnils,
64-6, Bg. Testa, D, 1st floor
Sant Cugat Vallès (08174),
Spain
mreyes@s21sec.com

Pompeu Casanovas
Institute of Law & Technology
U. Autònoma de Barcelona
Faculty of Law, Campus UAB
Bellaterra (08193), Spain
pompeu.casanovas@uab.cat

ABSTRACT

The Internet and Information Systems evolution have dramatically increased the amount of information hold by governments and companies. This information can be very sensitive, specially regarding personal data, so governments and industries promote acts and guidelines in order to ensure privacy and data security. Thus, companies have to consider legal and Information Technology (IT) compliance. Nevertheless, compliance assessment is still a manual task performed by experts, but steps towards an automated compliance assessment, both in IT and legal, are in progress. In this paper we introduce the Neurona framework, a software application based on legal and security ontologies that aims at providing organizations with legal compliance support.

1. INTRODUCTION

Internet Information Systems have grown in complexity and performance featuring real time transactions, high bandwidth data flows and large databases. Furthermore, remote connections and distributed processes increase the risk of network attacks and accidental data losses. In this scenario, compromising information security may have critical consequences for customers and companies[1].

Governments and industries follow instruments from regulatory bodies and standardization institutions to ensure information security. Thus, companies face compliance from two perspectives: on the one hand, IT compliance of industry best practices and guidelines and, on the other hand, compliance of legal regulations. Currently, IT and legal compliance are verified mostly by experts, usually auditors or consultants, and it is still a manual task. This compliance assessment process can be extraordinarily expensive.

In the Information Era, one can think of an automated process that could perform some compliance assessment steps automatically, reducing associated costs. In [7] a logical formalism that specifies privacy policies is depicted; these policies can be verified in a federated digital identity scenario. Security companies such as RSA[2] and Cornerstone OnDemand[3] also offer some tools as a proposal to solve the IT compliance problem, with emphasis on policies and guidelines that usually emerge from industry best practices. Proposals for solving the legal compliance perspective are scarce or focused on access to data [3].

The aim of this work is to describe the Neurona[4] framework, a software application that uses OWL ontologies modeling legal knowledge to generate legal compliance reports of a company's state regarding privacy regulations, specifically the Spanish Personal Data Protection Act[5] (LOPD).

The paper is organized as follows. In section 2, we briefly introduce the legal knowledge methodologies applied, and some non-functional requirements found. In section 3, the system behaviour and its main use cases are described. Finally, in section 4, we give a set of conclusions.

[1]In 2009, the Spanish Data Protection Agency (*Agencia Española de Protección de Datos, AEPD*) imposed penalties for a total of 24.8M€[1].

[2]SIEM Automatic Compliance Reports, `http://www.rsa.com/node.aspx?id=3182`

[3]Enterprise Compliance Reporting, `http://www.cornerstoneondemand.com/compliance-reporting-tools`

[4]The Neurona project is funded by the Spanish Ministry of Industry, Tourism and Commerce and is developed by the Institute of Law and Technology (IDT-UAB) and S21sec.

[5]Ley Orgánica 15/99 de 13 de Diciembre de Protección de Datos de Carácter Personal.

2. LEGAL KNOWLEDGE

There are deep semantic differences between legal regulations and guidelines or best practices. There are existing or on-progress solutions for the IT compliance problem, such as UCF[6] or SCAP[7]. In IT regulations, very deterministic concepts such as *controls* or *safeguards* are specified, often in a logical formalism that can be checked in a real scenario with an algorithm. On the other hand, in legal regulations, like LOPD, more uncertain and open-textured concepts are found. These are much more difficult to implement in a way they can be checked by a validation algorithm. Ontologies were found suitable for this legal compliance scenario because concepts described in them can be defined in an expressive and more relaxed way that avoids subjective interpretations of legal regulations. Basics for ontology construction [5], legal requirements for compliance [6] and legal knowledge representations [2, 4] were applied.

In order to maintain reusable and changeable knowledge, the domain representation was split into two ontologies. The first one, DPCO[8], would define legal concepts contained in LOPD and relationships between them. Changes in this ontology may occur rarely, and its contents may be used only as an organizational taxonomy. The second one, DPRO[9], would specify a classification of possible desired or undesired situations regarding the application of the legal regulation, and their rules and constraints. DPRO imports DPCO for entity relationship discovery, but user instances and reasoning processes are entirely done in DPRO.

3. KNOWLEDGE MANAGEMENT SYSTEM

With the data structure depicted in section 2, we developed an OWL API-based tool to perform three basic use cases required for an automated ontology-based legal compliance assessment: **operative**, **knowledge management** and **intelligence**. The *operative* use cases gather information from company's assets and use it to generate ontology instances, which represent the company's current state regarding its assets and the dependency relationships between them. The *knowledge management* use cases require a maintainer role to load different versions of DPCO & DPRO in the OWL format when necessary (*e.g.* after a change in the Act). The *intelligence* use cases generate a legal compliance report, after having accessed OWL ontologies in a transparent way and having run the *Pellet* reasoner, which performs a classification of individuals in situations modeled in DPRO.

The starting scenario consists of a company that holds some files containing personal data of employees or customers (such as *name*, *ID*, *address*, *salary*, *account number* and *purchase history*), and wants to know its compliance state regarding those files and the LOPD normative.

First, a system administrator runs some *knowledge management* use case, in which a pair of OWL ontologies are loaded in the system and become the active legal knowledge base. Second, an operative-level user (*e.g.* a security controller) runs some *operative* use case, in which instances of some company assets (*e.g.* files or employees) and its state are created transparently into active ontologies. This can be performed manually or automatically, filling forms or executing net bots for data discovery, respectively. Finally, a strategic-level user (*e.g.* a Chief Compliance Officer) runs some *intelligence* use case, and reasoner classification results are shown in a report (*e.g.* if security measures of files are whether appropriate or not regarding the LOPD act).

[6]Unified Compliance Framework, `http://www.unifiedcompliance.com`

[7]The Security Content Automation Protocol, `http://scap.nist.gov`

[8]*Data Protection Conceptual Ontology*

[9]*Data Protection Reasoning Ontology*

4. CONCLUSIONS

We have shown a summary of the Neurona project, focusing our interest on legal compliance assessment of the LOPD act, applicable to most companies in Spain. We briefly discussed differences between existing IT compliance implementations based on *control tables* and *policy specifications*, and suitability of *ontologies* for the legal compliance.

In spite of system accuracy inherited from legal texts' open-textured concepts, ontologies allow us extracting basic concepts contained in legal texts without falling in interpretations and judicial decisions. Moreover, the use of ontologies has provided desirable software quality features: **reusability** (concept ontologies, for instance DPCO, can be used in a number of contexts outside the original application), **changeability** (changes in the domain only imply changes in ontologies, not in the program source code) and **ease of use** (almost any critical stakeholder can perform updates in the data model). With the use of ontologies, this tool could provide organizations with up-to-date monitoring of data protection regulations compliance. The work to evolve this system into a continuous report system for the company's legal compliance situation is still in progress.

5. REFERENCES

[1] Memoria 2009. Agencia Española de Protección de Datos.

[2] N. Casellas. *Modelling Legal Knowledge through Ontologies. OPJK: the Ontology of Professional Judicial Knowledge*. PhD thesis, Universitat Autònoma de Barcelona, 2008.

[3] D. el Diehn I. Abou-Tair, S. Berlik, and U. Kelter. Enforcing privacy by means of an ontology driven xacml framework. *Proceedings, 3rd International Symposium on Information Assurance and Security*, 2007.

[4] A. Gangemi, A. Prisco, M.-T. Sagri, G. Steve, and D. Tiscornia. Some ontological tools to support legal regulatory compliance, with a case study. *OTM Workshops*, LNCS 2889:607–620, 2003.

[5] A. Gómez-Pérez, M. Fernández-López, and O. Corcho. *Ontological Engineering*. Springer Verlag, 2003.

[6] A. K. Massey, P. N. Otto, L. J. Hayward, and A. I. Anton. Evaluating existing security and privacy requirements for legal compliance. *Requirements Engineering*, 2010.

[7] A. Squicciarini, M. C. Mont, A. Bhargav-Spantzel, and E. Bertino. Automatic compliance of privacy policies in federated digital identity management. *IEEE Workshop on Policies for Distributed Systems and Networks*, 2008.