A Legal Ontology to Support Privacy Preservation in Location-Based Services

Hugo A. Mitre, Ana Isabel González-Tablas, Benjamín Ramos, and Arturo Ribagorda

Computer Science Department, Carlos III Technical University of Madrid
28911 Leganes, Madrid
100059676@alumnos.uc3m.es, {aigonzal, benja1, arturo}@inf.uc3m.es

Abstract. During the last years many laws have been promulgated in diverse countries to protect citizens' privacy. This fact is due to the increase of privacy threats caused by the tendency of using information technologies in all scopes. Location Based Services (LBS) compose a situation where this privacy can be harmed. Even there exist mechanisms to protect this right in LBS, generally this services have not been developed over regulatory norms or if so, it has been in a partial way or interpreting those norms in a particular form. This situation could be a consequence of the lack of a common knowledge base representing the actual legislation in matters of privacy. In this paper an ontology of the main Spanish privacy norm is presented as well as the method used to construct it. The ontology is specifically aimed and applied to the preservation of privacy in LBS.

Keywords: Ontology, regulation, privacy, location-based services.

1 Introduction

Nowadays, the use and expansion of digital information technologies to most fields have caused an increase in the number of threats to the citizens' privacy. The evolution of mobile and positioning technologies has allowed the recent development of Location Based Services (LBS). LBS offer value-added services based on the geographic position of mobile devices. In [1] an extensive classification of LBS applications can be found. Some examples of their applications are emergency, security and medical services; navigation and information services; m-commerce; fleets management; proximity and entertainment services. Concrete scenarios of LBS are a taxi company that drives the requested taxis to the place the clients have called from (clients' position is obtained during the call), a user that wants to locate a friend, and a provider of local information that personalizes it depending on the position of the requesting user (e.g., the weather).

Although LBS can provide great benefits, implications for users' privacy arise because of their utilization of user's location information. Location can be considered private from regulation's point of view when this information is associated with any identified or identifiable natural person. Furthermore, the monitoring and position

R. Meersman, Z. Tari, P. Herrero et al. (Eds.): OTM Workshops 2006, LNCS 4278, pp. 1755 – 1764, 2006.

tracking of individuals allows the construction of locations profiles, making possible to recognize consumption habits, preferences, private costumes, and behavior of an individual. Alastair and Stajano define location privacy in [5] as the ability to prevent other parties from learning one's current or past location. Nowadays regulations of several countries include specific laws to protect citizens' privacy (e.g., most EU countries [19] and Japan [20]). Location privacy is a specific case also addressed by existing privacy regulations, either implicitly or explicitly. Therefore, it is necessary to provide in LBS enough guarantees to protect individuals' location privacy as required by regulations at the same time that the evolution of LBS is not limited [14].

Problem Description and Goal. Designing and integrating mechanisms that guarantee LBS's compliance with privacy regulation is a hard work, as some authors point out in [10]. On the one hand, designers should become experts in privacy legislation in order to interpret their precepts. On the other hand, they should know which security mechanisms are best for protecting location privacy, adapt them if necessary or design new ones in order to make LBS compliant with existing regulations. Although several researchers have proposed different location privacy mechanisms [2, 3, 4, 5, 6, 7, 9, 10, 11], they have not been developed using as base any regulatory norm or, if so, it has been in a partial way or interpreting the norms from a particular point of view. This situation can be a consequence of the lack of a common knowledge base that represents the current legislation in matters of privacy.

Christopher Welty and Nicola Guarino pointed out an important difference between an ontology and a data model. A conceptual model is an implementation that has to satisfy the engineering trade-offs of a running application, while The design of an onlotogy is to specify the conceptualization of the world underlaying such application [23]. In our case the conceptualization of the world can be represented by a specific domain. Ontologies can be used to reason about the objects in that domain and their relations. Ontology languages, such as OWL (Web Ontology Language [21]), allow the encoding of ontologies. Ontologies have been already used to represent several law domains such as in the TRACS system [15], and the E-POWER and E-COURT projects [17, 18]. An ontology for the privacy legislation domain will provide a common knowledge base that can be used to support privacy legislation interpretation and LBSs compliant with privacy regulations. In this paper a first approach to the development of an ontology for the privacy legislation domain is presented. In particular, the proposed ontology is based on the main Spanish norm for data privacy protection [13].

Organization of the paper. Next, in Section 2, some related works are analyzed. In Section 3, the construction method is presented and, in Section 4, the proposed ontology is described. Section 5 comprises the conclusions and future works.

2 Related Works

There exist several proposals that aim to solve location privacy problem. They can be classified in three groups according to the main security mechanism they are based on. A first group is composed by location privacy mechanisms based on attribute certificates [4, 9]. An attribute certificate is a digital token that binds privileges or

specific information (group membership, role, category, etc.) to entities. Proposals in [4, 9] use attribute certificates to authorize users several actions such as location information obtaining, storage, processing or third-party communication. The second group of location privacy mechanisms is based in policy systems [2, 3, 7-11]. Policies can be defined as a set of rules that specify the behaviour of a system. This kind of location privacy mechanisms use the policies to define which actions are allowed to whom, which obligations are derived, etc. The third group dissociates the user's identity from the information related to them [5, 6], for example using a pseudonym instead or controlling the granularity of the revealed information.

Privacy legislation are huge and complex, making quite difficult its interpretation by non-experts. On one hand, proposals in [2, 3, 7, 11] have been developed taking into account privacy norms or the precepts inferred from them. However, the resulting location privacy mechanisms are partially compliant with privacy regulatory norms. From the point of view of the authors of this paper this fact is due to that the legislation has been interpreted in a simplified way and following personal criterion. On the other hand, proposals in [4, 9] do not consider legislation at all. Finally, proposals in [5, 6] are compliant with current privacy regulations as dissociating the identity from data is excluded of the norm scope.

3 The Construction Method

The ontology construction method is based in the four main activities of the TERMINAE method [12]: corpus constitution, linguistic study, normalization and formalization. This process is described following.

The *documental corpus* is the main Spanish law for privacy protection [13], comprising 12 pages. Its main goal is to guarantee and protect personal data processing, public liberties and fundamental rights of physical persons, taking special care of their honour and personal and familiar intimacy.

A *linguistic study* was made on the lecture of the document:

- The candidate terms were extracted without considering if they were concepts, relations or instances.
- From the candidate terms plus the documental corpus, the main terms were extracted, that is, the concepts. It was made by using comprehensive lecture of repeated terms and inference of the concept.
- From the candidate terms, relations between main concepts were extracted considering the specific cases that make sense in legislation.

Each concept's *semantic is normalized*, searching again in the documental corpus to extract structure properties (e.g., date, position, number) and functional properties (e.g., enacted by).

Last, the ontology was *formalized* and refined using LRI-Core [16]. LRI-Core is a core ontology that covers the main concepts that are common to all legal domains. New concepts were aggregated according to necessities of a legal domain. The concept of mental-world was excluded as it is not necessary to model human thinking, wish or motivation but only a representation of the facts happening during LBS scenarios. As well, quality concept was also discarded as energy, strength or

substance concepts are not used. The ontology resulting from this step is based in roles, processes, physical concepts, abstract concepts and occurrences. It is described in next Section.

4 Description

LegLOPD (The Legal Ontology Domain) is compound by five top concepts extracted from the LRI-Core. They are: physical concept, occurrence, process, role, and abstract concept. To formalize the LegLOPD an useful subset of classes (Fig. 1) were taken from LRI-Core which defines a core ontology covering all legal domains.

Fig. 1. Class Hierarchy from LRI-Core taken as base for our legal ontology

In [16], the physical world evolves two main classes: physical objects and processes. Objects have mass, extension. Processes can change objects and consume energy. Energy is a current concept that could be interpreted as a metaphor, as the process of burning is a kind of energy. The same energy can be saved inside objects (batteries, petrol, etc.). In our ontology is not conceived any kind of energy, nor any kind of objects containing or consuming this energy, because is not necessary to represent simple facts.

Our approach will be explained related to LRI-Core (fig. 1) in next paragraphs of this section.

The objects render the physical world stable and observable. Inside the term **physical concept** we find time and space, both can be part of a process where any object can participate. Concepts added under the branch of physical concept (Concepto_Físico) in the figure 2 are explained in next sentences. Artefact (Artefacto) class represents all physical objects that treat digital information. Physical medium (Medio_Físico) represents the physical part of services (could be a LBS), giving also information treatment. We decide to add physical support (Soporte_físico) in order to save information into files (Fichero) or repositories (Repositorio). The concept natural object represents physical persons (Persona_física) and groups (Grupo). A group of persons could be an Organization, Administrative organ, etc.

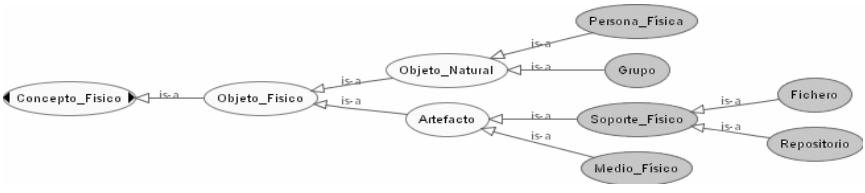

Fig. 2. Classes added to the physical concept of LRI-Core (Concepto_Físico)

The **abstract concept** is the clearest concept in common sense. A few mathematical concepts are known such as collections, sequences, and count numbers. Manipulating quantities or data structures is not common in the legislation.

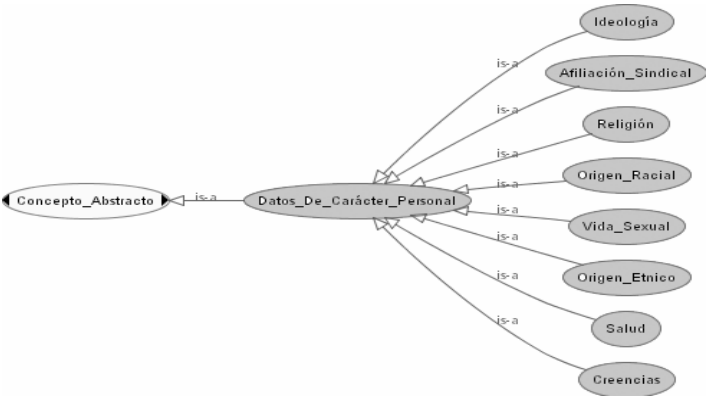

Fig. 3. Classes added to the Abstract concept of LRI-Core (Concepto_Abstracto)

The abstract concept sets numbers and structures to make up references associated to files, dates, repositories, etc. We added the concepts private data (Datos_De_Carácter_Personal)(fig 3), which is the essential structure to be protected. Sexual life (Vida_sexual), ideology (Ideología), health (Salud), syndical affiliation

(Afiliación_Sindical), racial origin (Origen_Racial), beliefs (Creencias), religion (Religión), and ethnic origin (Origen Etnico) are considered private data structures.

Roles in [16] are concepts that can be played by entities. It happens when the entity fits the role behavior, where instances are players. In addition, roles are properties and have dynamic properties. E. g. a role can be played by several entities, simultaneously or in different times. An entity can change its role or play multiple roles. Let's suppose that a physical person is responsible for private files treatment, being outside European territory. The responsible person has the obligation to asign a representative like an administrative organization. At the same time, this administrative organization could be responsible for other privacy file. This example shows two roles played by the same administrative organization, being at the same time representative and responsible of distinct private files.

We consider that roles can be assigned to physical objects as an artefact or natural objects. An object can take more than one role, and it also can change its role according to its behaviour. The subclasses added to artefact class can play the roles included in subclasses of the concept object role (Rol_Del_objeto). These are several examples: a *file* can take the role of private file (Fichero_de_titularidad_privada), public file (Fichero_de_Titularidad_Pública), forces and safety body file (Fichero_de_fuerzas_y_Cuerpo_de_seguridad), etc.; a *repository* (Repositorio) can play the role of the data protection general register (Registro_General_De_protección_De_Datos); and the *physical medium* as data treatment medium (Medio_de_Tratamiento_de_Datos). Furthermore, physical persons and groups, subclasses of natural objects playing roles such as legal roles or actors, e. g.: a person could be an actor (Actor) playing the role of affected (Afectado); a group could be an administrative organization (Organo_Administrativo).

Our ontology uses the **process** concept as an action or an activity. In LRI-Core, process can be classified according to two views: the formal kind of change (transformation, transduction, and transfer) and the kinds of objects involved (e.g. movements are the change of position of objects). [22] mentions that action communication has not the meaning of physical influence, but rather influenced by mental state. A person transfers his/her intentions for motives, personal plans, and beliefs to other persons. We discuss about the decision of this affirmation, and we conclude telling that they infer that an action is caused by intention. This is useful to find person culpability, but in our case, the aim is not to discover reasons for personal intentions, but rather to take the reason as a simple fact. E. g. personal intentions to eat with friends could be a cause to use located services, and locate the nearest restaurant. This information is too much for us, if we only want to know that a user makes use of located services to find a restaurant. It is a question of informing the user of the rights and obligations that imply the type of LBS that offers the service, and not to find guilts.

Actions can be provoked or caused by actors. These are some examples: the action of Reclaim (Reclamación) could be provoked by the affected (Afectado); Inscription can be caused by a representative.

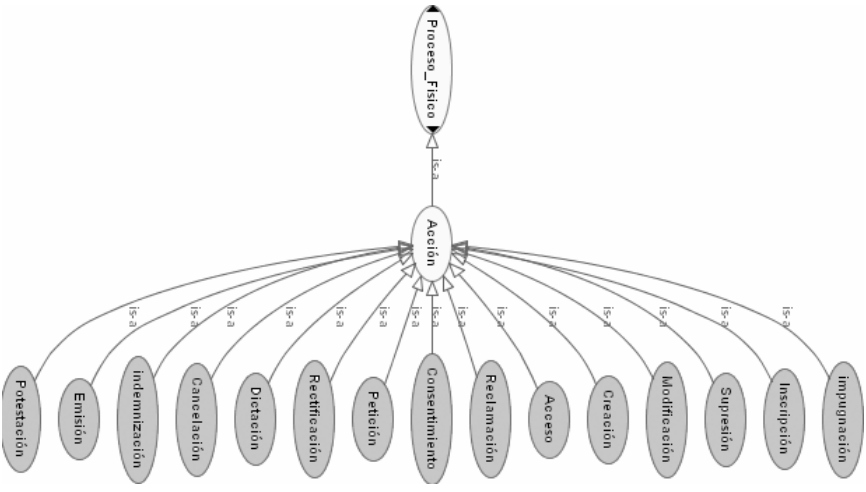

Fig. 4. Classes added to the process concept of LRI-Core (Proceso)

Occurrences capture all these aspects related to the execution of the scenes, and cover objects and processes. In a scene there may be events, states, situations, space and time references, and moments.

Events and States are occurrences. Both have a moment of execution in time. An event is caused by some physical object, and could happen after and before a situation (e.g. a service *transmits* inform). The state for us is, when an object is kept in some activity until any object provokes an event (e.g. a service keep in *treatment* information). The situations are caused by the actions of the actors on objects in a moment, and among the situations events can pass before and after this situation (e.g. between *verifies* register and *transmit* inform events is a information collection situation).

Occurrence concept is the most important concept to us, because it can help to represent the execution of a scene. The legal domain ontology LegLOPD is being constructed to represents scenarios. In the next figure, we describe a teorical scenario in order to understand the occurrence world in LBS. At the left side there are the physical objects and its roles: physical person as affected, and other physical person as interested; file as private file; and physical medium as data treatment medium. This scenario shows two actors, the interested that tries to collect privacy information of the affected by data treatment medium (e.g. an LBS). The physical medium makes use of the privacy file.

The order of execution is the following one: (Event 1) the interested makes a *request* register; (State 1) the interested is then in *data transferring* state; (State 2) the data medium treatment is in *treatment in recording* state; (Event 2) the private file is also an automated file (role) and *generates* an inform; (Event 3) the data medium treatment *transmits* inform; (State 3) the data medium treatment is in *treatment notification* state; (Event 4) the interested *consents* a treatment; (Event 5) the private file *registers* data; (Event 6) the interested makes a data *request*; (State 4) the data medium treatment, staying in *consulting data*; (Event 7) the private file *verifies* the

register; (event 8) the data medium treatment *transmits* an inform; (State 5) the data medium treatment *consents requesting*; (Event 9) the affected *revokes* inform; (State 6) the affected is in the state *treatment in blockade*; (Event 10) the data medium treatment *revokes* inform; (State 7) finally the affected is in state *not authorized access*.

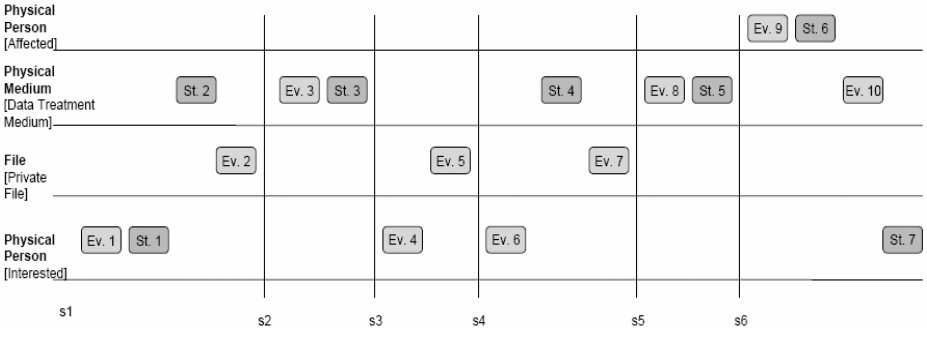

Fig. 5. This example shows a scene representing the occurrence world

To give an automatic reasoning of the legal domain can be necessary to be based on the occurrences. In [22] an approximation that allows an automatic analysis of cases described in terms of ontology. The proposed system is known as DIRECT. This system has to find casual and intentional relations in order to enable automatic analysis. This approach could be considered for our legal ontology, but oriented to causal relations.

5 Conclusions and Further Work

Nowadays, the use and expansion of digital information technologies to most fields have caused an increase in the number of threats to the citizens' privacy. Therefore, regulations of several countries have included specific laws to protect citizens' privacy. The evolution of mobile and positioning technologies has allowed the recent development of Location Based Services (LBS). Although LBS can provide great benefits, implications for users' privacy arise because of their utilization of user's location information.

Although several researchers have proposed different location privacy mechanisms [2, 3, 4, 5, 6, 7, 9, 10, 11], they have not been developed using as base any regulatory norm or, if so, it has been in a partial way or interpreting the norms from a particular point of view. This situation can be a consequence of the lack of a common knowledge base that represents the current legislation in matters of privacy.

In this paper a first approach to the development of an ontology for the Spanish main privacy law [13] has been presented and the method used in its construction (based on the TERMINAE method [12]). The proposed ontology has been developed using as base LRI-Core.

Future plans include finishing the legal domain ontology LegLOPD, and investigating other techniques for automated analysis in order to build an agent. Location-based systems and location services will be able to collaborate with the agent on the requests of the user in order to fulfil the law on protection of personal data for both parts.

References

[1] G. M. Giaglis, P. Kourouthanassis, and A. Tsamakos.: Towards a Clasification Framework for mobile location services, pages 67-85. Idea Group Publishing.

[2] Ana Isabel González-Tablas Ferreres.: Arquitectura de servicios de acreditación y sellado espacio-temporal .PhD thesis, Universidad Carlos III de Madrid, 2005.

[3] M. Langheinrich.: A privacy awareness system for ubiquitouscomputing enviroments. In the proceedings of 4th International Conference on Ubiquitous Computing (Ubicomp'02), pages 237-245, 2002.

[4] C. Hauser and M. Kabatnik.: Towards privacy support in a global location service. In the proceedings of the IFIP Workshop on IP and ATM Traffic Management (WATM/EUNICE 2001), 2001.

[5] A. R. Beresford and F. Stajano.: Location privacy in pervasive computing. IEEE Pervasive Computing, 2003.

[6] M. Gruteser and D. Grunwald.: Anonymous usage of location-based services through spatial and temporal cloaking. In the Proceedings of ACM/USENIX International conference on Mobile Systems, Applications, and Services (MobiSys), 2003.

[7] A. I. Gozález-Tablas, L. M. Salas, B. Ramos and A. Ribagorda.: Providing personalization and automation to spatial- temporal stamping services. In the Proceedings of the 16th Internacional Workshop on Database and Expert Systems Applications (DEXA'05), 2005.

[8] E. Snekkenes.: Concepts for personal location privacy policies. In the proceedings of the 3rd ACM conference on Electronic Commerce. ACM Press, 2001.

[9] U. Hengarther and P. Steenkiste.: Implementing access control to people location information. In the Proceedings of the Ninth ACM Symposium on Access Control Models and Technologies (SACMET'04), 2004.

[10] G. Myles, A. Friday, and N. Davies.: Preserving privacy in enviroments with location-based applications. IEEE Pervasive Computing, 2003.

[11] A. Gajparia, C. J. Mitchell and C. Y. Yeun.: Information Preference Authority: Suppoting user privacy in location based services. In the proceedings of the 9th Nordic Workshop on Secure IT-Systems (NordSec 2004), 2004.

[12] Sylvie Despres, Sylvie Szulman.: Construction of a legal Ontology from a European Community Legislative Tex., In T. Gordon (ed.), Legal Knowñedge and Information Systems. Jurix 2004: The Seventeenth Annual Conference. Amsterdam: IOS Press, 2004, pp. 79-88.

[13] Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal.

[14] Dr. Robert P. Minch.: Privacy Issues in Location-Aware Mobile Devices. Proceedings of the 37th Hawaii International Conference on System Sciences, 2004.

[15] Joost Breuker and Radboud Winkels.: Use and reuse of legal ontologies in knowledge engineering and information management. ICAIL 2003 Workshop on Legal Ontologies & Web Based Legal Information Management, April 2003.

[16] Joost Breuker.: Constructing a legal core ontology: LRI-Core. University of Amsterdam, 2004.
[17] Joost Breuker, Abdullatif Elhag, Emil Petkov, and Radboud Winkels.: IT Support for the Judiciary: Use of Ontologies in the e-Court Project, 2002.
[18] Boer, R. Hoekstra, R. Winkels, and T. van Engers.: ^{META}*lex*: Jurisdiction and Language. In Monica Palmirani, Tom van Engers, and Maria A. Wimmer, editors, Proceedings of the E-Government Workshop in conjunction with JURIX 2003, pages 54-66. Universitätsverlag Rudolf Trauner, December 2003.
[19] European Union (EU) Data Protection Directive of 1995, Rebecca Herold, CISM, CISSP, CISA, FLMI, May 2002.
[20] Japanese Personal Information Protection Law, Tokyo Aoyama Aoki Law Office,Baker & McKenzie, Attorney at Foreign Law Office, Registered Associated Offices, June 17, 2004.
[21] OWL Web Ontology Language Overview. W3C Recommendation 10 Feb 2004.
[22] Joost Breuker and Rinke Hoekstra.: 'DIRECT: Ontology-based Discovery of Responsibility and Causability in Legal Case Descriptions' in T. Gordon (ed.), Legal knowledge and Information Systems. Jurix 2004: The Seventeenth Annual Conference.Amsterdan: IOS Press, 2004, pp. 59-68.
[23] C. Welty et N. Guarino.: Supporting Ontological Analysis of Taxonomic Relationships. Data and Knowledge Engineering, 2001. Également : LADSEB-CNR Int. Rep. 08/2001.