Proceedings of the International Multiconference on
Computer Science and Information Technology, pp. 813 – 820

ISBN 978-83-60810-14-9
ISSN 1896-7094

A Semantic Framework for Privacy-Aware Access Control

Georgios V. Lioudakis, Nikolaos L. Dellas, Eleftherios A. Koutsoloukas,
Georgia M. Kapitsaki, Dimitra I. Kaklamani, Iakovos S. Venieris
National Technical University of Athens, Heroon Polytechniou 9, 15773, Athens, Greece
Email: {gelioud, ndellas, lefterisk, gkapi}@icbnet.ntua.gr, dkaklam@mail.ntua.gr, venieris@cs.ntua.gr

***Abstract*—The issue of privacy is constantly brought to the spotlight since an ever increasing number of services collects and processes personal information from users. In fact, recent advances in mobile communications, location and sensing technologies and data processing are boosting the deployment of context-aware personalized services and the creation of smart environments but, at the same time, they pose a serious risk on individuals' privacy rights. Being situated in the realms of legal and social studies, the notion of privacy is mainly left, concerning its protection, to legislation and service providers' self-regulation by means of privacy policies. However, all laws and codes of conduct are useless without enforcement. Based on this concept, this paper presents a framework conceived on the basis of privacy legislation. It uses a semantic model for the specification of privacy-aware data access rules and a middleware system which mediates between the service providers and the data sources and caters for the enforcement of the regulatory provisions.**

I. Introduction

ON *the Internet, nobody knows you are a dog*, according to the famous 1993 Pat Steiner cartoon in The New Yorker, which has been very frequently cited in order to emphasize the potential for anonymity and privacy that the Internet was supposed to offer. However, the reality seems to be rather different; in fact, more than a century after the first essay identifying that privacy as a fundamental human right was endangered by technological advances [1], never before in history the citizens have been more concerned about their personal privacy and the threats by the emerging technologies [2].

The potential impact of contemporary Information and Communication Technologies on the privacy rights of the users is regarded as being among their most evident negative effects. The advances in mobile communications, location estimation and sensing technologies, along with data storage and processing technologies, have expanded the sphere of electronic services' provision and digital facilities from the Web to pervasive smart environments. They create impressive perspectives of rich, highly personalized and coherent services, information and computation ubiquity and, thus, spur an information revolution that brings significant improvements of the citizens' quality of life. On the other hand, they pose serious risks on the privacy rights of the data subjects; the personal data collection scale is augmented, information access, processing, aggregation, combination and linking are facilitated and new, sometimes even more sensitive, types of data are collected. A stream of data about individuals pours into data warehouses, while personal information is increasingly viewed as a valuable financial asset which is a subject of trading.

The service providers usually express their practices by means of privacy policies. Privacy policies concern the formal specification of an organization's business practices regarding the collection and the consequent use of personal data. The privacy policies are supposed to be restricted according to fair information principles and to comply with the relevant legal framework. Privacy legislation dictates how personal data should be treated after their provision by the data subjects to service providers and other processing entities, defining in essence the requirements for the privacy-aware management of personal data through their whole life cycle.

The Platform for Privacy Preferences (P3P) W3C specification [3] has been the first initiative towards this direction, providing a way for a web site to encode its relevant practices and to communicate them to the users that visit the site. Since proposed, P3P has received broad attention from both industry and research community, but it has also been subject of criticism from the current technical work, e.g., [4]. The major issue with P3P is the lack of the mechanisms for the enforcement of the specified privacy policies. In essence, P3P formalizes privacy promises given to the users for fair information practices; nevertheless, after their disclosure to a service provider, there is no guarantee about the fate of a user's personal data. Besides, there are numerous cases where the real practices contradict to well-stated privacy policies, e.g. [5], [6].

The challenge of enforcing a privacy policy has been thoroughly examined and several different solutions have been proposed, e.g., by IBM [7], OASIS [8] and Hewlett Packard [9]. These frameworks mainly focus on enterprise environments and provide the means for the automation of the privacy policies enforcement. The means for achieving this is to apply privacy-aware access control mechanisms which enhance traditional Role-Based Access Control (RBAC) mod els with additional, privacy-related aspects, such as the pur

This work was partially supported by the EU IST project DISCREET.

pose for data collection, retention periods, users' consents, notifications, etc.

However, all these solutions have their weak points. First, although they manage to address the issue of privacy policies internal enforcement within an organization to a great extent, they fail in providing the necessary guarantees for fair information practices to the users. In fact, since an organization possesses some personal data, their use or abuse by means of processing and disclosure are still based on good intents. Misuse may occur by a malicious employee with legitimate access to sensitive data or by any form of direct access to the data that bypasses the privacy protecting system. Second, the privacy policies specified in the context of these frameworks cannot be efficiently audited and verified as far as their regulatory compliance and consistency is concerned. Even an organization with the best intentions may specify a privacy policy that is not legislation-proof. Third, the specification of complex privacy policies and the continuous process of keeping them up-to-date introduce significant economical, operational and administrative overhead to an organization.

In the light of the above issues, this paper proposes a framework for the enforcement of privacy policies by the providers of e-services that are based on the legislation. The main concept behind the framework is the formal and detailed codification and specification of the regulatory provisions by a Privacy Authority into a single privacy policy document and its automatic dissemination to the providers. This unique privacy policy constitutes the technical translation of the privacy principles and regulations and overrides any other privacy policy defined by a provider. For its enforcement, a middleware architecture is introduced, that acts as a three way privacy mediator between the law, the users and the service providers. Its main component is the D-Core Box, a privacy proxy installed at the service provider's premises but totally controlled by the Privacy Authority. The D-Core Box stores any personal data, keeping them separated from the provider. Access to the data is granted based on the legislation originated privacy policy, as well as the relevant preferences expressed by the users via an associated mechanism that the framework offers.

The remainder of this paper is structured as follows. Section II provides some insights on the legal aspects of privacy; it codifies the legal privacy principles that form the requirements for the proposed framework. Section III describes the Ontology of Privacy, the semantic information model that constitutes the basis of the proposed approach, since it contains the rules stemming from the legislation. In Section IV, the means for enabling the users to specify their privacy preferences are outlined. Section V describes the middleware architecture that undertakes the task of enforcing the privacy rules, both regulations- and user- originated. The paper concludes in Section VI.

II. Personal Data Protection Legislation

The starting point to obtain a formal modeling of the privacy legislation and also to design a technology capable of being privacy compliant and at the same time ensuring enforcement of privacy law provisions is identifying the regulatory requirements to be complied with.

A significant milestone in the privacy literature has been the codification of the fundamental privacy principles by the Organization for Economic Co-operation and Development (OECD), in 1980 [10], as this codification lays out the basis for the protection of privacy. The OECD principles are reflected in the European Directive 95/46/EC [11], which enforces a high standard of data protection and it is the most influential piece of privacy legislation worldwide, affecting many countries outside Europe in enacting similar laws. It is particularized and complemented with reference to the electronic communication sector by the Directives 2002/58/EC [12] and 2006/24/EC [13], which impose explicit obligations on network and service providers to protect the privacy of users' communications. The European Directives constitute the basis for the following summary of the main regulatory data protection requirements to be taken into account for the specification of the proposed framework.

A. Regulatory Requirements

The following requirements are stemming from the European personal data protection legislation:

Lawfulness of the data processing : The system should be able to examine whether the data processing complies with applicable laws and regulations.

Purposes for which data are processed : The system should provide the means for identifying the data processing purposes, which must be lawful and made explicit to the data subject (namely the subject whose data are processed). Moreover, it should be able to check these purposes to avoid that data processed for a purpose may be further processed for purposes that are incompatible with these for which data have been collected.

Necessity, adequacy and proportionality of the data processed : The system should be able to guarantee that only the data functional, necessary, relevant, proportionate and not excessive with regard to the sought processing purpose are processed.

Quality of the data processed : The system should provide that the data processed are correct, exact and updated. Inaccurate data must be deleted or rectified; outdated data must be deleted or updated.

Identifiable data : The system should provide the means for keeping the data processed in identifiable form only for the time necessary to achieve the sought processing purpose.

Information to the data subjects; consent and rights of the data subject : The system should be able to provide for informing the data subject that his data are processed according to applicable data protection legislation. Moreover, the system should guarantee that when requested by applicable data protection legislation, the data subject's consent to the data processing is required, and that the data processing is performed according to the preferences expressed by the data subject. In addition, the system should enable the data subject to exercise the rights acknowledged by applicable data protection legislation in relation to intervention in the data processing (for example the right to access data, to ask for data rectification, erasure, blocking, the right to object the data processing, etc.).

Data security and confidentiality : The system should be secure in order to guarantee the confidentiality, integrity, and availability of the data processed. Moreover, the system should provide that the listening, tapping, storage or other kinds of interception or surveillance of communications and the related traffic data may be performed only with the data subject's consent or when allowed by applicable legislation for public interest purposes.

Traffic data and location data other than traffic data; special categories of data : The system should be able to guarantee that the processing of special categories of data (for example traffic or other location data, sensitive and judicial data) is performed in compliance with the specific requirements that the applicable data protection legislation sets forth for said categories of data.

Access limitation : The system should provide for an authorization procedure that entails differentiated levels of access to the data and for recording the accesses to the data.

Data storage : The system should be able to automatically delete (or make anonymous) the data when the pursued processing purpose is reached or in case of elapse of the data retention periods specified under applicable legislation.

Notification and other authorizations from competent Data Protection Authority : The system should be able to monitor compliance with the notification requirement and with the provisions on the authorizations of competent Data Protection Authority. Moreover, the system should provide for means that allow communications between the system and the competent Data Protection Authority.

Supervision and sanctions : The competent Data Protection Authority should be provided with the means for supervising and controlling all actions of personal data collection and processing.

Lawful interception : The competent public authority should be provided with the means to perform interception only when this is allowed by applicable laws and regulations and according to the conditions therein set forth. The necessary "hooks" for the lawful interception should under no circumstance become available to other not authorized third parties.

B. The Regulations in a Nutshell

From the regulatory requirements presented above, the following facts are extracted:

The role of the users : The users are granted certain rights; the right to be informed regarding the collection or processing of personal data, to be asked about their explicit consent, to access their data. Additionally, they should be able to specify their privacy preferences and affect this way the service provision procedure, with respect to privacy.

The role of the Authorities : The legislation grants the Data Protection Authorities with certain rights and competences. These include the notification of the Authority, the supervision of the procedures and the means for performing Lawful Interception. That is, the Authority should be able to interact with the system.

The role of semantics: The semantics play a very crucial role in what can be characterized as "privacy context". On the one hand, each data item should be treated according to its special type. On the other hand, very important is the purpose for which the data are collected and processed.

Access control: The access to the data should be controlled. Beyond the legacy Role-Based Access Control models, decisions concerning access to personal data should take into consideration the semantics characterizing each privacy session.

Complementary actions: Access to the data should be accompanied by certain behavioral norms of the system. These include the information of the users or the request for their explicit consent, the notification of the Authorities, the automatic enforcement of data retention periods, as well as the adjustment of the detail level of the data.

Security: Naturally, in order for the personal data to be protected, the means for securing their transmission and storage should be taken; security always constitutes the bottom line for privacy protection and this paper takes as granted the availability of the corresponding means.

III. Semantic Information Model

The modeling of the privacy legislation is achieved using a semantic information model that associates personal data, services and actors with explicitly defined regulatory rules. In that respect, the approach taken is to express any related information by means of an ontology, namely the Ontology of Privacy, which is implemented using the W3C Web Ontology Language (OWL) [14]. The vision is that the ontology should be as detailed as possible in terms of the various types of personal data and the types of services, so that the widest range of services and situations when personal data are involved can be covered. This is similar to what the Common Procurement Vocabulary (CPV) [15] represents for public procurement in Europe; it provides an exhaustive –almost semantic– list of several thousands of products that can constitute subject of public procurement.

In order to associate the personal data with specific processing tasks, the identification of the particular type of each personal data item is necessary. Moreover, in order to define the appropriate rules that will regulate the processing of a personal data item with respect to the purpose for which the information is provided by the user or requested by the service provider, a similar taxonomy of the provided services must be present. These taxonomies constitute separate subgraphs of the ontology. Therefore, the Ontology of Privacy provides a detailed vocabulary of personal data types and services' types, structured in an hierarchical way with well defined inheritance rules, that enables the system to associate all privacy related decisions to semantically specified notions. An equivalent taxonomy is needed for the involved actors; however, here we consider a very simple model, comprised of three actors: the user, the service provider and the Privacy Authority.

Regarding the personal data subgraph, all the types are defined as instances of the `PersonalData` OWL class. Inheritance hierarchies, as well as other relationships between personal data are defined using OWL properties. The first hierarchy specifies the inheritance of characteristics, referring to legislation-originating rules that regulate the collection and processing of personal data. The "root" personal data

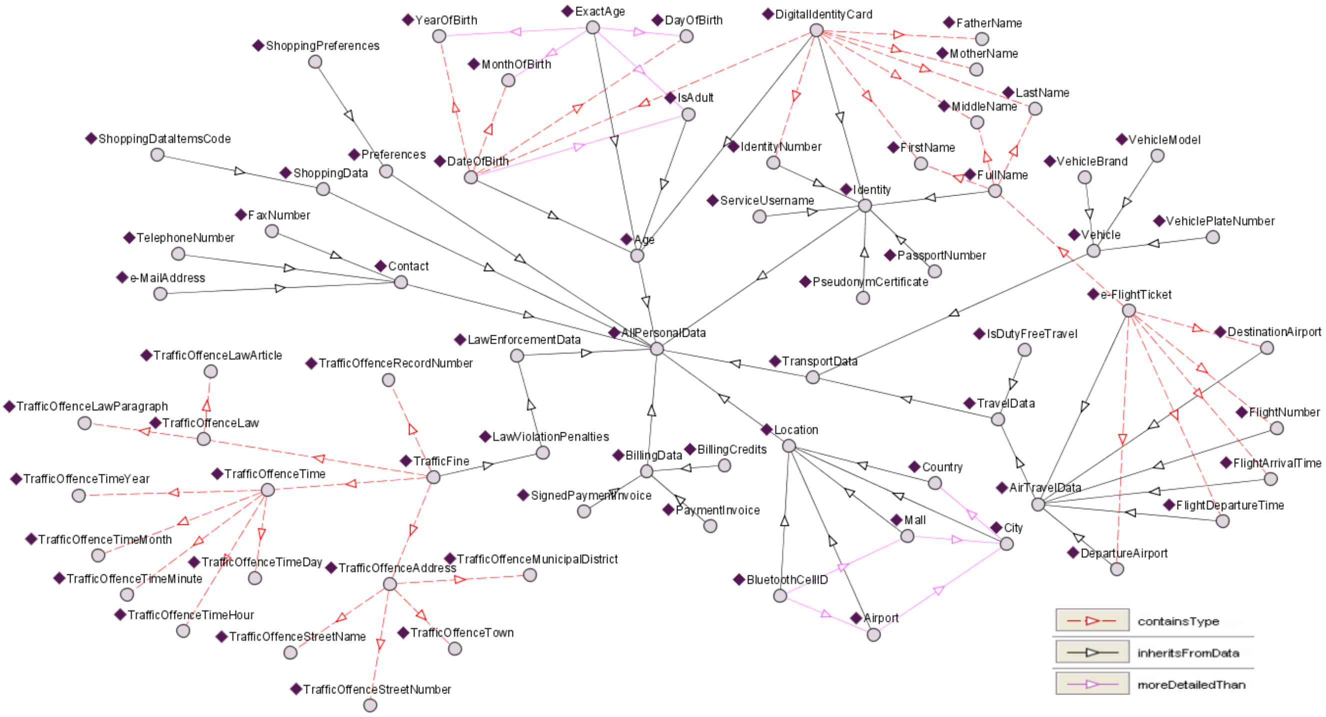

Fig 1: Ontology of Privacy – Personal Data Subgraph

type is the `AllPersonalData` type, from which all the other data types inherit, while "first level" children of `AllPersonalData` type instance include `Age`, `BillingData`, `Contact`, `Identity`, etc. These types constitute general data types, in essence categories of personal data types. This hierarchy is implemented by means of the `inheritsFromData` object OWL property.

The second hierarchy defined inside the `PersonalData` class deals with the detail level of personal data types. For this purpose, two properties are defined, `lessDetailedThan` and `moreDetailedThan`, being the one inverse to the other. In that respect, the `ExactAge` personal data types is `moreDetailedThan` the `YearOfBirth`, while the `Country` is `lessDetailedThan` the `BluetoothCellID`, with respect to the data subject's location.

The last relationship between the instances of the `PersonalData` class is the one that defines complex types resulting from simpler ones. In that respect, the data subject's

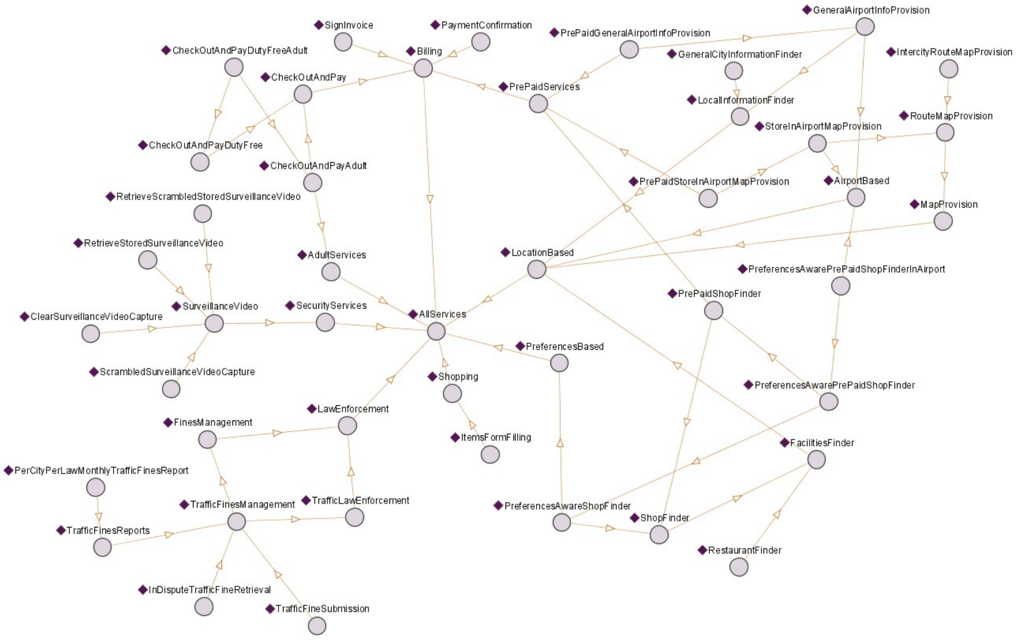
GeneralAirportInfoProvision
SignInvoice
PaymentConfirmation
PrePaidGeneralAirportInfoProvision
CheckOutAndPayDutyFreeAdult
Billing
GeneralCityInformationFinder
IntercityRouteMapProvision
CheckOutAndPay
PrePaidServices
LocalInformationFinder
RouteMapProvision
StoreInAirportMapProvision
CheckOutAndPayDutyFree
CheckOutAndPayAdult
PrePaidStoreInAirportMapProvision
AirportBased
RetrieveScrambledStoredSurveillanceVideo
MapProvision
RetrieveStoredSurveillanceVideo
AdultServices
LocationBased
PreferencesAwarePrePaidShopFinderInAirport
PrePaidShopFinder
SurveillanceVideo
SecurityServices
ClearSurveillanceVideoCapture
AllServices
PreferencesBased
ScrambledSurveillanceVideoCapture
Shopping
PreferencesAwarePrePaidShopFinder
LawEnforcement
FinesManagement
ItemsFormFilling
FacilitiesFinder
PerCityPerLawMonthlyTrafficFinesReport
TrafficFinesManagement
TrafficLawEnforcement
PreferencesAwareShopFinder
ShopFinder
TrafficFinesReports
RestaurantFinder
InDisputeTrafficFineRetrieval
TrafficFineSubmission

Fig 2: Ontology of Privacy – Subgraph of Services

FullName contains the FirstName, LastName and MiddleName data types. The containsType property and its inverse isContainedToType implement the corresponding relationships.

Fig. 1 illustrates part of the personal data subgraph, along with the OWL properties that implement the three types of relationships between the personal data instances, as described above.

The different services' types are organized as a hierarchy that defines inheritance of characteristics. All the defined types constitute instances of the Services OWL class. The "root" service type is the AllServices type, from which all the other services' types inherit, while "first level" children of AllServices type instance include AdultServices, Billing, LawEnforcement, LocationBased, etc. These types constitute general services' types. This hierarchy is implemented by means of the inheritsFromService object OWL property and its inverse one. It is noted that multiple inheritance is possible. As an example, a service can be location-based, while targeting adults; in this case, the service should inherit from both the LocationBased and the AdultServices types. Fig. 2 illustrates part of the services' subgraph; the arks represent inheritance associations, with the source node inheriting from the destination node.

As afore-mentioned, regarding the actors involved in the service provision chain, a very simple model has been considered. So far, the actors that have been defined are the DataSubject, the PrivacyAuthority and the ServiceProvider. However, this assumption can be easily removed with the extension of the Actors class to constitute a very detailed hierarchy of roles and –therefore– render the model fully role-based.

Access control rules are defined as instances of the Rules class of the ontology, in order to regulate the provision of services. Every rule is associated with a {personal data type, service type, actor} triad, using the corresponding refersToData, refersToService and refersToActor OWL object properties, and defines one or more properties that specify the permitted/forbidden actions of the *actor* over the *personal data type*, in the context of the provision of the *service type* under consideration, possibly along with certain complementary actions that must be additionally performed by the system.

With the use of OWL Annotation Properties, every rule contains the following information:

- DisclosureOfData: it defines whether the data of the specified type should be disclosed or not to the specified actor in the context of the provision of the specified service.
- RetentionPeriod: it specifies the period for which the data of the type under consideration should be retained.
- ModificationPermission: it defines if the specified actor should be granted with write/modify rights on the data of the specified type.

While the information above define the "core" of the rule, additional properties specify the complementary actions that should be potentially executed:

- DataSubjectInformation: it refers to the right of the user to be informed when the rule is applied (i.e., when in the context of the specified service, the personal data of the specified type are disclosed to the specified actor, or their modification takes place).
- DataSubjectConsent: it enables the user to be asked about explicit consent, prior enforce the body of the rule.
- AuthorityNotification: it forces the notification of the Authority when the rule is applied.

Finally, a rule may be characterized by certain meta-properties that serve for resolving conflicts between contradictory rules:

- appliesToPersonalDataDescendants: this binary property specifies whether the rule is inherited to the descendants of the specified data type, with respect to the corresponding subgraph of the ontology and the inheritance relationships.
- appliesToServiceDescendants: similarly to the case above, this binary property specifies the inheritance of the rule to the service type descendants.
- appliesToActorDescendants: although redundant since the corresponding actors' subgraph has not been defined yet, it refers to the inheritance of the rule to the descendants of the actor's type.
- OverrideDataSubjectPreferences: in certain cases, the user may have specified privacy preferences that contradict with the rules of the ontology; this property serves for defining which rule dominates over the other.

In Fig. 3, an example of an access control rule is illustrated. What this rule states is that "when the service under consideration is an adult service (AdultServices), and when the service provider (ServiceProvider) requests access to the personal data of IsAdult type (a binary data type, reflecting whether the data subject is an adult or not), the data should be given to the provider, while the data should not be further retained. The rule applies for the descendants of the AdultServices service type, while it does not apply for the descendants of the IsAdult personal data type and of the ServiceProvider actor type."

IV. Specification of Privacy Preferences

While regulations-originated policies as specified in the Ontology of Privacy may determine the access permission to data up to some extent, the users should be able to determine the fate of their personal data. In that respect, a technical problem to be approached is how to enable the user to that direction, i.e., to control the disclosure, storage and processing of personal information, when the information is traveling through the various system and service components. To face this issue, it is necessary to associate to the data additional information which is communicated and stored with the data and brings information aimed at enforcing the specific treatment desired for the considered data.

Therefore, prior to leaving the user's terminal, the user's personal data are encapsulated into a data structure with the descriptive name Privacy Lock. The purpose behind its use is twofold: to make certain metadata (i.e., the user's preferences) available along with the respective data and to ensure the safe transmission of the data.

In essence, the Privacy Lock constitutes a secure shell that encapsulates the personal data transmitted by the terminal to the D-Core Box and vice versa, along with their metadata into an encrypted and optionally digitally signed object that ensures the safe communication. Moreover, the Privacy Lock can be used for the transmission of metadata solely, that express user preferences as far as either already stored data or data that will be processed in the future are concerned.

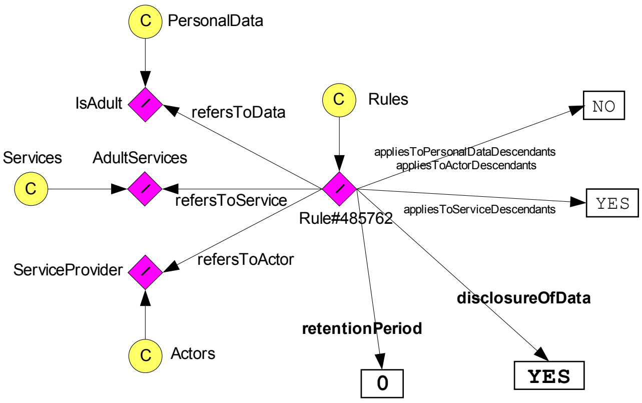

Fig 3: Ontology of Privacy – Example of Access Control Rule

An essential and mandatory attribute that is defined inside the Privacy Lock is the data type, which constitutes a critical parameter for their treatment in terms of disclosure, retention and processing. The type of the data is semantically specified with respect to the personal data subgraph of the Ontology of Privacy.

Apart from reflecting the data type, the metadata assigned to the data define certain properties for {personal data, services} [1] associations. These properties include:

- Whether the data should be disclosed for the provision of a certain service.
- The level of abstraction when the data are disclosed for the provision of a specified service.
- The expression of the user's preference to be informed or asked for consent whenever some processing or disclosure of personal data is about to take place.
- The determination of the desired retention period.
- Issues concerning data and metadata administration and management.

The metadata are formally expressed using a proprietary XML-based language, namely the Discreet Privacy Language (DPL). The DPL is used for the definition of all the necessary elements for the specification of all the afore-described types of metadata. Additionally, it is used for structuring the personal data when communicated from the user's terminal to a D-Core Box and vice-versa into a Privacy Lock, along with their privacy-related metadata.

The DPL's syntax is XML-based and contains the appropriate elements for the specification of the personal data attributes. In that respect, the DPL defines elements for the specification of:

- The personal data types, data items and service types that are regulated by the considered rule.
- Whether the considered rule is inherited by the personal data or service type descendants, with respect to the Ontology of Privacy.
- The rules themselves. The different rules' types defined include the disclosure level along with the corresponding level of precision, the access rights to the personal data, the demand for notifications and consents and the retention/validity period of the data.
- Meta-rules for the resolution of conflicts that naturally occur (e.g., contradictory user and regulations originated rules, rules overriding, etc.).

The detailed description of the DPL is beyond the scope of this paper; the normative definition of the DPL by means of Augmented Backus-Naur Form (ABNF) [16] specification is provided in [17].

V. Middleware Architecture

The privacy protecting system takes the form of a distributed middleware that regulates the diffusion of personal data from the user towards the service provider, using legislative input. This "privacy broker" is comprised of three high level entities, each assigned to one of the three actors, i.e., the users, the service providers and the Privacy Authority. These entities form a privacy domain, the D-Core, inside which personal data handling is subject to both legislative requirements regarding privacy and user privacy preferences. The high level entities and the privacy domain they define are illustrated in Fig. 4 .

The entity that delivers the core system functionality is the D-Core Box (Fig. 5). This is an intelligent privacy proxy, which, despite the fact that it is logically and physically deployed at the service provider's premises, is being managed by the Privacy Authority and not by the service provider. It constitutes the "edge" module of the D-Core and the border between the service provider's applications and the D-Core.

Any personal data provided by the user to the service provider are stored by the D-Core Box inside the Personal Data Repository. That is, the personal data are kept isolated from the service provider, which has no direct access to them. The storage may be short-time (e.g., immediate service provision) or long-time (e.g., services that require information archives). The data are stored together with the associated privacy preferences of the user, which are either transmitted with the data by means of a Privacy Lock, or defined/updated at any time through the corresponding interface that the D-Core Box offers. The same interface enables user to maintain complete control over the data, i.e., to update or delete them.

[1] It is noted that the metadata specification is actors-unaware; in fact, it is impossible for the user to refer to the internal structure of an organization.

When a service provider submits a request for users' personal data, the request is evaluated by the D-Core Box. All related decisions concerning personal data handling are taken by the Policy Engine which uses two sources of rules. The first source, the Ontology of Privacy is deployed to the D-Core Box by the Privacy Authority and is stored in the Regulations Repository. It provides the legislation-originated rules that are translated internally into a set of concrete DPL rules prior to be provided to the Policy Engine. The second source is the set of user defined privacy preferences which are provided by the Personal Data Repository, expressed as DPL rules. Through this prism, the Policy Engine examines the request and decides about the disclosure of the personal data and the potential execution of associated actions, such as the obfuscation of the data, the information of the user, the request for the user's consent, the notification of the Privacy Authority.

The key idea for the operation of the D-Core Box is to minimize the amount of personal data that are delivered to the application, without degrading the service. Moreover, the data should be disclosed pseudonymized or anonymized. In that respect, data that identify the user should not be disclosed, unless absolutely essential for the provision of the service. Therefore, in order to further minimize the amount of disclosed data, the D-Core Box incorporates modules that execute internally either simple data processing tasks or whole services' parts. These are, respectively, the Embedded Operators and Embedded Services components. Typical Embedded Operators functionalities are the filtering of the data precision prior to their disclosure (e.g., the translation of exact location to more abstract terms, or the transformation of the `ExactAge` data type to the `IsAdult` one.). Embedded Services undertake the execution of standard service components internally, mainly involving data that identify the user (e.g., e-mail sending or service charging mediation). This way, typical processing procedures that concern critical personal data, such as someone's identity or credit card number, are executed inside the D-Core Box and the need for the re spective personal data disclosure is eliminated. The Embedded Operators and Services are invoked by the Session Manager, a “stateful” component of the D-Core Box which orchestrates the functionality of the other components and manages every privacy session.

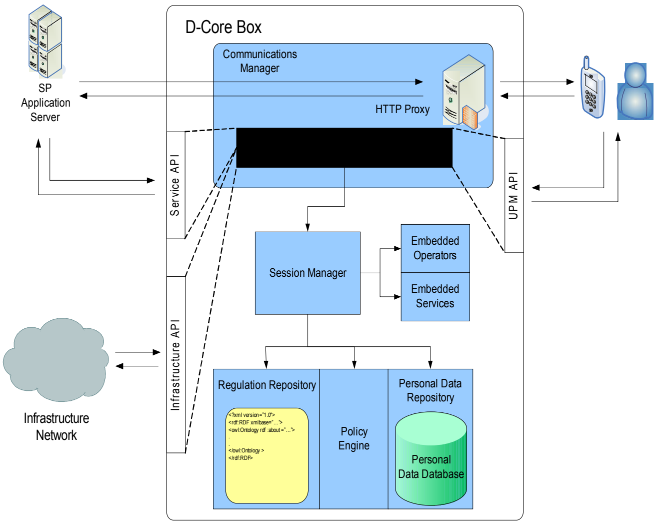

Fig 5: The D-Core Box

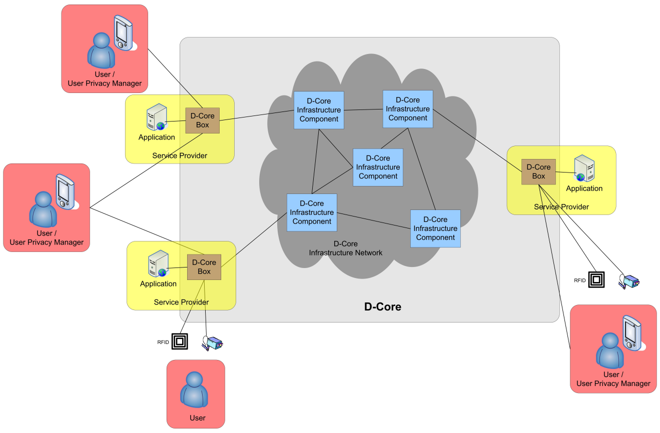

Fig 4: High level entities and the D-Core domain.

The D-Core extends at the user side with the User Privacy Manager (UPM). The UPM is a privacy agent for the user. It manages user identities for different services, it provides a console to edit privacy preferences and UIs to insert/edit personal data inside each identity and it creates Privacy Locks when personal data need to be delivered to the D-Core. The UPM is the peer entity of the D-Core Box for functions like informing the user when a service requests access to a specific personal data type, requesting user permission for this access and when creating and sending Privacy Locks containing personal data along with associated metadata.

The third entity in the D-Core domain is the Infrastructure Network. This is comprised of Infrastructure Components that constitute the Privacy Authority’s entry point to the domain. It provides to the Privacy Authority the means for the management of the Ontology of Privacy, the monitoring, and management of the system and the conduction of Lawful Interception. When the Ontology of Privacy is updated (e.g., due to legislation modification), then the updated version is transmitted through the Infrastructure Network to all the D-Core Boxes in order to consider the new requirements in the subsequent personal data requests. Regarding system management and monitoring, each Infrastructure Component undertakes the administrative responsibility regarding a number of D-Core Boxes and fulfills typical functions like collecting log data, checking status, generating error alarms, etc.

The communication of the D-Core Box with the other components of the D-Core domain as well as the applications is performed through a messaging framework, based on SOAP. The patterns defined for these interactions are presented in detail in [17], along with the detailed specification of the D-Core, its components and the respective interfaces between them and with the providers’ applications.

VI. Conclusion

In this paper, a framework defining a protection domain for personal data was presented. Conceived on the basis of the legislation provisions, it provides the means for their enforcement and, therefore, the applications' commitment to adhere to privacy requirements. It presents two innovative features.

The first is the formal modeling of the data protection legislation in terms of the Ontology of Privacy. Using the Ontology of Privacy as a powerful tool for expressing the respective notions, a mere dictionary of terms is defined and shared by all system components and actors, starting from a Privacy Authority that specifies and configures the Ontology of Privacy on a constant basis and ending in the data subject and the service providers that make use of it. In that respect, all personal data are semantically marked and their type affects their consequent treatment by all the involved entities. Similarly, the specification of each service type provides the means for disclosure and processing purposes' specification and binding. The policies inside the ontology not only determine the necessity of a personal data type for a service's provision, but also constitute a complete set of regulations that are translated to access rights, services' flows and other rules for the protection of the personal data.

The second contribution of the proposed framework is the explicit separation of the personal data from the service providers' applications. With the mediation of D-Core Box, a service provider cannot gain access to personal data other than the one specified by the legislation and the user's preferences. The incorporation of several privacy-critical processing functionalities in the D-Core Box eliminates further the danger of data misuse.

References

[1] S. D. Warren and L. D. Brandeis, "The Right to Privacy", *Harvard Law Review,* Vol. IV, No. 5, pp. 193–220, Dec. 1890.

[2] The European Opinion Research Group, "European Union citizens' views about privacy", *Special Eurobarometer 196,* Dec. 2003.

[3] The World Wide Web Consortium (W3C), The Platform for Privacy Preferences (P3P) Project, online: http://www.w3.org/P3P/.

[4] E. Bertino, J. Byun and N. Li, "Privacy-Preserving Database Systems", *Foundations of Security Analysis and Design III,* Lecture Notes in Computer Science 3655, Springer-Verlag, 2005.

[5] U.S.A. Federal Trade Commission, "Eli Lilly Settles FTC Charges Concerning Security Breach", FTC File No. 012 3214, Jan. 2002.

[6] U.S.A. Federal Trade Commission, "FTC Sues Failed Website, Toysmart.com, for Deceptively Offering for Sale Personal Information of Website Visitors", FTC File No. 002 3274, Jul. 2000.

[7] P. Ashley, S. Hada, G. Karjoth, C. Powers, M. Schunter, "The Enterprise Privacy Authorization Language (EPAL), *EPAL* 1.2 Specification", IBM Research Report, 2003.

[8] Organization for the Advancement of Structured Information Standards, "eXtensible Access Control Markup Language TC", 2004.

[9] M, Casassa Mont and R. Thyne, "A Systemic Approach to Automate Privacy Policy Enforcement in Enterprises, in *Proc. of 6th Workshop on Privacy Enhancing Technologies,* Lecture Notes in Computer Science, Vol. 4258, Springer-Verlag, 2006.

[10] Organization for Economic Co-operation and Development (OECD), "Guidelines on the Protection of Privacy and Transborder Flows of Personal Data", Sep. 1980.

[11] European Parliament and Council, "Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data", OJEC, No. L 281, pp. 31-50, Nov. 1995.

[12] European Parliament and Council, "Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector", OJEC, No. L 201, pp. 37-47, Jul. 2002.

[13] European Parliament and Council, "Directive 2006/24/EC on the retention of data generated or processed in connection with the provision of publicly available electronic communications services or of public communications networks and amending Directive 2002/58/EC", OJEC, No. L 105, pp. 54-63, Apr. 2006.

[14] The World Wide Web Consortium (W3C), "Web Ontology Language (OWL)", online: http://www.w3.org/2004/OWL/.

[15] European Parliament and Council, "Regulation 2195/2002/EC on the Common Procurement Vocabulary (CPV)", Official Journal of the European Communities, No. L 340, pp. 1–562, December 2002.

[16] D. Crocker, P. Overel. "Augmented BNF for Syntax Specifications: ABNF," RFC2234, IETF, Nov. 1997.

[17] Georgios V. Lioudakis et. al., "Implementation Report on the Core System", IST DISCREET Deliverable D3102, January 2008.