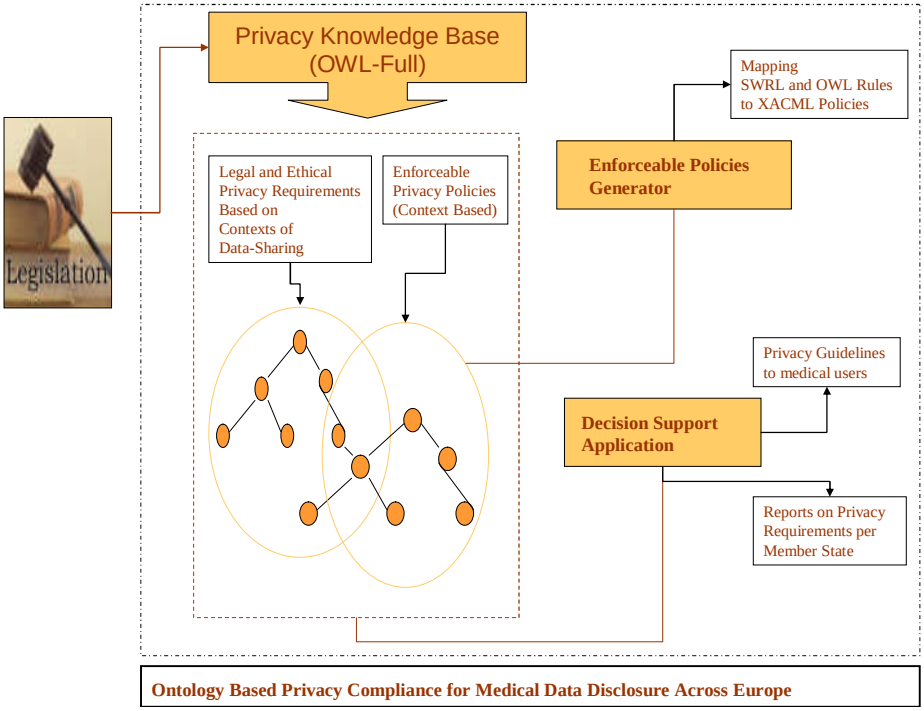
Ontology-Based Privacy Compliance on European Healthgrid Domains

Hanene Boussi Rahmouni[1a], Tony Solomonides [a],
Marco Casassa Mont [b], Simon Shiu [b]
[a] *Bristol Institute of Technology, UWE, Bristol/BS16 1QY*
[b] *HP Labs, Stoke Gifford, Bristol/BS34 8QZ*

1. Introduction

The diversity and complexity of legislations on data protection across Europe might be an obstacle towards the deployment of integrated European medical research and healthcare systems. Our research is aiming to help enhancing privacy compliance when sharing medical data across Europe. Our approach is based first on bridging the gap between high level legislation on data protection and operational level controls by means of semantic modeling and second on the use of semantic applications to provide decision support for medical users on how to handle sensitive patient data and also for generating enforceable privacy policies.



2. Modeling Privacy requirements: RDF/OWL plus Rules

The diversity, complexity and dynamicity of the rules governing privacy protection in Europe explains the need for a modelling approach that is able to abstract this complexity and facilitate its automation and enforcement at the process level. We mean by privacy requirements all the obligations that must be fulfilled by all parties involved in the process of sharing and processing sensitive patient data for medical purposes including healthcare and medical research to preserve the patient privacy. This includes patient consent, data anonymity, and the rights of the data subject including their right to de-sent and to be notified [1]. Our model should reflect any conflicts between the EU member states in the specification and the provision of these requirements. In the following paragraphs we are presenting our attempt to model and automate privacy requirements in the context of medical data disclosure in Europe.

[1] Hanene2.Rahmouni@uwe.ac.uk

Our approach uses the Protégé 3.4 OWL-Full [2] features to represent privacy obligations in the context of medical data disclosure. OWL allowed us to model the conceptual domain of "data sharing" or "data disclosure" and its components as a hierarchy of classes, subclasses and a hierarchy of properties to represent the relationships between them. As shown in figure2: privacy requirements such as patient's consent requirements could be modeled as OWL classes and assigned to the "dataSaharing" class as object properties. Additional class expressions including restrictions, Boolean expressions, enumerated classes and value partitions were also useful. For example, consider a cardinality constraint such as a data item must have one and only one place of origin.

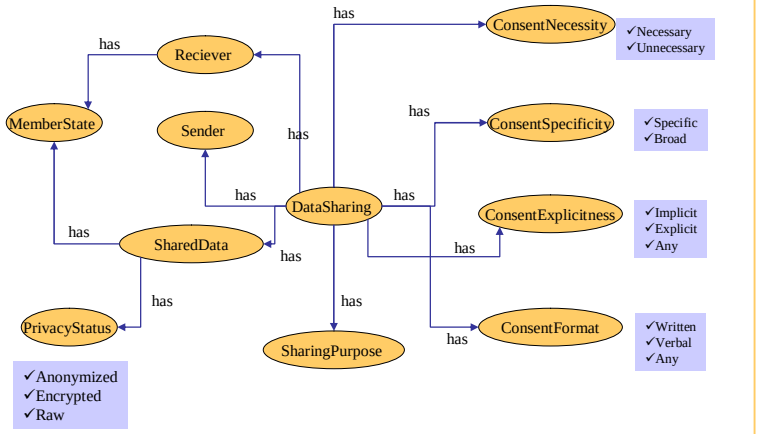


Figure2: Ontology Model of Data Sharing Requirements

Moreover, through the use of OWL properties that have additional capabilities (transitive, symmetric or functional), and special types of OWL classes i.e. OWL Equivalent Class we allowed overlapping models of a concept to be merged, even when different naming have been used for the same resource; for example, Explicit Consent might be named Express Consent in another model but both concepts have the same meaning.

With complex legal domains, we need to model relationships that cannot be expressed in OWL because the logic for describing porperties is not rich enough. Legal rules usually expressed in the form of *if –then*-like rules. For example, we want to model a rule stating that if the data belongs to the UK then patient concent is necessary for the processing. Expressing this kind of rule requires the use of a semantic web rule language to allow building sets of rules in terms of the different concepts of the sharing process already described in the ontology and their properties. This will allow us to run reasoning operations on the relevant set of rules and the ontology classes to infer privacy requirements for different possible instances of sharing in the real world.

As a rule language we have relied on a promising approach based on the Rule Markup Language (RuleML) that is the Semantic Web Rule Language (SWRL) [3]. The protégé SWRL tab/editor allowed as to directly incorporating OWL entities from the ontology into the rules we are building. Switching between SWRL rule editing and the editing of OWL entities was effectuated in a seamless way. The following example is a SWRL representation to the rule stating that "the patient consent is necessary for the sharing of a UK medical data item that is anonymized".

DataSharing(?x) ^ hasSender(?x, ?s) ^hasReceiver(?x, any) ^
locatedIn(?s, UK) ^ hasSatus(?d, Anonymised)
→ hasConsentNecessity(?x, Necessary)

Our rules were tested and executed by invoking a JESS [4] engine through the use of the provided Protégé SWRL Bridge. Once executed, JESS finds the relevant OWL knowledge from the ontology model and asserts some inferred knowledge to them. If the rule stated above is executed JESS will assert the value Necessary to the *hasConsentNecessity* property of the matching individuals of the *DataSharing* class.

3. Decision support for clinicians and medical technicians to enhance compliance with privacy regulations

Our system should reason on the model described in the previous section to generate guidelines or protocols for medical users, to guide them through the different processing tasks required. For this purpose we developed a semantic web application that allows users to specify details of the different entities that constitute a sharing process. The data entered by the user are uploaded to the OWL model of "data sharing contexts" through Protégé OWL API calls. Methods from The SWRL-Bridge API were called in order to invoke a JESS rule engine which would fire up the relevant SWRL rules from our model. The result is a set of new inferred axioms that are added to the model as attributes of the instance of the "Data Sharing" class in question. These attributes will be returned to the user as the set of privacy requirements necessary to allow the sharing of the data. For example if the rule engine has decided to fire up the following rule:

DataSharing(?x) ∧ hasSender(?x, ?s) ∧ hasReceiver(?x, any) ∧
locatedIn(?s, UK) ∧ hasSatus(?d, Anonymised) → hasConsentNecessity(?x, Necessary) ∧ hasConsentSpecificity(?x, Specific) ∧ hasConsentExplicitness(?x, Any)

The system then indicates to the requestor that Specific Consent is required for the sharing of the data of interest and the consent could be either Explicit or Implicit. Our system also allows users to generate reports on privacy safeguards for each member state. These reports help by informing the users of possible conflicts that might exist between the regulatory framework of the member state owning the data and other frameworks across Europe.

4. Extending the ontology to enable the specification of enforceable privacy policies to insure compliance

For better governance of European integrated health systems, legal and ethical requirements for privacy must be enforced at operational level as formal privacy policies. Through the use of OWL- Full on Protégé 3.4 we were able to represent the concept of a "Policy" and edit policy instances in a seamless way.

Privacy policies are divided into two main categories according to their order of enforcement compared to the access control policy associated with them. The first category of policies is the ones that might affect a system access control decision (could cause access permissions to rollback). For example if patient consent is required for a specific context of sharing, the system will check for availability of the patient consent before allowing the user to access their data. The second category are Privacy Obligations [5]that are rather post conditions of an access control decision that deals with issues such as: disclosure of data to third parties, data deletion and retention.

In order to be easily enforced at the system level we suggest that our policies should be specified in a way that conforms to a widely adopted policy language standard that has proven efficiency in the enforcement of privacy policies. Our choice was the extendable access control markup language (XACML). Privacy policies in XACML are specified using some standard XML elements including (Policy, Target, and List of rules) where the Target refers to the resource we are controlling access to, and the rules attached to the policy are described in terms of other standard elements of XACML including: Rule Effect (permit, deny…), Rule Target (Subject or Requester, Resource, Requested Action) and Rule Conditions [5]. The conditions attached to each rule are specific constraints on (the subject or requestor, resource, and others (depends on the context). XACML also allows users to add more user defined components or elements to the traditional vocabulary [6].

On our attempt to semantically model privacy policies while conforming to the XACML specification, we have chosen to model the different components of the policy vocabulary as classes in our ontology. OWL does not allow privacy control conditions to be directly mapped to classes, but the use of DL restrictions on the different entities that constitute a rule target were good enough to solve this problem. Instances of the Rule class will be mapped to some SWRL access control rules defined on our knowledge base. The following example shows an attempt to formalize the rule stating that: "A user may access a patient mammogram if patient has provided informed consent for a specific purpose of processing and the processing purpose is compatible with the purpose consented for":
Obtained (?Subject, InformedConsent) Λ hasCollectionPurpose(?Object, CollectionPurpose) Λ CompatibleWith(?Purpose, ConsentPurpose)
–> Allow(?Subject,Access)

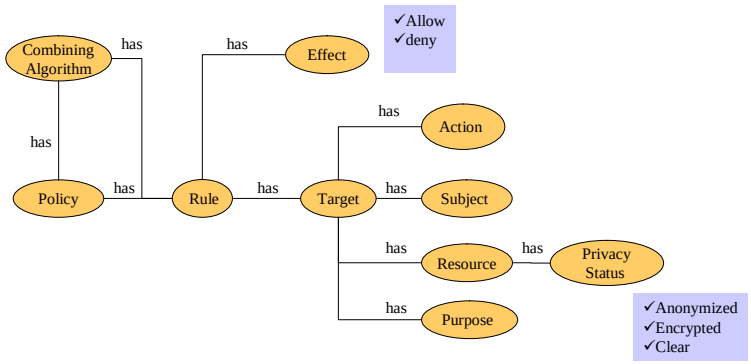


Figure2: Ontology Model of Privacy Policies

The case of modeling "Obligations" is even more challenging. As mentioned above obligations are to be enforced once a policy decision was made. Therefore these obligations are assigned to a specific policy with a specific effect (allow, deny…) this component will be difficult to model using OWL description logic restrictions as they are only consequences to a specific condition. We believe SWRL rules are a more suitable approach to model Privacy Obligations.

Our model captured all the components that constitute the XACML privacy policy specification and extends the Rule target component to allow setting constraints according to the purpose of data processing, the member state to which the data belongs.

5. Conclusion and Future Work

Throughout our research we have managed to model high level policies interpreted from European and national data protection law as privacy requirements for data disclosure. We have been able to capture similarity and possible conflict between the different frameworks across Europe through the use of SWRL rules, JESS and the protégé' API. We have also specified by ontology means the concept of "Enforceable Privacy Policy" (Conforming to the XACML Standard). Privacy policies could therefore be created as instances of the Policy class and could be assigned to an equivalent privacy requirement. Linking between privacy requirements that is generated by the data disclosure decision support application and the policies that are enforced at the operation level is the basis for privacy compliance assurance on integrated medical systems. For future work we are looking at developing a semantic application to construct XACML privacy policies and obligations from the policies specified in our OWL Full ontology model of enforceable privacy policy.

References

[1] D. Beyleveld, D. Townend, S. Rouillé-Mirza and J. Wright. "Implementation of the Data Protection Directive in Relation to Medical Research in Europe".ISBN: 0754623696

[2] OWL Web Ontology Language Reference. www.w3.org/TR/owl-ref, 2004.

[3] SWRL: A Semantic Web Rule Language Combining OWL and RuleML. Joint US/EU adhoc Agent Markup Language Committee, November 2003. http://www.daml.org/ 2003/11/swrl/

[4] E. Friedman-Hill. JESS in Action: Java Rule-based Systems, Manning Publications Company, June 2003, ISBN 1930110898, http://herzberg.ca.sandia.gov/JESS/

[5] OASIS eXtensible Access Control Markup Language Technical Committee: eXtensible Access Control Markup Language (XACML). http://www.oasisopen.org/committees/tc_home.php?wg_abbrev=xacml

[6] Yuri Demchenko, Cees de Laat, Oscar Koeroo, Hakon Sagehaug. "Extending XACML Authorisation Model to Support Policy Obligations Handling in Distributed Applications ". *ACM Conference Paper*.