Medical Informatics in a United and Healthy Europe
K.-P. Adlassnig et al. (Eds.)
IOS Press, 2009

doi:10.3233/978-1-60750-044-5-695

Modelling and Enforcing Privacy for Medical Data Disclosure across Europe

Hanene BOUSSI RAHMOUNI [a,1], Tony SOLOMONIDES [a],
Marco CASASSA MONT [b], Simon SHIU [b]
[a] *Bristol Institute of Technology, University of the West of England, Bristol, UK*
[b] *Trusted Security Lab, HP Labs, Bristol, UK*

Abstract. The harmonization of data protection legislation in Europe has been theoretically achieved by means of the EU directive on data protection. In practice the harmonization is not absolute and conflicts and inconsistencies continue to exist in the way Member States are implementing the directive. The integration of different European medical systems by means of grid technologies will continue to be challenging if technology does not intervene to enhance interoperability between national regulatory frameworks on data protection. In this paper we present an approach to automate privacy requirements for the sharing of patient data across Europe on a healthgrid domain and ensure its enforcement internally and within external domains where the data might travel. This approach is based on the semantic modelling of privacy obligations that are of legal, ethical or cultural nature. These requirements are for the sharing of personal data between different European Member States. Our model reflects both similarities and conflicts, if any, between the different Member States. This allows us to reason on the safeguards a data controller should ask from an organization belonging to another Member State before disclosing medical data to them. The system will also generate the relevant set of policies to be enforced at the process level of the grid to ensure privacy compliance before allowing access to the data.

Keywords. privacy, EU data protection directive, health-grid, semantic web technologies

1. Introduction

When sharing medical data between different health organizations in Europe, it is important that the different parties involved in the sharing handle the data in the way indicated by the legislation of the Member State where the data was originally collected since the requirements may differ from one state to another. Privacy requirements, such as patient consent, may be subject to conflicting conditions between different national frameworks as well as between different legal and ethical frameworks of the single Member State. Whilst most EU Member States are now governed by similar personal data protection rules, harmonization remains more apparent than real. This is due first to the fact that subject to the provision of suitable safeguards the European data protection directive [1] leaves some space for Member States to lay down simplifications and exemptions to some of the obligations that are dictated [1] i.e., the

[1] Corresponding Author: Bristol Institute of Technology, University of the West of England, Bristol, UK, BS16 1QY; E-mail: Hanene2.Rahmouni@uwe.ac.uk.

obligation to notify the data subject of the processing of their data. Also for reasons of substantial public interest, Member States may lay down exemptions to the ban of the processing of sensitive personal data in addition to those laid down in the directive, either by national law or by decision of the supervisory authority [1]. Second, as specified by some studies [2], the definitions used do not lead to a uniform understanding of the key concepts underpinning the directive. Focusing on the concept of "Personal Data", many Member States find it difficult to interpret. The UK found that in some cases data is not easily classified as personal or non personal. And this classification could be relative according to the circumstances. Overlaps in the interpretation of "Personal Data" have also resulted in different ways of governing anonymized and pseudonymized data [2]. Consequently, the frameworks in some Member States such as the UK [3] tend to be less favourable to the processing of personal data for medical research compared to other frameworks including the Italian data protection framework. The latter seems to grant more privileges to medical researchers in allowing consent for the processing of medical data across different healthcare organizations to be given in a single, one-off statement [4]. This raises ethical concerns on handling secondary usage of the data [5].

These issues explain the diversity, complexity and dynamicity of the rules governing privacy protection. We believe modelling could simplify and abstract the complexity of rules from the real world to allow their automation and enforcement at the organizations' process level. For this paper our ideas will be structured as follows: in section two, we present our technical solution to the modelling and automation of privacy requirements. Section three presents a proof of usability of the model for building decision support applications to help the healthgrid's [6] medical users to share medical data while complying with privacy obligations. Finally we conclude and hint to future tasks that look at enforcing privacy obligations on the grid system.

2. Modelling Privacy Requirements: 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, anonymization or pseudonymization, the rights of the data subject including their right to dissent and to be notified. Our approach deals only the requirements that could be enforced using a policy-based approach and does not include the cases where the intervention of ethical committees is essential. Our model should rather reflect similarity and possible conflicts between the EU Member States in the specification and the provision of these requirements. In the following paragraphs we present our attempt to model and to automate privacy requirements in the context of medical data disclosure in Europe.

Our approach uses the Web Ontology language (OWL) [7] to represent privacy obligations in the context of medical data disclosure. OWL allows us to model the conceptual domain of "data sharing" or "data disclosure" and its components as hierarchies of classes/subclasses and of properties to represent the relationships

between them. Privacy requirements such as consent requirements may be modelled as OWL classes and assigned to the "dataSharing" resource as object properties.

Moreover, OWL provides additional abilities to allow overlapping models of a concept to be merged, even when different naming conventions 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.

In complex legal domains we need to model relationships that cannot be expressed in OWL because the logic for describing properties is not rich enough. Legal rules are usually expressed as *if–then*-like rules. For example, we want to model a rule stating that if the data belongs to the UK then patient consent is necessary for any 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 allows us to reason on the relevant set of rules and the ontology classes in order to infer privacy requirements for different possible instances of sharing from the real world. The Semantic Web Rule Language (SWRL) [8] satisfies our requirements for this task. The following example is a SWRL representation of the rule stating that *patient consent is necessary for the sharing of a UK medical data item that is anonymized*. Thus,

$$\text{dataSharing}(?x) \wedge \text{hasSender}(?x, ?s) \wedge \text{hasReceiver}(?x, \textit{any}) \wedge \text{locatedIn}(?s, \text{UK}) \wedge \text{hasSatus}(?d, \text{Anonymized}) \rightarrow \text{hasConsentNecessity}(?x, \text{Necessary})$$

In the next section we describe how OWL ontology and the semantic rules we have created could be used to provide decision support for medical users to help them to behave in a privacy-aware manner when sharing patient data on the grid.

3. Decision Support for Clinicians 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 both uploading data to a European grid and for accessing and downloading data from the grid to be processed within an external IT environment. For this purpose we developed a semantic web application that allows users to specify details of the different entities that constitute a sharing process and invoke a Jess rule engine [9] to fire up the relevant SWRL rules from our model. The result will be 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:

$$\text{dataSharing}(?x) \wedge \text{hasSender}(?x, ?s) \wedge \text{hasReceiver}(?x, \textit{any}) \wedge \text{locatedIn}(?s, \text{UK}) \wedge \text{hasSatus}(?d, \text{Anonymized}) \rightarrow \text{hasConsentNecessity}(?x, \text{Necessary}) \wedge \text{hasConsentSpecificity}(?x, \text{Specific}) \wedge \text{hasConsentExplicitness}(?x, \textit{any})$$

(in practice this is divided into a set of rules), then the system indicates to the requestor that Specific Consent is required for the sharing of the data in question 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. Figure 1 shows how the decision support application fits on the architecture of the privacy compliance framework we are working on.

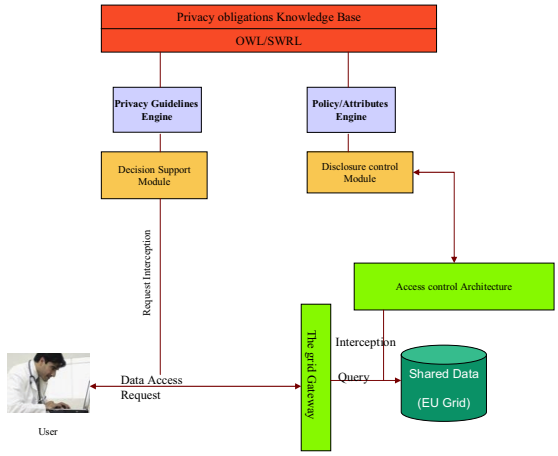

Figure 1. Architecture for enhancing privacy compliance on the grid domain

3.1. Uploading Data on the Grid

When a user requests to upload data from the hospital database to the federated grid database, the system must first generate the set of privacy obligations that the user needs to comply with before the data is uploaded to the grid. These requirements are generic and do not depend on the geographic location of the entities that would have access to it or share it in future. In other terms the national legal and ethical framework would be the primary reference for identifying privacy requirements for this task. Requirements could include anonymization, pseudonymization, data de-identification including image scrambling, consent for storing the data in the grid and obligations related to the quality of the data including data provenance, accuracy and relevance. To achieve this goal, a local version of the framework must be deployed as part of the local resources at each hospital or medical research centre participating in the grid.

3.2. Downloading Data from the Grid

In our application, the grid system is not fully open and data may be shared only on request. When a user within Member State A requests to access data belonging to another Member State B, the system should generate the relevant set of requirements which are just the additional safeguards that Member State B would usually ask users at Member State A to guarantee before sharing medical data with them. Allowing

access to the data would be subject to some security policies that are not part of our focus and also to the privacy assurance the user provides when requesting the access.

In order to control data disclosure when downloading data from the grid, a distributed version of the framework is required. This allows the management of sharing requests coming from all nodes participating on the grid in an appropriate manner. The following example represents the rule indicating that in order to process some UK data for secondary purposes and when consent is necessary, a researcher who is a member of the grid must contact the general practitioner (GP) of the data subject and wait for them to collect *consent for consent* from the patient (the patient's permission to collect their consent).

$$\begin{aligned} &\text{dataSharing}(?x) \wedge \text{concerning}(?x, ?d) \wedge \text{belongsto}(?d, \text{UK}) \wedge \text{hasPurpose}(?x, ?p) \\ &\quad \wedge \text{isa}(?p, \text{SecondaryPurpose}) \wedge \text{hasRequestor}(?x,?r) \wedge \text{generalPractitioner}(?g) \\ &\qquad \rightarrow \text{mustContact}(?r, ?g) \wedge \text{mustObtain}(?g, \text{ConsentforConsent}) \end{aligned}$$

4. Conclusion and Future Work

Privacy requirements for the sharing of medical data between European Member States can be described within a semantic model. Once it is rich enough, the model could form a knowledge base for inference engines to reason about the duties of medical users as imposed by different European and national legislation in order to preserve patient privacy. The new inferred knowledge generated by the inference engine can provide guidelines and protocols to help clinicians and other medical users across Europe to share medical data while complying with regulations. Our work has mainly focused on the requirement of patient consent but we believe other requirements could be modelled in the same way, including anonymization, role-roaming, etc. In future work we will extend our semantic model of privacy requirements to allow the specification of privacy obligations as enforceable policies conforming to a standard access control policy language.

References

[1] EU Directive 95/46/EC The Data Protection Directive, http://ec.europa.eu/justice_home/fsj/privacy/docs/95-46-ce/dir1995-46_part1_en.pdf.

[2] McCullagh, K. (2006) A study of data protection: Harmonization or confusion? In *Proceedings of the 21st BILETA Conference on Globalisation and Harmonisation in Technology Law*, http://www.bileta.ac.uk/Document%20Library/1/Data%20protection%20-%20harmonisation%20or%20confusion.pdf.

[3] Iversen, A., Liddell, K., Fear, N., Hotopf, M., Wessely, S. (2006) Consent, confidentiality, and the Data Protection Act. *British Medical Journal* 332(7534):165–169.

[4] Italian Personal Data Protection Code, Legislative Decree No. 196, 30 June 2003.

[5] Beyleveld, D., Townend, D., Rouillé-Mirza, S., Wright, J. (2004) *Implementation of the Data Protection Directive in Relation to Medical Research in Europe*. Ashgate, Aldershot.

[6] Breton, V. et al. (2005) The HealthGrid White Paper. In *"From Grid to Healthgrid", Proceedings of the Third HealthGrid Conference 2005*, IOS Press, *Studies in Health Technology and Informatics* 112:249–318. (Published online: http://initiative.healthgrid.org/the-initiative/healthgrids-concept/white-paper.html in 2004).

[7] McGuinness, D.L., van Harmelen, F. (2004) OWL Web Ontology Language Overview, W3C Recommendation, 10 February 2004, http://www.w3.org/TR/owl-features/.

[8] Joint US/EU ad hoc Agent Markup Language Committee (2004) SWRL: A Semantic Web rule language combining OWL and RuleML, http://www.w3.org/Submission/2004/SUBM-SWRL-20040521/.

[9] Friedman-Hill, E. (2003) *Jess in Action: Java Rule-Based Systems*. Manning Publications Company, Greenwich.