ENDORSE

Malone, P.; McLaughlin, M.; Leenes, R.E.; Ferronato, P.; Lockett, N.; Guillen, P.B.; Heistracher, T.; Russello, G.

Published in:
Governance of technology, information and policies

Document version:
Publisher's PDF, also known as Version of record

Publication date:
2011

[Link to publication](#)

ENDORSE: A Legal Technical Framework for Privacy Preserving Data Management

Paul Malone
Waterford Institute of
Technology
Waterford
Ireland
pmalone@tssg.org

Mark McLaughlin
Waterford Institute of
Technology
Waterford
Ireland
mmclaughlin@tssg.org

Ronald Leenes
University of Tilburg
Tilburg
The Netherlands
r.e.leenes@uvt.nl

Pierfranco Ferronato
Soluta.Net
via Edificio 2
31030, Caselle D'Altivole (TV).
Italy
pferronato@soluta.net

Nick Lockett
DL-Legal LLP
Rahere House
59 Broomfield Avenue
London, UK
nicklockett@dllegal.com

Pedro Bueso Guillen
University of Zaragoza
Zaragoza
Spain
pbueso@unizar.es

Thomas Heistracher
Salzburg University of Applied
Sciences
Urstein Süd 1
A-5414 Puch/Salzburg, Austria
thomas.heistracher@fh-salzburg.ac.at

Giovanni Russello
CREATE-NET
Via alla Cascata 56/D Povo
38123 Trento
Italy
giovanni.russello@create-net.org

ABSTRACT

The ENDORSE project is concerned with providing assurances for data protection for both data controllers and data subjects. The project will define a rules based language called PRDL (Privacy Rules Definition Language) which can be used to express legislative requirements, organizational privacy policy as well as user consent. ENDORSE will provide a rules engine to ensure that privacy policies expressed in this language are compliant with legislative requirements for the applicable jurisdictions. In addition a set of technology adapters will be developed which will provide transformations from PRDL to target access control and policy configuration instances, which in turn can be used by organizations to ensure that internal data handling practices are in turn compliant. In parallel to this effort a certification methodology will be developed to provide a means of generating a privacy seals. This paper describes an overview of the project, the motivation behind the initiative, its aims and objectives as well as an introduction to the approach taken and technologies foreseen to achieves these aims. The paper also provides a discussion of how the results of the project can be applied in different scenarios.

Categories and Subject Descriptors

K.6.5 [**Management of Computing and Information Systems**]: Security and Protection; K.4.1 [**Computers and Society**]: Public Policy Issues—*privacy*

General Terms

Legal Aspects, Management, Reliability, Verification

Keywords

Data Protection, Privacy, Data Management, Compliance

1. INTRODUCTION

Privacy and data protection are of concern to many stakeholders, including the data subjects (end-users), the data controllers (organizations) as well as legislative bodies, data protection agencies, consumer rights organizations and human rights advocates. End-users require assurances that their personal data are being fairly and correctly collected and managed for purposes for which they, ideally, have given explicit consent and it is done so in a transparent manner. Organizations collecting personal data need to ensure the data management practices employed are in compliance with legal requirements and not subject to misuse by its employees. These data protection requirements introduce an overhead (both financial and operational). For European SMEs(Small and Medium Enterprises)[1], the need to ensure

---

[1] A legal concept in an EU context in: Commission Recommendation 96/280/EC of 3 April 1996 concerning the definition of small and medium-sized enterprises (Official Journal L107, 30/04/1996, pp. 4-9)

compliance can lead to a disproportionate cost when compared to core activities, inhibiting growth and opportunities in competitive global markets.

What can help is an open source toolset which allows organizations to ensure that their personal data management policies are compliant with the appropriate legislation.

The contribution of this paper is to highlight the goals of ENDORSE, a European Union funded project. The ENDORSE project brings together a consortium of data protection legal experts, academic computer science partners, software implementors and interested industry players to deliver an open source toolset to create legally compliant privacy policies which can be deployed in organizational infrastructure.

This paper is organized as follows. In Section 2, we provide the motivations driving our research. Section 3 presents an overview of the ENDORSE project, providing the project aims and objectives. Section 4 is dedicated to the ENDORSE architecture. The application areas where the ENDORSE outcomes will be applied are described in Section 5. We conclude our discussion with a summery presented in Section 6

2. MOTIVATION

Privacy and data protection concepts are of considerable concern to many stakeholders, including the data subjects, the data controllers as well as legislative bodies, data protection agencies, consumer rights organizations and human rights advocates. On the one hand, end-users require assurances that their personal data is being fairly and lawfully collected and managed for the purposes specified in accordance with art. 6 of the Data Protection Directive 95/46/EC [1] and is done so in a transparent manner. Ideally, end-users also provide informed consent regarding the processing (cf. art. 7 para 1 95/46EC[1]). On the other hand, organizations collecting personal data are keen to ensure that the data management practices employed are in compliance with legal requirements and not subject to misuse by its employees. A slack attitude to data protection and security can be costly for organizations, not just in terms of conceding ground in the market to competitors that better cater to customer concerns, but also in terms of the high cost to organizations of a data breach.[2] From an organizational point of view, these data protection requirements introduce an overhead (both financial and operational), which needs to be managed and minimized. The ultimate goal of ENDORSE is to address these issues as follows:

1. Provide the end-user (data subject) with the assurance that the data management policies of the data controller are in compliance with the appropriate legislation and that the control of data access is in line with those policies. These data management policies are expressed in the company's privacy operating policies and the ENDORSE toolset ensures that infrastructural data management procedures are in line with that policy. A certification methodology will enable the organization to produce a certificate of compliance, pro-

viding a means of verifying this adherence to the data subject.

2. Provide a means for organizations to ensure compliance by providing a privacy rules definition language to define data collection and management policies. ENDORSE will also provide a representation of data protection legislation expressed in that language ensuring compliance of the organization's policies with the appropriate regulation. This language, together with the legal representation of the legislation in the language and the toolset to create policies and technical data access control specifications and privacy policies will be made available as open source components.

An open source approach lowers the costs of compliance for SMEs and provides an open and transparent framework for data protection and privacy compliant practices.

3. ENDORSE OVERVIEW: AIMS AND OBJECTIVES

The primary aim of ENDORSE is to create an open and freely available legal technical toolset for privacy preserving data management that can be adopted by public bodies and enterprises to offer solid guarantees to service subscribers regarding the range of use of personal information on their systems. This toolset will prevent the accidental or unauthorized manipulation of sensitive personal information. The framework will describe how personal information can be stored and accessed in a compliant and secure manner on public and private data stores, and how it is exposed to services and authorized personnel using a privacy preserving rule based modeling approach.

This framework will consist of a legal and a technical component:

- The *legal component*, informed by social science, the principles of human rights, data protection law and the limitations of technology, will create a specification for data access and manipulation within digital systems that can be adhered to by data controllers. This component will also provide a roadmap for how this specification could be adopted as a standard for privacy preserving personal information storage in law and/or by voluntarily compliant parties.

- The *technical component* will provide an architecture, a privacy rule definition language and a toolset for management of data access and manipulation that complies with the specification produced by the legal component, which provides a definition of a filtered scheme of access to data according to role-based policies, respecting data collection rationale, and utilizing the state of the art in secure communication and encryption technologies and methods.

A major outcome of this project will be the enforcement of data protection compliant data access logic, clear definitions for responsibilities of compliant data controllers and processors, with additional specification for web applications, such as definition and generation of comprehensible privacy policies and consistent interface for data subjects. This effort to standardize and harmonize data management practices and ensure legal compliance will be facilitated and enforced by the technological component of this project.

---

[2]According to the recent Ponemon report[7], the average cost of a data breach was $3.4m, or $142 per record per data breach, in surveyed countries (US, UK, Germany, France and Australia). 44% of the cost was accrued due to lost business.

3.1 Requirements

ENDORSE is concerned with addressing the following requirements for organizations seeking compliance and user acceptance:

1. Data should, with as few exceptions as possible, only be gathered for a particular purpose and the framework will ensure that this purpose is explicit and that data is not accessed or manipulated outside of that scope.

2. Data should be accessible via a policy-driven data access interface, taking into account factors such as the accessing party's role and the scope of data availability for the given access purpose.

3. The data access interface should not admit direct access to raw data and should instead provide only data segments to fulfill agreed 'data needs' between data holder and service subscriber allowed for by the data subject's consent.

4. Personal data should remain accessible and alterable to subscribers, and personal information entered into a digital database should be limited to that which is sufficient for the subscriber to participate in the service they have subscribed to.

5. Data stores should only be merged with explicit consent from subscribers according to a new contract agreeing the new 'data needs' between service provider and subscriber.

6. Data store access should be determined by role, and the 'data needs' and/or rights associated with that role is a system wide concept.

ENDORSE will achieve this by bringing together a consortium of data protection legal experts with academic computer science partners and interested industry players. The project will produce a privacy rule definition language which will be used to express data access and data processing requirements derived from the appropriate European directives[3] together with the national implementations of the directives. The language and these legislative instances along with the toolset to create legally compliant privacy policies that can be enforced by IT systems will be released as open source. Two industry players will perform trials using this toolset. One of these partners is a large multi-national insurance organization and the other a start-up web based organization providing communications services online for end users. The methodology for this validation will be defined at an early stage of the project, with a view to evaluating the successful achievements of the project requirements.

3.2 Objectives

The ENDORSE objectives are:

1. To provide a toolset that enables organizations storing personal data to ensure that their data gathering, data access and data manipulation policies and subsequent implementation of those policies is compliant with data protection legislation. The toolset will be released as open source.

---

[3]These include the Data Protection Directive 95/46/EC, the ePrivacy Directive 2002/58/EC, the eCommerce Directive 2003/31/EC and the Data Retention Directive 2006/24/EC

2. To define a Privacy Rules Definition Language capable of expressing legal requirements, privacy policies and data subject consent. The language will be published as open source and submitted as a standard to the appropriate bodies.

3. To provide a collection of Rule Sets that represent data protection provisions derived from the relevant EU directives on data protection and data retention as well as a subset of the national implementations of the relevant provisions. These Rule Sets will be made publicly available.

4. To provide data subjects with the assurance that their personal data is stored, accessed and forwarded in a manner that is compliant with the appropriate data protection legislation and industry best practices.

5. To define a certification methodology which can provide assurances for the end-user that the organization's privacy operating policies are in compliance with appropriate legislation and that the infrastructural data management procedures are derived from these policies.
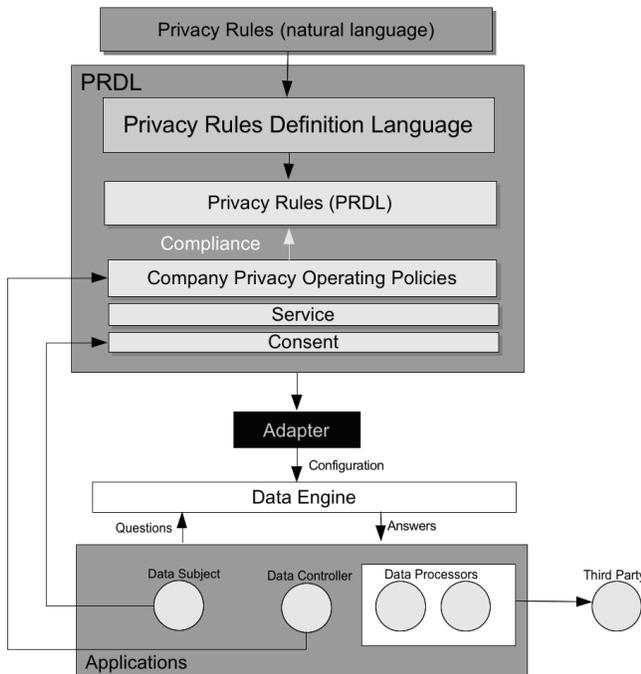
4. THE ENDORSE APPROACH

The following section introduces the architecture overview as well as the domain-specific notation suggested before relevant technologies are covered.

4.1 Architecture

The ENDORSE architecture, as shown in Figure 1, revolves around the concept of a privacy Domain Specific Language (DSL) called Privacy Rules Definition Language (PRDL). This is used to express instances of privacy rules from natural language privacy rule sets (e.g., EU Directive 95/46/EC [1], national implementations of this, industry sector specific privacy guidelines). The PRDL Privacy Rules are written by privacy experts using editing tools developed in the project. These PRDL instances of the rules will be used to ensure compliance with the Company Privacy Operating Policies also expressed in PRDL. The rules are to be stored in the library of rules and maintained in the Rule Repository; here they can be versioned, tested, documented, deleted and maintained. Other actions, such as dependency, import or export features, shall be implemented here. Instances of service offerings, in combination with data subject consent definitions, will comply with this company privacy policy. The PRDL needs to be able to express these concepts as well as mapping data fragment access to roles, scope of usage and consent. A set of adapters will be developed which will have various outputs such as XACML[6], P3P[10], Natural Language Privacy Policies, contracts, data retention triggers, etc. Some of these outputs are used to configure data engines (for data collection, data access, etc,.) with which the actors can interact via a 'questions and answers' type interface using appropriate applications.

A challenge exists with regard to establishing which legal concepts can and can't be easily represented in machine interpretable code. It is often the case where certain terms in legislation are open to interpretation. A good example is the use of the term "informed consent" in European data protection legislation. It is not the role of ENDORSE to provide interpretations of these concepts, rather to provide

Figure 1: ENDORSE Architecture.

a language rich enough to enable organizations to express their own interpretations and definitions of these concepts. With the open source approach of ENDORSE, it is envisioned that generally accepted definitions of such terms will emerge and become "standardized" across the user base.

The specification of the rules are to be written in PRDL using a rule editing tool. PRDL will be a domain specific language based on an existing rules language like RuleML [9] or the open source implementation Business Rules Management System [4]. PRDL will be used to express rules using privacy specific terms and verbs. This semantically rich language abstracts from the target platform technology, such as EJB, Spring, Java or C# and focus on the prime goal, which is how to specify rules for privacy access policy. This approach can guarantee that the rule specification for a specific sector or country can scale to the specific technology used in the organization deploying ENDORSE.

4.2 Privacy Rules Definition Language

The scope of the PRDL will encompass clauses from legislation and service contracts that must be codified. Separate sets of data management and control rules, rendered from operationally different categories of legislative and contractual clauses, may be enforced very differently within the ENDORSE platform. One set of rules may be rendered into XACML for the purpose of vetting queries against access permissions, another set might be rendered into a schedule document for moderating the timing of duties that the data controller must perform. In each case, there is an appropriate adapter for transforming the rules set and an appropriate data engine for processing and enforcing them.

A preliminary content analysis of important data protection and privacy legislation, such as Directives 95/46/EC and 2002/58/EC, and common privacy terms of service from service providers, allows us to come to some provisional con-

clusions on the typical categories of legal and contractual clauses that must be enforced by ENDORSE[4]. These categories of clauses will be enforced either via assurances implicit within the ENDORSE platform, or via a PRDL rule set. Prominent examples of rule sets are:

- clauses that govern conditional access to data by data processors,

- clauses that oblige data controllers to perform certain duties at certain times or under certain conditions,

- clauses that govern the type of data that can be gathered by the controller from data subjects,

- clauses that determine when consent or notification is required from the data subject,

- clauses that moderate how sensitive data is transferred to third parties or across jurisdictions.

Adapters and engines will be developed for these and other important rule sets. The MDA[5] and rules-based approach taken by ENDORSE allows us to express complementary rules sets together within a single framework to achieve a level of systemic integration and assurance not currently possible with today's piecewise solutions. For example, there is a strong logical relationship between data gathering rationale and data processing criteria, since data may only be gathered for the purpose, or purposes, for which it is eventually used. By modeling data gathering rationale, the roles of employees within an organization that process personal information, and the (minimum) set of questions that they are allowed to put to the system, the ENDORSE platform can validate the processing criteria rules against the data gathering rationale and vice versa. In this way, an appropriate ontology and rules language can add significant value to a privacy preserving data management framework, by making it more comprehensive and more efficient.

4.3 Technologies and Tools

4.3.1 Rule Engine

ENDORSE will provide a rule engine that is capable of applying the rules concept - already successfully used for enforcing business rules in enterprises - to data and its privacy. The rule engine (RE) executes the privacy rules formulated in PRDL to create run-time transformation via interfaces to the different technology adapters, providing controlled access to the underlying persistence layer. The rule engine is responsible for consistency checks of rule definitions, conflict management, appropriate data deletion strategies and other specific tasks to be defined in course of the requirements and design phase. The rule engine decouples the privacy data stored in its respective containers (e.g., repositories) from the privacy enforcement mechanisms, thus externalization from application code is guaranteed and flexible extension to third-party systems possible.

---

[4]The PRIME project has, for instance, delivered a set of (legal) requirements that provides a useful starting point (Günter Schumacher (ed.), Requirements for Privacy Enhancing Tools version 3, Deliverable 1.1.d, 20 March 2008, the PRIME Consortium)

[5]Model Driven Architecture, http://www.omg.org/mda/

4.3.2 Technology Adapters

ENDORSE will implement a set of technology adapters as well as a set of Authentication, Access Control and Accountability infrastructural components. The technical adapters will take instances of privacy operating procedures expressed in PRDL as an input and provide various document formats as outputs. One of the most important of these adapters will be an access control policy adapter, of which XACML with a privacy profile is an initial candidate[6]. Other adapters will produce human readable privacy policy statements, machine readable privacy policy statements, data gathering logic and GUI specifications. Standards based specification instances will be the preferred output of these target adapters.

4.3.3 End User Tool

ENDORSE will provide the data subject with a tool for policy inspection, consent inspection, personal data requests and provide audit trails via access to accountability data which can be related to policy and consent instances. Authentication interfaces will be developed to aid in the deployment of the tool in heterogenous environments.

4.3.4 Rules Editor

ENDORSE will to design and develop an open source toolkit that allows users to specify rules. Existing open source modeling tools (e.g., Eclipse Modeling Framework) will be considered as an input to this task. The output of this task is a rule editor. The ultimate goal is to allow users to precisely define rules and behavior in an assisted way.

4.3.5 Open Source Strategy

An open source strategy provides some significant advantages to the ENDORSE approach. Firstly the open source approach lowers the barriers to adoption through free access, providing European SMEs with lower costs when expending resources for data protection compliance. Secondly, this approach encourages wider involvement of the legal, data protection and interested software development communities to contribute to the mutual goals. Thirdly, the open APIs provided by the rules engine for technology adapters offers opportunities for third party software vendors to develop proprietary solutions for integration with enterprise software systems. Items that will be freely available are:

- The PRDL specification.

- PRDL editing environment.

- PRDL engine for execution

- A selection of European legislative rulesets expressed in PRDL.

- A selected set of technology adapters for access control, privacy policy expression, service contract creation, data retention scheduling, and other related data protection control or information revealing modules.

- An example set of rules to demonstrate how privacy policies can be expressed in compliant PRDL instances

and transformed to access configuration and policy statement objects using the technology adapters.

- Instances of actual company privacy procedure policies will not be mandatorily available as open source although creators might deem it advantageous to do so for transparency purposes. Suitable licenses for privacy policies would be Creative Commons[7] or GFDL (GNU Free Documentation License)[8].

4.4 Key Innovations

4.4.1 Legal-Technical data protection compliance

A key innovation of ENDORSE is to provide a technical solution to a known operational legal problem, i.e. the management of legally compliant data access and retention procedures to ensure that organizations operate within the law with respect to data protection and privacy issues and that the data subject can be assured that its data are being used for the the purpose for which it has given its explicit consent and/or is consistent with the law. ENDORSE will use the relevant European directives on data privacy as well as selected set of the national legislation implementing those directives to build a set of rules which the data controller can reference to validate its own policies in terms of data collection, storage, retrieval, manipulation, retention and transmission. The project will operate in this EU context where there exists a harmonization of data protection legislation. We are thus provided with a minimum common "set of rules". ENDORSE will address and express a selection of the national implementations of these rules. It is these national implementations of the Data Protection Directive that needs to be applied as these are what provide the legal obligations on organizations. This means that users can be assured that their data is being used consistently with the applicable national data protection law. It is important that the language used to express these legislative instances are founded in fundamentals of privacy concepts in order that future rule-sets (e.g. for US Law) can be also expressed. Such further rule sets beyond the candidate rule-sets will be created outside of ENDORSE which will provide the tools and fundamentals to do so. These rules are processed via an adapter to create concrete instances of various standard document formats to define data access, privacy policy, user interface forms for data entry and data retention triggers for the data subject, data controller and data processors. Examples of these outputs are XACML[6] instances for data access, P3P [10] instances for web privacy policy creation, XForms[11] documents for the generation of user interfaces. These outputs will be underpinned by the rule set describing the organization's privacy policies which are in turn automatically checked for compliance with the relevant legislation and industry codes of conduct. An example of how this can be done using current standards is the generation and use of privacy policies P3P, which has been criticized for its lack of conformance to the highest standards of data protec-

---

[6]The EU FP7 PrimeLife project is currently extending XACML with privacy policy expression components. This extension is called Primelife Policy Language (PPL). ENDORSE may further expand on this work if the PrimeLife results turn out to be fruitful.

[7]Creative Commons, a non-profit organization providing "some rights reserved" copyright licenses, http://www.creativecommons.org/

[8]A form of copyleft intended for use on a manual, textbook or other document to assure everyone the effective freedom to copy and redistribute it, http://www.gnu.org/licenses/fdl.html

tion and privacy and its lack for enforcement of compliance in the data management infrastructure[3].

4.4.2 Privacy as a cross cutting concern

The key innovation, besides an efficient technical solution to the legal framework problem, is the extraction of privacy issues from the business domain creating a single point of maintenance of personal and private information. This approach intends to isolate privacy and consider it as a cross cutting concern, in the same way as it is currently for authentication, authorization, logging, monitoring, etc,. Legal obligations are often implemented in IT systems as an afterthought and as such can appear in different locations of the software without a common linkage. In many corporate applications this leads to severe difficulty in maintenance when rules and legal specifications change. A separation of concerns can be successfully applied with the help of the proper modeling architecture, design patterns and technical development. It can scale both technically and functionally and as a consequence it can be managed, implemented and deployed in different business domains with no side effect on other organizations' software. This project will define the founding framework and specification for considering the legal framework as a functional cross cutting concern; we expect that ERP, CRM and other corporate software could use ENDORSE to provide privacy and data protection support for the organizational legal framework.

4.4.3 Increased transparency for end users

One of the tools that ENDORSE will provide is a powerful end user verification tool which provides transparency for end users with regard to details of the personal data that is being stored on them, how this data may be processed and by whom and for what purposes. It will also be possible for end users to request corrections and alterations to their personal data via the tool. The tool will act as a means of policy and consent inspection and through the integration of the accountability module will allow users to examine who issued requests to access their data, for what purpose and whether the request was granted or not. These audit trails can be linked to appropriate policy and consent (or exception) instances providing complete transparency to the personal data management life cycle.

One initial goal of the project is promote standardized delivery of rules to the user in terms of both terminology and presentation, thus facilitating clarity and transparency of business practices. Standardized terms of service, for example, would greatly increase user readability and accessibility in itself. These gains can be extended by further translating rules into language that may be more friendly to the average user.

4.4.4 Certification

One of the outputs of the ENDORSE approach is a toolset and technologies with the ability create digitally signed Privacy Seals and software objects to enable the provision of a "certificate of data protection compliance", which on one hand provides the data subject with assurances of compliance, while on the other allows the data controller to ensure that its procedures are in line with national data protection legislation and also sectoral codes of practice. The advantage for the data subject of this certification is the assurance that its personal data is stored, accessed and updated not only in a manner that is in line with the organization's privacy policy but also with the appropriate national legislation and European directives. The advantage for the data controller is the assurance that its data processing and storage procedures are automatically in line with its own privacy policy and also the national legislation and European directives on data protection. ENDORSE will provide a means of verifying that configuration of infrastructural components is compliant with appropriate data protection legislation. The project will develop a certification methodology and tools to produce a digitally signed seal which can verify that the company operating privacy policy is compliant with a rule set representing data protection legislation and industry codes of practice. This methodology will specify how such a seal can be verified by a trusted third party.

5. APPLICATION AREAS

ENDORSE will directly address two primary application areas in which the legal-technical framework will be deployed:

1. Web application service provider dealing with personal information including credit card details and contact details.

2. Large scale public or private databases dealing with sensitive information. e.g., large scale databases used in healthcare and private sector bank/insurance databases.

In each case the concerns for data subjects and controllers are similar, but the deployment of our framework might be quite different. Both of these are discussed below together with a discussion on complex layers of compliance exemplified in the healthcare sector, which will also be considered by ENDORSE.

5.1 Web application service provider

The project is concerned with intentional and unintentional data disclosures, such as the sale of data-sets for marketing purposes, whether anonymized or not (anonymization has been shown to be not as effective as is generally thought [8][5]), non-standard, opaque privacy terms of service, and the ad-hoc update of these terms that retrospectively affect data already passed from data subject to controller. The situation is complicated by the cross border nature of web services, where users are often not protected to the extent of the laws in their own jurisdiction (e.g., safe harbor agreement between the US and EU[2]). It is also true that web application providers may be inadvertently non-compliant with the relevant data protection legislation. There are currently no set of rules and/or software components that aid application providers in maintaining data protection compliance.

To address these points, ENDORSE will create and promote:

1. standard and compliant privacy policy format as part of terms of service for web companies,

2. open source software components to ensure compliance,

3. a mechanism for mapping data usage to data gathering rationale to avoid scope creep in the use of sensitive data. Often collected personal data can currently be

accessed by any arbitrary SQL query and new data can be synthesized.

A means of certification will be a powerful legal component of the framework for standardizing good practices in the above areas. The certification methodology will be crafted to ensure legal compliance across a number of jurisdictions, encapsulate jurisdiction specific terms and provide a high-water mark for privacy and data protection to be adopted by 'privacy friendly' application providers. If a de-facto privacy/data protection terms of service and policy statements (and set of software components enforcing these) could find adoption amongst existing and emerging web applications, this would constitute a major advance in user privacy and data protection compliance and a significant potential deliverable of this project.

5.2 Large scale public and private databases

The main source of accidental data disclosure appears to be poor data handling practices (e.g., copying databases to CDs unencrypted, where they can potentially be brought outside of the organization and subsequently lost or stolen), inadequate internal security (often arising from poor workflows, e.g., sticky note with administrator password on computer monitor), and hacking. Intentional database disclosure occurs when organizations exchange their subscribers' sensitive data or a wider range of personnel gain access to more data when databases are merged following organizational mergers unless strict controls are applied. What is most important in this case is the granularity of data access controls by employee, context, data collection rationale and the principle that only the minimum amount of data be revealed in order to facilitate the task in hand. Additionally, these access controls must be preserved as large data-sets are merged, when the likelihood of additional data items becoming synthesized and revealed to a greater number of individuals in a broader range contexts becomes higher.

In relation to this, ENDORSE will:

1. Promote dynamic, granular data access policies at all times. Access should be restricted by employee, context and data collection rationale and never be presented wholesale in a form that is amenable to duplication and distribution,

2. Ensure that databases are accessed only in a manner that facilitates minimum data disclosure,

3. Data access could be based on a set of 'questions and answers' that are relevant to an employee in a given context. These Q&As are defined and agreed by in-house data protection officers and facilitated by database admins.

4. Produce a procedure and set of tools for merging databases in such a way that no additional privacy or data protection concerns arise for the composite database and access policies than were there for both databases individually.

Currently there exists no standard means of ensuring that databases are accessed according to data protection logic, and in particular there has been no attempt to map employee-context access to data gathering rationale, which must occur to ensure data protection compliance in the EU and to promote maximum user privacy in an environment where an ever greater number of individuals have access to ever larger databases.

5.3 Layers of data protection issues, particularly exemplified in healthcare

The Data Protection Directive distinguishes three layers: general data protection provisions, provisions with respect to sensitive data, and provisions regarding data disclosure to third countries. Within the different member states additional layers can be distinguished. In many sectors, the general data protection regulation is supplemented by sector-specific regulation. This is particularly the case in healthcare where strict provisions are imposed on the collection and use of personal data. Here, for instance, we may find provisions stating that only practitioners with a contract to treat a particular patient may access this patient's data. Another set of sector-specific data protection provisions may be found in the work sphere where labour law contains specific provisions regarding the rights and obligations of employers regarding their employees. An example here is workplace monitoring which may be subject to different requirements in the different EU member states. All in all, these different layers result in a complex mesh of provisions that see to the collection and use of personal data, which are difficult to assess for businesses, especially in the case of cross border services.

ENDORSE will create:

1. A rule language that allows the representation of relevant provisions regarding the collection and use of personal data from the different legal sources (e.g., EU directives, national implementations and national legislation, sectoral provisions).

2. Develop a rule engine that is capable of combining the applicable rule sets for a particular context (consisting of a set of one or more service providers in potentially different jurisdictions, the application domain, and the jurisdiction of the data subjects) into a comprehensive, and ideally non-conflicting, set of requirements. In the case of rule conflicts, the engine will reveal these conflicts.

3. Develop an end-user tool that shows the applicable rules in their context as distilled by the rule engine, in a comprehensible manner.

6. SUMMARY

The ENDORSE project aims to ease overheads of compliance with regard to data protection of personal data stored in organizational databases. This is motivated by the needs of data controllers and data subjects. ENDORSE will use an MDA approach to transform privacy policies expressed in a rules based language to target policy and access control languages and as such provide assurances of legislative compliances in a manner that is transparent to and easily understood by data subjects. The resulting open source technologies will be trialled in two different real world commercial settings and can be reused by any organization with concerns for data protection compliance.

7. ACKNOWLEDGMENTS

8. REFERENCES

[1] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. Official Journal L No. 281, November 1995.

[2] European Commission. Commission Staff Working Document - The implementation of Commission Decision 520/2000/EC on the adequate protection of personal data provided by the Safe Harbour privacy Principles and related Frequently Asked Questions issued by the US Department of Commerce SEC(2004)1323. http://ec.europa.eu/justice/policies/privacy/docs/ adequacy/sec-2004-1323_en.pdf.

[3] European Commission, DG XV, Working Party on the Protection of Individuals with regard to the processing of Personal Data. OPINION 1/98, Platform for Privacy Preferences (P3P) and the Open Profiling Standard (OPS). http://epic.org/privacy/internet/ec-p3p.html, June 1998.

[4] JBoss. JBoss Enterprise BRMS. available at http://www.jboss.com/products/platforms/brms/.

[5] B. Malin, L. Sweeney, and E. Newton. Trail Re-identification: Learning Who You are From Where You Have Been, Data Privacy Laboratory Technical Report. Technical Report LIDAP-WP12, Carnegie Mellon University, School of Computer Science, Pittsburgh, February 2003.

[6] Organization for the Advancement of Structured Information Standards(OASIS). eXtensible Access Control Markup Language. http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=xacml.

[7] Ponemon Institute, LLC. 2009 Annual Study: Global Cost of a Data Breach. available at http://www.securityprivacyandthelaw.com/uploads/ file/Ponemon_COB_2009_GL.pdf, 2010.

[8] C. Soghoian. The Problem of Anonymous Vanity Searches. *I/S: A Journal of Law and Policy for the Information Society*, 3(2), 2007.

[9] The RuleML Initiative. Schema Specification of RuleML 1.0. available at http://ruleml.org/1.0/, August 2010.

[10] World Wide Web Consortium (W3C). Platform for Privacy Preferences. http://www.w3.org/P3P/.

[11] World Wide Web Consortium (W3C). XForms. http://www.w3.org/MarkUp/Forms/.