A Systemic Approach to Automate Privacy Policy Enforcement in Enterprises

Marco Casassa Mont[1], Robert Thyne[2]

[1] Hewlett-Packard Laboratories, Trusted Systems Lab,
Bristol, United Kingdom
marco.casassa-mont@hp.com
[2] Hewlett-Packard, Software Business Organisation,
Toronto, Canada
robert.thyne@hp.com

Abstract. It is common practice for enterprises and other organisations to ask people to disclose their personal data in order to grant them access to services and engage in transactions. This practice is not going to disappear, at least in the foreseeable future. Most enterprises need personal information to run their businesses and provide the required services, many of whom have turned to identity management solutions to do this in an efficient and automated way. Privacy laws dictate how enterprises should handle personal data in a privacy compliant way: this requires dealing with privacy rights, permissions and obligations. It involves operational and compliance aspects. Currently much is done by means of manual processes, which make them difficult and expensive to comply. A key requirement for enterprises is being able to leverage their investments in identity management solutions. This paper focuses on how to automate the enforcement of privacy within enterprises in a systemic way, in particular privacy-aware access to personal data and enforcement of privacy obligations: this is still open to innovation. We introduce our work in these areas: core concepts are described along with our policy enforcement models and related technologies. Two prototypes have been built as a proof of concept and integrated with state-of-the-art (commercial) identity management solutions to demonstrate the feasibility of our work. We provide technical details, discuss open issues and our next steps.

1 Introduction

Privacy management is important for enterprises and organisations that handle identities and personal data of customers, employees and business partners: it has implications on their compliance with regulations, their reputation, their brand and customers' satisfaction [19,20].

Privacy laws [1,2] and privacy guidelines, such as OECD [3], dictate that enterprises should clearly state the purposes for which they are collecting personal data and should take into account the consent given by data subjects to use their data for these purposes. In addition, personal data should be deleted once its retention is not required anymore. Openness and transparency over how data is processed, manipulated and disclosed to third parties are also key requirements. Data subjects (users) should be notified of changes affecting the management of

their personal data and they should retain a degree of control over it. Compliance to all these aspects must be monitored and any violations promptly reported and addressed. Furthermore large enterprises that are geographically distributed across different nations might need to comply with different privacy laws. Privacy policies can be used to represent privacy laws and guidelines: they describe data subjects' rights on their personal data, permissions given to enterprises and obligations that enterprises need to fulfil when handling personal data.

On one hand, enterprises have been investing in identity management solutions to automate the management of personal and identity information. This includes solutions to store personal and confidential data and use it for access control and authorization purposes. On top of this single-sign-on mechanisms and federated identity management solutions have been developed to simplify and enable multi-party interactions. Provisioning and user account management solutions have also been developed to simplify users' self-registration process and provision users' information to various enterprises' systems and data repositories. On the other hand, in terms of privacy management, much is still done by means of manual processes, which make them difficult and expensive to comply. Simplification of the involved processes and better control are key enterprise requirements: this leads towards the need to introduce automation also for privacy management.

Most of the technical work currently done in this space focuses on auditing and reporting solutions to analyse logged events and check them against privacy policies. This addresses compliance aspects of privacy management. However, also operational aspects of privacy need to be addressed. In particular, the enforcement of privacy policies is very important to guarantee that personal data is accessed, used, disclosed and managed according to these policies. Often privacy policies are hardcoded into enterprise applications and services or managed with very vertical, ad-hoc solutions, in specific contexts. This approach is not adaptive to changes and does not scale. The enforcement of privacy rights, permissions and obligations on confidential and personal data requires the mapping of these concepts into rules, constraints and access control, the meaning of which must be unambiguous so that it can be deployed and enforced by software solutions. This still requires following best practices and good behaviours. However, automating aspects of the enforcement of privacy policies can really help enterprises to improve their practice and simplify the overall management. This paper describes our systemic approach to automate the enforcement of privacy policies (inclusive of obligations). Our technology can be integrated with enterprise middleware solutions, in particular identity management solutions.

2 Addressed Problem

This paper focuses on the problem of how to automate the enforcement of privacy policies within enterprises by keeping into account privacy laws, enterprise guidelines and data subjects' privacy preferences. As anticipated, privacy policies dictate privacy rights, permissions and obligations. Addressing the problem of automating their enforcement requires dealing with: (1) *privacy-aware access to personal data*; (2) *enforcement of privacy obligations*.

Our goal is to address this by developing a privacy enforcement framework and a systemic approach that can be leveraged by current enterprise identity management solutions.

3 Important Issues and Requirements

In the remaining part of this paper, for simplicity, we will use in an interchangeable way the terms: "data subjects", "people" and "users". We consider scenarios where users are asked by enterprises (e.g. a service provider) or other organisations to disclose their personal information in order to access services, engage in transactions or access information.

We want to enable *users* to specify their privacy preferences and dictate obligations on how their data should be managed, give explicit consent and specify limitations about the usage of their data. We want to provide them with degrees of control on their personal data. We also want to enable *enterprises* to: keep into account users' privacy preferences and enforce them; explicitly author privacy policies and obligations, deploy and enforce them during accesses, manipulations and transmission of personal data. Enterprises need tools to achieve this but at the same time ideally they would like to leverage their investments in identity management solutions.

The (technological) enforcement of privacy permissions and rights (on stored personal data) requires extended access control and authorization mechanisms that check these privacy permissions against data requestors' credentials, check the consistency of data requestors' intent against stated purposes and take into account the consent given by data subjects [19]. This applies, for example, to enterprise services or applications that need to access and manipulate personal data for various reasons. Traditional access control systems are necessary but not sufficient to enforce privacy policies on personal data. They are mainly based on "access control lists" and enforcement mechanisms that keep into account only the identities of data requestors, their rights and permissions and the types of actions that are allowed/disallowed on the involved resources. These systems do not keep into account additional aspects relevant to privacy enforcement: the *stated purposes* for collecting data and data subjects' consent - i.e. properties usually associated to collected data - the *intent of data requestors* and any additional enterprise or customized data subjects' *constraints*. To address the above issues and move towards privacy-aware access control systems to protect personal data, it is important to satisfy the following core requirements: (1) *Explicit modeling of personal data stored by enterprises*; (2) *Explicit definition, authoring and lifecycle management of privacy policies*; (3) *Explicit deployment and enforcement of privacy policies*; (4) *Integration with traditional access control and identity management systems*; (5) *Simplicity of usage of all the involved system*; (6) *Support for auditing*. A more comprehensive analysis of these aspects can be found in [19].

Even more complex is the case of dealing with the enforcement of privacy obligations [20,21]. Privacy obligations dictate criteria for a privacy-aware information lifecycle management. They might require the deletion or transformation of confidential data after a predefined (potentially very long) period of time, periodic notifications and requests for authorization to data subjects, fulfilment of opt-in/opt-out choices made by data owners, ongoing compliance with laws' obligations and internal guidelines. Privacy obligations can have ongoing aspects that need to be monitored and satisfied over a long period of time. All these tasks are challenging for enterprises because of the need for specific IT infrastructures and processes able to manipulate confidential data as dictated by privacy obligations. It is important that privacy obligation management solutions address the following core requirements: (1) *Explicit modeling and representation of privacy obligations*; (2) *Association of obligations to data*; (3) *Being able to timely enforce privacy obligations*; (4) *Mapping obligations into enforceable actions*; (5) *Compliance of refined obligations to high-level policies*; (6) *Tracking the evolutions of*

obligation policies; (7) *Dealing with long-term obligation aspects*; (8) *Accountability management*; (9) *Monitoring obligations*; (10) *User involvement*. A comprehensive analysis and discussion of these aspects can be found in [20,21].

4 Our Work

This section describes our work to automate the enforcement of privacy policies and privacy obligations on personal data stored by enterprises. Our approach consists of researching and building solutions that can be leveraged by current enterprise identity management solutions. In particular, our approach focuses on the following (typical) enterprise identity management processes (already supported by current identity management solutions), which occur when a new user wants to access services or applications that might require financial or business transactions:

1. The user (data subject) is asked to access a self-registration web site and provide their personal information and other requested data. Some privacy preferences might also be asked to the user and stored. The user later on will be allowed to change their information and preferences;
2. Provisioning and user account management solutions are used to manipulate user's information and store (parts of) it within relevant enterprise data storages. The same provisioning solutions will take care of creating user accounts across enterprise' relevant systems and set proper access control on these resources. These provisioning tool will track changes happening on stored information and ensure that information is kept aligned and consistent;
3. As an effect of the previous provisioning step, authorization and access control systems have been provisioned (by means of access control constraints, new user accounts, etc.) and will be able to and grant (or deny) access to services.

The above steps usually focus only on the automation of identity management aspects. Privacy aspects are either not included or their enforcement is not automated. In addition, personal data is stored in enterprise data repositories subject only to security aspects. As summarised in Figure 1, our work wants to:

1. Enable users to explicitly define their privacy preferences and customise them during their self-registration phase;
2. Use users' privacy preferences, during the provisioning phase, to:
    a. Configure extended access control systems to provide privacy-aware access to personal data: this includes ensuring that these systems can keep track of stated purposes, data subjects' consent and other privacy constraints;
    b. Turn parts of these privacy preferences (such as deletion date of data, notification choices, etc.) into explicit privacy obligations to be enforced by enterprises.
3. Allow enterprises to author, deploy and enforce "enterprise-side" privacy policies and privacy obligations derived from privacy laws and internal guidelines.

Section 4.1 describes our work on privacy-aware access control. We introduce our privacy-aware access control model. We illustrate a prototype that we have built by leveraging and extending HP Select Access [14] (a state-of-the-art access control solution) to deal with privacy policy enforcement on personal data.

Section 4.2 describes our work on privacy obligation management within enterprises. It provides details of our obligation management model along with our prototype of an obligation management system. We also describe how we have successfully integrated it with HP Select Identity [23], a state-of-the-art provisioning and user account management solution.

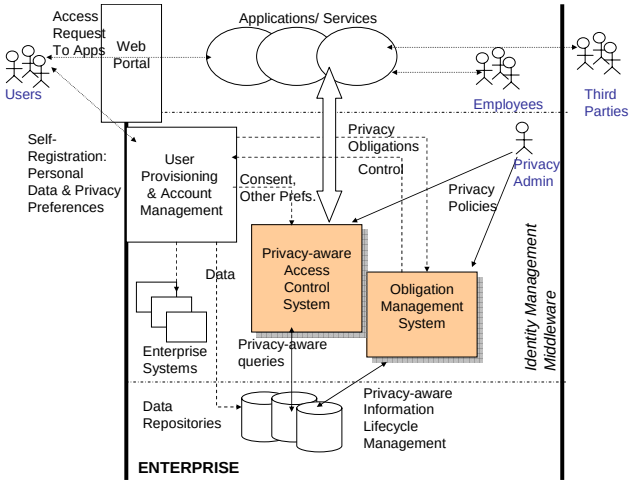

Figure 1: Automation of Privacy Enforcement within Identity Management Solutions

4.1 Privacy Policy Enforcement

Our approach to enforce privacy policies on stored personal data is based on a privacy-aware access control model. This model extends traditional access control models (based on users/groups, users' credentials and rights, access control lists and related policies) by explicitly dealing with the *stated purposes* for which data is collected, checking - at the access request time - the *intent* of requestors against these purposes, dealing with data subjects' *consent* and enforcing additional access conditions and constraints defined by data subjects and/or enterprise administrators [1,2,3] – see Figure 2. The main aspects of this model are:

a) **A mechanism for the explicit modelling of personal data that are subject to privacy policies**: this mechanism provides a description of data including the type of the data repository (database, LDAP directory, etc.), its location, the schema of these data, types of attributes, etc.;

b) **An integrated mechanism for authoring privacy policies along with traditional access control policies**: it is a Policy Authoring Point (PAP) to allow privacy administrators to describe and author privacy policy constraints and conditions (including how to check consent and data purpose against requestors' intent and how to deal with data filtering and transformation, etc.) along with more traditional access control policies based on security criteria (e.g. who can access which resource, given their rights and permissions);

c) **An integrated authorization framework for deploying both access control and privacy-based policies and making related access decisions**: it is an integrated Policy Decision Point (PDP);

d) **A run-time mechanism –referred to as the "Data Enforcer" - for intercepting attempts to access personal data and enforcing decisions based on privacy policies and contextual information**, e.g., intent of requestors, their roles and identities, etc. It is a Policy Enforcement Point (PEP). This mechanism is in charge (among other things) of dealing with the transformation of queries to access personal data (e.g. SQL queries) and filtering part of the requested data, if their access is not authorised for privacy reasons.

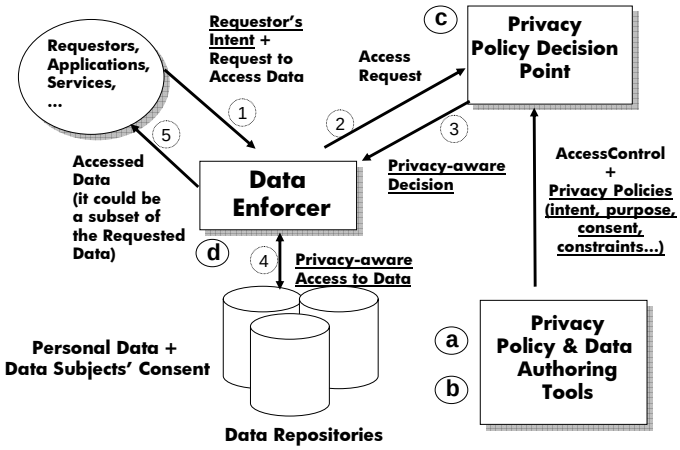

Figure 2: Model of our Privacy-aware Access Control System

The Data Enforcer component plays a key role to enforce privacy policies on personal data. At "run-time", attempts to access personal data are intercepted and managed in the following way - Figure 2:

1. A request from a data requestor to access personal data is intercepted by the data enforcer. Available information about the requestor (credentials, identity, etc.) is collected, along with their *intent* (that can be explicitly passed as a parameter or could be predefined in the application/service making the request);
2. The data enforcer interacts with the privacy policy decision point by passing information about the request (including the intent) and the requestor;
3. The privacy policy decision point makes a decision, based on available privacy policies and the context (request, requestor's information, etc.). This decision is sent back to the data enforcer. It can be any of the following types:
    - ***Deny***: access to data is denied;
    - ***Deny + conditions***: access to data is denied. Some conditions are sent back to the requestors. The satisfaction of these conditions (for example passing the intent or authenticating) could change the outcome of the decision;
    - ***Allow*****:** access to data is granted;
    - ***Allow + conditions****:* access to (part of the) data is allowed, under the satisfaction of the attached conditions. Among other things, these conditions might require data filtering, transformations and manipulations.

4. The data enforcer enforces this decision. In particular, if the decision is "*Allow + conditions*" the data enforcer might have to manipulate the query (query pre-processing) and/or transform the requested personal data (result post-processing), before returning the result to the data requestor;
5. Data (or alternatively no data) is returned to the data requestor, based on the enforced decision.

Figure 3 shows a simple example based on this model where an attempt to access personal data is made by an enterprise employee.

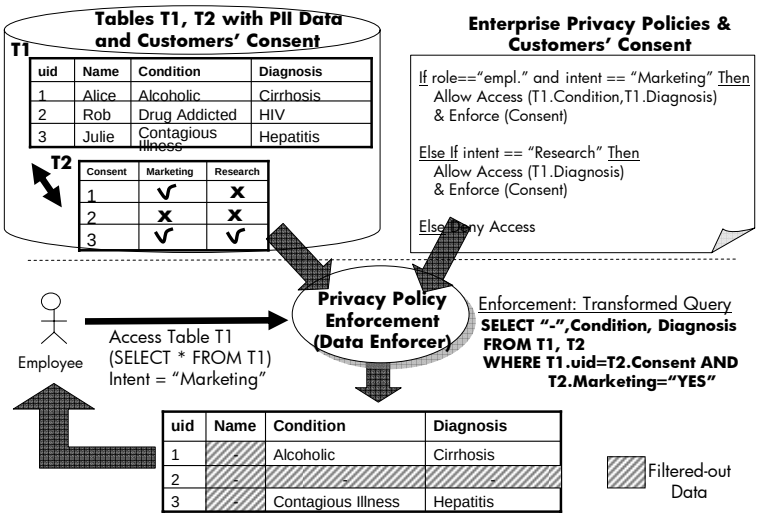

Figure 3: Example of Privacy Policy Enforcement

In this example, the employee's declared *intent* (i.e. *marketing*) is consistent with the declared *purposes* of data (*marketing, research*). However the employee is trying to access – via a SQL query - more data than she is allowed to. The SQL query is intercepted by the enforcement point (data enforcer) and transformed on-the-fly (before being submitted to the database) in a way to include constraints based on data subjects' consent and the filtering of data. The transformed query is then submitted to the database. In this example privacy is achieved by pre-processing and transforming the query before actually interacting with the database.

We implemented our privacy enforcement model by leveraging and extending HP Select Access. HP Select Access [14] is a leading-edge access control product. It provides policy authoring, policy decision and policy enforcement capabilities via the following components:

- **Policy Builder**: it is a graphical tool to author access control policies (PAP) on resources managed by the system;
- **Validator**: it is a Policy Decision Point (PDP). It makes access control decisions based on the access control policies (authored with the Policy Builder) and contextual information, such as the identity of a requestor;

- **Web Enforcer plug-in**: it is a Policy Enforcement Point (PEP) for web resources.

The current commercial version of HP Select Access does not handle data as managed resources: it only deals with traditional access control policies on web resources. Additional functionalities have been added to HP Select Access (HP SA) in our prototype, to explicitly deal with privacy-aware access control on personal data, as shown in Figure 4:

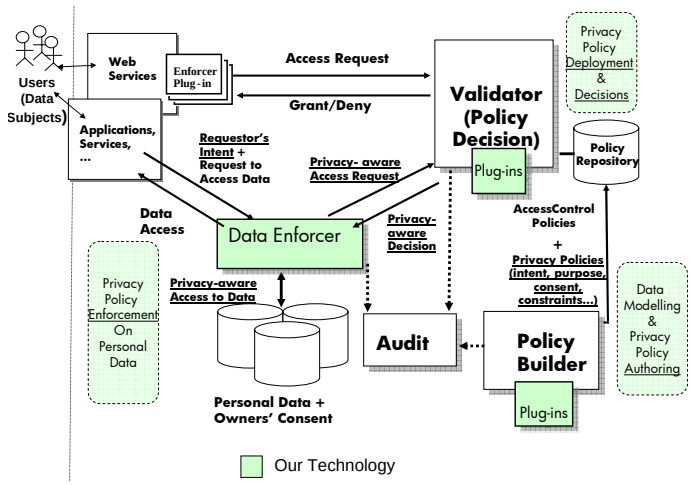

Figure 4: Extended HP Select Access to deal with Privacy Policy Enforcement

The specific extensions are:

- The HP SA Policy Builder has been extended to represent "data resources" (databases, LDAP directories, virtual-directories, their schemas, etc.) in addition to traditional IT resources (such as web resources);
- The HP SA Policy Builder has been extended to *graphically* author privacy policies on "data resources" in addition to traditional access control policies: a set of additional plug-ins has been implemented to allow checking (at the enforcement time) the requestor's intent against the stated data storage purposes, take into account data subjects' consent and data retention policies and describe how the accessed personal data must be filtered, obfuscated or manipulated, etc. By using this tool administrators can manage the lifecycle of both privacy and security policies, in an integrated environment – based on the same principles and GUI. This simplifies the overall policy management process and differentiates our approach from related work (see section 5);
- The HP SA Validator has been extended to make privacy-aware decisions. Plug-ins, correspondent to the ones used in the Policy Builder, have been implemented. This enhanced-version of the Validator can now also make "*Allow + conditions*" decisions as described in our model;
- A Data Enforcer has been built and added to the framework: this is a new functionality added to HP Select Access. It is in charge of enforcing privacy decisions made by the Validator, as previously described in our model. The data enforcer proxies managed data repositories (e.g. databases, LDAP directories, virtual directories, etc.): we envisage that a

family of data enforcers (sharing a common logic but differentiated by add-ons dealing with different types of data resources) need to be built, because of the different semantic of different data repositories. As a proof of concept, we implemented a data enforcer as a JDBC proxy for RDBMS databases.

The above functionalities address and satisfy the core requirements described in section 3 for privacy enforcement on personal data. Policy authoring and enforcement processes are audited by the HP Select Access's Audit Server, for accountability and compliance management.

Figure 5 provides additional details about the data enforcer that we developed to intercept SQL queries for RDBMS databases and enforce privacy policies on the requested data.

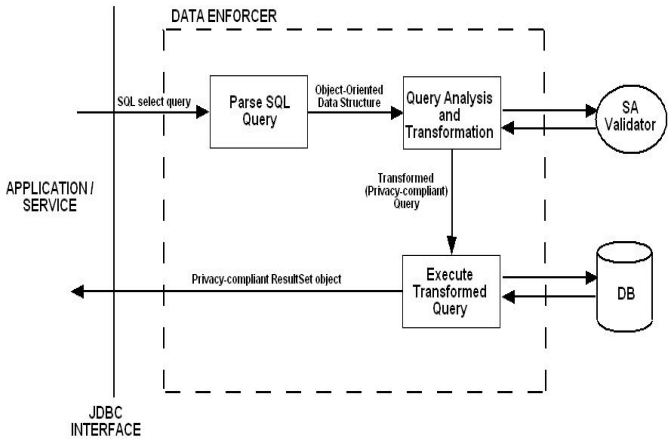

Figure 5: Internal architecture of the Data Enforcer

The data enforcer is based on a JDBC Proxy (JDBC driver). Applications and services do not need to be modified apart from having to use this JDBC driver. Standard JDBC APIs are used. The data enforcer intercepts applications' SQL queries and processes them.

The *intent* (reason for accessing data) of a data requestor (e.g. application) could be implicit in its role: this ensures the most transparent interactions between applications and our data enforcer. In case the *intent* information has to be explicitly passed by the application to the data enforcer, we support two mechanisms to achieve this: (1) the *intent* information is added by the application at the end of its SQL query – before submitting it; (2) the *intent* information is passed as a property object by the application, via the JDBC API getConnection method.

The "**Parse SQL Query**" component intercepts incoming SQL queries (SELECT, UPDATE, CREATE, etc.), parses them and generates an explicit tree-based (object-oriented) representation of these queries. This representation clearly identifies, given an arbitrary SQL query, what the involved *data resources* are (e.g. DB tables, fields, etc.), *intent* information, SQL conditions on data, etc. The "**Query Analysis and Transformation**" component - for each involved data resource - checks with the Validator if any privacy policy applies. In doing this it will pass relevant contextual information (requestor's identity, *intent*, etc.) to the Validator. If privacy policies apply, related decisions are recorded. They might include the filtering of some of the data associated to specific fields, the fact that consent has to be enforced, etc. The transformation of the SQL query happens on-the-fly: for example, if specific fields need to be filtered out (because a privacy policy says so), these fields are replaced in the query representa-

tion with default values (as described in the policies). If data subjects' consent has to be enforced, additional JOIN conditions are added into the query representation to check for data subjects' consent information. See Figure 3 for an example. The outcome of this module is a transformed SQL query that keeps into account all the stated privacy constraints and is still compatible with the original stated SQL query. This query is sent from the "**Execute Transformed Query**" to the RDBMS system and executed by the real SQL engine. The result of this privacy-compliant query is sent back to the application/service.

4.2 Privacy Obligation Management

Our work in this area focuses on the explicit management and enforcement of privacy obligations on personal data stored by enterprises. In our model, privacy obligations are "first class" entities, i.e. they are explicit entities that are modeled and managed to provide a privacy-aware lifecycle management of personal data: this includes data deletion, data transformation, dealing with notifications, etc. A related obligation management framework is introduced to manage these privacy obligations. In our vision their management and enforcement must be independent from the management and enforcement of privacy-aware access control policies [20,21]. For example, deletion of personal data has to happen independently from the fact that this data has ever been accessed. This differentiates our approach from related work (see section 5).

A privacy obligation is an "object" that includes (at least) the following aspects: *Obligation Identifier*; *Targeted Personal Data*; *Triggering Events* (e.g. time-based events); *Actions* (e.g. data deletion, sending notifications). Different categories of privacy obligations need to be managed and enforced by enterprises: *transactional obligations*; *data retention and handling obligations*; o*ther types of event-driven obligations.* A complementary classification of our managed privacy obligations is based on their activation timeframe and period of validity: *short-term obligations*; *long-term obligations*; *ongoing obligations*.

In our obligation management framework (a) data subjects can explicitly define their privacy preferences on their personal data at the disclosure time (e.g. during a self-registration process) or at any subsequent time; (b) privacy preferences are automatically turned into privacy obligations based on supported privacy obligation templates; (c) enterprise privacy administrators can associate other privacy obligations, for example dictated by laws or internal guidelines.

Our obligation management framework handles these obligations by providing the following core functionalities: (1) ***scheduling the enforcement of privacy obligations***; (2) ***enforcement of privacy obligations***; (3) ***Monitoring the fulfilment of privacy obligations***.

These functionalities can be accessed by enterprise privacy administrators and potentially also by data subjects, for example to monitor their personal data and check for privacy compliance. Figure 6 shows the high-level architecture of our obligation management system.
A comprehensive description of this obligation management system components can be found in [20,21]. A working prototype has been implemented in the context of the EU PRIME project [22], as a proof of concept, providing the core functionalities: scheduling, enforcement and monitoring of privacy obligations. At the moment the managed obligations are restricted to handling time-based and access based events. The supported actions include deletion of data and notifications. Short-term, long-term and ongoing obligations are supported. Our work addresses the core issues and requirements described in section 3.

This obligation management can be considered as an additional component of current enterprises' identity management solutions. In particular it can be integrated with the self-registration, customization and account management capabilities of identity provisioning systems to allow users and administrators to describe and handle privacy preferences and turn them into privacy obligations for the enterprise. In this context our system allows for the explicitly representation and management of privacy obligations, along with the coordination of their overall enforcement and monitoring.

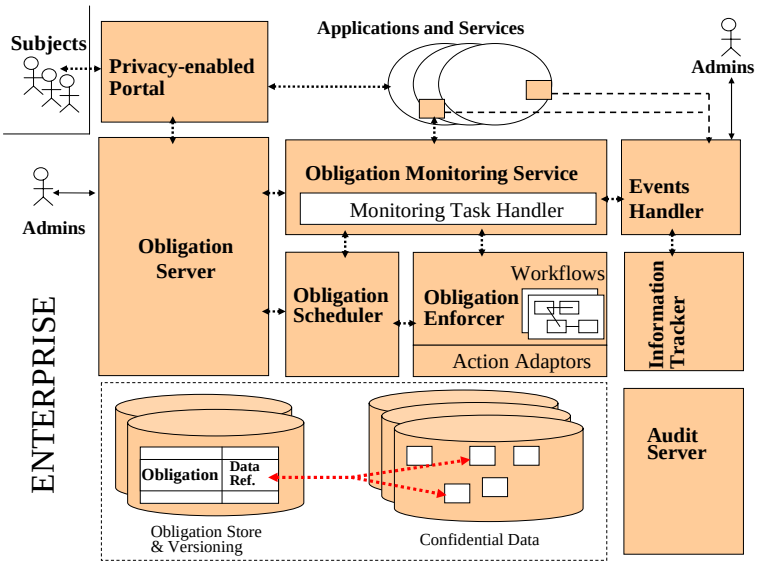

Figure 6: High-level Architecture of our Obligation Management System

To demonstrate how this can be achieved for real, we integrated our Obligation Management System (OMS) with HP Select Identity, as shown in Figure 7. HP Select Identity [23] is a state-of-the-art solution to manage digital identities within and between large enterprises..It automates the process of provisioning, managing and terminating user accounts and access privileges by keeping all this information consistent and synchronised across provisioned platforms, applications and services (within and between enterprise boundaries). Interactions with these third party systems (i.e. data repositories, legacy applications, services, etc.) are achieved via *Connectors*. These third parties can provide feedback to HP Select Identity (via agent-based mechanisms) about changes to their local copies of provisioned data, by calling its Web Service API.

As shown in Figure 7, in our integrated prototype we use (1) HP Select Identity self-registration capabilities to allow users to specify their privacy constraints and preferences along with required personal data. Personal data is provisioned by HP Select Identity to various enterprise systems and data repositories (2). Please notice that at this stage external systems – *such as our privacy-aware access control system* – can be configured with privacy preferences and related constraints. Specifically, privacy preferences are also processed by our OMS connector (2), turned into privacy obligations (based on predefined templates) and pushed to the OMS system (3). Privacy obligations are then scheduled, enforced and monitored by our OMS sys-

tem (4). We leverage the workflow and user/identity management capabilities of HP Select Identity to enforce aspects of privacy obligations (5). Our system retains control of the supervision of obligations and their monitoring (6). HP Select Identity enforces obligations constraints, such as deletion of identities, data transformation, etc. At the moment the deletion of personal data (as the effect of enforcing obligations) is achieved by triggering HP Select Identity workflows, whilst the obligation management system handles the notifications to users.

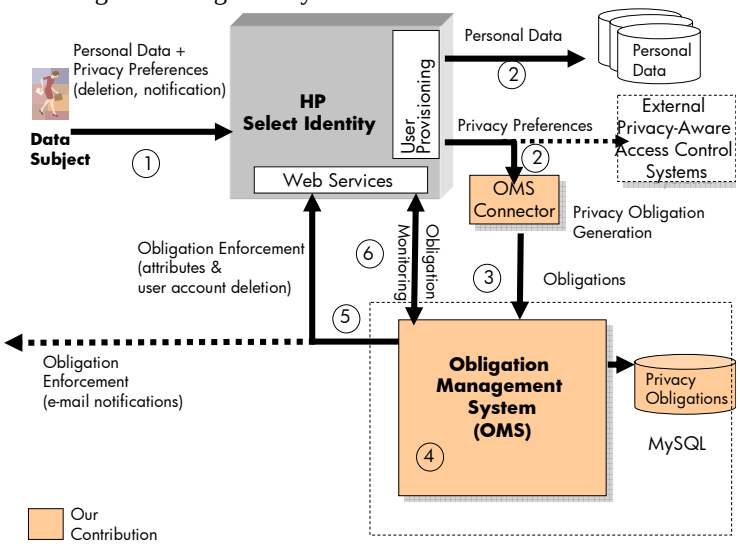

Figure 7: High-level Architecture: Integration of OMS with HP Select Identity

HP Select Identity audits the overall lifecycle of managed personal data. The Audit Server within the OMS system can be used to specifically audit how privacy obligations are authored, managed and enforced.

5 Related Work

A common approach adopted by enterprises to enforce privacy-aware access control policies on personal data consists of hardcoding them within applications and services or building ad hoc solutions. This approach is suitable for very simple and static environments: it shows all its limitations and maintenance costs in case of complex and dynamic organizations that need to adapt to changes. As described in the requirements section, to explicitly address the automation problem, a model of the relevant personal data is required. Privacy policies need to be authored, deployed, enforced and audited. This requires the definition of a comprehensive privacy-aware access control model and systems that implement it. Relevant work in this direction, for privacy management and enforcement in enterprises is described in [4,5,6,7]. An Enterprise Privacy Architecture (EPA/E-P3P) is introduced and described in [7]. This approach is further refined and described in the Enterprise Privacy Authorization Language (EPAL) specification [8]. However these papers mainly provide general guidelines and do not describe an overall deployable solution within current identity management solutions.

The content of Figure 1 is compatible with [4,5,6,7]. Our work differentiates from this because: (1) we do not focus our effort in defining a new privacy-oriented access control language (such as EPAL). Instead, we ensure that privacy aspects (e.g. dealing with data purposes, consent, etc.) can be managed by current identity management system, by leveraging and extending their capabilities; (2) our approach to obligation management is not subordinated to access control, as instead recommended by [4,5,6,7]; (3) we described how our privacy management capabilities can actually be integrated with state-of-the art identity management solutions.

Important related work on actual privacy enforcement on personal data has been done on Hippocratic databases [9] and similarly on Oracle databases (Private Virtual database/Privacy Manager component). The drawback of this approach is that it mainly focuses at the database level, specifically on RDBMS data repository architectures and related data schemas. The enforcement of privacy policies can span across a broad variety of data repositories and legacy systems to include LDAP directories, meta/virtual directories, file systems and legacy systems. It might need to incorporate higher-level views and perspectives than just the database-level perspective.

In terms of commercially available solutions, IBM Tivoli Privacy Manager [10, 11] provides mechanisms for defining fine-grained privacy policies and associating them to data. On one hand this solution provides the required privacy enforcement functionalities. On the other hand this approach dictates strong constraints on how applications need to be developed and how personal data has to be stored and administered: it might require some duplications of administrative and enforcement frameworks (e.g. it requires the parallel usage of Tivoli Access Manager) and it is vertically-based on other IBM products and solutions. Other products, such as HP Select Federation [12] and ePok [13], focus on single-sign-on and related privacy aspects: they enforce privacy rules on personal data in federated environment when these data are disclosed by an organization (or an identity provider) to other parties.

Our work on privacy-aware access control specifically addresses the problem of enforcing privacy policies on personal data stored in a broad variety of data repositories within enterprises. This is a major differentiator compared to related work. Personal data can be accessed by different types of requestors, including people, applications and services. It includes related aspects of modeling the managed data and authoring privacy policies. Our work aims at not being invasive for applications and services: privacy policies are managed in an explicit way, in conjunction with traditional access control policies and not hardcoded in applications and services. We avoid duplication of efforts by providing a single, integrated framework for authoring, administering and enforcing both traditional access control and privacy policies. This has been demonstrated in the way we leveraged and extended HP Select Access [14] to enforce privacy policies on personal data – along with security policies.

In terms of managing and enforcing privacy obligations, relevant work is described in [4,5,6,7,8], in particular the EPAL specification. As previously described, their approach to handling privacy obligations is driven from an authorization and access control perspective. However, privacy obligations typically cannot be managed solely from an authorization-based perspective. Similar observations apply to XACML.

Our approach addresses this issue. In our work obligation policies are first-class entities with their explicit and independent management. Our architecture has high-level commonalities with the architecture described in [4,5,6,7] but in our work we further refine the concept of obligations and their enforcement. We split the enforcement mechanisms in two parts by including a

scheduling mechanisms and an enforcement mechanism allowing for workflow automation and human intervention.

Approaches to deal with (privacy) obligations have already been implemented in products, in particular for data retention [15] and in a variety of document management systems. Nevertheless, these approaches are very specific, focused on particular domains and handle simple obligation policies on files and documents, not really on personal data. Our work aims at pushing the barrier even further to create an obligation management framework that can be leveraged in multiple contexts, for different purposes and that can be actually integrated with state-of-the-art identity management solutions. The feasibility of our work in the real world has been demonstrated by integrating it with HP Select Identity [23] – in a context of user provisioning and user account management.

A lot of work has been done in representing privacy policies, including obligations such as [16,17]. Paper [24] provides a formal definition and classification of obligations, in a data protection context. Relevant work on mechanisms to associate policies to data is described in [4,5,6,7,18]. We can leverage aspects of this work to provide a stronger association of obligation policies to confidential data.

6. Discussion and Next Steps

Our prototypes are proof of concepts. However they show the feasibility of our work to address the enforcement of privacy policies and obligations in a systemic way, integrated with state-of-the-art identity management solutions. We are refining and extending them for their potential productisation. It is important to highlight the fact that our models and technologies are general purpose: they can be leveraged, integrated and deployed in other identity management contexts, beyond HP identity management solutions.

We believe that, by leveraging and extending current identity management solutions, we reduce the barrier that enterprises might have in adopting our privacy management solutions - if compared to other approaches where new, additional privacy solutions need to be used. We are currently exploring opportunities for technological trials with HP customers to further investigate this point.

Another important aspect characterising our work is the way we manage privacy policies. As anticipated, our management of privacy-aware access control policies is integrated with the management of traditional (security-based) access control policies. This simplifies administrators' tasks that only need to use one tool and a related GUI. We also automate the creation of privacy obligations, based on predefined templates – at least in context of user provisioning and account management contexts. However additional work needs to be done in terms of implementing a more comprehensive lifecycle management of privacy obligations. "User studies" can help to show how data subjects and administrators deal with the overall system. We are collaborating with Karlstad University on this topic.

At the moment the enforcement of privacy policies in HP Select Access mainly consists in enforcing data subjects' consent, constraints on data purposes and data expirations via data filtering. Current performance tests and analysis (done on databases of sizes from 100K to 500K records) are promising. No noticeable loss of performance (i.e. the time spent between sending a query to a RDBMS and retrieving the last returned record) has been registered so far, on common SQL queries. More tests and experiments are in progress on different varieties of

SQL queries. We are also planning to: (1) explore the implications of post-processing queries (post-processing of query results) to extend the current set of managed privacy constraints; (2) explore the enforcement of privacy policies on LDAP repositories and virtual directories.

In terms of privacy obligation enforcement, we are currently refining the integration of our obligation management system with HP Select Identity, specifically to leverage as much as possible the provisioning and workflow capabilities of HP Select Identity to enforce obligations' actions. Additional work and research in the space of privacy obligations is going to be done in PRIME [22]: in particular we plan to research on how to make the obligation management system scalable to cope with large amounts of personal data. A promising research topic to explore is the management of parametric obligations that apply to a large subset of personal data subject to similar privacy preferences.

7. Conclusions

Privacy management is becoming more and more important for enterprises to ensure their compliance to regulation, their governance objectives and address customers' preferences and rights. This paper focuses on how to automate the enforcement of privacy policies and privacy obligations on personal data, stored and accessed by enterprises. We discussed a privacy-aware access control model to enforce privacy policies on personal data - including handling the purpose of data, checking data requestors' intent against data purposes and enforcement of data subjects' consent. We also analysed aspects and concepts related to privacy obligations, considered in our model as "first-class" entities (i.e. not subordinated to access control) and introduced our obligation management framework to schedule, enforce and monitor them.

Working prototypes have been implemented and integrated with state-of-the art identity mangement solutions: specifically we described our work to add privacy policy enforcement to HP Select Access and the integration of obligation management and enforcement capabilities with HP Select Identity, in a context of user provisioning. These technologies are ready for commercial exploitation. Research and development work continues to refine our technolgies and implement adiditional functionalities, in particular in the context of the PRIME project.

References

1. C. Laurant, "Privacy International: Privacy and Human Rights 2003: an International Survey of Privacy Laws and Developments, Electronic Privacy Information Center (EPIC)", Privacy International. http://www.privacyinternational.org/survey/phr2003/ 2003
2. Online Privacy Alliance, "Guidelines for Online Privacy Policies", http://www.privacyalliance.org/, Online Privacy Alliance, 2004
3. OECD, "OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data.", http://www1.oecd.org/publications/e-book/9302011E.PDF, 1980
4. G. Karjoth, M. Schunter, "A Privacy Policy Model for Enterprises", IBM Research, Zurich. 15th IEEE Computer Foundations Workshop, 2002
5. G. Karjoth, M. Schunter, M. Waidner, "Platform for Enterprise Privacy Practices: Privacy-enabled Management of Customer Data", 2nd Workshop on Privacy Enhancing Technologies, Lecture Notes in Computer Science, Springer Verlang , 2002

6. M. Schunter, P. Ashley, "The Platform for Enterprise Privacy Practices", IBM Zurich Research Laboratory, 2002
7. G. Karjoth, M. Schunter, M. Waidner, "Privacy-enabled Services for Enterprises", IBM Zurich Research Laboratory, TrustBus 2002, 2002
8. IBM, "The Enterprise Privacy Authorization Language (EPAL), EPAL 1.1 specification", http://www.zurich.ibm.com/security/enterprise-privacy/epal/, IBM, 2004
9. R. Agrawal, J. Kiernan, R. Srikant, Y. Xu, "Hippocratic Databases", http://www.almaden.ibm.com/cs/people/srikant/papers/vldb02.pdf, IBM Almaden Research Center, 2002
10. IBM Tivoli Privacy Manager, "Privacy manager main web page", http://www-306.ibm.com/software/tivoli/products/privacy-mgr-e-bus/, 2005
11. IBM Tivoli Privacy Manager, "online technical documentation", http://publib.boulder.ibm.com/tividd/td/PrivacyManagerfore-business1.1.html, 2005
12. HP, "HP Select Federation - Product and Solution Overview", http://www.managementsoftware.hp.com/products/slctfed/, 2005
13. ePok, "identity management solution - Trusted Data Exchange Server", http://www.epokinc.com/, 2005
14. HP, "HP OpenView SelectAccess - Overview and Features", http://www.openview.hp.com/products/select, 2005
15. IBM, "IBM Tivoli Storage Manager for Data Retention", 2004
16. C. Bettini, S. Jajodia, X. Sean Wang, D. Wijesekera, "Obligation Monitoring in Policy Management", 2002
17. N. Damianou, N. Dulay, E. Lupu, M. Sloman, "The Ponder Policy Specification Language", 2001
18. M. Casassa Mont, S. Pearson, P. Bramhall, "Towards Accountable Management of Privacy and Identity Information", ESORICS 2003, 2003
19. M. Casassa Mont, R. Thyne, Pete Brmhall, "Privacy Enforcement with HP Select Access for Regulatory Compliance", HPL-2005-10, 2005
20. M, Casassa Mont, "Dealing with Privacy Obligations: Important Aspects and Technical Approaches", TrustBus 2004, 2004
21. M. Casassa Mont, "Dealing with Privacy Obligations in Enterprises", ISSE 2004, 2004
22. PRIME, "Privacy and Identity Management for Europe, European RTD Integrated Project under the FP6/IST Programme", http://www.prime-project.eu.org/, 2004
23. HP, "HP OpenView Select Identity – Overview and Features", http://www.openview.hp.com/products/slctid/index.html, 2005
24 M. Hilty, D. Basin, A. Pretschner, "On Obligations", Information Security, ETH Zurich, Switzerland, 10th ESORICS, 2005