Privacy Enforcement with HP Select Access for Regulatory Compliance

Marco Casassa Mont, Robert Thyne, Pete Bramhall
Trusted Systems Laboratory
HP Laboratories Bristol
HPL-2005-10
January 21, 2005*

E-mail: marco.casassa-mont@hp.com, robert.thyne@hp.com, pete.bramhall@hp.com

privacy, privacy enforcement, access control, privacy-aware access control, regulatory compliance, data governance, policy management

Regulatory compliance is a hot topic for enterprises. The increasing number of laws, including SOX, GLB, HIPAA and various governmental directives on data protection require enterprises to put in place complex processes to comply with related policies. Among other things, this involves the analysis, modeling, deployment, enforcement and audit of these policies. Privacy management is a core aspect of regulatory compliance. Enterprises store large amounts of personal (confidential) data about their employees, customers and partners. Failure to comply with privacy policies can have serious consequences for their reputation and brand and have negative legal and financial impacts. Most of the solutions in this space address auditing and reporting issues. However, being able to enforce privacy policies on personal data by means of flexible, integrated and adaptive solutions is also very important: at the moment this aspect is still a green field, open to research. This paper describes work done at HP Labs to address this problem and develop a privacy-aware access control system to enforce privacy policies on personal data. A working prototype and a related demonstrator have been implemented, as a proof of concept, by leveraging the HP Select Access product: privacy policies are authored with an extended version of the HP Select Access Policy Builder (via standard plug-ins); related decisions are made by an extended version of the HP Select Access Validator (via standard plug-ins). A brand new "Data Enforcer" has been implemented and integrated with HP Select Access to enforce fine-grained privacy decisions on personal data stored in data repositories. The management of traditional access control policies is integrated with the management of privacy policies. This brings simplicity and rationalises the required set of management and enforcement tools.

* Internal Accession Date Only
Approved for External Publication

Privacy Enforcement with HP Select Access for Regulatory Compliance

Marco Casassa Mont, Robert Thyne, Pete Bramhall

Hewlett-Packard Laboratories, Filton Road, Stoke Gifford
BS34 8QZ, Bristol, UK
marco.casassa-mont@hp.com, robert.thyne@hp,com,
pete.bramhall@hp.com

Abstract. Regulatory compliance is a hot topic for enterprises. The increasing number of laws, including SOX, GLB, HIPAA and various governmental directives on data protection require enterprises to put in place complex processes to comply with related policies. Among other things, this involves the analysis, modeling, deployment, enforcement and audit of these policies. Privacy management is a core aspect of regulatory compliance. Enterprises store large amounts of personal (confidential) data about their employees, customers and partners. Failure to comply with privacy policies can have serious consequences for their reputation and brand and have negative legal and financial impacts. Most of the solutions in this space address auditing and reporting issues. However, being able to enforce privacy policies on personal data by means of flexible, integrated and adaptive solutions is also very important: at the moment this aspect is still a green field, open to research. This paper describes work done at HP Labs to address this problem and develop a privacy-aware access control system to enforce privacy policies on personal data. A working prototype and a related demonstrator have been implemented, as a proof of concept, by leveraging the HP Select Access product: privacy policies are authored with an extended version of the HP Select Access Policy Builder (via standard plug-ins); related decisions are made by an extended version of the HP Select Access Validator (via standard plug-ins). A brand new "Data Enforcer" has been implemented and integrated with HP Select Access to enforce fine-grained privacy decisions on personal data stored in data repositories. The management of traditional access control policies is integrated with the management of privacy policies. This brings simplicity and rationalises the required set of management and enforcement tools.

1 Introduction

Regulatory compliance is a hot topic for enterprises. The list of regulations to be compliant with is fast and growing and cuts across a variety of areas, including:

- **Financial compliance**: Basel III, GLB, SOX, etc.;
- **Health Records**: HIPPA, etc.;
- **IT compliance**: Calif. SB 1386, SEC 17A-B, etc.;
- **Manufacturing Compliance**: DOT mandates, FDA 21 CFR Part 11, MSDS, OSHA Mandates, etc.;
- **Antiterrorism**: Foreign Corrupt Practices Act, Homeland Security Act, Patriot Act, etc.

The increasing number of laws and legislation and additional governmental directives on data protection require enterprises to put in place complex processes to comply with related policies. Among other aspects, this involves the analysis, modeling, deployment, enforcement and audit of these policies.

Privacy is a core aspect of regulatory compliance. Enterprises store large amounts of personal (confidential) data about their employees, customers and partners. Failure to comply with privacy policies can have serious consequences for the reputation and brand of organizations and have negative legal and financial impacts. Enterprises are heavily investing in identity management solutions: in this context, being able to enforce privacy policies on personal data via systemic and verifiable manner is becoming a core requirement.

This paper describes basic privacy concepts and aspects related the problem of dealing with privacy in enterprises: in this context, it discusses the relevance of privacy management for data governance.

It focuses on the enforcement of privacy policies: it describes work done at HP Labs to develop a privacy-aware access control model and a related system to enforce privacy policies on personal data stored in data repositories. The goal is to demonstrate how privacy policies, dictating constraints and conditions on personal data, can be integrated with enterprise access control policies by leveraging a common authoring, deployment and enforcement framework. A technical solution is described along with the underlying mechanisms to rationalize and simplify the management and enforcement of these policies.

As a proof of concept, a working prototype has been implemented by leveraging and extending the HP Select Access product: privacy policies are authored with an extended version of the HP Select Access Policy Builder (via standard plug-ins); related decisions are made by an extended version of the HP Select Access Validator (via standard plug-ins). A brand new “Data Enforcer” has been implemented and integrated with HP Select Access to enforce fine-grained privacy decisions on personal data stored in data repositories. The core aspects of this prototype are presented along with a related demonstrator in a healthcare scenario.

2 Privacy Management in Enterprises

Enterprises store, manage and process large amounts of personal and confidential data related to their customers, employees and partners. These data have to be managed in a privacy-compliant way. Privacy is a complex topic as shown by figure 1. Different aspects need to be kept into account:

- People, their personal data, their preferences and requirements;
- Legislations and laws;
- Social and business aspects;
- Technologies.

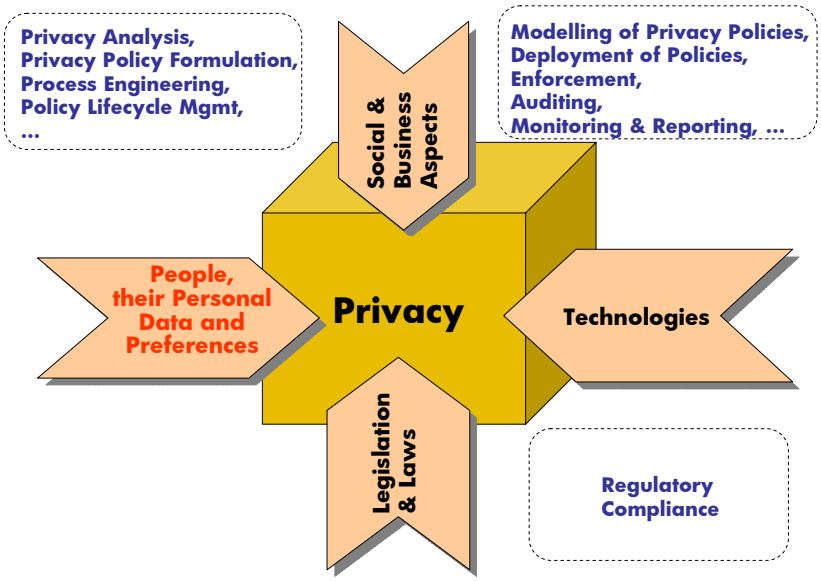


Fig. 1. Privacy: a Complex Topic

At the very base there are people, their personal data, their requirements and expectations. There are conflicts on requirements and interests when considering personal, social and business perspectives. On one hand, personal data has to be disclosed by people (data subjects) to enable enterprises' business processes, interactions and transactions. On the other hand, personal data should be accessed and used only for the purposes for which it has been disclosed and with the consent of the data subjects. Enterprises increasingly recognise that dealing correctly and honestly with privacy matters can have a beneficial return in terms of branding, trust, customers' satisfaction and business opportunities. Figure 2 shows various aspects that enterprises need to keep into account (including customers' expectations, laws and guidelines) and the effects of their behaviours.

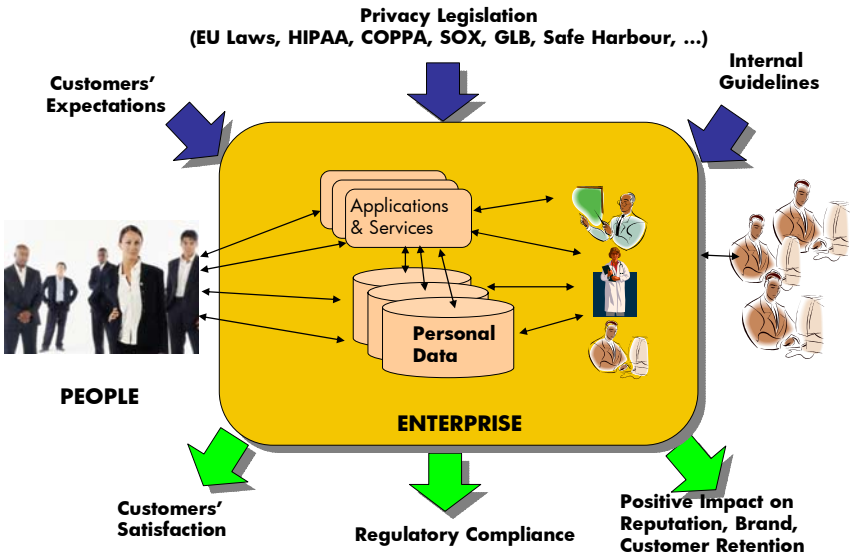


Fig. 2. Privacy: Impact on Enterprises and Opportunities

As anticipated in the introduction of this paper, enterprises must be compliant with privacy laws and related requirements. A lot of work has been done in terms of privacy legislation, often driven by local or geographical needs. This includes European Community data protection privacy laws, various US privacy laws and more specific national privacy initiatives [1]. Guidelines are also available on the protection of privacy and flows of personal data, including OECD guidelines [2] that describe concepts such as collection limitation, data quality and purpose specification principles and online privacy policies [3]. Large enterprises that are geographically distributed across different nations might need to comply with different privacy laws.

Figure 3 shows a simplified version of the overall data governance and compliance process that enterprises might need to go through when handling data. At the "centre" of this process there are data, people and their various roles, systems, applications and services that use these data.

Policies are developed and modelled to describe how data has to be stored, accessed, manipulated, processed, managed, transferred and eventually deleted. Inventories of data are created and subsequently kept up-to-date: gap and risk analysis tasks are performed to check for the suitability of processes, frameworks and behaviours against these policies and identify risks and gaps. Eventually policies are deployed and enforced. Their enforcement is monitored and audited: reports are generated to spot anomalies and violations. All these phases are not linear and can involve various refinement loops.

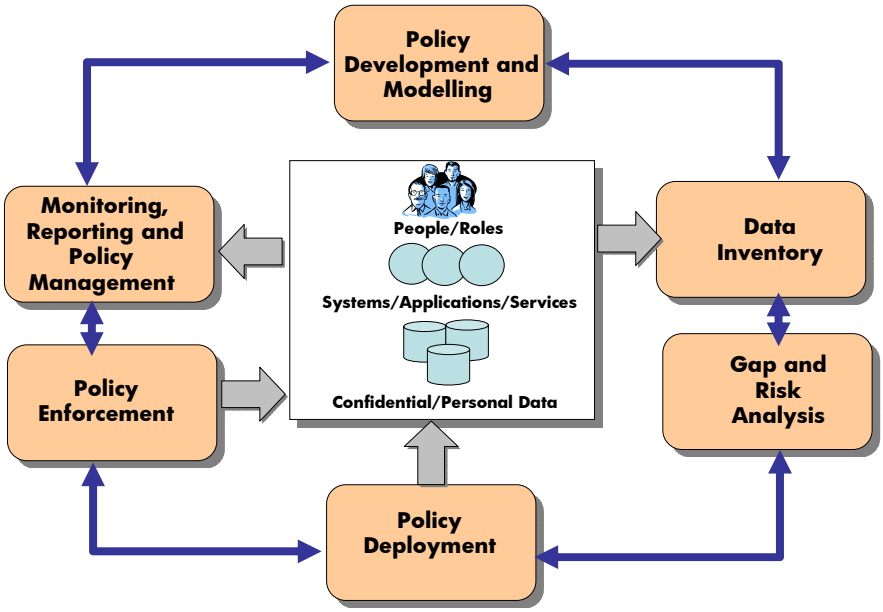


Fig. 3. Data Governance and Policy Life-cycle Management

Privacy management has an impact on this data governance process. Privacy requirements need to be analysed and privacy policies formulated. Privacy policies are a suitable tool to represent and describe privacy laws, guidelines and privacy statements. At the very base they can express rights, permissions and obligations. Privacy policies need to be deployed, enforced, monitored, audited and reports generated about their overall compliance.

Emerging technologies are shaping how these privacy management aspects can be carried out by enterprises. In most cases, once privacy policies have been formulated, they are enforced with ad-hoc, vertical approaches or “embedded” into applications and systems. Most of the solutions available today only explicitly address the auditing and reporting aspects. Section 5 provides more details and insights about related work.

The explicit management and enforcement of privacy policies is still a green field and an important problem that is worth exploring and researching, especially in complex contexts such as enterprises: identity management solutions, currently used by enterprises to deal with personal data and profiles, need to be extended to include privacy management and enforcement capabilities.

3 Addressed Problem

The key problem addressed by this paper is the enforcement of privacy policies for personal data stored by enterprises. Closely related to this problem are the issues of modeling personal data, authoring and deploying privacy policies.
The enforcement of these privacy policies is not just a matter of traditional access control. Additional aspects need to be considered, such as managing the stated purposes for which personal data is collected, stored and used, evaluating the intent of access requestors against these purposes and taking in account the consent given by "data subjects".

Privacy policies might impose additional conditions and constraints (dictated by data subjects and legislation) on which personal data or any subset of these data can actually be accessed, given a specific context. This might involve the *manipulation*, *transformation* and *filtering* of these data before being accessed by a requestor.

How are all these aspects to be taken into account when accessing personal data? Entities, applications and services must be able to retrieve personal data by searching data repositories, in a way that is consistent with privacy policies and compliant with regulations. This process must be flexible and adaptive to changes of business needs, privacy policies and data subjects' preferences.

What does a privacy policy enforcement framework look like? How can attempts to access confidential data be intercepted and related privacy policies enforced? How transformation or filtering of accessed data can be performed in an efficient way?

Section 4 describes important related issues and requirements. Section 5 presents related work and section 6 introduces and describes our solution.

4 Issues and Requirements

This section describes issues and requirements related to the enforcement of privacy policies on personal data. It also introduces the terminology used in the remaining part of this paper.

4.1 Terminology

When dealing with privacy management for personal data, it is important to keep into account important aspects such as the data purposes, consent and intent. This terminology is introduced along with a description of the principal entities involved in the disclosure, management and processing of personal data: data subjects, data requestors and enterprise privacy administrators – see figure 4.

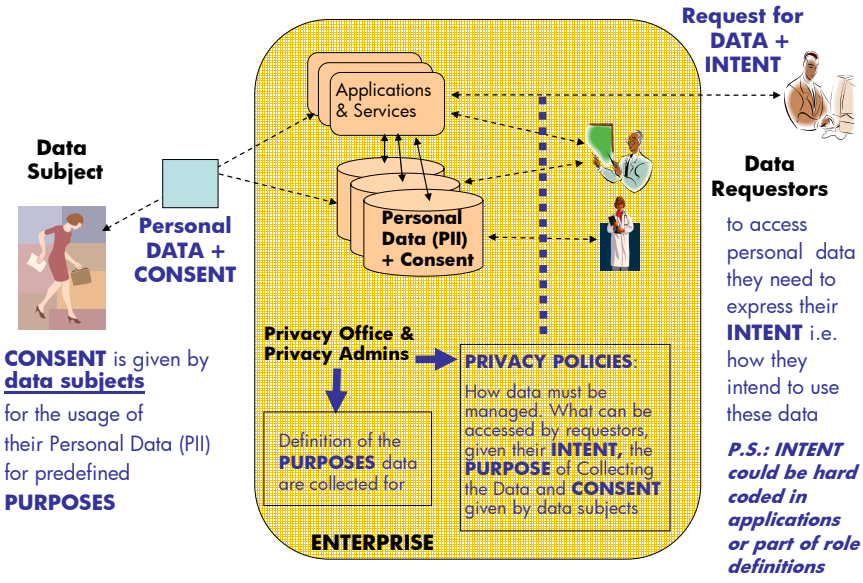


Fig. 4. Terminology: Consent, Intent and Data Purpose

The ***purpose*** of personal data is the "reason" for which data is collected. Enterprises should clearly state the purpose of collecting personal data from people.

Consent is given by *data subjects* to enterprises (organizations) to use their personal data for predefined purposes. Consent can include constraints to be satisfied, obligations and requirements on how data should be managed, disclosed and retained/deleted by the enterprise.

The ***intent*** is expressed by *data requestors* when trying to access personal data. In other words it is their "reason" for accessing these data. Requestors can be enterprise entities or external entities (such as partners, other people, etc.)

At the very base a ***privacy policy*** describes how data has to be managed and accessed: in particular which data can be accessed by requestors, given their *intent*, the data *purpose* and the *consent* given by data subjects.

Figure 5 shows how all these aspects come together when dealing with ***privacy policy enforcement***: given personal data (provided by a data subject) and related consent, given a request to access these data by a requestor, the "*privacy policy enforcement point"* has to ensure that: the requestor's intent is consistent with the specified data purposes; data subject's consent is kept into account, along with any preference and constraint; only the "legitimate" parts of data are accessed by the requestor, according to the privacy policies.

At the very base, privacy policies describe constraints and conditions. During the enforcement of these policies some of these constraints might or might not satisfied: this dictates which data (if any) can actually be accessed by a requestor and the kinds

of data “transformation” that are required to be enforced, including data filtering, data encryption, data transformation, statistical modification, etc.

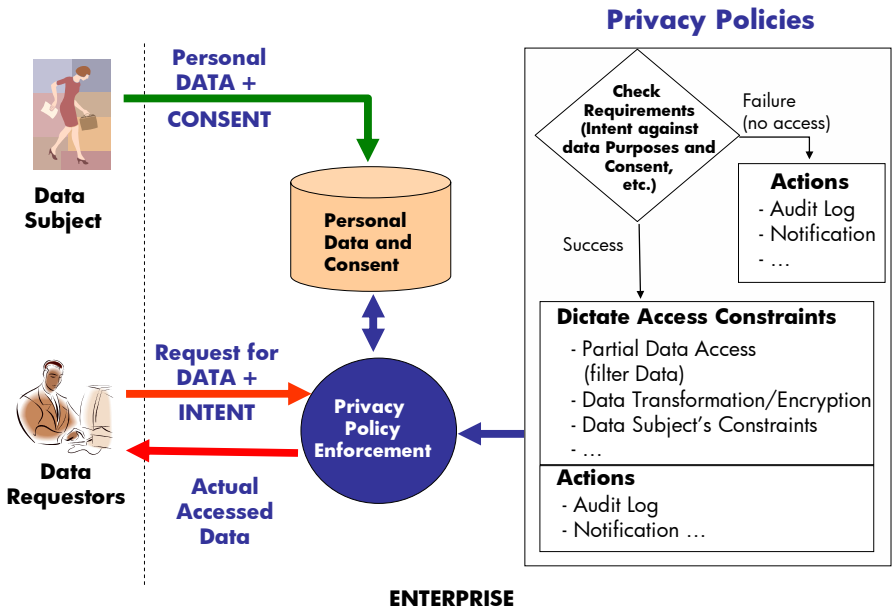


Fig. 5. Terminology: Privacy Policy and Privacy Policy Enforcement

The logging and auditing of all these activities are an important part of this process.

4.2 Important Issues

Current privacy laws and legislations have in common a few core principles and aspects that need to be kept into account by enterprises when enforcing privacy policies on personal data:

- **Purpose specification;**
- **Consent;**
- **Limited collection of data;**
- **Limited use of data;**
- **Limited disclosure of data;**
- **Limited retention of data.**

Privacy management systems need to be compliant with all these principles: they have implications on how personal data can be accessed and used.

Specifically, *privacy enforcement* on personal data has access control implications: the data purpose and the consent given by data subjects impose limitations on how data is accessed. Similarly, the limitations on data usage, disclosure and retention dictate conditions and constraints that need to be satisfied before accessing personal data - see figure 6.

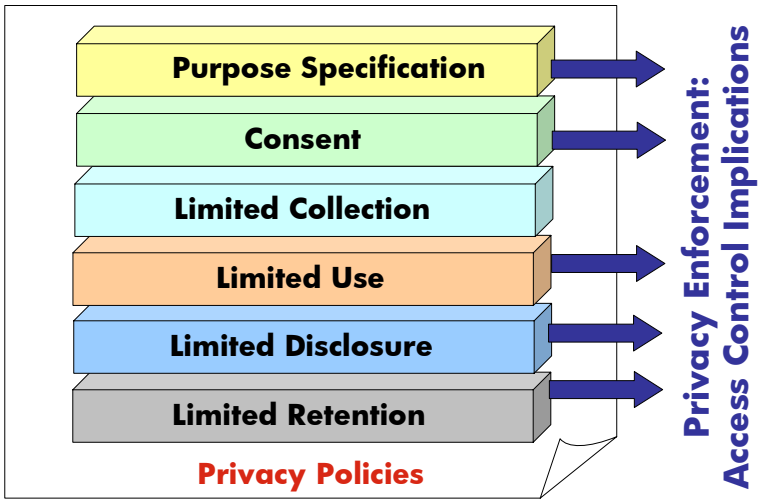


Fig. 6. Privacy Enforcement and Access Control implications

We argue that traditional access control systems are necessary but not sufficient to enforce privacy policies on personal data.

Figure 7 compares a traditional access control system (7a) against a privacy-aware access control system (7b).

Traditional access control systems (7a) are mainly based on “access control lists” and enforcement mechanisms that keep into account the identities of data requestors, their rights and permissions and the types of actions that are allowed/disallowed on the involved resources (data resources).

These systems do not keep into account additional aspects relevant to privacy enforcement (7b): the stated data purpose and data subjects’ consent - i.e. properties usually associated to collected data - the intent of the requestors and any additional enterprise or customized data subjects’ constraints.

An important issue to be addressed is how to build “privacy extensions” of traditional access control systems to move towards ***Privacy-Aware Access Control*** systems and enforce privacy policies.

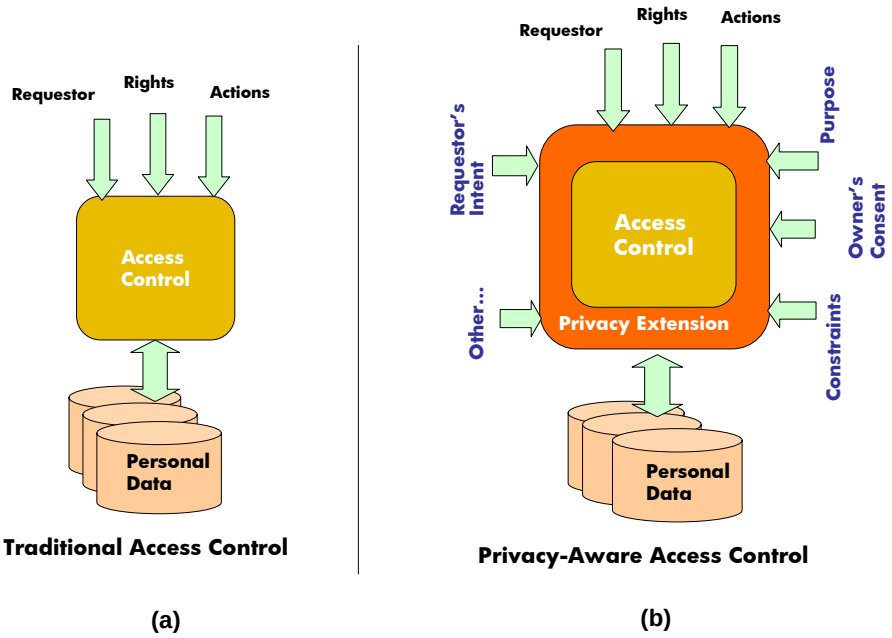


Fig. 7. Traditional Access Control (a) vs. Privacy-Aware Access Control (b)

4.3 Requirements

To address the above issues and move towards privacy-aware access control systems, it is important to satisfy the following core requirements:

- **Modeling of Personal Data**: in order to explicitly define privacy policies on personal data, a detailed, fine grained model of these data is required. This model must describe which types of data are subject to privacy constraints, their schema, semantic and properties;
- **Explicit definition, authoring and lifecycle management of privacy policies**: it must be possible to explicitly author, modify and manage privacy policies in a fine grained way (by including all its constraints, conditions, actions, etc.). Authoring and management tools are required to support the lifecycle management of these policies;
- **Extensible privacy policies**: privacy policies needs to be described in a language and format that is easy to change, extend and adapt to new business, legislative and security needs;
- **Explicit deployment and enforcement of privacy policies**: it must be possible to explicitly deploy and enforce privacy policies. Framework support-

ing these aspects must guarantee the separation of privacy policies from business logics (within applications and services) in order to allow rapid and adaptive redeployments of these policies in case of changes;

- **Support for auditing**: all the phases involving the management of privacy policies (authoring, deployment, decision-making, enforcement, etc.) need to be logged and audited. Auditing tools need to be provided for regulatory compliance;
- **Integration with traditional access control system**: system dealing with privacy policies should be integrated with more traditional system handling access control policies. This is important to guarantee the rationalization and simplification of the overall management and enforcement tasks;
- **Simplicity**: privacy-aware access control systems need to be simple to use, both for administrators and data subjects, especially when authoring and defining privacy policies on personal data. Simple and intuitive graphical UIs are required to underpin these tasks.

5 Related Work

The problem of enforcing privacy policies on personal data can be addressed and solved in different ways, each of them with pros and cons.

A common approach to address this problem consists of hardcoding privacy policies within applications and services or building ad hoc solutions that work in specific contexts and for specific purposes. 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 problem, a model of the relevant personal data is required. Privacy policies dictating how these data must be accessed 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 the space of privacy management and enforcement for enterprises is described in [4,5,6,7]. Enterprise Privacy Architecture is introduced and described in [7], encompassing a policy management system, a privacy enforcement system and an audit console. This concept is framed in the context of privacy rules defined for authorization purposes. This approach is further refined and described in the Enterprise Privacy Authorization Language (EPAL) specification [8].

The above work makes important advancements in exploring and addressing the problem of privacy management in enterprises and the explicit representation of privacy policies: it focuses on the authorization and access control perspectives and provides general guidelines.

More explicit solutions dealing with the management and enforcement of privacy policies only operate in well defined contexts and by using vertical technology.

In this direction, important work on privacy enforcement on personal data has been done by IBM with their research on Hippocratic databases [9], i.e. databases that

include mechanisms for preserving the privacy of the data they manage. Privacy metadata is associated to personal data, within a data repository along with mechanisms to enforce privacy. 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 might need to span across a broad variety of data repositories and legacy systems, not only RDBMS databases: additional data repositories might include LDAP directories, meta and 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 (it requires the parallel usage of Tivoli Access Manager) and it is vertically-based on other IBM products and solutions.

Products such as HP Select Federation [12] and ePok [13] focus on single-sign-on aspects: they manage personal data in federated environments. The enforcement of simple privacy policies is provided in federated contexts, when personal data has to be disclosed by an organization (or an identity provider) to other parties in a way that is compliant to constraints or "contracts" agreed with data subjects. These products and solutions provide pragmatic approaches specifically to address identity federation issues: they can be deployed within enterprises in addition to more traditional access control and authorization solutions.

Other recent products, including Synomos [14], InTrust [15], NetForensics [16], SenSage [17], provides tools to explicitly model, audit and report compliance to policies, including privacy policies. They require administrators to use additional set of management tools to deal with access control issues. They do not address or only partially address the problem of enforcing privacy policies.

Our work specifically addresses the problem of enforcing privacy policies on personal data stored in a broad variety of data repositories within enterprises. Personal data can be accessed by different types of requestors, including people, applications and services. It includes the 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 and not hardcoded in applications and services. We want to avoid duplications of efforts by providing a *single, integrated framework* for authoring, administering and enforcing both traditional *access control* and *privacy policies*.

We believe that this work can provide an important competitive advantage for HP customers if current HP investments in the identity management space are leveraged and extended to enforce privacy policies, in particular HP Select Access [18,19,20].

6 Our Solution

This section provides details about our model to enforce privacy policies on personal data and describes how this model can be implemented by extending the HP Select Access product [18,19,20].

6.1 Model

Our model for a *privacy-aware access control system* extends the traditional access control model (based on users/groups, their credentials and rights, access control lists and policies) by explicitly dealing with data purposes, checking - at the enforcement time - the intent of requestors against these purposes, dealing with data subjects' consent and enforcing additional access conditions and constraints on personal data. Figure 8 shows the main aspects of this model:

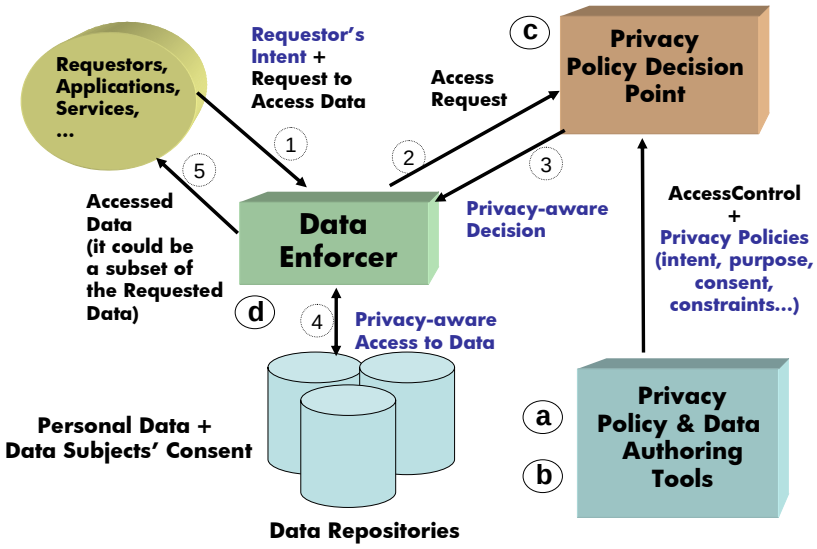


Fig. 8: Model of our Privacy-Aware Access Control System

Our model of a privacy-aware access control system consists of:

a) **A mechanism for the explicit *modelling of personal data* subject to privacy policies**: this mechanism provides a description of these 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 describing 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 (such as who should 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) to make decisions based on these policies;

d) **A run-time mechanism –referred with the "*data enforcer*" term - 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 and filtering of part of the requested data.

Our model leverages traditional/standard access control models, based on Policy Authoring Points (PAPs), Policy Decision Points (PDPs) and Policy Enforcement Points (PEPs). We extend this model with access control capabilities on personal data driven by privacy policies.

At "run-time" attempts to access personal data are intercepted and managed in the following way:

1. A request from a data requestor to access personal data stored in a data repository 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:
    - ***No***: access to data is denied;
    - ***No & conditions***: access to data is denied. Some conditions are sent back to the requestors. The satisfaction of these conditions (for example to pass the intent or to authenticate) could change the outcome of the decision;
    - ***Yes***: access to data is granted;
    - ***Yes & conditions***: access to (part of the) data is allowed, under the satisfaction of the attached conditions. Among other things, these conditions might require data transformations and manipulations.

4. The data enforcer enforces this decision. In particular, if the decision is *"Yes & conditions"* the data enforcer might have to manipulate and transform the accessed personal data, 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.

Two examples are presented to provide additional details about this model. In a first very simple example, shown in figure 9, personal medical data is stored by an enterprise (for example a healthcare service provider) in a data repository - for example a RDBMS database. This database contains patients' data stored as records in a relational table, having three fields: *Name*, *Condition*, *Diagnosis*.

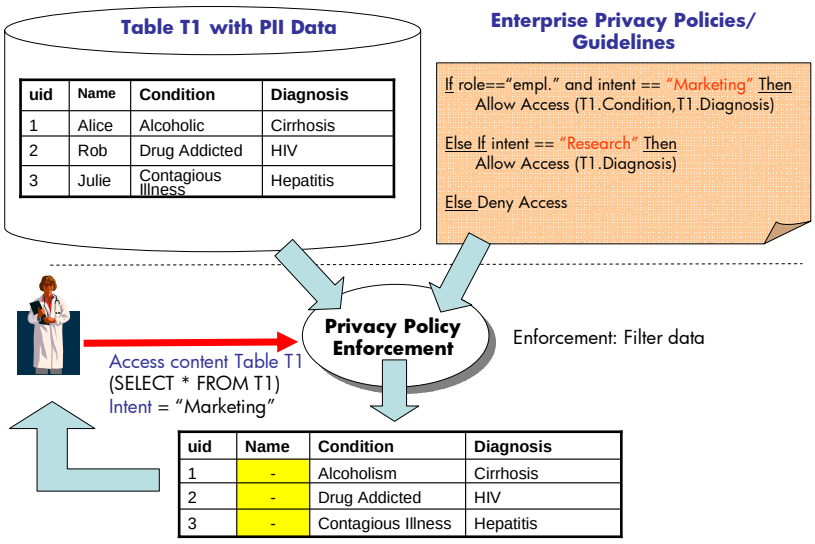


Fig. 9: Example: Privacy-aware Access Control – Purpose and Intent Management

The enterprise defines a few privacy policies that can be summarized as follow:

- Patients' medical data can only be collected and accessed for "Research" and "Marketing" purposes;
- Medical data can be accessed for "Marketing" purposes only by enterprise's employees. In this case, only the "Condition" and "Diagnosis" fields of each customer's record can be accessed. The content of the "Name" field must not be visible;
- Medical data can be accessed for "Research" purposes, but only the "Diagnosis" field of each customer's record is visible. All the other fields are not visible.

- Any other attempt to access to data, for purposes beyond research and marketing must be denied.

A data requestor - in this example an enterprise' employee - tries to access the entire content of the database, with "Marketing" intent. Because of the privacy policies described above, the *privacy policy enforcement point* will intercept this request and "transform" data in a way that the actual data returned to the requestor is compliant to these policies. In the example, all patients' records are returned but the content of the "Name" field is removed (i.e. it contains the "null" value).

A slightly more complex example, shown in figure 10, is based on the same scenario, with the same kind of personal data. In this example the *consent* given by patients (data subjects) is stored in the database, along with their personal data.

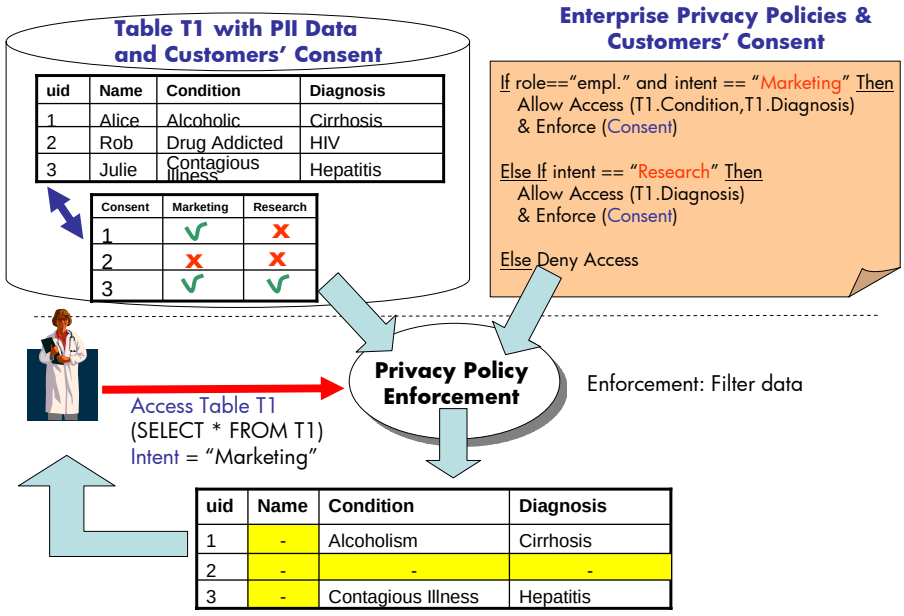


Fig. 10: Example: Privacy-aware Access Control – Consent, Purpose and Intent Management

In this example, consent is given at a "macro" level i.e. to allow/disallow access of the entire personal data of a data subject, for marketing or research purposes (of course more fine-grained consent can be handled by the system). The patient "Rob", for example, has given no consent to use his data for any purpose.

The privacy policies are extended to include the enforcement of *consent* given by patients (data subjects).

In this example, the same data requestor, trying to access the content of the database with "Marketing" intent, will actually retrieve a different "portion" of the data (if compared to the previous example), as this time consent's constraints are kept into

account. For example, “Rob’s record” will be completely “removed” i.e. no information about his data will be returned to the requestor, according to his preferences.

These two examples have been deliberately kept very simple to show the main concepts.

All these queries could have been hardcoded in applications and services: this would work in case of static environments that are not subject to changes. However, in the real world the situation is much more complex, especially for medium-large enterprises that need to run thousands of applications and services to underpin their businesses, have thousands/millions of customers and need to cope with ever changing business and legal needs.

In real-world scenarios, data repositories storing personal data could be heterogeneous, including relational databases, LDAP directories, meta and virtual directories and legacy storage systems. Within the same data repository, personal data could be stored in different “tables”, and be accessible either directly or via different “views”.

Privacy policies might be much more complex that the one shown in the examples. They might include data retention requests and specific customer’s or enterprise constraints such as requests for notifications and requests for authorizations. They might require different types of data manipulation and transformation, in addition to filtering, depending on the content and the specific types of requests. This might include data encryption, statistical modification of data, etc.

The remaining part of this section describes a practical implementation of this model by using and extending the HP Select Access product. A section will follow discussing current results, compliance to requirements and next steps.

6.2 Technical Approach

The key drivers for the implementation of our privacy-aware access control model have been to: (1) leverage as much as possible the investments made by HP customers on HP access control and identity management solutions to handle also privacy aspects; (2) differentiate HP identity management solutions by adding simple and integrated privacy enforcement capabilities. The goal is to avoid duplications of customers’ investments and efforts: we aim at providing an integrated and simple approach to author, deploy and enforce privacy and access control policies. We want to avoid incurring in the same problems afflicting existing solutions in this space (see the “related work” section).

The HP Select Access (HP SA) framework [18,19,20] can provide the infrastructural components needed to author, deploy and enforce privacy policies. We will briefly illustrate the basic HP Select Access functionalities and how they have been leveraged and extended to implement our model.

6.2.1 HP Select Access

HP Select Access is a leading-edge access control product [18,19,20]. Figure 11 shows its core components:

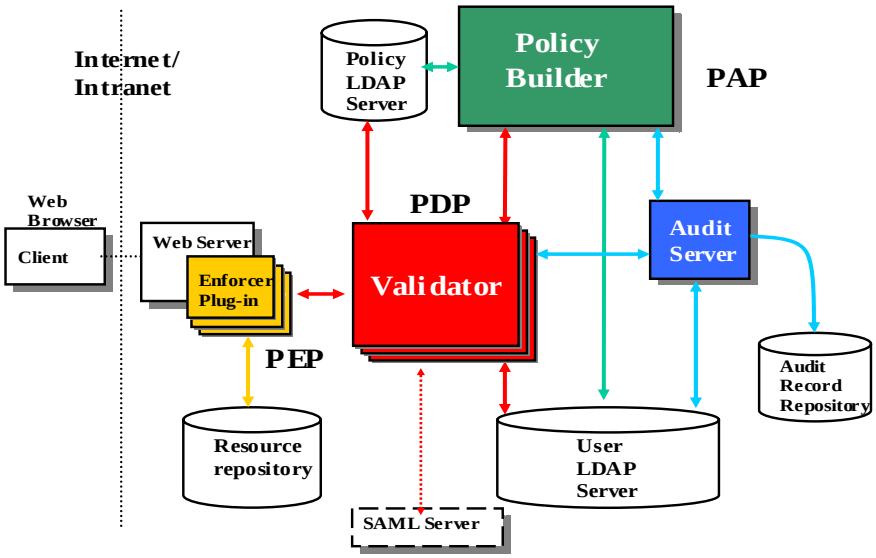


Fig. 11: HP Select Access – Architecture and Main Components

HP Select Access includes the following core components:

- **Policy Builder**: it is a graphical tool to author access control policies (PAP) on resources managed by the system. In the current product these resources are basically web resources (web files, HTML pages, etc.);
- **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). In the current product this enforcer is mainly for web resources;
- **Audit Server**: it stores logged information in a tamper evident storage. Log records can be sent to this component from all the other HP Select Access components.

An important characteristic of HP Select Access, that differentiates this product from others, is the Policy Builder i.e. the graphical tool to author and manage access control policies on resources, at different levels of granularity. Figure 12 shows the Policy Builder along with an example of managed web resources and “user” information:

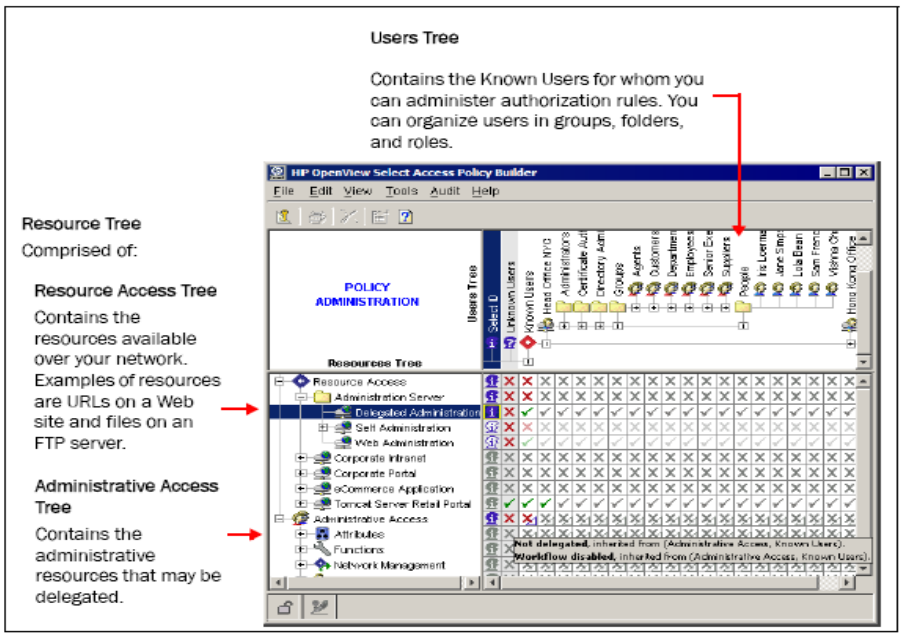


Fig. 12: HP Select Access – Policy Builder

The HP Policy Builder allows administrators to define access control rights (allow/deny access) on administered resources for given enterprise users. In the current product these resources are mainly web resources, accessible via web servers. This can be achieved via an intuitive matrix-based graphical UI.

Hierarchies of resources and users (via groups and sub-groups) can be defined. Administrator rights can be delegated.

In addition to this, an administrator can define fine-grained access control constraints and conditions to grant/deny access to resources. This is done via a "Rule Editor".

Figure 13 shows an example of a simple rule (policy) authored with the Rule Editor. In this figure the access control rule/policy requires the user to provide stronger credentials for authentications and checks for their validity.

A rule can be composed by assembling rule components (shown in the Rule Editor's toolbar): HP Select Access provides a built-in set of these components and APIs and classes to create new ones (via plug-ins).

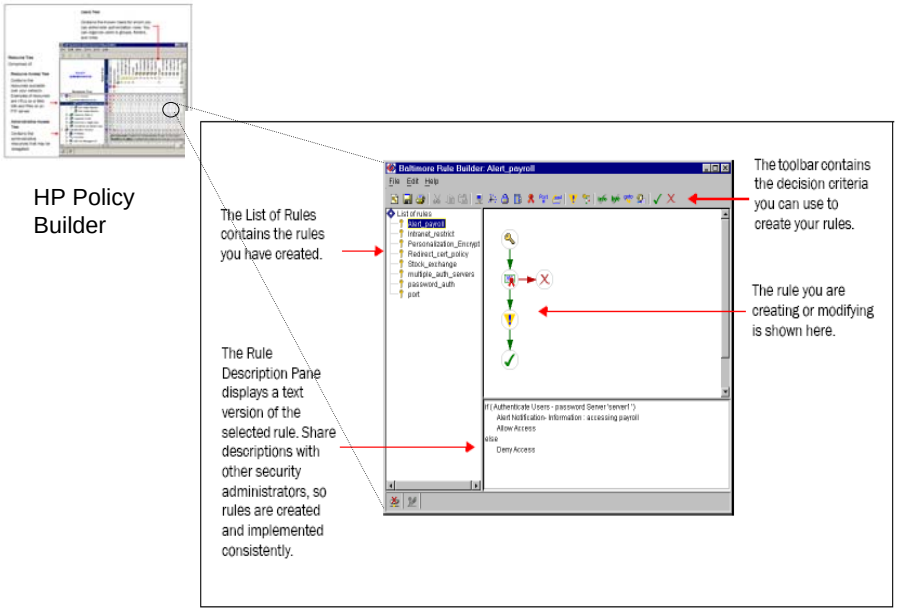


Fig. 13: HP Select Access – Rule Editor

6.2.2 Extending HP Select Access to Enforce Privacy Policies

As anticipated in section 6.1, our privacy-aware access control model requires the implementation of the following four mechanisms:

a) A mechanism for the explicit modelling of personal data, subject to privacy policies;
b) An integrated mechanism for authoring privacy policies along with traditional access control policies;
c) An integrated authorization framework for deploying both access control and privacy-based policies and making access decisions based on them;
d) A run-time mechanism - based on a "data enforcer" - for intercepting attempts to access personal data and enforcing decisions based on privacy policies and contextual information.

HP Select Access provides the basic functionalities that can be leveraged to enforce privacy policies on personal data: its policy authoring, policy deployment and policy enforcement capabilities. However, the current version of the HP Select Access product has to be extended to implement the following required functionalities:

- **Modeling of personal data**: the current version of HP Select Access only deals with web resources, not data resources;
- **Possibility to express privacy constraints**: the current version of HP Select Access only deals with access control constraints;
- **Enforcement of decisions based on privacy policies**: the current version of the Web Enforcer enforces decisions on web resources. These resources are considered as "black boxes" and accesses to them are either granted or denied on the whole resource. This enforcer is not appropriate to enforce privacy constraints on data as these constraints are fine-grained. Data cannot be considered as a "black box", as manipulation of data components might be required, for example to filter out or transform part of these data.

New functionalities have been added to HP Select Access (HP SA) to deal with the above aspects. Specifically:

1. **The HP SA Policy Builder has been extended to represent "data resources" in addition to traditional IT resources (such as web resources)**. Although the concept of "data resource" is not native of HP SA, it is easy to add descriptions of data resources to the system in a way they can be manipulated by the Policy Builder and Validator;
2. **The HP SA Policy Builder has been extended to author privacy policies on "data resources" in addition to access control policies**: this has been done via the definition and implementation of a set of additional plug-ins, including the ones that check (at the enforcement time) the requestor's intent against the stated data storage purpose, take into account data subjects' consent & data retention policies and describe how the accessed personal data has to be filtered, obfuscated or manipulated, etc.;
3. **The HP SA Validator has been extended to make privacy-aware decisions**. This has been done via the definition and implementation of additional plug-ins correspondent to the ones used in the Policy Builder. An important side-effect of our work is that the enhanced-version of the Validator can now make *"Yes & constraints"* decisions, i.e. decisions where access to data is allowed subject to the satisfaction of further privacy constraints - such as filtering out/obfuscating or statistically transforming part of these data;
4. **A Data Enforcer has been built and added to the framework**: this is a new functionality added to HP Select Access. The data enforcer is in charge of enforcing privacy decisions made by the Validator. It intercepts incoming calls to data resources, interacts with the Validator, performs fine grained manipulation of data resources and deals with the interpretation and enforcement of additional constraints as defined by the privacy policies. The data enforcer sits nearby 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. Further R&D work has to be done in this space.

Figure 14 shows the high level architecture of the extended version of HP Select Access and highlights the correspondences with our privacy-aware access control model for privacy enforcement (shown in figure 8):

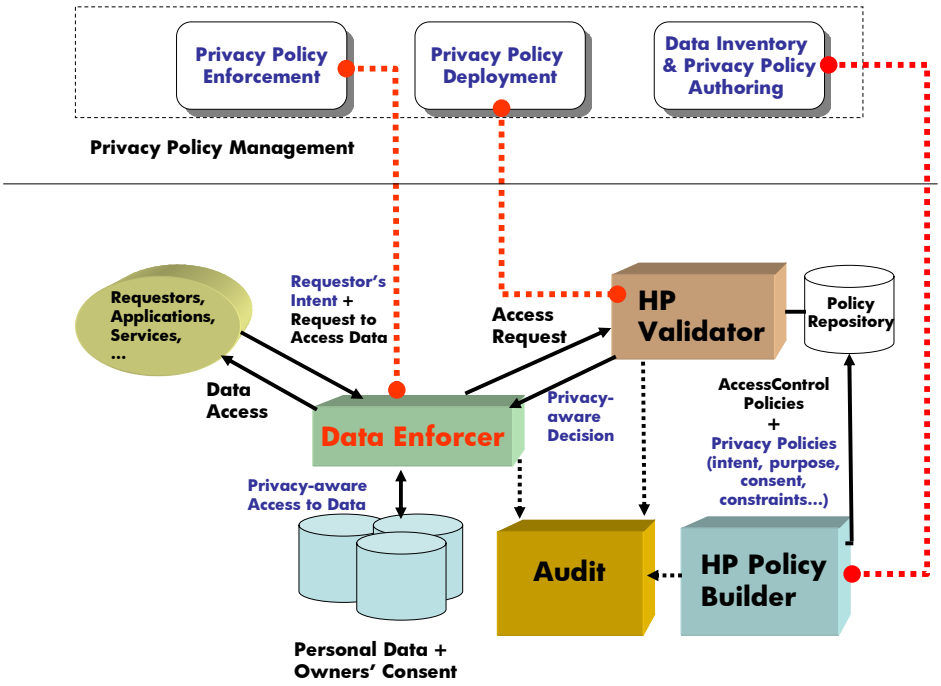


Fig. 14: Extended HP Select Access – a Privacy-Aware Access Control System

The *auditing* capability of HP Select Access is leveraged to log (among other things) requests to access data and related decisions made by the enforcement system.

The remaining part of this section provides more technical details about our extensions of HP Select Access to handle and enforce privacy policies.

6.2.2.1 Data Modelling: Extension of HP SA Policy Builder to Represent Data Resources

It is simple to extend the set of resources managed by HP Select Access (web resources) to also include "data resources". The *resource trees* shown by the HP Policy Builder are stored in a local LDAP repository (part of the HP Select Access framework). It is possible to inject additional information in this LDAP repository about the managed data resources. Tools can be built to explore the schemas of data repositories containing personal data and insert this information in the HP Select Access LDAP repository.

Figure 15 shows an example where “data resources” related to a RDBMS database (specifically a table in the database) are added to the system and are displayed by the Policy Builder, along with traditional web resources:

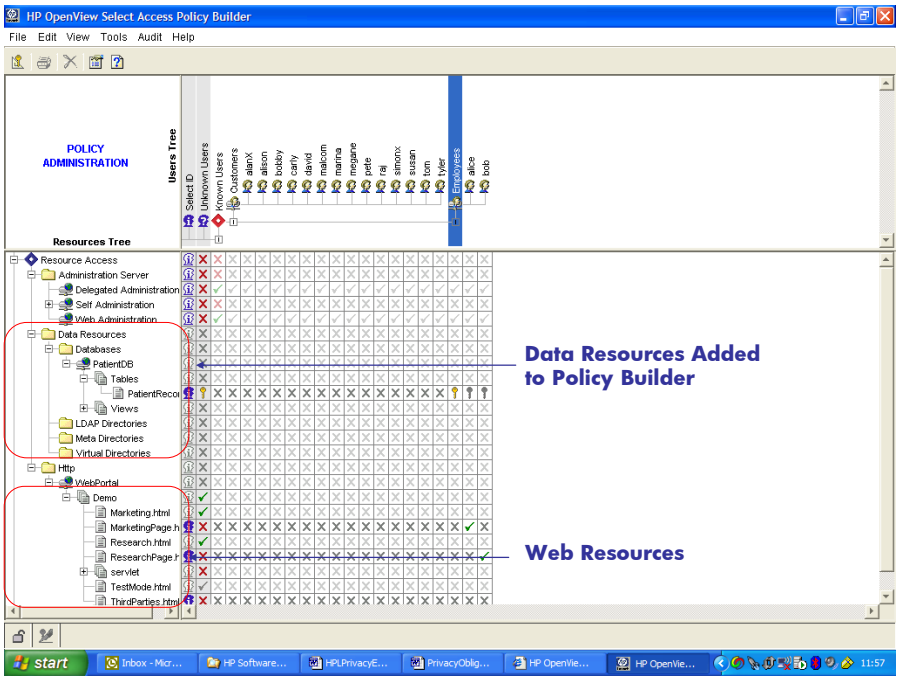


Fig. 15: HP Policy Builder with Data Resources

6.2.2.2 Authoring Privacy Policies in HP Policy Builder via Rule Editor

Privacy policies can be composed graphically by assembling plug-in components in the Rule Editor. These policies are then evaluated at runtime (by the HP Validator), to make decisions.

New plug-in components have been added to the Rule Editor to allow administrators to express different kinds of privacy constraints. A simple example of privacy policy authored by using our plug-ins is shown in figure 16. This policy defines the criteria to check the requestor’s intent against two data purposes: *marketing* and *research*. In case the requestor’s intent matches any of these purposes, the policy specifically defines how to: manipulate and filter the personal data to be returned to the requestor, enforce data subject’s consent and enforce data retention constraints.

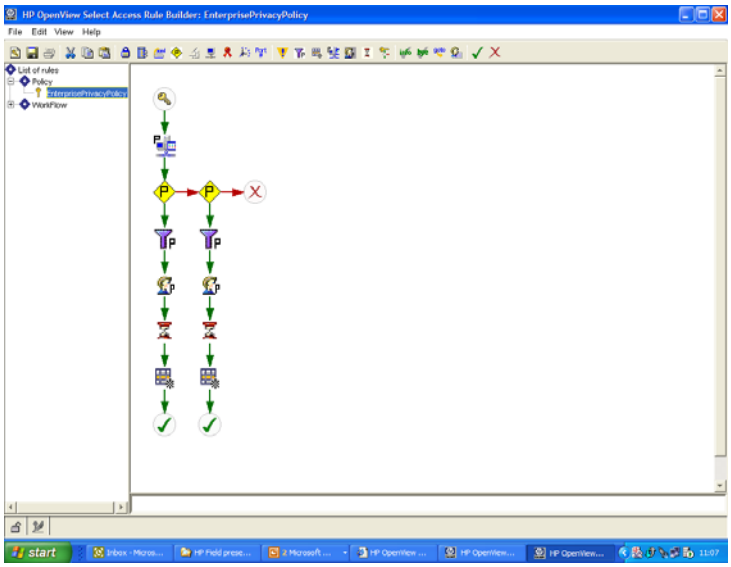


Fig. 16: HP Rule Editor – Example of Privacy Policy

The privacy policy plug-in components currently implemented provide the following capabilities:

- **Privacy decision plug-in**: it is a *privacy-related decision point*. At run time this plug-in allows the Validator to check the intent of the requestor against one of the allowed purposes for handling data. Depending on the result, alternative sets of other plug-ins are activated. Figure 17 shows an example of this plug-in;
- **Data transformation plug-in**: it describes in more details the schema/structure of the personal data to be accessed along with the types of transformations these data has to go through before being returned to the requestor. This includes: filtering out part of the data, encrypting data, doing statistical transformation of data, etc. This information is used at run-time by the data enforcer. Figure 18 shows an example of this plug-in;
- **Consent management plug-in**: it describes how to retrieve consent information (provided by the data subject) and how to link it to data subjects' personal data. This information is used at run-time by the data enforcer, to enforce data subjects' preferences and constraints. Figure 19 shows an example of this plug-in;
- **Data retention plug-in**: it describes how to retrieve specific data retention information and how to link it to data subject's personal data. Data retention constraints can be defined by data subjects, for example to impose deletion obligation, at predefined period of times. This information is used at run-time by the data enforcer to prevent access to data that is "expired". Figure 20 shows an example of this plug-in.

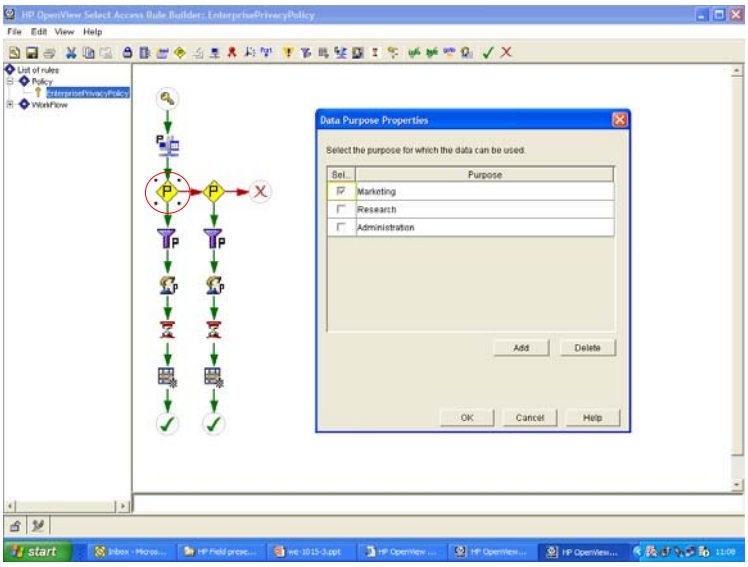


Fig. 17: HP Rule Editor – Privacy Decision plug-in

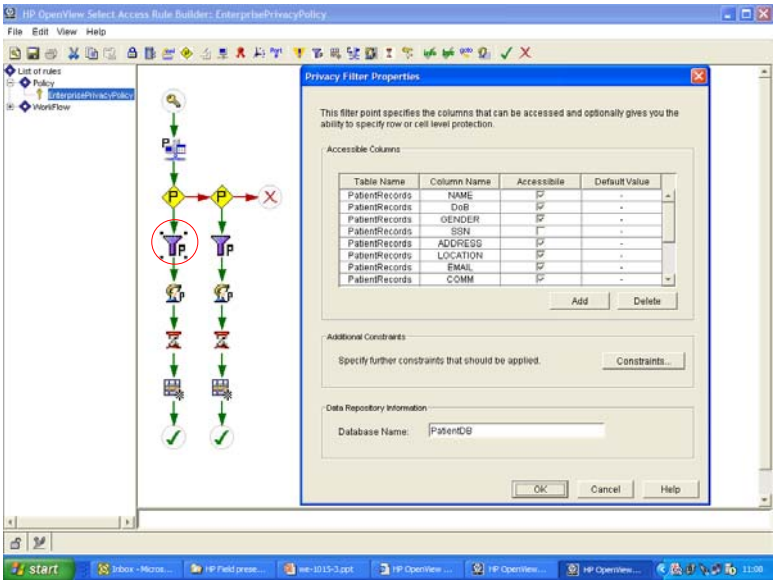


Fig. 18: HP Rule Editor – Data Transformation plug-in

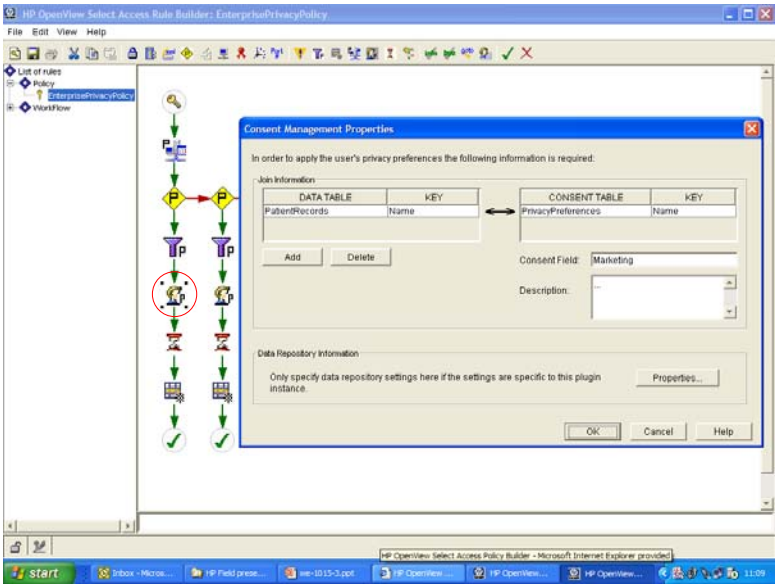


Fig. 19: HP Rule Editor – Consent Management plug-in

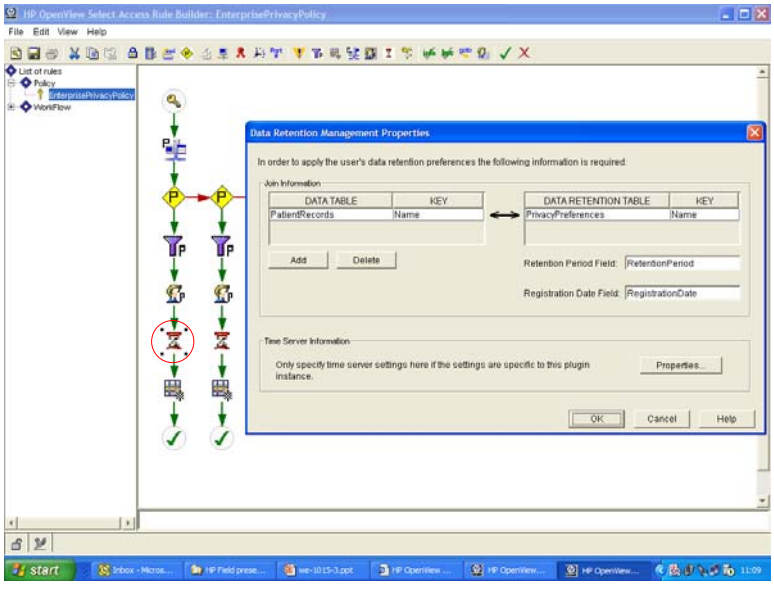


Fig. 20: HP Rule Editor – Data Retention plug-in

The graphical representation of privacy policies in the Policy Builder, along with all its plug-in components, is automatically mapped into an internal XML format.

Despite the fact that the examples described above only focus on privacy policies, it is important to notice that the same Policy Builder and Rule Editor will be used to author both traditional access control and privacy policies, in an integrated way.

The set of plug-in components described above has been implemented to demonstrate core privacy policy functionalities. Their actual definition can be modified or extended. New plug-in components can be added to the system depending on needs. All of them have been implemented by using standard HP Policy Builder APIs and related classes. Figure 21 provides details about the classes and APIs used to build these plug-ins.

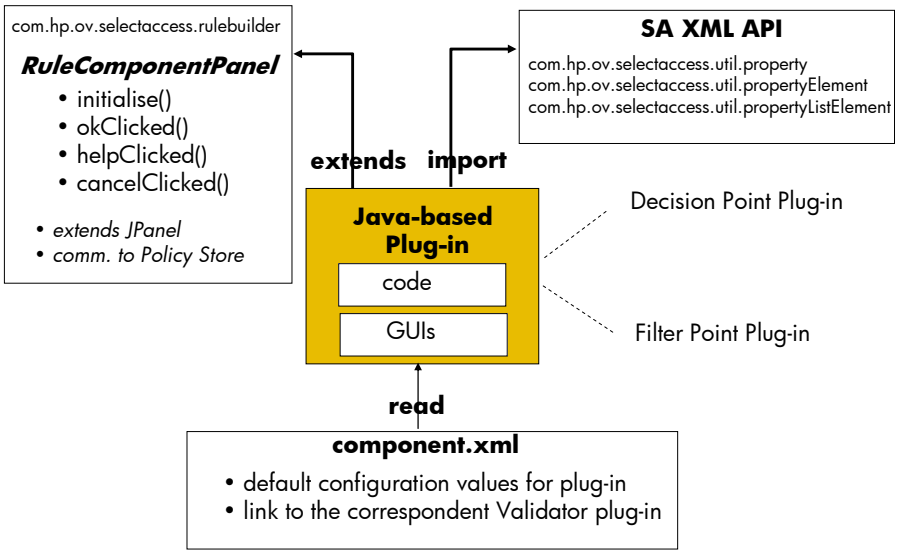


Fig. 21: HP Policy Builder –Plug-in APIs and Classes

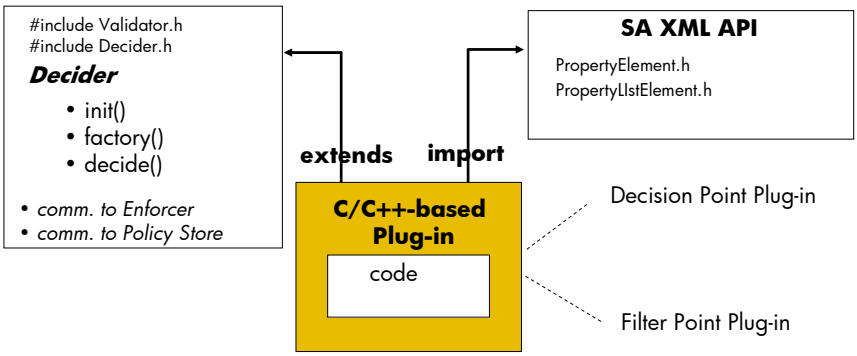
6.2.2.3 HP Validator: Making Decisions based on Privacy Policies

The HP Validator has been extended to make decision based on the privacy policy constraints and conditions described in the Policy Builder. The following types of access decisions on personal data are supported by this extended version of the Validator:

a) **No**: access to data is denied;
b) **No & conditions**: access to data is denied but some conditions are sent back to the requestors;
c) **Yes**: access to data is granted;
d) **Yes & conditions**: access to data is allowed, under the satisfaction of the attached conditions. Among other things, these conditions might require to perform data transformations and manipulations.

The support for decisions of type (d) – *Yes & condition* – has been explicitly added to HP Select Access thanks to our extensions: it allows the *data enforcer* to handle the access to personal data in a fine grained way, according to constraints and conditions described by the relevant privacy policies.

For each plug-in introduced in the HP Policy Builder, a correspondent plug-in has been implemented for the HP Validator, to convey to the system its semantic at the decision-making time: this includes which data has to be transformed/filtered, how to retrieve the consent given by data subjects, how to deal with data retentions' constraints. All these plug-ins have been implemented by using standard HP Validator APIs and classes. Figure 22 provides details about the classes and APIs used to build them.



init()	register plug-in
factory()	retrieve portion of privacy policy from LDAP. Create plug-in instance
decide()	•decision point plug-in: decide path to follow based on Enforcer's request •filter point plug-in: add constraints to Validator's reply

Fig. 22: HP Validator - Plug-in APIs and Classes

6.2.2.3 Data Enforcer: Enforcing Privacy Decisions on Personal Data

The *data enforcer* is a new component added to HP Select Access to:

- Intercept attempts (of applications and services) to access personal data stored in data repositories;
- Interact with the HP Validator to obtain a privacy-based decision;
- Enforce this decision.

The data enforcer is a "data repository proxy". Applications, services and requestors believe they are still interacting with the required data repository via standard protocols or mechanisms (ODBC/JDBC, LDAP, etc.). However, they actually interact with the data enforcer that will act as a proxy to enforce decisions based on privacy policies. The data enforcer component satisfies the following requirements:

- It is located nearby data repositories for performance reasons;
- It knows how to access/handle data and related "queries". More than one data enforcer might be required, depending on the different types of data repositories that are managed: RDBMS database, OO-DBMS database, LDAP repository, meta/virtual directories, etc.;
- It knows how to enforce privacy constraints on specific types of data;
- It can provider "query rewriting" capabilities (i.e. filtering, etc.).

The data enforcer has been designed to have a general purpose engine in order to interact with the HP Validator: it can have one or more "constraint enforcement engines" to interact with specific data repositories and enforce related policy constraints. Figure 23 summarises its design principles and architecture.

At the moment, a data enforcer has been implemented for RDBMS databases as a proof of concept. It is a JDBC proxy that intercepts JDBC calls from applications and services, interacts with the HP Validator and enforces its privacy decisions on personal data stored in the database. These data can be filtered or obfuscated, depending on the privacy constraints, before being returned to the data requestor.

Figure 24 shows the architecture of two versions (currently available) of this data enforcer: (1) standalone and (2) client/server. The standalone version can be deployed within the application/service by using the appropriate classes. Its advantage is the overall simplicity: its disadvantage is that it causes the potential loss of independence of the privacy-enforcement framework from applications/services and the exposure of any database secrets. The client/server version addresses these problems. In this case, the data enforcer runs in a standalone server. The applications/services use "lightweight and extended" JDBC classes (client part): this "client part" will contact the standalone server. This mechanism is transparent to applications/services, i.e. the way they access databases via JDBC is exactly the same.

In both versions, the requestor's "intent" is transmitted as an additional parameter during the database connection phase (it is passed either by the requestor or the application/service). These parameters are managed by the data enforcer engine.

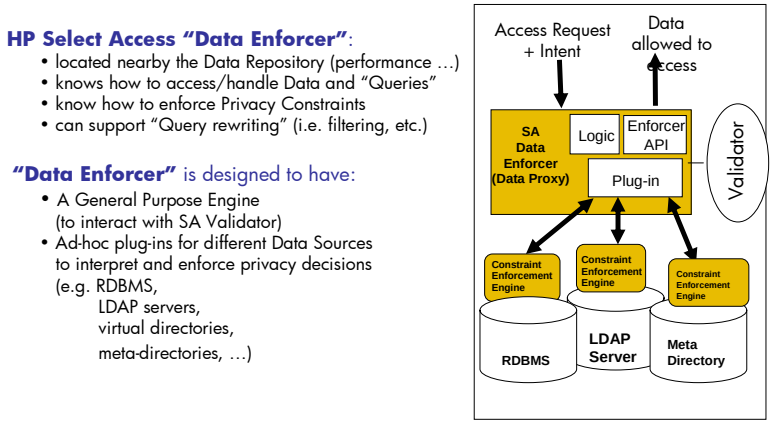


Fig. 23: Data Enforcer – Design Principles and Architecture

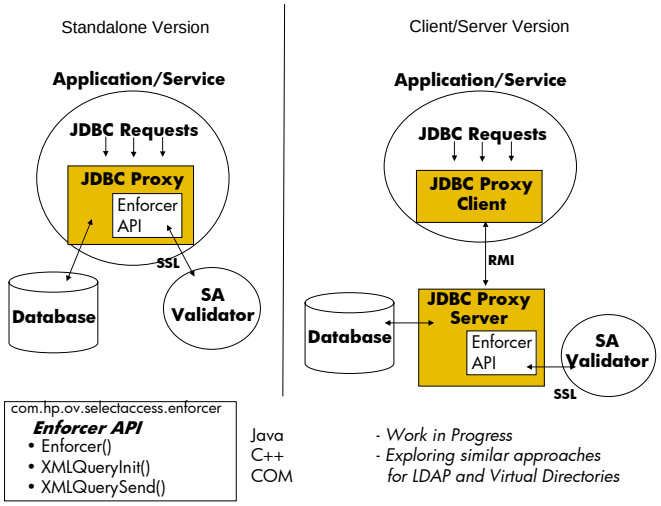


Fig. 24: Data Enforcer for RDBMS databases – Available Implementations

The same principles and approach can be used to implement *data enforcers* for other types of data repositories, including LDAP repositories, meta/virtual directories, etc.

7 Discussion

Our extension of the HP Select Access product addresses the requirements described in Section 4. It addresses the need to explicitly model personal data, explicitly author and manage privacy policies, deploy and enforce them. Privacy policies are extensible, given the plug-in capabilities of the Policy Builder and Validator and their internal XML representation. Additional privacy constraints and conditions can be managed by adding additional plug-in components to these Select Access modules.

The current HP Select Access system already provides some auditing capabilities: additional audit log information can be generated, if required, during the evaluation of privacy policies by the Validator and by adding further instrumentation to the Data Enforcer.

An important capability provided by the extended version of HP Select Access is its integrated management of access control policies with privacy policies. In addition to privacy policies, traditional access control policies on web resources, applications and services can be authored, deployed and enforced by using the same HP Select Access framework. Specifically, it is possible to use exactly the same tools - i.e. the Policy Builder and Validator - to author and deploy these policies. Different instances of enforcers, e.g. web enforcer and data enforcer (all of them based on the same template – i.e. APIs and interaction mechanisms with the Validator) can be used to enforce decisions. An advantage of this approach, if compared against other solutions, is that administrators can author and manage security and privacy policies by using the same solution without having to swap between different specialized and vertical UIs and tools. The integration of the managed resources (web resources, data resources, etc.), policies, authoring, deployment and enforcement framework introduces an overall rationalization and simplification of the policy management and enforcement process.

We believe that the approach based on “data enforcers” can really minimize the impact on applications and services and make the process of enforcing privacy policies as much transparent as possible. This is achievable via the “proxy-based approach” where common protocols and mechanisms (ODBC, JDBC, LDAP, etc.) used to access data are kept with potential minimal changes (extensions).

We have built a fully working prototype and a related demonstrator to show the feasibility of the proposed privacy model and how to leverage and extend HP Select Access. Of course, our current work is a proof of concept. This means that the current prototype cannot be used, as is, in a production environment. Additional work needs to be done to achieve this objective. In particular, our data enforcers need to be further refined and related issues explored in terms performance, in particular when handling time-intensive and complex query transformations.

Further work has to be done to explore how to automatically populate the “data model” in the Policy Builder. Using data repository introspection or similar techniques could be a viable way to tackle this problem.

We also need to fully understand the implication of our privacy enforcement on the “business logic” of applications and services. Despite the fact that these applications and services will continue to interact with data repositories by using the same protocols and mechanisms, the outcome of their data requests (queries) is now af-

fected by data transformations and filtering, as dictated by privacy policies. Applications and services need in some way to be "robust" to these results. This aspect need to be addressed in a wider and more holistic way, as part of "Design for Privacy" initiatives.

8 Current Status

A full working prototype and a demonstrator - based on HP Select Access 6.0 - have been implemented as a proof of concept to show the feasibility of our privacy enforcement model and related solution.

Our demonstrator is based on a healthcare scenario where patients' personal data is collected and can be used for different purposes (including research and marketing). In this scenario, a healthcare service provider collects personal data via a web portal. Patients, whilst disclosing their personal data (or self-administering these data), can define their consent choices and data retention constraints. The service provider uses our extended version of HP Select Access to author, deploy and enforce privacy policies on these data. Figure 25 shows the high level architecture of our demonstrator.

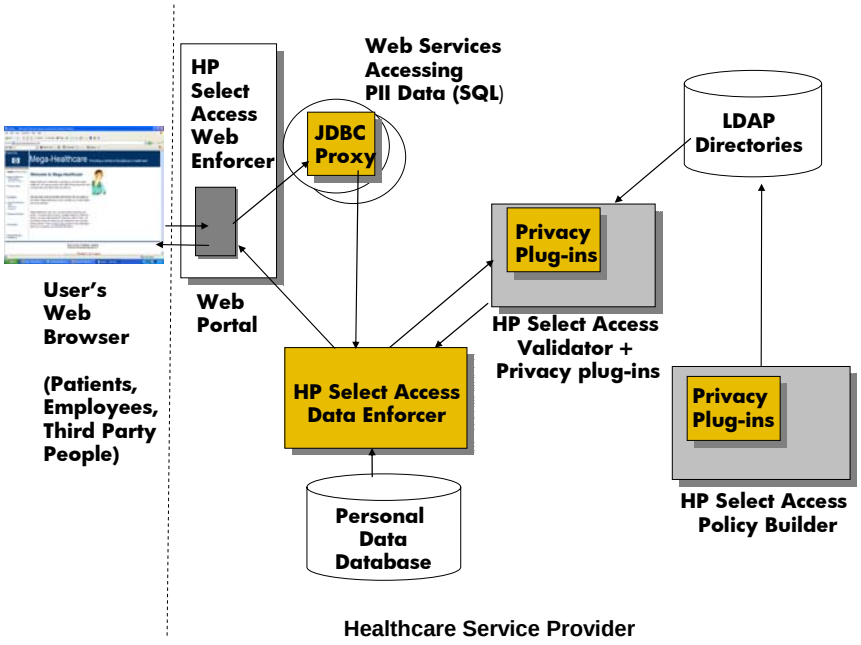


Fig. 25: Demonstrator – High Level Architecture

This demonstrator shows how different "views" of personal are actually retrieved and accessed by data requestors, depending on their roles, their intents, the personal data,

the specific consent given by patients and the privacy policies defined by the service provider. If the intent of a requestor or the consent given by patients or service providers' privacy policies are modified, the returned result will change accordingly, without having to modify the involved applications and web services.

Specifically we can show that: (1) it is possible to compose on-the-fly SQL queries on personal data stored in a database; (2) these queries are intercepted by the data enforcer; and (3) data are transformed (by filtering out part of the attributes) and returned to the data requestor as an effect of privacy enforcement.

Figure 26 shows a possible way to compose SQL queries in our demo, via the web portal. In this example a service provider's employee, with "Marketing" intent is trying to access all personal data stored in the RDBMS database. The "intent" of the data requestor is contextual to the web service: our demo can also show how the requestor can explicitly declare his/her intent and pass it along with the query (of course, all these steps are audited).

Fig. 26: Demonstrator – Query Composition from Web Portal

Figure 27 shows the actual result returned to the data requestor, after the query shown in figure 26 is sent to the RDBMS database and is intercepted by our data enforcer. The result is determined by enforcing patients' consent constraints and privacy policies defined by the service provider. Specifically, the personal data of patients that gave no consent for "Marketing" have been filtered out (horizontal yellow lines). In addition the enforcement of the privacy policies defined by the service provider is

reflected by filtering out a few fields in each patient's record (vertical yellow lines): these fields include SSN, GP, and others containing medical information.

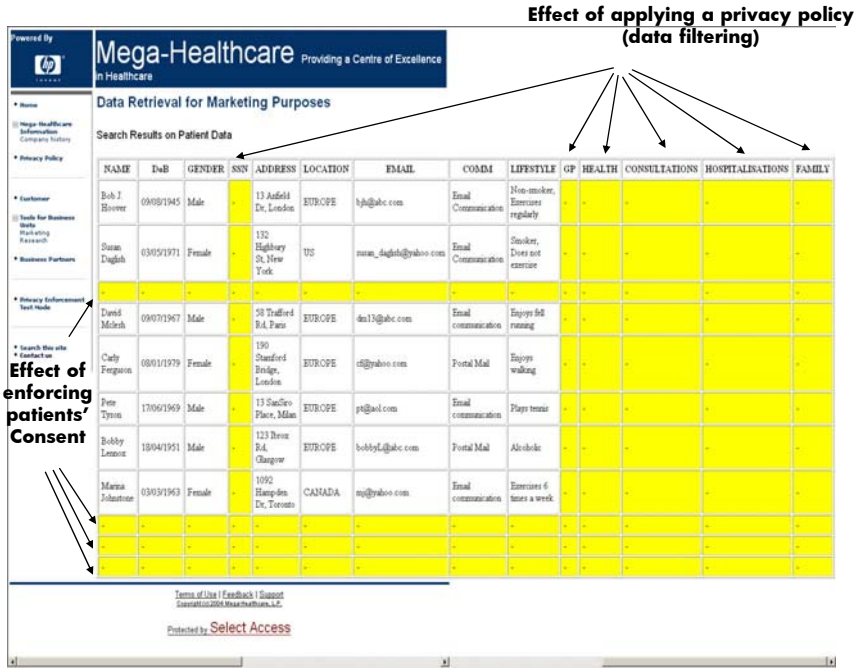


Fig. 27: Demonstrator – Returned Result after Privacy Enforcement

9 Next Steps

We plan to extend our prototype by adding a fine-grained management of the consent given by data subjects on their personal data, additional constraints on the usage of these data (plug-ins) and more complex privacy policies. An evolution of this work could be commercialized as a "privacy add-on" for HP identity management solutions.

We recognize that our work on privacy enforcement has to be considered in a more comprehensive regulatory compliance context, in particular by including privacy obligation management and enforcement capabilities [21,22], extended auditing capabilities [23,24], policy violation analytics and reporting capabilities. Work done in other HPL projects could be leveraged and integrated with this work.

Further research and development is also required to address aspects related to the "data enforcer" component, in terms of the flexibility of the query interception mechanism, its efficiency and scalability.

10 Conclusions

Privacy is important for enterprises: privacy management is required for regulatory compliance. The emphasis of most of the current solutions is on auditing and reporting aspects. The explicit management and enforcement of privacy policies on personal data (stored and processed by enterprises) is another important aspect but still a green field. This aspect is currently addressed with ad-hoc or very vertical solutions.

Dealing with the problem of privacy enforcement includes: modeling personal data, dealing with data purposes, checking requestors' intent against data purposes and customers' consents, defining and enforcing enterprises and customers' constraints as transparently as possible to applications and services. Privacy-aware access control systems are required.

In this paper we specifically address these issues: we propose a solution based on an extension of the HP Select Access product. Specifically, we describe a privacy enforcement model and a technical approach to model personal data, author privacy policies and customers' consent, deploy and enforce them in an integrated framework. The management of access control policies is integrated with the management of privacy policies. This brings simplicity and rationalises the required set of management and enforcement tools.

A fully working prototype and a demonstrator have been built as a proof of concept, to demonstrate the feasibility of our ideas.

Acknowledgements

A special thank to Adrian Baldwin (HP Labs, Bristol) for his help to refine the involved concepts and his input during the design of our privacy-aware access control system.

References

1. Laurant, C., 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. OECD: OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data. http://www1.oecd.org/publications/e-book/9302011E.PDF (1980)
3. Online Privacy Alliance: Guidelines for Online Privacy Policies. http://www.privacyalliance.org/, Online Privacy Alliance (2004)
4. Karjoth, G., Schunter, M.: A Privacy Policy Model for Enterprises. IBM Research, Zurich. 15th IEEE Computer Foundations Workshop (2002)
5. Karjoth, G., Schunter, M., Waidner, M.: Platform for Enterprise Privacy Practices: Privacy-enabled Management of Customer Data. 2nd Workshop on Privacy En-

hancing Technologies, Lecture Notes in Computer Science, Springer Verlang (2002)
6. Schunter, M., Ashley, P.: The Platform for Enterprise Privacy Practices. IBM Zurich Research Laboratory (2002)
7. Karjoth, G., Schunter, M., Waidner, M.: 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. Agrawal, R., Kiernan, J., Srikant, R., Xu, Y.: 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/
11. IBM Tivoli Privacy Manager: online technical documentation - http://publib.boulder.ibm.com/tividd/td/PrivacyManagerfore-business1.1.html
12. HP: HP Select Federation - Product and Solution Overview - http://www.managementsoftware.hp.com/products/slctfed/
13. ePok: identity management solution - Trusted Data Exchange Server - http://www.epokinc.com/
14. Synomos: Align 3.0 suite - http://www.synomos.com/
15. Quest Software: InTrust Auditing solutions - http://wm.quest.com/products/intrust/
16. NetForensics: Compliance and Audit solutions - http://www.netforensics.com/
17. AddaMark: SenSage solutions - http://www.addamark.com/
18. HP: HP Openview Select Access - Overview and Features - http://www.openview.hp.com/products/select/
19. HP: HP Openview Select Access - Product Manuals - http://ovweb.external.hp.com/lpe/doc_serv/
20. HP: HP Identity Management Solutions and HP Select Access - http://www.managementsoftware.hp.com/solutions/im/index.html
21. Casassa Mont, M.: Dealing with Privacy Obligations: Important Aspects and Technical Approaches, TrustBus 2004 (2004)
22. Casassa Mont, M.: Dealing with Privacy Obligations in Enterprises, ISSE 2004 (2004)
23. Baldwin, A., Shiu S.: Enabling shared audit data, IJIS (to appear) (2004)
24. Baldwin, A.: Enhanced accountability for electronic processes, 2nd international conference on trust management. Lecture notes in computer science, vol. 2995, Springer (2004)