The Platform for Enterprise Privacy Practices

Matthias Schunter
IBM Zurich Research Laboratory
Zurich, Switzerland
mts@zurich.ibm.com

Paul Ashley
IBM Software Group - Tivoli
Gold Coast, Australia
pashley@au1.ibm.com

Abstract

Enterprises collect personal data while promising fair information practices to their customers. The Platform for Enterprise Privacy Practices (E-P3P) enables an enterprise to keep the privacy promises made. It formalizes the privacy promises into policies and associates a consented policy to each piece of collected data. This consented policy can then be used in access control decisions to enforce the privacy promises made.

Introduction

Consumer privacy[1] is a growing concern in the marketplace. While the concerns are most prominent for e-commerce, the privacy concerns for traditional transactions are increasing as well. Some enterprises are aware of these problems and of the market share they might loose if they do not implement proper privacy practices. As a consequence enterprises publish privacy statements that promise fair information practices[2].

Unfortunately, companies usually face the following problems:

- Enterprises create stockpiles of personally identifiable information (PII[3]). Larger enterprises may not know what types of PII are collected and where it is stored.
- Enterprises may neither know the consent a customer has given nor the legal regulations that apply to a specific customer record.

There are a number of risks to an enterprise if it does not manage its PII correctly. The main risks are:

[1] Privacy is the right of individuals to determine for themselves when, how, and to what extent information about them is communicated to others (Alan Westin).
[2] The OECD defined a set of privacy principles in 1980. The document "OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data" is considered to contain the core requirements for managing privacy today.
[3] Information is considered PII or Personally Identifiable Information if it can be linked to a person. Information that has been de-personalized or anonymized would not be considered PII - unless there are ways of linking it back to the individual through re-personalization or inference.

1. *Legislative Penalty*: Recently there has been new privacy legislation enacted in many parts of the world. Most of these laws incorporate rules governing collection, use, retention and distribution of PII. It is up to an organization to ensure that it is compliant with any legislative requirements or industry regulations that apply to it.
2. *Brand and Reputation Erosion*: Business relationships are built on trust. Organizations that demonstrate good privacy practices can build trust. Organizations with poor privacy practices will alienate customers.
3. *Lawsuits*: Lawsuits against organizations that violate privacy regulations or promises are becoming more common. A quick search of the United States Federal Trade Commission Web Site [1] will find a number of companies that have both been charged by the FTC on privacy violations and that are under a class action lawsuit from their customers.

As a first step towards managing privacy effectively organizations publish privacy promises as text or P3P [2]. The P3P statements can be used by a P3P client (e.g. the Internet Explorer 6 web browser) to notify the user automatically whether the privacy policy of the enterprise matches that configured by the user. Although this is beneficial for stating an enterprise's privacy promises, enterprises do not have the privacy technology to enforce the promises throughout the enterprise. This has resulted in privacy violations being a common occurrence today, even from well meaning companies.

This article describes a new approach towards enterprise-wide enforcement of the privacy promises. Its core is a new framework for managing collected personal data in a sensitive, trustworthy way. The framework, called the Platform for Enterprise Privacy Practices (E-P3P), enables enterprises to formalize and enforce privacy practices and to manage the consent of their customers.

The paper is set out as follows. The next section describes the requirements for managing private data in an enterprise. This includes five steps that we consider to be the core of privacy management. The following section outlines our new privacy policy language. The core of this language is the ability to define an enforceable privacy policy and implement it across an enterprise. We finish with a summary.

Implementing Privacy Management

For an enterprise to enforce its privacy promises and act as a custodian of their customer's PII they need to implement the following phases:

1. Define an enterprise privacy policy. We have defined a new privacy policy language called E-P3P for this purpose.
2. Deploy a policy to the IT systems that contain privacy sensitive information.
3. Record consent of end users to advertised privacy policy when they submit privacy sensitive data.
4. Enforce the privacy policy and create an audit trail of access to privacy sensitive information.

5. Generate both enterprise wide and individualized reports showing accesses to privacy sensitive information and their conformance to the governing privacy policy.

Figure 1 shows a high level architecture for implementing privacy management. The CPO (or one of her staff) defines the high level privacy policy using E-P3P. This policy is at a high level to map to regulatory and industry requirements. Once the privacy policy is defined it is deployed across the IT systems that deal with PII data. This way the privacy policy can be enforced.

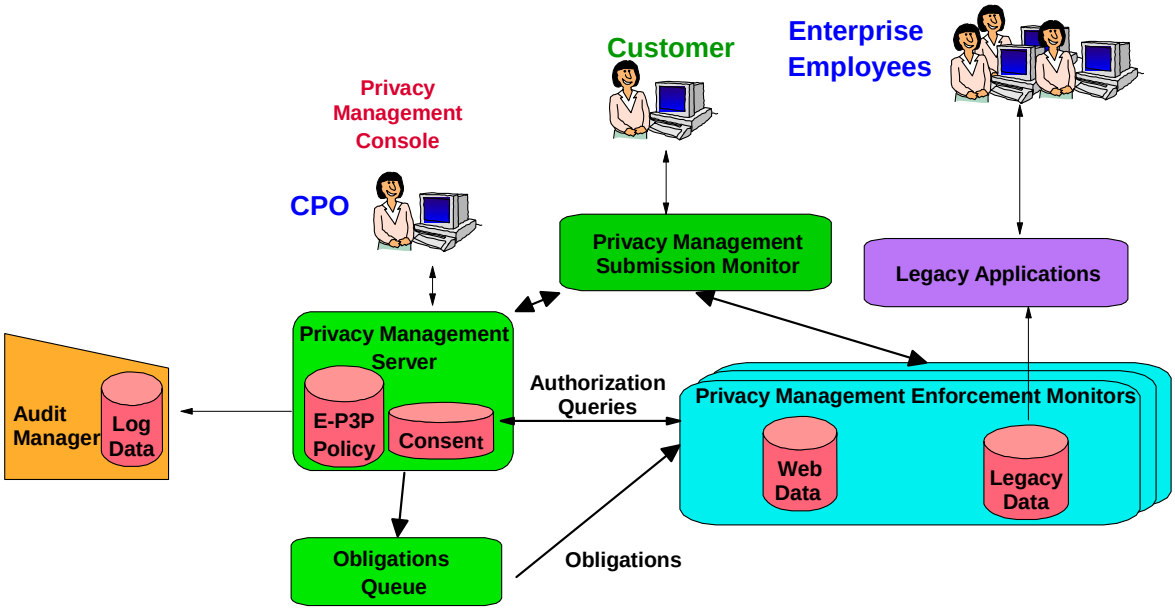

Figure 1: Implementing Privacy Management Using E-P3P

During submission of PII data, the privacy management system must create a submission record. Submission monitors will create the submission records. This data is a permanent record of when PII data was submitted by a person, what privacy policy version was in place at that time, and what the user's preferences were.

Later, when the PII data is to be accessed for use by the enterprise, the privacy management enforcement monitors ensure that only data accesses are allowed that conform to the privacy policy. They also create access records that record which user accessed the data and for what purpose.

The combination of submission and enforcement monitors allow the enterprise to demonstrate that it is a good custodian of data and gives the enterprise some assurance that it is enforcing its stated privacy policy.

Defining the Privacy Policy

The first step in implementing a privacy management solution is to allow for the Chief Privacy Officer (CPO) or her staff to create an enterprise privacy policy. An enterprise privacy policy states the rules about the collection and use of PII. Privacy policies are defined by people who understand the business and legal environment of the organization and typically express conceptual requirements from the applicable law and business strategy. Privacy policies do not refer to specific applications or systems in the IT infrastructure, nor do they refer to specific technologies.

The discussion of our new privacy policy language called E-P3P which has been created to define an enterprise privacy policy is left until later in the paper.

Deploying a Policy to the IT Systems

After the CPO or an equivalent person has created an enterprise privacy policy using E-P3P, the IT staff can now deploy this to the IT systems within the enterprise. Deploying a policy consists of three steps:

1. Mapping the Data Types defined in the privacy policy to the PII that is stored in the IT systems.
2. Mapping the Data Users defined in the privacy policy to enterprise roles that are defined in the IT systems.
3. Mapping the tasks that IT systems and applications perform into policy defined business purposes.

The mapping allows for the rules engine to *resolve* a physical data access on an IT system with the privacy policy that has been defined. Deploying the privacy policy also involves integrating submission and enforcement monitors into the enterprise.

Recording Consent of End Users to Advertised Privacy Policy

At the heart of managing PII is to ensure that a user has consented to use of their data before its used within the enterprise. The user should explicitly consent to the privacy policy advertised, and to each and every purpose of use in the enterprise. An enterprise should not accept any PII until the user has consented to the privacy policy in place and consented (positively or negatively) to the use of the data for each purpose. Besides recording the collected data, this requires the following privacy management data:

- An identifier of the person whose data is being submitted
- The PII types being submitted

- The storage of user consents
- The time of the data submission
- The applicable version of the privacy policy

This comprises all information necessary to govern all future usage of the data. An important point to note, is that the data is submitted under a particular privacy policy, and should be linked to that policy. We call this approach to managing PII " ***the sticky policy paradigm***". If the enterprise's privacy policy is updated (which it will in time), it is important that the user's data is managed under the policy at consent time, and not to this new policy. Only if the user explicitly consents to the new policy should that data be treated under this new policy.

The submission monitors create the detailed records of PII submissions.

Enforce the Privacy Policy and Create and Audit Trail of Access

The other key task is to watch for applications accessing privacy data in a protected system. This requires identifying whose data is being accessed, its PII type, who is accessing the data, the purpose for the access and the time the access occurred. This information is used to retrieve the submission record corresponding to the data that is being accessed, the governing privacy policy and the user's consents, and finally to decide whether access shall be granted or not.

The enforcement monitors make sure that the access complies with the enterprise privacy policy.

Generate Both Enterprise Wide and Individual Reports

Being able to report on activities relating to PII is an essential part of privacy management. This requires generation of reports both at an enterprise-wide perspective and at an individual perspective. For example, an individual may make a request to an organization, "What data do you have stored on me, who has been accessing it and for what purpose?". An auditor may ask "Please show me a report showing any PII accesses that were outside of privacy policy or user consents". Because the privacy management system has kept very detailed audit records of submissions and accesses, both of these questions can be answered.

E-P3P a New Privacy Policy Language

The Platform for Enterprise Privacy Practices (E-P3P) is a language for formalizing enterprise-internal privacy practices[4]. Its core is an authorization scheme that defines whether certain operations are allowed or not.

[4] These privacy promises can be formalized using P3P [1]. Note that P3P is coarser than the enterprise privacy practices of E-P3P. While P3P may promise ``we do not disclose data", E-P3P may define the exact flows that are authorized in an enterprise.

Requirements of Privacy Policies

The exact syntax of the set of rules defining a privacy policy will depend on the language used. However, in general the policy contains the following elements:

- *Data Users*: Data Users are used to classify individuals who are accessing or receiving data. Data Users are required in a privacy context, as privacy policies will depend on the relationship between the individual requesting data and in the individual who the data is about. For example, one type of Data User might be *physician* while another might be *primary care physician*.

- *Operations*: Some privacy policies make distinctions about who can perform activities based on the action begin performed. For example, a policy might say that anyone in the company can *create* a customer record, but that only certain Data Users are allowed to *read* that record.

- *Data Types:* Privacy policies must define the types of data which the enterprise will be holding. Typically, the data types used in privacy policies are high-level descriptions of data, such as *customer contact information*. Detailed, low-level descriptions are not typically required in privacy policies.

- *Purposes*: Data access requests are made for a specific purpose or purposes. This represents how the data is going to be used by the recipient. For example, the data may be used for *marketing* or for *fulfilling* the individual's order.

- *Conditions*: Rules can be qualified based upon additional conditions. Often, legislation or privacy policies make statements based on specified conditions. For example, COPPA [3] imposes requirements on data received from persons less than 13 years of age. Another common condition is that the user must have consented before PII can be used for a particular purpose.

- *Obligations*: A privacy policy may also state that when a certain access is allowed, the enterprise must take some additional steps. An example is that all accesses against a certain type of data for a given purpose must be logged. Another might be that PII must be deleted if its owner has not performed business with the enterprise for one year.

The elements are then used as the terminology to express privacy rules expressing what requested data accesses are allowed or denied, and under what conditions:

ALLOW [Data User]
TO PERFORM [Operation] ON [Data Type]
FOR [Purpose] IF [Condition].
CARRY OUT [Obligation].

Or

DENY [Data User]
TO PERFORM [Operation] ON [Data Type]
FOR [Purpose] IF [Condition].
CARRY OUT [Obligation].

Details of the Language

The E-P3P policy language categorizes the data an enterprise holds and the rules which govern the usage of data of each category. Since E-P3P is designed to capture privacy policies in many areas of responsibility, the language cannot predefine the elements of a privacy policy. Therefore, an E-P3P document either starts with definitions of the elements which are used to build the policy or else imports them from a base-policy.

An E-P3P policy is essentially a set of privacy rules. A rule is a statement that includes a ruling, a data user, an action, a data category, and a purpose. A rule may also contain conditions and obligations. Each rule contains a precedence level.

A E-P3P policy document consists of three main sections:

1. *Policy Information*: This is used to identify the policy. It consists of information such as Issuer, Version Number, Start Date, End Date, Replacement Policy Name, Replacement Policy Version.
2. *Definitions*: This defines all of the possible components that can be used in the following rules. Here is where Data Users, Data Categories, Purposes, Actions, Context Models, Conditions and Obligations are defined.
3. *ALLOW or DENY*: Rules to define whether Data Users are ALLOWed or DENYed to perform Action on Data Category for Purpose under Conditions.

The following examples show how E-P3P is used to express privacy policy rules:

privacy policy (informal)	*Allow a sales agent or a sales supervisor to collect a customer's data for order entry if the customer is older than 13 years of age and the customer has been notified of the privacy policy. Delete the data 3 years from now.*
ruling	allow
data user	sales department
action	store
category	customer-record
purpose	order-processing
condition	the customer is older than 13 years of age

obligation	delete the data 3 years from now

Rules are used to determine if a request is allowed or denied. A request contains a data user, an action, a category, and a purpose. Continuing with the same enterprise as above, consider the following request.

request (informal)	*A person acting as a sales agent and an employee requests to collect a customer's email for order entry.*
data user	sales department
action	store
data type	customer-record
purpose	order-processing

The above rule allows the request, so the sales agent would be permitted to store the customer's contact information. Additional rules can then govern how this stored data may be used.

Here are some other examples showing the rules defined in XML:

1. Email can be used for the book-of-month club only if consent has been given and age is more than 13:

   ALLOW data-user = “borderless-books”operation = “read”data-category = “e-mail”purpose = “book-of-month-club”condition = “/Owner/Consent/BookClub && /Owner/Age > 13”

   In E-P3P XML:

```
<rule id="rule1" precedence="5" ruling="allow">
        <data-user id="borderless-books"/>
        <data-category id="email"/>
        <purpose id="book-of-the-month-club"/>
        <operation id="read"/>
        <condition id="consentToBookClub"/>
        <condition id="olderThan13"/>
        <obligation id="retention">
        <parameter id="days">5</parameter>
        </obligation>
</rule>
```

2. Parents are allowed to read any data of their children:

ALLOW data-user = “parent”, data-category = “any-category” purpose = “current”, operation = “read”, condition = “true”

In XML:

```
<rule precedence=“1” ruling=“allow”>
        <data user id = “parent”/>
        <data category id = “p3p: any category”/>
        <purpose id = “p3p: current”/>
        <operation id = “read”/>
</rule>
```

E-P3P is Complementary to P3P

There is an existing privacy policy language called P3P [2] or the Platform for Privacy Preferences. This language (version 1.0) is a W3C recommendation (at April 16, 2002).

P3P is used to express high level privacy promises on web sites:

- It allows web sites to express their privacy promises in a standard format.
- Can be retrieved and interpreted by web browsers.
- Web browsers can alert users if privacy settings at web sites are different to user settings in a browser.

Although P3P is well suited for expressing high level web site policies, it is not as suitable for expressing an internal enforceable privacy policy. E-P3P on the other hand is designed specifically to express an internal privacy policy that can be enforced by an enterprise privacy management system.

A high level view of the differences between E-P3P and P3P are shown in Table 1:

- *Categories*: P3P has a pre-defined list of data categories. E-P3P allows for an enterprise to define its own list of data categories and these may be hierarchical.
- *Data-Users*: P3P has a pre-defined list of data users. E-P3P allows for an enterprise to define its own list of data users and these may be hierarchical.
- *Purposes*: P3P has a pre-defined list of purposes. E-P3P allows for an enterprise to define its own list of purposes and these may be hierarchical.
- *Operations*: P3P only defines the operation “use”. E-P3P allows for a definable list of operations.
- *Conditions*: P3P does not define a condition language. E-P3P defines a generalized condition language using XSL Transformations [4].
- *Obligations*: P3P only defines the operation “retention”. E-P3P allows for a definable list of obligations.
- *Choices*: P3P only allows for simple opt-in/opt-out choices. E-P3P allows for a more generalized set of choices.

P3P is an excellent language for expressing high level privacy promises on web sites. However, E-P3P should be used to define an enforceable privacy policy within an enterprise. [5]

Table 1: E-P3P vs P3P

Elements	E-P3P	P3P
Categories	hierarchy	list; predefined
Data-Users	hierarchy	list; predefined
Purposes	hierarchy	list; predefined
Operations	list	'use'
Conditions	XSLT	none
Obligations	list	'retention'
Choices	generalized	+/- purpose
Conclusion	✓ **Flexible** ✓ **Hierarchical** ✓ **Access Control** ✓ **Enforceable** ↘ **Non-Interoperable**	✓ **Simple** ✓ **Interoperable** ↘ **Non-extensible** ↘ **No enforcement** ↘ **Limited** ↘ **Web-specific**

Summary

For an enterprise to be a good custodian of their customer's PII data they must be able to define an enterprise privacy policy that is enforceable across the enterprise. To do this we propose five steps in privacy management that are required to manage the PII data and enforce the privacy policy. We also define a new privacy policy language called E-P3P, which is specifically designed to be an internally enforceable privacy policy language for an enterprise. E-P3P is complementary to W3C's P3P.

[5] We are currently investigating how to project a P3P policy from E-P3P.

Bibliography

[1] Federal Trade Commission, Available at http://www.ftc.gov (August 2002)

[2] W3C, The Platform for Privacy Preferences (P3P),W3C Recommendation, 16 April 2002, Available at http://www.w3.org/TR/P3P (August 2002).

[3] COPPA, Children's Online Privacy Protection Act of 1998, October 1998

[4] W3C, XSL Transformations (XSLT), W3C Recommendation, 16 November 1999, Available at http://www.w3.org/TR/xslt (August 2002).