

A Comparative Study of Privacy Mechanisms and a Novel Privacy Mechanism [Short Paper]

Gunmeet Singh and Sarbjeet Singh

University Institute of Engineering and Technology, Panjab University
Sector-14 Chandigarh, India
gunmeetsingh@gmail.com, sarbjeet@pu.ac.in

Abstract. Privacy of PII(Personally Identifiable Information) on the Internet is a major concern of a netizen. On the Internet different service providers are supposed to publish their own privacy policies but understanding of these policies is a major problem. Standards like Platform for Privacy Preferences(P3P), provide a computer readable format and a protocol for allowing web browsers to retrieve and process privacy policies. In this paper we studied the various privacy mechanisms in place and compared them on the basis of their architecture and third party intervention. We also proposed an alternative privacy mechanism that introduces the concept of a third party whose role is to verify the privacy policy and keep a proactive check on the use of specified PII's. In case of a violation the third party, informs the users of the breach. The implementation of the proactive check on the PII has been done through software agents. The requirement of granting legal status to transactions of the PII by the use of Digital Signatures and PKI has also been proposed,thereby legally binding the web entity to use the PII as per the agreed terms.

Keywords: Privacy, Trusted Third Party, Security,P3P, EPAL, Digital Signatures, Personally Identifiable Information (PII), Security.

1 Introduction

The growth of web services which require the use of PII's has increased manifold. Hence the use and distribution of the PII's between business entities have increased. The misuse of the PII which has resulted in crimes like spim , spam and junk mails. As PII itself is an identity of a netizen on the Internet. Therefore the use of PII itself should be checked and verified by the user at the service providers end. The interchanging of the PII's between business entities should also be notified to the user. . In case the entities commit a misuse of the PII they should be legally held for such a breach of confidence. Paradoxically the netizen will express very strong concerns about privacy of their PII, but be less than vigilant about safeguarding it [1]. Thereby requiring the inclusion of third party to safeguard the PII.

Web Services are associated with a Privacy Policy which states the objectives and aims of using the PII's [2][3]. Trade practices require that the privacy policies should be stated by the web services and laws like GLBA(Gramm Beach Bliley Act) state that the language of the policy stated should be in very simple and in clear terms[3].But the users of these services do not use the policies as they find these policies complex and difficult to understand. This does not allows user to make an informed choice about sharing their PII's.W3C provided a mechanism which allows the web services to state their policies in XML encoded form called P3P(Platform for Privacy Preferences)[5].This XML encoded form makes it easier to understand and provide a standardized way of stating privacy policy. And as compared to the privacy policies stated in Human Readable languages ,it is much easier to understand. Different other mechanisms like E-P3P, EPAL provide and enforce privacy policies inside Web Service entity. These restrict the access to the PII's to different groups inside an entity. But this whole approach to privacy is an inactive one ,once the PII is given to the Web Service there is no check to how the data is being used. We propose a privacy mechanism that solves this problem by introducing the concept of trusted third party which monitors the use of the PII by the web service. And in case there is a unauthorized use of the PII which violates agreed terms between the user and the web service. Thereby keeping a proactive check on the use of the PII.

2 Different Privacy Mechanisms

Most of the transactions for any service e.g. Setting up an email account on a email service, on the Internet is never complete without the exchange of PII. Hence the need to ensure the privacy of the PII is very important and must be addressed technically. Hence various mechanisms were developed and adopted the most popular being the P3P and EPAL. P3P was the first to be introduced and helped the privacy policy of the company to be stated in the machine readable form. The next step is to allow user to specify its privacy preference with the P3P document of the service. This is implemented with the use of user agents which allow the comparison of the policy and the preference e.g. AT& T Privacy Bird[6].The natural extension to this is to enforce the privacy policy of the company through out the the organization. EPAL is the mechanism that implemented the privacy policy through out the organization. This allows transparency and the synchronization of the privacy policies and the internal PII usage practices.

2.1 P3P (Platform for Privacy Preferences)

P3P is a standard defined by W3C,that allows a Web Service to state the privacy policy in a standardized machine readable form. The privacy policy states all the objectives of a web service regarding the information collected in a semi-structured XML form. The P3P specification[5] has a standard vocabulary to describe data practices which states the entities that will access the data and the purpose. Base data Schema is used for collecting information. An overview of the P3P vocabulary as stated in the P3P Specification[5] is described in the Table 1.

Table 1. Overview of the P3P vocabulary

P3P Policy Element	Detail
Entity	The web service or website which collects the information. This element stores the contact information of the entity which will collect the users PII
Access	This element specifies the which sub-entities can access the PII
Dispute	Describes how to resolve related disputes with the web service
Data	The kind of data collected by the web service
Purpose	This element states the purpose for which the PII collected will be used
Recipient	States the entities with which the data will be shared
Retention	Describes the retention policies of the information collected
Consequences	Human Readable element and explains the web services data practices

The privacy policy in the P3P Specification uses the above vocabulary to state their privacy policy. The P3P specification also has a protocol,built on the HTTP protocol to transmit and receive the privacy policies

2.1.1 Overview of P3P

The Privacy Preferences is the XML format defined under P3P which provides for the user to state the privacy preferences and also provides the algorithms for matching of the user privacy preferences with the web services privacy policy which are stated in the W3C APPEL. The overview of P3P and the various steps involved are shown in the Figure 1.

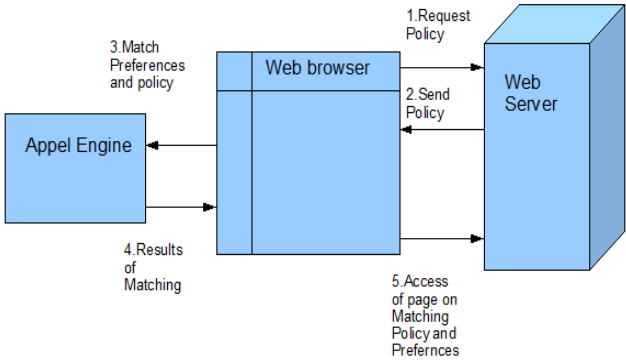







Fig. 1. Overview of P3P

2.1.2 Overview of User Agents

User Agents are tools that fetch the P3P policy of the site. The user agents like AT&T privacy bird, Microsoft IE 6 and Privacy Companion allow users to specify their privacy preferences [6]. They can either be built into the web browser or can be end user applications. User agents also compare the privacy policies to privacy preferences and gives advice to users to whether to exchange PII with the specified web service. If the privacy policy is found in conflict with the user privacy preference the user agent takes the appropriate action e.g. blocking the page, Displaying a warning on the status bar. Hence allowing the user to make a choice to whether share the information e.g. AT&T Privacy Bird a very popular user agent uses the following Symbols are showed in the header of the Web Browser [6],[7] to show different status. Different symbols are described in Table-2.

Table 2. Symbols and Descriptions of AT&T Privacy Bird [6]

Header Symbols	Details
	This symbol indicates that the privacy policy and the preferences matches.
	Indicates that privacy policy and the preferences match but the site contains some frames, pictures etc.
	Indicates that Web Service\Site does not have P3P policy.
	Indicates that the Privacy policy does not Match
	Indicates that the Tool has been turned of

Thereby this gives adequate warning and information about the site to the user. Hence user can make a informed choice about the site.

P3P also provides mechanism for specifying cookie related data practices. These P3P policies are referred to as “compact policies”. These are included in HTTP-Response headers and provide a quick way for a user agent to compare the policy with the preferences without referring to another document.

2.1.3 Drawbacks of P3P

P3P though provides a machine readable format for the specification of privacy policy, but the mechanism to ensure the privacy of the user is not present. The P3P privacy policy is a formal document that states the usage of the PII, but it is not enforceable throughout the web service. P3P documents are difficult to write as compared to EPAL Policy documents [8]. Due to its complicated syntax it has been not adopted widely [9].

2.2 EPAL(Enterprise Privacy Authorization Language)

EPAL[10] [11] is a XML based privacy specification language that is used by organizations to specify their internal privacy policies. EPAL is basically used by web services or entities to define how the data is accessed inside them,it also synchronizes the policies of the entities and their business partners so that compliance is assured between the respective business policies. Developed by IBM,it was defined as a formal language to specify internal privacy policies which unlike P3P are enforceable and automated across entities systems. The privacy policy in EPAL is made up of elements which are analogous to P3P's policies elements,these defines access to the data. These are shown in table 3.

Table 3. Overview of EPAL elements

EPAL Policy Element	Detail
Ruling	Specifies one rule which can be “allow” or “deny”
User Category	This element specifies which User Group can access the specified data
Action	Models how the collected data is used.
Data-category	The different kind of data collected by the web service e.g. medical record or Contact information
Purpose	This element states the purpose for which the PII will be used. Can serve as a ruling whether to allow or deny the use of PII
Obligation	This element states certain actions to be completed
Condition	This element defines external conditions

The EPAL policy stated is enforceable throughout the organization. This is done by the Enforcement Engine[10] which parses the policy stated and controls the access of the user groups to the data store. The overview of the EPAL is given below in the Figure 2.

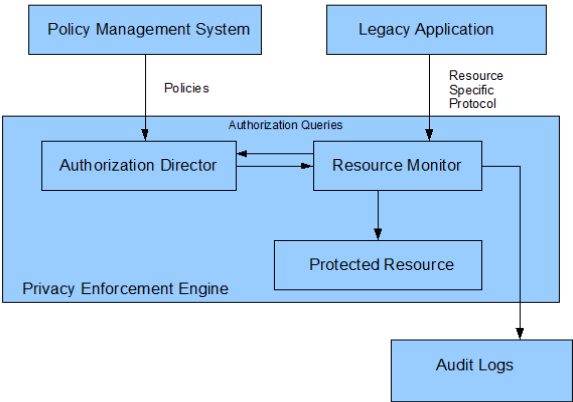


Fig. 2. Overview of the EPAL architecture

The EPAL policy is made up of rules which define the access conditions e.g. consider the EPAL policy given below

```
<epal-policy>
  <rule id="Email_Privacy_Rule_1" ruling="allow">
    <user-category refid="Subsidiaries"/>
    <data-category refid="Email"/>
    <obligation refid="Check opt-out list"/>
  </rule>
</epal-policy>
```

Consider the rule Email_Privacy_Rule_1. The ruling is “allow” which means allow access to the Email address. The user group allowed access is the Subsidiaries. The Obligation that the user group needs to perform is the checking of the opt-out list.

Another important concept of EPA called “Sticky Policy Paradigm” [10] which states that the terms and conditions which were promised or agreed upon will be applicable even if it is transferred from one entity to another.

2.3 Limitations of Current Privacy Mechanisms

The Privacy Mechanisms discussed above have the following disadvantages

1. No third party evaluation:- In both of the privacy mechanism there is no third party involved. The concept of third party ensures the impartiality of the system.
2. Non Proactive Approach towards protection of PII:-The PII is the identity of a netizen. And the user/netizen requires that the web service which has its PII to disclose the way it has been used. This can help in solving lots of problem related to proliferation of data from unscrupulous web services e.g. spam, spim, fraud.
3. Legal Status for PII transactions and its use:- PII is the identity of a netizen on the INTERNET .The privacy policies purpose is to ensure user that the PII will be safe. But to give legal recognition to this transaction of is the need of the hour.

Limitations are summarized in Table 4

Table 4. Limitations of the Current Privacy Mechanisms

Limitations	P3P	EPAL
Enforceability of privacy policy throughout the web service	No	Yes
Easily understandable syntax	No	Yes
Third Party Audit and Evaluation of PII use	No	No
Ability of user to monitor use of PII	No	No
Legal binding on exchange of PII	No	No
Legal Binding on the use of PII as per the stated privacy policy	No	No
Ability to link policy with data(Sticky Policy paradigm)	No	Yes

3 Proposed Privacy Framework

The proposed framework is a scheme for privacy-enabled management of netizen's data. Its core is an authorization scheme that defines how collected data may be used. Also a platform which facilitates third party evaluation is added to the proposed framework. The Proposed Privacy framework has four parties involved

1. User-The user represents a netizen who accesses and shares his/hers PII with the web service.
2. Web Service-The web service is an organization /entity which provides service and requires the PII from the user. Web service is also required to share some access details/logs with the trusted third party.
3. Trusted Third Party-The trusted third party is an entity which monitors the use of PII given to the web service for the user. This party legally binds the web services to follow the conditions agreed on by the user and the web service.
4. Controller For Privacy Insurers- This party will control the Trusted third party. It is a regulatory body and will control and regulate the trusted third party. Each country will have one. Since its role is similar to the CCA (Controller of Certifying Authorities) ,it can perform the role of CPI.

3.1 Prerequisites and Application Model

The Web Service run applications that collect and use PII's. Each application performs some tasks. For example a "view record" displays the record of a certain user.

Therefore there is a privacy policy that controls the access of the PII throughout the organization. Privacy policy may be informal rules that are applicable throughout the web service. These privacy policy is implemented in the form of EPAL policies in the system.

3.2 Policies Definition and Conversion

The Privacy policies state how the PII will be used. The privacy policy is defined using the EPAL Rules. These rules define the usage of a certain field in the form of information. The EPAL policy thus stated is thus converted into the P3P format [12] and hosted by the web service.

3.3 Collection and Transfer of PII

This is the first step when a user interacts with the web service for the first time and agrees to use the service and transfer the PII. The following steps take place

1. The user accepts the web services terms and conditions also stating the Opt-in and Opt-out choices and sends its privacy preferences in the XML form as stated in the privacy preferences in P3P.
2. The Web Service replies back to the user, the XML form is converted into a SQL query and it is compared with the web services privacy policy. The reason why the user sends the web service also returns its privacy policy.

- The user transfers the web services privacy policy to the trusted third party to analyze the privacy policy further. The trusted third party can store the history of the defaulter web services hence advice the user about the web services previous record
4. The trusted third party returns its analysis to the user.
5. The user and Web service transfers the PII and sign a contract which states the users preferences and the web services privacy policy with their private keys.
6. The document is further attested by the third party and a clause is added that the third party will be allowed full access to access details .The web service passes a “key” to the third party which will identify the specified PII in the web service.
7. The user is informed that the transaction is complete.

The overview of the collection and transfer of the PII is given in the Figure 3 below

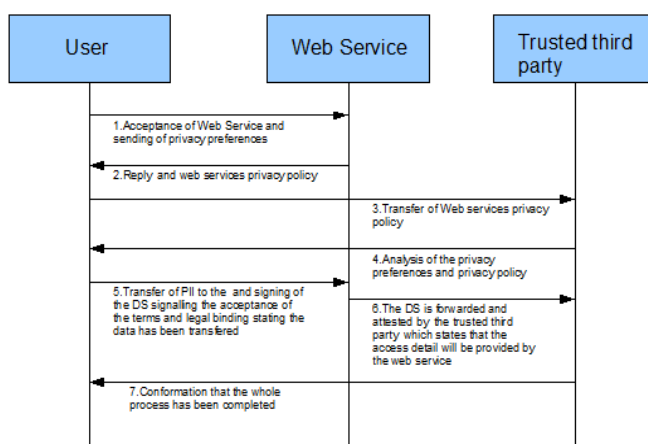


Fig. 3. Overview of the collection and transfer of the PII

3.4 Granting or Denying Access

The form and associated Opt In and Opt Out choices is used to decide whether the access will be granted or not. Any access to the PII is checked and verified as per the agreed terms and a separate audit log is maintained.

3.5 Privacy Insuring Mechanism

The privacy of the user is ensured by an evaluation of the audit logs that contains the access details to the PII. The access details includes the following fields of information

1. Access Purpose-States the purpose of PII access.
2. Access Time
3. Access Fields-The fields of information that were accessed
4. Request By-By whom the access request is generated. This field will store the information uniquely identifying user or the program that accessed the PII.

5. Sub-Entity name-Every employee of the web service is part of certain sub entity which can be the different departments of the company for e.g. Marketing, Finance.

The audit of the the access details is carried by a software agent which is deputed by the third party that was involved in the transfer of PII phase. The access detail of the specific PII is searched by a unique key that was transferred to the third party in the previous steps. The following steps are involved in whole process.

1. The user feels that the PII has been leaked e.g. unsolicited communication ,the user informs the trusted third party.
2. The Trusted third party asks the Web services for access details/logs. This step can occur at regular intervals of time. Inability of the web service to provide access to Trusted Third Party leads to the breach of contract. Hence making the Web Service accountable and need to be compliant .The access of audit Logs can be implemented by the use of Software Agents which is discussed later.
3. The Access detail is provided by the Web service to the Trusted Third Party.
4. The Analysis of the Access detail is done and the terms and conditions agreed by the user and the Web Service and check if any breach occurred. If any breach of terms agreed upon the user and the web service,user is informed and the user is entitled to take Legal remedies.

3.6 Access Classes

The access details reveal the day to day activities of the web service to the trusted third party. Therefore there is a abstraction layer required for hiding which sub entities access the PII's, at the same time not hiding the logs from the trusted third party.

This is implemented by Access Classes which classifies the different sub entities of the Web services into access classes. Classification is done as per the data proliferation risk each sub entity presents for e.g.

1. Least Data Proliferation Risk Class:-This class consists of sub business entities like the Maintenance Department,which requires very frequent access to the data and the risk for the proliferation of an individual PII is very less.
2. Medium Data Proliferation Risk Class:-This class consists of those entities that require less frequent access to the data or require data for purposes that include marketing.
3. High Data Proliferation Risk Class:-This class consists of the entities from where the proliferation of the PII is a major issue. This class should only contain entities that the organization wants to restrict the access of the PII's for e.g. the subsidiaries and other partner companies lie in this class.

The exact sub entity in the access details is mapped to the access class thus protecting the integrity of web services.

3.7 Framework Overview

The whole framework is Consists of the following Sub Modules

1. Privacy Policy Sub-Module:-This sub module states the privacy policy of the organization and helps in stating the Privacy policy which can be converted from

the Internal Privacy Policy as stated by Karjoth et AL (Conversion of EPAL[12] to P3P).This Submodule will produce a P3P document that will state all the data usage terms which the user can compare with his/her privacy preferences. Thus this Sub Module will represent the privacy system to the user when he first accesses the web services and wants to transfer his/her PII.

2. **Digital Signature Storage Sub-Module:-**This Module stores the agreements signed by the user and the trusted Third party with the web service .This component will act as a repository and since the UNICITRAL's Model IT law the Digital Signatures has the same status as a signed contract this protects the web service from any fraud and gives legal recognition to transfer to the PII.
3. **Internal Privacy Enforcement Engine:**This engine is similar to EPAL's Enforcement Engine. It is responsible for the internal enforcement of the privacy policies agreed upon and stated in the privacy policy sub module ,it further consists of
 1. **Authorization Director:-**It parses the policies and authorizes any request to access the protected resource.
 2. **Resource Monitor:-**Its role is to get permission of access to the protected resource and make audit of every request. This component plays an important role in the whole system as it is responsible for the maintenance of the audit logs.
 3. **Policy Management System:-**This system defines all the privacy policies.
 4. **Audit logs:-**An important part stores every access detail. It contains the entity which accessed the data,Purpose for the access, date, etc.
4. **External audit Engine:-**This engine implements the concept of access classes and the access map. And it is responsible for the mapping the access class to the audit logs. Also an agent platform for the Software agents is the part of the submodule. This allows for the Software agents of the third party to access the audit logs.

The Overview of the whole framework is given in Figure 4.

3.8 Hierarchy of Trusted Third Party and Their Regulators

The regulator of the trusted third party plays an important role of monitoring and regulating the trusted third party. Since the trusted third parties play such a critical role in the security and are entrusted with access to audit logs of the web services. The CPI(Controllor for Privacy Insurers) ensure that the third parties are non-biased,and in any case this is compromised,take severe action against it.

Each country has a CPI and all the trusted third parties are required to get themselves registered with the CPI in which country they have a presence. The role of CPI will be similar to the role of CCA (Controllor of Certifying Authorities) as recommended in UNICITRAL's Model IT law and implemented in Information Technology Act 2000 Chapter 6[13] .

Since the laws and rules are in place the role of the CPI can be assigned to the CCA. The control structure is given below in Figure 5.

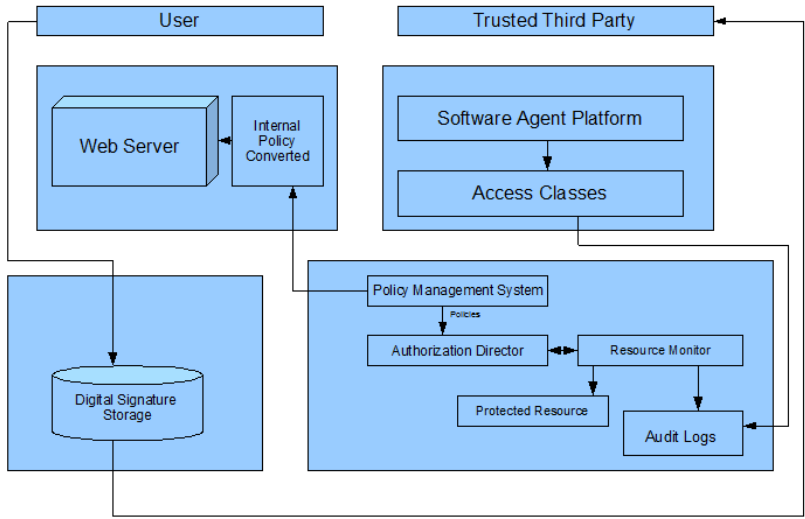


Fig. 4. Overview of the framework

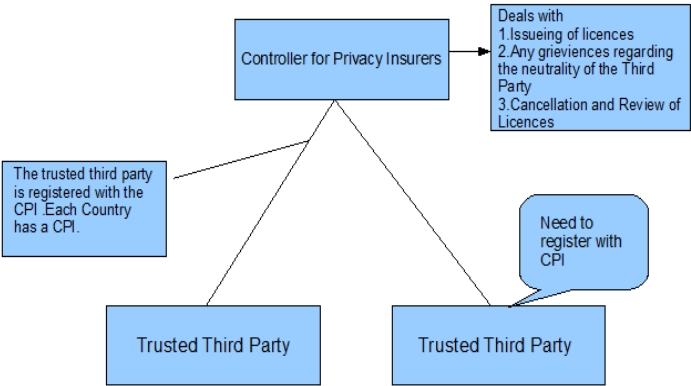


Fig. 5. Hierarchy of CPI and Trusted Third Parties

3.9 Advantages and Disadvantages of the Proposed Privacy Framework

The proposed privacy framework provides the following advantages

- 1. Third party evaluation-the third party plays an active role in auditing and evaluating the users PII usage.
- 2. Ability of the user to monitor PII use.
- 3. Legal Binding of transfer of PII.
- 4. Availability of legal remedies to address the problems of data proliferation by web services
- 5. Legal Framework of CPI and trusted third parties.

6. Integrity of the Web Service protected through Access Classes.

Disadvantages of the Proposed Privacy Policy

1. Complex technologies involved for example Software Agents .
2. Uses Data Intensive auditing to enforce PII's integrity.

4 Implementation Details

The whole system is implemented using Microsoft Visual Studio 2005 and the database used is Microsoft SQL Server 2005. The object oriented programming methodology is used in the implementation of the model. The policies are written using EPAL vocabulary. EPAL to P3P converter [12] is used to convert the EPAL policy to P3P and host it on the web service. The database is used to store all the data and users preferences. Access map is also in the form of a XML file. The Software Agent platform has not been implemented.

5 Future Scope and Conclusion

The proposed privacy platform solves many of the problems identified (Trusted Third Party, Legal Recognition to PII). The whole problem of privacy of PII will take even more and more significance, as the Internet penetration increases. The problems of privacy will be magnified in such a scenario, therefore the concept of third party is bound to make comebacks as it provides impartiality. The framework will also compel the web services to manage the PII in a better and secure manner.

Future scope lies in developing the Software Agent Platform and the standardization of the framework. The concept of Access Classes can be further refined and extended.

References

1. Awad, N.F., Krishnan, M.S.: The Personalization Privacy Paradox: An Empirical Evaluation of Information Transparency and the Willingness to be Profiled Online for Personalization. *MIS Quarterly* 1(30), 13–28 (2006)
2. Federal Trade Commission. Privacy online: A report to congress, <http://www.ftc.gov/reports/privacy3/>. 1998
3. Privacy Leadership Initiative. Privacy Notices Research Final Results. Conducted by Harris Interactive (December 2001), <http://www.ftc.gov/bcp/workshops/glb/supporting/harris%20>
4. Antón, A.I., Earp, J.B., Bolchini, D., He, Q., Jensen, C., Stufflebeam, W.: The Lack of Clarity in Financial Privacy Policies and the Need for Standardization. *IEEE Security & Privacy* (2004), <http://ieeexplore.ieee.org>
5. W3C: The Platform for Privacy Preferences 1.0 (P3P1.0) Specification (2002)
6. Cranor, L.F., Arjula, M., Guduru, P.: Use of a P3P User Agent by Early Adopters Workshop on Privacy In The Electronic Society. In: *Proceedings of the 2002 ACM workshop on Privacy* (2002)
7. Web privacy with P3P LF Cranor (2002) ISBN 81-7366-521-4

8. An Assessment of P3P and Internet Privacy, EPIC (Electronic Privacy Information Center) (June 2000)
9. Cranor, L.F., Egelman, S., Sheng, S., McDonald, A.M., Chowdhury, A.: P3P deployment on websites. In: Electronic Commerce Research and Applications. Elsevier, Amsterdam (2008)
10. Karjoth, G., Schunter, M., Waidner, M.: Platform for enterprise privacy practices: Privacy-enabled management of customer data. In: Dingledine, R., Syverson, P.F. (eds.) PET 2002. LNCS, vol. 2482, pp. 69–84. Springer, Heidelberg (2003)
11. (EPAL1.1) Specification. IBM Research Report, <http://www.zurich.ibm.com/security/enterprise-privacy/epal>
12. EPAL to P3P converter, <http://sourceforge.net/projects/policyconverter>
13. Sharma, V.: Information Technology Law and Practice Law of emerging Technology. Cyber Law and E-commerce (2007) ISBN-978-81-7534-619-2