Privacy Policy Compliance for Web Services

Yee, George; Korba, Larry

National Research Council Canada
Conseil national de recherches Canada

Institute for Information Technology
Institut de technologie de l'information

NRC · CNRC

Privacy Policy Compliance for Web Services *

Yee, G., and Korba, L.
July 2004

* published in Proceedings of the IEEE International Conference on Web Services (ICWS 2004). San Diego, California, USA. July 6-9, 2004. NRC 46566.

Canada

Privacy Policy Compliance for Web Services[1]

George Yee and Larry Korba

Institute for Information Technology
National Research Council Canada
1200 Montreal Road, Building M-50
Ottawa, Ontario, Canada K1A 0R6
{George.Yee, Larry.Korba}@nrc-cnrc.gc.ca

Abstract

The growth of the Internet has been accompanied by the growth of web services (e.g. e-commerce, e-health). This proliferation of web services and the increasing regulatory and legal requirements for personal privacy have fueled the need to protect the personal privacy of web service users. We advocate a privacy policy negotiation approach to protecting personal privacy [1,2]. We provided semi-automated approaches for deriving personal privacy policies in [3]. However, it is evident that approaches are also needed to ensure that providers of web services comply with the privacy policies of service users. In this paper, we examine privacy legislation to derive requirements for privacy policy compliance systems. We then propose an architecture for a privacy policy compliance system that satisfies the requirements and discuss the strengths and weaknesses of our proposed architecture.

1 Introduction

An avalanche of web services targeting consumers has accompanied the rapid growth of the Internet. Web services are available for banking, shopping, learning, healthcare, and Government Online. However, each of these services requires a consumer's personal information in one form or another. This leads to concerns over privacy.

In order for web services to be successful, privacy must be protected. An effective and flexible way of protecting privacy is to manage it using privacy policies. Where the privacy policy of a web service consumer conflicts with the privacy policy of a web service provider, we have advocated a negotiations approach to resolve the conflict [1,2]. We have provided two semi-automated approaches for consumers to create personal privacy policies [3]. We turn now to the problem of privacy policy compliance. Assuming that the service provider has agreed to uphold the consumer's privacy policy, how can the consumer be assured that the provider does indeed comply with the policy? A promising approach is to give the consumer a measure of control over her private information through the use of a Privacy Policy Compliance System (PPCS). In this paper, we derive the requirements for such a system by looking at privacy legislation. We then propose an architecture for a PPCS that satisfies the requirements and discuss the strengths and weaknesses of the architecture.

IBM's definition of web services [4] is "web services are self-contained, modular applications that can be described, published, located, and invoked over a network, generally, the World Wide Web." We concur with this definition but we would add the following: a) the World Wide Web is today's platform of choice, but web services can evolve or be adapted to other platforms, b) emerging web services employ XML (eXtensible Markup Language), WSDL (Web Service Definition Language), SOAP (Simple Object Access Protocol), and UDDI (Universal Description, Discovery, and Integration) [4] but we also include as web services previous generations of web-based applications that involve web browsers interacting with web servers that do not employ XML, WSDL, SOAP or UDDI. For the purposes of this paper, it is not necessary to consider the details of service operation. Our approach for privacy policy compliance is applicable to all web services.

We now explain how privacy policies are employed to protect consumer privacy. A provider has a privacy policy stating what private information it requires from a consumer and how the information will be used. A consumer has a privacy policy stating what private information she is willing to share, with whom it may be shared, and under what circumstances it may be shared. An entity that is both a provider and a consumer has separate privacy policies for these two roles. A privacy policy is attached to a software agent that acts for a consumer or a provider as the case may be. Prior to the activation of a particular web service, the agent for the consumer and the agent for the provider undergo a privacy policy exchange, in which the policies are examined for compatibility. The web service is only activated if the policies are compatible (i.e. there are no conflicts), in which case we say that there is a "match" between the two policies. If service is initiated, the provider is expected to comply with the consumer's privacy policy.

Privacy polices may of course be applied to other systems as a way of expressing privacy preferences (e.g. manually when one visits one's dentist). However, they are especially applicable to web services, not only due to the need to preserve privacy, but also due to the existence of web services compatible technology based on XML, that can be used to express privacy policies (e.g. APPEL [15]).

Section 2 derives requirements for a PPCS by examining privacy legislation. Section 3 presents an architecture for a PPCS that satisfies the requirements of Section 2 and reviews related works in the literature. Section 4 gives our conclusions and directions for future work.

2 Requirements for Privacy Policy Compliance Systems

2.1 Privacy Legislation

To protect consumer privacy, legislative bodies in many countries have enacted legislation that define personal information and spell out the obligations of the service provider with respect to consumer privacy. In Canada, privacy legislation is enacted in the *Personal Information and Electronic Documents Act* [5] and is based on the Canadian Standards Association's *Model Code for the Protection of Personal Information* [6] recognized as a national standard in 1996. This Code consists of ten Privacy Principles [6] that for convenience, we label as CSAPP. Data privacy in the European Union is governed by a very comprehensive set of regulations called the Data Protection Directive [7]. In the United States, privacy protection is achieved through a patchwork of legislation at the federal and state levels. However, privacy has been recognized as a constitutional right and there exists a highly developed system of privacy protection under tort law for the past century [8]. The CSAPP (Table 1) is representative of principles behind privacy legislation in many countries, including the European Union. We will examine it to obtain requirements for PPCSs.

Table 1. CSAPP - The Ten Privacy Principles from the Canadian Standards Association [6]

Principle	***Description***
1. Accountability	An organization is responsible for personal information under its control and shall designate an individual or individuals accountable for the organization's compliance with the privacy principles.
2. Identifying Purposes	The purposes for which personal information is collected shall be identified by the organization at or before the time the information is collected.
3. Consent	The knowledge and consent of the individual are required for the collection, use or disclosure of personal information, except when inappropriate.
4. Limiting Collection	The collection of personal information shall be limited to that which is necessary for the purposes identified by the organization. Information shall be collected by fair and lawful means.
5. Limiting Use, Disclosure, and Retention	Personal information shall not be used or disclosed for purposes other than those for which it was collected, except with the consent of the individual or as required by the law. In addition, personal information shall be retained only as long as necessary for fulfillment of those purposes.

6. Accuracy	Personal information shall be as accurate, complete, and up-to-date as is necessary for the purposes for which it is to be used.
7. Safeguards	Security safeguards appropriate to the sensitivity of the information shall be used to protect personal information.
8. Openness	An organization shall make readily available to individuals specific information about its policies and practices relating to the management of personal information.
9. Individual Access	Upon request, an individual shall be informed of the existence, use and disclosure of his or her personal information and shall be given access to that information. An individual shall be able to challenge the accuracy and completeness of the information and have it amended as appropriate.
10. Challenging Compliance	An individual shall be able to address a challenge concerning compliance with the above principles to the designated individual or individuals accountable for the organization's compliance.

2.2 Attributes of Consumer Private Information

In Table 1, we interpret "organization" as "provider" and "individual" as "consumer". In the following, we use CSAPP.n to denote Principle n of CSAPP. Principle CSAPP.2 implies that there could be different providers requesting the information, thus implying a *collector* attribute. Since on a provider policy, the collector is always the provider, *collector* is only used in consumer privacy policies. Principle CSAPP.4 implies that there is a *what* attribute, i.e. what private information is being collected. Principles CSAPP.2, CSAPP.4, and CSAPP.5 state that there are *purposes* for which the private information is being collected. Principles CSAPP.3, CSAPP.5 and CSAPP.9 imply that the private information can be disclosed to other parties, giving a *disclose-to* attribute. Principle CSAPP.5 implies a *retention time* attribute for the retention of private information. Thus, from the CSAPP we derive 5 attributes of consumer private information, namely *collector*, *what*, *purposes*, *retention time*, and *disclose-to*.

Based on the above examination of CSAPP, the contents of a privacy policy should, for each item of private data, identify a) *collector* - who wishes to collect the information (for consumer policies only), b) *what* - the nature of the information, c) *purposes* - the purposes for which the information is being collected, d) *retention time* – the amount of time for the provider to keep the information, and e) *disclose-to* – the parties to whom the information will be disclosed. A privacy policy can be considered as a machine-readable document that lists each item of private information with corresponding description of *collector (for consumer policies only)*, *what*, *purposes*, *retention time*, and *disclose-to* [3].

2.3 Requirements for Privacy Policy Compliance Systems

The Privacy Principles also prescribe certain operational requirements that must be satisfied between provider and consumer, such as identifying purpose and consent. Some of these requirements lead directly to PPCS requirements. Principle CSAPP.1 can be satisfied by the provider clearly displaying the name(s) and contact information for the individual(s), called Privacy Compliance Officer(s), accountable for compliance on its web page. Principles CSAPP.2 and CSAPP.3 are automatically satisfied by our use of privacy policies, including the exchange of privacy policies between consumer and provider. For example, consider CSAPP.3. Since the consumer is in control of her privacy policy, the matching of this policy with the provider's privacy policy implies the consumer's knowledge and consent for the ensuing collection, use, and disclosure of the consumer's private information. Principles CSAPP.4, CSAPP.5, CSAPP.6, CSAPP.7, CSAPP.8, CSAPP.9, and CSAPP.10 are satisfied by the provider's PPCS. They lead directly to requirements for a PPCS, as follows (we discuss CSAPP.7 at the end):

- CSAPP.4, Limiting Collection: for each purpose for which private information is collected, the PPCS must provide consumers with an

explanation of what information is necessary in order to accomplish the purpose; this explanation must be open and retrievable by the general Internet community for scrutiny (to ensure that providers do not request information beyond what is necessary for the stated purpose); furthermore, for each purpose, the collection of private information must be securely logged and this log must be available to the owner of the private information or her designate for examination (to ensure that data is collected by fair and legal means).

- CSAPP.5, Limiting Use, Disclosure, and Retention: for each purpose for which private information is collected, the PPCS must provide consumers with an explanation of how it intends to use or disclose the consumer's private data; this explanation must be open and retrievable by the general Internet community for scrutiny; furthermore, for each purpose, the use and disclosure of the consumer's private data must be securely logged and this log must be available to the owner of the private data or her designate for examination and comparison against the previous explanation of use and disclosure (to ensure that providers do not use the consumer's private information for other than the stated purpose). In addition, the PPCS must ensure that all copies (including copies disclosed to other parties) of the consumer's private information is deleted at the earliest of a) the time when the data is no longer needed for the fulfillment of the purpose, or b) the expiration of the data's retention time. This deletion must also be securely logged and the log accessible by the owner of the private information or her designate.

- CSAPP.6, Accuracy: the PPCS must provide a facility with which consumers can access, check the accuracy, update, and add to their private data, as necessary for the corresponding purposes. These actions should also be securely logged and accessible to the provider or the data owner for verification purposes.

- CSAPP.8, Openness: upon request, the PPCS must display the provider's specific information about its policies and practices relating to the management of private information.

- CSAPP.9, Individual Access: upon a consumer's request, the PPCS must inform the consumer of the existence, use, and disclosure of her personal information, and give her access to that information; upon review of the information, the consumer can perform the actions of CSAPP.6.

- CSAPP.10, Challenging Compliance: upon request, the PPCS must allow the consumer or her designate to review the secure log to verify compliance to her privacy policy. In case of non-compliance, the consumer can take action outside the scope of the PPCS, i.e. notify the provider's Privacy Compliance Officer(s) of the non-compliance and take legal action if necessary.

- CSAPP.7, Safeguards: it is apparent from the above that the PPCS contains:
  a) the provider's explanations of what private data it requires for particular purposes,
  b) the provider's explanations of how it uses or discloses private data for particular purposes,
  c) the provider's specific information about its policies and practices relating to the management of private information,
  d) the provider's privacy policies,
  e) the consumer's privacy policies,
  f) the consumer's private data,
  g) the log entries.

  The PPCS needs to apply the following protection to these information groups: groups a), b), c), and d) can be viewed by anyone in the Internet community but need to be protected from unauthorized tampering; groups e) and f) must be viewable only by the provider, the party receiving the private information as a disclosure (view only the information disclosed and corresponding privacy policy), and the consumer owner of the private information; groups e) and f) can only be modified, deleted, or added-to by the consumer owner of the private information, except for deletion, where the provider or the party receiving the information as a disclosure can delete the information, either because the corresponding purpose has been accomplished, or the information's retention time has expired; group g) must be viewable only by the consumer owner of the corresponding private information, the consumer owner's designate, the provider, or the party receiving a disclosure of corresponding

private information; group g) information once written by the PPCS, must not be modifiable by any party. Storage and transfer of the data referred to above will be access-controlled and use cryptographic techniques to protect data integrity and limit the release of the information.

3 An Architecture for Privacy Policy Compliance Systems

Figure 1 presents an architecture for a PPCS that satisfies the requirements of Section 2.3.

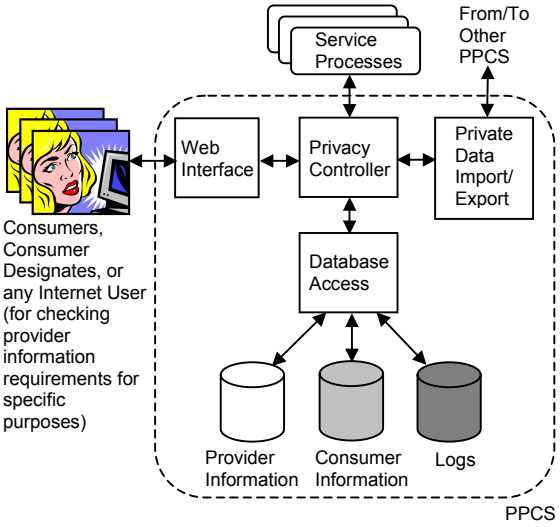

Figure 1. Privacy policy compliance system architecture

Descriptions of the architecture components in Figure 1 follow:

- Web Interface: web user interface for interactions with the consumer, consumer designate, or any Internet user (for checking provider information requirements for specific purposes); specific actions include: a) provides interface for user access to update private information or to examine logs, b) upon request, displays provider information regarding names and contact information for Privacy Compliance Officers, provider specific policies on the management of private information, and provider explanations of what information is required for various purposes as well as how the private information will be used, c) establishes a secure channel to the consumer or consumer delegate and authenticates them (see Section 3.1).

- Privacy Controller: controls the flow of provider and consumer information and requests to fulfill the consumer's privacy policy; specific actions include: a) make log entries, b) delete private information upon completion of purpose or information expiry, c) grant access for consumer update of private information (including the update of information that has been provided to third party data processors), d) grant access for the examination of logs and comparisons of information, e) upon request, inform the consumer of the existence, use, and disclosure of her private information.

- Database Access: provides read/write access to the databases as requested by the Privacy Controller, handles write protection for the Provider Information database, handles data encryption/decryption for the Consumer Information database, and read/write protection for the Logs database (see Section 3.1) to meet the requirements expressed in CSAPP.7.

- Private Data Import/Export: sends private information disclosures to other providers; receives private information disclosures from other providers, sets up secure channel to other providers for sending information disclosures, authenticates these providers.

- Provider Information Database: contains provider information items a) to d) inclusive as given in the CSAPP.7 bullet of Section 2.3.

- Consumer Information Database: contains consumer information items e) and f) as given in the CSAPP.7 bullet of Section 2.3; segmented for each consumer.

- Logs Database: contains log entries for PPCS-consumer actions such as information collection, information use and disclosure, information access and update, information deletion; segmented for each consumer.

- Service Processes: represent the services offered by the provider; the arrow going out of these processes represents private information collected by the services; the arrow going in to these processes represents private information required to carry out the services.

We need to clarify how parties who have received private information disclosures can be expected to delete the information upon completion of purpose or information expiry. Such parties are considered to be subcontractor providers of the first provider and provide services to the first provider that are needed to complete the purposes of the first provider. In this case, the first provider is actually a consumer. As a consumer, the first provider has negotiated a consumer privacy policy with each subcontractor provider, containing the required purposes and information retention times reflecting the wishes of the original consumer. The PPCS of each subcontractor provider then deletes the original consumer's private information upon completion of the purposes in the privacy policy agreed with the first provider or upon information expiry.

3.1 Security

Table 2 identifies security requirements and implementations for the above PPCS architecture. Standard protection such as firewalls and intrusion detection systems are assumed in place. Although we have not specified it in Table 2, some consumers may wish to be anonymous, requiring authentication through blind certificates.

Table 2. Security requirements and implementations for the proposed PPCS architecture

Architecture Component or Location	**Security Requirement**	**Security Implementation**
Database: Provider Information	Write Protection	Operating System Directory Protection (e.g. Linux)
Database: Consumer Information	Read Protection	Public Key Encryption / Decryption (e.g. RSA) in conjunction with SSL
Database: Logs	Read /Write Protection	Operating System Directory Protection (e.g. Linux)
Communication Channel: To Consumer, Consumer Designate, or any Internet User	Secure Channel and 2-way authentication for Consumer or Consumer Designate	SSL for secure channel and authentication of provider; digital certificate to authenticate consumer or consumer designate
Communication Channel: To Other PPCS	Secure Channel and 2-way authentication	SSL for secure channel and authentication of providers at both ends of the channel

3.2 Discussion of Strengths and Weaknesses

The strengths of the proposed architecture include:

- The provider's explanations of what information it requires for specific purposes as well as how the information will be used and disclosed is open to scrutiny by the entire Internet community, giving assurance that the provider is honest.
- Private information deletion by parties receiving disclosures is handled simply and elegantly.
- The consumer can verify privacy policy compliance by accessing a secure log. This gives the consumer first hand assurance of compliance, which is a psychologically higher level of assurance than having the consumer rely on automatic or programmed compliance.

The weaknesses of the proposed architecture include:

- Lack of scalability. The PPCS could be overwhelmed if the number of consumers is very large. A possible solution is for the provider to load share the consumers among a number of PPCSs. This load sharing could be based on geography (where the consumer lives) or on the volume of consumer business.
- Consumers may not bother or lack the know-how to check the secure logs for compliance. In this case, there may be a business opportunity for Internet firms such as Certificate Authorities to offer consumers a compliance verification service.
- A malicious provider may tamper with its PPCS so that fallacious logs are recorded. First, PPCSs may need to be standardized and certified by a

privacy authority (e.g. privacy commissioner belonging to a province or state). Second, critical PPCS components may be made tamperproof by incorporating them in hardware.

- Providers may not want to install PPCSs due to the costs. In this case, consumers can choose to do business with providers that do have installed PPCSs. Such providers would have a higher reputation and attract more customers. Eventually the providers that don't have PPCSs will realize that it's a cost of doing business and come on board. In some jurisdictions, the law may require using something like a PPCS, which may be operated by data protection authorities, or their representatives.

3.3 Implementation

We are leaving the implementation of the above architecture to future work. Each service provider is expected to offer a PPCS for the service(s) that it provides. The PPCS may be one that is implemented on the provider's premises for its sole use or one that is provided by a PPCS service provider (e.g. data protection authority) for use by multiple providers, whose services may be individually too lightweight (either in size or number of customers) to justify the cost of maintaining a PPCS. Perhaps provision of the PPCS by a service provider that is a data protection authority is the better approach, since that would answer some of the weaknesses noted in Section 3.2 and would undoubtedly result in a higher level of consumer confidence regarding privacy policy compliance.

3.4 Related Work

The closest related work is [9] where the authors proposed similarities between a system for digital rights management and a system for privacy rights management. The authors go on to examine the feasibility of turning a digital rights system into a privacy rights system. We did not find any other work in the literature dealing specifically with privacy policy compliance. However, we found works on security policy compliance or general e-contract enforcement. These works (e.g. [10,11,12,13,14]) differ mainly from ours in that they deal with the enforcement of complex security policies or business contracts that require automatic program verification of rules expressed in a suitable language - we deal with simpler privacy policies with enforcement via secure logs and legal recourse. We believe our approach is better for privacy policies since privacy is more personal and people are more inclined to verify compliance personally.

4 Conclusions and Future Work

We began by examining representative privacy legislation to derive requirements for privacy policy compliance systems. This ensured that the resulting requirements are core to any PPCS. We then presented an architecture that satisfies the requirements and discussed its strengths and weaknesses. Web services can only succeed if consumers are confident that their privacy is protected. PPCSs are essential for giving consumers this confidence. As future work, we plan to realize our proposed architecture in a prototype to explore any potential usability and performance issues.

5 References

[1] G. Yee, L. Korba, "Bilateral E-services Negotiation Under Uncertainty", Proceedings, The 2003 International Symposium on Applications and the Internet (SAINT2003), Orlando, Florida, Jan. 27-31, 2003.

[2] G. Yee, L. Korba, "The Negotiation of Privacy Policies in Distance Education", Proceedings, 14th IRMA International Conference, Philadelphia, Pennsylvania, May 18-21, 2003.

[3] G. Yee, L. Korba, "Semi-Automated Derivation of Personal Privacy Policies", Proceedings, 15th IRMA International Conference, New Orleans, Louisiana, May 23-26, 2004.

[4] M. O'Neill et al, Web Services Security, McGraw-Hill/Osborne, 2003.

[5] Department of Justice, Privacy Provisions Highlights, http://canada.justice.gc.ca/en/news/nr/1998/attback2.html

[6] Canadian Standards Association, "Model Code for the Protection of Personal Information", retrieved Sept. 5, 2003 from: http://www.csa.ca/standards/privacy/code/Default.asp?articleID=5286&language=English

[7] European Union, "Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data",

unofficial text retrieved Sept. 5, 2003 from: http://aspe.hhs.gov/datacncl/eudirect.htm

[8] Industry Canada, "Privacy and the Information Highway, Regulatory Options for Canada", chapter 6, retrieved Sept. 5, 2003 from: http://strategis.ic.gc.ca/SSG/ca00257e.html#6

[9] S. Kenny and L. Korba, "Adapting Digital Rights Management to Privacy Rights Management", Computers & Security, Vol. 21, No. 7, November 2002, 648-664.

[10] D.K.W. Chiu et al, "A Three-Layer Architecture for E-Contract Enforcement in an E-Service Envronment", Proceedings of the 36th Hawaii International Conference on System Science (HICSS'03), 2002.

[11] X. Ao et al, "A Hierarchical Policy Specification Language, and Enforcement Mechanism, for Governing Digital Enterprises", Proceedings of the Third International Workshop on Policies for Distributed Systems and Networks (POLICY'02), 2002.

[12] P. McDaniel and A. Prakash, "A Flexible Architecture for Security Policy Enforcement", Proceedings of the DARPA Information Survivability Conference and Exposition (DISCEX'03), 2003.

[13] I. Kao and R. Chow, "Enforcing Complex Security Policies for Commercial Applications", Proceedings of the Nineteenth Annual International Computer Software and Applications Conference, 1995.

[14] J. Burns et al, "Automatic Management of Network Security Policy", Proceedings of the DARPA Information Survivability Conference & Exposition II (DISCEX '01), Volume 2, 2001.

[15] W3C, "A P3P Preference Exchange Language 1.0 (APPEL 1.0)", retrieved April 22, 2004 at: http://www.w3.org/TR/P3P-preferences/

[1] NRC Paper Number: NRC 46566