Usable Security and Privacy: A Case Study of Developing Privacy Management Tools

Carolyn. Brodie, Clare-Marie Karat, John Karat
IBM T. J. Watson Research Center
19 Skyline Drive
Hawthorne, NY 10532
1-914-784-7237
cbrodie, ckarat, jkarat@us.ibm.com

Jinjuan Feng
University of Maryland Baltimore County, Information Systems Dept.
1000 Hilltop Circle
Baltimore, MD 21250
1-410-455-3888
Jfeng2@umbc.edu

ABSTRACT

Privacy is a concept which received relatively little attention during the rapid growth and spread of information technology through the 1980's and 1990's. Design to make information easily accessible, without particular attention to issues such as whether an individual had a desire or right to control access to and use of particular information was seen as the more pressing goal. We believe that there will be an increasing awareness of a fundamental need to address privacy concerns in information technology, and that doing so will require an understanding of policies that govern information use as well as the development of technologies that can implement such policies. The research reported here describes our efforts to design a privacy management workbench which facilitates privacy policy authoring, implementation, and compliance monitoring. This case study highlights the work of identifying organizational privacy requirements, analyzing existing technology, on-going research to identify approaches that address these requirements, and iteratively designing and validating a prototype with target users for flexible privacy technologies.

Categories and Subject Descriptors

H5.2. Information interfaces and presentation: User Interfaces. K4.1. Public policy issues: Privacy

General Terms

Management, Design, Experimentation, Security, Human Factors.

Keywords

Privacy, privacy policies, security, social and legal issues, design.

Symposium On Usable Privacy and Security (SOUPS) 2005, July 6-8, 2005, Pittsburgh, PA, USA.

1. INTRODUCTION

As organizations come to rely on the collection and use of personal information in order to provide quality services to their customers, patients, constituents and return on investment to their share holders, the ability to protect that information and enforce privacy polices becomes more important. The increasing number of reports of privacy violations due to external break-ins as well as accidental and malicious misuse of personal information by individuals within an organization is only exacerbating the problem. While an increasing amount of research concentrates on identifying security and privacy weaknesses and how to address them, making this technology usable remains an important issue. The Computing Research Association (CRA) Conference on Grand Research Challenges in Information Security and Assurance has identified the ability to "give end-users security controls they can understand and privacy they can control for the dynamic, pervasive computing environments of the future" as a major research challenge [14].

As Whitten and Tygar [35] point out, "security mechanisms are only effective when used correctly" and this is often not the case due to usability issues with security software. In this paper we present the design of a set of privacy utilities that are intended to assist organizations with the creation, implementation, and internal auditing of privacy policies. We will discuss how we used knowledge of organizational user needs gained during an earlier phase of the project as the basis for analysis of current privacy technologies and on-going research to create an abstract architecture for an organizational privacy solution. Then based on that architecture, we have designed and prototyped a privacy management workbench to assist organizations in creating and managing their privacy policies.

We chose the domain of organizational privacy policy creation and enforcement because use and misuse of personal information (PI) is an area of increasing concern in many geographies and domains around the world. Organizations need usable methods to ensure that the information policies they put in place are enforced correctly without negatively affecting their business processes. Research has shown that many invasions of privacy are not intentional [1]. When designing systems that use personal information, we must not only secure them so that information cannot be accessed by unauthorized users but also from authorized users for unauthorized purposes.

Privacy can mean many things to many people. In the context of our research and this paper, we define privacy as the right of an individual to control information about themselves rather than as the right to individual isolation [27, 30]. The OECD principles [30] provide high level privacy standards for dealing with personal information and have widespread consensus. These have provided input to legislation in many parts of the world that requires organizations to have privacy policies and constrains organizational collection and use of personal information to differing degrees. This legislation varies by both geography and domain [24]. These variations as well as the inherent differences between domains [10,11] and between the business practices of different organizations means that it is not likely that a single privacy policy can be created to cover all personal information. The research reported in this paper has focused on how technology can be used by organizations to create and enforce the range of privacy policies needed to meet the varied requirements.

2. RELATED WORK

There are many aspects of privacy that have been the subject of research, including research on the public perceptions of the need to protect PI, research and development of many types of privacy preserving technologies, as well as research into the current approaches that are being used by organizations to protect the PI of their customers, constituents, patients, and employees. In this section we will discuss recent research into the public perceptions of privacy within organizations and how they affect individual willingness to share data, technological approaches for enforcing privacy policies, and finally how organizations are protecting PI today.

Research has identified high levels of consumer concerns regarding privacy [17, 18, 31] in a large range of geographies and domains. A multi-national consumer privacy survey in 1999 investigated US, German, and UK consumers' attitudes toward privacy in different industries [18]. Seventy-eight percent of the people in the survey reported that they have refused to provide information in the past due to concerns about PI misuse. A privacy and business survey in 2000 conducted for the Australian government revealed that 95% of the respondents think it is necessary to implement laws to protect PI and also documented that approximately 50% of the respondents routinely and intentionally provide inaccurate PI [31]. A more recent Forrester report found that 97% of North American consumers believe that online privacy concerns are real and 94% reported that they believe the benefits they receive for sharing personal information do not outweigh their concerns [13]. In the health care domain, physicians and practitioners are concerned about serious threats to patient privacy due to information gathering methods, record accuracy and access, and unauthorized secondary use [11]. In the education sector, a Stanford University report reveals that PI is not effectively protected [34].

Researchers have responded to these concerns through the development and analysis of machine readable privacy policies and the development of mechanisms for helping end-users to understand the policies and organizations to enforce the policies. One area of research is on the development and use of machine readable privacy policy schemas for enabling privacy functionality. P3P [15] is one of the first privacy policy languages that has been standardized by an international standards body, the W3C. P3P is an XML based language that allows organizations with Websites to create machine readable versions of their privacy policies. Generally, P3P allows organizations to specify rules that contain the type of data, the type of use, the user of the data, the purpose of the use, and how long the data will be retained. From the end-user or client point of view, automated agents, such as the AT&T Privacy Bird [8] and browsers such as Microsoft's Internet Explorer [26] can use the P3P policies to provide individual users with the ability to quickly determine if the Website's privacy policies match their privacy preferences. Other proposed schemas, such as APPEL, have expanded on the goal of helping individuals to quickly determine if a Website's policies match their preferences by allowing the user to define rule sets for describing acceptable organizational privacy policies [36].

While the ability to quickly understand a site's privacy policy and determine if the site conforms to their preferences is helpful to end-users, it is important to understand that there is no guarantee that the policy is actually implemented as specified within the organization. This fact has lead to research into how machine readable (XML schema languages) privacy policies can be used by organizations to enforce policies. Karjoth and Schunter [22] analyzed how enterprise privacy policies differ from security policies and how well P3P can express an enterprise privacy policy. Based on this analysis, they propose a privacy policy model that can be used for internal access control within an enterprise. New XML schemas designed to enforce privacy policies include, the Enterprise Privacy Authorization Language (EPAL) [7] and XACML with a privacy profile [29]. These allow more expressive policies that include hierarchical policy elements, conditions on rules, and a user definable set of obligations. EPAL is being considered by the W3C standards body and XACML with a privacy policy profile is being considered by OASIS. The ability to use a language like EPAL to capture and logically enforce the privacy policies of large, complex organizations has been studied and formalized by Backes, Pfitzmann and Schunter [9].

In addition to policy analysis, researchers have been exploring enforcement mechanisms for some time. Anderson [4,5] proposed a security policy model for the British Medical Association that described how to implement and manage compartmented security in health care. In an update in 2000, he reported that it had been implemented successfully in three British Hospitals [5]. Since that time there has been research into how machine readable policies can be used internally by organizations to enforce their privacy policies. Some approaches have concentrated on allowing policies defined by individuals to dictate how their information is used [12], while many others have concentrated on enforcing privacy policies created at the organizational level. An example of this is the Hippocratic Database [3] in which P3P is used to define access rules that are then enforced by the Hippocratic Database. IBM's Tivoli Privacy Manager is another example of an approach that has used P3P to define privacy policies which are then enforced by deploying monitoring software around data stores that sends requests for PI to a server which then determines if the access conforms to the privacy policy and logs both the attempt and the enforcement decision [19].

Even with all of the research that indicates that there is growing concern about privacy issues and the possible technical approaches that have been developed to protect PI, most organizations that depend on the use of personal information in their business processes have done little to implement the policies through technology [21, 33]. Privacy policy enforcement is still often accomplished through predominately manual procedures. According to a 2003 study conducted by Ponemon for the IAPP [32] only 19% of the organizations sampled report that they are currently using any privacy enabling technology. This confirms the situation described by Forrester with respect to privacy [17]. This Forrester report describes differences between consumer and executive views of privacy practices in industry. According to this report, the majority of executives who participated in the study (58%) believe that their companies are doing a good job of addressing privacy issues while customer concerns about privacy remains high. In fact, the majority of executives did not know whether their customers even checked the privacy policies or not and few see the need to enhance their privacy practices. These results were echoed by research in the Asia-Pacific region [31].

More recent research indicates that many organizations recognize that privacy is an issue for them. They currently do not know how to use technology to help them enforce their privacy policies. The Ponemon study [32] reported that although 98% of the companies in their survey have a privacy policy, 52% believe they do not have the resources to adequately protect privacy. Furthermore, most organizations store PI in heterogeneous server system environments and currently they do not have a unified way of defining or implementing privacy policies that encompass data collected and used by both Web and legacy applications across different server platforms [6]. This makes it difficult for organizations to put in place proper management and control of PI, for the data users to access and work with the PI inline with the privacy policies, and for the data subjects to understand rights regarding use of their PI. It has been suggested that one reason that organizations are not employing new privacy enabling technologies to protect PI is that these technologies are currently very difficult to use [14,35]. In practice user-centered design techniques have contributed to the development of some highly usable security systems [20, 37]. Based on this evidence, our emerging focus has been on applying HCI-based research techniques to answering how organizations could create policies, and how technology might be used to enforce the policies and provide audit capabilities to ensure compliance within the organization. We believe that this focus complements the diverse range of privacy research that is being conducted by making privacy technologies accessible to organizations so that technology can enable the protection of privacy and not just be a force which reduces individual rights.

3. PROJECT BACKGROUND

The research presented in this paper builds on our team's previous research in which we identified privacy needs within organizations through email survey questionnaires and then refined the needs through in-depth interviews with privacy-responsible individuals in organizations. A more complete description of this work can be found in [21]. In this research fifty-one individuals who were responsible for either the creation and/or implementation of privacy policies within their organizations responded to an email survey. The participants came from industry and government organizations in North America, Europe, and Asia Pacific. The participants were asked to identify their top privacy concerns, the types of functionality they felt would be valuable to them in addressing these concerns, and what actions their organizations were currently taking to address privacy issues.

We then held in-depth interviews with a subset of thirteen of the survey participants. The goals of these interviews were to build a deeper understanding of the participants' and their organizations' views regarding privacy, their privacy concerns, and the value they perceived in the desired privacy technology they spoke of in the context of scenarios of use involving PI in their organizations. The majority of the interview sessions were centered on discussion of a scenario of use provided by the respondent regarding PI information flow in their organization and follow-up questions related to it. We wanted to identify and understand examples of how PI flowed through business processes in the organization, the strengths and weaknesses of these processes involving PI, which of these processes are automated and which are manual, and the additional privacy functionality they need in the context of these scenarios.

The participants reported that protecting their customers, patients, constituents, and employees PI requires a multifaceted approach. The organization must develop an implementable privacy policy, educate employees and the people they serve on that policy and the importance of privacy in general, identify where PI is stored and used within their business processes, and then develop both manual procedures and technological solutions to enforce the policy they have created. One of the main goals with this research was to help organizations in their efforts by identifying how technology could be used to assist them in protecting the PI they collect and use. Using the survey and interview data that we collected, we developed a set of five key privacy concepts that are important to meeting the needs of organizational users of privacy protecting technologies. They include:

1. It is important to provide users with **one integrated solution for an organization's heterogeneous configuration** even if it consists of a set of utilities that provide users with a similar set of functionality and interaction methods for systems that are implemented differently on different technologies.
2. The **privacy functionality must be separated from the application code** for cost, consistency, and flexibility reasons – users do not want to have to modify all of their applications individually to ensure that PI is protected.
3. There needs to be the ability to **support an appropriate level of granularity** for applying the privacy policy. For example, the ability to control access at the field level in a database.
4. There must be the ability to work with both **structured and unstructured information**. This

includes protecting field level data and handling PI within documents in appropriate ways.

5. There must be **simple and flexible privacy functionality** that is designed to meet the needs of the user community that owns each subtask in the privacy process. For example, CPO's and/or business process owners often write the privacy policies. They must be able to author policies that will end up in machine readable form without having IT skills.

4. Architectural Analysis of Privacy Functionality

Using the set of key design concepts for any privacy solution that we identified in earlier phases of this research, we analyzed existing privacy architectures to identify areas in which user-centered design techniques could be applied to best meet the needs of organizational privacy users. To facilitate the description of this analysis we have created a generalization of many approaches to protecting the privacy of PI which is shown in Figure 1. In this figure a privacy policy authoring utility is used to create privacy policies that are stored in a machine readable format. This machine readable privacy policy is then used by a privacy enforcement mechanism that is positioned between applications and data stored within the organization's configuration. The architecture also dictates that the enforcement mechanism should create a log of privacy events which can be analyzed by the organization's audit mechanism in order to report on compliance with the privacy policy. The generalized architecture drawing in Figure 1 is purposefully abstract so that it can be used to describe the common elements and mechanisms in a variety of possible privacy implementation approaches.

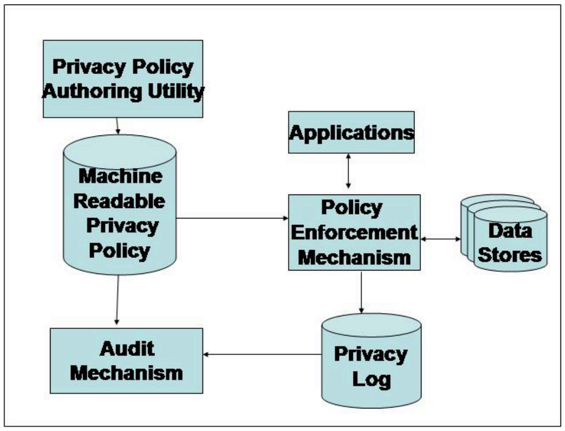

Privacy Policy Authoring Utility
Applications
Machine Readable Privacy Policy
Policy Enforcement Mechanism
Data Stores
Audit Mechanism
Privacy Log

Figure 1. Abstract Privacy Architecture

Different types of machine readable policies have been proposed and are at different points in the standardization process. P3P is currently a privacy language standard and is used to define privacy policies in some approaches [3, 19]. Other standards that allow for more expressive policies, such as EPAL and XACML with a privacy profile are also being considered. Likewise, there are many different approaches to privacy policy enforcement that have been proposed including query re-writing [2], data access monitoring and the use of a rules based enforcement engine [19], and the application of a modified access control mechanism [4,5]. Also, not all approaches in the literature include all components in the drawings. For example, the model proposed by Anderson uses an enforcement mechanism based on the concepts for multilevel security research as well as an audit mechanism, but does not address the use of machine readable privacy languages. While we recognize that each of these types of solutions do have the potential to be valuable to organizational users, we have found that all share some high level strengths and weaknesses in terms of the key privacy design concepts we described in our earlier research [21].

Based on our analysis we found that the technologies that are being researched and developed can be used to meet three of the five key privacy design concepts identified. In considering concept 1, we compared the user scenarios that we collected from the organizations that participated in our interview research and the range of privacy solutions that we found in the literature. We did not find one solution that obviously met all of the users needs for providing a single solution that would protect data within large organizations' highly heterogeneous and widely distributed configurations. Nor does it seem likely that one could be designed anytime soon. However, there are at least two approaches to addressing this problem. One approach is the creation of a common set of privacy utilities that provides users with a single method for creating, visualizing and auditing privacy policies that could then be enforced using the appropriate range of technologies. Another possible approach is for a set of utilities to be provided to a central PI store on a single platform that has a privacy policy enforcement mechanism. This would create a PI "vault". Other distributed applications would then request data from that system.

We recognize that there are privacy enabling technologies that address concepts 2 and 3. Many of the privacy approaches that have been identified allow the privacy enforcement to be separated from the application. For example, the Hippocratic Database [2] allows applications to query the database as they always have. The query re-writing done by the JDBC layer ensures that only PI accesses or updates allowed by the policy occur. Likewise, data store monitoring approaches such as that employed by Tivoli Privacy Manager [19] separate the application from the privacy auditing and/or enforcement. Each of these approaches also has the potential to allow privacy enforcement at the database field level.

Although we found approaches that can address the first three key privacy concepts, we have not found any approach that addresses either of the last two concepts. In the case of concept 4, the representatives of the organizations that we interviewed told us that they needed to be able to provide privacy protection for information within unstructured documents. Perhaps text analytics research combined with a privacy enforcement mechanism may be able to address this need in the future. Finally, while there has been research into the design of interaction methods to allow end users to define privacy policies with their preferences regarding sharing data with e-commerce companies [16] and with pervasive devices [23], none of the privacy technologies we analyzed addressed the last key privacy design concept (concept 5) that we identified. Organizational users have a need for simple and flexible interaction methods for dealing with complex, organizational privacy policies that can be used by individuals who do not have IT skills. Therefore, this is the need that we decided to address in our research. We identified three areas where highly usable privacy utilities were needed. The first is a utility to assist users in creating and understanding privacy policies. The second is a utility to assist users in implementing the privacy policy. The design of this utility is partially dependent on the choice of enforcement engines used. Finally the third utility enables organizations to conduct internal audits of their privacy policies. While our research has focused on all three areas, our work in the privacy policy creation area is the most mature and is the least dependent on a particular enforcement engine. Therefore, we will concentrate on this utility in this paper.

During the survey and interview research, many of the participants indicated that privacy policies in their organizations were created by committees made up of business process specialists, lawyers and security specialists as well as information technologists. Based on the range of skills generally possessed by people with these varied roles, we hypothesized that different methods of defining privacy policies would be necessary. Figure 2 shows the abstract architecture updated with a more detailed privacy policy creation utility. The figure shows the privacy policy creation utility divided into three parts. There is a privacy policy authoring utility that uses and stores natural language policies, a transformation utility for translating the policy into machine readable policies, and a visualization utility for helping users understand the implications of new and existing policies. The architectural view of this utility was used to guide the design of a prototype privacy management tool.

5. Designing and Evaluating a Privacy Policy Prototype

Using the completed survey and interview research and the architectural analysis, we designed and developed a prototype of a privacy policy management workbench called SPARCLE (Server Privacy ARchitecture and CapabiLity Enablement). SPARCLE is written in dynamic HTML and is a "Wizard-Of-Oz" prototype. By this, we mean that the prototype allows users to see how the functionality would operate, but that it is not fully functional. The use of this prototype allowed the team to obtain user feedback on the types of functionality included in the prototype before a fully functional version was developed.

The overall goal in designing SPARCLE was to provide organizations with tools to help them create understandable privacy policies, link their written privacy policies with the implementation of the policy across their IT configurations, and then help them to monitor the enforcement of the policy through internal compliance audits. Once we designed the prototype, we conducted a series of walkthrough sessions in which we utilized the prototype to discuss an appropriate scenario with representatives of health care, government, and finance organizations. In this paper, we will concentrate on the techniques we designed and developed for authoring privacy policies and assisting organizations in understanding the policies that have been created.

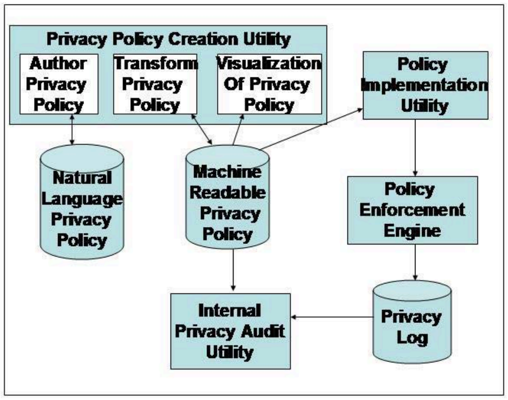

Privacy Policy Creation Utility
Author Privacy Policy
Transform Privacy Policy
Visualization Of Privacy Policy
Policy Implementation Utility
Natural Language Privacy Policy
Machine Readable Privacy Policy
Policy Enforcement Engine
Internal Privacy Audit Utility
Privacy Log

Figure 2. Abstract Privacy Architecture with Privacy Policy Creation Utility Expanded

5.1 Authoring Privacy Policy

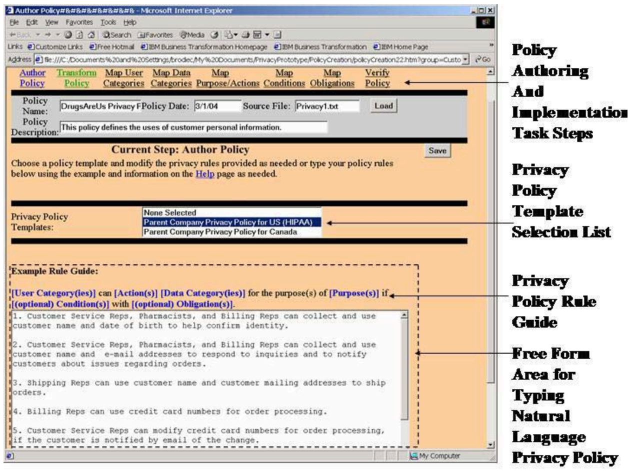
Based on the architectural drawings above and building on research into using natural language processing for policy development [25], SPARCLE was designed to support users with a variety of skills by allowing individuals responsible for the creation of privacy policies to define the policies using natural language or using a structured format to define the elements and rule relationships that are then directly used in the creation of a machine readable version of the policy. SPARCLE keeps the two formats synchronized. For users who prefer authoring with natural language, SPARCLE transforms the policy into a structured form so that the author can review it and then transforms it into a machine readable format such as EPAL [7], XACML [29] or other appropriate privacy languages. SPARCLE translates the policies of organizational users who prefer to author rules using a structured format into both a natural language format and the machine readable version. During the entire privacy policy authoring phase, users can switch between the natural language and structured views of the policy for viewing and editing purposes. Once the machine readable policy is created, it is possible to employ any enforcement engine that is capable of using the elements of the standardized privacy policy language to ensure the policy is enforced for data stored in the organization's on-line data stores.

Figure 3 contains a screen capture of SPARCLE's natural language interface for defining privacy policies. SPARCLE supports a set of privacy tasks that were identified from the data collected using the survey and interview research. The identified tasks include: authoring the policy in natural language (step shown in Figures 3), transforming the policy into policy elements (step shown in Figure 4), mapping the user categories, mapping the data categories, mapping the purposes and actions, mapping the conditions, mapping the obligations, and verifying the policy. The mapping steps are used to associate policy elements with system objects, and enable the separation of high level and detailed policy specification. The verify step allows users to confirm that all parts of the policy have been mapped. In SPARCLE these tasks are represented by the tabs shown at the top of Figure 3. The page also contains general information about the policy, (the name, date created, and file source of the policy, and a description of the policy authoring task to be performed) a list of privacy policy templates that could be either provided by the tool for particular domains and geographies based on laws or created by the organization for customization and use by its divisions, and an Example Rule Guide describing the elements that make up a privacy policy rule. The guide is based on analyses of privacy policy rules specified in [7].

The guide defines the basic components that are necessary in an enforceable privacy policy rule including user categories, allowed actions, data categories, purposes, as well as optional components such as conditions and obligations. Finally, a text entry area is provided for the actual privacy policy. When the user begins the process of creating a new policy, she can create the policy from scratch by typing into the text entry area, copying an existing policy into that area, or selecting one of the templates provided and modifying it.

When the author is satisfied with the policy, he clicks on the save button shown in Figure 3. This causes the text policy to be passed to a shallow parser [28]. The natural language policy is analyzed using the shallow parser and an associated dictionary in order to identify the policy elements (the strings which describe the User Categories, Actions, Data Categories, Purposes, Conditions, and Obligations) in each rule. Then when the user chooses the Transform Policy tab (shown in Figure 3), the natural language entry field area is replaced with a structured privacy policy creation view (shown in Figure 4). The page also contains the policy information and the list of policy templates that was available on the policy authoring page. Next, the user is provided with a list containing the parsed rules from the current policy.

Whenever a parsed rule is selected in the transformed view, the original unparsed text is also displayed and the elements of the rule that have been identified are highlighted in individual policy element selection lists as shown in Figure 4. There is one policy element selection list for each of the 6 types of rule elements. There were two original purposes for this part of the prototype. First, while the natural language parsing technology in a limited domain such as privacy policy creation has promising accuracy, it is not perfect.


Author Policy#&#&#&#&#&#&#& - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Links Customize Links Free Hotmail IBM Business Transformation Homepage IBM Business Transformation IBM Home Page
Address file:///C:/Documents%20and%20Settings/brodiec/My%20Documents/PrivacyPrototype/PolicyCreation/policyCreation22.htm?group=Custo
Go
Author Policy
Transform Policy
Map User Categories
Map Data Categories
Map Purpose/Actions
Map Conditions
Map Obligations
Verify Policy
Load
Save
My Computer
Policy Authoring And Implementation Task Steps
Privacy Policy Template Selection List
Privacy Policy Rule Guide
Free Form Area for Typing Natural Language Privacy Policy

Policy Name: DrugsAreUs Privacy F Policy Date: 3/1/04 Source File: Privacy1.txt

Policy Description: This policy defines the uses of customer personal information.

Current Step: Author Policy

Choose a policy template and modify the privacy rules provided as needed or type your policy rules below using the example and information on the Help page as needed.

Privacy Policy Templates:
None Selected
Parent Company Privacy Policy for US (HIPAA)
Parent Company Privacy Policy for Canada

Example Rule Guide:

[User Category(ies)] can [Action(s)] [Data Category(ies)] for the purpose(s) of [Purpose(s)] if [(optional) Condition(s)] with [(optional) Obligation(s)].

```
1. Customer Service Reps, Pharmacists, and Billing Reps can collect and use
customer name and date of birth to help confirm identity.

2. Customer Service Reps, Pharmacists, and Billing Reps can collect and use
customer name and  e-mail addresses to respond to inquiries and to notify
customers about issues regarding orders.

3. Shipping Reps can use customer name and customer mailing addresses to ship
orders.

4. Billing Reps can use credit card numbers for order processing.

5. Customer Service Reps can modify credit card numbers for order processing,
if the customer is notified by email of the change.
```

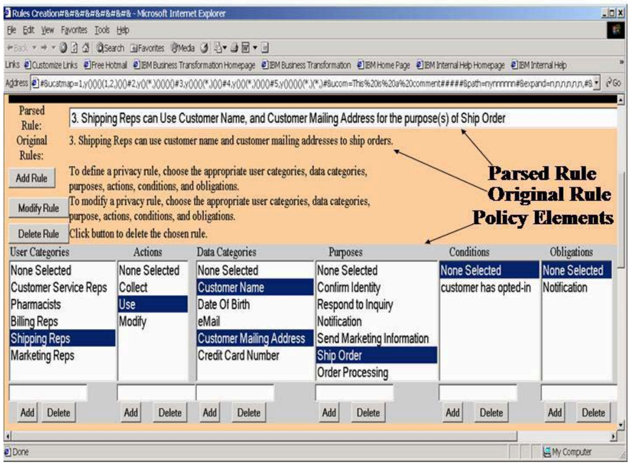
Figure 3. SPARCLE Natural Language Privacy Policy Creation Screen

This page allows users who have created the policy using the natural language technique to confirm that the parsing technology has identified all parts of the rules correctly and to correct anything that is in error. Second, for users who prefer the more structured method for privacy policy creation, this method can be used to create the entire policy. The organization or user can define policy element lists and then rules can be created by selecting the appropriate elements from each of the policy element selection lists and selecting "Add Rule". Likewise, a rule can be modified or deleted by highlighting the rule in the rule selection list, modifying the selected elements as appropriate and selecting "Modify Rule" or "Delete Rule". Any modification to rules or rules added or deleted using the structured approach is automatically reflected in the natural language version of the policy as well. Therefore, the author is able to go back and forth between the two methods to view the policy either in natural language or the parsed format with the elements identified.

During the course of the scenario-based sessions with target users, an additional use of the combined natural language and structured methods was identified. The users indicated that the ability to parse policy rules into policy elements would be valuable to them for assessing the completeness of their existing privacy policies. Several participants were excited about the possibility of using SPARCLE to analyze their existing natural language privacy policies and then viewing the elements and rules identified in order to identify gaps and inconsistencies in the policies. For example, if an existing privacy policy rule fails to identify the purpose for which a particular user group is allowed to use a particular piece of data, the parsed rule would contain "none found" where purpose would usually be. The organizational users felt that this would be a valuable tool for ensuring the quality of the privacy policies used by the organization and helping them to educate their organizations regarding their privacy policies.

5.2 Understanding Privacy Policies

Based on the data collected from interviews with organizational users responsible for the creation of privacy policies, they often find it difficult to understand the policies that they create in order to ensure that policies are complete, able to be implemented, and consistent.


Rules Creation#&#&#&#&#&#&#& - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Links Customize Links Free Hotmail IBM Business Transformation Homepage IBM Business Transformation IBM Home Page IBM Internal Help Homepage IBM Internal Help
Address #&ucatmap=1,y000(1,2,)00#2,y0(*,)0000#3,y000(*,)00#4,y00(*,)000#5,y0000(*,)(*,)#&ucom=This%20is%20a%20comment#####&path=nynnnnn#&expand=n,n,n,n,n,#&
Go
Add Rule
Modify Rule
Delete Rule
Add
Delete
Done
My Computer
Parsed Rule
Original Rule
Policy Elements

Parsed Rule: 3. Shipping Reps can Use Customer Name, and Customer Mailing Address for the purpose(s) of Ship Order

Original Rules: 3. Shipping Reps can use customer name and customer mailing addresses to ship orders.

To define a privacy rule, choose the appropriate user categories, data categories, purposes, actions, conditions, and obligations.
To modify a privacy rule, choose the appropriate user categories, data categories, purpose, actions, conditions, and obligations.
Click button to delete the chosen rule.

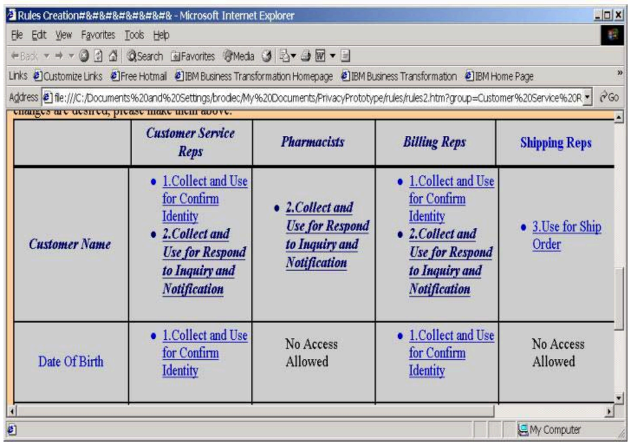
User Categories	Actions	Data Categories	Purposes	Conditions	Obligations
None Selected	None Selected	None Selected	None Selected	None Selected	None Selected
Customer Service Reps	Collect	Customer Name	Confirm Identity	customer has opted-in	Notification
Pharmacists	Use	Date Of Birth	Respond to Inquiry		
Billing Reps	Modify	eMail	Notification		
Shipping Reps		Customer Mailing Address	Send Marketing Information		
Marketing Reps		Credit Card Number	Ship Order		
			Order Processing		

Figure 4. SPARCLE Structured Policy Creation Screen

Figure 5 shows our design to provide users with easy ways of viewing the privacy policy. The Figure contains a table in which two of the policy element types are used as axes and the other privacy rule elements that are associated with each row and column are shown in the cells. In the example that is shown, user categories are used as column labels and data categories are used as row labels. The cells in the table contain the purposes, conditions, and obligations for rules that apply to that user and data category. Using this table, users can see at a glance what type of data users are allowed to access each data element and also see which user groups are never allowed to access particular data items. While the table format was well received by users, we are not yet sure how well a two dimensional table scales up to real organizational policy complexity. Scaling and visualization will be the subject of our future research.

5.3 Validation of Prototype with Target Users

We conducted scenario-based usability walkthrough sessions of two iterations of SPARCLE with people who were responsible for the creation, implementation, and auditing of privacy policies within large organizations in the domains of health care, banking, and government. During the course of the 90 minute sessions, each with 1 to 4 participants, we gathered verbal and written feedback on the usability, design, and value of the privacy tool. For the first iteration of the prototype, walkthrough participants (7 participants in 5 sessions) rated the prototype positively (an average rating of 2.6 on a 7-point scale with 1 indicating “highest value” and 7 indicating “no value”). We present this summary result since it communicates the overall response to the prototype. However, the primary purpose for the sessions was to gather more qualitative responses from the participants about the value of the system to their task of managing privacy policies.


Rules Creation#&#&#&#&#&#&#& - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Links Customize Links Free Hotmail IBM Business Transformation Homepage IBM Business Transformation IBM Home Page
Address file:///C:/Documents%20and%20Settings/brodiec/My%20Documents/PrivacyPrototype/rules/rules2.htm?group=Customer%20Service%20R
Go
My Computer

changes are desired, please make them above.

	Customer Service Reps	*Pharmacists*	*Billing Reps*	Shipping Reps
Customer Name	• 1.Collect and Use for Confirm Identity • *2.Collect and Use for Respond to Inquiry and Notification*	• *2.Collect and Use for Respond to Inquiry and Notification*	• 1.Collect and Use for Confirm Identity • *2.Collect and Use for Respond to Inquiry and Notification*	• 3.Use for Ship Order
Date Of Birth	• 1.Collect and Use for Confirm Identity	No Access Allowed	• 1.Collect and Use for Confirm Identity	No Access Allowed

Figure 5. Table Showing Privacy Policy Rules that Apply to Each User and Data Category

After analyzing the qualitative feedback we received during the evaluation of the first iteration of SPARCLE, we made the following changes: 1) We added the ability to import pre-existing privacy policies into the natural language policy authoring condition to allow SPARCLE to highlight gaps and inconsistencies in the policies, 2) we added the ability to use privacy policy templates as a starting point for authoring privacy policies using either the natural language or structured policy authoring methods, and 3) we improved the readability of the table view of the privacy policy by bulletizing entries and making it scrollable. Additional improvements were made to the mapping and auditing functionality which we will not discuss here. During the second iteration of walkthrough sessions, the participants (15 participants in 6 sessions) also rated the revised prototype very positively (an average rating of 2.5 on the same scale).

During the evaluations we asked the participants to rate 20 features. Figure 6 summarizes the evaluation results over the two iterations of the prototype for 5 of these features which were included in both versions of the prototype and one feature that was added for the second iteration. While the data presented here only represent a small sample, we think that it provides the reader with a good picture of how the users responded to the prototype. The added feature was the ability to import policy files from other sources and to modify those files. This would enable localization of larger corporate policies or laws. This was seen as a highly valuable feature in itself, and we also believe that it led to a more positive evaluation of the natural language entry in the second iteration of SPARCLE. While structured rule entry seemed to be preferred in the first iteration, Natural Language and Structured Entry had equal ratings in the second iteration (these features were not altered substantially between iterations). It was also important to hear from the target users that they felt there was considerable value in the fairly simply policy table that we included in the prototype. We had viewed this two-dimensional representation as an initial design which we might need to change substantially, but found that users actually found it to be very clear and a powerful tool for understanding policy coverage. Additionally, target users responded very positively to the incremental authoring process which allowed high level specification in natural language followed by detail specification (possibly by a different person at a different time). Finally, the participants reported that the compliance checking capabilities we included in the prototype are likely to meet many of their needs regarding monitoring the use of PI within their organizations. Organizations in many domains and geographies now have the requirement of having to investigate and report on the use of individual data subject PI as well as to confirm compliance to policies. The reports provided assist with both of these issues.

6. DISCUSSION AND CONCLUSION

Organizations are increasingly recognizing the need to enforce privacy policies to protect the personal information entrusted to them by their customers, patients, constituents, and employees [32]. This recognition is being driven by an increasing amount of privacy legislation in many geographies as well as changing perceptions regarding privacy in the general public [13,17,18]. Our research contributes to the understanding of the needs of organizations regarding privacy and explores methods for developing usable privacy and security technologies. This research provides insight about developing privacy perspectives, concerns, and the needs of organizations.

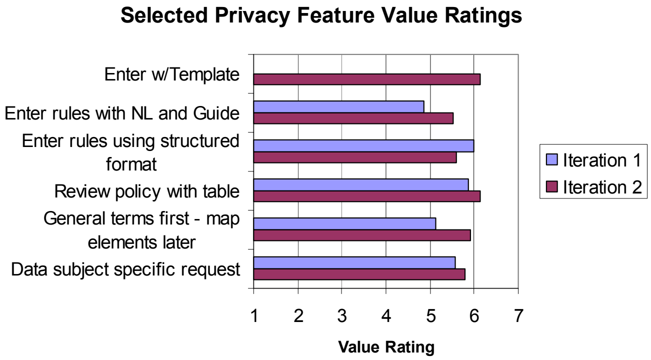


Figure 6. Quantitative Results for Top 5 Rated Features

In this paper, we have discussed research into identifying the needs of organizations regarding the creation, implementation, enforcement and auditing of privacy policies. We then described how those needs were used to drive an analysis of existing and emerging technological approaches to protecting the privacy of PI. Based on that analysis, we have focused our research on the development of a highly usable privacy policy management workbench. This workbench allows users to create privacy policies using interactions methods that are well suited to their skills and background and to visualize the policies that they have created to help ensure that they accomplish their intended goals. We explored and iterated on the design with target users and were able to obtain valuable feedback well before we could complete a full implementation of the prototype. While work on the natural language parsing and mapping components of SPARCLE are still underway, we think we have gained a solid understanding of organizational requirements that is needed for the project's successful completion. The results of these design feedback sessions have provided evidence that the types of functionality highlighted in the two versions of the prototype will be valuable to organizations in helping them to manage the privacy of the PI they must collect and use to provide value to the customers, patients, and constituents and return on investment to their shareholders.

Given the positive results to date, the next step is to create a fully functional version of SPARCLE that will allow us to test our approach to privacy policy creation and management with complex, real-world, organizational privacy policies. Working with policies that contain hundreds of rules may create complications that have not emerged when using smaller sets of rules. For example, two areas of future research into the scalability of the approach include studying the accuracy and the reliability of the natural language parsing of complex privacy policies and determining the effectiveness of visualizations of complex privacy policies in assisting organizational users in understanding the policies that they have created.

We are also interested in determining the degree to which the approach to policy creation used in SPARCLE is able to be generalized to other policy domains. While we hypothesize that the SPARCLE approach may work well for many security and system management areas, additional domain analysis and empirical testing are needed to determine the characteristics of domains and users within the domain which make this approach effective.

While the results of our research into understanding and addressing organizational user needs for privacy will be useful to organizations in helping them protect the privacy of the PI they collect and store, we believe that a secondary value of this work is as an example of how to create more usable security software. Multiple researchers [14, 35] have identified usability as one of the grand challenges for security and privacy research. The application of user-centered methods and HCI research techniques described in this paper could serve as a model for the design of interaction methods for many security projects. In a world with more and more reports of security and privacy risks and breaches, the importance of creating usable security and privacy solutions is increasing. HCI research and the application of user-centered design techniques can help the security and privacy community step up to the challenge of creating interfaces and interaction methods that reduce the complexity in defining, implementing, and managing privacy policies and security solutions for the benefit of all parties.

7. REFERENCES

[1] Adams, A. and Sasse, A. (2001) Privacy in Multimedia Communications: Protecting Users, Not Just Data. In A. Blandford, J. Vanderdonkt & P. Gray [Eds.]: People and Computers XV - Interaction without frontiers. *Joint Proceedings of HCI2001 and ICM2001*, Lille, Sept. 2001. pp. 49-64. Springer.

[2] Agrawal, R., Kiernan, J., Srikant, R., and Xu, Y. Hippocratic Databases. *Proceedings of the 28th Very Large Database Conference (VLDB),* Hong Kong, China, 2002.

[3] Agrawal, R., Kiernan, J., Srikant, R., and Xu, Y. Implementing P3P Using Database Technology. *Proceedings of the 19th International Conference on Data Engineering*, Bangalore, India, 2003.

[4] Anderson R. J. A Security Policy Model for Clinical Information Systems. *In the Proceedings of the 1996 IEEE Symposium on Security and Privacy*, 30-43.

[5] Anderson R. J. Privacy Technology Lessons from Healthcare. In the Proceedings of the 2000 IEEE Symposium on Security and Privacy.

[6] Anton, A., He, Q., and Baumer, D. (2004) The complexity underlying JetBlue's privacy policy violations. *IEEE Security & Privacy*. August/September, 2004.

[7] Ashley, P., Hada, S., Karjoth, G., Powers, C., and Schunter, M. (2003). *Enterprise Privacy Architecture Language (EPAL 1.2).* W3C Member Submission 10-Nov-2003. http://www.w3.org/Submission/EPAL/

[8] AT&T Privacy Bird (2003). http://privacybird.com/

[9] Backes, M., Pfitzmann, B., and Schunter, M. A Toolkit for Managing Enterprise Privacy Policies. *In the Proceedings of the 8th European Symposium on Research in Computer Security (ESORICS)*, Springer-Verlag, Berlin, 2003.

[10] Ball, E. (2003). Patient privacy in electronic prescription transfer. *IEEE Security and Privacy*, 1, 2, 77-80.

[11] Baumer, D., Earp, J.B., and Payton, F. C. (2000). Privacy in medical records: IT implications of HIPAA. *Computers and Society*, December, 2000, 40-47.

[12] Bohrer, K., Levy, S., Liu, X., and Schonberg, E. Individual Privacy Policy Based Access Control. In Proceedings of the 6th International Conference on Electronic Commerce Research (ICECR-6), October, 2003, Dallas, Texas.

[13] Chatham, B. (2004). Online Privacy Concerns: More than Hype. *The Forrester Report*, 2004

[14] CRA Conference on "Grand Research Challenges in Information Security and Assurance". http://www.cra.org/Activities/grand.challenges/security/. November 16-19, 2003.

[15] Cranor, L. (2002). *Web Privacy with P3P*. Cambridge: O'Reilly.

[16] Cranor, L. F., Arjula, M., and Guduru, P., Use of a P3P user agent by early adopters. In the Proceedings of the 2002 ACM workshop on Privacy in the Electronic Society, 1-10.

[17] Hagen, P. (2000). Personalization versus privacy. *The Forrester Report*, Nov., 2000, 1-19.

[18] IBM. *IBM Multi-National Consumer Privacy Survey*. http://www.ibm.com/services/files/privacy_survey_oct991.pdf

[19] IBM Tivoli Privacy Manager for eBusiness (2004). http://www-306.ibm.com/software/tivoli/products/privacy-mgr-e-bus/.

[20] Karat, C. Iterative Usability Testing of a Security Application. *In Proceedings of the Human Factors Society 33rd Annual Meeting*, 1989.

[21] Karat, J., C. Karat, C. Brodie, and J. Feng, "Privacy in information technology: Designing to enable privacy policy management in organizations", *International Journal of Human Computer Studies,* in press.

[22] Karjoth, G. and Schunter, M.(2002) A Privacy Policy Model for Enterprises. *Proceedings of the 15th IEEE Computer Security Foundations Workshop*, 271-281.

[23] Lederer, S., Hong, J. I., Dey, A., and Landay, J. A., Personal Privacy through Understanding and Action: Five Pitfalls for Designers. Personal and Ubiquitous Computing 2004. 8(6): p. 440-454.

[24] Manny, C. H. (2003). European and American privacy: Commerce, rights, and justice. *Computer Law and Security Report*, 19, 1, 2003, 4-10.

[25] Michael, J.B., V.L. Ong, and N.C. Rowe, "Natural-language processing support for developing policy-governed software systems", *39th International Conference on Technology for Object-Oriented Languages and Systems,* IEEE Computer Society Press, pp. 263-274.

[26] Microsoft Internet Explorer (2004). Help Safeguard your privacy on the web. http://www.microsoft.com/windows/ie/using/howto/privacy/config.mspx.

[27] National Research Council. (2003). *Who goes there? Authentication through the lens of privacy*. Washington, D.C: National Academies Press.

[28] Neff, M., Byrd, R., and Boguraev, B. (2003) The Talent system: TEX-TRACT architecture and data model. In Proceedings of HLT-NAACL *Workshop on Software Engineering and Architectures of Language Technology Systems*, Edmonton, Alberta, Canada.

[29] OASIS (2004). Privacy Policy Profile of XACML draft 01. http://docs.oasis-open.org/xacml/access_control-xacml-2_0-privacy_profile-spec-cd-01.pdf

[30] OECD (1980). OECD guidelines on the protection of privacy and transborder flows of personal data. http://www.oecd.org/home/

[31] Office of the Federal Privacy Commissioner of Australia. (2000). *Privacy and Business (2000).* http://www.privacy.gov.au.

[32] Ponemon Institute and IAPP. (2004). 2003 Benchmark Study of Corporate Privacy Practices.

[33] Smith, J. (1993). Privacy policies and practices: Inside the organizational maze. *Communications of the ACM*, 36, 12, 105-122.

[34] The Stanford Student Computer and Network Privacy Project. A study of student privacy issues at Stanford University. *Communications of the ACM*, 45, 3, 2002, 23-25.

[35] Whitten, A. and Tygar J. D. (1999) Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0. *In Proceedings of the 9th USENIX Security Symposium*, August, 1999.

[36] W3C (2002) A P3P Preference Exchange Language 1.0 (APPEL 1.0). http://www.w3.org/TR/P3P-preferences/

[37] Zurko, M. E., Simon, R., and Sanfilippo, T. (1999) A User-Centered, Modular Authorization Service Built on an RBAC Foundation. *In Proceedings of the 1999 Symposium on Security and Privacy*, May, 1999, 57-71.