Examining Internet Privacy Policies Within the Context of User Privacy Values

Julia B. Earp, Member, IEEE, Annie I. Antón, Senior Member, IEEE, Lynda Aiman-Smith, and William H. Stufflebeam

Abstract—Internet privacy policies describe an organization’s practices on data collection, use, and disclosure. These privacy policies both protect the organization and signal integrity commitment to site visitors. Consumers use the stated website policies to guide browsing and transaction decisions. This paper compares the classes of privacy protection goals (which express desired protection of consumer privacy rights) and vulnerabilities (which potentially threaten consumer privacy) with consumer privacy values. For this study, we looked at privacy policies from nearly 50 websites and surveyed over 1000 Internet users. We examined Internet users’ major expectations about website privacy and revealed a notable discrepancy between what privacy policies are currently stating and what users deem most significant. Our findings suggest several implications to privacy managers and software project managers. Results from this study can help managers determine the kinds of policies needed to both satisfy user values and ensure privacy-aware website development efforts.

Index Terms—E-commerce, privacy management, privacy policy, software engineering.

I. INTRODUCTION

INFORMATION privacy has been recognized as an important issue in management, and its significance will continue to escalate as the value of information continues to grow [37], [48], [50]. Understanding and protecting personal privacy in information systems is becoming increasingly critical with widespread use of networked systems and the Internet. These technologies provide opportunities to collect large amounts of personal information about online users, potentially violating those users’ personal privacy [8], [11]. Organizations have taken to displaying their website privacy policies online. This is in part, a way of limiting potential legal liability via disclaimers. Approximately half of new visitors to a given e-commerce site say they read the organization’s privacy policy [16]. Site visitors can make inferences about the organization from the website privacy policy. Some observers have criticized these policies as being too feeble, or too convoluted [6], [33].

Researchers have noted that consumers consider whether a given organization is one that they would feel good about interacting with as part of their transaction decisions [16]. Moreover, consumers make that assessment based on “signals” from the organization [53]. For example, the organization’s published privacy policies may be seen as a signal about the trustworthiness of an organization [35], [42]. If the privacy policies are clearly and explicitly stated, then the visitor/consumer perceives the organization as more trustworthy [28]. This, in turn, helps the organization to attract new customers and retain existing customers. The converse is also true, so that an organization with more suspect policies might have trouble securing new customers or retaining previous ones. Previous customers who become mistrustful may defect to a competitor; feel reticent about disclosing any additional, subsequent information; or give the organization a bad word-of-mouth review [15]. Practices, such as policy statements, that address a customer’s concern about divulging personal information result in positive experiences with an organization, and as a result increase the customer’s perception that the organization can be trusted [15]. Thus, tailoring privacy policies and policy statements to customer interests can likely influence an organization’s bottom line. Because these policies can be so important, privacy, information, marketing, and public relations managers are tasked with the increasing challenge of balancing customer interests with organizational objectives.

These managers are not the only professionals impacted by privacy issues. Software engineering managers and project managers need to make sure that the system functionality matches the privacy statement’s claims. For example, two recent studies found discrepancies between privacy statements and the actual privacy practices in organizations [5], [25]. Bringing policy claims and functionality into alignment is a complex activity; bringing both into alignment with what users value is even more difficult. System functionality is a direct result of system requirements, therefore, accomplishing such alignment begins with consideration at the requirements phase of website development. If the development effort is going to be successful, then users, analysts, developers, and managers must work together during the software requirements phase.

In this paper, we present research bridging the gap between management and software requirements engineering. We address three research questions.

1) What are the most stringently regulated organizations (health care related organizations including health insurance, pharmaceutical, and drugstores) saying in their privacy policy statements?

2) What do consumers value regarding information privacy?

3) Do the privacy policy statements provide the information that consumers want to know?
Results from this study can help managers determine the kinds of policies needed to both satisfy user values and ensure privacy-aware website development efforts.

This paper is organized as follows. First, we discuss relevant research on privacy, policy analysis, and software requirements engineering. Next, we cover the research methodologies of content analysis and survey development, and then the survey results. Finally, we discuss the results and implications of this work for privacy managers and software project managers.

II. BACKGROUND

This section describes fair information practices, the relationship between privacy policies and organizational trust, the relevant work in privacy policy, the role of requirements engineering in policy analysis, and Internet users’ privacy surveys.

A. Background—Fair Information Practices in the United States

U.S. Congressional hearings in the 1970s, where privacy advocates sought to ban credit bureaus from using centralized computer databases, lead to the recognition that organizations have certain responsibilities and individuals have certain rights, regarding information collection and use. Since 1973, the Fair Information Practice (FIP) principles [18] have served as the basis for establishing and evaluating U.S. privacy laws and practices. The FIP principles consist of: 1) notice/awareness; 2) choice/consent; 3) access/participation; 4) integrity/security; and 5) enforcement/redress [18]. U.S. government agencies, Internet users, and industry leaders all agree that organizational privacy policies—particularly those belonging to organizations using electronic transactions—should reflect the FIPs [18]–[20]. Several studies, however, have found that often they do not [5], [14], [17].

In 1980, the Organization for Economic Cooperation and Development (OECD), an international organization, issued the Guidelines on the Protection of Privacy and Transborder Flows of Personal Data [46]. The OECD guidelines are the current best-practice global standard for privacy protection and are the recommended model for legislation in member countries. Although not legally binding, the guidelines are recognized by all OECD members, including the European Union (EU) and the U.S. They are implemented differently among individual nations, suggesting privacy views differ between countries [7]. The U.S. FIPs do not include all of the OECD guidelines, but reflect a subset of them. The EU Directives are even more comprehensive with respect to privacy, and provide the legal foundation for those countries. Since the FIP Principles are fundamental to U.S. privacy guidelines, they are incorporated in this research. For clarity of results, the study presented here is limited to U.S. Internet users. We also focus our study on health care websites because these websites are required by law [the Health Information and Portability Accountability Act (HIPAA)] to post a website privacy policy.

B. Background—Relationship Marketing: Privacy Statements as Signals for Trust

In the late 1990s, organizations began to pay more attention to relationship marketing, where marketing activities are customized down to the individual level. Organizations can track data about an Internet visitor’s behavior [51]: which links they select, in which order, and how long they spend on each webpage. If the user transacts a purchase, then the organization might store the user’s name, address, phone number, e-mail, password, bank account, or credit card information. Organizations can then use these data to personalize and enrich the user’s shopping experiences [15], [33], in the hopes of increasing sales. Some consumers worry, however, the information might be traded or sold. When Internet users consider engaging in e-commerce with an organization, they look for signals from the website to develop an assurance of the organization’s trustworthiness. These signals include posted statements or privacy policies that inform website visitors what a given organization will do with any personal information provided [32].

When an organization adopts fair procedures to protect individual privacy, customers tend to develop trust in the organization. As a result, they are likely to be more willing to disclose personal information and have that information used. Such trust is essential in business to both customers and businesses, because research shows that trust contributes substantially to customers’ satisfaction and commitment to continuing to do business with a particular organization [24], [38], [40], [43]. In recent surveys of Internet user attitudes, users consistently cited privacy as a primary concern [13], [30]. In particular, the results from these surveys suggest that the security and privacy of business-to-consumer websites are a driving factor in user purchasing decisions [49]. Therefore, it is important to align website privacy policies and practices with users’ needs and concerns regarding privacy [4].

C. Background—Research on Privacy-Policy Analysis

Business success requires the technology businesses use fit the needs of the human consumer. One stream of our research is in the software engineering perspective of privacy, establishing a core set of privacy related software requirements intended for web-based systems [3]. Therefore, we discuss policy analysis from both a software engineering perspective and a managerial perspective.

In making online consumer privacy recommendations to the U.S. Congress, the Federal Trade Commission (FTC) has relied on four studies assessing organizational awareness of and adherence to the U.S. FIP principles [2], [14], [19], [20]. The four studies all employed expert web surfers to analyze the privacy-policy statements for the given sample of U.S. websites. The content analysis process used varying questions and focused on different aspects of the FIP principles. Three studies concentrated on notice/awareness, choice/consent, access/participation, and integrity/security issues [14], [19], [20]. The other explored notice/awareness, choice/consent, and integrity/security issues [2]. The content analysis in all four studies considered the entire website privacy policy as the unit of analysis, rather than individual statements contained within the policies. Such high-level analyses provided by these four studies are insufficient from an engineering perspective, because software engineers who are developing web-based systems need to develop functionality at the detailed, per-statement, level.
Although the aforementioned study provides some information and insight, it did not address the issue of what consumers, the users, value. Yet, what users’ value is critical for driving manager policy decisions and actions, and for justifying engineering plans. Examining either of these issues (privacy policies or user values) in isolation of each other thwarts managers’ ability to efficiently plan and direct systems to meet users concerns and increase their trust.

D. Background—Research on Internet Users: Privacy Opinions

Managerial researchers investigating user privacy opinions have identified dimensions of privacy concerns, developed descriptive categories of users, determined factors relating to user confidence, and established that users are looking for privacy policies with specific content. Some of these studies are described in this section.

In 1996, Smith et al. published a privacy survey [52], in the management of information systems literature, measuring four dimensions of privacy concerns about overall organizational practices: 1) concerns about the collection and storage of information; 2) concerns regarding errors in data collected about individuals; 3) concerns about data being collected for one purpose, and subsequently used for another purpose without the individual’s consent; and 4) concerns about individuals’ data being available to unauthorized viewers.

Two research studies, one in 1999 [1] and another in 2001 [54], asked U.S. Internet users about their online privacy concerns regarding specific online scenarios. The Internet users were then categorized according to three privacy categories used by Harris-Equifax [30].

1) Privacy fundamentalists are individuals who are extremely concerned about their privacy; they rarely reveal any private information about themselves, even when privacy protection measures are in place.

2) The pragmatic majority, the bulk of Internet users. These users are concerned about privacy, but if policies address their specific concerns, they can be appeased. The pragmatic majority can be divided into two specific clusters: the identity concerned are most concerned about revealing personally identifiable information such as their name, address, e-mail address, etc., while the profiling averse are more concerned about revealing information about their hobbies, interests, health, etc.

3) Finally, the marginally concerned are those individuals who are willing to provide personal information to websites under almost any circumstances.

A 2002 study on Internet privacy examining health care, retail, and financial websites identified three factors as being the most influential on user confidence [16]. Most important was the company name. Next important was providing users the option to “opt-out” of data collection. Third important was the presence of a site privacy policy. Fifty-four percent of respondents said they would read a website’s privacy policy on the first visit; 66% of respondents indicated increased confidence in a website if a privacy policy was present. Although the organization’s brand name is the most influential characteristic of user confidence and loyalty, a website lacking brand-recognition could boost user confidence by displaying a user-centric privacy policy [16].

A study of users and businesses found that 87.5% of surveyed users expect to see comprehensive information regarding privacy practices when visiting a commercial website [23]. Similarly, another found that 59% of users say they have read privacy notices, while 91% thought it important to post privacy notices [30]. These findings further emphasize the importance of site privacy policies and privacy-policy content.

These research studies, and the surveys they used, did not address organizational website signaling to the Internet user, nor did they address specifics to aid privacy managers, project managers, and requirements engineers in the development of web-based systems and privacy policies that align with what users want.

Information and privacy managers, project managers, and engineers have a professional responsibility to consider the social impacts of the systems that they design and implement. Managers must ensure they establish and maintain the trust of the public who read the policies on the organization’s websites. User trust can be attained if strategic data management decisions and practices align with what they value. Therein lies the significance and novelty of this work, in that we examine the claims that the most stringently regulated organizations (health care related) are making in their privacy-policy statements. We also investigate what users value regarding information privacy. We then determine—do the privacy-policy statements provide the information that users want to know?

III. RESEARCH METHODS

A. Content Analysis—What are Websites Saying

We addressed the first of our three research questions (what are health care organizations saying in their privacy-policy statements?) using content analysis. Content analysis is a grounded theory approach [26] for making valid inferences from text-based data [36]. We used a diverse sample in our website analysis, as well as a purposeful sampling strategy to capture heterogeneity [39]. We identified 24 websites from a variety of industries (e.g., health care, retail, Internet service providers, and travel agencies) and chose three websites from each industry. We chose the websites that received the most visitors in that industry, because of their popularity. We used the privacy policy statements from those 24 websites for exploratory coding into categories. We then identified 23 highly visited health care websites for data analysis in this study. We chose the health care industry because we reasoned given regularity conditions those websites would have the most comprehensive privacy policies. Additionally, those organizations are required by HIPAA to have a privacy policy.

In contrast to the high-level policy analyses done by previous researchers (see [2], [14], [19], and [20]), our content analysis examined each statement within an organization’s policy at a micro level. That is, we scrutinized each individual policy statement, parsing it into sentences and words, rather than considering the overall meaning that flowed from the policy. This process of coding entailed selective reduction of a large volume of text.
We coded by reading through the text and manually noting statement occurrences. We could not use an automated computer program in this research because we did not have a strict set of *a priori* categories and decision rules. Hence, the coding process was time-intensive. For example, the analysis of each website policy by one person averaged two full workdays (i.e., 16.5 hours), since one website might yield 12–20 pages of privacy-policy text. Three separate researchers working independently coded each website. We originally used the five U.S. FIP categories in coding, but quickly noted that statements in the actual privacy policies did not fit—other categories emerged from the content analysis process. All three researchers coded, compared categories, discussed the categorization processes and findings with acknowledged privacy experts, and reiterated this process until 100% categorization agreement was reached. This intense content coding and discussion process resulted in 12 categories of privacy policies (see Table I). Developing this 12 category taxonomy is explained more fully in [3].

The 12 categories in the taxonomy can be described either in terms of Privacy-Protection (desired protection of user privacy rights) or Privacy-Vulnerability (potential for invasions of privacy). An example of a Privacy-Protection statement would be “(organization name) does not sell or rent customer lists.” An example of a Privacy-Vulnerability statement would be “(Organization name) retains all customer information as a business asset.” This taxonomy provides a framework for comparing and analyzing privacy-policy statements, and can be used by researchers, by managers in organizations, and by website developers to ensure that their stated and actual policies are consistent with each other and that they reflect what online consumers value.

For this current research study, we used this 12 category taxonomy and extended the work to identify more fully what privacy policies in health care organizations are saying. Two researchers (one was not involved in developing the taxonomy) did a separate content analysis on another set of 23 health care

---

### Table I

<table>
<thead>
<tr>
<th>Privacy Protection Goal Classifications</th>
<th>Privacy Vulnerability Goal Classifications</th>
</tr>
</thead>
</table>
| **Notice/Awareness**  
Assert that users should be notified and/or made aware of an organization’s information practices (e.g., via an organization’s privacy policy) before any information is actually collected from them. | **Information Monitoring**  
Organizations’ tracking practices (e.g., what users do on their site through means such as cookies). Could be for the user’s benefit (e.g., when an e-commerce application maintains a shopping cart for a user), or for the organization’s benefit, be it for purely statistical use or for profit (e.g., via selling of aggregated information to 3rd parties). |
| **Choice/Consent**  
Ensure that users are given the option to decide what personal information collected about them is to be used and whether it may be used for secondary purposes. | **Information Aggregation**  
Aggregation practices as when organizations combine previously gathered data in such a way that they create non-identifying statistical data often used for marketing and promotional purposes. |
| **Access/Participation**  
Allow or restrict access to a particular site or functionality based upon whether or not the user provides their PII (Personally Identifiable Information). Address the ability for users to access or correct their PII. | **Information Storage**  
What and how records are stored in an organization’s database. |
| **Integrity/Security**  
Ensure that data are both accurate and secure. Security and accuracy come from both the user and the organization collecting the personal information. Goals in this category range from vague ones stating only that personal information is securely kept to specific technical descriptions of what security protocols will be used to transfer personal information over the Internet. | **Information Transfer**  
Any transfer of information. Privacy by its very definition means insurance that others cannot find something out, that information must not be transferred. Addresses the transfer of information, as well as to whom what information is transferred. |
| **Enforcement/Redress**  
Address the mechanisms in place to enforce privacy. Prescribe general guidelines that companies and their employees should follow. These include both self-imposed and government imposed work restrictions. Redress includes possible actions for consumers harmed by a violation of the policy. | **Information Collection**  
How and what information is being collected. Collection occurs when an organization collects information from a user either by directly requesting that they enter the information, or by collecting information without user consent, such as browser information. |
| **Information Personalization**  
Actions that reflect the customization of a website to a specific visitor, thus affecting the functionality/content offered to individuals. This may be as simple as greeting the website visitor by name (e.g., “Welcome, George.”) | **Contact**  
How and for what purpose organizations contact users using their personal information. This could be helpful, such as contacting customers to validate an email address, or annoying, such as sending out unwanted promotions based on past patterns. |
related websites. These sites included pharmaceutical companies and health insurance companies, collectively referred to as the health care industry. These websites were all well established; they reflected large traffic patterns, significant revenues, or name-recognition. Two researchers systematically identified 404 privacy statements, then separately classified them into the 12 content categories of the taxonomy. Cohen’s kappa [12] was computed to measure the interrater reliability, and was 0.93, thus indicating a high level of agreement. Table II shows the frequency of statements by category. This enabled us to address our first research question, what are health care organizations saying in their website privacy policies?

B. Developing a Survey Instrument to Measure Users’ Privacy Concerns

A web-based survey seemed the most appropriate means for data collection to accomplish the objectives of this study since the population of interest is general Internet users. In this section, we provide an overview of the survey measurement items, discuss our pilot study, and describe our final survey distribution.

1) Measurement Items: Following a careful item development process, as suggested by Nunnally [45], drawing from the literature [13], [52], and statement language from the taxonomic categories, items were created to tap into what users’ value in terms of privacy policies. These items were reviewed by survey experts and a panel of privacy experts to assure we had adequately tapped into the content area, and were given to a focus group of 26 Internet users to evaluate the items for clarity. Feedback from these processes resulted in a draft survey with 60 items on privacy concerns, user attitudes about providing personal information, and online experiences. We also developed 16 demographic items. We then pilot tested this draft survey.

2) Pilot Testing: We pilot tested the survey using a written format with a sample of 386 students from a large southeastern university. Of these, 347 respondents returned usable surveys (90% response). Based on exploratory analysis and Nunnally’s reliability heuristics [45], we deleted some items and reworded some others, resulting in a revised survey instrument having 36 scale items. As a reasonable tradeoff between respondents’ concerns about the survey length and our desire to collect additional information, the final survey was shortened for a random selection of respondents (using the capabilities of online surveying).

3) Survey Distribution: The final survey (see Appendix) was distributed online to Internet users worldwide. Respondents were solicited through a variety of outlets, including links to the survey from various university web pages, general news sites, computer enthusiast sites, e-commerce sites, and privacy sites, as well as offline newspapers, e-mail, and word of mouth. The survey was available April 5, 2002 to May 31, 2002 via the Web at an NSF-sponsored project site. To reduce potential occurrences of incomplete survey responses due to the survey length, the online survey was coded so that 80% of respondents would receive the shortened survey version having three sections totaling 61 items (36 scale items, 1 open-ended item, and 24 usage and demographic items). The other 20% of respondents would receive the original revised survey having four sections totaling 73 items (36 scale items, 1 open-ended item, 24 usage and demographics, and 12 additional questions). The selection process for each respondent was carried out in a random fashion. During the one-month time period 1005 usable surveys were returned.

C. Survey Results

Respondents represented 30 countries with 82% stating their country of residence as the United States. Due to differences in social views on privacy between the U.S. and other countries, we chose for the sake of clarity to use only the 827 respondents who indicated they were U.S. residents.

Approximately 67% of U.S. survey respondents were male, 55% were between the ages of 18 and 35, and 82% had earned a college degree or higher. These demographics are comparable to profiles reported in other Internet user studies [34], [47]. Additionally, 89% of the U.S. respondents began using the Web over four years ago and 47% currently use the Web more than 20 hours a week. See full demographic descriptions in Table III. Because privacy is more of a concern with transaction-based websites than other types of websites, it is important to note that 67% of the respondents had made an online purchase within the past 30 days.

D. Exploratory Factor Analysis

We split the total data set in two; and used one part for exploratory work, and the other for confirmatory work [31]. We factor analyzed a randomly selected subset (n = 300) of the 827 usable responses. Those items that loaded less than 0.40 or cross-loaded were discarded.

E. Confirmatory Factor Analysis

Construct validity focuses on the extent to which data exhibit evidence of convergent validity, discriminant validity, and
The confirmatory analysis showed good fit of the measures in those six factors. All factor parameter estimates were significant. We termed these factors Personalization, Collection, Transfer, Notice/Awareness, Storage, and Access/Participation. These factors in terms of underlying content and language relate to many of the dimensions we had extracted from the websites privacy policies. Table IV shows descriptive statistics, as well as reliabilities for these measures using the entire data set. Reliabilities range from 0.74 (Access/Participation) to 0.93 (Transfer). The factor correlation matrix (see Table IV), indicated discrimination between these constructs. The full instrument is included in the Appendix.

### IV. Results

Because of their inherent dissimilarities, the content analyses of the website data and the survey data from users could not be quantitatively compared. We, therefore, applied a qualitative comparative approach to our third research question: do the privacy-policy statements provide the information that consumers want to know?
As previously noted, two researchers independently coded privacy statements into categories (see Table II). The final data set of coded statements enabled us to count the frequencies of each category. An implication we can draw from analyzing all the texts is that the more often a type of statement is made, the more emphasis, or importance, the organization’s privacy policy is placing upon it. Since a privacy policy acts as a signal to the user, it should relay information that is of importance to users. The category with the most frequently made statements was Integrity/Security, or statements assuring security about data collection and transfer. It is reasonable, therefore, to interpret that privacy statements in these websites are placing highest importance on informing consumers of the integrity and security of data during collection and transfer over the Internet.

The category with the next highest number of statements made was Collection, or statements about how data are collected—either by direct entry of user, or by unobservable means such as using browser cookies. Organizational websites were emphasizing the explicit and hidden processes about data collection. Although this is an important concept, the majority of personal data collected by websites is done through direct means of user entry. Therefore, users may already be aware of much of the information that is being collected.

The third highest frequency of statements fell in the category of Choice/Consent, or statements about the user being able to decide what information about them can be used. For several years, practitioners and researchers have debated the usefulness and effectiveness of opt-in and opt-out alternatives [41]. As a result, this topic is heavily emphasized in the content of our analyzed policies.

What users value reflect clearly different privacy priorities. In contrast to the areas emphasized in the websites, the survey data showed that users are most concerned with: 1) transfer, or concerns that their data will be shared, lent, or sold to other entities (mean = 4.77, stddev = 0.58); 2) notice/awareness, or concerns about having full information about how their data might be used before providing data to the organization (mean = 4.65, stddev = 0.51); and 3) storage, or how the organization means to store and maintain the user data (mean = 4.54, stddev = 0.76). Thus, our third research question—are website privacy-policy statements aligned with user privacy values—is answered in the negative (see Table V).

V. DISCUSSION AND CONCLUSION

To date, common practice suggests the five Fair Information Practice Principles (Access/Participation, Choice/Consent, Enforcement/Redress, Integrity/Security, and Notice/Awareness) should be the basis of privacy-policy content [19]. However, our content analytical research found many website statements that lie outside of the FIP principles; in fact, we developed seven categories of statements that we term vulnerabilities to users in that they pose potential invasions of privacy. This suggests that the FIP principles provide an incomplete basis for developing and analyzing privacy-policy statements. We also found in the development of our survey research that some FIP principles, surprisingly, were not valued highly by Internet users.

One of our conclusions from this research is that the protection statements of the FIP principles will not buffer the user against potential privacy vulnerabilities. Consider an example protection statement, under the Choice/Consent category of the taxonomy (Table I), and how such a statement would affect vulnerabilities. The following appeared in one single privacy policy in our analysis.

- **Choice/Consent Statement:**
  (Organization) will not disclose PII without consent.
- **Transfer Vulnerability:**
  (Organization) can disclose PII to designated pharmacy for service.

The intent of the policy writer may have been to protect the user’s privacy by ensuring that information disclosure to pharmacies occurs only with user consent; however, the presence of both these statements introduces a potential conflict into the system. A user of that website would not know whether the policy refers to implicit or explicit consent. If this website privacy policy had simply addressed the FIPs, then that statement
about disclosure to a designated pharmacy would have been omitted from the policy, and the user would be unaware that their PII was being disclosed.

Software engineers and managers need to go beyond the FIPs. They need to pay attention to practices that open vulnerabilities to their website users, and they need to consider what those users value when developing website privacy policies. As well, project managers should use the same information when designing and managing consumer-based systems. Our research suggests systems are being developed to satisfy the businesses’ objectives, but not those of the consumers/users. If systems are developed that do not reflect what users’ value, then users may not develop trust in the organizations.

A. Implications for Management

Internet use is on the rise [44] and e-commerce is becoming increasingly pervasive. As the transition to a more technology-based society takes place, privacy of personal information is becoming increasingly important. Research has shown that Internet users are aware of the privacy-policy concept and are increasingly beginning to read privacy-policy content when visiting websites [16], [29]. We have revealed that the privacy practices currently disclosed in website privacy policies are not aligned with the privacy practices users want to see. This study, then, serves as an initial step toward proactive privacy management by organizational managers and other professionals. By carefully evaluating their own organization’s website and privacy policy using the categories in the privacy taxonomy, while noting what consumers value, managers can identify potential problems with the visible website privacy-policy. By taking a proactive approach toward managing both privacy-policy content and website implementation, managers could influence the development of consumer trust. This further suggests managers should be aware and in control of the practices disclosed in their organizations’ website privacy policy.

One explanation of the disparity between user privacy values and website privacy-policy content is that website organizations and users have different objectives. Website organizations, in general, undoubtedly have an objective to be legally protected from potential lawsuits; therefore, their website privacy policies are often written in such a way to protect the organization. This objective interferes with user needs by causing the privacy policy to assist the organization rather than the user. In some cases, there are legal requirements for privacy notices, e.g., Gramm–Leach–Bliley [27] and Children’s Online Privacy Protection Act (COPPA) [10].

Previous research has found that the FIP principles as a whole are valued by users and should be adopted by organizations as a means to facilitate user trust [15]. Our study revealed that organizations need to look beyond the FIP principles and disclose privacy practices that also relate to vulnerabilities such as Transfer (when data is given to other entities), and Notice/Awareness (providing users with full information about how the organization stores and maintains user data). A website privacy policy can easily address the FIP principles, while neglecting those areas that appear to be of greatest concern to users. Simply adhering to the FIP principles can mean that privacy-invasive practices are not included in the policy statement, which can be dangerous. For example, if the policy does not conform to its practices, the organization may be sued by the FTC for deceptive trade practice (e.g., see the cases of Toysmart [21] and Eli Lilly [22]). This emphasizes the importance of software engineering alignment and managerial awareness. In particular, managers must be able to ensure that their systems are compliant with policies and legislation, while reflecting the values of those who use the systems.

Website implementation should support the respective policy of the organization; therefore, it is important to establish an effective approach for designing appropriate and accurate privacy policies. Further exploration into privacy policy may reveal the reasons that privacy-policy content does not align with user concerns. Whether it is because the website organizations are not aware of user concerns, or because they have ulterior motives to protect themselves, our results provide motivation for researchers to now explore the reasoning behind current privacy-policy development and establish novel approaches to developing website privacy policies.

B. Limitations and Conclusion

This work has resulted in a validated survey instrument containing 36 scale items pertaining to the privacy statements for future use by privacy researchers and organizational managers. The main contribution of this study is the establishment of a gap between what user value and what website privacy-policies emphasize. As noted previously, little attention has been given to alignment between website privacy policies and user concerns [6]. This study has shown that website privacy policies do not address the same dimensions of website privacy that users are concerned with and want to see. The survey data reveal users are most concerned with transfer, notice/awareness and storage; therefore, researchers addressing alignment in the future now have a substantiated rationale for focusing on those dimensions.

Given the nature of the data in the current study, our results were based upon an effort to quantify and perform numerical analysis of qualitative statements. Although we have presented a qualitative comparison that illustrates the divergence between what users are concerned about and the privacy practices websites disclose, further privacy-policy research could strengthen these results. Further research could also demonstrate whether or not users are indeed more likely to visit, develop trust in, or purchase from a website that contains what they want to see in a privacy-policy.

A limitation of this research is that we used data only from the U.S. People often think of privacy as a legal or moral right [11], but attitudes toward privacy differ among cultures and different jurisdictions afford greatly differing levels of legal and regulatory protections. These legislative variations from one country to the next have implications for globally focused organizations and for users who deal with organizations from across the world—and the Internet is by nature, global. Because of these differences, it is important that organizations develop and display comprehensive privacy policies detailing organization-specific practices with regard to personal information. It is also
imperative that global users who read and evaluate these policies fully understand what they are reading, and that they are familiar with the privacy management (and privacy invasive) technologies that are used by the organization. Future research could include administering the survey in a manner that will result in more non-U.S. respondents and performing analyses of non-U.S. privacy policies.

The results show that there is a notable discrepancy between what privacy policies are currently stating and what users deem most significant. This finding provides great managerial implications given that privacy policies act as a signal of trustworthiness. This suggests that website privacy managers have an opportunity to align their privacy policies better with user concerns, potentially resulting in a better relationship with their customers. To better address the void between the information users want to receive and the information organizations provide, organizations should be certain to include statements that address users’ concerns as revealed by this study. Specifically, organizations should be certain to include statements that address: how users’ data is shared, lent, or sold to others (e.g., “We may share your account profile information with our business partners and subsidiaries for marketing analysis and other collaborative efforts amongst our organizations.”); inform users’ how their sensitive information might be used before providing it (e.g., “If you choose to provide us with your e-mail address, we reserve the right to use it to contact you regarding changes to our website, this policy, and to provide you with up-to-date information on our product and service offerings.”); and how user data will be securely stored and maintained (e.g., “All information collected via this website is encrypted using SSL during the collection process. It is then stored both locally and at an offsite facility, access to both of which is restricted to employees who need access to the information, and the accesses are restricted solely to those necessary for the specific task at hand.”). Organizations should make a concerted effort to include statements similar to these examples to ensure they have addressed those items of greatest concern to their users: data transfer, notice/awareness, and storage of data.

APPENDIX

FINAL SURVEY INSTRUMENT. EACH ITEM USES A FIVE-POINT LIKERT SCALE ANCHORED BY “STRONGLY DISAGREE” (1) AND “STRONGLY AGREE” (5)

Factor 1: Personalization (alpha = .86)
I mind when a Web site uses my (PII) to customize my browsing experience.
I mind when a Web site uses cookies to customize my browsing experience. (A cookie is information that a Web site puts on your hard disk so that it can remember something about you at a later time).
I mind when a Web site uses my purchasing history to personalize my browsing experience (e.g. by suggesting products for me to purchase).
I mind when my PII (Personally Identifiable Information) is used for marketing or research activities.
I mind when a Web site monitors my purchasing patterns.

Factor 2: Notice / Awareness (alpha = .82)
I want the option to decide how my PII is used.
I want a Web site to disclose security safeguards used to protect my PII.
I want a Web site to disclose how my PII (Personally Identifiable Information) will be used.
I want a Web site to inform me before using my PII in a manner that it had not previously disclosed to me.
I want a Web site to keep me informed of changes to its privacy practices.

Factor 3: Transfer (alpha = .93)
I mind when a Web site discloses my buying patterns to third parties.
I mind when my information is shared with third parties.
I mind when my PII (Personally Identifiable Information) is traded with or sold to third parties.

Factor 4: Collection (alpha = .87)
I mind when a Web site that I visit collects (without my consent) information about my browsing patterns.
I mind when a Web site that I visit collects (without my consent) information about my browser configuration.
I mind when a Web site that I visit collects (without my consent) information about my IP address (a number that uniquely identifies your computer from all other computers on the Internet).
I mind when a Web site that I visit collects (without my consent) information about the type of computer/Operating System I use.
I mind when a Web site records the previous Web site I visited.

Factor 5: Information Storage (alpha = .82)
I am concerned about unauthorized employees getting access to my information.
I am concerned about unauthorized hackers getting access to my information.

Factor 6: Access / Participation (alpha = .74)
I want a Web site to allow me to check my PII (Personally Identifiable Information) for accuracy.
I want a Web site to allow me to modify my PII.
ACKNOWLEDGMENT

The authors gratefully acknowledge G. Dickson, A. Hevner, and R. Barkhi for their helpful comments in the preparation of this manuscript. They also appreciate T. Alspaugh, H. To, and A. Reese for their assistance with the content analysis and S. Scullen for his assistance with the data analysis. The authors also acknowledge D. Baumer, C. Jensen, and C. Potts for their thought-provoking discussions that assisted in the completion of this work, as well as C. Benavente for assisting in the initial draft of survey items. They also appreciate the suggestions provided by the department editor and reviewers throughout the revision process.

REFERENCES


Julia B. Earp (M’00) received the B.S. degrees in mathematics and in statistics from North Carolina State University, Raleigh, in 1991 and 1991, respectively, the M.S. degree in information technology from the Virginia Polytechnic Institute and State University (Virginia Tech), Blacksburg, in 1991 and 1997, respectively.

She is an active Senior Research Collaborator with theprivacyplace.org. She is an Associate Professor with the College of Management, North Carolina State University. Her research interests focus on Internet security and privacy issues from several different perspectives, including data management, consumer values, systems development, and policy. She has published articles on information privacy and security in several journals such as Communications of the ACM, IEEE Security and Privacy, Requirements Engineering Journal, Computers and Security and the International Journal of Organizational Computing and Electronic Commerce. Dr. Earp is a member of the Association for Computing Machinery (ACM) and AIS.

Lynda Aiman-Smith received the B.S. and M.S. degrees in organizational behavior and information systems from the University of San Francisco, San Francisco, CA, in 1984 and 1985, respectively, and the Ph.D. degree in organizational behavior and technology management from Purdue University, West Lafayette, IN, in 1996. Prior to her doctoral work, she was with Raychem Corporation in supply chain and a Manufacturing Plant Manager. She is an Associate Professor of Business Management with the College of Management, North Carolina State University, Raleigh. She also serves as an Associate Faculty for the Center for Innovation Management Studies, North Carolina State University. She is an active Senior Research Collaborator with theprivacyplace.org and is currently doing research and writing on innovation and organizational culture.

William H. Stufflebeam received the B.S. degree in business management, the M.S. degree in computer science, in 2001 and 2004, respectively, and the Professional Computer Programming Certificate, all from North Carolina State University, Raleigh. He is currently working towards the Ph.D. degree in computer science at North Carolina State University.

He is a member of The Privacy Place (http://theprivacyplace.org) and The Cyber Defense Laboratory. His research interests include requirements engineering, information security and privacy, and the applicability of software engineering techniques to a broader spectrum of systems development.

Mr. Stufflebeam is a member of the Association for Computing Machinery (ACM).

Annie I. Antón (S’91–M’92–SM’03) received the B.S., M.S., and Ph.D. degrees in computer science from the Georgia Institute of Technology, Atlanta, GA, in 1990, 1992, and 1997, respectively. She is an Associate Professor with the College of Engineering, North Carolina State University, Raleigh, where she is Director of The Privacy Place (http://theprivacyplace.org) and a Cyber Defense Laboratory member. Her research interests include software requirements engineering, information privacy and security, software evolution, and process improvement.

Dr. Antón is a member of the IFIP Working Group 2.9, the 2004–2005 IDA/DARPA Defense Science Study Group, the Association for Computing Machinery (ACM), and the International Association of Privacy Professionals (IAPP). She is presently an Associate Editor for the IEEE TRANSACTIONS ON SOFTWARE ENGINEERING and cognitive issues Area Editor for the Requirements Engineering Journal.