Examining the Usability of Web Privacy Policies

Robert W. Proctor  M. Athar Ali  Kim-Phuong L Vu
Department of Psychological Sciences, Purdue University, West Lafayette
Department of Computer Technology, Purdue University, West Lafayette
Department of Psychology, California State University, Long Beach

Goals

• To understand what types of information are included in privacy policies
• To obtain measures of users’ comprehension of privacy policies and attitudes toward the policies and their host sites

Analysis of Web Sites

Last year, we reported an analysis of 7 Web site categories, with 6 sites in each, that examined the types of personal information that each site requested users complete for on-line transactions (Ali, Proctor, & Vu, 2005). Last year, we reported an analysis of 7 Web site categories, with 6 sites in each, that examined the types of personal information that each site requested users complete for on-line transactions (Ali, Proctor, & Vu, 2005).

We found that considerable differences exist between site categories. More importantly, there were significant differences within categories with respect to the type and amount of information that a user is asked to provide. We found that considerable differences exist between site categories. More importantly, there were significant differences within categories with respect to the type and amount of information that a user is asked to provide.

The within-category variability indicates that many sites are asking for information that is not necessary to conduct a transaction. At least 80% of the Web sites in all categories except on-line gaming had links to the site’s privacy policy. These links were typically located at the bottom of the page, centered or to the right.

Content Analysis

Privacy policies were analyzed to determine the number of protection and vulnerability goals included in each.

Method

Privacy policies from 25 Web sites in each of the 4 most widely used categories (pharmacy, retail, banks, and insurance) from the first phase were selected for content analysis. Goal mining analyses were performed to determine the number of privacy goals contained in each policy and whether they consisted of mainly protection or vulnerability goals (Anton et al., 2004).

Results

Across the four categories, we found that the privacy policies contained more vulnerability goals than protection goals when goals addressing both issues were not taken into account.

For the pharmacy category, protection goals were almost equal in number to vulnerability goals, but the number of goals of each type was lower for that category than for the other three categories. Also, online pharmacies, retailers, and banks were more likely to be low on both protection and vulnerability goals or high on both.

Readability analysis

The same 25 Web sites in each category were also analyzed for readability. The readability analysis was performed using MS Word and ReadMe tool to determine the word count. Flesch Reading Ease Scores (FRES) and Flesch Grade Level (FGL) were obtained to score the readability of the policies.

Results

The reading analysis showed that the mean grade level needed to understand these policies is at least 13 years. Given that not everyone using the Web has a college education, many users may not be able to understand the privacy policies. Moreover, many privacy policies were too long to be read by most users. Overall, the length of the policy did not affect its readability (r = 0.10). So, making the policy more concise will increase the likelihood that the policy will be read without necessarily decreasing its readability.

Reading Comprehension Experiment

We conducted an experiment to examine comprehension of privacy policies by college students, who should be at the 13-year reading level at which the average policy is written.

Method

20 students enrolled in Introductory Psychology, all experienced with Web-based transactions, participated. The students were at the educational level at which the average policy is written.

Privacy policies were selected from the bank and retail categories because college-age students are most likely to use Web sites within these categories. Four privacy policies from those used in the content analysis were selected that were of similar reading level, with two being LL and two being HH on goals. The company names were removed to eliminate any possible influence of previous experience.

Each person read all policies, with the order counterbalanced across subjects. Each policy was read one-at-a-time with no time limit. When the subject entered a command to continue, s/he was asked 9 questions about the content of the policy (“Does this website share information with other websites/outside companies?”) and 9 questions about perceptions of the policy (e.g., “Do you think that the company collects more customer information than is required?”).

Results

<table>
<thead>
<tr>
<th>Website Type</th>
<th>% correct</th>
<th>Website Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short length (2100 words) &amp; LL on goals</td>
<td>59</td>
<td>Policy A: Bank</td>
</tr>
<tr>
<td>Long length (1106 words) &amp; HH on goals</td>
<td>59</td>
<td>Policy B: Bank</td>
</tr>
<tr>
<td>Long length (2100 words) &amp; LL on goals</td>
<td>54</td>
<td>Policy C: Retail</td>
</tr>
<tr>
<td>Medium length (1106 words) &amp; LL on goals</td>
<td>37</td>
<td>Policy D: Retail</td>
</tr>
</tbody>
</table>

References
