Both privacy policies and end-user licensing agreements (EULAs) are ubiquitous in today’s computing environment. Users are frequently prompted to agree to privacy policies and EULAs, and often do so without even looking at them. This is in seemingly direct contradiction to the documented concern consumers feel regarding the capture, storage and handling of their personal information by websites and mobile applications. This study explores why users may choose to disregard the single document that describes the level of privacy they can expect from websites and software companies. In particular, the interface, by which these policies are communicated to the user, is addressed, as previous research has noted its many deficiencies. The survey results indicate that users have a desire for increased control over their own personal information and that barriers, such as long blocks of legal text, should be addressed in order to increase policy notice readership on the web.
POLICY NOTICE READERSHIP ON THE WEB

by

Michelle M. Lopez Miller

A Master’s paper submitted to the faculty of the School of Information and Library Science of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Science in Information Science.

Chapel Hill, North Carolina

November 2012

Approved by

_________________________
Bradley Hemminger
Table of Contents

Abstract .............................................................................................................................................. i
Introduction ........................................................................................................................................... 2
Literature Review ................................................................................................................................. 4
  Privacy Policy Backstory .................................................................................................................. 4
  Policy Presentation ........................................................................................................................... 4
  Readability of Documents .............................................................................................................. 6
  Readability of Policy Notices .......................................................................................................... 7
  Tactics to Encourage Systematic Choice ......................................................................................... 8
Method .................................................................................................................................................. 10
  Sample ............................................................................................................................................ 10
  Survey Instrument .......................................................................................................................... 11
Results .................................................................................................................................................. 12
  Demographics ............................................................................................................................... 12
  Readership and Accessibility .......................................................................................................... 13
  Understanding, Control and Trust .................................................................................................. 18
  Potential Improvements to Interface .............................................................................................. 19
  Priority of Privacy .......................................................................................................................... 23
  Data Tracking ................................................................................................................................ 23
Conclusions .......................................................................................................................................... 25
References ............................................................................................................................................ 28
Appendices .......................................................................................................................................... 32
  Appendix A: Email to GISA Personnel ....................................................................................... 32
  Appendix B: Email to SILS-Masters Listserv ............................................................................... 33
  Appendix C: Survey Instrument .................................................................................................... 34
INTRODUCTION

Both privacy policies and end-user licensing agreements (EULAs) are ubiquitous in today’s computing environment. Consumers are frequently being prompted to provide consent to websites’ or software’s terms of use. Despite consumers reporting high levels of concern over internet privacy, very few people have been found to actually read EULAs or privacy notices (Bohme & Kopsell, 2010; Cottrill, 2011). This creates an interesting question of how consumers truly regard their information privacy and what, if anything, should be done to increase policy notice readership.

Enormous amounts of consumers’ personal information is captured each day on the web. As part of the phenomena known as Big Data, every click, login, search and purchase can, and is, being collected by the hosts of mobile applications (apps), e-commerce sites, search sites, e-mail providers, and even the manufacturers of the smartphones themselves. The Wall Street Journal reported that a criminal investigation had been launched into the information gathering practices of several smartphone applications (Efrati, Thurm & Searcey, 2011). In fact, mobile apps can capture contact list information, make phone calls and track user location, all without the user knowing (Gralla, Sacco, & Faas, 2011).

The collection of user data utilizing web browser’s local storage, known as “cookies,” or similar technologies, does allow for a level of customization that many users have come to enjoy and expect. This includes personalized shopping recommendations, search results, radio playlists, news and more. At the same time, this
enables companies to track large amounts of personal information through cookies. Additionally, millions of people have joined social networks, voluntarily broadcasting personal details about themselves, engendering a colossal assemblage of user data; Facebook’s 950 million users generate approximately 500 terabytes of data daily (Kern, 2012). And although many individuals modify the privacy settings of their social media accounts to mirror their privacy intentions, and perhaps limit the dissemination of their personal details, Madejski, Johnson, & Bellovin (2011) found that most are not able to effectively control their privacy settings.

As the web becomes increasingly personalized, so does advertising. User data is collected or bought, and used to build profiles. Frequently, the result is ads that are specifically directed towards certain individuals or user groups. In 2010, online targeted advertising was expected to generate $25.8B dollars in revenue (Broder, 2011). Clearly, a lucrative market exists for amassing user data, and as the role of privacy on the web continues to be murky, it does not appear that there will be any deceleration in the harnessing of user data anytime soon.

Despite all of this, privacy advocates, researchers, legislators, as well as concerned users, continue to call for improvements in privacy within the digital landscape. User privacy on the web has been the subject of numerous studies (Friedman, Lin & Miller, 2005; McDonald, Reeder, Kelley, & Cranor, 2009; Madejski et al., 2011; Wu, Huang, Yen, & Popova, 2012), with the general consensus that consumers are indeed concerned about the use and collection of their information, but that the presentation of policy notices has failed to clearly communicate information practices, leaving consumers with an asymmetrical choice between spending large amounts of time
reading and attempting to understand complex and legalistic policies, or simply entrusting second parties with their personal information, and perhaps in the process, de-prioritizing their privacy concerns.

**LITERATURE REVIEW**

*Privacy Policy Backstory*


Although the FTC laid out their ideals surrounding data use on the internet, they did not provide a regulatory framework for fear it would inhibit the new marketplace. Instead, the FTC promoted self-regulation in the industry and some companies adopted a strategy of voluntary disclosure in an effort to satisfy the “notice” principle of the FIPs. However, due to the voluntary nature of these notices, no restrictions or guidelines for format, readability, content, or quality, have ever been put into place (Pollach, 2004; McDonald et al., 2008).

*Policy Presentation*

Policy notices have historically been presented using lengthy blocks of natural language and are filled with legalese. The content within EULAs and privacy policies has
been found to be written at a reading level equivalent to college level or higher (Reeder, 2008; Cottrill & Thakuriah, 2011). In the case of EULAs, they are almost always appear as a pop-up dialog box, whereas privacy policies are most often a web page or document. Theses notices can span multiple pages, such that the user has to scroll or click next to view the entire document, and are almost always black text, sometimes capitalized, on a white background. There have been little to no changes made to this mechanism over the last two decades of increasingly prevalent internet use by everyday people, in spite of various studies detailing intriguing design variations, such as Kay and Terry’s textured agreements (2010). According to Pollach, this is because companies are writing and distributing information policy disclosures not as a matter of consumer protection and fair data handling, but to diminish the threat of privacy litigation (2007).

Bohme et. al (2010) studied EULAs and notices as “interception dialogs” (p.2403). Since EULAs are often presented as a popup dialog box, a user’s task at hand is interrupted. In fact, both types of policy notices appear before the user is allowed to receive the product or services they wanted when they began the process that resulted in the policy notification. One example of this is when a user is downloading and installing software, a EULA pops up asking for user consent when the user’s primary concern in not reading a long document, but rather, completing the software installation. This, the authors argue, has played a significant role in user habituation; that is, users have become accustom to overlooking policy disclosures because they are seen as an interruption, not a useful or informative text.
**Readability of Documents**

Gribble (1999) studied the readability of informed consent documents. Much like privacy policy notices, informed consent documents are likely to be written by lawyers or subject matter experts. In the case of informed consent, the authors are also often doctors, scientists or researchers. This creates an education gap between the authors of the documents and the typical readers of the documents. This renders informed consent unobtainable, if in fact, readers are unable to comprehend the language of the documents.

While the research on presentation is somewhat focused on informed consent documents, Jane Watson, who trains business leaders in writing and communication, clearly lays out “12 Ways to Increase the Readability of Your Business Documents” (jwatsontraining.com). She emphasizes the importance of page layout, and details such as font and print density. Some of Watson’s recommendations include: a short first introductory paragraph; smart use of headings, bulleted lists, bolding, italicizing, and white space; and 5-line paragraph length. She is not alone: newspapers, book publishers and web page developers all pay attention to layout, white space and font, in addition to content and word choice.

Doak and Doak (2010) documented a variety of writing strategies to effectively communicate with people of all levels of reading. Their work was particular to the medical field, but their findings seem to be generally applicable to writing for any diverse audience, as the title of the article suggests (“Writing for readers with a wide range of reading skills”). They advocate using common vocabulary and simple sentence structure. If a less-common word must be used, the author should explain the meaning of that word with examples, interpretations or pictures. Before new information is given, context
should be presented, as this allows readers to transition from one section of information to another (Figure 1). Doak et. al also cite the importance of making the document look inviting by appropriate use of white space and a large-enough font size (e.g., 12). To further facilitate reader understanding, reader interaction should be encouraged. In the case of EULAs or privacy policies, reader interactions could potentially be achieved with a link to a FAQ or a “Contact us” page.

Figure 1: Context before new information.
The second example is more effective (Doak et. al, 2010, p. 151)

<table>
<thead>
<tr>
<th>Information given before context.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eating broccoli, spinach, collard greens, lettuce greens, rhubarb, Brussels sprouts, kale, beet greens, and carrots can reduce your risk of certain kinds of cancers.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Context given before information:</th>
</tr>
</thead>
<tbody>
<tr>
<td>The foods you eat affect your health. You can reduce your risk of certain cancers by eating broccoli, spinach, collard greens, lettuce greens, rhubarb, Brussels sprouts, kale, beet greens, and carrots.</td>
</tr>
</tbody>
</table>

**Readability of Policy Notices**

(2010), as described in the previous section. Namely, there are six PL guidelines: reader-centric organization of text; use of the word “you;” use of active voice; short sentences; use of common, everyday words; and “easy-to-read design features.” Although PL is a federal set of guidelines, some states have adopted similar strategies for improving document communications to the general public. In Washington state, the authors of its public policy documents are equipped with a “Reader-Friendly Document Toolkit” (Derthick et. al, 2007, p. 1). Specifically, the toolkit assists writers to effectively use headings, graphics, such as tables and sidebars, and lists.

The literature review of readability for various types of documents, including medical documents for patients, informed consent documents, and public policy documents shows common themes of what makes a document user-friendly and easy to read. It is the website and software privacy notices that continue to violate essentially all of the document presentation research recommendations and resist adoption of any potential improvements.

**Tactics to Encourage Systematic Choice**

Bohme et. al (2010) explored the policy notice features of coercive language and interruption by conducting a study using a live website, AN.ON/JonDonym (http://anon.inf.tu-dresden.de), and an optional client software update. They wanted to test the idea that modifying the EULA-like consent dialog message to a more voluntary style (in contrast to typically coercive language) would increase the participation level and the response latency (the time users spent with the form before consenting). The study used three specific variations of the dialog box: a neutral heading versus a polite heading using the wording “Please help us…”, typical button text (“I accept” and “I
decline”) and more voluntary language (“I take part” and “I do not take part”) and variations in the default button between the consent, objection and no selection.

All results were statistically significantly given the large data set (81,920 users), but among the interpretations of the results was that users responded positively to the voluntary nature of the “I take part” button text. Perhaps more interesting, the polite heading of “Please help us...” had a slightly negative response. The authors interpret this as a sign of the habituation of users accepting anything that looks like a EULA; when the form deviates from the typical language, users do not reflexively accept, but instead seem to think twice about what they really want to do. This depicts the broader idea of heuristic versus systematic users, where heuristics dominate the policy notice domain and users mindlessly accept due to lack of true choices and the deluge of notices constantly being presented to them.

Reinforcing this idea were the response latency results. Sensitivity to all three types of variations decreases as response latency increases, indicating a systematic path is strongly related to users taking more time with the form. In fact, at a response latency between 30 and 60 seconds (in contrast to less than 5 seconds on the low end), a polite heading becomes a positive aspect (participation increases), which is in line with social psychology research on altruism and helping tendency. That is, when users systematically process the EULA-like form, as indicated by an increase in response time, they respond positively to the altruistic implications of a polite request for help. Whereas when operating on heuristics, users respond negatively to the deviation of what they expect from a EULA and respond negatively, in contrast to established research on persuasion and helping tendency (Bohme et. al, 2010).
METHOD

This study was designed with the intention to generate data that describes how many people within the two distinct samples read policy notifications, as well as to better understand how users interact with, and comply with, user policy notifications and legal use agreements for software and websites.

Sample

Two convenience samples were utilized in the execution of this study. The first sample was individuals who worked at the Ground Intelligence Support Activity (GISA) agency. GISA is a government agency located in Fort Bragg, North Carolina that employs a high number of retired military personnel and veterans. The majority of GISA employees worked within the Signal or Intelligence branches of the military and now, as Department of the Army civilians or contractors, fulfill technical duties such as system administration, network management and help desk support.

The other sample was derived from the SILS Masters Listserv. As of October 2012, there were 394 masters students subscribed to the listserv, which includes both full and part-time students. Respondents were not asked if they were a part of the Information Science (IS) program or the Library Science (LS) program and no distinction was made between the two.

Although each of the samples has uniquely identifying characteristics that are almost certainly a deviation from a random sample taken from the general public, it was thought the inclusion of the two different populations might provide additional insight on the subject matter as compared to just surveying one narrow subject population like university students.
**Survey Instrument**

The survey was created and distributed over the web, using Qualtrics survey software. It was comprised of 22 questions, including two demographic questions: “Are you male or female?” and “What is your age?” The survey was designed to investigate several areas of interest. First, although numerous studies have reported low readership numbers for policy notices, this study sought to establish its own independent data to that effect or otherwise, so users were asked directly whether or not they read EULAs or privacy policies. Other questions then addressed accessibility of privacy settings, users’ perception of their own comprehension of privacy policies, users’ desire for control of their own information, users’ trust in second parties to handle their data, and interest in potential changes that could be made to the policy notice interface to increase the likelihood that they would read it.

The survey was distributed via two separate instances, one for GISA persons (Appendix A) and another for SILS students (Appendix B), each in October 2012. The tone of each e-mail varied slightly due to the nature of each audience. SILS students are likely to be familiar with surveys conducted within academic research, whereas a friendlier tone was more appropriate for GISA personnel who may not feel as comfortable with a research survey request. They survey was open for one week for GISA participants and five days for SILS participants.

All participants were given the opportunity to enter an e-mail address that would be entered into a drawing for one of two $25 Visa gift cards. The e-mail addresses were not affiliated with the survey responses in any way.
RESULTS

Of the 122 responses recorded by Qualtrics, 118 were complete surveys. (A completed survey does not indicate that all questions were answered, but instead that the respondent navigated through each question and reached the end of the survey, regardless of the number of questions answered.) Data was analyzed using the Qualtrics Reporting (beta) feature. Data was also exported to JMP, a SAS data analysis software, and Microsoft Excel.

**Demographics**

The first two questions were demographic questions regarding age and sex. Approximately two-thirds of the respondents were female and the average age was 29. The SILS sample was a younger sample, with the average age between 23 to 27, while 66.7% of GISA respondents were older than 41 years of age. This reflects the expected age range of Master’s students for SILS and the post-military retirement (or post-military service) age of GISA. Many individuals who work for GISA are in their second career.

<table>
<thead>
<tr>
<th>Age</th>
<th>Number of Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>18-22</td>
<td>8</td>
</tr>
<tr>
<td>23-27</td>
<td>51</td>
</tr>
<tr>
<td>27-31</td>
<td>19</td>
</tr>
<tr>
<td>32-36</td>
<td>9</td>
</tr>
<tr>
<td>37-41</td>
<td>6</td>
</tr>
<tr>
<td>&gt;41</td>
<td>25</td>
</tr>
</tbody>
</table>
Overall, the respondents were primarily female. 76% of SILS responses were from females, with 69% of total responses coming from female participants. This is likely a natural result of the high female to male ratio within SILS. More females naturally occur in the SILS convenience sample, while GISA has a more balanced gender distribution, both in the sample and the responses.

Table 1: Gender distribution of SILS responses

<table>
<thead>
<tr>
<th>Gender</th>
<th>Count</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>21.00</td>
<td>23.6%</td>
</tr>
<tr>
<td>Female</td>
<td>68.00</td>
<td>76.4%</td>
</tr>
<tr>
<td>Total</td>
<td>89.00</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 2: Gender distribution of GISA responses

<table>
<thead>
<tr>
<th>Gender</th>
<th>Count</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>16.00</td>
<td>53.33%</td>
</tr>
<tr>
<td>Female</td>
<td>14.00</td>
<td>46.67%</td>
</tr>
<tr>
<td>Total</td>
<td>30.00</td>
<td>100%</td>
</tr>
</tbody>
</table>

**Readership and Accessibility**

The next questions were to establish a benchmark of readership for EULAs and privacy policies. That is to say, what are the specific numbers of individuals in these samples who self-report to read policy notices? The types of notices were separated into two distinct questions because although they share several characteristics in terms of style, content, and challenges, EULAs are primarily presented for software installation as a pop-up dialog box and may be treated differently by users than privacy policies for websites that they may visit regularly. The data does in fact indicate that users are less likely, by 30 percentage points, to read EULAs than privacy policies, with nearly 90%
responding that they do not read EULAs when installing software. SILS students are less likely than GISA respondents to read EULAs by 9 percentage points.

Figure 3: Do users read privacy policies or EULAs?

Over one-third of all respondents reported that they read privacy policies for social media, e-mail or online banking. 40% of GISA individuals reported that they read privacy policies, although this is higher than regularly cited numbers (McDonald et al., 2009). Log file analysis showed that out of 55,158 sessions, 131 (0.24%) people visited the privacy policy during the registration of a website (Jensen & Potts, 2004).

Following the privacy policy readership question, display logic was implemented so that users who responded yes were routed to a question about how long they spent reading and those who responded no, were asked why. For those individuals who self-reported to read privacy policies, 95% reported spending longer than 5 seconds reading. Approximately half reported spending greater than 30 seconds with the policy notice.

The results, when analyzed by age groups, demonstrate that the mean falls within the 5-30 second range, with a standard deviation of .59. While almost half of the
responses are normally distributed across the 5-30 second range, with most of the represented (i.e., non-zero) age groups reporting a 10-25% reading time for all three choices, there are two outliers. A very high percentage of all respondents older than 41 self-report to spend longer than 30 seconds reading. Similarly, nearly 40% of all 18-22 year olds reported to spend at least 5-30 seconds reading the policies (Figure 4).

In 2008, McDonald et al. conducted a study using the privacy notices of the 75 most popular websites, and an average reading time of 250 words per minute, to calculate the time required to read a privacy policy in full as 10 minutes. It is a potential limitation of this study that the selections presented did not measure reading times at a more granular level above 30 seconds, such that a determination could be made as to how many respondents skim policy notices and how many read the notices in their entirety.

Figure 4: When users self-report to read policies, how much time do they spend?

Several choices were presented to those who do not read privacy policies. Users were asked to select the number one reason they choose not to read policy notices. The
choices were drawn from literature citing the most commonly named barriers to user readership (Anton, Earp, He, Stufflebeam, Bolchini, & Jensen, 2004; Cranor et al., 2006; McDonald et al., 2008; Reeder, 2008). These include long blocks of text, legal language, cost of time, inability to affect policy, and the ability to locate the policy (accessibility). Procrastination was also added, although not cited by the authors above. The respondents were also given the opportunity to write in an answer and six people (8%) did so. Four of those referred to reasons that were included in the choices, but either desired to list more than one reason or add a slight variation on the given choice (e.g., “writing is very small and the content is too long” and “I feel they are too complex and full of legal loopholes for me to truly understand what the implication of these policies are”). Two exceptions to the given choices were “I’ll hear online if there’s a big problem with them” and “Do not care.”

Figure 4: Reasons users do not read privacy policies

<table>
<thead>
<tr>
<th>Reason</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Long blocks of text</td>
<td>36%</td>
</tr>
<tr>
<td>Legal language</td>
<td>17%</td>
</tr>
<tr>
<td>I am too busy to read</td>
<td>8%</td>
</tr>
<tr>
<td>Why read when you can’t change them</td>
<td>29%</td>
</tr>
<tr>
<td>Plan to read it but never get around to it</td>
<td>3%</td>
</tr>
<tr>
<td>Cannot find it</td>
<td>0%</td>
</tr>
<tr>
<td>Other - please specify</td>
<td>8%</td>
</tr>
</tbody>
</table>
The only choice not selected by any of the respondents was that they could not locate the privacy policy in order to read it. Echoing this trend, accessibility was not reported to be an issue by the majority of respondents in the next question regarding the accessibility of privacy settings. Meanwhile, the majority of users stated that policy settings, while accessible, are not easy to understand. This follows with the notion that simply fulfilling the notice FIP has no guarantee with respect to quality or content (Pollach, 2004; McDonald et al., 2008). However, despite the potential of a comprehension barrier, nearly all respondents (108 of 118) reported that they had modified privacy settings of their social media or other types of online accounts at least once, with 100% of the youngest generation polled having modified their privacy settings (n = 8). These results are aligned with the findings of Madden (2010), where the majority (71%) of social network users had modified privacy settings to restrict their profile’s visibility.

Figure 5: Are privacy policy settings accessible? Are privacy policy settings easy to understand?
Understanding, Control and Trust

The next three questions pertained to the users’ perception of their own understanding of how second parties handle their personal information. With regard to social media, a little over one-third reported that they felt they understood how their personal information is used. Slightly less reported the same for their e-mail providers and less still for mobile applications, with an 11 point difference between understanding of use by social media and mobile applications.

Table 3: Users’ understanding of how web service providers handle their personal information

<table>
<thead>
<tr>
<th>Understand personal information handling</th>
<th>Social media</th>
<th>E-mail provider</th>
<th>Mobile application</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>38%</td>
<td>31%</td>
<td>27%</td>
</tr>
<tr>
<td>No</td>
<td>61%</td>
<td>69%</td>
<td>73%</td>
</tr>
</tbody>
</table>

Although 38% of respondents reported to understand how social media uses their personal information, only 16% said that they trust social media with their personal information. There appears to be some uncertainty, even amongst the users who feel comfortable in their understanding of social media’s data policies, that the companies are actually doing what they say they are doing with user data.
91% of survey takers believe their personal information is sold to third parties and an equally high number wish they could have greater control over how their information is treated (93%). This corresponds to Westin’s findings that the ability to control one’s own data is paramount (Madejski et al., 2011). However, when asked about a variety of changes that could be made to the policy notice interface in order to potentially increase the likelihood that the user would read the policy, results were mixed.

**Potential Improvements to Interface**

Four options were given, within individual questions, as possible improvements to the current design and presentation of policy notices: translations from legal language into easily understood everyday language, color in the interface, graphics in the interface and e-mailed policies. Generally speaking, the need for everyday language in order for users to be able to understand content is well documented (Gribble, 1999; Doak et. al 2010; Derthick et. al, 2009). If jargon or legal language is foreign to readers, and no alternate text is present, that is a significant obstacle to reader comprehension, even for high-level readers.
In the case of graphics used in documents used to communicate to medical patients, research has found that reader understanding, recall and compliance can increase up to 500% when pictures are included with text, helping to interpret its meaning (Delp & Jones, 2008; Houts, Doak, C., Doak, L., & Loscalzo, 2006).

The option of policy notices being e-mailed to the user was introduced in order to test whether a facilitation of accessibility would encourage greater numbers of individuals to read and spend time with the policies. This would disrupt the interruptive nature of policies and allow the user to not only read the policy at a time of their choosing, but also to maintain a copy either within their e-mail inbox or on their personal computer. The results of this question further validate earlier described results (Figure 4, Figure 5) which strongly indicate that for these samples, accessibility is a non-issue. Zero respondents stated that the reason they do not read policies is that they cannot find them and 84% said that privacy settings are accessible.

Of the four options, translations for legal language garnered the greatest number of yes responses with 93. An interface that included some graphics earned slightly less than 50% of yeses, but the majority of survey takers did not feel that the other two options of color and e-mails would encourage increased readership.
When the results are examined by sample group, there are two significant disparities. Both differences in opinion had to do directly with visual representation: color and graphics. While 70% of GISA personnel indicated that the addition of graphics to a policy notice would make it more likely that would read them, only 42% of SILS students felt the same way. Color in the interface had an even wider gap: 70% positive from GISA and only a 27% positive response from SILS. Further research would be necessary to parse out the reasons for this difference, as well as to determine whether these two samples are representative of the general population. However, there is a good probability that the difference is reflective of a gap in education level between the two groups. Although education level data was not collected from GISA, the National Assessment of Adult Literacy found that 40% of American adults were at proficiency level 1 (http://nces.ed.gov, 2003). According to Doak et. al, proficiency level 1 is approximately equivalent to a 6th or 7th grade reading level. It is highly likely that even if most GISA respondents do not fall within the 40% of literacy proficiency level 1, they
have a lower education level than SILS Master’s students. As reading ability goes down, the importance of supplementary ease-of-reading tactics is more critical.

Both GISA and SILS survey takers responded positively to the possibility of the inclusion of legal translations for the dense legalese of which policies are currently comprised. This corresponds to the reasons users rated as the greatest barriers. Although length and inability to change the policies were the top reasons users selected, the length of the policy is not easily altered without compromising its content and the level of control may be better tackled via policy changes. It is the legalese that could be reformed within the presentation design, as demonstrated by Kelley, Bresee, Cranor, and Reeder with their Nutrition Label design (2009), and Kay et al.’s Textured Agreements (2010). Creative Commons is an example of a working framework for a standardized, yet customizable, document that is layered to maintain its legal integrity and at the same time, also provide a “human readable” layer (creativecommons.org, 2012). The three layers (public, legal and machine readable) make for easy utilization by the non-expert public for declaring rights for creative works.

**Figure 8: Which changes would increase likelihood of policy readership by sample groups**
**Priority of Privacy**

In order to gauge users’ priority of privacy with regard to social media, survey takers were asked directly if the ability to use their social media account overrides their privacy concerns. The results were almost split down the middle, with 65 individuals reporting that yes, it did, and 52 reporting that it did not. When broken down by age, a distinct trend emerged: nearly 80% of those between the ages of 18 and 22 prioritized social media over privacy while barely 20% of those older than 41 felt the same way. This appears to indicate that the role and understanding of privacy is evolving dramatically with new generations.

![Figure 9: Privacy prioritized against the ability to use social media by age](image)

**Data Tracking**

Three questions were asked in order to establish how comfortable users are with website and application data tracking. The first question asked if it was ok if an application knew the user’s location if the purpose of the application was to assist him to
find a restaurant. The majority (67%) responded that location tracking was acceptable for that purpose.

The next question asked if it was ok if e-mail providers knew the purchasing habits of their users. The implication behind this question is the ongoing practice of Google’s Gmail to derive the “concepts” of an e-mail by examining the e-mail header, body and addressing information for the purpose of targeting ads (epic.org, 2004, 2.2). Only 22% agreed that it is ok if e-mail providers knew their purchasing habits.

Users were nearly split when data tracking is anonymized and purely demographic, with 56% reporting that it is an acceptable practice. It is possible users uncomfortable with anonymous demographic data tracking are not confident in anonymization efficacy. Sweeney (2000) established that 87% of all Americans could be uniquely identified when only zip code, birth date and sex are known. Ohm (2009) and others confirm that reidentification or deanonymization are much easier done than is often assumed by the general population.

Figure 10: Users report comfort with data tracking methods
CONCLUSIONS

In an 1890 *Harvard Law Review* article, “The Right to Privacy,” Samuel Warren and Louis Brandeis argued that individuals’ “right to be let alone” should be recognized by the law (Friedman et al., 2005). In today’s age of social media and boundless electronic connections, it seems that many people’s “right to be let alone” has completely transformed into something else entirely. A delicate balance is required for a fast-paced, high tech, personalized web to coexist with some evolving semblance of digital privacy, which is at least in part defined by the actions, or inaction, of consumers.

The gap between consumer concern over online companies’ information practices and consumers’ recurrent choice to disregard policy notices, is certainly indicative that policy notices are unreasonably prohibitive (Vila, Greenstadt, & Molnar, 2003; Jensen et al, 2004). Privacy advocates, including customers themselves, need to recognize that any modifications proposed need to result in a lower cost for users to read policy notices, as the current value users are placing on their data, however superficial, is far lower than the value of the 10 minutes, or more, of their time it would take to read and analyze privacy policies. Users simply do not equate the value of a virtually invisible and innocuous exchange of their personal information for a quality web experience as an unreasonable deal. This is especially true with the younger generations, as their value of the privacy of their personal data appears to be declining. For the youngest of users, sharing even the most intimate information about themselves online is the norm. And although 100% of 18-22 year olds surveyed stated that they would like greater control of their information, they are unwilling to forgo the social media experience that is such an integral part of their everyday lives (Figure 9).
Consumer personal information is being collected in a very quiet and non-threatening way, and as long as that continues, consumers will continue to “sell” their personal information to websites and software companies without full understanding of the implications of that lopsided exchange. But how does that collection of valuable, often personally identifiable, information remain so quiet? It is because the information that users want and need to know is literally hidden underneath a cloak of impenetrable legalese and an uninviting display of text. Yet, there has been essentially no adoption of modified privacy notice presentations or proposed standardization mechanisms (e.g., P3P). One of the foremost reasons for this is that there is very little incentive, economic or otherwise, for software and online companies to provide users with a better policy notice experience. They are checking the FTC boxes of notice and choice with little to no pushback from consumers. Although there have been lawsuits questioning the information practices of some of the online companies, there has not yet been any single significant legislative action with the ability to change the status quo. It is unclear what could an impetus for change, other than legislature coming as a result of legislative action or a push from privacy advocates that spurs action in Congress.

The results of this study show there are at least two improvements that could be made to the current policy presentation to increase readership by making policies more user friendly: the inclusion of translations for legal language and the inclusion of graphics. Other industries, such as the medical field and the public sector, have implemented various versions of user-friendly document practices, but they have public interest ties that are dissimilar to the business interests of online social networks and
software companies. In fact, better communication in the medical and public sectors is likely to lower costs and reduce mistakes. This is not necessarily true for the web.

The study’s data shows a clear desire for greater control of personal information. The question remains what is the appropriate mechanism to allow for greater control, while making economic sense to both the consumer and the market, but a first step in increasing control is to provide users with an actual choice of who they allow to steward their data. One obstacle to the innovation of such a mechanism is that there is ostensible distrust in social media and other second parties’ information practices (Figure 6), but there is no corresponding reluctance to use those parties’ products.

Westin’s 21st century definition of privacy is the right of an individual to affect what information about himself is known to others (2003). Privacy disclosures in their current state are severely inhibiting consumers’ ability to protect and guard their digital privacy and all reasonable effort should be made to redress what Greenstadt et al. have determined to be a lemon market of privacy (2003). Relatively easy to implement presentation changes have the potential to completely reshape users’ interactions with digital privacy documents. Consumers should not have to settle for unknowing consent and unreadable privacy notices on the web.
REFERENCES


Creative Commons (2012). Retrieved from http://creativecommons.org/licenses/.


APPENDIX A: EMAIL TO GISA PERSONNEL

Hello all,

I am conducting a survey as part of my Master’s thesis. The survey is about reading privacy policy notices on the web. It is a short 5 minute survey. Any volunteers will be entered into a drawing for a chance to win one of two $25 VISA gift cards. This survey will remain open until Friday, October 22, 2021 at midnight.

Click here for the survey: Privacy Policy Survey [https://arc.qualtrics.com/MyQualtricsSurveyEngine/?Q.s=8e1Feh8mW7y7t3Jm7V3uwV8snu6_MfRZr2VbI74OaQa4v8.a>]

Your time and consideration are greatly appreciated!

Thank you!

Michelle

Note: If you would like to pass this email forward to others, respond to this e-mail and I will send you the open link.

If link in the body of the email does not work, you can copy and paste the URL below into your internet browser:
https://arc.qualtrics.com/MyQualtricsSurveyEngine/?Q.s=8e1Feh8mW7y7t3Jm7V3uwV8snu6_MfRZr2VbI74OaQa4v8.a

Follow the link to opt out of future emails:
Click here to unsubscribe [https://arc.qualtrics.com/CP/Register.php?Op=Out+true&d=Mj8=ZryY76W6i678v8R1OobW+47753whMqcomKQ8.a]
APPENDIX B: EMAIL TO SILS-MASTERS LISTSERV

Invitation to participate in a Masters Paper Study about privacy policies on the web

Michelle Lopez [noreply@qemailserver.com]
Monday, October 08 2012 10:31 AM
To: [SILS-masters@listserv.unc.edu]

- You replied on 10/8/2012 10:56 AM.

Hello fellow SILSers,

Please consider taking part in a study about privacy policy readership on the web. This survey will take approximately 5 minutes of your time and is open until this Friday, Oct. 12, at midnight.

All participants will be registered in a drawing to win one of two $25 Visa gift cards.

Click here for the survey: Privacy Policy Survey

Best regards,
Michelle Lopez
SILS MSIS Candidate
Primary Investigator
(919) 816-1828

If the embedded link does not work, copy the following into your browser:
https://unc.qualtrics.com/WQRQualtricsSurveyEngine/?Q_DL=86ekcDBdkdablec9H_bh5TV3ug5b1xweV_MLRP_43exx13385e977&_=1

Follow this link to opt out of future emails:
Click here to unsubscribe

- The information from this list is official communication for SILS masters students.
APPENDIX C: SURVEY INSTRUMENT

Informed Consent

Time: This survey consists of 22 short questions, which might take between 5 and 10 minutes of your time.

Risks: There are no anticipated risks in participating in this study beyond those you may encounter in day-to-day living.

Benefits: You may feel satisfaction in knowing you are contributing to research, but there are no direct benefits to you.

Compensation: You will receive no payment for your participation in this study, however, all volunteers will have the opportunity to enter themselves in a drawing for one of two $25 Visa gift cards.

Privacy: By participating in this survey, you are giving your consent for your responses to be used in this study. There will be no identifiers collected that may be used to associate any responses to an individual. This means that no names, IP addresses, e-mail address or any other identifier, will be stored or used as any part of this study, all data will be anonymous. Any data that is collected will be used in a responsible manner.

Voluntary: Your participation in this survey is completely voluntary. You may skip any questions you do not wish to answer. You may change your mind at any time. Should you decide not to participate, you may exit the survey even if you have already begun.

Questions: The researchers conducting this study are Michelle Lopez Miller and Dr. Brad Hemminger. If you have questions at any time regarding this study, you can reach Michelle Lopez Miller at lopezmm@ive.unc.edu or (919) 816-1828. You can reach Dr. Hemminger at bhemngt@unc.edu or (919) 966-2818. Additionally, if you have questions or concerns about your rights as a research subject, contact the University of North Carolina at Chapel Hill Institutional Review Board (IRB) at IRB_subjects@unc.edu or (919) 966-3113.

Consent: If you would like to continue and take this survey, please click "I agree" below. You may also print this page for your reference.

Print

- I agree

Quit

0% 100%
What is your age?
- 18-22
- 23-27
- 27-31
- 32-36
- 37-41
- >41

What is your sex?
- Male
- Female

When installing software or applications do you read end user licensing agreements?
- Yes
- No

Do you read privacy policies for social media, e-mail or online banking?
- Yes
- No
Are privacy settings accessible to you?
- Yes
- No

Are privacy settings easy to understand?
- Yes
- No

Have you ever modified privacy settings for your social media or other types of online account?
- Yes
- No

Do you feel you understand how your personal information is used by social media companies?
- Yes
- No

Do you feel you understand how your personal information is used by e-mail providers?
- Yes
- No

Do you feel you understand how your personal information is used by mobile applications?
- Yes
- No
Do you trust social media with your personal information?
- Yes
- No

Do you think your personal information is being sold to third parties?
- Yes
- No

Do you wish you had greater control in how your information is treated?
- Yes
- No

I would be more likely to read policy notices if there were some color in the interface
- Yes
- No

I would be more likely to read policy notices if there were some graphics in the interface
- Yes
- No

I would be more likely to read policy notices if they were e-mailed to me
- Yes
- No

I would be more likely to read policy notices if there were translations for the legal language
- Yes
- No
The ability to use my social media account overrides my privacy concerns
- Yes
- No

It is ok if an application knows my location if it helps me find a restaurant
- Yes
- No

It is ok if my e-mail provider knows my purchasing habits
- Yes
- No

If it is ok if an application knows my demographic information if it is anonymous
- Yes
- No

We thank you for your time spent taking this survey.
Your response has been recorded.