An Analysis of Web Privacy Policies Across Industries

An Interactive Qualifying Project Report

Submitted to the faculty of

Worcester Polytechnic Institute

In partial fulfillment of the requirements for the

Degree of Bachelor of Science

IQP-CEW-1201

By

Joseph D. Brown

Mohammad A. Ghani

Muyeedul Hoque

Umair A. Rehman

Project Advisor: *Professor Craig Wills*

Abstract

We undertook the task of evaluating aspects of privacy policies for online industries. We assessed a total of hundred websites using a fixed methodology, which we created based on our background research, and then applied that methodology to each privacy policy. There are different kinds of privacy policies and they all dealt with the factors of collection and protection of data and the magnitude of disclosure of information differently. We investigated the relation between user information and the role of third-parties, affiliates and cookie-based targeted advertisements. We found many trends in online privacy policies like the location of the link of privacy policy, the mention of last revision date, etc. Our analysis also showed that user information is quite vulnerable as most companies pass information to both affiliates and third-parties, which in most cases includes personal information. There are still companies present online without online privacy policies. Readability scores are quite consistent across websites, and they have not varied significantly over the years. However, the privacy policies of healthcare and dating websites are particularly difficult to read. Evaluating specific results and analyzing various online industries truly conveyed how a specific privacy policy behaved within its industry and in the web ecosystem as a whole.

Table of Contents

1. Introduction

Websites typically have partnerships with or are attached to larger or smaller companies and other websites; these are considered as their affiliates. Third parties are parties not directly owned by the websites but might have a contract or agreement with the website. Affiliates and third parties play a significant role in governing a website's privacy policy and are mentioned in its text time and time again. The privacy policy of a website is a statement or a legal document that discloses some or all of the ways affiliates and third parties use, disclose and manage a customer or user's personal and non-personal information. Personal information can be used to identify an individual. It is not limited to but includes; name, address, date of birth, marital status, contact information, credit information, financial records, medical history, travel information, and intentions to acquire goods and services. In the case of a business it is often a statement that declares a party's policy on how it stores, releases and uses the personal information it collects. It informs the client what specific information is collected, and whether it is kept confidential, shared with partners, or sold to other firms, enterprises or third parties.

Most company websites are obliged to have a privacy policy because grounded in it is the user's physical and moral autonomy. For this reason it is worthy of protection. The exact contents of the privacy policy depend on laws and legal jurisdictions, which may be influenced by geographical location. Most countries have their own set of laws and guidelines governing the collecting and sharing of information. But in most cases it is essential for privacy policies to address key issues regarding the use of personally identifiable information in order to rightfully inform the users of how and why their information is being disclosed.

As the world rapidly heads towards a digital era, our interactions, transactions and relaxations are heavily dependent on the use of the Internet. Here on most occasions users are prompted for a username and password to access some of the featured services a website offers.

While disclosing this “log in” information, users also grant the website authority to do whatever it claims it can, under its privacy policy with the shared personal information. Many customers and users tend to disregard this fact and use the services in oblivion. But with most companies heavily relying on their online customer base and an increasing amount of information being shared over the web, this fact is becoming an alarming realization for most faithful Internet users.

Website privacy policies, while playing an important role in relaying critical information between the website and user, are not always available or easily accessible. Since it has been established that privacy is a key issue, it is important for these policies to be thorough, well organized and easily understandable. A privacy policy can be governed through numerous factors. While deeming a privacy policy to be “good” is subjective, there are certain crucial aspects that it must entail to effectively convey its message to the user. Analyzing the effectiveness and clarity of these aspects in a privacy policy forms the basis of our project.

The purpose of this report is to provide a road map of our approach and results for this project. Chapter 2 discusses the types of researches that were conducted in order for the team to get a better idea of what privacy policies are and how are they analyzed. Chapter 3 sums up the goals that were determined through research while in Chapter 4 a methodology of analyzing privacy policies is determined and the questions are listed. Chapter 5 talks about the type of data that was used for analysis. The data comprises of 100 websites. 10 websites are selected from 10 different website industries. These industries include Financial, News, Adult, Social, Blogging, Search Engines, Job, Gaming, Healthcare and Dating. Chapter 6 displays our results individually for each question in the methodology while in Chapter 7 a discussion is made on these results on a categorical and overall basis. Chapter 8 highlights the takeaways from the research conducted

and the conclusions that can be drawn from our project. Chapter 9 is a follow up to our project which discusses the possible future work that can be done if a different group is to undertake a line of work similar to this project.

2. Background

2.1. Introduction

The project was initiated with a basic idea of understanding the approach of web privacy policies toward data protection and usage of user information. For us to understand how user information is dispensed we had to study in detail the existing privacy policy management systems and available information related to privacy policies. To serve the purpose we started of studying information on websites and articles. Then we came up with article reviews and website reviews for further analysis and setting out the direction for the nature of our research. We investigated the relation between user information and the role of third-parties, affiliates and cookie-based targeted advertisement, which later became a key element of our research.

2.2. Article Analysis

The following articles and papers were read as part of our study to get a better understanding of previous work done in our research field:

2.2.1. **Privacy in information technology: Designing to enable privacy policy management in organizations** (Karata, July 2005)

This article studies the characteristics that are imperative in designing of a privacy policy management system in general organizations. The article repeatedly highlights the fact that with the increasing diffusion of information, the need for a robust privacy policy is indispensable to cater to privacy concerns. The article elucidates the basic privacy policy structure which includes the four elements (data user, data element, purpose, use, condition, and obligations) that are crucial in giving shape to a generalized privacy policy. The question it inquires as a part of its primer is why companies are enthused to do sweeping privacy research and answers it by providing a simple case study that fortifies the fact that every organization has different interests

to protect and privacy policy is used as a deterrent in most cases. It is important that all the concerns of an organization are fully dealt with or else due to the interlinking of paradigms; a privacy policy could end up being futile in the long run.

The designing of privacy policy is initiated by understanding the needs of the organization followed by the research-based detailed interviews that are essential in the development of the privacy policy workbench prototype that is the third step. In the third step the team designed and established a prototype with the general objective of providing organizations a tool to help them create reasonable privacy policies, tie their written privacy policies with the enactment of the policy across their IT configurations, and then support them to scrutinize the execution of the policy through internal audits. In Step Four the team steered an empirical usability laboratory test of two methods of authoring policies. According to them the laboratory test was an important section of the overall examination in helping to justify the value and improve the overall design of each authoring approach. These were the basic steps discussed in the preparation of the privacy policy followed by minute details of how system architecture could be changed in synchrony with the privacy policy and how research and evaluation could overall increase the quality of the policy.

Since this article deals with the designing of a privacy policy management system it is a little unconnected to the kind of research we would do for our IQP but this article does give great insights into the main elements of a privacy policy and is a great resource of information to keep in observance while we would be forming a methodology and assessing privacy policies.

2.2.2. **What's wrong with online privacy policies?** (Pollach, September 2007)

This paper examines the imperfections in online privacy policies. It illuminates the data privacy concerns that have augmented amongst Internet users due to the easy accessibility of assortment and scrutiny of data. The article explains how people actually draw online

transactions by having a rich privacy policy regarding data handling operations, which can diminish customer uncertainties and hence actually attract more business. The article is trying to expound that customer 'trust' is deeply related to how the privacy policy is exercised and 'credibility' plays a critical role in the process. The article then surveys the weaknesses in online privacy policies. Missing a privacy seal, inadequate coverage of privacy issues and scanty use of language were some hitches that were ascertained by the author. A crucial point that was raised by the author was that some online privacy policies were conscripted with the risk of privacy litigation in mind rather than a sincere effort to improve the data handling practices. Many of such companies have defectively defined policies that negatively affect the quality of the overall organization.

This article is related to our IQP since it informs us about imperfections in Privacy Policies. It educates us in the sense that it apprises us about the crucial factors we should be looking for once we evaluate privacy policies. Our methodology would definitely discuss some of the information provided in the article. Apart from this, the article also discusses some data handling operation which were not related to our area of focus.

2.2.3. **A Privacy Policy Model for Enterprises** (Schunter, 2002)

This article studies the privacy policy model for enterprises. It serves as the foundation of a control system that has to manage received data in accord with the privacy policy henceforth the data subject providing his/her personal data has the assurance that the enterprise receiving the information will handle it according to the stated privacy policy. The enterprise as well can authenticate that its business practices are not in conflict with the privacy policies declared by them. The privacy policy model described by them combines user consent, obligations, and distributed administration.

The privacy policy by them does not necessitate a transformation of the IT infrastructure. It is a simple and easy implementable policy that deals with data in mostly single forms. The policy description, which is allied from data systems such as 'relational databases or object-oriented systems', also allows personal data to be condensed with its related privacy policy when passed to other companies, such that the admittance control system of the delivery company can interpret and reinforce the policy.

This article discusses an ideal privacy policy model for small companies. It is not directly related our IQP, but discuses some of the key components that a privacy policy should address. It provides great information on accessibly, comprehensibility, user approval and channeling of information which would be some of the key components that we would be discussing in the later part of the project.

2.2.4. Privacy Policies as decision-making tools: an evaluation of online privacy notices (Jensen & Potts, 2004)

This paper gives a comprehensive outline of the issues related to usability of online privacy policies, gathered from two different sets of data. One set of data is from high-traffic websites and the other is from healthcare websites. The authors gathered massive data from reliable sources and analyzed the results in accordance to usability issues. It seemed most websites had accessible links to their privacy policy page, even if it was not on the home page. However, based on a readability standard, they found only a handful of privacy policies which were completely understandable by adults without high-school education, for both sets of data. In fact 54% of the policies were beyond the grasp of 56.6% of the Internet population based on their average level of education. The number of regulations which are enforced continue to increase, which generally make these privacy policies *harder* to read and grasp. Regarding policy content, it appeared many websites did not notify users of policy change and 69%

required users to check the page to notice changes. Amongst the latter group, many did not even post a modification policy or date. Generally there was not a communication forum to discuss matters and users were probably unaware of several policy issues and changes over time and how their previous and current data were used.

Although the content of this paper is derived from outdated data, the material presented in this paper is relevant to our IQP. Accessibility and usability are two of the most important aspects of any privacy policy and how it is communicated to the users. However, our project will probably go far beyond these issues merely and also look into *how* the information derived from users is being transmitted and utilized, an area that is out-of focus for this paper. However, the data analyzed in this paper is a great starting point for our project.

2.2.5. An Analysis of P3PEnabled Web Sites among Top 20 Search Results

(Egelman, Cranor, & Chowdhury, 2006)

The Platform for Privacy Preferences (P3P) is a standard format developed by W3C around 2002 recommending websites to release intended use of user's information. It is also the only format that encodes machine readable privacy policies. This paper analyzed data from popular search engine searches to find relative websites who use P3P privacy policies. In addition, they used a customized search engine developed by the authors called Privacy Finder. This resulted in searches which also indicated if each website's P3P privacy policy was in accordance with the user's preference. So a user was aware of the privacy guidelines before actually going on the website. But only about 10% of the searches had a P3P privacy policy, which meant the users did not have a lot to choose from. Certain industries like government and e-commerce websites used it more often due to regulations and industries like pornographic websites barely used it as customers were probably indifferent about it. Around 6% of all P3P

based privacy policies contained critical errors. But overall there has been an increase of P3P based privacy policies over the years.

As this paper mainly deals with search engines' output related to P3P privacy policy data, it is not too relevant to our IQP. Some of this research may be helpful to understand the more recent regulations in place and where technology is driving towards. Our focus should be on how much information is collected and distributed by these P3P privacy policy websites and possibly the errors they contain within their policies.

2.2.6. Privacy and Rationality in Individual Decision Making

(Acquisti & Grossklags, 2005)

This paper dives into the psychological factors that affect users' attitude on privacy issues. The authors conducted experiments to find more detail on what affects individuals' perceptions to privacy policies. Several factors like incomplete information, bounded rationality and systematic psychological deviations from rationality seem the dominant factors. Although most seemed worried about current privacy policies, it seemed most were not aware of implications of sharing basic information and the rights websites have in transmitting that information. Everyone perceives privacy differently and how much trouble they are willing to go through to protect their privacy. Many confuse security with privacy as well. Even if individuals are aware of the existence of privacy policies, they tend not to go through great trouble to read the whole material. Individuals are willing to trade-off sharing of private information with small rewards. And most of the educated surveyors suggested a need for government regulation over current privacy issues.

This paper, although is extremely interesting, is not relevant to our IQP. The factors considered in this research are purely behavioral or economic, none of which pose any

significant relevance to our project. However, this angle may be helpful for us to consider if we deal with individual behavior and actions regarding privacy policies.

2.2.7. The Lack of Clarity in Financial Privacy Policies and the Need for Standardization (Anton, Earp, He, Stufflebeam, Bolchini, & Jensen, 2004)

This paper discusses the lack of standardization and some of its implications in the financial industry. The authors analyze the privacy policies in accordance to the Gramm-Leach-Bliley Act (GLBA) – the most comprehensive financial privacy legislation till 2003.

The authors employed a goal-mining technique to analyze all the privacy documents. The used Goal-Mining to extract goals from privacy policies and listed out all the distinct goals from all the policies. Analyzing certain keywords noticed in privacy documents, they reached a conclusion that several of these are either not fully defined, which results in confusion amongst users, or present conflicts with other keywords and goals. They also hypothesized that the number of privacy protection goals (notice and awareness; choice and consent; access and participation; integrity and security; and enforcement and redress) will be greater than the number of privacy vulnerabilities, e.g. privacy invasions, and this was confirmed. This result could be due to the presence of growing financial regulation concerning privacy policies.

The analysis also showed several conflicts regarding privacy issues presented in the several privacy policies, even of those by the popular websites. Concerning readability, the conclusions made in this paper were similar to the ones made in "Privacy policies as decision-making tools: an evaluation of online privacy notices". Both the papers had used a common standard to evaluate – The Flesch Reading Ease Score (FRES) model. It seemed that the general privacy policy required a higher education level than that of the average Internet user.

2.2.8. Reading Level of Privacy Policies on Internet Health Websites

(Graber, D'Alessandro, & Johnson-West, 2002)

Online privacy policies while proving easy to understand for some can be difficult for others. Most policies require a college reading level and an ability to crack through confusing phrases and meaningless jargon. Therefore the reading level of any privacy policy is quite important to its comprehensibility and effectiveness. This paper examines the reading level of privacy policies on different health websites. These websites range from the top 25 Internet health sites to less used sites which users stumble upon while researching common health issues. The readability level of these websites was calculated using the Flesch, the Fry and the SMOG readability levels.

The relevance of this paper to our project is that we should analyze the reading levels of the privacy policies we study which as some point we surely will, as it is a key factor in conveying the privacy policy of a website to the user. And also that our data that is collected should be as diverse as possible, ranging from popular and highly rated websites to ones that are less commonly used. This approach will take into account the difference and also provide a comparable analysis of a diverse set of data incorporating various types of websites.

2.2.9. Consumer Trust, Perceived Security and Privacy Policy: The Three Basic Elements of Loyalty to a Website

(Flavian & Guinaliu, 2006)

This paper analyzes the effect of a good privacy policy and perceived security on the level of trust shown by the consumer on the Internet. As one can predict that an individual's loyalty to a website is closely linked to the levels of trust. This development of trust not only affects the intention to buy but also affects the purchasing behavior and frequency of visits of the

user to the websites, therefore the level of profitability each consumer brings. This paper deals with some psychological aspects and influences that privacy policies can have on users. Good privacy policies would be clear and readable in language for the user to easily comprehend; frequent updates when changes in policies occur via email/messaging/call could ensure a higher level of satisfaction; a direct link of the policy on the main page of the website could grasp the user's attention and psychologically force the user to think that the website has a ubiquitous and effective privacy policy.

This paper also discusses how the relationship of a user with a website is affected by the privacy policy and how it can be made stronger using various methods ensuring a greater trust. The user/website relation is a complex one and can be tampered with in ways that can determine higher or lower profitability to the website. As privacy policies differ from website to website, the concern for these policies also differs from individual to individual, depending on the user's level of education, the website content and type and other various factors, all adding a whole different perspective to this relationship.

This article although interesting has little relevance to our project. It focuses more on the user/website relationship and how this relationship is influenced by different factors. This article can be helpful for companies in order for them to determine what aspects of privacy policies influence users and how these aspects can be modified to attract a greater crowd of users.

2.2.10. Self-Disclosure on the Web: The Impact of Privacy Policy, Reward and Company Reputation

(Kaltcheva, Andrade, & Weitz, 2001)

This paper examines the approaches online companies take to encourage self-disclosure over the web and constitutes the exploratory effect of these approaches through an experimental design. It proposes that self-disclosure, like any other inter-personal behavior, depends upon and

is interpreted through the costs and benefits to the individuals, therefore also applies to electronic interaction. Some approaches companies take to alter the consumer's cost-benefit analysis are: developing a reputation for trustworthiness, providing a detailed ubiquitous privacy policy and offerings rewards in exchange. The authors hypothesize about each approach before they begin the testing. They predict that with a more complete and comprehensible privacy policy, the self-disclosure should increase. This is the only hypothesis of our concern. They test their hypothesis with an experimental design conducted on subjects.

The completeness of the policy was measured with three seven-point semantic-differential scales: *detailed vs non detailed, complete vs incomplete, informative vs non-informative.* The subjects indicated their concern of self-disclosure through seven types of information–name, social security number, email address, medical information, product preferences, and interests and hobbies. At the conclusion of the experiment subjects were asked how a websites privacy policy influences their readiness to disclose sensitive information. In conclusion the hypothesis for the privacy policy of a website proved to be true: A complete privacy policy alleviated concern over disclosure. So according to this paper most people do care about the privacy policies companies have to offer and keep them in mind before committing to self-disclosure.

This article takes up a practical approach to privacy policies through experimental analysis. Although our project does not involve conducting an experiment, there are a few takeaways from this article. Firstly, it highlights one key aspect of privacy policies, readability. Readability plays a significant role in determining the quality of a privacy policy. This article also addresses the disclosing of sensitive information to third parties and the corresponding response of the subjects to this issue. These are some of the factors that should be taken into account while analyzing privacy policies as users do care about their privacy and personal information.

2.3. Website Analysis

Following the article research, we went on to analyze privacy policies from websites across several industries to get a better perception of the similarities and differences amongst them.

2.3.1. **Bank of America** (Bank of America)

The privacy policy overview link is presented in the home page of the website which is useful and recommended. The overview page is broken down into several segments: Consumer Privacy Notice, Online Privacy Notice, Privacy Choices, and Affiliate Companies.

The Consumer Privacy Notice page talks about why the bank extracts customer information, for what purposes and how it is generally used. There is a breakdown of the reasons for collecting user information, e.g. social security and employment information, and whether the bank shares any of this information and whether a customer can limit the sharing if it does. For example, a user can limit sharing of such information – "For non-affiliates to market to you — for all credit card accounts". The page also provides general information about when this information is usually collected, how it is protected, what regulations (grouped by state laws as well) prevent a customer blocking usage of all information and some basic definitions.

The Online Privacy Notice page describes how information is collected from users online and how it is generally distributed or used. The last time the page was updated is mentioned on the top of the page. Firstly, it lists all the ways the bank reportedly uses personal information. Then the page lists some of the personal information that the bank collects, e.g. browser information. Then it lists all the ways it collects such and other information – through the browser, using cookies, Using pixel tags, web beacons, clear GIFs or other technologies, flash objects, IP address, and Aggregated and De-identified Data. It describes in detail how they work

and what kind of information is usually stored. Then the company addresses the kind of advertising it conducts or sponsors. The user, it seems, can opt-out of behavioral advertising from third-party advertisers whenever they please, although the browser has to be set to allow cookies to be accepted. Even then, the customers will still receive online advertising from Bank of America. The company passes personal information to affiliated companies and some non-affiliated ones, although the latter can be regulated. The company, however, does not hold themselves liable for any misuse of information if the user follows links from the website to other websites, as it is completely the user's responsibility be aware of their policies.

The Privacy Choices page explains how the customer can login and specify preferences regarding matters like marketing policies and behavioral advertising. The Affiliate Companies page lists the companies the bank is affiliated with.

2.3.2. **JPMorgan Chase & Co.** (JP Morgan Chase & Co.)

JPMorgan Chase & Co. is one of the largest American multinational banking institutions. The privacy policies of JPMorgan Chase and Co. are separated into two segments – one for JPMorgan and for Chase. The link to the privacy policy home page and the Terms and Conditions page is located on the bottom of the home page of the JPMorgan Chase and Co.'s website.

The privacy policy of JPMorgan describes in detail how the client information is protected and how such information is used for appropriate purposes. It also describes the reasons for sharing such information with affiliates and what sort of choices the company provides to clients regarding this matter. There is also a small section on certain laws which the company complies with regarding the privacy of client information. The company also provides come online security tips section and links to other pdf files concerning privacy policies in

different countries regulations. The page, however, does not have a date for the last update of the policy.

The format for the privacy policy for Chase bank is similar to that of Bank of America's. The last update of the policy is listed on the right hand corner which was in October 2011. The policy consists of two pages listing what kind of information they collect, how they usually use and share such information, the reasons behind using such information and how much freedom a client possesses in sharing their own information. Some laws are referred when mentioning why a client cannot completely block all sharing of one's information.

The company also has separate link to its Cookies Policy and Terms and Conditions. The cookies policy page first describes what cookies actually are and what types of cookies the company mainly uses to store information – session cookies and persistent cookies. Most first party and third party sites cookies can be blocked by the user's browser, but the company highly recommends not to block its first party site cookies as the data stored in them is used to enhance user's functionality in the websites. A complete list of third party sites is not accessible anywhere in the privacy policy home page.

2.3.3. **Citigroup** (CitiGroup)

Citigroup is also one of America's largest financial multinational financial services corporation. Citigroup's link to its online privacy policy is provided at the bottom of the website's home page. The privacy policy at this link does not describe the policies for collecting and sharing information from clients for all of its firms. The page ensures customers that everything is being done to keep their information private as it is considered a high priority by the company. It then talks about email encryption and how the company uses it to protect the

clients' information, especially regarding emails. The last update for the page is mentioned at the end, which was 26th October, 2011.

As the privacy policy was not comprehensive at the home page, we further looked into the websites of the separate firms, e.g. Citibank. The privacy policy listed at the bottom of that home page is completely different and goes in depth of the bank's policies, which we assume is the same case with all the other firms' websites. At first, the page describes the non-personal information that it collects, e.g. IP address. Then it describes the personal information it collects. It also provides links to other important privacy statements – Consumer Privacy Notice, Personal Wealth Management Privacy Notice, Citi Cards Privacy, Internet Privacy Policy and Statement. Then the page describes how Citigroup generally uses the information it directly collects and collects through advertisements and how this information is used. They, then, indirectly recommend not blocking cookies as the company may not be able to identify the client sometimes. The page also mentions how the client can opt-out of third-party targeted advertising specifically. However, there is no list of affiliates like it was provided in the Bank of America's website. The last effective update is mentioned at the bottom, which in this case was in 31st July, 2011.

The Consumer Privacy Notice statement is similar to that of the other banks', showing in columnar form of how and what kind of information is usually collected from consumers and how it is utilized. Then it mentions what choices the consumer has in limiting the share of one's information. The degree of freedom seems to be constant throughout all banks, most probably due to the adherence of the same laws. The last update for this is mentioned - July 2012. The Online Privacy Statement describes mainly what kind of information is passed or not passed to third-parties. The page also talks in detail about how information is protected through various

forms of encryptions and recommends a consumer certain protection settings. The last update for this page is not mentioned.

2.3.4. **Goldman Sachs** (Goldman Sachs)

Goldman Sachs is one the largest American investment banking firms that engages in several global services. The links to the Privacy and Security, Terms of Use and Regulatory Disclosures pages are all located on the bottom of the website's home page.

The Privacy and Security page is divided into several sections – Goldman Sachs Global Privacy Policy, How You Can Protect Yourself, Our Use of Cookies and Your Options, Other Sources of Information, Privacy Policy – Australia and New Zealand, Privacy Policy – India. The Global Privacy Policy page mentions some of the sources of information that is collected, e.g. through applications forms. Then it lists the typical information that the company collects and stores, and how the company uses such information to facilitate the efficient running of the firm. Then it mentions how information can be disclosed to all firms and entities within Goldman Sachs and how the only information passed onto third parties is described within this privacy policy. Then the page lists the methods the company is adopting to ensure all customer information is secure, as prescribed by the law, e.g. having an Information Security department. Then the page mentions some other important aspects about privacy and security in the company, e.g. information about cookies, how 'clickstream' data is collected, stored and shared within the company, and about third-party sites collecting information in general. The site also mentions that regulations will differ amongst countries and thus certain policies of using information will differ too. The policy mentions that this version is effective since 1st July, 2010.

The ‘How You Can Protect Yourself’ section provides security tips to customers. It talks about Phishing, Firewalls and Viruses. The page also lists some ways customers can protect themselves appropriately. The Cookies page lists the ways the company generally uses information collected from cookies. The page also talks about various ways this information may be used, e.g. to authenticate a user and to perform research and analytics. However, it strictly mentions that that the cookies don’t store any information which can be read and understood by others. A list of affiliates was not provided anywhere.

2.3.5. **The New York Times** (The New York Times)

The link to the privacy policy of New York Times is listed at the bottom of the webpage, including links of their Terms and Services and Terms of Sales. The font, however, for these links is significantly lower than the font used for the rest of the material in the homepage.

The privacy policy starts off with a list of affiliates and the date the page was last updated in. The privacy policy has also been certified by a third-party company called TRUSTe proving that complies with all acceptable policies and standards.

The first section of the policy is ‘What Information Do We Gather From You?’. This includes personal information – registration information, billing and credit card information, user generated content and public activities (for which the New York Times is not held accountable for), surveys and market research, events and promotions, and mobile messaging service. The section also includes details about non-personal information collected using technology – device information (e.g. browser information), cookies, analytics technology (New York Times sponsors third-parties to provide report and analyses on browsing patterns and other information, but the information is ultimately their property), IP addresses, reading history, and location

information (mainly via mobile apps). The third sub-heading is about 'Information Collected By Third Parties Using Technology'. Third parties may extract information using cookies for which the company will not be held responsible for. The company provides separate links to how Google uses cookies to extract information and how a user can opt-out of behavioral advertising.

The next section deals with what the information collected is used for. The company uses the personal and non-personal information to fulfill requests, to perform necessary statistical analysis, to customize user experience, to perform effective advertising, and to send email newsletters. The next section describes who the company shares this information with – affiliates, third parties (only relevant information) and service providers (for their specific needs only).

The next few sections describes in detail the email usage policies for the company, how a user can update personal information, and how a user can opt-out from receiving promotional offers from the company. In the last few sections, the policy describes how all information is completely secure with the company, and how the company complies with all laws when dealing with information related to users (including Children protection laws). The last section notes that significant changes to the policy will be posted on the webpage thirty days prior to taking effect and registered users will also be notified about it via email.

2.3.6. **CNN** (CNN)

CNN's website provides links to their Terms and Services, Privacy Policy and Ad Choices pages at the bottom of their home page. The font size for the links are significantly lower than the ones present in most other websites visited.

The privacy statement is a descriptive page with details about the different policies. The first section describes the information the website collects, which is presented and operated by a third-party named ‘Turner Broadcasting System, Inc.’. The website collects personal information through activities like registration or subscriptions. The site also collects non-personal information, such as type of operation system and domain name of Internet provider. The information that is collected can be used to provide exclusive content, personalize online experience and contact through social networking sites. However, information collected by third-parties from customer visits to their websites will not be regulated by CNN.

The next section describes how the information collected is being used. It is used to fulfill requests, connect to the customer, deliver relevant marketing strategies and improve design of services. The next section informs to whom this information is shared. The information is shared with law enforcements, business units and entities, agents and contractors of the company. Some personally identifiable information is also shared with third-parties for their marketing needs, however, this share of information can be opted-out of.

The page moves on to Cookies. It describes what cookies are and what kinds of technologies are used to track information through cookies. The section also informs the user that blocking all cookies may prevent access to several features in the website. Some third-parties sponsor several features in the website. The company has no control on how these third-parties use the collected information. The user will be notified if these third-parties want access to personal information. The company also passes on information to third-parties who generate reports and perform site analytics. The privacy policies for these companies are all distinct.

The next section describes how ads are generated for users using information collected. However, users can opt-out of interest-based advertising. Turner is a participant of the industry's self-regulatory program Digital Advertising Alliance (the "DAA") and adhered to all its principles. The page also mentions that policies will differ amongst countries as information is passed on. The end of the page lists contact information for queries and how a user can access personally identifiable information that is collected from them. The page also lists the date of last update, and mentions that any major changes in the privacy policy will be notified to the customer via email.

2.3.7. **The Wall Street Journal** (The Wall Street Journal)

The links to the privacy policy, data policy, copyright policy, and subscriber agreement and terms and use are all provided, with appropriate font sizes, at the bottom of the homepage of The Wall Street Journal.

The last update for the privacy policy is listed at the top of the page of the privacy policy. Dow Jones and Company Inc. operates the website of The Wall Street Journal. The website does not collect any personal information from children below 14 years of age. The privacy policy applies to all consumers and users of the webpage, and by providing personal information a user complies with all material presented in the privacy policy.

The next section describes the personal information the website collects, either by submission from user, e.g. registration, or from third-parties, e.g. purchase history. The website also collects other information, i.e. non-personal information, through user input, third-parties and other information the website collects when the user is online, e.g. IP address. The company uses cookies to track information, and third parties that provide services to the company may use

similar technologies too to extract information. However, the company cannot be held accountable for what they do with such information as their privacy policies are different.

The page then lists how the company uses this information. The information is mainly used to manage and provide primary services more effectively, to improve that service over time and personalize the online user experience. The information collected is also used to deliver relevant content and recommendations, to deliver relevant online advertising, and to contact the customer. The next section describes how the company shares this information. Certain information is made public through services that the consumer uses. The company also shares all information with affiliates (the link to the affiliates website is provided but I haven't been able to locate a complete list). The company may also share personal information to manage corporate, third-party, and student subscriptions, to allow third-party service providers to assist us in providing and managing services, to permit third-parties to send you promotional materials, to protect the rights of Dow Jones and the customers', and to complete a merger or sale of assets.

The next section lists some other notices necessary, including data retention of information as long as required by the law, cross-border transfer information, and privacy policy changes which will be notified to the customer. The company mentions it is not liable for the conduct of third-parties that may have links present on their website. The last two sections describe how security is vital to the company, and how a customer can update information on the website.

The Data Policy basically summarizes critical information from the privacy policy. It mainly notes that no unaffiliated party of the company can directly use or extract information about customers from the website without prior permission from the company.

2.3.8. **Google** (Google)

The privacy policy for Google is comprehensive and verbose. The page is easily accessible. (google.com/policies). Essentially it goes over the following points:

1. What information the website collects.

2. How the information is used.

3. The choices that are offered for personalization and how a user can access these choices.

The website keeps track of how the user uses the services provided by Google like the search history, device information, internet protocol addresses and phone information, if the service is accessed via a phone. Location information and application numbers for unique services are also kept in record. Because Google is such a huge service provider, offering a variety of ubiquitous applications to its enormous base of users, it has to track each and every one of them. It does a good job in telling them how it's done.

The information that is collected is intended to provide, improve and develop new services. The user is offered a variety of choices to personalize the disclosed and upon deletion of this information the policy specifically mentions that residual copies of the information may not be deleted immediately from the databases or backup systems. In regards to sharing information the policy states that if the Google Account is managed by a domain administrator (in Google Apps) then the domain administrator will have access to the Google Account information which includes email information and other data. The personal information that Google collects is offered to affiliates and 'trusted' businesses for external processing and is shared in compliance to the privacy policy. This information also is shared with companies, organizations or individuals outside of Google if Google has a good-faith belief that the information being used is necessary to meet any applicable law, enforce Terms of Service, detect

and prevent fraud or to protect against harm against the rights or safety of Google. Non personal information is shared publicly with Google's partners- like publishers, advertisers or connected sites.

For information security, the policy also mentions the different kinds of services it provides to protect the information like encryption with SSL, a two-step verification process to access a Google Account and a Safe Browsing feature while using Chrome. The ways these services are implemented are explained in further detail through various links offered on the website. There is also a link to the Terms of Service which goes over the rules which a user must agree to abide by in order to use Google's service.

The privacy page has a separate tab dedicated specifically to Google advertising and explaining what policies are used and how they're implemented. It describes in detail how ads are served to the user and can be personalized based on user interests. There also exits an "Opt Out" option in Google's Ad Preferences Manager, which once selected, ensures that Google does not collect interest category information and therefore stops interest-based ads to pop up via Google when accessing the web. But ads still may show up based on non-personal information or the content of a web page. However, the user is ensured the user that information is not leaked to third parties without user consent. The policy also states that Google works with regulatory authorities including local data protection authorities, to resolve any complaints regarding disclosure of any personal data that cannot be resolved with the user directly. The Google privacy policy is regularly reviewed and modified and the page shows when the last update was made. The policy is also available in over 60 languages. One piece of missing information is a contact link for customer support.

2.3.9. **Bing** (Bing)

The Bing privacy policy is well organized and simplistic therefore does not go into as much detail as Google's policy. The web page is not easily accessible; one as to go through Yahoo!'s privacy policy to find a link to Bing's privacy policy. The policy follows a similar pattern to Google's by firstly explaining what the website does with the collected information. It also goes over the usage, storage and sharing of the information that is searched. So once a search is made on Bing, Microsoft collects the user IP address, the cookies and browser configuration and then attempts to derive the approximate location.

Bing uses the search queries to serve ads of user interest. Bing stores cookie ID's from account information that directly identifies the user. According to the policy, Microsoft uses built-in technological and procedural safeguards to prevent the overlapping of this data. And also removes the entire IP address after 6 months and other cross section identifiers after 18 months which is not a feature of Google's privacy policy. About the sharing of search information the policy states that the website might share search queries with selected third parties for research purposes. But before this is done the IP addresses and identifiers are removed. Bing provides search services to Yahoo! and also to non-Microsoft services that use it. The privacy policy mentions Facebook Personalization which explains how Bing uses Facebook Connect to validate a user's Facebook identity to personalize the search experience. At the end of the page there is a "Contact Us" option for customer support unlike Google's privacy page. There is no instant way to change the language of the policy.

2.3.10. **Dogpile** (Dogpile)

Dogpile is a low traffic search engine operated by InfoSpace, Inc. The webpage for its privacy policy has a low level of readability. The font size is a source of discomfort to the eyes.

It follows a similar pattern to Google and Bing's policies by going over what the website does with the collected information. According to the policy the website has two types of information – Anonymous User Information and Personal User Information. When the website is accessed as an Anonymous User, the IP address and browser configurations are automatically stored in the server records. When using Dogpile as a Personal User, most of the personal information, in addition to the IP address and browser settings, is used. Dogpile might receive personal user information from a broad range of activities including surveys, contests, program registration and general correspondence.

Dogpile conducts voluntary surveys, contests or promotions from time to time, information collected from these events are specifically used only for informing the user about the events. The policy informs the user that the website engages third party advertisement networks that place cookies on the user's computer when using Dogpile services. This information is stored and used to identify and target related advertising to the user of the computer, when on Dogpile or other websites within the third party's network.

The policy goes over the opt-out option where the user can choose to opt-out to some third-party advertising (members of the Network Advertising Initiative) but not all. It also makes the user aware of the Inter-Company Transfer of information which means that Dogpile has the right to transfer information internally and indeed does so, when accessing the website the user automatically consents to this policy. The privacy policy ensures the user that the information stored is safeguarded but doesn't disclose the methods or ways it protects the information. Changes to the privacy policy are made, according to the website but the date of change is not mentioned. The contact information at the end is provided for customer support.

2.3.11. **Excite** (Excite)

Excite is another low-traffic less popular search engine operated by Mindspark Initiative Network, Inc part of the IAC family of businesses including websites like Ask.com, Girlsense.com, Match.com, and many others. The website has a high level of readability. The privacy policy is a brief one taking up the shape of a list rather than a long description with a 'read more' option on the side.

The website collects the same information as the other search engines, including IP addresses, browser types and language, through which it tracks users' movements on the web to derive patterns and trends. Information can be collected to identify a user when using an Excite Product or from any other IAC businesses. The website may assign cookies to the user's computer which can pick up information to personalize the online experience for users. The policy says that personal information is not provided to any websites, but Excite may allow third-party advertising companies or ad networks to display advertisements on the website. The user may opt-out of ad targeting if the company is a member of the Network Advertising Initiative. Account and search personalization is offered. The user is given a choice of not to disclose personal information, but this choice renders the user void of certain features provided by the website. Cookies can be deleted or declined through browser settings.

Similar to Dogpile's policy, Excite's policy also mentions information security in a general way. It includes that the policy is subject to changes and if any material change in a way that personal information is collected or used then the user is notified via email. There is a contact option at the end of the page for any questions or support. The privacy policy webpage had ads relating to the different IAC businesses at the end.

2.3.12. **Monster** (Monster)

The privacy policy for Monster.com briefly highlights the important aspects of the actual privacy statement that the user should be aware of. There is a separate link to the detailed privacy statement that covers Monster's privacy practices more extensively. It goes over what kind of information is collected, which includes basic personal and demographic information and how this information is used and retained.

The policy says that information is also acquired from third parties in an effort to personalize the user experience. The policy does not go into specifications about who "the others" are. It states that the website tries its best to limit access to this database to only legitimate users but does not guarantee that unauthorized parties are entirely limited to access. The information provided is used to deliver the services offered by the website. This information is shared with third parties that help deliver these services to the user. The policy claims that the information is only used to assist Monster in providing the services. Information might be disclosed if the business is sold.

Monster's privacy policy also goes over the personalization of information and explains the steps one should take to modify the account. It states that Monster has no control over the information that Ad Networks or third parties may have about the user. The opt-out option does exist for some of these ad networks. It is mentioned that turning off cookies might affect the use of the website. It also goes over the information security policy explaining how the information is protected and the methods Monster uses like SSL and JavaScript etc. At the end of the page the contact information of Monster's privacy office is provided for customer assistance. The website is secured by TRUSTe which is a company best known for its online privacy seals certifying more than 5,000 websites.

2.3.13. **Jobing** (Jobing)

The privacy policy for Jobing.com is relatively short. It does not provide a link to a detailed privacy statement like Monster does. It explains how user resumes are stored in databases. If the user wishes the information to be confidential then the resume can be specified to not be searchable by employers or recruiters but still can be used to apply for jobs. The level of confidentiality of the users resume can be changed.

The website collects demographic information from users who wish to provide it, this data is used to for EEO (Equal Employment Opportunity) reporting to the employers. Like most websites cookies are used to keep track of user activity to improve and personalize the user experience. Newsletters are automatically sent to the users email but one can chose not to receive them. Contests are run on the website where the user is prompted for contact information. This information is used to send out promotional material from some of Jobing's partners (not specified) and additional information about the company. The opt-out option is available for some of the ad networks and a feedback link is provided. It also goes over information security and assures the user that the information is protected but does not specify how the website does so. The policy is subject to change and any changes made to the policy are to be posted on the webpage, the date of change however is not mentioned.

2.3.14. **Simplyhired** (Simplyhired)

The privacy policy for Simplyhired is a comprehensive one. It covers in detail general aspects of most privacy policies like how information is collected and used. The policy states that the user can access the website without disclosing any type of information. However some parts of the site require the user to create an account to access the services. It also explains how information collected from surveys, questionnaires or contests is used to deliver specific

services, verify authority to protect a user's account and to send notices for services that may interest the user.

The policy mentions how Facebook can be used to receive information like friends, location, education, etc. to improve and customize the services provided. Simplyhired, like most websites, places cookies to track activity and personalize user experience. The policy states that advertising partners of the company can also download cookies to the user's computer to deliver target advertisements. The site does provide links to third party websites and the user acknowledges theses links as a convenience and agrees that Simplyhired is not responsible for the content of such websites. Information security is mentioned but not specified what measures are taken. The policy is updated and the date of change is provided. The website provides an email address to address any issues.

2.3.15. Facebook (Facebook)

Facebook's privacy policy comes in the form of a very clearly organized page divided into six categories deemed the most important parts of privacy policy: How information they receive is used, sharing and finding your Facebook (getting to help control your privacy settings), other websites and applications (what other websites do with your info), how their advertising works, cookies pixels and other systems, and general info. Each of these subjects is broken into other segments with very detailed descriptions of everything that is going on. This is a simple and very easy to understand privacy policy using full size pictures and other details. Being that the site is also popular the site is also relatively consistently updated.

An interesting look at Facebook is also the many sub policies that are formed from the number of Facebook games that are out there. As of 2011 there are over two thousand extension games or apps for Facebook, some with their own polices and some without. Facebook is also

linked to many third party sites that give the user the option of using their Facebook login information to sign into the third party site. Facebook seems to cover a lot of these sites using generally broad terms rather than specifying out names.

2.3.16. **MySpace** (Myspace)

MySpace's privacy policy was last updated in 2008 and is not as uniform as other social websites. Its appearance is more of a jumble of text with a few larger font/bold/orange lettered titles. The language is also not simple nor is there any help from pictures illustrating where and how to access features of the policy. It seems to just throw all information at you and hopes you can understand the technical jargon of PII and other acronyms.

It is strange to see MySpace in this state as given its previous status as one of the most accessed websites of all time. The privacy policy that was presented was neither tabbed nor formatted in an easy and understandable way and was therefore rather confusing.

2.4. Summary

The objective of our background research was to create a practical foundation and understanding of the current privacy policies in effect across the US and also find out the human factor involved in the policies. Some of the key trails that we followed were looking to understand the privacy policy management systems, finding the ideal policy, as well as finding the imperfections of policies. This would help solidify our overall goal of analyzing the privacy policies around the net.

Privacy Policies in the US are founded on the principle that the consumer should be protected. The purpose of the privacy policy is to notify the user of what the website can and

cannot do and most if not all websites use privacy policies. However despite this the issue websites run into is the human factor. The human factor is unpredictable and inconsistent. This makes perfecting the privacy policy difficult. We researched how privacy policies were regulated to fit to the user interface. Regulations produced the Federal Trade Commission, the California Online Privacy Protection Act, the Children's Online Privacy Protection Act, Trust Guard privacy guidelines, and the CAN-SPAM Act. These acts center on bringing the privacy policies to understand through near uniformity. For example the FTC's policy is organized such that there are five key points plus two extra for children. Those points are notice/awareness, choice/consent, access/participation, integrity/security, and enforcement/redress. (Federal Trade Comission , 2012) The two extra for children are parental notice/awareness choice/consent and access participation integrity/security. These guidelines provide stable ground for privacy uniformity and can be called key components of an ideal privacy policy.

There was a dedicated article we came across that actually pointed out the need for clarity and standardization in privacy policies. The article focused on analyzing sites through the Gramm-Leach-Bliley Act (GLBA) and found various inconsistencies in 2003 legislation. Also through use of the Flesch Reading Ease Score (FRES) they found that the reading level necessary for most sites policies was greater than the education of the average Internet user. (Anton, Earp, He, Stufflebeam, Bolchini, & Jensen, 2004) This was documented in 2004 and has such been rectified by modern standards, for the most part.

Privacy policies can also be used as decision-making tools. (Jensen & Potts, 2004) Issues with privacy policies were gathered as data sets from high-traffic websites and healthcare websites. What the information found was that part of the population could not understand the policies based on the average level of education. There were also some sites that did not notify

users of policy change nor post a modification date. Another problem was the lack of a communication forum to discuss issues and previous regulations.

Our research also touched human decision-making. Privacy and Rationality in Individual Decision Making (Egelman, Cranor, & Chowdhury, 2006) is a paper that looks into the psychology of the user. Despite all the findings in the information including how individuals confuse security with privacy as well as the lack of user knowledge about the existence of privacy policies, this article was not deemed relevant to our research.

Our research also had us familiarize ourselves with various different website's policies and layouts. We stuck to different genres for this part of our research that lead to the analysis of banking sites, news sites, search engines, and social networks. From this part of the background research we were able to gather the information that group the sites contained in each genre together. The information would prove to be a founding structural block for the latter part of our work.

3. Research Goals

The project was initiated to evaluate web privacy policies of various sectors of the online network. Data privacy concerns are a serious issue to hundreds of Internet users and hence the major objective of the project is to trace down where user information could be at a serious risk and what correlation it has with its privacy policy. After thorough research, we had a tentative idea what major privacy policies incorporate within them. Our goals evolved over time as we continued to refine the direction for our project. Once we had a clear image of the kind of research that had been previously done in our desired field of work, we shaped our goals so that we could use previous ideas to conduct our research.

One of our initial concerns was the positioning of the privacy policy and terms and conditions links on the home page of a website. In addition, we also wanted to find out whether these links would directly take the user to the desired page or to a page with links to specific selections. We were also interested to know how many companies actually mentioned the date of latest revision on the privacy policy, and whether any of these companies notify the customers when changes are made to the policies. We thought of looking further into the Terms and Conditions page, specifically whether companies mention the availability of the privacy policy and refer to any of the major privacy policies in their terms and conditions page.

We planned to investigate how successful a privacy policy is in apprising to the user how the information provided is utilized. The role of third parties and affiliates in sharing of data is another major research goal as we intend to determine how personal and non-personal information stored in web databases gets shared by other websites. Analyzing what roles third-party sites and affiliates play when it comes to user information is one of the primary objectives

of our project. We want to explicitly investigate what first-party sites specify about collecting and sharing of user information, the level of choice the customers possess in regulating these policies and the level of freedom third-party and affiliate sites possess in exploiting the user information. It is also important to note what kind of information is shared.

Understanding the role of targeted advertisement as well as cookie based trickle of information in respect to the whole situation is another area of attention which we would retort to by our research. It is widely recognized that first-party sites utilize cookies to track information. However, we wanted to investigate whether first-parties allow third-parties to track information through cookies placed from either party from the first-party website. In addition, we intended to examine the number of companies that explicitly mention the option of modifying browser setting to not accept cookies. We want to figure out how the privacy policy addresses behavioral and targeted advertisements and what alternatives it provides in cases where users may want to opt out of receiving such advertising.

We wanted to inspect whether companies care to provide customer service contact regarding privacy policy issues. We also believed it would be interesting to find how many online privacy policies are actually certified by a regulating body and what standards these privacy policies followed.

One of our other principal research goals was to analyze the readability of privacy policies. This method has been used in previous research and we wanted to incorporate it in ours in order to connect all aspects together. The readability level of privacy policy sheds an interesting light on how these policies are designed and whether companies actually reflect upon their customer base when forming these policies.

Once all our research goals tied into to our methodology, we had a better understanding on how to apply our goals to online privacy policies. The design, readability and transparency of these privacy policies were our prime investigation concerns when we went forward to form our methodology.

4. Methodology

4.1. Introduction

After performing necessary background research and diving into privacy policies from several industries, we formulated a list of attributes, and possible outcomes for each, to apply on various privacy policies. In devising our methodology, we incorporated all of our research goals in the methodology by investigating several aspects of privacy policies. We formed our methodology in a manner that would provide unbiased results. We integrated factors which that we deemed to be important, regardless of which companies and which industries incorporate them in their privacy policies. The methodology went through several redrafts, till we settled on a list of effective questions and standardized choices for each on which to base our analysis.

4.2. Methodology Questions

1. *Location of the link to the privacy policy home page on the front page of the company – Top, Bottom, Not Available*

   The purpose of this question is to notice whether there are differences across websites on where they locate the privacy policy link.

2. *How accessible the link to the privacy policy page is on the host website? – Number of clicks to go to the policy from the front page of the website.*

   This question is to determine how easy it is for a user to reach the actual privacy policy page.

3. *Whether the date of last update is explicitly mentioned? – Yes/No.*

   This question is to determine whether a company informs a user how up-to-date its policy is.

4. *The readability of the privacy policy – We used one of the popular standards to analyze this factor - The Flesch-Kincaid Test* (Flesch Readability Score).

The formula for the Flesch Reading Ease Score (FRES) test is:

$$206.835 - 1.015\left(\frac{Total\ Words}{Total\ Sentences}\right) - 84.6\left(\frac{Total\ Syllables}{Total\ Words}\right)$$

The total number of words, sentences, and syllables are collected from the main page, if it is directly available, or from all relevant links for which our methodology applies. For example, the main page of Facebook's privacy policy is a list of seven links that go over the different aspects of the policy. To calculate the readability of the entire policy, the readability (FRES score) is calculated for each link separately and then the average score is taken.

The FRES readability score is an industry standard. Many researchers have used this score to determine the readability level. It takes into account most factors which account for the level of difficulty in reading a text. There are a few ways to interpret the score, and we have used the ones which are most appropriate. The readability score determines the difficulty of the text; the higher the score the easier the text is to read. This conclusion can be inferred from Tables 1 and 2. However, it is crucial to note it does not consider several factors that could affect one's ability to comprehend the text, e.g. font, font size, text color, background color, etc.

(Table References):

Table 1. Score Interpretation

Flesch Reading Ease Score	Readability Level
0 - 29	Very difficult
30 - 49	Difficult
50 - 59	Fairly difficult
60 - 69	Standard
70 - 79	Fairly easy
80 - 89	Easy
90 - 100	Very easy

Table 2. Score Educational Interpretation

SCORE	NOTES
90.0–100.0	easily understood by an average 11-year-old student
60.0–70.0	easily understood by 13- to 15-year-old students
0.0–30.0	best understood by university graduates

5. *Location of the link to the Terms and Conditions or User Agreement page on the front page of the company – Top, Bottom, Not Available*

   This question is to determine whether the Terms and Conditions link is placed in a different location compared to the privacy policy.

6. *Whether the link to the Terms and Conditions or User Agreement page is accessible? – Number of clicks to get there from the front page of the website.*

   This question is to determine the level of ease to reach the Terms and Condition page. We can compare this with the number of clicks it requires to reach the privacy policy link.

7. *Does the Terms and Conditions or User Agreement mention the availability of the website's Privacy Policy? – Yes/No*

8. *Does the material of the Terms and Conditions or User Agreement discuss any of the policies regarding collecting or sharing of information? – Yes/No, Unclear*

The previous two questions will determine the extent to which the Terms and Conditions page refers to the privacy policy. The Terms and Conditions page is probably read by more customers than the privacy policy, so it was important to know how many of them refer to the policies regarding collecting or sharing of information.

9. [1]*Does the company pass on information to third-parties? – Yes/No, Unclear*

   Third-parties include non-trusted, trusted third parties, and non-affiliated third parties.

[1] The answer will be deemed 'Yes' even if only one particular case satisfies the situation. For example: the company may distribute only non-personal information to third-parties and not share personal information, yet the answer to Question 9 will be 'Yes'.

10. [1]*If Yes to Question 9, is the consent of the user needed to pass the information? – Yes/No, Unclear.*

11. *If Yes to Question 10, what kind of information? – Non-personal, Personal, Personal and Non-Personal, Unclear.*

    Non-personal – Browser information, Information collected through cookies, Aggregated and De-identified data.

    Personal – Name, Address, Billing information, telephone number, credit card, SSN.

12. *If Yes to Question 9, what are the third-parties allowed to do with the information? – In accordance to host company's privacy policy, In accordance to third party's privacy policy, Not Addressed.*

    In this case, Not Addressed addresses all other options, including Unclear and N/A.

The previous four questions are equally important and help us determine what each company specifically mentions about passing information to third parties. This analysis can give a picture on what these companies actually declare in their privacy policies and the level of transparency they disclose.

13. [1]*Does the company pass information to affiliates? – Yes/No, Unclear.* Affiliates are companies which are part of a conglomerate and affiliated third-parties.

14. [1]*If Yes to Question 13, is the consent of the user needed to pass the information? – Yes/No, Unclear.*

15. *If Yes to Question 14, what kind of information? – Non-personal, Personal, Personal and Non-Personal, Unclear.*

Non-personal – Browser information, Information collected through cookies, Aggregated and De-identified data.

Personal – Name, Address, Billing information, telephone number, credit card, SSN.

16. *If Yes to Question 13, what are the affiliates allowed to do with the information? – In accordance to host company's privacy policy, In accordance to affiliate's privacy policy, Not Addressed.*

In this case, Not Addressed addresses all other options, including Unclear and N/A.

The previous four questions are of equal importance as well, and help us determine what the companies declare about sharing information to affiliates. Most companies share, or are at least expected to share, customer information to affiliates. This question will reveal how many companies explicitly mention that they do.

17. *What is the number of affiliates/partners? – Number.*

This question will help us find the number of affiliates a company lists to the users. That will also allow us to calculate the average number of affiliates per industry.

18. *Are third parties allowed to store cookies from the host company's website? – Yes/No, Unclear.*

Many companies store information through cookies. This question is to determine how many of them allow third parties to store cookies in a user's computer from the host company's website.

19. *Does the privacy policy address the enabling or disabling of browser settings to accept cookies? – Yes/No, Unclear.*

This question determines if companies mention to users the choice of not accepting cookies by modifying browser settings, as many users may not be aware of this option.

20. *Can a user opt-out of behavioral or targeted advertising or marketing from the first party? – Yes/No, Unclear.*

This question indicates whether a company explicitly mentions the option of opting out of first-party behavioral or targeted advertising.

21. *Does the privacy policy address the option of opting out of third-party behavioral or targeted advertising? – Yes/No, Unclear.*

This question indicates whether a company explicitly redirects the user to another website to opt-out of third-party behavioral or targeted advertising.

22. *Whether a registered customer will be notified for significant changes made in the privacy policy? – Yes/No, Unclear.*

Many companies update the privacy policy without directly notifying the customer about it, e.g. via email. This question is to track how many companies will directly contact the customer in such a situation.

23. *Whether there is customer service and contact available for queries and questions? – Yes/No, Unclear.*

This question is to find out how many companies are willing to clarify confusion amongst customers regarding the privacy policy.

24. *Whether the policy is certified or verified by a regulating body or company? – Yes/No.*

This question is to find out how many companies' privacy policies are certified by a third-party.

25. *If Yes to Question 24, what is the name of the company? – Name.*

This question lists the names of the regulating companies which verified the privacy policy.

4.3. Summary

These attributes provide us with meaningful results and provide us with a strong base for our analysis and interpretation of results. We tried to incorporate all possible answers or results in our list of choices. We kept in mind the fact privacy policies differ significantly across websites due to industry standards, laws and regulations. We believe that such a methodology would lead us to find comprehensive results relating to privacy policies. Our next step was to decide a set of categories and websites to pick on which we could apply our methodology on. We discuss our reasoning behind that in the following chapter.

5. Selection of Websites

5.1. Introduction

The aim of this project was to accomplish an overall picture on the operation of web-based privacy policies as well as to highlight certain anomalies that were particular to individual websites or specific industries. Henceforth, we decided to evaluate major industries in the web environment. After conducting a bit of research and using web information tools like Alexa (Alexa - The Web Information Company) and Quantcast (Quantcast), we came up with ten different widely used industries over the Internet.

The following ten industries were chosen:

1) Financial
2) News
3) Adult
4) Social
5) Blogging
6) Search Engines
7) Job
8) Games
9) Healthcare
10) Dating

Using the same tools we choose ten websites within the specific industry that confirmed a total of hundred websites that were to be evaluated. Our goal was both to conduct an exhaustive survey of the major sites, and include several representatives of other common sites from different industries so that we could conclude an overall analysis as well as an individual analysis within the respective industry.

First, we collected general information about the site, such as its launch date, estimated user count and traffic ranks, country of operation, and ownership. Next, we examined the publicly viewable sections of the webpage which are presented to non-members who visit the website. These offered the most valuable insight and on the basis of the statistics we choose the websites.

Global Rank was an estimate we used to determine the website's popularity. The rank was calculated using a combination of average daily visitors to the website and page views over the past 3 months. Rank in country is an estimate of the websites popularity in a specific country. The rank by country was calculated using a combination of average daily visitors to website and page views from users from that country over the past month. Reputation was the number of links to the website from sites visited by users in the traffic panel. Analyzing this criterion we came up with the list of websites that were to be evaluated. This ensured that the quality of the study was augmented because the research was conducted on those sites that were not only widely used but also had retained a specific standing in the online environment.

5.2. Categories

The following categories were chosen:

5.2.1. Financial

WEBSITE	GLOBAL RANK	RANK IN COUNTRY	REPUTATION	COUNTRY
1. Bank of America	129	26	14,306	USA
2. Chase Bank	143	29	4,865	USA
3. Wells Fargo	169	38	9,335	USA
4. PayPal	39	23	87,104	USA
5. American Express	352	97	18,034	USA
6. HDFC Bank Ltd.	448	26	3,704	INDIA
7. Capital One Financial Corp.	460	87	4,730	USA
8. ICICI Bank	525	33	3,192	INDIA
9. Citibank U.S.	707	159	6,293	USA
10. Goldman Sachs	19,578	8,003	6,276	USA

5.2.2. News

WEBSITE	GLOBAL RANK	RANK IN COUNTRY	REPUTATION	COUNTRY
1. CNN Interactive	68	18	320,067	USA
2. The Huffington Post	80	20	211,437	USA
3. BBC News	48	5	435,835	UK
4. New York Times	109	25	463,308	USA
5. Wall Street Journal	228	80	2220,878	USA
6. Forbes Magazine	284	121	156,408	USA
7. Fox News	152	30	103,487	USA
8. The Guardian	192	16	256,956	UK
9. The Weather Channel	118	24	94,045	USA
10. NBC News	215	58	15,743	USA

5.2.3. Adult

WEBSITE	GLOBAL RANK	RANK IN COUNTRY	REPUTATION	COUNTRY
1. LiveJasmin.com	66	52	18,563	USA
2. Youporn	104	85	8,699	USA
3. XNXX Galleries	112	86	9,173	USA
4. Adult Friendfinder	225	218	23,488	USA
5. Streamate.com	915	537	2,336	USA
6. FreeOnes	974	505	8,011	USA
7. Literotica.Com	1,537	471	1,806	USA
8. Adam4Adam	1,627	323	356	USA
9. Manhunt	1,829	876	905	USA
10. GayRomeo	1,855	311	1,406	DENMARK

5.2.4. Social

WEBSITE	GLOBAL RANK	RANK IN COUNTRY	REPUTATION	COUNTRY
1. Facebook	2	2	8,126,001	USA
2. Myspace	198	135	976,452	USA
3. LinkedIn	12	11	1,475,711	USA
4. Hi5	618	610	35,637	INDIA
5. Tagged	264	268	19,949	USA

6. Twitter	8	8	6,452,783	USA
7. Pinterest	38	15	644,721	USA
8. LiveJournal	121	10	378,062	RUSSIA
9. Orkut	679	40	73,931	BRAZIL
10. Badoo	135	37	7,978	ITALY

5.2.5. Blogging

WEBSITE	GLOBAL RANK	RANK IN COUNTRY	REPUTATION	COUNTRY
1. photobucket.com	173	108	333,845	USA
2. flickr.com	58	45	1,280,133	USA
3. tumblr.com	36	21	543,068	USA
4. instagram.com	92	49	43,644	USA
5. wordpress.com	95	126	3,458,021	USA
6. reddit.com	131	64	685,628	USA
7. imgur.com	98	44	84,236	USA
8. TypePad	345	212	86,258	USA
9. Blog.com	710	159	87,027	INDIA
10. HubPages	428	366	110,020	USA

5.2.6. Search Engines

WEBSITE	GLOBAL RANK	RANK IN COUNTRY	REPUTATION	COUNTRY
1. Google	1	1	5,099,963	USA
2. Bing	21	14	177,372	USA
3. Yahoo!	4	4	2,025,170	USA
4. Ask	46	32	58,954	USA
5. Dogpile	3,434	965	14,497	USA
6. Excite	4,408	1,134	20,337	USA
7. Goodsearch	6,829	1,321	12,139	USA
8. HotBot	46,089	20,436	13,111	INDIA
9. Webopedia	5,940	1,355	11,370	INDIA
10. DuckDuckGo	1,974	1,057	11,176	USA

5.2.7. Job

WEBSITE	GLOBAL RANK	RANK IN COUNTRY	REPUTATION	COUNTRY
1. Monster	501	116	26,83	USA
2. Jobing	24,124	4,788	2,929	USA
3. Simplyhired	1,522	439	13,172	USA
4. Careerbuilder	636	140	23,241	USA
5. Indeed	254	81	26,134	USA
6. EmploymentGuide	58,799	11,128	2,088	USA
7. Vault	38,439	13,755	3,617	USA
8. Dice	4,622	1,150	5,929	USA
9. Snagajob	3,487	734	2,981	USA
10. The Ladders	6,842	1,237	2,366	USA

5.2.8. Games

WEBSITE	GLOBAL RANK	RANK IN COUNTRY	REPUTATION	COUNTRY
1. Electronic Arts	1,134	1,156	32,454	USA
2. IGN	416	196	56,126	USA
3. GameSpot	803	571	33,362	USA
4. Steam	879	896	27,965	USA
5. MiniClip	635	817	15,370	USA
6. Battle.net	361	386	23,791	USA
7. Xbox	841	364	29,478	USA
8. PlayStation	1,378	1,201	33,805	USA
9. Pogo	617	131	4,327	USA
10. League of Legends	1,098	667	6,974	USA

5.2.9. Healthcare

WEBSITE	GLOBAL RANK	RANK IN COUNTRY	REPUTATION	COUNTRY
1. Univera Healthcare	694,019	170,180	80	USA
2. NIH	351	174	204,222	USA
3. WebMD	500	160	71,416	USA
4. MedicineNeT	1,445	597	24,952	USA
5. MayoCliniC	1,492	427	42,541	USA

6. Drugs	2,396	727	14,486	USA
7. everydayHealth	2,073	550	17,580	USA
8. MedHelp	4,596	1,679	4,364	USA
9. HealthGrades	3,210	673	7,021	USA
10. RealAge	8,377	2,042	5,786	USA

5.2.10. Dating

WEBSITE	GLOBAL RANK	RANK IN COUNTRY	REPUTATION	COUNTRY
1. Match	356	99	9,565	USA
2. PlentyofFish	15,019	5,483	3,349	USA
3. Zoosk	907	472	1,011	USA
4. eHarmony	3,178	691	4,228	USA
5. Singlesnet	8,220	2,029	560	USA
6. Ashley Madison	2,225	978	3,248	USA
7. True	66,806	18,122	571	USA
8. ChristianMingle	3,665	823	682	USA
9. Cupid	10,282	5,794	2,365	USA
10. DateHookup	3,510	748	2,087	USA

5.3. Summary

We wanted to evaluate a wide variety of industries so that we can later come up with industry specific results and comment on anomalies that were only witnessed in certain industries. Having a wide variety of web-based industries also ensured that our results would be more realistic and reliable since they would portray the overall picture and not just results pertaining to one online sector.

6. Results

6.1. Introduction

We applied our methodology to the privacy policies of the selected websites. After recording the data, we compiled the results into groups, which allowed us to directly compare them across all industries. This form of analysis gave us the scope of finding the differences and similarities that separate industries have regarding the various aspects of a privacy policy we investigated in our methodology.

6.2. Comparison

The following tables represent our tabulated results for each question in our methodology. The first column represents the options available for each question. The row counts represent the number of companies, per industry, for which the option is applicable. All the counts add up to one hundred. We analyze our findings at the attribute level in this chapter, before diving into deeper analysis on a per-category basis in the next chapter.

Legend:

FI – Financial, NW – News, AD – Adult, SO – Social, BL – Blogging, SE – Search Engines, JB – Job, GM – Games, HE – Healthcare, DT – Dating, Ttl – Total.

1. *Location of the link to the privacy policy home page on the front page of the company – Top, Bottom, Not Available*

		FI	NW	AD	SO	BL	SE	JB	GM	HE	DT	Ttl
1	Bottom	10	10	8	9	9	10	10	10	9	8	**93**
	Top	0	0	0	1	0	0	0	0	0	1	**2**
	Not Available	0	0	2	0	1	0	0	0	1	1	**5**
	Total	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**100**

93 websites, across all industries, post the link of the privacy policy link at the bottom of the home page. However, 2 Adult, 1 Blogging, 1 Healthcare and 1 Dating website did not post the link at the home page at all. This practice clearly seems to be an online trend, as all industries are following it. However, 5 websites in total did not post the privacy policy link on the bottom of the home page.

2. *How accessible the link to the privacy policy page is on the host website? – Number of clicks to go to the policy from the front page of the website.*

		FI	**NW**	**AD**	**SO**	**BL**	**SE**	**JB**	**GM**	**HE**	**DT**	**Ttl**
2	1	7	9	8	8	9	8	8	10	9	9	**85**
	2	3	1	1	2	1	2	1	0	1	1	**13**
	>2	0	0	0	0	0	0	1	0	0	0	**1**
	N/A	0	0	1	0	0	0	0	0	0	0	**1**
	Total	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**100**

85 privacy policies are one click away from the home page. However, 14 websites across all industries, apart from Gaming websites, lead the users to a page containing several links regarding the privacy. The user must then navigate to the appropriate privacy page from the links provided. One website in the Adult category did not possess a privacy policy page whatsoever. It appears that websites in general prefer to keep the privacy policy page only one click away from the home page.

3. *Whether the date of last update is explicitly mentioned? – Yes/No*

		FI	NW	AD	SO	BL	SE	JB	GM	HE	DT	Ttl
3	Yes	9	10	3	7	9	7	5	7	6	7	**70**
	No	1	0	6	3	1	3	5	3	4	3	**29**
	N/A	0	0	1	0	0	0	0	0	0	0	**1**
	Total	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**100**

In 70 out of 100 cases across all industries, the date of last update is mentioned in the privacy policy. All the News websites mentioned the latest update date. The stand out factor is that a majority of websites, six to be exact, in the Adult industry did not post the date of their last update. When Adult and News websites are compared, it can be clearly seen that News websites pay far more attention to detail and are more up-to-date than Adult websites. In fact, generally websites do prefer to post the latest update date.

4. *The readability of the privacy policy – We can use one of the popular standards to analyze this factor - Readability of the Websites using the Flesch-Kincaid Test:*

		FI	NW	AD	SO	BL	SE	JB	GM	HE	DT	Ttl
4	Average	36	34	36	43	43	47	43	39	23	22	**37**
	Median	36	33	34	41	43	46	45	41	24	23	**37**

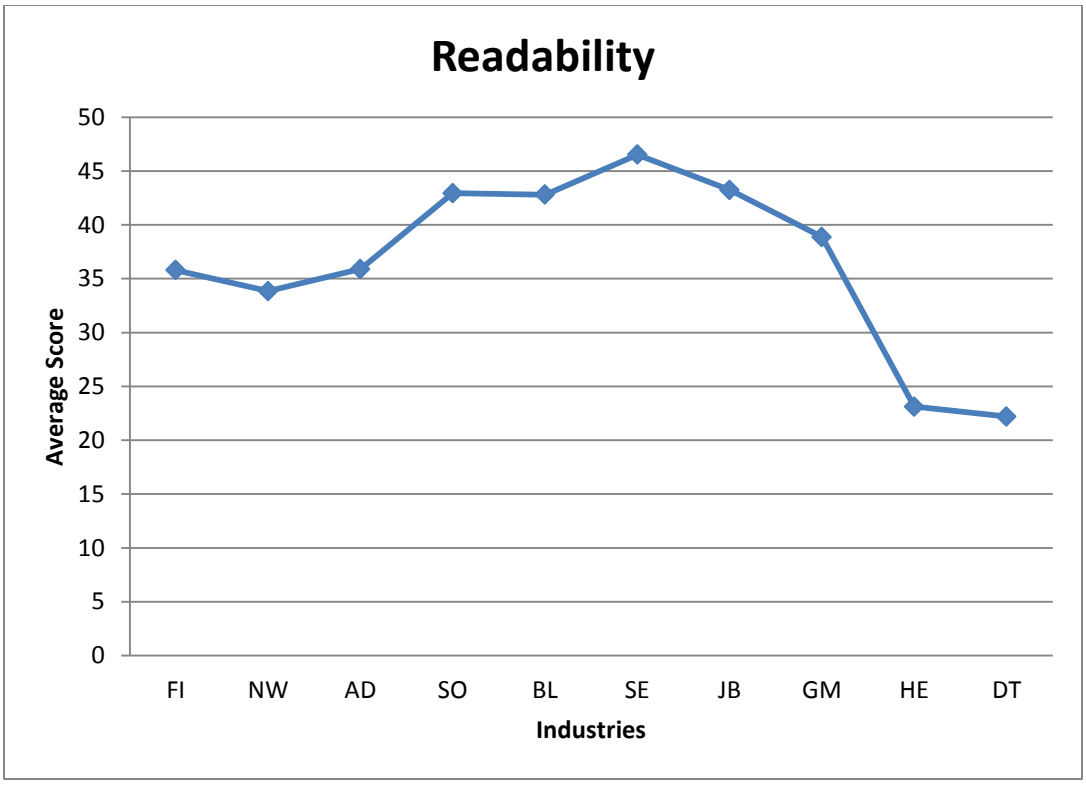

Figure 1. Readability Score Pattern Across Industries

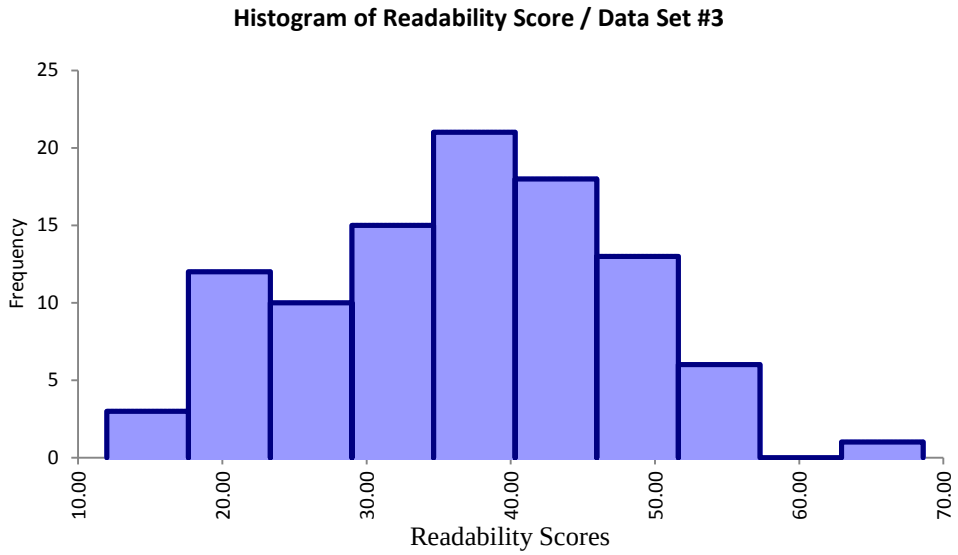

Figure 2. Histogram of Readability Scores

The averages vary quite a bit for the readability score across the different sectors. The different averages for different sectors are graphically represented in Figure 1. The overall standard deviation of our sample is 11, which is far less than the overall mean of 37. This shows that in general the readability scores are relatively close. When looking at the histogram of the readability scores across all websites, in Figure 2, we can notice that the scores are almost normally distributed. That result explains why the data has low variance and most data points are clustered towards the overall mean.

In comparison, the readability of Social, Blogging, Search Engines, and Job websites require lesser education than the other websites, which can be derived from Tables 1 and 2. Most of these websites require large amounts of user information for the website itself to run on a day-to-day basis. This may be one of the reasons why these websites have formed policies that are relatively easier to read and understand. The averages for Healthcare and Dating websites are extremely low, indicating privacy policies which are difficult to read and require high levels of education to fully comprehend them. In fact, the averages are more than one standard deviation away from the mean, indicating their great variance. It is not surprising that Healthcare websites are relatively difficult to read as they represent companies which deal with scientific matters and complicated user information. However, it is surprising to find privacy policies of dating websites difficult to read as the websites deal with social interactions and attract all crowds and age groups. Not all users who visit dating websites will possess a college degree, which is required to fully understand the privacy policies of these websites. In addition, not all healthcare website audience may be college-level educated either. In fact, these websites should design their websites and privacy policies for everyone, not just for the highly educated.

5. *Location of the link to the Terms and Conditions or User Agreement page on the front page of the company – Top, Bottom, Not Available*

		FI	NW	AD	SO	BL	SE	JB	GM	HE	DT	Ttl
5	Bottom	9	10	7	9	10	7	10	10	8	10	**90**
	Top	0	0	0	1	0	2	0	0	0	0	**3**
	Not Available	1	0	3	0	0	1	0	0	2	0	**7**
	Total	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**100**

90 websites post the link to the Terms and Conditions page at the bottom of the home page, which is consistent with the privacy policy link placement. Once again, this seems to be an overall online trend. However, a good proportion of websites, 7 of them, did not post the link of the terms and conditions page on the home page. This includes 1 website from Financial, 3 from Adult, 1 from Search Engines, 2 from the Healthcare industry. Most of the websites are from the Adult industry. It is important to note that 2 Adult sites had not posted the privacy link at the bottom of the home page either. These Adult sites may not be too keen to follow the standard.

6. *Whether the link to the Terms and Conditions or User Agreement page is accessible? – Number of clicks to get there from the front page of the website.*

		FI	NW	AD	SO	BL	SE	JB	GM	HE	DT	Ttl
6	1	8	10	7	9	10	7	9	10	8	7	**85**
	2	1	0	1	1	0	2	1	0	1	2	**9**
	>2	0	0	0	0	0	0	0	0	0	1	**1**
	N/A	1	0	2	0	0	1	0	0	1	0	**5**
	Total	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**100**

85 of the websites require one click to access the Terms and Conditions page from the home page. There are 9 websites that require two clicks and only 1 website in the dating industry

that requires three clicks. All News, Blogging and Gaming websites post the link at the bottom of the home page. The Terms and Conditions page was not available anywhere in the website for 1 Financial, 2 Adult, 1 Search Engine and 1 Healthcare Website. Once again, the adult websites don't seem to consistently follow the online standards.

7. *Does the Terms and Conditions or User Agreement mention the availability of the website's Privacy Policy? – Yes/No*

		FI	**NW**	**AD**	**SO**	**BL**	**SE**	**JB**	**GM**	**HE**	**DT**	**Ttl**
7	Yes	6	7	7	9	10	9	10	9	9	7	**83**
	No	4	3	0	1	0	0	0	1	1	3	**13**
	N/A	0	0	3	0	0	1	0	0	0	0	**4**
	Total	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**100**

From the table above it can be seen that most of the websites, 83 in total, mention the availability of their privacy policy in their Terms and Conditions page. There are a few websites, 13 in total, that do not do this. These websites are mostly from the Financial, News and Dating sectors. It is surprising to find that financial websites are not mentioning the availability of their privacy policies in their terms and conditions page, as these companies adhere to strict policies. The companies should direct the customers to their privacy policies so that they are aware how their financial information is protected by the company.

8. *Does the material of the Terms and Conditions or User Agreement discuss any of the policies regarding collecting or sharing of information? – Yes/No, Unclear*

		FI	NW	AD	SO	BL	SE	JB	GM	HE	DT	Ttl
8	Yes	4	6	5	9	10	3	8	5	9	3	**62**
	No	6	4	2	1	0	6	2	3	0	1	**25**
	Unclear or N/A	0	0	3	0	0	1	0	2	1	6	**17**
	Total	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**100**

The terms and conditions for 62% of the websites discuss policies regarding the collecting and sharing of information while 25% of them do not. The websites that do not discuss this issue are mostly financial and search engines. This question was important in the sense that a yes means that the websites privacy policy is somewhat consistent with its terms and conditions page.

9. Does the company pass on information to third-parties? – Yes/No, Unclear

		FI	NW	AD	SO	BL	SE	JB	GM	HE	DT	Ttl
9	Yes	9	10	7	10	8	9	8	9	9	9	88
	No	1	0	1	0	2	1	1	1	1	1	9
	Unclear	0	0	2	0	0	0	0	0	0	0	2
	N/A	0	0	0	0	0	0	1	0	0	0	1
	Total	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**100**

This question was fairly straight-forward and 88% of the websites do pass on information to third parties. This is an important factor as users have the right to know how their information is being treated, therefore adds to the transparency of the privacy policies. 9% of the total websites did not pass on information to third parties.

10. If Yes to Question 9, is the consent of the user needed to pass the information? – Yes/No, Unclear.

		FI	NW	AD	SO	BL	SE	JB	GM	HE	DT	Ttl
10	Yes	6	5	7	10	4	4	2	6	4	2	**50**
	No	2	1	0	0	3	2	1	1	5	7	**22**
	Unclear	1	4	0	0	1	3	5	2	0	0	**16**
	N/A	1	0	3	0	2	1	2	1	1	1	**12**
	Total	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**100**

The results to this question are somewhat spread out between the Yes, No and Unclear options. 50% of the websites require user consent. A good portion of the websites (22%) did not require user consent and it was unclear in many (16%) as well. User consent can be a big deal for some people and mentioning if the website requires consent or not to pass on information can change the mind of a user to continue using the services of the website.

11. If Yes to Question 9, what kind of information? – Non-personal, Personal, Personal and Non-Personal, Unclear.

		FI	NW	AD	SO	BL	SE	JB	GM	HE	DT	Ttl
11	Personal & Non-Personal	5	5	4	7	6	0	0	2	5	5	**39**
	Personal	1	3	3	2	0	5	8	6	2	2	**32**
	Non-Personal	2	1	0	0	2	4	0	1	2	0	**12**
	Unclear	1	1	0	1	0	0	0	0	0	2	**5**
	N/A	1	0	3	0	2	1	2	1	1	1	**12**
	Total	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**100**

This is also a significant aspect of privacy policies for which the results are slightly spread out. 39% of the websites do pass on both personal and non-personal information to third parties. A fair number of websites (32%) pass on personal information while 12% of the websites only share non-personal information. The results are evenly split for each category as well but one notable factor is that 8% of the job websites pass on personal information to third parties.

12. If Yes to Question 9, what are the third parties allowed to do with the information? – In accordance to host company's privacy policy, In accordance to third party's privacy policy, Not Addressed.

		FI	NW	AD	SO	BL	SE	JB	GM	HE	DT	Ttl
12	In accordance to host company's privacy policy	6	5	2	6	5	5	4	3	0	0	**36**
	In accordance to third party's privacy policy	0	4	0	3	3	3	4	3	3	3	**26**
	Not Addressed	4	1	8	1	2	2	2	4	7	7	**38**
	Total	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**100**

This question deals with the processing of the collected information. Most users would want to know what happens to their information after it is passed on to third parties. 36% of the websites deal with it according to their privacy policy, there is a fair percentage of the websites that leave it up to the third party (26%). Also in some of the cases this issue is not addressed, mostly coming from health and dating websites.

13. Does the company pass information to affiliates? – Yes/No, Unclear.

		FI	NW	AD	SO	BL	SE	JB	GM	HE	DT	Ttl
13	Yes	9	9	4	2	3	10	10	7	9	9	**72**
	No	0	0	2	0	1	0	0	1	0	1	**5**
	Unclear	1	1	4	8	5	0	0	2	1	0	**22**
	N/A	0	0	0	0	1	0	0	0	0	0	**1**
	Total	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**100**

72% of the websites pass on information to their affiliates. This was expected given the contracts and agreements most websites have with their partners and affiliate companies. 5% of the websites did not pass on information to affiliates. Some of the privacy policies are ambiguous and subtle regarding this matter so it was difficult to tell whether information is passed or not.

14. If Yes to Question 13, is the consent of the user needed to pass the information? – Yes/No, Unclear.

		FI	NW	AD	SO	BL	SE	JB	GM	HE	DT	Ttl
14	Yes	8	0	2	1	1	0	2	4	3	0	**21**
	No	0	2	1	0	1	5	2	2	5	7	**25**
	Unclear	1	7	1	1	1	5	6	1	1	2	**26**
	N/A	1	1	6	8	7	0	0	3	1	1	**28**
	Total	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**100**

The results for this question are widely spread out in different industries. While there is a significant percentage of websites that require user consent, there also is a fair share of websites where this issue is either Unclear or N/A. There seem to be no valid justification for this irregularity. The financial websites gave the most liberty to the user as eight of them had this option of user consent available whereas none of the websites had "No" for the answer. None of the news websites gave this option to the user neither did any dating website or search engines. Five health sites and five search engines did not have this option available which was the highest number in the category but on the other hand, for six job websites, eight blogging websites and nine social websites it was either unclear or the question did not apply which made the overall results look very haphazard.

15. If Yes to Question 13, what kind of information? – Non-personal, Personal, Personal and Non-Personal, Unclear.

		FI	NW	AD	SO	BL	SE	JB	GM	HE	DT	Ttl
15	Personal & Non-Personal	8	6	3	0	3	0	0	2	2	1	**25**
	Personal	1	1	1	1	0	7	10	4	2	3	**30**
	Non-Personal	0	0	0	0	0	1	0	1	1	2	**5**
	Unclear	0	2	0	1	0	2	0	0	4	3	**12**
	N/A	1	1	6	8	7	0	0	3	1	1	**28**
	Total	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**100**

Most of the websites that pass on information to their affiliates share both personal and non-personal information. There are a few websites that share only personal information mostly by job websites and search engines. The majority of the websites did not have a definite answer in the privacy policy for the question or the question was inapplicable to some privacy policies. Only Non-Personal information was shared by just one search engine, one gaming site and one health site. Surprisingly neither job site nor search engine or social websites revealed both kinds of information to the user.

16. *If Yes to Question 13, what are the affiliates allowed to do with the information? – In accordance to host company's privacy policy, In accordance to affiliate's privacy policy, Not Addressed.*

		FI	NW	AD	SO	BL	SE	JB	GM	HE	DT	Ttl
16	In accordance to host company's privacy policy	6	3	1	1	2	10	8	2	0	0	**33**
	In accordance to affiliate's privacy policy	0	2	0	0	1	0	2	3	3	0	**11**
	Not Addressed	4	5	9	9	7	0	0	5	7	10	**56**
	Total	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**100**

This question also has displays direct results showing that most of the websites deal with the shared information according to their privacy policy. This issue is not addressed in some of

the websites even through it is a fairly important factor. None of the dating website addresses this question and most of the adult, social, blogging and health sites. Search engines do a good job addressing this issue as all of them explicitly declare that their privacy policy would govern the channeled information to affiliates. Also most of the financial and job websites address this question as they generally keep high privacy standards.

17. What is the number of affiliates/partners? – Number.

		FI	NW	AD	SO	BL	SE	JB	GM	HE	DT	Ttl
17	**Average**	9	1	0	0	0	1	161	0	2	0	
	0	3	8	10	10	9	9	8	9	9	10	**85**
	>0	7	2	0	0	1	1	2	1	1	0	**15**
	Total	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**100**

The results for this question are significantly skewed. As most of the websites do not mention their affiliates the ones that do have a very large number of affiliates mentioned. Some Job websites mentioned a lot of partners and affiliates so that the public can trace where its private information could be channeled. Similarly most financial websites have reported affiliates; this testifies to the high standards the financial websites maintain but the majority of the websites analyzed did not mention the affiliates. None of the financial news and adult sites mentioned any affiliates.

18. Are third parties allowed to store cookies from the host company's website? – Yes/No, Unclear.

		FI	NW	AD	SO	BL	SE	JB	GM	HE	DT	Ttl
18	Yes	6	8	2	10	7	7	6	9	7	9	**71**
	No	1	0	1	0	0	2	0	1	1	0	**6**
	Unclear	3	2	7	0	3	1	4	0	1	1	**22**
	N/A	0	0	0	0	0	0	0	0	1	0	**1**
	Total	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**100**

Most of the websites are allowed to store cookies from the host company's website. This is a key issue as important information is tracked and collected through these cookies. This is the basis of behavioral targeting and many privacy concerns are associated with it. For a lot of adult websites inadequate information is provided regarding this question hence for more than half of them it was unclear to determine the answer. All social websites analyzed allowed third parties to store cookies and so did majority of the other websites. The privacy policy of only two search engines explicitly annulled this question.

19. Does the privacy policy address the enabling or disabling of browser settings to accept cookies? – Yes/No, Unclear.

		FI	NW	AD	SO	BL	SE	JB	GM	HE	DT	Ttl
19	Yes	9	10	7	8	4	3	8	5	0	2	**56**
	No	1	0	1	1	2	3	2	4	6	6	**26**
	Unclear	0	0	1	1	4	4	0	1	3	2	**16**
	N/A	0	0	1	0	0	0	0	0	1	0	**2**
	Total	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**100**

In most cases the websites do address the enabling and disabling of browser settings to accept cookies. This also is a significant factor that a user might pay close attention to. Most users are unaware that they can modify their browser settings to disable the accepting of cookies in order to protect their information. All news websites had this option available. The categories

which ended up being an anomaly were dating and health websites as six of them did not have this option available. For a high number of blogging and search engines the answer was hard to conclude after analyzing their privacy policy.

20. Can a user opt-out of behavioral or targeted advertising or marketing from the first party? – Yes/No, Unclear.

		FI	NW	AD	SO	BL	SE	JB	GM	HE	DT	Ttl
20	Yes	3	6	2	8	6	6	9	7	7	6	**60**
	No	1	0	0	1	2	3	0	0	2	3	**12**
	Unclear	6	4	8	1	2	1	1	3	1	1	**28**
	Total	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**100**

On sixty percent of the websites users can opt-out of behavioral or targeted advertising of marketing from the first party. The job websites gave users the most choice followed closely by the social networks at 90% and 80% respectively. It is interesting to see the fact that 28% of the websites are unclear with giving users the opt-out choice. The two leaders in this category are the adult sites and 80% unclear and 60% of the financial websites. This leaves only 12% of the websites being entrapping. With this being a preference based choice it is good to see that more than have the options give users the choice of whether or not they want advertisements.

21. Does the privacy policy address the option of opting out of third-party behavioral or targeted advertising? – Yes/No, Unclear.

		FI	NW	AD	SO	BL	SE	JB	GM	HE	DT	Ttl
21	Yes	2	8	1	4	3	8	7	6	7	4	**50**
	No	8	2	9	5	5	2	1	2	2	4	**40**
	Unclear	0	0	0	1	2	0	2	2	0	2	**9**
	N/A	0	0	0	0	0	0	0	0	1	0	**1**
	Total	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**100**

Of the choices, half of the websites addressed the option of opting out of third party behavioral or targeted advertising, led by news and search engines. Forty percent of the sites trap the user, with 90% of the adult sites leading in that category. The websites are generally clearer about third-party behavior relative to first party behavior, with a percentage difference of 28% to 9% respectively. Overall, the results were almost evenly split.

22. Whether a registered customer will be notified for significant changes made in the privacy policy? – Yes/No, Unclear.

		FI	NW	AD	SO	BL	SE	JB	GM	HE	DT	Ttl
22	Yes	0	7	6	9	9	4	5	7	9	10	**66**
	No	9	2	4	1	1	5	3	3	1	0	**29**
	Unclear	1	1	0	0	0	1	2	0	0	0	**5**
	Total	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**100**

Sixty-six percent will notify a registered customer for significant changes made in the privacy policy. There was quite the spread with all of the dating websites notifying you along with 90% of the social, blog and health websites. At the opposite end of the spectrum none of the financial websites notified you of policy updates. With all the information that goes into financial sites one would think those sites would notify the user. Twenty-nice percent of the websites flat out do no notify users lead by 90% of the financial sites. This could be a key issue for users that are concerned about the contents of the privacy policy who would like to be notified if the contents are modified in any way.

23. Whether there is customer service and contact available for queries and questions? – Yes/No, Unclear.

		FI	NW	AD	SO	BL	SE	JB	GM	HE	DT	Ttl
23	Yes	8	6	9	9	8	9	9	10	10	9	**87**
	No	2	4	1	1	0	1	1	0	0	1	**11**
	Unclear	0	0	0	0	2	0	0	0	0	0	**2**
	Total	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**100**

Eighty-seven percent of the websites have customer service contact information available with all the websites greater than 50% and both health and gaming websites maxed out at 100%. The news sites had the fewest customer service areas at 60%. On the other end the news site lead no with 40% of the sites. Twenty-percent of the blog sites represented the unclear sites in regard to having customer service.

24. Whether the policy is certified or verified by a regulating body or company? – Yes/No.

		FI	NW	AD	SO	BL	SE	JB	GM	HE	DT	Ttl
24	Yes	1	3	0	1	1	4	3	3	4	3	**23**
	No	9	7	10	9	9	6	7	7	6	7	**77**
	Total	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**100**

This question was a straightforward yes or no question. When it came to privacy policy certification 77% of the websites were a resounding no to regulation leaving only twenty-three websites regulated. The adult websites were 100% no with social, blog, and financial websites at a close 90%. The remaining website types are either 60% or 70% no. This can be a concern for users who would want some sort of assurance that the website they use is protected and regulated by an officially governed website.

25. If Yes to Question 24, what is the name of the company? – Name.

		FI	NW	AD	SO	BL	SE	JB	GM	HE	DT	Ttl
25	TRUSTe (only)	1	3	0	0	1	2	1	3	1	1	**13**
	Others	9	7	10	10	9	8	9	7	9	9	**87**
	Total	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**100**

This question was based on the previous question and saw to keeping track of the names of the regulating companies. Of the twenty companies that were regulated only 13 were regulated strictly by TRUSTe. This left 87 websites that fell under other for either having a mix of regulators or none at all.

6.3. Summary

After looking at the results for attributes across all categories, we can conclude that privacy policies are not designed in the same manner. There are similarities present in certain cases, e.g. placement of privacy policy link, but otherwise the websites tackle the key issues in various manners. We dive into further analysis and discussion of our results in our next chapter. There, we not only discuss our findings across the hundred websites, we also discuss the results on a per-category basis.

7. Discussion

7.1. Introduction

After the methodology was applied to the 10 different categories we had, the results were determined. These results were computed categorically as well as on an overall basis. Therefore these results were discussed category wise and a final overall analysis was also made for the 100 websites. This helped develop consistent trends within a category and served as a basis for comparison, while the overall analysis served to summarize the categorical results through a broader perspective.

7.2. Category Analysis

7.2.1. Financial

The average readability FRES score is 35.80, which is lower than the average for all companies. This suggests that it is relatively more difficult to understand the privacy policies for the financial industry. This result is expected due to legal and financial terminology that must be used in these privacy policies. This indicates even greater importance to high-school or college education to fully comprehend what some of these policies truly define.

90% of companies mention that they pass information to third parties, which is quite a high percentage. 67% of these companies require user consent and the same percentage pass personal information. Most interestingly, none of the companies suggested that when such information is passed, the information shall be governed by the third party's privacy policy. This result is quite different from the overall findings and suggests that the financial industry adheres to strong industry and legal norms when specifying how customer information is utilized.

90% of companies mention that they pass information to affiliates. 89% of them require user consent, and the same percentage of companies pass personal information. These values are higher than the overall values, suggesting greater transparency in the financial sector again. In addition, none of them indicate that the information passed will be governed according to the affiliates' privacy policy. This, again, is quite surprising and displays that these companies specify the needs of using customer information.

The average number of affiliates listed is 8.5. Although this value is lower than the overall average, it is still relatively high when considering that 85% of overall companies did not list any affiliates. It is still necessary to note that just because these companies did not list any affiliates does not confirm that, in reality, they do not have any affiliates.

60% of companies allow third parties to store cookies from their own website, whereas 90% companies mention the option of disabling browser settings to block cookies. Almost all companies inform the users that they can choose not to accept cookies, indicating greater transparency again.

30% of companies mention the option of opting out of first party advertising and only 20% of companies mention the option of opting out of third party behavioral or targeted advertising. Both these values are much lower than the average. This result suggests greater advertising importance in the financial sector.

Interestingly, no company from our financial sample will directly notify the user of changes made in the privacy policy. This finding is quite unexpected as this sector seemed to possess greater transparency than the overall sample. This policy is definitely an industry norm but the reasons behind this cannot be confirmed. It could be due to the fact that some of these

companies have a huge customer base which is not only spread across several countries but it is also changing every minute. In addition, not every aspect of the privacy policy applies to all customers; some parts of the policies only affect residents in specific geographical areas. It may not be possible to contact every active customer at a particular time. The companies may have decided to not inform all the customers explicitly to avoid hassle.

7.2.2. News

The analysis of the news websites bore some interesting results as well. All of the websites have their privacy policy link located at the bottom of the page adding to the ease of accessibility. 90% of the websites require 1 click to access the privacy policy, 10% of them require 2 clicks to access it. All of the websites have the date of last update mentioned. This ensures the user that the policy is regularly updated and can easily find out when this happened last. The average readability of the score of the websites is 33.85 which falls in the score range of difficult (using Table 2). This was a comparatively low readability score. And given that news websites rely heavily on the texts incorporated on their website this result was somewhat surprising. But also it is likely that people visiting these websites are mostly college graduates or at least have an educational level higher than high school.

All of the websites pass on information to third-parties, out of these websites, half of them pass this information with user consent. 30% of the websites pass on personal information to third parties. Out of all the websites, 50% of them deal with the sharing of information according to their privacy policy.

90% of these websites pass on information to their affiliates. Since news websites would have a greater number of affiliates this was expected. 22% of these websites do not require user

consent when passing on this information. 66% of these websites pass both personal and non-personal information indicating that there is no consistency.

The number of affiliates is not mentioned for 80% of the websites, only NBC has a list of 1 affiliate mentioned and Wallstreet journal had a list of 5 affiliates mentioned. In 80% of the websites, third parties are allowed to store cookies from the host company's website. A majority of the websites are not certified by a regulating company while a few of the websites are certified by TRUSTe.

7.2.3. Adult

The adult side had some peculiar tendencies to them. For starters two of the ten sites had no dedicated privacy policy tab at all instead having just a terms tab. The two sites were XNXXgalleries and Gayromeo while all the other sites had their links located at the bottom. The eight with the click at the bottom were also standard one clicks while gayromeos privacy policy is a two click located in it's terms page and xnxxgalleries doesn't have a privacy page at all. When it came to mentioning the date of the last update only a measly 30% were a yes while another 60% were no and one N/A. The overall average readability was significantly more difficult than the readability of the social sites with a score of 35.94. The hardest one to read was youporn with 29 and the easiest to read being LiveJasmin with a 47.78 which was rather on par with the social networks. This creates a massive range of 18.78, which by the Fleisch-Kincaid is a huge standard even larger than the social networks swing. The sites also all seem to have terms and conditions at the bottom of the page save Literotica who didn't seems to have a terms but only a privacy page. Seventy-percent of had simple one click terms and conditions while 10% were two-clicks and the last two sites were N/A's. Seventy-percent of the terms and conditions

mention privacy policies while 30% have non-applicable answers those three sites being Manhunt.com, Litterotica.com, and XNXXgalleries.com.

When it came to mentioning the collecting and sharing of information 50% of the sites did mention the information while 20% didn't, and thirty-percent were still non-applicable. Regarding the sharing of information to third-parties 80% were had a confirmed answer split 70% yes ten percent no leaving twenty percent unclear with those two sites being youporn and xnxx galleries. Those same numbers were recorded regarding user consent except that the two unclear ones were XNXX galleries and adult friend finder. The division of the information was such that 40% were personal/non-personal, 30% was just personal, 20% was unclear, and 10% was not available. Only two websites had regulations based on their own privacy policy, those two being gayromeo.com and realitykings.com, while the remaining 80% didn't address anything along those lines. 40% of the companies pass information to affiliates, 20% do not, and the remaining 40% are unclear. When it comes to user-consent of affiliate information 20% were yes, 10% were no, 50% was unclear, and the last 20% was not applicable. For data type 30% was both personal and non-personal, 20% were just personal, 30% was unclear. 10% of affiliate regulation was by the host company while 90% was not addressed. The average was zero, with 100% of the companies reflecting the zero stat for mentioned affiliates.

70% of the companies allow third-parties to use information in cookies, while the remaining 30% is distributed among the no, unclear, and N/A categories. Eighty percent are unclear if you even need consent while two are confirmed to need user consent. The user cannot opt-out of first-party targeted advertising for ninety-percent of the companies, leaving only one company where you can opt-out, the site being Adam4adam.com. For third-party targeting advertising it's split 60% yes and 40% no. 90% of the websites will notify you of updates of their

privacy policy leaving only XNXXgalleries not too. 100% of the websites do not have customer service, nor a verified policy.

7.2.4. Social

The social networks seem to have a general uniformity to them when it comes to their privacy policies. Most have their terms and policy in the same area with their icons located at the bottom of their respective home pages. It is through these icons where they can be accessed via a one-click link. Out of the ten pages only one was different, amounting to 10%, and had its privacy policy top instead of bottom.

Only 20% of those websites were more than one click. Comparatively speaking 93% of our analyzed websites had their policies on the bottom with 13% being more than one click. The site that differentiated itself from a bottom shelf link was Pinterest.com where scrolling down would only cause more pictures to appear. This site was also one of the two click sites, the other being Facebook, though it separated itself by not having a direct terms or policy link instead having them hidden under an about tab.

When it came to the date of the last update being listed, seventy percent of the sites had a last update date displayed while the rest fell into no. The average readability score based on the Flesch Kincaid system was 42.82 where 0 is hardest to read and 90 is the easiest. This put the scores average closer to the harder to read side of things. The hardest of these sites to read was LinkedIn with a flat 36 while on the opposite end of the spectrum was Pinterest with 51.7 making for a difference of 15.7. Having a range of 15 is something rather interesting given that that is about a sixth of the scale. Ninety percent of the sites had their privacy policies mentioned in their terms and agreements with only Xing not following the form. Xing was also the only website not to refer to the policies regarding the collecting and sharing of data.

All of the sites passed on information to third parties with user consent. The data shared was relatively well spread among our categories of personal/non-personal, strictly personal, strictly non-personal, unclear, or not available. The dispersion was 40%, 30%, 0%, 20%, and 10% respectively. How the policies were regulated was generally not addressed, and there was 20% that was regulated by the host. When it came to passing information to affiliates 40% of them did, 20% didn't, and the last 40% were unclear. Overall having an affiliate was something special as most sites didn't even mention it. One of sites had the information passed to affiliates regulated by their host rules everyone else was not addressed. None of the sites managed to mention the names of any of their affiliates and as a result the average was 0 with a hundred percent being zero. Another shared uniformity was that a hundred percent of the sites allowing third parties to use information stored in cookies. All sites save Facebook and MySpace need consent for the cookie usage where Facebook was unclear and MySpace just did it without consent.

The option to opt-out from first-party stood for all save Xing and MySpace where MySpace was once again a flat out no and Xing was unclear. When it came to opting-out for third party there a was a bit more of an even divide with 50% not allowing opting 40% allowing and 10% that was unclear. Surprisingly Xing was the only site that did not notify the user of an update to its privacy policy, it was also the only site without some form of customer service which seemed to have become a uniform standard among the social sites. Also surprising was the fact that LiveJournal.com was the only certified site that was backed by the Safe Harbor principles.

7.2.5. Blogging

For all the blog sites there seemed to be a relative uniformity similar to the layout of the social networks. For starters 90% of the sites had their buttons for privacy policy on the bottom

leaving only *Imgur* not following this protocol. The site interestingly enough didn't even have a dedicated privacy policy tab. *Imgur* was actually a two click site due to it's privacy policy being located within its terms, everything else was a one click. 90% of the sites notify users of the updates of privacy policy, the only site that doesn't is imgur. The average is 42.8 in terms of Flesch-Kincaid score and is neither as hard to read as the adult sites nor as easy to read as the social networks.

The locations of the terms and agreements section all align at the bottom of the websites. They're all one-click terms and conditions websites as well as all mention the websites privacy policy in the terms. The also all have policies that refer to collecting and sharing of information. 80% of the websites do pass on information to third parties leaving only word press and blog.com that don't pass information to other parties. Only 40% of those sites are guaranteed to have user consent when passing, while 30% is not consented. 20% is unclear and ten-percent is not applicable. 60% of the websites are personal/non-personal with the remaining 40% split between unclear and not applicable.

Half the websites comply with the host's privacy policy when it comes to allowances, 30% adhere to third-party principle, and the remainder is not addressed. Only three of the ten companies pass information to affiliates, half are unclear about what they do and the rest are split not applicable and definitely do not pass info to affiliates. 60% is not applicable when it comes to personal data. 30% is personal/non-personal and the last ten-percent is just personal. 20% is the host company privacy policy regulation, 10% follows the affiliate privacy standard and the rest are not addressed. The average for this set of website was .1 as blog.com provides one affiliate. 30% of the websites are unclear about whether or not they allow third party websites to

use store and use cookies, while the remaining seventy percent allow such a process. 40% need user consent, while twenty percent think the opposite.

The remaining forty are unclear about their intentions. 60% of the websites allow you to opt-out of first party behavioral targeted advertising while the remaining 40% is split unclear or no. For third-party advertising the information is a little more interesting in that only 30% allow you to opt out while half the websites will not let you opt out. 20% are unclear. 90% notify you of significant changes while the differentiated website is Squarespace. 80% have customer service and the remaining two are unclear. In terms of certification 90% sites are uncertified, with Squarespace being the only one certified by Safe Harbor.

7.2.6. Search Engines

The evaluation of the privacy policies of search engine websites bore some interesting results. The methodology was applied to 10 different search engines. Despite observing a high level of consistency for most of the attributes, there were a few discrepancies.

All of the links to the privacy policies of these websites are located at the bottom of the homepage. In most cases a single click suffices to navigate to the privacy policy meaning that the websites have a high level of accessibility. 80% of the websites mention the date of the last update. Most of the privacy policies were updated recently, which is expected of search engines as these websites are used frequently and have a large user base; therefore there is a greater need for updates.

The average readability score of the websites is calculated to be 46.52. The value falls in the score range of difficult and a minimal high school level is required to comprehend the policy. The score is relatively higher in comparison to the readability score of the other categories that were analyzed.

The terms of service pages for these websites are highly accessible, with most of the links being located at the bottom of the website and mostly requiring one click to access them. 90% of the websites mention the privacy policy in their terms of service page while only 30% of the websites refer their privacy policy when dealing with the collecting and sharing of information. This was somewhat surprising as one would expect the terms of services to be consistent with the privacy policy.

90% of the websites pass on information to third-parties, out of these websites, 44% of them pass it with user consent and more than half of them pass on personal information to the third parties. This is a significant portion of the total websites that share personal information. 33% of these websites deal with the sharing of information according to the third party's privacy policy which means that a significant percentage of these websites has minimal or no control over what happens to the information when it is passed on to the third parties.

All of these websites pass on information to their affiliates. Half of them require user consent while 70% of them pass on personal information. The privacy policies were clear in mentioning what kind of information is passed on to affiliates and what happens to the information once it is passed on. All of these websites deal with the sharing of information according to the host company's privacy policy.

In 70% of the websites, third parties are allowed to store cookies from the host company's website and 70% of them address the enabling and disabling of browser settings to accept. This is an important issue as most users are unaware of how browser settings can affect third party cookies.

7.2.7. Job

After evaluating the job website category a few interesting factors came into perspective. All of the websites in this category have their privacy policy link located at the bottom. In most cases, the privacy policy page can be accessed in one click. This adds to the accessibility of the privacy policy page making it easier for a user to locate and navigate to it.

The mean readability score for this category is 43.24, which corresponds to a minimal high school education level for comprehension. The score is relatively high in comparison to the other categories but still falls in the score range of difficult.

The accessibility of the terms of service page for all of the websites was significantly high, with 90% of the websites requiring just one click to access. 80% of the websites mention their privacy policy in the terms of service page and the same percentage refers to the policy in regards to the sharing of information. This was an important factor that was constant for most of the cases. This indicated that at a certain level the privacy policies were consistent with the corresponding terms of service.

A significant percentage of the websites pass on information to third-parties and out of these websites all of them pass on personal information to the third parties. In terms of sharing information the privacy policies in this category were transparent and more direct. Since a great amount of personal information is collected through resumes, leakage of information is expected. All of the websites in this category pass on information to their affiliates which is expected. It was unclear for 60% of the websites if this information was passed on with user consent. It was clear that all of the websites pass on personal information to their affiliates. This result can be expected as most websites have contracts and agreements with their affiliates.

7.2.8. Games

The evaluation of the privacy policy of gaming websites some some pretty unexpected results to some questions while most answers were consistent to what was expected. The methodology was exercised on the privacy policy of 10 different kinds of gaming sites. As a whole most of the gaming websites had pretty similar privacy policies.

All of the websites assessed had the location of the privacy policy link at the bottom on the homepage and was pretty distinct in most occasions. A single click would lead straight to the privacy policy in all cases with no pop ups or advertisements. A good number of the websites assessed had the date of the last update mentioned whereas what was startling was that popular websites like IGN and Gamespot did not have the date of last update mentioned which is normally imperious to a well-tailored privacy policy.

The mean readability score was reasonable mounting up to an almost 39 which is higher than the average for all companies. This suggested that it was easy to understand the privacy policy of most gaming websites and was probably due to the fact that people of all ages interacted with these websites and they wanted to make the articles of the policy well-defined for the younger generation.

In most of the cases, the Terms and Conditions or User Agreement did mention about the availability of the websites privacy policy however strikingly only one popular gaming website the League of Legends ended up being an incongruity. Except for one website all the websites passed on information to third parties and 60% of these websites asked for user consent so this indicated that the top tier gaming websites do observe some legal standards regarding this issue but a large number of gaming websites did not address the issue that under which privacy policy

will the revealed information be governed and this gave an indication that many gaming websites might not have control over the revealed information to third parties which could end up being a privacy issue.

70% of the gaming websites passed on information to affiliates and a reasonable number asked for user consent before revealing information, but again one of the most striking technical flaws discovered while investigating the privacy policy of this industry was the inability of gaming websites to properly address the imperative issue that discusses which privacy policy will govern the revealed information to third parties or affiliates. This, again, is quite surprising and shows that these companies should appropriately embark upon this issue since a large number of the companies were disclosing Personal Information to third parties and affiliates.

Only one website gave a list of affiliates and this was not a good outcome bearing in mind that user information was being transferred to affiliates. 90% of the gaming websites allowed third parties to store cookies from the host website which was very staggering however only 50% of the websites mentioned in their privacy policy about enabling or disabling browser setting to accept or reject cookies and only 60% addressed that how a user can opt-out of behavioral/targeted advertising or marketing from the first-party, indicating lesser transparency overall and uncovering further obscurity to the reader of the policy.

Almost all gaming websites would notify the users about changes they made to the policy and most of them also had the date of the last update mentioned. The date of update was of the present year in most websites so this indicated that most gaming websites had been updating their privacy policy over time and some of them were even regulated by regulating bodies like TRUSTe.

7.2.9. Healthcare

The evaluation of the privacy policy of health websites uncovered some pretty standard results to most questions. Reasonable conclusions were drawn and apart from certain things nothing inexplicable was deduced from the assessment.

All the websites assessed had the location of the privacy policy link at the bottom of the homepage whereas only one website was missing the link which was alarming to discover. The accessibility to the link was easy and unlike the websites of other industries the privacy policy icon found in the homepage was larger than usual indicating that the website wanted the user to view the policy prior to accessing health related. 90% of the websites evaluated had the date of the last update mentioned and seeing recent dates on most of the privacy policies suggested that the policies had been frequently updated.

The mean readability score was unusually low up to an almost 23. This is almost 13 points lower than the overall average readability score of all the websites from all the industries indicating that the reading ease of a privacy policy of a health website was best understood by a university graduate. The reason for this might be the complicated use of medical terminology and the health privacy guidelines that many of the assessed websites were conforming.

In exactly 80% of the cases the Terms and Conditions or User Agreement did mention about the availability of the websites privacy policy. Except for one website, all of them disclosed information to third parties and affiliates but a high number enquired for user consent before disclosing information. Unlike dating websites, health websites did not have a huge user bases and most of the health sites did not serve networking purposes where people could connect hence

the inability of health websites to address the issue of governance of revealed information was partly justified. Moreover not huge percentage of health website revealed personal information.

More than half of the health websites allowed third parties to store cookies from the host website which was staggering. None of the websites mentioned in their privacy policy about enabling or disabling browser setting to accept or reject cookies which we believe should have been addressed however a huge percentage addressed that how a user can opt-out of behavioral/targeted advertising or marketing from the first-party, indicating greater transparency.

As many as 70% of the websites said that they would notify the user of any significant changes made to the privacy policy. All the privacy policies of health websites had a customer service or contact information available and almost half of the privacy policies were verified by regulating companies that suggested that the health industry observes specific standards and regulations when dealing with web-based privacy policies.

7.2.10. Dating

After the assessment of dating websites different kinds of conclusions were construed. Most articles however were in synchrony with the results of other websites. Unlike gaming websites, within dating websites they were some pretty distinct results and hence gave a wider picture to the overall analysis.

80% of the websites assessed had the location of the privacy policy link at the bottom of the homepage whereas one website had it on top and one website was missing the link which was startling to discover. Pof.com was missing the link of the privacy policy and strikingly it was one of the most widely used dating websites. They had a single page that served as their Terms of Condition and Privacy Policy and its accessibility was complicated as it took three clicks to go to that page after figuring out its location using a search engine.

The mean readability score was unusually low up at almost 22. This is almost 14 points lower than the overall average readability score of all websites from all industries. The reason for this remains unclear because unlike the financial sector the use of hard terminology in the privacy policy of a dating website remains largely unjustified.

In most of the cases the Terms and Conditions or User Agreement did mention about the availability of a websites privacy policy, in one case however the terms and conditions were the same as the privacy policy and one website was missing the terms and conditions page. Except for one website all the websites passed on information to third parties however half of these websites asked for user consent before disclosing information. A large number of dating websites revealed both personal and non-personal information. Similar to gaming websites a large number of websites failed to address the imperative issue that discussed which privacy policy will govern the revealed information to third parties and the user had no clue that the provided personal information that is being shared to unknown third parties is under the control of the host website or not. A privacy concern that we believe needs to be resolved.

Again except for one website all dating websites revealed information to affiliates however for half of these websites it was unclear which information they revealed. Since dating websites do hold private user information, this issue should have been addressed explicitly in the privacy policy. Another shocking thing was that none of the websites gave a list of their affiliates; this means that the user do not know with which website is its personal or non-personal information being shared.

Almost all the dating websites allowed third parties to store cookies from the host website which was staggering but only 20% of the websites mentioned in their privacy policy about

enabling or disabling browser setting to accept or reject cookies. Many websites did not address how a user can opt-out of behavioral/targeted advertising or marketing from the first-party indicating that dating websites were heavily perpetrated with behavioral advertisements and cookie based mining of information for advertisement was extensive.

Almost all dating websites would notify the users about changes they made to the privacy policy since many of the websites had a huge user base. For better user satisfaction, some of them were even regulated by regulating bodies like Trustee or Spark.

7.3. Overall Analysis

Only 70% of our 100 sample websites mentioned the date of last update of the privacy policy. This result is unexpected, as one would expect almost all websites to mention the time of latest update. It suggests that almost one-third of websites are not paying attention to detail regarding their privacy policies. Users cannot even be certain whether the policies described within the privacy policies are up-to-date or not.

The average FRES readability score was 36.53. This result indicates that an average privacy policy is fairly difficult to understand, and one would at least require high-school degree or some college level education to understand the material presented in these policies. It is safe to assume that a high percentage of online users may not possess a high-school degree or some college level education. Our background research has indicated that the average education level of online users is slightly higher than the average education level in the nation. In that regard, it may not be too difficult for an average reader to understand these privacy policies.

The paper **"**Privacy policies as decision-making tools: an evaluation of online privacy notices", which was published in 2004 calculated the average FRES score for online privacy

policies to be 34.2. That score is lower than our result, indicating the difficulty in reading privacy policies has been decreasing over time. However, the previous paper only looked at 64 online privacy policies, which were not equally distributed amongst various industries. Therefore, the trend in privacy policy readability cannot be accurately confirmed from that comparison. The paper also suggested that most online privacy policies are outside the grasp of most internet users. We do not know the current education level of all (or mostly American) internet users to compare our results to previous findings regarding that matter. But if the readability level of online privacy policies is still hovering around the same range, the addition of internet users from all education backgrounds may still ensure that most privacy policies will be beyond the comprehension level of a significant proportion of internet users. We haven't found any other research paper that has looked into 10 different industries and compared the average readability scores amongst them. That is why our project gives a fresh look into current readability levels of online privacy policies.

A significantly high portion, 83% of websites, mentioned the availability of its privacy policy in their terms and conditions page, which was expected as the two policies are somewhat correlated. Surprisingly, 62% websites referred in some manner to at least one of the policies in its privacy policy. This result was not expected as the goals of the Terms and Conditions page differ from the goals of the Privacy Policy page. However, it appears that most websites prefer to discuss or mention at least one of the policies presented in their privacy policy in the Terms and Conditions page.

88% of companies indicated that they pass information to third-parties. 57% of these companies suggested that this information is passed with user consent at least in some cases (the rest either indicated that it was not, or did not directly specify this information). In 81% of cases

where the company passes information to third-parties, the information that is passed contains personal information. Surprisingly, approximately 30% of companies indicated that the third party is liable to use some portion of this information based on the third party's privacy policy. This is a relatively high figure as almost 30% of websites suggest that when information is passed to third parties, the third parties are allowed to use the customers' information according to the third parties' policies. A user may not be aware of this fact when he/she authorizes a first-party website to collect and share personal information. We cannot conclude whether these are the only companies that hold no liability for shared information, as many others might follow the same path without declaring it.

72% of websites indicate that they pass information to affiliates, which in hindsight seems a relatively low percentage as one would assume that all websites do. 21% of all websites directly mentioned that they pass information with user consent, therefore approximately 29% of the websites who pass information to affiliates require user consent. Approximately 79% websites that mention passing information to affiliates specified that the shared information includes personal information, which is surprisingly quite high. Most importantly, the policy under which the shared information is governed is not mentioned in most cases. This is not as troublesome as it is with information passed to third parties, mainly because affiliates usually deal with customer information based on similar privacy rules compared to the host party. In addition, the affiliates usually need customer information for specified purposes, and in most cases to carry out functions required by the host party.

The average number of affiliates provided in a website is 17. However, this number is heavily skewed as one particular website listed approximately 1600 affiliates. That conclusion can also be inferred from the fact that the median of the data set is 0. In 85% cases, a website

didn’t provide any names of affiliates whatsoever. Therefore, in most cases a company does not feel necessary or important to provide a list of affiliates.

Although 71% of websites mention that third parties are allowed to store cookies from the first party website, only 56% overall companies mention the availability of disabling cookies from the browser. This is crucial as many users may not be aware of this option, and almost half the companies do not want to convey the message to users. Even in cases when the company mentions the option, it generally follows with the message that disabling browser settings may affect a user’s functionality, interactivity and experience in the website. This policy indicates most companies do not want customers to disable browser settings to accept cookies. Half the overall websites directly mention how to opt-out of third party behavioral or targeted advertising, usually by directing the user to another website. This is not strange as a company does not regulate third-party targeted advertising. That may be why about half the companies choose to ignore mentioning this option to customers. In cases where a third party can perform targeted advertising on a host company’s website, the host company should mention how to opt-out of such advertising. We do not, however, know how many of them exist in in our analysis, as we didn’t dive into such granularity.

Most companies will notify the customer directly if there are any significant changes made to the privacy policy. Most companies will also provide customer service for assistance regarding the privacy policy, indicating that companies are placing importance on these policies and are willing to discuss customers’ queries and resolve confusion.

7.4. Additional Factors

As we were collecting data through analysis of website privacy policies, there were many factors that stood out in our research. Some of these factors could not be incorporated in our results as they did not directly impact our findings. These additional factors about the privacy policies are noted below:

- HDFC Bank's privacy policy was significantly smaller compared to the other policies in the financial sector.
- We could not find the Terms and Conditions page for US Bank.
- The default text font size for the privacy policy of The Weather Channel is extremely small. It was difficult to read such a policy. It is important to note that the FRES readability score does not take into account the font size. This could affect the ability of a customer to comprehend the meaning of the overall text.
- Both the privacy policy and the terms and conditions pages for NBC News and The Washington Post contained advertisements within. The privacy policy for The Washington Post was broken down into four separate pages, each containing advertisements.
- For the adult website, LiveJasmin, the privacy policy and the terms and conditions page is combined together and broken down into several clauses. The entire text was used to calculate the readability score.
- The adult website YouPorn's privacy policy is extremely small.
- The adult website XNXX Galleries did not contain a privacy policy.

7.5. Summary

The evaluation differed greatly over the respective industries. There are many crucial factors to take away from our findings and analysis. We notice in that in some cases, there seems to be an online trend, e.g. posting Privacy Policy and Terms and Conditions link at the bottom of a page. In other cases, there are industry trends where specific sectors seem to address various privacy policy issues differently. Moreover, the readability differs within a specific range across industries, indicating different demands of education level to comprehend the policies appropriately. We summarize and focus on the most critical results of our analysis in the next chapter.

8. Conclusion

Our project led to various interesting results. Analyzing a large set of data proved to be an extremely practical approach in order to draw solid conclusions. A significant part of our work involved analyzing 100 web privacy policies. Although the findings resulting from this analysis may not extend to all web privacy policies, we believe that some of the issues that were brought up may be common and applicable to a wide range of privacy policies. These issues include, but are not limited to, the disclosure of personal and non-personal information to third parties, the readability level of the text of a privacy policy and the accessibility of the privacy policy page on a webpage. The project has helped us to draw conclusions on how privacy policies are designed and the discrepancies they may have.

The evaluation differed greatly over the respective industries. Certain aspects of privacy policies need to be amended for user satisfaction and information security. Websites should give a list of affiliates and should explicitly address whose privacy policy would be administering the revealed information to third parties or affiliates. In this way users can predict how their information is being channeled and which websites would be holding on to the information.

Many health, dating and gaming websites do not address the imperative issue of which privacy policy would administer the information revealed to third parties or affiliates. Since most of them do not reveal the list of affiliates, the user cannot know where the user's provided information is being channeled. Cookie-based mining of information for advertisements purposes is also extensive. Most companies allow third parties to utilize information stored in cookies from the first-party website.

Through this IQP a great deal has been learned about the inner workings of Internet privacy policy. The policies on social networks, blogging sites, and adult sites all held interesting trends. The social networks tended to be much more user agreeable and thus easier to read among the privacy policies. Blogging sites tended to be a little more self-contained when it came to their policies. Adult sites were among some of the shortest policies, if they had a policy, and were rather linear in format. These three distinct styles tend to reflect the social group that they are aiming for and thus is an interesting aspect about privacy policies.

The privacy policies were seen to vary from industry to industry depending on what type of information the industry uses and how this information is further channeled to affiliates or third parties. For job websites, most users are prompted to upload a resume to the website database. Therefore most of the privacy polices under the job industry address how this information is shared and made available to employers around the world. This factor adds greater transparency to the category. Search engines predominantly track user information like browser type, location, recent searches through cookies and direct advertisements accordingly. This result was mainly expected from this particular industry. News websites had a relatively low readability score given the fact that most users from this category have a comparatively higher educational level. In conclusion there were takeaways from each category that could be justified through the nature and type of industry analyzed.

The financial industry seemed to be more transparent in structuring their privacy policies than most other industries. The privacy policies seem to adhere to strict industry standards when it comes to declaring how user information is utilized. However, none of the companies will notify the customers directly when changes are made to the privacy policies. This practice may

be due to the huge customer base some of these companies have, which spreads over international borders.

Most websites tended to post the links of privacy policy and terms and conditions pages at the bottom of the home page. The access to these pages usually requires one click from the home page. Although, in some cases, companies either did not have a privacy policy, or had a combined page for their privacy policy and terms and conditions. Not all privacy policies are willing to mention the date of last revision for the privacy policy. Most companies mention that they pass information to affiliates and third-parties, and a significantly large portion of them pass personal information. When information is passed to third-parties, the third-parties, in some cases, have the complete authority to utilize that information for their own purposes. Almost half the companies do not convey the message that the user has the choice and ability to disable browser settings of accepting cookies. Most companies will notify the customers when changes are made to the privacy policy and most of them also provide customer service for any queries.

The average readability FRES score for all websites is 36.53. This score indicates that a user must have a high-school degree or some college level education to understand an average privacy policy. The average score has not changed significantly over the past years, which is inferred from the results found in previous work. Although new regulations and standards have come in practice, companies have found ways to keep the content in the privacy policies around the same difficulty level. However, when previous research is taken into account, they are still too difficult for most Internet users to comprehend. The scores across all websites are almost normally distributed, which show that they do not have significant variance. However, the averages for dating and healthcare websites are way below the industry average. Healthcare websites may have to deal with greater standards and laws, and may have to refer to medical

terminology if necessary which could explain this phenomenon. However, it is surprising that interactive websites like the dating industry have formed such difficult to read privacy policies. This trend is not beneficial to customers as vital personal information is provided to both these industries, and customers will demand privacy policies which are easily understood by all populations.

Our research and conclusions have given us a better sense of the level of importance of web privacy policies. This IQP project has enabled us to analyze online privacy policies with greater depth and we hope our work will aid researchers in this field. In the next chapter, we discuss what implications our project may have on future research.

9. Future Work

In essence our work aims to help shed light on Internet privacy policies. From readability to location and formatting, we hope that our work may help be the springboard for further exploits into privacy policy researching.

With our data showing a clear lacking of certified and regulating bodies for 87% of our websites, it is a blatant indicator of the need for more research into why this number occurs. We did not have the time to get into the specifics of this particular point. It remains an interesting foundation for research. It may be worthwhile to compile research for this topic.

Another take on our data will be to apply it into researching user feedback about the 100 analyzed sites. Having user feedback applied to our data may produce some interesting views on our work. For example knowledge of website users found compared to readability scores may provide insight to the accuracy of the Fleisch-Kincaid system. This approach would either strengthen or weakened the system.

We also hope our information, after more research, would help make sure websites clarify the opt-out privileges in regards to first and third party policies. There should be a dedicated way for each type to be opted out of straightforwardly without having to worry about caches and such. With an installed opt-out program the last issue that customers will need to be worried about is lack of transparency and misuse of information. That addition would be a great exploration for those in the future. Our work overall could be used as a solid ground for those wishing to explore Internet privacy policies. It is our hope that our work will be expanded by future projects into finding meaningful and effective results.

10. Bibliography

(n.d.). Retrieved from Bank of America: https://www.bankofamerica.com/

(n.d.). Retrieved from JP Morgan Chase & Co.: http://www.jpmorganchase.com/corporate/Home/home.htm

(n.d.). Retrieved from CitiGroup: http://www.citigroup.com/citi/

(n.d.). Retrieved from Goldman Sachs: http://www.goldmansachs.com/

(n.d.). Retrieved from The New York Times: http://www.nytimes.com/

(n.d.). Retrieved from CNN: http://www.cnn.com/

(n.d.). Retrieved from The Wall Street Journal: http://online.wsj.com/home-page

Federal Trade Comission . (2012, 11 23). Retrieved 10 22, 2012, from http://www.ftc.gov/reports/privacy3/fairinfo.shtm

Acquisti, A., & Grossklags, J. (2005). Privacy and rationality in individual decision making. *Security and Privacy, IEEE - Volume 3: Issue 1*, 26-33.

Alexa - The Web Information Company. (n.d.). Retrieved from http://www.alexa.com/

Anton, A., Earp, J., He, Q., Stufflebeam, W., Bolchini, D., & Jensen, C. (2004). Financial privacy policies and the need for standardization. *Security and Privacy, IEEE*, 36-45.

Bing. (n.d.). Retrieved from http://www.bing.com/

Dogpile. (n.d.). Retrieved from http://www.dogpile.com/info.dogpl/search/home

Egelman, S., Cranor, L. F., & Chowdhury, A. (2006). An analysis of P3P-enabled web sites among top-20 search results. *International Conference on E-Commerce* (pp. 197 - 207). New York, NY: ACM.

Excite. (n.d.). Retrieved from http://excite.com/

Facebook. (n.d.). Retrieved from https://www.facebook.com/

Flavian, C., & Guinaliu, M. (2006). *Consumer trust, perceived security and privacy policy: Three basic elements of loyalty to a website.* Emerald Group Publishing Limited.

Flesch Readability Score. (n.d.). Retrieved from http://rfptemplates.technologyevaluation.com/readability-scores/flesch-reading-ease-readability-score.html

Google. (n.d.). Retrieved from https://www.google.com/

Graber, M. A., D'Alessandro, D. M., & Johnson-West, J. (2002). *Reading level of privacy policies on Internet health Web sites.* Iowa City, Iowa.

Jensen, C., & Potts, C. (2004). Privacy policies as decision-making tools: an evaluation of online privacy notices. *Conference on Human Factors and Computing Systems* (pp. 471-478). New York: ACM.

Jobing. (n.d.). Retrieved from http://phoenix.jobing.com/

Kaltcheva, V., Andrade, E. B., & Weitz, B. (2001). *Self-Disclosure on the Web: The Impact of Privacy Policy, Reward, and Company Reputation.* Florida.

Karata, J. (July 2005). Privacy in information technology: Designing to enable privacy policy management in organizations. *International Journal of Human-Computer Studies*, 153–174.

Monster. (n.d.). Retrieved from http://www.monster.com/

Myspace. (n.d.). Retrieved from http://www.myspace.com/

Pollach, I. (September 2007). What's wrong with online privacy policies? *COMMUNICATIONS OF THE ACM*.

Quantcast. (n.d.). Retrieved from https://www.quantcast.com/

Schunter, G. K. (2002). A Privacy Policy Model for Enterprises. *Proceedings of the 15th IEEE Computer Security Foundations Workshop*.

Simplyhired. (n.d.). Retrieved from http://www.simplyhired.com/

Table References. (n.d.). Retrieved from http://lisemcclendon.wordpress.com/tag/flesch/