CrowdVerify: Using the Crowd to Summarize Web Site Privacy Policies

Submitted in partial fulfillment of the requirements for

the degree of

Master of Science

in

Information Technology - Information Security

Angelia

B.A., Asia Pacific Management, Ritsumeikan Asia Pacific University

Carnegie Mellon University
Pittsburgh, PA

December, 2014

Acknowledgements

I would like to take this chance to express my gratitude towards people that helped me making this possible.

My research advisor, Jason Hong for his ideas, patience, and guidance throughout the whole process. Danny Fernandes, my reader, for the advises that helped on clearing my sight on the problem. Carolyn P. Rose, for her insight on the research truly helped me in so many occasions. Annabel Sun, whom I discuss the research problems I had with. Janusz Szczypula, for the statistical advises. CHIMPS Research Group members for their input about my research. Information Networking Institute staffs for their help and support. And finally my parents and friends who never stop believing in me.

This research was funded in part by the National Science Foundation (TWC-1422018). The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressly or implied, of the U.S. Government.

Abstract

The goal of this project is to design and evaluate crowd-based techniques to highlight unusual and unexpected parts of privacy policy on web sites, so users do not need to spend a long time finding parts of privacy policy that might be of concern to them. The proposed method is to rank parts of privacy policies by comparing them and make people choose the one that they think to be more important to know. With this method, we obtained several lists of privacy policy statements sorted according to their importance. By combining this method along with other sorting techniques, we managed to find the method that is most effective, time and money wise. We also found that we cannot consistently reach an absolute ordering of statements according to its importance with this method, but we deem that absolute ranking is not necessary since our main goal is to summarize the privacy policy. This finding suggests that crowdsourcing combined with ranking method can be used to summarize long documents such as privacy policies.

Table of Contents

List of Tables

List of Figures

1

Introduction

1.1 Privacy Policy

Privacy is the state of being free from unwanted or undue intrusion or disturbance in one's private life or affairs [6]. However, it is impossible to have complete privacy in the Internet due to websites tracking their users for building up detailed profiles for pinpoint ad-targeting[9]. For this reason, users need a document that can tell them what the websites do with the information they collect and here is where privacy policy comes to play its part. One study indicates that 87.5% of users in the study expects comprehensive information regarding the site's security and privacy policy [8] and another study indicates that both completeness of privacy policy and reputation of the company reduce the level of concern over self-disclosure [2].

However, most users never bother to read the policy. It is almost ten times more time consuming than completing income taxes [13], and even if users do read the policy, it is likely that they will misunderstand it due to the fact that often privacy policies in a convoluted way that makes them difficult to comprehend [15]. A study by Annenberg Public Policy Center of the University of Pennsylvania reported 66 percent of the 1,200 American adults interviewed mistakenly believed that "sites with

a privacy policy won't share data" even though they claimed that privacy policies are easy to understand [14]. As such, we need a time efficient method for users to understand the most important content of website's privacy policies.

Related Work

Having a complete and comprehensive privacy policy is necessary, but it means that the document will turn out pretty long. This is one of the main reasons why it takes so much time to read everyone of them other than its difficulty to be comprehended. Previous studies has attempted to tackle this problem in various ways such as labeling [11] and rating by crowd [5].

In 2013, Kelley has studied a method of summarizing privacy policy by labeling the information contained in privacy policy and displaying the result in a grid [11]. This proposed method uses Platform for Privacy Preferences (P3P) as its foundation. P3P is a standard machine-readable format for encoding the online privacy policy of a company or organization created by World Wide Web Consortium [11].

The study indicates that labeling makes it easier and pleasurable to understand privacy policy given it is displayed in simple to understand grid. The color coded labels also makes it easier for people to find which privacy policy is stronger. However, this method hides many specific details that might be present in the privacy policy's original text form, which might prevent people from fully understanding the privacy policy. We aim to address this issue by having people read the highlighted parts of privacy policies.

Next is a study done by Chaianuchittrakul [5]. In this study, the approach is to have crowd rate segments of privacy policy according to its importance to know, comfort/discomfort caused by its existence, and easiness/difficulty to be understood, and then visualize the result in a heat map.

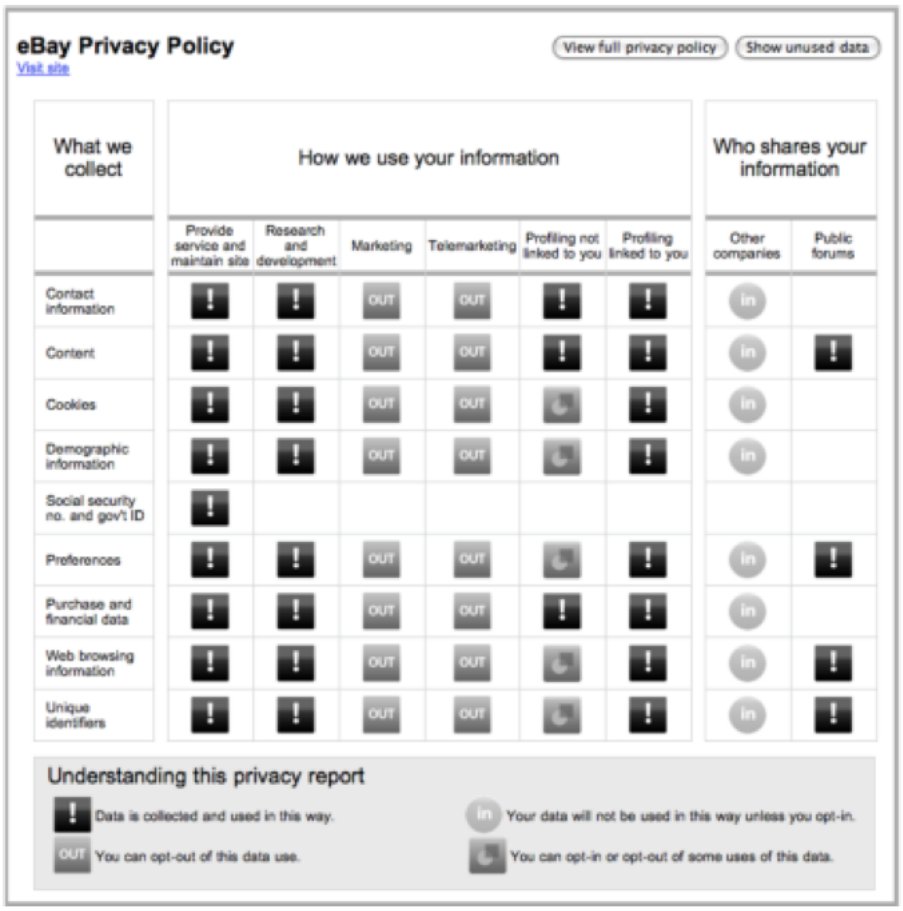


Figure 1.1: The Grid used in the study.

The study does not manage to find good enough result to generate a heat map of privacy policy due to the crowd constantly rating the segments as important, easy to read, and not surprising. It becomes clear in this study that without comparison, it is hard to determine which part to be highlighted in a privacy policy. To avoid the same result, we decided to make people compare the segments side by side and choose which one is more important instead of asking them to rate it one by one.

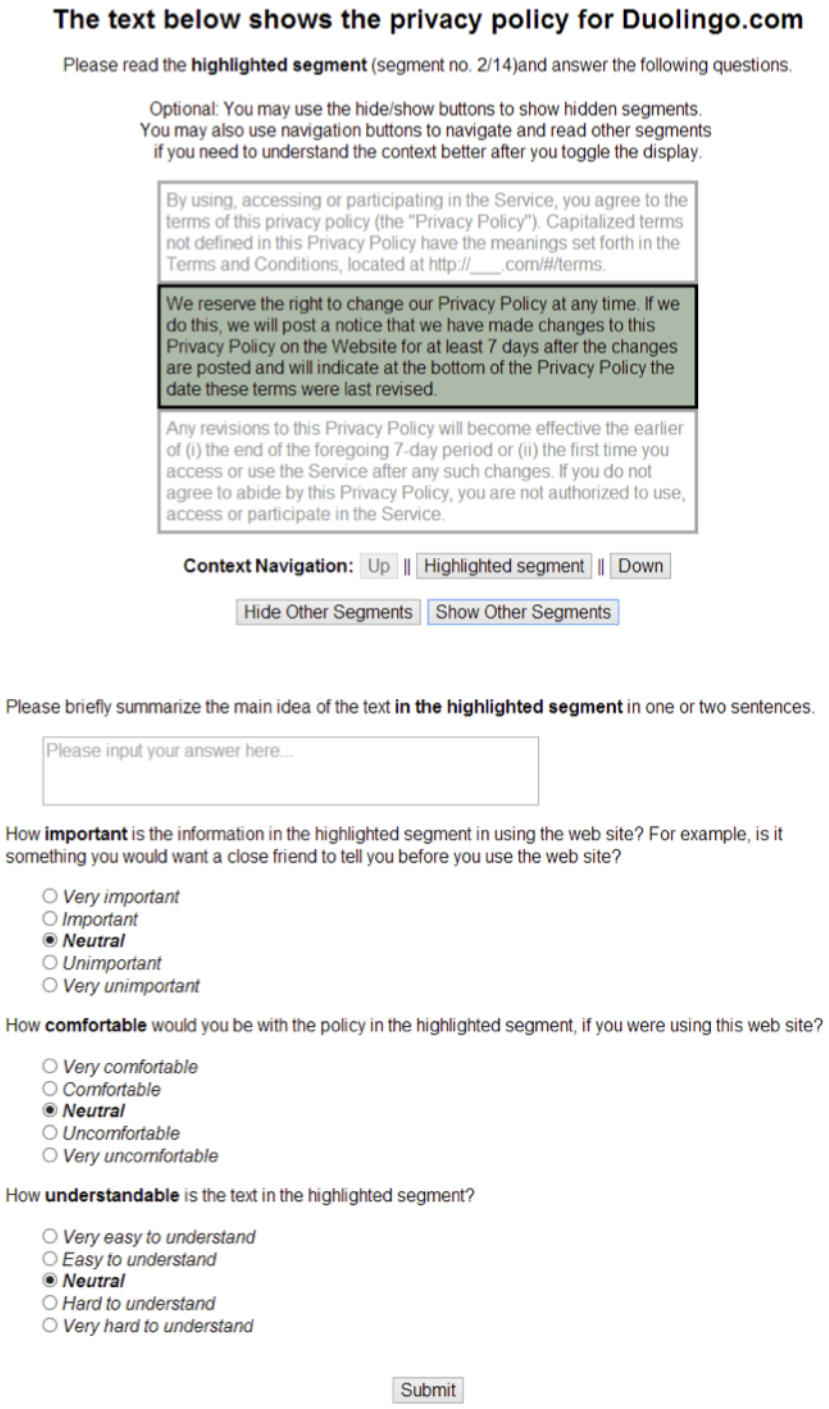


Figure 1.2: The interface of the experiment in the study.

1.2 Crowdsourcing

Crowdsourcing word itself is a fusion of 'crowds' and 'outsourcing', a term coined by Mark Robinson and Jeff Howe in an article in Wired magazine in June 2006 [7]. According to the article, the term refers to the act of outsourcing a function performed by employees to a large network of people in the form of an open call [7]. In business, crowdsourcing is an excellent mean of gathering on demand talent capacity and paying the workers per performance [4]. In research, crowdsourcing is used as a mean to collect data from underutilized resources such as patients of lung disease [1] or resources that is hard to find such as participants for survey other than university students [3]. In terms reliability, the survey data from crowdsourcing sample is as good as or better than the corresponding university sample [3].

Related Work

Our choice of crowdsourcing platform is Amazon Mechanical Turk. Amazon Mechanical Turk (MTurk) is a crowdsourcing marketplace for work that requires human intelligence. Mechanical Turk has proven to be cheap and has fast turn-around times in studies [10; 12], but it is not without disadvantages. Kittur et al. describes several challenges in using MTurk, including lack of demographic information and the possibility of gaming or cheating the system [12]. While our research does not involve any demographic information except for the location (participants need to be located inside United States), we are concerned on acquiring accurate results. We address this issue, by using highly accurate workers provided by MTurk.

1.3 Research Objective

Our objective is to develop a new user-friendly approach design and evaluate crowd-based techniques to highlight unusual and unexpected content of a website's privacy policy. We divided the privacy policy into statements and utilized crowd-workers from Amazon Mechanical Turk to work on this method by choosing which statement is more important to know over the other statement, one at a time. The aggregated result will show us a list of statements of privacy policy ranked by their importance. Statements that score high are deemed more important to know over the other, and they will be our 'highlighted' content. The results we aim to produce are the following:

- a group of highlighted statements, and
- a working method to highlight privacy policies.

1.4 Overview of Results

We conducted three experiments in total and one pre-test before the experiments. Each used the same tournament method we propose (introduced in Chapter 2), but are handled differently to test for robustness and efficiency. The test and experiments are as follows:

1. The testing phase is done to find out if the method shows any promise of working. We find that even with small number of people and input, and even errors, we still get enough scores to determine the strongest/weakest statements.
2. The first experiment is to test the difference between normal and master (high accuracy) workers. We find no statistically significant difference in results, but it is shown that the master workers take more time on choosing which

statement is more important over the other. Taking time into consideration, we decided to employ master workers on the next experiments.

3. The second experiment is to turn all the statements into one big group of pairing and perform the survey multiple times to see if the scores will stabilize over time. We also take one statement away from a pair to test for the method's robustness. We find that the method is indeed robust, but the scores do not stabilize over time. We also find this method is unreliable on whether the statements are paired with different statement each time. We also find this method is too time consuming and expensive, which lead us to the last experiment.

4. The last experiment combines the proposed method with round-robin method. Instead of putting all pairs into one big group like in second experiment, we split them into groups of five and have lesser number of people working on each group. We find that this method is more reliable in pairing statements and more efficient in converging the score of the strongest/weakest statement. It is also cheaper compared to the second experiment.

The method we propose cannot achieve a stable score or absolute ranking in any experiment. Despite that, we can still approximately group the statements according to their importance because the statements' score does not change by much. We also find unexpected data in our results, such as how the workers choose a statement which information should be common sense as the most important to know statement.

2

Methodology

In this section we introduce the survey method as well as the tools we used to perform the survey.

2.1 Principle of the Method

www.kittenwar.com is a cutest kitten polling website. The way they do it is by displaying two cat pictures and have the visitors click on the picture they think is cuter. Visitors can refresh the page if they cannot decide which cat is cuter. As time goes by, more match between kittens is done, and the cat with most wins and loses will be revealed in the 'Winningest Kittens' and 'Losingest Kitten.'

Using this kittenwar website as the base, we devised a summarizing method which let people compare parts or a statement of a privacy policy with each other and chose according to the criteria of our choice in a survey. The statements that were chosen more would be the 'summary' of the privacy policy according to the chosen criteria. In this study, however, we will use only one criteria; the importance to know. In other words, all 'top scorers' in these experiments are the statements which most survey participants thinks to be more important to know than the other.

Figure 2.1: The homepage of www.kittenwar.com. The visitors click on the cat they think is cuter and refresh the page if they can't decide.

2.2 Data Collection

2.2.1 Tools

Amazon Mechanical Turk

We used Amazon Mechanical Turk as a platform for our survey. Amazon Mechanical Turk (MTurk) is a crowdsourcing marketplace for work that requires human intelligence, such as answering surveys, identifying objects in a video, or transcribing audio recordings. These kind of tasks are called Human Intelligence Task or HIT for short. The workers of MTurk is paid for each HIT they do.

The reason we chose MTurk over other online survey platforms is because we can select the participants according to their accuracy rate. The high accuracy

participants are called Master Workers in MTurk. These Masters have demonstrated high accuracy across a variety of Requesters and must pass statistical monitoring in order to remain as Masters. All workers will be paid after we review and approve their work.

Microsoft Excel

MTurk's input and output is in csv format, so we use Microsoft Excel, to work with these files. We also use it to handle other research tasks such as analyzing the data and graphing.

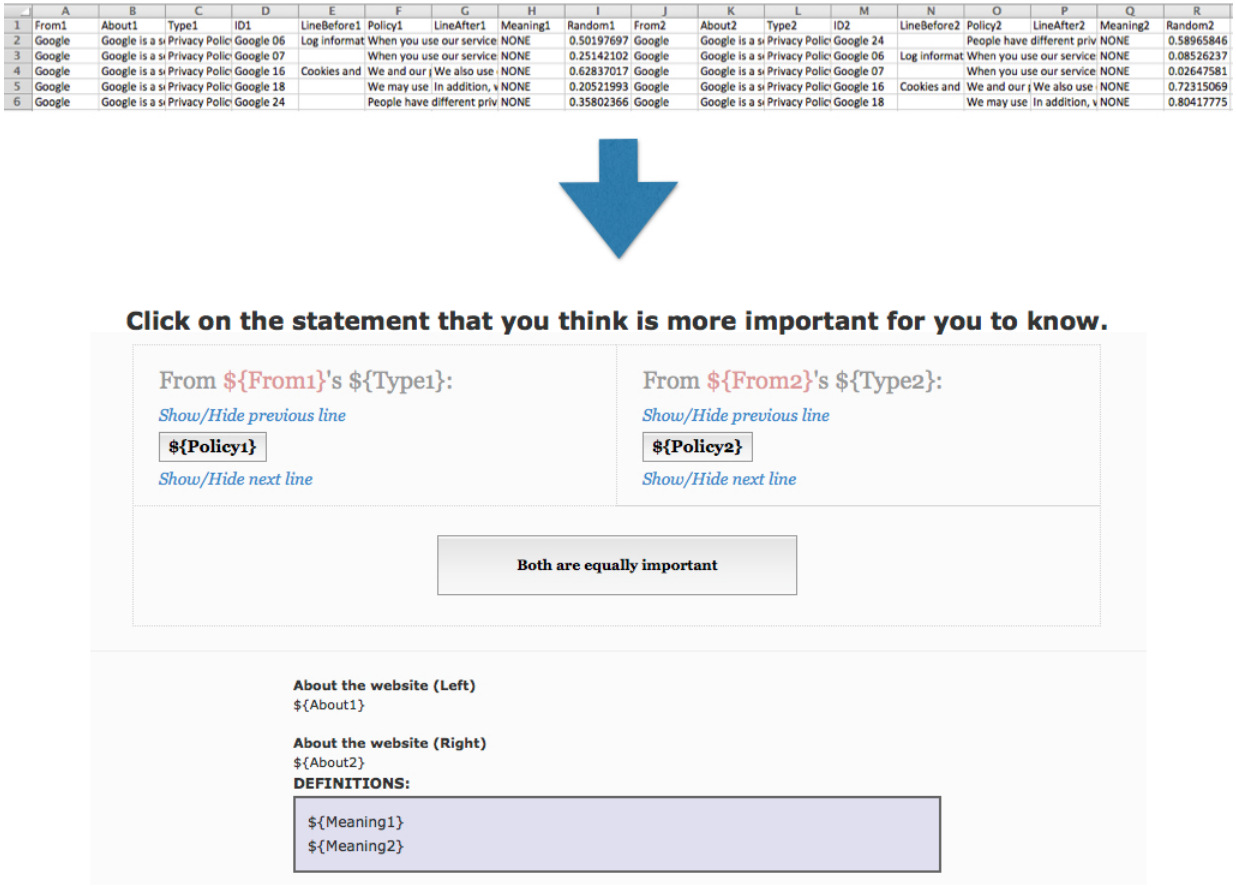
Python Script

We assign scores to each chosen statement. However, the output we receive from MTurk does not automatically aggregate the score, so we wrote a small python script to automatically aggregate the score and automatically sort the statements according to the highest scorer.

2.2.2 Input

We convert the privacy policy to input file by manually selecting the statements or sentences that seems to be important or representative for the whole paragraph (the goal is to summarize the policy, so we chose not to display everything). We give an ID to each statements according to their order of appearance for ease of process. For example, the sentence of the first paragraph will be given ID Google01, while the sentence of the next paragraph will be given ID Google02. Then we put those statements together in a csv file that will be uploaded to MTurk, which will then be converted to a survey according to the template we previously designed.

The privacy policy we chose as test input for the whole experiment is Google Privacy policy. We chose it as the test input due to the website's fame and the variability of services the privacy policy has to cover.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	From1	About1	Type1	ID1	LineBefore1	Policy1	LineAfter1	Meaning1	Random1	From2	About2	Type2	ID2	LineBefore2	Policy2	LineAfter2	Meaning2	Random2
2	Google	Google is a s	Privacy Polic	Google 06	Log informat	When you use our service		NONE	0.50197697	Google	Google is a s	Privacy Polic	Google 24		People have different priv		NONE	0.58965846
3	Google	Google is a s	Privacy Polic	Google 07		When you use our service		NONE	0.25142102	Google	Google is a s	Privacy Polic	Google 06	Log informat	When you use our service		NONE	0.08526237
4	Google	Google is a s	Privacy Polic	Google 16	Cookies and	We and our p	We also use	NONE	0.62837017	Google	Google is a s	Privacy Polic	Google 07		When you use our service		NONE	0.02647581
5	Google	Google is a s	Privacy Polic	Google 18		We may use	In addition, v	NONE	0.20521993	Google	Google is a s	Privacy Polic	Google 16	Cookies and	We and our p	We also use	NONE	0.72315069
6	Google	Google is a s	Privacy Polic	Google 24		People have different priv		NONE	0.35802366	Google	Google is a s	Privacy Polic	Google 18		We may use	In addition, v	NONE	0.80417775

Figure 2.2: The csv file and MTurk template. The template will be filled according to the header in the csv file.

2.2.3 Survey

On the survey, workers will see two statements besides each other. They will have to click to choose which one is more important to know. The participants can also see the statement's previous line and next line to understand the context of the statement. If they don't know anything about the website of the privacy policy, or the meaning of a particular word, they can refer to the about section and the definitions box below the line.

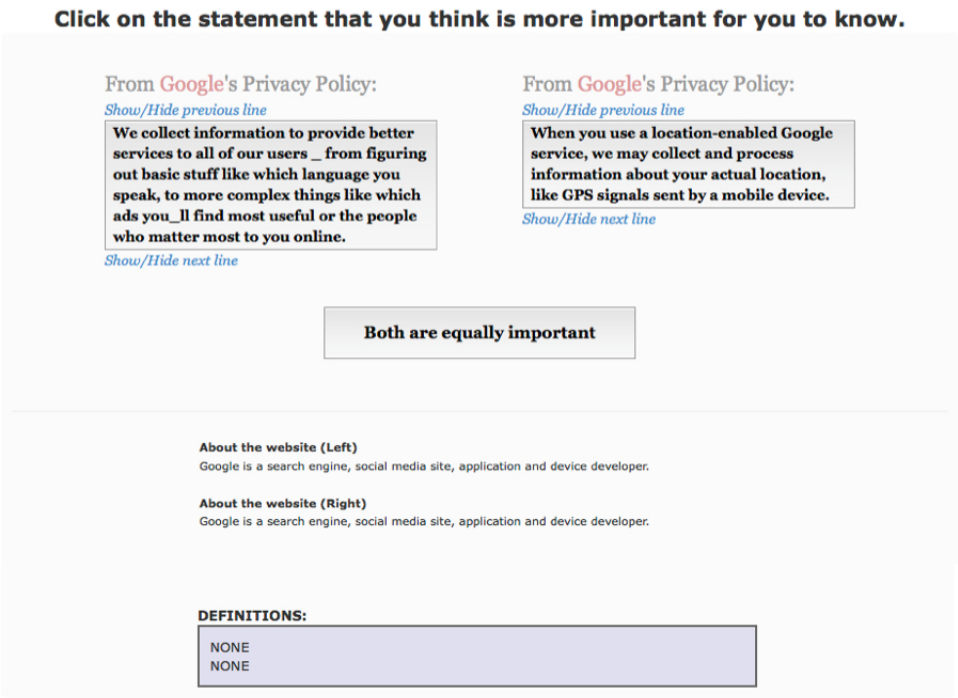


Figure 2.3: This is how the survey looks like to the MTurk workers. The workers choose by clicking the statement or the 'Both are equally important' button.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	From1	About1	Type1	ID1	LineBefore1	Policy1	LineAfter1	Meaning1	Random1	From2	About2	Type2	ID2	LineBefore2	Policy2	LineAfter2	Meaning2	Random2
2	Google	Google is a s	Privacy Polic	Google 06	Log informat	When you use our service		NONE	0.50197697	Google	Google is a s	Privacy Polic	Google 24		People have different priv		NONE	0.58965846
3	Google	Google is a s	Privacy Polic	Google 07		When you use our service		NONE	0.25142102	Google	Google is a s	Privacy Polic	Google 06	Log informat	When you use our service		NONE	0.08526237
4	Google	Google is a s	Privacy Polic	Google 16	Cookies and	We and our	We also use	NONE	0.62837017	Google	Google is a s	Privacy Polic	Google 07		When you use our service		NONE	0.02647581
5	Google	Google is a s	Privacy Polic	Google 18		We may use	In addition, v	NONE	0.20521993	Google	Google is a s	Privacy Polic	Google 16	Cookies and	We and our	We also use	NONE	0.72315069
6	Google	Google is a s	Privacy Polic	Google 24		People have different priv		NONE	0.35802366	Google	Google is a s	Privacy Polic	Google 18		We may use	In addition, v	NONE	0.80417775

Figure 2.4: This is how the end result of pairing looks like. This file will be uploaded to MTurk, which then will be converted to the survey according to the template.

2.2.4 Pairing

We refer the two statements displayed at the same time as a pair. In order to create pairs of statements, we used Excel's RAND() function to produce random numbers and sort a list of statements according to it. We repeat the process to the same list of statements to produce another randomly ordered list of statements to pair with the previous list of randomly ordered statements. This method produces a list of randomly paired statements that will be displayed on each HIT.

2.2.5 Scoring and Cost

Scoring

On the survey, the chosen statement will be given a score of 1. The one that is not chosen will not receive any score. If both statements are equally important, then both will receive half score. The number of workers participating in the survey will determine the maximum score a statement can get. The maximum score will be:

$$Max_{score} = 2w \tag{2.1}$$

where w is the number of unique workers participating in the survey. The maximum score is the number of workers times two because the statement appear two times on each survey (refer figure 3.4 ID1 and ID2 column to see how a statement can appear two times). For example, if there are five workers working on a survey, Google06 can get ten points if all five people choose Google06 whenever it appears. The score can get as low as zero if the statement is never chosen by workers.

Cost

Workers are paid per task or HIT. One pair is counted as one HIT and the survey will be often held in a batch of HITs, or multiple pairs. The file in figure 3.4 have 5 pairs so it will produce 5 HITs when uploaded to MTurk. Workers will be paid \$0.1 per HIT, so the cost of each batch will be:

$$Cost = i \times g \times n \times w \times \$0.1 \tag{2.2}$$

where i is the number of iterations, g is the number of groups, n is the number of pairings in each group, and w is the number of workers.

2.2.6 General Process

This is the general workflow of how we do our experiment. The experiment varies, but this is the general workflow on each experiment.

1. Generate input by creating randomly paired statements.
2. Upload the input to MTurk. MTurk Will convert it to survey format according to the template.
3. Workers work on the survey.
4. Review the workers' work. We can approve or reject their work on MTurk.
5. Download and process the data with the Phyton script.
6. Analyze the results using Microsoft Excel.
7. Create new input based on the previous result. How the input is created depends on each experiment.
8. Analyze the whole data and determine if the experiment is successful or not.

2.3 Experiments

2.3.1 Testing Phase

We want to know if the proposed method works before we proceed, so we tested the method with MTurk Sandbox. The MTurk Sandbox allows us to test the survey in the exact same Amazon MTurk environment without the need to pay the workers. The participants of the test were non-native English speakers coming from various countries outside US recruited through personal network.

We used a small sample for this test. We have prepared ten pairs of statements, created from ten statements chosen out of thirty. We required fifteen number of

participants and each person will work on ten pairs of statements, so the total work to be submitted will be 150. The statements used are statement ID Google11 to Google20. We chose this range because of they contain both interesting and boring statements such as 'When you use our services or view content provided by Google, we may automatically collect and store certain information in server logs. This may include: Location information' (Google12) and 'We use information collected from cookies and other technologies, like pixel tags, to improve your user experience and the overall quality of our services' (Google20).

Hyphothesis

Even though the participants are non-native English speakers, we have a reasonable amount of confidence that they will understand the policy as they were all international students. Google's privacy policy are also easy to read, so it should pose no problem to the participants. However, this is their first time working with MTurk, so we expect them to have questions and inputs regarding the survey. Regarding the survey result, we expect to have a winner and a loser statement even with such a small number of statements and people.

Result

We gathered 12 out of 15 people needed for the test. Out of the 120 HITs that are supposed to be completed, there are only 116 HITs submitted. In total, there are only 77% completed HITs out of 150. The highest scorer is Google 13 and the lowest scorer is Google 20. The difference between the top scorer and the second top scorer is small.

Discussion

As we expected, participants have no problem understanding the policy. However, some have expressed confusion regarding the instruction itself. Apparently, it is

Statement ID	Score
Google 13	15.5
Google 12	14
Google 14	12
Google 16	11.5
Google 11	11.5
Google 19	11
Google 18	11
Google 15	11
Google 17	10
Google 20	7.5

Table 2.1: The result of the testing phase, sorted from the highest to the lowest score.

unclear what is the base of the importance criteria, so we changed the instruction from 'Click on the statement that you think is more important' to 'Click on the statement that you think is more important for you to know.'

The participants also reported various problems such as blank buttons which produced errors upon click. The cause of it was misplacement of the statement on the input file which leave the button blank. This blank boxes was also the primary cause of incomplete submission of HITs because the participants opted to skip the HIT with blank buttons.

However, despite these setbacks, the result of the survey was satisfactory. As can be seen on table 3.1, the statements received enough score from the survey to determine the winner and loser. We deemed that this approach showed a promising result to continue the experiment.

2.3.2 Normal and Master Workers

The main difference between normal and master workers are their number and accuracy level. Normal workers are larger in number while master workers are smaller, but has higher accuracy. For our first experiment we wanted to test if there was a

difference of result between normal and master workers. For that purpose, we made the normal and master workers work on the same survey.

We use 15 workers located in US per survey and put all 30 statements we prepared on the survey.

Hypothesis

There should be a difference of result between normal and master workers due to difference in accuracy. The method should also still work in larger number of statements.

Result

There are differences of result between the normal and master worker result-wise (Table B1). However, the t-test (Table B2) shows that there is no statistically significant difference between the two groups. The normal workers take 35 seconds to do one HIT on average while master workers take 61 seconds on average. Like the testing phase, the result of the survey statements gives clear winner and loser out of the statements.

Discussion

Most of the statements get very different scores from the normal and master workers, but it is not statistically significant. There is also not enough evidence to show that the difference is caused by the two group's difference in accuracy level and/or difference in the time they take to choose, but we can say that master workers do take time in reading and choosing the statements. With this result in consideration, we choose to do the experiment with the master workers for the next experiments.

2.3.3 Frequency Experiment

For the next experiment, we chose to do the simplest method by performing the survey multiple times, each time with different randomly paired statements. Our goal was to see if the statements' rank will stabilize as more surveys are done. We designed the experiment to be as random as possible. The previous result would carry no weight over the next survey, and the workers also could not see the result of the previous survey. This experiment was done with the same set of policies and same number of workers with the first experiment.

On this experiment we also further tested the robustness of this method. We removed one statement from a pair to see if there's a significant difference with the results. The workers would see a blank box in the place of the removed statement in this broken survey and can still click on it (their choice is counted).

We spent a grand total of $225 for this experiment (B.1).

Hyphothesis

We expected a change of ordering across iterations, but we assume it would eventually show a trend over time. We also expected that there would be a difference in both broken and normal survey.

Result

As can be seen on Table 2.2, the ordering changed each time, and it didn't converge even after fifth iteration. Though it might seems like the ordering is very different between broken and normal iterations, the t-Test calculation on the standard deviations showed no significant difference (Table B5) between the two grups.

Broken Iteration						Normal Iteration			
1^{st} Iteration		2^{nd} Iteration		3^{rd} Iteration		4^{th} Iteration		5^{th} Iteration	
Statement ID	Score	Statement ID	Score	Statement ID	Score	Statement ID	Score	Statement ID	Score
Google 07	29.5	Google 04	29	Google 30	28.5	Google 30	27	Google 08	25
Google 08	29	Google 08	28.5	Google 10	26	Google 13	26.5	Google 13	24
Google 13	28	Google 12	27.5	Google 11	25.5	Google 07	26.5	Google 03	24
Google 30	25.5	Google 07	27.5	Google 12	24.5	Google 04	24.5	Google 12	23
Google 04	22	Google 13	26.5	Google 08	23.5	Google 08	24	Google 10	22
Google 10	21	Google 16	22.5	Google 13	23	Google 10	21.5	Google 07	22
Google 09	19.5	Google 11	22.5	Google 09	22.5	Google 01	21	Google 16	20.5
Google 18	18.5	Google 10	22.5	Google 05	21.5	Google 05	19	Google 11	20
Google 12	18	Google 03	20	Google 07	18.5	Google 02	18.5	Google 23	19.5
Google 16	17.5	Google 28	18	Google 16	18	Google 16	16.5	Google 28	18
Google 03	17.5	Google 23	18	Google 14	18	Google 09	16.5	Google 05	18
Google 05	17	Google 30	17.5	Google 03	18	Google 28	16	Google 30	17.5
Google 28	16	Google 06	17.5	Google 18	16.5	Google 27	15	Google 06	17
Google 21	16	Google 02	17	Google 21	16	Google 12	15	Google 02	16.5
Google 01	15.5	Google 05	15.5	Google 15	15	Google 20	14.5	Google 04	16
Google 27	15	Google 24	15	Google 02	15	Google 18	14.5	Google 24	15
Google 14	14.5	Google 09	15	Google 01	14.5	Google 14	14.5	Google 09	15
Google 06	14.5	Google 21	13	Google 04	14	Google 06	14.5	Google 21	14.5
Google 23	14	Google 14	12	Google 23	13.5	Google 03	14.5	Google 14	14.5
Google 15	12.5	Google 19	11.5	Google 06	13.5	Google 15	14	Google 19	13.5
Google 02	12.5	Google 15	11	Google 28	13	Google 11	13	Google 15	12.5
Google 11	11.5	Google 18	10.5	Google 17	10.5	Google 21	12.5	Google 18	11.5
Google 20	9.5	Google 01	7.5	Google 27	9.5	Google 24	12	Google 20	9
Google 24	9	Google 20	5.5	Google 20	8.5	Google 23	9.5	Google 27	8.5
Google 29	7.5	Google 26	4.5	Google 19	8.5	Google 22	8.5	Google 01	8.5
Google 25	5.5	Google 29	4	Google 24	6.5	Google 29	6	Google 29	6.5
Google 17	4.5	Google 27	4	Google 22	4	Google 17	5.5	Google 17	6.5
Google 26	3.5	Google 25	3	Google 26	2	Google 25	4.5	Google 26	5
Google 22	3.5	Google 17	3	Google 25	1.5	Google 19	3.5	Google 22	3.5
Google 19	2	Google 22	0.5	Google 29	0.5	Google 26	1	Google 25	3

Table 2.2: The result of the first experiment, sorted from the highest to the lowest score on each iteration. As can be seen, the ordering is not absolute across the iterations.

Discussion

Although the result varies at a glance (Table B3), it is not statistically significant according to our calculation Table B5. We deem that this method is robust enough against a lack of one statement error.

As for the trend, although we observed a trend growing over time, we found little graphically observable difference between the top scorer and the second highest scorer and between the lowest scorer and the second lowest scorer (Figure 2.6).

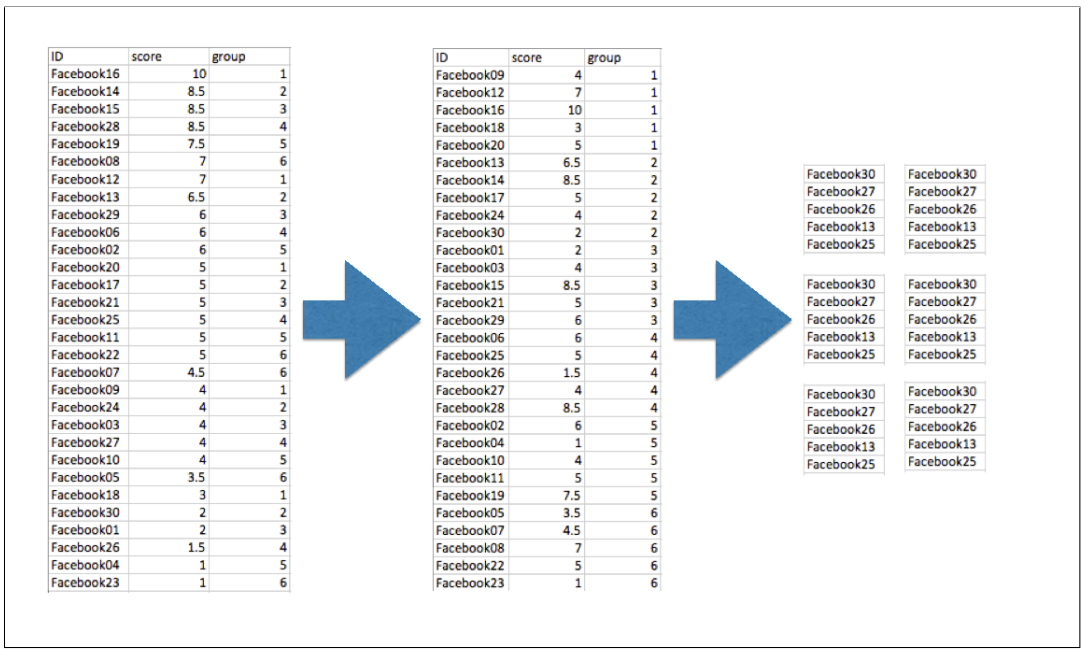
However, with the frequency method, we have no guarantee if all pairs are thor-

oughly tested with each other. A statement can be strong against one statement, but weak for when going against the other statement. For example, there's a possibility that Google 04 was able to receive a high score and reach the top at second iteration because it's only against weak statements on that iteration. As we spent a grand total of $225 for this experiment (B.1), we went to look for another less costly and more reliable method to achieve the same result, which led us to the round robin method.

2.3.4 Round-Robin Method

In the third experiment, we use the round-robin method to make sure that the statements meet other statements in other iterations. The steps of the method is as follows:

1. Convert Privacy Policy into a list of statements.
2. Randomize the order of the privacy policy statements.
3. Split them into groups of n statements (each group must have the same number of statements).
4. Run the survey.
5. Using the result of the survey, order the statements from the highest to the lowest scorer.
6. Assign numbers to the statements and group them according to the assigned numbers. The assigned numbers must be no bigger than the number of groups. For example, if there are six groups, then the statements must be given number from one to six and repeat (Figure 2.5).
7. Run the survey again with the new groups and repeat from step 5.



ID	score	group
Facebook16	10	1
Facebook14	8.5	2
Facebook15	8.5	3
Facebook28	8.5	4
Facebook19	7.5	5
Facebook08	7	6
Facebook12	7	1
Facebook13	6.5	2
Facebook29	6	3
Facebook06	6	4
Facebook02	6	5
Facebook20	5	1
Facebook17	5	2
Facebook21	5	3
Facebook25	5	4
Facebook11	5	5
Facebook22	5	6
Facebook07	4.5	6
Facebook09	4	1
Facebook24	4	2
Facebook03	4	3
Facebook27	4	4
Facebook10	4	5
Facebook05	3.5	6
Facebook18	3	1
Facebook30	2	2
Facebook01	2	3
Facebook26	1.5	4
Facebook04	1	5
Facebook23	1	6

ID	score	group
Facebook09	4	1
Facebook12	7	1
Facebook16	10	1
Facebook18	3	1
Facebook20	5	1
Facebook13	6.5	2
Facebook14	8.5	2
Facebook17	5	2
Facebook24	4	2
Facebook30	2	2
Facebook01	2	3
Facebook03	4	3
Facebook15	8.5	3
Facebook21	5	3
Facebook29	6	3
Facebook06	6	4
Facebook25	5	4
Facebook26	1.5	4
Facebook27	4	4
Facebook28	8.5	4
Facebook02	6	5
Facebook04	1	5
Facebook10	4	5
Facebook11	5	5
Facebook19	7.5	5
Facebook05	3.5	6
Facebook07	4.5	6
Facebook08	7	6
Facebook22	5	6
Facebook23	1	6

Facebook30	Facebook30
Facebook27	Facebook27
Facebook26	Facebook26
Facebook13	Facebook13
Facebook25	Facebook25

Facebook30	Facebook30
Facebook27	Facebook27
Facebook26	Facebook26
Facebook13	Facebook13
Facebook25	Facebook25

Facebook30	Facebook30
Facebook27	Facebook27
Facebook26	Facebook26
Facebook13	Facebook13
Facebook25	Facebook25

Figure 2.5: An example of how the statements are grouped according to the result of the previous survey.

Like the frequency experiment, we use Master workers located in United states. We randomly assigned five Masters to work on each group. The number of workers per group will not change even if the number of statements per group or the number of group itself increases. We also use not only Google privacy policy, but also Facebook privacy policy as input.

Hyphothesis

With this method, we split the rank ordered statements into segments where the split is right at where the numbering repeats. This method ensures the statements to be pitted against other statements with different strength, but also similar at the same time. For example, group 1 is filled with the strongest statements of each segments and group 6 is filled with the weakest statements of each segments. Again, we expected the ranking to converge at some point after several iterations in a rate faster than the frequency method.

Result

At a glance, there seems to be no particular difference between the frequency method (Table B4) and the round-robin method (Table B6). Google 08 is still the top scorer just as shown in while Google 26 just moved one up from the bottom. However, when we sum all their scores across the iterations, we can see a clear difference.

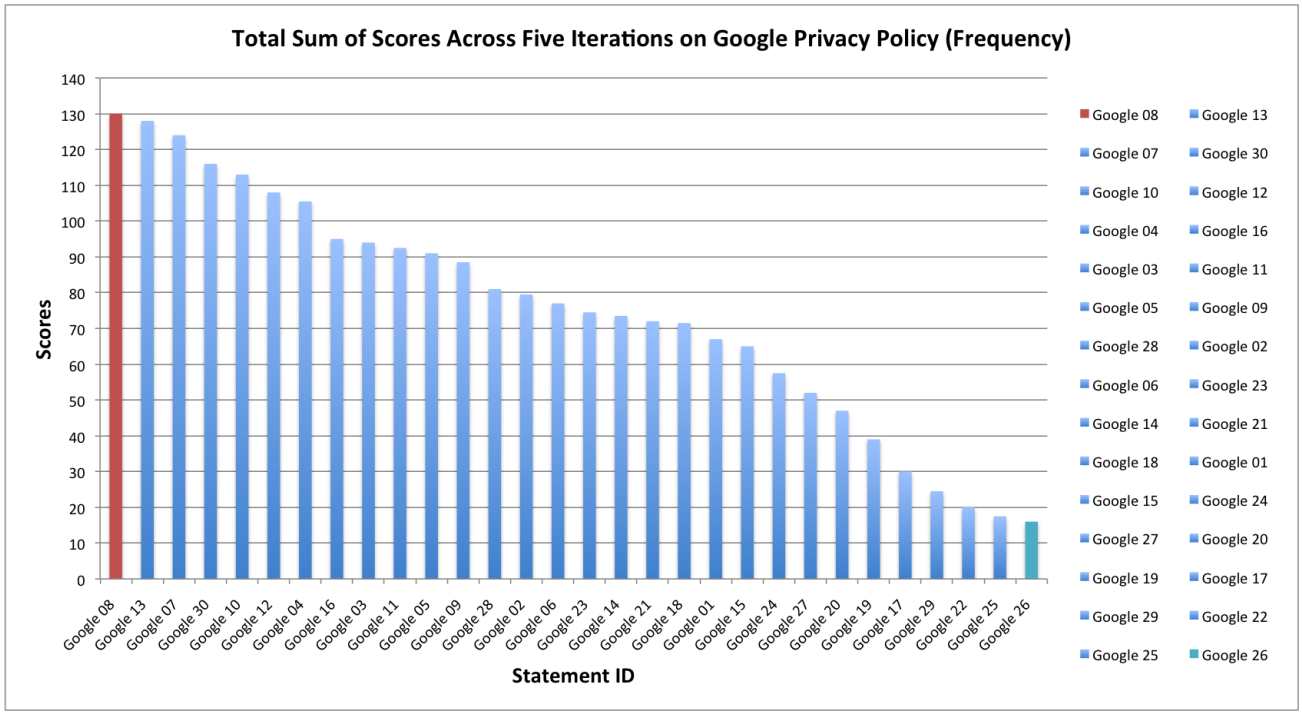


Figure 2.6: The scores of five iterations summed, sorted according to number of scores. Both top and bottom scorer are color coded. (Frequency)

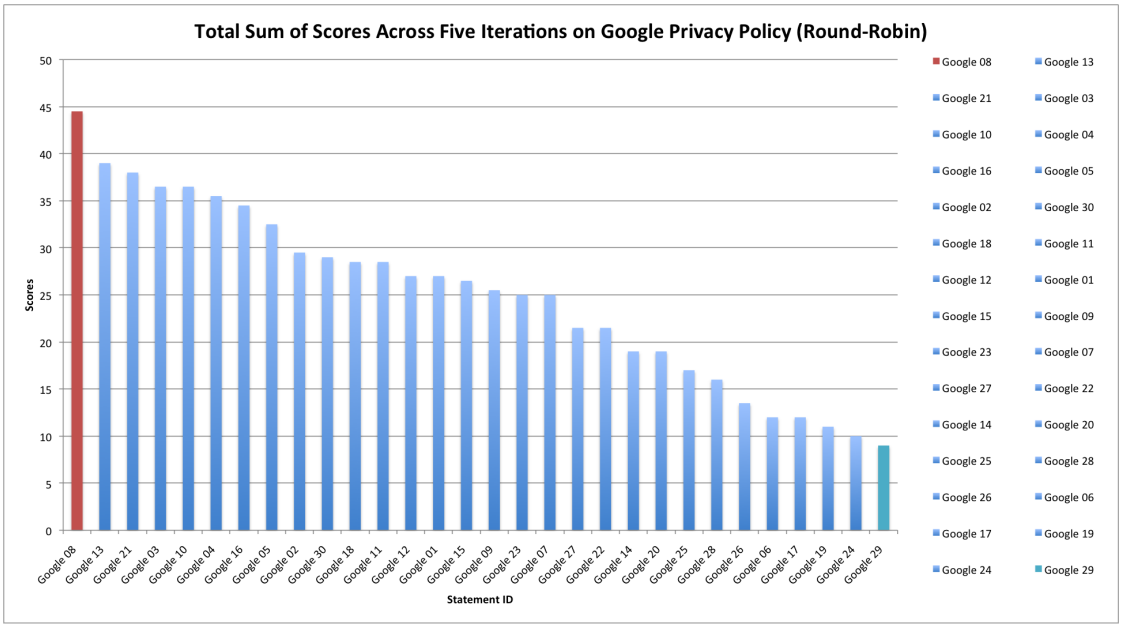


Figure 2.7: The scores of five iterations summed, sorted according to number of scores. Both top and bottom scorer are color coded. (Round-Robin)

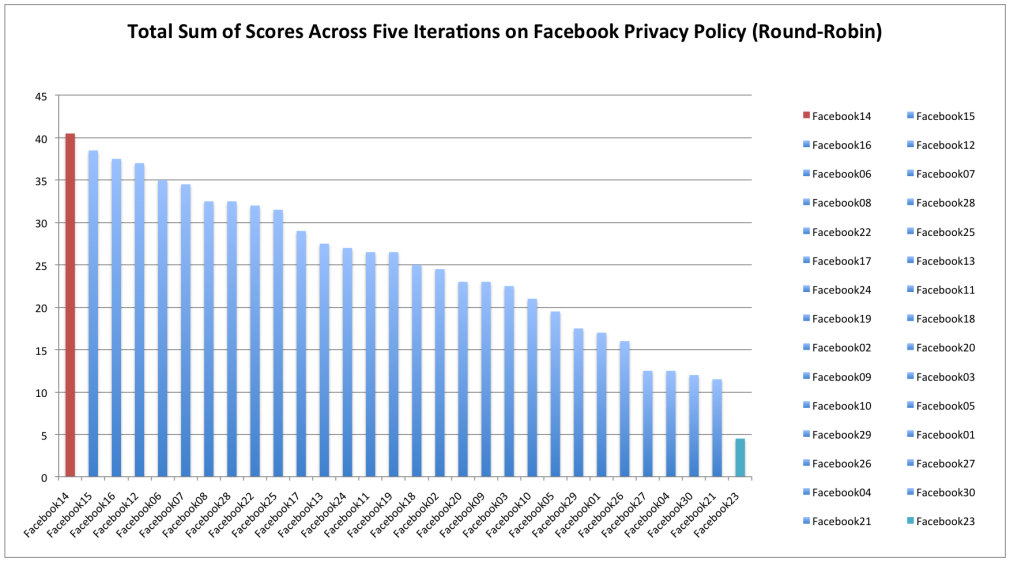


Figure 2.8: The scores of five iterations summed, sorted according to number of scores. Both top and bottom scorer are color coded. (Round-Robin)

Facebook shows similar graph (Figure 2.8), with the only difference is that the distance between the top scorer and second top scorer is smaller but the distance between bottom scorer and second bottom scorer is larger.

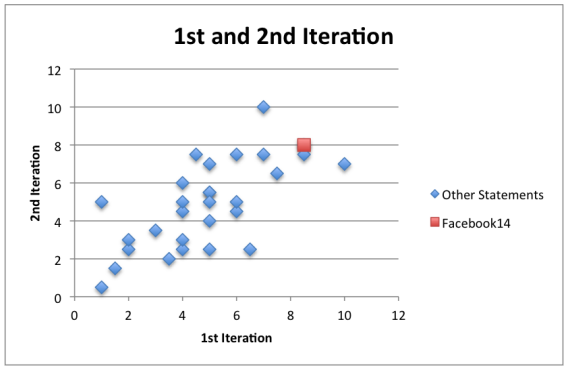


Figure 2.9: A scatterplot of the scores between the 1st and 2nd iteration.

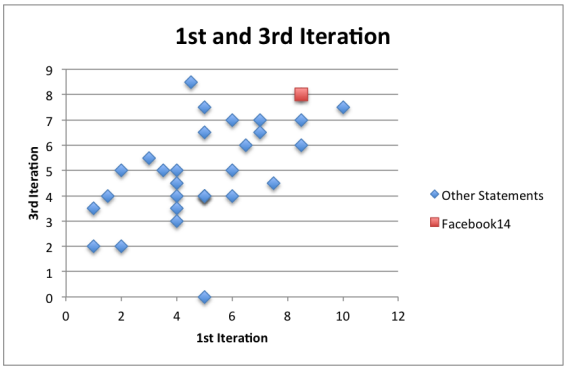


Figure 2.10: A scatterplot of the scores between the 1st and 3rd iteration.

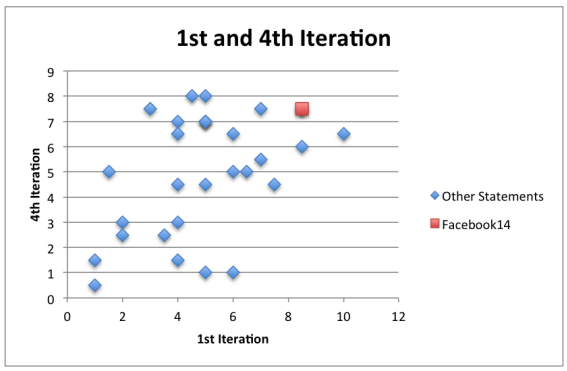


Figure 2.11: A scatterplot of the scores between the 1st and 4th iteration.

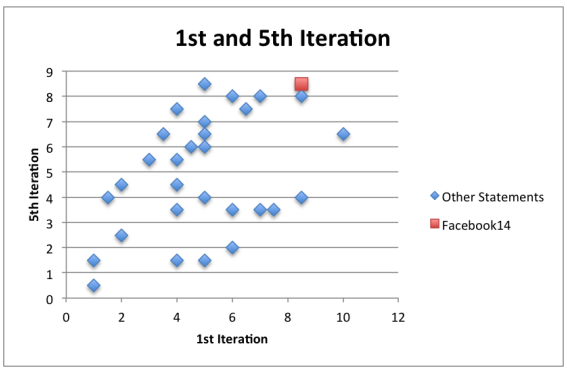


Figure 2.12: A scatterplot of the scores between the 1st and 5th iteration.

On Facebook privacy policy, we plot scatterplot graphs and mark one statement to detect its movements across iterations. Facebook14 moves around the graph, but it stays at top right corner. The graph shape is also showing a trend.

The statements are also paired with other statements most of the time as can be seen in the sample in Table B8.

Discussion

By looking at Figure 2.6, Figure 2.7, and Figure 2.8, it is clear that the round robin method is better at converging the top and bottom result sum-wise. The wide gap of difference shows that the strongest/weakest statement keep winning/losing with other statements, and thus showing us that the strongest/weakest is indeed, strong/weak against many. For more reliable result, we assume that we will need at least fifteen iterations to pit one statement against the other 29 since five iterations pit the statements against up to 10 different statements. We will need to do more experiments to determine how many iterations we need for reliable result.

The movement of statement monitored on the scatterplot shows us that the statement scores changes across iterations. However, the change is not that big or extreme since Facebook14, one of the strong statements keep moving in the same upper right corner. In other words, the method can approximately group the statement according to its strength.

As for the cost, it is much cheaper than frequency experiment even if the number of workers is more than the frequency experiment. There are five per group and we have six groups, so there are thirty workers per iteration in round-robin, while frequency employs fifteen per iteration. The reason is because each workers only need to work on five pairs instead of thirty pairs like in frequency experiment. We spent \$75 for five iterations (B.2), a large contrast against \$225 on previous experiment.

2.3.5 Surprising Results

We have found several surprising data on our experiment results and compiled it here.

Unexpected Most Important Statement

In experimenting using Facebook privacy policy, we expected statement like *"That means that when you visit Facebook-enhanced applications and websites you are making your Facebook information available to someone other than Facebook"* (Facebook16) will rise up to the top three or five. While it does rise to the top three, it is unexpectedly less important than another statement; *"Even after you remove information from your profile or delete your account, copies of that information may remain viewable elsewhere to the extent it has been shared with others, it was otherwise distributed pursuant to your privacy settings, or it was copied or stored by other users"* (Facebook14). We find it surprising because we thought that statement like Facebook14 is a common knowledge, since it basically talks about what has been uploaded to the Internet, stays on the Internet. But it seems that the workers think that it is important to know that Facebook does not guarantee total deletion of information.

Unexpected Score Fluctuation on a Statement

We did not expect statement *"Many of our services let you share information with others. Remember that when you share information publicly, it may be indexable by search engines, including Google"* (Google 28) would get a fluctuation in score. We thought this statement was pretty important, so we expected this score would be stable, but it is not so.

2.3.6 Result Summary & Discussion

We did not manage to get absolute ranking or steady score from our experiments. There are too many factors such as website's fame or nature of service that can affect a statement's importance so it will probably never reach convergence. However, the experiments shows us that this method is usable for determining the strongest/weakest statement and approximately grouping the statements according to their strength. This result is satisfactory, because our goal is to summarize privacy policies, not to determine the absolute ranking of their parts.

3

Conclusion & Future Work

3.1 Conclusion

Privacy policy is an important document that explains the users how the websites handle the information they collect from the users. For this reason, it needs to be complete and comprehensive. However, the privacy policies' completeness turns the document into long pages of explanations which takes time to be read. In an attempt to address this problem, our project proposes a method that can highlight parts of the privacy policies.

We pre-selected statements of privacy policies, paired each statement with another statement from the same policy, and asked people to choose which one is more important to know via Amazon Mechanical Turk platform. With this method, we conducted three experiments using Google privacy policy only and also used Facebook privacy policy on the last experiment to find out if the method works on another privacy policy. Lastly, we graphed the results to see the statements' movements across experiments.

We find that the system is quite robust against errors in the testing phase and the first experiment. The accuracy of the workers does not have statistically significant

effect on the result, but we need more experiments to prove it. When we did the round-robin experiment, we find that simply repeating the method in the frequency experiment is expensive and less efficient in terms of converging the result. We never reach absolute rank order in any experiment, but we observe that the top and bottom statements move up and down in limited area. We also find some winning statements (the most important statements) are the statements we did not expect to win.

Our method has succeeded in generating result that can highlight parts of a privacy policy. The unexpected results of our research also shows that this method can be useful for finding out what people think to be important to know instead of expect from expert's view. This might be an advantage over methods that use our expectations as input such as text analysis.

But this method is not without disadvantages. We do not know if this method is feasible outside the study. The survey participants in this study are paid so they are willing to take their time to read and choose which statement that is more important for them, but there is no guarantee that the real users will do the same. If we use the normal MTurk workers as indicator, we can expect the users taking a short time to read and chose the statements. We will to work more on this method, and here is our suggestions for future work.

3.2 Future Work

We have yet to find out if a statement is deemed important because of the statement's meaning itself or if it is because it sounds important or it is because the website whose privacy policy contains the statement. We also need to know if the length and understandability of the statements affects people's choice.

This method also has the potential to test and predict how impactful a policy statement is. We can insert a fake statement and see how people react to it. Website owners and lawyers may want to use this method to check if the policy has a surprising

element unknown to them. And as time goes by, we can generate enough data to create a prediction model and move from manual to automated process. After this, we can apply this method to other long documents such as Terms of Service, End User License Agreement (EULA), or even online contracts.

Bibliography

[1] *CenterWatch Weekly*, p. 2, 2014. [Online]. Available: http://go.galegroup.com/ps/i.do?id=GALE%7CA234715794&v=2.1&u=cmu_main&it=r&p=ITOF&sw=w&asid=7d42ced170acfe1f77ded16ad3430065. [Accessed 7 Dec. 2014].

[2] E. B. Andrade, V. Kaltcheva, and B. Weitz, "Self-disclosure on the web: the impact of privacy policy, reward, and company reputation," *Advances in Consumer Research*, vol. 29, no. 1, pp. 350–353, 2002.

[3] T. Behrend, D. Sharek, A. Meade, and E. Wiebe, "The viability of crowdsourcing for survey research," *Behavior Research Methods*, vol. 43, no. 3, pp. 800–813, 2011. [Online]. Available: http://dx.doi.org/10.3758/s13428-011-0081-0. [Accessed 6 Dec. 2014].

[4] A. Bingham, "Why crowdsourcing is the next cloud computing | wired," 2013. [Online]. Available: http://www.wired.com/2013/10/why-crowdsourcing-is-the-next-cloud-computing/. [Accessed 4 Dec. 2014].

[5] C. Chaianuchittrakul, "Crowdsourcing privacy policy analysis: Evaluating the comfort, readability and importance of privacy policies," Ph.D. dissertation, Carnegie Mellon University, 2013.

[6] Dictionary.com, "the definition of privacy," 2014. [Online]. Available: http://dictionary.reference.com/browse/privacy. [Accessed 5 Dec. 2014].

[7] C. Doyle, "crowdsourcing," 2013. [Online]. Available: http://www.oxfordreference.com.proxy.library.cmu.edu/view/10.1093/acref/9780199590230.001.0001/acref-9780199590230-e-0473. [Accessed 7 Dec. 2014].

[8] S. Furnell and T. Karweni, "Security implications of electronic commerce: a survey of consumers and businesses," *Internet Research*, vol. 9, no. 5, pp. 372–382, 1999. [Online]. Available: http://dx.doi.org/10.1108/10662249910297778. [Accessed 15 Nov. 2014].

[9] C. Hoffman, "Htg explains: Learn how websites are tracking you online," June 2012. [Online]. Available: http://www.howtogeek.com/115483/htg-explains-learn-how-websites-are-tracking-you-online/. [Accessed 6 Dec. 2014].

[10] P. G. Kelley, "Conducting usable privacy & security studies with amazon's mechanical turk," in *Symposium on Usable Privacy and Security (SOUPS)(Redmond, WA*. Citeseer, 2010.

[11] ——, "Designing privacy notices: Supporting user understanding and control," Ph.D. dissertation, Carnegie Mellon University, 2013.

[12] A. Kittur, E. H. Chi, and B. Suh, "Crowdsourcing user studies with mechanical turk," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '08. New York, NY, USA: ACM, 2008, pp. 453–456. [Online]. Available: http://doi.acm.org/10.1145/1357054.1357127. [Accessed 3 Dec. 2014].

[13] A. M. McDonald and L. F. Cranor, "The cost of reading privacy policies," *ISJLP*, vol. 4, p. 543, 2008.

[14] P. Piazza, "Consumers read, misunderstand privacy policies," *Security Management*, vol. 47, no. 10, p. 40, October 2003. [Online]. Available: http://search.proquest.com/docview/231137707?accountid=9902. [Accessed 28 Nov. 2014].

[15] K.-P. Vu, V. Chambers, F. Garcia, B. Creekmur, J. Sulaitis, D. Nelson, R. Pierce, and R. Proctor, "How users read and comprehend privacy policies," in *Human Interface and the Management of Information. Interacting in Information Environments*, ser. Lecture Notes in Computer Science, M. Smith and G. Salvendy, Eds. Springer Berlin Heidelberg, 2007, vol. 4558, pp. 802–811. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-73354-6_88. [Accessed 28 Nov. 2014].

Appendix A

Input Statements

This appendix contains the input statements we used in our experiments. You can look at the results table and refer to this table to see what is the content of the 'Statement ID' mentioned in the results table. The 'Statement' is the policy statement that is displayed to the workers, while the 'Line Before' and 'Line After' is the supporting statements to help placing the displayed 'Statement' into context.

Statement ID	Line Before	Statement	Line After
Google 01	Information we collect	We collect information to provide better services to all of our users _ from figuring out basic stuff like which language you speak, to more complex things like which ads you_ll find most useful or the people who matter most to you online.	

Google 02	We collect information in two ways: Information you give us. For example, many of our services require you to sign up for a Google Account.	When you do, we_ll ask for personal information, like your name, email address, telephone number or credit card.	If you want to take full advantage of the sharing features we offer, we might also ask you to create a publicly visible Google Profile, which may include your name and photo.
Google 03	Information we get from your use of our services.	We may collect information about the services that you use and how you use them, like when you visit a website that uses our advertising services or you view and interact with our ads and content.	
Google 04	Device information	We may collect device-specific information (such as your hardware model, operating system version, unique device identifiers, and mobile network information including phone number).	Google may associate your device identifiers or phone number with your Google Account.
Google 05	We may collect device-specific information (such as your hardware model, operating system version, unique device identifiers, and mobile network information including phone number).	Google may associate your device identifiers or phone number with your Google Account.	
Google 06	Log information	When you use our services or view content provided by Google, we may automatically collect and store certain information in server logs.	

Google 07		When you use our services or view content provided by Google, we may automatically collect and store certain information in server logs. This may include: details of how you used our service, such as your search queries.	
Google 08		When you use our services or view content provided by Google, we may automatically collect and store certain information in server logs. This may include: telephony log information like your phone number, calling-party number, forwarding numbers, time and date of calls, duration of calls, SMS routing information and types of calls.	
Google 09		When you use our services or view content provided by Google, we may automatically collect and store certain information in server logs. This may include: Internet protocol address.	

Google 10		When you use our services or view content provided by Google, we may automatically collect and store certain information in server logs. This may include: device event information such as crashes, system activity, hardware settings, browser type, browser language, the date and time of your request and referral URL.	
Google 11		When you use our services or view content provided by Google, we may automatically collect and store certain information in server logs. This may include: cookies that may uniquely identify your browser or your Google Account.	
Google 12		When you use our services or view content provided by Google, we may automatically collect and store certain information in server logs. This may include: Location information	
Google 13		When you use a location-enabled Google service, we may collect and process information about your actual location, like GPS signals sent by a mobile device.	We may also use various technologies to determine location, such as sensor data from your device that may, for example, provide information on nearby Wi-Fi access points and cell towers.

Google 14	Unique application numbers	Certain services include a unique application number. This number and information about your installation (for example, the operating system type and application version number) may be sent to Google when you install or uninstall that service or when that service periodically contacts our servers, such as for automatic updates.	
Google 15	Local storage	We may collect and store information (including personal information) locally on your device using mechanisms such as browser web storage (including HTML 5) and application data caches.	
Google 16	Cookies and anonymous identifiers	We and our partners use various technologies to collect and store information when you visit a Google service, and this may include sending one or more cookies or anonymous identifiers to your device.	We also use cookies and anonymous identifiers when you interact with services we offer to our partners, such as advertising services or Google features that may appear on other sites.
Google 17	How we use information we collect	We use the information we collect from all of our services to provide, maintain, protect and improve them, to develop new ones, and to protect Google and our users.	We also use this information to offer you tailored content _ like giving you more relevant search results and ads.

Google 18		We may use the name you provide for your Google Profile across all of the services we offer that require a Google Account.	In addition, we may replace past names associated with your Google Account so that you are represented consistently across all our services. If other users already have your email, or other information that identifies you, we may show them your publicly visible Google Profile information, such as your name and photo.
Google 19		When you contact Google, we may keep a record of your communication to help solve any issues you might be facing.	We may use your email address to inform you about our services, such as letting you know about upcoming changes or improvements.
Google 20		We use information collected from cookies and other technologies, like pixel tags, to improve your user experience and the overall quality of our services.	For example, by saving your language preferences, we_ll be able to have our services appear in the language you prefer. When showing you tailored ads, we will not associate a cookie or anonymous identifier with sensitive categories, such as those based on race, religion, sexual orientation or health.
Google 21		We may combine personal information from one service with information, including personal information, from other Google services _ for example to make it easier to share things with people you know.	We will not combine DoubleClick cookie information with personally identifiable information unless we have your opt-in consent.

Google 22		We will ask for your consent before using information for a purpose other than those that are set out in this Privacy Policy.	
Google 23	Google processes personal information on our servers in many countries around the world.	We may process your personal information on a server located outside the country where you live.	
Google 24		People have different privacy concerns. Our goal is to be clear about what information we collect, so that you can make meaningful choices about how it is used.	
Google 25	People have different privacy concerns. Our goal is to be clear about what information we collect, so that you can make meaningful choices about how it is used.	For example, you can: Control who you share information with.	
Google 26	People have different privacy concerns. Our goal is to be clear about what information we collect, so that you can make meaningful choices about how it is used.	For example, you can: Take information out of many of our services.	
Google 27	People have different privacy concerns. Our goal is to be clear about what information we collect, so that you can make meaningful choices about how it is used.	For example, you can: You may also set your browser to block all cookies, including cookies associated with our services, or to indicate when a cookie is being set by us.	However, it_s important to remember that many of our services may not function properly if your cookies are disabled. For example, we may not remember your language preferences.

Google 28	Information you share	Many of our services let you share information with others. Remember that when you share information publicly, it may be indexable by search engines, including Google.	Our services provide you with different options on sharing and removing your content.
Google 29		Whenever you use our services, we aim to provide you with access to your personal information.	If that information is wrong, we strive to give you ways to update it quickly or to delete it – unless we have to keep that information for legitimate business or legal purposes. When updating your personal information, we may ask you to verify your identity before we can act on your request.
Google 30		We will share personal information with companies, organizations or individuals outside of Google when we have your consent to do so.	We require opt-in consent for the sharing of any sensitive personal information.

Table A1: The Google Privacy Policy Statements used in the experiments.

Statement ID	Line Before	Statement	Line After
Facebook01	If you are under age 13, please do not attempt to register for Facebook or provide any personal information about yourself to us.	If we learn that we have collected personal information from a child under age 13, we will delete that information as quickly as possible.	
Facebook02	One of the primary reasons people use Facebook is to share content with others. Examples include when you update your status, upload or take a photo, upload or record a video, share a link, create an event or a group, make a comment, write something on someone_s Wall, write a note, or send someone a message.	If you do not want us to store metadata associated with content you share on Facebook (such as photos), please remove the metadata before uploading the content.	
Facebook03	We may retain the details of transactions or payments you make on Facebook.	However, we will only keep your payment source account number with your consent.	
Facebook04	We offer contact importer tools to help you upload your friends_ addresses so that you can find your friends on Facebook, and invite your contacts who do not have Facebook accounts to join. If you do not want us to store this information, visit this help page.	If you give us your password to retrieve those contacts, we will not store your password after you have uploaded your contacts_ information.	

Facebook05		When you share your location with others or add a location to something you post, we treat that like any other content you post (for example, it is subject to your privacy settings).	If we offer a service that supports this type of location sharing we will present you with an opt-in choice of whether you want to participate.
Facebook06		We keep track of the actions you take on Facebook, such as adding a friend, becoming a fan of a Facebook Page, joining a group or an event, creating a photo album, sending a gift, poking another user, indicating you _like_ a post, attending an event, or authorizing an application	In some cases you are also taking an action when you provide information or content to us
Facebook07		When you access Facebook from a computer, mobile phone or other device, we may collect information from that device about your browser type, location, and IP address, as well as the pages you visit	
Facebook08	We do not own or operate the applications that you use through Facebook Platform (such as games and utilities) or the websites that you interact with through Facebook Connect. We refer to them as _Facebook-enhanced_ applications and websites because they use our Platform to provide you with social features.	Whenever you authorize a Facebook-enhanced application or website, we will receive information from them, including information about actions you take.	In some cases, in order to personalize the process of connecting, we may receive a limited amount of information even before you authorize the application or website.

Facebook09	We may institute programs with advertising partners and other websites in which they share information with us:	We may ask advertisers to tell us how our users responded to the ads we showed them (and for comparison purposes, how other users who didn_t see the ads acted on their site).	This data sharing, commonly known as _conversion tracking,_ helps us measure our advertising effectiveness and improve the quality of the advertisements you see.
Facebook10	We may institute programs with advertising partners and other websites in which they share information with us: 1) We may ask advertisers to tell us how our users responded to the ads we showed them (and for comparison purposes, how other users who didn_t see the ads acted on their site). This data sharing, commonly known as _conversion tracking,_ helps us measure our advertising effectiveness and improve the quality of the advertisements you see. 2) We may receive information about whether or not you_ve seen or interacted with certain ads on other sites in order to measure the effectiveness of those ads.	If in any of these cases we receive data that we do not already have, we will _anonymize_ it within 180 days, meaning we will stop associating the information with any particular user.	If we institute these programs, we will only use the information in the ways we explain in the _How We Use Your Information_ section below
Facebook11		We may collect information about you from other Facebook users, such as when a friend tags you in a photo or video, provides friend details, or indicates a relationship with you.	You can limit who can see that you have been tagged in a photo or video _ which we refer to as photos or videos _of me_ _ in your privacy settings.

Facebook12	We designed our privacy settings to enable you to control how you share your information on Facebook. You should review the default privacy settings to make sure they reflect your preferences. Here are some specific things to remember:	Certain categories of information such as your name, profile photo, list of friends and pages you are a fan of, gender, and networks you belong to are considered publicly available, and therefore do not have privacy settings. (We will soon stop using regional networks, but your geographic region will still be considered publicly available).	You can limit the ability of others to find this information on third party search engines through your search privacy settings.
Facebook13	We designed our privacy settings to enable you to control how you share your information on Facebook. You should review the default privacy settings to make sure they reflect your preferences. Here are some specific things to remember:	Some of the content you share and the actions you take will show up on your friends_ home pages and other pages they visit.	
Facebook14	We designed our privacy settings to enable you to control how you share your information on Facebook. You should review the default privacy settings to make sure they reflect your preferences. Here are some specific things to remember:	Even after you remove information from your profile or delete your account, copies of that information may remain viewable elsewhere to the extent it has been shared with others, it was otherwise distributed pursuant to your privacy settings, or it was copied or stored by other users.	

Facebook15		Information set to _everyone_ is publicly available information, may be accessed by everyone on the Internet (including people not logged into Facebook), is subject to indexing by third party search engines, may be associated with you outside of Facebook (such as when you visit other sites on the internet), and may be imported and exported by us and others without privacy limitations.	The default privacy setting for certain types of information you post on Facebook is set to _everyone._ You can review and change the default settings in your privacy settings. If you delete _everyone_ content that you posted on Facebook, we will remove it from your Facebook profile, but have no control over its use outside of Facebook.
Facebook16	As mentioned above, we do not own or operate Facebook-enhanced applications or websites.	That means that when you visit Facebook-enhanced applications and websites you are making your Facebook information available to someone other than Facebook.	To help those applications and sites operate, they receive publicly available information automatically when you visit them, and additional information when you formally authorize or connect your Facebook account with them.
Facebook17		We may make information about the location of your computer or access device and your age available to Facebook _enhanced applications and websites in order to help them implement appropriate security measures and control the distribution of age-appropriate content.	

Facebook18		We don_t share your information with advertisers without your consent.	(An example of consent would be if you asked us to provide your shipping address to an advertiser to receive a free sample.) We allow advertisers to choose the characteristics of users who will see their advertisements and we may use any of the non-personally identifiable attributes we have collected (including information you may have decided not to show to other users, such as your birth year or other sensitive personal information or preferences) to select the appropriate audience for those advertisements.
Facebook19		We occasionally pair advertisements we serve with relevant information we have about you and your friends to make advertisements more interesting and more tailored to you and your friends.	For example, if you become a fan of a Page, we may display your name and profile photo next to an advertisement for that Page that is displayed to your friends. We only share the personally identifiable information visible in the social ad with the friend who can see the ad. You can opt out of having your information used in social ads on this help page.

Facebook20		We use your profile information, the addresses you import through our contact importers, and other relevant information, to help you connect with your friends, including making suggestions to you and other users that you connect with on Facebook.	If you want to limit your visibility in suggestions we make to other people, you can adjust your search visibility privacy setting, as you will only be visible in our suggestions to the extent you choose to be visible in public search listings. You may also block specific individual users from being suggested to you and you from being suggested to them.
Facebook21		If we are notified that a user is deceased, we may memorialize the user_s account.	In such cases we restrict profile access to confirmed friends, and allow friends and family to write on the user_s Wall in remembrance. We may close an account if we receive a formal request from the user_s next of kin or other proper legal request to do so.
Facebook22		When you enter into transactions with others or make payments on Facebook, we will only share transaction information with those third parties necessary to complete the transaction and will require those third parties to agree to respect the privacy of your information.	

Facebook23		We generally limit search engines_ access to our site.	We may allow them to access information set to the _everyone_ setting and your public search listing (but you can turn off your public search listing in your privacy settings).
Facebook24		We may disclose information pursuant to subpoenas, court orders, or other requests (including criminal and civil matters) if we have a good faith belief that the response is required by law.	This may include respecting requests from jurisdictions outside of the United States where we have a good faith belief that the response is required by law under the local laws in that jurisdiction, apply to users from that jurisdiction, and are consistent with generally accepted international standards. We may also share information when we have a good faith belief it is necessary to prevent fraud or other illegal activity, to prevent imminent bodily harm, or to protect ourselves and you from people violating our Statement of Rights and Responsibilities. This may include sharing information with other companies, lawyers, courts or other government entities.

Facebook25	Facebook Beacon. [We have announced a settlement of a lawsuit related to the Beacon product: the Beacon product will be discontinued and this language removed from the privacy policy upon approval of a settlement by the court.] Facebook Beacon is a means of sharing actions you have taken on third party sites, such as when you make a purchase or post a review, with your friends on Facebook.	In order to provide you as a Facebook user with clear disclosure of the activity information being collected on third party sites and potentially shared with your friends on Facebook, we collect certain information from that site and present it to you after you have completed an action on that site.	You have the choice to have us discard that information, or to share it with your friends. To learn more about the operation of the service, we encourage you to read the tutorial here. To opt out of the service altogether, click here. Like many other websites that interact with third party sites, we may receive some information even if you are logged out from Facebook, or that pertains to non-Facebook users, from those sites in conjunction with the technical operation of the system. In cases where we receive information from Beacon sites on users that are not logged in, or on non-Facebook users, we do not attempt to associate it with individual Facebook accounts and will discard it.
Facebook26		If the ownership of all or substantially all of our business changes, we may transfer your information to the new owner so that the service can continue to operate.	In such a case, your information would remain subject to the promises made in any pre-existing Privacy Policy.

Facebook27	Viewing and editing your profile. You may change or delete your profile information at any time by going to your profile page and clicking _Edit My Profile._ Information will be updated immediately.	While you cannot delete your date of birth, you can use the setting on the info tab of your profile information page to hide all or part of it from other users.	
Facebook28		Even after you remove information from your profile or delete your account, copies of that information may remain viewable elsewhere to the extent it has been shared with others, it was otherwise distributed pursuant to your privacy settings, or it was copied or stored by other users.	However, your name will no longer be associated with that information on Facebook. (For example, if you post something to another user_s profile, and then you delete your account, that post may remain, but be attributed to an _Anonymous Facebook User._) Additionally, we may retain certain information to prevent identity theft and other misconduct even if deletion has been requested.
Facebook29		Removed and deleted information may persist in backup copies for up to 90 days, but will not be available to others.	
Facebook30		If a user provides your email address to us, and you are not a Facebook user but you want us to delete your address, you can do so on this help page.	However, that request will only apply to addresses we have at the time of the request and not to any addresses that users provide to us later.

Table A2: The Facebook Privacy Policy Statements used in the experiments.

Appendix B

Tables and Calculations

Statement ID	Normal Workers	Master Workers
Google 01	21	21
Google 02	15	18.5
Google 03	6	14.5
Google 04	25.5	24.5
Google 05	12.5	19
Google 06	15.5	14.5
Google 07	27	26.5
Google 08	23	24
Google 09	18.5	16.5
Google 10	19.5	21.5
Google 11	14.5	13
Google 12	14.5	15
Google 13	24.5	26.5
Google 14	12.5	14.5
Google 15	19.5	14
Google 16	18	16.5
Google 17	8	5.5
Google 18	15.5	14.5
Google 19	5.5	3.5
Google 20	18	14.5
Google 21	19.5	12.5
Google 22	3.5	8.5
Google 23	12	9.5
Google 24	14	12
Google 25	5	4.5
Google 26	3	1
Google 27	15	15
Google 28	17	16
Google 29	3	6
Google 30	24	27

Table B1: Normal and master workers's result for the same survey.

	Normal Workers	Master Workers
Mean	15	15
Variance	47.74137931	48.12068966
Observations	30	30
Pearson Correlation	0.887776734	
Hypothesized Mean Difference	0	
df	29	
t Stat	0	
P(T<=t) one-tail	0.5	
t Critical one-tail	1.699127027	
P(T<=t) two-tail	1	
t Critical two-tail	2.045229642	

Table B2: The paired t-Test between the normal workers' iteration and master workers' iteration. The result is not significant at $p < 0.05$.

Statement ID	1st Iteration	2nd Iteration	3rd Iteration	4th Iteration	5th Iteration
Google 01	15.5	7.5	14.5	21	8.5
Google 02	12.5	17	15	18.5	16.5
Google 03	5.5	2	1	14.5	24
Google 04	20	20	14	24.5	16
Google 05	17	15.5	21.5	19	18
Google 06	14.5	17.5	13.5	14.5	17
Google 07	29.5	27.5	18.5	26.5	22
Google 08	29	28.5	23.5	24	25
Google 09	19.5	15	22.5	16.5	15
Google 10	21	22.5	20	21.5	22
Google 11	11.5	22.5	25.5	13	20
Google 12	18	27.5	24.5	15	23
Google 13	28	26.5	23	26.5	24
Google 14	14.5	12	18	14.5	14.5
Google 15	12.5	11	15	14	12.5
Google 16	17.5	22.5	18	16.5	20.5
Google 17	4.5	8	10.5	5.5	6.5
Google 18	18.5	10.5	16.5	14.5	11.5
Google 19	2	11.5	8.5	3.5	13.5
Google 20	14.5	5.5	8.5	14.5	9
Google 21	16	13	16	12.5	14.5
Google 22	3.5	8.5	4	8.5	3.5
Google 23	14	18	13.5	9.5	19.5
Google 24	9	15	6.5	12	15
Google 25	5.5	3	1.5	4.5	3
Google 26	3.5	4.5	2	1	5
Google 27	15	4	9.5	15	8.5
Google 28	16	18	13	16	18
Google 29	7.5	4	7.5	6	6.5
Google 30	25.5	17.5	28.5	27	17.5

Table B3: The result of the frequency experiment with master workers, sorted according to the statement ID. The red column is the ones with a broken statement (blank choice box) and the cyan column is the normal one without a broken statement.

Statement ID	1st Iteration	2nd Iteration	3rd Iteration	4th Iteration	5th Iteration	SUM	STD.DEV (1-3)	STD.DEV (4-5)
Google 08	29	28.5	23.5	24	25	130	3.041381265	0.707106781
Google 13	28	26.5	23	26.5	24	128	2.56580072	1.767766953
Google 07	29.5	27.5	18.5	26.5	22	124	5.859465277	3.181980515
Google 30	25.5	17.5	28.5	27	17.5	116	5.686240703	6.717514421
Google 10	21	22.5	26	21.5	22	113	2.56580072	0.353553391
Google 12	18	27.5	24.5	15	23	108	4.856267428	5.656854249
Google 04	22	29	14	24.5	16	105.5	7.505553499	6.01040764
Google 16	17.5	22.5	18	16.5	20.5	95	2.753785274	2.828427125
Google 03	17.5	20	18	14.5	24	94	1.322875656	6.717514421
Google 11	11.5	22.5	25.5	13	20	92.5	7.371114796	4.949747468
Google 05	17	15.5	21.5	19	18	91	3.122498999	0.707106781
Google 09	19.5	15	22.5	16.5	15	88.5	3.774917218	1.060660172
Google 28	16	18	13	16	18	81	2.516611478	1.414213562
Google 02	12.5	17	15	18.5	16.5	79.5	2.254624876	1.414213562
Google 06	14.5	17.5	13.5	14.5	17	77	2.081665999	1.767766953
Google 23	14	18	13.5	9.5	19.5	74.5	2.466441431	7.071067812
Google 14	14.5	12	18	14.5	14.5	73.5	3.013856887	0
Google 21	16	13	16	12.5	14.5	72	1.732050808	1.414213562
Google 18	18.5	10.5	16.5	14.5	11.5	71.5	4.163331999	2.121320344
Google 01	15.5	7.5	14.5	21	8.5	67	4.358898944	8.838834765
Google 15	12.5	11	15	14	12.5	65	2.020725942	1.060660172
Google 24	9	15	6.5	12	15	57.5	4.368447474	2.121320344
Google 27	15	4	9.5	15	8.5	52	5.5	4.596194078
Google 20	9.5	5.5	8.5	14.5	9	47	2.081665999	3.889087297
Google 19	2	11.5	8.5	3.5	13.5	39	4.856267428	7.071067812
Google 17	4.5	3	10.5	5.5	6.5	30	3.968626967	0.707106781
Google 29	7.5	4	0.5	6	6.5	24.5	3.5	0.353553391
Google 22	3.5	0.5	4	8.5	3.5	20	1.892969449	3.535533906
Google 25	5.5	3	1.5	4.5	3	17.5	2.020725942	1.060660172
Google 26	3.5	4.5	2	1	5	16	1.258305739	2.828427125

Table B4: The result of the frequency experiment with master workers, sorted according to the Sum, along with the standard deviation calculation. The red column is the ones with a broken statement (blank choice box) and the cyan column is the normal one without a broken statement.

	Broken	Normal
Mean	3.36998952	3.064129385
Variance	2.245808705	6.227011494
Observations	30	30
Pearson Correlation	0.494725521	
Hypothesized Mean Difference	0	
df	29	
t Stat	0.766837342	
P(T<=t) one-tail	0.224688142	
t Critical one-tail	1.699127027	
P(T<=t) two-tail	0.449376284	
t Critical two-tail	2.045229642	

Table B5: The paired t-Test between the broken iterations' standard deviation and normal iterations' standard deviation. The result is not significant at $p < 0.05$.

Frequency experiment cost calculation:

$$\begin{aligned} Cost &= 5 \times 1 \times 30 \times 15 \times \$0.1 \\ &= \$225 \end{aligned} \tag{B.1}$$

Statement ID	1st Iteration	2nd Iteration	3rd Iteration	4th Iteration	5th Iteration	Sum
Google 08	9.5	10	8.5	9	8.5	45.5
Google 21	8.5	6	9	7.5	7	38
Google 13	7	9	7	7.5	7.5	38
Google 04	8	8	6.5	7	5.5	35
Google 16	6.5	8.5	8	6	6	35
Google 03	8	7.5	6.5	8	5	35
Google 10	9.5	6	6.5	6	4.5	32.5
Google 05	4.5	7.5	5.5	8.5	4.5	30.5
Google 12	6.5	6.5	5.5	4	7	29.5
Google 02	7.5	3	6.5	5	7	29
Google 30	5.5	6.5	6	6	4.5	28.5
Google 18	7.5	7	5.5	4.5	4	28.5
Google 01	8.5	3	6	3.5	7.5	28.5
Google 09	5	7	6.5	4	5.5	28
Google 11	3.5	5	5.5	7	5.5	26.5
Google 23	5	5.5	4.5	5	5	25
Google 15	4.5	3.5	6	6	5	25
Google 07	4	5	4.5	3.5	7.5	24.5
Google 27	3	5	6	4.5	5.5	24
Google 14	3.5	3.5	4.5	4.5	6	22
Google 22	6	4	3	4.5	3.5	21
Google 28	1.5	3	3.5	6.5	2.5	17
Google 20	4	4.5	3	2.5	3	17
Google 06	3	2	4	2	5	16
Google 25	2.5	1	3.5	6	3	16
Google 19	1.5	4	2	2	4	13.5
Google 17	3	3.5	1	2.5	2.5	12.5
Google 29	2	0	1.5	3	5	11.5
Google 26	1	2	2.5	3	2	10.5
Google 24	0	3	1.5	1	1	6.5

Table B6: The result of round-robin experiment on Google Privacy Policy, ordered by the highest sum of score.

Statement ID	1st Iteration	2nd Iteration	3rd Iteration	4th Iteration	5th Iteration	Sum
Facebook14	8.5	8	8	7.5	8.5	40.5
Facebook15	8.5	7.5	7	7.5	8	38.5
Facebook16	10	7	7.5	6.5	6.5	37.5
Facebook12	7	7.5	7	7.5	8	37
Facebook06	6	7.5	7	6.5	8	35
Facebook07	4.5	7.5	8.5	8	6	34.5
Facebook08	7	10	6.5	5.5	3.5	32.5
Facebook28	8.5	8	6	6	4	32.5
Facebook22	5	5	6.5	7	8.5	32
Facebook25	5	5.5	7.5	7	6.5	31.5
Facebook17	5	7	4	7	6	29
Facebook13	6.5	2.5	6	5	7.5	27.5
Facebook24	4	5	3.5	7	7.5	27
Facebook11	5	5.5	4	8	4	26.5
Facebook19	7.5	6.5	4.5	4.5	3.5	26.5
Facebook18	3	3.5	5.5	7.5	5.5	25
Facebook02	6	5	5	5	3.5	24.5
Facebook20	5	2.5	4	4.5	7	23
Facebook09	4	4.5	4.5	4.5	5.5	23
Facebook03	4	6	5	3	4.5	22.5
Facebook10	4	3	4	6.5	3.5	21
Facebook05	3.5	2	5	2.5	6.5	19.5
Facebook29	6	4.5	4	1	2	17.5
Facebook01	2	3	5	2.5	4.5	17
Facebook26	1.5	1.5	4	5	4	16
Facebook27	4	2.5	3	1.5	1.5	12.5
Facebook04	1	5	3.5	1.5	1.5	12.5
Facebook30	2	2.5	2	3	2.5	12
Facebook21	5	4	0	1	1.5	11.5
Facebook23	1	0.5	2	0.5	0.5	4.5

Table B7: The result of round-robin experiment on Facebook Privacy Policy, ordered by the highest sum of score.

Google 02 paired with	Google 30 Google 08	Google 25 Google 05	Google 26 Google 07	Google 27 Google 03	Google 05 Google 28
on xth Iterations	1st	2nd	3rd	4th	5th

Table B8: Sample of pairings across iterations on round-robin experiment. Notice that Google 02 pairs with other statements each time except for 2nd and 5th iteration.

Round Robin Experiment Cost Calculation:

$$\begin{aligned} Cost &= 5 \times 6 \times 5 \times 5 \times \$0.1 \\ &= \$75 \end{aligned} \tag{B.2}$$