The Usable Privacy Policy Project:

Combining Crowdsourcing, Machine Learning and Natural Language Processing to Semi-Automatically Answer Those Privacy Questions Users Care About

Norman Sadeh, Alessandro Acquisti, Travis D. Breaux, Lorrie Faith Cranor, Aleecia M. McDonald[a], Joel R. Reidenberg[b], Noah A. Smith, Fei Liu, N. Cameron Russell[b], Florian Schaub, Shomir Wilson

December 2013
CMU-ISR-13-119

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

[a]The Center for Internet and Society, Stanford Law School, Stanford University, Stanford, CA 94305

[b]Center on Law and Information Policy, School of Law, Fordham University, New York, NY 10023

Natural language privacy policies have become a de facto standard to address expectations of "notice and choice" on the Web. However, users generally do not read these policies and those who do read them struggle to understand their content. Initiatives aimed at addressing this problem through the development of machine-readable standards have run into obstacles, with many website operators showing reluctance to commit to anything more than what they currently do. This project builds on recent advances in natural language processing, privacy preference modeling, crowdsourcing, formal methods, and privacy interface design to develop a practical framework based on websites' existing natural language privacy policy that empowers users to more meaningfully control their privacy, without requiring additional cooperation from website operators. Our approach combines fundamental research with the development of scalable technologies to (1) semi-automatically extract key privacy policy features from natural language privacy policies, and (2) present these features to users in an easy-to-digest format that enables them to make more informed privacy decisions as they interact with different websites. This work will also involve the systematic collection and analysis of website privacy policies, looking for trends and deficiencies both in the wording and content of these policies across different sectors and using this analysis to inform public policy. This report outlines the project's research agenda and overall approach.

This research was supported in part by the National Science Foundation under grant CNS-1330596.

--- Keywords: Privacy, online privacy, usability, law, public policy, behavioral economics, natural language processing, privacy policies, privacy notices, notice & choice, privacy policy analysis, privacy decision making, crowdsourcing, machine learning, privacy preferences, privacy preference modeling, cognitive biases.

1. Introduction

Natural language privacy policies have become a de facto standard to address expectations of "notice and choice" on the Web [FTC10]. Yet, there is ample evidence that users generally do not read these policies and that those who occasionally do struggle to understand what those policies contain. Studies have shown that, if users were to read the privacy policies of every website they access during the course of a year, they would end up spending a substantial amount of their time doing just that and would often still not be able to answer basic questions about what these policies really say [MC08, MRK+2009]. This situation can only be expected to get worse as we interact with a growing number of online services and access these services from devices such as smartphones, which are even less conducive to reading privacy policies [KCC+12].

This challenge was first recognized in the mid nineties. Over the years, it has prompted the launch of initiatives aimed at codifying privacy practices in an attempt to reduce user burden and make it more practical for users to gain relevant information about and sufficient control over data collection and usage practices of sites they visit. This included the development of the Platform for Privacy Preferences (P3P) standard between 1996 and 2003, as well as more recent initiatives such as "Do Not Track" (DNT) [CDE+06, BHN+01]. These initiatives have made significant progress in helping codify data collection and usage policies and in initiating dialogues between relevant stakeholders, but they have also run into significant obstacles [McD13]. While the vast majority of prominent website operators have natural language privacy policies (some required by legal regulation, e.g. [CA03]), many of these entities have shown reluctance in adopting standards and machine-implementable solutions that would force them to further clarify their privacy practices and/or commit to more stringent practices.

Recent advances in natural language processing, privacy preference modeling, crowdsourcing, and privacy interfaces suggest that it may be possible to overcome these challenges and develop practical solutions that rely on existing natural language privacy policies rather than imposing new requirements on website operators. Our project aims to do just that through a combination of fundamental research along with the development and large-scale deployment of novel tools that build on recent advances in these areas.

Specifically our goal is to develop, evaluate and deploy new technologies in the context of a novel, practical framework that empowers users to more meaningfully control their privacy without any additional cooperation from website operators other than the natural language privacy policies that they already have in place. Our work is organized around several mutually supporting strands:

- **Semi-Automated Understanding of Privacy Policies:** Building on recent advances in natural language processing (NLP), machine learning, and crowdsourcing, we are developing algorithms that automatically interpret privacy policies along dimensions of greatest concern to users or policymakers [AWS+12]. Our approach combines linguistic representations provided by natural language semantic analyzers [DCM+13] with statistical learning algorithms that generalize from human-labeled examples obtained from crowdworkers – initially from experts and over time from a broader population of crowdworkers – with the help of basic text analysis [SOJN08]. We identify suitable representations into which privacy policies can be robustly parsed, and that also support automated inference about a policy's content. Importantly, we consider that to support the aim of usable privacy disclosure, a complete meaning representation of a policy's contents is not required. Our work here brings together two complementary strands: formal modeling of policies, laws, and regulations and data-driven natural language processing. Our goal is to develop a "semi-automated solution," where humans provide annotations for machine learning and also help supplement natural language processing techniques when these techniques are not sufficient. As part of our research, we explore different divisions of labor between human experts, crowdworkers, and algorithms and see how they can best help extract different privacy policy features.
- **Privacy Preference Modeling for Usable Privacy Disclosures:** Recent research by the authors and their collaborators has shown that, while people's privacy preferences are complex and diverse, their privacy decisions are often driven by a somewhat limited number of aspects (e.g. [KBC+09, BKS+11]). As part of previous work in this area, Kelley and Cranor have developed and experimented with privacy nutrition labels intended to simplify the presentation of privacy policies by focusing on those elements most important to users

and relying on standardized presentation formats that are both succinct and easy to interpret [KBC+09]. Sadeh's recent work in location privacy and mobile app privacy has shown that it is possible to quantify the benefits of exposing different privacy settings to users [BKS+11] and use crowdsourcing to identify particularly salient privacy policy features [LAH+12]. As part of research in this area, we are exploring ways of simplifying and dynamically adjusting the presentation of privacy policies to users, including the use of machine learning to identify privacy profiles that can be used to tailor presentation of privacy notices to different groups of individuals [RBK+09, MSS11,LLS14]. This will likely also include experimentation with solutions that represent policy elements as a small number of tags or icons, customizing short-form notices to highlight conflicts with a user's privacy preferences, or highlighting changes in a site's policy since a user's last visit to that site, and exploring drill-down formats that allow users to request more details when they want to.

- **Mitigating Deleterious Cognitive and Behavioral Biases in Privacy Disclosures:** As we develop models of user privacy preferences and experiment with different privacy interfaces, we extend our investigation to encompass the study of cognitive and behavioral biases known to often influence people's privacy decisions, building on recent work by Acquisti, Cranor, Sadeh and others [AG07, Acq09, BAL12, JAL11,AJL12, BLA+11,SCK+13]. Results from this research will inform the refinement of our models and interfaces with the objective of countering deleterious effects associated with these biases. In particular we will investigate the gap between user *ex ante* and *ex post* privacy preferences and decisions. By "ex ante," we refer to the user's state before a certain privacy sensitive event occurs or a certain privacy behavior is exhibited. By "ex post," we refer to the user's state following that outcome or behavior. For instance, individuals may claim to have certain privacy preferences *ex ante* ("I want to share my information with search engines in return for personalized search results"), but may actually realize *ex post* that they regret their decision ("I wish I had not opted for personalized search results, as I now realize this includes sharing sensitive medical information with search engines"). Rather than limiting ourselves to the design of privacy disclosure interfaces based on ex ante preferences, we wish instead to develop interfaces that will nudge users to make decisions consistent with their expected ex post preferences, thereby reducing the chances of future regret. This perspective will include experimenting with different levels of granularity at which users can configure their privacy preferences, the impact of default profiles and the deleterious effects associated with giving users too much control or information. This will include experimenting with just-in-time disclosures, where rather than disclosing large and difficult to understand privacy policies in "one shot", relevant policy features are incrementally disclosed to users based on their particular interactions with a website. For instance, data practices relating to contact information would only need to be shown when a users provides contact information to the website.
- **Privacy Policy Analysis:** A challenge for privacy policy authors is to find an appropriate balance between retaining business flexibility, while providing sufficient details to users; this tension can lead to inconsistent policies. Breaux, et al. have shown that semi-formal [BA05, BVA06] and formal models [BAD09] can be applied to natural language policies to analyze policies for ambiguities and conflicts [Bre09]. Based on this work, we plan to explore ways to use automated reasoning over privacy policy models to help website operators debug their policies and to help website users reliably answer questions about their personal information. To this end, we aim to better understand how to semi-automate formalization in combination with advances in NLP and how to design user interfaces that human analysts can use to consistently and reliably formalize policy. This formalization aims to answer the most relevant questions for users and website operators within reasonable computational limits for modern day theorem provers. Overall we expect our project to yield a body of annotated policy corpora, privacy summaries for user displays and formalized policies that we will use to build on our prior work and develop new statistical analysis to discern trends in privacy practices as they evolve across business sectors over time. This can include which types of policies adhere to the fair information principles or various consent mechanisms and how these constructs interact to meet evolving user privacy preferences. Statistically significant policy trends based on reliably coded data sets can be used to further inform policy makers about public policy implications in the United States and Europe. By studying correlations between trends and different regulatory regimes, we plan to inform public policy makers about the extent to which various regulatory models are likely to impact privacy practices as observed in privacy policies.

These research strands are intended to build on each other, with progress in one area contributing to research in another. For instance, work on privacy preference modeling is expected to help identify key policy features to be

extracted through semi-automated analysis of natural language policies. Our hope is that results from this work will in turn inform the design, evaluation and refinement of privacy disclosure interfaces, contribute to the study of cognitive and behavioral biases and the development and evaluation of solutions to mitigate these biases. Privacy policy models built using semi-automated natural language processing can provide a basis for the analysis of policies (e.g. in different sectors), including their evolution over time. In turn we hope that this research will help inform relevant public policy discussions. It could also help refine disclosure interfaces (e.g. highlighting particularly unusual policy features) or natural language processing techniques (e.g. in response to changes in the way in which policies are written, or to help clarify partial NLP results).

Our work combines multiple methodologies. This includes the use of highly iterative user-centered design principles in the development and refinement of privacy displays as well as crowdsourcing mechanisms and interfaces. Machine learning and natural language processing techniques are playing a key role in developing and validating semi-automated solutions to extract salient privacy policy fragments and features. Crowdsourcing and machine learning techniques are also contributing to building and refining models of user privacy preferences, with these models informing the identification of salient policy features as well as the design and personalization of disclosure interfaces. Methodologies from behavioral economics will be used as part of our work to better understand and mitigate biases, while statistics and formal methods will help analyze models of privacy policies.

2. Overall Approach

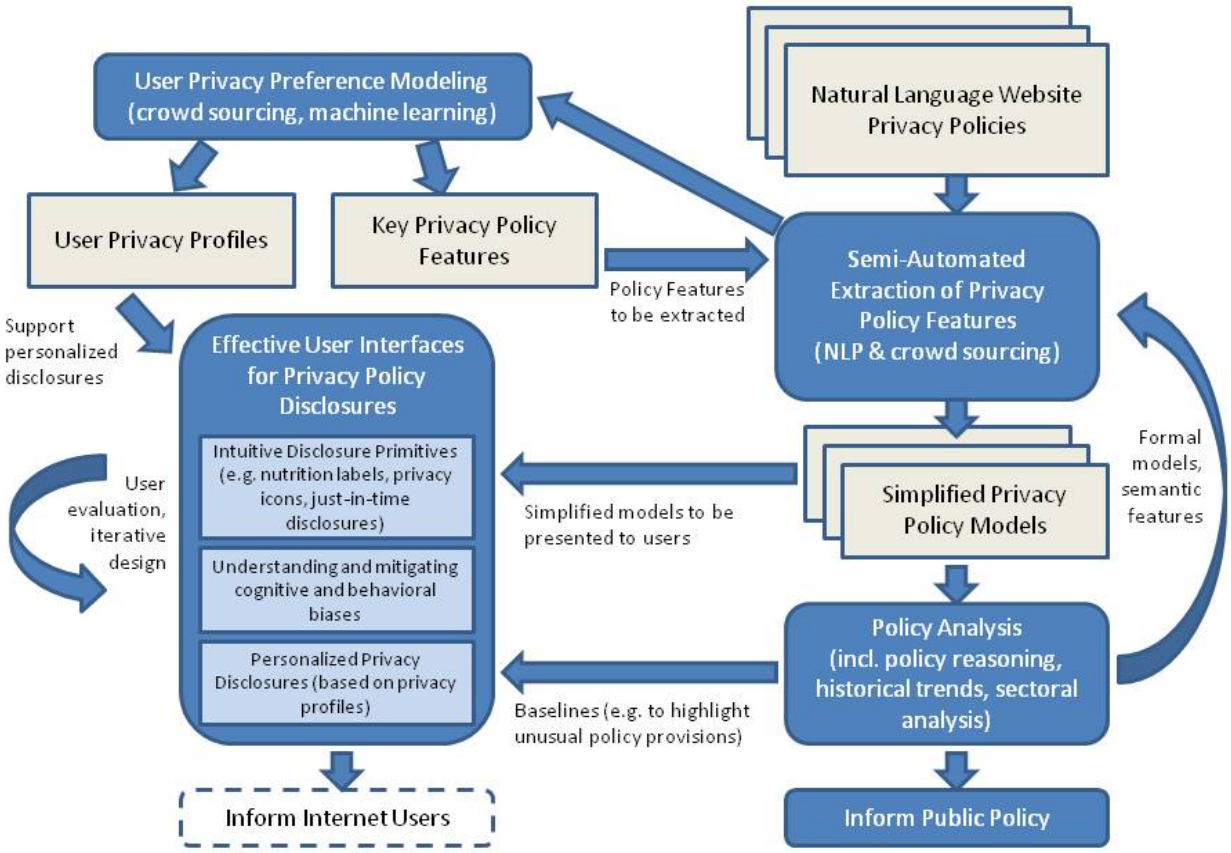

Figure 1. Overall Approach. Darker boxes identify the project's main areas of research and highlight key elements of functionality and/or results each of these areas is expected to produce.

Our current work includes the development of an infrastructure for collecting and interpreting key privacy features and for developing models of people's privacy preferences to help identify those policy features to be extracted and presented to users. Figure 1 outlines those research areas on which this project focuses.

This includes the development and evaluation of semi-automated functionality to process natural language website privacy policies and extract relevant policy features. This research combines and extends recent advances in natural language processing along with the development of a crowdsourcing framework intended to complement natural language processing. This work itself is driven in part by concurrent research aimed at developing models of user privacy preferences. These models aim to identify those policy features (*"Key Privacy Policy Features"* in Figure 1) that are most critical to informing people's decisions when it comes to interacting with different websites. The identification of these features, including the amount of detail people care about (e.g. knowing whether a website shares one's shopping preferences with other entities versus knowing with which types of entities these preferences might be shared), is driving work on semi-automated natural language processing.

Our research in user privacy preference modeling is also exploring differences between categories of websites. Different categories of sites have different kinds of data requirements (e.g. a pure marketing site versus an e-tailing site), which in turn can be expected to entail different types of user preferences (e.g. Amazon needs to collect one's street address to ship items purchased at the site, but users are less likely to feel comfortable sharing this same information with a site that does not need it). Also, while not all users feel the same way when it comes to disclosing a particular piece of data to a particular website, prior research suggests that it might be possible to organize them in a small number of clusters of like-minded users (e.g. see [RBK+09, MSS11,LLS14] in the case of location privacy preferences and mobile app privacy preferences). As part of our work in *User Privacy Preference Modeling*, we are studying to what extent it is possible to identify such categories and associated *User Privacy Profiles*, and use this information to personalize and/or simplify privacy disclosures. This would mean showing (or highlighting) different *key privacy policy features* to different groups of users based on the *privacy profiles* that seem to best match their preferences.

Results from this research will in turn influence the design and evaluation of *user interfaces for privacy policy disclosures* (Figure 1). Here *simplified privacy policy models* consisting of collections of *key privacy policy features* semi-automatically extracted from natural language website privacy policies will be presented in succinct and possibly personalized formats to users. We employ user-centered design to experiment and and evaluate different types of *disclosure primitives* (e.g. privacy nutrition labels, privacy icons, score/letter grades), drill-down formats that allow people to access more detailed disclosures, if they want to, as well as *just-in-time policy disclosures* that notify users of relevant policy features as they interact with different parts of a website. By regularly revisiting privacy policies and automatically identifying changes, it will also be possible to automatically highlight changes to users, taking into account when they last visited a site. Another important part of this research involves studying deleterious *cognitive and behavioral biases* that can lead users to make privacy decisions they may later regret, as well as researching disclosure interfaces that can effectively mitigate these biases (e.g. [Acq09,BLA+11, TKD+09, LAH+12]). We expect our *policy disclosure interfaces* to leverage results form our work analyzing website privacy policies. This includes both reasoning about policies (e.g. checking for compliance with relevant law and regulations, looking for inconsistencies) as well as statistical analysis (e.g. identifying unusual policy features that should be highlighted to users). Our expectation is that results from these analyses will also lead to the publication of website privacy reports, looking at trends in different sectors, non-compliance metrics and other relevant observations likely to be of interest to policy makers.

As our work progresses, we expect to address increasingly rich collections of privacy policy features. Initial features include looking at the collection and use of sensitive data (e.g. contact information, location, health, financial information), including whether use of collected information is internal only or also external, as well as looking at security policies and data deletion policies. Over time, the selection of these features will be informed by increasingly sophisticated models of privacy preferences. We also anticipate developing and experimenting with increasingly sophisticated combinations of natural language processing techniques to extract these policy features and increasingly sophisticated techniques to analyze the resulting policy models.

As work progresses, we will refine our crowdsourcing mechanisms. We have started with relatively small-scale prototypes and conducted experiments involving between tens and hundreds of crowdworkers (e.g. law students, privacy professionals, Amazon Mechanical Turkers). Over time, our objective is to scale to significantly larger and more diverse groups of crowdworkers. Specifically, crowdsourcing is being used for two different purposes: (1) to develop models of people's privacy preferences, in a manner similar to that demonstrated by Sadeh's work in mobile app privacy [LAH+12] and as further discussed in Section 4 ("Privacy Preference Modeling"), (2) to

assist with the extraction of policy features from natural language website privacy policies. At the time of writing, initial small-scale crowdsourcing experiments are being conducted and are helping us develop high quality corpora of policy annotations, which we are using as training samples to develop initial classifiers (e.g. beyond the original work reported by Smith, Sadeh and Wilson in [AWS+12]). These corpora will also help us evaluate the viability of fully automated functionality aimed at extracting an initial set of relatively simple policy features. They are also being analyzed to develop NLP techniques aimed at identifying and extracting text fragments (e.g. sentences or paragraphs) that pertain to different key policy features, as a way of reducing the amount of work to be crowdsourced in the context of semi-automated solutions. The goal here is for NLP to help reduce manual labor, enabling crowdworkers to zoom in on relevant text fragments rather than requiring them to read (and re-read) entire policies. This includes experimenting with different crowdsourcing interfaces and task workflows, as well as different combinations of NLP techniques. As work progresses, we hope to open up our crowdsourcing platform to accommodate a broader variety of crowdworkers and experiment with different crowdsourcing mechanisms to motivate them (e.g. [CK12]) and manage quality [ABI+13], including the introduction of roles such as reviewers and mechanisms to keep track of the quality of individual workers (e.g.[BLM+10]).

Considering that many websites have privacy policies that can amount to five or more pages of text, this type of semi-automated approach can already be expected to significantly reduce the total amount of work to be crowdsourced. It could also lead to higher levels of accuracy, which in turn could help further reduce the number of crowdworkers required – lower levels of accuracy require using more crowdworkers to disambiguate policy interpretations. As our natural language processing techniques improve, enabling us to achieve greater levels of automation in the extraction of policy features, we hope to scale our crowdsourcing efforts, effectively tackling a larger number of privacy policies (and websites) and identifying a greater number of policy features, at an increasingly lower cost. We are also currently experimenting with approaches and mechanisms aimed at identifying high quality crowdworkers with the eventual goal of further enticing them to produce quality annotations. We envision mechanisms that rely primarily on volunteer workers (e.g. possibly enticed through a game-based crowdsourcing framework), with perhaps a small group of paid experts responsible for disambiguating policy features that are more difficult to extract and to help assess the quality of the work produced by others.

As we make our results available to the public at large through the release of our privacy disclosure interfaces (e.g. web browser plug-in), we also hope to eventually entice website operators to verify our interpretation of their policies and offer them mechanisms to correct these interpretations or clarify their policies. This would effectively amount to yet another layer of crowdsourcing. Developing such mechanisms would obviously entail authenticating website operators and most probably holding them accountable for the changes they submit, effectively turning their submissions into extensions of their natural language privacy policies. If we can produce results that are of high enough quality to entice website operators to provide such corrections or clarifications, we would effectively have turned around the current state of affairs. We would have managed to move from a situation where website policies are ambiguous because they are primarily intended to protect website operators (rather than inform the public), to a situation where website operators feel compelled to clarify their policies and make them more transparent and meaningful to users. While such an outcome is far from guaranteed, a project like this one has the potential of getting us a lot closer to it and can also help inform public policy debates about what legal or regulatory changes might be needed to get us there.

In the following, we discuss our efforts in each of the involved research strands in more detail.

3. Natural Language Analysis for Semi-Automated Policy Feature Extraction

One of the longstanding goals of the field of natural language processing (NLP) is to develop robust algorithms that can interpret text into data structures that are "actionable" by machines. Among NLP's recent successes: automatic translation systems good enough for an English speaker to get a basic understanding of documents written in select other languages, and web-based question answering accurate enough to win Jeopardy! when competing against the best human players. Our aim is to develop NLP algorithms that interpret website privacy policies well enough to populate data structures with key privacy policy features that capture the most important aspects of a policy that people are interested in. This is a new problem for NLP; we note some related natural language analysis problems and their solutions.

- **Text document categorization.** Many problems in NLP involve mapping a document to one of a relatively small set of simple, atomic labels. These problems include: assigning topical labels to news stories [Seb02], measuring the polarity of an author's sentiment in a product or restaurant review [PL08], identifying the author of a historical document [MW63], and predicting reader response to a document, such as whether a Congressional bill will pass [YSW12]. Text categorization is typically solved using machine-learned classifiers that make use of relatively simple representations of text, such as word frequency histograms. In preliminary work, Smith, Sadeh and their collaborators explored text categorization-inspired techniques for a single privacy policy feature and already obtained promising results [AWS+12].
- **Semantic parsing.** One classic view of natural language meaning is that the central aspect of a sentence's meaning is the real-world conditions under which the sentence is true (i.e., truth-conditional semantics). Building on Montague's idea that natural languages like English can be treated in much the same way as formal ones like programming languages, the technical challenge for semantic parsing is to map natural language strings to expressions in logic (typically first-order logic) using syntactic analysis and the lambda calculus; see [Car98] for a complete discussion. Treating natural language the same way we treat programming languages is appealing to those seeking artificial intelligence programs that can go beyond representing what is meant to make inferences. Methods for representing truth conditions have found success in narrow domains (e.g., relative to a simple database of geographic relations [ZM96]) where the portion of the world that might be discussed is relatively small. In recent developments, machine learning continues to play a central role in exploiting various kinds of data (text, logical expression annotations, and databases) to learn the mapping from sentences to logical expressions.
- **Predicate-argument analysis.** Between the two extremes of "shallow" text categorization and "deep" semantic parsing are myriad alternatives. Incomplete representations of text meaning have attracted considerable attention: this includes identifying events and relevant participants in those events, and the roles they play [GJ02]. Our recent work has explored how rich linguistic resources and statistical machine learning from human-annotated and raw text can be combined to build a state-of-the-art statistical frame parser [DCM+13]. In specific domains, there are also ways to model regularities in document structure [EB08]; for example, in privacy policies, there is often a section on information collected, followed by how it is used.

The NLP analysis of privacy policies must do more than simple text categorization; the output will be multidimensional data structures that answer a range of privacy policy questions that are important to a user. Initially, we will consider one question at a time. This will allow us to better understand what makes some questions more or less difficult to automatically answer, and to focus effort on automating natural language analysis where it will be most helpful (e.g., directly answering questions about policies, or supporting human annotators to assist with selecting from a list of prospective answers, as we discuss below). Our prior work includes extensive manual analysis of over 100 privacy policies to map natural language policy statements into frame-based and first-order logic representations [BA05, BAD09]. Our related work analyzing regulatory text has yielded natural language heuristics, some expressible as simple regular expressions, that can be used to identify frame-based representations of actions (see Breaux and Anton [BA08] for study results extracting over 300 access control requirements) and whether actions on information are permitted, required or prohibited with various conditions, exceptions and purposes [Bre09]. These heuristics can be used to identify statements and phrases that answer common questions, e.g., for what purposes a user's personal information may be used, or with whom this information can be shared. More recently, we conducted a preliminary experiment to use NLP methods to consider a single question: whether a privacy policy is considered clear (by humans) about a particular set of procedures pertaining to sensitive user data [AWS+12]. Using crowdsourced annotations from the "Terms of Service; Didn't Read" project (http://tosdr.org) [ToS12], we trained a statistical classifier that achieved 84% held-out accuracy, which is 20% better than a baseline of always guessing that the policy is unclear.

We believe significant improvement is possible, based on two observations. First, the model used only a frequency histogram on sequences of one, two, or three words; there are many richer representations possible. Second, this model was built and evaluated in a leave-one-out setup using only 19 instances (only 7 marked "transparent"). Indeed, when considering a related question, whether a privacy policy gives a user the right to voluntarily cancel, terminate or delete his or her account, we had only 18 policies, 3 positive, and were unable to improve

over the baseline. This highlights the potential for improvements by acquiring more data. We consider these two directions, representation and data acquisition, below.

3.1 Text representations for NLP on privacy policies

Given a privacy policy and a particular question (e.g., "are the procedures an individual must follow to seek correction of erroneous data clear?"), we seek to discover a representation that supports answering the question. One approach is to find an appropriate passage of text that addresses this question (the phrase erroneous or a synonym or hypernym will likely be found in such a passage) and then more deeply analyzing the text to infer the clarity of the procedures. In some cases, simply matching text segments against similar "boilerplate" found in annotated examples might suffice. In others, abstract representations and inferences will likely be required. We therefore require a flexible machine learning framework that can accommodate such representations and textual cues at many granularities. Smith's recent book synthesizes appropriate techniques in this area of research [Smi11].

One desideratum that drives our NLP development is the importance of interpretability of any NLP model, for researchers on the project team (especially those who are not NLP experts), and for anyone wishing to understand why a conclusion is being drawn about a privacy policy. This requirement drives the design of both our linguistic representations and the algorithms that make use of them to answer questions. We have extensive experience with lightweight syntactic and semantic analyses that are intuitive, such as dependency parses [MSA+11a] and frame-semantic parses [DCM+13]. Further, we have developed statistical learning models specifically for NLP that are driven by the need for interpretability, emphasizing sparse linear models that use a small number of clues in the final predictor (e.g., [MSA+11b]).

We note that, while the application of automated analysis of privacy policies is our primary goal, such applications tend to drive innovation in basic NLP research. We expect that the techniques developed within this project will be of great interest in the core NLP community as well, and will have broader impact in suggesting techniques for other text analysis problems. For example, in other work, we have explored financial disclosures [KLR+09], Congressional bills [YSW12], scientific articles [YHO+11], and many others.

3.2 Acquiring annotations and feedback to support statistical NLP

While some volunteer efforts for annotating privacy policies along various dimensions exist (e.g., the "Terms of Service; Didn't Read" project noted above and a dataset of several thousand annotated instances from PrivacyChoice), there is currently insufficient data (lacking in both quantity and nuance) to learn highly accurate models to answer a wide range of key privacy questions. For example, tacit knowledge is often needed to complete a partial, semi-formal representation acquired from a privacy policy. In our prior work, we discovered manual techniques that analysts can use to categorically identify and resolve natural language ambiguities and to infer implied rights from obligations statements [BVA06]. In this project, we employ a range of approaches to gathering better data. The first of these is to simply train and hire annotators to answer key questions about policies based on a close reading and by using some of our previously discovered techniques. We believe that we can make the annotation task much easier using some simple automation. For example, if for a given question we can locate the relevant passages in a (long) privacy policy, we can speed up the job of an annotator. If we can simplify the task enough through automated preprocessing into independent subtasks, the result may be a candidate for crowdsourcing to non-expert annotators. In this case, we can use multiple non-experts to simulate an expert judgment [SOJN08]. We are leveraging our prior experience evaluating human analysts and their abilities to apply annotations and reach consensus about the meaning of annotated policy statements in this project [Bre09].

In addition to the time-tested annotation approach, we plan to explore feedback mechanisms. We foresee that some website owners will notice our algorithm's interpretation of their policy, and wish to provide a clarification or request modifications. By developing a platform that allows third-party modification to annotations, and that provides the ability to show which passages are relevant to a correct understanding of the policy, we can augment annotations or direct additional crowdsourced effort to reconsider an earlier annotation. We emphasize that annotation must always be treated as evidence, not ground truth, since humans can make mistakes and even behave

adversarially. Current statistical methodology helps in identifying anomalies in such settings [Car08] and we aim to apply these techniques to detect inconsistency and disagreement.

3.3 Automatic alignment for new feature detection.

Over time, once a balance between human annotation (at that stage, primarily through crowdsourcing) and automated NLP has been struck, we plan to revisit the representation question to tackle a more difficult challenge: as software and hardware evolves, new privacy issues will inevitably surface. The key privacy questions of the future may differ substantially from those of today, and we expect this to be reflected in the policies themselves. We therefore plan to explore new statistical analysis methods that align privacy policies to each other (i.e., unaligned portions of the text), both within and across time frames. This will allow us to detect new segment types as they begin to appear. The attention of analysts and, if appropriate, annotators, can then be directed to these sections in aggregate, enabling the identification of new key privacy features to be semi-automatically detected.

4. Privacy Preference Modeling for Usable Privacy Notices

In order for privacy policy information to be useful to end users, it must be displayed in a format that is accessible to them and addresses their needs. In our work, we explore novel ways of simplifying and dynamically adjusting the presentation of privacy policies to users, including the use of machine learning to identify privacy profiles that can be used to tailor presentation to different groups of individuals.

4.1 Usable Privacy Notice Interfaces: Privacy Nutrition Labels and Beyond

Previous research by Cranor, Acquisti, Sadeh, and their collaborators on different privacy notice interfaces has shown that it is possible to develop succinct privacy disclosures that reduce the amount of time required from users to observe and understand them, yet retain key elements of privacy policies. For example, the Privacy Bird plugin for Internet Explorer displayed a bird icon in the corner of the user's browser window that changed colors to serve as a persistent indicator as to whether the website the user was visiting had a policy that matched her privacy preference settings [CGA06]. Users could click on the bird icon to get information about any mismatches with their privacy preferences as well as a summary of the site's privacy policy. Privacy Bird met the needs of users who wanted information about the privacy policies associated with the websites they visited, but it required users to visit a website in order to get this information. To address this problem, Cranor, Acquisti, and collaborators developed and tested Privacy Finder, a search engine that annotates search results with privacy information, similar to the information provided by Privacy Bird [CGA06]. Their laboratory studies found that users were able to make use of Privacy Finder information to help them choose more privacy protective vendors from which to purchase products online. Many users were shown to be willing to pay a premium to shop at the more privacy protective websites [TEC+11]. A subsequent study found that the timing and placement of the privacy information (e.g. in search results, in an interstitial, or at the top of the vendor's website) were significant factors in whether or not users selected the more privacy protective vendors [ETC+09]. Similar results were recently observed in the context of mobile app privacy disclosures [KCN13].

Cranor and collaborators have also developed and experimented with privacy nutrition labels intended to simplify the presentation of privacy policies by displaying those elements most important to users in a standardized format that is both succinct and easy to interpret. Inspired by food nutrition labels and work on financial privacy notices [Kle08], Kelley et al. used an iterative design process to develop and test a privacy nutrition label. User evaluations suggest that the latest privacy nutrition label iteration allows consumers to find information more quickly and accurately than when relying directly on plain natural language privacy policies [KCB+10]. Privacy nutrition labels are shorter and easier to read and interpret than natural language policies. Their standardized tabular format allows users to learn where to look for answers to particular questions (e.g. whether some information is shared with marketers or not) and facilitates comparison between policies. In addition, the use of colored symbols allows users to get an overview of a policy at a glance from observing the overall color intensity of a policy [KCB+10].

While effective for users, these previous efforts faced the challenge that they all relied on the availability of machine-readable privacy policies (e.g. P3P policies) or, more generally, the willingness of website operators to provide the necessary information (e.g. nutrition labels), which has not materialized. In contrast, our current work on semi-automated understanding of natural language privacy policies offers the prospect of being able to extract the information required by these new types of interfaces without reliance on cooperation by the website provider. Specifically, our experience with interfaces such as privacy nutrition labels provided us with an initial set of questions we are in the process of answering with semi-automated approaches that combine natural language processing and crowdsourcing. In addition, as we develop scalable solutions to semi-automatically extract this information from existing natural language privacy policies, we are further refining our understanding of which policy features are most important to users. This information will inform further design iterations of privacy nutrition labels and provide a basis for experimenting with variations of these interfaces. Beyond privacy nutrition labels, we also plan to explore the effectiveness of "privacy scores" or "privacy grades" (e.g., as used in the Tos;DR browser plugin [ToS12]) that reflect key concerns of users and their reactions to different policy statements. Such privacy scores can act as extremely valuable tools for companies seeking to compete on trustworthy practices, which is an important international regulatory concern.[3]

We are in the process of conducting a series of online surveys to gain a more detailed understanding of what type of privacy information is most important to display to users [LUW+2013]. For example, while data sharing has been identified previously as something users find important, more work is needed to determine what distinctions users make about the type of data sharing. Is it sufficient to classify data sharing into two categories: data sharing only necessary to complete a transaction vs. data sharing for any purpose? Or is it important to also distinguish the type of data to be shared (e.g. personally identifiable or non-identifiable), the purpose of sharing (e.g. online targeted ads, telemarketing, unrestricted), or with whom the information is being shared (e.g., payment providers versus advertisers)? We recently conducted a study that presented various web browsing scenarios to users in the context of online behavioral advertising, highlighting the privacy practices of the websites used in each scenario. Users were asked about their willingness to provide various types of information in each scenario. By observing which practices result in changes in users' willingness to share data we were able to gain deeper insights into the distinctions that users find important. Our results indicate that the willingness to share is impacted by the perceived utility and necessity of sharing information, and whether participants believed a specific type of information should be used for targeted advertising.

4.2 User Privacy Preference Modeling to Identify Key Policy Features

Our research on privacy nutrition labels as well as on modeling people's location privacy preferences and their mobile app privacy preferences suggests that whether a user is comfortable with a website's privacy practices is often determined by a relatively small number of privacy policy features. To support the development of privacy notice interfaces that are both succinct and informative, we aim to develop deeper, more systematic models of those policy features that are most important to users.

In prior work in location privacy, Sadeh, Cranor, and collaborators have developed methodologies to quantify the benefits of exposing different combinations of privacy settings to users [BKS+11]. Most recently, Sadeh and collaborators developed a crowdsourcing methodology to identify those Android app permissions that matter most to users. This research showed that many permissions requested by Android apps are actually expected by users. For instance most Android users expect Google Maps to require access to their location. In contrast, some mobile app-permission pairs are less likely to be expected by users, such as Angry Birds requiring access to a smartphone user's location, or Pandora requiring access to the user's contacts list. When adequately disclosed to users, those unexpected app-permission pairs are also the ones most likely to determine whether these users feel comfortable downloading the app on their smartphones or not.

With similar intentions, we are developing models to capture those website privacy policy features that are most important to disclose to users. This includes identifying common policy features that users generally expect to see

[3] For instance, the proposed European Draft Data Protection Regulation seeks explicitly to encourage the development of privacy seals to simplify compliance with privacy policies and legal obligations.

at different categories of websites and that they generally do not find objectionable, as well as policy features that are more unusual or more likely to be viewed as objectionable. Currently, our investigation focuses on relatively coarse policy features, i.e., whether a policy makes statements about collection and sharing of specific information types and whether users are able to access and delete data provided to or collected by a website. Over time, we plan to explore finer nuances in privacy policies. Similarly, we started by assessing fairly broad categories of websites (e.g., financial, health, e-tailing, news, entertainment, social). As our work progresses, we anticipate differentiating between finer categories of websites and also between different groups of users, as it is well known that not all users express the same level of sensitivity to different privacy considerations.

In our research, we initially focused on small sets of manually annotated privacy policies and are currently expanding our efforts by crowdsourcing information about how users feel about different policy features (e.g. level of surprise, level of comfort). This approach is informed by prior experience using Amazon Mechanical Turk to crowdsource mobile app privacy preferences [LAH+12], the evaluation of privacy nutrition labels [KCB+10], and the evaluation of password policies [KSG+11]. With respective studies we also study the effectiveness and quality of policy annotations provided by experts or skilled annotators (e.g., law or public policy students) compared to untrained crowdworkers.

As our natural language processing techniques mature, we expect to be able to scale up this research by sampling significantly more diverse sets of privacy policies (e.g. more diverse categories of websites and a broader set of policies) and extracting more diverse policy features as well. We employ a mixed methods approach to data analysis that combines quantitative results with qualitative data analysis based on semi-structured interviews. This approach allows us to not only learn about user preferences but also gain a deeper understanding of what shapes these preferences and drives the privacy decision making processes of our study participants. We hope that our results will lead to and inform the development of more dynamic privacy disclosure interfaces. For instance, privacy nutrition labels that are not static but highlight different privacy policy features based on expectations for the particular website being considered and the particular policy at that site. Such interfaces require a nuanced understanding of privacy preferences as well as perceptions of privacy policy features.

As our ability to semi-automatically extract policy features matures, enabling us to process a larger number of websites, we also plan to use clustering techniques to identify categories of websites that entail different sets of privacy preferences (e.g. sites where users expect to see and generally appear to be comfortable with some data practices versus sites where the same practices are less expected or are perceived as being more objectionable). The identification of these categories (or clusters) is expected to further boost our ability to scale up our approach, effectively making it possible to predict those website/policy feature pairs that are most important for a given category of websites without having to crowdsource privacy preferences for every single website in that category. Annotation and analysis of whole categories of websites will further allow us to model expectations of data practices or features typically found in privacy policies of websites from the same category, which in turn could serve to identify positive or negative deviations of a given policy from those expectations.

4.3 Privacy Profiles for Personalized Privacy Notices

Another aspect we are considering in the context of user preference modeling is that when it comes to privacy, not all users share the same preferences [BGS05, Ray10, KC05]. In their work on Privacy Bird and Privacy Finder, Cranor et al. used a small number of website privacy preference settings to allow users to quickly select among different sensitivity levels when evaluating the privacy policies of websites [CGA06]. More recently, Sadeh and collaborators have been working on clustering techniques to automatically identify privacy profiles based on crowdsourced privacy preference information in the context of location sharing [RBK+09, MSS11]. The results suggest that, while people's privacy preferences are diverse, it is often possible to identify a relatively small set of profiles, which collectively can capture the preferences of a diverse population of users with relatively high accuracy. Essentially, every user in the population (or at least the vast majority of these users) can be closely associated with one of those privacy profiles. We plan to use similar clustering techniques to analyze people's crowdsourced website privacy preferences. This will involve identifying clusters of users that express similar levels of comfort and surprise with respect to policy features as they relate to different categories of websites or similar levels of concern about privacy issues such as "protection of children's privacy." Assuming that such clusters

and associated privacy profiles can be identified, we further plan to experiment with techniques that use this information to customize and personalize the presentation of privacy nutrition labels (and other notice displays) to individual users. Essentially, the profiles will help us highlight those policy features that are most important to a given user (e.g. policy features that seem to conflict with the user's expectations or preferences), based on the profile that best matches that user's preferences.

5. Mitigating Cognitive and Behavioral Biases

Acquisti, Cranor, Sadeh, and their collaborators have further applied theories and methodologies from a variety of disciplines (including usability research, HCI, behavioral decision research, and behavioral economics) to the understanding of privacy decision making and the hurdles that hamper it [AG07]. One of the insights gained from that line of research has been the recognition that problems of asymmetric or incomplete information, which privacy policies aim to address, are but one of the many obstacles individuals face when trying to make the "right" privacy decision – that is, a decision they will not regret later on.

However, even when information is actually available about how one's personal data will be collected and used, problems of bounded rationality, such as the inability to consider the full consequences of revealing certain information [SIM82], and an array of behavioral and cognitive biases may affect and hamper the decision to disclose or protect personal information. Such biases constitute systematic deviations from canonically rational economic decision making [CL03]. Typical examples are hyperbolic discounting [RO00], or problems with lack of foresight and will power [LH07]. An ever-expanding stream of studies has investigated how those biases may in fact affect privacy valuations and decision making [BAL12, AJL12, JAL11]. A major part of our prior and ongoing research has been to not only identify the effects of such hurdles on privacy decision making, but also investigating ways to address or mitigate their impact. For instance, privacy nutrition labels resulted in measureable improvements in user comprehension [KCB+10] and we found that people are willing to pay more for enhanced privacy when presented with a simple indicator for privacy [TEC+11]. We also have and continue to investigate the role of privacy "nudges," which are soft paternalistic interventions [TS08], in ameliorating privacy decision making and helping individuals avoid disclosures (or, inversely, lack of disclosures) that they may later regret [BLA+11].

With regard to mitigating cognitive and behavioral biases in the context of users accessing websites, we are conducting research in multiple directions.

5.1 Ex Ante and Ex Post Privacy Preferences

Our research loosely draws inspiration from behavioral economist George Loewenstein's theory of hot-cold empathy gaps [Lo05], to investigate the gap between ex ante and ex post privacy preferences, valuations, and decisions. By "ex ante," we refer to the state before a certain privacy-sensitive event occurs or behavior is exhibited; accordingly, by "ex post," we refer to the state following said behavior or outcome. We conjecture that individuals may express and act upon different privacy preferences and valuations in ex ante versus ex post states. For instance, before or after the decision to disclose certain personal information is taken or a privacy invasion has occurred. It would follow that individuals may claim to desire certain privacy features in a system ex ante (for instance, "I want to have [an interface that gives me] certain controls over who can get access to my personal information"), but may actually realize ex post that a different set of features may have reduced their privacy regrets (for instance, "I would have liked to have different default settings for the visibility of my information"). Accordingly, in order to develop effective tools for notification and control, we are in the process of planning and conducting user studies and experiments to compare ex ante and ex post privacy preferences, valuations, and objectives. For instance, we plan to contrast ex ante claims by individuals about the importance of different features in privacy policies to ex post considerations about which features may have decreased user regret. We further plan to develop auditing interfaces to crowdsource the daily collection of data on information disclosures, following the studies we conducted in [WKL+11] and [SCK+13]. A desired outcome of this investigation would be the analysis of the "optimal" amount of information to provide to users about a privacy policy – less information may be

sometimes needed in order to overcome problems of bounded rationality, but more data may at other times be needed to guide informed decision making.

5.2 The Impact of Default Settings on Privacy Preferences and Information Disclosure

More expressive settings often require more initial decisions by the designer, such as which information to protect or disclose by default in a location-tracking service [SHC+09,RBK+09,BKS+11]. These initial decisions may not only be interpreted by the end-user as vested of intrinsic value, but may also create path-dependent dynamics where two individuals with similar privacy preferences may end up using the system with completely different disclosure settings. Indeed, status quo bias – the tendency to stick with default settings – has been extensively investigated by behavioral economists [SZ88]. Much less explored, however, is the impact of default settings in an expressive environment of privacy preferences. We are investigating whether default settings end up overriding idiosyncratic individual privacy preferences by assessing self-reported attitudes towards privacy and disclosure ex ante, and then manipulating the default settings assigned to subjects with similar attitudes across the three variants of the mechanisms presented above. Across individuals with ex ante similar sensitivity and attitudes, we will observe, ex post, the relative influence of the initial attitudes versus the settings chosen by the researchers.

5.3 Effective Privacy Controls

Research on ex ante and ex post privacy preferences also informs novel research on effective tools for privacy control. Ongoing research has uncovered paradoxical effects associated with giving individuals more control over their personal information: control can create a sense of protection or even overconfidence, leading individuals to take more risks with their personal information. For instance, sharing more sensitive information with strangers [BAL12]. While users typically derive pleasure from control [Wh59, De68], too much control may therefore daunt users, raising their cognitive costs and the likelihood of mistakes, or (as in the above mentioned study) making them overconfident.

We are investigating if and how such control paradoxes can be avoided. We plan to test and contrast various mechanisms for control, for instance by randomly assigning sets of users to one of several experimental conditions in the context of a browser add-on that highlights different sets of privacy policy features to users:

- **High control:** a condition in which users of the browser add-on are presented with a high number of privacy scenarios (for instance, highly granular privacy "nutrition labels" information), and are asked to specify rules indicating the conditions under which they would be willing to share their information with others under those scenarios;
- **Medium control:** a condition in which users of the browser add-on are presented with a lower number of scenarios, and then, based on pattern comparison between the users' selections and those that other similar users have already found to be desirable, the system completes the set of rules for the user;
- **Low control:** a version where users are simply asked to choose between a few default settings.

By observing accuracy of, and satisfaction with, rules and preferences expressed through the various conditions, along the finer-grained lines discussed above, we will be able to identity which mechanisms for control achieve more desirable combinations of ex ante preferences and ex post satisfaction.

5.4 Just-in-Time Disclosure Notices

Our previous research has uncovered some fundamental, systemic limits in the ability of disclosures and notices to affect privacy behavior [AAB12]. One of the goals of this project is to investigate whether nudging techniques could address and overcome those limitations. In the non-privacy literature on disclosure, one can find examples both of notifications that can be quite effective in influencing behavior (for instance, cleanliness inspection ratings visibly posted outside restaurants) and notifications that are less effective (for instance, health warnings on cigarette packages). We leverage lessons learnt from that literature in order to investigate the effectiveness of "just in time" disclosures. In particular, we plan to investigate whether it would be preferable to move away from static, monolithic privacy nutrition labels and instead disaggregate these labels (or other representations of privacy poli-

cy features) to disclose them in a more incremental, just-in-time manner. This will involve presenting different elements of a site's policy to the user in a just-in-time fashion, based on the particular interactions a particular user is having with the website. The idea is that privacy notifications would essentially appear at the moment when an individual is about to make a certain privacy sensitive decision. Such just-in-time notices may take multiple forms, including icons, nutrition labels, or short textual notices. Prior work on nudging can inform which specific moments may be most effective to present salient information to users in actionable ways. We plan to apply these ideas to the privacy domain. For instance, as a user interacts with different parts of a website and possibly provides different types of information, we would inform the user about the specific privacy policy practices that pertain to that particular interaction (e.g. a particular online form or even a particular field in such a form). This includes modifying the language, the graphical representation, the frequency (and so forth), of these interventions, and evaluating to what extent this impacts regrets or satisfaction.

6. Privacy Policy Analysis

As the principal instrument to improving individual awareness of privacy practices, privacy policies have evolved with new technologies. Privacy policy evolution is the result of multiple forces: governments have introduced laws to shape privacy policy formats, wording and what practices must be described in policies and companies may introduce new formats aimed at improving awareness, simplifying their business practices or changing their business practices over time. Government regulations impose constraints on privacy policies. For example, the GLBA and HIPAA Privacy Rules governing finance and healthcare, respectively, the COPPA Rule governing children's information, the Video Privacy Protection Act (VPPA) and several U.S. national security laws all impose constraints on how privacy policies are written. Similarly, European data protection laws as well as other foreign laws impose notice requirements for data processing. These constraints require precise descriptions of legally compliant data practices and explicit statements of specific consumer rights, such as the right to request copies of personal information or the right to opt-out of marketing communications. By controlling the language in a privacy policy, these laws and regulations compel organizations to engage in certain practices and to communicate information about those practices with their consumers.

On the business side, companies have competing pressures for their privacy policies. In late 2012, for example, Google consolidated their privacy policies from over 60 services into a single policy that collectively governs their user's data. Google's policy consolidation made it easier for Google to share data about the same user across their services (e.g., Gmail, Google Search, Google Maps, etc.) by introducing numerous textual ambiguities in the notice. The public and government reaction to this consolidation was largely negative, because the increased information sharing among services was contrary to most users' expectations for their data and came without commensurate opportunities to opt-out, among other concerns [Art12]. Google, however, believes that the aggregation of data across services will enable the company to explore more innovative offerings that employ advanced analytics across the services. While we highlight Google, other companies, including Amazon, Apple, Facebook, Microsoft, and many more must contend with the same issue. Due to regulatory enforcement actions covering privacy policies, the policy is both a communication tool between companies and users as well as a general specification about how businesses handle personal information.

In addition, companies with an international presence must comply with foreign law as well as federal and state law in the United States. For data coming to U.S. organizations, this includes considering a "Safe Harbor" agreement to reconcile gaps between US and EU law, or complying directly with a foreign regulation. Multiple forces can lead to highly complex privacy policies that serve different industry, government and consumer communities, simultaneously. The challenge for users is assessing what portions of the policy are relevant to their preferences and, moreover, what are the logical implications of these policy statements in regard to how their information will be used. Longer term, policy makers and researchers are interested in how these policies change over time and what types of policy statements lead to improving privacy and regulatory harmony.

6.1 Modeling and Reasoning about Privacy Policies

Formal and semi-formal notations can be used to model privacy policies and legal regulation and to extract precise specifications of consumer, company, and government rights and obligations [Bre09]. Using empirically validated methods [Bre09], a human analyst can extract privacy requirements from policies and legal regulation into a canonical form consisting of semantic roles that answer privacy-relevant questions [BA08], such as: who is the data subject, what data is collected, used and disclosed and to whom, for what purpose and whether these practices are permitted, required or prohibited. Semi-formal encoding is used to systematically identify ambiguities, such as missing role values (e.g., purposes or data recipients) and uncertain attachments for clauses that would otherwise clarify to which information types a given purpose or data recipient applies [BVA06]. Semi-formal encodings can be refined further in Description Logic (DL) [BCM03], which we have used to formally reason about possible interpretations [BAD08]: do company obligations imply consumer rights, or are there conflicts between government-imposed obligations and company rights? For very large policies (more than a few pages), these types of inferences are difficult for a human to trace without this formalization. Frequently, privacy policies cover multiple overlapping stakeholder categories and rich data environments, such as Google's 60+ services covering e-mail, chats, video, the Web, and so on. Furthermore, privacy policies can be written from multiple viewpoints, e.g., the consumer's, company's, third party's or government's viewpoint, simultaneously.

Due to the large number of privacy policies in the wild and the rate at which policies are changing, we need new tools and techniques to automate the encoding and analysis process. We propose to integrate our empirically validated, manual methods with emerging NLP-based techniques to further automate the extraction of formally encoded privacy practices. The first goal led by Breaux is to identify and formalize a subset of privacy policy semantics to logically answer questions most relevant to users about their personal information. The research challenge includes demonstrating that multiple analysts can consistently encode statements into the same formalism, and that the questions to be answered are both useful and difficult to infer by simply reading the text alone. For example, consider the excerpt from a Google Privacy Policy in Figure 2; ellipses are used for brevity.

1 We do not share personal information with companies, organizations and individuals outside of Google un-
2 less one of the following circumstances apply:
3 With your consent…We require opt-in consent for the sharing of any sensitive personal information.
4 With domain administrators… Your domain administrator may be able to:
5 • … access or retain information stored as part of your account.
6
7 • receive your account information in order to satisfy applicable law, regulation, legal process or enforce-
8 able governmental request.
• restrict your ability to delete or edit information or privacy settings.

Figure 2: Sample privacy policy excerpt from the Google Privacy Policy, June 27, 2012

This policy can be impenetrable and often meaningless to a user who has a specification question: are my personal web searches available to my employer who administers my Google e-mail account? To answer such a question, consider the technical challenges. First, a "policy" can consist of multiple hyperlinked documents, including separate policies for specific services or business practices (e.g., advertising), lists of definitions (see blue hyperlink Figure 2, line 3), illustrating examples, and so on. Information from these documents must be integrated into a single context, and definitions often include categories of things (e.g., sensitive information includes sexual orientation, but not IP address or server logs.) Second, permissions can exist alongside potentially conflicting prohibitions: in Figure 2, lines 1-2, a prohibition is indicated by the keywords "do not share," which has a broad set of exceptions (or permissions) indicated by "unless…" and contained in the subsequent paragraphs. Third, meaningful policy statements often span multiple paragraphs and language features that a reader needs to trace across these paragraphs: in Figure 2, lines 3-4, the two separate conditions appear that permit sharing with consent or without consent, when sharing with domain administrators. When encoding these statements in logic, the meaning of certain statements may have different implications. First, exceptions indicate priorities between what is permitted and what is prohibited in order to de-conflict a policy interpretation [BA08]. Second, opt-in consent assumes that a covered sharing activity is initially prohibited; only with consent, does the activity become permitted, which illustrates how policies describe multiple worlds depending on a user's preferences. Opt-out consent has asym-

metric semantics. Third, the semantics of data retention and deletion (see line 8, Figure 2) is temporal: data must be collected before it can be retained or deleted, and premature disposal prevents obligations to share or retain data. In our approach, we aim to build on the extensive prior work to formalize privacy policy [CGA06, Kag04, May08, PS03, UBL+08] and privacy laws [BDM+06, Bre09, DGJ10, HBK+07], while we plan to more principally study how formal policy implications affect user privacy preferences and privacy policy authorship.

Our first goal in formalization will yield empirically validated heuristics for use by human analysts to map natural language policy statements into logic. This validation will consist of measures of repeatability and reliability across different policy types and business sectors. In addition to internal consistency (i.e., do privacy policies contain conflicting statements?), we aim to formally check external consistency as well: are privacy policies consistent with specific regulatory requirements? Formalization can detect the conflicts based on the axioms of Deontic Logic, which allow us to check for conflicts between permissions and prohibitions [Hor93], and temporal logic, which allows us to check for conflicts among data retention policies. To this end, we will select policies that are governed by portions of relevant privacy law in the U.S. and Europe. We plan to coordinate these results with our NLP efforts to automate the encoding process and, in particular, to identify the degree of hybridized machine automation and human analysis that yields the best outcomes with regard to heuristic reliability. We expect that automation will greatly reduce human effort by pairing select policy statements with relevant heuristics while maintaining high recall (i.e., few to no false negatives, but with potential false positives).

The second goal led by Breaux in collaboration with Cranor, Sadeh and Reidenberg is to further integrate our privacy policy formalization with privacy displays. Under this goal, we aim to develop logically comparable formalizations of privacy practices expressed in logic. We have done this informally for arbitrary regulations [GB12] and more formally for simple privacy statements in prior work [BAD08], but we now aim to formalize more complex policy statements with richer implications on user privacy preferences. The ability to formally compare data practices enables relative scoring between data practices and user privacy preference: e.g., what is the relative dissimilarity between "marketing," "targeted marketing" and "third-party marketing." This in turn can be used to personalize displays for particular users by comparing the user's privacy preferences to a particular policy, and by comparing a particular policy to a community "average" across multiple policies: e.g., this company's policy is more or less lenient in obtaining consent for data re-purposing, or this company shares information with a broader category of third-parties. We envision that the policy formalization can be updated as policies change, which in turn yields new changes to privacy displays that use the formalization to measure levels of privacy; the deltas of which can be monitored against user thresholds based on surveys and personas. Finally, we expect to incorporate findings from our work on cognitive and behavioral biases. If users are likely to assume certain terms exclude emerging variants (e.g., marketing is likely to mean direct-marketing from a first-party provider, but not include online behavioral advertising by third parties), then we can modify our DL axioms of inference to treat these interpretations consistently with prevailing user assumptions. Thus, we aim to check conflicts between user expectations and privacy policies using our formalization.

Finally, the third goal led by Breaux in collaboration with Reidenberg and McDonald aims to evaluate current regulatory regimes for their efficacy in shaping privacy policy. By mapping relevant regulatory requirements from laws to formalized privacy policy statements, we aim to examine the extent to which privacy policies and the logical implications contained therein are consistent with the broader policy and legal goals of the regulatory regime. In prior work, Breaux has worked on the HIPAA Privacy Rule and discovered the outcome of nuanced policy discussions had produced complex, multi-tier exception hierarchies in the HIPAA's information access regime. Reidenberg wrote the first published paper proposing the mapping of privacy law to formalized coding and highlighted the disconnect between legal statements and expressible machine readable code [REI97]. Such findings can serve to inform privacy policy writers about the rationale and implications of certain privacy statements.

6.2 Statistical Analysis of Privacy Policies

As this project proceeds, we expect to build up a collection of privacy policies that have been parsed, modeled, and analyzed. Besides its utility as a data source for user privacy tools, this collection will also provide a wealth of information about trends in both privacy disclosures and privacy practices. This will enable statistical analyses about the types of information included in or omitted from privacy policies and the frequency with which particu-

lar types of information or data practices are mentioned in privacy policies. It will also enable comparisons within and between industry sectors, geographic regions, and legal jurisdictions.

Conducting such statistical analyses of privacy policies is rarely done because it is labor intensive for analysts to gather the necessary information from a large number of policies. Some limited analysis was done by the FTC in a series of privacy sweeps conducted in the 1990s, and a more detailed analysis was done by a policy think tank in 2001 [AEL02, FTC98, FTC00]. Most subsequent large-scale privacy data collection efforts have taken more automated approaches, focusing only on those data points that were feasible to collect automatically. For example, [JSJ+07] used a web crawler to collect information about the use of cookies, privacy seals, and P3P policies. Cranor and McDonald and their collaborators investigated the prevalence of P3P policies on a variety of different types of websites and analyzed the collected P3P policies for errors and to determine the types of practices disclosed [CES+08]. In a subsequent study they used automated tools to look more closely at P3P policy errors and discovered that inaccurate P3P policies had become widespread [LCM+10].

Without accurate computer-readable privacy policies, conducting large-scale analysis of privacy policies has been difficult. The small-scale analyses that have been done suggest that larger-scale analyses would help inform policy making. For example, Cranor and collaborators analyzed financial privacy notices before and after the implementation of a new US financial privacy law and were able to observe improvements in the readability of financial privacy notices, but did not observe improvements in privacy practices [SC06]. The tools we are building in this project will allow us to build and maintain a large collection of privacy policies, modeled so as to facilitate this sort of analysis longitudinally, and on a larger scale. We aim to build a front-end to allow for querying the privacy policy collection in a variety of ways, including industry sector, geographic location, data practice, or period in time.

6.3 Public Policy Implications

We hope that our research will have broad public policy implications within the United States and Europe. In the United States, we have relied heavily upon the idea that informed users will make rational, individual privacy decisions, informed by companies' privacy policies. In practice, the notice and choice approach has failed users. When they read notices, they do not understand them and believe they are protected by laws that do not exist [TKH+09]. Users also believe reputational pressure will prevent companies from engaging in practices that are, to the contrary, quite widespread. The failure of the notice and choice model has resulted in numerous Congressional bills around data practices [Co11, Co12]. Some of these bills envision that details will be worked out by the Federal Trade Commission on the national level, or the state Attorney General's office on the state level. One of the persistent obstacles to legislation and regulation is the concern that policy makers do not have enough information to solve privacy problems. We hope that our work will contribute to fact-based policy decisions by establishing what information is currently available in privacy policies and what information is required to be in privacy notices if users are to make decisions. One of three outcomes is possible:

- We find that nearly all information needs are already met by publically available information. At that point, public policy should focus on how to make that information salient and actionable for users. This might take the form of calling upon web browser companies to improve their offerings, and could leverage our work creating browser extensions as one practical approach to enhanced decision making.
- We find that nearly no information needs are met by publically available information. This would suggest an insurmountable gap between what companies publish, and what users need to know. Public policy should then focus on improving the quality of information provided, perhaps by creating example best practices, or even by directly mandating elements that must be included in privacy policies.
- We find that some information needs are met, but a substantial portion is not. The public policy response here might be to bring increased focus on specific domains where companies are not providing the data users need.

7. Broader Impact

A Compelling Vision and a Novel Approach: Natural language privacy policies have become a de facto standard to address expectations of "notice and choice" on the Web [FTC10]. Yet, there is ample evidence that users generally do not read these policies and that those who occasionally do struggle to understand what they read. As new opportunities to collect, exploit and share data continue to emerge, this problem is only getting worse. Building on recent advances by the authors in complementary research areas, this project offers the prospect of dramatically changing this situation. Specifically, we will combine linguistic analysis and crowdsourcing technologies to develop a scalable infrastructure for semi-automatically extracting key features from existing natural language website privacy policies and use novel interface technologies to effectively present users with depictions of those policy features likely to matter most to them, thereby empowering them to make effective privacy decisions.

Work in this project combines several complementary strands of research: natural language analysis for semi-automated privacy policy extractions, user privacy preference modeling coupled with the design and evaluation of novel privacy disclosure technologies, research on mitigating critical cognitive and behavioral biases in privacy decisions, privacy policy analysis work combining formal methods, statistical analysis and work in public policy. This project is conducted by a multi-disciplinary team of experts and directly builds on recent results from their respective research. Short of just producing scientific and public policy publications, an important part of our research involves the development and refinement of practical tools and a scalable infrastructure to extract key privacy policy features and present them in an effective manner to everyday Web users. For this purpose, we are developing a respective browser extension. We also plan to release a tool to help website operators clarify their privacy policies, leveraging advances in formal methods to model and reason about privacy policies (e.g. identification of conflicts).

We further hope that our efforts in curating a corpus of privacy policies and tracking these policies over time will be beneficial to the broader privacy and NLP research communities. This corpus is of practical relevance to our project for updating our models; indeed, to the extent that our NLP algorithms are interpretable, websites may start rewriting policies in response to our findings to make their policies less ambiguous, either to humans or to our algorithms, or to make their policies less explicit. Such a corpus of privacy policies will enable further research on the substantive evolution of privacy policies over time. More generally, we hope that the research and tools developed in this project will positively impact privacy policy decision-making. Our project should yield critical information on the functional capacity of the notice and choice framework to operate effectively with online data gathering. This framework is currently being challenged in the US and Europe in the context of Big Data. Semi-automated analysis can likely assist participants in the policy process – government regulators, stakeholders, advocates and scholars – in identifying strengths and weaknesses in the reliance on web privacy policies. From a regulatory standpoint, these techniques also offer regulators new ways to monitor more effectively the concordance between legal obligations and the stated practices of companies on an industry-wide basis. At present, the task is neither feasible nor manageable. By creating a mechanism to understand large sets of privacy policies, we hope that we can help regulatory oversight to become more effective.

References

[AAB12] Adjerid, Idris, Acquisti, A., Brandimarte, L. "Sleight of Privacy". CIST, 2012.

[ABI+13] Allahbakhsh, M., Benatallah, B., Ignjatovic, A., Motahari-Nezhad, H. R., Bertino, E., Dustdar, S. "Quality Control in Crowdsourcing Systems: Issues and Directions," IEEE Internet Computing, vol. 17, no. 2, pp. 76–81, Mar. 2013.

[Acq09] Acquisti, A. "Nudging Privacy". The Behavioral Economics of Personal Information", IEEE Security and Privacy, Nov.-Dec. 2009.

[AEL02] Adkinson, W., Eisenbach, J., Lenard, T. "Privacy online: A report on the information practices and policies of commercial web sites". Tech report, Progress & Freedom Foundation, 2002.

[AG07] Acquisti, A., Grossklags, J. “What Can Behavioral Economics Teach Us About Privacy?” In Alessandro Acquisti, Sabrina De Capitani di Vimercati, Stefanos Gritzalis, Costas Lambrinoudakis (eds), Digital Privacy: Theory, Technologies and Practices, Auerbach Publications (Taylor and Francis Group), 363-377, 2007.

[AJL12] Acquisti, A., John, L., Loewenstein, G. “The Impact of Relative Judgments on Concern about Privacy”. Journal of Marketing Research, 49(2), 2012.

[Art12] Arthur, C. “Google privacy policy slammed by EU data protection chiefs”. The Guardian, 16 October 2012.

[AWS+12] Ammar, W., Wilson, S., Sadeh, N., Smith, N. “Automatic Categorization of Privacy Policies: A Pilot Study”. School of Computer Science, Language Technology Institute, Technical Report CMU-LTI-12-019, December 2012.

[BA05a] Breaux, T. D., Antón, A. I. “Deriving Semantic Models from Privacy Policies”. 6th IEEE International Workshop on Policies for Distributed Systems & Networks, pp. 67-76, 2005.

[BA05b] Breaux, T. D., Antón, A. I. “Analyzing Goal Semantics for Rights, Permissions and Obligations”. In Proc. IEEE 13th International Requirements Engineering Conference (RE'05), Paris, France pp. 177-186, Aug. 2005.

[BA08] Breaux, T. D., Antón, A. I. “Analyzing Regulatory Rules for Privacy and Security Requirements”. IEEE Transactions on Software Engineering, Special Issue on Software Engineering for Secure Systems (IEEE TSE), 34(1), pp. 5-20, 2008.

[BAD08] Breaux, T. D., Antón, A. I., Doyle, J. “Semantic Parameterization: A Process for Modeling Domain Descriptions”. ACM Transactions on Software Engineering Methodology (ACM TOSEM), 18(2), Article 5, 2008.

[BAL12] Brandimarte, L., Acquisti, A., Loewenstein, G. “Misplaced Confidences: Privacy and the Control Paradox”. Social Psychological and Personality Science, forthcoming, 2012. (Winner, Best Doctoral Student Paper Award, CIST 2010; Runner-up, Best Paper Award, CIST 2010; Leading paper, 2010 Future of Privacy Forum's Best "Privacy Papers for Policy Makers" Competition.)

[BCM03] Baader, F., Calvenese, D., McGuiness, D. (eds.). The Description Logic Handbook: Theory, Implementation and Applications. Cambridge University Press, 2003.

[Bea02] Beales, H. Remarks on the Privacy Notices and the Federal Trade Commission’s 2002 Privacy Agenda. http://www.ftc.gov/speeches/other/privacynotices.shtm, 2002.

[BDM+06] Barth, A., Datta, A., Mitchell, J. C., Nissenbaum, H. “Privacy and Contextual Integrity: Framework and Applications”. IEEE Symposium on Security and Privacy, 2006, pp. 184-198.

[BGS05] Berendt, B., Günther, O., Spiekermann, S. “Privacy in E-Commerce: Stated Preferences vs. Actual Behavior”. Communications of the ACM (CACM) 48 (3): 101-106, 2005.

[BHN+01] Brookman, J., Harvey, S., Newland, E., West, H. “Tracking Compliance and Scope”. W3C Working Draft, October 2012.

[BKS+11] Benisch, M., Kelley, P.G., Sadeh, N., Cranor, L.F. “Capturing Location-Privacy Preferences: Quantifying Accuracy and User-Burden Tradeoffs”. Journal of Personal and Ubiquitous Computing 15(7), October 2011 (published online December 7, 2010).

[BLA+11] Balebako, R., Leon, P. G., Almuhimedi, H., Kelley, P. G., Acquisti, A., Cranor, L. F., Sadeh, N. “Nudging Users towards Privacy on Mobile Devices”. In Proceedings of the Workshop on Persuasion, Nudge, Influence and Coercion, Computer-Human Interaction Conference (CHI), 2011.

[BLM+10] M. S. Bernstein, G. Little, R. C. Miller, B. Hartmann, M. S. Ackerman, D. R. Karger, D. Crowell, and K. Panovich. Soylent: a word processor with a crowd inside. In Proceedings of the 23nd annual ACM symposium on User interface software and technology, UIST ’10, pages 313–322, New York, NY, USA, 2010. ACM.

[Bre09] Breaux, T.D. Legal Requirements Acquisition for the Specification of Legally Compliant Information Systems. Ph.D. Thesis, North Carolina State University, Apr. 2009.

[BVA06] Breaux, T.D., Vail, M.W., Antón, A.I. “Towards Compliance: Extracting Rights and Obligations to Align Requirements with Regulations”. In Proc. IEEE 14th International Requirements Engineering Conference (RE'06), Minneapolis, Minnesota, pp. 49-58, Sep. 2006.

[CA03] Official California Legislative Information. The Online Privacy Protection Act of 2003. California Business and Professional Code §§22575-22579.

[Car98] Carpenter, B. Type-Logical Semantics. MIT Press, 1998.

[Car08] Carpenter, B. “Multilevel Bayesian models of categorical data annotation”. Unpublished manuscript, 2008.

[CDE+06] Cranor, L., Dobbs, B., Egelman, S., Hogben, G., Humphrey, J., Langeinrich, M., Marchiori, M., Presler-Marshall, M., Joseph, R., Schunter, M., Stampley, D.A., Wenning, R. “Platform for Privacy Preferences (P3P) Specification”. W3C Working Group Note, 2006.

[CES+08] Cranor, L. F., Egelman, S., Sheng, S., McDonald, A., Chowdhury, A. “P3P deployment on websites”. Electron. Commer. Rec. Appl. 7(3), pages 274-293, Nov. 2008.

[CGA06] L. Cranor, P. Guduru, and M. Arjula. “User Interfaces for Privacy Agents”. ACM Transactions on Computer-Human Interaction, pp 135-178, June 2006.

[CK12] Chandler, D. and Kapelner, A., “Breaking Monotony with Meaning: Motivation in Crowdsourcing Markets”, University of Chicago Papers in Economics, version of Oct. 2012 - http://www.danachandler.com/files/Chandler_Kapelner_BreakingMonotonyWithMeaning.pdf

[CL03] Camerer, C., Lowenstein, G. “Behavioral economics: Past, present, future”. In Advances in Behavioral Economics, pages 3–51. 2003.

[Co11] S.799, 112th Congress (2011) (Commercial Privacy Bill of Rights Act of 2011)

S. 1223, 112th Congress (2011) (Location Privacy Protection Act of 2012)

H.R. 654, 112th Congress (2011) (Do Not Track Me On- line Act)

H.R. 1528, 112th Congress (2011) (Consumer Privacy Protection Act of 2011)

[Co12] S. 3515, 112th Congress (2012) (Protect America's Privacy Act of 2012)

H.R 5817, 112th Congress (2012) (Eliminate Privacy Notice Confusion Act)

H.R. 6377, 112th Congress (2012) (Mobile Device Privacy Act)

[DCM+13] Das, D., Chen, D., Martins, A., Schneider, N., Smith, N. A. “Frame-Semantic Parsing”. To appear in Computational Linguistics, 2013.

[De68] DeCharms, R. Personal causation: The internal affective determinants of behavior. New York: Academic Press, 1968.

[DLW99] Downs, J. S., Loewenstein, G., Wisdom, J. “Strategies for Promoting Healthier Food Choices”. American Economic Review 99:2, 2009.

[EB08] Eisenstein, J., Barzilay, B. “Bayesian Unsupervised Topic Segmentation.” In Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2008.

[ETC+09] Egelman, S., Tsai, J., Cranor, L., Acquisti, A. “Timing Is Everything? The Effects of Timing and Placement of Online Privacy Indicators”. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2009.

[FTC98] Federal Trade Commision. “Privacy online”. Jun. 1998.

[FTC00] Federal Trade Commision. “Privacy online: Fair information practices in the electronic marketplace”. May 2000.

[FTC10] Federal Trade Commision. “Protecting consumer privacy in an era of rapid change: A proposed framework for business and policymakers”. Technical report, Preliminary FTC Staff Report, Dec. 2010.

[GB12] Gordon, D. G., Breaux, T. D. “Reconciling Multi-Jurisdictional Requirements: A Case Study in Requirements Water Marking”. 20th IEEE International Requirements Engineering Conference, pp. 91-100, 2012.

[GJ02] Gildea, D., Jurafsky, D. “Automatic Labeling of Semantic Roles”. Computational Linguistics 24(3), 2002.

[HBK+07] Hanson, C., Berners-Lee, T., Kagal, L., Sussman, G. J., Weitzner, D. "Data-purpose algebra: modeling data usage policies". 8th IEEE Work. Pol. Dist. Sys. & Nets., 2007, pp. 173-177.

[Hor93] Horty, J. F. "Deontic logic as founded in non-monotonic logic". Annals of Mathematics and Artificial Intelligence, 9: 69-91, 1993.

[JAL11] John, L., Acquisti, A., Loewenstein, G. "Strangers on a Plane: Context-dependent Willingness to Divulge Personal Information". Journal of Consumer Research, 37(5), 858-873, 2011.

[JSJ+07] Jensen, Carlos, Sarkar, C., Jensen, Christian, Potts, C. "Tracking website data-collection and privacy practices with the iWatch web crawler". In Proceedings of the 3rd symposium on Usable privacy and security (SOUPS '07), pages 29-40, New York, NY, USA, 2007.

[Kag04] Kagal, L. A Policy-Based Approach to Governing Autonomous Behavior in Distributed Environments. Ph.D. Thesis, University of Maryland, Baltimore County, Sep. 2004.

[KBC+09] Kelley, P. G., Bresee, J., Cranor, L., and Reeder, R. "A 'Nutrition Label' for Privacy". In Proceedings of the Symposium on Usable Privacy and Security (SOUPS), 2009.

[KC05] Kumaraguru, P., Cranor, L. F. "Privacy Indexes: A Survey of Westin's Studies". Technical Report CMU-ISRI-5-138, Carnegie Mellon University, 2005.

[KCB+10] Kelley, P. G., Cesca, L.J., Bresee, J., Cranor, L. F. "Standardizing Privacy Notices: An Online Study of the Nutrition Label Approach". In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2010.

[KCC+12] Kelley, P. G., Consolvo, S., Cranor, L. F., Jung, J., Sadeh, N., Wetherall, D. "A Conondrum of Permissions: Installing Applications on an Android Smartphone". In Proceedings of USEC2012: Financial Cryptography and Data Security Workshop on Usable Security, March 2012.

[KCN13] Kelley, P., Cranor, L. F., Sadeh, N. Privacy as part of the app decision-making process. Forthcoming: CHI 2013.

[Kle08] Kleimann Communication Group Inc. "Evolution of a prototype financial privacy notice". http://www.ftc.gov/privacy/privacy initiatives/ftcfinalreport060228.pdf. February 2006.

[KLR+09] Kogan, S., Levin, D., Routledge, B. R. R., Sagi, J. S., Smith, N. A. "Predicting Risk from Financial Reports with Regression". In Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics, 2009.

[KSG+11] Komanduri, S., Shay, R., Kelley, P. G., Mazurek, M. L., Bauer, L., Christin, N. Cranor, L. F., Egelman, S. "Of passwords and people: Measuring the effect of password-composition policies". In Proceedings of the SIGCHI Conference on Human Factors in Computing System, May 2011.

[LAH+12] Lin, J., Amini, S., Hong, J., Sadeh, N., Lindqvist, J., Zhang, J. "Expectation and Purpose: Understanding Users' Mental Models of Mobile App Privacy through Crowdsourcing". In Proceedings of the 14th ACM International Conference on Ubiquitous Computing, pp 501-510, Pittsburgh, USA, Sept. 2012.

[LBS+12] Lin, J., Benisch, M., Sadeh N., Niu J., Hong, J.I, Lu, B. and Guo, S. "A Comparative Study of Location-sharing Privacy Preferences in the US and China". Journal of Personal and Ubiquitous Computing, published online October 2012. Also available as CMU CyLab Technical Report CMU-CyLab-12-003, January 2012.

[LCM+10] Leon, P. G., Cranor, L. F., McDonald, A., McGuire, R. "Token attempt: the misrepresentation of website privacy policies through the misuse of P3P compact policy tokens". In Proceedings of the 9th annual ACM workshop on Privacy in the electronic society (WPES '10), pp. 93-104, ACM, New York, NY, USA, 2010.

[LH07] Loewenstein, G., Haisley, E. "The economist as therapist: Methodological ramifications of 'light' paternalism". In Perspectives on the Future of Economics: Positive and Normative Foundations. Oxford University Press, 2007.

[LLS14] B. Liu, J. Lin, N. Sadeh, "Reconciling Mobile App Privacy and Usability on Smartphones: Could User Privacy Profiles Help?", 23rd Interntional Conference on the World Wide Web (WWW2014). Also available as Tech Report CMU-CS-13-128 orCMU-ISR-13-114.

[Lo05] Loewenstein, G. “Hot-cold empathy gaps and medical decision making”. Health Psychology 24(4, Suppl), S49-S56, Jul. 2005.

[LUW+2013] P.G. Leon, B. Ur, Y. Wang, M. Sleeper, R. Balebako, R. Shay, L. Bauer, M. Christodorescu, L.F. Cranor. What Matters to Users? Factors that Affect Users' Willingness to Share Information with Online Advertisers. In Proceedings of the Eight Symposium On Usable Privacy and Security (SOUPS ’13), Newcastle, United Kingdom, 2013.

[May08] May, M. J. Privacy APIs: Formal Models for Analyzing Legal and Privacy Requirements. Ph.D. Thesis, University of Pennsylvania, 2008.

[Moz11] Mozilla. “Mozilla Firefox 4 Beta, now including ‘Do Not Track’ capabilities”. Febuary 2011. http://blog.mozilla.org/blog/2011/02/08/mozilla-firefox-4-beta-now-including-do-not-track-capabilities/

[MSA+11a] Martins, A. F. M., Smith, N. A., Aguiar, P. M. Q., Figueiredo, M. A. T. “Dual Decomposition with Many Overlapping Components”. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2011.

[MSA+11b] Martins, A. F. M., Smith, N. A., Aguiar, P. M. Q., Figueiredo, M. A. T. “Structured Sparsity in Structured Prediction”. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2011.

[MC08] McDonald, A. M., and Cranor, L. F. “The cost of reading privacy policies”. I/S – A Journal of Law and Policy for the Information Society 4(3), 2008.

[McD13] McDonald, A. M. “Browser Wars: A New Sequel?” in The Technology of Privacy, technical report of Silicon Flatirons, University of Colorado Law School, presented Jan. 11, 2013.

[Moz12a] Mozilla. “Introducing Collusion: Discover Who’s Tracking You Online”. https://www.mozilla.org/en-US/collusion/, 2012.

[Moz12b] Mozilla. “Collusion Demo”. https://www.mozilla.org/en-US/collusion/demo/, 2012.

[MSS11] Mugan, J., Sharma, T., Sadeh, N. “Understandable Learning of Privacy Preferences Through Default Personas and Suggestions”. Carnegie Mellon University’s School of Computer Science Technical Report CMU-ISR-11-112, http://reports-archive.adm.cs.cmu.edu/anon/isr2011/CMU-ISR-11-112.pdf, August 2011.

[MW63] Mosteller, F., Wallace, D. L. “Inference and Disputed Authorship: The Federalist”. Journal of the American Statistical Association 58(302), 1963.

[PL08] Pang, B., Lee, L. Opinion Mining and Sentiment Analysis. Foundations and Trends in Information Retrieval, 2008.

[Pon13] Ponemon Institute. “2012 Most Trusted Companies for Privacy”. January 2013. http://www.ponemon.org/local/upload/file/2012%20MTC%20Report%20FINAL.pdf

[PS03] Powers, C., Schunter, M. “Enterprise Policy Authorization Language”. Version 1.2, W3C Member Submission, Nov. 2003.

[Ray10] Raynes-Goldie, K. “Aliases, creeping, and wall-cleaning: Understanding privacy in the age of Facebook”. First Monday 15(1-4), 2010.

[RBK+09] Ravichandran, R., Benisch, M., Kelley, P. G., and Sadeh N. “Capturing Social Networking Privacy Preferences: Can Default Policies Help Alleviate Tradeoffs between Expressiveness and User Burden?” In Proc. 2009 Privacy Enhancing Technologies Symposium, August 2009.

[REI97] Reidenberg, J., “The Use of Technology to Assure Internet Privacy : Adapting Labels and Filters for Data Protection”. LEX ELECTRONICA, III:2, http://www.lex-electronica.org/articles/v3-2/reidenbe.html, 1997.

[RO00] Rabin, M., O’Donoghue, T. “The economics of immediate gratification”. Journal of Behavioral Decision Making, 13(2): 233–250, 2000.

[SC06] Sheng, X., Cranor, L. F. “An Evaluation of the Effectiveness of US Financial Privacy Legislation Through the Analysis of Privacy Policies”. I/S: A Journal of Law and Policy for the Information Society 2(3), pp. 943-979, Fall 2006.

[SCK+13] Sleeper, M., Cranshaw, J., Kelley, P. G., Ur, B., Acquisti, A., Cranor, L. F., Sadeh, N. "I read my Twitter the next morning and was astonished: A conversational perspective on Twitter regrets". To appear in CHI 2013.

[SHC+09] N. Sadeh, J. Hong, L. Cranor, I. Fette, P. Kelley, M. Prabaker, and J. Rao, "Understanding and Capturing People's Privacy Policies in a Mobile Social Networking Application", Journal of Personal and Ubiquitous Computing, Vol. 13, No. 6, August 2009

[Seb02] Sebastianini, F. "Machine Learning in Automated Text Categorization". In ACM Computing Surveys 34(1):1-47, 2002.

[Sim82] Simon, H. A. Models of bounded rationality. MIT Press, Cambridge, MA, 1982.

[Smi11] Smith, N. A. Linguistic Structure Prediction. Morgan & Claypool, 2011.

[SOJN08] Snow, R., O'Connor, B., Jurafsky, D., Ng, A. "Cheap and Fast - But is it Good? Evaluating Non-Expert Annotations for Natural Language Tasks". In Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2008.

[SZ88] Samuelson, W., Zeckhauser, R. "Status Quo Bias in Decision Making". Journal of Risk and Uncertainty 1(1), pages 7-59, Mar. 1988.

[TEC+11] Tsai, J., Egelman, S., Cranor, L., Acquisti, A. "The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study". Information Systems Research, 22, 254-268, 2011.

[ToS12] Terms of Service, Didn't Read project http://tosdr.org/

[TKD+09] Tsai, J., Kelley, P. G., Hankes P., Cranor, L. F., Hong, J., Sadeh, N. "Who's Viewed You? The Impact of Feedback in a Mobile Location Applications". In Proceedings of the 27th annual SIGCHI Conference on Human Factors in Computing Systems (CHI 2009), April 2009.

[TKH+09] Turow, Joseph, King, Jennifer, Hoofnagle, Chris, Jay, Bleakley, Amy and Hennessy, Michael. "Americans Reject Tailored Advertising and Three Activities that Enable It". Available at SSRN: http://ssrn.com/abstract=1478214 or http://dx.doi.org/10.2139/ssrn.1478214, 2009.

[TS08] Thaler, R. H., Sunstein, C. R. Nudge: Improving Decisions About Health, Wealth, and Happiness. Yale University Press, March 2008.

[UBL+08] Uszok, A., Bradshaw, J. M., Lott, J., Breedy, M., Bunch, L. "New Developments in Ontology-Based Policy Management: Increasing the Practicality and Comprehensiveness of KAoS". Proc of the IEEE Work. on Pol. Dist. Sys. & Nets., pp. 145-152, 2008.

[Wh59] White, R. W. "Motivation reconsidered: The concept of competence". Psychology Review 66, pages 297-333, Sep. 1959.

[Wiki12] Wikimedia Traffic Analysis Report – Browsers (Oct 1, 2012 – Oct. 31, 2012) http://stats.wikimedia.org/archive/squid_reports/2012-10/SquidReportClients.htm

[WKL+11] Wang, Y., Komanduri, S., Leon, P. G., Norcie, G., Acquisti, A., Cranor, L. F. "I regretted the minute I pressed share: A Qualitative Study of Regrets on Facebook". In Proceedings of the 7th Symposium on Usable Privacy and Security (SOUPS2011).

[YHO+11] Yogatama, D., Heilman, M., O'Connor, B., Dyer, C., Routledge, B. R. R., Smith, N. A. "Predicting a Scientific Community's Response to an Article". In Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2011.

[YSW12] Yano, T., Smith, N. A., Wilkerson, J. "Textual Predictors of Bill Survival in Congressional Committees". In Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics, 2012.

[ZM96] Zelle, J. M., Mooney, R. J. "Learning to Parse Database Queries Using Inductive Logic Programming". In Proceedings of the National Conference on Artificial Intelligence, 1996.