The Seventh AAAI Conference on Human
Computation and Crowdsourcing (HCOMP-19)

Going against the (Appropriate) Flow: A Contextual Integrity Approach to Privacy Policy Analysis

Yan Shvartzshnaider,[*1,2] **Noah Apthorpe,**[*2] **Nick Feamster,**[3] **Helen Nissenbaum**[4]
[1]New York University, [2]Princeton University, [3]University of Chicago, [4]Cornell Tech
yansh@nyu.edu, apthorpe@cs.princeton.edu, feamster@uchicago.edu, helen.nissenbaum@cornell.edu

Abstract

We present a method for analyzing privacy policies using the framework of contextual integrity (CI). This method allows for the systematized detection of issues with privacy policy statements that hinder readers' ability to understand and evaluate company data collection practices. These issues include missing contextual details, vague language, and overwhelming possible interpretations of described information transfers. We demonstrate this method in two different settings. First, we compare versions of Facebook's privacy policy from before and after the Cambridge Analytica scandal. Our analysis indicates that the updated policy still contains fundamental ambiguities that limit readers' comprehension of Facebook's data collection practices. Second, we successfully crowdsourced CI annotations of 48 excerpts of privacy policies from 17 companies with 141 crowdworkers. This indicates that regular users are able to reliably identify contextual information in privacy policy statements and that crowdsourcing can help scale our CI analysis method to a larger number of privacy policy statements.

1 Introduction

Federal and state regulations require online services to notify consumers about information collection and sharing practices through privacy policies (Federal Trade Commission 1998). Researchers have demonstrated that privacy policies are vague and often incomplete, contributing to "misunderstanding among stakeholders, wherein stakeholders have different interpretations regarding the incomplete information" (Bhatia and Breaux 2018; Bhatia et al. 2016a). In this paper, we use the theory of contextual integrity (CI) (Nissenbaum 2010) to synthesize these existing privacy policy evaluation methods and provide a new formal approach for detecting specific types of ambiguities that interfere with readers' ability to understand the information collection practices described in privacy policies.

Our CI-based analysis method (Section 3) involves identifying and annotating contextual parameters of information flows described in privacy policies, specifically the *senders, recipients* and *subjects* of information, information types *(attributes)*, and the conditions under which information may be transferred or collected (*transmission principles*). The resulting annotations allow descriptive and normative analyses based on a core principle of contextual integrity: Understanding and assessing the privacy implications of an information flow requires knowing the full context of the flow (i.e., all five contextual parameters). This assertion allows us to evaluate privacy policies for specific issues that hinder understandability, including information flow incompleteness, parameter bloating, and vagueness.

Incomplete information flows, which omit one or more contextual parameters, invite readers to interpret the missing parameters according to their own expectations, which may not match the actual practices of the company (Bhatia and Breaux 2018; Martin and Nissenbaum 2016). Parameter bloating, or specifying more than one instance of a contextual parameter, increases the cognitive load required for readers to decipher which combinations of five parameters define fully-specified information flows that are actually allowed by the policy (Micheti, Burkell, and Steeves 2010). Finally, vague information flows contain language that makes it unclear which actors share the information or under what conditions the data collection practice described by the flow actually takes place.

Analyzing privacy policies on the basis of a consistent set of CI parameters also allows for seamless and rigorous comparison between policy versions and across many policies from different companies. Finally, the use of CI ties our method to an existing body of research using CI for descriptive and normative analyses of privacy implications in other settings (Apthorpe et al. 2018; Guinchard 2017; Hull, Lipford, and Latulipe 2011; Shvartzshnaider et al. 2016; Wijesekera et al. 2015; Zimmer 2008).

We present a method for annotating privacy policies using the contextual integrity framework (Section 3). The use of a structured framework allows rigorous analysis of difficult information policy statements and is applicable to policies across companies and sectors.

We demonstrate a range of analytical methods enabled by our approach through two applications: a comparative analysis of Facebook privacy policy updates (Section 4) and crowdsourced annotations of 48 privacy policy ex-

[*] These authors contributed equally to this work

cerpts (Section 5).

To support future research and policymaking efforts, we have made the privacy policy annotations performed for this work publicly available for the wider community.[1]

2 Related Work

Prior efforts by the research community have analyzed privacy policies in order to identify statements that are uninformative or potentially confusing to the reader. These works fall into two main categories: 1) Detecting textual ambiguity and vagueness in privacy policies, and 2) Privacy policy annotations.

Ambiguity and Vagueness. Bhatia et al. (2016a) proposed a formal "theory of vagueness for privacy policy statements based on a taxonomy of vague terms" to show that statements with vague language affect readers' perceptions of privacy risk from the described data collection practices. More recent work by Bhatia and Breaux (2018) used frame semantics (Fillmore 1976) to identify incomplete privacy statements that omit relevant contextual information. Textual ambiguity in privacy policies has also been the focus of work performing lexical analysis to extract hypernyms, meronyms, and synonyms in information type descriptions (Bhatia et al. 2016b; Evans et al. 2017; Hosseini et al. 2016). These projects have aimed to build a concise ontology of information types described in privacy policies.

Our CI-based analysis benefits from these insights; however, we capture a more complete picture of data collection practices described in privacy policies including and beyond issues of textual ambiguity. We are able to evaluate privacy policy statements with respect to a broader space of issues that make it difficult for readers to assess whether the practices being described respect or violate privacy norms.

Using CI to analyze privacy policies is also supported by recent work showing the importance of contextual factors to users' privacy expectations. Rao et al. (2016) compared users' privacy expectations to existing companies' practices to show that users' privacy expectations depend on website types and the types of information being exchanged. Martin and Nissenbaum (2016) showed that when confronted with a privacy-related scenario that was missing some contextual information, respondents mentally supplemented the information, essentially generating a different version of the scenario. Specifically, the "context of information exchange – how information is used and transmitted, the sender and receiver of the information – all impact the privacy expectations of individuals." Bhatia and Breaux (2018) reported similar results: Specifically, users' willingness to share information significantly increased with addition of statements describing the purpose and provision of choice.

Privacy Policy Annotations. Wilson et al. (2016a) recruited law students to hand-annotate privacy policies with metadata tags such as "first party collection/use," "user choice/control," "data retention," and "data security." They then used the hand-labeled policies to train a machine learning algorithm for annotating policies with the same tags. This labelling taxonomy was used in more recent work (Harkous et al. 2018) to train a neural network classifier to automatically annotate segments of privacy policies and to build a Question-Answering system that supports free-form querying of the privacy policy content. Wilson et al.; Wilson et al. (2016b; 2018) also showed that the answers of the crowdworkers agreed with those of skilled annotators over 80% of the time, indicating that crowdsourcing can be used to identify paragraphs describing specific practices in privacy policies.

These existing techniques (Harkous et al. 2018; Wilson et al. 2016a; 2016b; 2018) have aimed to make users aware of a wide variety of information handling practices, such as third party data collection and data retention. In contrast, we use CI to annotate five information flow parameters, rather than a large labelling taxonomy. This allows us to directly evaluate privacy policies for specific properties, such as excessive or missing details from a CI perspective that are difficult to detect using previous annotation methods.

3 CI Analysis Method

We use the framework provided by CI to identify and annotate information flows and their component parameters described in privacy policy statements.

3.1 CI Overview

In contrast to other theories of privacy, Contextual Integrity (CI) defines privacy as the appropriateness of information flows determined by conformance with existing legitimate, informational norms specific to given social contexts (Nissenbaum 2010). In other words, a person's privacy is prima facie violated when a transfer of information deviates from established norms in a particular context. For example, someone might view sharing Fitbit data with their doctor as appropriate but sharing the same data with an insurance company as a privacy violation. Changing the recipient of the information alters the flow, and as a consequence, could violate a contextual norm. The sources of these contextual norms can vary, ranging from law and regulation to societal beliefs and family values.

To facilitate analysis, CI offers a framework to describe information flows using 5-parameter tuples. These five parameters capture specific actors (*senders*, *recipients*, and *subjects*) involved in an information flow, the type (*attribute*) of information in the flow, and the condition (*transmission principle*) under which the information flow occurs. *Importantly, all five parameters must be specified in order to understand the context of an information flow, and changing even one parameter can affect a flow's overall appropriateness*. This is a central premise of CI; without stating all parameters characterizing an information flow, its context is underspecified and its implications are ambiguous.

3.2 Privacy Policy Annotation

We use the following definitions to identify and label information flows and contextual parameters in privacy policy text. These annotations are the raw data for the analyses

[1] https://ci-annotations-project.github.io

described in the following section. Annotation can be performed manually or formulated as crowdworking task for scalable application of the CI analysis method.

Information Flow. Any self-contained description of a transfer of information. Information flows are typically single sentences or short paragraphs, but are also presented as bulleted lists in some privacy policy formats.

Sender. Any entity (person, company, website, device, etc.) that transfers or shares information. This may be a pronoun or a specific entity, such as "Company A," "strategic partners," or "publisher."

Recipient. Any entity (person, company, website, device, etc.) that ultimately receives information. This may be a pronoun or a specific entity, such as "third party," "developer," "other users," or "Company B and its affiliates."

Transmission principle. Any clause describing the "terms and conditions under which [...] transfers ought (or ought not) to occur" (Nissenbaum 2010). This includes descriptions of how information may be used or collected. Examples include "if the user gives consent," "when an update occurs," or "to perform specified functions."

Attribute. Any description of information type, instance, and/or example, such as "date of birth," "credit card number," "photos," or, more generally, "personal information."

Subject. Any subjects of information exchanged in a flow. Subjects may be explicitly stated or implicitly described using pronouns and possessives. For example, the following annotated statement from the Facebook privacy policy describes a single information flow:

> **[We (Facebook)]**^{recipient} *also collect* **[contact information]**^{attribute} *that* **[you]**^{sender} *provide* **[if you upload, sync or import this information (such as an address book) from a device].**^{TP}

This flow contains an explicit sender, recipient, attribute, and transmission principle (TP). The subject parameter is not included, but is implicitly the user agreeing to the privacy policy.

3.3 Information Flow & Parameter Analyses

We can use annotated information flows and parameters in privacy policy texts for a variety of analyses, including, but not limited to, the following.

Comparing Privacy Policy Versions. We can compare snapshots of a privacy policy across updates or get an aggregated view across different privacy policies. This offers insights into the general nature of the policy differences, including which parameters were preferentially added, removed, or modified.

Identifying Incomplete Flows. In order to understand the privacy implications of an information flow, it is important to provide a complete description with all five contextual parameters specified. Otherwise, consumers are left uninformed about company behavior (Martin and Nissenbaum 2016). Identifying privacy statements that underspecify information flows can reveal problematic sections of privacy policies.

Diagnosing Vague Statements. The use of vague and ambiguous terminology in privacy policy statements makes it increasingly difficult for readers to reason about information flow appropriateness and privacy implications. Building on prior work (Bhatia et al. 2016a; Reidenberg et al. 2016), we can use CI annotations to identify specific privacy statements that describe such ambiguous flows. This also makes it easier for regulators and policymakers to monitor the appearance of such statements across privacy policy updates and privacy policies from different companies.

Recognizing CI Parameter Bloating. CI parameter bloating occurs when a single information flow contains two or more semantically different CI parameters of the same type (e.g., two senders or four attributes) without a clear indication of how these parameter instances are related to each other. This creates an information flow with a combinatorial number of possible contexts. It is difficult for readers or regulators to determine which combinations of parameters describe contexts in which information flows actually take place. Previous research indicates that "eliminating connectives that clarify the relationship between ideas makes sentences harder to understand because readers are left to infer the relationship" (Micheti, Burkell, and Steeves 2010). CI parameter bloating is a specific example of this phenomenon.

4 Detecting Privacy Policy Ambiguities

Revelations about the misuse of consumer data by Facebook and Cambridge Analytica (Frier 2018) rekindled the debate around users' privacy and informed consent on such platforms. In response to public outcry, Facebook worked to rectify the situation by updating its privacy policy (data policy) on April 19, 2018.

We apply our CI analysis technique to the Facebook privacy policy from immediately before and after this update. We used the Brat rapid annotation tool,[2] and the annotation guidelines in Section 3 to manually annotate information flows and CI parameters in the previous and updated policy versions. Two of the authors separately annotated both versions of the policy and performed statement by statement comparison to produce the final annotation.

From a legal perspective, the new document discloses more about the company's information sharing practices. However, our CI analysis method reveals fundamental ambiguity issues present in both versions. These issues prevent users from interpreting new details in the updated version to fully understand how their data is being collected and shared. Of course, Facebook's privacy policy was unlikely written with contextual integrity in mind. We therefore intend the following analysis not as a criticism of Facebook per se, but as an opportunity to demonstrate our method and to point out issues common across privacy policies from many companies.

4.1 Information Flow Updates

We used our CI annotations to compare numbers (Figure 1) and specifics of each information flow parameter described in the previous and updated Facebook privacy policies. The

[2] Brat Rapid Annotation Tool. http://brat.nlplab.org.

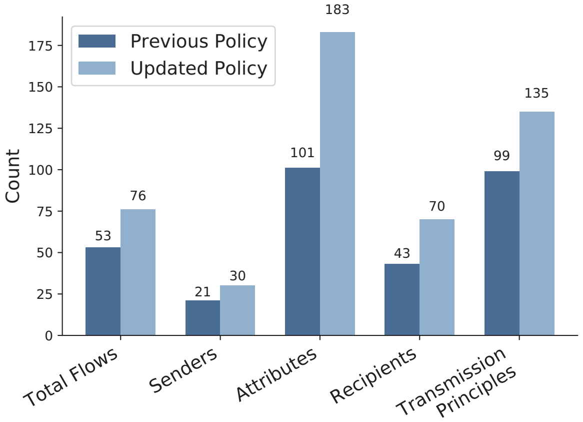

Figure 1: Distribution of unique CI parameters identified in the previous and updated Facebook privacy policies.

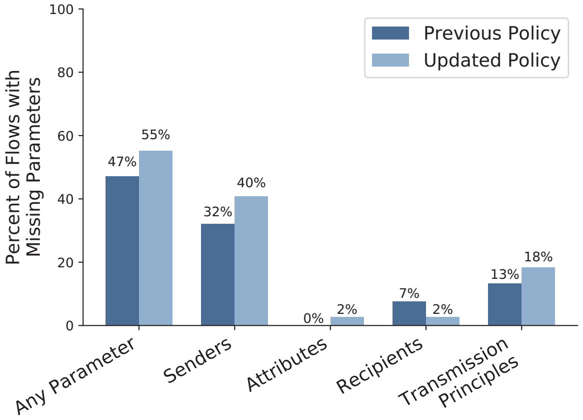

Figure 2: Percentage of incomplete information flows in Facebook's previous and updated privacy policies with missing CI parameters.

updated Facebook privacy policy has about 50% more information flows than the previous policy (Figure 1). However, more information flows does not necessarily equal less confusion. Our analysis shows that many of the newly introduced information flows are incomplete (Section 4.2), are overloaded with CI parameters (Section 4.3) and/or use vague language (Section 4.4).

4.2 Incomplete Information Flows

Our analysis of the Facebook privacy policy versions finds many described information flows with missing (non-specified) parameters (Figure 2).

In the previous privacy policy, 47% (25/53) of flows are missing one or more parameters. In the updated policy, this number increases to 55% (42/76), including 16 incomplete flows from the previous policy and 27 new incomplete flows.

Missing Recipient. The previous policy has three flows without an explicit recipient while the updated policy has two. Not stating information recipients forces users to infer what entities will have access to their information from other sources, often leading to incorrect notions of company behavior (Turow, Hennessy, and Draper 2018; Martin and Nissenbaum 2016).

Missing Sender. The sender parameter is not specified in 17 (32%) flows in the previous policy nor in 31 (40%) flows in the updated policy. Many of the statements with missing senders describe "use-of-data," i.e., they inform the consumer how the collected information will be used but not from where it is collected. Missing senders can easily lead to misinterpretations and false privacy expectations. For example, the source of the information in the following statement is unclear: *"We collect information about the people, Pages, accounts, hashtags and groups you are connected to and how you interact with them."* Without knowing which of Facebook's various services collect and send this information, users are unable to take specific action to avoid this data collection or adjust their behavior on the platform.

Missing Transmission Principle. We identified 7 information flows in the previous policy where the transmission principle is missing. For example, the statement "*We share information we have about you within the family of companies that are part of Facebook*" does not specify under what conditions/constraints the information is being shared. Previous research (Martin and Nissenbaum 2016) shows that in these instances consumers will end up guessing when and for what reason information is collected.

The updated policy contains even more (14) flows with missing transmission principles. Without a transmission principle, flows like "*We also receive information about your online and offline actions and purchases from third-party data providers who have the rights to provide us with your information"* become ambiguous because it is not clear when or why this information is being collected.

4.3 CI Parameter Bloating

We used our annotations of the Facebook policy versions to identify information flows that suffer from CI parameter bloating, including the flow in Figure 3. At first glance, this statement seems transparent and informative. It explicitly specifies the type of information that is being exchanged, among what actors (sender, recipient, subject) and under what conditions. However, this is an example of CI parameter bloating. Taking into account all the possible permutations results in total of 3 (senders) × 1 (subject) × 6 (attributes) × 1 (recipient) × 7 (TPs) = 126 possible flows.

How should the consumer reason about this privacy policy statement? Do all listed senders transfer all of these information types to Facebook or does each particular sender transmit a specific information type? Do flows with each sender/attribute pair occur under each listed TP or only specific ones? Even technically-savvy users will have difficulty reasoning about the many possible information flows with all combinations of each parameter type.

Our CI annotation analysis identifies several statements in both previous and updated policies that suffer from parameter bloating. The previous policy has 15 statements (28% of all flows) with multiple instances of two or more CI param-

[Advertisers, app developers and publishers]^{senders} can send **[us]**^{recipient} information **[through Facebook Business Tools that they use, including our social plug-ins (such as the Like button), Facebook Login, our APIs and SDKs or the Facebook pixel]**^{TP}. These partners provide information about **[your]**^{subject} **[activities off Facebook including information about your device, websites you visit, purchases you make, the ads you see and how you use their services]**^{attributes}**[whether or not you have a Facebook account or are logged in to Facebook]**^{TP}.

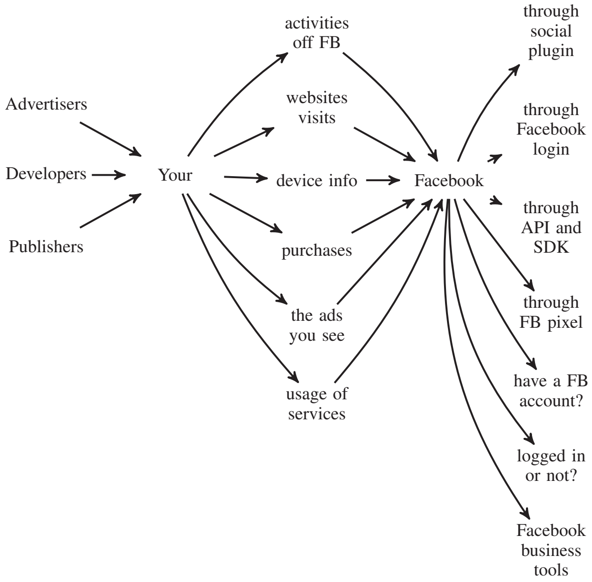

Figure 3: Example of CI parameter bloating in privacy policy text (*top*) and mapped into possible interpretations (*bottom*).

eters. These statements have up to 4 senders, 20 attributes, 10 recipients, and 7 transmission principles and describe 4 to 180 total information flow permutations each (Figure 4). The updated policy has 30 statements (39% of all flows) with multiple instances of two or more CI parameters. These statements have up to 7 senders, 41 attributes, 8 recipients, and 8 transmission principles and describe 4 to 492 total information flow permutations each (Figure 4).

Given that an average consumer today spends little to no time reading privacy policies (Schaub, Balebako, and Cranor 2017), it is unreasonable to assume that the even the most privacy-concerned citizen will dissect all possible combinations of this many multi-parameter flows. Instead, we believe that privacy policies should list all prescribed information flows explicitly, with each including all five parameters. This will increase the length of the policy and might initially be construed to decrease readability. However, adopting a regular 5-tuple structure for all policy statements will increase machine interpretability and allow user interfaces that can provide "different notices for different audiences" (Schaub et al. 2015) by automatically parsing, filtering, and categorizing privacy policy statements.

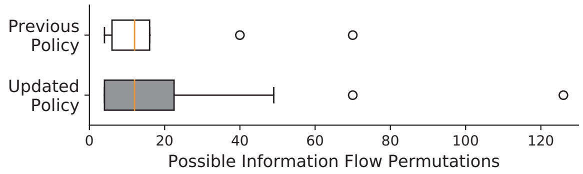

Figure 4: Extent of CI parameter bloating in privacy policy statements with multiple instances of at least two different CI parameters. *Not shown:* one outlier flow with 180 possible permutations in the previous policy and one outlier flow 492 possible permutations in the updated policy.

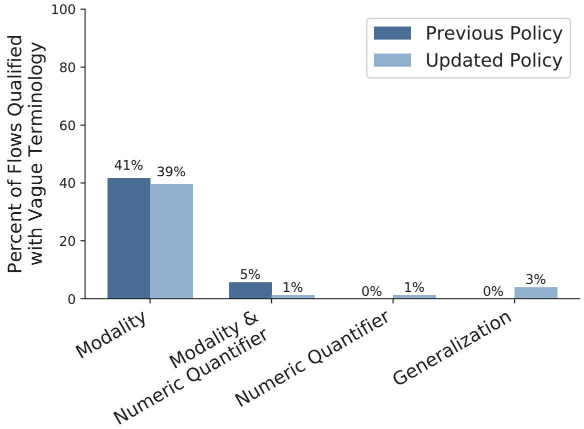

Figure 5: Percentage of information flows in Facebook's previous and updated privacy policies qualified with various categories of vague terminology.

4.4 Vague Information Flows

We used the annotations to identify flows which are prescribed using one or more combinations of vague terms from the vague terms taxonomy defined by (Bhatia et al. 2016a). As discussed in Section 2, vague information flows affect readers' ability to accurately interpret whether the described data collection practice violates or respects their privacy. Figure 5 shows the percentage of vague information flows in Facebook's previous and updated policies. In both policies, "modality" vagueness dominates, occurring in close to 40% of all flows. The updated policy does not represent a reduction in vague terminology from the previous version. Rather, the percentage of flows with vague terminology remains the same. This supports our initial claim that the updated policy does not contribute to clarity. The widespread occurrence of flows qualified by vague terminology further supports the problem that privacy policies are too often "obtuse and noncommittal [and] make it difficult for people to know what information a site collects and how it will be used" (Turow, Hennessy, and Bleakley 2008).

5 Crowdsourcing CI Privacy Policy Analysis

We also test our method to see whether crowdworkers are able to identify CI parameters in privacy policy statements. Specifically, we created an Amazon Mechanical Turk (AMT) Human Intelligence Task (HIT) to annotate 48 privacy policy excerpts. These included 16 excerpts from the Google privacy policy circa October 2017 and 16 pairs of excerpts from the privacy policies of 16 well-known companies[3] before and after May 2018 updates. These excerpt pairs describe information flows with differences in parameters between the versions. The excerpts are also self-contained and do not require additional information from the policy to correctly annotate. The excerpts range from 21 to 113 words[4] and from 1 to 4 sentences for a total of 2621 words over 103 sentences.

We compared aggregated crowdworker annotations to ground-truth annotations from the authors (Section 5.4). The crowdworker annotations had an average precision of 0.96 across CI information flow parameters, indicating that the crowdworkers understood the relatively complex notion of information flow parameters and were able to correctly identify them in real privacy policy text. These results show that crowdworking can be an effectual tool for scaling CI annotation. We have made the crowdworker annotations available as a public dataset for future research.[5]

5.1 Annotation Task Design

We developed the annotation task as a Qualtrics[6] survey deployed on AMT. The task was designed to optimize annotation accuracy while minimizing cost.

Consent and Instructions. The first page of the annotation task was a consent form. Participants who did not consent were prevented from proceeding. The annotation task collected no personal information about crowdworkers and was approved by our university's Institutional Review Board. The task next presented annotation instructions, including a description of each information flow parameter that should be annotated (sender, attribute, recipient, and transmission principle) and an example annotated flow. The information flow parameter descriptions matched those described in Section 3.2.

Screening Questions. Each crowdworker was asked to annotate (highlight and label) all words and phrases corresponding to CI information flow parameters in three privacy policy excerpts. These excerpts served as screening questions to identify workers who are able to perform high-accuracy annotations. Workers whose annotations had an F_1 score of at least 0.7 compared to ground-truth expert annotations on the first screening question (for which the correct answer was given) and either of the next two screening questions were allowed to proceed with the task.

Annotations. Each worker who passed the screening questions was asked to annotate 5 excerpts selected randomly from the 48 excerpts of interest. Annotations from multiple workers were collected, analyzed, and processed into the final crowdsourced annotation for each privacy policy excerpt (Section 5.3).

5.2 Task Deployment

We deployed the annotation task as a HIT on AMT using TurkPrime (Litman, Robinson, and Abberbock 2017), an online tool for researchers to easily manage AMT tasks. We limited the HIT to AMT workers in the United States with a HIT approval rating of 90–100% and at least 100 HITs approved. We did not collect or place any other criteria on the demographics or technical background of the AMT workers. 141 total workers accepted the HIT. Of these workers, 99 passed the screener questions. All 48 excerpts were annotated by between 7 and 12 workers (mean 10.2). AMT workers who did not pass the screening questions were automatically reimbursed \$0.25. AMT workers who passed the screening test and completed the entire annotation task were reimbursed \$1.50. Collecting all responses took approximately 4 hours from HIT launch until completion and cost a total of \$198 (including AMT fees).

5.3 Majority Vote Annotations

We were ultimately interested in acquiring the single highest-accuracy annotation for each privacy policy excerpt independent of individual workers. We therefore combined multiple annotations of each privacy policy excerpt into a "majority vote" annotation, which assigned each word in an excerpt to the CI parameter annotated by at least 50% of the workers presented with that excerpt. If fewer than 50% of workers labeled a word with the same parameter, then the word is given no label in the majority vote annotation.

The majority vote method reduced the influence of unreliable or adversarial crowdworkers who passed the screening questions. Assuming that such crowdworkers were a minority of those assigned to an excerpt, their annotations (or lack thereof) did not affect the final annotation.

5.4 Crowdworker Annotation Accuracy

Two of the authors annotated all excerpts prior to seeing the crowdworker results. These authors compared their independent annotations and manually resolved minor differences to create a single set of ground truth expert annotations.

These authors then found all discrepancies between the crowdworker and expert annotations and divided them into six categories: correct parameters, skipped parameters, ambiguous parameters, overlapping parameters, true errors, and expert errors (Figure 6, Section 5.5). This comparison was performed manually to ensure accuracy and avoid the need for string matching heuristics. Categorizing the discrepancies allowed us to count the number of true positives (correct parameters), false negatives (skipped parameters), and false positives (true errors) and compute precision, recall, and F_1 scores[7] for the crowdworker annotations (Table 1).

[3] Amazon, Fitbit, Indiegogo, LinkedIn, The New York Times, Microsoft, Shapeways, Slack, Spotify, Steam, Stripe, Tinder, Twitter, Uber, WhatsApp, Yelp

[4] Mean: 55 words/excerpt, SD: 23 words/excerpt

[5] https://ci-annotations-project.github.io

[6] www.qualtrics.com

[7] Precision = $\frac{TP}{TP+FP}$, Recall = $\frac{TP}{TP+FN}$, $F_1 = 2 \cdot \frac{\text{Precision}\cdot\text{Recall}}{\text{Precision}+\text{Recall}}$

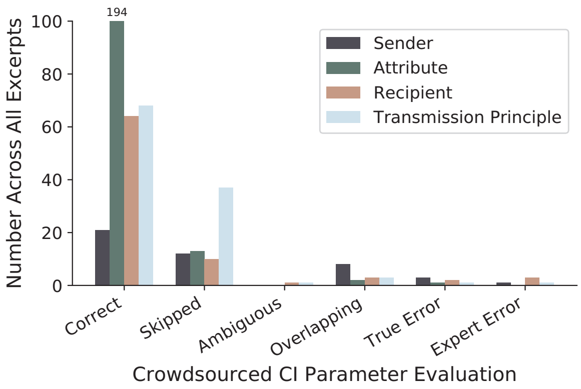

Figure 6: Comparison of crowdworker majority vote annotations to expert ground truth. Correct parameters are labeled in both annotations. Skipped parameters are only labeled by the expert. All other categories are described in Section 5.5.

	Precision	**Recall**
Attribute	0.99	0.94
Sender	0.88	0.64
Recipient	0.97	0.86
TP	0.99	0.65

Table 1: Precision and recall scores of crowdworker majority vote annotations for each CI parameter across all excerpts.

Overall, the high precision of the majority vote crowdworker annotations indicates that the majority of crowdworkers understood the CI annotation task and were able to correctly identify and highlight CI parameters in short privacy policy excerpts. A closer look at the flows where the majority of crowdworkers missed some parameters (Section 5.5) provides interesting insight into the reasons for the moderately lower recall numbers.

5.5 Evaluating Annotation Discrepancies

Analyzing the crowdworker annotations raised the question "What causes particular excerpts or CI parameters to be more difficult for crowdworkers to annotate than others?" We evaluated the discrepancies between crowdworker and expert annotations to better understand their underlying causes.

Ambiguity. The annotated excerpts include the various types of ambiguities found in the Facebook privacy policy evaluation (Section 3). 32 excerpts describe incomplete information flows, 20 excerpts describe bloated information flows, and 27 excerpts include vague language. We used the Mann-Whitney U test to compare excerpts with and without incomplete information flows, parameter bloating, and vague language. We found no significant difference in F_1 scores based on these conditions ($p > 0.05$).

This supports using crowdworking to scale CI analysis of privacy policies, because it indicates that crowdworkers can identify individual CI parameters even in privacy policy excerpts with semantic ambiguities that hinder interpretation of complete information flows, allowing post-annotation analysis to detect and evaluate these ambiguities.

Readability. We calculated Spearman correlations of the crowdworker majority vote annotation F_1 scores for each excerpt versus word count, Flesch-Kincaid Reading Ease (Kincaid et al. 1975), FOG Index, and number of CI parameters. However, all of the resulting correlation coefficients had absolute values < 0.5 and $p \gg 0.05$, indicating no significant correlations with F_1 score. This suggests that crowdworker difficulties with annotating certain excerpts were due to more nuanced factors than length or readability, which we explore by looking at each category of discrepancy in more detail.

Skipped Parameters. The most common type of discrepancy occurred when the crowdworkers simply neglected to annotate some or all instances of a given parameter. These discrepancies were the primary contributor to lowering recall scores without affecting precision.

The skipped parameters offer a glimpse into how a majority of the crowdworkers interpret the privacy policy excerpts. For example, we noted that the majority did not annotate a sender in the information flow beginning with *"We may display your Profile name. . . "* presumably because they don't see an "act of displaying" as sharing information. Additionally, in the information flow *"We collect information when you sync non-content like your email address book, mobile device contacts, or calendar with your account,"* both the expert and the crowdworkers labeled "email address book," "mobile device contacts," and "calendar" as attributes. However, the expert also labeled "information" as an attribute, while the majority of crowdworkers did not. From the CI analysis perspective, it is important to label "information" as an attribute because it acts as a superset, while the provided examples are merely selected instances. This is another type of privacy policy ambiguity that we would like to investigate in future work. Alternatively, the crowdworkers may have found a few instances of each parameter and then moved on to the next excerpt without double-checking to ensure that none were missed. The crowdworkers may also have intentionally skipped parameters. This could be due to cognitive fatigue or the fact that crowdworkers are incentivized to finish the annotations as quickly as possible to optimize their hourly compensation rate.

Ambiguous Parameters. Ambiguous parameter discrepancies occurred when a CI parameter was mislabeled compared to the expert annotation, but the correct labeling is ultimately open to interpretation. Consider the sentence *"If you want to take full advantage of the sharing features we offer, we might also ask you to create a publicly visible Google Profile, which may include your name and photo."* In this sentence, "publicly" could be interpreted as a recipient, i.e. the public would receive the data in the Google Profile. However, "publicly" could also be interpreted as a transmission principle i.e. the flow is from "you" to your "Google Profile" and the condition on the flow is that it is public. The expert labeled "publicly" as a recipient, while the crowdworker majority did not. We only identified 2 such ambiguous parameter discrepancies, indicating that CI information flow descriptions map naturally to privacy policy texts.

Overlapping Parameters. Overlapping parameter discrepancies occurred when a CI parameter was mislabeled compared to the expert annotation, but the text in question is part of two or more CI parameters simultaneously. We identified 16 overlapping parameters. Consider the excerpt *"When you use our services or view content provided by Google, we automatically collect and store certain information in server logs."* The first clause (before the comma) could be interpreted as a single transmission principle, but the "you" could also be a sender. Variations on this issue were the primary cause of discrepancies for the "sender" parameter, i.e. the expert annotated an entire clause as a transmission principle but the majority vote annotation instead labeled a single word in the clause as a sender. The presence of overlapping parameter discrepancies is due to a tradeoff in our implementation of the CI annotation task. We chose to allow only one CI parameter annotation per word in each excerpt to simplify the task for workers.

True Errors. True errors occurred when the crowdworkers unambiguously misannotated a CI parameter. Fortunately, we only observed 7 true errors across all annotations. This implies that when a label made it into the majority vote annotation (with sufficient workers contributing to the vote), it was most likely correct. The low frequency of true errors indicates that, with improvements to reduce the number of skipped parameters, crowdworking can be a high-accuracy method of obtaining CI annotations of privacy policies.

Expert Errors. Finally, we identified 5 cases where the crowdworker majority vote annotation was correct while the "ground-truth" expert annotation was incorrect. Most of these cases were due to the expert annotation missing a one-word sender or recipient, e.g. "we." We did not adjust recall or precision scores to reflect the incorrect expert annotations, as these judgments were made after, and could have been influenced by, viewing the crowdworker annotations.

6 Discussion & Future Work

In this paper, we argue that the notion of information flow appropriateness in the CI framework lends itself well to data collection practices described in privacy policies. Requiring that privacy policies have distinct five-parameter information flow descriptions for all data collection practices would complement ongoing efforts to improve interpretability of privacy policies, move towards an efficient auditing of devices and services, and understand how privacy policies relate to societal privacy norms.

6.1 Auditing Privacy Policies

The FTC and other regulatory bodies recommend that privacy policies include specific components, including the type of information collected, the entities that receive or store the information, uses of the information, and the conditions governing data acquisition and handling (Federal Trade Commission 2012). Our CI analysis method would enable a scalable auditing technique to check whether such requirements on the information flow descriptions in privacy policies are followed. The CI analysis method would also simplify continued auditing of privacy policies across updates by only requiring annotation of the differences between versions rather than each version in its entirety. This would indicate the CI parameter and information flow changes between versions, providing enough information for detecting ambiguous flows while requiring minimal annotation overhead.

6.2 Comparing Privacy Policies to Norms

Our analysis method adopts the notions of contextual integrity. On one hand, privacy policy statements made by a company should be compliant with existing regulation and legal statues. On the other hand, they need to be informed by the context in which they operate. In other words, company privacy practices should not only be about legal compliance but also about respecting users' privacy expectations and societal privacy norms. This challenge is particularly relevant to modern technosocial systems and platforms that operate in a myriad of social contexts.

Fortunately, the research community has already taken steps towards addressing this challenge that can be furthered by our CI privacy policy analysis method. Previous efforts (Apthorpe et al. 2018; Shvartzshnaider et al. 2016) have used the CI framework as a practical tool to discover privacy norms. These methods could be combined with CI annotations to determine whether the practices described in privacy policies align with users' privacy expectations and societal norms. This combination of CI annotations and survey data could inform company behavior, as data collection practices aligned with user norms are less likely to cause consumer backlash. It could also enable longitudinal ethnographic insight into how user norms are changing vis-a-vis privacy policies over time.

7 Conclusion

We present a privacy policy analysis method, based on the theory of contextual integrity, for detecting specific ways that privacy policies make it difficult for readers to assess whether the described practices respect or violate privacy norms (Section 3).

We demonstrated the utility of the method in two settings: First, we analyzed versions of Facebook's privacy policy from before and after the Cambridge Analytica incident in April 2018 (Section 4). Our analysis shows that the updated policy describes more information flows than the previous version, but that the updates do not improve the percentage of flows that contain vague language, omit parameters, or allow for many possible interpretations by including several parameters of the same type. Second, we showed that non-expert users can help scale the CI analysis method by successfully crowdsourcing annotations of 48 privacy policy excerpts from 17 companies (Section 5). In summary, our method complements existing privacy policy research and offers a new, scalable, approach to help study and protect user privacy.

Acknowledgements

This work is supported by NSF Award CPS-1739809, CCF-1350574, CNS-1801501, and NSA H98230-18-D-006.

References

Apthorpe, N.; Shvartzshnaider, Y.; Mathur, A.; Reisman, D.; and Feamster, N. 2018. Discovering Smart Home Internet of Things Privacy Norms Using Contextual Integrity. In *Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)*, 59:1–59:23. ACM.

Bhatia, J., and Breaux, T. 2018. Semantic Incompleteness in Privacy Policy Goals. In *International Requirements Engineering Conference (RE)*, 159–169. IEEE.

Bhatia, J.; Breaux, T.; Reidenberg, J.; and Norton, T. 2016a. A Theory of Vagueness and Privacy Risk Perception. In *International Requirements Engineering Conference (RE)*, 26–35. IEEE.

Bhatia, J.; Evans, M. C.; Wadkar, S.; and Breaux, T. D. 2016b. Automated Extraction of Regulated Information Types Using Hyponymy Relations. In *International Requirements Engineering Conference Workshops (REW)*, 19–25. IEEE.

Evans, M. C.; Bhatia, J.; Wadkar, S.; and Breaux, T. D. 2017. An Evaluation of Constituency-based Hyponymy Extraction from Privacy Policies. In *International Requirements Engineering Conference (RE)*, 312–321. IEEE.

Federal Trade Commission. 1998. Privacy Online: A Report to Congress. *Washington, DC*.

Federal Trade Commission. 2012. Protecting Consumer Privacy in an Era of Rapid Change: Recommendations for Businesses and Policymakers. *Washington, DC*.

Fillmore, C. J. 1976. Frame Semantics and the Nature of Language. *Annals of the New York Academy of Sciences* 280(1):20–32.

Frier, S. 2018. Facebook Updates Policies After Privacy Outcry, Limits Data Use. *Bloomberg*.

Guinchard, A. 2017. Contextual Integrity and EU Data Protection Law: Towards a More Informed and Transparent Analysis. *SSRN*.

Harkous, H.; Fawaz, K.; Lebret, R.; Schaub, F.; Shin, K.; and Aberer, K. 2018. Polisis: Automated Analysis and Presentation of Privacy Policies Using Deep Learning. In *USENIX Security Symposium*. USENIX.

Hosseini, M. B.; Wadkar, S.; Breaux, T. D.; and Niu, J. 2016. Lexical Similarity of Information Type Hypernyms, Meronyms and Synonyms in Privacy Policies. In *AAAI Fall Symposium Series*. AAAI.

Hull, G.; Lipford, H. R.; and Latulipe, C. 2011. Contextual Gaps: Privacy Issues on Facebook. *Ethics and Information Technology* 13(4):289–302.

Kincaid, P.; Fishburne Jr, R.; Rogers, R.; and Chissom, B. 1975. Derivation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel. Technical report, NATTC.

Litman, L.; Robinson, J.; and Abberbock, T. 2017. TurkPrime.com: A Versatile Crowdsourcing Data Acquisition Platform for the Behavioral Sciences. *Behavior research methods* 49(2):433–442.

Martin, K., and Nissenbaum, H. 2016. Measuring Privacy: An Empirical Test Using Context to Expose Confounding Variables. *Colum. Sci. & Tech. L. Rev.* 18:176.

Micheti, A.; Burkell, J.; and Steeves, V. 2010. Fixing Broken Doors: Strategies for Drafting Privacy Policies Young People can Understand. *Bulletin of Science, Technology & Society* 30(2):130–143.

Nissenbaum, H. 2010. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford University Press.

Rao, A.; Schaub, F.; Sadeh, N.; et al. 2016. Expecting the Unexpected: Understanding Mismatched Privacy Expectations Online. In *Symposium on Usable Privacy and Security (SOUPS)*, 77–96. USENIX.

Reidenberg, J.; Bhatia, J.; Breaux, T.; and Norton, T. 2016. Ambiguity in Privacy Policies and the Impact of Regulation. *The Journal of Legal Studies* 45(S2):S163–S190.

Schaub, F.; Balebako, R.; and Cranor, L. F. 2017. Designing Effective Privacy Notices and Controls. *Internet Computing*.

Schaub, F.; Balebako, R.; Durity, A. L.; and Cranor, L. F. 2015. A Design Space for Effective Privacy Notices. In *Symposium On Usable Privacy and Security (SOUPS)*, 1–17. USENIX.

Shvartzshnaider, Y.; Tong, S.; Wies, T.; Kift, P.; Nissenbaum, H.; Subramanian, L.; and Mittal, P. 2016. Learning Privacy Expectations by Crowdsourcing Contextual Informational Norms. In *Conference on Human Computation and Crowdsourcing*. AAAI.

Turow, J.; Hennessy, M.; and Bleakley, A. 2008. Consumers' Understanding of Privacy Rules in the Marketplace. *Journal of Consumer Affairs* 42(3):411–424.

Turow, J.; Hennessy, M.; and Draper, N. 2018. Persistent Misperceptions: Americans' Misplaced Confidence in Privacy Policies, 2003–2015. *Journal of Broadcasting & Electronic Media* 62(3):461–478.

Wijesekera, P.; Baokar, A.; Hosseini, A.; Egelman, S.; Wagner, D.; and Beznosov, K. 2015. Android Permissions Remystified: A Field Study on Contextual Integrity. In *USENIX Security Symposium*, 499–514.

Wilson, S.; Schaub, F.; Dara, A. A.; Liu, F.; et al. 2016a. The Creation and Analysis of a Website Privacy Policy Corpus. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, volume 1, 1330–1340.

Wilson, S.; Schaub, F.; Ramanath, R.; Sadeh, N.; et al. 2016b. Crowdsourcing Annotations for Websites' Privacy Policies: Can It Really Work? In *International Conference on World Wide Web (WWW)*, 133–143. WWW.

Wilson, S.; Schaub, F.; Liu, F.; et al. 2018. Analyzing Privacy Policies at Scale: From Crowdsourcing to Automated Annotations. *Transactions on the Web (TWEB)* 13(1):1.

Zimmer, M. 2008. Privacy on Planet Google: Using the Theory of Contextual Integrity to Clarify the Privacy Threats of Google's Quest for the Perfect Search Engine. *J. Bus. & Tech. L.* 3:109.