Cloak and Swagger: Understanding Data Sensitivity Through the Lens of User Anonymity

Sai Teja Peddinti*, Aleksandra Korolova†, Elie Bursztein†, and Geetanjali Sampemane†
*Polytechnic School of Engineering, New York University, Brooklyn, NY 11201
Email: psaiteja@nyu.edu
†Google, 1600 Amphitheatre Parkway, Mountain View, CA 94043
Email: korolova, elieb, geta@google.com

***Abstract*—Most of what we understand about data sensitivity is through user self-report (e.g., surveys); this paper is the first to use behavioral data to determine content sensitivity, via the clues that users give as to what information they consider private or sensitive through their use of privacy enhancing product features. We perform a large-scale analysis of user anonymity choices during their activity on Quora, a popular question-and-answer site. We identify categories of questions for which users are more likely to exercise anonymity and explore several machine learning approaches towards predicting whether a particular answer will be written anonymously. Our findings validate the viability of the proposed approach towards an automatic assessment of data sensitivity, show that data sensitivity is a nuanced measure that should be viewed on a continuum rather than as a binary concept, and advance the idea that machine learning over behavioral data can be effectively used in order to develop product features that can help keep users safe.**

I. Introduction

As the world moves to an ever-connected paradigm, online interactions are increasingly shaping how we interact with others and are perceived by them. The rise of services such as Facebook, Twitter, Google+, and YouTube that empower individuals to share their thoughts and experiences instantly and easily have opened the flood gates of user-generated content. This content deeply influences many aspects of our culture: from the creation of new dance styles [1] to the way breaking news are reported [2], to the rise of self-published authors [3].

A risk of this *always-on sharing* culture is that it may push users to share or express things that can harm them. The web is full of stories of careless or mistaken sharing of information or opinions that led to embarrassment or harm, from getting fired because of ranting about job frustrations [4] to public relations catastrophes due to "tweeting" under the influence [5].

The approach taken by online services to address this challenge to date has taken two directions: the first one defines what content users may consider sensitive and attempts to prevent its sharing without explicit confirmation. The second one introduces granular privacy controls in order to empower users to choose the desired privacy settings for each item they share. Both face scalability issues. Hand-crafted or survey-based definitions of sensitivity can hardly keep up with differences in preferences and expectations due to the context in which they are being applied or due to cultural, geographic, and demographic factors [6]. The second approach may be overwhelming due to diversity of privacy choices available.

In this work we explore whether it is possible to perform a large-scale behavioral data analysis, rather than to rely on surveys and self-report, in order to understand what topics users consider sensitive. Our goal is to help online service providers design policies and develop product features that promote user engagement and safer sharing and increase users' trust in online services' privacy practices.

Concretely, we perform analysis and data mining of the usage of privacy features on one of the largest question-and-answer sites, Quora [7], in order to identify topics potentially considered sensitive by its users. The analysis takes advantage of the Quora privacy feature that allows users to choose whether to answer each question anonymously or with their names attached. To learn what topics are potentially sensitive for Quora users, we analyze 587,653 Quora questions and 1,223,624 answers that span over 61,745 topics and 27,697 contexts. We find evidence in support of sensitivity of the oft-cited topics, such as those related to race, religion, sex, drugs, and sexual orientation [8], [9], [10], [11], and discover topic groups that are not typically included in such lists, many of them related to emotions, relationships, personal experiences, education, career, and insider knowledge. We use the obtained knowledge to build a machine learning model that is able to predict the sensitivity of particular questions with 80.4% accuracy and anonymity of answers with 88% accuracy, demonstrating that data on users' use of privacy-enhancing features can be used to develop policies and product features that enable safer sharing. Finally, we run a 1,500 person user survey on the US population via Google Consumer surveys [12] and compare our user activity-driven inferences with those obtained via a self-report. As far as we know, we are the first to use large-scale data analysis of users' privacy-related activity in order to infer content sensitivity and leverage the data towards building a machine learning model designed to help service providers design better privacy controls and foster engagement without a fear of over-sharing.

The remainder of the paper is organized as follows: in Section II we review the current approaches towards defining content sensitivity and concerns related to data sharing. In Section III we introduce Quora and its features, and describe the dataset we collected based on it. In Sections IV and V we present and discuss the results of our data analyses based on users' usage of Quora's anonymity features in terms of topics

and words indicative of sensitivity. In Section VI we present the results of our attempts to predict question-level and answer-level anonymity based on their content. Section VII discusses limitations of our approach and the challenges of relying on a purely data-driven analysis for identifying sensitivity, and presents a comparison of our findings with those based on an online survey. Section VIII describes related work on inferring users' privacy preferences, privacy risks, and efforts related to helping users minimize regret from sharing. We conclude by summarizing our contributions in Section IX.

II. Background

In this section we discuss the notions of content sensitivity adopted by several popular online services and data protection authorities, the potential negative consequences of over-sharing, and the positive impact that product features cognizant of data sensitivity can have on engagement with a product.

A. What is Sensitive Content?

There is no universally adopted definition of what constitutes sensitive content. Each online service provider defines sensitive data independently, describes it in the service's privacy policy or Terms of Service, and then develops functionalities or policies to observe this definition. For example, Google's privacy policy defines sensitive personal information as "confidential medical facts, racial or ethnic origins, political or religious beliefs or sexuality" [8] and Google does not associate cookies or anonymous identifiers with sensitive advertising categories as "those based on race, religion, sexual orientation or health" [13]. Facebook's advertising guidelines prohibit targeting users based on their personal characteristics from the categories of "race or ethnic origin; religion or philosophical belief; age; sexual orientation or sexual life; gender identity; disability or medical condition (including physical or mental health); financial status or information; membership in a trade union; and criminal record" [10]. Similarly, Microsoft's advertising policy states that ads cannot be related to prohibited and restricted categories such as adult content, firearms and weapons, gambling, surveillance equipment, suffering, violence and exploitation, dating/personals, health, political and religious content, etc. [9]. Quora considers adult content to be sensitive, as evidenced by their decision to disable the *views* feature (Section III) on questions related to that type of content [14].

Legal and data protection authorities have also proposed definition of sensitive content that is similar but not identical to those of online service providers. For example, CNIL [15], the French administrative regulatory body whose mission is focused on data privacy, defines sensitive data as "any type of data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, or data concerning health or sex life" [11]. A directive of the European Parliament delineates special categories of data for whom stricter processing laws apply and includes in it: racial or ethnic origins, views on politics, labour relations or religion, health, sex life, penalties, convictions and national ID [16].

User perceptions of content sensitivity may differ from the categories defined by online services and policy makers. Privacy experts call for definitions that would allow the concept of sensitive data to "*evolve over time and [depend] on the context, technologies, the use made of the data and even the individuals themselves*", highlighting that most current definitions are restricted to "*data that may give rise to discrimination*", and that "*even trivial data can become sensitive owing to the way it accumulates on the internet and the way in which it is processed*" [6]. The goal of this work is to deepen our understanding and capture the nuances of users' content sensitivity perception through a data-driven study of users' privacy-related actions. The data-driven approach could complement that based on user surveys and serve as a more scalable way to understand the evolution and differences in perception over time and across cultures.

B. Dangers of Over-Sharing

Social networks such as Facebook and Google+ offer controls to limit the visibility of one's posts to a specific user group. However, users often make mistakes when posting or sharing information that can lead to trouble. For example, sharing location information on Twitter facilitated a successful burglary [17], and inadvertent exposure of sexuality on Facebook resulted in threats to sever family ties [18]. Mistakes or unforeseen consequences of sharing can also cause discomfort or embarrassment [19] and there's ample evidence that users often regret their online postings [20].

Furthermore, data shared online has become one of the most-often checked sources of information about a person. For example, colleges routinely assess the social network profiles of their applicants [21], employers look at potential candidates' online profiles [22], and people of both genders research their potential dates online [23]. Thus the potential consequences of sharing mistakes (due to lack of understanding of sharing impact, inattentiveness, lack of privacy controls, or spur-of-the-moment decisions) are constantly increasing, which is starting to cause fear of engaging, sharing, or expressing one's opinion online [24], [25].

C. Impact of Better Privacy Tools and Features

On the other hand, development of many privacy-enhancing technologies and product features has enabled many people to engage and share online with more confidence and less risk [26], [27], [28], [29]. Recently, smart privacy features have become one of the core enablers of success for social networks and sharing services [30], [31].

In particular, though anonymity as a feature can reduce accountability and accuracy of information shared [32], it often enables people to be more open about their views [33], increases the effectiveness of leadership and group transactions [34], and enables user engagement [35].

Overall, a better understanding of user privacy fears and sensitivity of the content shared, would enable many online service providers to improve their products and engagement with their products. A feature on Facebook, LinkedIn, or

Google+ that double-checks the user's intention to share a drunken rant publicly or with their work colleagues, or alerts them that a post would likely make others aware of their religious or sexual preference, could help avoid sharing mistakes and build user confidence in sharing and engaging online. A feature that double-checks the intention to share a sensitive piece of information, with sensitivity evaluated from a user's perspective rather than a legal or one-size-fits-all perspective, would be even more impactful.

III. Quora

We use Quora [7], a popular question-and-answer site, in order to perform our proof-of-concept data-driven analysis of user perceptions of content sensitivity. Quora is a particularly fertile data source for such an analysis since it has a rich and prominent set of privacy features actively utilized by its user base when sharing about or expressing interest in a particular topic. We describe Quora functionality, its core privacy features, incentives for sharing anonymously or non-anonymously, and the characteristics of Quora as a study dataset next.

A. About Quora

Quora [7] is a *question-and-answer* website founded in 2009, somewhat similar to the once-popular Yahoo! Answers [36]. It has functionality that enables users to ask and answer questions on a variety of topics, as well as to "follow" or subscribe to updates on activity by other users or activity by all users related to a particular topic.

An example Quora page is shown in Figure 1, with several core features, present in each question page, highlighted. Every *Quora Question* page has three main information blocks that we are interested in – *Question*, *Answer* and *Follower* blocks. The *Question Block* has five pieces of information. It has two sets of tags (*Quora Context* and *Quora Topics*), the actual question text, additional question details and comments. Quora Context and Quora Topics are highlighted by a (violet) rectangular box at the top of Figure 1. Each question is assigned at most one context and zero or more topics by Quora moderators. Users can choose to follow individual topics or questions, in which case they receive notifications about new activity related to them and their follow choice gets shared with other users.

Each question has zero or more answers. Each answer has four pieces of information – the answerer details, a partial list of voters (who upvoted the answer), the answer text and comments. The answerer field contains the name of the person who answered and a short description of the person (highlighted by green box in Figure 1). If the answerer prefers to answer anonymously, she can use the *Make Anonymous* option available above the answer text box (highlighted by a red box in Figure 1). When the *Make Anonymous* option is exercised, others see *Anonymous* instead of the name in the answerer field (as highlighted by the red box in Figure 1).

Every question has zero or more followers, who are interested in the question and would like to be notified about new answers being posted. A partial grid of followers (indicated by their pictures) is provided at the bottom right of the question webpage (highlighted in an orange rectangular box in Figure 1). Only a max of 45 pictures are shown, even when the number of followers is much higher for a question. Similar to the option to answer anonymously, Quora provides an option to follow a question anonymously. The anonymous question followers are indicated by grey icons in the follower grid (as highlighted by the last icon in the followers list in Figure 1).

For every question, Quora also keeps track of the number of Quora users who viewed the question. The number of views is shown to all users above the grid of followers (as highlighted by a violet rectangular box at the bottom right in Figure 1). Clicking on that number provides the list of Quora users who viewed this particular Quora question.

B. Quora Privacy Features

Although Quora has a strict real-names policy [37], similar to that of other online social networks such as Facebook and Google+, it provides several privacy-related product features that are core to its functionality and are heavily utilized by its users [38]. Specifically, each user can choose whether to follow a question anonymously or non-anonymously, and whether to write an answer anonymously or non-anonymously by using the *Make Anonymous* feature. Users are also provided an option to hide their question views, so they are not listed in the user group who viewed a particular question (however the shown view count includes all page visits). Furthermore, Quora provides protection to its users against crawlers using a feature called *Search Engine Privacy* [39]. The default option is to **Allow** search engines to index the name. If indexing is disallowed by a user, Quora prevents crawlers, search engines, and other not-logged in users from seeing that user's profile page information, his activity and, renders any activity performed by that user to be indistinguishable from anonymous user activity for anyone except other logged in users.

C. Incentives for Anonymity and Non-Anonymity

Quora users describe several motivations for following and answering questions anonymously [40]. Some users prefer not to identify themselves in their answers when they relate to personal experiences, or experiences of friends and family members, or contain information about sensitive topics such as medical history. Others answer anonymously to avoid embarrassing or unfavorable situations that their answer can lead to [21], [22], [23], or to avoid trouble when sharing potentially sensitive or confidential information about companies about which they have insider knowledge. Others, who are striving to build a reputation in a certain domain, prefer to answer anonymously and reveal their identity later if the answer gains recognition or popularity, as indicated by up-votes, another Quora feature. Finally, since Quora is akin to a social network where people follow others, answering anonymously prevents the answer from appearing in followers' feeds.

There are several strong incentives for answering questions with one's real name, as pointed out in [41]. Providing one's identity along with an answer may lend it credibility [32], as readers can verify the provided information with the help

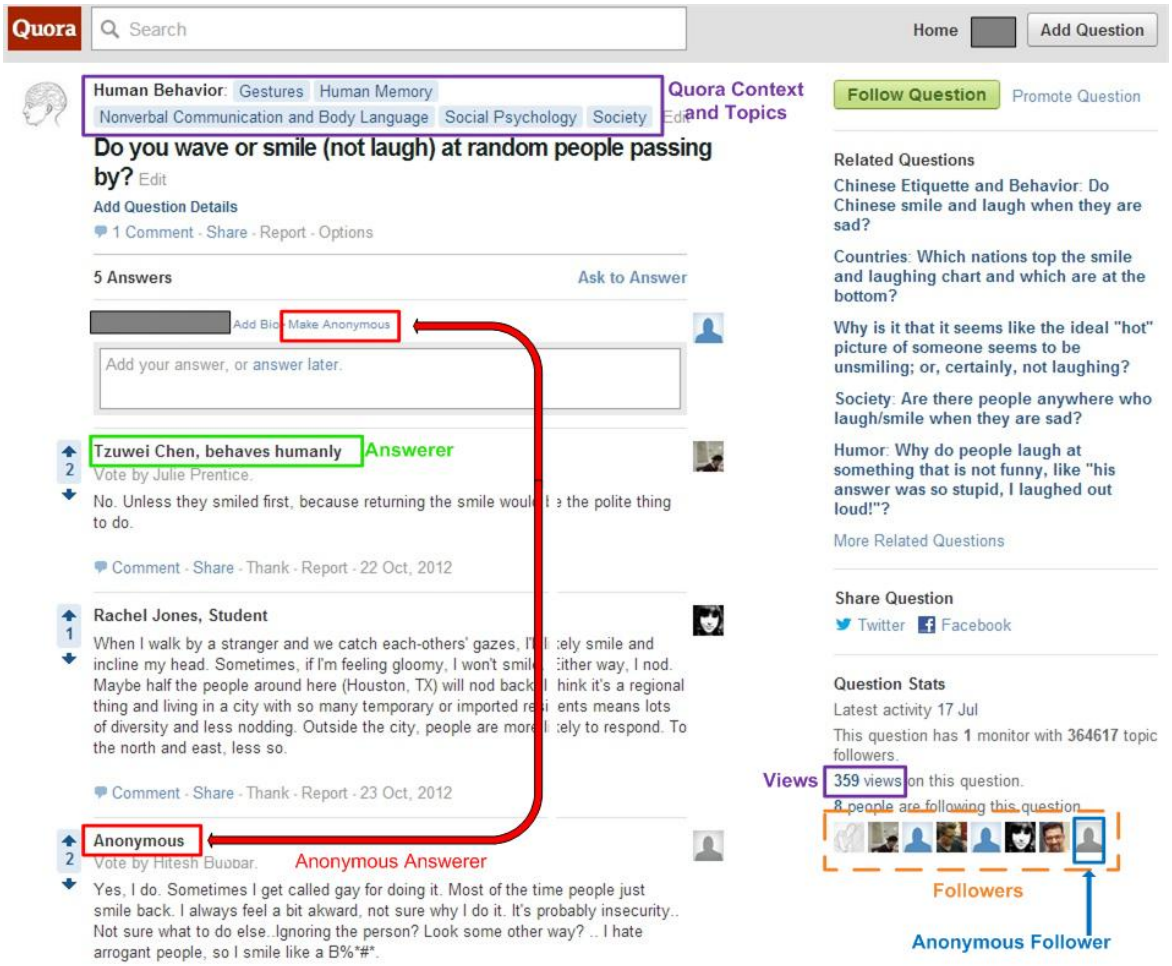

Fig. 1. Sample Quora Webpage with Interesting Components Highlighted

of details provided in the answerer's profile. Non-anonymous answers help build reputation, popularity, and social capital. It also helps build new connections, and it is revealed that people are using votes as social signals to draw attention of influential people.

Irrespective of an individual's reasons for answering anonymously or non-anonymously, one can argue that it is better for the Quora eco-system if many answers are provided non-anonymously. Such answers promote user interaction and engagement, as non-anonymous answers appear in the answerer's followers feeds. Since users may take more care when answering non-anonymously as they want to appear knowledgeable, the quality of such content increases. Hence, enabling users to share non-anonymously while avoiding undesirable situations would be a desirable outcome for Quora.

D. Dataset Characteristics

We crawled the Quora website using our own custom crawler during the period of August - October, 2012. We follow a similar approach as outlined in [42] for crawling Quora. Our crawler observed the Quora's *robots.txt* as well as rate-limited our access. Furthermore, in order to limit the request load, we only crawled the Quora question pages, and omitted all other pages, such as answer pages, follower pages, activity pages, views pages, and user profile pages. As a result, the information we obtained about question followers is limited to the followers listed at the bottom right of the Quora question page (Figure 1), and does not include all question followers. The question pages list up to 45 followers, and our manual inspection suggests those are chosen at random (with caching). Furthermore, we have observed that the answers of users who have enabled the "Search Engine Privacy" feature on Quora appear as "Anonymous" to non logged-in users, regardless of whether that answer was written anonymously or not [39]. Since our crawler did not possess the credentials of a logged in Quora user, our dataset does not distinguish between answers that were written anonymously and those that were labeled by Quora as "Anonymous" due to users' "Search Engine Privacy" settings, which is an important limitation. We discuss the implications of these crawl limitations in Sections IV-B2 and VII-A.

Our obtained dataset contains 587,653 Quora questions. Of these, 437,622 (74.47%) have at least one answer, and 563,954 (95.97%) have at least one follower. The number of Quora questions containing at least one anonymous answer is 138,576 (23.58%), while number of questions with at least one anonymous follower is 336,551 (57.27%). Since the effort

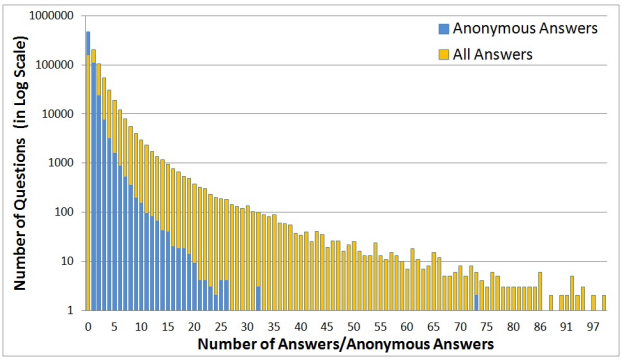

Fig. 2. Quora Anonymous and All Answers Distribution

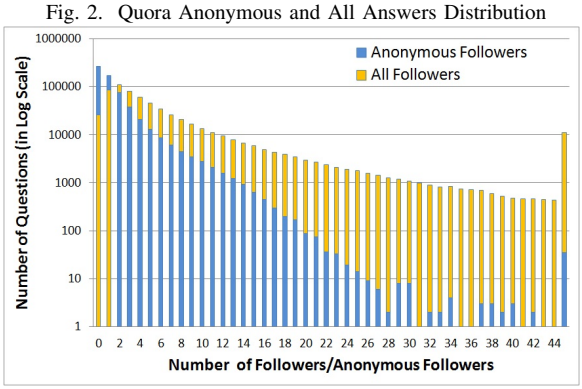

Fig. 3. Quora Anonymous and All Followers Distribution

required to answer a question is significantly higher than the effort required to follow a question, and, furthermore, answering a question requires some knowledge about the question's topic, whereas following a question is merely an expression of curiosity or interest about it and its potential answers, it is not surprising that there are more questions with at least one follower than with at least one answer.

The distribution of the number of answers and anonymous answers for questions in our dataset is shown in Figure 2; the x-axis shows the number of answers or anonymous answers and the y-axis – the number of questions with that many answers (in log scale). The figure omits a handful of questions that have more than 100 answers for readability. The question with the most answers and the most anonymous answers in our dataset "*What is the most useful, shortest, and generally applicable piece of wisdom?*"[1] has 440 answers, 73 of them anonymous.

The distribution of the number of followers for the questions in our collected dataset is shown in Figure 3; the x-axis shows the number of followers or anonymous followers and the y-axis – the number of questions with that many followers[2].

IV. Sensitive Contexts and Topics

In this section, we present two methodologies for identifying which contexts and topics may be more sensitive than others based on Quora user anonymity actions. Both approaches give support for sensitivity of topics typically considered sensitive, such as sex, religion, etc., but also suggest themes outside of the typical set as potentially sensitive. In addition to identifying topics that are considered potentially sensitive by Quora users, our findings lend support to the feasibility of our proposed approach – identifying user sensitivities and privacy preferences via behavioral data analysis of users' use of privacy-enhancing product features.

A. Measuring Context-level and Topic-level Anonymity Ratios

In our first methodology, we measure a topic's sensitivity level by considering all questions belonging to the topic and computing the fraction of answers to those questions that are anonymous among the total number of answers posted for those questions. We perform this analysis separately for Quora topics and Quora contexts. We choose to use the Quora-assigned topics and contexts, rather than classify the questions and answers into our own topic hierarchy, because Quora's human moderators have spent significant effort in order to hand-label each of the questions with a corresponding context or set of topics, and thus we expect their label quality to be higher than what can be derived based on a short snippet via unsupervised machine learning techniques. We exclude topics and contexts for which we do not have sufficient data from consideration in order to avoid making erroneous conclusions. We also exclude questions that did not receive a single answer from consideration in all our analyses, as they do not provide information about the anonymity/non-anonymity user choices that are of interest for this research.

1) Context Data: There are 21,232 contexts on Quora that have at least 1 question with at least 1 answer and at least 1 follower[3]. The average number of questions per context among these is 10.85, so we consider only the most popular contexts, i.e., those that have at least 11 questions, in our analysis. The 3,129 contexts with at least 11 questions belonging to them contain 188,121 questions with a total of 512,225 answers, 86,213 of which are anonymous, suggesting a 0.17 overall anonymity rate. We further limit our analysis to the 1,525 contexts that, in addition to being comprised of at least 11 questions each, contain at least 66 answers. The motivation for choosing contexts with at least 6 answers per question on average stems from the desire to focus on contexts that have generated sufficient engagement. Additionally, a question with 6 answers one of which is anonymous, has an anonymity rate of 0.17, which is the overall average anonymity rate for contexts with 11 questions; furthermore, a question with 6 answers two of which are anonymous, would have an anonymity rate of more than twice the average, enabling even the discrete answer counts to distinguish between average and above average [43]. These 1,525 contexts have a total of 159,884 questions with 452,221 answers, 76,947 of which are anonymous.

For each context C of the 1,525 contexts that contain at least 11 questions and at least 66 answers, we compute its answer anonymity ratio, $A(C)$, as the fraction of anonymous answers

[1]http://www.quora.com/Advice/What-is-the-most-useful-shortest-and-most-generally-applicable-piece-of-wisdom

[2]We use overlapped rather than stacked bars in Figures 2 and 3 for ease of comparison.

[3]A large number of questions, 256,270, are not labeled with any context and are, therefore, excluded from this part of the analysis.

Sex, *Penises*, *Political Thinking (1986 book)*, *Indian Muslims*, **Attractiveness and Attractive People**, *Cheating (relationship and marital infidelity)*, **Anonymity on Quora**, **Patent Law**, **Greece**, **Palantir Technologies**, *Pick-Up Artists*, *Prostitution*, *Interracial Dating and Relationships*, **Intellectual Property**, *Pornography*, *Sexism*, *Secrets*, *Bipolar Disorder*, **Patents**, *Asian Americans*, **California Institute of Technology**, **Hacking (computer security)**, *Abortion*, **Bridge (card game)**, *OkCupid*, **Topics (Quora feature)**, *Women*, *Racism*, **Recipes**, **Boards (Quora feature)**, *Investment Banking*, *Ethnic and Cultural Differences*, *LGBTQ Issues*, **Baby Names**, **Square, Inc.**, *British-American Differences*, *Judaism*, *Depression*, **The Ivy League**, **Views on Quora (feature)**, *Salaries*, **What Does It Feel Like to X?**, *Race and Ethnicity*, *Humor on Quora*, **Harvard College**, **Interpersonal Interaction**, **Friendship**, **Hard Disk Drives (HDD)**, *Taxes*, *Gender Differences*, *Dating and Relationships*, *American Express*, **Menlo Park, CA**, *Men*, **Table Tennis**, **Airlines**, **Mitt Romney's Taxes and Related Debate (Summer 2012)**, **Higgs Boson**, **Joke Question**, *Indian People*, **Middle East**, *Feminism*, *Asian People*, *Cannabis*, **Hackers**, *LGBTQ*, **Civil Engineering**, *Armchair Philosophy*, *Dating Advice*, *God*, **Trolling on the Internet**, **IQ**, *Suicide*, *Same Sex Marriage*, **Management Consulting and Management Consulting Firms**, **Quora Etiquette**, **Sparrow (mail app)**, **Quora Moderation**, **Foreign Policy**, **Social Advice**, *Self-Defense*, *Rush Limbaugh*, *Christianity*, **Quora Promote Feature**, **Harry Potter Book 7 Deathly Hallows (2007 book)**, **Expressions (language)**, *Breakups*, **Trains (transportation service)**, **Names and Naming**, **Downtown Palo Alto**, **Quora (product)**, *Iranian Nuclear Threat and Potential Israeli Attack*, *Homosexuality*, **German (language)**, *Jewish People*, **Flying**, **Wealthy People and Families**, **Zynga**, **Product Naming**, **Air Travel**

Fig. 4. Top 100 Contexts by Anonymity Ratio. In italics – those that belong to themes already considered sensitive by Facebook, Google and CNIL; in bold – those that do not, according to manual categorization by hired workers.

received to questions within that context to the total number of answers received to questions within that context. Similarly, we compute its follower anonymity ratio, $F(C)$, as the fraction of anonymous followers among the total number of followers. We observe that: $\text{mean}\big(A(C)\big) = 0.165, \text{stdev}\big(A(C)\big) = 0.078$; $\text{mean}\big(F(C)\big) = 0.172, \text{stdev}\big(F(C)\big) = 0.044$. Furthermore, answer anonymity ratios and follower anonymity ratios are highly correlated, $\text{corr}(A, F) = 0.84$.

2) Context-Based Results: Our findings are presented in Table I and Figure 4. The former presents statistics for those 14 contexts whose answer anonymity ratio, $A(C)$, is three standard deviations above the mean, while the latter lists the top 100 Quora contexts in the decreasing order of their $A(C)$s.

We manually analyze each of the 243 contexts whose answer anonymity ratio exceeds one standard deviation above the average. For each context we attempt to assess whether it belongs to one of the typically considered sensitive categories, and if not, identify its broader theme. We do so by hiring 5 workers[4] and tasking them with labeling each context with one or two of the most appropriate 41 themes we provide (or "None"). The themes we provide are comprised of the top-level content themes utilized by Google AdWords[5] and the categories typically considered sensitive as described in Section II-A.

Several observations based on this analysis strike us as noteworthy. First, the majority of sensitive categories described by Google, Facebook, Microsoft, and CNIL, such as *racial or ethnic origins; political, philosophical, or religious beliefs; sexual orientation or sex life; gender identity; disability or medical condition (including physical or mental health); financial status or information; dating/personals; weapons*, have supporting evidence among the selected Quora contexts. For example, supporting evidence for the sensitivity of *financial information* are contexts such as *Salaries, Taxes, Investment Banking* and *American Express*. The only exception for which we did not find supporting evidence among the ones used by these four entities is: *criminal record*. The absence of evidence in favor of this is likely due to selection biases among Quora users and questions.

Second, as is visually clear from Figure 4, which distinguishes the contexts that belong to typically considered sensitive categories (as judged by at least two workers) from those that do not, although several of the contexts with the highest anonymity ratio belong to the categories of data typically considered sensitive, many do not. Specifically, 120 contexts out of 243 considered, were not associated with a conventional sensitive category by any worker. We loosely group the contexts whose answer anonymity ratio exceeds one standard deviation but do not belong to any of the conventionally considered sensitive categories into themes, and present the themes and the contexts supportive of them in Table II.

Our findings based on this analysis methodology support the hypothesis that data sensitivity is quite nuanced, and that sensitive topics include but are not limited to the ones typically considered.

3) Topic data: We repeat a similar analysis to the one performed for Quora contexts above, for Quora topics. Specifically, there are 53,551 topics on Quora that have at least 1 question with at least 1 answer and at least 1 follower. The average number of questions per topic is 22.6, larger than the average per context, since each question is labeled with at most one context but may be labeled with many topics. We consider only the most popular topics, i.e., those that have at least 23 questions, in our analysis. The 6,799 topics with at least 23 questions each, contain 418,575 questions with a total of 1,027,549 answers, 178,038 of which are anonymous, suggesting 0.17 overall anonymity rate. We further limit our analysis to the 4,067 topics that, in addition to being comprised of at least 23 questions each, contain at least 138 answers, i.e., on average, 6 answers per question, as in the analysis above. These 4,067 topics have a total of 408,828 questions with a total of 1,014,300 answers, 175,986 of which are anonymous.

For each topic T of the 4,067 topics that contain at least 23 questions and at least 138 answers, we compute its answer anonymity ratio, $A(T)$, as the fraction of anonymous answers received to questions within that topic to the total number of answers received to questions within that topic. Similarly, we compute its follower anonymity ratio, $F(T)$, as the fraction of anonymous followers among the total number of followers. We

[4]Workers hired and paid using the outsourcing service Premier [44].

[5]https://support.google.com/adwords/answer/156178?ctx=tltp

Context name	# Questions	# Answers	# Anonymous Answers	# Followers	# Anonymous Followers	$A(C)$	$F(C)$
Sex	449	1067	561	3062	1584	0.526	0.341
Penises	28	81	41	194	88	0.506	0.312
Political Thinking (1986 book)	26	91	42	180	79	0.462	0.305
Indian Muslims	40	109	49	224	85	0.450	0.275
Attractiveness and Attractive People	85	313	140	926	378	0.447	0.290
Cheating (relationship & marital infidelity)	51	156	69	453	206	0.442	0.313
Anonymity on Quora	78	172	74	702	320	0.430	0.313
Patent Law	52	88	37	191	60	0.420	0.239
Greece	36	74	31	209	71	0.419	0.254
Palantir Technologies	58	97	40	1268	441	0.412	0.258
Pick-Up Artists	32	114	46	371	150	0.404	0.288
Prostitution	24	72	29	298	126	0.403	0.297
Interracial Dating and Relationships	37	110	44	431	202	0.400	0.319
Intellectual Property	54	95	38	213	69	0.400	0.245

TABLE I
QUORA CONTEXTS WITH HIGH ANONYMITY RATIO, $A(C)$

Theme	Example Context Support	Example Topic Support
Quora Product	*Quora Etiquette, Quora Moderation, Quora Credits, Upvoting and Downvoting (Quora feature)*	*Quora User Tips, Quora User Feedback*
Law & Government	*Patent Law, U.S. Foreign Policy and Foreign Relations,Intellectual Property Law, Freedom of Speech, U.S. Supreme Court, Justice*	*International Law and Legal Institutions, Legislation, Central Intelligence Agency, Freedom of Speech, Tax Law, U.S. Constitutional Law*
Personal Experiences	*What Does it Feel Like to X?, High School, The College and University Experience, Teenagers and Teenage Years*	*Life Decisions, What Would You Do if X?, What Does It Feel Like to X?*
Companies	*Zynga, Square Inc., Sparrow, Palantir Technologies, Management Consulting and Management Consulting Firms*	*The Boston Consulting Group, Bain and Company, Blekko, What Is It Like to Work At X?*
Education and Educational Institutions	*The Ivy League, Graduate School Admissions, Harvard College, Law School*	*Yale University, Cornell University, Exams and Tests*
Relationships	*Interpersonal Interaction, Friendship, Meeting New People, Family and Families*	*Relationship Counseling, Interpersonal Conflicts, Working and Dealing with Difficult People*
Emotions & Emotional Experiences	*Love, Bullying, Breakups, Humor*	*Sadness, Embarrassment, Crying, Annoyances, Jealousy, Revenge*
Career	*Work-Life Balance, Unemployment, Workplace and Professional Etiquette, People Skills*	*Women in Technology, PhD Careers*
Humor	*Jokes, Pranks, Joke Question*	*Jokes, Laughing*
Science	*Civil Engineering, Higgs Boson, NASA, Architecture, Materials Science, Space Shuttle*	*Aerospace Engineering, Polymaths, String Theory, Scientific Explanations, Anthropology, Mathematical Optimization*
Arts & Entertainment, Celebrities	*Royalty, Celebrities, Fictional Characters, Rush Limbaugh*	*Celebrity Gossip, Art Collecting, Music Videos, Warren Buffett, Michael Arrington, Marissa Mayer*
Travel & Transportation	*Flying, Airlines, Trains (transportation service), Air Travel, Google Self-Driving Cars*	*Airfares, Airports, Subways, Flights, Highways*
Social issues	*Sexism, Economic Inequality, Manners and Etiquette, Human Behavior, Social Advice*	*Social Customs, Human Rights, Civil Society, Immigration Policy, Pro-Life Movement*
History and Historical Events	*Apollo 11: 1969 Moon Landing, U.S. Presidents, Conspiracy Theories*	*Ancient Greece, World War I*
Psychology and Philosophy	*Child Psychology, Gender Differences, Human Behavior, IQ, Intelligence, Child Psychology, Morals and Morality*	*Emotions, Self-Esteem, Ethics, Existentialism, Bad Habits*
Popular Culture	*The Hunger Games (2012 Movie), Harry Potter Book 7 Deathly Hallows (2007 book)*	*Voldemort (Harry Potter character), Doomsday*
Internet Privacy and Security	*Trolling on the Internet, Hacking (computer security), Hackers, Anonymity on Quora, Search Engines, Computer Security*	*Hacker Culture, Hackers, Internet Etiquette*
Food & Drink	*Drinking (alcohol), Eggs, Vegetarianism, Cheese, Recipes*	*Flavors, Grilling, Sauces, Steak, Ice Cream, Tastes of Meat*

TABLE II
POPULAR THEMES AMONG THE CONTEXTS AND TOPICS THAT ARE NOT COVERED BY CONVENTIONAL SENSITIVITY ASSUMPTIONS

Topic name	# Questions	# Answers	# Anonymous Answers	# Followers	# Anonymous Followers	$A(T)$	$F(T)$
Orgasms	137	363	198	1408	748	0.545	0.347
Masturbation	107	246	133	706	398	0.541	0.361
Genitalia	67	173	90	468	223	0.520	0.323
Penises	108	292	147	880	427	0.503	0.327
Female Sexuality	50	170	84	541	262	0.494	0.326
Adult Content on Quora	184	646	317	2050	994	0.491	0.327
LSD	70	154	73	650	270	0.474	0.293
Sex	1514	4258	1930	13314	6429	0.453	0.326
Sexuality	542	1520	671	4715	2107	0.441	0.309
What Does It Feel Like to Be in a Relationship with X?	33	162	71	674	236	0.438	0.259
Male Sexuality	80	205	89	515	272	0.434	0.346
Internet Pornography	61	147	62	434	189	0.422	0.303
Rape	92	338	142	881	374	0.420	0.298
Abuse	39	168	70	327	122	0.417	0.272
Psychology of Sexuality	59	184	76	605	279	0.413	0.316
Investment Banks	65	170	70	707	242	0.412	0.255
Anonymity on Quora	207	526	216	2038	934	0.411	0.314
Interracial Dating and Relationships	103	479	196	1409	554	0.409	0.282
Sexual Ethics	110	443	181	1239	538	0.409	0.303
Drug Effects	88	225	89	664	284	0.396	0.300
Seduction	34	210	83	587	245	0.395	0.294
Lady Gaga	112	228	90	627	166	0.395	0.209
Game (seduction technique)	47	166	65	575	198	0.392	0.256
Why Do Women Do or Like X?	74	473	185	1292	541	0.391	0.295
Sexual Orientation	61	202	79	535	209	0.391	0.281
Sex Workers & Prostitution	88	205	80	737	297	0.390	0.287
Transgender	80	191	74	481	163	0.387	0.253
Bathroom Etiquette	38	165	63	266	113	0.382	0.298
Cannabis	310	849	323	2077	792	0.380	0.276
Being Single	57	266	101	682	275	0.380	0.287
Sexual Attraction	75	277	105	749	321	0.379	0.300
University of Pennsylvania	33	177	67	381	140	0.379	0.269
Pornography	224	450	170	1588	680	0.378	0.300
Bipolar Disorder	72	228	86	664	262	0.377	0.283
Casual Sex	54	154	58	525	216	0.377	0.291

TABLE III
QUORA TOPICS WITH HIGH ANONYMITY RATIO, $A(T)$

observe that: mean$\big(A(T)\big) = 0.173$, stdev$\big(A(T)\big) = 0.068$; mean$\big(F(T)\big) = 0.181$, stdev$\big(F(T)\big) = 0.039$. Furthermore, answer anonymity ratios and follower anonymity ratios are highly correlated, corr$(A, F) = 0.88$.

4) Topic-based Results: Table III presents statistics for those 35 topics whose $A(T)$ is three standard deviations above the mean. As becomes immediately clear from the table, the most sensitive topics are dominated by the adult themes of *sex, sexuality, sexual orientation, pornography*, and by the theme of *drugs*. However, even among these, there are outliers: *Lady Gaga, Bathroom Etiquette, University of Pennsylvania*, confirming our hypothesis from the study of contexts that topics considered sensitive by users are not limited to the obvious ones, and that education, celebrities, and personal experiences may be important exception themes.

Similar to the manual analysis done for contexts, we hired five workers to label all the 596 Quora topics whose answer anonymity ratio, $A(T)$, exceeds one standard deviation above the mean. As was the case for contexts, a high number of topics, namely 188, were not associated with any of the conventionally considered sensitive categories by any of the workers. Our loose categorization of these topics into themes is presented in Table II.

The analysis based on topics lends support for all typically considered sensitive categories, including *criminal record*, via high answer anonymity ratios for topics such as: *Capital Punishment, Organized Crime, When the Police Arrest You or Pull You Over*.

B. Discussion of Approach and Findings

1) The Approach: Although we present the results based only on answer anonymity ratio, A, the results based on the follower anonymity ratio, F, are quite similar for both contexts and topics. This is not unexpected, based on previously mentioned high correlation between the two measures (0.84 in the case of contexts and 0.88 in the case of topics), and the common sense that given the Quora features, someone who prefers not to associate one's interest in a topic with their real name, would also prefer to answer questions in that topic anonymously, and vice versa. Several notable exceptions to this, where A is significantly higher than F are the topics and contexts of: *Patent Law, Orgasms, Genitalia*; whereas the situation is reversed for *Interviews (Behavioral), Student Loans and Debt, Immigration*.

We have explored two methodologies for inferring sensitivity: one based on contexts (Section IV-A1) and another based on topics (Section IV-A3), and both yield similar and consistent results, which adds confidence to the methodology and robustness of findings.

2) Search Engine Privacy Impact Mitigation: Furthermore, it is important to remember that due to the method used for data collection, we are not able to distinguish between answers and followers that are truly anonymous versus those that are marked as such due to "Search Engine Privacy" settings by those users. Although this is a potentially significant limitation, we believe it has a limited impact on the conclusions made in the preceding analyses for the following reasons.

Firstly, the "Search Engine Privacy" setting is not enabled by default on Quora, which likely implies its limited utilization since users rarely change defaults [45]. Secondly, users who seek out and choose to enable this setting likely do so because of the nature of the questions they are following or answering, and their desire to protect their privacy while doing so, advancing an argument that actions by users whose "Search Engine Privacy" setting is enabled should be viewed as a weaker, but also possibly valid, indicators of sensitivity. Finally, if the previous argument is not correct, then the limitation due to search engine privacy should affect all topics and contexts at an equal rate in expectation, making the absolute anonymity ratios for topics and contexts higher than the true ones, but doing so equally, and therefore, enabling correct conclusions based on the relative comparisons between the average ratios.

To verify the previous two hypotheses, we randomly sampled 100 question URLs which include a context and have at least 6 answers from each of the following groups of questions: our entire crawl, the 14 contexts with the highest anonymity ratio (Table I), our crawl excluding the 14 contexts with the highest anonymity ratio, the contexts ranked 15-28 according to the anonymity ratio. For each of the 100 questions, we manually loaded the corresponding Quora page while being signed-in (and thereby, bypassing the crawl limitation) and noted the total number of answers and number of anonymous answers for it. Table IV presents the anonymity ratio computed for each of the four groups of questions based on the data not subject to the "Search Engine Privacy" limitation and the data subject to it. As expected, the true anonymity ratios are lower than the ones computed based on our crawl, but the relative magnitudes are unchanged, with questions from contexts ranked 1-14 and 15-28 based on our crawl exhibiting significantly higher true anonymity ratios than the average.

Set from Which Questions Chosen	True $A(C)$	$A(C)$ w/ "Search Engine Privacy"
All data	0.08	0.17
Contexts ranked 1-14 based on $A(C)$	0.30	0.48
All data excluding contexts ranked 1-14	0.06	0.19
Contexts ranked 15-28 based on $A(C)$	0.18	0.38

TABLE IV
ANONYMITY RATIOS COMPUTED ON CRAWL DATA SUBJECT TO "SEARCH ENGINE PRIVACY" CONSTRAINT VS MANUALLY OBTAINED DATA NOT SUBJECT TO IT

These findings lend credibility to our hypothesis that the impact of "Search Engine Privacy" on our conclusion is limited, as long as we rely on relative, rather than absolute values of anonymity ratios when comparing contexts and topics for sensitivity. We base most of our analyses in the subsequent sections on the data from identified contexts and topics with high anonymity ratio, and therefore, hope to further mitigate the impact of our crawl limitation due to "Search Engine Privacy".

3) Surprising Findings: Although, arguably, many readers would have predicted that the themes of relationships, law & government, and personal experiences would be among the ones for which Quora anonymity features are highly utilized, there are several themes among our findings whose prominence among the topics and contexts for whom anonymity is utilized is quite unexpected. In particular, we speculate on the reasons for some of the unexpected findings:

- Answers to education and educational institution related questions are often anonymous spurred by questions such as "*What are the downsides of attending Harvard as an undergrad?*"[6]
- Answers to questions related to particular companies are often anonymous due to possibility of disclosing information that only insiders of the company have access to, e.g., "*How do Zynga employees feel about the company's summer 2012 stock price drop?*"[7]
- Humor makes the list because of answers or questions that are not politically correct or may hurt someone's feelings, e.g., "*What's the most offensive joke ever?*"[8]
- Celebrities – because users may be interested in the gossip but not eager to admit it, e.g., "*Who are famous people who had/have relationships with dogs?*"[9]
- Several topics related to online privacy and security also elicit a high rate of anonymous answers and followers.

One hypothesis for the unifying reason for these seemingly surprising sensitive themes is that they combine a topic with feelings, personal experiences or thoughts, or insider information. This suggests one avenue for possible future work in order to develop better privacy-preserving features that would enable users to share without regrets or negative consequences – to rely not only on a set of pre-identified sensitive topics, but to also evaluate whether the question or its answers may include personal experiences, feelings, judgements, emotions, or insider information. Another possible conclusion is one that supports the main thesis of this research – content sensitivity is quite nuanced, and one of the core methods to understand and accommodate users' preferences should be based on a data-driven analysis of user actions related to the use of privacy-enhancing features in the product for which the sensitivity policies are to be set.

V. SENSITIVE WORDS

In this section, we perform an analysis that compares vocabulary of anonymous answers with the vocabulary of non-anonymous answers. As was the case for topics and contexts,

[6]http://www.quora.com/Harvard-College/What-are-the-downsides-of-attending-Harvard-as-an-undergrad, 9 answers, 8 of them anonymous

[7]https://www.quora.com/Zynga-Stock-Price-Collapse-Summer-2012/How-do-Zynga-employees-feel-about-the-companys-summer-2012-stock-price-drop, 21 answers, 17 of them anonymous

[8]https://www.quora.com/Whats-the-most-offensive-joke-ever, 56 answers, 30 of them anonymous

[9]http://www.quora.com/Celebrities/Who-are-famous-people-who-had-have-relationships-with-dogs, 4 answers, 1 of them anonymous

the words that are more prominent in anonymous answers are not limited to the expected set.

A. Word Data

We limit our analysis of sensitive words to answers from questions that belong to one of the 243 contexts whose anonymity ratio, $A(C)$, exceeds the overall average by at least one standard deviation identified in Section IV-A2. Such a choice allows us to partially mitigate the impact of "Search Engine Privacy" limitation, as the anonymity ratio is higher in these contexts regardless (see discussion in Section IV-B2). We do not use word stemming in order to preserve ability to easily reason about findings rather than have to guess the word a particular root form is arising from. Hence, we observe some root word repetitions in the reported results, e.g., both singular and plural forms of the same word.

Among the answers analyzed, 60,912 distinct words occur 1,952,979 times. For every word, we calculate its number of occurrences in anonymous answers and its number of occurrences in non-anonymous answers. The average number of occurrences of a word in anonymous answers is 10.2 and in nonanonymous – 21.8, with the latter being (unsurprisingly) higher than the former since there are more nonanonymous answers than anonymous ones. To avoid making statistically spurious observations, we exclude words with less than 32 ($= 10.2 + 21.8$) occurrences in total among all answers from consideration. Among the remaining, reasonably frequent, 5,396 words, we manually identify and remove 114 so-called stop words (such as "like", "the", "and", "or", etc.). The remaining 5,281 words occur a total of 939,849 times. We analyse these words to identify strong indicators of answer anonymity and content sensitivity.

B. Analysis Methodologies

We explore two methodologies – one statistical and another natural language processing based – for identifying words that are strong indicators of anonymity. We do not claim one method is better than the other, but only highlight the fact that multiple analysis approaches exist and may offer slightly differing perspectives. An online service provider may choose to combine several such techniques in practice.

1) Statistical Analysis: In our first methodology, for each word, we divide its number of occurrences in anonymous answers by the total number of all word occurrences in anonymous answers to obtain its normalized rate of occurrence in anonymous answers, $R_A(W)$. Similarly, we compute $R_N(W)$ based on number of the occurrences in non-anonymous answers. We then compute each word's anonymity ratio, $A(W)$, as R_A/R_N. We observe that, $\text{mean}\big(A(W)\big) = 1.05, \text{median}\big(A(W)\big) = 0.98, \text{stdev}\big(A(W)\big) = 0.69$.

The intuition behind such choice of measurements is that a word W that is not relevant to the outcome of whether the answer is anonymous or not will have approximately the same rate of occurrence in both types of answers, i.e., $R_A(W) \approx R_N(W)$, whereas for a word relevant to the outcome, R_A will significantly exceed R_N. Confirming this, the average and median of A are both close to 1; whereas there are 159 words with word anonymity ratio, $A(W)$, at least three standard deviations above the mean of A. We present the top 100 words based on their anonymity ratio in Figure 5, formatted analogously to the coding of contexts in Section IV-A2.

proverbs, **verifone**, *transgender*, *leviticus*, *revelation*, *breasts*, *asians*, **queue**, *vagina*, **merchants**, **boiling**, **gorgeous**, *orgasm*, *vietnamese*, **gulf**, **turkey**, **apology**, **boson**, **reader**, **borderline**, **lift**, **modeling**, **merchant**, *mastercard*, *bidding*, **laughing**, *payment*, *girlfriends*, **sue**, *testament*, **arthur**, **square**, **arabic**, **ashamed**, *commission*, **loop**, **aggressively**, **clearance**, **affirmative**, *feminists*, **astronauts**, **righteous**, *lds*, **bedroom**, **relatives**, **faithful**, *pregnancy*, **saudi**, *medication*, **retail**, **witness**, **grandfather**, **denied**, **admissions**, **lane**, **secretly**, **leg**, **api**, **nerd**, **orbiter**, **translations**, **bird**, **immigration**, *rape*, *reproduction*, **bond**, **pitch**, **wet**, **officers**, **tuna**, *kissing*, **stereotype**, **gate**, *transaction*, **colleges**, **card**, **wash**, **jack**, *lover*, **spoon**, *christ*, **governments**, **sour**, **faculty**, **nervous**, **dress**, **dorm**, **graduates**, **sticking**, **academics**, **crossing**, **forgiveness**, **partial**, **neighbors**, *girlfriend*, *quran*, **terribly**, **acquiring**, **customers**, **grandmother**

Fig. 5. Words with High Anonymity Ratio, $A(W)$. In italics – those that belong to conventionally sensitive themes; in bold – those that do not, according to manual categorization by the authors.

square, **quora**, **answer**, **content**, *nondual*, *asians*, *sex*, **card**, **proverbs**, **merchants**, **merchant**, *testament*, **reader**, *gay*, **verifone**, *payment*, **user**, *christ*, **questions**, *girlfriend*, **boson**, **palantir**, *leviticus*, *asian*, **question**, *revelation*, **woman**, *transgender*, **going**, *bible*, **english**, **committee**, *messiah*, **college**, **world**, **israel**, **women**, **turkey**, **marines**, *gods*, **anon**, **date**, **site**, **lift**, *orgasm*, **story**, **dress**, **feel**, **friends**, **queue**, **gorgeous**, **eyed**, **charlie**, **zynga**, **followers**, **girl**, *judas*, **ryan**, **customers**, **night**, *pregnancy*, *transaction*, **higgs**, **cheese**, **men**, **jack**, *jesus*, *feminists*, *vagina*, **admins**, **relatives**, *atheists*, **deeper**, **france**, *rape*, **parents**, *girlfriends*, *breasts*, **modeling**, **apology**, **posts**, **speech**, *jewish*, **lane**, **gps**, **fiction**, **another**, **feet**, **morality**, **partner**, *aging*, **technical**, **science**, **jon**, **form**, **beef**, **leaf**, **boiling**, **gulf**, **vietnamese**

Fig. 6. Sensitive Words based on Likelihood Ratio Test. In italics – those that belong to conventionally sensitive themes; in bold – those that do not, according to manual categorization by the authors.

2) Collocation Analysis: In our second methodology, we apply the likelihood ratio test, typically used in word collocation discovery in natural language processing [46], to our problem.

We model our sensitive word discovery problem as a collocation discovery problem, where instead of attempting to discover a word's collocation with another word, we look for significant collocations between a word and a label – "anonymous" or "nonanonymous", in a corpus obtained by converting each occurrence of a word w in an anonymous answer into an instance of w with label "anonymous", and each occurrence of w in a nonanonymous answer – into an instance of w with label "nonanonymous". The likelihood ratio test with each label then quantitatively evaluates two alternative hypotheses – the word being independent or dependent of the label, with the log likelihood ratio of the maximum likelihood estimates of those hypotheses enabling ranking of the words (and their co-occurrence with the label) by their significance.

As is standard in NLP [46], we rank collocations in the decreasing order of -2 times the log of their likelihood ratios. Since we are interested in identifying sensitive words, we present the top 100 words co-occurring with the "anonymous" label in Figure 6.

C. Discussion of Findings

As the two methodologies rely on different underlying principles, a top ranked word identified using one methodology might not appear in the top 100 words obtained using the other methodology. However, this does not mean that the second method did not identify any correlation between the specific word and anonymity. In fact, though the ordering is different, we observe a significant overlap among the words identified as anonymity indicators by the two methods. Even among the top 100 words listed in Figures 5 and 6, there are several overlaps, such as *transgender*, *proverbs*, *verifone*, *leviticus*, etc.

As is evident from Figures 5 and 6, the proportion of words that are not typically considered sensitive among those identified as sensitive via our data-driven analysis is quite high. We manually group the words not typically considered sensitive and identify several noteworthy themes:

- Law & Government, such as *sue, witness*
- Companies, such as *verifone, zynga, quora, square, acquiring, palantir*
- Education and Educational Institutions, such as *admissions, colleges, graduates, faculty, dorm, academics, committee*
- Relationships, such as *relatives, grandfather, neighbors, grandmother, parents, followers, friends, customers*
- Emotions & Emotional Experiences, such as *apology, laughing, ashamed, aggressively, affirmative, secretly, feel, denied*
- Career, such as *modeling, astronauts, officers, admins*
- Science, such as *boson, api, site, technical, science, orbiter*
- Arts & Entertainment, such as *fiction, story*
- Travel & Transportation, such as *gate*
- Social Issues, such as *immigration*
- Food & Drink, such as *cheese, beef, spoon*
- People Qualities, such as *gorgeous, righteous, faithful, forgiveness, morality, stereotype*

Many of these themes echo the ones identified in Section IV and described in Table II, with the exception of the last – related to *People Qualities*. There were no analogues for these in context and topic analyses likely due to the absence of context and topic labels conveying this theme among those created by the Quora moderators.

As in the previous section, our findings support the hypothesis that sensitivity is quite nuanced, and not limited to the typically considered sensitive topics and words. Concretely, among the top 200 words identified using the above methodologies (100 words from each technique), nearly 73% of words, evoking the themes of emotions, relationships, career, etc., would be missed if we relied only on the conventional assumptions.

Not all the words identified as characteristic of anonymous answers, and therefore potentially sensitive, carry a negative connotation. There are several positive words, such as: *laughing, gorgeous, righteous, faithful, forgiveness*, and several neutral words, such as *acquiring, feel, admissions, committee, bidding*. This suggests that purely sentiment analysis-based methods [47] that rely on the sentiment of the item being shared would not be successful at predicting the item's sensitivity.

Finally, besides serving as an additional confirmation of the hypothesis that sensitivity is nuanced, for which we found evidence via the analysis of topics and contexts in Section IV, the ability to build a vocabulary of potentially sensitive words is valuable in its own right. For example, in scenarios when users are sharing posts for which an accurate topic inference is not feasible (e.g., due to the short length of a post or lack of time or resources for manual labeling of its topic), having a vocabulary of potentially sensitive words for that application can power a cheap and easy-to-implement "Are you sure?"-type feature with high potential gain for user privacy.

VI. Towards Automated Sensitivity Prediction

In this section, we explore the possibility of training a machine learning classifier capable of warning users when they are about to follow a potentially sensitive question or to share or disclose something sensitive.

A. Question Sensitivity Prediction

To evaluate the possibility of predicting a question's potential sensitivity, we consider questions from contexts identified in Section IV-A2 whose answer anonymity ratio, $A(C)$, exceeds the average by 2 standard deviations. We further limit the set of questions to those 15,466 that have at least 6 answers, since our goal is to predict a question's sensitivity, and an accurate computation of the anonymity rate among the question's answers is unlikely for questions with few answers. We label a question as sensitive if the fraction of anonymous answers to its total answers is at least 0.32, i.e., 2 standard deviations above the average. The label was chosen in such a way as to roughly correspond to a 95% confidence interval [43].

Following the common machine learning practice, we randomly partition the data into two datasets: one for training and one for evaluation. The evaluation dataset consists of 1,000 questions in order to allow for a 0.1% precision in the evaluation. The training dataset contains the remaining 14,466 questions. We note that given our question sensitivity labeling, 21.5% of the questions in the evaluation dataset are considered sensitive, which establishes the baseline at 78.5%[10].

We experiment with soft-margin classifiers, linear and SVM classifiers, as they have been shown to be the most effective on NLP tasks that involve short text, such as Twitter sentiment analysis [47]. We use exhaustive search to evaluate the best method to convert the words in the dataset into features (e.g., with or without stop word removal, with or without stemming, using unigrams or bigrams, etc.), converging on *no stop word removal, no stemming, and use of bigrams* as the transformation that yields the best accuracy when used in conjunction with a linear classifier.[11] We experimented with four distinct types of

[10] An algorithm that always predicts that the question is not sensitive will achieve a 78.5% accuracy.

[11] We did not perform an analogous exhaustive search for the SVM classifier due to its prohibitive computational cost.

the bigram feature representations, namely: *binary, occurrence count, term frequency, and TF-IDF*, and concluded that the frequency representation works best. For the linear classifier, we tested various regularization modes, including L1 and L2. For the SVM classifier with an RBF kernel we performed a grid search to determine the optimal gamma and cost.

Table V presents the outcome of attempts to predict a question's sensitivity when each of the trained models is tested on the evaluation set. Overall, the best accuracy achieved is 80.4%, which represents a slight improvement relative to the baseline of 78.5%. Even with a small training sample and noise due to "Search Engine Privacy", our machine learning predictions of question sensitivity outperform the baseline. However, our results also suggest that relying purely on the content may not be sufficient and more information needs to be factored when evaluating the potential sensitivity of sharing something. We discuss several candidates for additional information, such as a person-specific sensitivity measure and the nuance of sensitivity depending on a person in Sections VII-A and VII-B.

Algorithm	Parameters	Accuracy
Linear classifier	–	80.4%
SVM linear kernel	c=0.0029	79.9%
SVM RBF kernel	c=850 g=0.01	80.2%

TABLE V
PERFORMANCE OF ALGORITHMS PREDICTING QUESTION SENSITIVITY

B. Answer Sensitivity Prediction

We run a set of experiments similar to the ones described in the previous section in order to assess whether it is possible to predict the sensitivity of an answer from its context and content. We limit our consideration to answers that contain at least 80 characters, which significantly decreases the number of answers, and experiment with two datasets. The first one, *S*, contains 3,660 answers to the questions that were labeled as sensitive in the question sensitivity experiment above. The second one, *A*, contains 151,825 answers to questions from the 1,525 contexts analyzed in Section IV-A1. As above, we randomly partition our data into a training and evaluation sets, with 1,000 answers in the evaluation datasets to allow for a 0.1% precision in the evaluation.

Algorithm	Parameters	Accuracy
Linear classifier	L1	62.3%
SVM linear kernel	c=385	63.1%
SVM RBF kernel	c=2 g= 0.00195	61.7%

TABLE VI
PERFORMANCE OF ALGORITHMS PREDICTING ANSWER SENSITIVITY, *S* CORPUS

Class	Precision	Recall
Anonymous	0.63	0.22
Non-anonymous	0.61	0.90

TABLE VII
PRECISION AND RECALL FOR THE VARIOUS CLASSES, *S* CORPUS

In the evaluation subset of *S*, the fraction of anonymous answers is 42.2%, setting a 57.8% baseline (using an algorithm that always predicts an answer will be non-anonymous). Table VI reports the performance accuracy of our answer anonymity predictor for the evaluation part of S, with the best algorithm[12] achieving an accuracy of 63.1%, which is 5.3% above the baseline. When evaluating precision and recall, reported in Table VII, the following conclusions emerge: first, predictions of anonymous and non-anonymous class have roughly the same precision which suggests that content provides information in both directions. Second, the weakest part of the prediction is the recall for the anonymous class: barely 2 out of 10 anonymous answers are correctly classified by the algorithm. This indicates that the biggest area of potential improvement lies in finding additional features to improve anonymous recall.

In the evaluation subset of *A*, the fraction of anonymous answers is 16.5%, setting a 83.5% baseline. Table VIII reports the performance accuracy of our answer anonymity predictor using the Linear classifier[13], with the algorithm achieving 88.0% accuracy, which is 4.5% above the baseline. As was the case in question sensitivity prediction, our answer sensitivity prediction results are able to beat the baseline performance even when given a small training set and in the presence of noise due to "Search Engine Privacy". The results highlight another important direction for improving classification quality: the need for additional training data, as the hypothesis that the quality of the prediction will improve with increase in the amount of data available for training is supported by the observation that our performance is better on the larger corpus, *A*, than on *S*.

Algorithm	Parameters	Accuracy
Linear classifier	L1	88.0%

TABLE VIII
PERFORMANCE OF ALGORITHM PREDICTING ANSWER SENSITIVITY, *A* CORPUS

Overall the experiments related to sensitivity prediction support our hypothesis that it is possible to use a data-driven approach of learning based on users' use of privacy-enhancing features, in order to provide better privacy protections for them. On the other hand, the accuracies of our classifiers also strongly suggest that predicting what is sensitive is a complex and nuanced problem that could benefit from additional features and better training data.

VII. DISCUSSION

A. Limitations of the Study due to Dataset Choice

The dataset we collected and used for our study of content sensitivity has several limitations, with implications for ability to generalize the conclusions made on its basis to other populations and other services and for the kind of statistical analyses and machine learning models that are feasible to perform on it.

Firstly, although Quora has a real name policy and many users answer questions on Quora in order to build their

[12]Removing stop words, performing stemming, using unigrams and representing words as binary features.

[13]SVM kernel models were not built due to their prohibitive computational costs on such a large corpus.

reputations as an expert in certain topics, and therefore, have an incentive to use a real name, some may be creating accounts using names other than their real one. Answering using an account with a fake name is analogous to answering anonymously from the perspective of risks we consider; hence, although anecdotal evidence suggests most users use their real name[14], our findings are limited by the extent to which Quora succeeds in enforcing the real names policy.

Secondly, although the Quora user base is fairly large and diverse[15], it may not be representative of users of other Internet services. Furthermore, the true privacy paranoids are unlikely to post on Quora or on any other online service. Therefore, our inferences can be effectively applied to improve the privacy of Quora's products, and can serve as a starting point for discussion on sensitivity, but would need additional service-specific research in order to be properly generalized to other services and other populations of users.

Thirdly, as discussed in Section III-C, the reasons for exercising anonymity choices on Quora may vary, and are not limited to data sensitivity. However, both previous work on user regrets about posting online [20] and Quora users' self-report on usage of anonymity [40] suggest that content sensitivity may be one of the significant motivating factors. Therefore, although one certainly cannot equate anonymity with sensitivity, we believe that anonymity is a strong indicator of potential sensitivity and our findings could serve as a starting point for further research on the topic.

Fourthly, unlike Quora itself, we do not have a user-level view of each user's anonymous and non-anonymous answers. Our inability to include user-specific features, such as gender or tendency for anonymous answering, likely significantly hampers the quality of the anonymity predictors we can build[16]. In practice, Quora has access to such information and would not be subject to the same limitations were it to attempt to learn privacy preferences based on its data or build features that could help prevent regret. Furthermore, lack of a user-level view prevents us from studying the potential differences in preferences due to gender, age, location, etc.

Finally, our dataset quality is limited by the quality and reach of the crawler we used. We cannot be sure that we collected a complete snapshot of Quora, that our parsing of the question page was perfect[17], or that the access we have to the followers of a question through the follower grid is representative of all its followers. Another limitation due to the crawler used relates to the *Search Engine Privacy* feature of Quora. The inferences we make are based on both truly anonymously written answers and those made anonymous due to the search engine privacy setting of the writing user. As described in Section IV-B2, we mitigate this limitation by choosing analyses whose inferences are minimally affected by such noise. We limit our word analyses and some of our machine learning analyses to data from contexts which exhibit elevated level of anonymity – firmly placing them above noise that may be due to search engine privacy.

In spite of these limitations, we are able to make informative inferences and develop sensitivity predictors which outperform the baseline prediction rates. This suggests that in practice, the service providers who are not constrained by the limitations we face, should be able to both better understand their users' privacy preferences and build predictors that enable them to improve users' privacy related experiences through introduction of appropriate nudges or defaults.

B. Content Sensitivity is Subjective

As pointed out by privacy experts in [6], determining content sensitivity is a complex problem. Content sensitivity depends not only on the content but also on the context, i.e., *who* is sharing the information and *when, where* and with *whom* they are sharing it, along with *what* they are sharing. Individuals may have widely differing anonymity and sensitivity preferences, depending on their personalities, cultural or religious backgrounds, experiences, etc. Consider the following examples of Quora questions and answers that illustrate that individual people may be making choices that differ from those that would be expected from most users:

- The question, "*Selfishness: What is the most selfish thing you have ever done?*"[18], has 12 (10 without search engine privacy) anonymous answers out of 18 total answers. However, one user gave the following very personal answer non-anonymously, and even provided a link to her Facebook account[19]: "*Thought that my husband and 2 young children could wait a year while I enjoyed, for the first time in my life, my job. At the end of the year, my marriage was in a shambles, and my eldest daughter was dead.*"
- The question, "*Why do homeless people wear so much clothing?*"[20], has 6 (2 without search engine privacy) anonymous answers out of 9 total answers. The following answer was provided anonymously, though there isn't anything obviously sensitive in it – "*I always assumed the reason they usually wear clothing in layers is because they have no storage facility to stash them. They always say: dress in layers in SF. Seriously, it can be 30 degrees in the morning (or colder) and 70 in the afternoon. In addition the extra layers are versatile, they can double as blankets and pillows. Also, many homeless people have issues with hoarding. Obviously you can't be a hoarder if*

[14]Many users link to their Facebook and Twitter accounts in their Quora profiles.

[15]A recent press interview suggests that Quora has been experiencing a healthy user growth in 2012-2013 [48]. Statistics provided by web traffic analytics companies Alexa [49], Compete [50], and trafficestimate [51], estimate that Quora has ∼1 million unique monthly visitors, and ∼30 million total monthly visits. The user base is dominated by visitors from India and the United States, who together account for more than 60% of the total traffic.

[16]In a related scenario of analyzing online content, [52] finds that author features have a strong discriminative power.

[17]We observed several answers that were blurred out with images or only partially collected by the crawl, which we omitted.

[18]https://www.quora.com/Selfishness/What-is-the-most-selfish-thing-you-have-ever-done

[19]We believe this user is not using her real name.

[20]https://www.quora.com/Homelessness/Why-do-homeless-people-wear-so-much-clothing

you are homeless but frequently they "collect" stuff and hang on to it."

- A question related to murders, "*What does it feel like to murder someone?*"[21], may be expected to have many anonymous answers. However, only 1 out of the 9 answers for it is anonymous.
- The question trying to understand reasons for anonymous answers, "*What drives people to contribute anonymous answers on Quora?*"[22], also contains many anonymous answers – 21 (17 without search engine privacy) of the 27.

C. Correlation with a User Survey

We initiate a study that aims to compare our behavioral data-driven findings with survey-based ones, via a short user survey using Google Consumer Surveys [12], a new public tool that enables anyone to quickly and cheaply run surveys online. To provide an (imperfect) parallel with our study of sensitivity based on Quora anonymity choices, we posed the questions:

1) Of the following topics, which ones would you be comfortable writing about online using your **real name**?
2) Of the following topics, which ones would you be comfortable writing about online **anonymously**?

The topics included in the choices were: *Prostitution, Recreational Drugs, Depression, Friendship, Government Leaders and Politicians, Religion and Beliefs* (high anonymity ratio according to analysis in Section IV), and *Mobile Phone* and *Superhero Films* (low anonymity ratio). Selection of more than one answer was permitted, along with the option "None of the above". The topic presentation order was randomized.

Figure 7 presents the results based on 1,500 responses received for each question, with respondents chosen to be representative of the US Internet population (via the quota method provided by [12]). The results highlight the difficulty of eliciting user privacy preferences and sensitivities, as although participation in online sharing platforms such as Twitter and Tumblr is skyrocketing, the vast majority of respondents indicated they would not be comfortable writing online even about the seemingly innocuous topic of *Mobile Phones*. On the other hand, they give support to the validity and promise of our proposed approach: firstly, for most topics, the respondents' indicated comfort level is higher when assuming they'd be answering anonymously rather than with their real name, supporting our hypothesis that anonymity choices may be indicative of sensitivity. Secondly, the ranking of topics by percentage of respondents who'd be comfortable writing about it is different, but not radically so, from the one derived in Section IV. This suggests that behavioral-data driven analyses and research based on user surveys could complement and support each other.

[21]https://www.quora.com/What-Does-It-Feel-Like-to-X/What-does-it-feel-like-to-murder-someone

[22]https://www.quora.com/What-drives-people-to-contribute-anonymous-answers-on-Quora/

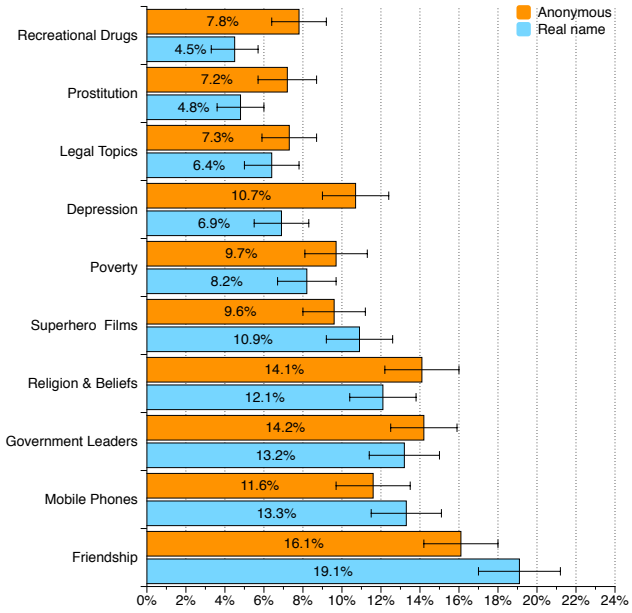

Fig. 7. Percentage of respondents who indicated they'd be comfortable writing about a topic anonymously vs with their real name.

Much more extensive research, which is beyond the scope of this paper, is needed in order to understand the relative merits of survey-based and behavioral data-based approaches to eliciting privacy preferences and learning about data sensitivity. For example, our survey results present some supportive evidence that gender may affect sensitivity perception and sharing comfort. The survey results suggest that women may be more comfortable writing about *Friendship* with their real names than men – *20.5%* vs *12.8%* (p-value 0.02), while men may be more willing to discuss *Prostitution* anonymously than women – *3.8/7* vs *3.3/7* (p-value 0.02). We leave an in-depth investigation of factors affecting sensitivity perceptions and comparison between approaches to future work.

VIII. Related Work

1) Understanding User Privacy Preferences via Surveys: Online surveys and personal interviews are the predominant method currently used for learning about user privacy expectations, and identifying problems in existing privacy-related offerings. They have been used to examine privacy preferences in e-commerce [53], understanding concerns related to disclosure of health information online [54], learn about privacy expectations in location based systems [55], [56], [57] and in social networks [58], [59]. Surveys have also been used to understand why people regret posting online and identify sensitive topic themes [20], or to capture user preferences about anonymity and data sensitivity online [60].

Survey-based learning of privacy preferences is more difficult and more expensive than our proposed approach when it needs to be adopted to a specific product, to potential cultural, location, language or demographic-based differences in preferences among users, or to shift in preferences over time. However, they are a useful complementary approach to that of the behavioral

data-driven one we propose, and can help in formulating questions or hypotheses, providing illustrative examples, etc. The survey-based work with the goal most similar to ours by Wang et al [20] identifies anecdotal examples of user sharing regrets from many of the same themes we discover: personal and family issues, religion and politics, work and company, sex, and illegal drug use.

2) Understanding Privacy Risks of Product Usage: Understanding how users use product features with potential privacy implications can help identify and mitigate potential privacy risks. [27] studies the security and privacy of private browsing modes in web browsers and how installed plugins and browser extensions may undermine the security of private browsing. [61] performs a large-scale quantitative analysis of *delete tweet* feature in Twitter, and highlights the privacy ramifications of availability of deleted tweets outside of Twitter. [62] performs a quantitative study of the amount and kinds of personally identifiable information disclosed in Twitter messages and potential privacy implications of such disclosures.

3) Minimizing Regret When Sharing: Recent work has begun to make steps towards understanding and minimizing user regret when sharing on social networks. Via a combination of surveys, interviews and log analyses, [26] investigates how active users organize and select audiences for sharing content on the Google+ social network. [63] studies deleted *bullying* tweets on Twitter, and proposes building a *regrettable tweets* predictor to warn users if a tweet might cause user regret later. [64] proposes a template for the design of a social networking privacy wizard, that builds a machine learning classier to learn from a small sample of user's privacy preferences, and then uses this classifier to configure the user's privacy settings automatically. Our work, particularly in Section VI, complements these efforts and provides evidence in support of feasibility of product features aimed at minimizing regret.

4) Other Work Studying Quora: Our work is not the first in using Quora to gain insights into individuals' behavior. For example, [42] makes an attempt to understand what drives the growth of *question-and-answer* websites like Quora, and how does it attract and motivate visitors to contribute, while [41] studies the reputation mechanisms in Quora, i.e., how users judge the authoritativeness of other users and content, build reputation, and identify and promote high quality content.

IX. Conclusion

We performed a large-scale analysis of user anonymity choices during their activity on Quora, a popular question-and-answer site, to determine user content sensitivity preferences. We enumerated different analysis methodologies on contexts, topics, and words, and identified sensitive themes that are not included in the common characterizations of content sensitivity. We built several machine learning models able to predict user anonymity choices better than a fixed guess would, suggesting a possibility for features improving user experience.

Although much more in-depth research is needed, our work makes the first step to show that data-driven analysis of users' use of privacy-enhancing product features can improve our ability to understand user privacy preferences and expectations at scale, and enable online services to develop policies and features that better protect their users.

X. Acknowledgements

We are grateful to Vasyl Pihur for help with statistical analyses and valuable feedback on the work. We thank Pern Hui Chia, Dorothy Chou, and Jessica Staddon for useful feedback on the paper drafts, and Úlfar Erlingsson for the title suggestion. We thank the anonymous reviewers for thoughtful comments and suggestions.

References

[1] "Harlem shake: Baauer cashes in on viral video's massive YouTube success," http://www.theguardian.com/technology/2013/feb/19/harlem-shake-baauer-youtube-success, Accessed: Nov 13, 2013.

[2] S. Choney, "Kony video proves social media's role as youth news source: Pew," http://www.nbcnews.com/technology/kony-video-proves-social-medias-role-youth-news-source-pew-455365, Accessed: Mar 14, 2014.

[3] "When the Self-Published Authors Take Over, What Will Publishers Do?" http://www.forbes.com/sites/jeremygreenfield/2013/04/30/when-the-self-published-authors-take-over-what-will-publishers-do/.

[4] "Facebook and Twitter Postings Cost CFO His Job," http://online.wsj.com/articles/SB10001424052702303505504577404542168061590, Accessed: Nov 11, 2013.

[5] "The Red Cross' Rogue Tweet: #gettngslizzerd On Dogfish Head's Midas Touch," http://www.huffingtonpost.com/2011/02/16/red-cross-rogue-tweet_n_824114.html, Accessed: Nov 11, 2013.

[6] "First issue of CNIL IP Reports: "Privacy towards 2020" – 42 experts share their visions of the future of privacy with the French regulation authority," http://www.cnil.fr/linstitution/actualite/article/article/first-issue-of-cnil-ip-reports-privacy-towards-2020-42-experts-share-their-visions-of-the, Accessed: Nov 11, 2013.

[7] "Quora," http://quora.com/.

[8] "Google Policies and Principles," http://www.google.com/policies/privacy/key-terms/#toc-terms-info, Accessed: Nov 9, 2013.

[9] "Microsoft Advertising Creative Acceptance Policy Guide," http://advertising.microsoft.com/en-uk/WWDocs/User/display/cl/content_standard/2007/global/Microsoft-Advertising-Creative-Acceptance-Policy-Guide.pdf, Accessed: Nov 10, 2013.

[10] "Facebook Advertising Guidelines," https://www.facebook.com/ad_guidelines.php, Accessed: Nov 9, 2013.

[11] "CNIL: Questionnaire to Google," http://www.cnil.fr/fileadmin/documents/La_CNIL/actualite/questionnaire_to_Google-2012-03-16.pdf, Accessed: Nov 10, 2013.

[12] "Google Consumer Surveys," http://www.google.com/insights/consumersurveys/home, Accessed: Nov 11, 2013.

[13] "Google Privacy Policy," http://www.google.com/policies/privacy/, Accessed: Nov 9, 2013.

[14] S. L. Huang, "Removing feed stories about views," http://blog.quora.com/Removing-Feed-Stories-about-Views, Accessed: Nov 9, 2013.

[15] "CNIL," http://www.cnil.fr/english/the-cnil/constitution-and-composition.

[16] "Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data," http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:en:HTML, Accessed: Nov 13, 2013.

[17] J. V. Grove, "Twitter your way to getting robbed," http://mashable.com/2009/06/01/twitter-related-burglary/, Accessed: Nov 5, 2013.

[18] G. A. Fowler, "When the most personal secrets get outed on Facebook," http://online.wsj.com/articles/SB10000872396390444165804578008740578200224, Accessed: Nov 4, 2013.

[19] K. Hill, "Oops. Mark Zuckerberg's sister has a private Facebook photo go public." http://www.forbes.com/sites/kashmirhill/2012/12/26/oops-mark-zuckerbergs-sister-has-a-private-facebook-photo-go-public/, Accessed: Nov 4, 2013.

[20] Y. Wang, G. Norcie, S. Komanduri, A. Acquisti, P. G. Leon, and L. F. Cranor, "'I regretted the minute I pressed share': a qualitative study of regrets on Facebook," in *Proceedings of the Seventh Symposium on Usable Privacy and Security (SOUPS)*, 2011, pp. 10:1–10:16.

[21] N. Singer, "They Loved Your G.P.A. Then They Saw Your Tweets," http://www.nytimes.com/2013/11/10/business/they-loved-your-gpa-then-they-saw-your-tweets.html, Accessed: Nov 13, 2013.

[22] L. Kwoh, "Beware: Potential Employers Are Watching You," http://online.wsj.com/articles/SB10000872396390443759504577631410093879278, Accessed: Nov 13, 2013.

[23] M. Miller, "Facebook Graph Search Will Find You The Perfect Date," http://www.forbes.com/sites/mattmiller/2013/01/29/facebook-graph-search-date/, Accessed: Nov 13, 2013.

[24] C. Breen, "Why I left Facebook," http://www.pcworld.com/article/196237/article.html, Accessed: Nov 4, 2013.

[25] J. V. Grove, "Why teens are tiring of Facebook," http://news.cnet.com/8301-1023_3-57572154-93/why-teens-are-tiring-of-facebook/, Accessed: Nov 4, 2013.

[26] S. Kairam, M. Brzozowski, D. Huffaker, and E. Chi, "Talking in circles: selective sharing in Google+," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI)*, 2012, pp. 1065–1074.

[27] G. Aggarwal, E. Bursztein, C. Jackson, and D. Boneh, "An analysis of private browsing modes in modern browsers," in *Proceedings of the 19th USENIX Conference on Security*, 2010, pp. 79–94.

[28] "Tor: Anonymity Online," https://www.torproject.org/.

[29] C.-Y. Chow, M. F. Mokbel, and X. Liu, "Spatial cloaking for anonymous location-based services in mobile peer-to-peer environments," *Geoinformatica*, vol. 15, no. 2, pp. 351–380, Apr. 2011.

[30] M. S. Bernstein, A. Monroy-Hernández, D. Harry, P. André, K. Panovich, and G. G. Vargas, "4chan and/b: An analysis of anonymity and ephemerality in a large online community." in *ICWSM*, 2011.

[31] M. McFarland, "Snapchat is thriving and that's a great sign for your privacy," http://www.washingtonpost.com/blogs/innovations/wp/2013/10/28/snapchat-is-thriving-and-thats-a-great-sign-for-your-privacy/, Accessed: Nov 14, 2013.

[32] Y. Lelkes, J. A. Krosnick, D. M. Marx, C. M. Judd, and B. Park, "Complete anonymity compromises the accuracy of self-reports," *Journal of Experimental Social Psychology*, vol. 48, no. 6, pp. 1291–1299, 2012.

[33] L. M. Jessup, T. Connolly, and J. Galegher, "The effects of anonymity on gdss group process with an idea-generating task," *MIS Q.*, vol. 14, no. 3, pp. 313–321, Sep. 1990.

[34] T. Postmes, R. Spears, K. Sakhel, and D. de Groot, "Social influence in computer-mediated communication: The effects of anonymity on group behavior," *Personality and Social Psychology Bulletin*, vol. 27, no. 10, pp. 1243–1254, 2001.

[35] R. Kang, S. Brown, and S. Kiesler, "Why do people seek anonymity on the internet?: informing policy and design," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI)*, 2013, pp. 2657–2666.

[36] "Yahoo! Answers," http://answers.yahoo.com/.

[37] "Quora Policies and Guidelines: Do I have to use my real name on Quora? what is Quora's Real Names policy?" http://www.quora.com/Quora-Policies-and-Guidelines/Do-I-have-to-use-my-real-name-on-Quora-What-is-Quoras-Real-Names-policy, Accessed: Nov 5, 2013.

[38] "How many anonymous answers have you posted on Quora," https://www.quora.com/Anonymity-on-Quora/How-many-Anonymous-answers-have-you-posted-on-Quora-And-if-tomorrow-due-to-some-glitch-those-answers-are-posted-under-your-name-what-would-be-your-reaction, Accessed: Mar 13, 2014.

[39] "Quora and Search Engines: What does enabling "Search Engine Privacy" under Account Settings do?" https://www.quora.com/Quora-and-Search-Engines/What-does-enabling-Search-Engine-Privacy-under-Account-Settings-do, Accessed: Nov 5, 2013.

[40] "What drives people to contribute anonymous answers on Quora?" https://www.quora.com/What-drives-people-to-contribute-anonymous-answers-on-Quora/, Accessed: Nov 5, 2013.

[41] S. A. Paul, L. Hong, and E. H. Chi, "Who is authoritative? Understanding reputation mechanisms in Quora," in *Collective Intelligence*, 2012.

[42] G. Wang, K. Gill, M. Mohanlal, H. Zheng, and B. Y. Zhao, "Wisdom in the social crowd: an analysis of Quora," in *Proceedings of the 22nd International Conference on World Wide Web (WWW)*, 2013.

[43] F. Andrews, L. Klem, S. Institute, and P. O'Malley, *Selecting Statistical Techniques for Social Science Data: A Guide for SAS Users*. SAS Institute, 1998.

[44] "Premier by mobileworks," https://premier.mobileworks.com/.

[45] W. E. Mackay, "Triggers and barriers to customizing software," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI)*, 1991, pp. 153–160.

[46] C. D. Manning and H. Schütze, *Foundations of statistical natural language processing*. MIT Press, 1999.

[47] A. Agarwal, B. Xie, I. Vovsha, O. Rambow, and R. Passonneau, "Sentiment analysis of Twitter data," in *Proceedings of the Workshop on Languages in Social Media (LSM)*. Association for Computational Linguistics, 2011, pp. 30–38.

[48] A. Tsotsis, "Quora Grew More Than 3X Across All Metrics In The Past Year," http://techcrunch.com/2013/05/28/quora-grows-more-than-3x-across-all-metrics-in-the-past-year/.

[49] "Alexa – How popular is quora.com?" http://www.alexa.com/siteinfo/quora.com, Accessed: Mar 7, 2014.

[50] "Compete – Quora site info," https://siteanalytics.compete.com/quora.com/#.UxmRj_mSzpo, Accessed: Mar 7, 2014.

[51] "trafficestimate – quora.com Website Traffic and Information," http://www.trafficestimate.com/quora.com, Accessed: Mar 7, 2014.

[52] B. Gelley and T. Suel, "Automated decision support for human tasks in a collaborative system: The case of deletion in Wikipedia." in *WikiSym*, 2013.

[53] M. S. Ackerman, L. F. Cranor, and J. Reagle, "Privacy in e-commerce: Examining user scenarios and privacy preferences," in *Proceedings of the 1st ACM Conference on Electronic Commerce (EC)*, 1999, pp. 1–8.

[54] G. Bansal, F. . Zahedi, and D. Gefen, "The impact of personal dispositions on information sensitivity, privacy concern and trust in disclosing health information online," *Decision Support Systems*, vol. 49, no. 2, pp. 138–150, 2010.

[55] S. Wilson, J. Cranshaw, N. Sadeh, A. Acquisti, L. F. Cranor, J. Springfield, S. Y. Jeong, and A. Balasubramanian, "Privacy manipulation and acclimation in a location sharing application," in *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp)*, 2013, pp. 549–558.

[56] E. Toch and I. Levi, "Locality and privacy in people-nearby applications," in *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp)*, 2013, pp. 539–548.

[57] S. Patil, G. Norcie, A. Kapadia, and A. J. Lee, "Reasons, rewards, regrets: privacy considerations in location sharing as an interactive practice," in *Proceedings of the Eighth Symposium on Usable Privacy and Security (SOUPS)*, 2012, pp. 5:1–5:15.

[58] L. Bauer, L. F. Cranor, S. Komanduri, M. L. Mazurek, M. K. Reiter, M. Sleeper, and B. Ur, "The post anachronism: The temporal dimension of Facebook privacy," in *Proceedings of the 12th Annual Workshop on Privacy in the Electronic Society*, Nov. 2013.

[59] Y. Liu, K. P. Gummadi, B. Krishnamurthy, and A. Mislove, "Analyzing Facebook privacy settings: user expectations vs. reality," in *Proceedings of the 11th ACM Internet Measurement Conference (IMC)*, 2011.

[60] L. Rainie, S. Kiesler, R. Kang, and M. Madden, "Anonymity, Privacy, and Security Online: Part 4: How Users Feel About the Sensitivity of Certain Kinds of Data," http://pewinternet.org/Reports/2013/Anonymity-online/Main-Report/Part-4.aspx, Accessed on Nov 15th 2013.

[61] H. Almuhimedi, S. Wilson, B. Liu, N. Sadeh, and A. Acquisti, "Tweets are forever: a large-scale quantitative analysis of deleted tweets," in *Proceedings of the 2013 Conference on Computer Supported Cooperative Work (CSCW)*, 2013, pp. 897–908.

[62] L. Humphreys, P. Gill, and B. Krishnamurthy, *How much is too much? Privacy issues on Twitter*. ACM Press, 2010, pp. 1–29.

[63] J.-M. Xu, B. Burchfiel, X. Zhu, and A. Bellmore, "An examination of regret in bullying tweets." in *HLT-NAACL*, 2013, pp. 697–702.

[64] L. Fang and K. LeFevre, "Privacy wizards for social networking sites," in *Proceedings of the 19th International Conference on World Wide Web (WWW)*, 2010, pp. 351–360.