

“I Didn’t Sign Up for This!” Informed Consent in Social Network Research

Luke Hutton and Tristan Henderson

School of Computer Science
University of St Andrews, St Andrews, Fife, United Kingdom
{lh49, tmhh}@st-andrews.ac.uk

Abstract

The issue of whether, and how, to obtain informed consent for research studies that use social network data has recently come to the fore in some controversial cases. Determining how to acquire valid consent that meets the expectations of participants, while minimising the burden placed on them, remains an open problem. We apply Nissenbaum’s model of contextual integrity to the consent process, to study whether social norms of willingness to share social network data can be leveraged to avoid burdening participants with too many interventions, while still accurately capturing their own sharing intent. We find that for the 27.7% of our participants ($N = 109$) who conform to social norms, contextual integrity can be used to significantly reduce the time taken to capture their consent, while still maintaining accuracy. Our findings have implications for researchers conducting such studies who are looking to acquire informed consent without having to burden participants with many interactions.

Introduction

In recent years, many scientific disciplines have become interested in using data collected from social network sites (SNSs), such as Facebook and Twitter. Such data are considered a rich source for understanding social dynamics, from analysing discussion of sleep disorders (Powell et al. 2012), to exploring how bereaved people use Facebook to grieve (Getty et al. 2011); a comprehensive review of how Facebook has been used in social science research is provided by Wilson et al. (2012). Such research can be of high value, but raises myriad ethical questions. Studies leveraging social network data may do so without the people who originally published it knowing. With such studies conducted remotely with no direct contact between the researchers and the content creators whose data are used, it is uncertain whether they should be considered participants, and experimental procedures subject to the same scrutiny as other human subjects research. There is disagreement about whether data derived from public sources such as Twitter should be fair game for researchers, or whether repurposing such data for research violates the expectations of content creators.

The debate came to a head when researchers at Facebook published a study exploring emotional contagion through manipulation of Facebook’s News Feed product (Kramer, Guillory, and Hancock 2014). When it emerged that Facebook did not seek informed consent from participants, there was widespread condemnation of the ethical standards of the experiment (Hill 2014). Since the study was conducted, Facebook has added the word “research” to a clause in their Data Use Policy regarding “how we use the information we receive”, one of three policy documents users are asked to agree to before registering.¹ While Facebook considers this clause to be an acceptable proxy for gaining informed consent for individual studies, it betrays the expectation that participants engage in research knowing what information is collected, and for what purpose, which Warrell and Jacobsen (2014) contend is essential when conducting such studies.

The attraction of such “big data” research can be difficult to reconcile with the logistics of gaining informed consent from participants. Big data that are public are not necessarily fair game for research (boyd and Crawford 2012). Even if someone agrees to participate in such an experiment, they may not be willing to share all of their social network data, some of which could be considered highly sensitive. The state of the art is to largely adopt one of two strategies for gaining informed consent. Building on Luger and Rodden’s (2013) recommendations for considering consent in ubiquitous computing, we term these *secured* and *sustained* consent. *Secured consent* is the traditional checkbox or “end-user license agreement (EULA)”, where people provide consent at a single point in time to an unqualified collection of data, without respect to the nature or purpose of the study. While this approach is quick and trivial for participants, it risks violating their expectations if data are collected and processed against their wishes. At the other extreme, we consider *sustained consent*, where participants are continuously probed about their willingness to share discrete pieces of data over the course of an experiment, characterised by studies such as Sleeper et al.’s (2013) examination of self-censorship on Facebook, where participants chose when to share information with researchers over a period of time. Such a consent strategy is likely to engender more

¹Facebook Data Use Policy: [facebook.com/policy.php](https://www.facebook.com/policy.php)

support from participants who have actively chosen what to disclose, at the risk of burdening them with a large number of interventions which may cause frustration, and ultimately lead to attrition, a problem for all longitudinal studies.

Previous social network studies have involved the collection and processing of large quantities of data from the users of social network sites. In a recent study, we have surveyed the literature to find that the extent to which people consented to their data being used is generally not collected or reported, with only 28 of 505 (5.5%) of papers describing their consent procedures (Hutton and Henderson 2015). Therefore, we have little understanding as to whether participants would have consented to their data being used for such purposes, or appreciate the implications of exposing their data to researchers. Studies which rely on a single checkbox in a boilerplate consent form risk exposing participants to inappropriate levels of disclosure, with the benefit of reduced cognitive and temporal load on the participant. Conversely, studies which ask the participant to explicitly sanction the disclosure of each piece of data may better meet the expectations of participants, but at the cost of a significantly greater burden on the participant.

In this paper, we propose a third way of acquiring consent which aims to maximise the accuracy of the sustained approach, while achieving a low burden on participants, as with the secured approach. We use Nissenbaum's (2004) model of contextual integrity to leverage context-specific social norms of willingness to share social network data in an experiment which asks participants to consent to disclosing a subset of their Facebook data for a hypothetical study. By detecting whether users conform to such norms, we aim to reduce the number of times we need to explicitly ask for their consent, with a minimal loss in accuracy.

Specifically, we hypothesise the following:

- **H1:** Acquiring consent to share social network data with researchers using contextual integrity reduces the burden compared to current methods of acquiring explicit sustained consent, while more accurately reflecting people's intent than secured consent methods which involve no such interventions.
- **H2:** Acquiring consent with contextual integrity is as robust to temporal changes in willingness to share data as sustained consent.

This paper makes the following contributions:

- We develop the first means of quantifying norms for sharing social network data.
- We apply contextual integrity to the study of informed consent, with the first study to measure whether contextual integrity provides an acceptable trade-off between accuracy and the burden on people to acquire their consent.

Contextual integrity

Nissenbaum (2004) proposes contextual integrity as a theoretical framework for considering information privacy. She argues that information is not inherently public or private, but governed by context-specific norms, which determine

to whom it is appropriate for information to be transmitted to, and for what purpose. Individuals, organisations, and sections of society each have their expectations about what constitutes the appropriate flow of information, and any actor can perceive a privacy violation if these expectations are not met. For example, in order to receive a diagnosis for a medical condition, a patient accepts that they must communicate sensitive information about their health to a doctor. The information might generally be considered "private", but both actors have an understanding of the norms governing this sharing of information, and thus there is no privacy violation. If, however, that doctor was to subsequently gossip about this condition to someone else, the expectations of the patient have been violated, and so has their privacy.

This concept, while simple, has profound implications for how we consider information privacy. Our specific focus is on whether we can accurately infer context-specific norms by sampling people's willingness to share social network data. We then identify if people conform to such norms, and use the "wisdom of the crowd" as a proxy for individual disclosure decisions, thus reducing the time spent collecting explicit consent from participants. Our aim is to demonstrate the usefulness of contextual integrity in this domain by illustrating what proportion of the population are norm-conformant, and can benefit from such inferences.

Method

We conducted a study to investigate whether acquiring sustained informed consent based on contextual integrity performs better than two other strategies, summarised as:

- **C1 Secured consent:** Participants provide up-front consent to data acquisition in an informed consent form
- **C2 Contextual integrity consent:** Participants are asked explicitly about their willingness to share each piece of data, unless they clearly conform to or deviate from a social norm for sharing such data.
- **C3 Sustained consent:** Participants are asked explicitly about their willingness to share each piece of data.

Deriving willingness-to-share norms

To evaluate contextual integrity (C2), we first need to derive social norms so that we can detect conformity and deviation. To do so, we use a dataset from a previous study that examined people's willingness to share social network data with researchers (McNeilly, Hutton, and Henderson 2013). In the study, participants were asked which of 100 pieces of their Facebook data they were willing to share with researchers, of the following types:

- Location check-ins by the user
- Names of Facebook friends
- "Liked" Facebook pages
- Photos uploaded by the user
- Photo albums created by the user
- Biographical attributes
- Status updates

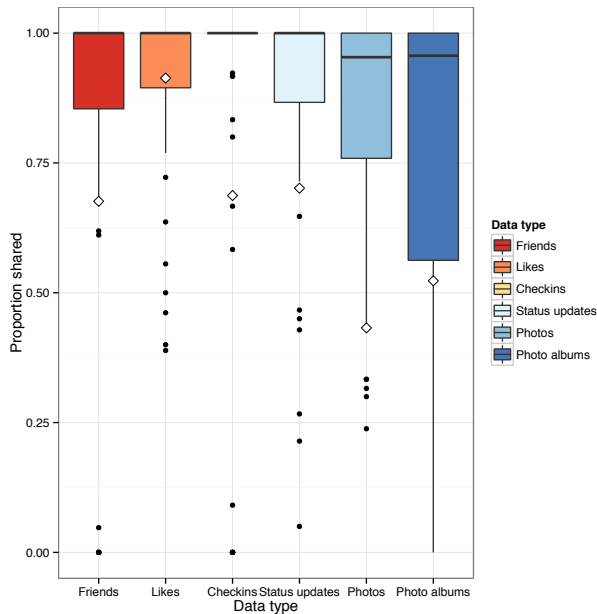


Figure 1: Rates of sharing of different types for participants in C3. The white diamonds represent the mean sharing rate of that type in the 2012 dataset, constituting the prevailing norm. Willingness to share data with researchers has increased on all fronts in this period.

We consider the proportion of shared content for each attribute to be the prevailing norm (Figure 1). We draw on these norms in our study to determine the extent to which our participants conform with them.

User study

Participants were recruited online (through advertisements in university mailing lists, Twitter, and Reddit) for a study which modelled the three consent strategies and evaluated their performance, with participants randomly assigned to one of these strategies. The study comprised two parts: *consent acquisition* and a *performance evaluation*.

In the consent acquisition phase, participants were asked about their willingness to share up to 100 of pieces of data randomly extracted from their Facebook profile with researchers conducting a hypothetical study into “how emotions spread on Facebook” (Figure 2). These data were of the types found in the norms dataset, with the exception of biographical attributes. The strategy the participant was assigned to affected the presentation of this step. The purpose of this step was to infer a “consent policy”, which represented the subset of the data the participant was willing to share, according to the rules of that strategy.

As we are unaware of other attempts to quantify the norms for sharing SNS data, we used a simple method for determining participants’ conformity to the norms derived from the source dataset. After each participant answered a question, we performed a series of chi-square equality of proportions tests, comparing the distribution of responses by the partic-

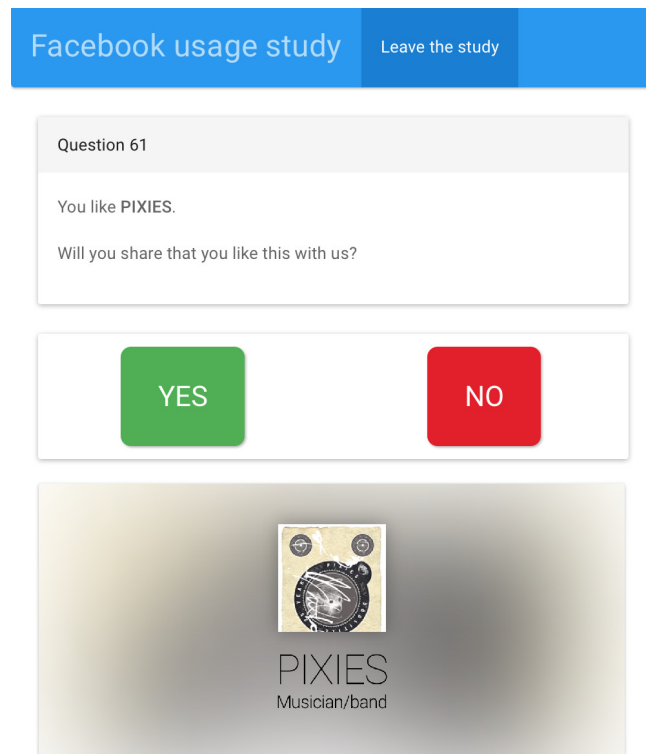


Figure 2: A screenshot of a question in the study. In this case, the participant is asked to share the fact that they like a Facebook page with the researchers.

ipant for each data type to the distribution of the norm. If $p \leq 0.1$, we considered the participant to conform to the norm, and all further questions of that type were removed. If $p \geq 0.9$, we considered the participant to deviate from the norm, and again all questions of that type were removed. As the chi-square test assumes there are at least five counts in each cell, we did not attempt to calculate norm conformity until the participant shared five items of a given attribute. Therefore, the method is not robust to those who share very small amounts of data. We chose to test at the 0.1 significance level, as early piloting of the method allowed a determination to be made about conformity within a small number of questions while maintaining accuracy. The results of this study will be used to vindicate this design decision.

Secured consent Participants in all conditions were asked to complete a boilerplate informed consent form, in accordance with current best practices and the ethical requirements of our institution. For participants in this condition, an additional question was injected in to the consent form, adapted from a clause in Facebook’s Data Use Policy: “I understand the researchers may use Facebook data they receive about me for data analysis, testing, and research.” For these participants, completing this form was treated as the participant giving consent to all 100 pieces of content being processed for the purpose of the study, forming the consent policy, and ending the consent acquisition phase.

Contextual integrity consent Participants were shown, in turn, each of the 100 pieces of data collected, and asked whether they were willing to share it with the researchers. This strategy aimed to reduce the number of questions, however, by constantly evaluating whether the participant conformed to the prevailing social norm of willingness to share that type of data. Each time a participant answered a question, their norm conformity was measured for each of the six data types they were asked about, removing redundant questions if conformity was established. After the participant completed this phase, the consent policy was based on the proportions of each data type the participant was willing to share. The final policy consisted of the content explicitly shared by the participant, augmented by additional content consistent with these proportions.

Sustained consent The presentation of this method was similar to the contextual integrity condition, in that all questions were shown in turn to the participant, but no questions were removed, as no norm conformity calculations were made. The consent policy contained the subset of attributes explicitly shared by the participant.

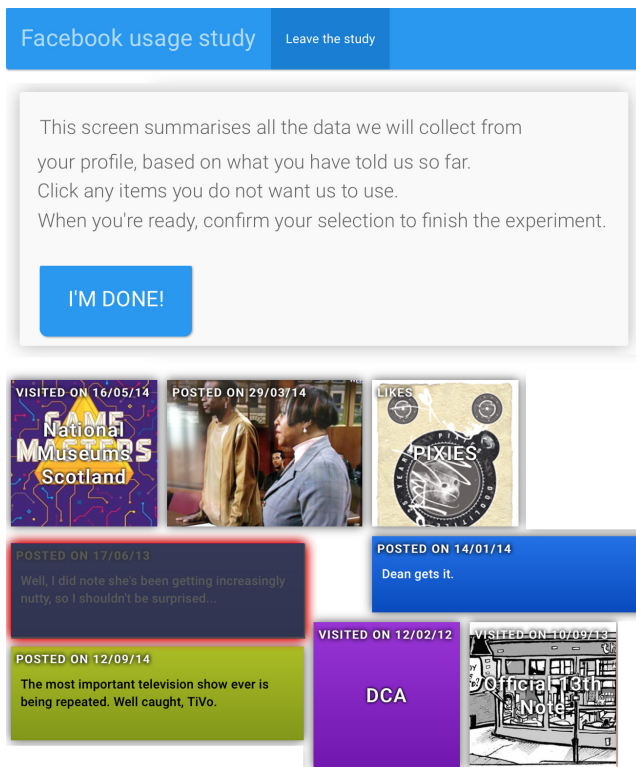


Figure 3: A screenshot of the performance evaluation step. Participants are shown the content collected according to their consent policy, and asked to remove any items they deem inappropriate to share. The proportion of items selected determines the accuracy of the method.

The second phase of the study, the performance evaluation, was the same for all conditions. Participants were shown all of the data in their consent policy, and asked to

click on any items which they do not wish to have shared (Figure 3).

A week after the study was completed, participants were asked to complete the study again. Assigned to the same condition, participants completed the same process, with a different random subset of data.

We evaluated our results using the following metrics:

- **Burden:** How long the participants spent participating in the study. This is measured as the percentage of potential questions that were presented to the participant. For secured consent this is 0%, as the method assumes that participants were willing to share everything. For sustained consent, this is 100%, as all participants were asked the maximum number of questions. The burden of the contextual integrity condition lies somewhere between these two extremes, with fewer questions asked depending on the participant's norm conformity throughout the study.
- **Accuracy:** How successfully the consent strategy meets the intent of users. This is measured as the percentage of data in the consent policy which the participant was also willing to share in the performance evaluation step. We expect the accuracy of the sustained consent condition to be very high as all participants explicitly chose which data to share, while in the secured condition, accuracy is likely to be much lower and more variable between participants, as they did not choose which content would be in the consent policy.
- **Robustness:** Previous work shows that willingness to share social network data is temporally sensitive (Bauer et al. 2013; Ayalon and Toch 2013), so we repeated the study after a week to observe the extent to which this effect manifests, and to determine which methods are more robust to it. The accuracy of using answers from the first week to predict responses to the second week provides this measure, as it will highlight whether each method is sensitive to changes in willingness to share information over time.

Ethical considerations

Before conducting the study, our experimental design was scrutinised and approved by the relevant ethics committee (our institution's equivalent of an Institutional Review Board). The experiment used our previously developed framework for privacy-sensitive handling of social network data (Hutton and Henderson 2013) which ensured only the data necessary to execute the study was collected, and that all data were appropriately sanitised and deleted during the lifecycle of the experiment.

Results

154 people began the study, of whom 109 completed participation in the first week. 71 of these participants also completed the second part of the study a week later. As shown in Table 1, participation was broadly equal across all three conditions. In our analysis, we consider the responses of all 109 participants, with only those who completed both

Condition	Started	Completed week 1	Completed week 2
Secured consent	41	32	22
Sustained consent	55	39	24
Contextual integrity consent	58	38	25

Table 1: Participants were assigned to one of the three conditions at random, and participation was broadly equal.

Category	Response	Study %	Facebook %
Gender	Male	33	43.6
	Female	66	46.3
Age	18-24	72.3	20.5
	25-34	21.8	24.7
	35-44	7	17.9
	45-54	2	14.2
	55+	0	13.7
Education	High School	3	20
	Undergraduate degree	66.3	32.1
	Postgraduate degree	31.7	3.6

Table 2: Comparison of demographics in our study to those of Facebook in the UK. Demographics are derived from data made available to Facebook advertisers, correct as of January 2015.²

weeks considered in our temporal analysis of the robustness of contextual integrity. Table 2 shows the demographics of our sample population, compared to data Facebook make available to their advertisers.² As 77.2% of our participants live in the UK, we compare our sample to the UK’s demographic make-up. A side-effect of primarily promoting the study in university mailing lists is that we oversample younger, university-educated people. 52.6% of participants were affiliated with our institution. To determine whether such participants would be more likely to share data with researchers at their own institution, and thus skew the results, we compared their overall willingness to share data with other participants, finding no differences.

Is there a relationship between burden and accuracy?

We are interested in the relationship between burden – the percentage of potential questions a participant is asked about their willingness to share data with researchers, and accuracy – whether inferences based on these questions satisfy the expectations of the individual.

As discussed earlier, participants in the *secured* consent condition were not asked any questions about their willingness to share individual pieces of data, representing a burden

of 0%. Conversely, participants in the *sustained* consent condition were asked whether they were willing to share each individual piece of data collected in the study, a potential burden of 100%.³ For users in the *contextual integrity* condition, we expected burden to lie somewhere in the middle. In the worst case, the burden would match that of sustained consent, however the more norm-conformant a participant was, the fewer questions they were asked.

We expected that participants in the *sustained* consent condition would yield the highest accuracy. As these participants were explicitly asked about their willingness to share each piece of data, the acquisition of these data should meet their expectations. Conversely, we expected the *secured* consent condition to exhibit lower, and more variable accuracy. As this method does not account for differences between participants’ willingness to share data, it is unlikely to meet most people’s expectations.

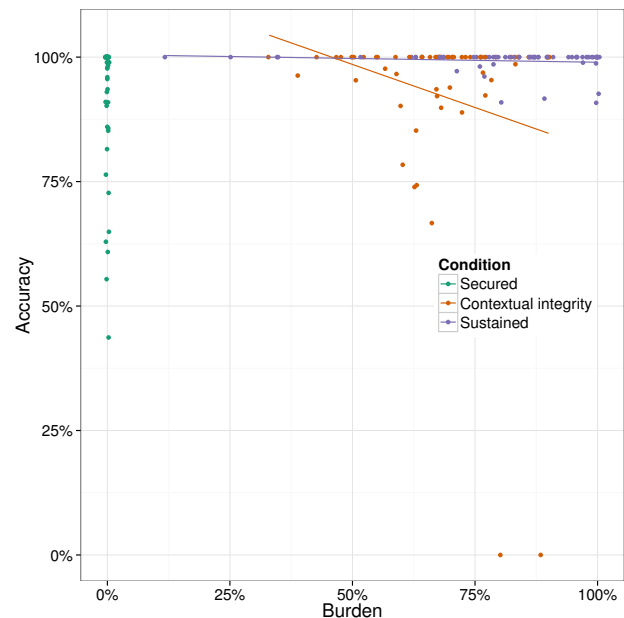


Figure 4: Scatterplot showing the relationship between accuracy and burden in all three conditions. Accuracy is the most variable for those in the secured consent condition, whereas in the case of contextual integrity, a small loss in accuracy is met with a greater time burden saving. Note that the points have been jittered to improve readability.

Figure 4 shows the relationship between burden and accuracy in all three conditions. As expected, participants in the sustained consent condition mostly saw perfect accuracy. While secured consent does indeed exhibit the most variable accuracy, it performs better than we originally anticipated; we discuss the implications of this later.

Behaviour in the contextual integrity condition is more variable. While there is a similar tendency towards higher accuracy, there is a notable drop in performance. As shown by the regression line, accuracy surprisingly drops with increased burden. This exposes an important detail about the

²Facebook Ads Manager: facebook.com/advertising

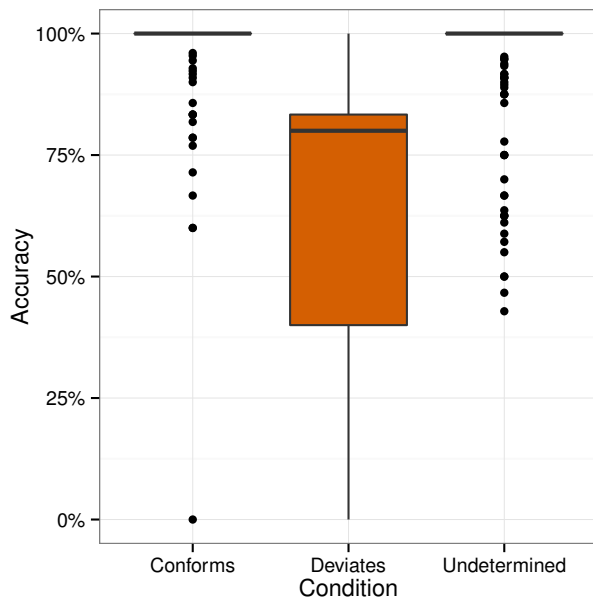


Figure 5: Boxplot showing how accuracy differs between participants in the contextual integrity condition depending on their norm conformity. Norm-conformant people achieve higher, and less variable, accuracy rates than those who deviate from norms.

applicability of the contextual integrity method. When we expand this condition to show the accuracy of participants based on their norm conformity, this trend is easier to understand. The technique attempts to detect the norm deviance and conformity of participants. As shown in Figure 5, the former was not useful for maximising accuracy. As deviance requires more questions to detect, this drags down overall accuracy as burden increases. Later in this section we discuss the implications of contextual integrity better serving certain types of people.

When combining responses from all conditions, a one-way ANOVA shows no statistically significant effect of burden on accuracy ($F(1,171)=0.903, p>0.1$), suggesting that as people’s behaviour is so diverse, improving the accuracy of consent acquisition simply through increasing the number of interventions may not be sufficient.

Does contextual integrity reduce participant burden?

Contextual integrity is used in the informed consent process to determine if people conform to social norms, and leverage this conformity to ask fewer questions about their willingness to share data. Participants in the sustained consent and contextual integrity conditions were asked a maximum of 100 questions.³ In the contextual integrity conditions, questions were not asked if the participant was found

³For some, this number was smaller if they did not have enough pieces of data in their Facebook profile of each type.

to be norm-conformant or deviant with respect to a particular data type. On average, participants in the sustained condition were asked 81.9 questions, while contextual integrity participants were asked 67.2, an 21.9% decrease in burden. When comparing the distribution of burden between the two conditions, a one-way ANOVA shows a statistically significant difference ($F(1,122)=25.15, p<0.05$). This significant reduction is useful when conducting longitudinal studies, as it suggests the technique may allow fewer disruptive interventions to acquire consent. This finding is only useful, however, if accuracy is not compromised.

Does contextual integrity significantly reduce accuracy?

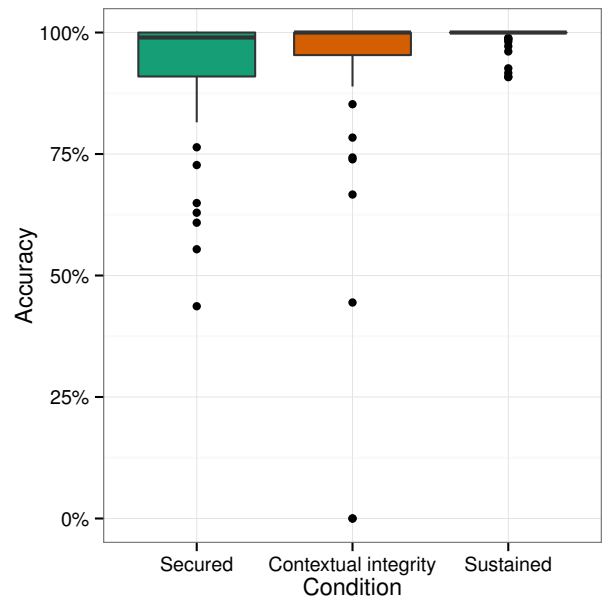


Figure 6: In all three conditions, median accuracy is high, although variability increases as the number of questions asked is reduced.

In all conditions, mean accuracy is very high, only dropping from 99.3% in the sustained condition, to 92.7% in the contextual integrity case, and 91.2% for secured consent. Accuracy is most variable in the contextual integrity and secured conditions, as depicted in Figure 6. Surprisingly, median accuracy in the secured consent condition is close to 100%, suggesting a large proportion of the sample were willing to share all of their data with researchers. While identifying the motivations for this are beyond the scope of this study, previous work has shown that people are more willing to share social network data with academic researchers depending on the purpose of the study (McNeilly, Hutton, and Henderson 2013).

As expected, variability for contextual integrity participants lies between the two extremes of the other conditions. Despite the similarly high medians, an ANOVA comparing accuracy between the contextual integrity and sustained con-

ditions suggests that the former exhibits a significantly lower accuracy distribution, failing one of our performance metrics ($F(1,120) = 7.45, p < 0.05$) The plurality of a high accuracy cluster, and very low-performing outliers merits further explanation, and goes some way to explaining this apparent failure.

Who does contextual integrity work for?

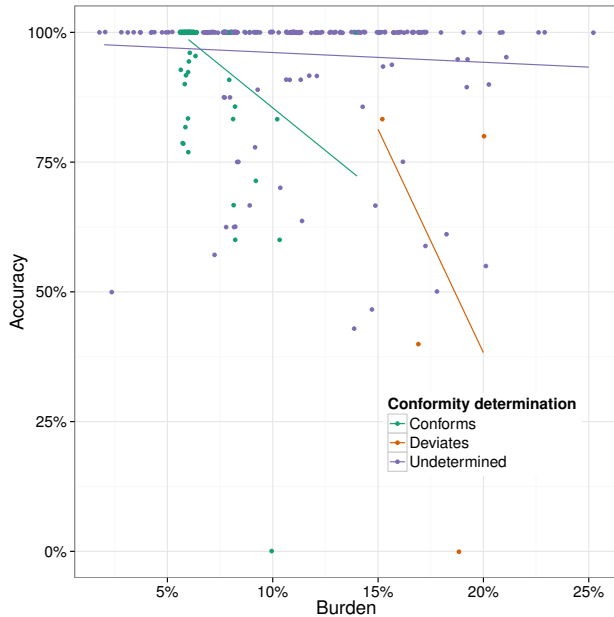


Figure 7: Scatterplot showing the relationship between norm conformity and accuracy. As indicated by the cluster after 5 questions, high accuracy can be preserved when norm conformity is detected quickly, although the technique is not useful for people who are highly norm deviant. Note that the points have been jittered to improve readability.

We classify participants in the contextual integrity condition based on their norm conformity. Figure 8 shows the time taken to determine which of these groups participants belong to. If a participant is norm-conformant, in the vast majority of cases this is identified within just 7 questions per attribute, allowing us to skip all further questions and use this conformity as a proxy for their willingness to share discrete items. Figure 7 illustrates how this purging of questions for norm-conformant and deviant participants affects accuracy, indicating the types of user that the contextual integrity technique can support. The technique used to detect norm conformity requires a person to agree to share at least 5 pieces of data before being able to calculate whether or not they conform to the norm. The large cluster of green points here with high accuracy indicates that people who are norm-conformant can be asked a small number of questions while maintaining accuracy. Indeed, when questions are removed at this point, the technique performs better than for people whose conformity could not be established (and for whom no questions were removed), suggesting that people

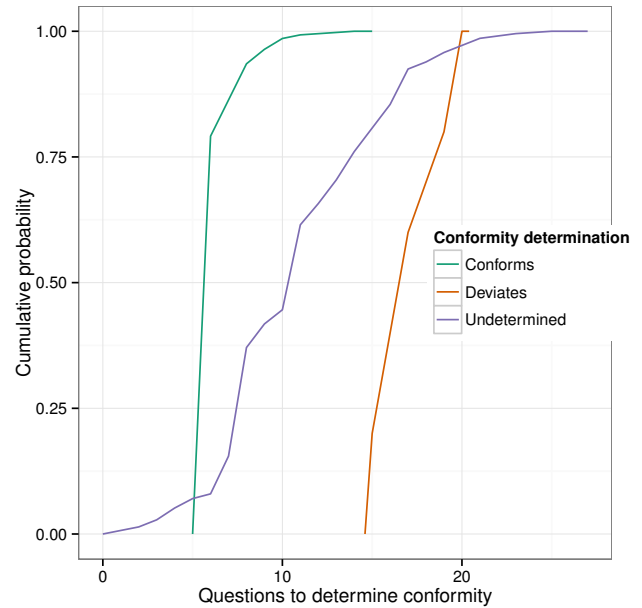


Figure 8: Cumulative distribution function of the number of questions needed to determine whether participants in the contextual integrity condition conform or deviate from the norm, or if this could not be determined. If people conform to norms, this is detected after a small number of interventions.

who behave similarly to their peers can be asked fewer questions and achieve higher accuracy.

27.7% of participants' conformity was detected within 6 questions while achieving more than 96.5% accuracy, the intercept with the undetermined regression line at this point. Across the whole condition, the same number of norm-conformant participants achieved this degree of accuracy or higher than undetermined participants, while on average achieving a 41.1% reduction in burden, indicating that for such people, contextual integrity achieves both a reduction in burden and an *increase* in accuracy. This is an important result, as although it indicates that contextual integrity is not be appropriate for all people, within just 6 questions we can detect whether a user's consent can be accurately captured, and apply an appropriate strategy based on their behaviour. Based on the slope of the regression lines, we can determine best practices for application of the contextual integrity technique. If a person's norm conformity can be detected within 6 questions, then trusting this as a proxy for their explicit performs very well. We also determined whether participants were significantly norm-deviant. That is, if their behaviour significantly deviated from the social norms, we would also remove questions and take their current sharing ratios as a proxy for willingness to share. We found that this approach does not perform well, as deviance requires at least 15 interventions to determine, and is a very low-performing proxy. As such, if people are significantly deviant from social norms, it is best to continue asking them

questions, as high accuracy is maintained.

Is contextual integrity robust to temporal changes in willingness to share?

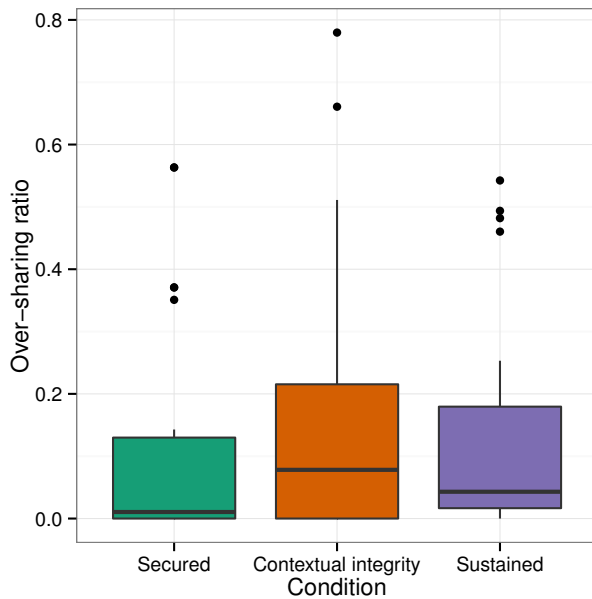


Figure 9: Boxplot showing how robust each condition is to temporal change by using the consent policy from the first week to predict the participant’s responses in the second week. All conditions exhibit over-sharing, highlighting the difficulty of capturing consent at a single point in time.

Previous work has shown that attitudes towards privacy and sharing social network data are temporally volatile (Bauer et al. 2013), and as such, decisions about willingness to share data with researchers may only represent thinking at a single point in time, and not a person’s “true” intent. This may cause regret, and perceived leakage of social network data (Wang et al. 2011). The consent methods we examine in this study consider the temporal issue differently. The secured consent method, perhaps the most common in social network studies, assumes that a participant’s willingness to participate in a study is *carte blanche* to collect any data associated with them, and disregards any temporal effects. Conversely, sustained consent relies on constant interventions to mitigate any drift in a person’s willingness to share data, which achieves high accuracy at the cost of a significant burden on the participant. As a goal of the contextual integrity method is to reduce the burden on participants, we hypothesise that leveraging social norms is more robust over time. If a user is found to highly conform to social norms at one point in time, we expect this to hold true, as a proxy for willingness to share discrete pieces of data, better than the secured consent method. As we have discussed earlier, we expected a small decrease in accuracy compared to sustained consent as the significant number of interventions ensures accuracy. By repeating the

study over a week, we capture changes in behaviour to determine the robustness of these techniques. To do this, we apply the consent policy of the first week’s results to predict the participant’s responses in the second week. These predictions are validated by the participant’s responses to the performance evaluation questionnaire, the same way accuracy is measured. Figure 9 illustrates the extent to which these predictions would have led to over-sharing of data. These results suggest that privacy attitudes do indeed change over the course of a week. Across all conditions, trying to use predictions from the previous week performs quite poorly in many cases. This is most problematic in the case of secured consent because there is no way of accommodating such changes in intent, suggesting that consent acquisition in a single moment in time is not sufficient. The sustained consent condition shows a very similar distribution, however in practice this would be mitigated by continuing to intervene to capture consent, dismissing the need to rely on old data, at the cost of higher participant burden. Surprisingly, the contextual integrity condition performs poorly by this measure of robustness, however this is understandable in the context of our previous result that the technique is only applicable for about a quarter of the population who are highly norm-conformant, and an ANOVA suggests that there is no significant difference in over-sharing between conditions ($F(2,65) = 0.168, p > 0.1$). As this condition includes participants of varying degrees of conformity, attempts to leverage this to make longitudinal predictions for non-conformant participants performs very poorly. In practice, users who have not been identified themselves as norm-conformant within 6 interventions would be excluded from a contextual integrity-based solution in favour of a sustained consent approach which would better capture their intent.

Related work

The ethical collection of data from Facebook and other public sources of online social network information has been widely discussed in the literature. Zimmer (2010) discusses problems with the well-known T3 Facebook study conducted at Harvard (Lewis et al. 2008), and concludes that simply because a social network user chooses to share data on an SNS, this does not give a researcher the right to collect such data; on the contrary, researchers must take even more care with such data. Neuhaus and Webmoor (2012) similarly argue that academic researchers should take more care with SNS data than commercial data processors, and propose a set of best practices termed “agile ethics”. Such agile ethics assume that data collection practices may change over the course of research, thus implying that traditional secured consent may be insufficient. Solberg (2010) considers that Facebook research has low risks, and that such research does not require any ethics approval. Moreno et al. (2008) examines SNS research involving adolescents, and argues that (secured) informed consent may be required but should be decided on a case-by-case basis. The British Psychological Society (2013) provide a detailed set of guidelines on obtaining “valid consent” from participants in Internet research.

Recent work by Morrison et al. (2014) is perhaps the closest to our own. They look at obtaining consent from partic-

ipants in mobile HCI experiments. They propose that once participants have shared a particular quantity of data with researchers, they should be presented with a personal representation of their shared data to determine whether they continue to grant the researchers consent to access these data. Such representations are found to be useful in a large user study. Our work differs from this in that we attempt to use contextual integrity as a more efficient way of detecting changes in consent.

Contextual integrity has been widely discussed since its introduction over ten years ago (Nissenbaum 2004). Empirical evaluations of contextual integrity include studies of Google Books (Jones and Janes 2010), cloud computing (Grodzinsky and Tavani 2011), and blogging (Grodzinsky and Tavani 2010), while Barkhuus (2012) argues that it should be more widely used in HCI. To the best of our knowledge, we are the first to apply contextual integrity in an empirical evaluation of informed consent.

Conclusions and future work

We have presented the first application of contextual integrity to the acquisition of informed consent for sharing social network data. We conducted a study with 109 participants to identify whether contextual integrity could reduce the burden on participants while maintaining accuracy.

Returning to our hypotheses, we find qualified support for H1. On average, contextual integrity reduces burden by 21.9%. While median accuracy is not significantly better than secured consent, for 27.7% of participants, contextual integrity delivered perfect accuracy with a 41.1% reduction in burden. We also find support for H2, as the contextual integrity method is not significantly less robust than sustained consent over time. Our results vindicate our decision to test norm conformity at the 0.1 significance level, but further work can demonstrate whether this boundary is appropriate in all cases.

We conclude that as human behaviour is so diverse, there is no one-size-fits-all approach to consent that achieves optimal results. We believe it is important for organisations acquiring consent for social network data to consider the implications of these results. The attraction of our use of contextual integrity is that as norm conformity can be quickly established, if a person clearly does not conform to such norms, it is possible to transparently change strategy to a sustained approach and maximise accuracy. We demonstrate that while the low-burden secured consent approach may be sufficient for some people, it can not be relied on to maintain accuracy in most cases.

We acknowledge that our measure of accuracy is not the sole means to determine that informed consent has been sought. This metric allows us to confirm that the participant disclosed the SNS data that they were willing to, which we believe is important to establish. It does not, however, investigate whether the participant understands the *implications* of sharing their data, or the purpose of the research. In biomedical studies, consent comprehension tests are commonly used to determine that participants' consent is informed, but their effectiveness has been questioned (2007). Investigating the wider implications of assessing consent

comprehension is important further work, where again we anticipate contextual integrity could be leveraged. For example, while we found that our semi-automated approach to determining consent was appropriate for some people, others might find it invasive, and striking this ethical balance is a sensitive topic.

We would like to investigate other applications to which we can apply our findings. For example, we are interested in applying contextual integrity to mobile HCI studies which would benefit from gaining informed consent in such a way that reduces the burden on participants. We are also interested in exploring other methodological implications of adopting contextual integrity, for instance studying whether the method is liable to introducing bias as in other forms of informed consent (Rothstein and Shoben 2013).

Acknowledgments

This work was funded by the Engineering and Physical Sciences Research Council (EPSRC) through a Doctoral Training Grant.

References

- Ayalon, O., and Toch, E. 2013. Retrospective privacy: Managing longitudinal privacy in online social networks. In *Proceedings of the 9th Symposium on Usable Privacy and Security*.
- Barkhuus, L. 2012. The mismeasurement of privacy: using contextual integrity to reconsider privacy in HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 367–376.
- Bauer, L.; Cranor, L. F.; Komanduri, S.; Mazurek, M. L.; Reiter, M. K.; Sleeper, M.; and Ur, B. 2013. The post anachronism: The temporal dimension of Facebook privacy. In *Proceedings of the 12th ACM Workshop on Privacy in the Electronic Society*, 1–12.
- boyd, d., and Crawford, K. 2012. Critical questions for big data. *Information, Communication & Society* 15(5):662–679.
- British Psychological Society. 2013. Ethics guidelines for Internet-mediated research. INF206/1.2013. Available online at <http://www.bps.org.uk/publications/policy-and-guidelines/research-guidelines-policy-documents/research-guidelines-poli>.
- Cohn, E., and Larson, E. 2007. Improving participant comprehension in the informed consent process. *Journal of nursing scholarship* 39(3):273–280.
- Getty, E.; Cobb, J.; Gabeler, M.; Nelson, C.; Weng, E.; and Hancock, J. 2011. I Said Your Name in an Empty Room: Grieving and Continuing Bonds on Facebook. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 997–1000.
- Grodzinsky, F. S., and Tavani, H. T. 2010. Applying the “contextual integrity” model of privacy to personal blogs in the blogosphere. *International Journal of Internet Research Ethics* 3(1):38–47.

- Grodzinsky, F. S., and Tavani, H. T. 2011. Privacy in “the cloud”: applying Nissenbaum’s theory of contextual integrity. *ACM SIGCAS Computers and Society* 41(1):38–47.
- Hill, K. 2014. Facebook Added ‘Research’ To User Agreement 4 Months After Emotion Manipulation Study. *Forbes*. <http://onforb.es/15DKfGt>.
- Hutton, L., and Henderson, T. 2013. An architecture for ethical and privacy-sensitive social network experiments. *SIGMETRICS Perform. Eval. Rev.* 40(4):90–95.
- Hutton, L., and Henderson, T. 2015. Towards reproducibility in online social network research. *IEEE Transactions on Emerging Topics in Computing*. To appear.
- Jones, E. A., and Janes, J. W. 2010. Anonymity in a world of digital books: Google Books, privacy, and the freedom to read. *Policy & Internet* 2(4):42–74.
- Kramer, A. D. I.; Guillory, J. E.; and Hancock, J. T. 2014. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences* 111(24):8788–8790.
- Lewis, K.; Kaufman, J.; Gonzalez, M.; Wimmer, A.; and Christakis, N. 2008. Tastes, ties, and time: A new social network dataset using Facebook.com. *Social Networks* 30(4):330–342.
- Luger, E., and Rodden, T. 2013. An informed view on consent for UbiComp. In *Proceedings of the ACM international joint conference on Pervasive and Ubiquitous computing*, 529–538.
- McNeilly, S.; Hutton, L.; and Henderson, T. 2013. Understanding ethical concerns in social media privacy studies. In *Proceedings of the ACM CSCW Workshop on Measuring Networked Social Privacy: Qualitative & Quantitative Approaches*.
- Moreno, M. A.; Fost, N. C.; and Christakis, D. A. 2008. Research ethics in the MySpace era. *Pediatrics* 121(1):157–161.
- Morrison, A.; McMillan, D.; and Chalmers, M. 2014. Improving consent in large scale mobile HCI through personalised representations of data. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational*, 471–480.
- Neuhaus, F., and Webmoor, T. 2012. Agile ethics for massified research and visualization. *Information, Communication & Society* 15(1):43–65.
- Nissenbaum, H. F. 2004. Privacy as contextual integrity. *Washington Law Review* 79(1):119–157.
- Powell, S. J.; Linehan, C.; Daley, L.; Garbett, A.; and Lawson, S. 2012. “i can’t get no sleep”: Discussing #insomnia on Twitter. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1501–1510.
- Rothstein, M. A., and Shoben, A. B. 2013. Does consent bias research? *The American Journal of Bioethics* 13(4):27–37.
- Sleeper, M.; Balebako, R.; Das, S.; McConahy, A. L.; Wiese, J.; and Cranor, L. F. 2013. The post that wasn’t: Exploring self-censorship on Facebook. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work*, 793–802.
- Solberg, L. 2010. Data mining on Facebook: A free space for researchers or an IRB nightmare? *University of Illinois Journal of Law, Technology & Policy* 2010(2).
- Wang, Y.; Norcie, G.; Komanduri, S.; Acquisti, A.; Leon, P. G.; and Cranor, L. F. 2011. “I regretted the minute I pressed share”: a qualitative study of regrets on Facebook. In *Proceedings of the 7th Symposium on Usable Privacy and Security*.
- Warrell, J. G., and Jacobsen, M. 2014. Internet research ethics and the policy gap for ethical practice in online research settings. *Canadian Journal of Higher Education* 44(1):22–37.
- Wilson, R. E.; Gosling, S. D.; and Graham, L. T. 2012. A review of Facebook research in the social sciences. *Perspectives on Psychological Science* 7(3):203–220.
- Zimmer, M. 2010. “But the data is already public”: on the ethics of research in Facebook. *Ethics and Information Technology* 12(4):313–325.