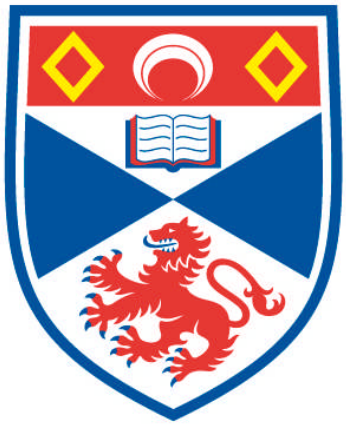
APPLYING CONTEXTUAL INTEGRITY TO THE STUDY OF SOCIAL NETWORK SITES

Luke Hutton

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



2015

Applying contextual integrity to the study of social network sites

Luke Hutton

University of
St Andrews

This thesis is submitted in
partial fulfilment for the
degree of Doctor of Philosophy
at the University of St Andrews

September 2015

Abstract

Social network sites (SNSs) have become very popular, with more than 1.39 billion people using Facebook alone. The ability to share large amounts of personal information with these services, such as location traces, photos, and messages, has raised a number of privacy concerns. The popularity of these services has enabled new research directions, allowing researchers to collect large amounts of data from SNSs to gain insight into how people share information, and to identify and resolve issues with such services. There are challenges to conducting such research responsibly, ensuring studies are ethical and protect the privacy of participants, while ensuring research outputs are sustainable and can be reproduced in the future.

These challenges motivate the application of a theoretical framework that can be used to understand, identify, and mitigate the privacy impacts of emerging SNSs, and the conduct of ethical SNS studies. In this thesis, we apply Nissenbaum's model of contextual integrity to the study of SNSs. We develop an architecture for conducting privacy-preserving and reproducible SNS studies that upholds the contextual integrity of participants. We apply the architecture to the study of informed consent to show that contextual integrity can be leveraged to improve the acquisition of consent in such studies. We then use contextual integrity to diagnose potential privacy violations in an emerging form of SNS.

Declarations

Candidate's declarations

I, Luke Hutton, hereby certify that this thesis, which is approximately 36,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in September 2011 and as a candidate for the degree of Doctor of Philosophy in September 2011; the higher study of which this is a record was carried out in the University of St Andrews between 2011 and 2015.

date _____ *signature of candidate* _____

Supervisor's declarations

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Doctor of Philosophy in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

date _____ *signature of supervisor* _____

Permission for publication

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. I have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

PRINTED COPY

No embargo on print copy

ELECTRONIC COPY

No embargo on electronic copy

signature of candidate _____ *signature of supervisor* _____

date _____

There's a path stained with tears,
Could you talk to quiet my fears?
Could you pull me aside,
Just to acknowledge that I've tried?
...
And as your last breath begins,
You find your demon's your best friend.
And we all get it in
The end.

Scott Matthew, In The End

Contents

List of Figures

List of Tables

Acknowledgements

Getting to this point hasn't always been easy, and nor should it be. It certainly wouldn't have been possible without the endless support of some amazing people.

First of all, I thank my supervisor, Tristan Henderson, for his belief that I would get here even when I was losing hope, and for every friendly nudge to submit submit submit. His supervision was methodical, sometimes intense, always appreciated. My thanks also to my second supervisors and reviewers, Saleem Bhatti, Per Ola Kristensson, and John Thomson for their feedback and support. I also wish to thank Judi Robertson in CS, and Ruth Unsworth for all their help.

I am lucky to be surrounded by so many wonderful friends. In St Andrews, the Taint: CJ Davies, Ruth Hoffmann, and Jaunty Ward have provided endless gallows humour, gin, and sloth GIFs, and they have no idea how much I love them for it. A special thanks to Davie and Ruth Letham for the chat, Pokémon, and vegan noms. To everyone else in CS and beyond, past and present, who's trawled to the pub and put up with me, thank you. Beyond the bubble, Sarah and Nathan Chapman and Alanna Kerr earn every bit of their Team FAB moniker every day, even surviving pilotgate (*"the greatest test of our friendship so far"*) while being forever fantastically awesome. This one's for Nickelback. I thank Lisa Campbell whose support and belief has been phenomenal, and for knowing that pizza and beer is the answer to most of my problems; and Kath and Emma Carrick, for wine, for great food, for understanding, and a place to play Civilization.

Finally, my family have been a rock, and my parents Stewart and Joanne have been a constant source of support and inspiration, whether bailing me out in times of need, or just giving me an ear to whinge at. I also want to thank my brothers, Ben, Neal and Colin; my sisters-in-law Hannah and Emily; and my awesome nieces, Mindy and Honor, for everything, whether it was a glass of wine, a video game, or to just hang out with an episode of Adventure Time. My love also goes to Grandma Fowler, my Grandma and late Grandad Hutton, Karen, Tracy, Sean, Kieran, Daniel, Josh, and Toby.

Chapter 1

Introduction

Social network sites (SNSs) have quickly become one of the widest-used applications on the Internet. A range of services has been enabled that allows people to share content with their peers, such as messages, photos, and their current location, with the structure of people's interpersonal relationships represented as a graph of connected peers. As of 2014, 63% of UK adults owned a smartphone [102], while 47% of adults used SNSs. Of those, 82% of 16-24 year olds are SNS users [101]. With such services achieving near-ubiquity among a generation of people, it is important to consider the implications they present.

The quantities of personal information that can be collected and shared with others using these services has raised many privacy concerns. The data shared on SNSs may not be considered appropriate or relevant to everyone in a person's social network, placing responsibility on the individual to correctly configure the privacy settings for each disclosure such that sensitive data are not inadvertently exposed to the wrong audience [76]. The protection of individuals' privacy is not just dependent on users' comprehension of these settings, but on the trustworthiness of their peers and the SNS itself. While people's attitudes towards their privacy may change, these settings are static, meaning that people may subsequently regret the posts they have previously made [136], again placing responsibility on the individual to police their historic disclosures. As well as trusting that the service does not inappropriately share their users'

data [71], they are trusted to protect data from external attack, which is considered a growing threat [108].

These issues, and the difficulty of defining privacy itself, has motivated many attempts to solidify privacy within a theoretical framework. Helen Nissenbaum's model of contextual integrity is one such effort [100], which determines privacy violations emerge when people's context-specific norms about acceptable information sharing are not met. The model can be used as a diagnostic tool to identify potential sources of privacy breaches by determining when existing norms are perturbed by a new process that changes what types of information are collected and transmitted.

The popularity of SNSs, and concern about such privacy challenges has opened new research directions concerning their design and the behaviour of the people who use them. These studies can provide useful insights, but it is important to consider how to conduct such research responsibly, by ensuring its outcomes benefit a variety of stakeholders, and society at large [119]. Acknowledging both the ethical challenges associated with collecting and processing such sensitive data, and the sustainability challenges of making such research reproducible in the future, can make progress towards making SNS research more responsible.

As there is no consensus about whether all studies using social network sites constitute human subjects research, there is a lack of consistency about the amount of ethical oversight such research receives, with previous studies attracting attention because of failures to preserve the anonymity of participants [147], or gaining their informed consent [49]. This places great responsibility on researchers to adequately protect their participants, specifically when considering sharing data with others, which could reveal sensitive information about participants, with it being possible to reidentify participants when data are not sufficiently sanitised [97, 22].

Conducting research responsibly can also be enhanced through steps to improve its sustainability: ensuring research artefacts continue to be usable and meaningful in the future. Making experiments reproducible, which has long

been considered a fundamental part of the scientific method [78], can significantly contribute to this. SNS research is particularly difficult to execute in a reproducible manner, however, as the design of services often change or are made obsolete, requiring constant maintenance of code that accesses SNS data, and in some cases, experiments simply cannot be run again. Additionally, limited reporting of methodological details such as how participants were sampled, or how informed consent was sought, can make it more difficult to reproduce such studies, with no standards for encoding and sharing the workflow of SNS experiments being adopted.

As the use of SNSs continues to rise, new services emerge, with at least 200 operating as of 2015 [140]. With outstanding privacy challenges with the services themselves, and concerns about the responsible conduct of SNS studies, it remains unknown whether a single conceptual framework can be leveraged to identify and address these issues in a holistic manner. While Nissenbaum's model of contextual integrity has been used to investigate a range of privacy concerns [42, 59, 75, 113], it has not been used to identify privacy breaches in emerging SNSs where existing norms might be disrupted. Furthermore, whether it can be used to evaluate and improve the conduct of SNS research remains unknown.

In this thesis, we address the following research questions:

RQ1: Is contextual integrity an appropriate framework for understanding and mitigating ethical concerns in SNS research?

RQ2: Can contextual integrity be used in the evaluation of SNSs to detect and mitigate their privacy impacts?

1.1 Thesis statement

We make the following thesis statement:

> Contextual integrity can be used to conduct reproducible and privacy-preserving experiments using social network sites, and can detect potential privacy violations in new services in order to mitigate their impact.

In support of this, we make three contributions. We demonstrate that:

1. Using an architecture informed by contextual integrity to conduct SNS studies can better protect the privacy of participants and lead to more reproducible experiments, than directly accessing the APIs provided by SNSs.
2. Contextual integrity can be leveraged to acquire consent in SNS studies to better meet the expectations of participants and mitigate the ethical issues associated with failing to gain informed consent.
3. Contextual integrity can be used to detect privacy breaches in emerging SNSs, and to identify how the design of these services can be altered to avoid such violations.

1.2 Outline

This thesis is structured as follows.

Chapters 2 and 3 outline the research context, and the state of the art.

- Chapter 2 introduces the range of social network sites we will study, the privacy challenges associated with these services, and the ethical challenges associated with conducting research into SNSs.

- Chapter 3 examines the recent literature that attempts to resolve some of the challenges we identified in Chapter 2, noting the open problems we will address.

Chapters 4, 5, and 6 describe our three contributions to support our thesis.

- In Chapter 4, we introduce an architecture we have developed to improve the ethical conduct of SNS studies, the effectiveness of which we demonstrate by reproducing an existing study.
- In Chapter 5, we tackle a specific ethical challenge in SNS consent, demonstrating that a method for acquiring consent for SNS studies that maintains contextual integrity can meet the expectations of participants while reducing the burden placed on them.
- In Chapter 6, we conduct two case studies to show how contextual integrity can be used to detect breaches in two types of SNSs, and how the framework guides the design of the services to mitigate these breaches.

  Finally, we conclude with a summary of the contributions we have made, and outline directions for future research.

1.3 Publications

This thesis is entirely my work, but has been supported by a number of collaborators. Throughout this thesis, I use the word "we" to acknowledge the contribution these collaborators have made to this work.

During the course of my PhD, I have contributed to the following publications. Where I am first author, I have been chiefly responsible for the core contributions of experimental design, implementation, and execution of studies and analyses, and it is this work that I present in this thesis.

- Luke Hutton and Tristan Henderson. "Towards reproducibility in online social network research". In: *IEEE Transactions on Emerging Topics in Computing* (2015). doi: 10.1109/tetc.2015.2458574
- Luke Hutton and Tristan Henderson. "Making social media research reproducible". In: *Proceedings of the ICWSM Workshop on Standards and Practices in Large-Scale Social Media Research*. Oxford, UK, May 2015
- Luke Hutton and Tristan Henderson. ""I didn't sign up for this!": Informed consent in social network research". In: *Proceedings of the 9th International AAAI Conference on Web and Social Media (ICWSM)*. May 2015. url: http://tristan.host.cs.st-andrews.ac.uk/research/pubs/icwsm2015.pdf
- Luke Hutton and Tristan Henderson. "An architecture for ethical and privacy-sensitive social network experiments". In: *SIGMETRICS Performance Evaluation Review* 40.4 (Apr. 2013), pp. 90–95. doi: 10.1145/2479942.2479954
- Luke Hutton, Tristan Henderson and Apu Kapadia. ""Here I Am, Now Pay Me!": Privacy Concerns in Incentivised Location-sharing Systems". In: *Proceedings of the 2014 ACM Conference on Security and Privacy in Wireless & Mobile Networks*. WiSec '14. Oxford, UK: ACM, July 2014, pp. 81–86. doi: 10.1145/2627393.2627416
- Sam McNeilly, Luke Hutton and Tristan Henderson. "Understanding ethical concerns in social media privacy studies". In: *Proceedings of ACM CSCW Workshop on Measuring Networked Social Privacy: Qualitative & Quantitative Approaches*. San Antonio, TX, USA, Feb. 2013. url: http://tristan.host.cs.st-andrews.ac.uk/pubs/mnsp2013.pdf

Chapter 2

Background

In this chapter, we outline the nature of the social network sites that we study in this thesis. We also discuss the prevailing methodological approaches adopted in the research of such systems, and the ethical challenges associated with these methods.

2.1 Social network sites

In its early years, the World Wide Web was a distributed repository of largely static information, where the majority of users only consumed content [19]. Beginning in the early 2000s, the development of technologies such as AJAX allowed more sophisticated applications to be delivered through Web browsers, encouraging users of the Web to become creators. These applications allowed people to share photos, write blog posts, and comment on the content generated by others, adding a social dimension that had been largely absent from the Web until then, in a movement widely known as "Web 2.0" [19].

The introduction of smartphones with large displays and touch capabilities has led to further developments in the Web. The computational power and relatively high-resolution displays of these devices reduced the gap between mobile phones and computers, with the ability to render websites at similar fidelity

to their desktop counterparts [15]. The quick adoption of smartphones coupled with acknowledgement of their limitations with respect to input modalities and display size, has led to efforts to optimise websites for such devices, and to deliver native applications which can use their full capabilities and achieve higher performance [14]. The additional capabilities of smartphones, such as location-sensing and cameras, coupled with the increasing availability of high-speed mobile Internet connections, has enabled new services which allow people to share photos, videos, and their current location [15]. These abilities have led to a change in how people access the Web, with more than 37% of all Web traffic coming from mobile devices as of February 2015 [121].

The social functionality which began to emerge in websites evolved into a set of applications that were designed around the concept of managing a set of friends with which information could be shared: the social network site. In 2007, boyd and Ellison defined social network sites as Web-based services which meet the following criteria [11]:

1. Users can create a profile, which might be public or visible to a subset of people.
2. Users can curate a list of other users of the service they are connected to.
3. Users can traverse these connections, and the connections made by others, to see their profiles and other content.

This definition is intentionally sparse. While individual SNSs vary in design and purpose they share some structural components. All provide a service over the Internet, and implement an eponymous social network; a graph where nodes are individual people, and the edges a semantically-appropriate social connection between them, such as friendship or physical co-location. In many cases, it is the semantics of the connections within that network which define the purpose of the service.

Launched in 2004 as an SNS for students at Harvard University, Facebook

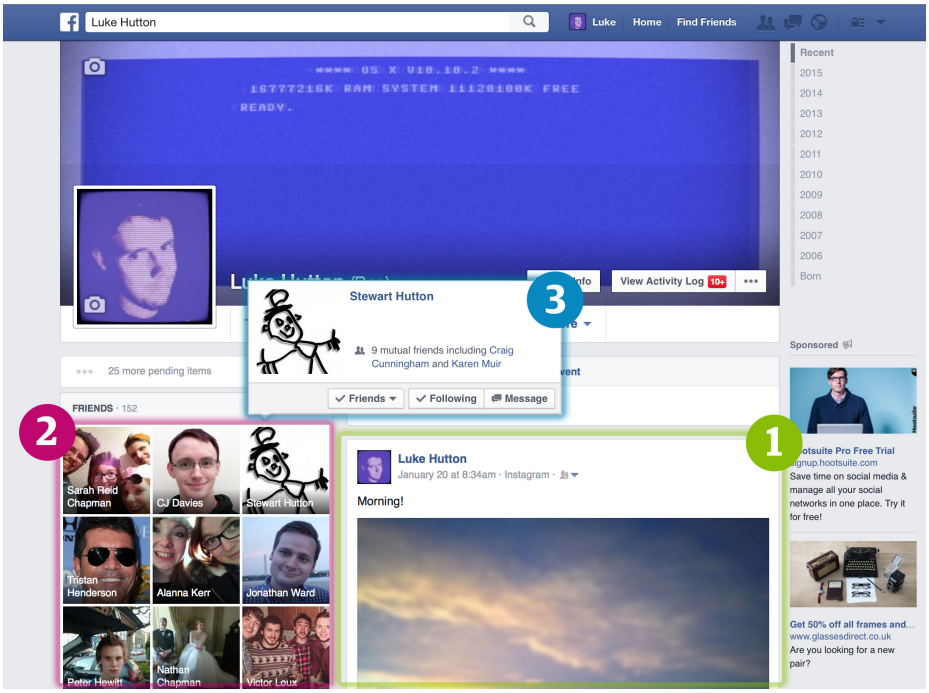


Figure 2.1: A screenshot of a Facebook profile, illustrating its implementation of the three fundamental SNS concepts: 1) A profile consisting of content decided by the user. 2) A list of friends curated by the user. 3) Connections between friends are exposed, allowing traversal between profiles.

has rapidly grown to become one of the world's most popular SNSs, with 1.39 billion active users as of December 2014.[1] It allows its users to share and engage with photos, videos, links to websites and text-based status updates posted by peers, via its website or mobile applications, the latter becoming increasingly popular, as we will discuss later. Facebook meets the SNS criteria identified by boyd and Ellison through one of its core products, Profile. As shown in Figure 2.1, the product aggregates content published by the user, which is visible to an audience determined by its author. The profile exposes the "friends list" of the user, an egocentric social network of their direct connections. This mechanism allows the user to traverse profiles of their friends, with mutual friends and recommended connections surfaced through the structure of the global social network. While initially designed to represent friendships between college students, this concept has extended to model familial relationships, dynamics between colleagues, or even between businesses and their customers. Des-

[1]Facebook company information: http://newsroom.fb.com/company-info/

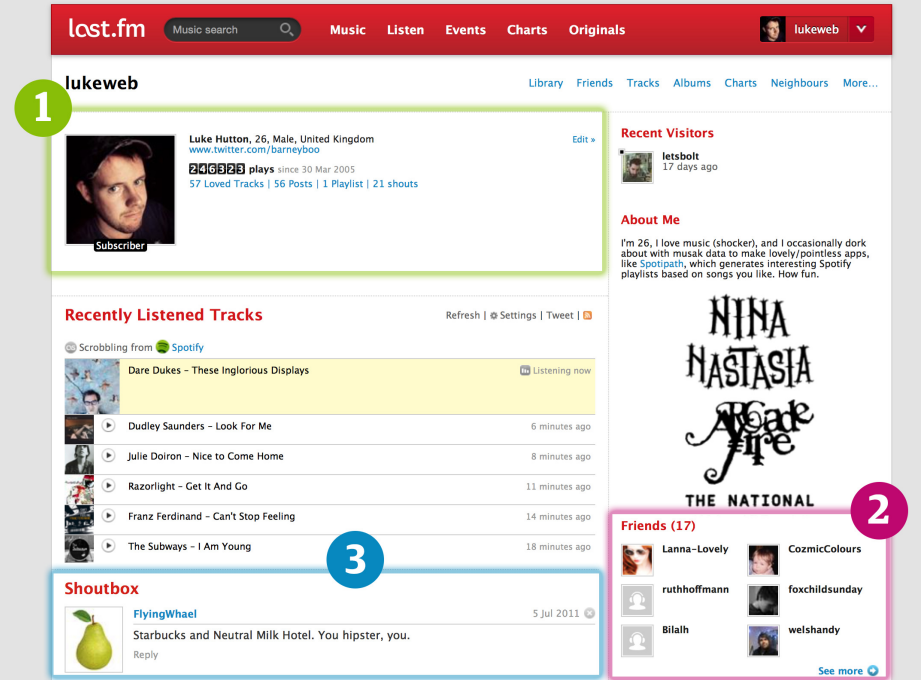


Figure 2.2: A screenshot of a Last.fm profile, illustrating how these same three SNS concepts manifest in more niche services, with: 1) A profile consisting of attributes disclosed by the user. 2) A list of friends curated by the user. 3) Friends can post messages on each other's wall, encouraging engagement between profiles.

pite the development of new products by Facebook, the fundamental purpose of the profile is remarkably similar to that of its predecessors, such as Friendster (founded in 2002) and MySpace (founded in 2003). These services, among many others which emerged in the early 2000s, have waned in popularity since Facebook's dominance took hold, but despite superficial advances, Facebook's structure deviates little from these services.

boyd and Ellison's SNS model holds true for services which deviate from the approach of Facebook, and its ancestors. One such example is Last.fm, which forges communities around shared music interests. Despite carving out a relative niche, its structure is again consistent with that of Facebook's Profile product, as shown in Figure 2.2. Last.fm's core offering is to track the music people listen to on their personal devices, and recommend new artists through collaborative filtering techniques, illustrating how SNS concepts can be embedded in a range of applications. The service identifies users with similar tastes and encourages them to become friends, making it easier to follow each other's

Figure 2.3: A screenshot of a Twitter profile, showing how it implements the three SNS concepts: 1) A profile consisting of tweets shared by the user. 2) The directed nature of the Twitter graph means that the indegree and outdegree of each user is reported independently, with the outdegree ("following") determined by the user. 3) The profiles of others can be traversed to view the tweets they have recently published.

activity on the service.

While boyd and Ellison do not make explicit references to the graph which underpins the connections between these services, the design of this graph has significant implications for the services built on top of them. The SNSs we have discussed so far all employ an undirected graph in their core network, where friendships are modelled as symmetric connections.[2] A friendship, initiated by one party, must be accepted by the other to forge a connection on the graph, and allow content to flow bidirectionally between peers. Another major SNS, Twitter, does not employ this approach.

Depicted in Figure 2.3, Twitter allows users to broadcast short text messages, known as tweets. Its underlying social graph is directed, allowing users to choose which of their peers to follow, without necessarily being reciprocated. This approach allows people to access a stream of tweets from the users they follow, enabling an asymmetry where a user's reach is dictated by their number of followers, and not by the size of their egocentric network, enabling

[2]Facebook's graph can also include directed connections, allowing users to follow updates from brands and celebrities, but not in its core frienship model.

celebrities and large organisations to reach a large audience. This subtle design choice has enabled a new class of communication, including more social forms of advertising which we will consider later, but again reflects the robustness of boyd and Ellison's definition.

Privacy challenges

The sudden popularity of SNSs, and their ability to connect large numbers of people around the sharing of potentially sensitive information has given rise to a number of privacy risks.

Bourdieu defines social capital as the potential resources that can be extracted from a network of mutually acknowledged relationships, allowing people to use such relationships to advance their own interests [10]. SNSs encourage the development of social capital between users by making it trivial to disclose personal information to a large social network, while being able to easily access the disclosures made by others. Koroleva et al. propose a conceptual model in which disclosures and reciprocity build social connectedness to create higher value social networks and provide emotional support between peers [69]. The SNS operator relies on its users to keep populating it with new content to avoid stagnancy, and people are compelled to disclose personal information on an ongoing basis to keep building social capital with their peers. This relationship can create a tension. The design of the SNS may compel its users to share more than they would otherwise want to [136], in order to satisfy its own objectives, while users may feel pressure from their peers to keep making disclosures to build social capital [124]. One challenge with managing disclosures to an SNS is that of "context collapse". When communicating face-to-face, people constantly adjust their self-presentation based on their context, particularly who they are with, perhaps adopting a more formal tone with their employer than their close friends, for example. Marwick and boyd note that most SNSs, however, are constructed around an egocentric social network in which all peers of the user are considered equal, collapsing these nuanced contexts into

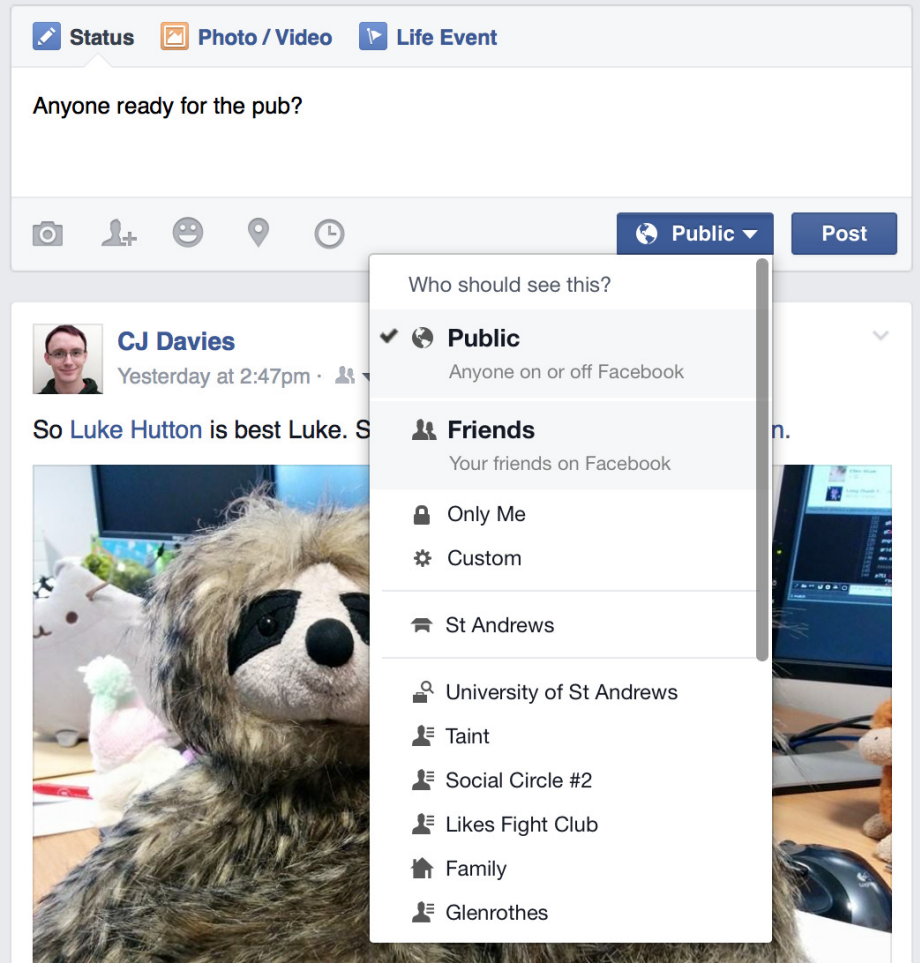


Figure 2.4: The Facebook audience selector allows a user to decide which subset of their social network can see the content they publish, or whether to make it publicly accessible. In this case, the user has created many friends lists that allow content to be targeted at predetermined subsets of people, however controlling the audience in such a way is not common [76].

one [85]. As SNSs depend on ongoing disclosures, they usually encourage users to broadcast disclosures to their entire social network, without consideration for whether that disclosure is relevant, of interest, or appropriate to everyone. As Marwick and boyd argue, this leads either to blandness where people only make lowest-common denominator disclosures, or over-sharing of information to a subset of the user's audience. Marder et al. argue that the difficulty of balancing these distinct social spheres can lead to social anxiety, leading to a withdrawal from making disclosures if over-exposure is anticipated [83].

The causes and problems associated with over-sharing have been widely

discussed in the literature. One issue are the user interfaces used by SNSs to let users choose the audience for their disclosures. In an attempt to mitigate context collapse, Facebook provides fine-grained control of the audience for individual disclosures, allowing users to explicitly choose which of their peers can and cannot see a post, or to create lists of people who are able to see subsets of content, as shown in Figure 2.4. Liu et al. found, however, that only a third of people change the default audience of their posts, which means for most people, all peers in the social network can see it. More concerningly, the actual audience is the same as the user's expected audience in only 37% of cases, with content exposed to a larger audience in most cases [76].

Trust is an important component governing disclosures in SNSs, and is considered to be a key factor in people's privacy calculus when using SNSs. Privacy calculus can be considered the risk-benefit analysis people perform when deciding whether to disclose information, balancing the potential effects of being exposed against the utility that might be gained [73]. Krasnova et al. model disclosure behaviours considering the trust people place in their peers, and in the integrity of the SNS itself, as a factor in the privacy calculus when choosing to make a disclosure [71]. Attacks have been demonstrated which can compromise this trust relationship, and risk the privacy of users. Nagle and Singh found that Facebook users are more likely to accept friend requests and share sensitive information with a stranger if they have a mutual friend in the social network [96], taking advantage of the implicit trust signal given by the mutual friend.

The temporal robustness of privacy settings on SNSs has also been highlighted as a source of concern. Decisions about the audience of content are made at the time of the disclosure, and are likely never revisited [81]. In the meantime, the audience may have changed, if the user has added new peers to their social network, and the user's attitude towards that content may have changed. Wang et al. found that people commonly regretted past Facebook posts, if the audience of their disclosures was greater than they intended, or if the posts dealt with sensitive issues such as drug or alcohol use which the

poster later felt embarrassed by [136], while Bauer et al. found that people often wanted to restrict the audience of their posts, even within a week of having shared them [5].

2.2 Location-based social networks

The increasing popularity of smartphones has enabled a new class of SNS, designed around the purpose of sharing location data with one's social network. Zheng defines location-based social networks (LBSNs) as a form of social network site, where people can record their current location, and accumulate a profile of historic location records. These are distinguished from SNSs which may add location-based elements, as the context and purpose of the system is defined by the dependency on generating and sharing location data [146].

These services build on the work of early location-based services (LBSs), which used self-reported locations or GSM localisation on mobile phones to deliver simple services over standard cellular technologies such as Wireless Application Protocol (WAP) or Short Message Service (SMS), such as recommendations for nearby points of interest [133], marketers sending location-appropriate adverts by SMS [4], or the ability to send messages to physically co-located peers. In recent years, encouraged by the growth of the smartphone market, LBSs have increased in sophistication, taking advantage of the improved location-sensing capabilities of phones equipped with GPS receivers, and the availability of higher-bandwidth cellular Internet connections to provide richer services such as high-resolution mapping with turn-by-turn directions and real-time road closure and traffic incident alerts, and the ability to share one's location with their peers via an existing SNS or one managed by the LBS provider [129].

LBSNs involve the sharing of location data between peers. Tang et al. consider such sharing to either be *purpose-driven*, where people choose to share their location to achieve a utilitarian purpose such as coordinating a meeting,

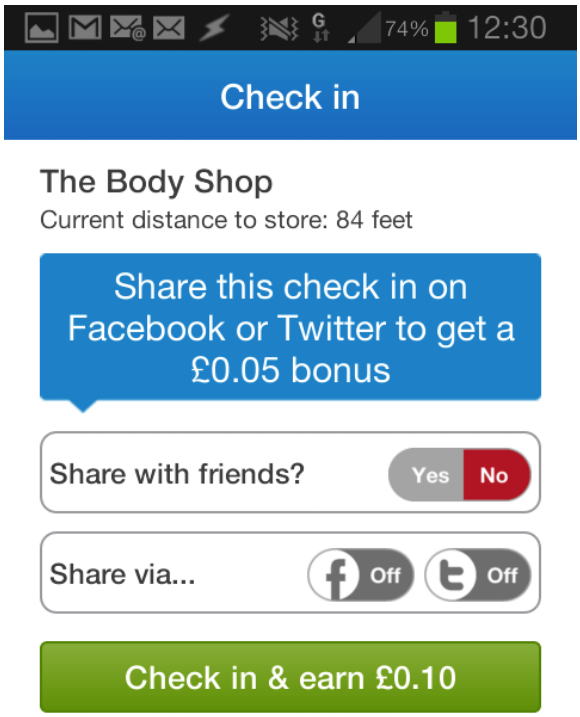


Figure 2.5: Screenshot of the cashback service Quidco that incentivises checking-in with commercial partners and sharing this with one's social network.

or *social-driven*, where people choose to share their location as an act of impression management, and not because the information is necessarily useful to someone else [126]. In recent years, a number of prominent social-driven LBSNs have emerged, such as Foursquare,[3] which allows a user to "check-in" to a place of interest they are currently visiting, and share this with their peers on the service, or to a third-party SNS such as Facebook. As well as different motivations for disclosures, the frequency of location-sharing often varies between these two forms. In purpose-driven services such as Apple's Find My Friends, a user's location is tracked continuously and made available to a subset of peers, while social-driven services often rely on users making discrete check-ins [142].

In recent years, the operators of LBSNs have looked for ways to make revenue from their services, which are usually provided for free. As with many other online enterprises, this has often meant incorporating advertising into

[3]Foursquare:http://www.foursquare.com

their services, and the commercialisation of their users' data. The current location of users is a rich commodity, which had been acknowledged long before such services became commonplace [133]. Using demographic data disclosed by users or drawn from their online presence, advertisers are able to reach highly-targeted groups of people when they are located in a commercially useful area. The perceived value of such services to its users might come from the ability to offer context-specific discounts. Some LBSNs later partnered with advertisers to offer discounts if a user checks-in to a specific business, thus making the user a complicit advertising agent by promoting the brand to their social network for compensation. Foursquare employed a gamification element which crowned users the "mayor" of a particular location if they checked in there the most. In 2010, Foursquare partnered with Starbucks to offer coupons to users who became mayor of a Starbucks outlet [118], a direct form of compensation for users who promoted the brand to their social network. Other examples include sending text messages to people who are near a particular business [89], and electronics retailer Radioshack's partnering with Foursquare to offer discounts to users who achieved a certain number of "badges" by checking in to a set of locations [110]. While these developments augment existing LBSNs with commercial elements, there have also been applications which are built around an entirely commercial premise. Quidco[4] is an online cashback provider, who developed a mobile application where users could see a list of nearby businesses and check-in for a small cash incentive or coupon at a retailer, or receive an additional incentive if this check-in was shared with the user's social network, as shown in Figure 2.5. In addition, Quidco shares user data with its partners when providing this service. [5] We term these new systems *incentivised location sharing* services.

[4]Quidco: http://www.quidco.com/
[5]Quidco Privacy Policy: http://www.quidco.com/privacy-policy

Privacy challenges

The ability to share one's location builds a historic record of a person's mobility, constituting a particularly sensitive aspect of context. Over-exposure of someone's current or historic locations can risk revealing their home address or mobility patterns to an unknown adversary.

Previous work has shown that the users of LBSNs primarily exhibit concern about the context in which someone is making a request for their location, and to a lesser extent, what the purpose of the request is [74]. Barkhuus and Dey found that services which employed continuous tracking by other people induced greater concern than services which detected location for the direct delivery of a service to the user, because of the lack of perceived utility for the person being tracked [3]. Similar concerns were articulated by participants in a user study by Consolvo et al., where location-sharing services which allowed others to request one's location felt "creepy" and could affect relationships with those making inappropriate requests [18].

As with other SNSs, the trustworthiness of LBSNs is a significant concern. Tsai et al.'s survey of LBSNs found that a third of services did not have a privacy policy to explain how users' data would be processed [131], and that the majority of services which did have policies retained permissions to store personally-identifiable information (PII) indefinitely, which includes any data which can be used to uniquely distinguish an individual [86]. The context-dependent sensitivity of location disclosures places great responsibility on the user to manage the interfaces used to control who can view check-ins. The authors also noted that while three-quarters of services expose some privacy controls, these are often not made obvious to users.

While much of the research concerning privacy risks of LBSNs have focused on whether the service provider can be trusted, and the social implications of over-sharing with one's peers, some services are explicitly designed to expose one's location to strangers, with potentially more damaging consequences. The

service Grindr is popular with men who have sex with men to discover and talk to others who are nearby. Fiebig et al. found significant security shortcomings which allowed user profiles and social networks to be exposed by fabricating requests to the service [29], while a report by Synack researchers found that the exact location of users could be identified through trilateration when making spoof requests [125]. This has implications in cultures where people may not be able to be forthcoming about their sexual identity and could be the subject of abuse or implicated criminally, with Toch and Levi noting that harassment of users of such services was not uncommon [130].

Concerns with LBSNs that employ mobile advertising have been identified. King and Jessen note that data protection regulation in the EU and US have gaps which could lead to abuses of behavioural advertising, specifically citing the lack of consent being sought before collecting or using personal data, surveillance of people's behaviour, and unwanted solicitation of commercial activity [67]. This is of particular concern considering the sensitivity of data being collected and shared between partners. Consent is found to be an important feature in acceptance of mobile advertising, with Scharl et al. observing that the intrusiveness of mobile advertising is liable to being perceived as spam if people do not feel in control of what they receive [112]. Similarly, Xu and Teo's theoretical study of privacy concerns in such applications found that the perceived benefit, particularly if the implementation of advertising was considered entertaining, was a significant predictor of people's acceptance of having advertisements targeted to them [145].

In this section, we have introduced some instantiations of social network sites and location-based social networks. The novelty of these services, and the blending roles between them, has significant implications for how people consider and use these services.

In the next section, we consider the challenges posed by attempts to research these services from ethical and privacy perspectives.

2.3 Studying social network sites

SNSs have become popular over a short period of time, around which a new field of research has quickly formed. SNSs are an attractive set of technologies to examine for myriad fields. For social network analysts and graph theorists, they provide an explicitly defined social network with no need to infer the structure, representing an advance on previous sources of social network data such as email correspondence and phone logs [25], allowing for rich analysis of the properties of the underlying network [132]. For social scientists, the wealth of data held by SNS operators allows for observation of social dynamics and human behaviour on an unprecedented scale, with the ability to target niche audiences that might otherwise be difficult to recruit for such studies. Wilson et al. conducted a comprehensive review of the range of social science research enabled by Facebook [141]. Examples of the research that has been enabled includes Getty et al.'s study of how people use Facebook to grieve [37] and Powell et al.'s study of how insomnia sufferers use Twitter to discuss their condition [106]. The insights such research yield can be of high value, but perhaps due to the relative novelty of the field, and its quick pace of development, there has been little consideration of the implications of conducting experiments which use data derived from SNSs, and a lack of best practices for researchers to follow.

In this section, we outline some fundamental aspects of conducting human subjects research, and consider how these relate to the study of SNSs.

2.3.1 Research ethics

Ethics is a broad area of philosophy, attracting study across many disciplines, and can be considered the norms for conducting one's self to allow acceptable and unacceptable behaviour to be detected [109].

There are two prevailing schools of thought concerning the application of ethics to the research context, framed as either *deontological* or *consequentialist*

in nature. The former suggests that individuals and groups have duties which must be upheld, irrespective of the actual consequences of executing these duties, and is articulated in Kant's formulation of the Universal Law, allowing an ethical judgement about the appropriateness of an action to be made if one can expect all reasonable people in the same context to make the same decision [62]. Conversely, consequentialist ethics holds that the appropriateness of an action is determined by its actual impacts. This utilitarian approach to ethics is articulated by Bentham, who suggests ethical decisions are reached based on maximising the benefit to the greatest number of people [7].

While both approaches may reflect common ethical thinking, they have been criticised. Hegel rejects Kantian deontological ethics by demonstrating situations in which universality cannot be reached, incorrectly judging an action as immoral [46]. Kagan criticised consequentalist ethics for suggesting that people would make dispassionate decisions which are likely to negatively impact their immediate social network in favour of the greater number of strangers who may be impacted by a decision, known as the demandingness objection [60].

The Belmont Report articulates a consensus of how these ideals can be applied to the research context, identifying four principles: autonomy, beneficence, non-maleficence, and justice [127]. In this thesis, we primarily consider autonomy, which the acquisition of consent and appropriate handling of personal data aim to uphold.

Resnik argues that research ethics ensure the following are true [109]:

1. research artefacts, such as papers, do not misrepresent what the data show, and truthfully reflect what was found.
2. collaborations can be achieved in a culture of mutual respect and trust.
3. publicly funded research can be held accountable by mandating protections for human subjects and animals, and to avoid misconduct.
4. public support for research is maximised by highlighting the responsible

manner in which it is conducted.

5. research promotes moral and social values by protecting human subjects, the welfare of animals, and being socially responsible.

In this thesis, we are particularly concerned with final objective, as we consider the wider social ramifications of SNS research which does not uphold appropriate ethical norms.

Funding agencies and professional bodies have provided guidance on how to ethically conduct research, which collectively forms the norms against which the conduct of research can be assessed. The Association of Internet Researchers provide a comprehensive process for considering and addressing ethical issues in online research [84] which provides more relevant guidance than policy which is optimised for printed materials and face-to-face interventions. These guidelines include some consideration of challenges in SNS research, such as the ability to reidentify participants by combining data from multiple SNSs, but does not specifically address the issues posed by SNSs. Next, we consider the protections and mechanisms which aim to support ethical research, and their applicability to the SNS domain.

2.3.2 Human subjects research

Walther defines human subjects research as:

> "That in which there is any intervention or interaction with another person for the purpose of gathering information, or in which information is recorded by the researcher in such a way that a person can be identified directly or indirectly with it." [135]

Until the late 1940s, there was little attention paid to the ethical conduct of human subjects research. The conduct of experiments on prisoners of war

and civilians during the Nazi regime in the Second World War, and subsequent Nuremberg trials, led to the formulation of a ten-point code which could be used to identify legitimate medical research [8]. While not immediately adopted, it led to international discussion on how to appropriately address the newly-acknowledged issue of research ethics. The inconsistent application of the Nuremberg Code by individual countries, and perceived restrictions on research by its absolute requirement for informed consent [6], led to the Declaration of Helsinki as an expression of international consensus regarding research ethics, acknowledging instances in which consent may not be necessary, or feasible to acquire, and permitting the use of proxies to make consent decisions [144]. As a result, most research institutions, and even some legislatures, have mandated ethical requirements for conducting human subjects research, administered by what are generally known as ethics committees or institutional review boards (IRB). These requirements were originally designed to ensure that subjects of biomedical research were sufficiently protected, and to act as a liability shield for institutions. Over time, the role of IRBs increased in scope, becoming enshrined in law, with IRB oversight required to secure federal funding for such research in the United States. Also, IRBs routinely began to monitor a wider range of human subjects research, including social sciences, legal studies, and the humanities [64]. Despite this evolution in purpose, however, IRB remains optimised towards conducting biomedical research, and the appropriateness of the ethical instruments it prescribes for certain kinds of research is unclear, with some information and communications technology (ICT) researchers considering IRB oversight to be an irrelevant hindrance, with overly bureaucratic procedures [120].

There is fundamental disagreement about whether Internet research, including the study of SNSs, can be considered human subjects research. As IRBs only scrutinise such studies, the distinction is important. Many online studies involve a researcher intervening in a participant's context, or observing their behaviour remotely, and the layers of abstraction presented by computer-mediated research can make the distinction less clear. McKee and Porter's ex-

amination of the decision charts used to determine whether a study is eligible for IRB scrutiny found they have ambiguous applicability to many forms of online research, raising concerns about the dehumanising effects this framing can have when conducting research online [87]. Solberg argues that data mining from SNSs is low-risk as the Internet is a public space and Facebook users have limited expectations of privacy. As a result, she suggests that this should not bes considered human subjects research and exempt from further scrutiny [116]. McKee and Porter contend, however, that even without direct intervention between researchers and participants, researchers will feel a greater sense of obligation to participants when considering the artefacts of research as coming from human participants, and not just mining digital texts. They argue that having material published on the Web should not be considered explicitly public in the way that a printed publication might, and that such policies do not sufficiently capture the nuances of online publishing [87].

2.3.3 Informed consent

The Nuremberg Code encoded the importance of consent in research studies. Subsequently, the importance of consent being *informed* has been adopted by the research community, which the Council for International Organisations of Medical Sciences defines as:

> "A decision to participate in research, taken by a competent individual who has received the necessary information; who has adequately understood the information; and who, after considering the information, has arrived at a decision without having been subjected to coercion, undue influence or inducement, or intimidation." [20]

The European Union's Data Protection Directive defines that in any data processing context, consent means:

> "...any freely given specific and informed indication of his wishes

by which the data subject signifies his agreement to personal data relating to him being processed." [28]

Friedman et al. propose a model for acquiring informed consent online, consisting of five components [33]:

1. **Disclosure:** Providing information about the potential benefits and harms of participating in research.
2. **Comprehension:** Determining that the participant understands what is being disclosed, perhaps by restating it in their own words, or being given the opportunity to ask questions of the researcher.
3. **Voluntariness:** Ensuring that participation has not been coerced.
4. **Competence:** Ensuring the participant has the emotional, mental, and physical faculties to give informed consent.
5. **Agreement:** Providing an explicit opportunity to accept or decline to take part.

We can consider the acquisition of consent to broadly take two forms: *secured*, and *sustained*, as defined by Luger and Rodden [80]. *Secured consent* might take the form of a single checkbox asking someone to agree to the terms of participation, such as in an end-user license agreement (EULA). In an experimental context, this might mean potential participants provide consent at the beginning of the study to the collection and processing of data, while removed from the context of the actual data collection. As the language used in such forms is often beyond what a literate adult can understand, it is not clear that securing consent in this fashion ensures the consent is meaningful [79]. While this approach is trivial for participants to complete, it risks violating their privacy if data are collected and processed in a way they did not expect. 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 [115] examination of self-censorship on Facebook, where participants chose when to share information with researchers over a period of time. This approach has the advantage of directly gaining consent in the context of the data collection, meaning participants are better placed to make an informed decision about individual disclosures, rather than sanctioning a wider, less well defined period of data collection. This comes at a cost, with participants required to spend longer making consent decisions, which may prove frustrating and contribute to attrition, which is a common challenge in experience-sampling experiments [47].

As with the ambiguity of whether studying SNSs constitutes human subjects research, the applicability of informed consent to the study of SNSs is contentious. Neuhaus and Webmoor suggest that obtaining informed consent when conducting large-scale SNS studies, which may involve hundreds of thousands of participants, is impractical, and that simply importing ethical practices from the "offline" world is impractical [98].

2.3.4 Ethics controversies in recent SNS research

The ethical issues associated with SNS research have been highlighted by two studies which have attracted debate.

In 2008, researchers at Harvard University released a dataset of ostensibly anonymised Facebook profiles from a cohort of undergraduate students, named "Tastes, Ties and Time" (T3). As explained by Zimmer, based on the information provided by the researchers to aid analysis of the dataset, it was possible to infer the college the dataset was derived from, and as some attributes were only represented by a single student, it was likely that individuals could be identified [147]. It is important to note that in this case, the researchers made good faith attempts to protect the identity of the members of their dataset, without appreciating the susceptibility of the data to trivial reidentification attacks. In addition, their data collection was approved by their IRB, with the

authors noting that because the researchers did not directly communicate with the people whose data were collected, it was not a source of concern [147]. This case highlights the limitations of the IRB model when those tasked with reviewing experiment protocols are not sufficiently aware of the risks to participants in online research. As discussed earlier, this returns to the fundamental question of whether the subjects of remote data collection are "participants" at all, and what responsibility researchers have to them. Zimmer concludes that the availability of information on an SNS does not make it fair game for researchers [147].

In 2012, researchers at Facebook's Data Science Team conducted a large scale experiment on 689,003 Facebook users. The study manipulated the presentation of stories in Facebook's News Feed product, which aggregates recent content published by a user's social network, to determine whether biasing the emotional content of the news feed affected the emotions people expressed in their own disclosures [70]. The study attracted attention because of several ethical concerns:

1. **Commercial exemption from IRB** As discussed earlier, academic research in the US, and many other countries, is subject to IRB oversight if the research is deemed to involve human subjects. Commercial researchers, including Facebook, have no such regulatory obligation. Although researchers at Cornell University were co-investigators, their institution did not require IRB review because the data collection was handled exclusively by Facebook.[6] Fiske and Hauser argue that as the intent of Facebook's study was to contribute to "generalizable knowledge", and not exclusively for internal product improvement, their responsibility should equal that of academic institutions [31].

2. **Asymmetry of commercial data collection:** Commercial organisations, such as SNS providers and telecommunications operators regularly publish research which has only been possible because of their ability to repurpose

[6] Cornell statement: http://mediarelations.cornell.edu/2014/06/30/media-statement-on-cornell-universitys-role-in-facebook-emotional-contagion-research/

vast amounts of internal proprietary data which no academic institutions can likely reproduce, because the data are both unique to such organisations, and generally not made available to other researchers. This makes studies such as Facebook's unreproducible, meaning the academic community is unable to validate their findings. Neuhaus and Webmoor reflect on this asymmetry from a different perspective, noting that academic researchers have a greater ethical obligation, because unlike commercial operators, they are not using data from people in exchange for a service which returns benefit to the individual [98].

3. **Informed consent:** As discussed earlier, acquiring informed consent is considered a key instrument when conducting human subjects research, as mandated by IRBs. The Facebook study was conducted, however, without participants' knowledge. As of January 2015, Facebook's Data Policy contains a clause stating "We use all of the information we have to help us provide and support our Services...We conduct surveys and research, test features in development, and analyze the information we have to evaluate and improve products and services, develop new products or features, and conduct audits and troubleshooting activities." [7] Before this, a similar clause declared that user data was used "for internal operations, including troubleshooting, data analysis, testing, research and service improvement". Harriman and Patel argue that this clause implies research was only for internal product development, and not for academic publication [45]. The intended reach of this clause is rendered irrelevant, as at the time the study was conducted, this policy did not include any reference to research [49], avoiding any semblance of consent from users.

2.3.5 Reproducibility

So far, we have considered the ethical issues concerning the relationship between researchers and the participants of SNS research. We now consider the respons-

[7] Facebook Data Policy: https://www.facebook.com/full_data_use_policy

ibility researchers have to other stakeholders including funding agencies, the wider academic community, and the general public who benefit from and often fund research activity. The conduct of research in a manner which satisfies these stakeholders, and considers its ethical and sustainability impacts, is considered responsible research and innovation (RRI) [119]. An important consideration is the reproducibility of experiments [128]; that a researcher ought to be able to take a previous experiment and perform it again, or build on it to create a new study, which is considered critical to the scientific method [26]. To ensure reproducibility, the entire experimental workflow must be considered, from the initial collection of data, its processing and analysis, and the dissemination of artefacts such as research papers. Thompson and Burnett [128] suggest that reproducibility consists of three elements:

1. Supporting computationally intensive research, by sharing source code, tools, and workflows for executing these tools.
2. Supporting structured analysis, by encoding the scripts that conduct analyses and produce components of publications such as tables and figures.
3. Allowing the dissemination of research artefacts, such as papers, and raw data. Rather than treating papers as a static piece of text, they should include, or provide access to executable code, and other resources needed for replication.

We can think of these elements as broadly encapsulating three key themes: code, methods, and data, respectively.

There are particular challenges to conducting SNS research in a reproducible manner, some of which arise from the tension between social science and systems-oriented approaches which manifest in much SNS work. The definition of reproducibility we use is optimised towards computational research in which methods can be explicitly encoded, analyses automated, and results disseminated. This workflow, however, is contrary to the way in which large amounts of social science is conducted, where ethnographic approaches are

reactive to the phenomenon being observed, and are often difficult to reproduce. The publication of Goffman's ethnographic study of an underprivileged neighbourhood in Philadelphia [39] drew attention to the challenge of validating ethnographic research because of attempts to anonymise participants, and insufficient documentation of people's accounts made it difficult for others to verify the accuracy of her findings [99]. When social scientists turn to computational methods, including the study of SNSs, the lack of an embedded reproducibility culture can impede progress towards sharing of methodological details. Expectations of reproducibility in different communities often relate to the inconsistent application of terminology and their implications. Drummond distinguishes between replicability and reproducibility, where the former aims to completely replicate a previous experiment without deviation, while the latter means designing a unique experiment which arrives at the same result as a previous study [27]. Again, while this realisation of reproducibility may be an attractive ambition in many computational and physical sciences, attempting to design an independent social science study which produces the exact same result as a previous study is unlikely, and arguably not particularly desirable. From these interdisciplinary disagreements, sharing of best practices between disciplines should be encouraged, however there are SNS-specific challenges which can make reproducibility difficult.

Most major SNSs, such as Facebook and Twitter, provide fettered access to their data through application programming interfaces (APIs), the use of which is subject to a license agreement. These providers assert control over how the data that they host are used, and actively disallow large datasets of their content to be published.[8,9] This may impede the third tenet of reproducible research, particularly when work concerns a specific corpus of content, such as Denef et al.'s examination of tweets during the 2011 London riots [24], rather than a random sample of content generated by a certain population. If SNS data cannot be directly shared, then it might be possible to instead repeat the experiment, but only if the sampling strategy of the original experiment can be

[8]Twitter Developers Terms of Service: https://dev.twitter.com/terms/api-terms
[9]Facebook Platform Policy: https://developers.facebook.com/policy

replicated. This is challenging, however, when papers do not sufficiently disclose how their participants were recruited, and data were collected. In user studies where users interact with SNSs, the range of variables make it difficult to replicate the participant's experience, from the text used in prebriefing and consent documentation, through to the implementation of user interface elements, which can affect how people engage with the study [88].

Where research is dependent on the use of APIs provided by SNSs, there are additional challenges to reproducibility. Code that evokes certain API endpoints is dependent on that API being online and its design remaining consistent, which may be an impractical expectation for actively developed services where new features are developed and retired over time. In 2013 alone, Facebook announced seven sets of "breaking" changes, where developers needed to amend their code if it used certain features, incorporating the change or withdrawal of 47 API endpoints.[10] As some of these changes concern the removal of features, some legacy code will not be usable even if actively maintained. This is a significant challenge to reproducing results dependent on live SNS data. More recently, Facebook has introduced an API versioning scheme which will go some way to improving this situation, but retired API versions will only receive support for one to two years[11], and such approaches are not common to all SNSs.

Twitter provides two APIs for accessing tweets that are published in real-time. Their "streaming" API, which is widely available, allows developers to access a small sample of the tweets that are generated, which might be appropriate for visualisations of overall activity, but without the full fidelity necessary to investigate smaller populations. A "firehose" API is made available to certain commercial partners, and provides unfettered access to the tweets being generated [12]. Morstatter et al. find that the firehose API provides a significantly different sample of tweets than the streaming API [93]. These challenges

[10]Facebook Platform Roadmap: https://developers.facebook.com/roadmap/completed-changes
[11]Facebook Platform Upgrade Guide: https://developers.facebook.com/docs/apps/upgrading
[12]Twitter Streaming APIs: https://dev.twitter.com/streaming/overview

restrict researchers' ability to replicate studies if they are not able to collect a similar distribution of content, depending on their license agreement with the SNS provider. For example, De Choudhury et al. leveraged their corporate firehose access to collect depression-indicative content which others might not be able to recreate [23]. In many cases, access to the original data is not necessary. Unlike some more theoretical fields where reproducibility may concern the replication of results by seeding a simulation with data, or evaluating a statistical model, many SNS papers consist of user studies which use SNSs as a conduit for examining behaviour of a population. In such instances, replication of methods is key. For example, even subtle changes in the presentation of consent forms can have an impact on how people interact with an experiment [88], and even the act of asking for consent may bias results [111]. Failure to encode such methodological details can make it difficult to accurately replicate studies and meaningfully compare results.

The difficulty of adequately anonymising sensitive SNS data is another challenge. Anonymisation has a temporal quality - what might be sufficiently obfuscated today may be deanonymised tomorrow. Narayanan and Shmatikov demonstrate how many apparently anonymised datasets simply replace names with random identifiers, rather than obfuscating uniquely identifying attributes, permitting re-identification [97]. Dawson surveyed 112 articles to show participants quoted from public Web sources could trivially be reidentified [22]. Sufficiently protecting the privacy of participants after their data have been released, while maintaining their utility in further studies, is a constant tension for SNS research.

Despite being a fundamental aspect of the scientific method, reproducibility in computational sciences has only recently been identified as a significant issue. Stodden surveyed researchers to understand attitudes towards reproducibility [123], finding that while researchers value the benefit to the scientific community that increased sharing of data and source code can yield, the time taken to prepare artefacts for distribution, as well as legal and administrative barriers, was a significant disincentive.

2.3.6 Operationalising privacy with contextual integrity

Privacy is an ever-evolving concept with myriad, and often conflicting, definitions and applications. Warren and Brandeis' late 19th century definition of privacy as "the right to be left alone" [137] informed substantial public debate about privacy throughout the next century, in a culture where technology risked exposing the lives of the individual to the masses. Rejecting the breadth of this definition, Gavison defines privacy as concern for limited access to the self: the extent to which we are known, physically accessible, or the subject of others' attention [35], and rejects the notion that privacy is a form of control over information, as articulated by Westin [138].

The difficulty in arriving at a definition of privacy which is sufficiently expressive but not overly broad remains an open problem. In this thesis, we adopt the principle of privacy as a form of maintaining limited access to the self, however this is insufficent to measure the privacy impacts of a process, and to recommend best practices to mitigate these impacts. Solove proposed a taxonomy of privacy which models the relationship between the data subject and data holders, and the role of information collection, processing, dissemination, and potential invasion [117]. Solove's framework is intended to help broaden the understanding of privacy violations and their impacts, but is not comprehensive enough to identify and mitigate issues in individual processes. We therefore assess the applicability of a conceptual framework for identifying and resolving privacy breaches. Our focus in studying SNSs is to consider the ethical and privacy implications of the study of such systems, as well as the issues associated with the design of the systems themselves. Therefore, we adopt contextual integrity, a theoretical framework proposed by Nissenbaum [100] for considering information privacy. Avoiding the narrow definitions of her contemporaries, Nissenbaum does not define privacy so much as propose a model for identifying the source and impact of privacy violations. 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.

Self-reported scales are often used to capture people's qualitiative concern about information-sharing practices. Westin's privacy indexes have been used in a number of studies [72] to show emerging concerns in domains such as consumer and medical privacy, since the early 1990s. Concern about the rise of direct marketing and, later, online information collection, led to the development of domain-specific models such as the Internet Users' Information Privacy Concerns (IUIPC) scale [82]. These scales are attractive because they provide a means of operationalising an inherently difficult concept such as privacy. In this thesis, however, we choose to adopt contextual integrity as our model for studying privacy issues rather than some of these extensively studied and validated scales, as they are focused on the individual's perception of privacy violation, which can be difficult to assess where information collection and processing is unseen or too complex for laypeople to articulate. In addition, these scales are not sufficient for identifying the source of a privacy breach, or the wider societal impacts of these practices. As such, contextual integrity's attraction is that it provides a vocabulary for examining and mitigating the issues with emerging systems.

Decision heuristic

To aid the analysis of information-sharing practices with contextual integrity, Nissenbaum provides a nine-step "decision heuristic" to analyse the significant points of departure created by a new process, thus determining if the new practice represents a potential violation of privacy. The first six steps involve modelling the existing and new contexts, allowing a *prima facie* judgement to be rendered as to whether the new process significantly violates the entrenched norms of the context. The final steps of the heuristic involve a wider examination of the moral and political implications of the process to make a recommendation as to whether the new practice should be adopted. These steps are as follows:

1. Describe the new practice in terms of its information flows.
2. Identify the prevailing context in which the practice takes place, which should be suitably broad such that the impacts of any nested contexts can be considered.
3. Identify the subjects, senders, and recipients of information.
4. Identify the transmission principles: the conditions under which information ought (or ought not) to be shared between parties. These might be social or regulatory constraints, such as the expectation of reciprocity when friends share news, or the obligation for someone with a duty of care to report when their ward is in danger.
5. Identify any applicable entrenched informational norms in the context, and identify any points of departure the new practice introduces.
6. Making a *prima facie* assessment: there may be a violation if there are discrepancies in these parameters, or if there are incomplete normative structures in the context to support the new practice.
7. Consider the moral and political factors affected by the new practice. Could

it affect people's freedom or autonomy, impact on power structures, justice, or the execution of democracy?

8. How does the new practice affect the goals or values of the prevailing context? If there were any implications identified in the previous step, how do they effect the goals of this context?

9. Finally, make a determination as to whether the new process violates contextual integrity based on a consideration of these wider factors.

As a diagnostic tool, the decision heuristic further supports our decision to study privacy with contextual integrity, as it provides a consistent mechanism for assessing the impacts of an emerging process, which consumer-focused privacy scales do not extend to.

2.4 Summary

In this chapter, we have introduced the range of social network sites that we will consider in the remainder of this thesis, and discussed the ethical issues associated with conducting research of such systems. We note the following:

- SNSs use knowledge of the relationships between people to deliver socially-appropriate services over the Internet.
- SNSs such as Facebook and LBSNs such as Foursquare are becoming increasingly popular with people; however each raises serious privacy implications.
- The increasing popularity of such services is encouraging research into how they are used, and how they can be improved, but such research is fraught with a number of ethical and methodological challenges.

In the next chapter, we consider recent research that aims to tackle the issues

with SNSs we identified, and examine proposed solutions to the methodological issues with researching such systems.

Chapter 3

State of the art

In this chapter, we explore recent research that attempts to resolve the privacy challenges we identified earlier with several forms of social network sites (SNSs). We also survey the SNS literature to determine the extent to which work in this domain tackles the issues of ethics and reproducibility discussed in Chapter 2.

3.1 Resolving privacy challenges in location-based social networks

In Chapter 2.2 we outlined a number of privacy challenges identified in the literature concerning LBSNs. In this section, we discuss recent efforts to resolve these issues and identify open problems.

We noted that the context and purpose of requests for location data was a significant predictor of concern, whether from a peer, or a third party such as the LBSN provider or an advertiser. Hoyle et al. propose using feedback about attempts to access data to mitigate such concerns. Their system changes the apparel of an avatar representing the user to illustrate how contextually inappropriate their exposure is [50]. Such an abstraction allows for a quick

comprehension of how exposed someone's context is, but does not include detail of discrete requests, which may limit someone's ability to judge the suitability of these requests.

The availability and ease-of-use of mechanisms to allow users to control who has access to their data was identified as an issue. Knijnenburg et al. find that simply providing more fine-grained controls does not lead to better privacy decisions, and note that application designers should consider the tangible privacy and benefit trade-offs each option represents [68]. Finnis et al. propose a policy-specification language that could be used to mitigate differences between services. Allowing applications to declare when location data are needed and at what precision allows restrictions on their capabilities to be enforced and communicated to users, who are otherwise given coarse controls for the permissions of applications [30]. Their work does not examine, however, whether users can comprehend the constraints the policy language enables and make informed decisions about the suitability of applications. Christin et al. propose "privacy bubbles" as location-specific spheres where only those within the bubble can access the content generated within it [16]. Such an approach tackles some of the usability challenges issues with ad-hoc sharing in colocated environments, such as allowing a group of tourists to share photos without granting each other permanent access to each others' content. This also makes important contributions towards exposing these settings in an easy to comprehend manner, but it is left to further work to determine whether such a metaphor would be appropriate for other types of sharing. Kapadia et al. propose "virtual walls" as a system for controlling access to contextual information, such as location, in pervasive sensing environments [63], using a metaphor of walls of differing transparency to communicate the types of information that can be exposed. While they found that the metaphor can be understood, its applicability in the wild has not yet been demonstrated.

As we introduced in Chapter 2.2, incentivised location sharing augments traditional LBSNs with the addition of financial incentives, and the sharing of personal information with third party merchants. This can compound con-

cerns with the trustworthiness of LBSNs. Moniruzzaman and Barker developed a form of ILS that requires users to check-in to a participating business a certain number of times before receiving an incentive. As disclosing an entire check-in history to a merchant in order to validate whether a deal should be offered could risk the privacy of users, their framework minimises the amount of information transmitted to merchants. The framework also employs a recommendation engine to suggest to users at what interval they should check-in to minimise the chance of mobility patterns being inferred by an adversary [90]. While such an approach resolves some of the risks to users when disclosing their location, it places responsibility on the end-user to take advantage of the system to maximise their benefit, and does not address the underlying issues with allowing users to control the use of their information in an ILS application. Kahl et al. proposed a privacy-preserving architecture for mobile advertising that used the provider of an SNS as a middleman for matching adverts to relevant groups of users, based on contextual knowledge of the users, and without transmitting data about individual users to advertisers [61]. As well as supporting the traditional business-to-consumer (B2C) advertising paradigm, the architecture also considered a consumer-to-consumer (C2C) model used in viral advertising campaigns, where one person advertises to their social network. While the framework considers how influential members of a social network could be identified with a hope to spreading the advertising message, and notes that such peers would likely need to be given an incentive to do so, this work does not directly consider the implications of incentivising such C2C advertising activity. Haddadi et al. propose a mobile advertising system in which determining the appropriateness of adverts to an individual user is performed on the user's own device, and does not rely on a middle-man with understanding of the users, unlike Kahl et al.'s model [44]. The system distributes adverts to users through cellular multicast protocols or from 802.11 hotspots, and leverages delay-tolerant networks to anonymously report engagement statistics for ads, adding additional layers of protection to users. Again, while the authors do consider that incentives would likely be needed to encourage adoption of such a system, consideration of their impacts was beyond the scope of their work.

The potential for sensitive location data to be intercepted by an adversary has drawn attention to the need to protect the privacy of location information, particularly when the LBSN operator, or a third-party merchant, is not trusted. Such information must be preserved while maintaining the utility of the service. Skvortsov at al. propose a system in which individual operators are sent low-precision location data, fusing the traces given to different services to allow individual operators to use higher-precision data [114]. While this is robust to a number of attacks, both external, and by an untrustworthy LBSN provider, it would need to be coupled with a usable model for controlling the rights given to each operator, which may be too abstract for end-users and divorced from the utility of such services.

In this section, we have outlined solutions to some of the privacy challenges with LBSNs. We note there are many strategies that aim to secure the transmission of location data when the provider of an LBSN, or other peers in the service, cannot be trusted. New interfaces which abstract users from the complexity of configuring privacy policies, and visualisations of how their information are being accessed, attempt to mitigate some of the concerns about how to control access to information, and feel more comfortable that the context of requests is appropriate. There has also been some consideration of how to manage the addition of commercial entities to the equation, mostly from trust perspectives. There are, however, open questions not yet resolved by the literature. When commercial actors are involved, do people comprehend the flows of information involved? Are they able to make informed decisions about what should be shared with whom, based on this understanding? As the addition of incentives to these services is still relatively novel, there has been little consideration of the privacy and usability impacts they represent. Later in this thesis, we will tackle these outstanding issues.

3.2 Informed consent

In Chapter 2.3.3 we outlined some of the issues with acquiring consent in SNS studies. In this section, we discuss recent approaches towards consent, and consider their appropriateness to SNS research.

As discussed earlier, informed consent is rooted in biomedical research, and it is in this domain that most recent discourse is centred. In the collection of samples for biobanks, consent has traditionally been considered *broad*, where donors give consent to their samples being used in future research within an agreed framework [122]. This is distinct from most SNS studies we consider to be self-contained, where a participant consents to an individual study, where data are collected or generated, and the participation ends. While similar to our notion of *secured* consent, it is not strictly analogous, as donors are expected to give consent again if any aspect of the framework changes. Kaye et al. suggest that this is insufficient, as it is too open-ended to allow donors to make informed decisions about what might happen to their samples in the future. They propose *dynamic* consent [65] as a solution to this, applicable to domains beyond biobanking. This approach is designed to engage participants in research over time, using the Internet to let them see how their samples have been used, to receive requests for consent for new studies and to request additional information and give consent, with the participant determining the extent to which they engage in the process. We consider this a manifestation of *sustained* consent, as it does not rely on consent sought at a single point in time. Despite its grounding in biomedical research, it is likely applicable to SNS studies. For studies where data are collected over a period of time, a dynamic approach to consent that considers the changing contexts in which the data are collected and the participant's wishes in those contexts might be more appropriate. In addition, researchers who wish to share datasets collected from SNSs with others could employ dynamic consent to determine whether participants sanction individual studies, and to give feedback about how their data are used.

Luger and Rodden discuss how such an approach might be applicable to ubiquitous computing, where devices are continuously tracking, processing, and sharing sensor data about people and their environment. They argue that consent needs to be sensitive to the context in which data are collected, and evolve beyond simply notifying people of data collection, and instead meaningfully inform them of how and why their data are being used over time [80]. Morrison et al. demonstrate how this could be adopted in mobile experiments, by presenting feedback to users about what data have been collected. Their finding that visualisation of these data increases the attrition rate in such studies raises the question of whether these participants would have declined to participate had they understood the implications of a study that employed a secured approach to consent [92]. Munteanu et al. advocate a "situational" approach to ethics in HCI research, arguing that "static" ethical approval is detached from the realities of conducting ethnographic studies in the field, and that traditional consent instruments can be difficult to administer in practice, particularly when working with vulnerable people. They recommend involving HCI practitioners in the creation of ethics policy, and allowing flexibility in ethics protocols to adapt to the research context as it unfolds [95]. Similarly, Neuhaus and Webmoor propose "agile ethics", arguing acquiring traditional forms of informed consent is not practical when conducting large scale experiments using SNSs. Informed by agile software engineering techniques, they suggest that consideration of ethics should be ongoing, and rooted in an understanding of all actors involved, rather than focusing on up-front documentation and fixed IRB protocols. They argue that an approach that considers the realities of data collection can mitigate the lack of consent acquisition and lead to more responsible research [98].

Steinsbekk et al. argue that a dynamic approach to consent is not necessary, because it burdens participants with the need to sanction each additional research project, even if there is no significant ethical departure, and transfers ethical responsibility from IRBs to individual participants, which reduces ethical oversight [122]. Friedman et al. also caution that overwhelming participants

with information and consent interventions can have the counter-productive result of disengaging participants and numbing them to the process [34]. In focus groups with medical researchers, Whitley et al. find concerns about the logistical burden of providing a framework for revoking consent, both in terms of the time and cost in administrating it, and the difficulty in planning trials when higher attrition might be likely [139]. Gomer et al. propose using a semi-automated agent to make consent decisions on behalf of a user to reduce the burden placed on individuals. These decisions are informed by preferences expressed by the user, and refined through periodic review [41]. Similarly, Moran et al. consider how to negotiate consent in multi-agent environments, suggesting that identifying interaction patterns can gain insight into appropriate moments to acquire consent [91].

From this tension, we can identify two spectra on which consent can be measured: *accuracy* and *burden*. That is to say, does the consent instrument capture the willingness of the participant to have their data used in a particular context, and how much time is spent and cognitive load placed on the participant to acquire consent? Kaye et al. argue that broad consent minimises burden at the cost of accuracy, while dynamic consent maximises accuracy at the cost of burden. Conversely, Steinsbekk et al. suggest that broad consent minimises burden while achieving acceptable accuracy, while dynamic consent substantially increases burden with no demonstrable improvement to accuracy.

Without any quantitative measurement of the interaction between these two variables, it is difficult to determine what form of consent is most appropriate to SNS studies. Attempting to formalise understanding of this remains an open problem.

3.3 Reproducibility in SNS research

In Chapter 2.3.5, we noted that reproducibility consists of three elements, which relate to the capture and sharing of code, methods, and data. In this section,

we discuss the state of the art in addressing these issues, and consider the applicability of these methods to the study of online social networks.

There are increasing efforts towards supporting the sharing of source code and data. One such example is the data-sharing repository FigShare[1], which allows researchers to upload research artefacts and generate digital object identifiers (DOIs), which allow such artefacts to be cited by others. Gent proposes a recomputation framework that allows legacy experimental code to be executed in the future by encapsulating the environment needed to execute it in a virtual machine image [36]. Such an approach is optimised towards experiments that are inherently recomputable, such as simulations, but its applicability to human subjects research, such as that involving SNSs, is unclear, where replications may depend on the availability of external services, and participants to provide new data.

Another area of interest is to encode the workflow of an experiment. This speaks to the second pillar of reproducibility: encapsulating the methodology of a study. Domain-specific tools for workflow capture have been proposed, but due to the specific challenges found in each field, these are not universally applicable, or appropriate for SNS research. The development of standards for sharing domain-specific data allows researchers to collaborate on shared problems For example, biologists have developed standards for representing proteomics data, such as the protein sequence database UniProt [57]. There are ongoing attempts by the W3C's Social Web Working Group to define standards to encourage interoperability between SNSs, but it remains to be seen whether such efforts will also benefit SNS researchers.[2] The development of appropriate standards enables the encoding and sharing of workflows, where there are domain-specific solutions. VisTrails is a system for encoding workflows, to allow the reconstruction of visualisations and plots from the original data [32]. Sumatra couples the execution of experiments with revisions in a version control system to allow the environment of a particular execution to be recreated

[1]FigShare: http://figshare.com/
[2]W3C Social Web Working Group Charter: http://www.w3.org/2013/socialweb/social-wg-charter

later [21]. There are domain-specific tools to support the execution of workflows, such as in biology, which has an embedded culture of sharing protocols such that other researchers can re-use and adapt these protocols in their own experiments. One such initiative is Taverna, which allows workflows that rely on several Web services and tools to be combined into a pipeline [143]. These workflows can be shared with others using repositories such as myExperiment, which adopts an SNS-like approach to sharing [38]. Such efforts have been found to support the building of social capital within the research community, with workflow-sharing perceived as a reputation-building exercise. When adopting the system, though, researchers found limitations with reproducing workflows "off-the-shelf", due to poor annotation and documentation of many workflows [107].

We are not aware of any solutions that aim to deal with the specific challenges of SNS research we identified in Chapter 2.3.5. The approaches we have discussed so far do not consider the sustainability challenge of reproducing experiments that may depend on APIs that have changed, or SNSs that no longer operate. These approaches are not optimised for human subjects research, with no consideration of how to encode the participant sampling strategy, the briefing and consent materials they were presented with, or the ethical and privacy-preserving handling of sensitive SNS data.

Noting the increased attention reproducibility has attracted across a range of disciplines, but considering the lack of tools to support the reproducibility of SNS research, we are interested in the extent to which the issue has been addressed in the SNS literature. To do so, we survey research papers that have recently used data from SNSs such as Facebook or Twitter.

Our survey of the SNS literature is not the first. Mullarkey developed a typology of SNS papers based on a sample of papers, to illustrate biases in the nature of SNS research [94]. Wilson et al. [141] look at 412 papers that use Facebook, to examine how Facebook has been used by social scientists. Caers et al. [13] find 3,068 papers in a broader search to identify the themes of such research,

but neither focuses on the reproducibility of SNS research. Golders and Macy conduct a wide-ranging survey of SNS research in sociology [40], and outline privacy as a research challenge, but not ethics, and discuss methodology but in the context of training sociologists in methods for collecting SNS data. Alim surveys SNS researchers about ethics concerns, and finds that 25% of respondents sought ethics approval for their studies [2], however there might be an element of selection bias, since researchers more interested in ethics might have responded to this particular survey.

To determine the state of reproducibility in the field, we examine 901 papers from 26 venues, published between 2011 and 2013. A range of venues were included to gain a diverse range of perspectives, including top-tier HCI conferences, network science workshops, and social science journals. We first collected all papers that satisfied the search terms shown in Table 3.1. For each paper, we then assessed whether the paper involved the handling of SNS data. If a paper's methodology concerned the collection or publication of data intended for an SNS, whether already established (such as Facebook or Twitter), or developed as a testbed for academic study, it was included. This was the case whether the authors directly processed the data themselves, or a previously crawled dataset was utilised.

Of the 901 papers examined, 505 met this criteria and were then tested against ten criteria we devised for assessing reproducibility, which we outline in Chapter 3.3.1.

To better understand trends across the literature, we categorised venues in one of two ways. Journals and magazines were grouped by field, using the publication's top category as listed by Thomson Reuters[3], while conferences were grouped by the best-fitting top-level category in the ACM Computing Classification System.[4] A summary of the venues, their classifications, and the number of papers examined is shown in Table 3.2. Finally, for each paper included in

[3]Thomson Reuters' Journal Citation Reports: http://thomsonreuters.com/journal-citation-reports

[4]ACM CCS: http://dl.acm.org/ccs.cfm

Field	Keywords
Abstract contains any of	Facebook
	Twitter
	Foursquare
	LinkedIn
	Friendster
	Weibo
	Flickr
	LiveJournal
	MySpace
	“Online social network”
	“Social network site”
	“Social networking site”
	SNS
	OSN
Publication date between	01-01-2011
	31-12-2013

Table 3.1: The semantics of the search term used to identify papers in the study (the exact syntax for expressing the search varied from source to source).

the survey, we conducted a citation analysis by querying Google Scholar to receive a citation count for each paper on July 8th 2014. While Google Scholar may not provide an exhaustive count of all citations, it allows us to study the relative performance of the papers we examine, in a similar fashion to other studies [105].

3.3.1 Explanation of criteria

Each of the 505 papers was tested against the following set of criteria. These align with the three aspects of reproducibility outlined earlier. For each criterion, a paper is assigned a binary flag to indicate satisfaction. This was determined by a manual reading of the papers, and not the result of an automated content analysis process.

Venue type	Venue	Total	Relevant
Computer science	IEEE Transactions on Mobile Computing	2	1
	Computing Networks	7	4
	Communications of the ACM	46	2
	Computer Communications	9	5
	IEEE Pervasive Computing	2	1
Security & Privacy	NDSS	2	1
	SOUPS	9	2
	S&P	3	1
	CCS	13	1
	WPES	4	2
Information systems	COSN	14	12
	EuroSys SNS	13	7
	WOSN	9	7
	WebSci	40	33
	ICWSM	200	177
	ASONAM	155	120
	HotSocial	9	6
Human-centered computing	CHI	82	39
	CSCW	73	45
	Pervasive	9	0
	UbiComp	23	9
Multidisciplinary	Nature	10	3
	Proceedings of the National Academy of Sciences	7	4
	Science	7	2
Communication	Journal of Computer-Mediated Communication	18	3
Psychology	Computers in Human Behaviour	129	13
Anthropology	Social Networks	6	5
Total		**901**	**505**

Table 3.2: Breakdown of the surveyed papers by venue. The "total" column indicates how many papers matched our search term, while the "relevant" column indicates how many used OSN data, meriting further study.

Methods

1. **Source SNS**: User behaviour is not identical across social network sites, so replications are dependent on knowing where data were collected, either to collect data from a similar population, or to show differences between SNSs. Thus we note whether the paper explicitly identify the SNS(s) from which data were collected or published to. If the authors note that data were collected from an SNS aggregation service such as FriendFeed[5] without clarifying which underlying SNSs were accessed, this criterion is not met.

2. **Sampling strategy**: Just as the choice of underlying SNS may indicate biases in the resulting data, the way participants in the research were chosen is an important consideration. When conducting user studies, it is important to know whether the authors were investigating a certain population, or whether they intend their findings to be generally applicable to a wider population, as this has implications for how participants are recruited for replications. Similarly, large-scale crawling exercises may be biased if, for example, user IDs are collected in increments from an arbitrary starting point. To satisfy this criterion, the paper must explain how participants were recruited, either explaining the sampling technique, or offering a breakdown of the participants' demographics. If the study used an existing dataset, the authors must explain how the underlying data were collected.

3. **Length of study**: As discussed in Chapter 2.1, SNSs exhibit a number of temporal effects, with the design of systems continuously evolving, and people's behaviours and attitude towards their data evolving. Accordingly, in order to replicate SNS studies, it is important to know the length of time over which data were collected, as this can affect user behaviour, and ideally at what time data were collected. To satisfy this criterion, the period of data collection must be identified.

4. **Number of participants**: As the number of participants will affect the number of results, and the effect size of analyses, it is important to disclose

[5]FriendFeed: http://friendfeed.com

how many were collected. To satisfy this criterion, the number of participants, or users whose data were crawled, must be identified. In user studies, if participants were in one of many experimental conditions, the distribution of participants among these conditions must be disclosed.

5. **Data processing**: Understanding how data are handled throughout an experiment is an important detail, from both reproducibility and ethical perspectives. Knowing precisely which attributes of sensitive SNS data were collected is important to both replicate the study, and ensure data collection is proportionate to requirements, especially as SNS APIs make it trivial to collect significant amounts of information. In addition, knowledge of how data were sanitised is important, particularly when releasing data that relates to sensitive SNS content. For example, have identifying characteristics been anonymised or aggregated, and how? To satisfy this criterion, the paper must have answered at least one of the following questions: Is the data handling strategy identified? Are the attributes of collected data enumerated? Were the data sanitised? How were they stored? Who had access to the data?

6. **Consent**: As discussed in Chapter 2.3.3, the issue of obtaining informed consent when conducting online research is contentious. Depending on its nature, SNS research may constitute human subjects research, in which case data-handling practices should be subject to the participants' informed consent. Understanding whether consent was sought is important for replications, as the process may have implications on the results. To satisfy this criterion, the authors must note whether the human subjects of the data collection provided consent to participate. The authors do not need to have sought consent to satisfy this criterion, but the issue must have been considered in the text.

7. **Participant briefing**: As with the acquisition of consent, the briefing and debriefing experience is an important ethical consideration when conducting human subjects research. These procedures ought to be explained in the text such that other studies can replicate the procedures for the most

consistent participant experience. To satisfy this criterion, the paper must disclose whether participants were briefed and debriefed to bookend their participation in the study.

8. **IRB/Ethics**: Alongside disclosure of consent and briefing procedures, studies should disclose whether the procedures of an experiment were approved by an Institutional Review Board (IRB), ethics committee, or equivalent. The need for such approval is dependent on what certain institutions or jurisdictions deem to be human subjects research, but disclosure can support replications, as IRB oversight may affect the ultimate data collection protocol of an experiment. To satisfy this criterion, the authors must note whether such bodies have approved the practices of the study.

Data

9. **Data shared**: The studies we examine may concern first-hand collection of data, perhaps by crawling an SNS, or conducting a user study to examine behaviour in an SNS. Alternatively, studies may use existing datasets, either provided through arrangement through a third-party, or by using a public dataset. Data sharing is acknowledged as an important aspect of reproducibility, but for all SNS research it is not essential, particularly where the data collection practices are sufficiently explained to allow other researchers to collect their own data. Nonetheless, we consider for each paper whether the data are shared with the research community, or if the authors explicitly entertain requests for access to the data. Where an existing dataset is used, the authors must explicitly cite it.

Code

10. **Protocol**: Another pillar of reproducibility concerns access to software artefacts necessary for collecting data, conducting analysis, and generating outputs such as plots. If a study concerns a bespoke visualisation, or the

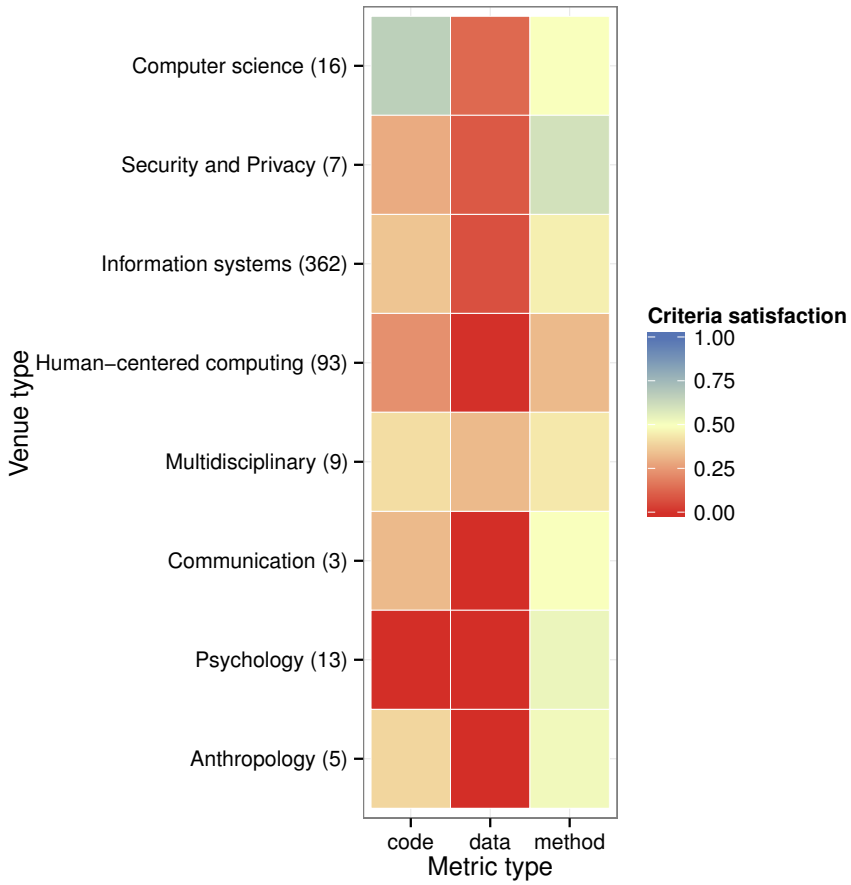


Figure 3.1: Heatmap showing how different fields achieve our three criteria types. Data-sharing is particularly poor across most disciplines, while reporting of methodologies is generally stronger.

development of a new SNS or alternative SNS interface, these should be accessible openly, and ideally the source should be available for others to use. To satisfy this criterion we check whether authors who develop their own software make this available to other researchers, and whether statistical analyses are explained in such a way that they can be replicated.

Our survey highlights differences in how well papers in different venues achieve reproducibility. Figure 3.1 shows a high-level summary of how different fields satisfy the three criteria types we introduced in Chapter 3.3.1.

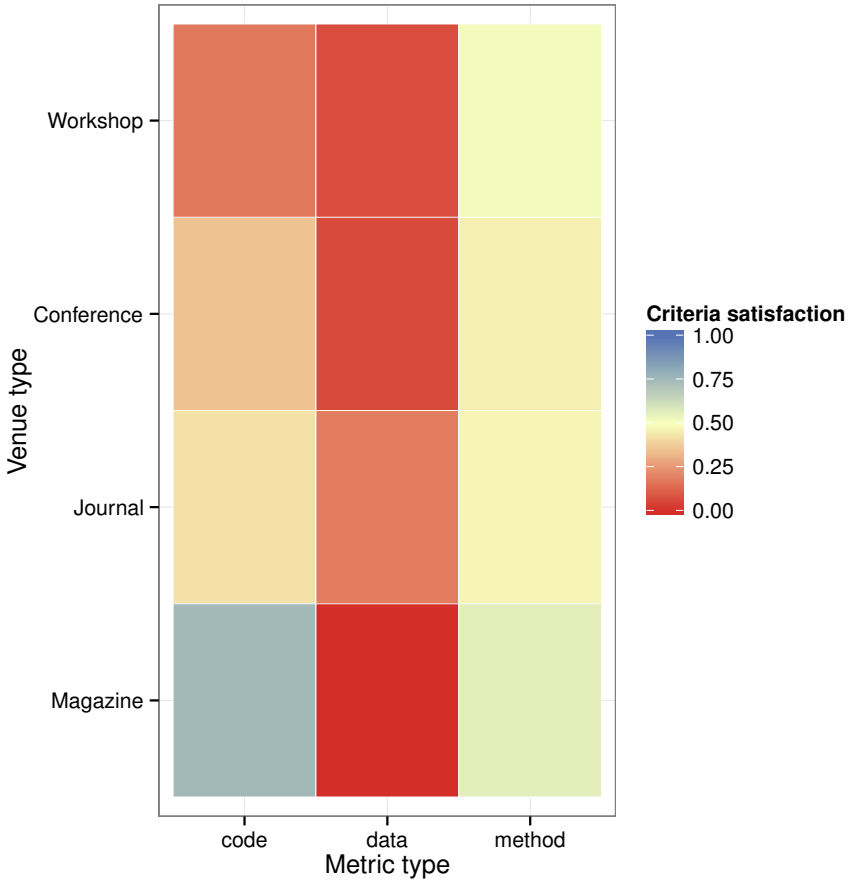


Figure 3.2: Heatmap showing how well each type of venue achieve our three criteria types. Data-sharing and methodology reporting are similar, however conferences and magazines are better at sharing code.

3.3.2 Few SNS researchers share their data

The most striking finding is that few papers share their data at all, with only 6.1% of papers in our survey doing so. Unsurprisingly, this is closely associated with the data-sharing policies of different venues. Multidisciplinary journals such as *Nature* and *Science* mandate authors to include data such that reviewers and other researchers can replicate results,[6] [7]and accordingly are a notable exception to this trend, with 40% of papers sharing their data. We are not aware of any conferences in our survey that mandate data necessary for replication must be shared, although conferences such as *SOUPS* do allow authors to include appendices that support replication.[8] Similarly, the *ICWSM* conference operates a data sharing initiative to encourage the sharing of datasets,[9] that may explain why 35.4% of the papers that shared data came from this venue. We note that papers at some information systems venues, such as *EuroSys SNS* and *COSN*, are moderately better at their data sharing practices, with authors at both sharing data twice as often as the venue average. This appears to be a side-effect of many papers using crawled social graphs, rather than datasets of content, such as tweets, that are licensed under terms which prohibit redistribution. As shown in Figure 3.2, papers in venues of all types are quite poor at routinely sharing their data. Journals fare better with 13.9% of papers sharing their data; however a chi-square test of independence does not suggest this is a significantly greater effect than other venue types ($x^2 = 4.38$, df = 2, $p = 0.11$).

3.3.3 Social scientists rarely share code for experiments and analyses

We find that code-sharing practices are generally better, which includes the distribution of theorems or algorithms that support replication, but notably no

[6]Nature data policy: http://www.nature.com/authors/policies/availability.html
[7]Science data policy: http://www.sciencemag.org/site/feature/contribinfo/prep/gen_info.xhtml
[8]SOUPS 2014 Call for Papers: http://cups.cs.cmu.edu/soups/2014/cfp.html
[9]ICWSM Data Sharing Initiative: http://icwsm.org/2015/datasets/datasets

venue types, except for multidisciplinary journals, include a majority of papers who satisfy this.

In this analysis, *Computers in Human Behavior* (CHB) was notable in that none of the papers we examined shared code. *CHB*'s simultaneous computational and social science focus attracts authors from diverse disciplines and may go some way to explaining this. Of the 13 papers that we examined, first authors are affiliated with computer science, communications, political science, management, humanities, psychology, and law faculties. For many such fields, there may be no expectation that quantitative methods are shared to allow replication. As multidisciplinary efforts like this gain traction, it is important that the strengths of social sciences – such as experience with qualitative methods – feed into computer science, just as traditional CS strengths – such as an emphasis on sharing code – are accepted by the wider computational social sciences community.

Code-sharing rates increase dramatically between publication types. As shown in Figure 3.2, protocols are shared in approximately a quarter of workshop and journal papers, while 41.4% of conference papers satisfy this. In Paek and Hsu's work to create phrase sets for text entry experiments from large corpora, the researchers made the phrase sets and code available, and included detailed algorithmic details within the paper [103]. As noted earlier, we attribute this trend towards sharing to more stringent requirements for supplementary materials in such publications. As workshops are often used for work in progress, it may be that researchers are reticent to share unfinished code. We would hope to see this change, however, to help engage the community in the development and re-use of software even in an unfinished state.

3.3.4 Reporting of core experimental parameters is strong

Reporting of the methodological attributes appears strong across all papers; however, the breakdown of these criteria in Figure 3.3 shows a more complex

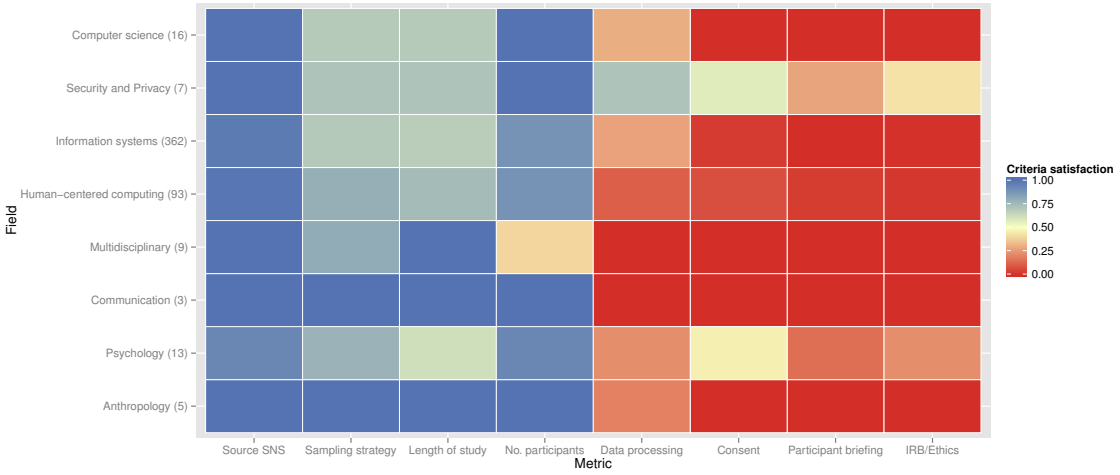


Figure 3.3: Breakdown of the eight criteria we assess for "methods". Generally, papers successfully report descriptive attributes of their study, but often, participant handling and data processing are not sufficiently explained.

dichotomous story. The first four criteria illustrate the extent to which studies report the core aspects of their data collection practices, critical to reproduce any such studies, including the source SNSs, how participants and their data were sampled, for how long data were collected, and the number of participants. Generally, papers are very good at reporting this information, with some notable exceptions. Just as studies that used existing datasets are inherently better at sharing the data they use, they tend to be worse at reporting the provenance of their datasets, such as the composition of the dataset's participant pool. These are crucial details that are required to replicate such studies, particularly if the original dataset is not to be used – such as aiming to replicate the findings of a user study with a different population.

3.3.5 Participant-handling and ethical considerations are not discussed

The final four criteria in our methodology breakdown concern data processing and participant ethics, two critical aspects of reproducibility, where consistently most papers do not report core methodological concerns: did participants give consent? Were procedures approved by an IRB? How were the collected data

handled? Again we see a divide in approaches between systems papers, and social sciences. Quantitative work, for example, is better at reporting how their data were handled, such as anonymisation practices, and which attributes of datasets were stored. As Figure 3.3 shows, the seven "Security and Privacy" papers we consider are better at reporting these concerns. We attribute this to a culture of reporting these details at SOUPS, while WPES allows appendices with supplementary information to be provided. Conversely, the social science background of many CHB papers is highlighted in the marginal improvement in reporting of ethical concerns, shown in the *Psychology* group. We were surprised to find that human-computer interaction (HCI) papers were not particularly strong in this regard. Indeed, such reporting is so uncommon that attention should be drawn to positive cases, such as Johnson et al.'s description of their recruitment material and consent procedures [58], and Ali et al.'s reporting of their study's participant briefing process [1]. Simply reporting the existence of briefing and consent procedures generally does little to support replication. Our concern with the lack of robust description of such methods, is that as previous work shows the briefing experience can affect people's disclosure behaviours in SNS experiments [88], it is important that researchers can replicate these procedures when conducting user studies using SNS data.

In this section, we have looked at how well the state of the art addresses reproducibility in SNS research. We find that venues from more technical backgrounds differ in their reporting from the social sciences. While some fields have developed domain-specific tools to capture workflows and aid with reproducibility, there are none that support the SNS-specific issues we have identified.

3.4 Studying SNSs with contextual integrity

In Chapter 2.3.6, we introduced Nissenbaum's model of contextual integrity. In the rest of this thesis, we apply this model to the design and study of SNSs, but

we first consider how others have applied the framework to the study of online privacy.

Barkhuus advocates contextual integrity as a method for studying privacy in HCI studies, arguing it provides a more useful lens for understanding people's privacy preferences and for detecting violations, compared to some commonly-used methods, such as the Westin-Harris privacy scales [77], which try to capture a person's overall level of privacy concern. Borcea-Pfitzmann et al. argue contextual integrity should be incorporated into all assessment about the potential impacts of systems, and that designers should seek to maximise the contextual integrity of user data [9].

Grodzinsky and Tavani applied contextual integrity to cloud computing, specifically adopting the decision heuristic, discussed in Chapter 2.3.6, to determine whether the cloud-based document editing service Google Docs might violate people's privacy [42]. Jones and Janes also apply the decision heuristic to Google Books, highlighting its usefulness for determining whether a new technology might perturb privacy norms where this is an established context, in this case physical libraries [59], finding that the loss of reader anonymity, and potential for surveillance, has privacy risks.

Lipford et al. note that SNSs such as Facebook may lead to violations of contextual integrity [75], because of the risk of context collapse, as discussed in Chapter 2.1. Hull et al. make a similar argument, noting that violations arise from the News Feed product and application programming interface (API), as they often involve the repurposing of user data beyond the context in which it was initially published [51]. Similarly, Shi et al. argue that the introduction of "friendship pages" to Facebook, which aggregates content shared between two peers, violates contextual integrity because the original authors of the content would not expect it to be aggregated in this way [113].

While contextual integrity has been widely used to study a range of online services, including SNSs, we have not observed its applicability to determining whether those conducting research on such services respect the contextual

integrity of their participants. Furthermore, the usefulness of contextual integrity to investigate specific disruptions to existing SNSs, such as the addition of incentives to LBSNs, has not yet been established.

3.5 Summary

In this chapter, we have surveyed the recent literature proposing solutions to privacy challenges in SNSs, and identified issues with the way researchers using SNS data report on the ethical conduct of their studies. We note the following points:

- There are many proposed solutions to some of the usability and trust issues with existing SNSs, but there remain open problems. When commercial features are added to existing services, do people understand how the existing context has been perturbed, and are they provided usable controls for deciding how their information is used?
- Acquiring dynamic consent may engage participants in the research process and reduce the risk of violating their expectations, but it is not yet known whether this can be applied to SNS research, and without placing an unacceptable burden on participants.
- While SNS research is very popular, researchers often do not report on the ethical conduct of their studies, which can cause difficulties when reproducing experiments, or establishing the ethical legitimacy of a study.
- There are many domain-specific workflow tools for managing research data, but so far none that consider the unique challenges of conducting SNS research.
- Contextual integrity is a framework that has been used to study some SNSs, but its applicability to examining the appropriateness of research methodologies, and the role of commercial actors in LBSNs, has not been established.

In the rest of this thesis, we apply contextual integrity to the study of privacy, consent, and reproducibility, which we have discussed in this chapter.

In the next chapter, we introduce a framework that has been designed to improve on the state of the art in reproducibility and ethics in SNS research, which we will evaluate in a number of case studies. Informed by contextual integrity principles, it enables the study of such systems in a manner that ought to protect the expectations of participants.

Chapter 4

A framework for ethical SNS research

In Chapter 2.3 we noted a range of challenges with conducting research using social network sites (SNSs), specifically with respect to gaining informed consent, protecting the privacy of participants, and conducting research in a reproducible manner. Our survey of the SNS literature in Chapter 3.3 identified limited progress towards enabling reproducibility through the reporting of key methodological details, and sharing of artefacts such as source code and data.

The challenges associated with conducting and reproducing SNS research raise many interesting research issues and dilemmas, particularly involving the sensitivity of the personal data that are communicated via SNSs. How can we collect and analyse SNS data while respecting the privacy of SNS users? Can we assume that data shared on an SNS are fair game for all researchers? Are results from one SNS or set of SNS users representative of all users or all SNSs? How can we share data with other researchers while maintaining the privacy of participants?

In this chapter, we make the following contributions:

- We propose an architecture for executing privacy-preserving SNS studies that encode the workflow needed to reproduce the experiment.
- We evaluate the architecture by recreating an existing SNS study from the

literature, to demonstrate how we encode the parameters of an experiment, and illustrate how it preserves the privacy of participants.

To meet the challenges identified in Chapter 2.3, we suggest that an appropriate solution is to develop a lightweight framework that acts as middleware between a researcher who wishes to make use of SNS data, and the providers of such data. We have distilled these challenges into the following requirements:

1. In our survey in Chapter 3.3, we found that SNS research was conducted using a wide variety of methodologies, including questionnaires, ethnographic studies, user studies, and crawling the data provided by SNSs. The architecture must enable all such studies, and we should be able to reuse system components for different methodological approaches.

2. In Chapter 3.3 we identified a number of tools that enable the encoding and sharing of experiment workflows. We did not find these sufficient for encoding the specific details of SNS studies. Therefore, the architecture must allow researchers to capture workflows, in terms of data acquisition, and privacy and ethical requirements, so that experiments can be repeated and experimental designs shared with other researchers.

3. Our survey noted that studies make use of a range of services, from currently popular services such as Facebook and Twitter, to now-obsolete SNSs such as Friendster and MySpace. The architecture must allow experiments to be run using data from a wide range of SNSs and other sources of social data, to allow experiments to be conducted against a range of data sources, without knowing details of their implementation. It should be sufficiently extensible such that support for new services can be added without any changes to the core framework.

4. As identified in Chapter 2.3.5, and illustrated in our survey, a barrier to reproducibility is that the SNSs may change their APIs at any point, which can break existing code, or indeed the whole service may cease to exist. Our framework should mitigate the impact of these changes, by handling

mismatches between new APIs and old code, and allowing the environment of a defunct SNS to be emulated where it is no longer available, allowing privacy-preserving requests to be made of legacy SNS datasets.

5. In Chapter 2.3.4, we noted how some previous SNS studies may have failed to uphold their participants' expectations of privacy. The theory of contextual integrity, which we introduce in Chapter 2.3.6, provides a framework for designing systems to meet such expectations. The architecture must maintain the contextual integrity of participants, by respecting the context-appropriateness of information transfers.

4.1 The PRISONER architecture

We now present our architecture for privacy-sensitive social network experiments, titled PRISONER (Privacy-Respecting Infrastructure for Social Online Network Experimental Research).

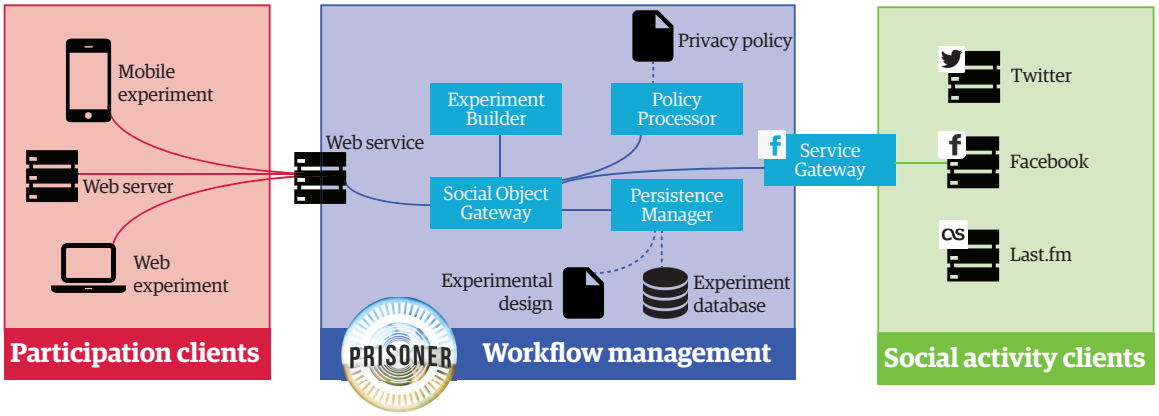


Figure 4.1: The PRISONER architecture, showing the flow of data between its three constitutent components.

The architecture consists of three components, as shown in Figure 4.1. We collect experimental data from a participant's activity on an SNS (social activity clients), and through participation in experimental interventions (participation clients). Both the social activity and participation clients feed and receive data from the workflow management component. The workflow manager is re-

sponsible for ensuring that all data generated and analysed in the system reflect privacy and ethical requirements; this includes data collected from third-party SNSs, data generated for experimental purposes, and data that are fed back to participants in interventions.

4.1.1 Social objects

At the core of PRISONER is an abstraction of SNSs, as a set of common objects, with service-specific extensions. Informed by the Activity Streams 2.0 standard [134], we implement a number of abstract types, which correspond to common entities found in various SNSs, as we discussed in Chapter 2.1. We term these *social objects*, which include:

- **Notes:** These encapsulate short text messages, which are analogous to status updates in Facebook, or a tweet in Twitter.
- **People:** In the most abstract sense, these can represent any other human, however each service can provide its own semantics about their significance within the context of that service. For example, a *person* might be tagged in a photo in which they appear, or "own" a collection of *notes*, such as a user's timeline of tweets.
- **Place**: This represents the attributes of a location. This may be used to plot a location on a map, or to provide location-specific context to a *note*. Such abstractions are useful as LBSNs may implement "check-ins"in a variety of ways, but can share a canonical representation of a location.

PRISONER contains a core implementation of eight social objects, which describes a common set of fields. References to these abstract social objects can be made by an experiment, with the SNS-specific implementation resolved at runtime.

Listing 4.1 A short privacy policy for a PRISONER experiment allowing a participant's username and sanitised user ID to be collected

```
<policy for="Facebook:User">
      <attributes>
         <attribute type="id">
            <attribute-policy allow="retrieve">
               <transformations>
                  <transform type="hash" level="sha224" />
               </transformations>
            </attribute-policy>
         </attribute>
         <attribute type="username">
            <attribute-policy allow="retrieve" />
         </attribute>
      </attributes>

      <object-policy allow="retrieve">
         <object-criteria>
               <attribute-match match="author.id"
               on_object="session:Facebook.id" />
         </object-criteria>
      </object-policy>
</policy>
```

4.1.2 Privacy policies

PRISONER uses privacy policies, written by the researcher, to determine what types of SNS data are appropriate for a study to handle, and what sanitisations should be applied to protect the privacy of participants. Privacy policies use an expressive XML schema to allow researchers to declare criteria for allowing objects to be collected, depending on how permissive their application is, and how reusable their policies need to be. Without a policy, an experiment is not allowed to process any data. As shown in Listing 4.1, a policy consists of *policy* elements, each corresponding to a social object. For maximum re-use, this can refer to a generic social object, or with the inclusion of a namespace, an object for an individual service, as in this example. Each object policy enumerates the attributes of that object that need to be collected, with a policy for each attribute declaring whether that attribute can only be retrieved or can also be stored, and whether it needs to be sanitised before being made available to the experiment. Sanitisations are requested by "transformations" elements, with a

range of semantically-appropriate transformations available. In this example, the user ID is hashed to avoid the researcher having direct access to it, while a location attribute in a check-in object might be coarsened from an exact coordinate to a city-level aggregation. Finally, the object policy contains object criteria, which dictates which objects the researcher is limited to retrieving. In this case, the criteria notes that only objects authored by the current participant can be collected, meaning that only the participant's own profile can be retrieved. This prevents the researcher from accidentally collecting any data beyond this scope.

Requiring researchers to explicitly specify the data they require in this manner serves several purposes:

- Researchers are encouraged to think about the scale of data collection required by their experiment, requiring consideration of the contribution each attribute of each object makes to answering their research questions.
- Policies can be written offline before experimental applications have been developed, encouraging the data-collection practices of an experiment to be iterated on independent of the application itself. This acts as a testable specification of the requirements of a study. If at the time of implementation, further changes are required, this invites further scrutiny of why the original policy was insufficient, and to iterate on creating a more appropriate policy, rather than simply collecting all possible data.
- Policies can be rendered in a human-readable format that can be used as supporting documentation for IRB review, or as part of informed consent forms to give to participants.
- While researchers can still choose to simply collect all attributes of all objects, the use of policy-driven consent forms should encourage better practices, as potential participants may be dissuaded from taking part in a study that demands significant amounts of data without justification.
- Policies can be shared with other researchers, as a way of distributing

workflows, even if the source code for an experimental application is not shared. Researchers who are reproducing or building on previous work can therefore recreate the same data-collection practices as the original study. Even if other researchers do not use PRISONER, the policy provides a human-readable description of the experiment's workflow.

We now explain how each of these components were designed, and how their implementation in PRISONER fulfils the requirements discussed earlier.

4.1.3 Participation clients

These are the experimental applications that interface with the PRISONER system. These can be of any form, including questionnaires, mobile applications that use the experience sampling method (ESM) [47], or alternative interfaces to existing social network sites. They interface with PRISONER using its API, which provides a consistent way of handling all the services that PRISONER supports. As PRISONER is able to resolve references to generic social objects to their service-specific implementations, this allows the same participation client to be re-used in experiments that target different services without substantial rewriting required.

4.1.4 Social activity clients

Social activity clients are systems that provide social data that a researcher might wish to use or generate. For the purposes of this thesis, they are likely to be SNSs such as Facebook, however ultimately any system or dataset of a social nature could be implemented. As Social activity clients take many forms, PRISONER mediates between their implementations with an API layer called service gateways.

Listing 4.2 A fragment of the Facebook service gateway, showing how "liked" pages are retrieved.

```
def Like(self, operation, payload):
   # handle requests to retrieve Likes
   if (operation == "GET"):
      try:
         # requests are keyed on the ID of the user who liked pages
         user_id = payload
         # populate a Person social object with the identity of the user
         author = SocialObjects.Person()
         author.id = user_id

         # make a request to the FB Graph API for the user's likes
         result_set = self.get_graph_data("/" + user_id + "/likes")
         like_obj_list = []

         # paginate through the collection of likes and add information
         # about each page to a list
         while ((result_set.has_key("paging")) and (result_set["paging"].has_key
         ("next"))):
            like_obj_list.extend(result_set["data"])
            result_set = self.get_graph_data(result_set["paging"]
            ["next"])

         likes = []

         # for each of the raw likes we retrieved, populate a
         # Page social object with the appropriate attributes
         for like in like_obj_list:
            this_like = Page()
            this_like.displayName = self.get_value(like, "name")
            this_like.id = self.get_value(like, "id")
            this_like.url = "https://www.facebook.com/" + this_like.id
            this_like.author = author
            this_like.category = self.get_value(like,"category")
            this_like.image = self.graph_uri + "/" + this_like.id +
            "/picture?type=large" + "&access_token=" + self.access_token
            likes.append(this_like)

         # any collection of objects is associated with the SNS it came from
         # to avoid any interop issues

         likes_coll = SocialObjects.Collection()
         likes_coll.author = author
         likes_coll.provider = "Facebook"
         likes_coll.objects = likes

         return likes_coll

      except:
         # if there are no likes to get, return an empty set
         return SocialObjects.Collection()
```

Service gateways

The role of a service gateway is to translate between the native responses provided by the API of a social activity client, and the social objects that PRISONER can parse. As with the rest of the architecture, service gateways are implemented in Python as modules conforming to a specification, in which methods are written for each type of social object the service supports. The architecture was written in Python as the language is well-supported, and its dynamic introspection capabilities were useful for resolving references to aspects of the system at runtime. The social objects gateways handle can include the generic social objects discussed earlier, or service-specific ones. Each method handles requests to retrieve or publish one or more social objects of that type.[1] GET requests return a fully-formed social object of the type named by the method, matching criteria included in the request, such as a user ID, as shown in Listing 4.2. POST requests create objects of the named type, given a payload provided in the request, and publish these to the relevant SNS. PRISONER includes a number of gateways for common SNSs, such as Facebook, Twitter, and Last.fm. If researchers need to add support for an additional service, they must first write a service gateway module to wrap the API of the required service.

4.1.5 Workflow management

This component represents PRISONER's core, and is responsible for delegating requests from participation clients to service gateways, validating policies for experiments, sanitising social objects, and exposing interfaces for researchers such as a Web service. It consists of the following modules:

[1] Adopting the parlance of the HTTP specification, these are referred to as *GET* and *POST* requests respectively.

Experiment Builder

When a participant begins using a PRISONER experiment, the system must be supplied with some metadata describing the experiment, including:

- The privacy policy, which dictates which data the experiment can collect.
- The experimental design, which describes the structure of any data that should be stored, such as the responses to a questionnaire.
- The name of the study, and contact details for the researcher.
- A list of services the participant needs to be authenticated with.
- A callback URL to redirect the participant to after authentication.

The experiment builder generates a front-end to the experiment using this information. All future requests can then be validated against the correct policies.

Policy Processor

This module is responsible for validating and sanitising all requests to retrieve and publish social objects. Using the privacy policy provided by the experiment builder, when an experimental application requests data from an SNS, it ensures an experiment is allowed to process the type of the object requested, and applies any sanitisation strategies that have been requested.

Persistence Manager

This module handles the storage of data retrieved or generated during the course of an experiment, including sanitised instances of social objects. The persistence manager infers and maintains the database schema based on the experiment's policy, and abstracts the researcher from maintaining the database itself, by providing APIs to retrieve data in line with the experiment's policy

without over-exposing sensitive information. It maintains the logical relationships between objects, which might be based on sensitive keys such as a user's ID, allowing data to be returned to the researcher without such details being known or exposed to them.

Social Object Gateway

This module is used to coordinate the activity of other parts of the system. It provides a single interface for initiating requests, dispatching requests and converting raw objects into a form that can be used by experiments. Experiments can access this interface directly, but are more likely to use the Web service.

Web service

The Web service is the primary interface to PRISONER, providing a Web-based wrapper around the Social Object Gateway. An experiment begins by providing PRISONER with the policies for the experiment, which internally uses the experiment builder to configure the parameters of the study. The participant visits a Web address provided by PRISONER to register for the experiment. This assigns a unique identifier for the participant, and uses the authorisation mechanics for the underlying service to allow the experiment to access the participant's account with that service. PRISONER maintains the credentials needed to access the underlying services in a session object, which are not directly exposed to individual experiments. From then on, the experiment makes requests for data to PRISONER, using the participant's session cookie for authentication. If, for example, an experiment needs to retrieve the photos from a participant's Facebook account, it can make a HTTP GET request to the following URL:

```
https://prisoner.cs.st-andrews.ac.uk/prisoner/get/Facebook/Photo/session
:Facebook.id
```

PRISONER recognises the "session" namespace at the end of the URL, and

substitutes this with the ID of the participant's Facebook account. Internally, the request is processed using the components discussed earlier, before returning a JSON array of sanitised Photo objects to the experiment.

4.1.6 Designing for contextual integrity

Our fifth design requirement is that our architecture maintains the contextual integrity of participants. As we discussed in Chapter 2.3.6, privacy violations can result when someone's expectations about the reasonable flow of information are not met. This can explain the sense of violation that were associated with the two studies we considered in Chapter 2.3.4. To avoid a repeat of such incidents, we consider how the design of PRISONER can mitigate this.

First, we consider transparency to be key to upholding contextual integrity. The specification of data-handling practices encoded in an experiment's privacy policy is useful, not just because it provides a way of encoding these details, but as the policy is enforced at runtime, there is no scope for the stated ethical practices of an experiment, whether communicated to IRBs or participants, to be inaccurate. Therefore, when a participant completes a consent form that has been generated from a PRISONER policy, they can trust that the consent they give is to an experiment as described. This is important, as even when researchers make good faith efforts to protect the privacy of participants, there is still the risk of violation, as we found in Chapter 2.3.4. To mitigate concerns about inappropriate flows of information, we require researchers to specify how information is managed throughout the lifecycle of the experiment. By communicating to participants what information is collected and stored, and how it is sanitised, they are given explicit information about the flow of information, which helps to uphold contextual integrity.

As we discussed in Chapter 2.3.3, the acquisition of consent in a secured fashion is divorced from the data collection practices of the study itself. Whether this consent is informed or meaningful can be difficult to determine, if the im-

plications of consenting to such data collection are difficult to appreciate. The approach we take with PRISONER aims to leverage people's understanding of the SNSs they already use to make their consent decisions more meaningful. By embedding these decisions within the context of the SNS itself, and not as an abstract process, we believe that this better maintains contextual integrity.

4.1.7 Dealing with API changes

The development of PRISONER has met with some challenges. Chiefly among these are dealing with changes made to the APIs provided by social network sites. As discussed in Chapter 2.3.5, these APIs are often volatile, with frequent changes to the interfaces that affect which endpoints developers can use, limitations on how the APIs are used, or changes to which parameters endpoints accept, or what they return. One of PRISONER's main goals is to improve the state of reproducibility in SNS research, which is made more difficult if researchers cannot be confident that code that targets a specific SNS will still be usable, even in the near future. While we cannot completely solve this issue, as the availability of an SNS and its API is out of our control, we have adopted a few strategies to mitigate its effects.

First, PRISONER includes a versioning scheme in service gateways. Each method can incorporate version-specific behaviours, either to target requests at an explicit version of the API provided by the SNS, or to translate requests targeted at an older API version, to the format expected by a new one. The intention of this approach is to be completely transparent to an individual experimental application, and require no adjustment on their part. An experiment's design policy can indicate which logical version of an API it is designed to target. At runtime, PRISONER will attempt to mediate between these changes by compensating for differences between the old and current API, which might be appropriate if an endpoint has changed name, or parameters are provided in a different, but semantically identical format. Only if an API call is no longer able to return a sensible result, for example if an endpoint has been completely

removed, does PRISONER raise an exception to indicate that the request can no longer be completed. This approach allows some "breaking" changes to be absorbed by the service gateway, without affecting the outward interface to applications. In addition, the service gateway can expose additional functionality to newer applications if an API endpoint has been enhanced, without breaking compatibility with older applications.

This approach was developed in response to a suite of API changes Facebook made between 2014 and 2015, in its transition from version 1.0 to 2.x of its Graph API. The Graph API is Facebook's primary interface for providing applications with access to the data and functionality exposed by the service. To determine the appropriateness of our strategy, we examined how many of the breaking changes in this update could be absorbed by changes to the service gateway, without any impact on old experimental applications. While changes in this version were far-reaching, affecting various platform-specific SDKs and login processes, we only consider the endpoints that related to social objects already implemented in our Facebook service gateway. We found that four endpoints were removed, four endpoints were modified, and twenty four API permissions were removed. Of the four removed endpoints, we were able to map requests for three to different endpoints with no loss in the accuracy of the data returned. One removed endpoint, `/me/notes`, coincided with the deprecation of a blogging application called "Notes", with no appropriate replacement. With the exception of the permissions for accessing notes, we were able to map all permissions to their replacements to maintain access to these data. Of the modified endpoints, we were able to mitigate their effects in all but one case. Applications are no longer able to access a full list of friends using the `/me/friends` endpoint, instead only accessing the friends who have installed the same app. As this significantly alters the semantics of this endpoint, PRISONER injects a confidence warning into the object returned by this endpoint to signal that is not representative. Older applications can ignore this and make use of the subset of data available, while newer applications can interpret this warning to make a determination of the quality of the data. Importantly, while this exercise does

not resolve all issues with API changes, in this example, we can demonstrate that PRISONER provides a significantly higher degree of reproducibility than applications that directly interface with the API, without any changes being made to individual experimental applications.

4.2 Reproducing an experiment with PRISONER

In Chapter 3.3, we identified issues with reproducibility in social network site research. We now explore the process of recreating an experiment with PRISONER, to illustrate how it improves on the state of reproducibility. To do this, we examine one paper identified in our survey [66], and reproduce its data collection procedures. We choose this paper as it is the only study in our analysis to meet all ten criteria, suggesting it should be possible to fully recreate its procedures.

Our chosen paper [66] studies attitudes towards information-sharing with third-party Facebook applications, by evaluating how well participants understand the data-handling practices of applications, and the differences between features operated by Facebook and applications provided by third-parties. The authors built a Facebook application to deliver a survey about privacy attitudes, which masqueraded as a personality quiz to encourage participation. Participants believed their responses would be used to classify them as one of a number of personality types. In reality, the application measured a participant's level of engagement with Facebook based on how many profile attributes they disclose (such as age, gender, and work history), and how many status updates they shared. This was used to provide a classification "for entertainment value" to the participant, while providing a quantitative measure of how much information they disclose on Facebook. To achieve this, the researchers collected significant amounts of information from a participant's profile using the Facebook API. The authors note that they:

"...collected data about each respondent's profile (but no actual profile data)

in order to compute measures of how much information people were sharing on Facebook. For most fields we computed a simple binary score (1 if the field contained data, 0 if blank) or a count if available (such as the total number of status updates and the number of status updates in the past 30 days)."

This suggests that at no stage were any sensitive data stored, but in order to compute these measures, requests for the data had to be made. In this instance, the authors make good faith efforts to protect the privacy of their participants, but in replications, such details are easily overlooked, and could easily lead to inappropriate quantities of information being stored. We note that the deception employed in the original study raises ethical questions about the appropriateness of misleading participants into the collection of data pertaining to them, however for the purposes of this analysis we are concerned only with the data collection exercise itself and not its purpose. The authors also justify the use of deception, and the procedures were approved by their IRB.

This study is ideal to model using PRISONER, as it relies on the collection of large quantities of data, while demonstrating a clear workflow that dictates how data should be sanitised and aggregated through the duration of the experiment. To recreate this workflow with PRISONER, we create a privacy policy that encodes the requirements we have discussed, in terms of which SNSs are accessed, which data types we require, and how they should be sanitised. We then write an exemplar Web-based application that supplies this policy to the PRISONER Web service, then makes requests to the PRISONER API whenever Facebook data are required.

Listing 4.3 A fragment of the privacy policy for the experiment we reproduce, showing how some of the attributes of the participant's profile are collected and immediately sanitised.

```
<policy for="Facebook:User">
   <attributes>
      <attribute type="id">
         <attribute-policy allow="retrieve">
            <transformations>
               <transform type="hash" level="sha224" />
            </transformations>
         </attribute-policy>
         <attribute-policy allow="store">
            <transformations>
               <transform type="hash" level="sha224" />
            </transformations>
         </attribute-policy>
      </attribute>
      <attribute type="gender">
         <attribute-policy allow="retrieve">
            <transformations>
               <transform type="reduce" level="bit" />
            </transformations>
         </attribute-policy>
         <attribute-policy allow="store">
            <transformations>
               <transform type="reduce" level="bit" />
            </transformations>
         </attribute-policy>
      </attribute>
   </attributes>
   <object-policy allow="retrieve">
      <object-criteria>
            <attribute-match match="author.id"
            on_object="session:Facebook.id" />
      </object-criteria>
   </object-policy>
   <object-policy allow="store">
      <object-criteria>
            <attribute-match match="author.id"
            on_object="session:Facebook.id" />
      </object-criteria>
   </object-policy>
</policy>
</p:privacy-policy>
```

To illustrate how this works in practice, we take each of the criteria in the *methods* category of our survey, and explain how we apply PRISONER to achieve that aspect of reproducibility. A fragment of the privacy policy that we created for this example is illustrated in Listing 4.3, and available in full online [48].

- **Source SNS** – To replicate this study, we need to collect data from Facebook.

PRISONER allows researchers to request generic *social objects* that exist on various SNSs. As this experiment only uses a single SNS, however, we make this explicit in the experiment's privacy policy. We create policies for each type of object our experiment needs to retrieve. Our policy for *Facebook:User* only allows us to retrieve profile information from that SNS as it is explicitly namespaced. This policy can be shared with others to ensure any further data collection comes from the same service, while a policy for a generic *Person* object could allow a replication to use data from any compatible SNS.

- **Length of study** – This study requires us to collect all of a participant's status updates in order to determine how many have been posted. PRISONER allows privacy policies to include temporal constraints on the data collected. This ensures that when the policy is reused, data from evolving sources, such as Facebook status updates, are only accessible from the same time period, or over the same duration. This study requires that a user's entire history of status updates is collected, so that the total number can be counted, so we did not provide an explicit time limit in this instance.

- **Data processing** – This study outlines some crucial data sanitisation requirements that must be preserved to both replicate the conditions of the study and preserve participant privacy. As described earlier, we do not need to collect the content of profile attributes or status updates, but rather a count of how many are accessible. When manually evoking the Facebook API to do this, it would be necessary to collect the sensitive data then manually sanitise it. While achieving the desired result, this is not ideal, due to the possibility that data may be inappropriately stored in an unsanitised form, especially when using third-party bindings that may implement their own clientside caching behaviours. This may risk participant privacy.

  By encoding these data-handling requirements in a declarative manner in the experiment's privacy policy, researchers do not need to be concerned

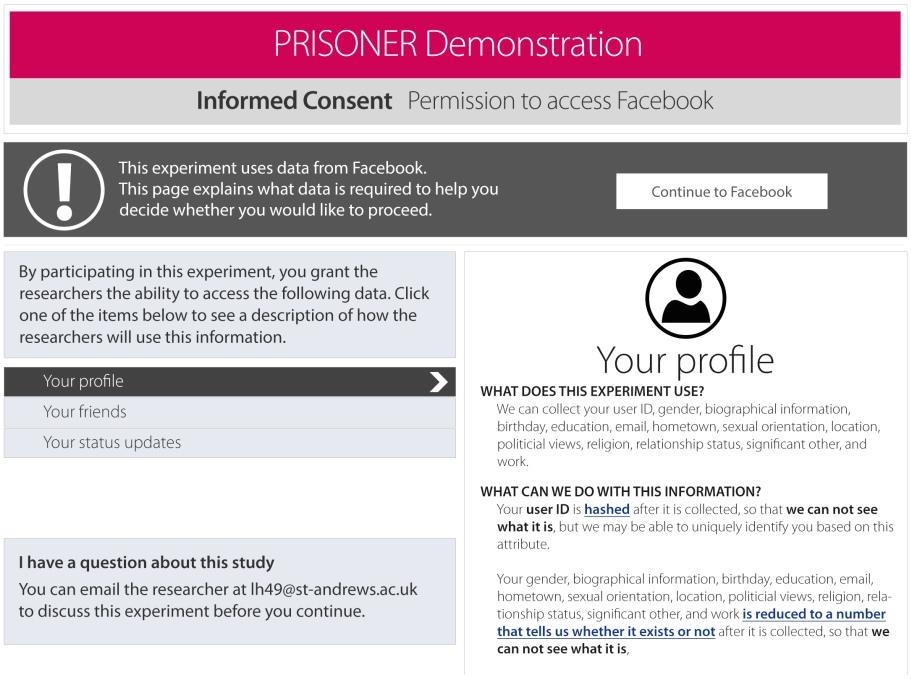


Figure 4.2: Part of the consent form that participants would be shown before taking part in the reproduced experiment. The form outlines the data collection practices of the study in a readable format. Participants who want to learn more about how their information is handled can read detailed descriptions of the sanitisations employed.

with such implementation details. To ensure we do not inadvertently collect too much information, we only request the *id* attributes from status updates, as shown in the *attributes* collection in Listing 4.1. On all other requests for sensitive attributes, such as work history or gender, we use *reduce* transformations whenever we retrieve data. The *bit* attribute immediately sanitises the response from the Facebook API to only return 1 if the attribute is present, or 0 if it is not, before the data are made available to the experimental application. As well as only collecting the number of profile attributes, the study requires that "respondents' Facebook user IDs were hashed for anonymization purposes". The transformation policy for the *User* object shows we hash the user ID using SHA-224 after retrieving it. Note that while this technique is commonly used to provide a degree of obfuscation, it is not impervious to attack. PRISONER does not provide any guarantees about the anonymity afforded by use of such techniques.

- **Consent** – The authors note that "a consent statement appeared on the first

page of the survey", but this is not sufficient to replicate the study, as language used to obtain consent can impact the results of SNS research [88]. As all attributes collected from SNSs are encoded in an experiment's policy, PRISONER can generate participant consent forms that explain which SNSs data are collected from, which attributes are collected, and how data are processed through the life of the experiment. This information is provided in a consistent, human-readable format that ensures a participant's informed consent is tied to the exact procedures of the experiment, as illustrated in Figure 4.2. When PRISONER workflows are replicated, the consent language is consistent.

- **IRB/Ethics** - The authors explain that "our design was reviewed and approved by our university's IRB." While it is encouraging to see this confirmed, the tendency to not share IRB protocols presents some challenges to reproducibility, particularly where the actual procedures of an experiment have drifted from the previously agreed protocol, so-called "ethical drift". While it is beyond the scope of PRISONER to resolve these challenges, allowing researchers to share a testable specification of the data-handling requirements of a study with their IRB when making an application, rather than a speculative protocol, constitutes an improvement on the state of the art.

- **Participant briefing** - The authors explain some of their briefing procedures, particularly "Our university's name and seal were featured prominently on every page of the survey and on the app's home page on Facebook.", which may have a priming effect and is important to be able to replicate. While researchers are responsible for conducting their own participant briefing, PRISONER provides a consistent "bookending" experience, including the presentation of consent forms, which explain the procedures of the experiment. This, when augmented by other cosmetic details, such as those outlined by the researchers in this study, provides a degree of consistency between replications.

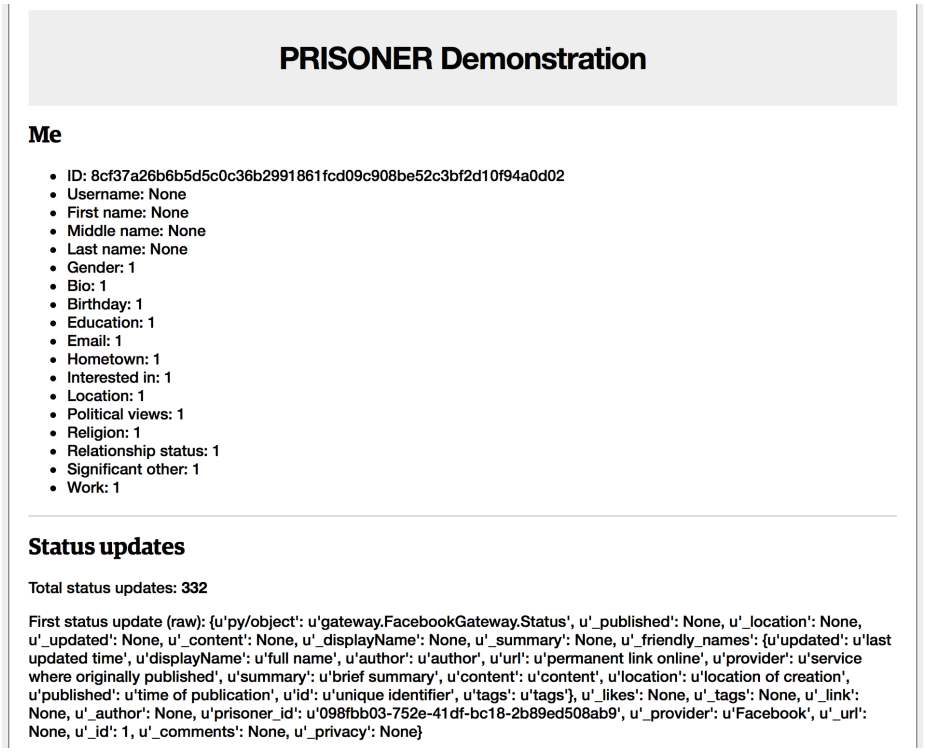


Figure 4.3: The result of making requests to Facebook with a PRISONER policy applied. Only the permitted objects and attributes are extracted, and are sanitised as outlined in the policy.

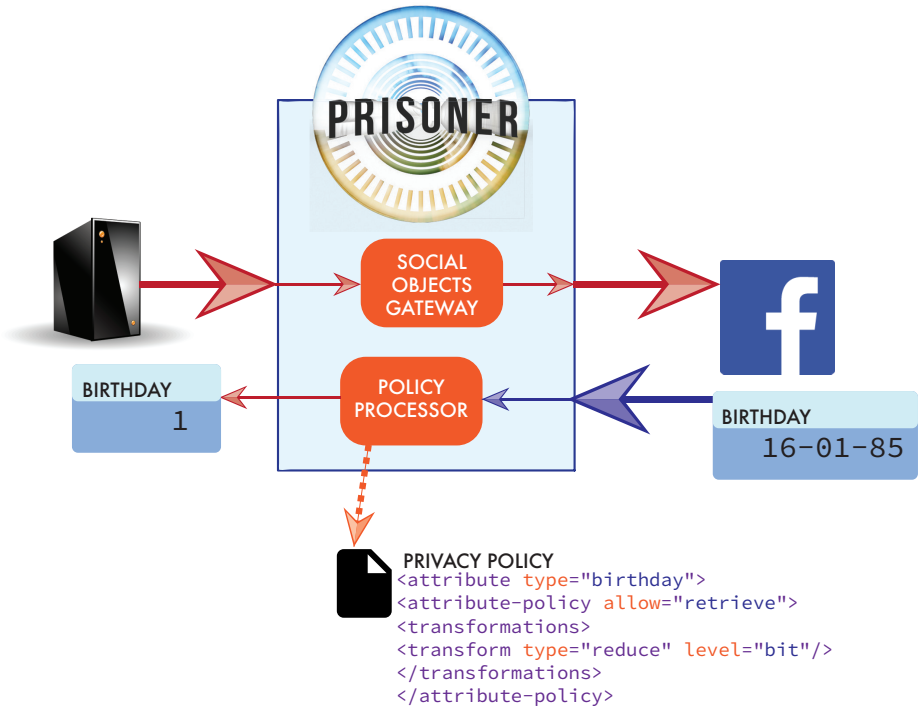


Figure 4.4: Illustration of how a request for data is handled by PRISONER. An experimental application makes a request to PRISONER's social objects gateway for an object, which is delegated to the appropriate social network. The object is returned and handled by PRISONER's policy processor, which invokes the privacy policy for the experiment to ensure the data are suitably sanitised. In this example, the participant's birthday has been reduced to a bit indicating its presence before being returned to the application.

We do not replicate the entire experiment in the study that we examine [66], but rather recreate its data collection requirements that we demonstrate in a simple example that attempts to retrieve a plethora of information about participants, with the sanitised results depicted in Figure 4.3.

As discussed earlier, this experiment requires access to a participant's Facebook profile to determime the presence of certain attributes. Figure 4.4 illustrates how data are handled by the framework for one such attribute. As shown, at the beginning of the experiment, the participant provides the PRISONER gateway with access to their Facebook account, binding this to their PRISONER session, and ensuring any requests for profile data are made via the PRISONER proxy. When the experimental application requests these data, PRISONER's policy processor consults the application's privacy policy for an appropriate "retrieve" clause to determine whether the application can access the attribute, and if any sanitisation should occur. In the example shown, the experimental application needs to determine whether the participant discloses their birthday. Thus, the policy processor sanitises the attribute before making it available to the application, and the sensitive attributes are discarded.

Having produced a policy file, it can be distributed to other researchers who can subsequently replicate the workflow. Even if researchers do not have access to the original code for the experiment, they can build an application against the same policy to make requests for data. They bootstrap an instance of PRISONER by pointing to the policy. PRISONER then generates all consent forms, briefing materials, and provides access to the SNS authentication flow and sanitisation API. In addition, if a researcher wished to run the same experiment using, e.g., Twitter as the source SNS, simply replacing any reference to "Facebook" with "Twitter" will provide this without any further modification. In this example, we have shown how an experiment can be managed by PRISONER in a reproducible and ethical manner. Even if we were to conduct this experiment and not share our application's source code, other researchers can replicate the experiment in their environment of their choice, but re-use our experiment's workflow to ensure data are collected under the same conditions. It is import-

ant to note, however, that workflow sharing alone is not sufficient to guarantee the accuracy of replications, particularly as this may not consider all possible corner cases that could affect the result of a replication. As we have discussed, a wider culture change is needed to achieve a higher degree of reproducibility.

The PRISONER tools have been made available to the academic community, [2] with examples and documentation to help researchers run their own experiments, and a Docker image which allows quick evaluation of a PRISONER instance and accompanying example application.

4.3 Addressing the requirements

Returning to the requirements we expressed at the beginning of this chapter, we summarise how the architecture we have outlined achieves these objectives, while discussing which aspects of the framework have yet to be realised.

1. *In our survey in Chapter 3.3, we found that SNS research was conducted using a wide variety of methodologies, including questionnaires, ethnographic studies, user studies, and crawling the data provided by SNSs. The architecture must enable all such studies, and we should be able to reuse system components for different methodological approaches.*

   PRISONER is agnostic about the nature of the experiments it is used to conduct. In Chapter 4.2, we recreated one type of user study. In the rest of this thesis, we will outline two other experiments that used PRISONER, demonstrating its versatility. In Chapter 5, we use the framework to collect information from participants' Facebook profiles to present to them in a questionnaire, while in Chapter 6, we will use the framework in a mobile user study to sensitively handle location data. The framework has also been used in a study of sharing preferences [88]. In further work, we will conduct further experiments to determine whether the framework's

[2] http://prisoner.cs.st-andrews.ac.uk

design and policy language is sufficiently expressive for all SNS experiments.

2. *The architecture must allow researchers to capture workflows, in terms of data acquisition, and privacy and ethical requirements, so that experiments can be repeated and experimental designs shared with other researchers.*

   We have introduced an expressive policy language that allows researchers to encode the requirements of their experiment in a human and machine-readable format. In Chapter 4.2, we showed that this was sufficient to reproduce an experiment from the literature. From the experiments conducted with the framework so far, we have found the policy language to be able to encode data acquisition requirements, and our ability to generate consent documentation constitutes an improvement on the state of the art with respect to encoding ethical requirements of a study. At this stage, however, the framework does not allow all methodological details to be encoded, and as such does not replicate all aspects of an SNS experiment. Specifically, we have not yet shown how PRISONER could be used to encode the sampling strategy of an experiment, and we have not yet developed the framework to be easily portable, such that another researcher can easily reproduce an experiment based on the policy of an experiment.

3. *The architecture must allow experiments to be run using data from a wide range of SNSs and other sources of social data, to allow experiments to be conducted against a range of data sources, without knowing details of their implementation. It should be sufficiently extensible such that support for new services can be added without any changes to the core framework.*

   In our current implementation, we have full support for Facebook, Twitter, and Last.fm, based on the requirements of the experiments we discuss in this thesis, and for other projects. These service gateways have been developed by others in accordance with our service gateway specification, which is testament to the expressiveness of the specification, and the ability for others to implement it. So far, however, we have only demonstrated the applicability of the specification to existing SNSs, and we have not yet

evaluated whether it is appropriate for querying other sources, such as adapters for existing datasets.

4. *As identified in Chapter 2.3.5, and illustrated in our survey, a barrier to reproducibility is that the SNSs may change their APIs at any point, which can break existing code, or indeed the whole service may cease to exist. Our framework should mitigate the impact of these changes, by handling mismatches between new APIs and old code, and allowing the environment of a defunct SNS to be emulated where it is no longer available, allowing privacy-preserving requests to be made of legacy SNS datasets.*

   Obsolescence is a significant challenge for SNS research, and not one our current implementation solves. The first issue, breaking API changes, is somewhat mitigated by the versioning scheme we discuss in Chapter 4.1.7, but this requires maintenance of the service gateway implementation, and may not be sufficient for all API changes. Supporting a completely defunct SNS is not an issue we have addressed in the current implementation.

5. *The architecture must maintain the contextual integrity of participants, by respecting the context-appropriateness of information transfers.*

   The design we have adopted aims to maintain contextual integrity by providing transparency to participants about how their data are processed, ensuring there is no disparity between what data collection a participant agrees to, and the data that are collected about them. Furthermore, consent acquisition is embedded into the same context as data collection, to remove ambiguity about the practices of a study.

4.4 Summary

In this chapter, we have introduced a framework for conducting ethical and privacy-preserving SNS experiments, informed by contextual integrity. We note the following:

- By using standards-based Social Objects and sanitising data before they are analysed or presented to participants, we have shown how we can run experiments that scale across a number of social network sites, without collecting or disclosing sensitive data.
- We have demonstrated how the framework supports reproducibility by re-creating an experiment from the literature, allowing us to share the protocol so it can be executed by others.
- We have discussed how our workflow management tools constitute an improvement on the state of the art, but acknowledge that the current implementation does not address a number of the requirements we believe such a framework should achieve.

In the next chapter, we discuss the first of two experiments that use the framework, applying contextual integrity a significant methodological challenges we identified earlier: acquiring informed consent for SNS studies.

Chapter 5

Improving SNS research methodologies with contextual integrity

In Chapter 2.3.3, we identified a number of challenges with obtaining informed consent in SNS studies, while in Chapter 3.3.5, we found that reporting of the ethical conduct of studies, including whether consent was sought from participants, is lacking. This raises the question of whether traditional forms of informed consent are sufficient for conducting research using SNSs, where lots of sensitive data may be collected. We investigate this by applying contextual integrity to the process of acquiring consent. Leveraging social norms of willingness to share data with researchers, we will determine whether this can be an appropriate proxy for asking people's explicit consent to share each piece of data.

In this chapter, we make the following contributions:

- We develop a means of quantifying norms for sharing SNS data with researchers.
- We introduce a new method for acquiring informed consent, informed by contextual integrity.
- We evaluate this method by applying our framework from Chapter 4, com-

paring the effectiveness of our method to two other methods often used in SNS research, assessing its performance by the accuracy of the method, and the burden placed on participants.

In Chapter 2.3.3, we characterised informed consent as being secured or sustained in nature, noting a tension between the accuracy of the method for gaining consent, and the burden put on participants. In this chapter, we propose a third way of acquiring consent that aims to achieve similar accuracy to the sustained approach, while placing a low burden on participants, as with the secured approach. This approach is informed by Nissenbaum's model of contextual integrity, which we introduced in Chapter 2.3.6. We leverage the notion of context-appropriate social norms, to determine what amount of information is deemed socially acceptable to be collected by researchers, as a proxy for asking participants to explicitly sanction the sharing of each individual piece of data, as in the sustained approach.

We evaluate this approach in the context of SNS research, in a study that asks participants to share data from their Facebook profiles with a researcher. By detecting whether participants conform to norms of willingness to share such data, we aim to reduce the number of times we need to explicitly ask for their consent, with minimal loss in accuracy. We compare its performance to the secured and sustained approaches based on the two measures of *accuracy* and participant *burden*.

We hypothesise the following:

- **H1:** Acquiring consent to share SNS 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 that involve no such interventions.
- **H2:** Acquiring consent with contextual integrity is as robust to temporal changes in willingness to share data as sustained consent.

5.1 Method

To test these hypotheses, we designed a study to investigate whether a method for acquiring sustained informed consent based on contextual integrity performs better than two other strategies. These can be summarised as:

- **Secured consent**: Participants provide up-front consent to data acquisition in an informed consent form.
- **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.
- **Sustained consent**: Participants are asked explicitly about their willingness to share each piece of data.

5.1.1 Willingness-to-share norms

Our contextual integrity method requires a set of social norms of willingness to share SNS data with researchers, in order to determine whether participants' behaviour can be considered norm-conformant.

A previous study produced a relevant dataset showing people's willingness to share SNS data of various types with researchers [88]. In this 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

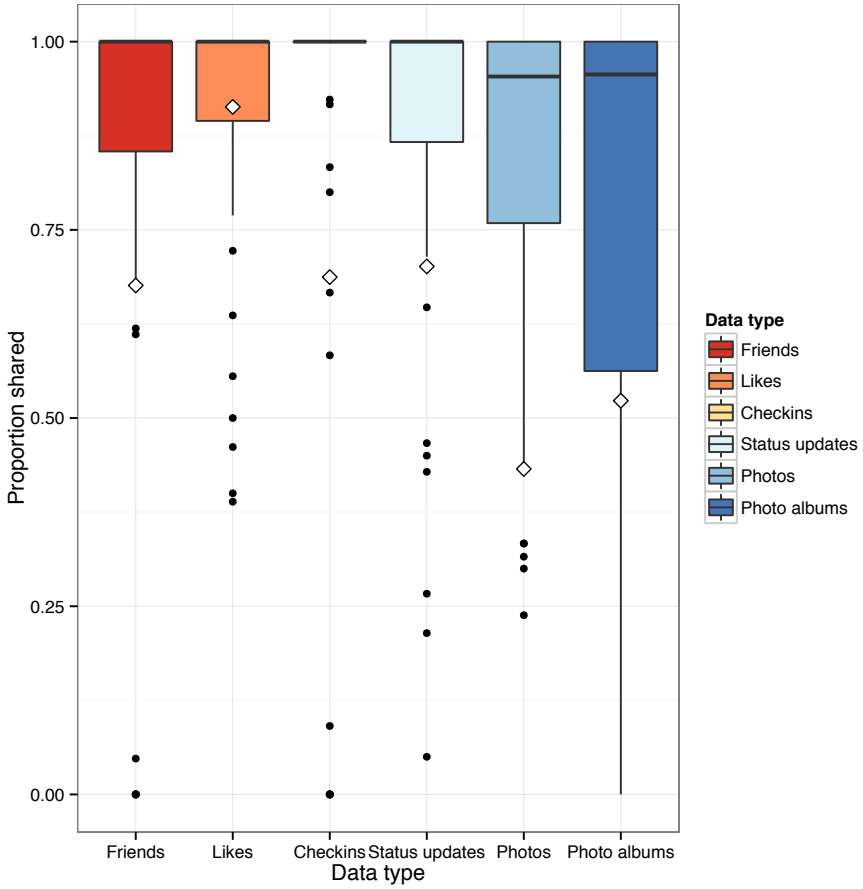


Figure 5.1: Boxplot showing sharing rates of different data types for users in the sustained consent condition in the 2014 study. For comparison, the white diamonds represent the mean sharing rate of that type in the 2012 dataset, constituting the prevailing norm. Willingness to share data with researchers has increased on all fronts in this period.

- Photo albums created by the user
- Biographical attributes
- Status updates

We consider the proportion of shared content for each attribute to represent the norm for that attribute, as depicted by the white diamonds in Figure 5.1. We draw on these norms in our study to determine the extent to which our participants conform with them.

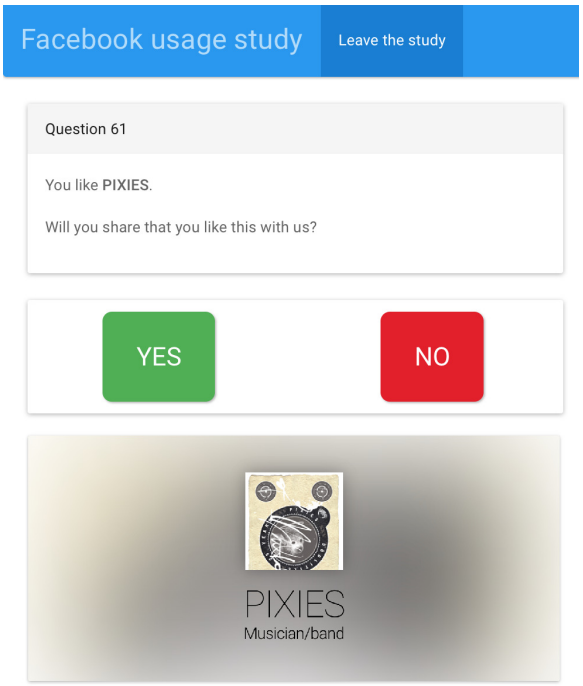


Figure 5.2: A screenshot of a question in the study. In this case, the participant is asked to share the fact that they like a Facebook page with the researchers.

5.1.2 User study

To evaluate our method of consent using contextual integrity, we conducted a user study. Participants were recruited online (through advertisements in university mailing lists, Twitter, and Reddit) for a study that modelled the three consent strategies and evaluated their performance in terms of accuracy and burden, 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 5.2). These data were of the types found in the norms dataset, with the exception of biographical attributes, which we excluded as these data are static in nature, and not likely to exhibit interesting temporal properties. 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, which we explain for each condition later.

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, which we use in our contextual integrity condition. After each participant answered a question, we performed a series of chi-square equality of proportions tests, comparing the distribution of responses by the participant 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 the secured consent 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.

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 that they did not wish to have shared, as shown in Figure 5.3.

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.

Our study allows the following metrics to be calculated:

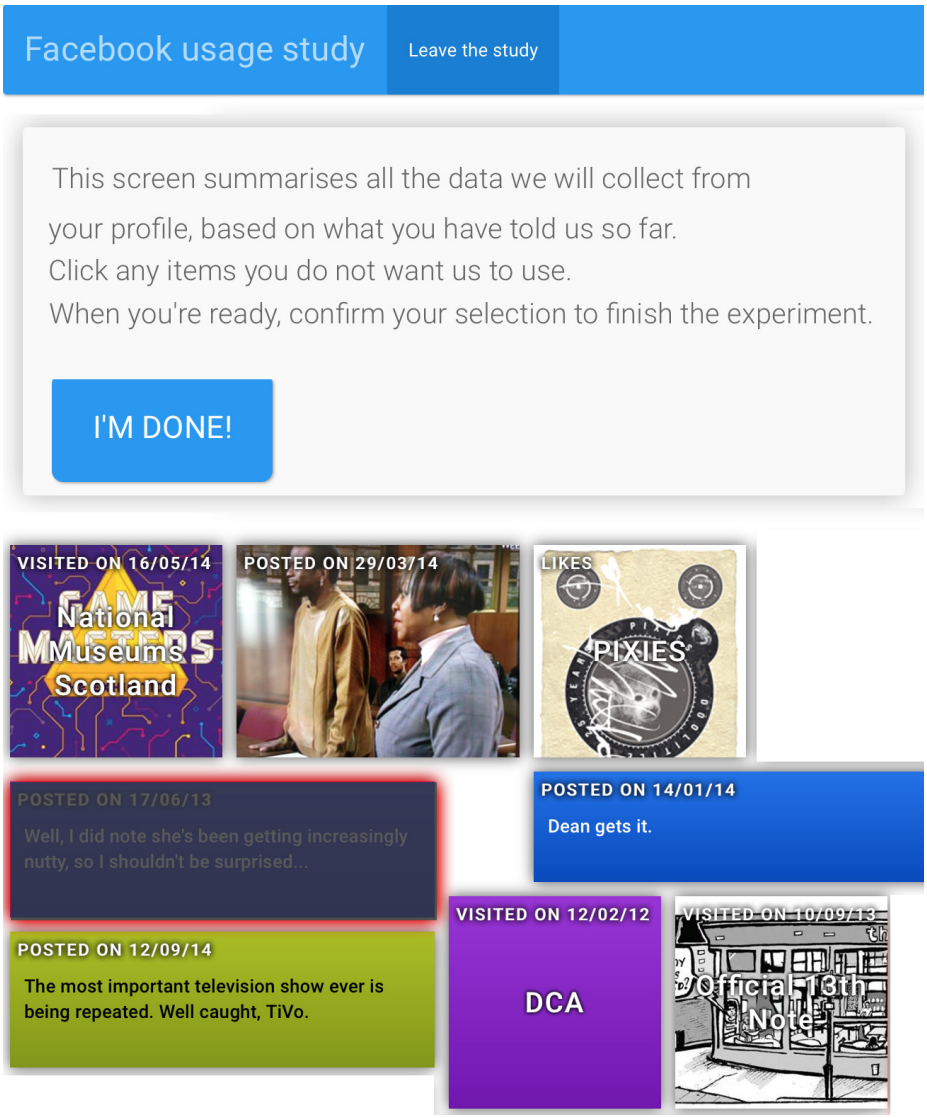


Figure 5.3: A screenshot of the performance evaluation step. Participants are shown the content collected according to their consent policy, and asked to remove any items they deem inappropriate to share. The proportion of items selected determines the accuracy of the method.

- **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 that 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**: As discussed in Chapter 2.1, willingness to share social network data is temporally sensitive, 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. We use this time period as the literature demonstrates that privacy decisions are able to change within a week [5].

5.1.3 Applying the PRISONER framework

When conducting this experiment, we used the PRISONER framework introduced in Chapter 4. As a Facebook service gateway had been developed for a previous study [88], it was trivial to develop the interactions with Facebook. A

policy was written that ensured the experimental application could only retrieve Facebook data for the purposes of presenting questions to participants. As the content of these data were irrelevant to the study, the policy ensured that this content was not stored by the experiment application.

Before conducting the study, our experimental design was scrutinised and approved by an ethics committee within the School of Computer Science.

5.2 Results

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 5.1: 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 5.2: 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.

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 5.1, participation was broadly equal across all three conditions. In our analysis, we consider the responses of all 109 parti-

cipants, with only those who completed both weeks considered in our temporal analysis of the robustness of contextual integrity. Table 5.2 shows the demographics of our sample population, compared to data Facebook make available to their advertisers.[1] 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. In our results, we did not find any significant relationships between willingness to share data and these demographic factors.

5.2.1 Evaluating the norms dataset

The effectiveness of using contextual integrity to reduce the burden of acquiring informed consent is incumbent on the quality of the social norms that we calculate. The 2012 dataset used for this similarly oversampled undergraduate students and so is suitable for comparison. To observe how behaviour has changed in the two years before this 2014 study, we compare the amount of data shared by participants in our *secured consent* condition to those in the 2012 dataset. We isolate these participants as they were asked the full set of possible questions, without norm conformity being used to remove questions from the set. Figure 5.1 shows that participants in our study are much more willing to share data than in 2012. For all data types, willingness to share is greater than in 2012, although in both sets, willingness to share photos and photo albums is lower. We believe the high variability in willingness to share photo albums is because some participants interpret this as sharing the album and all photos within, while others consider it to only be sharing knowledge of the album, which would result in the exposure of much less data. While our study does not address the cause of this trend towards increased sharing, it may be due to the increased adoption of SNSs, and the sharing of content on these services.

[1] Facebook Ads Manager: facebook.com/advertising

5.2.2 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%. 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 5.4 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 per-

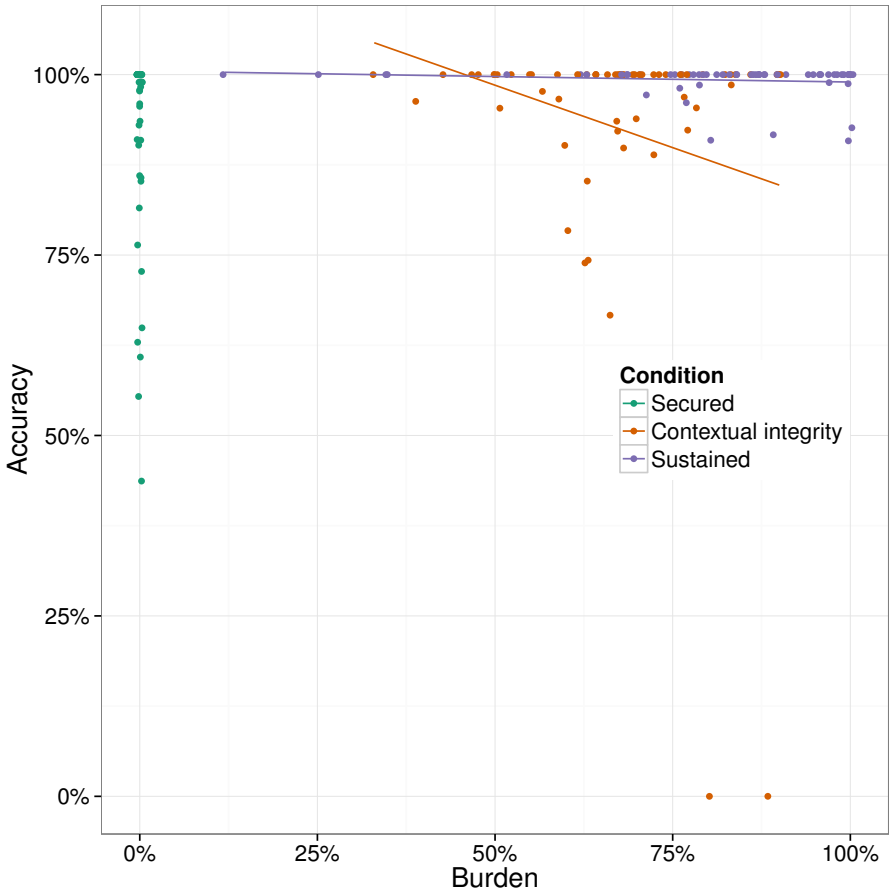


Figure 5.4: Scatterplot showing the relationship between accuracy and burden in all three conditions. Accuracy is the most variable for those in the secured consent condition, whereas in the case of contextual integrity, a small loss in accuracy is met with a greater time burden saving. Note that the points have been jittered to improve readability.

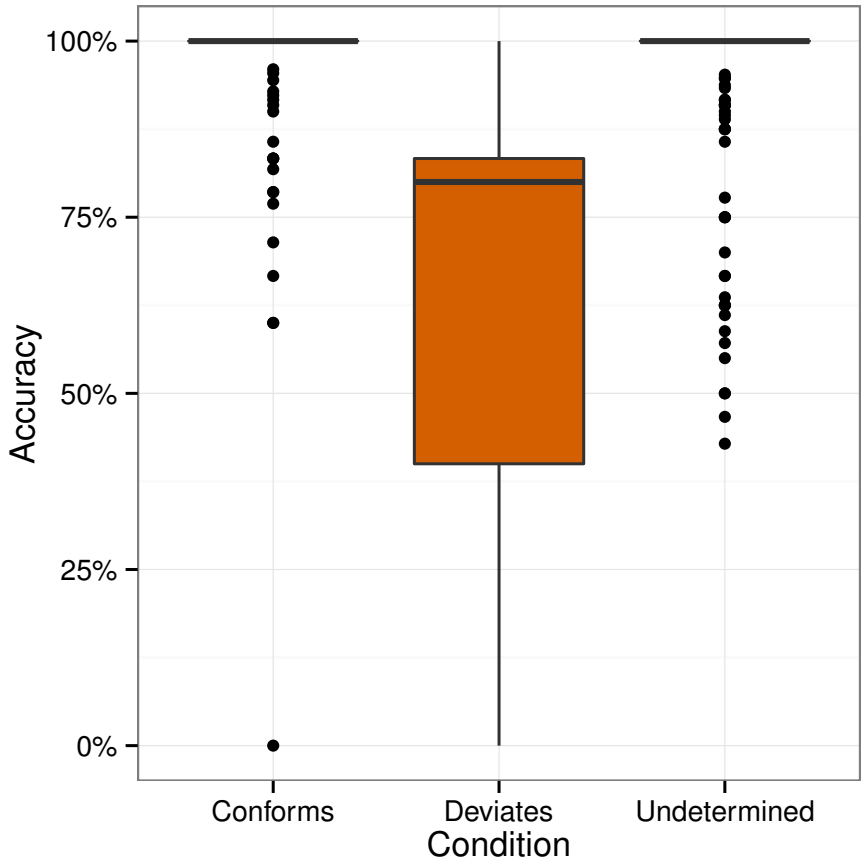


Figure 5.5: Boxplot showing how accuracy differs between participants in the contexutal integrity condition depending on their norm conformity. Norm-conformant people achieve higher, and less variable, accuracy rates than those who deviate from norms.

formance. As shown, surprisingly accuracy drops with increased burden. This exposes an important detail about the 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 5.5, the former was not useful for maximising accuracy. As deviance requires more questions to detect, this drags down overall accuracy as burden increases. In addition, this highlights that where norm conformity was identified, accuracy improved slightly compared to those whose conformity could not be determined, with the exception of one outlier. 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.

5.2.3 Does contextual integrity reduce participant burden?

We use contextual integrity in the 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.[2] In the contextual integrity condition questions were not asked if the participant was found 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, a 21.9% decrease in burden. When comparing the distribution of burden between the two condi-

[2]For some participants, this number was smaller in practice if they did not have enough pieces of data in their Facebook profile of each type.

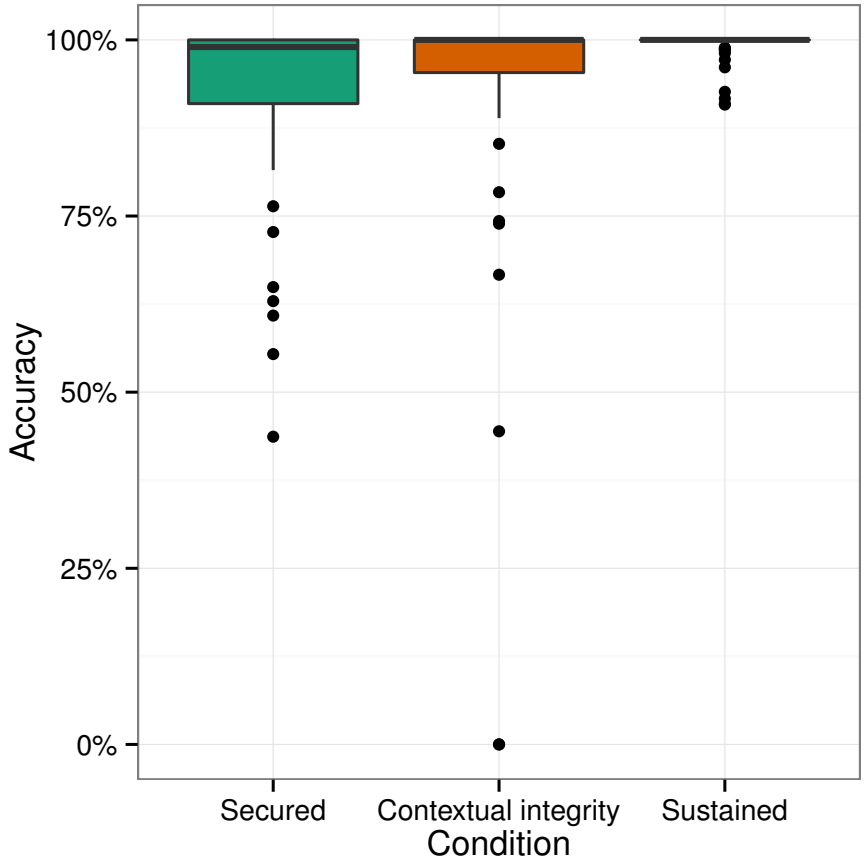


Figure 5.6: In all three conditions, median accuracy is high, although variability increases as the number of questions asked is reduced.

tions, 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.

5.2.4 Does contextual integrity significantly reduce accuracy?

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 5.6. 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, the study from which the 2012 norms are derived similarly found that people are more willing to share social network data with academic researchers depending on the purpose of the study [88]. 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 conditions 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.

5.2.5 Who does contextual integrity work for?

We classify participants in the contextual integrity condition based on their norm conformity. Figure 5.8 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 5.7 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),

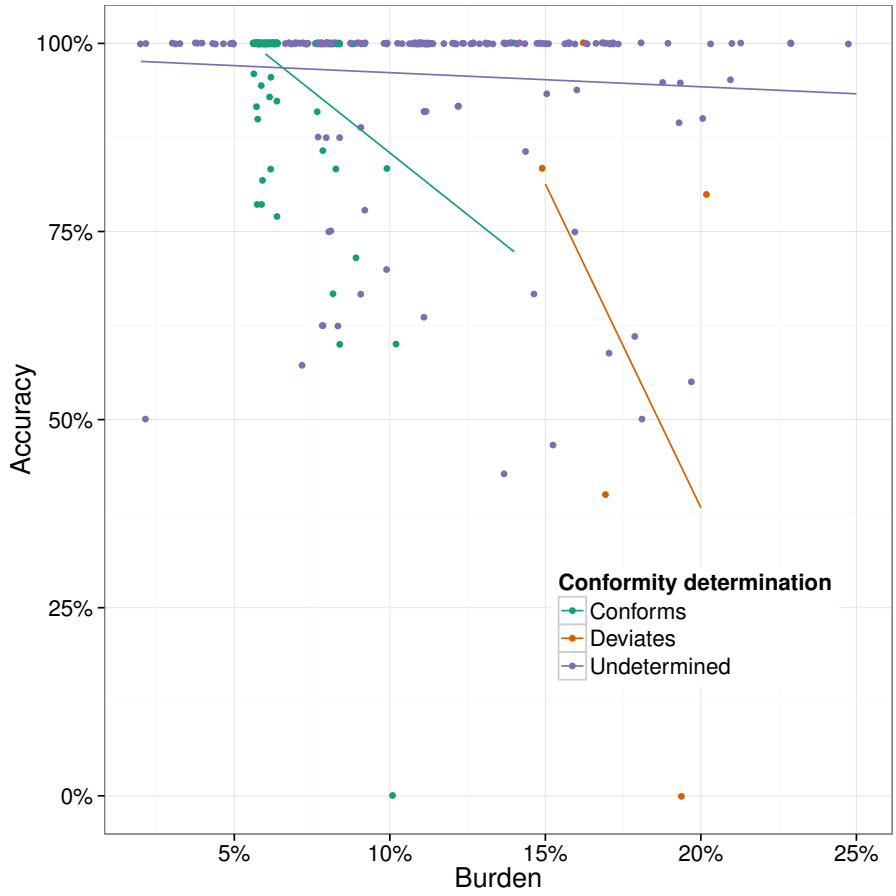


Figure 5.7: Scatterplot showing the relationship between norm conformity and accuracy. As indicated by the cluster after 5 questions (5% burden), high accuracy can be preserved when norm conformity is detected quickly, although the technique is not useful for people who are highly norm deviant. Note that the points have been jittered to improve readability.

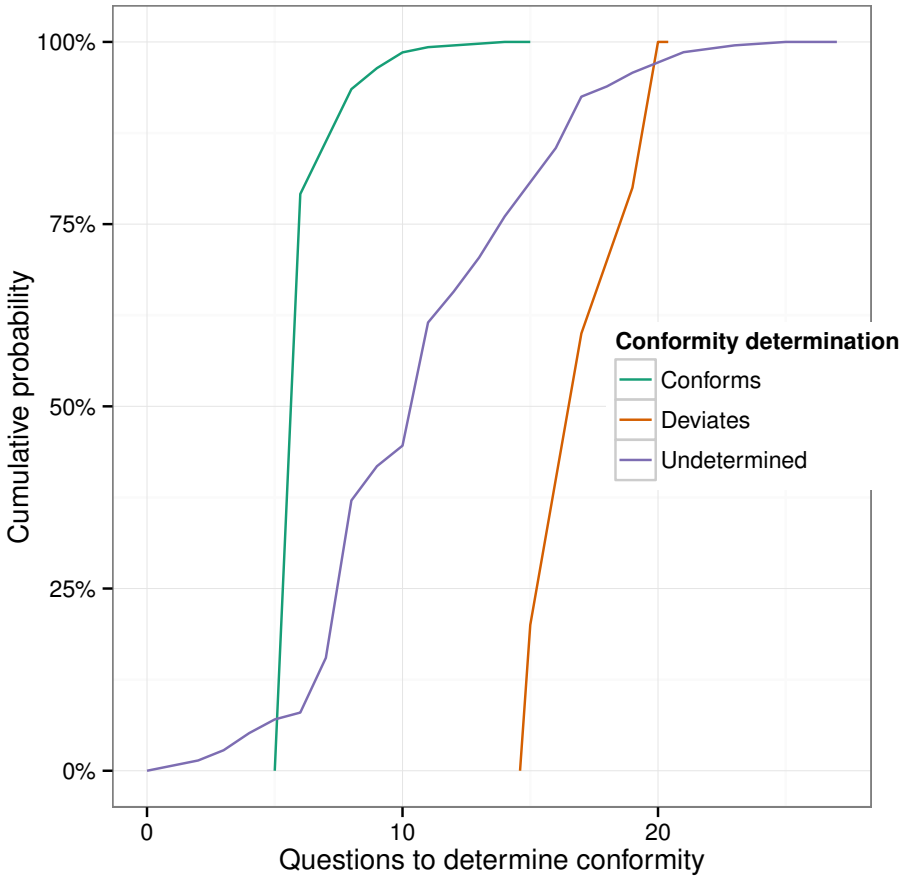


Figure 5.8: Cumulative distribution function of the number of questions needed to determine whether participants in the contextual integrity condition conform or deviate from the norm, or if this could not be determined. If people conform to norms, this is detected after a small number of interventions.

suggesting that people who behave similarly to their peers can be asked fewer questions and achieve higher accuracy. Users who are norm-deviant often need to answer more questions, with deviance detected within 15 to 20 questions. We found, however, that removing questions for norm-deviant participants actually hurt accuracy. A very small number of participants behaved in such a way that deviance-detection was employed, but the poor performance of this technique has excluded using it in future applications of the technique.

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, 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 relative to the sustained condition. 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 consent 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 to maintain high accuracy.

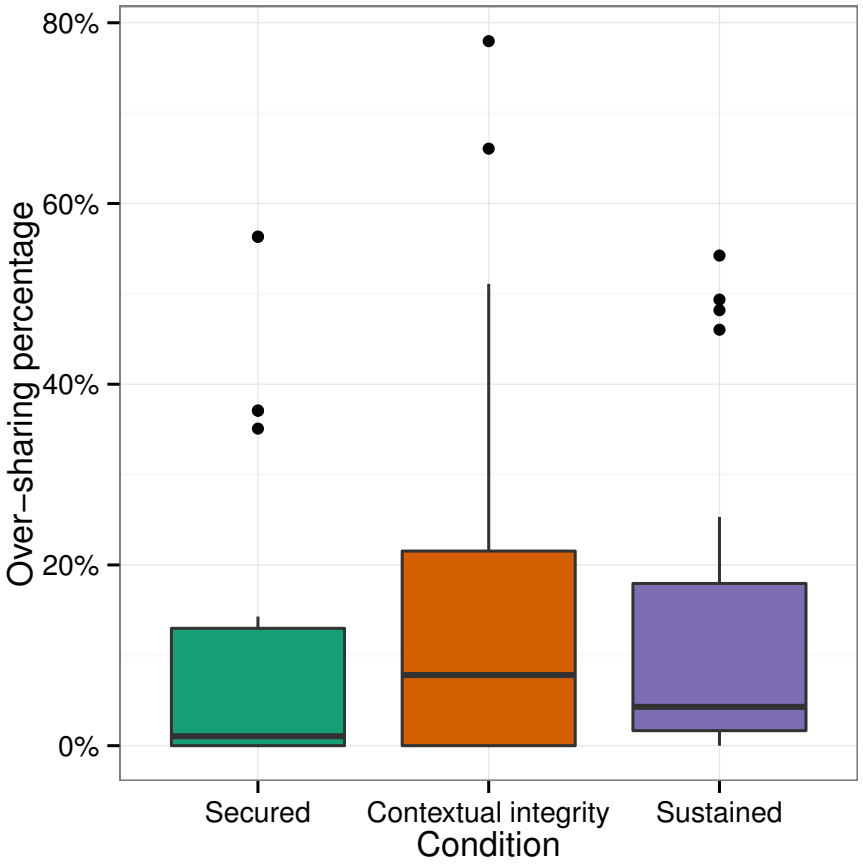


Figure 5.9: Boxplot showing how robust each condition is to temporal change by using the consent policy from the first week to predict the participant's responses in the second week. All conditions exhibit over-sharing, highlighting the difficulty of capturing consent at a single point in time.

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

Attitudes towards privacy and sharing social network data are temporally volatile, 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. As discussed in Chapter 2.1, this may cause regret, and perceived leakage of social network data. 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. 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, just as how accuracy is measured. Figure 5.9 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 over-sharing is not significantly higher in this case ($F(2, 65) = 0.168, p > 0.1$). 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 that would better capture their intent.

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 compared to the sustained condition. We also find support for H2, as the contextual integrity method is not significantly less robust than sustained consent over time.

We note that as human behaviour is so diverse, there is no "one-size-fits-all" approach to consent that achieves optimal results. A benefit of the method we introduce 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 found 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, determine 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 [17]. 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.

5.3 Summary

In this chapter, we have presented the first application of contextual integrity to the acquisition of informed consent for sharing SNS data. We note the following:

- Contextual integrity can be leveraged to reduce the burden placed on participants to acquire consent on average by 21.9%.
- For 27.7% of participants who are highly conformant to social norms, contextual integrity can deliver accuracy paralleling that of burdensome sustained consent, while reducing burden by 41.1% compared to the sustained condition.
- Using norm conformity to determine consent is temporally robust over a week, but further work is needed to determine whether this holds true for longer periods.

Having shown how we can apply contextual integrity to a framework for studying SNSs, and demonstrating that it can constitute an appropriate means of acquiring informed consent in SNS studies, in the next chapter we consider

how the framework can be applied to the study of emerging SNSs, to examine their potential privacy impacts.

Chapter 6

Identifying privacy breaches in emerging SNSs with contextual integrity

In this chapter, we demonstrate how contextual integrity can be used as a diagnostic to determine the potential privacy impacts of emerging SNSs. We conduct a user study to investigate the potential risks when adding financial incentives to LBSNs.

In this chapter, we make the following contributions:

- We conduct the first application of contextual integrity to identify the norms governing the use of incentivised location sharing systems.
- We conduct a user study of 22 smartphone users to understand expectations and motivations in an incentivised location sharing system, and to see whether feedback affects willingness to disclose locations.

The rise of smartphones and mobile sensing smart devices is enabling vast amounts of personal data to be collected or generated, and optionally shared with other people, services, and businesses, as we discussed in Chapter 2.2. Such self-tracking is the logical extension of context-sharing applications such as Foursquare. Such services are increasingly delivering financial incentives to

encourage people to act as an advertising agent on behalf of a business to their social network. The introduction of incentives raises several questions. Do they affect privacy concerns, or people's uses of such services? Do people's decisions change for the worse as a result of incentives, and how, or indeed should, we improve this? In this section, we look at ILS services (Chapter 2.2), to determine whether they may constitute a risk to individual privacy. While the use of LBSNs has been well studied, this is the first user study to examine the potential risks in embedding financial incentives in traditionally social-driven online sharing.

To determine whether incentives may introduce new privacy violations, we use Nissenbaum's model of contextual integrity, introduced in Chapter 2.3.6. Specifically, we use the contextual integrity *decision heuristic*, a diagnostic tool for determining whether a new process risks violating the expectations within an entrenched context, as described in Chapter 2.3.6. We discuss a preliminary user study in which 22 people use an ILS application for one week, and receive financial incentives to share their location with businesses and their social network. The user study allows us to better understand the expectations of people using such an application, their behaviour and motivations for disclosing their location for a financial incentive, and how the design of the application affects how people use it. We use these findings to complete the decision heuristic and decide whether ILS constitutes a privacy violation. In addition, we recommend a number of best practices for application developers, to provide services that deliver incentives for disclosures in a way that preserves people's comfort and privacy, while delivering benefits to advertisers and developers.

This study aims to address the following research questions:

1. Do users of LBSNs have different expectations of privacy when their disclosures are financially motivated?
2. Do incentives perturb privacy norms in an LBSN to the extent that contextual integrity is violated?
3. Does greater transparency about the flow of personal information when

users share their location for money affect behaviour, and reduce the risk of privacy violations?

6.1 Applying the decision heuristic to ILS

As introduced in Chapter 2.3.6, the first stage of the decision heuristic requires the understanding of information flows through the application in question. We therefore first examine Quidco, one of the commercial ILS services described in Chapter 2.2. This application allows users to share their location with their social network, with the addition of a financial incentive provided to the user whenever they disclose their patronage at a participating business, who in return receive some personal information about the user [1]. We now step through the decision heuristic in turn.

The first six steps of Nissenbaum's nine step heuristic allow us to identify whether ILS systems generate changes in the transmission principles present in the location-sharing context, raising an immediate "red flag" as violating entrenched informational norms. This *prima facie* assessment about the potential privacy impacts of this transformed context motivates further exploration of the context, without considering the wider moral factors [100] beyond the scope of this study. This assessment is informed, where possible, by previous work that has identified the expectations, behaviours, motivations, and norms of the location-sharing context, although we identify the aspects of the ILS context not yet studied, which we capture in this user study. We structure our analysis in the same vein as Jones and Janes' application of contextual integrity to Google Books [59].

[1] Quidco Privacy Policy: http://www.quidco.com/privacy-policy

6.1.1 Information flows

The first step in applying the decision heuristic is to identify the flows of information that manifest in the context. In an ILS service, there are three primary information flows during use of the application:

- A user's identity and location is disclosed to the provider of the ILS service so it may return context-appropriate adverts to display to the user.
- If the user checks in to any of the locations advertised, some personal information is disclosed to the relevant advertiser.
- When checking in, a user's location data may be disclosed to their social network on a service such as Facebook or Twitter.

Nested contexts

ILS systems, as with many other LBSNs, disclose location data to other SNSs, which facilitate further engagement around the location disclosure, such as commenting or soliciting "likes". Therefore, we consider the SNS a nested context, subject to its own entrenched norms, although we do not consider this in detail in this analysis. There is a risk of impacts from these contexts, particularly where the expectations associated with the LBSN do not align with those of the SNS. Users may exhibit boundary regulating behaviours on the LBSN independently of the SNS, and where these do not align, a privacy violation may occur [104]. We do not expect the introduction of an incentivised components to LBSNs to dramatically alter the incidence of these collisions.

6.1.2 Information subjects, senders, and recipients

When discussing the flows of information, we must consider who is responsible for sending and receiving information, and to whom the information relates

(the *subject*).

In the ILS context, the subject of the information is the system's users. The flows of information within the context are designed to enable the disclosure of locations by these users to their social network, and the disclosure of some personal information in exchange for a financial incentive.

The *senders* of information are the users whenever they request relevant adverts from a set of advertisers, or disclose their location to their social network and PII to an advertiser.

The *recipients* are the advertiser, when they receive a request for advertisements, and when they learn the identity of a user who has checked in to their business. In addition, the user's social network, and the provider of the ILS service, receive identifiable location updates from the user.

6.1.3 Information attributes

Information attributes refer to the types of information shared within a context.

When a user initially registers with the ILS service, they provide PII such as their name and email address, as well as demographic information of likely value to advertisers such as their age.

Users transmit their current location and a personal identifier to the service provider in order to receive location-specific adverts. The service provider uses the identifier to authenticate and log the user's activity, and the user's coordinates are used to find relevant advertisers. At this point, there is no sharing of PII with an individual advertiser.

When a user checks in to a location, their identity and patronage at that location is disseminated to their social network, or delegated to another SNS to increase its reach. In the case of Facebook, this is transformed into a status update, outlining the name and location of the business the user is visiting,

and a current timestamp. In addition, the user's identity is disclosed to the relevant advertiser, who delivers a financial incentive to the user. The provider of the mobile advertising system may also deliver demographic information to the advertiser to assist their advertising targeting practices.

6.1.4 Transmission principles

Transmission principles are the rules that govern people's information sharing in a particular context, and are a critical aspect of any context-relative informational norm. In an LBSN, as in other SNSs, disclosures are motivated by impression management and to build social capital, without an expectation of reciprocity, as discussed in Chapter 2.1. Disclosures to the social network are generally accompanied by an expectation of confidentiality, in so much as users would not expect the location to be relayed to an audience beyond their social network. The provider of the LBSN is trusted to not sell or otherwise share these data to third parties.

We do not have an exhaustive understanding, however, of all the transmission principles that manifest in the ILS context. We might expect that location-sharing transmission principles may apply to some extent, as incentivised location disclosures are made in much the same way as in a traditional LBSN. In the ILS context, however, where users receive a financial incentive for their disclosures, different transmission principles may emerge. The additional flow of PII to advertisers may introduce new data protection requirements, depending on jurisdiction, or regulations governing the handling of marketing data. The extent to which users appreciate these differences when deciding to disclose their location data is unknown. In addition to the social motivations outlined earlier, we also do not know whether financially-incentivised users have different expectations about when their social network should be able to access their location, or how they should consume this information, with the context of requests for location data identified as a common concern in Chapter 2.2.

We hypothesise that the ILS context may perturb some of these principles, by introducing a financial, rather than social, motivation for disclosing locations. The myriad unknown transmission principles motivate a user study in which we examine motivations for sharing and withholding location data, and how the design of an ILS application can affect the incidence of these motivations. We need to additionally capture users' expectations of how their data are transmitted to other parties. The prevailing motivations and expectations will constitute our set of transmission principles. If this set is significantly divergent from motivations identified in the location-sharing literature, this may constitute an unacceptable departure from the norms established in the location-sharing context.

6.1.5 Entrenched informational norms

Entrenched informational norms describe the existing sets of practices that prevail in a given context, encompassing the flows of information and expectations of the actors involved.

The ILS context inherits the entrenched norms governing location sharing, specifically the same context, senders, subjects, and attributes of information. The significant points of departure involve the introduction of new recipients of information and transmission principles.

As in a traditional LBSN, ILS users disclose their own location selectively to their social network. While an LBSN motivates disclosure through perceived social benefit, ILS augments this by introducing a financial incentive for disclosing certain locations. The substance of the location disclosure is the same, but some personally identifiable information is provided to the owner of the business whenever this happens, in return for the financial incentive. The introduction of this additional exchange of information is conflated with the social benefit of the location disclosure, so it is not known to what extent a disclosure was motivated by traditional location-sharing norms of building social

capital, or the financial incentive, and how willingness to disclose is reduced by concerns about sharing personal identifiable information with a commercial third-party.

In an LBSN, the user directly controls the disclosure of their location, and is able to review the content and audience of this disclosure on the social network site to which it was published. This is maintained in the ILS context, however some personal information is also disclosed to an advertiser, without similar in situ feedback. In our user study, we explore whether different degrees of transparency about secondary use of personal identifiable information affects willingness and motivations for sharing. Through the lens of contextual integrity, we will determine whether some of these disclosures may be considered inappropriate, if the user's mental model of the flows of information is not reconciled with reality, and if the addition of an incentive significantly increases disclosure rates without a clear understanding of the implications of such disclosures.

Table 6.1 summarises the steps of the decision heuristic. As shown, while we are able to comprehensively illustrate the fundamental aspects of the ILS context, we do not know all the transmission principles, particularly the expectations of users regarding how their information is used. Our user study allows us to identify these transmission principles and complete the decision heuristic. In this study, we focus on the steps needed to render a *prima facie* assessment regarding a potential breach of contextual integrity. Steps 7 and 8 invite an in-depth study of the wider moral and political factors, and the extent to which these factors impinge on the values and goals of the context. We choose not to examine these factors in this work as they are dependent on this initial assessment having been made.

This initial analysis shows that an ILS service constitutes an extension of the norms within an existing LBSN, as a user's interaction with the system, the actors, and information flows are largely inherited from a social-driven LBSN. The introduction of financial incentives itself does not affect the flow of

Step	What we know
1. Information flows	• Location disclosure to platform operator • PII disclosure to advertiser • Location disclosure to social network
2. Prevailing context	Location-sharing services
3. a) Subject b) Senders c) Recipients	Users' PII and disclosed locations Users Users' social network and advertisers
4. Transmission principles	To be identified in user study
5. Entrenched informational norms	• Location disclosures are voluntary • Disclosures try to build social capital
6. Prima facie assessment	To be made after user study
7. Moral and political factors	*Out of scope in this study*
8. Impingement on contextual values	*Out of scope in this study*
9. Recommend for or against the practice	Initial recommendation to be made after user study

Table 6.1: A summary of Nissenbaum's decision heuristic, showing what we know at each step prior to the user study.

information, but may affect the motivations for disclosures. This in turn may constitute a violation of privacy if people's mental models are not sufficiently adapted to this adjusted context.

6.1.6 Method

To complete Nissenbaum's contextual integrity decision heuristic, we conducted a week-long user study with 22 participants, designed to identify the prevailing norms and expectations necessary to determine whether ILS constitutes a *prima facie* violation of contextual integrity. We chose to run the study for seven days based on recommendations from the experience sampling literature [47] and from running such studies in the past.

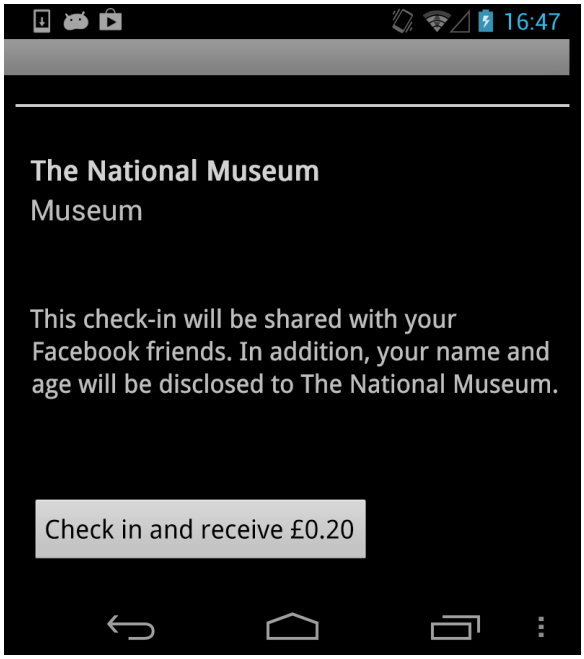


Figure 6.1: Screenshot of the incentivised location sharing application created for our user study. One group of participants are shown information about the flow of PII before confirming a check-in.

For this user study, we developed an application for Android smartphones that closely resembled the interface and feature-set of existing commercial applications such as Quidco and Foursquare. The application consisted of a widget that used the Google Places API to periodically update and display the names of businesses close to the participant's current location. From the widget, the participant could select a nearby business and check in, thus creating a Facebook status update, in exchange for a small financial incentive. At the start of the study, participants chose six of their Facebook friends who would be able to view these stories, representing a cross-section of close friends, acquaintances, and colleagues. By choosing a small subset of people to share locations with, we mitigate potential adverse effects of over-sharing in the study, considering our interest in potentially inappropriate disclosures, while still making participants consider the social impact of their disclosures to a diverse audience. Participants could pause the application for a short period of time if they did not want location data to be collected.

Study design

Participants were randomly assigned to one of three conditions, which affected the feedback displayed to participants immediately before they checked in, and the value of the cash incentive. These conditions were:

- Low incentive, no feedback (**LowNo**): Participants received £0.10 for each check-in, and were not actively reminded that PII would be disclosed to the business.
- High incentive, no feedback (**HighNo**): Participants received £0.20 for each check-in, and were not actively reminded that PII would be disclosed to the business.
- High incentive, feedback (**HighYes**): Participants received £0.20 for each check-in, and were reminded that their name and age would be disclosed to the business.

Participants in the HighYes condition were shown the information depicted in Figure 6.1, while other participants were only informed that their check-ins would be disclosed to their Facebook friends before checking in. These incentive levels were chosen based on the distribution of incentives we find in commercial applications such as Quidco. The high incentive level was set at £0.20 because we were interested in seeing whether differences would manifest even between marginally different levels of micro-incentives, and to avoid compelling lower-income participants to check in out of financial need, which we are not investigating in this study. Participants were not told that there were other conditions, nor how the incentives were chosen.

Before joining the study, all participants were asked to read the application's privacy policy, which specified that the business indicated would receive some PII in return for the financial incentive. Our feedback conditions aimed to determine whether increased visibility of this sensitive information flow sig-

nificantly affected willingness to check in, and the motivations for completing a check-in.

Before beginning to use the application, participants completed a pre-briefing questionnaire, consisting of 15 questions drawn from the 'collection', 'control', 'awareness', and 'secondary use' dimensions of the Internet Users' Information Privacy Concerns (IUIPC) scale [82]. To these we added a question identifying expectations in the ILS context. This questionnaire is shown in full in Appendix C. We did not include questions pertaining to the 'errors', 'improper access', or 'global information privacy concerns' dimensions as they were tangential to this study. Immediately after completing the study, participants were asked to complete the same questionnaire, allowing us to identify a relationship between different feedback conditions and a change in privacy attitudes.

The application was instrumented to record all completed check-ins, and *abandoned* check-ins (when the participant accesses the check-in interface, shown in Figure 6.1, but does not complete a check-in). In addition to recording the participant's activity during their seven days of participation, participants received an automatically-generated end-of-day questionnaire each night, based on their activity during the preceding day. This allowed us to capture qualitative data about the motivations for activity within the application, and to clean anomalous outliers (such as accidental interface taps) within hours of the activity occurring.

Recruitment

Participants were recruited through advertisements on Facebook, mailing lists aimed at university students and staff, and viral messaging on social networks such as Twitter and Facebook. Participants were not screened, with the only requirement being possession of an Android smartphone and a Facebook account. 39 participants were recruited in total, of whom 22 completed, and 17 prematurely left the study. The majority of those who did not complete the

study installed the application to their mobile device but did not complete the registration and consent process. 9 participants were in the LowNo condition, 6 in the HighNo condition, and 7 in the HighYes condition. To reduce the impact of cultural differences in privacy expectations, we recruited all participants from the United Kingdom. Recruitment for the experiment positioned our system as a new commercial application, to closely align participant expectations with that of existing commercial applications.

Ethical considerations

Due to the sensitivity of the location data collected during the study, and the deception employed, we took care to ensure the experiment was conducted in an ethical manner. While participants were told that some personal information would be disclosed to advertisers, no such disclosure occurred, and location data were only transmitted to the system to allow the application to generate relevant adverts.

Data collection used our PRISONER framework for privacy-sensitive handling of social network data, which we introduced in Chapter 4. The study made limited use of Facebook, with the need to retrieve the names of the participant's Facebook friends to allow them to choose which would receive location disclosures, and the limited ability to publish these disclosures to Facebook. A policy was written that allowed the names of the participant's friends to be temporarily retrieved so this could be displayed to the participant, and the friends' user IDs were stored so that the visibility of disclosures could be restricted to them. This raised a design limitation of PRISONER's current policy language. It is possible to place constraints on the retrieval and storage of SNS data, but in this situation, a distinction between the storage of data for the execution of the experiment, and storage for the benefit of the researcher would be useful. For example, although we need to store the user IDs of friends to restrict the audience of location disclosures, the collection of these data is not relevant to our interests. The current policy language, however, would allow us to access

these data anyway. Although not developed for this experiment, this has highlighted the need to make our policy language more expressive, to allow data to be stored and sanitised securely without providing the researcher direct access to it, if such access is not necessary.

The deception within the experimental design was handled sensitively. As per best practice [12], all participants were informed of the deception after the study closed in an email providing their remuneration, explaining the motivation for the study, and participants were given the opportunity to ask any further questions about how the experiment was conducted.

Participants were told they would earn money for sharing their locations. Rather than provide the exact amount promised by the application, all participants were given an Amazon voucher of equal value at the end of the study, surpassing the value any participant accrued during normal use of the application. This strategy was employed due to ethical concerns about financially rewarding some participants more than others.

The experimental design and all recruitment materials were approved by our ethics committee.

6.1.7 Results

In total, our 22 participants completed 212 check-ins, abandoned 471, and the most active user checked in 15 times in one day. Before completing the decision heuristic, we examine the overall differences between our groups of participants to understand the effects of our feedback and incentive conditions.

Less feedback induces greater sharing

Figure 6.2 shows the proportion of completed check-ins and abandoned check-ins. We hypothesise that users who were offered more money would check-in

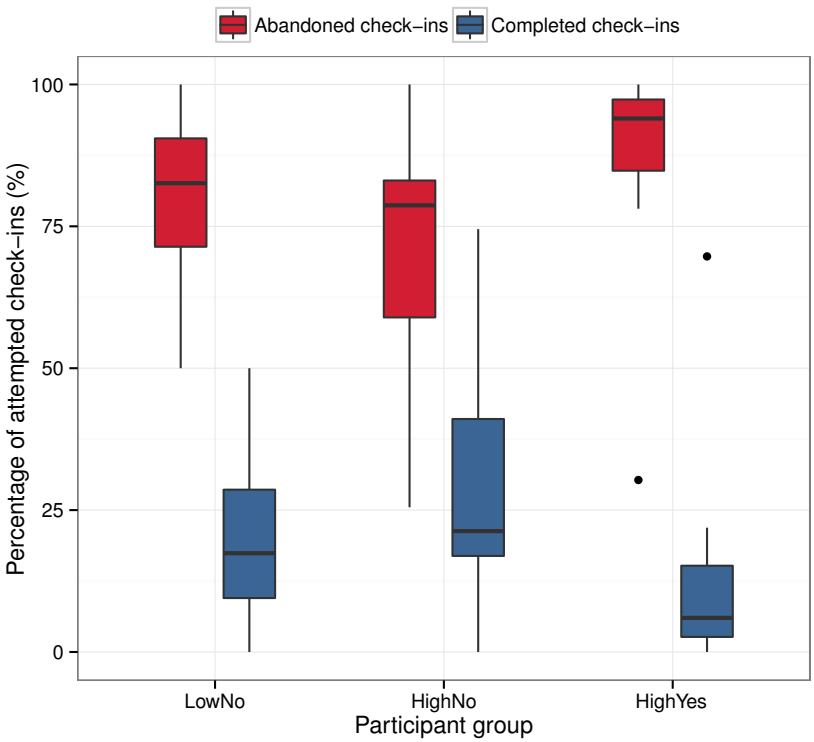


Figure 6.2: Participants who receive higher incentives but no additional feedback (HighNo) check in more often, whereas those given feedback about PII flows (HighYes) check in the least often.

more often (H1), unless more feedback was shown about how their personal information would be used (H3).

We found that participants in the HighNo condition exhibited the most variable behaviour. A number of participants performed more check-ins than lower-paid participants, but this was not consistent across the group. A one-way ANOVA shows no significant differences in the overall rate of check-ins $[F(2,13) = 0.807, p > 0.05]$, so we can not accept H1. This is not surprising, as users did not know that other participants were being offered different sums of money, therefore the distinction between 'low' and 'high' incentives is more subjective than the arbitrary values we selected for this study.

We did, however, note a reduction in disclosure rates for participants in the HighYes condition, where most users only completed less than 10% of check-ins. Behaviour was the most consistent within this group, and the higher variance among non-feedback groups indicates that the absence of such feedback generally induced more sharing. Although overall numbers of abandoned check-ins were not significantly different between groups $[F(2,13) = 0.596, p > 0.05]$, those who received more feedback abandoned more relative to the number of completed check-ins. We note that higher incentives without feedback is associated with greater variance in check-in rates, indicating an influence on behaviour. Some participants in all conditions did not complete any check-ins throughout the study, despite continuing to interact with the application.

Feedback may engender support for ILS

Figure 6.3 shows the results of the IUIPC survey before and after the study. Participants exposed to more feedback about the flow of PII during the study reported greater agreement on most dimensions in our post-brief survey, particularly "secondary use" (concern about information being used for reasons not originally sanctioned), "control" (loss of control leads to privacy violation), and "awareness" (of privacy practices), however this increase is not statistic-

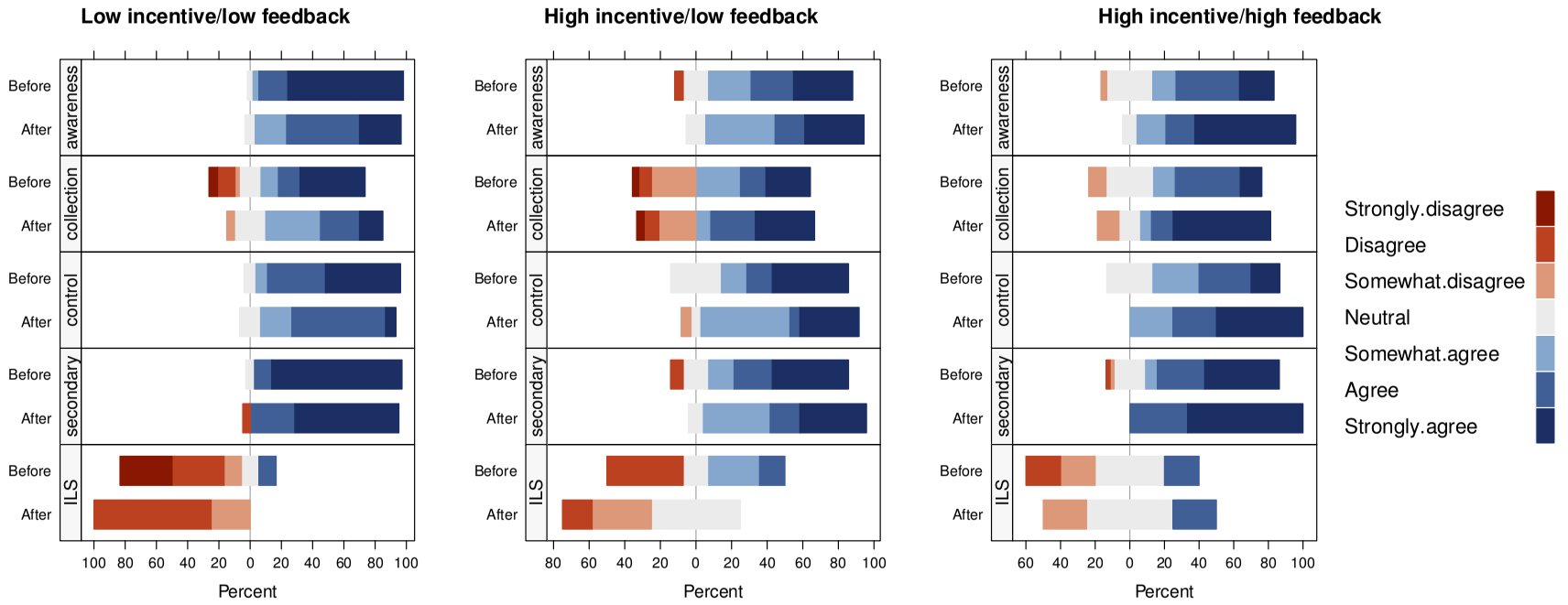


Figure 6.3: Diverging stacked bar chart showing how participants responded to a questionnaire about online privacy concerns before and after participating in the study. For each dimension tested, positive answers indicate increased concern or awareness about relevant privacy practices. Participants reported high concern and awareness about online privacy issues generally, and these were reinforced when more feedback about PII flows was provided by the application. Many believed, however, that companies are "entitled" to personal information in exchange for money, and comfort with this practice increased with use of a mobile advertising check-in application, as shown by the positive tendency on the ILS dimension.

ally significant. This is consistent with our expectations, as participants in this condition were provided with more information about how their information would be used, but were not provided with additional tools for managing this flow, other than not using the application. Participants in low feedback conditions did not significantly alter their responses in the debrief questionnaire.

We found that participants who received less feedback were less comfortable with ILS at the end of the study, although we saw increased comfort with ILS services for those in the higher feedback conditions, represented by the greater positive tendency for this question in Figure 6.3. This is surprising, as we expected those receiving a higher material benefit to feel more affinity for the practice as they agree that businesses are *entitled* to personal information in exchange for money. Participants in both high incentive conditions reported greater concern about secondary use of their PII, a fundamental aspect of ILS, while lower-paid participants' concern did not change as much. This suggests neither incentives nor feedback significantly impact people's concern about the use of their information.

Our examination of attitudes before and after the study reveals that concern and awareness of privacy issues generally increases through participating in the study. Interestingly, only participants who were provided feedback about the flows of their information believed that ILS was a more legitimate practice than at the start of the study, while confidence in ILS dropped for our other participants. This mirrors our finding that high feedback participants are much more comfortable with the disclosure of their PII. We believe this can be attributed to a combination of such participants feeling more empowered by the transparent explanation of how their information is used, while other participants, without the same *in situ* assurances, may have exhibited a priming effect from the study being bookended by our IUIPC questionnaire. Participants did not frequently report such concerns in our end-of-day questionnaires, lending further support to this theory.

These results give us some insight into the expectations of people before

adopting an ILS service, and their values after having experienced such a system. We observe paradoxical results, with participants' general privacy concerns hardened, but their attitudes towards ILS more relaxed. Contextual integrity suggests if a new process perturbs the values of an existing context, there is a risk of privacy breaches. This appears to manifest in our results, as people are reporting greater concern, yet appear to be placated by the introduction of a financial incentive. Failure to reconcile these behaviours risks people feeling compelled to make disclosures they might otherwise consider inappropriate. Our analysis of the IUIPC study indicates that there may be a relationship between exposure to feedback and people's privacy attitudes, but further work is needed to assess the significance of this effect.

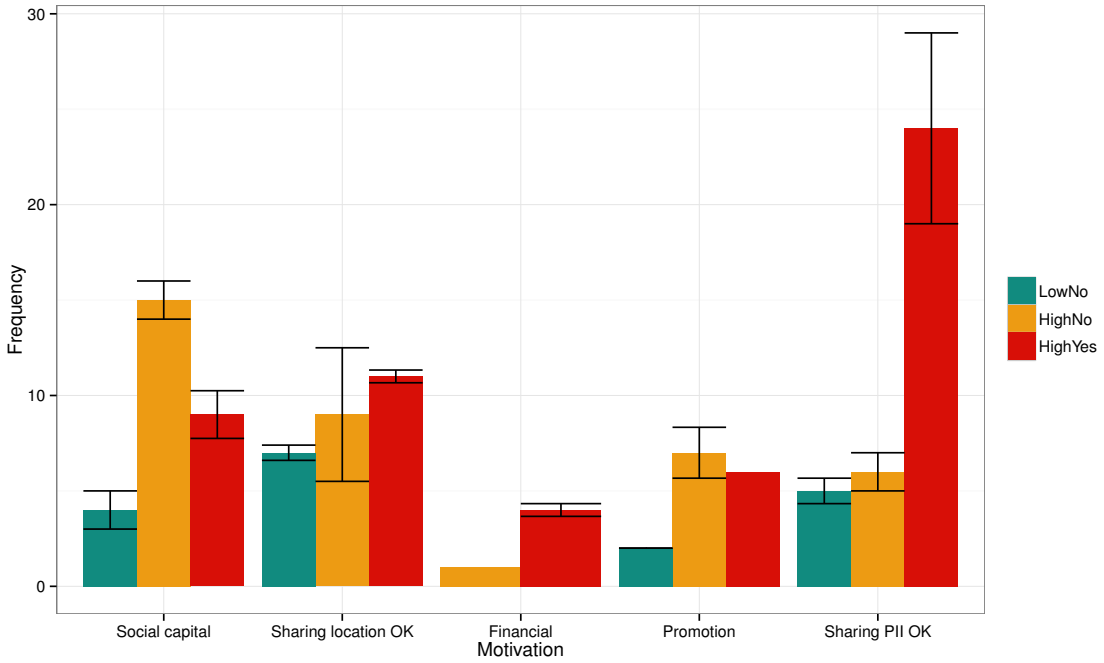


Figure 6.4: The frequency of self-reported motivations for checking in to our system. Consistent with other LBSNs, most disclosures were motivated by impression management and to build social capital. Participants who received higher incentives (HighNo and HighYes) would often explicitly promote the businesses they were visiting.

Higher incentives change check-in motivations

We now examine the motivations for these check-ins and abandoned check-ins. Understanding the motivations of people using a system provides insight into

the norms governing people's expectations for the appropriate flows of information. This is critical for identifying the transmission principles in the context, to see whether feedback and incentives has an effect on people's expectations.

As discussed in Chapter 2.2, disclosures in traditional LBSNs are often motivated by attempts to build social capital, and this was a recurring theme in our study too. When asked in end-of-day questionnaires to explain why they checked in to certain locations, participants in all conditions frequently reported "wanting my friends to see that I was there". This motivation did not often coincide with those directly pertaining to the introduction of incentives, such as "I don't mind sharing my personal information in exchange for cash" and "wanting to promote the business", which were commonly reported independently. Interestingly, in 69.2% of cases where social capital was a motivation, it did not coincide with any other motivations, suggesting people may be motivated on two largely independent fronts — appealing to their social network, and promoting businesses for money.

Participants who were offered a higher incentive were more likely to intentionally promote local businesses to their social network. We also find that when participants cited promotion as a motivation, it coincided with no other motivation in 57.1% of cases, suggesting that participants treated these two use cases somewhat independently. Lower-paid participants did not often exhibit this motivation, suggesting they may have not considered the value of the incentive sufficient to deliberately act as an advertising agent for the business, even if the actual exposure of their check-ins was the same.

Participants who received less feedback cited social capital-building motivations most often when disclosing their location, suggesting they treated the service in the same manner as other LBSNs. Similarly, the lowest-paid participants were unlikely to cite financial or promotional motivations, suggesting they also considered the application to be much like any other LBSNs, despite the additional PII flows our system introduces.

Identifying transmission principles

Examining motivations reveals that many of the same transmission principles that have been previously identified in location-sharing applications also manifest in our ILS service, specifically attempts to build social capital by being positively associated with certain locations.

Users who received less feedback cited social capital-building motivations most often when disclosing their location, suggesting they treated the service in the same manner as other location-sharing services. Similarly, the lowest-paid participants were unlikely to cite financial or promotional motivations, suggesting they also considered the application to be much like any other location-sharing system, despite the additional PII flows our system introduces.

6.1.8 Implications

Incentivised location sharing may violate contextual integrity

In order to make a *prima facie* judgement as to whether ILS risks violating contextual integrity, we consider the transmission principles which were absent from our initial analysis in Chapter 6.1.5, but identified in our user study. Based on these transmission principles, we make a *prima facie* judgement that ILS risks violating contextual integrity, supporting H2, which hypothesises that incentivised location sharing will perturb norms in the location-sharing context.

Our concern is that the prominence of traditional location-sharing motivations such as social capital building and impression management, particularly among users who received less feedback, suggests users are treating the application in the same manner as any other social-driven LBSN, despite sensitive information being disclosed to advertisers, and the possibility that their social

network will perceive incentivised check-ins to be of lower value. When users receive clear feedback about the use of their personal information at the point of disclosure, they make slightly fewer disclosures, and while disclosures are still often socially motivated, they often constitute a deliberate effort to advertise that business to their social network. Current commercial applications often do not deliver this level of feedback, burying information about the flows of personal information within unread privacy policies, and we argue that if such feedback is provided, participants are able to make more informed decisions about when and why their location will be disclosed. Users who receive more feedback also report slightly better awareness of online privacy issues.

We do not suggest that the introduction of new motivations for disclosure themselves constitute a breach of contextual integrity. In our user study, many participants managed the social and promotional aspects of the ILS context independent of each other. The relationship between the feedback within the application, and people's behaviour and wider attitudes towards online privacy, suggests the design of such an application can have a significant impact on people's relationship with technology. Users who received more feedback about the flow of PII were the most comfortable with the practice of ILS.

We did not see strong evidence that people felt pressured to disclose their location for money, or that a desire to earn money was motivating poor privacy-preserving behaviours. Therefore, in this prima facie assessment, we do not believe that ILS represents a major threat to personal autonomy. In our user study, one participant appeared to game the system by checking in to a wide variety of locations in an effort to maximise their income, but this behaviour was unusual.

Our results highlight interesting differences in how feedback and incentives affect behaviours and motivations in our application, but this preliminary study has some limitations. Our study is limited to a small number of participants and levels of incentives, and the impact of demgraphic variables has not been studied. This work indicates some interesting implications, to be investigated

and validated in further work.

People value incentives over privacy

Despite high levels of online privacy concern, our results suggest many people may be comfortable with the notion of disclosing their location and personally identifiable information for a cash incentive. When our participants were exposed to feedback about the flows of their personal information, their overall privacy concern increases, but they become even more comfortable with incentivised location sharing information flows, believing companies are "entitled" to their personal information if they are paid. This result is consistent with previous findings that people generally value money over their PII [43], and has implications for further study of ILS. We find that location disclosures were instigated by a mix of social and financial motivations that often did not coincide, which may suggest people are attempting to reconcile the context of a traditional LBSN with ILS. This is cause for concern as, in our system, all location disclosures reached the same audience, leading to potentially inappropriate disclosures to one's social network. Application designers can address this issue by avoiding the conflation of social and incentivised disclosures through distinct interfaces for each, and allowing people to choose different audiences for alternative types of disclosures.

Feedback does not discourage sharing

Participants who received more feedback about the flow of their personal information were more comfortable with the practice of ILS at the end of the study, without making a significantly reduced number of location disclosures compared to other participants. This is an important finding for application designers to note, as it contradicts any intuition that full disclosure about how people's information is used might dissuade users. Rather, our results suggest our participants may be empowered by understanding how their information is

used. While they are more discerning about when to share their information, they are also the most confident that they are aware of the privacy implications of using such services, and that the disclosure of PII is an appropriate outcome. Furthermore, the changes in motivations among our participants should be noted by designers, who ought to design services that satisfy both the social and financial reasons for their use. Where insufficient feedback was provided, we were concerned by the conflated motivations for sharing one's location, as participants struggled to reconcile the distinct use cases. Among participants who understood how their personal information was used, however, distinct social and financial thought processes were observed. Designers should provide *in situ* disclosures of how personal information is used, as our results suggests this satisfies people's privacy concerns, without severely affecting their willingness to use such services. In addition, incentivised and non-incentivised disclosures should be represented distinctly, to ensure people's motivations for making disclosures are aligned with the exposure of their information.

6.2 Summary

In this chapter, we have applied our framework for conducting SNS studies to illustrate how contextual integrity can be used to detect potential privacy issues in emerging SNSs. We note the following:

- In a user study with 22 smartphone owners, we determined that the addition of incentives to an LBSN may constitute a *prima facie* violation of contextual integrity.
- While monetisation does not change the frequency of location disclosures, people's motivations for sharing their location are changed, and their privacy concerns increase.
- Application designers can use feedback to show people how their information is used to mitigate concerns, and increase confidence in the practice

of ILS.

In the next chapter, we consolidate and discuss the implications of the findings we have made in this thesis, and outline directions for further work.

Chapter 7

Conclusion

Social network sites such as Facebook and Twitter, and location-based social networks such as Foursquare, have enabled the widespread collection, processing, and sharing of sensitive personal information. The nature of these data exposes people to various privacy risks. Meanwhile, the popularity of such services is encouraging researchers across many disciplines to study SNSs and the people who use them, raising ethical concerns about how people's data are handled, and whether participants understand the implications of consenting to such research. We have addressed the following thesis:

> Contextual integrity can be used to conduct reproducible and privacy-preserving experiments using social network sites, and can detect potential privacy violations in new services in order to mitigate their impact.

To test this, we considered the following questions:

RQ1: Is contextual integrity an appropriate framework for understanding and mitigating ethical concerns in SNS research?

RQ2: Can contextual integrity be used in the evaluation of SNSs to detect and mitigate their privacy impacts?

To address the first question, in Chapter 4 we proposed an architecture for conducting privacy-preserving SNS studies, informed by contextual integrity. In Chapter 5 we applied the architecture to investigate consent in SNS studies, and found that an approach to consent that aims to maintain contextual integrity was appropriate for many people.

To address the second question, in Chapter 6 we used contextual integrity to investigate an emerging form of LBSN, to detect whether the addition of financial incentives to the existing LBSN context could perturb entrenched norms and violate privacy. In a user study, we found that the lack of clarity about how people's information might be used by advertisers may violate contextual integrity, and allowed us to identify best practices that could mitigate the impact of this in future applications.

7.1 Contributions

In Chapter 4, we introduced a framework to conduct privacy-preserving and reproducible SNS experiments. Its design was informed by the need to uphold the contextual integrity of participants, and to support the encoding and reporting of key methodological details, which in Chapter 3 we found was often absent in the state of the art. We demonstrated how our architecture met these requirements by recreating a recent study from the literature, producing a shareable protocol for the data-collecting practices of the study, and minimising the amount of data available to researchers who implement the protocol of that study.

In Chapter 5, we investigated a significant methodological challenge: the acquisition of informed consent. Leveraging contextual integrity, we developed a new method of acquiring consent that aims to reduce the burden of asking people lots of questions about their willingness to share data, while still capturing their intent. In a user study we found that our method performs very well for a subset of the population who conform to social norms. The method is able

to quickly detect norm conformity and switch to an explicit form of gathering sustained consent if someone does not conform to norms.

Finally, in Chapter 6, we used contextual integrity to examine potential privacy impacts in an incentivised location sharing system. In a user study that examined the effects of incentives and feedback on people's behaviour, we found that some people over-share their location data when they don't understand that additional actors, such as advertisers, may receive personal information about them in exchange for a financial incentive. The lack of clarity about these new information flows risks violating contextual integrity, as it perturbs the entrenched norms that already existed in the location-sharing context, without clearly conveying the impact to users.

7.2 Discussion and further work

In Chapter 4, we proposed a framework for conducting ethical and reproducible SNS studies. We have demonstrated how we can use the framework to encode the data collection practices of a previous experiment, producing an artefact that can be archived or shared with other researchers. As we have argued, this is an important advance on the state of the art, and enables reproducibility of SNS experiment workflows, and supports the ethical conduct of studies in a way no previous tool has allowed. We do acknowledge, however, that this tool is not a "silver bullet" to resolving ethical and reproducibility challenges. PRISONER encodes the data collection practices of a study as a policy that can be communicated to other stakeholders, such as IRBs or participants, and archived or shared with other researchers. PRISONER can enforce this policy to restrict the data available to an experiment, and even if in the future a researcher does not have access to PRISONER, the human-readable nature of the policy allows the requirements it encodes to still be upheld. However, PRISONER is not able to execute an entire experiment with just a policy, which defines constraints for data collection, but does not encode when or why data are collected. The

sharing of the code and the environment needed to execute an experiment is a widely-acknowledged problem that is outwith the scope of this work, and in Chapter 3.3 we discussed recent work to improve this situation. These solutions, coupled with PRISONER, mitigate some of these challenges. For example, we can anticipate researchers sharing a virtual machine image that includes the source code for conducting the experiment and analyses, the operating system and other environmental dependencies needed to execute the code, including a PRISONER instance and experiment policies to enforce the constraints on data collection.

As we acknowledged in Chapter 4.1.7, the design and availability of APIs provided by SNSs is a significant barrier to reproducibility. We also consider that PRISONER could evolve to meet this challenge. In the future an archived experiment image, as we conceived of earlier, could include the original dataset collected from a particular SNS. Even if the source SNS is no longer available, or its API has significantly deviated from when the data collection took place, researchers could interrogate the dataset using the original policy and PRISONER instance to still apply the same constraints to the processing of that dataset. We also anticipate that a trusted third party could store a repository of datasets from SNS experiments, and expose a PRISONER API such that other researchers can conduct privacy-preserving replications of experiments without unfettered access to the sensitive underlying data.

In this thesis, we have discussed PRISONER's contribution to improving the state of reproducibility in SNS research. While our approach allows workflows to be encoded, shared and reproduced, further work is needed to address concerns about the sustainability and archiving of SNS research. PRISONER could be expanded to provide an environment that effectively emulates an SNS's API at a point in time. Therefore, even if an SNS has long been made obsolete, PRISONER could be used to provide an interface to the underlying data of a previous experiment, and enforce the same policy.

The design of the PRISONER framework has been informed by its grounding

in contextual integrity, and shaped by its subsequent application in a number of studies. This body of work has identified useful extensions to the framework. Our consent method discussed in Chapter 5 could be integrated into PRISONER itself. Abstracting researchers from having to provide a dataset of social norms and detect norm conformity, PRISONER could be developed to enforce this automatically, asking whether participants consent to data being shared when requests are made by experimental applications, and using norm conformity as a proxy for asking explicitly every time, where appropriate.

In Chapter 5, we proposed a method for acquiring consent in SNS studies based on contextual integrity. We found that a quarter of the population conform to social norms of willingness to share SNS data with researchers. For these people, our method could be used as a proxy for asking them on an ongoing basis about their willingness to share data, while still meeting their expectations of what would be collected. This result has implications for researchers conducting similar studies. With the academic community confronting the issue of how to acquire consent for "big data" experiments, we have demonstrated that a large proportion of the population can be served by such a method. In addition, the method can detect if it is not appropriate after a small number of interventions, automatically regressing to a traditional sustained consent approach where participants are asked to give explicit consent to data sharing.

This work has demonstrated the feasibility of such a method, but there are limitations to the current implementation. First, our method depends on the availability of existing social norms data that can be used to detect whether participants are norm-conformant or not. Our study showed that a dataset that was collected more than two years earlier and with a different population was robust to these differences and was still a useful measure of norms. We cannot, however, guarantee that this would be the case in all contexts, and researchers would need to evaluate the source of their social norms before determining whether it is appropriate for their study. To simplify this process, we anticipate that future implementations of the contextual integrity consent method could

collect and determine social norms while the study is running. As more participants take part, the precision of the inferred norms would improve. Further work is needed to assess the effectiveness of this, and the impact of bootstrapping where no norms data are available.

In Chapter 6, we used contextual integrity to identify potential privacy breaches in an incentivised location sharing system. Our application of the decision heuristic focused on the first six steps necessary to render a *prima facie* judgement as to whether contextual integrity may be violated. As this study focuses on the relationship between the individual user, the operator of the ILS service, and third-party advertisers, we believe this is an appropriate application of the heuristic, and was sufficient for us to make a judgement. The remaining three steps consider the wider moral, political, and social ramifications of the process. One outcome of our study questioned whether people might feel compelled to over-share their location out of financial need. While we did not see evidence of this, it may be appropriate to examine these wider implications to identify whether ILS could impact on other aspects of life.

We have shown that contextual integrity provides a framework that enables the holistic examination of social network sites and their users. From the development of tools to examine SNSs, through to the development of new methodologies, and identifying and mitigating potential breaches in actual systems, contextual integrity has provided a vocabulary for understanding the appropriateness of data-handling throughout the process.

Appendix A

Glossary

The following terms are explained throughout this thesis. For convenience, their definitions are summarised here.

- **Check-in**: A colloquialism that refers to sharing one's current location with their social network on an LBSN.
- **Context-relative informational norms**: Describes how information appropriately flows in a given context, in terms of the senders and recipients, the types of information, and the transmission principles that govern its transmission.
- **Context collapse**: An emergent phenomenon on many social network sites, where the real-life contextual distinctions between peers are removed, making it difficult to target communications to meaningful subsets of people.
- **Contextual integrity**: A theoretical framework for considering information privacy that suggests privacy violations occur when people's context-relative informational norms are violated.
- **Decision heuristic**: A diagnostic that allows the application of contextual integrity to determine whether a new process may violate privacy.

- **Facebook**: One of the most-used social network sites, established in the US in 2004.
- **Foursquare**: A popular location-based social network, which offers gamified rewards to encourage people to share their location.
- **Global Positioning System (GPS)**: A system that uses the position of satellites relative to a receiver to be used to determine that receiver's location.
- **Global System for Mobile Communications (GSM)**: A set of protocols used to deliver digital cellular communications, also known as 2G.
- **Human subjects research**: Research where there is any intervention or interaction with other people in order to gather information, or where information that may identify someone else is collected.
- **Incentivised location sharing (ILS)**: A practice on some location-based social networks where users are offered a financial or tangible reward to share their location, rather than a purely social motivation.
- **Informed consent**: A decision to participate in research taken by a competent individual who has received and understood all information, without being coerced or induced into consenting.
- **Institutional review board (IRB)**: An ethics committee that has oversight over human subjects research to ensure it meets institutional and regulatory requirements.
- **Location-based service (LBS)**: Any service that uses people's current location to deliver utility to users.
- **Location-based social network (LBSN)**: A form of social network site dedicated to the recording and share of one's location with their peers.
- **Mobile advertising**: The use of wireless data protocols and location-sensing technologies to deliver advertisements to users on their mobile devices.

- **Privacy calculus**: The risk-benefit analysis performed when deciding whether to disclose information, balancing the potential effects of being exposed against the utility that might be gained.
- **Purpose-driven sharing**: A form of location-sharing motivated by utilitarian purposes such as coordinating a meeting.
- **Quidco**: An incentivised location sharing service that offers small cash incentives to people who check-in to a business and share this with their social network.
- **Reproducibility**: The ability to repeat the procedures of an experiment, and to understand the provenance of results.
- **Secured consent**: Consent that is provided at a single point in time, such as at the start of an experiment.
- **Short message service (SMS)**: A standard for sending short text messages between mobile phones over GSM.
- **Smartphone**: A class of mobile phone often characterised by large displays, significant computational power, sensing capabilities and the ability to connect to the Internet.
- **Social capital**: The resources that can be extracted from a network of mutually acknowledged relationships, allowing people to use such relationships to advance their own interests.
- **Social-driven sharing**: A form of location-sharing motivated by social capital building and impression management, rather than the utility of the disclosed location.
- **Social network site (SNS)**: Web-based services that allow people to share content with a curated set of peers.
- **Sustained consent**: When consent is reacquired at regular intervals, ie. whenever new data are collected.

- **Transmission principles**: The rules that govern the appropriate flow of information, such as whether confidentiality or reciprocity is expected.
- **Twitter**: A popular social network site, distinguished from some competitors such as Facebook by its underlying directed graph that distinguishes between who a person follows, and is followed by.
- **Wireless Application Protocol (WAP)**: An early protocol for transmitting simplified Web content to mobile phones over GSM, superseded by the ability to directly access the Internet from modern smartphones.

Appendix B

Ethics approval

Two of the experiments discussed in this thesis involved human participation and were thus scrutinised and approved by the University of St Andrews' Teaching and Research Ethics Committee (UTREC). Confirmation of approval for both of these experiments, discussed in Chapters 5 and 6 respectively, are included on the following pages.

University of St Andrews
Scotland's first university – 1413

University Teaching and Research Ethics Committee Sub-committee

16th October 2014
Luke Hutton
School of Computer Science

Ethics Reference No: *Please quote this ref on all correspondence*	**CS11185**
Project Title:	**Investigating informed consent in social network research**
Researchers Name(s):	**Luke Hutton**
Supervisor(s):	**Dr Tristan Henderson**

Thank you for submitting your application which was considered at the Computer Science School Ethics Committee meeting on the 16th October 2014. The following documents were reviewed:

1. Ethical Application Form	29th September 2014
2. Participant Information Sheet	29th September 2014
3. Consent Form	29th September 2014
4. Debriefing Form	29th September 2014
5. Advertisements (*as necessary*)	29th September 2014

The University Teaching and Research Ethics Committee (UTREC) approve this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years. Projects, which have not commenced within two years of original approval, must be re-submitted to your School Ethics Committee.

You must inform your School Ethics Committee when the research has been completed. If you are unable to complete your research within the 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' https://www.st-andrews.ac.uk/utrec/guidelines/ are adhered to.

Yours sincerely

UNIVERSITY OF ST. ANDREWS SCHOOL OF COMPUTER SCIENCE

Convenor of the School Ethics Committee

Ccs Supervisor
School Ethics Committee

University of St Andrews

University Teaching and Research Ethics Committee
Sub-committee

29th April 2013
Luke Hutton
School of Computer Science

Ethics Reference No: *Please quote this ref on all correspondence*	**CS9799**
Project Title:	**Understanding behaviour in mobile advertising systems**
Researchers Name(s):	**Luke Hutton**
Supervisor(s):	**Tristan Henderson**

Thank you for submitting your application which was considered at the School of Computer Science School Ethics Committee meeting on the 7th March 2013. The following documents were reviewed:

1. Ethical Application Form 10/01/2013
2. Participant Information Sheet 10/01/2013
3. Consent Form 10/01/2013
4. Debriefing Form 10/01/2013

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years. Projects, which have not commenced within two years of original approval, must be re-submitted to your School Ethics Committee.

You must inform your School Ethics Committee when the research has been completed. If you are unable to complete your research within the 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' https://www.st-andrews.ac.uk/utrec/guidelines/ are adhered to.

Yours sincerely

pp Convenor of the School Ethics Committee

Ccs Supervisor
School Ethics Committee

Appendix C

IUIPC Questionnaire

The following questionnaire was presented to participants at the beginning and end of the experiment discussed in Chapter 6, based on the Internet Users' Information Privacy Concerns (IUIPC) scale [82]. Participants were asked to answer on a seven-point Likert scale, with responses ranging through "strongly disagree", "disagree", "somewhat disagree", "neutral", "somewhat agree", "agree", and "strongly agree".

1. Online privacy is really a matter of my right to exercise control and autonomy over decisions about how my information is collected, used and shared.
2. Control of my personal information lies at the heart of privacy.
3. I believe online privacy is invaded when control is lost or unwillingly reduced as a result of a marketing transaction.
4. Companies seeking information online should disclose the way data are collected, processed, and used.
5. A good online privacy policy should have a clear and conspicuous disclosure.
6. It is very important to me that I am aware and knowledgeable about how my personal information will be used.

7. It usually bothers me when online companies ask me for personal information.
8. When online companies ask me for personal information, I sometimes think twice before providing it.
9. It bothers me to give personal information to so many online companies.
10. I'm concerned that online companies are collecting too much personal information about me.
11. Online companies should not use personal information for any purpose unless it has been authorised by the individuals who provided information.
12. When people give personal information to an online company for some reason, the online company should never use the information for any other reason.
13. Online companies should never sell the personal information in their computer databases to other companies.
14. Online companies should never share personal information with other companies unless it has been authorized by the individuals who provided the information.
15. If I receive a financial incentive to share my location, the business operating that location is entitled to receive some personal information about me.

Bibliography

[1] Abdallah E. Ali, Sicco N. A. van Sas and Frank Nack. "Photographer Paths: Sequence Alignment of Geotagged Photos for Exploration-based Route Planning". In: *Proceedings of CSCW 2013*. San Antonio, TX, USA: ACM, 2013, pp. 985–994. doi: 10.1145/2441776.2441888.

[2] Sophia Alim. "An initial exploration of ethical research practices regarding automated data extraction from online social media user profiles". In: *First Monday* 19.7 (6th July 2014). doi: 10.5210/fm.v19i7.5382.

[3] Louise Barkhuus and Anind K. Dey. "Location-Based Services for Mobile Telephony: a Study of Users' Privacy Concerns." In: *INTERACT*. Vol. 3. 2003, pp. 702–712. url: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.10.527&rep=rep1&type=pdf.

[4] Patrick Barwise and Colin Strong. "Permission-based mobile advertising". In: *Journal of Interactive Marketing* 16.1 (Jan. 2002), pp. 14–24. doi: 10.1002/dir.10000.

[5] Lujo Bauer, Lorrie F. Cranor, Saranga Komanduri, Michelle L. Mazurek, Michael K. Reiter, Manya Sleeper and Blase Ur. "The Post Anachronism: The Temporal Dimension of Facebook Privacy". In: *Proceedings of the 12th ACM Workshop on Workshop on Privacy in the Electronic Society*. WPES '13. Berlin, Germany: ACM, 2013, pp. 1–12. doi: 10.1145/2517840.2517859.

[6] H. K. Beecher. "Ethics and clinical research". In: *Bulletin of the World Health Organization* 79.4 (1966), pp. 367–372. url: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2566401/.

[7] Jeremy Bentham. *An introduction to the principles of morals and legislation*. Elibron Classics, 2005. isbn: 9781402185649.

[8] Jessica W. Berg and Paul S. Appelbaum. *Informed consent legal theory and clinical practice*. Oxford University Press, 2001. isbn: 9780195126778.

[9] Katrin Borcea-Pfitzmann, Andreas Pfitzmann and Manuela Berg. "Privacy 3.0 := Data Minimization + User Control + Contextual Integrity". In: *it - Information Technology* 53.1 (Jan. 2011), pp. 34–40. doi: 10.1524/itit.2011.0622.

[10] Pierre Bourdieu. *The Forms of Capital*. Blackwell Publishers Ltd, 2008, pp. 280–291. doi: 10.1002/9780470755679.ch15.

[11] danah boyd and Nicole B. Ellison. "Social Network Sites: Definition, History, and Scholarship". In: *Journal of Computer-Mediated Communication* 13.1 (Oct. 2007), pp. 210–230. doi: 10.1111/j.1083-6101.2007.00393.x.

[12] British Psychological Society. *Code of Ethics and Conduct*. 2009. url: http://www.bps.org.uk/system/files/Public%20files/bps%5C_code%5C_of%5C_ethics%5C_2009.pdf.

[13] Ralf Caers, Tim De Feyter, Marijke De Couck, Talia Stough, Claudia Vigna and Cind Du Bois. "Facebook: A literature review". In: *New Media & Society* 15.6 (Sept. 2013), pp. 982–1002. doi: 10.1177/1461444813488061.

[14] Andre Charland and Brian Leroux. "Mobile Application Development: Web vs. Native". In: *Communications of the ACM* 54.5 (May 2011), pp. 49–53. doi: 10.1145/1941487.1941504.

[15] Andrew Charlesworth. "The ascent of smartphone". In: *Engineering & Technology* 4.3 (Feb. 2009), pp. 32–33. doi: 10.1049/et.2009.0306.

[16] Delphine Christin, Pablo Sánchez López, Andreas Reinhardt, Matthias Hollick and Michaela Kauer. "Share with strangers: Privacy bubbles as user-centered privacy control for mobile content sharing applications". In: *Information Security Technical Report* 17.3 (Feb. 2013), pp. 105–116. doi: 10.1016/j.istr.2012.10.004.

[17] Elizabeth Cohn and Elaine Larson. “Improving participant comprehension in the informed consent process.” In: *Journal of Nursing Scholarship* 39.3 (2007), pp. 273–280. url: http://view.ncbi.nlm.nih.gov/pubmed/17760802.

[18] Sunny Consolvo, Ian E. Smith, Tara Matthews, Anthony LaMarca, Jason Tabert and Pauline Powledge. “Location disclosure to social relations: why, when, & what people want to share”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’05. Portland, Oregon, USA: ACM, 2005, pp. 81–90. doi: 10.1145/1054972.1054985.

[19] Graham Cormode and Balachander Krishnamurthy. “Key differences between Web 1.0 and Web 2.0”. In: *First Monday* 13.6 (25th Apr. 2008). doi: 10.5210/fm.v13i6.2125.

[20] Council for International Organizations of Medical Sciences. “International ethical guidelines for biomedical research involving human subjects.” In: *Bulletin of medical ethics* 182 (Oct. 2002), pp. 17–23. url: http://view.ncbi.nlm.nih.gov/pubmed/14983848.

[21] Andrew Davison. “Automated capture of experiment context for easier reproducibility in computational research”. In: *Computing in Science and Engineering* 99.PrePrints (2012). doi: 10.1109/mcse.2012.41.

[22] Phillip Dawson. “Our anonymous online research participants are not always anonymous: Is this a problem?” In: *British Journal of Educational Technology* 45.3 (May 2014), pp. 428–437. doi: 10.1111/bjet.12144.

[23] Munmun De Choudhury, Scott Counts and Eric Horvitz. “Social Media As a Measurement Tool of Depression in Populations”. In: *Proceedings of WebSci 2013*. Paris, France: ACM, 2013, pp. 47–56. doi: 10.1145/2464464.2464480.

[24] Sebastian Denef, Petra S. Bayerl and Nico A. Kaptein. “Social Media and the Police: Tweeting Practices of British Police Forces During the August 2011 Riots”. In: *Proceedings of CHI 2013*. Paris, France: ACM, 2013, pp. 3471–3480. doi: 10.1145/2470654.2466477.

[25] Jana Diesner, TerrillL Frantz and KathleenM Carley. “Communication Networks from the Enron Email Corpus ”It’s Always About the People. Enron is no Different””. In: *Computational & Mathematical Organization Theory* 11.3 (14th Oct. 2005), pp. 201–228. doi: 10.1007/s10588-005-5377-0.

[26] David L. Donoho, Arian Maleki, Inam U. Rahman, Morteza Shahram and Victoria Stodden. “Reproducible Research in Computational Harmonic Analysis”. In: *Comput Sci Eng* 11.1 (Jan. 2009), pp. 8–18. doi: 10.1109/mcse.2009.15.

[27] Chris Drummond. “Replicability is not Reproducibility: Nor is it Good Science”. In: *Proc. of the Evaluation Methods for Machine Learning Workshop at the 26th ICML*. Montreal, QC, Canada, 2009. url: http://cogprints.org/7691/.

[28] European Parliament and the Council of the European Union. “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”. In: *Official Journal of the European Union* L 281 (1995), pp. 0031–0050. url: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:31995L0046%5C&%5C#38;from=en.

[29] Tobias Fiebig, Wouter Katz and Jeroen van Beek. *Grindr application security evaluation report*. 2013. url: https://www.os3.nl/_media/reports/grindr.pdf.

[30] Joshua Finnis, Nalin Saigal, Adriana Iamnitchi and Jay Ligatti. “A location-based policy-specification language for mobile devices”. In: *Pervasive and Mobile Computing* 8.3 (June 2012), pp. 402–414. doi: 10.1016/j.pmcj.2010.11.003.

[31] Susan T. Fiske and Robert M. Hauser. “Protecting human research participants in the age of big data”. In: *Proceedings of the National Academy of Sciences* 111.38 (23rd Sept. 2014), pp. 13675–13676. doi: 10.1073/pnas.1414626111.

[32] Juliana Freire, David Koop, Fernando Chirigati and Cláudio T. Silva. "Reproducibility Using VisTrails". In: *Implementing Reproducible Research*. Ed. by Victoria Stodden, Friedrich Leisch and Roger D. Peng. Chapman and Hall/CRC, 2014. url: http://osf.io/c3kv6/.

[33] Batya Friedman, Edward Felten and Lynette I. Millett. *Informed consent online: A conceptual model and design principles*. Tech. rep. CSE Technical Report, 2000. url: ftp://ftp.cs.washington.edu/tr/2000/12/UW-CSE-00-12-02.pdf.

[34] Batya Friedman, Peyina Lin and Jessica K. Miller. "Informed consent by design". In: *Security and Usability* (2005), pp. 495–521. url: http://hornbeam.cs.ucl.ac.uk/hcs/teaching/GA10/lec9extra/ch24friedman.pdf.

[35] Ruth E. Gavison. "Privacy and the Limits of Law". In: *Social Science Research Network Working Paper Series* (18th May 2012). url: http://ssrn.com/abstract=2060957.

[36] Ian P. Gent. *The Recomputation Manifesto*. 12th Apr. 2013. url: http://arxiv.org/abs/1304.3674.

[37] Emily Getty, Jessica Cobb, Meryl Gabeler, Christine Nelson, Ellis Weng and Jeffrey Hancock. "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*. Vancouver, BC, Canada, 2011, pp. 997–1000. doi: 10.1145/1978942.1979091.

[38] Carole A. Goble, Jiten Bhagat, Sergejs Aleksejevs, Don Cruickshank, Danius Michaelides, David Newman, Mark Borkum, Sean Bechhofer, Marco Roos, Peter Li and David De Roure. "myExperiment: a repository and social network for the sharing of bioinformatics workflows". In: *Nucleic Acids Research* 38.suppl 2 (1st July 2010), W677–W682. doi: 10.1093/nar/gkq429.

[39] Alice Goffman. *On the run : fugitive life in the American ghetto*. University of Chicago Press, 2014. isbn: 9780226136714.

[40] Scott A. Golder and Michael W. Macy. “Digital Footprints: Opportunities and Challenges for Online Social Research”. In: *Annu Rev Sociol* 40.1 (July 2014), pp. 129–152. doi: 10.1146/annurev-soc-071913-043145.

[41] Richard Gomer, M. C. Schraefel and Enrico Gerding. “Consenting Agents: Semi-autonomous Interactions for Ubiquitous Consent”. In: *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*. UbiComp ’14 Adjunct. Seattle, Washington: ACM, 2014, pp. 653–658. doi: 10.1145/2638728.2641682.

[42] F. S. Grodzinsky and H. T. Tavani. “Privacy in “the cloud”: applying Nissenbaum’s theory of contextual integrity”. In: *ACM SIGCAS Computers and Society* 41.1 (Oct. 2011), pp. 38–47. doi: 10.1145/2095266.2095270.

[43] Jens Grossklags and Alessandro Acquisti. “When 25 Cents is too much: An Experiment on Willingness-To-Sell and Willingness-To-Protect Personal Information”. In: *Proceedings of WEIS 2007*. 2007. url: http://weis2007.econinfosec.org/papers/66.pdf.

[44] Hamed Haddadi, Pan Hui, Tristan Henderson and Ian Brown. “Targeted Advertising on the Handset: Privacy and Security Challenges”. In: *Pervasive Advertising*. Ed. by Jörg Müller, Florian Alt and Daniel Michelis. Human-Computer Interaction Series. London: Springer London, Sept. 2011. Chap. 6, pp. 119–137. doi: 10.1007/978-0-85729-352-7_6.

[45] Stephanie Harriman and Jigisha Patel. “The ethics and editorial challenges of internet-based research”. In: *BMC Medicine* 12.1 (15th July 2014), pp. 124+. doi: 10.1186/s12916-014-0124-3.

[46] Georg W. Hegel. *Elements of the philosophy of right*. Cambridge University Press, 1991. isbn: 9780521348881.

[47] Joel M. Hektner, Jennifer A. Schmidt and Mihaly Csikszentmihalyi. *Experience sampling method: measuring the quality of everyday life*. Thousand Oaks, CA, USA: SAGE Publications, 2007. isbn: 1-4129-2557-6.

[48] Tristan Henderson and Luke Hutton. *Data for the paper "Towards reproducibility in online social network research"*. Aug. 2014. doi: 10.6084/m9.figshare.1153740.

[49] Kashmir Hill. "Facebook Added 'Research' To User Agreement 4 Months After Emotion Manipulation Study". 30th June 2014. url: http://onforb.es/15DKfGt.

[50] Roberto Hoyle, Sameer Patil, Dejanae White, Jerald Dawson, Paul Whalen and Apu Kapadia. "Attire: Conveying Information Exposure Through Avatar Apparel". In: *Proceedings of the 2013 Conference on Computer Supported Cooperative Work Companion*. CSCW '13. San Antonio, Texas, USA: ACM, 2013, pp. 19–22. doi: 10.1145/2441955.2441961.

[51] Gordon Hull, Heather Richter Lipford and Celine Latulipe. "Contextual gaps: privacy issues on Facebook". In: *Ethics and Information Technology* 13.4 (1st Dec. 2011), pp. 289–302. doi: 10.1007/s10676-010-9224-8.

[52] Luke Hutton and Tristan Henderson. "An architecture for ethical and privacy-sensitive social network experiments". In: *SIGMETRICS Performance Evaluation Review* 40.4 (Apr. 2013), pp. 90–95. doi: 10.1145/2479942.2479954.

[53] Luke Hutton and Tristan Henderson. ""I didn't sign up for this!": Informed consent in social network research". In: *Proceedings of the 9th International AAAI Conference on Web and Social Media (ICWSM)*. May 2015. url: http://tristan.host.cs.st-andrews.ac.uk/research/pubs/icwsm2015.pdf.

[54] Luke Hutton and Tristan Henderson. "Making social media research reproducible". In: *Proceedings of the ICWSM Workshop on Standards and Practices in Large-Scale Social Media Research*. Oxford, UK, May 2015.

[55] Luke Hutton and Tristan Henderson. "Towards reproducibility in online social network research". In: *IEEE Transactions on Emerging Topics in Computing* (2015). doi: 10.1109/tetc.2015.2458574.

[56] Luke Hutton, Tristan Henderson and Apu Kapadia. ""Here I Am, Now Pay Me!": Privacy Concerns in Incentivised Location-sharing Systems".

In: *Proceedings of the 2014 ACM Conference on Security and Privacy in Wireless & Mobile Networks*. WiSec '14. Oxford, UK: ACM, July 2014, pp. 81–86. doi: 10.1145/2627393.2627416.

[57] Eric Jain, Amos Bairoch, Severine Duvaud, Isabelle Phan, Nicole Redaschi, Baris Suzek, Maria Martin, Peter McGarvey and Elisabeth Gasteiger. "Infrastructure for the life sciences: design and implementation of the UniProt website". In: *BMC Bioinformatics* 10.1 (2009), pp. 136+. doi: 10.1186/1471-2105-10-136.

[58] Maritza Johnson, Serge Egelman and Steven M. Bellovin. "Facebook and Privacy: It's Complicated". In: *Proceedings of SOUPS 2012*. Washington, DC, USA: ACM, July 2012. doi: 10.1145/2335356.2335369.

[59] Elisabeth A. Jones and Joseph W. Janes. "Anonymity in a World of Digital Books: Google Books, Privacy, and the Freedom to Read". In: *Policy & Internet* 2.4 (Jan. 2010), pp. 42–74. doi: 10.2202/1944-2866.1072.

[60] Shelly Kagan. "Does Consequentialism Demand too Much? Recent Work on the Limits of Obligation". In: *Philosophy & Public Affairs* 13.3 (1984). url: http://www.jstor.org/stable/2265413.

[61] Christian Kahl, Stephen Crane, Markus Tschersich and Kai Rannenberg. "Privacy Respecting Targeted Advertising for Social Networks". In: *Information Security Theory and Practice. Security and Privacy of Mobile Devices in Wireless Communication*. Ed. by Claudio A. Ardagna and Jianying Zhou. Vol. 6633. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin / Heidelberg, 2011. Chap. 26, pp. 361–370. doi: 10.1007/978-3-642-21040-2_26.

[62] Immanuel Kant. *Groundwork for the metaphysics of morals*. Yale University Press, 2002. isbn: 9780300128154.

[63] Apu Kapadia, Tristan Henderson, Jeffrey Fielding and David Kotz. "Virtual Walls: Protecting Digital Privacy in Pervasive Environments". In: *Proceedings of the 5th International Conference on Pervasive Computing*. LNCS 4480. Toronto, Canada, May 2007, pp. 162–179. doi: 10.1007/978-3-540-72037-9_10.

[64] Jack Katz. "Toward a Natural History of Ethical Censorship". In: *Law & Society Review* 41.4 (1st Dec. 2007), pp. 797–810. doi: 10.1111/j.1540-5893.2007.00325.x.

[65] Jane Kaye, Edgar A. Whitley, David Lund, Michael Morrison, Harriet Teare and Karen Melham. "Dynamic consent: a patient interface for twenty-first century research networks". In: *European Journal of Human Genetics* 23.2 (7th May 2014), pp. 141–146. doi: 10.1038/ejhg.2014.71.

[66] Jennifer King, Airi Lampinen and Alex Smolen. "Privacy: Is There an App for That?" In: *Proceedings of SOUPS 2011*. Pittsburgh, PA, USA: ACM, 2011. doi: 10.1145/2078827.2078843.

[67] Nancy J. King and Pernille W. Jessen. "Profiling the mobile customer – Privacy concerns when behavioural advertisers target mobile phones – Part I". In: *Computer Law & Security Review* 26.5 (Sept. 2010), pp. 455–478. doi: 10.1016/j.clsr.2010.07.001.

[68] Bart Knijenburg, Alfred Kobsa and Hongxia Jin. "Preference-based Location Sharing: Are More Privacy Options Really Better?" In: *Proceedings of CHI 2013*. Paris, France, 2013. url: http://www.ics.uci.edu/%5C~%7B%7Dkobsa/papers/2013-CHI-kobsa.pdf.

[69] Ksenia Koroleva, Hanna Krasnova, Natasha Veltri and Oliver Günther. "It's all about networking! Empirical investigation of social capital formation on social network sites". In: *Proceedings of ICIS 2011*. Dec. 2011. url: http://aisel.aisnet.org/icis2011/proceedings/onlinecommunity/24.

[70] Adam D. I. Kramer, Jamie E. Guillory and Jeffrey T. Hancock. "Experimental evidence of massive-scale emotional contagion through social networks". In: *Proceedings of the National Academy of Sciences* 111.24 (17th June 2014), pp. 8788–8790. doi: 10.1073/pnas.1320040111.

[71] Hanna Krasnova, Natasha F Veltri and Oliver Günther. "Self-disclosure and Privacy Calculus on Social Networking Sites: The Role of Culture". In: *Business & Information Systems Engineering* 4.3 (2012), pp. 127–135. doi: 10.1007/s12599-012-0216-6.

[72] Ponnurangam Kumaraguru and Lorrie F. Cranor. *Privacy Indexes: A Survey of Westin's Studies*. Tech. rep. CMU-ISRI-5-138. Pittsburgh, PA, USA: Institute for Software Research International, School of Computer Science, Carnegie Mellon University, Dec. 2005. url: http://www.cs.cmu.edu/%5C~%7B%7Dponguru/CMU-ISRI-05-138.pdf.

[73] Robert S. Laufer and Maxine Wolfe. "Privacy as a Concept and a Social Issue: A Multidimensional Developmental Theory". In: *Journal of Social Issues* 33.3 (1st July 1977), pp. 22–42. doi: 10.1111/j.1540-4560.1977.tb01880.x.

[74] Scott Lederer, Jennifer Mankoff and Anind K. Dey. "Who wants to know what when? Privacy preference determinants in ubiquitous computing". In: *CHI 2003: CHI 2003 extended abstracts on Human factors in computing systems*. Ft. Lauderdale, Florida, USA: ACM, 2003, pp. 724–725. doi: 10.1145/765891.765952.

[75] Heather R. Lipford, Gordon Hull, Celine Latulipe, Andrew Besmer and Jason Watson. "Visible Flows: Contextual Integrity and the Design of Privacy Mechanisms on Social Network Sites". In: *2009 International Conference on Computational Science and Engineering*. Vol. 4. Vancouver, BC, Canada: IEEE, Aug. 2009, pp. 985–989. doi: 10.1109/cse.2009.241.

[76] Yabing Liu, Krishna Gummadi, Balanchander Krishnamurthy and Alan Mislove. "Analyzing Facebook privacy settings: User expectations vs. reality". In: *Proceedings of the 11th ACM/USENIX Internet Measurement Conference (IMC'11)*. Nov. 2011. url: http://www2.research.att.com/%5C~%7B%7Dbala/papers/imc11.pdf.

[77] Louis Harris & Associates and Alan F. Westin. *E-Commerce and Privacy: What Net Users Want*. Hackensack, NJ, USA, June 1998. url: http://www.privacyexchange.org/survey/surveys/ecommsum.html.

[78] Ardie Lubin. "Replicability as a publication criterion". In: *American Psychologist* 12.8 (Aug. 1957), pp. 519–520. doi: 10.1037/h0039746.

[79] Ewa Luger, Stuart Moran and Tom Rodden. "Consent for all: revealing the hidden complexity of terms and conditions". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '13. Paris, France: ACM, 2013, pp. 2687–2696. doi: 10.1145/2470654.2481371.

[80] Ewa Luger and Tom Rodden. "An Informed View on Consent for UbiComp". In: *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. UbiComp '13. Zurich, Switzerland: ACM, 2013, pp. 529–538. doi: 10.1145/2493432.2493446.

[81] Michelle Madejski, Maritza Johnson and Steven M. Bellovin. "A Study of Privacy Settings Errors in an Online Social Network". In: *Proceedings of SESOC 2012*. Lugano, Switzerland, Mar. 2012. url: http://www.cs.columbia.edu/~smb/papers/fb-violations-sesoc.pdf.

[82] Naresh K. Malhotra, Sung S. Kim and James Agarwal. "Internet Users' Information Privacy Concerns (IUIPC): The Construct, the Scale, and a Causal Model". In: *Information Systems Research* 15.4 (Dec. 2004), pp. 336–355. doi: 10.1287/isre.1040.0032.

[83] Ben Marder, Adam Joinson and Avi Shankar. "Every Post You Make, Every Pic You Take, I'll Be Watching You: Behind Social Spheres on Facebook". In: *Proceedings of HICSS 2012*. Maui, HI, USA: IEEE, Jan. 2012, pp. 859–868. doi: 10.1109/HICSS.2012.12.

[84] Annette Markham and Elizabeth Buchanan. *Ethical Decision-Making and Internet Research*. Tech. rep. Dec. 2012. url: http://aoir.org/reports/ethics2.pdf.

[85] Alice Marwick and danah boyd. "I Tweet Honestly, I Tweet Passionately: Twitter Users, Context Collapse, and the Imagined Audience". In: *New Media and Society* (2010). url: http://nms.sagepub.com/content/early/2010/06/22/1461444810365313.full.pdf+html.

[86] Erika McCallister. *Guide to protecting the confidentiality of personally identifiable information*. Diane Publishing, 2010. url: http://csrc.nist.gov/publications/nistpubs/800-122/sp800-122.pdf.

[87] Heidi McKee and James E. Porter. “The Ethics of Digital Writing Research: A Rhetorical Approach”. In: *College Composition and Communication* 59.4 (June 2008), pp. 711–749. url: http://www.inventio.us/ccc/2008/06/heidi-mckee-and-james-e-porter.html.

[88] Sam McNeilly, Luke Hutton and Tristan Henderson. “Understanding ethical concerns in social media privacy studies”. In: *Proceedings of ACM CSCW Workshop on Measuring Networked Social Privacy: Qualitative & Quantitative Approaches*. San Antonio, TX, USA, Feb. 2013. url: http://tristan.host.cs.st-andrews.ac.uk/pubs/mnsp2013.pdf.

[89] Claire Cain Miller. “Take a Step Closer for an Invitation to Shop”. In: *New York Times* (Feb. 2010), B4. url: http://www.nytimes.com/2010/02/23/business/media/23adco.html.

[90] Md Moniruzzaman and Ken Barker. “Redeem with Privacy (RWP): Privacy Protecting Framework for Geo-social Commerce”. In: *Proceedings of the 12th ACM Workshop on Privacy in the Electronic Society*. Berlin, Germany: ACM, 2013, pp. 189–200. doi: 10.1145/2517840.2517858.

[91] Stuart Moran, Ewa Luger and Tom Rodden. “Exploring Patterns as a Framework for Embedding Consent Mechanisms in Human-Agent Collectives”. In: *Active Media Technology*. Ed. by Dominik Ślzak, Gerald Schaefer, SonT Vuong and Yoo-Sung Kim. Vol. 8610. Lecture Notes in Computer Science. Springer International Publishing, 2014, pp. 475–486. doi: 10.1007/978-3-319-09912-5_40.

[92] Alistair Morrison, Donald McMillan and Matthew Chalmers. “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*. NordiCHI ’14. Helsinki, Finland: ACM, 2014, pp. 471–480. doi: 10.1145/2639189.2639239.

[93] Fred Morstatter, Jürgen Pfeffer, Huan Liu and Kathleen M. Carley. “Is the Sample Good Enough? Comparing Data from Twitter’s Streaming API with Twitter’s Firehose”. In: *Proceedings of ICWSM 2013*. 2013. url: http://www.aaai.org/ocs/index.php/ICWSM/ICWSM13/paper/view/6071.

[94] Matthew T. Mullarkey. "Socially Immature Organizations: A Typology of Social Networking Systems [SNS] with Organizations As Users [OAU]". In: *Proceedings of CSCW 2012*. Seattle, WA, USA: ACM, 2012, pp. 281–292. doi: 10.1145/2141512.2141604.

[95] Cosmin Munteanu, Heather Molyneaux, Wendy Moncur, Mario Romero, Susan O'Donnell and John Vines. "Situational Ethics: Re-thinking Approaches to Formal Ethics Requirements for Human-Computer Interaction". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Seoul, South Korea, 2015. doi: 10.1145/2702123.2702481.

[96] Frank Nagle and Lisa Singh. "Can Friends Be Trusted? Exploring Privacy in Online Social Networks". In: *2009 International Conference on Advances in Social Network Analysis and Mining*. Athens, Greece: IEEE, 22nd July 2009, pp. 312–315. doi: 10.1109/asonam.2009.61.

[97] Arvind Narayanan and Vitaly Shmatikov. "De-anonymizing Social Networks". In: *Proceedings of 30th IEEE Symposium on Security and Privacy*. Oakland, CA, USA: IEEE, May 2009, pp. 173–187. doi: 10.1109/sp.2009.22.

[98] Fabian Neuhaus and Timothy Webmoor. "Agile Ethics for Massified Research and Visualization". In: *Information, Communication & Society* 15.1 (Feb. 2012), pp. 43–65. doi: 10.1080/1369118x.2011.616519.

[99] Leon Neyfakh. "The Ethics of Ethnography". In: *Slate* (18th June 2015). url: http://www.slate.com/articles/news%5C_and%5C_politics/crime/2015/06/alice%5C_goffman%5C_s%5C_on%5C_the%5C_run%5C_is%5C_the%5C_sociologist%5C_to%5C_blame%5C_for%5C_the%5C_inconsistencies.single.html.

[100] Helen Nissenbaum. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford, CA, USA: Stanford Law Books, 2009. isbn: 0804752362.

[101] Ofcom. *Communications Market Report 2014*. Aug. 2014. url: http://stakeholders.ofcom.org.uk/binaries/research/cmr/cmr14/2014_UK_CMR.pdf.

[102] Ofcom. *The Consumer Experience 2014*. Jan. 2015. url: http://stakeholders.ofcom.org.uk/binaries/research/consumer-experience/tce-14/TCE14_research_report.pdf.

[103] Tim Paek and Bo J. Hsu. "Sampling Representative Phrase Sets for Text Entry Experiments: A Procedure and Public Resource". In: *Proceedings of CHI 2011*. Vancouver, BC, Canada: ACM, 2011, pp. 2477–2480. doi: 10.1145/1978942.1979304.

[104] Xinru Page, Alfred Kobsa and Bart P. Knijnenburg. "Don't Disturb My Circles! Boundary Preservation Is at the Center of Location-Sharing Concerns". In: *Sixth International AAAI Conference on Weblogs and Social Media*. Dublin, Ireland, June 2012. url: http://www.aaai.org/ocs/index.php/ICWSM/ICWSM12/paper/view/4679.

[105] Heather A. Piwowar, Roger S. Day and Douglas B. Fridsma. "Sharing detailed research data is associated with increased citation rate." In: *PLoS ONE* 2.3 (21st Mar. 2007), e308+. doi: 10.1371/journal.pone.0000308.

[106] Sue J. Powell, Conor Linehan, Laura Daley, Andrew Garbett and Shaun Lawson. ""I Can't Get No Sleep": Discussing #Insomnia on Twitter". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Austin, Texas, USA, 2012, pp. 1501–1510. doi: 10.1145/2207676.2208612.

[107] R. Procter, M. Poschen, W. Lin Y-, C. Goble and D. De Roure. "Issues for the Sharing and Re-Use of Scientific Workflows". In: *Proceedings of 5th International Conference on e-Social Science*. 2009. url: http://www.escholar.manchester.ac.uk/uk-ac-man-scw:117546.

[108] Lee Rainie, Janna Anderson and Jennifer Connolly. *Cyber Attacks Likely to Increase*. Tech. rep. Pew Research Center, Oct. 2014. url: http://www.pewinternet.org/files/2014/10/PI%5C_FutureofCyberattacks%5C_102914%5C_pdf.pdf.

[109] David B. Resnik. "What is Ethics in Research & Why is it Important?" May 2011. url: http://www.niehs.nih.gov/research/resources/bioethics/whatis.cfm.

[110] Mattias Rost, Louise Barkhuus, Henriette Cramer and Barry Brown. "Representation and Communication: Challenges in Interpreting Large Social Media Datasets". In: *Proc. CSCW*. San Antonio, Texas, USA: ACM, 2013, pp. 357–362. doi: 10.1145/2441776.2441817.

[111] Mark A. Rothstein and Abigail B. Shoben. "Does Consent Bias Research?" In: *The American Journal of Bioethics* 13.4 (2013), pp. 27–37. doi: 10.1080/15265161.2013.767955.

[112] Arno Scharl, Astrid Dickinger and Jamie Murphy. "Diffusion and success factors of mobile marketing". In: *Electronic Commerce Research and Applications* 4.2 (June 2005), pp. 159–173. doi: 10.1016/j.elerap.2004.10.006.

[113] Pan Shi, Heng Xu and Yunan Chen. "Using contextual integrity to examine interpersonal information boundary on social network sites". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI 2013. Paris, France: ACM, 2013, pp. 35–38. doi: 10.1145/2470654.2470660.

[114] Pavel Skvortsov, Frank Dürr and Kurt Rothermel. "Map-Aware Position Sharing for Location Privacy in Non-trusted Systems". In: *Pervasive Computing*. Ed. by Judy Kay, Paul Lukowicz, Hideyuki Tokuda, Patrick Olivier and Antonio Krüger. Vol. 7319. Lecture Notes in Computer Science. Newcastle, UK: Springer Berlin Heidelberg, June 2012. Chap. 24, pp. 388–405. doi: 10.1007/978-3-642-31205-2_24.

[115] Manya Sleeper, Rebecca Balebako, Sauvik Das, Amber L. McConahy, Jason Wiese and Lorrie F. Cranor. "The Post That Wasn't: Exploring Self-censorship on Facebook". In: *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*. CSCW 2013. San Antonio, Texas, USA: ACM, 2013, pp. 793–802. doi: 10.1145/2441776.2441865.

[116] Lauren Solberg. "Data Mining on Facebook: A Free Space for Researchers or an IRB Nightmare?" In: *University of Illinois Journal of Law, Technology & Policy* 2010.2 (2010). url: http://www.jltp.uiuc.edu/works/Solberg.htm.

[117] Daniel J. Solove. "A Taxonomy of Privacy". In: *University of Pennsylvania Law Review* 154.3 (Jan. 2006), pp. 477–560. doi: 10.2307/40041279.

[118] Erica D. Spiegler, Christian Hildebrand and Florian Michahelles. "Social Networks in Pervasive Advertising and Shopping". In: *Pervasive Advertising*. Ed. by Jörg Müller, Florian Alt and Daniel Michelis. London, UK: Springer, 2011. Chap. 10, pp. 207–225. doi: 10.1007/978-0-85729-352-7_10.

[119] Bernd C. Stahl. "Responsible research and innovation: The role of privacy in an emerging framework". In: *Science and Public Policy* (19th Sept. 2013), sct067+. doi: 10.1093/scipol/sct067.

[120] Bernd Carsten Stahl, Damian Okaibedi Eke and Christine Fidler. "Understanding the relevance of ethics reviews of ICT research in UK computing departments using dialectical hermeneutics". In: *Journal of Information, Communication and Ethics in Society* 13.1 (2015), pp. 28–38. doi: 10.1108/JICES-03-2014-0015.

[121] StatCounter. *StatCounter Global Stats*. Feb. 2015. url: http://gs.statcounter.com/#all-comparison-ww-monthly-201402-201502.

[122] Kristin S. Steinsbekk, Bjorn Kare Myskja and Berge Solberg. "Broad consent versus dynamic consent in biobank research: Is passive participation an ethical problem?;" in: *European Journal of Human Genetics* 21.9 (9th Jan. 2013), pp. 897–902. doi: 10.1038/ejhg.2012.282.

[123] Victoria Stodden. *The Scientific Method in Practice: Reproducibility in the Computational Sciences*. Tech. rep. 4773-10. MIT Sloan School of Management, 9th Feb. 2010. doi: 10.2139/ssrn.1550193.

[124] Fred Stutzman, Jessica Vitak, Nicole B. Ellison, Rebecca Gray, Cliff Lampe and H. John Heinz. "Privacy in Interaction: Exploring Disclosure and Social Capital in Facebook". In: *Proceedings of ICWSM 2012*. 2012. url: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.227.7574.

[125] Synack. *The Do's and Don'ts of Location Aware Apps; A Case Study*. Sept. 2014. url: https://www.synack.com/labs/projects/the-dos-and-donts-of-location-aware-apps-a-case-study/.

[126] Karen P. Tang, Jialiu Lin, Jason I. Hong, Daniel P. Siewiorek and Norman Sadeh. "Rethinking location sharing: exploring the implications of social-driven vs. purpose-driven location sharing". In: *Proceedings of the 12th ACM international conference on Ubiquitous computing*. Ubicomp 2010. Copenhagen, Denmark: ACM, 2010, pp. 85–94. doi: 10.1145/1864349.1864363.

[127] "The Belmont Report : Ethical Principles and Guidelines for the Protection of Human Subjects of Research". 1979. url: http://ohsr.od.nih.gov/guidelines/belmont.html.

[128] Paul A. Thompson and Andrew Burnett. "Reproducible Research". In: *CORE Issues in Professional and Research Ethics* 1.6 (2012). url: http://nationalethicscenter.org/content/article/175.

[129] Eran Toch, Justin Cranshaw, Paul H. Drielsma, Jay Springfield, Patrick G. Kelley, Lorrie Cranor, Jason Hong and Norman Sadeh. "Locaccino: a privacy-centric location sharing application". In: *Proceedings of the 12th ACM international conference adjunct papers on Ubiquitous computing*. Ubicomp 2010. Copenhagen, Denmark: ACM, 2010, pp. 381–382. doi: 10.1145/1864431.1864446.

[130] Eran Toch and Inbal Levi. "Locality and privacy in people-nearby applications". In: *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*. UbiComp 2013. Zurich, Switzerland: ACM, 2013, pp. 539–548. doi: 10.1145/2493432.2493485.

[131] Janice Y. Tsai, Patrick G. Kelley, Lorrie F. Cranor and Norman Sadeh. "Location-Sharing Technologies: Privacy Risks and Controls". In: *TPRC 2009: Proceedings of the 37th Research Conference on Communication, Information, and Internet Policy*. Arlington, VA, USA, Sept. 2009. url: http://ssrn.com/abstract=1997782.

[132] Johan Ugander, Brian Karrer, Lars Backstrom and Cameron Marlow. *The Anatomy of the Facebook Social Graph*. 18th Nov. 2011. url: http://arxiv.org/abs/1111.4503.

[133] Christopher Ververidis and G. Polyzos. “Mobile marketing using a location based service”. In: *Proceedings of the First International Conference on Mobile Business, Athens, Greece*. 2002. url: http://www.123seminarsonly.com/Seminar-Reports/020/6550915-Mobile-Marketing-Using-a-Location-Based-Service.pdf.

[134] W3C. *Activity Streams 2.0*. Jan. 2015. url: http://www.w3.org/TR/activitystreams-core/.

[135] Joseph B. Walther. “Research ethics in Internet-enabled research: Human subjects issues and methodological myopia”. In: *Ethics and Information Technology* 4.3 (1st Sept. 2002), pp. 205–216. doi: 10.1023/a:1021368426115.

[136] Yang Wang, Gregory Norcie, Saranga Komanduri, Alessandro Acquisti, Pedro G. Leon and Lorrie 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. Pittsburgh, Pennsylvania: ACM, 2011. doi: 10.1145/2078827.2078841.

[137] Samuel D. Warren and Louis D. Brandeis. “The Right to Privacy”. In: *Harvard Law Review* 4.5 (1890). url: http://www.jstor.org/stable/1321160.

[138] Alan F. Westin. *Privacy and freedom*. New York: Atheneum, 1970. isbn: 9780370013251.

[139] E. A. Whitley, N. Kanellopoulou and J. Kaye. “Consent and Research Governance in Biobanks: Evidence from Focus Groups with Medical Researchers”. In: *Public Health Genomics* 15.5 (2012), pp. 232–242. doi: 10.1159/000336544.

[140] Wikipedia. *List of social networking websites — Wikipedia, The Free Encyclopedia*. [Online; accessed 25-March-2015]. 2015. url: http://en.wikipedia.org/w/index.php?title=List_of_social_networking_websites&oldid=652834760.

[141] Robert E. Wilson, Samuel D. Gosling and Lindsay T. Graham. “A Review of Facebook Research in the Social Sciences”. In: *Perspectives on Psychological Science* 7.3 (May 2012), pp. 203–220. doi: 10.1177/1745691612442904.

[142] Shomir Wilson, Justin Cranshaw, Norman Sadeh, Alessandro Acquisti, Lorrie F. Cranor, Jay Springfield, Sae Y. Jeong and Arun 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. Zurich, Switzerland: ACM, 2013, pp. 549–558. doi: 10.1145/2493432.2493436.

[143] Katherine Wolstencroft, Robert Haines, Donal Fellows, Alan Williams, David Withers, Stuart Owen, Stian Soiland-Reyes, Ian Dunlop, Aleksandra Nenadic, Paul Fisher, Jiten Bhagat, Khalid Belhajjame, Finn Bacall, Alex Hardisty, Abraham Nieva de la Hidalga, Maria P. Balcazar Vargas, Shoaib Sufi and Carole Goble. "The Taverna workflow suite: designing and executing workflows of Web Services on the desktop, web or in the cloud". In: *Nucleic Acids Research* 41.Web Server issue (2nd May 2013), gkt328–W561. doi: 10.1093/nar/gkt328.

[144] World Medical Association. "Declaration of Helsinki: Ethical principles for medical research involving human subjects". In: *JAMA* 310.20 (2013), pp. 2191–2194. doi: 10.1001/jama.2013.281053.

[145] Heng Xu and Hock-Hai Teo. "Privacy Considerations in Location-Based Advertising". In: *Designing Ubiquitous Information Environments: Socio-Technical Issues and Challenges*. Ed. by Carsten Sørensen, Youngjin Yoo, Kalle Lyytinen and Janice I. DeGross. Vol. 185. IFIP International Federation for Information Processing. New York: Springer Boston, 2005. Chap. 8, pp. 71–90. doi: 10.1007/0-387-28918-6_8.

[146] Yu Zheng. "Location-Based Social Networks: Users". In: *Computing with Spatial Trajectories*. Ed. by Yu Zheng and Xiaofang Zhou. Springer New York, 2011, pp. 243–276. doi: 10.1007/978-1-4614-1629-6_8.

[147] Michael Zimmer. ""But the data is already public": on the ethics of research in Facebook". In: *Ethics and Information Technology* 12.4 (Dec. 2010), pp. 313–325. doi: 10.1007/s10676-010-9227-5.