Limits of Individual Consent and Models of Distributed Consent in Online Social Networks

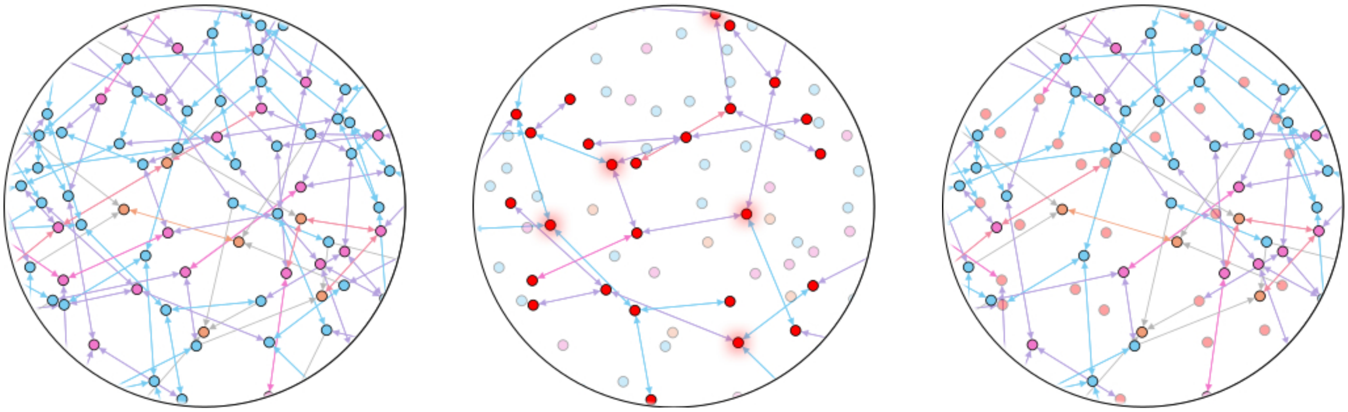
Juniper Lovato
Vermont Complex Systems Center,
University of Vermont
Burlington, Vermont, USA
juniper.lovato@uvm.edu

Antoine Allard
Département de physique, de génie
physique et d'optique, Université
Laval
Québec, Canada
antoine.allard@phy.ulaval.ca

Randall Harp
Department of Philosophy, University
of Vermont
Burlington, Vermont, USA
randall.harp@uvm.edu

Jeremiah Onaolapo
Department of Computer Science,
University of Vermont
Burlington, Vermont, USA
Jeremiah.Onaolapo@uvm.edu

Laurent Hébert-Dufresne
Department of Computer Science,
University of Vermont
Burlington, Vermont, USA
laurent.hebert-dufresne@uvm.edu



a. Information network **b.** Observed network **c.** Protected network

Figure 1: (a.) Information flow across a network with our basic implementation of distributed consent. Blue nodes have the lowest security settings and are susceptible to surveillance from third-party applications or websites. Purple nodes have stricter security settings but share their posts and data with all their neighbors. Orange nodes follow a distributed consent model and only share their data with purple nodes or other orange nodes. (b.) The same network where a third party directly observes a handful of low-security accounts is highlighted in red with shading. All nodes sharing their data with directly observed accounts are also de facto observed and shown in red. Nodes at a distance $L > 1$ can also be observed if the third party leverages some statistical procedure, inferring data up to a distance of two from directly observed nodes. (c.) We show the remaining unobserved or protected subnetwork.

ABSTRACT

Personal data are not discrete in socially-networked digital environments. A user who consents to allow access to their profile can expose the personal data of their network connections to non-consented access. Therefore, the traditional consent model (informed and individual) is not appropriate in social networks where informed consent may not be possible for all users affected by data processing and where information is distributed across users. Here, we outline the adequacy of consent for data transactions. Informed by the shortcomings of individual consent, we introduce both a platform-specific model of "distributed consent" and a cross-platform model of a "consent passport." In both models, individuals and groups can coordinate by giving consent conditional on that of their network connections. We simulate the impact of these distributed consent models on the observability of social networks and find that low adoption would allow macroscopic subsets of networks to preserve their connectivity and privacy.

FAccT '22, June 21–24, 2022, Seoul, Republic of Korea
© 2022 Association for Computing Machinery.
ACM ISBN 978-1-4503-9352-2/22/06…$15.00
https://doi.org/10.1145/3531146.3534640

CCS CONCEPTS

• **Security and privacy** → **Social aspects of security and privacy**; **Privacy protections**; • **Computing methodologies** → *Model development and analysis*; • **Social and professional topics** → *Privacy policies.*

KEYWORDS

Limits of consent, Data privacy, Data ethics, Social networks, Privacy models

ACM Reference Format:

Juniper Lovato, Antoine Allard, Randall Harp, Jeremiah Onaolapo, and Laurent Hébert-Dufresne. 2022. Limits of Individual Consent and Models of Distributed Consent in Online Social Networks. In *2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22), June 21–24, 2022, Seoul, Republic of Korea.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3531146.3534640

1 INTRODUCTION

One key focus of the burgeoning field of data ethics concerns how big data and networked systems challenge classic notions of privacy, bias, transparency, and consent [33]. In particular, the traditional privacy model, which relies on individual self-determination and individual consent, we argue, is no longer appropriate for the digital age. First, the traditional privacy model requires that consent be *informed*, which may not be possible in the context of large data sets and complicated technologies. Second, the traditional privacy model presumes *individual* control over personal information, even though the flow of information in networked systems precludes anyone from having such control over any piece of data. While the modern information environment shows both conditions as problematic, and while we briefly discuss the information condition, we focus most of our attention on the individuality condition.

Individual consent (by which we mean requiring that individual end-users consent in order for some action or outcome to be permissible) has many limitations—notably, we live in a highly networked [37] and advanced technological society, where digital decisions and actions are interconnected and affect not just ourselves but our digital community as a whole. In a digital age, individual consent is flawed [11] and ineffectual when protected class data and social profiles can be easily inferred via our social networks [6, 10, 24, 38, 52]. The individual consent model works most effectively in a physical space with accepted boundary norms [48, 62], linear contacts between two discrete parties, and no externalities. This, however, does not translate well to a digital realm where personal data boundaries are fuzzy and interwoven. The current over-reliance on individual consent online has also led to a negative externality of less legitimate consent due to consent desensitization, in part, because users are now faced with a deluge of consent requests [53]. Thus, a new approach to data privacy and consent in this context is needed.

A new data privacy model will need to consider several factors: the networked virtual space that we occupy; integration of group consent; and a mechanism for distributed moral responsibility when data privacy is breached or data are processed, combined, or manipulated in unethical manners [22]. In this paper, we will focus on distributed consent in particular and evaluate, in a mathematical model, its potential to increase online social networks' general privacy. We aim to cover the latter data privacy concerns in future work. In addition, future work could explore the potential for early adopters of distributed consent to influence their network neighbors (e.g. via cascade effects towards a contagious taste for privacy).

1.1 A Critique on the Adequacy of Individual Consent for Data Transactions

We will call the means by which information is shared *personal data transactions*. Broadly speaking, a personal data transaction is any transaction in which one party gives or reveals personal information to another, so the category is vast; it includes every behavior and every speech act that imparts information of some kind to another.

We can narrow the broad category of personal data transactions in three ways for our purposes here. First, we are interested in those transactions for which the *primary purpose of the actions or transaction is the transfer of information*. We will not attempt to give a complete conceptual framework here, but we can provide some indication of what we mean. If *A* gives *B* money, and *B* gives *A* a shirt, information has been exchanged, but the primary purpose of *A*'s giving money, and *B*'s giving the shirt was not the information exchange—it was the exchange of goods. If *A* asks *B* how their day was and *B* replies "pretty good," information has been exchanged, but it is possible that the primary purpose was not the information exchange but rather the demonstration of caring or the strengthening of solidarity. So we will focus on data transactions where the primary purpose behind the transaction is the transfer of information itself—though it is important to know that the information might be valued because of further downstream uses of the information. If *A* asks *B* about where *B* grew up, what *B*'s mother's maiden name is, and the name of B's childhood pet, then *B*'s responses would be data transactions—even if *A*'s ultimate reason in asking the questions were to get access to *B*'s online banking accounts.

Second, we are interested in those transactions which are about *personal information.* As with the previous case, we cannot precisely conceptualize personal information here. The general idea is that personal information is information about an identifiable living person [43]. If *A* downloads all of the 1911 Eleventh Edition of the *Encyclopaedia Britannica* and gives it to *B*, that is a data transaction in the broad sense but is likely not a transfer of personal information in the sense that we are interested in here (assuming that B is not an entry in the encyclopedia). If *A* gives *B* information about *A*'s whereabouts over the past 24 hours or information about *A*'s food preferences, that is personal information. It is important to note that personal information need not be information about the person giving it to someone else; *A* can give *B* information about *C*'s food preferences or whereabouts, which would constitute a data transaction. Here again, the boundaries are difficult to layout precisely.

Where do the limits of personal data lie? Suppose *A* were to give *B* a copy of a biography of Barack Obama: A's giving *B* the biography is outside of the scope of data transactions that we are interested in because the personal information about Obama contained in the biography is presumed to be *public.* Data transactions are valuable insofar as some party is gaining something of value, and discrete information that is previously known is not valuable. Moreover, that suggests a third way to restrict the domain of data transactions that we are interested in: we are concerned only with data transactions that deal with *non-public information.*

So when we talk about personal data transactions, we are talking about exchanges of personal information between two parties where the exchange of information is the (or a) primary purpose of the exchange, where the information is personal, and where the information is non-public. Personal data transactions of this form make up a significant and increasing portion of our modern lives. The following examples are just two of many but highlight personal data exchanges that have the potential to be highly impactful on our moral lives.

As one example, consider our use of social media. When we make an account on social media, we potentially engage in three different kinds of personal data transactions. First, there are the data transactions between the person with the account (the "end-user") and the social media company or platform. Second, there are the data transactions between one end-user and other end users on the platform between whom there are some network connections. Third, there are data transactions between end-users and third-party auxiliaries (e.g., individuals, apps, bots) that receive or process data to enhance the end user's social network experience. As examples of these, consider Facebook: anyone creating a Facebook account shares with the site all of the information they intentionally put on the site (their profile information, their network connections/friends, their posts). There is also a host of information that they might not be aware they are putting on the site (their location, the amount of time they spend reading posts or watching ads). The Facebook user also builds networks of social connections (nodes) on the platform.

Selected information is shared among various nodes on the friendship network in accordance with the end user's preferences. End-user A might share a lot of information with friends B and C. User A is close friends with B and C and may want to share information such as profile data, location data, and the content of all of their posts. User A is an acquaintance with users D and E and may want data sharing to be limited to only some posts or only some photos, or no location information or information about A's friend network. The third kind of information flow is that which is shared with third parties. For example, third parties might be designers of games or quizzes that can be run on the Facebook platform, such that those third parties can then collect information Facebook possesses about the end-user playing the game. (Cambridge Analytica is an example of a third party who acquired information from end-users through games or apps on the Facebook platform; Facebook has since changed some of its policies on third-party data acquisition.) Third-party websites can also allow users to sign in with their Facebook accounts, and those third parties can then get some user information when people use their Facebook accounts to log in. While not every social media site offers the same options for how information is shared across those three modalities, having an account on any social media site entails sharing information in each of those three different ways.

1.2 A Theory of Consent

The fact that people have putatively consented to all of the personal data transactions that we identified above is supposed to be doing a lot of normative work: it takes information gathering actions that would have been impermissible and supposedly makes them permissible. As we have seen, our modern lives are filled with personal data transactions. And while it has been argued that the ubiquity of these personal data transactions is such as to entail that we are living within a *de facto* surveillance state [67], there is at least one important *prima facie* difference: surveillance states are typically imposed on subjects without their consent, so that personal information in a surveillance state is gathered without any consideration to the preferences of the surveilled, whereas (it is claimed) we freely consent to the personal data transactions that we are subject to. (Note that we are not here endorsing this argument; we present it merely in order to motivate our analysis of the role of consent in personal data transactions.) In this way, consent for personal data transactions functions analogously to how consent functions in medical ethics and how consent functions in sexual ethics.

To borrow a phrase used by both Heidi Hurd and Larry Alexander, consent is a kind of "moral magic": "it transforms acts from impermissible to permissible" [2, 28]. This is true in generic ways; A's entering B's house can be either a trespass or a permissible visit based on whether B has consented to A's entering. It is particularly true in matters of sexual ethics, where the moral status of a sexual act crucially depends on whether the act is consented to at the time that it is performed [19]. Furthermore, it is equally valid for medical ethics, where invasive medical procedures and treatments are impermissible unless consented to by the patient.

Regardless of the legality, it is reasonable to think that if A were to persistently and systematically surveil B (going through B's trash, compiling every public record about B, recording all of B's public movements), and B were not a public figure who might be an appropriate target of community scrutiny, then that would be an impermissible form of information gathering. Of course, if B were to consent to their information being gathered in this way by A (say they are the willing subject of a documentary), then the information-gathering would thereby become permissible. So consent would have the same kind of moral magic as it does in other contexts [42].

Likewise, personal information transactions open up the risk of one's autonomy or integrity being violated. After all, personal information is personal—and as such, it potentially enables others to identify them and predict or control one's behavior in a way that compromises one's autonomy and capacity to act on one's intentions and one's own conception of the good. It is true that we share personal information with others around us all the time; the mere sharing of personal information does not compromise one's autonomy. However, we typically share personal information with those we trust to use the information correctly. We share information that is anodyne enough to not threaten our ability to pursue our own goals; we do not typically share personal information with those we know intend to use that information to circumvent our autonomy or act contrary to our interests. This is why privacy is important even for those who do not think themselves to have a strong taste for privacy: autonomy matters because it is the means by which we pursue the good, and privacy is connected to autonomy. Getting legitimate consent to personal data transactions helps to ensure that one's autonomy is safeguarded.

Finally, personal data transactions often involve commercial parties, either as one of the parties to the data transaction (as happens when you upload your information to Facebook) or as the platform in which data transaction occurs (as when you share your data with your friends through Facebook). Commercial parties have an interest in limiting their liability. To this end, obtaining putative consent to acquire and process data helps limit the legal liability they face. Companies and commercial agents have a vested interest in preventing the end users of social media sites and data technologies from complaining. Obtaining consent helps support the argument that the end-users have, indeed, forfeited their right to complain about any consequences arising from the personal data transactions.

1.3 Consent in Data Transactions

Before discussing the limitations of the individual consent model for personal data transactions, it is worth addressing one question: Why would we ever have thought that consent was relevant for data transactions at all? After all, the argument goes, data transactions are just one species of transaction. Moreover, transactions are necessarily mutually consensual; if they were not, they would not be transactions. An exchange in which A gets a beer and B gets a dollar is a transaction if they both agree, but it is robbery or coercion if either A or B does not consent to the exchange. (The fact that both A and B receive something is irrelevant; if A breaks into B's house and steals B's property, it is no less a robbery just because A left something behind in exchange.) Likewise, the objection goes, that if A and B are engaged in a personal data transaction, then it is irrelevant to ask whether the transaction is consensual; if it were not, it would not be a personal data transaction but would instead be a data theft or something similar. If this is right, it is as unnecessary to ask whether a data transaction is consensual as it would be for the cashier at a clothing store to explicitly ask every patron whether they consent to trade their money for the clothes they wish to buy.

One reply to this objection is to say that we actually *do* care about obtaining consent for some exchanges precisely because we want to ensure that the exchange is a "transaction" rather than a "theft." When the stakes are high. There is a possibility of future risk. In other words, we do seek to clarify the consensual nature of the exchange, but when the stakes are low and there is a low perception of risk, we are less concerned [1, 21, 55]. This reply is correct, but we think an important point is in danger of being obscured. In the case of ordinary transactions, there is less danger of the transaction being non-consensual precisely because there is little danger of the parties to the exchange not knowing what they are exchanging; when A gives B money, and B gives A a shirt, both parties are aware that the transaction happened because both parties have clearly gained something and have lost something.

Moreover, there are positive actions that A and B both perform, without which the transaction cannot take place; A must hand over the money, and B must hand over the shirt. However, we can easily imagine transactions in which these conditions are not met—perhaps, for instance, it is not clear among the parties precisely what has been gained and what has been lost after a transaction. (One might think of the "sale" of the island of Manhattan by Native people that Peter Minuit orchestrated on behalf of the Dutch; what exactly is one selling if the land is still there after the transaction is done?) Alternatively, imagine transactions for which no positive actions need to be taken, like arrangements that automatically deduct money from a bank account or other store of value without anyone needing to do anything. It is certainly not challenging to sway our intuitions towards thinking of these "transactions" as theft or, at a minimum, theft-adjacent. The same issues arise with data transactions (whether legitimate or not). We might wonder whether the exchange of data between A and B is actually a transaction if one or both parties are not clear on precisely what has been gained or lost—but, as we will see, this is a common feature of most modern personal data exchanges. It is certainly not like a transaction for a shirt, where one minute you have a dollar and the next minute you do not; after a personal data exchange is complete, you have just as much of your own personal data as you started with. Personal data exchanges often do not require any positive action on behalf of the parties; we lose data in personal data exchanges all the time, through no action of our own. So while it is acceptable to say that valid transactions are consensual, it is also true that not every data exchange is a data transaction in that strict sense. It is reasonable to ask for explicit consent to data exchanges so that all parties are confident that it is a data transaction and not a mere exchange.

It is a reasonable strategy, but it is nevertheless a failed one. As we will see, there are systematic reasons why personal data exchanges cannot be justified by individual consent.

2 RESULTS

2.1 A General Overview of the Problems with Individual Consent

We can now provide a very general overview of the problems with individual informed consent when applied to data transactions. Note that we are assuming for the sake of this discussion that the relevant data are, in fact, adequately subject to control by individuals. This is a problematic assumption; data often implicate multiple individuals or are otherwise 'co-owned' and thus are not properly the things that individual consent can govern. This important point requires a fuller discussion, but we make this simplifying assumption here because social networks cannot function in anything like their current form without it.

Consent, in this context, should not be mistaken for a state of mind or an attitudinal event [31, 64] and it must meet certain criteria in order to be considered a valid. The legitimacy of consent hinges on a number of criteria [9]:

(1) the subject has sufficient accurate information and understands the nature of the agreement,
(2) the agreement is entered into without coercion,
(3) the agreement is entered into knowingly and intentionally,
(4) the agreement authorizes a specific course of action.

In the context of personal data transactions, digital consent also rests on the four criteria mentioned above. The user agrees to the specified service terms and privacy policy outlined by the data processor. Notably, the four criteria listed above fail in the context of personal data transactions and classic Privacy Policies and Terms of Service (ToS) agreements.

First, most users entering into consent agreements know very little about data processing or the risks of handing over their data. The dense legal and technical nature of ToS agreements task non-experts to consent to something they do not understand [17]. This dynamic takes advantage of asymmetry in technical and legal knowledge.

Second, it is difficult to opt-out of these services since online platforms are an important social ecology where people form personhood, maintain personal relationships, and build valuable networked counter-publics [23, 29, 30, 50]. Nevertheless, there is little to no power on the part of the individual to negotiate the ToS with these companies, as consent in these ToS is typically presented on a take-it-or-leave-it basis and offers no conditions of choice [44, 54]. Online privacy then turns into an unfortunate social optimization problem [62], where the user must choose between the pressures of disclosing too much personal information (being digitally crowded) and being socially isolated [4].

Third, the volume of consent requests a user faces has led to a troublesome externality where the user is fatigued and habitually agrees to everything due to consent desensitization [17]. This delegitimizes the premise that each act of putative consent reflects the individual user's autonomous judgment.

Fourth, the language in ToS agreements is typically so broad and open-ended that data processors have the flexibility to manipulate the data in many ways. The consent scope cannot be so broad as to allow actions that the user could not have considered or would otherwise not have consented to. An adequately limited scope of consent also implies that there should be some mechanism for a user to check if their data are indeed following the agreed-upon course of action. However, data processors often make it very difficult [32], if not impossible, to track personal data, know what they have collected or how it is being processed, and hold them accountable for misuse [13, 57].

A fundamental assumption for individual consent is that the user has power over their personal data and can trade their personal privacy in exchange for using an online service [14]. Perhaps more importantly, a significant concern with the individual consent model is that personal data, in this context, are distributed information that contains information about more than a single individual and spans a broader communication boundary than the user is aware [48]. In reality, these data may not belong wholly to the individual. Therefore, it is not appropriate for the individual to act alone in controlling the course of action or the flow of these data. Perhaps the first step to understanding the impact of this issue in an online social media context is to understand how different levels of consent impact the flow of networked data and observability in the first place. It will be important for us to understand if the network effect has a strong influence on privacy to justify group-level consent settings.

2.2 A Threat Model for Leaky Individual Data in Social Networks

The densely interconnected nature of online social ecology creates a significant problem with the model of individual consent. When users share personal information online, they are also leaking personal information about others in their social network (digital or otherwise) [6, 24]. According to Bagrow et al., "due to the social flow of information, we estimate that approximately 95% of the potential predictive accuracy attainable for an individual is available within the social ties of that individual only, without requiring the individual's data" [6].

One example of leaky data is when a user attempts to sign on to a new online service. They may be prompted to skip the hassle of entering their personal information manually and instead opt to use an existing account to act as a secured access delegation [63] in order to gain quicker access to the new third-party online service. The online service can ask to gain access to the user's online social network, phone or email contacts, location, and other personal data. Through these leaky data, third parties can be granted access to a wealth of knowledge about people who never consented to share their information with that particular service.

Similarly, attackers can breach social accounts via various methods, including phishing attacks, malware, and data breaches [18, 27, 59]. They can also create fake social accounts and then use those fake accounts to befriend real accounts. To extend their reach, those attackers can then monitor the activity of other accounts that are directly connected to compromised or fake accounts. By leveraging network effects, attackers can further indirectly observe the social activity of groups of accounts that are several hops away from the captive accounts, starting with accounts that are one hop away [16, 60].

Leveraging those captive accounts, the attackers traverse the social graph or segments of it and record profile information and social activity that would later be used in future phishing and spam attacks, influence manipulation attempts [65], and disinformation campaigns [5], among others. Given the massive size and connected nature of online social networks, the potential reach of attackers and the resulting harm both have the capacity to rise to catastrophic levels. We describe examples of real-world incidents that demonstrate the severity of this problem in our discussions.

2.3 A Model of Distributed Consent and Network Observability

To account for the distributed nature of personal data (i.e., the distributed online self), we consider a simple model of distributed consent. Imagine a social network platform where individuals have the following privacy options:

0. Individuals share their data with all their connections and are vulnerable to third-party surveillance (similar to Facebook accounts with access for "Apps, Websites and Games" turned on).
1. Individuals share their data with all their connections but are not directly vulnerable to third-party surveillance.
2. Individuals only share their data with their connections whose privacy levels are set at least 1.

N. Individuals only share their data with connections whose privacy levels are set at least to $N - 1$.

Options 2 and greater are currently unavailable on popular social media platforms but are a first-order implementation of what we call distributed consent. Simple implementations of this concept could be an attractive setting to adopt if the platforms wish to address privacy concerns and keep users. Individuals who pick this option are stating that they want to be part of a local group that agrees on minimal privacy settings. It is a consent conditional on the consent of their neighbors in the network structure.

Imagine now that a third party wishes to observe this population, either by releasing a surveillance application on the social network or by explicitly gaining control of their accounts through similar malware. Say they can directly observe a fraction φ of individuals with privacy level set to 0 through this attack. They can then leverage these accounts to access neighboring individuals' data with the privacy level set to 1 or 0, therefore using the network structure to observe more nodes indirectly. They can further leverage all of these data to infer information about other individuals further away in the network, for example, through statistical methods, facial recognition, and other data sets.

Deep surveillance processes are relevant to other network systems where it is possible to indirectly observe nodes within a certain distance from directly observed nodes. Deep surveillance allows us, for example, to monitor an entire power grid without monitoring the voltage and line currents everywhere in the system [66]. It was recently shown that the surveillance process itself could be generally modeled through the concept of depth-L percolation [3]: Monitoring an individual allows one to monitor their neighbors up to L hops away. Depth-0 percolation is a well-studied process known as site percolation. The third-party would then be observing the network without the help of any inference method and by ignoring its network structure. With depth-1 percolation, they would observe nodes either directly or indirectly by observing neighbors of directly observed nodes, e.g., by simply observing their data feed or timeline. Depth-2 percolation would allow one to observe not only directly monitored nodes but also their neighbors and neighbors' neighbors, e.g., through statistical inference [6]—and so on, with deeper observation requiring increasingly advanced methods. The model is illustrated in Fig. 1 and detailed in our Methods section.

To study the interplay of consent and observability on social networks, we combine our distributed consent and depth-L percolation models on subsets of Facebook friendship data informed by empirical work on the demographic population's taste for privacy. We assume that one-third of the population has a taste for privacy [34]. These data are anonymized with all metadata removed and are simply used to capture the density and heterogeneity of real online network platforms. We use distributed consent with a security level up to $N = 2$ and an observation process with $L = 2$ (observing nodes two hops away from the compromised account). As we will see, those values mean that the third party is more sophisticated than our distributed consent mechanism, and no one is *guaranteed* to be unobservable. We then set 1% of accounts with the lowest security setting to be compromised and directly observed such that that between 90% and 100% of the population will be observed given the default security settings. We then ask to what extent distributed consent can preserve individual privacy even when a large fraction of nodes can be directly observed and third parties can infer data of unobserved neighbors. How widely should distributed consent be adopted to guarantee connectivity *and* privacy of secure accounts?

The results of our simulations are shown in Fig. 2 and Fig. 3. We focus on the number of observed nodes of different security levels and on the size of the giant component of unobserved nodes. This last quantity refers to the size of the largest subpopulation of accounts that are not observed and maintain connected pathways of any lengths between one another, thus preserving both their individual privacy and the global connectivity objective of the social network. In classic percolation theory, only one giant component can span the entire system [58], meaning only one subset of nodes can scale with the total size of the social network. Yet, from recent results on network observability [3], we also know that giant *observed* components can co-exist with giant *unobserved* components. This is where distributed consent can play a large role: Even if a third-party surveillance system scales with the size of the social network, it is theoretically possible for accounts to maintain their individual privacy and global connectivity at scale.

Figure 2 shows the results of our model in populations facing either a very strong attack (1% of compromised accounts, chosen to observe almost the full population) or a more modest attack (0.05% of compromised accounts, chosen for about 50% observation). Against a strong attack, we find that while extremely low adoption levels of distributed consent have little impact on the observability of the system, moderate adoption (roughly 1 in 5 users) can lead to a transition where observability now drops sharply with the adoption of distributed consent; see Fig. 2(a, e). There are few unobserved nodes at a low adoption rate of distributed consent. All are mostly disconnected from each other and therefore observable through compromised neighbors. At higher levels of adoption rate, the system transitions to an unobservable and connected phase where privacy can co-exist with connectedness and information flow; see Fig. 2(b, f). With large-scale adoption of distributed consent (say one-third of users), we find that close to half of all accounts are now protected even against very strong attacks, even if their privacy settings only prevent about 22% of data flow around them.

Against the more modest attack, increases in the adoption of distributed consent lead to an increase in smaller but smoother and more reliable protection. With 33% adoption, populations can halve the size of their observed population and conversely double the size of their unobserved component.

To understand these results, notice that any user with privacy settings set to a greater value than the percolation depth will be unobservable. If we had set simulated naive attackers who observe at a depth $L = 1$ only, adopters of distributed consent would have been fully protected. Indeed, users using the security setting N will only share their data with users using settings of $N - 1$ or more, and this statement holds for all N. We thus know that users using setting N will be at least N steps away from users using the lowest setting, which are the only directly observable nodes. Users with security level set to $1 < N < L$ can however be observed indirectly through their relationships. At low levels of adoption of distributed consent, a large amount of luck is required to remain unobservable (e.g., having zero connections with low-security users). At higher levels

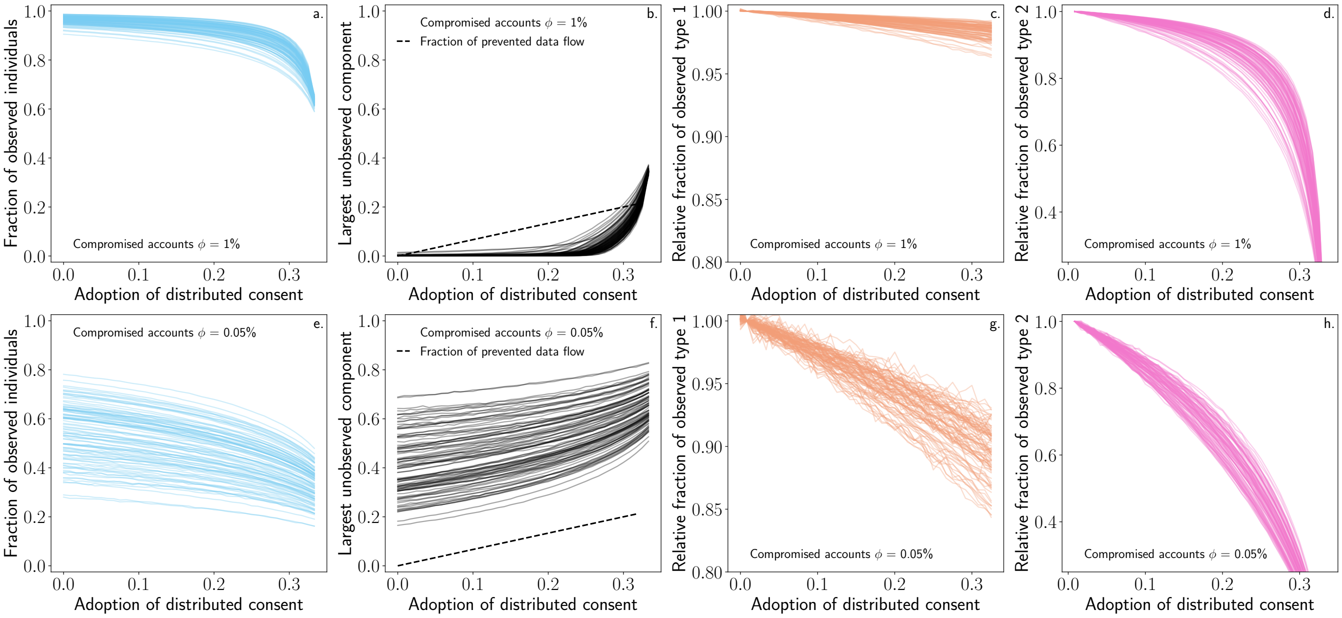


Figure 2: We use the anonymized Facebook100 dataset [61]. We assume that one-third of the population has a taste for privacy [34], split between security options 1 and 2 (i.e., classic or distributed consent) according to the adoption rate of distributed consent shown on the horizontal axis. At the same time, the remaining two-thirds will use the default setting with the lowest security, option 0. We set 1% (top row) or 0.05% (bottom row) of accounts with security option 0 to be directly observable by a third-party app, which can also observe neighbors up to two hops away in the network. These values are chosen to model attacks that observe nearly the entire population (top row) or about a half of it (bottom row). We then vary the adoption rate and measure the total fraction of observed accounts (blue curves), the relative size of the largest unobserved connected component (black curves), and the fraction of observed individuals with security option 1 (orange curves) or two (pink curves).

of adoption, users of distributed consent connect to and therefore protect, one another. However, these connections are localized and do not spread throughout the entire system. We find that when roughly 25% of nodes with a taste for privacy adopt distributed consent, a large macroscopic component of connected unobservable nodes emerge even against the strongest attack. This component reflects a parallel, protected community that is unobservable but still connected to the rest of the social network.

Even though a phase transition in connected unobservable nodes occurs at a fairly low level of distributed consent adoption, these nodes provide secondary protection to other users. The pervasive adoption of group consent is required to fully protect a network. Again, a single observable neighbor is all it takes for one vulnerable node to be indirectly observed. Because of this and because of the density of most online networks platforms, it is extremely hard to completely protect vulnerable nodes even if distributed consent provides some secondary protection to all nodes. We thus see the coexistence of both observed and unobserved connected components at the medium adoption level of distributed consent. Interestingly, these components are interconnected, with data flowing both ways across observable and unobservable components, yet the users in the latter remain fully protected from statistical inference of their data.

Importantly, the macroscopic but unobservable component that we see emerge with increased adoption of distributed consent does not only contain adopters of distributed consent. Early adopters of distributed consent provide some low amount of *herd privacy* to the population, protecting otherwise vulnerable users; see Fig. 2(c, g). Users with lower privacy settings can thus also benefit since the adoption of distributed consent in one's neighborhood reduces the probability that one of their neighbors is directly or indirectly observed, thereby reducing the probability that they are themselves observed. However, as long as a majority of users rely on the default lax security settings, this effect will be limited as a single compromised neighbor is sufficient to observe a node. However, we do find much stronger herd privacy effects against more moderate attacks. Moreover, similar effects provide non-linear returns on relative protection of distributed consent users, Fig. 2(d, h), consistent with our previous observation of the emergence of an unobserved component.

Finally, in Fig. 3, we reproduce the large attack against a population with a stronger taste for privacy and compare our previous results with those obtained by halving the fraction of unprotected nodes (two thirds to one third). Unsurprisingly, the stronger the taste for privacy, the stronger the effects of distributed consent; both at the macroscopic level (the fraction of observed nodes and the size of the unobserved component in panels a, b, e, and f) and at the microscopic level (herd privacy effects shown in panels c, d, g, and h). To make the comparison easier, we plot all quantities against the fraction of users with a taste for privacy (any security

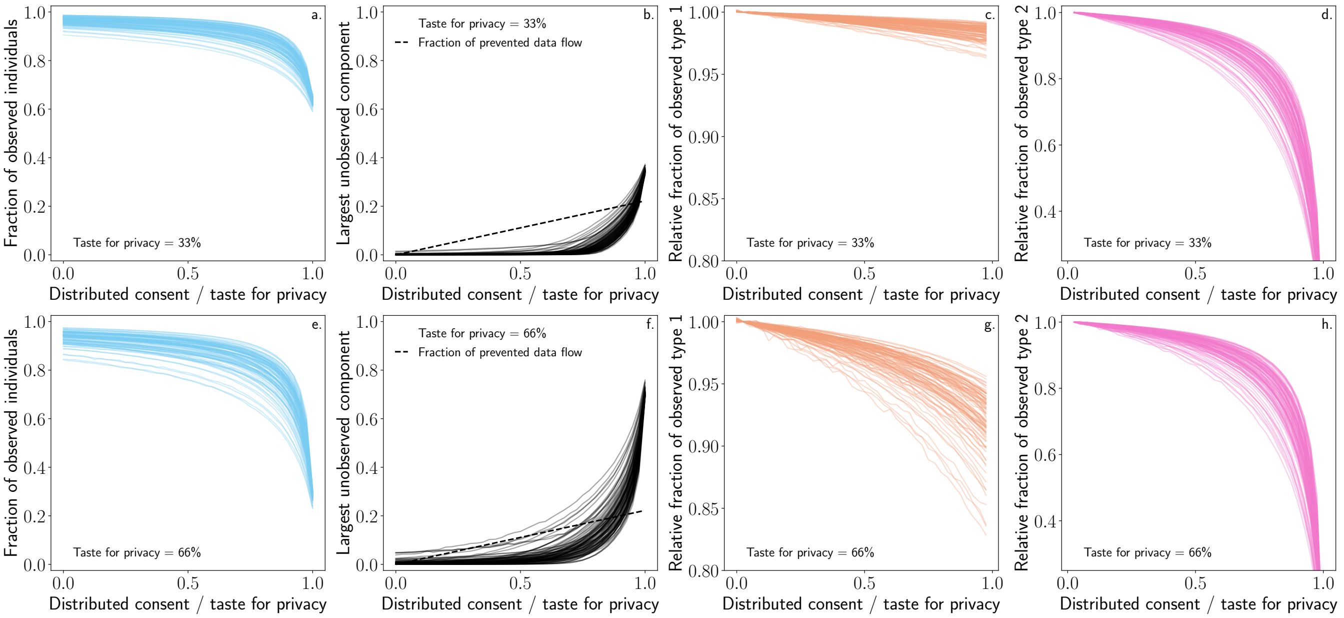


Figure 3: We reproduce results from Fig. 2 but using populations with different taste for privacy. The results in the top row of this figure match the top row of Fig. 2 but are now shown as a function of the nodes with a taste for privacy (security level greater than 0) that opt for distributed consent (security level greater than 1). We then change the fraction of the population with a taste for privacy from 33% (top row) to 66% (bottom row). Qualitatively, the results are very similar. Therefore, the key quantity that drives the macroscopic effects of distributed consent is not its total adoption but its relative adoption within individuals that cannot be directly observed.

level other than the lowest) that opt for distributed consent. This ratio collapses all results on similar curves and therefore appears to be the critical quantity in determining the privacy level of a population.

2.4 A Model of Coordinated Consent Across Platforms

Currently, there is a David and Goliath problem [36]; the inequality of knowledge and power between the user and the data collector functionally takes away the individual's ability to control their personal information realistically. Part of the problem we believe is that the terms of the agreement are predicated solely on the platform's norms. Users do not have the option to opt-out and pay for the service in exchange for limiting information flow for most online social networks. There is a high cost on the user's side to read and understand all ToS agreements presented to them [40].

There is little to no power on the part of the individual to negotiate the ToS with tech companies. Online consent in its current state creates an unfortunate social optimization problem, where the user must choose between the pressures of disclosing too much personal information (being digitally crowded) and being socially isolated [4, 8]. Moreover, this also implies that most platforms have little power to change the digital ecosystem on their own if users can be exposed to other platforms with laxer security and privacy policies. Multiple platforms create a multilayer network where information flows through social connections and across platforms through different layers of the network. This complex network structure exposes users to whichever platform offers them the least privacy.

Our second model focuses on coordinated consent across platforms. Inspired by previous work on automated ToS tools [7, 26, 35, 41, 45, 47, 51], we envision a *consent passport* model where instead of relying on every platform to offer advanced privacy settings and therefore asking users to adjust their settings on every platform individually, users could use a consent passport stating their desired privacy settings before they join a platform. This could take the form of a key enabling a user to set their privacy baseline criteria based on their taste for privacy. The key will act to shift the burden [15, 25, 56] away from the user to read and understand the legalese of the platform's privacy policy and ToS. The consent key would restrict login and present a warning [12] when the user attempts to enter sites that do not meet a minimum privacy criteria in their ToS and privacy settings. This key is intended for users who have a taste for privacy but do not want to fall prey to consent desensitization. The consent passport will need to be dynamic in nature so that users can remain autonomous in their decision-making and easily opt to enter a platform with discordant privacy settings if they trust the site or decide that the privacy cost is worth the risk. Importantly, this ensures that a given user's security settings will be coordinated across platforms, which might be the only way to confidently protect them in a complex multilayer ecosystem.

To include a consent passport within our simulation model, we turn our network data into a multilayer infrastructure by duplicating the Facebook networks' structure to represent two platforms

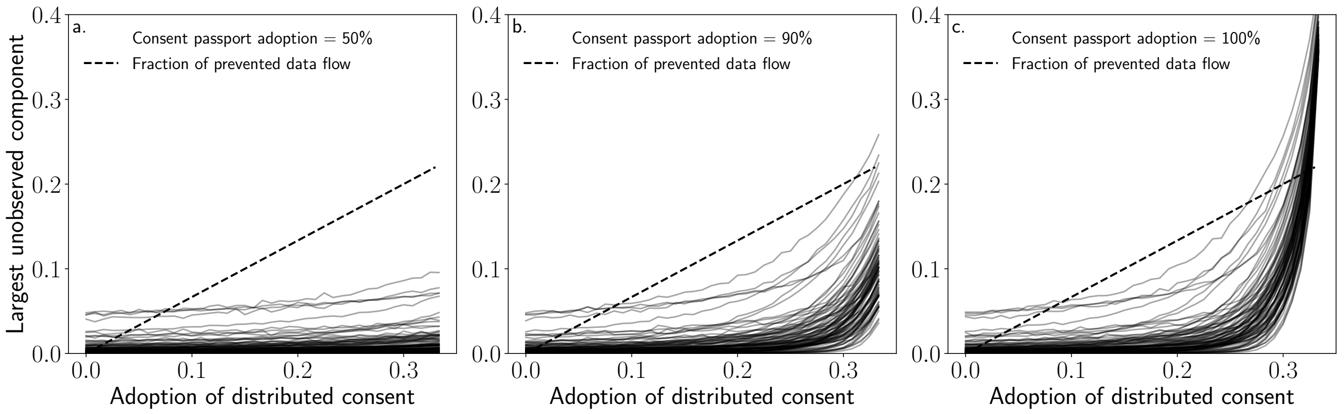


Figure 4: We again use the anonymized Facebook100 dataset [61]. We now create a multilayer network by doubling the original data, mimicking a two-platform ecosystem. We use the same parameters as in Fig. 2, assuming that for each platform, only one-third of accounts have changed their default security setting to options 1 or 2 (i.e., classic or distributed consent) according to the adoption rate of distributed consent shown on the horizontal axis. The remaining two-thirds will use the default setting with the lowest security, option 0. To account for the doubled network density and the fact that users can be observed on either platform, we now set 0.25% of accounts with security option 0 to be directly observable by a third-party app that can also observe neighbors up to two hops away through any layer of the network. We then vary the adoption rate of distributed consent (horizontal axis) and, within that subset of the population, vary the adoption of consent passport (different panels). In the resulting systems, we measure the relative size of the largest unobserved connected component. Randomly protecting users on a single platform does not protect anyone if spyware can jump network layers, but distributed consent coordinated across platforms through a consent passport can restore our ability to create an unobservable component within the multilayer ecosystem.

where users are randomly assigned security setting independently on each platform. However, a subset of users adopts a consent passport which guarantees that they will follow our previous model of distributed consent on *both* platforms. We again use depth-L percolation with $L = 2$ in the resulting systems to simulate a third party's ability to observe network users. As a worst-case scenario, we allow the depth-L percolation process to be able to jump between layers of the networks freely, e.g., observing a Facebook neighbor of an individual directly observed through Instagram. As detailed in our Methods section, we classify any user as observed if they are observed on either platform.

Figure 4 shows that distributed consent alone cannot protect you if your security settings are not coordinated. Indeed, we parametrize the system such that the networks are roughly equally observable. However, there is no emergence of a giant unobservable component even with medium adoption of consent passports (50% adoption among distributed consent users, left panel). It is only at very high adoption of consent passports that we start seeing a non-linear benefit in unobservability (90% adoption, middle panel) and only at near-perfect adoption that we protect more users than we prevent data flow (95% and above, right panel). These results demonstrate that multi-platform ecosystems are a much more complex beast to protect; users participating in multiple platforms are only as secure as their weakest security settings. Therefore, coordinating privacy settings across platforms is a critical part of the solution. We discuss this problem further in the next section.

While our current results illustrate how mathematical and computational models can contribute to the study of consent and privacy policies, it is critical to keep all of their assumptions and approximations in mind before drawing conclusions about real-world applications. In the case of our model, we have assumed that all users generate and share similar amounts of data, that all users with a given security setting areas are susceptible to privacy breaches, and that security settings are uncorrelated with the connectivity in the network. In practice, one might imagine that high-profile accounts tend to have higher security settings but that they might also face more frequent attacks. Accounting for correlations between network structures, taste for privacy, and susceptibility to breaches could be included in the model, but these mechanisms first require further empirical studies.

3 DISCUSSION

3.1 Implications

In recent real-world incidents, attackers reportedly leveraged online social networks to target groups of people. In 2020, BBC News reported that a foreign intelligence agent allegedly used LinkedIn, a prominent online social network, to locate and befriend "former US government and military employees" [49]. Additionally, the same BBC article reported that Germany's intelligence agency stated that foreign agents "used LinkedIn to target at least 10,000 Germans" in 2017. And another example comes to mind: the Cambridge Analytica case [65], in which political entities made attempts to sway the political stance of groups of people via Facebook, another large online social network.

In view of these real-world examples, the need for a model of distributed consent, as presented in this paper, becomes more apparent. Although the proposed model would not completely stop the attackers, it offers a better level of protection to users of online social accounts than the status quo. In other words, the broad reach of attackers within the threat model presented previously could be restricted with the deployment of the model of distributed consent. We hope that online social network platforms will consider and adopt models where users can have more power to coordinate their privacy with their network neighbors and across platforms. We also hope that policymakers would actively push for adopting similar models to help make online social networks safer for all users.

3.2 Conclusion

Altogether, in this work, we provided a philosophical critique of individual consent in the context of data transactions and used a modeling framework to suggest potential solutions.

As part of our philosophical critique, we listed four criteria for the legitimacy of informed consent. We argued that none of the four criteria are met by individual consent within online media's complex ecosystem. Further, a fundamental problem is that if personal data are distributed across individuals, so should be their consent.

Our results based on computational models and simulations suggest that even the simplest implementation of distributed consent could allow users to protect themselves and the flow of their data in the network. They do so by consenting to share their data conditionally on the consent or security settings of their contacts, thereby not sharing their data with users who might, in turn, make them available to third parties. This simple condition allows users to authorize a specific course of action for their own personal data (criterion 4).

While this protection disconnects them from some other users, only a relatively low level of adoption of distributed consent is required to create a connected macroscopic sub-system within existing online network platforms. This sub-system consists of different individuals, including some that are granted secondary protection despite their low-security settings and remain connected to the rest of the system such that information still flows throughout the entire population of users. Via this protected sub-system, distributed consent removes the *de facto* coercion (criterion 2) involved in forcing individuals to choose between relinquishing control of their data or simply not participating in a platform.

Beyond the actual protection mechanism, this new consent model may also have interesting behavioral impacts on the users. Exposing users to this type of coordinated privacy setting might prompt them to reflect on their personal data's distributed nature and its flow through online media. This realization may encourage users to openly voice their social boundaries to their social network or restrict sending sensitive information to social neighbors who do not share their taste for privacy [34]. Imagine a user publishing a post to their social network before enacting the new privacy settings, urging those who want to remain connected to change their settings as well. Beyond the utility of limiting the social network's observability, this measure could also serve as an important educational tool on the interconnectedness of personal data (criterion 1).

In a modest form, distributed consent could allow concerned users to protect themselves without entirely leaving a platform. It would also let platforms maintain a large critical mass of observable users that chose to remain vulnerable and who are not granted sufficient protection through their contacts.

That being said, legitimate consent criterion 1 (understanding the consent agreement) and criterion 3 (or consent fatigue) remains an issue that will need additional consideration in future work. An important caveat is that a useful implementation of distributed consent might require platforms to provide additional education regarding data privacy. Regarding criteria 1 and 3, the consent passport attempts to shift the agreement's burden to platforms rather than users. In doing so, it may provide additional protection in complex multi-platform ecosystems. There are many types of privacy violations that are not solved by distributed consent. These data are still leaky; individual users can still aggregate information about their neighbors without their explicit consent. Finally, while the distributed consent model goes beyond the strict individuality of the traditional privacy model, it does so modestly; it still models the agents, choices, and values as fundamentally individual. Obviously, there is no silver bullet to solve this multi-scale complex problem; data privacy is a significant societal issue with multi-level interdependencies that must be considered thoughtfully and ethically. Much work remains to be done in this area.

Future work should extend to look at possible collective behavior around the adoption of new consent models. Indeed, the greatest hurdle to herd-like immunity against network observability is our assumption that only one-third of the population has a taste for privacy such that two-thirds of users will never deviate from the default lax security settings. Users signaling their adoption of distributed consent and potentially influencing their network neighbors to do the same could then spark a contagious taste for privacy whose co-evolution with observability could be modeled using tools from network epidemiology [46].

Beyond new notions of consent, effective data privacy measures will need to take a systems-level approach and integrate a mechanism for distributed moral responsibility [22] that will simultaneously involve both top-down and bottom-up interventions. Doing so will involve a synergy between increased governmental and professional regulation, technological intervention, distributed consent, and citizens' empowerment. Increasing data privacy and protection is not only an essential public service but a democratic imperative [20, 23, 50]. Access to data privacy and protection is a growing global issue [39], and it must be investigated through further multidisciplinary collaboration.

ACKNOWLEDGMENTS

The authors would like to thank James Bagrow for his early discussions and feedback on this project. This work is supported by Google Open Source under the Open-Source Complex Ecosystems And Networks (OCEAN) project, MassMutual under the MassMutual Center of Excellence in Complex Systems and Data, the Natural Sciences and Engineering Research Council of Canada, and the Sentinelle Nord program from the Canada First Research Excellence Fund. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the aforementioned financial supporters.

REFERENCES

[1] Alessandro Acquisti, Leslie K. John, and George Loewenstein. 2013. What Is Privacy Worth? *The Journal of Legal Studies* 42, 2 (June 2013), 249–274. https://doi.org/10.1086/671754

[2] Larry Alexander. 1996. The Moral Magic of Consent (II). *Legal Theory* 2, 3 (September 1996), 165–174. https://doi.org/10.1017/s1352325200000471

[3] Antoine Allard, Laurent Hébert-Dufresne, Jean-Gabriel Young, and Louis J. Dubé. 2014. Coexistence of Phases and the Observability of Random Graphs. *Phys. Rev. E* 89 (February 2014), 022801. Issue 2. https://doi.org/10.1103/PhysRevE.89.022801

[4] Irwin Altman. 1977. Privacy regulation: Culturally universal or culturally specific? *Journal of social issues* 33, 3 (1977), 66–84.

[5] Gregory Asmolov. 2018. The disconnective power of disinformation campaigns. *Journal of International Affairs* 71, 1.5 (2018), 69–76.

[6] James P. Bagrow, Xipei Liu, and Lewis Mitchell. 2019. Information Flow Reveals Prediction Limits in Online Social Activity. *Nat. Hum. Behav.* 3, 2 (January 2019), 122–128. https://doi.org/10.1038/s41562-018-0510-5

[7] Vinayshekhar Bannihatti Kumar, Roger Iyengar, Namita Nisal, Yuanyuan Feng, Hana Habib, Peter Story, Sushain Cherivirala, Margaret Hagan, Lorrie Cranor, Shomir Wilson, Florian Schaub, and Norman Sadeh. 2020. Finding a Choice in a Haystack: Automatic Extraction of Opt-Out Statements from Privacy Policy Text. In *Proceedings of The Web Conference 2020*. Association for Computing Machinery, New York, NY, USA, 1943–1954. https://doi.org/10.1145/3366423.3380262

[8] Susan B. Barnes. 2006. A privacy paradox: Social networking in the United States. *First Monday* 11i9 (2006).

[9] Tom L. Beauchamp and Ruth R. Faden. 1986. *A History and Theory of Informed Consent.* Oxford University Press, Oxford, United Kingdom.

[10] Frederik J Zuiderveen Borgesius. 2016. Singling out people without knowing their names–Behavioural targeting, pseudonymous data, and the new Data Protection Regulation. *Computer Law & Security Review* 32, 2 (2016), 256–271.

[11] Frederik Zuiderveen Borgesius. 2015. Informed consent: We can do better to defend privacy. *IEEE Security & Privacy* 13, 2 (2015), 103–107.

[12] Ryan Calo. 2011. Against notice skepticism in privacy (and elsewhere). *Notre Dame L. Rev.* 87 (2011), 1027.

[13] Fred H Cate. 1999. Principles of internet privacy. *Conn. L. Rev.* 32 (1999), 877.

[14] Julie E. Cohen. 2000. Examined Lives: Informational Privacy and the Subject as Object. *Stan. L. Rev.* 52, 5 (May 2000), 1373–1438. https://heinonline.org/HOL/P?h=hein.journals/stflr52&i=1392

[15] Jessica Colnago, Yuanyuan Feng, Tharangini Palanivel, Sarah Pearman, Megan Ung, Alessandro Acquisti, Lorrie Faith Cranor, and Norman Sadeh. 2020. Informing the Design of a Personalized Privacy Assistant for the Internet of Things. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376389

[16] Ana-Maria Creţu, Federico Monti, Stefano Marrone, Xiaowen Dong, Michael Bronstein, and Yves-Alexandre de Montjoye. 2022. Interaction data are identifiable even across long periods of time. *Nature Communications* 13, 1 (2022), 1–11.

[17] Bart Custers, Simone van Der Hof, Bart Schermer, Sandra Appleby-Arnold, and Noellie Brockdorff. 2013. Informed consent in social media use — the gap between user expectations and EU personal data protection law. *SCRIPTed* 10, 3 (December 2013), 435–457. https://doi.org/10.2966/scrip.100113.1

[18] Rachna Dhamija, J. D. Tygar, and Marti Hearst. 2006. Why Phishing Works. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Montréal, Québec, Canada) *(CHI '06)*. Association for Computing Machinery, New York, NY, USA, 581–590. https://doi.org/10.1145/1124772.1124861

[19] Tom Dougherty. 2014. Fickle consent. *Philosophical Studies* 167, 1 (2014), 25–40.

[20] Ritam Dutt, Ashok Deb, and Emilio Ferrara. 2018. "Senator, We Sell Ads": Analysis of the 2016 Russian Facebook Ads Campaign. In *International conference on intelligent information technologies* (Chennai, India) *(Advances in Data Science. ICIIT 2018)*. Springer, New York, NY, USA, 151–168. https://doi.org/10.1007/978-981-13-3582-2_12

[21] J Doyne Farmer and John Geanakoplos. 2009. Hyperbolic discounting is rational: Valuing the far future with uncertain discount rates. *Cowles Foundation Discussion Paper* No. 1719 (2009).

[22] Luciano Floridi. 2016. Faultless Responsibility: On the Nature and Allocation of Moral Responsibility for Distributed Moral Actions. *Philos. Trans. Royal Soc. A* 374, 2083 (December 2016), 20160112. https://doi.org/10.1098/rsta.2016.0112

[23] Nancy Fraser. 1990. Rethinking the Public Sphere: A Contribution to the Critique of Actually Existing Democracy. *Soc. Text* 25/26 (1990), 56–80. https://doi.org/10.2307/466240

[24] David Garcia. 2017. Leaking Privacy and Shadow Profiles in Online Social Networks. *Sci. Adv.* 3 (2017), e1701172. https://doi.org/10.1126/sciadv.1701172

[25] Elias Grünewald and Frank Pallas. 2021. TILT: A GDPR-Aligned Transparency Information Language and Toolkit for Practical Privacy Engineering. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (Virtual Event, Canada) *(FAccT '21)*. Association for Computing Machinery, New York, NY, USA, 636–646. https://doi.org/10.1145/3442188.3445925

[26] Alfonso Guarino, Nicola Lettieri, Delfina Malandrino, and Rocco Zaccagnino. 2021. A Machine Learning-Based Approach to Identify Unlawful Practices in Online Terms of Service: Analysis, Implementation and Evaluation. *Neural Comput. Appl.* 33, 24 (dec 2021), 17569–17587. https://doi.org/10.1007/s00521-021-06343-6

[27] William G Halfond, Jeremy Viegas, Alessandro Orso, et al. 2006. A classification of SQL-injection attacks and countermeasures, In Proceedings of the IEEE international symposium on secure software engineering. *IEEE* 1, 13–15.

[28] Heidi M. Hurd. 1996. The Moral Magic of Consent. *Legal Theory* 2, 2 (1996), 121–146. https://doi.org/10.1017/S1352325200000434

[29] Sarah J. Jackson and Sonia Banaszczyk. 2016. Digital Standpoints: Debating Gendered Violence and Racial Exclusions in the Feminist Counterpublic. *J. Commun. Inq.* 40, 4 (September 2016), 391–407. https://doi.org/10.1177/0196859916667731

[30] Sarah J Jackson and Brooke Foucault Welles. 2015. Hijacking #myNYPD: Social Media Dissent and Networked Counterpublics. *J. Commun.* 65, 6 (November 2015), 932–952. https://doi.org/10.1111/jcom.12185

[31] John Kleinig. 1982. The Ethics of Consent. *Can. J. Philos.* 12, sup1 (January 1982), 91–118. https://doi.org/10.1080/00455091.1982.10715825

[32] Issie Lapowsky. 2019. One Man's Obsessive Fight to Reclaim His Cambridge Analytica Data. Retrieved October 5, 2020 from https://www.wired.com/story/one-mans-obsessive-fight-to-reclaim-his-cambridge-analytica-data/

[33] Peter G. Leonard. 2018. Emerging Concerns for Responsible Data Analytics: Trust, Fairness, Transparency and Discrimination. *SSRN Electronic Journal* 941 (April 2018), 151–168. https://ssrn.com/abstract=3154269

[34] Kevin Lewis, Jason Kaufman, and Nicholas Christakis. 2008. The Taste for Privacy: An Analysis of College Student Privacy Settings in an Online Social Network. *J. Comput.-Mediat. Commun.* 14, 1 (October 2008), 79–100. https://doi.org/10.1111/j.1083-6101.2008.01432.x

[35] Marco Lippi, Przemysław Pałka, Giuseppe Contissa, Francesca Lagioia, Hans-Wolfgang Micklitz, Giovanni Sartor, and Paolo Torroni. 2019. CLAUDETTE: an automated detector of potentially unfair clauses in online terms of service. *Artificial Intelligence and Law* 27, 2 (2019), 117–139.

[36] Marco Loos and Joasia Luzak. 2016. Wanted: a bigger stick. On unfair terms in consumer contracts with online service providers. *Journal of consumer policy* 39, 1 (2016), 63–90.

[37] Alice E Marwick and Danah Boyd. 2014. Networked privacy: How teenagers negotiate context in social media. *New media & society* 16, 7 (2014), 1051–1067.

[38] Tobias Matzner. 2014. Why Privacy is Not Enough Privacy in the Context of "Ubiquitous Computing" and "Big Data". *Journal of Information, Communication and Ethics in Society* 12, 2 (2014), 93–106. https://doi.org/10.1108/jices-08-2013-0030

[39] Gibson Mba, Jeremiah Onaolapo, Gianluca Stringhini, and Lorenzo Cavallaro. 2017. Flipping 419 Cybercrime Scams: Targeting the Weak and the Vulnerable. In *Proceedings of the 26th International Conference on World Wide Web Companion* (Perth, Australia) *(WWW '17 Companion)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 1301–1310. https://doi.org/10.1145/3041021.3053892

[40] Aleecia M McDonald and Lorrie Faith Cranor. 2008. The cost of reading privacy policies. *Isjlp* 4 (2008), 543.

[41] Hans-W Micklitz, Przemysław Pałka, and Yannis Panagis. 2017. The empire strikes back: digital control of unfair terms of online services. *Journal of consumer policy* 40, 3 (2017), 367–388.

[42] Helen Nissenbaum. 2004. Privacy as contextual integrity. *Wash. L. Rev.* 79 (2004), 119.

[43] Helen Nissenbaum. 2009. *Privacy in context: Technology, policy, and the integrity of social life.* Stanford University Press, Stanford, CA, USA.

[44] Jonathan A. Obar and Anne Oeldorf-Hirsch. 2017. Clickwrap Impact: Quick-Join Options and Ignoring Privacy and Terms of Service Policies of Social Networking Services. In *Proceedings of the 8th International Conference on Social Media & Society* (Toronto, ON, Canada) *(SMSociety17)*. Association for Computing Machinery, New York, NY, USA, Article 50, 5 pages. https://doi.org/10.1145/3097286.3097336

[45] Alessandro Oltramari, Dhivya Piraviperumal, Florian Schaub, Shomir Wilson, Sushain Cherivirala, Thomas B Norton, N Cameron Russell, Peter Story, Joel Reidenberg, and Norman Sadeh. 2018. PrivOnto: A semantic framework for the analysis of privacy policies. *Semantic Web* 9, 2 (2018), 185–203.

[46] Romualdo Pastor-Satorras, Claudio Castellano, Piet Van Mieghem, and Alessandro Vespignani. 2015. Epidemic processes in complex networks. *Rev. Mod. Phys.* 87, 3 (Aug 2015), 925–979. https://doi.org/10.1103/RevModPhys.87.925

[47] Siani Pearson and Prodromos Tsiavos. 2014. Taking the Creative Commons beyond copyright: developing Smart Notices as user centric consent management systems for the cloud. *International Journal of Cloud Computing 2* 3, 1 (2014), 94–124.

[48] Sandra Petronio. 1991. Communication boundary management: A theoretical model of managing disclosure of private information between marital couples.

Communication theory 1, 4 (1991), 311–335.

[49] Kevin Ponniah. 2020. How a Chinese agent used LinkedIn to hunt for targets. Retrieved Accessed: 2021-10-29 from https://www.bbc.com/news/world-asia-53544505

[50] Antoinette Rouvroy and Yves Poullet. 2009. The Right to Informational Self-Determination and the Value of Self-Development: Reassessing the Importance of Privacy for Democracy. In *Reinventing Data Protection?*, Serge Gutwirth, Yves Poullet, Paul De Hert, Cécile de Terwangne, and Sjaak Nouwt (Eds.). Springer Netherlands, Dordrecht, 45–76.

[51] Cristiana Santos, Midas Nouwens, Michael Toth, Nataliia Bielova, and Vincent Roca. 2021. Consent Management Platforms Under the GDPR: Processors and/or Controllers?. In *Privacy Technologies and Policy*, Nils Gruschka, Luís Filipe Coelho Antunes, Kai Rannenberg, and Prokopios Drogkaris (Eds.). Springer International Publishing, Cham, 47–69.

[52] Emre Sarigol, David Garcia, and Frank Schweitzer. 2014. Online Privacy as a Collective Phenomenon. In *Proceedings of the Second ACM Conference on Online Social Networks* (Dublin, Ireland) *(COSN '14)*. Association for Computing Machinery, New York, NY, USA, 95–106. https://doi.org/10.1145/2660460.2660470

[53] Bart Willem Schermer, Bart Custers, and Simone van der Hof. 2014. The Crisis of Consent: How Stronger Legal Protection May Lead to Weaker Consent in Data Protection. *Ethics Inf. Technol.* 16 (March 2014), 171–182. https://doi.org/10.1007/s10676-014-9343-8

[54] Paul M. Schwartz. 2000. Internet Privacy and the State. *Conn. L. Rev.* 32, 3 (May 2000), 815–859. https://doi.org/10.2139/ssrn.229011

[55] Michael Warren Skirpan, Tom Yeh, and Casey Fiesler. 2018. What's at Stake: Characterizing Risk Perceptions of Emerging Technologies. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3173644

[56] Robert H Sloan and Richard Warner. 2014. Beyond notice and choice: Privacy, norms, and consent. *J. High Tech. L.* 14 (2014), 370.

[57] Daniel J. Solove. 2004. *The Digital Person: Technology and Privacy in the Information Age.* New York University Press, New York, NY, USA.

[58] Dietrich Stauffer and Ammon Aharony. 2018. *Introduction to percolation theory.* Taylor & Francis, London.

[59] Brett Stone-Gross, Marco Cova, Lorenzo Cavallaro, Bob Gilbert, Martin Szydlowski, Richard Kemmerer, Christopher Kruegel, and Giovanni Vigna. 2009. Your Botnet is My Botnet: Analysis of a Botnet Takeover. In *Proceedings of the 16th ACM Conference on Computer and Communications Security* (Chicago, Illinois, USA) *(CCS '09)*. Association for Computing Machinery, New York, NY, USA, 635–647. https://doi.org/10.1145/1653662.1653738

[60] Sam Thielman. 2015. Surveillance reform explainer: can the FBI still listen to my phone calls. Retrieved Accessed: 2021-09-29 from https://www.theguardian.com/world/2015/jun/03/surveillance-reform-freedom-act-explainer-fbi-phone-calls-privacy

[61] Amanda L. Traud, Peter J. Mucha, and Mason A. Porter. 2012. Social Structure of Facebook Networks. *Physica A* 391, 16 (August 2012), 4165–4180. https://doi.org/10.1016/j.physa.2011.12.021

[62] Zeynep Tufekci. 2008. Can you see me now? Audience and disclosure regulation in online social network sites. *Bulletin of Science, Technology & Society* 28, 1 (2008), 20–36.

[63] Na Wang, Heng Xu, and Jens Grossklags. 2011. Third-Party Apps on Facebook: Privacy and the Illusion of Control. In *Proceedings of the 5th ACM Symposium on Computer Human Interaction for Management of Information Technology* (Cambridge, Massachusetts) *(CHIMIT '11)*. Association for Computing Machinery, New York, NY, USA, Article 4, 10 pages. https://doi.org/10.1145/2076444.2076448

[64] Peter Westen. 2017. *The logic of consent: The diversity and deceptiveness of consent as a defense to criminal conduct.* Routledge, Oxfordshire, England, UK.

[65] Christopher Wylie. 2019. *Mindf*ck: Cambridge Analytica and the Plot to Break America.* Random House, New York, NY.

[66] Yang Yang, Jianhui Wang, and Adilson E Motter. 2012. Network observability transitions. *Physical Review Letters* 109, 25 (2012), 258701.

[67] Shoshana Zuboff. 2019. *The age of surveillance capitalism: the fight for a human future at the new frontier of power.* Public Affairs, Boston, MA, USA.

A RESEARCH METHODS

A.1 Data.

We use network data from the anonymous Facebook100 [61] data set without any associated metadata and for the sole purpose of having realistic network structures from a social media platform. The original data set presents 100 complete and independent networks of Facebook "friendships" from 100 American colleges and universities collected as a single-day snapshot in September 2005. Figures 2 and 4 show simulations of our models independently on the 95 Facebook networks with more than 2000 nodes, showing the individual averages obtained from each set of parameters on each network.

A.2 Observability Model.

Our observability model runs on a directed (or undirected[3]) network of potential data flow where a link from i to j means that user j receives data from user i. We simulate an observability process by selecting a fraction φ of users whom a third party directly observes. For an observability process of depth $L = 0$, the simulation is now over, and a fraction φ of users have been observed. For an observability process of depth $L = 1$, all currently unobserved users whose data are received by directly observed users are now also observed (call those users the first generation of indirectly observed users). For an observability process of depth $L = 2$, all currently unobserved users whose data are received by the first generation of indirectly observed users are now also observed. While the model can be extended to any depth l, Figs. 2 and 4 both use $L = 2$. In these figures, we measure the fraction of observed individuals (directly and indirectly observed), the largest component of unobserved individuals connected through uninterrupted data flow, and the fraction of observed individuals with a given security setting.

A.3 Distributed Consent Model.

Our distributed consent passport constrains data flow in social media networks and the possibility of users being directly or indirectly observed. We define a general distributed consent model but implement it using only three distinct security settings: Users at level 0 share their data with all network neighbors and can be directly observed by a third party; users at level 1 share their data with all network neighbors but can *not* be directly observed by a third party; users at level 2 only share their data with neighbors at level 1 or 2 and can *not* be directly observed by a third party. Security levels are randomly assigned to nodes in a network (uniformly at random) at the start of every run of the model with the following probabilities given an adoption frequency x of distributed consent: 2/3 of users are assigned to level 0, 1/3 - x are assigned to level 1, and x are assigned to level 2. In Fig. 2, a fraction $\varphi = 1\%$ of users at security level 0 are then directly observed, and the observability model then runs as defined above but with the directionality of data flow limited by security level 2.

A.4 Consent Passport Model.

We extend the distributed consent passport to a general multilayer network where users are part of multiple platforms. In Fig. 4 we do this by doubling the original network to obtain a two-platform system where social connections exist on two platforms at once. Given an adoption frequency y of consent passport, we force a fraction y of the fraction x of adopters of security level 2 to have level 2 on both platforms. On both platforms independently, we then distribute uniformly at random the remaining $(1 - y)x$ fraction of adopters of level 2 without the passport, then the 1/3-x fraction of level 1 users, and the 2/3 fraction of level 0 users. In Fig. 4, we then select $\varphi = 0.25\%$ of users at security level 0 on at least one platform to be directly observed. The observability model then runs as usual but indirectly observing the network neighbors j of observed node i if data flows from j to i on *at least* one platform.

A.5 Data and Code Availability.

All data and all codes for model implementation and figure replication are available from https://github.com/antoineallard/distributed-consent.