Studies in Philosophy Politics and Economics

John Gordon (2020) *When Should Data-Processing Agents Be Allowed to Collect Our Personal Information? The Case against Reliance on Individuals' Consent in the Age Of Big Data*, Studies in Philosophy, Politics and Economics, 2(1), 55-60

When Should Data-Processing Agents Be Allowed to Collect Our Personal Information? The Case against Reliance on Individuals' Consent in the Age of Big Data

John Gordon*

* Philosophy, Logic and Scientific Method, London School of Economics, Houghton Street, London, WC2A 2AE, UK. Email: john.jacques.gordon@gmail.com

Abstract

In discussing the type of individual consent which allows data-processing agents to collect personal information, most of the privacy literature implicitly assumes that individual consent can be a sufficient protection of the right to privacy. In this article, I argue against this assumption. I explain how Big Data analytics has made our privacy depend on the choices of others, by introducing the problem of cross-persons data aggregation. In this context, I show how any type of individual consent systematically fails to uphold our common right to privacy, unless coordinating action is taken.

Keywords: Privacy; Big Data; Consent

Privacy Self-Management (hereafter PSM) is the system whereby individuals make their own decisions on whether to consent to the collection of their information by data-processing agents. In this paper, I agree with Solove (2013) that PSM is an inadequate system to govern the processing of personal data, but for different reasons. Whereas he argues that consent is not reached meaningfully under PSM, I contend that individual consent - even if meaningful - is an insufficient protection of the right to privacy. I first summarise Solove's argument and that of his main critics. Then, I argue both have fallen behind recent developments in data analytics, which have shown that people's right to privacy can no longer be upheld by individual consent alone. I finish by discussing potential objections to my argument.

Solove argues that PSM is inadequate because it relies on privacy notices, allowing data-processing agents to collect data if individuals consent to them. However, he claims that these notices systematically fail to deliver morally transformative consent - to be understood as a consent that makes an otherwise immoral transaction moral (such as the collection of personal data). As shown in the next paragraph, his notion of morally transformative consent aligns with the requirements of the autonomous authorization model (hereafter, AA) (Solove 2013). AA requires that consent be given (1) freely, (2) intentionally, (3) by a competent

LSE Houghton St Press

decision-maker and (4) with appropriate understanding (Beauchamp 2009). Although many of his premises are implicit, an interpretation of his argument is the following:

> P1: Individuals have a right to privacy - defined as meaningful control over one's personal information - which should be protected.
>
> P2: An individual's morally transformative consent is the necessary and sufficient condition for data-processing agents to collect personal data without infringing on the right to privacy.
>
> P3: Morally transformative consent arises when the conditions outlined by AA (Beauchamp 2009) are fulfilled.
>
> P4: PSM, as a legal paradigm, does not allow individuals' consent to fulfil the conditions of the autonomous authorization model.
>
> ---
>
> C: PSM is an inadequate paradigm to determine the rules governing privacy.

1. Evaluating P4

Solove focuses on justifying the fourth premise. Firstly, the competence requirement of AA, (3), is often violated: individuals are subject to cognitive faults which impair their rational decision-making (Solove 2013: 1886) and discount the downstream harms of data disclosure (Solove 2013: 1891). Secondly, individuals are not able to fully comprehend the consequences of data disclosure, which corresponds to AA's condition of understanding, (4). For example, they rarely read and understand the privacy notices PSM relies on, and there are too many consent-requiring transactions to understand the subtleties of each (Solove 2013: 1883). It is also very difficult to assess the harms of data disclosure: downstream uses can be complex, and individual decision-making rarely accounts for the societal effects of large-scale data disclosure (Solove 2013: 1891). A key issue with assessing harm, which will be relevant for the rest of this paper, is the problem of aggregation. Pieces of data can be combined to reveal potentially harmful information - including the identity of individuals (Solove 2013: 1889).

Solove's defence of the fourth premise is authoritative. His arguments on cognitive faults are well-documented (Kahneman 2011; Acquisti *et. al.* 2015), and problems such as that of aggregation are empirically validated (Barbaro and Zeller 2006). However, most critics take issue with the third, rather than fourth, premise.

2. Evaluating P3

AA is criticised as both an unnecessary and insufficient model of morally transformative consent. A popular alternative is the Fair Transactions model, whereby morally transformative consent arises in a transaction which is deemed 'fair'[1]. The main criticism of AA is that it is insufficient because it ignores the bilateral nature of a consent transaction: if one party gains too much benefit over another, a consented transaction could still be deemed exploitative. By evaluating the permissibility of a transaction as a whole, this is what the Fair Transactions model addresses particularly well (Miller and Wertheimer 2009; Schermer *et. al.* 2014). Further criticisms are leveraged against AA: Miller & Wertheimer

[1] What constitutes a fair transaction is widely discussed, but goes beyond the scope of this paper, see Miller & Wertheimer 2009; Schermer et al. 2014.

(2009) argue that AA is unnecessary because it prevents insufficiently informed individuals from engaging in beneficial, but consent-requiring, transactions. Indeed, many consent transactions may not reach the stringent condition of full understanding, yet they argue that individuals should not necessarily be prevented from engaging in potentially beneficial transactions for that reason. Schermer *et. al.* (2014) also argue AA is too demanding of individuals. If individuals really had to fully understand each consent transaction, this would require too much time and effort. Ultimately it would make individuals consent simply by cognitive fatigue and would in fact be self-defeating.

By placing less emphasis on individual competence and more on the structural conditions (fairness) of the consent transaction, the Fair Transactions model addresses these points successfully. Nevertheless, I find that it fails to protect people's right to privacy for the same reason that AA fails. Namely, both these frameworks view morally transformative consent as a sufficient protection of people's right to privacy (hereafter, I will call such frameworks consent-reliant). However, I argue against this: I disagree with the second premise.

3. Evaluating P2

My objection rests on an extension of Solove's problem of aggregation, namely that it can also occur across persons. This emerged with Big Data analytics, and in particular the Cambridge Analytica-Facebook scandal. Cambridge Analytica started with a study of 58 000 volunteers who submitted psychometric tests and data including their Facebook likes (Kosinski *et. al.* 2013). This enabled Cambridge Analytica to establish correlations between an individual's Facebook likes and his psychological profile. After harvesting the Facebook information of millions of users, they were thus able to determine the psychological profile of those users based on their likes. In this case, they used it to design political ads to which the users would be most receptive, with evident success (Rosenberg *et. al.* 2017).

This scandal initially highlighted how Cambridge Analytica were able, mostly legally, to harvest the personal data of millions of Facebook users. This reinforced the idea that people's understanding of the downstream uses of their data should be enhanced (Solove 2013). More importantly however, this exemplified how large-scale data analysis can successfully aggregate data across persons. Just as an individual's data can be combined to reveal more information about himself (Solove 2013), data from different individuals can be aggregated to reveal more information about one of those individuals (in this case the psychological profile of the Facebook users). This principle is central to Big Data analysis, which tries to extrapolate from existing data and accurately predict personal information - ranging from consumer preferences (Bradlow *et. al.* 2017) to the predisposition to crime (Chan and Moses 2015).

In this context, the volunteers' decision to provide psychometric tests along with their Facebook likes imposed additional data disclosure onto the Facebook users, which corresponds to a violation of the users' right to privacy. Indeed, the users may have provided an appropriate form of consent, as all required conditions could have been met during the consent transaction. In addition, they may have consented knowing that data could be aggregated across persons, and the risks associated with that. Yet, their right to privacy was violated nonetheless. Because the users' personal information may or may not have been revealed depending on the choices of the volunteers, the users have lost a substantial amount of control over their data. Meaningful control over one's data is the definition of the right to privacy. Hence, if an individual's data is revealed through cross-persons aggregation (hereafter CPA), this can be seen as a violation of their right to privacy. This new inter-

dependency in privacy choices is an issue that consent-reliant frameworks cannot resolve, and P2 should be rejected

To illustrate this, I will show that even under idealised circumstances which satisfy both AA and the Fair Transactions model, the problem of CPA leads to violations of the right to privacy.

Imagine a game in which all players choose the utility-maximising level of privacy they would want to enjoy. They do so by balancing how much utility they derive from privacy and from the benefits of disclosing data. Let us assume perfectly rational agents and perfect information, so that all players understand how their data can be used and aggregated, and they also know what data others will reveal and how this can be aggregated with theirs. This satisfies the AA model. Say that the data-disclosing transactions are all deemed 'fair', so that the Fair Transactions model is also satisfied.

When each player reveals the utility-maximising amount of data, individuals with extremely low privacy-derived utility – or extremely high data disclosing-derived utility – will have shared large amounts of information (the volunteers). This will have imposed a negative externality unto individuals who would have enjoyed higher levels of privacy (the Facebook users). Indeed, they could choose to reveal the same amount of information, but have to accept the additional data disclosure imposed unto them. Or, they could adjust by revealing less information, but would then incur a loss in data disclosing-derived utility. In any case, the optimal choice they would have made, had other individuals not shared so much information, is no longer available, and they must incur a loss in utility. This equates to a violation of the right to privacy because it either restricts the usage of one's personal information or imposes additional data disclosure: it reduces control over one's data.

As such, morally transformative consent, even fulfilled under idealised assumptions, is unable to resolve the privacy violations imposed by CPA. In reality, the violation always consists of additional data disclosure, as people don't know what others will share and never adjust. The following objections will help clarify further why consent-reliant frameworks necessarily fail.

One can first object that an appropriate consent-reliant framework could actually resolve these types of privacy violations. Indeed, by the same logic as above, one could argue that the same problem arises under the simple problem of aggregation: additional data disclosure is also imposed on individuals, as some of their data may be revealed, which they did not reveal themselves. As such, they face the same constrained choice of either adjusting to this additional disclosure - by revealing less information to data-processing agents - or simply accepting it and incurring a loss in privacy-derived utility. Yet, the Fair Transactions model can resolve this: by requiring that the transaction be deemed fair, it could require that the loss of utility due to aggregation be adequately compensated by data-processing agents.

However, this comparison is unsuitable because CPA imposes an externality, while the problem of aggregation imposes a loss within the consent transaction. Indeed, this loss arises due to the capacity of data-processing agents - who are party to these transactions - to combine data. As such, a consent-reliant framework can appropriately address this through the rules governing that consent transaction. In contrast, the loss of utility arising from CPA is imposed unto a third party, outside of the consent transaction. Consent-reliant frameworks cannot account for this without coordinating action, such as imposing external constraints or incentives for the sake of this third party. It is therefore the nature of the loss as an externality which makes consent-reliant frameworks fail.

One may finally object that negative externalities do not necessarily invalidate consent-reliant frameworks. Indeed, the negative externalities of individual privacy choices are already well-documented, and Solove reviews them in his paper (2013). For example, the reduction in privacy that arises from data disclosure harms intellectual freedom (Richards 2008; Cohen 2013) and the functioning of democratic society (Schwartz 1999). Yet, none of these authors conclude that consent-reliant frameworks are inadequate. Additionally, there are positive externalities of data disclosure, such as data being used to improve consumer products, which could potentially counter-balance the negative externalities. As such, the fact that they allow some negative externalities is not a definitive argument against consent-reliant frameworks.

However, these other externalities do not harm (or benefit) the right to privacy, but a set of separate common goods (be it intellectual freedom or consumer products). Hence, arguments based on them undermine (or support) consent-reliant frameworks for the sake of these other goods, but not for the sake of privacy. In contrast, CPA renders consent-reliant frameworks fundamentally inadequate because they fail to protect privacy itself: their protection of the right to privacy at the individual level ignores the inter-dependencies which undermine the right to privacy at the aggregate level. With CPA, consent-reliant frameworks harm the very social good they are designed to protect: P2 should indeed be rejected.

To conclude, I have argued that PSM's – and the wider privacy literature's – reliance on morally transformative consent as a sufficient protection of the right to privacy is inadequate. I have showed that the CPA problem, introduced by Big Data analytics, has created inter-dependencies in our privacy choices. Consequently, relying on individual consent allows the choices of some individuals to constrain those of others, and fails to protect the right to privacy at the aggregate level. Ultimately, consent may still be at the core of an effective framework, but coordinating action - which accounts for third parties outside the consent transaction - will be necessary.

References

Acquisiti, A., Brandimarte, L. and Loewenstein, G. 2015. "Privacy and human behaviour in the age of information". *Science* 347(6221): 509-14.

Barbaro, M. and Zeller, T.Jr. 2006. "A Face Is Exposed for AOL Searcher No. 4417749". *The New York Times*. August 9 2006. URL: https://www.nytimes.com/2006/08/09/technology/09aol.html

Bradlow, E.T., Gangwar, M., Kopalle, P.K. and Voleti, S. 2017. "The role of Big Data and predictive analytics in retailing". *Journal of Retailing* 93(1): 79-95.

Chan, J. and Moses, B.L. 2016. "Is Big Data challenging criminology?". *Theoretical Criminology* 20(1): 21-39.

Cohen, J.E. 2000. "Examined lives: informational privacy and the subject as object". *Stanford Law Review* 52: 1373-1428.

Beauchamp, T. 2009. "Autonomy and consent". In *The Ethics of Consent: Theory and Practice*, edited by Miller, F. G. and Wertheimer, A. 55-78. Oxford: Oxford University Press.

Kahneman, D. 2011. *Thinking, Fast and Slow*. London: Penguin Press.

Kosinski, M., Stillwell, D. and Graepel, T. 2013. “Private traits and attributes are predictable from digital records of human behaviour”. *PNAS* 110(15): 5802-5805.

Miller, F.G. and Wertheimer, A. 2010. “Preface to a theory of consent: beyond valid consent”. In *The Ethics of Consent: Theory and Practice*, edited by Miller, F. G. and Wertheimer, A. 79-106. Oxford: Oxford University Press.

Richards, N.M. 2008. “Intellectual Privacy”. *Texas Law Review* 87: 387.

Rosenberg, M., Confessore, N, and Cadwalladr, C. 2018. “How Trump Consultants Exploited the Facebook Data of Millions”. *The New York Times*. March 17 2018. URL: https://www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html

Schermer, B.W., Custers, B. and Van Der Hof, S. 2014. “The crisis of consent: how stronger legal protection may lead to weaker consent in data protection”. *Ethics and Information Technology* 16(2): 171-182.

Schwartz, P.M. 1999. “Privacy and democracy in cyberspace”. *Vanderbilt Law Review* 52: 1609-1613.

Solove, D. 2013. “Privacy self-management and the consent dilemma”. *Harvard Law Review* 126: 1880-1903.