CUSPE Cambridge University Science and Policy Exchange

CJSP Cambridge Journal of Science & Policy

Rethinking Informed Consent in the Context of Big Data

Communication | Editorial | Invited contribution | **Perspective** | Report | Review

Anna Bruvere
Department of Philosophy, Logic and Scientific Method
London School of Economics

Victor Lovic
Department of Physics
Imperial College London

Abstract

A widely accepted method for addressing digital privacy concerns is the use of informed consent: asking users to agree to privacy policies and consent to the use of their personal data. This approach has come under strain with the emergence of "big data" in which large datasets are collected and analysed. This paper argues that since individuals do not understand or even read the privacy policies they agree to, informed consent ultimately fails to protect privacy. Following the work of Solon Barocas and Helen Nissenbaum, this paper proposes an updated definition of informed consent and argues that the responsibility of protecting privacy should be shifted from individuals to organisations.

Science ⇒ Policy

Informed consent currently consists in agreeing to long privacy policies which few people read, and even fewer understand. To ensure the privacy of individuals is respected, it is necessary to rethink our conception of informed consent. The burden of guaranteeing privacy should be shifted from individuals to organisations.

Keywords Privacy · Informed Consent · Big Data

Introduction

Big data refers to the collection and analysis of large datasets for the purposes of finding patterns or otherwise extracting value from them [1, 2]. For example, private corporations collect data about their users to generate tailored product recommendations or to provide targeted advertisements [3, 4]. Big data is a promising avenue to improve many aspects of our lives, from increasing our scientific and social knowledge, helping improve our healthcare systems, and developing our public decision-making processes. However, when making use of personal information, big data can conflict with the privacy rights of individuals.

One method for addressing privacy concerns is the use of informed consent; asking individuals to consent to the collection and use of their personal information. In this way, individuals may waive their privacy rights and allow for their personal information to be used under certain specified conditions. Following the work of Solon Barocas and Helen Nissenbaum this paper argues that, as currently conceived, informed consent is problematic and does not succeed in protecting individuals' privacy [5].

Informed Consent in the Context of Social Norms

Informed consent typically takes the form of asking users to agree to privacy policies which specify what information will be recorded and in what ways it will be used. These privacy policies are long documents which exhaustively list all the ways in which the personal data might be used.

The problem with this "notice and consent" approach to privacy lies in the fact that very few people read privacy statements. Furthermore, there is abundant evidence that even the few people who do read privacy policies do not understand them [6]. Ticking boxes under long and incomprehensible privacy policies does not constitute informed consent.

The think tank Brookings points out that while "the shortcomings of consent are often acknowledged, the response is often a push for more and better consent" [7]. Clearly, longer and more thorough privacy policies are not the answer. One of the most widely discussed alternatives to the "notice and consent" approach is shifting the focus from the collection of personal data to its usage [8]. This puts the burden on companies, rather than individuals, to protect privacy. Brookings argues that companies should be restricted to using user data for "legitimate purposes [...] consistent with reasonable expectations formed in their relationships with [users]." However, for uses of personal data outside of these reasonable expectations, informed consent is still required. The question then becomes, 'which actions require informed consent, and which do not?'

Addressing this issue in the context of healthcare, bioethicists Neil Manson and Onora O'Neill [9] argue that informed consent should be understood with reference to a background of social norms, ethical standards and legal obligations. In this way, consent is only relevant to cases which depart from these norms. Barocas and Nissenbaum argue that this idea helps to explain why our current approach to consent fails. By asking users to consent to an exhaustive list of all the different ways their personal data might be used, the real privacy issues at stake are drowned out by irrelevant detail. For example, consider the following excerpts from the privacy policies of Facebook, Google, and Snap (formerly Snapchat) Inc. respectively:

"[...] when you use Messenger or Instagram to communicate with people or businesses, those people and businesses can see the content you send" [10]

"[...] if you contact Google, we'll keep a record of your request in order to help solve any issues you might be facing." [11]

"When you interact with our services, we collect information that you provide to us." [12]

The listing of such details leads to lengthy privacy policies that are difficult to understand, even for users who are intent on reading them. By following the principle described by Manson and O'Neill, that consent only applies to uses of information that deviate from expected social norms, privacy policies would be drastically shortened and would focus on the real issues that individuals are consenting to.

This leaves open the question of specifying what the norms are. This paper argues that the government should determine and specify the norms of information usage. In the UK, the Information Commissioner's Office (ICO, an independent public body sponsored by the Department for Digital, Culture, Media & Sport) is tasked with offering guidance, advice and promoting good practice for data protection. They could compile a list of social norms and expected behaviours to serve as a standard against which privacy policies can be formulated. The UK government already uses this approach within the context of healthcare – the National Data Guardian (sponsored by the Department of Health and Social Care) provides guidelines as to what constitutes a "reasonable expectation" when sharing healthcare data [13]. These guidelines can be extended into other domains relevant to privacy. The burden then lies with each company or organisation seeking the consent of its users to explain the ways in which their use of personal information deviates from the guidelines set out by these public bodies. This will alleviate the problem of long and incomprehensible privacy policies while ensuring that users are informed and are able to give consent. Although the line between reasonable and unreasonable cannot be sharply defined, an attempt must be made. Disagreements that arise from this (necessarily) vague definition can be solved through the justice system.

Finally, this paper considers a possible objection: by having a public body compile a list of privacy norms, the problem of lengthy and incomprehensible privacy policies is simply shifted from private companies to the government. This objection fails for three reasons: first, the list compiled by the government is aimed at organisations rather than individuals. Its purpose is to serve as a reference against which privacy policies can be drafted. Organisations have more resources to devote to understanding lengthy policies and furthermore, they only have their own policy to consider. On the other hand, individuals can be asked to agree to different privacy policies on a near daily basis, so it makes sense to put the burden on organisations. Second, the government's list of privacy norms should be in line with the reasonable expectations of the wider public. This could be ensured by polling public opinion; for example, the ICO makes use of a "citizen reference panel" to gain insight into people's attitudes to information rights issues [14]. Third, a government list would represent a single source of information, applicable to all organisations and individuals. This would be a significant improvement to the current situation where each company compiles its own lengthy privacy policy.

Conclusion

With the advent of the digital age and big data comes new challenges to protect the privacy of individuals. In particular, informed consent has been misused and overused by organisations by asking users to agree to lengthy privacy policies which most do not understand or read. This article has argued that the focus of privacy should be shifted from the collection of information to its usage, putting the burden on organisations rather than individuals to protect privacy rights. Organisations handling personal data should be required to comply with a set of norms and "reasonable expectations" set out by a government agency like the Information Commissioner's Office (ICO) in the UK. It is then the responsibility of organisations to follow these guidelines where possible, and otherwise seek the consent of users when the usage of their personal data goes beyond reasonable expectation. This will result in shorter privacy policies, focused on the real privacy issues at stake so that users are able to make informed decisions about their personal and private data.

Acknowledgments

The authors wish to thank the anonymous reviewers for their excellent suggestions.

References

[1] SAS, "Big data: What it is and why it matters," N.D. [Online]. Available: http://bit.ly/37dmAyN

[2] T. Segal, "Big data," July 2019. [Online]. Available: https://www.investopedia.com/terms/b/big-data.asp

[3] B. Marr, "Amazon: Using big data to understand customers," 2020. [Online]. Available: http://bit.ly/3ajnfAp

[4] J. Peterson, "A guide to google analytics behavioural targeting," 2013. [Online]. Available: https://www.scripted.com/seo/guide-google-analytics-behavioral-targeting

[5] S. Barocas and H. Nissenbaum, "Big data's end run around anonymity and consent," in *Privacy, Big Data and the Public Good*, S. B. J. Lane, V. Stodden and H. Nissenbaum, Eds. New York: Cambridge University Press, 2014, pp. 44–76.

[6] F. M.-W. Y. Bakos and D. R. Trossen, "Does anyone read the fine print? testing a law and economics approach to standard

form contracts," 2009. [Online]. Available: https://bit.ly/3kEzBGP

[7] D. Medine and G. Murthy, "Companies, not people, should bear the burden of protecting data," December 2019. [Online]. Available: http://brook.gs/3pfdcAB

[8] F. H. Cate and V. Mayer-Schönberger, "Notice and consent in a world of big data," *International Data Privacy Law*, vol. 3, no. 2, 2013. [Online]. Available: Available: https://doi.org/10.1093/idpl/ipt005

[9] N. C. Manson and O. O'Neill, *Rethinking Informed Consent in Bioethics.* Cambridge: Cambridge University Press, 2012. [Online]. Available: https://doi.org/10.1017/CBO9780511814600

[10] Facebook, "Data policy," April 2018. [Online]. Available: https://www.facebook.com/policy.php

[11] Google, "Privacy policy," March 2020. [Online]. Available: https://policies.google.com/privacy

[12] S. Inc., "Privacy policy," December 2019. [Online]. Available: https://www.snap.com/en-US/privacy/privacy-policy

[13] N. D. Guardian, "Sharing patient data: exploring consensus on reasonable expectations," November 2017. [Online]. Available: http://bit.ly/302VXZe

[14] ICO, "Citizen reference panel," May 2016. [Online]. Available: https://ico.org.uk/about-the-ico/research-and-reports/citizen-reference-panel/

About the Authors

Anna is an MSc Philosophy and Public Policy student at the London School of Economics. She is interested in technology and environmental policy and in new ethical challenges posed by emerging technologies. Previously, she finished an MA in Philosophy at University College London and an MA in English & Philosophy at the University of Glasgow. She can be contacted at `a.bruvere@lse.ac.uk/anna`.

Victor is a PhD student in the Department of Physics at Imperial College London. His research is in the field of quantum cryptography, investigating the security of quantum key distribution and quantum random number generation. He is interested in the potential of emerging technologies to change the world, in good or bad ways, and thinks that good policymaking is key to ensuring good outcomes and mitigating any risks. Previously, Victor studied Physics at the University of Glasgow He can be contacted at `v.lovic19@imperial.ac.uk`.

Conflict of interest The Authors declare no conflict of interest.