

Protecting health privacy through reasonable inferences

Brent Mittelstadt¹

In the digital age individuals face key choices about whether and how to share intimate details of their lives, “images of the body, biological data in general and diagnostic information” (Pyrrho et al. 2022). These are ultimately choices about privacy; or “the control that a person has over the extent, nature and use of their personal information.” The stakes of privacy in health practices in the digital age are “a space of freedom of choice and self-determination of the subjects” (Pyrrho et al. 2022).

In their article ‘Privacy and Health Practices in the Digital Age’, Pyrrho et al. rightly recognise a tendency to oversimplify privacy in health, portraying it as an inevitable conflict between individual and collective interests. This approach ignores how individual privacy interests, for example freedom of personal choice in disclosing intimate details of one’s life, can promote collective interests by enabling communal belonging and democracy. But their treatment of collective interests misses a key type of collective privacy interest which equally deserves attention.

Pyrrho et al. discuss two distinct categories of collective interests. The first is public welfare, where a conflict between individual and collective interests is seemingly unavoidable in many cases, seen for instance in refusals to share individual medical data for public health surveillance purposes (Mittelstadt et al. 2018). The second type of collective interest considered is community belonging, in particular belonging to social and political organizations. In these terms privacy can be seen as a fundamental enabler of democratic societies; this much has been reflected in the treatment of privacy as a special type of human right that enables the exercise of free expression and other human rights (Wachter 2017). This type of collective interest is at the heart of Pyrrho et al.’s conclusion that privacy cannot be viewed solely as a conflict between individual and collective interests.

What is missing from this discussion of collective interests is the existence of collective or group privacy interests (Bloustein 1976; Mantelero 2016; Mittelstadt 2017). The authors seemingly recognise group privacy interests to some extent, saying “the effects of a breach of privacy are not solely personal. Geolocalized identification of a concentrated incidence of diseases, even those which are not communicable, can generate perverse collective effects of stigma and discrimination” (Pyrrho et al. 2022). They rightly conclude that mere de-identification is insufficient to protect against such effects which target groups (in their example, people residing in a particular geographical location) rather than unique (identifiable) individuals.

But this example misses the potential for individual privacy interests to enhance the privacy of a group; the refusal to share genetic data, for example, can protect the privacy of relatives. Ascriptive and ad hoc or algorithmically assembled groups can similarly benefit from individual members making privacy enhancing choices about data sharing and usage (Mittelstadt 2017). Group privacy interests have taken on a new significance in the age of surveillance capitalism (Zuboff 2019), where learning

¹ Corresponding author. Oxford Internet Institute, University of Oxford, 1 St. Giles, Oxford, OX1 3JS, United Kingdom. E-mail: brent.mittelstadt@oii.ox.ac.uk. This work has been supported through research funding provided by the Wellcome Trust (grant nr 223765/Z/21/Z), Alfred P. Sloan Foundation (grant nr G-2021-16779), the Department of Health and Social Care (via the AI Lab at NHSx), British Academy Postdoctoral Fellowships (grant nr PF2\180114 and grant nr PF\170151), Luminate Group, and the Miami Foundation.

technologies can assemble extremely granular groups at scale, derive knowledge about them, and make consequential decisions (Mittelstadt 2017; Wachter and Mittelstadt 2019).

‘Inferential analytics’, or technologies designed to find small patterns and give meaning to large and diverse datasets, can draw highly invasive inferences about intimate details of private life. Numerous cases have come to light (Wachter and Mittelstadt 2019). Recent research has found that voice input data collected by Amazon smart speakers are used to draw sensitive inferences about users for advertising purposes (Iqbal et al. 2022). Facebook, operating on a public health justification, openly promotes its usage of artificial intelligence (AI) to detect and intervene on posts suggesting suicidal intent (TechCrunch 2017; Gomes de Andrade et al. 2018). A variety of other health conditions are alleged to be detectable and predictable on Internet platforms including depression on Twitter (Nadeem 2016), flu outbreaks on Google (Olson et al. 2013; Lazer et al. 2014), and neurodegenerative disorders on Microsoft Bing (White et al. 2018). The accuracy of these methods is a separate and important question, but also arguably irrelevant so long as their results are treated as reliable or actionable.

Making informed choices about personal data in the digital age is undoubtedly difficult. Pyrrho et al. argue that free, meaningful, and well-informed choices about when and how to share intimate health information are difficult in the digital age. They trace this difficulty to opacity and asymmetry in information flows, or “because the flow of information does not allow to foresee what the uses of this data will be. This dimension of unpredictability of future uses resulting from a present lack of control in the flow of information” is recognised as a key challenge for data subjects and health professionals in the digital age (Pyrrho et al. 2022).

Attributing this uncertainty to the opacity and asymmetry of information flows is correct, but ultimately incomplete. Opaque and imbalanced information flows undoubtedly create uncertainty about when and how intimate data is used, but so too do the analytic technologies used to make sense of data in the digital age.

Inferential analytics are opaque, highly complex, asymmetrical, and ultimately unpredictable. Inferences take the form of assumptions about a person’s past behaviour and personal characteristics, or predictions about future behaviour. It is impossible to predict with certainty the patterns, correlations, or novel ad hoc groupings that will be discovered and applied to individual subjects and groups. Inferences can be drawn that are privacy-invasive, sometimes counterintuitive and, in any case, cannot be verified at the point they are drawn. While subjects are often unable to predict, understand or refute intimate inferences drawn about their lives, they nonetheless impact on identity, reputation, and self-determination (Wachter and Mittelstadt 2019).

Privacy risks of the digital age do not stem solely from how personal data is collected, observed, or intentionally shared. Rather, the ability to draw intimate inferences about people from their data arguably poses the greatest risk. These inferences determine how they are viewed, evaluated, and ultimately treated by third parties. Pyrrho et al. (2022) recognise the critical risks of unintended disclosure of sensitive health information, which can include algorithmically generated inferences about health, noting that genetic or diagnostic data can compromise a person’s “chance of getting a job, health plan coverage, or financing a home (Roessler and Mokrosinska 2013).”

Complicating matters, these risks are not limited to inferences drawn from medical data sources. Rather, seemingly benign datasets can be combined and used to draw a variety of privacy-invasive inferences. Health-related inferences can be drawn from seemingly irrelevant data. Weather patterns, for example, could be used to infer future illness based on geolocation (Purtova 2018). Inferences

ultimately pose the greatest risk, and the technologies responsible for generating them are all too often opaque, complex, inaccessible, and immune to challenge.

Regulations designed to provide oversight and control over how data is collected and processed are therefore insufficient on their own. Regulating data flows and collection is not enough. Rather, to address the privacy risks posed by inferential analytics in health, regulations must provide meaningful protection against not only the inputs, but the outputs of data processing. Protections are needed that enable individuals to constrain how their data is used and re-used even when de-identified and aggregated, and not solely at the point of collection or anonymization (Wachter and Mittelstadt 2019).

In Europe, data protection regulation currently fails in this regard. The General Data Protection Regulation (GDPR) provide little protection against the novel risks of inferential analytics. Compared to other types of personal data, inferences receive the least protection. To ensure individuals can meaningfully exercise meaningful choice over sharing intimate health data in the digital age, I have argued elsewhere that a new data protection right is required, a ‘right to reasonable inferences’ (Wachter and Mittelstadt 2019).

In cases where algorithms draw ‘high-risk inferences’ about individuals, such as medicine and health practices, this right would require ex-ante justification to be provided by the data controller to establish whether an inference is reasonable. This disclosure would address (1) why certain data is a relevant basis to draw inferences; (2) why these inferences are relevant for the chosen processing purpose or type of automated decision; and (3) whether the data and methods used to draw the inferences are accurate and statistically reliable. The ex-ante justification is bolstered by an additional ex-post mechanism enabling unreasonable inferences to be challenged (Wachter and Mittelstadt 2019).

Going forward, data protection regulations need to look beyond limitations on data collection and pay greater attention to how, why, and for which purposes data is being processed over its lifecycle. Just as it was necessary to create a ‘right to be forgotten’ in a Big Data world, it is now essential to create a ‘right of how to be seen’ in the age of inferences.

References

- Bloustein, E. J. 1976. Group Privacy: The Right to Huddle. *Rutgers Camden Law Journal* 8: 219.
- Gomes de Andrade, N. N., D. Pawson, D. Muriello, L. Donahue, and J. Guadagno. 2018. Ethics and artificial intelligence: suicide prevention on Facebook. *Philosophy & Technology* 31(4). Springer: 669–684.
- Iqbal, U., P. N. Bahrami, R. Trimananda, H. Cui, A. Gamero-Garrido, D. Dubois, D. Choffnes, A. Markopoulou, F. Roesner, and Z. Shafiq. 2022. Your Echos are Heard: Tracking, Profiling, and Ad Targeting in the Amazon Smart Speaker Ecosystem. *arXiv:2204.10920 [cs]*.
- Lazer, D., R. Kennedy, G. King, and A. Vespignani. 2014. The Parable of Google Flu: Traps in Big Data Analysis. *Science* 343(6176): 1203–1205. doi: 10.1126/science.1248506.
- Mantelero, A. 2016. Personal data for decisional purposes in the age of analytics: From an individual to a collective dimension of data protection. *Computer Law & Security Review* 32(2): 238–255. doi: 10.1016/j.clsr.2016.01.014.
- Mittelstadt, B. 2017. From Individual to Group Privacy in Big Data Analytics. *Philosophy & Technology* 30(4): 475–494. doi: 10.1007/s13347-017-0253-7.
- Mittelstadt, B., J. Benzler, L. Engelmann, B. Prainsack, and E. Vayena. 2018. Is there a duty to participate in digital epidemiology? *Life Sciences, Society and Policy* 14(1): 9.
- Nadeem, M. 2016. Identifying Depression on Twitter. *arXiv:1607.07384 [cs, stat]*.
- Olson, D. R., K. J. Konty, M. Paladini, C. Viboud, and L. Simonsen. 2013. Reassessing Google Flu Trends Data for Detection of Seasonal and Pandemic Influenza: A Comparative Epidemiological Study

- at Three Geographic Scales. *PLOS Computational Biology* 9(10). Public Library of Science: e1003256. doi: 10.1371/journal.pcbi.1003256.
- Purtova, N. 2018. The law of everything. Broad concept of personal data and future of EU data protection law. *Law, Innovation and Technology* 10(1): 40–81. doi: 10.1080/17579961.2018.1452176.
- Pyrrho, M., L. Cambraia, and V. F. de Vasconcelos. 2022. Privacy and Health Practices in the Digital Age. *The American Journal of Bioethics* 0(0). Taylor & Francis: 1–10. doi: 10.1080/15265161.2022.2040648.
- Roessler, B., and D. Mokrosinska. 2013. Privacy and social interaction. *Philosophy & Social Criticism* 39(8). SAGE Publications Ltd: 771–791. doi: 10.1177/0191453713494968.
- TechCrunch. 2017. Facebook rolls out AI to detect suicidal posts before they're reported. *TechCrunch*.
- Wachter, S. 2017. *Privacy: Primus Inter Pares — Privacy as a Precondition for Self-Development, Personal Fulfilment and the Free Enjoyment of Fundamental Human Rights*. SSRN Scholarly Paper ID 2903514. Rochester, NY: Social Science Research Network.
- Wachter, S., and B. D. Mittelstadt. 2019. A right to reasonable inferences: re-thinking data protection law in the age of Big Data and AI. *Columbia Business Law Review* 2019(1).
- White, R. W., P. M. Doraiswamy, and E. Horvitz. 2018. Detecting neurodegenerative disorders from web search signals. *npj Digital Medicine* 1(1): 8.
- Zuboff, P. S. 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power: Barack Obama's Books of 2019*. Main edition. London: Profile Books.