PRIMER II: ON PRIVACY AND CONTACT TRACING

The concept of privacy is not an intrinsic feature of all technical and legal systems. In the United States, for example, the idea of privacy developed during the Reconstruction Era (1865–1877) as Americans confronted new aspects of modern life and social organization after the Civil War and the declared end of racial slavery.¹ Historically, claims to privacy helped define citizenship in the US, and elsewhere, by establishing who could demand freedom from intrusion and scrutiny, and under what circumstances.² Conversely, citizens who were denied privacy were denied full social standing. As a corrective to this distinct form of inequality, the United Nations included an explicit “right to privacy” as Article 12 of the Declaration of Human Rights (1948).³ In the intervening decades, with the emergence of technological developments that access and deploy personal information in new ways, concerns arose about digital privacy, in particular. More recently, in 2013, the UN General Assembly passed resolution 68/167, “the right to privacy in the digital age.”⁴ By 2015, the first UN special rapporteur for privacy had been appointed “to study trends, developments and challenges in relation to the right to privacy and to make recommendations to ensure its promotion and protection, including in connection with the challenges arising from new technologies.”⁵

Privacy (or lack thereof) not only defines civic membership, it also drives inequality. For example, the Allegheny Family Screening Tool, an algorithm used by the state of Pennsylvania, employs private, personal information about parents’ use of public services and compliance with directives from health and social workers to allegedly predict risk of child abuse or neglect.⁶ This system subjects poor families to greater surveillance and punitive measures. Similarly, “predictive policing” amplifies longstanding patterns of bias rather than predicting crime, resulting in disproportionate surveillance of Black and brown communities.⁷ Such surveillance via interactions with punitive institutions can suppress support-seeking from other vital social institutions such as the public health and medical care systems.⁸

Contact tracing, a public health process that requires the collection and use of personal information to identify and assess people who may have been exposed to an infectious disease,⁹ is inherently about privacy.¹⁰ In countries in which contact tracing does not operate with the backing of the state (and its sanctioning powers), or in which there is relatively low trust in institutions, there are added concerns of state surveillance. Many aspects of contact tracing systems may raise privacy concerns, including: (1) what information is collected and
by whom, (2) how long the information is retained, (3) who has access to the information, (4) how the information is used, and (5) how much agency individuals have within the system of contact tracing.

Contact tracing requires the collection of personal information that may be sensitive, including health status, personal contacts, and daily whereabouts. Revealing one’s health status may make one vulnerable to discrimination across a variety of social and economic institutions. Revealing interpersonal relationships can create perceptions of “guilt by association.” Individuals identified through contact tracing could be labeled as “at risk” without their knowledge, and aggregating contact lists could index risk scores for social groups. Social network data can become the scaffolding to make inferences about other sensitive traits with high accuracy, amplifying potential harm. Place carries cultural meaning, and is highly structured by historical social and political processes. Like network data, longitudinal and temporal location information can be used to make inferences about its subjects, including their daily rhythms, income, race, and potentially stigmatizing information. Further, publishing information about transmission “hot spots” could stigmatize community sites and businesses. Contact tracing itself can create sensitive information—for example, whether one does or does not install a contact tracing app.

Decisions to share information are highly contextual—one shares specific information at a given time, to a specific entity, for a given purpose. Failing to account for and respect that context can violate privacy and undermine trust. Which organizations facilitate contact tracing affect whether people are willing to share their information, and which organizations people trust may vary within a population. While sharing information with additional actors may differ from the context the subject consented to, it happens often, as demonstrated by an app developed by Alibaba Group that appears to share information with the police. Information can also be improperly shared through hacks or data leaks. Data management and analysis can violate initial context. For example, information could be retained for longer than is necessary for intended public health interventions. Analytical techniques could reveal additional information about subjects including their identities, and could link to other administrative records. Data can misrepresent people, and information collected via Bluetooth, GPS location, or QR codes can all create inaccurate information about one’s behavior or potential exposure.

Data sharing and sophisticated analytical technology invite mission creep, the process by which an intervention expands to purposes other than its original scope, which may also violate the context initially intended. For example, data collected through public health surveillance may be valuable to advertisers for targeted advertising and price discrimination. The resale of ZIP-Code-level fever data by the smart thermometer company Kinsa demonstrates
the potential for simultaneous surveillance and profit, which could be amplified if technology like wearables prove useful for public health surveillance and are harnessed by advertising platforms. Responses to public health and other crises often involve measures and interventions that would not be otherwise possible, through which power may be centralized and individual liberties curtailed. To prevent the magnitude of potential harm via loss of privacy, academics and nonprofit organizations alike have called for data collection to be proportional to the seriousness of the crisis, be limited to what is necessary to achieve public health objectives, cease when the crisis abates, and be scientifically justified.

The people and communities who share information are at the heart of understanding privacy. Powerful social actors, like governments and large companies, tend to have more control over the systems and programs that affect privacy and more knowledge of how those systems work. Digital data collection tends to be more covert and continuous than historic forms of surveillance, with more seamless, greater accumulation of information than in the past. Though many develop an understanding of these systems, the scale and structure of digital data collection can obscure privacy-relevant information from the people who are subject to observation, creating an imbalance of power that inhibits their participation in their own privacy. Creating more transparent systems can both improve those systems and create more public trust. The agency people have over what information they do or do not share is also central to enabling privacy, but binary opt-in/opt-out structures do not create genuine agency when opting-in is necessary for daily life, or if opting out is met with social sanctions. For example, many Indian citizens have been met with the decision to install the Aarogya Setu Covid-19 app or face fines or job loss. Technical infrastructures can be changed to grant users more control over their information, and citizens can be included in the governance of these technologies.

Lacking real agency, people may change their behavior in search of privacy and autonomy. Responses to crises often involve measures that would not be possible in the absence of a crisis, which can consolidate state or corporate power and curtail individual liberties. Given that surveillance has historically, disproportionately targeted vulnerable groups, some people may have legitimate reasons to hide information, for example, in attempts to avoid undeserved or disproportionate nonlegal punishment. Beyond action within particular systems, surveillance can also suppress agency in other ways; for example, it can discourage free speech and free association. Configuring systems of information collection and surveillance involves political and normative judgments that allocate power, access, cost, and opportunity across actors, often unevenly. Though the goals of contact tracing—to identify and assess people who may have been exposed to an infectious disease—is consistent, methods of contact tracing vary in form. Given the availability and potential of new technologies, states are also exploring the use of digital methods, in part or in full. Manual contact tracing involves human
actors contacting those possibly exposed to infection. *Digital-hybrid* contact tracing involves human tracers using technology to locate contacts, with the work guided by human interaction and decision-making. *Automated* contact tracing uses technology to find the contacts and notify people about exposure automatically. Among the most common methods proposed for this pandemic, manual contact tracing has the longest history.

Manual contact tracing involves identifying and interviewing a person who is carrying a virus to identify those with whom that person had recently been in contact, informing those contacts, and repeating the process. Manual intervention has been effective at controlling the spread of Ebola, SARS, HIV, and other infectious diseases, but has limitations as it can be time-consuming, requires some expertise and therefore training contact tracers, and relies on human memory. The reliance on individuals to report their histories can give individuals more agency over their information, but sometimes to the detriment of containing a disease. In New York City, for example, it is estimated that only two out of every five people with the virus shared information with contact tracers. Nonetheless, manual contact tracing does not preclude privacy risk. Information collected can be disclosed to other people and institutions, creating material harm and undermining trust. The imbalance of power is especially apparent when individuals are required to participate in manual contact tracing efforts.

Digital technologies have created new methods for public health surveillance and inspired hope that these new methods may improve health outcomes. In contact tracing, many hope that combining traditional manual methods of contact tracing with digital records will result in more accurate information, as digital records supplement human memory, and individuals can review their own data and identify potential false positives. Like fully manual tracing, the digital-hybrid contact tracing poses data sharing, reuse, and overreach risk. These risks may be heightened through technical infrastructure, as detailed diaries and storing data along with identifying information like phone numbers can make it easier to identify people, link to other records, and target those individuals. Sharing information may be mandated—in Thailand and Malaysia, participation in digital-hybrid contact tracing was a requirement to end lockdowns. In Singapore, participating in digital-hybrid contact tracing is required to enter places like supermarkets, schools, and offices. Surveillance across sites of daily life has the potential to reshape behavior as individuals seek privacy. Though digital-hybrid contact tracing promises higher accuracy than manual or fully automated methods, these methods are not error-free. For example, one approach to digital-hybrid tracing involves scanning QR codes when one enters buildings. However, two people who've scanned the same QR code...
within a short window may not have come into contact. Such methods create the potential for more privacy from digital trackers, but also potential for misinterpretation.

Given the enthusiasm for digital approaches to public health, many governments and public health authorities have also explored the potential of fully automated approaches to contact tracing. Automated contact tracing (also called exposure notification) uses technologies like smartphones or wearables to passively record where someone has been or with whom they have been in contact, and uses that information to identify potential contacts if that person becomes infected with the novel coronavirus. Proponents of automated approaches argue they have the potential to scale to large populations quickly, identify at-risk individuals quickly, and could be more accurate than human memory. They also argue data may be better kept confidential than in manual contact tracing systems. However, automated systems may create the same harm at greater scale. Indeed, the granularity, extensiveness, and scale of digitally collected data is a double-edged sword, as it enables broader surveillance and amplifies potential privacy intrusions. Detailed and comprehensive data may be easier to reidentify, and centralized data may be easier to repurpose. Given measurement issues with Bluetooth and GPS, small errors can have an outsized effect in predictive models or contact identification and can therefore erode public trust. Automated systems do not require active participation by those surveilled, and so passive or hidden data collection may also invite data and infrastructure use past the original intent of these systems. Automated systems may also collect more information about their users than is necessary—not only because they can record granular data, but also because researchers lack conclusive evidence that digital-hybrid and automated systems lead to more effective interventions than manual contact tracing. Decentralized contact tracing apps, which store information on individuals’ phones rather than a central database, may abate many concerns about privacy at the expense of additional data for public health surveillance. Some are concerned that decentralized systems could also enable surveillance from attackers and introduce data integrity issues. Like other forms of contact tracing, use of automated technologies may be mandated directly or as a condition of participation in daily life, undermining consent. Automated approaches may also exclude the most vulnerable populations, as less than half of the world’s population has a smartphone, and about 25 percent of smartphones, do not have the software necessary to run automated contact tracing.

There is no clear evidence that automated contact tracing is more effective than manual and digital-hybrid methods. If true, automated methods could constitute overreach, the collection of more information than is necessary and, therefore, raise privacy concerns. New Zealand health officials deemed it “a legitimate policy decision” to conclude that automated contact tracing was unnecessary when the manual process would suffice. Norway reached a similar
decision and discontinued its app, concluding that risks to privacy and liberties outweighed the benefit of this form of automated tracing.

Contact tracing also requires the ability to intervene at scale. Although automated contact tracing can notify those exposed at scale, in some countries the next steps after individual notification remain underspecified or undelivered upon. To whom or what institution should the exposed individual turn? How do you test and isolate all who might have been exposed if the data is not held centrally by the state? And what of those who cannot or will not use the app? Automated contact tracing raises privacy issues in its outcomes (e.g., further use of data, inequality in the community of users) even if privacy-preserving measures are in place.

Controlling the spread of Covid-19 requires interventions that break the chain of human-to-human transmission. Identifying and assessing people who have been exposed to the novel coronavirus through contact tracing is one critical public health intervention, in addition to others like testing, isolation, and care. Yet collecting the information necessary for various forms of contact tracing poses lingering threats to privacy, including access to information about one’s health status, personal connections, whereabouts, and behavior that may make a person vulnerable to stigma, punitive measures, and other social harm.

Clara Hanson
Alondra Nelson

June 2020
ENDNOTES


17 Redmiles, “User Concerns.”


28 Ibid.


33 Greenwood et al., “Privacy Principles.”

34 Ada Lovelace Institute, Rapid, Online Deliberation on COVID-19 Technologies (Ada Lovelace Institute, 2020).


36 Mark Tunick, “Privacy and Punishment,” Social Theory and Practice 39, no. 4: 643–668, JSTOR.
37 Solove, *Nothing to Hide*.


39 WHO Team, “Context Tracing.”

40 Our thanks to Dr. Andrew Chen for this helpful typology.


42 Ibid.


44 Doughty, “The Confidentiality.”

45 Ibid.

46 Shanahan, “Party Guests.”


49 Tokmetzis and Meaker, “We Were Told.”

50 Tokmetzis and Meaker, “We Were Told.”

51 Hart et al., “Outpacing the Virus.”

52 Hart et al., “Outpacing the Virus.”


55 Tokmetzis and Meaker, “We Were Told.”


58 Daalder, “Doubts Rise.”


60 WHO Team, “Context Tracing.”