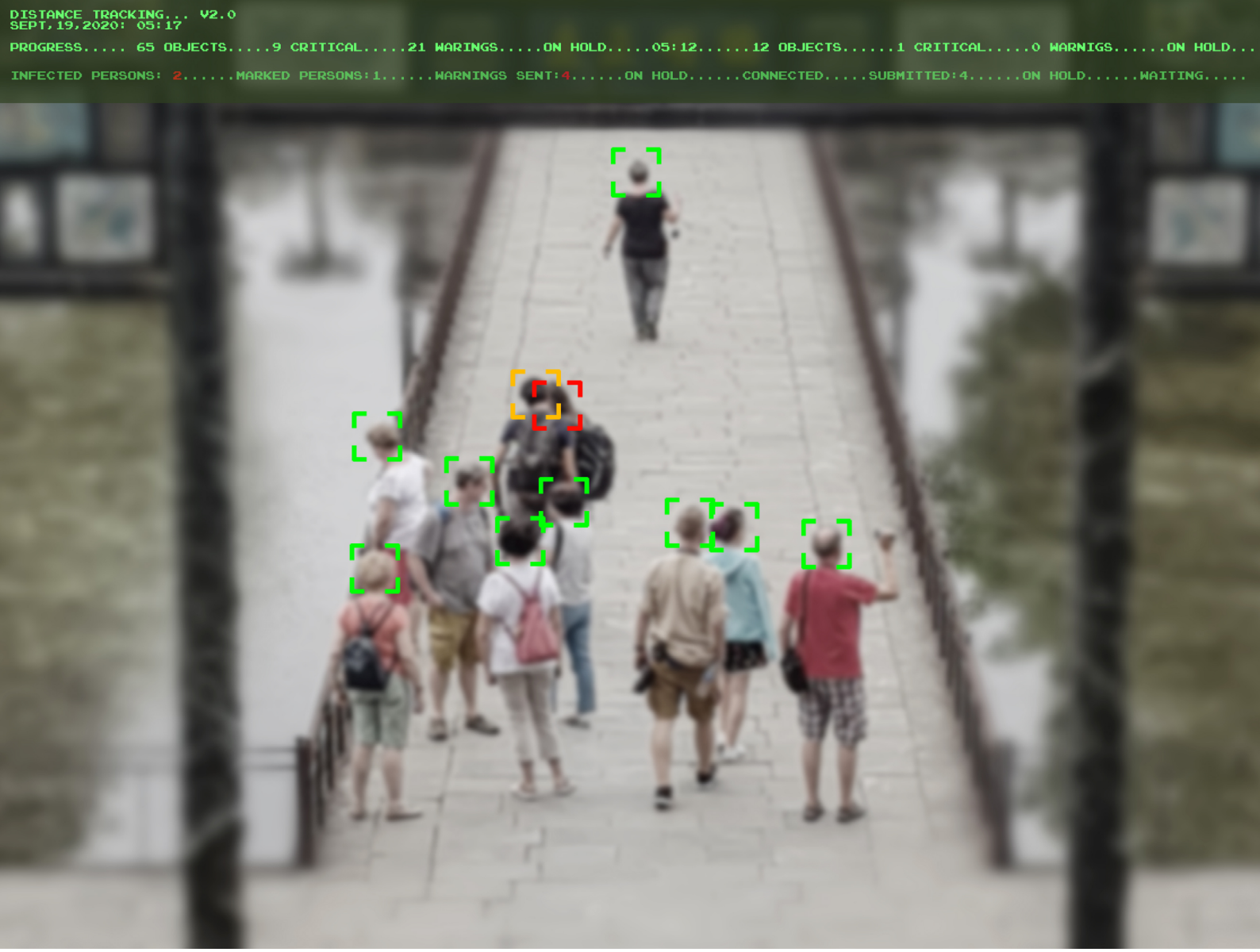


Balancing Privacy with Technology Use: Lessons from Covid-19

Digital Pathways Paper Series

Harshita Agrawal, Prakhar Misra and Ananth Padmanabhan

BLAVATNIK SCHOOL OF GOVERNMENT
UNIVERSITY OF OXFORD
Data Governance Network
Anchored by IDFC Institute
IDFC INSTITUTE
Digital Pathways at Oxford

Harshita Agrawal is Associate at IDFC Institute.
Prakhar Misra is Senior Associate at IDFC Institute, and leads the Data Governance Network.
Ananth Padmanabhan is Dean of Academic Affairs and Associate Professor at Sai University, Chennai.
The Data Governance Network, anchored by IDFC Institute, is a multi-disciplinary community of researchers tackling issues of data-enabled governance and the digital economy in India.
IDFC Institute has been set up as a think/do tank to investigate India's ongoing transition from a low income, state-led country to a prosperous market-based economy.

Paper 8
December 2020

Digital Pathways at Oxford is a research programme based at the Blavatnik School of Government, University of Oxford. It produces cutting-edge research across the fields of public policy, law, economics, computer science, and political science to support informed decision-making on the governance of digital technologies, with a focus on low- and middle-income countries.

This paper is part of a series of papers on technology policy and regulation, bringing together evidence, ideas and novel research on the strengths and weaknesses of emerging practice in developing nations. The views and positions expressed in this paper are those of the author and do not represent the University of Oxford.

Citation:
Agrawal, H., Misra, P. and Padmanabhan, A. (2020). *Balancing Privacy with Technology Use: Lessons from Covid-19.* Digital Pathways at Oxford Paper Series; no. 8. Oxford, United Kingdom

https://www.bsg.ox.ac.uk/research/research-programmes/digital-pathways

@DigiPathOxf
Cover image: Shutterstock

Abstract

There has been an uptick in the use of technology in fighting the COVID-19 pandemic. Governments have had to be swift and agile with response strategies. A variety of technology applications have been helpful in this regard. Yet, data governance and privacy principles have taken a backseat. In this paper, we derive a set of eight principles that can be followed by governments to balance individual privacy while deploying technology for governance purposes in emergency situations. We look at three technologies used during COVID-19: contact tracing, AI use for drug development and disease prevention, and syndromic surveillance for our analysis. We highlight informational privacy harms that these technologies perpetrate through a framework of privacy proposed by Koops et al (2017), in order to give a set of eight principles that governments should enforce.

Acknowledgements

The authors would like to thank Arindrajit Basu, Rahul Matthan, Malavika Raghavan and Shweta Reddy for inputs on the idea and the framework, and an anonymous reviewer and Jonathan Klaaren for comments on an early draft of the paper. We would also like to thank Vikram Sinha for editorial inputs. A version of this paper was presented at the Quarterly Roundtable of the Data Governance Network held in August 2020.

We would also like to thank Digital Pathways at the Blavatnik School of Government, University of Oxford for funding this work, with special thanks to Toby Philips and Beatriz Kira. Any errors are the authors' own.

Table of Contents

1. Introduction

The Covid-19 pandemic has highlighted the importance of technology in governance more than any other public policy crisis of the 21st century. The speed and agility required from policymakers has pushed technological application in nearly every conceivable sphere of action. Some of this technology use is non-controversial, like Artificial Intelligence (AI)-powered chatbots to provide information on the symptoms of the infection. However, other uses of technology — such as cell phone tracking, drone surveillance and facial recognition systems — are more problematic due to potential privacy violations and security breaches.

Due to the urgency to save lives, data governance principles have taken a backseat in jurisdictions across the world. Governments have relaxed data localisation norms or flouted them[1] in order to use telemedicine applications[2] for augmenting healthcare capacity during this crisis. Technology players like Qure.Ai[3] are changing accountability and transparency paradigms in decision making while genomics companies like 23andMe[4] are using sensitive personal data with few purpose and storage limitation safeguards (notwithstanding Health Insurance Portability and Accountability Act guidelines and Genetic Information Non-discrimination Act). Ownership rights issues may arise as data collected during antibody testing might be considered proprietary information of the individuals being tested.[5] There have been several instances of authorities sharing individuals' contact and residential details on WhatsApp and Twitter. These issues, among others, show that data governance principles need to be considered before deploying technology.

Saving lives is of paramount importance, yet there should be a way to balance data governance principles with such imperatives. The right to privacy is a fundamental right in many countries and unless constitutionally suspended in a state of emergency,[6] a middle path needs to be sought. We draw the analogy, only in this respect, to war where conventions regulate behaviour of individuals (like prisoners of war). We realise that in this crisis, the trade-off can be costly on either side: over-regulation of technologies may protect individual privacy more effectively but can come in the way of saving lives. Meanwhile, a *laissez faire* approach can lead to deeper entrenchment of governments and corporations in individuals' lives. Such overreach is difficult to roll back.

[1] Since the virus is ongoing, data transfers for medical research may not be occasional and non-repetitive, the only two conditions for cross-border sharing of data for this purpose under Europe's General Data Protection Regulations (GDPR). Such examples complicate technology use for governance.

[2] Consider the example of a US medical centre that would collect the body temperature, heart rate, blood pressure and respiratory rate of the research subjects residing in Africa transmitting through a wearable device, or that of a hospital in Paris getting direct treatment from one in Wuhan, China (Ferreira, W., & Rosales, A.).

[3] The technology scans x-rays and uses algorithms and machine learning to monitor the rate of viral infections.

[4] It collects genetic samples from individuals all over the US to study genetic variants associated with the disease.

[5] In some jurisdictions like the US, the health care provider gets property rights over the data and becomes its legal custodian. But if the health care provider de-identifies the data, it is not protected by HIPAA anymore (Sharma, 2018). The debate around ownership is complicated in such regimes and in some- it is non-existent.

[6] For instance, the emergency provisions of the Indian constitution explicitly mention the instances when Fundamental Rights can be suspended. Similarly, the Inter-American Court of Human Right and the European Court of Human Rights state conditions such as times of war, emergency that threatens the independence of security of a State party, situations affecting the entire population and so on for suspension of rights and freedoms.

To find the right balance, we attempt to establish eight principles regulating use of technology by looking at technology deployment in three areas during COVID-19: contact tracing applications, syndromic surveillance, and AI use for drug discovery and disease prevention. This paper follows a methodology where we look at aspects of technology that amplify privacy harms and give solutions that redress those.

We focus on privacy-related harms and keep other issues outside the ambit of this paper. Differences in digital infrastructures and readiness of developed vs. developing countries, the evolution of data protection frameworks and their compatibility with said laws and other data governance violations pertaining to data localisation norms, data sharing norms, and content moderation are all important issues, but they fall outside the scope of our paper. We also want to qualify that this is an attempt to look for guiding frameworks that promise the optimal balance between privacy protection and informed policy decisions in the current context of the pandemic.

This paper is divided into five sections. Section 2 details our conceptual framework through which we derive these principles. Section 3 analyses three specific technologies against our proposed framework. Section 4 prescribes principles that governments should follow to prevent those privacy violations while still allowing technology deployment to take place. The last section concludes the paper.

2. Conceptual Framework for Balancing Privacy with Technology Use

Technology is a horizontal that cuts across sectors in offering solutions and posing regulatory challenges. For this paper, we focus on privacy harms which are immediate and direct.[7] We focus on privacy for three key reasons as discussed below.

First, privacy is intrinsic to an individual's freedom, dignity and social standing. It is a vital human right, necessitating a first-order ethical duty on individuals to respectfully value their own privacy and second-order duty to protect one's own privacy in order to protect others. For instance, while it is an individual's duty to protect their own genome data, it is also incumbent on the individual to protect their genome data since it belongs to their family as well. Thus, in modern day societies, especially in a democracy, there is intrinsic value in upholding the idea of privacy where people must be able to think and act independently (Allen, 2013). Benn (1971) uses Kant's principle of respect for another person to state the importance of privacy. He suggests that recognising an individual's agency is recognising that the person has the liberty to act in his own accordance. Similarly, Post (2000) also suggests that privacy is intrinsic to an individual's dignity and which is based on the underlying norm of mutual respect between individuals. While there are other conceptions of privacy,[8] we think the intrinsic value of privacy to an individual's being makes it necessary to protect it.

Second, privacy has instrumental value and that value can be defined using the concepts of negative freedoms and autonomy. Specifically, in the realm of data protection, individual autonomy can be exercised through the option of informed consent. It is in place to protect individuals from abuse of power and infringement by the government. As we have seen during this crisis, privacy violations can have second and third order effects in society. Access to personal data like location data to monitor and track individuals gives the state extraordinary power to monitor its citizens, which, along with monopoly over coercive machinery, can be dangerous. Data aggregation of geocoded health records can be used to ostracise entire communities or geographical zones to further illegitimate political ends. Given that extraordinary powers extended to the government are rarely rescinded, such risks are intensified in the long run. For instance, the Patriot Act, which bolstered the federal government's surveillance powers in the United States in the aftermath of the 9/11 terrorist attacks, continues to be in effect after 19 years. Thus, checking the actions of the government is important to maintain negative freedom and autonomy.

[7] For instance, monitoring technologies like China's health code application — which combines data on travel history and health records and assigns a health code which determines whether an individual is fit to travel or not — or the Government of Karnataka's instructions — to upload selfies of its citizens on an app for purposes of detecting their location and confirming observance with quarantine restrictions — regulating behaviour of its citizens do not follow basic protocols of storage limitation, purpose limitation and data minimisation.

[8] Post (2000) also states Jeffrey Rosen's view that privacy connects to knowledge in the way that privacy is important to avoid misjudgments about an individual based on out-of-context information rather than on genuine knowledge. According to him, a third concept of privacy is that of freedom. Privacy as freedom is the liberty to act and speak as one pleases without any explanations. A private space thus becomes where social norms are suspended, people are autonomous and are free to discuss diverse opinions. Post (1991) mentions Warren and Bradeis's one such conception that it is a 'right to one's personality' that is revealed during private correspondence.

Third, as a legal and constitutional matter, privacy is a fundamental right with over 130 countries providing for constitutional edicts concerning its protection.[9,10] Such constitutional protection manifests in domestic legal frameworks as data protection laws. For example, the European Union's General Data Protection Regulation (GDPR), strengthens the right to privacy to an extent where rights such as the right to be forgotten, where an individual can request her data to be deleted and/or made unavailable online, are explicitly included (Sloot, 2014). Greenleaf (2019) states that, as of 2018, 132 countries had data privacy laws encompassing both the public and private sectors. Many others are at various stages of introducing Bills for these laws or updating and replacing existing laws.

All of this points to the importance of preserving privacy as a constitutional guarantee even during trying times such as this pandemic. We now turn towards the main framework to analyse privacy violations.

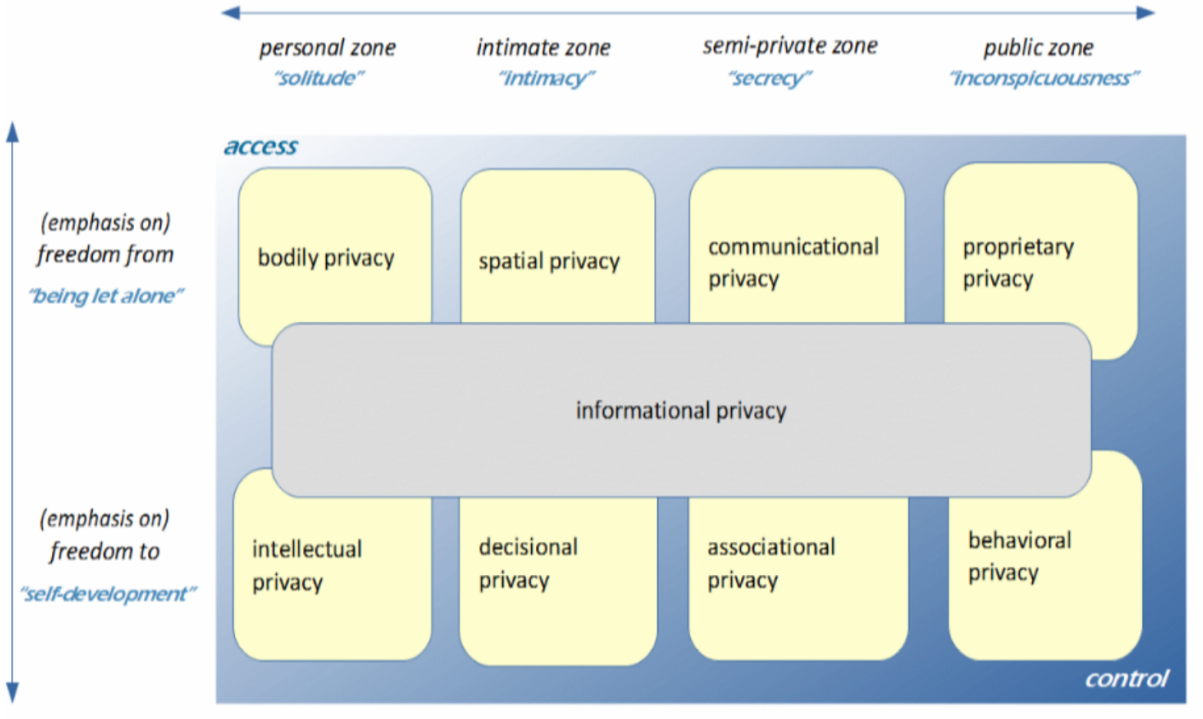
2.1 Main framework

We rely on the framework given by Koops et al (2017) to outline a typology of privacy and use that to analyse implications of technology deployment during COVID-19. The typology itself is built on the works of Nissenbaum (2009), Moore (2010), Cohen (2012), Solove (2008) and Westin (1967) among others. A snapshot of the framework is shown in Figure 1.

[9] While there are varied versions of data protection and privacy laws, the GDPR is considered a global norm-setter. It outlines specific grounds on which privacy can take a back-seat. The specified exemptions — an entity does not operate within or monitor the personal data of people within the European Union; it is not processing personal data but only making use of anonymised data; it is processing unstructured paper records manually; personal data is being processed for domestic use like personal correspondence, keeping an address book or for social networking; law enforcement by police and secret services; journalistic purposes and educational purposes — do not accommodate within them, at least in a neat manner, the range of privacy intrusions that have taken place in the name of deploying technology to combat the raging pandemic.

[10] In India too, the Supreme Court has clearly listed a few exceptions to the right to privacy and have a bearing on interventions during Covid-19 as it remains to be seen whether State action during these times have been confined within the realm of such permissible exceptions. Further, information processing that leads to identification, aggregation and secondary use of data have the potential for intrusion, violating individual dignity and putting lives at risk.

Figure 1: A Typology of Privacy



Source: Koops, B. J., Newell, B. C., Timan, T., Skorvanek, I., Chokrevski, T., & Galic, M. (2016). A typology of privacy. U. Pa. J. Int'l L., 38, 483

The framework works on two axes. The horizontal axis looks at various zones of a person's life and the types of privacies that exist in each zone ranging from personal to public.

• Personal zone is the deepest space for privacy expectations, where there are no external stimuli interfering with the person's thoughts and actions. Koops et al. quote Mantovani (2013) describing this as a state where "no one knows" and individuals can be most autonomous.

• Intimate zone refers to areas where the person is in close proximity to other individuals but is also allowed seclusion to achieve a more 'frank' relationship with them. It is a state of 'limited information flow within trusted relationships.'

• In the semi-private zone, some degree of anonymity is still sought after. 'Reputation' and 'identity building' start affecting actions of individuals like when an individual is interacting in a public place but not expecting to be monitored.

• The public zone is the place where privacy expectations exist even when private actors are conducting themselves in full public view. This zone concerns restricting access to property and personal data in public environments. Individuals should be inconspicuous - allowing themselves to be themselves even when in public.[11]

[11] For a more detailed understanding of these zones, refer to Koops, B. J., Newell, B. C., Timan, T., Skorvanek, I., Chokrevski, T., & Galic, M. (2016). A typology of privacy. U. Pa. J. Int'l L., 38, 483.

The vertical axis looks at the negative and positive aspects of freedom — freedom from and freedom to exercise something. The authors do recognize that a neat distinction between the two often doesn't exist and it is possible that both conceptions of freedom may be invoked at the same time. They also echo MacCallum's (1967) triadic relationship between conceptions of freedom: "freedom is thus always of something (an agent or agents), from something, to do, not do, become, or not become something."

Embedded within the framework is a spectrum ranging from "access" to "control" of the various types of privacies elaborated. Access is closely linked with the idea of exclusion – can an individual expect to restrain others from intruding into deeply private and intimate zones of her life. Naturally, as we move to semi-private and public zones, this expectation weakens. Yet, not all is lost as the individual can still expect to maintain some control. In other words, privacy expectations do not end merely because information is in the public domain. Rather they manifest through elements of control that individuals can exercise over unchecked disclosure of such data.

The framework outlined above allows us to capture nine different types of privacies that exist and how technologies used during the pandemic may invade those. We briefly go over each of these types below and expand upon their interaction with Covid-19 technologies in the next section.

• Bodily privacy - This type generally refers to privacy of the physical space. It is the freedom to exclude people from touching one's body without one's consent. Compulsory immunisation or compulsory sterilisation or blood transfusion without consent could amount to privacy violations.

• Spatial privacy - This kind of privacy generally refers to the privacy of personal space which could be a home, office, car etc. Interventions like smart CCTVs, wi-fi tracking and facial recognition software would be liable to violate this privacy.

• Communicational privacy - This type of privacy is damaged by interception of personal communications. Eavesdropping, phone tapping and access to emails or stored communication without consent would abrogate it.

• Proprietary privacy - This type refers to reputation and includes a kind of privacy which will be violated if property is used without consent.

• Intellectual privacy - Privacy of a person's own thoughts and beliefs, with freedom to generate opinions of their own.

• Decisional privacy - This type includes privacy of making decisions related to most intimate aspects of life, primarily sexual or procreative in nature.

• Associational privacy - This refers to individuals having freedom to associate with any groups of people they'd choose. Recording of public meetings could alter such aspects of privacy and take away agency of individuals.

• Behavioural privacy - This relates to activities that happen in public spaces like religious activities or political views being aired.

• Informational privacy - Every type of the privacy outlined above reflect some of this type of privacy which is why it is not a separate category but looked at as "the other side of the coin." It is defined as "the interest in preventing information about one-self to be collected and in controlling information about one-self that others have legitimate access to."[12]

Within each of these privacy harms, we specifically look at technologies that amplify some of these risks and further suggest ways in which that could be mitigated. Most of the risks we discuss are ones that directly impact privacy. Specifically, we focus on the privacy harms caused due to collection of data and information. There are possibilities of harms to privacy even without the collection of data, such as a spatial privacy being violated when a drone is following an individual, irrespective of collecting data, or violence inflicted on individuals due to their identity being revealed thereby harming bodily privacy. However, that is outside the purview of this study. Apart from this, there are a few other considerations as discussed next.

2.2 Privacy risks, amplified and mitigated

In our analysis, we keep in mind three additional variables which do not get directly covered in the above framework but could have an impact on regulation. While risk-based and outcome-based approaches to regulation are important, these additional concerns play a role in designing regulatory responses as well. These may not be explicitly called out in all aspects of analysis, but it is important to keep these in mind to have effective principles of regulation.

2.2.1 Entity processing the data

The nature and kind of entity processing the data is an important factor. We specifically distinguish between a public and a private entity.[13] There are five reasons why we think this distinction is important.

First, right to privacy is a fundamental right in many countries. For this very reason, the vertical application of this right within the context of State action over private individuals must be treated on a separate footing from its horizontal application, i.e. actions between private individuals (Corrin, 2009).[14]

[12] Ibid.

[13] We look at the ownerships of servers on which the data is stored as the barometer to classify entities. So, even if a private company is collecting data for a government entity but the data is being stored on government servers, we will consider that as a public/government entity.

[14] While some recent scholarship has contested this distinction, particularly because big technology companies are as influential and dominant in our daily lives as State actors, we believe that the distinction is at least important enough to warrant considering the possibility of differentially regulating public and private entities.

Second, private and public sectors differ in their mission and in their approach. Public sector companies, most times, are not seeking to maximise profit but to improve service delivery and governance outcomes whereas the private sector uses personal data for profit-making purposes, and societal benefit is secondary to their mission. Similarly, the dangers would be different too. For instance, public sector surveillance could be legitimate to protect lives but illegitimate to persecute sections of society. However, the concern with the private sector in this case would be information collection and flow and not direct concerns with surveillance.

Third, given that the possibility of governmental and executive decision-making is greater with a public entity than with private, the meaningfulness of stipulated checks and balances would be heavily affected in cases where the government is both the umpire and an actor in the data processing domain. A second and more ominous concern in this regard would be the possibility of political interference — again, higher in the case of a public sector entity — thus necessarily demanding different safeguards for data use by the two agencies.

Fourth, the architecture and accountability expectations are different. In the private sector, there would be a limited number of shareholders and stakeholders that the firm is accountable to. But in the case of public sector entities, the government, in democratic societies, would be accountable directly to all people within its governance umbrella. Further, the operational architectures would be different with limited central and local government bodies for the public sector but a more diversified and complex set for private entities. For instance, Facebook and Google may even own a few small companies further complicating their architectures and data flows. This directly affects the tracing of data and identification of data controllers and by implication, obligations under data protection laws.

Fifth, the state is a special entity with specific coercive powers outlined in the constitution and thus its regulation differs from that of private players. A widely articulated scholarly framing of the state is that it enjoys monopoly over violence. Therefore, controls over such an entity should be thought through differently than other entities using such technology and the data that is generated.

2.2.2 Kind of data being processed

Most data protection laws categorise data into two categories of personal and sensitive personal data. Sensitive personal data is a subset of personal data.[15] This distinction is made keeping in mind the higher risk of harm that may be caused to an individual dependent on the kind of data being processed. The distinction is also made to prevent discrimination, ill-treatment, humiliation and embarrassment to individuals. Personal data like name, address, real-time location, identification number and so on can be attributed to an individual and can be used for daily tracking and

[15] Some laws treat this distinction more stringently than others. Like, the Data Privacy Act of Philippines prohibits any processing of sensitive personal data except in a few cases. The Australian Privacy Act places stringent obligations on data controllers by including strict consent requirements using an explicit opt-in approach when dealing with sensitive personal data both for collection and disclosure.

monitoring purposes. Sensitive personal data like health records, sexual orientation, religious beliefs, ethnicity, etc. can be used in the longer term by political parties to discriminate against certain groups or by private corporations to capitalise on needs based on association with a certain group.[16]

This distinction further becomes important in times of a pandemic. Personally identifiable information like real-time location may be harmful to an individual in the immediate time frame whereas blood samples, plasma samples and other sensitive personal data may be harmful in the long run. Thus, it is important to consider the purposes for which such data is collected and processed before it is allowed.

2.2.3 Time for which the data is stored

It is important to keep in mind the time duration permissible for storage of certain kinds of data. This becomes crucial when collecting and processing data in a pandemic. The need for innovation must be balanced against potential harms due to breach of privacy, making the data retention period crucial. This aspect of the framework refers to the principle of storage limitation.[17]

During this crisis, we see many applications using protein data, analysing samples of genes of individuals, and processing blood and plasma samples without an end result in sight. On the other hand, even geo-locations and recording of daily activity of individuals, though collected for immediate purposes, can be potentially retained for a longer period keeping in mind broader public interest. Thus, the legitimacy of the duration will have to be decided on a case by case basis. Drawing out principles informing retention policies of various technologies explicitly will be useful in this regard.

[16] For a detailed discussion on spectrums of data sensitivity, refer to Rumbold, J. M., & Pierscionek, B. K. (2018). What are data? A categorization of the data sensitivity spectrum. *Big data research*, 12, 49-59.

[17] The International Association of Privacy Professionals (IAPP) describes storage limitation as: "the principle that personal data must be kept in a form that permits identification of data subjects for no longer than is necessary for the purposes for which the personal data is processed."

3. Technology use during Covid-19

In this section, we look at three technologies that are used in the fight against Covid-19. Each of these are selected keeping in mind the variation in the three tertiary variables identified above: a) entity processing the data (public/private); b) kind of data being used (personal/sensitive personal); c) time for which the data is stored.

Using these we identify contact tracing applications, AI use for drug discovery and disease prevention, and syndromic surveillance in this section.[18] At the time of writing this paper, these three technologies were the most widely deployed by governments and private entities in an attempt to contain the spread of the virus. Out of them, the use of contact tracing and AI use has been more extensive than syndromic surveillance. Thus, the strength of our analysis is directly informed by the evidence available of the use of these technologies. Further, within them we look at a two-pronged approach of risk amplification and risk mitigation that can be enabled to find this balance.

In our methodology, we have in particular focused on the direct privacy types that are implicated by the technology under examination. This is for the following reason: these privacy types are not watertight compartments. Therefore, one kind of privacy harm could easily metamorphose or escalate into another kind of harm over time, or because the technology under examination is used alongside some new technology or modified using some incremental capability. To give an example, a contact tracing application could be designed to gain camera access as well, thereby violating bodily privacy. Or the communication that it reads could be an intimate one, thereby allowing the State to take further steps to harm an individual's decisional privacy. But if the analysis were to be conducted by factoring in all such possibilities, then naturally every technology would possibly implicate every kind of privacy interest. Instead, we have kept our analysis tighter and more rigorous by confining the identification of privacy interests to those directly implicated by the technology as it stands today. Hypothetical possibilities, while worrisome, cannot guide any constructive attempt at reforming the privacy and data governance structure in respect of any technology, hence this specific approach in our analysis.

[18] While we chose these three technologies to analyse the privacy violations, the principles for regulation that will emanate from these can be applied to any other technology that has the same outcome. So, while we analyse six contact tracing applications, the purpose of the technology that needs to be regulated is contact tracing. Therefore, any technology that fulfils that outcome is an indication that privacy protection is necessary as it is likely that its features will have similar privacy implications as laid out in the paper and will have to be similarly regulated.

Table 1: Technology Use During Covid-19 and Privacy Violations

Zone:	Personal	Intimate	Semi Private	Public	Personal	Intimate	Semi Private	Public	Informational Privacy
Harms:	Bodily Privacy	Spatial Privacy	Communicational Privacy	Proprietary Privacy	Intellectual Privacy	Decisional Privacy	Associational Privacy	Behavioural Privacy	
Contact Tracing									
AI for disease prevention and drug discovery									
Syndromic Surveillance									

3.1 Contact tracing

Contact tracing is one of the most widely used methods for monitoring the spread of the disease. Using this technology, one can track the individuals that may be potentially exposed to another infected person and prevent them from further spreading the virus. Depending on the uptake of the technology, it has the potential to reduce the R0, which is the rate of transmission that determines how many individuals one person is capable of infecting. Since R0 depends on population size and density, isolating infected individuals and potential cases reduces their exposure to non-infected individuals, thereby reducing the rate of transmission.

We look at six contact tracing applications — three from public agencies and three private. Among public applications, we look at Aarogya Setu in India, HaMagen in Israel and Singapore's TraceTogether. Among private applications, we consider the Private Kit: Safe Paths developed by MIT, Pan European Privacy Preserving Proximity Tracing (PEPP-PT) and Decentralised Privacy Preserving Proximity Tracing (DP3T).

3.1.1 Spatial privacy

At the outset, four of these contact tracing applications by default infringe upon spatial privacy since they collect location data to track an individual's movements. TraceTogether and DP3T are the exceptions.

Aargoya Setu expects users to enable Bluetooth and location on the device, violating spatial privacy. Signing up with the application was voluntary when it was launched,[19,20] and so once the application is installed and information is entered, it is assumed to be consensual. While proximity data may be seen as essential to tracking the individual to know if he/she has been in contact with an infected user, it is not necessary to know the real time location. Personally identifiable information can amplify the risk of surveillance and spatial harm. It seems similar to Jeremy Bentham's model prison, the Panopticon, where the cells are organised in a manner that a central authority can view, in this case surveil citizens, but the prisoner (citizen in this case) can't see the central authority.

It is a similar case for Singapore's TraceTogether which collects Bluetooth data to approximate distance to other phones which have the same application. TraceTogether requires user consent at the time of registration.[21] However, once tested positive, users have a legal obligation to disclose that information for containment measures and the data will be uploaded to the servers. While it only collects proximity data through Bluetooth to track a user's data which is then encrypted, the data can eventually be identified by the Ministry of Health which holds the decryption key. While consent is taken for some kinds of data, disclosure of information on test results, location, logs and data stored on the applications is legally mandated. HaMagen is quite similar to these as well. Upon installation, it requires authorisation to access location and the internet, but collects all data stored on the phone upon activation. There is no explicit consent taken from users.

The private applications, Safe Paths, PEPP-PT and DP3T, all use the technology for proximity tracing and informing other users of their proximity to an infected user. Safe Paths collects location data via GPS, PEPP-PT collects it using radio signals and DP3T collects Bluetooth signal data. Again, there is violation of spatial privacy. They, however, do a better consent management job than the government applications. Safe Paths uses a Quick Response (QR) Code for persons to voluntarily report their results. For PEPP-PT too, if a user tests positive, they are contacted by the health authorities that provide a code that prevents incorrect information from being injected into the PEPP-PT system. They can then voluntarily provide information to the national trust service. Since the application does not ask for personally identifiable information and giving information about tests to authorities is voluntary, it improves consent collection mechanisms. However, in PEPP-PT, there is no mention of consent for radio signal usage.

There is a lot of variance on how the applications handle the duration of data storage. TraceTogether is clear with anonymised data being stored on the user's phone only for 25 days. HaMagen does not specify any details on this front. Aarogya Setu has a complicated process in this regard. Data collected upon registration is stored as long as the application is in use or as required under law. Other data is stored on the phone for 45 days. If the user tests positive, the DiD is linked

[19] It is voluntary in the sense that it is not being pushed directly to phones by the government, but mandatory in that its use was made compulsory in certain contexts.

[20] On 1 May, 2020 the Union Ministry issued guidelines mandating the application to be downloaded by all individuals residing in a containment zone. However, this mandate was diluted with new guidelines released on 17 May, 2020 stating that to the best possible efforts employers must ensure installation of the application by employees to maintain safety.

[21] Consent is taken for (a) storing mobile numbers in a secured TraceTogether registry and (b) receiving information on contacts and risk. The user can request for identification data to be deleted from the servers.

with personal information, which is uploaded and will be retained for 60 days on a centralised government server of the National Informatics Centre (NIC). The duration of storage does not seem justified in this case, since an infected individual recovers much sooner.

There is variance among the private sector applications as well. On Safe Paths, the location data is stored on the user's device and is deleted after 28 days. PEPP-PT takes an epidemiological path. It mentions that when two users are sufficiently epidemiologically proximate, the encrypted proximity history is stored locally on the device. It is deleted when it becomes epidemiologically unimportant.

To mitigate the risks, these applications must be used strictly for contract tracing purposes. However, it still causes spatial harm, which may be necessary to a certain degree given the nature of the crisis. In that case, taking the individual's consent for data collection is important.[22] While taking consent does not resolve the privacy harm caused due to use of this technology, it reduces information asymmetry for the user regarding the widespread use of varied kinds of technology and possibly, their consequences.

To reduce this information asymmetry, consent in this instance is necessary. It is widely accepted in literature and from users that granular consent is a more optimal solution for maintaining individual privacy, especially with regard to health data (Caine and Hanania, 2012; Wachter, 2018).[23] While bundled consent is an acceptable form of collecting consent, it does not further privacy interests in a meaningful manner. Bundled consent, as a condition of a service being provided, may be acceptable in cases when the purpose of the data use is explicitly stated and there is no possibility of other uses. However, as the GDPR states, when consent is bundled with other terms and conditions, it is not free and more or less useless, if not invalid.[24] Thus, consent for each type of data or of categories of data with complete information on location of storage and duration of storage should be taken from the individual.

Two additional protections need to be embedded given the above scenarios. First, if a protocol is available with better consent collection mechanisms and is not being adopted, governments must be obligated to issue explanatory reasons for the same. Second, even when all the information on purpose of use and data processing is revealed, consent should also be dynamic as per the user's request as it allows for granular control over data. The user should be able to withdraw consent at any point in time. While dynamic consent brings up issues like increased user responsibility or exclusion of some individuals or groups who are unable to fully comprehend the implications of opting in or out, it does provide increased transparency and flexibility to the user (Williams et al, 2015).

---

[22] The goal of a consent mechanism is to make an individual aware of the harm being caused and having an option to opt out.

[23] Granular consent allows applications to offer separate options to consent for different purposes of technology use and different types of data processing by the technology.

[24] An instance of bundled consent is disallowing purchase of a product or service without providing a phone number or restricting use of an application without approval on collection of various kinds of data makes consent void. While in the first case true consent would mean allowing for purchase of the product or service even without the provision of a phone number by the customer and this would further privacy, the second scenario would call for granular consent.

The duration of location data storage must be based on epidemiological studies verified by health experts, keeping contact tracing in mind. Else the risk of spatial harm may be very high, as in the case of Chinese use of contact tracing applications. The Chinese government, with help from a subsidiary of Alibaba, launched the Alipay Health Code system. The system gives users a colour code based on travel history and health status by relying on information from the government, personal medical records and self-reporting. The code green allows a person to travel freely, yellow advises home isolation and red indicates a confirmed infection of Covid-19 and mandates quarantine. There is no way to know how the code changes. Residents of Hubei, even after being quarantined for weeks, were stuck and unable to travel because their health code remained red. Unchecked deployment can allow it to become a norm to show the results on these applications before entering grocery stores, using public services or be mandated by your employer as is gradually happening in China and Hong Kong. This leads to spatial harms as well.

3.1.2 Communicational privacy

HaMagen requires internet access for cellular surfing and collects all data stored on the phone once the application is downloaded. Since this may include information regarding correspondence via email and messaging platforms, it raises the risk of communicational harms. This information is not required for immediate contact tracing purposes and may be used for other purposes during or after the pandemic has subsided.[25]

In the case of contact tracing applications, there is an overlap of communicational privacy and spatial privacy concerns. While spatial privacy was previously mentioned in the context of physical space, an individual's online space is also very much a space that requires protection. Increasingly, we live our lives online and so the information there is subject to privacy violations as well (Cohen, 2008). Accessing all information on the mobile device and accessing the internet for cellular surfing, as in the case of HaMagen, makes users susceptible to online spatial harm as there may be information on private correspondence with other individuals, parts of a persona that the individual wants to keep private, that the government may now be able to access and use for unknown purposes.

To mitigate this risk, since the purpose of contact tracing applications is to track the movement of individuals to determine if they have come in contact with infected individuals, the bare minimum data required for that purpose must be collected. Since there are technologies like the decentralised applications created by private entities that can perform the same tasks without collecting all the information on the phone, governments must avoid collecting such data. Additionally, there need to be safeguards regarding who has access to the data.

[25] For instance, the NHSX, which is the unit responsible for setting national policy and standards of data and technology used by the National Health Service in the United Kingdom, planned to centrally store de-anonymised data of infected individuals and of those who the infected individual may have been in contact with. This opens up the possibility of further use of data for other purposes even once the pandemic is over.

3.1.3 Proprietary privacy

Proprietary privacy is also violated in some cases. In the case of Aarogya Setu, the added threat is when a user tests positive, their data is uploaded on the server. Similarly, in Singapore, once tested positive, users of TraceTogether have a legal obligation to disclose that information for containment measures and the data will be uploaded on the servers. In the case of TraceTogether there is a safeguard in place. It does not mention any secondary use and adheres to the privacy principle of purpose limitation.[26] So it is unlikely that reputational harm is caused due to databases being linked. However, in the case of Aarogya Setu, there are no specifications provided on whether the data will be linked to other databases or not. Identification of infected users can lead to discrimination and ostracisation. For instance, the South Korean government disseminated data on movement of diagnosed individuals leading to stigmatisation.

Risk mitigation would involve moving to decentralisation of data. Decentralised applications do not store data on centralised servers and do not collect personally identifiable information. Centralised applications, on the other hand, store anonymised data on a centralised database where the data is matched with other contacts in case a user develops symptoms or has been infected and the computer server sends alerts. The decentralised versions store data on the phone, the matches with potential infected persons are made on the phone itself and the phone sends alerts in case of a positive result. Given that there exists a technology that does not require personally identifiable information to be collected that could lead to reputational harm, there should be no reason to use centralised applications that carry that risk.

That said, if centralised applications must be deployed, personally identifiable data must be anonymised. In TraceTogether, keyed in mobile numbers are substituted by a random permanent ID.[27] Since the temporary ID is refreshed at regular intervals, there is no possibility of identification of a user by a third party. All user data is encrypted and stored on the user's device, and data will not be accessed unless the user has been in close contact with a confirmed Covid-19 case. However, since the Ministry of Health has the identification key, it is possible to decrypt the data, identify an infected user and use the data for other purposes. Decryption of data can cause breach of privacy. Last year, the NSO Group, an Israeli technology company, was sued by WhatsApp for using its spyware software, Pegasus, to surveil individuals in 20 countries, including India, by hacking into phones, decrypting information and targeting them.

The private applications fare better on protecting proprietary privacy of an individual. Safe Paths does not share data with any third party, even the government. Users are only notified if they are in close proximity to diagnosed carriers without revealing their identity. In the case of PEPP-PT, if a user is identified as being associated with another country than the other proximate user,

[26] The Personal Data Protection Commission of Singapore released some advisories which state that personal data that is collected by organisations for contact tracing must follow the Data Protection Provisions of the Personal Data Protection Act (PDPA), 2005, to ensure personal data is protected and it is not used for any other purpose than that mentioned in the application.

[27] When close to another TraceTogether enabled device, the Bluetooth exchanges a temporary ID generated by encrypting the permanent ID with a private key held by the health ministry. This temporary ID can only be decrypted by the health ministry but the privacy protocol states that it does not reveal anyone's identity.

information associated with the anonymous ID of one of them is transmitted to the national trust service of the other country. This transmission is fully encrypted and digitally signed. Further processing is done by the national trust service of the country that issued the app.

One way to anonymise data is the way DP3T does it. It creates a frequently-changing Ephemeral Identifiers (EphID) which is an anonymous ID with no opportunity to identify an individual. Safe Paths encrypts the location data it collects.

However, there is enough evidence that no technology to anonymise data is foolproof. A 2013 study has revealed that 95% of the people from a database containing high frequency location data of 1.5 million people, can be identified using only four data points.[28] The more the number of variables, the higher the chances are of identifying individuals and less likelihood of anonymised protection. Additionally, we think governments should implement strict rules, protocols and conditions under which decryption of the data would be allowed - violation of which should lead to commensurate punishments. A second thing that can be done here is to build an architecture which leaves audit trails whenever data is decrypted, noting identification of metadata on who decrypted the data. An audit log can contain information on each operation of the database, whether it is access or processing. The log can include account details of the user for each transaction. In case of suspicious tampering, all access and operation can be analysed and illegal or unauthorised access or operations can be detected (Elmasri & Navathe, 2004).

3.1.4 Informational and associational privacy

Aarogya Setu, at the time of registration, collects information like name, phone, age, sex, profession, countries visited in last 30 days and location data. Signing up with the application is voluntary in most cases. Therefore, once the application is installed and this information is entered, it is assumed to be consensual. HaMagen is similar with regards to the amount of data collected. Aarogya Setu does not mention the purpose of collecting sex and profession of an individual. Firstly, from the use of both applications, informational harm is immense due to the amount of information collected. It is more than is required for the purposes of curbing the spread of the disease. Association of an individual to groups, like the ones mentioned here, can lead to serious harms. Aarogya Setu does mention an exhaustive set of uses: contact tracing, quarantine, care, location sanitisation, identification of clusters, generation of heat maps and containment. However, Tiwari et al (2020) mention that uses like legal requirements stated in the terms of Aarogya Setu create ambiguity and can be misused. Identification of clusters and generation of heat maps are vague uses and data can be illegitimately misused to identify clusters other than those of just infected individuals, in order to ostracise or discriminate against certain groups. As Raskar et al (2020) note, this kind of tracking has privacy implications for various people. In this case, diagnosed users are at greatest risk of their privacy being violated due to public identification.

[28] For more details of the study, see: https://www.livemint.com/Opinion/h9A1vfAHPgSKZCMEvmqO7M/Opinion--The-utter-meaninglessness-of-anonymizing-telecom-d.html

While TraceTogether does collect personally identifiable information and health data when an individual is infected, they do not collect additional information that is not directly related to their efforts of contact tracing.

All three private applications, on the other hand, are almost completely free from information and associational privacy harm since they do not collect any personally identifiable information. The only data collected is location data which can be deemed to be necessary and proportionate to the need of the hour.

As in the case of these private applications, the data collected by public solutions should be limited to the immediate purpose of the contact tracing in order to mitigate this risk. The collection of data should be in alignment with the purpose and any additional information must not be collected. Additionally, aggregation of data can also lead to greater possibilities of surveillance and associational harm. Therefore, linking of data collected during the pandemic with other data sets for any use other than those specified should be prohibited.

3.2 Artificial intelligence for disease prevention and drug development

AI applications have been very helpful in the fight against Covid-19. However, they infringe on informational and proprietary privacy. Before detailing how these privacy violations manifest, we explore the applications and their functions. Vaishya et al. (2020) document seven applications for which AI has been used during COVID-19: detection and diagnosis of infection, monitoring treatment, contact tracing, projecting cases of mortality, development of drugs and vaccines and prevention of disease.

All AI applications are not a threat to privacy. Consider Benevolent AI that attempts to solve pharmacological puzzles by using a "knowledge graph" tool, which uses information culled from scientific literature and identifies drugs that could be moved to clinical trial phases. It uses AI to change the way medicines are discovered and developed. For Covid-19, it identified Baricitinib which had significant advantages over other drugs used for combination therapies. The drug maker was contacted and Baricitinib replicated results, as Benevolent AI had predicted. Its clinical trial was expedited, and is underway at the time of writing this paper.

We discuss three applications which do in fact infringe on individuals' privacy rights. It must be noted that the threat to privacy comes from the kind of data they use for the applications. Thus, the regulatory prescriptions are based on the data used rather than the specific AI technology. That said, we do think that the technology solutions are incentivising large scale data collection, a concern in itself.

First, consider AbCellera which focuses on antibody testing. It scans antibodies in a human which can be used for treatment of diseases, and works with biotech firms to develop those. During Covid-19, it used the proprietary technology it developed to scan blood samples from recovered

Covid-19 patients, and evaluate antibodies that may carry curative potential. AbCellera has been successful and has attracted investment from the Canadian government to further scale its work in the fight against the pandemic. The results are promising and the company is establishing itself in the AI ecosystem.

Second, Google Deepmind is building advanced AI or general artificial intelligence. It has built an AI tool called AlphaFold that focuses on protein folding of genomic data to identify structures of proteins. They uncover what the structure of the protein would look like in 3D by modelling the distance between pairs of amino acids and the angle of chemical bonds connecting the amino acids. For the Covid-19 virus, it uncovered the spikey proteins on the surface of the virus. These proteins combined with another complementary component allows the virus to enter human cells. Further, some cells — like in the lungs — are more susceptible to being infected from this kind of a virus. It looks at the data from GISAid and Protein Data Bank to make its predictions and train its AI.

Third, NextStrain works with pathogen genome sequence data and strives to make it readily available for use. Their motivations and objectives as indicated on their website are: "Current scientific publishing practices hinder the rapid dissemination of epidemiologically relevant results. We think an open, online system that implements robust bioinformatic pipelines to synthesize data from across research groups has the highest capacity to make epidemiologically actionable inferences." Thus, NextStrain is an open-source platform that supports genomic analyses which can be epidemiologically relevant. It also has a visualisation tool that can be used to get a "real-time snapshot" of evolving pathogen populations around the world. They source their data from GISAid which consists of 73,000 viral genomic sequences of Covid-19 shared by labs across the world.

3.2.1 Informational and bodily privacy

The problems with using AI applications for disease prevention and drug discovery primarily stem from the nature of data that these applications use. The possibility of informational privacy breach is very high. Most applications mentioned above use some variant of sensitive personal data for these purposes, mostly linked to genomic data obtained from open source data banks. A person's genome contains very sensitive and personal information such as proclivities towards contracting diseases of a certain type. In its digital form, the genome itself is queried against known variations to determine susceptibility for diseases like Alzheimer's and diabetes. While useful for research, this poses a high risk of breach of privacy and, potentially, of harm caused to individuals. For instance, if a person is known to carry mutation on a gene associated with cancer, they may be denied health insurance (Akgun et al., 2015). Gottleib (2001) notes that 22 states in the US had banned genomic testing for employment on the grounds that workers might face "illegal discrimination and invasion of privacy".

Broadly speaking, AI use during Covid-19 has impacted informational privacy. Google Deepmind uses data from open data banks which hold vast amounts of genomic information to process such data. NextStrain itself is an open-source platform with direct and derived information about individuals. Given the fluid nature of information flow across various entities, in a space where use

of such data is largely unregulated, control over information by the individual is nearly impossible. One could argue that bodily privacy may be infringed during the process as blood samples, throat/cheek swabs may be collected for some purposes but individuals may not be aware that the data could then be stored in centralised databanks, open-sourced and shared for research purposes. This claim, however, merits further investigation.

We give mitigation strategies to prevent this later in the section given that they have a significant overlap with strategies to mitigate other privacy harms as well.

3.2.2 Proprietary and associational privacy

Whether one's genomic information counts as one's property or not could be a matter of philosophical debate.[29] John Locke, the 17th century philosopher, argued about self-ownership with "every man has a Property in his own Person" in Two Treatises of Government (1681). Given that collection of information about one's body is personal to an individual, we think harms resulting from access to this information would qualify as infringement of proprietary privacy. As an example, consider the case of John Moore who sued after a commercial cell line was developed by University of California based on tissue used from him when he had leukaemia. Although the court did not consider the human tissue or DNA being subjected to commodification, they did acknowledge that at least some part of the harm could be mitigated by enforcing informed consent on the use of such data.

Similarly, Skloot and Turpin (2010) note that Ms. Henrietta Lacks' genome was published by researchers after her death, endangering the privacy of her family. A few of her cancerous cells were removed for research purposes, but this ended up revealing the family's genome sequence to the world. No tangible damage was recorded because of this, but given the possibilities of genomic data use in the long-term, there is potential danger. People's genomic data being linked back to their families may compromise the privacy of their family members too.

While this has not yet happened in a post Covid-19 world, we do see a possibility of this arising due to the amount of data being collected and processed. Consider the AbCellera and Eli Lilly collaboration which identified LY-CoV555, an antibody, developed by obtaining the first blood sample from a patient who recovered from Covid-19. We do think the possibility of proprietary harms arising are immense given the involvement of patients in the testing and development of drugs and other solutions to fight Covid-19. Since artificial applications use and share genomic data, there is a high possibility of amplification of harms. It could also be magnified as a result of widespread commercialisation of what is arguably property belonging to an individual. One way to mitigate this harm is to take consent from individuals. Even if this is not possible at the sample collection stage, companies should attempt to involve individuals in the process in some manner to be more transparent.

[29] This debate has many contemporary references in case the reader wants to explore in detail. Property Rights in Genetic Information by Spinello (2004), Property and Human Genetic Information by Neilsen et al. (2019); Genomic Knowledge Sharing: A Review of Ethical and Legal Issues by Francis (2014); and Who Should Hold Property Rights to the Human Genome? An Application of the Common Heritage of Humankind by Sturges (1999) are a few to name.

Among the mitigation techniques, anonymisation has not worked too well in the past. Genomic data is, by definition, a unique identifier when it comes to individuals, thus carrying with itself a higher burden to be anonymous. Sweeney et al. (2013) have demonstrated why this is problematic. They took the profiles of people who uploaded their DNA information on the website of the Personal Genome Project and linked that with publicly available profiles using voter data and public records to 84-97% of the people being reidentified. Gymrek et. al. (2013) also performed a similar exercise linking surname data to genomic data and triangulating on an individual's identity. Furthermore, not only does genetic data give information about the person's past but also about future possibilities of contracting diseases making the person vulnerable to various kinds of harms from ill-motivated individuals. Unlike other test results (blood samples/sugar levels/blood pressure etc.), genomic data doesn't change very much. For an individual, their genotype is unique and stable. Once the test results are revealed, the time period for which the inferences are known is almost infinite, which brings with itself challenges of safeguards that need to be built during collection and data sharing phases before the data is actually processed.

One way to mitigate such risks is to structure laws in a manner such that interlinking of genomic data with other datasets is performed only in extremely rare cases. One could also think of implementing strong penalties if some health data is linked to other datasets, thereby disincentivising the risk of re-identification. The state will have to take the burden of drafting this. We do envision a different set of checks and balances between the private and public entities owing to the purpose for which they would be interlinking datasets.

There is associational privacy infringement as well. Such data analysis does not merely harm an individual but also reveals information about the family members of such individuals — people who were not tested directly — thus making the question of privacy even more pertinent than in other data points. This very quality, of course, serves well for it in the fight against Covid-19 as well. Yet, the dangers to privacy are equally exacerbated. Given what we know of anonymisation techniques, it may be hard to prevent such information from being revealed and individuals being uniquely identified.

Over and above all this, it is relatively easy to make inferences based on people's genomic data. Genetic data can be derived from tissue samples, throat swabs, blood samples etc. Further, machines like Oxford Nanopore's MinIon[30] makes sequencing DNA and RNA fast and easy, making this very cost-effective and reducing barriers to processing such data.

Some other mitigation strategies too may be considered. First, allowing unbundling and revoking of consent during the time of the pandemic would allow companies and governments to be held accountable to certain obligations outlined in individual laws and regulations. Second, collection of consent to use genomic data for uses in artificial intelligence should not be bundled with consent collected for other data points. We realise withdrawing consent for research purposes may be hard to enforce, especially in a 21st century paradigm where data is widely shared internationally and with various kinds of entities. But, we do think that individuals are within their legitimate right to opt out of such research should they feel a significant amount of threat to their privacy.

[30] The MinIon is a portable device developed by Oxford Nanopore for DNA and RNA sequencing which makes it possible to stream the collected data in real-time.

3.3 Syndromic surveillance

Syndromic surveillance refers to collecting health data like symptoms and drug purchases, and mapping these to a location. This could flag a potential outbreak. The identification of a cluster of infections or diseases in a certain geographical area would then direct efforts for containment.

Detection of disease clusters could go a long way in aiding governments. For instance, the firm BlueDot deployed an AI algorithm that was able to predict the Covid-19 outbreak a week before WHO announced it. It detected an unprecedented number of pneumonia cases shooting up in a market in Wuhan and that was how it predicted the outbreak of the virus. It uses natural language processing and machine learning techniques with its data sources lying outside of healthcare. Official public health organisations' statements, global airline ticketing data, livestock health reports, population demographics, foreign language news reports, animal, insect and plant driven networks, climate data from satellites etc. constitute sub-elements of the vast data repository at its disposal. Using these data sources, it predicted the virus spread and identified a list of cities most likely to be impacted. Earlier, it had also predicted the outbreak of ZIKA in South Florida six months before it actually happened.

While BlueDot is a private entity, the Singapore and Japanese governments used such modes of surveillance too. The Singapore General Hospital, a public tertiary care hospital, used syndromic surveillance for healthcare workers. They checked temperatures upon arrival and devised a plan in case other symptoms arose. This personal data, along with data from the Human Resources (HR) department, was aggregated and analysed by the department of infection prevention and epidemiology (IPE). Heat maps were made to track the onset of symptoms arising and location of healthcare workers.

Consider Japan where the Ministry of Health, Labour and Welfare convened specialists to identify clusters of the Covid-19 outbreak. They used web searches of symptoms from smartphones and personal computers and their location data from the Yahoo! JAPAN App. The application itself is used for web searches, news and weather reports. Aggregation of this data allowed the authorities to successfully detect two clusters on the island of Hokkaido in Japan.

These applications hold the potential to infringe upon bodily, informational, spatial and associational privacy.

3.3.1 Bodily and informational privacy

In all three cases, some form of health data is collected, be it direct health reports in the case of Singapore's syndromic surveillance or through proxies like public health organisations' statements and web searches of symptoms by BlueDot and the Japanese syndromic surveillance study. As Koops et al (2017) suggest, bodily and informational privacy is compromised when one loses access to health data or genetic information. First, when one is made to provide samples of bodily fluids

or even allow for checking body temperature, one loses autonomy. Second, it gives someone else access to an aspect of one's life that can then be used for further analysis, which may not end up being favourable for the individual. Based on the analysis of someone's health records, they could be refused medical insurance due to the moral hazard problem, as they may be more likely to claim insurance.

3.3.2 Spatial and associational privacy

While syndromic surveillance can help in predicting disease outbreaks, there can be misuse as well. Considering that aggregation is a technique for anonymising data, de-anonymisation of data or re-identification is not impossible. While the aggregated data may not have been harmful, once de-anonymised, it can be used to identify individuals and target them for adverse reasons. As Luk Arbuckle states, reconstruction of aggregated data could leave only one or two people being represented.[31] For instance, if income data of individuals is being aggregated and only one individual's income represents 80% of the total distribution, then it can be used perversely. Such data processing could reveal habits or susceptibilities of a certain group from certain locations. The locations may be a place of residence for vulnerable groups or of a certain ethnic community. Healthcare infrastructure or services provision could be restricted or disproportionately provided in these places for political reasons. Such possibilities point to the spatial and associational harm that an individual or a group may be subject to.

The spatial harm may also have worse consequences. As in the case of Covid-19, once an area was deemed a cluster of disease outbreak, movement of individuals was restricted as seen in Mumbai. While it may be justified in extreme circumstances, this opens up the possibility for extreme surveillance by the State.

The ways to mitigate such risks are varied. To reduce bodily harm, explicit, granular and alterable consent mechanisms should be built with immediate effect. To some extent, technologies like account aggregators do the job of grievance redressal. Second, since the purpose of syndromic mapping or surveillance is to identify geographical disease clusters, anonymisation to the highest degree must be undertaken. Third, storage limitation mechanisms should be built into the architecture of the system. Since anonymised but aggregated data can also cause harms, a storage limitation clause is extremely important. Linking of databases must only be performed with explicit consent in hand and with authorisation from an independent authority. While the results can be used for disease surveillance which can inform policy responses, no data must be linked to sensitive information like ethnicity, religious and political beliefs and so on.

In the United States, the laws pertaining to syndromic surveillance vary from state to state. However, all healthcare institutions follow rules and procedures stipulated in the Health Insurance Portability and Accountability Act (HIPAA). They have different protocols based on the different uses of health data. We could borrow from these rules of rigorous protocols and accountability measures based

[31] For more information on issues with aggregation of data see: https://iapp.org/news/a/aggregated-data-provides-a-false-sense-of-security/

on the degree and kind of data use and processing. In the case of internal disease surveillance in a hospital to inform staffing demands, no institutional review board (IRB) approval is required. If a research project for disease surveillance is undertaken and it uses personally identifiable information, then the IRB approval from the Federal Office for Human Research Protections is required. If the data stored does not include any of the personal identifiers that are listed in HIPAA, the data can be shared with researchers under a data use agreement. Lastly, if a single hospital is to share disease surveillance data with public health authorities for public health operations, then transfer of data is permitted but with encryption and other safety measures in place.

The merits of these three types of technologies do not need to be debated. Contact tracing has worked as an immediate response technology aiming to curb the spread of the virus on-ground. Syndromic surveillance on the other hand has proven to predict the location of outbreaks and inform containment strategies. Whereas, AI has been used to identify the features of the virus to help create drugs for the virus. However, due to the widespread use of these technologies and the sheer amount of data collected, they have also been the most harmful for individual privacy. While in their current form, the applications cause significant privacy harms such as spatial, bodily, associational, proprietary, among others, we do think the dynamic nature of innovative technology will enable less privacy intrusive methods of contact tracing. In the case of AI use, although the implications of harms are not completely known, the possibility of harm due to the kind of health data collected can be disastrous. Given that the data collected during the pandemic does not pertain only to real time information, the implications are wider which need to be mitigated. Lastly, for syndromic surveillance, since the kind and amount of data collected is much less, the harms appear to be fewer than with the use of contact tracing applications.

We think the protocols set in place to mitigate risks of privacy harm caused by contact tracing and AI applications should be useful for syndromic surveillance as well. Such technology use is the first of many instances when privacy invading technologies will be used. Following protocols set for normal times will not be helpful since they operate with the view of being non-privacy invasive. However, given that lives are at stake during crises, a trade-off of saving lives over maintaining complete privacy needs to be made. Given that possibilities of technology use are endless, a different set of principles are required. We detail them out in the following section.

4. Principles for Mitigation of Privacy Risks

In the previous three sections, we have outlined the various privacy harms that can be caused due to technology and some solutions for mitigating them. We primarily rely on the framework suggested by Koops et al (2017) and embedded three additional factors: public/private, personal/sensitive personal data and time for which the data is stored. Based on this analysis, we suggested a few ways in which risks to privacy can be mitigated provided governments follow a set of protocols towards this.

We come up with eight principles to mitigate privacy harms. It must be noted that while these are principles, we do envision them as solutions that can and should be operationalised by governments in real-time. While there is little by way of affecting technology deployment during COVID-19, we do think future emergencies should use at least some variant of these principles to regulate technology deployment and protect privacy of individuals. A detailed mapping of these principles is available in Table II.

1. Purpose limitation is essential. This needs to be governed by the 'analog' components of the system. Governments should look at collection, sharing and processing of data during emergencies with an explicit adherence to collecting only required data to fight the pandemic. To do this, governments should have a line by line explanation on what data points are being collected and what purposes they would serve during emergencies such as COVID-19.

2. Consent and collection mechanisms need to be changed. The operating principle here should be granular consent over bundled consent. Even if this causes fatigue, we do think the trade-off is worth it for furthering informed consent. Further, depending on what the technology under consideration permits, alterable consent mechanisms with revoking of consent should be an in-principle guarantee. Thus, where these are not being operationalised, justifications should be given by both public and private entities.

3. Governments should hold themselves accountable for adoption of technological developments. If a new protocol is being released which preserves privacy more than other protocols, then it should be incumbent on governments to issue clarifications on why the new technology is not being adopted. This should be embedded in disaster management legislation and laws across governance regimes via amendments, ordinances or other mechanisms at disposal.

4. Duration of storage of data must be limited to the duration of the crisis and must be informed by consultations with experts who can opine on crisis mitigation strategies. For COVID-19, this process must involve epidemiological experts, who would approximate the time for which certain data points such as geo-location should be stored.

5. Internal safeguards for authorisation and access to data should be specified by collection and processing agencies, whether government or private. Encoding audit trails would help in tracking who accessed what data. Again, this should be done by default and any reason to veer away from this course should be backed by reasonable justifications.

6. Decentralised data storage should always be chosen over centralised data storage. This is mostly a feature of the technology being deployed. Yet, governance and regulatory mechanisms should find ways to incentivise this feature so as to achieve the goal of unnecessary storage of data and linking of datasets for future use.

7. Keeping in mind potentialities of surveillance and the fear of a big brother state, governments should come up with a 'negative list' of databases which will require special approval for linking with the data collected by COVID-19 applications. Protocols and exceptions to link those databases should be specified clearly. Penalties should be appropriately tailored to the severity of the infringement.

8. All data should be anonymised and encrypted by default. Strict and explicit conditions under which data would be de-anonymised and decrypted and by whom should be specified as a part of FAQs of the technology use. This should be available to the citizens of the country as well.

Table II: Mapping of Privacy Principles to Technology Use

Technology use	**Type of Privacy Harm**	**Reason for Risk Amplification**	**How to Mitigate the Risk?**
Contact Tracing	Spatial Privacy	Collection of Bluetooth and GPS data	- Limit use only for contact tracing - Where possible opt for granular consent over bundled consent (and allow withdrawal where needed) - If a better mechanism for collecting consent/ storing data is discovered, the government should issue a memo acknowledging that and specifying reasons why it is not adopting it - Duration of data storage must be specified by epidemiological experts
		Time for which the data is stored is longer than needed or unspecified	
	Communicational Privacy	Collection of data on the device	- Minimum data should be collected and any additional data collected should be followed up with the purpose for the data. - Strict safeguards on who can access what kind of data should be enforced.
	Proprietary Privacy	Uploading of data on servers	- Enforce technical safeguards like decentralisation of data storage i.e. storing data on the phone rather than centralised servers. - If, for some reason data needs to be stored centrally, it should be done in an anonymised manner like using Ephemeral Identifiers. - Design strict rules and protocols under which data would be decrypted. - Build in privacy by design by leaving audit trails whenever a type of data is accessed.
	Associational and Informational Privacy	Collection of extra information	- Prohibit linking of data with some other datasets that may increase potential of surveillance and cause harm

ctd.

AI for Disease Prevention and Drug Discovery	Proprietary and Bodily Privacy	Collects sensitive information about individuals like their genomic data samples	- Consent should be free and informed and should specify that data may be used for such purposes. - If involving people at the sample collection stage is hard, they should be involved in other parts of the process.
	Informational and Associational Privacy	Genomic data has a long shelf life	- Prohibit interlinking of genomic data with other datasets and put in place penalties if such violations occur.
		It gives out information about the families as well	- Unbundle consent with respect to genomic data from other data uses. - Allow revoking of consent for specific uses of this data.
Syndromic Surveillance	Bodily and Informational Privacy	Mandating collection of samples	- Alterable consent mechanisms through account aggregators should be enforced.
	Spatial and Associational Privacy	Aggregating of data	- Given that the purpose is to have aggregate level outcomes, all data should be anonymised and not be allowed to be de-anonymised at any stage of processing. - Linking of databases with other information of individuals should be prohibited. - Borrow some of the HIPAA guidelines on sharing of data

5. Conclusion

The COVID-19 crisis made clear what we already knew – that technology use is going to become pervasive and rapidly so. Naturally, governance will attempt to adopt technology as well, as it should. Thus, we need better frameworks to balance technology use against critical human rights such as privacy. While this is a hard task, a start needs to be made to regulate such technology – one that this paper attempts to do by examining three prominent uses of technology for fighting the pandemic that come bundled with potential privacy implications.

Based on governance preferences and intent, some technologies follow data protection principles and privacy measures, while others don't. In a jurisdiction that lacks regulation, a set of principles that instill privacy-preserving mechanisms is essential. Further, these principles need to be as much a part of both the digital and analog realms. Thus, not only do we need to talk about anonymisation, but also institutional safeguards that will prevent misuse of information by government and private actors. By looking at contact tracing, AI for drug discovery and disease prevention and syndromic surveillance applications, we reach some conclusions on this front.

On the institutional side, governments need to practise self-restraint. Limiting the use of data collected, generating notices informing the public about technological deployments, issuing protocols around decrypting data and limiting access to the information collected will go a long way towards helping balance privacy with technology use. Legal and regulatory frameworks must move in this direction.

On the technology side, consent-based collection mechanisms are key to informing individuals about how their data will be used. Unbundling consent requirements for each purpose, making consent alterable and revoking consent should all be key elements of any technology used in such a crisis. Anonymisation, no matter how flawed or difficult, will bring us a step closer to a more privacy-respectful reality. Other technical safeguards like decentralising storage of data and embedding audit trails on who has accessed what kind of data will also be helpful in achieving the right balance.

It is important to note that not all technologies violate privacy. As we showed, contact tracing technologies like PEPP-PT and DP3T have better protocols in place in some regards than others. Similarly, some uses of AI like Benevolent AI do not pose a threat to privacy while others do. Our principles have been informed by this in determining what is possible when defining the minimum viable protection.

In this paper, we focused on technologies that were directly catering to the healthcare sector and suggested principles that governments could use to regulate their use in a crisis. However, governments deployed a wider range of technologies in their Covid-19 responses. For instance, autonomous bots or vehicular systems witnessed an uptick during this pandemic. These autonomous systems were used for logistical purposes like transporting medical supplies and other items to infected areas. Autonomous wheeled robots have been used for labour intensive tasks like disinfecting hospitals, removing contaminated materials from hospitals and for

telemedicine in order to reduce risk of transmission to frontline healthcare workers by running diagnostic and treatment procedures for patients. In India, autonomous drones are also being used to warn people about violations of lockdown protocol. Draganfly, a private sector US company, has developed 'pandemic drones' which are soon to be deployed for detecting fever, sneezing, heart rate, respiratory rate and even social distancing. Due to the varied uses of autonomous bots and vehicles and the variety and volume of data being collected, regulating this technology is tougher, but crucial. We hope future research will build on the principles proposed to incorporate these technologies in a privacy-friendly manner.

We end with where we began, that extraordinary measures should not outlast extraordinary circumstances. We are confident that the principles outlined here, if complied with, will give governments the freedom to use technology to fight pandemics such as the present Covid-19 crisis while at the same time safeguarding the privacy rights of individuals.

References

Akgün, M., Bayrak, A. O., Ozer, B., & Sağiroğlu, M. Ş. (2015). Privacy preserving processing of genomic data: A survey. *Journal of Biomedical Informatics*, 56, 103–111. https://doi.org/10.1016/j.jbi.2015.05.022

Allen, Anita L., "An Ethical Duty to Protect One's Own Information Privacy?" (2013). Faculty Scholarship at Penn Law. 451.

Arbuckle, L. (2020, April 27). *Aggregated data provides a false sense of security*. IAPP. https://iapp.org/news/a/aggregated-data-provides-a-false-sense-of-security/

Ardron, M., et al. (n.d.). *Unified research on privacy-preserving contact tracing and exposure notification*. Retrieved 26 August 2020, from https://docs.google.com/document/u/1/d/16Kh4_Q_tmyRh0-v452wiul9oQAiTRj8AdZ5vcOJumgY/edit?usp=embed_facebook

Bart van der Sloot, Privacy as human flourishing: Could a shift towards virtue ethics strengthen privacy protection in the age of Big Data?, 5 (2014) JIPITEC 230 para 1.

Benn, S. I. (1971). Respect for Persons'(1971) in JR Pennock & JW Chapman, eds., NOMOS XIII, Privacy.

Cafaggi, F., & Renda, A. (2012). Public and Private Regulation: Mapping the Labyrinth. In *CEPS Working Document* (Issue 370).

Caine, K., & Hanania, R. (2013). Patients want granular privacy control over health information in electronic medical records. *Journal of the American Medical Informatics Association*, 20(1), 7–15. https://doi.org/10.1136/amiajnl-2012-001023

Choudhury, S. R. (2020, March 25). Singapore says it will make its contact tracing tech freely available to developers. *CNBC*. https://www.cnbc.com/2020/03/25/coronavirus-singapore-to-make-contact-tracing-tech-open-source.html

Claerhout, B., & DeMoor, G. J. E. (2005). Privacy protection for clinical and genomic data: The use of privacy-enhancing techniques in medicine. *International Journal of Medical Informatics,* 74(2–4), 257–265. https://doi.org/10.1016/j.ijmedinf.2004.03.008

Clifford, C. (2020, April 3). COVID-19 pandemic proves the need for 'social robots,' 'robot avatars' and more, say experts. *CNBC*. https://www.cnbc.com/2020/04/03/covid-19-proves-the-need-for-social-robots-and-robot-avatars-experts.html

Cohen, J. E. (2012). *What privacy is for*. Harv. L. Rev., 126, 1904.

Cohen, J. (2008). Privacy, Visibility, Transparency, and Exposure. *The University of Chicago Law Review*, 75(1), 181-201. Retrieved September 5, 2020, from http://www.jstor.org/stable/20141904

Consent. (2020, July 20). ICO. https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/lawful-basis-for-processing/consent/

Corrin, J. (2009). From Horizontal and Vertical to Lateral: Extending the Effect of Human Rights in Post Colonial Legal Systems of the South Pacific. *The International and Comparative Law Quarterly*, 58(1), 31–71. JSTOR.

Cossy-Gantner, A., Germann, S., Schwalbe, N. R., & Wahl, B. (2018). Artificial intelligence (AI) and global health: How can AI contribute to health in resource-poor settings? *BMJ Global Health,* 3(4), 1–7. https://doi.org/10.1136/bmjgh-2018-000798

COVID-19 and Digital Rights. (n.d.). *Electronic Frontier Foundation*. Retrieved 26 August 2020, from https://www.eff.org/issues/covid-19

Criddle, C., & Kelion, L. (2020, May 7). Coronavirus contact-tracing: World split between two types of app. *BBC News*. https://www.bbc.com/news/technology-52355028

Das, S. (2019, November 20). 'Empowered' by law to monitor, decrypt digital information, says govt. *Livemint.* https://www.livemint.com/news/india/govt-says-authorized-to-intercept-monitor-digital-activity-11574178058575.html

Elmasri, R., & Navathe, S. B. (2004). Fundamentals of database systems:[4-th edition].

Ferreira, W., & Rosales, A. (n.d.). *International Telemedicine: A Global Regulatory Challenge*. Lexology. Retrieved 2 September 2020, from https://www.lexology.com/library/detail.aspx?g=f2d9946b-e5c3-43f5-b813-9528e23afbda

Google Flu Trends and Privacy. (n.d.). *Electronic Privacy Information Center*. Retrieved 2 September 2020, from https://epic.org/privacy/flutrends/

Gottlieb S. (2001). US employer agrees to stop genetic testing. *BMJ (Clinical research ed.)*, 322(7284), 449.

Greenleaf, Graham, Global Tables of Data Privacy Laws and Bills (6th Ed January 2019) (February 9, 2019). (2019) Supplement to 157 Privacy Laws & Business International Report (PLBIR) 16 pgs, Available at SSRN: https://ssrn.com/abstract=3380794

Gymrek, M., McGuire, A. L., Golan, D., Halperin, E., & Erlich, Y. (2013). Identifying personal genomes by surname inference. *Science*, 339(6117), 321-324.

Heaven, W. D. (2020, March 17). A new app would say if you've crossed paths with someone who is infected. *MIT Technology Review*. https://www.technologyreview.com/2020/03/17/905257/coronavirus-infection-tests-app-pandemic-location-privacy/

Hern, A. (2019, July 23). 'Anonymised' data can never be totally anonymous, says study. *The Guardian*. https://www.theguardian.com/technology/2019/jul/23/anonymised-data-never-be-anonymous-enough-study-finds

Hern, A. (2014, June 27). New York taxi details can be extracted from anonymised data, researchers say. *The Guardian*. http://www.theguardian.com/technology/2014/jun/27/new-york-taxi-details-anonymised-data-researchers-warn

High Level Expert Group on Artificial Intelligence. (2019). *Policy and Investment Recommendations for Trustworthy AI*. 52. https://ec.europa.eu/digital-single-market/en/news/policy-and-investment-recommendations-trustworthy-artificial-intelligence

Koops, B.-J., Newell, B. C., Timan, T., Chokrevski, T., & Gali, M. (2017). A Typology of Privacy. *Penn Law: Legal Scholarship Repository*, 2017, 38:2, 93.

Krishnaswamy, S. (2007). Horizontal application of fundamental rights and state action in India. *Human Rights, Justice, & Constitutional Empowerment,* May, 47–73. http://ezp-prod1.hul.harvard.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2007-10043-003&site=ehost-live&scope=site

Lazer, D., & Kennedy, R. (2015, January 18). What We Can Learn From the Epic Failure of Google Flu Trends. *Wired*. https://www.wired.com/2015/10/can-learn-epic-failure-google-flu-trends/

Linder, C. (2020, March 18). This MIT App Tracks the Spread of Coronavirus While Protecting Your Privacy. *Popular Mechanics*. https://www.popularmechanics.com/technology/apps/a31742763/covid-19-app-private-kit-safe-paths/

Lomas, N. (2020, April 29). UK privacy and security experts warn over coronavirus app mission creep. *TechCrunch*. https://techcrunch.com/2020/04/29/uk-privacy-and-security-experts-warn-over-coronavirus-app-mission-creep/

Lomas, N. (2019, July 24). Researchers spotlight the lie of 'anonymous' data. *TechCrunch*. https://techcrunch.com/2019/07/24/researchers-spotlight-the-lie-of-anonymous-data/

M, A. (2020, January 14). Has India's Privacy Bill Considered the Dangers of Unrestricted Processing of 'Anonymised' Data? *The Wire*. https://thewire.in/government/privacy-bill-anonymous-data

MacCallum Jr, G. C. (1967). Negative and Positive Freedom, 76 Phil. *REV*, 312, 314-19.

Mandl, K. D., Overhage, J. M., Wagner, M. M., Lober, W. B., Sebastiani, P., Mostashari, F., Pavlin, J. A., Gesteland, P. H., Treadwell, T., Koski, E., Hutwagner, L., Buckeridge, D. L., Aller, R. D., & Grannis, S. (2004). Implementing Syndromic Surveillance: A Practical Guide Informed by the Early Experience. *Journal of the American Medical Informatics Association: JAMIA*, 11(2), 141–150. https://doi.org/10.1197/jamia.M1356

Mankotia, A. S., & Mandavia, M. (2019, November 2). Whatsapp Privacy: After Pegasus spying row, India asks WhatsApp to explain privacy breach. *The Economic Times*. https://economictimes.indiatimes.com/tech/internet/after-pegasus-spying-row-india-asks-whatsapp-to-explain-privacy-breach/articleshow/71851802.cms?from=mdr

Mantovani, F. (2008). Diritto penale: Delitti contro la persona. Parte speciale, 1. Cedam.

Matthan, R. (2019, January 2). The utter meaninglessness of anonymizing telecom data sets. *Livemint*. https://www.livemint.com/Opinion/h9A1vfAHPgSKZCMEvmqO7M/Opinion--The-utter-meaninglessness-of-anonymizing-telecom-d.html

Matthan, R. (2020, May 12). Aarogya Setu and the value of syndromic surveillance. *Livemint*. https://www.livemint.com/opinion/columns/aarogya-setu-and-the-value-of-syndromic-surveillance-11589304017285.html

Moore, A. D. (2010). *Privacy rights: Moral and legal foundations.* Penn State Press.

Nissenbaum, H. (2009). *Privacy in context: Technology, policy, and the integrity of social life*. Stanford University Press.

Nissenbaum, H. (2004). Privacy as Contextual Integrity. *Washington Law Review,* 79, 119.

Nyholt, D. R., Yu, C. E., & Visscher, P. M. (2009). On Jim Watson's APOE status: Genetic information is hard to hide. *European Journal of Human Genetics*, 17(2), 147–149. https://doi.org/10.1038/ejhg.2008.198

Panday, J. (2017, August 28). India's Supreme Court Upholds Right to Privacy as a Fundamental Right—And It's About Time. *Electronic Frontier Foundation*. https://www.eff.org/deeplinks/2017/08/indias-supreme-court-upholds-right-privacy-fundamental-right-and-its-about-time

Pohl, K., Opdahl, A., & Rossi, M. (2001). Sixth international workshop on requirements engineering: Foundation for software quality (REFSQ'00). *Requirements Engineering*, 6(1), 1–2. https://doi.org/10.1007/PL00010353

Post, R. C. (2000). Three concepts of privacy. Geo. LJ, 89, 2087.

Post, R. C. (1990). Rereading Warren and Brandeis: Privacy, property, and appropriation. *Case W. Res. L. Rev.*, 41, 647.

Raskar, R., et al (2020). Apps Gone Rogue: Maintaining Personal Privacy in an Epidemic. *Private Kit: MIT*, arXiv:2003.08567 [cs]. http://arxiv.org/abs/2003.08567

Roratto, R., & Dias, E. D. (2014). Security information in production and operations: A study on audit trails in database systems. *Journal of Information Systems and Technology Management*, 11(3), 717–734. https://doi.org/10.4301/S1807-17752014000300010

Scherer, M. U. (2015). Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies. *SSRN Electronic Journal*, 29(2). https://doi.org/10.2139/ssrn.2609777

Scott, C. (2016). *Private Regulation of the Public Sector : A Neglected Facet of Contemporary Governance* Author(s): Colin Scott Source: Journal of Law and Society, Vol. 29, No. 1, New Directions in Regulatory Theory. Published by: Wiley on behalf of Cardiff University. 29(1), 56–76.

Sharma, R. (2018, April 23). Council Post: Who Really Owns Your Health Data? *Forbes*. https://www.forbes.com/sites/forbestechcouncil/2018/04/23/who-really-owns-your-health-data/#34ccc3176d62

Skloot, R., & Turpin, B. *The immortal life of Henrietta Lacks*. 2010. London, England.

Solove, Daniel J., Understanding Privacy. Daniel J. Solove, Understanding Privacy, Harvard] University Press, May 2008, GWU Legal Studies Research Paper No. 420, GWU Law School Public Law Research Paper No. 420, Available at SSRN: https://ssrn.com/abstract=1127888

Soltani, A., Calo, R., & Bergstrom, C. (2020, April 27). *Contact-tracing apps are not a solution to the COVID-19 crisis*. Brookings. https://www.brookings.edu/techstream/inaccurate-and-insecure-why-contact-tracing-apps-could-be-a-disaster/

Sweeney, L., Abu, A., & Winn, J. (2013). Identifying Participants in the Personal Genome Project by Name (A Re-identification Experiment). *SSRN Electronic Journal*, 1–4. https://doi.org/10.2139/ssrn.2257732

Takashima, K., Maru, Y., Mori, S., Mano, H., Noda, T., & Muto, K. (2018). Ethical concerns on sharing genomic data including patients' family members. *BMC Medical Ethics*, 19(1), 1–6. https://doi.org/10.1186/s12910-018-0310-5

Teo, Y. (2018). *Regulating Artificial Intelligence: An Ethical Approach.*

Tiwari, P. B., Tripathi, A., Bajpai, H., Venkatesh, K., Tripathy, A., & Rizvi, K. (2020, May 6). *Privacy Framework for the Aarogya Setu App.* The Dialogue. https://thedialogue.co/wp-content/uploads/2020/05/Privacy-Framework-for-the-Aarogya-Set-App.pdf

Tripathy, A. (2020, May 1). Contact Tracing for Covid-19: A Global Snapshot of Utility-Privacy Trade-Off. *PSA Legal Counsellors*. http://www.psalegal.com/contact-tracing-for-covid-19-a-global-snapshot-of-utility-privacy-trade-off/

Troncoso, C. et al (2020). *Decentralized Privacy-Preserving Proximity Tracing.* https://github.com/DP-3T/documents/blob/master/DP3T%20White%20Paper.pdf

Vaishya, R., Javaid, M., Khan, I. H., & Haleem, A. (2020). Artificial Intelligence (AI) applications for COVID-19 pandemic. *Diabetes and Metabolic Syndrome: Clinical Research and Reviews*, 14(4), 337–339. https://doi.org/10.1016/j.dsx.2020.04.012

Wachter, S. (2018). Normative challenges of identification in the Internet of Things: Privacy, profiling, discrimination, and the GDPR. *Computer Law & Security Review*, 34(3), 436–449. https://doi.org/10.1016/j.clsr.2018.02.002

Walsh, B. (2014, March 13). Google Flu Trends Failure Shows Drawbacks of Big Data. *Time*. https://time.com/23782/google-flu-trends-big-data-problems/

Wang, Z., & Tang, K. (2020). Combating COVID-19: health equity matters. *Nature Medicine*, 26(4), 458. https://doi.org/10.1038/s41591-020-0823-6

Westin, A. F. (1967). *Privacy and freedom*. Atheneum. New York, 7, 431-453.

What is valid consent? (2020, July 20). ICO. https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/consent/what-is-valid-consent/

When Does the GDPR Not Apply? (2020, February 13). TermsFeed. https://www.termsfeed.com/blog/gdpr-exemptions/

Williams, H., Spencer, K., Sanders, C., Lund, D., Whitley, E. A., Kaye, J., & Dixon, W. G. (2015) Dynamic Consent: A Possible Solution to Improve Patient Confidence and Trust in How Electronic Patient Records Are Used in Medical Research. *JMIR Medical Informatics*, 3(1), e3. https://doi.org/10.2196/medinform.3525

Data
Governance
Network
Anchored by IDFC Institute

Digital
Pathways
at Oxford