IDFC Institute

DGN Policy Brief 12

March 2021

Data Governance Network

Anchored by IDFC Institute

Balancing Privacy with Technology Use: Lessons from COVID-19

Harshita Agrawal, Prakhar Misra and Ananth Padmanabhan.

Privacy is a non-negotiable right at any given time. Yet, governments may use crises to justify a significantly lower degree of individual privacy to be maintained. However, saving lives and maintaining privacy are not mutually exclusive tasks. Given that many countries which have legislation that allows for compromising data privacy for public good, lack implementation frameworks for minimising privacy loss in crises or lack data protection laws altogether, a different set of principles must be followed by the government while deploying technology during national crises. While they are based on the more stringent principles of preservation of privacy, introducing specific features of the crisis and of the technologies used are useful to create a set of new principles.

Context

Agility and a solutions-oriented approach are necessary in crises. Governments have demonstrated this during the COVID-19 pandemic through the use of technology to curb the spread of the virus. They cite the pandemic as justification for collecting personal and sensitive personal data with few privacy safeguards. For instance, autonomous robots have been used for surveillance and taking body temperatures (Ferreira, W., & Rosales, A.). Health data has been collected via telemedicine and contact tracing with little or no safeguards, impeding not just privacy but also freedom of movement by tracking infected individuals and restricting their scope of movement.

While a crisis like this pandemic may allow governments leeway to protect lives, they must operate with defined boundaries. Ad hoc policy making can lead to both a disproportionate government response to the crisis and entrenched measures that extend beyond the crisis. Moreover, preserving privacy also serves the purpose of operationalising autonomy and protecting individuals from coercion. It is intrinsic to an individual's freedom and dignity, and must be upheld to allow one to act independently (Post, 2000).

In this policy brief, we state the implications of three types of technology use – contact tracing, Artificial Intelligence (AI) for drug discovery and disease prevention, and syndromic surveillance – on individual privacy. We use the framework laid out by Koops et al.[1] which lays out nine types of privacy harms caused due to technology use. We present the privacy implications of the three technologies and then elaborate on a set of eight actionable principles that may regulate technology use in crises such as these. We only focus on informational privacy harms caused due to data collection and technology use, rather than all harms such as physical harm too.

Technology use

Contact tracing

Contact tracing has been used to track the location and movement of individuals who may have been exposed to an infected person and are likely to infect others. We base our recommendations on three public applications (Aarogya Setu, HaMagen and TraceTogether), which are centralised and three decentralised private applications (Private Kit: Safe Paths by MIT, Pan European Privacy Preserving Proximity Tracing (PEPP-PT) and Decentralised Privacy Preserving Proximity Tracing (DP3T)). While they have been successful in containing the spread of the virus to an extent, they degrade spatial, communicational, proprietary, informational and associational privacy.[2] This is a result of various kinds of data collected such as

[1] For a more detailed understanding of the framework, 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.

[2] Spatial privacy refers to the privacy of personal physical space; communicational privacy is when personal communication is intercepted; proprietary privacy is when property is used without consent and one's reputation is tarnished; informational privacy is the control of any information about an individual; and associational privacy refers to freedom of being associated with any group.

location, age, sex, profession and so on, the arbitrary duration of storage, and extensive processing without audits.

AI for Drug Discovery and Disease Prevention

Artificial intelligence has been widely used during COVID-19. We look at three applications for our analysis. First, Abcellera uses proprietary technology to scan blood samples from recovered COVID-19 patients, and evaluate antibodies that may carry curative potential. Second, Google Deepmind uses genomic data from GISAid and Protein Data Bank to determine protein structures of the virus. Third, NextStrain, an open-source platform, synthesises pathogen genome sequence data and makes it open source to be epidemiologically useful. While these applications are useful in advancing scientific discovery, the collection and use of sensitive data such as individual health data causes informational, bodily, proprietary and associational privacy violations.

Syndromic Surveillance

During COVID-19, syndromic surveillance has been used to inform containment strategies. Using health data like symptoms and drug purchases, and mapping these data points to location to be able to identify a potential disease outbreak is known as syndromic surveillance.

BlueDot was the first AI algorithm to predict the COVID-19 outbreak by looking at a rise in the number of pneumonia cases in an area in Wuhan, even before the World Health Organisation (WHO). The Ministry of Health, Labour and Welfare in Japan used web searches for symptoms along with their location from the Yahoo! JAPAN App to identify two clusters on the Hokkaido island in Japan. Similarly, the Singapore General Hospital, a public hospital, used body temperature, location and other data from Human Resources to track potential infections in healthcare workers. Syndromic surveillance is known to reduce bodily, information, spatial and associational privacy.

Recommendations

Analysing the three technologies using the Koops framework, we propose a list of eight actionable principles to mitigate various harms to privacy.

1. Purpose limitation should be maintained: Data collection, sharing and processing during emergencies must be solely for fighting the crisis.[3] To do this, governments should have detailed explanations on what data points are being collected and what purposes they would serve during such emergencies. This needs to be governed by the 'analog' components of the system such as legal frameworks, regulation, disincentives for collecting, storing and processing such data and so on.

2. Consent mechanisms need to be changed: The operating principle here should be granular consent over bundled consent.[4] Taking granular consent for each data point collected, its purpose, duration and location of storage should be protocol.[5] Even if this causes user fatigue, 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 as it gives the user more autonomy and allows transparency. Moreover, where these are not being operationalised, justifications should be provided by both public and private entities.

3. Accountability of government action should be maintained: Governments should hold themselves accountable for adoption of technological developments. If a new technology is being released which preserves privacy more than other technologies, 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.

4. Duration of storage of data must be informed by expert consultation: Data storage 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-

[3] Singapore's Data Protection Provisions of the Personal Data Protection Act (PDPA), 2005, states that personal data must be used only for the purpose stated in the application.

[4] Bundled consent refers to umbrella consent taken for all data rather than for each data type.

[5] Caine and Hanania, 2012, and Wachter, 2018 in their research have advocated for granular consent mechanisms to preserve individual privacy, especially with regards to health data.

19, this process must involve epidemiological experts, who would approximate the time for which health-related data points or geo-location should be stored.

5. Internal safeguards for authorisation and access to data should be specified: Encoding audit trails would help in tracking who accessed what kind of data and for what purpose. Collection and processing agencies, whether government or private, will have to take the lead on this. This must be done by default and any decision not to do so should be backed by reasonable justifications.

6. Decentralised data storage data on the phone and not on central servers: This is mostly a feature of the technology being deployed. Yet, governance and regulatory mechanisms must create incentives for this feature to be built in, so as to avoid unnecessary storage of data and linking of datasets for future use.
7. Governments should come up with a 'negative list' of databases: There should be a list of datasets which require special approval for linking with the data collected by COVID-19 applications, keeping in mind the potential of surveillance and the fear of a big brother state. Protocols and exceptions to link those databases should be specified clearly. In case of exceptions, consent from individuals must be taken. If there is a possibility that individuals don't comprehend the uses and its consequences, consent for linking databases must be taken from an independent authorised entity such as a Data Protection Authority, like the one that Europe's General Data Protection Regulation (GDPR) suggests. If linking is done without consent, penalties should be appropriately tailored to the severity of the infringement.

8. All data should be anonymised and encrypted by default: Even in cases when centralised applications are used, personally identifiable data must be anonymised with frequently refreshed temporary IDs.[6] Strict and explicit conditions under which data is de-anonymised and decrypted, and by whom should be specified as a part of Frequently Asked Questions of the technology use. This should be available to citizens.

Conclusion

While technological intervention has proven successful in containing the spread of this virus and saving lives, it is impossible to ignore its negative implications for individuals' right to privacy. As with other fundamental rights, individual privacy must be protected even in times of crises, so that the abrogation does not spill over to normal times as well. The brief lays out the various kinds of privacy harms that emerge due to use of various technologies during COVID-19 and principles that governments can follow in such times to minimise them -- even if eliminating them or complete protection may not be possible due to the imperative to save lives.

[6] Singapore's TraceTogether application generates temporary IDs which decrypt an individual's data.

References

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

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

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.

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

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

Data Governance Network

The Data Governance Network is developing a multi-disciplinary community of researchers tackling India's next policy frontiers: data-enabled policymaking and the digital economy. At DGN, we work to cultivate and communicate research stemming from diverse viewpoints on market regulation, information privacy and digital rights. Our hope is to generate balanced and networked perspectives on data governance - thereby helping governments make smart policy choices which advance the empowerment and protection of individuals in today's data-rich environment.

About Us

IDFC Institute has been set up as a research-focused think/do tank to investigate the political, economic and spatial dimensions of India's ongoing transition from a low income, state-led country to a prosperous market-based economy. We provide in-depth, actionable research and recommendations that are grounded in a contextual understanding of the political economy of execution. Our work rests on three pillars – 'State and the Citizen', 'Strengthening Institutions', and 'Urbanisation'. The State and the Citizen pillar covers the design and delivery of public goods, ranging from healthcare and infrastructure to a robust data protection regime. The Strengthening Institutions pillar focuses on improving the functioning and responsiveness of institutions. Finally, the Urbanisation pillar focuses on the historic transformation of India from a primarily rural to largely urban country. All our research, papers, databases, and recommendations are in the public domain and freely accessible through www.idfcinstitute.org.

About the Authors

Harshita Agrawal is an Associate at IDFC Institute. Her research focuses on urban planning and governance, the link between infrastructure investment and job creation, and data governance. She has a postgraduate diploma from Meghnad Desai Academy of Economics, affiliated to University of Mumbai. She holds a Bachelor of Arts (Hons.) from St. Xavier's College (Autonomous), majoring in Economics and Sociology, with a minor in English Literature.

Prakhar Misra is a Senior Associate at IDFC Institute and leads Data Governance Network, for which IDFC Institute is the Secretariat. He is also an Advisory Board Member of Commonwealth Drone Partnership. He holds a Master in Public Policy from the Blavatnik School of Government, University of Oxford where he was a Chevening Scholar and a Postgraduate Diploma in Economics and Finance from the Meghnad Desai Academy of Economics where he was a Chanakya Scholar.

Acknowledgments

The authors would also like to thank Vikram Sinha for valuable feedback, comments and editorial inputs. We are indebted to Digital Pathways at University of Oxford for funding this study, without whom this work would not have been possible.

Disclaimer and Terms of Use

The views and opinions expressed in this paper are those of the author and do not necessarily represent those of the IDFC Institute.

IDFC Institute
301, 3rd Floor, Construction House 'A', 24th Road, Off Linking Road,
Khar West, Mumbai 400052

/idfcinstitute @idfcinstitute /IDFCInstitute