mPRIVACY.ORG
AN OPENGROWTH.ORG INITIATIVE

JUNE 2020

Privacy-Preserving Data Exchange

Standards for Telecommunications & Mobile Network Operators

Sai Sai Sathya (S20.AI) Alka Asthana (DeZaView.ai)
Dr. Ramesh Raskar (MIT Media Lab)
Dr. Santanu Bhattacharya (Indian Institute of Science)
Gunjan Sinha (MetricStream) Vidya Phalke (MetricStream)

mPrivacy.org is an OpenGrowth.org (EIN: 47-1195860) initiative
Comments or feedback can be sent to info@mPrivacy.org

Applying state of the art privacy enhancing technologies and safeguards such as encryption helps mitigate privacy risks, which reinforces consumer trust.

Jade Nester
Privacy Expert, FIP, CISSP

Nextgeneration tools for epidemic containment at their core would be privacy first solutions that abstract the beneficial value of data while preserving privacy.

Dr. Ramesh Raskar
MIT Media Lab

Fast paced innovation helps solve critical problems for society; comprehensive standards ensure global adoption & scalability.

Gunjan Sinha
Chairman, MetricStream

Extracting deep insights from data can be the tip of the spear in war against pandemics.

Steven Ferguson
Deputy Director NIH, Office of Technology Transfer

Innovative uses of data, including telecommunications data, will prove critical to responding to the world's future crises, but privacy is a key consideration.

Shellye Archambeau
Board Member Verizon, Okta

Foundation for resilience is tied to continuous assurance on risk management - something that works comprehensively and at scale. Towards that end standards play a vital role.

Vidya Phalke
Chief Innovation Officer, MetricStream

Executive Summary

As COVID-19 spread throughout the world, governments, academia, the private sector, and other stakeholders quickly began to explore how to leverage Information and Communications Technology to address the pandemic. Data about where people travelled, and with whom they socialised piqued the interest of researchers and led to the creation of myriad COVID-19 apps. Mobile network operators leveraged anonymised, aggregated data about the locations and movement patterns of their users to help health authorities model the spread of the virus and assess the effectiveness of containment measures. As governments work to contain the virus, stakeholders continue to raise important questions about how data can be leveraged to help in a privacy-protective way.

While some of the contact tracing apps under development include privacy-preserving technologies and architectures, these apps are out of reach for people who do not own smartphones. There are over 5 billion mobile networks subscribers in the world, and as governments seek to restore economic and social life, near real-time analysis of mobile data could provide critical insights. However, it has proven difficult to aggregate data across countries, regions or even globally. One reason for this is that sharing data- even anonymised and aggregated data- can raise privacy, security, and accountability concerns.

There is a need to develop data exchange standards that are privacy-preserving, risk aware and designed in a way that they can deliver mobile data insights at scale. We need to think of new standards to help the world react to crises throughout the next decade and beyond. Developing a standardised approach could unlock great potential. An extremely poignant example, from over 100 years ago, is that of a Chief Electricity Officer – who played a key role morphing organizations from steam power to electricity power. These “CEOs” tackled significant obstacles due to the lack of standards in grid voltages, equipment, and training. As those standards developed, it became much easier to use electricity, leading to exponential societal and economic benefits.

At mPrivacy we seek to build on existing work of organisations like the GSMA, NIST, and ENISA to explore how to develop guidelines, standards, reference architectures, and frameworks that will catalyse the agile privacy-preserving data exchange that is needed - not just for the coming year but for decades to come. Clearly the tailwinds are with us with advancements with Big Data, Artificial Intelligence, and Digitization forging ahead. And as the pandemic has demonstrated, our global interconnectedness presents both risks and opportunities. The world has experienced the risks of interconnectedness before- for example during the 2008 financial crisis. As we look toward the future- a climate crisis looms, and it will drive systemic changes impacting everyone around the world. It is incumbent on us to prepare. We can use the opportunity presented by near-ubiquitous mobile phone connectivity to rise to this challenge. The interconnectedness of the virtual world creates an opportunity to develop interventions in the physical world, benefiting society.

In the accompanying papers we describe how a framework harnessing state of the art privacy enhancing technologies (PETs) like homomorphic encryption, secure multi-party computation, and differential privacy can be used to minimize privacy risk while maximising social good benefits. These PETs can enable governments to query telecommunications operators’ (telco) data without accessing any underlying telco data. The framework helps governments receive timely information in a way that protects privacy and sustains trust- both between telcos and their users and between individuals and the government. These benefits could exponentially increase if the framework is widely adopted through a standardized approach.

Developing an open, voluntary standard based on the framework proposed in the accompanying papers benefits:

- Consumers, through the application of enhanced privacy safeguards
- Governments, through timely delivery of insights
- Telecommunications operators, through mitigation of privacy, security and accountability risks, as well as through the development of a standardized platform for data exchange that can also be leveraged for new, sustainable data-based business models.

We recognize that many telcos are currently developing or building their data analytics teams and producing relevant insights for customers, including governments. The proposed framework does not preclude or diminish these offerings. Instead, as a non-profit, mPrivacy intends to work with telcos and other stakeholders to develop an open, voluntary set of standards, technical architectures, and frameworks. Future plans may include cross-sectoral data exchange, to enable new, sustainable data-based business models.

We plan to convene stakeholders such as telecommunications engineers, data scientists, coders, developers, privacy advocates, academia, NGOs and others to join this open, voluntary effort and help achieve the goal of creating data exchange standards that are privacy-preserving, risk aware, and will help respond to global challenges now and in the future.

Table of Contents

1. Introduction: Leveraging mobile big data analytics responsibly to tackle the COVID-19 crisis

COVID-19 is a global calamity on an unimaginable scale, with devastating consequences. Private companies, particularly telecommunications operators (Telcos), are producing significant amounts of data every day which can help answer key questions about current and emerging hotspots. While these data may be used for valuable analysis and model building, the data may also reveal sensitive information about individuals, depending on the level of anonymity of the data. Additionally, most countries have requirements around who can directly access this data. To sustain trust between telcos, citizens, and governments, and to address concerns from civil society and other important stakeholders, analysis of telco data should be conducted in a privacy and risk aware manner.

One way of responsibly leveraging Telco data to research COVID-19 is to draw valuable insights from aggregated, anonymised mobility data, reflecting the movements of groups of individuals, where no one user can be identified or singled-out. Existing studies reflecting the efficacy of social distancing measures to fight COVID-19 utilize aggregated data —both GPS [Lai et al., 2020] data and location data [Reiter, 2020] derived from call detail records (CDRs).[1]
In addition to the research utilizing aggregated mobility data, a number of governments, companies, and non-governmental organisations have developed specialised COVID-19 mobile apps designed, for example, to provide users with information about whether they may have crossed paths with an infected individual, and to provide other information, such as the location of the nearest testing center. However, instead of telco data, these apps generally leverage GPS data and/or Bluetooth data, with varying degrees of privacy protections.

While these apps may reach a wide number of users across many of the countries most affected by COVID-19, they are designed for smartphones[2], rather than feature phones[3]. Some of the privacy protective elements of the apps, such as the ability to store and query a user's GPS data on their own device, without sharing data with government authorities, require computing and storage capacities beyond what is available on most feature phones. In low and middle-income countries (LMICs)[4], where COVID-19 is beginning to make an impact, most users will not have access to COVID-19 apps developed for smartphones. For example, in a survey conducted across the 18 LMICs, the GSMA found that 45% of those surveyed did not own a phone that is capable of connecting to the internet [GSMA, 2019a].
In Sub-Saharan Africa, 39% of mobile connections are made using a smartphone, and in South Asia, 49% of connections involve a smartphone [GSMA, 2019b]. As a result, many

[1] Call Detail Records (CDR) data- refers to a record of a voice call or an SMS generated by a mobile network operator that includes the mobile number of both the person making and receiving the call, date, time and call duration, and low resolution location information (nearest cell tower).
[2] A smartphone is defined as a mobile handset enabling advanced access to internet-based services with computer-like functions.
[3]While both smartphones and feature phones provide mobile Internet access, feature phones use closed platforms that do not support native development, although downloadable applications are often supported using Java.
[4] This is a classification used by the World Bank: Low & middle income

mobile users in LMICs use feature phones, or basic phones that allow voice call, SMS and USSD features. Basic phones may have Bluetooth but not internet/WiFi capability [GSMA, 2017].

In addition to existing efforts by telecommunications operators to deliver critical connectivity services during the COVID-19 crisis and to leverage their data for research and analysis, there is also an opportunity for telecommunications operators to develop a privacy-protective way to provide users with relevant, trusted information about the pandemic via basic phones and feature phones, without telecommunications operators sharing identifiable user data with governments.[5] To the extent that telecommunications operators are required to share user location data with governments, privacy-protective measures may be implemented to mitigate risks, while meeting government objectives to stem the spread of disease.

The application of Privacy by Design [Cavoukian, 2009] helps to identify and mitigate privacy risk, and the utilization of Privacy Enhancing Technologies ('PETs') can further protect privacy while also enabling analysis for mobile big data. Applying sophisticated PETs and artificial intelligence ('AI') to mobile big data can responsibly create products that would aid the government, health care providers, emergency responders and the public to better prepare to deal with short and long-term consequences of the disease. To ameliorate privacy and security concerns, we propose a framework that will support this urgently needed analysis in compliance with the laws, while minimizing risk to individuals or groups to the greatest extent possible using suitable privacy, security and risk management techniques.

The use of telecommunications operators' mobile big data to help address the COVID-19 crisis has been the subject of great interest from the research community, as well as scrutiny from the public, regulators, and privacy advocates. In addition to their existing research[6] and connectivity initiatives[7] focused on COVID-19 response, there is an ongoing opportunity for telecommunications operators' to move with speed and certainty to work with Governments around the world to address this crisis. It seems clear that social distancing will continue to be necessary until safe vaccines and therapies become available to address this public health crisis. Furthermore, for the future, it is important for the world to be prepared for other similar public health pandemics. Therefore, telecommunications operators' should develop a long-term strategy to better leverage their data in privacy protective ways. This white paper is intended to generate discussion, and to outline a privacy-protective framework that could be utilized by telecommunications operators and governments.

[5] In some cases, relevant laws may require telecommunications operators to share user data with governments to prevent the spread of disease. These legal requirements should be necessary, proportionate and time-bound. See also [GSMA 2020].
[6] E.g., Telenor; Deutsche Telekom; Telefónica.
[7] E.g., Airtel; Orange; Safaricom.

2. How would it help?

We start this section with the description of some use cases for Mobile Network Operator data in a range of countries around the world, depending on their applicable legal frameworks.

First, the individual level information has been used in the context of contact tracing, for example:

- Once infected people have been identified, to trace people living in the vicinity of infected individuals.
- To try to identify people who have come in contact with the infected individuals.
- To trace particular individuals who are not contactable post treatment at hospitals.

Second, is the case around localized group-level information for ensuring a successful lockdown. Across the world, many areas have been designated as high-risk containment areas or the 'red-zones' either by the governments or research institutes. These could be municipal areas or localities. Residents of such areas are required to follow protocols and practice precautions to reduce spread of the disease and eventually, qualify for gradual opening of their localities, businesses and restring livelihoods.

To ensure and monitor that the localities are practicing COVID-19 safety measures, the local authorities would need to acquire certain information:

- Aggregate number of violations of lockdown, for example, on an hourly basis.
- Aggregate number of people entering a containment area from the outside.
- Aggregate number of people are venturing out of the containment zones.

These use-cases require data at a locality or area basis and often in real time. Such information may be useful for monitoring and containing the progress of the disease.

Finally, the third use-case looks at the aggregated data for ensuring a smooth "un-lockdown" or "re-opening". To manage and map the spread of COVID-19 and enable smooth un-lockdown region-by-region, on an aggregated level, the government would require:

- Information on population movements during lockdown, areas with high and low mobility.
- A standard set of data for data modeling or for social good from which one can derive insights, for example, proxy data for economic prosperity post-COVID, based on information such as mobility, relative amount spent on telecom services, number of transactions, etc.

Using these use cases as a background, the conceptual framework proposed by this white paper would enable governments to query telco data for lawful purposes without directly accessing any MNO or telco data. More specifically, the framework would enable national health authorities to share encrypted infected patient phone numbers via a trusted third-party middleware and an on-premises, privacy-enabled and secure computation platform. The trusted third party (for example, a university) can then run queries on the Telco data to provide the national health authority with encrypted, differentially private responses of areas potentially at high-risk of an outbreak. The health authority could access the list of risk areas via a web-based dashboard. The Telco could then send SMS/USSD/or flash messages to mobile users in high-risk areas providing them with authentic and relevant information about their health

risks. The Telco could also send users questions posed by the health authority, e.g., about reported symptoms, and the trusted third party would send the health authority encrypted, differentially private responses to the questions, to help the health authority monitor the potential spread of the disease. This solution can be provided to users with their consent, on an opt-in basis, and would provide all mobile users with a privacy-protective means to obtain critical information about the pandemic, regardless of whether they have a basic phone, a feature phone, or a smartphone.

The use of mobile notifications and/or warnings related to geographic areas experiencing outbreaks may have originated during SARS outbreaks in the early 2000s. In 2003, a mobile network operator in Hong Kong launched an opt-in service enabling customers to be notified via SMS which buildings within a kilometre of their location had experienced SARS infections, according to the Hong Kong Department of Health. The most recent and widely-publicised geography-based notification system for mobile phones has been deployed in South Korea, where the government provides anonymized data about infected peoples' movements to the public, via cell broadcast notifications to people within the range of cell towers where the infected person travelled. The information is also posted online, in a form that provides precise information about time and location but does not include the individual's name (although the precise time and location information has led to some people being identified). Two apps- Corona 100m and Coronamap- have been developed that use the publicly available government database to provide individuals with warnings if they step within 100 meters of where an infected patient has travelled.

In South Korea, the warnings have resulted in people avoiding certain areas for fear of becoming infected. However, the precise location and time noted in the warnings has led to identification and harassment of some individuals. The solution presented in this paper seeks to provide individuals with similar geography-based warnings about possible COVID-19 infection, without 1) providing mobile data directly to the government; and 2) inadvertently leading to the re-identification of individuals based on granular but de-identified personal data; and 3) without requiring cell broadcast technology, which is only available in a minority of countries. As we stated earlier, the solution also seeks to provide useful information on potential outbreaks to users who do not have access to smartphones.

Beneficiaries

The responsible use of Telco data through Big Data, Artificial Intelligence technologies would have two key direct beneficiaries: 1) The government: In a way of developing concrete systems planning and information gathering and; 2) The citizens, whose data would help the government processes while keeping their privacy intact. These key benefits would indirectly enable the health care professionals, on-ground workers, emergency responders and the private sector to better align themselves in accordance with the requirements of dealing with crisis situations. Below are the important ways in which the government, the citizens, healthcare professionals and the corporates (private sector) would benefit:

2.1 Governments

Governments will improve information gathering and planning significantly:

- Knowing high-risk zones would aid in planning testing locations, medical care and supplies facilities, etc. in addition to alerting people.
- Feedback from the users and other collaborators will improve the data that's being collected at the test centers.

2.2 Citizens

Without sharing Telco user information with governments, citizens will get more authenticated information, leading to potentially better behaviour inspired by concrete understanding of why they need to behave a certain way:

- Increase social distancing with additional region-specific information.
- Prompt to get tested if they are showing any symptoms or have been in contact with people who have shown symptoms.
- SMS, USSD and flash messages would cover not only the smartphone users but also feature phones users, especially in areas with limited internet access.
- The framework could be provided on an opt-in basis- consent could be obtained from users prior to receiving notifications.

2.3. Health Care Professionals and Care-givers

Health care workers and on-ground care providers would be better placed to deal with patients and citizens trying to get themselves tested:

- They would know which areas to provide health care.
- They would have a better idea for allocation of resources, better estimates of the number of healthcare equipment that they might need.
- The solution could directly link the citizens, most likely to have infection with corresponding health facilities to reduce time lags.

2.4 Corporates

Businesses would be better placed to find a transition towards an un-lockdown phase:

- They would be able to identify branches, offices where they can start reviving work and where they still need to practice caution.
- This would also help in identifying areas/sectors of business opportunities.

3. Challenges

The major challenge in the implementation of such a solution would be to strike a balance between maintaining the privacy of the citizens and providing a robust solution that could enable all stakeholders (including the citizens) to be able to obtain relevant information regarding the spread of the virus and take appropriate measures.

Successful models that have been able to track COVID-19 spread efficiently, have made use of radical measures of surveillance apart from building capacity for the manufacture of test kits, medical facilities and other resources. China [WHO et al., 2020] and South Korea [Sonn, 2020] have made use of heavy implementation of surveillance technology to identify who to

test, this includes: CCTV, tracking of bank cards and mobile phone usage. South Koreans are also notified when a person in their district contracts COVID-19 [Buchwald, 2020] and they are given highly detailed information about their whereabouts (without releasing their names) — including the exact bus they may have taken and whether or not they wore a mask. Singapore, Taiwan and Hong Kong contained the virus through slightly less strict surveillance but efficient tracking [Cowling and Lim, 2020].

While considering the best approach for a large democracy, for example India, several challenges are likely to arise:

1) **The laws pertaining to data usage:** In order to adopt successful models around the globe, societal structure and the local laws are important considerations. For example, in India, privacy is a fundamental right of the citizens and unlawful use of data by data fiduciaries is a serious concern. The country is well on its way to implementing the Data Protection Bill [Haritas, 2020] which will impose limitations to the ways in which citizen data is used by public and private entities.
2) **Meeting government objectives while avoiding the implementation of measures that may be considered to be intrusive:** This white paper proposes a highly intelligent solution that solves the problems of tracking COVID-19 patients and their whereabouts, without compromising on the ethics of data use, and mitigating privacy risks, at each stage of the process. It would also be important to ensure government and regulatory approval of such a novel approach.
3) **Protecting human rights**: While the lack of granular location information involved will help prevent re-identification, other risks to individuals and groups remain. For example, users could attempt to single out an infected individual, either correctly or falsely, leading to harassment. Additionally, high-risk locations could become targets for harassment, vandalism, or other crime. It is critical that the government provides high-risk areas with protection and helpful interventions.

Based on the above challenges, a solution that can take countries such as India out of the clutches of this pandemic and prevent similar situations in the future needs to be based on the following principles.

4. Principles

The framework proposed in this document for providing the above-mentioned outcomes was developed on the basis of three key considerations: security, privacy and risk. It is aimed to make the solution privacy-friendly and risk averse to avoid any conflict between the interest of the government and the citizens. A secure and privacy-friendly solution would safeguard the personal data of the citizens and the risk-awareness of it would identify the practical models of operation for the governments.

4.1 Secure and Privacy-friendly

At a high-level, governmental requests for telecommunications data fall into two categories: 1) requests for aggregated, anonymised data; and 2) requests for identifiable personal data related to specific end-users. The legal requirements for both types of requests differ, and the

latter type of request are much more sensitive because it can reveal sensitive personal data about end users. In the context of COVID-19, both types of data have been legally requested by governments (see, e.g., the European Commission request for aggregated, anonymised data [EDPS 2020] and the South Korean use of identifiable mobile data under Article 76-2(2) of South Korea's Infectious Disease Control and Prevention Act). The solution proposed in this paper seeks to create a middle ground- where governments can query telco data without accessing telco data or obtaining telco data.

Privacy enhancing technologies-specifically homomorphic encryption and multi-party computation enable this privacy-protective querying of telco data. Under the EU General Data Protection Regulation, encryption is designated as a "safeguard" to mitigate privacy risks. Under the GDPR (and other privacy laws based on the GDPR), the application of safeguards by themselves does not render processing of data lawful. Another valid legal basis, such as consent, must apply. But comprehensive privacy protections require a holistic approach, by implementing Privacy by Design, continuously assessing privacy risks and seeking to mitigate those risks.

The proposed solution also aligns with other commonly accepted privacy principles. For example, accountability is an important pillar of privacy frameworks around the world. This principle refers to organisations protecting personal data and being able to demonstrate to regulators their organisational and technical measures to protect privacy. Accountability also requires that privacy protections continue to apply if data is shared with any third parties. The solution we propose limits the necessity of data sharing with governments (including health authorities), while still helping them meet valid public policy objectives. Implementing privacy enhancing technologies to safeguard personal data helps minimise data processing and storage, prevents data breaches, and makes it easier for telcos to remain accountable for protecting and respecting end-user privacy. Privacy enhancing technologies continue to advance and develop, and it will prove increasingly important to leverage these tools to enhance privacy protections.

The underlying technology, product architecture as well as the service framework should support telco alignment with privacy principles and meet the local laws and international laws while enabling multiple parties to collaborate and provide a secure two-way communication between telecommunications operators and users with the least amount of friction. Further, it should provide authenticated information from the government and a means to verify the authenticity so as to ensure people can trust the information, act on it and have an easy way to provide responses if they chose to. Given that the solution would provide users about potential high-risk areas for infections, it would also be important to consider the potential human rights impacts, to ensure that people in high-risk areas are not stigmatized. Additional privacy and security elements include:

- End-to-end security will be ensured during any kind of data transfers using state of the art cryptographic techniques and a trusted middleware between the parties.
- To meet privacy requirements, differential privacy [Dwork et al., 2006], secure multi-party computation [Bogetoft et al., 2009], homomorphic encryption [Sathya et al., 2018] and other emerging privacy-preserving technologies[8] and frameworks[9] [Raskar

[8] Split Learning: distributed deep learning and inference without sharing raw data
[9] Private Kit: Safe Paths; Privacy-by-Design Contact Tracing

et al., 2020, Gupta and Raskar, 2018] would be appropriately used with the help of domain experts and partners[10]. Data would only be processed for specified purposes; no secondary processing would be permitted.

4.2 Risk-aware and Risk-managed

The system should be risk aware to identify the practical limits of what kind of models and data sharing are acceptable within the Government and telecommunications operators' risk appetite and managed using a systemic framework of audits, policies, issue and control management and 3rd party management[11]. This systemic awareness will provide for proactively manage the risk and decision making at various levels and stages. Risk management should further include:

- Computing risk appetite using suitable frameworks defined and accepted by the local Government and telecommunications operators. Both NIST and ENISA have published and widely adopted frameworks that can be leveraged. [NIST, 2011], [NIST 2017], [ENISA 2015]
- Identifying emerging patterns of risks, how they affect the system and tracking issues raised by stakeholders and users across the globe that could be part of the risk profiles.
- Controlled testing using predefined conditions to ensure expected behavior of the system and updating the control parameters based on feedback and regular audits. Continuous testing of controls will be extremely relevant here in this context.
- Auditing and testing data flows and processes to ensure compliance to the underlying privacy, security, data governance laws and policies.
- Federated policy management to enable state level or regional policies that is properly governed by central policies of the system.
- Overall policy management that defines all the policies that govern our approach and system globally.
- Methods to identify and onboard 3rd parties for risk management, assessment and governance of the system.

4.3 Consideration of Human Rights

The human rights of individuals in high-risk areas must be protected at all times. A local advisory committee on ethics and human could be developed to assess ethical concerns before the implementation of the framework, and during its use. Additionally, partnership with a health-focused non-governmental organisation or international organisation could provide an additional accountability mechanism to ensure ethical conduct. Such a partnership could also provide helpful epidemiological expertise.

5. Solution

Following are the key components of the solution:

[10] S20.AI: Secure and Privacy enabled AI Collaboration Platform
[11] MetricStream: Governance, Risk and Compliance

- **Government data sources** to collect COVID-19 infected patient list from central and state ministries and various health bodies.
- **Telco data sources** to utilize the mobility patterns and contact graph of patients to identify high risk zones in the country.
- **Models** developed by the Telcos and various collaborators including but not limited to Government bodies, academic institutions, independent research groups, approved public and private organizations.
- **Trusted middleware** to provide access controls and ensure secure, privacy preserving and risk aware transfer of data, models and responses between the government, Telcos and other collaborators.
- **Secure computing infrastructure** to manage data transformations and transfers, query executions and model inferences.

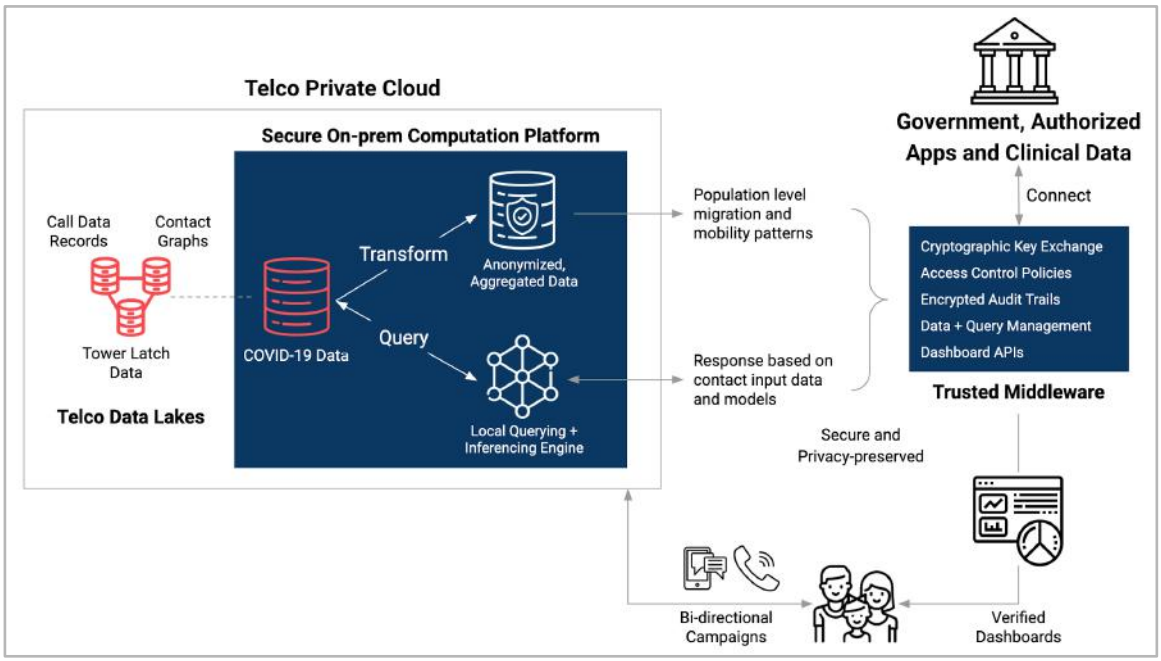

Figure 1: The schematic diagram of the solution using secure, risk-aware and privacy-enabled technologies and frameworks

The system further comprises of a digital dashboard and a mobile solution to enable bi-directional campaigns with the users:

5.1 Digital Dashboard

A web-based dashboard that provides country, state, district and locality level information on COVID-19. Features to be considered:

- Detailed statistics at different levels of granularities - geographic, stages of disease (e.g., in treatment, released).
- Feeds - data sources from governments and medical bodies to help people get valid and authenticated information.

- A visual map of demarcation of zones with different levels of COVID-19 breakout e.g., RED vs. GREEN, so that people receive relevant information and make decisions about the relative safety of their geographic movements.

5.2 Mobile Solution

Use Telco data to determine risk zones and create an infrastructure to send bi-directional, mass notifications for enabling information sharing and campaigns by the government and relevant authorities to people. Combined with a proper notification system that is simple to construct, ubiquitous and operates with ease to aid the humanities battle with COVID-19. Design considerations are:

- Should be feature phone and basic phone friendly
- Should allow all channels for communications, i.e. SMS, USSD and flash message.

5.3 Additional Considerations

Security: The trusted middleware provider becomes a target for cyberattacks, and there is potential for a data breach via the middleware provider. State of the art cybersecurity controls should be implemented to mitigate security risks.

Granularity: The geocodes identified via the middleware query engine would be based on CDR location data, e.g., information about user connections to specific cell towers. This data is less granular than GPS or Bluetooth data, so the geocoded information sent back to the health authority is less specific than the data they would obtain via more granular methods.

Battery preservation: Given that this solution does not require use of GPS or Bluetooth, it may help preserve the user's device battery, while also providing valuable consumer-facing information.

6. Recommendations for Implementation

While it will be easier for the government to take the lead, in the interest of citizen privacy, a US 501(c)(3) Non-profit public organization would be best to solve this problem. Large non-profits have been very successful in the health-care domain. The advantages of Non-Profit are clear – as they create an environment of least friction and most participation for all the entities – public, private or governmental agencies.

- The Non-profit public organization can have its own Governing Board and legal entity to help steer this initiative.
- Pro-bono services from top corporate lawyers can be availed to address issues of legal compliance. The independent body should have a separate management team to enable long-term success of this program, which will be well governed by a world-class Board of Directors representing its stakeholders.

The independent legal entity could solicit primary financial support from a range of stakeholders, including Telcos, foundations and federal governments around the world through a consortium model.

- Corporates interested in solving global health issues using capital and resources could also be members of the consortium.

- Philanthropic organizations and CSR Foundations could support such initiatives
- The independent body would also require support of advisories that can help raise the required capital from various sources and assist in the proper utilization of the funds.

7. Conclusion

The crisis that the world is facing is unprecedented. But there is no reason to believe that we cannot face such situations in the future. Countries around the world have faced healthcare crises at smaller scales intermittently over the past few decades, and it is only now that we are facing a modern pandemic at a scale that makes countries rethink their models of prevention and care.

If anything, the current situation should prompt us to implement robust preventive and healthcare delivery measures, much like the way we prepare ourselves for the nuclear wars. A pandemic has a higher chance of repeating itself in the future than a third world war and thus demands significant investment of time, resources and research from all countries.

The responsible use of data analytics can go a long way in safeguarding our societies from such calamities and has the potential of building models that function in silo with multiple entities to build products and organizations that troubleshoot such problems at the point of origin. Telecom companies have an advantage of being in possession of data that is extremely valuable in tracing citizens' whereabouts. However, it is critical to protect the privacy of this data, and to prevent unethical use. This is the time for us to adopt solutions that enable us to implement data analytics and AI safely, consistent with privacy laws and best practices, while at the same time saving millions of lives now and in the future.

The paper has developed very concrete measures with respect to the use of Big Data Analytics with adherence to privacy laws and addressed the logistical framework that can be used to bring this into operation. The model presented in the paper can be used by any country in ways that does not compromise any local legal requirements or internationally accepted privacy best practices. See [OECD 2013] and [GSMA 2016]

8. Acknowledgements

The authors are indebted to Ms. Jade Nester (Director, Consumer Policy - GSMA, acting in her personal capacity) for her valuable inputs in writing this paper and sharing her privacy and policy expertise, and the numerous discussions she has led on this topic. In addition, thanks goes to Niyaatii Swami for proofreading the paper and providing editorial remarks.

9. References

[Bogetoft et al., 2009] Bogetoft, P., Christensen, D. L., Damg˚ard, I., Geisler, M., Jakobsen, T., Krøigaard, M., Nielsen, J. D., Nielsen, J. B., Nielsen, K., Pagter, J., et al. (2009). Secure multiparty computation goes live. In International Conference on Financial Cryptography and Data Security, pages 325–343. Springer.

[Buchwald, 2020] Buchwald, E. (2020). What we can learn from south korea and singapore's efforts to stop coronavirus (besides wearing face masks). MarketWatch.

[Cavoukian, 2009] Cavoukian, A. (2009). Privacy by design. Take the challenge. Information and privacy commissioner of Ontario, Canada.

[Cowling and Lim, 2020] Cowling, B. J. and Lim, W. W. (2020). They've contained the coronavirus. Here's how. The New York Times.

[Dwork et al., 2006] Dwork, C., McSherry, F., Nissim, K., and Smith, A. (2006). Calibrating noise to sensitivity in private data analysis. In Theory of cryptography conference, pages 265–284. Springer.

[GSMA, 2017] GSMA (2017). Accelerating affordable smartphone ownership in emerging markets. Mobile for Development, page 25.

[GSMA, 2019a] GSMA (2019a). The state of mobile internet connectivity 2019. Mobile for Development, page 32. https://www.gsma.com/mobilefordevelopment/resources/the-state-of-mobile-internet-connectivity-report-2019/.

[GSMA, 2019b] GSMA (2019b). The state of mobile internet connectivity 2019. Mobile for Development, page 25.

[Gupta and Raskar, 2018] Gupta, O. and Raskar, R. (2018). Distributed learning of deep neural network over multiple agents. Journal of Network and Computer Applications, 116:1–8.

[Haritas, 2020] Haritas, B. (2020). Personal data protection bill: Is your enterprise prepared? The Economic Times.

[Lai et al., 2020] Lai, S., Ruktanonchai, N. W., Zhou, L., Prosper, O., Luo, W., Floyd, J. R., Wesolowski, A., Zhang, C., Du, X., Yu, H., et al. (2020). Effect of non-pharmaceutical interventions for containing the covid-19 outbreak: an observational and modelling study. medRxiv.

[Raskar et al., 2020] Raskar, R., Schunemann, I., Barbar, R., Vilcans, K., Gray, J., Vepakomma, P., Kapa, S., Nuzzo, A., Gupta, R., Berke, A., et al. (2020). Apps gone rogue: Maintaining personal privacy in an epidemic. arXiv preprint arXiv:2003.08567.

[Reiter, 2020] Reiter, J. (2020). Correct use of telecom data can help in this crisis. https://www.politico.eu/sponsored-content/correct-use-of-telecom-data-can-help-in-this-crisis/.

[Sathya et al., 2018] Sathya, S. S., Vepakomma, P., Raskar, R., Ramachandra, R., and Bhattacharya, S. (2018). A review of homomorphic encryption libraries for secure computation. arXiv preprint arXiv:1812.02428.

[Sonn, 2020] Sonn, J. W. (2020). Commentary: South korea succeeded in controlling covid-19 panic buying, thanks to tracking and surveillance. CNA.

[WHO et al., 2020] WHO et al. (2020). Report of the who-china joint mission on coronavirus disease 2019 (covid-19). Geneva: World Health Organization.

[NIST, 2011] Joint Task Force Transformation Initiative, Managing Information Security Risk: Organization, Mission, and Information System View, NIST Special Publication 800-39, March 2011 - https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-39.pdf

[NIST 2017] National Institute of Standards and Technology 2017 An Introduction to Privacy Engineering and Risk Management in Federal Systems - https://nvlpubs.nist.gov/nistpubs/ir/2017/NIST.IR.8062.pdf

[ENISA 2015] ENISA Privacy and data protection by design – from policy to engineering (2015) - https://www.enisa.europa.eu/publications/privacy-and-data-protection-by-design

[GSMA 2020] The GSMA COVID-19 Privacy Guidelines April 2020, GSMA COVID-19 Privacy Guidelines.

[OECD 2013] OECD Privacy Framework (2013) OECD Privacy Guidelines

[GSMA 2016] Mobile Privacy Principles Promoting consumer privacy in the mobile ecosystem (2016) GSMA Mobile Privacy Principles.

[EDPS 2020] Memo from European Data Protection supervisor on Monitoring spread of COVID-19 https://edps.europa.eu/sites/edp/files/publication/20-03-25_edps_comments_concerning_covid-19_monitoring_of_spread_en.pdf

ANNEXURE I

Data Flows

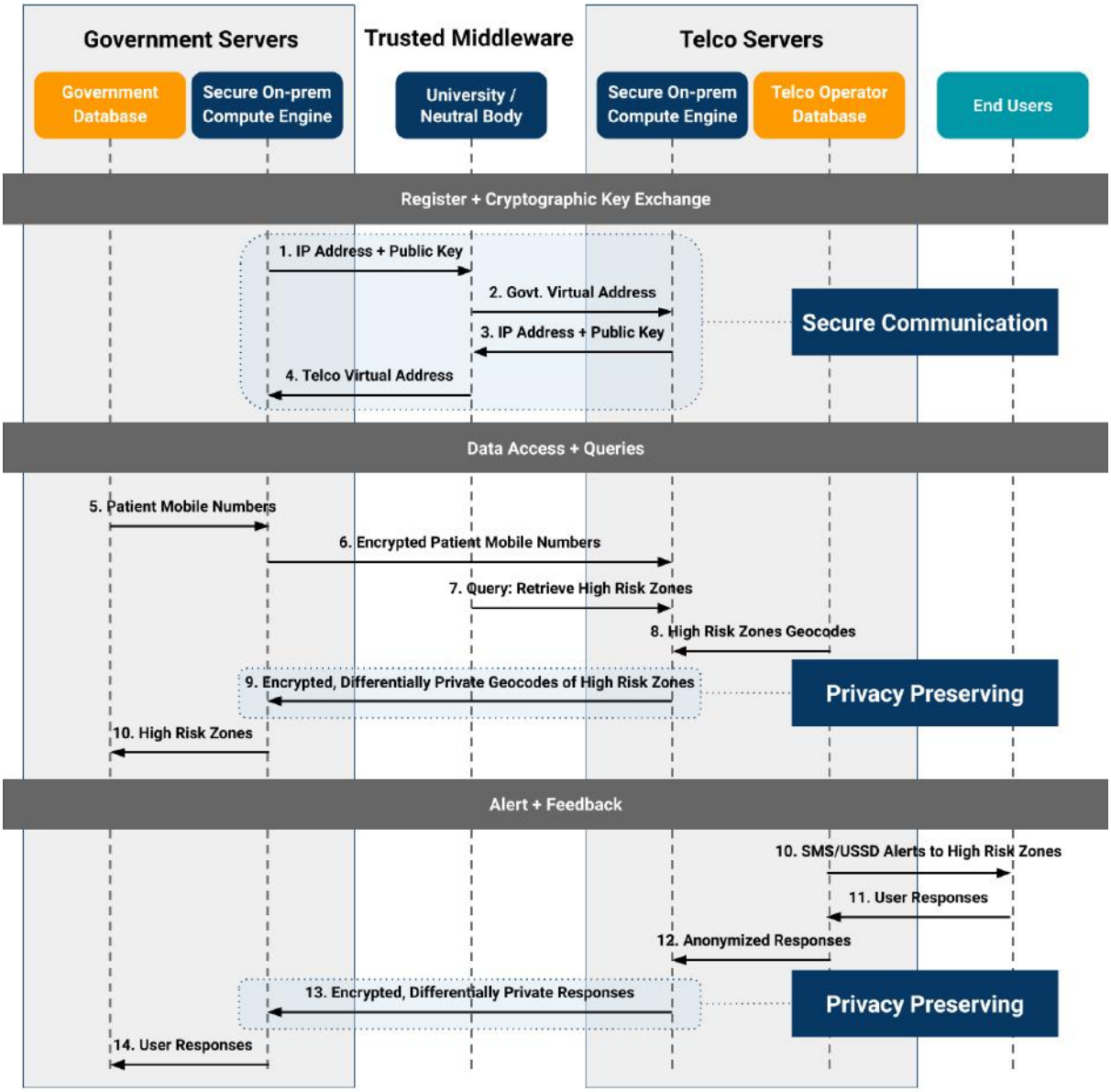

Figure 2: Privacy-preserving data flow between the Government, Telco and S20.AI (trusted middleware and on-premises secure computation platform)