

Policy Brief

Data Access and Openness for the Covid-19 Response

Xenia Rasul

Summary

- Decisions to battle a global pandemic in the digital age should place due importance on the role of data as the backbone that supports a timely and impactful response.
- Pakistan managed to avert a catastrophic blow-up of the Covid-19 pandemic as a result of a data-driven response strategy.
- For an effective data regime, standardization, granularity, openness, accuracy, and timeliness are important factors. Pakistan could benefit from improving on all fronts.
- High-quality open data not only develops a data-first culture, but also enriches the data landscape by bringing together insights from diverse groups that enhances and improves decision-making by pooling expertise and perspectives.
- An agreement on processes and technical aspects of the data eliminates room for confusion and increases reliability on the system and related decisions.

Identifying the problem

Covid-19 has disrupted the world system as we knew it. State institutions have faced immense pressure to maintain order and trust. Economy, health, education, law and order, emergency and essential services have all been affected in a way that will leave a lasting impact through human and financial losses. Given the importance of data to understand the corona virus: its spread, impact, evolution and possible treatments, the need for a data regime that positions humans at a vantage point offering a holistic view is significant. Decisions to battle a global pandemic in the digital age should place due importance on the role of data as the backbone that supports a timely and impactful response.

Pakistan's success in beating the expectations of a catastrophic blow-up of the Covid-19 pandemic amidst concerns of weak state capacity to manage and contain the virus points towards a holistic, robust and effective data-driven response strategy and structure that has not only prevented the loss of life but has been able to open-up the economy gradually to sustain livelihoods. The 'Smart Lockdown' strategy relies heavily on micro and nano-level data from health facilities, travel entry and exit points, contact tracing and administrative oversight of compliance. The initiating and easing of lockdown in localised hotspots is demonstrative of the state's capacity to not only source and process data but to convert it into high-quality intelligence to shape an agile and highly effective national response.

While the administrative data systems overseen by the National Command and Operations Center (NCOC) and managed by the Ministry of National Health Services, Regulation and Coordination, provincial health departments and the National Information Technology Board (NITB) are serving the purpose to contain the pandemic, the public facing side of these systems merits analysis and much improvement. In addition to establishing an architecture for real and near real-time reporting of Covid-19 cases and associated indicators, federal and provincial governments have both used a range of public reporting and dissemination tools. The most notable of these are:

Data Source	Format	Organisation	Status
National Covid-19 Dashboard	Online Dashboard (daily)	NCOC / MoNHSRC & NITB	Public
Daily Situation Report	PDF report (daily)	MoNHSRC / NEOC	On-demand
Field Epidemiology Report	PDF report (weekly)	National Institute of Health (NIH)	Public
Punjab Covid-19 Dashboard	Online Dashboard (daily)	Urban Unit	Public
Daily Situation Report	PDF report (daily)	Sindh Health Department	Public
Daily Situation Dashboard of Covid-19	Social Media Posts (daily)	Health Department, Gilgit-Baltistan	Public
Khyber Pakhtunkhwa Corona Virus Covid-19 Daily Update	Social Media Posts (daily)	Health Department, KP	Public
Covid-19 Summary	Social Media Posts (daily)	Office of Deputy Commissioner, Islamabad	Public

The national data portal for Covid-19 (www.covid.gov.pk) is assessed below based on the core attributes of a strong data regime that relies on data to be standardized, granular, timely, accurate and open. This analysis aims to make a case for improving the public-facing elements of the data regime to enhance quality and accessibility of the data around Covid-19 and its use by public and private actors.

Standardized	Documented process for data collection, processing and reporting – entities, data flows, timelines and rules	No
	Definitions and taxonomy for reported indicators that is clear and universal across different departments and agencies	No
	Notes to changes, exceptions, variances, anomalies in reported data	No
	Aggregation of data according to defined templates	Yes
	Anonymity of data reported	Yes
Granular	Data reporting by province/region	Yes
	Data reporting by sub-provincial/regional	No
	Data reporting by gender	No
	Data reporting by age group	No
	Data reporting by demographic groups	No
	Data reporting disaggregated by source: public/private facilities	No
Timely	Real-time data reporting	No
	High frequency of data refresh – daily or higher	Yes
	Lag in data reporting from source data – less than 24 hours	Yes
Accurate	Data reporting is valid across different dimensions	No
	Data reporting is consistent with other sources	No
	Data reporting is valid and consistent over time	No
	General accuracy in descriptions, indicators, values, operations and representation of data	Partial
Open	Data is reported on a single and stable URL	Yes
	Data is legally open (licensed) to use, re-use and distribute	Yes
	Data is technically open – accessible in machine-readable formats (not pdfs and images)	No
	Download / Export functionality	No

Some of the deficiencies in the Covid-19 dashboard are addressed by reports developed and published/circulated by Ministry of National Health Services, Regulation and Coordination which provide more granularity, sub-provincial/regional data, age-groups but still have a long way to go to cover the kind of insights required for an equitable and inclusive response. These reports, despite having a wider range of indicators, are not accessible openly and published in PDF formats that do not comply with usability standards for being machine-readable.

Implications

An agreement on processes and technical aspects of the data eliminates room for confusion and increases reliability on the system and related decisions. Strong internal systems that produce accurate, credible and timely administrative data can result in public interfaces that allow data to be made open without additional time, cost and effort. The significance of open data in promoting transparency, accountability and value creation making government data available to all cannot be emphasized enough. Making data public makes institutional practices and performance transparent thereby incentivizing responsible practices to uphold standards of accountability.

The pandemic has been tailed by an infodemic making it easy to amplify general levels of apprehension, worry and dismay through disinformation and fake news. Given the propagation of such disinformation through social platforms and its viral uptake, one of the most important role the government will have to play is to build and enhance public trust ranging from accepting guidelines and precautions to support for policy decisions and corresponding outcomes. Bearing in mind the pitfalls of the Polio Eradication Initiative (PEI), it is important that the government quashes parallel information dissemination and monopolizes the knowledge and data circulated in the public to ensure its uniformity and credibility.

Trust in processes that shape responses, understanding of local conditions, and national benchmarking with regional and international performance should all support and strengthen the narratives for success. Having transparent and robust systems can help minimize the disruptive potential of disinformation and fake news because primary sources of information can be relied on. Pakistan's success thus far, must be anchored not only in credible data but an open and transparent mapping of how the response was made possible and the drivers of success.

The use, reuse and free distribution of data can open avenues for generation of information and intelligence at a compounded pace which in turn promotes innovation aimed at improving outcomes for all. High-quality open data not only develops a data-first culture, but also enriches the data landscape by bringing together insights from diverse groups that enhances and improves decision-making by pooling expertise and perspectives. While confidentiality, privacy and security do not allow all datasets to be made public, a policy framework that promotes data openness and emphasizes quality can go a long way in transforming governments.

Applications developed by the National Information Technology Board (NITB) like 'Pass Track App' and 'Covid-19 Gov PK' rely on accurate and timely data and information to ensure quality output. In the absence of granular, timely, accurate, open and standardized data, the evaluation of conditions by these applications will be far from reality. Moreover, the dearth of updated and accurate information through applications and other platforms contributes to a lack of trust in data by the public. Applications like 'Pass Track App' that rely on crowd sourcing to gather data will be impacted negatively by a culture of aversion towards and mistrust in data.

Recommendations

1 Platforms, Capabilities & Sophisticated Analytics

- All Covid-19 reporting should be consolidated on a central platform and interface that aggregates data from all sources and covers all parameters and indicators.
- The platform should allow for more sophisticated representation of data and integrate use-cases like interactive mapping, drill-downs, cross-tabulation and alternative visualization of data.
- Setting up of a collaboration platform that serves as a data exchange for private sector, non-profit, community-based organizations and other entities to contribute and add to the data pool for virus tracking, local conditions, research outputs and benefiting from shared expertise around data collection processes and protocols.

2 Pivot to Open Data

- Institute a truly open data regime for Covid-19 by ensuring that all public and private data, associated back-end processes and structures are publicly available in a machine-readable and user-friendly manner – allow for downloads, standard formats, aggregation of all datasets on a single interface and enrich data from multiple sources across the board.
- Improve capacity of the public and private sector by developing and delivering state-of-the-art courses on data systems, analytics, science, visualizations etc. Existing initiatives like DigiSkills, Presidential Initiative on Artificial Intelligence and Kamyab Jawan should be used as platforms to not only attract and train human resources but also incubate the economic potential of a highly trained data economy workforce. The same workforce – within and outside the public sector can work on improving the Covid-19 data regime. Public Private Partnerships (PPPs) can be used to enhance state capacity in designing and delivering these programmes.
- Use Covid-19 response as a pilot programme for a whole-of-government transition to Open Data by designing, understanding and learning from experimentation and experiences around technical capacity, organisational coordination and resistance, federal-provincial coordination and the technology backbone to transform data regimes across the board. Success and findings from the Covid-19 efforts should inform a fast-tracked national plan for Open Data.

Recommendations

3 Enrich Virus Tracking

- Data collection processes from provinces, public and private facilities, forms and instruments, definitions and protocols, timelines and processing stages to be defined and made public for ensuring that understanding, analysis and interpretation of data is consistent and accurate. For example, it has been reported that the Sindh Health Department used demarcations of union councils (UCs) that were no longer valid in the compilation of data on Covid-19 cases.
 - The data should go deeper and more granular by providing a breakdown in terms of age group, gender, region and sub-region (as granular as possible) in order to improve understanding of the virus patterns. Currently, the district level breakdowns are available on the social media handles and websites of provincial health departments and period reports, but there is no structure and standardization of consolidating information in the national data dashboard.
-

4 Beyond Virus Tracking

- A structured programme on commissioning research on different aspects of the pandemic and the response should be initiated whereby non-government entities can get access to source data, open interactions with key agencies and allow for a universal collation of already underway research that is being done at local levels in pockets. Economic, social, psychological, cultural, demographic and a host of other factors will be required to tackle the expected long-drawn multi-cycle pandemic.