

A Data Ecosystem to Defeat COVID-19

Bapon Fakhruddin¹

1. Technical Director, Disaster Risk Reduction and Climate Resilience, Tonkin + Taylor, New Zealand; Co-Chair, Disaster Loss DATA, IRDR/ISC and LODGD/CODATA (Committee on Data of the International Science Council)

The novel coronavirus disease (COVID-19) has created a human crisis globally, which has demanded an array of drastic, immediate responses. The United Nations (UN) Secretary-General has swiftly called for action, "*for the immediate health response required to suppress transmission of the virus to end the pandemic and to tackle the many social and economic dimensions of this crisis*"¹. The pandemic also requires thinking and decision making supported by a data ecosystem that is more complete than currently, and which looks much further into the future than previous short-term approaches.

The COVID-19 outbreak has led to the proliferation of initiatives to facilitate open access to scientific research and databases and encourage research collaboration through digital platforms. However, there are concerns about the quality of data and publications provided in near real-time, leading to potentially poor decision making. These issues include comparability and interpretation of data, notably between countries, insufficient specification of methodology, and political acceptance of invalid results potentially biasing scientific methods. A call for data and research is necessary in relation to the discussion of the transmission of the disease.

When it comes to predicting or early warning of pandemics and other cascading risks, modelling and situational analysis using historical and current data are the baselines. There have been several lessons learned from the past severe acute respiratory syndrome (SARS) and Middle East respiratory syndrome (MERS), which argued for more cross-disciplinary research. In fact, there are a wealth of data that are under-utilized or unutilized at local, regional and global level, that could greatly assist current and future pandemic-wave responses². Big data such as social media data (e.g., from Facebook, WhatsApp, Twitter, etc.) and local data (e.g., from lab test records, mobile phone users, flight records, etc.) can enable modellers to develop scenarios to better understand and predict the spread of disease as well as its cascading impacts.

COVID-19 virus is primarily transmitted between people through respiratory droplets and contact routes, e.g., when a person is in close contact with someone who has respiratory symptoms such as coughing or sneezing. Transmission may also occur through fomites within the environment around the infected person. Thus, transmission of COVID-19 can occur by direct contact with infected person(s) and/or indirect contact with surfaces in the immediate environment of the infected person(s). This transmission mode makes it difficult to track and understand the complexities of the propagation of the virus.

Social distancing and quarantine are the optimal measures to reduce the spread of COVID-19 at an exponential rate. However, compliance rates are variable, and in many cases, complete social distancing is practically impossible and solely relies on voluntary civic participation. This is often true due to both cultural and infrastructural factors. Social distancing actions include those of: a) refraining from going outdoors and avoiding physical contact with others b) keeping in touch with people through social media instead of meeting in person and c) being proactive with personal hygiene by regularly washing hands. This makes developing countries vulnerable both now and, on an ongoing

¹ Shared responsibility, global solidarity: responding to the socio-economic impacts of COVID-19, UN, 2020

² Wave approach includes other cascading hazards or additional natural hazard within pandemic period

basis, to future waves. Hand-washing becomes hard if there is lack of, or inadequate, access to running water. Governments may require people not to go out to work, but if that means that their families will not eat, people will likely go out anyway to get what they need to survive (consider too the situation in cyclone-ravaged countries). COVID-19 has started spreading in the Pacific countries (e.g., in Fiji) and the healthcare systems of these countries are in no position to cope. The combination of COVID-19 and the cyclone season puts extra pressure on essential services and resources. Countries whose health systems suffer will fail to replicate other developed countries' success in slowing out the outbreak.

There will be no standard response to recovery from COVID-19 due, at least in part, to the lack of a properly formed multi-dimensional data ecosystem to support consistent and well-founded decision making. COVID-19 will likely bounce back to the developed countries as these countries enter the recovery and mitigation stages. An unexpected area that needs more understanding is the matter of quarantine and the compliance rates. Thus, in order to have improved policy decisions, there is a critical and urgent need for a future-focused data-driven approach. Countries are already utilizing data from COVID-19-affected countries, including from their neighbouring countries, for a better policy decision to the response. Every sector within the countries should be using tools and techniques consistently to understand their sectoral impact and develop business continuity plans or pandemic response plans.

The use of data becomes most fraught when it moves beyond modelling to the direct tracking of individuals to identify disease transmission trajectory. For example, as the outbreak took off in China in early January 2020, international travel continued as usual. By 31 January, outbreaks were already growing in over 30 cities across 26 countries, most seeded by people that had travelled from Wuhan³. The virus started spreading locally, moving quickly in confined spaces like religious places and restaurants, and infecting people who had not travelled to China — the start of a pandemic. By March, thousands of cases were reported in Italy, Spain, USA, Iran and South Korea. China was no longer the main 'epicentre' of the outbreak (Figure 1). New cases had started to climb dramatically in countries like Italy, USA and Iran. People travelling to those countries subsequently brought cases to their resident countries as far away as in other continents. The virus has now spread to every continent except Antarctica. Using this type of advance information could be utilized to prevent the COVID-19 spread. Effective response to such spread relies on timely intervention, ideally informed by all available sources of data.

³ Daily The New York Times, 26 March 2020 edition

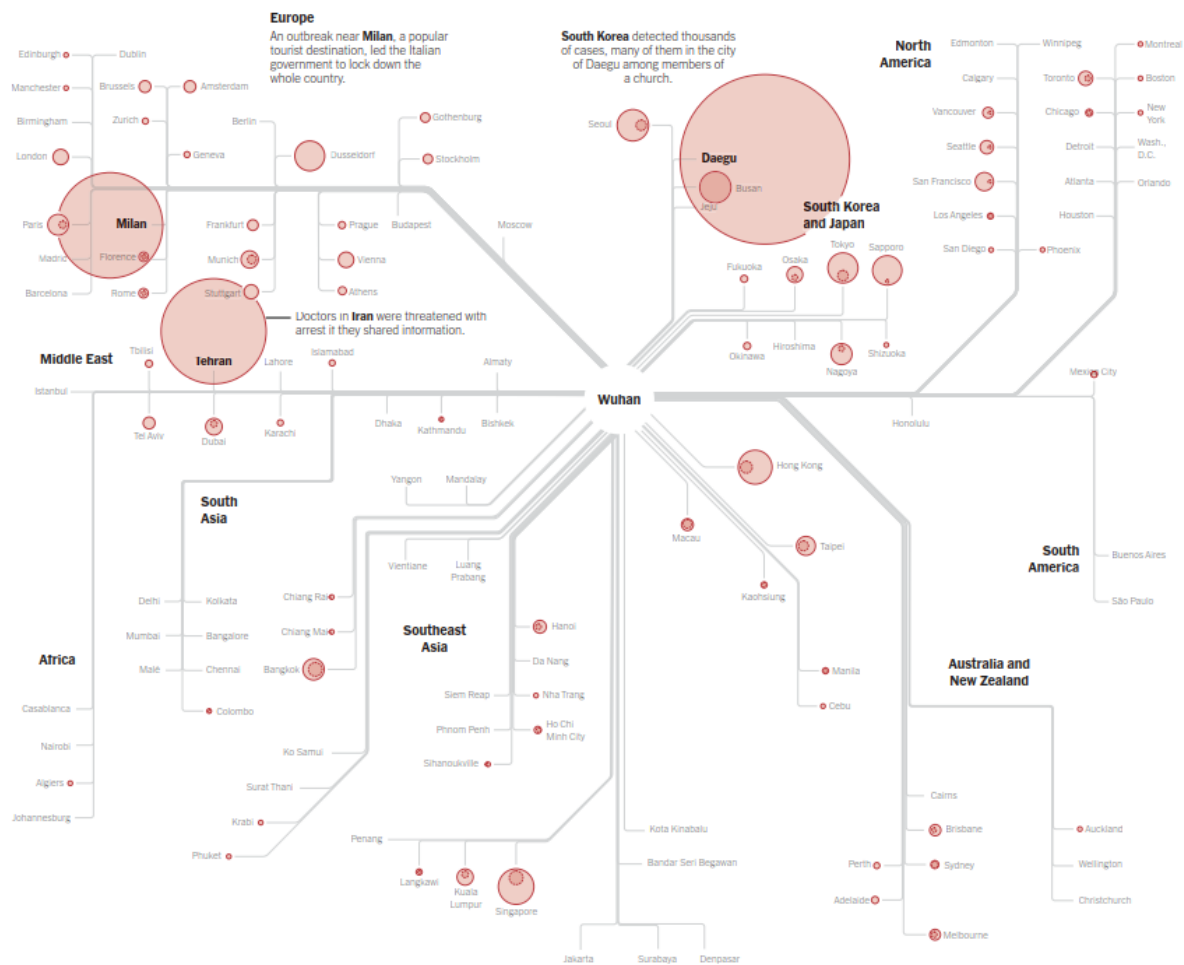


Figure 1: Local outbreaks grew after travel was stopped (The New York Times, 26 March 2020)

Using data for predicting any long-range impact of the pandemic outbreak is complex. It requires a range of inter-connected tasks, multiple disciplines and experts to work together to develop a holistic response and recovery plan. The data governance also requires setting up for COVID-19 response and recovery nationally. Cross-domain research, maximising the utility of data while also ensuring controlled and proportionate access to data is key to understanding, mitigating and responding to the outbreak and preparing for future events. For example, quantitative and qualitative data that are used to understand human behaviour, movement and interaction can be utilized to help predict how and where COVID-19 is going to spread. Emerging technologies are becoming increasingly important in the fight against the disease and attempt to stop COVID-19.

Enhanced surveillance and contact tracing is can be seen as necessary to minimise widespread transmissions within communities. These steps and the technologies that support them also present risks. Governments around the world have implemented a range of digital tracking, physical surveillance and censorship measures (i.e. Governments across Asia have implemented COVID-19-related censorship more than any other region, while European countries have introduced the most digital tracking measures⁴). In March, 20 new digital tracking measures were implemented in continents across the world including Europe, Asia and South America.⁵ Such tracking measures varied from targeted contact tracing apps to large-scale acquisition of aggregated and anonymised location data. Advanced tracing technology options (such as automated rapid mass tracing including the use

⁴ <https://www.top10vpn.com/news/surveillance/covid-19-digital-rights-tracker/>

⁵ Top10VPN: COVID-19 Digital Rights Tracker (<https://www.gpsworld.com/19-countries-track-mobile-locations-to-fight-covid-19>)

of GPS location data and Bluetooth data) will allow officials to accurately trace and monitor their populations to inform decision making and implement measures to slow the spread of the virus. An example of this comes from a study in Singapore where scientists analysed data from Bluetooth using the TraceTogether app on mobile phones to see how many days it took on average to get in contact⁶. Results revealed that the authorities were able to contact people within 3-4 days. Another recent example of South Korean digital systems easing the load on human contact tracers indicates how it is possible to define contact tracing quickly (i.e. within 10 minutes⁷) to reduce the spread of the virus. Other current examples of digital tracking initiatives include:

- Geolocation tracker through the use of location map
- Cell site location information and call detail record data acquisition methods
- GPS satellite and Bluetooth data
- Development of location apps.

However, it remains essential to maintain the limitations on how such data can be accessed and used during this pandemic period. This includes ensuring the security of personal information and privacy violations, promoting scrutiny and ensuring that these measures do not continue longer than necessary.

A wide range of approaches could be applied to understand transmission, outbreak assessment, risk communication, cascading impacts assessment on essential and other services. The network-based modelling of System of Systems (SOS), mobile technology, frequentist statistics and maximum-likelihood estimation, interactive data visualization, geostatistics, graph theory, Bayesian statistics, mathematical modelling, evidence synthesis approaches and complex thinking frameworks for systems interactions on COVID-19 impacts could be utilized. An example of tools and technologies that could be utilized to act decisively and early to prevent the further spread or quickly suppress the transmission of COVID-19, strengthen the resilience of health systems and save lives and urgent support to developing countries with businesses and corporations are shown in Figure 2. There are also WHO guidance on '[Health Emergency and Disaster Risk Management](#)⁸', [UNDRR supported 'Public Health Scorecard Addendum](#)⁹', and other guidelines (e.g. [WHO practical considerations and recommendations for religious leaders and faith-based communities in the context of COVID-19](#)¹⁰) that could enhance pandemic response plan. It needs to be ensured that any such use is proportionate, specific and protected and does not increase civil liberties' risk. It is essential therefore to examine in detail the challenge of maximising data use in emergency situations, while ensuring it is task-limited, proportionate and respectful of necessary protections and limitations. This is a complex task and the COVID-19 will provide us with important test cases. It is also important that data is interpreted accurately. Otherwise, misinterpretations could lead each sector down to incorrect paths.

⁶ TraceTogether- <https://www.healthhub.sg/apps/38/tracetgether-app>

⁷ <http://www.koreaherald.com/view.php?ud=20200326000987>

⁸ WHO Health Emergency and Disaster Risk Management - <https://www.who.int/hac/techguidance/preparedness/health-emergency-and-disaster-risk-management-framework-eng.pdf?ua=1>

⁹UNDRR Public Health Scorecard Addendum

https://www.unisdr.org/campaign/resilientcities/assets/toolkit/documents/Disaster%20Resilience%20Scorecard_Public%20Health%20Addendum%20Ver1%20Final_July%202018.pdf

¹⁰ <https://www.who.int/publications-detail/practical-considerations-and-recommendations-for-religious-leaders-and-faith-based-communities-in-the-context-of-covid-19?fbclid=IwAR0GtTGRHqvgrDd7KiRLH6Sza8bJ7aQP40cSsyFju3w-HFRQBiy7Yic9eU8>

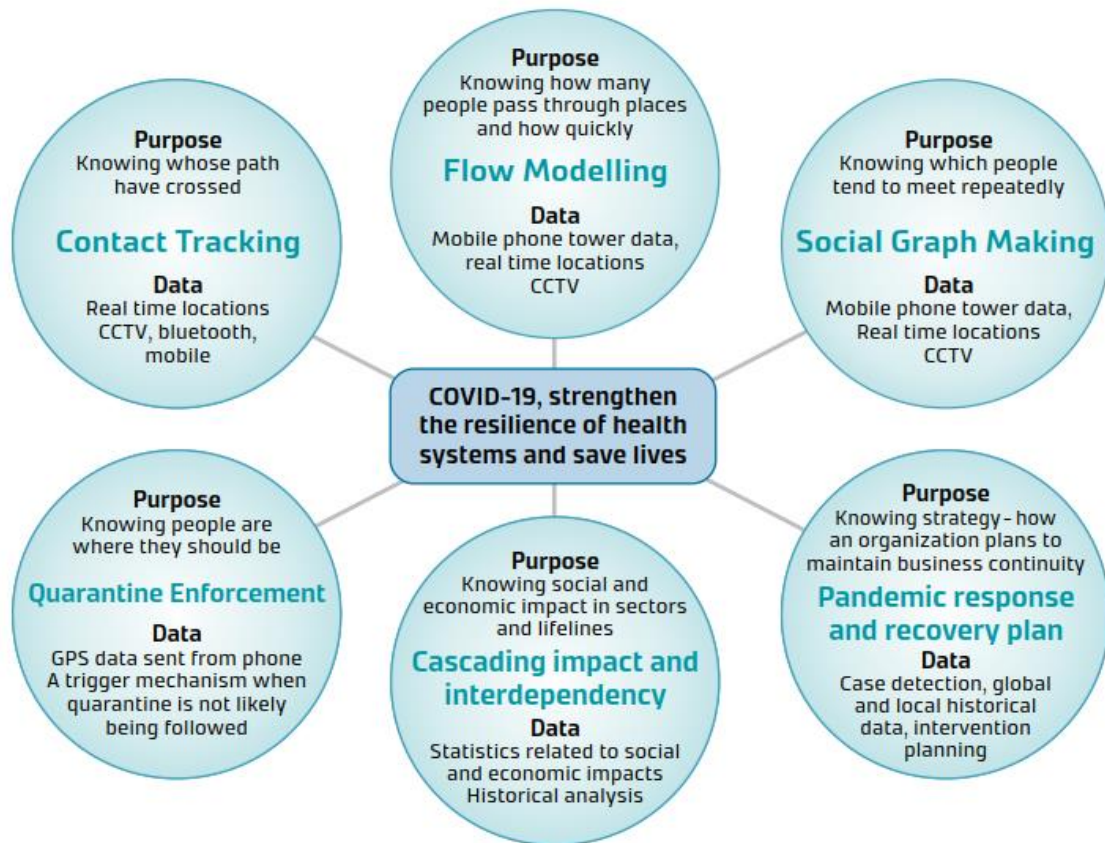


Figure 2: Tools to strengthen resilience for COVID-19

Many countries are still learning how to make use of data for their decision making in this critical time. The COVID-19 pandemic will provide important lessons on the need for cross-domain research and on how, in such emergencies, to balance the use of technological opportunities and data to counter pandemics against fundamental protections. Lessons learned from this devastating outbreak may provide significant improvements in preparation to fight potential pandemic in the future. The [CODATA Task Group on FAIR Data for Disaster Risk Research is preparing a series of policy briefs on a number of DRR issues](#). In collaboration with other experts and actors in the space this will consider in further detail policy issues in relation to data to inform pandemic response.

Acknowledgement: The author acknowledge editorial support and valuable comments received from Simon Hodson, Executive Director, CODATA.