COMBATING COVID-19 WITH DATA: WHAT ROLE FOR NATIONAL STATISTICAL SYSTEMS?

POLICY BRIEF, 14 APRIL 2020

PARIS21
COVID-19 RESPONSE

ABSTRACT

The COVID-19 pandemic has brought data to the centre of policymaking and public attention. A diverse ecosystem of data producers, both private and public, report rates of infection, fatality and recovery on a daily basis. However, a proliferation of data, which is at times contradictory, can also lead to confusion and mistrust among data users.

Meanwhile, policymakers, development partners and citizens need to take quick, informed actions to design interventions that reach the most vulnerable and leave no one behind. As countries comply with lockdowns and other containment measures, national statistical systems (NSSs) face a dual effect of growing data demand and constrained supply. This in turn may squeeze NSSs beyond their institutional capacity.

At the same time, alternative data sources such as mobile phone or satellite data are in abundance. These data could potentially complement traditional sources such as censuses, surveys and administrative systems. However, with scant governance frameworks to scale and sustain their use, policy action is not yet based on a convergence of evidence.

This policy brief introduces a conceptual framework that describes the adverse effects of the crisis on NSSs in developing countries. Moreover, it suggests short and medium-term actions to mitigate the negative effects by:

1. Focusing data production on priority economic, social and demographic data.
2. Communicating proactively with citizens, academia, private sector and policy makers.
3. Positioning the NSO as advisor and knowledge bank for national governments.

NSSs contribute significantly to robust policy responses in a crisis. The brief thus calls on national statistical offices to assume a central role as coordinators of the NSSs and chart the way toward improved data ecosystem governance for informing policies during and after COVID-19.

This report was prepared by Archita Misra, Julia Schmidt and Lauren Harrison. Comments and factual corrections welcome (please send to archita.misra@oecd.org , julia.schmidt@oecd.org , lauren.harrison@oecd.org).

The authors are thankful for inputs from Liliana Sudocholska, Yu Tian and Salma Dahir.

1. WHY QUALITY DATA IS INTEGRAL TO ADDRESS THE COVID-19 CRISIS

In early 2020, the world began to experience a massive outbreak of COVID-19 – a novel, highly contagious coronavirus with strong respiratory symptoms. By the end of January, the World Health Organization declared the COVID-19 outbreak a global health emergency, upgrading it to a global pandemic on 11 March. With every passing day, the crisis has revealed and amplified various fault lines of our socio-economic systems – from under-capacitated health services, brittle financial markets to deficient social policy and debilitated politics and governance.

1.1 A pandemic in the age of data revolution - the role of trusted, reliable statistics

The COVID-19 pandemic has emerged at a time of unprecedented digitalisation and data abundance. While official statistical agencies once served as the primary source of data, today we witness the proliferation of data actors, both private and public. These actors are part of a modern data ecosystem that produces and disseminates valuable information to monitor the reach and impact of the pandemic.[1]

However, the proliferation of these actors, including data aggregators and analysts, is accompanied by an increasing uncertainty and dispute over these statistics.[2] Data quality and reliability have been a persistent problem since the pandemic began, making it nearly impossible to determine the true scale of the outbreak in affected countries. John Ioannidis, an Epidemiologist at Stanford University, has characterised the outbreak as a "once-in-a-century evidence fiasco", adding, "at a time when everyone needs better information, [...] we lack reliable evidence on how many people have been infected with SARS-CoV-2 or who continue to become infected. Better information is needed to guide decisions and actions of monumental significance and to monitor their impact" (Ioannidis, 2020[1]).

1.2 Pathways to scale up alternative approaches – the role of data governance

There is heightened global interest in harnessing alternate data sources held with private companies to track population movement and obtain a high-resolution picture of the outbreak. Early examples from different regions[3] show the combining of official data with alternative sources, such as satellite and mobile phone data, to identify high-risk populations and support the modelling of the geographical spread of the epidemic (Zaimova, 2020[2]).

However, there is still scarce strategic and systematic usage of these alternative data sources such as mobile phone data in the COVID-19 response (Oliver et al., 2020[3]). This mirrors past trends, notably in the Ebola outbreak of 2014-2016, where several pilot or one-off activities to leverage big data were also initiated, although with limited effects on the status-quo and standardised procedures in the long term. [ibid, (The Economist, 2014[4]), (McDonald, 2016[5])]. Concerns around the applications of these technologies remain, including ethical issues of data sensitivity and anonymity, representation of vulnerable groups, and future considerations around civil rights and privacy ((Oliver et al., 2020[3]); (Gray, 2020[6]); (McDonald, 2020[7]); (Stein, 2020[8])). These concerns are heightened in the context of developing countries where lack of institutional mechanisms for coordination, data-sharing, regulatory frameworks, and privacy laws present obstacles to sustainably scaling up these approaches.

1 See Table 1 in Annex for a select list of popular COVID-19 data sources.

2 Some indicative examples: National Post, Fortune, Time, The Economist, The Guardian, Sky News, The Local.

3 See Nigeria, Ghana

1.3 Beyond immediate data – the role of national statistical systems

The data requirements for COVID-19 go far beyond the figures on cases and deaths. Policy makers and development partners supporting the most vulnerable people in least developed societies require foundational data that is timely, accurate and disaggregated to design interventions that leave no one behind. For LMICs, this poses a unique set of challenges. Limited statistical capacity and existing structural data deficits, including weak administrative data systems such as Health Management Information Systems and Civil Registration and Vital Statistics (CRVS), further reinforce the need for well-governed data systems to close critical data gaps and strengthen flows of quality data and statistics to inform pandemic response, recovery and future resilience (PARIS21, 2019[9]). The crisis thus holds far-reaching effects for society and the economy that will be felt long after the case counts decline.

In addition, the crisis has brought out the need to close structural development data gaps that form the bedrock of any inclusive policy intervention, on the one hand, while underscoring on the other hand the need for new approaches to responsible data collection and use in modern data ecosystems fraught with noise and heterogeneity. This has had the effect of emboldening the call for data governance frameworks and sustainable data stewardship. Hence, the crisis presents an opportunity for independent official statistical agencies to chart the way for a well-governed data ecosystem with quality and credible statistics and inform policies for response and recovery.

2. UNDERSTANDING THE IMPACT OF COVID-19 ON NATIONAL STATISTICAL SYSTEMS

In low-and middle-income countries, policy makers face dire challenges in deciding on aligned response and recovery operations – not least due to severe constraints in institutional capacity. Under these conditions, the crisis heavily affects NSSs in the short-and medium term, disrupting the supply of quality data.

2.1 Government responses in LMICs

Early indications show COVID-19 will affect developing and least developed countries in unique ways. While some of these countries have younger populations (mitigating some COVID-19 risk factors) and more time to prepare, the effects of the crisis may last over longer time horizons than for OECD counterparts (Lempinen, 2020[10]). In least developed countries and fragile contexts with larger informal economies, poorer infrastructure and limited social safety nets, health and fiscal capacity, options for responding to the crisis are more constrained and come with starker trade-offs (The Economist, 2020[11]).

Since the outbreak of the crisis, governments in LMICs have employed a wide range of containment measures to encourage 'social distancing' or to augment public health provision. However, the timeliness and stringency of these measures vary significantly across governments. The Colombian government, for example, began public information campaigns in January and pre-emptively closed businesses on 25 February, before registering a single infected case. Additional measures followed in March, including public transport closures, making Colombia one of the most stringent countries according to the Oxford COVID-19 Government Response Tracker. Rwanda, on the other hand, started only on 11 March with a public information campaign. However, work places and public transport closed on 16 March, later upgrading to a full lockdown by the end of the month with only 81 infected cases (Hale Thomas et al., 2020[12]).

As COVID-19 continues to spread, these containment measures result in adversarial effects on NSSs that now find themselves facing unprecedented challenges.

2.2 Short-term effects on national statistical systems

As more developing countries adopt containment measures, the COVID-19 crisis has created a dual shock affecting data demand and supply that precipitates disruptions in the data value chain for official statistics (see Figure 1).

On the demand side, governments call for timely, ad hoc data to inform containment and mitigation efforts. Ministries of Health are asked to monitor and release information on COVID-19 infections and deaths to track the pandemic. Information on vulnerable populations, businesses and the labour market are required to help policy makers support citizens in this time of crisis. In addition, information on public transport, events and internal mobility is needed to implement efficient containment measures. However, in many countries, robust and complete administrative records of vital events and domestic migration are not available to inform valid policy responses at the national and sub-national levels. Moreover, available socio-economic and price data may rapidly fall out of date due to evolving conditions on the ground.

While demand is increasing, the operational space for NSSs is becoming more constrained, limiting the supply of official statistics. The full or partial closure of workplaces forces national statistical offices (NSOs) and other NSS institutions to shift to teleworking arrangements, often without sufficient IT solutions in place to sustain ongoing activities. Furthermore, many NSOs rely on technical assistance and capacity development programmes supported by development partners whose staff are repatriated in times of crisis. In response, most NSOs have bootstrapped their data production and postponed field-based data collection. While some institutions are looking for options to conduct remote data collection via phone, NSOs in Niger and Rwanda have stopped all field missions until further notice, indefinitely delaying survey and census production. Paraguay's NSO has assured the continuity of the ongoing household survey by limiting its scope, focusing only on employment dynamics and reducing the survey questionnaire to 40% of its original length.

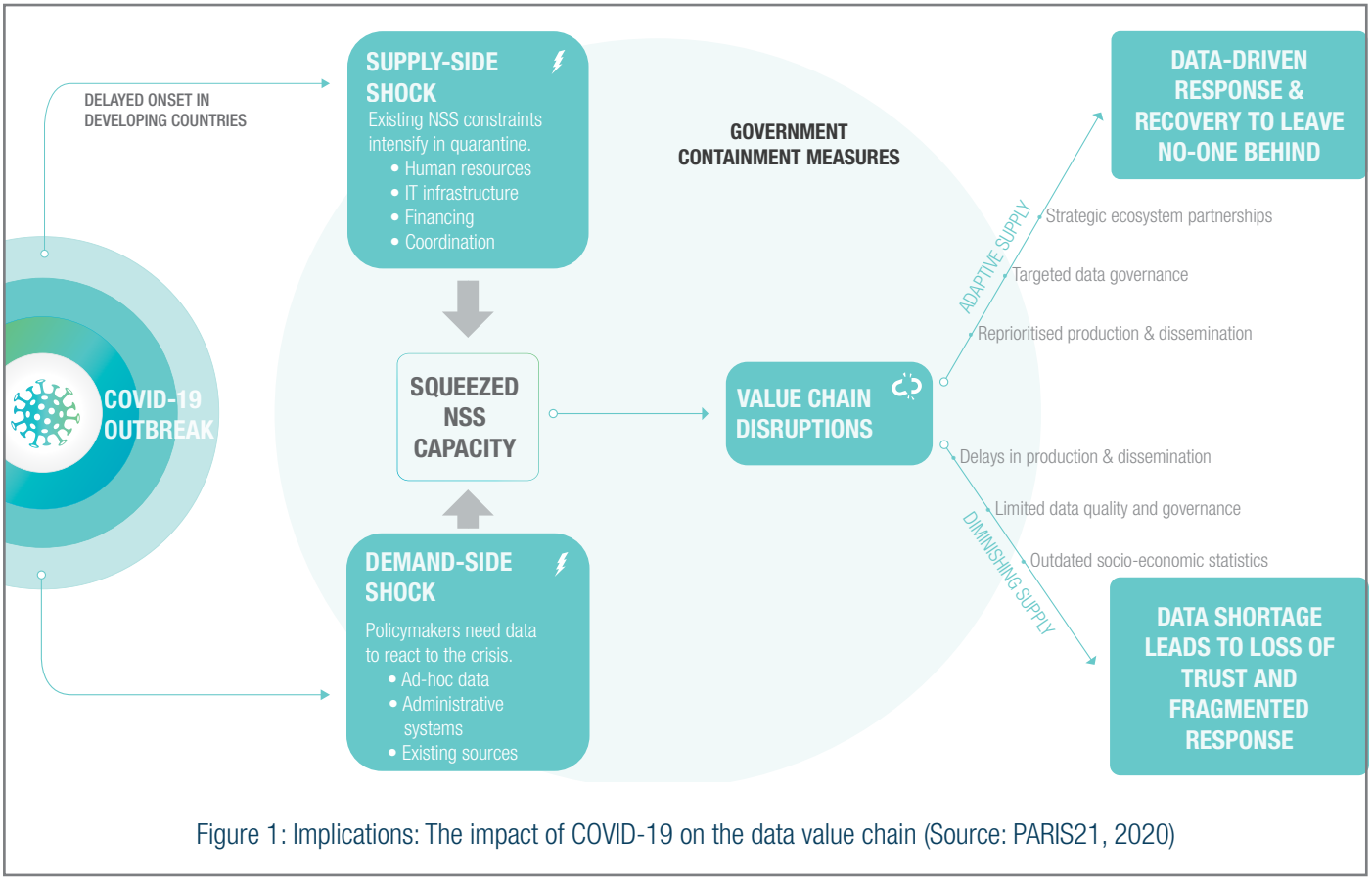
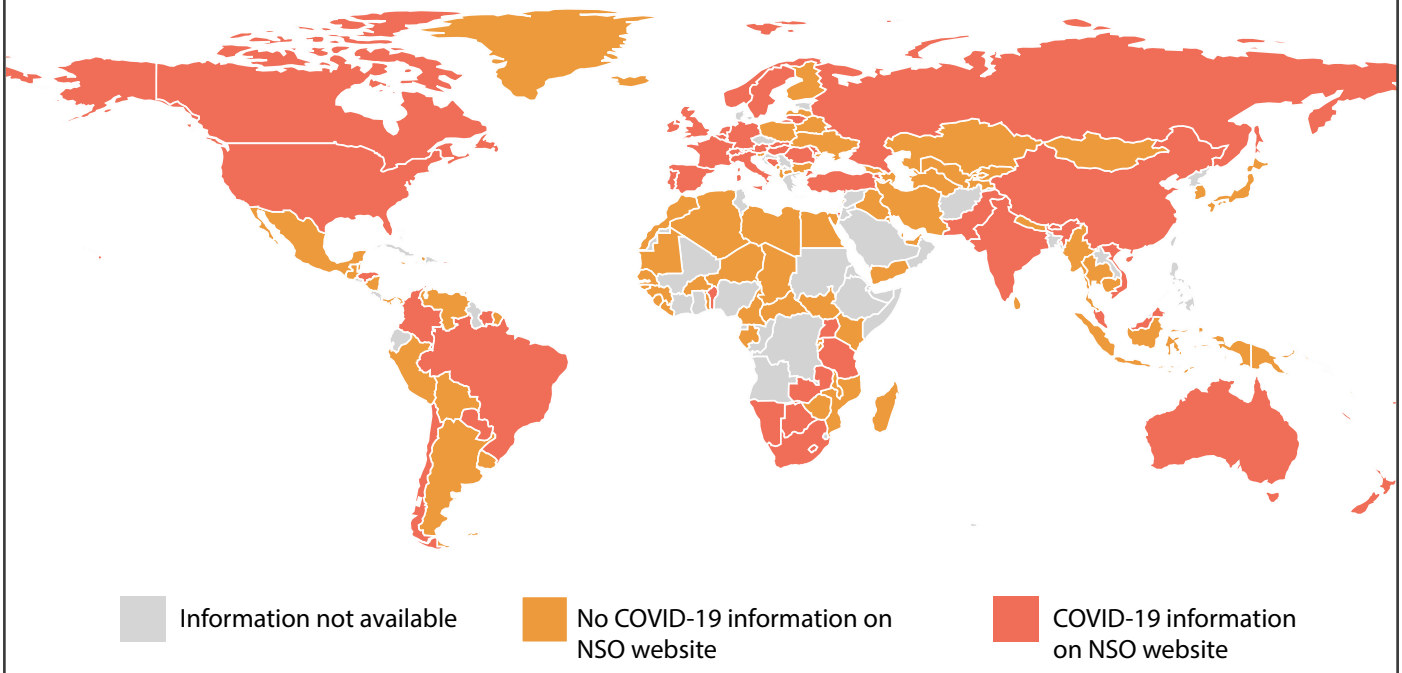


Figure 1: Implications: The impact of COVID-19 on the data value chain (Source: PARIS21, 2020)

Box: Communicating in times of crisis – National statistical offices fall behind in providing information

Demand and supply shocks to NSSs can also be reflected in the communication patterns of NSOs in developing countries. Only 49 out of 151 NSO websites analysed by PARIS21 currently publish any information related to the pandemic in their homepages. Even as governments increasingly adopt more stringent containment measures, the majority of NSOs in Africa, Latin America and Asia do not provide any COVID-specific information on available data sources or operational effects on data production.



Note: This map is for illustrative purposes and is without prejudice to the status of or sovereignty over any territory covered by this map. The data is web-scraped from NSO websites as of 03.04.2020.

Figure 2: NSOs communicate heterogeneously on COVID-19 (Source: PARIS21, 2020)

2.3 Medium- and long-term effects on the statistical value chain

Statistical capacity plays a decisive role in the way NSSs in developing countries are coping with containment measures. The short-term dual effect of rising demand and deteriorating supply squeezes NSSs beyond their institutional limits. Whether they are able to mitigate those negative effects in the medium-term does not only rely on financial, technical and human resources, but more so on strong leadership, a flexible organisational culture and the ability to coordinate and collaborate effectively with partners. Earlier PARIS21 assessments show that NSSs struggle to develop these capabilities even when operating in less severe circumstances (PARIS21, 2019[9]).

Depending on their respective statistical capacity, NSSs might either adapt or diminish their data supply chains, as visualised in Figure 1[4]. Under an adaptive model, NSSs will reprioritise data production and dissemination based on emerging data demands, balancing urgency with feasibility. Moreover, NSSs will coordinate with other data producers to manage data responsibly and enter into strategic partnerships with other ecosystem actors. In doing so, NSSs can effectively contribute to response and recovery operations by delivering data that enables decision-makers to identify and prioritise those at risk of being left behind.

4 It must be noted that these medium and long-term scenarios rely on extrapolations from the observed shocks and close consultations with partner countries; and will require further investigation as the crisis spreads.

In a diminishing model – likely to occur in low-capacitated NSSs – delays in production and dissemination gradually lead to outdated socio-economic statistics, hampering effective response and recovery operations over time. In Honduras, for example, policy makers now rely on outdated census data to identify populations at risk. Under these conditions, users turn to alternative data sources that have been emerging during the pandemic. In the absence of effective coordination and governance, these data are noisy, unfettered and lack quality assurances. Fake news, threats to privacy and state censorship might flourish (Mchangama and McLaughlin, 2020[13]). Ultimately, the second scenario could lead to a fragmented data ecosystem, leaving policy makers with unverified sources to base their policy responses on.

In a crisis more than in any other time, NSSs need to reach beyond their traditional mandate as a data producer, and engage in data stewardship across the data ecosystem (GOVLAB, 2020[14]). Data stewardship will become particularly pivotal in the long-term, since the observed shocks might spill over to the whole data ecosystem and cause chronic disruptions along the data value chain. Delays in production and dissemination of economic statistics might result in poorly fitted policy responses to economic shocks in the aftermath of the crisis, causing collateral damage to the economy with potential backlashes on government funding. In this context, private data producers, civil society organisations and academia might step in and provide complementary data supply to policy makers. In the future, NSOs should thus re-focus their capacity development efforts on coordination, governance and quality management.

3. WAYS FORWARD IN THE CRISIS – RECOMMENDATIONS

National statistical systems are the backbone of an informed and robust policy response in a crisis. Many NSSs face similar challenges in catering to heightened policy demand under such circumstances. This section thus presents a set of short- and medium-term recommendations and mitigation strategies employed in PARIS21 partner countries.

Short-term recommendations: Early action to reconcile supply and demand pressures will be key to manage disruptions to the data value chain.

1. Mitigate delays in data production

    a. Scale up teleworking arrangements: Wherever possible, NSS institutions should work with policymakers and development partners to improve operational conditions for remote work. Without these capabilities, risks of critical data shortages will increase over time. The Colombia National Administrative Department of Statistics (DANE) took early steps to ensure staff connectivity such as VPN access, free antivirus licenses and increased data-processing capacities.

    b. Bootstrap data collection instruments: NSS institutions can reduce survey burden by limiting sample sizes or scope of questionnaires (as observed in Paraguay). Any of these measures should also take into account current and forward-looking data requirements for pandemic response.

    c. Accelerate innovation and digitalisation: As opportunities for traditional, field-based approaches to data collection wane, NSOs and other NSS institutions should prioritise transitions to new technologies for official statistics. The Senegal National Agency of Statistics and Demography (ANSD) is pursuing options to replace field operations with phone and web-based survey production. The Lesotho Bureau of Statistics (BOS) solicited support from the World Bank to develop alternative approaches to annual survey production using big data.

2. Expand dissemination and communication

    a. Package existing data sources for decision-makers: While options to collect new data are constrained, policy makers and partners can use existing data sources to develop evidence-based approaches to pandemic response and identify populations at greatest risk. The Dominican Republic Office of National Statistics (ONE) responded to data demands by establishing a central data repository of all information related to COVID-19 in their country. The Maldives National Bureau of Statistics (NBS) is collaborating with the United Nations Development Program to undertake a rapid livelihood assessment of populations most at risk in the context of the epidemic.

    b. Engage in the global discussion and communicate to citizens: Even under limited operational conditions, NSOs and NSSs serve an important function in communicating trusted information to the public. Under this mandate, press releases and regular website updates related to COVID-19 are essential to communicate anticipated effects on data production and dissemination and identify opportunities for engagement. The Niger National Institute of Statistics (INS) produces a daily note on the evolving situation globally. The note is regularly submitted to the Office of the President, and could serve as a framework for regular communication to the public.

3. Strengthen mechanisms for collaboration and coordination

    a. Engage with data users and other producers: In addition to acting as a central data repository, NSOs and NSSs are positioned to offer advisory services to policy makers and other ecosystem stakeholders – as part of a crisis committee, for example. To provide actionable insights, NSOs need to understand information requirements associated with COVID-19 response and recovery. Colombia (DANE) created an administrative data committee with members from the National Planning Department and the Ministry of ICT to understand the data needs of their policy makers.

Medium-term recommendations: Going forward, the difference between adaptive and diminishing data supply chains will be shaped largely through effective governance and ecosystems engagement. Actions taken by NSOs and their partners should be directed toward developing coordination, quality management and data governance capacities.

1. Reprioritise data production and dissemination: NSSs need to prioritise the production and dissemination of foundational data such as census, household surveys, CVRS data and administrative data. Most importantly, the data needs to be disaggregated by gender, income, employment status, age to enable policy makers to take meaningful action to leave no one behind.

2. Strengthen data quality and governance: NSOs have to step up beyond their role as a data producer and engage in quality management, data governance and coordination. Efforts to develop capacities to support these efforts should follow a holistic approach taking into account the whole data ecosystem as suggested by the CD4.0 framework .

3. Develop strategic partnerships: In a modern data ecosystem, new data producers such as the private sector or civil society, or knowledge banks, such as academic institutes provide an opportunity to extend the coverage and scope of quality data production for the public good. NSSs thus need to engage in sustainable and complementary partnerships that offer a path toward coordinated flows of high quality data to inform response and recovery.

ANNEX: WHERE IS THE DATA ON COVID-19 COMING FROM?

POPULAR DATA SOURCES/ AGGREGATORS	LINK/INFORMATION	UNDERLYING DATA SOURCES
European Center for Disease Prevention and Control (ECDC)	https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases The European CDC publishes daily statistics on the COVID-19 pandemic. It collects aggregates and harmonizes data from countries around the world. The lag between nationally available data and the ECDC data is not very long as the ECDC publishes new data daily.	https://www.ecdc.europa.eu/en/covid-19/data-collection Key sources include: • Regular updates from EU/EEA countries through the Early Warning and Response System (EWRS), The European Surveillance System (TESSy), the World Health Organization (WHO) and email exchanges with other international stakeholders. • Screening of up to 500 sources every day to collect COVID-19 figures from 196 countries, including o websites of ministries of health (43% of the total number of sources) o websites of public health institutes (9%) o websites from other national authorities (ministries of social services and welfare, governments, prime minister cabinets, cabinets of ministries o websites on health statistics and official response teams) (6%), o WHO websites and WHO situation reports (2%) o official dashboards and interactive maps from national and international institutions (10%) • Screening of social media accounts maintained by national authorities, for example Twitter, Facebook, YouTube or Telegram accounts run by ministries of health (28%) and other official sources (e.g. official media outlets) (2%).

Johns Hopkins University's Center for Systems Science and Engineering (CSSE)	The site tallies the total cases by country or hotspot, the number of deaths, and more optimistically, the number of people who have recovered. https://coronavirus.jhu.edu/map.html	• World Health Organization (WHO): https://www.who.int/ • DXY.cn. Pneumonia. 2020. http://3g.dxy.cn/newh5/view/pneumonia • BNO News: https://bnonews.com/index.php/2020/02/the-latest-coronavirus-cases/ • National Health Commission of the People's Republic of China (NHC) http://www.nhc.gov.cn/xcs/yqtb/list_gzbd.shtml • China CDC (CCDC): http://weekly.chinacdc.cn/news/TrackingtheEpidemic.htm • Hong Kong Department of Health: https://www.chp.gov.hk/en/features/102465.html • Macau Government: https://www.ssm.gov.mo/portal/ • Taiwan CDC: https://sites.google.com/cdc.gov.tw/2019ncov/taiwan?authuser=0 • US CDC: https://www.cdc.gov/coronavirus/2019-ncov/index.html • Government of Canada: https://www.canada.ca/en/public-health/services/diseases/coronavirus.html • Australia Government Department of Health: https://www.health.gov.au/news/coronavirus-update-at-a-glance • European Centre for Disease Prevention and Control (ECDC): https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases • Ministry of Health Singapore (MOH): https://www.moh.gov.sg/covid-19 • Italy Ministry of Health: http://www.salute.gov.it/nuovocoronavirus • 1Point3Arces: https://coronavirus.1point3acres.com/en • WorldoMeters: https://www.worldometers.info/coronavirus/
Our World in Data (Global Change Data Lab, and University of Oxford)	https://ourworldindata.org/coronavirus	• European Center for Disease Prevention and Control (ECDC). • Until March 18 they relied on the World Health Organization (WHO) as their source, however no longer do, citing reasons of errors in the daily Situation Reports, loss of comparability and unreliable time-series information. https://ourworldindata.org/coronavirus-source-data

World Health Organization (WHO)	https://www.who.int/	https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/ WHO COVID-19 Situation Reports present official counts of confirmed COVID-19 cases, thus differences between WHO reports and other sources of COVID-19 data using different inclusion criteria and different data cutoff times are to be expected. Latest situation report, as of April 6: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200405-sitrep-76-covid-19.pdf?sfvrsn=6ecf0977_2
Worldometers	https://www.worldometers.info/coronavirus/ A general reference site that shows country-by-country breakdowns of aggregates case numbers.	It primarily draws data from government agencies and media reports, and in turn serves as one of the sources for Johns Hopkins' map. https://www.worldometers.info/coronavirus/#countries

REFERENCES

1 Ioannidis, J. (2020), A fiasco in the making? As the coronavirus pandemic takes hold, we are making decisions without reliable data

2 Zaimova, R. (2020), How data can help fight a health crisis like the coronavirus, World Economic Forum

3 Oliver, N. et al. (2020), Mobile phone data and COVID-19: Missing an opportunity

4 The Economist (2014), Ebola and big data - Waiting on hold

5 McDonald, S. (2016), "Ebola: A Big Data Disaster. Privacy, Property, and the Law of Disaster", Centre for Internet and Society.

6 Gray, S. (2020), A Closer Look at Location Data: Privacy and Pandemics

7 McDonald, S. (2020), No, we don't need an app for this

8 Stein, S. (2020), How to restore data privacy after the coronavirus pandemic, World Economic Forum

9 PARIS21 (2019), Statistical Capacity Development Outlook

10 Lempinen, E. (2020), Africa faces grave risks as COVID-19 emerges, says Berkeley economist

11 The Economist (2020), COVID-19 presents stark choices between life, death and the economy

12 Hale Thomas, A. et al. (2020), Variation in Government Responses to COVID-19

13 GOVLAB (2020), Wanted: Data stewards. (Re-)defining the roles and responsibilities of data stewards for an age of data collaboration

14 Mchangama, J. and S. McLaughlin (2020), Coronavirus Has Started a Censorship Pandemic

www.paris21.org
@ContactPARIS21
info@paris21.og

PARIS21
COVID-19 RESPONSE