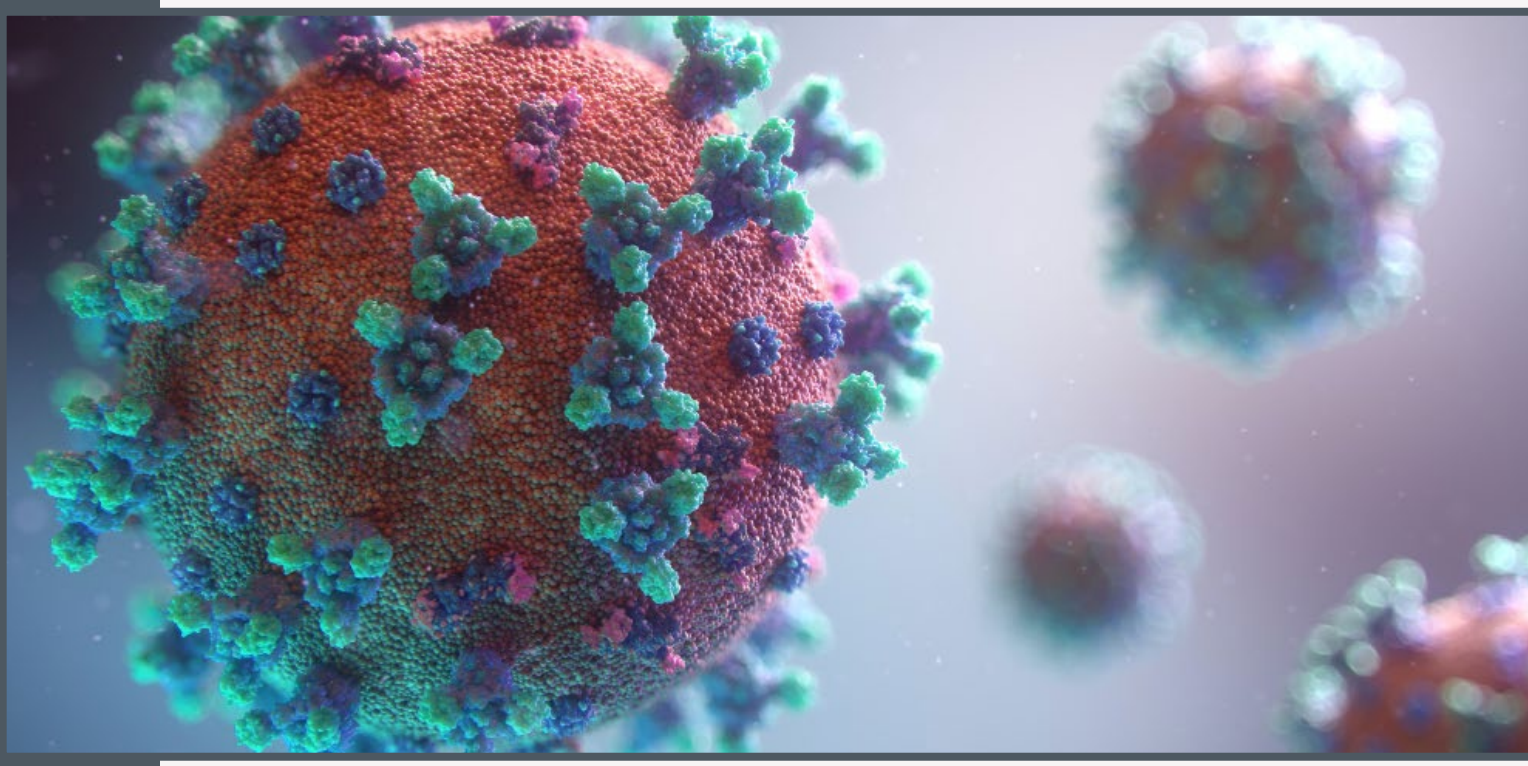
The #Data4COVID19 Review

Assessing the Use of Non-Traditional Data During a Pandemic Crisis



Hannah Chafetz, Andrew J. Zahuranec, Sara Marcucci, Behruz Davletov, and Stefaan Verhulst

GOVLAB

KNIGHT FOUNDATION

Authors

Hannah Chafetz, Andrew J. Zahuranec, Sara Marcucci, Behruz Davletov, and Stefaan Verhulst

Acknowledgements

This work would not have been possible without the support and guidance of many individuals throughout the research process, in particular Lilian Coral of the Knight Foundation, who was behind the initial idea of conducting this study and provided the resources to do so. Salwa Mansuri, Andrew Young, Lisa Talia Moretti, Marine Ragnet and all at The GovLab were instrumental in contributing to earlier versions and provided important research and editorial assistance. Finally, the authors are extremely grateful to the following individuals for their time and dedication to this effort.

International Advisory Group

- **Dr. Ana Beduschi**: Professor of Law and Personal Chair at the University of Exeter, United Kingdom
- **Dr. Caroline Buckee:** Professor of Epidemiology, and the Associate Director of the Center for Communicable Disease Dynamics at the Harvard T.H. Chan School of Public Health
- **Dr. Ciro Cattuto:** Principal Scientist at ISI Foundation, Associate Professor in the Computer Science Department of the University of Torino
- **Dr. Leo Ferres:** Associate Research Professor, Data Science Institute, Facultad de Ingeniería, Universidad del Desarrollo
- **Sir Geoff Mulgan:** Professor of Collective Intelligence, Public Policy and Social Innovation at University College London
- **Dr. Chinwe Lucia Ochu**: Director, Prevention Programmes & Knowledge Management and Head of Research at Nigeria Centre for Disease Control
- **Dr. Nuria Oliver:** Scientific Director and Co-Founder of ELLIS Alicante Foundation, Chief Data Scientist at the Data-Pop Alliance, Chief Scientific Advisor at the Vodafone Institute, Board Member at ELLIS, Fellow at the Spanish Royal Academy of Engineering, Member of the Industrial Advisory Board at King's College

Peer Reviewers

- **Jaimie Boyd:** National Digital Government Leader at Deloitte Canada
- **Dr. Richard Benjamins:** Chief AI and Data Strategist at Telefonica, Co-founder and Vice President of OdiseIA, Member Board Of Directors of CDP, Deputy Board Member of IndesIA
- **Malcolm Fraser:** Executive Chair at the i4 Institute, Chartered Fellow at The Chartered Governance Institute
- **Lauren King:** Director of Training and Evaluation at ImpactTulsa, Board Member of Oklahoma Access to Justice Foundation
- **Brennan Lake:** VP of Social Impact at Spectus
- **Dr. Oliver Marsh:** Founder at The Data Skills Consultancy
- **Sergio Andrés Morales-Barreto**: Lawyer, Master in Constitutional Law, Part-time lecturer at the Department of Constitutional and Legal Theory at University of La Sabana (Colombia, Member of the Editorial Team of the Dikaion Law Journal
- **Dr. Daniela Paolotti:** Research Leader at ISI Foundation - Digital Epidemiology Laboratory and Data Science for Social Impact
- **Dr. Andrew Schroeder:** Vice President of Research and Analysis at Direct Relief, Co-founder of WeRobotics.org, Co-Director of CrisisReady
- **Eric Sutherland:** Executive Director, pan-Canadian Health Data Strategy at the Public Health Agency of Canada

Interviewees

- **Dr. Melissa Haendel:** Co-Lead of The National COVID Cohort Collaborative (The N3C)
- **Aparna Keshaviah:** Principal Researcher at Mathematica
- Ben Purple: Director of GEOLYTIX
- **Camila Salazar:** Lead Data Analyst at the Open Contracting Partnership
- **Dr. Lucy Lu Wang:** Assistant Professor, Information School, University of Washington; Visiting Scientist, Allen Institute for AI (AI2)

Table of Contents

Table of Contents

List of Figures

List of Tables

Photo by Mufid Majnun on Unsplash

Executive Summary

The COVID-19 pandemic has had a catastrophic impact on the world—inflicting enormous (and often uneven) health, economic, political, and cultural costs. Leaders from around the world have increasingly sought to mitigate these complex problems through the use of data, including data from non-traditional sources. In this report, we refer to that as "non-traditional data" and we define it as data that is "digitally captured [...], mediated [...] or observed," using new instrumentation mechanisms.[1] As documented elsewhere, non-traditional data (NTD) has the potential to expand a decision maker's toolbox to respond to not only pandemics but all kinds of dynamic crises, from climate disasters to energy emergencies. However, the use of non-traditional data for public good purposes remains nascent and there is much to be learned about how to utilize it ethically and effectively.

With the support of the Knight Foundation, The GovLab's #Data4COVID19 Review assesses if and how NTD was used during the different waves of the COVID-19 pandemic and provides guidance for how future data systems may be more effectively employed in future dynamic crises. The Review does this with four briefings that document and evaluate the most prominent uses of NTD: health, mobility, economic, and sentiment analysis. These four uses were synthesized from an assessment of The GovLab's #Data4COVID19 Data Collaborative Repository–a crowdsourced list of almost 300 data collaboratives, competitions, and data-driven efforts that aimed to address the pandemic response.

Based on these briefings and our review of the current literature, we identified the following findings and recommendations for how decision-makers might better use NTD in future crisis management efforts.

---

1 Albanna, Basma, Andreas Pawelke, Jeremy Boy, and Andreas Gluecker. "The Data-Powered Positive Deviance Handbook." The University of Manchester Global Development Institute, November 2021. https://www.gdi.manchester.ac.uk/research/publications/di/dd-wp92/.

Findings

- Throughout the COVID-19 pandemic, officials increasingly called for the use of NTD to answer questions where and when traditional data such as surveys and case data were not sufficient or could not be leveraged. At the same time, given the limitations of NTD as it relates to representativeness, the collection and use of traditional data was needed to ground truth the insights;

- NTD sources were primarily used across four main areas of interest:

    - non-traditional health data (describing the physical health of individuals or populations as it relates to COVID-19 diagnoses and risk factors);
    - non-traditional mobility and geolocation data (referring to the physical location of an individual or object in relation to one another or geography);
    - non-traditional economic data (indicating the economic activity of individuals, groups, or organizations); and
    - non-traditional sentiment data (explaining the attitudes and perceptions of individuals and groups regarding developments related to COVID-19).

- Many projects were focused on responding to COVID-19 rather than increasing pandemic readiness, recovering from the pandemic or reforming systems in response to it. This pattern reflected an overall lack of preparedness for the pandemic and need for the rapid development of initiatives to address the consequences of it.

- NTD initiatives frequently took the form of cross-sectoral data partnerships or collaborations. These initiatives were developed to respond to specific, bounded needs, but the failure to adopt a wider scope often contributed to problems later on. Many institutions did not have the systems and infrastructure in place for these collaborations to be sustainable.

- Several NTD initiatives were developed with the purpose of collecting granular data about populations that was not readily available from traditional data initiatives. However, many of these initiatives were implemented without the necessary social license to do so, leading to public concerns about ethics and hindering public trust in NTD.

- In comparison with previous dynamic crises, COVID-19 was a watershed moment in terms of access to and re-use of NTD in those four areas.

- At the same time, the majority of efforts were fragmented and uncoordinated, reflecting the larger fragmented response to COVID-19, limiting their impact and potential to scale.[2]

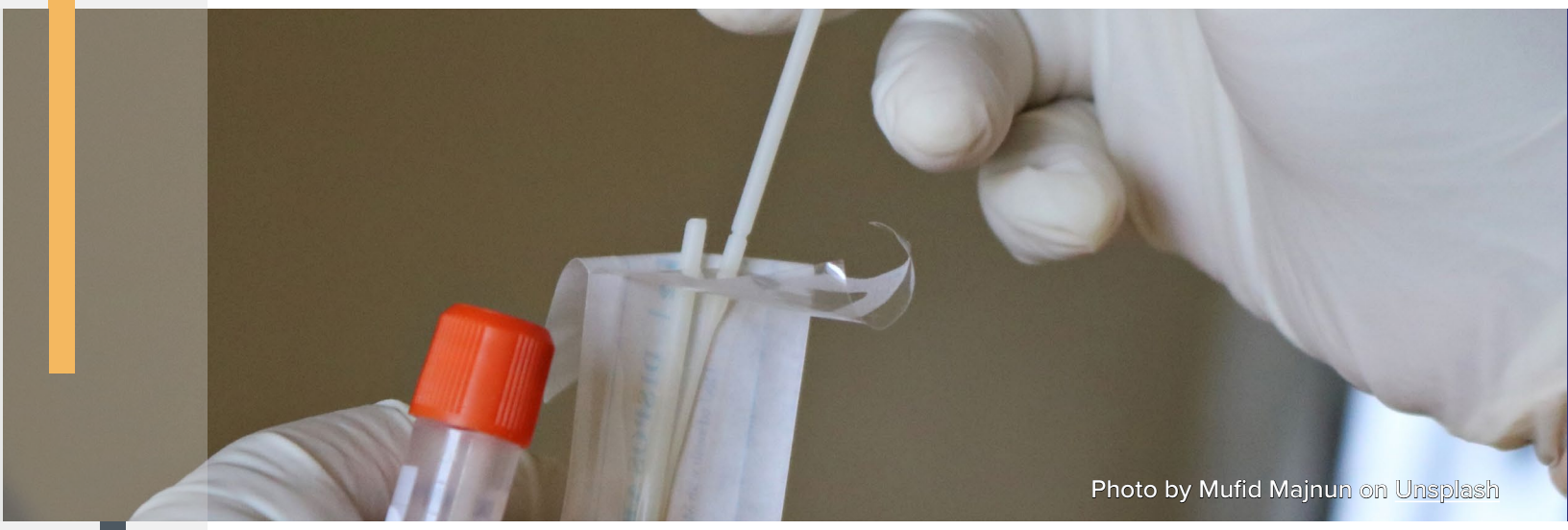


Photo by Mufid Majnun on Unsplash

Recommendations

Based on the findings from the briefings and the analysis of the other information gathered, we have several recommendations that could advance the systematic and responsible use of NTD for crisis response. A more detailed explanation of how to operationalize each category can be found in the recommendations section of the report.

1. **Increasing evidence and awareness about the value proposition of NTD**: There is a need for a stronger evidence base that can generate awareness among decision-makers and the public of current NTD practices–including how personal data is being protected. This basis can help organizations develop the value proposition of and demand for the appropriate use of NTD during crisis situations, demand based

2 Schechtman, Kara, and Sara Simon. "America's Entire Understanding of the Pandemic Was Shaped by Messy Data." The Atlantic, May 25, 2021, sec. Science. https://www.theatlantic.com/science/archive/2021/05/pandemic-data-america-messy/618987/; Ubaldi, Barbara, Arturo Rivera Perez, and Cecilia Emilsson. "Open Data & Covid-19: Looking Forward towards Government Readiness & Reform: Summary Record." Paris, France: OECD Expert Group Meeting on Open Government Data, June 11, 2020. https://www.oecd.org/gov/digital-government/6th-oecd-expert-group-meeting-on-open-government-data-summary.pdf.

on effective and meaningful use. Creating a more data-driven approach to using NTD (that is, having more data about how to use data) can broaden NTD initiatives and ensure they can grow beyond current pilot programs and proofs of concept.

2. **Advancing trust, ethics, and equity within NTD initiatives:** There is a need for public and private organizations to build trust within society by explaining how and why NTD is used. Prioritizing trust, ethics, and equity at the start of and throughout NTD initiatives can increase the legitimacy of its use and more meaningfully resolve crisis-driven challenges based on a combination of NTD and lived experience. It might also be useful to articulate the circumstances under which NTD should not be used to set appropriate expectations of what NTD can and cannot do.

3. **Strengthening collaboration and institutionalization of NTD use**: There is a need for multi-stakeholder partnerships to increase the quality and speed of implementation of NTD initiatives in a systematic, standardized, and scalable way. Institutionalizing internal professional functions and external partnerships through data stewards can accelerate the responsible and efficient use of NTD during crises. Such efforts must necessarily take into account the context and incentives that organizations involved to be effective. When done successfully, institutionalization can ensure that knowledge is preserved and that advancements in NTD capacity are not lost due to "pandemic fatigue."

4. **Preventing fragmentation, and improving readiness and coordination**: There is a need for a more coordinated approach—across all crisis management efforts and functions—to address evolving needs throughout the data and pandemic lifecycle phases. Minimizing fragmentation within jurisdictions and at an international scale through the development of NTD standards and other common practices has the potential to increase readiness for future dynamic crises. However, this reduction of fragmentation must be done in a way that pro-actively addresses concerns about surveillance and institutional ethics.

1. Introduction

Photo by Taylor Brandon on Unsplash

1.1. The Challenge

Pandemics are major crises on all levels, with cataclysmic and unequal effects on society. These effects force governments and businesses to fundamentally change how they operate. They impact the way we interact with families and friends by limiting who we can meet and how we can interact. They aggravate other crises, such as the mental health crisis,[3] which affect people around the world. As the last two years of the COVID-19 pandemic have demonstrated, these impacts can be catastrophically large if institutions responsible for public welfare lack relevant data and are unable to manage COVID-19's consequences.[4] Without data-driven insights,

3 World Health Organization. "COVID-19 Pandemic Triggers 25% Increase in Prevalence of Anxiety and Depression Worldwide," March 2, 2022. https://www.who.int/news/item/02-03-2022-covid-19-pandemic-triggers-25-increase-in-prevalence-of-anxiety-and-depression-worldwide.

4 Baird, Matthew D., David G. Groves, Osonde A. Osoba, Andrew M. Parker, Ricardo Sanchez, and Claude Messan Setodji. "Don't Make the Pandemic Worse with Poor Data Analysis," May 6, 2020. https://www.rand.org/blog/2020/05/dont-make-the-pandemic-worse-with-poor-data-analysis.htm.; Ionnidis, John P.A. "A Fiasco in the Making? As the Coronavirus Pandemic Takes Hold, We Are Making Decisions without Reliable Data." STAT. First Opinion (blog), March 17, 2020. https://www.statnews.com/2020/03/17/a-fiasco-in-the-making-as-the-coronavirus-pandemic-takes-hold-we-are-making-decisions-without-reliable-data/.

decision-makers could find themselves unable to address supply shortages,[5] develop treatments for those infected,[6] or assess the efficacy of disease control measures and adopt better policies.[7] They could struggle to prepare for a crisis, respond to it, recover from it, or reform institutions in its aftermath (see Figure 1).[8]

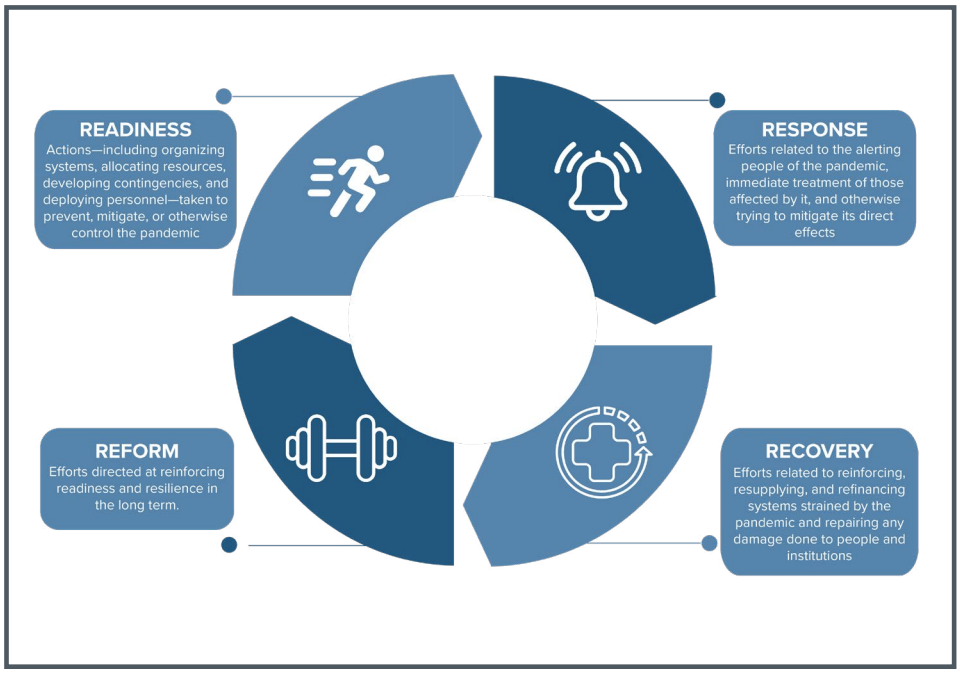


Figure 1: Pandemic Lifecycle Phases adapted from the OECD Risk Management Cycle From the OECD's Risk Management Cycle, this figure demonstrates the different phases of the pandemic response cycle.[8]

We have written about these challenges elsewhere, as have other organizations. In our Open data in action: initiatives during the initial stage of the COVID-19 pandemic, which we co-wrote with the OECD Digital Government and Data Unit we assessed the wide-ranging health, social, and economic impact of the pandemic and the various efforts adopting open

5 Rubin, Rita. "Tracking COVID-19 Supply Shortages." JAMA 326, no. 12 (September 28, 2021): 1138. https://doi.org/10.1001/jama.2021.15924.

6 National Institutes of Health. "COVID-19 Treatment Guidelines," April 29, 2022. https://www.covid19treatmentguidelines.nih.gov/about-the-guidelines/guidelines-development/.

7 Wines, Michael. "Florida Struggled with Bad Data in the Pandemic's First Months, an Audit Finds." The New York Times, June 13, 2022, sec. U.S. https://www.nytimes.com/2022/06/09/us/florida-covid-data-audit.html.

8 For more information on how the OECD Risk Management Cycle can be adapted to the pandemic, see: Felipe González-Zapata, Jacob Arturo Rivera Pérez, Lucia Chauvet, Cecilia Emilsson, Andrew J. Zahuranec, and Andrew Young. "Open Data in Action: Initiatives during the Initial Stage of the COVID-19 Pandemic." OECD and The GovLab, March 2021. https://www.oecd.org/gov/digital-government/use-of-open-government-data-to-address-covid19-outbreak.htm.

government data to address it.[9] The Data-Pop Alliance's C-19 Global South Observatory has also written extensively about the unique data challenges facing countries in the Global South.[10] In "The Lancet Commission on lessons for the future from the COVID-19 pandemic" Professor Jeffrey D Sachs, et al. explain there was a "lack of timely, accurate, and systematic data on infections, deaths, viral variants, health system responses, and indirect health consequences" at an international scale and a need for standardization throughout the data value chain.[11]

1.2. What is Non-Traditional Data?

The University of Manchester Global Development Institute's "The Data-Powered Positive Deviance Handbook" defines Non-Traditional Data (NTD) as follows: "*data that is digitally captured (e.g. mobile phone records and financial data), mediated (e.g. social media and online data), or observed (e.g. satellite imagery)*",[12] using new instrumentation mechanisms, often privately held. NTD is the result of increased digitalization and the accompanying datafication—where rapid technological innovation has led to the transition of more aspects of our lives to the digital world, creating new sources of data that are seen as an asset to different applications or use cases.[13] Consequently, NTD is "based around repurposed assets"[14] meaning that it is typically reused for new use cases beyond its original intention. For example, telecommunications data is generated to inform the activities of telecommunications companies. Telecommunications is not generated to support public health but can be repurposed to fill this function.

---

9 Felipe González-Zapata, Jacob Arturo Rivera Pérez, Lucia Chauvet, Cecilia Emilsson, Andrew J. Zahuranec, and Andrew Young. "Open Data in Action: Initiatives during the Initial Stage of the COVID-19 Pandemic." OECD and The GovLab, March 2021. https://www.oecd.org/gov/digital-government/use-of-open-government-data-to-address-covid19-outbreak.htm.

10 Data-Pop Alliance. "C19 Global South Observatory." Data-Pop Alliance, April 2, 2020. https://datapopalliance.org/covid19/c19globalsouthobservatory/.

11 Sachs, Jeffrey D., Salim S. Abdool Karim, Lara Aknin, Joseph Allen, Kirsten Brosbøl, Francesca Colombo, Gabriela Cuevas Barron, et al. "The Lancet Commission on Lessons for the Future from the COVID-19 Pandemic." The Lancet 0, no. 0 (September 14, 2022). https://doi.org/10.1016/S0140-6736(22)01585-9.

12 Albanna, Basma, Andreas Pawelke, Jeremy Boy, and Andreas Gluecker. "The Data-Powered Positive Deviance Handbook." The University of Manchester Global Development Institute, November 2021. https://www.gdi.manchester.ac.uk/research/publications/di/dd-wp92/.

13 Mejias, Ulises A., and Nick Couldry. "Datafication." Internet Policy Review 8, no. 4 (November 29, 2019). https://policyreview.info/concepts/datafication.

14 Advisory Group Meeting. Zoom, September 9, 2022.

This characteristic also means that NTD often serves as a proxy, an attempt to extrapolate data describing one phenomenon as having insight into another. As the effort's Advisory Group pointed out: when we use NTD, "we are measuring proxies for the real behavior we want to measure. This proxy data and how it can speak about reality requires us to have some sense of the "ground truths," that can improve meaningful interpretation of the information."

When leveraged responsibly, NTD has the potential to enable new ways of addressing dynamic crises such as a pandemic.[15] NTD, including, for instance, telecommunications data and credit card transaction data, are increasingly made available through collaborations with the private sector, academic institutions, and civil society to bolster traditional approaches to public health surveillance.[16] They have complemented case counts and official statistics, generated through traditional survey and statistical methods, with digital technologies that promise to expand practitioners' reach and accuracy.[17] NTD produced several insights that were then tested using traditional data sources. As cited by our Advisory Group, COVID-19 accelerated the digital maturity of many organizations–particularly in the public sector and as result enabled them to provide and leverage NTD beyond COVID-19.

The value proposition of using NTD amid crises is motivated by organizations viewing traditional data as insufficient to respond to dynamic situations and seeing NTD as a valuable complement to existing resources.[18]

---

15 Ramachandran, Selva, and Deepali Khanna. "Leveraging Non-Traditional Data For The Covid-19 Socioeconomic Recovery Strategy." Forbes, February 1, 2022. https://www.forbes.com/sites/deepalikhanna/2022/02/01/leveraging-non-traditional-data-for-the-covid-19-socioeconomic-recovery-strategy/.

16 See Non-Traditional Mobility Data and Non-Traditional Economic Briefings

17 See, for example: The Government of Hong Kong Special Administrative Region. "Government Follows up on Positive Sewage Testing Results and Urges Public to Undergo Testing," January 27, 2022. https://www.info.gov.hk/gia/general/202201/27/P2022012700041.htm;
Ramesh, Aditi, Stefaan Verhulst, Andrew Young, Andrew J. Zahuranec, Brennan Lake, Ben Snaith, and Oliver Thereaux. "The Use of Mobility Data for Responding to the COVID19 Pandemic: Data4COVID19 Deep Dive." Open Data Institute, The GovLab, and Cuebiq, March 24, 2021. http://theodi.org/wp-content/uploads/2021/04/Data4COVID19_0329_v3.pdf;
Meta. "Data For Good COVID-19 Trends and Impact Survey Request Data Access." Accessed September 29, 2022. https://dataforgood.facebook.com/dfg/docs/covid-19-trends-and-impact-survey-request-for-data-access.

18 For this report, we define traditional data as, "those datas that come from sources such as surveys, censuses and administrative records with statistical potential. This is connected with the work of government organizations that generate and use statistical information and are within the framework of the National Statistical Systems coordinated by Statistical Offices." See: Rodríguez, Fredy. "Traditional and Non-Traditional Data Sources for Sustainable Development." Data Republica, May 2021. https://datarepublica.org/en/noticias/84.

In terms of limitations, traditional data can present several challenges. First, the source can have long timelines from data collection to dissemination due to bureaucracy and resource costs—slowing response to a crisis.[19] Traditional data also can have insufficient scale, leading to an inaccurate evidence base.[20] Lastly, traditional data efforts often have barriers that limit the data quality (linkability, completeness, and timeliness) and granularity. This lack of quality and granular data can create challenges in understanding how problems resonate at the local level.[21]

Previous research at The GovLab has documented that NTD use offers a growing potential to drive meaningful impact during crisis situations. Our past work suggests NTD can complement traditional data because of the increased volume, variety, veracity, and velocity[22] of NTD:

- Has the potential to advance our understanding of the current state of our society and identify the underlying causes of large-scale problems;[23]

- Provides new types of assessments that can be used to address different aspects of the

---

19 For examples, see: Banco, Erin. "Inside America's Covid-Reporting Breakdown." Politico, August 15, 2021. https://www.politico.com/news/2021/08/15/inside-americas-covid-data-gap-502565;
Ramachandran, Selva. "Leveraging Non-Traditional Data for COVID-19 Recovery." UNDP Philippines, February 17, 2022. https://www.undp.org/philippines/news/leveraging-non-traditional-data-covid-19-recovery;
The Center for Effective Global Action. "Global Research Insights Using Non-Traditional Data." CEGA (blog), May 12, 2020. https://medium.com/center-for-effective-global-action/global-research-insights-using-non-traditional-data-23f8d8c7b5e4.

20 Fast, Austin. "Millions Of People Are Missing From CDC COVID Data As States Fail To Report Cases." NPR, September 1, 2021, sec. Investigations. https://www.npr.org/2021/09/01/1032885251/millions-of-people-are-missing-from-cdc-covid-data-as-states-fail-to-report-case.;
"Thinking beyond Crisis: Using the Pandemic to Advance High Quality, Timely and Inclusive Data — SDG Indicators." United Nations Statistics Division. Accessed September 7, 2022. https://unstats.un.org/sdgs/report/2022/thinking-beyond-crisis;
Weber, Ingmar, Muhammad Imran, Ferda Ofli, Jennifer Colville, Fouad Mrad, Mehdi Fathallah, Alissar Chaker, and Wigdan Ahmad. "Non-Traditional Data Sources : Providing Insights Into Sustainable Development." Communications of the ACM 64 (March 1, 2021): 88–95. https://doi.org/10.1145/3447739.

21 Holtzman, Geoffrey S., Neda A. Khoshkhoo, and Elaine O. Nsoesie. "The Racial Data Gap: Lack of Racial Data as a Barrier to Overcoming Structural Racism." The American Journal of Bioethics 22, no. 3 (March 4, 2022): 39–42. https://doi.org/10.1080/15265161.2022.2027562;
Piller, Charles. "Data Secrecy Is Crippling Attempts to Slow COVID-19's Spread in U.S., Epidemiologists Warn." Science, July 16, 2020. https://www.science.org/content/article/us-epidemiologists-say-data-secrecy-covid-19-cases-cripples-intervention-strategies.

22 There are several examples of NTD initiatives that align with the 4 Vs of Big Data: Volume, Variety, Veracity, and Velocity. BBVA. "The Five V's of Big Data," May 26, 2020. https://www.bbva.com/en/five-vs-big-data/.

23 Young, Andrew, Andrew Zahuranec, Stefaan G. Verhulst, and Kateryna Gazaryan. "The Third Wave of Open Data Toolkit: Operational Guidance on Capturing Institutional and Societal Value of Data Re-Use." Brooklyn, New York: Open Data Policy Lab, March 2021. http://files.thegovlab.org/The-Third-Wave-of-Open-Data-Toolkit.pdf.

crisis in real-time and develop targeted policy interventions;[24]

- Increase the speed in which we respond to crises, including by supporting data-driven service design and delivery;[25]

- Expands our capabilities across the crisis lifecycle–from preparing for what is possible to recovering from an event and implementing learnings moving forward;[26] and

- Expand the solution set–involving new multidisciplinary actors that can help accelerate operations and ways of working in different domains.[27]

1.3. Limitations of Non-Traditional Data

However, there are several limitations relating to the use of NTD. First, NTD often does not have a specific procedure that outlines the variables of interest before data is collected. As a result, formulating research questions and priorities that can leverage the available metrics can be challenging. Second, NTD can include personally identifiable information and requires data protection throughout the data lifecycle. Lastly, NTD has its own gaps.

As discussed, NTD often provides proxies for other activities (e.g. cell phone location for human movement, wastewater data for infection rates), and needs to be complemented by other data sources to be reliable and representative. Consequently, NTD frequently needs to be complemented or verified with other data sources like well-constructed surveys designed in a traditional setting when measuring characteristics such as symptoms and opinion. Surveys can be vital in understanding how people perceive and understand developments related to COVID-19, as evidenced by the COVID19impactsurvey.org in helping researchers understand symptoms and the overall attitudes of people throughout

---

24 Barnett, Aleise, David Dembo, and Stefaan Verhulst. "Toward Metrics for Re(Imagining) Governance: The Promise and Challenge of Evaluating Innovations in How We Govern." SSRN Scholarly Paper. Rochester, NY, April 18, 2013. https://doi.org/10.2139/ssrn.2563794.

25 The GovLab. "A Call for Action." Medium. Data Stewards Network (blog), March 16, 2020. https://medium.com/data-stewards-network/a-call-for-action-813669f32244.

26 Felipe González-Zapata, Jacob Arturo Rivera Pérez, Lucia Chauvet, Cecilia Emilsson, Andrew J. Zahuranec, and Andrew Young. "Open Data in Action: Initiatives during the Initial Stage of the COVID-19 Pandemic." OECD and The GovLab, March 2021. https://www.oecd.org/gov/digital-government/use-of-open-government-data-to-address-covid19-outbreak.htm.

27 Zahuranec, Andrew J. "New Report: 'Leveraging Private Data for Public Good.'" The Govlab Blog, October 31, 2019. https://blog.thegovlab.org/archive/new-report-leveraging-private-data-for-public-good.

the crisis.[28] institutions, including the Nigeria Centre for Disease Control, similarly, used surveys to assess compliance with non-pharmaceutical interventions and understand public opinion.[29] Similar initiative, such as Influenzanet, an online flu surveillance system supported by voluntary surveys, has a long history of use predating the COVID-19 pandemic,[30] and increased in value during the pandemic.

1.4. Research Objectives

This report seeks to provide greater clarity on the use and role of NTD during the COVID-19 pandemic.[31] While analyzing events leading up to and between crises is valuable in assessing the preparation for the pandemic response, this interrogation is specifically motivated by a desire to test the suggestion that COVID-19 was a watershed moment in accessing and re-using NTD. As David Buckeridge (professor at McGill University Medical School) stated, "This pandemic is unprecedented, and the amount of data sharing is unprecedented."[32] To date, limited research has been undertaken to evaluate whether and what types of NTD initiatives were most effective during COVID-19. This report further seeks to understand how organizations might improve future preparation, response, recovery, and reform from dynamic crises based on lessons from initiatives during COVID-19.[33] It supplements prior research and analysis by The GovLab through its #Data4COVID19 initiative.[34]

---

28 It should be acknowledged, however, that these and other surveys often have qualities of NTD sources by virtue of their use of social media platforms to achieve scale. See ELLIS Alicante Foundation. "Covid19 Impact Survey Results." Accessed September 13, 2022. https://ellisalicante.org/en/covid19impact-survey.;Martínez-Garcia, Marina, Emilio Sansano-Sansano, Andrea Castillo-Hornero, Ruben Femenia, Kristof Roomp, and Nuria Oliver. "Social Isolation during the COVID-19 Pandemic in Spain: A Population Study." Scientific Reports 12, no. 1 (July 22, 2022): 12543. https://doi.org/10.1038/s41598-022-16628-y.

29 Nigeria Centre for Disease Control. "Data4COVID19 Africa Challenge: Inception Meeting of the Nigerian Project," June 25, 2021. https://datachallenge.africa.

30 Data-Pop Alliance. "C19 Global South Observatory." Data-Pop Alliance, April 2, 2020. https://datapopalliance.org/covid19/c19globalsouthobservatory/.

31 This definition is drawn from: Albanna, Basma, Andreas Pawelke, Jeremy Boy, and Andreas Gluecker. "The Data-Powered Positive Deviance Handbook." The University of Manchester Global Development Institute, November 2021. https://www.gdi.manchester.ac.uk/research/publications/di/dd-wp92/.

32 Rocha, Roberto. "The Data-Driven Pandemic: Information Sharing with COVID-19 Is 'Unprecedented.'" CBC News, March 17, 2020. https://www.cbc.ca/news/canada/coronavirus-date-information-sharing-1.5500709.

33 Albanna, Basma, Richard Heeks, Andreas Pawelke, Jeremy Boy, Julia Handl, and Andreas Gluecker. "Data-Powered Positive Deviance: Combining Traditional and Non-Traditional Data to Identify and Characterise Development-Related Outperformers." Development Engineering 7 (January 1, 2022): 100090. https://doi.org/10.1016/j.deveng.2021.100090.

34 The GovLab's #Data4COVID19 effort includes its "Call For Action to Build A Data Infrastructure and

This work is critical given the continuing cost of the COVID-19 pandemic. Governments and all institutions responsible for the public well-being urgently need to identify data-driven solutions that can meaningfully improve people's lives. We believe this work is essential not only to understand what happened during the largest public health crisis of the century but also to ensure public institutions are aware of its lessons about data-driven crisis management.

At the same time, NTD, and its reuse, can pose issues, such as the interoperability of repositories or data privacy and anonymity concerns.[35] It is imperative to assess how these efforts played a role and if they can be integrated in the prevention of (or response to) future disease outbreaks.

There is a growing literature on data (re)use for pandemic response, some of which this report cites.[36] This report seeks to be distinctive by supplementing these resources with real-world examples and interviews from practitioners in the field. It seeks to contribute to the ongoing discussion on the value and opportunities of NTD and other data sources.

Ecosystem Capable Of Addressing Dynamic Crises," its data collaborative repositories tracking innovative data reuse efforts for pandemic response, and its #Data4COVID19 Africa Challenge which funded teams across Africa to use innovative data sources to address COVID-19. The #Data4COVID19 effort began initially as an effort to track data-driven responses to COVID-19 and has developed into a larger effort to assess the use of data and act upon those insights. Read more at: https://data4covid19.org/.

35 Boté-Vericad, Juan-José, and Miquel Termens. "Reusing Data: Technical and Ethical Challenges." DESIDOC Journal of Library & Information Technology 39 (November 1, 2019): 329–37. https://doi.org/10.14429/djlit.39.6.14807.

36 For examples, see: John P.A. Ioannidis. "A Fiasco in the Making? As the Coronavirus Pandemic Takes Hold, We Are Making Decisions without Reliable Data." STAT, March 17, 2020. https://www.statnews.com/2020/03/17/a-fiasco-in-the-making-as-the-coronavirus-pandemic-takes-hold-we-are-making-decisions-without-reliable-data/;
Joseph Menzin and Peter Neumann. "How Real-World Data Can Help Us Better Prepare for the Next Pandemic." Scientific American, April 22, 2021. https://www.scientificamerican.com/article/how-real-world-data-can-help-us-better-prepare-for-the-next-pandemic/;OECD. "Enhancing Access to Research Data during Crises: Lessons Learned from the COVID-19 Pandemic," April 23, 2021. https://www.oecd.org/sti/inno/enhance-access-research-data-during-crises.htm.;
Pillai, Priyanka. "How Do Data Bolster Pandemic Preparedness and Response? How Do We Improve Data and Systems to Be Better Prepared?" Patterns 2, no. 1 (January 8, 2021): 100190. https://doi.org/10.1016/j.patter.2020.100190.;
Thomas H. Davenport, A. Blanton Godfrey, and Thomas C. Redman. "To Fight Pandemics, We Need Better Data." MIT Sloan Management Review, August 25, 2020. https://sloanreview.mit.edu/article/to-fight-pandemics-we-need-better-data/;
Rositsa Zaimova. "How Data Can Help Fight a Health Crisis like the Coronavirus." World Economic Forum, March 31, 2020. https://www.weforum.org/agenda/2020/03/role-data-fight-coronavirus-epidemic/.

1.5. Methodology

#Data4COVID19 is a series of projects undertaken by The GovLab—with partners such as the OECD, Open Data Charter, Luce Foundation, Microsoft, Rockefeller Foundation, SDSN TRENDS, World Economic Forum, University of Washington, and others—to identify, collect, and analyze the value data can provide to the ongoing COVID-19 pandemic. As a part of this work, The GovLab developed the #Data4COVID-19 repository, a crowdsourced list of almost 300 data collaboratives, competitions, and data-driven efforts that aim to address different aspects of the pandemic response. Initiatives included in this list were identified and curated by The GovLab #Data4COVID19 team or suggested through an intake form shared with the Data Stewards Network.

As The GovLab's #Data4COVID19 repositories demonstrate, there have been a variety of different efforts to use NTD for the COVID-19 response.[37] Comprehensively describing and analyzing all possible uses in all possible contexts would be a complex, if not impossible, task. As such, the research team focused on only those uses of NTD that could provide instructive lessons to decision-makers. These included data efforts that recurred in different contexts and settings and for which there was sufficient information available to assess the effort. A more detailed description of the inclusion and exclusion criteria is shown in Appendix 2.

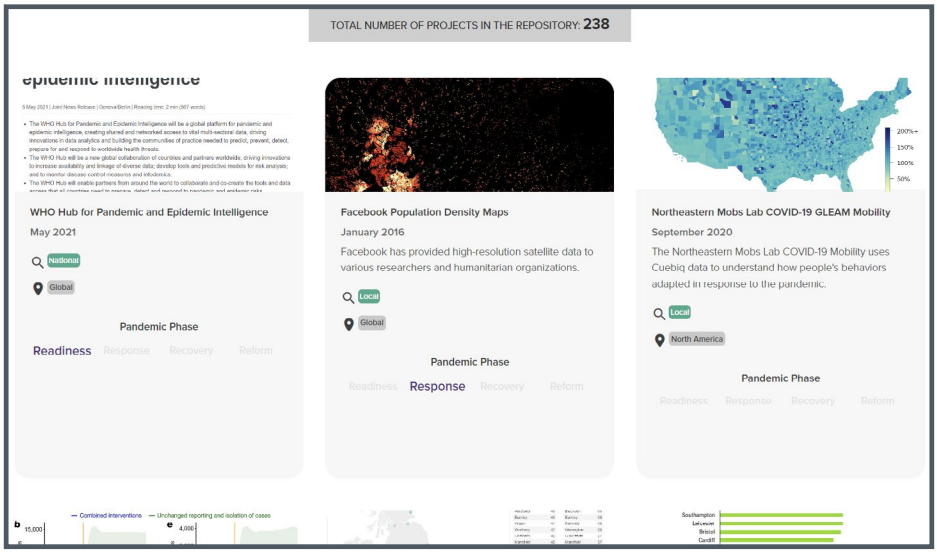


Figure 2: Screenshot of the #Data4COVID19 Data Collaborative Repository. See: https://list.data4covid19.org/

37 The GovLab. "A Living Repository for Data Collaboratives Seeking to Address the Spread of COVID-19." #Data4COVID19. Accessed September 14, 2022. https://list.data4covid19.org/.

Subsequently, the team developed a multi-staged and mixed-method approach with the aim of developing recommendations for future data-driven crisis management efforts. It achieved this goal by developing briefings (short form, digestible descriptive analyses for decision makers and their staff) that explain how different types of NTD were used during COVID-19. This approach adopts the following steps:

1. **Desk Research:** The GovLab rapidly reviewed existing research around data (re)use for COVID-19 and other crises.

2. **Research Plan:** From the desk research, it developed a proposal for a frame of analysis (relying on pandemic phases and data sources) and a standardized assessment framework to evaluate the findings from different use cases or NTD initiatives (See Appendix 3).

3. **Expert Group:** The GovLab assembled an international Advisory Group of seven experts in data science, public health, human rights, and innovation to understand what kinds of examples and issues our briefings should prioritize. The experts, a global cohort identified on pg. 1, participated in a call in which they discussed The GovLab's research, frame of analysis, and assessment framework. They also identified several topics for The GovLab to research that had not been previously included.

4. **Literature Review and Briefing Development:** Based on feedback provided by this group, The GovLab conducted a review of the literature on data-driven responses to the COVID-19 pandemic–drawing from the #Data4COVID19 Data Collaborative Repository. It collected and grouped responses by data type, data source, and pandemic lifecycle phase and narrowed the information gathered based on the criteria within the frame of analysis. The four briefings were selected based on the four largest categories of uses within the repository. It subsequently developed briefings on each category, conducting interviews with individuals who helped deploy particular types of responses.

5. **Briefing Analysis:** Following the development of the briefings, The GovLab evaluated the findings of specific use cases or NTD initiatives using the assessment framework criteria (see Figure 3) and developed several cross-cutting lessons

learned across all briefings. The assessment framework breaks down different components of a project—inputs and planning and governance and outputs. Using this approach, we could assess how one aspect of an initiative was realized well while another poorly. This process aimed to provide the evidence base for recommendations for future data-driven crisis management efforts as well as test how our assessment framework could best be utilized.

6. **Peer Review:** The GovLab shared initial findings and updated assessment framework with the expert group for discussion. After receiving their feedback into the draft and hosting a virtual call to discuss it, the team revised the report. During this time, The GovLab also submitted an open call for expert review on its Data Stewards Network Newsletter and reached out to several industry experts. The experts subsequently conducted a supplementary review of this final draft.

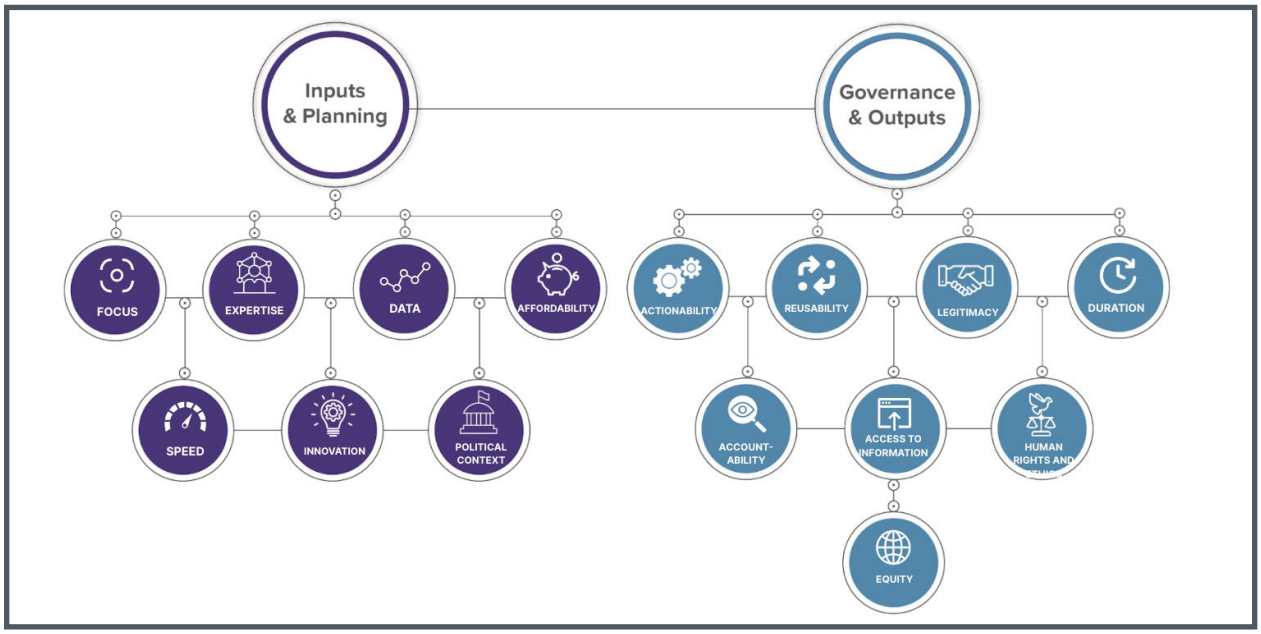


Figure 3: Assessment Framework
This figure includes a summary of the assessment framework criteria used to evaluate specific NTD initiatives. A more detailed explanation of the assessment framework criteria can be found in Appendix 3.

1.6. How to Use This Report

As mentioned, this document seeks to provide decision-makers and their staff with more clarity on four types of NTD sources for pandemic response—non-traditional health, mobility, and economic, and sentiment data. The four briefings describe what is involved in using a particular type of NTD for pandemic response and the considerations that officials should be cognizant of should they pursue that application themselves. Each briefing includes six parts:

- **Problem at Hand:** What was the overarching need that data practitioners sought to fill before deciding to pursue this data type?

- **Data Sources:** What are the types of data (or data sources) included in this data type and how are they generated or collected?
    - What are real-world examples of this data in use?
    - During which phases of the pandemic lifecycle was this data used?

- **Design and Planning:** How were efforts to use these different data sources organized and what types of questions did it seek to answer?

- **Implementation and Outcomes:** Generally, how were these efforts deployed in the field and what tended to be the result of these efforts?

- **Additional Examples:** Are there specific, instructive examples of that work that decision-makers and their staff can learn from?

- **Vignette:** What was a noteworthy use case or NTD initiative within this data category?
    - Why was this initiative developed?
    - How was the initiative designed and planned?
    - How was it implemented and what were the outcomes?
    - How did the initiative perform across our assessment criteria?

Each briefing intends to help a practitioner understand how data-driven projects are developed from design to implementation. We do not make a value judgment about any data type in the abstract, as the responsibility and effectiveness of any deployment depends on context-specific factors. However, using the vignettes, our briefings provide a model for assessing specific projects. This assessment, developed in coordination with our Advisory Group, centers on the dual pillars of "Inputs and Planning" and "Governance and Outputs," both of which we believe to be necessary for a well-managed initiative.

The latter part of this report includes the analysis of the briefings. First, it identifies findings that apply to all of the briefings. Second, it evaluates those findings across the assessment framework criteria for specific NTD initiatives along with the information gathered throughout the research process and provides top-level lessons learned. Lastly, it provides recommendations for decision-makers looking to incorporate NTD in future crisis response efforts systematically. A detailed explanation of our assessment framework as well as a toolkit that can be used to assess future crisis NTD initiatives can be found in Appendix 3.

2. A Review of NTD Initiatives for the COVID-19 Response

Photo by Jakayla Toney on Unsplash

In what follows, we build on the value proposition of NTD (both potential and realized) in the context of COVID-19 within four data types. Each of the four data types have been used in efforts led by the public sector to respond to the pandemic. The briefings describe several data sources that most stood out within each of the four categories during the literature review, review of the #Data4COVID19 Data Collaborative Repository, and speaking with our Advisory Group. A summary of the data types and sources is shown in Table 1.

Unstructured NTD can be difficult to categorize. As such, there were initiatives that did not fit perfectly into one of the four given categories and encapsulated more than one, or all four categories.[38]

38 Oliver, Nuria, Xavier Barber, Kirsten Roomp, and Kristof Roomp. "Assessing the Impact of the COVID-19 Pandemic in Spain: Large-Scale, Online, Self-Reported Population Survey." Journal of Medical Internet Research 22, no. 9 (September 10, 2020): e21319. https://doi.org/10.2196/21319.

Briefing	Description	Data Sources Included
Non-Traditional Health Data	NTD that describes, or can be used as a proxy for, the physical health of individuals or populations, especially as it relates to COVID-19 diagnoses.	• Digital patient data • Platform-facilitated symptoms data • Wastewater data • Wearables and biometrics
Non-Traditional Mobility Data	NTD that describes the location of a device relative to people and landmarks in the physical world or to other devices.	• Telecommunications data: call detail records, x-data records • Software development kit data: first-party SDK-derived data, third party SDK-derived data • GPS and other resources: vehicle GPS, bluetooth
Non-Traditional Economic Data	NTD that describes the economic activity of individuals, groups, or organizations.	• Credit and debit card transactions • Open contracting • Supply chain data
Non-Traditional Sentiment Data	NTD that describes how individuals and groups perceive and understand developments related to COVID-19.	• Social media data • Crowdsourced data

Table 1: Four Briefings
The table above outlines the data types and sources included in each briefing. There is a certain maturity level where NTD becomes traditional. While these NTD sources may be considered traditional in different sectors due to the maturity of use, they involve new instrumentation methods in the context of crisis management.

Each of the following briefings includes a vignette–a description of how one project evolved from design to implementation. The quantitative assessments of the vignettes are subjective and not based on mathematical calculations. They were developed leveraging the information gathered for each criteria and then critiqued through an open peer review.

Photo by Steven Cornfield on Unsplash

2.1. Non-Traditional Health Data

Main Takeaways

- **Main use cases:** Non-traditional health data sources were primarily used to validate, supplement, or direct traditional COVID-19 testing data and understand how the virus impacted different population segments.
- **Agile methods and indicators:** Non-traditional health data initiatives that were successful in tracking the progression of symptoms in the long-term were designed in an agile manner with flexible indicators.
- **Cross-sectoral collaboration:** Non-traditional health data initiatives typically had significant resource costs. Multidisciplinary partnerships—specifically for waste-water and digital patient data initiatives—increased data capacity and accelerate timelines.
- **Combining with traditional data sources:** The transition from data to decision intelligence was a challenge across all non-traditional health data sources due to the necessity to integrate the data gathered and findings within traditional data systems.
- **Governance approach:** Government support for non-traditional health data efforts—specifically digital patient data initiatives—entailed suspending privacy policies to streamline external access to data.

2.1.1. Problem at Hand

During a media briefing on March 16th, 2020, Tedros Adhanom Ghebreyesus, Director-General of the World Health Organization (WHO), stated: *"You cannot fight a fire blindfolded. And we cannot stop this pandemic if we don't know who is infected. We have a simple message for all countries: test, test, test."*[39]

Since the start of the pandemic, testing has been at the core of government response efforts. However, many clinical testing initiatives—whether they be PCR (polymerase chain reaction), typically the most reliable tests even in the absence of symptoms, antigen or serologic tests[40]– were not available to everyone, were delayed, or inconsistently delivered.[41]

"You cannot fight a fire blindfolded. And we cannot stop this pandemic if we don't know who is infected. We have a simple message for all countries: test, test, test."

These problems stem from various sources, including limited government capacity and capabilities; issues with how labs process data; and the reliance on physical tools and systems that are labor intensive.[42] In lieu of widely available clinical tests and concerns about accuracy, many organizations sought proxies that could validate or provide additional case counts.[43] These proxies, used from the local to national levels, included

---

39 World Health Organization. "WHO Director-General's Opening Remarks at the Media Briefing on COVID-19 - 16 March 2020," March 16, 2020. https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---16-march-2020.

40 As stated by the WHO, these tests are types of clinical COVID-19 tests that detect genetic materials, viral proteins and antibodies– all signals of current or previous COVID-19 infection "An Update on Testing Strategies for COVID-19." World Health Organization, January 15, 2021. https://www.who.int/docs/default-source/coronaviruse/risk-comms-updates/updates46-testing-strategies.pdf.

41 Apuzzo, Matt, and Selam Gebrekidan. "Can't Get Tested? Maybe You're in the Wrong Country." *The New York Times*, March 20, 2020. https://www.nytimes.com/2020/03/20/world/europe/coronavirus-testing-world-countries-cities-states.html.

42 Pandit, Jay A., Jennifer M. Radin, Giorgio Quer, and Eric J. Topol. "Smartphone Apps in the COVID-19 Pandemic." *Nature Biotechnology* 40, no. 7 (July 2022): 1013–22. https://doi.org/10.1038/s41587-022-01350-x.

43 Yaraghi, Niam. "The US Lacks Health Information Technologies to Stop COVID-19 Epidemic." *Brookings* (blog), March 13, 2020. https://www.brookings.edu/blog/techtank/2020/03/13/the-u-s-lacks-health-information-technologies-to-stop-covid-19-epidemic/.

"various degrees of precision, convenience, and cost."[44] While privacy advocates have raised concerns about the need for additional data protections and ambiguity surrounding how privately run companies may repurpose this data,[45] the public has generally supported increased access to non-traditional health data for the public good–specifically digital patient data–and been willing to supply their data.[46]

2.1.2. Data Sources

Non-traditional health data refers to the types of data that describe the physical health of individuals or groups as it relates to COVID-19 diagnoses. This category does not include traditional epidemiological data sources such as clinical COVID-19 testing data (i.e. PCR tests, rapid antigen tests). In what follows, we describe four sources of non-traditional health data used during COVID-19: digital patient data, platform-facilitated symptom data, wastewater data, and wearables and biometrics.

Digital patient data: Digital patient data refers to electronic health records and metrics of COVID-19 patients. While the use of patient data during a pandemic is not new and varied between countries, COVID-19 accelerated data collaborations by combining digital patient data from medical facilities, private sector organizations, and other sources within large-scale databases or "warehouses."[47] Generally, the efforts explored in this research aimed to broaden the understanding of the virus, design targeted policy interventions, and accelerate research. They may also be used for public health surveillance and to better manage the movement of people, such as in the case of digital immunity certificates.[48]

44 Rosenberg, Amy. "Wastewater Surveillance for COVID-19: It's Complicated." Tufts Now, June 21, 2022. https://now.tufts.edu/2022/06/21/wastewater-surveillance-covid-19-its-complicated.

45 Institute of Medicine (US) Roundtable on Value & Science-Driven Health Care. *Clinical Data as the Basic Staple of Health Learning: Creating and Protecting a Public Good: Workshop Summary. Clinical Data as the Basic Staple of Health Learning: Creating and Protecting a Public Good: Workshop Summary*. Summary. Washington, DC: National Academies Press (US), 2010. https://www.ncbi.nlm.nih.gov/books/NBK54290/.

46 It should be noted however that, with the expansion of home-based testing, there has been a growing challenge of lack of reporting, which has impacted epidemiological assessments at later phases of the pandemic. *See* also: Tosoni, Sarah, Indu Voruganti, Katherine Lajkosz, Shahbano Mustafa, Anne Phillips, S. Joseph Kim, Rebecca K. S. Wong, et al. "Patient Consent Preferences on Sharing Personal Health Information during the COVID-19 Pandemic: 'The More Informed We Are, the More Likely We Are to Help.'" *BMC Medical Ethics* 23, no. 1 (December 2022): 53. https://doi.org/10.1186/s12910-022-00790-z.;
Gerdon, Frederic, Helen Nissenbaum, Ruben L. Bach, Frauke Kreuter, and Stefan Zins. "Individual Acceptance of Using Health Data for Private and Public Benefit: Changes During the COVID-19 Pandemic." *Harvard Data Science Review*, April 6, 2021. https://doi.org/10.1162/99608f92.edf2fc97;
David Betts, Leslie Korenda, and Shane Giuliani. "Are Consumers Already Living the Future of Health?" Deloitte Insights, August 13, 2020. https://www2.deloitte.com/us/en/insights/industry/health-care/consumer-health-trends.html.

47 Madhusoodanan, Jyoti. "Unlocking the Potential of Health Data to Help Research and Treatments." *Nature* 605, no. 7908 (May 3, 2022): 182–83. https://doi.org/10.1038/d41586-022-01205-0.

48 "Immunity Certificates: If We Must Have Them, We Must Do It Right." White Paper. COVID-19 Rapid

- ***Example:*** *In 2018,[49] the Multi-Regional Clinical Trials Center (Brigham and Women's Hospital and Harvard University) in collaboration with other actors launched Vivli–a non-profit organization that functions as a data broker for clinical trial data.[50]*

  In March 2020, Vivli launched a portal for COVID-19 clinical trial sharing. [51] Using this portal, data providers must fill out a "Data Contribution Form,"[52] sign a "Data Contributor Agreement",[53] and anonymize their data prior to publishing.[54] D-wise,[55] a partner organization, provides support in anonymizing data that could be personally identifiable and sharing the data.[56] Researchers access the data by signing up for a Vivli account,[57] searching for studies on the platform, and filling out a request form–describing the study to which the data will be used and who is involved.[58] Their request is analyzed by representatives from Vivli and the data provider as well as external reviewers familiar with the research area[59] based upon criteria set forth by Vivli and the data provider.[60] Following this assessment, the researcher must sign a data-sharing agreement that the data provider approved

REsponse Impact Initiative. Edmond J. Safra Center for Ethics, April 20, 2020. https://ethics.harvard.edu/files/center-for-ethics/files/12immunitycertificates.pdf.

49 Li, Rebecca, Julie Wood, Amrutha Baskaran, Stanley Neumann, Elizabeth Graham, Marcia Levenstein, and Ida Sim. "Timely Access to Trial Data in the Context of a Pandemic: The Time Is Now." *BMJ Open* 10, no. 10 (October 2020): e039326. https://doi.org/10.1136/bmjopen-2020-039326.

50 The GovLab. "Vivli Platform." #Data4COVID19 Repository, 2021. https://list.data4covid19.org/projects/vivli-platform.

51 Regan, Gabby. "Share Your COVID-19-Related Trials on the Vivli Portal." Vivli, April 13, 2020. https://vivli.org/vivli-covid-19-portal/.

52 "Vivli Data Contribution Form." Vivli Center for Global Clinical Research Data, September 22, 2020. https://vivli.org/wp-content/uploads/2020/11/2020_09_22_Data-Contribution-Form.pdf.

53 Vivli. "How to Share Your Clinical Research Data on the Vivli COVID-19 Portal," June 2021. https://vivli.org/how-to-share-your-research-on-the-vivli-covid-19-portal/.

54 Vivli. "How to Share Your Clinical Research Data on the Vivli COVID-19 Portal," June 2021. https://vivli.org/how-to-share-your-research-on-the-vivli-covid-19-portal/.

55 Regan, Gabby. "Share Your COVID-19-Related Trials on the Vivli Portal." Vivli - Center for Global Clinical Research, April 13, 2020. https://vivli.org/vivli-covid-19-portal/.

56 d-wise. "Clinical Trial Transparency for Clinical Data & Document Disclosure | D-Wise," 2022. https://www.d-wise.com/transparency.

57 "Quick Start Guide: Vivli Users." Vivli Center for Global Clinical Research Data, January 14, 2022. https://vivli.org/wp-content/uploads/2022/01/2022_01_14-Vivli_quick_start_users_training.pdf.

58 "How-To: Request Studies on Vivli." Vivli Center for Global Clinical Research Data, October 6, 2021. https://vivli.org/wp-content/uploads/2021/12/2021_12_13-How-to-Request-Studies-on-Vivli-2.0.pdf.

59 Vivli. "Data Request Process Overview." Accessed August 22, 2022. https://vivli.org/about/data-request-review-process/.

60 Li, Rebecca. "Vivli Clinical Research Data Sharing: Share. Discover. Innovate." Vivli - Center for Global Clinical Research, September 2019. https://vivli.org/wp-content/uploads/2020/05/2019_9_19_OHRP.pdf.

or designed.[61] In 2021, Vivli data supported around 50 scientific publications and continues to operate to date.[62]

***Tag:** Response and Recovery*

Platform-facilitated symptom data: This category includes the use of new survey nstrumentation approaches such as internet searches, social media, privately run websites, and mobile applications to gather information from the public about COVID-19 symptoms. Platform-facilitated symptom data was used within both short- and long-term projects and relied upon the self-identification of symptoms.[63] Key initiatives studied in this report aim to accelerate research and find new indicators of infection to guide policy solutions.

- ***Example:** Professor Tim Spector at King's College London and ZOE, a digital nutrition company,[64] launched an App called ZOE in March 2020 to collect self-reported COVID-19 symptoms of different population segments using self-reported symptoms surveys.[65] To participate, users download the app and provide consent for their data to be repurposed for uses including and beyond COVID-19 research.[66] Following data collection, The ZOE team anonymizes the data and shares it with research institutions that signed their data-sharing agreement.*

  Since its launch, the initiative has been funded by United Kingdom's Department of Health and Social Care grants among others.[67] Findings from ZOE have contributed to approximately 50 academic papers. The initiative has also impacted government decision-making; the loss of smell and taste were identified as early indicators of COVID-19 in April 2020 and subsequently added to the United Kingdom's official

---

61 "How-To: Request Studies on Vivli." Vivli Center for Global Clinical Research Data, October 6, 2021. https://vivli.org/wp-content/uploads/2021/12/2021_12_13-How-to-Request-Studies-on-Vivli-2.0.pdf.

62 Vivli. "Highlights Milestones Revenue & Expenses." Accessed August 22, 2022. https://vivli.org/annual-report/.

63 Vermicelli, Silvia, Livio Cricelli, and Michele Grimaldi. "How Can Crowdsourcing Help Tackle the COVID-19 Pandemic? An Explorative Overview of Innovative Collaborative Practices." *R&D Management* 51, no. 2 (March 2021): 183–94. https://doi.org/10.1111/radm.12443.

64 "Understand How Food Affects Your Body." Accessed September 1, 2022. https://joinzoe.com/.

65 Wakefield, Jane. "Zoe Covid-Tracking App Loses Government Funding." *BBC News*, March 11, 2022, sec. Technology. https://www.bbc.com/news/technology-60708330.

66 "Frequently Asked Questions - COVID Symptom Study." Accessed August 26, 2022. https://health-study.joinzoe.com/faq.

67 "ZOE Health Study." Accessed July 22, 2022. https://health-study.joinzoe.com/.

list of COVID-19 symptoms in May 2020.[68] ZOE is still in operation and currently examining long-COVID[69] and other indicators of health beyond COVID-19 (e.g., mental health).[70]

Tag: *Response*

Wastewater Data: Wastewater—which is a combination of household sewage, industrial run-off, and, in some places, storm water—can contain trace amounts of the SARS-CoV-2 virus and detect COVID-19 among asymptomatic and untested individuals, allowing for earlier detection of outbreaks.[71] Samples of wastewater from sewer systems have been used to provide a more holistic view of COVID-19 risk and transmission at the local and regional levels, in a cost-effective manner.[72] Wastewater data has been used to identify communities in need of additional pandemic response measures, target testing and vaccination, and inform public health communications.[73]

- ***Example:*** *In October 2020, the Government of Hong Kong launched its wastewater surveillance initiative with the goal of identifying areas in need of additional clinical testing.[74] The initiative began as a pilot program in collaboration with the Drainage Services Department and the Environmental Protection Department[75] and was then*

---

68 Spector, Tim, Sebastien Ourselin, and Claire Steves. "ZOE COVID Study App: How King's Researchers Slowed the Spread of COVID-19." King's College London, May 7, 2022. https://www.kcl.ac.uk/news/spotlight/zoe-covid-study-app-kings-researchers-slowed-the-spread-covid-19.

69 Spector, Tim, Sebastien Ourselin, and Claire Steves. "ZOE COVID Study App: How King's Researchers Slowed the Spread of COVID-19." King's College London, May 7, 2022. https://www.kcl.ac.uk/news/spotlight/zoe-covid-study-app-kings-researchers-slowed-the-spread-covid-19.

70 Reynolds, Matt, and Morgan Meaker. "How Covid Tracking Apps Are Pivoting for Commercial Profit." *Wired UK*, June 23, 2022. https://www.wired.co.uk/article/covid-19-data-switch.

71 Manuel, Doug, Carlo Alberto Amadei, Jonathon R. Campbell, Jean-Martin Brault, and Jeremy Veillard. "Strengthening Public Health Surveillance Through Wastewater Testing: An Essential Investment for the COVID-19 Pandemic and Future Health Threats." Washington, DC: World Bank, January 19, 2022. https://openknowledge.worldbank.org/handle/10986/36852.

72 Keshaviah, Aparna. Interview re: wastewater data. Zoom, July 11, 2022.

73 Keshaviah, Aparna, Ruchir N Karmali, Divya Vohra, Xindi C Hu, and Megan B Diamond. "The Role of Wastewater Data in Pandemic Management." Washington, DC: Mathematica, April 2022.

74 The Government of Hong Kong Special Administrative Region. "Government Follows up on Positive Sewage Testing Results and Urges Public to Undergo Testing," January 27, 2022. https://www.info.gov.hk/gia/general/202201/27/P2022012700041.htm.

75 Deng, Yu, Xiawan Zheng, Xiaoqing Xu, Ho-kwong Chui, Wai-kwan Lai, Shuxian Li, Hein Min Tun, et al. "Use of Sewage Surveillance for COVID-19: A Large-Scale Evidence-Based Program in Hong Kong." *Environmental Health Perspectives* 130, no. 5 (May 2022): 057008. https://doi.org/10.1289/EHP9966.

widely expanded using technology developed by the University of Hong Kong– which uses Geospatial data collection methods within the sewage system.[76]

In June 2021, a wastewater sample from an area representing 33,000 residents, where viral levels had plateaued for some time, registered the presence of COVID-19.[77] *This sample became the first evidence of the spread of the Delta variant in Hong Kong.*[78] *From there, the government of Hong Kong undertook more localized wastewater surveillance in that area to identify blocks and buildings of focus formandatory clinical testing.*[79] *Wastewater testing uncovered the positive Delta variant case 3 days earlier than clinical testing.*[80]

Tag: *Readiness and Response*

---

76 Tang, Winnie. "Digitising the Information of Underground Drainage Network." Ejinsight, March 7, 2022. ejinsight.com/eji/article/id/3061663/20220307-Digitising-the-information-of-underground-drainage-networ.

77 Deng, Yu, Xiaoqing Xu, Xiawan Zheng, Jiahui Ding, Shuxian Li, Ho-Kwong Chui, Tsz-Kin Wong, Leo L. M. Poon, and Tong Zhang. "Use of Sewage Surveillance for COVID-19 to Guide Public Health Response: A Case Study in Hong Kong." *The Science of the Total Environment* 821 (May 15, 2022): 153250. https://doi.org/10.1016/j.scitotenv.2022.153250.

78 Ibid.

79 Ibid.

80 The University of Hong Kong. "Sewage Testing Uncovered Delta Variant in Community Sewage before Hong Kong's First Delta Case Was Identified," July 4, 2021. https://www.hku.hk/press/news_detail_22978.html.

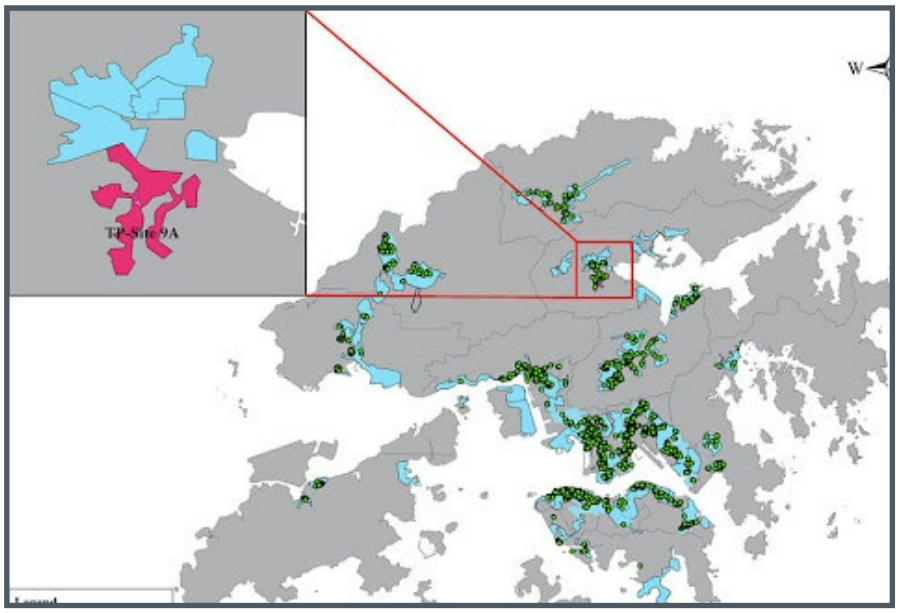


Figure 4: Wastewater surveillance in Hong Kong
This figure shows the June 21, 2021 results from the Hong Kong wastewater surveillance initiative. As shown, a positive case was detected in the "TP- Site 9A" which led to the decision for more segmented wastewater surveillance in that area.[81]

Wearables and Biometric Data: This category refers to data about the human body collected using wearable technologies (e.g. smart watches), sensors, and other devices. Throughout the pandemic, biometric data has been used to identify individuals with physical characteristics that could indicate COVID-19 such as body temperature,[82] heart rate,[83] and breathing patterns[84] as well as monitor how COVID-19 impacts the human body.[85]

---

81 Deng, Yu, Xiaoqing Xu, Xiawan Zheng, Jiahui Ding, Shuxian Li, Ho-Kwong Chui, Tsz-Kin Wong, Leo L. M. Poon, and Tong Zhang. "Use of Sewage Surveillance for COVID-19 to Guide Public Health Response: A Case Study in Hong Kong." *The Science of the Total Environment* 821 (May 15, 2022): 153250. https://doi.org/10.1016/j.scitotenv.2022.153250.

82 Ghosh, Shona. "Police in China, Dubai, and Italy Are Using These Surveillance Helmets to Scan People for COVID-19 Fever as They Walk Past and It May Be Our Future Normal." *Business Insider*, May 17, 2020. https://www.businessinsider.com/coronavirus-italy-holland-china-temperature-scanning-helmets-2020-5.

83 See, for example: Robert Koch Institute & Research on Complex Systems of Humboldt University of Berlin. "Corona Data Donation." Corona-Datenspende, 2020. https://corona-datenspende.de/science/en/.

84 See, for example: Faezipour, Miad, and Abdelshakour Abuzneid. "Smartphone-Based Self-Testing of COVID-19 Using Breathing Sounds." *Telemedicine Journal and E-Health: The Official Journal of the American Telemedicine Association* 26, no. 10 (October 2020): 1202–5. https://doi.org/10.1089/tmj.2020.0114.

85 See, for example: Sherburne, Morgan. "Wearables Can Track COVID Symptoms, Other Diseases." University of Michigan News, April 19, 2022. https://news.umich.edu/wearables-can-track-covid-symptoms-other-diseases/.

- ***Example:*** *In 2012 Kinsa Health, a medical technology company in the United States, launched an internet enabled thermometer that collects and shares body temperature data with smart phones.[86] Prior to COVID-19, Kinsa Health had been publishing online maps forecasting the spread of the flu across the country based on the data collected from the network of smartthermometers.[87] In March 2020, Kinsa Health began looking into atypical fevers from the data collected from their smart thermometers and predicted outbreaks in different areas across the U.S. ahead of the Center for Disease Control and Prevention (CDC).[88]*

  In addition to body temperature and symptoms data, Kinsa Health collects several types of personally identifiable data, including names, locations, and COVID-19 diagnoses.[89] With the user's consent, de-identified data is aggregated and published within interactive maps. Also, both identifiable and de-identified data is shared with stakeholders such as pharmacies, schools, and other organizations.[90] Workplaces and schools that partnered with Kinsa Health to track COVID-19 spread could negotiate different consent options based on their needs.[91] Kinsa Health data has been used within multiple states by groups such as policy makers, schools, and the NBA. For example, the city of New Orleans distributed Kinsa thermometers with low income households that may not have access to clinical testing to help fill gaps in clinical testing data.[92]

  Tag: *Response, Recovery*

---

86 Wilks, Jeremuy, and Pascale Davies. "How a Smart Thermometer Saw COVID-19 Coming before the Experts." euronews, March 30, 2021. https://www.euronews.com/next/2021/03/30/how-a-smart-thermometer-saw-covid-19-coming-before-the-experts.

87 Etherington, Darrell. "Kinsa's Fever Map Could Show Just How Crucial It Is to Stay Home to Stop COVID-19 Spread." *TechCrunch*, March 23, 2020. https://techcrunch.com/2020/03/23/kinsas-fever-map-could-show-just-how-crucial-it-is-to-stay-home-to-stop-covid-19-spread/?guccounter=1.

88 McNeil, Donald G. "Can Smart Thermometers Track the Spread of the Coronavirus?" The New York Times, March 18, 2020, sec. Health. https://www.nytimes.com/2020/03/18/health/coronavirus-fever-thermometers.html.

89 Kinsa Health. "Privacy Principle." Accessed September 22, 2022. https://home.kinsahealth.com/privacy.

90 Ibid.

91 "Privacy Policy." Accessed September 2, 2022. https://home.kinsahealth.com/privacy-policy.

92 Winn, Zach. "Taking the Pandemic's Temperature." *MIT Technology Review*, February 23, 2021. https://www.technologyreview.com/2021/02/23/1016861/taking-the-pandemics-temperature/.

Figure 5: Kinsa Health
The image above provides a summary of the Kinsa Health data collected through its smart thermometer and surveys in New York County, NY on October 27, 2022.[93]

Vignette: The National COVID Cohort Collaborative (N3C)

At the onset of the pandemic, the demand for COVID-19 digital patient data among policymakers, researchers, medical professionals, and other stakeholders rapidly accelerated to address many challenges, including understanding the indicators of infection, identifying the population segments most at risk, and developing targeted interventions.[94] In the United States, a decentralized healthcare system—where data is captured and analyzed differently across the country—was a barrier to combining and analyzing data about how COVID-19 patients experienced the virus.[95] Thus, there was a need for a central database that combined medical data across the country.

The National COVID Cohort Collaborative (N3C), a public-private partnership between the National Center for Advancing Translational Sciences (NCATS) (under the National Institutes of Health (NIH)) and several other organizations,

93 Kinsa HealthWeather. "New York County, New York (NY)," October 27, 2022. https://healthweather.us/map/new-york-NY/new-york-county-36061.

94 Ferguson, Cat. "It Took a Pandemic, but the US Finally Has (Some) Centralized Medical Data." *MIT Technology Review*, June 21, 2021. https://www.technologyreview.com/2021/06/21/1026590/us-covid-database-n3c-nih-privacy/.

95 National Center for Advancing Translational Sciences. "About the National COVID Cohort Collaborative," June 9, 2022. https://ncats.nih.gov/n3c/about.

was launched in September 2020 with the goal of creating a national-level database of COVID-19 patient data that could be leveraged to advance research and decision-making (known as the N3C Enclave).[96] This initiative aimed to build upon the NCATS existing mandate to the Center for Data to Health to collaborate to improve data governance and quality for machine learning purposes across the 60 Clinical and Translational Science Awards program institutions.[97] The public-private partnership now consists of over 300 participating institutions from the public, private, and academic sectors.[98]

2.1.3. Design and Planning

Digital patient data

Since the onset of the pandemic, policy makers, researchers, and other stakeholders have used COVID-19 patient data–now available digitally–to accelerate research[99] and inform decision-makers how to best support the pandemic response.[100] These efforts sought to address a range of technical and policy crisis management questions such as:

- *What population segments are most vulnerable to severe infection?*[101]
- *What are our health challenges and where is government support needed?*[102]
- *What could the future of the pandemic look like?*[103]

---

96 Ibid.

97 Haendel, Melissa. Interview re: The N3C. Zoom, July 1, 2022.

98 Sharafeldin, Noha, Benjamin Bates, Qianqian Song, Vithal Madhira, Yao Yan, Sharlene Dong, Eileen Lee, et al. "Outcomes of COVID-19 in Patients With Cancer: Report From the National COVID Cohort Collaborative (N3C)." *Journal of Clinical Oncology* 39, no. 20 (July 10, 2021): 2232–46. https://doi.org/10.1200/JCO.21.01074.

99 Robbins, Rebecca. "Hospital Records Hold Valuable Covid-19 Data. Making It Usable Is Time-Consuming Work." *STAT* (blog), May 27, 2020. https://www.statnews.com/2020/05/27/mass-general-brigham-covid19-genetics-biobank/.

100 Madhavan, Subha, Lisa Bastarache, Jeffrey S Brown, Atul J Butte, David A Dorr, Peter J Embi, Charles P Friedman, et al. "Use of Electronic Health Records to Support a Public Health Response to the COVID-19 Pandemic in the United States: A Perspective from 15 Academic Medical Centers." *Journal of the American Medical Informatics Association* 28, no. 2 (February 15, 2021): 393–401. https://doi.org/10.1093/jamia/ocaa287.

101 Sittig, Dean F., and Hardeep Singh. "COVID-19 and the Need for a National Health Information Technology Infrastructure." *JAMA* 323, no. 23 (June 16, 2020): 2373. https://doi.org/10.1001/jama.2020.7239.

102 Stokel-Walker, Chris. "How Health Data Have Been Used during Covid-19, and Whether the Changes Are Here to Stay." *BMJ* 372 (March 23, 2021): n681 https://doi.org/10.1136/bmj.n681.
Lewis, Paul, David Conn, and David Pegg. "UK Government Using Confidential Patient Data in Coronavirus Response." *The Guardian*, April 12, 2020, sec. World news. https://www.theguardian.com/world/2020/apr/12/uk-government-using-confidential-patient-data-in-coronavirus-response.

Countries such as Iceland, Sweden, and New Zealand had pre-existing national databases containing COVID-19 patients' digital records that could be tapped into for COVID-19 response efforts.[104] For example, in March 2020, Iceland leveraged its existing infrastructure to collect digital COVID-19 patient records in a centralized national database, which helped accelerate response timelines.[105] Another example was in the United Kingdom, where existing National Healthcare System (NHS) data was repurposed for the NHS COVID-19 Data Store—a database used to develop data visualizations for decision-making purposes.[106] Some countries that did not have this kind of infrastructure and needed patient data established public, private, non-profit, and academic partnerships to create a large-scale evidence base of patient level data. A more detailed explanation of how the United States leveraged partnerships to develop its infrastructure is included in Vignette 1.

Generally, digital patient data initiatives combined de-identified data from hospitals, clinics, and other health organizations within large databases or "warehouses"[107] where researchers could request access to leverage it within their own research and artificial intelligence diagnosis systems.[108] These efforts typically required emergency data sharing agreements—including data protections–and had processes for researchers to request the data for different research purposes.

Besides funding, governments have provided legislative support to encourage access to digital patient data for the public good. Several governments considered COVID-19 the type of scenario where digital patient data could be shared with third parties by default and used regulations to expedite the process. For example, the United Kingdom government

---

103 Lewis, Paul, David Conn, and David Pegg. "UK Government Using Confidential Patient Data in Coronavirus Response." *The Guardian*, April 12, 2020, sec. World news. https://www.theguardian.com/world/2020/apr/12/uk-government-using-confidential-patient-data-in-coronavirus-response.

104 Scudellari, Megan. "How Iceland Hammered COVID with Science." *Nature* 587, no. 7835 (November 25, 2020): 536–39. https://doi.org/10.1038/d41586-020-03284-3.

105 Ibid.

106 Blauer, Beth. "Universal Healthcare for Unified Data." Johns Hopkins University & Medicine Coronavirus Resource Center, May 2, 2022. https://coronavirus.jhu.edu/pandemic-data-initiative/news/universal-healthcare-for-unified-data.

107 Madhusoodanan, Jyoti. "Unlocking the Potential of Health Data to Help Research and Treatments." *Nature* 605, no. 7908 (May 3, 2022): 182–83. https://doi.org/10.1038/d41586-022-01205-0.

108 Madhusoodanan, Jyoti. "Unlocking the Potential of Health Data to Help Research and Treatments." *Nature* 605, no. 7908 (May 3, 2022): 182–83. https://doi.org/10.1038/d41586-022-01205-0.;
Patino, Marie. "The Rise of the Pandemic Dashboard." *Bloomberg*, September 25, 2021, sec. CityLab. https://www.bloomberg.com/news/features/2021-09-25/why-every-government-needs-a-covid-dashboard.;
Zoabi, Yazeed, Shira Deri-Rozov, and Noam Shomron. "Machine Learning-Based Prediction of COVID-19 Diagnosis Based on Symptoms." *Npj Digital Medicine* 4, no. 1 (January 4, 2021): 1–5. https://doi.org/10.1038/s41746-020-00372-6.

suspended the Common Law Duty of Confidentiality, typically requiring patient consent to share data,[109] to accelerate access.[110] Additionally, the Netherlands stopped asking for consent from critically ill patients to address bandwidth constraints of front-line workers and advance research.[111] These decisions were aligned with existing principles such as The Caldicott Principles[112] and with logistical requirements (receiving patient consent from all patients simply was not feasible given the volume of cases and need for data). However, privacy advocates, medical professionals, and other stakeholders raised concerns about how non-government actors might repurpose personal health records—specifically, researchers at privately run companies.[113] To address these concerns, the United Kingdom and other governments eventually offered citizens the opportunity to not have their own healthcare data shared with private companies.[114]

Several governments considered COVID-19 the type of scenario where digital patient data could be shared with third parties by default and used regulations to expedite the process.

For efforts involving digital immunity certificates, data design and planning were handled slightly differently given their different aims (i.e., controlling mobility) and different users (i.e., businesses, governments, and individuals seeking to control access to a set location). As "Digital COVID Credentials: An Implementation Process" notes, to receive a digital immunity certificate, an individual must present themselves to a certified vaccination or testing center.[115] At this location, their identity is verified with an official ID prior to testing,

109 Department of Health. "The Common Law Duty of Confidentiality." Accessed September 22, 2022. https://www.health-ni.gov.uk/articles/common-law-duty-confidentiality.

110 Kolstoe, Simon. "Coronavirus: Researchers No Longer Need Consent to Access Your Medical Records." The Conversation, May 31, 2020. http://theconversation.com/coronavirus-researchers-no-longer-need-consent-to-access-your-medical-records-138567.

111 Fleuren, Lucas M., Tariq A. Dam, Michele Tonutti, Daan P. de Bruin, Robbert C. A. Lalisang, Diederik Gommers, Olaf L. Cremer, et al. "The Dutch Data Warehouse, a Multicenter and Full-Admission Electronic Health Records Database for Critically Ill COVID-19 Patients." *Critical Care* 25, no. 1 (August 23, 2021): 304. https://doi.org/10.1186/s13054-021-03733-z.

112 United Kingdom Government. "The Caldicott Principles," December 8, 2020. https://www.gov.uk/government/publications/the-caldicott-principles.

113 Lovell, Tammy. "Privacy Fears over NHS Plans to Share GP Medical Records with Third Parties." Healthcare IT News, June 2, 2021. https://www.healthcareitnews.com/news/emea/privacy-fears-over-nhs-plans-share-gp-medical-records-third-parties.

114 Murgia, Madhumita, and Max Harlow. "NHS Shares English Hospital Data with Dozens of Companies." Financial Times, July 27, 2021. https://www.ft.com/content/6f9f6f1f-e2d1-4646-b5ec-7d704e45149e.

115 Nehme, Mayssam, Laurent Kaiser, Philippe Gillet, Philippe Thevoz, Silvia Stringhini, and Idris Guessous. "Digital COVID Credentials: An Implementation Process." *Frontiers in Digital Health* 3 (2021). https://www.frontiersin.org/articles/10.3389/fdgth.2021.594124.

vaccination, or determination of recovery. Upon completing their appointment, the individual is either able to self-report their status as possessing immunity (either by taking a picture of their vaccination record or by simply indicating they are vaccinated) or are able to connect with a centralized database run by a relevant healthcare authority (e.g. a national, regional or local agency).[116] Once connected and verified, the person receives a digital proof of their status via an app, usually accompanied by a QR code and other content to limit the possibility of fraud.

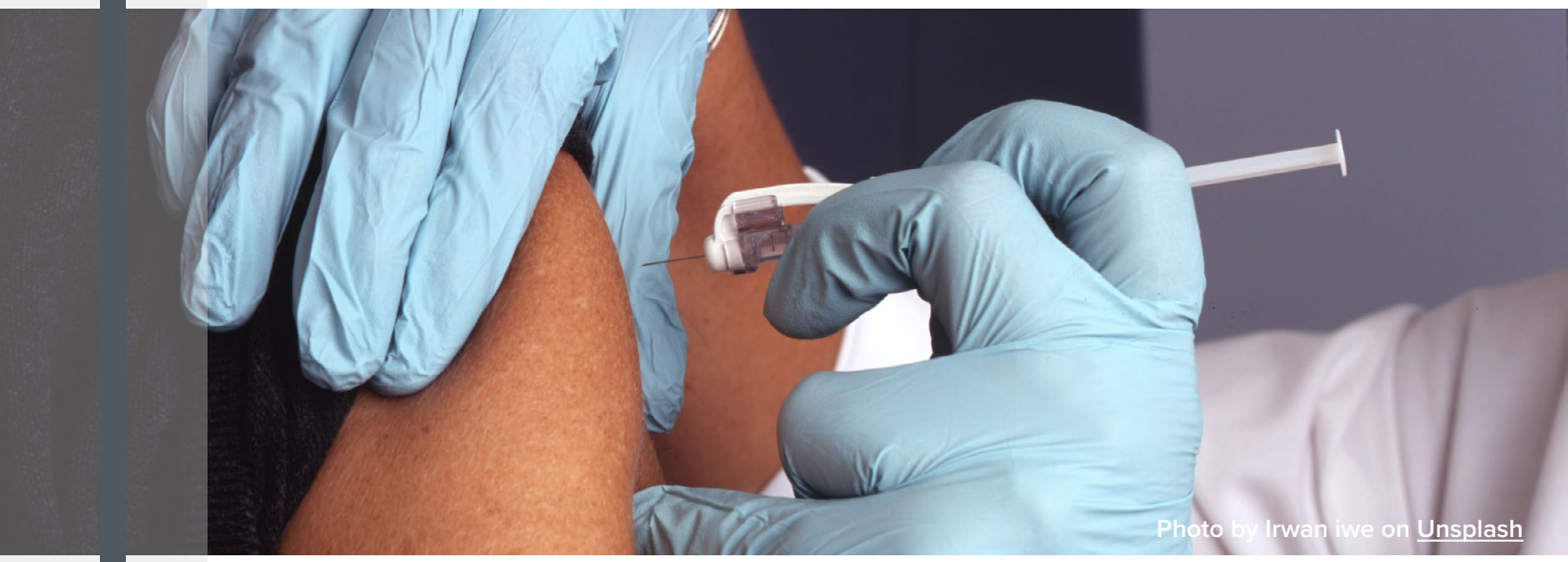


Photo by Irwan iwe on Unsplash

116 Shachar, Carmel, and Chloe Reichel. "Fake Vaccine Cards and the Challenges of Decentralized Health Data." Bill of Health, April 27, 2021. http://blog.petrieflom.law.harvard.edu/2021/04/27/vaccine-data-fake-cards/.

Platform-facilitated symptom data

The examples studied in this report and the #Data4COVID19 Data Collaborative Repository indicate that platform-facilitated symptom data initiatives rapidly expanded during the initial months of the pandemic, involving new instrumentation methods to gather non-traditional health data at scale, such as aggregated data from Facebook surveys, Google Search Trends, surveys on mobile Apps, and targeted online surveys advertised on news platforms and other privately run websites.

COVID-19 symptom data collection efforts using surveys on mobile Apps and websites tended to be driven by academic institutions or governments in collaboration with policymakers, private companies, and other stakeholders.[117] These initiatives were designed to track the health impacts of COVID-19, including the indicators of infection and mental health implications in the short- and long-term.[118] Using ZOE as a model, such initiatives tended to require a data sharing agreement, user consent for repurposing their personal health data, and a data anonymization process.[119] In instances like the FluTracking platform, flu monitoring mobile App in Australia and New Zealand, and Influenzanet in Europe,[120] the initiative was broadened to include additional COVID-19 metrics.[121]

Initiatives such as Facebook's COVID-19 Symptom Survey conducted micro-agreements with academic partners. These agreements were often provided by the academic institutions during the data access request process.[122] The final aggregated outputs were typically published on platform websites in the form of visualizations and dashboards.[123] The intention of these efforts was for the data to be combined with other traditional data sources, such

---

117 "Coronavirus Surveys, Apps to Track COVID-19." Harvard T.H. Chan School of Public Health, April 14, 2020. https://www.hsph.harvard.edu/coronavirus/covid-19-response-public-health-in-action/surveys-apps-to-track-covid-19/.

118 Ibid.

119 "Frequently Asked Questions - COVID Symptom Study." Accessed August 28, 2022. https://health-study.joinzoe.com/faq.

120 InfluenzaNet Analytics. "Project Information." Accessed September 29, 2022. https://influenzanet.info/#page/info.

121 Newcastle Herald. "FluTracking Needs Community Support to Help Track Any Potential Coronavirus Outbreaks," February 24, 2020. https://www.newcastleherald.com.au/story/6646474/flutracking-needs-community-support-to-help-track-any-potential-coronavirus-outbreaks/.

122 University of Waterloo. "Facebook - COVID-19 Symptom Survey – Request for Data Access," May 31, 2020. https://uwaterloo.ca/research/facebook-covid-19-symptom-survey-request-data-access.

123 Meta. "COVID-19 RESPONSE. We Have Tools That Can Help Organizations Respond to the COVID-19 Pandemic." Data for Good, 2022. https://dataforgood.facebook.com/dfg/covid-19.

as clinical test results.[124] Given the scale of these initiatives, the data collected was typically analyzed using machine learning[125]—"a subfield of artificial intelligence, which is broadly defined as the capability of a machine to imitate intelligent human behavior"[126]—and anonymized.

COVIDCast,[127] an initiative led by the Delphi Group (a research group at Carnegie Mellon University focused on epidemiological forecasting) and other multidisciplinary partners, launched in March 2020 with the objective of forecasting COVID-19 hotspots.[128] This initiative was made possible through collaboration[129]—leveraging Carnegie Mellon's existing machine learning capacity along with the resources of other organizations, including the US Centers for Disease Control and Prevention (CDC).[130] This effort combined Facebook survey data about patient symptoms with Google search trends, insurance claim data, antigen test data (clinical testing), and mobility data. Carnegie Mellon researchers combined the data sources and developed dashboards about indicators of infection and the spread of the virus.[131] This initiative was made publicly accessible using the COVIDcast API.[132] COVIDCast became "the nation's largest public repository of diverse, real-time indicators of COVID-19 activity"[133] and continues to publish data.

---

124 Kelly, Heather. "Those Symptoms You Googled Could Help Researchers Better Understand Coronavirus." *Washington Post,* September 2, 2020. https://www.washingtonpost.com/technology/2020/09/02/google-covid-data/.

125 Mackey, Tim, Vidya Purushothaman, Jiawei Li, Neal Shah, Matthew Nali, Cortni Bardier, Bryan Liang, Mingxiang Cai, and Raphael Cuomo. "Machine Learning to Detect Self-Reporting of Symptoms, Testing Access, and Recovery Associated With COVID-19 on Twitter: Retrospective Big Data Infoveillance Study." *JMIR Public Health and Surveillance* 6, no. 2 (June 8, 2020): e19509. https://doi.org/10.2196/19509; Vermicelli, Silvia, Livio Cricelli, and Michele Grimaldi. "How Can Crowdsourcing Help Tackle the COVID-19 Pandemic? An Explorative Overview of Innovative Collaborative Practices." *R&D Management* 51, no. 2 (March 2021): 183–94. https://doi.org/10.1111/radm.12443.

126 Brown, Sara. "Machine Learning, Explained." MIT Management - Sloan School, April 21, 2021. https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained.

127 The GovLab. "COVIDcast Real-Time COVID-19 Indicators," 2021. https://list.data4covid19.org/projects/covid19-realtime-indicators.

128 Carnegie Mellon University. "Delphi Research Group." Accessed September 22, 2022. https://delphi.cmu.edu/about/.

129 Aupperlee, Aaron. "CMU's COVIDcast Offers Lessons Learned for the Future of Pandemic Forecasting." Carnegie Mellon University, February 2, 2022. http://www.cmu.edu/news/stories/archives/2022/february/covidcast-lessons.html.

130 Spice, Byron. "Delphi Enhances COVIDcast With Change Healthcare Claims Data." Carnegie Mellon University, December 8, 2020. http://www.cmu.edu/news/stories/archives/2020/december/covidcast.html.

131 Carnegie Mellon University. "COVIDcast Dashboard," September 20, 2022. https://delphi.cmu.edu/covidcast/.

132 Delphi Epidata API. "COVIDcast Epidata API," 2020. https://cmu-delphi.github.io/delphi-epidata/api/covidcast.html.

133 Carnegie Mellon University. "Delphi Research Group." Accessed September 22, 2022. https://delphi.cmu.edu/about/.

Wastewater data

Wastewater initiatives designed to detect the presence of COVID-19 and quantify infection levels within different areas address questions such as:[134]

- *Is there COVID-19 in this jurisdiction?*
- *Is there a new COVID-19 variant spreading?*
- *Do COVID-19 clinical testing data accurately reflect changing infection levels in the community?*
- *Did community activities (such as large sporting events, a return to in-person school/university classes, or an influx of tourists) increase COVID-19 levels?*
- *Where are there hotspots of infection?*
- *What buildings, blocks, communities, cities, or regions require additional clinical testing, vaccination, and/or policy interventions?*

Some COVID-19 wastewater surveillance initiatives were adapted from previous wastewater monitoring for drug use—reflecting a shift in prioritized targets rather than development of a new program.[135] However, in many communities, the pandemic prompted the first time wastewater surveillance was undertaken, thus requiring new infrastructure and programs.[136] Initiatives often began as pilot programs led by academic and independent researchers manually collecting wastewater at different sites and testing it at labs and other facilities.[137] These initiatives involved non-personal data only, thus minimizing personal data protection requirements.[138] Often, these initiatives were initially funded by government grants made to academic institutions[139] and aimed to set up the processes and infrastructure, or data collection sites and automate testing,

134 Keshaviah, Aparna. Interview re: wastewater data. Zoom, July 11, 2022.

135 Margetts, Miranda, Aparna Keshaviah, Xindi C. Hu, Victoria Troeger, Jordan Sykes, Nicholas Bishop, Tammy Jones-Lepp, Marisa Henry, and Deborah E. Keil. "Using Wastewater-Based Epidemiology with Local Indicators of Opioid and Illicit Drug Use to Overcome Data Gaps in Montana." Preprint. Public and Global Health, April 23, 2020. https://doi.org/10.1101/2020.04.18.20064113.

136 Ladyzhets, Betsy. "The National Fight Against COVID-19 Isn't Ready To Go To The Sewers." *FiveThirtyEight*, April 20, 2022. https://fivethirtyeight.com/features/the-national-fight-against-covid-19-isnt-ready-to-go-to-the-sewers/.

137 Barber, Gregory. "One Way to Potentially Track Covid-19? Sewage Surveillance." *Wired*, April 7, 2020. https://www.wired.com/story/one-way-to-potentially-track-covid-19-sewage-surveillance/.

138 John Hopkins - Bloomberg School of Public Health. "How COVID-19 Created a 'Watershed' Moment for Wastewater Surveillance," May 13, 2022. https://publichealth.jhu.edu/2022/how-covid-19-created-a-watershed-moment-for-wastewater-surveillance.

139 Keshaviah, Aparna. Interview re: wastewater data. Zoom, July 11, 2022.

before expanding to other areas.[140] As the pandemic progressed, many communities used COVID-19 relief funds or other federal government grants to support wastewater monitoring programs.[141]

Aparna Keshaviah, Principal Researcher at Mathematica explains: "One of the strengths and challenges of wastewater surveillance is it requires a lot of collaboration in an inter-agency fashion."[142] The World Bank Group identified four areas of expertise required in the design of wastewater initiatives: city personnel to collect and transport data from municipalities, regional personnel to collect data from larger wastewater and sewage sites, data and IT specialists to analyze and disseminate results, and public health staff (primarily at the local level) to implement and share findings for the public health response.[143] These collaborations often arose as a result of efforts (i.e. meetings, webinars, and published articles) led by wastewater experts to raise awareness of this data source and its promise to fill gaps in traditional data sources, based on evidence and lessons learned from previous applications to address the opioid crisis[144] and the polio epidemic.[145]

"One of the strengths and challenges of wastewater surveillance is it requires a lot of collaboration in an inter-agency fashion."

---

140 Manuel, Doug, Carlo Alberto Amadei, Jonathon R. Campbell, Jean-Martin Brault, and Jeremy Veillard. "Strengthening Public Health Surveillance Through Wastewater Testing: An Essential Investment for the COVID-19 Pandemic and Future Health Threats." Washington, DC: *World Bank*, January 19, 2022. https://openknowledge.worldbank.org/handle/10986/36852.

141 Keshaviah, Aparna, Ruchir N Karmali, Divya Vohra, Xindi C Hu, and Megan B Diamond. "The Role of Wastewater Data in Pandemic Management." Washington, DC: Mathematica, April 2022.

142 Keshaviah, Aparna. Interview re: wastewater data. Zoom, July 11, 2022.

143 Manuel, Doug, Carlo Alberto Amadei, Jonathon R. Campbell, Jean-Martin Brault, and Jeremy Veillard. "Strengthening Public Health Surveillance Through Wastewater Testing: An Essential Investment for the COVID-19 Pandemic and Future Health Threats." Washington, DC: *World Bank*, January 19, 2022. https://openknowledge.worldbank.org/handle/10986/36852.

144 Keshaviah, Aparna, and Miranda Margetts. "Want to Test More People for Covid-19? Look At the Sewage." Route Fifty, May 6, 2020. https://www.route-fifty.com/health-human-services/2020/05/test-covid-19-sewage/165177/.; Keshaviah, Aparna, Xindi C. Hu, and Marisa Henry. "Developing a Flexible National Wastewater Surveillance System for COVID-19 and Beyond." *Environmental Health Perspectives* 129, no. 4 (April 2021): 045002. https://doi.org/10.1289/EHP8572.

145 Hovi, T., L. M. Shulman, H. Van Der Avoort, J. Deshpande, M. Roivainen, and E. M. De Gourville. "Role of Environmental Poliovirus Surveillance in Global Polio Eradication and Beyond." *Epidemiology & Infection* 140, no. 1 (January 2012): 1–13. https://doi.org/10.1017/S095026881000316X.; Brouwer, Andrew F., Joseph N. S. Eisenberg, Connor D. Pomeroy, Lester M. Shulman, Musa Hindiyeh, Yossi Ma
nor, Itamar Grotto, James S. Koopman, and Marisa C. Eisenberg. "Epidemiology of the Silent Polio Outbreak in Rahat, Israel, Based on Modeling of Environmental Surveillance Data." *Proceedings of the National Academy of Sciences* 115, no. 45 (November 6, 2018). https://doi.org/10.1073/pnas.1808798115.

Wearables and biometrics

Data about the human body collected using wearable technologies and other devices provided rapid, frequent monitoring to quickly identify, trace, and isolate cases of the virus.[146] Such efforts primarily emanated from private companies that had developed the infrastructure for data collection–and in some cases, were using those technologies to detect other diseases.[147]

One such example is Corona Datenspende (The Corona Data Donation Project), a partnership between the Robert Koch Institute, the data research company Thryve, and other partners in Germany. The partnership was formed after Thryve shared its research with the Robert Koch Institute demonstrating how data from wearables could be used in the COVID-19 response.[148] This initiative consists of a mobile app, developed over the course of one month,[149] that collects data from fitness-wristbands or smartwatches on physical activity and pulse and combines it with survey data and other data sources.[150] Using the app, users are asked to provide information such as their postal code and physical characteristics.[151] The data collected is pseudonymized–not personally identifiable—and stored on a cloud based server.[152] The data is linked to geodata at the postal code level and analyzed using artificial intelligence (AI) methods to identify possible Covid-19 symptoms. The initiative continues to operate and has expanded to study the impact of vaccines.[153]

---

146 Gadaleta, Matteo, Jennifer M. Radin, Katie Baca-Motes, Edward Ramos, Vik Kheterpal, Eric J. Topol, Steven R. Steinhubl, and Sensors and Explainable Machine Learning Algorithms." *NPJ Digital Medicine* 4 (December 8, 2021): 166. https://doi.org/10.1038/s41746-021-00533-1.

147 Winn, Zach. "Real-Time Data for a Better Response to Disease Outbreaks." MIT News, August 21, 2020. https://news.mit.edu/2020/kinsa-health-0821.

148 Busvine, Douglas. "Germany Launches Smartwatch App to Monitor Coronavirus Spread." Financial Post, April 7, 2020. https://financialpost.com/pmn/business-pmn/germany-launches-smartwatch-app-to-monitor-coronavirus-spread.

149 Institute of Medicine (US) Roundtable on Value & Science-Driven Health Care. *Clinical Data as the Basic Staple of Health Learning: Creating and Protecting a Public Good: Workshop Summary. Clinical Data as the Basic Staple of Health Learning: Creating and Protecting a Public Good: Workshop Summary*. Summary. Washington, DC: National Academies Press (US), 2010. https://www.ncbi.nlm.nih.gov/books/NBK54290/.

150 The GovLab. "Living Repository for Data Collaboratives Seeking to Address the Spread of COVID-19." Data4COVID19. Accessed September 13, 2022. https://list.data4covid19.org/.

151 Robert Koch Institut. "Corona Data Donation App 2.0." Accessed September 22, 2022. https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Corona-Datenspende-allgemein.html.

152 Corona-Datenspende. "FAQ." Accessed August 2, 2022. https://corona-datenspende.de/faq/.

153 Wiedermann, Marc, Annika H. Rose, Benjamin F. Maier, Jakob J. Kolb, David Hinrichs, and Dirk Brockmann. "Evidence for Positive Long- and Short-Term Effects of Vaccinations against COVID-19 in Wearable Sensor Metrics -- Insights from the German Corona Data Donation Project." *arXiv*, April 6, 2022. https://doi.org/10.48550/arXiv.2204.02846.

Vignette: The National COVID Cohort Collaborative (N3C)

The coordination of participating research networks was a fundamental component of the design of The N3C.[154] In the United States, research networks tend to follow a distributed structure; meaning that they do not share patient-level data and their data quality and management practices are executed locally.[155] While the distributed network has benefits from a privacy perspective, it creates challenges in applying machine learning methods.[156] The N3C team led negotiations with key distributed research networks (e.g. Odyssey Research,[157] Trinetx[158]) to develop a framework for them all to collaborate and document their data models. From there, The N3C team in collaboration with government partners, could develop a data governance structure and processes for data holders to share and use digital patient data on The N3C Enclave.[159]

The N3C data governance processes focus on data quality, security, and attribution. These processes include: A partner signs off on the NCATS standardized "Data Transfer Agreement" indicating they can legally share the data;[160] and a partner employee confirms that they will act in accordance with the N3C data use policies for the type of data they are working with.[161] Given that partner organizations are not sharing information that could directly identify patients, they are not required to ask for patient consent

154 Haendel, Melissa A, Christopher G Chute, Tellen D Bennett, David A Eichmann, Justin Guinney, Warren A Kibbe, Philip R O Payne, et al. "The National COVID Cohort Collaborative (N3C): Rationale, Design, Infrastructure, and Deployment." *Journal of the American Medical Informatics Association* 28, no. 3 (March 1, 2021): 427–43. https://doi.org/10.1093/jamia/ocaa196.

155 Ibid.

156 Haendel, Melissa. Interview re: The N3C. Zoom, July 1, 2022.

157 Odyssey. "Home." Accessed September 1, 2022. https://www.odysseyresearch.org.

158 TriNetX. "Home." Accessed August 29, 2022. https://trinetx.com/.

159 Haendel, Melissa. Interview re: The N3C. Zoom, July 1, 2022.

160 National Center for Advancing Translational Sciences. "NIH COVID-19 Data Warehouse Data Transfer Agreement ('Agreement')," July 4, 2021, 7.

161 Haendel, Melissa. Interview re: The N3C. Zoom, July 1, 2022.

ahead of sharing their data.[162] Researchers interested in accessing the data must submit a "data access request" for The N3C to review, indicating how they plan to use the data and confirming the approval from an Institutional Review Board for data that has dates or geocodes.[163]

Dr. Melissa Haendel, Co-Lead of The N3C explains: "I think one of the most important things has been the fact that not only does the system provide high security and provenance, but because of that it also provides for robust attribution. We can track every single person who's contributed in any way. And we have really strong policies for how to attribute contributions to any manuscript that goes out."[164]

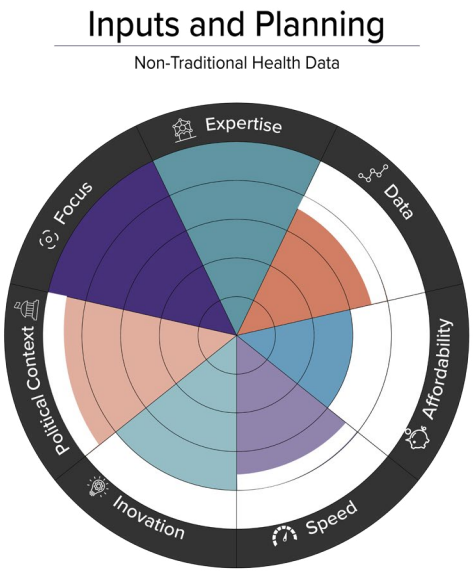


Figure 6. Assessment: Non-Traditional Health Data, Inputs and Planning
The figure above shows our assessment of the design of The N3C using the assessment criteria.

162 National Center for Advancing Translational Sciences. "About the National COVID Cohort Collaborative," February 1, 2022. https://ncats.nih.gov/n3c/about.

163 Haendel, Melissa. Interview re: The N3C. Zoom, July 1, 2022.

164 Ibid.

2.1.4. Implementation and Outcomes

Digital patient data

While in many cases, researchers volunteered to participate in the pandemic response,[165] hosts of digital patient data initiatives developed challenges, competitions, or hackathons to incentivize the use of their data within research projects. This work sought to make use of changing attitudes toward data sharing.

Recent studies suggest that patients became more comfortable sharing their health records during COVID-19 compared to before the pandemic and some patients were more willing to share their health records for public efforts than other data sources such as social media data.[166] However, this growing trust in providing external access to digital patient data did not always transfer to the private sector. For example, during COVID-19 the United Kingdom increased private sector access to digital patient data[167] and developed new collaborations with tech companies such as Google and Palantir to understand healthcare system capacity and provide targeted aid.[168]

Advocacy groups criticized these initiatives for not providing information about how third parties would protect their personal data, what was involved in the agreement and the

---

165 Park, Jay J. H., Robin Mogg, Gerald E. Smith, Etheldreda Nakimuli-Mpungu, Fyezah Jehan, Craig R. Rayner, Jeanine Condo, et al. "How COVID-19 Has Fundamentally Changed Clinical Research in Global Health." *The Lancet Global Health* 9, no. 5 (May 1, 2021): e711–20. https://doi.org/10.1016/S2214-109X(20)30542-8.

166 Tosoni, Sarah, Indu Voruganti, Katherine Lajkosz, Shahbano Mustafa, Anne Phillips, S. Joseph Kim, Rebecca K. S. Wong, et al. "Patient Consent Preferences on Sharing Personal Health Information during the COVID-19 Pandemic: 'The More Informed We Are, the More Likely We Are to Help.'" *BMC Medical Ethics* 23, no. 1 (December 2022): 53. https://doi.org/10.1186/s12910-022-00790-z.;
Gerdon, Frederic, Helen Nissenbaum, Ruben L. Bach, Frauke Kreuter, and Stefan Zins. "Individual Acceptance of Using Health Data for Private and Public Benefit: Changes During the COVID-19 Pandemic." *Harvard Data Science Review*, no. Special Issue 1 (April 6, 2021). https://doi.org/10.1162/99608f92.edf2fc97.;
Grande, David, Nandita Mitra, Raghuram Iyengar, Raina M. Merchant, David A. Asch, Meghana Sharma, and Carolyn C. Cannuscio. "Consumer Willingness to Share Personal Digital Information for Health-Related Uses." *JAMA Network Open* 5, no. 1 (January 24, 2022): e2144787. https://doi.org/10.1001/jamanetworkopen.2021.44787.;
PEW. "Most Americans Want to Share and Access More Digital Health Data," July 27, 2021. https://pew.org/3B9DJXq.; David Betts, Leslie Korenda, and Shane Giuliani. "Are Consumers Already Living the Future of Health?" Deloitte Insights, August 13, 2020. https://www2.deloitte.com/us/en/insights/industry/health-care/consumer-health-trends.html.

167 Murgia, Madhumita, and Max Harlow. "NHS Shares English Hospital Data with Dozens of Companies." Financial Times, July 27, 2021. https://www.ft.com/content/6f9f6f1f-e2d1-4646-b5ec-7d704e45149e.

168 Ibid.

possible allowance to extend beyond COVID-19.[169] These concerns were heightened following reports of private organizations using digital patient data analyses (including government data) for profit and that companies planned to use the data beyond the initial purpose.[170]

For digital immunity certificates, the purpose has been to allow individuals to more easily prove they have been vaccinated or otherwise immunized to prove their status for travel purposes. As Harvard's Edmond J. Safra Center for Ethics' notes, it is also intended to help workplaces, medical facilities, airlines, businesses, and others to more easily verify an individual's COVID-19 status before granting access to a facility.[171] These certificates have been used by the European Union to allow visitors and residents to more easily allow individuals to travel within their borders.[172] It has also been developed by private associations such as the International Air Transport Association[173] and International Chamber of Commerce[174] to manage access to the locations that they and their members manage.

These platforms have been used in many contexts around the world, though they have become the source of political controversy in some contexts and the subject of bans.[175] The reasons for this opposition vary, with some expressing concern over privacy, adherence with existing regulations or individual agency, though many experts do not believe these objections have merit in most contexts.[176] In "Taking stock of COVID-19 health status

---

169 Kathleen Hall. "The Tech Firms Getting Their Hands on NHS Patient Data to Fight Coronavirus." The Bureau of Investigative Journalism, May 7, 2020. https://www.thebureauinvestigates.com/stories/2020-05-07/the-tech-firms-getting-their-hands-on-nhs-patient-data-to-fight-coronavirus.

170 Murgia, Madhumita, and Max Harlow. "NHS Shares English Hospital Data with Dozens of Companies." Financial Times, July 27, 2021. https://www.ft.com/content/6f9f6f1f-e2d1-4646-b5ec-7d704e45149e.

171 "Immunity Certificates: If We Must Have Them, We Must Do It Right." White Paper. COVID-19 Rapid Response Impact Initiative. Edmond J. Safra Center for Ethics, April 20, 2020. https://ethics.harvard.edu/files/center-for-ethics/files/12immunitycertificates.pdf.

172 Montanari Vergallo, Gianluca, Simona Zaami, Francesca Negro, Pietro Brunetti, Alessandro Del Rio, and Enrico Marinelli. "Does the EU COVID Digital Certificate Strike a Reasonable Balance between Mobility Needs and Public Health?" *Medicina* 57, no. 10 (October 9, 2021): 1077. https://doi.org/10.3390/medicina57101077.

173 IATA. "IATA Travel Pass Initiative." Accessed September 22, 2022. https://www.iata.org/en/programs/passenger/travel-pass/.

174 ICC - International Chamber of Commerce. "ICC mobilises network to help end pandemic in the Caribbean," April 30, 2021. https://iccwbo.org/media-wall/news-speeches/icc-mobilises-network-to-help-end-pandemic-in-the-caribbean/.

175 Elliott Davis Jr. "Which States Have Banned Vaccine Passports?" US News & World Report, June 1, 2021. //www.usnews.com/news/best-states/articles/which-states-have-banned-vaccine-passports.

176 Schafer, Arthur. "Lots of Opposition but 'a Striking Absence of Good Arguments' against Vaccine Passports, Says Ethicist." *CBC*. August 1, 2021, sec. Opinion. https://www.cbc.ca/news/canada/manitoba/opinion-national-covid-19-vaccine-passports-arthur-schafer-1.6122386.

certificates: Legal implications for data privacy and human rights" Dr. Ana Beduschi explains: "technical solutions for the verification of COVID-19 health status do not suffice on their own. Because technologies do not evolve in a legal vacuum, the existing laws and regulations must be respected. The risks of implementing such technologies must be anticipated and mitigated as much as possible before their deployment on a global scale."[177]

Platform-facilitated symptom data

Three common factors tended to impact the effectiveness of platform-facilitated symptom data initiatives: bias, timing, and agility. First, platform-facilitated symptom data initiatives often required access to a smart phone and internet. They also required users to opt in by downloading an app, which often limited participation.[178] This fact is particularly challenging for initiatives that aim to track symptoms longitudinally–requiring daily survey entries over several months as it requires both a larger user base and those same users to remain dedicated over a long period of time to voluntarily supply data. This further raises the risk of collecting data only representative of specific population segments.

This issue of representativeness has been a theme in the literature for some time. In "Unrepresentative big surveys significantly overestimated US vaccine uptake", Valerie C. Bradley, et al. explain that despite collecting large volumes of data, Delphi-Facebook's COVID-19 Symptom Tracker overestimated the national vaccine count by 17% in May 2021.[179] The authors explain that surveys collecting big data often do not adjust the margins for error for a larger sample size and do not reflect the sample pool of the Census. The Delphi-Facebook survey collected data from active Facebook users only. "Delphi–Facebook's weighting scheme

Surveys collecting big data often do not adjust the margins for error for a larger sample size and do not reflect the sample pool of the Census.

177 Beduschi, Ana. "Taking Stock of COVID-19 Health Status Certificates: Legal Implications for Data Privacy and Human Rights." *Big Data & Society* 9, no. 1 (January 1, 2022): 20539517211069300. https://doi.org/10.1177/20539517211069300.

178 Spector, Tim, and Andrew Chan. "Coronavirus: Research Reveals Way to Predict Infection – without a Test." *The Conversation*, May 11, 2020. http://theconversation.com/coronavirus-research-reveals-way-to-predict-infection-without-a-test-138284.

179 Bradley, Valerie C., Shiro Kuriwaki, Michael Isakov, Dino Sejdinovic, Xiao-Li Meng, and Seth Flaxman. "Unrepresentative Big Surveys Significantly Overestimated US Vaccine Uptake." *Nature* 600, no. 7890 (December 2021): 695–700. https://doi.org/10.1038/s41586-021-04198-4.

does not adjust for race/ethnicity, and hence their weighted sample still overrepresents white adults by 8 percentage points, and underrepresents the proportions of Black and Asian individuals by around 50% of their size in the population."[180] The authors demonstrate that these inaccuracies contributed to its overestimation.[181]

Second, initiatives that aimed to predict the spread of the virus and future outbreaks were often not implemented and launched until the virus was already present and had spread. For example, Fluwatchers, a weekly online survey program supported by the Public Health Agency of Canada, had been collecting self-reported symptoms for its "national influenza surveillance program" since 2015.[182] In April 2020 Fluwatchers transitioned from monitoring symptoms from October to May to the entire year.[183] This decision was made four months following the first case of COVID-19 in Canada and after the virus had already spread.[184]

Boston Children's Hospital's initiative, HealthMap, was launched in 2006 to track where diseases are spreading around the world. The project combines several data sources including search trends and social media data using artificial intelligence and publishes the results on their online dashboard.[185] HealthMap tracked a news article on December 30, 2019 reporting pneumonia-like cases in Wuhan and included it on its dashboard. This reporting was made possible because the infrastructure was already in place to track disease on an ongoing basis.[186]

Lastly, some initiatives did not update their research questions, indicators, and metrics to reflect the evolving state of the virus—hindering the accuracy and ability to scale. COOPERA, a crowdsourcing application for COVID-19 symptoms in Japan, asked for known symptoms at the onset of the pandemic and did not collect data about emerging symptoms

---

180 Ibid.

181 Ibid.

182 Public Health Agency of Canada. "FluWatchers." Education and awareness. Government of Canada, June 28, 2017. https://www.canada.ca/en/public-health/services/diseases/flu-influenza/fluwatcher.html.

183 Lee, Liza, Shamir Mukhi, and Christina Bancej. "Crowdsourced Disease Surveillance Success Story: The FluWatchers Program." *Canada Communicable Disease Report* 47, no. 9 (September 10, 2021): 354–56. https://doi.org/10.14745/ccdr.v47i09a01.

184 Canadian Institute for Health Information. "COVID-19 Intervention Timeline in Canada," June 9, 2022. https://www.cihi.ca/en/covid-19-intervention-timeline-in-canada.

185 Cho, Adrian. "Artificial Intelligence Systems Aim to Sniff out Signs of COVID-19 Outbreaks." Science, May 12, 2020. https://www.science.org/content/article/artificial-intelligence-systems-aim-sniff-out-signs-covid-19-outbreak.

186 Ibid.

such as "loss in taste (ageusia) and/or smell (anosmia);"[187] thus providing limited data to meaningfully supplement traditional data as the pandemic progressed.

More broadly, platform-facilitated symptoms data is one of many data sources that was used to detect and predict COVID-19 outbreaks. In "An early warning approach to monitor COVID-19 activity with multiple digital traces in near real time" Kogan et al. contend that this data source was an effective early warning signal in combination with other data sources such as mobility data and biometrics in the United States.[188] The authors explain: "From 1 March to 30 September 2020 [...] We observe that increases in digital data stream [or data sources] activity anticipate increases in confirmed cases and deaths by 2 to 3 weeks."[189] The authors recommend combining these data sources in forecasting future outbreaks.[190]

Photo by Roin Worrall on Unsplash

Wastewater data

To date, wastewater surveillance has predominantly been implemented in highly developed economies where governments and academic institutions could provide

187 Desjardins, Michael R. "Syndromic Surveillance of COVID-19 Using Crowdsourced Data." *The Lancet Regional Health – Western Pacific* 4 (November 1, 2020). https://doi.org/10.1016/j.lanwpc.2020.100024.

188 Kogan, Nicole E., Leonardo Clemente, Parker Liautaud, Justin Kaashoek, Nicholas B. Link, Andre T. Nguyen, Fred S. Lu, et al. "An Early Warning Approach to Monitor COVID-19 Activity with Multiple Digital Traces in near Real Time." *Science Advances* 7, no. 10 (March 5, 2021): eabd6989. https://doi.org/10.1126/sciadv.abd6989.

189 Ibid.

190 Ibid.

funding,[191] though some low- and middle-income countries expanded wastewater surveillance systems initially established to detect poliovirus and antimicrobial resistance to detect SARS-CoV-2.[192] Given that wastewater surveillance often requires access to maintained public infrastructure (e.g. municipal wastewater treatment systems), many efforts during COVID-19 have tended to be led or at least partially funded by governments.[193] Having government ownership facilitates the data to be directly inputted in decision-making processes–which during COVID-19 often took the form of dashboards or aggregated databases.[194]

For example, the Netherlands includes select wastewater data supplied by the National Institute for Public Health and the Environment[195] collected across the country in its Coronavirus Dashboard which also holds visualizations of traditional data (e.g. clinical testing data). Within this dashboard, wastewater data is classified as an "early indicator"[196] of infection. Nonetheless, this process of combining traditional data with wastewater data for decision-making purposes has had its challenges.

Having government ownership facilitates the data to be directly inputted in decision-making processes–which during COVID-19 often took the form of dashboards or aggregated databases.

In September 2020, The United States Center for Disease Control and Prevention, US Department of Health and Human Services and other federal agencies developed the National Wastewater Surveillance System (NWSS), a centralized database of wastewater data collected across the country.[197] One of the many reasons why wastewater surveillance

---

191 Naughton, Colleen C., Fernando A. Roman, Ana Grace F. Alvarado, Arianna Q. Tariqi, Matthew A. Deeming, Kyle Bibby, Aaron Bivins, et al. "Show Us the Data: Global COVID-19 Wastewater Monitoring Efforts, Equity, and Gaps." Preprint. *Public and Global Health*, March 17, 2021. https://doi.org/10.1101/2021.03.14.21253564.

192 Street, Renée, Shirley Malema, Nomfundo Mahlangeni, and Angela Mathee. "Wastewater Surveillance for Covid-19: An African Perspective." *Science of The Total Environment* 743 (November 15, 2020): 140719. https://doi.org/10.1016/j.scitotenv.2020.140719.;
Nelson, Bryn. "What Poo Tells Us: Wastewater Surveillance Comes of Age amid Covid, Monkeypox, and Polio." *BMJ* 378 (July 29, 2022): o1869. https://doi.org/10.1136/bmj.o1869.

193 Keshaviah, Aparna. Interview re: wastewater data. Zoom, July 11, 2022.

194 Example: Centers for Disease Control and Prevention. "COVID Data Tracker," September 21, 2022. https://covid.cdc.gov/covid-data-tracker.

195 Rijksoverheid. "Coronadashboard data explained." Accessed September 22, 2022. https://coronadashboard.government.nl/verantwoording.

196 Ibid.

197 Centers for Disease Control and Prevention. "National Wastewater Surveillance System (NWSS)," May 16, 2022. https://www.cdc.gov/healthywater/surveillance/wastewater-surveillance/wastewater-surveillance.html.

has expanded in the United States is that "80 percent of U.S. households are served by municipal wastewater collection systems"[198] as opposed to septic systems which are not included in wastewater monitoring. Using the NWSS, wastewater is collected and tested at the local, state, tribal, and territorial levels and then shared with national counterparts using the NWSS Data Collation and Integration for Public Health Event Response (DCIPHER) portal.[199] The NWSS DCIPHER then analyzes the data[200] and combines it with epidemiological data to be used in federal COVID-19 response efforts.[201]

This process has been challenging due to wastewater data being collected using different methodologies across the country, and because metrics are lacking to reliably flag an alert when viral concentrations first rise about background levels.[202] Parts of the data are then disseminated within a public facing dashboard–the COVID Data Tracker.[203] The process of combining wastewater data with traditional data sources within dashboards has had its technical challenges due to different data systems. Some countries have chosen not to make their wastewater data publicly accessible.

The challenge of combining and analyzing NTD with traditional data was not unique to wastewater data. Sir Geoff Mulgan, professor at University College London, explains that there is a larger problem in the public sector's ability to combine and use scientific information during COVID-19: "Input from evidence synthesis is crucial for policymaking. But the capacity of governments to absorb such evidence is limited, and syntheses for decisions must go much further in terms of transparently incorporating assessments of political or practical feasibility, implementation, benefits and cost, among many other factors. The gap between input and absorption is glaring."[204]

198 Ibid.

199 Ibid.

200 Ibid.

201 "Building COVID-19 Wastewater Surveillance with the National Wastewater Surveillance System." *Centers for Disease Control and Prevention*, July 3, 2022, 1.

202 Keshaviah, Aparna. Interview re: wastewater data. Zoom, July 11, 2022.

203 Centers for Disease Control and Prevention. "COVID Data Tracker," September 21, 2022. https://covid.cdc.gov/covid-data-tracker.

204 Mulgan, Geoff. "COVID's Lesson for Governments? Don't Cherry-Pick Advice, Synthesize It." *Nature* 602, no. 7895 (February 1, 2022): 9–9. https://doi.org/10.1038/d41586-022-00212-5.

Despite the challenges with data synthesis and standardization, and the ongoing need to identify optimal sampling schemes and risk metrics that provide officials with sentinel warnings,[205] wastewater data has provided valuable insights at the local and regional levels. Wastewater surveillance helped contain infections in university dorms across the United States,[206] led to the early detection of the Omicron variant prior to clinical case reports,[207] and informed hospitals on how to target treatments based on the relative abundance of Omicron versus Delta variants in community wastewater.[208]

Globally, wastewater surveillance initiatives have provided valuable insights for responding to the COVID-19 pandemic. The infrastructure developed for such initiatives could potentially be leveraged to track "viral pathogens, antimicrobial resistance (AMR), pharmaceutical consumption, or exposure to chemical pollutants"[209] during and inpreparation for the next crisis, depending on the research limitations. Maximizing the impact of such initiatives requires a systematic approach where data can be shared among municipalities, regions, and countries quickly and efficiently.[210]

Wastewater data has provided valuable insights at the local and regional levels

205 Keshaviah, Aparna. Interview re: wastewater data. Zoom, July 11, 2022.

206 Betancourt, Walter Q., Bradley W. Schmitz, Gabriel K. Innes, Sarah M. Prasek, Kristen M. Pogreba Brown, Erika R. Stark, Aidan R. Foster, et al. "COVID-19 Containment on a College Campus via Wastewater-Based Epidemiology, Targeted Clinical Testing and an Intervention." *Science of The Total Environment* 779 (July 20, 2021): 146408. https://doi.org/10.1016/j.scitotenv.2021.146408;
Gibas, Cynthia, Kevin Lambirth, Neha Mittal, Md Ariful Islam Juel, Visva Bharati Barua, Lauren Roppolo Brazell, Keshawn Hinton, et al. "Implementing Building-Level SARS-CoV-2 Wastewater Surveillance on a University Campus." *The Science of the Total Environment* 782 (August 15, 2021): 146749. https://doi.org/10.1016/j.scitotenv.2021.146749.

207 Kirby, Amy E. "Notes from the Field: Early Evidence of the SARS-CoV-2 B.1.1.529 (Omicron) Variant in Community Wastewater — United States, November–December 2021." MMWR. *Morbidity and Mortality Weekly Report* 71 (2022). https://doi.org/10.15585/mmwr.mm7103a5.

208 Diamond, Megan B., Aparna Keshaviah, Ana I. Bento, Otakuye Conroy-Ben, Erin M. Driver, Katherine B. Ensor, Rolf U. Halden, et al. "Wastewater Surveillance of Pathogens Can Inform Public Health Responses." *Nature Medicine,* September 8, 2022. https://doi.org/10.1038/s41591-022-01940-x.

209 Robins, Katie, Anne F. C. Leonard, Kata Farkas, David W. Graham, David L. Jones, Barbara Kasprzyk-Hordern, Joshua T. Bunce, et al. "Research Needs for Optimising Wastewater-Based Epidemiology Monitoring for Public Health Protection." *Journal of Water and Health*, August 12, 2022, jwh2022026. https://doi.org/10.2166/wh.2022.026.

210 Ibid.

Wearables and biometrics

There are several examples[211] of COVID-19 pilot programs using various devices to detect biometrics, but the majority of these initiatives were not actualized beyond proof of concept. The initiatives that were implemented tended to involve temperature checks using thermal imaging along with other methodologies. These temperature check initiatives were implemented in a range of settings, including shops, airports, workplaces, and other settings.[212]

One such example is KC Wearable's (a technology firm) smart helmet "equipped with an ARM processor, an augmented reality display screen, an infrared camera, and a visible light camera"[213] that can detect the body temperature of people nearby. The helmet has an alarm when people who walk by have a high body temperature indicative of a fever, and the data collected is stored within the helmet.[214] As of June 2020, the helmets were used by 35 countries, including Italy, South Africa and Chile by public servants such as police officers, transit operators, medical staff, and others.[215] While the smart helmet has been used to rapidly detect high temperatures among large groups, these initiatives have been widely criticized by the public for the necessity of this data, ambiguity surrounding how it will be repurposed, and racial biases.[216] As the indicators of infection evolved, body temperature was no longer as reliable as a COVID-19 indicator. Limited information is available about how this initiative is still in use.

---

211 For examples, see: Mishra, Tejaswini, Meng Wang, Ahmed A. Metwally, Gireesh K. Bogu, Andrew W. Brooks, Amir Bahmani, Arash Alavi, et al. "Pre-Symptomatic Detection of COVID-19 from Smartwatch Data." *Nature Biomedical Engineering* 4, no. 12 (December 2020): 1208–20. https://doi.org/10.1038/s41551-020-00640-6;
Reuters Staff. "Chinese Startup Rokid Sees Opportunity with COVID-Fighting Smart Glasses." *Reuters*, May 1, 2020, sec. Healthcare & Pharma. https://www.reuters.com/article/us-health-coronavirus-china-detection-gl-idUSKBN22D4TQ.

212 Lewis, Nicole. "Biometric Technology Use During the Pandemic Can Pose Ethical Problems." SHRM, November 9, 2020. https://www.shrm.org/resourcesandtools/hr-topics/technology/pages/biometric-technology-use-during-pandemic-can-pose-ethical-problems.aspx;
Davies, Rob. "'Conditioning an Entire Society': The Rise of Biometric Data Technology." *The Guardian*, October 26, 2021, sec. Technology;
Burt, Chris. "Contactless Biometrics and Temperature Screening Launched and Deployed to Prevent COVID-19 Spread," October 1, 2020. https://www.biometricupdate.com/202010/contactless-biometrics-and-temperature-screening-launched-and-deployed-to-prevent-covid-19-spread.

213 Ghosh, Shona. "Police in China, Dubai, and Italy Are Using These Surveillance Helmets to Scan People for COVID-19 Fever as They Walk Past and It May Be Our Future Normal." Business Insider, May 17, 2020. https://www.businessinsider.com/coronavirus-italy-holland-china-temperature-scanning-helmets-2020-5.

214 Ibid.

215 KC Wearable. "KC Wearable Announces SMART Helmet Is Being Used in 35 Countries Globally to Help Tackle COVID-19 and End Lockdowns." Cision PR Newswire, June 16, 2020. https://www.prnewswire.com/news-releases/kc-wearable-announces-smart-helmet-is-being-used-in-35-countries-globally-to-help-tackle-covid-19-and-end-lockdowns-301077028.html.

216 Maxwell, Tom. "Police Are Using High-Tech Helmets to Check People for Fevers from COVID-19." Input, May 18, 2020. https://www.inputmag.com/tech/police-in-china-dubai-italy-are-using-these-helmets-to-scan-people-for-covid-19-fevers.

Vignette: The National COVID Cohort Collaborative (N3C)

The N3C was supported through NCATS funding for different purposes including developing infrastructure, supporting data providers. However, much of the implementation can be attributed to the network of volunteers driven to impact the pandemic response.[217] It is unclear how this initiative will evolve in the long-term if crisis-driven volunteerism plateaus.

Since its launch, The N3C has become one of the most comprehensive, public databases for individual patient health data in the United States consisting of around 14 billion data rows.[218] Additionally, The N3C has led to outcomes including: "identification of the clinical characteristics of acute COVID-19 for risk prediction, assisted in providing clinical care for immunocompromised adults, revealed how COVID infection affects children, and documented that vaccines appear to reduce the risk of developing long COVID."[219] The N3C is continuing to operate and looking to expand to include other diseases beyond COVID-19.

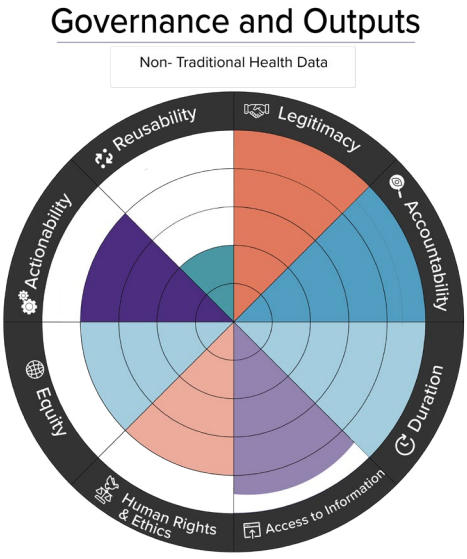


Figure 7. Assessment: Non-Traditional Health Data, Governance and Outputs
The figure above shows our assessment of the implementation and governance of The N3C using the assessment criteria.

217 Haendel, Melissa. Interview re: The N3C. Zoom, July 1, 2022.

218 National Center for Advancing Translational Sciences. "About the National COVID Cohort Collaborative," February 1, 2022. https://ncats.nih.gov/n3c/about.

219 Lloyd, Jay. "Building a Data Infrastructure for the Bioeconomy." *Issues in Science and Technology* (blog), May 18, 2022. https://issues.org/building-data-infrastructure-bioeconomy-sarma-haendel/.

2.1.5. Additional Examples

Data source	Additional Examples
Digital patient data	• COVID-NET: Researchers with the University of Waterloo, Canada and DarwinAI Corp. launched a study that uses an AI system to aid in the screening for COVID-19 within chest radiology images. • Carbon Health Clinical Data Repository: Carbon Health, a medical technology company, and Braid Health, a medical diagnostic imaging company, developed a collection of the physical indicators of COVID-19 patients–including those who have and have not tested positive on a clinical test. • Israel's Green Pass: a Green Pass proving one's vaccination as a precondition for entering certain businesses and public spheres. • Italy's COVID Passport: a COVID Passport for workers to attest they have been vaccinated and are able to work.
Platform facilitated symptom data	• World Bank COVID-19 & Google Trends: The World Bank developed a global dashboard, consisting of almost 200 countries, including internet search trends that correlate with COVID-19 symptoms from February 1 through October 31, 2020. • IoConto: IoConto is an open civic platform in Italy where hospitals and individuals could report COVID-19 infections, hospitalizations and deaths across the country. • How We Feel: How We Feel is an app developed by Harvard Chan faculty that collects self-reported symptoms of COVID-19. • HealthMap: HealthMap is a social listening platform that collects data from local newspapers, social media, and other sources to generate visualizations of the spread of disease.
Wastewater data	• South Africa: South Africa compiles their wastewater data collected across the country into a dashboard. • Northern Italy: A study testing wastewater data in Northern Italy demonstrated that COVID-19 had been present in the area since December 2019.
Wearables and biometrics	• School lunch payments: A school in Gateshead, United Kingdom implemented a pilot program using facial recognition technology to collect lunch payments without physical contact. • Smartphone testing: A study was conducted looking into whether breathing patterns collected from smart-phones could indicate a COVID-19 infection.


Photo by Zhang Kenny on Unsplash

2.2 Non-Traditional Mobility and Geolocation Data

This briefing is based, in part, on The GovLab's previous report "The Use of Mobility Data for Responding to the COVID19 Pandemic" (2021). For more information on specific applications of mobility data, please refer to that report.

Main Takeaways

- **Main use cases:** Throughout the pandemic, public officials and health experts mainly used mobility data to understand the spread of the virus and the impact of control measures.
- **Mobility Data Is Complex:** Mobility data used for pandemic response tended to come from large telecommunications companies, software companies, and start-ups who already had experience gathering and using mobility data in their daily operations. While this data could be rich and granular, these same attributes could make it difficult for other actors to meaningfully use it without support from data suppliers.
- **Agreements Were Key in Organizing Collaborations:** Many efforts to use telecommunications data or data derived from software development kits relied on agreements between governments, researchers, and data suppliers outlining how and where data would be used and what each party's responsibilities were. These enabled parties to work toward a common goal and address potential sources of risk.
- **There is Little Transparency on Pathways to Impact:** Several projects had clear pathways to impact decision-making, having signed agreements to provide insights directly to government leaders and their staff. However, there is little public information on how these insights led to meaningful decisions. In several cases where organizations did not have a direct relationship with policymakers, it is unclear who the intended audience was and how insights were meant to reach them.

2.2.1. Problem at Hand

COVID-19 is a disease spread via proximity. When people breathe "in air when in close contact with an infected person who is exhaling small droplets and particles that contain the virus" or have "these small droplets and particles land on their eyes, nose, or mouth," they are susceptible to infection.[220] Because of this relationship, decision-makers and epidemiologists turned to new data methods to understand the movements of different people in relation to their environment and to other people.

When paired with case counts and other metrics, some data practitioners have argued that aggregated mobility data—acquired from telecoms, tech companies, and advertisers—could guide them on when and where to institute stay-at-home orders, social-distancing mandates, and other requirements without the labor-intensive and memory-dependent work of contact tracing.[221] They could also evaluate the effectiveness of existing measures to control disease spread (e.g. stay-at-home orders, social distancing requirements). In settings where traditional sources may be limited or inaccessible, aggregated mobility data may be essential for humanitarian response.[222]

While some privacy advocates had previously expressed concern[223] about these methods potentially exposing identities, a variety of organizations responded to these arguments by releasing aggregated mobility data or insights to the public, researchers, or public policymakers.[224] Many privacy advocates, academics, and members of the public continue to have concerns about the potential of these sources to be used by surveillance, particularly after reports that countries such as Singapore shared mobility data with

---

220 Centers for Disease Control and Prevention. "COVID-19 Vaccines for People Who Are Moderately or Severely Immunocompromised," September 14, 2022. https://www.cdc.gov/coronavirus/2019-ncov/vaccines/recommendations/immuno.html.

221 Halpern, Sue. "Can We Track COVID-19 and Protect Privacy at the Same Time?" *The New Yorker*, April 27, 2020. https://www.newyorker.com/tech/annals-of-technology/can-we-track-covid-19-and-protect-privacy-at-the-same-time.

222 Open Data Institute, The GovLab, and Cuebiq. "COVID-19 Mobility Data Network's Use of Facebook Data for Good Mobility Data," n.d. https://theodi.org/wp-content/uploads/2021/04/5-COVID-19-Mobility-Data-Networks-Use-of-Facebook-Data_v2.pdf.

223 Kelsey Campbell-Dollaghan. "Sorry, Your Data Can Still Be Identified Even If It's Anonymized." *Fast Company*, October 12, 2018, sec. Smart Cities. https://www.fastcompany.com/90278465/sorry-your-data-can-still-be-identified-even-its-anonymized.

224 The GovLab, The Open Data Institute, and Cuebiq. "A Living Repository of Mobility Data Collaboratives Eeking to Address the Spread of COVID-19." #Data4COVID19 Mobility Repository. Accessed September 22, 2022. https://mobility.data4covid19.org/.

police.[225] In countries such as Canada, questions over the handling and processing of mobility data have fueled parliamentary hearings.[226] Human Rights Watch, for its part, has expressed concerns about the lack of transparency surrounding tracking apps, the potential for individuals in anonymized datasets to be "re-identified", among other concerns.[227]

2.2.2. Data Sources

"Mobility data" is a general term referring to any type of data describing the location of a device relative to people and landmarks in the physical world or to other devices. It is produced through the device's normal activity and can include data collected from telecommunications, software, GPS, and other assets. As Kishore, et al. in *The Lancet* note, the sources of mobility data "are aggregated in different ways by different providers, and have not been rigorously shown to reflect local contact rates or behaviour that are proportional to risk of transmission. Furthermore, the heterogeneous processing steps these data providers take to compute mobility metrics obscure what different proxies mean with respect to transmission."[228]

As The GovLab's "What is Mobility Data? Where is it Used?" briefing notes, mobility data can include telecommunications data, software development kit data, and GPS and other resources.[229]

Telecommunications Data

- ***Call Detail Records (CDR)*** are records of each time a mobile phone connects to a network when sending or receiving a voice call or SMS. A typical CDR contains the timestamp, call duration, caller identifier, recipient identifier, and the origin and

225 Illmer, Andreas. "Singapore Reveals Covid Privacy Data Available to Police." *BBC News*. January 5, 2021, sec. Asia. https://www.bbc.com/news/world-asia-55541001.

226 House of Commons - Chambre Des Communes - Canada. "Collection and Use of Mobility Data by the Government of Canada," November 22, 2021. https://www.ourcommons.ca/Committees/en/ETHI/StudyActivity?studyActivityId=11471238.

227 Human Rights Watch. "Mobile Location Data and Covid-19: Q&A," May 13, 2020. https://www.hrw.org/news/2020/05/13/mobile-location-data-and-covid-19-qa.

228 Kishore, Nishant, Aimee R. Taylor, Pierre E. Jacob, Navin Vembar, Ted Cohen, Caroline O. Buckee, and Nicolas A. Menzies. "Evaluating the Reliability of Mobility Metrics from Aggregated Mobile Phone Data as Proxies for SARS-CoV-2 Transmission in the USA: A Population-Based Study." *The Lancet Digital Health* 4, no. 1 (January 1, 2022): e27–36. https://doi.org/10.1016/S2589-7500(21)00214-4.

229 The GovLab and Cuebiq. "What Is Mobility Data? Where Is It Used?," April 2021. https://files.thegovlab.org/COVID-19_Mobility_Data_One-Page_Brief_v3_041521.pdf.

destination of cell towers when a call between devices occured. CDR's spatial precision can be limited because the coverage of cell towers tends to be large. They typically benefit from large user counts, which can represent a robust sample of the population.

CDRs can also present privacy challenges as, in its raw and unprocessed form, data refers to the behavior of specific devices (and in turn the users of those devices). The Groupe Spéciale Mobile Association (GSMA) privacy guidelines advise that any analysis on CDR be done on de-identified data and that individual data should not leave host servers.[230]

- ***Example:*** *In the TELUS Data for Good initiative (which was launched in April 2020 in response to the pandemic),*[231] *a Canadian national telecom company gave supervised and guided access to its network data to help the Natural Sciences and Engineering Research Council (NSERC) of Canada (which funds and provides other support to researchers across Canada) to NSERC researchers to study the COVID-19 crisis. The company provided free "supervised and guided access to strongly de-identified and aggregated network mobility data."*[232] *The TELUS Data for Good initiative informed several peer reviewed papers produced by researchers at Canadian universities. Their findings were published in academic journals, though the datasets were not included.*[233] *It also informed work conducted by the Public Health Agency of Canada, who sought CDR starting in December 2020 to understand disease spread.*[234]

  In advance of these collaborations TELUS contacted Canada's Office of the Privacy Commissioner that it intended to share "de-identified, aggregate

230 "Mobile Privacy Principles - Promoting Consumer Privacy in the Mobile Ecosystem." GSMA, n.d.

231 "TELUS Program Receives Prestigious Global Privacy Recognition." TELUS, n.d. https://www.telus.com/en/about/news-and-events/media-releases/telus-program-receives-prestigious-global-privacy-recognition.

232 "TELUS Data for Good Program to Provide De-Identified Network Mobility Data and Insights to the Natural Sciences and Engineering Research Council of Canada in Support of COVID-19 Research." TELUS, n.d. https://www.telus.com/en/about/news-and-events/media-releases/telus-data-for-good-program-to-provide-de-identified-network-mobility-data-and-insights.

233 Long, Jed A., and Chang Ren. "Associations between Mobility and Socio-Economic Indicators Vary across the Timeline of the Covid-19 Pandemic." *Computers, Environment and Urban Systems* 91 (January 2022): 101710. https://doi.org/10.1016/j.compenvurbsys.2021.101710.

234 Raisa Patel. "Is Public Health a Good Enough Reason for Ottawa to Look at Your Cellphone Data?" *The Toronto Star*, January 13, 2022, sec. Federal Politics. https://www.thestar.com/politics/federal/2022/01/13/is-public-health-a-good-enough-reason-for-ottawa-to-look-at-your-cellphone-data.html.

data with governments, health authorities and academic researchers in an effort to support work to respond to the COVID-19 crisis."[235] *While the Commissioner's office provided comments and guidance, TELUS opted not to enter a formal advisory engagement to determine whether "adequate safeguards had been adopted" to preserve individual privacy.*[236] *Following several complaints, the Commissioner's office subsequently opened an investigation into the effort to ensure the work did not collect or use personal information*[237]*—leading to further constraints in this type of data use moving forward. This led to media reporting calling to reevaluate Canada's privacy laws and for more transparent communication with the public.*[238]

Tag: Response

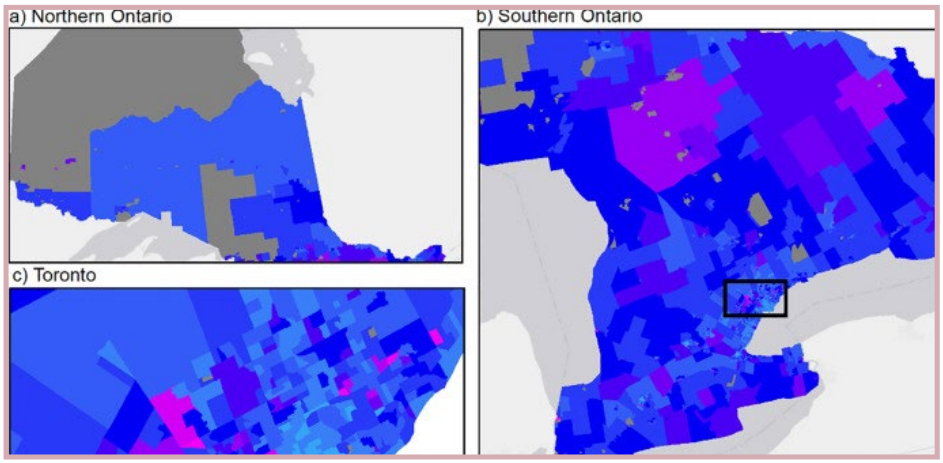


Figure 8. TELUS Data For Good Initiative - Mobility Patterns
This map demonstrates how mobility patterns from April 1st-8th in the province of Ontario changed from the initial lockdown period at the onset of the pandemic in Canada using data provided from the TELUS initiative.[239]

---

235 Office of the Privacy Commissioner of Canada. "Letter to the Standing Committee on Access to Information, Privacy and Ethics on Their Study of the Collection and Use of Mobility Data by the Government of Canada," March 3, 2022. https://www.priv.gc.ca/en/opc-actions-and-decisions/advice-to-parliament/2022/parl_sub_220301/.

236 Ibid.

237 Zafar, Nida. "Privacy Commissioner Speaks to Ethics Committee on Government Using Mobile Data to Track COVID." *MobileSyrup*, February 7, 2022. https://mobilesyrup.com/2022/02/07/privacy-commissioner-speaks-to-ethics-committee-on-government-using-mobile-data-to-track-covid/.
Office of the Privacy Commissioner of Canada. "Letter to the Standing Committee on Access to Information, Privacy and Ethics on Their Study of the Collection and Use of Mobility Data by the Government of Canada," March 14, 2022. https://www.priv.gc.ca/en/opc-actions-and-decisions/advice-to-parliament/2022/parl_sub_220301/.

238 Karadeglija, Anja. "Public Health Agency's Data Tracking of Canadians Spurs Calls for New Laws." National Post, January 13, 2022. https://nationalpost.com/news/politics/public-health-agencys-data-tracking-of-canadians-spurs-calls-for-new-laws.

239 Long, Jed A., and Chang Ren. "Associations between Mobility and Socio-Economic Indicators Vary across the Timeline of the Covid-19 Pandemic." *Computers, Environment and Urban Systems* 91 (January 1, 2022): 101710. https://doi.org/10.1016/j.compenvurbsys.2021.101710.

- ***x-Data Records (x-DRs)*** refers to internet connection detail records generated when a device (such as a cell phone or tablet) connects to a mobile internet. x-DRs provide more limited information than CDRs, often only describing when the connection happened and the coordinates of the telecom antenna used to download the content. However, it tends to be more time-specific.

  x-DRs face many of the same privacy risks as CDRs and, as GSMA recommends, any analysis should be conducted on de-identified data while individual level data remains restricted to network operator servers.[240]

  - ***Example:*** *In May 2020, researchers at the Universidad del Desarrollo, University of Turin, and University of Greenwich used x-DRs provided by the telecom provider Telefónica to assess mobility in Santiago de Chile as Chile responded to COVID-19.*[241] *The company processed its data to anonymize mobile phones included in the dataset. This engagement was a one-off research publication intended to inform decision-making in the city at the time of the publication (though it is unclear if and how this analysis informed decision-makers locally and nationally). Researchers noted that "the data will not be made publicly available," though they expressed an openness to collaboration with academia.*

    Tag: *Response*

Software Development Kit Data

- ***First-party, SDK-derived data*** is collected directly from a smartphone application without intermediary brokers. As it is data about one's own audience that an organization owns and manages itself, first-party SDK-derived data allows a company to exercise more control over how it collects data (including whether it allows individuals to opt in or out) and what it collects. It can also have improved spatial accuracy and precision. However, it is limited to the scale and breadth of the organization's own operations and audience. It might also be limited by privacy

240 "Mobile Privacy Principles - Promoting Consumer Privacy in the Mobile Ecosystem." GSMA, 2016. https://www.gsma.com/publicpolicy/wp-content/uploads/2012/03/GSMA2016_Guidelines_Mobile_Privacy_Principles.pdf.

241 Ferres, Leo, Rossano Schifanell, Nicola Perra, Salvatore Vilella, Loreto Bravo, Daniela Paolotti, Giancarlo Ruffo , and Manuel Sacasa. "Measuring Levels of Activity in a Changing City, A Study Using Cellphone Data Streams," 2020. http://datascience.udd.cl/covid_ids_tef_01.pdf.

legislation that curbs what can be collected by companies from apps and who they can share it with this type of data is unregulated and opaque in many places around the world, making it difficult for users to know how their data is used. While most organizations will have a "Privacy Policy" or "Terms of Service" associated with their app that describes how data is collected, analyzed, and shared, many of these apps can collect or share data in ways that were not anticipated. Moreover, many organizations seek to compile as much data as possible to create profiles on users, which may have privacy implications.[242]

- ***Example:*** *In May 2020, a research team that included the University of Southampton and the Wuhan Center for Disease Control and Prevention used first-party, SDK-derived data from the Chinese search engine Baidu to measure the changes in movement patterns and COVID spread from non-pharmaceutical interventions in China in response to COVID-19. The team released the models (the simulation used to arrive at its insights) it used for its analysis to inform future research and decision-making on GitHub.*[243]

  Tag: *Response*

- ***Third-party, SDK-derived data*** is data collected from smartphone apps by many sources aggregated into one dataset. This data can be rich (in that it describes all the items collected by the source organizations) and broad (in that it includes all organizations' audiences). However, the source organization has little to no control over collection practices and may be unable to trace back data to a reliable source that collected data with the knowledge and consent of data subjects.

  Third-party SDK data faces many of the same challenges as first-party DSK data, with the additional challenge of data holders potentially not knowing how the data was initially collected and users having difficulty knowing which organizations hold their data.

---

242 Klosowski, Thorin. "How Mobile Phones Became a Privacy Battleground—and How to Protect Yourself." *Wirecutter: Reviews for the Real World* (blog), September 29, 2022. https://www.nytimes.com/wirecutter/blog/protect-your-privacy-in-mobile-phones/.

243 Lai, Shengjie, Nick W. Ruktanonchai, Liangcai Zhou, Olivia Prosper, Wei Luo, Jessica R. Floyd, Amy Wesolowski, et al. "Effect of Non-Pharmaceutical Interventions to Contain COVID-19 in China." *Nature* 585, no. 7825 (September 2020): 410–13. https://doi.org/10.1038/s41586-020-2293-x.

- ***Example:*** *In the Teralytics project the Zurich-based mobility data company provided dashboards to health authorities, emergency responders, mobility providers, and transportation planners based on its collected third-party SDK-derived data. These dashboards informed decision-makers on how people move through their communities.*[244]

Tag: *Response*

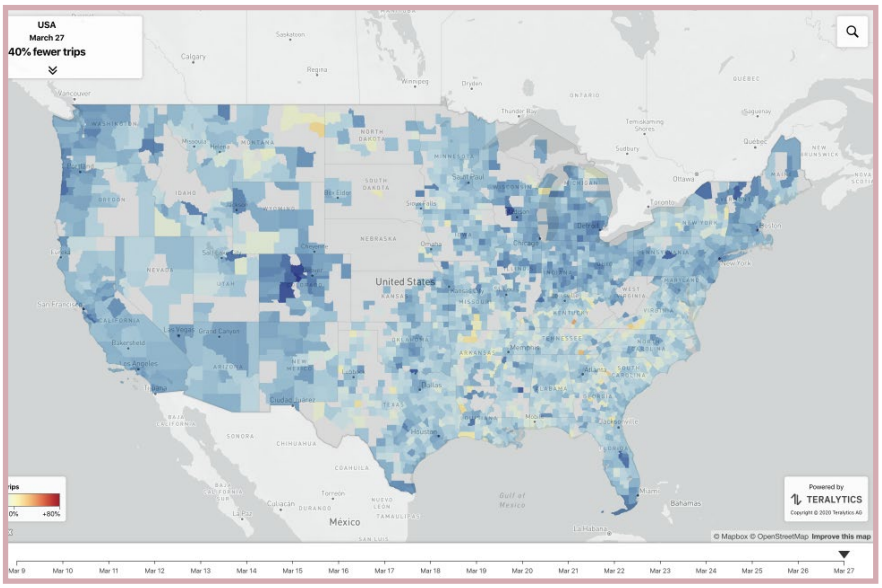


Figure 7. Teralytics
The map below shows how mobility patterns were changing in the United States using SDK-derived data.[245]

GPS and Other Resources:

- ***Vehicle GPS*** data describes the location of an automobile, ship, airplane, or other vehicle as indicated by portable devices placed in vehicles or hardware built into them.

  As with all GPS and location tracking, vehicle GPS has privacy implications. When aggregated, it can yield patterns on types of driver needs, preferences, and behavior but individually it can be present challenges. Moreover, many consumer vehicles may include GPS tracking without explicitly informing drivers or providing the ability to opt-out.[246]

244 Teralytics. "Mobility and COVID-19." Accessed August 29, 2022. https://www.teralytics.net/mobility-and-covid-19/.

245 Ibid.

246 The Parallax View. "How to Protect What Your Car Knows about You," April 3, 2018. https://www.the-parallax.com/protect-connected-car-data/.

- ***Example:*** *Geotab, a Canadian company that connects commercial vehicles to the internet and provides tools to analyze fleet activity, manages a COVID-19 Mobility Impact dataset that depicts changes in mobility due to the pandemic. This dataset supports decision makers planning to transport goods.*[247]

  Tag: *Response, Recovery*

- ***Bluetooth*** allows for short-range wireless communication between devices. It can be used to determine when and for how long a Bluetooth device came into proximity with another, though distance, object interference, and signal strength affect accuracy and precision.[248]

  As Privacy International notes, Bluetooth-facilitated tools have the potential to be one of the "least invasive" resources, especially when combined with encryption and decentralized designs.[249] However, the data remains very detailed (because proximity is so small) and apps with centralized databases may still be able to deanonymize individuals.[250]

  - ***Example:*** *In April 2020, Google and Apple announced they would support health agencies by building cross-device Bluetooth-based contact tracing functionality into their devices.*[251] *This functionality, known as Exposure Notifications,*[252] *was used by various health departments to develop apps that would notify users when they came in contact with someone who had*

247 #Data4COVID19 Mobility Data Collaborative Repository. "Geotab COVID-19 Mobility Impact," 2020. https://mobility.data4covid19.org/projects/geotab-covid-19-mobility-impact.

248 "Digital Contact Tracing Technology: Overview and Considerations for Implementation." Congressional Research Service: Informing the Legislative Debate Since 1914, May 2020. https://crsreports.congress.gov.

249 Privacy International. "Bluetooth Tracking and COVID-19: A Tech Primer," March 31, 2020. http://privacyinternational.org/explainer/3536/bluetooth-tracking-and-covid-19-tech-primer.

250 Ibid.

251 Apple Newsroom (United Kingdom). "Apple and Google Partner on COVID-19 Contact Tracing Technology," April 10, 2020. https://www.apple.com/uk/newsroom/2020/04/apple-and-google-partner-on-covid-19-contact-tracing-technology/.

252 Exposure Notifications: Helping fight COVID-19 - Google. "Exposure Notifications: Helping Fight COVID-19, with One Step on Your Phone." Google. Accessed August 29, 2022. https://www.google.com/intl/en_us/covid19/exposurenotifications/.

indicated they were sick with COVID-19.[253] *This later evolved into an in-system setting that users could opt in or out of manually on their device.*[254]

Tag: *Response*

Vignette: Mobility and Proximity in Canada During the COVID-19 Pandemic: First Party Software Development Kit Data

In early 2020, as COVID-19 began spreading across Canada, researchers at the University of Toronto decided to research whether there was a relationship between economic activity, and social distancing measures in Canada and elsewhere. To support this work, the research team collaborated with Cuebiq, a location-intelligence company that specializes in analyzing consumer behaviors for companies and marketers through apps that rely on its software development kit. Cuebiq collects data directly from smartphone users who have agreed to share their locations ("opting in") with these apps. Per its website, the company's proprietary software development kit lets Cuebiq "know exactly where the data is coming from and can ensure users are providing consent to data collection."[255]

2.2.3. Design and Planning

Mobility data encompasses a variety of different types of data and, necessarily, efforts to manage and use manifested in different ways. Many of these efforts centered on assessing adherence to government-imposed restrictions on movement and proximity to help decision-makers or (less frequently) the general public.

253 Singer, Natasha. "Why Apple and Google's Virus Alert Apps Had Limited Success." *The New York Times*, May 27, 2021, sec. Business. https://www.nytimes.com/2021/05/27/business/apple-google-virus-tracing-app.html.

254 Apple Developer. "Exposure Notification." Accessed August 22, 2022. https://developer.apple.com/documentation/exposurenotification.

255 Chan, Lawrence, and Cuebiq Marketing Team. "Cuebiq on Location Data Advancements and How Privacy Plays a Role." Cuebiq *(blog)*. https://www.cuebiq.com/resource-center/resources/cuebiq-on-location-data-advancements-and-how-privacy-plays-a-role/.

Telecommunications Data

Data collaborations around telecommunications data emerged in the early days of the pandemic. In the European Union, for example, a collaboration between the European Commission and telecom operators emerged in April 2020 following a request for data from telecom operators by the European Commissioner for the Internal Market.[256] By supplying aggregated and anonymized data on a pro bono basis[257] to the European Joint Research Centre, decision-makers hoped to better understand the impact of movement restrictions on mobility patterns. On a national level, Statistics Estonia coordinated several national telecoms for the same purpose.[258]

These projects are indicative of the general approach to using telecommunications data. Government institutions (such as national statistical agencies)[259] or educational centers (such as universities)[260] spearhead efforts to understand the impact of government orders on public mobility and disease spread. This insight is communicated to policymakers to inform their decisions. While specifics can vary depending on the data-sharing agreement signed, the data holder (which can include one or more telecommunications companies)[261] frequently conducts aggregation and/or anonymization of data in house (to minimize potential accidental exposure of sensitive information) before providing data to the external research party.[262] This data can be provided on a one-off basis or a continual basis.[263] Data use is further regulated by the individual data privacy and protection laws in effect in the

256 Open Data Institute, The GovLab, and Cuebiq. "European Commission Joint Research Center's Mapping Mobility Functional Areas (MFA) with Telecom Mobility Data," n.d.

257 The collaboration included -3 Group - part of CK Hutchison, A1 Telekom Austria Group, Altice Portugal, Deutsche Telekom, Orange, Proximus, TIM Telecom Italia, Tele2, Telefonica, Telenor, Telia Company and Vodafone.

258 #Data4COVID19 Mobility Data Collaborative Repository. "Statistics Estonia COVID-19 Mobility Analysis," 2020. https://mobility.data4covid19.org/projects/statistics-estonia-covid-19-mobility-analysis.

259 Ghana Statistical Service. "Ghana Covid 19 Hub." Accessed August 29, 2022. https://ghcovid19-stats-ghana.hub.arcgis.com/.;
Benjamins, Richard, Jeanine Vos, and Stefaan Verhulst. "Mobile Big Data in the Fight against COVID-19." *Data & Policy* 4 (ed 2022): e9. https://doi.org/10.1017/dap.2021.39.

260 The GovLab and Cuebiq. "COVID-19 Mobility Project." #Data4COVID19 Mobility Data Repository, 2021. https://mobility.data4covid19.org/projects/covid-19-mobility-project.

261 The GovLab and Cuebiq. "TELUS Data for Good." #Data4COVID19 Mobility Data Repository, 2021. https://mobility.data4covid19.org/projects/telus-data-for-good.

262 Open Data Institute, The GovLab, and Cuebiq. "European Commission Joint Research Center's Mapping Mobility Functional Areas (MFA) with Telecom Mobility Data," 2021. https://theodi.org/wp-content/uploads/2021/04/2-European-Commission-Joint-Research_v2.pdf.

263 The GovLab and Cuebiq. "Corona Prediction." #Data4COVID19 Mobility Data Repository, 2021. https://mobility.data4covid19.org/projects/corona-prediction-deutsche-telekom;
The GovLab and Cuebiq. "Statistics Estonia COVID-19 Mobility Analysis." #Data4COVID19 Mobility Data Repository, 2021. https://mobility.data4covid19.org/projects/statistics-estonia-covid-19-mobility-analysis.

countries where companies collect data and practitioners use data. Arrangements may further stipulate procedures and practices to preserve individual privacy (e.g. requiring that individual-level data not leave the data holder's server).


Photo by Nadir sYzYgY on Unsplash

Software Development Kit Data

With the exception of efforts with pre-established infrastructure (such as Meta's Population Density Maps[264] or Cuebiq's Data for Good program,[265] both of which were repurposed for COVID-19 response), many efforts reliant on data collected from apps launched between the Summer and Fall of 2020, several months after the World Health Organization declared COVID-19 a pandemic. These efforts, including the Curve Flattening Project[266] and "Effect of non-pharmaceutical interventions to contain COVID-19 in China"[267] sought to help the reader understand the impact of social distancing interventions imposed by governments. As these examples suggest, the general goal of efforts using Software Development Kit data was to understand human movement patterns in response to restrictions and disease spread.

Efforts involving large or well-established data providers like Meta, Cuebiq, or Safegraph connected directly with decision-makers at a local or national level or academics who

264 Meta's Population Density Maps are a series of high-resolution and highly accurate population maps developed using satellite imagery. It is estimated to be accurate up to a 30 meter resolution. *See*: Meta. "Data For Good at Meta High Resolution Population Density Maps." Facebook, 2022. https://dataforgood.facebook.com/dfg/tools/high-resolution-population-density-maps.

265 The GovLab and Cuebiq. "Cuebiq COVID-19 Data Collaboratives." #Data4COVID19 Mobility Data Repository, 2021. https://mobility.data4covid19.org/projects/covid19-mobility-data-collaborative.

266 The GovLab and Cuebiq. "Curve Flattening Project." #Data4COVID19 Mobility Data Repository, 2021. https://mobility.data4covid19.org/projects/curve-flattening-project.

267 Lai, Shengjie, Nick W. Ruktanonchai, Liangcai Zhou, Olivia Prosper, Wei Luo, Jessica R. Floyd, Amy Wesolowski, et al. "Effect of Non-Pharmaceutical Interventions to Contain COVID-19 in China." *Nature* 585, no. 7825 (September 2020): 410–13. https://doi.org/10.1038/s41586-020-2293-x.

might influence decision-making, typically in a data collaborative arrangement. Many of these efforts operated similarly to efforts involving telecommunications data. Government officials signed a data-sharing agreement to receive aggregated and anonymized data on an as-needed basis. One such example is the partnership between Cuebiq and the ISI Foundation–a research institution in Turin, Italy–which sought to understand how lockdowns in Italy impacted mobility patterns through mobility data collected by Cuebiq.[268]

Efforts involving large or well-established data providers like Meta, Cuebiq, or Safegraph connected directly with decision-makers at a local or national level or academics who might influence decision-making, typically in a data collaborative arrangement.

This effort included anonymized and aggregated user data and adhered to the GDPR.[269] The NYC Recovery Data Partnership, for example, relied on a set of agreements between the City of New York and participating data providers, with city agencies made aware of their obligations under citywide agreements.[270] Staff of the Mayor's Office of Data Analytics, First Deputy Mayor's Office, and Mayor's Office of Information Privacy jointly oversaw implementation of these agreements (and requests for data), informed by the city's Chief Privacy Officer and the Mayor's Office of Information Privacy as well as an outside expert advisory board.[271] This advisory board of experts was meant to meet on a bimonthly basis for as long as the partnership remains active and provide advice on specific proposals for data use.

Efforts involving smaller or less established providers often processed, analyzed, and published insights of their data with no defined audience. These efforts, such as the Enel X City Analytics - Mobility Map[272] and Unacast COVID-19 Social Distancing Scoreboard,[273]

268 Livaccari, Anna. "Real-Time Location Data Reveals Effect of Lockdown on Mobility in Italy Due to COVID-19." *Cuebiq* (blog). Accessed September 22, 2022. https://www.cuebiq.com/resource-center/resources/real-time-location-data-reveals-effect-of-lockdown-on-mobility-in-italy/.

269 Pepe, Emanuele, Paolo Bajardi, Laetitia Gauvin, Filippo Privitera, Brennan Lake, Ciro Cattuto, and Michele Tizzoni. "COVID-19 Outbreak Response, a Dataset to Assess Mobility Changes in Italy Following National Lockdown." *Scientific Data* 7, no. 1 (July 8, 2020): 230. https://doi.org/10.1038/s41597-020-00575-2.

270 The GovLab and Cuebiq. "NYC Recovery Data Partnership." #Data4COVID19 Mobility Data Repository, 2021. https://mobility.data4covid19.org/projects/nyc-recovery-data-partnership.

271 NYC Recovery Data Partnership. "Recovery Data Partnership." *NYC Analytics*, January 11, 2021. https://www1.nyc.gov/site/analytics/initiatives/recovery-data-partnership.page.

272 The GovLab and Cuebiq. "Enel X City Analytics - Mobility Map." #Data4COVID19 Mobility Data Repository, 2021. https://mobility.data4covid19.org/projects/city-analytics-mobility-map.

273 The GovLab and Cuebiq. "COVID-19 Social Distancing Scoreboard." #Data4COVID19 Mobility Data Repository, 2021. https://mobility.data4covid19.org/projects/covid19-socialdistancing-scoreboard.

manifested as dashboards or visualizations that could hypothetically be used by a member of the general public to make their own decisions. These organizations do not provide direct access to their data or specifics on how it was gathered.

GPS and Other Resources:

Given the diversity of data types included in this category, differences in project design appeared. These included:

- ***Vehicle GPS:*** Data-driven efforts making use of vehicle GPS tended to be conducted by companies that produced their own GPS data. TomTom[274] and Geotab[275] used GPS in their products and commercial vehicles to assess vehicle activity. These organizations subsequently created dashboards and visualizations meant to be consumed by the public to assess the impact of the pandemic on mobility.[276] There is often little public indication of the privacy standards used in these projects.

- ***Bluetooth:*** As early as April 2020, university researchers[277] and large tech companies[278] proposed new tools and technologies be used to scale and improve the accuracy of contact tracing, a common practice in public health to identify persons who may have come in contact with a person infected with a disease.[279] While several digital contact tracing methods were used during COVID-19 (see Table 2), this method, reliant on Bluetooth, was unique in that it theoretically would make individuals themselves data providers and data users. Using the capacity of individual consumer devices (primarily smartphones), people could transmit signals to one another while in proximity. If one individual tested positive and logged it in

274 The GovLab and Cuebiq. "Is Life Returning to Normal as Lockdown Lifts?" #Data4COVID19 Mobility Data Repository, 2021. https://mobility.data4covid19.org/projects/tomtom-congestion-analysis.

275 #The GovLab and Cuebiq. "Geotab COVID-19 Mobility Impact." #Data4COVID19 Mobility Data Repository, 2021. https://mobility.data4covid19.org/projects/geotab-covid-19-mobility-impact.

276 #Data4COVID19 Mobility Data Collaborative Repository. "Is Life Returning to Normal as Lockdown Lifts?," 2020. https://mobility.data4covid19.org/projects/tomtom-congestion-analysis.

277 "PACT: Private Automatic Contact Tracing," April 9, 2020. https://web.archive.org/web/20200409225425/https://pact.mit.edu/.

278 Apple Newsroom. "Apple and Google Partner on COVID-19 Contact Tracing Technology," April 10, 2020. https://www.apple.com/newsroom/2020/04/apple-and-google-partner-on-covid-19-contact-tracing-technology/.

279 Centers for Disease Control and Prevention. "What to Do If You Were Exposed to COVID-19," August 22, 2022. https://www.cdc.gov/coronavirus/2019-ncov/your-health/if-you-were-exposed.html.

their device, all people they were in contact with would be notified.

Several Bluetooth-enabled contact tracing efforts have been attempted using various frameworks. Singapore's TraceTogether app,[280] for instance, has centralized report processing in which users' contact data is uploaded to a health authority-administered server. This centralized approach has led to some privacy concerns, exacerbated when Singapore officials revealed that TraceTogether data was being used by police for criminal investigations and declared use of the app mandatory for all residents.[281] Similar challenges have affected Hong Kong's LeaveHomeSafe.[282] In the United States, disparate health authorities have launched their own contact-tracing apps with varying standards and practices and interoperability.[283] These applications are based on a collaboration between Google and Apple, which has incorporated additional privacy protections and keeps data decentralized so that users' locations cannot be continuously tracked.[284] This effort later evolved into a system built into the mobile phone software where users could participate without downloading an app.[285] The decentralized nature means that there is no one "data holder", data instead being held on individual devices and no data-sharing agreements. Ireland's Covidtracker app is similarly fully decentralized.[286] Overall responsibility of these initiatives tends to rest with the app developer, whether that be a company or government.

---

280 A Singapore Government Agency Website. "TraceTogether," August 30, 2022. https://www.tracetogether.gov.sg.

281 Andreas Illmer. "Singapore Reveals Covid Privacy Data Available to Police." *BBC News*. January 5, 2021. https://www.bbc.com/news/world-asia-55541001.

282 App Store Preview. "LeaveHomeSafe." Accessed September 21, 2022. https://apps.apple.com/us/app/leavehomesafe/id1536377801.

283 Singer, Natasha. "Why Apple and Google's Virus Alert Apps Had Limited Success." *The New York Times*, May 27, 2021, sec. Business. https://www.nytimes.com/2021/05/27/business/apple-google-virus-tracing-app.html.

284 "Exposure Notifications Frequently Asked Questions Preliminary — Subject to Modification and Extension ." *Apple & Google*, September 2020. https://covid19-static.cdn-apple.com/applications/covid19/current/static/contact-tracing/pdf/ExposureNotification-FAQv1.2.pdf.

285 Brandom, Russell. "Apple and Google Announce New Automatic App System to Track COVID Exposures." *The Verge*, September 1, 2020. https://www.theverge.com/2020/9/1/21410281/apple-google-coronavirus-exposure-notification-contact-tracing-app-system.

286 Health Service Executive. "Privacy and How We Use Your Data," November 26, 2020. https://www2.hse.ie/services/covid-tracker-app/privacy-and-how-we-use-your-data.html.

Data source	Description	Methods Used	Consumer Technologies	Examples
Bluetooth	Individual consumer devices can transmit signals to one another while in proximity and log it in their device–allowing all devices that were in proximity of someone who tested positive to be notified.	Proximity tracking Google and Apple API Decentralized privacy-preserving proximity tracing (DPT-3)	Primarily smart-phones	Estonia's HOIA App New Zealand's NZ COVID Tracer App Switzerland's SwissCovid App
Telecommuni-cations data	When individuals make phone calls, they transmit a signal to cell phone towers. These signals transmitted can be used to locate individuals.	Cell phone tower triangulation	Cell phones or smartphones	Cyprus' CovTracer App Ghana's GH COVID-19 Tracker Taiwan's location tracking during quarantine periods
GPS	All phones are equipped with a GPS system that send signals to satellites and wifi towers indicating the location. Previous locations can be used to track individuals who have tested positive and identify others in close proximity.	Satellites Wifi Towers	All phones	Bulgaria's Virusafe Israel's HaMagan App

Table 2: Digital Contact Tracing - Data Sources
The table above summarizes how different sources of mobility data were used for digital contact tracing.[287]

287 The information in Table 2 is drawn from the following sources: Patrick Howell O'Neill, Tate Ryan-Mosley, and Bobbie Johnson. "A Flood of Coronavirus Apps Are Tracking Us. Now It's Time to Keep Track of Them." *MIT Technology Review*, May 7, 2020, sec. Tech Policy. https://www.technologyreview.com/2020/05/07/1000961/launching-mittr-covid-tracing-tracker/;
Elise Poillot, Gabriele Lenzini, Giorgio Resta, and Vincenzo Zeno-Zencovich. *Data Protection in the Context of Covid-19. A Short (Hi)Story of Tracing Applications*. Romatre-Press, 2021. https://romatrepress.uniroma3.it/en/libro/data-protection-in-the-context-of-covid-19-a-short-history-of-tracing-applications/;
In Conversation with Audrey Tang. Video. Chatham House, London, 2020. https://www.chathamhouse.org/events/all/members-event/conversation-audrey-tang.

Vignette: Mobility and Proximity in Canada During the COVID-19 Pandemic: First Party Software Development Kit Data

The researchers at the University of Toronto decided to focus on three basic questions:

- How did Canadians react to the restrictions on mobility that were implemented in March, 2020, and subsequently to contain [the] spread of COVID? How did Canadians respond to the lifting of restrictions? Has mobility recovered to pre-pandemic levels?
- How did the proximity of Canadians to one another change with restrictions on mobility? Has contact recovered to pre-pandemic levels with the lifting of restrictions?
- What evidence do we have of changes in air travel?

They then began a conversation with Cuebiq, securing a data donation through the company's Data for Good Program. This donation was regulated through a data-sharing agreement between the two parties. Through a legal operative appointed by the University of Toronto, the research team negotiated with Cuebiq on the detail, scope, and quality of the data it would provide to the researchers as well as the physical conditions under which it could be stored and analyzed. To facilitate these negotiations, Cuebiq provided researchers with a "test dataset" containing entries of fake inputs that could demonstrate the types of information Cuebiq could provide.

Ultimately, the two parties came to an agreement in which Cuebiq would provide fully de-identified and anonymized data with differential privacy for 40 consecutive weeks—from January 1, 2020 to October 6, 2020. This dataset included the locations of iOS and Android devices that used any of the over 100 smartphone applications using Cuebiq's software development kit, with location aggregated by census division. The dataset (narrowed from Italy, the United States, and Canada to Canada alone) reflected 700,000 devices, a number which increased to 870,000 by mid-March due to additional downloads.[288]

288 Cavalli, Gabriel, Brennan Lake, Anita M McGahan, and Emanuele Pepe. "Mobility and Proximity in Canada During the COVID-19 Pandemic." Innovation Policy Lab. Toronto: Munk School of Global Affairs & Public Policy. University of Toronto, October 27, 2020.

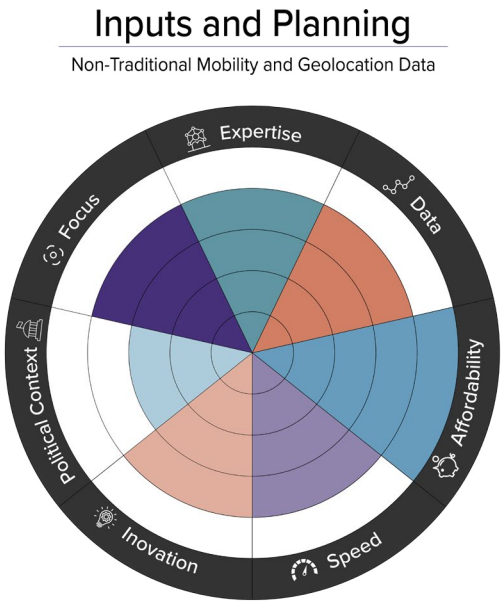


Figure 9. Assessment: Non-Traditional Mobility Data, Inputs and Planning
The figure above shows our assessment of the design of First Party Software Development Kit Data using the assessment criteria.

2.2.4. Implementation and Outcomes

Mobility data was predominantly used to understand the spread of COVID-19 and the impact of mobility restrictions by assessing patterns of people's movements. What separated approaches was whether they adopted a centralized solution, that being those approaches that focused on identifying patterns of behavior that could inform public health strategies, or decentralized, those being approaches that provided less fine-grained insights to individuals in a more privacy protective manner.

Telecommunications data, SDK data, and vehicle GPS data all took the first approach, seeking to understand these patterns. For efforts where governmental institutions were data users, projects fed directly into decision-making on how and where further restrictions were imposed. For efforts where researchers and academics were data users, projects fed into reports meant to be consumed by other researchers and, potentially, those advising decision-makers on policy. Other projects using these types, published on the open web, claimed that they were intended for public use and consumption. These efforts alleged they "informed millions on the changes in how people moved," though there is

little verifiable evidence they were widely used or that they meaningfully impacted public decision-making.[289] As Caroline Buckee, Satchit Balsari, and Andrew Schroeder note in their article, *Making data for good better,* "most efforts [to use private-sector data] did not yield consistently useful information for decision making, particularly in low resource settings, where capacity limitations in the public sector are most acute."[290]

Bluetooth, however, took a decentralized approach. It was not intended to feed into the decisions of high-ranking officials or contribute to large, internally held datasets. Rather, it sought to allow individuals to improve their decision-making by providing them with alerts on a voluntary basis when they had been in close contact with those who tested positive for COVID-19.[291] This approach could be seen in how many of these apps operated.[292] According to Technology Review's COVID Tracing Tracker, 40 of 49 contact-tracing applications from around the world (including both GPS-based and Bluetooth) were voluntary and only 18 stored data in a centralized repository.[293] An additional 31 of applications overall destroyed any data they did collect. In short, bluetooth contract tracing aimed to encourage individuals to understand when they had been exposed and self-isolate. There is little public reporting suggesting Bluetooth-facilitated contact tracing apps were used widely and significant reporting of problems with these apps stemming from physical restrictions[294] and low adoption[295] (an estimated 17.3%, for example, downloaded such apps in Italy). In Singapore, where residents were initially required to download the contact tracing app TraceTogether, the app's use by police for criminal investigations contributed to low levels of trust[296] and the app has gradually been phased out.[297]

289 Unacast. "Social Distancing Data Schema." Accessed August 29, 2022. https://www.unacast.com/covid19/docs/schema-for-covid-19-social-distancing-scoreboard.

290 Buckee, Caroline, Satchit Balsari, and Andrew Schroeder. "Making Data for Good Better." *PLOS Digital Health* 1, no. 1 (January 18, 2022): e0000010. https://doi.org/10.1371/journal.pdig.0000010.

291 Ministry of Health – Manatū Hauora. "Bluetooth Tracing." Accessed August 29, 2022. https://www.health.govt.nz/covid-19-novel-coronavirus/covid-19-resources-and-tools/nz-covid-tracer-app/using-nz-covid-tracer-app/bluetooth-tracing..

292 Patrick Howell O'Neill, Tate Ryan-Mosley, and Bobbie Johnson. "A Flood of Coronavirus Apps Are Tracking Us. Now It's Time to Keep Track of Them." *MIT Technology Review*, May 7, 2020, sec. Tech Policy. https://www.technologyreview.com/2020/05/07/1000961/launching-mittr-covid-tracing-tracker/.

293 Ibid.

294 Biddle, Sam. "The Problem With Using Bluetooth for Coronavirus Contact Tracing." *The Intercept*, May 5, 2020. https://theintercept.com/2020/05/05/coronavirus-bluetooth-contact-tracing/.

295 Guazzini, Andrea, Maria Fiorenza, Gabriele Panerai, and Mirko Duradoni. "What Went Wrong? Predictors of Contact Tracing Adoption in Italy during COVID-19 Pandemic." *Future Internet* 13, no. 11 (November 15, 2021): 286. https://doi.org/10.3390/fi13110286.

296 Han, Kristen. "Broken Promises: How Singapore Lost Trust on Contact Tracing Privacy." *MIT Technology Review*, January 11, 2021, sec. Pandemic Technology Project. https://www.technologyreview.com/2021/01/11/1016004/singapore-tracetogether-contact-tracing-police/.

297 Wei, Low De. "Singapore Phases Out the Use of a Controversial Covid Contact Tracing App." *Bloomberg*, April 22, 2022, sec. Political Prognosis. https://www.bloomberg.com/news/articles/2022-04-22/singapore-phases-out-use-of-controversial-contact-tracing-app.

Vignette: Mobility and Proximity in Canada During the COVID-19 Pandemic: First Party Software Development Kit Data

The University of Toronto researchers compiled their findings into a policy report available on the website of the Innovation Policy Lab at the Munk School of Global Affairs & Public Policy at the University of Toronto. Though the project had contracted in scope after its initial conception, the researchers expressed hope the piece would make an impact on the policymakers, scholars, and the public for whom the effects of restrictions in Canada were unclear.[298]

Following publication, the report received press coverage. The research team's findings were reported on coverage on television[299] and radio.[300] The media focused particularly on the privacy-protected approach, which lent the project additional legitimacy. The team lead, Anita McGahan, published an op-ed in the Toronto Star, Canada's highest circulation newspaper.[301] It also received attention from organizations based in major Canadian cities, stemming from the report's inclusion of estimates of mobility and proximity across Canada's 30 largest cities. The team presented its findings to members of city government in Ontario and Quebec.[302] In total, the team estimated subsequent attention necessitated three weeks of intensive media work.

Though the published report led to interactions with media personalities, former national leaders, academics, and city leaders, the team argued that their report's greatest impact was on the Canadian public. All of this activity around the report allowed the findings to be disseminated broadly and helped demonstrate the importance and effectiveness of social distancing measures. As this was intended to be a short-term research initiative, work concluded at the end of this effort.

---

298 McGahan, Anita. Interview re: Mobility and Proximity in Canada. Zoom, December 8, 2020

299 *It's Been Months since the COVID-19 Pandemic Started. How Have Our Behaviours Changed?* Video. Toronto: Global News Morning, 2020. https://globalnews.ca/video/7444297/its-been-months-since-the-covid-19-pandemic-started-how-have-our-behaviours-changed/.

300 "Training Judges Could Have Downfalls, How's Our Distancing? & Form Toy Making to the First Recyclable Masks." Radio. *The Kelly Cutrara Show*. 640 Toronto, October 28, 2020. https://omny.fm/shows/kelly-cutrara/training-judges-could-have-downfalls-hows-our-dist.

301 McGahan, Anita M. "In the Battle against COVID, Canadians Have Accomplished a Lot — and We Need to Do More." *The Toronto Star*, November 7, 2020, sec. Opinion. https://www.thestar.com/opinion/contributors/2020/11/07/in-the-battle-against-covid-canadians-have-accomplished-a-lot-and-we-need-to-do-more.html.

302 McGahan, Anita. Interview re: Mobility and Proximity in Canada. Zoom, December 8, 2020.

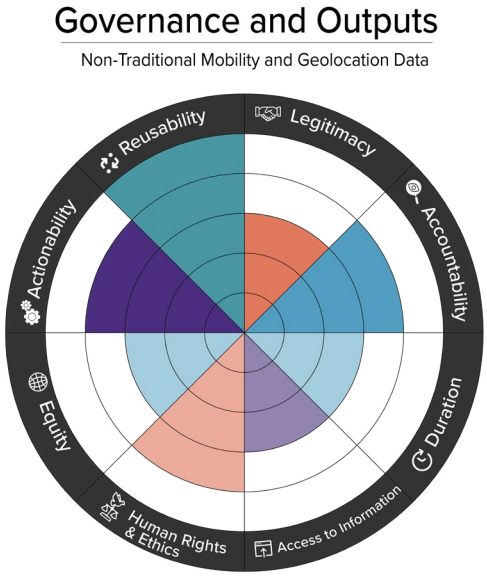


Figure 10. Assessment: Non-Traditional Mobility Data, Governance and Outputs. The figure above shows our assessment of the implementation and governance of First Party Software Development Kit Data using the assessment criteria

2.2.5. Additional Examples

Data source	Additional Examples
Telecommunications	• Taiwan digital fence - a mobile phone-based "electronic fence" that uses location-tracking to ensure people who are quarantined stay in their homes. • Measuring Levels of Activity in a Changing City (Chile) - researchers with UDD, the University of Turin, University of Greenwich, and Telefónica Investigación y Desarrollo analyzed call detail records in Chile, focusing on the Region Metropolitana, to assess mobility as the country responded to COVID-19. • European mobile network operators (MNOs) - the European Commission asked MNOs to share anonymized and aggregate mobile positioning data. The aim was to provide mobility patterns of population groups and serve the following purposes in the fight against COVID-19.

Software Development Kit Data	• Google Community Mobility Reports - to support public health officials internationally, Google has created a series of reports, based on aggregated, anonymized data from its products, that chart changes in movement due to the pandemic. • GLA COVID-19 Mobility Report - This report from the Greater London Authority provides evidence of changes in movement before and after city-wide control measures for COVID-19 using Apple, Google, and OpenStreetMap. • Cuebiq - Cuebiq, a private data company, assessed the movement of people via GPS-enabled mobile devices across the U.S. • SafeGraph COVID-19 Data Consortium – SafeGraph, a location and business intelligence firm collaborated with several researchers, nonprofits, and public agencies to study the COVID-19 pandemic. Safegraph obtained GPS data by regularly pinging 18 million smartphones with certain apps each day. It shared with its partners aggregated, anonymized data related to people's mobility patterns and foot traffic to businesses. • Data For Good program - a large-scale dataset provided by Facebook to the research community to offer a representation of the evolution of Italian mobility during the 2020 lockdown.
Vehicle GPS	• COVID-19 Mobility Impact dataset - a public dataset that offers insights to changes in mobility and the impact on wait times and mobility flow at critical points in the transport of goods. • Enel X City Analytics - Mobility Map - a dashboard free to access for citizens wanting to see the effects of the travel containment and social distancing measures adopted in response to the emergency.
Bluetooth-Based Proximity Data	• TraceTogether - Singapore's Bluetooth-enabled contact tracing app. The city slowly began to phase out use. It was considered controversial because it shared records with police. • COVID Alert NY - A "voluntary, anonymous, exposure-notification smartphone app" used in New York City that relies on Bluetooth technology to determine if the user was in proximity of someone who later indicated they were infected with COVID-19. • CoronApp: CoronApp is an application from the Government of Colombia that asks for personal and health data from the person who registers and allows users to track the well-being of registered users within their family. The app can transmit location data and use bluetooth to identify proximity to someone who has indicated infection with the disease.
Geospatial data	• Google Community Mobility Reports (Global) - to support public health officials internationally, Google has created a series of reports, based on aggregated, anonymized data from its products, that chart changes in movement due to the pandemic. • TomTom Traffic Index (Global) - an annual measure of worldwide traffic congestion using data from TomTom's navigation and mapping tools. This information can be used to estimate the impact of the COVID-19 pandemic on activity within these regions.

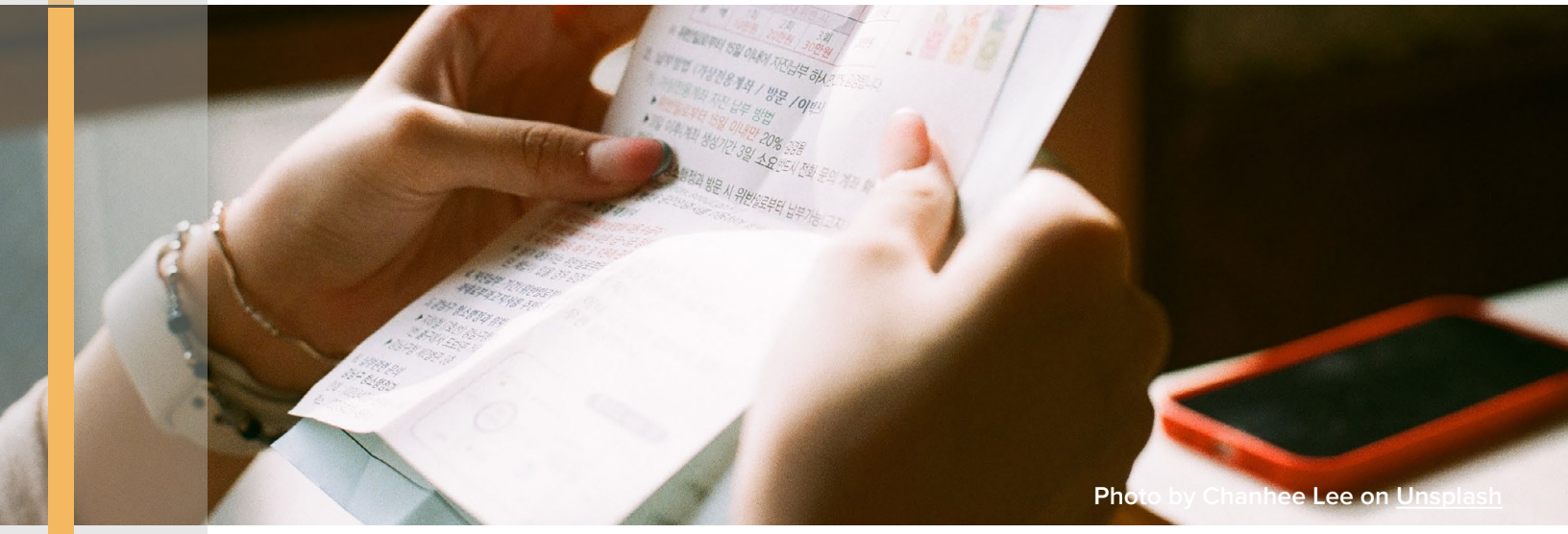


Photo by Chanhee Lee on Unsplash

2.3. Non-Traditional Economic Data

Main Takeaways:

- **Main use cases:** Having access to non-traditional economic data allowed policymakers to gain further insights into how businesses were reacting to social lockdowns as well as how citizens' spending habits were changing as the virus and associated health policies were evolving.

- **Non-traditional economic data was made accessible by repurposing existing data:** At times, credit card companies provided access to data that had already been made accessible in the past for research purposes. In other instances, COVID-19 provided a watershed moment to be perceived as an insight provider.

- **The data contained biases:** Policy makers cited that non-traditional economic data was timely, but often incomplete and contained biases making it unable to stand on its own.

- **COVID-19 acted as a wake-up call for more transparency for governments:** Various governments had already been implementing open procurement platforms, COVID-19 also served as a wake-up call for more transparency because of the increases in spending. The research highlighted that many of these issues stem from procurement authorities' need to work quickly and efficiently during a pandemic, meaning that normal procedures and safeguards are set aside.

- **Transparency in procurement practices led to more efficiency**: Countries with transparent and data-driven procurement systems in place at the start of 2020 have pursued efforts towards inviting scrutiny, tracking supply and demand and getting resources to people who need them.[303]

303 Open Contracting. "Open for Business: Colombia's Data-Driven Procurement Reforms Increase Competition." *Open Contracting Partnership* (blog), July 16, 2020. https://www.open-contracting.org/2020/07/16/open-for-business-colombias-data-driven-procurement-reforms-increase-competition/.

2.3.1. Problem at Hand

The COVID-19 pandemic's economic effects were unprecedented—dramatically undermining the financial solvency of individuals, businesses, and governments.[304] In the first few weeks of lockdowns, stories emerged about mass firings and hiring freezes at companies, communities in danger of bankruptcy and individuals unable to pay rent and other bills.[305] These stories contributed to an overall feeling of crisis. The effects of the pandemic were everywhere, immediate, and overwhelming–demonstrating the need to rapidly understand the state of economic health to design targeted policy solutions.

The traditional metrics used to measure economic health, such as the GDP or inflation rates, lagged behind the needs of policy-makers due to either slower collection and reporting or because they were too broad to address local issues meaningfully.[306]

To fill these unmet needs, policy-makers and businesses began looking for new approaches and methods that could supplement the traditional assessments conducted by economic agencies, financial institutions and statistical agencies. They began exploring ways to employ new, non-traditional, big data sources to better understand the financial health of various actors at scale and in near-real time.[307] In addition to using mobility data as a proxy for economic activity, these actors turned to (re)using new types of economic data about credit card activity, and contracts.

---

304 Jackson, James, Martin Weiss, Andres Schwarzenberg, Rebecca Nelson, Karen Sutter, and Michaela Th. Mayrhofer. "Global Economic Effects of COVID-19." Washington, DC: *Congressional Research Service*, November 10, 2021. https://sgp.fas.org/crs/row/R46270.pdf.

305 Bhattarai, Abha, Heather Long, and Rachel Siegel. "The First U.S. Layoffs from the Coronavirus Are Here." *Washington Post*, March 11, 2020, sec. Business. https://www.washingtonpost.com/business/2020/03/11/layoffs-coronavirus/.

306 Cajner, Tomaz, Laura J. Feiveson, Christopher J. Kurz, and Stacey Tevlin. "Lessons Learned from the Use of Nontraditional Data during COVID-19." *Brookings* (blog), April 27, 2022. https://www.brookings.edu/essay/lessons-learned-from-the-use-of-nontraditional-data-during-covid-19/.

307 Mehta, Nishita, and Sharvari Shukla. "Pandemic Analytics: How Countries Are Leveraging Big Data Analytics and Artificial Intelligence to Fight COVID-19?" *SN Computer Science* 3, no. 1 (November 9, 2021): 54. https://doi.org/10.1007/s42979-021-00923-y.

2.3.2. Data Sources

Non-traditional economic data refers to data about the economic activity of individuals, groups, or organizations. Sources of non-traditional economic data include credit card transactions, real-time supply chain data and open procurement contracting data. During the pandemic, governments and organizations were able to detect trends and make NTD-driven decisions by using methodologies such as statistical algorithms, machine learning, and predictive analytics capabilities.[308]

Credit and Debit Card Transactions: Credit and debit card transactions generate data about where, when, and how people spend money as well as how much money they are spending. Statistical organizations have tried to use these transactions as a proxy for economic health amid the pandemic.[309] This work included point of sale transactions to assess payment habits during the pandemic.

- ***Example:*** *The Bureau of Economic Analysis of the United States used daily card data from Fiserv, a financial technology company, to measure the reduction in revenue of local businesses around the time of the pandemic.[310] This data source was intended to help understand economic activity at the local level. The aggregate Fiserv card series included merchants that utilized Fiserv services. This means that all associated card transactions (for example, credit card, debit card, and gift cards) go through their system, where the number of transactions and the amount of each transaction is documented. All data was aggregated in an attempt to make both merchants and individuals unidentifiable. This data was used to track consumer spending, mobility and change in shopping habits as a result of containment measures.[311]*

  Tag: *Response, Recovery*

---

308 Beasley, Kevin. "Covid-19 Altered The Supply Chain—Does Your Forecasting Strategy Reflect The New Reality?" *Forbes*, July 14, 2022. https://www.forbes.com/sites/forbestechcouncil/2022/07/14/covid-19-altered-the-supply-chain-does-your-forecasting-strategy-reflect-the-new-reality/.

309 Dunn, Abe, Kyle Hood, and Alexander Driessen. "Measuring the Effects of the COVID-19 Pandemic on Consumer Spending Using Card Transaction Data." *BEA Working Paper Series, WP2020-5*, April 24, 2020, 22.

310 Ibid.

311 Ibid.

Supply Chain Data: A supply chain is broadly defined as "the processes involved in producing and distributing a product or service, including the organizations, people, tools, and information necessary to get an entity into consumers' hands."[312] The supply chain data, then, are all the data referring to such processes,[313] that inform the production and distribution of goods by analyzing trends and providing organizations with necessary insights to create an optimal supply plan.[314]

- ***Example****: Using maritime traffic data collected via a global network of Automatic Identification System (AIS) receivers, a study from the Marine Traffic Research Lab[315] investigates how the COVID-19 pandemic and the consequential social containment measures related to the shipping industry, which accounts alone for more than 80% of the world trade. They did so by comparing global maritime mobility levels in 2020 to those of previous years, and found that there was an unprecedented drop in maritime mobility across the world, and for all categories of commercial shipping. During times of crises, information as such proved to be key for organizations and managers to coordinate their production and distribution activities.[316]*

  Tag: *Response, Recovery*

Open Contracting: Open contracting refers to the practice of "publishing and using open, accessible and timely information on public contracting to engage citizens and businesses to fix problems and deliver results."[317] Amid the pandemic and the large amounts of money spent for response and recovery, some institutions used or otherwise expanded open contracts to prevent fraud, identify misuse of funds, make medical supplies more efficiently available, and mitigate corruption.[318]

---

312 Devane, Heather. "What Is the Data Supply Chain?" *Immuta* (blog), August 11, 2021. https://www.immuta.com/blog/what-is-the-data-supply-chain/.

313 Ibid.

314 Koshulko, Alex. "Overcoming Supply Chain Challenges With AI: What Large Manufacturers Need To Know." *Forbes*, April 12, 2022, sec. Innovation. https://www.forbes.com/sites/forbestechcouncil/2022/04/12/overcoming-supply-chain-challenges-with-ai-what-large-manufacturers-need-to-know/.

315 MarineTraffic. "Research at MarineTraffic," 2022. https://www.marinetraffic.com/research/research-at-marinetraffic/.

316 Spiliopoulos, Giannis. "COVID-19 Impact on Global Maritime Mobility." *MarineTraffic Research* (blog). Accessed September 6, 2022. https://www.marinetraffic.com/research/publication/covid-19-impact-on-global-maritime-mobility/.

317 Open Contracting Partnership. "What Is Open Contracting." Accessed August 31, 2022. https://www.open-contracting.org/what-is-open-contracting/.

318 Open Contracting Partnership. "Impact." Accessed August 31, 2022. https://www.open-contracting.org/impact/.

- ***Example:*** *The Open Contracting Data Standard "is a free, non-proprietary open data standard for public contracting, [...] that describes how to publish data and documents at all stages of the contracting process."*[319] *The Open Contracting Standard uses unique identifiers to make data accessible across countries using a common data model. The steps of public procurement are as follows: planning, tender, award, contract, and implementation.*[320] *The Open Contracting Standard provides various forms of information, data, and documents for each of these processes, allowing potential users to examine the data for several objectives. Interested citizens or government partners can then reference their procurement data against indicators of potential corruption referred to as "red flags."*[321]

Tag: *Readiness, Response and Recovery*

Vignette: Opening up data on emergency procurement in Ecuador

At the beginning of the pandemic, Ecuador had to make a quick decision on whether or not to publicly publish data on emergency procurement. As the pandemic escalated, medical supplies had to be sourced quickly and efficiently, and the need for transparency in public procurement grew exponentially.

Less than a month after the WHO declared the coronavirus a pandemic, Ecuador's National Public Procurement Service (SERCOP)[322] launched a platform that gathered all information about emergency buying procedures.[323] The platform gathered and showed data about open contracts involving both national and municipal agencies, making it easier for the supply and demand sides to connect, ultimately allowing medical supplies to be used more efficiently.[324]

319 Open Contracting Partnership. "Data Standard." Accessed August 30, 2022. https://www.open-contracting.org/data-standard/.

320 "Guide to Covid19 Procurement Data Collection, Publication & Visualization #public." Open Contracting Data Standard. Accessed September 21, 2022. https://docs.google.com/document/d/1VTqbBRuxEH3N1w-Lo5ozDFqLYNzMhxqOMh-bD6U1lYq0/edit?usp=embed_facebook.

321 "RED FLAGS for Integrity: Giving the Green Light to Open Data Solutions." Open Contracting Partnership & Development Gateway, August 2022. https://www.open-contracting.org/resources/red-flags-integrity-giving-green-light-open-data-solutions/.

322 "Gobierno Del Encuentro," n.d. https://portal.compraspublicas.gob.ec/sercop/.

323 El Universo. "En plataforma, se encuentran las contrataciones por la emergencia sanitaria," May 7, 2020. https://www.eluniverso.com/noticias/2020/05/07/nota/7835031/plataforma-se-engloban-contrataciones-emergencia-sanitaria.

324 Colman, Romina. "How COVID-19 and Collective Intelligence Transformed Procurement Risks into Opportunities for Transparency in Ecuador." Open Contracting Partnership (blog), December 20, 2020. https://www.open-contracting.org/2020/12/20/how-covid-19-and-collective-intelligence-transformed-procurement-risks-into-opportunities-for-transparency-in-ecuador/.

2.3.3. Design and Planning

Credit and Debit Card Transactions

During the initial months of the pandemic, some researchers and policymakers believed that the private sector held data on credit card transactions that could help them understand and assess the state of the economy.[325] Some sought collaborations with companies to gain access to credit card transaction data with the goals of providing rapid insights to evaluate the consumption habits, movements and the effects of containment on shopping habits during the early stages of the pandemic.[326] While some design efforts emanated from the private sector, launched by organizations such as the JP Morgan Chase Institute–an institute operating within the JP Morgan Bank[327]—others were developed by

academic institutions, such as Harvard University, using private sector data.[328] Finally, other initiatives originated from government actors, such as the Fiserv example discussed in the previous section.[329]

During the first few months of the pandemic, when policy-makers realized there was a gap in traditional data, several private sector actors decided to share the data they held on credit card transactions. These datasets were sometimes shared by the private company itself, with little transparency on the data or the conditions under which it would be disseminated, or at times shared with or purchased by researchers. These initiatives necessitated the creation of dedicated teams with specific technical expertise who could develop and analyze the data. The resources needed to develop the human capital were

---

325 U.S. Bureau of Economic Analysis (BEA). "COVID-19 and Recovery: Estimates From Payment Card Transactions," September 15, 2022. https://www.bea.gov/recovery/estimates-from-payment-card-transactions.

326 Bounie, David, Youssouf Camara, and John W. Galbraith. "The COVID-19 Containment Seen through French Consumer Transaction Data: Expenditures, Mobility and Online Substitution." *CEPR*, May 26, 2020. https://cepr.org/voxeu/columns/covid-19-containment-seen-through-french-consumer-transaction-data-expenditures.

327 The GovLab. "JPMorgan Chase Institute COVID-19 Research." #Data4COVID19 Repository, 2021. https://list.data4covid19.org/projects/jpmorgan-chase-institute-covid19.

328 "Real-Time Economics: A New Platform to Track the Impacts of COVID-19 on People, Businesses, and Communities Using Private Sector Data - Raj Chetty & Nathaniel Hendren." Accessed August 30, 2022. https://www.hks.harvard.edu/centers/mrcbg/programs/growthpolicy/real-time-economics-new-platform-track-impacts-covid-19-people.

329 Dunn, Abe, Kyle Hood, and Alexander Driessen. "Measuring the Effects of the COVID-19 Pandemic on Consumer Spending Using Card Transaction Data." *BEA Working Paper Series, WP2020-5*, April 24, 2020, 22.

therefore high, as they required specific human capital investments.[330] These investments were the result of a desire from the private sector to be seen as "insight providers" for either businesses or policy-makers, and the belief that launching such initiatives would either result in increasing business activity or visibility.[331] The JP Morgan Institute, for example, aggregated existing data with the purpose of informing private and public sector stakeholders.[332]

Efforts to assess the economy through non-traditional data were often combined with traditional data sources to complement policy-makers' understanding of COVID-19's impact on the economy. Harvard Business School, for example, created a publicly accessible dashboard that "tracked economic activity at a high frequency, granular level."[333] The dashboard[334] provided statistics on several indicators such as consumer spending, employment rates, business revenues, and job postings which were anonymized and aggregated by a consulting firm.[335] This included data from credit card processors, payroll firms, job posting aggregators, and financial services firms.[336] The researchers presented these statistics in real-time by using an "automated pipeline" that ingests data from these firms and publicly reports statistics on the dashboard, often with a three-day latency after the relevant transactions occur.[337]

---

330 Cajner, Tomaz, Laura J. Feiveson, Christopher J. Kurz, and Stacey Tevlin. "Lessons Learned from the Use of Nontraditional Data during COVID-19." *Brookings* (blog), April 27, 2022. https://www.brookings.edu/essay/lessons-learned-from-the-use-of-nontraditional-data-during-covid-19/.

331 Jérôme, Coffinet, Delbos Jean-Brieux, Kaiser Vojtech, Kien Jean-Noël, Kintzler Étienne, Lestrade Ariane, Mouliom Michel, Nicolas Théo, Bricongne Jean-Charles, and Meunier Baptiste. "Tracking the Economy during the Covid-19 Pandemic: The Contribution of High-Frequency Indicators." *Banque de France Bulletin* 5, no. 231 (March 12, 2020): 14.

332 The GovLab. "JPMorgan Chase Institute COVID-19 Research." #Data4COVID19 Repository, 2021. https://list.data4covid19.org/projects/jpmorgan-chase-institute-covid19.

333 Chetty, Raj, John Friedman, Nathaniel Hendren, Michael Stepner, and The Opportunity Insights Team. "The Economic Impacts of COVID-19: Evidence from a New Public Database Built Using Private Sector Data." Cambridge, MA: National Bureau of Economic Research, June 2020. https://doi.org/10.3386/w27431.

334 Track the Recovery. "Opportunity Insights - Economic Tracker." Accessed August 30, 2022. https://tracktherecovery.org/.

335 Dunn, Abe, Kyle Hood, Alex Driessen, and Peter Shieh. "COVID-19 and Recovery: Estimates From Payment Card Transactions." U.S. Bureau of Economic Analysis, 2022. https://www.bea.gov/recovery/estimates-from-payment-card-transactions.

336 Chetty, Raj, John Friedman, Nathaniel Hendren, Michael Stepner, and The Opportunity Insights Team. "The Economic Impacts of COVID-19: Evidence from a New Public Database Built Using Private Sector Data." Cambridge, MA: National Bureau of Economic Research, June 2020. https://doi.org/10.3386/w27431.

337 Ibid.

Supply Chain Data

Following the onset of the pandemic, supply chain disruptions have significantly increased all around the world.[338] Indeed, demand suddenly decreased for some goods and soared for others,[339] affecting a variety of different goods, from toilet paper[340] to ventilators,[341] furniture[342] and groceries.[343] This resulted in skyrocketing shipping rates, missing equipment, demand and supply fluctuations and worker shortages,[344] among other challenges.[345]

To address the circumstances, companies used NTD and automated systems to build a resilient, flexible and efficient supply chain.[346] Many of these efforts employed AI, which easily and accurately detect demand and supply, dynamically simulating supply chains to show any fluctuation and inform companies on how to manage them.[347]

For instance, in the UK, a consortium of major industrial and technology businesses from the medical sector came together to produce ventilators for the NHS shortly after

---

338 Beasley, Kevin. "Covid-19 Altered The Supply Chain—Does Your Forecasting Strategy Reflect The New Reality?" *Forbes*, July 14, 2022. https://www.forbes.com/sites/forbestechcouncil/2022/07/14/covid-19-altered-the-supply-chain-does-your-forecasting-strategy-reflect-the-new-reality/.

339 Alicke, Knut, Ed Barriball, and Vera Trautwein. "How COVID-19 Is Reshaping Supply Chains." *McKinsey & Company*, November 23, 2021. https://www.mckinsey.com/capabilities/operations/our-insights/how-covid-19-is-reshaping-supply-chains.

340 Narishkin, Abby, Steve Cameron, and Victoria Barranco. "Why Toilet-Paper Demand Spiked 845%, and How Companies Kept up with It." *Business Insider*, September 28, 2020. https://www.businessinsider.com/why-toilet-paper-demand-spiked-845-how-companies-kept-up-2020-5.

341 Bhaskar, Sonu, Jeremy Tan, Marcel L. A. M. Bogers, Timo Minssen, Hishamuddin Badaruddin, Simon Israeli-Korn, and Henry Chesbrough. "At the Epicenter of COVID-19–the Tragic Failure of the Global Supply Chain for Medical Supplies." *Frontiers in Public Health* 8 (November 24, 2020): 562882. https://doi.org/10.3389/fpubh.2020.562882.

342 Wood, Zoe. "The Great Furniture Delay: 'We'll Be Eating Christmas Dinner on Our Camping Tables.'" *The Guardian*, December 11, 2021, sec. Business. https://www.theguardian.com/business/2021/dec/11/the-great-furniture-delay-well-be-eating-christmas-dinner-on-our-camping-tables.

343 Gregory, Xanthe. "Bare Supermarket Shelves in NSW as COVID Causes More Shortages." *ABC News*, January 4, 2022. https://www.abc.net.au/news/2022-01-04/food-shortages-at-major-supermarkets-covid-rises/100737066.

344 Koshulko, Alex. "Overcoming Supply Chain Challenges With AI: What Large Manufacturers Need To Know." *Forbes*, April 12, 2022, sec. Innovation. https://www.forbes.com/sites/forbestechcouncil/2022/04/12/overcoming-supply-chain-challenges-with-ai-what-large-manufacturers-need-to-know/.

345 UNCTAD. *Review of Maritime Transport 2021*. Geneva: United Nations, 2021. https://unctad.org/system/files/official-document/rmt2021_en_0.pdf.

346 Koshulko, Alex. "Overcoming Supply Chain Challenges With AI: What Large Manufacturers Need To Know." *Forbes*, April 12, 2022. https://www.forbes.com/sites/forbestechcouncil/2022/04/12/overcoming-supply-chain-challenges-with-ai-what-large-manufacturers-need-to-know/.

347 Ibid.

the pandemic was declared as such by the WHO.[348] The consortium used the Planning Optimization Add-in for Microsoft Dynamics 365 Supply Chain Management to speed up master planning and reduce performance loads.[349] The Planning Optimization Add-in for Dynamics 365 Supply Chain Management enables companies to perform data-driven production planning, informing them of fluctuations in customer demand, availability of production materials, and capacity constraints.[350]

The companies that managed to incorporate new ways of analyzing and acting upon such a dynamic supply chain using NTD were 2.5 times more likely to report they had preexisting advanced-analytics capabilities.[351] As previously noted throughout this report, indeed, it seems that design efforts for this non-traditional data source relied heavily on traditional data and previously established processes.

Open Contracting

Since the start of the pandemic, the demand for open contracting efforts increased significantly, driven by a governmental need for greater flexibility and actionability, and a call from the public for higher levels of transparency and accountability.[352] In many cases, governments had already implemented open contracting practices before the pandemic and were expanded or expedited during COVID-19 due to a need to upgrade procurement systems.[353]

The companies that managed to incorporate new ways of analyzing and acting upon such a dynamic supply chain using NTD were 2.5 times more likely to report they had preexisting advanced-analytics capabilities.

348 Microsoft News Centre UK. "Global companies come together to make ventilators for the NHS." March 30 2020. https://news.microsoft.com/en-gb/2020/03/30/global-companies-come-together-to-make-ventilators-for-the-nhs/

349 Swaminathan, Manoj. "Reduce Supply Chain Disruptions with AI, IoT, and Mixed Reality." Microsoft Dynamics 365 Blog (blog), April 8, 2020. https://cloudblogs.microsoft.com/dynamics365/bdm/2020/04/08/reduce-supply-chain-disruptions-with-ai-iot-and-mixed-reality/.

350 Ibid.

351 Alicke, Knut, Ed Barriball, and Vera Trautwein. "How COVID-19 Is Reshaping Supply Chains." *McKinsey & Company*, November 23, 2021. https://www.mckinsey.com/capabilities/operations/our-insights/how-covid-19-is-reshaping-supply-chains.

352 Open Contracting Partnership. "Emergency Procurement for COVID-19: Buying Fast, Smart, and Open," 2022. https://www.open-contracting.org/what-is-open-contracting/covid19/.

353 Open Contracting Partnership. "Advocating for open contracting: How to use the growing open contracting momentum for change." *Open Contracting Partnership* (blog). Accessed September 6, 2022. https://www.open-contracting.org/advocacy/.

Pre-existing partnerships and infrastructure were fundamental for emergency open contracting initiatives to succeed.[354] In Moldova, for instance, it was possible to build a campaign for COVID-19 monitoring in only 60 days thanks to a group of civil society organizations and government ministries[355] that had been previously established to face HIV and tuberculosis.[356]

Additionally, many of these initiatives relied on collaborations between governments, research institutions, and private sector organizations.[357] For instance, in Ukraine, the Ministry of Health and the Medical Centralized Procurement Agency developed a tool for hospitals to publicly share information on their demand and current stocks.[358] Moreover, a coalition made of members from civil society, business and government created a transparent, open source, online system called Prozorro,[359] with the aim to allow government agencies to conduct procurement deals electronically and fully transparently.[360]

The public also played a crucial role in overseeing and improving e-procurement efforts. In Paraguay, for instance, "ordinary people started taking screenshots of procurement data, which led to calls for certain tenders being canceled as a result of public pressure or processes being halted"[361] said David Riveros Garcia, from reAcción.[362]

---

354 Frauscher, Kathrin. "5 Procurement Strategies for Navigating the COVID-19 Crisis from around the World." *Open Contracting Partnership* (blog), April 8, 2020. https://www.open-contracting.org/2020/04/08/5-procurement-strategies-for-navigating-the-covid-19-crisis-from-around-the-world/.

355 Tender.health. "Coaliția Organizațiior Societății Civile." Google Docs, April 9, 2020. https://docs.google.com/document/u/3/d/1DSI4oN3zq0F4O63vHAMuSWGIS60UFUi6JYIZtngBrBg/edit?fbclid=IwAR1dLcVDVi9f_AGCiBNs79p0MPRkiFIdFEzOdzp6vXavchAdBwx3-kbLrGQ&usp=embed_facebook.

356 Sklar, Kaye. "Building a Campaign for COVID Monitoring in Moldova in 60 Days." *Open Contracting Partnership* (blog), May 8, 2020. https://www.open-contracting.org/2020/05/08/building-a-campaign-for-covid-monitoring-in-moldova-in-60-days/.

357 Open Contracting Partnership. "Emergency Procurement for COVID-19: Buying Fast, Smart, and Open," 2022. https://www.open-contracting.org/what-is-open-contracting/covid19/.

358 "Home Page." MedData, 2022. https://meddata.com.ua/.

359 ProZorro. "ProZorro Is a Hybrid Electronic Open Source Government E-Procurement System Created as the Result of a Partnership between Business, Government and the Civil Society." Accessed August 8, 2022. http://prozorro.gov.ua/en.

360 Open Contracting Partnership. "Ukraine: Everyone Sees Everything." *Open Contracting Partnership* (blog), 2022. https://www.open-contracting.org/impact-stories/impact-ukraine/.

361 Colman, Romina. "Calling for Accountability: How Paraguay's Open Emergency Procurement Can Help Restore Public Trust." *Open Contracting Partnership* (blog), May 3, 2021. https://www.open-contracting.org/2021/05/03/calling-for-accountability-how-paraguays-open-emergency-procurement-can-help-restore-public-trust/.

362 Ibid.

Vignette: Opening up data on emergency procurement in Ecuador

Thanks to partnerships that were already solidly organized, civil society organizations and government officials collaborated to plan and design the efforts needed to open up data about emergency contracting. Indeed, prior to the arrival of COVID-19 in Ecuador, Ecuador's National Public Procurement Service was already collaborating with a collection of civil society organizations and the Open Contracting Partnership to establish a more open, transparent, and responsive procurement environment. This existing collaboration laid the groundwork for SERCOP to be able to show what was happening in the procurement system as the different contracts were taking place.

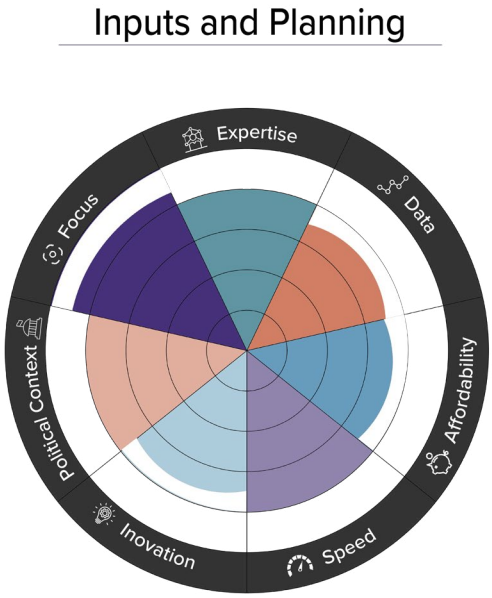


Figure 12. Assessment: Non-Traditional Economic Data, Inputs and Planning
In this figure, we demonstrate our assessment of the design of this initiative using our assessment framework.

2.3.4. Implementation and outcomes

The proliferation of data collection technologies and applications produced much economic data that became available to policy makers during the response and recovery stages of the pandemic. These sources, alongside mobility data research, provided new insights into COVID-19's economic effects.[363] However, the rapidly evolving nature of the pandemic created a need to respond quickly to the economic crisis caused by the COVID-19 pandemic. During the response and recovery stages of the pandemic, non-traditional data such as credit card transactions, open procurement and supply chain data emerged as a way to more quickly and specifically assess the state of the economy and specific stakeholders within it.

Credit and Debit Card Transactions

Researchers who used credit card transaction data held by private companies have stated that it holds enormous potential for assessing economic activity with a more localized heterogeneity, "as different places, sectors, and subgroups are often hit by different shocks and pursue different local policy responses."[364]

When it comes to crisis situations like COVID-19 where rapid responses are necessary, credit card transactions brought granular, detailed data to policy-makers that could in turn provide a better understanding of real-time developments, making it easier to "get around the problem of the long publication times for official data, and rapidly measure the impact of the economic shock".[365] This, in turn, could also lead to a more targeted policy recommendation, responses and inform follow-on policy actions.

For example, the United Kingdom used credit card transaction data along with several other NTD sources–including real-time data sources–to address speed and granularity

363 Mena, Gonzalo E., Pamela P. Martinez, Ayesha S. Mahmud, Pablo A. Marquet, Caroline O. Buckee, and Mauricio Santillana. "Socioeconomic Status Determines COVID-19 Incidence and Related Mortality in Santiago, Chile." Science 372, no. 6545 (May 28, 2021): eabg5298. https://doi.org/10.1126/science.abg5298.

364 Chetty, Raj, John Friedman, Nathaniel Hendren, Michael Stepner, and The Opportunity Insights Team. "The Economic Impacts of COVID-19: Evidence from a New Public Database Built Using Private Sector Data." Cambridge, MA: *National Bureau of Economic Research*, June 2020. https://doi.org/10.3386/w27431.

365 OECD. "How Will COVID-19 Reshape Science, Technology and Innovation?" *OECD Policy Responses to Coronavirus (COVID-19)*, June 23, 2021. https://www.oecd.org/coronavirus/policy-responses/how-will-covid-19-reshape-science-technology-and-innovation-2332334d/.

gaps of traditional data sources. Credit card transaction data showed that spending on "essentials" was consistent throughout the pandemic, in contrast to non-essential purchases which decreased.[366] Insights from these NTD sources were used within government decision making processes. The United Kingdom Office of National Statistics received "generally positive reviews for its performance during the crisis [...], in contrast to the 2008 financial crisis" where they primarily relied upon traditional data.[367]

However, it seems important to note that initiatives of this type were frequently found in countries where credit card usage is the norm.[368] Researchers have indeed noted the biases in these datasets. The Bank of Canada notes, for example, that findings show that estimates of economic activity based only on card data may be skewed and may overestimate consumer spending increases since payment preferences appeared to shift during the pandemic. Others have noted the existence of certain biases and limitations in the use of this data.[369] The Brookings Institution notes that "unlike government statistics, most alternative data sources are not designed with the purpose of generating statistics but are instead a byproduct of another use (such as card transactions). As such, the data are not designed to be representative of consumers or firms and may be hard to interpret or, worse, misleading".[370]

Non-traditional data such as credit card transactions, open procurement and supply chain data emerged as a way to more quickly and specifically assess the state of the economy and specific stakeholders within it.

Finally, initiatives emanating solely from the private sector demonstrated a willingness to publicize involvement in the COVID-19 response and recovery phases but few actual

366 Rosenfeld, David. "Using Real-Time Indicators for Economic Decision-Making in Government: Lessons from the Covid-19 Crisis in the UK." The Open Data Institute, September 13, 2022. https://odi.org/en/publications/using-real-time-indicators-for-economic-decision-making-in-government-lessons-from-the-covid-19-crisis-in-the-uk/.

367 Ibid.

368 There were, however, a variety of initiatives that examined the effect of digital cash transfers in places like Kenya. See, for example: Banerjee, Abhijit, Michael Faye, Alan Krueger, Paul Niehaus, and Tavneet Suri. "Effects of a Universal Basic Income during the Pandemic." Innovations for Poverty Action Working Paper, 2020. https://www.povertyactionlab.org/evaluation/effects-universal-basic-income-during-covid-19-pandemic-kenya.

369 Dahlhaus, Tatjana, and Angelika Welte. "Payment Habits During COVID-19: Evidence from High-Frequency Transaction Data," 2021, 35. https://www.bankofcanada.ca/2021/09/staff-working-paper-2021-43/.

370 Cajner, Tomaz, Laura J. Feiveson, Christopher J. Kurz, and Stacey Tevlin. "Lessons Learned from the Use of Nontraditional Data during COVID-19." *Brookings* (blog), April 27, 2022. https://www.brookings.edu/essay/lessons-learned-from-the-use-of-nontraditional-data-during-covid-19/.

outcomes. While claims were made about opening up and sharing data, most results were shared via reports and press briefings put out by the companies themselves and the accessibility of the data remains unclear.

Supply Chain Data

A study from Oden Technologies found that 71% of the production companies that have aimed to innovate their supply chain management are doing it by adopting NTD and digital analytical systems.[371] The use of NTD related to the supply chain was not only aimed at planning a more efficient and effective inventory; indeed, many organizations adopted predictive analytics to easily recognize new business opportunities.[372]
Overall, NTD proved to be key for businesses not only to increase operational efficiency, but also to exponentially improve cost savings. For instance, shipment-processing time can be significantly reduced thanks to new NDT-based analysis systems, managers have greater control and a more accurate understanding of the supply chain, and can more easily identify potential operational bottlenecks.[373]

Bearing in mind the above-mentioned benefits that NTD has provided in facing the sudden disruption of global supply chains stemmed from the COVID-19 pandemic, it seems important to note that the full potential of this data source often seems to be yet unlocked.[374] Indeed, whereas right now NTD are mainly used to provide insights and learnings to managers that are then to make decisions, they do offer significant opportunities to automate many decision making processes. Those decisions could be based on continuously learning systems that steadily improve their prediction accuracy from the outcomes of each singular decision and consequential event.[375]

---

371 Koshulko, Alex. "Overcoming Supply Chain Challenges With AI: What Large Manufacturers Need To Know." *Forbes*, April 12, 2022, sec. Innovation. https://www.forbes.com/sites/forbestechcouncil/2022/04/12/overcoming-supply-chain-challenges-with-ai-what-large-manufacturers-need-to-know/.

372 Beasley, Kevin. "Covid-19 Altered The Supply Chain—Does Your Forecasting Strategy Reflect The New Reality?" *Forbes*, July 14, 2022. https://www.forbes.com/sites/forbestechcouncil/2022/07/14/covid-19-altered-the-supply-chain-does-your-forecasting-strategy-reflect-the-new-reality/.

373 Beremski, Matei. "COVID-19: Challenges to Supply Chains, and the AI Solution." Supply Chain Digital, September 11, 2020. https://supplychaindigital.com/digital-supply-chain/covid-19-challenges-supply-chains-and-ai-solution.

374 Rodriguez, Pepe, Stefan Gstettner, Ashish Pathak, Ram Krishnan, and Michael Spaeth. "Why AI-Managed Supply Chains Have Fallen Short and How to Fix Them." *Boston Consulting Group*, September 1, 2022. https://www.bcg.com/publications/2022/benefits-of-ai-driven-supply-chain.

375 Ibid.

Open Contracting

Open contracting data sources were replicated across numerous countries during COVID-19 and were positively received by civil society, citizens and the media in those countries, as it provided further transparency on the spending of public money and possible acts of corruption.[376] Some of the findings were also used to take concrete actions such as ending certain types of contracts and prosecuting government officials involved.[377] These cases, however, remain quite rare. In Lithuania, the government had simplified their procurement negotiations procedures during the emergency to conclude contracts for PPE quicker, but speeding up the process also meant putting aside the usual checks-and-balances that maintain accountability.[378] Recognizing the value of CSO and journalist monitoring, it shared all the information as open data on a publicly accessible dashboard.[379] The dashboard offered visualizations of pandemic procurement information to show who brought what, from whom, and when.

A study from Oden Technologies found that 71% of the production companies that have aimed to innovate their supply chain management are doing it by adopting NTD and digital analytical systems.

Open procurement initiatives aimed to increase access to information and did not include any personally identifiable information, therefore, the need to protect the content of the datasets was not central to the implementation of the program. In most cases, the project was continued throughout various phases of the pandemic and beyond.[380] Indicators to measure success, however, were limited beyond transparency. While the programs were

376 Open Contracting Partnership. "Monitoring COVID-19 emergency procurement with data." *Open Contracting Partnership* (blog), March 25, 2020. https://www.open-contracting.org/2020/03/25/monitoring-covid-19-emergency-procurement-with-data/.

377 Hrytsenko, Yevhenii. "Tracking COVID-19 Procurement in Kazakhstan: From Red Flags to Cancellations." *Open Contracting Partnership* (blog), March 11, 2021. https://www.open-contracting.org/2021/03/11/tracking-covid-19-procurement-in-kazakhstan-from-red-flags-to-cancellations/.

378 Open Contracting Partnership. "Buy Open, Buy Fast: How Open Contracting Helped Lithuania's Coronavirus Response," March 30, 2021. https://www.open-contracting.org/2021/03/30/buy-open-buy-fast-how-open-contracting-helped-lithuanias-coronavirus-response/.

379 Viešųjų pirkimų tarnyba. "Kovai su COVID-19 sudarytos sutartys." Mano vyriausybė, October 15, 2021. https://vpt.lrv.lt/lt/powerbi/kovai-su-covid-19-sudarytos-sutartys.

380 Colman, Romina. "How COVID-19 and Collective Intelligence Transformed Procurement Risks into Opportunities for Transparency in Ecuador." *Open Contracting Partnership* (blog), December 20, 2020. https://www.open-contracting.org/2020/12/20/how-covid-19-and-collective-intelligence-transformed-procurement-risks-into-opportunities-for-transparency-in-ecuador/.

effective in making procurement efforts available online in several countries (see the example in the vignette below), pointing to other specific outcomes seems more difficult.

Vignette: Opening up data on emergency procurement in Ecuador

Ecuador managed to act quickly during the pandemic by publishing information on all emergency contracts as timely open data, thus increasing both civil society and public officials' capacity to tackle the emergency, as well as residents' awareness of public emergency procurement. Over 50 legal lawsuits have been filed, and nearly 24,000 officials have been instructed in how to handle emergency procurement more effectively and transparently.[381] The data also allowed civil society organizations to establish a public procurement observatory, aimed to issue recommendations on how to make buying processes more efficient.[382]

Since launching their initial public procurement platform, SERCOP has maintained their efforts to increase transparency and make it easier for both civil society and citizens to oversee and track what is happening regarding public contracts. An example of this is the real-time "citizens' dashboard" tool, created by SERCOP based on their initial platform to monitor emergency buying.[383]

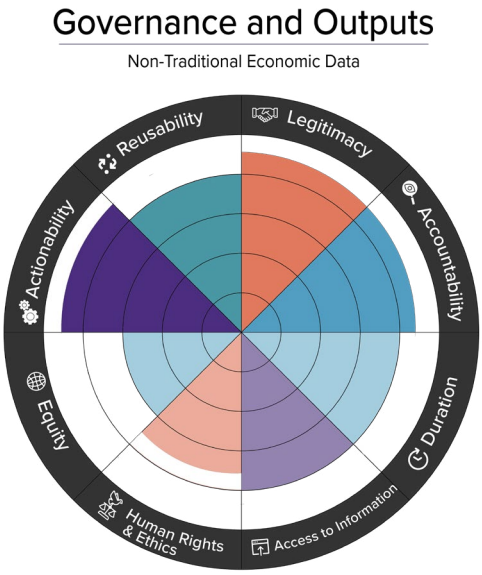


Figure 13. Assessment: Non-Traditional Economic Data, Governance and Outputs
In this figure, we demonstrate our assessment of the governance and outputs of this initiative using our assessment framework.

381 Ibid.
382 Ibid.
383 Ibid.

2.3.5. Additional Examples

Data source	Additional Examples
Credit card and debit card transactions	• The US Department of Commerce's Bureau of Economic Analysis and Fiserv: The US Department of Commerce's Bureau of Economic Analysis is using card transaction data obtained from the global financial services tech company Fiserv to assess economic health amid the pandemic. • Opportunity Insights: a nonprofit based at Harvard University used anonymous data from private companies – credit card and payroll processors to track small business activity, consumer spending on a daily basis.
Open contracting	• Open procurement in Paraguay: Open data has helped Paraguay's National Public Procurement Agency (DNCP) to create an early warning system with real-time monitoring of all emergency procurement processes. Publishing open data on public contracts has allowed journalists, activists, and civil society organizations to monitor COVID-19 procurement closely and highlight any irregularities. • Moldova contract tracking dashboard: The tool was launched thanks to a collaboration between government and civil society, and it is aimed at increasing "...transparency about the use of public money is guaranteed, increasing the government's social responsibility and citizens' trust in government actions,"

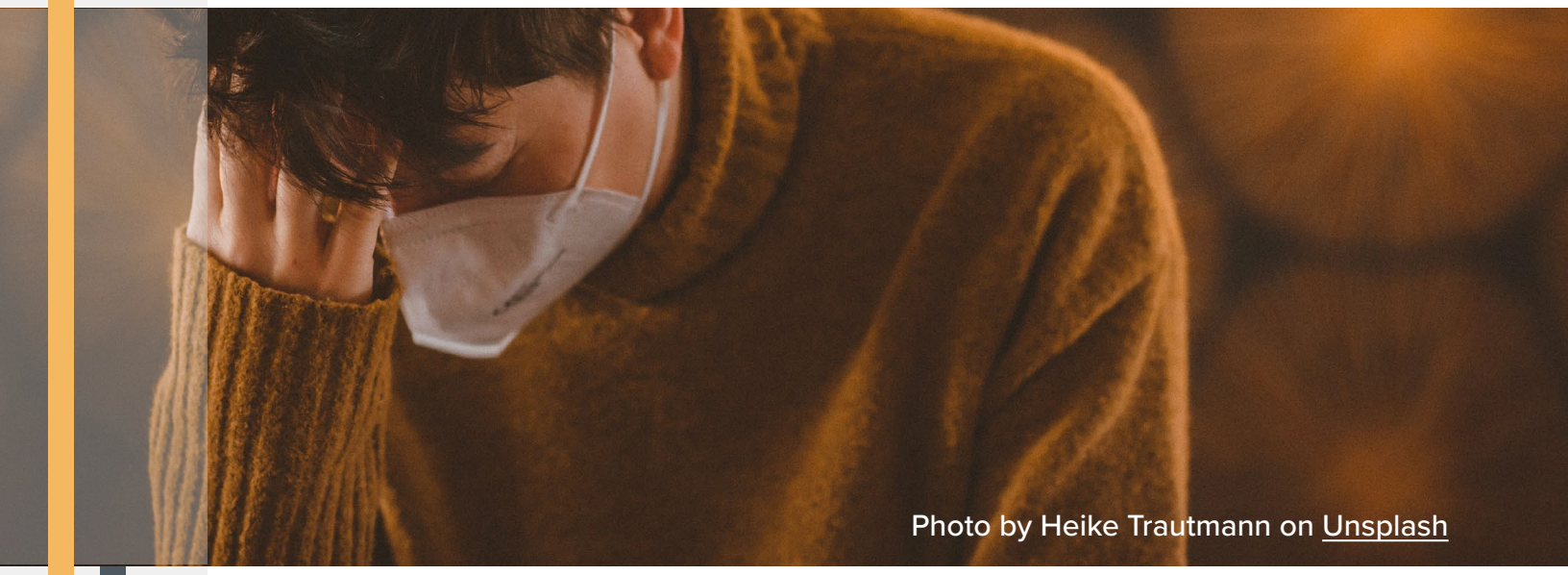


Photo by Heike Trautmann on Unsplash

2.4. Non-Traditional Sentiment Data

Main Takeaways

- **Main use cases:** Non-traditional sentiment data was mainly used to understand the public's perceptions and attitude about pandemic response measures, the impact of misinformation and supplement traditional data sources such as surveys.
- **Participation and Volunteerism increased in response to the crisis**: The public's willingness to participate in crowdsourcing initiatives increased significantly during the pandemic.
- **Open government practices supported crowdsourcing efforts:** Crowdsourcing initiatives during COVID-19 benefited from support from national governments which often took the form of guidance and funding.
- **Commercial tools played an important role in measuring sentiment:** Tools developed by the private sector for commercial purposes played an important role in helping to solve problems during COVID-19. Social listening tools, for example, were mostly commercial tools that were used leading to actionable insights.

2.4.1. Problem at Hand

To earn public trust and address misinformation, government agencies need to understand how individuals feel about the pandemic, the response to it, and its related policies. This understanding can help officials design targeted public communications, distribute information more effectively, and make sure that critical services are reaching those who need them most.

To address this need, government officials and agencies have increasingly turned to "sentiment data"—data that captures how people feel about events, issues, and more. As societies move through different stages of the pandemic, online information seeking and sharing patterns can provide valuable clues about people's concerns and how their attitude is changing in response to the pandemic. While this type of work has been in practice for quite some time, it was accelerated as a result of the pandemic.[384]

Public sentiment in response to COVID-19 changed rapidly, and this shift was driven by the widespread panic and uncertainty about the virus, as well as by a desire to be informed about and protect oneself from the disease.[385] In order for policymakers to develop relevant and effective policies to tackle the emergency, it was important to identify ways of understanding people's sentiments in relation to the challenges posed by COVID-19. Knowing people's attitudes toward the pandemic and the containment measures was indeed a valuable element to design and adjust government action and policy, helping policymakers to "figure out how and when to reopen schools and stores and how to get people vaccinated", among other things.[386]

2.4.2. Data sources

Non-traditional sentiment data is a general term referring to any type of data that captures how people feel about a topic. This data can come from a variety of sources, including social media posts and crowdsourcing platforms. Key sources of non-traditional sentiment data used in the pandemic response are described below:

---

384 "My G7." n.d. Eimdemo.com. Accessed September 28, 2022. http://magellan.eimdemo.com/g7/en/index.html.

385 World Health Organization. "Social Listening: Finding the Signal Through the Noise," August 18, 2021. https://www.who.int/news/item/18-08-2021-social-listening-finding-the-signal-through-the-noise.

386 James F. Smith. "Tracking attitudes and behavior on COVID in all 50 states, week by week". February 23, 2021. Harvard Kennedy School. https://www.hks.harvard.edu/faculty-research/policy-topics/health/tracking-attitudes-and-behavior-covid-all-50-states-week-week.

Social media data: Social media data, gathered from social media platforms like Twitter, Facebook, and Instagram, can provide insights into how people are feeling about different aspects of the pandemic and the response to it. Social media data is generally more detailed and timely than online search data such as Google Trends, but it can be more difficult to obtain and analyze.

- ***Example:*** *CivicLytics, a South American initiative launched by IDB Group, used data from social media and applied AI to understand how people were feeling about the pandemic and the response to it. The data has been obtained from public posts, using the Twitter API and aggregate sources of available information such as forums, and message boards. All posts anonymized removing personal identifiable information such as names. Opinions are presented in summary form only without any raw text shared publicly. The data gathered from social media was contrasted with survey results and structured reports. This data was then used to inform government decision-making around the pandemic.*[387]

Tag: Response, Recovery

Photo by Bernie Almanzar on Unsplash

387 "CivicLytics (Big Data Analytics Cívica)." CivicLytics (big data analytics cívica). Accessed September 22, 2022. https://bidciviclytics.citibeats.com/#/en.

Figure 14. Public sentiment analysis in South America
This graphic from CivicLytics illustrates public sentiment in South American countries regarding various social issues during the pandemic.[388]

Crowdsourced Data: Crowdsourced data, via mobile apps or online surveys, can also provide insights into how people are feeling. These data are generally most useful for understanding how people feel about specific issues or policies. Platforms for data analytics like Kaggle have begun crowdsourcing data related to the pandemic and the response to it.[389] This data can be used to understand the pandemic at a granular level and identify patterns or trends that would be otherwise difficult to discern. Crowdsourced data can provide more detailed and specific insights, but it may not be as timely as social media data.

- ***Example:*** *CoMix, is an initiative coordinated by Hasselt University & University of Antwerp, Belgium and funded by the European Commission, used crowdsourcing tools and data analytics to quickly gather public sentiment data from across the EU.*[390] *Participants were recruited through social media ads and email campaigns*

388 Ibid.

389 Vermicelli, Silvia, Livio Cricelli, and Michele Grimaldi. "How Can Crowdsourcing Help Tackle the COVID-19 Pandemic? An Explorative Overview of Innovative Collaborative Practices." *R&D Management* 51, no. 2 (2021): 183–94. https://doi.org/10.1111/radm.12443.

390 Hasselt University. "The CoMix Study," 2022. https://www.uhasselt.be/en/aparte-sites-partner-en/epi-pose/the-comix-study.

by the market research company Ipsos MORI. Data collected from this study was used to inform government decision-making regarding the COVID-19 pandemic.[391] *It was first conducted in March 2020 in Belgium, the Netherlands, and the United Kingdom and had been running every week during the pandemic. Later additional 17 European countries were included in the study.*[392]

Tag: *Response*

Photo by Kate Trifo on Unsplash

391 "An infodemic is too much information including false or misleading information in digital and physical environments during a disease outbreak. It causes confusion and risk-taking behaviours that can harm health. It also leads to mistrust in health authorities and undermines the public health response." "Infodemic." Accessed September 29, 2022. https://www.who.int/health-topics/infodemic;
Gimma, Amy, James D. Munday, Kerry L. M. Wong, Pietro Coletti, Kevin van Zandvoort, Kiesha Prem, CMMID COVID-19 working group, et al. "Changes in Social Contacts in England during the COVID-19 Pandemic between March 2020 and March 2021 as Measured by the CoMix Survey: A Repeated Cross-Sectional Study." Edited by Megan B. Murray. *PLOS Medicine* 19, no. 3 (March 1, 2022): e1003907. https://doi.org/10.1371/journal.pmed.1003907.

392 Jarvis, Christopher. "The CoMix Study." UHasselt, 2022. https://www.uhasselt.be/en/aparte-sites-partner-en/epipose/the-comix-study.

Vignette: Social Listening in Eastern and Southern Africa

Online platforms, such as social media and instant messaging applications, have contributed to the spreading of false information about COVID-19 and fuelled the 'infodemic'.[393] During the COVID-19 pandemic, there was a need for real-time information on how the virus was spreading and what people's concerns and needs were.[394] To help address this issue, UNICEF worked to build social listening capacity at the regional and country levels.[395] Social listening is the process of monitoring online conversations to identify trends and insights about specific topics or issues.[396] It can be used to inform decision-making in emergency situations like the COVID-19 pandemic, as well as for preparedness and response planning.

UNICEF worked with the Ministry of Public Health in Madagascar, and a number of other African countries such as Kenya, Malawi, Comoros, and Zambia, to support an emergency response.[397] The goal of this project was to collect social media data and analyze the sentiment of the public about the COVID-19 response in order to improve the effectiveness of it. This work included identifying narratives of misinformation, monitoring online conversations for trends and issues related to the response effort, and providing support to countries in incorporating social listening data into their decision making.[398]

---

393 Cinelli, Matteo, Walter Quattrociocchi, Alessandro Galeazzi, Carlo Michele Valensise, Emanuele Brugnoli, Ana Lucia Schmidt, Paola Zola, Fabiana Zollo, and Antonio Scala. "The COVID-19 Social Media Infodemic." *Scientific Reports* 10, no. 1 (October 6, 2020): 16598. https://doi.org/10.1038/s41598-020-73510-5.

394 World Health Organization. "Social Listening: Finding the Signal Through the Noise," August 18, 2021. https://www.who.int/news/item/18-08-2021-social-listening-finding-the-signal-through-the-noise.

395 Sommariva, Silvia, Jenna Mote, Helena Ballester Bon, Herisoa Razafindraibe, Domoina Ratovozanany, Vanou Rasoamanana, Surangani Abeyesekera, et al. "Social Listening in Eastern and Southern Africa, a UNICEF Risk Communication and Community Engagement Strategy to Address the COVID-19 Infodemic." *Health Security* 19, no. 1 (February 1, 2021): 57–64. https://doi.org/10.1089/hs.2020.0226.

396 Chaney, Sarah Cunard, Peter Benjamin, and Patricia Michael. "Finding the Signal through the Noise. A Landscape Review and Framework to Enhance the Effective Use of Digital Social Listening for Immunisation Demand Generation.," June 2021, 44.

397 Sommariva, Silvia, Jenna Mote, Helena Ballester Bon, Herisoa Razafindraibe, Domoina Ratovozanany, Vanou Rasoamanana, Surangani Abeyesekera, et al. "Social Listening in Eastern and Southern Africa, a UNICEF Risk Communication and Community Engagement Strategy to Address the COVID-19 Infodemic." Health Security 19, no. 1 (February 1, 2021): 57–64. https://doi.org/10.1089/hs.2020.0226.

398 Ibid.

2.4.3. Design and Planning

Social Media Data

Social media data has been used by multilateral organizations such as the United Nations and World Health Organization and governments to support their response to COVID-19 and understand the public attitude. Data analytics is conducted on various online sources, including Twitter posts and forums, to study the trends and spread of misinformation and sentiment analysis about specific events flagged by the World Health Organization.[399] The analysis has been conducted on an ongoing basis and shared through UN Global Pulse weekly reports. The results have been used by the World Health Organization to issue recommendations and solutions for timely course correction of communications and for better targeted engagement strategies.[400]

The two main topics that were studied by researchers early on in 2020 were misinformation and understanding public sentiment about the pandemic. The focus later shifted to using social media data to measure the social consequences of COVID-19 and how vaccine perceptions have changed.[401]

Social media data was often collected through web scraping or via application programming interfaces (APIs) which allow for automated data collection.[402] This data was then used to generate maps of misinformation, analyze the sentiment around the pandemic measures, and understand how people coped with the pandemic.[403]

Securing access from social media platforms is not immediately possible for many organizations, and to scrape this data properly off the platforms can be complex and time

399 World Health Organization. "Social Listening: Finding the Signal Through the Noise," August 18, 2021. https://www.who.int/news/item/18-08-2021-social-listening-finding-the-signal-through-the-noise.

400 Pulse Lab New York. "Understanding the COVID-19 Pandemic in Real-Time." *UN Global Pulse* (blog). Accessed September 21, 2022. https://www.unglobalpulse.org/project/understanding-the-covid-19-pandemic-in-real-time/.

401 Nathalia. "How Researchers Studied COVID-19 on Twitter." Twitter - Developer Platform, March 16, 2021. https://developer.twitter.com/en/blog/community/2021/how-researchers-studied-covid-19-on-twitter.

402 Lan, Hai, Dexuan Sha, Anusha Srirenganathan Malarvizhi, Yi Liu, Yun Li, Nadine Meister, Qian Liu, Zifu Wang, Jingchao Yang, and Chaowei Phil Yang. "COVID-Scraper: An Open-Source Toolset for Automatically Scraping and Processing Global Multi-Scale Spatiotemporal COVID-19 Records." *IEEE Access* 9 (2021): 84783–98. https://doi.org/10.1109/ACCESS.2021.3085682.

403 Purnat, Tina D., Paolo Vacca, Christine Czerniak, Sarah Ball, Stefano Burzo, Tim Zecchin, Amy Wright, et al. "Infodemic Signal Detection During the COVID-19 Pandemic: Development of a Methodology for Identifying Potential Information Voids in Online Conversations." *JMIR Infodemiology* 1, no. 1 (July 28, 2021): e30971. https://doi.org/10.2196/30971.

consuming without support from an outside organization.[404] Data collected from social media platforms is often unstructured. The collected data then needs to be cleaned and organized before they can be analyzed. Importantly, the analysis will also require machine learning and natural language processing (NLP) techniques. This process can be time-consuming and requires a certain amount of technical expertise.[405] In addition, the methods used to collect and analyze social media data such as web scraping are not without ethical concerns as they can violate the terms of service of these platforms and the privacy of users.[406] Since the Cambridge Analytica scandal in 2018, many social media platforms have restricted access to data and made it more difficult for researchers and developers to collect data through their API. This has made it harder for researchers to collect social media data around several topics, including COVID-19.[407]

Crowdsourced Data

Crowdsourcing refers to the process of collecting data, ideas or obtaining services from a large number of people through the Internet.[408] Crowdsourcing efforts can be administered using different types of tools depending on the intended outcome. These tools include surveys, challenges and innovation contests, hackathons, and online collaboration tools.[409] Platforms like Challenge.com, Battle of Concepts, HeroX, InnoCentive, JOGL, Kaggle, and Innoget have hosted challenges and data related to COVID-19. Also, new platforms have been developed with the purpose of hosting COVID-19 challenges.[410] The datasets made

404 Several companies, such as Dataminr, do this work as part of their core business model. They conducted analysis throughout the COVID-19 pandemic. See: Dataminr. "Dataminr Report: COVID-19 U.S. Study." Accessed October 5, 2022. https://www.dataminr.com/covid19-us..

405 Lan, Hai, Dexuan Sha, Anusha Srirenganathan Malarvizhi, Yi Liu, Yun Li, Nadine Meister, Qian Liu, Zifu Wang, Jingchao Yang, and Chaowei Phil Yang. "COVID-Scraper: An Open-Source Toolset for Automatically Scraping and Processing Global Multi-Scale Spatiotemporal COVID-19 Records." *IEEE Access* 9 (2021): 84783–98. https://doi.org/10.1109/ACCESS.2021.3085682.

406 Bhatt, Paras, Naga Vemprala, Rohit Valecha, Govind Hariharan, and H. Raghav Rao. "User Privacy, Surveillance and Public Health during COVID-19 – An Examination of Twitterverse." *Information Systems Frontiers*, January 31, 2022. https://doi.org/10.1007/s10796-022-10247-8.

407 Bastos, Marco, and Shawn T. Walker. "Facebook's Data Lockdown Is a Disaster for Academic Researchers." *The Conversation*, April 11, 2018. http://theconversation.com/facebooks-data-lockdown-is-a-disaster-for-academic-researchers-94533.

408 Colovic, Ana, Annalisa Caloffi, and Federica Rossi. "Crowdsourcing and COVID-19: How Public Administrations Mobilize Crowds to Find Solutions to Problems Posed by the Pandemic." *Public Administration Review* 82, no. 4 (2022): 756–63. https://doi.org/10.1111/puar.13489.

409 Vermicelli, Silvia, Livio Cricelli, and Michele Grimaldi. "How Can Crowdsourcing Help Tackle the COVID- 19 Pandemic? An Explorative Overview of Innovative Collaborative Practices." *R&D Management* 51, no. 2 (March 2021): 183–94. https://doi.org/10.1111/radm.12443.

410 Yadav, Gaurav. "COVID-19 Huge Resources Collection." Kaggle, 2020. https://www.kaggle.com/general/197225;Colovic, Ana, Annalisa Caloffi, and Federica Rossi. "Crowdsourcing and COVID-19: How Public Administrations Mobilize Crowds to Find Solutions to Problems Posed by the Pandemic." *Public Administration Review* 82, no. 4 (2022): 756–63. https://doi.org/10.1111/puar.13489.

available through challenges can be used by anyone to develop predictive models and analyze the existing data to look for public sentiment trends and outliers. The sponsors for these challenges are usually governments, organizations or businesses that are looking for ways to use data to understand public sentiment.[411]

Governments have used crowdsourcing to conduct surveys and collect data to measure public sentiment. Statistics Canada, for instance, used a web-based crowdsourcing application, a program that allows people to access and interact with online content, in an effort to gather information from citizens about their opinion regarding the government response to COVID-19.[412] Other Canadian government agencies have used the data and findings from the survey in order to effectively respond to the needs of their citizens.

Crowdsourcing projects and hackathons were organized in collaboration between multilateral organizations and private companies. One example is the CodeTheCurve hackathon, which is an online, immersive environment for two weeks in April 2020, where nearly 200 teams from various countries met with expert mentors to create data models and digital prototypes for deployable solutions. One of the important themes for the hackathon was misinformation and disinformation surrounding the COVID-19 pandemic. The hackathon was created in collaboration with UNESCO, IBM, and SAP.[413]

Vignette: Social Listening in Eastern and Southern Africa

Social listening is a resource-intensive process and can be time-consuming when dealing with a rapidly changing social media environment. Therefore the initiative needed to be incorporated into UNICEF Eastern and Southern Africa Regional Office (ESARO)'s crisis response strategy to support Risk communication and community engagement (RCCE) on COVID-19. The goal was to provide guidance to country offices on social media monitoring

411 Colovic, Ana, Annalisa Caloffi, and Federica Rossi. "Governments' Use of Crowdsourcing during COVID-19: Some Insights from a Study of Crowdsourcing Challenges." *Centre for Innovation Management Research* (blog), March 16, 2022. http://www7.bbk.ac.uk/cimr/2022/03/16/governments-use-of-crowdsourcing-during-covid-19-some-insights-from-a-study-of-crowdsourcing-challenges/.

412 Statistics Canada. "Crowdsourcing: Impacts of COVID-19 on Canadians' Trust in Others Public Use Microdata File," July 15, 2021. https://www150.statcan.gc.ca/n1/pub/45-25-0005/452500052020001-eng.htm.

413 UNESCO. "CodeTheCurve: Youth Innovators Hacking Their Way to Surmount COVID-19 Challenges," April 21, 2020. https://www.unesco.org/en/articles/codethecurve-youth-innovators-hacking-their-way-surmount-covid-19-challenges.

activities applied to the COVID-19 response."[414]

The initiative had three objectives related to the COVID-19 response: "(1) mitigate potential effects of misinformation on adoption of recommended behaviors, demand, and uptake of services by implementing a system to track, analyze, and manage COVID-19-related rumors at the regional level; (2) inform the design and implementation of high-quality digital communication for content that meets information needs and responds to concerns and rumors shared by different audiences; and (3) reinforce country-level capacity in rumor tracking and management by providing guidance to country offices on social media monitoring activities applied to the COVID-19 response."[415]

The data was collected from various sources, such as social media, radio, TV, newspapers, and SMS. Data collected was analyzed by the Information Watch and Rumor Mitigation working group within the National Communication Sub-Committee for the Fight Against Epidemics."[416]

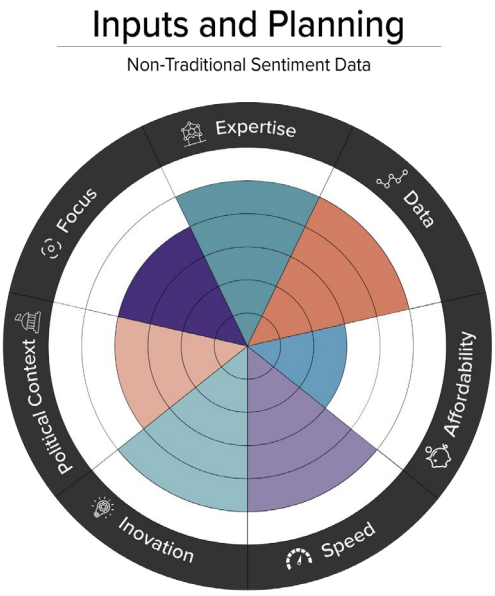


Figure 15. Assessment: Non-Traditional Sentiment Data, Inputs and Planning
The figure above demonstrates our assessment of the design of this initiative using our assessment framework.

414 Sommariva, Silvia, Jenna Mote, Helena Ballester Bon, Herisoa Razafindraibe, Domoina Ratovozanany, Vanou Rasoamanana, Surangani Abeyesekera, et al. "Social Listening in Eastern and Southern Africa, a UNICEF Risk Communication and Community Engagement Strategy to Address the COVID-19 Infodemic." *Health Security* 19, no. 1 (February 1, 2021): 57–64. https://doi.org/10.1089/hs.2020.0226.

415 Ibid.

416 Ibid.

2.4.4. Implementation and outcomes

Social media data

Commercial tools, platforms developed by private companies for commercial purposes, played an important role in helping to solve problems during the pandemic. Social listening tools, which gather data from social media platforms, were mostly commercial tools that were used by organizations like WHO and UNICEF, leading to actionable insights.[417] These tools were initially developed by private companies for marketing purposes, to understand their consumers' needs.[418] Social listening tools are powerful, but they have limitations, mostly due to their heavy reliance on Twitter and Reddit as the data sources. For example, it is difficult to track conversations taking place on Facebook or Instagram due to restrictive access to the private data. In addition, the Twitter population in the United States, the largest user base of the platform, is relatively younger, wealthier, and more educated than the average American.[419] This makes it difficult to use social listening as the sole source of information. To get a complete picture, this data source needs to be supplemented with other research methods such as traditional surveys, or NTD sources like crowdsourcing and online search data.[420]

To get a complete picture, this data source needs to be supplemented with other research methods such as traditional surveys, or NTD sources like crowdsourcing and online search data.

Social media platforms started sharing COVID-19 related data and tools in April 2020.[421] Since then, social media platforms have become more willing to provide access to data through various initiatives such as Data for Good, creating new APIs or changing their existing APIs to make them more accessible.[422] These changes may

417 Ibid.

418 Social Listening: Covid-19, Social Media, and The Path to a Better Safety Net. (2021). Retrieved 20 July 2022, from "Social Listening: Covid-19, Social Media, and The Path to a Better Safety Net." *Aspen Institute*, January 2021. https://www.aspeninstitute.org/wp-content/uploads/2021/01/AspenFSP_Social-Listening_2021.pdf.

419 Hirose. "24 Twitter Demographics That Matter to Marketers in 2022." *Social Media Marketing & Management Dashboard* (blog), September 20, 2022. https://blog.hootsuite.com/twitter-demographics/.

420 Bari, Anasse, Aashish Khubchandani, Junzhang Wang, Matthias Heymann, and Megan Coffee. 2021. "COVID-19 Early-Alert Signals Using Human Behavior Alternative Data." Social Network Analysis and Mining 11 (1): 18. https://doi.org/10.1007/s13278-021-00723-5.

421 Tornes, Adam. "Enabling Study of the Public Conversation in a Time of Crisis." *Twitter Developer Platform Blog* (blog), April 29, 2020. https://blog.twitter.com/developer/en_us/topics/tools/2020/covid19_public_conversation_data.

422 Meta. "COVID-19 RESPONSE. We Have Tools That Can Help Organizations Respond to the COVID-19 Pandemic." *Data for Good*, 2022. https://dataforgood.facebook.com/dfg/covid-19.

have been driven by new privacy policies during COVID-19,[423] adaptations to existing policies in the United States and European Union, or rapid guidance within the existing open data policies.

As demonstrated by the previous examples, while sentiment data from social media platforms proved valuable in gathering data about public sentiment, the existence of this data accelerated the spreading of online mis-/dis-information and social media companies did not address it effectively. Researchers at the Center for Digital Social Research, Aarhus University identified six types of health-related misinformation during COVID-19: conspiracy theories, fake cures, political measures, fake news about vaccines, fake news about the virus, and other health related disinformation.[424]An analysis of 12 leading social media and messaging platforms' responses to COVID-19 misinformation found that most platforms stated that they prohibited COVID-19 misinformation, but their responses lacked clarity and transparency.[425] The outbreak of COVID-19 and the use of social media to disseminate information about the disease have led to a rapid spread of misinformation.

An analysis of 12 leading social media and messaging platforms' responses to COVID-19 misinformation found that most platforms stated that they prohibited COVID-19 misinformation, but their responses lacked clarity and transparency.

Crowdsourced Data

There has been an increase in the availability of crowdsourcing platforms sponsored by governments, private sector, and academia. This growth has been spurred, in part, by collaborations with international organizations, such as WHO, to combat COVID-19.[426] The increasing number of platforms is also fueled by increased willingness from the public to

423 Chen, Adrian. "Building a Data Stream to Assist with COVID-19 Research." Twitter - Engineering, March 16, 2021. https://blog.twitter.com/engineering/en_us/topics/insights/2021/how-we-built-a-data-stream-to-assist-with-covid-19-research.

424 Charquero-Ballester, Marina, Jessica G Walter, Ida A Nissen, and Anja Bechmann. "Different Types of COVID-19 Misinformation Have Different Emotional Valence on Twitter." Big Data & Society 8, no. 2 (July 2021): 205395172110412. https://doi.org/10.1177/20539517211041279.

425 Krishnan, Nandita, Jiayan Gu, Rebekah Tromble, and Lorien C. Abroms. "Research Note: Examining How Various Social Media Platforms Have Responded to COVID-19 Misinformation." *Harvard Kennedy School Misinformation Review*, December 15, 2021. https://doi.org/10.37016/mr-2020-85.

426 COVID-19 Global Hackathon 1.0. "COVID-19 Global Hackathon 1.0." Devpost, 2022. https://covid-global-hackathon.devpost.com/.

participate and volunteer for crowdsourcing initiatives, even in absence of clear language around data sharing and privacy.[427]

Crowdsourcing data for COVID-19 presents a number of challenges, including how to ensure volunteers respect data agreements, how to make different datasets interoperable, and where to share data. Throughout the pandemic, organizations have most frequently relied on trusted third parties to manage these challenges. These organizations, when deployed effectively, can make sure the data is ethically open and secure.[428]

According to a study of several global repositories, more than half of the COVID-19 related challenges across various platforms have been funded by national governments.[429] During the pandemic public administrations used crowdsourcing for four main purposes: e-procurement, release of open data, eliciting ideas, and open co-creation. Public administrations have used all four types of crowdsourcing, but open co-creation represents only a small share of all challenges.[430]

Crowdsourcing can also be an effective way to quickly find solutions to problems that arise during a crisis and build trust between the government and the public. These values were demonstrated when the Scottish Government used a crowdsourcing platform Dialogue to gather public opinion on appropriate response frameworks across Scotland during COVID-19.[431] The crowdsourcing exercise received over 4,000 ideas and 16,000 comments in one week. The government widely promoted the exercise and remained transparent throughout, updating participants on progress and findings.[432]

---

427 Bowser, Anne, Alison Parker, and Alex Long. "Citizen Science and COVID-19: The Power of the (Distanced) Crowd." Wilson Center. *Science and Technology Innovation Program* (blog), June 22, 2020. https://www.wilsoncenter.org/blog-post/citizen-science-and-covid-19-power-distanced-crowd.

428 Ibid.

429 Colovic, Ana, Annalisa Caloffi, and Rossi Federica. "Governments' Use of Crowdsourcing during COVID-19: Some Insights from a Study of Crowdsourcing Challenges." *Centre for Innovation Management Research* (blog), March 16, 2022. http://www7.bbk.ac.uk/cimr/2022/03/16/governments-use-of-crowdsourcing-during-covid-19-some-insights-from-a-study-of-crowdsourcing-challenges/.

430 Colovic, Ana, Annalisa Caloffi, and Federica Rossi. "Crowdsourcing and COVID-19: How Public Administrations Mobilize Crowds to Find Solutions to Problems Posed by the Pandemic." *Public Administration Review* 82, no. 4 (2022): 756–63. https://doi.org/10.1111/puar.13489.

431 Delib. "Customer Story: The Scottish Government," 2022. https://www.delib.net/dialogue/customer_stories/scottish_government.

432 Ibid.

Vignette: Social Listening in Eastern and Southern Africa

Between May and December 2020, UNICEF ESARO digital social listening tracked over 2,500 rumors, questions, or concerns about COVID-19 expressed by online users through social media or digital news comment spaces. The misinformation posts tracked by this initiative has generated at least half a million engagements on social media platforms. Insights have been disseminated through biweekly and monthly reports (20 weekly/biweekly reports and 8 monthly reports between May and December 2020) with national partners.[433]

In July and August 2020, an internal survey of UNICEF ESARO country offices found that most offices were using information from digital social listening activities to share with national partners or internally to support decision-making on COVID-19 and the broader response (e.g. adapting communication strategies).[434]

This data collected and analyzed through this initiative was used to update the communication plan, for example, through key messages that are regularly updated based on key findings from social listening.[435] In order to communicate the results of social listening analysis to all country offices and response partners, UNICEF ESARO has adopted a model that combines periodic reporting of social listening insights with interactive dashboards that display the results of the analysis of rumors, trends, and questions on a weekly and monthly basis.[436]

433 Sommariva, Silvia, Jenna Mote, Helena Ballester Bon, Herisoa Razafindraibe, Domoina Ratovozanany, Vanou Rasoamanana, Surangani Abeyesekera, et al. "Social Listening in Eastern and Southern Africa, a UNICEF Risk Communication and Community Engagement Strategy to Address the COVID-19 Infodemic." *Health Security* 19, no. 1 (February 1, 2021): 57–64. https://doi.org/10.1089/hs.2020.0226.

434 Ibid.

435 Adebisi, Yusuff Adebayo, Adrian Rabe, and Don Eliseo Lucero-Prisno III. "Risk Communication and Community Engagement Strategies for COVID-19 in 13 African Countries." *Health Promotion Perspectives* 11, no. 2 (May 19, 2021): 137–47. https://doi.org/10.34172/hpp.2021.18.

436 Sommariva, Silvia, Jenna Mote, Helena Ballester Bon, Herisoa Razafindraibe, Domoina Ratovozanany, Vanou Rasoamanana, Surangani Abeyesekera, et al. "Social Listening in Eastern and Southern Africa, a UNICEF Risk Communication and Community Engagement Strategy to Address the COVID-19 Infodemic." *Health Security* 19, no. 1 (February 1, 2021): 57–64. https://doi.org/10.1089/hs.2020.0226.

This initiative helped government agencies, such as the Ministry of Health, to provide regular updates on the virus to the public and disseminate messages addressing misinformation and important behaviors through social media channels (Facebook, Twitter, Instagram, and LinkedIn). Dedicated websites and phone lines have also been set up so that people can speak to a trained health educator and get accurate information about the virus and the vaccines.[437]

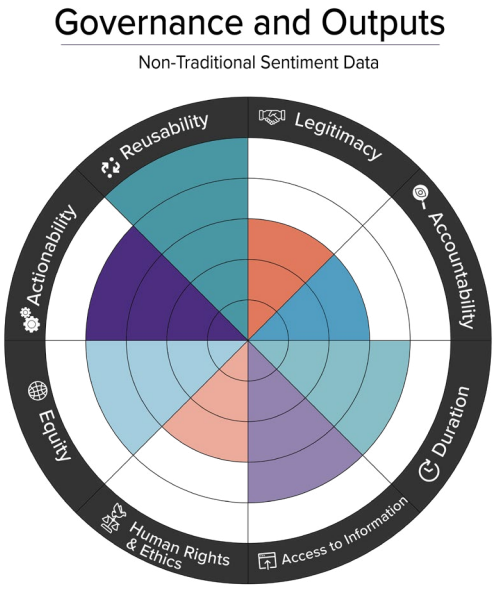


Figure 16. Assessment: Non-Traditional Sentiment Data, Governance and Outputs
In the figure above, we provide our assessment of the governance and outputs of this initiative using the assessment framework.

437 Adebisi, Yusuff Adebayo, Adrian Rabe, and Don Eliseo Lucero-Prisno III. "Risk Communication and Community Engagement Strategies for COVID-19 in 13 African Countries." *Health Promotion Perspectives* 11, no. 2 (May 19, 2021): 137–47. https://doi.org/10.34172/hpp.2021.18.

2.4.5. Additional Examples

Data source	Additional Examples
Social media data	• Online Citizen Perceptions: COVID-19 Pandemic - a platform to understand how people across Latin America are responding to the COVID-19 pandemic. • COVID19 Infodemics Observatory - a digital platform for the interactive visualization of the infodemic risk due to exposure to unreliable information on Twitter. • COVID-19 Reddit Algo-Tracker - The Citizens and Technology Lab at Cornell University has developed a service monitoring the algorithms on the social media platform Reddit during the COVID-19 pandemic.
Surveys and Crowdsourced data	• CoMix Study - study that follows households across Europe in real-time over the course of the COVID-19 pandemic to understand people's awareness, attitudes and behaviors in response to COVID-19 and measures how these change over time. • Sudan - mobile phone surveys - the World Bank and Central Bureau of Statistics jointly collected data through a mobile phone survey to assess the impact of COVID-19 on households. • Nigeria's Federal Government Response - Nigerian government experts with United Nations agencies conducted weekly national polls to assess residents' perceptions, and behaviors to help fine-tune the response. • Impacts of COVID-19 on Canadians - Statistics Canada conducted a survey on the impacts of the COVID-19 on Canadians, and the insights used by government organizations to evaluate the response.

3. Findings

Photo by Viki Mohamad on Unsplash

Our analysis of these briefings find several commonalities in how NTD was used throughout the different pandemic waves and lifecycle phases. Below we provide a summary of those findings. Our assessment of the findings together with the individual briefings and how they can be operationalized are explained in the following sections.

3.1. Why NTD was used during COVID-19

Generally, NTD aimed to complement rather than compete with traditional data. NTD aimed to fill gaps in traditional data methods, validate findings from traditional data methods, and respond quickly to a dynamic crisis. More specifically, NTD was used for syndromic surveillance (gathering data about COVID-19 symptoms), understanding mobility patterns, identifying the economic impacts of the pandemic, and collecting public sentiments. At the same time, given the limitations of specific types of NTD as it relates to representativeness, the collection and use of traditional data and crowdsourcing was often needed to ground truth the insights.

NTD initiatives were often developed with the purpose of generating access to data about populations that were not readily available through traditional data systems. While these NTD projects helped address equity related topics, they were often developed without the necessary social license to do so. This led to widespread concerns about ethics hindering trust in NTD initiatives.

3.2. NTD sources that stood out during COVID-19

Several NTD sources were used within the four data types throughout COVID-19. However, many of such data sources were not operationalized beyond proof of concept or pilot programs. Figure 17 summarizes the data sources that stood out to The GovLab research team throughout different waves of the pandemic based on the information gathered through the literature review, #Data4COVID19 Data Collaborative Repository, interviews, and Advisory Group meetings.

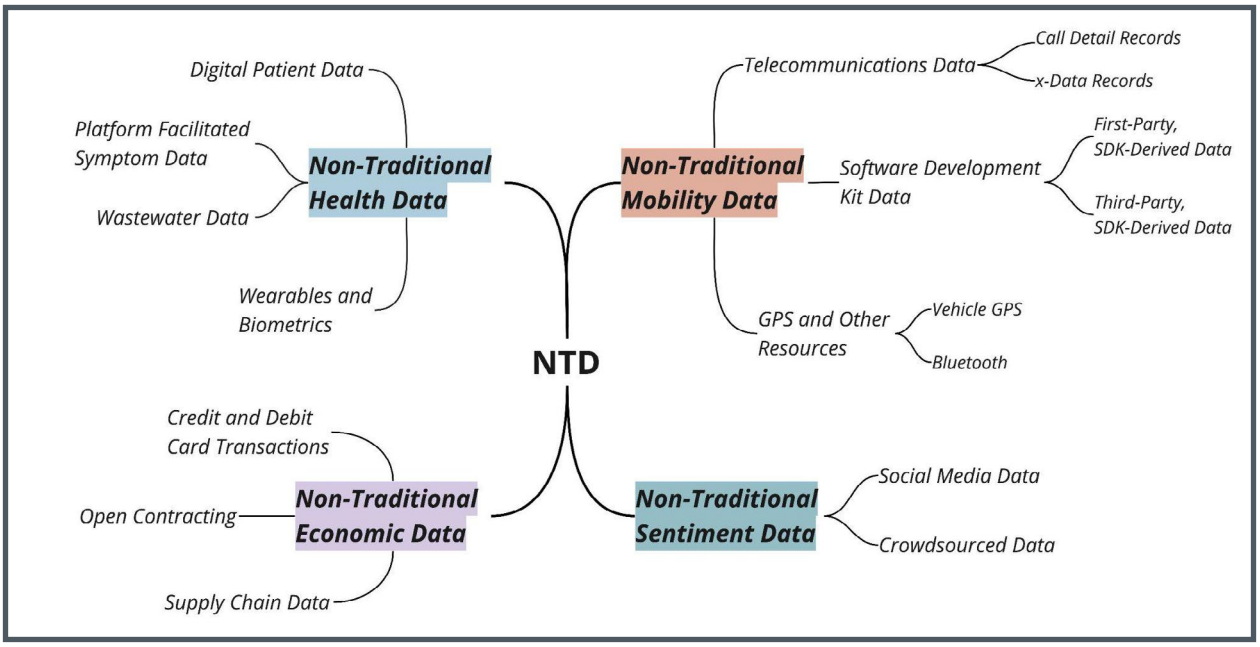


Figure 17. Data Sources
This figure summarizes the most prominent data sources within each briefing category.

3.3. Who called for NTD

The briefings and #Data4COVID19 Data Collaborative Repository demonstrated that three primary stakeholder groups sought NTD during COVID-19. However, there was limited clarification of the types of questions NTD could be used to effectively address.

- **Policy-makers:** This stakeholder group consists of policy-makers at all levels of government who called for NTD to inform their decisions about the crisis. These were the individuals who defined their government's approach to the pandemic and they existed at the:

- ***National and International Levels:*** National and international decision makers typically used NTD to supplement or validate data already collected through national level data programs (e.g. National Statistics Offices) or by public health authorities. This data was used within national and international pandemic response decisions and policy-making processes;
    - **Example:** *In September 2020, The United States Federal Government developed the National Wastewater Surveillance System (NWSS), a centralized database of wastewater data collected within municipal and state-level initiatives to be combined with clinical testing data to inform decision-making.*[438]

- ***Regional and City Level:*** Regional and city level decision-makers, sub-national leaders, looked to NTD to decide how and where policies and programs should be implemented. This work was often to supplement resources provided by national authorities or to supplement existing traditional health approaches, such as identifying specific neighborhoods that needed more testing.
    - **Example***: The New York Recovery Data Partnership developed to give New York City public agencies access to data that could directly support and feed into existing agency priorities as it related to the COVID-19 pandemic in New York City.*[439]

- **Health Researchers & Disease Specialists:** This stakeholder group includes members of the research and epidemiological community that used NTD to understand and address specific challenges related to COVID-19. This group's efforts were often directed toward developing publishable material that could reduce uncertainty around the pandemic and its manifestations.
    - **Example***: Using the COVID-19 Open Research Dataset (CORD-19) developed by the White House in collaboration with the Allen*

438 Centers for Disease Control and Prevention. "National Wastewater Surveillance System (NWSS)," March 21, 2022. https://www.cdc.gov/healthywater/surveillance/wastewater-surveillance/wastewater-surveillance.html.

439 The GovLab. "NYC Recovery Data Partnership." Data4COVID19, 2021. https://list.data4covid19.org/projects/nyc-recovery-data-partnership.

Institute for AI, Kaggle, and other partners, researchers globally undertook new COVID-19 research initiatives to advance the understanding of the virus and support decision-making.[440]

- **General Public (Non-Government Actors):** The general public, more broadly all other stakeholders outside of the government and research community, looked for NTD to inform research, business, and their own personal decisions related to COVID-19.
    - **Example**: *Various open-contracting initiatives emanated during the COVID-19 pandemic as the general public, including civil society and the media, urged governments to open up their procurement initiatives in a bid to allow more transparency on public spending.*[441]

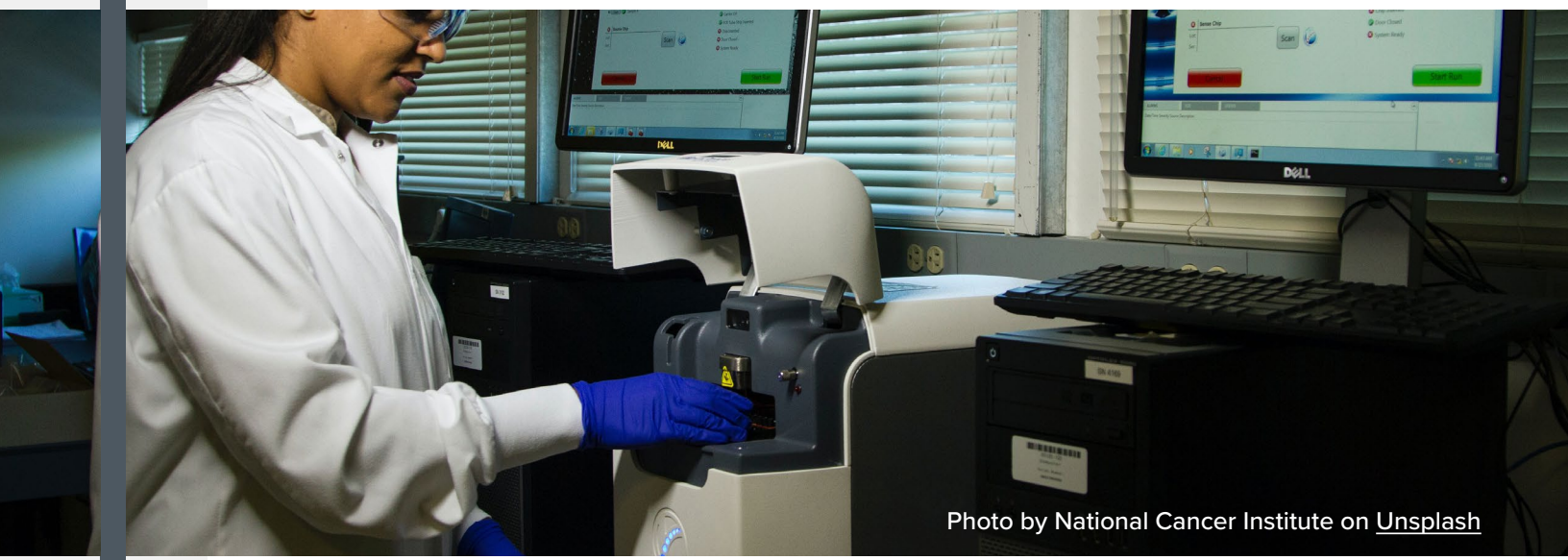


Photo by National Cancer Institute on Unsplash

440 Lu Wang, Lucy, Kyle Lo, Yoganand Chandrasekhar, Russell Reas, Jiangjiang Yang, Darrin Eide, Kathryn Funk, et al. "CORD-19: The Covid-19 Open Research Dataset." *ArXiv*, April 22, 2020, arXiv:2004.10706v2.

441 Open Contracting Partnership. "Emergency Procurement for COVID-19: Buying Fast, Smart, and Open," 2022. https://www.open-contracting.org/what-is-open-contracting/covid19/.

3.4. When NTD initiatives took place

Project launch date

Most applications of NTD for COVID-19 arose in the first few months after the World Health Organization declared the outbreak a pandemic. Of the 238 projects in The GovLab's #Data4COVID19 Data Collaborative Repository, 156 launched between March and May 2020.[442] Energy and enthusiasm for these projects declined sharply in the months that followed, with many efforts subsequently becoming inactive, poorly maintained, or inoperable. However, there were a few examples of projects being scaled and supported by governments.

In engagements with interviewees and this initiative's Advisory Group, participants cited "pandemic fatigue" as the reason behind this shift. As the crisis has continued, many organizations have decided against maintaining or launching initiatives that appear to add little value to their core interests. Private-sector companies who initially provided data out of a sense of corporate responsibility have seen their commitments stretch far beyond what they initially imagined or intended. Governments and civil society organizations have had their attention diverted by new crises around the world and have had to make decisions about where to apply their limited resources.[443] Members of the public who at first intensely engaged with dashboards, news analysis, and other resources have struggled to sustain interest over time.[444] Outside of a few select partnerships (e.g. the CrisisReady partnership between Direct Relief and Harvard) and already well-established Data for Good programs (e.g. those managed by Meta, Cuebiq, Safegraph, and Orange), there was a failure to maintain business and funding models to sustain this work.[445]

---

442 The GovLab. "A Living Repository for Data Collaboratives Seeking to Address the Spread of COVID-19." #Data4COVID19. Accessed September 14, 2022. https://list.data4covid19.org/.

443 Advisory Group Meeting. Zoom, September 9, 2022.

444 "Pandemic Fatigue - Reinvigorating the Public to Prevent COVID-19. Policy Framework for Supporting Pandemic Prevention and Management." Copenhagen: WHO Regional Office for Europe, 2020. https://apps.who.int/iris/bitstream/handle/10665/335820/WHO-EURO-2020-1160-40906-55390-eng.pdf.

445 "CrisisReady." Accessed September 29, 2022. https://www.crisisready.io/.

Pandemic phase

The pandemic life cycle includes readiness, response, recovery, and reform. However, there was an extremely large focus on responding to COVID-19 as it spread (200 of 238 projects) over preparing for the pandemic, recovering from it or reforming in response to it. This emphasis on pandemic response reflected a general lack of preparedness for the pandemic and the need for the rapid development of initiatives to understand and address specific consequences of it.

3.5. Where NTD was accessed and used

Examining the almost 240 projects represented in the #Data4COVID19 Data Collaborative Repository found significant disparities in where NTD efforts launched. Large, highly developed (often national) institutions in the Global North, were among the most common users of NTD.

- 51% of all data collaborative efforts took place in Europe or North America. Only 13 projects total specifically addressed sub-Saharan Africa.
- Only 13% of all data collaboratives had a local focus.
- Many of those projects that did launch tended to be connected with large, well-financed governments, companies, and research institutions.

3.6. How NTD was accessed

NTD was mainly made available through data collaborations. These data collaborations primarily took the form of:

- *Paired and Pooled Data Collaboratives:* These tended to be cross-sectoral arrangements that used the existing capabilities of stakeholder groups across different domains to generate new knowledge in the public interest. These partnerships were typically used to accelerate research timelines and to increase the data capacity.

- *Trusted Intermediaries:* Intermediaries from academia, government, and the private sector provided access to data to those who requested it or as a part of corporate social responsibility initiatives. These initiatives tended to be of smaller scale than data partnerships and have a specific focus.

These initiatives were frequently developed to respond to specific needs, but often contributed to problems later on. Many institutions did not have the systems and infrastructure in place for these collaborations to be sustainable and needed to rapidly develop several components during the implementation process.

Generating access to data through these types of collaborations typically involved several administrative components. First, these initiatives often included an emergency data-sharing agreement that stipulated the who (the parties involved), what (the data made available), where and when (under what conditions the data was available), why (the purpose for the collaboration), and how (the approach taken with the data). These frameworks, once established, allowed leaders to understand the roles and responsibilities of those involved in a data collaborative effort.

Second, initiatives included processes for researchers to access the data. Instead of providing broad, unrestricted access to sensitive or proprietary datasets, organizations sought to restrict use to only those datasets that they considered valuable and relevant. This process often involved a request form. In many cases, researchers had to have an intended use or project to access the data that had already been through an ethics review (sometimes conducted by an institutional review board).

Lastly, the initiatives had a dedicated data storage system that supported the types of data included. This data storage system might be controlled by the research team or, more frequently, by the data supplier to guarantee control over access.

4. Lessons Learned

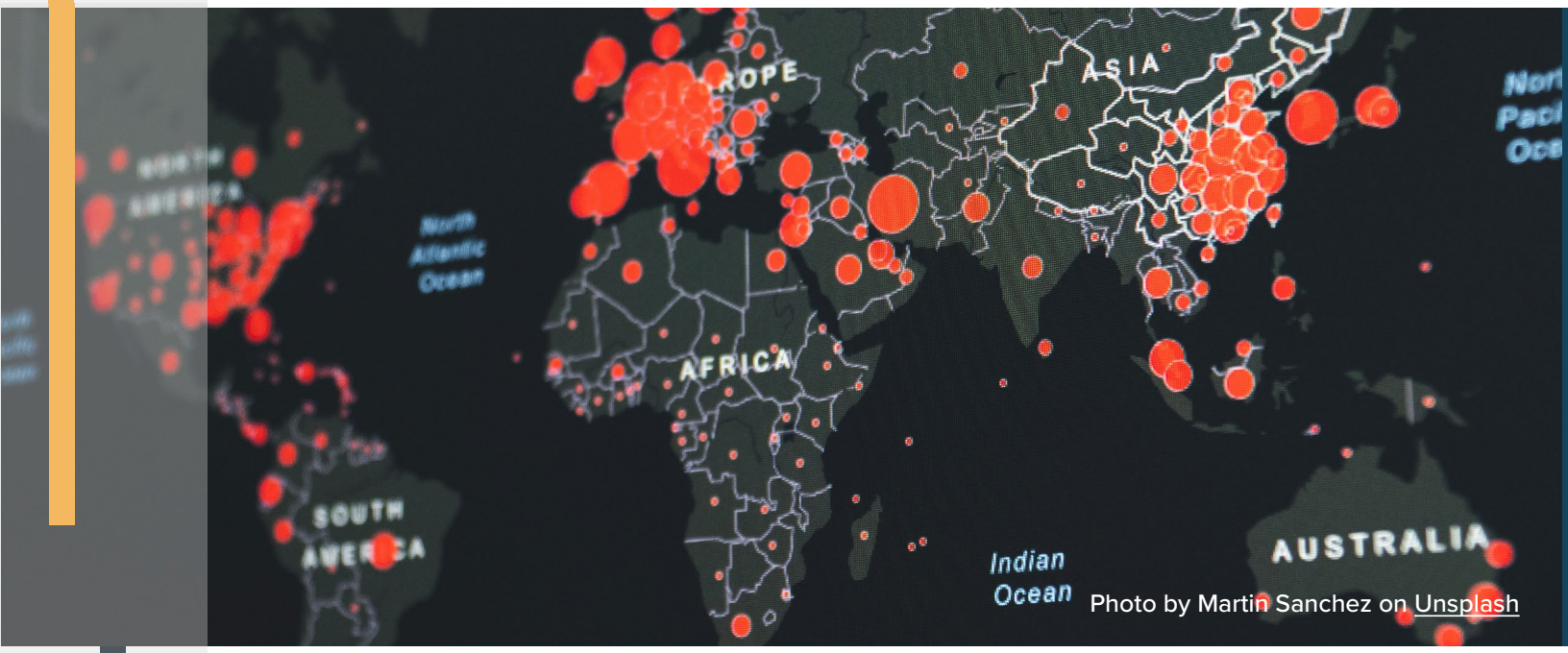


Photo by Martin Sanchez on Unsplash

The GovLab's analysis of the briefings along with the interviews, #Data4COVID19 Data Collaborative Repository, and Advisory Group meetings demonstrated the following ten lessons learned about how NTD initiatives were designed and implemented–across the four data types. We organize these takeaways into three general categories:

- **Public Input and Legitimacy**
    - **The Need for a Social License:** Organizations need a social license—the degree to which an initiative aligns with expectations of the general public and other stakeholders. Public and private sector organizations frequently developed NTD initiatives without the necessary social license, the degree to which an initiative aligns with expectations of the general public and other stakeholders. The speed of design and implementation was often prioritized over assessing possible human rights, ethics, and equity implications. While the reasons for this fact are understandable in context, organizations rarely had any frameworks—developed in consultation with the public—before the outbreak of the crisis to guide the activities they undertook.

- **The Need for a Clear Value Proposition:** When purpose and value proposition were clearly defined, trust and impact were higher and partners were better aligned. Articulating this in a way that was broadly understandable but applicable to a specific aspect of the crisis proved essential. Articulating a value proposition (and how it aligns with an organization's existing interests or can feed into other activities) demonstrated to be a crucial way of ensuring projects can be sustained long-term without becoming victim to "pandemic fatigue." If people know how an initiative is valuable, they will be more likely to support it long-term.

- **Design and Planning**

  - **The Need for Systematized Data Collaboratives:** Most collaboratives were developed haphazardly and in an ad hoc manner due to the lack of clear methods and tools. There was also an absence of dedicated centers of expertise in setting up a data collaborative around NTD. As a result many were short lived and limited in impact.

  - **The Need to Plan for Political Realities:** Political realities often determined what insights from NTD were used and which ones were deemed less relevant. Organizations needed to understand what decision-makers and the public needed as well as the geopolitical environment before jumping into action to produce insights that could meaningfully inform decisions. As noted by the Advisory Group, data holders also needed to be aware of how their work might be politicized in ways they do not intend. This includes understanding how the media is communicating NTD efforts with the public.

  - **The Need to Prioritize Equity and Data Rights:** Organizations often did not seem cognizant of the ways that benefits of NTD initiatives could be unevenly distributed, exacerbating pre-existing inequalities. For those that did recognize the importance of ethics and equity, they lacked a systematized approach in how ethics and equity would be integrated throughout the implementation phase. Organizations will need to prioritize equity and data rights to ensure they do not undermine either.

- **The Need for Crowdsourcing:** NTD developed non-traditional means for accommodating gaps in traditional methods–whether it be granularity or scale. Crowdsourcing through well-targeted surveys was needed to address these gaps.

- **Implementation and Outputs**

  - **The Need for Coordination:** There was little coordination between the projects leading to various instances of fragmentation and misalignment among NTD actors across the data lifecycle and pandemic phases. The Advisory Group explained that competition among data holders and NTD initiatives hindered coordination within and across countries. Many data holders aimed to be the first to publish some new insight instead of contributing to existing efforts, which resulted in an immense volume of outputs that did not address public needs. This development created logistical challenges in finding and analyzing data that could address public problems as relevant information could be buried or rendered duplicative by other efforts.

  - **The Need for an Evidence Base:** The absence of evidence base on the value of NTD for pandemic response hampered a more sophisticated use (and conversation). Organizations did not have models to follow and did not know in advance which types of data initiatives would yield the best results.

  - **The Need for Non-Personal Data:** Non-personal data–data that does not include personally identifiable information–provided rapid insights about the state of COVID-19 of different communities and populations. For example, as shown in the non-traditional health data briefing, the speed of implementing wastewater surveillance proved valuable in detecting COVID-19 outbreaks ahead of clinical testing while protecting personal privacy. Additional guidance is needed on how to implement these initiatives and combine them with personal data sources.

- **The Need to Continue Growing NTD Initiatives:** As cited by our Advisory Group, COVID-19 increased the maturity cycle, or speed of implementation and final use, of public sector digitization efforts. While the robustness and volume of COVID-19 NTD initiatives have declined, these efforts have accelerated new ways of working in the public sector. There is an opportunity to apply these initial investments to new use cases beyond the COVID-19 response.

5. Conclusion and Recommendations


Photo by Shubham Sharan on Unsplash

Overall, the use of data during the COVID-19 pandemic has been vital in understanding the spread of the virus and informing public health interventions. A variety of NTD sources have been used, including credit card transactions, telecommunications data, social media data, and wastewater data. Researchers have relied on these data sources to generate maps of risk factors, study the impact of social distancing measures, and understand how people have coped with the pandemic. COVID-19 presented a unique challenge that required the use of NTD on a scale never before seen. The successful use of data during the pandemic has demonstrated the potential for data to be used in other pandemics and crises from a local to global scale.

The use of data during the COVID-19 pandemic has informed a variety of policy interventions at the local, national, and international level. Decision makers have used data for several purposes including tracking the spread of the virus, studying the impact of social distancing measures, and understanding the socio economic factors that may increase the risk of infection. Data has also been used to inform the development of contact tracing apps, support the rollout of vaccines, and monitor compliance with quarantines. However, many NTD initiatives have had their challenges.

While many NTD initiatives have led to meaningful outcomes, others did not meet their potential. Several NTD initiatives have had skewed outcomes, been costly to implement, did not advance beyond proof of concept, or were not combined with traditional data when needed. Additionally, many initiatives relied upon the public to act on them and were not integrated in government response efforts.

There is an opportunity for decision makers to evaluate these successes and failures and take action towards effective and meaningful NTD-driven crisis management. Leveraging key findings from the briefings above, we have identified a set of recommendations that could advance the systematic and responsible use of NTD for crisis response. While these recommendations can be executed independently, we propose jointly enacting them to maximize the impact. These recommendations are tailored to public health crises, but can be applied to other types of crises at all scales (e.g. a natural disaster). The recommendations are categorized along the following four clusters:

1. ***Increasing evidence and awareness about the value proposition of NTD:*** There is a need for a stronger evidence base that can generate awareness of current NTD practices and support the value proposition of NTD during crisis situations. Creating a more data-driven approach to using NTD (that is, having more data about how to use data) can broaden NTD initiatives and ensure they can grow beyond current pilot programs and proofs of concept during future crises.

2. ***Advancing trust, ethics, and equity within NTD initiatives:*** There is a need for public and private organizations to establish trust among parties and with society on how NTD is being used. Prioritizing trust, ethics, and equity at the start of and throughout NTD initiatives can increase the legitimacy of its use and more meaningfully resolve crisis-driven challenges based on a combination of NTD and lived experience.

3. ***Strengthening collaboration and institutionalization of NTD uses:*** There is a need for multi-stakeholder partnerships to increase the data capacity and speed of implementation of NTD initiatives in a systematic and scalable way. Institutionalizing internal professional functions and external partnerships through data stewards can accelerate the responsible and efficient use of NTD during crises.

4. ***Preventing fragmentation, and improving readiness and coordination:*** There is a need for a more coordinated approach–across all crisis management efforts and functions–to address needs both as health systems evolve and as institutions face future public health emergencies. Minimizing fragmentation has the potential to increase readiness for future dynamic crises.

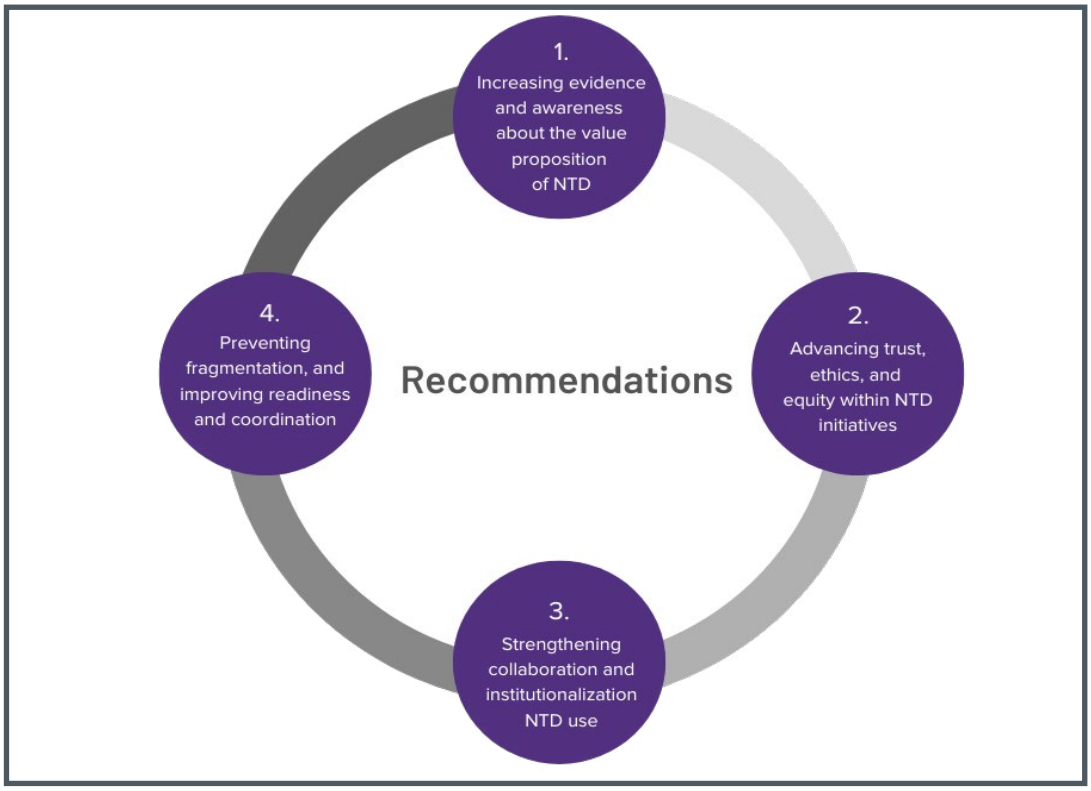


Figure 18. Recommendations Summary
This figure shows the four clusters of recommendations towards advancing the responsible and effective use of NTD during crisis situations.

5.1. Increasing evidence for and awaress of the value proposition of NTD for crisis response

There is a need for a stronger evidence base to support the value proposition and build awareness of the use of NTD for crisis response. Our and other studies have documented that NTD does have value in crisis decision-making, especially in increasing the specificity and timeliness of insights. However, more evidence could better define the value proposition of NTD and provide leaders with the information they need to act. Such evidence can also generate a more sophisticated understanding of its potential and challenges. In short, we need more data about which NTD data practices work.

This same evidence-gathering can support additional efforts needed to generate broader public awareness of current practices–including personal data protections–and successes to enable a more informed public debate. The value and challenges of NTD are often not communicated by the institutions that use it. This lack of awareness and transparency is likely one of the reasons contributing to the limited public and private support that could help scale the practice beyond current pilots.

Toward that end, we recommend a concerted set of efforts to generate and share evidence and awareness on the value of NTD initiatives during times of crisis. Possible (interconnected) pathways forward include

- ***Develop a DATA4CRISIS Alliance***: Build an alliance of diverse NTD holders that can share lessons learned and develop the value proposition collectively. This alliance could facilitate peer learning and expand the evidence base of what works under what conditions. This alliance could also orchestrate various outreach efforts to promote awareness of NTD's value proposition among the public through dedicated webinars, podcasts, and other communication channels. It could also establish standards and practices for stakeholders, assist in expediting data-sharing agreements, and provide resources to support collaborations in emergency settings.

- ***Develop a Global NTD Research Agenda***: Develop and strengthen the evidence on how NTD provides value at times of crisis by creating a cross-country, cross-institutional research agenda on NTD for crisis response. This agenda could be supported by new "crisis data fellows" hosted at different research organizations across sectors who could advance publications about the value proposition of NTD during crises.

- ***Develop broadly accepted metrics of success***: In collaboration with policy-makers, identify the metrics and criteria that NTD initiatives need to meet to be useful for decision-makers in crisis situations. Use these metrics to guide after-action reports and reviews documenting the factors contributing to NTD initiatives.

- ***Develop an NTD Observatory:*** Build a living repository of practices and examples of NTD initiatives. Such an observatory could provide broader transparency,

enable research on NTD's value and challenges, and act as a "broker" of proven useful methodologies in response to different types of crises. Such a collection of examples and models could also generate key insights on how NTD generated insights and informed decisions. Initiatives such as the Data Pop-Alliance's C-19 Global South Observatory, which seeks to monitor pandemic conditions and provide decision-makers with country-specific briefings based on ongoing developments, could be used as a model for this work.[446]

- ***Develop NTD Transparency Index:*** Designing and implementing a template for NTD data holders and users to use when sharing publicly, which NTD data was shared with whom, along with a more public record on who sought access to data and how those requests for data were handled.

5.2 Advancing trust, ethics, and equity within crisis NTD initiatives

There is a need for an increased emphasis on trust, ethics, and equity among parties and with society on how NTD is being used. Throughout COVID-19, public and private sector organizations frequently developed NTD initiatives without the necessary social license, the degree to which an initiative aligns with expectations of the general public and other stakeholders.[447] As described in the briefings, this was partly due to the urgency to respond to the crisis and led to several instances of data protections being temporarily dissolved. People had no way of providing input into projects and ensuring they were not exploitative or detrimental to their interests.

There is an opportunity for public and private organizations to more proactively develop principles, practices, and processes—guidance that can advance privacy, security, and equity through an intersectional lens. This includes defining clear roles and responsibilities for short-and long-term implementation. Additionally, legal protections and frameworks, including robust data protection mechanisms, should be included throughout the data value chain from data collection to transfer. These protections are needed even if they

446 Data-Pop Alliance. "C19 Global South Observatory," April 2, 2020. https://datapopalliance.org/covid19/c19globalsouthobservatory/.

447 Kenton, Will. "Social License to Operate (SLO)." Investopedia, May 31, 2021. https://www.investopedia.com/terms/s/social-license-slo.asp.

are not the same across countries. Such approaches could integrate public norms and expectations and foster data-driven insights more closely aligned with public needs and expectations, ultimately increasing the legitimacy of the initiative.

In what follows we suggest possible pathways towards advancing trust, ethics, and equity in future crisis NTD initiatives:

- ***Include Public Advocates in Contractual Negotiations:*** One way to ensure the public's interests are represented in data NTD collaborative arrangements (and to boost the legitimacy of said efforts) is to provide the public a seat at the table in discussions about how their data will be used. This seat can be provided by granting the public a designated representative in contractual negotiations. Organizations might seek input from officials in government (either elected or appointed), the leaders of particularly relevant civil society organizations, or some other well-trained and well-positioned figure.

- ***Establish Common Policy Frameworks:*** Co-design an ethics and equity framework—including responsible data principles—with key stakeholders to ensure local values and norms are integrated into the design, planning, and implementation of NTD initiatives. This process can be facilitated through consultative conversations and working sessions with key stakeholders and beneficiaries across sectors. Organizations might provide options for asynchronous contributions (e.g. public comment) to improve accessibility. They might also provide an open peer review for contributions to the framework.

- ***Institute Legislation or Regulation to Address Specific Crisis Context:*** In certain contexts, officials in the public sector might want to pursue legislative or administrative action to set standards around NTD in crisis scenarios. Any requirements should be developed with inclusion of the needs of the public, the risks and opportunities involved in NTD, and the incentives and disincentives that might motivate data users and suppliers.

- ***Publish the Contracts Impacting Data Reuse:*** To guarantee oversight and allow people to provide input into how their data is used by others, organizations might also seek to publish draft agreements in an online repository open for public

contributions. These repositories, such as those provided by the Contracts for Data Collaboration initiative, could ensure that organizations are clear and transparent about what it is they intend to do and what their collaboration is intended to address.[448] This publication could also minimize the opportunity for collaborative projects to be misinterpreted or misconstrued and increase their overall legitimacy.

- ***Conduct Regular Engagements to Understand Needs***: Identify public concerns and priorities on an ongoing basis through regular engagements—such as with Data Assembly mini-publics, surveys and public comment, and workshops with local officials involved in an effort. Develop an open repository of public concerns and use cases where there is public consensus that can be added to on an ongoing basis. Within the repository, policy makers can add supporting information about public needs to foster new initiatives, identify and support initiatives aligned with the public's needs, and clearly identify the social risks of the initiative (e.g. over-reliance on empirical models).

5.3 Strengthening collaboration and institutionalization for NTD use during crises

COVID-19 demonstrated the need for multi-stakeholder partnerships to implement and sustain NTD initiatives. NTD initiatives are resource-intensive, requiring significant funding, support, and expertise to get them off the ground. This fact is reflected in how NTD initiatives launched during the COVID-19 pandemic were frequently cross-sectoral partnerships. Organizations needed the additional capacity and resources to reduce costs and accelerate implementation.

However, these initiatives were frequently implemented haphazardly, contributing to problems later on. Organizations struggled to rapidly integrate data systems and infrastructure, create emergency legal agreements, develop personal data protections and design processes, all of which slowed implementation.

448 "Contracts for Data Collaboration," August 25, 2022. https://contractsfordatacollaboration.org/.

Many NTD partnerships were formed without consideration for how the data would be used, where it could be implemented within the policy decision-making cycle, and how it could be scaled beyond initial programs. For example, wastewater data provided valuable insights into the state of COVID-19 from a holistic perspective, but many governments did not have systems to integrate it within traditional clinical testing databases.

There is an opportunity for organizations to minimize these challenges in future crises and promote more well-organized, multi-stakeholder partnerships by advancing data stewardship. Strong data stewardship can accelerate the responsible and efficient use of data across organizations. Below we provide recommendations for institutionalizing data collaborations through data stewardship as the core function.

- ***Establish Data Stewards:*** Drawing on organizational incentives to produce knowledge and insights, improve brand equity, or acquire a license to operate, appoint Chief Data Stewards within organizations with data assets to create and grow a network of data stewards across private, public, and non-profit organizations.[449] The Chief Data Stewards can be the point of contact to streamline the creation of data collaboratives and host regular training on advancing the responsible use of NTD across organizations. The network of data stewards can initiate new collaborations in a systemized format, exchange best practices, help develop plans to address NTD driven mis-/dis-information and advance the value of NTD across their organization.

- ***Create an Institutional Data Governance Framework:*** Facilitate meetings with leadership across government functions, public engagements, and co-design workshops to develop common data standards, principles, and practices for efficient and ethical internal and external data collaborations. Unlock funding and resources to support the implementation of the framework–this could be a crisis data fund or an emerging group of trusted technology and research partners that can be leveraged during crises.

449 Zahuranec, Andrew J. "The '9Rs Framework': Establishing the Business Case for Data Collaboration." *Data Stewards Network* (blog), November 9, 2021. https://medium.com/data-stewards-network/the-9rs-framework-establishing-the-business-case-for-data-collaboration-26585455ccc0.

5.4 Preventing fragmentation, and improving readiness and coordination

Across the briefings, we observed various instances of fragmentation and misalignment among NTD actors. These differences could be large or minor. They could appear during the data lifecycle, when the actors involved in planning have little idea how others will use their insights, or as the pandemic evolves, when a failure to accommodate shifting resources, needs, and information leads to disruptions. Organizations either developed their initiatives in silos from other crisis management efforts or failed to consider how pandemic response efforts would need to evolve, a problem that affiliated health governance throughout the pandemic.

While the unprecedented nature of the COVID-19 pandemic clarifies some aspects of the challenges, it also indicates a broader level of unpreparedness and lack of coordination. Dashboards on case counts were launched without a sense of how others might use them or what already existed in the field. Organizations announced plans to analyze specific types of biometric data without considering what new insights this could provide or how this could supplement existing data sources, leading to few efforts moving beyond the pilot phase. Crowdsourced efforts to collect symptoms neglected to revise their app surveys as the pandemic evolved and new research emerged about new symptoms associated with COVID-19. In short, this study found many examples of NTD projects that were duplicative, redundant, or unusable by decision-makers. Other institutions have observed that these problems contributed to a breakdown in data-driven decision-making.[450]

Consequently, there is an acute need to improve preparedness and coordination across the data and pandemic lifecycle. The gaps identified in this section can be filled by developing infrastructure that can sustain and promote the cohesion of NTD initiatives over an extended timeframe as well as by developing shared research agendas across public and private organizations—leveraging big and thick data—to increase crisis readiness. Pathways to accomplish this goal can include:

---

450 MPs on the House of Commons Public Administration and Constitutional Affairs Select Committee. "Covid-19 and Data Driven Decision Making." House of Commons Committees, March 15, 2021. https://houseof-commons.shorthandstories.com/data-transparency-coronavirus-PACAC-report/.

- ***Coordinate with Third Parties and Launch a #Data4COVID19 Fellowship Program:*** Proactively develop relationships with technical and private technology experts that can be leveraged during crisis situations to address NTD needs rapidly. By initiating consultative conversations with industry experts and providers, organizations can develop a list of the resources and capabilities needed for certain types of crises and identify pathways for attaining those resources. They can use this understanding of needs to develop partnerships with experienced organizations that can, in turn, provide staff and other assistance on an as-needed basis. Organizations might also seek to develop draft agreements to continue the program beyond crisis situations to increase readiness.

- ***Create an International Data Innovation for Pandemics Council:*** Anticipate future pandemics by funding an International Council of data, public health, policy, foresight, and innovation specialists that aim to advance NTD initiatives for increasing pandemic readiness. The Council would fund and support NTD initiatives around the world, particularly in the Global South, that identify and track disease outbreaks (e.g. wastewater data programs) as well as initiatives that combine NTD with thick data to anticipate low probability, high impact events requiring further preparedness. The Council might provide training programs for advancing internal data innovation capacities, and organizations might assign a point of contact within crisis management data teams to participate and lead collaborations.

- ***Institute a National NTD Localization Process:*** Develop standardized processes for local policy-makers to contribute to national NTD initiatives and decisions. This work could include inviting local representatives to a series of design sprints. Organizations might also fund efforts to localize NTD initiatives or explicitly seek participation from local communities in larger NTD initiatives. These efforts could be guided by requirements that all types of national NTD initiatives include guidelines on how local and subnational governments could expand upon NTD.

Recommendation	Pathway	Actors	Timeline
1. Increasing evidence and awareness about the value proposition of NTD	Develop a DATA4CRISIS Alliance	Multilateral organizations, private data holders, and research institutions	Medium-term
	Develop a Global NTD Research Agenda	Research institutions, private data holders, and multilateral organizations	Medium-term
	Develop broadly accepted metrics of success	Private data holders, national and subnational policymakers, and research institutions	Short-term
	Develop an NTD Observatory	Multilateral organizations, private data holders, and research institutions	Short-term
	Develop NTD Transparency Index	Private data holders, and national policy makers	Short-term
2. Advancing trust, ethics, and equity within NTD initiatives	Include Public Advocates in Contractual Negotiations	Private data holders, national policy makers, and local policymakers	Medium-term
	Establish Common Policy Frameworks	Multilateral organization, national, subnational, and local policy makers, private data holders, privacy experts, and advocacy groups	Long-term
	Institute Legislation or Regulation to Address Specific Crisis Context	National policymakers, sub-national policymakers, local policymakers, privacy experts, advocacy groups, and community organizations	Short-term
	Publish the Contracts Impacting Data Reuse	National policymakers, sub-national policymakers, local policymakers, privacy experts, advocacy groups, and community organizations	Short-term
	Conduct Regular Engagements to Understand Needs	National, subnational, and local policy makers, private data holders, non-profits, and research institutions	Short-term
3. Strengthening collaboration and institutionalization NTD uses	Establish Data Stewards	National policy makers, private data holders, and multilateral organizations	Short-term
	Create an Institutional Data Governance Framework	National policymakers, data rights advocates, research institutions, private sector data holders	Medium-term

4. Preventing fragmentation, and improving readiness and coordination	Coordinate with Third Parties and Launch a #Data4COVID19 Fellowship Program	National policy makers (including legal departments), private data holders, research institutions, and multilateral organizations	Medium-term
	Create an International Data Innovation for Pandemics Council	Multilateral organizations, foundations, private data holders, national policy makers, research institutions, and non-profit	Long-term
	nstitute a National NTD Localization Process	National, subnational, and local policy makers, and community groups	Medium-term

Table 3: Pathways Forward

The above table provides the possible pathways and actors involved in making the recommendations the reality.

Appendix 1: Glossary of Term

Aggregation	"Aggregation is the combination of related categories, usually within a common branch of a hierarchy, to provide information at a broader level to that at which detailed observations are taken"[451] "Aggregate data therefore is data obtained by aggregation, as distinct from unit record data"[452]
Artificial Intelligence	"Artificial Intelligence (AI), a term coined by emeritus Stanford Professor John McCarthy in 1955, was defined by him as 'the science and engineering of making intelligent machines'. Much research has humans program machines to behave in a clever way, like playing chess, but, today, we emphasize machines that can learn, at least somewhat like human beings do"[453]
Credit and debit card transactions	Credit and debit card transactions generate data about where, when, and how people spend money as well as how much money they are spending. Statistical organizations have tried to use these transactions as a proxy for economic health amid the pandemic.[454]
Crowdsourced Data	"Crowdsourced data collection is a participatory method of building a dataset with the help of a large group of people."[455]
Datafication	Datafication refers to the "quantification of human life through digital information."[456]

451 OECD. "Aggregation Definition." Glossary of Statistical Terms, June 10, 2013.

452 OECD. "Aggregation Definition." Glossary of Statistical Terms, June 10, 2013. https://stats.oecd.org/glossary/detail.asp?ID=68.

453 Manning, Christopher. "Artificial Intelligence Definitions." Stanford University Human-Centered Artificial Intelligence, September 2020.

454 Dunn, Abe, Kyle Hood, and Alexander Driessen. "Measuring the Effects of the COVID-19 Pandemic on Consumer Spending Using Card Transaction Data." BEA Working Paper Series, WP2020-5, April 24, 2020, 22.

455 Hunt, Amelia, and Doug Specht. "Crowdsourced Mapping in Crisis Zones: Collaboration, Organization and Impact." Journal of International Humanitarian Action 4, no. 1 (December 2019): 1. https://doi.org/10.1186/s41018-018-0048-1.

456 Mejias, Ulises A., and Nick Couldry. "Datafication." Internet Policy Review 8, no. 4 (November 29, 2019). https://policyreview.info/concepts/datafication.

Data Governance	"Data Governance is a system of decision rights and accountabilities for information-related processes, executed according to agreed-upon models which describe who can take what actions with what information, and when, under what circumstances, using what methods."[457]
Data Source	A data source refers to "a specific data set, metadata set, database or metadata repository from where data or metadata are available."[458] See also Metadata: "It is pieces of information that have some meaning in relation to another piece of information. It can be created, managed, stored, and preserved like any other piece of data"[459]
Data warehouses	"A data warehouse is a repository for storing data which may have been gathered from a source or multiple sources, manually or automatically, via an integration layer that transforms data to meet the criteria of the warehouse. Data warehouse can be conceptualized as a one stop information center large volume of data which is designed under a common framework"[460]
Granularity	"Granularity concerns the ability to represent and operate on different levels of detail in data, information, and knowledge that are located at their appropriate level."[461]

457 The Data Governance Institute. "Defining Data Governance." Accessed October 4, 2022. https://data-governance.com/defining-data-governance/.

458 OECD. "Data Source Definition." Glossary of Statistical Terms, February 1, 2006. https://stats.oecd.org/glossary/detail.asp?ID=7045.

459 UNC University Libraries. "Metadata for Data Management: A Tutorial: Definition," July 26, 2022. https://guides.lib.unc.edu/metadata/definition.

460 Dhillon, Sarinder. "Data Warehouse - an Overview." ScienceDirect, 2017. https://www.sciencedirect.com/topics/immunology-and-microbiology/data-warehouse.

461 Keet, C. Maria. "Granularity." In Encyclopedia of Systems Biology, edited by Werner Dubitzky, Olaf Wolkenhauer, Kwang-Hyun Cho, and Hiroki Yokota, 850–53. New York, NY: Springer, 2013. https://doi.org/10.1007/978-1-4419-9863-7_65.https://guides.lib.unc.edu/metadata/definition.

Machine Learning	Machine learning refers to "a subfield of artificial intelligence, which is broadly defined as the capability of a machine to imitate intelligent human behavior."[462]
Non-Traditional Data (NTD)	"Data that is digitally captured (e.g. mobile phone records and financial data), mediated (e.g. social media and online data), or observed (e.g. satellite imagery)",[463] using new instrumentation mechanisms, often privately held.[464]
Non-Traditional Economic Data	NTD that describes the economic activity of individuals, groups, or organizations.
Non-Traditional Health Data	NTD that describes, or can be used as a proxy for, the physical health of individuals or populations, especially as it relates to COVID-19 diagnoses.
Non-Traditional Mobility Data	NTD that describes the location of a device relative to people and landmarks in the physical world or to other devices.
Non-Traditional Sentiment Data	NTD that describes how individuals and groups perceive and understand developments related to COVID-19.
Open Contracting	Open contracting refers to the practice of "publishing and using open, accessible and timely information on public contracting to engage citizens and businesses to fix problems and deliver results."[465]

462 Mackey, Tim, Vidya Purushothaman, Jiawei Li, Neal Shah, Matthew Nali, Cortni Bardier, Bryan Liang, Mingxiang Cai, and Raphael Cuomo. "Machine Learning to Detect Self-Reporting of Symptoms, Testing Access, and Recovery Associated With COVID-19 on Twitter: Retrospective Big Data Infoveillance Study." JMIR Public Health and Surveillance 6, no. 2 (June 8, 2020): e19509. https://doi.org/10.2196/19509.; Vermicelli, Silvia, Livio Cricelli, and Michele Grimaldi. "How Can Crowdsourcing Help Tackle the COVID- 19 Pandemic? An Explorative Overview of Innovative Collaborative Practices." R&D Management 51, no. 2 (March 2021): 183–94. https://doi.org/10.1111/radm.12443.

463 Albanna, Basma, Andreas Pawelke, Jeremy Boy, and Andreas Gluecker. "The Data-Powered Positive Deviance Handbook." The University of Manchester Global Development Institute, November 2021. https://www.gdi.manchester.ac.uk/research/publications/di/dd-wp92/.

464 Ibid

465 Open Contracting Partnership. "What Is Open Contracting." Accessed August 31, 2022. https://www.open-contracting.org/what-is-open-contracting/.

Platform-facilitated symptom data	Platform-facilitated symptoms data refers to the use of new survey instrumentation approaches such as internet searches, social media, privately run websites and other resources to gather information from the public about COVID-19 symptoms. Platform-facilitated symptom data was used within both short- and long-term projects and relied upon the self-identification of symptoms.[466]
Social media data	Social media data, gathered from social media platforms like Twitter, Facebook, and Instagram, can provide insights into how people are feeling about different aspects of the pandemic and the response to it.
Software Development Kit (SDK) data	"A Software Development Kit (SDK) is a piece of software integrated into a smartphone application and enables the app to collect location data by using the device's hardware, such as GPS, Bluetooth and Wi-Fi. It can either be collected directly or indirectly."[467]
Supply chain data	A supply chain is broadly defined as "the processes involved in producing and distributing a product or service, including the organizations, people, tools, and information necessary to get an entity into consumers' hands."[468] The supply chain data, then, are all the data referring to such processes,[469] that inform the production and distribution of goods by analyzing trends and providing organizations with necessary insights to create an optimal supply plan.[470]

466 Vermicelli, Silvia, Livio Cricelli, and Michele Grimaldi. "How Can Crowdsourcing Help Tackle the COVID-19 Pandemic? An Explorative Overview of Innovative Collaborative Practices." R&D Management 51, no. 2 (March 2021): 183–94. https://doi.org/10.1111/radm.12443.

467 The GovLab and Cuebiq. "What Is Mobility Data? Where Is It Used?," April 2021. https://files.thegovlab.org/COVID-19_Mobility_Data_One-Page_Brief_v3_041521.pdf.

468 Devane, Heather. "What Is the Data Supply Chain?" Immuta (blog), August 11, 2021. https://www.immuta.com/blog/what-is-the-data-supply-chain/.

469 Ibid.

470 Koshulko, Alex. "Overcoming Supply Chain Challenges With AI: What Large Manufacturers Need To Know." Forbes, April 12, 2022, sec. Innovation. https://www.forbes.com/sites/forbestechcouncil/2022/04/12/overcoming-supply-chain-challenges-with-ai-what-large-manufacturers-need-to-know/.

Telecommunications data	"Telecommunications companies (telecoms) are frequent providers of mobility data due to the wide availability of portable cellular devices and the wide coverage of their networks."[471]
Vehicle GPS data	Vehicle GPS data describes the location of an automobile, ship, airplane, or other vehicle as indicated by portable devices placed in vehicles or hardware built into them.
Wastewater surveillance	"The virus can then be detected in wastewater, enabling wastewater surveillance to capture presence of SARS-CoV-2 shed by people with and without symptoms. This allows wastewater surveillance to serve as an early warning that COVID-19 is spreading in a community."[472]
Wearables and Biometric Data	'"Biometric data means personal data resulting from specific technical processing relating to the physical, physiological or behavioral characteristics of a natural person, which allow or confirm the unique identification of that natural person, such as facial images or dactyloscopic data."[473]

471 The GovLab and Cuebiq. "What Is Mobility Data? Where Is It Used?," April 2021. https://files.thegovlab.org/COVID-19_Mobility_Data_One-Page_Brief_v3_041521.pdf.

472 Centers for Disease Control and Prevention. "National Wastewater Surveillance System (NWSS)," May 16, 2022. https://www.cdc.gov/healthywater/surveillance/wastewater-surveillance/wastewater-surveillance.html.

473 Information Commissioner's Office. 2022. "What Is Special Category Data?". Ico.Org.Uk. https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/special-category-data/what-is-special-category-data/.

Appendix 2: Scope

The COVID-19 pandemic and the responses to it have affected almost every individual on the planet and, inherently, no single report can cover every aspect of the crisis. The use of "data" alone is too broad to be adequately discussed given the extensive history of data analysis in public health and epidemiology.[474]

Instead, this report focuses on a very specific issue: the use of NTD for pandemic response. The introduction to the report includes our definition of NTD. A more detailed explanation of NTD is outlined below:

- Non-epidemiological data sources that are not typically used in government public health systems. Frequently, these are data from social media and online searches, advertising, telecommunications, biometrics, and satellites. This data is frequently the result of collaborations with private sector entities, civil society organizations, and individuals themselves; or

- Extremely large datasets that can be analyzed by and used by automated systems such as artificial intelligence. This includes the amalgamation of electronic personal health records within large scale datasets.

Exclusion Criteria

The research team also restricted our research by imposing the following minimum requirements on the sources we gathered. By focusing on NTD, the intention of this work is to highlight elements of data-driven crisis response that could aid future crisis management.

- The action described must detail the production, analysis, or use of a NTD source during the COVID-19 pandemic (March 2020 to July 2022);

- There must be sufficient independent sources/stories or potential interviewees for

474 Soucie, J. Michael. "Public Health Surveillance and Data Collection: General Principles and Impact on Hemophilia Care." Hematology (Amsterdam, Netherlands) 17, no. 0 1 (April 2012): S144–46. https://doi.org/10.1179/102453312X13336169156537.

The GovLab team to evaluate each briefing using the assessment framework (inputs and planning and governance outputs);

- The online resources available must be available in or translatable to English.

- The NTD included in the scope of this research contributed to government decision-making to promote the public good and mitigate the consequences of the COVID-19 pandemic.

Inclusion Criteria

Given that various, disparate data sources meet the minimum requirements, the research team (with the input of the Advisory Group) developed an additional set of requirements to ensure the briefings focused on only those data sources with the most potential for impact during future crisis management efforts. The model includes three categories of criteria—universality, purpose-driven, and operationalization—all of which aim to assess whether the action could be effective in crisis decision making processes beyond COVID-19.

1. **Universality:** The action can be applied in alternative contexts beyond COVID-19;
    a. The types of data and data holders are present in multiple countries;
    b. Multiple countries have the institutions or computing capacity in place to replicate it;
    c. There are comparable examples in multiple countries (within or beyond COVID-19).

2. **Purpose-driven:** The action was designed to impact one or more aspects of the COVID-19 response;
    a. The action aimed to address specific COVID-19 related challenges.
    b. The action aimed to impact the COVID-19 response by providing information to decision makers at a minimum.

3. **Operationalization:** The action was implemented as a part of the crisis management efforts;
    a. The action was actualized beyond pilot programs and proof of concept (academic purposes);
    b. The action was implemented within one or more of the pandemic lifecycle phases: Readiness, Response, Recovery, and Reform.

The data sources were selected using this criteria while also considering diversity of geographic location, demographics, intention of using data, and degree of impact of potential examples. Our evaluation of potential briefings using the selection criteria is summarized in the table below.

Briefings and data sources	Universality			Purpose-driven		Operationalization	
	Data holders are present in multiple countries	Institutions or computing capacity to replicate it in other countries	Comparable examples in multiple countries	Aimed to address specific COVID-19 problems	Provided information to decision makers at a minimum	Actualized beyond pilots programs and proof of concept	Implemented during one or more pandemic lifecycle phases
Briefing 1: Non-Traditional Health Data							
Digital patient data	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Platform facilitated symptoms data	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wastewater data	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wearables and biometric data	Yes	Yes	Yes	Yes	Yes	No	No
Briefing 2: Non-Traditional Mobility Data							
Telecommuni-cations Data	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Software Development Kits	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Vehicle GPS	Yes	Yes	Yes	Yes	Yes	Yes	No
Bluetooth based proximity data	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geospatial data	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Briefing 3: Non-Traditional Economic Data							
Credit and debit card transactions	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Open contracting	Yes	Yes	Yes	Yes	Yes	No	Yes
Supply chain data	Yes		Yes	Yes	Yes	Yes	Yes
Briefing 4: Non-Traditional Sentiment Data							
Social media data	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Crowdsourced data	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Limitations

In terms of its scope, the authors are aware that this report is limited by the extent to which decision makers have publicized their data use or had their efforts reported on by third parties (e.g. press). The authors are also aware that its research into uses of NTD relied on searches on the open web in English and that they may have missed sources not generally available or available in other languages. Understanding that the pandemic is complex and rapidly evolving, it may have also missed those efforts outside the time scope, March 2020 to June 2022. Many of the examples are drawn from The GovLab's #Data4COVID19 Data Collaborative Repository, which developed from a crowdsourced listing of notable data-driven initiatives.[475]

In terms of insights generated, this work aims to address how data impacted the COVID-19 response by analyzing correlations between NTD use for decision-making and the main outcomes of the pandemic. This work does not intend to show causalities between data use and the pandemic outcomes. The manifestation of the virus varied regionally because of several variables beyond data. Data was one of many factors that contributed to the COVID-19 response. In many cases, the analysis here cannot identify specific impact nor can it assess impact on specific subsets of the population. We encourage further research and exploration to identify these effects.

475 The GovLab. "A Living Repository for Data Collaboratives Seeking to Address the Spread of COVID-19." #Data4COVID19. Accessed September 14, 2022. https://list.data4covid19.org/.

Appendix 3: Assessment Framework

The GovLab developed an assessment framework to evaluate how different NTD initiatives were organized and implemented. The goal is to compare how different NTD initiatives performed across each of the assessment criteria. Below we provide the criteria for evaluating NTD initiatives as well as a toolkit–leveraging spider graphs–to visualize how NTD initiatives performed.

Inputs and planning: All the work and resources used to organize and launch the data source**.**

Assessment criteria	Description
Focus	The extent to which there was a well defined problem that could be resolved through the use of the data source.
Expertise	The degree to which organizations or groups of organizations used their internal resources or external partnerships to bring relevant personnel who could contribute to the effort.
Data	The degree to which organizations or groups of organizations are able to collect new data or gain access to useful and relevant datasets for their work through collaborative methods.
Affordability	The expenses—specifically related to financial and organizational resources—associated with using the data source.
Speed	The extent to which an initiative launched by an organization or group of organizations can produce actionable insights in an immediate and responsive fashion.
Innovation	The extent to which innovative analytical and other tools were used.
Political Context	The extent to which alignment with existing political priorities and regulations impacted the ability to use the data source.

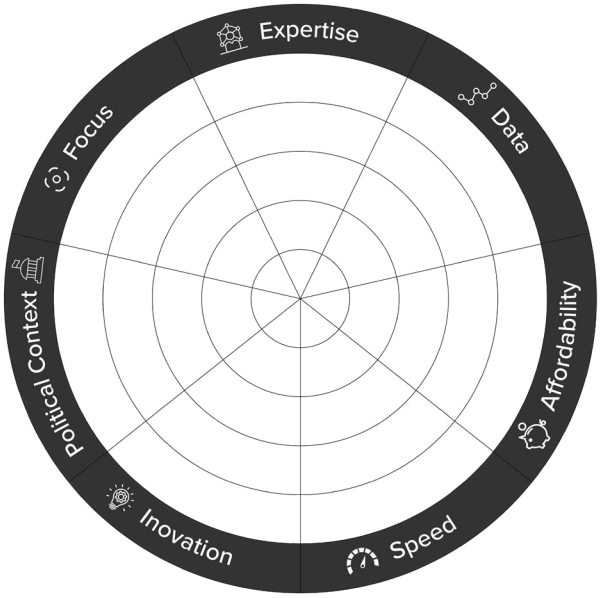
Governance and outputs: The degree to which the use of the data source can accomplish its goals in a responsible, legitimate, and accountable fashion.

Assessment criteria	Description
Actionability	The extent to which the work undertaken delivers results that can be meaningfully acted upon in a way that yields benefits for specific target populations.
Reusability	The degree to which the initiative could be scaled or reused by others in different contexts.
Legitimacy	The action is seen as legitimate and has the social license to operate by the groups it intends to benefit and abides by the laws, norms, and standards of the community in which it is based.
Accountability	The extent to which the efforts have been managed in a transparent, participatory, and representative fashion and ensured actors involved can be held accountable in the event of negligence or wrong-doing.
Duration	The length of the data lifecycle (Did the project continue throughout multiple pandemic phases? Or was it discontinued after the Readiness phase?).
Access to Information	The extent to which processes and outputs of using this data source where made publicly available before, during, and following the intervention.
Human Rights and Ethics	The degree to which the use of the data source proactively promotes the rights (e.g. privacy) of data subjects and beneficiaries and seeks to mitigate potential harm that might result from the work.
Equity	The inclusion of and impact on vulnerable groups.

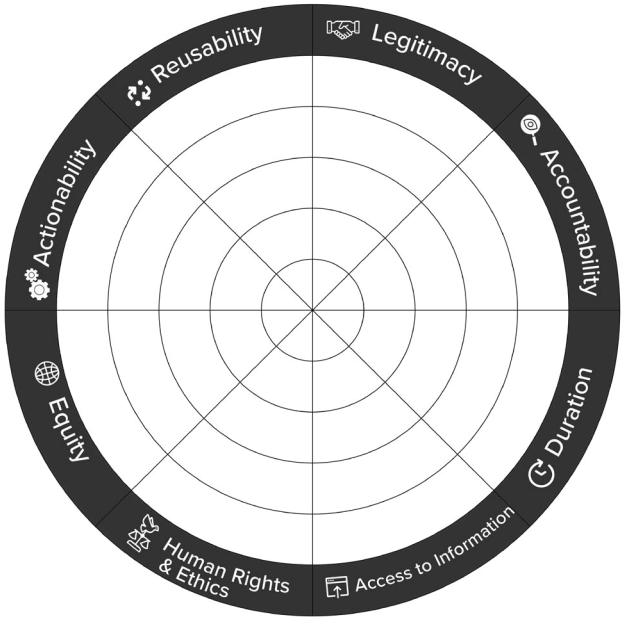
Toolkit

The GovLab has developed the following templates that can be used to visualize how different NTD initiatives performed across each criteria of the assessment framework. These templates aim to help decision makers demonstrate which facets of the initiative were fully actualized while other were not.

Inputs and Planning



Governance and Outputs



Appendix 4: Assessment of NTD Initiatives

Below we provide the insights developed across the assessment framework criteria used to develop the lessons learned in the report.

Inputs and Planning

- **Focus:** NTD was most effectively applied when it sought to fill a gap in existing traditional methods and not replace them outright.
- **Expertise:** Public sector institutions frequently relied on partners in the private sector to provide the expertise needed to analyze NTD and generate meaningful insights.
- **Data:** Large companies, start-ups, and research institutions tended to be the largest data holders of NTD, though many organizations did not provide meaningful access to their data.
- **Affordability:** NTD initiatives are resource intensive. Multi-stakeholder partnerships were frequently used to increase the capacity.
- **Speed:** Initiatives based on pre-existing relationships and agreements between institutions tended to be faster to start work and produce insights.
- **Innovation:** While innovative analytical tools provided valuable insights, they often did not align with traditional data systems creating technical challenges in using the data in the pandemic response
- **Political Context:** Initiatives that aligned more closely with existing political priorities were able to receive government funding and support.

Governance and Outputs

- **Actionability:** Most initiatives targeted some aspect of COVID-19's effects but many did not identify how the insights could meaningfully resolve these challenges.

- **Reusability:** Many NTD sources were used for the first time during COVID-19 and required new infrastructure to get them off the ground. This was particularly challenging for initiatives at the local level. Access to infrastructure helped make initiatives more reusable.

- **Legitimacy:** Public and private sector organizations frequently implemented NTD initiatives without the necessary social license, the degree to which an initiative aligns with expectations of the general public and other stakeholders.

- **Duration:** The majority of NTD initiatives were undertaken in the response phase of the pandemic. Limited research was available about noteworthy initiatives from the readiness and reform stages.

- **Accountability:** Few projects were built with mechanisms that enable clear and direct public oversight. In many cases, there is little public information available on how initiatives were governed or how they used data.

- **Access to Information:** In many cases, processes and outputs of using the data were not made publicly available until the initiative concluded or at all due to limited bandwidth and capacity constraints.

- **Equity:** NTD initiatives are resource intensive and were primarily implemented by stakeholders in the Global North who had the resources to do so.

- **Human Rights and Ethics:** The speed of implementation was often prioritized over evaluating the human rights and ethical implications.

Appendix 5: Selected Readings

Below we provide a curated list of recommended works to supplement this report. The focus of this list is advancing NTD use during crisis situations across the four categories of recommendations.[476] These readings were referred to throughout the research process and were discovered through various methods including recommendations by interviewees, peer reviewers, and our Advisory Group. This list is a part of our ongoing effort to build a knowledge base about NTD. We hope it provides clarity on some of the topics discussed in this report and inspires future responsible and effective NTD initiatives. Additional readings about several data related topics can be found at The Living Library Selected Readings.[477]

Abidoye, Babatunde, Joanna Felix, Serge Kapto, and Laurel Patterson. "Leaving No One Behind: Impact of COVID-19 on the Sustainable Development Goals (SDGs)." New York, NY and Denver CO: United Nations Development Programme and Frederick S. Pardee Center for International Futures, 2021. https://www.undp.org/publications/leaving-no-one-behind-impact-covid-19-sustainable-development-goals-sdgs.

Beduschi, Ana. "Taking Stock of COVID-19 Health Status Certificates: Legal Implications for Data Privacy and Human Rights." *Big Data & Society* 9, no. 1 (January 1, 2022): 20539517211069300. https://doi.org/10.1177/20539517211069300.

Benjamins, Richard, Jeanine Vos, and Stefaan Verhulst. "Mobile Big Data in the Fight against COVID-19." *Data & Policy* 4 (ed 2022): e9. https://doi.org/10.1017/dap.2021.39..

Buckee, Caroline, Satchit Balsari, and Andrew Schroeder. "Making Data for Good Better." Edited by Leo Anthony Celi. *PLOS Digital Health* 1, no. 1 (January 18, 2022): e0000010. https://doi.org/10.1371/journal.pdig.0000010.

Cajner, Tomaz, Laura J. Feiveson, Christopher J. Kurz, and Stacey Tevlin. "Lessons Learned from the Use of Nontraditional Data during COVID-19." *Brookings* (blog), April 27, 2022. https://www.brookings.edu/essay/lessons-learned-from-the-use-of-nontraditional-data-during-covid-19/.

476 See 6. Conclusion and Recommendations

477 The GovLab. "The Living Library Selected Readings." The Living Library, 2022. https://thelivinglib.org/selected-readings/.

Chakravorti, Bhaskar. "Why AI Failed to Live Up to Its Potential During the Pandemic." Harvard Business Review, March 17, 2022. https://hbr.org/2022/03/why-ai-failed-to-live-up-to-its-potential-during-the-pandemic.

Colovic, Ana, Annalisa Caloffi, and Federica Rossi. "Crowdsourcing and COVID-19: How Public Administrations Mobilize Crowds to Find Solutions to Problems Posed by the Pandemic." *Public Administration Review* 82, no. 4 (2022): 756–63. https://doi.org/10.1111/puar.13489.

Diamond, Megan B., Aparna Keshaviah, Ana I. Bento, Otakuye Conroy-Ben, Erin M. Driver, Katherine B. Ensor, Rolf U. Halden, et al. "Wastewater Surveillance of Pathogens Can Inform Public Health Responses." *Nature Medicine*, September 8, 2022. https://doi.org/10.1038/s41591-022-01940-x.

Empowering Society by Reusing Privately-Held Data for Official Statistics: A European Approach : Final Report Prepared by the High Level Expert Group on Facilitating the Use of New Data Sources for Official Statistics, 2022 Edition. Luxembourg: European Union, 2022. https://data.europa.eu/doi/10.2785/948477.

Felipe González-Zapata, Jacob Arturo Rivera Pérez, Lucia Chauvet, Cecilia Emilsson, Andrew J. Zahuranec, and Andrew Young. "Open Data in Action: Initiatives during the Initial Stage of the COVID-19 Pandemic." OECD and The GovLab, March 2021. https://www.oecd.org/gov/digital-government/use-of-open-government-data-to-address-covid19-outbreak.htm.

Halpern, Sue. "Can We Track COVID-19 and Protect Privacy at the Same Time?" *The New Yorker*, April 27, 2020. https://www.newyorker.com/tech/annals-of-technology/can-we-track-covid-19-and-protect-privacy-at-the-same-time.

LaFraniere, Sharon. "'Very Harmful' Lack of Data Blunts U.S. Response to Outbreaks." New York Times, September 20, 2022. https://www.nytimes.com/2022/09/20/us/politics/covid-data-outbreaks.html?partner=slack&smid=sl-share.

Lowe, Christy. "The Digitalisation of Social Protection before and since the Onset of Covid-19: Opportunities, Challenges and Lessons." ODI: Think change, June 23, 2022. https://odi.org/en/publications/the-digitalisation-of-social-protection-before-and-since-the-onset-of-covid-19-opportunities-challenges-and-lessons/.

Mehta, Nishita, and Sharvari Shukla. "Pandemic Analytics: How Countries Are Leveraging Big Data Analytics and Artificial Intelligence to Fight COVID-19?" *SN Computer Science* 3, no. 1 (November 9, 2021): 54. https://doi.org/10.1007/s42979-021-00923-y.

Mulgan, Geoff. "COVID's Lesson for Governments? Don't Cherry-Pick Advice, Synthesize It." *Nature* 602, no. 7895 (February 1, 2022): 9–9. https://doi.org/10.1038/d41586-022-00212-5.

Pandit, Jay A., Jennifer M. Radin, Giorgio Quer, and Eric J. Topol. "Smartphone Apps in the COVID-19 Pandemic." *Nature Biotechnology* 40, no. 7 (July 2022): 1013–22. https://doi.org/10.1038/s41587-022-01350-x.

Patino, Marie. "The Rise of the Pandemic Dashboard." *Bloomberg*, September 25, 2021, sec. CityLab. https://www.bloomberg.com/news/features/2021-09-25/why-every-government-needs-a-covid-dashboard.

Ramachandran, Selva, and Deepali Khanna. "Leveraging Non-Traditional Data For The Covid-19 Socioeconomic Recovery Strategy." Forbes, February 1, 2022. https://www.forbes.com/sites/deepalikhanna/2022/02/01/leveraging-non-traditional-data-for-the-covid-19-socioeconomic-recovery-strategy/.

Sachs, Jeffrey D., Salim S. Abdool Karim, Lara Aknin, Joseph Allen, Kirsten Brosbøl, Francesca Colombo, Gabriela Cuevas Barron, et al. "The Lancet Commission on Lessons for the Future from the COVID-19 Pandemic." *The Lancet* 0, no. 0 (September 14, 2022). https://doi.org/10.1016/S0140-6736(22)01585-9.

Scotti, Francesco, Francesco Pierri, Giovanni Bonaccorsi, and Andrea Flori. "Responsiveness of Open Innovation to COVID-19 Pandemic: The Case of Data for Good." *PLOS ONE* 17, no. 4 (April 26, 2022): e0267100. https://doi.org/10.1371/journal.pone.0267100.

Trautman, Lawrence J., and Mohammed T. Hussein. "The Internet of Things (IoT) in a Post-Pandemic World." SSRN Scholarly Paper. Rochester, NY, July 16, 2022. https://doi.org/10.2139/ssrn.4149477.

Tušl, Martin, Anja Thelen, Kailing Marcus, Alexandra Peters, Evgeniya Shalaeva, Benjamin Scheckel, Martin Sykora, et al. "Opportunities and Challenges of Using Social Media Big Data to Assess Mental Health Consequences of the COVID-19 Crisis and Future Major Events." *Discover Mental Health* 2, no. 1 (June 27, 2022): 14. https://doi.org/10.1007/s44192-022-00017-y.

Young, Andrew, Andrew Zahuranec, Stefaan G. Verhulst, and Kateryna Gazaryan. "The Third Wave of Open Data Toolkit: Operational Guidance on Capturing Institutional and Societal Value of Data Re-Use." Brooklyn, New York: *Open Data Policy Lab*, March 2021. http://files.thegovlab.org/The-Third-Wave-of-Open-Data-Toolkit.pdf.

https://review.data4covid19.org