CRISISREADY

Annual Report [2020]

Covid-19 Mobility Data Network

Center for Communicable Disease Dynamics | Harvard T.H. Chan School of Public Health | Harvard Medical School | Direct Relief

February 2021

In March 2020, as the COVID-19 pandemic grew out of control, we reached out to colleagues from around the world to serve as critical links between technology companies seeking to aid pandemic response by providing vast amounts of behavioral data from mobile phones, and policy makers who were overwhelmed with their responsibilities to contain the epidemic and did not have the capacity to synthesize and absorb the deluge of new data streams available to them. Given our collective years of experience in the application of mobility data to epidemic modeling and emergency response, we recognized that our ability to help researchers access and analyze data rigorously to inform and support policy makers would be a valuable contribution to the global response.

The Covid-19 Mobility Data Network that came together, comprising 100+ researchers from around the world, met 21 times in the course of 2020, to examine and exchange rapidly evolving epidemiological approaches to applying human mobility data generated from cell phones to pandemic modelling, response and mitigation. Network members provided daily analyses to city and state emergency response officials around the world, often connected through the network, and supported each other collaboratively through the sharing of code, methods, advice, and connections in industry and government.

In late 2020, we were called upon to apply these tools to the evacuation shelter planning in the Gulf states during the annual hurricane season in the US. Network members expanded the approach to plan health system resilience in response to the wildfires in California.

We have observed that the application of novel data streams in public health emergencies requires strengthening access and quality of data *prior* to the emergencies, capacity building to address the methodological challenges associated with combining and integrating these novel information sources with traditional data, and ongoing engagement with public agencies to understand and embed these insights into their response plans. Data Readiness, Methods Readiness and Translational Readiness therefore serve as the organizing principles for CrisisReady, a response-research platform that we now transition to, building on the CMDN Network. This platform, based at Harvard and Direct Relief, will work with academics, technology companies, public agencies and local communities to advance data-driven decision making for crisis mitigation, response and recovery.

We are grateful for the support provided by the vast network of scientists from around the world, and to collaborators at Camber Systems, Facebook Data for Good, Mapbox, Cubeiq, and colleagues at the Harvard Data Science Initiative, Institute for Quantitative Social Sciences, the Mittal Institute, and the Center for Communicable Disease Dynamics at Harvard. Nishant Kishore, a doctoral student at HSPH, deserves special mention for his enormous contribution to the Network's efforts. We thank Ankur Goel, Alissa Scharf, and Clay Heaton for their timely and generous help at the peak of the pandemic.

We thank our sponsors - the Harvard Data Science Initiative, Schmidt Futures, Facebook Data for Good, and Google.org - for support that will help us to sustain our vision for CrisisReady.

Sincerely,

Caroline Buckee

Harvard TH Chan

Satchit Balsari

Harvard Medical School

Andrew Schroeder

Direct Relief

Table of Contents

Introduction **5**

COVID19mobility.org **7**

CMDN Outputs **9**

Impact Evaluation **15**

Key Findings 15

Automated Situation Report 15

Translating Data into Action 17

A Path Forward 20

2021: Crisis Readiness **21**

Looking Ahead: 2021-2022 **23**

Appendix 1 - CMDN Network **26**

Appendix 2 - Sample Blogs from CMDN Meeting Series **30**

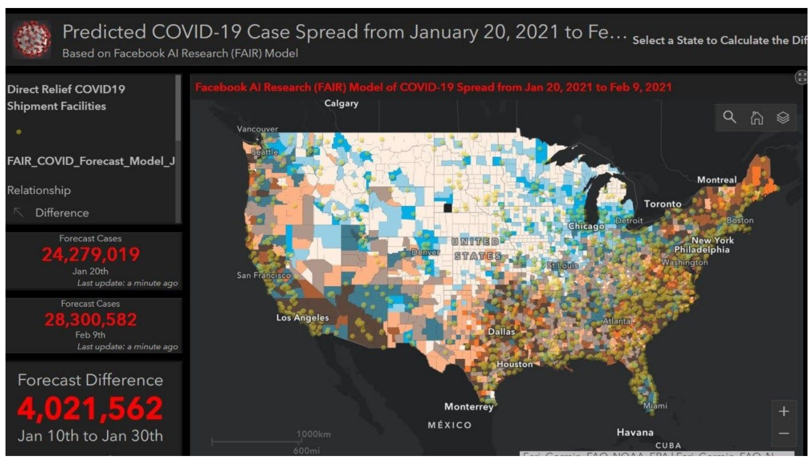
Appendix 3 - Panel Themes and Speakers **38**

Appendix 4 - Key Academic Outputs **41**

Appendix 5 - Our Team Now **44**

Introduction

In the early phases of the COVID-19 pandemic when much was not known about the SARS-CoV-2 virus, and neither drug therapies nor vaccines were on the horizon, non-pharmaceutical interventions like masking, handwashing, physical distancing and lockdowns were the only meaningful mitigation tools available. Societies around the world embraced each of these interventions with varying levels of acceptance. South Korea, China, and India, for example, were able to enforce drastic lockdowns. Smaller jurisdictions like the Cayman Islands and Singapore were able to track and monitor the movement of individuals, restricting privacy in the interest of public good. Populations in East and Southeast Asia have been comfortable using masks, when needed, since the SARS epidemic, while in the United States the mere act of using masks was intensely politicised.



Against this backdrop, technology companies made available vast amounts of anonymized aggregated human mobility data and analytics. These data were provided freely to researchers, policy makers and private corporations to help manage the pandemic. Foremost amongst the various applications of these data, were dashboards that scored how various regions responded to lockdowns or physical distancing orders.

TWO IMPORTANT SETS OF QUESTIONS NEEDED TO BE RESOLVED:

1] What do human mobility data not tell us? What is their representativeness? Who do they not count and who is left behind? And how do these data compare across data streams from say, Facebook, Google, Mapbox, Cubeiq, and Unacast, for example?

2] How do these data correlate with the actual epidemic spread? It was quite clear to infectious disease epidemiologists, from the beginning, that there would not be a linear relationship between human mobility (as captured by these data) and the pandemic - context was key. It was important to establish baseline mobility; examine the context within which interactions or movement occurred (private vehicles versus crowded public transport, density, local prevalence of the disease, masking practices, and so on); and have access to reliable clinical data, as both the scale and accuracy of testing gradually improved.

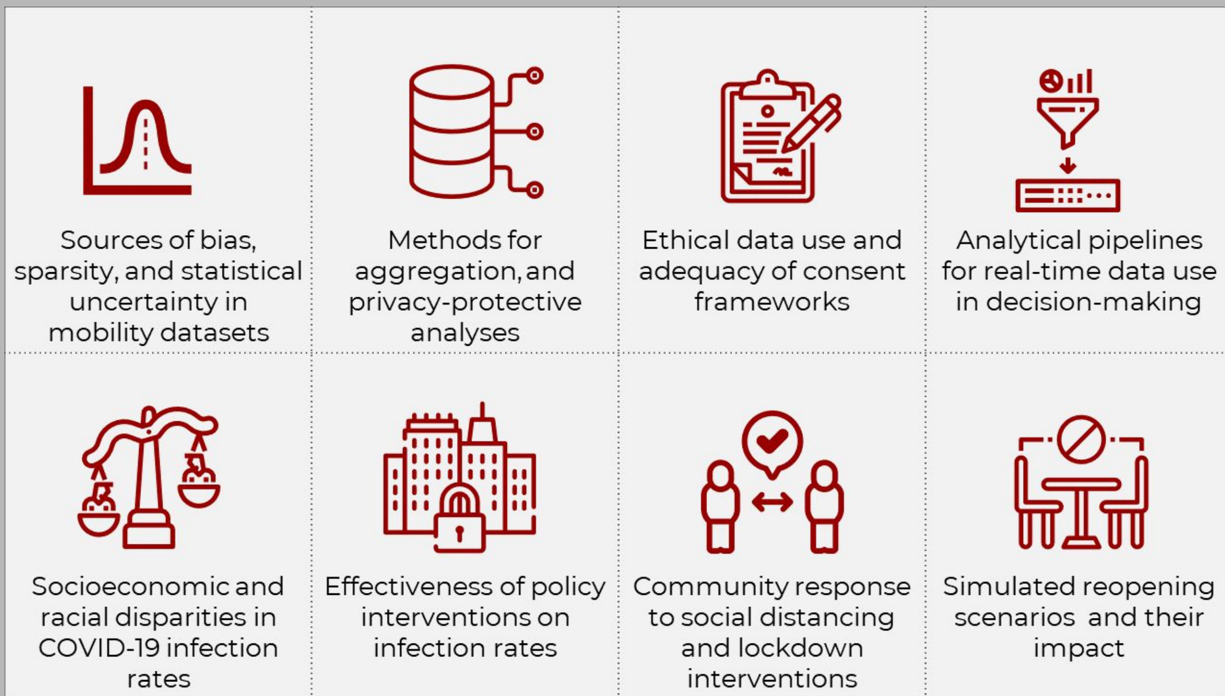
The **COVID-19 Mobility Data Network (CMDN)** was established in March 2020 to answer these questions, and facilitate the responsible and meaningful use of these widely available human mobility data by policy makers. We share here, a summary of our activities in 2020, and plans for 2021 and beyond.

COVID19mobility.org

CMDN's intent was to *"...establish routine analytic pipelines between tech companies and policy makers, providing meaningful policy-relevant information supported with scientific evidence and methodological rigor. These data will prove useful to communities, businesses and governments as restrictions are relaxed and policies to resume lives and livelihoods are planned, as national lockdowns and travel restrictions have resulted in highly unusual migration patterns."*

Starting as a group of five researchers based at the Harvard TH Chan School of Public Health and Direct Relief, CMDN expanded to 150+ researchers (Appendix 1) and policy makers around the world, serving as trusted intermediaries between technology companies and city and state governments across globally. CMDN provided analysis to or supported translational work for members of city governments, governors' offices, emergency operational centers (EOCs), public health offices, and multinational organizations.

CMDN convened 21 weekly workshops from April to December where experts examined evolving methodologies to apply human mobility data to pandemic response planning. Detailed blogs from our workshops are included in Appendix 2, and can also be accessed online at crisisready.io/news. In addition, CMDN also hosted four public panels, each attended by scientists from around the world. Panel themes and speakers are listed in Appendix 3.



A sample of themes discussed during the network-wide CMDN meetings

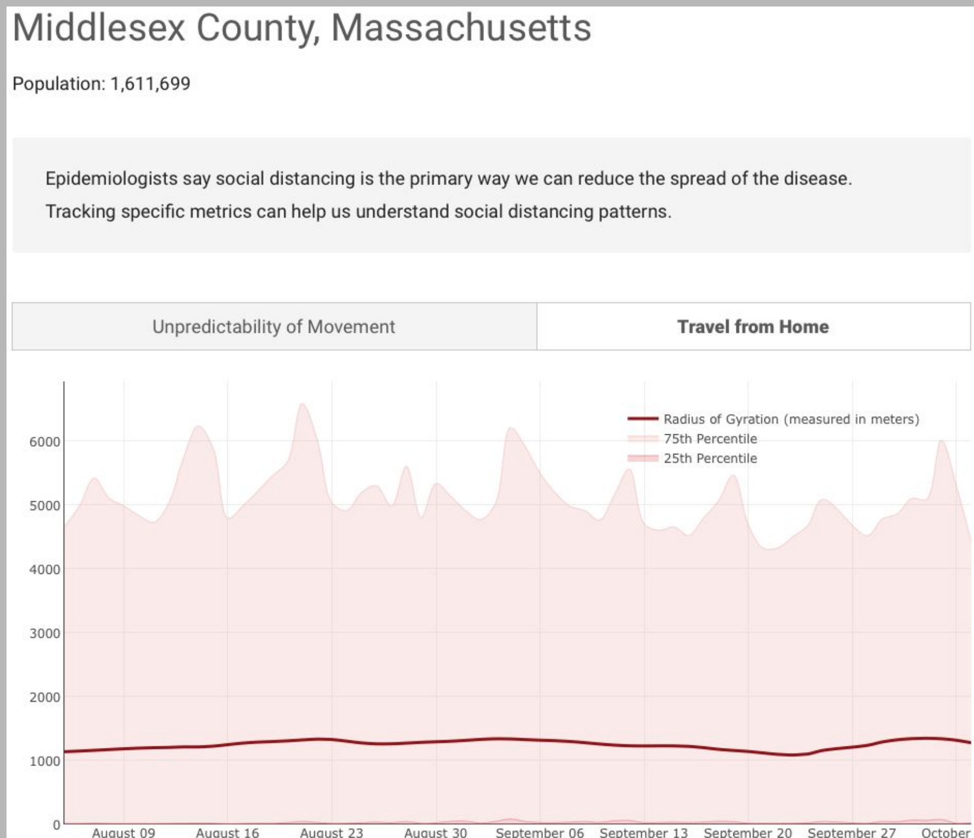
CMDN Outputs

The CMDN website hosted technical resources made freely available to Network Researchers, including the following:

- **Camber Systems Social Distancing Reporter Visualization Tool**
- **Facebook Data for Good Mobility Dashboard**
- **Code Repositories**
- **Mobility Data Metrics**
- **Technical Guidance Document / Data Source Comparisons**
- **Daily Debriefs**
- **Scientific Publications**

Camber Systems Social Distancing Reporter Visualization Tool

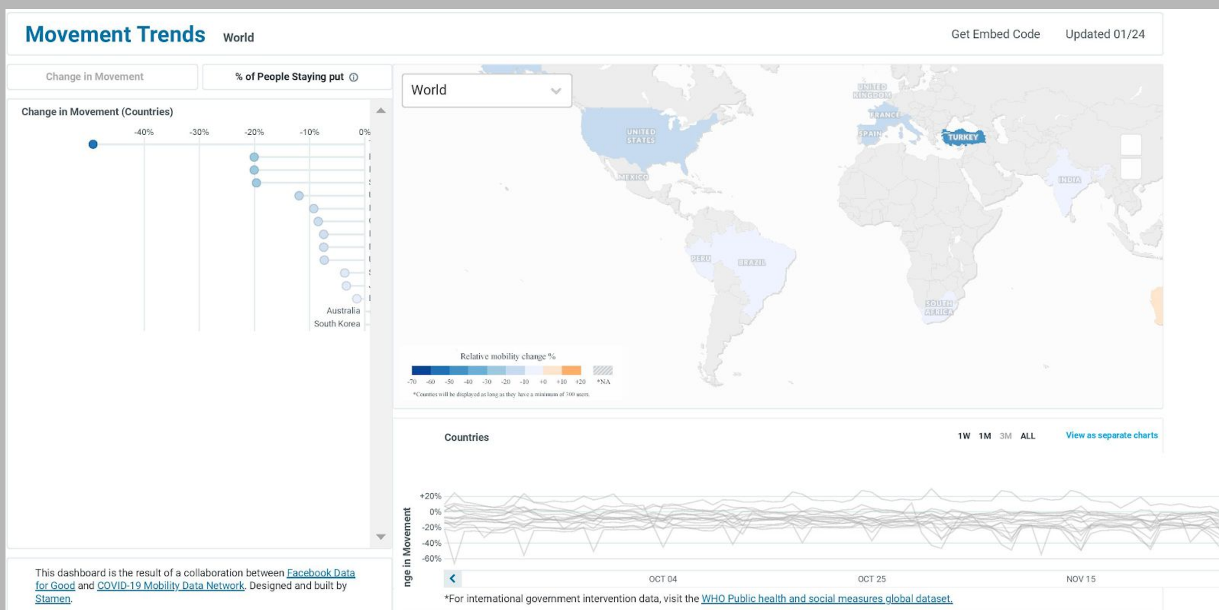
This website provides a simple interface through which individuals can easily monitor two mobility metrics that Camber Systems has evaluated at the county level with the help of researchers. The two metrics provided by Camber Systems openly are entropy and radius of gyration. In this situation, entropy defines how "predictable" the GPS traces are from a given device, and the radius of gyration provides an idea of how much these devices move on a given day. When aggregated to the county level, these metrics provide an idea of change in mobility over time. These metrics are derived from an aggregation of ad-tech data in the form of device-generated GPS points. The website also provides a single location for interested parties to better understand the mobility metrics and request access to these data directly. This website was built to be minimalist with a focus on searching for metrics of a single county rather than direct comparisons.



Camber Systems Social Distancing Reporter

Facebook Data for Good Mobility Dashboard

This dashboard provides a detailed interface to explore data generated by the Facebook Data for Good team. The final metrics and user interface were designed in conjunction with researchers from the CMDN. The website uses data generated by Facebook users around the world and provides a percent change in movement and a percent of Facebook users staying home metric. Both metrics depend on a categorization of a “home tile”, or a 600 by 600-meter tile in which a given individual spends most of their time at night. Percent change in movement captures how many of these 600m squared tiles users in a given region visit compared to a baseline time period conditional on the day of the week. Percent of people staying put evaluates the proportion of individuals in a given region who do not register any activity outside of their classified "home tile." This platform is designed to allow for ease of comparison between different regions and spatial scales, allowing a user to evaluate a region of interest while referring to similar or surrounding areas.



Facebook Data for Good Mobility Dashboard

Code Repositories

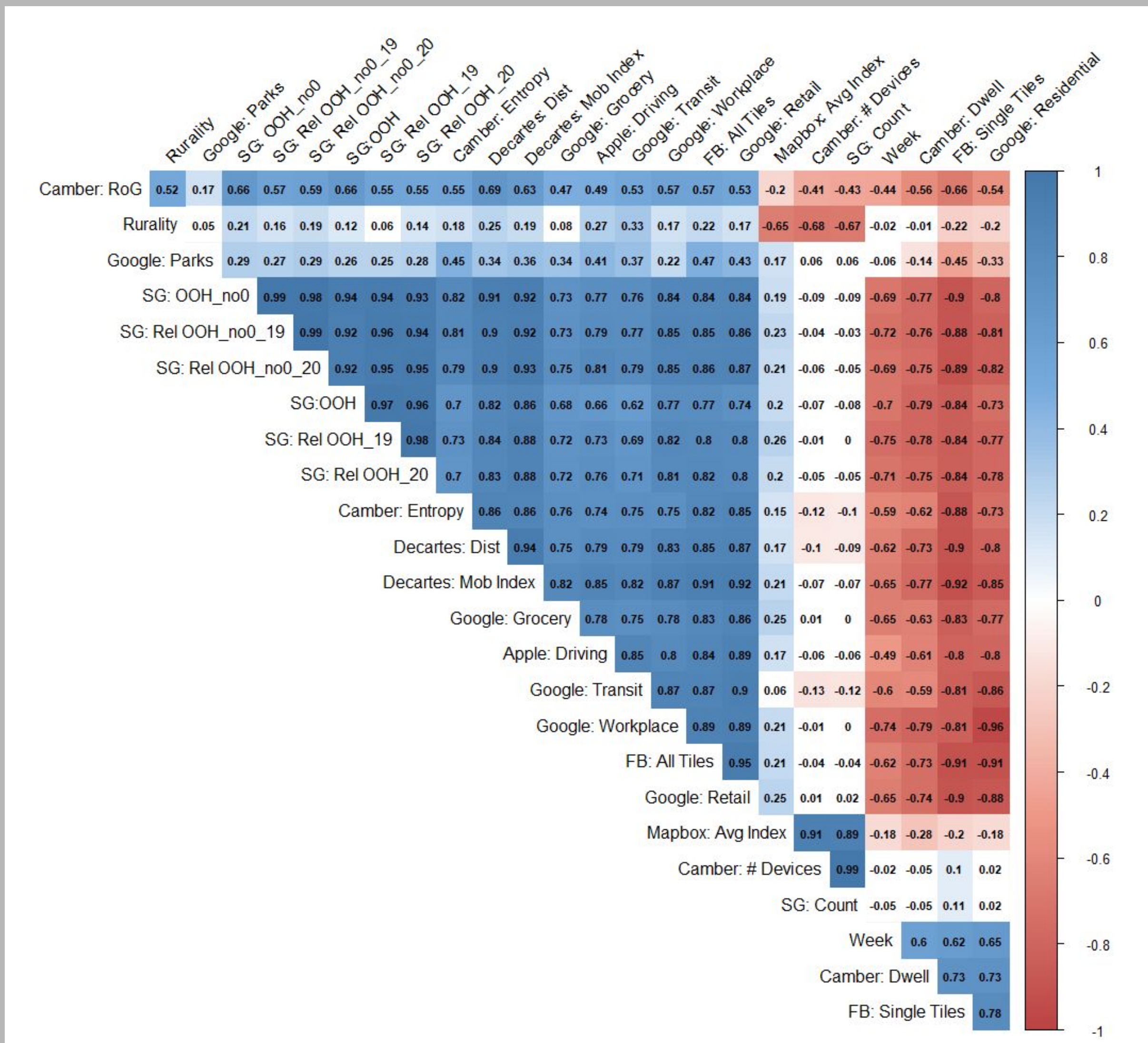
Early in the pandemic many researchers were conducting ad-hoc analyses using available mobility data from various providers. We consolidated many salient analyses under one GitHub organization and provided a single source for using these data sources. Network members shared code repositories and analytic tools and provided an online community to regularly develop new methods and analytic pipelines and policymakers expressed new needs in response to the pandemic.

Mobility Data Metrics

Despite several years of research in this area, standard frameworks for aggregating and making use of different data streams from mobile phones are scarce and difficult to generalise across data providers. Here, we examine aggregation principles and procedures for different mobile phone data streams and describe a common syntax for how aggregated data is used in research and policy. We argue that the principles of privacy and data protection are vital in assessing more technical aspects of aggregation and should be an important central feature to guide partnerships with governments that make use of research products. We provide a systematic review summarizing a description of commonly used metrics to measure mobility, used by policy makers, media, and communities.

Technical Guidance Document / Data Source Comparisons

Through the course of the pandemic a variety of providers have generated mobility-related data products to assist policy makers. We maintain a document that outlines considerations for public health use and privacy aggregation methods and sources of uncertainty methods for the analysis of aggregated data. Given the heterogeneity in data sources that are used to create these products, it is important that end-users are able to compare between metrics for different regions and better understand why they may agree or disagree. We are in the process of publishing a living document that reviews these data products and evaluates how well they correlate or don't so that end-users are better able to incorporate mobility metrics into their analyses and decisions. As more data become available, Network researchers expect to assist the development of guidelines on how these metrics can inform the relaxation of restrictions, serological surveillance, and community-based testing.



A comparison of how a variety of mobility metrics relate to each other, and the change in associations as the spatial and temporal granularity increases.

Daily Debriefs

Most of the collaborations commenced with direct virtual conversations, with over 60 public agencies around the world. These interactions diversified to meet the needs of each agency's capacity, workflow and availability.

Network researchers provided services, in the form of near real time support in:

- Familiarizing government teams with how mobility data could support their response activities (often on virtual calls and emails);
- Providing routine situation reports on county level data and co-designing metrics to meet the local context (Arapahoe, CO and Los Angeles, CA);
- Adapting visualizations for team requirements (Los Angeles, CA and Chile);
- Providing direct data sources (Syracuse, NY) or integrating new data sources (Kansas and Chile);
- Providing state or county-level automated reports, which were shared with decision-makers (Massachusetts, New York, NY).

Most teams met virtually early in their agency collaborations, which were followed by ad hoc meetings as they became more familiar with the data and outputs. From April until summer of 2020, many groups connected daily or weekly via email or shared file repositories with updated maps or situation reports. During pivot points of collaboration, such as integrating a new data source, groups reconvened to resolve challenges. One group maintained a near real-time communication channel through WhatsApp, while some groups continue to work collaboratively.

Scientific Publications

To further support the the application of human mobility data for pandemic response and planning, network researchers have also published over 100+ manuscripts in pre-print servers and leading peer-reviewed scientific journals around the world, developed tutorials on accessing mobility datasets for research, proposed common frameworks for mobility data analyses, and released a number of statistical packages for epidemiological analyses. Key papers are listed in Appendix 4.

Impact Evaluation

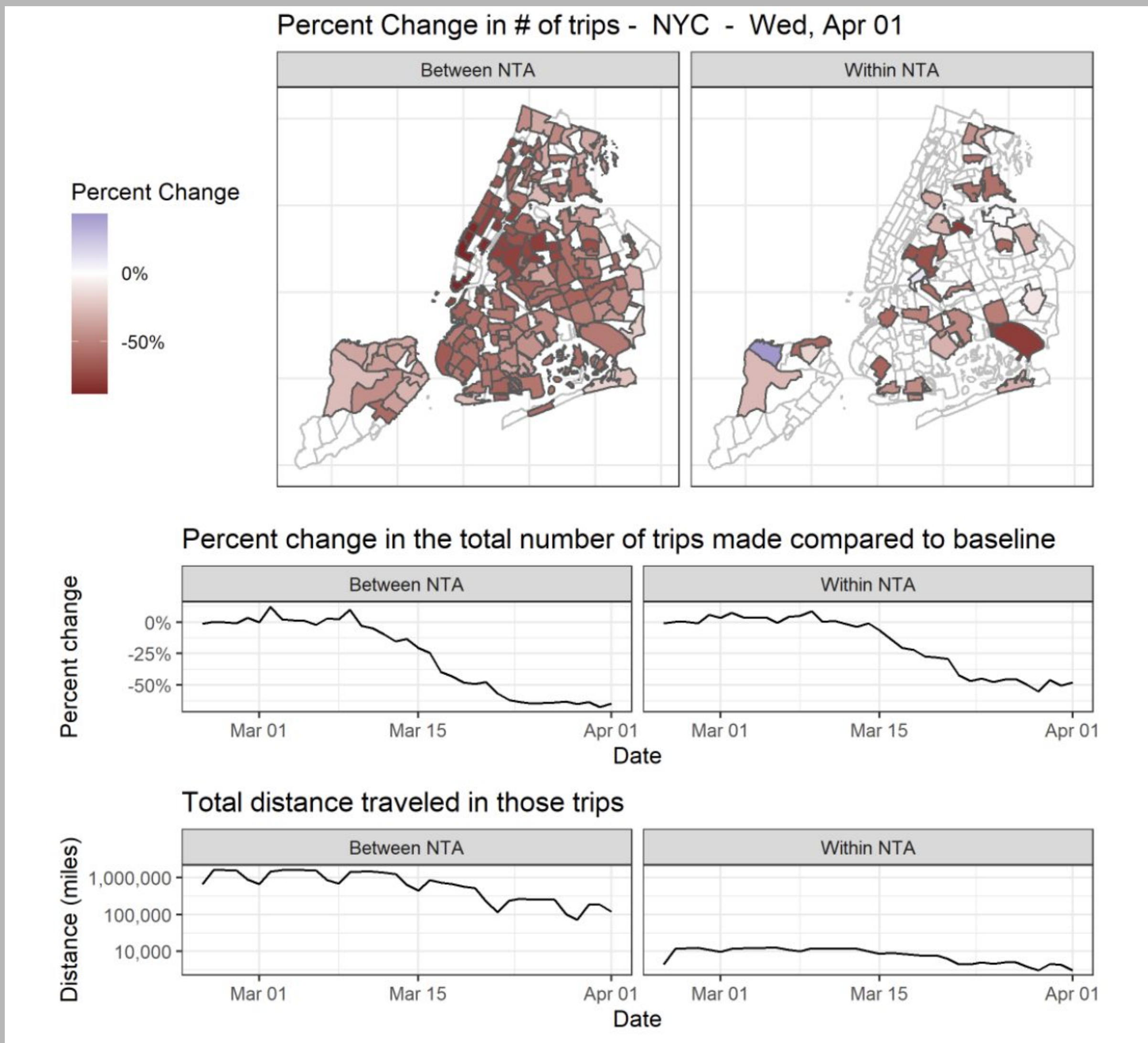
In order to understand how the data were used by both researchers and policy makers, Professor Jennifer Chan from Northwestern University, with intramural support, conducted an extensive qualitative survey across 70 Network participants from the US and abroad.

The goal of this exercise was to understand how the data were used, by whom, whether it was possible to gauge impact, and what barriers precluded more efficient use of these data. CMDN collaborators in government offices included emergency managers, chief technology officers, analysts, and consultants. There was neither a prescriptive process nor a set of requirements of how to engage and act within the network.

Key Findings

1. Automated Situation Report

The COVID-19 Mobility Data Network (CMDN) provided an automated situation report template, with the aim to provide easy access to Facebook mobility data in response to COVID-19. This provided an entry point for researchers to quickly generate reports and a way for practitioners to learn more about mobility data by further understanding its meaning for COVID-19 response activities. The situation reports also included “pointers on evaluating the data”, which recommended how to interpret the metrics, along with examples. It also described limitations of analysis, such as how to interpret short and long-distance travel across boundaries of analysis.



Jurisdiction-specific report template components - An example from New York City

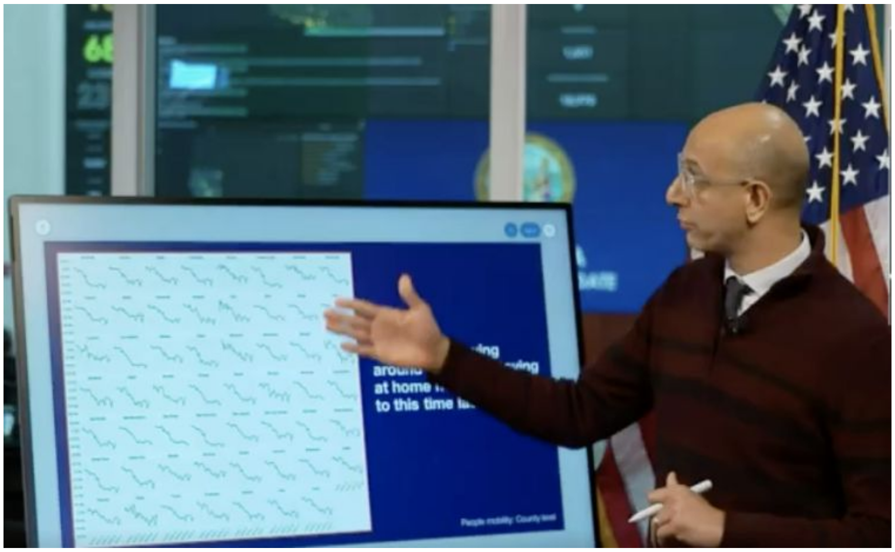
Templates showed practitioners data visualization for their specific jurisdiction, enabling conversations between researchers and practitioners by moving from the abstract view of mobility data to something tangible. This further built working relationships and a level of trust around the potential value of the data. In many jurisdictions, these templates, situation reports, and anonymized data were utilized in situation reports with government officials. We continue to provide these services even now.

2. Translating Data into Action

Information overload is a well-known challenge in disasters; too much data and information creates navigational challenges for users, often resulting in non-use. Data literacy constraints, time and capacity limitations, and increasing demand for a data-driven response create a difficult environment for practitioners to effectively use new data sources.

One early question was "*How do practitioners achieve a deeper and more targeted understanding of how this data could meet the purpose of their COVID-19 response work?*"

Specifically, how could the results of the data fulfill a purpose or need for their city, county, or region?



Dr. Mark Ghaly discusses mobility reports from the COVID19 mobility data network built with Facebook data during a California press conference on April 10, 2020

Aligning data for utilization requires a common understanding among collaborators about whether the data-derived insight corresponds to a group's specific purpose. As expected, government collaborators easily understood the relationship between population movement and the spread of COVID-19. Social distancing policies, from travel restrictions to quarantine, were a familiar lived experience around the world. Theoretically, looking at data about movement made sense, but getting to the next step was a challenge for many collaborators.

Understanding the meaning of mobility data in specific COVID-19 contexts was a much deeper and more complex question. Particular questions regarding Thailand's national response were different from questions that arose in Arapahoe, Colorado, United States. Another team in Barcelona, Spain aimed to support the European Commission with general updates, but did not intend to influence regional or local activities. Teams in Bogota, Colombia faced questions about the utility of strict quarantine measures when perceived cases were lower than other parts of the world.

Data translators use their technical skills to best align relevant data to practitioners' specific and evolving priorities. They act as a bridge between data, people, purpose, and action over a wide range of activities. They often emerge from multidisciplinary teams who are working with data in new environments or are tackling challenges with new collaborators.

Successful data translators were able to explain not only mobility data, but communicate how the results were meaningful (or not). They were often able to relate the results to the purpose and priorities of the potential key decision - makers. They were frequently able to adapt, willing to iterate, and in particular, manage the uncertainty and fluidity of the social environment.

"Translators play a critical role in bridging the technical expertise of data scientists with the operational expertise of marketing, supply chain, manufacturing, risk, and other frontline managers. In their role, translators help ensure that the deep insights generated through sophisticated analytics translate into impact at scale in an organization."

Henke, N., Levine, J. McInerney, P.
"You don't have to be a Data Scientist to Fill this Must-have Analytics Role." Harvard Business Review, 5 February 2018.

Many teams were initially limited by the expectation that data alone can provide meaningful, applicable results. Team adaptation can address this initial hurdle through dynamic learning and communication, which can help both researchers and practitioners build data literacy. In the future, a deeper understanding of why data doesn't always equate to purpose and meaning can help teams further iterate and adapt their collaborative work.

4. A Path Forward

Trust was apparent throughout CMDN and there was a collective commitment among participants to use anonymized data, as well as follow Facebook's Data Use Agreement and the Network Data Use Policy. Among researchers, the Network was able to coordinate actions in a distributed manner. Weekly network calls were opt-in and researchers convened during these virtual calls to share their ways of working, often posing challenges that other members of the groups would help with sharing ideas for potential solutions. Many practitioner interviewees expressed a strong interest and willingness to learn more about the data and how to best use it in their circumstances.

There are future opportunities for the Network to expand its scope to include some of the most crucial members of the environment for which it serves.

The structure, design, and engagement of a shared learning environment for this group will likely be different than a research network, as the demands, culture, and work environment are different than academia.

2021: Crisis Readiness

CMDN played a critical role in providing peer support to scientists when academic incentives are not necessarily aligned with humanitarian imperatives. Through this extraordinary experience, our scientists and policy makers learned important lessons about the potential for these data, their limitations, and the societal context in which they are applied.

During the 2020 hurricane season in the United States CMDN members also worked with federal partners in the US to apply human mobility data analysis to planning evacuation-shelter distribution across the Gulf Coast. We began looking at similar applications in California, in the aftermath of wildfire-related public safety power-shutoffs and evacuations, and concluded that the data streams required for critical decision-making during public health emergencies - especially novel data from social media, AdTech companies, satellite imagery, etc. -- are sequestered across public and private agencies, hard to acquire and secure in a timely manner without pre-negotiated access, sometimes across international jurisdictions. These new and large data sets are not easy to compare or combine with other data sources; yet they provide information of significant public health consequence. They also raise concerns about individual and group privacy and security. The analytic methods required to efficiently but responsibly use these data are evolving rapidly and must balance the potential for public good with privacy-preserving principles and societal norms. And finally, the capacity to then use these insights varies widely from jurisdiction to jurisdiction, certainly around the world, but even in the United States.

We observe that issues with data translation can be addressed by involving policy makers and practitioners in the early design and analytic process to ensure that scientists (and technology companies) are providing data and outputs that are most relevant to field operations. This could be done by further establishing a network model where researchers and practitioners can build a knowledge sharing environment together before, during, and after disasters.

To apply the lessons learned from CMDN, and to scale our operations to disasters beyond the pandemic, we launched **CrisisReady.io**, a response-research platform based at Harvard and Direct Relief (see our current team members in Appendix 5).

CrisisReady collaborates with academic partners, technology companies, and response agencies around the world to embed data-driven decision-making into local disaster planning. We are committed to developing scalable expertise in Data Readiness, Methods Readiness, and Translational Readiness, by securing data pipelines that provide actionable analyses to meet response agencies and policy makers

CRISISREADY: ORGANIZING PRINCIPLES

1. **DATA READINESS:** The data required to respond to disasters are pre-identified, and access is pre-negotiated across stakeholders, even if the data streams are only activated during crises.
2. **METHODS READINESS:** The data vary in nature of origin, representativeness, temporal and spatial scales, and there is an urgent need to develop standard frameworks for the analysis and interpretation of these disparate data during disasters and humanitarian emergencies.
3. **TRANSLATIONAL READINESS:** Even when high-quality data analysis is ready for near-real-time disaster response, public health departments often do not have embedded local capacity to drive data-driven response, and our network of researchers and public health professionals works to support the actionability of these data.

Looking Ahead: 2021-2022

As we chart our roadmap for the first five years of CrisisReady, seeking to become the preeminent global resource for applied data science during public health emergencies, the following projects have already been launched:

1. **Monthly Global Seminars**
2. **California Wildfires: Health System Resilience Mapping**
3. **OpenDP: Applying Differential Privacy to CrisisReady Datasets**
4. **White Paper on Human Mobility Data: Science and Society**
5. **World Bank Global Facility for Disaster Reduction and Recovery (GDRR)**
6. **Small Grants to Support Emergency Epidemiology**

1. Monthly Global Seminars

Monthly seminars open to CMDN members and a wider audience of scientists, technology companies and field practitioners seeking.

January 29, 2021 - Did data make a difference? Lessons from Facebook Data for Good in 2020

The role of Facebook Data for Good tools for COVID-19 response, routine immunizations, and development programs and what this means for COVID-19 vaccination in 2021

- **Moderator** - Andrew Schroeder, VP of Research and Analysis at Direct Relief
- **Speakers** -
  - Laura McGorman, Policy Lead for Data for Good at Facebook
  - Dr. Alex Reinhart, Assistant Professor in Statistics & Data Science at Carnegie Mellon University
  - Ayesha Durrani, Communication for Development Specialist, UNICEF/Pakistan
  - Justus Wambayi, Program Specialist, Cadasta
- **Respondent** - Dr. Jennifer Chan, Humanitarian Advisor, NetHope

2. California Wildfires: Health System Resilience Mapping

In collaboration with a range of public agencies in California, including CalOES, CDPH, CCLHO, and others, our researchers are developing decision tools for health systems to predict and respond to surges in acute and chronic care needs as a result of the wildfires, and related power outages and evacuations.

In collaboration with: Harvard, Direct Relief, Stanford, Chapman and CA public agencies

3. OpenDP: Applying Differential Privacy to CrisisReady Datasets

Our team will collaborate with Professors Gary King [Institute for Quantitative Social Sciences] and Salil Vadan [SEAS] at Harvard, to apply and OpenDP's open suite of tools to CrisisReady datasets to allow privacy preserving but wider access to our datasets. In 2021, we will first test OpenDP on Camber System's human mobility data sets. This project is supported by a grant from the Harvard Data Science Initiative.

4. White Paper on Human Mobility Data: Science and Society

Sponsored by the Radcliffe Institute of Advanced Studies, we will convened a series of four workshops, inviting domain experts from industry, applied sciences (academia), law and policy, and society and practice, to write a white paper on the current state-of-the-art and anticipated challenges in applying human mobility data (and related data streams) to public health emergencies. The publicly released report will then guide a 12 month consultation process in partnership with the World Bank described below.

6. World Bank Global Facility for Disaster Reduction and Recovery (GDRR)

In collaboration with GDRR and Northwestern University, CrisisReady will lead a 12-month global consultation process to develop a guidance document and open source tools for response agencies. The guidance document and toolkits is expected to provide: minimum standards in data, methods and translational readiness, by providing hardware, software, and training specifications for the responsible and meaningful application of such data.

7. Small Grants to Support Emergency Epidemiology

Recognizing that rapid funding for applied research during unpredictable emergencies is almost impossible to obtain using traditional routes, the small grants program aims to support and incentivize scientists to help crisis response efforts in a flexible way when needed. In order to sustain the work of key researchers in the Covid-19 Mobility Data Network who have continued to respond to the ongoing pandemic emergency CrisisReady has made two initial small sub-award grants. These grants have been issued to Dr. Pamela Martinez at the University of Illinois and Dr. Amy Wesolowski at Johns Hopkins University. Dr. Martinez's work focuses on understanding of mobility patterns and non-pharmaceutical interventions in Chile through travel surveys, while Dr. Wesolowski deepens understanding of risk and behavioral health factors in the US in combination with mobility data analysis.

In the future CrisisReady will aim to provide regular small grant support to researchers involved in emergency epidemiology in order to ensure that they are able to devote time and attention in relation to their ongoing academic work. These two grants, in addition to being valuable research in their own right, establish the framework of that future support. We hope to raise funds in 2021 to make 10-15 such small grants per year for a range of annual crises globally, for the next three years.

Appendix 1 - CMDN Network

This is a non-exhaustive list of Network researchers, practitioners, and other participants.

- Abhishek Bhatia, Harvard University
- Aditya Gopalan, Indian Institute of Technology Tirupati
- Alejandro Feged, Bogota
- Alex Perkins, Notre Dame University
- Alexandra Wood, Harvard Law School
- Ali Mostafavi, Texas A&M University
- Alpan Raval, Wadhwani AI
- Amy Wesolowski, Johns Hopkins University Bloomberg School of Public Health
- Anand A R, Indian Institute of Technology Madras
- Andy Tatem, University of Southampton
- Ankit Bhardwaj, Wadhwani AI
- Ankit Chaurasia, Wadhwani AI
- Arban Bhadra, Indian Institute of Technology Tirupati
- Arnab Jana, Indian Institute of Technology Bombay
- Ashley Xavier, Indian Institute of Technology Madras
- Ayesha Mahmud, University of California at Berkeley
- Bailey Fosdick, Colorado State University
- Balaram Ravindran, Indian Institute of Technology Madras
- Borworn Panklang, Mahidol University
- Brooke Nichols, Boston University
- Bryan Grenfell, Princeton University

- Christine Tedijanto, Harvard T.H. Chan School of Public Health
- Christophe Fraser, Oxford Big Data Institute
- Clay Heaton, HTEC, Beth Israel Deaconess Medical Center
- D. Manjunath, Indian Institute of Technology Bombay
- Daniel Klein, IDM
- Daniel Larremore, University of Colorado at Boulder
- Daniel Weinberger, Yale School of Public Health
- Dennis Chao, IDM
- G Srinivasaraghavan, Indian Institute of Technology Bangalore
- Gitakrishnan Ramadurai, Indian Institute of Technology Madras
- Hao Hu, Bill and Melinda Gates Foundation
- Harshavardhan P K, Indian Institute of Technology Madras
- Helen Wearing, University of New Mexico
- Himanshu Sinha, Indian Institute of Technology Madras
- Himanshu Tyagi, Indian Institute of Science
- Hsiao-Han Chang, National Tsing Hua University
- Jahnvi Patel, Indian Institute of Technology Madras
- Jennifer Chan, Northwestern University
- Jerome White, Wadhwani AI
- Jessica Metcalf, Princeton University
- Jigar Doshi, Wadhwani AI
- Jon Zelnor, University of Michigan
- Justin Lessler, Johns Hopkins University Bloomberg School of Public Health
- Kalidas Yeturu, Indian Institute of Technology Tirupati
- Karthik Raman, Indian Institute of Technology Madras
- Kenth Engo-Monsen, Telenor Research
- Krishna Prapoorna, Indian Institute of Technology Tirupati
- Kulchada Pongsoipetch, Mahidol University

- Lerato Magosi, Harvard T.H. Chan School of Public Health
- Linus Bengtsson, Flowminder
- Mathew Kiang, Stanford University School of Medicine
- Mauricio Santillana, Harvard Medical School
- Mihir Kulkarni, Wadhwani AI
- Mike Famulare, IDM
- Mitesh Khapra, Indian Institute of Technology Madras
- Mortiz Kraemer, University of Oxford
- Naga Siva Pavani Peraka, Indian Institute of Science
- Neeraj Agrawal, Wadhwani AI
- Nick Geard, University of Melbourne
- Nidhin Koshy, Indian Institute of Science
- Nihesh Rathod, Indian Institute of Science
- Nishant Kishore, Harvard T.H. Chan School of Public Health
- Pamela Martinez, Harvard T.H. Chan School of Public Health
- Paul Doherty, NAPSG Foundation
- Pavani Naga Siva Peraka, Indian Institute of Technology Tirupati
- Pratyush Kumar, Indian Institute of Technology Madras
- Preetam Patil, Indian Institute of Science
- Raghu Dharmaraju, Wadhwani AI
- Raghunathan Rengaswamy, Indian Institute of Technology Madras
- Rajesh Sundaresan, Indian Institute of Science
- Rebecca Harned, NAPSG Foundation
- Rebecca Kahn, Harvard T.H. Chan School of Public Health
- Rene Neihus, Harvard T.H. Chan School of Public Health, ECDC
- Richard Maude, MORU
- Rolando Acosta, Harvard T.H. Chan School of Public Health
- Roy Burstein, IDM

- Ryan Layer, University of Colorado at Boulder
- Saket Kanth, Wadhwani AI
- Sam Brand, University of Warwick
- Sam Scarpino, Northeastern University
- Sarah Cobey, University of Chicago
- Sheetal Silal, University of Cape Town
- Shekar Sivasubramanian, Wadhwani AI
- Shenyue Jia, Chapman University
- Shrisha Rao, Indian Institute of Technology Bangalore
- Siddharth Nishtala, Indian Institute of Technology Madras
- Soham Raste, Indian Institute of Technology Madras
- Srinath Srinivasa, Indian Institute of Technology Bangalore
- Stephen Riley, Imperial College
- Urs Gasser, Harvard Law School
- V Sridhar, Indian Institute of Technology Bangalore
- Vinit Mehta, Indian Institute of Technology Madras
- Yogesh Tripathi, Indian Institute of Technology Madras

Appendix 2 - Sample Blogs from CMDN Meeting Series

Covid-19 Mobility Data Network (CMDN) Meeting Series

From April 17, 2020 - December 4, 2020, we hosted a total of 21 global network-wide meetings. These meetings were archived. Blogs summarizing the discussions are publicly available on the website (https://crisisready.io/).

We have included a few examples below.

BLOG: Mobility Data Analysis and Application in the United Kingdom

The work discussed in this blog was presented to the COVID-19 Mobility Data Network on August 14, 2020, as part of regular weekly meetings that aim to share information, analytic tools, and best practices across the network of researchers.

First, we heard from Caroline Walters, from the MRC Centre for Global Infectious Disease Analysis at Imperial College London, on leveraging **crowd level mobility data to assess compliance with social distancing measures across populations in the United Kingdom.** Using anonymized data from the Facebook Data for Good program, the research team calculated the daily percentage change from baseline in the number of trips starting from within each UK local authority district. The analysis was repeated on a separate dataset from a mobile phone operator.

They found that the two mobility datasets produced similar trends, showing that mobility began to decrease around one week before the UK lockdown was enforced on March 24, 2020, and that there was little variation in trend across the 4 countries (see figures 1 and 2). They also found that there were larger reductions and greater variation in mobility in high population density areas relative to low population density areas. After the sharp decrease immediately following the lockdown, they observed a gradual and continuing increase in mobility across the UK.

The research team is currently working with the UK Government by monitoring mobility data trends in local outbreak areas and sending weekly .csv data files to aid both other researchers and policymakers in the COVID-19 response. Their manuscript is available here, with additional information on their data sources, and the code used to run the discussed analyses and plot the figures are made available through Github.

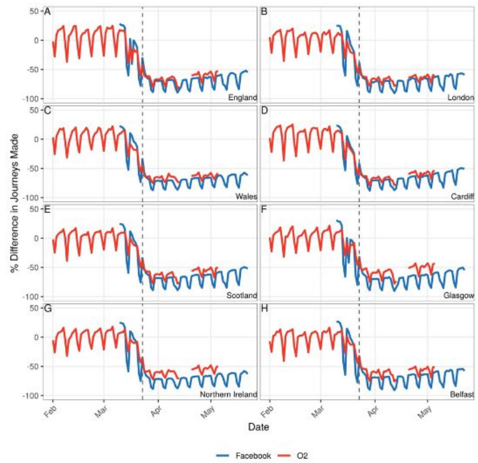


Figure 1: Change in movement over time as a percentage of baseline movement for the four home countries within the UK and their largest city for Facebook data (blue) and mobile phone data (red)

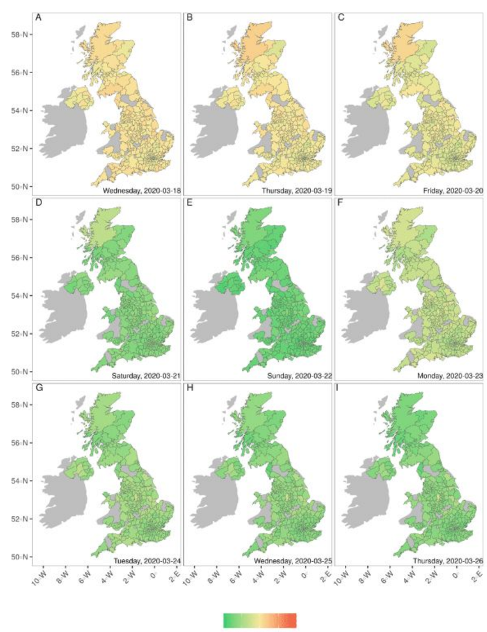


Figure 2: Reduction in mobility from March 18-26, 2020 relative to the baseline

BLOG: Mobility Data and the Limits of Data Protection Frameworks

The work discussed in this blog was presented to the COVID-19 Mobility Data Network on August 14, 2020, as part of regular weekly meetings that aim to share information, analytic tools, and best practices across the network of researchers.

The near real-time information about human movement provided by aggregated population mobility data has tremendous potential to help refine interventions when appropriate legal, organizational, and computational safeguards are in place. As the private sector, policymakers, and academia work together to leverage novel sources of data to track the spread of the pandemic, Randall Harp, Laurent Hébert-Dufresne, and Juniper Lovato from the University of Vermont argue that anonymization at the individual level is insufficient-- the notion of privacy should extend to communities as well. Information that identifies characteristics at the community level can also lead to “finger-pointing”, and disproportionately highlight marginalized groups in a large population. Ideally, one would want to add noise at the community-level to anonymize, for example, towns, balancing obfuscation of sensitive, identifiable data, with the loss of other information that may also compromise the information of public health interest.

The team observed that the presumption of *individual* control over personal data, and the assumption that consent is *informed* may no longer be valid in the context of large data sets and sophisticated, networked systems. The evolving complexities of data processing, the volume of consent needed to completely opt-out of personal data use within a complex system, and the broad-stroke terms of service that users typically consent to, often permit flexible data use that individuals may not have considered. Furthermore, in instances where the collected anonymized individual-level data are leaky, frameworks in place for individual privacy stop being protective at the population level -- populations that haven’t consented to their data being collected may still be implicated in its use.

Academia is a common source of leaky data. As researchers from CMDN discuss the different procedures they need to complete to comply with the requirements of traditional research ethics, it is evident that the existing systems of ethical review by institutional review boards (IRB) need to be updated to fit the kind of data being generated right now. The UVM team calls for the academic community to think about the regulations and structures required to equip the IRB with the mechanisms to determine the suitability of proposed analytics for these new large, complex, networked datasets. The current practice of requesting “Non-human subjects Research” exemptions/waivers in health data science, is no longer appropriate.

Network participants explored alternative models: As mobility data morph through the analytical pipeline, there is a need to consider alternatives to the heavy, content-driven architectures, and shift toward Privacy by Design frameworks, or architectures that incorporate a fiduciary responsibility within the system. Should there be a legal framework for data processors to act in a fiduciary capacity? For this to happen, there is a need to create professional guidelines and field-specific standards that would make this a feasible alternative. Current lacunae require that the use of such data retain a human in the loop, to avoid the pitfalls of black-box algorithms in machine learning.

You can find more information about the team's work on distributed consent frameworks for data privacy in their recently published manuscript here, and learn more about the other research taking place at The Vermont Complex Systems Center here.

--

With a focus on highlighting the regulatory gaps in the systems currently in place in the United States, the Brennan Center for Justice closed out the last part of our meeting with a presentation by Laura Hecht-Felella and Harsha Panduranga titled "Government Access to Mobile Phone Data for Contact Tracing - A Statutory Primer". Within the Brennan Center's scope of research, media advocacy, and legislative advocacy work, the Liberty and National Security team focuses on promoting government accountability and assuring that government use of new technologies does not violate fundamental rights. Their report examined the statutory protections that exist at the federal level in the United States to safeguard the privacy of the data collected through digital contact tracing tools, specifically cell phone location and proximity data, though their findings bear on the collection and dissemination of such data in the context of public health initiatives more broadly.

They first emphasized the lack of a comprehensive data protection law at the federal level in the United States. The patchwork of laws governing the disclosure of location data to the government by cell phone companies or smartphone app developers and data brokers does not adequately protect civil liberties. While cell phone carriers may be heavily regulated with regards to sharing individually identifiable data, gaps in this regulatory framework may permit workarounds for governments seeking proximity data without consent, and governmental entities may buy similar data from data brokers who may be legally able to purchase these data from the app developers who collect it. Within the United States, their review concluded that existing law falls short in regulating contract tracing proposals that are based on voluntary adoption by users, or aggregate data. As a whole, there are also few limitations on the sharing of aggregate data or the sharing of location information among government agencies (instead, several laws promote it). While the GDPR set the precedent for robust data protection and privacy regulations in the EU, the absence of a federal data privacy law in the United States places the onus on states to develop their own protections and make up for the gaps at the federal level, like the California Consumer Privacy Act that went into effect earlier this year.

As states turn to digital contact tracing tools, the Brennan Center believes that COVID-19 could transform the kind of data that the government is able to routinely access, with severe public health and civil rights implications. In some contexts, they point out that law enforcement has pushed to use cell phone data for policing purposes or to undermine democratic participation by reaching out to researchers to get access to data to monitor protests across the United States. Minority communities that are harder hit by COVID-19, and are also subject to higher rates of policing, will be less likely to participate in contact tracing if the data is shared with law enforcement or immigration enforcement agencies. There is an opportunity now for academics working with mobility data to chime in while app developers continue to work with states to launch these contact tracing apps and advocate for privacy frameworks to minimize civil rights violations and improve participation in digital interventions across populations. The full report by the Brennan Center for Justice can be found here.

BLOG: OpenDP: Open-source software tools for privacy-protective statistical analysis of sensitive personal data

The work discussed in this blog was presented to the COVID-19 Mobility Data Network on September 25, 2020, as part of regular weekly meetings that aim to share information, analytic tools, and best practices across the network of researchers.

The COVID-19 pandemic has underscored the importance of developing analytical pipelines that leverage data streams generated by individuals and communities to drive data-driven responses to crises. As these data increasingly exchange hands within agencies and are shared publicly for broader research, Salil Vadhan and Annie Wu from Harvard SEAS, and Navin Vembar from Camber Systems discuss the inadequacy of traditional methods for the protection of sensitive personal data.

Conventional methods of removing "personally identifiable information" or aggregating data above the individual level are antiquated, inadequate, and vulnerable to attacks. There are now numerous examples that demonstrate that the removal of PII still leaves data vulnerable to re-identification of individuals through the use of auxiliary datasets, shown through examples in medical data, and even through the Netflix database. Even with aggregate statistics, it is now possible to reconstruct almost the entire underlying dataset, demonstrated using publicly available Census data, or even to determine whether a target individual is in a dataset, for example using genomic data.

As the sophistication of these attacks on privacy grew, Differential Privacy (DP) evolved as a means of ensuring that individual-level information cannot leak when releasing statistical information. The goals of differential privacy, as detailed by the presenters, are to enable the collection, analysis, and sharing of a broad range of statistical estimates based on personal data, such as averages, contingency tables, and
synthetic data, while protecting the privacy of the individuals in the data. This is achieved by injecting small amounts of random noise into statistical computations to hide the effect of each individual subject while still allowing useful inference and statistical analysis of the population data. There is a growing base of mathematical literature showing that, in principle, DP is compatible with almost all forms of statistical analysis of populations.

The OpenDP Project is a community effort from the Institute of Quantitative Social Sciences (IQSS) at Harvard University to build a trustworthy and open-source suite of DP tools that can be easily adopted by custodians of sensitive data to make it available for research and exploration in the public interest. The goal of OpenDP is to open otherwise siloed and sequestered sensitive data to support scientifically oriented research and exploration in the public interest, including data shared from companies, government agencies, and research data repositories.

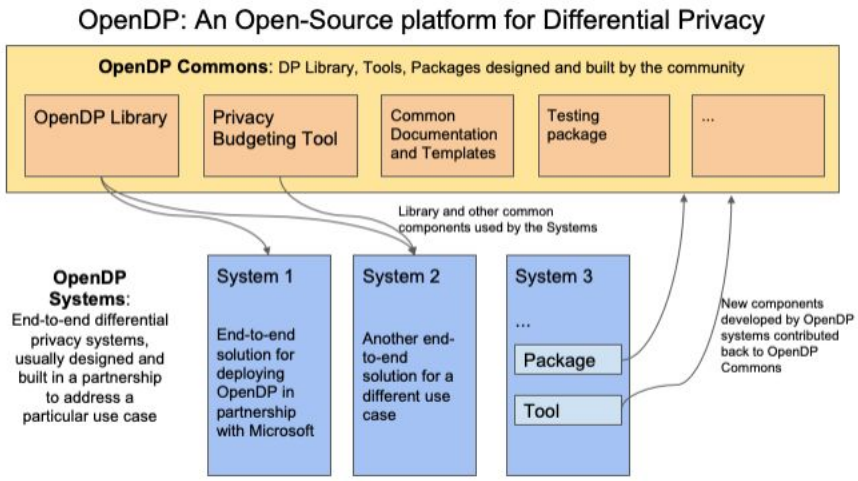


Figure 1: End-to-end privacy-preserving systems, utilizing the OpenDP Commons components.

OpenDP software aims to provide statistical functionality that is useful for the researchers who will analyze data while exposing measures of utility and uncertainty that will help researchers avoid drawing incorrect conclusions due to the noise introduced for privacy. It was developed to channel the collective advances in the DP community on science and practice, enable wider adoption to address compelling use cases, and identify important research directions for the field. Some of the high-priority use cases identified by OpenDP include:

1. Archival data repositories to offer academic researchers privacy-preserving access to sensitive data.
2. Government agencies to safely share sensitive data with researchers, data-driven policymakers, and the broader public.
3. Companies to share data on their users and customers with academic researchers or with
   institutions that bring together several such datasets.
4. Collaborations between government, industry, and academia to provide greater access to and
   analysis of data that aids in understanding and combating the spread of disease.

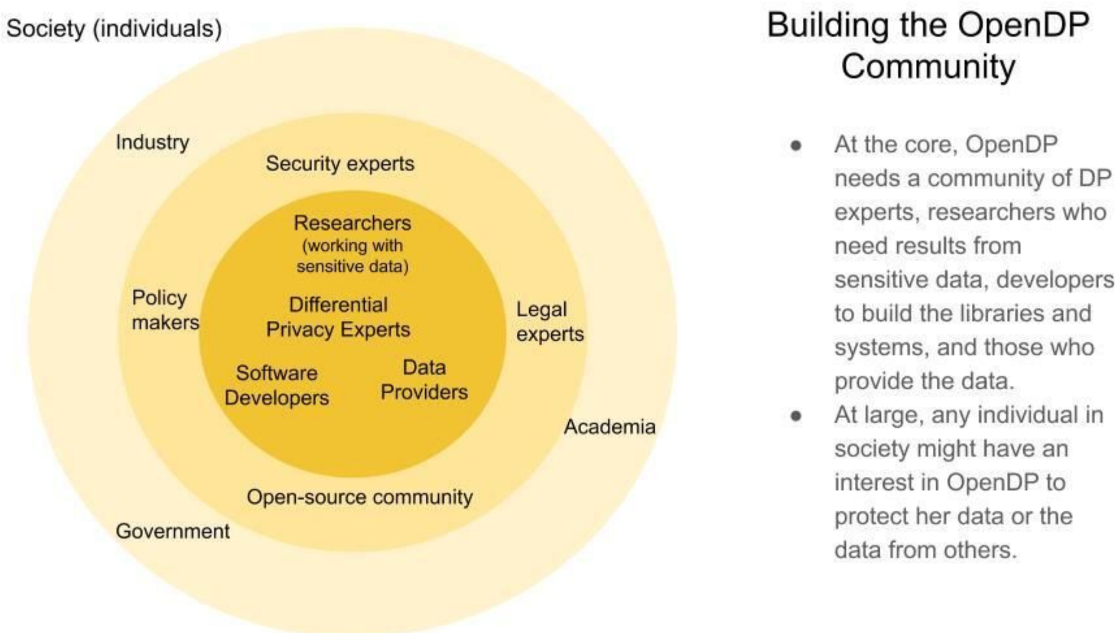


Figure 2: Definition of the OpenDP Community

The future collaboration between IQSS and CrisisReady will address the challenges of enforcing privacy-protective analysis at scale, by embedding differential privacy by design into developed data pipelines.

As the OpenDP community continues to grow, you can stay involved with the team's work by looking through their whitepapers, videos, and software at http://opendp.io/, exploring their Github, subscribing to their blog, joining the community mailing list, or get in touch at info@opendp.io.

Appendix 3 - Panel Themes and Speakers

- **Stanford Data Science: COVID-19 Data Forum**
- **Harvard Data Science Initiative Special Event: Trust in Science, Trust in Democracy**
- **Using Human Mobility Data to Inform Pandemic Response Theory and Practice**
- **CrisisReady Inaugural Panel - Data, Equity, and Wildfires: California 2020**

Stanford Data Science: COVID-19 Data Forum on Dec 10, 2020, 12:00 - 1:30 PM (ET)

- Speakers:
  - Caroline Buckee - Associate Professor of Epidemiology and Associate Director of the Center for Communicable Disease Dynamics at the Harvard T.H. Chan School of Public Health
  - Andrew Schroeder, PhD- Vice-president Research & Analytics for Direct Relief
  - Luca Ferretti, Senior Researcher in Statistical Genetics and Pathogen Dynamics, Big Data Institute, Oxford University, UK.
- Moderator: Chris Volinksky, PhD- Associate vice-president, Big Data Research, ATT Labs.
- More details on the COVID-19 Data Forum can be found here: https://covid19-data-forum.org

Harvard Data Science Initiative Special Event: Trust in Science, Trust in Democracy on Dec 7, 2020, 1:00 - 2:00 PM (ET)

- Speakers:
  - Marc Lipsitch, Professor of Epidemiology Director of the Center for Communicable Disease Dynamics. Associate Director of the Interdisciplinary Concentration in Infectious Disease Epidemiology at the Harvard T.H. Chan School of Public Health
  - Ken Prewitt, Carnegie Professor of Public Affairs, School of International and Public Affairs (SIPA) at Columbia University
  - Ruha Benjamin, professor of African American studies at Princeton University, founding director of the IDA B. WELLS Just Data Lab and author of two books, People's Science and Race After Technology, and editor of Captivating Technology.
  - Chris Mooney, Reporter covering climate change, energy and the environment at The Washington Post.
- Led by Sheila Jasanoff, Pforzheimer Professor of Science and Technology Studies at Harvard Kennedy School. The HDSI's Trust in Science Project is a collaboration between the Harvard Data Science Initiative and the Program on Science, Technology & Society.

Using Human Mobility Data to Inform Pandemic Response Theory and Practice on Dec 4, 2020, 10:00 - 11:30 AM (ET)

- Speakers:
  - Rafael Araos, MD, MMSc, Assistant Professor, Clínica Alemana, Universidad del Desarrollo
  - Greg Wellenius, ScD, Professor, Dept. of Environmental Health, Boston University School of Public Health, Visiting Scientist, GoogleHealth

- Jaimie Shaff, MPH, MPA, Director: Integrated Data Team Emergency Response Group, New York City Department of Health and Mental Hygiene
- Jure Leskovec, PhD, Associate Professor, Dept. of Computer Science, Stanford University, Chief Scientist, Pinterest, Investigator, Chan Zuckerberg Biohub. Author of the recent Nature article Mobility network models of COVID-19 explain inequities and inform reopening

- Hosted by Caroline Buckee, DPhil, Associate Professor of Epidemiology at Harvard School of Public Health, and co-founder of CrisisReady

CrisisReady Inaugural Panel - Data, Equity, and Wildfires: California 2020 on Oct 23, 2020 11:00 AM - 12:00 PM (ET)

- Speakers:
  - L. Vance Taylor, Chief of the Office of Access and Functional Needs, California Governor's Office of Emergency Services
  - Jason Vargo, Research Scientist, California Department of Public Health
  - Joan Casey, Environmental Health Sciences, Columbia Mailman School of Public Health
  - Mathew Kiang, Department of Epidemiology and Population Health, Stanford University School of Medicine
- In discussion with Andrew Schroeder, VP of Research and Analysis at Direct Relief and Satchit Balsari, Assistant Professor, Emergency Medicine at Harvard Medical School.

Appendix 4 - Key Academic Outputs

The studies below represent only the publications arising from Harvard/CCDD-affiliated researchers using mobility data, and is not a comprehensive list of all academic papers arising from the CMDN network. On our CrisisReady website, we will be maintaining and curating a list of resources, including academic publications, for researchers, public health practitioners, and the public, on these topics. The academic outputs below will form part of this resource list.

1. Balsari, S., Buckee, C., & Khanna, T. (2020, May 8). Which Covid-19 Data Can You Trust? *Harvard Business Review*. https://hbr.org/2020/05/which-covid-19-data-can-you-trust
2. Balsari, S., Sange, M., & Udwadia, Z. (2020). COVID-19 care in India: The course to self-reliance. *The Lancet Global Health*, *8*(11), e1359–e1360. https://doi.org/10.1016/s2214-109x(20)30384-3
3. Buckee, C. (2020). Improving epidemic surveillance and response: Big data is dead, long live big data. *The Lancet Digital Health*, *2*(5), e218–e220. https://doi.org/10.1016/s2589-7500(20)30059-5
4. Buckee, C. O., Balsari, S., Chan, J., Crosas, M., Dominici, F., Gasser, U., Grad, Y. H., Grenfell, B., Halloran, M. E., Kraemer, M. U. G., Lipsitch, M., Metcalf, C. J. E., Meyers, L. A., Perkins, T. A., Santillana, M., Scarpino, S. V., Viboud, C., Wesolowski, A., & Schroeder, A. (2020). Aggregated mobility data could help fight COVID-19. *Science (New York, N.Y.)*, *368*(6487), 145–146. https://doi.org/10.1126/science.abb8021
5. Chan, J. L., & Purohit, H. (2020). Challenges to Transforming Unconventional Social Media Data into Actionable Knowledge for Public Health Systems during Disasters. *Disaster Medicine and Public Health Preparedness*, *14*(3), 352–359. https://doi.org/10.1017/dmp.2019.92
6. Chang, M.-C., Kahn, R., Li, Y.-A., Lee, C.-S., Buckee, C. O., & Chang, H.-H. (2020). Variation in human mobility and its impact on the risk of future COVID-19 outbreaks in Taiwan. BMC Public Health, 2021 Jan 27;21(1):226. doi: 10.1186/s12889-021-10260-7. PMID: 33504339.

7. Choudhury, P., Koo, W. W., Li, X., Kishore, N., Balsari, S., & Khanna, T. (2020). *Food Security and Human Mobility During the COVID-19 Lockdown* (SSRN Scholarly Paper ID 3600376). Social Science Research Network. https://doi.org/10.2139/ssrn.3600376

8. De Salazar, P. M., Niehus, R., Taylor, A., Buckee, C. O., & Lipsitch, M. (2020). Identifying Locations with Possible Undetected Imported Severe Acute Respiratory Syndrome Coronavirus 2 Cases by Using Importation Predictions. *Emerging Infectious Diseases*, *26*(7), 1465–1469. https://doi.org/10.3201/eid2607.200250

9. Feehan, D., & Mahmud, A. (2020). Quantifying population contact patterns in the United States during the COVID-19 pandemic. *MedRxiv*, 2020.04.13.20064014. https://doi.org/10.1101/2020.04.13.20064014

10. Grantz, K. H., Meredith, H. R., Cummings, D. A. T., Metcalf, C. J. E., Grenfell, B. T., Giles, J. R., Mehta, S., Solomon, S., Labrique, A., Kishore, N., Buckee, C. O., & Wesolowski, A. (2020). The use of mobile phone data to inform analysis of COVID-19 pandemic epidemiology. *Nature Communications*, *11*(1), 4961. https://doi.org/10.1038/s41467-020-18190-5

11. Greene, S. K., McGough, S. F., Culp, G. M., Graf, L. E., Lipsitch, M., Menzies, N. A., & Kahn, R. (2021). Nowcasting for Real-Time COVID-19 Tracking in New York City: An Evaluation Using Reportable Disease Data From Early in the Pandemic. *JMIR Public Health and Surveillance*, *7*(1), e25538. https://doi.org/10.2196/25538

12. Kiang, M. V., Santillana, M., Chen, J. T., Onnela, J.-P., Krieger, N., Engø-Monsen, K., Ekapirat, N., Areechokchai, D., Prempree, P., Maude, R. J., & Buckee, C. O. (2021). Incorporating human mobility data improves forecasts of Dengue fever in Thailand. *Scientific Reports*, *11*(1), 923. https://doi.org/10.1038/s41598-020-79438-0

13. Kishore, N., Kahn, R., Martinez, P. P., Salazar, P. M. D., Mahmud, A. S., & Buckee, C. O. (2020). Lockdown related travel behavior undermines the containment of SARS-CoV-2. *MedRxiv*, 2020.10.22.20217752. https://doi.org/10.1101/2020.10.22.20217752

14. Kishore, N., Kiang, M. V., Engø-Monsen, K., Vembar, N., Schroeder, A., Balsari, S., & Buckee, C. O. (2020). Measuring mobility to monitor travel and physical distancing interventions: A common framework for mobile phone data analysis. *The Lancet. Digital Health*, *2*(11), e622–e628. https://doi.org/10.1016/s2589-7500(20)30193-x

15. Kishore, N., Mitchell, R., Lash, T. L., Reed, C., Danon, L., Sigmundsdóttir, G., & Vigfusson, Y. (2020). Flying, phones and flu: Anonymized call records suggest that Keflavik International Airport introduced pandemic H1N1 into Iceland in 2009. *Influenza and Other Respiratory Viruses*, *14*(1), 37–45. https://doi.org/10.1111/irv.12690

16. Kissler, S. M., Kishore, N., Prabhu, M., Goffman, D., Beilin, Y., Landau, R., Gyamfi-Bannerman, C., Bateman, B. T., Snyder, J., Razavi, A. S., Katz, D., Gal, J., Bianco, A., Stone, J., Larremore, D., Buckee, C. O., & Grad, Y. H. (2020). Reductions in commuting mobility correlate with

geographic differences in SARS-CoV-2 prevalence in New York City. *Nature Communications*, *11*(1), 4674. https://doi.org/10.1038/s41467-020-18271-5

17. Layer, R. M., Fosdick, B., Bradshaw, M., Larremore, D. B., & Doherty, P. (2020). Case Study: Using Facebook Data to Monitor Adherence to Stay-at-home Orders in Colorado and Utah. *MedRxiv*, 2020.06.04.20122093. https://doi.org/10.1101/2020.06.04.20122093
18. Mahmud, A. S., Chowdhury, S., Sojib, K. H., Chowdhury, A., Quader, M. T., Paul, S., Saidy, M. S., Uddin, R., Engo-Monsen, K., & Buckee, C. O. (2020). Participatory syndromic surveillance as a tool for tracking COVID-19 in Bangladesh. *MedRxiv*, 2020.08.28.20183905. https://doi.org/10.1101/2020.08.28.20183905
19. Martinez de Salazar Munoz, P., Niehus, R., Taylor, A., Buckee, C., & Lipsitch, M. (2020). *Using predicted imports of 2019-nCoV cases to determine locations that may not be identifying all imported cases*. https://doi.org/10.1101/2020.02.04.20020495
20. Menkir, T. F., Chin, T., Hay, J. A., Surface, E. D., De Salazar, P. M., Buckee, C. O., Watts, A., Khan, K., Sherbo, R., Yan, A. W. C., Mina, M. J., Lipsitch, M., & Niehus, R. (2021). Estimating internationally imported cases during the early COVID-19 pandemic. *Nature Communications*, *12*(1), 311. https://doi.org/10.1038/s41467-020-20219-8

Appendix 5 - Our Team Now

Abhishek Bhatia

Abhishek Bhatia is the IDHN Data Science Fellow at the Mittal Institute, where his research spans across digital health, health data science, and disaster response. His current work with Prof. Satchit Balsari and Prof. Caroline Buckee focuses on (1) assessing data readiness, methods readiness, and translational readiness for equitable and crisis planning and response, in the context of COVID-19 and the California Wildfires, (2) analyzing epidemiological and population-level data to assess the public health burden of war and disaster in the context of the floods in Kerala, India, and the Syrian Crisis, and (3) researching the combined application of task-shifting, training, and technology in primary care settings in India.

Andrew Schroeder

Dr. Andrew Schroeder is the Vice President of Research and Analysis for Direct Relief. He leads Direct Relief's work in GIS mapping, epidemiological analysis and humanitarian informatics. His work has been published or featured in publications including Science, The Lancet, The New York Times, The Washington Post, Fast Company, Motherboard Vice, Wired, The New Humanitarian, Prehospital and Disaster Medicine, and the International Journal of Cancer. He has worked in a consulting and advisory capacity for the United Nations Development Program (UNDP) and the World Food Programme (WFP). Andrew co-founded CMDN.

Caleb Dresser

Caleb Dresser is the 2019 to 2021 Fellow in Climate and Human Health of the LCF Consortium on Climate Science and Health Policy through the Department of Emergency Medicine at Beth Israel Deaconess Medical Center, Harvard C-Change and the FXB Center for Health and Human Rights at Harvard, and and practices emergency medicine as a member of Harvard Medical Faculty Physicians.. Caleb's current work focuses on the hazards posed by extreme heat events and weather-related electrical outages for patients in communities near Boston, including the threat that these can pose to patients with specific medical vulnerabilities. He is also examining the long-term health impacts of hurricanes and other climate-related disasters, including issues of prolonged loss of access to medical services and temporary and permanent migration of affected populations.

Caroline O' Flaherty Buckee

Dr. Caroline Buckee is Associate Professor of Epidemiology, and Associate Director of the Center for Communicable Disease Dynamics, at Harvard Chan. The Buckee lab uses mathematical models and data science to understand the mechanisms driving the spread of infectious diseases, with a focus on pathogens like malaria that affect vulnerable populations in low income countries. After receiving a D.Phil from the University of Oxford, Caroline worked at the Kenya Medical Research Institute to analyze clinical and epidemiological aspects of malaria as a Sir Henry Wellcome Postdoctoral Fellow. Her work led to an Omidyar Fellowship at the Santa Fe Institute, where she developed theoretical approaches to understanding malaria parasite evolution and ecology. In 2013 Dr. Buckee was named one of MIT Tech Review's 35 Innovators Under 35, a CNN Top 10: Thinker, and Foreign Policy Magazine's Global Thinkers. Her work has appeared in high profile scientific journals such as Science and PNAS, as well as being featured in the popular press, including CNN, The New Scientist, Voice of America, NPR, and ABC.

Jennifer Chan

Dr. Jennifer Chan is an Associate Professor at Northwestern University, and hybrid researcher-practitioner with over 15 years of experience working with NGOs and UN agencies as a public health, crisis informatics and humanitarian data specialist. She co-authored *Disaster 2.0*, describing the challenges of information sharing and technology in Haiti, and built the Crisis Informatics program at NetHope during the West Africa Ebola Response. Her transdisciplinary research in the information and computer sciences, engineering, and epidemiology, focus on the translation of scientific methods into practice. She collaborates with research teams and practitioners to help strengthen the theoretical and methodological assumptions of how scientific applications translate into practice with a focus on operational impact and ultimately decision-making. She currently advises both research and humanitarian organizations including NetHope, and formerly the United Nations Center for Humanitarian Data, and the International Federation of the Red Cross.

Mathew Kiang

Dr. Mathew V. Kiang is an Instructor in the Department of Epidemiology and Population Health at Stanford University School of Medicine. He also works with Stanford's Center for Population Health Sciences and is an FXB Fellow at the Harvard François-Xavier Bagnoud Center for Health and Human Rights. His research lies at the intersection of computational social science and social epidemiology. He works with individual-level, non-health data (e.g., GPS, accelerometer, and other sensor data from smartphones), traditional health data (e.g., health systems or death certificate data), and third-party data (e.g., cellphone providers or ad-tech data). To do this, he uses a variety of methods such as joint Bayesian spatial models, traditional epidemiologic models, dynamical models, and demographic analysis.

Mercedes D. Erdey

Mercedes' career in geo-spatial analysis has taken her many places from marine applications and seafloor analysis to building datasets for mapping applications. She has been able to apply the combination of her analytical insight, extensive project management skills, knowledge, and experience to the development and deployment of multiple effective geo-spatial data products and internet applications. Her experience with building and strengthening cross-functional teams has established team cohesion that ensures successful product delivery.

Navin Vembar

Navin Vembar is the Chief Technology Officer and Co-founder of Camber Systems, lending nearly 20 years of experience in government tech and a background in mathematics, data, and software to help make the world safer and better at every level of government. Navin got his start in government as a GS-7 software developer at the Bureau of Labor Statistics, and then moved on to run large programs at the FAA, including one that modernized the exchange of geospatial aeronautical information throughout the national airspace. Most recently, he served as the Chief Technology Officer of the General Services Administration.

Nishant Kishore

Nishant is a doctoral student in the Epidemiology concentration working with Dr. Caroline Buckee at the Center for Communicable Disease Dynamics, Harvard T.H. Chan School of Public Health. Nishant is interested in developing and implementing epidemiological tools that leverage high-throughput and "big" data for the evaluation of and intervention for infectious disease transmission dynamics in an increasingly urbanized and connected world. He is also the CEO and Co-Founder of EpiTech Consultants, a public health technology consulting firm that specializes in the development of analytic pipelines for complex public health problems.

Sarah Tsay

Sarah Tsay is a Doctor of Public Health (DrPH) student and Prajna Leadership Fellow at Harvard T.H. Chan School of Public Health. Sarah is the Monitoring and Evaluation Program Manager at NYC Emergency Management (NYCEM), where she guides internal policies in emergency response and monitoring and evaluation. She has created innovative public health planning efforts at NYCEM by integrating response operations with public health research methods, such as a mixed-methods approach in addressing healthcare surge for nursing homes and a tool to estimate citywide EMS transportation time during large-scale evacuations. Ms. Tsay served as a Health and Medical Emergency Support Function Coordinator in NYC's response to the COVID-19 pandemic, addressing immediate needs across the healthcare sector (e.g., healthcare staffing, alternate care sites, procurement and distribution of critical resources).

Satchit Balsari

Dr. Satchit Balsari is Assistant Professor in emergency medicine at Harvard Medical School and Beth Israel Deaconess Medical Center, and in Global Health and Population, at the Harvard TH Chan School of Public Health. His interdisciplinary research at the intersection of data science, disaster response, and population health has resulted in innovative applications of mobile, cloud-based technology to address public health challenges in disasters and humanitarian crises around the world, including at two of the world's largest mass gatherings. His work has been featured in major leading news outlets in the US and India. Satchit co-founded CMDN.

Shenyue Jia

Dr. Shenyue Jia is the Data Scientist for California Wildfires Project@CrisisReady. She is also an adjunct professor at Schmid College of Science and Technology, Chapman University. Her work lies at the convergence of wildfire ecology and geographical information sciences (GIS). As a remote sensing specialist, she conducts wildfire risk modeling using satellite imagery and hydrological data. Her interest in GIS and spatial data analysis drives her recent study to utilize crowdsourced geolocation data for wildfire and post-disaster recovery. She uses satellite fire progression data and population data with time and geolocation information to track hotspots of population under the evacuation order and identify the most vulnerable communities.