

Overview of Data Collection

Version 2.0 | Created June 16, 2020 | Last Updated Feb 22, 2021

Primary Objective

We are creating a database of the digital contact tracing technologies (DCTT) being used by public health authorities (PHA) in the United States to respond to the spread of SARS-COV-2, the virus responsible for COVID-19. While we collect data on many aspects of each individual DCTT, we are particularly interested in the geospatial technologies used in these systems, the location data they produce, and how that data is managed, stored, transferred and shared. Our focus is also on automated contact tracing systems that typically take the form of apps that use the geospatial technologies embedded in mobile devices to gather location or proximity data.

In the United States, there is no single federal system for COVID-19 contact tracing. Instead contact tracing is administered at the state and local level. For this reason, we are creating a list of DCTT that have been, or are scheduled to be, adopted by state governments. In selected instances (e.g., large urban centers with high caseloads), we will also gather data from locally deployed DCTT. State and local PHA are responsible for creating the apps used in DCTT, but build these apps using application programming interfaces (API) designed to function within the operating systems of mobile devices. The apps then use geospatial technologies (e.g., Bluetooth, GPS) embedded in those devices to gather geospatial data. For this reason, our database seeks to capture details about the design and administration of these apps, but also supporting information about the geospatial technologies they employ.

Our data collection effort is designed around the evaluation criteria and general principles for the assessment of DCTT presented by the Center for Disease Control and Prevention (1), the American Civil Liberties Union (2), the Johns Hopkins Project on Ethics and Governance of Digital Contact Tracing Technologies (3), and the Harvard Edmond J Safra Center for Ethics (4).

Data Collection Overview

Data collection proceeds in three phases.

- I. Initial Contact and System Scoping:** During this initial phase members of the research team collect data about the DCTT a state or municipality is planning on using and the system the state presently intends to develop. Data collection at this stage is designed to provide an overview of state plans and technology usage. A secondary, but essential, goal is to establish lines of communication with key officials for potential future data collection.
- II. System Details:** In this second phase of data collection research team members will gather data on the DCTT used by states and operational details of their systems. Data collection will focus on the elements outlined in the DCTT evaluation document. DCTT details will only be collected for states and municipalities that signaled their intention to develop DCTT during phase 1 of data collection. If in future repeat evaluations (see phase 3) as a state signals the development of a DCTT, detailed information will be collected following the protocols of this phase. We will also attempt to collect summary information about why selected states chose to not pursue DCTT.
- III. Repeat Evaluations:** At ~6 week intervals, the team will i) update the progress and details of in development DCTT, ii) update if a state that initially signalled that they were not developing a DCTT has subsequently altered course and begun development of a DCTT, or iii) ceased DCTT development. The purpose of repeat evaluations is to reaffirm initial data collection and track changes through time.

Data Collection Protocol - Phase I

Research team members create an overview of (1) the development status of DCTT for each state and (2) key technologies and platforms involved in DCTT development. Development status data will be generated by contacting state and municipal PHA and gathering the data outlined in *Protocol-Phase I*.

PHAContact. PHA will be provided with the *Information Sheet for Public Health Authorities* to ensure that they are informed of the purpose and protocols of the project. PHA responses will be kept in a data table that is linked to their contact information by an anonymous key. No identifying information will be solicited or included with PHA responses and respondents, no personal information will be gathered, nor will their opinions be asked. Key technologies will be reviewed by gathering and analyzing publicly available documents that outline the specifications of those technologies.

Project leaders will review notes taken during contact attempts to ensure no information about the actions of individual participants or their opinions was asked or recorded, and that no identifying information is included.

Data Collection Protocol - Phase II

Research team members will conduct an in-depth analysis of the state DCTT identified in Phase I using the DCTT Evaluation Document to perform a review of the technologies and individual follow-up. Team members will first attempt to gather necessary information using publicly available data. The team will then ask PHA to complete the same document to confirm the initial evaluation and address missing data. As in Phase I, no identifying information will be solicited or included with PHA responses and respondents, no personal information will be gathered, nor will their opinions be asked.

At the same time, a subset of research team members will gather technical details about the geospatial technologies, OS, API used by DCTT. Information will be gathered from a review of 1) available technical documentation, 2) academic and industry literature on technical capabilities, and, if necessary, 3) information requests from individual developers. This information will be organized and stored in a common repository and, when possible under licensing restrictions, shared publicly.

Data Collection Protocol - Phase III

At ~6 week intervals, the team will i) update the progress and details of in development DCTT, ii) update if a state that initially signalled that they were not developing a DCTT has subsequently altered course and begun development of a DCTT, or iii) ceased DCTT development. Repeat evaluations will follow protocols adapted from Phase I. If a change in DCTT development is signaled, in depth data collection on the technology will be reconducted using an adapted Phase II protocol. Updated information will be stored as new entries in original data tables to facilitate change tracking and analysis.

References

1. Center for Disease Control and Prevention, "Preliminary Criteria for the Evaluation of Digital Contact Tracing Tools for COVID-19 – Version 1.2" (U.S. Center for Disease Control and Prevention, 2020).
2. D. Gilmor. Principles for Technology-Assisted Contact-Tracing. *American Civil Liberties Union* (2020) (May 28, 2020).
3. J. Kahn, Ed., *Digital Contact Tracing For Pandemic Response* (Johns Hopkins University Press, 2020) <http://doi:10.1353/book.75831>.
4. V. Hart, *et al.*, Outpacing the Virus: Digital Response to Containing the Spread of COVID-19 while Mitigating Privacy Risks, COVID-19 Rapid Response Impact Initiative – White Paper 5 (Harvard University - Edmond J. Safre Center for Ethics, 2020).