Statement of Brian McClendon
Chief Executive Officer
CVKey Project

Privacy, Disclosure, and Opt-in Data collection in the age of COVID-19

Before the
Task Force on Artificial Intelligence
US House Committee on Financial Services
United States House of Representatives
July 8, 2020

Chairman Foster, Ranking Member Loudermilk, and Members of the Task Force, my name is Brian McClendon, and I'm the CEO of CVKey Project and a Research Professor at the University of Kansas. Previously, I spent 10 years leading the teams that built Google Maps, Google Earth, Streetview, and many other geo-related services. Thank you for giving me the chance to speak before the Task Force today. In my testimony, I will describe how privacy, disclosure, and opt-in data collection impact our ability to identify, locate, and isolate those who have been exposed or infected by COVID-19.

Since March of this year, I have been leading a 501(c)(3) called "**CVKey Project**", focused on helping America "re-open responsibly" given the on-going pandemic. Together with my team of world class engineers and product developers, we believe that user-data privacy is key to app adoption by Americans and our first app, CVKey focuses on individual's symptom checking, policy communication across communities, and access control into venues such as universities, workplaces, schools, and other venues. As we have recently seen in the media, there is significant need for improved policies and public communication if we have hope of operating at any where close to normal for the next 18 months.

Artificial Intelligence (AI) and Machine Learning (ML) require large amounts of data collected for training purposes. This "ground truth" data helps algorithms figure out how to make better predictions. The most valuable data to combat COVID-19 can be found in the contact-tracing interviews of infected and exposed people. Early detection, reporting, and testing leads to quick self-isolation and quarantine and can shut down the spread of the virus faster than any other method but only if it is resourced sufficiently and executed well.

The phone calls that contact-tracers make are often not answered because most people don't pick up from an unknown number these days. When the contact-tracer does reach a potential

“case”, they have to make a decision about whether to ask the person to quarantine, based on what they learn about that person’s risk during their conversation. Without quick information from the infected or exposed, the virus is quietly spreading.

What exposure notification provides is a way to use cellphone data to detect (after the fact) whether you were near someone who later tests positive for COVID-19. By notifying you, and informing you of next steps (usually to call the contact-tracing team, get tested, and/or quarantine), these new apps offer a way to help contact-tracing scale with less effort, more accuracy, and more coverage than is possible otherwise.

There are at least two ways these apps can work:

1. Using GPS location information logs to compare where the infected person was relative to everyone else. The first problem with this is that GPS locations are not accurate enough, even outdoors, to ensure that someone was really within 6’ for 10 minutes. The other, more serious problem with location data is that it contains personally-identifiable-information (PII) that can be impossible to algorithmically remove. If I can guess where someone works and lives, I can easily figure out who that person is, usually by pulling additional data from commercially available data sources like Equifax. Naive implementations of this method were deployed in UK, North Dakota, and Utah and did not meet with much success for multiple reasons:
    a. The biggest one was fear that either “big gov” or “big tech” was tracking anyone who installed it.
    b. Always-on GPS location collection has a material impact on phone battery life, so folks turned off or uninstalled it.
    c. Early implementations **did** upload data to the government and without sufficient protection that data could then be exposed/stolen by others, even if it wasn’t misused by the government.
2. A better solution is to use low-energy Bluetooth (BTLE) signals to allow phones to record when they are near other phones. A naive implementation could still allow tracking, but luckily Apple & Google came to an agreement and a privacy-protecting method called Google/Apple Exposure Notification (GAEN) was developed that worked between Android and iOS.
    a. Apple and Google then asked each country’s public health agency to build/release and app using this tech. In the United States, they weren’t delusional enough to assume a single federal solution, so they are working with one group per state.
    b. Like Dr. Raskar’s Safepaths group, my team is building an app using this technology, which we believe preserves privacy **and** works better than GPS solutions.
    c. The goal would be to get as many people as possible in a state to download and install this app, opting-in to sending/receiving this data between phones.

d. It uses rotating random keys broadcast from each phone and recorded by other phones, to be the connecting data. Think of it as a private, low-power version of Bluetooth “pairing mode”.
e. The data being recorded is **not** uploaded to the cloud or to Google or to Apple!
f. When a user of the app tests positive for COVID-19, they receive a call from the contact-tracing team. One of the action-items is to ask the infected person to upload their random keys to a state-run cloud server. **These keys contain no PII.** Even to upload the data, the infected person needs to **choose to opt-in again**.
g. Next, everyone else’s app downloads those keys and compares them to what they have recorded. This comparison occurs on each person’s phone. Ag**ain, nothing else is uploaded.** Based on adjustable algorithms, the app determines whether a possible exposure has occurred. If so, it provides and in-app notification to the user with the appropriate next steps which, again, could be:
    1. Call contact-tracing team
    2. Get tested
    3. Self-quarantine
h. By giving quick contact and information to the potentially exposed person, this system has the opportunity to greatly increase coverage and quality of contact tracing.
i. Alabama, South Carolina, Virginia, and the state of Washington have publicly announced they’ll be deploying apps based on GAEN technology. North Dakota is switching their app to use GAEN.

As with any app, but even more so here, the challenge is to get enough people to install and opt-in. These apps only work when both parties have it installed, so even if a uniform 40% of people install it, only 16% of exposures will be identified. Targeting specific audiences like universities, specific towns, or military bases with well-worded motivations to install could get a much higher penetration and, therefore, coverage.

Again, the message the user’s data is private to their phone and anything more will require an additional opt-in, is a very strong marketing message for Americans.

Thank you.

---------- end of spoken ----------

COVID-19 Top priorities should be:

- Sufficient testing so that every symptomatic person and most exposed people can get tested, with results back in less than 24 hours
- Sufficient staffing of contact-tracing teams
- Quick self-isolation of the infected, quick quarantine of those exposed
- Everyone wears a mask indoors
- Match reopening policies in a county to the current risk of spread in that county using science and metrics

Legislative Recommendations:

- Reselling of data without opt-in is a serious privacy problem. Only aggregates that use differential-privacy should be shared/sold outside the company. GDPR and CCPA are interesting prototypes, but the USA should design a system that is easier to comply with, but has no loopholes.
- Fund contact-tracing for all 50 states, 3007 counties, and 6 territories. It requires hiring (employing(!)) a lot of people to do important work to reduce the spread of the virus.
- Support digital exposure notification app distribution. It reduces cost and increases effectiveness of contact-tracing.
- Addressing COVID-19 requires leadership, from our executive branch, legislative branch, state government and county government. Without the leadership, we will not get a handle on this virus. I believe my governor, Gov. Laura Kelly of Kansas, is making the hard decisions necessary to reduce the spread.