

# CDC Tracked Millions of Phones to See If Americans Followed COVID Lockdown Orders

08.12.21

CDC

The Centers for Disease Control and Prevention (CDC) bought access to location data harvested from tens of millions of phones in the United States to perform analysis of compliance with curfews, track patterns of people visiting K-12 schools, and specifically monitor the effectiveness of policy in the Navajo Nation, according to CDC documents obtained by Motherboard. The documents also show that although the CDC used COVID-19 as a reason to buy access to the data more quickly, it intended to use it for more general CDC purposes.

Location data is information on a device's location sourced from the phone, which can then show where a person lives, works, and where they went. The sort of data the CDC bought was aggregated—meaning it was designed to follow trends that emerge from the movements of groups of people—but researchers have repeatedly raised concerns with how location data can be deanonymized and used to track specific people.

The documents reveal the expansive plan the CDC had last year to use location data from a highly controversial data broker. SafeGraph, the company the CDC paid \$420,000 for access to one year of data to, includes Peter Thiel and the former head of Saudi intelligence among its investors. Google [banned the company from the Play Store](#) in June.

***Do you work in the location data industry? We'd love to hear from you. Using a non-work phone or computer, you can contact Joseph Cox securely on Signal on +44 20 8133 5190, Wickr on josephcox, OTR chat on [jfcox@jabber.ccc.de](mailto:jfcox@jabber.ccc.de), or email [joseph.cox@vice.com](mailto:joseph.cox@vice.com).***

The CDC used the data for monitoring curfews, with the documents saying that SafeGraph's data "has been critical for ongoing response efforts, such as hourly monitoring of activity in curfew zones or detailed counts of visits to participating pharmacies for vaccine monitoring." The documents date from 2021.

Zach Edwards, a cybersecurity researcher who closely follows the data marketplace, told Motherboard in an online chat after reviewing the documents that "The CDC seems to have purposefully created an open-ended list of use cases, which included monitoring curfews, neighbor to neighbor visits, visits to churches, schools and pharmacies, and also a variety of analysis with this data specifically focused on 'violence.'" (The document doesn't stop at churches; it mentions "places of worship.")

Motherboard obtained the documents through a Freedom of Information Act (FOIA) request with the CDC.

The documents contain a long list of what the CDC describes as 21 different "potential CDC use cases for data." They include:

- "Track patterns of those visiting K-12 schools by the school and compare to 2019; compare with epi metrics [Environmental Performance Index] if possible."
- "Examination of the correlation of mobility patterns data and rise in COVID-19 cases [...] Movement restrictions (Border closures, inter-regional and night curfews) to show compliance."
- "Examination of the effectiveness of public policy on [the] Navajo

Nation."

At the start of the pandemic, cell phone location data was seen as a potentially useful tool. Multiple media organizations, including the *New York Times*, used location data provided by companies in the industry to show [where people were traveling to once lockdowns started to lift](#), or highlight that [poorer communities were unable to shelter in place as much as richer ones](#).

The COVID-19 pandemic as a whole has been a flashpoint in a broader culture war, with conservatives and anti-vaccine groups protesting against government mask and vaccine mandates. They've also expressed a specific paranoia that vaccine passports would be used as a tracking or surveillance tool, framing vaccine refusal as a civil liberties issue. Robert F. Kennedy Jr.'s Children's Health Defense, one of the more influential and monied anti-vaccine groups in the U.S., has promoted fears that digital vaccine certificates could be used to surveil citizens. QAnon promoter Dustin Nemos wrote on Telegram in December that vaccine passports are "a Trojan horse being used to create a completely new type of controlled and surveilled society in which the freedom we enjoy today will be a distant memory."

Against that inflamed backdrop, the use of cell phone location data for such a wide variety of tracking measures, even if effective for becoming better informed on the pandemic's spread or for informing policy, is likely to be controversial. It's also likely to give anti-vaccine groups a real-world data point on which to pin their darkest warnings.



A screenshot of the use cases proposed by the CDC. Image: Motherboard.

The procurement documents say that "This is an URGENT COVID-19 PR [procurement request]," and asks for the purchase to be expediated.

But some of the use cases are not explicitly linked to the COVID-19 pandemic. One reads "Research points of interest for physical activity and chronic disease prevention such as visits to parks, gyms, or weight management businesses."

Another section of the document elaborates on the location data's use for non-COVID-19 related programs.

"CDC also plans to use mobility data and services acquired through this acquisition to support non-COVID-19 programmatic areas and public health priorities across the agency, including but not limited to travel to parks and greenspaces, physical activity and mode of travel, and population migration before, during, and after natural disasters," it reads. "The mobility data obtained under this contract will be available for CDC agency-wide use and will support numerous CDC priorities."

The CDC did not respond to multiple emails requesting comment on which use cases it did deploy SafeGraph data for.

SafeGraph is part of the ballooning location industry, and SafeGraph has [previously shared datasets containing 18 million cell phones from the United States](#). The documents say this acquisition is for data that is geographically representative, "i.e., derived from at least 20 million active cellphone users per day across the United States."

Generally, companies in this industry ask, or pay, app developers to include location data gathering code in their apps. The location data then funnels up to companies who may resell the raw location data outright or package it into products.

SafeGraph sells both. On the developed product side SafeGraph has several different products. "Places" concerns points of interest (POIs) such as

where particular stores or buildings are located. "Patterns" is based on mobile phone location data that can show for how long people visit a location, and "Where they came from" and "Where else they go," [according to SafeGraph's website](#). More recently SafeGraph has started offering aggregated transaction data, showing how much consumers typically spend at specific locations, under the "Spend" product. SafeGraph sells its products to a wide range of industries, such as real estate, insurance, and advertising. These products include aggregated data on movements and spends, rather than the location of specific devices. Motherboard previously bought a set of SafeGraph location data for \$200. The data was aggregated, meaning it was not supposed to pinpoint the movements of specific devices and hence people, but at the time, Edwards said "In my opinion the SafeGraph data is way beyond any safe thresholds [around anonymity]." Edwards pointed to a search result in SafeGraph's data portal that displayed data related to a specific doctor's office, showing how finely tuned the company's data can be. Theoretically, an attacker could use that data to then attempt to unmask the specific users, [something which researchers have repeatedly demonstrated is possible](#).

In January 2019, the Illinois Department of Transportation bought such data from SafeGraph that related to over five million phones, [activist organization the Electronic Frontier Foundation \(EFF\) previously found](#).

The CDC documents show that the agency bought access to SafeGraph's "U.S. Core Place Data," "Weekly Patterns Data," and "Neighborhood Patterns Data. That last product includes information such as home dwelling time, and is aggregated by state and census block.

"SafeGraph offers visitor data at the Census Block Group level that allows for extremely accurate insights related to age, gender, race, citizenship status, income, and more," one of the CDC documents reads.

Both SafeGraph and the CDC have previously touched on their partnership, but not in the detail that is revealed in the documents. The CDC [published a study in September 2020](#) which looked at whether people around the country were following stay-at-home orders, which appeared to use SafeGraph data.

[SafeGraph wrote in a blog post in April 2020](#) that “To play our part in the fight against the COVID-19 health crisis—and its devastating impact on the global economy—we decided to expand our program further, making our foot traffic data free for non-profit organizations and government agencies at the local, state, and federal level.” Multiple location data companies touted their data as a potential mitigation to the pandemic during its peak in the United States, and provided data to government and media organizations.

A year later, the CDC purchased access to the data because SafeGraph no longer wanted to provide it for free, according to the documents. The Data Use Agreement for the in-kind provided data was set to expire on March 31, 2021, the documents add. The data was still important to access as the U.S. opened up, the CDC argued in the documents.

“CDC has interest in continued access to this mobility data as the country opens back up. This data is used by several teams/groups in the response and have been resulting in deeper insights into the pandemic as it pertains to human behavior,” one section reads.

Researchers at the EFF separately obtained documents concerning the CDC’s purchase of similar location data products from a company called Cubeiq as well as the SafeGraph documents. The EFF shared those documents with Motherboard. They showed that the CDC also asked to speed up the purchase of Cubeiq’s data because of COVID-19, and intended to use it for non-COVID-19 purposes. The documents also listed the same

potential use-cases for Cubeiq's data as in the SafeGraph documents.

Google [banned SafeGraph from its Google Play Store in June](#). This meant that any app developers using SafeGraph's code had to remove it from their apps, or face having their app removed from the store. It is not entirely clear how effective this ban has been: SafeGraph has [previously said it obtains location data via Veraset](#), a spin-off company which interfaces with the app developers.

SafeGraph did not respond to multiple requests for comment.

***Subscribe to our cybersecurity podcast, [CYBER](#). Subscribe to [our new Twitch channel](#).***