

[bloomberg.com](https://www.bloomberg.com)

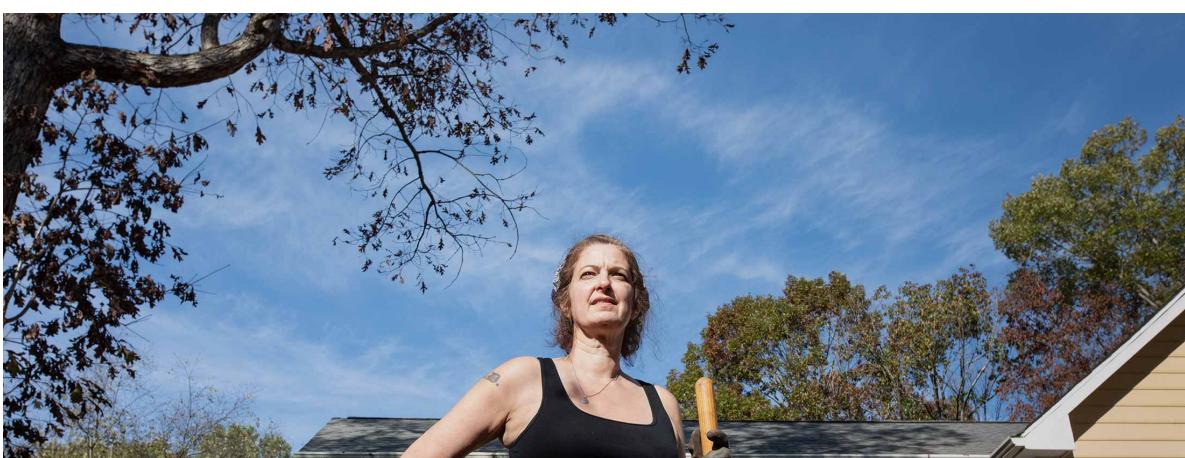
Data Heroes of Covid Tracking Project Are Still Filling U.S. Government Void

Drew Armstrong

13-17 minutes

At the start of 2020, Amanda French was in between academic jobs. Her mother had died about a year earlier, and she'd taken time off to help settle her affairs. Then the pandemic hit, interrupting her employment search, and she was alone outside Raleigh, N.C., with little to do but doomscroll through Twitter, as she described it.

Someone she followed online put out a call for volunteers to assist with a new project tallying how many Covid tests were being run across the U.S.—something the public wasn't getting a straight answer on from the federal government. "I had nothing to do. I was at home alone, anyways," French says. "It was much healthier than reading all the scary news." So on March 18, she signed up.





Amanda French

Photographer: Phyllis B. Dooney for Bloomberg Businessweek

Since then, the [Covid Tracking Project](#)—run by a small army of data-gatherers, most of them volunteers—has become perhaps the most trusted source on how the pandemic is [unfolding in the U.S.](#) The website has been referenced by epidemiologists and other scientists, news organizations, state health officials, the White House Coronavirus Task Force, and the Biden transition team. There are other reliable sources for pandemic statistics, but the project stands out for its blend of rich, almost real-time data presented in a comprehensible way. “I think they’ve done extraordinary work and have met an important need,” says Jennifer Nuzzo, a senior scholar at the Johns Hopkins Center for Health Security, which publishes its own set of pandemic data (and draws some information from the Covid Tracking Project). “They’re tracking things that aren’t being tracked.”

This critical repository of health information started, improbably, with three journalists, a data scientist/biotech investor, and a couple of spreadsheets. Back in late February, the coronavirus was still a sleeper threat in the U.S., with new cases popping up in ones and twos around the country and signs of hidden spread on the West Coast. Officials in the Trump administration held briefings touting the government’s rapid rollout of testing. But they couldn’t

answer one important question: How many tests were being done?

[Alexis Madrigal](#), a technology writer at the *Atlantic*, and [Robinson Meyer](#), an environmental writer at the magazine, decided to call every state and find out how many tests had been performed, plugging the numbers into a spreadsheet. While federal officials were talking about having distributed millions of tests, the two journalists [reported](#) on March 6 that fewer than 2,000 people in the U.S. had been checked for Covid.

Soon after the story published, Madrigal heard from an old friend, Jeff Hammerbacher. He and Madrigal had attended Harvard together. Madrigal got an English degree and became a journalist; Hammerbacher studied mathematics and went on to start the data team at Facebook.

Unbeknownst to Madrigal, Hammerbacher—who now helps found biotechnology companies but previously worked in medicine, applying data-science techniques to research—had his own sheet. He posted it online, kept updating it, and got feedback from readers. “I thought, ‘I guess I’ll keep doing this,’” Hammerbacher recalls.

When Madrigal’s first analysis was published at the *Atlantic*, Hammerbacher emailed him: “Hey, did you guys use my spreadsheet for this?” Madrigal and Meyer’s sheet was full of quotes from health department officials, while Hammerbacher’s was set up to become a proper database. They decided to team up until the data from the government got better. (Co-founder [Erin Kissane](#), an editor and community manager who works in journalism and technology, joined around this time.) They thought the project would last a week or two. “We just sort of figured, of

course the CDC would put out this information,” Madrigal says.

“But it just never happened.”

Search for the Covid Tracking Project on Google Scholar, which compiles academic literature, and you’ll get more than 500 results, a sign of its standing in the scientific community. The project has helped force states to improve their disclosure of Covid data: In April, it started giving states letter grades on the quality of the data they reported. At first only 10 states got an A or A+; now 40 states and territories have reached that grade.

The project is a demonstration of citizen know-how and civic dedication at a time when the country feels like it’s being pulled apart. Yet it’s confounding that, almost a year into the pandemic, the Covid Tracking Project is doing what might be expected of the U.S. government. “It’s kind of mind-boggling that it’s fallen to a group of volunteers to do this,” says Kara Schechtman, one of the project’s early volunteers, who’s since become the paid co-lead for data quality.

For decades the U.S. Centers for Disease Control and Prevention has tracked the flu and other illnesses in the nation. But its systems weren’t designed for real-time surveillance of a new pandemic. Typically, states get information from health-care providers and put it into their systems, and that data is then sent to the CDC, a process that can take several days, according to the agency. “It’s become abundantly clear that our systems of surveillance, both acquiring data and tracking data, are woefully inadequate,” says Nuzzo. (Johns Hopkins’s school of public health has received significant funding from Michael Bloomberg, the majority owner of *Bloomberg Businessweek*’s parent company.)

Racing to Keep Up With Covid Data

U.S. daily cases

Data: Covid Tracking Project, Bloomberg Reporting

In late spring the CDC, which is part of the U.S. Department of Health and Human Services, created a team of more than a dozen people to scrape state health websites overnight and then confirm the information with the states in the morning. This is, essentially, what the Covid Tracking Project started doing in early March. The CDC's information flows into [HHS Protect Public Data Hub](#), a tool launched in April to aggregate different sources of data on tests, cases, and hospitalizations.

If you visit HHS Protect to find Covid data for, say, Virginia, you might click through to the CDC's Coronavirus website, and from there to the CDC Covid Data Tracker homepage. It has state total cases and new cases over the past seven days, but new cases over the past 24 hours are under a different tab. Click on Virginia on the map, and you're redirected to the Virginia Department of Health. By contrast, the Covid Tracking Project has an easily findable Virginia page that tells you new cases today, patients in ICUs, patients on ventilators, and other granular data points. (HHS Protect also has information on hospital capacity, but separate from the CDC pages; in July, the government directed hospitals to send their data to the HHS instead of the CDC, causing concern among some health experts.) A search on Google Scholar for HHS Protect yields only about 30 results, compared with the Covid Tracking Project's 500.

Ryan Panchadsaram, an adviser to the Covid Tracking Project who was deputy chief technology officer in the Obama White

House, says the CDC is well-positioned to be a Covid data hub. “They can ask every hospital in the country for data, and they can ask all the labs and they can aggregate it,” he says.

The CDC does, in fact, gather most of the same data as the Covid Tracking Project, and gets quite a bit more from its own channels. But that hasn’t translated into data that’s fully public, easily usable, or transparent. “The part which is frustrating is that only a few of those datasets are being made public,” Panchadsaram says. “They’ve got all the pipes, they’ve got all of the relationships; they just aren’t following through on the other side.”

Paula Yoon, director of the CDC’s Division of Health Informatics and Surveillance, talks about a shortage of resources and the difficulty of attracting talent to the public sector. There also needs to be a tighter relationship, she says, between public health bodies and the hospitals and doctors—and their records and IT systems—dealing with diseases on the front lines. “Until those two pieces come together and use common data standards, common messaging standards for moving data from one place to another—until we get to that point it’s going to continue to be really hard,” Yoon says.

In April, CDC Director Robert Redfield said that the pandemic would strengthen the agency: “The core capabilities are going to be finally brought to where they need to be.” But the agency hasn’t gotten the resources it needs. Earlier this year, the Trump administration pulled \$700 million meant for the CDC’s vaccine distribution planning and redirected it to [Operation Warp Speed](#), a White House-run vaccine research and development effort. In October, Bloomberg reported that most of a \$1 billion package meant to help the CDC with surveillance, testing, improvements to

data systems, and other measures was stalled for five months.

In addition to 300 active volunteers, the Covid Tracking Project now has 30 paid staffers. The *Atlantic* provides it with legal, communications, and technology support and, crucially, has been the conduit for \$1.42 million in donations from such groups as the Chan Zuckerberg Initiative, the Rockefeller Foundation, and the Robert Wood Johnson Foundation. But the magazine is otherwise hands-off. “I don’t think [the project] could have been built inside the walls of a research institution or a media organization,” says Kissane, the project’s managing editor.

It is in some ways the quintessential pandemic-era startup, existing almost entirely on Slack, the workplace messaging platform. Most of the staff and volunteers have never met in person and know each other only through digital avatars and Zoom chats. They have a wide range of professional backgrounds: French, who is now a paid team leader, has a doctorate in English.

I was recently allowed to sit in on a volunteer training session, watch a data shift, and do an afternoon’s worth of data entry myself. Volunteers start as “checkers” who take a first pass at pulling data from the state websites. “Double-checkers” review it, and more-experienced staff help sort out problems and make judgment calls. Every afternoon the group enters and checks close to 800 data points from 56 states, territories, and the District of Columbia.

On my observation shift in early October, I see how fragmented state reporting systems are. Hawaii’s hospital data, for example, is usually posted first on Twitter or Instagram by Lieutenant Governor Josh Green. But the afternoon is dragging on and Green hasn’t

posted yet—he's been doing a live online chat with Anthony Fauci.

"I think he's too excited about Fauci," someone says on Slack.

Eventually, the day's data is found in his Instagram stories, the disappearing posts on that platform.

"Lt. Gov Green is sneaky with his hospital data. Always got to check the stories!" one shift veteran posts. It's charming, but depressing: Instead of a data feed sent directly to the CDC, the Covid Tracking Project, and other organizations, the state put crucial public health data up on a social media site better known for vacation and food photos. Hawaii has since begun posting the data on its website.

When it's my turn to work a shift the next night, states with data at the ready begin to show up in a queue on the spreadsheet. I spot an error in the Kansas numbers, and one of the shift leads flags my post with a disco ball emoji. "Nice catch!" a person tells me on Slack. I make a mistake with Maine's data, transposing a digit, that's quickly caught.

From interviews with half a dozen staff and volunteers, and after spending time on the project's Slack, I pick up on a sense of disappointment at the decline of technocratic competence in government. I ask French if there are things about the project that depress her or make her angry, given that it can be seen as a fill-in for the government. "What gets me is the failure of the CDC to provide data standards to the states," she says, and then adds a few moments later, "That sounds like a very nerdy thing to be activist about."

There are signs the CDC is making changes. On Nov. 2, the agency posted a request for information seeking a company to

help create a centralized platform for reporting Covid-19 test results. The agency is at work on the Covid Electronic Laboratory Reporting System, which will collect detailed information on each test performed in the U.S. It's already operational in more than 40 states, Yoon says, and gives the agency a closer look at how the virus is spreading: "We've been able to set up, in a very short period, a new system where the electronic laboratory data that is being sent to the states is in turn shared with us at CDC." The agency is also building new automated systems that can pull data directly from states and front-line health providers, according to Yoon.

President-elect Joe Biden has said his administration will create a "Nationwide Pandemic Dashboard" with real-time, local data down to the ZIP code level. The Covid Tracking Project has published a set of recommendations for the transition team. Madrigal says there are "many people in the CTP network who are in contact with the Biden team" but that the project doesn't plan on working directly with the new administration.

On Nov. 17 the project reported a grim statistic: 76,830, the number of people in the U.S. hospitalized with Covid-19, more than at any other point in the pandemic. As cases have surged, states have renewed restrictions on businesses, and officials are urging people to hold Thanksgiving on Zoom. President Trump has been touting [vaccine candidates](#) while Biden has cautioned that "we are still months away" from the end of the pandemic.

Because the country's response to the disease has become so politicized, Panchadsaram says, it's crucial to have a standard set of facts to agree on. "The shared reality we have binds us," he says. "If we can't all agree to what's happening with Covid, we

can't agree what happens next."