2021

# Protecting Privacy in COVID-19 Digital Contact Tracing

Nicolas M. Turner
*Colby College*

# *Protecting Privacy in COVID-19 Digital Contact Tracing*

*Nicolas M. Turner*

*Colby College*
*Honors Thesis*
*Science, Technology, and Society*

*May 19, 2021*

| | |
|---|---|
| _____ | _____ |
| Lijing Jiang, Advisor | Kara Kugelmeyer, Reader |

# *Abstract*

Digital contact tracing applications are a necessary and potentially dangerous tool used to combat the spread of COVID-19. Due to the potentially sensitive nature of personal information gathered by contact tracing applications, there is high potential for privacy issues to arise. This thesis explores the dangers of digital contact tracing or contact tracing via a mobile application. Focusing on the United States, this paper studies how contact tracing works and how a contact tracing application might collect different types of data. The paper then studies the effects of giving up location data and health data and what potential ramifications that may hold. Finally, the thesis concludes with the idea that there needs to be federal regulations on health and personal data beyond that of which is provided by HIPAA, noting that Europe's General Data Protection Regulation can be a good model for the United States.

# *Acknowledgments*

First, I would like to thank Professor Lijing Jiang for the guidance she provided the class and the tireless efforts she sustained throughout the year. Always responsive with good feedback, Professor Jiang was instrumental in making my thesis what it is today.

Second, I would like to thank Professor Kara Kugelmeyer for two semesters of one-on-one work. Through an independent study and many hours of discussion, she helped shape this thesis with her expertise in privacy and information science and her seemingly endless knowledge about other related topics. Professor Kugelmeyer has been improving my writing and guiding me throughout my college career and this paper is the culmination of many of her efforts.

Third, I would like to thank my peers for their support, encouragement and help. I would particularly like to call out Heather Jahrling and Ben Steib for their edits and constant camaraderie throughout this process. Thank you to everyone else who has aided me along the way.

# *Table of Contents*

# *Chapter 1: Introduction*

Disease, a biological disorder of structure or function of a biological system or entity, pre-dates the human species and chances are it will continue to endure long after humanity is gone. Biological life forms (humans, plants, animals, etc.) have long attempted to slow down the spread of disease or thrive through it. Some of these forms have adapted, some have died out, and some simply suffer through disease and persist as they are. Humans are one of the latter, surviving through disease and maintaining their place on the planet. Some diseases have cut deeper than others, some are hardly noticeable, and some cause noticeable marks on human history. The Black Plague killed an estimated 30-50% of Europe's population over the course of eight years (DeWitte, 2014). One could make the claim based upon past and current human experiences that humanity and disease are two sides of the same coin, always connected but never seeing eye to eye.

Humans have tried over the ages and with varying degrees of success, to slow down or use novel ways to cure disease. During the Black Plague, doctors recommended the consumption of Mercury and Arsenic as cures (BBC, 2019). Today, we know that both elements are toxic to the human body, but in a time of desperation in a fight against a disease killing one in two people, people will often resort to drastic measures. In the efforts to understand, cure, or eliminate disease, breakthrough ideas or "cutting edge" ideas have offered key insights. Some of these insights come in the form of social advances, some are technical advances, and many are medical.

Since 1920, one of the trusted ways to trace the spread of an infectious or highly transmissible disease was to break the chain of transmission by finding everyone an infected person has been in contact with, testing them if they were exposed and isolating them if they, too, were infected. Known as contact tracing, this method has been a staple of infectious disease control. Contact tracing has been a known effective method in fighting disease for many years. Along with social distancing, or the practice of keeping people physically separated to make it difficult for disease to spread, contact tracing has remained one of the two best tools for preventing disease. Social distancing will not be discussed in this paper as the paper focuses on contact tracing, but it is important to mention.

Contact tracing was implemented in the United States in response to continuous syphilis outbreaks. Anonymous contact tracing was developed for syphilis patients, "'Upstanding' patients with conditions like syphilis need not be reported by name to state health departments, only by a code" (Fairchild et al., 2020). Codified contact tracing is the modern practice, especially used in systems that will be discussed later in this paper. The stigmatized nature of syphilis forced physicians to keep people protected behind codes, if they were considered worthy of that protection. People who were not 'upstanding', or wealthy with white skin were not extended the same courtesy. Health departments would then communicate with the contacts of those patients to inform them they were potentially infected. In later overwhelming outbreaks, these rules were diminished a little bit in that everyone was reported by name. Collecting names to handle widespread outbreaks became the norm in the 1960s. Contact tracing centers were set up inside of health departments with ample resources (Fairchild et al., 2020).

In the 1980s, during the AIDS epidemic, trust in contact tracing centers began to suffer. Due to the illegal nature of homosexuality in the 1980s, many people refused to disclose personal

information about sexual encounters, including partners. Given the sensitive nature of a person's health information, states and the federal government created and enacted laws that would protect who could access a person's private health information. An example of this in the United States is the Health Insurance Portability and Accountability Act (HIPAA). HIPAA "is a federal law that requires the creation of national standards to protect sensitive patient health information from being disclosed without the patient's consent or knowledge" (CDC, 2018). These laws apply to healthcare providers, health plans, healthcare clearinghouses, and business associates. HIPAA revolutionized personal medical data protections in the United States. Globally, each country has their own stipulations about patient privacy. The range of protection of health-related information is from very strict in favor of the individual, like the European Union, to no privacy laws for the regulation for health systems data like in the case of China (Gong et al., 2020), (HIPAA Journal, 2018).

Coronavirus disease (COVID-19) is an infectious disease caused by a newly discovered coronavirus (World Health Organization, 2021). By March of 2020, COVID-19 had spread across the globe and the World Health Organization (WHO) had declared a world-wide pandemic. In countries that were prepared for a pandemic like Singapore and Hong Kong, contact tracing was able to go in full effect very quickly as they are technocratic societies where people were looking for technical solutions to the public health crisis due to the lack of medical solutions available. Technological solutions that were already deployed to the public were ideal to aid in the fight against COVID-19. Systems like smartphones, Bluetooth, GPS, and constant connectivity are omnipresent, more widespread even than COVID-19. These technologies can be used by people to trace and slow the spread of COVID-19 without inventing entirely new systems and going through the painstaking process required to deploy the systems globally.

The smartphone is a ubiquitous system. An example of the smartphone's adoption and reach is that over 75% of the US population has a smartphone, and that number is constantly growing (Pew Research Center, 2019). Smartphones are always connected because they have GPS, Cellular, WIFI, Bluetooth, and other systems that make them connected to the outside world (IBM, 2019). Smartphones can communicate with other phones locally, sensing each other through Bluetooth proximity and are considered by many the supercomputers in the pockets of everyday people with unparalleled abilities to act as a social technology.

To build on the social technology of contact tracing, the technological connections and prominence of smartphones, and the knowledge of what other devices, and by extension, people, are around them, people created contact tracing applications that run on smartphones. This contact tracing technology collects information on where the user is, who the user is near, and then when the user tests positive, and the applications notify everyone the user has encountered. Singapore first successfully implemented this in mid-March 2020, limited in effectiveness by its adoption rate (Koh, 2020). In March 2020, many tools to slow down the spread of COVID-19 were built and deployed, though with such powerful and ubiquitous tools like smartphones, there are always massive potential downsides.

For smartphone contact tracing applications to be effective, they need access to personal data like location, personal information, and health information. This information is worth billions of dollars, highly protected, and considered of the utmost importance in privacy. The collection of all this personal information opens a pandora's box of issues. Privacy is handled differently in different sectors of application development. For example, personal user data in the private sector is desired because personal user data is worth money, generated by selling more targeted advertisements. Personal data in the government can be desired for intelligence

purposes, or methods to essentially spy on its citizens, among other things. Protecting this data is of the utmost importance. In Qatar, the government can track the real time location of every one of their citizens using their contact tracing application, *Ehteraz* (*Bahrain, Kuwait and Norway contact tracing apps a danger for privacy*, 2020). This is a major violation of privacy on Qatari citizens.

Another hypothetical, yet plausible scenario surrounds the issue of insurance discrimination. Fortunately, the United States has laws protecting against insurance discrimination from preexisting health conditions, but insurance discrimination is a good hypothetical case study to show the power of personal health data. Early estimates suggest 35% of patients report lasting effects from COVID-19 after recovery (Tenforde, 2020). As of December 5, 2020, there have been approximately a total of 15 million cases in the US since the beginning of the pandemic. When you multiply this by 35%, there are roughly 5.25 million people with lasting effects from COVID-19 in December 2020. In the long run, this might lead to other health issues or complications that insurance companies or employers would be liable to pay for. These millions of long term COVID-19 patients could have a difficult time getting insurance or a job, all because a contact-tracing list was created, and the patients were on it. While insurance companies know if a client was in the hospital, most COVID-19 cases do not require hospitalization thus insurance companies would not know if their client had been infected.

This is the information age where data, but more specifically personal data, is gold, and gold is generated by the devices people hold in their hands. Facebook generates four petabytes (One petabyte is equal to a million gigabytes, or 250,000 HD movies) of data each day (Osman, 2020). This data, often which is personal data, is what makes Facebook money because it can use

the personal data to sell targeted advertisements. Collecting constant location data and knowing who is in proximity with who is worth billions of dollars to Facebook and its advertisers. The private sector would love to collect personal data on the general population because it would benefit the private sector business models greatly. Contact tracing applications and their mandates for use are a perfect vector for this information collection to happen.

It is imperative that the technology behind mobile contact tracing is thoroughly studied. New updates are released regularly and there are tens of contact tracing applications on the market. Discerning between safe and unsafe applications can be nigh on impossible. However, differences can be found by studying the background of the applications, for example who made the apps, why they made them, and the technologies the applications apply. If a government stores everything in a centralized database, holding onto data for more than thirty days, and requires GPS data, then they are likely to be abusing their power for unnecessary purposes. However, if companies like Google and Apple provide a transparent and published framework for anonymized Bluetooth tokenization, or a way to trade Bluetooth signals with another phone anonymously, like they did with their framework *Exposure Notifications*, then that is likely to be a valid and trustworthy service.

Understanding the concerns around the loss of personal data privacy is important too. Americans can build massive, personal data intensive, technological systems but there is no telling who is to use them. America greatly values privacy and independence, which are one in the same. Through Alan Westin, the father of American privacy law, and others, we can apply American privacy culture to digital COVID-19 contact tracing. Understanding privacy in America is paramount to understanding the way that contact tracing is conducted. Studying how the American private sector views privacy and privacy law is key as well, as the law rarely keeps

up with the forefront of American technology, and technology around contact tracing in its current form is brand new. America's data privacy is in danger of being compromised but through study and regulation can it be protected.

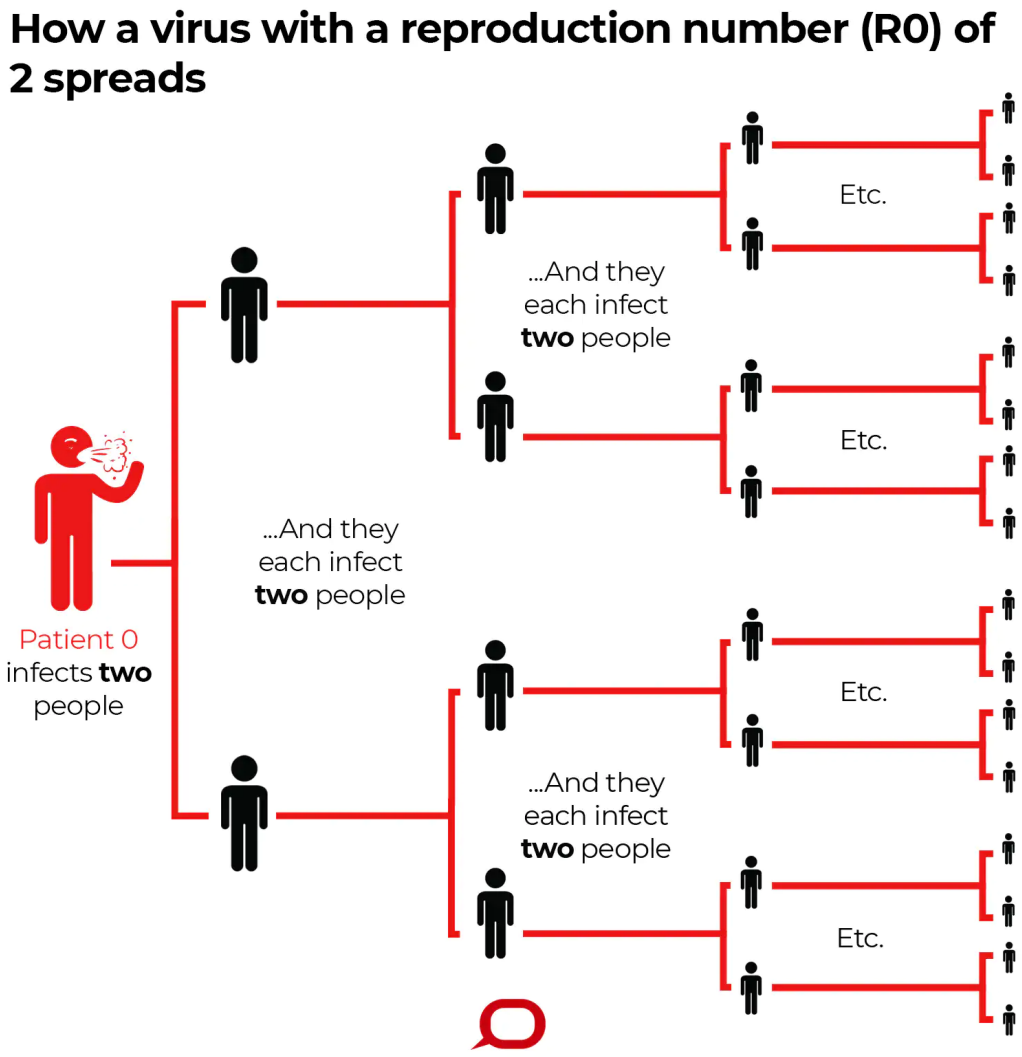
The essential questions that this paper will explore is, "With the advent of mobile application-based contact tracing, how is the online privacy of society at risk? How can those risks be mitigated when tracing is conducted by the private and public sector?" The thesis of this paper is as follows, "The private sector and public sector provide two different sets of risks, but by being diligent and implementing best-practices in regard to cybersecurity, creating and enforcing data protection regulations, being transparent in application implementation and privacy, and adopting new contact tracing applications carefully, we can reduce the dangers of losing privacy to contact tracing applications."

The structure for this paper will begin with a discussion of privacy and the history of privacy in America. Then a discussion of the history and current state of contact tracing will occur. The next chapter will be a discussion of why location data is important to keep private. This will center around phone-based location data. The fifth chapter will cover health data, why that is important and a light overview of the laws protecting it now. The sixth chapter will nestle this paper into the Science, Technology, and Society area of academic study and give detailed outlines for the next steps forward in the future as well as an overarching summary of the rest of the paper.

# *Chapter 2: What is Contact Tracing?*

The need to have reliable contact tracing is essential for tackling COVID-19. By tracing people who are connected to people who test positive with COVID-19, potentially sick people can be quarantined before infecting anybody else. Contact tracing, when done properly, can vastly inhibit the spread of disease, and with a virulent illness like COVID-19, this is extremely important. COVID-19 with regular social distancing has an $R_0$ value (the average number of people that COVID-19 will be transmitted to per case), pronounced R-naught, of 1.2, or a value that allows it to spread. With proper, and rapid contact tracing, followed by quarantine that $R_0$ value drops to 0.8, or a disease that is dying (Kretzschmar et al., 2020).

While $R_0$ value encompasses only part of how a disease spreads, it can be used as a basic metric to reiterate how contact tracing is effective. $R_0$ value is defined by the reproductive rate of a disease, so an $R_0$ value in March 2020 in the US was between two and five. This means that every person with COVID-19 on average spread it to two to five people. When social distancing was implemented, $R_0$ moved closer to one in the US (Adam, 2020). While the $R_0$ value is above one, it means that a disease is still spreading because more than one person is being infected on average. For example, Joey gets COVID-19, and spreads it to Lamar and Kristina, who each spread it to two people, so with an $R_0$ value of two, the disease is spreading exponentially. With an $R_0$ value of 0.8, if five people have COVID-19, then only four of them spread and they only spread it to one person each. So, after one iteration of the disease, only 80% of the people have COVID-19. This is how COVID-19 grows and falls exponentially.




(Eisenberg, 2020) The chart above is representative of how COVID-19 travels through society based on the scientific phenomenon of $R_0$ value. The image uses an $R_0$ value of two, or an exponential spread rate.

To flatten the curve, contact tracing must be done in an efficient and timely manner. A lag of more than three days between when a person tests positive and the notification to their close contacts does not lower the $R_0$ value below one. This means that from the moment a person tests positive, contact tracing must begin. To do contact tracing efficiently, a massive logistics

and staffing operation is required. In the US, there were nearly 200,000 new cases every day in November 2020 (CDC, 2020), overwhelming contact tracing systems. A "close contact is defined by the CDC as someone who was within 2 meters of an infected person for at least 15 minutes within a 24-hour period starting from 2 days before illness onset (or, for asymptomatic cases 2 days prior to positive specimen collection) until the time the patient is isolated" (CDC, 2020). Imagine that each person infected has five close contacts. There would be a million people that contact tracers in the US would need to call every day. Imagine each contact tracing call takes ten minutes, and tracers are working ten hours each day, the US would need 17,000 contact tracers to contact all contacts. It is not hard to see how some contact tracing centers were overwhelmed. This was well before the peak of cases as well, which was seen at over 315,000 new cases in a day in January 2020 (CDC, 2021).

Creating a purely digital contract tracing solution to help reach many of these people would massively help the fight against COVID-19. Instead of forcing contact tracers to call each person individually, a smartphone could report directly a user's close contacts and alleviate the stress on the overtaxed contact tracing system. There are many methods of doing digital contact tracing apps, which this thesis has already touched on, but each will be explored in more depth. Location based tracking uses GPS locations and compares them to other people. Bluetooth systems use the strength of a Bluetooth signal between multiple phones to see how close they are. Both methods are complicated and difficult processes (Dehaye, 2020) with varying results in accuracy.

Location based tracking is both ineffective and overly invasive. GPS, or the Global Positioning System, uses satellites to triangulate a position on the planet. The user's phone contacts the satellites to figure out where the phone is. GPS trackers on the best phones are only

accurate to within about two meters (important Safety Technologies, 2020). When a "close contact" is defined by someone within six feet of a user for more than fifteen minutes, this is insufficient accuracy. It is difficult to determine when someone is a close contact if applications are unable to define where a person is in proximity to the close contact.

Thus, location-based tracking is more useful for understanding where people are, rather than who is nearby for medical purposes. However, location tracking can also result in government surveillance. As mentioned before, Qatar can track everyone on their contact tracing app, *Ehteraz,* in real-time meaning the government can search a specific person and see their whereabouts. If this occurred in the United States, the backlash would be bigger than the results of the Snowden leaks. In the 1980s people did not trust the government to not compile a list of gay men, but now the government could potentially compile lists of all citizens with COVID-19. While a list of COVID-19 positive people appears to be harmless, a nefarious actor could hack these databases and collect information on millions of people. In a state where human rights can be negated by the Emir, this is dangerous for the people. Consider a situation where a person speaks out against the Emir's handling of COVID-19. The Emir of Qatar may order for this person to be silenced. To track this person, the government simply identifies the person's location on the *Ehteraz* database. While *Ehteraz* was built to save lives, it could just as easily be used as a tool to end them.

In the United States in May 2021, there were twenty-four states with contact tracing apps of some sort (Sato, 2020). All those states use only an *Exposure Notification* based system from Apple and Google except for Wyoming, North Dakota, Rhode Island, and South Dakota. Wyoming and North Dakota use both *Exposure Notification* and a system called *Care19 Diary* which uses GPS data. South Dakota only uses *Care19 Diary* while Rhode Island has *Crush*

*Covid RI* (Sato, 2020). It is curious that Wyoming and North Dakota both use multiple applications to track COVID-19. This seems to be unnecessary, suggesting that *Care19 Diary* is used for GPS tracking not for contact tracing purposes. The redundancy, especially using a GPS system in addition to the status quo *Exposure Notifications* system, is unneeded and potentially dangerous. *Care19 Diary* says it uses location data "to help in contact tracing and forecasting the pandemic's progression with accurate, real-time data" (North Dakota State Government, 2020). However, this does not mean that the location data is limited to just the government. *Care19 Diary* was caught selling user information to other partners (Fowler, 2020). This is exactly the problem with companies holding access to dangerous personally identifiable information.

Contact tracing using GPS is less focused on identifying the people the user is around, but more focused on the locations the user travels through. For example, a user spends time in McDonalds. At the same time, another user travels into the same McDonalds and spends an overlapping twenty minutes. The GPS app would identify the location of both these users and assume they might need to be contact traced. Yet, inaccuracies for GPS inside buildings make it nearly impossible to identify if the two users sat near each other or were across the establishment. Obtaining a two meter accuracy from GPS is difficult but defining a two meter distance indoors is impossible. This is where Bluetooth accuracy has an advantage.

Consider the same scenario, two users separately enter a McDonalds establishment. They do not know each other, but they are both using the same Bluetooth contact tracing application. Without prompting from the users, the phones connect and start measuring Bluetooth signal strength and passing information to each other. Both users get their food and sit down at tables near each other. Their phones recognize an increased signal strength and measure it to be less than two meters in distance. After fifteen minutes of this proximity, the phones pass anonymous

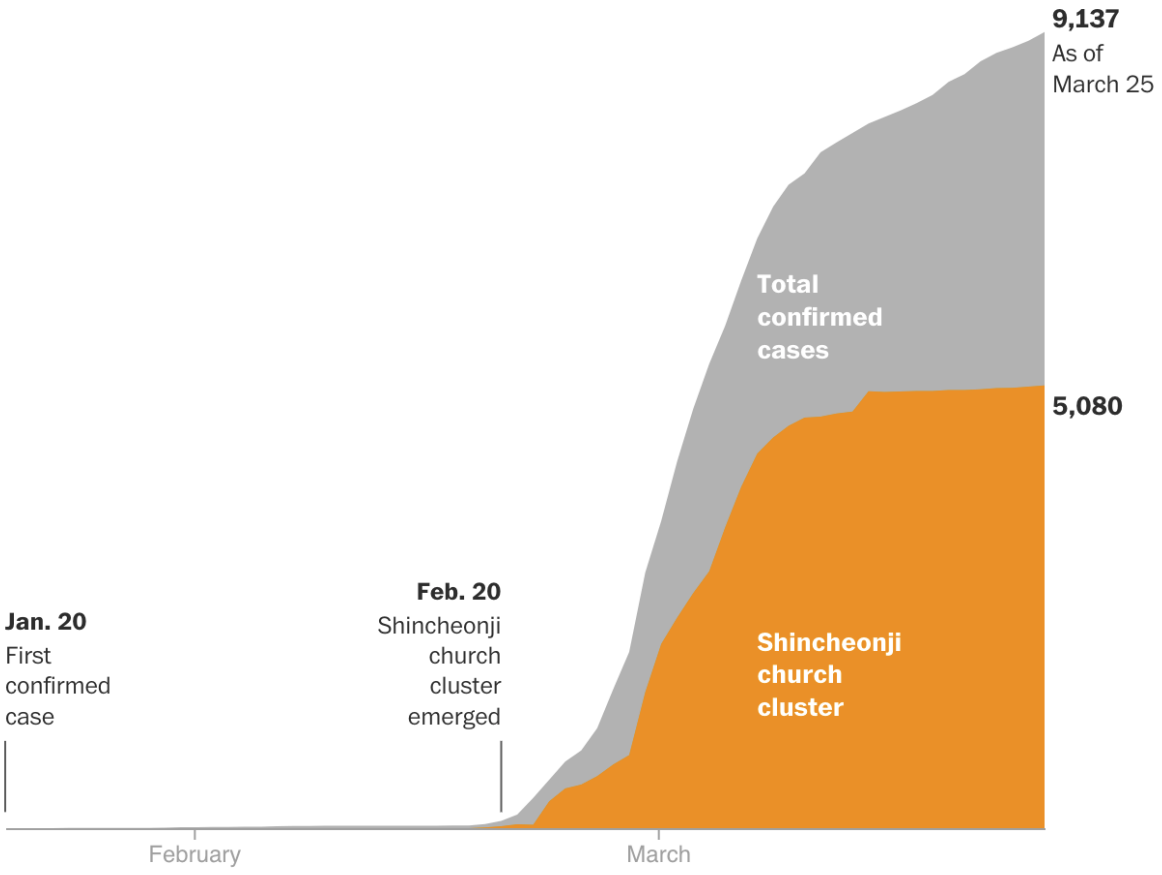
information, and the real time contact tracing process is complete. If one of those users tests positive in the next few days, the other user will be notified they may have been exposed to COVID-19.

GPS tracking does not work nearly as well for contact tracing as a Bluetooth method. Inside the McDonalds, the GPS system would be blind and may contact trace people who should not have been, a false positive, or fail to contact trace a user who should have been, a false negative. False negatives can be deadly. A false positive might be an inconvenience, but a false negative where someone is infected and failed to be notified might create an instance of community spread. Furthermore, a GPS system is invasive and altogether unnecessary.

The speed in which people can be notified is a major benefit to digital contact tracing. There are not enough human contact tracers to do all the work necessary to contact trace every case. Using applications such as *Exposure Notifications* can be a major aid in contact tracing. Consider the previous scenario in McDonalds. The person who will proceed to later test positive sits in the center of the room near the line. Assume that person is there for an hour. In that time, twenty people come in, wait for their food and spend time within two meters of the future infectious case. Furthermore, assume all twenty of those people also use the same contact tracing application that the positive case does. Instead of people spending hours calling each of these people, or none of them ever receiving a notification whatsoever, they are notified immediately that they may have been exposed to COVID-19. In this scenario, there are now up to twenty less cases, which could turn into a massive community spread event. A similar real-life case like this occurred in South Korea early in the COVID-19 pandemic.

On February 18, 2020, a lady in the Shincheonji Church in South Korea tested positive for COVID-19. Designated as “Patient 31,” she was the beginning of one of the world's largest

outbreaks tied back to a single person. Within two days, fifteen more church congregants were testing positive. One month later, 5,080 cases were tied back to Patient 31 and the church she was a part of (Shin et al., 2020). In this real-world scenario, there were no contact tracing applications available to digitally trace people. Imagine if each person in the church was contact traced based on location and proximity to others. If Patient 31 infected fifteen people right away, those fifteen people may have been contact traced digitally before they could infect any others as soon as Patient 31 tested positive. Contact tracing Patient 31 could potentially have prevented thousands of infections. Instead, human tracers were responsible for the situation and every minute they were not in contact with people, COVID-19 was spreading.



(Shin et al., 2020) In the chart above, within a month, Patient 31 was responsible for more than half of South Korea's cases. Digitally contact tracing Patient 31 before they attended church or contact tracing her close contacts could have saved many lives.

It is understandable how contact tracing can lower the $R_0$ value of a virus. Before Patient 31, spread was localized to family members and house mates, in other words there was no community spread and the pandemic in South Korea was under control. Patient 31 changed the state of the pandemic for South Korea, creating an $R_0$ value far higher than one. Contact tracing is paramount to handling a pandemic properly, and with the tools available today, it is necessary to use a digital contact tracing system to keep cases to a minimum. Contact tracing is an invaluable tool in the hands of authority in preventing community spread and lowering rates of spread to manageable levels. Without contact tracing, COVID-19 will spread through society, even if people are socially distant and using other mitigation strategies. Contact tracers are overwhelmed by larger numbers of people who are ill but can be aided by a digital system. The importance of a Bluetooth-based digital contact tracing system cannot be overstated, it is the difference between people who live and die from COVID-19.

# *Chapter 3: Privacy in America*

America was built on the desire for independence. Independence takes many forms but controlling one's own personal information is central to remaining independent. Personal information is a part of oneself, and by that standard, constitutes a choice. Alan Westin was a legal scholar who defined the modern right to privacy at the beginning of the computer age, almost prophetically foreseeing the reach that major companies would have. His key belief was that privacy is to be controlled by the individual. He believed, "Consumers were entitled to withhold such data…but were equally entitled, if they wished, to have it used to alert them to products and services targeted to their interests" (Fox, 2013). Westin believed that privacy was a choice, and that choice was for the consumer to make. This belief contradicts the way many of America's policies work today. Additionally, American privacy law and culture are deeply pertinent to the way the digital contact tracing is conducted.

Contact tracers must collect large amounts of information. To be effective, they must know the user's phone number and everyone around the user for the last fourteen days and their phone numbers. According to the US Department of Labor, phone numbers are personally identifiable information (*Guidance on the Protection of Personal Identifiable Information | U.S. Department of Labor*, 2019) as well. Some contact tracers may ask about locations, family members, and other contact information. Arbitrarily assuming the contact traced is in contact with thirty different people in fourteen days, that is thirty-one phone numbers that are collected. Suddenly, a profile has been constructed filled to the brim with information.

Companies like Facebook already collect this information. A contact tracing phone app knows the user tested positive for COVID-19. Eventually, the private company can create a list of users who tested positive for COVID-19. This list could be compiled and sold to other private

corporations for vast sums of money for reasons discussed later in this thesis, and users would not have a choice. They would have agreed to some clause buried in the terms and conditions that no one reads.

Data and information related privacy law, practices, and adoption of these practices in America is highly sectorial and seems to always be a step behind the technology at hand. While the Snowden Leaks showed us that "the NSA sifts through vast amounts of Americans' email and text communications going in and out of the country," Alan Westin condoned the use of wiretapping when national security was at stake. The Edward Snowden leaks were data dumps of vast amounts of documents from inside the NSA. The Snowden Leaks also showed us that "NSA analysts revealed to have sometimes spied on love interests, with the practice common enough to have coined the term LOVEINT, or love intercepts" (Szoldra, 2016). This breach of privacy rocked America, particularly the LOVEINT revelation. These leaks were arguably a turning point in modern American privacy culture. Even though this information was collected generally under the Patriot Act, it's a stunning example of the lack of oversight of personal data collection in America and the lack of laws to govern it. Not only does the government prolifically collect data on users, but companies do as well.

Facebook appears to be the largest and most pervasive company that collects the most personal information on users, especially in the social media world (Atamaniuk, 2020). Facebook users build profiles full of information including birthdays, things they like, and create a network of friends. They are encouraged to tag themselves and friends in images allowing Facebook to use a user to collect data on other users as well. That's not all by any means, but every click and message on Facebook is collected (Facebook, n.d.). Westin argued that consumers need their own choice about their data, but this is not the case. In the 2018 Cambridge

Analytica Scandal, Cambridge Analytica bought Facebook data on tens of millions of unapprised Americans to build a "psychological warfare tool," which was used to help sway US voters towards electing Donald Trump, president (Lapowsky, 2019). The data breach was vast, "Cambridge Analytica harvested data on 50 million US Facebook users, a number far larger than the 270,000 accounts Facebook initially cited. Facebook says it knew about the breach but had received legally binding guarantees from the company that all of the data was deleted" (Lapowsky, 2018). These consumers were offered no choice in this instance of data privacy. Their privacy, and therefore their independence was gone, all sold for the profit of a massive company.

It is instances like the Snowden leaks and the Cambridge Analytica scandal that the internal workings of data in America can be seen. Privacy laws do not encompass the extent of the data that is collected and are always falling behind the latest technology. This provides a terrifying blueprint for contact tracing applications. Apps that can potentially collect health data, location data, and personal information are worth billions of dollars, and selling that data would be catastrophic for American privacy. Alan Westin predicted this, and America can solve this problem. For starters, it can model the European General Data Protection Regulation (GDPR) internet privacy restrictions.

In GDPR, personal data is defined as any personally identifiable information, which is information that can be traced back to a person such as a phone number with a name attached to it, or a social security number. Sensitive personal data is one step further, defined as any genetic, biometric, and health data. The collection of personal data must be transparent, and explanations of the collection must "be provided in a concise, transparent, intelligible and easily accessible form, using clear and plain language" (Bohan & Bollard, 2020). The data that is collected must

be kept for “no longer than is necessary for the purposes for which the personal data are processed.” In the context of contact tracing, this means that the data collected must be listed in a clear manner and deleted after about fourteen days. Furthermore, even if a data subject (user) gives permission to data usage, it can be withdrawn at any time. Any data collection scheme is required to have a data protection officer to which questions of data protection can be addressed. Data is also required to be protected by “appropriate security measures” regarding cybersecurity, though these appropriate measures are not defined (Bohan & Bollard, 2020).

The US can undertake these steps at a federal level. The California Consumer Privacy Act, or CCPA, entitles consumers to unprecedented levels of data protection in the US. The Office of the Attorney General of California says CCPA allows for the following:

> “1. The right to know about the personal information a business collects about them and how it is used and shared;
>
> 2. The right to delete personal information collected from them (with some exceptions);
>
> 3. The right to opt-out of the sale of their personal information; and
>
> 4. The right to non-discrimination for exercising their CCPA rights.” (Becerra, 2018).

This means that people can control their information collected on them in California, giving people unprecedented control over their personal information. In a move that follows the European Union’s General Data Protection Regulation, California now leads the way in the US for personal data protection and control.

While the CCPA is effective, it is still having growing pains. Any company that has a revenue of more than $25 million each year or collects personal information on more than 50,000 people is required to follow CCPA in California (Besinger, 2020). However, the interpretations of CCPA have been broad. Some companies deliver far too little information about what

personal data is collected, and some companies bury the important information about key personal data collection under a mountain of other personal data. The result of this is a lack of clarity, however, over time these issues will be ironed out.

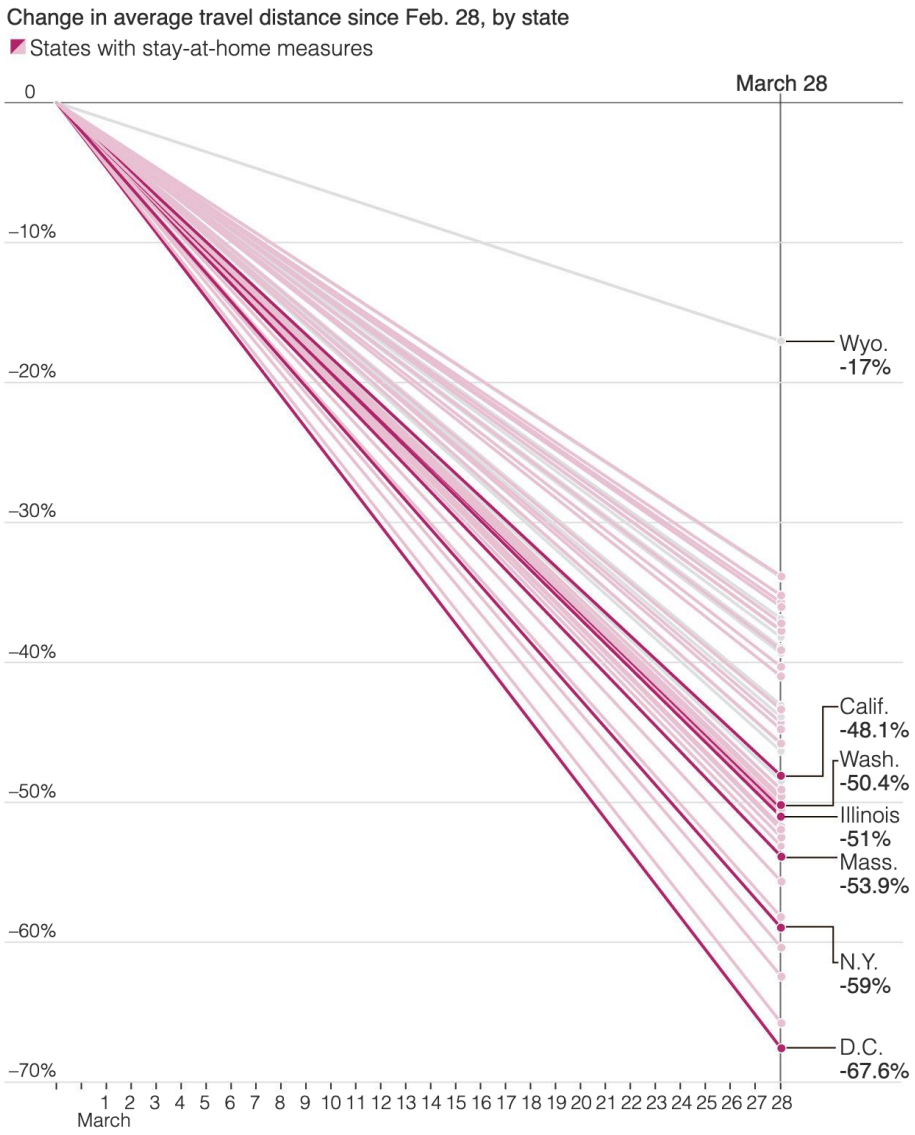
The beauty of CCPA is that it is the first true privacy regulation of its kind in America, and it can act as a guideline for other states and the rest of the country. With a federal version of CCPA, all Americans would be protected by law. However, the business-friendly elements of the federal government would be hard-pressed to approve something this stringent and far-reaching. Companies lobbied heavily against CCPA (Besinger, 2020) because of the threat it represented for their ability to make money and stay ahead of competitors. Data collection is the "secret sauce" to a viable and competitive business model in today's online world. Without something like the CCPA for the US, there will never be enough personal data protection for Americans using technology. CCPA may require contact tracing apps to follow its guidelines in California, but a user in Alabama has no such protections. It is time to extend personal data protection to everybody in the US.

# *Chapter 4: Why Location Data Privacy Matters*

The world before digitization of information maps were about paper, memorization, and handwritten directions. Born around 2004, Google Maps revolutionized the way that people get around (Gibbs, 2016). Before 2004, paper maps were far more prevalent, keeping physical locations of people generally a secret from the private sector. People could go to a specific grocery store, hairdresser, or shoe shop and not receive related advertisements later on the internet. Location data was largely nonexistent before the advent of smartphone GPS. With smartphones, every piece of location information is saved when the user allows it, which is nearly always because apps do not work properly without it.

Location information is the connection between the virtual world and the physical. Often, people click and make purchases on the web, but it is never something physical until the purchase arrives. They are at a web portal, never moving from that location. Companies only see the online information, not real-world information on people. In effect, the profiles that companies built never moved, they were stationary. Companies could never figure out where people ventured, just the clicks they made online. With GPS, that all changed. People were passively tracked, their locations known constantly, each spot measured for time they could be there.

This location data has enabled people to track the amount of movement that people did in the aftermath of state lockdowns due to COVID-19. By tracking the distance people traveled, Unicast was able to create charts looking at the drawdowns in travel as effects of lockdowns and stay-at-home orders between February 28 and March 28.




(Rust et al., 2020) Above is a chart defining the change in travel by people between March 1, 2020, and March 28, 2020. Some places like Washington D.C., reduced travel by 67.6% while people in Wyoming only reduced travel by 17%. This chart highlights the power of location data tracking.

People voluntarily gave up this travel information, but even then, the power of being able to measure how 25 million people moved is extraordinary. Couple that personal location information with IP addresses and companies can connect location with other user searches. Suddenly, companies know how far people are willing to travel to what stores, quantifying how

much people want to buy goods from those stores. The information that companies have is infinitely more valuable than before.

Private companies collect this consumer data to sell it. The ad targeting model is built around having better information on the users, therefore delivering more targeted ads, making the targeted user buy more of a certain advertised product. If a user wanted to use Target instead of Walmart, they might drive to a Target three times over the course of a month, and Walmart only once. When people were using paper maps, user's location information was privy only to the user. Nowadays, GPS aggregators will see that personal location data. Companies then can connect that information with an online profile, so when a user surfs the web, companies know the user likes Target better than Walmart, showing more ads for Target items. This means that the user is more likely to click on Target ads than a Walmart advertisement.

Target will pay more to put advertisements for their products in front of people who will click on those ads. The result of this advertising scheme is that Target will gain revenue from more purchases and advertisers will make more money off Target ads simply through better, more personalized advertising. This may not seem like a large problem, but according to the Pew Research Center, 90% of Americans used their cellular devices for location-based tools in America (Pew Research Center, 2016). If 75% or more of the US population has a cellphone (Electronic Privacy Information Center, 2010), there were roughly 218 million Americans whose location data was being collected.

COVID-19 apps that require the use of location are contributing to this problem. Rhode Island, South Dakota, Wyoming, and North Dakota all require the use of GPS tracking on their COVID-19 applications. This puts them in a constitutionally suspect position. The Fourth Amendment says it is the "right of the people to be secure in their persons, houses, papers and

effects, against unreasonable searches and seizures, shall not be violated" (Legal Information Institute, 2017). The use of GPS by the US Government has created several court cases. The current ruling on GPS tracking and violations of the Fourth Amendment are defined in *United States v. Jones,* 132 S. Ct. 945 (2012). *United States v. Jones* upholds a definition set in *Katz v. United States,* 389 U.S. 347 (1967), such that the US Government installing a GPS device on a target vehicle violates the definition of a "search" because it trespasses on private property (Electronic Privacy Information Center, 2010). A cellular device is also private property, so applications that track GPS run by the government, especially without given consent could violate the Fourth Amendment.

Consider the situation where law enforcement wants to find a suspect. In the era before prevalent GPS, it was up to them to find them via standard police work. Now, law enforcement agencies are buying GPS data on people without warrants (Morrison, 2020). As governments provide contact tracing apps for COVID-19, this violation of constitutional ambiguity is amplified. The rights of American citizens may be negated, all because of a COVID-19 contact tracing application that asks for too much access to connective mobile technologies. Location data is of the utmost importance to keep private if people are to retain their independence.

At minimum, location sharing must be an opt-in scenario. In contact tracing apps that are built around location sharing, access to location should be opt-in and both Android and iOS operating systems force that. However, some applications can bypass those restrictions. Facebook has been caught tracking location data even if it was turned off, though Facebook's data history is abysmal. *Wired* writes, "Using Facebook comes with a cost, even if it's not paid up front in dollars and cents" (Nield, 2020). This is true for all for-profit, free companies. The

companies collect personal data, then turn around to sell that data, and sometimes the buyer is the government, infringing on the rights of its private citizens, even in the United States.

# *Chapter 5: Why Health Data Privacy Matters*

A person's health can be considered one of their most cherished possessions. To have one's health allows a person to truly live their lives to fullest extent. The value of human health cannot be understated, on a personal level and a commercial one. The healthcare industry was worth nearly 18% of the US GDP in 2019 (Stasha, 2020). It is a massive and lucrative business. The US on average spends a little bit over $10,000 dollars a year per person on healthcare, double what most countries spend, yet there is no universal healthcare system as found in many developed countries around the globe. The US healthcare system is highly commercialized, and the health of people is invaluable to those commercial entities.

People's health data is worth vast sums of money. In this next case study, all references to insurance reference pre-2014 insurance, or insurance laws defined before the Affordable Care Act because the Affordable Care Act implemented protections against insurance discrimination. The healthcare insurance industry in the US is also massive, with some sources pointing towards about 1.2 trillion dollars (*IBISWorld - Industry Market Research, Reports, and Statistics*, 2021). Insurance is liable to save large amounts of money on people by understanding the health of their clients. For example, insurance companies can charge less money for someone who will generally be healthier. If a client is regularly injured or sick, that client costs the insurance company more than a client who is never in the hospital. Over time, understanding which clients are more likely to cost the insurance companies more will give the insurance companies larger profit margins. Insurance companies try to gather as much information on the clients as possible, because insurance is a gamble and user data improves the risk margins the companies face.

Pre-2014 insurance is the most straightforward example for why user data is valued by companies. One of the possible routes for insurance companies to get this data is through

COVID-19 contact tracing applications. COVID-19 affects people for extended periods of time. Recent studies from late 2020 say "50% to 80% of patients continue to have bothersome symptoms three months after the onset of COVID-19" (Komaroff, 2020). Some of those symptoms include "fatigue, body aches, shortness of breath, difficulty concentrating, inability to exercise, headache, and difficulty sleeping" (Komaroff, 2020). These are all symptoms that an insurance company may have to pay for eventually. Considering the costs of COVID-19 on human health, people who may have contracted COVID-19 are likely a more high-risk individual for an insurance company than people who did not. Leading up to the Affordable Care Act, people could be charged different premiums for different health histories (Norris, 2020), thus, people who have had COVID-19 could have been charged higher premiums by insurance companies. Since the Affordable Care Act, the health histories of clients are no longer allowed to be considered when creating insurance premiums (Norris, 2020).

It is in contact tracing applications that personal health data can be found regarding COVID-19. Contact tracing applications know which users have tested positive for COVID-19 and which have not. The company behind the user's COVID-19 contact tracing application has personally identifiable health data that affects the long-term health of a user, which is invaluable information. When more than 10,000 dollars are spent each year on health per person in the US, lowering that cost for insurance companies is key.

Protecting the personal health data desired by interested parties comes down to a law called the Health Insurance Portability and Accountability Act, or HIPAA. HIPAA is a national standard which "protects individuals' medical records and other personal health information" (Office for Civil Rights (OCR), 2015). The US Department of Health and Human Services states that HIPAA sets the rules for which health care providers must follow to protect health

information. HIPAA also lays the penalties for the violation of patients' rights. The act gives patients the ability to control their health information, and a key clause is that HIPAA sets the rules for public responsibility of data disclosure, like public health, as seen in the current COVID-19 pandemic (Office for Civil Rights (OCR), 2015). HIPAA also allows patients to understand how their data is being used, limits information release to only the necessary information, and gives individuals control over their own health information.

HIPAA is a comprehensive privacy law in health data that gives people unprecedented control over their personal information. All medical records fall under the HIPAA umbrella, and this includes COVID-19. COVID-19 health data is certainly health information that should not be released except for the public good, and even then, with minimal personal information. However, entities like Apple and Google are not covered under HIPAA, meaning their data collection efforts are not punishable or regulated by HIPAA (Shachar, 2020). This means that people must trust private entities like contact tracing companies to collect data in good faith and delete it when they are done with it, instead of selling it to interested parties. Companies are notoriously bad at this, and contact tracing applications are such a recent invention that people do not understand the implications of the power they are given over their personal information.

One contact tracing application named *Aura* was required by Albion College for all students. *Aura* was given vast power over student's lives on campus. With no opt-out, students were required to share location data constantly. If a student's location was listed as having left campus, they would have their key cards shut down and be unable to access campus resources. *Aura* provided test results and a myriad of other services to enable Albion to keep COVID-19 under control.

*Aura* was subsequently lambasted in the news for the security vulnerabilities and glitches it had. Early in the Albion fall 2020 semester, people learned that the *Aura* app sometimes signed students out of the application without their knowledge, halting the collection of location data (DeWeerd, 2021). There was no transparency to location data collection as well, only a mention of when it would be used, such as a contact tracing event, or a time when a student left campus (Whittaker, 2020). However, the most egregious fault revolved around health data. The local newspaper, the *Grand Valley Lanthorn* said that "data like students' full names and COVID-19 test results could be seen by anyone clever enough to maneuver around the app's code and QR code generator website" (DeWeerd, 2021). This means that people could access the test results of any number of students as it was not protected data. In other words, personal data connected to health data was accessible without hacking or any other nefarious means, and anyone could access the health data of COVID-19 test takers on the *Aura* application. Albion College drastically violated the health privacy of the students it was trying to protect.

The results from *Aura* were lackluster. Even with other mitigation and testing strategies, less than 1,500 students on campus, and *Aura*, there were still at least fifty-five student cases in the 2020 Fall semester at Albion (Albion College, 2021). In the same timeframe, Colby College received fifteen positive student cases (Colby College, 2021) with a significantly larger number of tests completed and students on campus. While these are products of different environments, Colby College students were allowed to leave campus to anywhere in the state of Maine and Albion students were confined to campus. *Aura's* extra protocols largely did not help the situation on campus. Oberlin College, another school in Michigan tested at least 2,000 students with only twenty-eight cases spread between both students and faculty (Oberlin College, 2020). Albion College traded student data for a consistent stream of cases and worse case statistics than

comparably sized schools. It is impossible to know if, without *Aura*, cases would have been worse on campus, but it appears that the app did not help prevent the spread of COVID-19.

Even with the personal health data information, people behind the creation of *Aura* would not be punished, at least not by HIPAA. They are not a public health entity, so the regulations of HIPAA do not apply, even though the application contained personally identifiable health information. This appears to be a major oversight in the safety of digital contact tracing in the United States, and one that needs to be remedied with haste. There is a group of lawmakers working on protecting COVID-19 contact tracing information, however the results are unlikely to pass. A group of Republican senators is looking to introduce the COVID-19 Consumer Data Protection Act, or an act that would help govern COVID-19 contact tracing applications created by entities not covered under HIPAA like Apple and Google (Shachar, 2020). The COVID-19 Consumer Data Protection Act would act more like the CCPA or GDPR from Chapter Two. The act would be an overarching data protection law that would govern any personal information relating to COVID-19, however it has some downsides as well. HIPAA does not allow covered entities to sell health data without the consent of all involved parties. The COVID-19 Consumer Data Protection Act does "allow covered entities to use consumer geolocation or personal health information for purposes beyond COVID-19 contact tracing, including selling data or using it for marketing purposes" (Shachar, 2020). This means that if people assume their health information is private, it still might be sold under this act, even though the act "protects" private user information.

This underlies the need for an expansion of covered entities under HIPAA. HIPAA's personal data protection is strong and unique for federal data protection in the United States but limited in scope. HIPAA would be far more effective if it governed all personal health data, not

just the data which was held by covered entities. It also underlies the need for COVID-19 contact tracing applications to be built under stricter guidelines by trustworthy companies with minimal data collection. All private, non-HIPAA covered entities pose the utmost risk to people's personal health data right now.

# *Chapter 6: Next Steps in Privacy, Data, Software, and STS*

At the core of digital contact tracing, it is the dilemma of the benefits contact tracing provides compared to the dangers posed to personal data privacy. The field of Science, Technology, and Society (STS) is uniquely positioned to study and assess the two sides of digital contact tracing. Often, society is lost when considering solutions to science and technology problems. Yes, with more data, science problems are easier to solve, but in the case of digital contact tracing, the personal data that is collected is people's lives. Personal data collection is like Wall Street traders selling and buying mortgage debts leading up to the 2008 Great Recession. To Wall Street, it was easy money, but at the heart of those trades were the livelihoods of people. When *Care19 Diary* sells user data to partners, ProudCrowd, the maker of *Care19 Diary*, sells numbers, not faces. But to users of the application, ProudCrowd is selling their personal lives and intimate secrets.

The human cost of misused data is hard to understate. The lives of people turn into ones and zeros. Each person can be represented by bits in a computer, with every action they take recorded. With the prevalence of data-collecting services, more and more of people's lives are recorded by machines, each using their information to build profiles of people based on observations made about their computer interactions. The personal data, once collected, cannot be taken back by the user. The user is often not privy to the fact that data was collected, nor what the personal data was, whether they were keystrokes, eye movements, mouse clicks or a bevy of other techniques. Companies know where users travel in the real world and the virtual, how old they are, whether they are single or married, what they look like and what they like to eat. Users will never get this information back from the companies. Some companies, like Google, allow

users to download all data that the company has on them. Google even allows users to delete all collected data on themselves (Smith, 2020). These changes, however, are relatively recent, happening over the past few years as society becomes more aware of the amount of personal data that companies collect.

This system likely needs to be broken, otherwise people will continue to be at the mercy of computers and big companies who know people better than people know themselves. A seismic shift needs to occur, and STS is well positioned to lead that shift as it bonds both the interests of technologists and society. Understanding who is at fault with data collection is key as well. Are the data collection policies a failure of the people who design the programs, or a failure of the programs themselves? In other words, is the seismic shift that is required a shift of people or a shift of the industry? When Apple and Google teamed up to build *Exposure Notifications*, it was with the best interests of society in mind. This may not be the same for the makers of *Care19 Diary* as they may be out to make money. Identifying those problems are key to understanding how to fix a broken system.

Looking to the future is where STS has the most impact. Google's business model is built on data collection and sale, as is much of the technology industry, but if people build the best tools with the best privacy, those tools will be more appealing and better for society overall. It is the personal data collection of today that is bringing the world to its knees. Facebook's data collection policies allow it to create personalized chambers that users are grouped into, generating echo chambers and helping create a more partisan and split world. If people never have discourse, people are never able to reconcile with the other ideologies. This goes beyond Facebook. Without proper personal data protection and the ability to create a way to protect the person that the user is not taking their data, then the darker side of technology will win out.

People's medical records often hold deep personal secrets that a public company does not need access too. The steps from publication of personal data to a situation where people are discriminated against based on personalized medical data are small. Imagine if companies deemed all those with Long-COVID, or a COVID-19 case that lasts more than a few weeks (Collins, 2021), too expensive to hire because of long term health issues. That would create a poverty cycle that hundreds of thousands or millions of people would never recover from. Protecting personal data in this brave new world full of malicious contact tracing is of the utmost importance.

STS is the field of study that defines that interconnectivity between the technology that aids humanity and the technology that ends humanity, or the technology that does both. STS scholars must focus on the rapidly changing world of technology to aid crises of scale, as technology gives humanity the tools to defend itself from the worst the world can give, but in the process may cost humanity dearly. There is a balance between the ability to protect society and the costs that may bring it to its knees. STS scholars must be on the forefront of this study because they are an important part of the shield to protect humanity.

This thesis has talked extensively about the dangers of COVID-19 digital contact tracing applications and how they may be a data-collection and privacy infringement pandemic rather than a panacea for COVID-19. COVID-19 contact tracing apps can be given an unreasonable amount of power in data collection when relating to all sorts of personally identifiable information. Location information and health information are the two primary candidates for data collection and sale. The driving force behind the collection of this personal data by the private sector is to sell it for profit to create better "information profiles" for users. These information profiles can be targeted to better market specific products and tailor advertising to

specific users, driving more profit and therefore making this information valuable. The profit-driven motivation to collect endless amounts of information and bypass the privacy of the individual is a staple in modern society and appears to be so in a technological future. If personal information is profitable, people will continue to try to obtain it at the cost of the user's privacy.

This begs the question, what must change to protect user privacy? Is it a matter of creating better regulations and privacy systems, or is this systematic failure an educational problem in both users and application developers? Oftentimes users are not educated about the personal data privacy they are giving up, though GDPR and CCPA are slowly changing this narrative. With required information panels on websites, users are fed information about a website's data collection policies. However, like a terms and conditions clause, there are no requirements that the information panels be read. Should developers be responsible to change this often-malicious system? Developers are paid by the companies that are benefiting from the same personal data collection epidemics. What data breach is so egregious that a developer turns down the hundreds of thousands of dollars they are being paid? In 2019, Google secretly acquired access to personally identifiable health information on around fifty million people in a program called Project Nightingale. Only after months of watching people fail to focus heavily on the personal data protection guidelines and being shocked by a lack of oversight was one of the 250 team members willing to blow the whistle on the project. Patients and doctors were not informed of the transfer of information from Ascension, the second largest health system in the US, to Google, one of the largest advertisers in the US (Copeland, 2019). If it takes something as monumental as this health data transfer for someone to speak up against the big tech giants from within, then there is something wrong with the system at large.

Langdon Winner's *Do Artifacts Have Politics* discusses how certain technologies can reinforce the same politics of the system they were created in, or how they reflect the biases of their creators (Winner, 1980). Technology is not created in a vacuum. The motivation behind a company creating a contact tracing application or collecting personal health data on fifty million people is unclear. It could be altruistic, but for-profit companies rarely do anything altruistically if it fails to help their business model. Winner argues that technologies must reflect all stakeholders, not just those who create it or the sociological model that it fits into. Stakeholders in modern technology include the users, so technology must not unknowingly disadvantage them.

This begs the question, what politics do contact tracing applications have? That depends on the creator of the contact tracing application, the regulations they fall under, and the motivation for creating such an application. Apple and Google were the two private sector companies best primed to create *Exposure Notifications* due to the prevalence of their operating systems worldwide and did so in a non-invasive manner because it suited them to return the world to normal as quickly as possible. However, other companies like the creators of the *Aura* app may have been focused on generating cash from the endeavor rather than the greater good of society. Governments might require contact tracing applications even if they are not as effective as hoped in order to obtain greater surveillance on their citizens. As a user of these contact tracing applications, it is imperative that one understands the implications, or politics, of the contact tracing application and the motivations of the creator to keep one's data safe.

The greater framework of the technology industry and how it generates money may be culpable for the entirety of the problems listed in this thesis. Since the industry revenue system is primarily based on trading advertising information for cash, it is the system that is at fault for

compromising user data. To be competitive, companies must gain better data and to do so, must further compromise user privacy. This cycle is endlessly perpetuating until nothing private remains at all, which is an unreasonable solution. Thus, something needs to change in the technology economic system overall. There needs to be a move from an advertising-based revenue model to something that users can control, or a different system that accounts for all stakeholders, not just those in power.

COVID-19 contact tracing applications are a case study on the wider systemic failures that the world faces online today. An altruistic digital contact tracing app would be nearly identical to the Apple and Google *Exposure Notification* system because it collects no data beyond what is required and maintains the privacy of the users. However, the key is Apple and Google built a framework, not an application, so other people must implement the framework themselves. Governments can step in to take this process over and out of the hands of the companies who must work in the system, thus breaking the for-profit system.

Ireland's contact tracing app, *COVID Tracker,* was commissioned by the Health Service Executive (HSE), but written by Nearform, an Irish software developer. *COVID Tracker* is built on the *Exposure Notification System* by Apple and Google. Every piece of the application is voluntary, from download to reporting a positive COVID-19 test. The HSE is given random IDs to match the users, though these are changed every 14 days. No names or personal information are identifiable from these anonymous random ids. The HSE voluntarily collects phone numbers to notify the user of need to quarantine in the event of a close contact for an extended period. All data on local smartphones is deleted after fourteen days, and most importantly, there is no use of GPS or other personal data requests. Users only need to leave Bluetooth turned on to take advantage of the contact tracing system (*Privacy and how we use your data*, 2020).

It is difficult to determine how well *COVID Tracker* has worked. As of October 21, 2020, "over 3,000 users of the COVID Tracker app in Ireland who have tested positive have uploaded their random IDs so that others can be alerted of close contacts. More than 5,800 people have been sent close contact alerts as a result of carrying the app" (Department of Health, 2020). Between July 7, when the app was launched, and October 21, when these figures were released, there were approximately 27,884 new cases in Ireland. This means that about 11% of cases were traced with *COVID Tracker*. This 11% of cases were traced without the loss of privacy by users because Ireland's government stepped in, and strong data privacy regulations set by GDPR were in place. It is possible for the private sector to do good with help from the public sector by setting strong regulations and disincentivizing personal data collection by paying the private sector for the work it does in cash, not data.

This thesis opened with the question, "With the advent of mobile application-based contact tracing, how is the online privacy of society at risk? How can those risks be mitigated when tracing is conducted by the private and public sector?" Over the course of the past six chapters, this thesis concludes with the understanding that the private sector and public sector provide two different sets of risks, but by being diligent and implementing best-practices regarding cybersecurity, creating and enforcing data protection regulations, being transparent in application implementation and privacy, and adopting new contact tracing applications carefully, we can reduce the dangers of losing privacy to contact tracing applications.

It is important to understand that regulations and diligent research when adopting contact tracing applications is a temporary fix for a larger problem, which is the revenue generation of the technology industry at large. Without a systematic change, whether it be another pandemic or crisis in the future in which technology is mass-distributed, there will be privacy breaches that

could prove detrimental to society. While COVID-19 is a pandemic that affects people's health, the downfall of data privacy is also a global pandemic that requires urgent attention. Until the greater problem is fixed, no personal data will be safe, but in the meantime, it is of the utmost importance that regulations are created and the creators of the personal data collecting tools like COVID-19 contact tracing apps are held accountable to those regulations.

# *References*


Adam, D. (2020). A guide to R — The Pandemic's Misunderstood Metric. *Nature*, *583*(7816), 346–348. https://doi.org/10.1038/d41586-020-02009-w

Alan Furman Westin. (2015). *Privacy and Freedom*. New York Ig Publishing. (Original work published 1967)

Albion College. (2021, April 23). *COVID-19 Testing Updates*. Albion College. https://web.albion.edu/together-safely/health-and-wellness/testing-updates?limit=4&start=20

Apple Inc. (2020). *Privacy-Preserving Contact Tracing - Apple and Google*. Apple. https://covid19.apple.com/contacttracing

Atamaniuk, M. (2020, October 14). *Which Company Uses the Most of Your Data?* Clario.co. https://clario.co/blog/columns/which-company-uses-most-data/#:~:text=Social%20media%20collects%20more%20data%20than%20anybody%20else&text=As%20well%20as%20the%20usual

*Bahrain, Kuwait and Norway contact tracing apps a danger for privacy*. (2020, June 16). Www.amnesty.org; Amnesty International. https://www.amnesty.org/en/latest/news/2020/06/bahrain-kuwait-norway-contact-tracing-apps-danger-for-privacy/

BBC. (2019). *The Black Death - Revision 4 - KS3 History - BBC Bitesize*. BBC Bitesize; BBC. https://www.bbc.co.uk/bitesize/guides/z7r7hyc/revision/4

Becerra, X. (2018, October 15). *California Consumer Privacy Act (CCPA)*. State of California - Department of Justice - Office of the Attorney General. https://oag.ca.gov/privacy/ccpa

Besinger, G. (2020, January 21). So far, under California's new privacy law, firms are disclosing too little data — or far too much. *The Washington Post*. https://www.washingtonpost.com/technology/2020/01/21/ccpa-transparency/

Bohan, A.-M., & Bollard, C. (2020, June 7). *Ireland: Data Protection Laws and Regulations 2020*. ICLG.com. https://iclg.com/practice-areas/data-protection-laws-and-regulations/ireland#

Brandom, R. (2020, September 1). *Apple and Google announce new automatic app system to track COVID exposures*. The Verge. https://www.theverge.com/2020/9/1/21410281/apple-google-coronavirus-exposure-notification-contact-tracing-app-system

Browne, R. (2020, July 3). *Why coronavirus contact-tracing apps aren't yet the "game changer" authorities hoped they'd be*. CNBC. https://www.cnbc.com/2020/07/03/why-coronavirus-contact-tracing-apps-havent-been-a-game-changer.html

*Campus Health Tracker for K-12 Schools*. (2020). https://www.ipc-global.com/campus-health-tracker

CDC. (2018, September 14). *Health Insurance Portability and Accountability Act of 1996 (HIPAA)*. Centers for Disease Control and Prevention. https://www.cdc.gov/phlp/publications/topic/hipaa.html

CDC. (2020a, February 11). *Coronavirus Disease 2019 (COVID-19)*. Centers for Disease Control and Prevention. https://www.cdc.gov/coronavirus/2019-ncov/global-covid-19/operational-considerations-contact-tracing.html#:~:text=Close%20contact%20is%20defined%20by

CDC. (2020b, March 28). *Coronavirus Disease 2019 (COVID-19) in the U.S.* Centers for Disease Control and Prevention. https://covid.cdc.gov/covid-data-tracker/#trends_dailytrendscases

Colby College. (2021, April 26). *Health Code and Testing Data*. Covid-19. https://covid19.colby.edu/health-code-and-testing-data/

Collins, F. (2021, February 23). *NIH launches new initiative to study "Long COVID."* National Institutes of Health (NIH). https://www.nih.gov/about-nih/who-we-are/nih-director/statements/nih-launches-new-initiative-study-long-covid

Constine, J. (2012, August 22). *How Big Is Facebook's Data? 2.5 Billion Pieces Of Content And 500+ Terabytes Ingested Every Day*. TechCrunch; TechCrunch. https://techcrunch.com/2012/08/22/how-big-is-facebooks-data-2-5-billion-pieces-of-content-and-500-terabytes-ingested-every-day/

Copeland, R. (2019, November 11). *Google's "Project Nightingale" Gathers Personal Health Data on Millions of Americans*. WSJ; Wall Street Journal. https://www.wsj.com/articles/google-s-secret-project-nightingale-gathers-personal-health-data-on-millions-of-americans-11573496790

Cullen, R. (2008). Culture, identity and information privacy in the age of digital government. *Emerald Group*, *33*(3), 405–421.

Dehaye, P.-O. (2020, October 19). *Inferring distance from Bluetooth signal strength: a deep dive*. Medium. https://medium.com/personaldata-io/inferring-distance-from-bluetooth-signal-strength-a-deep-dive-fe7badc2bb6d

Department of Health. (2020, October 21). *Ireland is one of the first countries to link contact tracing apps with other EU Member States*. Www.gov.ie. https://www.gov.ie/en/press-

release/2dc55-ireland-is-one-of-the-first-countries-to-link-contact-tracing-apps-with-other-eu-member-states/

DeWeerd, J. (2021, January 25). *Despite thousands of new daily COVID-19 cases, students struggle to trust contact tracing apps*. Grand Valley Lanthorn. https://lanthorn.com/79663/news/despite-thousands-of-new-daily-covid-19-cases-students-struggle-to-trust-contact-tracing-apps/

DeWitte, S. N. (2014). Mortality Risk and Survival in the Aftermath of the Medieval Black Death. *PLoS ONE*, *9*(5), e96513. https://doi.org/10.1371/journal.pone.0096513

Donohue, M. J. (2020, July 20). *New Procedures, Same Old FERPA: How the Return to School Impacts Student Records | Lexology*. Www.lexology.com. https://www.lexology.com/library/detail.aspx?g=9fc60673-e9ea-4b4f-a721-31ac3c6921ce

Eisenberg, J. (2020, February 5). *R0: How scientists quantify the intensity of an outbreak like coronavirus and predict the pandemic's spread*. The Conversation. https://theconversation.com/r0-how-scientists-quantify-the-intensity-of-an-outbreak-like-coronavirus-and-predict-the-pandemics-spread-130777

Electronic Privacy Information Center. (2010). *EPIC - Locational Privacy*. Epic.org. https://epic.org/privacy/location/

Facebook. (n.d.). *Benefits of adding click and impression tags*. Facebook Business Help Center. Retrieved April 26, 2021, from https://www.facebook.com/business/help/1051490305007669?id=399393560487908

Fairchild, A., Gostin, L., & Bayer, R. (2020, July 20). *Contact tracing's long, turbulent history holds lessons for COVID-19*. Contact Tracing's Long, Turbulent History Holds Lessons

for COVID-19. https://news.osu.edu/contact-tracings-long-turbulent-history-holds-lessons-for-covid-19/

Farronato, C., Iansiti, M., Bartosiak, M., Denicolai, S., Ferretti, L., & Fontana, R. (2020, July 15). *How to Get People to Actually Use Contact-Tracing Apps*. Harvard Business Review. https://hbr.org/2020/07/how-to-get-people-to-actually-use-contact-tracing-apps

Fowler, G. A. (2020, May 21). Perspective | One of the first contact-tracing apps violates its own privacy policy. *Washington Post*. https://www.washingtonpost.com/technology/2020/05/21/care19-dakota-privacy-coronavirus/

Fox, M. (2013, February 22). Alan F. Westin, Who Transformed Privacy Debate Before the Web Era, Dies at 83. *The New York Times*. https://www.nytimes.com/2013/02/23/us/alan-f-westin-scholar-who-defined-right-to-privacy-dies-at-83.html

Foy, K., & Kahn, A. (2020, June 16). *Contact tracing without Big Brother*. MIT Technology Review. https://www.technologyreview.com/2020/06/16/1002982/contact-tracing-without-big-brother/

Free, M. (2019). *MyBib: Free Citation Generator*. Google.com. https://chrome.google.com/webstore/detail/mybib-free-citation-gener/phidhnmbkbkbkbknhldmpmnacgicphkf/related

Gibbs, S. (2016, October 13). *Google Maps: a decade of transforming the mapping landscape*. The Guardian; The Guardian. https://www.theguardian.com/technology/2015/feb/08/google-maps-10-anniversary-iphone-android-street-view

Gong, M., Wang, S., Wang, L., Liu, C., Wang, J., Guo, Q., Zheng, H., Xie, K., Wang, C., & Hui, Z. (2020). Evaluation of Privacy Risks of Patients' Data in China: Case Study. *JMIR Medical Informatics*, *8*(2). https://doi.org/10.2196/13046

*Guidance on the Protection of Personal Identifiable Information | U.S. Department of Labor*. (2019). Dol.gov; Department of Labor. https://www.dol.gov/general/ppii

HIPAA Journal. (2018, April 25). *Comparison of European and American Privacy Law*. HIPAA Journal. https://www.hipaajournal.com/comparison-of-european-and-american-privacy-law/

*IBISWorld - Industry Market Research, Reports, and Statistics*. (2021, April 26). Www.ibisworld.com. https://www.ibisworld.com/industry-statistics/market-size/health-medical-insurance-united-states/

Igo, S. E. (2008). *The Averaged American Surveys, Citizens, and the Making of a Mass Public*. Harvard University Press.

Igo, S. E. (2020). *The Known Citizen: A History of Privacy in Modern America*. Harvard University Press.

important Safety Technologies. (2020, June 4). *We Tested Mobile GPS/GNSS Accuracy and Found Some Surprising Results*. Medium; Medium. https://medium.com/@importanttech/we-tested-mobile-gps-gnss-accuracy-and-found-some-surprising-results-b9ec35873e2e

Jasanoff, S. (2016). *The Ethics of Invention : Technology and the Human Future*. W.W. Norton & Company.

Koh, D. (2020, March 20). *Singapore government launches new app for contact tracing to combat spread of COVID-19*. MobiHealthNews.

https://www.mobihealthnews.com/news/asia-pacific/singapore-government-launches-new-app-contact-tracing-combat-spread-covid-19

Komaroff, A. (2020, October 15). *The tragedy of the post-COVID "long haulers."* Harvard Health Blog. https://www.health.harvard.edu/blog/the-tragedy-of-the-post-covid-long-haulers-2020101521173

Kretzschmar, M. E., Rozhnova, G., Bootsma, M. C. J., van Boven, M., van de Wijgert, J. H. H. M., & Bonten, M. J. M. (2020). Impact of delays on effectiveness of contact tracing strategies for COVID-19: a modelling study. *The Lancet Public Health*, *5*(8), e452–e459. https://doi.org/10.1016/s2468-2667(20)30157-2

Lane, F. S. (2011). *American Privacy: The 400-Year History of Our Most Contested Right*. Beacon. (Original work published 2009)

Lapowsky, I. (2018, March 17). *Trump Campaign Data Consultants Cambridge Analytica Took 50 Million Facebook Users' Data*. Wired; WIRED. https://www.wired.com/story/cambridge-analytica-50m-facebook-users-data/

Lapowsky, I. (2019, March 17). *How Cambridge Analytica Sparked the Great Privacy Awakening*. WIRED; WIRED. https://www.wired.com/story/cambridge-analytica-facebook-privacy-awakening/

Leswing, K. (2020, May 6). *Companies could require employees to install coronavirus-tracing apps like this one from PwC before coming back to work*. CNBC. https://www.cnbc.com/2020/05/06/pwc-is-building-coronavirus-contact-tracing-software-for-companies.html

MacAskill, E., Dance, G., Cage, F., Chen, G., & Popovich, N. (2014, March 23). *NSA files decoded: Edward Snowden's surveillance revelations explained*. The Guardian; The

Guardian. https://www.theguardian.com/world/interactive/2013/nov/01/snowden-nsa-files-surveillance-revelations-decoded#section/1

Marr, B. (2015, February 25). *A brief history of big data everyone should read*. World Economic Forum. https://www.weforum.org/agenda/2015/02/a-brief-history-of-big-data-everyone-should-read/

Menand, L. (2018, June 11). *Why Do We Care So Much About Privacy?* The New Yorker. https://www.newyorker.com/magazine/2018/06/18/why-do-we-care-so-much-about-privacy

Morrison, S. (2020, December 2). *A surprising number of government agencies buy cellphone location data. Lawmakers want to know why.* Vox. https://www.vox.com/recode/22038383/dhs-cbp-investigation-cellphone-data-brokers-venntel

Nield, D. (2020, January 12). *All the Ways Facebook Tracks You—and How to Limit It*. Wired. https://www.wired.com/story/ways-facebook-tracks-you-limit-it/

North Dakota State Government. (2020). *Care19*. ND Response. https://ndresponse.gov/covid-19-resources/care19

Oberlin College. (2020, December 15). *Fall 2020 Campus Testing Statistics*. Oberlin College and Conservatory. https://www.oberlin.edu/obiesafe/statistics/fall-2020

Office for Civil Rights (OCR). (2015, October). *What does the HIPAA Privacy Rule do*. HHS.gov. https://www.hhs.gov/hipaa/for-individuals/faq/187/what-does-the-hipaa-privacy-rule-do/index.html

Osman, M. (2020, November 6). *Wild and Interesting Facebook Statistics and Facts (2019)*. Kinsta; Kinsta. https://kinsta.com/blog/facebook-statistics/

Pew Research Center. (2016, January 29). *Americans increasingly use smartphones for more than voice calls, texting*. Pew Research Center. https://www.pewresearch.org/fact-tank/2016/01/29/us-smartphone-use/ft_01-27-16_smartphoneactivities_640/

Pew Research Center. (2019, June 12). *Mobile Fact Sheet*. Pew Research Center: Internet, Science & Tech; Pew Research Center: Internet, Science & Tech. https://www.pewresearch.org/internet/fact-sheet/mobile/

*Privacy and how we use your data*. (2020, December 31). Www2.Hse.ie. https://www2.hse.ie/conditions/coronavirus/covid-tracker-app/privacy-and-how-we-use-your-data.html

Rust, M., Haggin, P., & Wu, Y. (2020, April 3). *Where America Is Staying Home*. WSJ. https://www.wsj.com/graphics/where-america-is-staying-home/

Sato, M. (2020, December 14). *Contact tracing apps now cover nearly half of America. It's not too late to use one.* MIT Technology Review. https://www.technologyreview.com/2020/12/14/1014426/covid-california-contact-tracing-app-america-states/

Servick, K. (2020, May 21). *COVID-19 contact tracing apps are coming to a phone near you. How will we know whether they work?* Science | AAAS. https://www.sciencemag.org/news/2020/05/countries-around-world-are-rolling-out-contact-tracing-apps-contain-coronavirus-how

Shachar, C. (2020). Protecting Privacy In Digital Contact Tracing For COVID-19: Avoiding A Regulatory Patchwork. *Health Affairs*. https://doi.org/10.1377/hblog20200515.190582

Shin, Y., Berkowitz, B., & Kim, M. J. (2020, March 25). *How a South Korean church helped fuel the spread of the coronavirus*. Washington Post. https://www.washingtonpost.com/graphics/2020/world/coronavirus-south-korea-church/

Smith, D. (2020, June 28). *Google collects a frightening amount of data about you. You can find and delete it now*. CNET. https://www.cnet.com/how-to/google-collects-a-frightening-amount-of-data-about-you-you-can-find-and-delete-it-now

Stasha, S. (2020, July 29). *The State of Health Care Industry (2020)*. PolicyAdvice. https://policyadvice.net/insurance/insights/healthcare-statistics/

Szoldra, P. (2016, September 16). *This is everything Edward Snowden revealed in one year of unprecedented top-secret leaks*. Business Insider. https://www.businessinsider.com/snowden-leaks-timeline-2016-9

Tenforde, M. W. (2020). Symptom Duration and Risk Factors for Delayed Return to Usual Health Among Outpatients with COVID-19 in a Multistate Health Care Systems Network — United States, March–June 2020. *MMWR. Morbidity and Mortality Weekly Report*, *69*. https://doi.org/10.15585/mmwr.mm6930e1

Whittaker, Z. (2020, August 19). *Fearing coronavirus, a Michigan college is tracking its students with a flawed app*. TechCrunch. https://techcrunch.com/2020/08/19/coronavirus-albion-security-flaws-app/

Winner, Langdon. "Do Artifacts Have Politics?" *Daedalus* 109, no. 1 (1980): 121-36. Accessed October 19, 2020. http://www.jstor.org/stable/20024652.