Principles for Technology-Assisted Contact-Tracing

By Daniel Kahn Gillmor
April 16, 2020

The current COVID-19 pandemic is a deadly crisis, but also an opportunity to pull together as a society. Regrettably, it's also an opportunity for would-be authoritarians and powerful corporations to expand their power. One way they may try to do so is through the use of technology and data to address the pandemic.

In the last few weeks we have seen many proposals for technology-assisted "contact tracing." Contact tracing is a longstanding public health technique that works by identifying everyone whom a sick person may have exposed, and helping them identify their risks and take appropriate action. But traditional contact tracing techniques — carried out through in-person interviews — are labor-intensive, and often slow compared to a fast-moving pathogen like SARS-CoV-2, the virus that causes COVID-19. The proposed systems would instead rely on location or proximity detection by mobile phones to selectively deliver alerts about potential exposures.

While some of these systems may offer public health benefits, they may also cause significant risks to privacy, civil rights, and civil liberties. If such systems are to work, they must be widely adopted. But that won't happen if they do not enjoy strong trust within the population. As a group of public health experts wrote in a Johns Hopkins report recently, "To increase the chances that [contact tracing] efforts will be effective, trusted, and legal, use of technology in the contact tracing space should be conceived of and planned with extensive safeguards to protect private information from the beginning." Without such safeguards, the authors explain, "individuals may be unwilling to participate."

In this paper, we outline general principles that should guide the consideration of any proposal for technology-assisted contact-tracing, or TACT. This document does not address fine-grained details, either technical or legal, but sets out principles to help evaluate any TACT proposal. Given the trans-jurisdictional nature of many of these schemes, specific legal protections may take different forms in different contexts. But architectural and design principles will have an impact anywhere such a system is deployed.

Before addressing those principles, two important points should be noted.

First, technologists, policymakers, and others should keep in mind that there is a very real chance these systems will simply not prove practical in real world conditions. That could be so for a number of reasons. For example, their accuracy rate may prove low enough and the complexity of human interactions high enough that they generate too many false alarms, sending healthy people to medical facilities where they might become infected, or into quarantine for no good reason.

Second, it is vital to recognize that on its own, a TACT scheme does nothing to help stem the spread of COVID. It is useful only if those who learn of possible exposures to COVID are able to do something about it: get tested, get counseling, get treatment, or take measures like self-isolation. But it is useless if those services are unavailable or unaffordable. And advice that encourages self-isolation is implausible if the user of the TACT system or their family cannot afford to do so. The lack of adequate and equitable social and public health support systems would limit the effectiveness of any TACT system — potentially risking people's privacy without bringing them benefits.

The basic principles we see for evaluating a TACT are:

- Not displacing non-technical measures
- Voluntary
- Non-punitive
- Built with public health professionals
- Privacy-preserving
- Non-discriminatory
- Minimal reliance on central authorities
- Data minimization everywhere
- No data leakage
- Measurable impact
- Have an exit strategy
- Narrowly-tailored to target a specific epidemic
- Auditable and fixable
- Sustainably maintained

Technical Background

Many groups have floated TACT proposals since the emergence of COVID-19, and some of those proposals are distinctly dangerous. In particular, some schemes would regularly supply information about everyone in a society — location information, movements, social encounters, phone numbers, health data, etc. — to a central authority (a government agency,

corporation, or any other entity). These schemes are privacy-unfriendly, and the societies that deploy them can become (or are already) surveillance states with dangerous powers of social control vested in the central authorities. In addition, the location data typically generated by cell phones is not precise enough to identify epidemiologically relevant contacts, i.e. such as those within the requisite distance or with the relevant type of exposure. We reject these privacy-unfriendly TACT proposals outright because they do not strike the right balance between effectiveness, necessity, and intrusion.

Any contact-tracing scheme, tech-assisted or otherwise, does risk exposing an infected person's medical condition (which is sensitive health data) to their potential contacts. But more privacy-friendly schemes do exist. As of the publication of this paper, there are several TACT proposals that aim to be privacy-friendly including DP-3T, PACT, TCN, and the Apple and Google proposal.

All these proposals follow roughly the same pattern, using mobile phones presumed to be carried by the majority of the population.

Each phone has a way to interact with neighboring phones — typically through Bluetooth Low Energy beacons. If you enroll in the proposed systems, your phone announces itself to its neighbors with a different large random number every few minutes. Your phone records the numbers it has announced, and the announcements it "hears" from neighboring phones, creating a history that goes back a few weeks.

If you are diagnosed as infected, you voluntarily upload the recent history of the numbers your phone has announced to a central server. We can call this the "exposure database," but it's just a list of random numbers posted to a website.

If you haven't yet been infected, your phone regularly fetches the list of numbers that other people's phones have uploaded to the exposure database, and compares it to the list of numbers that your phone has "heard" from its neighbors in past few weeks.

If some numbers that your phone saw show up in the exposure database, then you know you've been in contact with someone who was infected. You can use that information to take action: get tested, self-isolate, make a call to your healthcare provider, etc.

The central server that hosts the exposure database doesn't know anything about your location, your movements, or whom you contacted. None of these systems need to know anything about your geographic location.

Such a Bluetooth system would be better than a location-tracking system. It would not be perfect at identifying when two phones were in close contact, but it would be more accurate and would result in fewer false positives. In addition, it would not use geolocation data, which can be incredibly revealing and privacy-invasive. Of course, it would record associations, which are also highly revealing, but — if done correctly — the identifiers that phones announced and heard would not be traceable back to an identified individual. In this way, such a system could be even more privacy-preserving than traditional contact tracing, because the public health authorities would not even need to directly handle data of an at-risk party before that person opted into a contact with the medical system.

But many risks remain.

Not displacing non-technical measures

No TACT scheme should divert resources from known, effective public health measures like testing, counseling, research, and treatment. Indeed, every TACT proposal is predicated on the availability of widespread, affordable, and prompt testing, so it would be pointless to deploy automated contact tracing at the expense of traditional medical and social interventions.

Voluntary

Any such application should be voluntary at each step. This principle merely acknowledges the reality of the situation: People will avoid participation in a privacy-sensitive scheme that seems compulsory and antagonistic and therefore untrustworthy, and people have fundamental rights to privacy and association that they might legitimately fear will be affected by a TACT scheme. In addition, the United States has never compelled people to carry a phone, much less to install a specific app on their phone, and doing so would represent an enormous and consequential step.

Note that voluntariness applies at many points in such a system. In the commonly proposed schemes described above, points where a user might exercise choice include:

- Whether to carry a Bluetooth-enabled phone on their person at all (this may not even be a choice for people of limited means).
- Whether to install such an app on their phone.
- Whether to leave the app operating in "broadcast and record mode" at any given time. Apps should at least have a "snooze" feature that allows the user to disable it for periods that they know they are not interacting with others, such as sleeping at home alone.
- If receiving an alert due to a match with the exposure database, whether and how to react to that alert.
- Freedom to select which medical providers or counselors to engage with, and how much of one's contact log to share with those counselors.
- If diagnosed as infected, whether to upload their log of contacts to the exposure database.
- The ability to selectively redact these logs before upload where an infected user retroactively identifies specific blocks of time that they do not want to upload. This can help to reduce false positives, for example if a user knows that identified contacts during that time were inaccurate (for example because they were in a car or wearing protective gear). It would also encourage people whose records include particularly sensitive contact information to at least volunteer the non-sensitive part of their records rather than failing to participate completely.

Note that social context matters for voluntariness as well. If installing and running the app is nominally "voluntary" but people are obliged to install and run the app in order to be able to (for example) go to work or shop for food, then it is not truly "voluntary."

Non-punitive

A TACT application must not be used for punitive measures such as criminal prosecution or immigration enforcement, or even for potentially health-related measures that the subject may view as punitive, like quarantine enforcement. Any TACT scheme is effective in proportion to its adoption by the general population. And adoption by the general population depends on trust that the application will not be used to harm the user. Both technical and legal safeguards must ensure that the TACT scheme warrants the trust that its effectiveness depends upon.

Some people will inevitably be bad actors in a pandemic, but a TACT is not an appropriate channel for dealing with such people because awareness of the tool as an enforcement mechanism and fear of repercussions could drive away many people, the vast majority of whom would most likely act responsibly.

Built with public health professionals

A TACT scheme built by technologists alone, without open consultation with public health professionals, experts in infectious disease, and other domain-specific experts is bound to fail.

First and foremost, we need experts to help decide when two phones have been close enough to each other to be medically relevant. Most technological methods can only measure rough physical proximity of mobile devices, which is not the same as an epidemiological "contact" opportunity for a virus to hop from one host to another. For example, Bluetooth devices can appear "nearby" even if they are on opposite sides of a wall that a virus cannot cross.

There are also types of viral transmission that cannot be detected at all by these radio signals; not only are there inevitably going to be people who don't carry such a device (or don't opt in), but the virus can jump hosts by temporarily resting on surfaces ("fomites") that are then handled by people. No use of phone data will capture that.

Experts in infectious disease need to be able to give guidance to application developers about what kinds of measurements come the closest to an actually risky contact, to help to decide what kinds of data should cause an alert to be shown, and to evaluate the ways that various TACT proposals implement such guidance.

Developers will also need to decide whether phones should alert people if they've been near someone who has experienced symptoms of COVID-19, or whether they should only issue alerts to people who have received a confirmed positive result from an accredited SARS-CoV-2 laboratory test — or something in between.

One answer might be "better safe than sorry, cast a wide net." But if the goal is to use the app to prioritize delivery of scarce medical resources or to avoid wasting medical resources on false alarms, alerting everyone could be no better than alerting no one at all.

These sorts of balancing decisions about how to effectively prioritize medical care based on technologically estimated risk can only be made well — and adjusted based on experience — in concert with experts who understand the underlying characteristics of disease transmission, what is happening on the ground in various communities, and the landscape of remedies available to people who might receive such an alert.

Experts should also lead systematic efforts to measure whether, once deployed, a TACT system is proving necessary, effective, and proportionate given its expense, harm to privacy, and any other costs.

Privacy-preserving

Some of these systems are decidedly privacy-unfriendly, leaking far more information about their users than is necessary to perform the public-health related function of stemming the epidemic. If these systems only endangered the fundamental rights of people who use them, that would be bad enough. But they also endanger their public-health mission, by causing people to mistrust and abandon them when they are needed.

A TACT system must not collect or transmit any data not strictly necessary for its public health purpose. It should adopt the strongest possible technical and legal safeguards for any data that is collected and transmitted.

For example:

- Data should stay local to devices controlled by the end user where possible.
- Any identifiers used by the system should be un-linkable to other identifiers, including not only such things as phone numbers and IP addresses, but also other anonymous identifiers. Any data that gets exposed can leak private information, and when multiple data points can be interlinked, it becomes significantly easier to tie them to specific people.
- The use of data that is difficult or impossible to anonymize (such as location data) should be avoided. A privacy-preserving TACT shouldn't be coupled with a system that collects problematic data.

All privacy-preserving measures that can be enforced technically should be enforced technically, and not by policy alone. Policy enforcement can be more easily overridden in the future — and worse, overridden in secret — which makes the system as a whole less trustworthy.

Non-discriminatory

We have repeatedly seen technological systems deployed that further entrench existing social inequities, such as recidivism risk scores that are more likely to keep people of color incarcerated, hiring algorithms that exacerbate existing gender disparities, and facial recognition systems that are more likely to misidentify people of color as criminals.

Any TACT that is worth considering should attempt to account for and mitigate these types of risks. Given that most tech-augmented systems rely at some level on the availability of both hardware and Internet access, the technological dimension of these issues already poses a challenge. For example, if accelerated access to testing is granted to people because they own a Bluetooth-capable device and have high-quality Internet access, that may amplify the existing disadvantages faced by poor communities (which are already at elevated risk of complications due to COVID-19).

The deployment of any TACT should be coupled with ongoing efforts to identify populations likely to be misrepresented or excluded by the system. It should also include funded measures to support these communities (such as expanded traditional contact tracing, subsidized Internet access, etc.) that are designed and led by relevant experts, including members of the affected populations themselves.

In particular, a TACT should be designed carefully to protect the rights of people who test positive for the targeted epidemic as well as people who are at elevated risk. Any scheme that instead makes life worse for such people is not only fundamentally unjust, but also likely less effective at encouraging the necessary participation.

Minimal reliance on central authorities

The most privacy-unfriendly TACTs ship huge amounts of data to central authorities, leaving users little to no control over what happens to that data once it leaves their device. There are several major risks here, the most obvious of which is that the central authority could decide to use this detailed information to punish or control people directly. Even in cases where the authority is well-intentioned toward the population it gathers data from, that authority could be compelled to turn over data to a more malicious authority in the future, or it could be technically compromised.

More privacy-preserving TACTs generally avoid sending detailed information like phone numbers or location history to central authorities. They might still rely on central authorities for some purposes, like serving as a gatekeeper to prevent flooding attacks on the system. A malicious party wanting to cause problems (a hostile nation, or someone wanting to interfere with Election Day, perhaps) could, for example, set up Bluetooth beacons around a city, broadcasting identifiers to ordinary people as they move around — and then falsely report themselves as infected in order to sow panic. A good gatekeeper could defend the system against such an attack. But a privacy-preserving TACT does not send sensitive data to those authorities to be stored.

Even these reduced dependencies on central authorities can potentially be problematic. For example, could a gatekeeping central authority deliberately exclude members of a particular community from notifications? Or, in a multi-jurisdictional context, could the authorities in one region disagree with the authorities in another region about how to execute their role in the TACT scheme? And what are the implications for users if cross-jurisdictional disagreements arise? Do users of the scheme need to be able to decide which central authorities they rely on? How can they realistically make that decision?

Any TACT scheme should be clear about which central authorities it expects its users to rely on, and for what purpose. It must make these decisions clear so that the public can know in whom they are placing their trust, and for what purpose.

Data minimization everywhere

Any TACT scheme must handle some data, no matter how privacy-preserving it aims to be. Different parts of the scheme will handle different types of data, and exchange them or expose them at different times.

Good data minimization is critical, and it must be employed at every point in the system: at the central servers that host an exposure database, as well as on the endpoint devices that might participate in such a system.

Data minimization has many aspects, but some of the basics include:

- **Keep data encrypted at rest where possible.** For example, a TACT phone-based app should not store an easily accessible database of time stamped location information: If a compromise of the app results in information leakage about the user, or if the app's data store becomes accessible to law enforcement, it will be rightly seen as a potentially punitive or dangerous tool.

- **Schedule data destruction.** Any data retained by any part of a TACT scheme should have a defined expiration date that is no later than the latest epidemiologically-relevant date. When it expires, it should be completely purged by any component that has access to it.

- **Avoid sharing fine-grained data.** For example, the exposure database that accepts submissions should batch submissions together in large numbers to increase the size of the anonymity set of infected users. It should not store logs of when smaller sets of published identifiers are received. It should retain only the aggregated, batched updates.

All these data minimization steps should be clearly documented and automatically enforced by both technical and legal measures.

No data leakage

While some data in a TACT scheme is available to central authorities, and other data is held by end-user devices, none of these parties should deliberately release data to uninvolved parties. In particular, there should be legal, procedural, and technical safeguards to prevent law enforcement agencies from accessing any of these data stores, combined with mechanisms to detect such access, and clear, enforceable penalties for doing so.

Information held by these systems is in a public trust for the purposes of limiting the spread of the pandemic within specific, well-defined parameters. Violations of that public trust put the whole system at risk.

Measurable impact

A TACT system must give people a way to know whether it is working or not. These system metrics need not be perfect or exact, but there must be a rough sense of whether the scheme as a whole is helping to reduce disease transmission.

This could be as simple as counting the number of exposures reported (which any of the more privacy-friendly systems will have access to). It could also include a count of the number of app installations, the number of app-facilitated contacts with the medical system, and the number of self-isolations, all of which could be self-reported voluntarily. Details of how the impact is measured should be published, and aggregate metrics should be made available to the general public on a timely basis so that we know what kinds of tradeoffs we are making by embracing a TACT scheme. A TACT scheme with these measurements should set pre-defined goals for what kind of an impact it expects to be able to make.

Have an exit strategy

A TACT scheme that targets a particular epidemic should not last beyond the particular disease it targets. Nor should a TACT scheme that is shown to lack effectiveness be continued. That means that from the outset, any responsible TACT should have built-in and publicly understood measures for phasing itself out. This means:

- Knowing when to "declare victory" and cease operations
- Knowing when to bow out due to lack of effective impact
- Knowing how to shut down any central servers or authorities safely
- Being able to uninstall itself or stop operation correctly on end user devices

There will be pressure to engage in "surveillance creep" from people who want to see a TACT system remain widely deployed for other purposes, benign or malicious. A responsible TACT scheme will be designed to resist those pressures.

Narrowly-tailored to target a specific epidemic

While lessons may be learned that might be applicable in future pandemics, it is unwise to attempt to devise a technical system that would work against all future pandemics. This is true because the characteristics of disease transmission and the public health response may vary widely depending on the disease.

A proximity-tracing tool that works well for the particular biophysical characteristics of COVID-19 transmission is unlikely to be a good fit for fighting other pathogens. Consider HIV, which spreads primarily by unprotected sex and sharing of needles. Bluetooth proximity measurements are entirely the wrong tool to approximate a potential exposure to HIV.

Furthermore, public health guidance on what to do in response to an alert of potential exposure might differ widely depending on prevalence in the population, types of transmission,

incubation period, potential symptoms, social stigma, etc. A TACT tool may not make sense for a COVID-like disease that remains relatively rare, for example.

A TACT scheme that aims to target all imaginable future pandemics would necessarily require gathering information that isn't needed to combat the spread of COVID-19. That would reduce trust in the system, and managing the extra data would divert engineering resources needed to make the most effective tool to counter the specific epidemic we are faced with.

Auditable and fixable

Users and communities faced with the choice of participating in a TACT need clear signals that the software involved is trustworthy and does what it intends to do.

Given the complexities of most software and how little experience most users have in evaluating it, this is a tough job. But there are techniques that can be used to demonstrate that software-based systems are more trustworthy than they might otherwise be. While governments can play a watchdog role here, a TACT should not depend for its integrity on any single auditing scheme. Rather, it needs to be transparent to review and improvement by the general public.

In particular, communities and users should be able to audit the tools themselves (or find people whom they trust to audit them), and should be able to fix any problems they find without reliance on unaccountable suppliers.

Those steps include:

- Having clearly stated design goals and implementations, along with means of being held accountable to them.
- Using exclusively free/libre and open source components.
- Ensuring that the software can be reproducibly built from source.
- Having all published versions (including source) of all components permanently available in an archive for research.
- Ensuring that the channel used to provide software updates is reliable and auditable.

Sustainably maintained

Despite having an exit strategy, a responsible and effective TACT will likely be deployed for months if not years.

Over the time that a TACT is deployed, problems will likely be found and lessons learned — in the underlying protocols, in the software, in the deployed services, and in our understanding of the disease and its social impacts. Meanwhile, a successful TACT will be deployed on a large number of devices, making it an attractive target for compromise.

The tooling needs to be able to change and adapt to these circumstances, and it needs to have the resources to do so responsibly. It needs a sustainable maintenance plan. This includes adequate funding and support for:

- Community liaisons
- Public health professionals
- UI/UX designers
- Cryptographers
- Security researchers (internal and external, including system audits and bug bounties)
- Software developers
- System administrators

Without sustainable, ongoing maintenance for the life of the system, a TACT scheme may pose additional serious risks to all parties who participate in it.

Conclusion

We may well see the development of a privacy-preserving, technology-augmented contact-tracing system that makes a significant contribution to fighting the COVID-19 pandemic. But there are risks to fundamental civil liberties posed by poorly designed systems in this space, and a poorly designed system may be ineffective or even make the pandemic worse. We need a sober consideration of the risks and tradeoffs of such a system so that it protects not only the fundamental right to health, but also our rights of privacy and free association.

In the coming weeks and months, we will see a real push to reopen the economy, a push that will rely heavily on public health measures that include contact tracing. TACT proposals are likely to become a central part of the discussion. Location tracking and massive centralized surveillance should be off the table, but proximity tracking could be useful. When looking at these proposals, the questions outlined here should be front and center.