Saving Zoë: Worldline self-monitoring can better limit the spread of diseases

Jeremy Levy[1,2]*, Akash Levy[3]

[1]University of Pittsburgh

[2]Pittsburgh Quantum Institute

[3]Stanford University

*Correspondence to: jlevy@pitt.edu.

Abstract: The emergence of the novel coronavirus COVID-19 has sparked intense interest in the use of mobile phone technology to aid in contact tracing. We demonstrate, using a simple model of infectious transmission, that voluntary offering of a richer dataset—worldline information—can reveal transmission paths that are missed through contact tracing alone. This information can be captured securely by a mobile device and kept confidential, thus allowing individuals to weigh their personal privacy concerns against the benefits to society, and make their own decisions. Worldline self-monitoring has other benefits for the adopter that are unrelated to health concerns, which could aid in widespread adoption. Because each worldline contributes to a more comprehensive model of objective reality, independent platforms can contribute information without coordination, similar to the way that scientific knowledge is constructed.

Testing and contact tracing are twin pillars of public health strategies to slow the spread of infectious diseases. Traditionally, contact tracing has been conducted by interviewing the infected person, gathering information about whom they might have been in contact with, and contacting those who may have been infected, and so on, to encourage self-isolation. While this approach is effective in situations where the spread of infection is relatively slow, it becomes impractical with highly infectious novel viruses such as SARS-CoV-2 (*1*), where the unchecked doubling time can be as short as three days in densely populated metropolitan areas such as New York City (*2*). Under those circumstances, calls for society to "lock down" to interrupt spread over a period of time sufficient to reduce the number of infected people. While lockdown can "hammer" (*3*) down the number of infected people, it must be followed by a "dance", meaning, the institutionalization of societal protocols that incorporate testing and contact tracing in order to maintain the effective reproduction rate R_e below or near unity (*4*).

The manner in which confidential medical information is shared, and the subsequent impact on privacy, is a central concern that must be balanced with all approaches to limiting the spread of infectious disease. Embedded within the Hippocratic oath (*5*) is the promise to respect privacy in order to encourage forthrightness of patients to disclose sensitive information. Contact tracers face privacy quandaries every working day, and they are trained to make decisions consistent with the oath that they take.

Human-powered contact tracing, while highly effective for slow-moving diseases, is not scalable for global pandemics like the one associated with COVID-19. Fortunately, there is a potential for using information technology to track and control the spread of disease (*6*). The strategy used in China, for example, requires citizens to report their location using their mobile phone which tracks location using GPS data, and this monitoring is often supplanted by other techniques such as facial recognition tracking (Figure 2A). When an infection is detected, possible contacts are required to self-quarantine. This form of direct, continuous surveillance of essentially all citizens is not currently tolerated in Western countries, where right-to-privacy laws protect citizens against being forced to share their location or even information about their contacts with other people.

Most app-development efforts are focused on recording proximity between people, leveraging technologies commonly found in mobile phones. Bluetooth low-energy transceivers are present in nearly all modern smart phones. This two-way wireless protocol allows keys to be exchanged between two parties when their phones come within range for a period of time (usually 15-30 minutes), and record information about the proximal interaction on both phones. Within this class of mobile applications ("apps"), there are two types. Apps currently being developed by governments of the UK (*7*), Singapore (*8*), France (*9*), and Norway (*10*), employ a centralized server that can find the matches, and correlate them with GPS, cell phone tower records, and other location cues to help track the spread of the disease. A second "decentralized" approach to contact tracing forgoes the need to maintain a central server. Apple and Google have jointly created an open specification (*11*) that lets users report data anonymously, without the need for a centralized database.

Both the centralized and decentralized approaches are closely modeled after traditional methods of contact tracing. However, these approaches miss important routes to contagion that can potentially help prevent the spread of infectious disease.

Imagine a world which is inhabited by four people: Zoë, Alice, Bob, and Charlie. A realistic world is three-dimensional, but the dimensionality does not play a critical role in the analysis. For sake of visualization, we can assume that the world is one-dimensional, and location is parameterized by a displacement coordinate x. The evolving lives (*autobiography)* of Zoë, Alice, Bob, and Charlie is completely represented by their time-dependent coordinates (worldlines) $x_Z(t), x_A(t), x_B(t), x_C(t)$, and additional information that includes a *healthchart* which can be in one of four distinct states: "susceptible", "infected", "diagnosed", or "deceased". In one spatial dimension, we can represent displacement along a horizontal axis and advancing time pointing upwards (by convention). Paths that are sketched along this coordinate system can visually represent a person's timeline. A specific point along the timeline has a unique position and time (x, t) and is called an *event*, but the full state of a person at time t also includes the *healthchart* information. By convention, we will represent the *autobiography* of a human to be the set of points $\boldsymbol{H} = (x_H(t), h_H(t))$ where we have excluded information that is not pertinent to tracking disease.

Consider the spacetime diagram for the four inhabitants, illustrated in Figure 1A. The first key event is the one in which Bob becomes infected at $t = 2$, and is diagnosed at $t = 4$. Let us assume that all citizens are running a contact-tracing app that allows two people who cross paths (events for which $x_H(t) = x_{H'}(t)$) to share evidence of this event. Charlie intersects Bob at $t = 3$, infecting him. Bob is diagnosed at $t = 4$, and reports to the contacting tracing app which allows Charlie to learn about the exposure. Alice gets infected at $t = 3$ by sharing the space

($x = 5$) that Bob inhabited while asymptomatic. Alice is unaware of her infection until $t = 5$, when she becomes symptomatic and is diagnosed. Meanwhile, Zoë has crossed into Alice's contaminated areas, causing her to become infected. Assuming a world where all infection leads to death, everyone dies.

Let us now consider a second scenario, illustrated in Figure 1B, and shown graphically in Figure 2. In this world, Bob again gets infected at $t = 2$, infects Charlie at $t = 3$, and is diagnosed at $t = 4$. At this point, Bob elects to anonymously publish his worldline for the relevant time interval $1 < t < 4$ (when *health ≠ susceptible*). Charlie compares the anonymously published autobiographical information with his own, and deduces that he was infected. Alice also determines that that she inhabited a contaminated space and self-quarantines. Zoë, who often shares space with Alice, is able to determine that the published autobiography has not impacted her, and she is also able to actively avoid infection through knowledge of the infected areas. In this world, Zoë survives.

This simple example clearly shows a superior outcome from a traceback approach that shares autobiographical information, compared with contact tracing alone. Knowledge of contact between any two people is derivable from the worldlines. The additional information of a worldline helps to prevent incorrect inferences from being made, and also helps to protect the lives of people who never came into direct contact with an infected person. The more precisely one knows one's whereabouts, the more readily one can assess the potential for transmission.

In a more realistic scenario, we know that the contamination does not last indefinitely, and the survival time of the pathogen without a human host can depend on many factors (*12*). Regardless of the magnitude of the risks involved, it is clear that knowledge of these secondary paths to infection can help in reducing the spread of infectious disease. To make the idea more realistic, imagine that an infected person enters an empty subway car in New York City, touches a stainless-steel subway pole, and exits the train. At the next stop, another person enters the train, touches the same subway pole, and leaves. No contact-tracing app will be able to capture this route of infection. In the subway example, if both the infected and susceptible person knows exactly which subway pole they were holding onto, or where their hand was placed precisely relative to the infected person, the ability to estimate the risk of infection could be made even more precise. The subway car example also shows how convection can lead to highly nonlocal transmission, with a reach of hundreds of meters, assuming that the airborne virus particles are riding the subway cars undetected.

Revealing one's autobiographical information (i.e., *worldline* and *healthchart*) raises important privacy issues. However, there is a difference between the approach described here and the one being adopted by China. First, all information is being divulged voluntarily, for the greater good of society. There is a robust literature on privacy tradeoffs associated with revealing more or less information about one's worldline (*13*). This literature can help guide the development of how to counsel an infected person, and allow them to edit out or blur parts of the autobiography that they do not wish to reveal. K-anonymity methods can be used to estimate the number of people with a similar enough worldline so as to obfuscate a patient's identity (*14*). It is also very helpful to involve a third trusted party, like a medical professional, to certify the *worldline* and *healthchart* of a person. Otherwise, it would be possible to cause hysteria by posting false narratives, like bomb threats relayed through an untraceable source (*15*). Even well-intentioned alerts, if improperly executed, either induce panic or cause people to ignore them. The existence of accurate worldline information enables an alert to be highly specific, providing no more concern than is necessary. Additionally, the ability to localize sites visited by infected people allows health authorities to neutralize contaminated locations so that they are safe for public access again.

Mobile phones are powerful data acquisition and communication devices. They can be harnessed to acquire a multitude of cues about location. Each sensing "channel" may each have varying degrees of accuracy; they can be combined to develop highly accurate models about timelines. This approach is similar to how different sources of evidence are used to constrain theories of physics. At the root of the challenge is the fundamental assumption that there is an objective reality. A typical mobile phone can detect sound, magnetic field, acceleration, light (cameras), Bluetooth receivers, GPS receivers, in addition to cellular reception. A phone can transmit information securely using Bluetooth, WiFi, or through USB connections. (It can also communicate insecurely. We assume that the phone is not afflicted with spyware, or displaying messages that reveal unwanted information.) A worldline logging app can continuously acquire all of the sensor data, and only analyze the data that is needed when a worldline in the general neighborhood is reported.

How hard would it be to detect the infection route in the subway car discussed earlier? Not hard at all, in fact. Using a combination of inexpensive Bluetooth beacons installed in each NYC subway car, a passenger could automatically record which subway car they traveled in, every time they rode the subway. With multiple beacons installed, it is possible to obtain meter-level three-dimensional location information, especially if the person is standing relatively still from one subway stop to the next. For higher resolution, the back-facing camera, which may be facing the ground as the owner is scrolling through news, could record a unique a pattern of dots printed on the floor of the subway car. In this way, the precise sub-footprint location of a person can be recorded multiple times per second.

Fig. 3A shows state-of-the-art facial recognition being used at scale in China. This information is used to control the behavior of its citizens. Fig. 3B shows how pattern recognition of *things* can be used to capture worldlines and potentially protect citizens. The following illustration uses technology originally developed to locate the position of a pen on paper. A "smart pen" with an embedded camera processes 6x6 dot arrays offset in one of four directions. The arrangement is unique over an area $4^{36}a^2$, where a is the mean dot spacing ($a\sim 300\ \mu$m for the commercial product). Patterns that can precisely locate the position of the ball point on a virtual sheet of paper comparable to the surface area of the planet (*16*). A free mobile app enables more distant patterns to be geolocated. Fig. 3C shows four images taken by one of the authors, just before the coordinates are returned (Fig. 3D). The coordinates returned have the format "$X: m\ ; Y: n$", where m and n are integers. We can rewrite these integers $m = m_0 + dm$, $n = n_0 + dn$, where $(m_0, n_0) = (5636306, 385876371)$ corresponds to an *absolute* reference marker with respect to the earth, and (dm, dn) represents a relative coordinate on the page (origin at the center). Using the 8" square background tile as a reference, and referencing the cross-hair marker of the software scanner, values for four points returned are listed in Table 1. The accuracy of this position detector falls within one inch.

The ability to precisely track one's location information has many benefits as well. Consider another scenario in which passenger Diana enters an empty subway car, and inadvertently leaves behind a valuable package. Diana could anonymously leave a message on an electronic "billboard" for that subway car. Evan, who enters the car at the next stop, might have elected to automatically subscribe to nearby billboards, perhaps opting for some kinds of ads but not others, as well as lost-item/"good Samaritan" alerts. Being a good citizen, Evan could safely and anonymously return the package to Diana by arranging for it to be picked up by transit authorities at the next stop. Both passengers would feel good, for different reasons (*17*). These "feel-good" stories may encourage more widespread adoption of this technology, which increases the effectiveness of the collective health benefits (*18*). This technology, a "find-my-phone" feature on steroids, would have many positive aspects, rather than having a reputation as a "sickness app". For the same reason that having the app run on one person's phone can be beneficial to that one person, having the app run on a single person's phone can help as well. In fact, it is straightforward to see that the effectiveness of worldline tracing scales linearly with the fraction of adopters, while for contact tracing it scales quadratically. If 10% of the users adopt contact tracing, then only 1% of interactions are captured, whereas if 10% of the users adopt worldline tracing, 10% of infected worldlines will be shared. The decision to help can be directly traced back to the person with the best understanding of the consequences of that decision, and the most power to make a decision that benefits others.

The technology for worldline tracing closely maps the methodology of experimental science. There is data acquisition, analysis, and interpretation. Data is neither intrinsically good nor bad. However, in a free society, it is important to limit information "feedback loops" that constrain desirable freedom of its citizens. Fortunately, there are solutions that can work, in principle. It is up to the ingenuity of hardware and software developers to produce compelling technologies that will benefit society without unnecessarily sacrificing privacy, breaching confidentiality, or restricting freedom.

References and Notes:

1. https://en.wikipedia.org/wiki/Severe_acute_respiratory_syndrome_coronavirus_2
2. https://www.npr.org/sections/coronavirus-live-updates/2020/03/24/820891370/coronavirus-n-y-infection-rate-is-doubling-about-every-3-days-cuomo-says
3. https://medium.com/@tomaspueyo/coronavirus-the-hammer-and-the-dance-be9337092b56
4. https://en.wikipedia.org/wiki/Basic_reproduction_number
5. https://www.medicinenet.com/script/main/art.asp?articlekey=20909
6. L. Ferretti *et al.*, Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science* **368**, eabb6936 (2020).
7. https://github.com/nhsx
8. https://www.tracetogether.gov.sg
9. https://www.aphp.fr/actualite/covidom-une-solution-de-telesuivi-domicile-pour-les-patients-porteurs-ou-suspectes-covid
10. https://helsenorge.no/smittestopp
11. https://www.apple.com/covid19/contacttracing
12. https://apps.who.int/iris/bitstream/handle/10665/332071/WHO-2019-nCoV-HCW_risk_factors_protocol-2020.3-eng.pdf
13. C. Dwork, A. Roth, The Algorithmic Foundations of Differential Privacy. *Foundations and Trends® in Theoretical Computer Science* **9**, 211-407 (2014).
14. L. Sweeney, k-Anonymity: A Model for Protecting Privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **10**, 557-570 (2002).
15. https://en.wikipedia.org/wiki/2012_University_of_Pittsburgh_bomb_threats
16. M. P. Pettersson, P. Ericson, USPTO, Ed. (US, 2002).
17. M. McLure Wasko, S. Faraj, "It is what one does": why people participate and help others in electronic communities of practice. *The Journal of Strategic Information Systems* **9**, 155-173 (2000).
18. https://www.research.ox.ac.uk/Article/2020-04-16-digital-contact-tracing-can-slow-or-even-stop-coronavirus-transmission-and-ease-us-out-of-lockdown
19. https://www.nytimes.com/2018/07/08/business/china-surveillance-technology.html
20. https://play.google.com/store/apps/details?id=com.anoto.adna

Acknowledgments: J.L. acknowledges support from a Vannevar Bush Faculty Fellowship (ONR N00014-15-1-2847). A.L. acknowledges support from a National Science Foundation Graduate Research Fellowship Program. **Funding:** include complete funding information, including grant numbers; **Author contributions:** J.L. conceived and wrote the manuscript, and performed the experiments shown in Figure 2B-E. A.L. provided a key insights related to validity of autobiographical information with respect to health status, and to the importance of differential privacy in deciding how best to sanitize autobiographical information prior to publication.; **Competing interests:** Authors declare no competing interests; and **Data and materials availability:** All data is available in the main text or the supplementary materials.

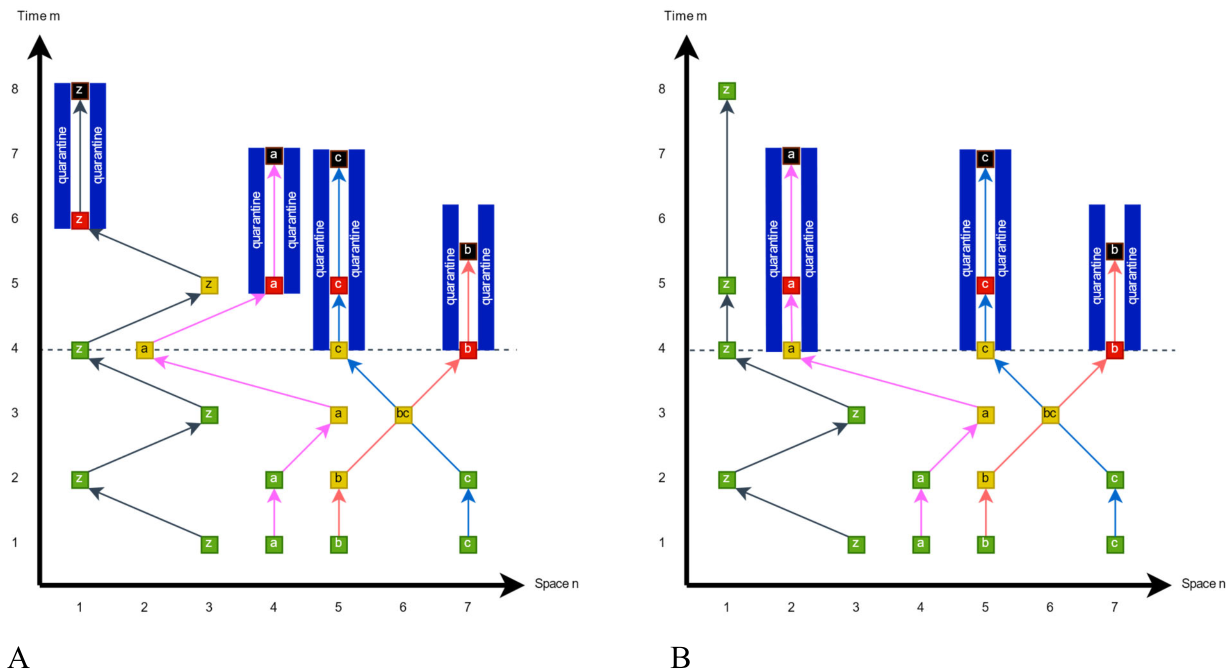

Fig. 1. Contact tracing versus worldline tracing. A simplified one-dimensional world has four inhabitants: Zoë, Alice, Bob, and Charlie. People are distinguished by their first-name initial. Their health is in one of four states: susceptible (green), infected (yellow), diagnosed (red), and deceased (black). Transmission can occur by direct contact, i.e., crossing paths with an infected person, or by occupying the space previously inhabited by an infected person. Self-quarantining or hospitalization is represented by the surrounding red walls representing a building. (A) In this world, Bob is infected at $t = 2$, who later infects Charlie at $t = 3$. The proximity sensor triggers an exchange of keys on a contact-tracing app. At $t = 4$, Bob is diagnosed, and the app publishes the matched pair of keys, informing Charlie anonymously about the infection so that Charlie can self-quarantine. Alice is also infected by Bob's trail at $t = 3$, and Zoë becomes infected between $t = 4$ and $t = 5$. Because the infection is 100% fatal, everybody dies. (B) In this world, Bob is infected at $t = 2$, infects Charlie at $t = 3$, and is diagnosed at $t = 4$. Bob publishes his worldline from $1 \leq t \leq 4$. Charlie finds a match, and self-quarantines. Alice sees Bob's and Charlie's worldline, and determines that she was infected by Bob, and not Charlie. Alice self-quarantines, saving Zoë from infection.

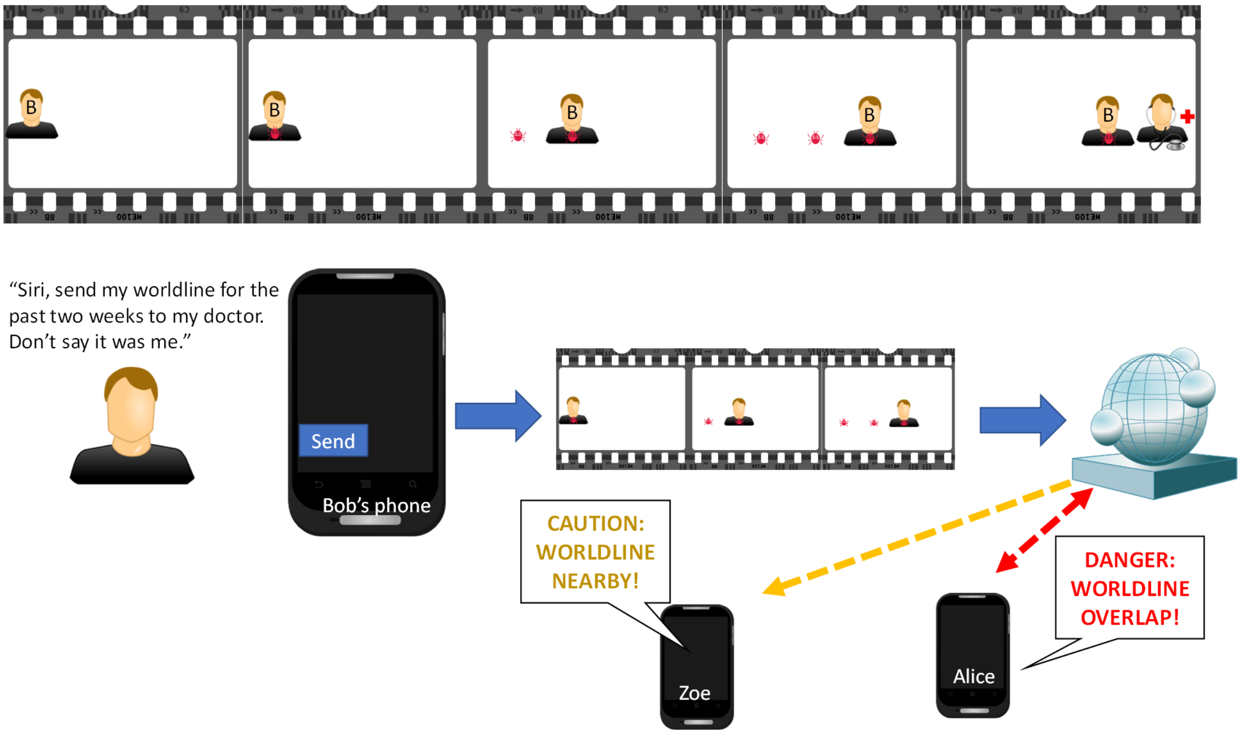

Figure 2. Illustration of interactions between Bob and his doctor, illustrating the choice he makes to send his worldline without his name. Alice has crossed paths with Bob and reports her timeline in a manner similar to Bob (process not shown). Zoe learns of the nearby areas contaminated by Bob and Alice and avoids them.

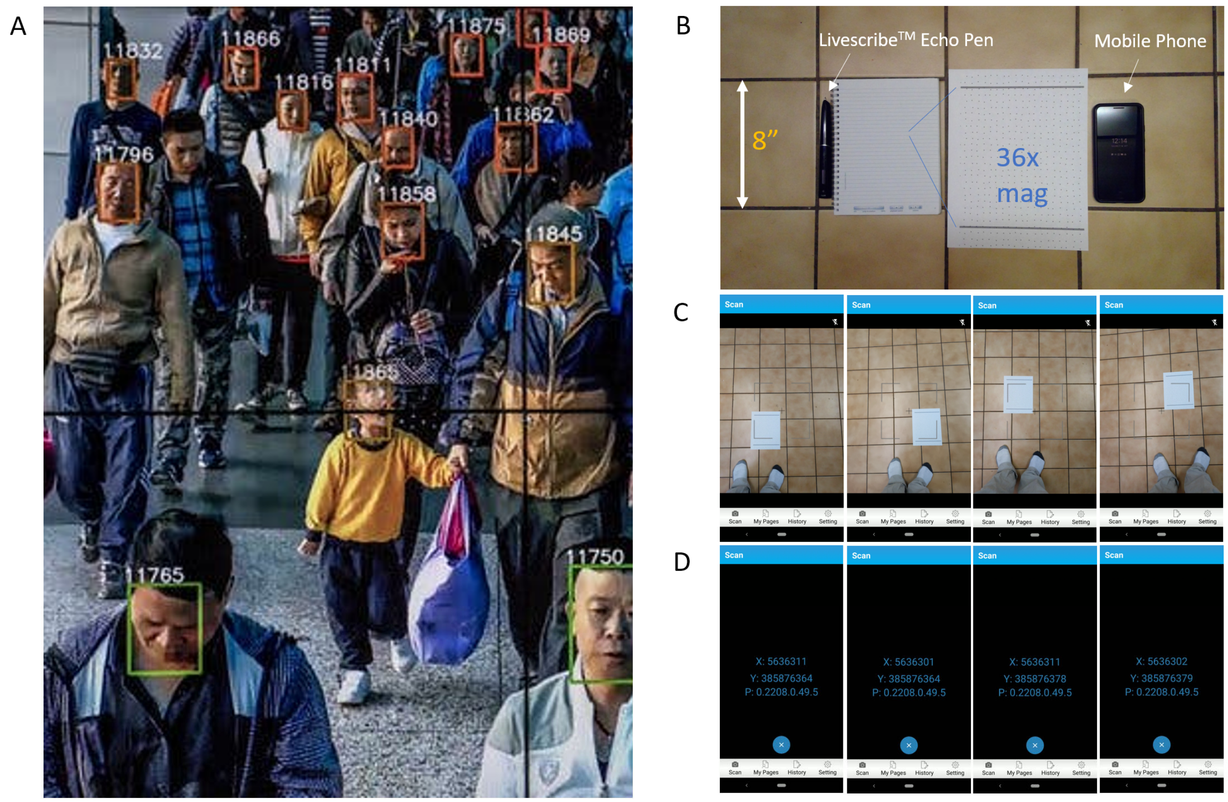

Figure 3. Precise geolocating of the masses versus precise geolocating for the masses. A. Face tracking in China (photo adapted from Ref. (*19*)). B. Visual positioning technology. Left-to-right: a pen that can stream its location on lightly patterned paper,within the diameter of its ball point at 70 Hz; notebook used with pen; 36x magnification of printed sheets; Android mobile phone capable of running a free mobile app (Anoto aDNA (*20*)). C. Four frames of aDNA Scan app at four edges of printed sheet. D. Integer cartesian coordinates returned by the app. One-inch resolution from waist-high distances is readily achievable at high frame rates using ubiquitous mobile phone technology. Still frames of C, D taken from a video captured by an Android phone running the aDNA app: https://youtu.be/9Bc0PBYm5-8 .

Table 1. *Table of absolute* (X, Y) *and relative* (dx, dy) *coordinates returned by mobile app, reporting the location of four corners of the sheet of paper on Earth's surface with a resolution of 1".*

X	Y	dx	dy
5636311	385876364	5	-7
5636301	385876364	-5	-7
5636311	385876378	5	7
5636302	385876379	-4	8