

MISSING THE FOREST FOR THE TREES

HOW WE MEASURE AND LOSE TRACK OF WHAT MATTERS

Ali Alkhatib

May 18, 2020

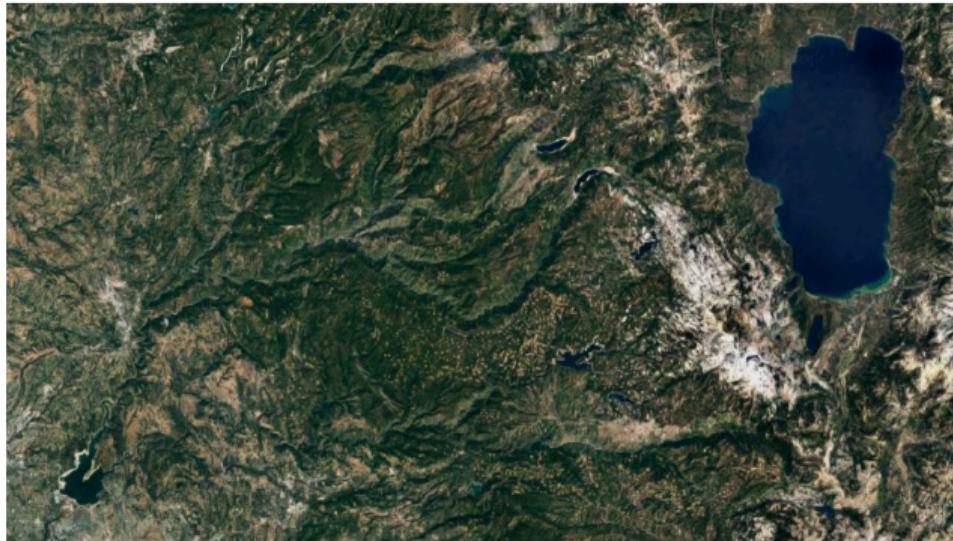
University of San Francisco Center for Applied Data Ethics



WE INTERPRET THE WORLD IN CONSTRAINED WAYS

**WE INTERPRET THE WORLD IN CONSTRAINED WAYS
IT'S HOW WE COPE WITH TOO MUCH INFORMATION**

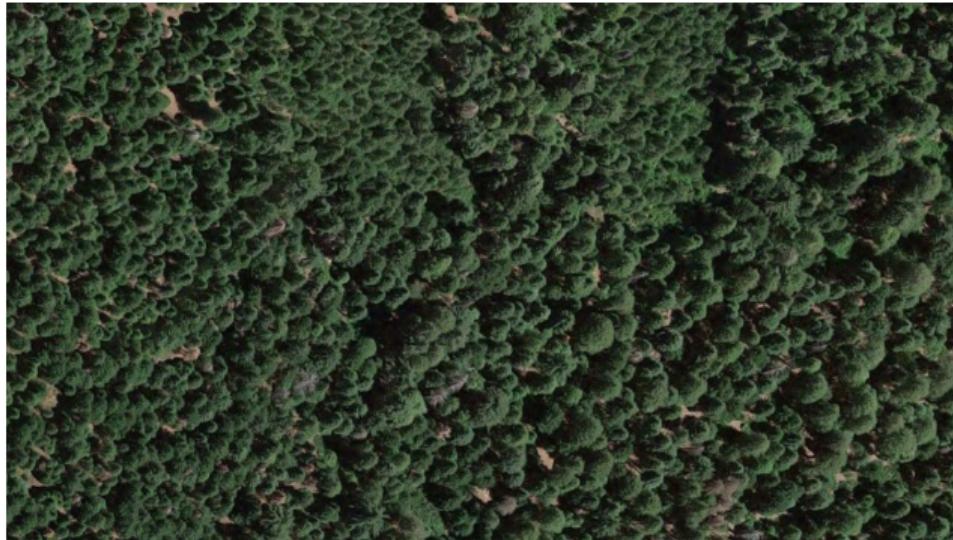
SCIENTIFIC FORESTRY



SCIENTIFIC FORESTRY



SCIENTIFIC FORESTRY



SCIENTIFIC FORESTRY



SCIENTIFIC FORESTRY



SCIENTIFIC FORESTRY



SCIENTIFIC FORESTRY



DELAYED DISASTERS

- Things were fine for a while

DELAYED DISASTERS

- Things were fine for a while
- Then they weren't

DELAYED DISASTERS

- Things were fine for a while
- Then they weren't

DELAYED DISASTERS

- Things were fine for a while
- Then they weren't

Pests

Storms

Hollowing out of the ecosystem

DELAYED DISASTERS

- Things were fine for a while
- Then they weren't

Pests

Storms

Hollowing out of the ecosystem

DELAYED DISASTERS

- Things were fine for a while
- Then they weren't

Pests

Storms

Hollowing out of the ecosystem

DELAYED DISASTERS

- Things were fine for a while
- Then they weren't

Pests

Storms

Hollowing out of the ecosystem

DELAYED DISASTERS

- Things were fine for a while
- Then they weren't
 - Pests
 - Storms
 - Hollowing out of the ecosystem

Waldsterben

DELAYED DISASTERS

- Things were fine for a while
- Then they weren't

Pests

Storms

Hollowing out of the ecosystem

Waldsterben

Forest death

ECOLOGICAL DIVERSITY

The forest was more than just the trees

ECOLOGICAL DIVERSITY

The forest was more than just the trees

It was everything that lived in and passed through

HOLLOWING OUT A SOCIAL ECOLOGY

DANGERS OF REDUCTIVE ALGORITHMS

“street-level algorithms” struggle with novel situations

DANGERS OF REDUCTIVE ALGORITHMS

“street-level algorithms” struggle with novel situations
how are they changing **our lives** both offline and online?

DANGERS OF REDUCTIVE ALGORITHMS

“street-level algorithms” struggle with novel situations
how are they changing our lives both offline and online?
...and how **will** they?

11:56 ↗



allure.com

≡ allure

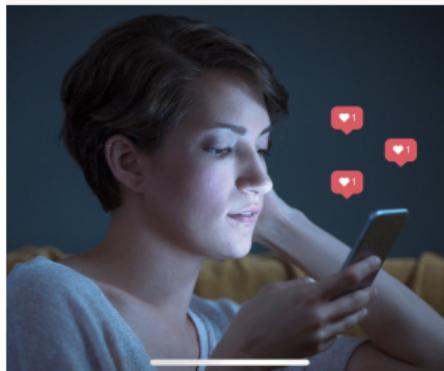
SUBSCRIBE

Why Finstas Are a Safe Sanctuary on Social Media

"I treat my finsta like a visual journal of my personal life."

BY ELLIE ABRAHAM

July 23, 2019



11:56 ↗
allure.com

≡ allure

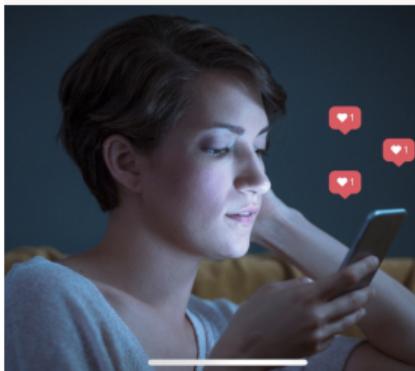
SUBSCRIBE

Why Finstas Are a Safe Sanctuary on Social Media

"I treat my finsta like a visual journal of my personal life."

BY ELLIE ABRAHAM

July 23, 2019



11:57 ↗
cnet.com

Teens have figured out how to mess with Instagram's tracking algorithm

Teenagers are using group accounts to flood Instagram with random user data that can't be tied to a single person.



Alfred Ng

Feb. 4, 2020 6:06 a.m. PT



2



Teens have been using group accounts on Instagram to feed randomized data to the social network and protect their privacy.

Alfred Ng / CNET

Like about a billion other people, 17-year-old Samantha Mosley spent her Saturday afternoon perusing Instagram.

11:56 ⓘ allure.com

≡ allure

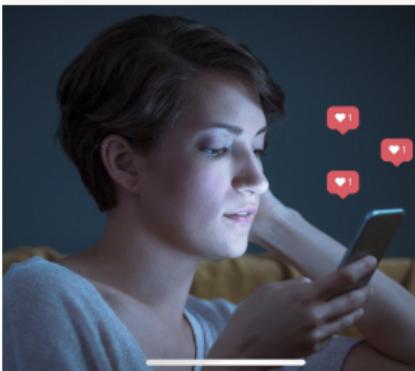
SUBSCRIBE

Why Finstas Are a Safe Sanctuary on Social Media

"I treat my finsta like a visual journal of my personal life."

BY ELLIE ABRAHAM

July 23, 2019



11:57 ⓘ cnet.com

Teens have figured out how to mess with Instagram's tracking algorithm

Teenagers are using group accounts to flood Instagram with random user data that can't be tied to a single person.



Alfred Ng ⓘ

Feb. 4, 2020 6:06 a.m. PT



Teens have been using group accounts on Instagram to feed randomized data to the social network and protect their privacy.

Alfred Ng / CNET

Like about a billion other people, 17-year-old Samantha Mosley spent her Saturday afternoon perusing Instagram.

12:02 ⓘ google.com

AA ⓘ urbandictionary.com



Type any word...



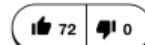
TOP DEFINITION

Backstreet Boys Reunion Tour

In the wake of the CO-VID 19 pandemic, which occurred in the spring of 2020, the Youtube gods attempted to **crack down** on the hysteria by demonetizing content that contained the word 'Corona', 'Virus' or '19' in any variation, especially in that specific succession. So in an effort to avoid divine punishment, humanity made the collective decision (headed by **Dan Avidan** of the Game **Grumps**) to henceforth refer to the virus only as the 'Backstreet Boys Reunion Tour'.

Seven people in my town died after attending the **Backstreet Boys reunion tour**.

by Marsakali April 02, 2020



11:56 ⓘ allure.com

≡ allure

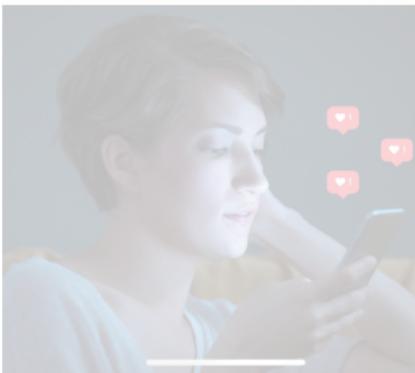
SUBSCRIBE

Why Finstas Are a Safe Sanctuary on Social Media

"I treat my finsta like a visual journal of my personal life."

BY ELLIE ABRAHAM

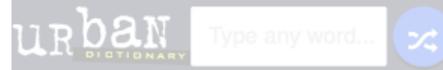
July 23, 2019



11:57 ⓘ cnet.com

12:02 ⓘ google.com

AA urbandictionary.com



TOP DEFINITION

Backstreet Boys Reunion Tour

In the wake of the CO-VID 19 pandemic, which occurred in the spring of 2020, the Youtube gods attempted to **crack down** on the hysteria by demonetizing content that contained the word 'Corona', 'Virus' or '19' in any variation, especially in that specific succession. So in an effort to avoid divine punishment, humanity made the collective decision (headed by **Dan Avidan** of the Game **Grumps**) to henceforth refer to the virus only as the 'Backstreet Boys Reunion Tour'.



Teens have been using group accounts on Instagram to feed randomized data to the social network and protect their privacy.

Alfred Ng / CNET

Like about a billion other people, 17-year-old Samantha Mosley spent her Saturday afternoon perusing Instagram.



allure.com 11:57 🔍

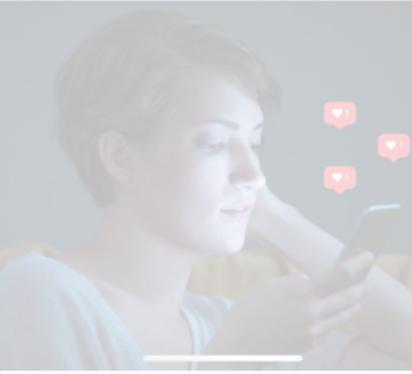
allure

SUBSCRIBE

Why Finstas Are a Safe Sanctuary on Social Media

"I treat my finsta like a visual journal of my personal life."

BY ELLIE ABRAHAM July 23, 2019



f **✉** **P** Digital Contact Tracing

Teens have been using group accounts on Instagram to feed randomized data to the social network and protect their privacy.

Alfred Ng / CNET

Like about a billion other people, 17-year-old Samantha Mosley spent her Saturday afternoon perusing Instagram.

cnet.com AA google.com

Teens have figured out how to mess with Instagram's tracking algorithm

Teenagers are using group accounts to flood Instagram with random user data that can't be tied to a single person.

Alfred Ng Feb. 4, 2020 6:06 a.m. PT



Teens have been using group accounts on Instagram to feed randomized data to the social network and protect their privacy.

Alfred Ng / CNET

Like about a billion other people, 17-year-old Samantha Mosley spent her Saturday afternoon perusing Instagram.

urban dictionary.com

urban

Type any word... X

TOP DEFINITION

Backstreet Boys Reunion Tour

In the wake of the CO-VID 19 pandemic, which occurred in the spring of 2020, the Youtube gods attempted to **crack down** on the hysteria by demonetizing content that contained the word 'Corona', 'Virus' or '19' in any variation, especially in the music genre. So in an effort to avoid punishment, humanity made the collective decision (headed by **Dan Avidan** of the Game Grumps) to henceforth refer to the virus only as the 'Backstreet Boys Reunion Tour'.

Seven people in my town died after attending the Backstreet Boys reunion tour.

by Marsakali April 02, 2020

72 0

9:58



AA

ali-alkhatib.com



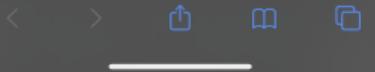
Ali Alkhatib

01 MAY 2020

We Need to Talk About Digital Contact Tracing

Over the past few weeks, Apple & Google have floated the idea of developing and distributing a digital contact-tracing app that will inform people when they've been exposed to someone who's contracted COVID-19, and communicate to people that they've been exposed to *you* if you later test positive yourself (edit: since writing this, Apple has released a beta of iOS 13 that includes the SDK necessary to begin using this system). Writing this in late April and early May, it feels like we're desperate for information and weary from not knowing who's caught COVID-19, who's still vulnerable, who gets it worse or why, or even how to treat it. We're desperate for any information we can get our hands on. This

proposal by Apple and Google promises to give



9:58

AA

ali-alkhatib.cc

Ali Alkhatib

01 MAY 2020

We Need to About Digital Contact Trac

Over the past few weeks, Apple floated the idea of developing a digital contact-tracing app that would let you know if you've been exposed to someone who has contracted COVID-19, and come into close proximity with them. If you later test positive yourself (edit: this, Apple has released a beta version of the app that includes the SDK necessary to build it), your phone would send out signals to other phones nearby that have also been exposed. Writing this in late April, it feels like we're desperate for information about who's been exposed to COVID-19, who's still vulnerable, who's been treated, why, or even how to treat it. We can't get any information we can get our hands on from the government or the media, so we're relying on Apple and Google to provide us with the information we need.

What this scheme will do (more or less, at the time of writing)

The simplest, most direct way to implement digital contact tracing would be to track where you are at any given moment. If you spend 20 minutes at a coffee shop, somewhere there'll be a record that **you** were at **Starbucks** from **12:00pm-12:20pm**. Somewhere, theoretically, there would be records for everyone participating in this scheme, such that if someone wanted to ask the system "who was at Starbucks at 12:10pm?", a list would come back with your name on it. To be honest, Google probably already has the means to produce a list like this from Google Maps. You can even request a record of all the places Google Maps knows you've been over a certain timeline.

Instead of recording **where** you are the way Google Maps does or other location-tracking apps you're familiar with, under this scheme your phone records **who** you're near. To illustrate what this means, let's make up and walk through an example.

Under this program, your phone would send out a signal every 5 minutes with a unique word. So from 12pm to 1pm, your phone would send out messages like this:

The image consists of three side-by-side screenshots of a mobile web browser displaying a blog post. The top bar of each screenshot shows the time as 9:58, signal strength, battery level, and the URL ali-alkhatib.com.

Screenshot 1: The title "Ali Alkhatib" and the date "01 MAY 2020" are visible. The main heading is "We Need to About Digital Contact Trac". Below it, a large block of text describes the concept of digital contact tracing, mentioning Apple's idea and how it would work. A small note at the bottom right says "powered by Apple and Google".

Screenshot 2: The title "What this schen do (more or less time of writing)" is displayed. Below it, a list of items with timestamps: "12:00: Apple", "12:05: Banana", "12:10: Orange", "12:15: Kiwi", "12:20: Potato", "12:25: Cherry", "12:30: Tomato", "12:35: Pear", "12:40: Coconut", "12:45: Orange", "12:50: Peach", "12:55: Eggplant", and "13:00: Kumquat".

Screenshot 3: The text continues with "Instead of recording where you Google Maps does or other loca apps you're familiar with, under your phone records who you're illustrate what this means, let's i walk through an example." It then provides an example: "Under this program, your phone a signal every 5 minutes with a from 12pm to 1pm, your phone messages like this:" followed by the timestamped list of items.



In this system, no other phone is sending out these exact words; if anyone ever hears **Kumquat**, it's because their phone was near your phone at 1pm. If my phone has **Kumquat** in its archive, it means that I was near you at 1pm.

Under this program, your phone a signal every 5 minutes with a from 12pm to 1pm, your phone messages like this:

I let's say I was in line behind you at a coffee

9:58

AA

ali-alkhatib.co

Ali Alkhatib

01 MAY 2020

We Need to Talk About Digital Contact Tracing

Over the past few weeks, Apple floated the idea of developing a digital contact-tracing app that would let you know if you've been exposed to someone who has contracted COVID-19, and communicate with people that they've been exposed to later test positive yourself (edit: this, Apple has released a beta includes the SDK necessary to build this system). Writing this in late April it feels like we're desperate for information we can get our weary from not knowing who's contracted COVID-19, who's still vulnerable, who's why, or even how to treat it. We can get our information we can get our from Apple and Google.

What this scheme does (more or less) is time of writing)

The simplest, most direct way to digital contact tracing would be: you are at any given moment. If minutes at a coffee shop, someone makes a record that you were at Starbucks between 12:00pm-12:20pm. Somewhere there would be records for everyone participating in this scheme, such as someone wanted to ask the sys Starbucks at 12:10pm?", a list would be made with your name on it. To be honest, probably already has the means list like this from Google Maps. request a record of all the places you've been over a certain period of time. Writing this in late April it feels like we're desperate for information we can get our weary from not knowing who's contracted COVID-19, who's still vulnerable, who's why, or even how to treat it. We can get our information we can get our from Apple and Google.

Instead of recording where you Google Maps does or other location apps you're familiar with, under your phone records who you've been in contact with. Let's illustrate what this means, let's walk through an example.

Under this program, your phone sends a signal every 5 minutes with a timestamp from 12pm to 1pm, your phone messages like this:

9:58

AA

ali-alkhatib.co

messages like this:

12:00: Apple

12:05: Banana

12:10: Orange

12:15: Kiwi

12:20: Potato

12:25: Cherry

12:30: Tomato

12:35: Pear

12:40: Coconut

12:45: Orange

12:50: Peach

12:55: Eggplant

13:00: Kumquat

In this system, no other phone has these exact words; if anyone ever says **Kumquat**, it's because their phone is in your phone at 1pm. If my phone is in your phone at 1pm, it means that I was in your phone at 1pm. Let's say I was in line behind you messages like this:

9:58

AA

ali-alkhatib.co

messages like this:

Quoting Arundhati Roy, "If we were sleepwalking into a surveillance state, now we are running toward it because of the fear that is being cultivated". We're gearing up to sacrifice substantial amounts of our privacy and anonymity in the world in exchange for a woefully dubious solution to our uncertainty.

Let's talk about why: first, this system isn't going to give us a better sense of the world in the most crucial places; and second, this system is going to undermine our privacy and dignity in ways that we can only begin to imagine right now.

Let's start with why this system won't work.

Digital contact tracing will exclude the poor, children, and myriad other uncounted groups

In the description I made earlier, I pointed out that Apple & Google's plan calls for some sort of proximity detection, probably involving Bluetooth and some of the fancier, newer technology that your smartphone may have if it was introduced in the last few years. Some of the technical requirements of this system preclude older devices from working, meaning the people with older smartphones won't be able to benefit from alerts to this system unless they upgrade to newer phones - something they almost certainly can't afford to do, nor should they be obliged

9:58

AA

ali-alkhatib.cc

Ali Alkhatib

01 MAY 2020

We Need to Talk About Digital Contact Tracing

Over the past few weeks, Apple floated the idea of developing a digital contact-tracing app that would let you know if you've been exposed to someone who has contracted COVID-19, and companies like Google have followed suit. Writing this in late April, it feels like we're desperate for information about what's going on. We're weary from not knowing who's been exposed to COVID-19, who's still vulnerable, who gets sick, why, or even how to treat it. We want to know if there's any information we can get our hands on.

Instead of recording **where** you've been, this system records **when**. Instead of Google Maps or other location-based apps you're familiar with, under your phone records **who** you've been near. To illustrate what this means, let's walk through an example.

Under this program, your phone sends a signal every 5 minutes with a timestamp from 12pm to 1pm, your phone messages like this:

9:58

ali-alkhatib.com

What this scheme does (more or less, time of writing)

The simplest, most direct way to participate in digital contact tracing would be to leave a record that you were at Starbucks between 12:00pm-12:20pm. Somewhere there would be records for everyone participating in this scheme, such as someone wanted to ask the sysadmin "Was I at Starbucks at 12:10pm?", a list would be generated with your name on it. To be honest, probably already has the means to do this from Google Maps. Request a record of all the places you've been over a certain period of time.

Instead of recording **where** you've been, this system records **when**. Instead of Google Maps or other location-based apps you're familiar with, under your phone records **who** you've been near. To illustrate what this means, let's walk through an example.

Under this program, your phone sends a signal every 5 minutes with a timestamp from 12pm to 1pm, your phone messages like this:

9:58

ali-alkhatib.com

messages like this:

- 12:00: Apple
- 12:05: Banana
- 12:10: Orange
- 12:15: Kiwi
- 12:20: Potato
- 12:25: Cherry
- 12:30: Tomato
- 12:35: Pear
- 12:40: Coconut
- 12:45: Orange
- 12:50: Peach
- 12:55: Eggplant
- 13:00: Kumquat

In this system, no other phone uses these exact words; if anyone ever says "I was in Starbucks at 1pm", it's because their phone is in your phone at 1pm. If my phone is in your phone at 1pm, it means that I was in Starbucks at 1pm.

Let's say I was in line behind you

9:59

ali-alkhatib.com

The spaces and the people we should be most concerned for, and paying the most attention to, won't show up in a digital contact tracing system like any of the ones proposed so far.

Digital contact tracing staggeringly misses the point of care that we should be most concerned with.

This system's exclusions will decisively undermine its accuracy and endanger everyone

The risks of acting on bad data - the result of excluding the poor, excluding children, and excluding other high-risk groups - are unspeakably high. This is different from working with a small sample size or even from working with no data. Building a dataset that excludes entire categories of people - as we're beginning to do with Native American COVID-19 patients - skews our vision of the world in ways we won't be able to account for, and with confidence.

That's unearned, ultimately steamrolling the groups that we leave out. This is why the CDC employs "sentinel surveillance" to track the flu - because acting only on the data we get from people who are wealthy enough to go to a doctor for the flu would give them a *dangerously misleading* picture of its spread every year. Everyone needs access to the same level of data, with no exclusions or user requirements.

9:58

AA

ali-alkhatib.co

Ali Alkhatib

01 MAY 2020

We Need to Talk About Digital Contact Tracing

Over the past few weeks, Apple floated the idea of developing a digital contact-tracing app that would let you know if you are at any given moment. If you were at a coffee shop, some minutes later a record that you were at Starbucks from 12:00pm-12:20pm. Somewhere there would be records for everyone participating in this scheme, such as someone wanted to ask the system "Was I at Starbucks at 12:10pm?", a list would be generated with your name on it. To be honest, probably already has the means to do this from Google Maps. Request a record of all the places you've been over a certain period of time.

Instead of recording where you have been, Google Maps does or other location apps you're familiar with, under your phone records who you've been near. To illustrate what this means, let's walk through an example.

Under this program, your phone sends a signal every 5 minutes with a timestamp from 12pm to 1pm, your phone messages like this:



9:58

ali-alkhatib.com

messages like this:

- 12:00: Apple
- 12:05: Banana
- 12:10: Orange
- 12:15: Kiwi
- 12:20: Potato
- 12:25: Cherry
- 12:30: Tomato
- 12:35: Pear
- 12:40: Coconut
- 12:45: Orange
- 12:50: Peach
- 12:55: Eggplant
- 13:00: Kumquat

In this system, no other phone uses these exact words; if anyone ever says "I was in line behind you at Kumquat", it's because their phone sent a signal from your phone at 1pm. If my phone is in its archive, it means that I was in line behind you at Kumquat at 1pm.

Let's say I was in line behind you at Kumquat at 1pm.

9:59

ali-alkhatib.com

Quoting Arundhati Roy, "If we went into a surveillance state, now we would be concerned for, and paying the price for, the people we won't show up in a digital contact like any of the ones proposed so far."

Digital contact tracing staggers the point of care that we should be with.

This system's exclusivity decisively undermines accuracy and endangers everyone

The risks of acting on bad data - excluding the poor, excluding children, excluding other high-risk groups - are unspeakably high. This is different from a small sample size or even no data. Building a dataset of entire categories of people - as to do with Native American COVID-19 cases - skews our vision of the world in ways that we can only begin to imagine.

Digital contact tracing excludes the poor, children, myriad other uncontrollable factors.

In the description I made earlier, that Apple & Google's plan calls proximity detection, probably involving GPS and some of the fancier, newer location technologies. Your smartphone may have it, in the last few years. Some of the requirements of this system prevent devices from working, meaning older smartphones won't be able to connect to this system unless they have newer phones - something they can't afford to do, nor should they be obliged to do.

9:59

ali-alkhatib.com

The spaces and the people we're concerned for, and paying the price for, the people we won't show up in a digital contact like any of the ones proposed so far.

Digital contact tracing systems that render the world as normally distributed space with spheres of influence and contact characterized by radio waves will consistently leave us with dangerously wrong pictures of our exposure.

Proximal tracing doesn't maintain privacy the way advocates seem to think it does

We also need to talk about what proximal contact tracing offers and what it doesn't, especially since we're adopting this contrived system to avoid "absolute" locations. I said earlier how absolute location tracking might say that we were both at some GPS coordinates. Or it might offer a street address. In either case, it would reveal that we were both at some semantically meaningful place from 8pm to 10pm, which could potentially be embarrassing or even damaging. Proximal tracking promises instead to reveal only that we were together between 8pm and 10pm.

The problem is that neither of these approaches is particularly helpful to me when I'm trying to maintain privacy about where I *meaningfully* am. While there are unique risks associated with my absolute location being revealed (for instance,

9:58 ↗ 9:58 ↗

Ali Alkhateeb @_alialkhateeb · May 1
I wrote a post - "We Need to Talk About Digital Contact Tracing".

We Need to Talk About Digital Contact Tracing
ali-alikhateeb.com

7 43 132

Ali Alkhateeb @_alialkhateeb · May 1
All I don't wanna do is zoom... · May 1

If you have at all been following the digital contact tracing discussions, or even if you are new to them, you really should read this by @_alialkhateeb

We Need to Talk About Digital Contact Tracing
ali-alikhateeb.com

5 64 98

Show this thread

12:25: Cherry most crucial places, and second going to undermine our privacy ways that we can only begin to i

12:30: Tomato

12:35: Pea Anna Lauren Hoffmann @annaeloverday

"It didn't help that their exact locations were obscured; they were 'nearby'. They were outed by their associations, not by their coordinates."

12:40: Coo

12:45: Orange

12:50: Pea This piece, from @_alialkhateeb 🌈

We Need to Talk About Digital Contact Tracing Over the past few weeks, Apple & Google have floated the idea of developing and distributing a digital contact-tracing ...
ali-alikhateeb.com

10:06 PM · May 1, 2020 · TweetDeck

In this system, these exact words your phone at 1pm. If my phone in its archive, it means that I was 1pm.

Kumquat, it's c your phone at 1pm. If my phone in its archive, it means that I was 1pm.

20 Retweets 58 Likes

It's a signal every 5 minutes with a from 12pm to 1pm, your phone messages like this:

Eric Topol @EricTopol

For much more in-depth dissection of the unaddressed problems with digital contact tracing, I think these are the 2 best:
ali-alikhateeb.com/blog/digital-c...
by @_alialkhateeb
brookings.edu/techstream/ina... @BrookingsInst by @ashk4n @rcalo @CT_Bergstrom

Private Kit

Contact-tracing apps are not a solution to the COVID-19 crisis Although Apple and Google's proposals for voluntary, anonymous contact tracing apps have been well received by some, they will likely serve as vehicle... brookings.edu

12:33 PM · May 2, 2020 · Twitter Web App

29 Retweets 53 Likes

would reveal that we were both at some

Kate Crawford @katecra... · May 2

Everyone is RTing this piece by @_alialkhateeb on contract tracing for good reason. It lays out the most serious structural concerns. Read it if you want a primer on the underlying problems that many of the 'tech solutions' don't address.

We Need to Talk About Digital Contact Tracing
ali-alikhateeb.com

4 121 239

WHAT DO WE MEASURE

**WHAT DO WE MEASURE
AND WHAT DOES IT MEAN WE OVERLOOK**