MOBILE DEVICE TRACKING

FTC: Spring Privacy Series
Washington, DC
Feb 19 2014

whoami

twitter: @ashk4n
ashkan.soltani@gmail.com
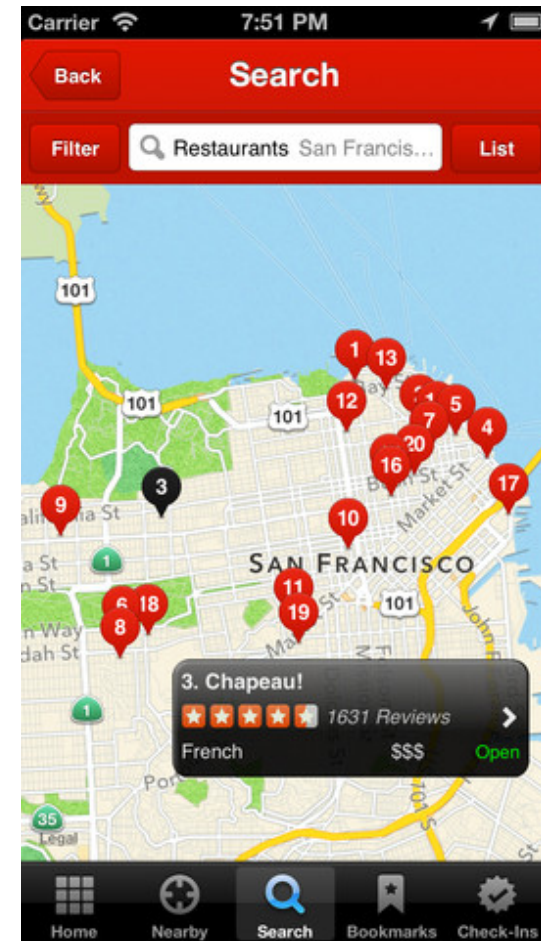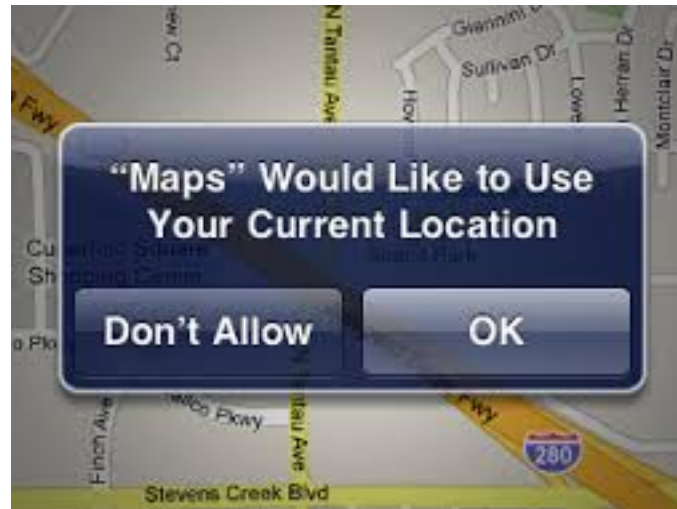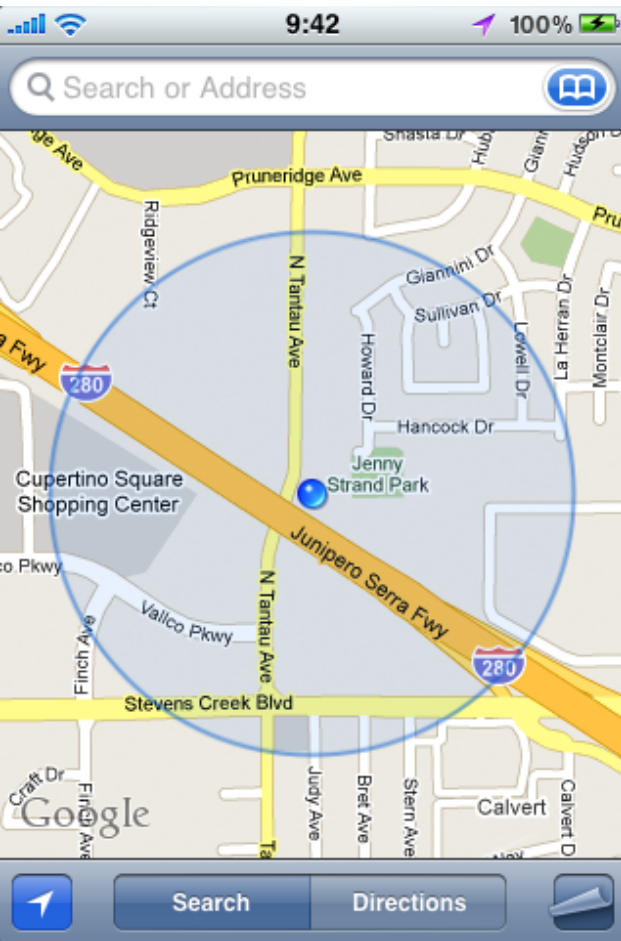independent researcher & consultant

mobile device tracking

- location aware devices
- device aware locations
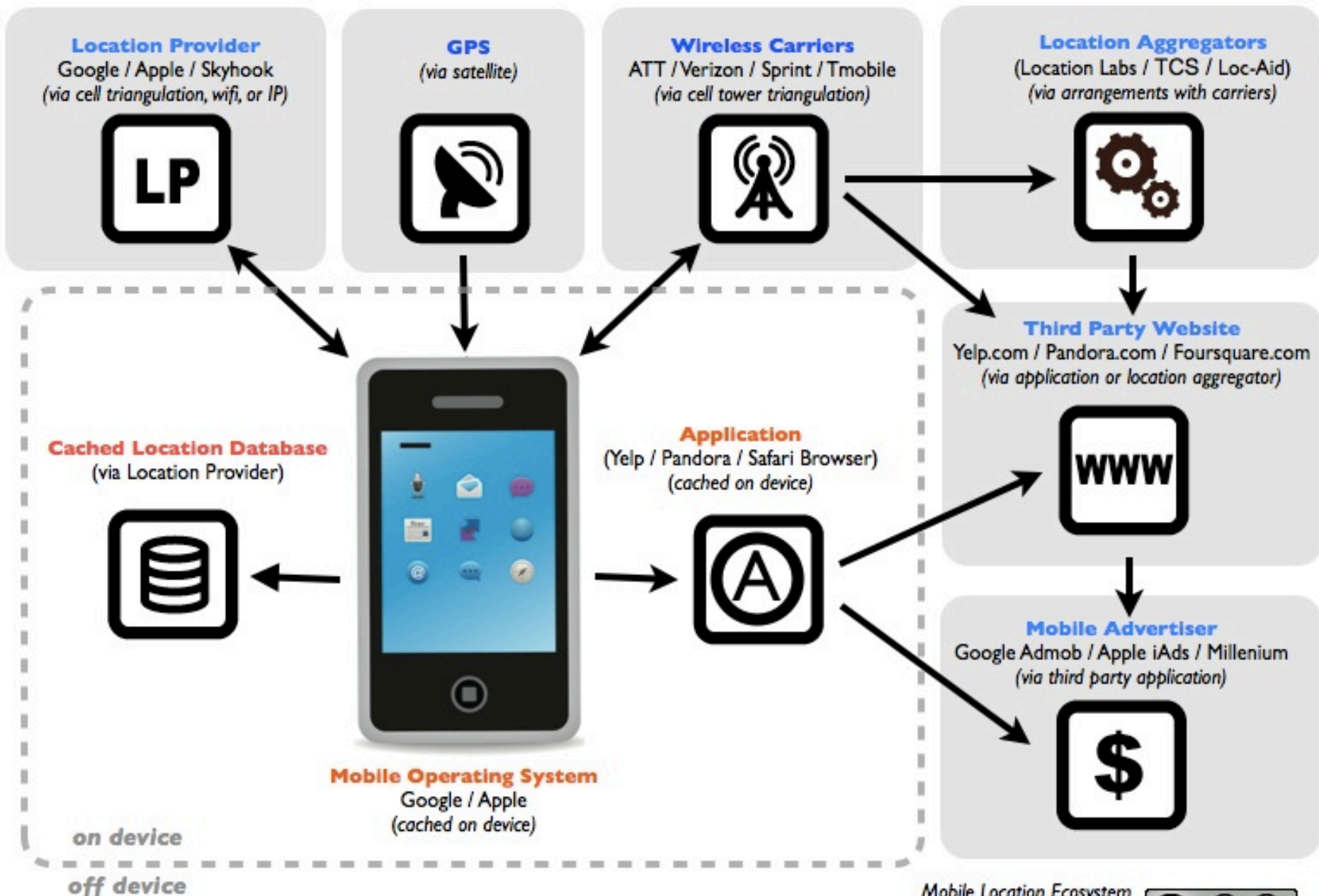- benefits
- concerns

mobile phones*

location aware devices

location aware devices

*the device tracks its location

Location Provider
Google / Apple / Skyhook
(via cell triangulation, wifi, or IP)

LP

GPS
(via satellite)

Wireless Carriers
ATT / Verizon / Sprint / Tmobile
(via cell tower triangulation)

Location Aggregators
(Location Labs / TCS / Loc-Aid)
(via arrangements with carriers)

Cached Location Database
(via Location Provider)

Application
(Yelp / Pandora / Safari Browser)
(cached on device)

A

Third Party Website
Yelp.com / Pandora.com / Foursquare.com
(via application or location aggregator)

WWW

Mobile Advertiser
Google Admob / Apple iAds / Millenium
(via third party application)

$

Mobile Operating System
Google / Apple
(cached on device)

on device

off device

Mobile Location Ecosystem
Prepared for Senate Judiciary Panel on Mobile Privacy - May 5, 2011
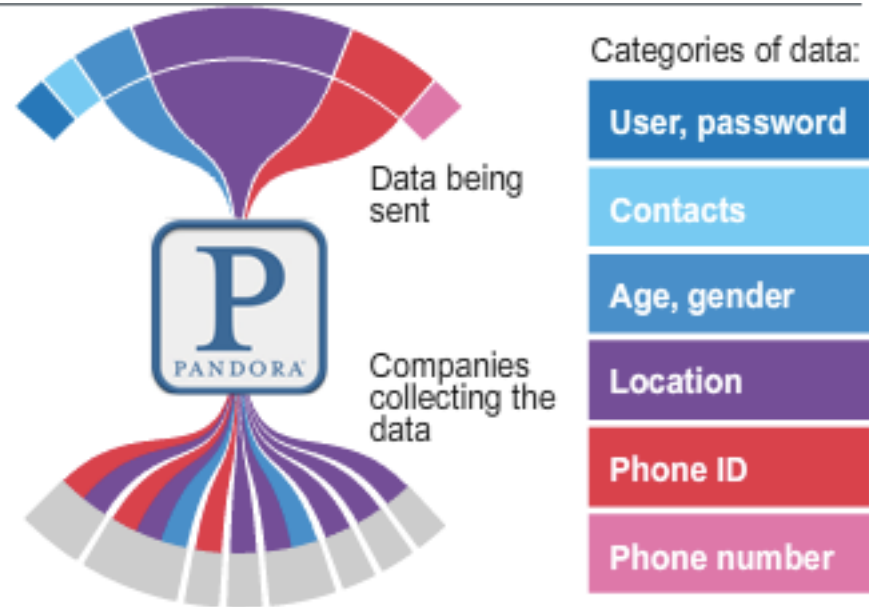Ashkan Soltani - http://ashkansoltani.org

What They Know - Apps

What we found on one app

The iPhone version of music app Pandora sent information to eight trackers. It sent location data to seven of these, a unique phone ID to three and demographic data to two.
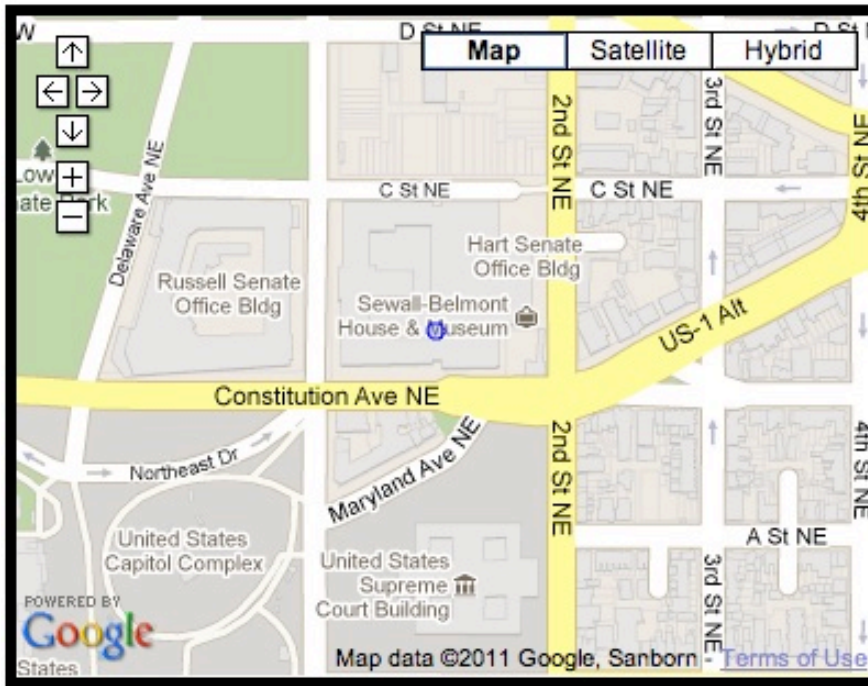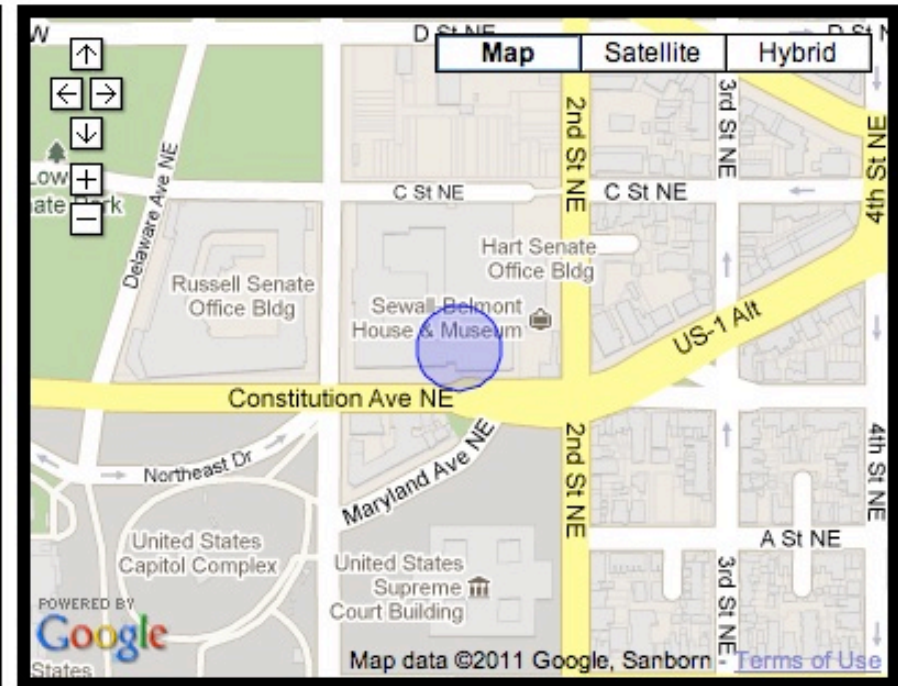
☐ **Click to explore data on all the apps**

Data being sent

PANDORA

Companies collecting the data

Categories of data:

User, password

Contacts

Age, gender

Location

Phone ID

Phone number
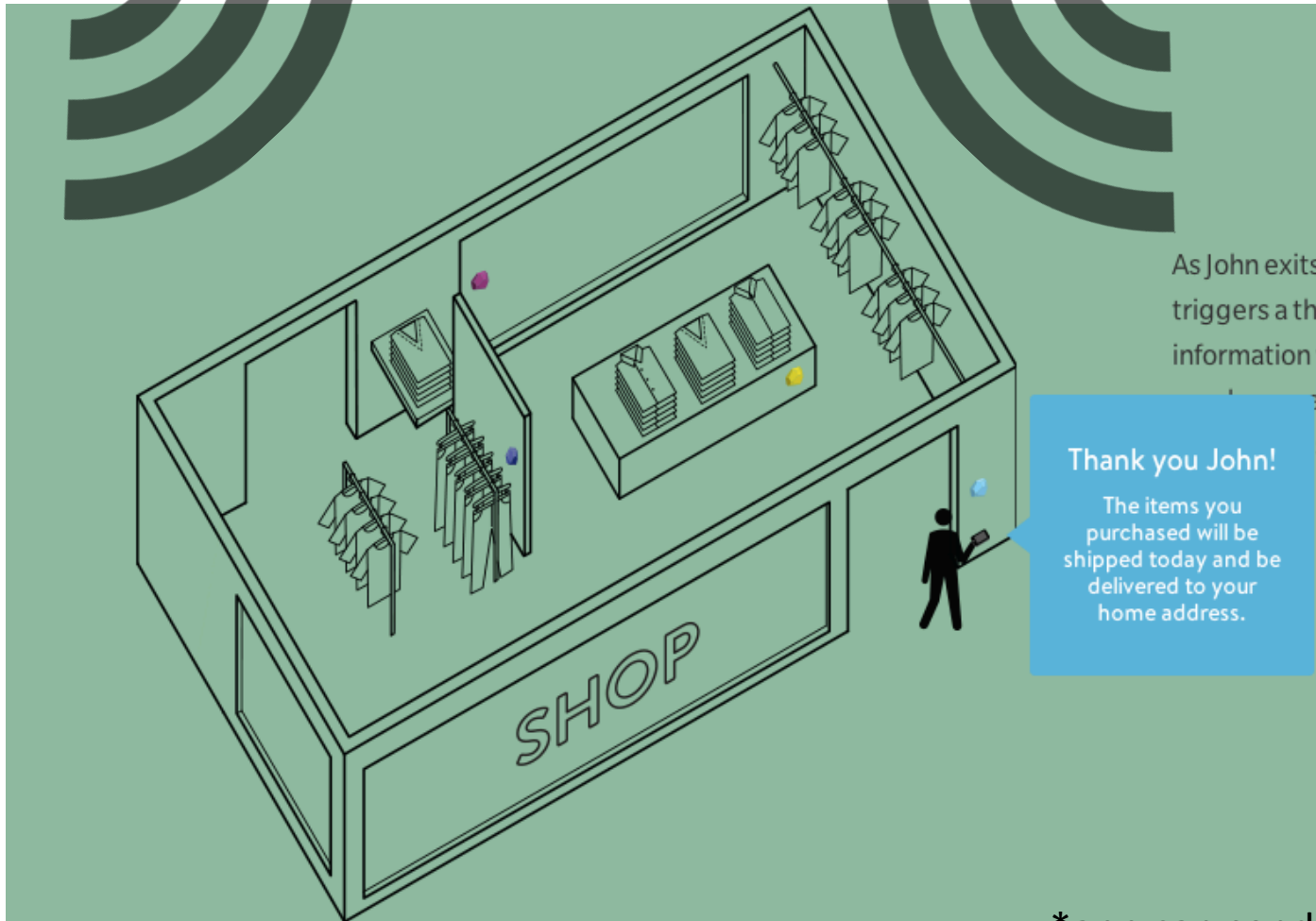
location
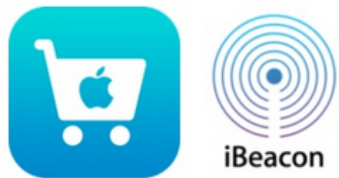
what do we mean by 'location'?

*different levels of accuracy

Bluetooth Low Energy (iBeacons)
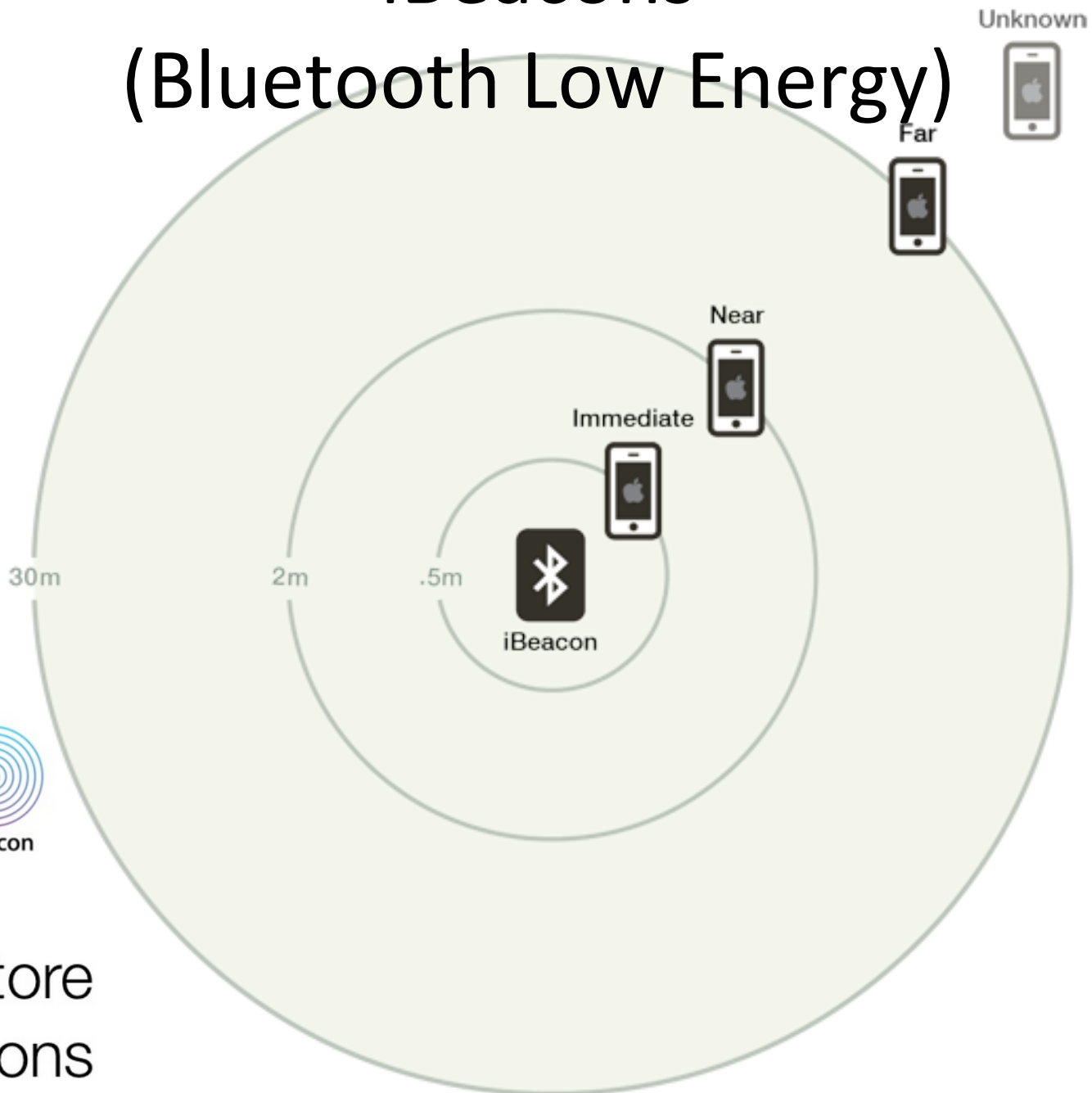

*app can send back info

iBeacons
(Bluetooth Low Energy)
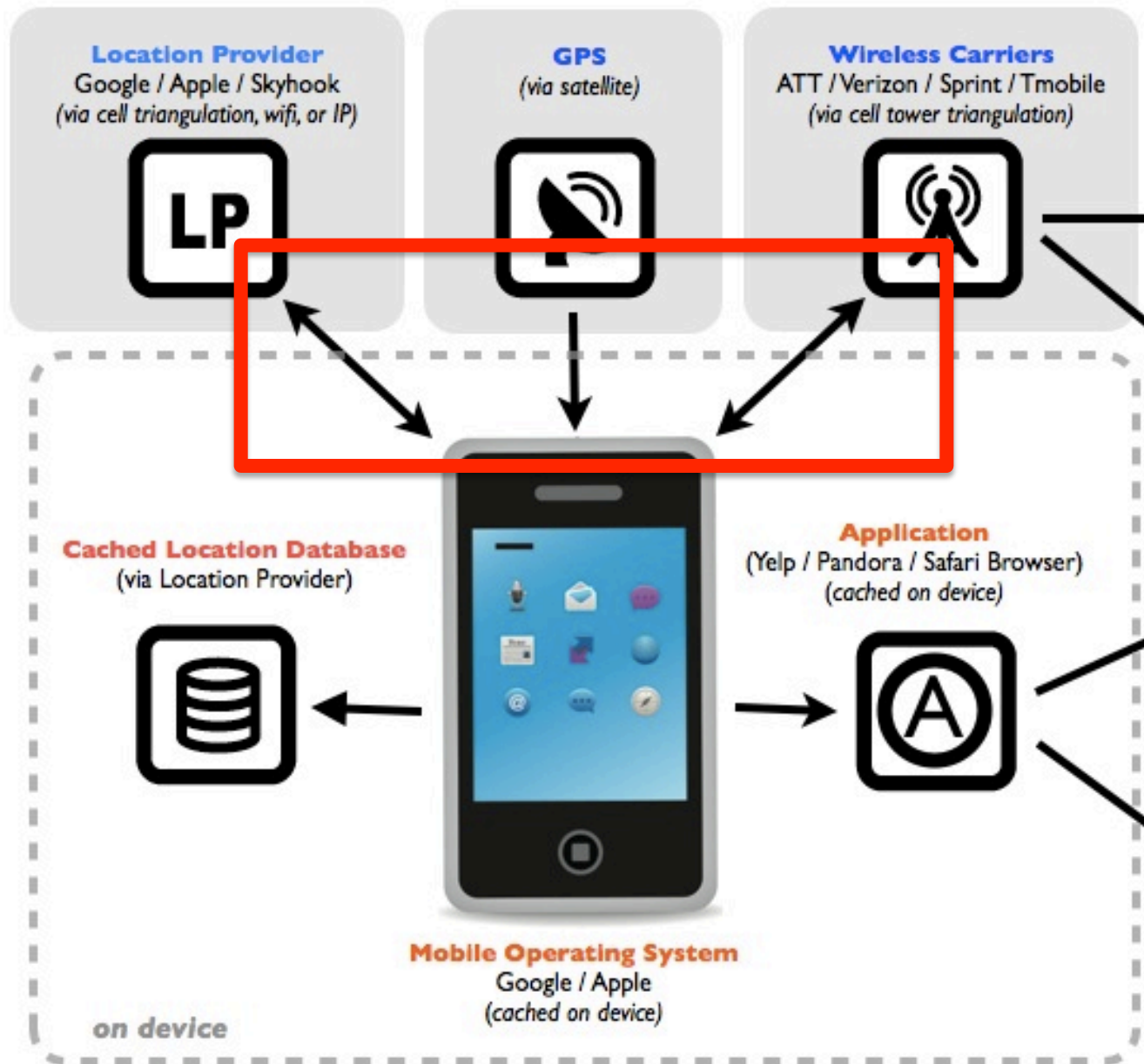
device aware locations

device aware locations

*the venue tracks the device

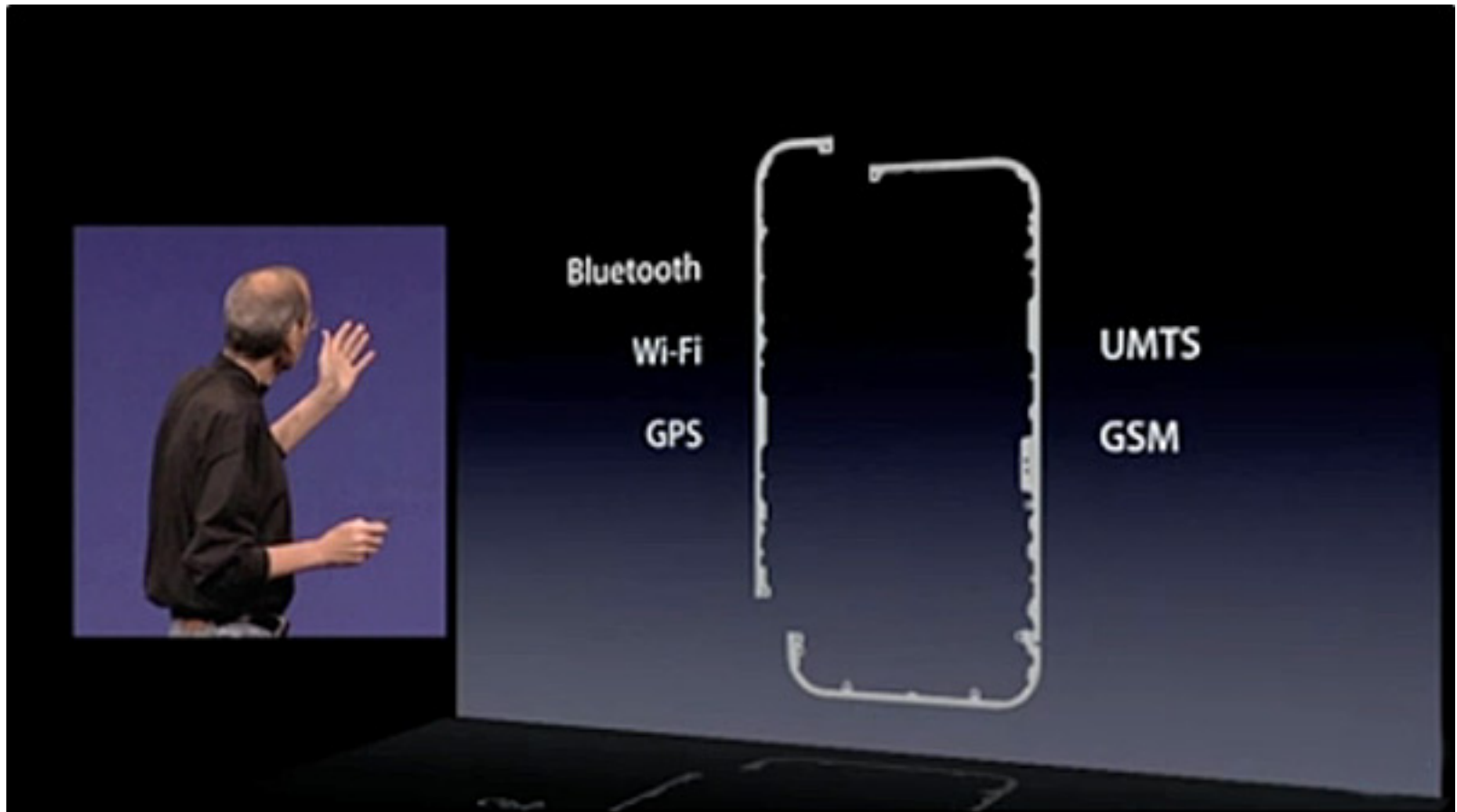# device aware locations

Location Provider
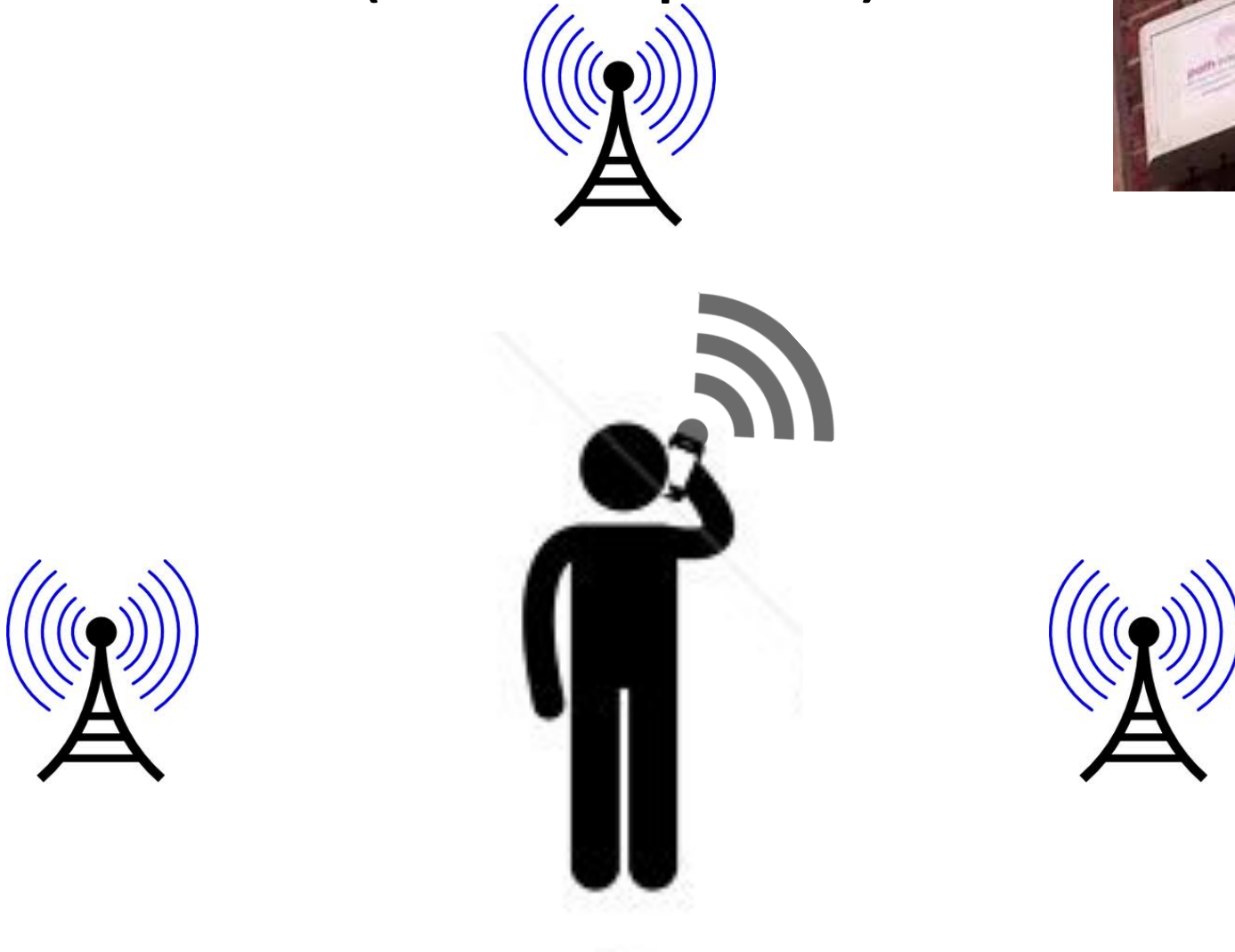Google / Apple / Skyhook
(via cell triangulation, wifi, or IP)

LP

GPS
(via satellite)

Wireless Carriers
ATT / Verizon / Sprint / Tmobile
(via cell tower triangulation)

Cached Location Database
(via Location Provider)

Application
(Yelp / Pandora / Safari Browser)
(cached on device)

A

Mobile Operating System
Google / Apple
(cached on device)

on device

mobile phones*
(antennas)

gsm

gsm
(interception)

pathintelligence

wifi/bluetooth (interception)

other methods
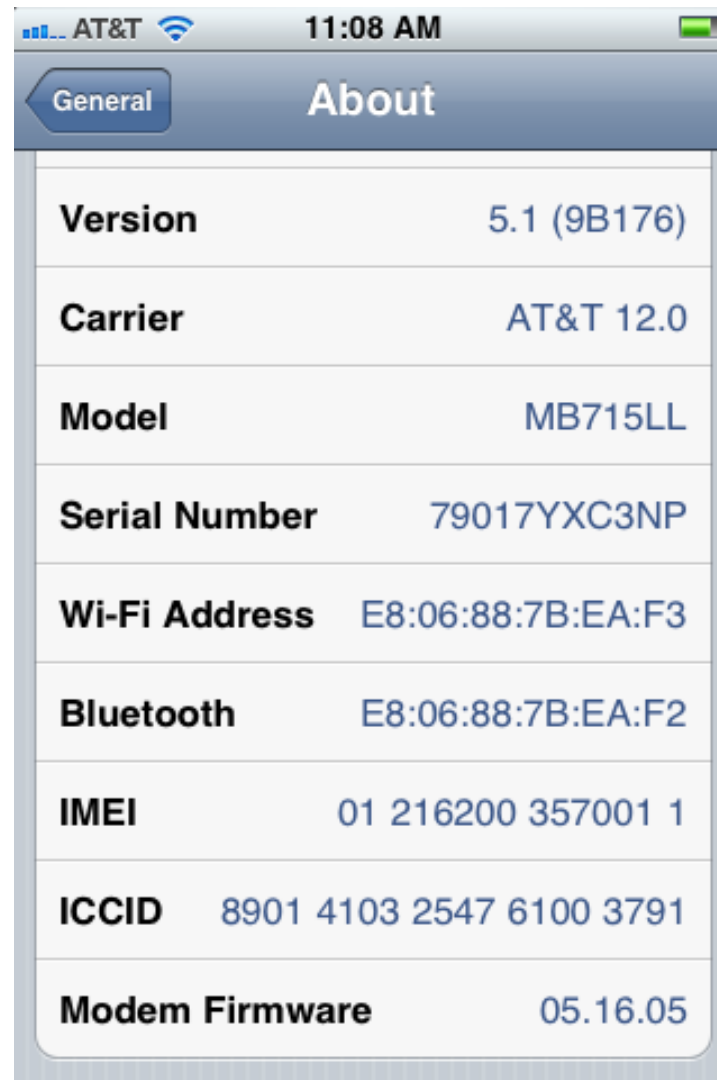
Figure 9: Array of Indoor Location Technologies

Cameras	WiFi	Bluetooth 4.0	RFID	Magnetic	LED Lighting	Sound/ Acoustic
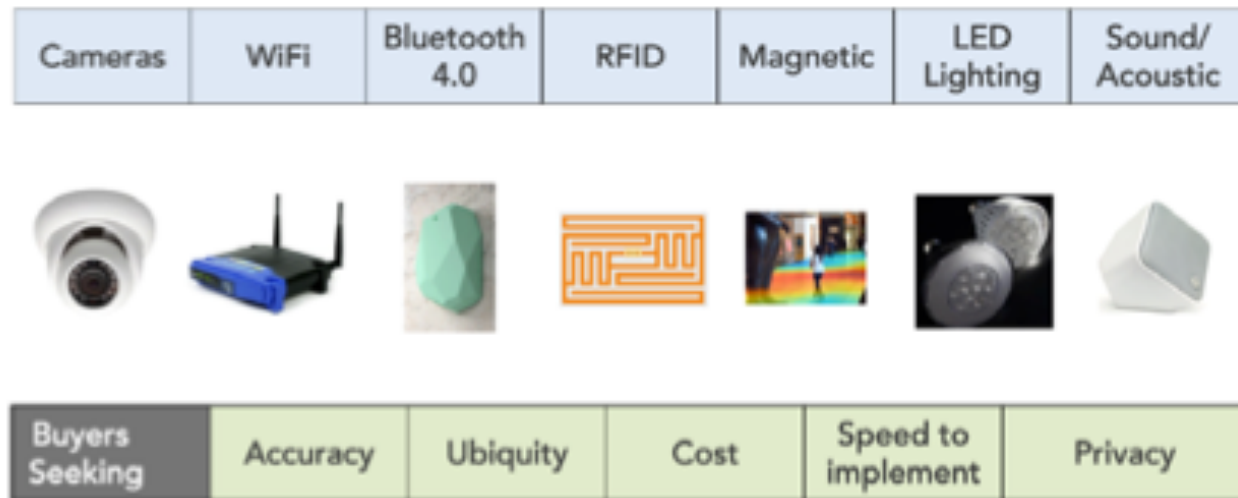
SOURCE: OPUS RESEARCH (2013)

Buyers Seeking	Accuracy	Ubiquity	Cost	Speed to implement	Privacy

identification
(tracking)

unique identifiers

unique identifiers

*refers to an individual person or device

hashing

Wi-Fi Address   E8:06:88:7B:EA:F3

➢ echo -n "E8:06:88:7B:EA:F3" | shasum -a 256
➢ 48fe4d4f1b4cc95567a8794830401081cd9ff1a79b644782129e5c51569b88aa

*guaranteed to be the same every time

'hashing'

« Transparency as a User Experience Problem | Are pseudonyms "anonymous"? »

Does Hashing Make Data "Anonymous"?
BY ED FELTEN

One of the most misunderstood topics in privacy is what it means to provide "anonymous" access to data. One often hears references to "hashing" as a way of rendering data anonymous. As it turns out, hashing is vastly overrated as an "anonymization" technique. In this post, I'll talk about what hashing is, and why it often fails to provide effective anonymity.

What is hashing anyway? What we're talking about is technically called a "cryptographic hash function" (or, to super hardcore theory nerds, a randomly chosen member of a pseudorandom function family–but I digress). I'll just call it a "hash" for short. A hash is a mathematical function: you give it an input value and the function thinks for a while and then emits an output value; and the same input always yields the same output. What makes a hash special is that it is as unpredictable as a mathematical function can be–it is designed so that there is no rhyme or reason to its behavior, except for the iron rule that the same input always yields the same output. (In this post I'll use a hash called SHA-1.)

It should be clear by this point that hashing an SSN does not render it anonymous. The same is true for any data field, unless it is much, much, much harder to guess than an SSN–and bear in mind that in practice the analyst who is doing the guessing might have access to other information about the person in question, to help guide his guessing.

Q Search...

Federal Trade Commission

Protecting America's Consumers

Recent Tweets

FTC Chairwoman Edith Ramirez announced the appointment of Latanya

Follow @TechFTC

'hashing'

12. If a law enforcement agency or a company told Euclid the MAC address for someone's smartphone and asked what stores the owner of that smartphone had previously walked past or visited, would Euclid be able to answer that question?

If the authorities provided the MAC address for a device, Euclid would only be able to determine whether the device had passed near one of its sensors by running it through the original hash function and then searching its databases for an identical hash result. Even if any matching information existed, we would only release it to a requesting agency if it complied with all necessary legal processes.

benefits

consumer benefits
(opt-in)

consumer benefits
(opt-in)

retailer benefits
(opt-out)

Our **FootPath** system calculates shopper locations to approximately 1-2m via transmissions from shoppers' mobile phones.
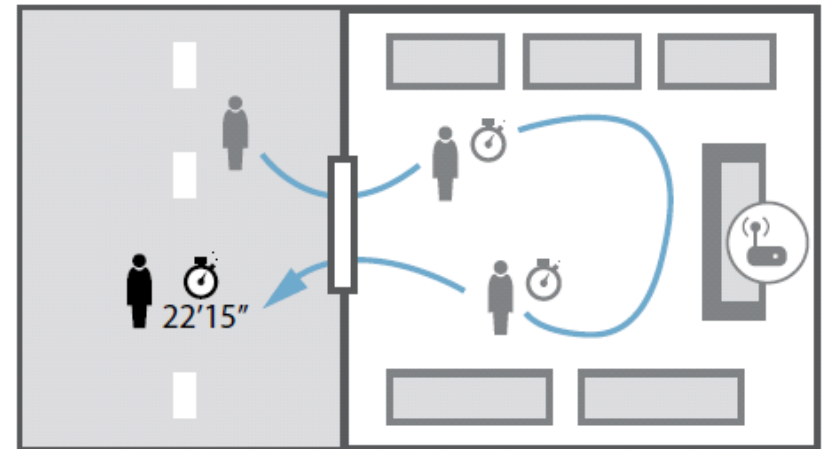
retailer benefits (opt-out)

concerns

retailer benefits
(opt-out)
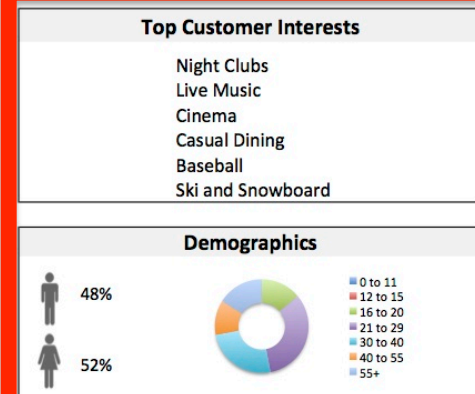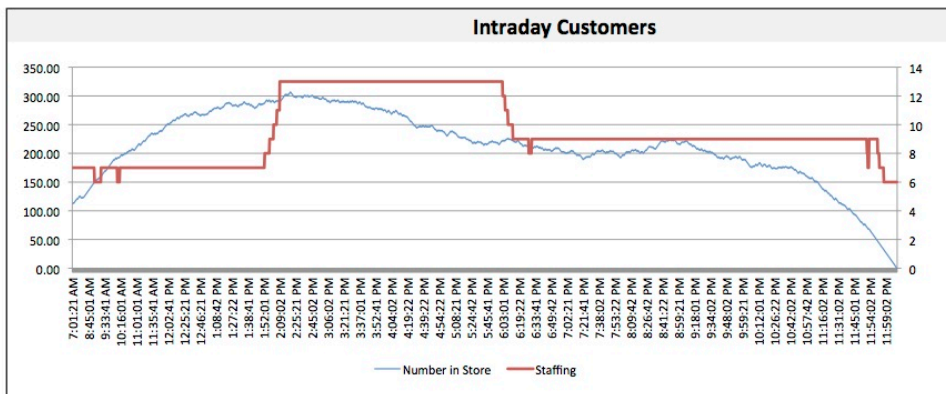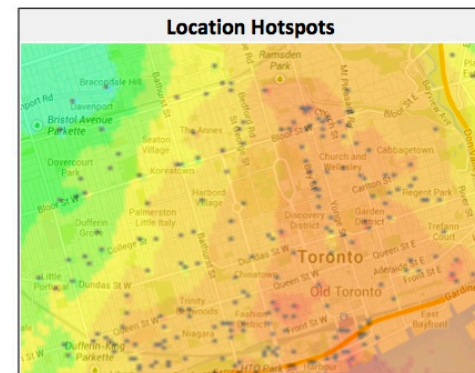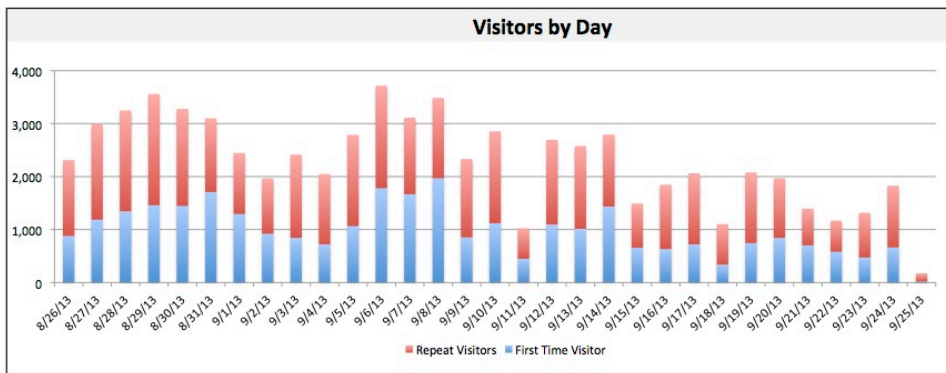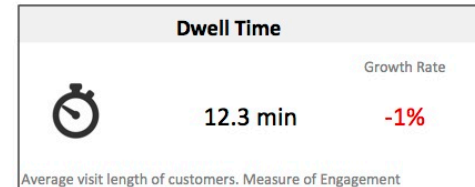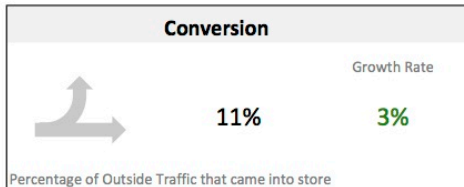
retailer benefits
(opt-out)

A TALE OF ONE CITY.
WHAT HAPPENS DURING A TYPICAL DAY IN PHOENIX?*

By identifying the top points of interest (POIs)—like restaurants, banks, gas stations and shops—we can see what keeps Phoenix busy.

| 6 AM | 10 AM | 3 PM | 7 PM | WEEKEND |

80% of restaurant POIs are fast foods.

For lunch, casual dining makes up **6** of the top **10** restaurants.

Fast food sees **MORE THAN DOUBLE** the traffic of casual dining.

Home improvement stores see a **16%** decrease in traffic.

Banks see **14%** more traffic now than between 10 and 3 PM.

Coffee shops get **80%** more traffic now than from 6 to 10 AM.

Gas stations see **16%** more traffic than from 10 to 3 PM.

TOP FIVE stores for moms:
Walmart
Ross Dress for Less
Target
Macy's
Children's Place

PRECISION MARKET INSIGHTS FROM VERIZON

* This report was created based on aggregate and anonymous insights provided by the Precision Market Insights platform and extrapolated to present a full view of a national population segment.

concerns

- invisible / passive collection
- opt-out vs opt-in
- difficulty of opt-out (i.e mac addr)
- pseudo-anonymous identifiers
- unclear/indefinite retention
- convergence

convergence

The most advanced in-store analytics solution available.

DATA SOURCES

- Video Cameras
- Wi-Fi & Bluetooth Devices
- Guest Wi-Fi
- Point-of-Sale Systems
- Staffing Systems
- Promotional Calendars
- Payment Cards
- Weather
- Plus More

RetailNext

OUTPUTS

- Web Dashboard
- Custom Reports
- Real-time Alerts
- Mobile App
- Predictive Analytics
- Data Exports
- Custom Analysis
- Maps
- 3rd Party Apps (API)
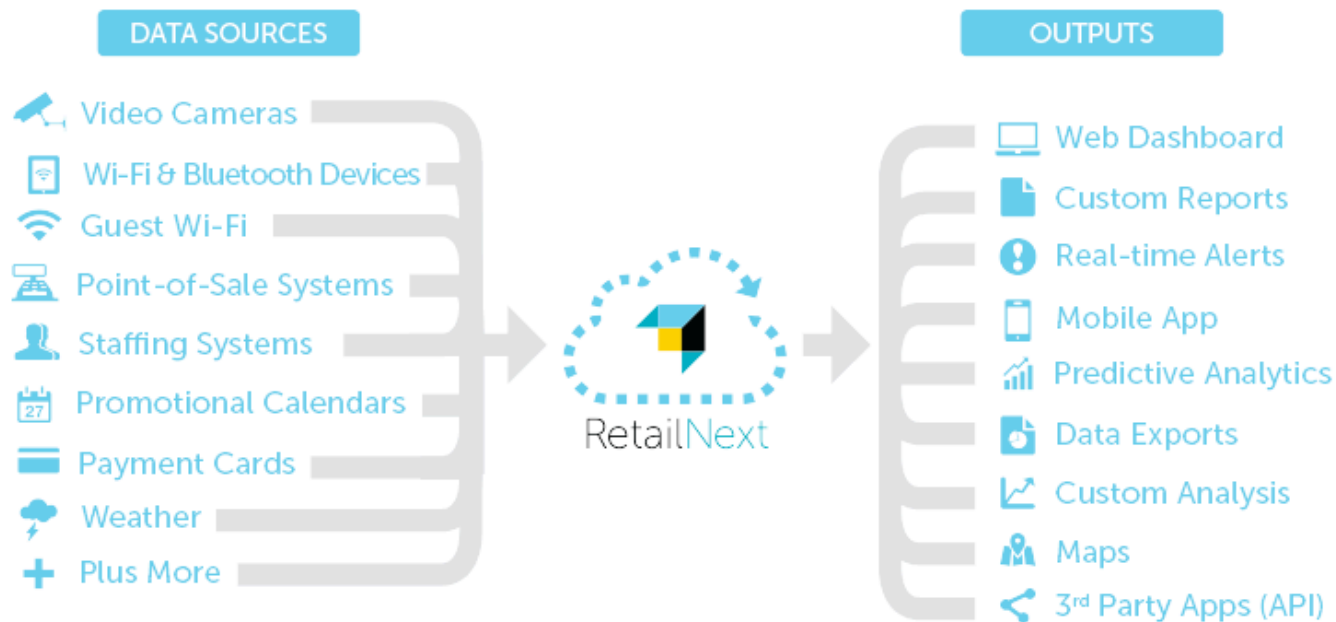
In-store analytics is a collection of systems working together to organize, analyze, and visualize massive amounts of diverse retail data.

convergence (hashing)

questions?

twitter: @ashk4n
ashkan.soltani@gmail.com