SPRING PRIVACY SERIES

Mobile Device Tracking

FEBRUARY 19, 2014

Welcome

Technological Overview

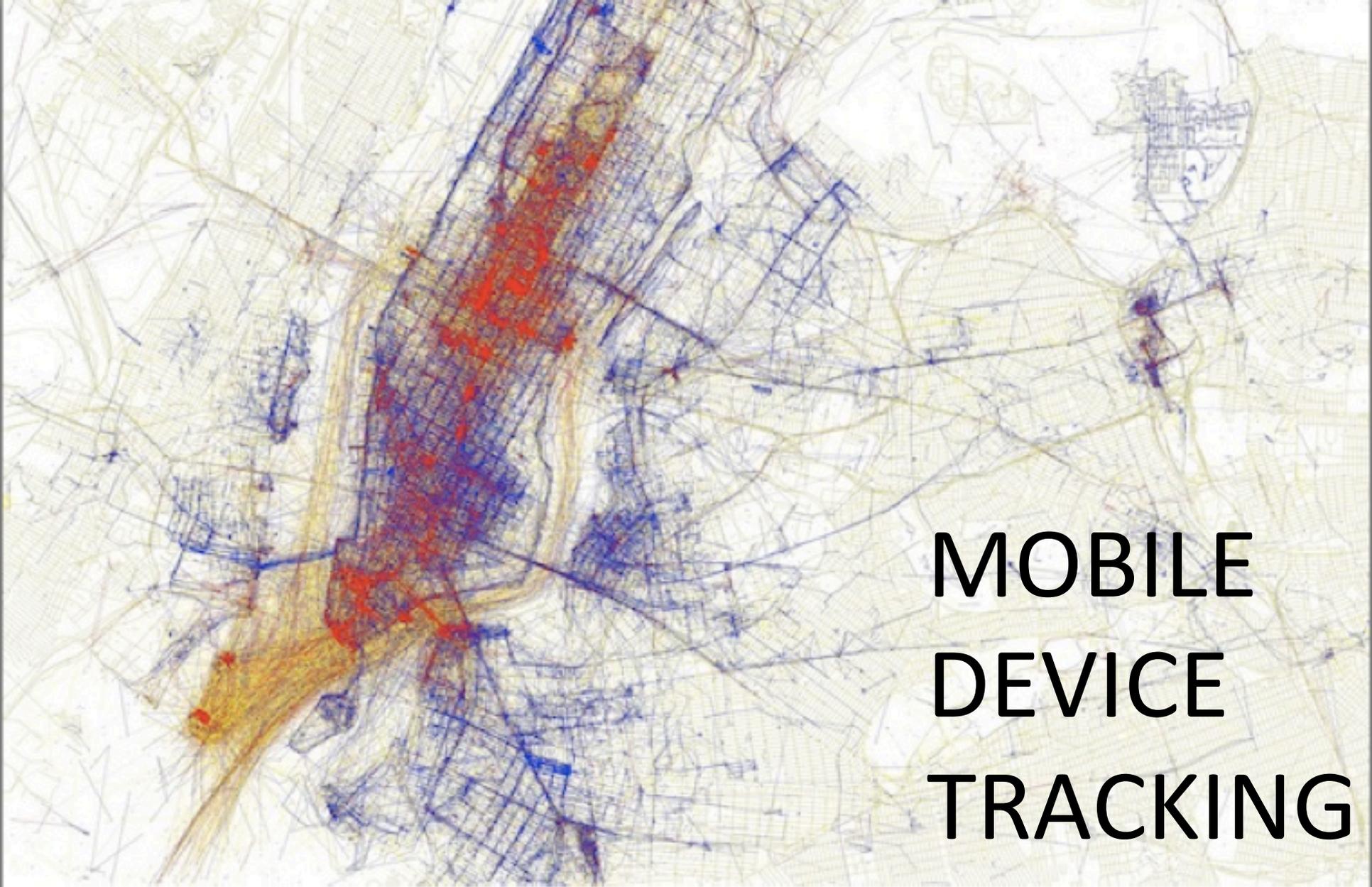## Ashkan Soltani

Independent Researcher and Consultant

MOBILE DEVICE TRACKING

FTC: Spring Privacy Series
Washington, DC
Feb 19 2014

whoami

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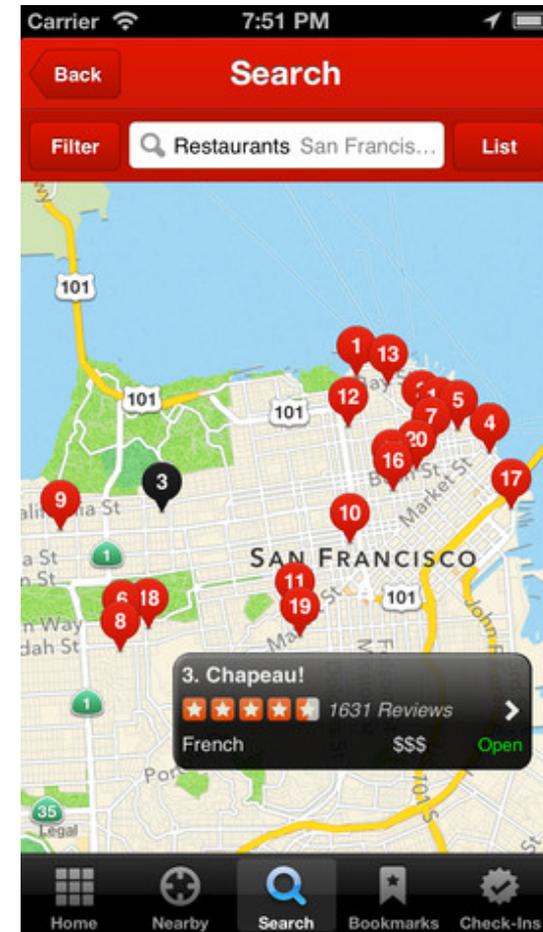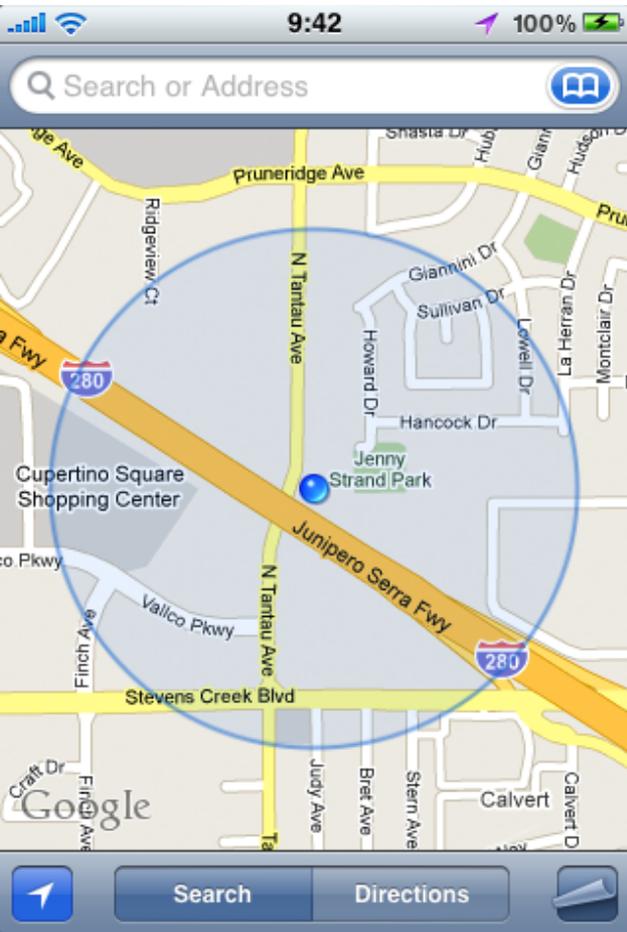
mobile device tracking

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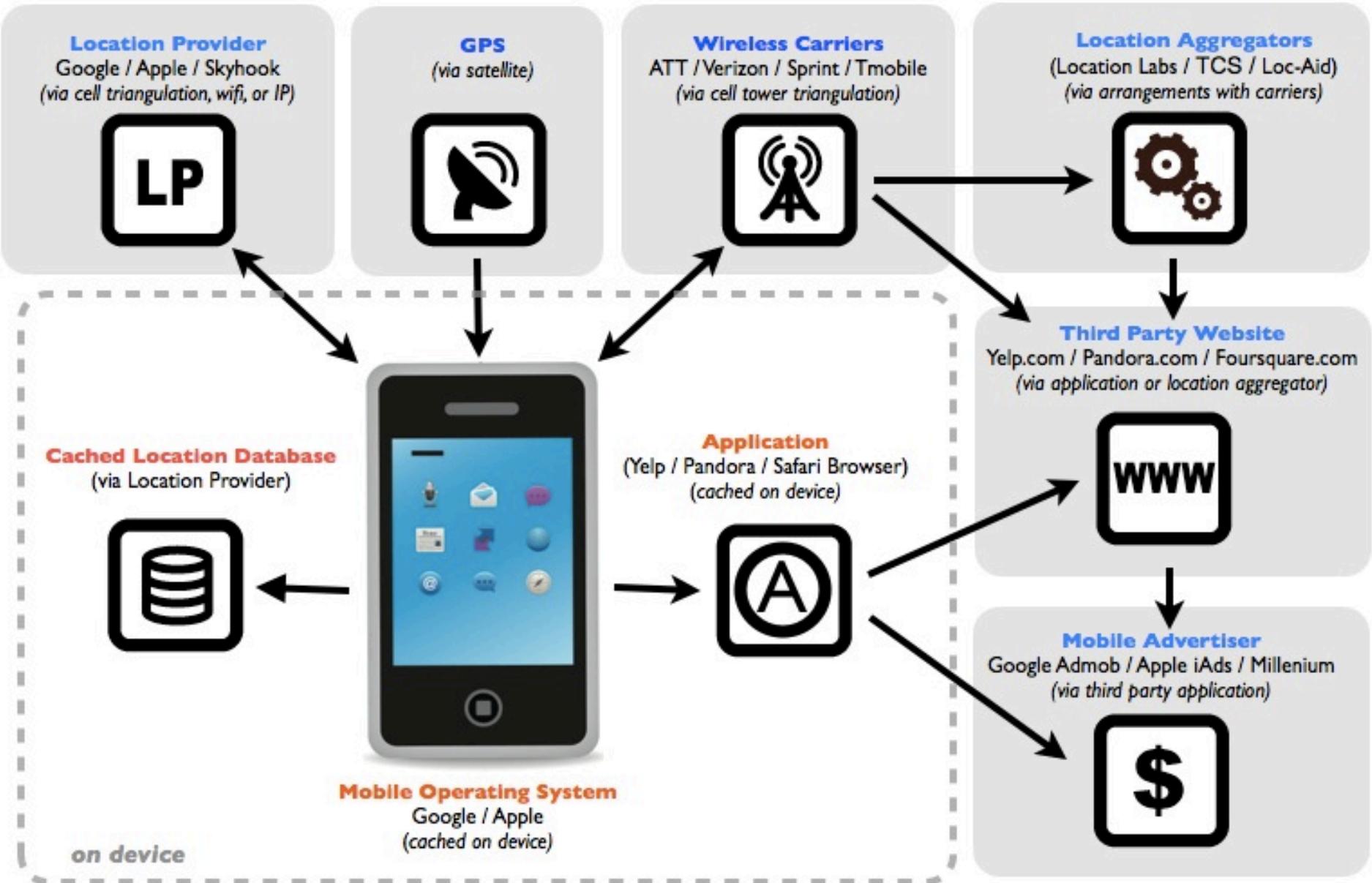- benefits
- concerns

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
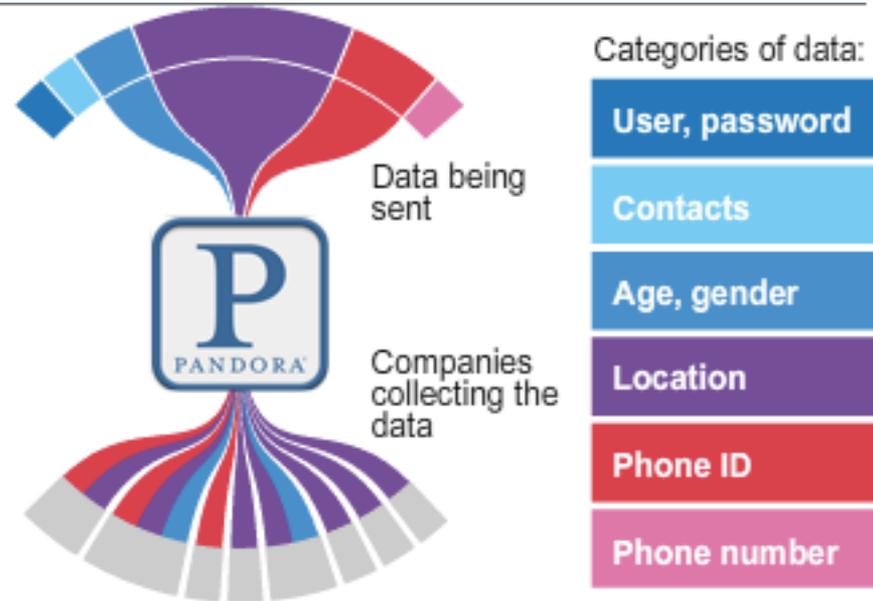Ashkan Soltani - http://ashkansoltani.org

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.
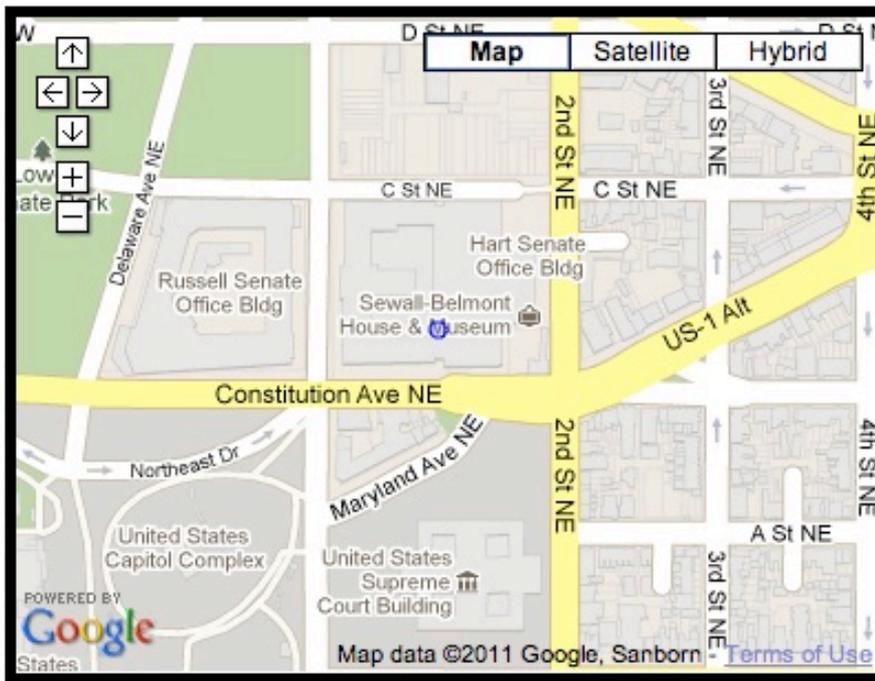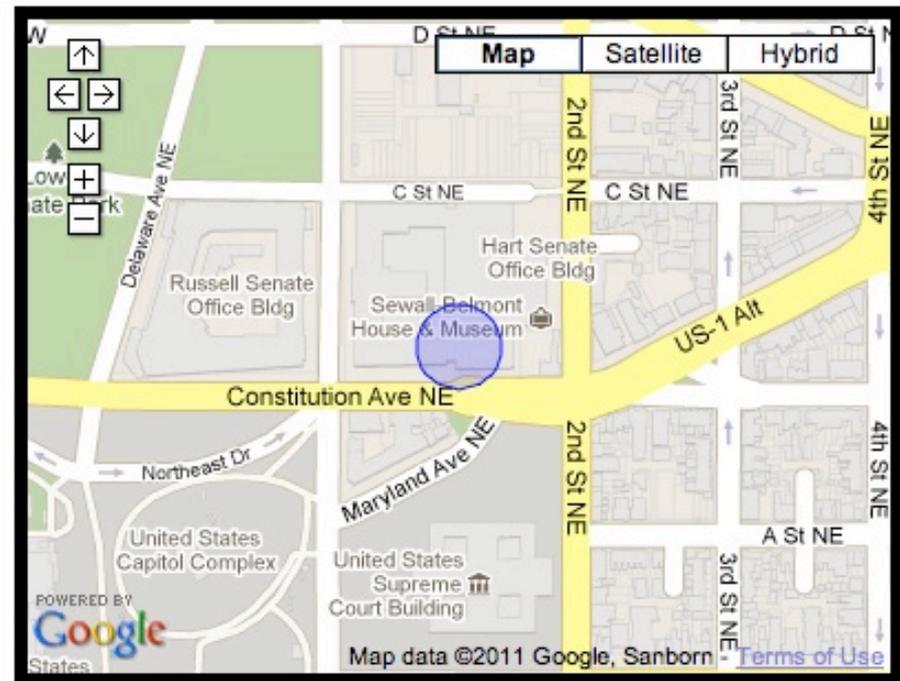
Click to explore data on all the apps

Data being sent

Companies collecting the data

Categories of data:
- User, password
- Contacts
- Age, gender
- Location
- Phone ID
- Phone number

location

what do we mean by 'location'?

GPS

Wi-Fi

*different levels of accuracy

Bluetooth Low Energy (iBeacons)

As John exits...
triggers a th...
information...

Thank you John!
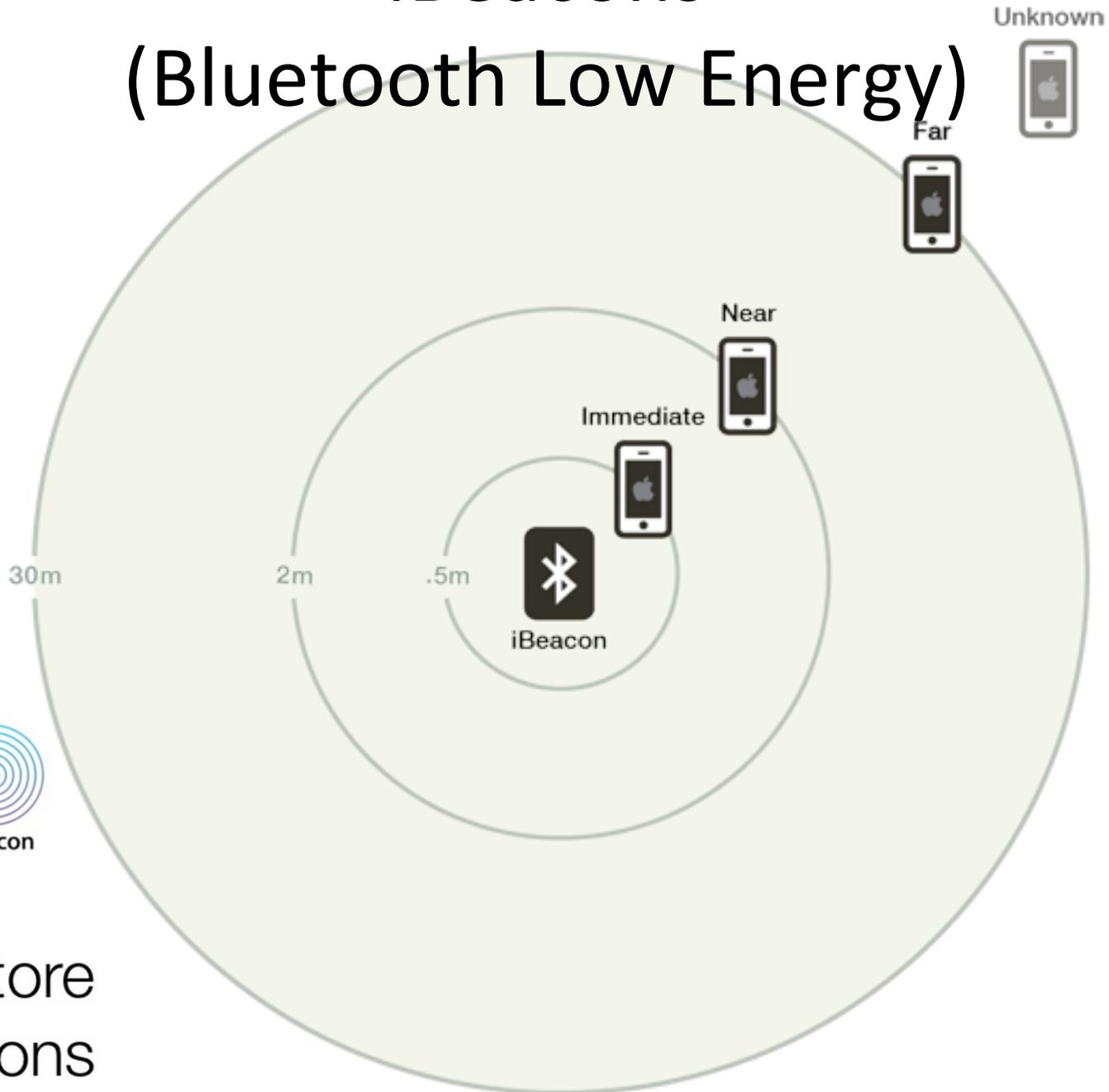The items you purchased will be shipped today and be delivered to your home address.
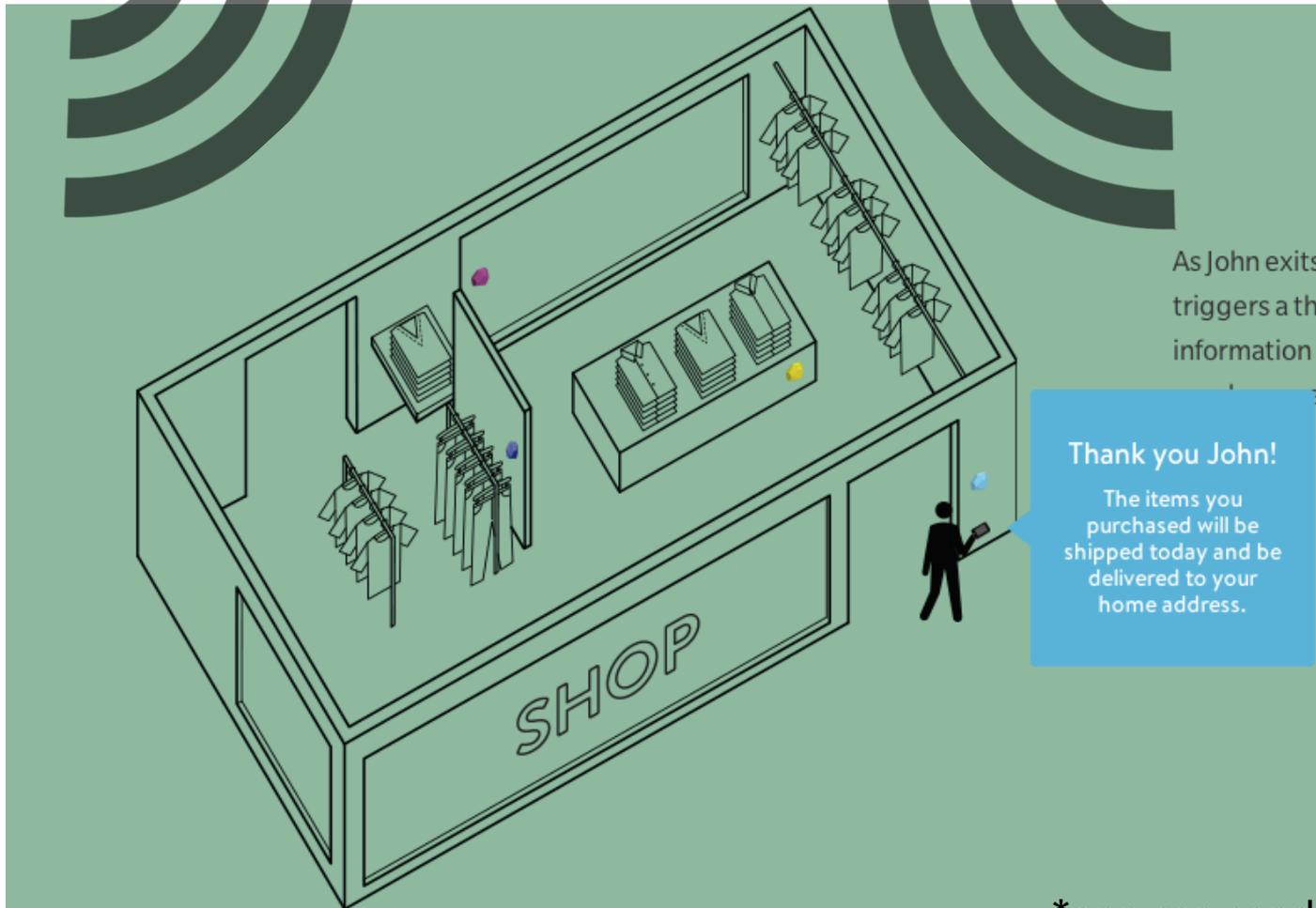
*app can send back info

iBeacons
(Bluetooth Low Energy)

Get In-Store
Notifications

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
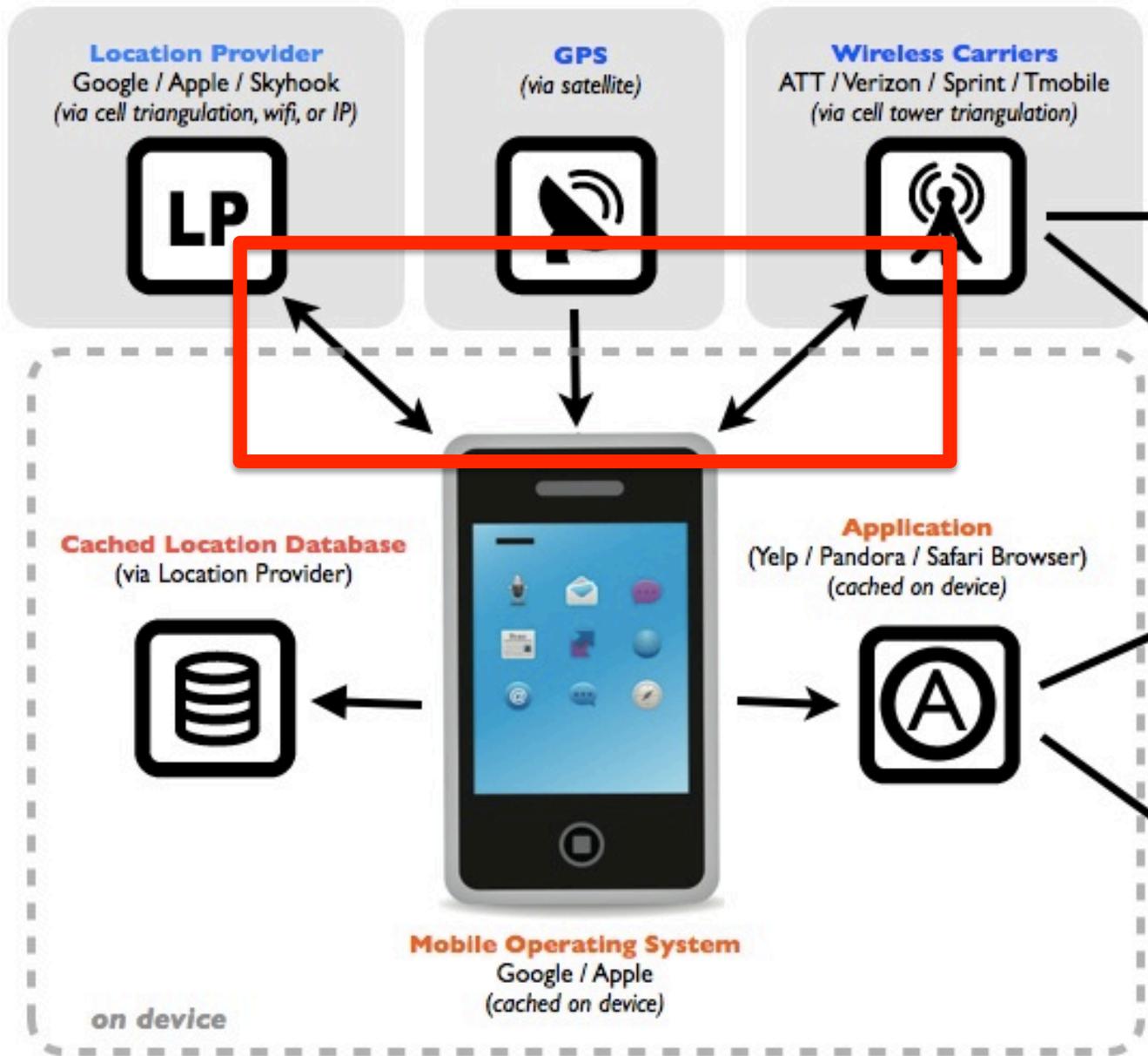
mobile phones*
(antennas)

gsm
(interception)

pathintelligence
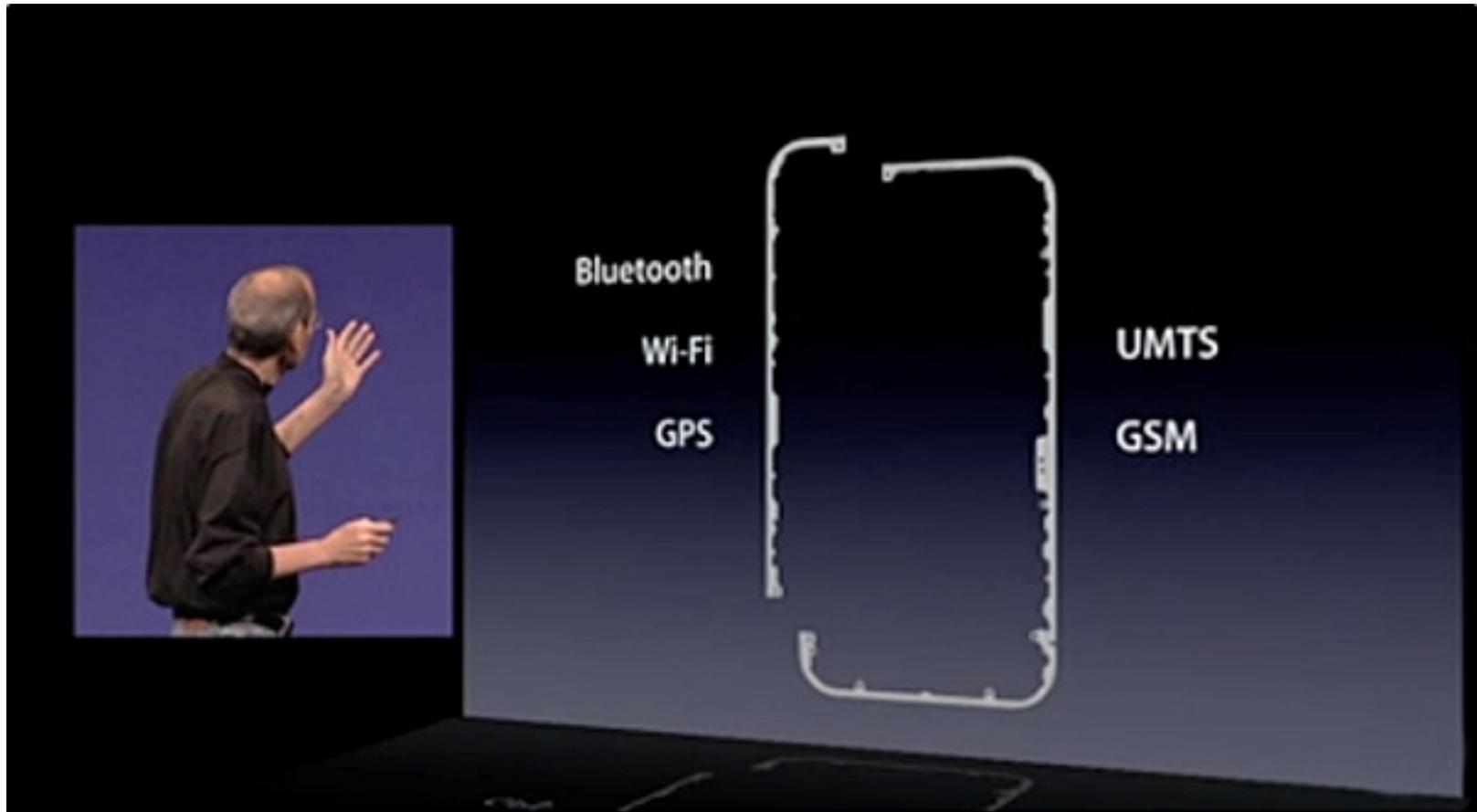
wifi/bluetooth
(interception)

other methods

Figure 9: Array of Indoor Location Technologies

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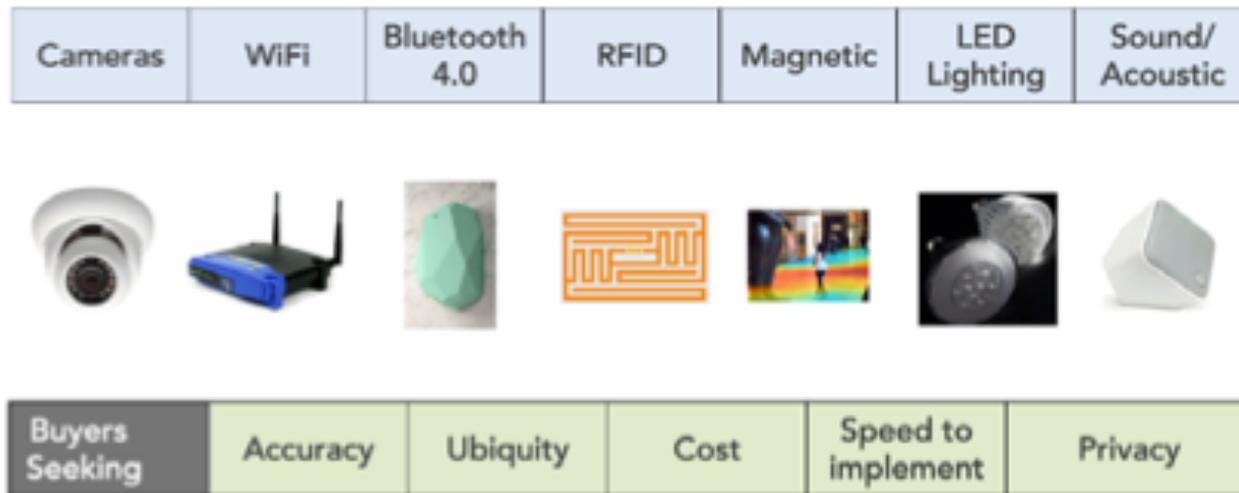
SOURCE: OPUS RESEARCH (2013)

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

➢48fe4d4f1b4cc95567a8794830401081cd9ff1a
79b644782129e5c51569b88aa

*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.

'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)
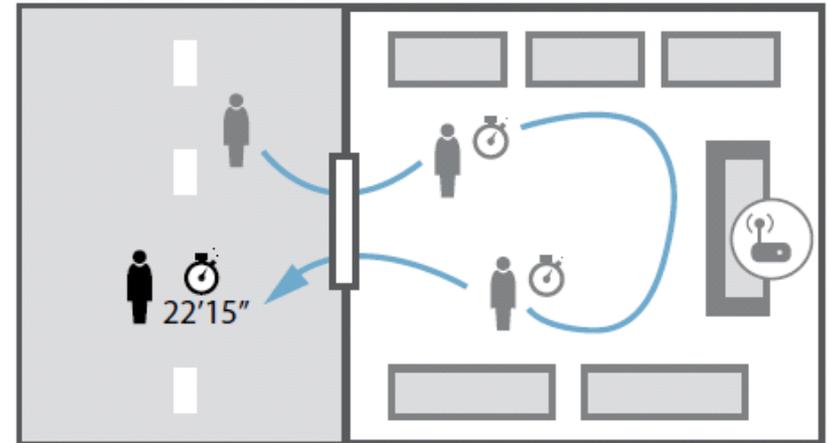
retailer benefits
(opt-out)
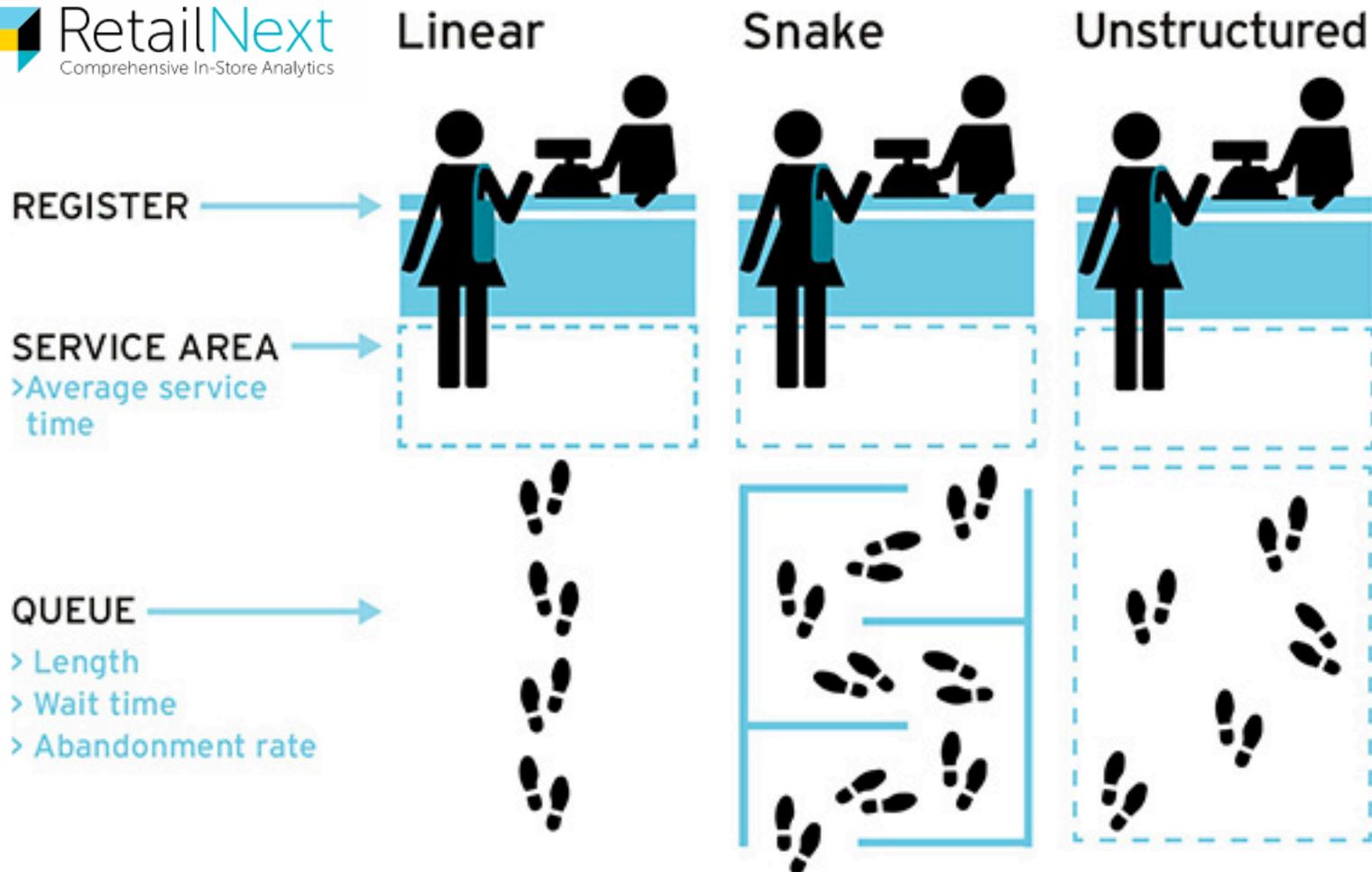
concerns

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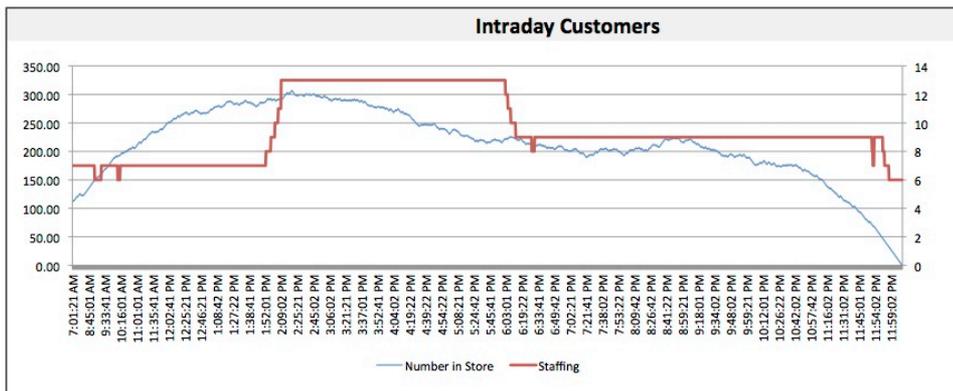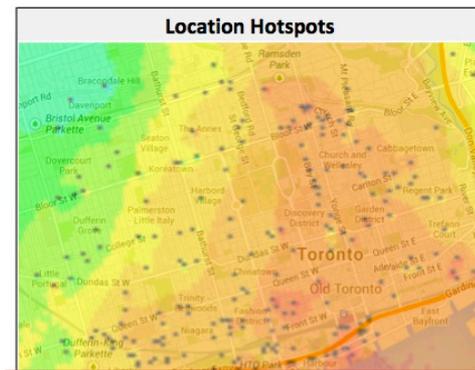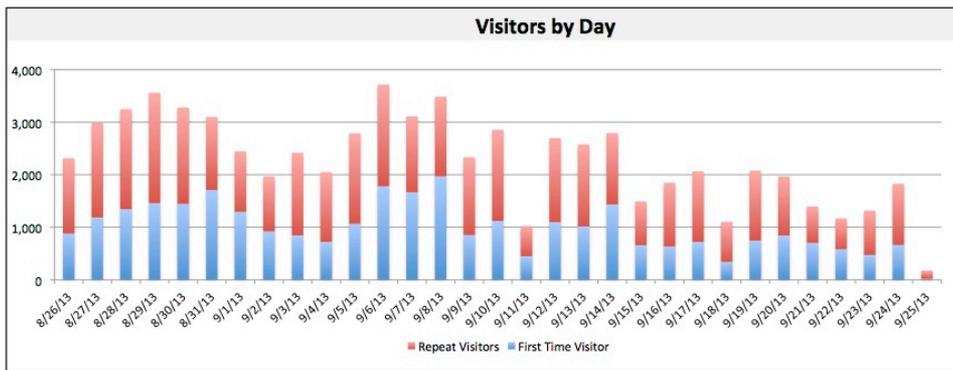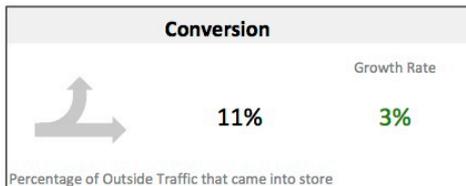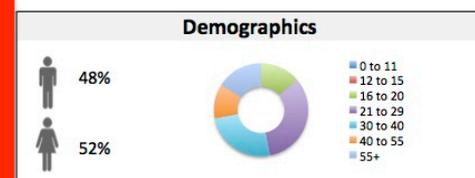
retailer benefits (opt-out)

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)
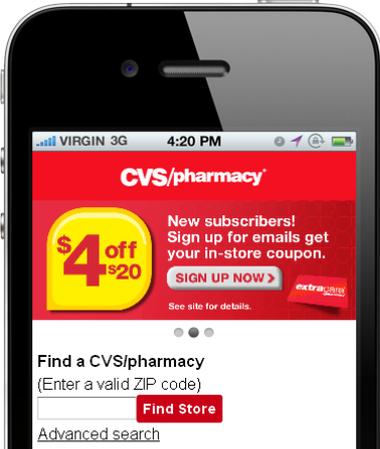
2014-02-17 00:10:40 POST http://analytics.localytics.com/api/v2/applications/ab98a2002763299a5a574cf-c0d33808
-7834-11e1-2223-00a68a4c01fc/uploads
← 202 text/plain [no content] 196.23kB/s

Request   Response
Content-Type: application/x-gzip
Content-Encoding: gzip
x-upload-time: 1392613041
x-install-id:
x-app-id:
x-client-version: android_2.16
User-Agent: Dalvik/1.6.0 (Linux; U; Android 4.4.2; Nexus 4 Build/KOT49H)
Host: analytics.localytics.com
Connection: Keep-Alive
Accept-Encoding: gzip
Content-Length: 631
[decoded gzip] Raw

:1392613760,"attrs":
3c99a3fc2b32b3"."iu"
.cvs.launchers.cvs"
04","aio":"T30dd66c9
0a68a4c01fc","nca":"
3","wmac":"735a5db5b20f5ca9d2f869ddd9dd28b6278a385e4a0c493fa55   ","dp

questions?

twitter: @ashk4n
ashkan.soltani@gmail.com

Panel Discussion

- **Mallory Duncan**, Senior Vice President and General Counsel, National Retail Federation

- **James Riesenbach**, Chief Executive Officer, iInside

- **Seth Schoen**, Senior Staff Technologist, Electronic Frontier Foundation

- **Glenn Tinley**, President and Founder, Mexia Interactive

- **Ilana Westerman**, Chief Executive Officer, Create with Context

We create the digital future.

Tuesday, February 18, 14

What creates trust?

"It is what I expected"   **"I can control it"**   **"I will do it"**

1. Transparency	2. Choice	3. Engagement
I am aware	I have options	It is easy to use
It is clear	I get to weigh the tradeoffs	Context is taken into account
I understand	I get to decide	On my terms
It is intuitive	… If I care	

4. Value

"I get something for it"

Tuesday, February 18, 14

createwithcontext

Good design is based on context, and the first step is to understand

Design is not art. Art is an expression of self, and comes from within. Art can be created without context.

Design, is about communication, it is about the receiver and their context. The first step to good design is to understand.

THE CONTEXT

Their environment

Cultural norms

Social context

Attention

Situation

Noise

The people

Knowledge
History
Values
Beliefs

Wants
Needs
Tradeoffs
Value

Tuesday, February 18, 14

createwithcontext
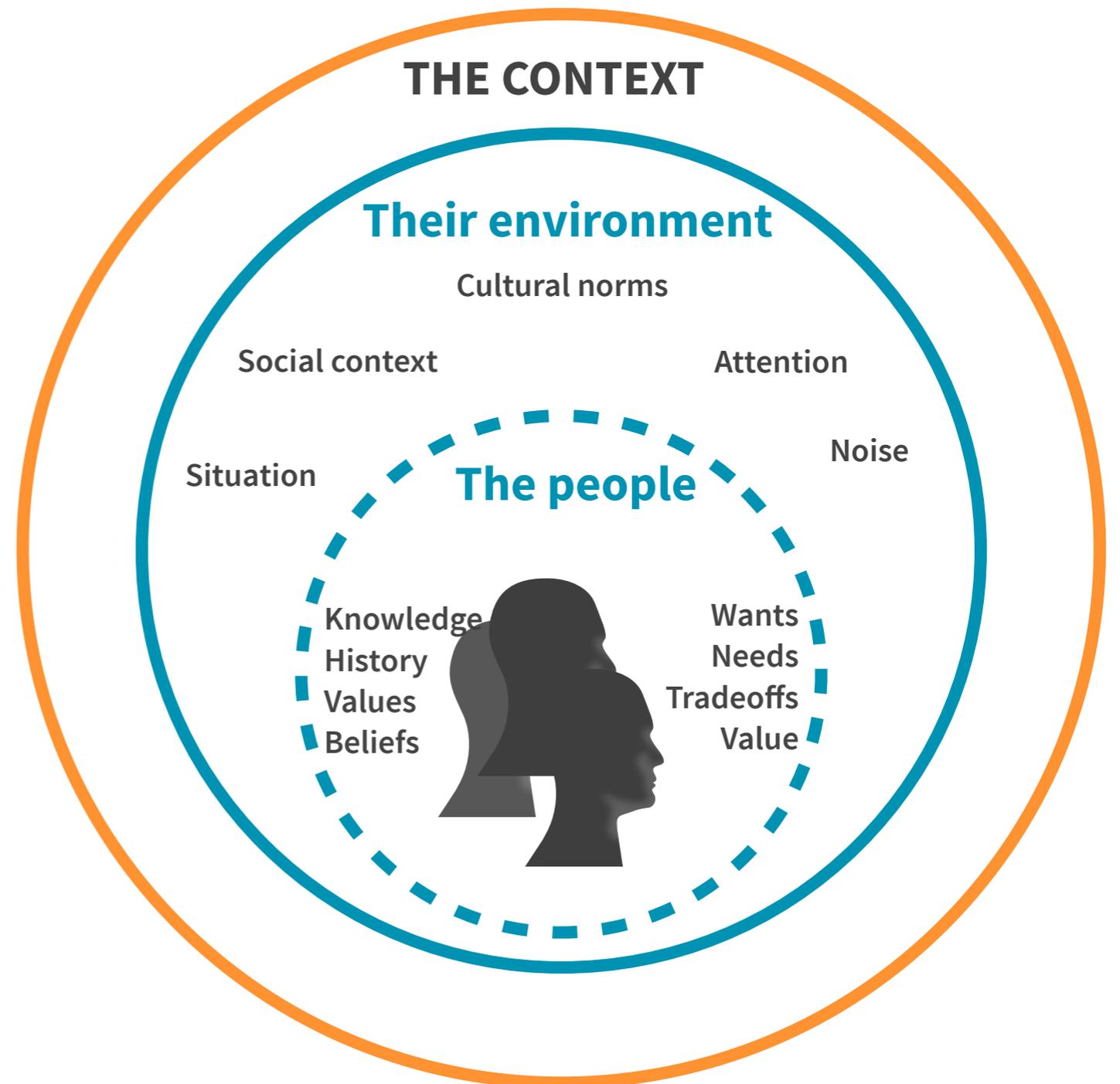
Creating trust for in-store mobile data collection

Research study, **large retailers**, US only.

Participants	Methodologies	Broad Sample
4624	In-store observation, secret shopper, 1:1 interviews, concept test, survey Day and evening, weekend and weekday	Urban, suburban, rural In-mall & standalone stores Bloomingdale's, Costco, Hollister, Walmart, Neiman Marcus, Pottery Barn, Macy's, Target, TJ Maxx, Walgreens

Tuesday, February 18, 14

createwithcontext

Americans trust retailers

On average, we trust retailers.

Doctors	State government	Teachers	CNN	HBO	AT&T	**Target**
Banks	Federal government	Health insurance	Apple	ESPN	Verizon	**Walgreens**
Cable companies	Local government	companies	Google	Yahoo!	**Macy's**	**Costco**
Phone companies	Your employer	New York Times	Facebook	Citibank	**Walmart**	**Sears**

Survey of **general US population**, rating trust of brands and organizations on scale 1 (distrust) to 7 (trust). (N=800) September 2013

5

Tuesday, February 18, 14

We will give up our information, if we get something in return

We asked people if they would provide information for 'a deal' …

97% would give up at least one piece of data to save money
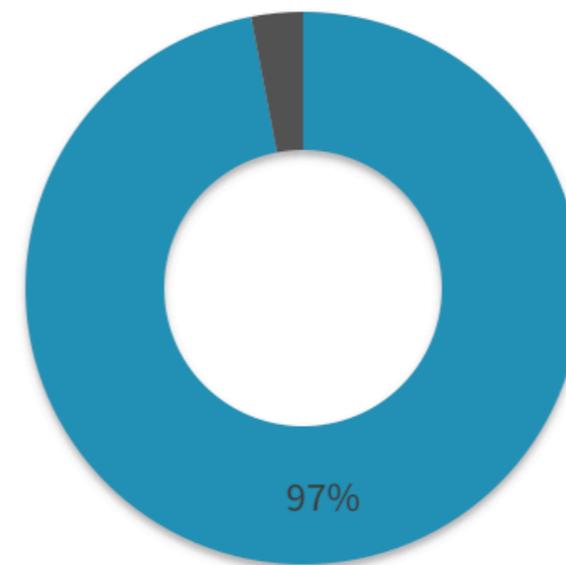
- Would give up at least one piece of data
- Would not give up any data

97%

Your phone number	Email messages you sent and receive (contents and addresses)	Your fingerprint
Which books and magazines you read		Your current location
What you search for online (on computer or mobile phone)	The pictures and files you store on your mobile phone	Your age
Your name	Where you have been (e.g., all locations in your phone)	Your income
What your interests are		Your social network connections
Your address	Your phone's address book	Your credit score
Which apps you use on your phone, and when	What you buy (both online and in stores)	Picture of your face/passport photo

Survey of general population, asking which information they would provide for 50% off of a gallon of milk, a large HD TV and a new car . (N=800) September 2013

Tuesday, February 18, 14

We will give up our data, for value

Tuesday, February 18, 14

create with context

We will give up our data, for value

We asked Alicia her thoughts about the article
"Attention Shoppers: Store Is Tracking Your Cell"

Tuesday, February 18, 14

People are **2.5x more willing** to give their data if it 'makes sense'

If it is clear WHY the data is needed, people are more willing to provide it

75%

would provide their location

When it **makes sense**, there is a **high** willingness to provide information

30%

would provide books and magazines they read

When it **doesn't** make sense, there is a **low** willingness to provide information

Survey of general US population, selecting from a list which information they would trade for benefit. (N=800)

Tuesday, February 18, 14

Some data is more important

% of people willing to provide:

'I have no right to share someone else's information'

5x more

Personal Digital Data
- Where have you been
- Your phone's address book
- Email messages you sent and receive
- Pictures and files you store on your mobile phone
- Social network connection

Sensitive/Financial Info
- Income
- Credit score
- Picture of your face/ passport photo
- Fingerprint

Interests Online
- What you buy
- What you search for online
- Which app you use on your phone and when

Interests Offline
- What your interests are
- Books and magazines you read

Directory Info
- Your name
- Your phone number
- Your age
- Your address
- Your current location

Survey of general population, asking which information they would provide for 50% off of a gallon of milk, a large HD TV and a new car . (N=800) September 2013

9

Tuesday, February 18, 14

There is a big difference in where I am and where I have been

% of people willing to provide:

Where you are now, versus where you have been

Personal Digital Data
- Where have you been
- Your phone's address book
- Email messages you sent and receive
- Pictures and files you store on your mobile phone
- Social network connection

Sensitive/Financial Info
- Income
- Credit score
- Picture of your face/passport photo
- Fingerprint

Interests Online
- What you buy
- What you search for online
- Which app you use on your phone and when

Interests Offline
- What your interests are
- Books and magazines you read

Directory Info
- Your name
- Your phone number
- Your age
- Your address
- Your current location

Survey of general population, asking which information they would provide for 50% off of a gallon of milk, a large HD TV and a new car . (N=800) September 2013

10

Tuesday, February 18, 14

create with context

I don't care about my location, but I do care about my personal data

Tuesday, February 18, 14

I don't care about my location, but I do care about my personal data

Marc understands why stores would want his personal data but he doesn't like it

Tuesday, February 18, 14

create
with context

There is low awareness of potential in-store data collection

While consumers trust retailers,

only 33%

of people believe that large retailers are collecting data such as: **history**, **contacts**, **apps installed, etc.**

In store

Stores access my data from my phone

50% think that **ONLY LOCATION** may be collected

Expectations   Reality
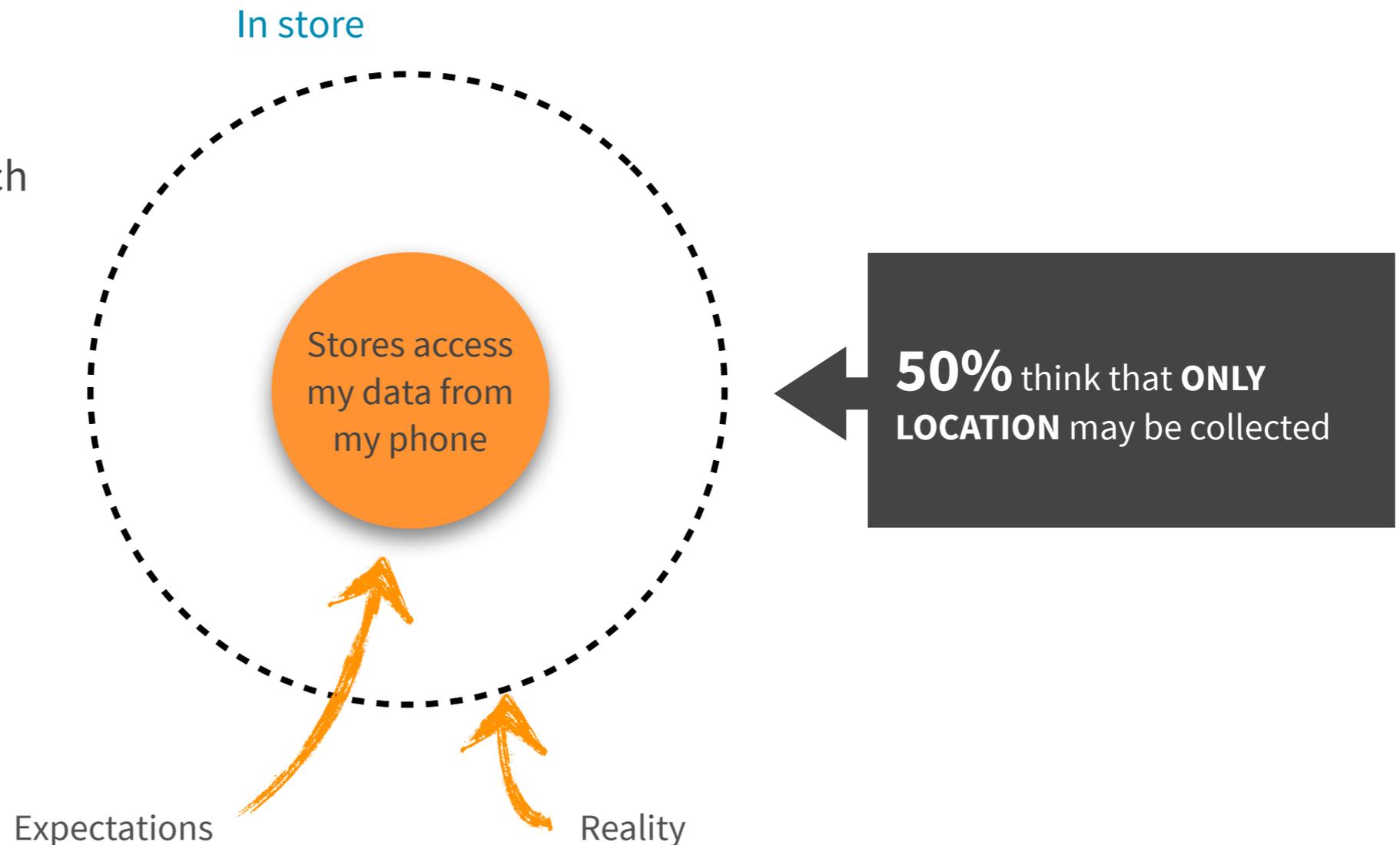
Survey of general population, rating how likely stores collect their data from their phone when they enter in the store. (N=1327) June, September 2013

Tuesday, February 18, 14

create
with context

People are confused

Tuesday, February 18, 14

People are confused

While Daphne thought the store 'might' be able to collect data, she wasn't sure
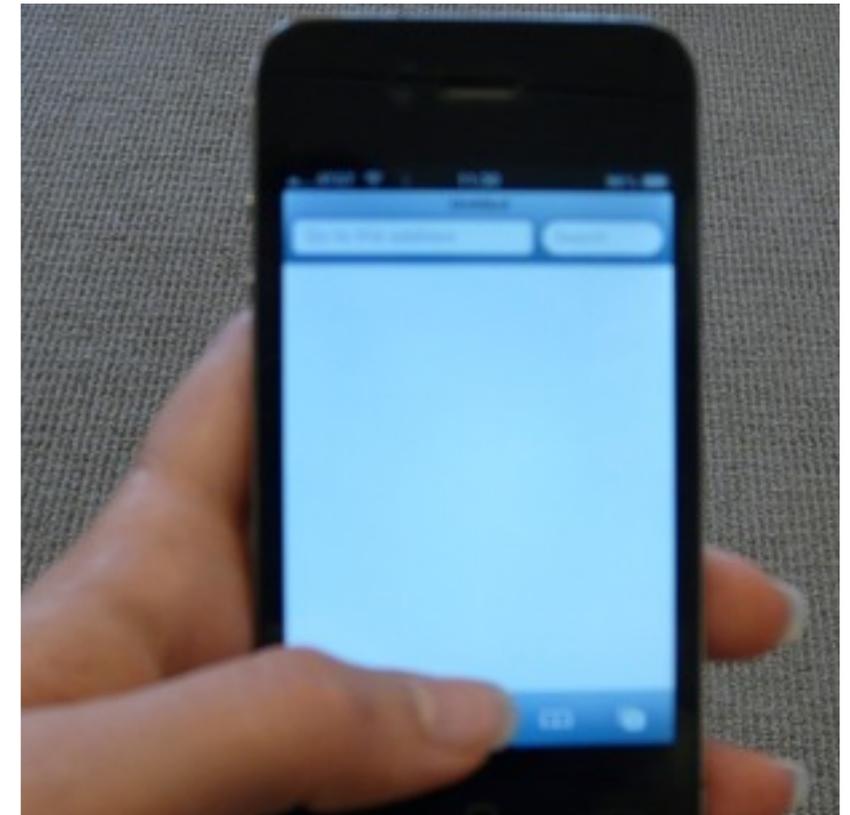
Tuesday, February 18, 14

Can we create awareness?

The 'easiest' way to create awareness is to use current channels, such as signs or mobile phones.
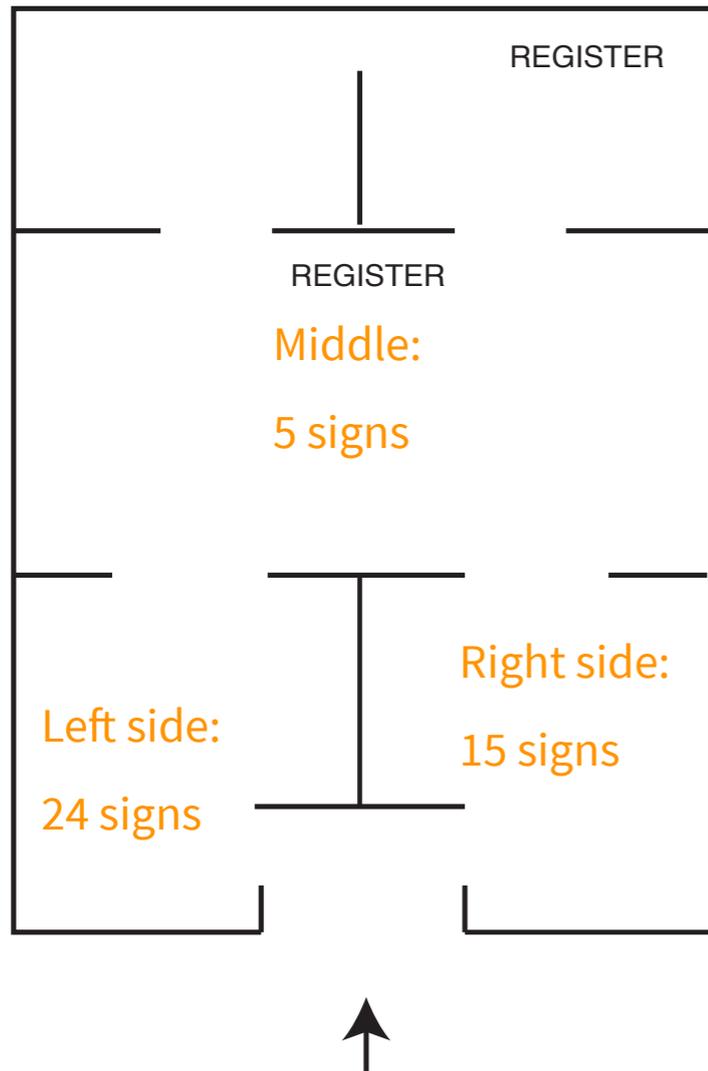
However, will this work?

Signs

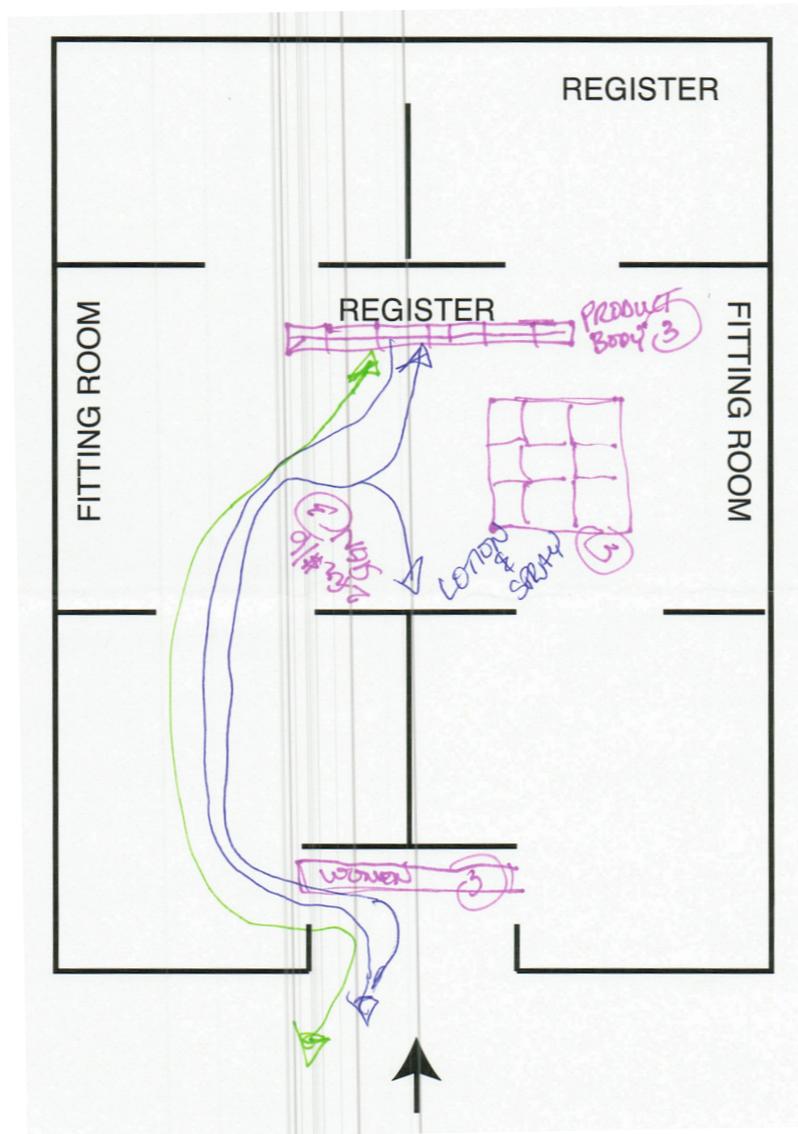**On device**

Tuesday, February 18, 14

Low awareness of in-store signage

Shoppers were asked to purchase and return an item. Later they were asked to draw the signs they saw.

TOTAL: 34 signs

RECALLED: 3

Recall of signage in retail stores
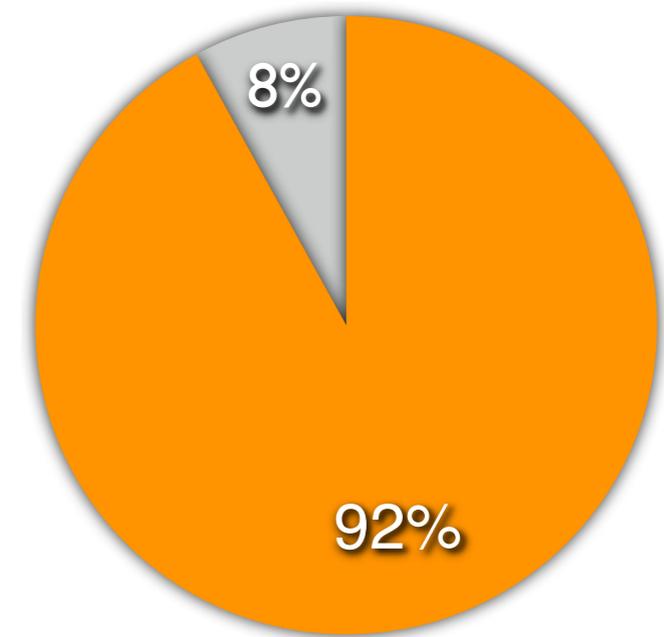
REGISTER

REGISTER

Middle:

5 signs

Left side:

24 signs

Right side:

15 signs

FITTING ROOM

REGISTER

FITTING ROOM

PRODUCT BODY (3

LOTION & SPRAY

WOMEN (3

8%

92%

● Didn't recall   ● Recalled

Tuesday, February 18, 14

create
with context

I don't pay attention

Tuesday, February 18, 14

create
with context

I don't pay attention

Tuesday, February 18, 14

Consumer notifications

0% of people recalled seeing the sign that is present at all registers in the county.
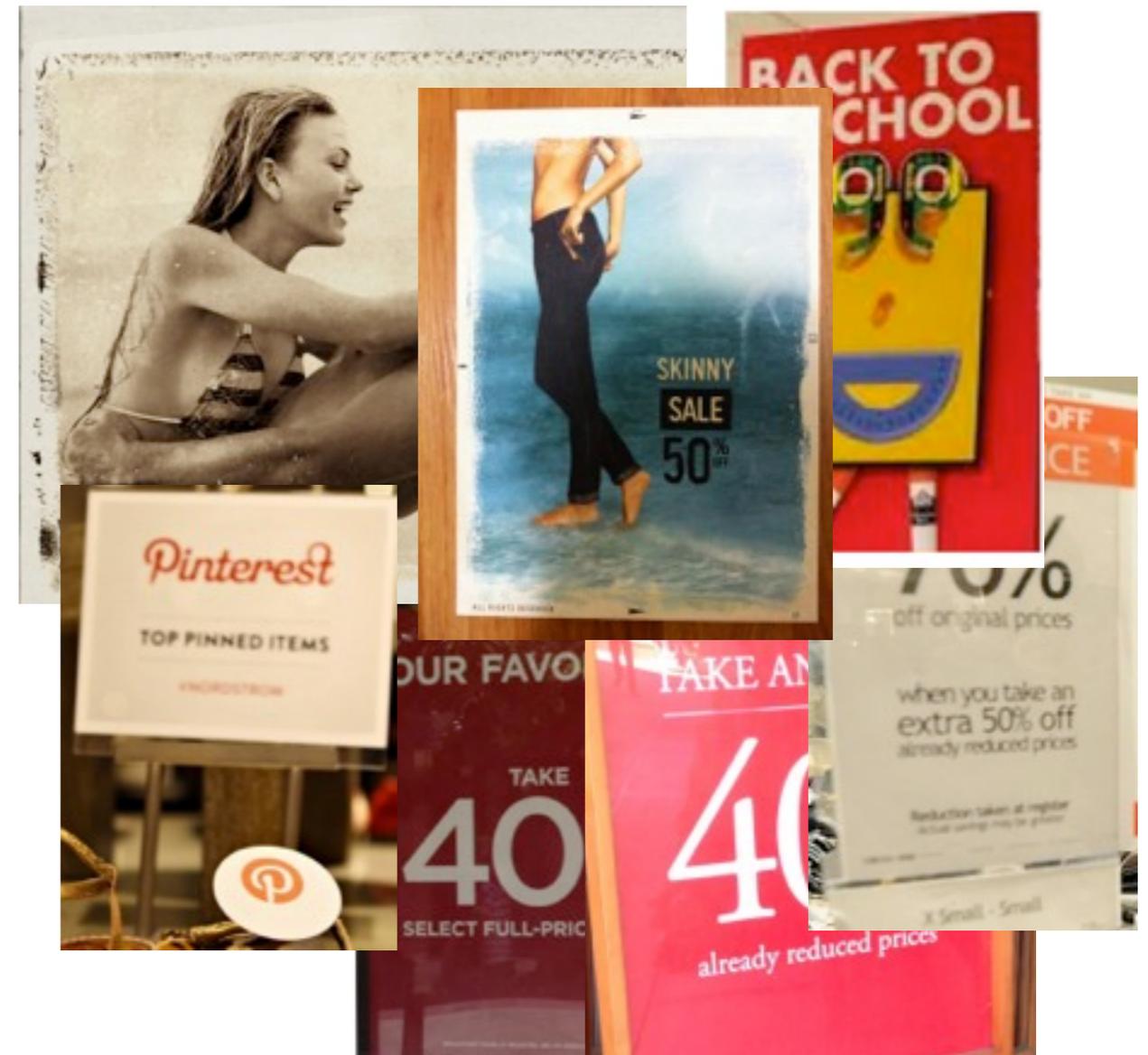
Tuesday, February 18, 14

Ambient awareness

We showed photos of signs that **were** at the stores and that **were not**.

Signs that were present

Signs that were not in-store

Tuesday, February 18, 14

create
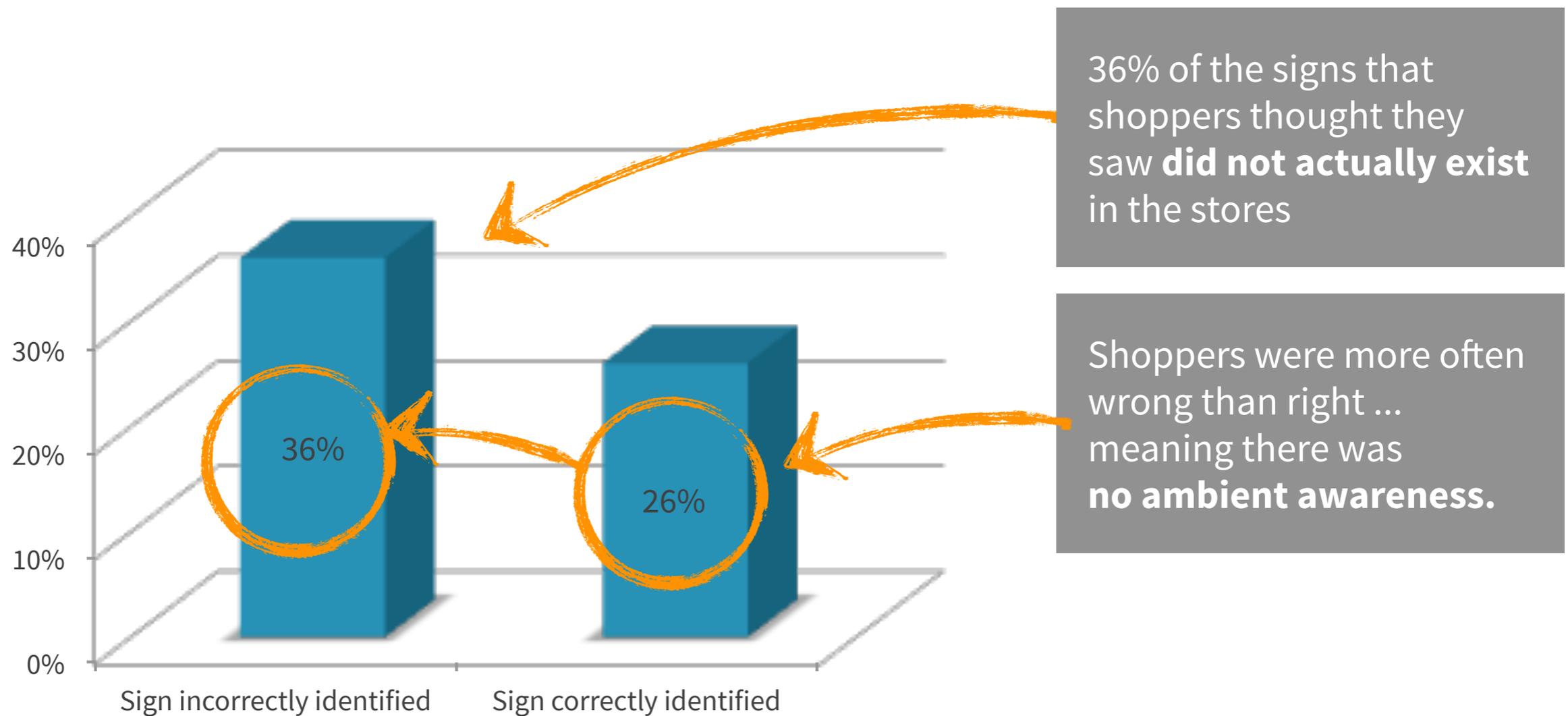with context

They are wrong more often than right

People were more likely to 'recall' signs that they had not actually seen in the store, compared to signs that were present in the store.

36% of the signs that shoppers thought they saw **did not actually exist** in the stores

Shoppers were more often wrong than right … meaning there was **no ambient awareness.**

40%

30%

20%   36%

10%   26%

0%

Sign incorrectly identified   Sign correctly identified

Tuesday, February 18, 14

The focus is not signs

Tuesday, February 18, 14

The focus is not signs

Ann is amazed by how much she didn't notice

Tuesday, February 18, 14

Why is signage ignored?

People ignore signs because they are in 'get in, get out' mode or they are enjoying the experience, focusing on the products and the process.

Get in and get out
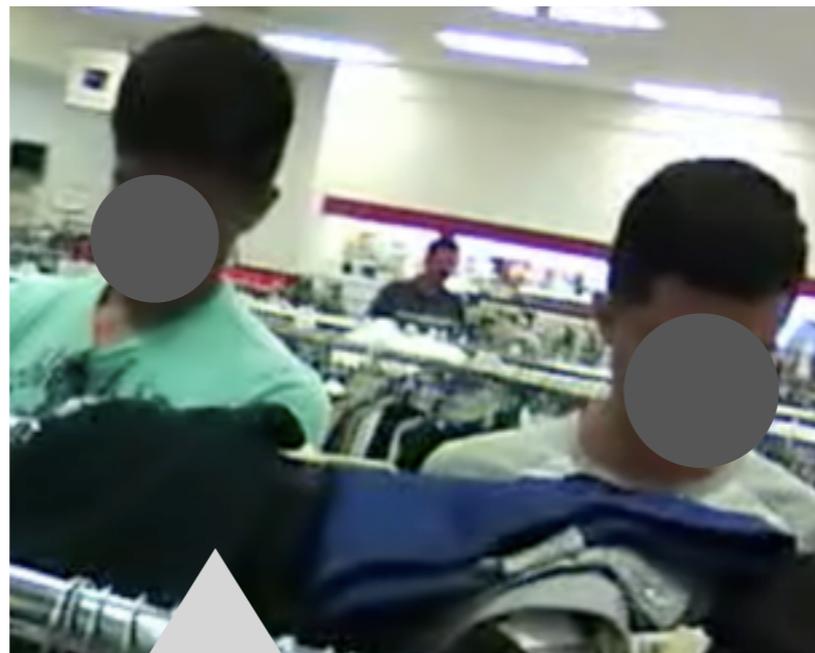
Focus is on more on the activity; low interaction with companion

Focus is on the products, the experience

Tuesday, February 18, 14

create
with context

Factors that increase attention

Part of shopping activity

(1) **Context**

Sign needs to be part of shopping activity

High repetition and easy to parse

(2) **High Opportunity**

When there is sign repetition and when it is at eye level

(3) **At a Glance**

It needs to be easy to parse

Tuesday, February 18, 14

create
with context

Should we message on the phone?

84% of smartphone shoppers* use their phones to help them shop while **in store**.

However, how **often** do they do this?

We researched the number of people who had **phones visible** in-store vs. in-mall to determine if this would be a solution.

*People who use a smartphone to assist with shopping at least once a month or more. Reference: http://www.google.com/think/research-studies/mobile-in-store.html

Tuesday, February 18, 14

Phones were visible for only 11% of people in the store

As a result, they most likely won't see any messages or alerts sent to their phones.

11%

Tuesday, February 18, 14
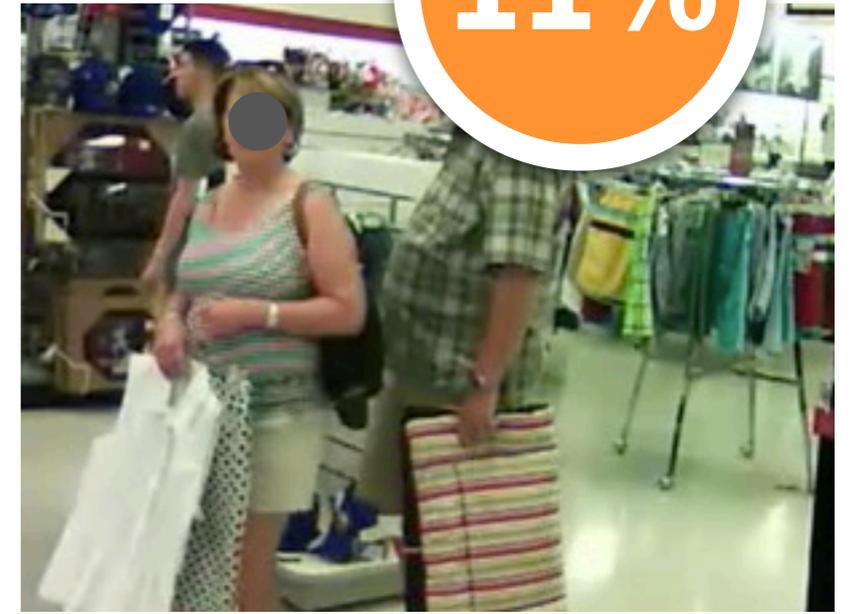
Phone were visible for 30% of people in mall areas

Higher chance for people to see messages or alerts sent to their phones.

30%

Tuesday, February 18, 14

create
with context
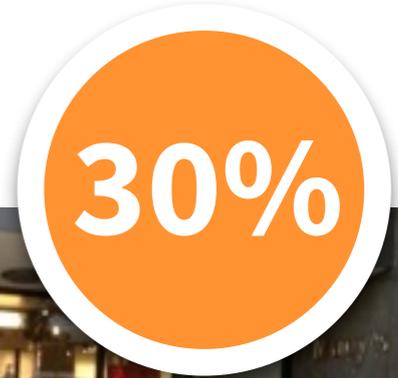
Why? Their hands were busy...

Tuesday, February 18, 14

So should we **not** provide notice on signs and devices?

Putting notices on signs and devices does not work to create initial widespread awareness of data collection.

Rather, once people become more aware of data collection, notices can reinforce what is happening – and notify people that they can interact when (and if) they want to.

Tuesday, February 18, 14

We can create awareness – over time

(1) **Implicit Awareness**

No need for notice; people understand that data is being collected.

> "Of course Maps has my location, how else would the service work?"

(2) **Explicit Awareness**

Direct communication from advertisement or just-in-time notification.

> "Yes, I want them to have my location so I can see movies that are playing close to me."

(3) **Ambient Awareness**

In the background to reinforce and maintain awareness from signs and symbols.

> "I don't recall seeing anything, I just know that they have my location which I am fine with."

Tuesday, February 18, 14

Implicit awareness

Hypothetical scenario

You download a wish list app, and friends and family subscribe and post the gifts they want.

The app alerts you when you're in a store that has a wanted item: "Retailer X has the perfume your Mom wants."

Then, when shopping online, you get an alert: "People who like the perfume you bought for your mother also bought this sweater."

Perfect, now you can surprise her with something she'll like!

Wish list alert

This store has the perfume that Esther Johnson wants

Show me

not now
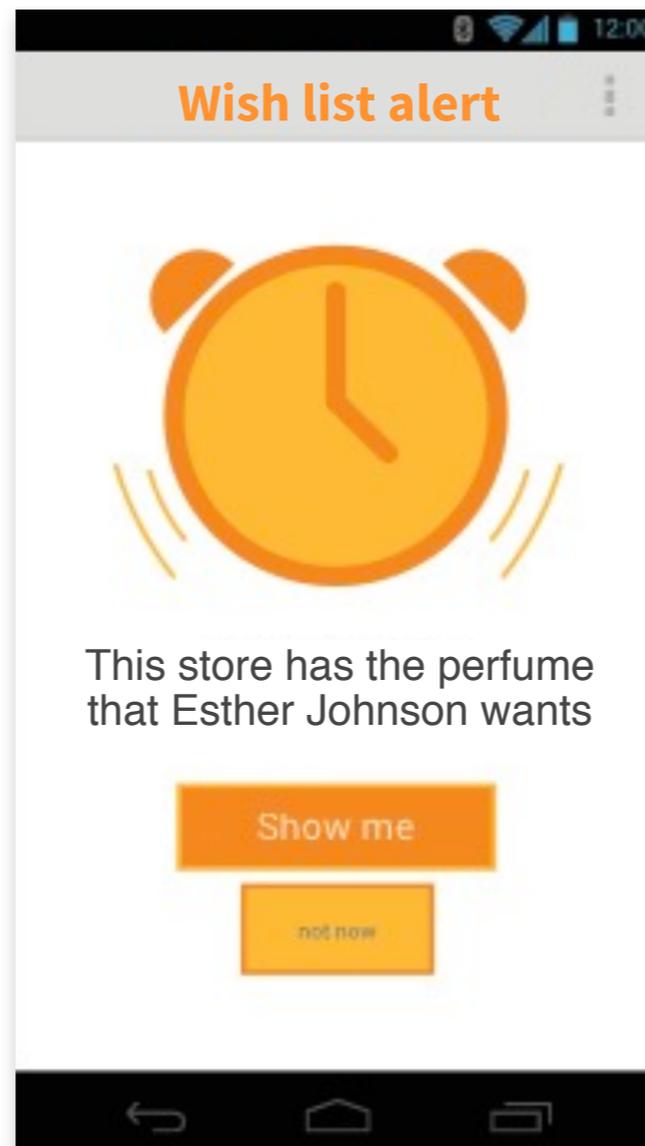
The outcome

Of course you **knows** my **location**, how else would the alert me.

Of course you **know** my **social network**, I asked them to join.

You **know** who I am, on my **phone** and on my **computer**.

I give you **permission** to use that **information** to provide me **value** in the **future**.

Tuesday, February 18, 14

"I want that"

Tuesday, February 18, 14

"I want that"

Ellen wants a reminder like that, it would save her time

Tuesday, February 18, 14

Explicit awareness

Notification can get in the way of an experience. However, asking for permission in-the-moment (when data is needed to complete the activity) has a higher likelihood of people paying attention.
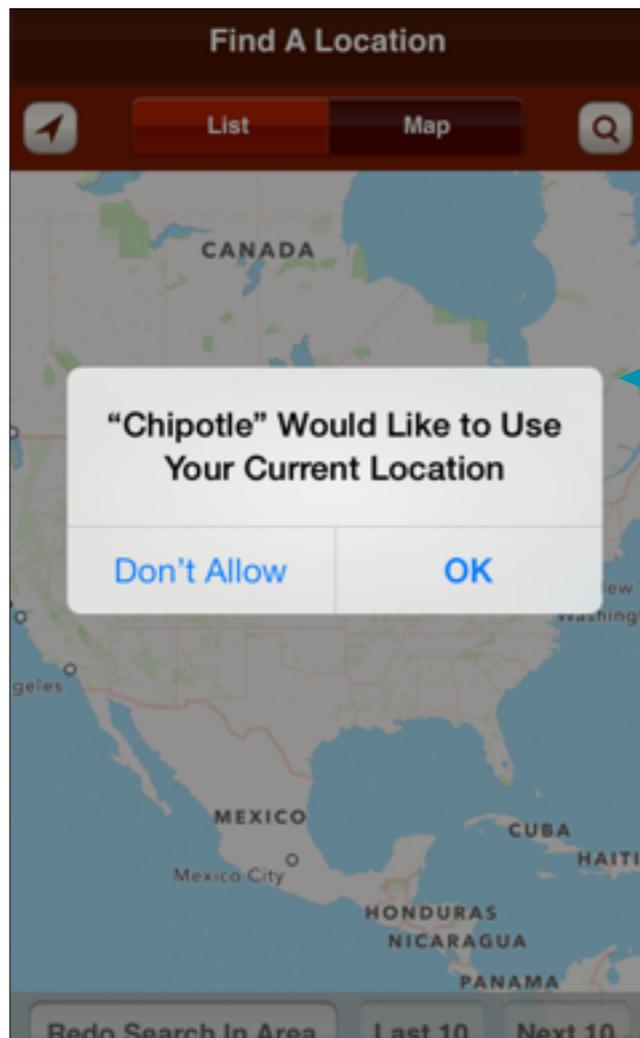
Example: When the user wants to place an order, ask them "in the moment" to use their location information

Example: Carts could have a dock which holds and charges the mobile phone. Shopper would be able to see alerts when they appear on the screen.

Tuesday, February 18, 14

Ambient Awareness

A suite of integrated, consistent notifications across platforms and contexts – using visual, auditory, and tactile feedback to create awareness.

Tactile

Vibration for wearables and mobile devices

Visual

Symbol for screens and signs
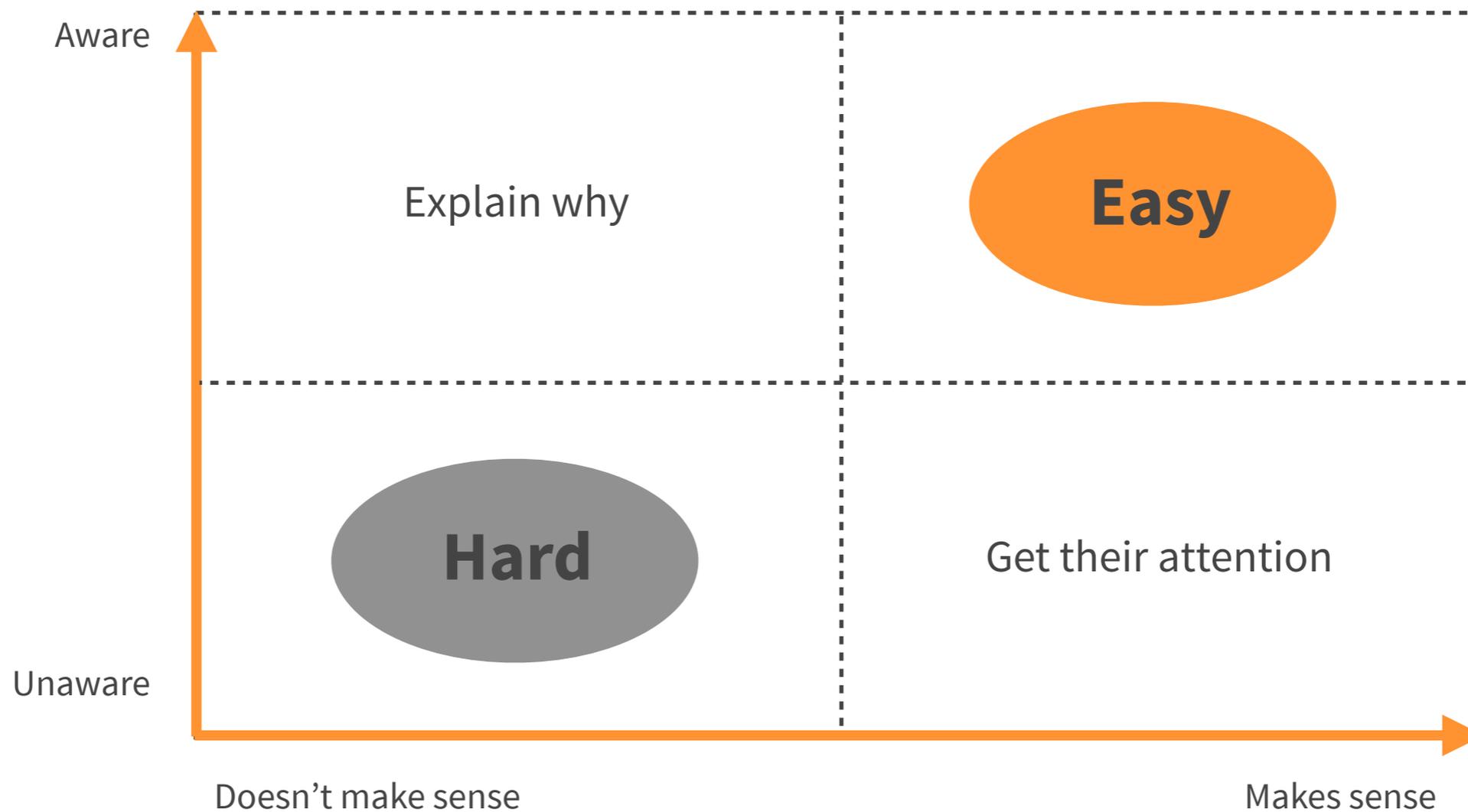
Auditory

Pleasant tone for physical contexts

Tuesday, February 18, 14

create
with context

Design challenge – Ambient awareness

Communicating a concept to people – when they don't know it's happening, it doesn't make sense to them, and you need to inform them – is a difficult design challenge.

Aware

Explain why

Easy

Unaware

Hard

Get their attention

Doesn't make sense

Makes sense

Tuesday, February 18, 14

Ambient awareness – My data symbol

Create a graphical symbol that reinforces that your personal data is being transmitted.

Communicate

Collecting

My data

Positive

Sending

Tracking

Transmitting

Ensure flexibility

Physical signage or on any screen, all sizes

What is collected now, and in the future

State on/off

Ability to signify that it is or is not happening

Immediate awareness

No need for labels

Learned awareness

Initial need for label, then becomes known

Tuesday, February 18, 14

create
withcontext

Good design doesn't just happen – It is a process
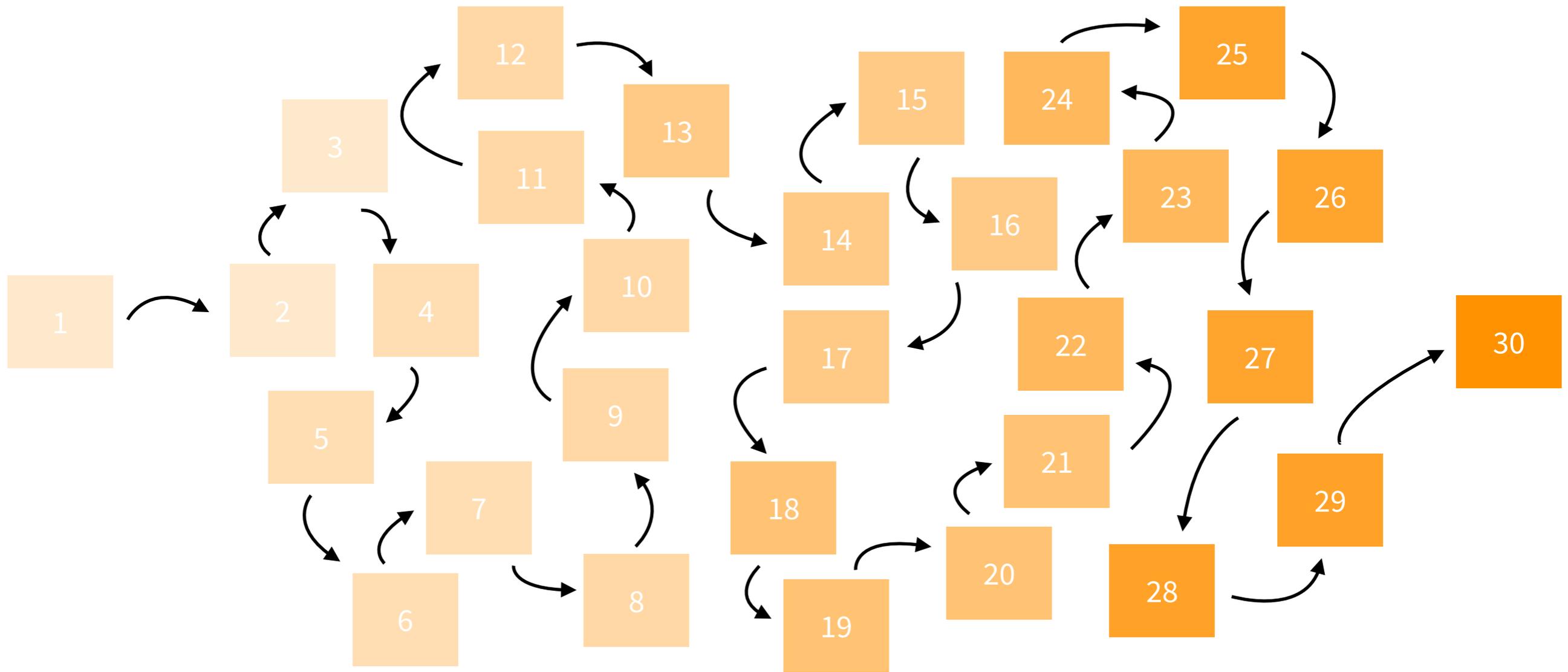
MODEL ➜ TEST ➜ ITERATE ➜ CREATE ➜ TEST ➜ ITERATE ➜ SOLUTIONS ➜ TEST ➜ REFINE ➜

Tuesday, February 18, 14

Design – Multiple concepts

Tuesday, February 18, 14

create with context

Design – Multiple concepts

Tuesday, February 18, 14

create
withcontext

Design – Multiple concepts

Tuesday, February 18, 14

create with context

Design – Multiple concepts

Tuesday, February 18, 14

create
with context

Design – Multiple concepts

Tuesday, February 18, 14

Design – Multiple concepts

Tuesday, February 18, 14

Design – Multiple concepts

Tuesday, February 18, 14

create with context

Design – Multiple concepts

Tuesday, February 18, 14

Design – Multiple concepts

Tuesday, February 18, 14

Design – Multiple concepts

Tuesday, February 18, 14

Design – Multiple concepts

Tuesday, February 18, 14

create
withcontext

Design – Multiple concepts

Tuesday, February 18, 14

Design – Multiple concepts

Tuesday, February 18, 14

Design – Multiple concepts

Tuesday, February 18, 14

Design – Multiple concepts

Tuesday, February 18, 14

Design – Multiple concepts

Tuesday, February 18, 14

Design – Multiple concepts

Tuesday, February 18, 14

Design – Multiple concepts

Tuesday, February 18, 14

Design – Multiple concepts

Tuesday, February 18, 14

Design – Multiple concepts

Tuesday, February 18, 14

create with context

Design – Multiple concepts

Tuesday, February 18, 14

Design – Multiple concepts

Tuesday, February 18, 14

Design – Multiple concepts

create
withcontext

Design – Multiple concepts

Tuesday, February 18, 14

Design – Multiple concepts

Tuesday, February 18, 14

Design – Multiple concepts

Tuesday, February 18, 14

Design – Multiple concepts

Tuesday, February 18, 14

Design – Multiple concepts

Tuesday, February 18, 14

Design – Multiple concepts

Tuesday, February 18, 14

create**with**context

Design – Multiple concepts

Tuesday, February 18, 14

Design – Multiple concepts

Tuesday, February 18, 14

create**with**context

Design – Multiple concepts

Tuesday, February 18, 14

Design – Multiple concepts

Tuesday, February 18, 14

Design – Multiple concepts

70

Tuesday, February 18, 14

My data symbol – Current iteration

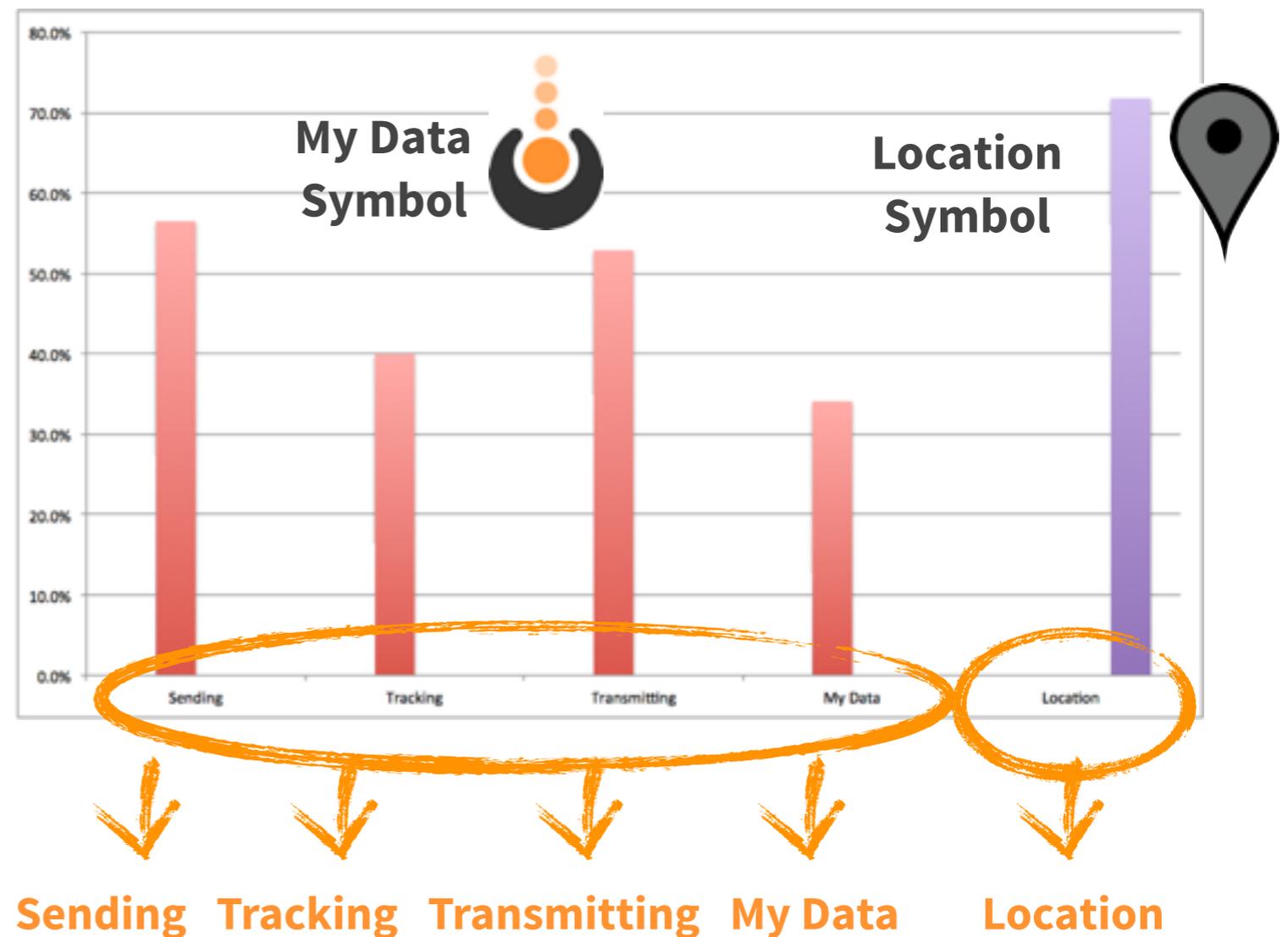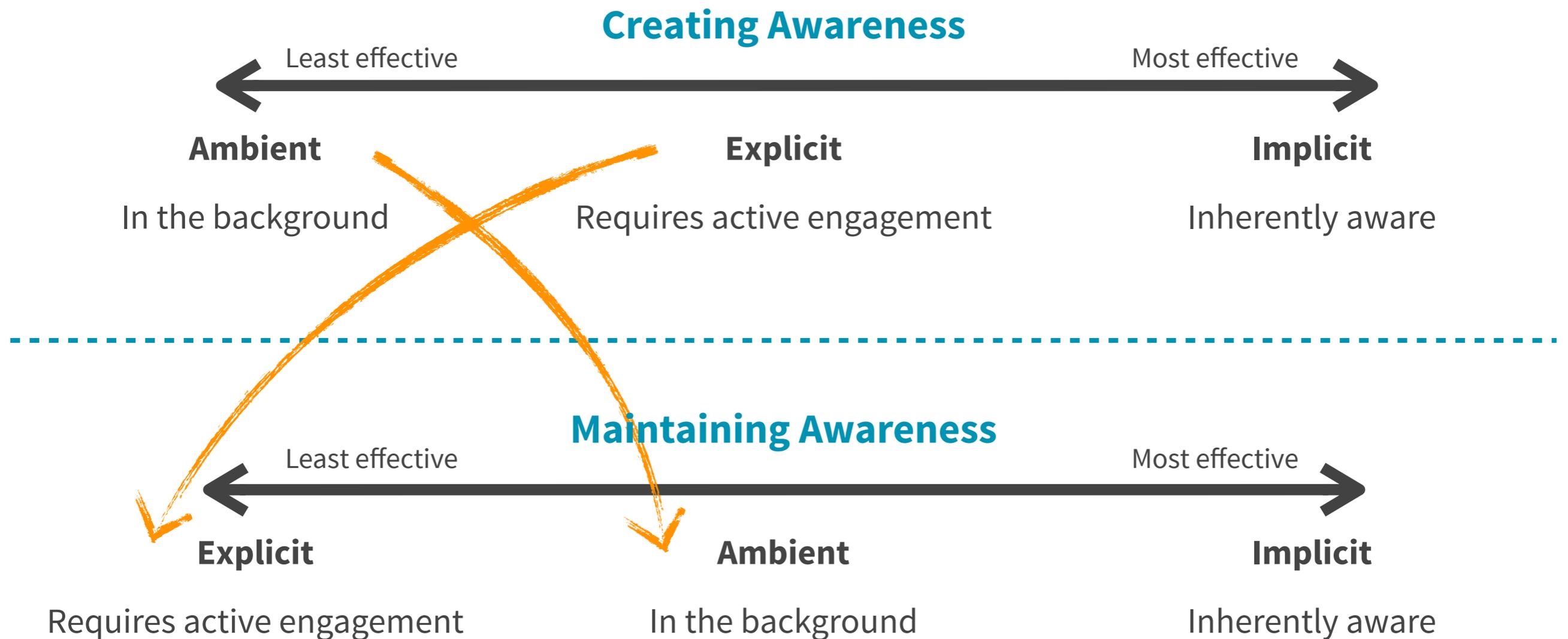People were asked to select which words best describe the visual from a list of words. A highly recognizable symbol is identified between 70-75% of the time.

Access	Dynamic	Irrelevant
Alive	Easy	**My data**
Appealing	Empowering	Not Valuable
Approachable	Encouraging	Overwhelming
Attractive	Energetic	Personal
Authentic	Engaging	Positive
Beneficial	Entertaining	Receiving
Busy	Exciting	Reliable
Choice	Friendly	Reputable
Collecting	Frustrating	Rigid
Community	Giving	Safe
Compelling	Harmful	**Sending**
Complex	Helpful	Signal
Confidential	Honest	Sharing
Confusing	Human	Taking
Connected	Impersonal	**Tracking**
Convincing	Ineffective	**Transmitting**
Coverage	Informative	Trustworthy
Credible	Innovative	Unattractive
Cutting edge	Inspiring	Understandable
Disconnected	Interesting	Undesirable
Disruptive	Intimidating	Unique
Distracting	Inviting	Useful

My Data Symbol   **Location Symbol**

Sending   **Tracking**   **Transmitting**   **My Data**   **Location**

71

Tuesday, February 18, 14

When people become aware, we shift to maintaining that awareness

Explicit messaging becomes **disruptive** once people are **actively aware**.

Creating Awareness

Least effective ⟵   Most effective ⟶

Ambient   **Explicit**   **Implicit**

In the background   Requires active engagement   Inherently aware

- -

Maintaining Awareness

Least effective ⟵   Most effective ⟶

Explicit   **Ambient**   **Implicit**

Requires active engagement   In the background   Inherently aware

Tuesday, February 18, 14

Thank you!

Contact

Website
createwithcontext.com

General inquiries
info@cwcmail.com

Media
media@cwcmail.com

Employment
jobs@cwcmail.com

Phone
+1 408 834-7601 main
+1 408 834-7602 fax

Silicon Valley office
Create with Context
1490 Lincoln Street
Santa Clara, CA 95050 USA

Panel Discussion

- **Mallory Duncan**, Senior Vice President and General Counsel, National Retail Federation

- **James Riesenbach**, Chief Executive Officer, iInside

- **Seth Schoen**, Senior Staff Technologist, Electronic Frontier Foundation

- **Glenn Tinley**, President and Founder, Mexia Interactive

- **Ilana Westerman**, Chief Executive Officer, Create with Context