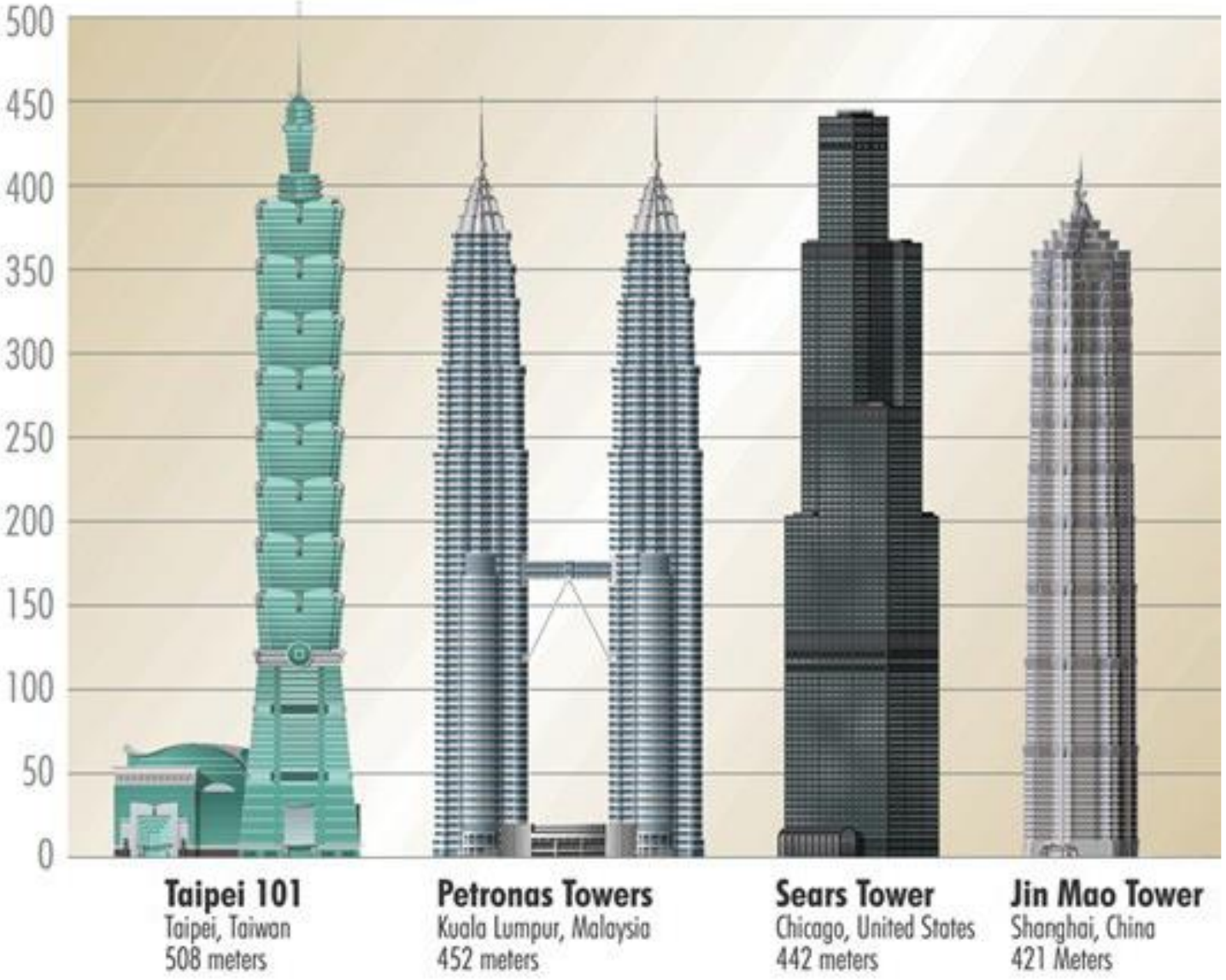
The Mobile Phone as a (Massive) Data Collection Platform

Jon Froehlich
jfroehli@cs.washington.edu
PhD Student
Computer Science and Engineering
University of Washington

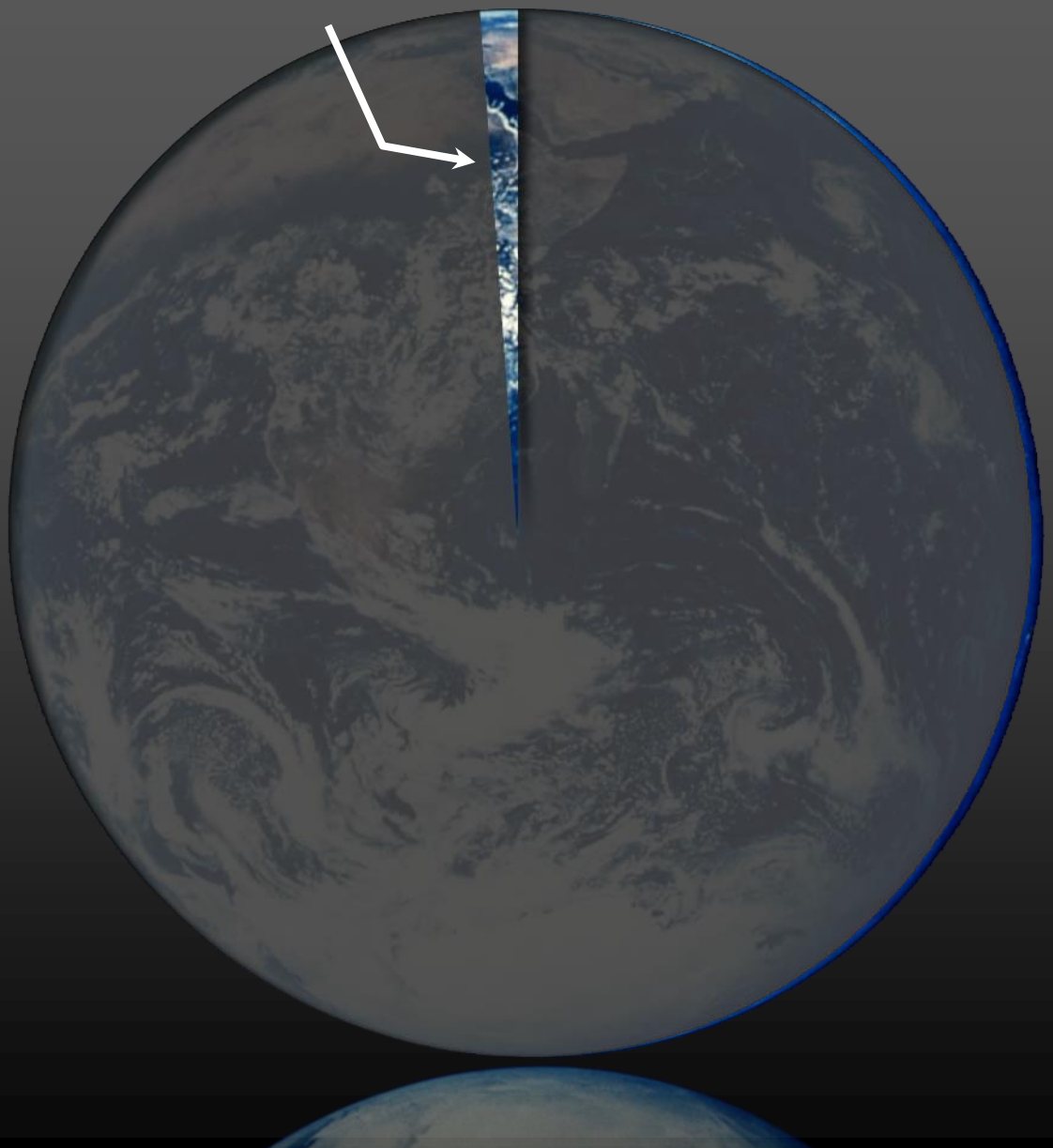
dub design: use: build:
university of washington

computer science & engineering
university of washington


500
450
400
350
300
250
200
150
100
50
0
Taipei 101
Taipei, Taiwan
508 meters
Petronas Towers
Kuala Lumpur, Malaysia
452 meters
Sears Tower
Chicago, United States
442 meters
Jin Mao Tower
Shanghai, China
421 Meters

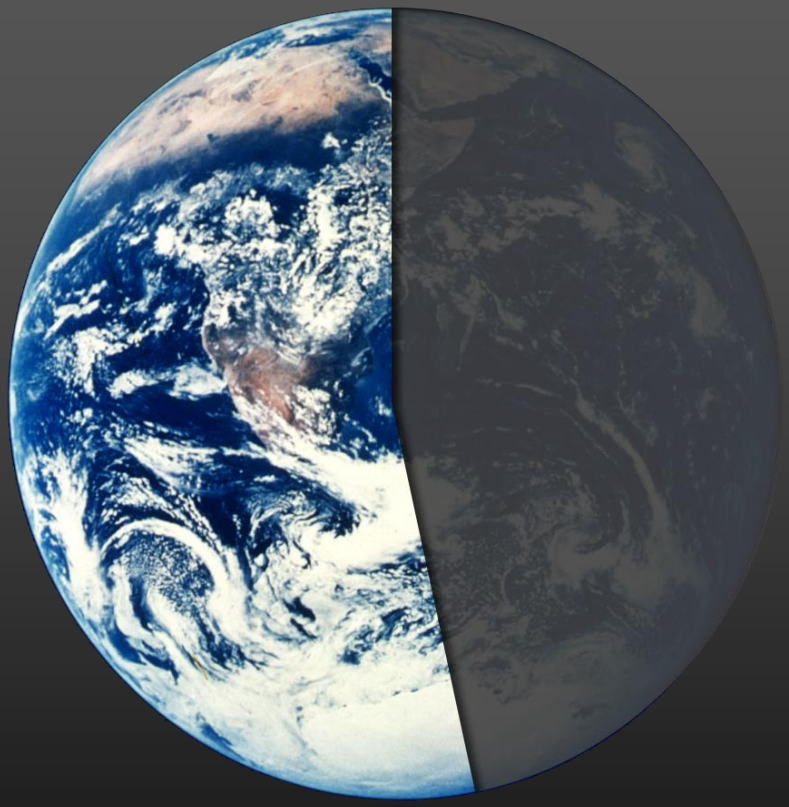
NOKIA
N95
L'Obélisque
Options
Exit
cingular
9:41 AM
Tuesday
9
Text
Calendar
Photos
Camera
Calculator
Stocks
Maps
Weather
Notes
Clock
Settings
Phone
Mail
Web
iPod
Start
Weather
30 minutes ago
London, UK
27°
High: 30°
Low: 25°
Weekly Meeting
1:00PM-2:00PM
Calendar
Contacts
htc

Mobile Phone Subscribers Worldwide 1995



2010 Worldwide Projections

51.47% Mobile Phone Subscribers



19.12% PC Users

The mobile phone is positioned to revolutionize the way we can collect data about people, their behavior and the environment.

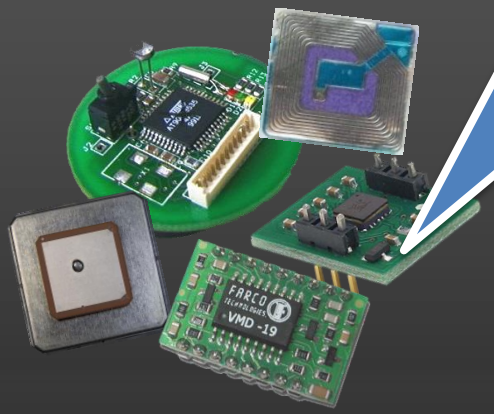
challenges

- How do we collect the data?
- How do we incentivize people to contribute data?
- How can we sustain participation over time?
- How do we handle privacy issues?
- How do we analyze/filter the data?

the myexperience tool

http://myexperience.sourceforge.net

sensors



Device usage and
states (*e.g.,* GPS) c
automatically sens

+ Technique scales
– Cannot capture
perception, r

Users respond to short context-triggered surveys on their mobile device.

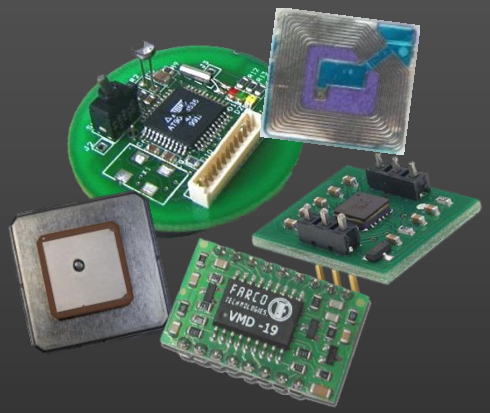
+ Can gather otherwise imperceptible data (both qual. and quant.)
– Lower sampling rate than sensors (increased user burden)

context | self-report | myexperience

MyExperience is open source software under the BSD license

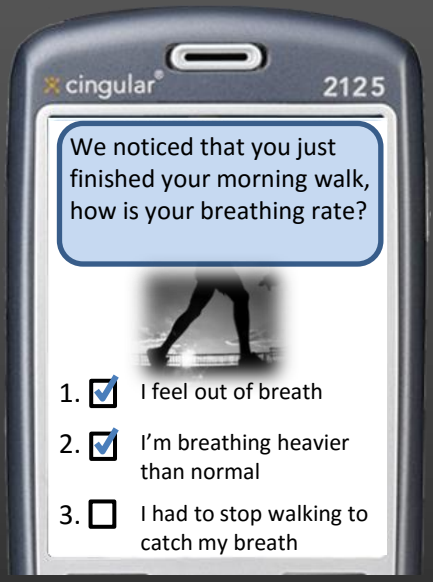
sensors, triggers, actions

Sensors



Triggers

Actions



Example Sensor:

- DeviceIdleSensor
- PhoneCallSensor
- RawGpsSensor
- SmsSentSensor
- HumanScaleActivitySensor

Example Triggers:

- DeviceIdle > 15 mins
- PhoneCall.Outgoing == true
- Gps.Longitude ==“N141.23”
- SmsSent == true
- Activity.StateExited == Walking

Example Actions:

- ScreenshotAction
- VibrationAction
- SmsSendAction
- DatabaseSyncAction
- SurveyAction

hardware sensors

GPS Sensors

GpsLatLongSensor
GpsRecordSensor
GpsSpeedSensor
GpsSpeedWindowSensor
GpsSustainedSpeedSensor

GSM Sensors

GsmCommonCellRatioSensor
GsmCellSensor
GsmMotionSensor
PhoneSignalStrengthSensor

MSP Sensors

ActivityProbabilitySensor
MspConnectionSensor

software sensors

Device Usage Sensors

ButtonSensor
ActiveApplicationSensor
ForegroundWindowSensor
DeviceIdleSensor
MediaPlayerSensor

Device State Sensors

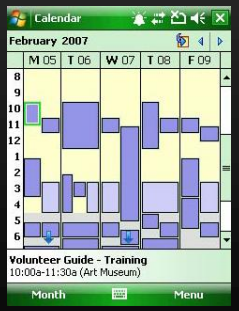
StorageCardFullSensor
PhoneProfileSensor
PowerLineSensor
PhoneRoamingSensor

Communication Sensors

IncomingCallSensor
OutgoingCallSensor
SmsSentSensor
SmsReceivedSensor

Meta-Sensors

ActionCompletedSensor
ActionStartingSensor
GlobalsSensor
MyExpStartingUpSensor
MyExpShuttingDownSensor



Calendar Sensors

CalAppointmentSensor
CalAppointmentLocationSensor
CalAppointmentSubjectSensor

Action Type	Summary Description
CreateProcessAction	Launches an additional process (this can be any executable that is local to the device).
KillProcessAction	Kills an existing process.
MessageAction	Displays a message in a dialog box to the user.
NotificationAction	Displays a notification in a dialog box to the user with a sound and/or vibration alert. User may respond “OK” or “Dismiss.”
PlayerAction	Plays a sound, vibrates the device, and/or flashes the device’s LEDs.
RecordAudioAction	Records audio in the background using the device’s microphone.
RestartDeviceAction	Restarts the device.
ScreenShotAction	Takes a screen shot of the current screen on the device.
SendSmsAction	Sends an SMS to a specified address.
SqlReplicationAction	If SQL Replication is setup, this action invokes a data replication with the master web server.
SurveyAction	Displays a survey to the user.

Action Type	Summary Description
CreateProcessAction	Launches an additional process (this can be any executable that is local to the device).
KillProcessAction	Kills an existing process.
MessageAction	Displays a message in a dialog box to the user.
NotificationAction	Displays a notification in a dialog box to the user with a sound and/or vibration alert. User may respond “OK” or “Dismiss.”
PlayerAction	Plays a sound, vibrates the device, and/or flashes the device’s LEDs.
RecordAudioAction	Records audio in the background using the device’s microphone.
RestartDeviceAction	Restarts the device.
ScreenShotAction	Takes a screen shot of the current screen on the device.
SendSmsAction	Sends an SMS to a specified address.
SqlReplicationAction	If SQL Replication is setup, this action invokes a data replication with the master web server.
SurveyAction	Displays a survey to the user.

Surveys do not necessarily have to ask simple closed form questions (e.g., Likert scale questions)

xml / scripting interface

- XML : Declarative
  - Define sensors, triggers, actions, and user interface
  - Set properties
  - Hook up events

```
<sensor name="Motion" type="GsmMotionSensor">
 <prop name="PollInterval">00:00:01</prop>
</sensor>
```

- Script : Procedural
  - Create fully dynamic behaviors between elements specified in XML
  - Interpreted in real time
  - New scripts can be loaded on the fly

```
<trigger name="Motion" type="Trigger">
  <script>
    motionSensor = GetSensor("Motion");
    if(motionSensor.StateEntered = "Stationary"){
      ... do some action ...
    }
  </script>
</trigger>
```

votewithyourfeet

Jon Froehlich[1,2], Mike Chen[2], Ian Smith[2], and Fred Potter[1,2]

dub **design: use: build:**

[1] university of washington

[2] Intel Research, Seattle

Amazon.com: Recommended for You - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://www.amazon.com/gp/yourstore/ref=pd_irl_gw/103-9546741-6627009

Google TiVo suggestions Search PageRank Check AutoLink AutoFill Subscribe Options TiVo

Recommended for Jon E. Froehlich (If you're not Jon E. Froehlich, click here.)

Recommendations Based on Activity

View & edit Your Browsing History

Recommendations by Category

Your Favorites Edit

Music
Electronics
Books
Computers & PC Hardware
Camera & Photo

More Categories

Apparel & Accessories
Baby
Beauty
Computer & Video Games
DVD
Gourmet Food
Health & Personal Care
Industrial & Scientific
Jewelry & Watches
Kitchen & Housewares
Magazine Subscriptions
Outdoor Living
Software

These recommendations are based on items you own and more.

view: All | New Releases | Coming Soon

More results

1.

Context and Consciousness: Activity Theory and Human-Computer Interaction
by Bonnie A. Nardi (Editor)
Average Customer Review: ★★★★☆
In Stock
Publication Date: November 21, 1995

Our Price: $68.00 **Used & new** from $47.92

Add to cart Add to Wish List

I Own It Not interested x|☆☆☆☆☆ Rate it

Recommended because you purchased Plans and Situated Actions and more (edit)

Explicit
Implicit

2.

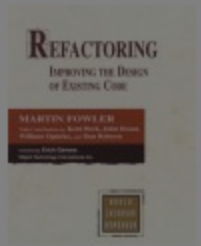
The Ecological Approach to Visual Perception
by James Jerome Gibson
Average Customer Review: ★★★★★
In Stock
Publication Date: January 1, 1987

Our Price: $46.45 **Used & new** from $32.93

Add to cart Add to Wish List

I Own It Not interested x|☆☆☆☆☆ Rate it

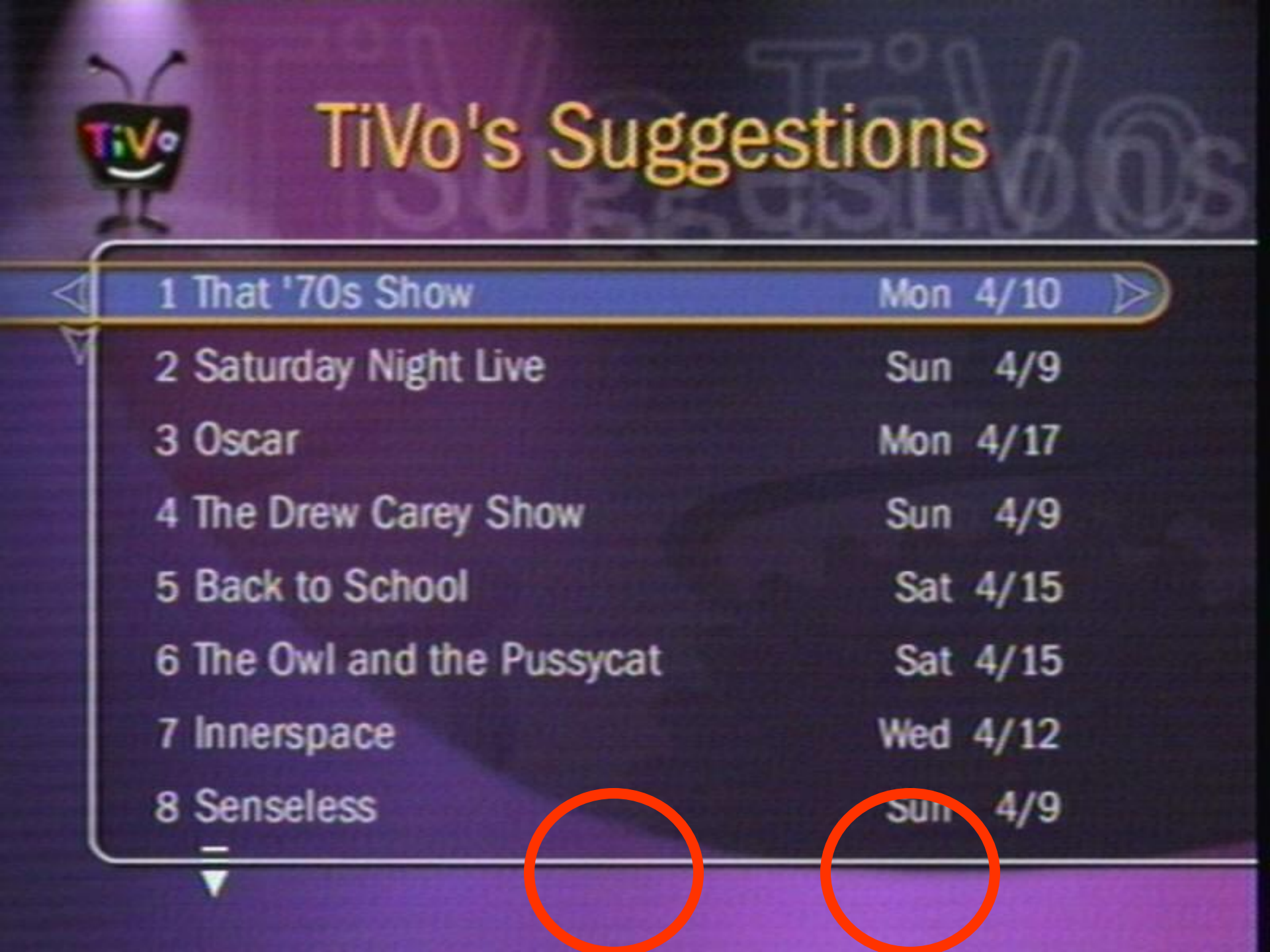
Recommended because you purchased Cognition in the Wild (Bradford Books) and more (edit)

3. 

Refactoring: Improving the Design of Existing Code
by Martin Fowler, et al.
Average Customer Review: ★★★★½
In Stock
Publication Date: June 28, 1999

Our Price: $46.79 **Used & new** from $27.99

Add to cart Add to Wish List


TiVo's Suggestions
1 That '70s Show Mon 4/10
2 Saturday Night Live Sun 4/9
3 Oscar Mon 4/17
4 The Drew Carey Show Sun 4/9
5 Back to School Sat 4/15
6 The Owl and the Pussycat Sat 4/15
7 Innerspace Wed 4/12
8 Senseless Sun 4/9

translation to physical world

Can we view place visit behaviors as an implicit form of expressing interest?

explicit vs. implicit indicators

VS.

Explicit Indicators

- Move about the world → Supply rating “tags”
- Requires device interaction

Implicit Indicators

- Location aware device → Observe travel patterns
- No device Interaction

cingular 2125
Google Local
Gmail
Inbox (213 unread)
T. Scott How is UbiComp?
J. Lestr Lunch tomorrow?
Suggested Places
Zoka's Coffee
Why was this suggested?
1.2 miles away
Opened last week
Route me (map)
Seth's Guitars
Why was this suggested?
2 blocks east
Opened Oct. 2005
cingular 2125
myspace.com
a place for friends
T. Scott
has 129 friends
Shares favorite public locations with you.
Top Favorite Places
1. Discovery Park
Last Visited: 2 days ago
Parks like this near you
2. Jackson Golf
Last Visited: 3 weeks ago
Golf courses near you
cingular 2125
You Are Here
Fashion Island
Newport Center Dr
Fashion Is
San Miguel Dr
Anacapa Dr
Farallon Dr
1 2ABC 3DEF
4GHI 5JKL 6MNO
7PQRS 8TUV 9WXYZ
*T9 0 + #

study overview

- Four week study
- Participant profile
  - 16 Participants
  - Gender balanced (8 male / 8 female)
  - Ages: 22-56 (median 29)
  - Various professions
    - Furniture designer, political consultant, bookseller, translator, …
- Tasked with
  - Carrying mobile phone for four weeks
  - Answering 11 *in situ* surveys a day about current place
  - Completing a minimum of 3-4 web diary entries a week
- Compensation
  - $1 per survey
  - Also compensated for interviews & paper questionnaires

survey triggers

Pseudo-Random Time Trigger

- No movement detected for 1 hr → trigger survey randomly within next hour
- Ensures consistent sampling in case of sensor failure

Sohn et al. *Mobility Detection Using Everyday GSM Traces.* UbiComp2006.

AUDIOVOX
Survey #1
Please click the OK button to take Survey #1. This message box will autosnooze in 1 minute.You have 1 reminder left before automatic dismissal.
OK
Menu
AUDIOVOX
1. Place name:
3 Pigs Bar-B-Q Bellevue
Key Arena
Kona Kitchen-Seattle
LexisNexis-Applied D...
Loews Theaters - Al...
McGrath's Fish Bar - ...
My House
Net Desk - Seattle
Back
Next
AUDIOVOX
2. Place category:
Restaurant
3. Please rate how much you like this place:
Back
Next
AUDIOVOX
4. How did you get here?
Car
5. How long did it take you to get here?
6 - 15 minutes
Back
Next

Fox Sports Grill on Thursday, September 8 ★★★☆☆ (3 stars)

1. What was the primary reason for your rating?

2. How did you find out about this place?

3. Why did you go to this place?

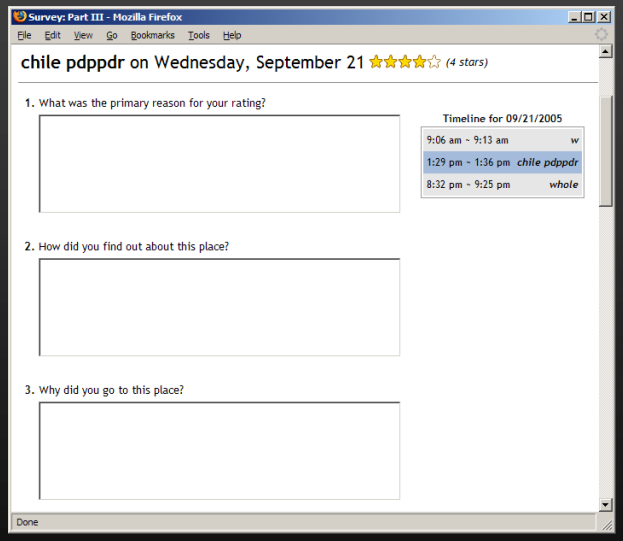
4. Would you recommend this place to others? Why or why not?

5. If you were with a group, how did the group decide to go to this place? If you were not with a group, type "N/A"

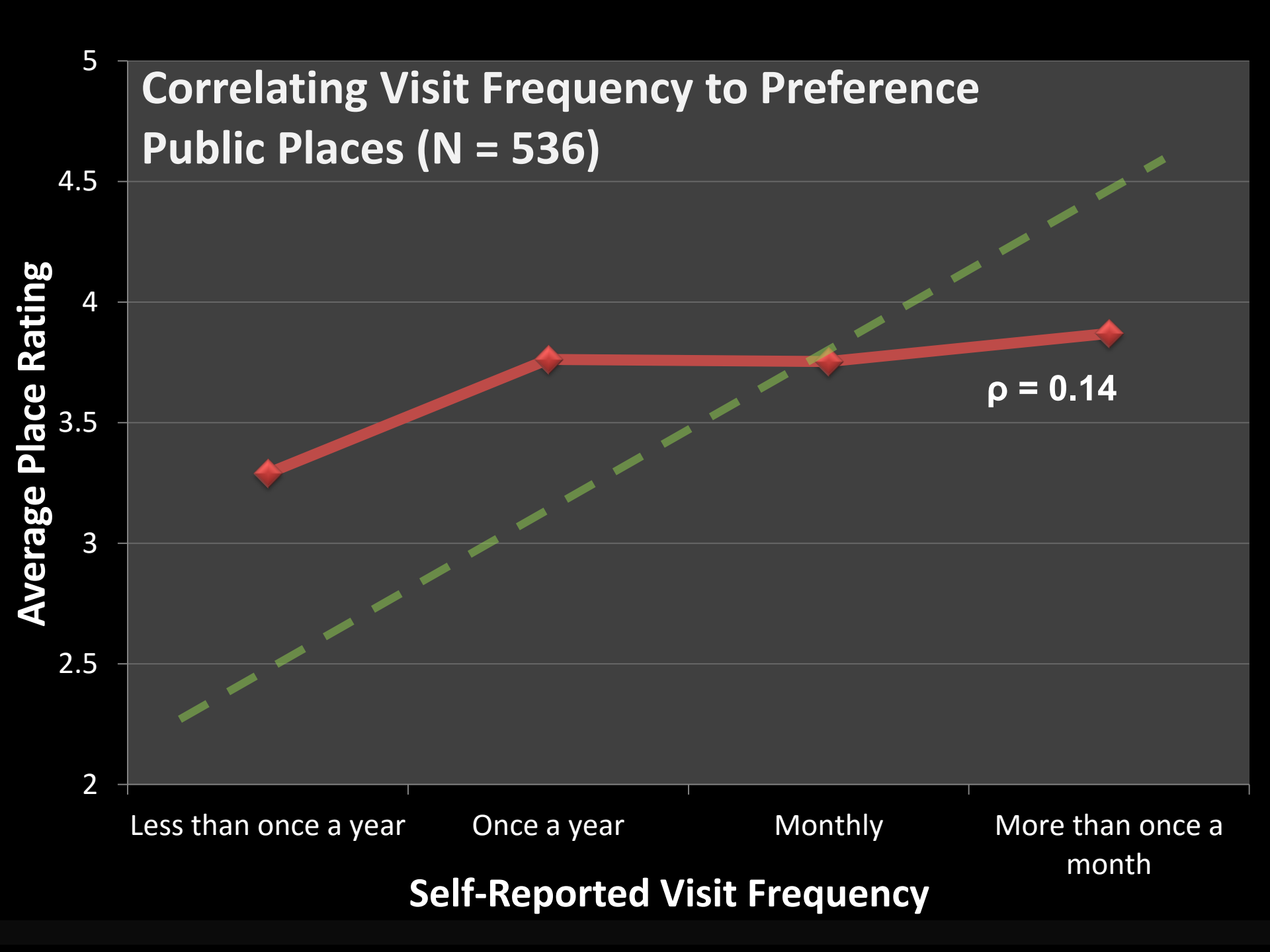
Timeline for 09/08/2005

5:32 pm (Wed) ~ 8:23 am	*My House*
8:30 am ~ 8:38 am	*Car*
9:12 am ~ 12:09 pm	*Net Desk - Seattle*
12:17 pm ~ 12:28 pm	*Walk*
12:28 pm ~ 12:46 pm	*Chez Dave - Union Square*
12:55 pm ~ 2:20 pm	*Walking*
2:20 pm ~ 3:48 pm	*Net Desk - Seattle*
3:48 pm ~ 3:50 pm	*Walking*
4:42 pm ~ 4:44 pm	*Rock Bottom-Seattle*
4:51 pm ~ 5:01 pm	*Walking*
6:35 pm ~ 6:39 pm	*Elephant and Castle*
6:46 pm ~ 7:34 pm	*Walking*
8:07 pm ~ 9:16 pm	*Fox Sports*

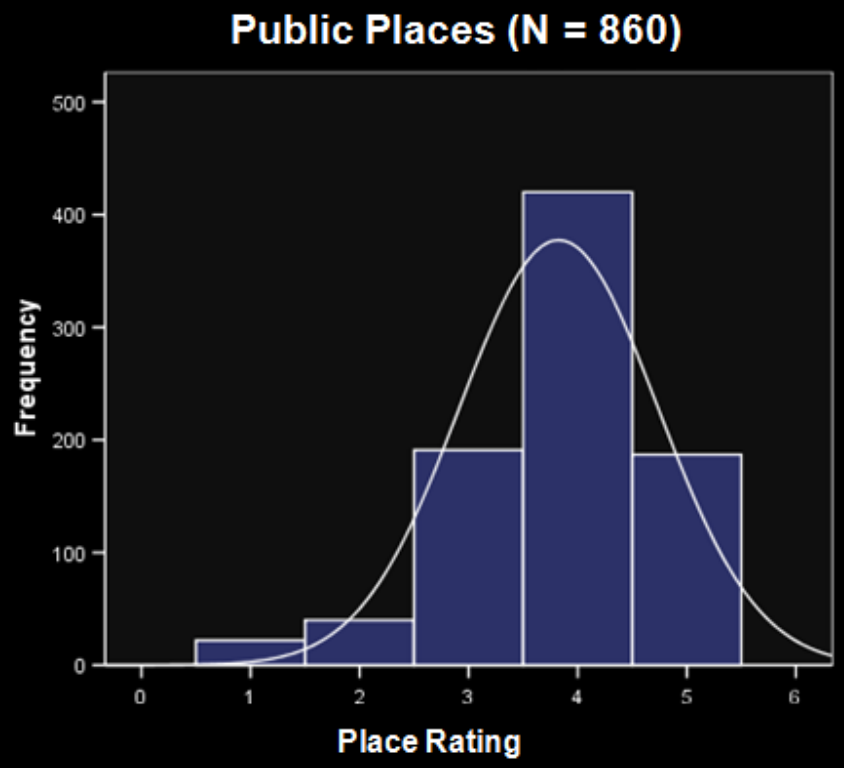
data collection stats



- ESM
  - 3,458 completed out of 4,295 (80.5%)
  - 216 surveys completed per person
  - 28 days average
  - 1.5 minute average completion time
- Web Diary
  - 368 web diary sessions completed
- Places
  - 1,981 individual place visits logged
  - 862 of which were public
    - ~2 a day per participant


Correlating Visit Frequency to Preference
Public Places (N = 536)
Average Place Rating
5
4.5
4
3.5
3
2.5
2
ρ = 0.14
Less than once a year
Once a year
Monthly
More than once a month
Self-Reported Visit Frequency

skewed distribution of ratings



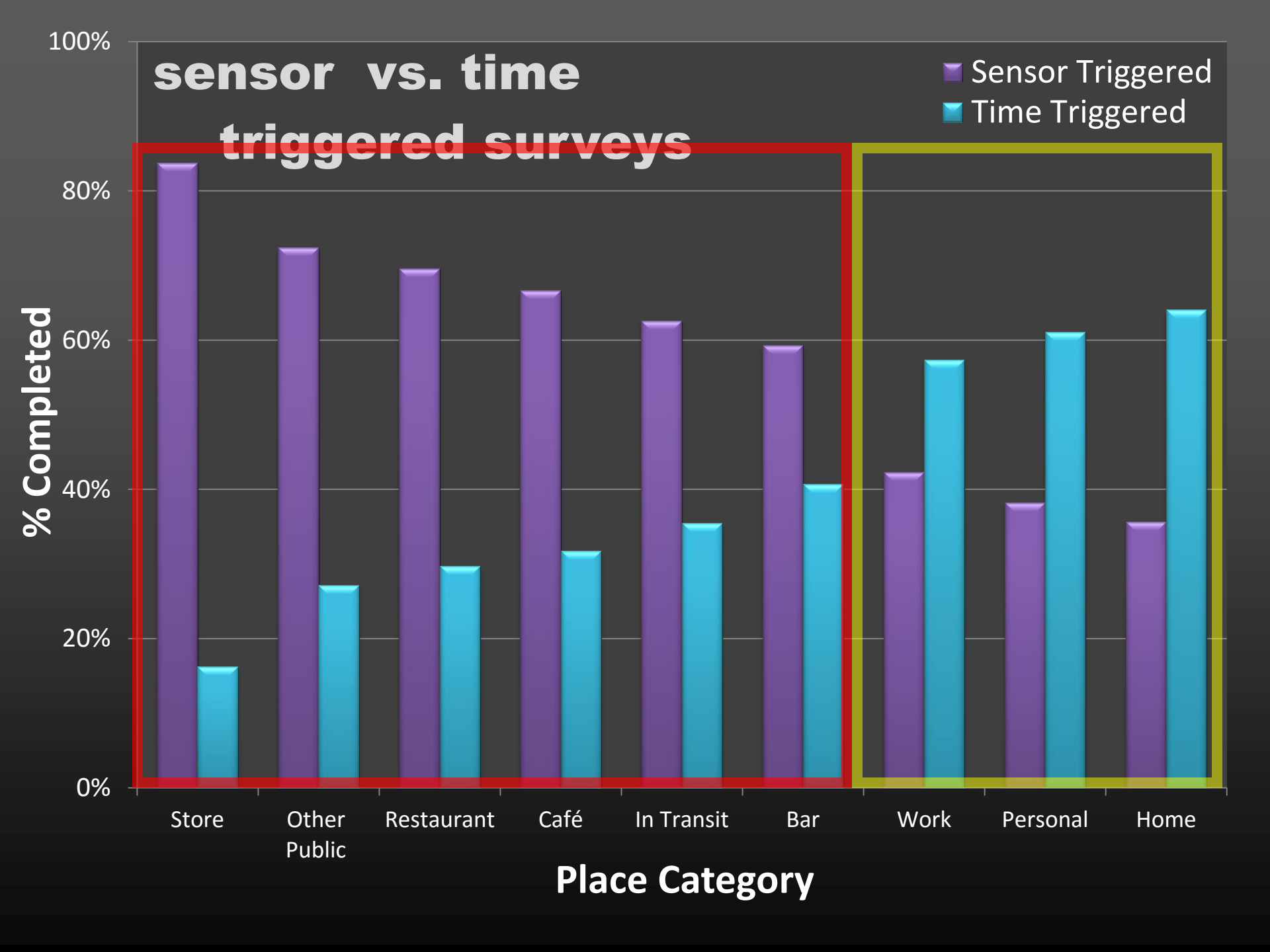
…by and large I go to places I've been to before and I already like.

-Participant #1

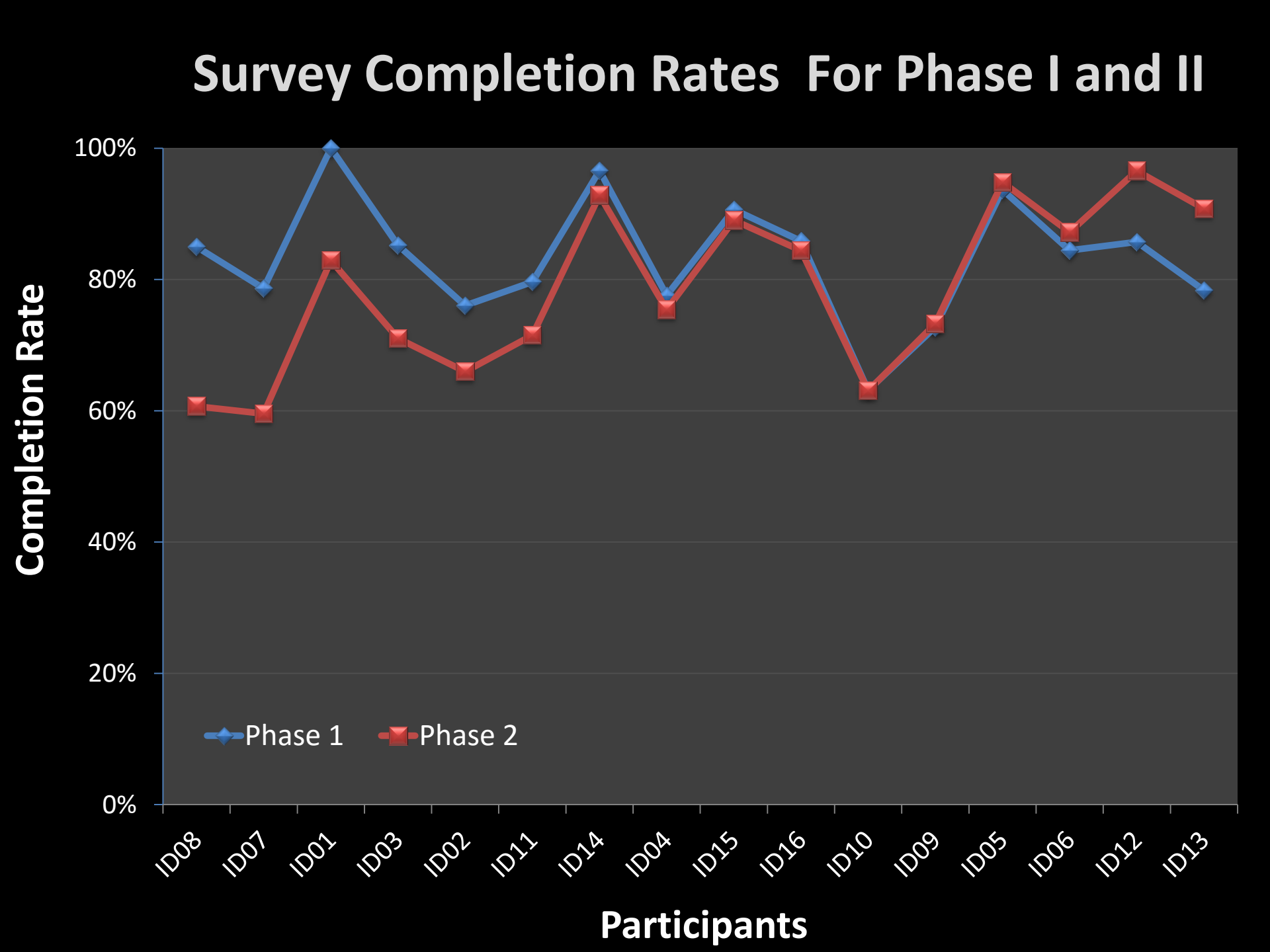
Most of them get pretty good ratings, 4's or 5's because I scrupulously avoid places that I've known I don't like and I always go back to the ones I do like.

- Participant #12

Distribution of ratings indicates that people tend to go where they like


sensor vs. time triggered surveys
Sensor Triggered
Time Triggered
100%
80%
60%
40%
20%
0%
% Completed
Store
Other Public
Restaurant
Café
In Transit
Bar
Work
Personal
Home
Place Category

Survey Completion Rates For Phase I and II



ongoing

ubifit

Sunny Consolvo[1,2], Jon Froehlich[1,2], Beverly Harrison[2], Pedja Klasnja[1,2], Anthony LaMarca, [2] James Landay[1,2], Louis Legrand[2], Ryan Libby[1,2], David McDonald[1], Ian Smith[2], Tammy Toscos[2]

dub design: use: build:

[1] university of washington

[2] Intel Research, Seattle

personal device as data collector

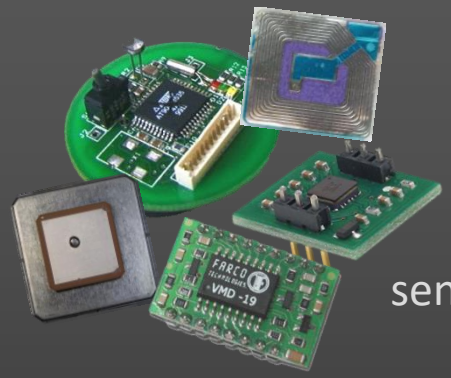
- MyExperience updated to run unobtrusively on user's own phone
  - Limit interruption
  - Idle sensing
  - Obey phone profile

mobile sensing platform

MSP Features

- Built on iMote2
- Linux OS
- 32MB RAM
- 2 GB Flash Storage
- Zigbee and Bluetooth
- 12-16 hours battery life

wearable msp



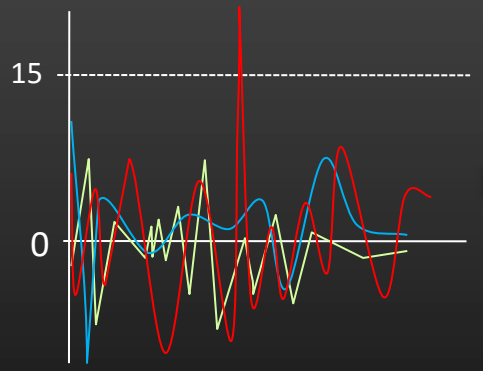
sensors

10 Built-in Sensors

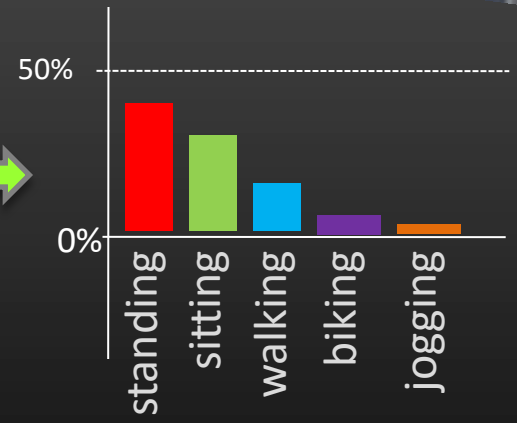
- 3D Accelerometer
- 2D Compass
- Barometer
- Humidity
- Visible light
- Infrared light
- Temperature
- UART, GPIO breakouts for additional sensors

msp & myexperience

Bluetooth



raw data



inference data

from raw sensor readings to meaningful activities

mean, median, range, etc.

measure of confidence for particular activities

collect raw sensor readings → **calculate features** → **produce margins**

generated at varying rates

generated at ¼ second intervals

smooth margins into meaningful activities

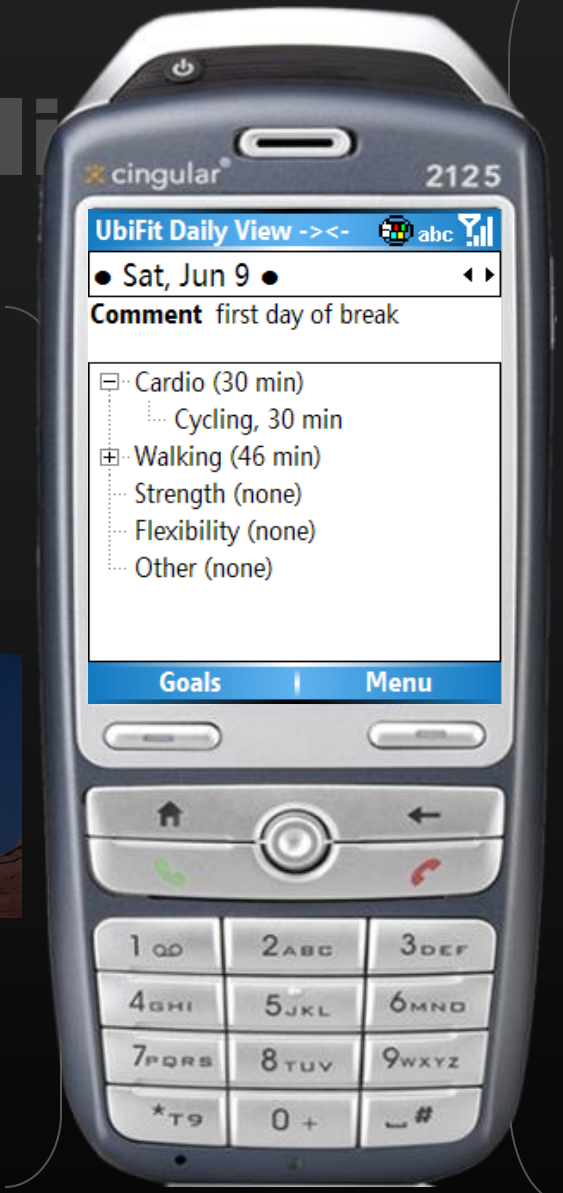
send margins to phone via bluetooth

this can be complicated, as you seldom do the activity 100% of the time you think you are doing it

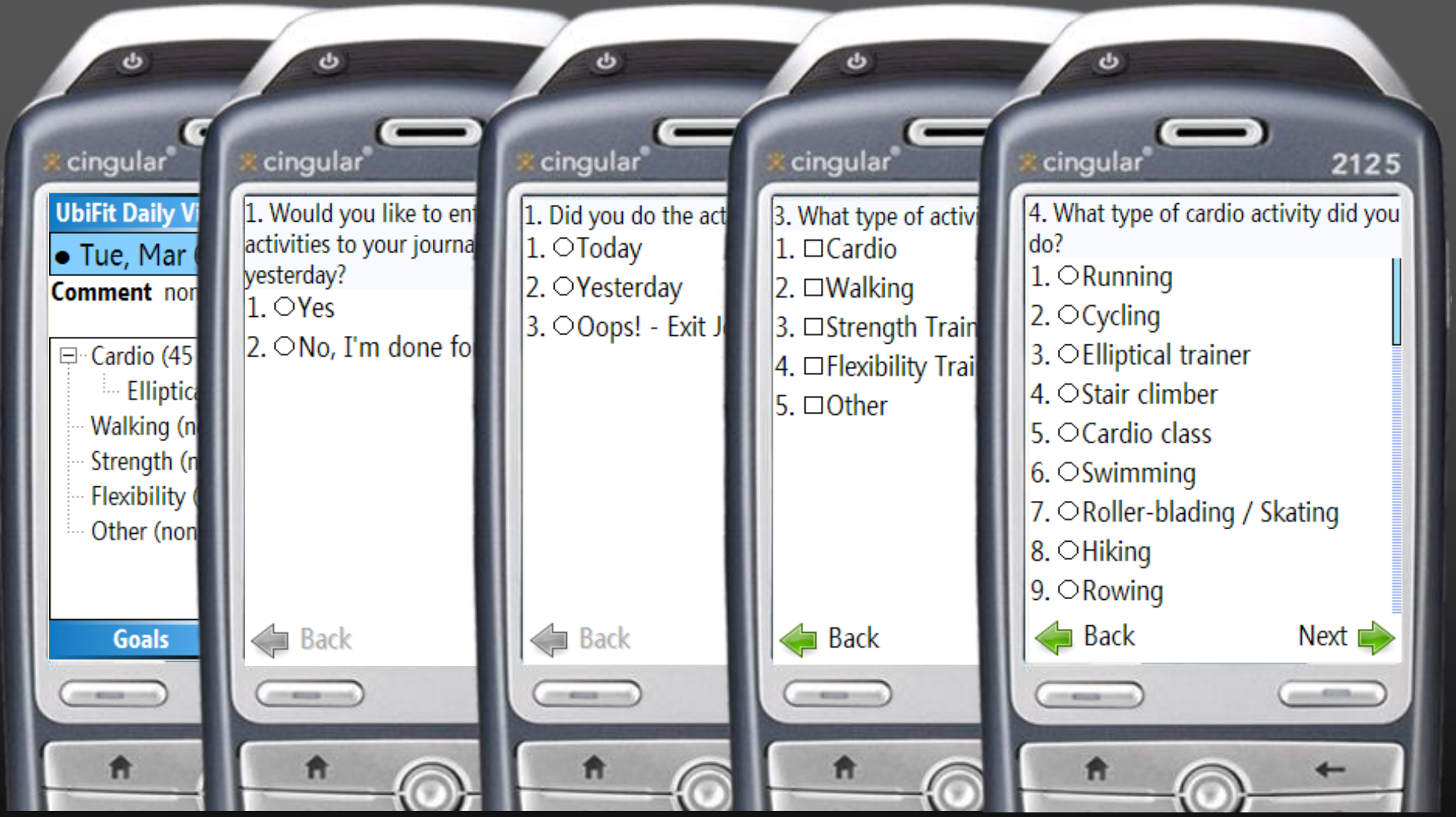
smoothing is defined by the application, not the msp

activity journali

any physical activity including those not inferred by the fitness device

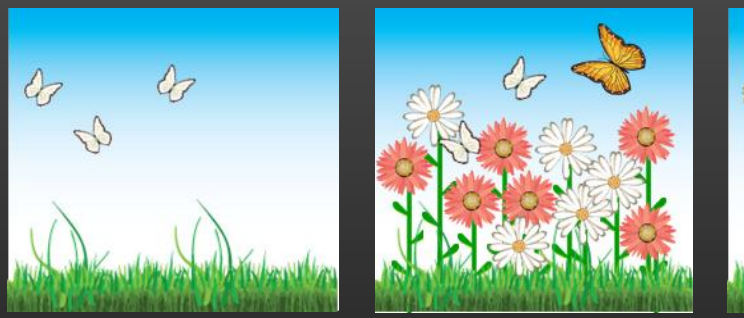


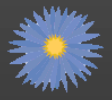
manual journaling



glanceable display

runs on the background screen of mobile phones, so it's frequently seen by the individual





strength

cardio

flexibility

walk

this week's goal met

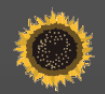
recent goal met

at-a-glance determination of:

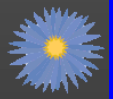
- active or inactive week,
- variety in routine,
- this week's goal met
- recent goal met

day view

the glanceable display in action


walk

cardio


strength

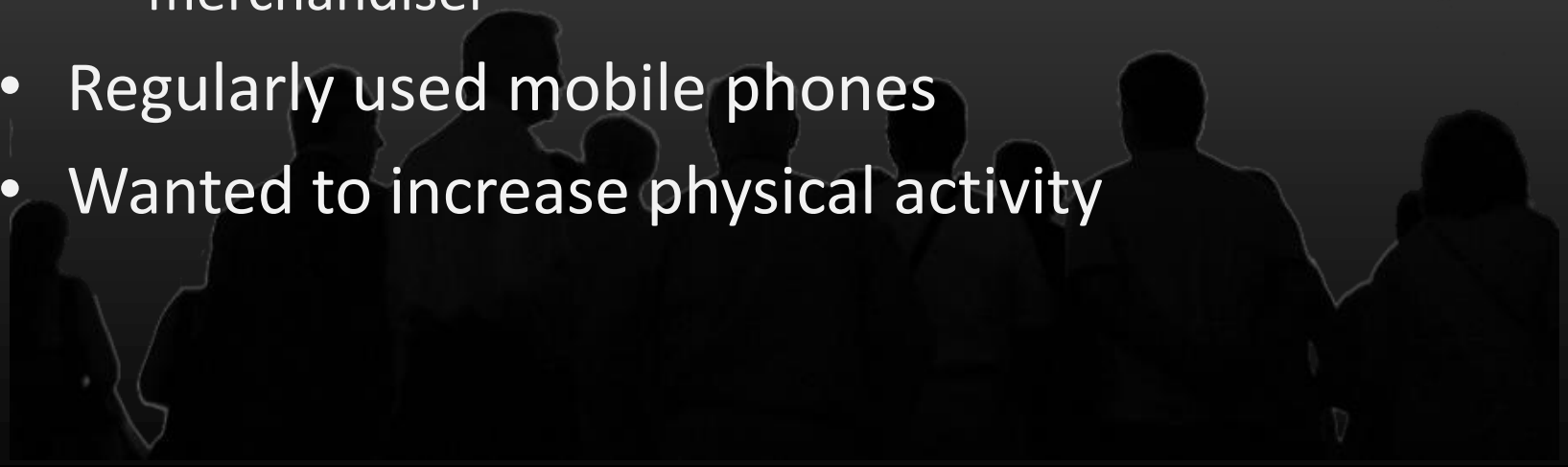
flexibility

this week’s goal met

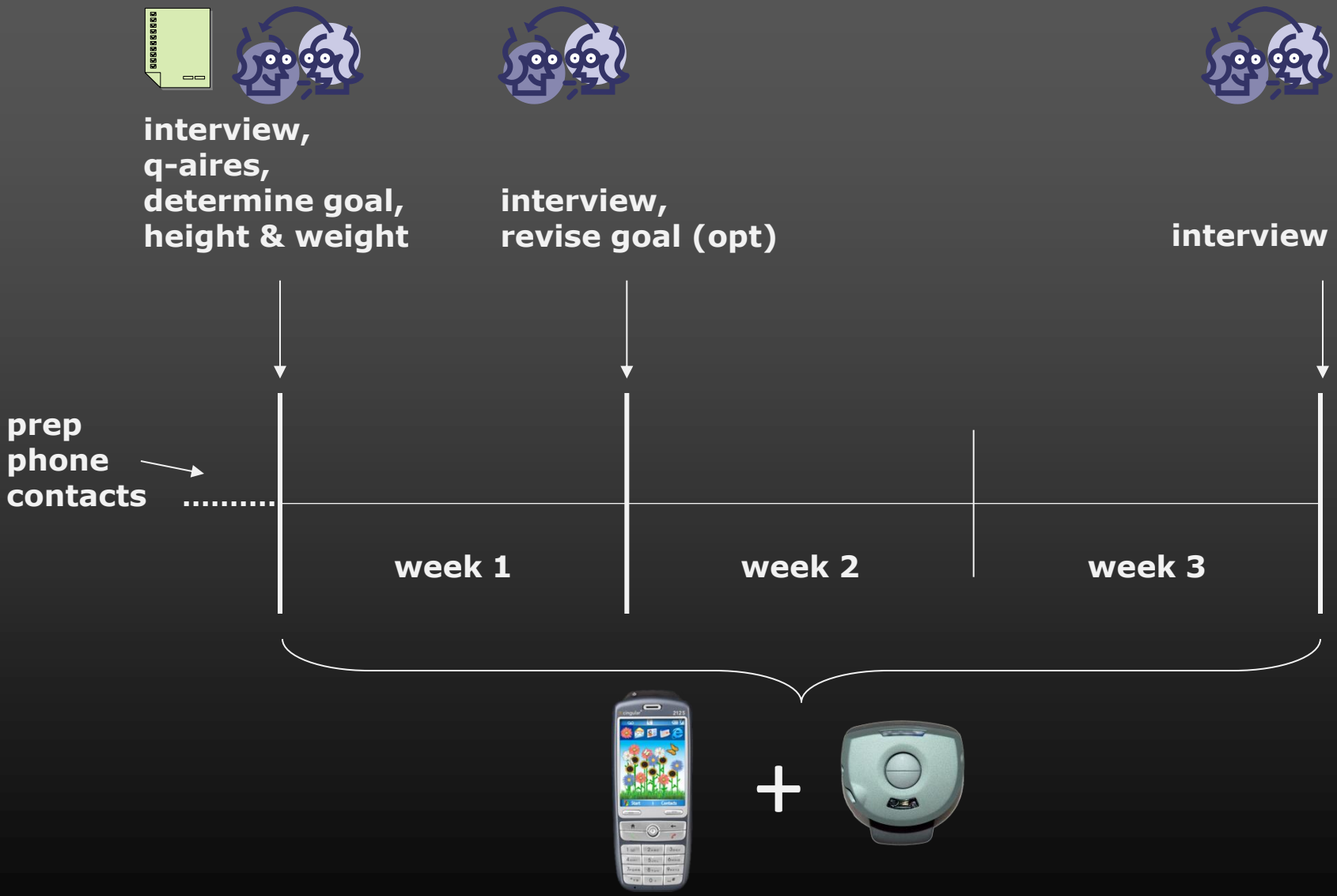
recent goal met

3-week field trial (n=12)

- Summer 2007
- 6 male, 6 female
  - wore fitness device & used UbiFit Garden everyday
- Aged 25-35, lived in seattle area
- Occupations (mostly non-technical)
  - receptionist, marketing, actor, copywriter, musician, merchandiser
- Regularly used mobile phones
- Wanted to increase physical activity



3-week field trial: overview of method



general reactions

- very positive, particularly r.e. the glanceable display (surprisingly so):

> “The silly flowers work, you know?…It’s right there on your wallpaper so every time you pick up your phone you are seeing it and you’re like, ‘Oh, look at this. I have all those flowers. I want more flowers.’ It’s remarkable, for me it was remarkably like, ‘Oh well, if I walk there it’s just 10 minutes. I might get another flower.’ So, sure, I’ll just walk.” {p5}

general reactions (contd)

"I think it's a great idea really…so you can physically see how much you're really moving around with the flowers growing and everything. …You kind of want to see more flowers grow or whatever opposed to working out or walking around and not seeing any results if you walk daily. I mean, it's going to take several months, but I feel like on the phone you can actually see that you're achieving something…"

{p12}

breakdown of activities performed

	inferred	journaled
cardio	34 (45%)	42 (55%)
walking	143 (55%)	115 (45%)
strength	0	79 (100%)
flexibility	1 (1%)	75 (99%)
other	29 (67%)	14 (33%)
totals	**207 (39%)**	**325 (61%)**

17 - 84 activities per participant (mean: 44, median: 39)

perceived errors of activity inference

- our analysis revealed 7 types of perceived errors with activity inference
- the fitness device could:
  1. make an error in start time
  2. make an error in duration
  3. confuse an activity it ***was*** trained to infer with another it ***was*** trained to infer
  4. confuse an activity it ***was not*** trained to infer with one it ***was*** trained to infer
  5. fail to detect an activity it ***was not*** trained to infer
  6. fail to detect an activity it ***was*** trained to infer
  7. detect an activity when none occurred

subset of ubifit triggers

Journal reminder

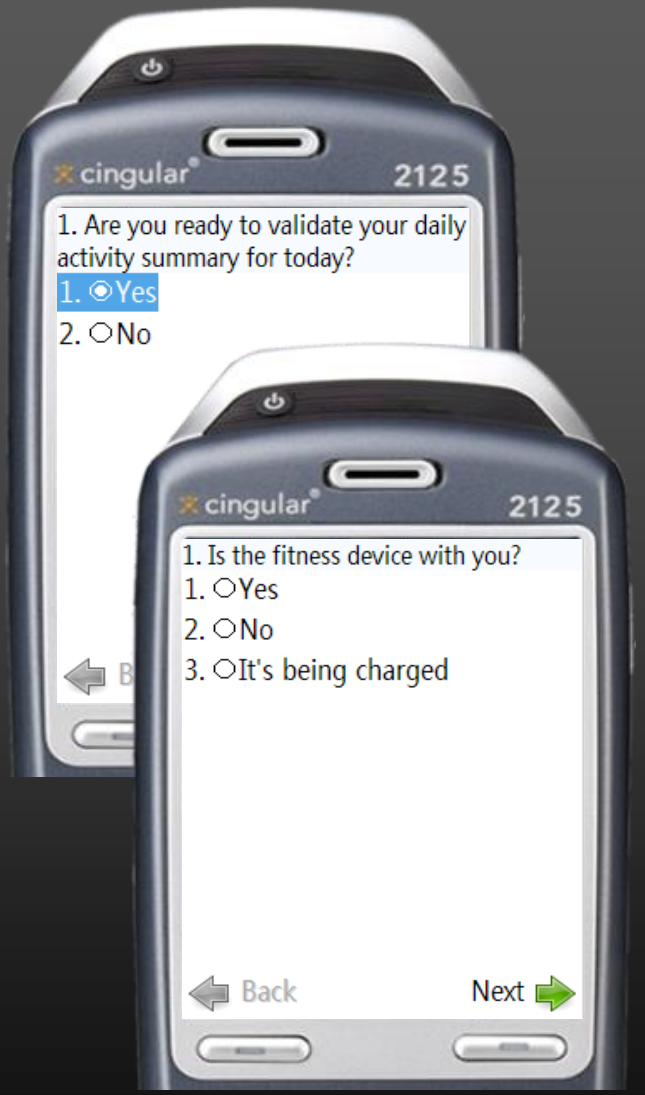
- If journal has not been used in ~2 days and it's past 8PM, launch journal reminder

Uncertain activity occurred

- If the system *knows* an activity occurred but couldn't determine the exact activity, a survey is launched

MSP troubleshooter

- If the MSP hasn't been seen in ~2 hrs and it's after 10AM, launch a troubleshooter



MobiSys2007, ongoing

mobilephoneusage

Jon Froehlich[1,2], Mike Y. Chen[2], Sunny Consolvo[1,2], James Landay[1,2]

dub design: use: build:

[1] university of washington

[2] Intel Research, Seattle

ZURIC
copyright 2000 philg@m

our goal

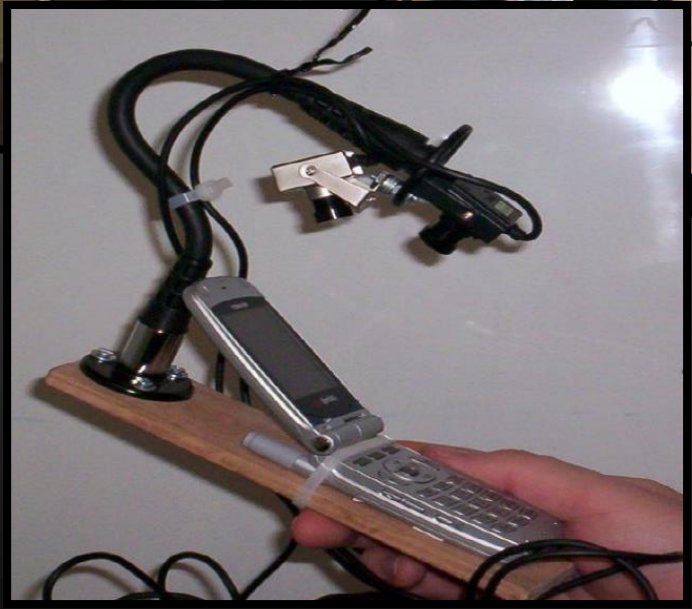
Collect data about *real* device usage & context *in the field* in a scalable fashion

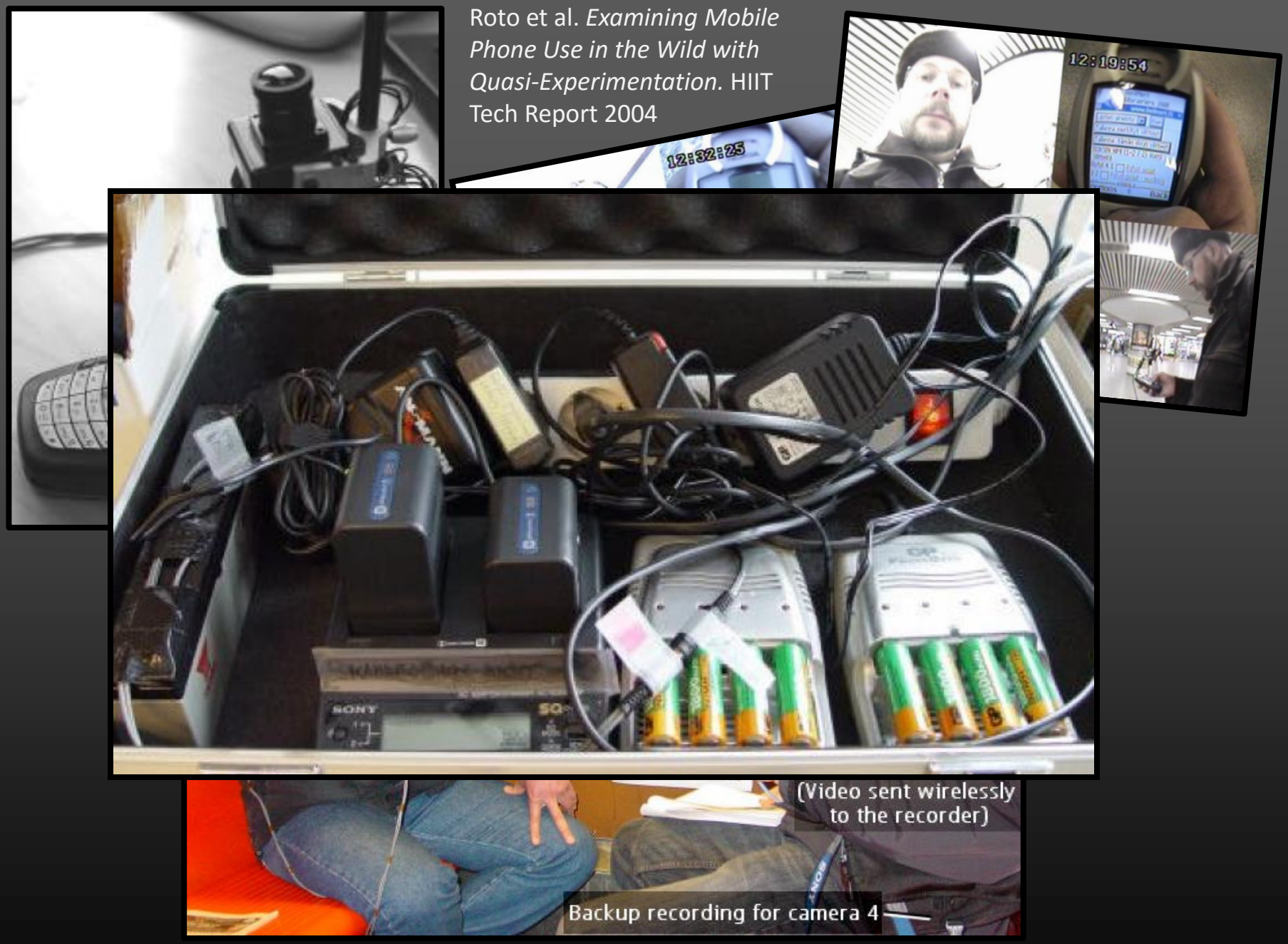
Data can be used to

- Better understand actual device/system usage
  - E.g., how mobility patterns affect access to WiFi
- Inform the design of future systems
  - E.g., optimize battery utilization algorithms based on learned charging behaviors

Kjeldskov, J. and Stage, J. *New Techniques for Usability Evaluation of Mobile Systems*. IJHCS2003

Schusteritsch *et al*. *Towards the Perfect Infrastructure for Usability Testing on Mobile Devices*. CHI2007.





Roto et al. *Examining Mobile Phone Use in the Wild with Quasi-Experimentation.* HIIT Tech Report 2004

Reichl et al. *The LiLiPUT Prototype: A Wearable Lab Environment for User Tests of Mobile Telecommunication Applications*. CHI2007

research challenges

1. Coverage: collect rich information about features of interest
2. Scale: collect large amounts of data over long periods of time
3. Extensible: easily add new data collecting capabilities
4. Situated: collect *real* usage data in its natural setting
5. Robustness: protect or backup data collected in the field

case study 1: charging behavior

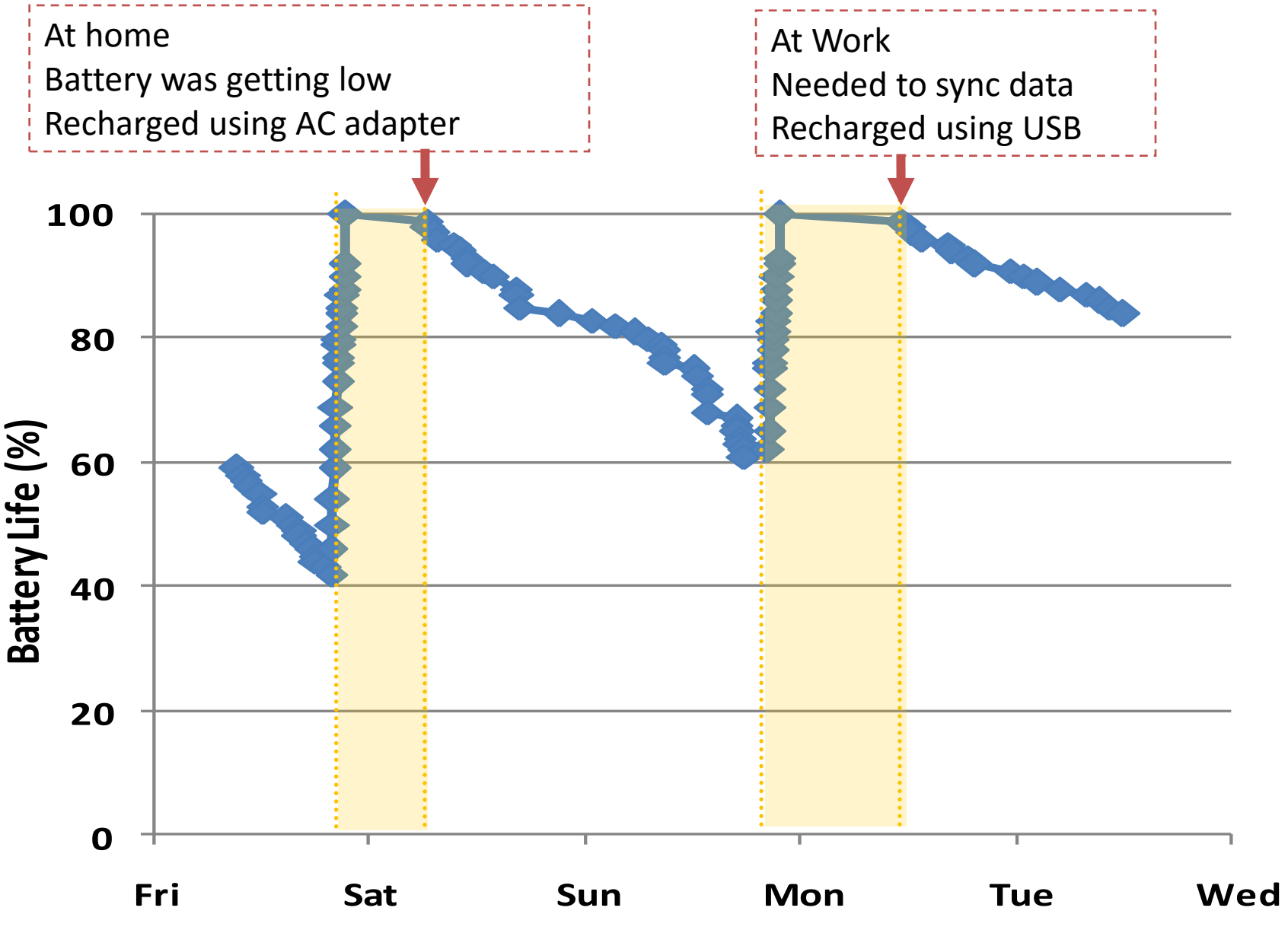
Motivation

- Battery life has long been a challenge in mobile computing
- Dependent on usage:
  - WiFi, video, length of calls

Study

- 2 week pilot study with 4 people
- Log device usage (e.g., phone calls, WiFi, active applications)
- Actively track battery life
- Survey at moments of charging

battery life & user response



case study 2: sms usage

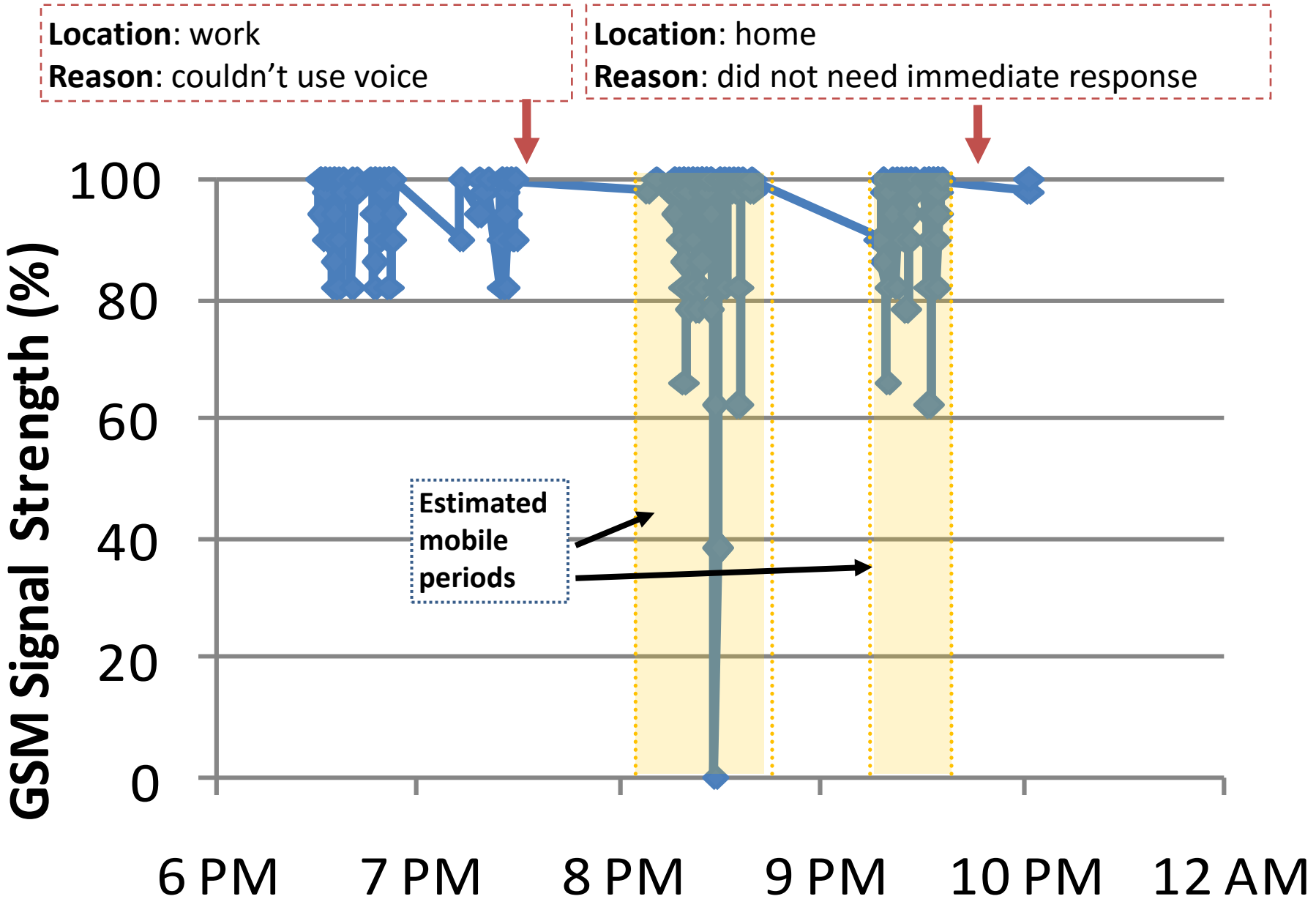
Motivation

- 1 trillion SMS messages sent worldwide in 2005
- Explosive growth begs research questions:
  - Why SMS vs. voice?
  - Where do people use SMS?

Study

- Similar setup as before
- Asked questions after SMS sent
  - User's location
  - Reason for using SMS

sms usage, mobility & self-report



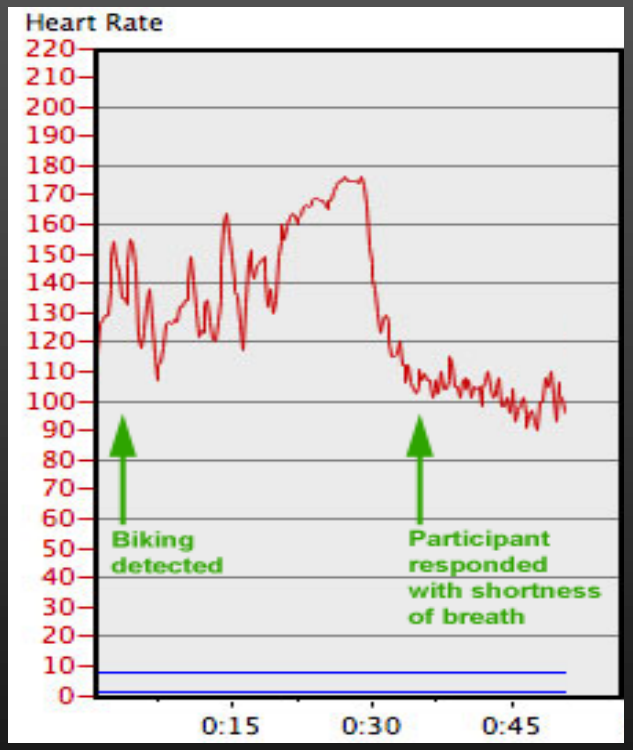
3-month Study Combining Activity Sensing with Device Usage Logging

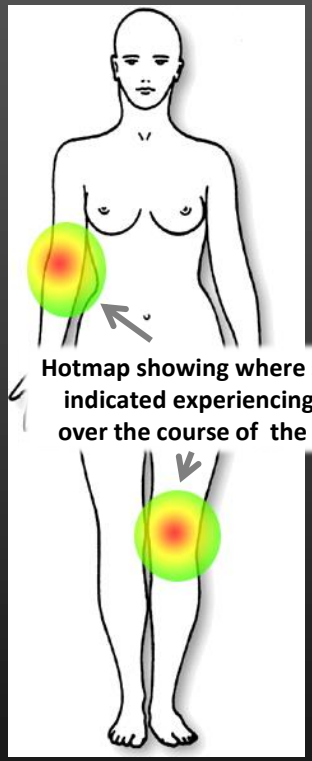
longitudinal study

- The data collected should allow us to explore:
  - The link between mobility patterns and application usage
    - Do people SMS more when stationary than moving?
    - Driving while texting
  - How often users suffer from low cell signal strength and how this affects voice vs. sms
  - Is application usage correlated to location, time, etc.
    - Could we start to prefetch data based on context?

analysis tools

How can we analyze gigabytes of sensor data per participant plus contextualized self-report data?





Hotmap showing where subject indicated experiencing pain over the course of the study

myexperience studies

Studying
Activity Recall

Stress, Behavior
Change and
Heart
Monitoring

UbiFit

Activities
and Mobile
Phone Usage

Mobile Heart
Health Study

Vote with
Your Feet

UbiGreen

Pain/Fatigue
Management
in Multiple
Sclerosis
Patients

Obesity, Activities
and Geography

http://myexperience.sourceforge.net/
myexperience
thankyou!
home
download
docs
source
contact
Quick Download
Release 2007.10.11
Recent SVN Commit Activity
Revision 130
Revision 129
Date: November 8, 11:07 am
Recency: 6.0 days ago
Author: jonfroehlich
Files: 0 add | 0 del | 2 mod
Message: updated the MyExperience Logger cab installer so that a shortcut is copied into WindowsStartUp
Please select your current location on this map.
the My Experience tool
Please rate the quality of that last last phone call.
Please take a picture of your current meal.
Next
download myexperience @
http://myexperience.sourceforge.net
MyExperience is a context-aware data collection platform for capturing objective and subjective data as it's experienced.
email: jonfroehlich@gmail.com
SOURCEFORGE.NET
My Experience is Open Source Software under the BSD license.
Webpage designed and developed by Jon Froehlich. Copyright (c) 2007

backup slides

ongoing

ubigreen

Jon Froehlich[1,2], Beverly Harrison[2], Pedja Klasnja[1,2], James Landay[1,2], Jen Mankoff[3]

dub design: use: build:

[1] university of washington

[2] Intel Research, Seattle

ubigreen

cingular
SurveyPanel: 1
It looks like you have r
changed your location.
correct?
1. Yes, take the trav
2. No, I didn't mov
BACK
cingular
SurveyPanel: 2
Where did you ju
1. home
2. work
3. grocery stor
4. shopping ce
5. restaurant
6. out of town
7. other
BACK
cingular
SurveyPan
How did yo
1. car (dr
2. carpoo
someo
3. bus
4. walk
5. bike
6. ferry
7. train
8. other
BACK
cingular
SurveyPanel: 4
How did you spe
while you were
1. read
2. worked
3. slept
4. listened t
5. played g
6. talked to
7. other
BACK
cingular
SurveyPanel: 2
Please take a picture of
something that illustrates h
you are experiencing your t
today.
Press any number key
to launch the
camera
BACK
NEX
cingular
2125
SurveyPanel: 2
Please take a picture of
something that illustrates how
you are experiencing your travel
today.
Your captured image
BACK
NEXT