

Mobile sensing systems: from ecosystems to human systems

Professor Deborah Estrin
NSF Science and Technology Center for Embedded Networked Sensing (CENS)
UCLA Computer Science Department
destrin@cens.ucla.edu

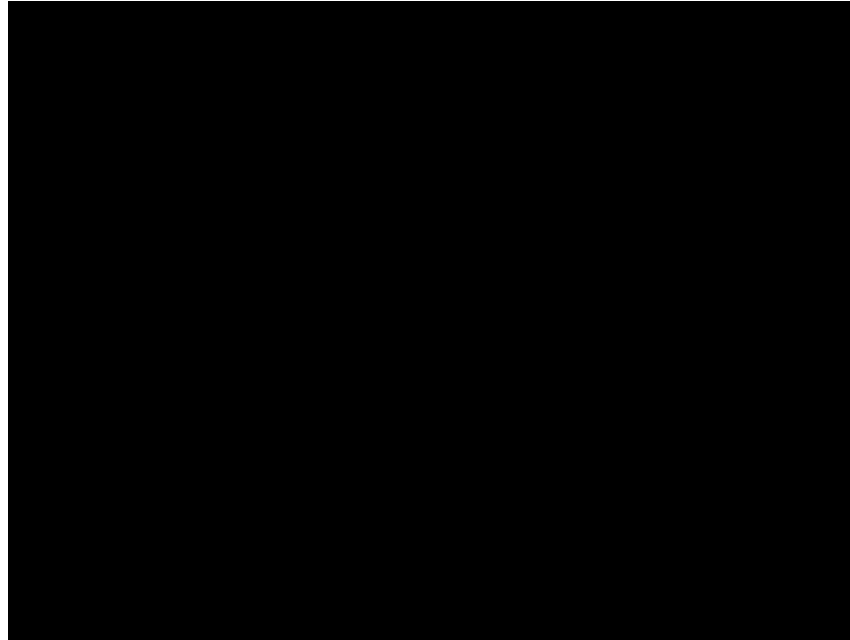
Co-PIs: Jeff Burke (CENS/REMAP), Jeff Goldman (CENS), Ramesh Govindan (CENS/USC), Eric Graham (CENS), Mark Hansen (CENS/Statistics), Mary Jane Rotheram-Borrus (Psychiatry/Semel), Mani Srivastava (CENS/EE), Ruth West (CENS)

Students and Staff: Peter Capone-Newton, Betta Dawson, Hossein Falaki, Taimur Hassan, Donnie Kim, Martin Lukac, Olmo Maldonado, Min Mun, Nicolai Petersen, Nithya Ramanathan, Sasank Reddy, Jason Ryder, Vids Samanta, Katie Shilton, Nathan Yau, Eric Yuen

Work summarized here is that of students, staff, and faculty at CENS. We gratefully acknowledge the support of our sponsors, including the National Science Foundation, Nokia, Intel Corporation, Cisco Systems Inc., Sun Inc., Google, Microsoft Research, UC Micro, Crossbow Inc., Agilent, Conservation International, and the participating campuses.



[YouTube Video on PEIR](#)



Many critical issues facing science, government, and the public call for high fidelity and real time observations of the physical world

Embedded sensing systems:

reveal the previously unobservable

help us understand and manage interactions with physical world, scarce resources, and one another



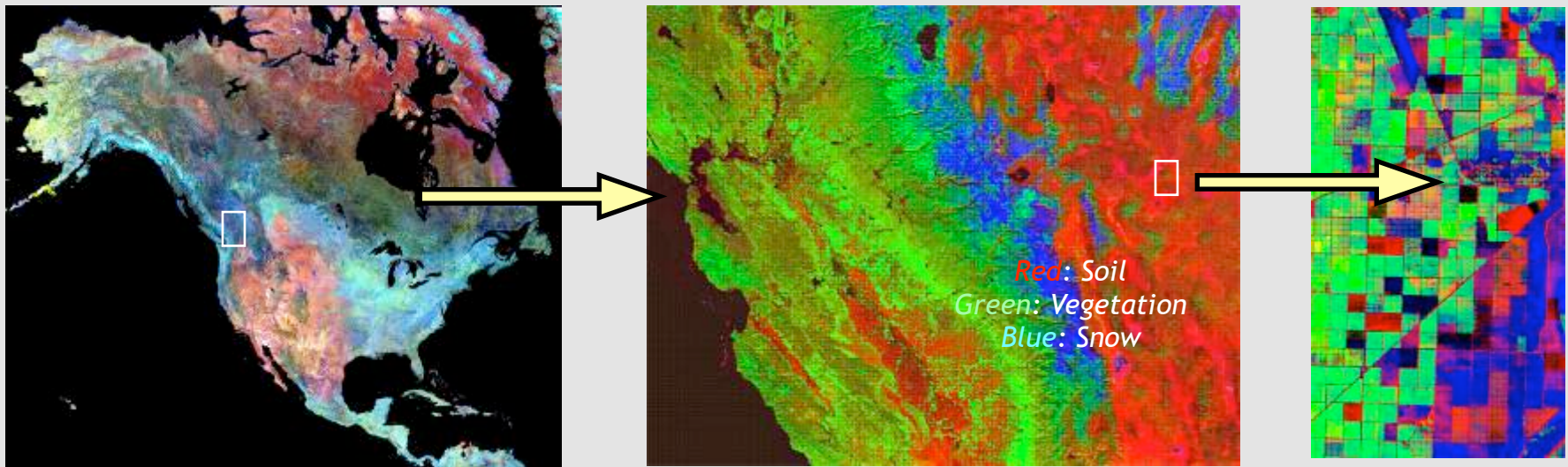
Why *embedded* sensing?

- Remote sensing transformed observations of large scale phenomena
- Embedded (in situ) sensing transforms observations of spatially rich processes

Embedded in the physical environment

Networked to share information/adapt function

Sensing physical world phenomena



San Joaquin River Basin

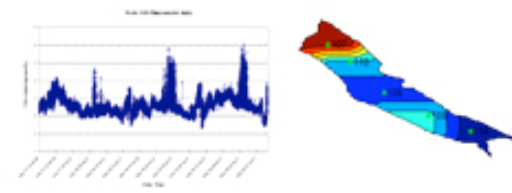
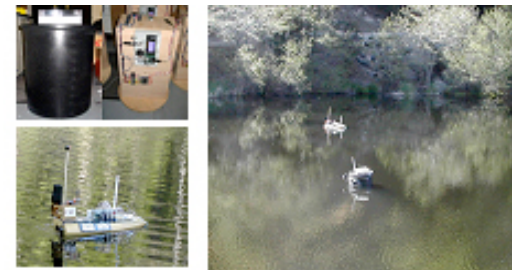
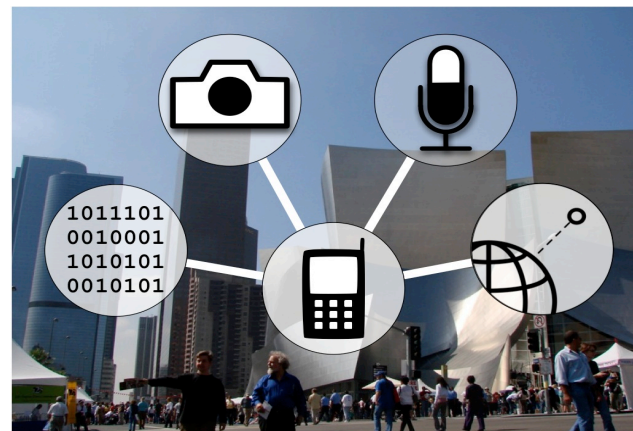
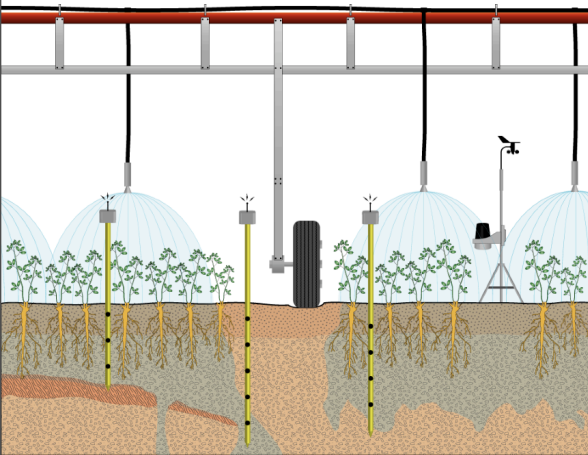
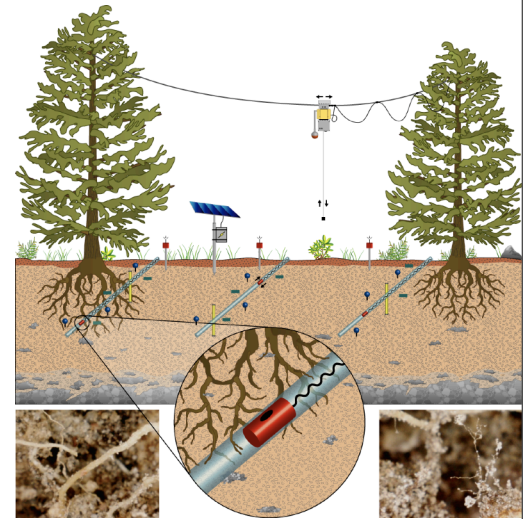
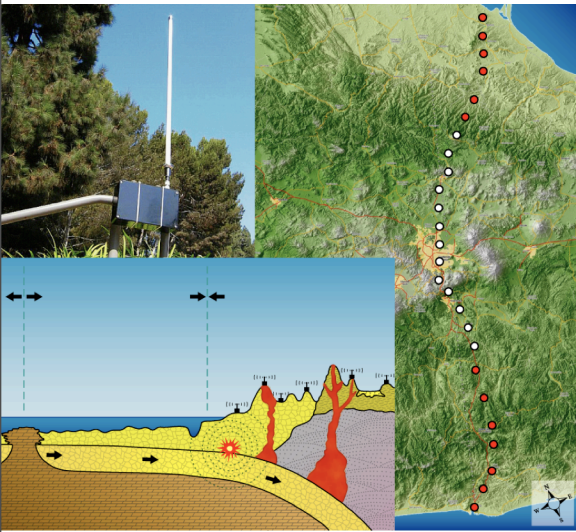
Susan Ustin-Center for Spatial Technologies and Remote Sensing

Center-wide focus: embedded networked sensing *to reveal the previously unobservable*

create programmable,
distributed, multi-modal, multi-
scale, multi-use observatories
to address compelling science
and engineering issues
...and reveal the previously
unobservable.

From the natural to the built
environment...

From ecosystems to human systems...



Lessons from the field lead to new approaches

Early Themes

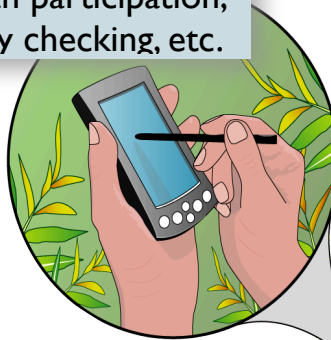
- Thousands of small devices
- Fully autonomous systems

Current Themes

- Mobility
- Server-side models, data, analysis
- Coupled human-observation systems

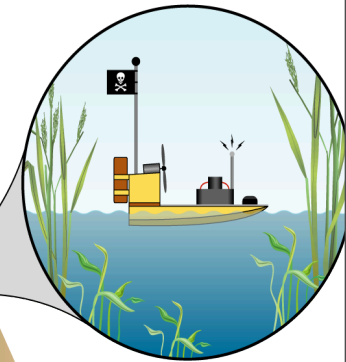
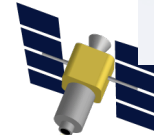
Handheld Sensing

human participation,
reality checking, etc.

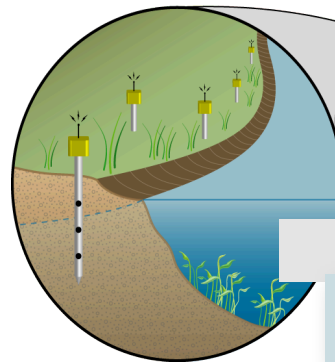


Remote Sensing

Overlaying the “big picture”
on local events

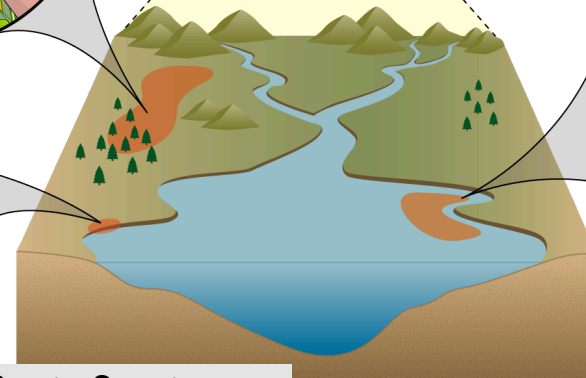


Robotic Mobility

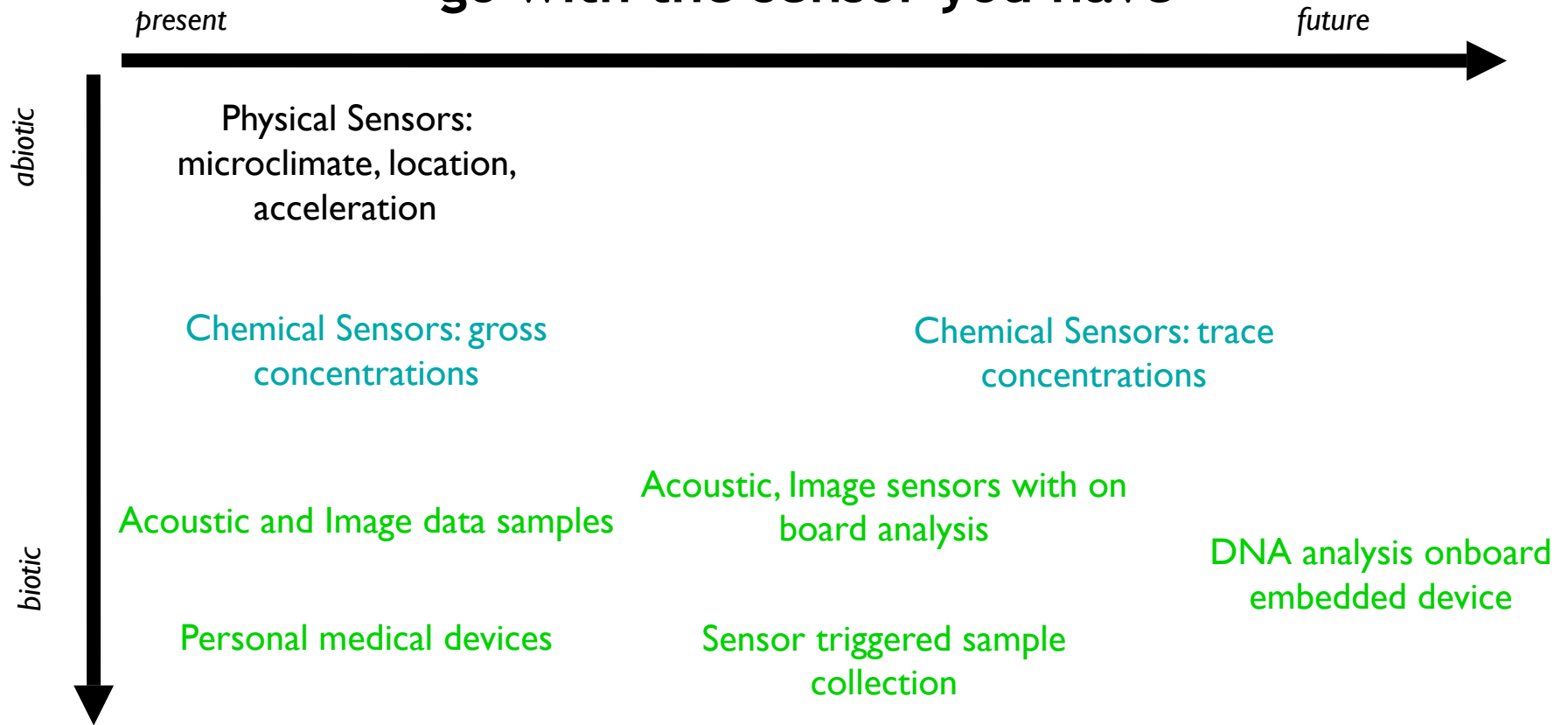


Static Sensing

Stationary sentinels,
continuous in time



If you can't go to the field with the sensor you want, go with the sensor you have



- System designs need to compensate for lack of sensor specificity, sensitivity, availability...particularly wrt biological response variables
- Leverage proxy sensors and model based interpretation/inference

Mobile Personal Sensing

<http://urban.cens.ucla.edu>

Enabled by $>3 \times 10^9$ mobile phone users, increasingly with...



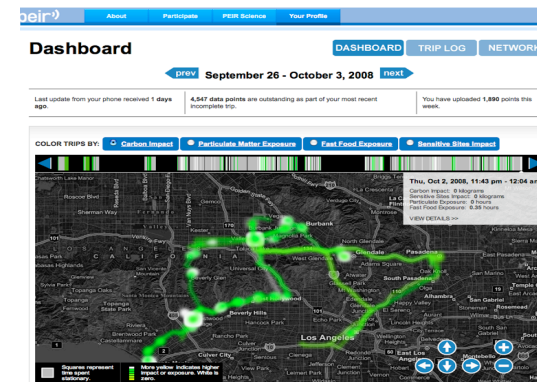
- Digital imagers, location, bluetooth-connected sensors
- Automatic-geocoding of data
- Programmed, user, and server-initiated capture
- Server-side/cloud processing/archiving/ presentation of personal



Motivated by 6×10^9 people on planet earth and their concerns...



- Individual health and wellness
- Public health, urban planning, epidemiology
- Civic concerns (transportation, safety, culture...)
- Resource management



Participatory Sensing: Campaign Model *leveraging real-time, geo-coded, images*

Distributed data gathering challenges as “Campaigns” -

Spatially and temporally constrained systematic data collection operations.

Exploring a single hypothesis, phenomena or theme.

Using human-in-the loop sensing to gather data.

With automatic and manual classification, auditing, and analysis.



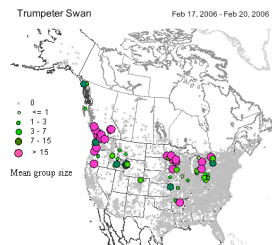
Precedent - Community-Based Participatory Research



PhotoVoice
Caroline Wang, 1996



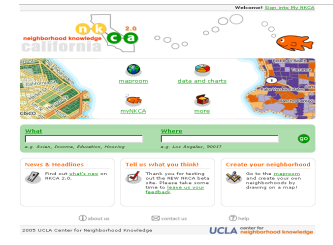
Citizen Science
World Water Quality Day



Citizen Science
Cornell e-Bird



Civic Participation
Video the Vote



Participatory GIS
Ctr for Neighborhood Knowledge

Burke, et al

Campus Sustainability Initiatives

GarbageWatch




<http://garbagewatch.com>



foodyouwaste

Lets take a bite out of food waste! Monitor your waste so you can do something about it!

type in your email address here [find](#)



Its easy to contribute! Just send an email or mms of the food that is left over that you are not planning to eat to the following address:

[mobile\[at\]foodyouwaste.com](mailto:mobile[at]foodyouwaste.com)

Then when you want to check out your food waste footprint, just fill out the form above with the email address you used to send the images!

Sponsored by: [CENS](#) | [IOE](#) | [USCLA](#) | [TakeFrost Labs](#)

[Home](#) | [About Us](#) | [Food Waste Links](#) | [Contact Us](#)

<http://foodyouwaste.com>

thecleanmachine@gmail.com

High Count: 7



Medium Count: 7



Low Count: 22



[challenge: automated campaign management, data integrity feedback]

CycleSense

Help bikers plan safe routes & collect data to improve routes.

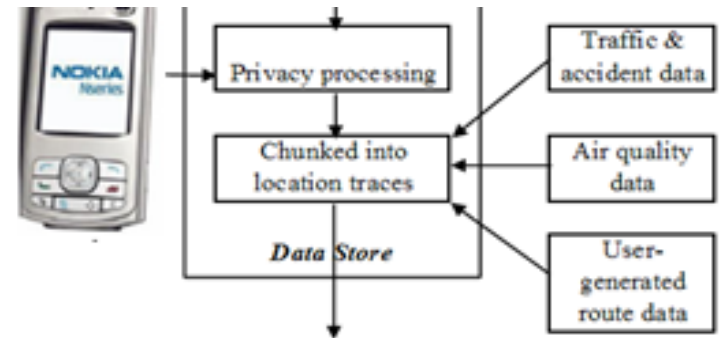
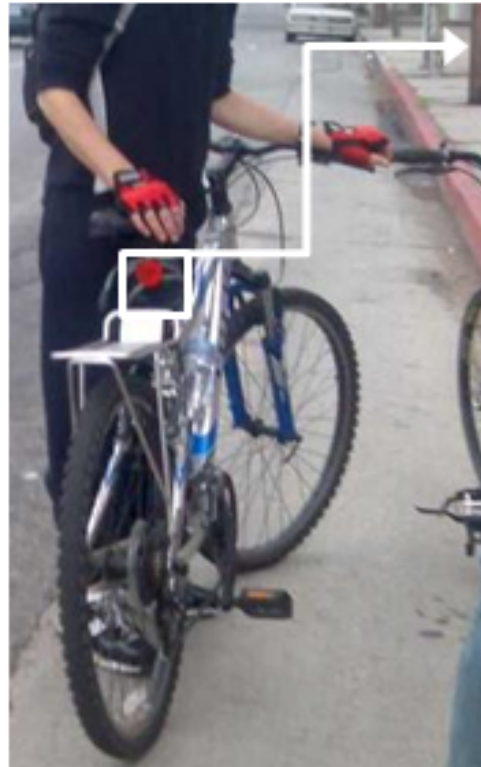
Capture/process/ share route data and features

location, duration,
stops/starts,
roughness, sound
samples, images of
route features

Web interface compares
routes' qualities

Mash up routes with air
quality, traffic
conditions & accidents

Community site for
sharing information
with LA bike community



Chinatown/Lincoln Heights/Cypress Park Bikeabout



[challenge: activity classification, usability]

Shilton, Reddy, Samanta, Ramanathan, et al

“...the total costs associated with HABs over past decades have been conservatively estimated at over \$1 billion.” - NOAA¹

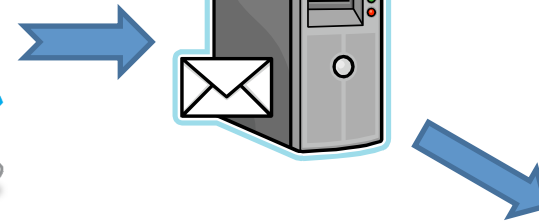
Mobile MERHAB



SMS / MMS



High spatial/temporal density data capture



SmartPhones



Networked Naturalist

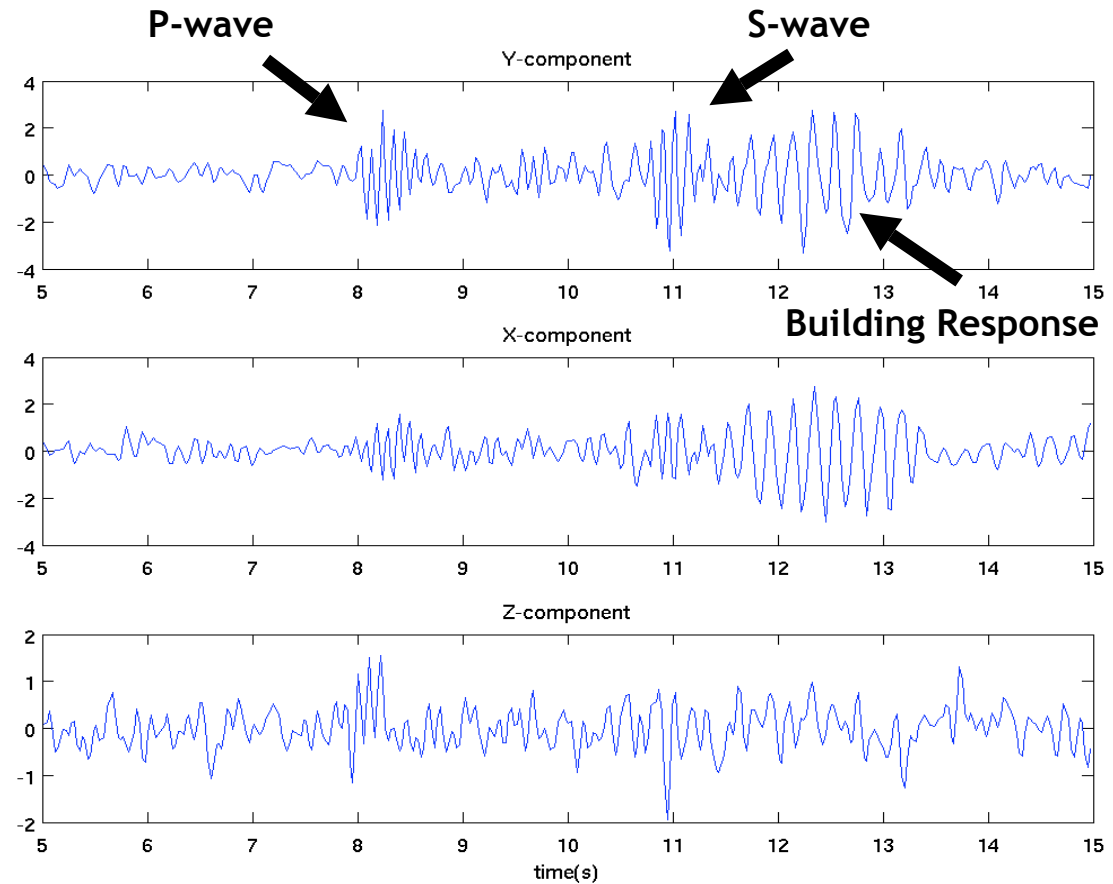


“... by monitoring changes in phenological events such as first bud, budburst, and flowering, scientists can detect climate change.”- Project BudBurst

[challenge: automated data integrity checks, usability]

Geocoded person-scale samples (accelerometer) for seismic

- First (we think) capture of magnitude 3.4 earthquake on Nokia n95 cell phone (8-bit ~ 40Hz 2G accelerometer) 5 miles away from epicenter (*has been done on laptops--UCR*)
- 1/23/09: Santa Monica, California. Magnitude 3.4. (ID ci10373093)
- P and S waves observed at ~ 8 and ~11 seconds, respectively.
- Ringing after 11 seconds typical of response of buildings to seismic waves--not in records from nearby (buried) seismic station.



Detrended and low-pass filtered (10Hz)

Traditional measures: <http://quake.wr.usgs.gov/recenteqs/Quakes/ci10373093.html>

[challenge: calibration, what does it mean?]

DietSense as alternative to self-reporting *leveraging real-time, automated, personal, images*

mobile phones

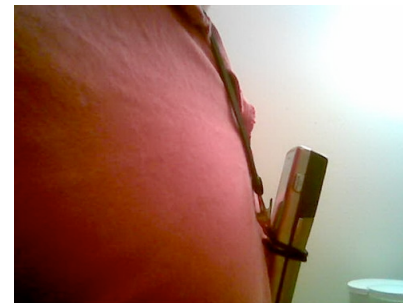
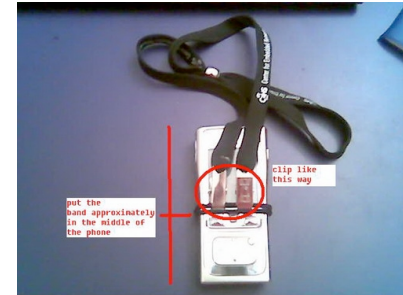
worn on a lanyard around the neck that automatically collect geo-coded, time-stamped images of food choices or purchases.

participant data repository

receives annotated media collected by the devices, allows individuals private access to their own data before they are available to others, and supports filtering and alerting based on upload patterns and basic analysis of received data.

annotation, filtering, and analysis tools

provide efficient mechanisms to navigate, annotate, filter, and analyze collected data, including capability to export reports to common statistical software packages.



[challenge: computer vision category recognition, wearable cameras]

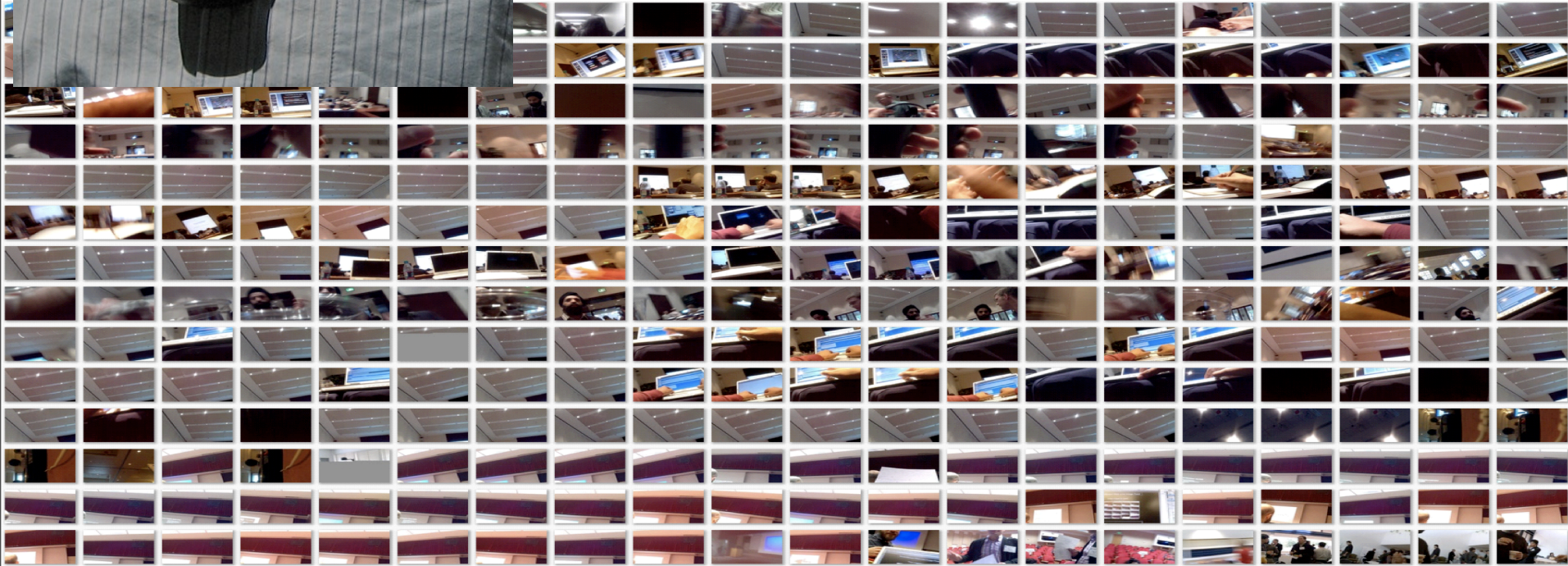
Privacy concerns dominate system design

Automated pre-filtering, clustering, to reduce number of viewed images

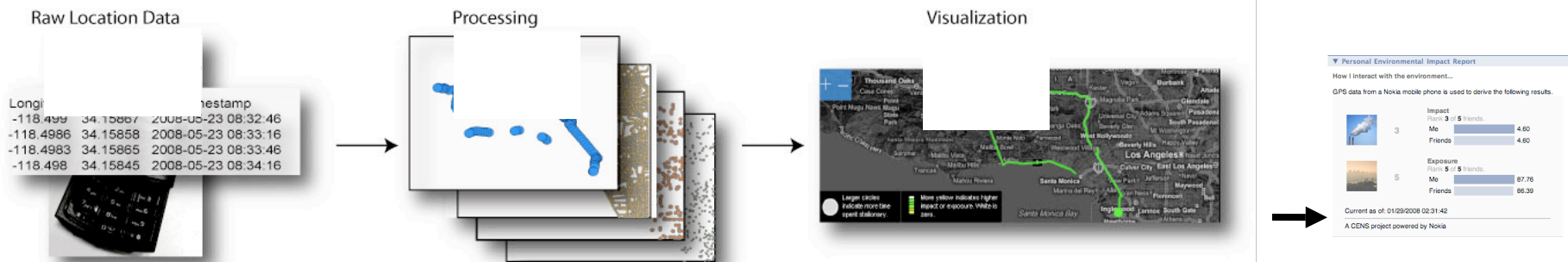
[challenge: computer vision, category recognition]

Privacy concerns dictate image viewing/tagging by individual, not by third party

[challenge: privacy filtering]



Geo-coding as primary (not just meta) data: *leveraging real-time location traces*



Device data capture and interaction:

- software on mobile prompts/captures and uploads
- data types: location image, audio, text tagging, worn sensors
- UI on phone

Processing

- activity classification
- mapping
- integrate with other GIS and realtime data about built and natural environment
- index into models
- privacy relevant filtering

Visualization

- for personal and professional insight
- legible, contextualized
- use/user configurable
- difficult to generalize
- projects need support
- platforms available for development

Sharing/Aggregating

- social networking
- web and device
- participatory
- privacy
- track data access for visibility/transparency

Imagine if....

Our **everyday cell phones** could show us *how we impact the environment, and how it impacts us*, just as they now alert us to traffic jams on the highway.



Carbon Emissions



Particulate Exposure



```

- <sensor type="tag">
  <prompt>Add An Annotation:</prompt>
- <list>
  <item>Indoor</item>
  <item>Idle</item>
  <item>Walking</item>
  <item>Running</item>
  <item>Biking</item>
  <item>Freeway Driving</item>
  <item>Street Driving</item>
  <item>Bus</item>
</list>
</sensor>

```



Exposure Assessment

Personal Environment Impact Report

▼ Personal Environmental Impact Report

How I interact with the environment...

GPS data from a Nokia mobile phone is used to derive the following results.



Current as of: 01/29/2008 02:31:42

A CENS project powered by Nokia

Exposure to Traffic-related Pollutants

Lifelong damage found in 13-year study of 3,600 Southland youngsters living within 500 yards of a highway.
The Los Angeles Times, 1/26/07

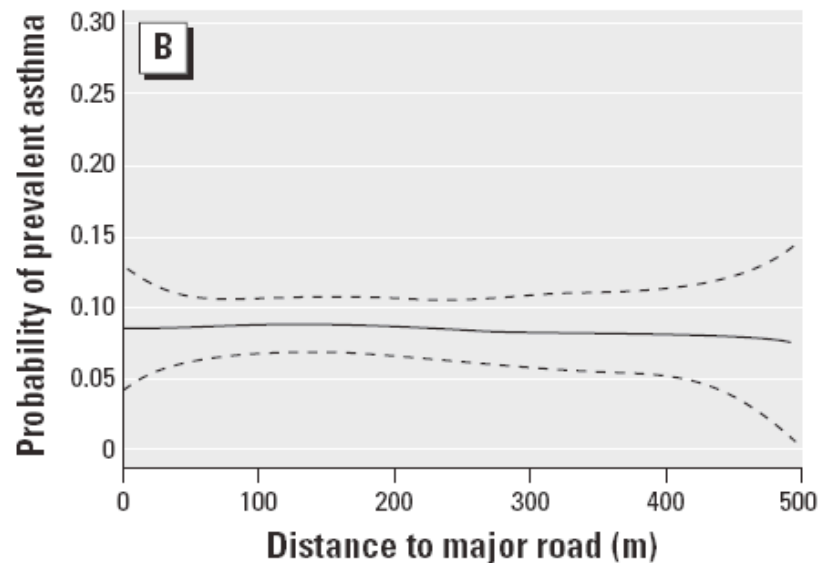
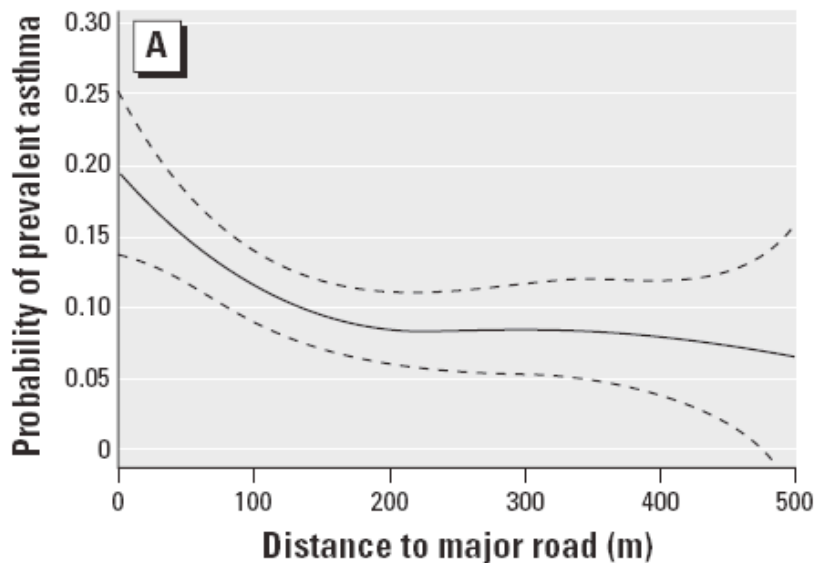


Figure 2. Prevalence of asthma by distance of residence to a major road within 500 m, among long-term (A) and short-term (B) residents with no family history of asthma. Dotted lines indicate 95% confidence interval.

Source: McConnell et al. *Traffic, Susceptibility, and Childhood Asthma. Environ Health Perspect* 114:766–772 (2006)

PEIR: Personal Environmental Impact Report

Personal, real-time, location traces....combined with micro-environmental models ...to provide personal exposure and impact assessment

Invite investigation of individual habits overtime:
...in relationship to others and the environment
...as seen in data and inferred from models.

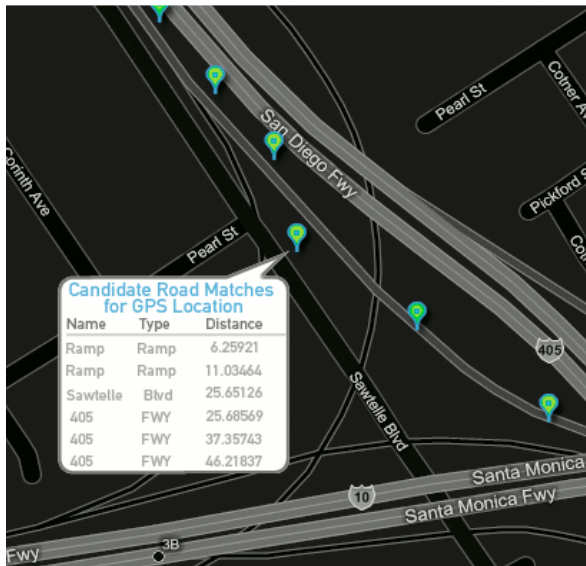
<http://peir.cens.ucla.edu>

The screenshot displays the PEIR web application. The main area is a map of Los Angeles with a green line representing a user's location trace. A tooltip for a specific point on the trace shows: "Tue, Jul 15, 2008, 2:53 pm - 4:59 pm", "Carbon Impact: 14.43 kilograms", "Sedimentation Impact: 8.84 mg/m³", "Particulate Exposure: 0.23 hours", and "Fast Food Exposure: 0 hours". The sidebar on the right contains the "peir" logo, the text "personal environmental impact report", and the question "How have you interacted with the environment lately?". Below this are "UPLOAD" and "SHARE" buttons, each with a corresponding icon. At the bottom of the sidebar, there is a section titled "Personal Environmental Impact Report" with a sub-header "How I interact with the environment...". It contains two bar charts: "Impact Rank 2 of 4 friends" and "Exposure Rank 1 of 4 friends".

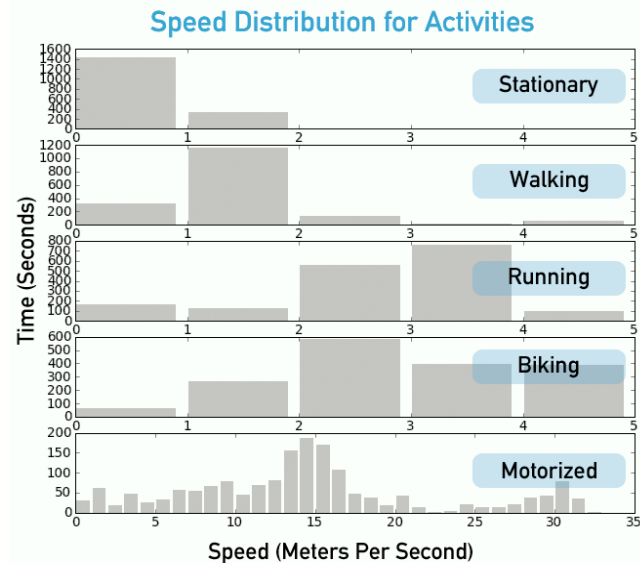
Category	Rank	Friends	Value
Impact Rank 2 of 4 friends	Mo	48.4	
	Friends	46.30	
Exposure Rank 1 of 4 friends	Mo	14.3	
	Friends	20.18	

Current as of: 06/24/2008 12:27:18
A CENS project powered by Nokia

Activity estimated from location trace on secure servers



Map Matching



Activity Classification

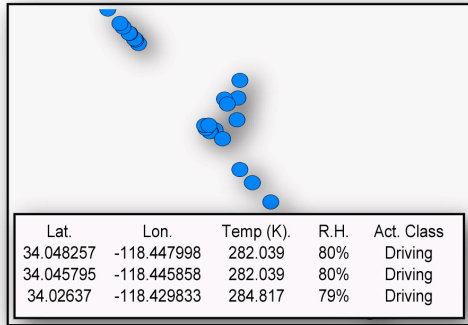
Automatically annotate GPS trace: still, walking, driving, cycling, bus, train
filter anomalous GPS points;
map match freeway;
GPS speed feature;
decision tree w/6 speed/freeway combinations;
HMM recognition;
trip chunking w/configurable dwell time (10 mins)

[challenge: robust and parsimonious activity classification]

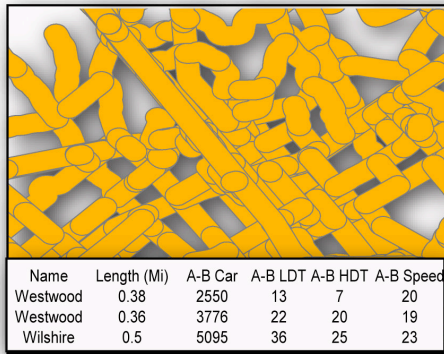
Location-Activity Trace processed through scientific models

Location Trace Processing

Location Trace + Weather
+ Activity Classification

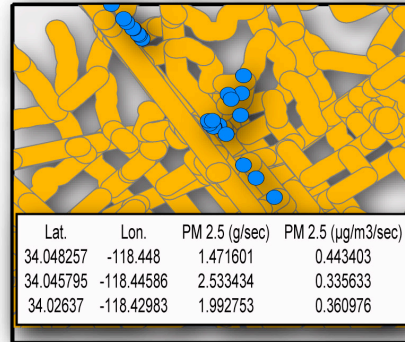


Road Buffers



EMFAC

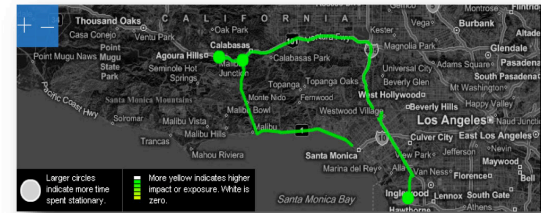
Trip Aggregation



Trip Summary

Trip	Avg. PM 2.5 Exposure	time spent over 0.112716 (hours)
2432	0.470002	0.15
2423	0.235333	0.27
2410	0.210576	0.04

PEIR U.I.



FILTERS → BEFORE OR DURING: mm/dd/yyyy TRIP TYPE: -- SORT BY: --

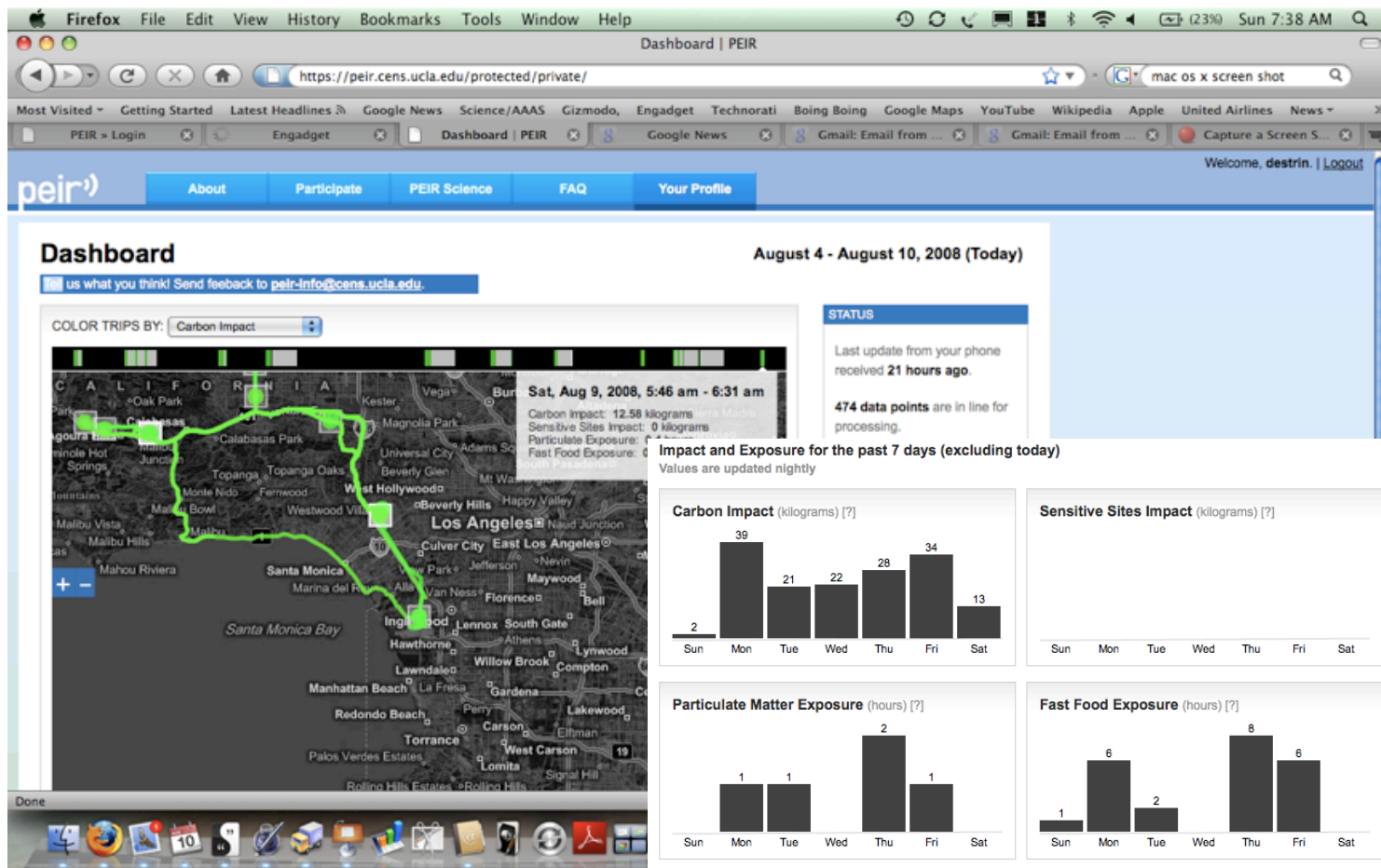
0 Trips selected [Back to top](#)

Trip Start Time	Trip Type	Duration (hours)	Carbon Impact (kilograms)	Particulate Exposure (hours)	Fast Food Exposure (hours)	Sensitive Site Impact (kilograms)
<input type="checkbox"/> May 17, 08:16:39am	stationary	2.48	0	0	0	0
<input type="checkbox"/> May 17, 08:16:26am	traveling	0.22	0	0	0.08	0
<input type="checkbox"/> May 17, 08:16:07am	stationary	0.31	0	0	0.3	0
<input type="checkbox"/> May 17, 08:5:28am	traveling	0.63	0	0	0.21	0
<input type="checkbox"/> May 17, 08:8:35am	stationary	0.11	0	0	0	0
<input type="checkbox"/> May 16, 08:16:45pm	stationary	4.03	0	0	0	0
<input type="checkbox"/> May 16, 08:16:02pm	traveling	0.71	0	0	0.22	0
<input type="checkbox"/> May 16, 08:8:22pm	stationary	0.66	0	0	0	0
<input type="checkbox"/> May 16, 08:8:08pm	traveling	0.21	0	0	0.03	0
<input type="checkbox"/> May 16, 08:8:17pm	stationary	0.66	0	0	0.65	0
<input type="checkbox"/> May 16, 08:7:27pm	traveling	0.9	0	0	0.03	0
<input type="checkbox"/> May 16, 08:0:14am	stationary	1.85	0	0	0	0

[Back to top](#)

[challenge: tracking uncertainty/confidence, web services]

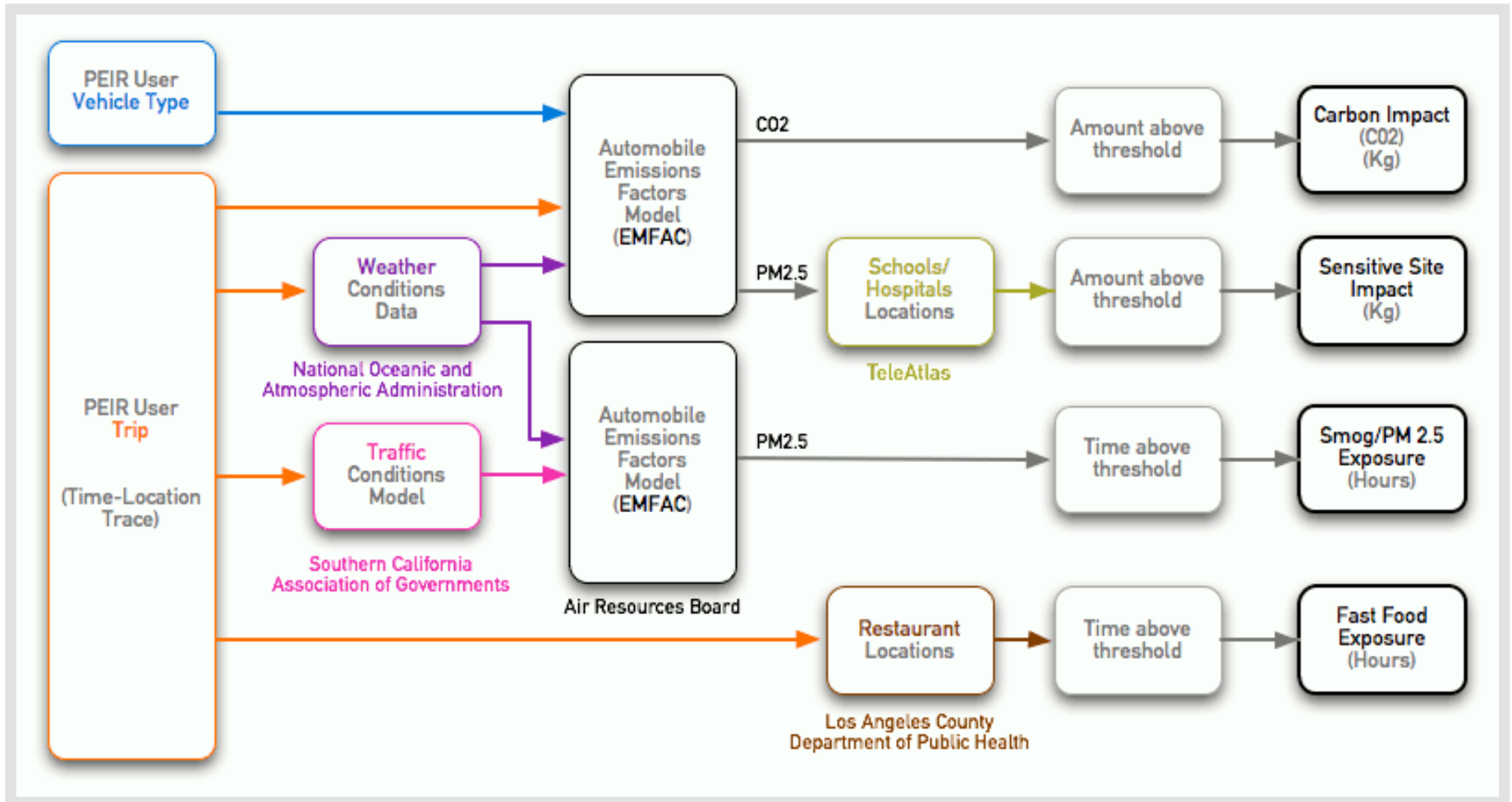
A week in PEIR



- User interface designed to promote data exploration and legibility
- User's data exploration begins with trip log
 - trip list sortable by model (e.g., most carbon impact/most particulate matter exposure)
 - calendar used to advance directly to specific points in time

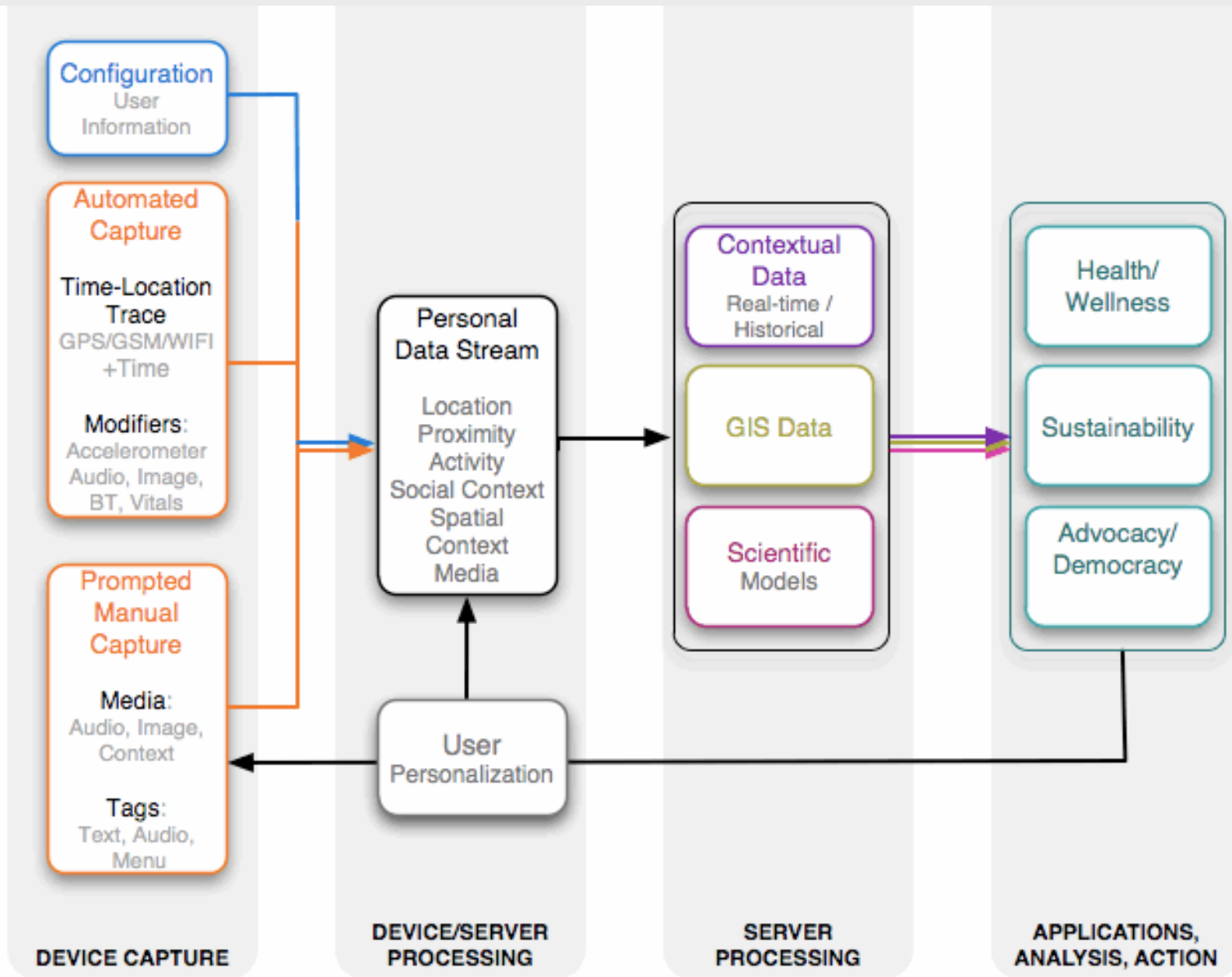
[challenge: visualization]

Non-trivial processing pipeline

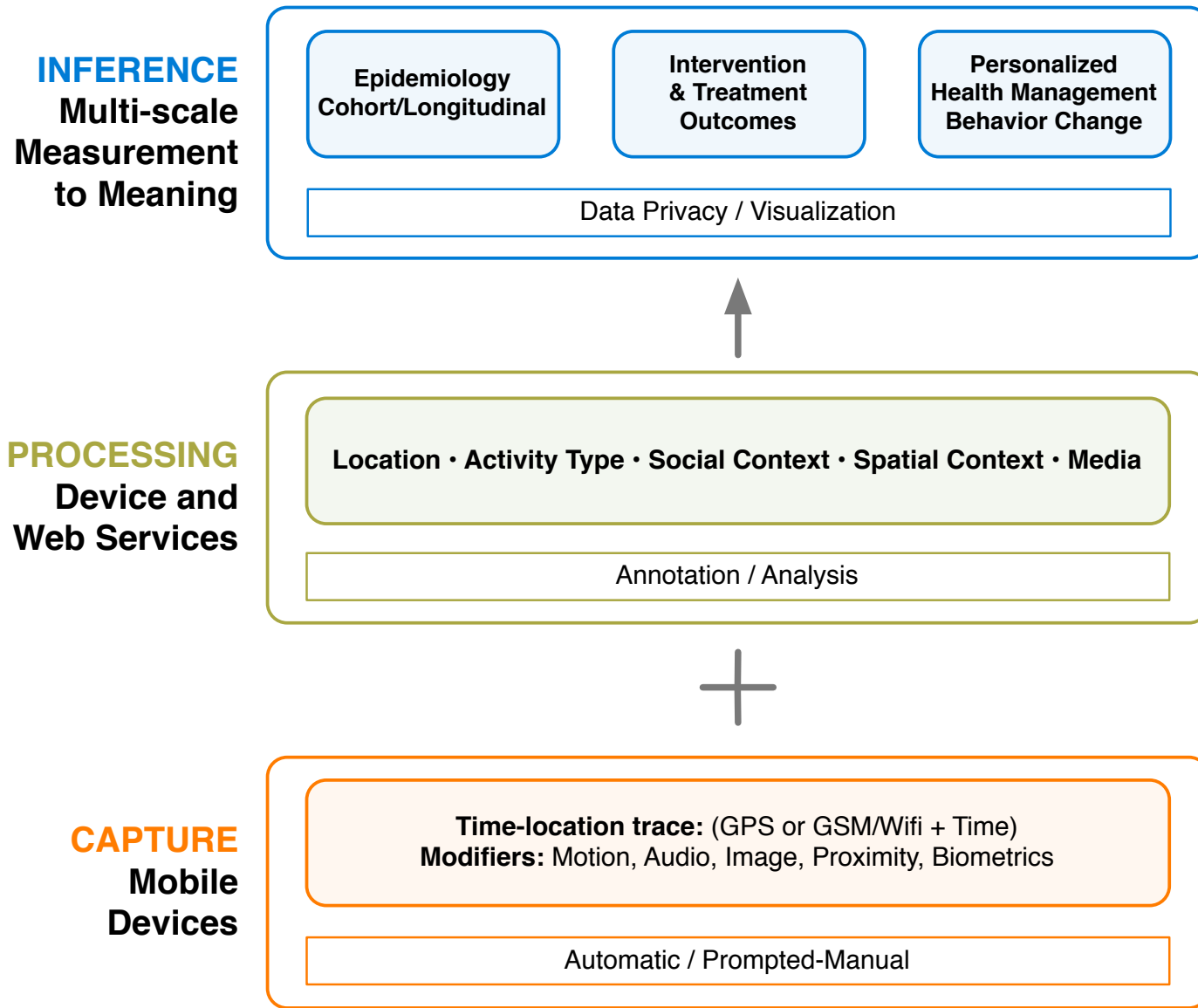


[challenge: robust modular processing pipeline and scalable web services]

Emerging mobile personal sensing system architecture

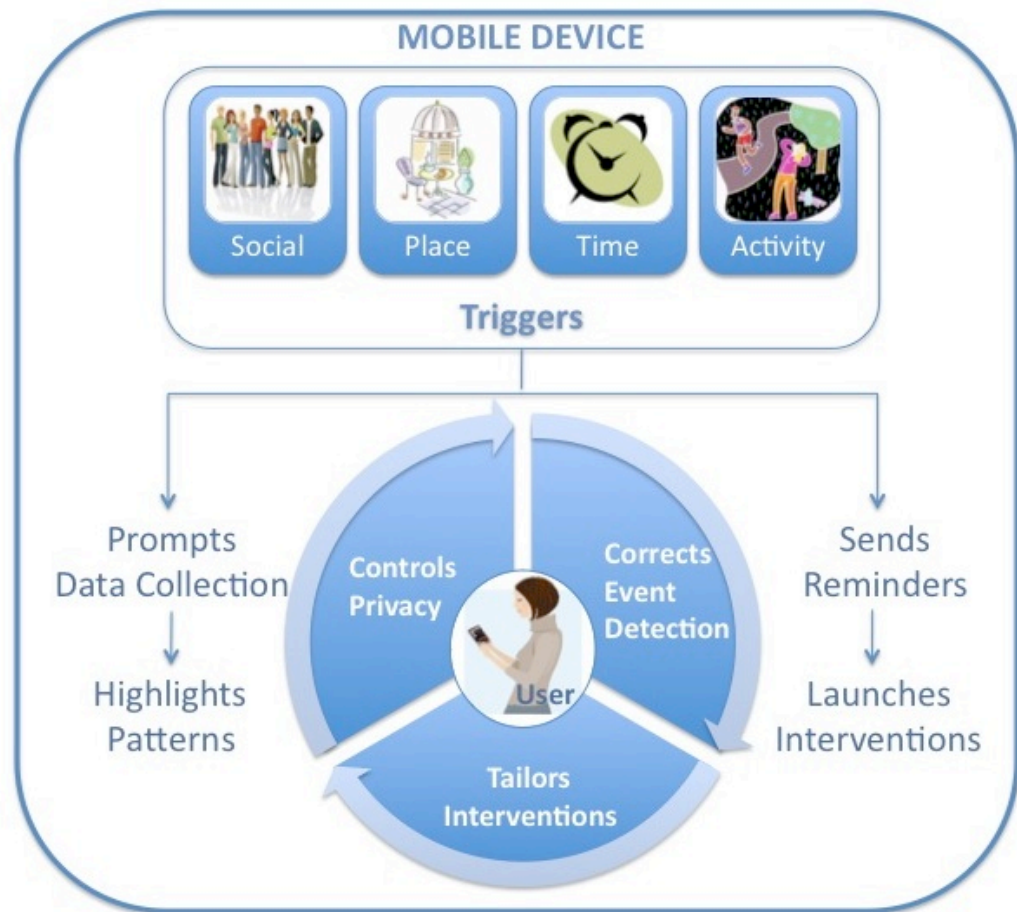
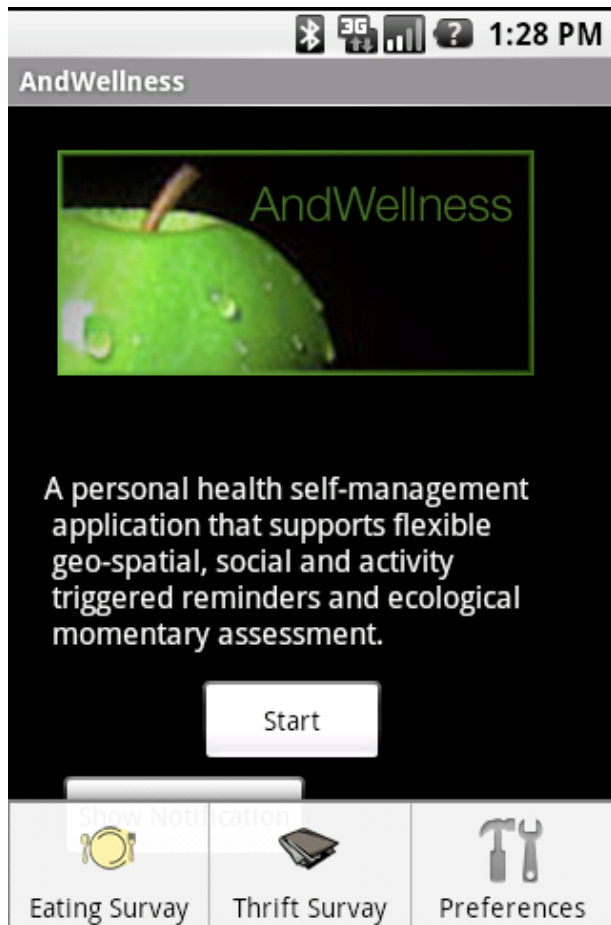


Mobile personal sensing for health and wellness



AndWellness

Personal health self-monitoring and management: powered by automated location and activity detection, adaptive interface, integration with web assets

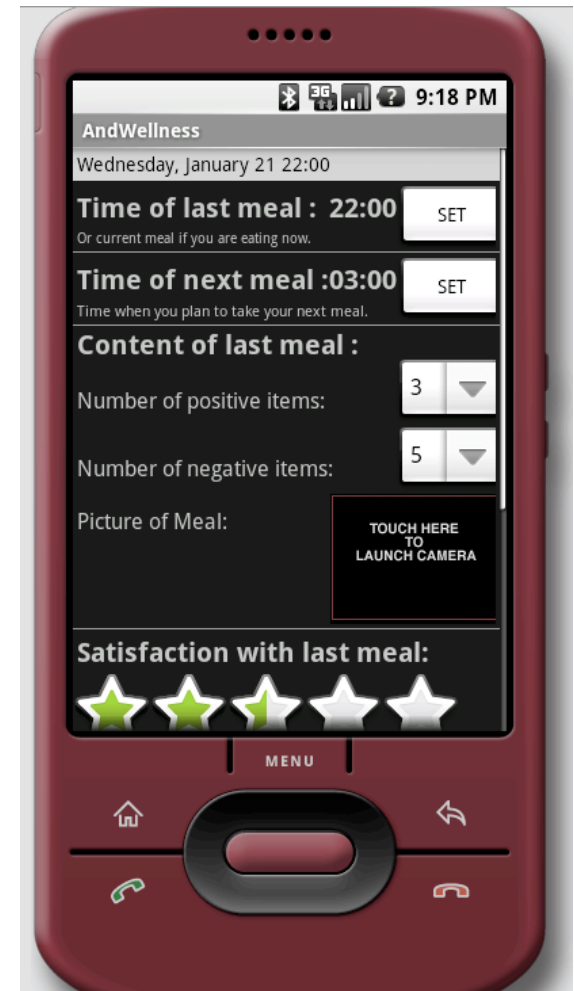


Adaptive event detection

- Automatically identify triggers to launch interventions or prompt user to record/journal.
 - Triggers based on the user's activity, place, temporal, or social context
- Incorporate user feedback to adapt to individuals and their environment.
 - Automatically groups similar events; only solicit feedback from users to label a group of unknown events, or to correct erroneous labels
 - Use lightweight clustering algorithms that are easy to modify

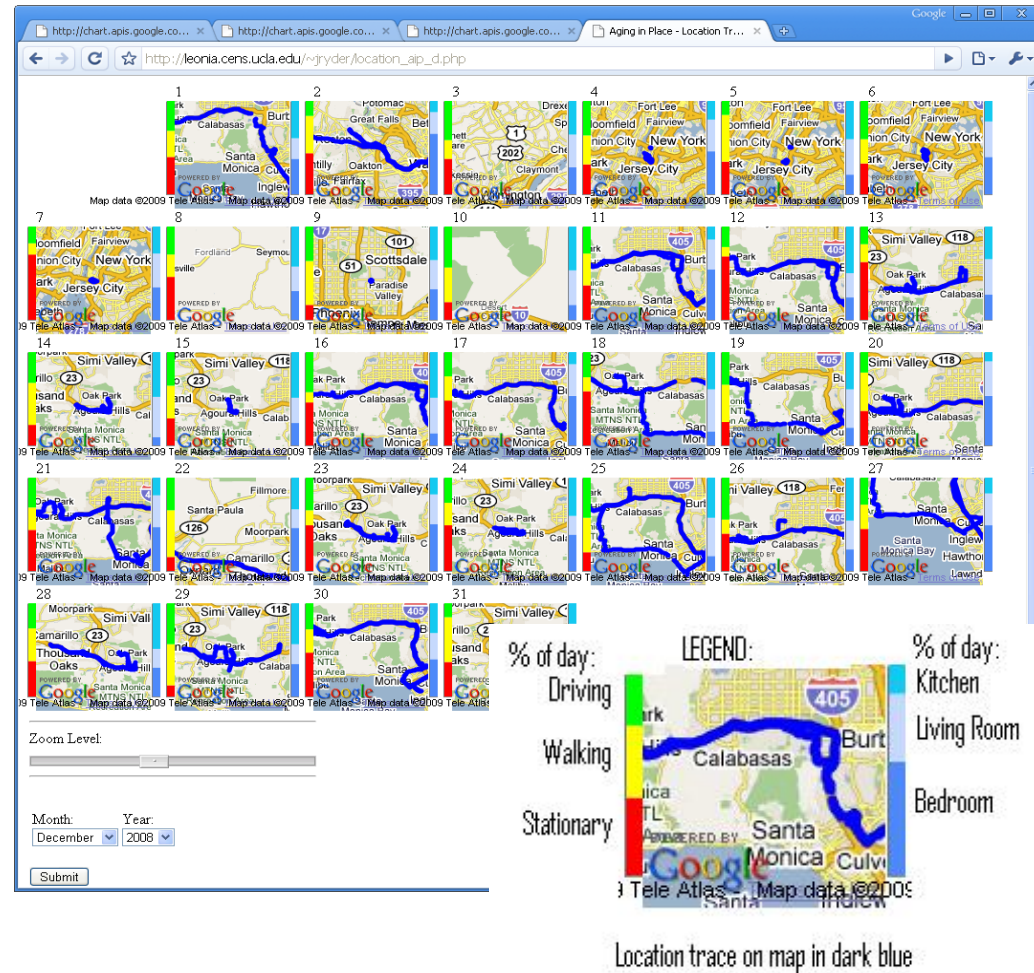
Partially automate services and incorporate limited user feedback to **shift the locus of control to the user while remaining non-invasive**

[challenge: adaptivity, learning, usability]



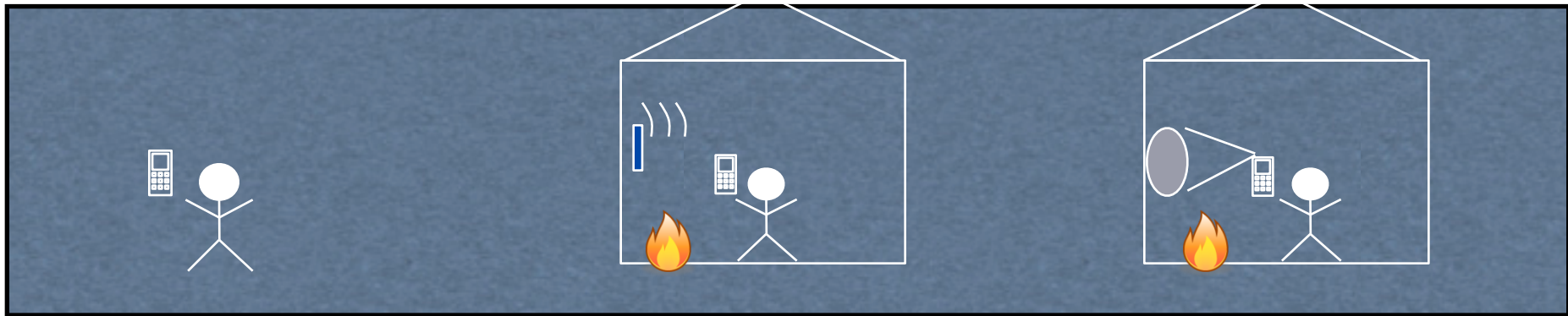
Related applications: activity and mobility profiles for those aging in place, or those managing chronic disease/conditions

- Observe patterns and trends in indicative activities of aging participants:
 - timing and frequency of trips to store, social activities, exercise routines
 - daily patterns of time spent in kitchen, dining area, TV room, bath/bedroom...
- Outdoor: time series of GPS and cell tower data points, combined with map matching
- Indoors: accelerometers and bluetooth stumbling



Community health application: monitor villagers' pollution exposure before and after introduction of clean cook stoves

profile participants' daily activities and exposure to indoor air pollution in unprecedented detail using mobile phone based location and proximity traces



Outdoor Activities are inferred from GPS and accelerometer data

Duration of exposure to cooking fires inferred when a user is in range of a Bluetooth temperature sensor in the kitchen.

Pollution Levels inferred from images of a special filter installed in the house



Epidemiologists at Sri Ramachandra University will deploy the cell-phone tool along with surveys and professional observation to evaluate Project Surya's impacts on the health of villagers.

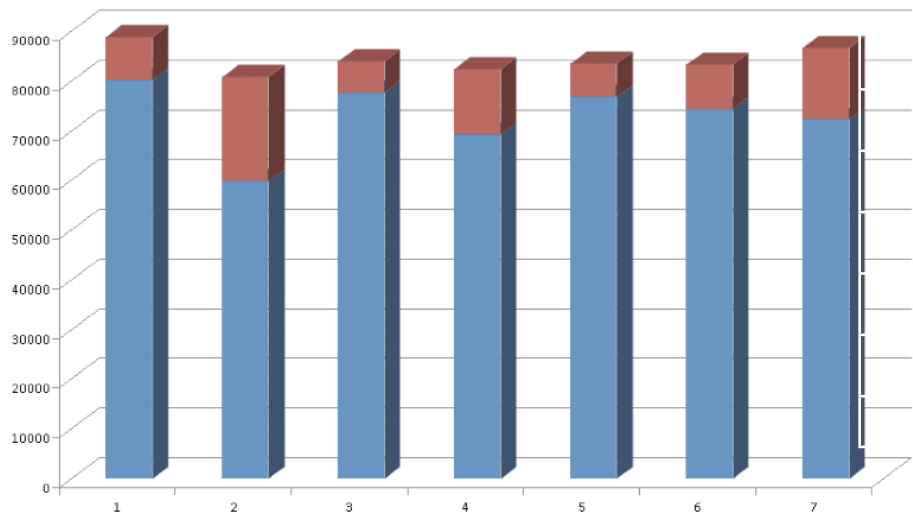
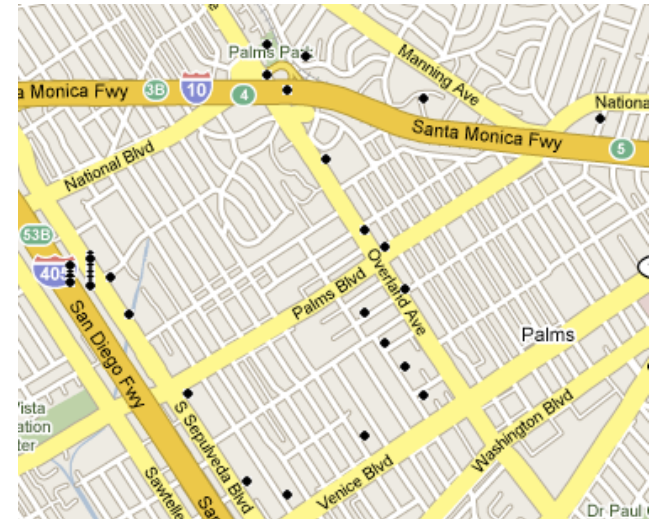
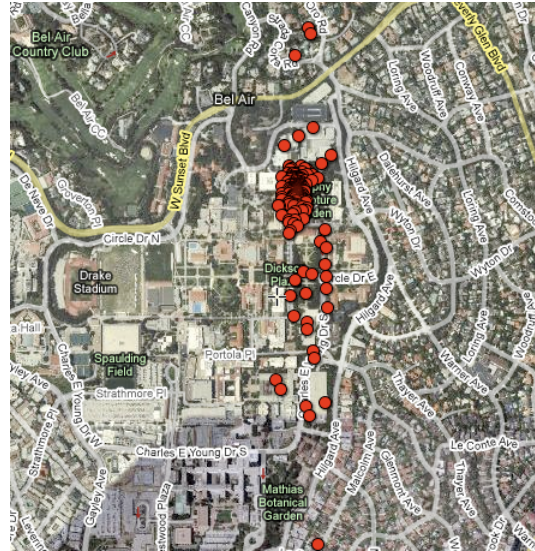
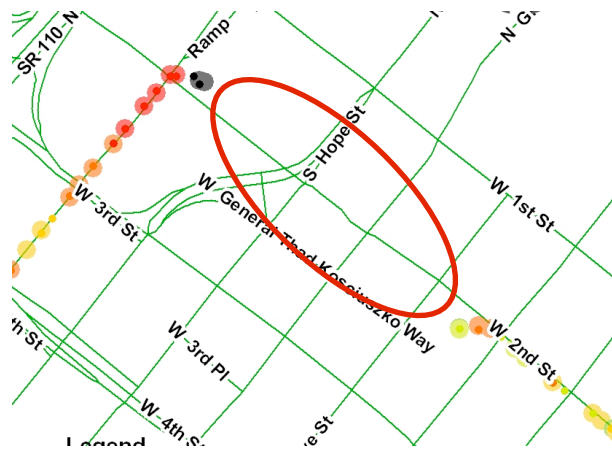
Key building block: activity classification

Mobile phones as a tool for introspection into the habits and situations of individuals and communities -- requires contextual information such as transportation mode

Extensive prior art

GPS, Contextual Models	Patterson 03 Liao 04,05,07 Zheng 08	<ul style="list-style-type: none">- Models are too complicated to perform other tasks- GIS data is not always readily available
GSM	Anderson 06 Sohn 06	<ul style="list-style-type: none">- A large portion of standard mobile phones does not release the information of multiple cell towers in range- They did not attempt to leverage smaller cell-size data such as Wi-Fi
Bluetooth	Tapia 04	<ul style="list-style-type: none">- Bluetooth data is inappropriate to infer mobility states with different speed values because it is not practical to have static Bluetooth sensors distributed ubiquitously in outdoor settings. Also, it is difficult to distinguish whether an individual is moving or if the environment around him or her is changing
Wi-Fi	Bahl 00(RADAR) Ladd 02 Krumm 04(LOCADIO) Griswold 02 Muthukrishnan 06	<ul style="list-style-type: none">- Wi-Fi data targets indoor environments with known access points and tower locations for localization.

Drawbacks of Using only GPS Data: coverage indoors/built areas, power draw



Activity	Power(Watts)
Phone Idle	0.054
GSM Sampling	0.056
GSM, WiFi Sampling	0.23
GPS Outdoor Sampling	0.407
Accelerometer Sampling	0.111

System Challenges: Phone Battery Life

Problem:

background applications should not jeopardize primary applications working for full charge cycle....

continuous gps sampling and uploading can reduce phone battery life to less than 12 hours

Approach:

adapt operation rate or fidelity based on user preferences or current battery level

Need models of:

battery life, user's charging behavior, energy consumed by "legacy applications"

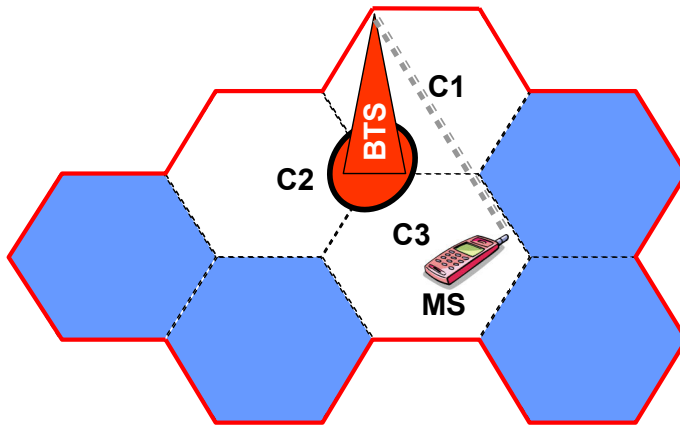
energy-performance trade-off of adaptive applications



Consider GSM and Wi-Fi

Cell tower locations can be used to **roughly indicate a user's location**.

Cell sizes in urban areas are **small/medium**; and density of BSs is **high** [cell-ID location technique, limits and benefits: an experimental study, WMCSA 04].



Features to leverage

Number of Unique Cell IDs (C unique, w)

Number of Cell ID Changes (C changes, w)

Residence Time in a Cell Footprint (C residence)

Duration of dominant wifi AP visibility

Proportion of duration of dominant wifi

Signal strength variance

- We do not try to find a user's exact location using location of WiFi access points. So neither a priori knowledge nor estimated location of access points are required.

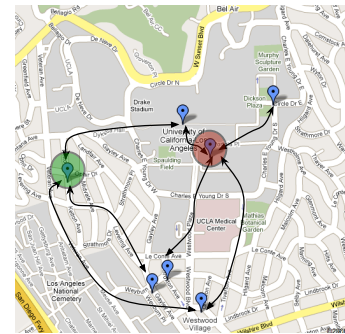
Comparing Mobility Profile for Similarity

- Build a base mobility profile from context information (location/activity).
- Compare periods of mobility information by calculating the similarity of eigenbehaviors for different time periods.

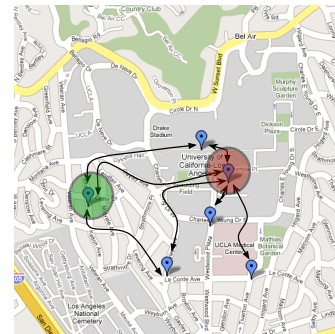
Amount of time spent at work.

	Mon.	Tue.	Wed.	Thu.	Fri.	Sat.	Sun.
8 a.m.	0.1	0.0	0.1	0.2	0.1	0.0	0.0
9 a.m.	0.1	0.0	0.1	0.3	0.1	0.0	0.0
10 a.m.	0.1	0.0	0.1	0.4	0.1	0.0	0.0
11 a.m.	0.1	0.0	0.1	0.1	0.1	0.0	0.0
12 p.m.	0.1	0.0	0.1	0.0	0.1	0.0	0.0

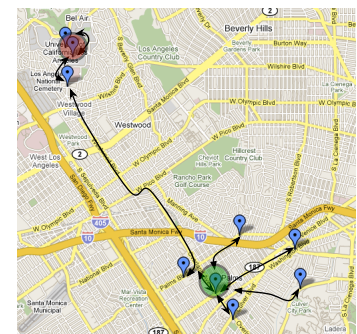
Base Profile



Natural Variations
(Participant at Different Restaurants/Stores)



Re-learning Needed
(Participant Moved)



Leveraging Accelerometer Data for Fine Grained Classification

Transportation mode classifier

- decision tree followed by discrete HMM
- distinguishes among stationary, walking, running, biking, motorized transport

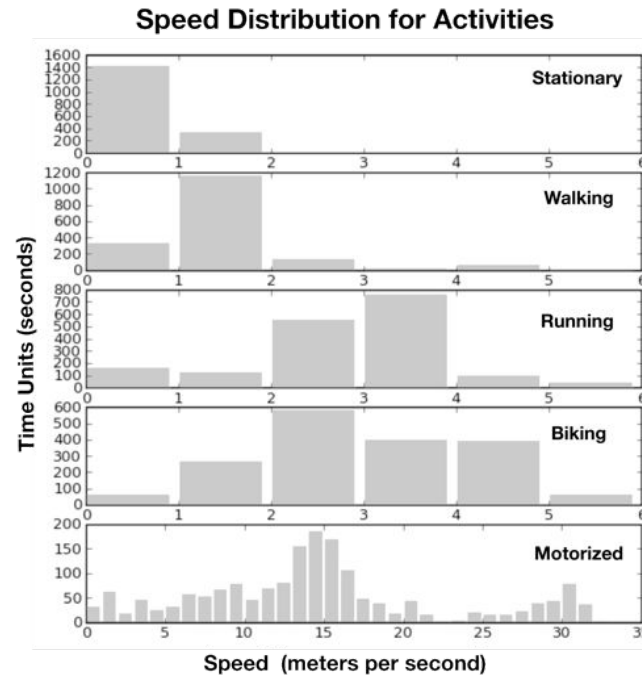
GPS receiver and 3-axis accelerometer as sensors

System does not have strict position/orientation requirements--worn outside or inside of clothes

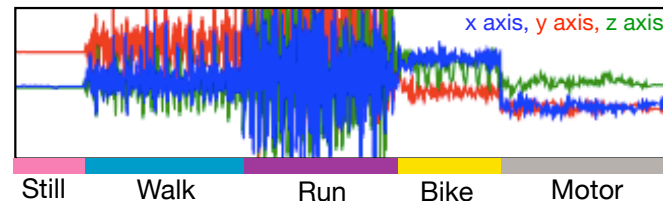
General classifier performance on par with user-specific and location-specific instances.

High accuracy levels in general

- greater than 93% - with user experiment with 16 individuals

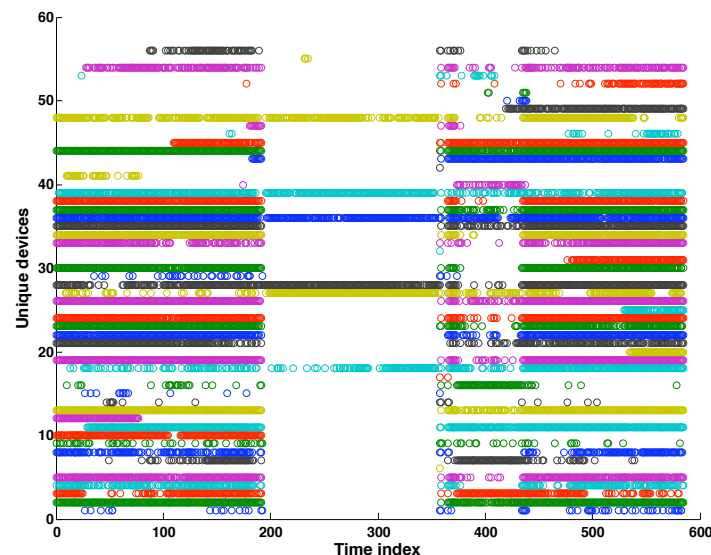


Accelerometer Data of User Carrying Cell Phone in their Pocket

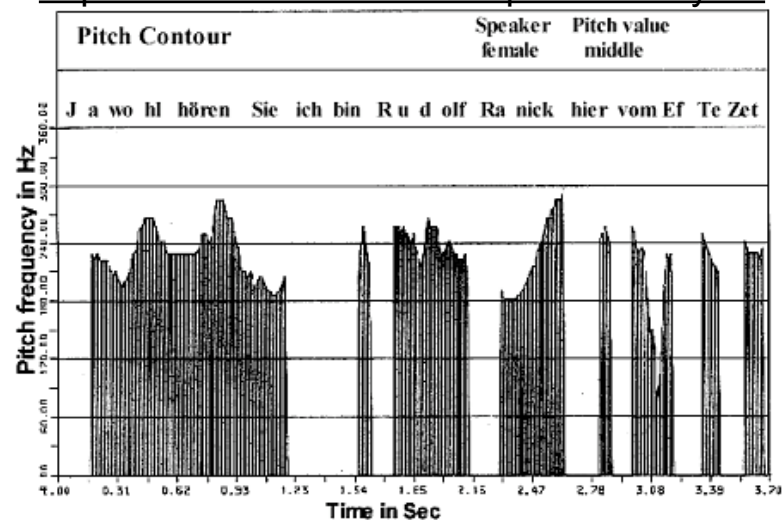


Social activity: an interesting indicator at all stages of life

- Co-location interaction patterns give insights for families
- Near term: use bluetooth proximity
- Mid term: Estimate frequency, duration, trends in human communication using audio samples
 - Programed automatic capture of short audio snippets (avoid content)
 - Processed locally/on-server to detect patterns of interactive communication (distinguish from TV, Radio; phone, in person)
- Observe aggregate data/trends to identify sudden or significant changes in social contact and interaction



<http://www.kt.tu-cottbus.de/speech-analysis/>



[challenge: robust speech feature extraction, training]

Activity classification future work

Adaptive mobility classification system

- different types of sensor data in various situations: e.g. when Wi-Fi APs are too sparse, only use GSM data; accelerometer when GPS speed and map matching makes inference ambiguous
- activate location and activity monitoring to capture outside events: avoid power draw of uniformly sampling when GPS has fix; trigger based on detected GSM-changes.

Opportunities to tune classification method.

- User input or monitoring usage improve accuracy
- Handle cases where features are not available
- Incorporate cost of capturing/processing features
- Use different device models
- Post-process to filter out unlikely series of activities

Key building block: privacy mechanisms and selective sharing

Share derived statistics instead of raw traces

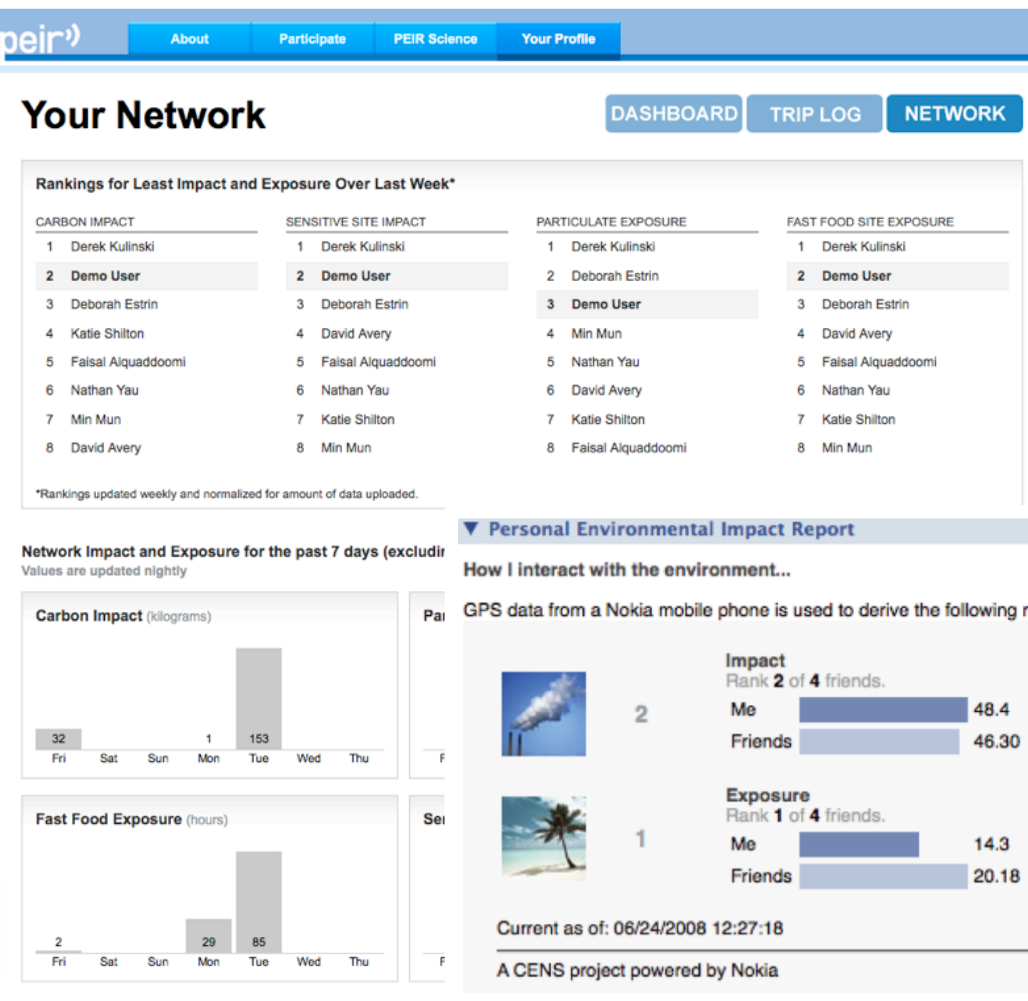
detailed data only accessible to individual

Simple example: peir Facebook app/widget

Research challenges
selective sharing and retention

model-equivalent substitute data

system transparency and audit-trail wrt data use and provenance



Personal Data Stream Control

Individual control of time/space accountability

Location traces quantify habits, routines, associations & are easy to mine

Why control?

Prevent discriminatory practices

Maintain safety/security

Respect social boundaries

Protect stigmatized activities

Full disclosure not inevitable

Selective sharing, hiding, archiving

Information flow control in supporting systems

Abuse never preventable

Need strong audit trails

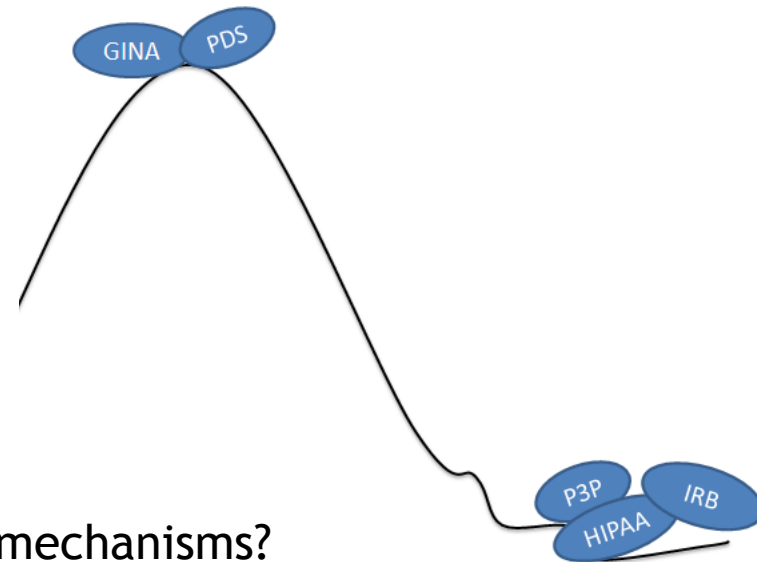
Legibility/transparency

Laws concerning fair use

Can we design lightweight, scalable, but effective mechanisms?

P3P, HIPAA, IRB are not embraced even if they are implemented

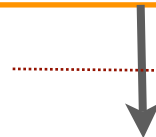
Could become as important as genetic discrimination (GINA)



Personal Data Stream Control Architecture

CAPTURE
PERSONAL
MOBILE DEVICES

Provide Expressive Default Policies
Enable Users to Configure Preferences

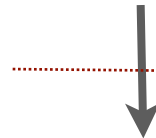


FILTER [Metadata]

*filters define selective
upload*

ARCHIVE/DISPATCH
PRIVATE
DATA VAULT

Decouple Conflicts of Interest
Foster Marketplace of 'Certified' Services



FILTER [Metadata]

*e.g., personal container
in web cloud*

*filters define selective
sharing with various
services*

ACT
APPLICATIONS
/SERVICES

Provide Transparency of Data Access, Uses, Inferences and Manipulations
Enable Audits by Users/Agents

*establish certified
services*

[challenge: scalable architecture, verifiable mechanisms]

'Private Data Vault' Functions

Technical precedents: Mail Servers/Webmail, Yahoo Fire Eagle, Cloud services



Alternatives--Limitations

Physically private store--robustness

Virtually-private store maintained by applications--conflict of interest

Private data vault model

Guards against conflicts of interest

Supports robust storage and archiving

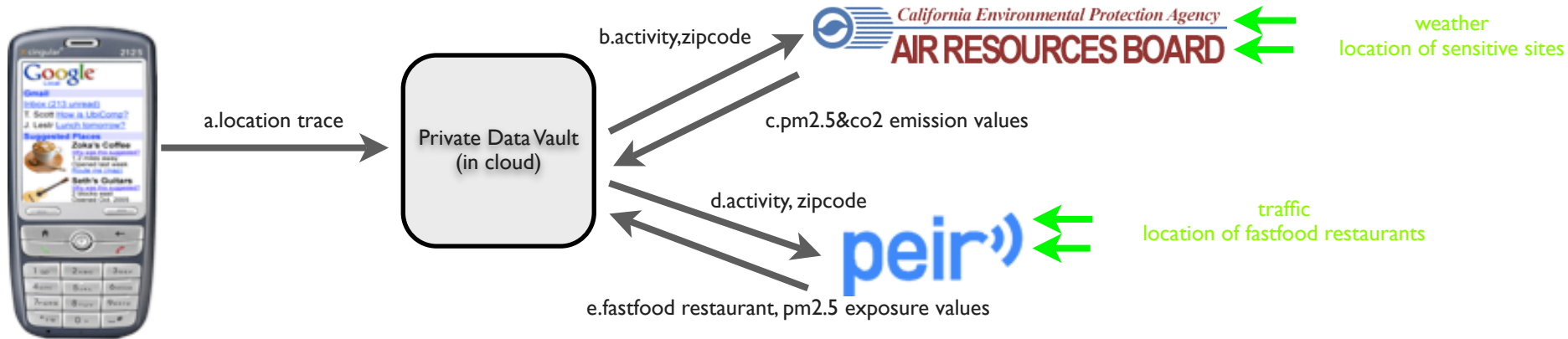
Enable a marketplace of 'certified' applications

But need a vault mechanisms that can export filtered/processed data

[challenge: scalable architecture, verifiable mechanisms]

traceaudit

Enable records of access, use, inference, and manipulation of data at multiple points in the system.

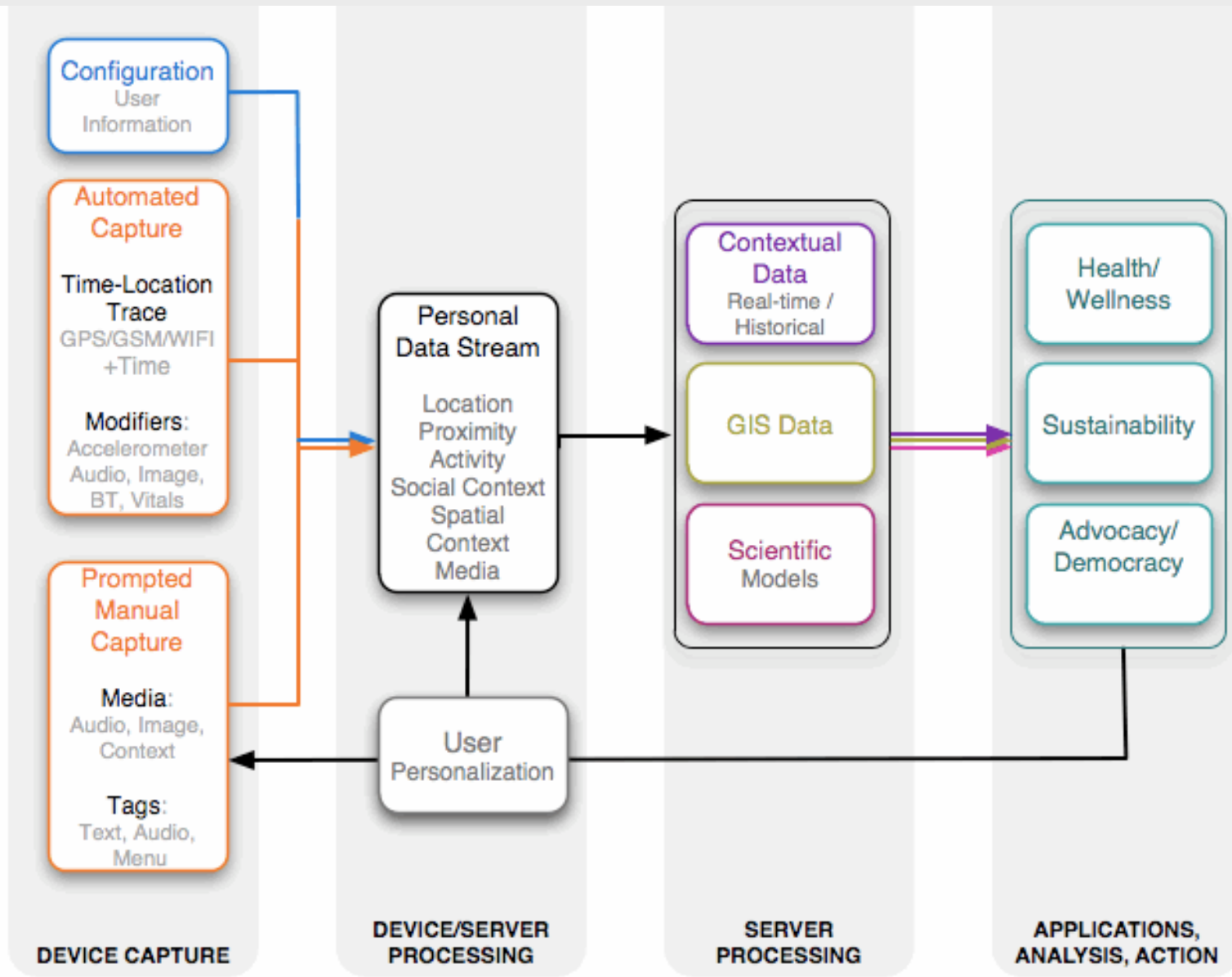


- traces
- outside data sources
- input data sources from users
- inferred data

	0	1	2	3	4	5
	datavault	ARB	datavault	PEIR		
1	locations of freeways	zipcode map				
2	location	activity	zipcode			
3	weather	locations of sensitive sites				
4	activity,zipcode	co2 emission	pm2.5 emission			
5	traffic	locations of fastfood restaurants				
6	activity,zipcode	pm2.5 exposure	fastfood exposure			

An example of metadata to support **traceaudit** and check mechanism services

Emerging mobile personal sensing system architecture



Conclusion

If you can't go to the field with the sensor you want...
go with the sensor you have! (Anon)

The power of the Internet, the reach of the phone (Voxiva)



Acknowledgments-Sponsors

- Excellent foundational and related work around the globe: Univ Washington, MIT, Dartmouth, Columbia, Nokia, Intel, MSR, Yahoo, ...
- NSF: OIA, NETS-FIND Program, CRI Program, CISE, Engineering, Bio
- Nokia, Cisco, Sun, Intel, Samsung, Google, MSR, Crossbow, Agilent
- UC Micro, Participating campuses (UCLA, UCR, UCM, USC, Caltech)
- Wilson Foundation, Conservation International

[YouTube Video on PEIR](#)

