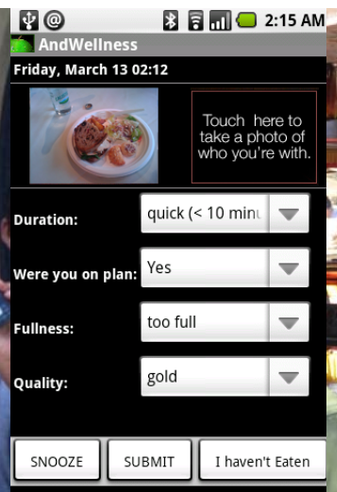
Participatory Sensing
Planning for Life-long Data Streams

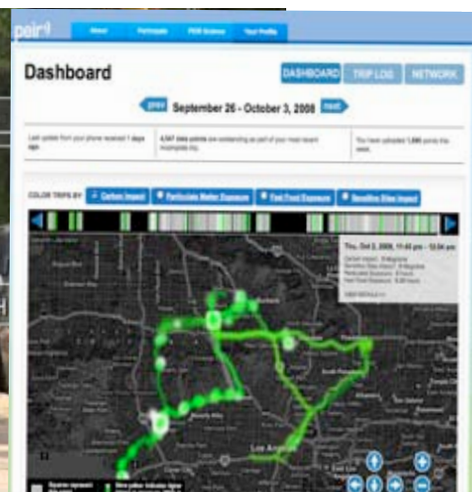
Enabled by over 3 Billion mobile phone owners,
with increasingly feature- and sensor-rich handsets,
and the wireless, cloud-connected infrastructure that grows daily.

Individuals and communities gather data about things that they care for, and make their own decisions about when and how to:

Capture → Store → Access → Analyze → Share

Jeff Burke
Work with Deborah Estrin, Mark Hansen, Mani Srivastava, and many others
UCLA Center for Embedded Networked Sensing (CENS)





Local and Personal Knowledge, *Collected Over a Lifetime*

Real Time (Always on) - **Real Place** (Always carried)

Real Context (Environmental, Social, Behavioral)

Prioritize widely available smartphone technology
rather than specialized sensors. **Data = GPS, cell tower, image, etc.**

***Local* problems & research questions, on a *global* scale**

- Explore personal patterns
- Investigate community issues
- Document neighborhood assets
- Make a case

Tailor interventions to specific individuals or groups

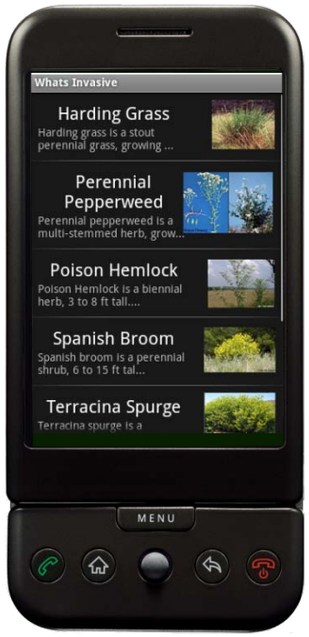
Formalize models of participation, phenomena, knowledge

Inspired by
CBPR in public health
Participatory GIS and P. Design
PhotoVoice
Community organizing

Applications
Health and wellness, epidemiology,
Urban planning, resource management,
Cultural documentation, creative expression,
Civic engagement, citizen science.

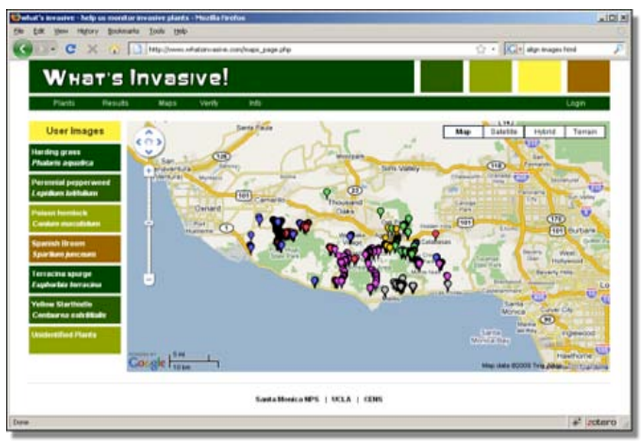
Citizen Data Campaigns: *What's Invasive!*

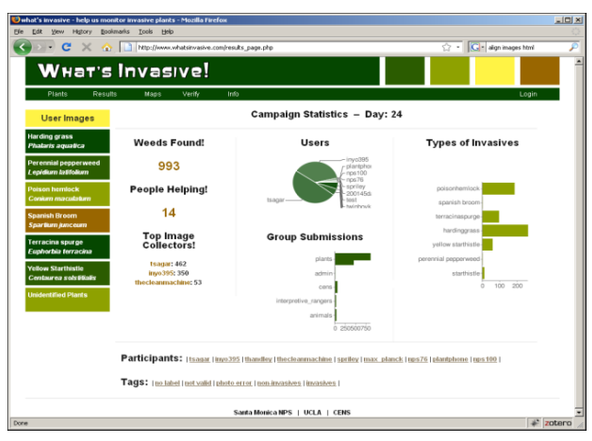
w/National Park Service, Santa Monica Mountains



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





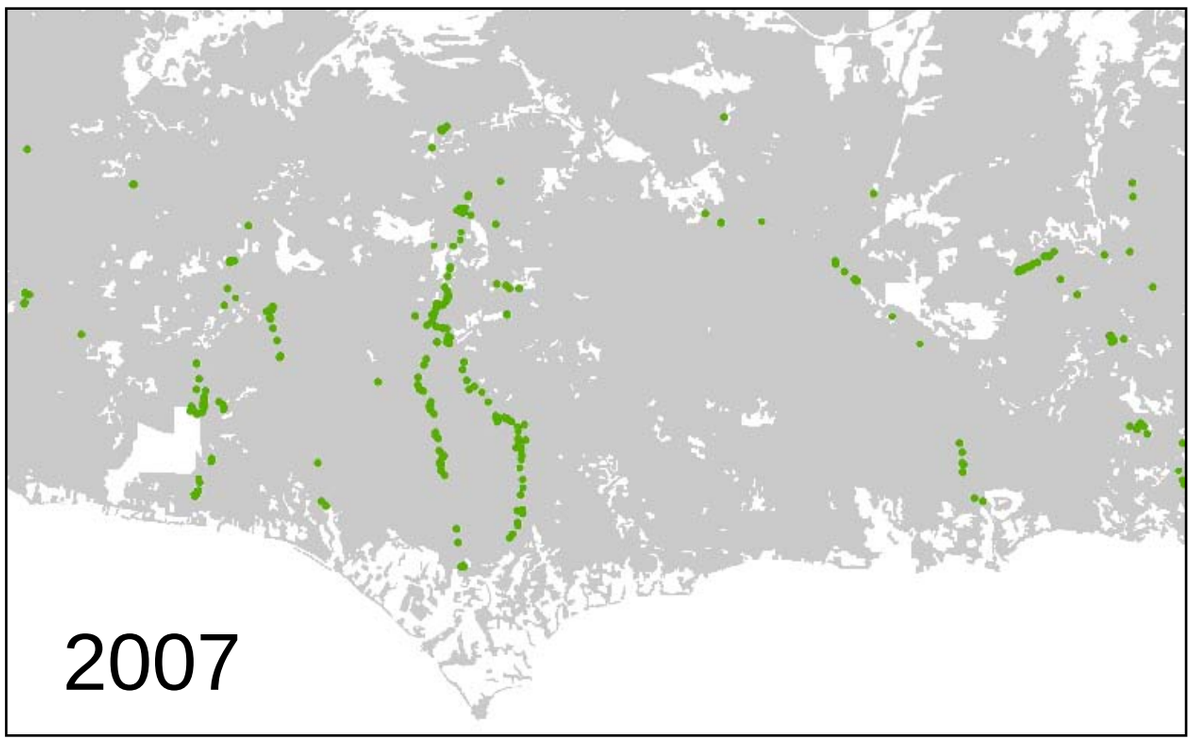
http://whatsinvasive.com

Mobile app available on Android Market and iPhone App store

What's Invasive! Preliminary Results

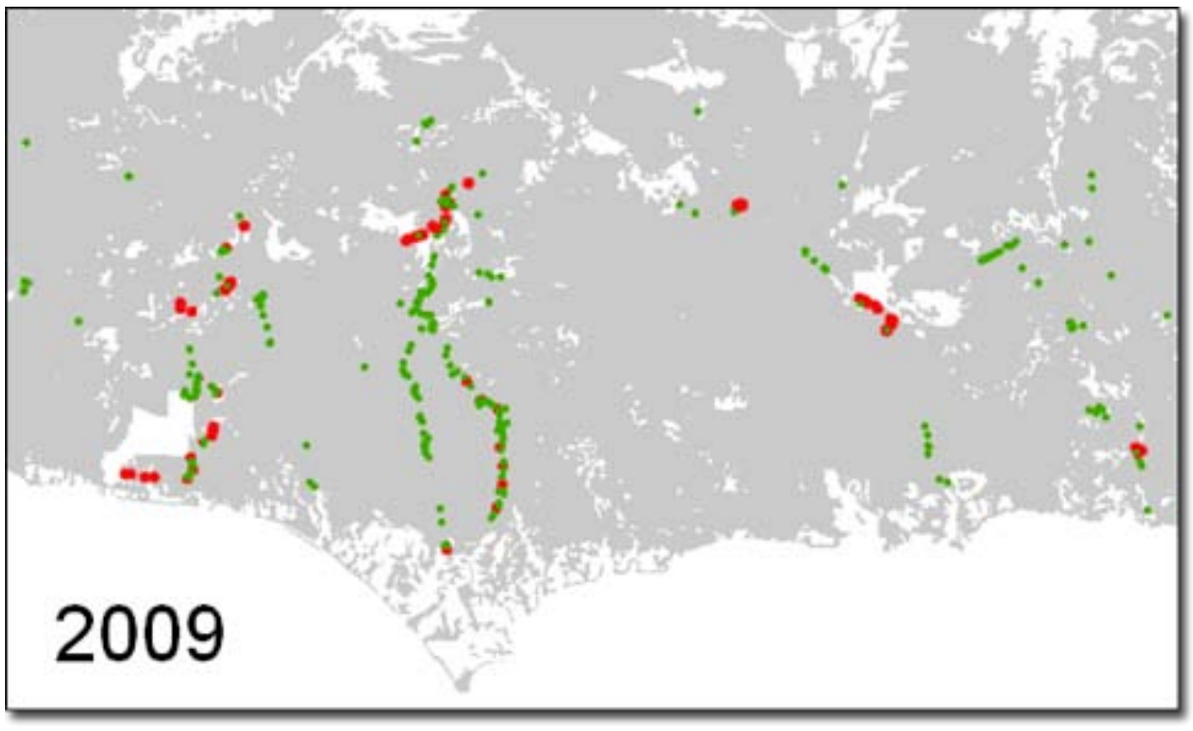
w/National Park Service, Santa Monica Mountains

Manual weed survey ended in 2007 took 2 years and thousands of person-hours to complete.



***What's Invasive!* two-week pilot** have results indicating coverage and quality is comparable.

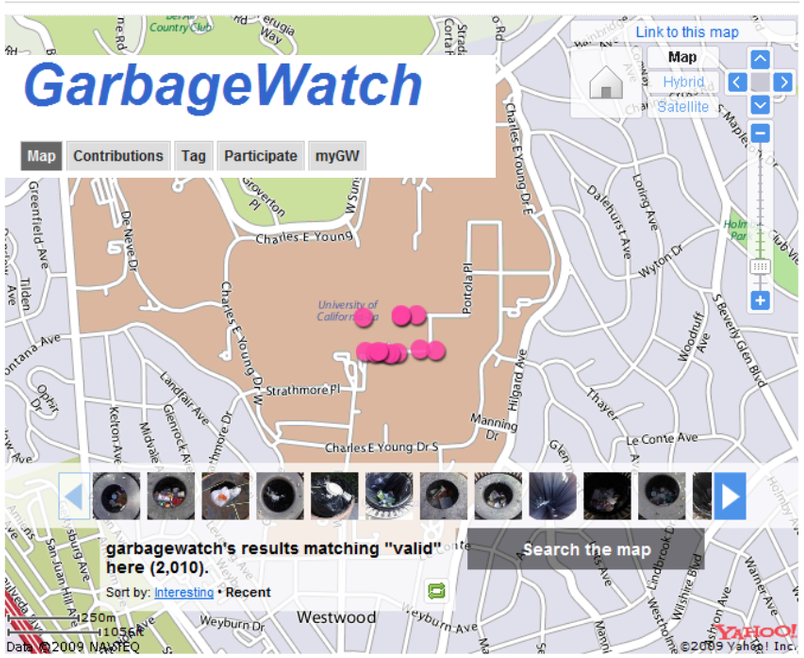
Pilot: 19 users, 1119 records to date
Total NPS Visitors: 272M in FY2007



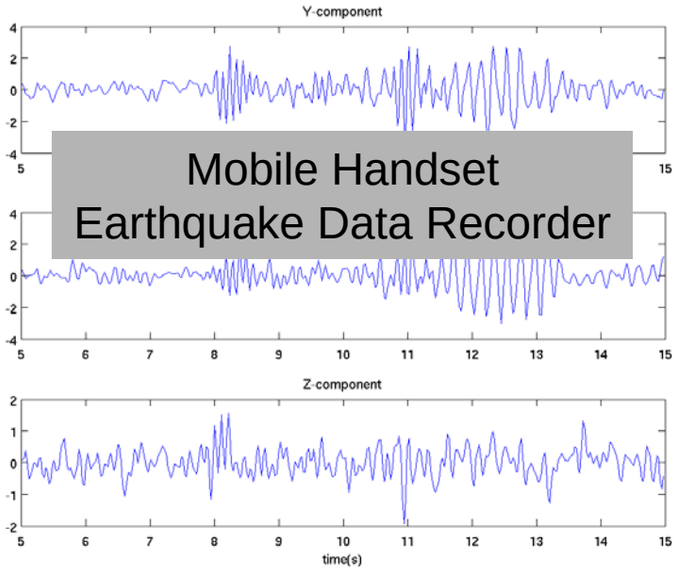
Hansen, Graham, Reddy, ...

Campaigns – Lightweight, easy to create / join

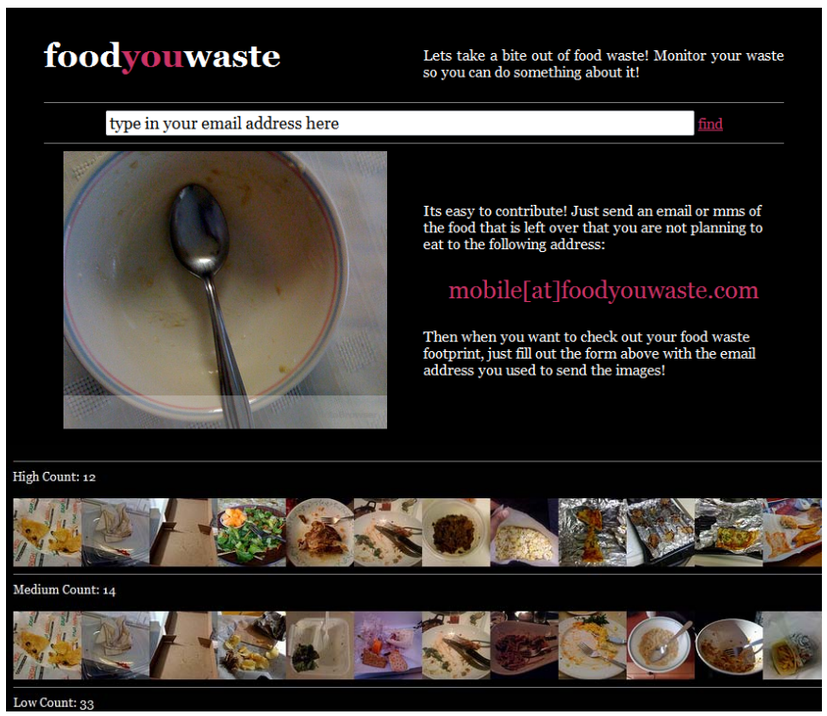
How to collect data only once? And how to manage user interest / availability?



http://garbagewatch.com/



Nature article: “Phoning in Data”



http;//foodyouwaste.com/

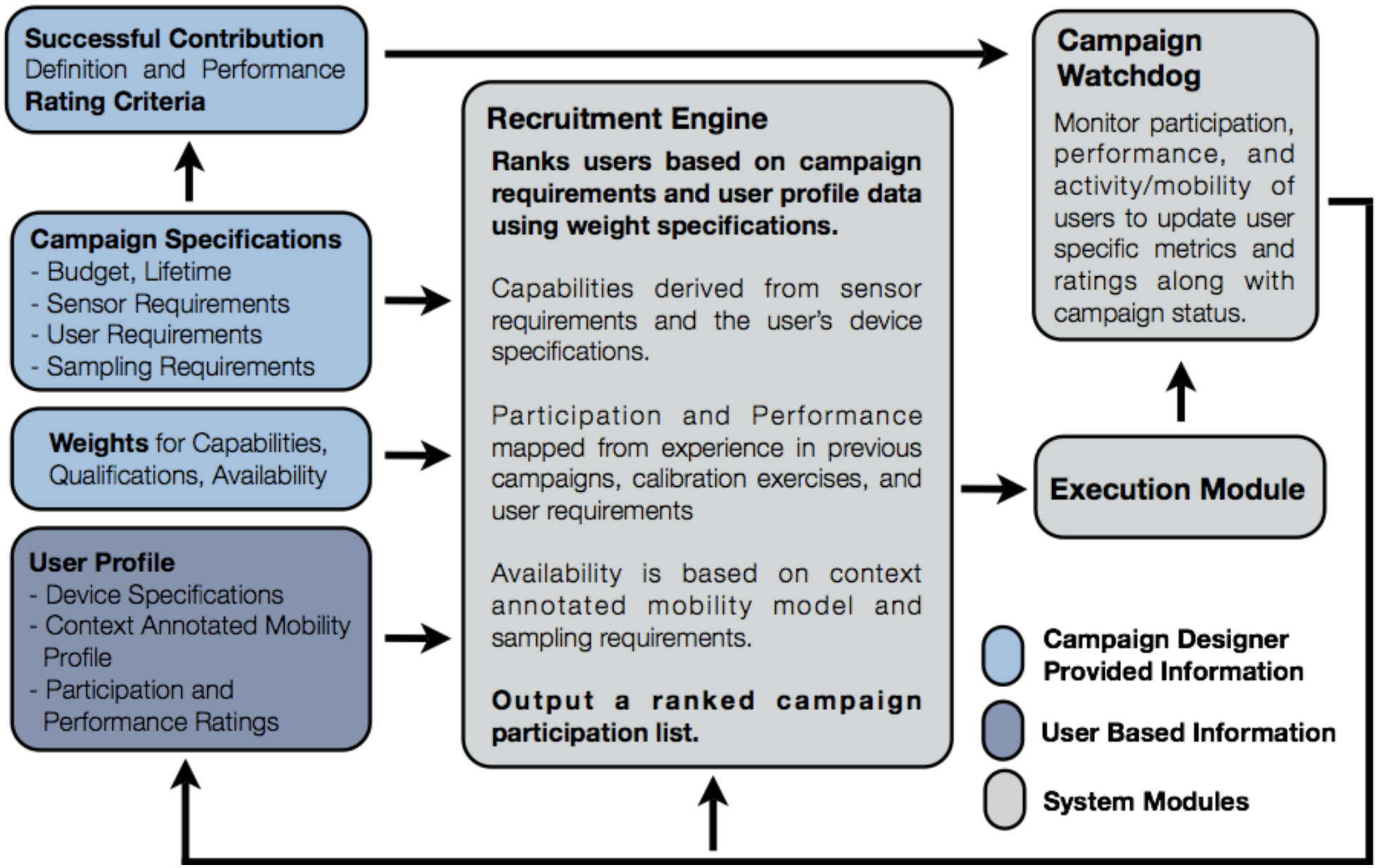
http://la.remap.ucla.edu/remappingla/hollywould

S. Reddy, O. Maldonado, M. Lukac, et al.

Recruitment and Reputation (Are all data equal?)

Balance goals of participation and data quality.

Three stage (qualifications, interview, online review) approach to using context-annotated mobility profiles as a basis for recruitment to campaigns - analogous to a job interview process. Ongoing monitoring of performance during a campaign.



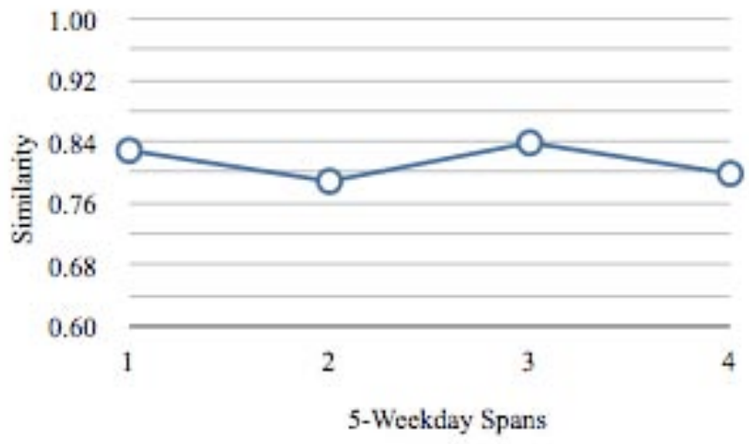
S. Reddy

From CNN to more advanced Availability Monitoring

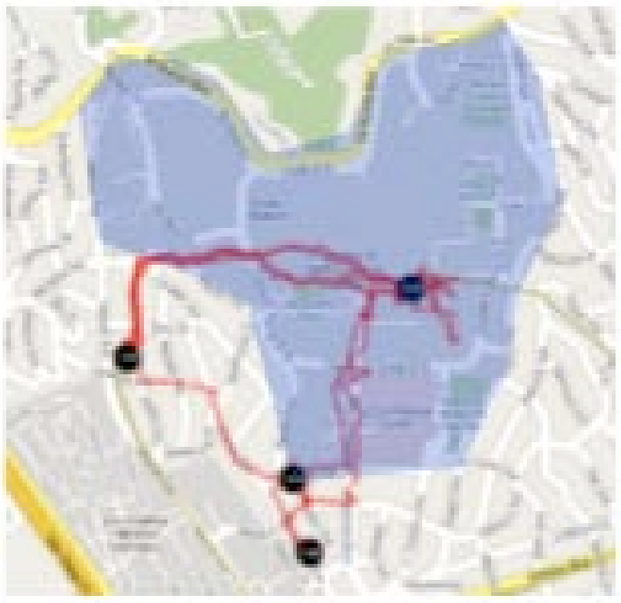
How to provide continual, ongoing feedback on what participatory sensing opportunities are near you? Continuous nearest neighbor, or more…

Profile Similar Over Weeks

Participant Works at CENs :-(

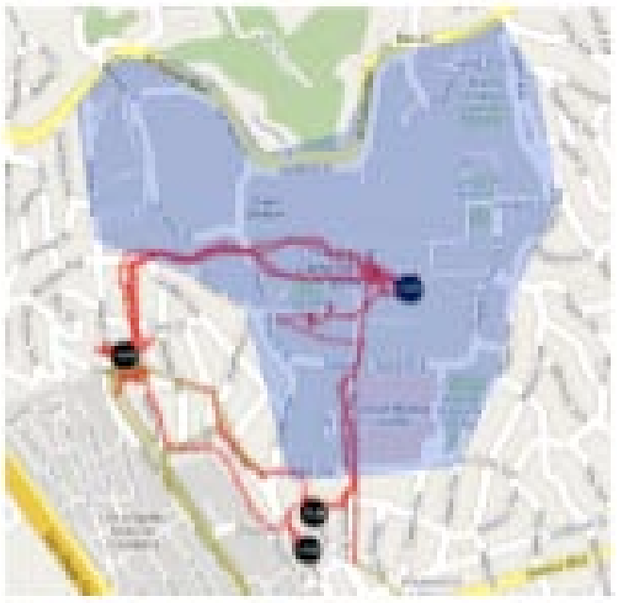


Base Profile



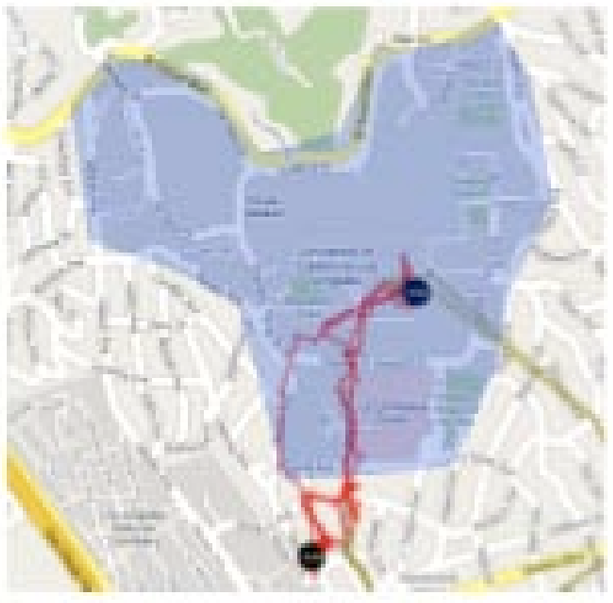
Natural Variations

(Participant at Different Restaurants/Stores)



Re-learning Needed

(Participant Moved)



S. Reddy

Health & Wellness: Ecological Momentary Assessments

Hybrid of time-location trace with media capture and self-report.

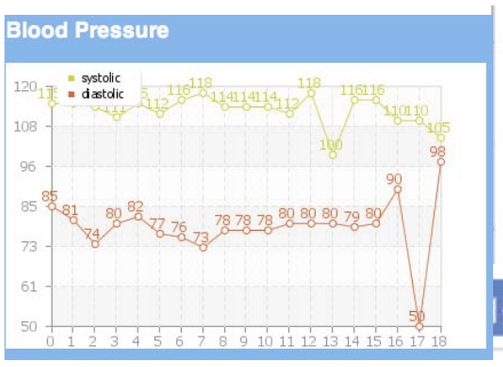
Our Actions

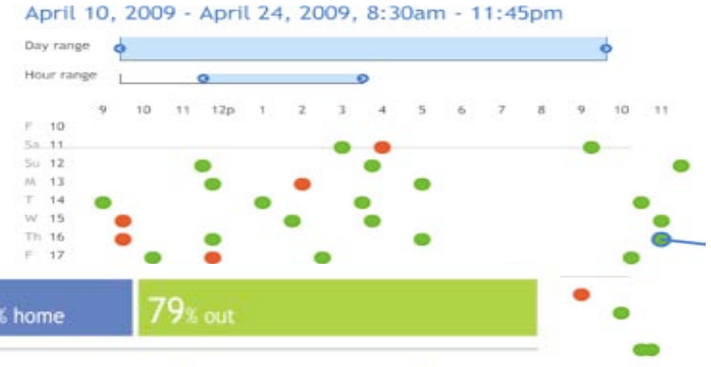
Our Self Report

Geocoded and time-stamped EMAs

Mobility traces

Private Data Storage





aggregate measures, trends, patterns

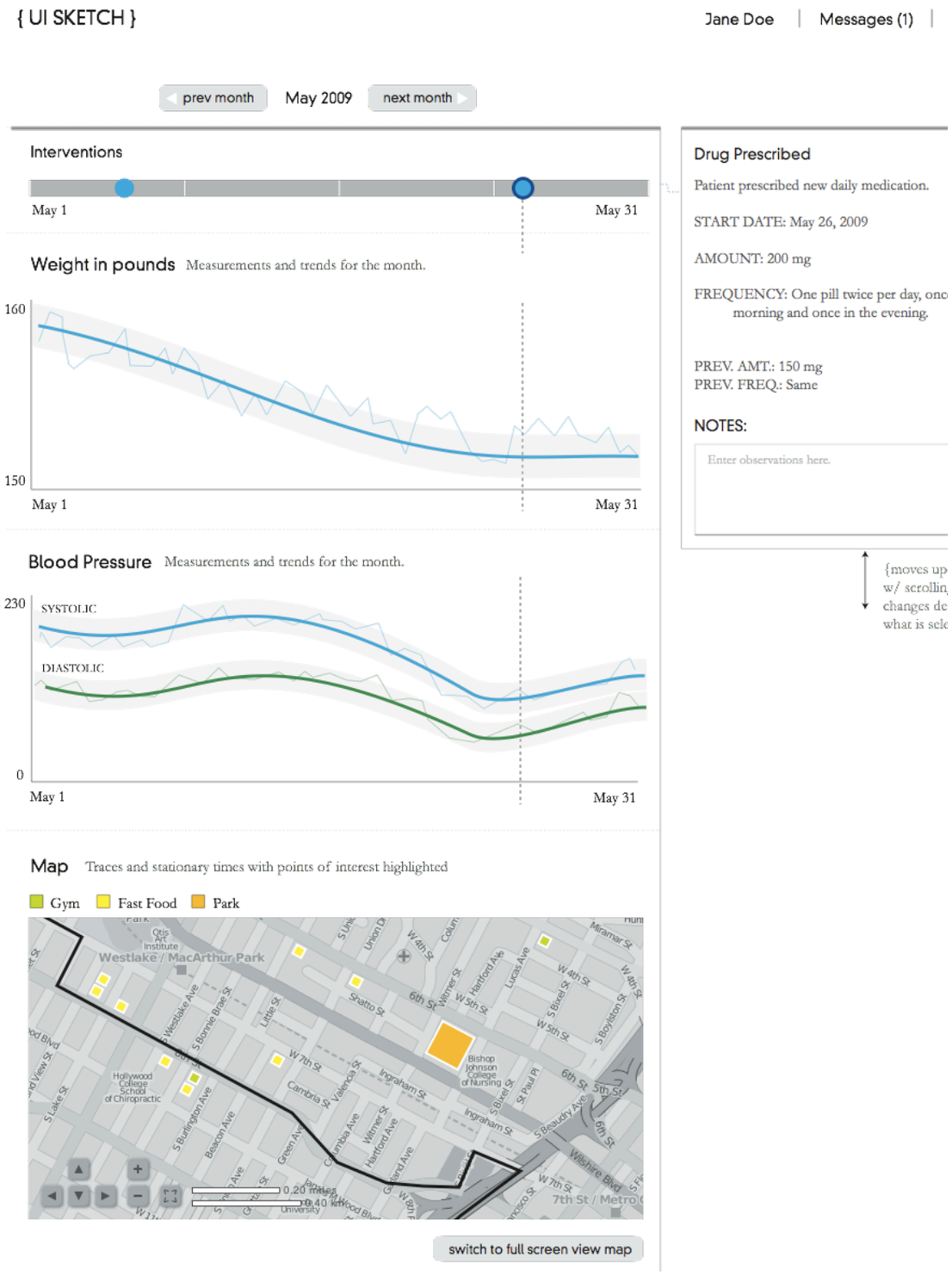
event detection

Processing

Visualization

Photo – Marshall Astor

Data Stream as a *Living Record*



Prompted, geocoded, uploaded:

- Physiological (BP, glucose...) data
- Patient reporting (medication, symptoms, stress factors)
- Location traces and co-location information
- Contextual, environmental, social factors

Applications

- Chronic disease management
- Acute health-behavior change
- Cancer survivor checkups
- Youth preventative health app (obesity/asthma)
- Clinical trial antidepressants
- Research study for stroke rehab protocols
- Promotora/CHW support
- Elder Caregivers

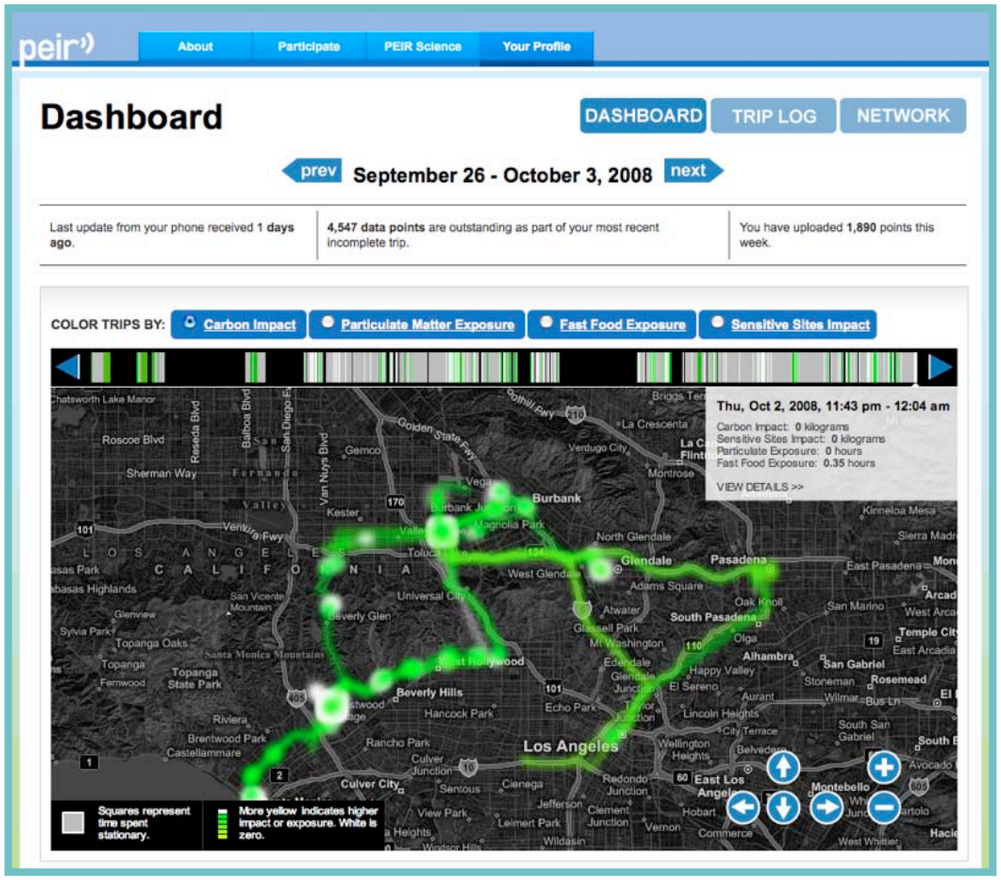
Doesn't require a smartphone: e.g., Twitter interface of http://your.flowingdata.com

PEIR: Personal Environmental Impact Report

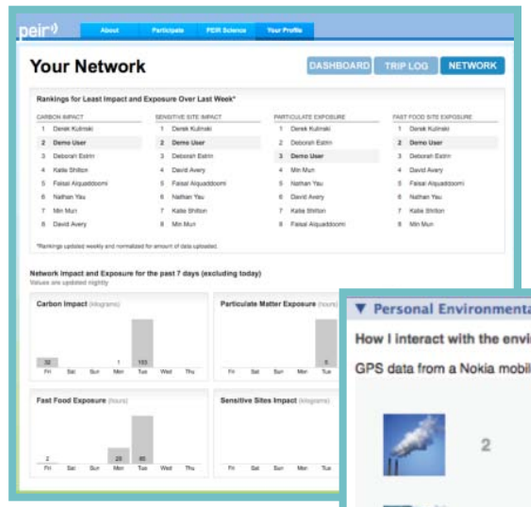
"Mobile-to-web" app using a GPS-equipped smartphone to explore and share how you impact the environment and how the environment impacts you.

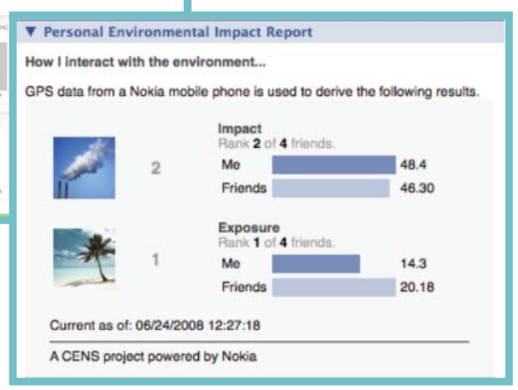
Models & data from: CalEPA/ARB, LA Dept. of Public Health, So Cal Assoc. of Govts

Background Mobile Capture on Symbian, Windows Mobile, Android



Web dashboard





Social network integration

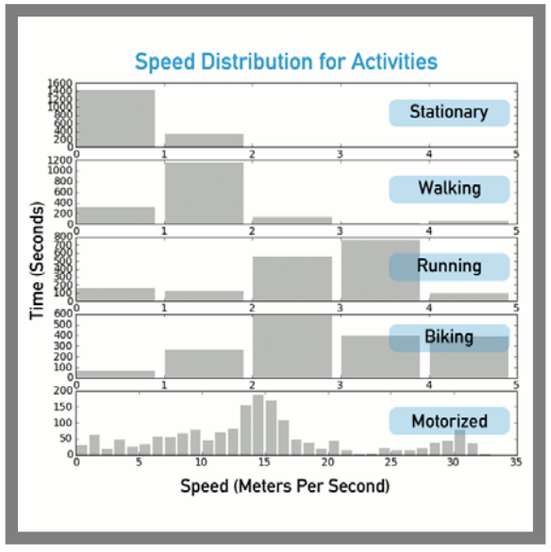
Pilot scale: 40-100 users.

J. Burke, D. Estrin, M. Hansen and many others. http://peir.cens.ucla.edu/

Supported by Nokia Research. http://theprogressproject.com/

Example of Processing: Activity Classification

Map Matching



Activity Classification

Location trace data is processed on server-side.

Multi-step activity classification:

- Filter out anomalous GPS points
- Match to verify if on a freeway.
- Calculate speed.
- Classify using decision tree.
- Classify using Hidden Markov Model (first order).
- Classify trips: “chunk” segments in between 10 minute dwell times.
- Annotate trace with activity. (e.g. still, walk, drive).

Then, data is processed through models like CalEPA/ARB EMFAC or CalLine4 – to go from individual traces to personalized estimates of, for example, PM2.5 exposure and impact.

(Performance objective: Users see fully processed data within a minute or less of a completed trip, as the system scales.)

Sharing PEIR Data with a Social Network

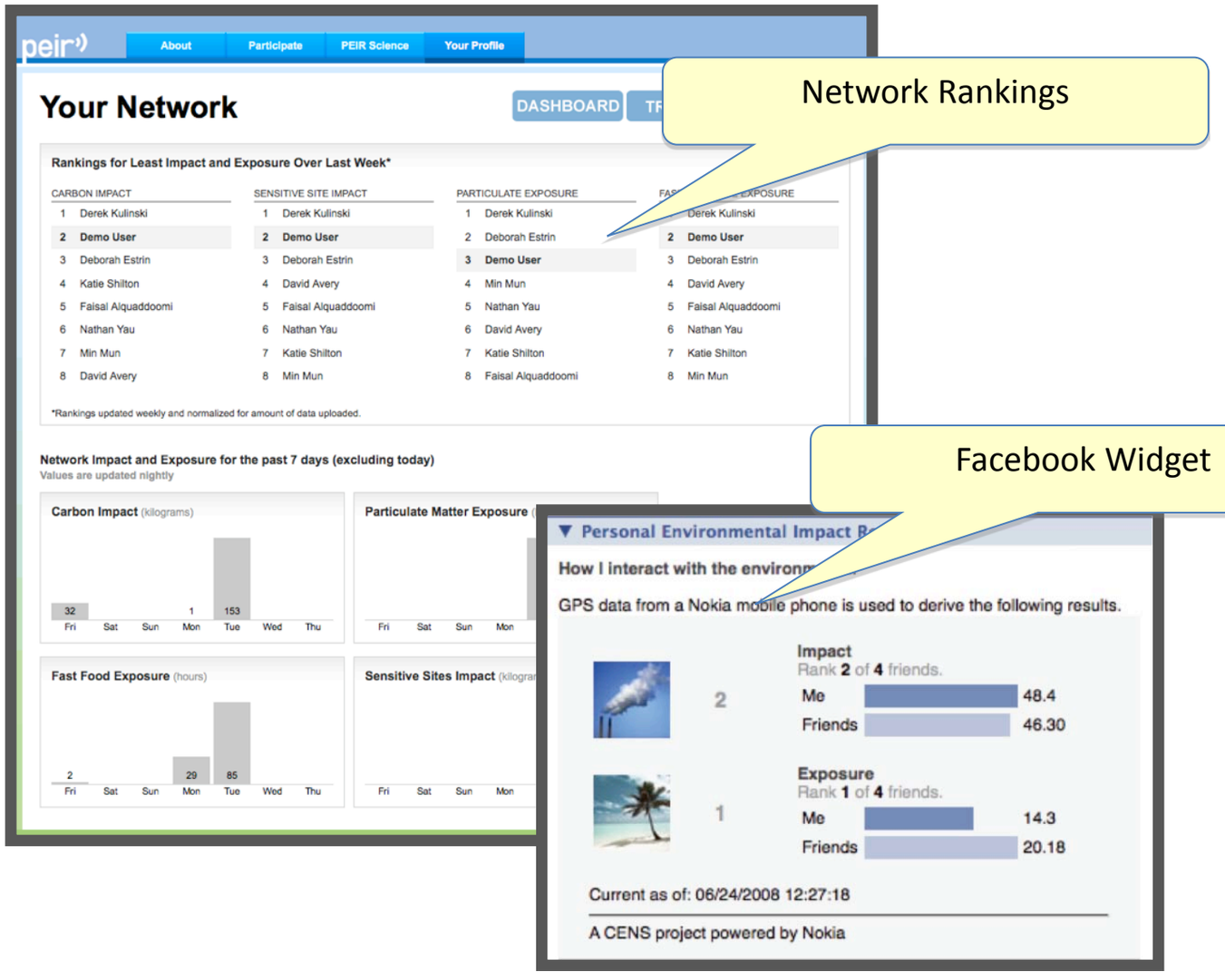
A user can compare impact and exposure to other PEIR users via network rankings and the PEIR Facebook application.

Privacy-sensitive system design.

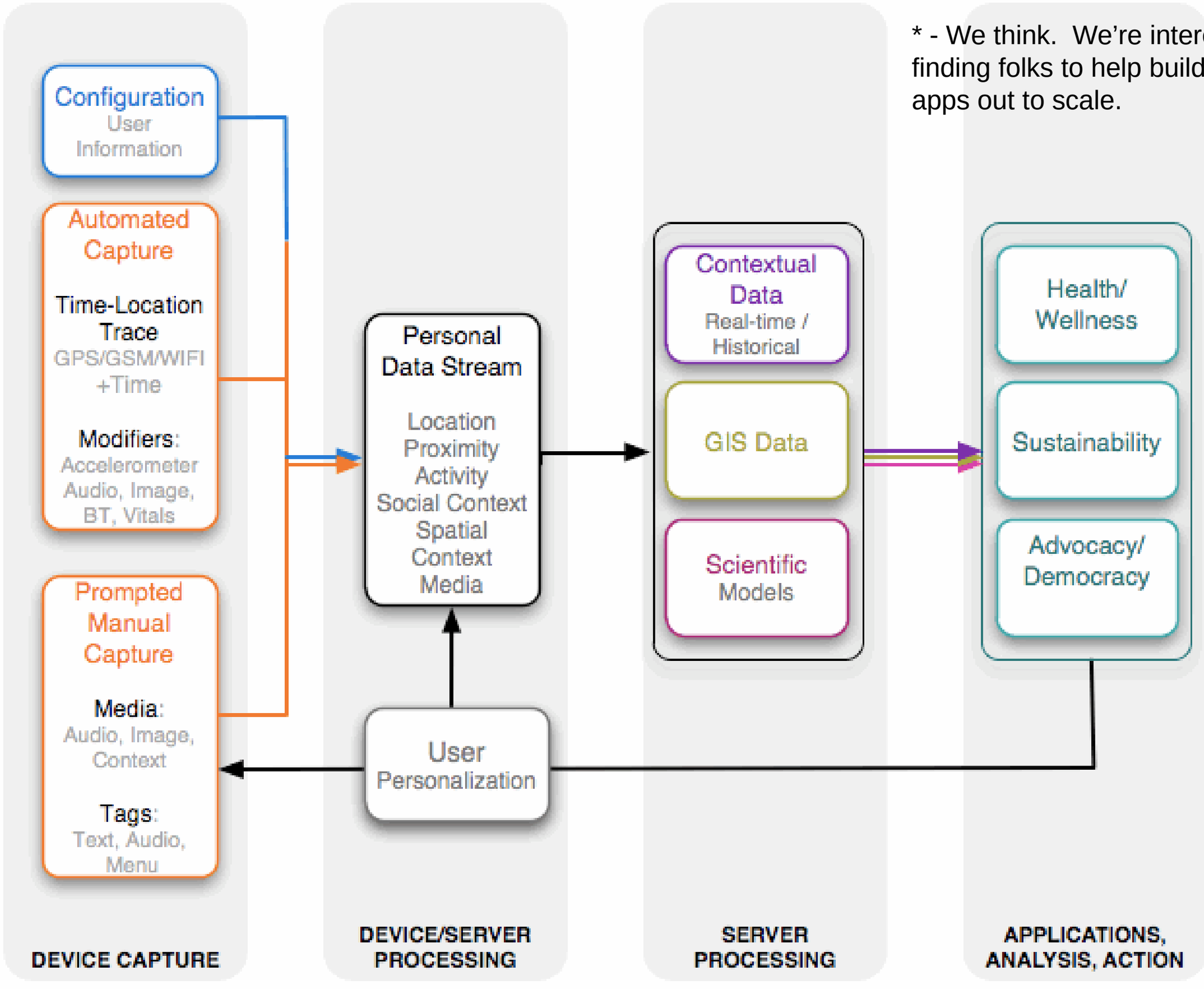
Users share only aggregates.

Individual's data is only visible to them via private account.

Local storage at the app seems like a stopgap.



Parallelizable* System Architecture



* - We think. We're interested in finding folks to help build these apps out to scale.

Privacy: New data streams are easy to mine, and personal

Knowledge@Wharton

Human Resources

From Incentives to Penalties: How Far Should Employers Go to Reduce Workplace Obesity?

Published: January 09, 2008 in Knowledge@Wharton

DOWNLOAD AUDIO PLAY AUDIO

This month, more than half of Americans probably made health-related New Year's resolutions, judging from past data, but few are likely to stick to them. Employees at CFI Westgate Resorts, an Orlando, Fla.-based vacation properties company, might consider themselves lucky: They have an incentive to get healthy. If they

Print

Get as a PDF

Send a comment

'TRACK' MAN IS SACKED

GPS NAILS ED. GUY

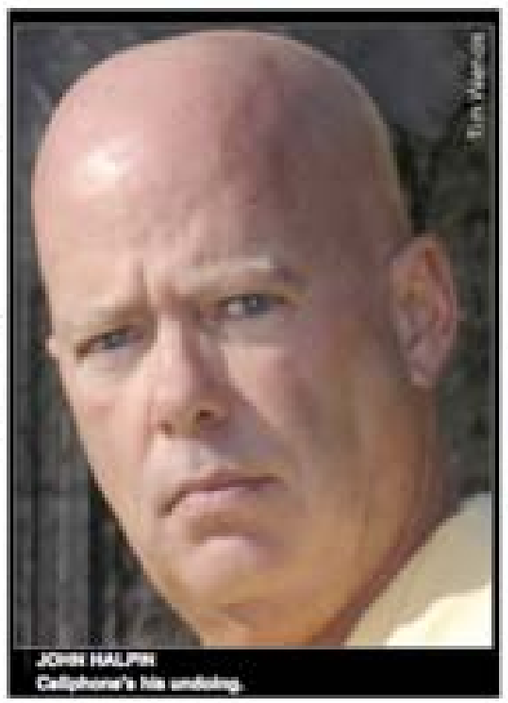
By DAVID SEIFMAN City Hall Bureau Chief

August 31, 2007

Schools Chancellor Joel Klein yesterday fired a veteran worker whose movements were tracked for five months through the GPS device in his cellphone, leading to charges that he was repeatedly cutting out early.

"This individual was getting paid for not working," said schools spokeswoman Margie Feinberg, explaining Klein's decision to accept an administrative law judge's recommendation to ax

for his 8 a.m.-to-3:30 p.m. shift.



JOHN HALPIN
Cellphone's his undoing.

Sponsored Links

The Quantified Self
Tools for knowing your own mind and body

PROTOTYPE

Predicting Where You'll Go and What You'll Like

BNET BUSINESS NETWORK: BNET | TECHREPUBLIC | ZDNET

BNET

The go-to place for management

Find Articles in free and premium articles

Today Management Strategy Work Life Insight Industries

Find Articles in: All Business Reference Technology News Lifestyle

Business Publications

Topic: RSS Feed

AIG Auto Insurance Launches GPS Based Teen Driver Pilot Program

Business Wire , April 9, 2007

EMAIL PRINT

NEW YORK -- AIG Auto Insurance today announced the AIG Teen GPS Program for auto insurance policyholders with teen drivers. According to the National Highway Safety Administration (www.nhtsa.gov), auto accidents are the leading cause of death for 16 to 20 year-olds, with roughly 6,000 young lives lost annually. The program will initially be piloted in Arizona, Illinois, New Jersey, Pennsylvania, South Carolina and Washington, and uses GPS technology to allow parents of teen drivers to monitor the location of the teen's car and driving.

After installing a small GPS unit, which is easily set up in

Children get first mobile phone at average age of eight

Eight is the average age at which children are given their first mobile phone, according to a survey.

By Stephen Adams
Published: 6:30AM GMT 18 Feb 2009

Child on mobile phone: A survey found that children as young as seven were offering to do chores in exchange for cash to buy ringtones Photo: IAN JONES

More than a third of children (35 per cent) own a mobile by the time they are that age, the charity Personal Finance Education Group (pfeg) discovered.

Its survey also found that three-quarters of all children aged seven to 15 owned "at least" one mobile.

Integrated Design of Technology & Policy

More than **privacy**. Beyond **consent**. Beyond **fair information practices**.
Towards *an ongoing process of engagement* with ***life-long*** **data repositories.**

Design Principles

Participant primacy: Users are in control. (And their choices matter.)

Data legibility: Users can understand what's being captured.

Longitudinal engagement: Users interact on an ongoing basis, and can change their minds.

Technical Innovation (Examples)

Data Vault: Secure, user-managed repository for data that’s intended to be shared with many apps

Filtering: “Easy”- On-the-fly subsampled streams; Harder- enforce principles of minimal data sharing

Auditing / Provenance: ”Contract” w/apps; infrastructure support for data traceroute

Identity: Range of identity options, from strong (tied to real life) to pseudonymous to anonymous

Network verification: Network vouches for spatiotemporal context of data injected into network

Supporting Social Structures

Legal privilege: Separation of concerns for data stored in the vault

Transparency of Services: “Fair Trade” data labeling for voluntary enforcement and monitoring

K. Shilton, J. Burke, D. Estrin, M. Hansen, J. Kang.

Personal Data Vault

Objectives

User-managed secure repository for time-location traces controlled by the person whose traces they are!

Granular, configurable sharing and archival

Audit trail enforcement and API

Controlled access by user-authorized third-party services

Migration from provider to provider over lifetime.

Technical Challenges

Support efficient federation of queries across PDVs

Defining “auditing” and its enforcement

Continual collection for availability monitoring, recruitment

Virtualization to isolate vaults and app processing

Privilege model – medical / legal

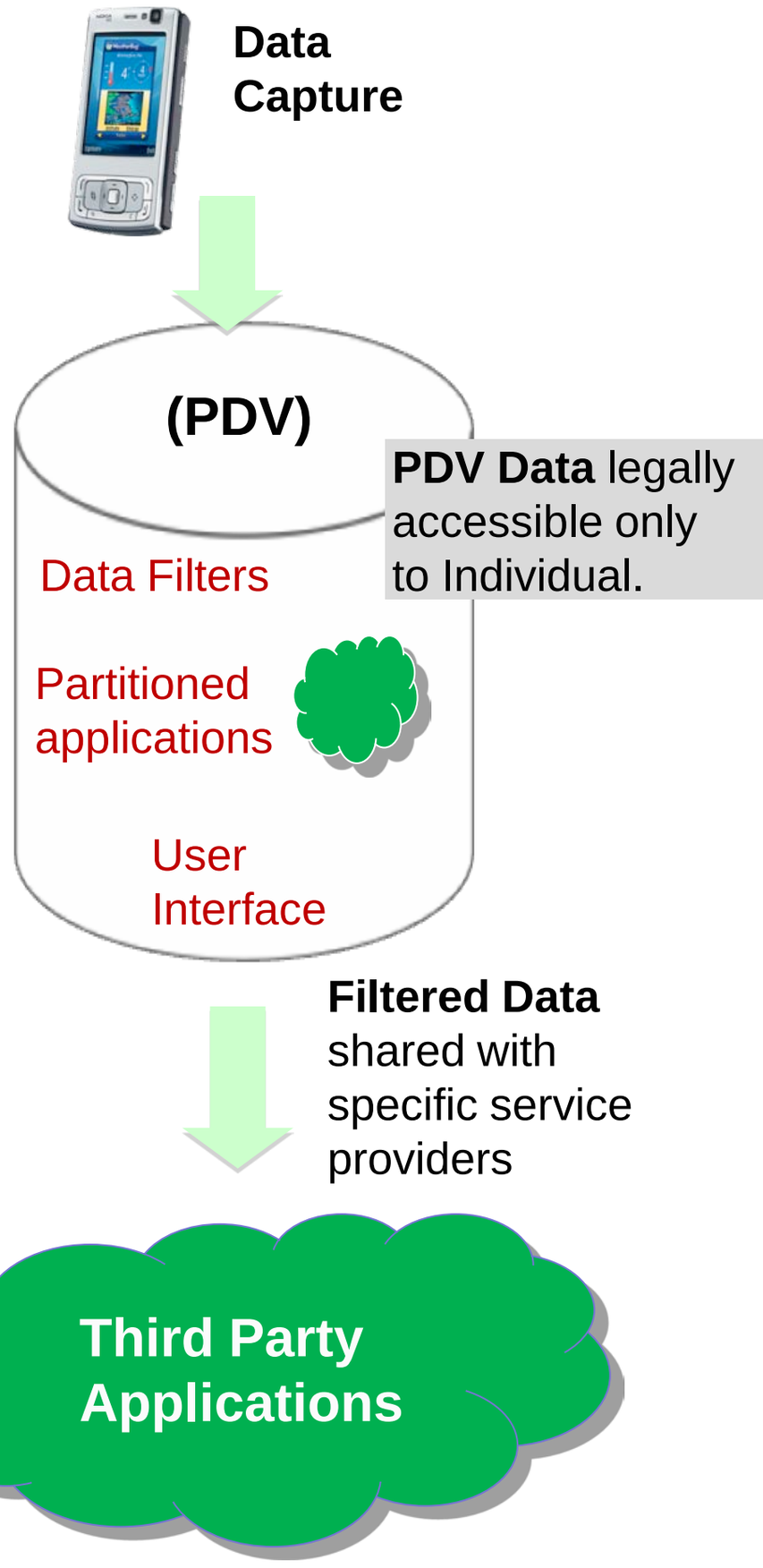
Reduces conflicts of interest

Limits circumstances for legal discovery

Enable a marketplace of ‘certified’ applications

Standards, best practices, and auditing required.

Precedents (sort of)**:** Yahoo Fire Eagle, Google Latitude



K. Shilton, M. Mun, J. Burke, D. Estrin, J. Kang, M. Hansen, et al.

Legibility of High Granularity, Life-Long Data

<u>Sampling</u>

- PEIR: Time-location (plus cell tower, battery, etc.) every 30 seconds;
- “Ambulation”: Time-location plus accelerometer data at higher rate;
- “Remapping LA” & Medical EMA: Time-location plus occasional images.
- Recruitment / availability monitoring requires ongoing sampling + analysis.

For life?

Visualization and management of data and transaction info are complex.

Personal data streams are the simple case.

- **Mixed ownership and access**. Personal, group, and external and/or sensitive and public data?
- **Recruitment / availability monitoring.** Approaches to managing this with minimum risk?
- Current search and visualization tools don’t do what we need yet -- though *Swivel*, *ManyEyes*, and others have very relevant related work.

=> Are these “solely” user interface problems?

Longitudinal Engagement with Life-Long Data Streams

Technical, legal and social implications of a life-long relationship between users and their sensed data.

Notwithstanding the engineering concerns of scaling, and deeper question of *how long should we remember* and the *social value of forgetting*, we have challenges:

Enable **informed redistribution decisions**, not just protect what's stored;

Embed **provenance** and offer sharing constraints based on it;

Provide **understandable representations of data and transaction history**;

Support **end-user participation the entire data life cycle**.

Challenges for this community

Provenance, "privacy", and auditing are needed at the data transaction level, especially for "mash-up" style applications that push data across services within different legal entities and system TOS.

They add overhead and require the creation of representations / management tools that are *accurate, granular,* and *at a variety of levels of abstraction*.

Recap of Provocations

3B+ Mobile Phone Users

All capturing, analyzing and selectively sharing data they care about…

Over their lifetime.

A few needs / questions

Scaling up these applications and corresponding engineering challenges.

Managing **Recruitment & Reputation** (Is all data equal?)

From Continuous Nearest Neighbor => **Availability Monitoring**

“Data vaults” (and federating queries across them), auditing, etc.

Integrated Design: Policy & Technology

Could policy enable **usable data provenance and auditing** for this data?

What about **varied concepts of identity** (and thus rights/ownership)?

And **network verification** of spatiotemporal context…

Beyond Privacy, Consent, and FIP: Legibility & Longitudinal Engagement.

Acknowledgements: Collaborators and Sponsors

Collaborators

Faculty/PIs: Jeff Burke, Deborah Estrin, Mark Hansen, Ramesh Govindan, Eric Graham, Jerry Kang, Nithya Ramanathan, Mary Jane Rotheram-Borus, Mani Srivastava, Michael Swiernik, Fabian Wagmister

Students/staff: Peter Capone-Newton, Patrick Crutcher, Betta Dawson, Joey Degges, Hossein Falaki, Brent Flagstaff, Cameron Ketcham, Donnie Kim, Keith Mayoral, Min Mun, Nicolai Munk Petersen, Nathan Yau, Sasank Reddy, Ruth West, Vids Samanta, Katie Shilton, Masanao Yajima, Eric Yuen

Sponsors and Partners

UCLA centers: CENS, REMAP, Global center for families and children

Federal funding: NSF: NETS-FIND Program, OIA

Corporate funding: Cisco, Google, Intel, MSR, Nokia, Sun, T-Mobile

Foundations/NGOs: Woodrow Wilson Center, Conservation International, Saint Johns Well Family and Childrens' Clinics, Project Surya

CENS Center for Embedded Networked Sensing

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