



# Privacy-preserving Ubiquitous Social Mining via Modular and Compositional Virtual Sensors

Evangelos Pournaras, Iza Moise, Dirk Helbing

# Motivation



THE MAGAZINE BLOGS VIDEO BOOKS CASES WEB

Guest

Subscribe today and get access to all current articles and

HBR Blog Network



## Big Data's Dangerous New Era of Discrimination

by Michael Schrage | 8:00 AM January 29, 2014

Comments (30)

Congratulations. You bought into Big Data and it's paying off Big Time. You slice, dice, parse and process every screen-stroke, clickstream, Like, tweet and touch point that matters to your enterprise. You now know exactly who your best — and worst — customers, clients, employees and partners are. Knowledge is power. But what kind of power does all that knowledge buy?

Big Data creates Big Dilemmas. Greater knowledge of customers creates new potential and power to discriminate. Big Data — and its associated analytics — dramatically increase both the dimensionality and degrees of freedom for **detailed discrimination**. So where, in your corporate culture and strategy, does value-added personalization and segmentation end and harmful discrimination begin?



## Existing social mining practices threaten social cohesion



*“surveillance has become increasingly privatized, commercialized and participatory”, Julie E. Cohen*

## First Degree Price Discrimination Using Big Data

Benjamin Reed Shiller, Economics Department, Brandeis University

**ZDNet**

White Papers Hot Topics Downloads Reviews Newsletters

US Edition Internet of Things Mobility Research Windows Enterprise Software

ARE YOUR HR & FINANCE SYSTEMS BASED ON TECH OLDER THAN YOU?

MUST READ *Hackers jump on the Shellshock Bash bandwagon*

Topic: Big Data Follow via: RSS Email

### Why big data evangelists should be sent to re-education camps

**Summary:** Big data is a dangerous, faith-based ideology. It's fuelled by hubris, it's ignorant of history, and it's trashing decades of progress in social justice.

By Stiglherrian for The Full Tilt | September 19, 2014 -- 07:13 GMT (00:13 PDT)

Follow @Stiglherrian Get the ZDNet Big Data newsletter now

Comments 27 Votes 7 Share 459 Tweet 692 LinkedIn Share more +

The last time I wrote about big data, in July, I called it a **big, distracting bubble**. But it's worse than that. Big data is an ideology. A religion. One of its most important gospels is, of course, at Wired.

In 2008, Chris Anderson talked up a thing called The Petabyte Age in *The End of Theory: The Data Deluge Makes the Scientific Method Obsolete*.

"The new availability of huge amounts of data, along with the statistical tools to crunch these numbers, offers a whole new way of understanding the world. Correlation supersedes causation, and science can advance even without coherent models, unified theories, or really any mechanistic explanation at all," he wrote.

Declaring the **scientific method** dead after 2,700 years is quite a claim. Hubris, even. But, Anderson wrote, "There's no reason to cling to our old ways." Oh, OK then.

Now, this isn't the first set of claims that correlation would supersede causation, and that the next iteration of computing practices would "make everything different".

Has anyone got a pin?

# Research Question



How to design an *open, decentralized, privacy-preserving & participatory* system to provide ubiquitous social **mining services** engineered as **public good**



**Social mining:** the process of discovering information from data sensed in one or more social environments so that a social phenomenon is understood or a societal problem is tackled.

# Approach

Introduction of a **modular & compositional** approach so that citizens **participate** during building lifecycle & later on



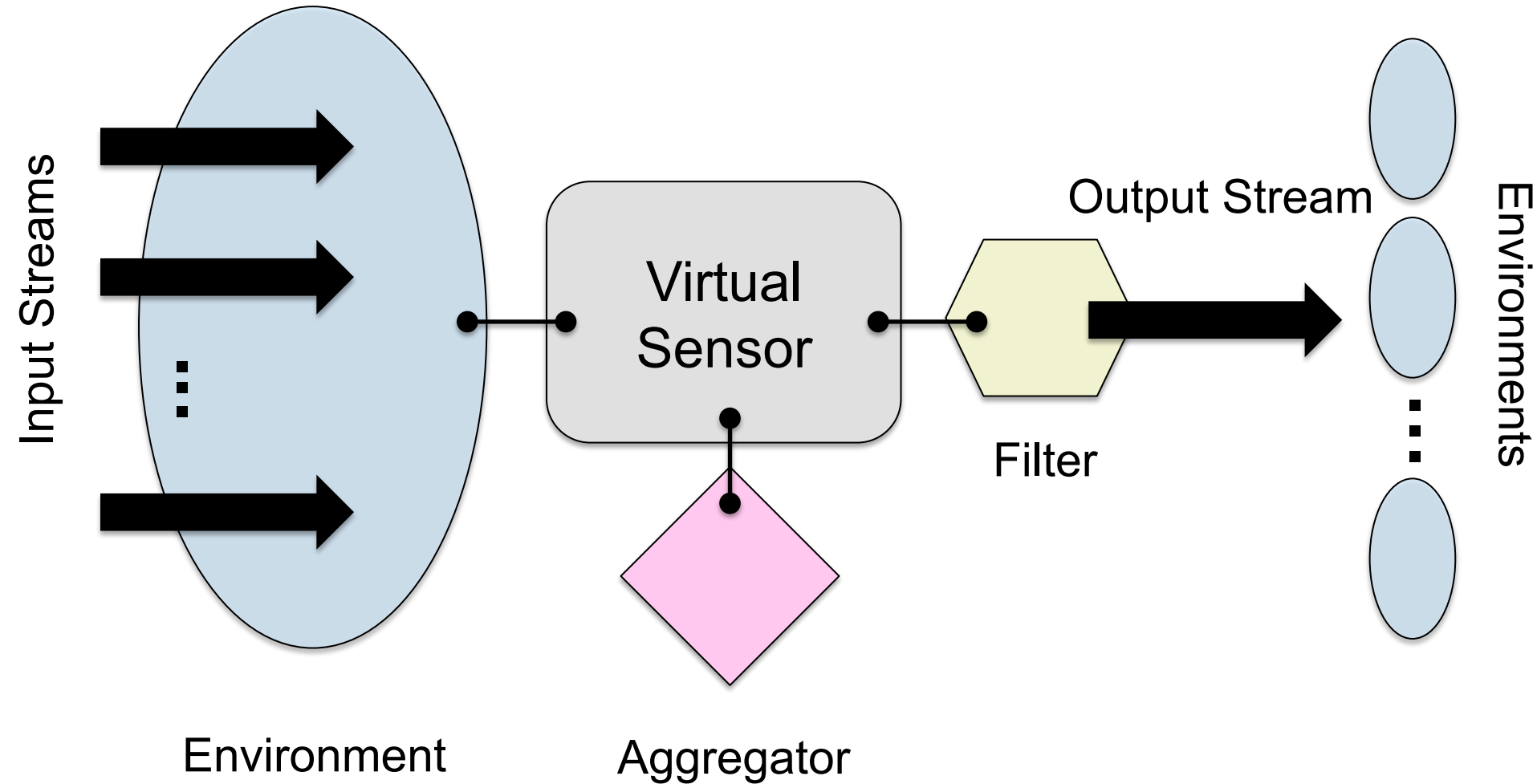
Why?

Designing **extensible & reusable** social mining processes via compositional data flow of sensors **simplifies application development**



A design principle with a potential to **simplify crowd-sourcing activities & increase engagement** of building communities.

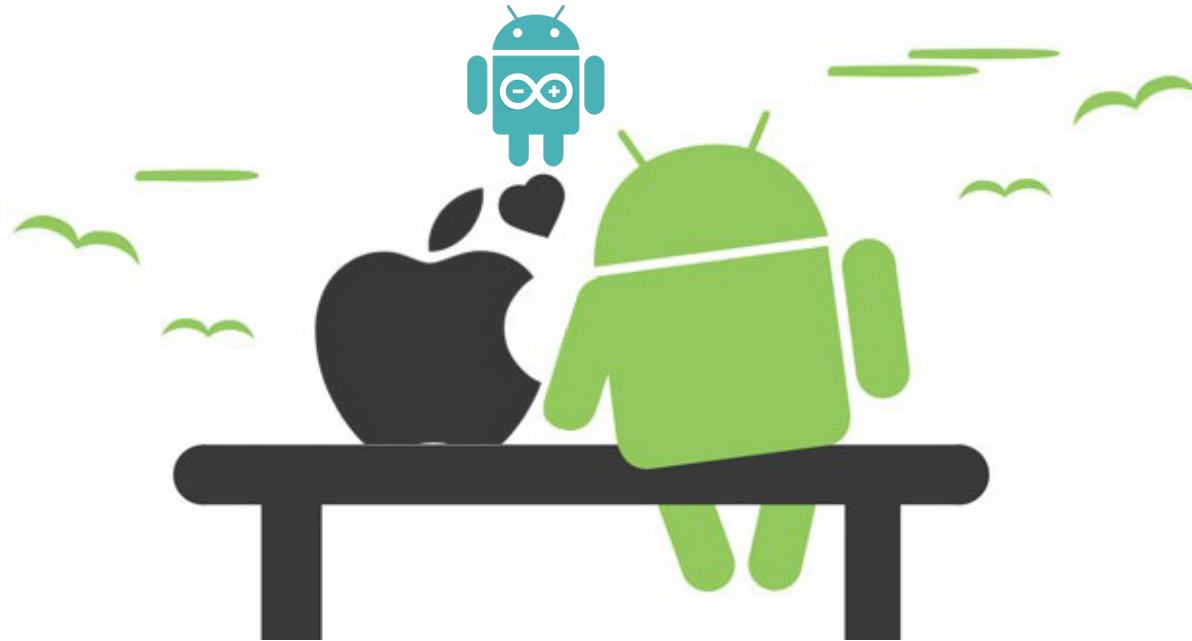
# The Virtual Sensor Model





***nervousnet***

# Portability – Internet of Things



We started in Android devices but...

... we move to iOS...

...and later on other embedded platforms, e.g. Arduino

# Planetary Nervous System

Accelerometer

Light

Temperature

Humidity

Gyroscope

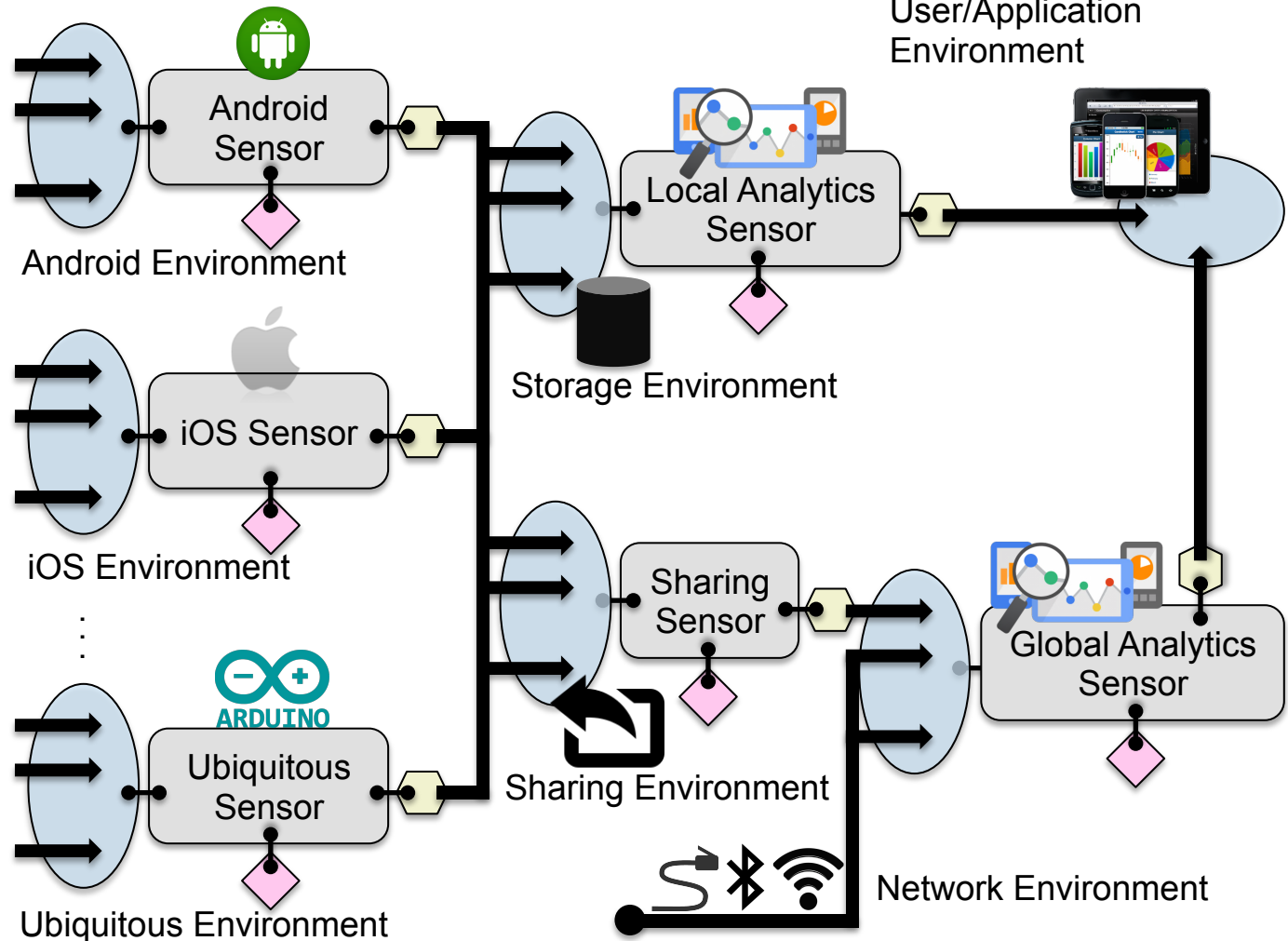
Proximity

Battery

Atm. Pressure

Magnetic

Noise



# Self-determination of Privacy



**nervousnet**

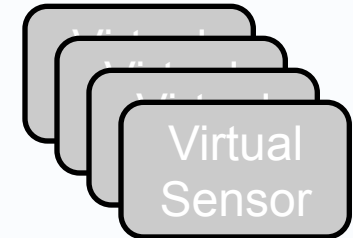
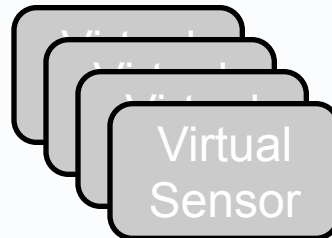
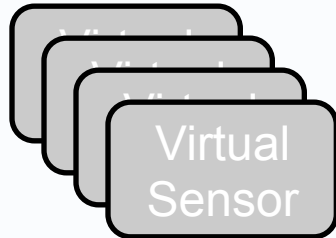
Accelerometer	<input checked="" type="checkbox"/> Log	<input checked="" type="checkbox"/> Share
Battery	<input checked="" type="checkbox"/> Log	<input checked="" type="checkbox"/> Share
BLEBeacon	<input checked="" type="checkbox"/> Log	<input checked="" type="checkbox"/> Share
Connectivity	<input checked="" type="checkbox"/> Log	<input checked="" type="checkbox"/> Share
Gyroscope	<input checked="" type="checkbox"/> Log	<input checked="" type="checkbox"/> Share
Humidity	<input checked="" type="checkbox"/> Log	<input checked="" type="checkbox"/> Share
Light	<input checked="" type="checkbox"/> Log	<input checked="" type="checkbox"/> Share
Magnetic	<input checked="" type="checkbox"/> Log	<input checked="" type="checkbox"/> Share
Noise	<input checked="" type="checkbox"/> Log	<input checked="" type="checkbox"/> Share
Pressure	<input checked="" type="checkbox"/> Log	<input checked="" type="checkbox"/> Share
Proximity	<input checked="" type="checkbox"/> Log	<input checked="" type="checkbox"/> Share
Temperature	<input checked="" type="checkbox"/> Log	<input checked="" type="checkbox"/> Share

Two privacy levels!

1. Local storage – Filter of Android/iOS Sensor

2. Sharing – Sharing Sensor

# Local & Global Analytics



Application Ecosystem



Planetary Nervous Middleware System

API Services

# Local & Global Analytics

**Local analytics:** Building new virtual sensors, e.g. noise sensor

A data-driven API for social mining

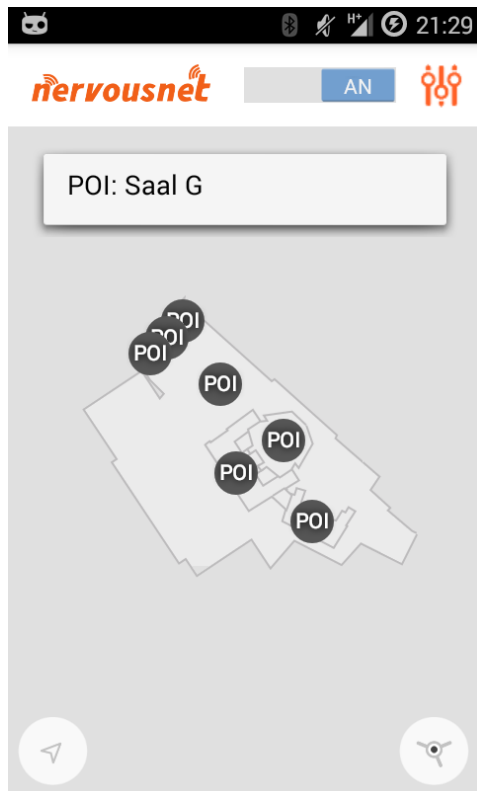
**Global analytics:** Building decentralized aggregation services

DIAS – Dynamic Intelligent Aggregation Service

Evangelos Pournaras, Martijn Warnier and Frances M.T. Brazier, *A Generic and Adaptive Aggregation Service for Large-scale Decentralized Networks*, Complex Adaptive Systems Modeling, 1:19, 2013 © SpringerOpen

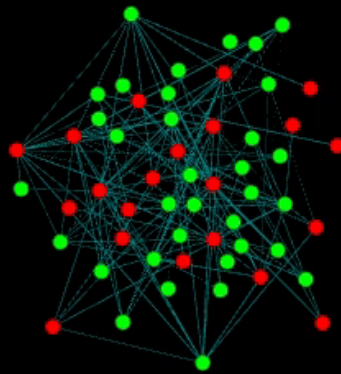
# Applications

## Real-time noise monitoring in Smart Cities



## Privacy-preserving social networking and navigation

# Demonstration



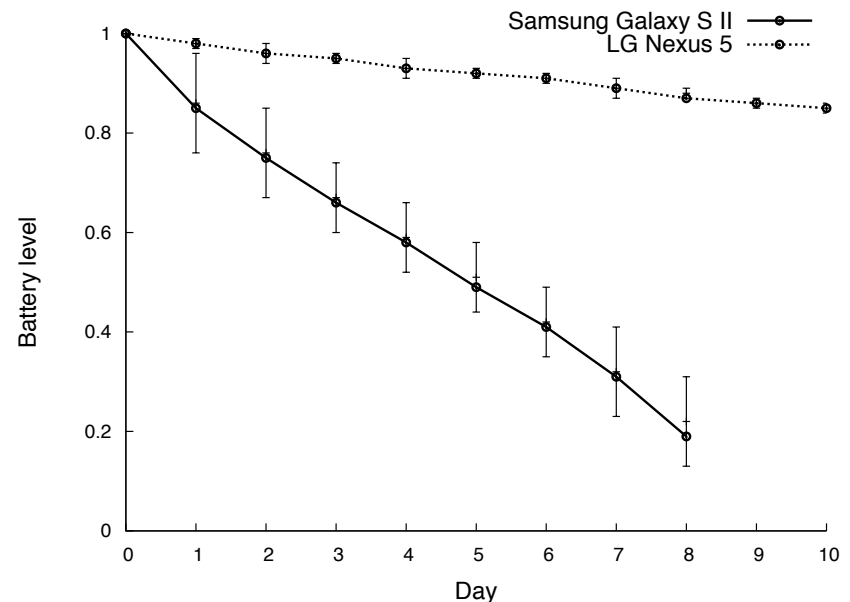
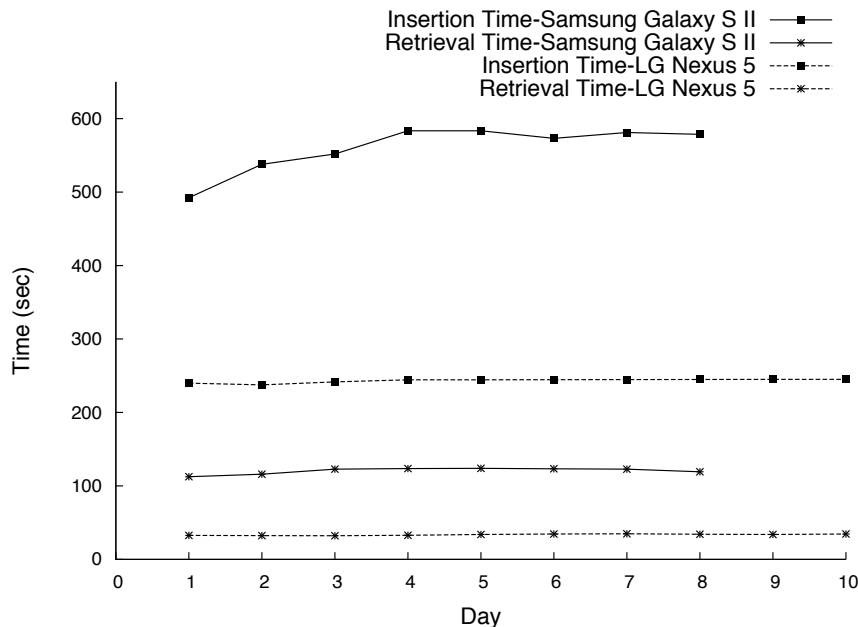
# Computational Results

How to access the cost-effectiveness of the nervous system?

Emulation!

Data generation: 2 phone users, 20 virtual sensors, log every 5 sec, 10 days

What are the storage, retrieval and battery consumption costs?



# Conclusions

Incremental development of social mining application

**Virtual sensors:** a promising design approach for building ubiquitous social mining services that are **by design** *open, decentralized, privacy-preserving & participatory*

**Modular and compositional approach:** stimulates engagement & innovation in crowd-sourcing activities

Performance evaluation confirms the feasibility of the introduced model

# Join our Collective Vision!



## Questions?

epournaras@ethz.ch

<https://github.com/mosgap/nervous>