Big data, industry and disease surveillance

Alessio Signorini

alessio-signorini@uiowa.edu
@a_signorini

How to Make Millions with Big Data

Alessio Signorini
@a_signorini

My Data is Bigger than Yours

Alessio Signorini

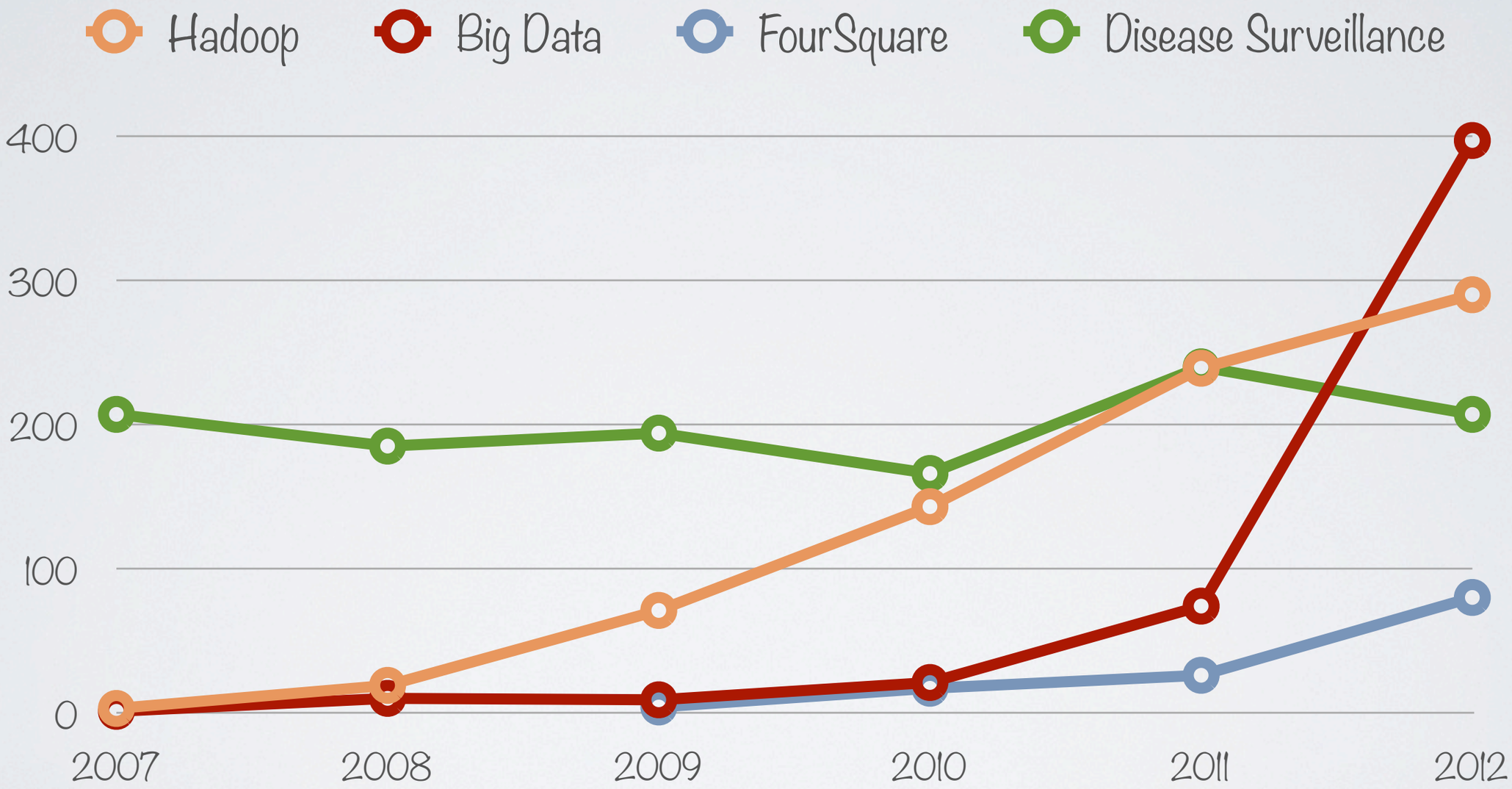
@a_signorini

Big data, industry and disease surveillance

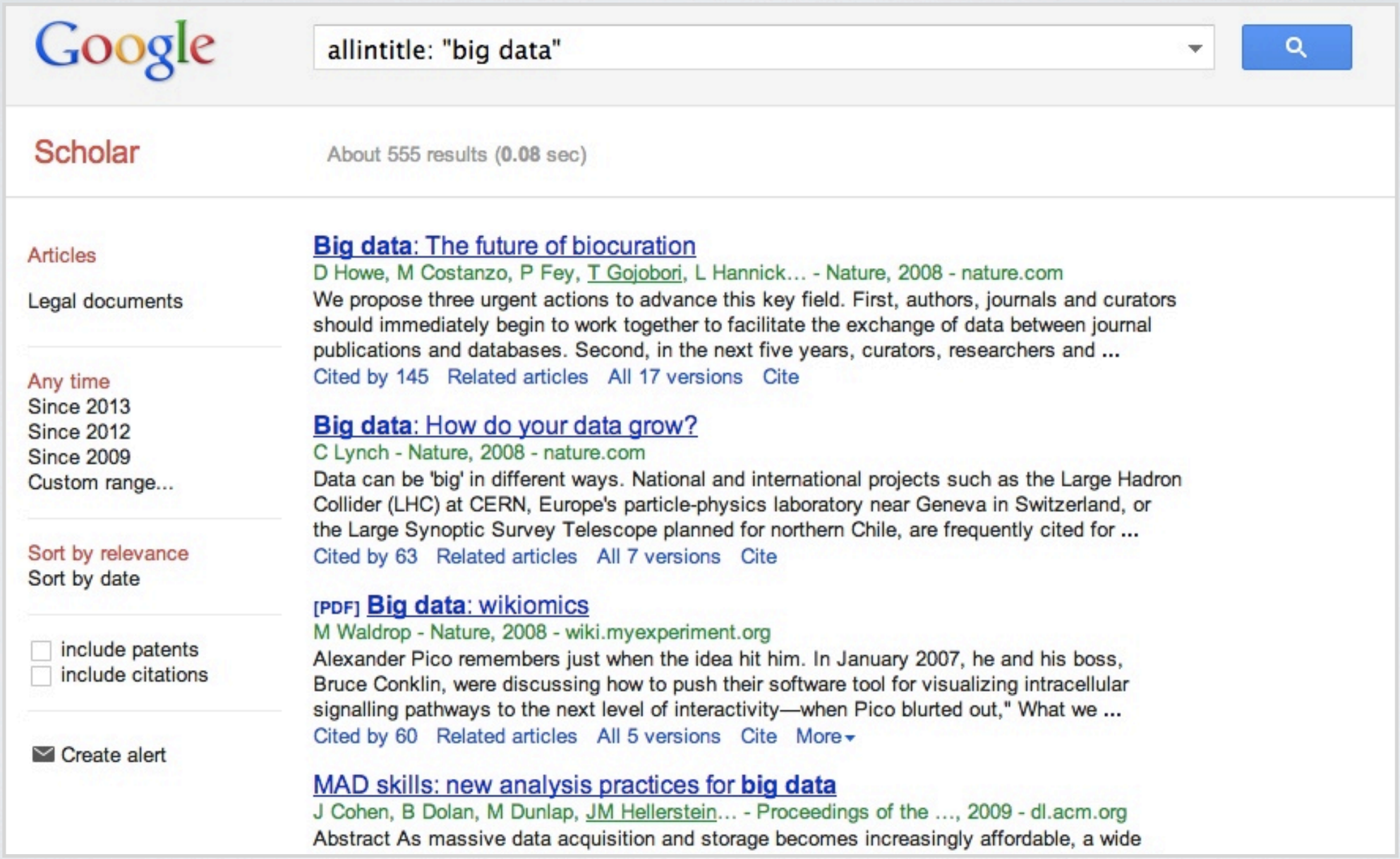
Alessio Signorini

alessio-signorini@uiowa.edu
@a_signorini

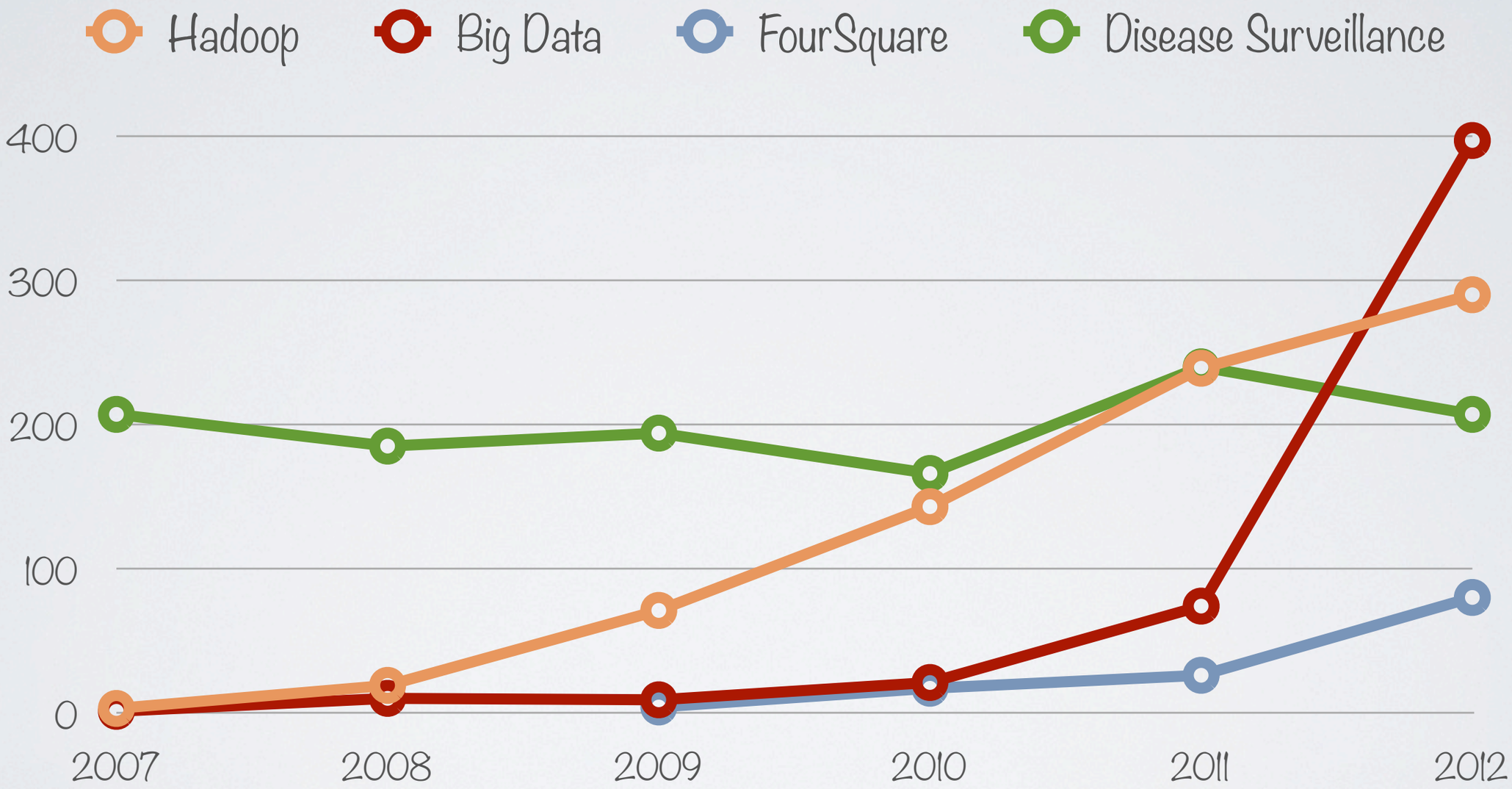
Number of Articles with Keyword in Title



Google Scholar + python



Number of Articles with Keyword in Title



17,000+

"Big Data" companies

SAILTHRU

$28M

Gigamon

$23M

Adyoulike

$2M

CIVITAS LEARNING

$4M

$80 Million

raised in last 30 days

Tykli

Let data speak

$70K

stremor

comprehending language

$2M

semetric

$5M

THINKBIG ANALYTICS

$3M

ffine

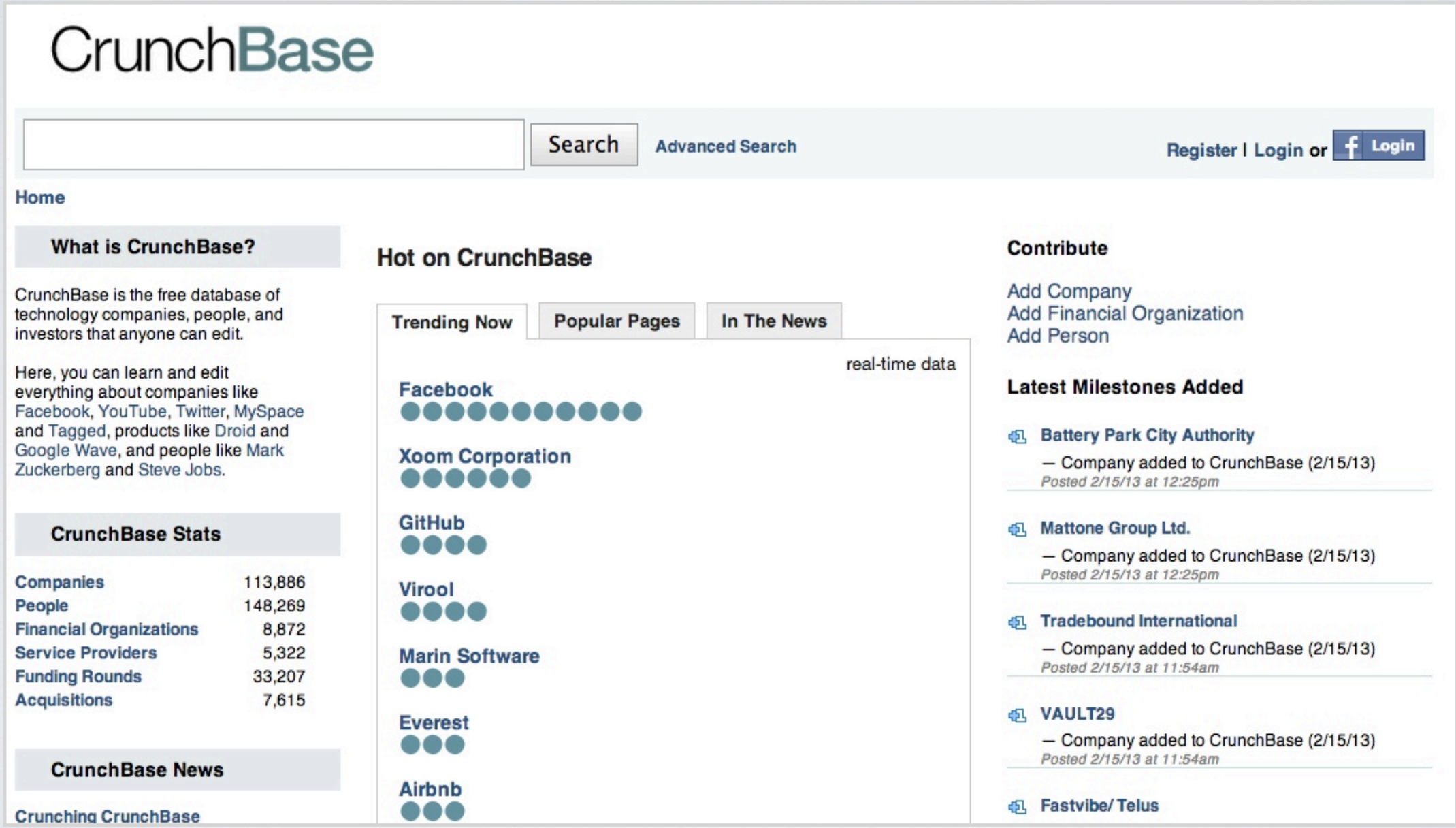
favorable transformations

$100K

AYASDI

$10M

CrunchBase + Perl 5



SAILTHRU

$28M

Gigamon

$23M

Adyoulike

$2M

CIVITAS LEARNING

$4M

$80 Million

raised in last 30 days

Tykli

Let data speak

$70K

stremor

comprehending language

$2M

semetric

$5M

THINKBIG ANALYTICS

$3M

ffine

favorable transformations

$100K

AYASDI

$10M

NORTH CAROLINA
ARIZONA
UTAH
CALIFORNIA
KENTUCKY

THE UNIVERSITY OF IOWA

PhD Candidate

Ask.com

Web Search Engine

OneRiot
Search for the pulse.

Social Search Engine

immersive

Digital Advertising

AchieveMint

Health Reward Platform

Data is Everywhere

I am only at Slide #10

and lots of you

are checking their phone

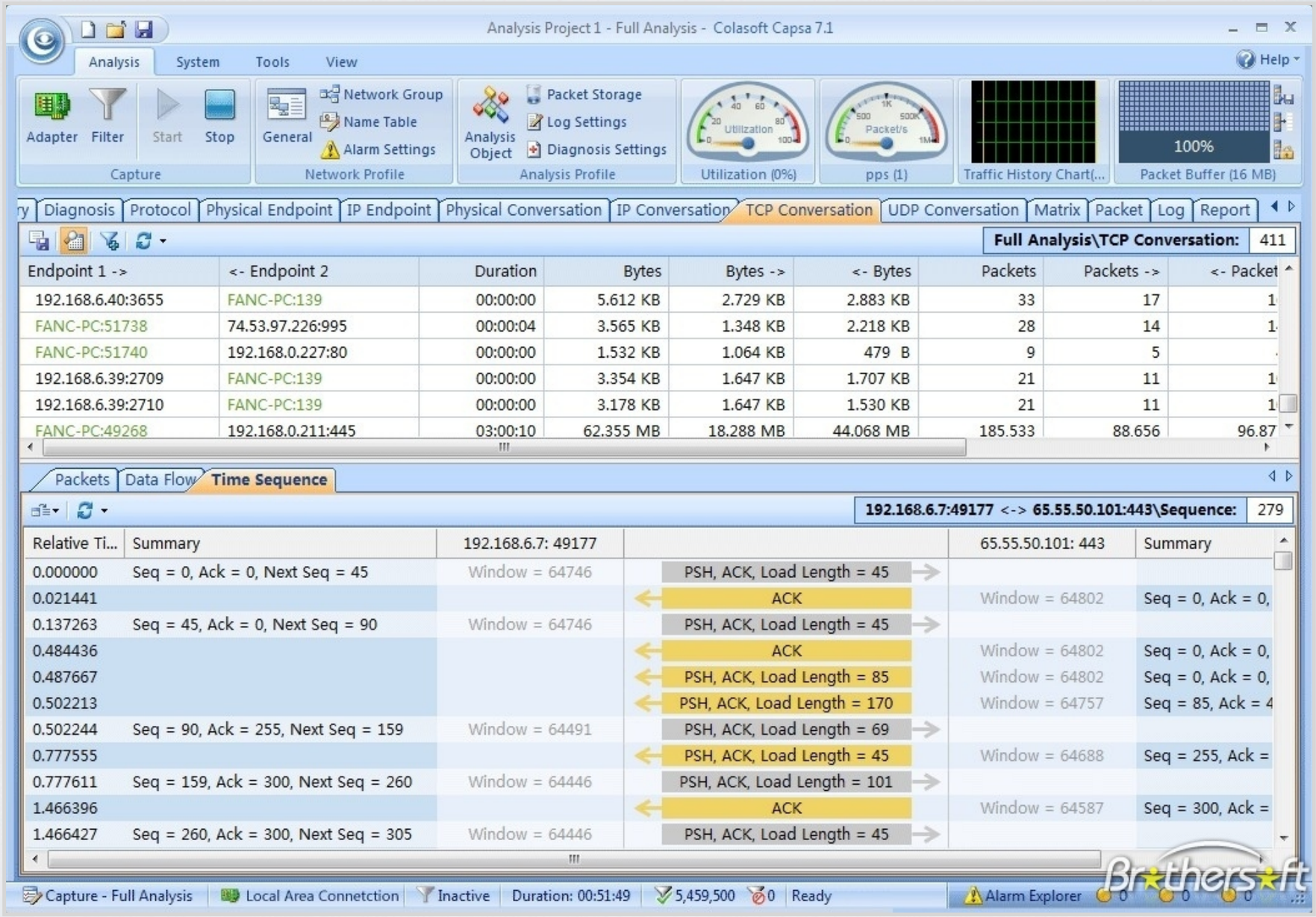
Bill Gates
@BillGates
Follow
@a_signorini talk at #IMED2013 is awesome! If you aren't at the conference today, you are really missing out...
bit.ly/WpSyyh
Reply
Retweet
Favorite
More
416
RETWEETS
269
FAVORITES
11:10 AM - 18 Feb 2013

Twitter

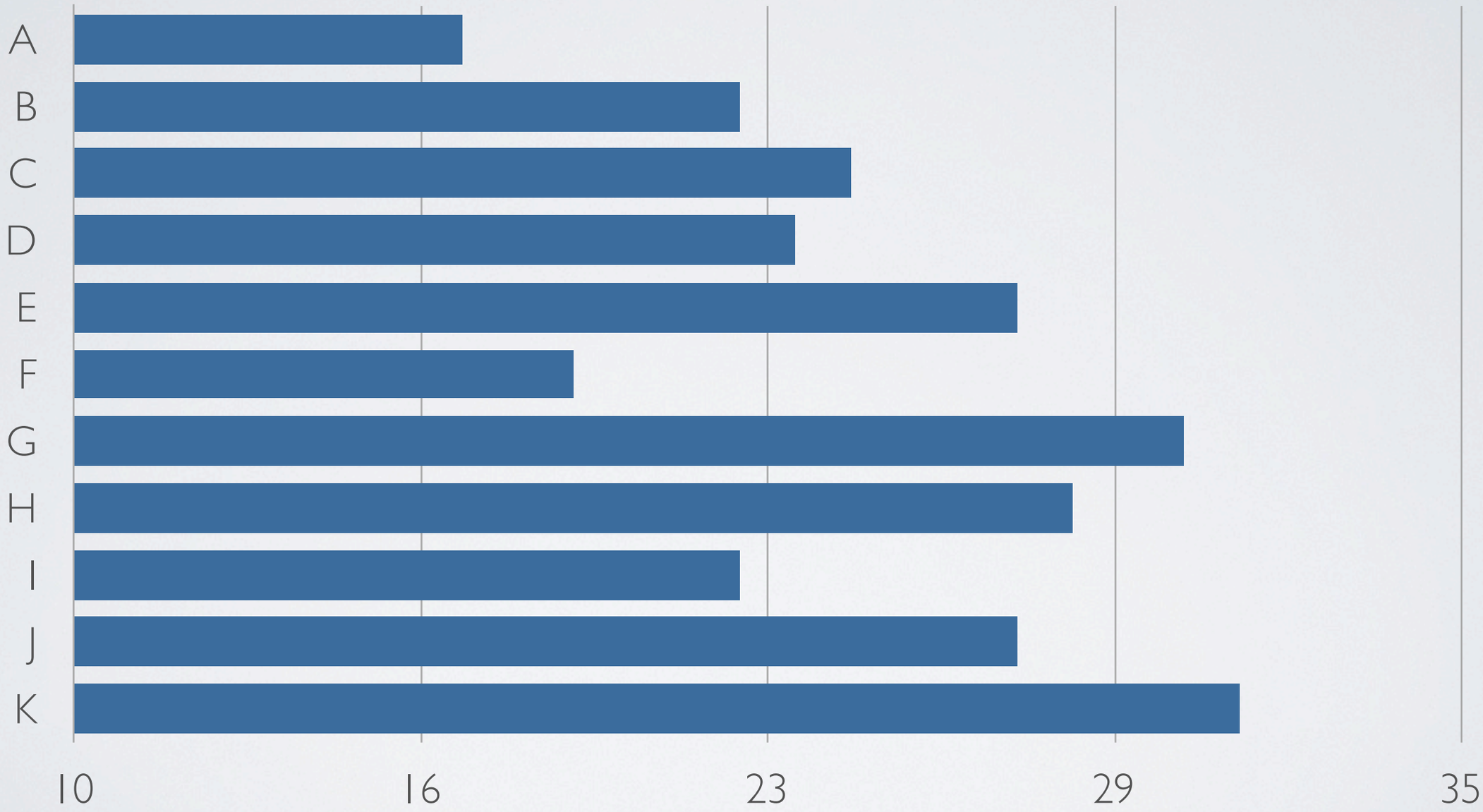
Twitter

Facebook/Email

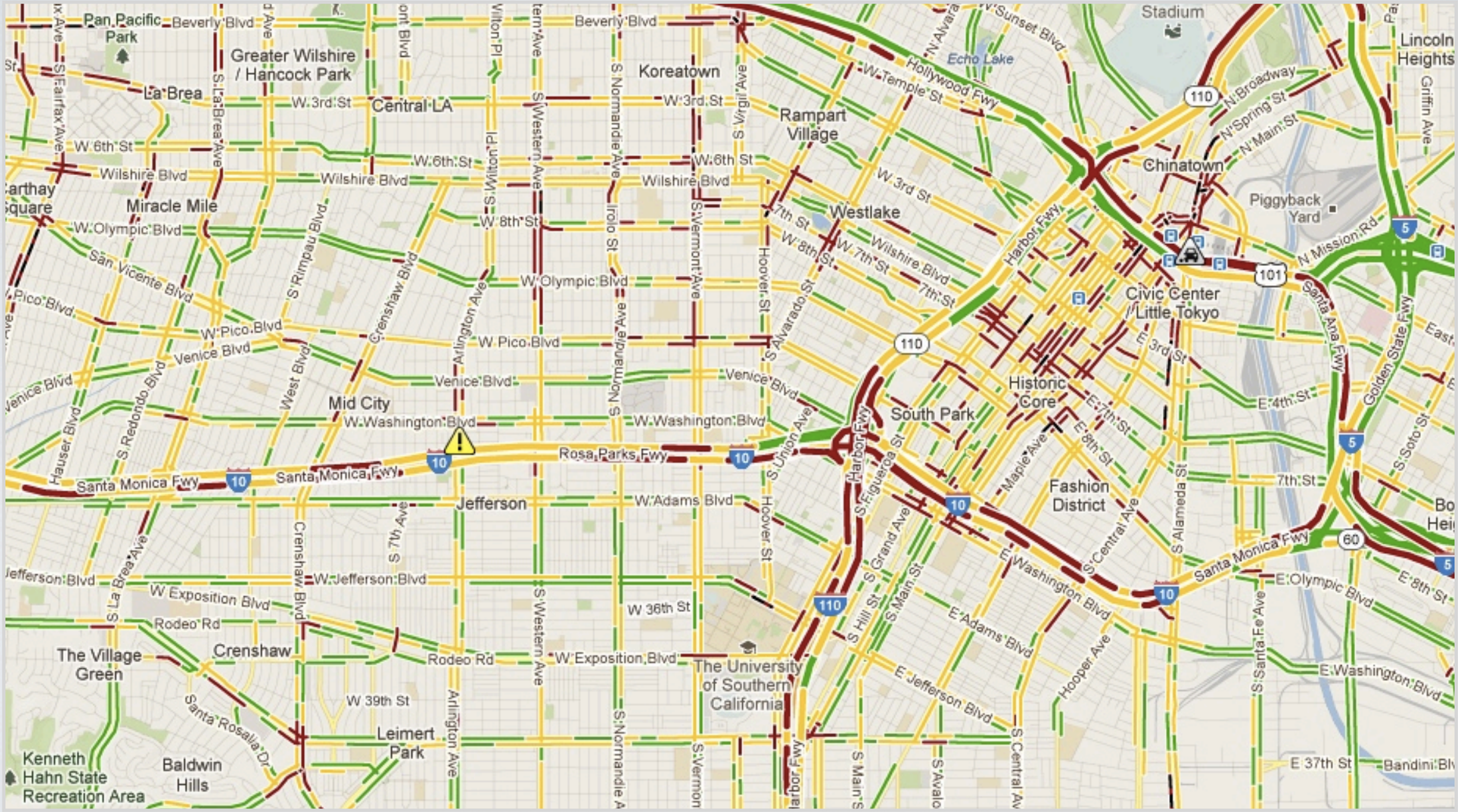
“Promiscuous” Mode



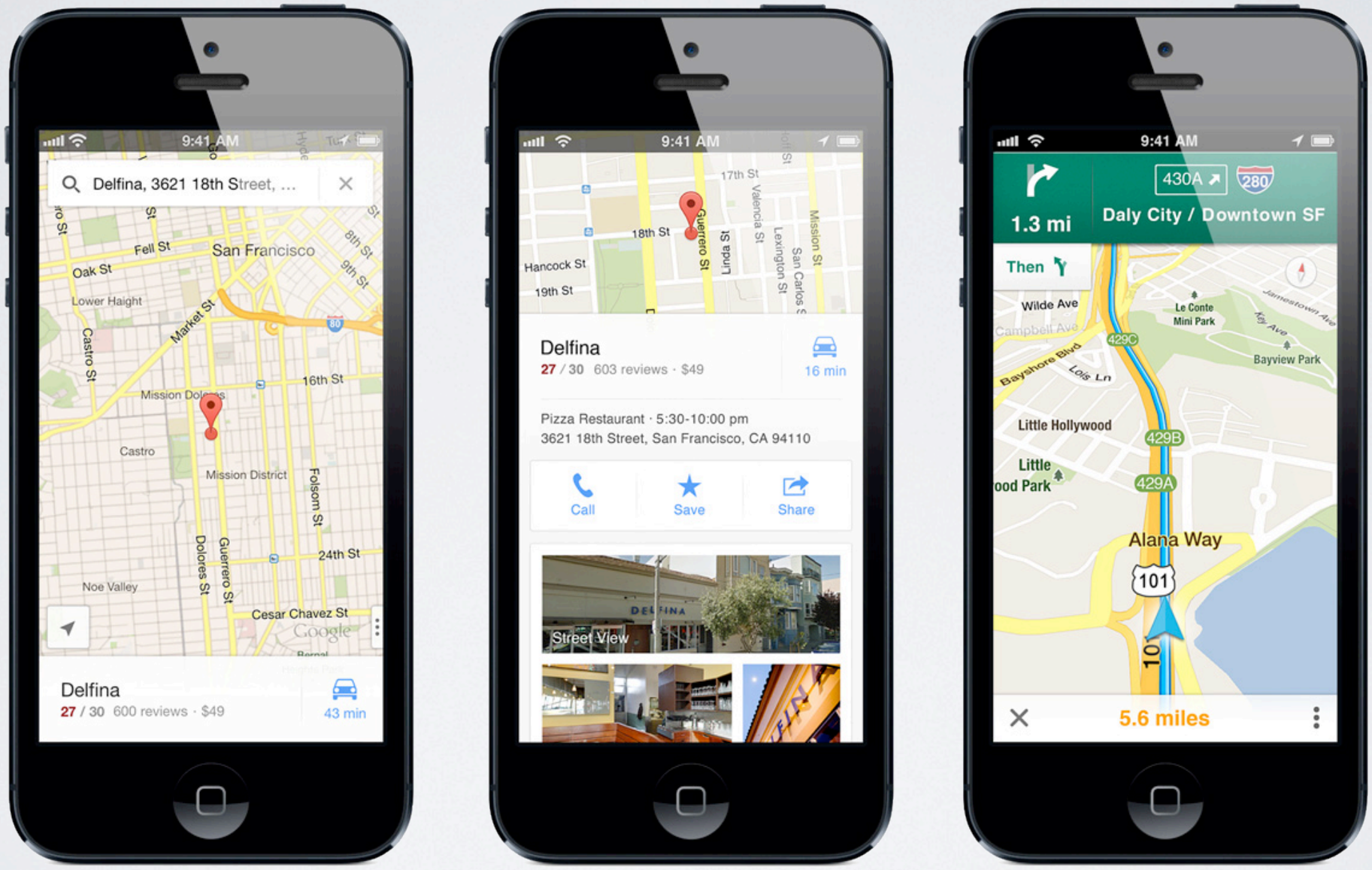
Number of Users on Wifi during Session 12



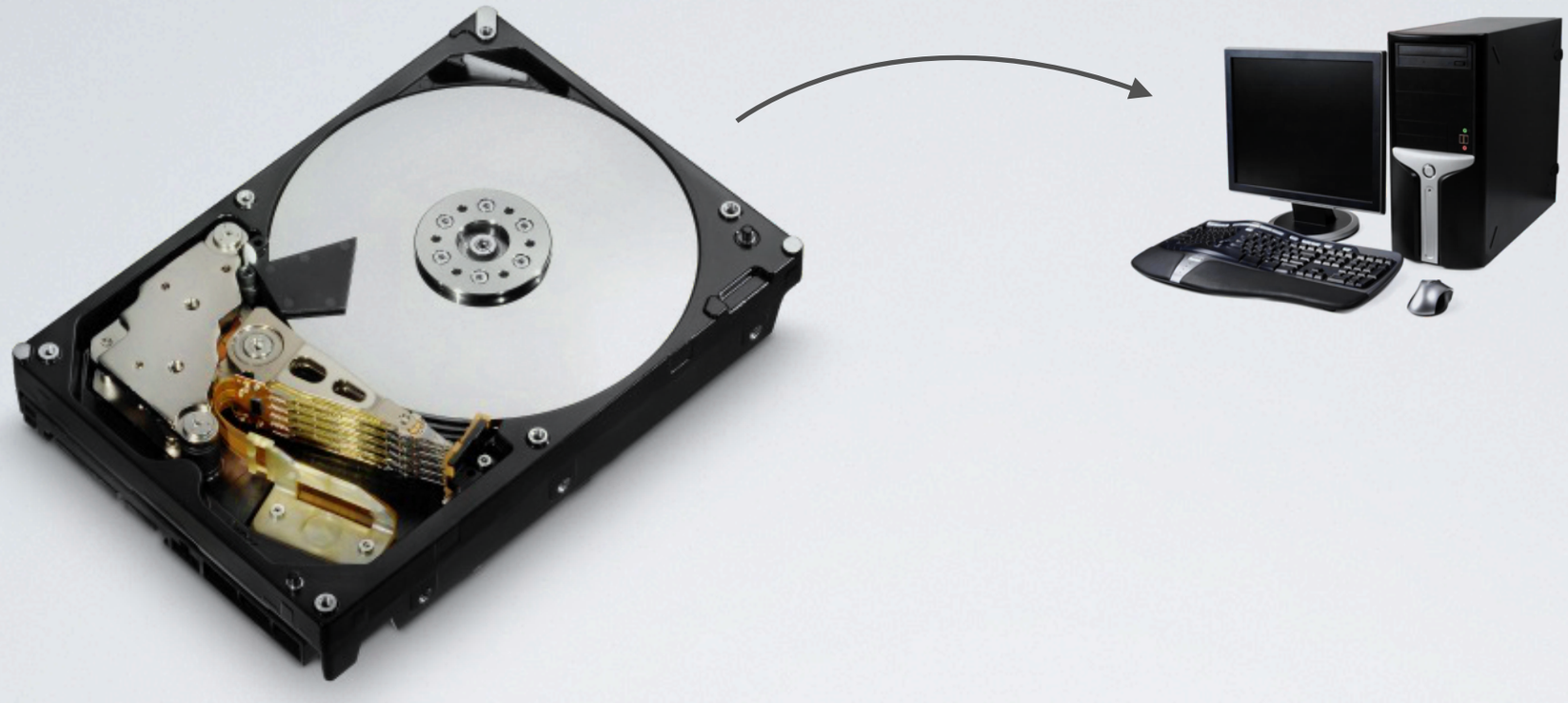
Google maps - Los Angeles Traffic



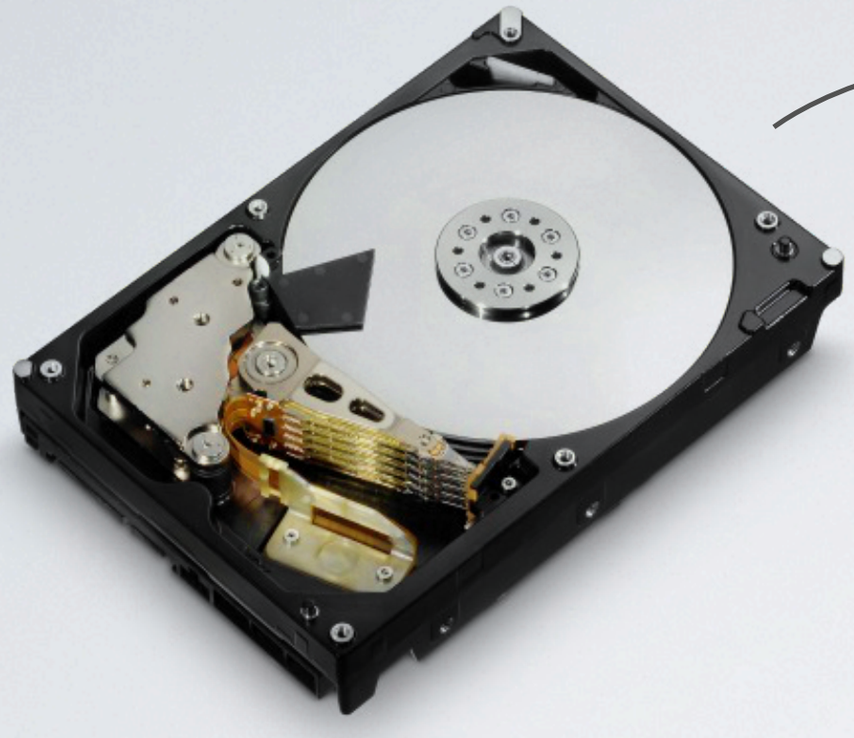
Google Mobile Apps sends back GPS data



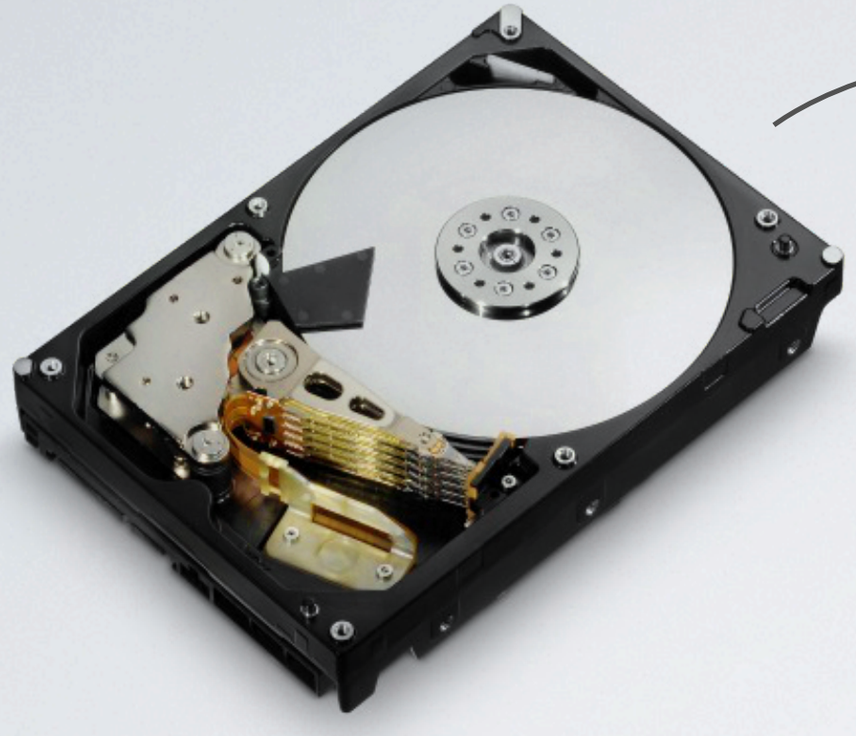
Hard Disk



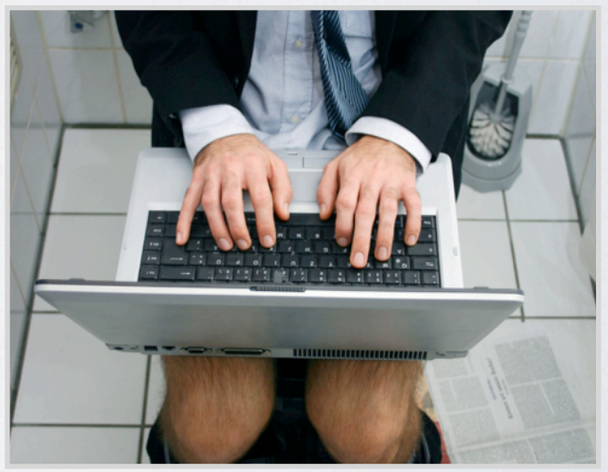
Hard Disk



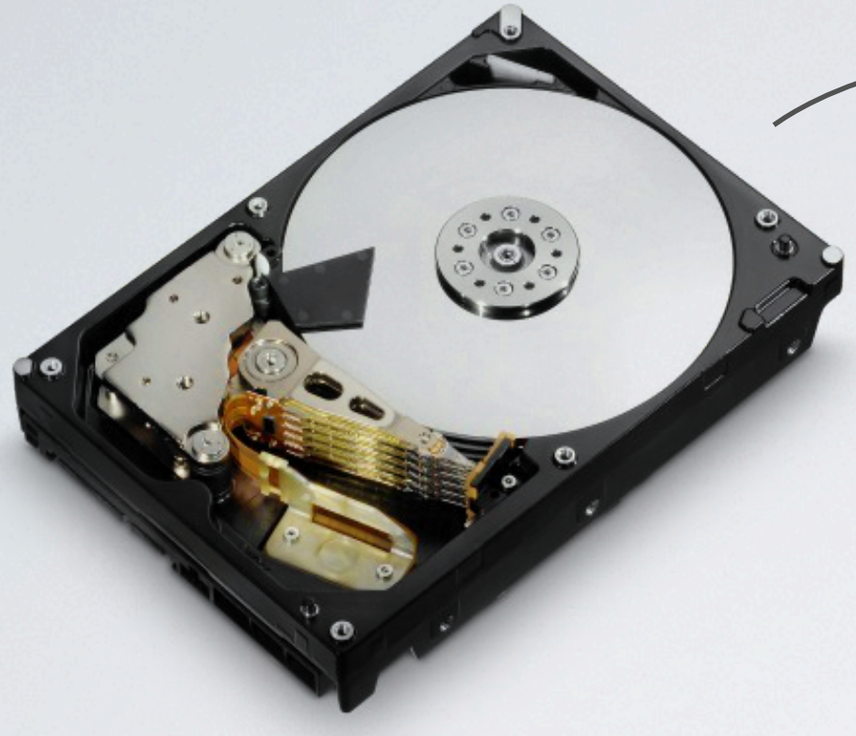
Hard Disk

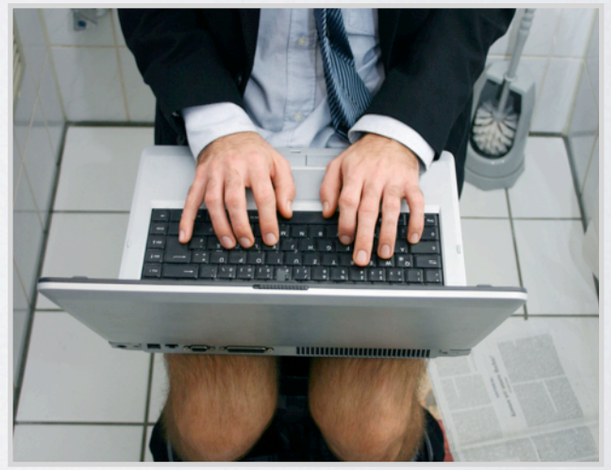


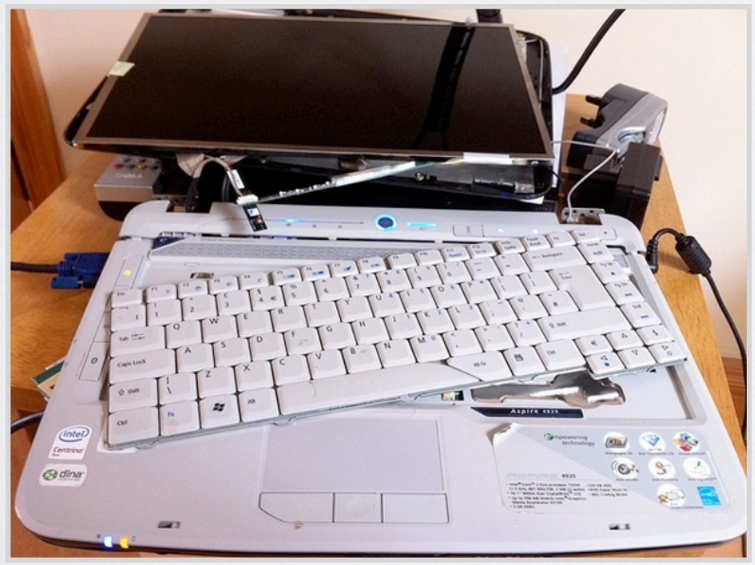
Hard Disk



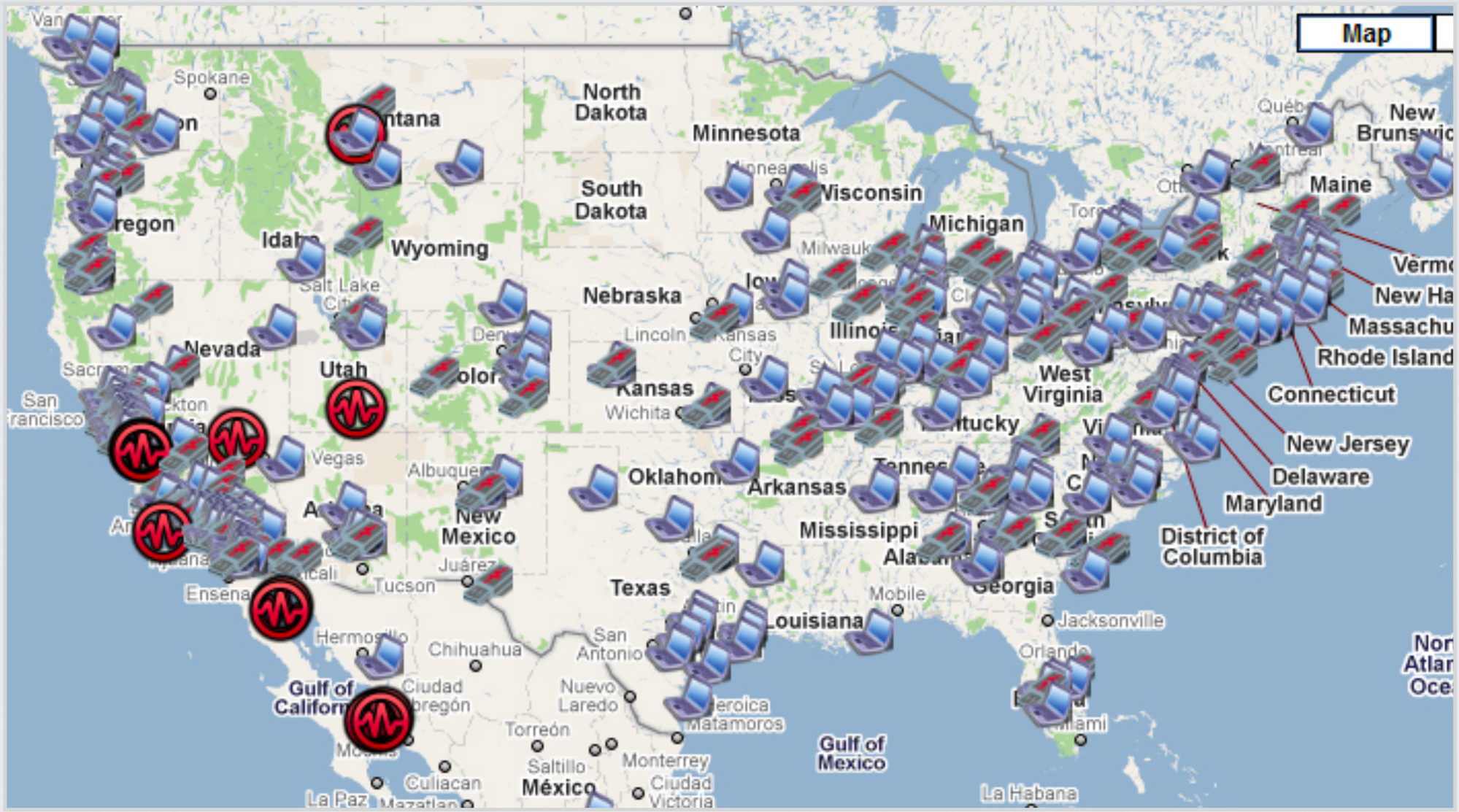
Hard Disk







Quake-Catcher Network - Stanford University



Data is Valuable

Big Data before anybody else...

AT&T

VISA

NASDAQ®

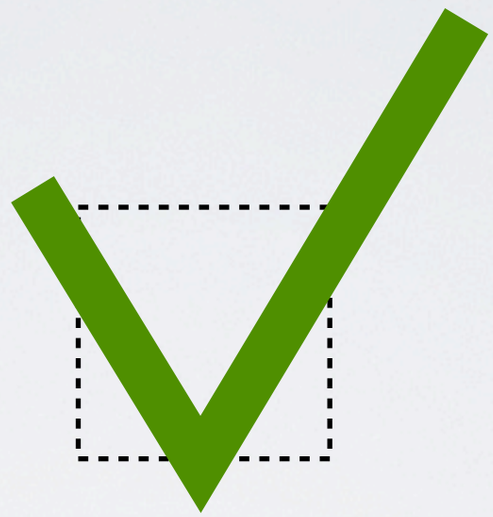
73% stock trades are made by algorithms

Oxford Health

$200/year discount if you connect FourSquare

PROGRESSIVE®

discounts if you install device to track driving habits



I have Read and Agree to the Term of Service

Men Keep Cellphones in their Pockets

Terms of Service

1) free but we can make you pay

2) if phone explodes not our fault

3) we may take your temperature

$5/year cost of Gmail Account

500 Million of Active Users

Free in-flight WiFi

Fast Browser

Free Wifi - Chelsea NYC

Google Analytics

Web Traffic Reports

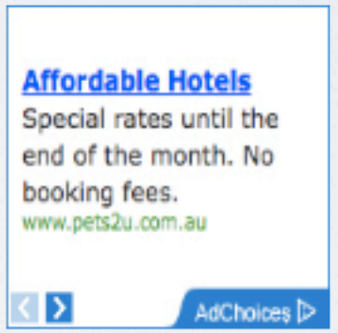
Google Adsense - 94% of Google Revenue

Leaderboard (728 x 90)

Banner (468 x 60)

Half Banner (234x60)

Button (125x125)



Free in-flight WiFi

Fast Browser

Free Wifi

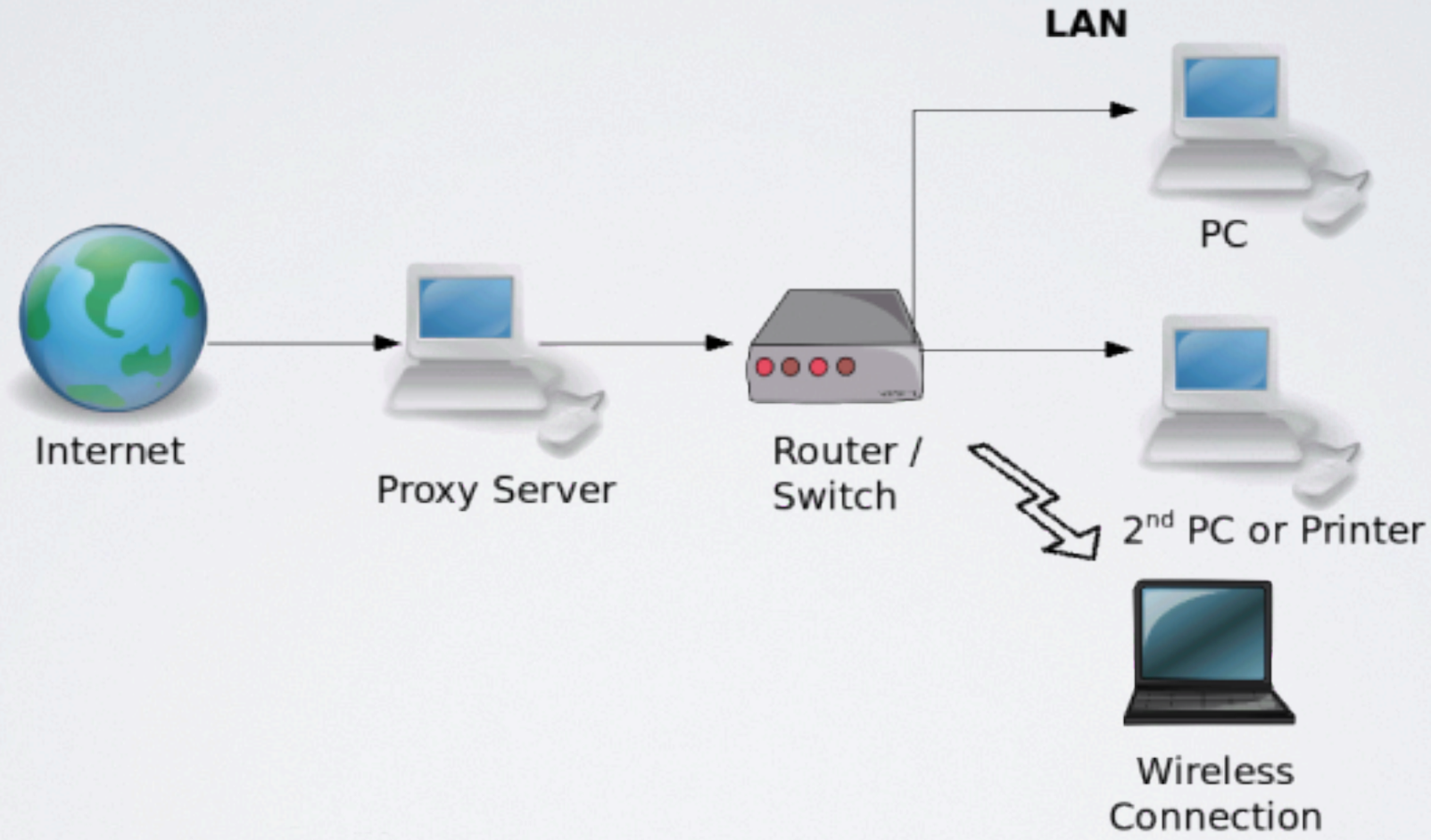
Web Traffic Reports

1 Billion Active Users

800 Million Posts per Day

80% users browse the web
while still logged on Facebook

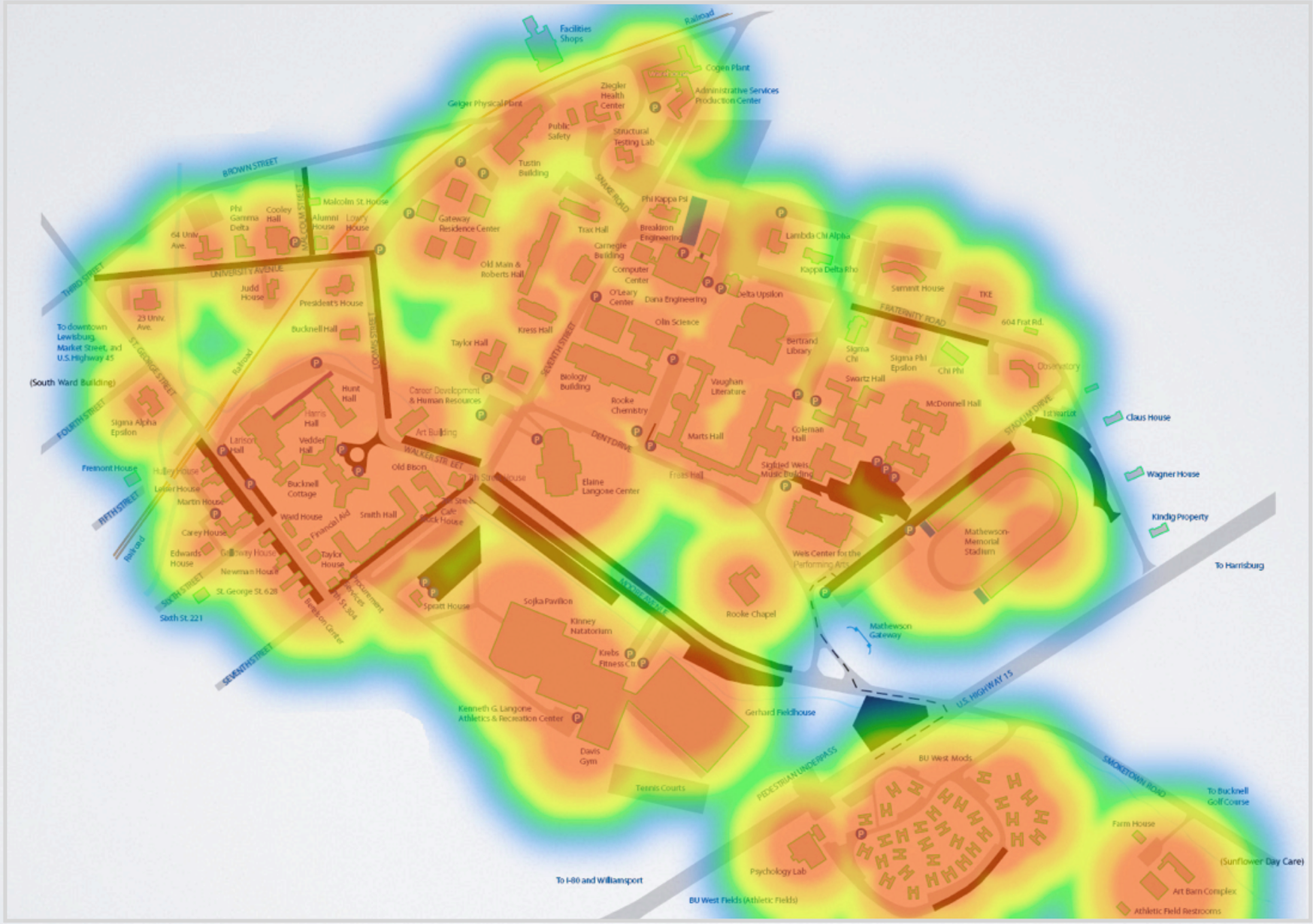
Get Proxy Logs from your IT Department



Guess what this guy was concerned about?

192.168.3.174 [21/Sep/2011:14:35:10]
"GET **http://www.espn.com/football** HTTP/1.1" 200 677
192.168.3.174 [21/Sep/2011:15:15:50]
"GET **http://www.facebook.com/** HTTP/1.1" 200 14287
192.168.3.174 [21/Sep/2011:15:28:41]
"GET **http://www.facebook.com/login** HTTP/1.1" 200 4567
192.168.3.174 [21/Sep/2011:15:43:00]
"GET **http://www.google.com/search?q=genital+rash** HTTP/1.1" 200 1677
192.168.3.174 [21/Sep/2011:15:44:33]
"GET **http://www.cdc.gov/std/herpes/stdfact-herpes.htm** HTTP/1.1" 200 18632
192.168.3.174 [21/Sep/2011:15:48:21]
"GET **http://www.cdc.gov/std/herpes/screening.htm** HTTP/1.1" 200 14632
192.168.3.174 [21/Sep/2011:15:49:27]
"GET **http://www.webmd.com/std/pictures-and-facts** HTTP/1.1" 200 17397
192.168.3.174 [21/Sep/2011:15:53:54]
"GET **http://www.google.com/search?q=dermatologist+52240** HTTP/1.1" 200 1693

Use Wireless Access Logs to Track Movements

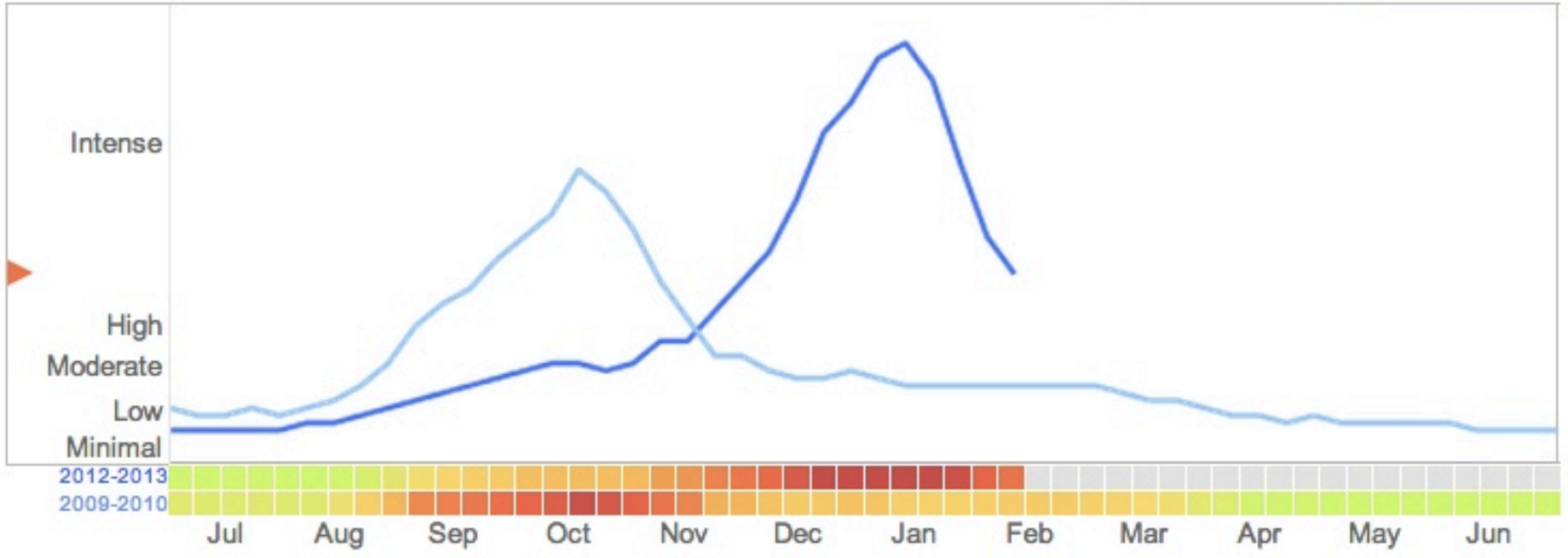


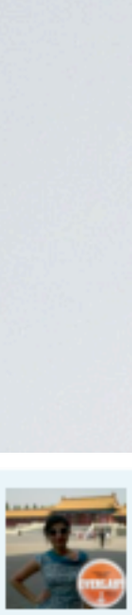
User Generated Data

Google Flu

National

2012-2013 2009-2010 ▼





lexarod Lexa Mendoza

Off to work going home now long day in work so happy I'm going home now so tired I'm still **sick** with **flu** ☹

48 minutes ago

jordanbby7717 Jordan Daniel

I seriously can't take working at **my job** anymore. So **sick of** dealing with the people I work with and customers #getmeoutofhere

34 minutes ago

viagraviagras cheap viagra

order **Viagra** buy dublin, buy cheap online Sildenafil Citrate generic ...: We offer **Viagra** online for lowest pric... bit.ly/etelnY

35 minutes ago

rubygreen1 ruby green

i don't think i have ever been this **sick** had the **flu** last night, woke up this morning with the biggest and most painful cramps in my tommy.

2 hours ago

SelenaSeesStars Selena Martinez

Eating a **hamburger** that my mommy made me. ^.^

1 hour ago

twitter

lexarod Lexa Mendoza
Off to work going home now long day in work so happy I'm going home now so tired I'm still **sick** with **flu** ☹
48 minutes ago

jordanbby7717 Jordan Daniel
I seriously can't take working at **my job** anymore. So **sick of** dealing with the people I work with and customers #getmeoutofhere
34 minutes ago

viagraviagras cheap viagra
order **Viagra** buy dublin, buy cheap online Sildenafil Citrate generic ...: We offer **Viagra** online for lowest pric... bit.ly/etelnY
35 minutes ago

SelenaSeesStars Selena Martinez
Eating a **hamburger** that my mommy made me. ^.^
1 hour ago

rubygreen1 ruby green
i don't think i have ever been this **sick** had the **flu** last night, woke up this morning with the biggest and most painful cramps in my tommy.
2 hours ago

twitter

lexarod Lexa Mendoza
Off to work going home now long day in work so happy I'm going home now so tired I'm still **sick** with **flu** ☹
48 minutes ago

jordanbby7717 Jordan Daniel
I seriously can't take working at **my job** anymore. So **sick of** dealing with the people I work with and customers #getmeoutofhere
34 minutes ago

viagraviagras cheap viagra
order **Viagra** buy dublin, buy cheap online Sildenafil Citrate generic ...: We offer **Viagra** online for lowest pric... bit.ly/etelnY
35 minutes ago

SelenaSeesStars Selena Martinez
Eating a **hamburger** that my mommy made me. ^.^
1 hour ago

rubygreen1 ruby green
i don't think i have ever been this **sick** had the **flu** last night, woke up this morning with the biggest and most painful cramps in my tommy.
2 hours ago

twitter

lexarod Lexa Mendoza
Off to work going home now long day in work so happy I'm going home now so tired I'm still **sick** with **flu** ☹
48 minutes ago

jordanbby7717 Jordan Daniel
I seriously can't take working at **my job** anymore. So **sick of** dealing with the people I work with and customers #getmeoutofhere
34 minutes ago

viagraviagras cheap viagra
order **Viagra** buy dublin, buy cheap online Sildenafil Citrate generic ...: We offer **Viagra** online for lowest pric... bit.ly/etelnY
35 minutes ago

SelenaSeesStars Selena Martinez
Eating a **hamburger** that my mommy made me. ^.^
1 hour ago

rubygreen1 ruby green
i don't think i have ever been this **sick** had the **flu** last night, woke up this morning with the biggest and most painful cramps in my tommy.
2 hours ago

twitter

lexarod Lexa Mendoza
Off to work going home now long day in work so happy I'm going home now so tired I'm still **sick** with **flu** ☹
48 minutes ago

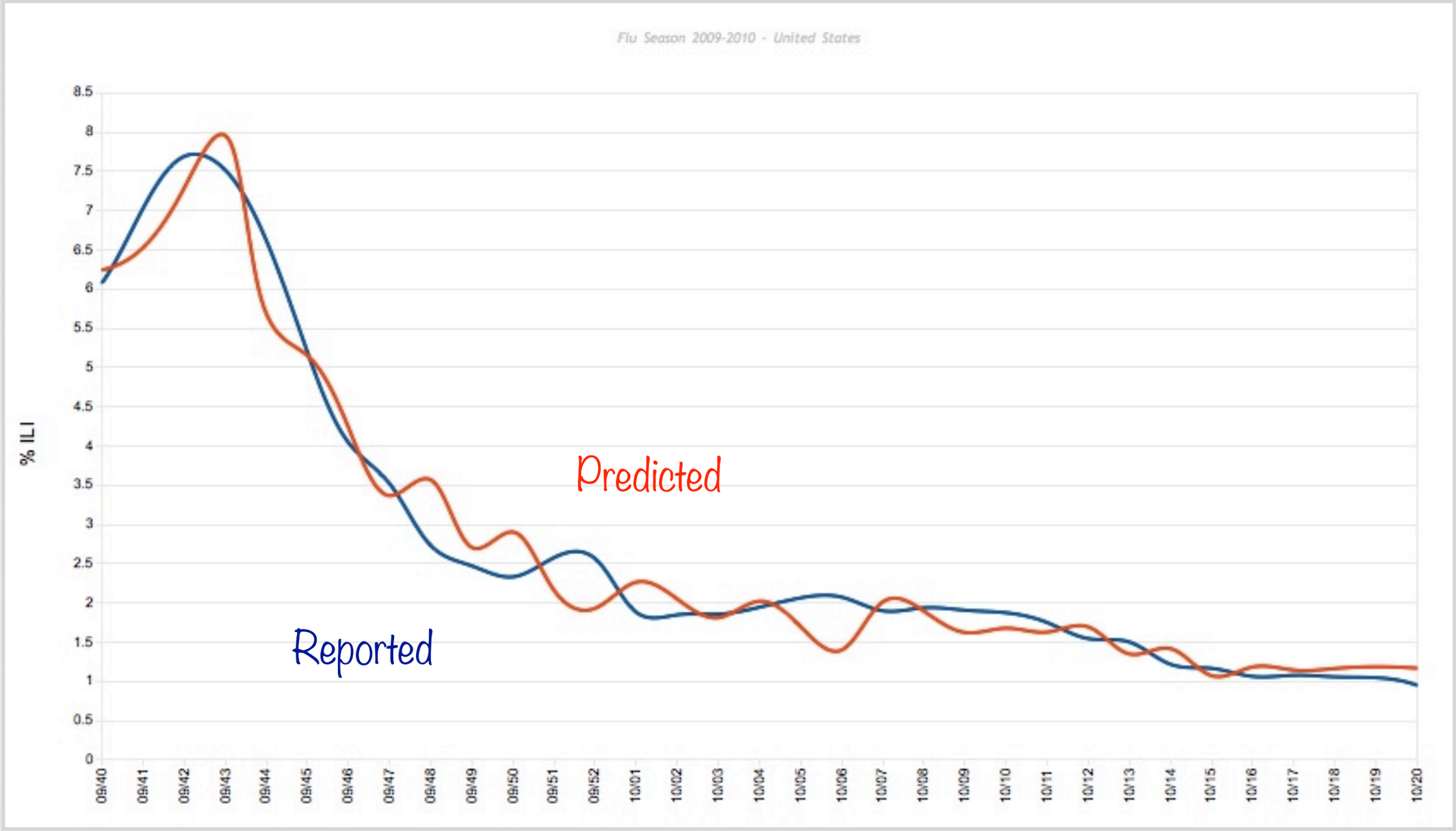
jordanbby7717 Jordan Daniel
I seriously can't take working at **my job** anymore. So **sick of** dealing with the people I work with and customers #getmeoutofhere
34 minutes ago

viagraviagras cheap viagra
order **Viagra** buy dublin, buy cheap online Sildenafil Citrate generic ...: We offer **Viagra** online for lowest pric... bit.ly/etelnY
35 minutes ago

rubygreen1 ruby green
i don't think i have ever been this **sick** had the **flu** last night, woke up this morning with the biggest and most painful cramps in my tommy.
2 hours ago

SelenaSeesStars Selena Martinez
Eating a **hamburger** that my mommy made me. ^.^
1 hour ago

Predicting ILI% using twitter



foursquare

Doctor

Gym

Restaurant

foursquare

Doctor

Gym

Restaurant

Drunken

Mayor

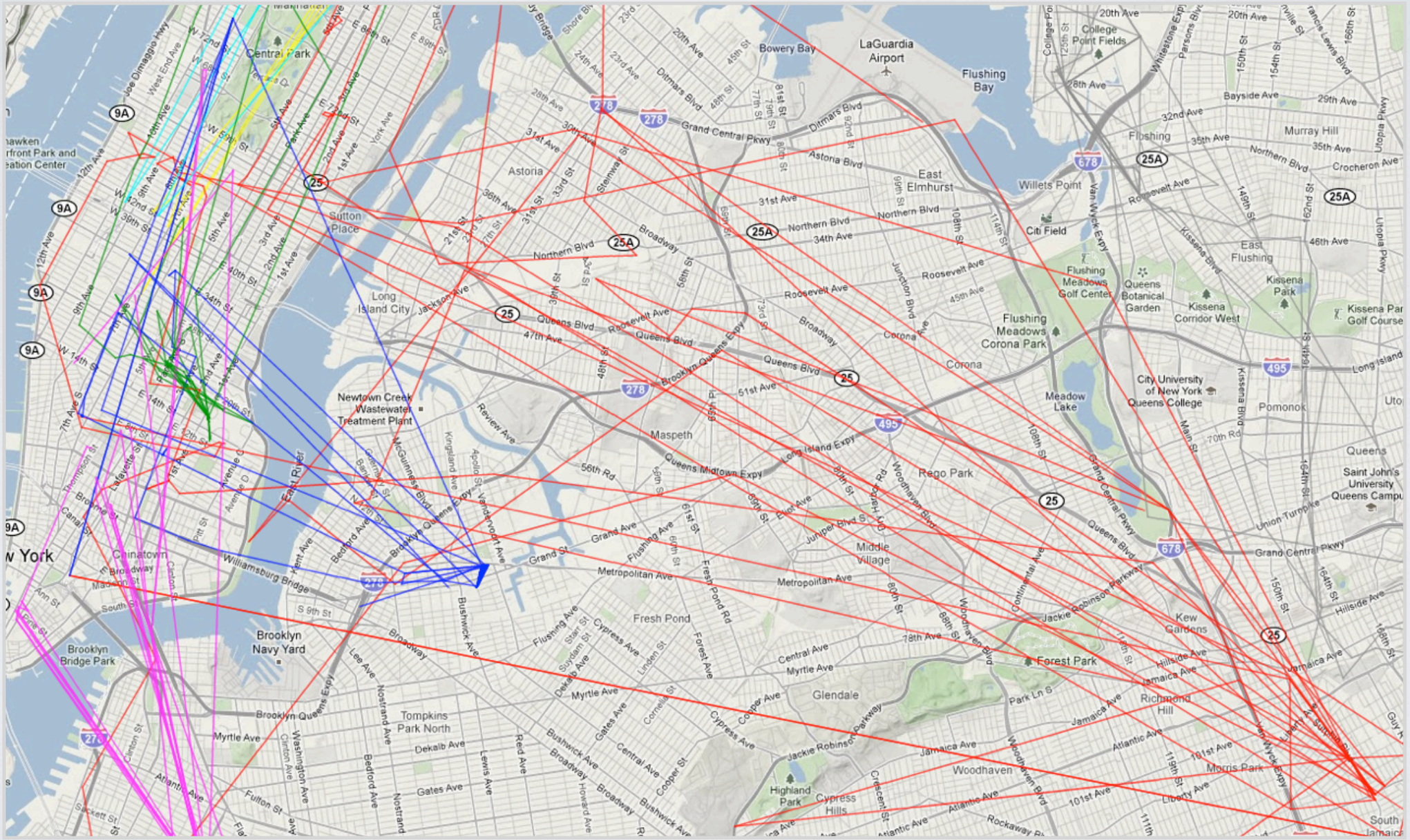
Movie

Pizza

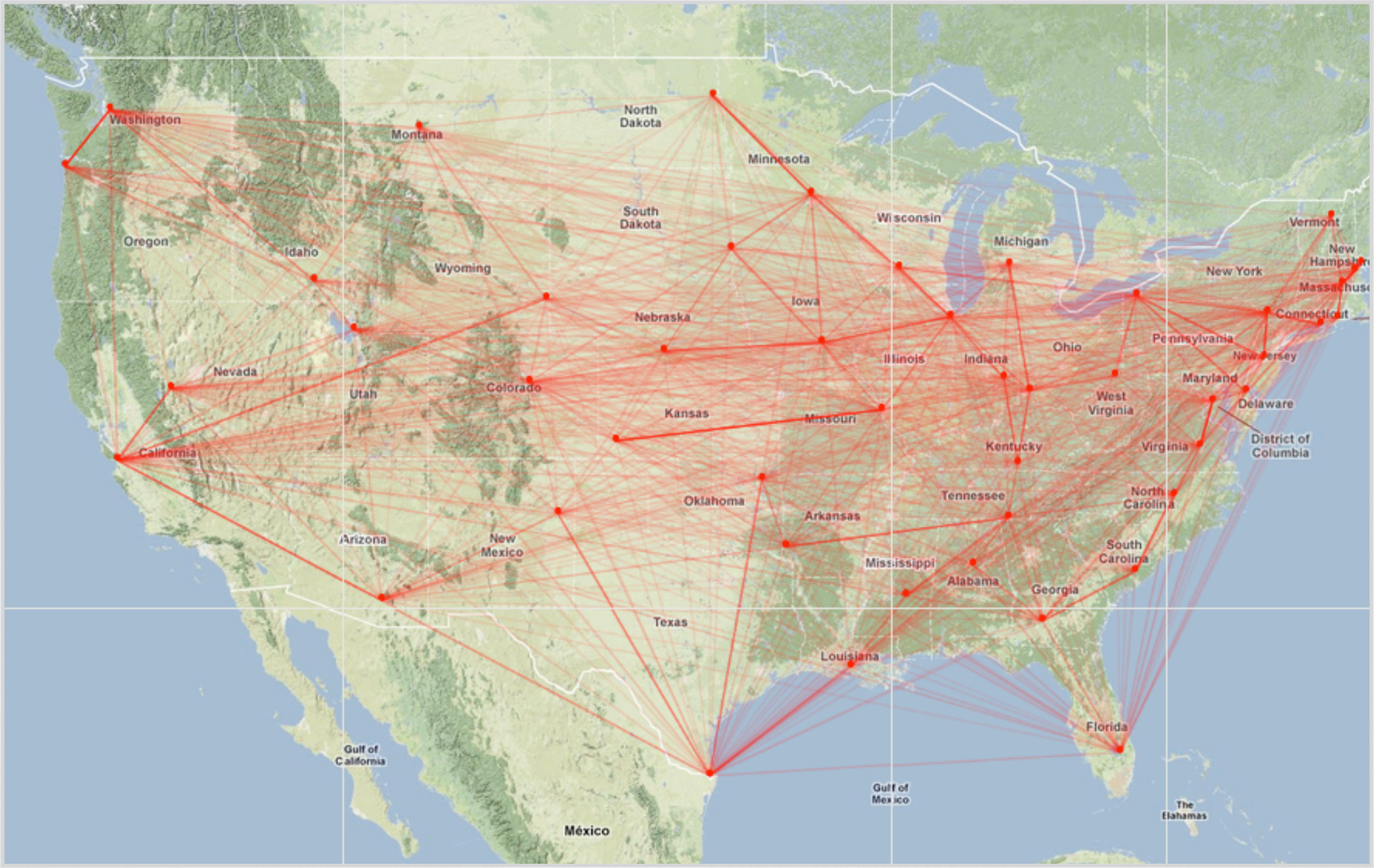
Traveler

Gym-rat

foursquare


Central Park
Bowery Bay
LaGuardia Airport
Flushing Bay
College Point Fields
Bayside Ave
Murray Hill
Flushing
Willets Point
Citi Field
East Elmhurst
Astoria
Sutton Place
Long Island City
Newtown Creek Wastewater Treatment Plant
Maspeth
Flushing Meadows Corona Park
Queens Botanical Garden
Kissena Corridor West
Kissena Park
East Flushing
Meadow Lake
City University of New York Queens College
Pomonok
Rego Park
Middle Village
Fresh Pond
Forest Park
Glendale
Richmond Hill
Kew Gardens
Woodhaven
Highland Park
Cypress Hills
Brooklyn Navy Yard
Brooklyn Bridge Park
Tompkins Park North
Chinatown
Williamsburg Bridge
Queens Midtown Expy
Long Island Expy
Grand Central Pkwy
Northern Blvd
Queens Blvd
Roosevelt Ave
Broadway
Jackie Robinson Parkway
Brooklyn Queens Expy
Ditmars Blvd
Astoria Blvd
Metropolitan Ave
Jamaica Ave
Atlantic Ave
Woodhaven Blvd
Utopia Pkwy
Union Turnpike
Hillside Ave
Main St
Kissena Blvd
Whitestone Expy
Van Wyck Expy
Cross Island Pkwy
Grand Central Pkwy
Saint John's University Queens Campus
Queens
Pomonok
Kissena Park
Kissena Park Golf Course
Flushing Meadows Golf Center
Corona
Jackson Ave
Review Ave
Kingsland Ave
McGuinness Blvd
Vandervoort Ave
Bushwick Ave
Flushing Ave
Cypress Ave
Forest Ave
Fresh Pond Rd
Myrtle Ave
Central Ave
Cooper Ave
Gates Ave
Lewis Ave
Red Ave
Nostrand Ave
Bedford Ave
Lee Ave
Broadway
Northern Blvd
Queens Blvd
Roosevelt Ave

foursquare


Washington
Oregon
Idaho
Montana
Wyoming
North Dakota
Minnesota
South Dakota
Nebraska
Iowa
Wisconsin
Michigan
Vermont
New Hampshire
Massachusetts
New York
Connecticut
Pennsylvania
New Jersey
Ohio
Illinois
Indiana
Maryland
Delaware
District of Columbia
West Virginia
Kentucky
Virginia
Nevada
Utah
Colorado
Kansas
Missouri
California
Oklahoma
Arkansas
Tennessee
North Carolina
South Carolina
Mississippi
Alabama
Georgia
Florida
Arizona
New Mexico
Texas
Louisiana
Gulf of California
Gulf of Mexico
México
The Bahamas

twitter

Matt @apple4ever 33 mins
My weight: 228.1 lb. 45.1 lb to go. Started at 234.6... withings.com
Expand

Brad Wood @randombrad 37 mins
My weight: 223.2 lb. 22.6 lb to go. #gettingInShape withings.com
Expand

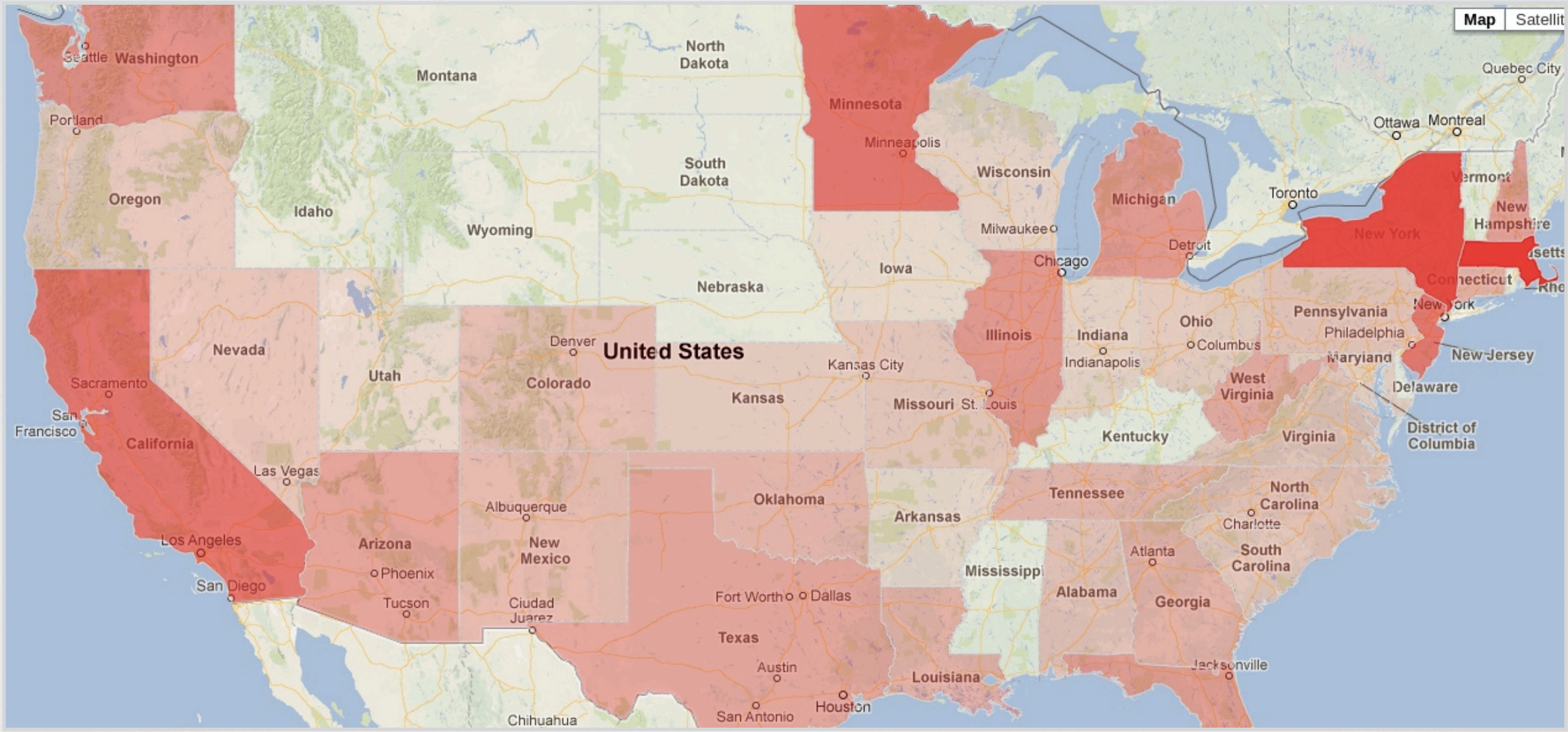
Shawn Hutchinson @Hutchbot1984 36 mins
I **ran** 6.30 **km** with @MapMyFitness. Check out my route in Carlisle ! #run #running #HWD #Sweat
Expand

bigbadbootyman @marckrajewski 22 mins
I **ran** with MapMyRun! Distance: 17.74km, time: 02:04:12, pace: 7:34min/**km**, speed: 7.93km/h. mapmyrun.com/view_route?r=6…
Expand

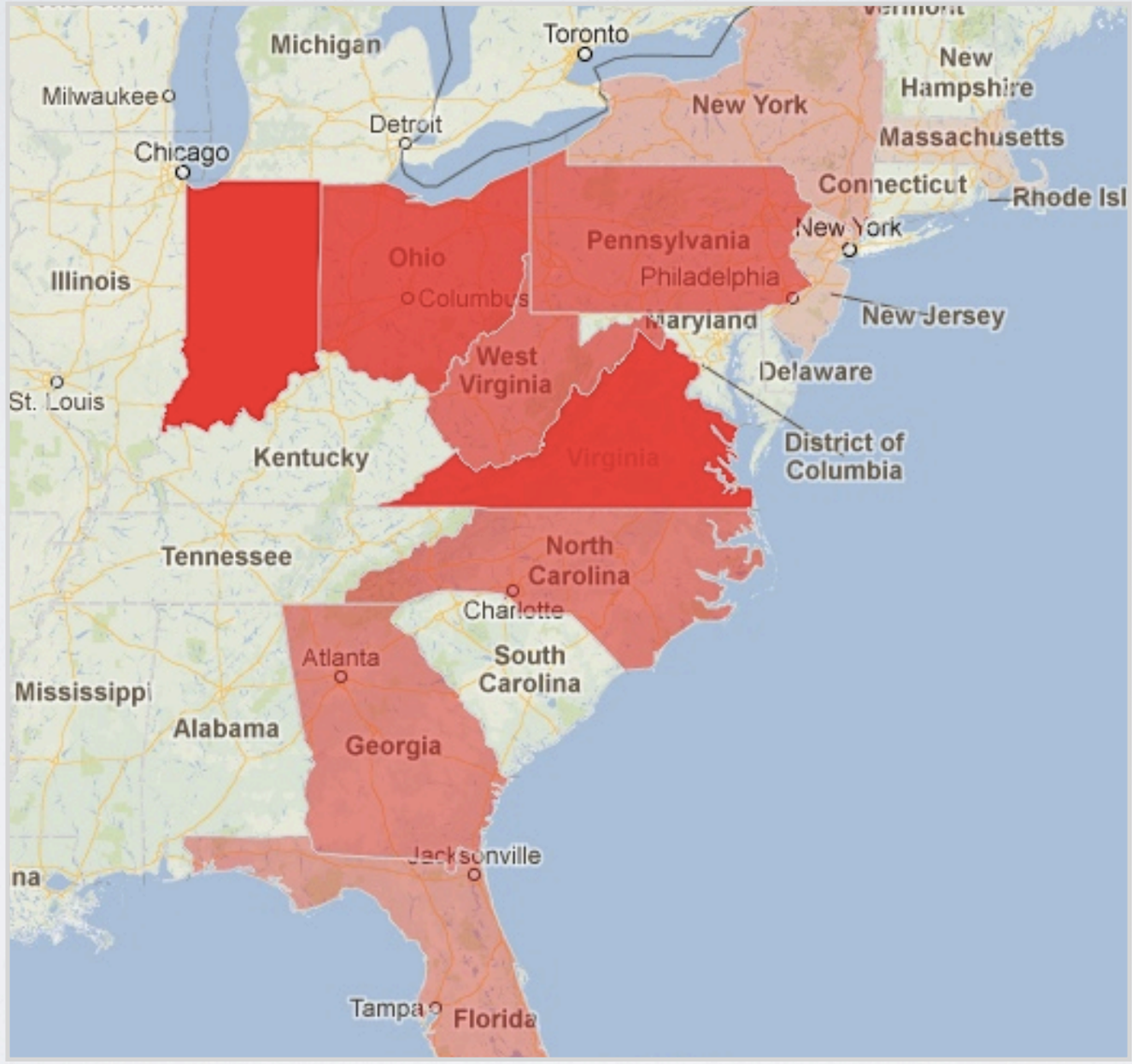
Chelsea Grothe. @ChelseaGrothe 16 mins
I just **ran** 8.4 **km**
Expand

John McWaldi @McWaldi 7 mins
Mein Gewicht: 66.2 kg. first try :) withings.com
Expand

Average Length of Runs - twitter



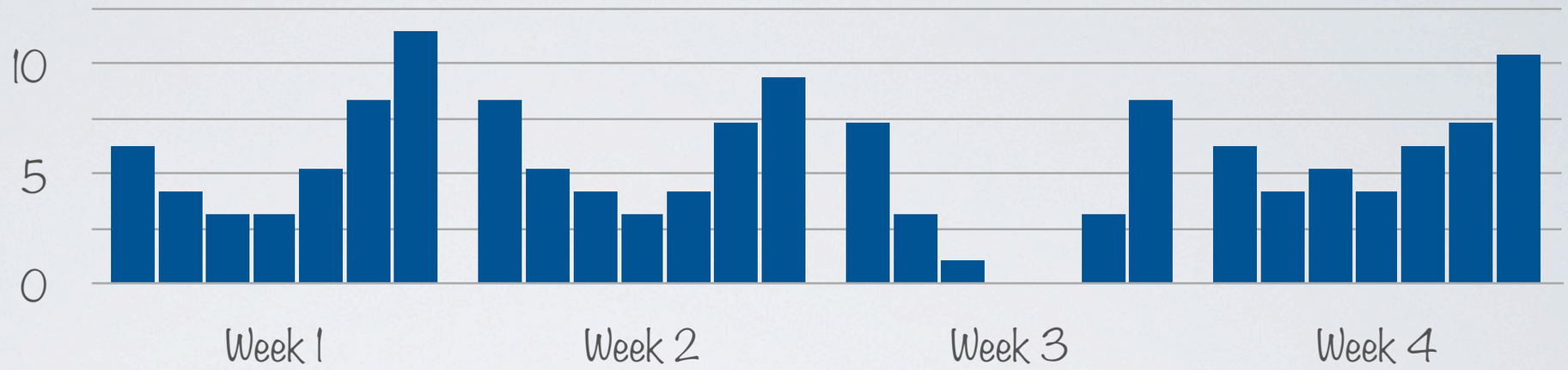
Average Weight - twitter



Social Media Monitoring

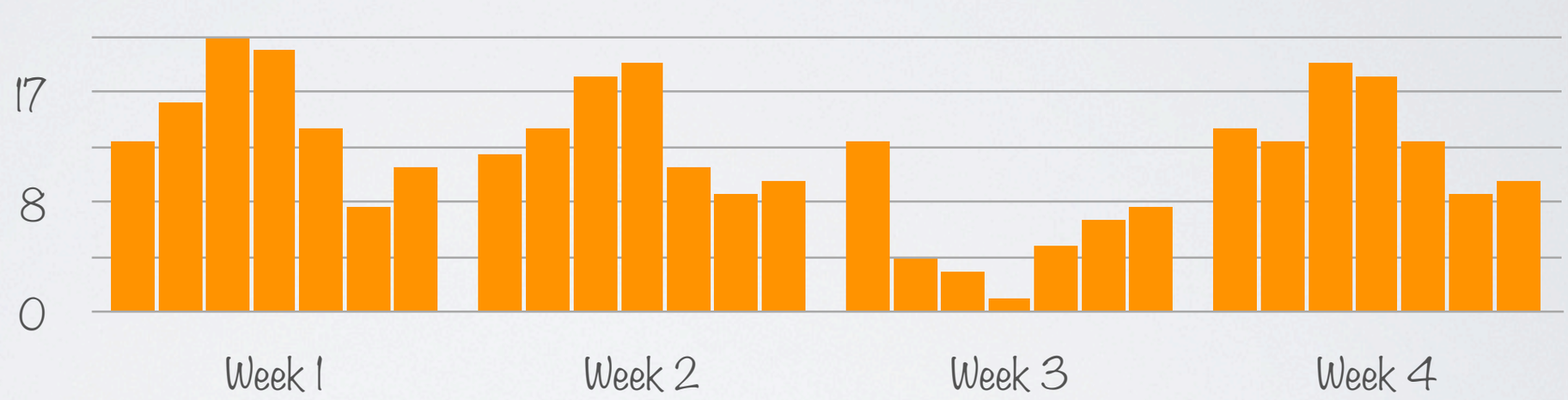
foursquare

checkins



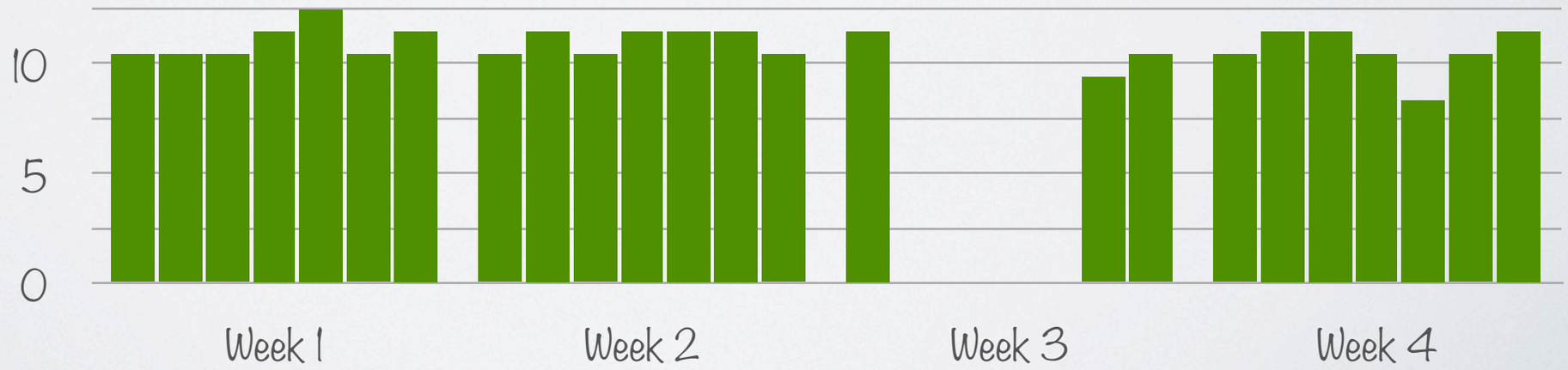
twitter

tweets



RunKeeper

km

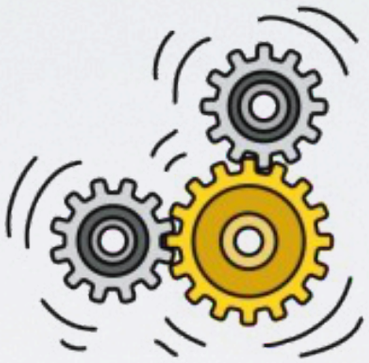


Social Media Monitoring

foursquare
checkins

twitter
tweets

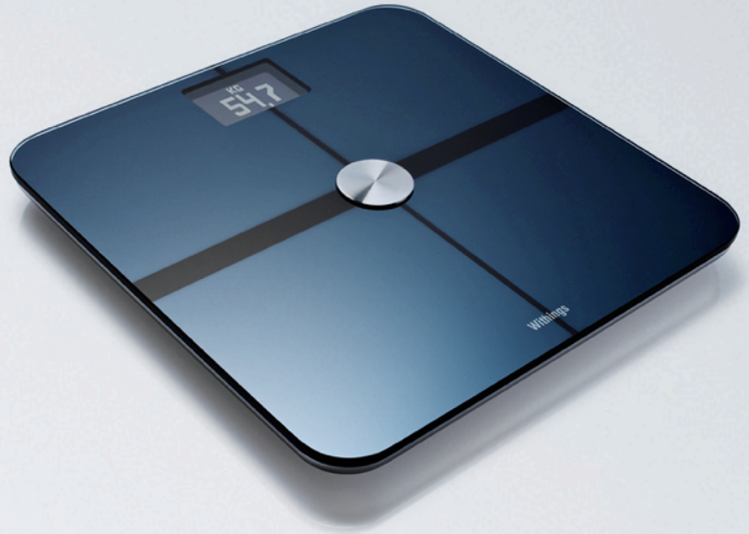
RunKeeper
km



New Devices

FitBit

BodyMedia



Withings

Nike Fuel

iHealth

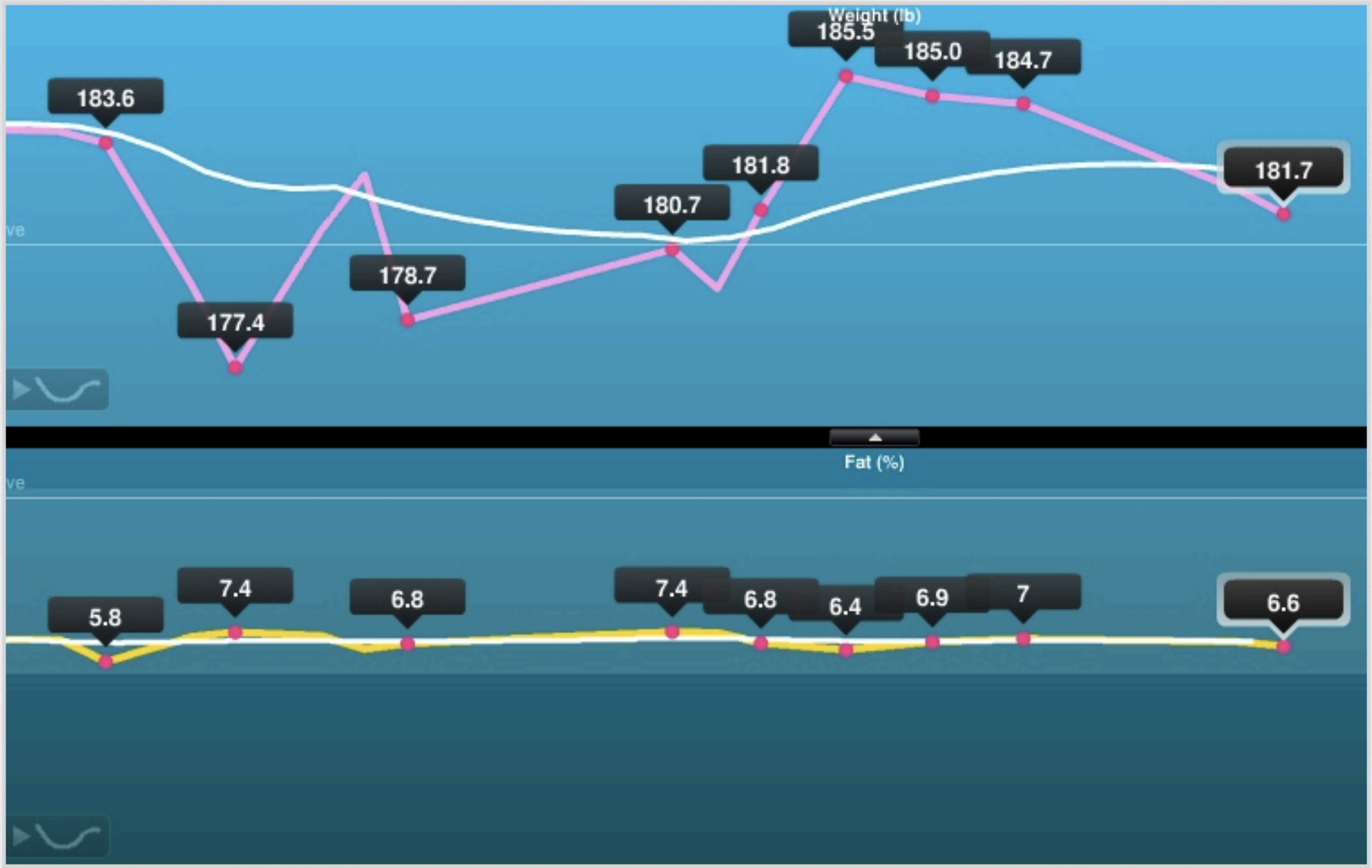
Ki

Zeo

Jawbone Up

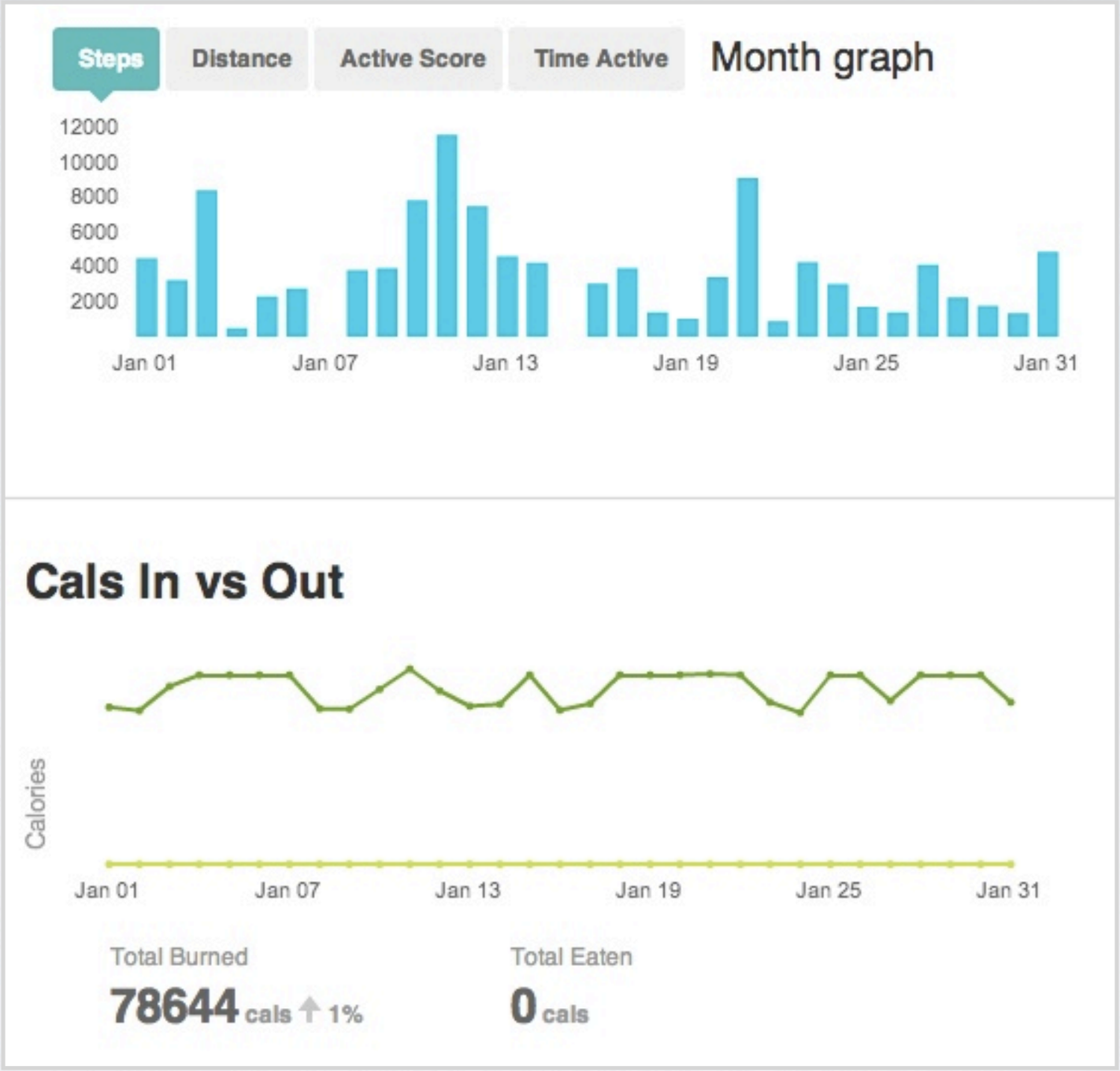
Withings Scale

Withings - Weight/Fat Statistics



FitBit Zip

FitBit - Daily Steps



iHealth Pressure Monitor

Zeo Sleep Trainer

iPhone

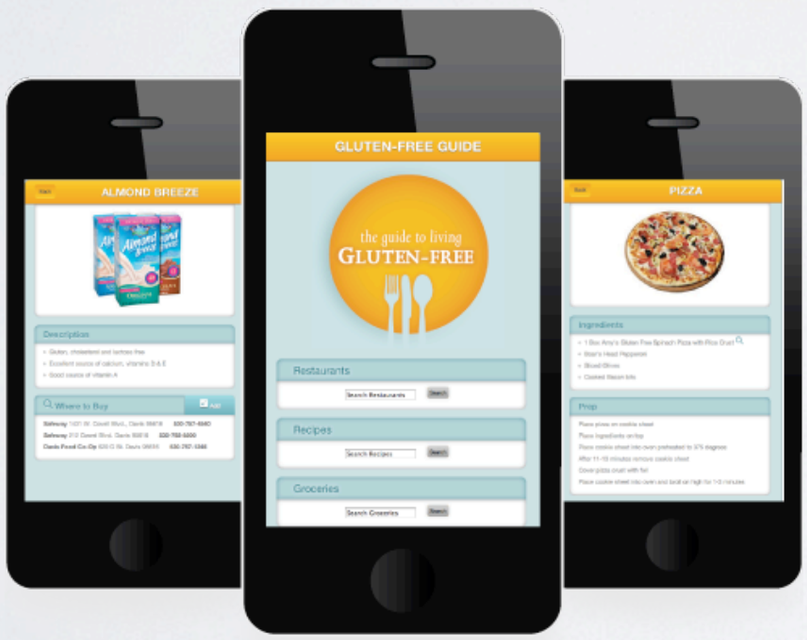
Android

Instant Heart Monitor

Cardiio

Calories Tracking

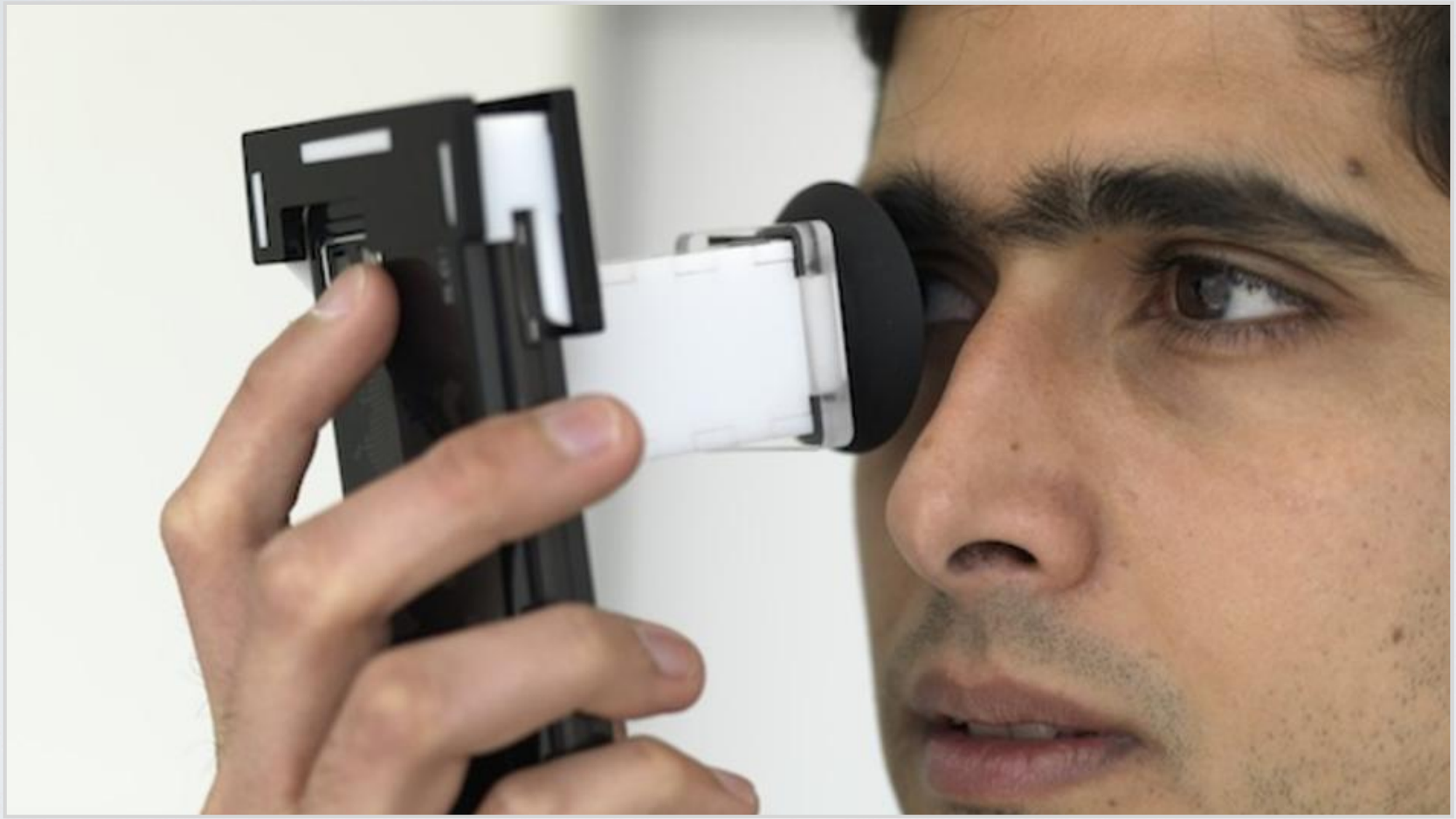
Running Partner



Dietary Restriction

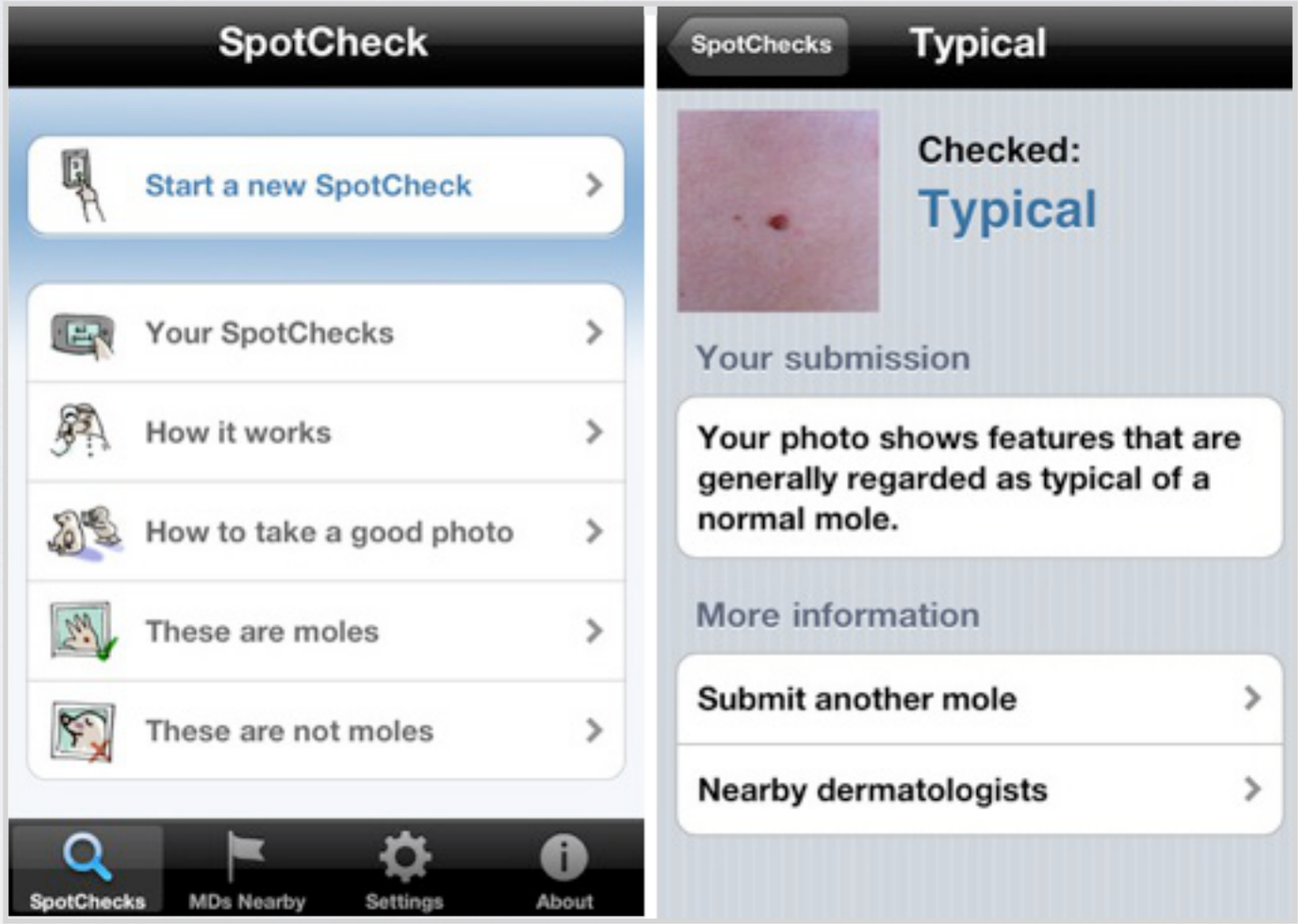
Addiction Management

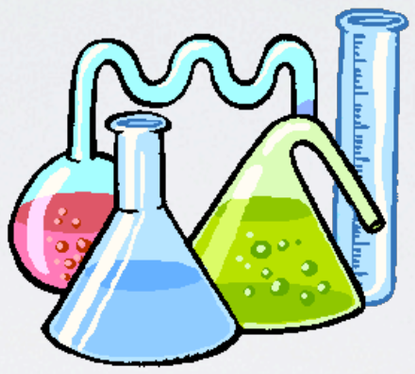
Eye Exam



Glucodock

Spotcheck


SpotCheck
Start a new SpotCheck
Your SpotChecks
How it works
How to take a good photo
These are moles
These are not moles
SpotChecks
MDs Nearby
Settings
About
SpotChecks
Typical
Checked:
Typical
Your submission
Your photo shows features that are generally regarded as typical of a normal mole.
More information
Submit another mole
Nearby dermatologists



New Diagnostic Paths

"Dystextia": Stroke diagnosed with help of garbled text message

ISTOCKPHOTO

Monitoring Computer Interactions to Detect Early Cognitive Impairment in Elders

Holly Jimison, *Member, IEEE*, Nicole Jessey, James McKanna, and Tracy Zitzelberger, Jeffrey Kaye

***Abstract*—Maintaining cognitive performance is a key factor influencing elders' ability to live independently with a high quality of life. We have been developing unobtrusive measures to monitor cognitive performance and potentially predict decline using information from routine computer interactions in the home. Early detection of cognitive decline offers the potential for intervention at a point when it is likely to be more successful. This paper describes recommendations for the conduct of studies monitoring cognitive function based on routine computer interactions in elders' home environments.**

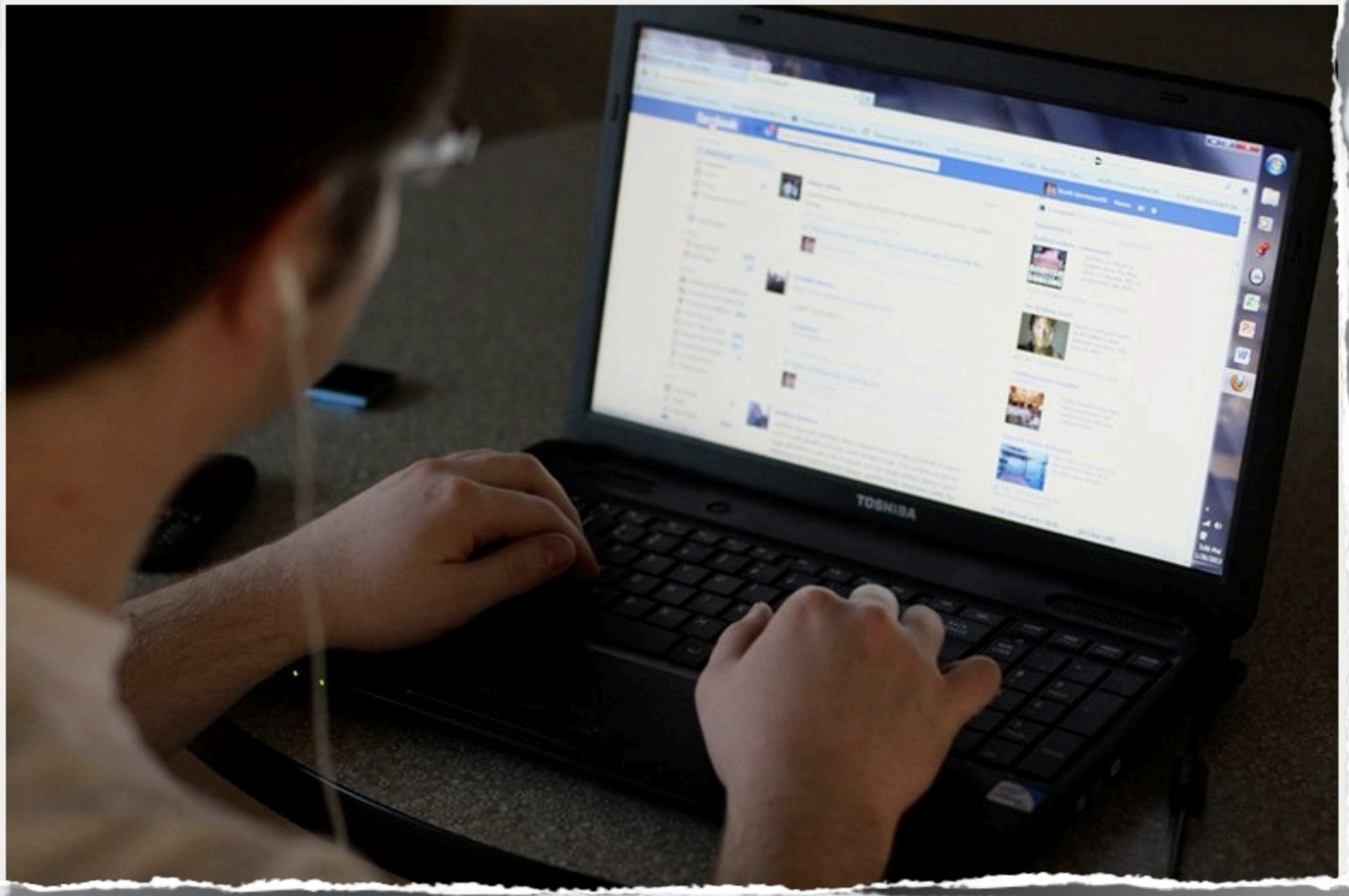
I. Introduction

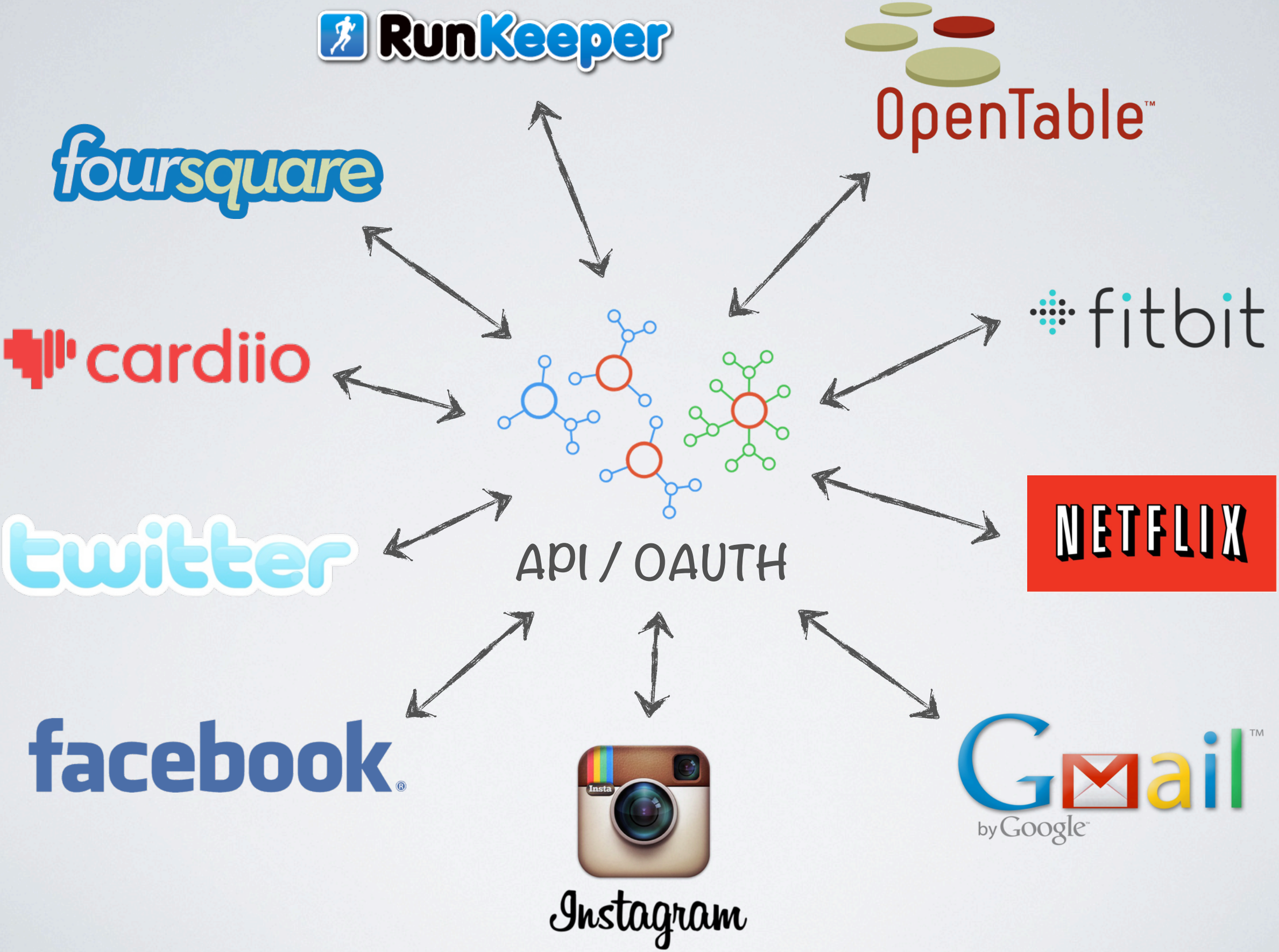
ELDERS are the fastest growing demographic in many countries, with a concomitant escalation of health care resources being spent on conditions associated with aging. One of the key functional losses at risk with aging is a decline in cognitive abilities. Estimates vary depending on assessment protocols and populations studied, but up to 50% of all individuals over age 85 are found to have measurable decline in cognitive function [1]. These individuals are at high risk for dementia, requiring increasing levels of assisted living. Even mild cognitive declines lead to degraded quality of life. Thus a major goal of seniors and their families is to optimize their quality of life and remain

useful in predicting an elder's degree of cognitive health. These metrics include relative typing speed and accuracy, relative mouse movement efficiency, and relative performance on computer games that have been designed for cognitive monitoring. Our continuous monitoring of cognitive indicators allows us to perform trend detection of individual performance, so that we are less susceptible to unwanted variability due to educational, language, and cultural backgrounds, as compared to traditional cognitive assessment tests. In this pilot project, we deployed our monitoring software in the homes of 15 elders in the Portland, Oregon metropolitan area for a period of 6 months and tested our ability to monitor computer activities and keep elders motivated and engaged in using the computer. This test was in preparation for a larger prospective long-term study that will determine our ability to use these types of metrics to predict future cognitive decline. Based on our pilot test of the monitoring software in the homes of elders, we developed a set of recommendations for effectively managing a large number of elderly participants in a long-term study, and eventually, as a routine service available to the public at large.

II. Background

Facebook activity may reflect mental health issues according to MU psychology study




RunKeeper
OpenTable™
foursquare
cardiio
fitbit
twitter
API / OAUTH
NETFLIX
facebook
Insta
Instagram
Gmail™
by Google

Alessio Signorini

alessio-signorini@uiowa.edu

@a_signorini

Computational Epidemiology Group

University of Iowa