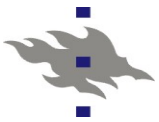
Big Data Applications in Programmatic Marketing Ecosystem

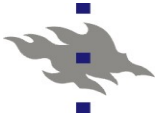
Ilkka O. Lavas
Helsingin Yliopisto
Tietojenkäsittelytieteen laitos



Big Data Applications

Deep Dive to Programmatic Marketing Ecosystem

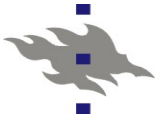
- **Why** this is interesting
- Summary
- How this relates to Big Data
- Summary and Conclusion



Big Data Applications

Big data have many applications in different areas

- science and research
- public health
- customer relation management
- machine and device performance analysis
- optimizing cities and countries
- finance and banking
- Advertising and data driven marketing

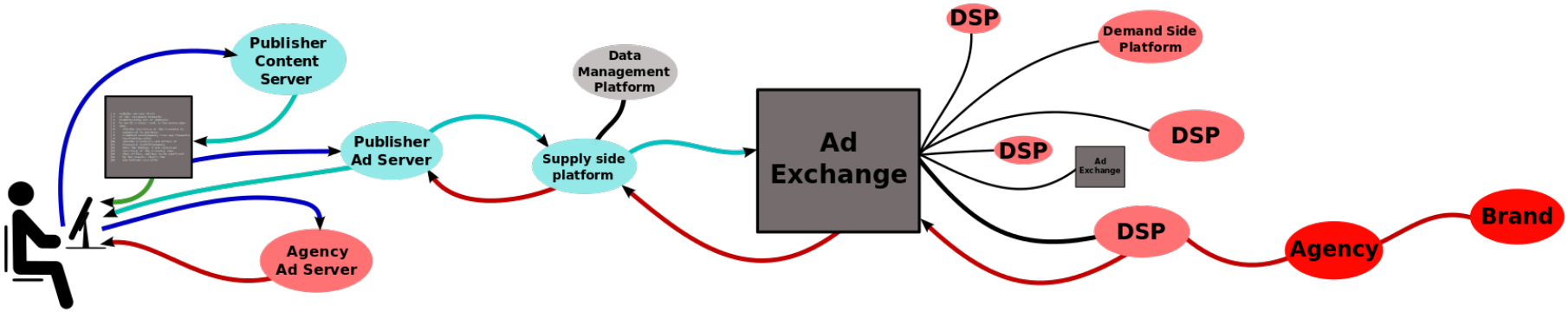


Programmatic Marketing Ecosystem examples

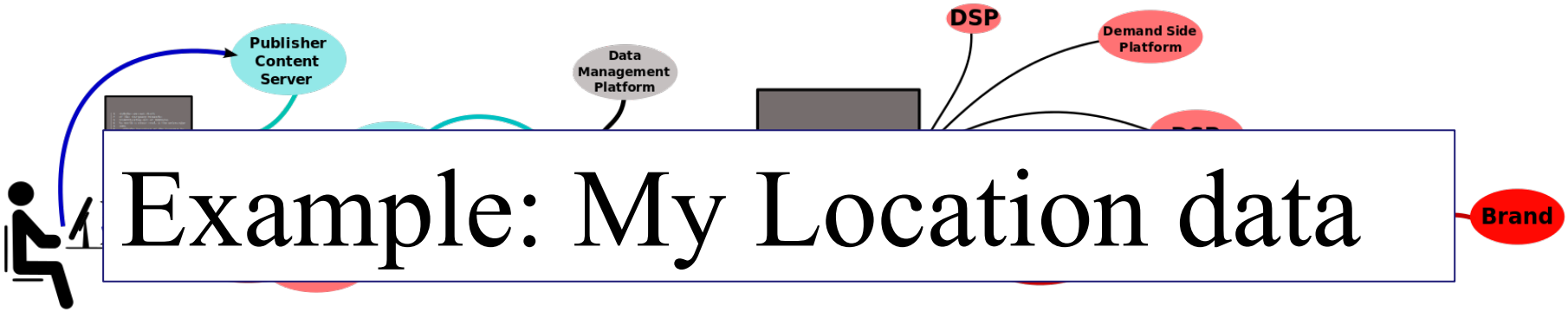
- Programmatic = Automating simple tasks
- Marketing = telling stories to existing and new potential customers
- Ecosystem = distributing of data between systems

- Hotels.com example (retargeting, email reminder)
- Zalando example

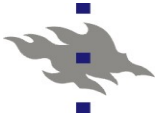
Advertising
Programmatic marketing ecosystem



Advertising
Programmatic marketing ecosystem



Example: My Location data

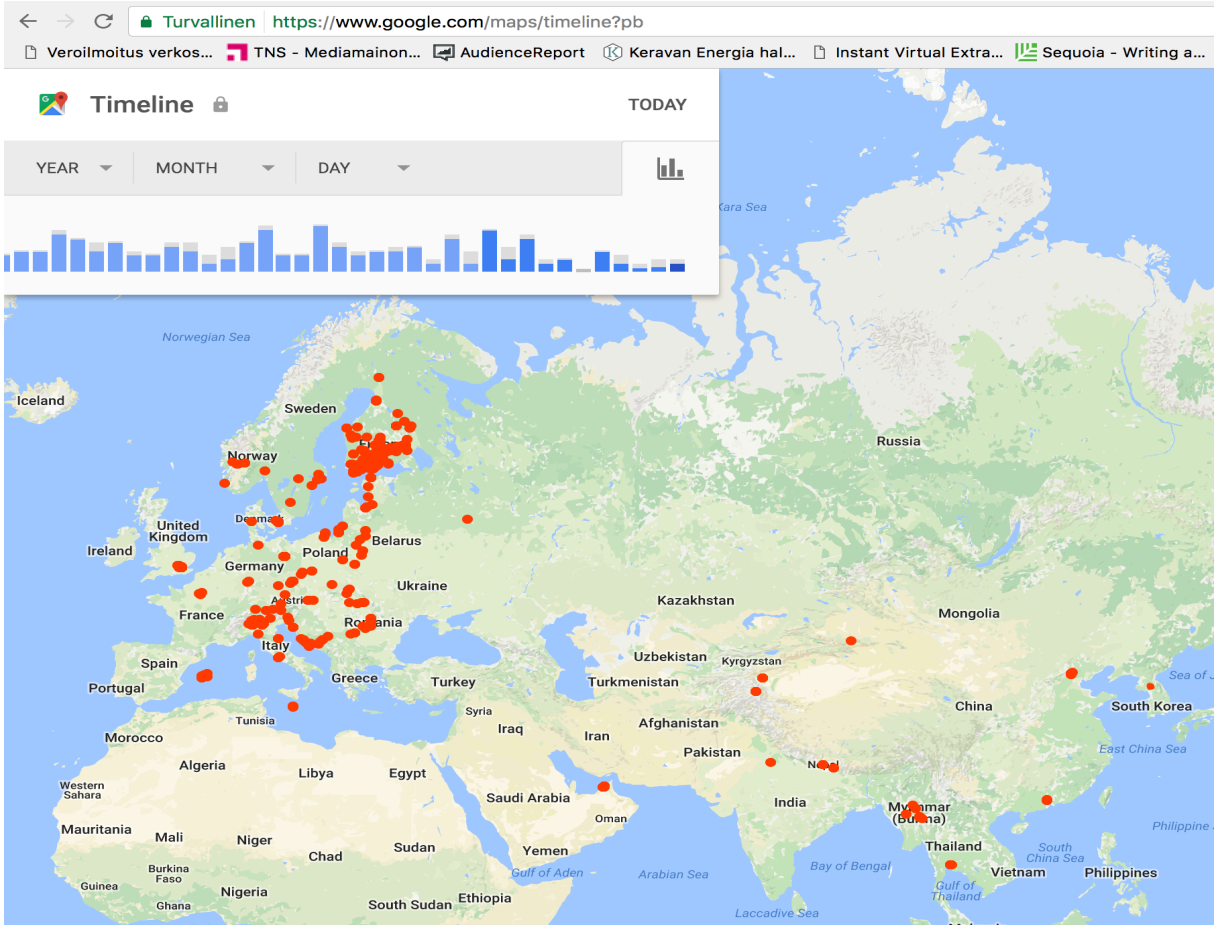


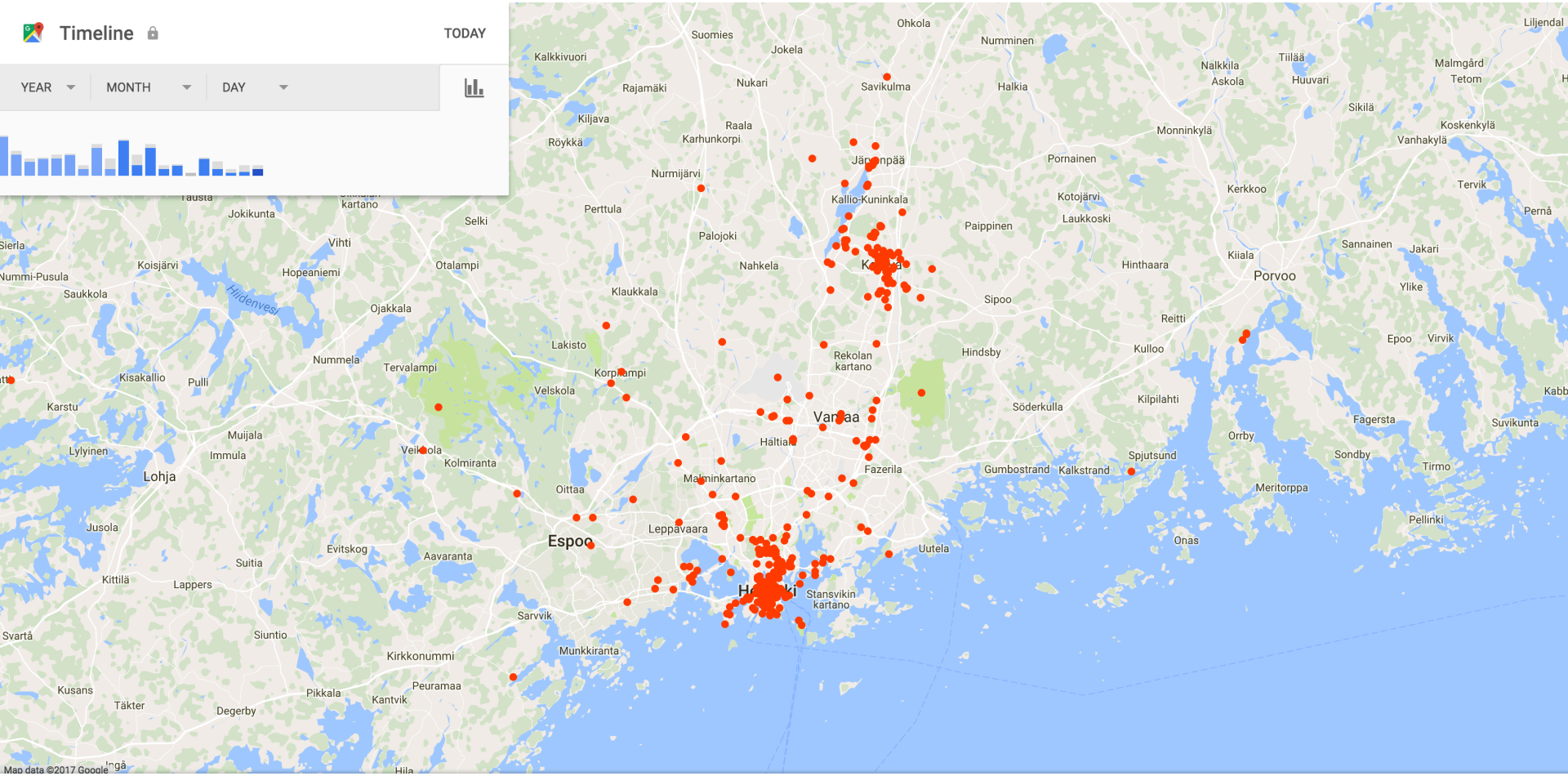
Example:

Geospatial data

combined with location data

is one form of data in programmatic marketing ecosystem

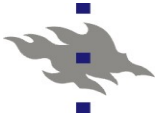

Turvallinen
https://www.google.com/maps/timeline?pb
Veroilmoitus verkos...
TNS - Mediamainon...
AudienceReport
Keravan Energia hal...
Instant Virtual Extra...
Sequoia - Writing a...
Timeline
TODAY
YEAR
MONTH
DAY
Norwegian Sea
Iceland
Sweden
Norway
Russia
United Kingdom
Ireland
Denmark
Belarus
Poland
Germany
Ukraine
France
Kazakhstan
Mongolia
Italy
Spain
Portugal
Greece
Turkey
Uzbekistan
Kyrgyzstan
Turkmenistan
China
South Korea
Tunisia
Syria
Iraq
Iran
Afghanistan
Pakistan
Morocco
Algeria
Libya
Egypt
Saudi Arabia
India
Western Sahara
Mauritania
Mali
Niger
Chad
Sudan
Yemen
Oman
Thailand
Vietnam
Philippines
Burkina Faso
Guinea
Ghana
Nigeria
South Sudan
Ethiopia
Gulf of Aden
Arabian Sea
Bay of Bengal
Gulf of Thailand
East China Sea
South China Sea
Laccadive Sea



Ilkka Lavas, Helsingin Yliopisto, Tietojenkäsittelytieteen laitos

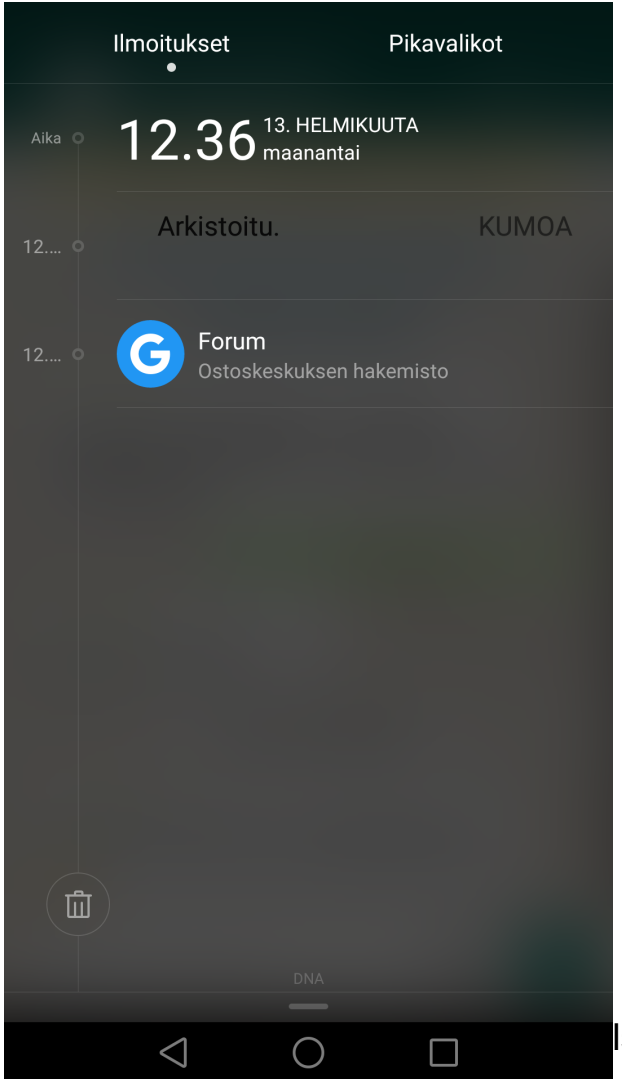
Why this is interesting?

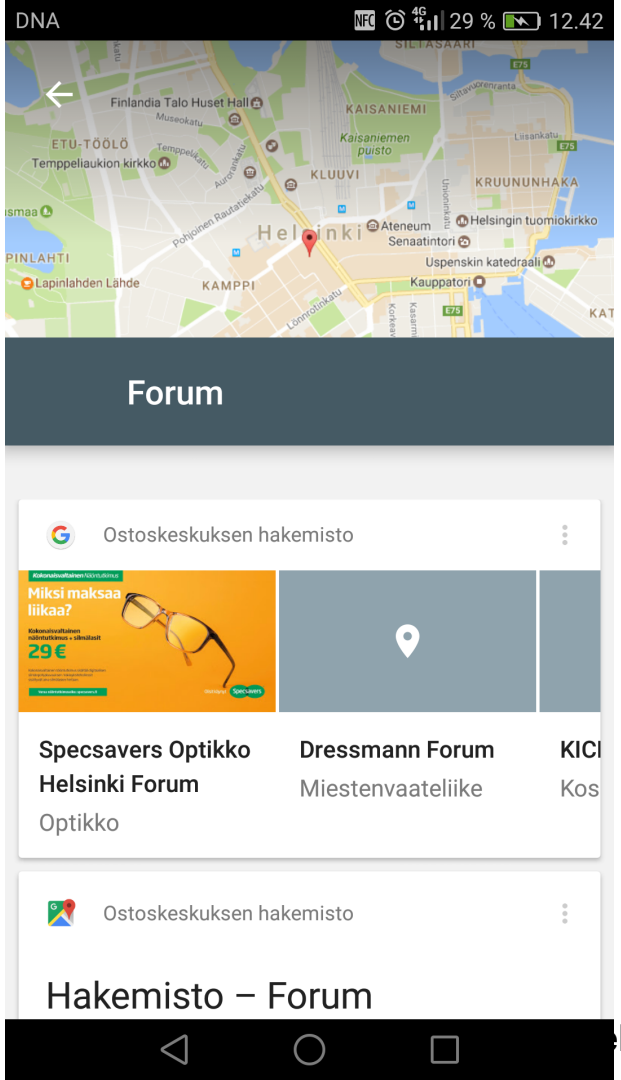
- geospatial big data and location data can give benefit to
  - save fuel
  - Save time
  - Increase revenues
  - Plan urban areas
  - Improve health care
  - Better marketing

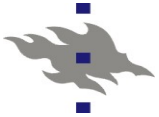


Applications

- Forum example (from mobile screenshot)


Ilmoitukset
Pikavalikot
Aika
12.36
13. HELMIKUUTA
maanantai
Arkistoitu.
KUMOA
12....
12....
Forum
Ostoskeskuksen hakemisto
DNA


DNA
29 %
12.42
Finlandia Talo Huset Hall
KAISANIEMI
ETU-TÖÖLÖ
Temppeliaukion kirkko
KLUUVI
KRUUNUNHAKA
Helsinki
Ateneum
Senaatintori
Helsingin tuomiokirkko
Uspenskin katedraali
Kauppatori
PINLAHTI
Lapinlahden Lähde
KAMPPI
Forum
Ostoskeskuksen hakemisto
Miksi maksaa liikaa?
29 €
Specsavers Optikko Helsinki Forum
Optikko
Dressmann Forum
Miestenvaateliike
Ostoskeskuksen hakemisto
Hakemisto – Forum

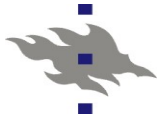


Case Netflix

- Netflix generated the data on the top-50rentals in 2009 in each zip code
  - Such patterns are very useful for recommending movies to their users

Marketers & brands want to talk to you personally and big data can help them to make you feel

”Oh how nice from you to think of me!”



Applying Big Data

1) Collecting data

Amount of data is increasing

- McKinsey Global Institute says that the pool of personal location data was in the level of 1 PB in 2009 and is growing at a rate of 20% per year
- + + + + New data from internet of things data stored to internal archives
- in Google, about 25 PB of data is being generated per day, and a significant portion of the data falls into the realm of spatio-temporal data

Rise of mobile

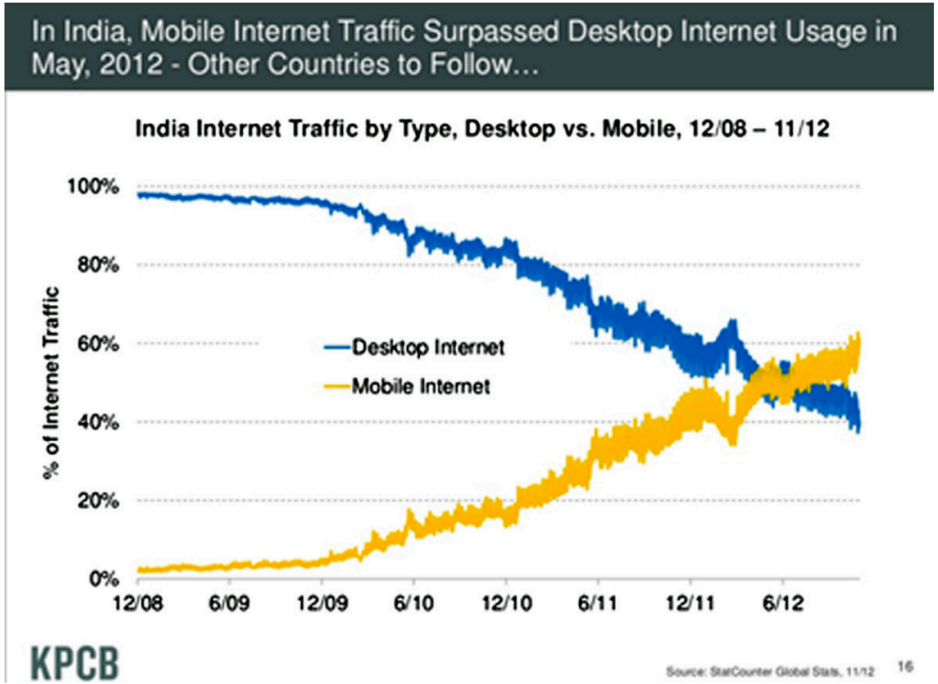


Fig. 1. Mobile internet traffic in India [3].

Not only hype anymore

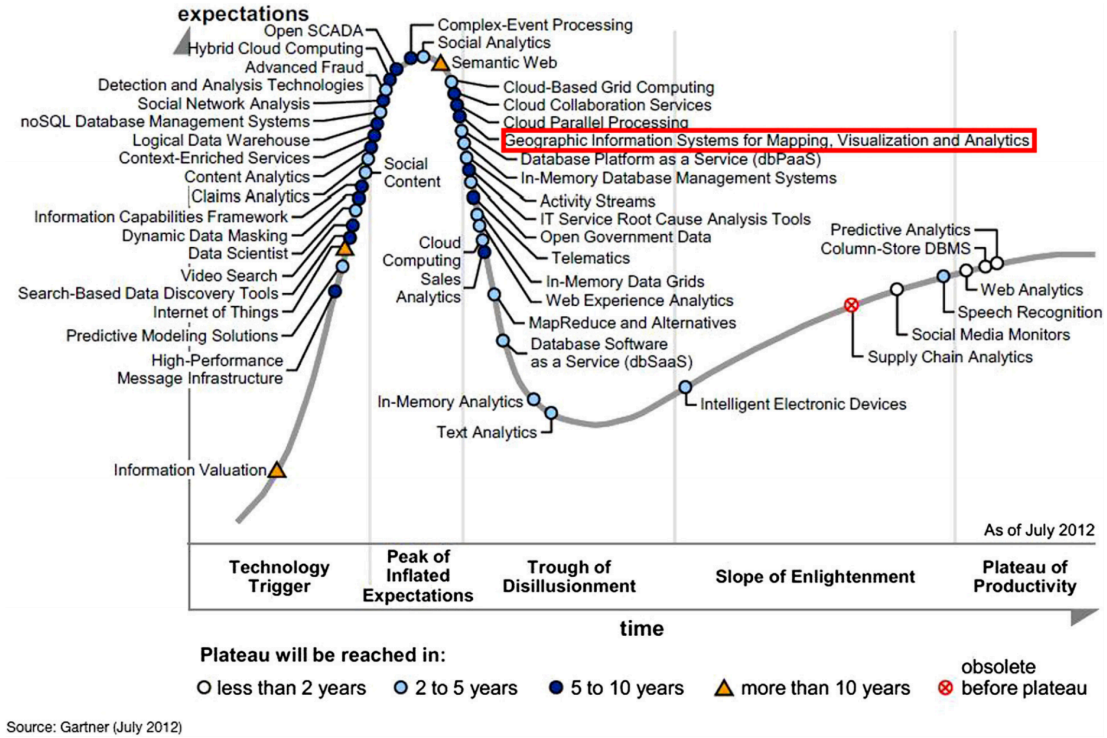


Fig. 2. Gartner's hype cycle (source: Gartner) [10].

Big Data is important for marketing

Example Location data, subset of Big Dta:

“Location targeting is holy grail for marketers.”

- Sir Martin Sorrell, the CEO of WPP Group

Applications

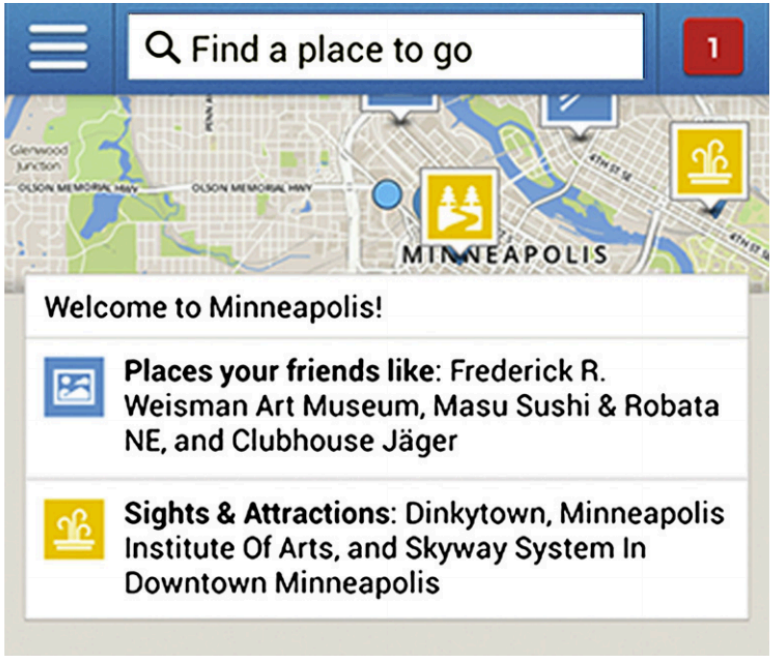
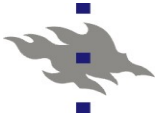


Fig. 6. Foursquare Android app with recommendations.

How to Apply big data for marketing

2) Analysing



Prediction

- The future location of a human can be predicted by analyzing his/her records of previous traces
- Analysis of human mobility can boost many applications ranging from epidemic modeling to traffic prediction and urban planning.

Case Walmart and Hurricane

- The executives at Walmart decided to adopt one of big data technologies—predictive analytics
- Linda M. Dillman, Walmart's chief information officer, asked her staff to predict what would happen soon based on what had happened when Hurricane Charley landed several weeks ago.
- By analyzing the transaction records stored in Walmart's data warehouse,the company could predict which items were bought justbefore or after an event (i.e., a hurricane) at a specific region.

Case Walmart and Hurricane

- People who had lived in Florida's Atlantic coast did not increasingly buy some products directly related to hurricanes, e.g., water andflash lights.
- Surprisingly, strawberry PopTarts increased in sales, by seven times compared with their usual sale rate, just before ahurricane. In addition, the top-selling item immediately before the hurricane was instead …

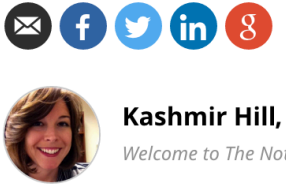
Top selling product before Hurricane was beer

Ilkka Lavas, Helsingin Yliopisto, Tietojenkäsittelytieteen laitos

FEB 16, 2012 @ 11:02 AM 3,112,802 VIEWS

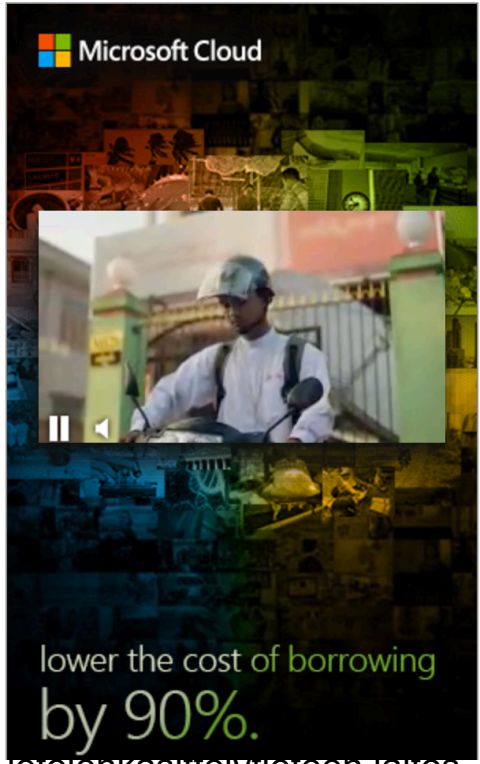
Peek Inside The Billionaire Portfolio Of Trump's Key Advisor

How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did



Kashmir Hill, FORBES STAFF
Welcome to The Not-So Private Parts where technology & privacy collide FULL BIO

Every time you go shopping, you share intimate details about your consumption patterns with retailers. And many of those retailers are studying those details to figure out what you like, what you need, and which coupons are most likely to make you happy. Target TGT +1.30%, for example, has figured out how to data-mine its way into your womb, to figure out whether you have a baby on the way long before you need to start buying diapers.



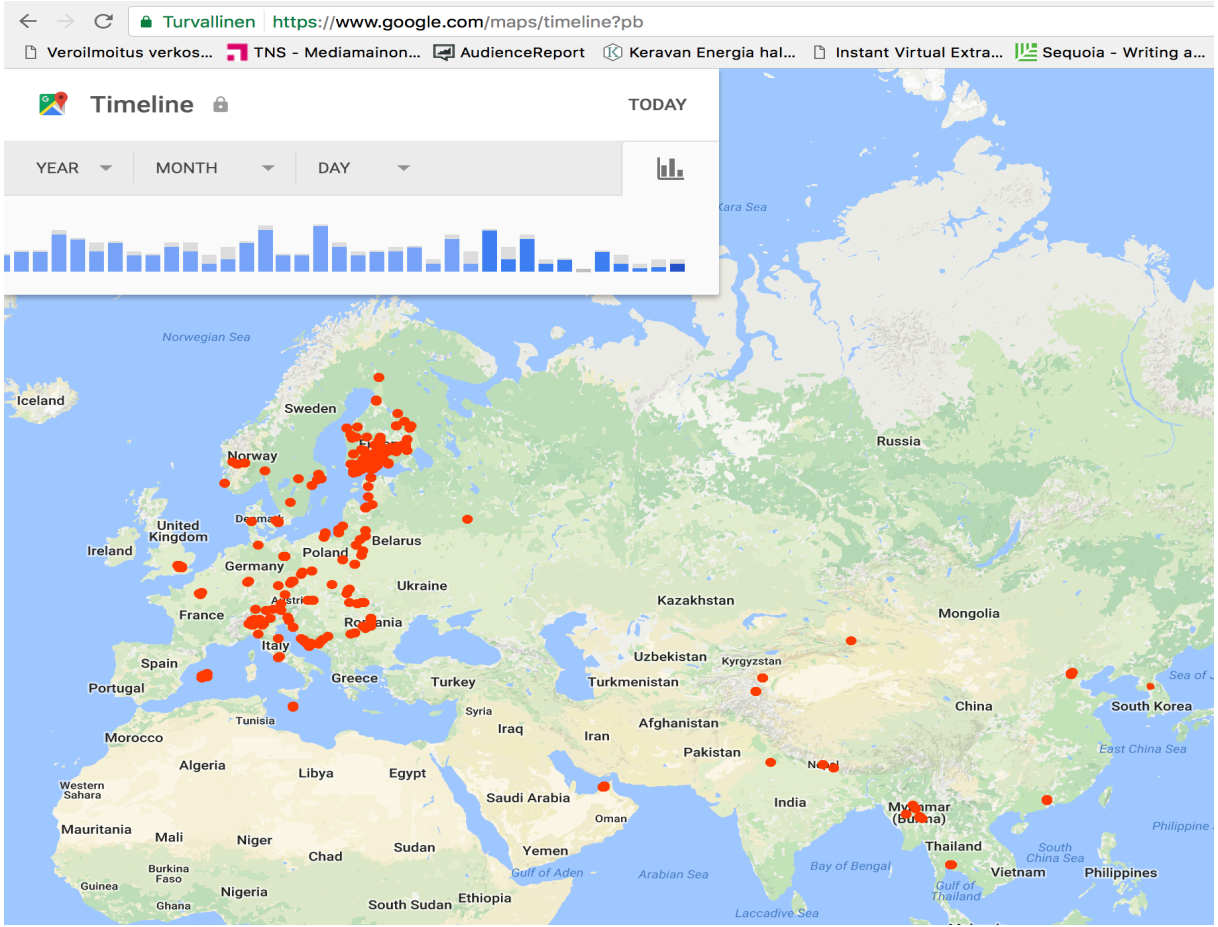
Ilkka Lavas, Helsingin Yliopisto, Tietojenkasittelytieteen laitos

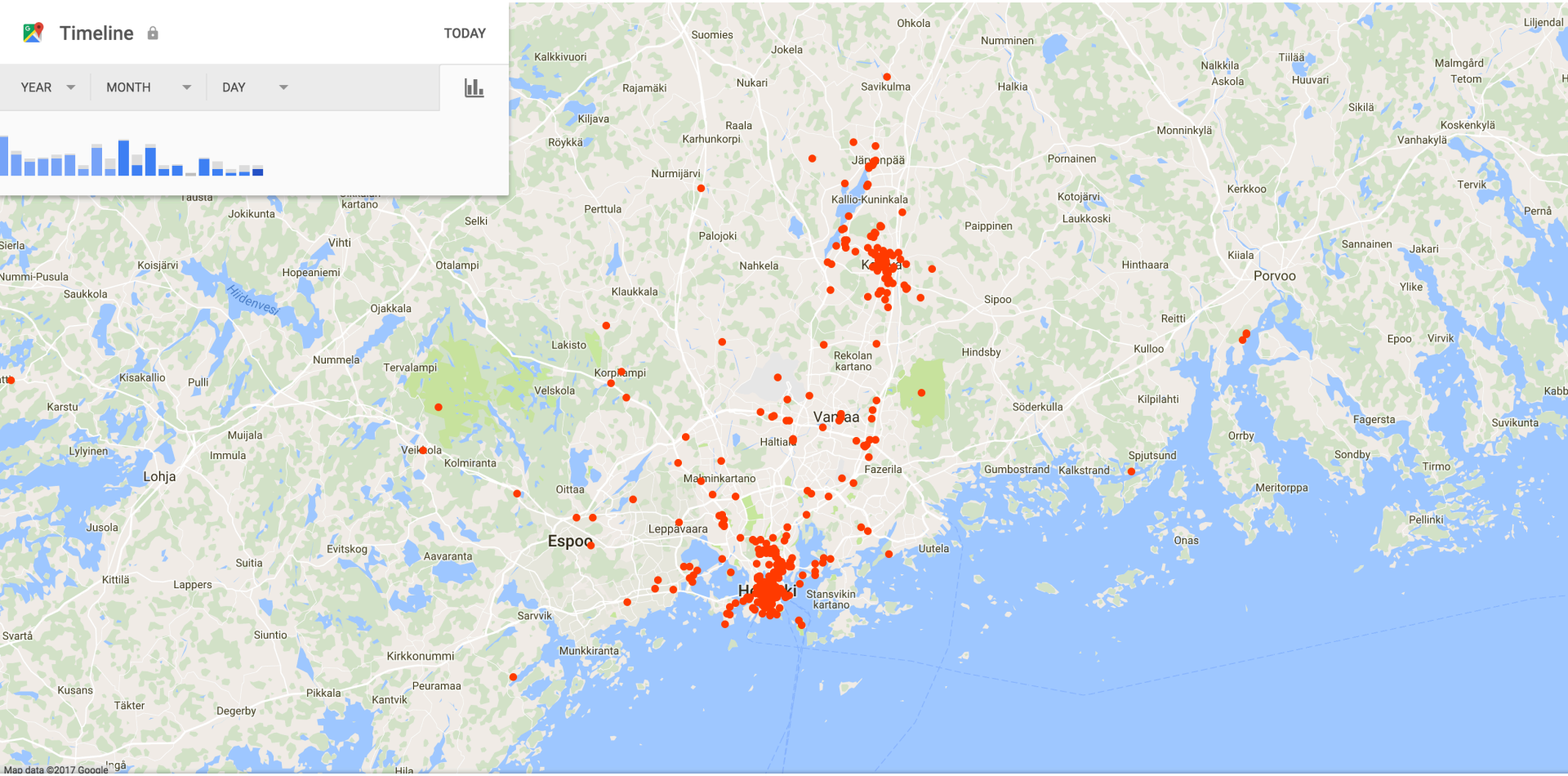
How we predict relevancy: example nearbyness

- Direct marketing.
  - This field is supported by Tobler's first law of geography

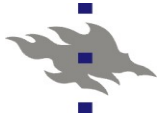
    "Everything is related to everything else,
    but near things are more related than distant things."

- suggestingthe services or stores close to the current location of a usershould be more effective than suggesting those far-away


Turvallinen
https://www.google.com/maps/timeline?pb
Veroilmoitus verkos...
TNS - Mediamainon...
AudienceReport
Keravan Energia hal...
Instant Virtual Extra...
Sequoia - Writing a...
Timeline
TODAY
YEAR
MONTH
DAY
Norwegian Sea
Iceland
Sweden
Norway
Russia
United Kingdom
Ireland
Denmark
Belarus
Poland
Germany
Ukraine
France
Kazakhstan
Mongolia
Italy
Spain
Portugal
Greece
Turkey
Uzbekistan
Kyrgyzstan
Turkmenistan
China
South Korea
Tunisia
Syria
Iraq
Iran
Afghanistan
Pakistan
Morocco
Algeria
Libya
Egypt
Saudi Arabia
India
Western Sahara
Mauritania
Mali
Niger
Chad
Sudan
Yemen
Oman
Thailand
Vietnam
Philippines
Burkina Faso
Guinea
Ghana
Nigeria
South Sudan
Ethiopia
Gulf of Aden
Arabian Sea
Bay of Bengal
Gulf of Thailand
East China Sea
South China Sea
Laccadive Sea



Ilkka Lavas, Helsingin Yliopisto, Tietojenkäsittelytieteen laitos

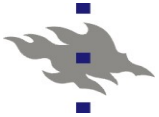


How data is collected?

```
Ilkka-MacBook-Pro-Iso:~ ilkka$ netstat
Active Internet connections
Proto Recv-Q Send-Q  Local Address          Foreign Address        (state)
tcp4       0      0  192.168.43.30.55445    api-fra.smoot.ap.https ESTABLISHED
tcp4       0      0  192.168.43.30.55442    46.189.56.88.http      CLOSE_WAIT
tcp6       0      0  dtb43h7ql6z7gg5v.55441 2001:14b8:1800:4.http  ESTABLISHED
tcp4       0      0  192.168.43.30.55439    40.76.1.176.https      ESTABLISHED
tcp4       0      0  192.168.43.30.55399    a23-43-140-72.de.http  ESTABLISHED
tcp4       0      0  192.168.43.30.55388    62-165-164-203.c.http  ESTABLISHED
tcp4       0      0  192.168.43.30.55387    62-165-164-203.c.http  ESTABLISHED
tcp4       0      0  192.168.43.30.55386    62-165-164-203.c.http  ESTABLISHED
tcp4       0      0  192.168.43.30.55385    62-183-170-34.co.http  ESTABLISHED
tcp4       0      0  192.168.43.30.55383    62-165-164-203.c.http  ESTABLISHED
tcp4       0      0  192.168.43.30.55382    62-165-164-203.c.http  ESTABLISHED
tcp4       0      0  192.168.43.30.55380    62-165-164-203.c.http  ESTABLISHED
tcp4       0      0  192.168.43.30.55350    62-183-170-19.co.http  ESTABLISHED
tcp6       0      0  dtb43h7ql6z7gg5v.55345 arn09s10-in-x03.https ESTABLISHED
tcp6       0      0  dtb43h7ql6z7gg5v.55343 arn09s11-in-x01.http  ESTABLISHED
tcp4       0      0  192.168.43.30.55321    arn06s07-in-f6.1.https ESTABLISHED
tcp6       0      0  dtb43h7ql6z7gg5v.55320 arn09s10-in-x02.http  ESTABLISHED
tcp4       0      0  192.168.43.30.55315    a23-43-140-72.de.http  ESTABLISHED
tcp4       0      0  192.168.43.30.55314    a23-43-140-72.de.http  ESTABLISHED
tcp4       0      0  192.168.43.30.55313    a23-43-140-72.de.http  ESTABLISHED
tcp6       0      0  dtb43h7ql6z7gg5v.55311 pbtng-new-nuggad.http  ESTABLISHED
tcp4       0      0  192.168.43.30.55310    112.245.178.107.https ESTABLISHED
tcp4       0      0  192.168.43.30.55301    a23-43-140-72.de.http  ESTABLISHED
tcp4       0      0  192.168.43.30.55298    62-183-170-26.co.http  ESTABLISHED
tcp4       0      0  192.168.43.30.55294    112.245.178.107.https ESTABLISHED
tcp4       0      0  192.168.43.30.55288    104.156.85.134.https   ESTABLISHED
tcp4       0      0  192.168.43.30.55285    151.101.84.134.http    ESTABLISHED
tcp4       0      0  192.168.43.30.55284    151.101.84.134.http    ESTABLISHED
tcp6       0      0  dtb43h7ql6z7gg5v.55273 arn09s11-in-x08.http  ESTABLISHED
tcp4       0      0  192.168.43.30.55272    a23-43-140-72.de.http  ESTABLISHED
tcp4       0      0  192.168.43.30.55271    a23-43-140-72.de.http  ESTABLISHED
tcp4       0      0  192.168.43.30.55268    a23-43-140-72.de.http  ESTABLISHED
tcp4       0      0  192.168.43.30.55267    a23-43-140-72.de.http  ESTABLISHED
tcp6       0      0  dtb43h7ql6z7gg5v.55266 cdn-vip47.hard.w.http  ESTABLISHED
tcp6       0      0  dtb43h7ql6z7gg5v.55265 cdn-vip47.hard.w.http  ESTABLISHED
tcp4       0      0  192.168.43.30.55264    62-183-170-18.co.http  ESTABLISHED
tcp4       0      0  192.168.43.30.55263    62-183-170-18.co.http  ESTABLISHED
tcp6       0      0  dtb43h7ql6z7gg5v.55260 cdn-vip47.hard.w.http  ESTABLISHED
tcp6       0      0  dtb43h7ql6z7gg5v.55259 cdn-vip47.hard.w.http  ESTABLISHED
tcp6       0      0  dtb43h7ql6z7gg5v.55258 cdn-vip47.hard.w.http  ESTABLISHED
tcp6       0      0  dtb43h7ql6z7gg5v.55257 cdn-vip47.hard.w.http  ESTABLISHED
tcp6       0      0  dtb43h7ql6z7gg5v.55256 cdn-vip47.hard.w.http  ESTABLISHED
tcp6       0      0  dtb43h7ql6z7gg5v.55255 li-in-x79.1e100.http   ESTABLISHED
tcp6       0      0  dtb43h7ql6z7gg5v.55253 cdn-vip47.hard.w.http  ESTABLISHED
tcp6       0      0  dtb43h7ql6z7gg5v.55252 cdn-vip47.hard.w.http  ESTABLISHED
tcp6       0      0  dtb43h7ql6z7gg5v.55251 cdn-vip47.hard.w.http  ESTABLISHED
tcp6       0      0  dtb43h7ql6z7gg5v.55250 cdn-vip47.hard.w.http  ESTABLISHED
tcp6       0      0  dtb43h7ql6z7gg5v.55249 cdn-vip47.hard.w.http  ESTABLISHED
tcp6       0      0  dtb43h7ql6z7gg5v.55248 cdn-vip47.hard.w.http  ESTABLISHED
tcp6       0      0  dtb43h7ql6z7gg5v.55245 cdn-vip17.hard.w.http  ESTABLISHED
tcp4       0      0  192.168.43.30.55240    104.41.207.73.https    FIN_WAIT_1
tcp4      31      0  192.168.43.30.55205    162.125.66.3.https     CLOSE_WAIT
tcp4       0      0  192.168.43.30.55196    a23-60-24-45.dep.https ESTABLISHED
tcp4       0      0  192.168.43.30.55195    a23-60-24-45.dep.https ESTABLISHED
tcp6       0      0  dtb43h7ql6z7gg5v.55184 2001:14b8:1800:4.https ESTABLISHED
tcp6       0      0  dtb43h7ql6z7gg5v.55149 xx-fbcdn6-shv-01.https ESTABLISHED
tcp4       0    794  192.168.43.30.55120    104.215.144.200.https  FIN_WAIT_1
tcp4       0      0  192.168.43.30.55070    138.68.120.27.https    LAST_ACK
tcp4      31      0  192.168.43.30.55059    d.v.dropbox.com.https  CLOSE_WAIT
tcp4       0      0  192.168.43.30.54675    162.125.18.133.https   ESTABLISHED
tcp4       0      0  192.168.43.30.54654    162.125.18.133.https   ESTABLISHED
tcp6       0      0  dtb43h7ql6z7gg5v.54568 arn09s10-in-x0e.https ESTABLISHED
tcp6       0      0  dtb43h7ql6z7gg5v.54562 arn09s11-in-x0e.https ESTABLISHED
tcp4       0      0  192.168.43.30.54417    ec2-52-45-60-77.https  CLOSE_WAIT
tcp4       0      0  192.168.43.30.54376    do-5.lastpass.co.https ESTABLISHED
```

```
Active Multipath Internet connections
Proto/ID  Flags      Local Address          Foreign Address        (state)
Active LOCAL (UNIX) domain sockets
Address          Type   Recv-Q Send-Q            Inode             Conn             Refs          Nextref Addr
 bbb9645ecdca872d stream      0      0                0 bbb9645ecdca8c5                0                0 /var/run/mDNSResponder
 bbb9645ecdca8c5 stream      0      0                0 bbb9645ecdc872d                0                0
 bbb9645ecdcb3b5 stream      0      0                0 bbb9645ecdcacad                0                0 /var/run/mDNSResponder
 bbb9645ecdcacad stream      0      0                0 bbb9645ecdcb3b5                0                0
 bbb9645ecdc8985 stream      0      0                0 bbb9645ecdc92e5                0                0
 bbb9645ecdc92e5 stream      0      0                0 bbb9645ecdc8985                0                0
 bbb9645ecdca66d stream      0      0                0 bbb9645ecdc9e9d                0                0
 bbb9645ecdc9e9d stream      0      0                0 bbb9645ecdca66d                0                0
 bbb9645ecdc921d stream      0      0                0 bbb9645ecdcb15d                0                0
 bbb9645ecdcb15d stream      0      0                0 bbb9645ecdc921d                0                0
 bbb9645e7a10475 stream      0      0                0 bbb9645e7a11cad                0                0 /var/run/mDNSResponder
 bbb9645e7a11cad stream      0      0                0 bbb9645e7a10475                0                0
 bbb9645e7a109ed stream      0      0                0 bbb9645e7a11735                0                0
 bbb9645e7a11735 stream      0      0                0 bbb9645e7a109ed                0                0
 bbb9645e7a10b7d stream      0      0                0 bbb9645e7a1198d                0                0
 bbb9645e7a1198d stream      0      0                0 bbb9645e7a10b7d                0                0
 bbb9645d3d4c605 stream      0      0                0 bbb9645d3d4cd0d                0                0 /var/run/mDNSResponder
 bbb9645d3d4cd0d stream      0      0                0 bbb9645d3d4c605                0                0
 bbb9645ecdc87f5 stream      0      0                0 bbb9645ecdc908d                0                0
 bbb9645ecdc908d stream      0      0                0 bbb9645ecdc87f5                0                0
 bbb9645ef2bb2ed stream      0      0                0 bbb9645ef2b9d0d                0                0
 bbb9645ef2b9d0d stream      0      0                0 bbb9645ef2bb2ed                0                0
 bbb9645ef2b9c45 stream      0      0                0                0                0                0
 bbb9645ef2ba1bd stream      0      0                0 bbb9645ef2b859d                0                0
 bbb9645ef2b859d stream      0      0                0 bbb9645ef2ba1bd                0                0
 bbb9645ef2bacad stream      0      0                0 bbb9645ef2b9605                0                0
 bbb9645ef2b9605 stream      0      0                0 bbb9645ef2bacad                0                0
 bbb9645ef2b99ed stream      0      0                0 bbb9645ef2babe5                0                0
 bbb9645ef2babe5 stream      0      0                0 bbb9645ef2b99ed                0                0
 bbb9645ef2bb3b5 stream      0      0                0 bbb9645ef2b9795                0                0
 bbb9645ef2b9795 stream      0      0                0 bbb9645ef2bb3b5                0                0
 bbb9645ef2bb895 stream      0      0                0 bbb9645ef2b9f65                0                0
 bbb9645ef2b9f65 stream      0      0                0 bbb9645ef2bb895                0                0
 bbb9645ef2b8efd stream      0      0                0 bbb9645ef2ba4dd                0                0
 bbb9645ef2ba4dd stream      0      0                0 bbb9645ef2b8efd                0                0
 bbb9645e7a11e3d stream      0      0                0 bbb9645e7a10605                0                0
 bbb9645e7a10605 stream      0      0                0 bbb9645e7a11e3d                0                0
 bbb9645e7a11b1d stream      0      0                0 bbb9645e7a10795                0                0
 bbb9645e7a10795 stream      0      0                0 bbb9645e7a11b1d                0                0
 bbb9645e7a11d75 stream      0      0                0 bbb9645e7a11a55                0                0
 bbb9645e7a11a55 stream      0      0                0 bbb9645e7a11d75                0                0
 bbb9645e7a1085d stream      0      0                0 bbb9645e7a122ed                0                0
 bbb9645e7a122ed stream      0      0                0 bbb9645e7a1085d                0                0
 bbb9645e7a10dd5 stream      0      0                0 bbb9645e7a117fd                0                0
 bbb9645e7a117fd stream      0      0                0 bbb9645e7a10dd5                0                0
 bbb9645e7a12225 stream      0      0                0 bbb9645e7a0fa4d                0                0
 bbb9645e7a0fa4d stream      0      0                0 bbb9645e7a12225                0                0
 bbb9645e7a102e5 stream      0      0                0 bbb9645e7a115a5                0                0
 bbb9645e7a115a5 stream      0      0                0 bbb9645e7a102e5                0                0
 bbb9645e7a0fca5 stream      0      0                0 bbb9645e7a10155                0                0
 bbb9645e7a10155 stream      0      0                0 bbb9645e7a0fca5                0                0
 bbb9645e7a0f7f5 stream      0      0                0 bbb9645e7a10c45                0                0
 bbb9645e7a10c45 stream      0      0                0 bbb9645e7a0f7f5                0                0
 bbb9645e7a0f59d stream      0      0                0 bbb9645e7a11f05                0                0
 bbb9645e7a11f05 stream      0      0                0 bbb9645e7a0f59d                0                0
```

```
kevt       0      0      1      1     11
kevt       0      0      1      6      1
kevt       0      0      1      6      1
kevt       0      0      1      1      7
kevt       0      0      1      1      1
kevt       0      0      1      6      1
kevt       0      0      1      6      1
kevt       0      0      1      6      1
kevt       0      0      1      6      1
kevt       0      0      1      6      1
kevt       0      0      1      6      1
kevt       0      0      1      1      2
kevt       0      0      1      6      1
kevt       0      0      1      1      2
kevt       0      0      1      6      1
kevt       0      0      1      1      0
Active kernel control sockets
Proto Recv-Q Send-Q   unit     id name
kctl       0      0      1      5 com.apple.net.necp_control
kctl       0      0      2      5 com.apple.net.necp_control
kctl       0      0      3      5 com.apple.net.necp_control
kctl       0      0      1      6 com.apple.net.netagent
kctl       0      0      2      6 com.apple.net.netagent
kctl       0      0      3      6 com.apple.net.netagent
kctl       0      0      4      6 com.apple.net.netagent
kctl       0      0      5      6 com.apple.net.netagent
kctl       0      0      6      6 com.apple.net.netagent
kctl       0      0      7      6 com.apple.net.netagent
kctl       0      0      8      6 com.apple.net.netagent
kctl       0      0      9      6 com.apple.net.netagent
kctl       0      0     10      6 com.apple.net.netagent
kctl       0      0      1      7 com.apple.net.utun_control
kctl       0      0      1      9 com.apple.netsrc
kctl       0      0      2      9 com.apple.netsrc
kctl       0      0      3      9 com.apple.netsrc
kctl       0      0      4      9 com.apple.netsrc
kctl       0      0      5      9 com.apple.netsrc
kctl       0      0      6      9 com.apple.netsrc
kctl       0      0      7      9 com.apple.netsrc
kctl       0      0      8      9 com.apple.netsrc
kctl       0      0      9      9 com.apple.netsrc
kctl       0      0     10      9 com.apple.netsrc
kctl       0      0     11      9 com.apple.netsrc
kctl       0      0     12      9 com.apple.netsrc
kctl       0      0     13      9 com.apple.netsrc
kctl       0      0     14      9 com.apple.netsrc
kctl       0      0     15      9 com.apple.netsrc
kctl       0      0     16      9 com.apple.netsrc
kctl       0      0     17      9 com.apple.netsrc
kctl       0      0     18      9 com.apple.netsrc
kctl       0      0     19      9 com.apple.netsrc
kctl       0      0     20      9 com.apple.netsrc
kctl       0      0     21      9 com.apple.netsrc
kctl       0      0     22      9 com.apple.netsrc
kctl       0      0     23      9 com.apple.netsrc
kctl       0      0     24      9 com.apple.netsrc
kctl       0      0     25      9 com.apple.netsrc
kctl       0      0     26      9 com.apple.netsrc
kctl       0      0     27      9 com.apple.netsrc
kctl       0      0     28      9 com.apple.netsrc
kctl       0      0     29      9 com.apple.netsrc
kctl       0      0     30      9 com.apple.netsrc
kctl       0      0      1     10 com.apple.network.statistics
kctl       0      0      2     10 com.apple.network.statistics
kctl       0      0      3     10 com.apple.network.statistics
kctl       0      0      4     10 com.apple.network.statistics
kctl       0      0      5     10 com.apple.network.statistics
kctl       0      0      1     12 com.apple.network.advisory
```



How data is collected?

```
Ilkka-MacBook-Pro-Iso:~ ilkka$ netstat
Active Internet connections
Proto Recv-Q Send-Q  Local Address          Foreign Address        (state)
tcp4       0      0  192.168.43.30.55445    api-fra.smoot.ap.https ESTABLISHED
tcp4       0      0  192.168.43.30.55442    46.189.56.88.http      CLOSE_WAIT
tcp6       0      0  dtb43h7ql6z7gg5v.55441 2001:14b8:1800:4.https ESTABLISHED
tcp4       0      0  192.168.43.30.55439    40.76.1.176.https      ESTABLISHED
tcp4       0      0  192.168.43.30.55399    a23-43-140-72.de.http  ESTABLISHED
tcp4       0      0  192.168.43.30.55388    62-165-164-203.c.http  ESTABLISHED
tcp4       0      0  192.168.43.30.55387    62-165-164-203.c.http  ESTABLISHED
tcp4       0      0  192.168.43.30.55386    62-165-164-203.c.http  ESTABLISHED
tcp4       0      0  192.168.43.30.55385    62-183-170-34.co.http  ESTABLISHED
tcp4       0      0  192.168.43.30.55383    62-165-164-203.c.http  ESTABLISHED
tcp4       0      0  192.168.43.30.55382    62-165-164-203.c.http  ESTABLISHED
tcp4       0      0  192.
tcp4       0      0  192.
tcp6       0      0  dtb4
tcp6       0      0  dtb4
tcp4       0      0  192.
tcp6       0      0  dtb4
tcp4       0      0  192.
tcp4       0      0  192.
tcp4       0      0  192.
tcp6       0      0  dtb4
tcp4       0      0  192.
tcp4       0      0  192.
tcp4       0      0  192.
tcp4       0      0  192.
tcp4       0      0  192.
tcp4       0      0  192.
tcp4       0      0  192.
tcp4       0      0  192.
tcp6       0      0  dtb4
tcp6       0      0  dtb4
tcp4       0      0  192.
tcp4       0      0  192.
tcp6       0      0  dtb4
tcp6       0      0  dtb4
tcp6       0      0  dtb4
tcp6       0      0  dtb4
tcp6       0      0  dtb4
tcp6       0      0  dtb4
tcp6       0      0  dtb4
tcp6       0      0  dtb4
tcp6       0      0  dtb4
tcp6       0      0  dtb4
tcp6       0      0  dtb4
tcp6       0      0  dtb43h7ql6z7gg5v.55249 cdn-vip47.hard.w.http  ESTABLISHED
tcp6       0      0  dtb43h7ql6z7gg5v.55248 cdn-vip47.hard.w.http  ESTABLISHED
tcp6       0      0  dtb43h7ql6z7gg5v.55245 cdn-vip17.hard.w.http  ESTABLISHED
tcp4       0      0  192.168.43.30.55240    104.41.207.73.https    FIN_WAIT_1
tcp4      31      0  192.168.43.30.55205    162.125.66.3.https     CLOSE_WAIT
tcp4       0      0  192.168.43.30.55196    a23-60-24-45.dep.https ESTABLISHED
tcp4       0      0  192.168.43.30.55195    a23-60-24-45.dep.https ESTABLISHED
tcp6       0      0  dtb43h7ql6z7gg5v.55184 2001:14b8:1800:4.https ESTABLISHED
tcp6       0      0  dtb43h7ql6z7gg5v.55149 xx-fbcdn6-shv-01.https ESTABLISHED
tcp4       0    794  192.168.43.30.55120    104.215.146.200.https  FIN_WAIT_1
tcp4       0      0  192.168.43.30.55070    138.68.120.27.https    LAST_ACK
tcp4      31      0  192.168.43.30.55059    d.v.dropbox.com.https  CLOSE_WAIT
tcp4       0      0  192.168.43.30.54675    162.125.18.133.https   ESTABLISHED
tcp4       0      0  192.168.43.30.54654    162.125.18.133.https   ESTABLISHED
tcp6       0      0  dtb43h7ql6z7gg5v.54568 arn09s10-in-x0e..https ESTABLISHED
tcp6       0      0  dtb43h7ql6z7gg5v.54562 arn09s11-in-x0e..https ESTABLISHED
tcp4       0      0  192.168.43.30.54417    ec2-52-45-60-77..https CLOSE_WAIT
tcp4       0      0  192.168.43.30.54376    do-5.lastpass.co.https ESTABLISHED
```

```
Active Multipath Internet connections
Proto/ID  Flags      Local Address          Foreign Address        (state)
Active LOCAL (UNIX) domain sockets
Address          Type   Recv-Q Send-Q            Inode             Conn             Refs          Nextref Addr
 bbb9645ecdc872d stream      0      0                0  bbb9645ecdca8c5                0                0 /var/run/mDNSResponder
 bbb9645ecdca8c5 stream      0      0                0  bbb9645ecdc872d                0                0
 bbb9645ecdcb3b5 stream      0      0                0  bbb9645ecdcacad                0                0 /var/run/mDNSResponder
 bbb9645ecdcacad stream      0      0                0  bbb9645ecdcb3b5                0                0
 bbb9645ecdc8985 stream      0      0                0  bbb9645ecdc92e5                0                0
 bbb9645ecdc92e5 stream      0      0                0  bbb9645ecdc8985                0                0
 bbb9645ecdca66d stream      0      0                0  bbb9645ecdc9e9d                0                0
 bbb9645ecdc9e9d stream      0      0                0  bbb9645ecdca66d                0                0
 bbb9645ecdc921d stream      0      0                0  bbb9645ecdcb15d                0                0
 bbb9645ecdcb15d stream      0      0                0  bbb9645ecdc921d                0                0
```

```
kevt       0      0      1      1     11
kevt       0      0      1      6      1
kevt       0      0      1      6      1
kevt       0      0      1      1      7
kevt       0      0      1      1      1
kevt       0      0      1      6      1
kevt       0      0      1      6      1
kevt       0      0      1      6      1
kevt       0      0      1      6      1
kevt       0      0      1      6      1
kevt       0      0      1      6      1
kevt       0      0      1      1      2
kevt       0      0      1      6      1
kevt       0      0      1      1      2
kevt       0      0      1      6      1
kevt       0      0      1      1      0
Active kernel control sockets
Proto Recv-Q Send-Q   unit     id name
kctl       0      0      1      5 com.apple.net.necp_control
kctl       0      0      2      5 com.apple.net.necp_control
kctl       0      0      3      5 com.apple.net.necp_control
kctl       0      0      1      6 com.apple.net.netagent
kctl       0      0      2      6 com.apple.net.netagent
kctl       0      0      3      6 com.apple.net.netagent
kctl       0      0      4      6 com.apple.net.netagent
kctl       0      0      5      6 com.apple.net.netagent
```

We all leave traces all the time from our laptops and mobiles

Spooky or good service?

```
bbb9645e7a12225 stream      0      0                0  bbb9645e7a0fa4d                0                0
bbb9645e7a0fa4d stream      0      0                0  bbb9645e7a12225                0                0
bbb9645e7a102e5 stream      0      0                0  bbb9645e7a115a5                0                0
bbb9645e7a115a5 stream      0      0                0  bbb9645e7a102e5                0                0
bbb9645e7a0fca5 stream      0      0                0  bbb9645e7a10155                0                0
bbb9645e7a10155 stream      0      0                0  bbb9645e7a0fca5                0                0
bbb9645e7a0f7f5 stream      0      0                0  bbb9645e7a10c45                0                0
bbb9645e7a10c45 stream      0      0                0  bbb9645e7a0f7f5                0                0
bbb9645e7a0f59d stream      0      0                0  bbb9645e7a11f05                0                0
bbb9645e7a11f05 stream      0      0                0  bbb9645e7a0f59d                0                0
```

```
kctl       0      0     26      9 com.apple.netsrc
kctl       0      0     27      9 com.apple.netsrc
kctl       0      0     28      9 com.apple.netsrc
kctl       0      0     29      9 com.apple.netsrc
kctl       0      0     30      9 com.apple.netsrc
kctl       0      0      1     10 com.apple.network.statistics
kctl       0      0      2     10 com.apple.network.statistics
kctl       0      0      3     10 com.apple.network.statistics
kctl       0      0      4     10 com.apple.network.statistics
kctl       0      0      5     10 com.apple.network.statistics
kctl       0      0      1     12 com.apple.network.advisory
```

Ilkka Lavas, Helsingin Yliopisto, Tietojenkäsittelytieteen laitos

How geospatial data is saved?

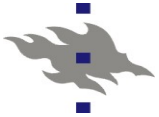
- 3 forms
  - Raster data
  - vector data
  - graph data

Data types 1/3: Raster data

- include geoimages typically obtained by unmanned aerial vehicles, security cameras, and satellites. Recently, the military is collecting huge amounts of raster data by utilizing drones, and the satellites keep providing us with the remote sensing data of the Earth
- The raster data is being provided by digital map services, e.g., Google Earth. Data analysts extract the tracks of moving objects or useful features from these raster data.
- Representative use cases include life pattern mining and change detection.

Data types 2/3: Vector data

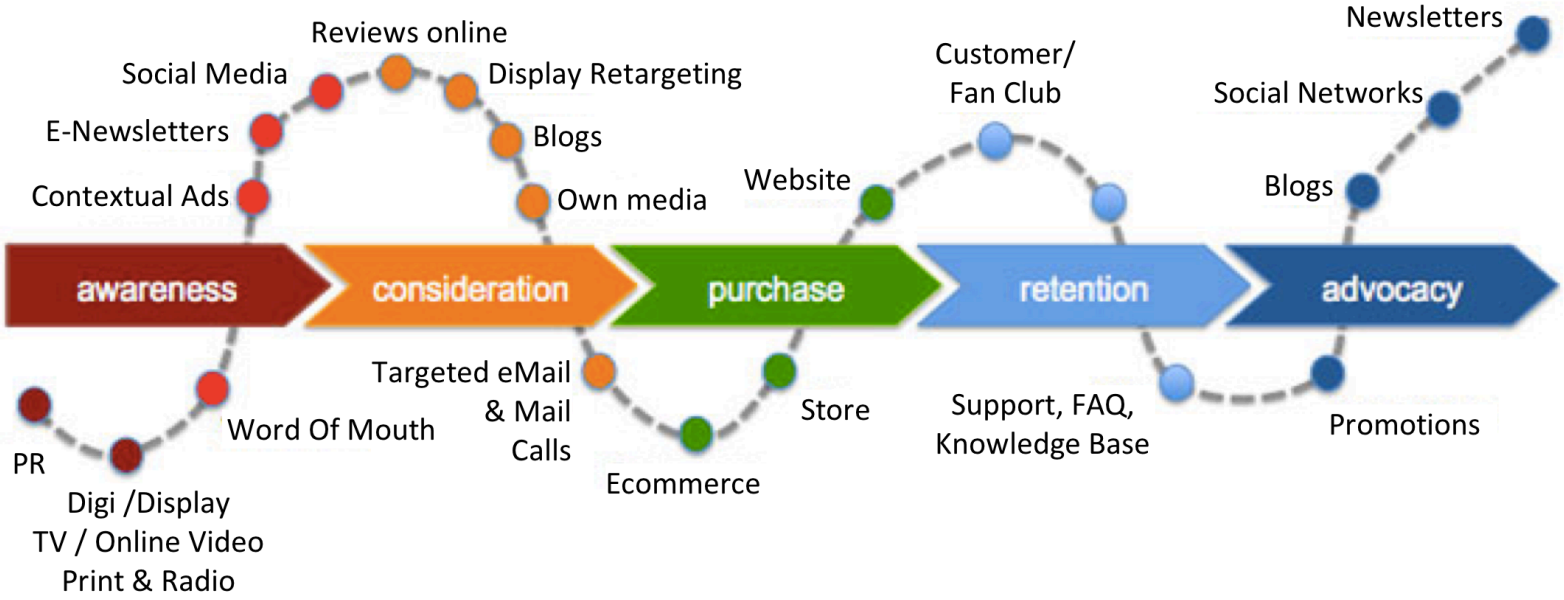
- vector data
  - consists of points, lines, and polygons.
  - For example, points can be collected through checkin's on
    - Foursquare /
    - Facebook /
    - Google Maps, and
    - lines and polygons correspond to roads in OpenStreetMap



Data types 3/3: Graph Data

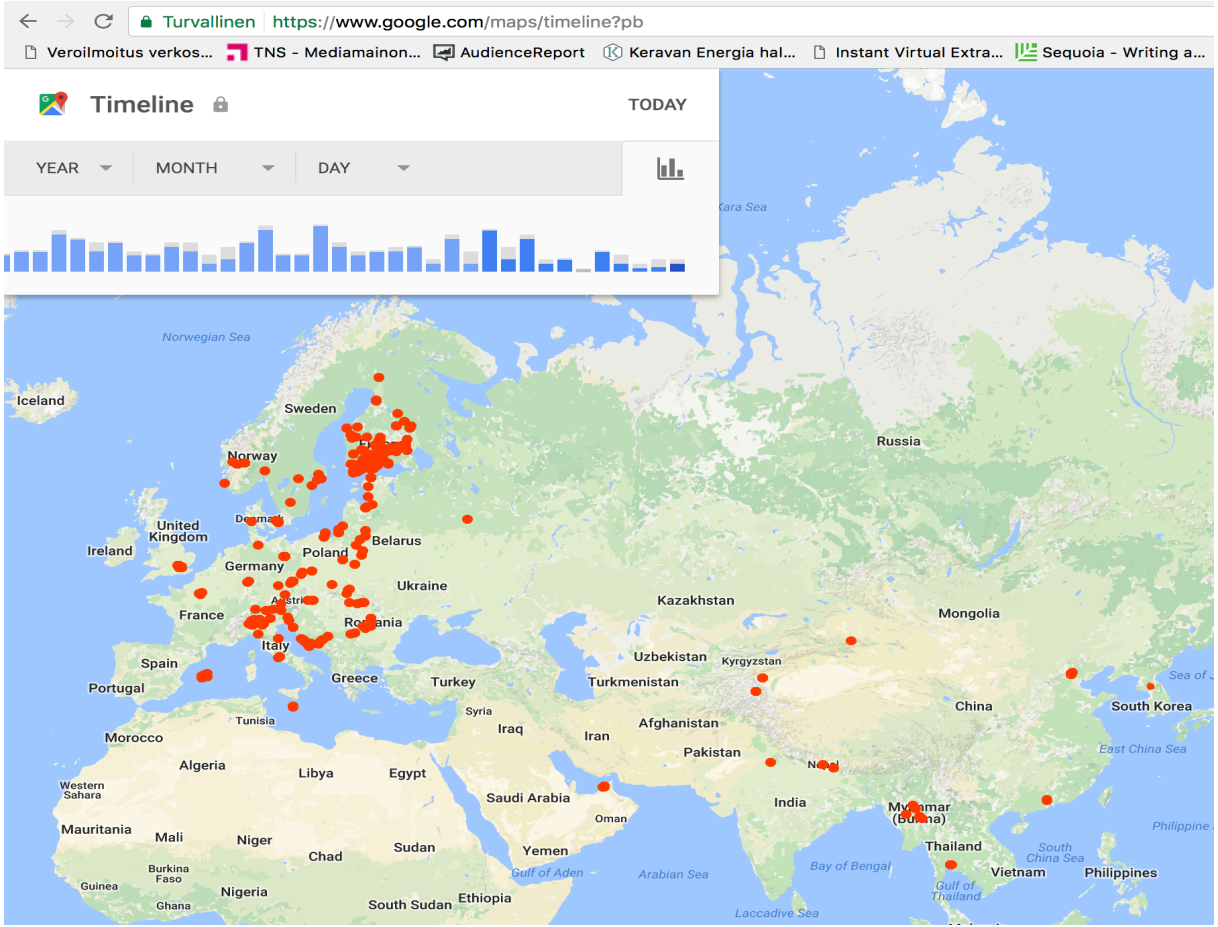
- graph data
  - mainly appears in the form of road networks.
  - Here, an edge represents a road segment, and a node represents an intersection or a landmark.
  - The trajectories of vehicles on the road network are
    - represented by sequences of road segments (edges)

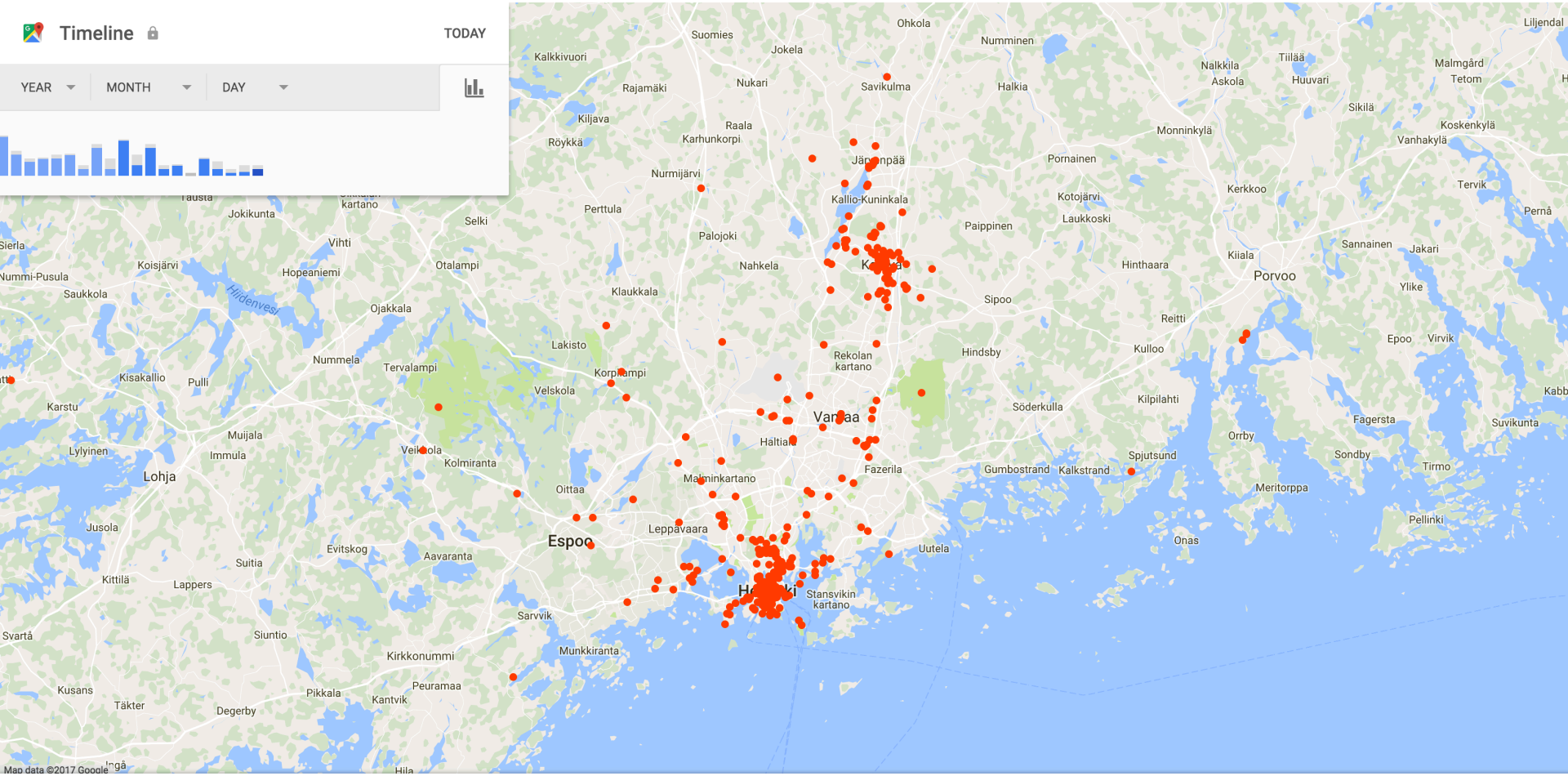
How we leave traces (=data)



Amount of data sources and real time data is increasing

- With the advancements of sensor and communication technologies, new sources of geospatial big data are emerging
  - sensors (or sensor networks) become more prevalent in these days.
    - loop detectors for detecting traffic in roads,
    - electrical grids,
    - environmental sensors for measuring air quality
  - Mobile


Turvallinen
https://www.google.com/maps/timeline?pb
Veroilmoitus verkos...
TNS - Mediamainon...
AudienceReport
Keravan Energia hal...
Instant Virtual Extra...
Sequoia - Writing a...
Timeline
TODAY
YEAR
MONTH
DAY
Norwegian Sea
Iceland
Sweden
Norway
Russia
United Kingdom
Ireland
Denmark
Belarus
Poland
Germany
Ukraine
France
Kazakhstan
Mongolia
Italy
Spain
Portugal
Greece
Turkey
Uzbekistan
Kyrgyzstan
Turkmenistan
China
South Korea
Tunisia
Syria
Iraq
Iran
Afghanistan
Pakistan
Morocco
Algeria
Libya
Egypt
Saudi Arabia
India
Western Sahara
Mauritania
Mali
Niger
Chad
Sudan
Yemen
Oman
Thailand
Vietnam
Philippines
Burkina Faso
Guinea
Ghana
Nigeria
South Sudan
Ethiopia
Gulf of Aden
Arabian Sea
Bay of Bengal
Gulf of Thailand
East China Sea
South China Sea
Laccadive Sea



Ilkka Lavas, Helsingin Yliopisto, Tietojenkäsittelytieteen laitos

Collecting data is not key issue anymore

The most important issue is how we use & exploit these geospatial big data

Applications

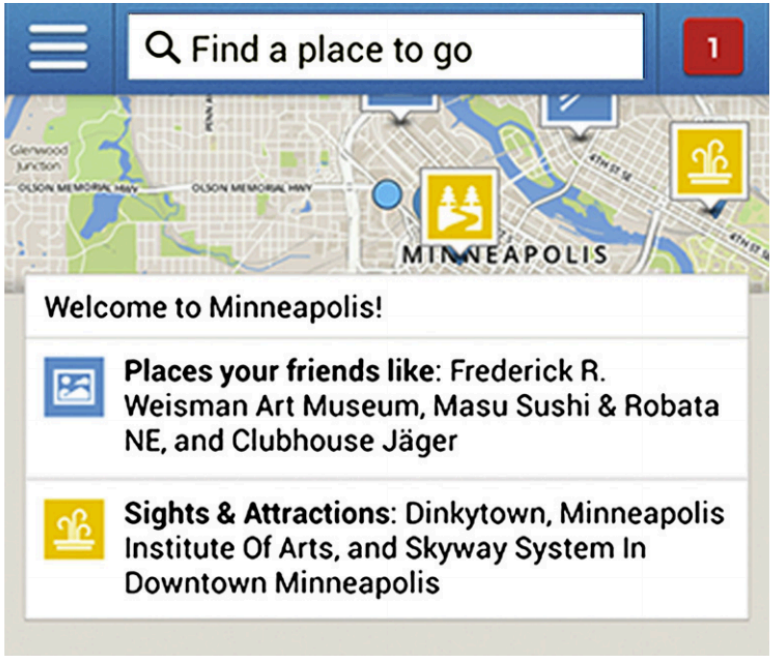


Fig. 6. Foursquare Android app with recommendations.

TILA

Kaikki

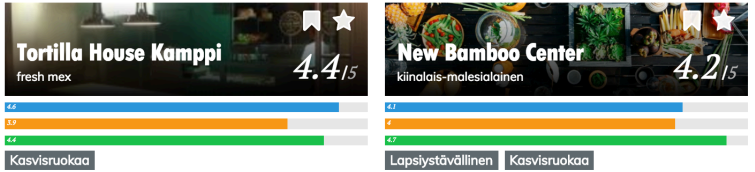
HINTA

Ei väliä

JÄRJESTÄ

Yleisarvosanan mukaan

Tyhjennä Näytä suodattimet

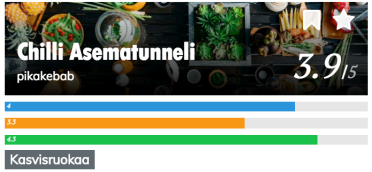


Kasvisruokaa

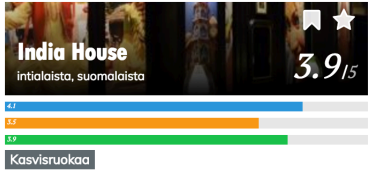
Lapsiystävällinen Kasvisruokaa

Kasvisruokaa

Lapsiystävällinen Kasvisruokaa



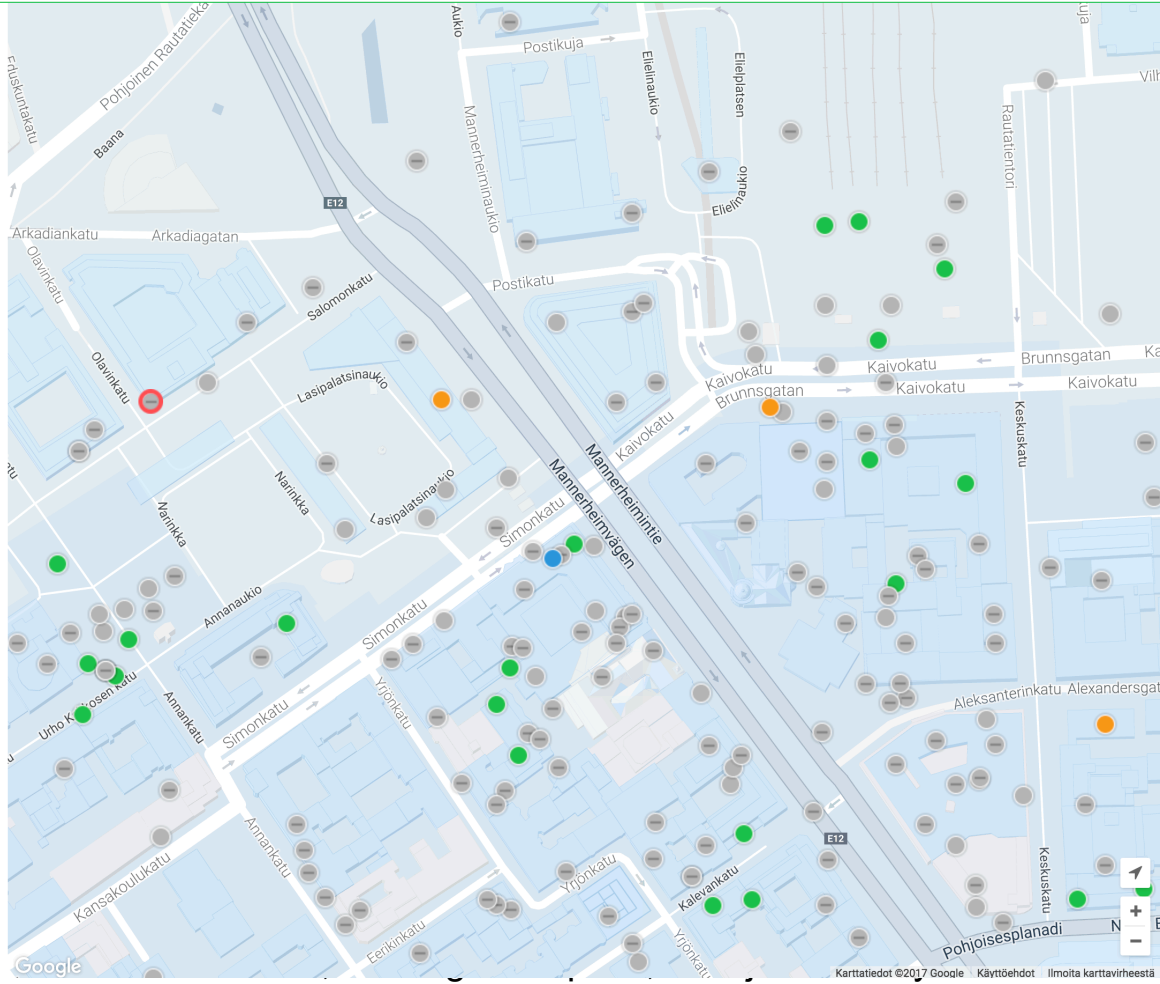
Kasvisruokaa

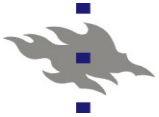


Kasvisruokaa

Lapsiystävällinen Kasvisruokaa Unicef

Kasvisruokaa

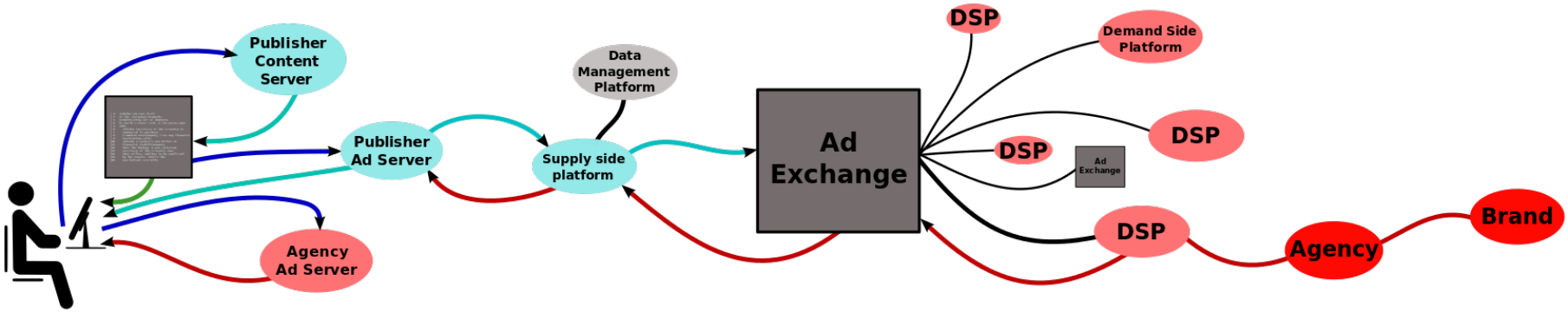




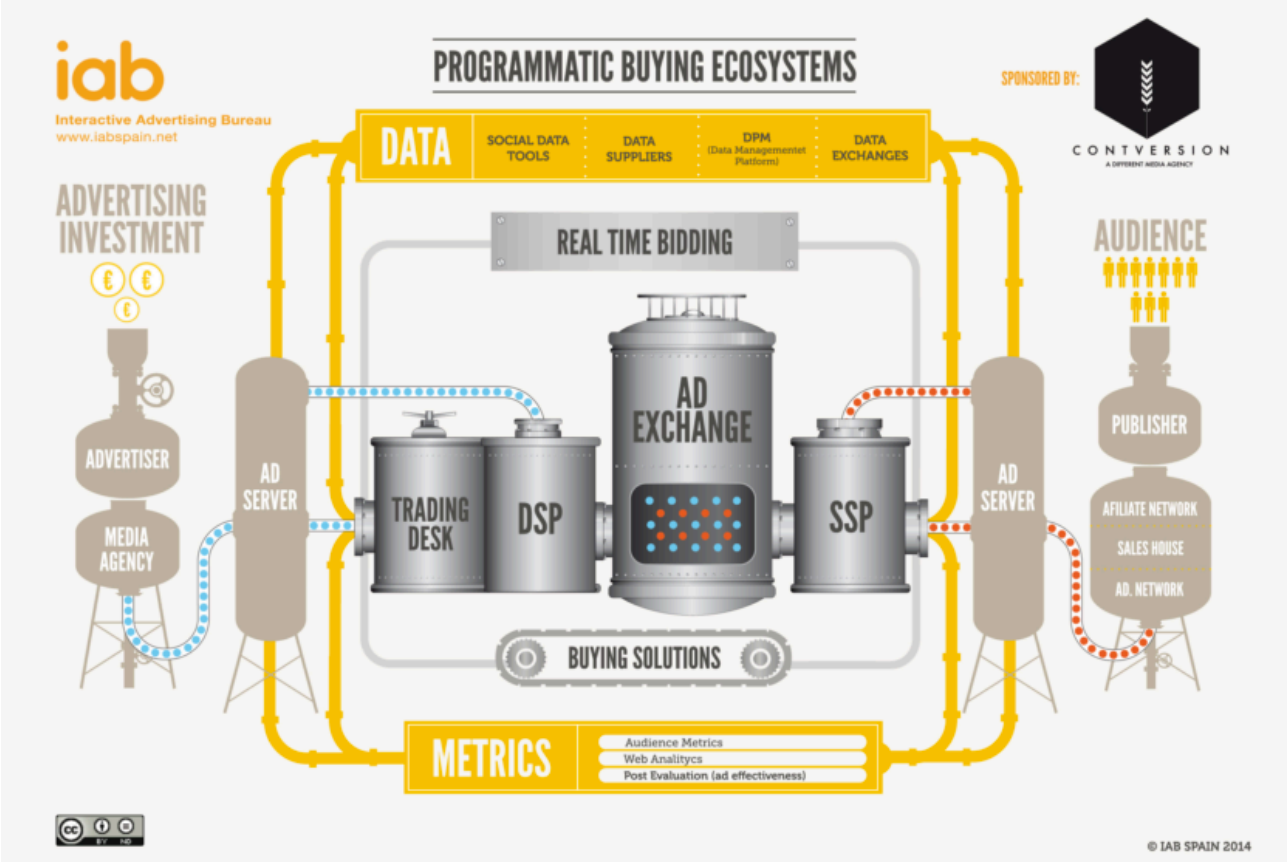
Advertisement Automation with Big Data

- Data can be used to provide better advertising

Advertising
Programmatic marketing ecosystem

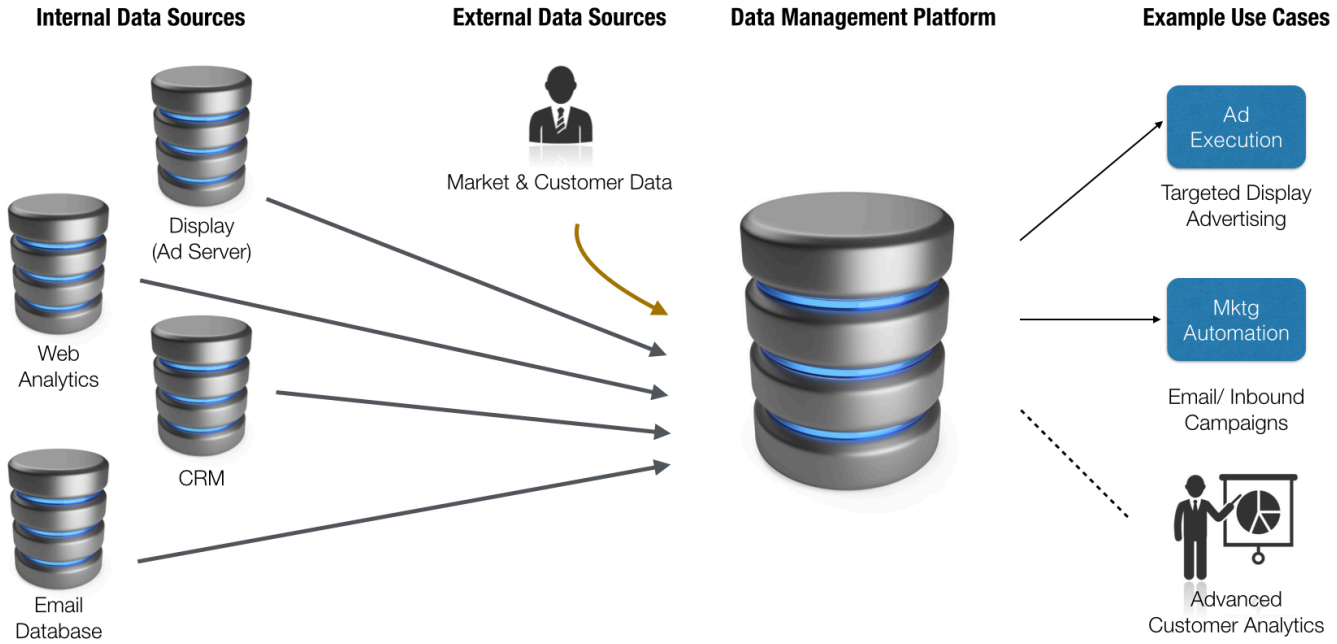


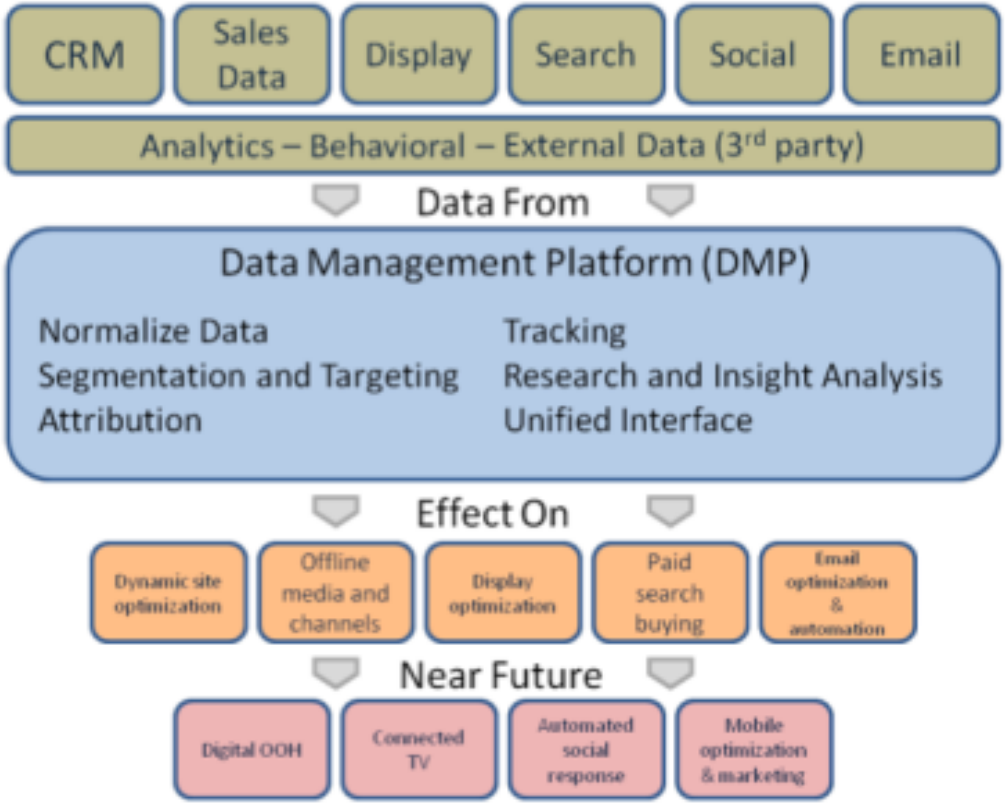
Advertising
Programmatic marketing ecosystem



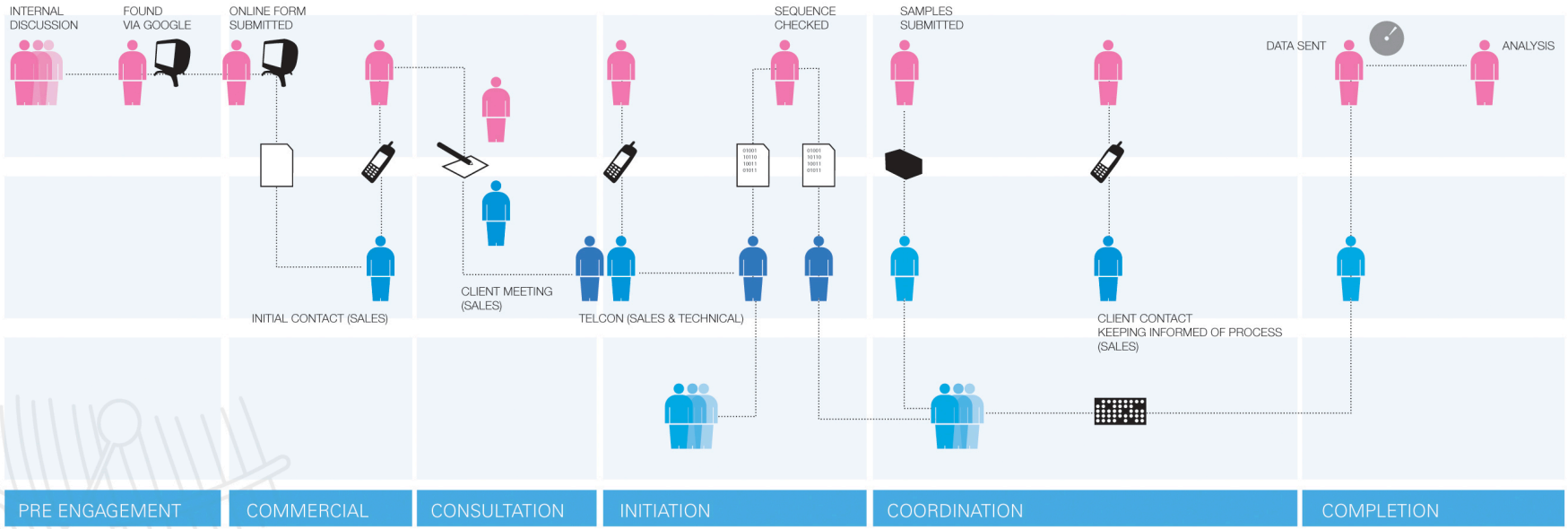
Data Management Platforms

Approach to collecting, organizing and activating customer data

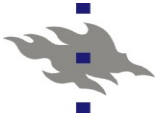



CRM
Sales Data
Display
Search
Social
Email
Analytics – Behavioral – External Data (3rd party)
Data From
Data Management Platform (DMP)
Normalize Data
Segmentation and Targeting
Attribution
Tracking
Research and Insight Analysis
Unified Interface
Effect On
Dynamic site optimization
Offline media and channels
Display optimization
Paid search buying
Email optimization & automation
Near Future
Digital OOH
Connected TV
Automated social response
Mobile optimization & marketing
Source: Econsultancy

Customer path



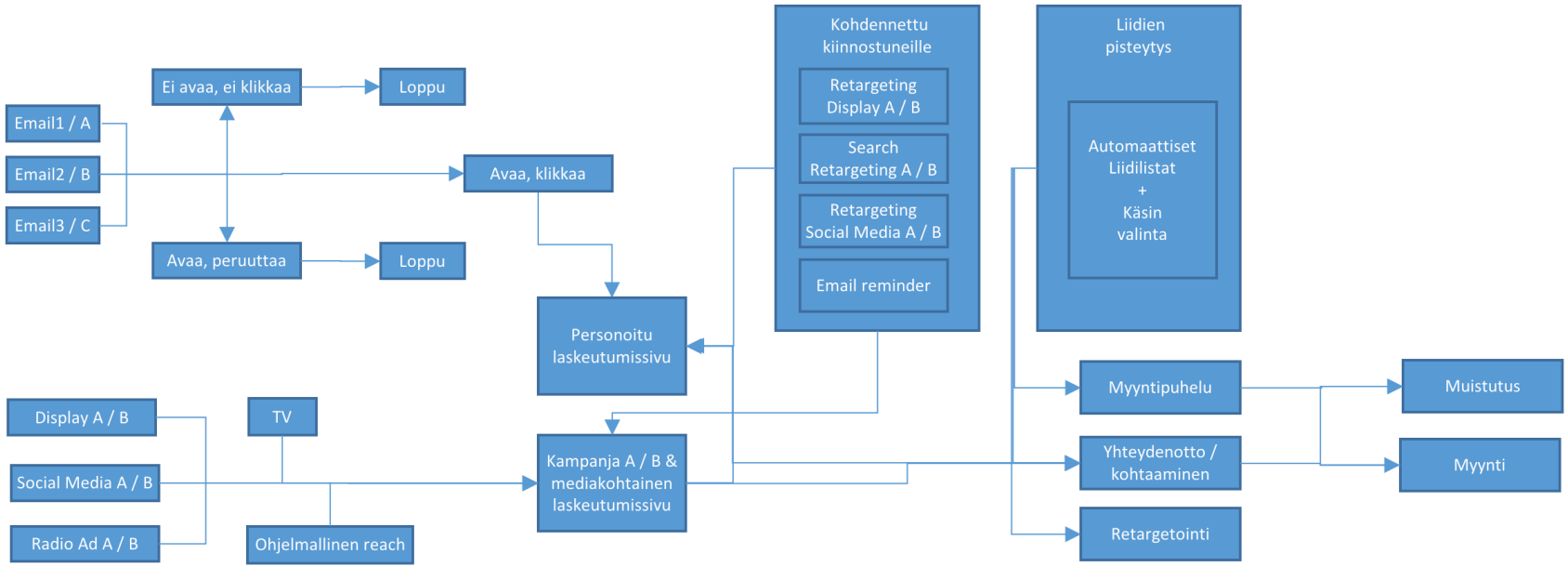
Ilkka Lavas, Helsingin Yliopisto, Tietojenkäsittelytieteen laitos



The future of advertising will be data driven

Ilkka Lavas, Helsingin Yliopisto, Tietojenkäsittelytieteen laitos

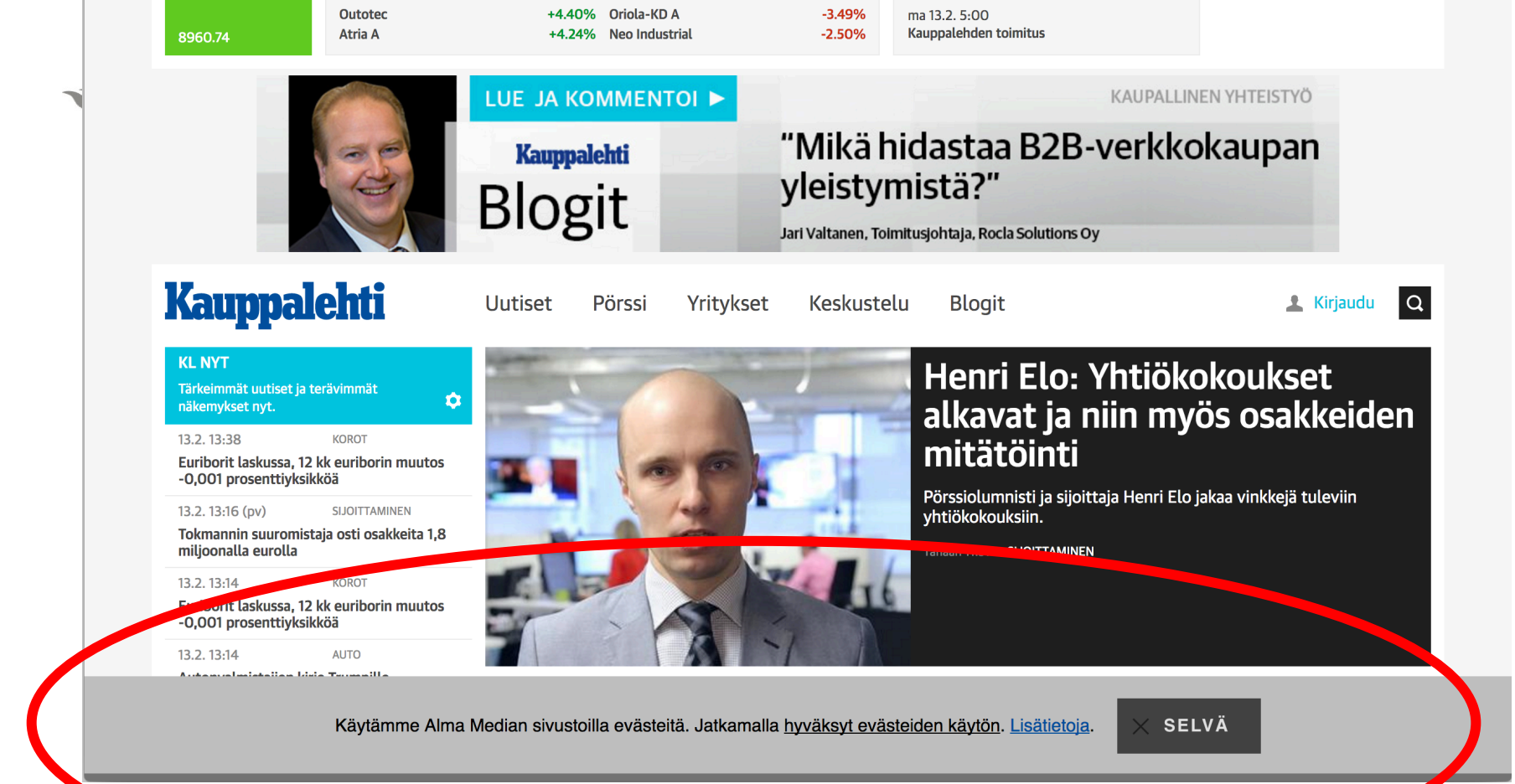
Continuous marketing w/ Big Data



Ilkka Lavas, Helsingin Yliopisto, Tietojenkäsittelytieteen laitos

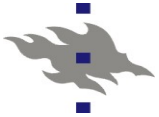
Open problems

- General Data Protection Regulation:
  REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC


8960.74
Outotec +4.40% Oriola-KD A -3.49%
Atria A +4.24% Neo Industrial -2.50%
ma 13.2. 5:00
Kauppalehden toimitus
LUE JA KOMMENTOI ▶
Kauppalehti
Blogit
KAUPALLINEN YHTEISTYÖ
"Mikä hidastaa B2B-verkkokaupan yleistymistä?"
Jari Valtanen, Toimitusjohtaja, Rocla Solutions Oy
Kauppalehti
Uutiset
Pörssi
Yritykset
Keskustelu
Blogit
Kirjaudu
KL NYT
Tärkeimmät uutiset ja terävimmät näkemykset nyt.
13.2. 13:38 KOROT
Euriborit laskussa, 12 kk euriborin muutos -0,001 prosenttiyksikköä
13.2. 13:16 (pv) SIJOITTAMINEN
Tokmannin suuromistaja osti osakkeita 1,8 miljoonalla eurolla
13.2. 13:14 KOROT
13.2. 13:14 AUTO
Henri Elo: Yhtiökokoukset alkavat ja niin myös osakkeiden mitätöinti
Pörssiolumnisti ja sijoittaja Henri Elo jakaa vinkkejä tuleviin yhtiökokouksiin.
Käytämme Alma Median sivustoilla evästeitä. Jatkamalla hyväksyt evästeiden käytön. Lisätietoja.
SELVÄ

Summary & Conclusion

- Big Data consist of collection, analysis and usage
- Collection is not anymore an issue
- Analysis needs a lot of work
- Right use gives more value to brand, customer, user and ecosystem
  - Better advertising
  - Less intrusive ads
  - More relevancy in ads
  - What before was spooky, will be new standard and we expect to be served according to data traces we leave: ”How nice that you think of me”



Sources

- Paul Suganthan G. C., Chong Sun, Krishna Gayatri K., Haojun Zhang, Frank Yang, Narasimhan Rampalli, Shishir Prasad, Esteban Arcaute, Ganesh Krishnan, Rohit Deep, Vijay Raghavendra, AnHai Doan: Why Big Data Industrial Systems Need Rules and What We Can Do About It. SIGMOD Conference 2015: 265-276 [pdf paper]
- https://www.cs.helsinki.fi/u/jilu/paper/bigdataapplication05.pdf

- (2) Javier Andréu Pérez, Carmen C. Y. Poon, Robert D. Merrifield, Stephen T. C. Wong, Guang-Zhong Yang: Big Data for Health. IEEE J. Biomedical and Health Informatics 19(4): 1193-1208 (2015) [pdf paper]
- https://www.cs.helsinki.fi/u/jilu/paper/bigdataapplication01.pdf

- (3) Jae-Gil Lee, Minseo Kang: Geospatial Big Data: Challenges and Opportunities. Big Data Research 2(2): 74-81 (2015) [pdf paper]
- https://www.cs.helsinki.fi/u/jilu/paper/bigdataapplication02.pdf

- (4) Taruna Seth, Vipin Chaudhary: Big Data in Finance. Big Data - Algorithms, Analytics, and Applications 2015: 329-356 [pdf paper]

- (5) Kesheng Wu, E. Wes Bethel, Ming Gu, David Leinweber, Oliver Rübel: A big data approach to analyzing market volatility. Algorithmic Finance 2(3-4): 241-267 (2013) [pdf paper]