Liikuntasovelluksella kerätyn big datan käyttö reitityksessä – case pyöräily

Cecilia Bergman

18.4.2016, Poligonin ja Suomen Kartografisen seuran kevätseminaari

MML PAIKKA-TIETO-KESKUS FGI

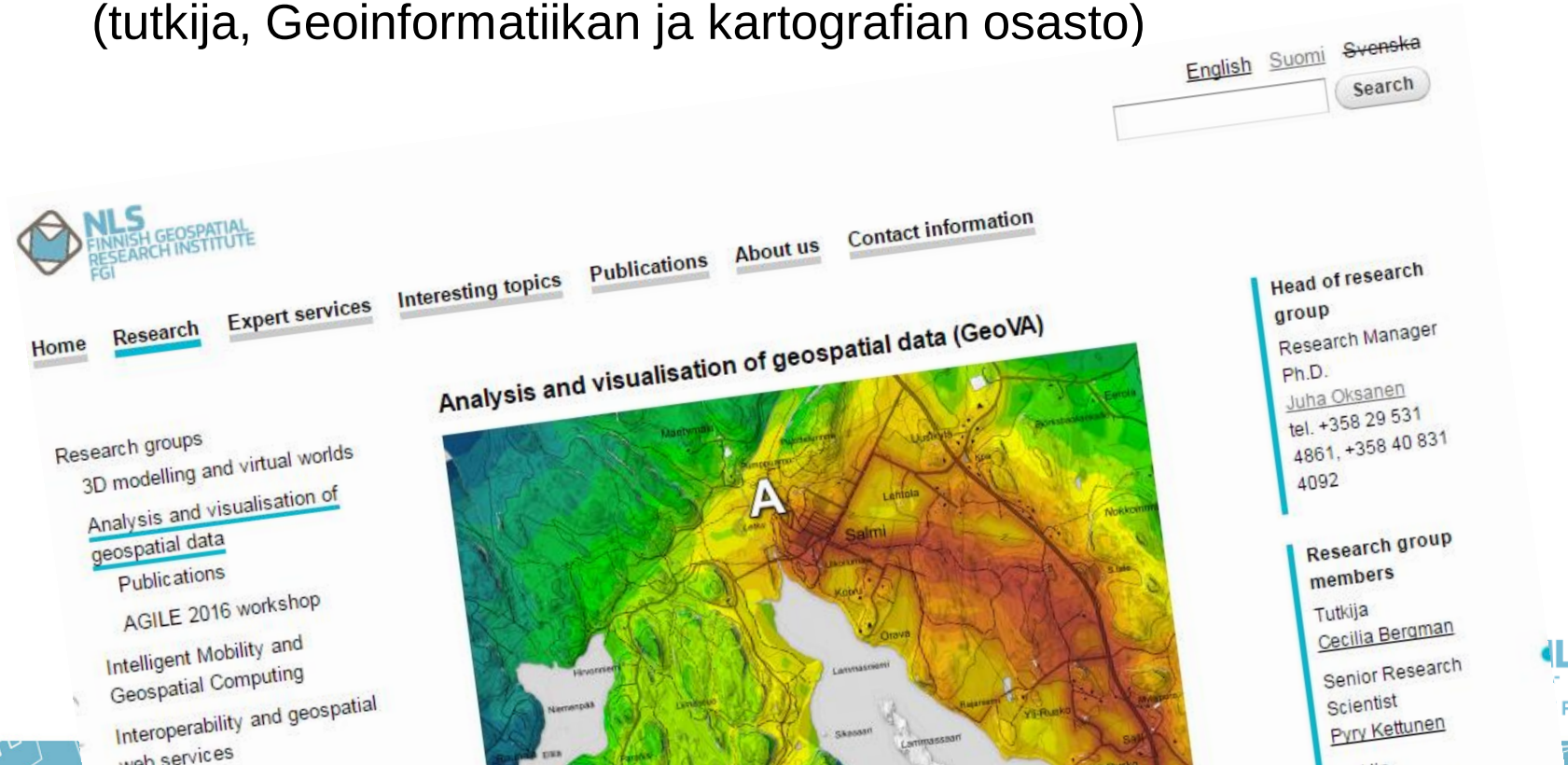
Sisältö

- Joukkoistettu data (crowd-sourced data)
  - Seuraavan sukupolven reitityspalvelut
  - Joukkoistettu data reitityksessä
- Sports Tracker -sovelluksella kerätyt liikuntasuoritukset
- Suosioon perustuva reititys OSM:n pyöräilyverkolla
- Haasteet ja mahdollisuudet

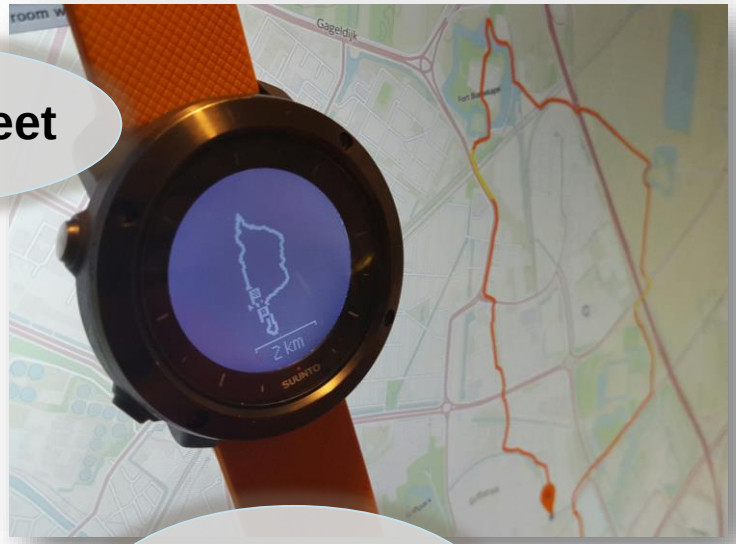
MML PAIKKA-TIETO-KESKUS FGI

Kuka?

- 2008 FM (Suunnittelumaantiede), HY
- 2008- TKK/Aalto (Geoinformatiikka)
- 2009-2014 WSP (suunnittelija, Liikenteen konsultointi)
- 2014 Paikkatietokeskus FGI, Maanmittauslaitos (tutkija, Geoinformatiikan ja kartografian osasto)



GPS-laitteet

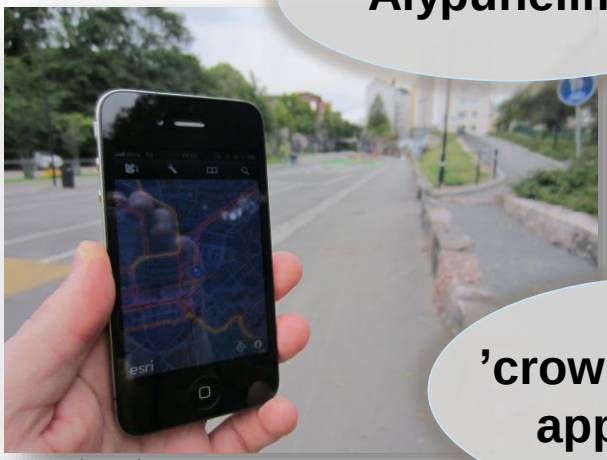


Sosiaalinen media

Digitaaliset sensorit

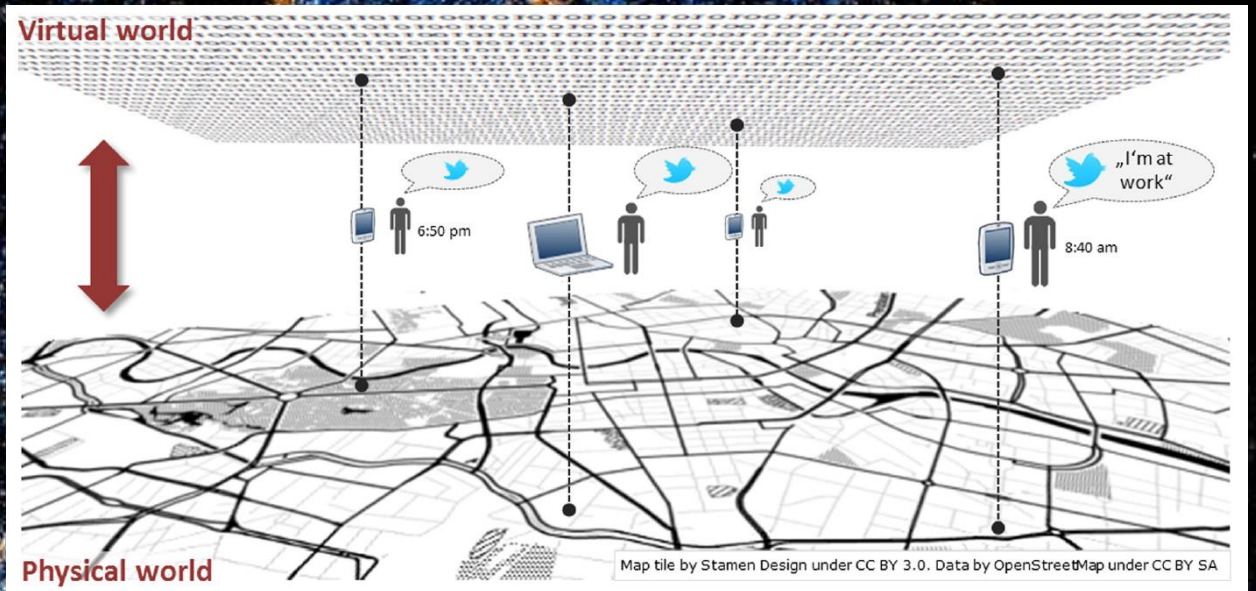
http://hylo.mygeotrust.org/

Älypuhelimet



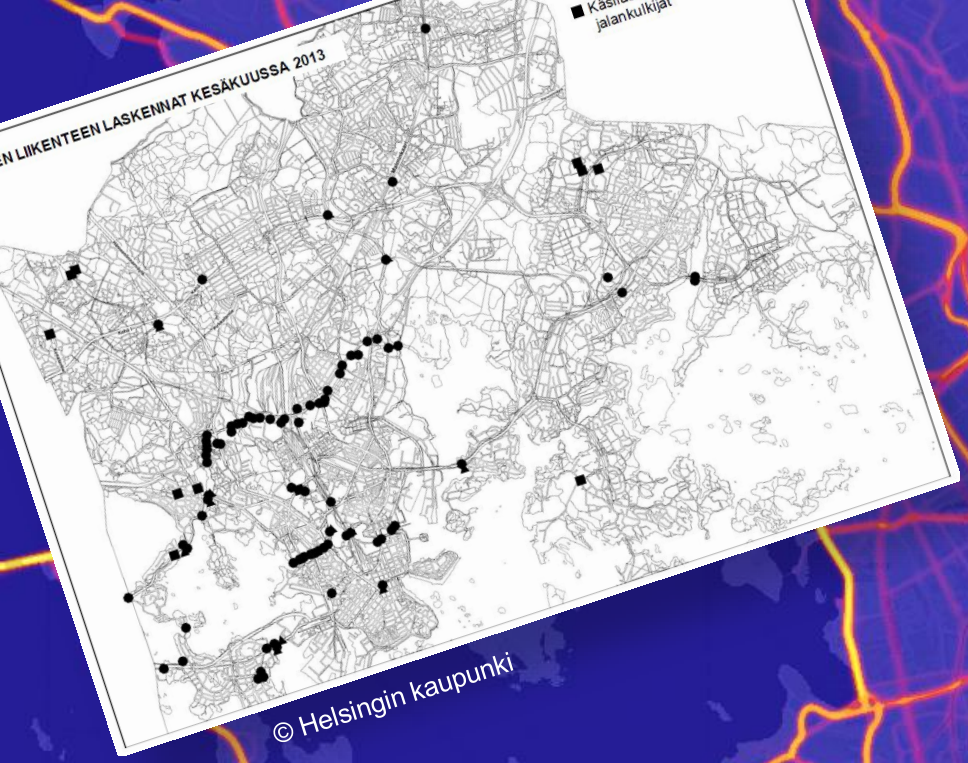
'crowdsourcing' applikaatiot

MML PAIKKA-TIETO-KESKUS FGI



Steiger, Enrico, et al. "Twitter as an indicator for whereabouts of people? Correlating twitter with uk census data." *Computers, Environment and Urban Systems* 54 (2015): 255-265.

- Älykäs kaupunki (smart city)
- Tilastot (esim. matkailu, liikkuminen)
  - täydentäjä / korvaaja / jotain aivan uutta?
- Ajantasaisuus / reaaliaikaisuus
- ?



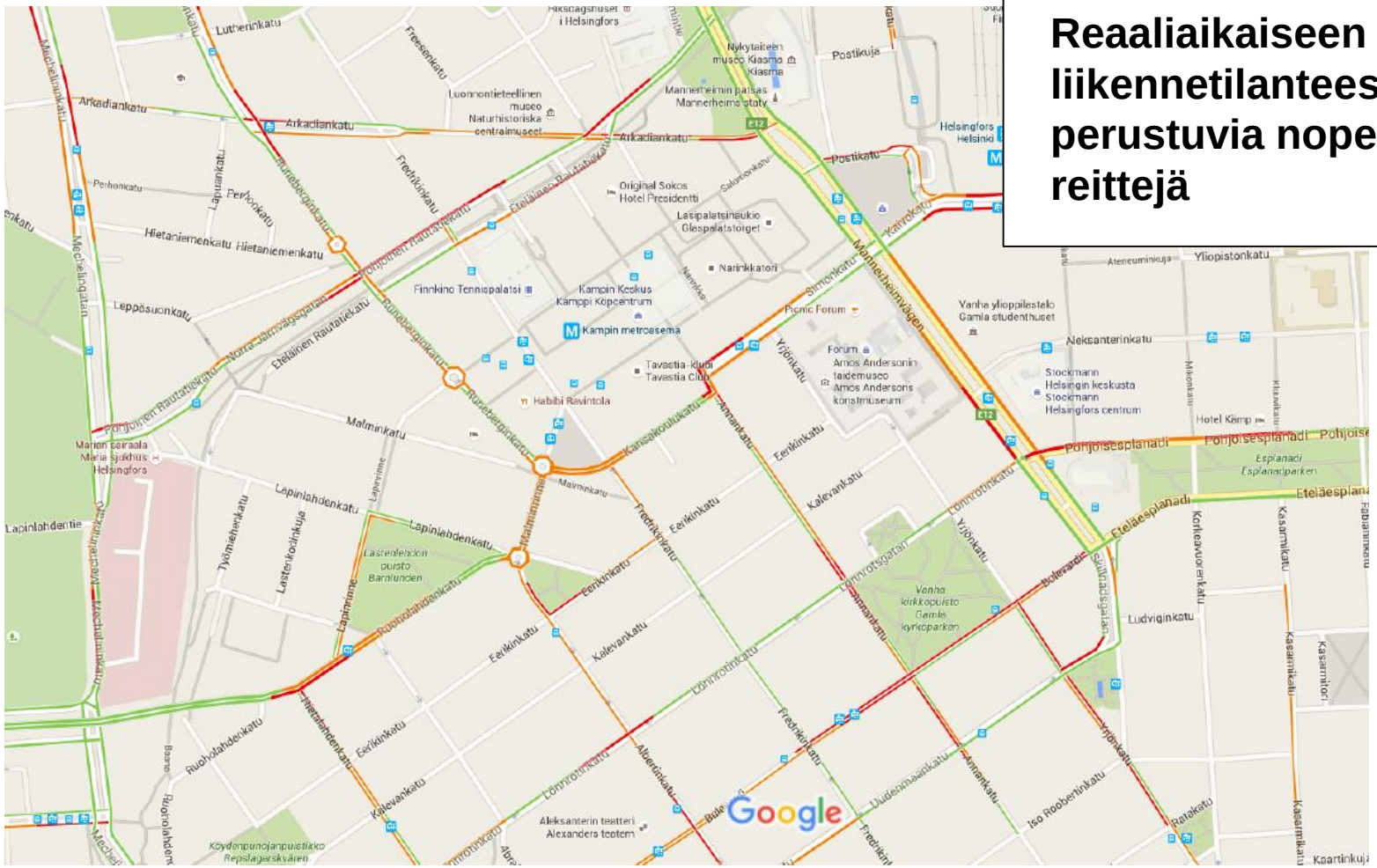
© Helsingin kaupunki

Supra

Density of workouts

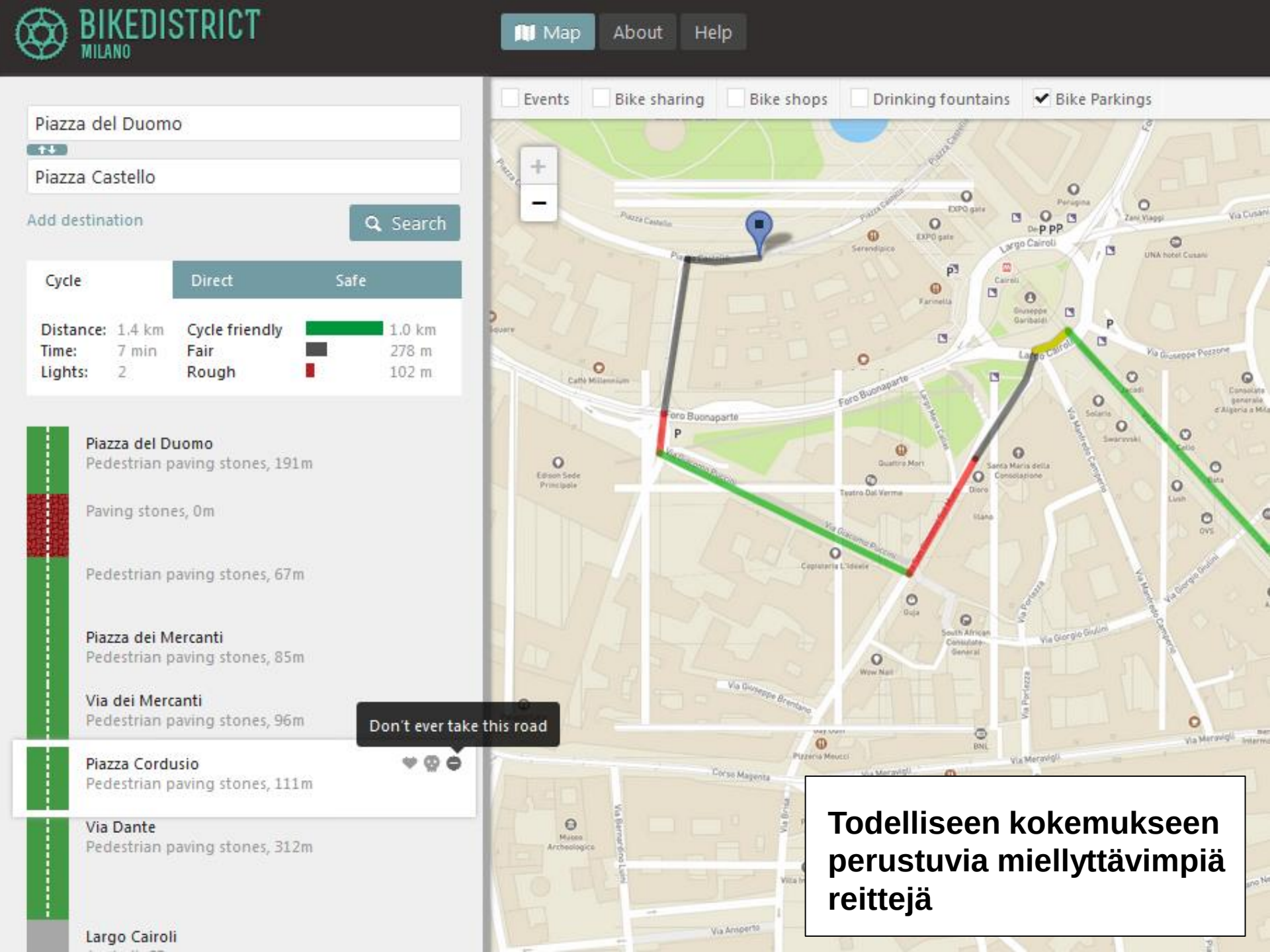
low high

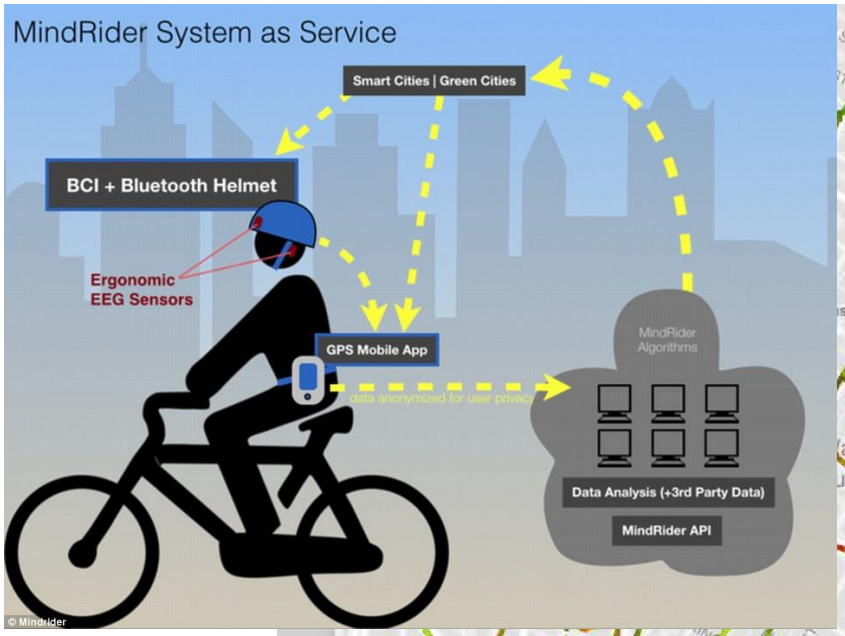
Joukkoistettu data reitityksessä

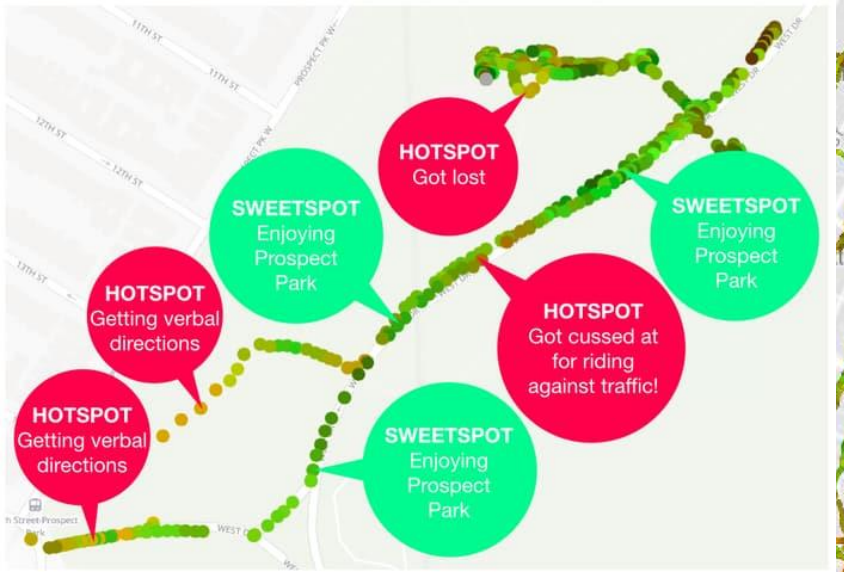


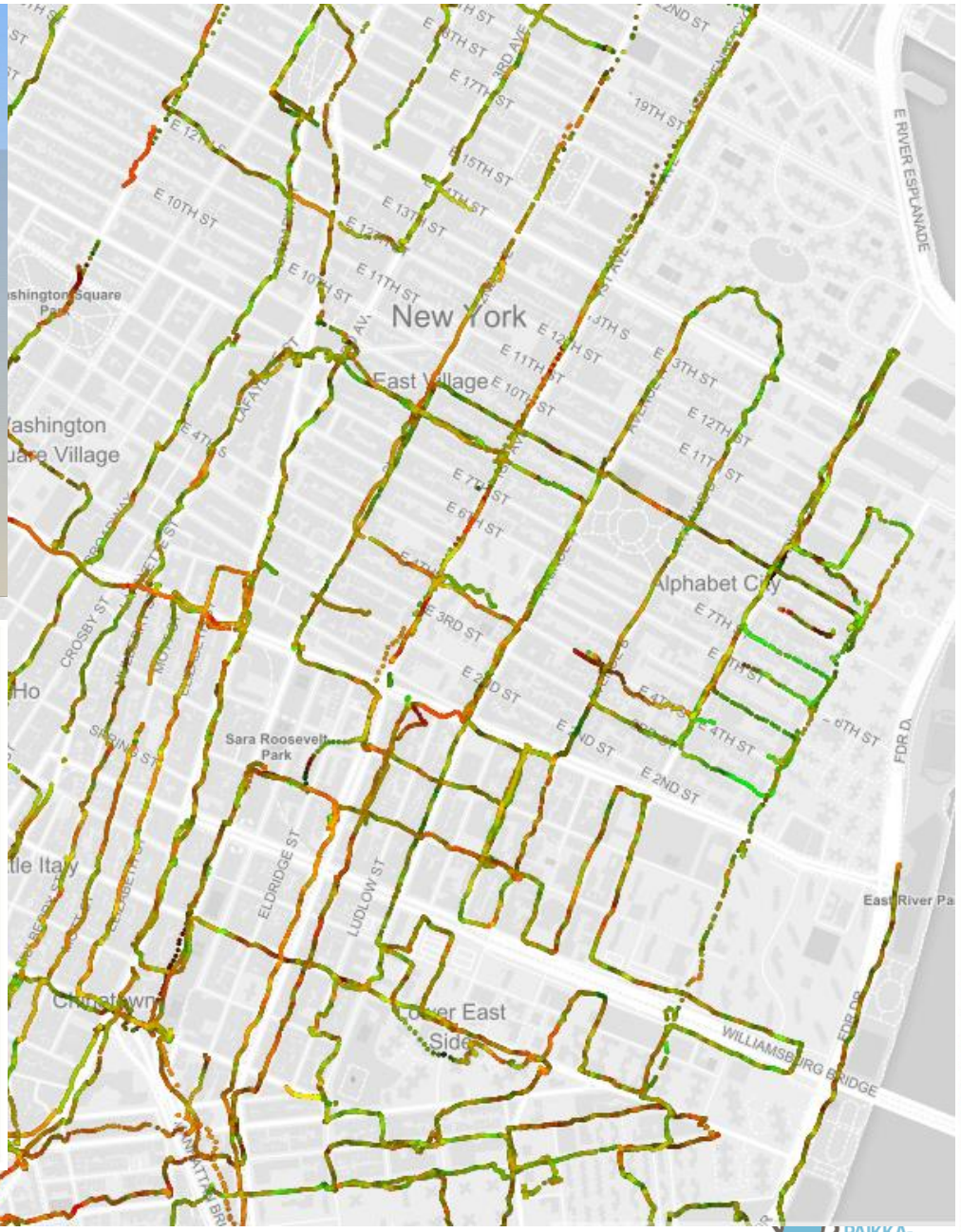
Reaaliaikaiseen liikennetilanteeseen perustuvia nopeampia reittejä

MML PAIKKA-TIETO-KESKUS FGI


BIKEDISTRICT
MILANO
Map
About
Help
Events
Bike sharing
Bike shops
Drinking fountains
Bike Parkings
Piazza del Duomo
Piazza Castello
Add destination
Search
Cycle
Direct
Safe
Distance: 1.4 km
Time: 7 min
Lights: 2
Cycle friendly 1.0 km
Fair 278 m
Rough 102 m
Piazza del Duomo
Pedestrian paving stones, 191m
Paving stones, 0m
Pedestrian paving stones, 67m
Piazza dei Mercanti
Pedestrian paving stones, 85m
Via dei Mercanti
Pedestrian paving stones, 96m
Don't ever take this road
Piazza Cordusio
Pedestrian paving stones, 111m
Via Dante
Pedestrian paving stones, 312m
Largo Cairoli
Todelliseen kokemukseen perustuvia miellyttävimpiä reittejä







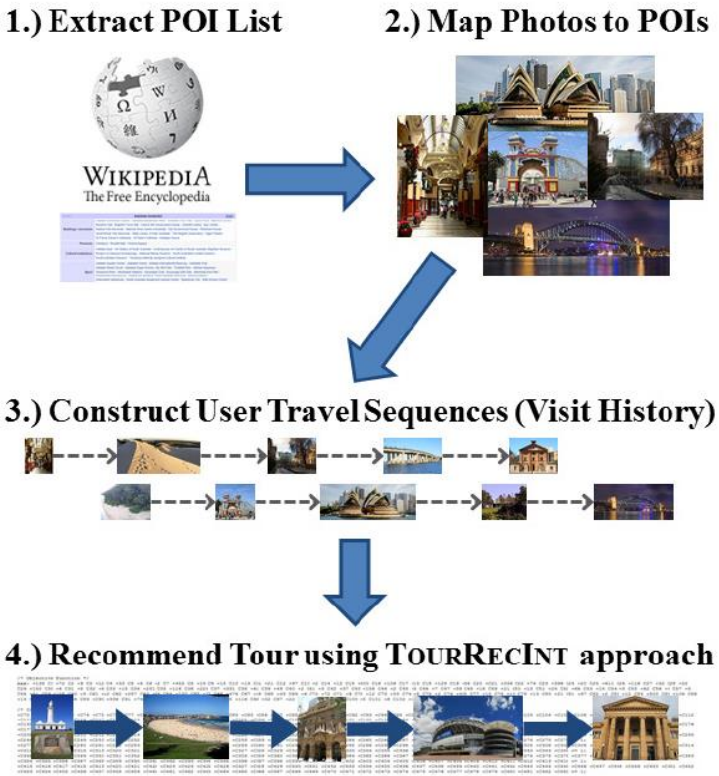
http://www.dailymail.co.uk/sciencetech/article-2936923/Stress-free-cycle-routes-revealed-using-BRAIN-WAVES-Mind-reading-helmet-highlights-relaxing-routes-New-York.html

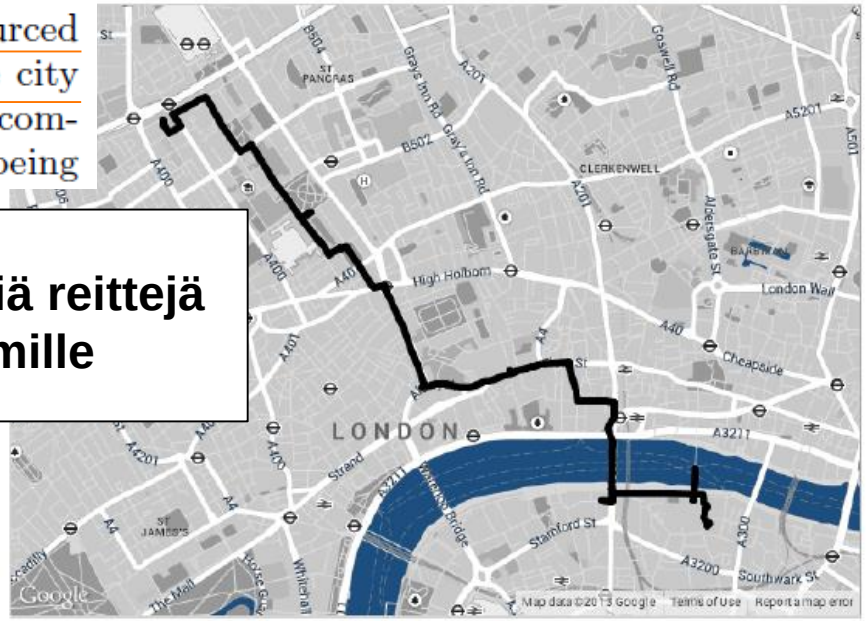
PAIKKA-
TIETO-
KESKUS FGI

only short but also emotionally pleasant. This goal is not algorithmic but is experimental. We rely on crowdsourced measurements of people's emotional experience of the city and use those measurements to propose new ways of recommending urban routes. Despite emotional responses being

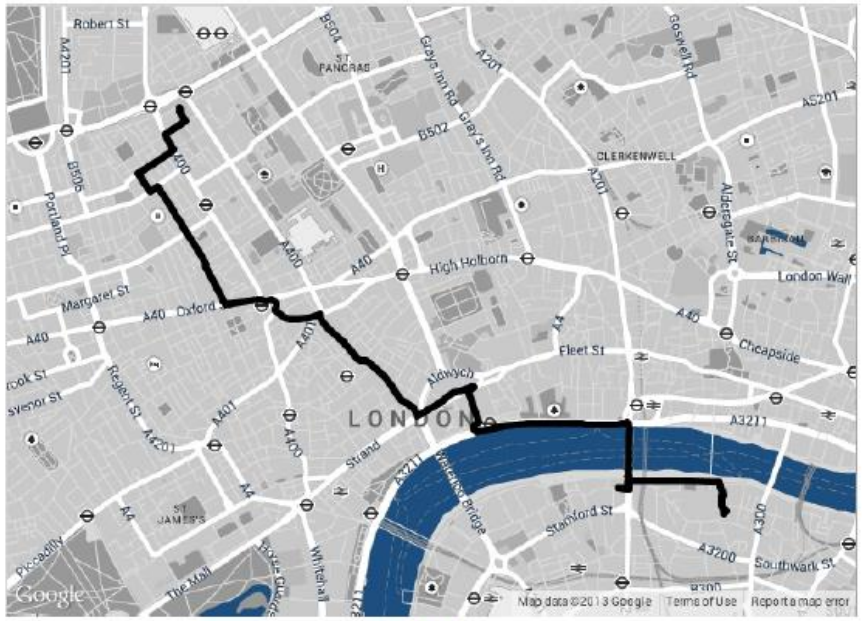
Miellyttävämpiä reittejä eri liikkujaryhmille

Lim, Kwan Hui. "Recommending tours and places-of-interest based on user interests from geo-tagged photos." *Proceedings of the 2015 ACM SIGMOD on PhD Symposium*. ACM, 2015.





(b) Beauty



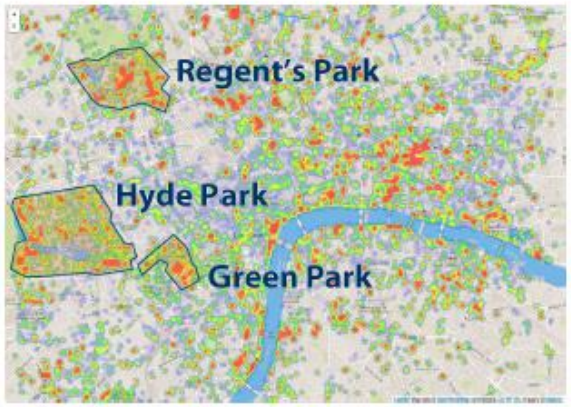
(d) Happy

Eri liikkujaryhmille räätälöityjä reittejä

Neis, Pascal. "Measuring the reliability of wheelchair user route planning based on volunteered geographic information." *Transactions in GIS* 19.2 (2015): 188-201.

Type of street	*highway*	[1]
Sidewalk(s) of the way	*sidewalk|footway*	*left|right|yes|no|both*
Width of the way	*width*	in [m]
Surface of the way	*surface*	[2]
Smoothness of the way	*smoothness*	[3]
Slope/Incline of the way	*incline*	in [%]
Curb/Kerb of the way	*sloped_curb*	in [m]
Lighting	*lit*	*yes|no*
General Access of a way	*foot|wheelchair*	*yes|no*

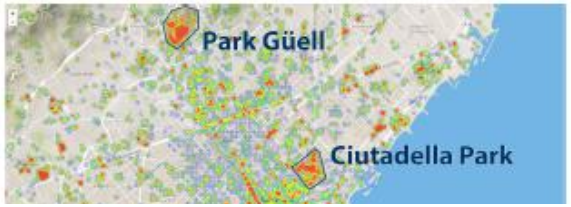
PAIKKA-TIETO-KESKUS FGI



(a) London, nature

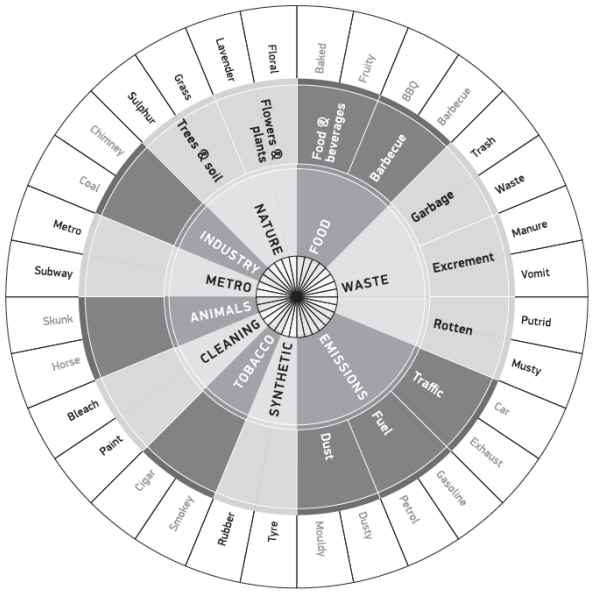
(b) London, emissions

(c) London, animals



(e) Barcelona, emissions

(f) Barcelona, animals



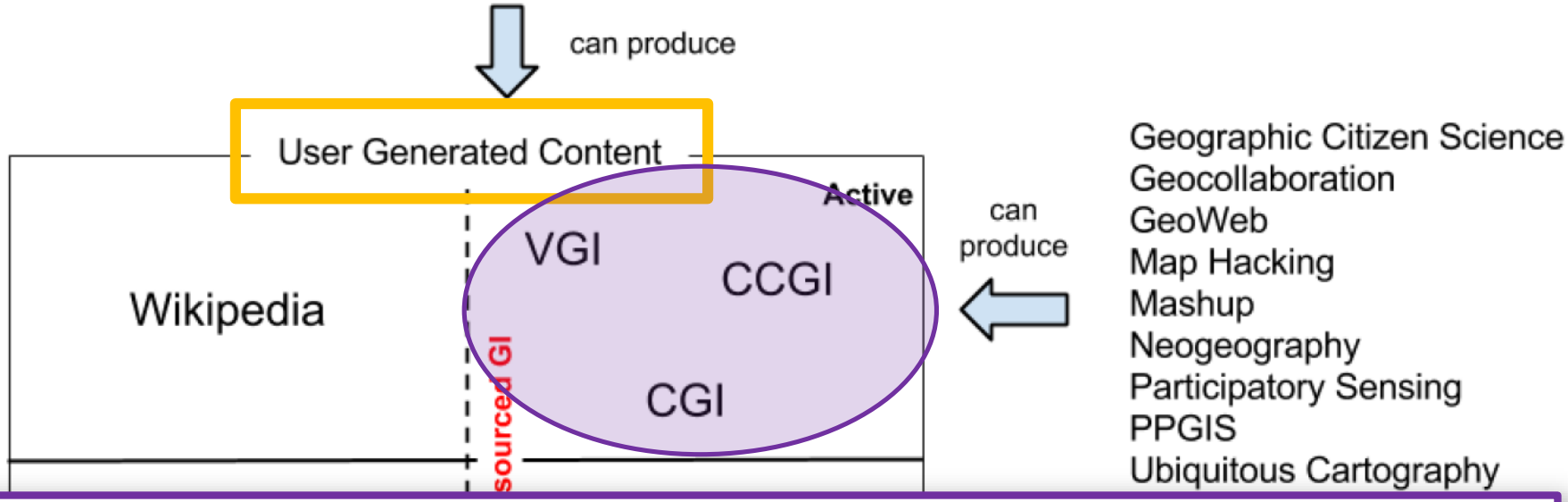
Urban Smellscape Aroma Wheel

e 7: Heatmaps of smell-related tag intensity

Quercia, Daniele, et al. "Smelly maps: the digital life of urban smellscapes."*arXiv preprint arXiv:1505.06851* (2015).

MML PAIKKA-TIETO-KESKUS FGI

(Extreme) Citizen Science, Crowdsourcing, PPSR, Science 2.0, Swarm Intelligence, Wikinomics



VGI (volunteered geographic information)
”geographic data provided voluntarily by individuals” (Goodchild 2007)
”opt-in”, esim. OpenStreetMap, Geocaching

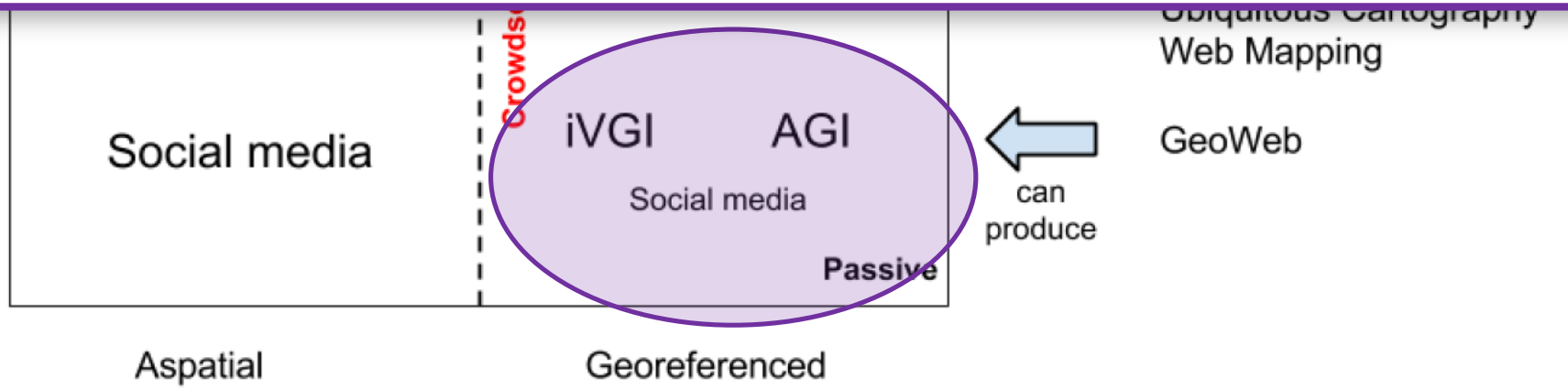
CCGI (collaboratively collected geographic information)

CGI (contributed geographic information)
“opt-out”, esim. cell-phone tracking

iVGI (involuntary geographic information)

AGI (ambient geographic information)

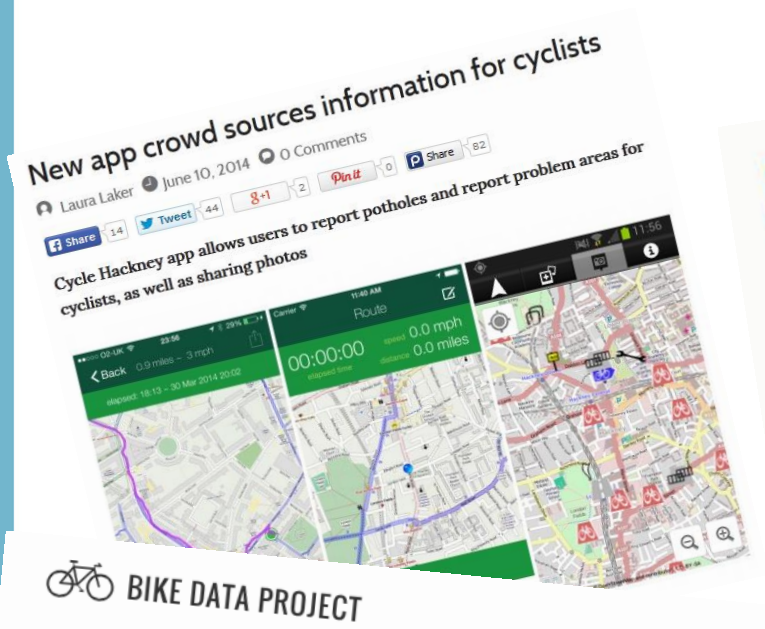
Social media

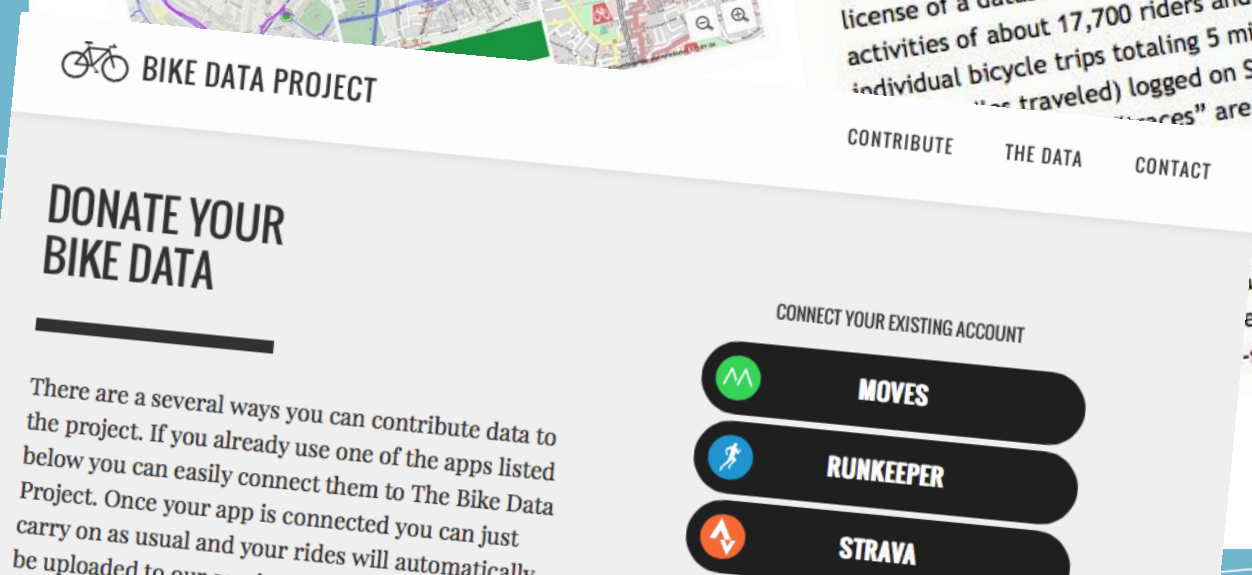


See et al. (2014). WG1 Review Document on Crowdsourced Geographic Information. Working Paper of COST Action TD1202: Mapping and the Citizen Sensor.

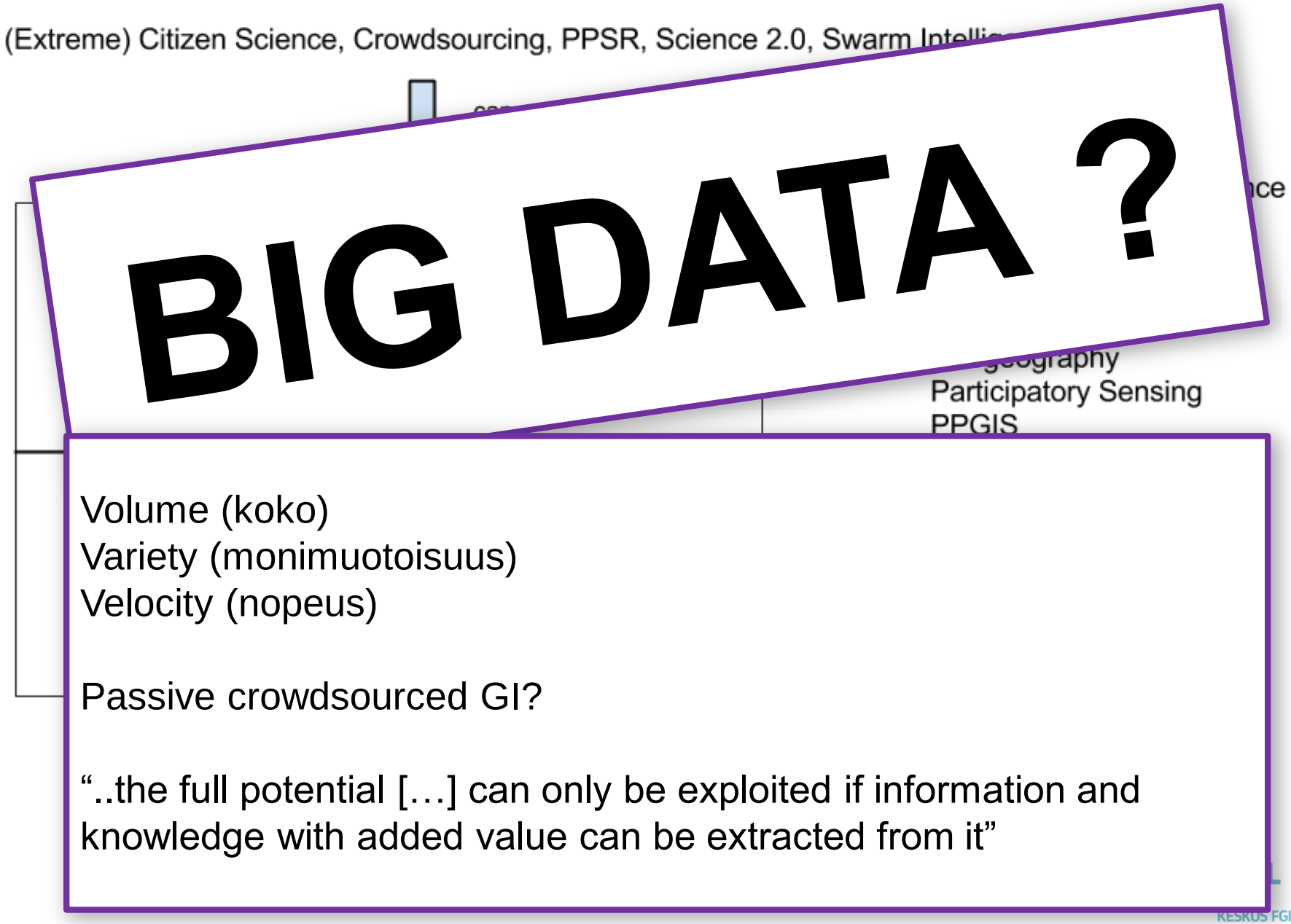
1. Datan keräämistä varten (esim. tutkimus, suunnittelu) tehdyt älypuhelisovellukset

2. Liikuntasovellukset, joiden ensisijainen tarkoitus liikuntasuoritusten seuranta ja jakaminen



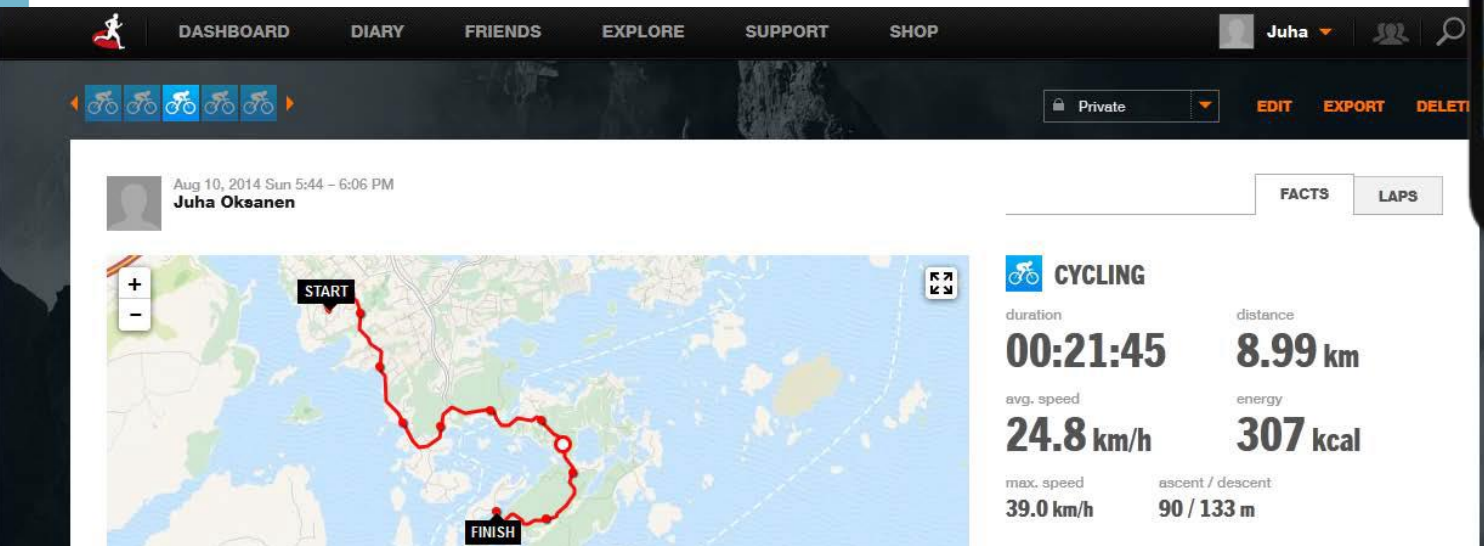


MML PAIKKA-TIETO-KESKUS FGI


(Extreme) Citizen Science, Crowdsourcing, PPSR, Science 2.0, Swarm Intelli
nce
BIG DATA ?
Participatory Sensing
PPGIS
Volume (koko)
Variety (monimuotoisuus)
Velocity (nopeus)
Passive crowdsourced GI?
“..the full potential […] can only be exploited if information and knowledge with added value can be extracted from it”
KESKUS FGI

Sports Tracker?

- Ilmainen mobiili liikuntasovellus, jolla voidaan pitää harjoituspäiväkirjaa
- Online palvelu *sports-tracker.com*
- Sosiaalinen verkosto
- Aloitti 2004 osana Nokiaa, 2009 itsenäinen yritys Sports Tracking Technologies Oy, 2015 osa Amer Sports Oyj:tä
- >25 miljoonaa sovelluslatausta kaikille yleisimmille alustoille, >2 miljoona käyttäjää kuukausittain >200 maassa



MML PAIKKA-TIETO-KESKUS FGI

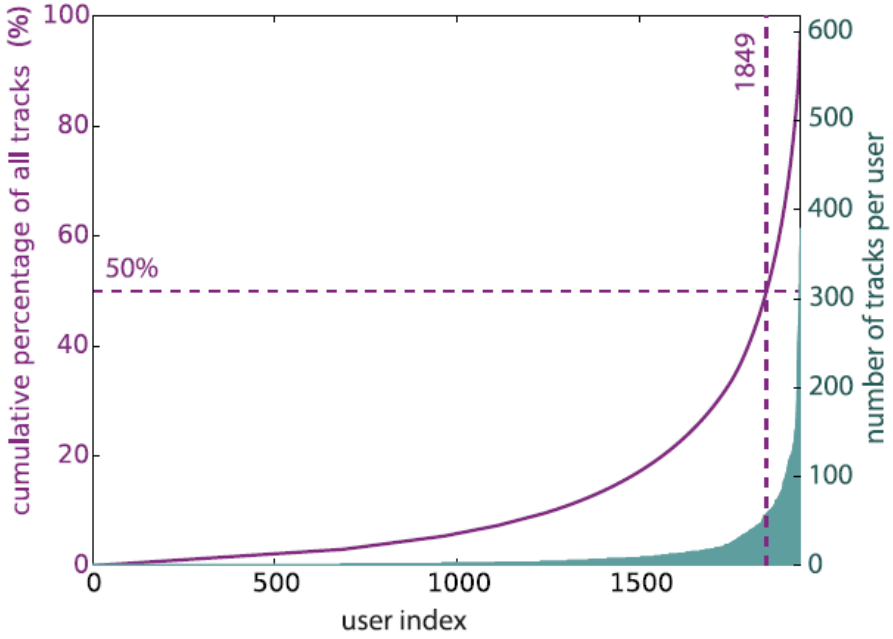
Sports Tracker data

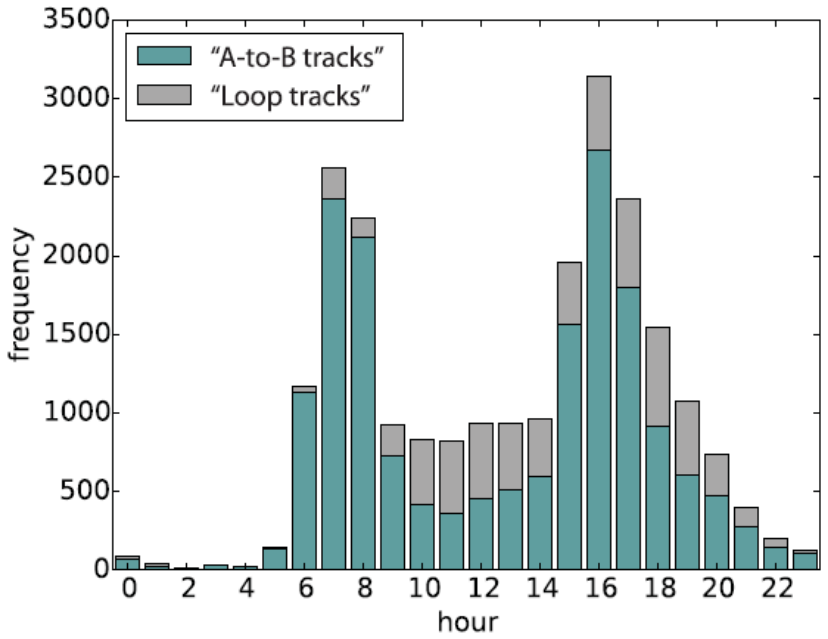
- ID/pseudo-ID
- Yksityisyyden taso
  - yksityinen
  - jaetaan ystäville
  - julkinen (näytetään ST web-sovelluksessa)

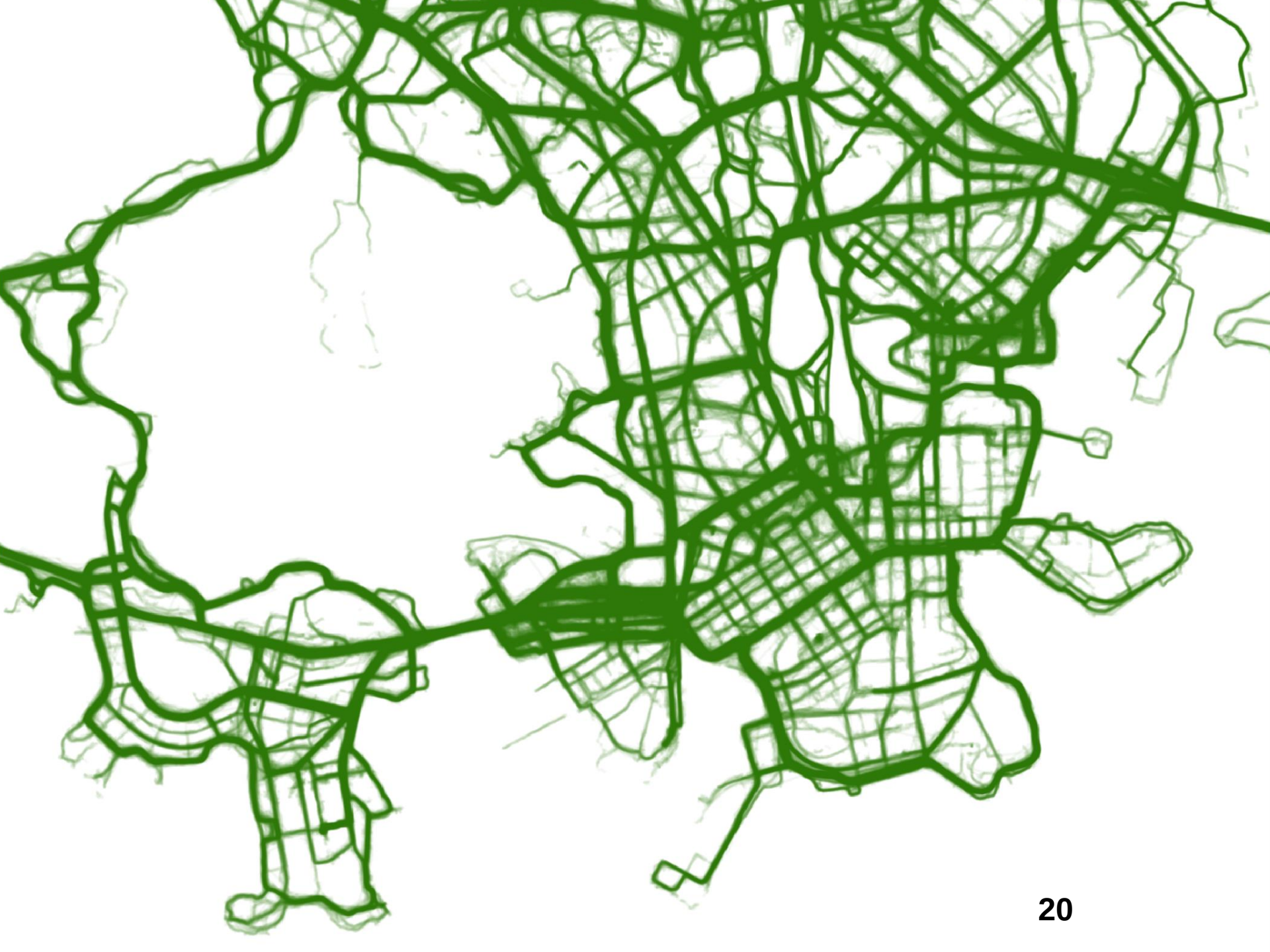
```
<metadata>
  <name>Cycling 24/05/2013/15:46:43.18</name>
  <desc></desc>
  <author>
    <name>Juha Oksanen</name>
  </author>
  <link href="www.sports-tracker.com">
    <text>Sports Tracker</text>
  </link>
</metadata>
<trk>
  <trkseg>
    <trkpt lat="60.16121833333333" lon="24.
      <ele>19</ele>
      <time>2013-05-24T15:46:43.18</time>
    </trkpt>
    <trkpt lat="60.16121833333333" lon="24.5458">
      <ele>19</ele>
      <time>2013-05-24T15:46:43.53</time>
    </trkpt>
    <trkpt lat="60.16117166666667" lon="24.545826666666667">…
```

Datan ominaisuuksia

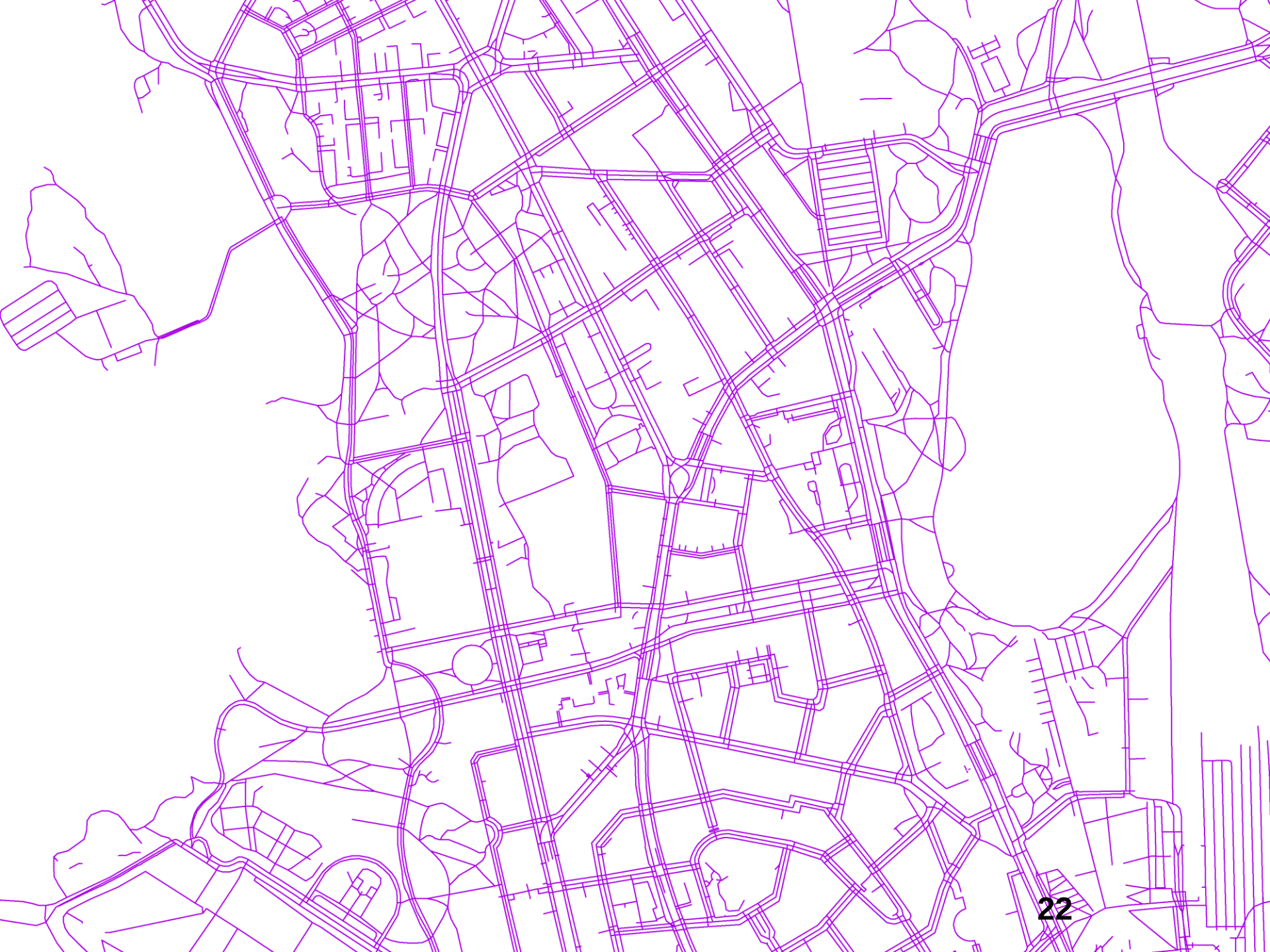
- 23 290 liikuntasuoritusta, 1994 käyttäjää (HKI 36 757, 2424)
- 17.4.2010–21.11.2012
- 80% validi aikaleima
- 65% käyttäjistä tallensi enintään 10 suoritusta ja 5 % käyttäjistä vastasi 50%:sta suorituksia







Suosioon perustuva reititys verkolla



OSM-verkon esikäsittely

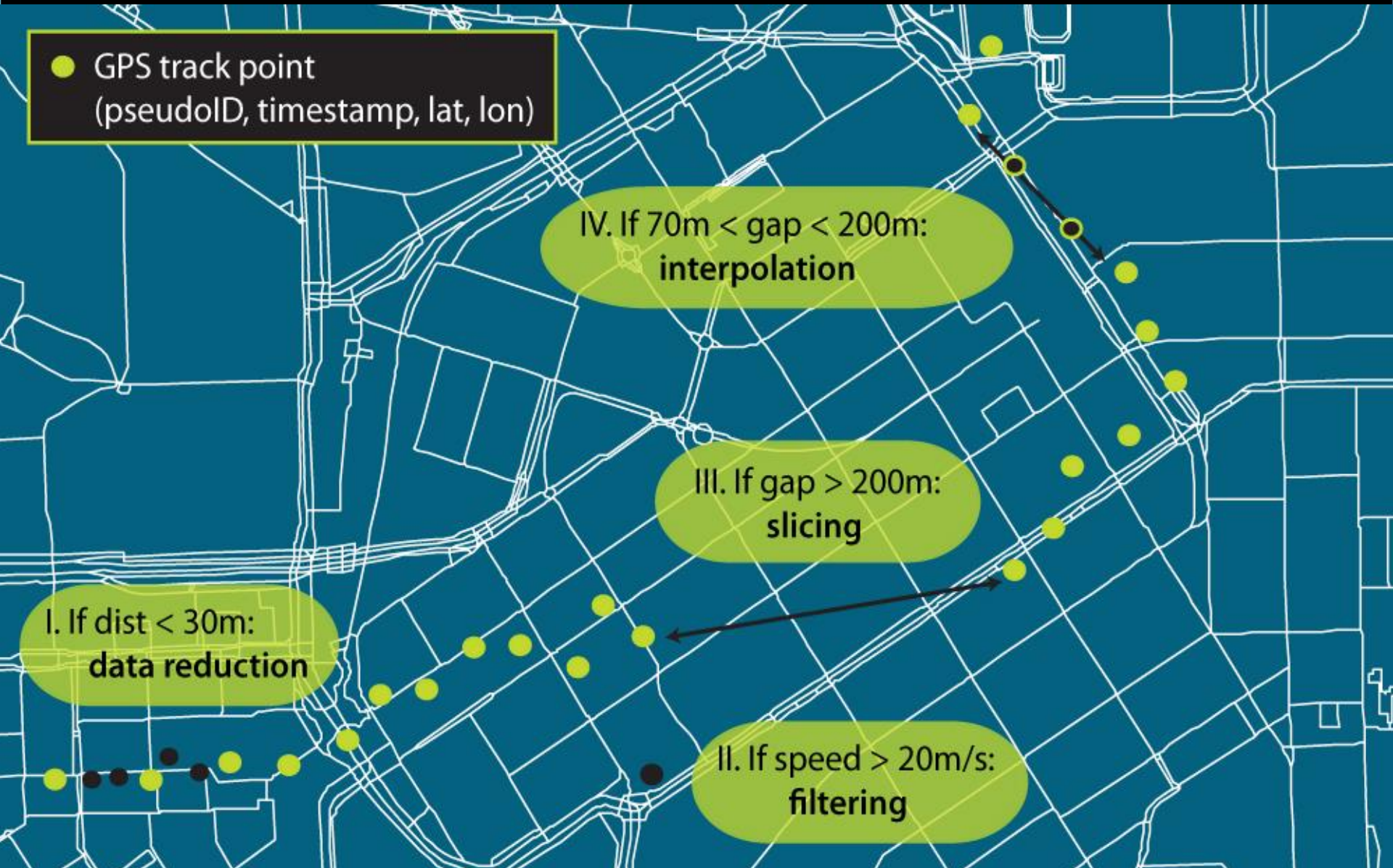
Highway type (key=highway)	Requirements related to additional tags of access permissions[a]
Cycleway	No additional requirements
(all)	Keep if Bicycle=yes |designated |official |lane
Primary[b], secondary[b], tertiary[b], trunk[b], residential, living_street, unclassified, track	Drop if Bicycle=no or (tunnel=yes and (access=no |private))[c]
Service	Drop if Bicycle=no or access=private |no or service=parking_aisle

[a] Key=value; if many possible values are related to the same key, they are separated by |

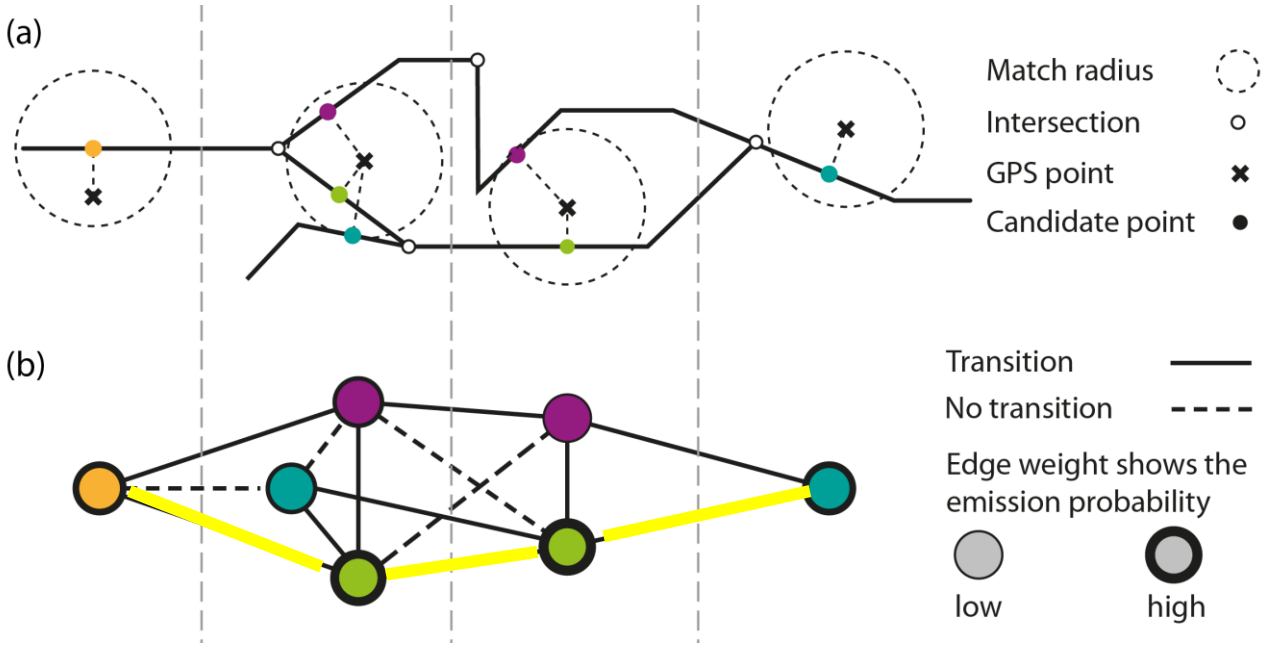
[b] Also with _link-tags

[c] The usage of additional tags access=private and access=no was very heterogeneous, and hence they have not been excluded more comprehensively

GPS liikeratojen esikäsittely



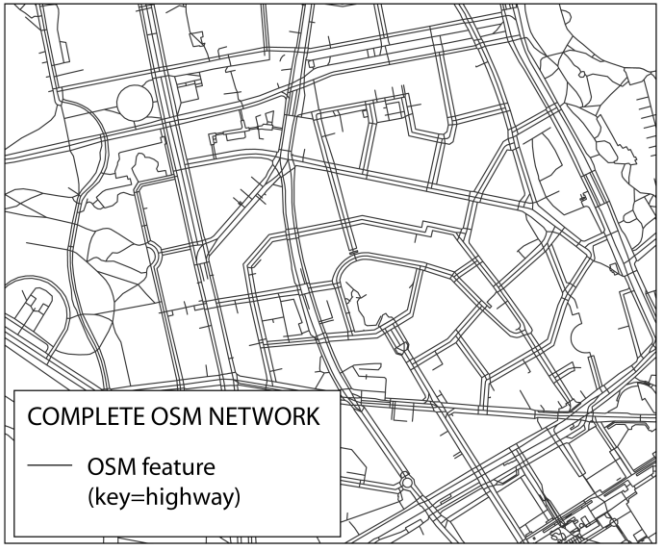
Map matching

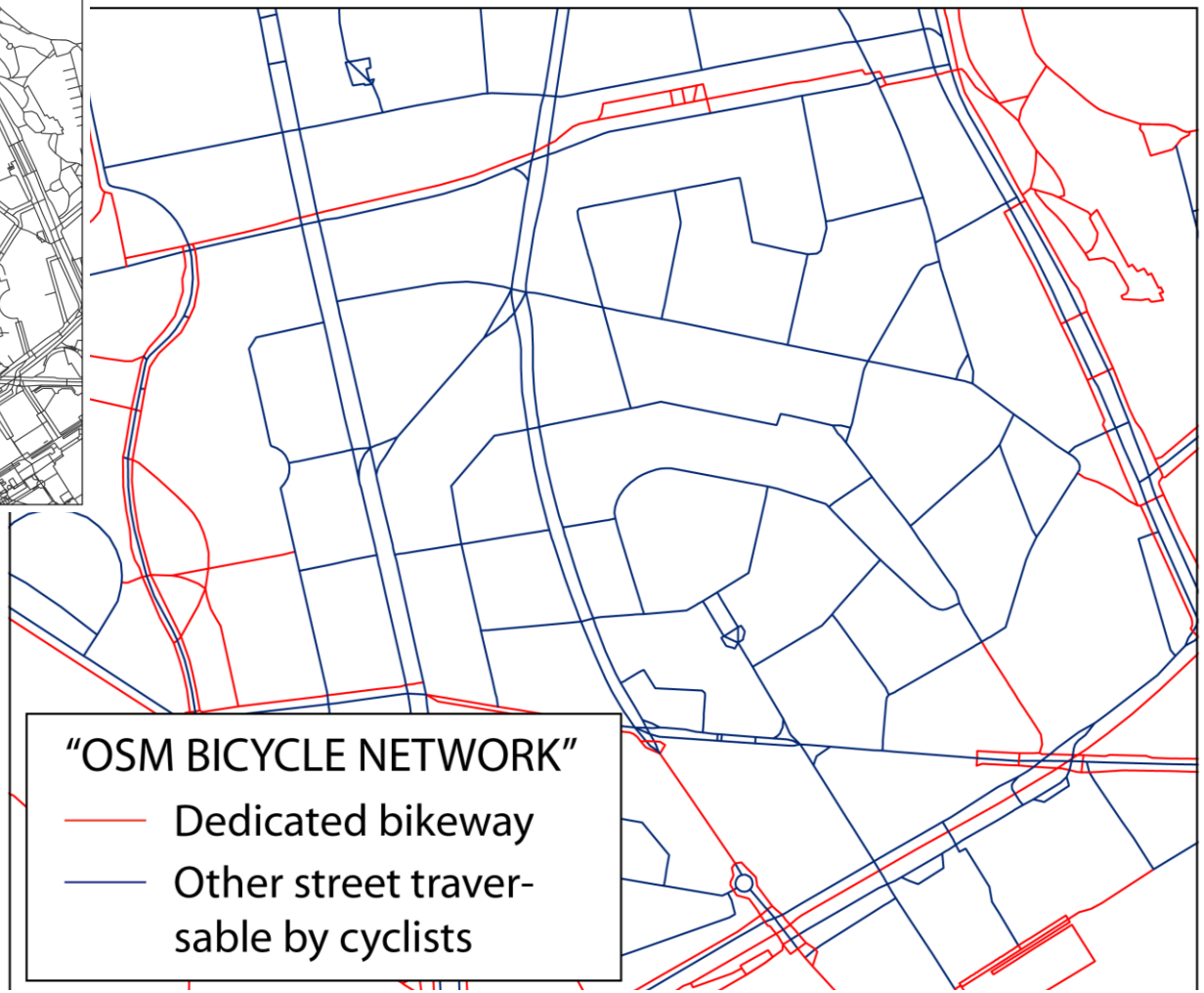


- **Markovin piilomalli (HMM, Hidden Markov Model)**
  - etsitään todennäköisin reitti hyödyntäen etäisyyden ohella verkon topologiatietoa

MML PAIKKA-TIETO-KESKUS FGI

(a)





- Painotetaan pyöräteitä osoittamalla näille suuremmat emissio- ja siirtymätodennäköisyydet

Table 2 Classification of bikeways in the study area based on OSM tagging

Highway type	Requirements related to additional tags[a]
Cycleway (all)	No additional requirements Cycleway[b]=lane |shared_busway |opposite_lane, bicycle=yes |designated |official

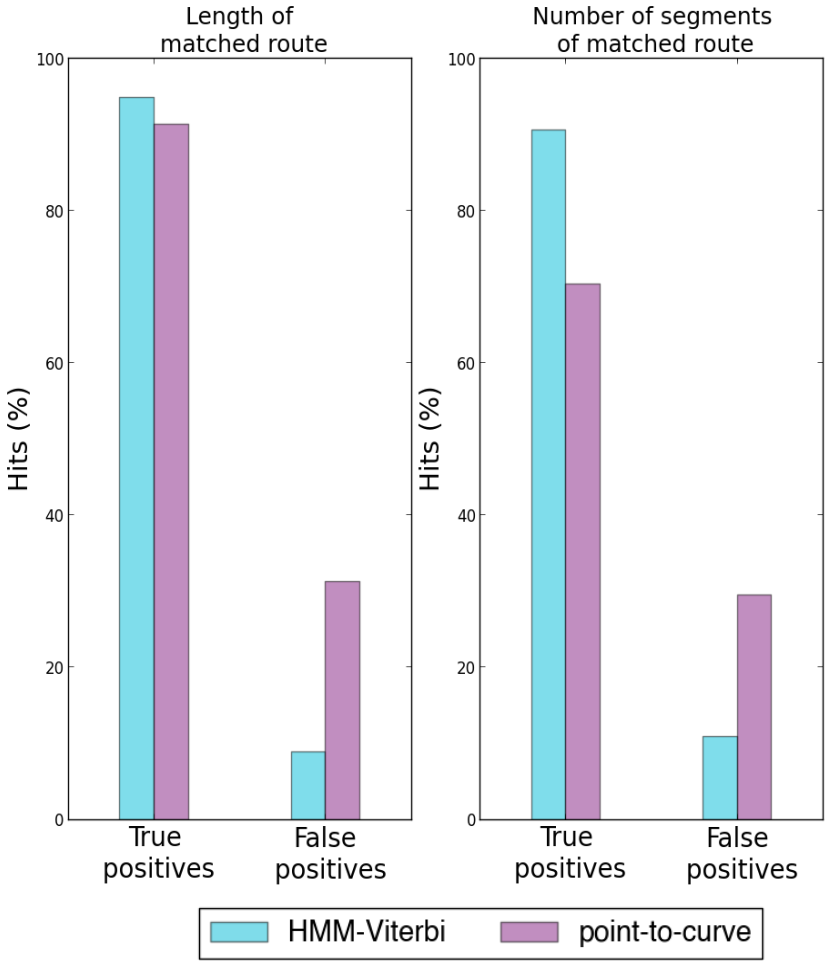
Map matching

- ”Point-to-curve”
  - Yhdistetään jokainen GPS piste lähimpään tiesegmenttiin 30 metrin säteellä
  - Ongelma: rinnakkaiset väylät
    - Poikkeus: jos 20 metrin säteellä on vähintään yksi pyörätie*, piste yhdistetään lähimpään pyörätiehen

* **highway = cycleway**
bicycle = yes or designated or official
cycleway* = lane or share_busway
cycleway = opposite_lane

*** mahdollinen lisämääre left/rigth/both**

Map matching

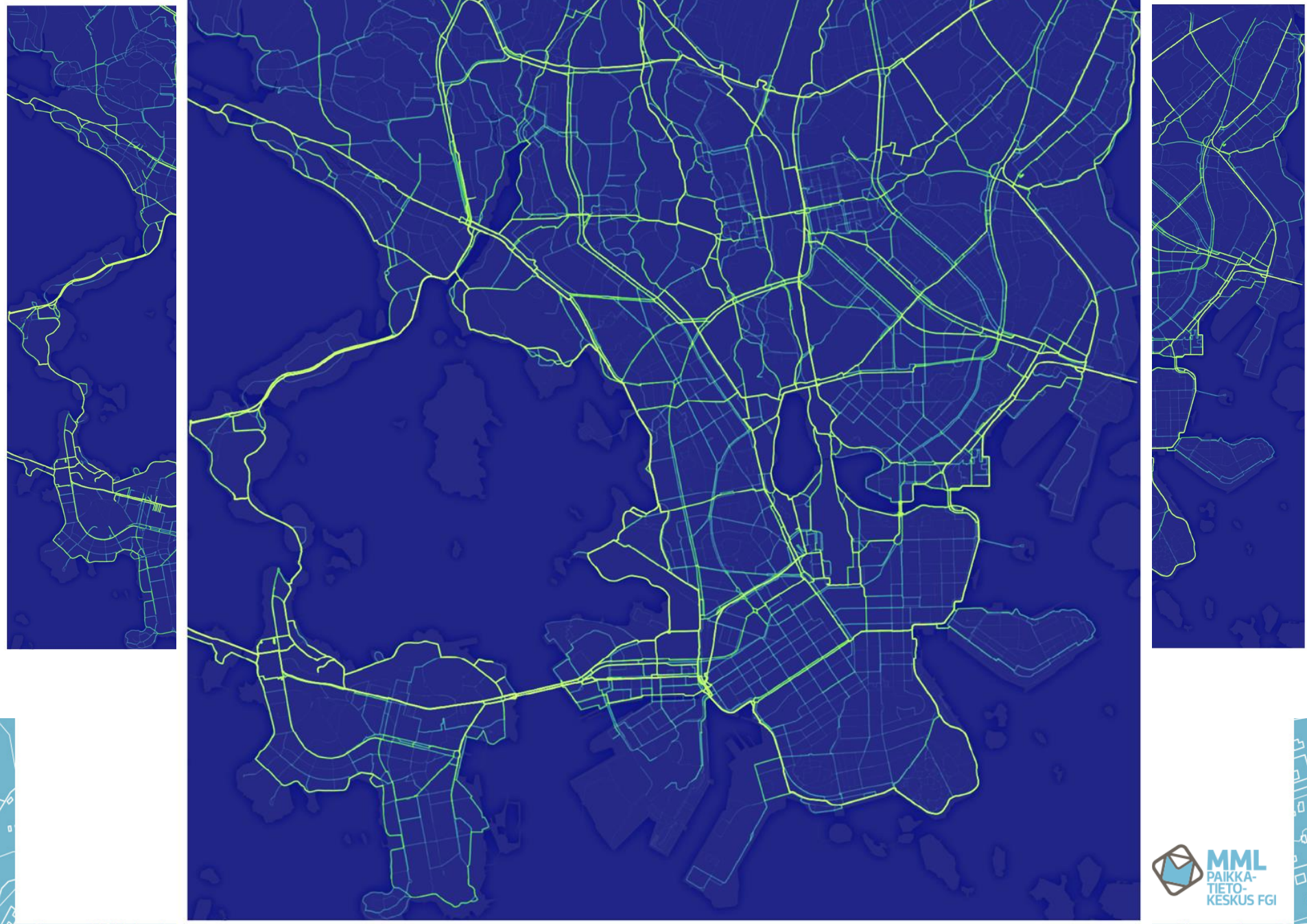


HMM - virheet

- Molemmin puolin katua rinnakkaiset pyörätiet, joilla pyöräily sallittu kahteen suuntaan
- Pyöräily jalkakäytävällä esim. vastavirtaan yksisuuntaisella kadulla
- Muutokset pyöräilyverkossa
- Risteykset - transitiotodennäköisyys

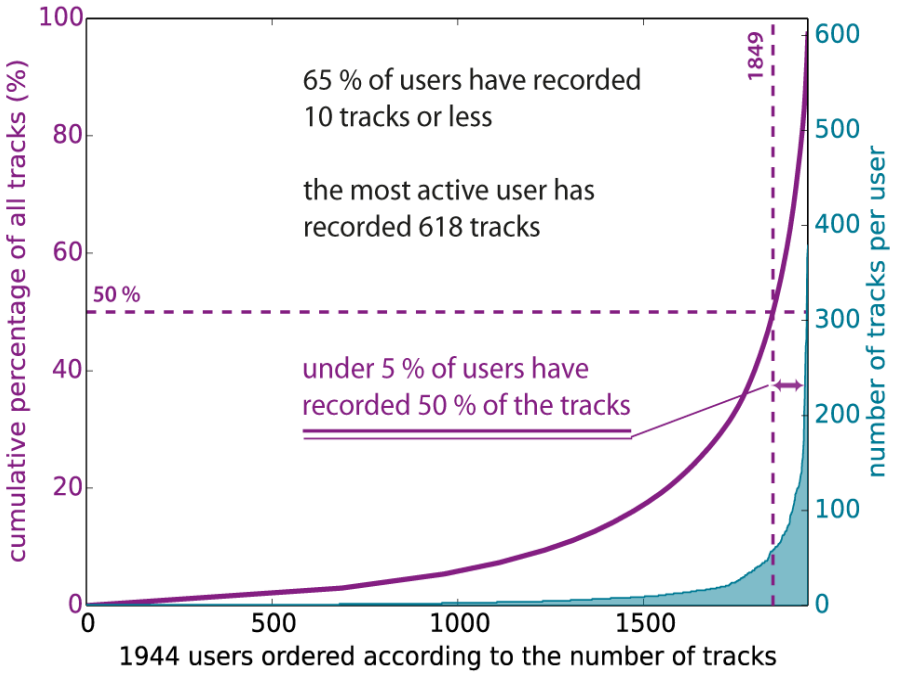
MML PAIKKA-TIETO-KESKUS FGI

(a)



MML
PAIKKA-
TIETO-
KESKUS FGI

Linkkien kustannukset



A. **Pyöräilijöiden** lkm

B. **Suoritusten** lkm

C. **Suoritusten** lkm kerrottuna Simpsonin **diversiteetti-indeksillä**

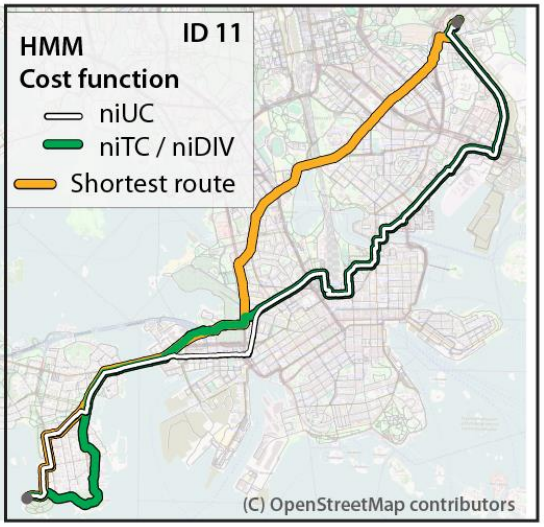
$$niUC\ (normalized\ inverse\ User\ Count) = \left(\frac{u}{l}\right)^{-1}$$

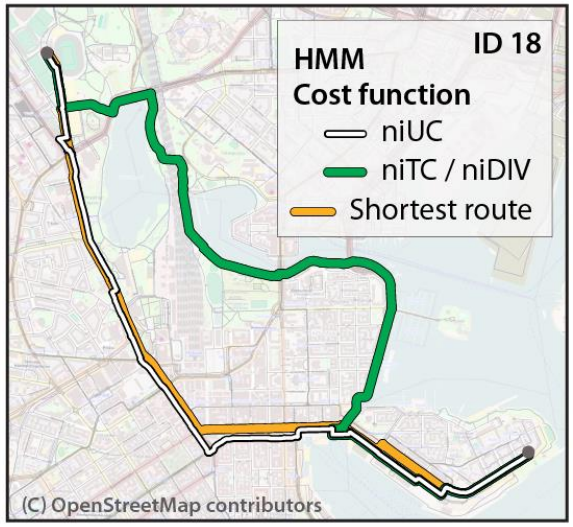
$$niTC\ (normalized\ inverse\ Track\ Count) = \left(\frac{t}{l}\right)^{-1}$$

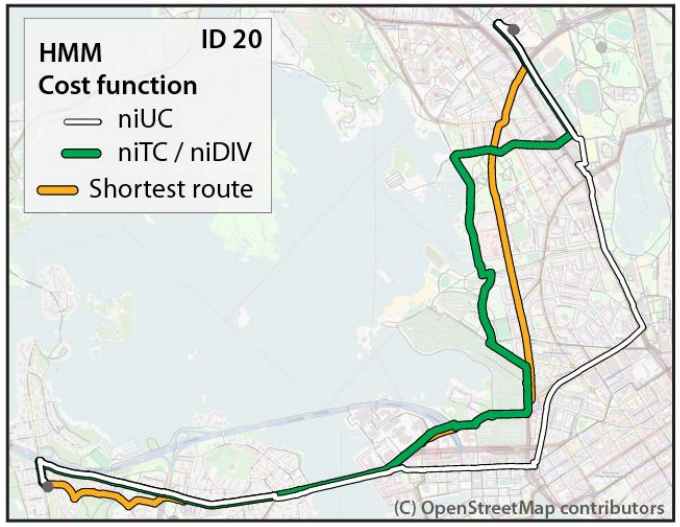
$$niTS\ (normalized\ inverse\ Tracks\ x\ Simpson) = \frac{l}{t\left(1-\sum p_i^2\right)}$$

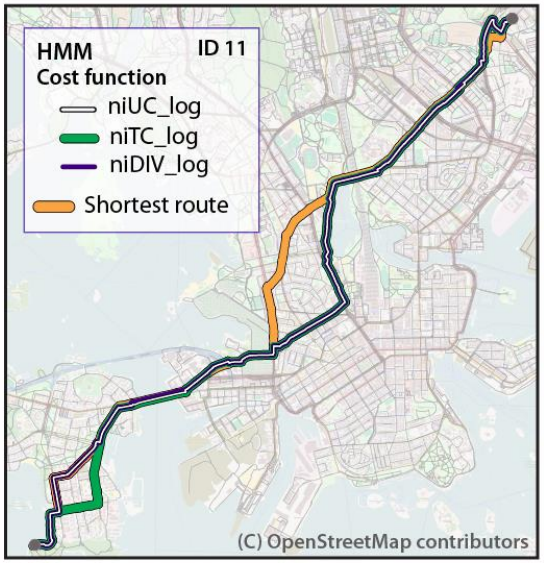
(kuvaa suoritusten jakautumista pyöräilijöiden kesken)

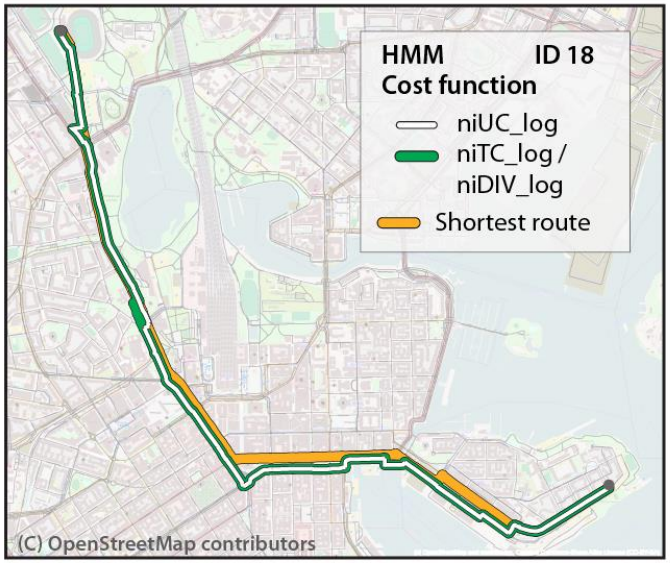
MML PAIKKA-TIETO-KESKUS FGI

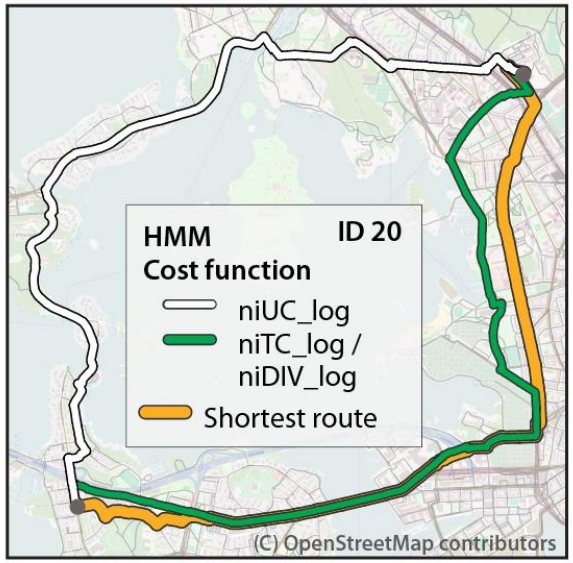












MML PAIKKA-TIETO-KESKUS FGI

Haasteita

- Luotettavuus
  - Ei formaaleja laatuvaatimuksia
- Heterogeenisyys
- Aineistojen yhteensopivuus
- Yksityisyyden suoja

- Koskevat yleisemminkin joukkoistetun datan käyttöä (esim. reitityksessä)

MML PAIKKA-TIETO-KESKUS FGI

Optimaalinen reitti – kenelle?

> Primarily the **generalizability** of MapMyFitness data must be thoughtfully considered before use in research.

> In-situ data may be […] **biased toward specific types of users**.

> This application captures a quite **small portion of the population** as Strava users.

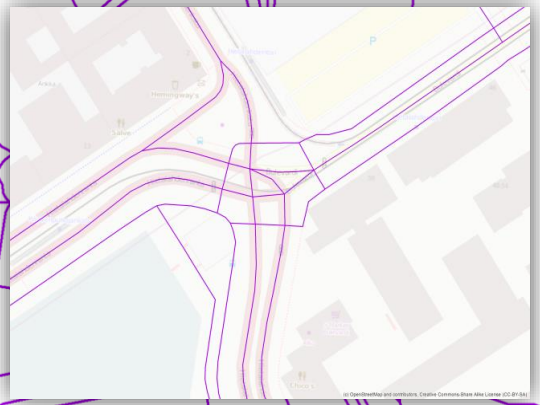
> Planners need to understand the **representation** of each dataset used, and its limits to a particular application.

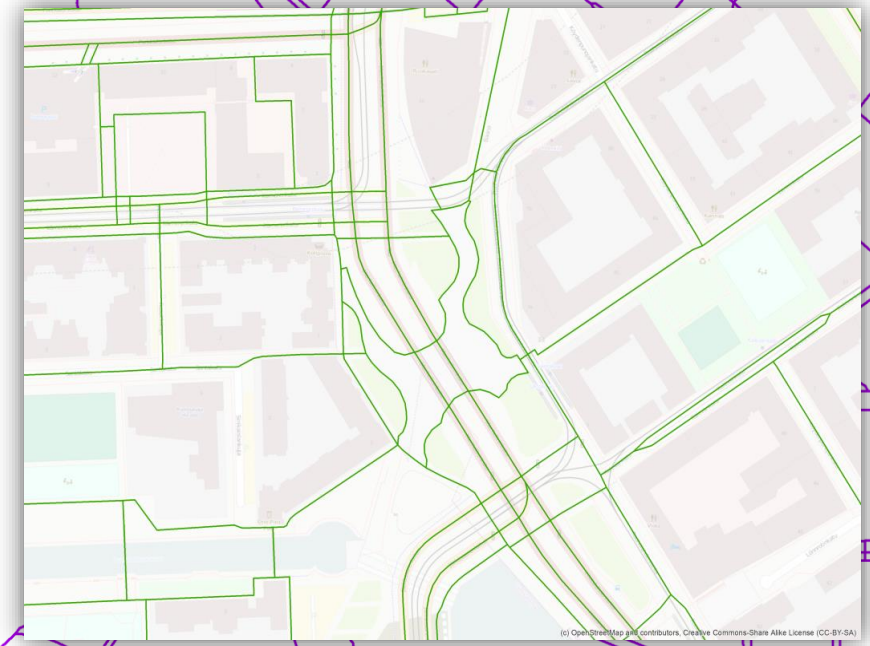
> […] where bicyclists ride, primarily for **fitness purposes**.

> Though the marketing focus of the Strava app is oriented towards fitness, **it is likely that its users log trips for other purposes, as well**.

Cortés et al. Sport Trackers and Big Data: Studying user traces to identify opportunities and challenges. Research Report RR-8636, INRIA Paris. 2014.
Ferrari and Mamei. Identifying and understanding urban sport areas using Nokia Sports Tracker. Pervasive and Mobile Computing 9(5): 616-628, 2013.
Griffin and Jiao. Where does bicycling for health happen? Analysing volunteered geographic information through place and plexus. Journal of Transport and Health, 2015.
Hirsch et al. Using MapMyFitness to Place Physical Activity into Neighborhood Context. Frontiers in Public Health 2:19, 2014.

”OSM pyöräilyverkko”





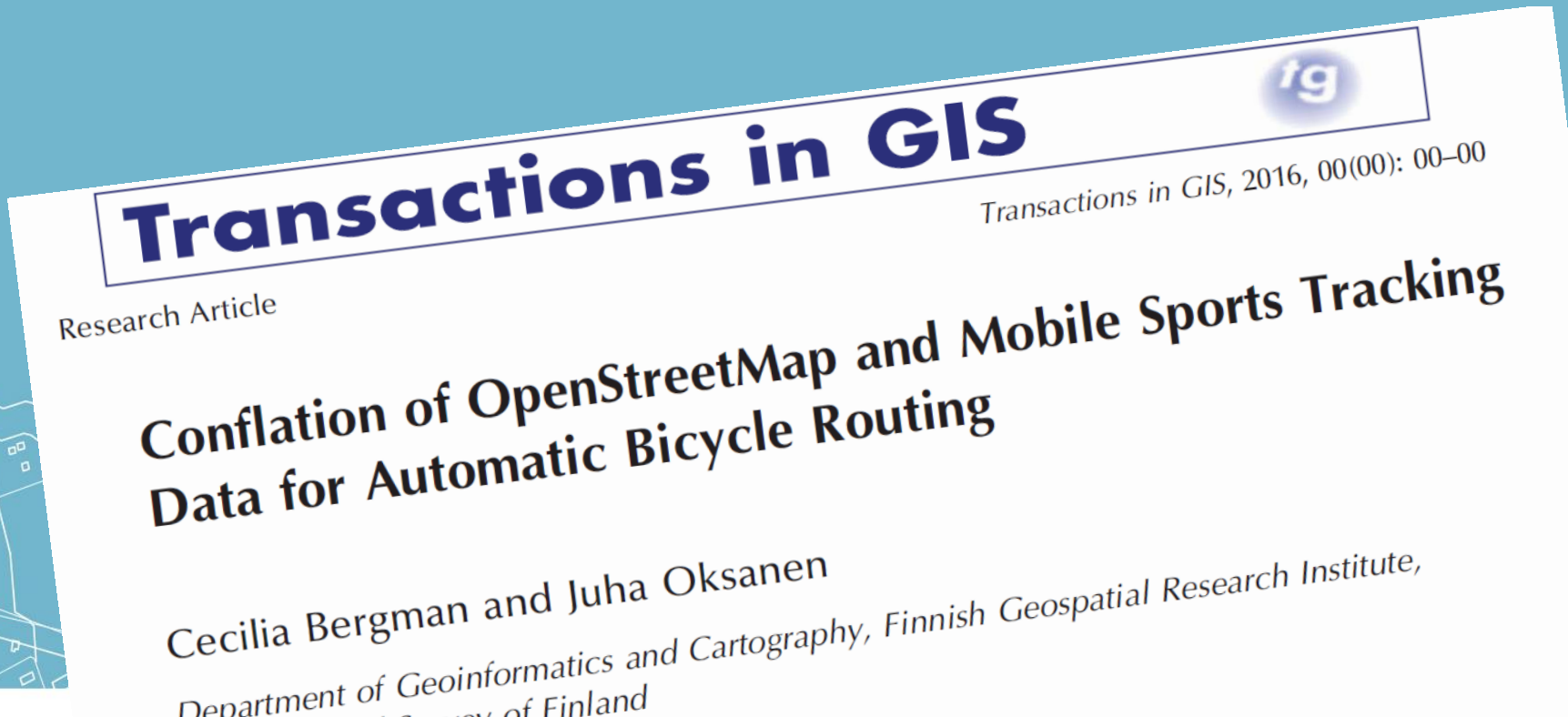
- Ilmainen
- Liikenneväylien luokitus (+ lisätiedot) voi mahdollistaa eri ryhmien tarpeiden mukaan personoidut reitityspalvelut
- Liikenneverkon päivitykset perustuvat pääosin GPS-liikeratoihin eli todelliseen liikkumiseen
- Spatiaalinen ja erit. semanttinen tarkkuus?

MML PAIKKA-TIETO-KESKUS FGI

Kiitos!

Juha Oksanen
Sports Tracker
Tekes

cecilia.bergman@nls.fi



Transactions in GIS

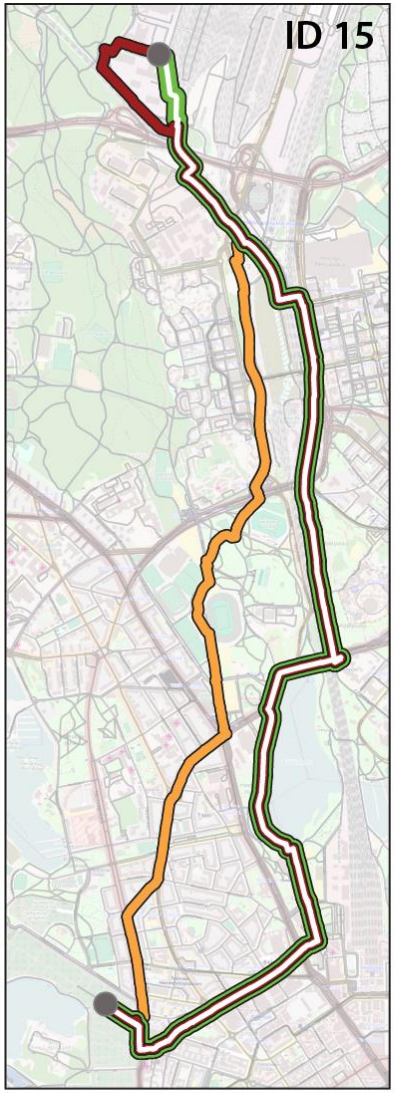
Transactions in GIS, 2016, 00(00): 00–00

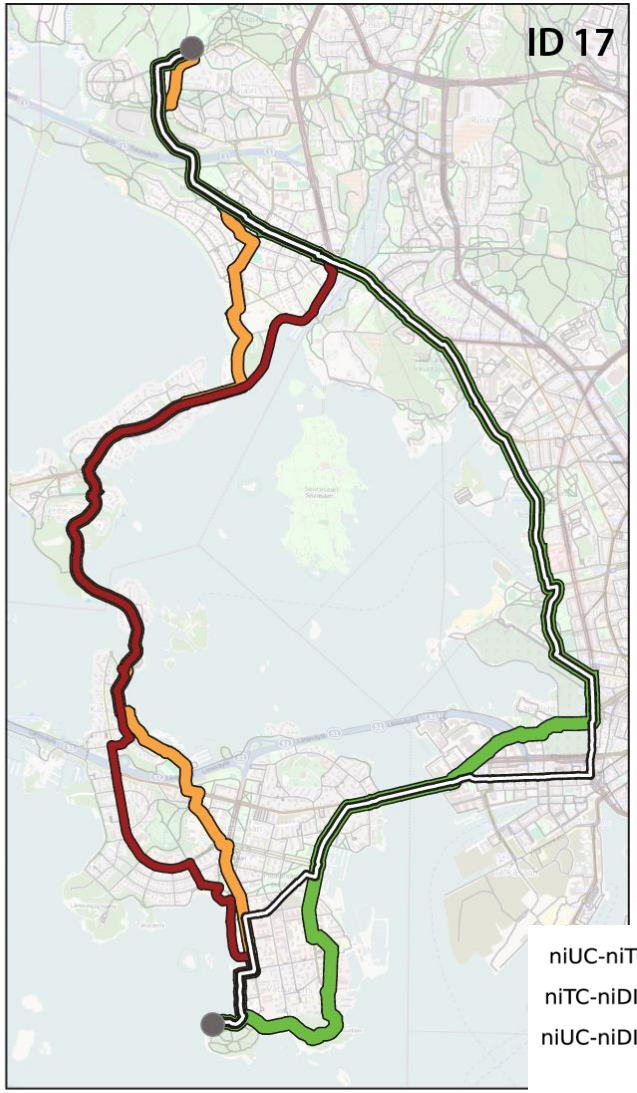
Research Article

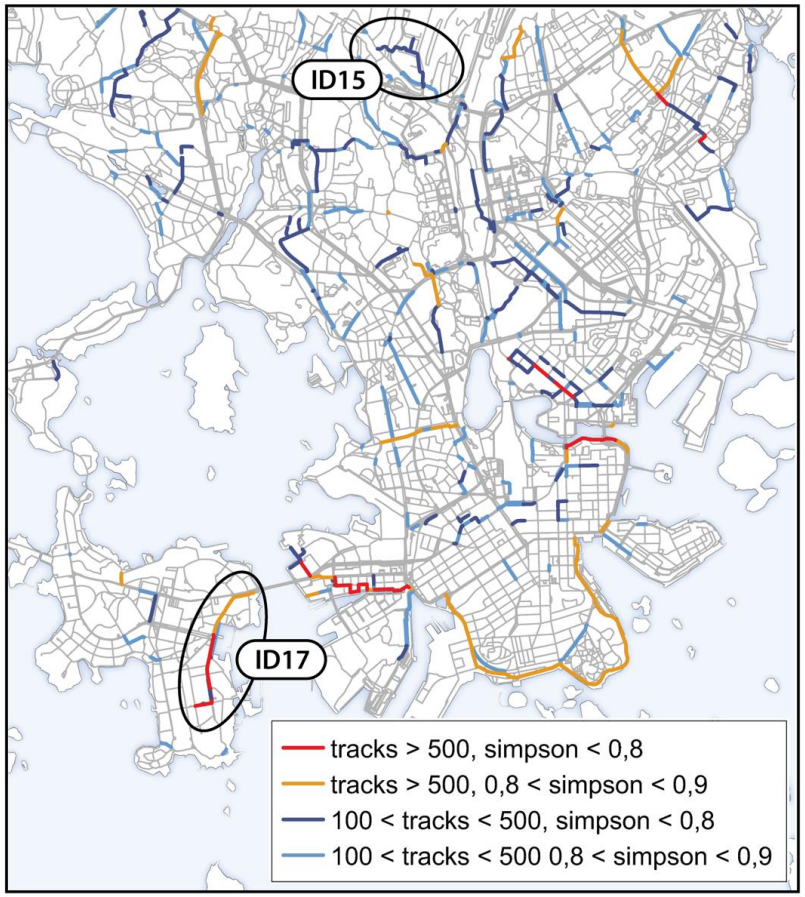
Conflation of OpenStreetMap and Mobile Sports Tracking Data for Automatic Bicycle Routing

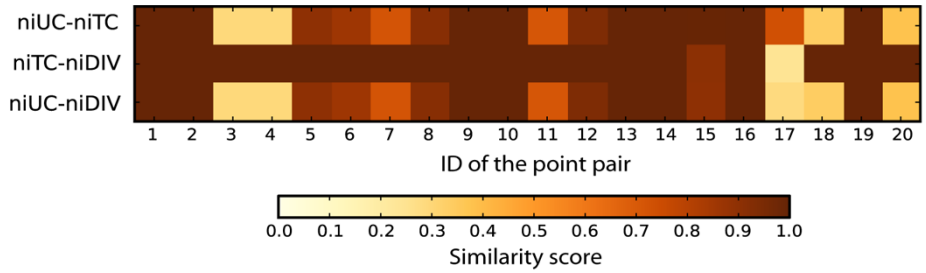
Cecilia Bergman and Juha Oksanen

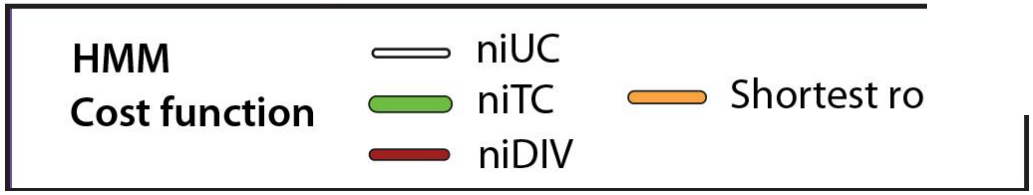
Department of Geoinformatics and Cartography, Finnish Geospatial Research Institute,


ID 15


ID 17


ID15
ID17
tracks > 500, simpson < 0,8
tracks > 500, 0,8 < simpson < 0,9
100 < tracks < 500, simpson < 0,8
100 < tracks < 500 0,8 < simpson < 0,9

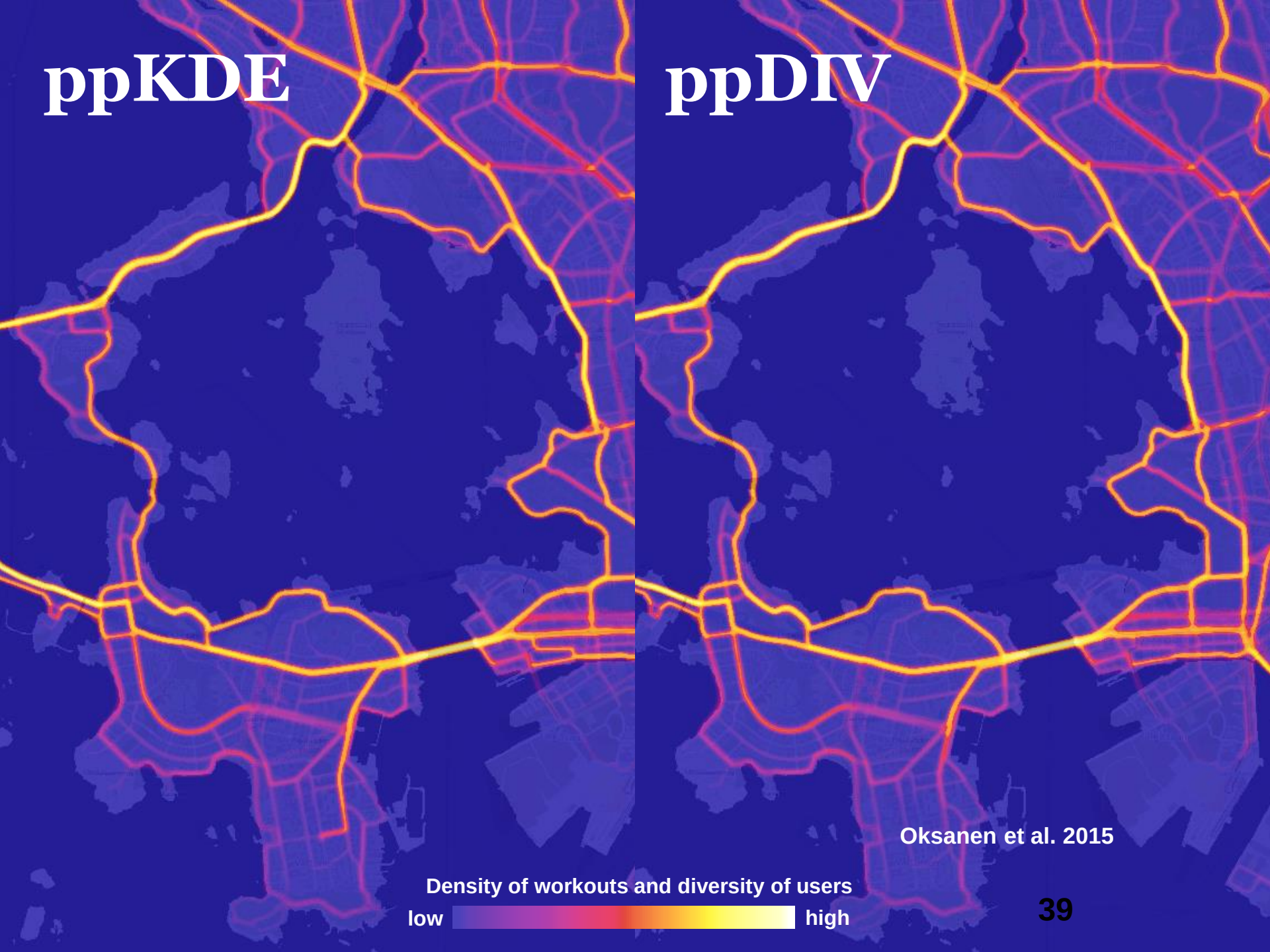

niUC-niTC
niTC-niDIV
niUC-niDIV
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
ID of the point pair
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
Similarity score


HMM
Cost function
niUC
niTC
niDIV
Shortest ro

Mahdollisuudet

- Monipuolisuus
- Yksityiskohtaisuus
- Paikallisuus
- Reaaliaikaisuus
- Päivitys nopeaa ja edullista

MML PAIKKA-TIETO-KESKUS FGI


ppKDE
ppDIV
Oksanen et al. 2015
Density of workouts and diversity of users
low
high