Applications with geo- and accelerometer data

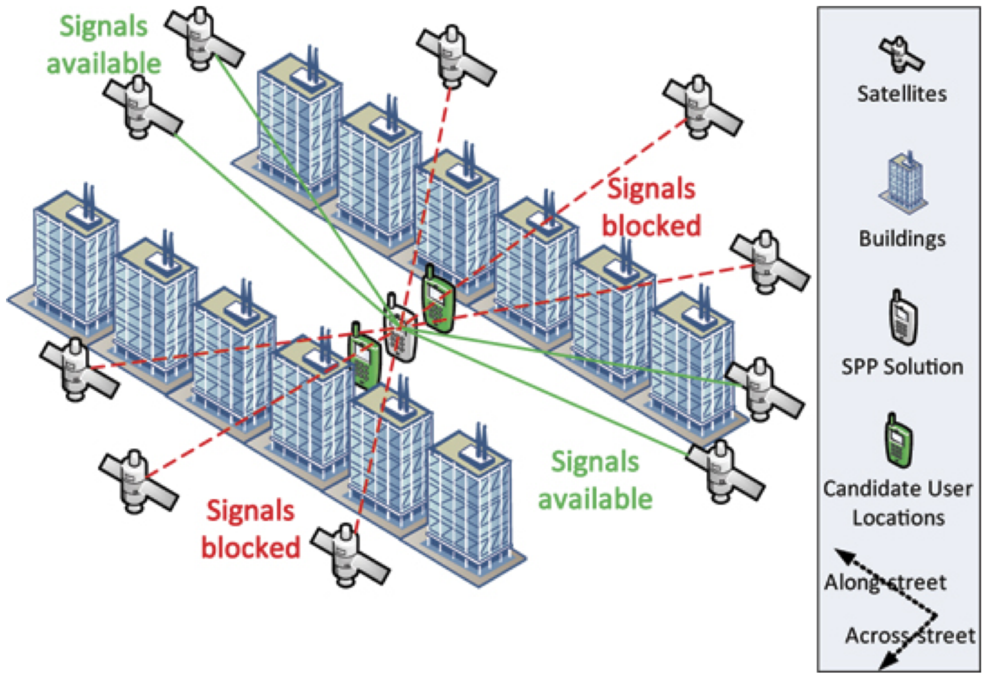
Day 4 Advanced Survey Design

Peter Lugtig

p.lugtig@uu.nl www.peterlugtig.com

Geolocation

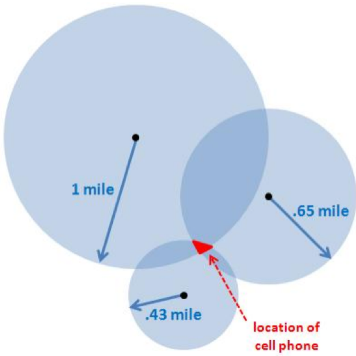
- GPS
  - Provides coordinates in longitude & Latitude
  - Based on distance (= rate x time) to at least 4 satellites
  - Newest generation has accuracy within 30 centimeters
  - Works without cell/Internet connection
  - Performs worse in ‘urban canyons’, indoors, & underground
  - Constant tracking is very battery-draining



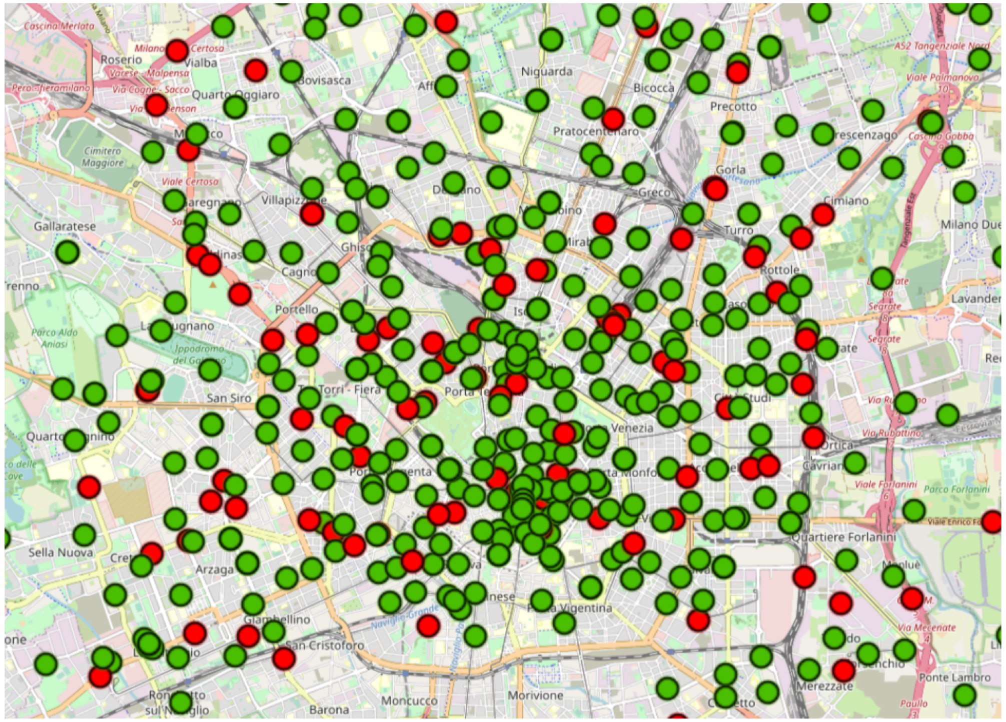
Source: https://www.gpsworld.com/wirelesspersonal-navigationshadow-matching-12550/

Geolocation

- GPS
- Cellular network
  - Multilateration of radio signals between (several) cell towers
  - Works even if GPS is turned off
  - If there is no signal then location information will be missing



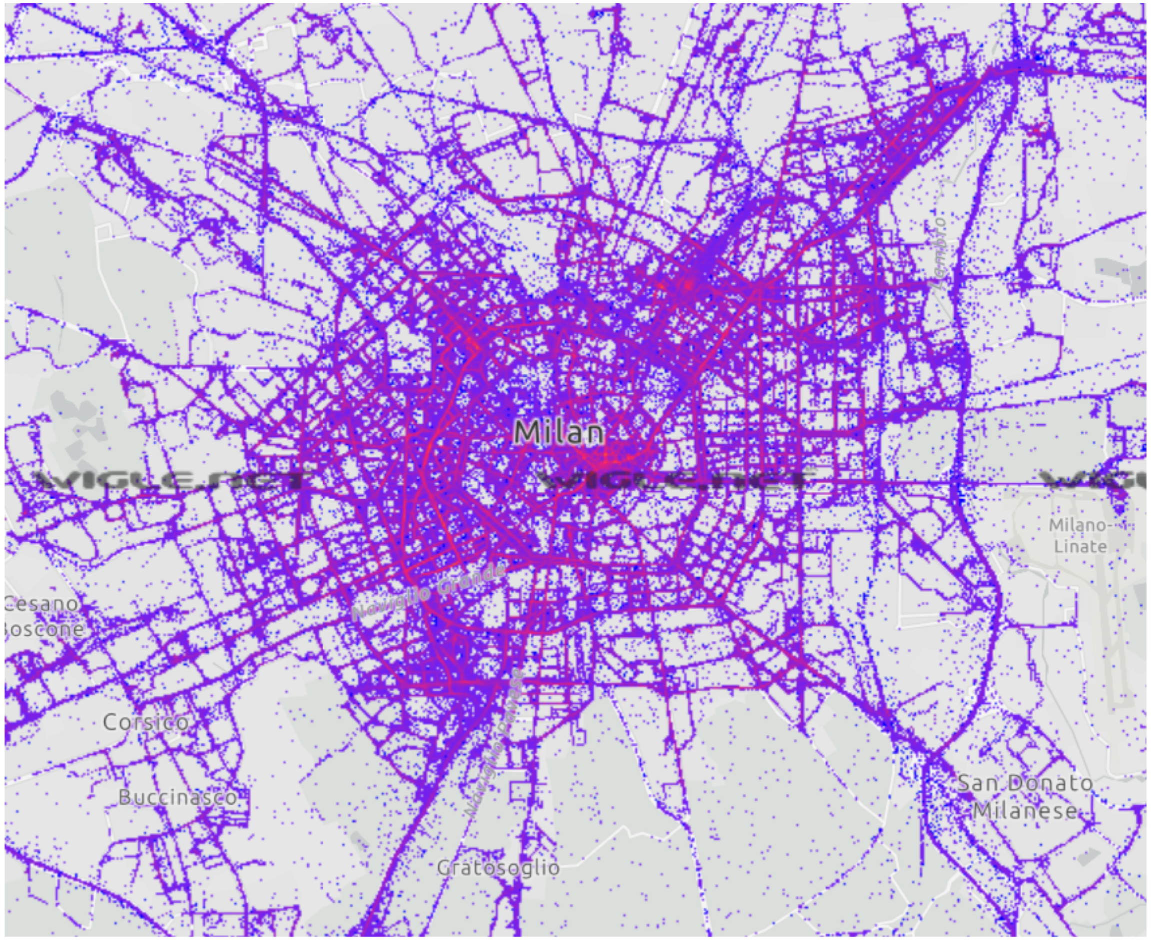
Source: https://searchengineland.com/cell-phone-triangulation-accuracy-is-all-over-the-map-14790



Source: https://www.cellmapper.net

Geolocation

- GPS
- Cellular network
- Wi-Fi
    - Inferring location from Wi-Fi access points (AP)
    - Can overcome problem of ‘urban canyons’ and indoor tracing

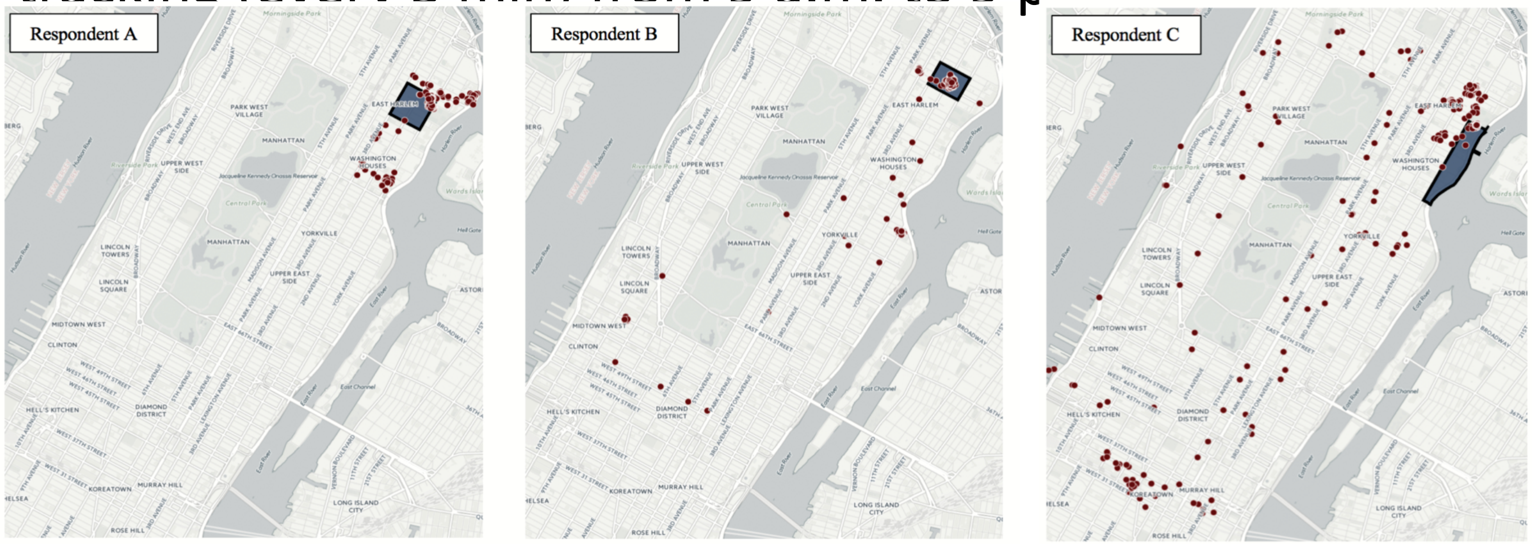


Source: https://www.wigle.net

Example: Aging in activity space

(York Cornwell & Cagney 2017, 2020)

- *Real-time Neighborhoods and Social Life Study* (RNSL)
- 60 participants aged 55+ in NYC provided with iPhones to carry for 7 days
- GPS-tracking (every 5 min) from 9 a.m. to 9 p.m and four EMAs per day



Example: Aging in activity space

(York Cornwell & Cagney 2017, 2020)

- Activity spaces vary considerably in size
- Participants spent ~40% of their time outside their residential tracts
  - On average >10 min in 9+ tracts
- Activity spaces larger among younger and more advantaged social groups (i.e., whites, those with college degree, car owners)
- Participants with less education and lower incomes spend more time outside
  of their residential tracts
- Four main activities outside of residential tracts
  - Shopping, exercising, socializing, participating in social groups or activities
- Poverty rates in nonresidential tracts lower than in residential tracts
- Higher concentrated disadvantage in an area associated with higher odds of self-reporting pain

Example: How do people find work after prison?

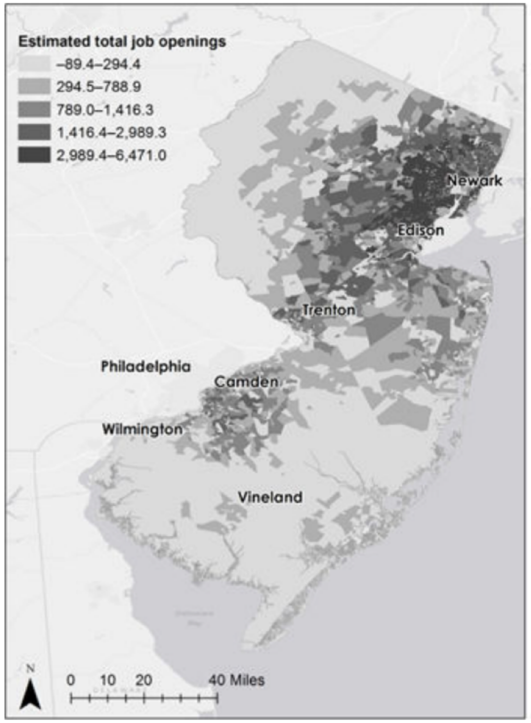
(Sugie 2018; Sugie and Lens 2017)

- Newark Smartphone Reentry Project (NSRP) 2012-2013
  - N = 133 with 8,000 daily observations (89% response, 1.5% noncompliance)
  - 3 months of data collection
- Men recently released from prison
  - Difficult group to follow due to unstable circumstances
- Loaner smartphones (Android)
- Surveys twice a day (EMA) about social interaction, job search & work, and emotional well-being
- Sensing
  - GPS location
  - Calls and messaging (encrypted)
- Survey triggered by calls/messages from new telephone numbers

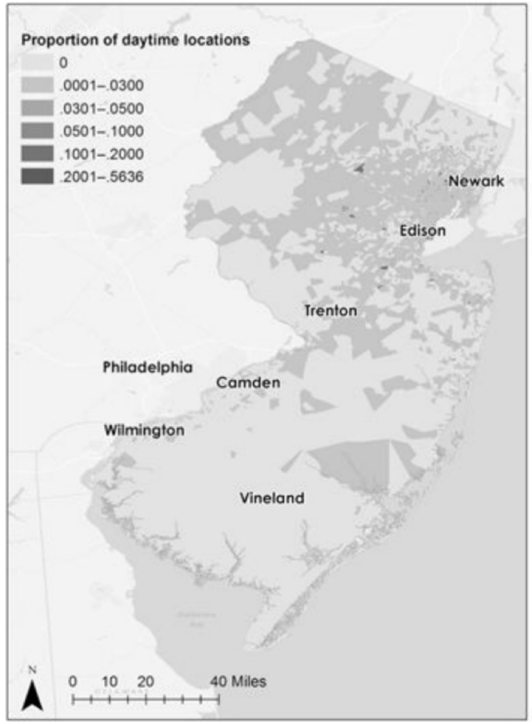
Example: How do people find work after prison?

(Sugie 2018; Sugie and Lens 2017)

- Spatial mismatch
  - Low-skilled, nonwhite job seekers within central cities, job opportunities in outlying areas
- Hypothesis
  - Parolees lack info on job openings, are geographically restricted, unable to travel to find work
- Findings
  - Residential mismatch lengthens time to employment
  - But mobility can compensate for residential deficits



Job openings

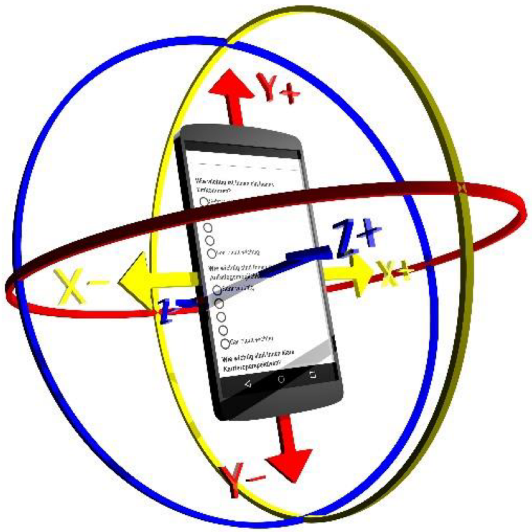


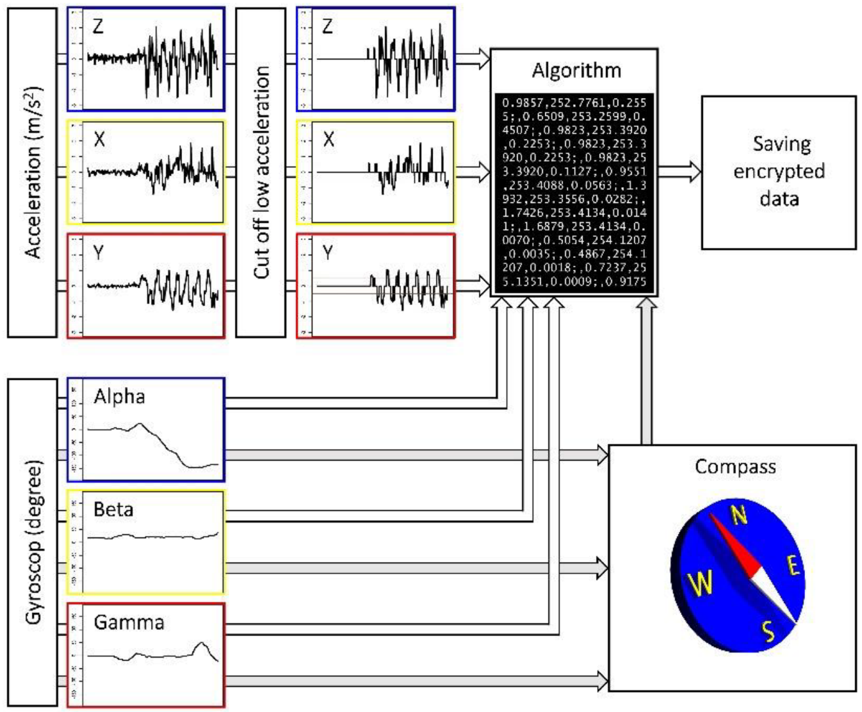
Daytime locations of parolees

Physical activity

- Accelerometer
- Gyroscope

Source: https://www.techradar.com/news/wearables/10-best-fitness-trackers-1277905





Schlosser et al. (2019)

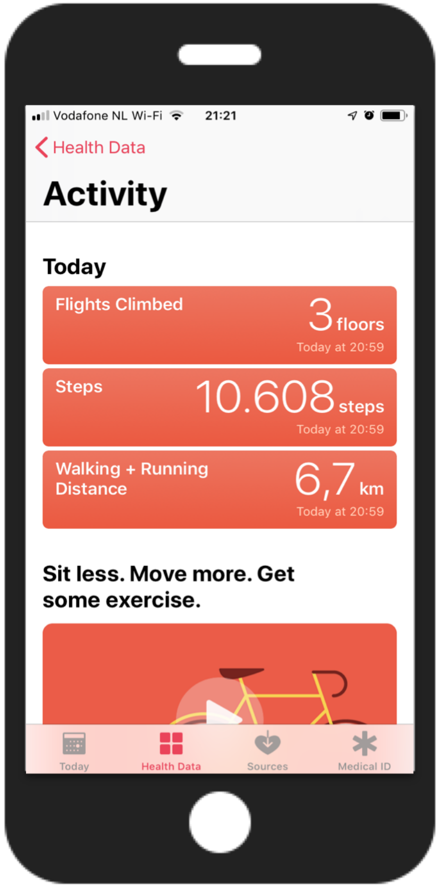
Sources: https://www.actigraphcorp.com/actigraph-wgt3x-bt/, https://www.activinsights.com/products/geneactiv/

Physical activity

- Accelerometer
- Gyroscope

  and

- Magnetometer
  - Serves as compass
- Barometer
  - Allows to track changes in elevation



physical activity

- Wearables
  - Wrist worn GENEActiv
  - Axivity ax3 at upper thigh
- (Total) Physical Activity
  - Time moderate-intense activity

- UK Millennium Cohort Study: Gilbert & Calderwood (2018)
- SHARE: Scherpenzeel, Angleys, & Weiss (2018)

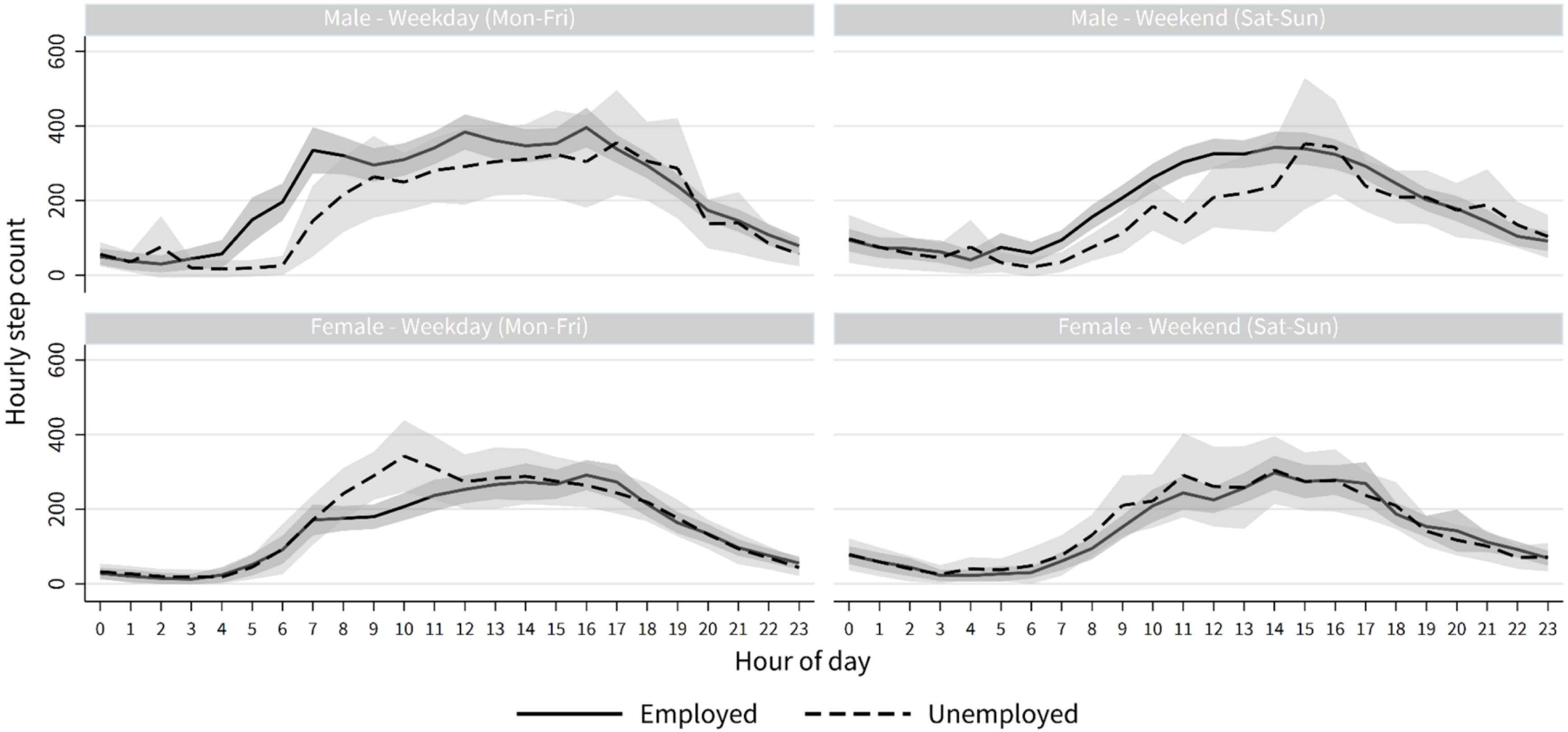
Example: What are the effects of unemployment?

(Kreuter et al. 2020)

- ~650 Android smartphone owners from German panel study "Labour Market and Social Security" (PASS) downloaded *IAB-SMART* app for 6 months
- Survey questions triggered by…
  - Schedule: *Q*s about affective impact of daily smartphone use, Big 5 personality, employment and job search activities, use of smartphones in everyday life, etc.
  - Geolocation: 400 job centers - *Q*s about visit to job center
- Five passive data collection modules:
  - Location using GPS, Wi-Fi, and cellular sensors every 30 min
  - Activity and means of transportation (e.g., walking, biking, riding in/on a motorized vehicle) using accelerometer and pedometer every 2 min
  - Call and texting behavior using phone and SMS logs
  - Use of apps installed on smartphone

Example: What are the effects of unemployment?

(Bähr et al. in preparation)



Kapteyn et al (2017)

Physical activity and obesity

Table 1 Subjective and objective distribution of physical activities: ages 50+

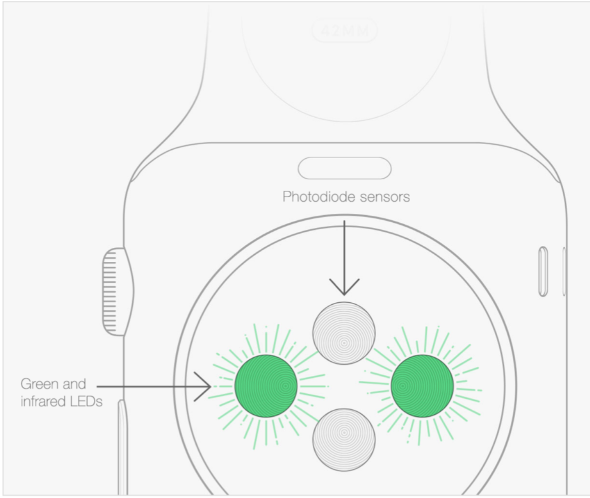
	Subjective			Objective		
	The Netherlands (LISS)	**The USA (UAS)**	**England (ELSA)**	**The Netherlands (LISS)**	**The USA (UAS)**	**England (ELSA)**
Inactive (%)	8	10	5	20	38	21
Mildly active (%)	21	32	23	20	22	18
Moderately active (%)	42	34	43	20	14	17
Active (%)	25	20	26	20	11	27
Very active (%)	3	5	3	20	14	17
Observations (n)	447	279	248	447	283	254
	$\chi^2(4)$	**P value**		**$\chi^2(4)$**	**P value**	
LISS vs UAS	13.54	0.0089		35.57	3.6e-07	
LISS vs ELSA	2.29	0.681		5.74	.2194	
UAS vs ELSA	13.42	0.010		32.6855	1.38e-06	

ELSA, English Longitudinal Study of Ageing; LISS, Longitudinal Internet Studies for the Social Sciences; UAS, Understanding America Study.

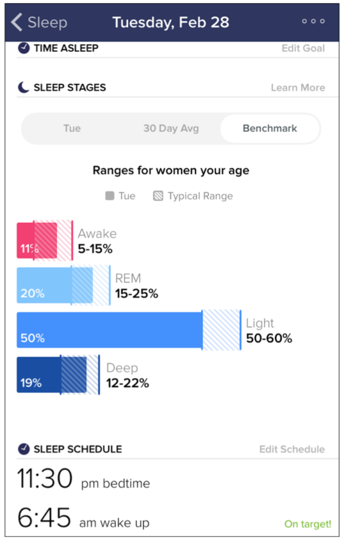
Kapteyn, A., Banks, J., Hamer, M., Smith, J. P., Steptoe, A., Van Soest, A., ... & Wah, S. H. (2018). What they say and what they do: comparing physical activity across the USA, England and the Netherlands. *J Epidemiol Community Health*, *72*(6), 471-476.

Heart-rate

- Most wristbands use LED-based system
  - Light “shines” onto skin, sensor detects blood volume changes
  - “... finely-tuned algorithms are applied to measure heart rate automatically and continuously...” (https://help.fitbit.com/articles/en_US/Help_article/1565)
  - Samsung Galaxy S uses similar system
- Used in combination with accelerometer determine sleep phases (e.g., on Fitbit)



Source: https://exist.io/blog/fitness-trackers-heart-rate/



Source: https://help.fitbit.com/articles/en_US/Help_article/2163

Accelerometer data structure

X	Y	Z	time
,0004	,0045	9,8158	1485181973469
,0012	,0047	9,7947	1485181973485

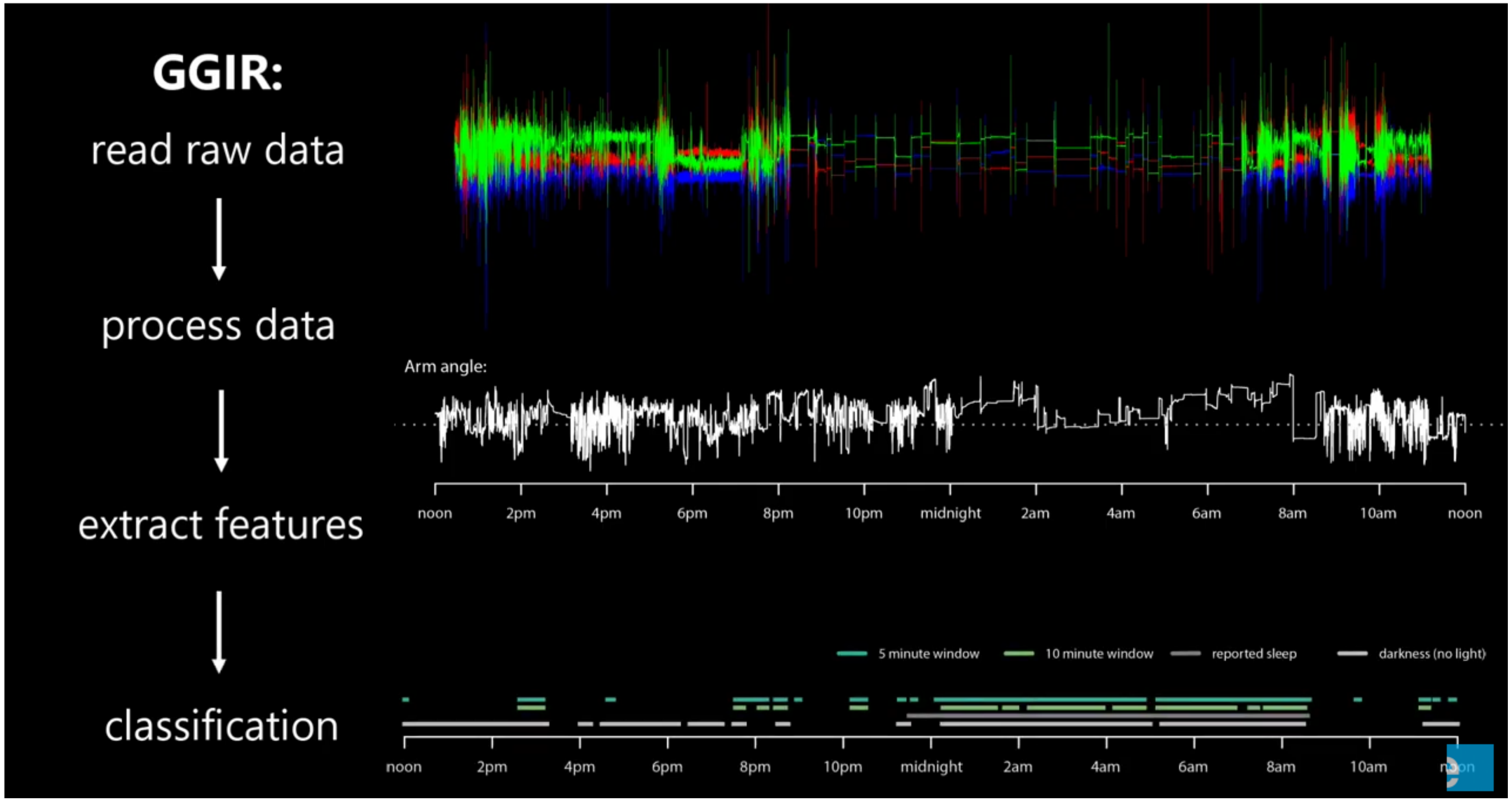
Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra and Jorge L. Reyes-Ortiz. A Public Domain Dataset for Human Activity Recognition Using Smartphones. 21st European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, ESANN 2013. Bruges, Belgium 24-26 April 2013.

Accelerometer data structure

X	Y	Z	time
,0004	,0045	9,8158	1485181973469
,0012	,0047	9,7947	1485181973485
Not sideways	Slow Forward	Up, down	16 ms diff (60 Hz)

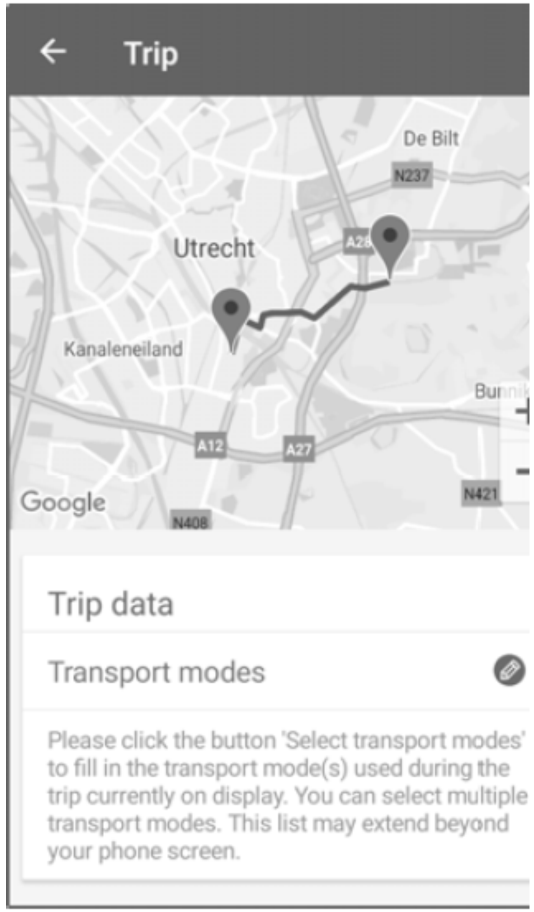
Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra and Jorge L. Reyes-Ortiz. A Public Domain Dataset for Human Activity Recognition Using Smartphones. 21st European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, ESANN 2013. Bruges, Belgium 24-26 April 2013.

Accelerometer data processing



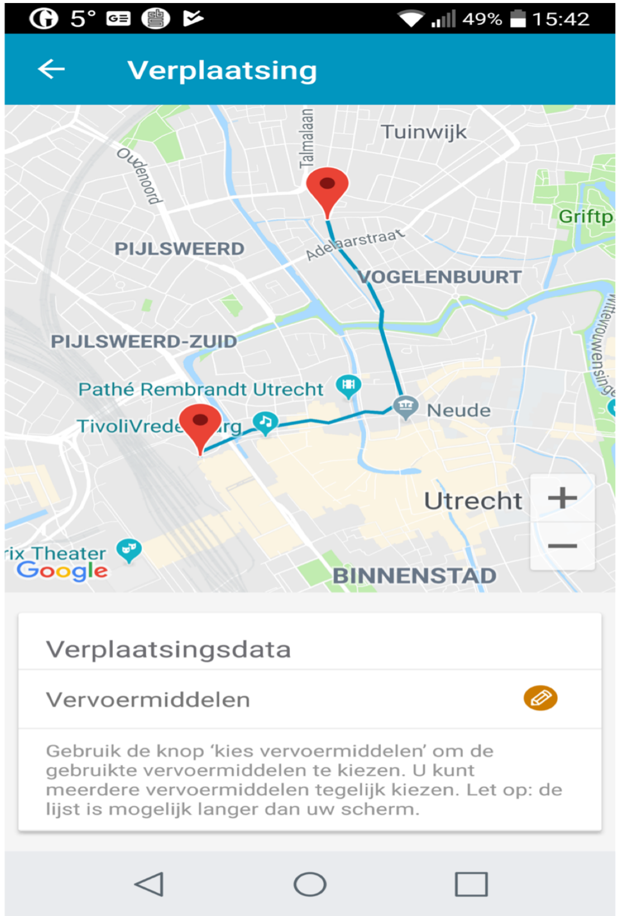
In more detail: smartphone travel study

- OLD: OdIN (On-route In the Netherlands)
  - Web-diary study
  - 2 days of travel data
  - start, endtime, location, travel modes, etc.
- New: TABI smartphone travel app
  - 1 week continuous tracking
  - Wifi/GPS
    - Every minute when stationary
    - Every second when moving

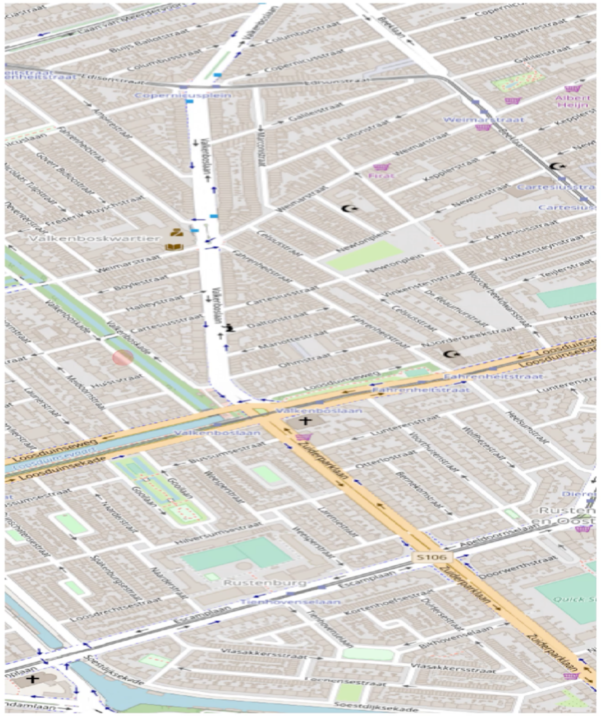


The Tabi app (1)

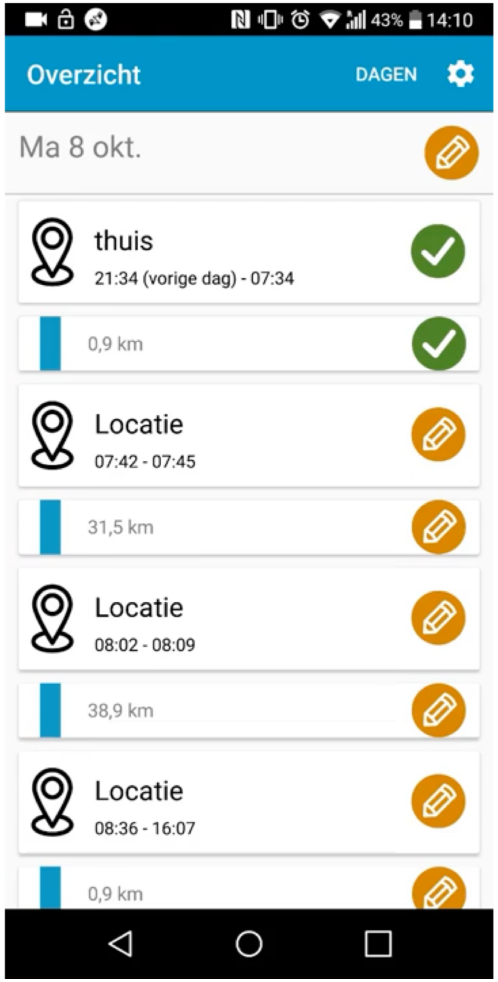
- Develop everything ourselves
  - Transparency, control
  - Open-source: https://gitlab.com/tabi/tabi-app
  - Modular design: parts can be recycled
- App should work on all smartphones
  - In practice: Android and ioS from ~2015
  - App should not drain battery
- Low respondent burden
  - Few (or no) questions

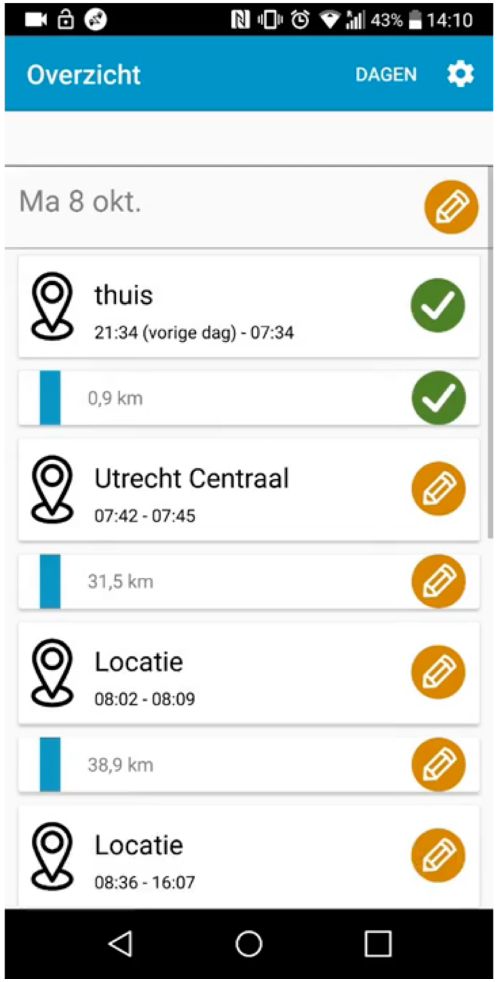


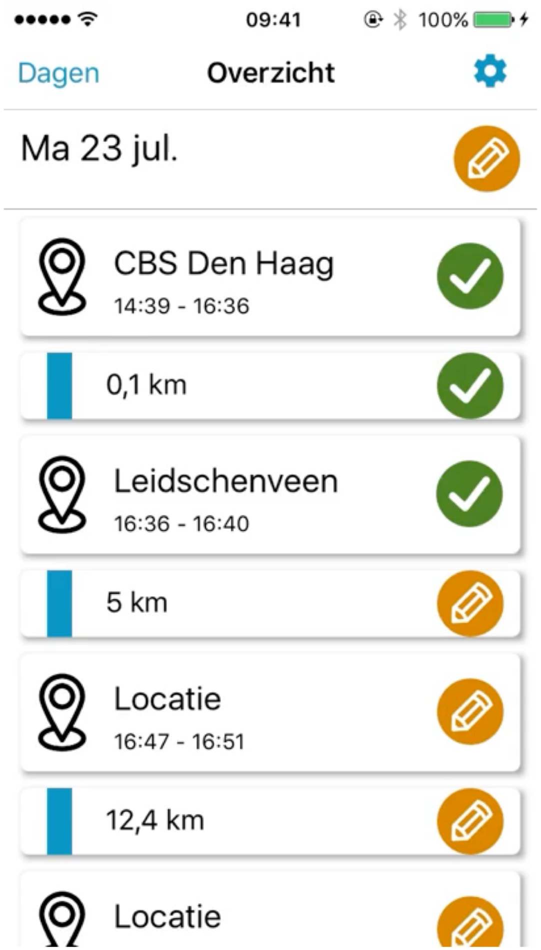
The app: Location measurements



The app: Diary and annotations







1 week travel in Dutch population - TABI app

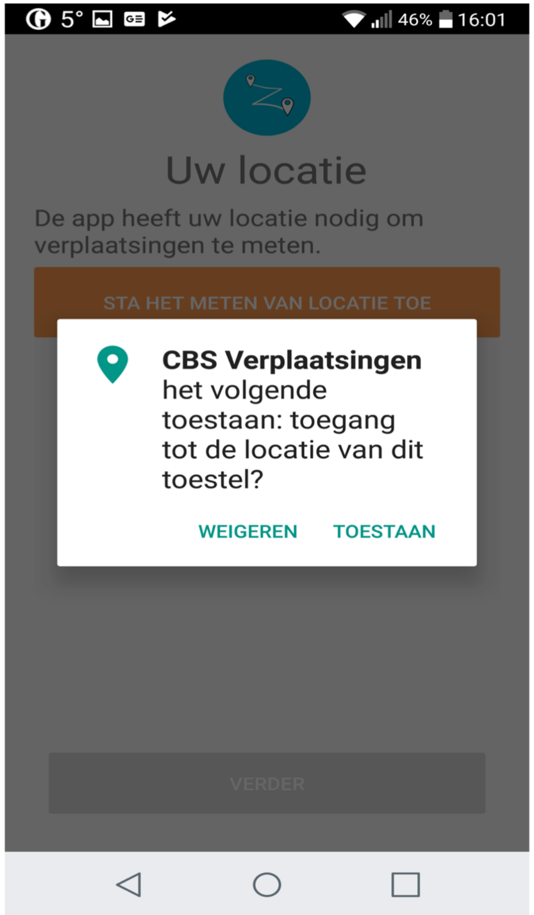
Skip to trips

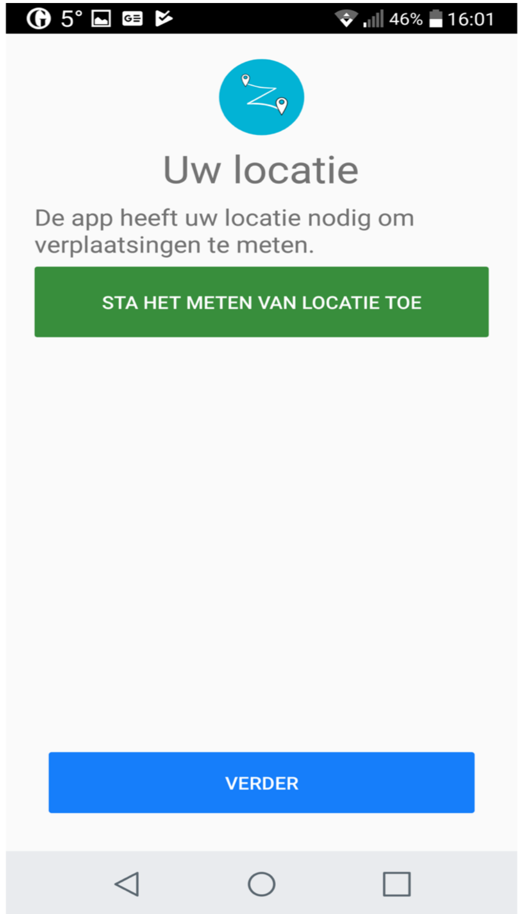
TABI: 3 experiments

- Experiment 1: Recruitment (n=1902)
  - Fresh cross-section taken from population register (n=951)
  - ODiN Web diary respondents in September 2018 (n=951)
- Experiment 2: Incentives
  - Incentive 5 + 5 + 5 (before, registering, after 7 days)
  - Incentive 5 + 10 (before, after 7 days)
  - Incentive 5 + 20 (before, after 7 days)
- Experiment 3: how to detect a stop? (not today)

Fieldwork

- 1. invitation letter
  - 1b. website
- 2. Download app
- 3. login
- 4. Allow location measurements



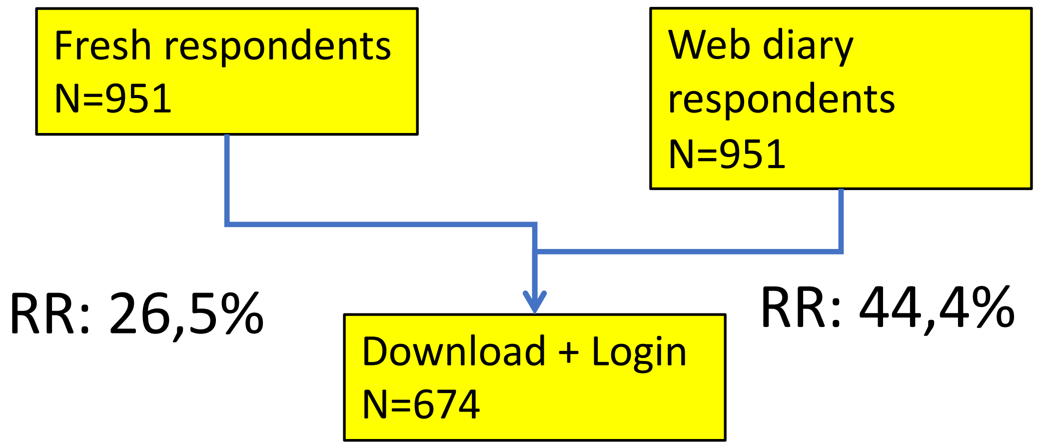


Results recruitment

Fresh respondents
N=951

Earlier Web diary respondents to ODIN study
N=951

Results recruitment (1)



Results recruitment (2)

Fresh respondents
N=951

Web diary respondents
N=951

RR: 26,5%

RR: 44,4%

Download + Login
N=674

dropout: 17%

Send data for at least 1 day
N=578

Results recruitment (3)

Fresh respondents
N=951

Web diary respondents
N=951

RR: 26,5%

RR: 44,4%

Download + Login
N=674

dropout: 17%

Send data for at least 1 day
N=578

dropout: 28%

Complete whole study
N=421

Complete Responses: 22,1% of sample

Main effect of incentives

% Installed app	sample	Incentive	
	5 + 5	0 + 10	0 + 20
n	191	231	252
%	30.1%	36.4%	39.7%

Effects on NR (conditional)

Stage 1: Registration (yes=674) Average Marginal effects	
Sample: ref= fresh	.16*
Incentive (ref=5+5+5)	
5 + 10	.07*
5 + 20	.10*
Age (ref =18-25)	
26-45	-.14*
46-65	-.20*
>65	-.28*
Drivers license (ref=no)	.05
Car owner (ref=no)	.02
Moped owner (ref=no)	.02
Highest level of education (ref = primary school)	
Lower secondary	-.01
Upper secondary/vocational	.09
Bachelor	.20*
Master	.19*
Unknown	.05
Marital status (ref=married)	-.06*
Origin (ref = Dutch)	
Non-westerm	-.11
Western	-.05
Income (ref = Q1-20)	
21-40	-.10*
41-60	.03
61-80	.08
81-100	.07
unknown	.02

Sample:

- 16% higher for web diary

Incentives:

- 5+10 7% higher than 5+5+5
- 5+20 another 3% higher

Covariates

- 20-28% lower for older age
- 20% higher for higher educated

Stage 1: Registration (yes=674) Average Marginal effects	
Sample: ref= fresh	.16*
Incentive (ref=5+5+5)	
5 + 10	.07*
5 + 20	.10*
Age (ref =18-25)	
26-45	-.14*
46-65	-.20*
>65	-.28*
Drivers license (ref=no)	.05
Car owner (ref=no)	.02
Moped owner (ref=no)	.02
Highest level of education (ref = primary school)	
Lower secondary	-.01
Upper secondary/vocational	.09
Bachelor	.20*
Master	.19*
Unknown	.05
Marital status (ref=married)	-.06*
Origin (ref = Dutch)	
Non-westerm	-.11
Western	-.05
Income (ref = Q1-20)	
21-40	-.10*
41-60	.03
61-80	.08
81-100	.07
unknown	.02
Female (ref = male)	[illegible]

NO interaction effects!

Nonresponse bias does not differ by:

- Incentive groups
- Sample groups

Roth, Lugtig & Schouten (in preparation) Nonresponse analysis in a longitudinal smartphone-based travel study

Can we do it properly?

- People spend all day on their mobile phones
  - Today: can you do a design-based general population study?
  - Response rates are good
    - 50% among 18-25 year olds…
  - But nonresponse bias in age and education strong

	Number of tracks	Number of unique locations per track	Average distance per trip (in meters)	Average speed per track (in m/s)	Average 95 percentile speed per track (in m/s)
Car	2738	444	5556	10.6	20.2
On foot	1873	30	150	1.4	3.4
Bike	995	260	1473	4.3	6.6
Train	197	526	23521	24.2	40.6
Electric bike	193	228	1729	5.0	8.1
Multiple transport modes	173	453	5607	4.4	16.6
Other	60	206	1322	3.4	6.8
Bus	59	668	5540	6.3	16.6
Wrong measurement	54	9	93	0.6	7.2
Truck	49	890	11236	11.6	21.3
Tram	35	419	3775	4.8	14.5
Scooter	34	375	2712	7.4	9.3
Metro	21	238	5400	9.4	26.0
Motor	7	18	23650	14.3	NA

Table 5.1: Modes of transport reported by respondents, along with characteristics of the trip

Some differences labeled/unlabeled trips

	Labelled trips	Unlabelled trips
median data points	347	238
median duration	553 sec	498 sec
median distance	2277 m	1707 m
% trips $\leq$ 500m	20.5%	27.4%
% E-bike trips	3.44%	7.72%*
% bike trips	17.84%	10.02%*
% car trips	50.98%	35.03%*
% metro trips	.35%	2.19 %*
% bus trips	1.04%	7.74%*
% scooter trips	.67%	1.66%*
% train trips	4.13%	5.98%*
% tram trips	.44%	2.23%*
% User errors	.65%	2.07%*
% walk trips	21.10%	27.46%*

Smeets, Lugtig & Schouten (in review) Automatic travel mode prediction in a national travel study. JRSS:A

App = better at short trips

	Labelled trips	Unlabelled trips	ODiN 2018 total
median data points	347	238	NA
median duration	553 sec	498 sec	900 sec
median distance	2277 m	1707 m	3000 m
% trips $\leq$ 500m	20.5%	27.4%	13.1%
% E-bike trips	3.44%	7.72%*	3.77%
% bike trips	17.84%	10.02%*	26.41%
% car trips	50.98%	35.03%*	43.00%
% metro trips	.35%	2.19 %*	.81%
% bus trips	1.04%	7.74%*	2.33%
% scooter trips	.67%	1.66%*	.72%
% train trips	4.13%	5.98%*	2.83%
% tram trips	.44%	2.23%*	.80%
% User errors	.65%	2.07%*	NA
% walk trips	21.10%	27.46%*	19.30%

App: more walks, fewer bike trips

	Labelled trips	Unlabelled trips	ODiN 2018 total
median data points	347	238	NA
median duration	553 sec	498 sec	900 sec
median distance	2277 m	1707 m	3000 m
% trips $\leq$ 500m	20.5%	27.4%	13.1%
% E-bike trips	3.44%	7.72%*	3.77%
% bike trips	17.84%	10.02%*	26.41%
% car trips	50.98%	35.03%*	43.00%
% metro trips	.35%	2.19 %*	.81%
% bus trips	1.04%	7.74%*	2.33%
% scooter trips	.67%	1.66%*	.72%
% train trips	4.13%	5.98%*	2.83%
% tram trips	.44%	2.23%*	.80%
% User errors	.65%	2.07%*	NA
% walk trips	21.10%	27.46%*	19.30%

Conclusions 1: smartphone travel app

- Large Nonresponse bias (not shown)
  - Age and level of education
  - Incentives work
- Respondents can be freshly recruited
- Measurement overall better with app
  - But some technical issues
  - And a lot of modeling effort