



How can innovative data collection look like to help reducing the response burden for time-use surveys?

Britta Gauckler

Eurostat

Workshop on Statistical Data Collection
“New sources and New Technologies”
Geneva, 14-16 October 2019



Background & objectives

- *2nd parallel Hackathon, 1st Hackathon in 2017*
- *Solve statistical problems by leveraging algorithms and available data*
- *Identify best of class European Data Scientists*
- *Promote and accelerate Big Data for Statistics initiatives in Europe*
- *Promote partnerships with research community and private sector*
- *To produce innovative products and tools, including visualisation*



Teams

- *Hackathon as part of the NTTS 2019 conference from 9-12 March 2019, before the NTTS 2019 conference*
- *Kick-off 11 March in the evening and announcement of the topic*
- **16 NSI teams** (Austria, Croatia, Czech Republic, Estonia, Finland, Germany, Greece, Ireland, Italy, Latvia, Netherlands, Poland, Romania, Slovenia, Sweden, United Kingdom)
- **One EMOS team**
(University of Pisa & Sant' Anna College of Pisa)





Resources & infrastructure

- *Technical infrastructure for the data collection and Hackathon provided by DIGIT's Big Data Technical Infrastructure (BDTI)*
- *Coaches during the hacking: 10 coaches from universities, AWS and BDTI*
- *One big data cluster for each team*



Data sources

- Data collected via smart phones
 - Sensors data + user contributed
 - Own data collection via i-Log () and myBigO ()
 - Dataset from VUB
- MOTUS
- HETUS



Eurostat

i-Log

Background questions

Sensors

User input (every hour)

Android status bar: 80% 17:41

i-Log PREVIOUS NEXT

Main activity status [17:40]

- Employed full-time
- Employed part-time
- On leave
- Unemployed
- Pupil, student, further training, unpaid traineeship
- In retirement or early retirement or has given up business;
- Fulfilling domestic tasks
- Other



What are you doing?

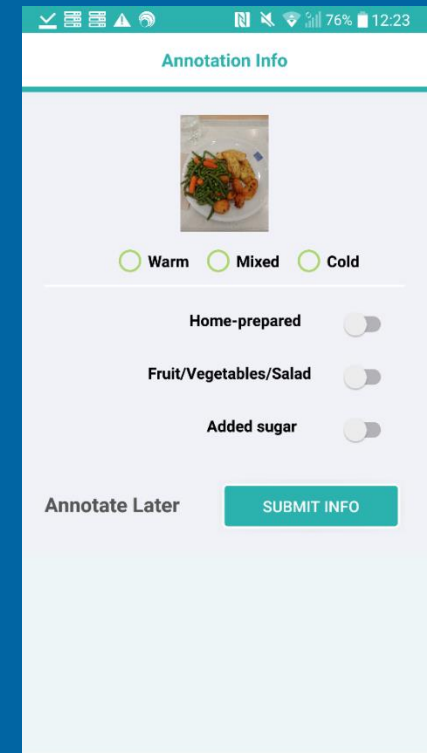
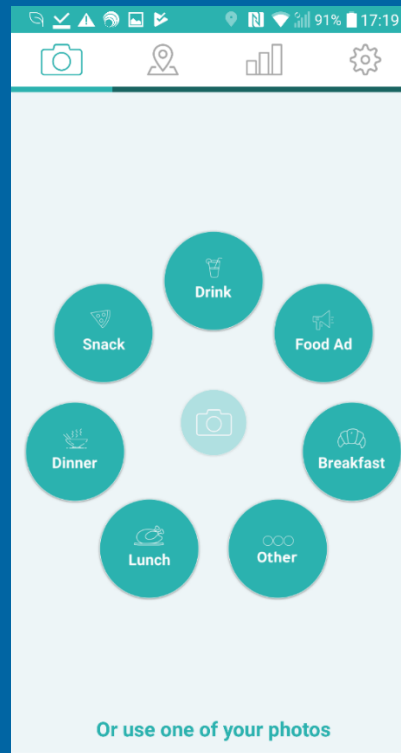
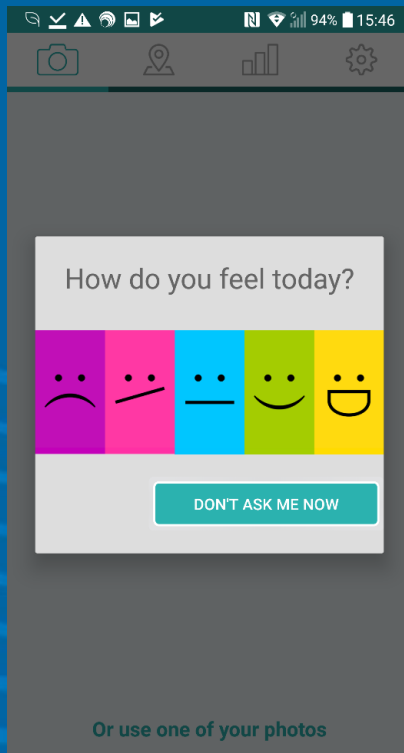
Where are you?

Which transport mode?

Who is with you?

What is your mood?

myBigO





Challenges for Time Use Surveys (TUS)

- The completeness to capture the complexity of social life
- The granularity of the information
- The cost (and time necessary to carry out a study)



Hackathon 2019 theme

How can innovative solutions for data collection reduce response burden and enrich or replace the statistical information provided by the time use survey?



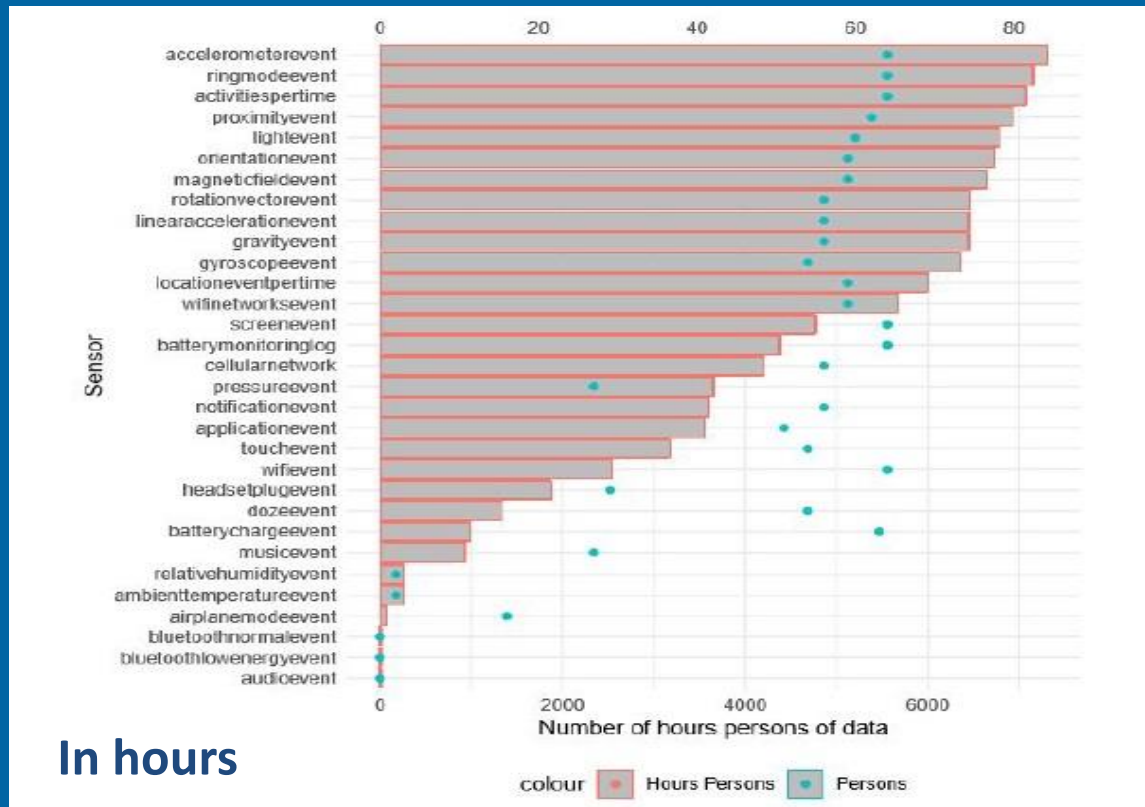
Hackathon 2019 theme (continued)

the solution could extend statistical information to additional areas or intensify information on existing areas, such as:

- *social well-being,*
- *work-life balance,*
- *use of information and communication technologies,*
- *mobility and travel,*
- *physical activity,*
- *social environment,*
- *geographical context,*
- *regularity and frequency of activities,*
- *recognition and registration of episodes,*
- *innovative time-space policy dashboards*
- *informing on relevant SDGs*

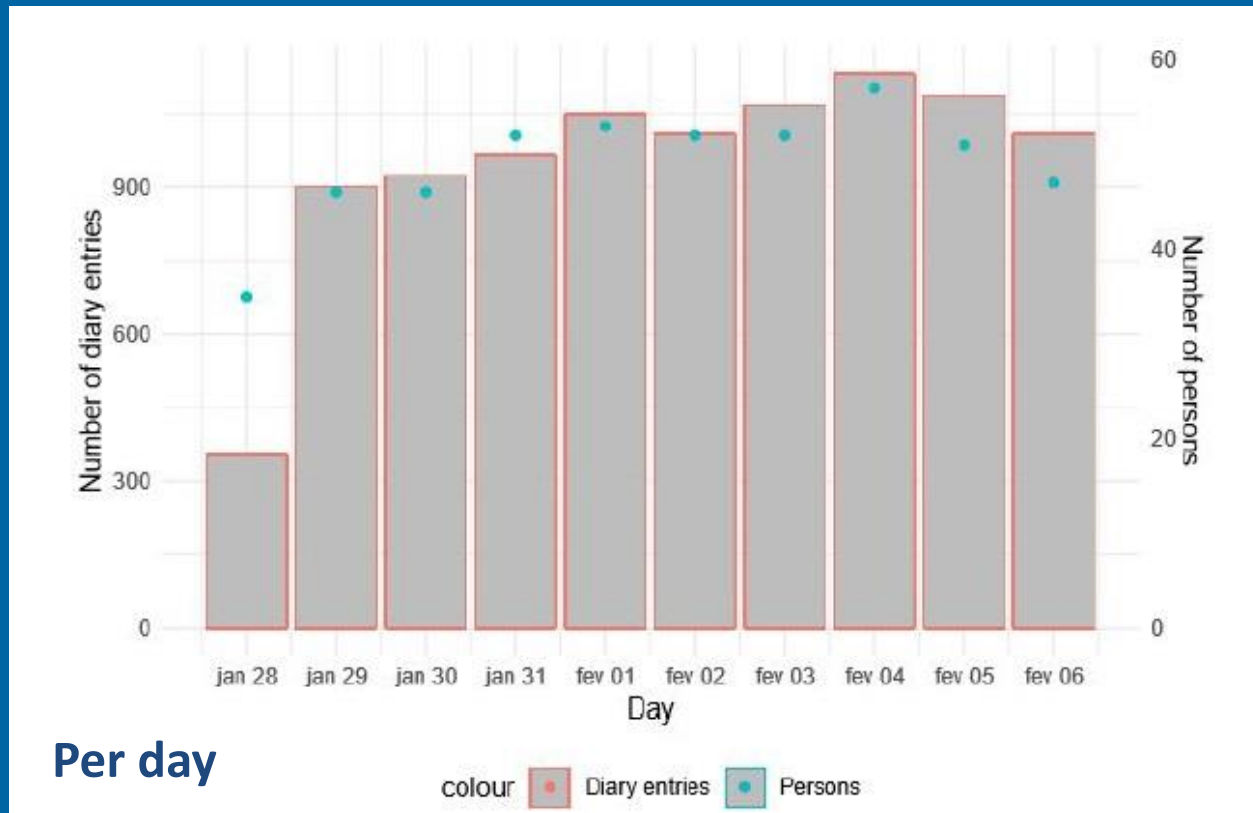


Amount of data collected from each sensor





Amount of time diaries answers collected



Per day

colour ■ Diary entries ● Persons



Selected results

.... from Sweden, Romania, Greece, Netherlands, Croatia, Latvia



Awarded winning teams

Big Data Hackathon award

- 1st Place -the team from Statistics Poland
 - created an open source prototype delivering a dashboard for the data analysis of the population time use
- 2nd Place -the team from ISTAT - Italy
 - created SMUTIS, an integrated open source environment for data analytics, visualisation and food classification
- 3rd Place -the team from the ONS – UK
 - developed a system to enrich the data collected via traditional questionnaire based surveys with an automatic processing of photos of meals taken by respondents



...what's next?

- *Promising results*
- *Challenges to overcome*