

LIFELOGS UNDERSTAND

EVERYDAY ACTIVITIES

GRADUATION PROJECT
DATA SCIENCE AND BIG DATA



CONTENT TABLE

The presentation contains the most important results of the graduation project of a postgraduate course *Introduction to Data Science and Big Data*, at Faculty of Mathematics and Computer Science, Universitat de Barcelona.

The data analysis, the data processing and insights into project's details can be found here:
<https://github.com/llevaNEUS/Lifelog-Project-Data-Science-UB-2016-2017>

June, 2017

NTCRI-12 LIFELOGS PROJECT	3
RESEARCH OBJECTIVES	4
LIFELOGS DATA OVERVIEW	5
YOLO IMAGE RECOGNITION	6
DATA ANALYSIS PROCESS	7
LIFELOGS ANALYSIS AND PREDICTION	8
CONCLUSIONS	25

NTCIR-12 LIFELOGS PROJECT

What is a lifelog?

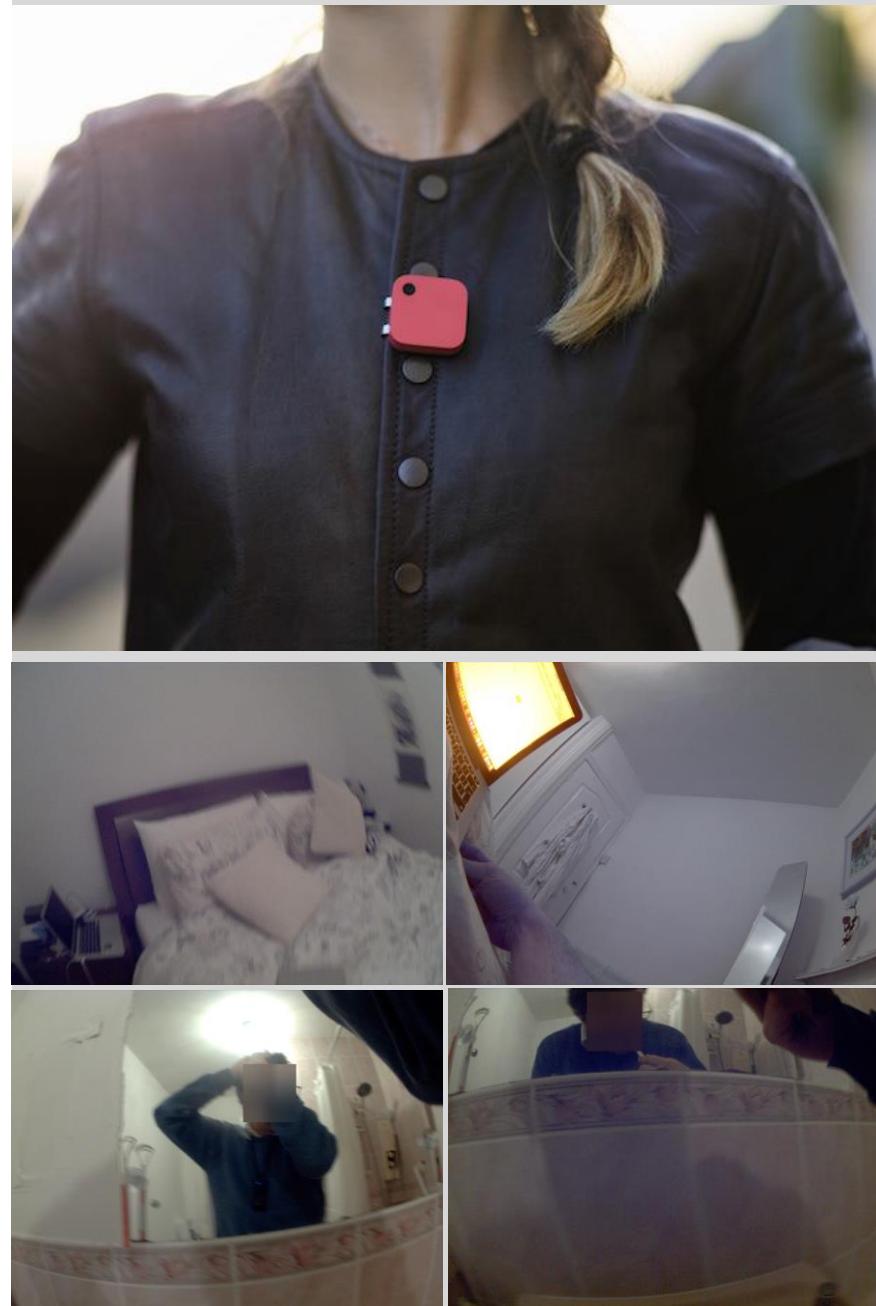
A lifelogging is defined as “*a form of pervasive computing, consisting of a unified digital record of the totality of an individual’s experiences, captured multi-modally through digital sensors and stored permanently as a personal multi media archive*” (Gurrin,. C., 2014)

Who does a research on lifelogs?

Lifelogs became an object of interest of researchers associated within NTCIR Evaluation Forum: professors and lecturers from **Dublin City University, University of Tsukuba, and University of Glasgow**. The NTCIR project started in 1997 when has promoted research efforts for enhancing information access (IA) technologies such as information retrieval (IR), text summarization, information extraction, and question answering techniques. In 2015 researchers decided to expose the data to external audience and release it as a challenge called NTCIR-12.

What is a research methodology?

Selected loggers have wear a narrative click camera which takes automatically one picture per 2 minutes. A camera stores images locally with assigned ID, time, and date. The total data collection is a large set of daily lifelogs.

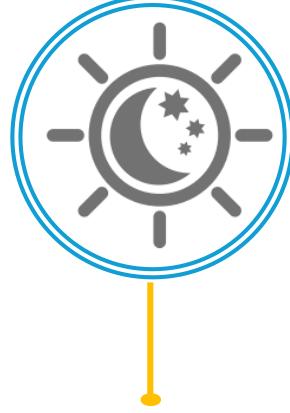


RESEARCH OBJECTIVES



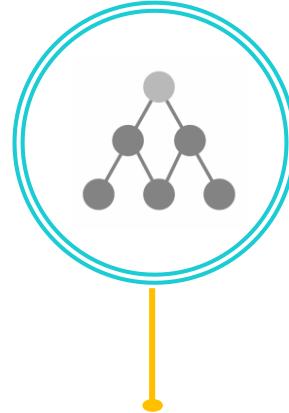
ARE LIFELOGGERS CONNECTED?

Identify the convergences in visited locations between lifeloggers.



WHAT IS A TYPICAL DAY IN LIFELOGGER'S LIFE?

Identify the typical daily schedule of a logger and its activities flow.



IS IT POSSIBLE TO PREDICT ACTIVITIES?

Identify the best and the most accurate model to predict activities by using images.



WHAT IS A LIFELOGGER'S LIFESTYLE

Identify the most common and significant ways of living and try to expose the impact on the logger's health.



ARE LIFELOGGERS PREDICTABLE?

Identify the final conclusions and find if lifeloggers are predictable.

LIFELOGS DATA OVERVIEW



18.6 GB of raw data



87 724 images



121 locations



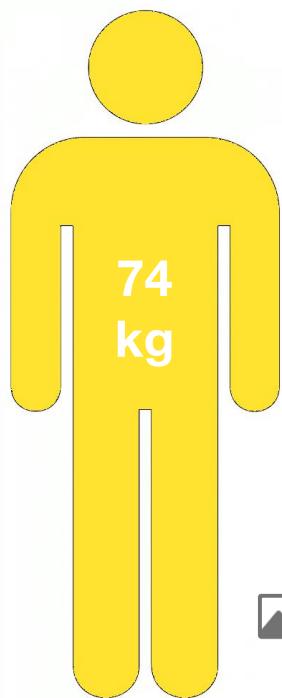
20 activities



40 years old

1.73 meter

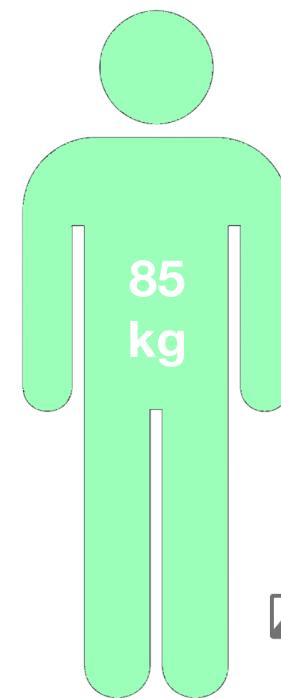
31740 images



33 years old

1.83 meter

24401 images



48 years old

1.76 meter

26583 images

2015

Feb

Mar

Apr

Jun

Jun

Aug

26 days

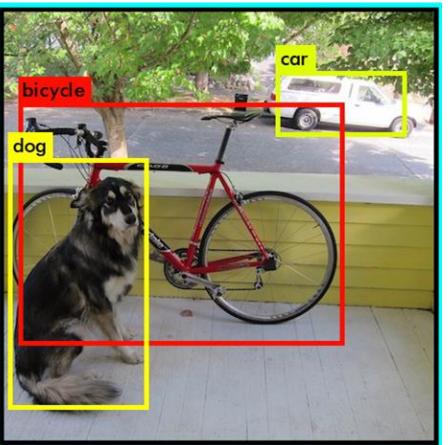
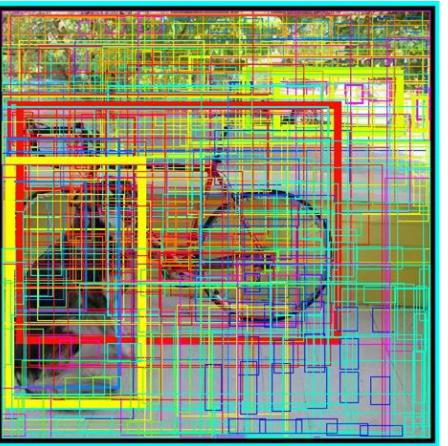
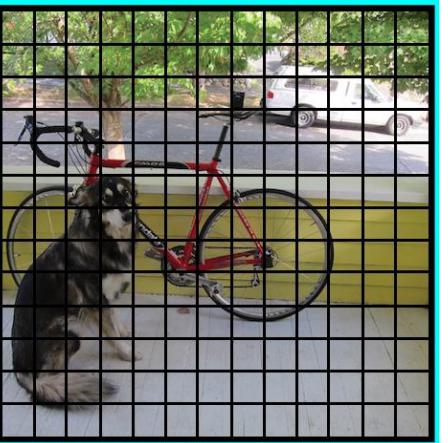
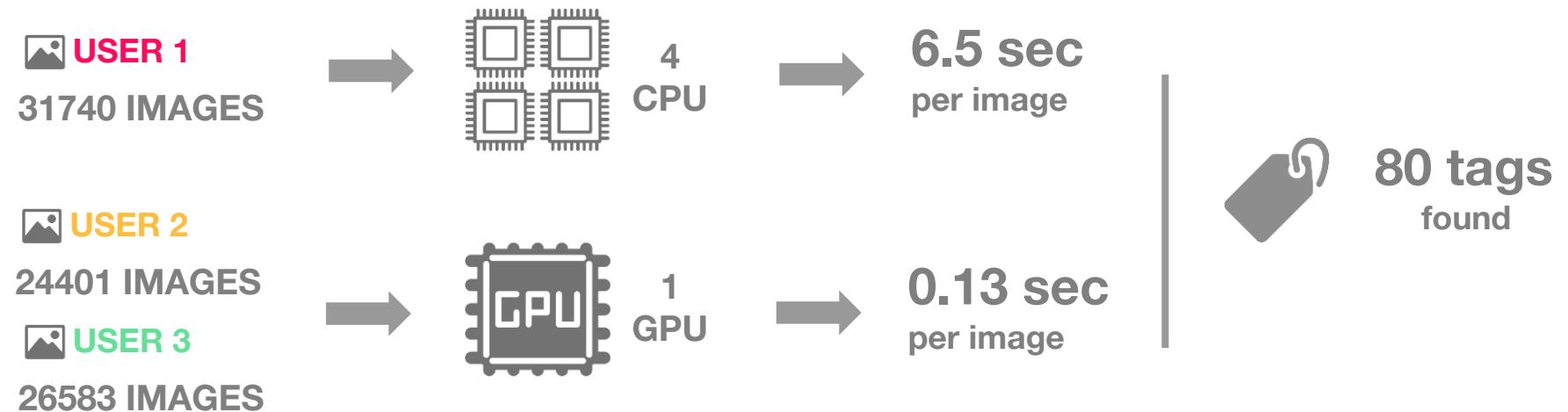
27 days

28 days

YOLO IMAGE RECOGNITION

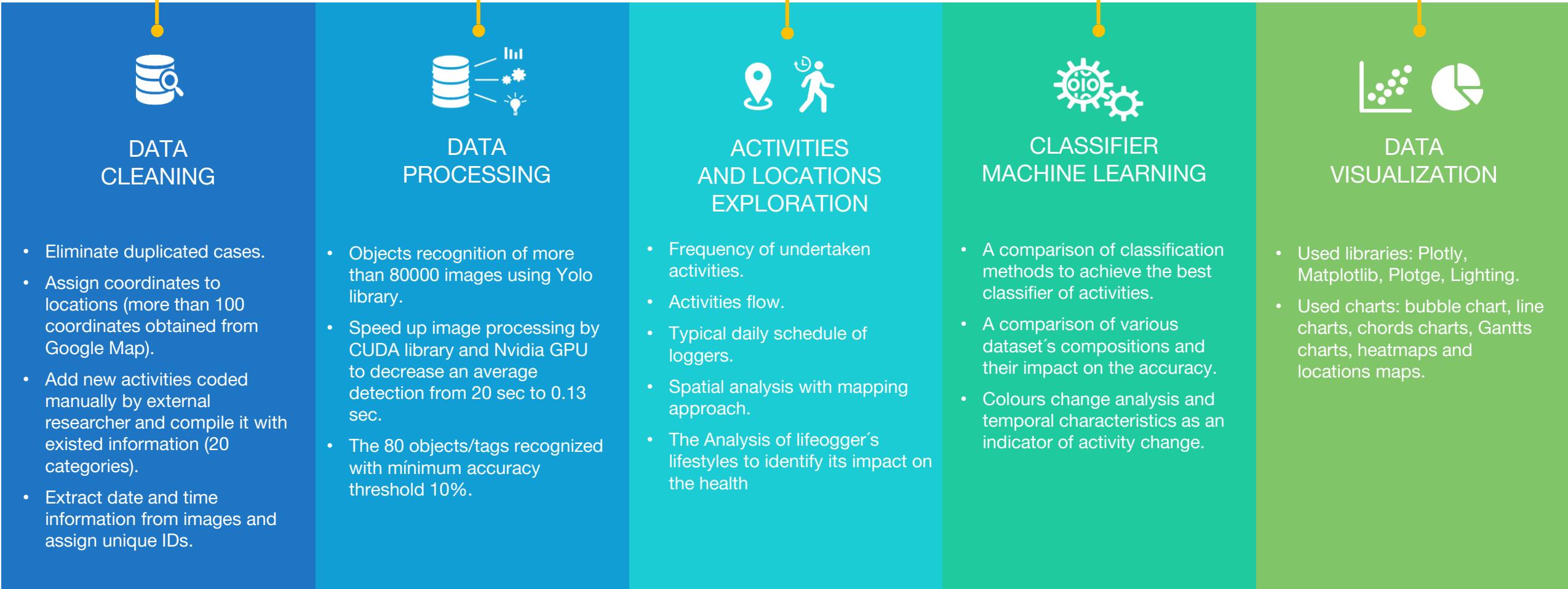
YOLO 9000 is a pretrained **neural network** based on Deep Learning ables to detect **9000 different objects**. A single neural network is applied to the full image and divides the image into regions predicting bounding boxes and probabilities for each region. These bounding boxes are weighted by the predicted probabilities and returned as detection's results.

LIFELOGGER'S IMAGES AND THE REAL COMPUTATION COST

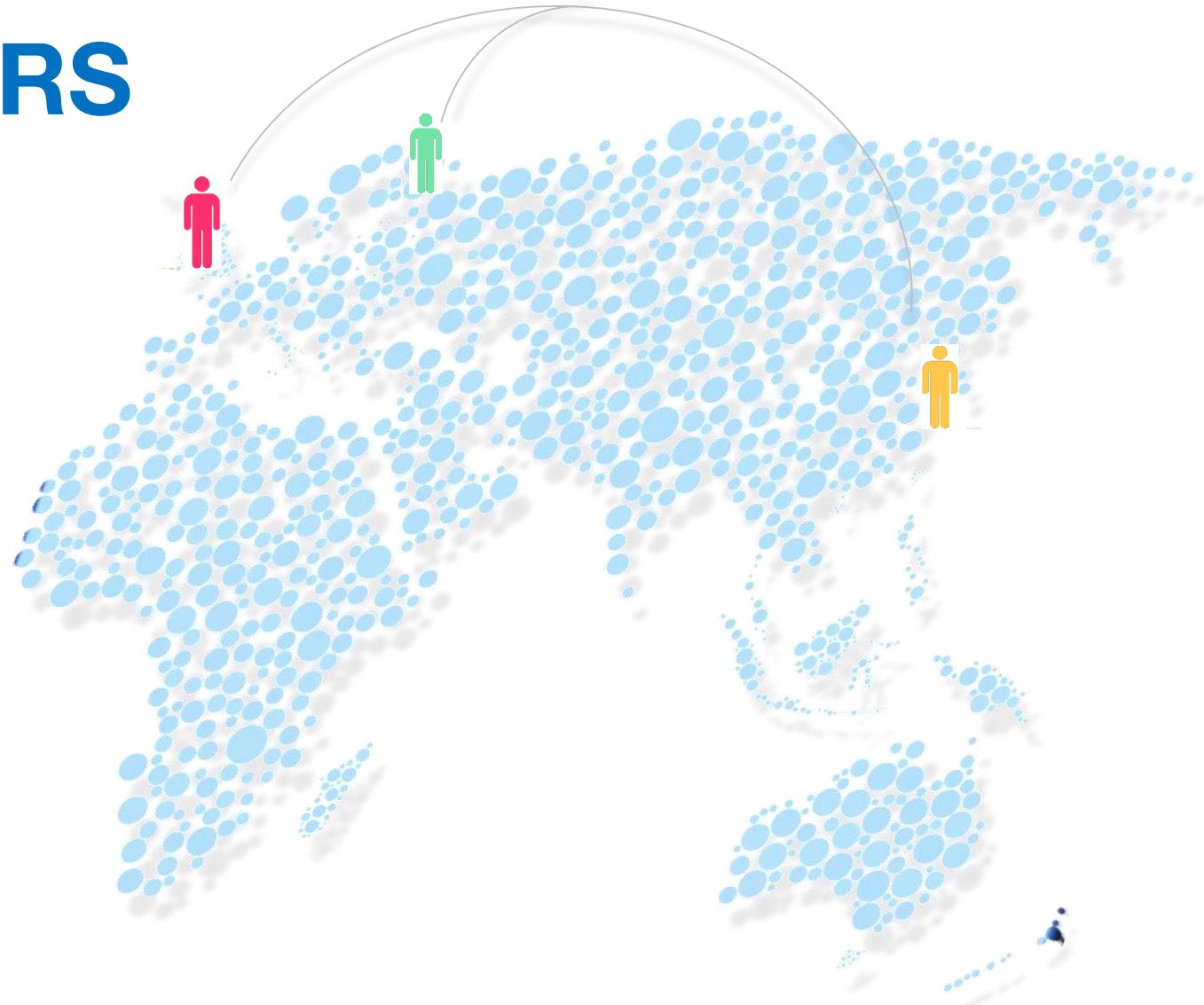




DATA ANALYSIS

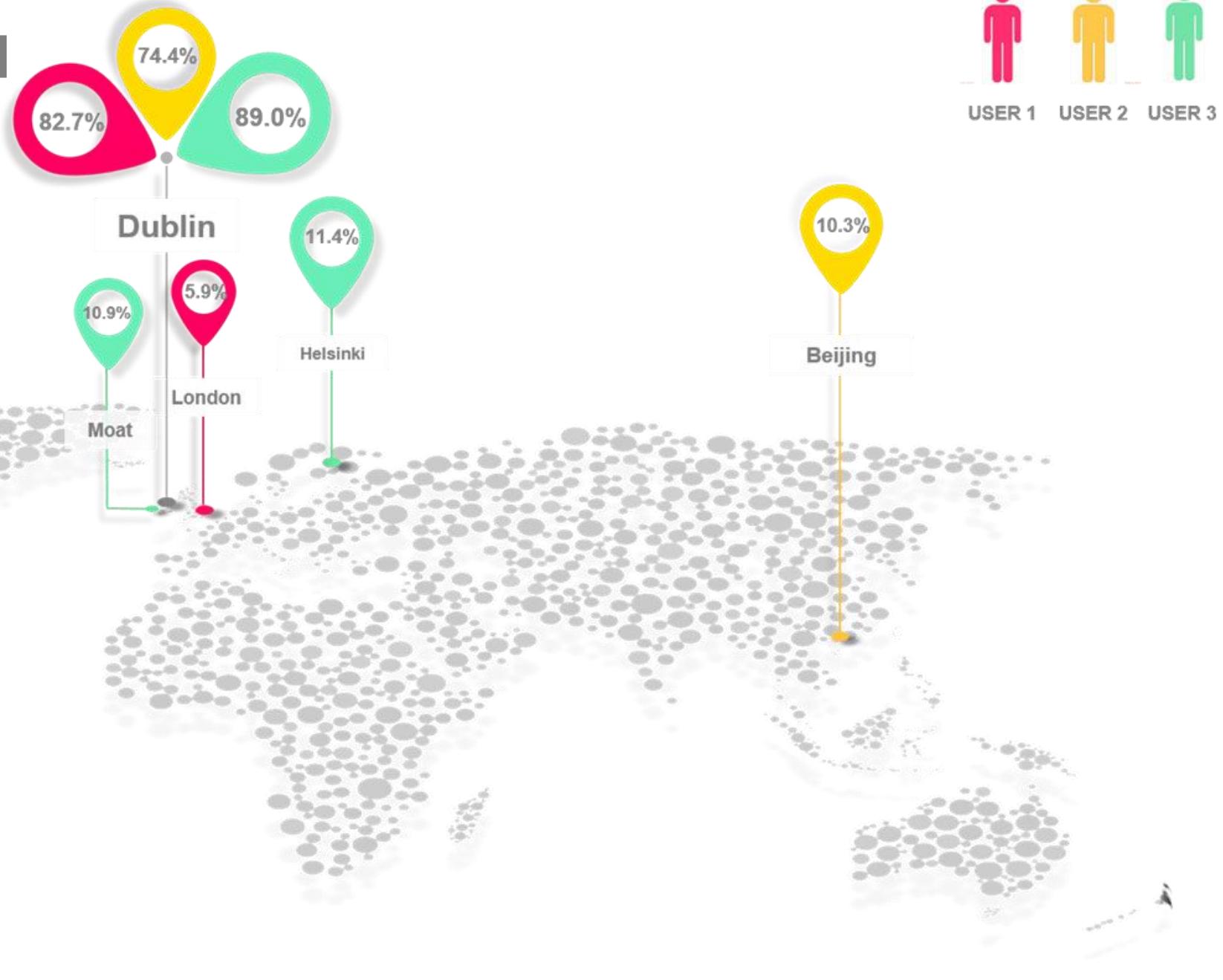


1 | ARE LIFELOGGERS CONNECTED?



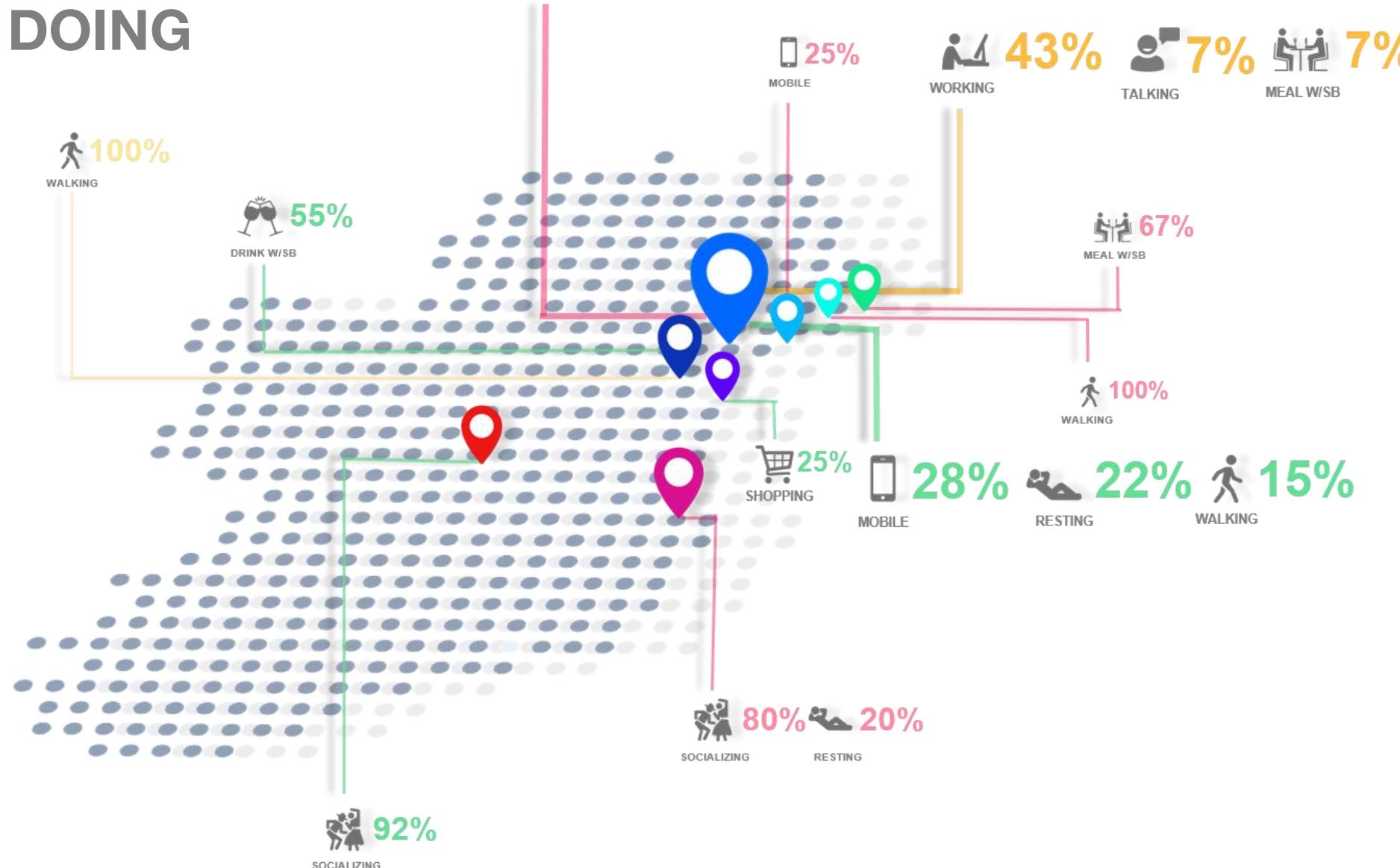
WHERE HAVE BEEN THE LOGGERS?

TOP LOCATIONS



WHAT HAVE BEEN THE LOGGERS DOING IN IRELAND?

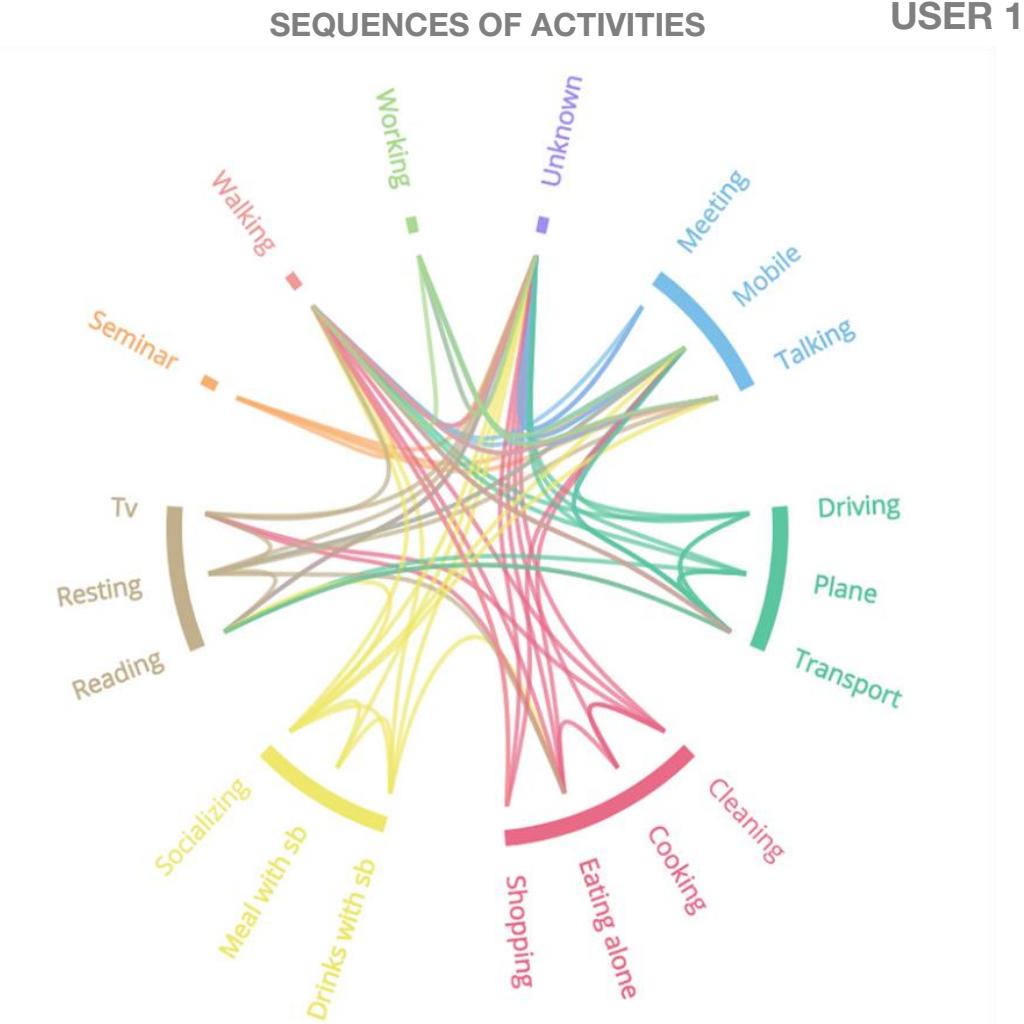
- Moate
- Enniscorthy
- Saggart
- Rathfarnham
- Dublin
- Kilbarrack
- Sword
- Malahide



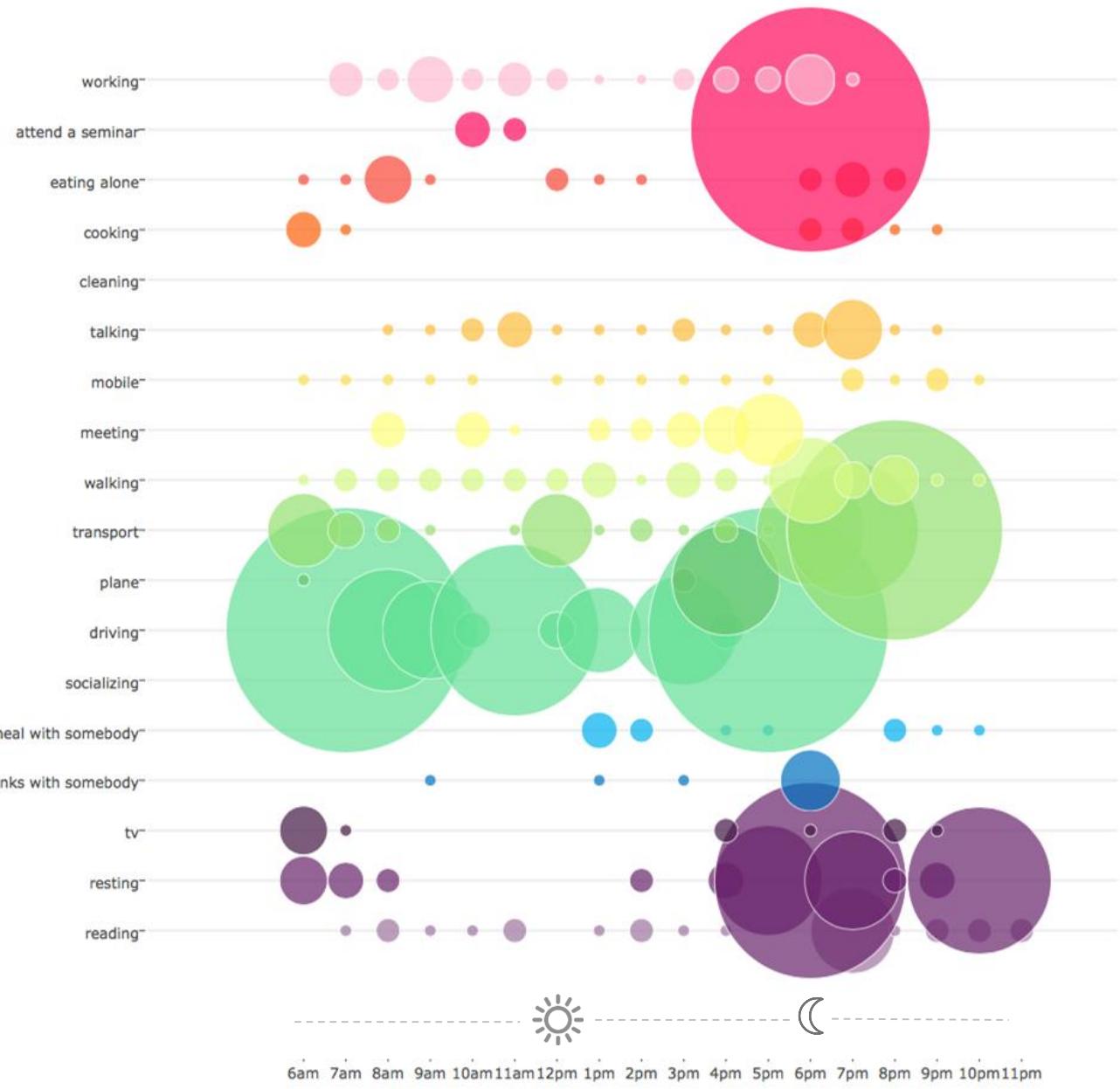
2 | WHAT IS A TYPICAL DAY IN LIFELOGGER'S LIFE?



HOW DOES LOOK LIKE THE EVERYDAY LIFE?

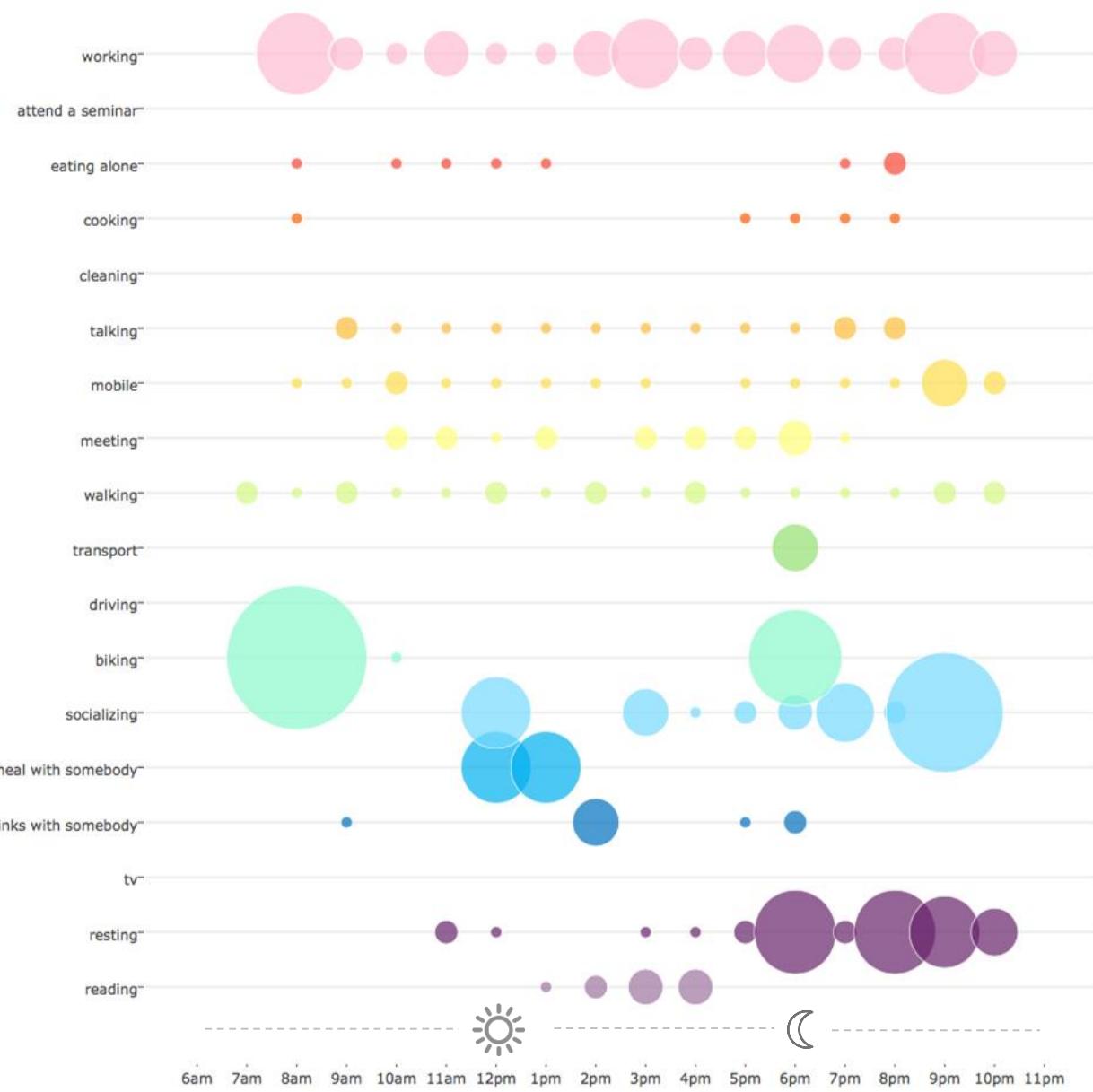
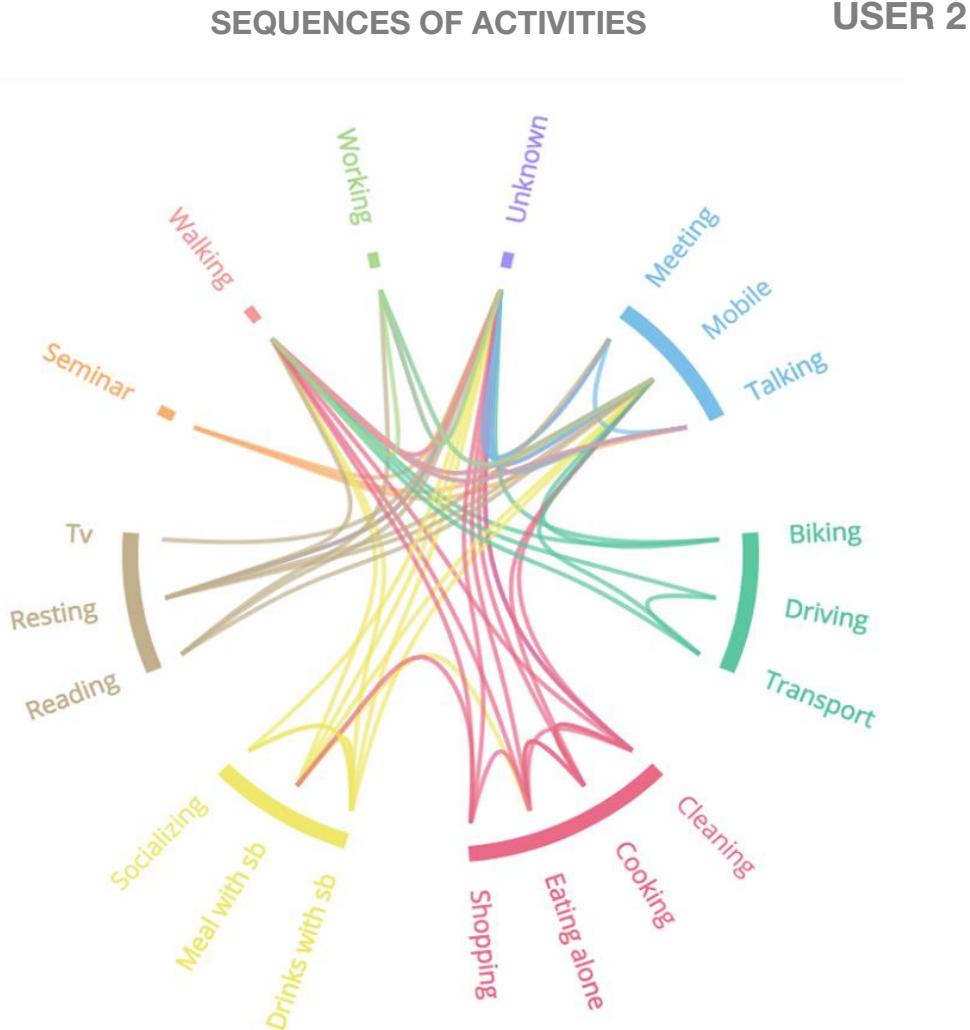


Repository: Lighting in: <http://public.lightning-viz.org/visualizations/8999361b-245f-433b-9f86-74a8ecd41413/public/>



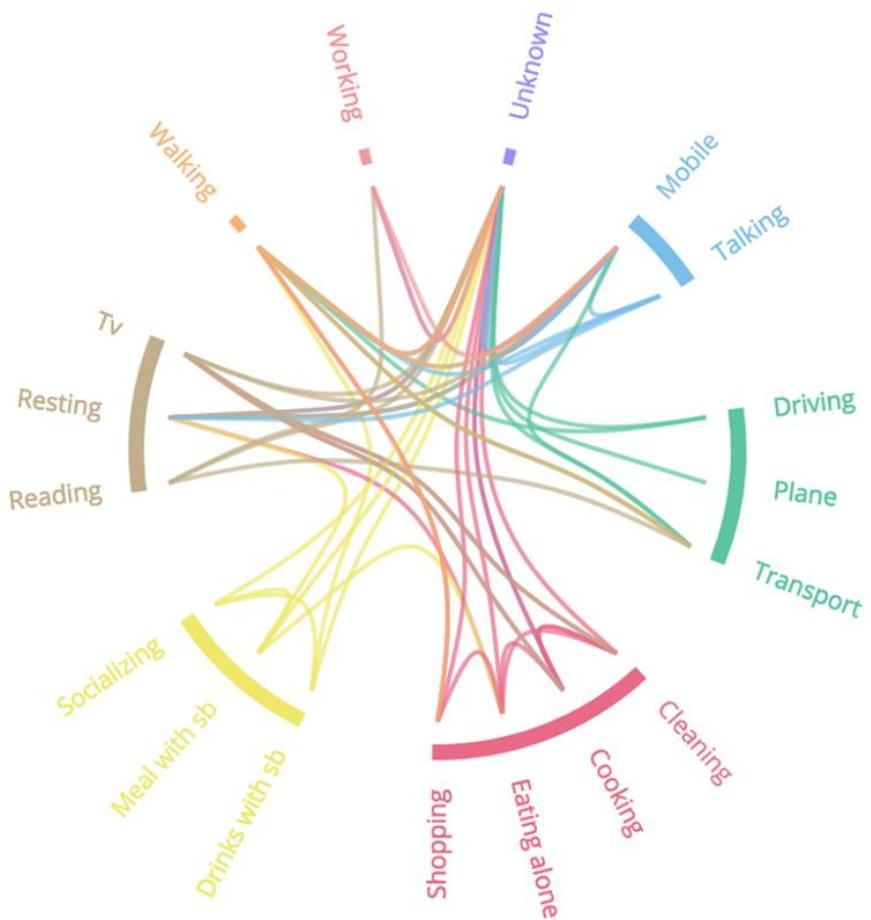
Repository: Plotly in: <https://plot.ly/~aneta.marcinkowska/114/activities-by-users-in-percentage/>

HOW DOES LOOK LIKE THE EVERYDAY LIFE?

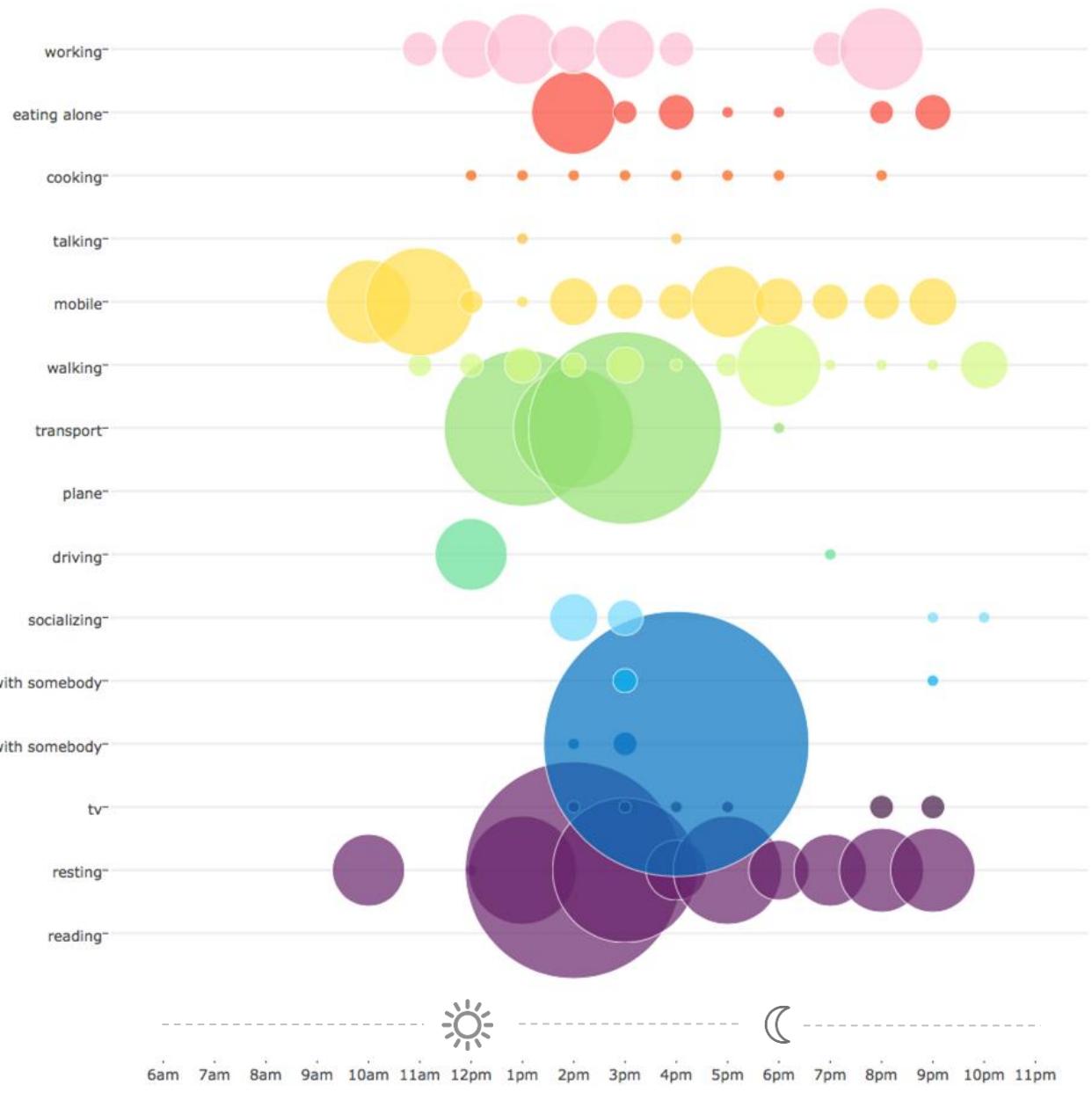


HOW DOES LOOK LIKE THE EVERYDAY LIFE?

SEQUENCES OF ACTIVITIES

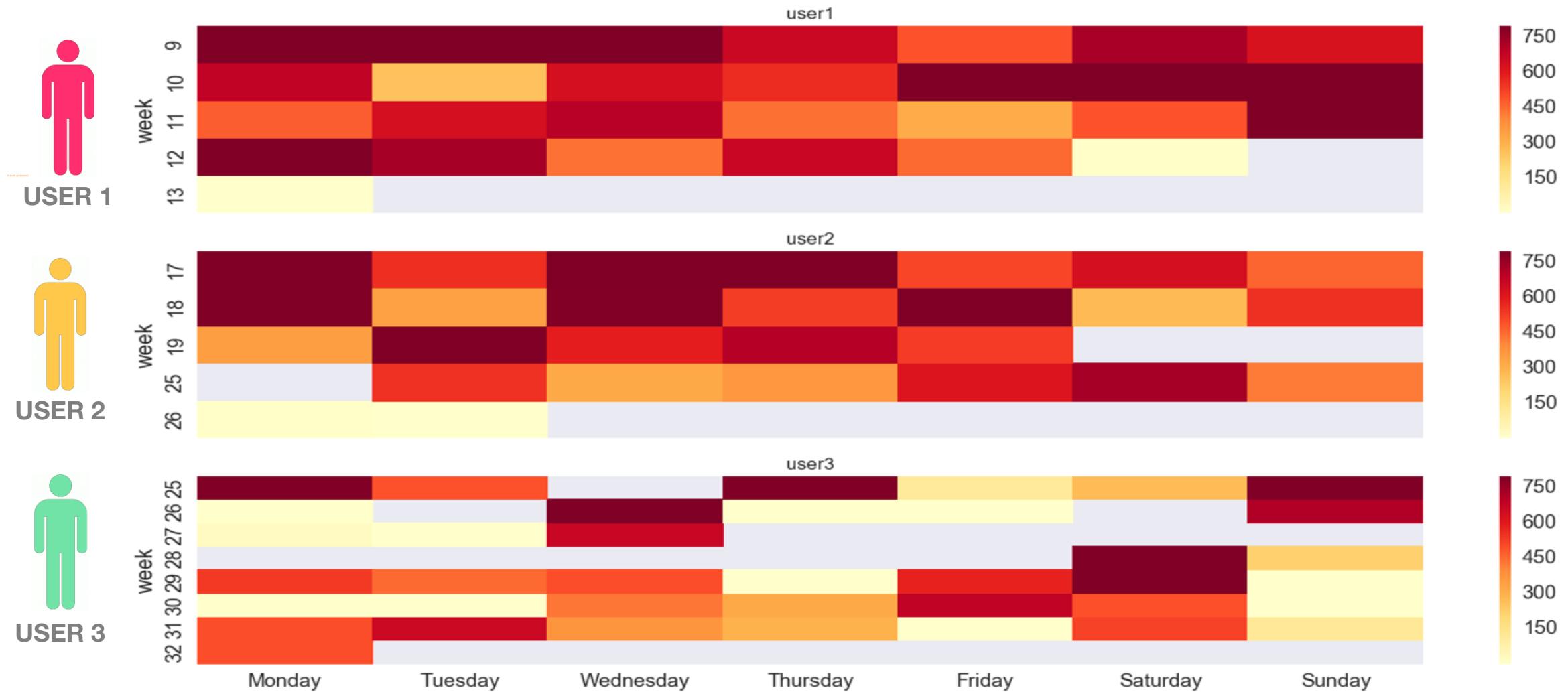


Repository: Lighting in: <http://public.lightning-viz.org/visualizations/a5ad452a-a474-4fc3-a9c8-8347dacc2691/public/>



HOW DOES LOOK LIKE THE EVERYDAY LIFE IN TERMS OF THE MOBILITY?

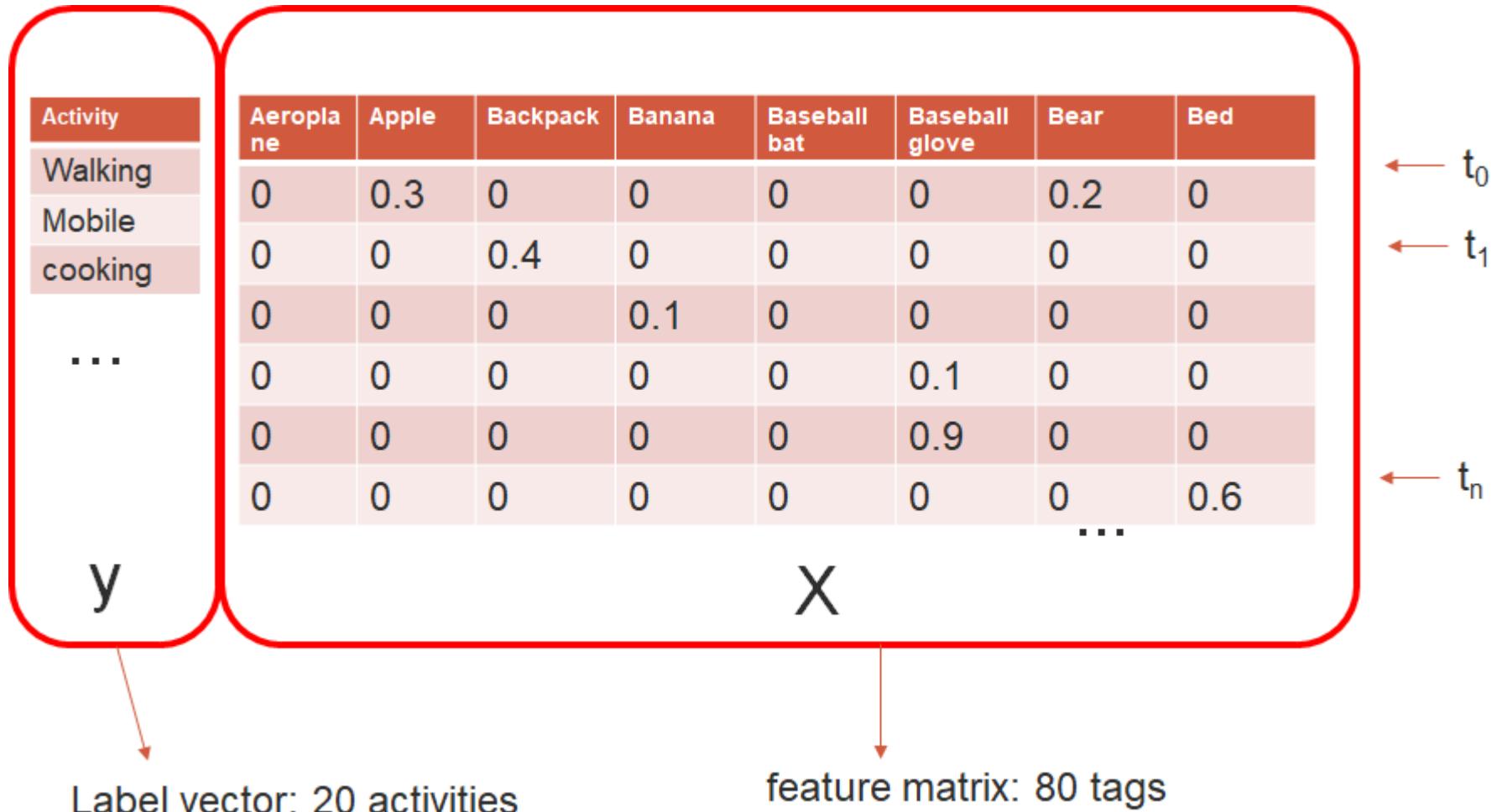
TRAVELING (num. of occurences per day)



3| IS IT POSSIBLE TO PREDICT ACTIVITIES?

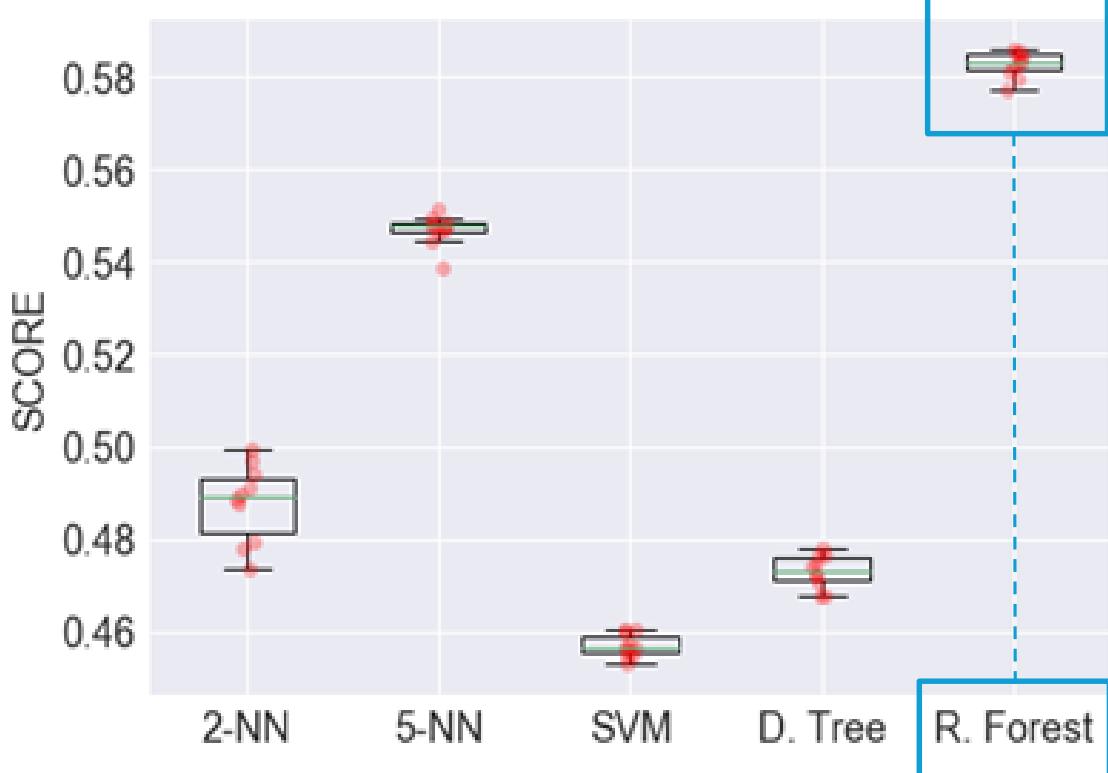


THE FEATURES MATRIX AS A RESULT OF YOLO PROCESSING



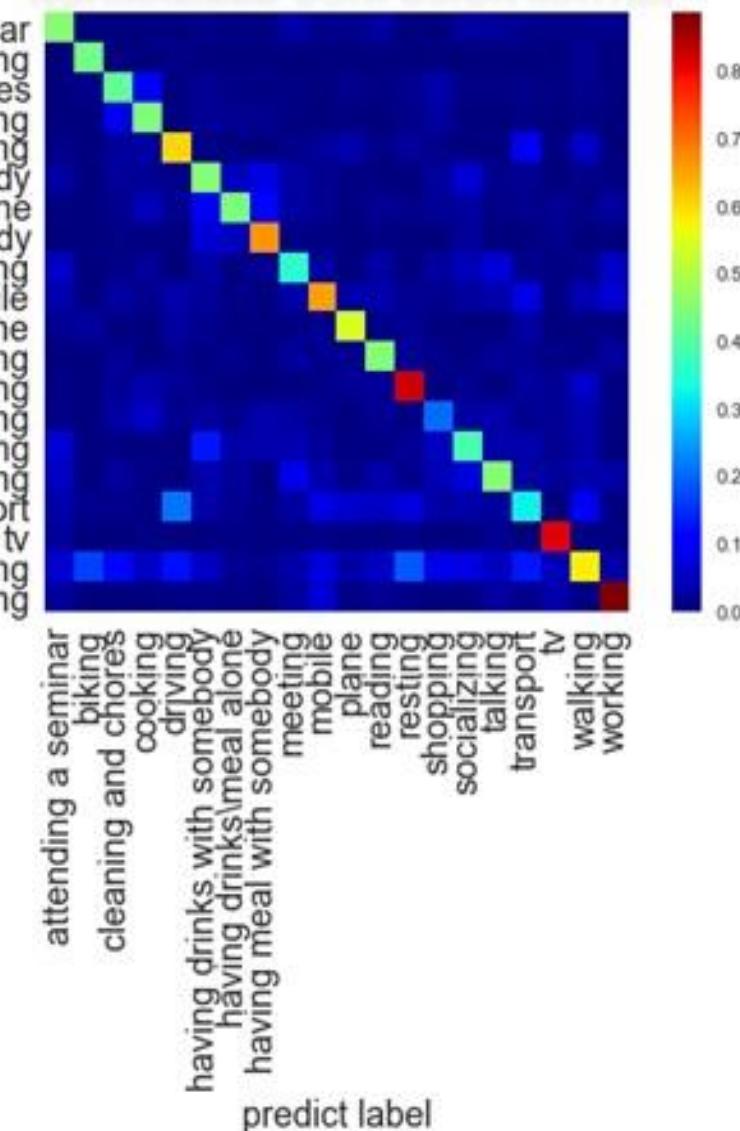
STEP 1: RAW AND SIMPLE PREDICTION

A COMPARISON OF CLASSIFIERS



true label

NORMALIZED CONFUSION MATRIX



CAN WE IMPROVE THE MODEL?

YES

BY ADDING ADDITIONAL
FEATURES TO THE MODEL

1 ADDING LOCATION AND TEMPORAL INFORMATION

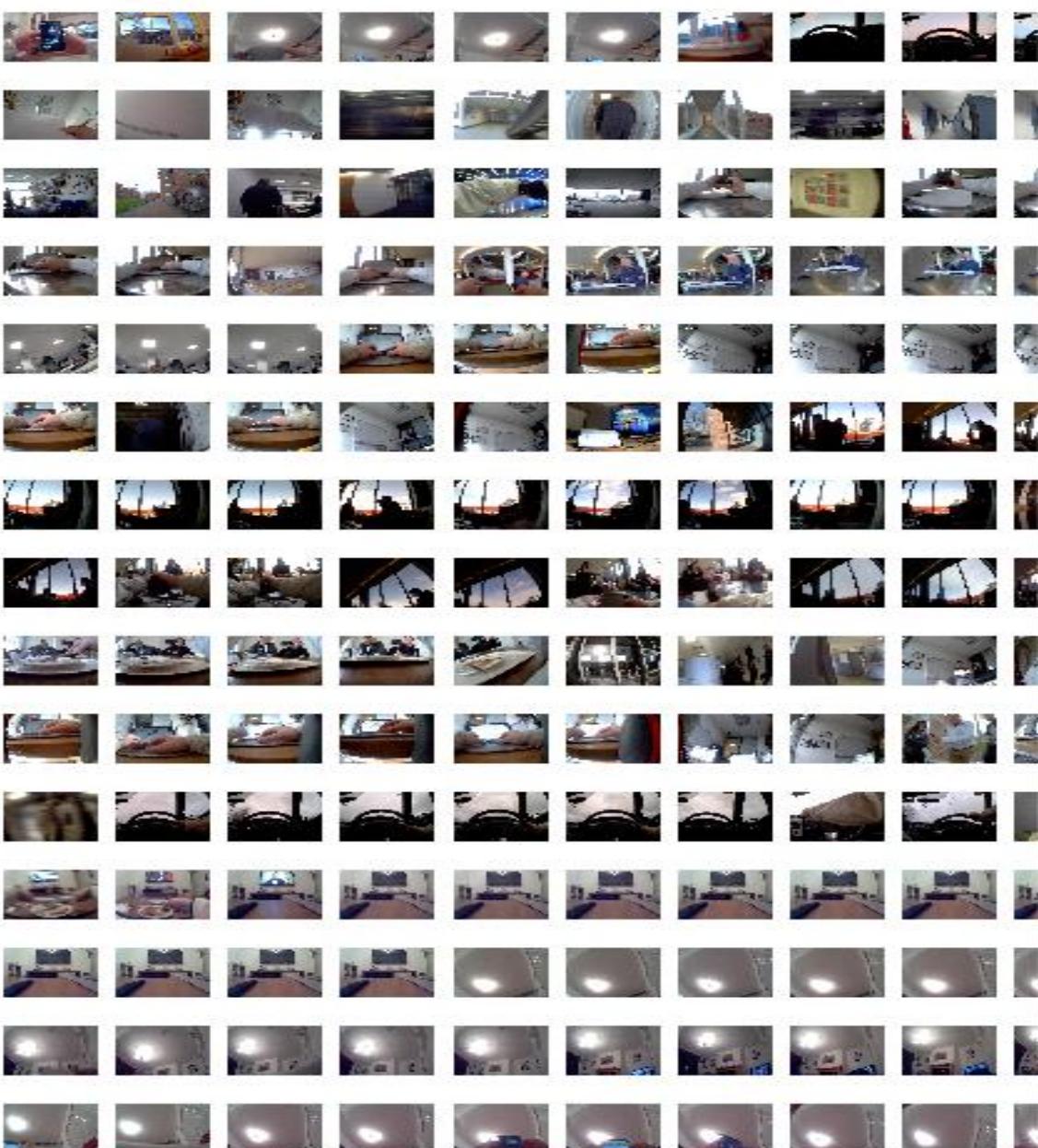
Binary variables based on **weekend** ('Was it weekend or not?') or whether the user is at **home** or not ('Is he/she at home or not?')

2 ADDING 8-BINS RGB FOR EACH IMAGE

The 8-bins RGB of each image, helps the classifier to 'see' how the logger's house, workplace or hometown looks like.

3 ADDING THE LAST ACTIVITY PERFORMED

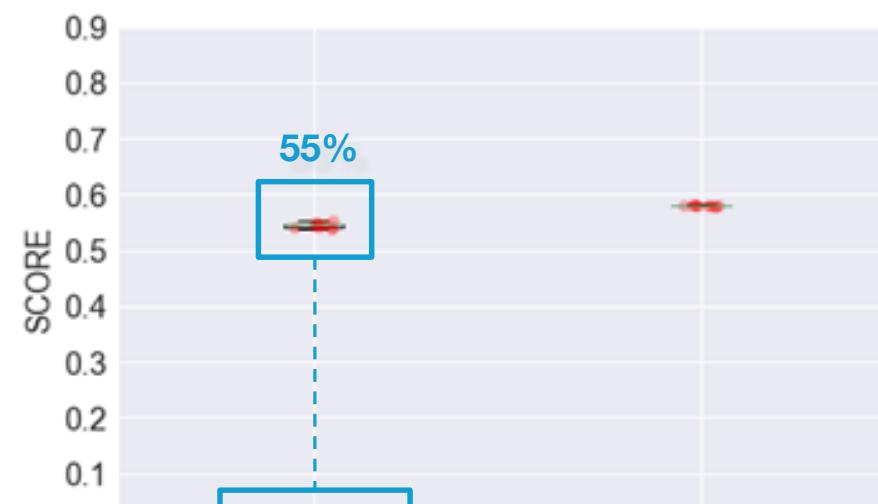
The variable with the previous activity (t_{-1}) allowed the classifier to associate an activity with previous behaviour.



STEP 2: THE MODEL IMPROVEMENT

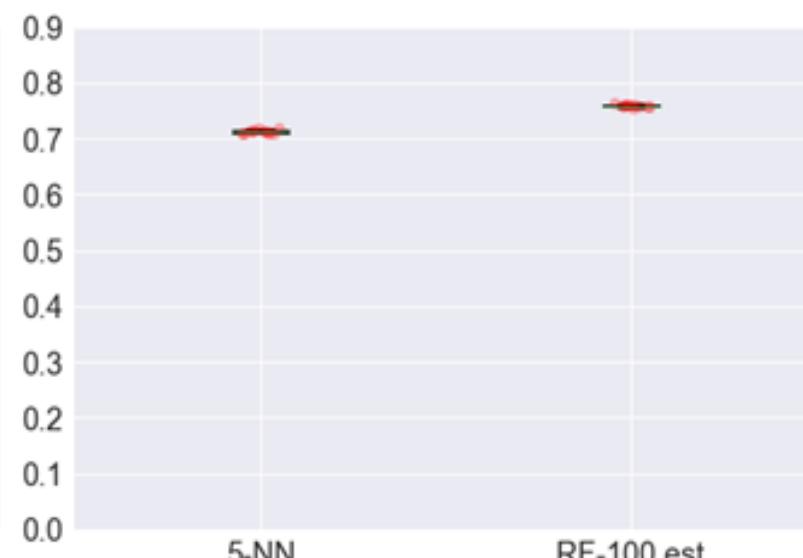
1

ADDING WEEKEND AND
HOME VARIABLE



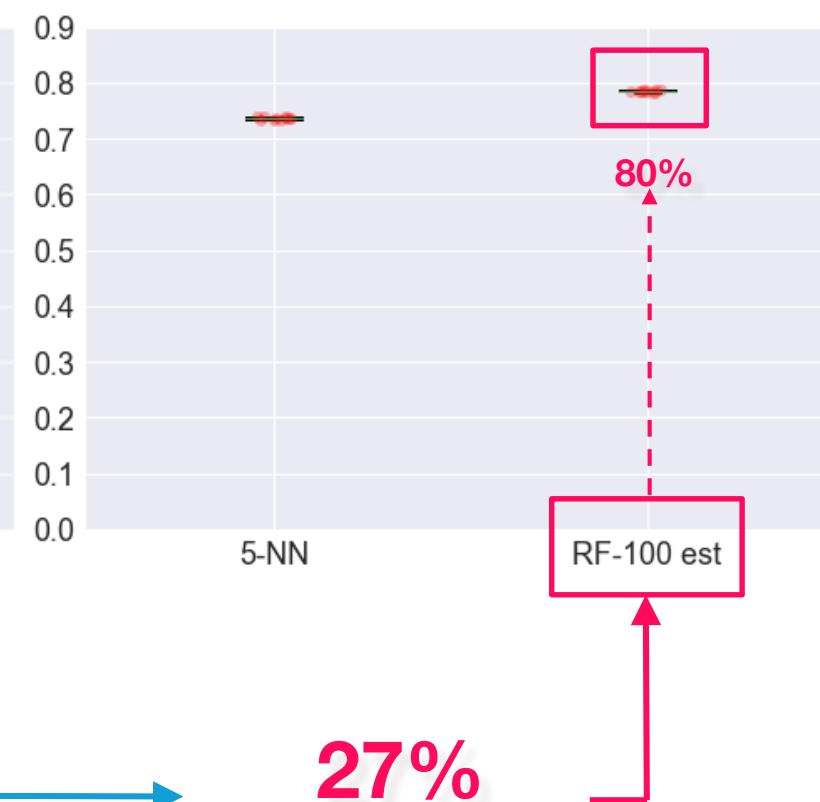
2

ADDING 8-BINS RGB
VARIABLE



3

ADDING PREVIOUS ACTIVITY



27%

IMPROVEMENT

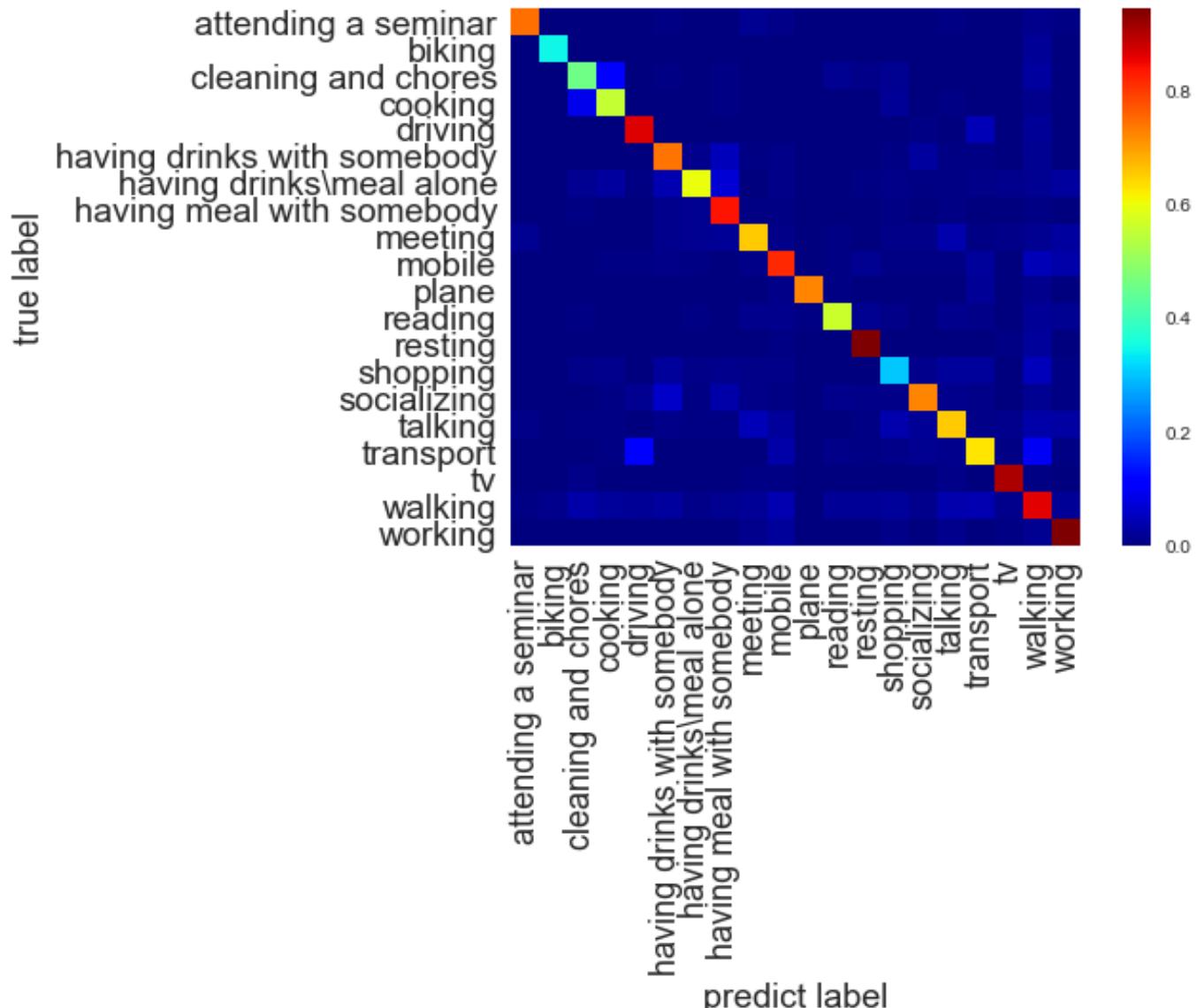
STEP 3: THE MODEL EVALUATION

TRUE LABELS

VERSUS

PREDICTED LABELS

NORMALIZED CONFUSION MATRIX



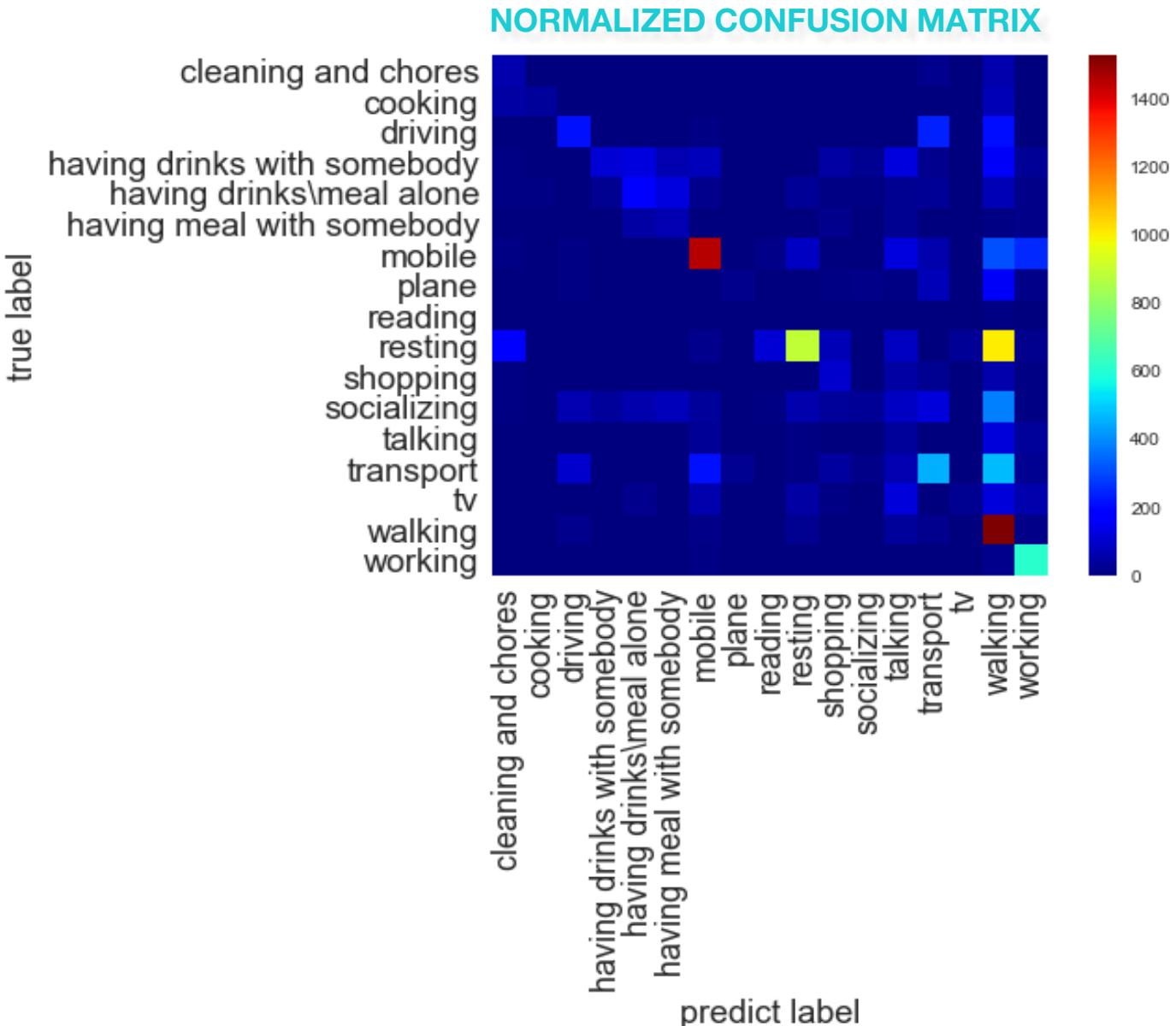
STEP 4: THE MODEL SCALABILITY

Can we train our system with one user and predict activities of for another user?

TRAIN MODEL BY USING DATA OF USER 1

TEST MODEL BY USING DATA OF USER 2

ACCURACY: 36% (-44%)

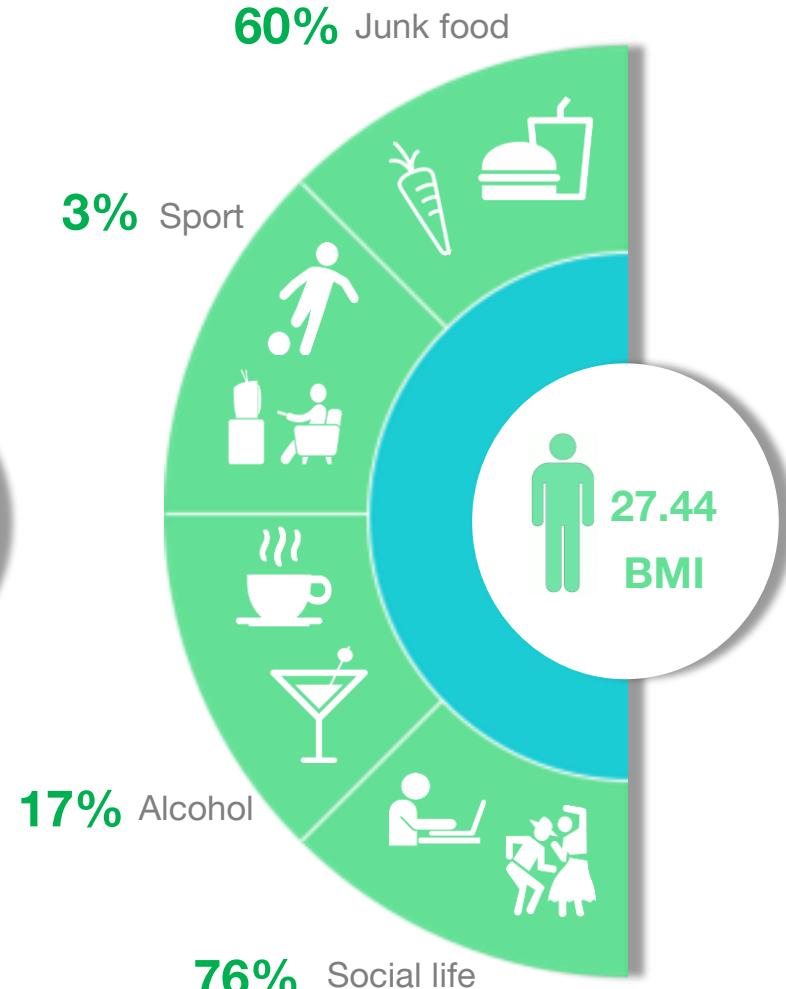
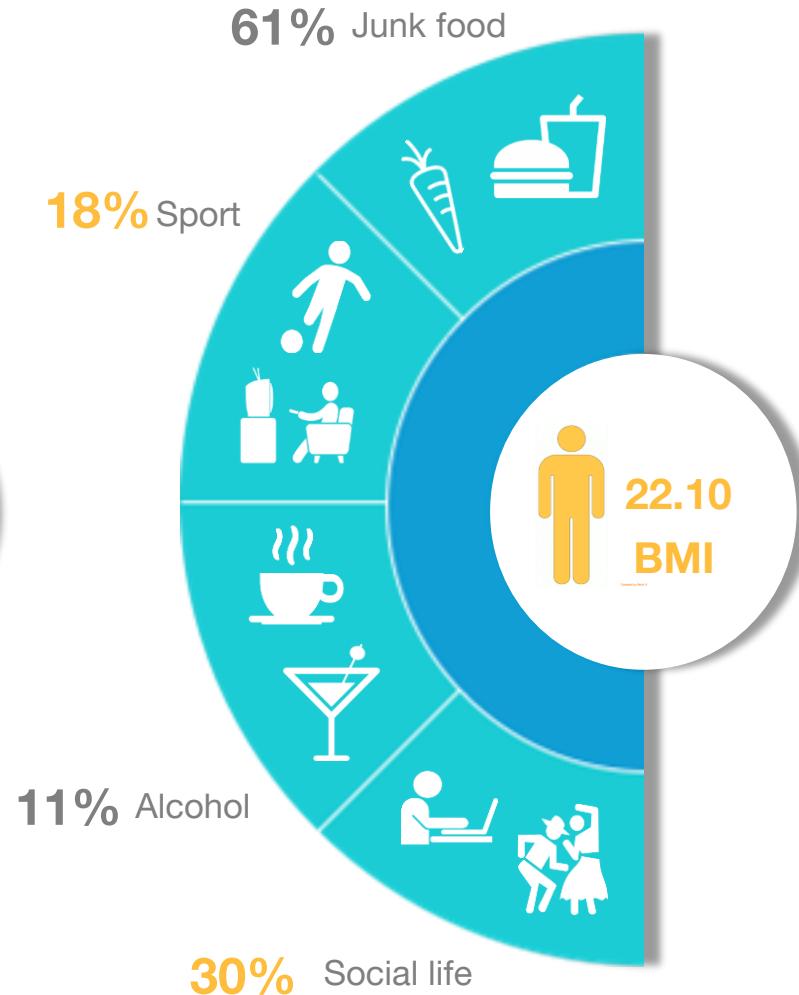
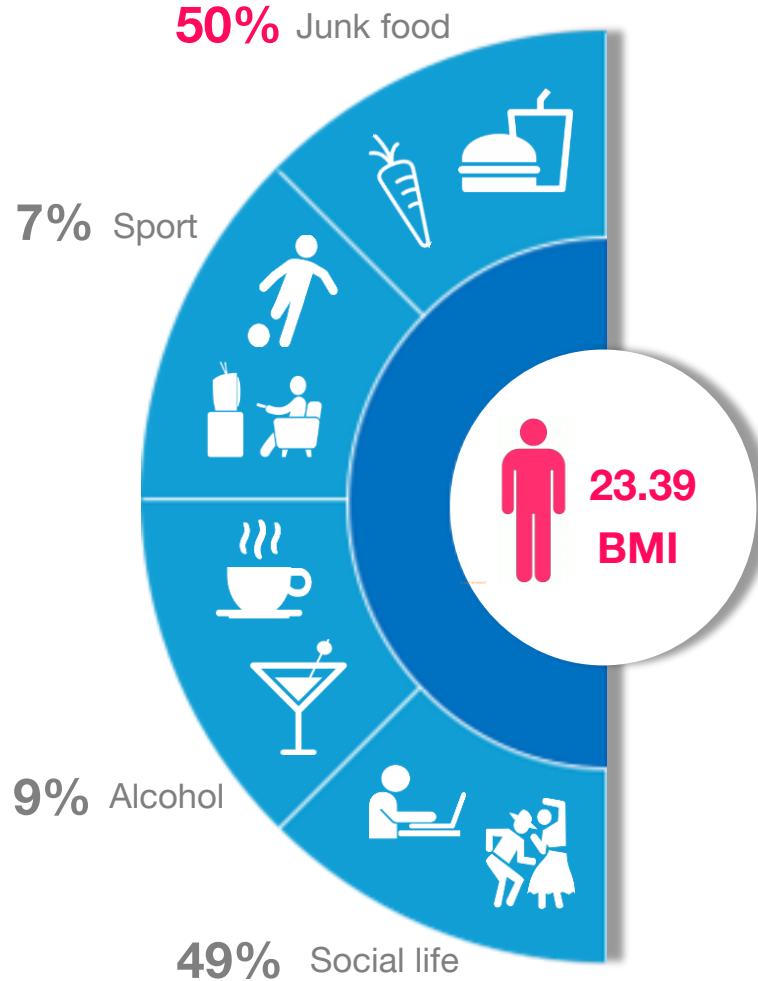


4 | WHAT IS A LIFELOGGER'S LIFESTYLE?



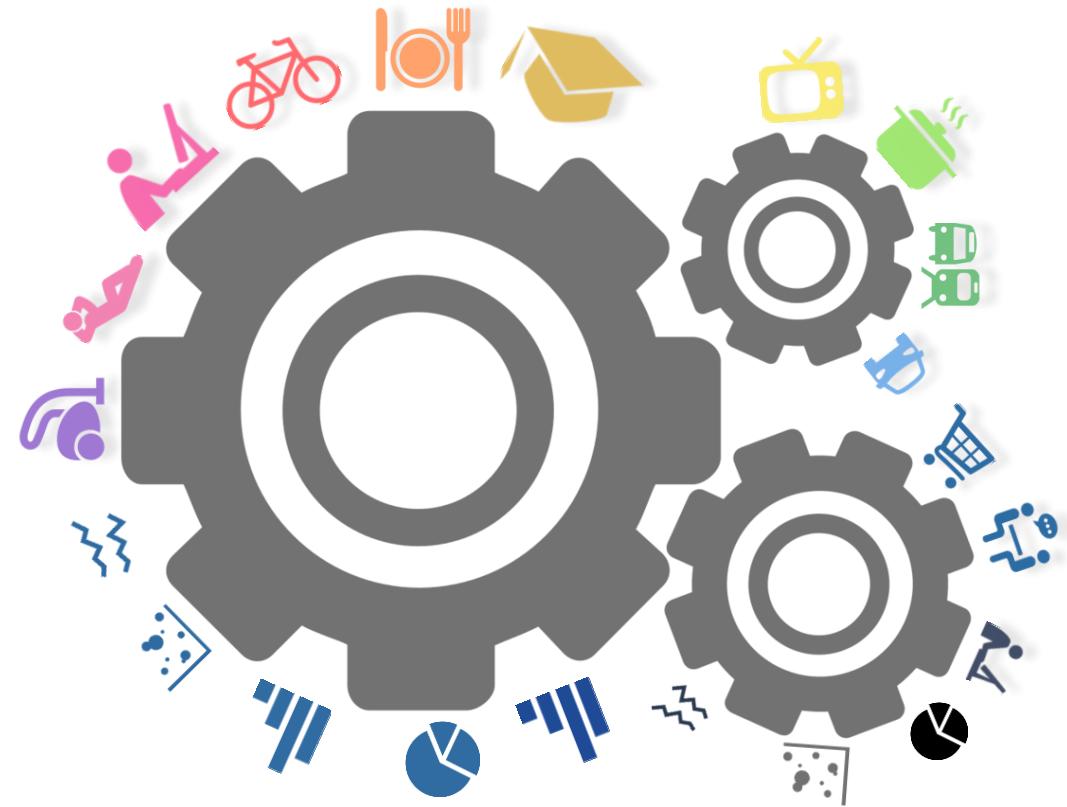
BMI AND LIFESTYLE

<18.5 18.5-24.9 25.0-29.9 >30.0



5 | ARE LIFELOGGERS PREDICTIBLE?

CONCLUSIONS





CONCLUSIONS

- 1 By using only images, it is possible to design a model which allows **to predict human activities** with **the accuracy close to 80%**.
- 2 Unfortunately, the **algorithm works** well when **is personalized** and applied to the same person. To extrapolate the algorithm from one user to another is not recommended.
- 3 **YOLO is an art-of-state solution.** However, Deep Learning tools and algorithms find many barriers to get round and be used widely.
- 4 The **data processing** and **data cleaning** is **70% of the success** in achieving a good prediction results.
- 5 The futuristic vision **of lifelogs become reality** thanks to the innovative technologies (e.g. smartwatches, smart wristbands)
- 6 **The lifelogs can find many useful applications**, for example: to increase health of the society, to build personal security tools, to improve human knowledge about themselves.



THANK YOU!

ALBERT COMELLAS DEL CASTILLO

MANUEL FORCADES

ALEIX LLENAS FARRÀS

ANETA MARCINKOWSKA

NÚRIA PUJOL VILANOVA



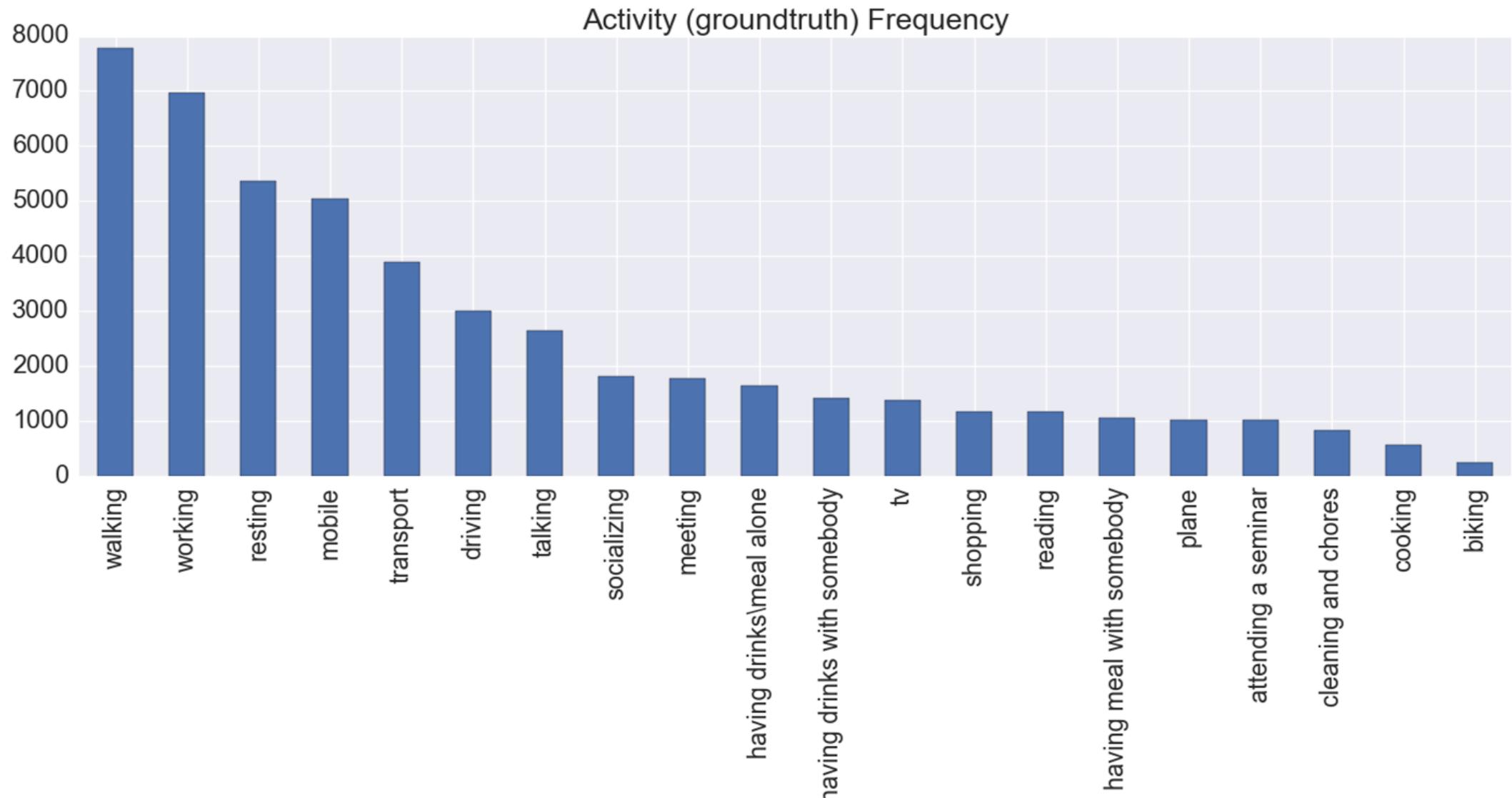
SUPERVISOR:
Dr. PETIA RADEVA



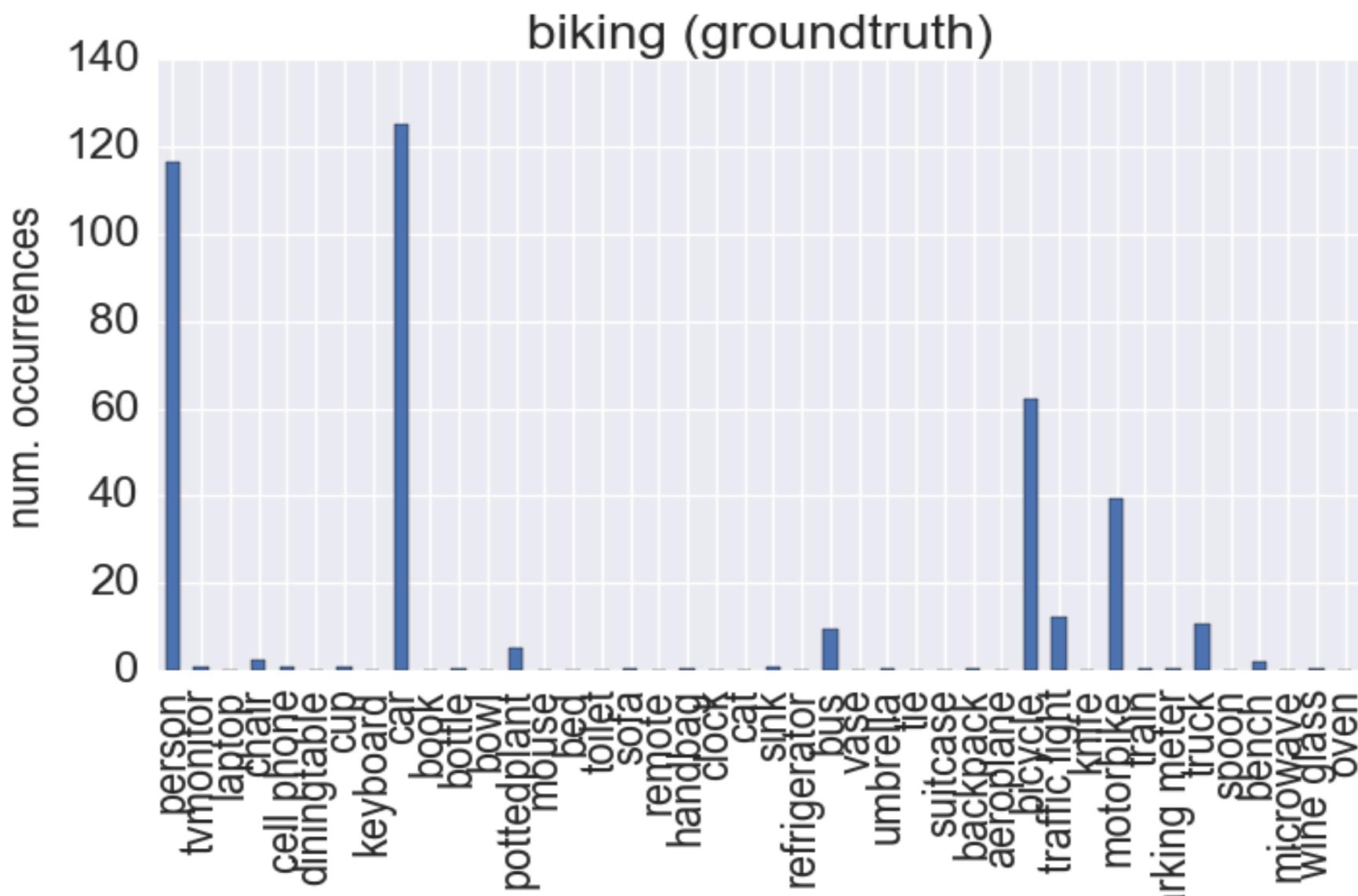


APPENDIX

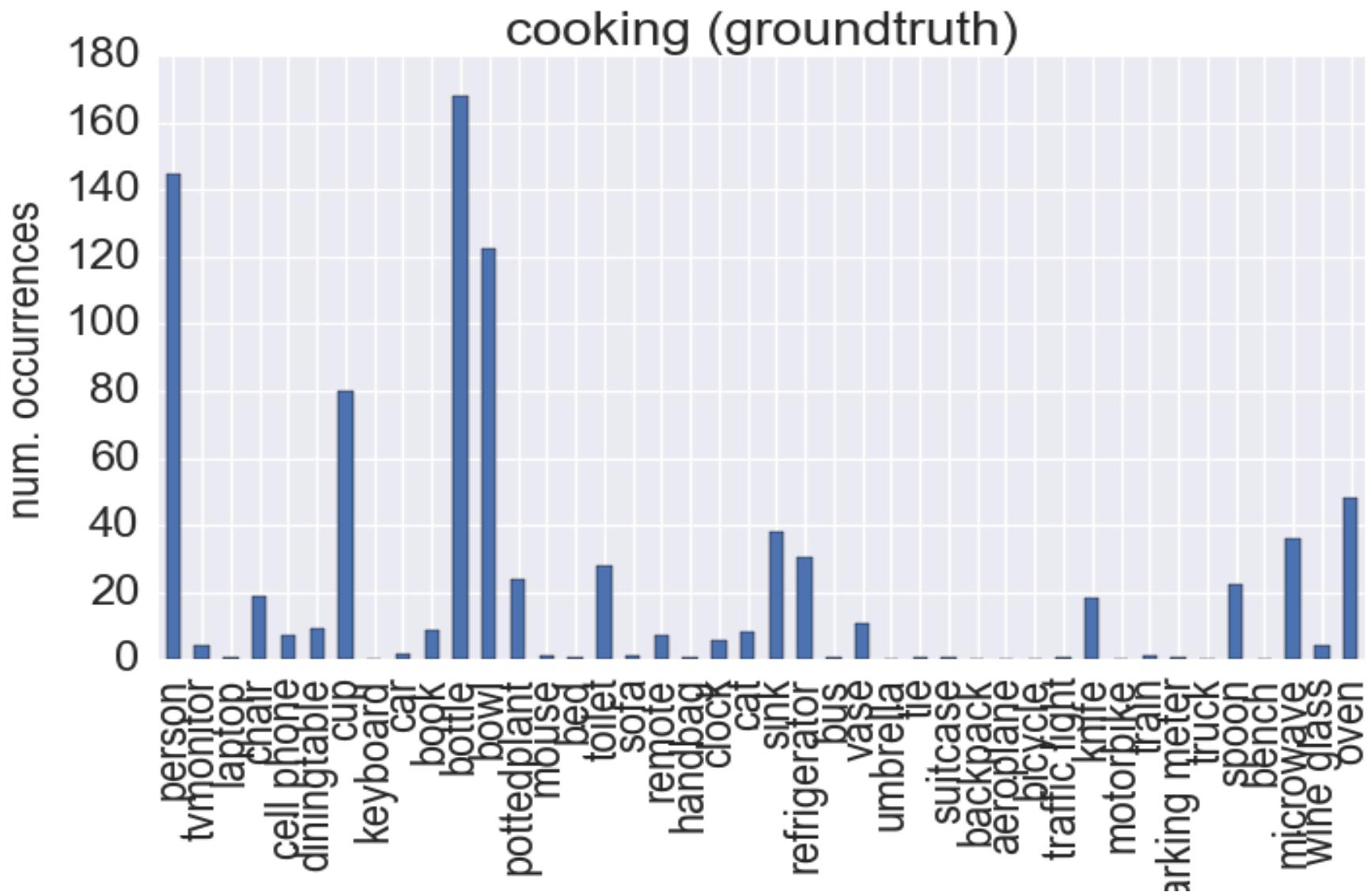
GROUND TRUTH: ACTIVITIES' FREQUENCY



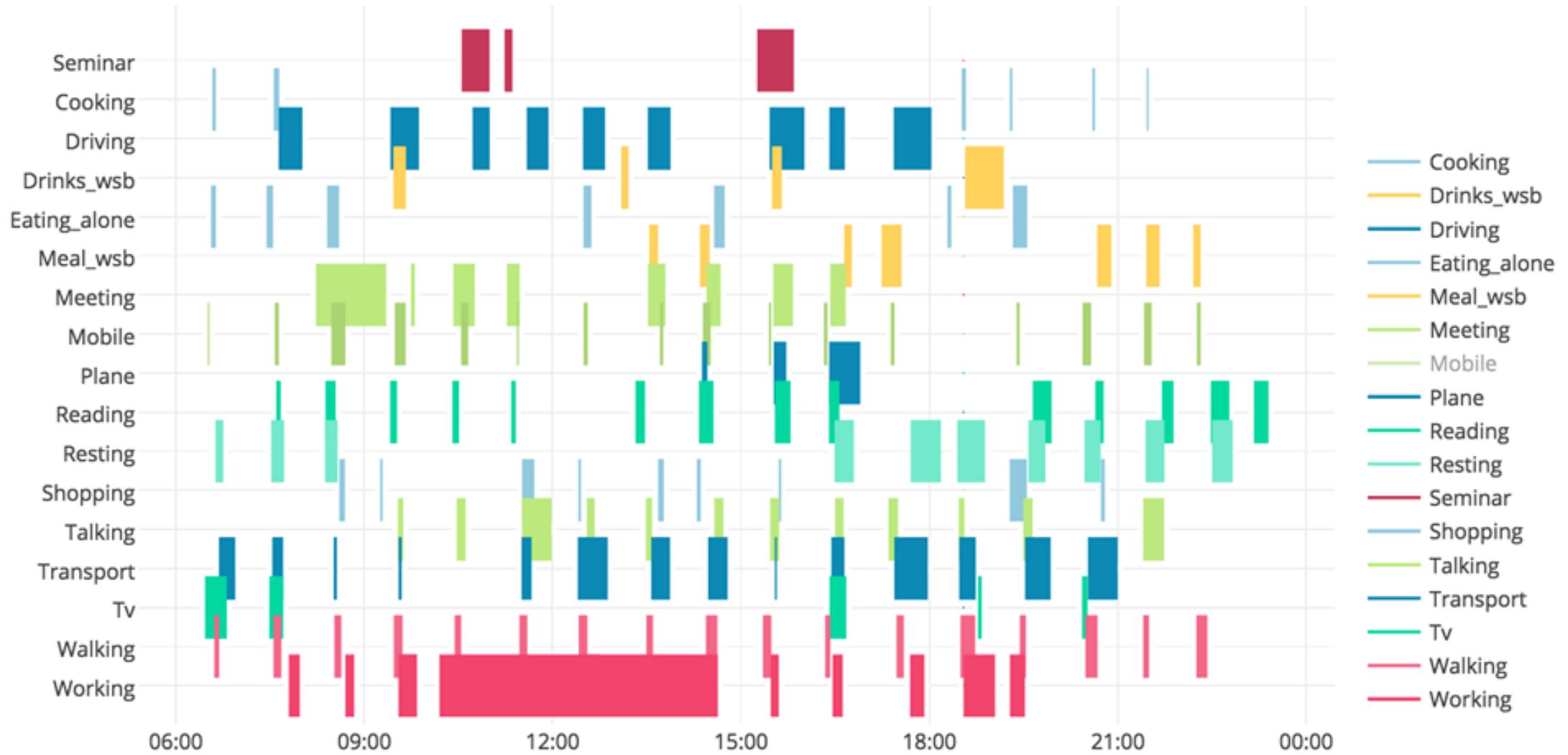
GROUND TRUTH: BIKING



GROUND TRUTH: COOKING



A DAILY SCHEDULE BY HOURS: USER 1



A DAILY SCHEDULE BY HOURS: USER 2

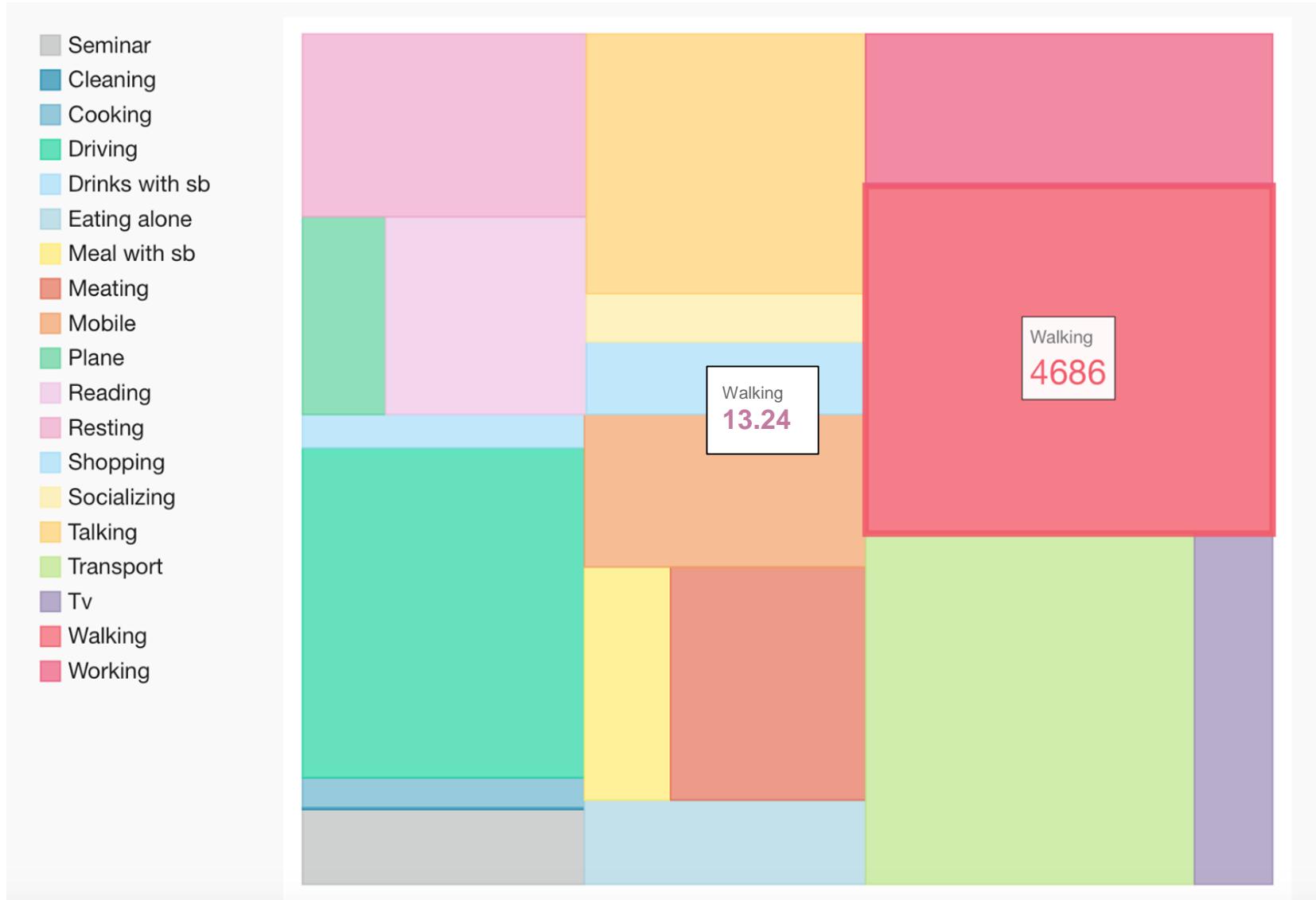


A DAILY SCHEDULE BY HOURS: USER 3



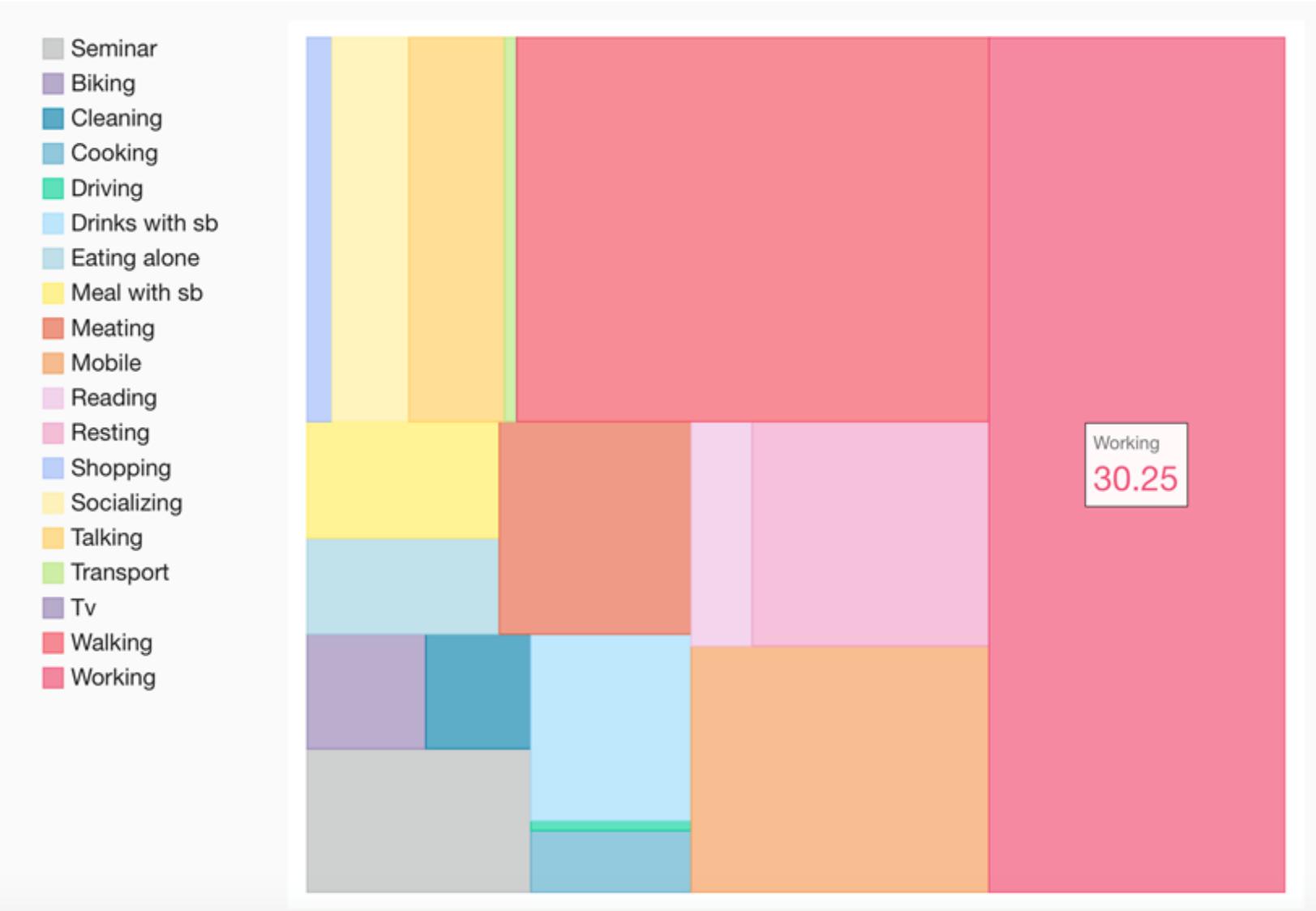
AN AMOUNT OF TIME SPENT ON AN ACTIVITY IN TOTAL (%)

USER 1



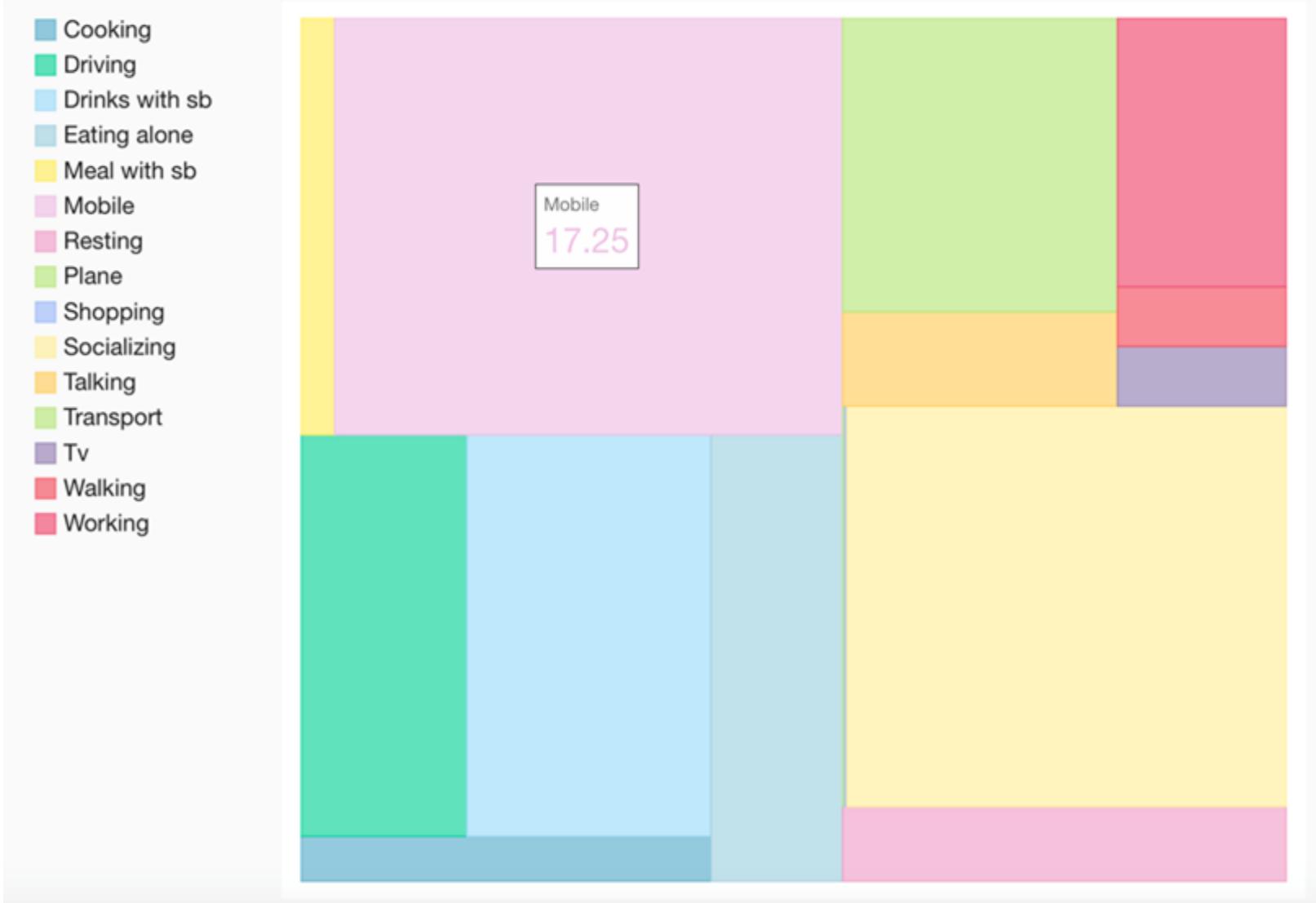
AN AMOUNT OF TIME SPENT ON AN ACTIVITY IN TOTAL (%)

USER 2

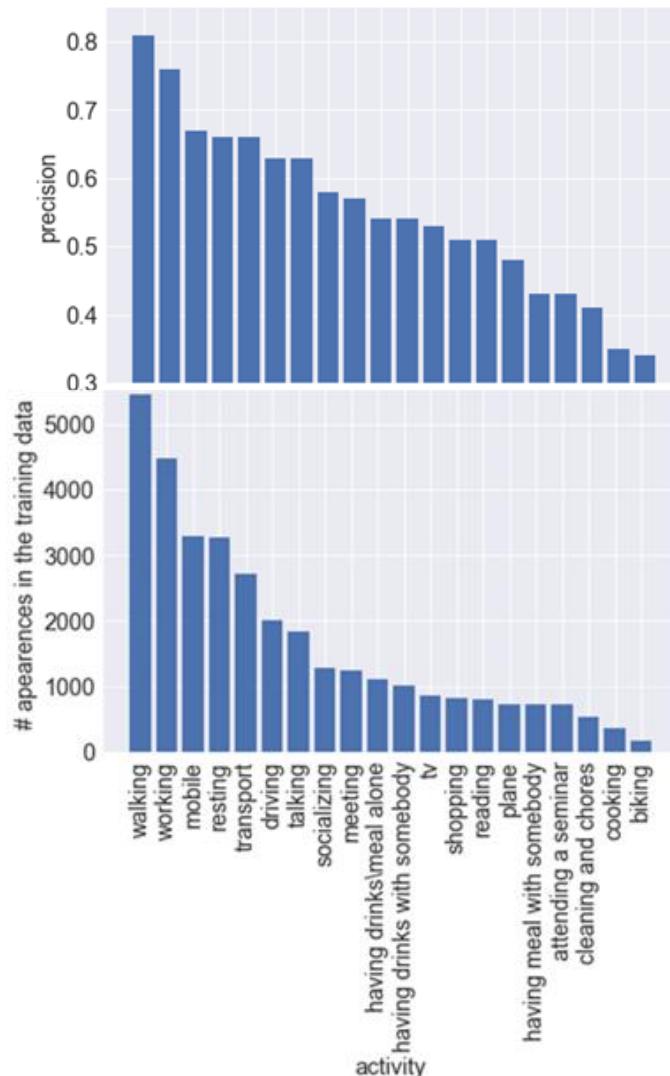


AN AMOUNT OF TIME SPENT ON AN ACTIVITY IN TOTAL (%)

USER 3



THE MODEL EVALUATION REGARDING N-SIZE AND PRECISION



THE PRECISION
INCREASE WITH THE
NUMBER OF
OBSERVATIONS IN
TRAINING SET

