Track & Know | Issue 7 | Date: October 2020

WELCOME TO THE SEVENTH TRACK & KNOW NEWSLETTER!

In this newsletter, you can find:

Project coordinator
Dr. Ibad Kureshi
Senior Research Scientist – Inlecom Systems
Square de Meeus 38/40
1000 Brussels
Belgium
Ibad.kureshi@inlecomsystems.com

ABOUT THE NEWSLETTER

This newsletter informs you about the results and activities of the EU H2020 research project Track& Know. The aim is to keep all relevant actors interested in managing big data, more specifically in the type of big data we focus on in the project and the tools/methods we develop to handle, analyse and visualize these datasets. T&K focuses on resolving key business cases for 3 test pilots, namely transport/mobility, insurance and health care. Business cases which will be explored in these pilots are as follows but not limited to: minimizing patients travel, carpooling and electric mobility potential, driver behaviour profiling etc.

Simulation of Electric Vehicles mobility and Individual mobility Networks - CNR (Italy)

Dr. Mirco Nanni and his team at CNR (Italy) developed a component to **simulate mobility of electric vehicles (EVs) using Individual mobility network.** This component is part of the Big Data Analytics (BDA) tool box and a key feature of the Insurance pilot of Track & Know project.

This component utilizes a **mix of mobility data analytics, ad hoc trip planning and performs a simulation to analyse the current fuel-based mobility of a user and quantitatively describes the impact of switching to EVs on a person's mobility life style.** Earlier similar studies that used mobility data investigated the typical mobility needs of the population and analysed how well they fit the typical characteristics of EVs mobility. Complete individual EV profiling was not studied in detail up to now. The objective of this work is to **find the Optimal EV Schedule for every user in order to waste as little time as possible during charging stops.** Starting with a set of information about the daily mobility routine of each driver we are going to investigate how to optimize their trips from the moment he or she owns an electric car instead of a common fuel car. The task can be defined as an **Electric Vehicle Route Planning With Recharging (EVRC**).

For the Tuscany region of Italy, the road network obtained from Open Street Map is **enriched with speed information and also altitude of nodes from Shuttle Radar Topography Mission (SRTM) data to drive slope of roads** which is a key attribute used in battery consumption model. Open Charge Map is used to obtain the information about recharging stations. Various recharge strategies are studied e.g.

- the user can recharge at the public station, at home and at her work place;
- the user can recharge at the public station and at home;
- the user can recharge at the public station and at her work place;
- the user can only recharge at the public station;

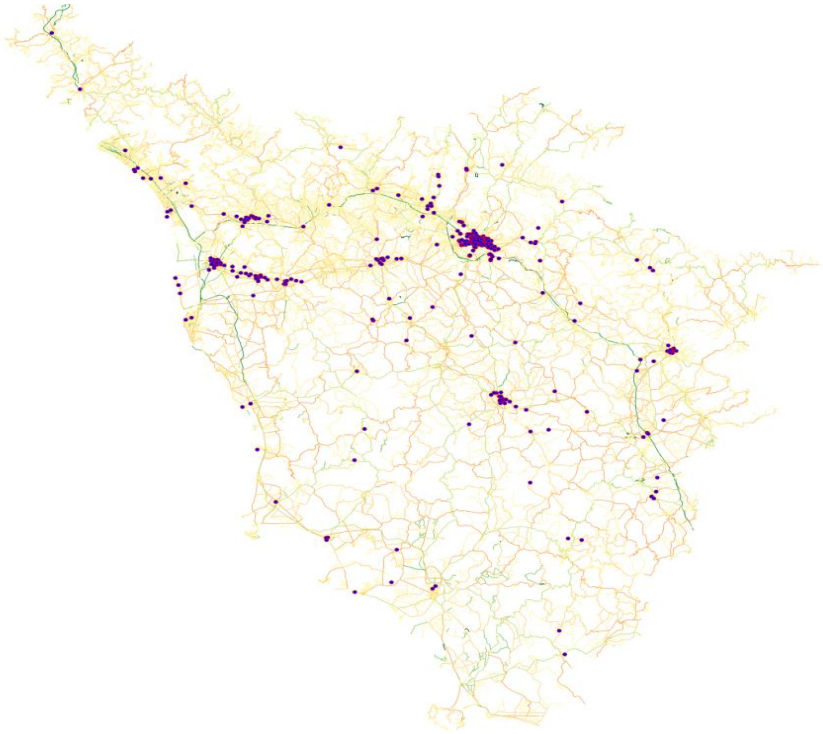


Fig. 1: Charging stations of Tuscany, Italy (source: Open Charge Map)

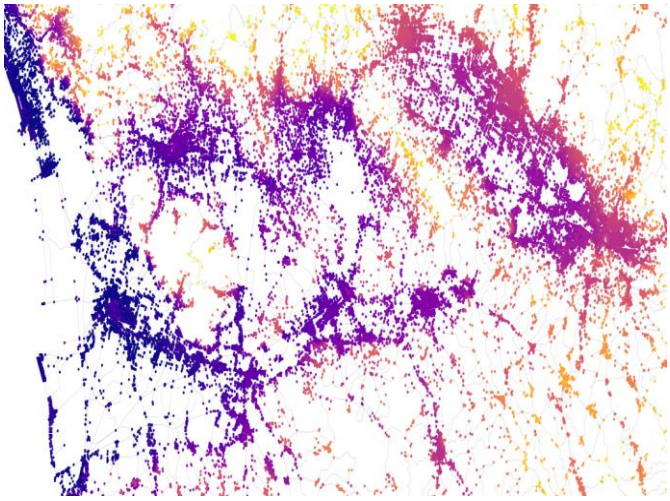


Fig. 2: Nodes altitude portion of Tuscany (source: SRTM)

For each scenario we calculate **how much battery the user has to charge in each charging option and how much time he has to wait for charging, as well as how much her original mobility (performed with a combustion engine) is affected by the limits of EVs, evaluating the expected increment in travel times and distances**. As an example, based on the mobility behaviour of the two users for a period January-February, 2017, the figure shows a timeline of a user (EV profiling) that recharges a little at home every day (green dots), and another one that has to rely only on public recharge stations and therefore recharges a lot more each time, yet more rarely (red dots).

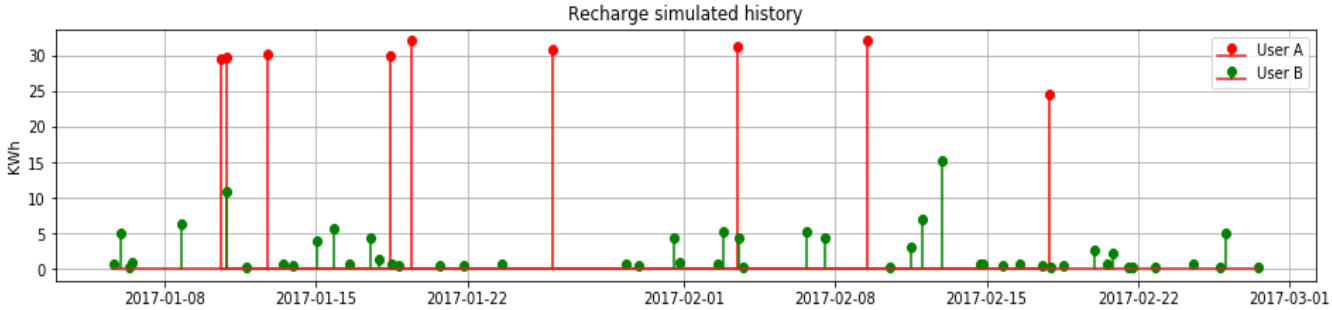


Fig. 3: Sample recharge timeline of one user who has access only to public stations (red), and one who can recharge at home (green).

The case study performed on a small sample of the Insurance Pilot dataset (obtained from Sistematica S.p.A (Italy)) shows that **the problematic situations for the users are rare, and the time spent for recharging can be very small if the options for recharging at home and/or at work are available.**

For more updates, you can contact **Dr. Mirco Nanni** via **mirco.nanni@isti.cnr.it**

Hot Spot Analysis - UPRC, Greece

Researchers from UPRC, Greece have developed **a method for Hot-Spot analysis with a view point that previous methods are based on spatiotemporal point data, not trajectories (sequences of points)**, and also parallel algorithms are designed and implemented in Apache Spark so they scale gracefully for Big Data.

Hot-spot analysis is the **problem of identifying statistically significant spatial clusters from an underlying data set**. A parallel and scalable algorithm (THS) for trajectory hot spot analysis is developed, using an adapted version of the Getis-Ord statistic tailored for trajectory data rather than point data. In brief, **they split the 3D spatio-temporal space in cells of user-defined granularity and then map the positions of moving objects to the cells. Then, they perform parallel processing of these cells to compute the Getis-Ord statistic for each cell, and therefore they are able to output the top-k cells according to their Getis-Ord value.** The method was tested with a qualitative test on 1GB of VFI Pilot data (fleet management pilot), and a scalability test was performed on a larger collection of approximately 90GB of mobility data. An example of hot-spot analysis over is illustrated in the figures.

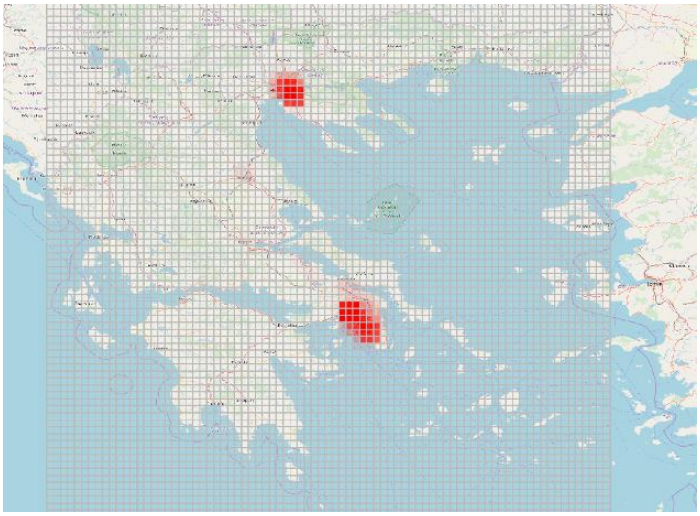


Fig. 4: Hot spots discovered in the wider area of Greece: most hot spot cells are located in Athens and Thessaloniki.

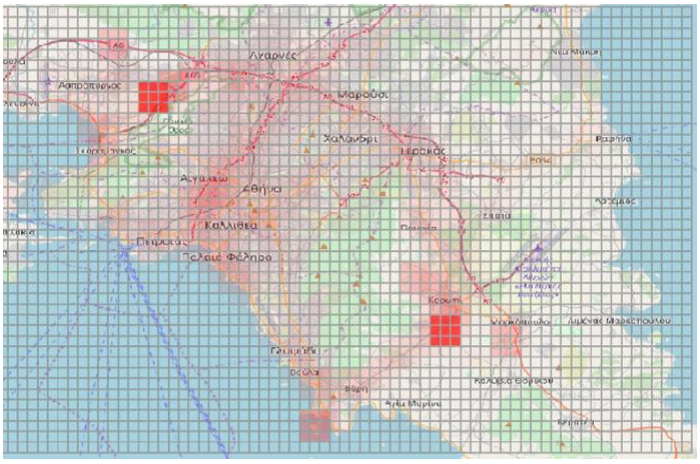


Fig. 5: Hot spots discovered when focusing on the area of Athens

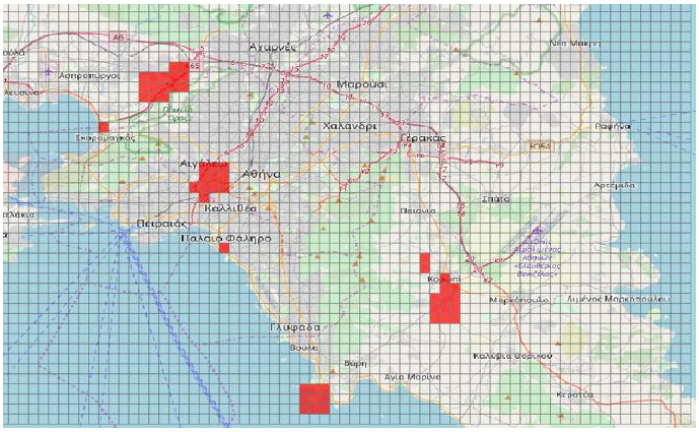


Fig. 6: Top-50 hot spots discovered when focusing on the area of Athens.

For more updates, you can contact **Dr. Christos Doulkeridis** via cdoulk@gmail.com

Identify driving patterns and categorize driving behaviour on-the-fly based on trajectory dynamics

- UPRC, Greece

Researchers from UPRC, Greece have developed a **method for driving behaviour profiling for online / offline trajectory analytics.** This is method is part of the BDA toolbox. The method is a completely unsupervised method (no ground truth required) that use sparse GPS location data (no accelerometer) and context-aware enrichments (local speed limit).

The **Driver Behavior Profiling (DBP) component** implements a data-driven approach to the challenge of analyzing, encoding and classifying driver behavioral patterns in the short- or the long-term, which in principle are of unknown categories. In this case, the core task of driver behavior profiling is addressed at the minimum level of pre-requisites, i.e., GPS-only trajectory data (no accelerometer or other sensors) of very low sampling rate (less than 0.1Hz). Additionally, the proposed approach is designed for online/streaming mode and lightweight yet powerful analytics, in order to be applicable to on-the-fly driver behavior profiling.

A dynamic temporal resampling algorithm is employed for transforming the sparse, variable-rate, GPS-only trajectory data into three distinct location-invariant time series, namely speed, acceleration and turn rate, after the raw GPS trajectories are map-matched to the underlying road network and noise-filtered for removal of artifacts. A wide range of statistical, time series and spectral methods are implemented as feature functions or `encoders' of various aspects of short-term mobility tracking.

An extensive real-world trajectory dataset is processed and transformed into such a feature-vector dataset, which is subsequently used in unsupervised training and adaptive category identification for the various driving behavior `states'. The results show that such an approach is feasible, despite this challenging context of constraints, providing a data-driven adaptive way to recognizing `normal' and `abnormal' driving patterns on-the-fly.

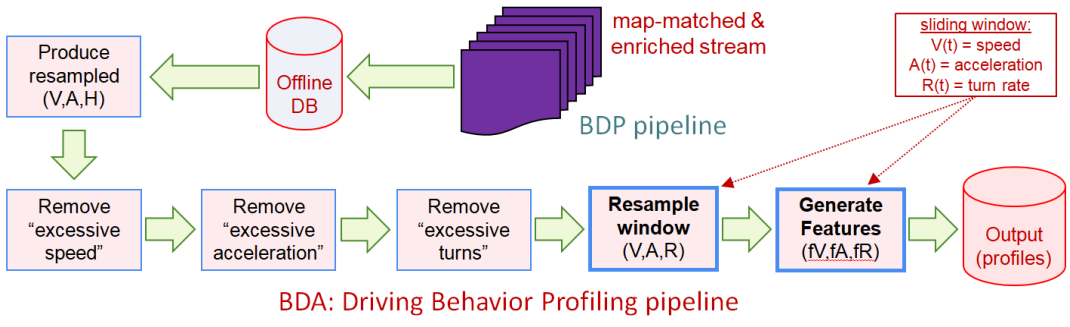


Fig. 6: Driver Behavior Profiling pipeline.

For more updates, you can contact **Dr. Harris Georgiou** via hgeorgiou@unipi.gr

Identifying business activity-travel patterns based on GPS data - UHASSELT, Belgium

Dr. Feng Liu at Hasselt University has developed a **method to identify business activity-travel patterns using GPS data.** This method is part of BDA tool box and aimed **to identify typical activity-travel patterns from business trips and characterize travel behaviour of specific companies or vehicles (and corresponding drivers) based on the obtained patterns.** The method and derived results will help uncover business activities and travel features, providing an improved behavioural mobility understanding, and exploring factors that would lead to addressing the increasing challenges related to business travel (e.g. environmental issues, driving safety, and travel demand management).

Currently, no methods exist yet to quantitatively analyse business travel behaviour in comparison to extensive body of knowledge and models aimed at analysing personal travel behaviour. Bridging this gap is important as business related travel also causing degradation of environment, and therefore, policies/strategies to curtail this travel are required to assess their impacts. The overall architecture of the method is illustrated in the below figure.

For more updates, you can contact **Dr. Feng Liu** via feng.liu@uhasselt.be

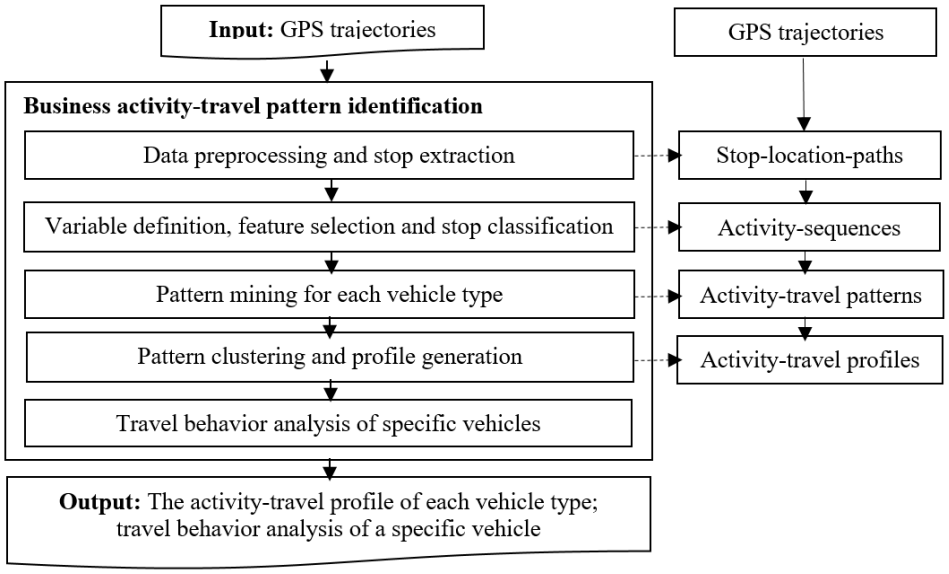


Fig. 7: Overall structure of the method

T&K Dashboards – UZH, Switzerland

T&K dashboard provides an **overview of some aggregated statistics, as well as the analytical results generated from the toolboxes**. In short, the dashboard does not serve as an interface of the analytical tools, but rather **visualizes the analytical results, which might be generated from off-line big data analytics (thus, it acts as a "strategic dashboard")**. It provides simple interactions (e.g., selection, filtering, and drilling down), allowing users to explore the analytical results on the browser side, partially as an analytical dashboard. It is implemented using web technologies and can be accessed via web browsers to maximize accessibility by end users from pilots.

In the background, it will connect to the database to receive **real-time updates of aggregated statistics and analytical results** and visualize them on the dashboard interface. Joint work with each pilot on configuring a dashboard with application-centered customized information is in the progress.

An intra-project data transferable format protocol for technical partners to deliver outputs to the dashboard is proposed. The outputs are planned to use the Kafka platform as the intermediate while the dashboard will consume these outputs from Kafka. Specifically, for each pilot the prelimany versions of the dashboard view are presented in Figure 8.

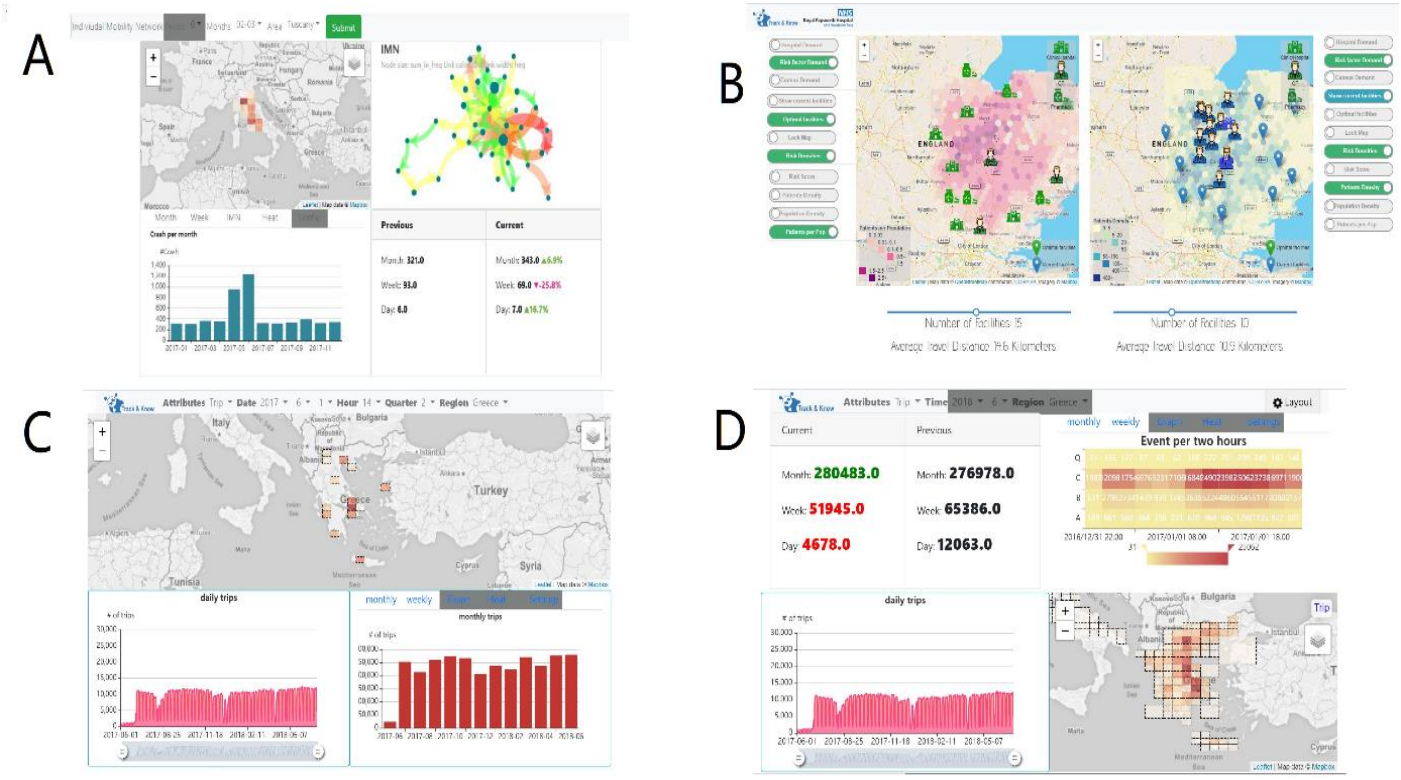


Fig. 8: Views of dashboards for SIS (A), PAP(B), and VFI (C for technicians, D for managers).

Big Data Pilot Demo Days

The new data-driven industrial revolution highlights the **need for big data technologies to unlock the potential in various application domains.** To this end, BDV PPP projects I-BiDaaS, BigDataStack, Track & Know and Policy Cloud deliver innovative technologies to address the emerging needs of data operations and applications. To enable data operations and data-intensive applications to fully exploit the sustainability and take full advantage of the developed technologies, the PDV PPP projects brought on board use cases that exhibit their applicability in a wide variety of sectors.

This series of webinars aimed at **showcasing the implementation of the Big Data technologies in the pilot studies and their applicability to an ever wider scope.** The webinars demonstrate the actual solutions implemented performing big data operations and applications to interested end-users from industry as well as technology providers for further adoption in their own solutions and projects. The projects jointly illustrated how they contribute to Europe's digital future.

The series included **3 Track & Know webinars which showcased the pilot studies**, introduced by **Jenny Rainbird**, representing project coordinator Inlecom. All 3 webinars started off with a comprehensive **presentation of the Track & Know Big Data Integration platform** by **Marios Logothetis** from project partner Intrasoft International.

On July 7th, **Athanasios Koumparos** from project partner Vodafone Innovus presented the fleet management pilot, focussing on **fleet management services: Quality and predictions in location data from GPS devices.**

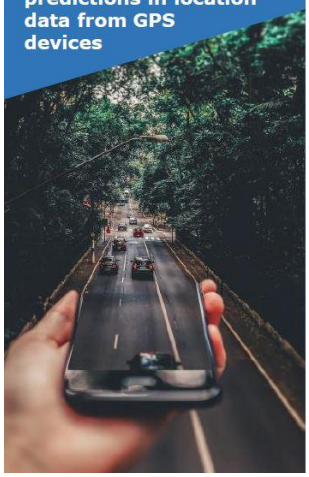
On July 14th, **Leonardo Longhi** showcased the insurance pilot and more specifically he elaborated on the **use of mobility data to understand and mitigate risky driving behaviour.**

On July 16th the health care pilot was put in the spotlights. **Toni Staykova, Ian Smith, Kieran Lee and Livio Brühwiler** explained **how effective use of patient mobility information can help to understand the provision of services across large rural and urban communities.**

The Track & Know webinars have been recorded and are currently available on the **Track & Know YouTube channel**. **Please click the below images to view the recordings**.

Big Data Pilot Demo Days: Track & Know

Fleet management services: Quality and predictions in location data from GPS devices

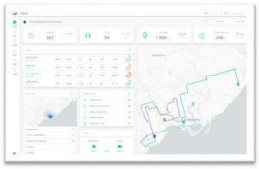


Biography

Athanasios Koumparos is a **senior software engineer** working on fleet management projects over a decade. Currently his main position is managing research projects at **Vodafone Innovus** and is interested in mobility in urban areas and big data analytics.

"Our mission is to select quality data and make valuable predictions in fleet management."

Athanasios Koumparos



Track & Know is a **Horizon2020 project**, with a focus on Big Data. More specifically, Track & Know will research, develop and exploit **a new software framework that aims at increasing the efficiency of Big Data**. This will be applied in the transport, mobility, motor insurance and health sectors.

Info: trackandknowproject.eu

Track & Know

Join this webinar LIVE!

Tuesday July 7th 2:00 pm – 3:00 pm CET
Register now by scanning this QR code.

This webinar is part of the Big Data Pilot Demo Days, hosted by BDV PPP.

Big Data Pilot Demo Days: Track & Know

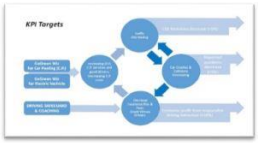
Insurance Sector: Using mobility data to understand & mitigate risky driving behaviour

Biography

Leonardo Longhi graduated in Biomedical, Electronic & Communication Engineering at the Roma Tre University. His M.Sc. thesis was about **Pattern Recognition methods for ML applications to recognize drowsiness & lack of focus by analysing facial expressions**. After collaborating in developing businesses for Italian start-ups in biomedical applications & in the signal processing business area, he shifted to **applied informatics in insurance & telematics products**. He is currently a consultant for a major telematics companies in Europe.

"Technologies for New Sustainable Mobility."

Leonardo Longhi



Track & Know is a **Horizon2020 project**, with a focus on Big Data. More specifically, Track & Know will research, develop and exploit **a new software framework that aims at increasing the efficiency of Big Data**. This will be applied in the transport, mobility, motor insurance and health sectors.

Info: trackandknowproject.eu

Track & Know

Join this webinar LIVE!

Tuesday July 14th 2:00 pm – 3:00 pm CET
Register now by scanning this QR code.

This webinar is part of the Big Data Pilot Demo Days, hosted by BDV PPP.

Big Data Pilot Demo Days: Track & Know

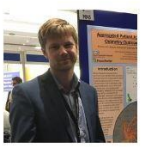
Healthcare Service: Using patient mobility information to understand the provision of services across large rural & urban communities

Healthcare Service: Using patient mobility information to understand the provision of services across large rural and urban communities - July 16th, 2 pm – 3 pm CET

Dr **Toni Staykova** is a **specialist physician and geriatrician** who has been actively involved in clinical innovation projects in Cambridge and across the EU over the past 10 years.

Dr **Ian Smith** is the **Clinical Director of the sleep service** at Royal Papworth Hospital. His research interests include designing new ways of delivering clinical services for patients with sleep disorders and neurological conditions in particular motor neurone disease, incorporating new pathways and monitoring devices to improve patient safety and maximise accessibility to specialist led care.

Livio Brühwiler is a **MSc Student in Geographic Information Science** at the University of Zurich. He is currently working on his Master Thesis on the topic of car accident risk prediction using GPS trajectories.



Kieran Lee is a **research practitioner within the R&D team** at Royal Papworth Hospital. He is involved in a range of projects within the field of sleep medicine, including drug and neurological profiling research, but with a primary focus on improving service accessibility and efficiency.

Track & Know is a **H2020 project**, with a focus on Big Data. The project will research, develop & exploit **a new software framework that aims at increasing the efficiency of Big Data**. This will be applied in the transport, mobility, motor insurance and health sectors. **Info: trackandknowproject.eu**

Join this webinar LIVE!
Register now by scanning the QR code.

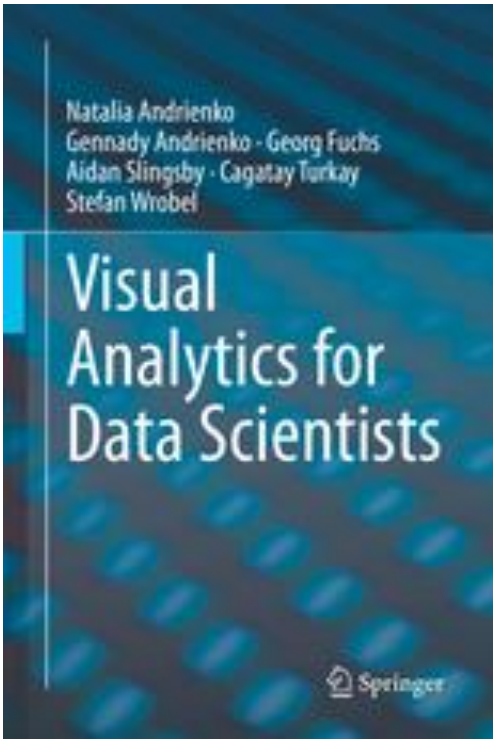
This webinar is part of the Big Data Pilot Demo Days, hosted by BDV PPP.

Book release

Recently, the book **Visual Analytics for Data Scientists** was released. Amongst the authors of this textbook are **Natalia Andrienko and Gennady Andrienko** from project partner Fraunhofer Instute IAIS.

This textbook presents the **main principles of visual analytics and describes techniques and approaches that have proven their utility and can be readily reproduced.** Special emphasis is placed on various instructive examples of analyses, in which the need for and the use of visualisations are explained in detail.

The content presented in various chapters of the book are resulted from detailed investigation of **various techniques/approaches that are used and incorporated in the Visual Analytics toolbox of the track & Know project**.



Natalia and Gennady Andrienko are lead scientists responsible for visual analytics research at the Fraunhofer Institute for Intelligent Analysis and Information Systems (IAIS) in Germany since 2007 and part-time professors at City, University of London since 2013. They co-authored monographs "Exploratory Analysis of Spatial and Temporal Data" (Springer, 2006) and "Visual Analytics of Movement" (Springer, 2013), and more than 100 peer-reviewed journal papers. Their research interests include geovisualization, information visualization with a focus on spatial and temporal data, visual analytics, interactive

RECENT PUBLICATIONS

- Adnan, M, Gazder U., Yasar A. H., Bellemans T., Kureshi, I (2020), Estimation of travel time distributions for urban roads using GPS trajectories of vehicles: a case of Athens, Greece, *Personal and Ubiqutious Computing*, DOI: https://doi.org/10.1007/s00779-020-01369-4, **Download here**
- Theodoridis Y. (2020) Learning from Our Movements – The Mobility Data Analytics Era. In: Tserpes K., Renso C., Matwin S. (eds) Multiple-Aspect Analysis of Semantic Trajectories. MASTER 2019. Lecture Notes in Computer Science, vol 11889. Springer, Cham. **Download here**
- Tampakis, P., Doulkeridis, C., Pelekis, N., Theodoridis, Y. (2020). "Distributed Subtrajectory Join on Massive Datasets." ACM Digital Library. Published version: https://dl.acm.org/doi/10.1145/3373642
- Nikitopoulos, P., Sfyris, G.A., Vlachou, A., Doulkeridis, C., Telelis, O. (2020) "Pruning Techniques for Parallel Processing of Reverse Top-k Queries." Distributed and Parallel Databases (Springer), 2020. Published version: DOI 10.1007/s10619-020-07297-9
- Guidotti, R., Nanni, M. (2020). Crash Prediction and Risk Assessment with Individual Mobility Networks. The 21st IEEE International Conference on Mobile Data Management (MDM 2020). Versailles, France, 2020, pp. 89-98. **Download here**
- Yeghikyan, G., Opolka, F., Lepri, B., Nanni, M., Lio, P. (2020) Learning Mobility Flows from Urban Features with Spatial Interaction Models and Neural Networks. IEEE International Conference on Smart Computing (SMARTCOMP). **To appear soon. Pre-print version here**.
- Liu, F., Andrienko, G., Andrienko, N., Chen, S., Janssens, D., Wets, G., Theodoridis, Y. 2020. "Citywide traffic analysis based on the combination of visual and analytic approaches." Download **here**
- Andrienko, N., Andrienko, G. "Spatio-temporal visual analytics: a vision for 2020s." Journal of Spatial Information Science, 2020, number 20, pp.87-95. **Download here**
- Tritsarolis A, Theodoropoulos GS, Theodoridis Y (2020). "Online discovery of co-movement patterns in mobility data." Int. J. Geographical Information Science, Taylor & Francis. **Pre-print available here**
- Andrienko, G., Andrienko, N., Patterson, F., Chen, S., Weibel, R., Huang, H., Doulkeridis, C., Georgiou, H., Pelekis, N., Theodoridis, Y., Nanni, M., Longhi, L., Koumparos, A., Yasar, A. and Kureshi, I. "Visual Analytics for Characterizing Mobility Aspects of Urban Context." Wenzhong Shi, Michael Goodchild, Michael Batty, Mei-Po Kwan, Anshu Zhang (Eds.) **Urban Informatics**. Springer, 2020. **Pre-print available here.**
- Andrienko, N., Andrienko, G. "Visual Analytics of Vessel Movement." Alexander Artikis and Dimitris Zissis (Eds.) **Maritime Informatics**. Springer, 2020. **Pre-print available here.**
- Andrienko, N., Andrienko, G., Fuchs, G., Slingsby, A., Turkay, C., Wrobel, S. **Visual Analytics for Data Scientists**. Springer, 2020. **Published version available here.**

Full list of publications 2018-2019-2020 is available HERE

Find & Follow us – www.trackandknowproject.eu - Twitter: @TrackandKnow - LinkedIn: Track & Know

The project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No 780754.