



EUROPEAN
**BIGDATA
VALUE** FORUM

**BERLIN + VIRTUAL
3-5 NOVEMBER 2020**

MacBook Pro

EUROPEAN
**BIG DATA
VALUE** FORUM
3 - 5 NOVEMBER . 2020 - BERLIN + VIRTUAL

Learning from our movements – the Track&Know EU project

Prof. Yannis Theodoridis
Data Science Lab., University of Piraeus (GR)



Outline

Movement data – the evolution

The Track&Know project

Objectives

Datasets

Architecture

Big Data Processing

Big Data Analytics

Complex Event Recognition

Visual Analytics

(all-in-one) Big Mobility Data Integration

Datasets

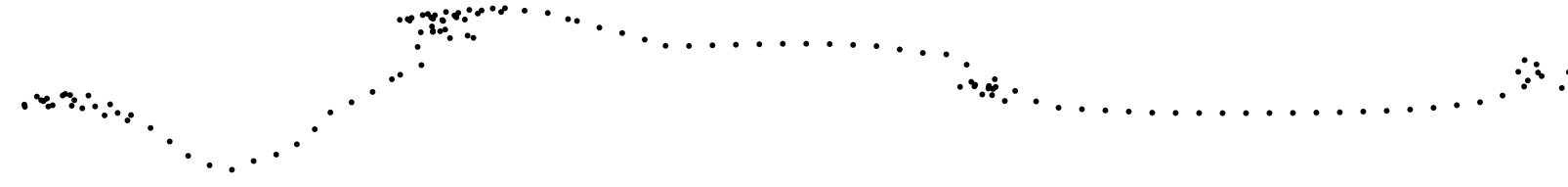


Movement data evolution

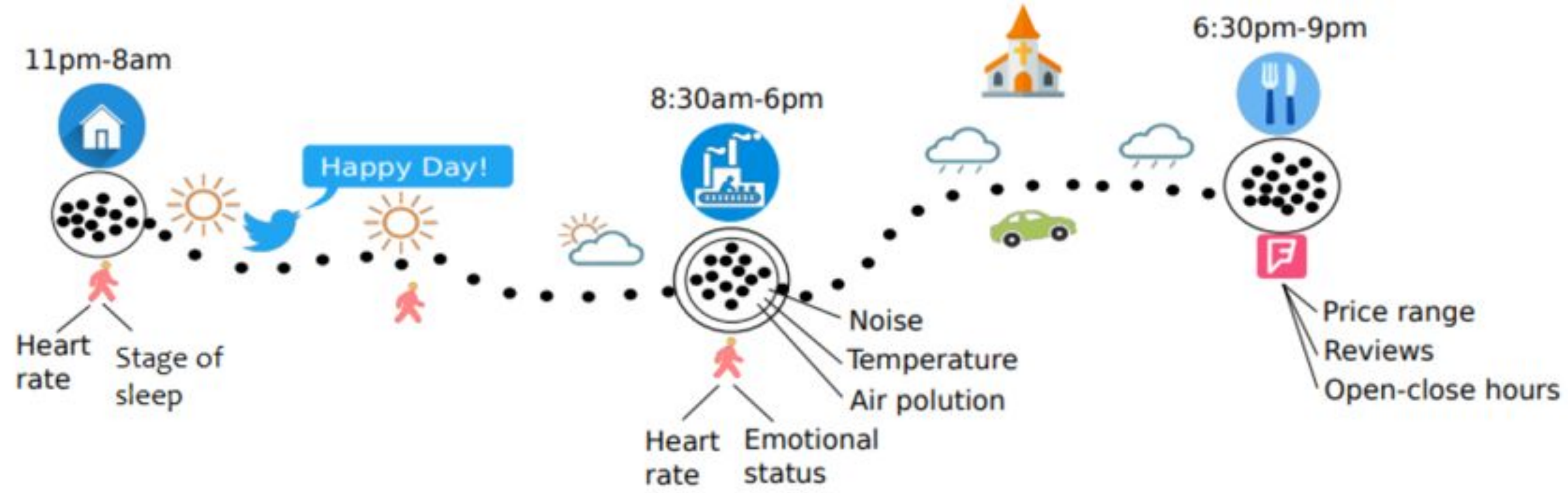


From **spatio-temporal DBs (back in 1990's)**.....

$$T = \{ \langle p_1, t_1 \rangle, \langle p_2, t_2 \rangle, \dots, \langle p_n, t_n \rangle \}$$



... to **Mobility timelines (today)**



Track&Know objectives



Big data for mobility tracking knowledge extraction in urban areas

*“... research, develop and exploit a new **software framework**
... that aims at increasing the efficiency of Big Data applications in
the **transport, mobility, motor insurance and health sectors**
... integrate multidisciplinary research teams from
Mobility Data management, Complex Event Recognition,
Geospatial Modelling, Complex Network Analysis,
Transportation Engineering and Visual Analytics
... to develop new models and applications. ...”*



Track&Know datasets -1

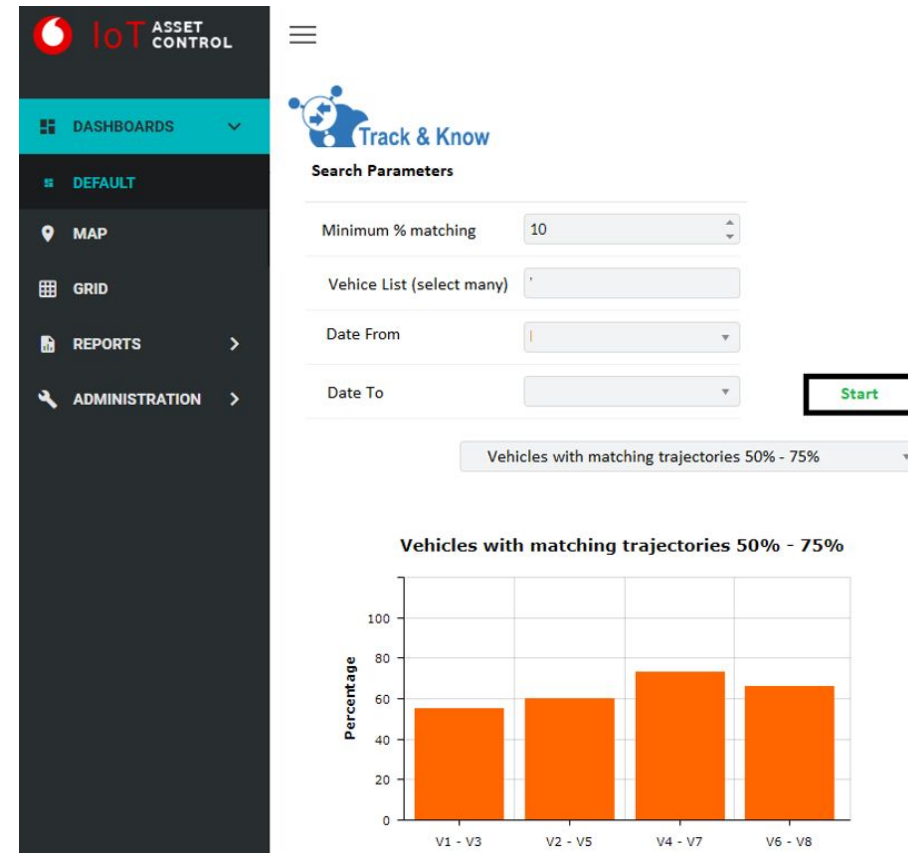
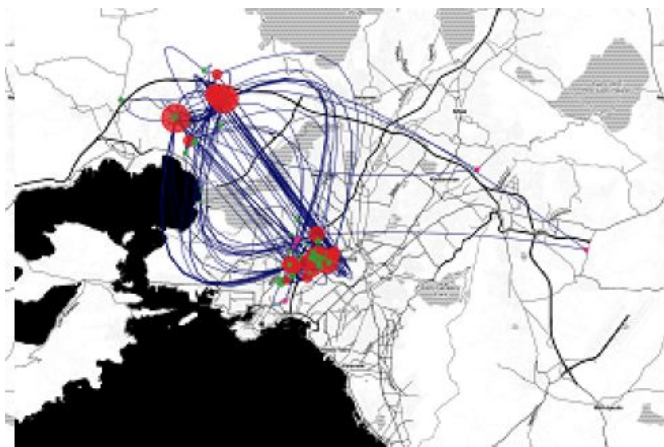


Transportation: Vodafone Innovus (VFI) pilot

- Millions of trips (mainly, trucks)
- Integrated with weather and points-of-interest info

- Objectives:

- Energy consumption monitoring,
- Individual Mobility Networks (IMN), etc.



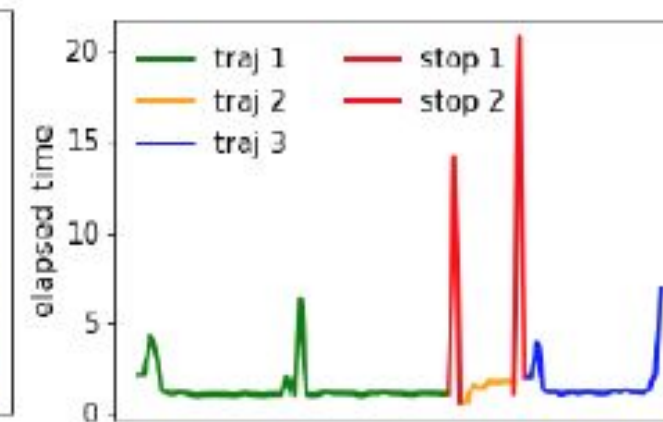
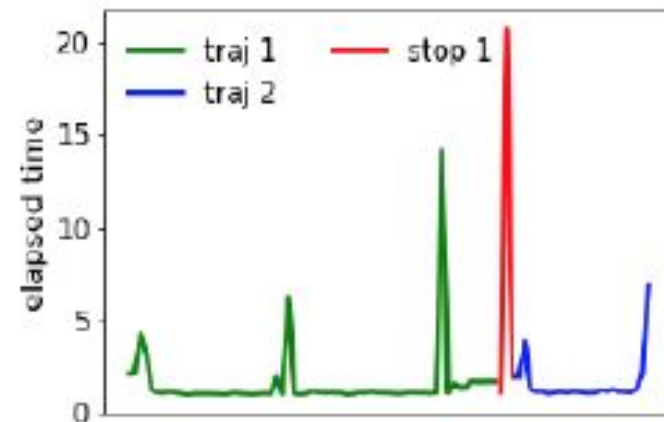
Track&Know datasets -2



Car Insurance: Sistematica (SIS) pilot

- Road traffic
- Activity concentration
- Origin-Destination matrix

- Objectives:
 - Crash risk prediction,
 - Car pooling, etc.

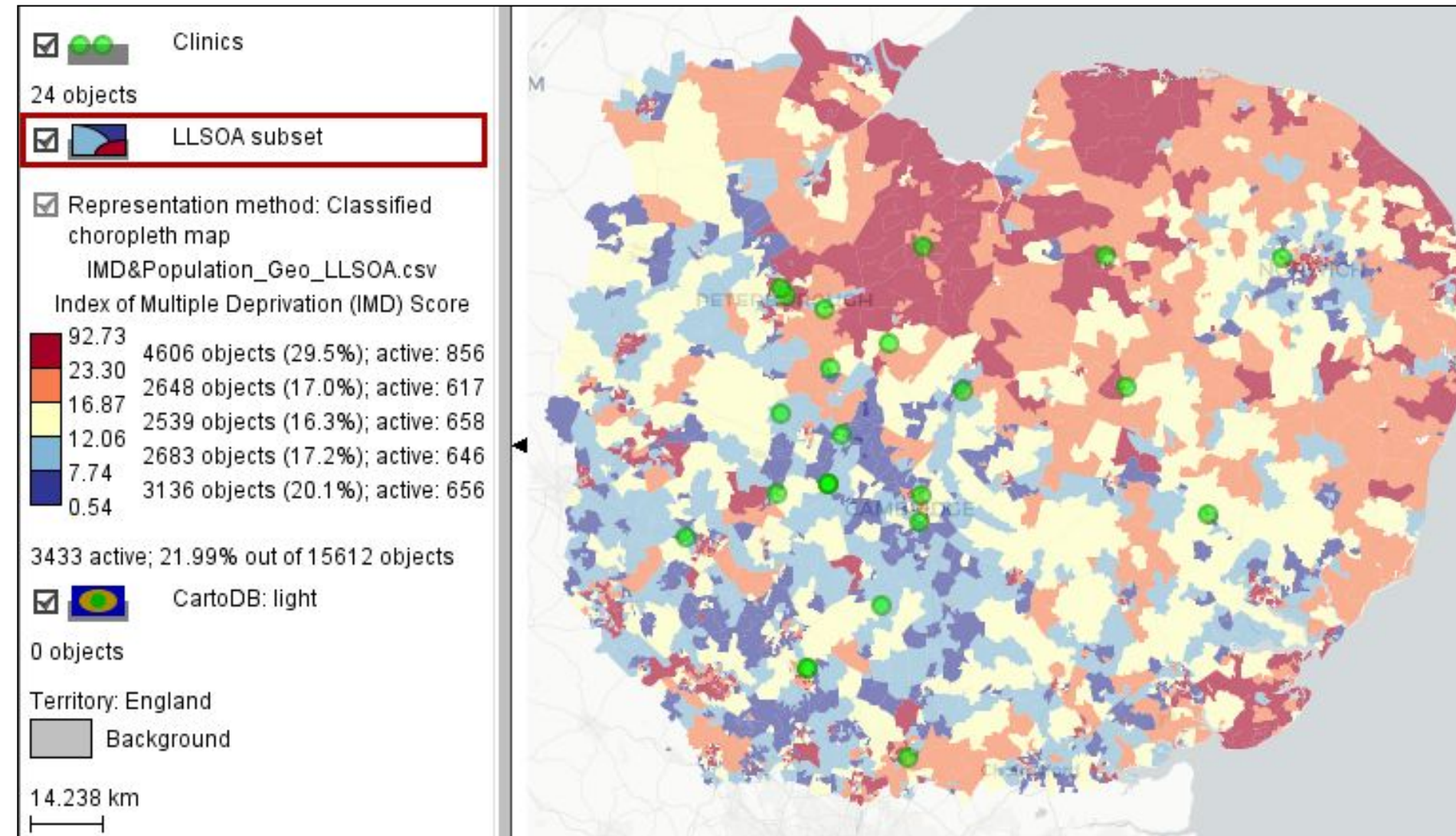


Track&Know datasets -3



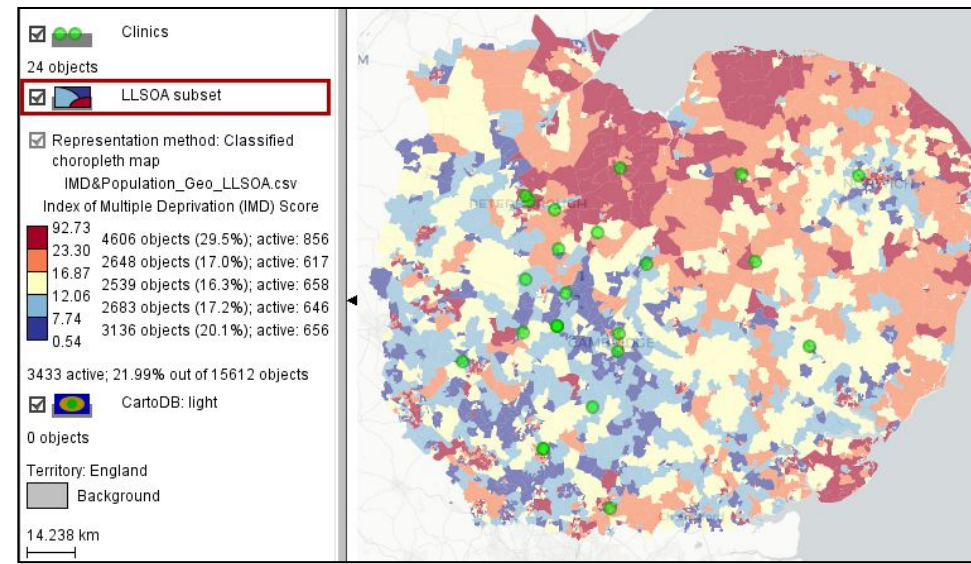
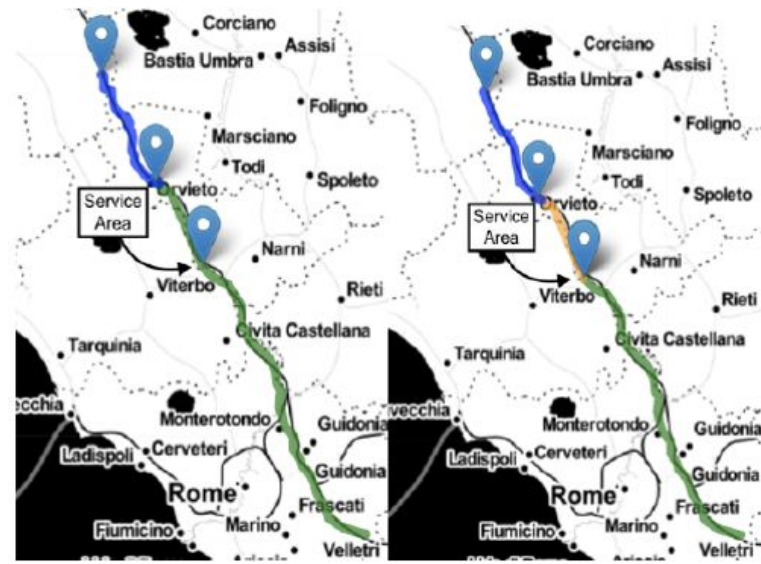
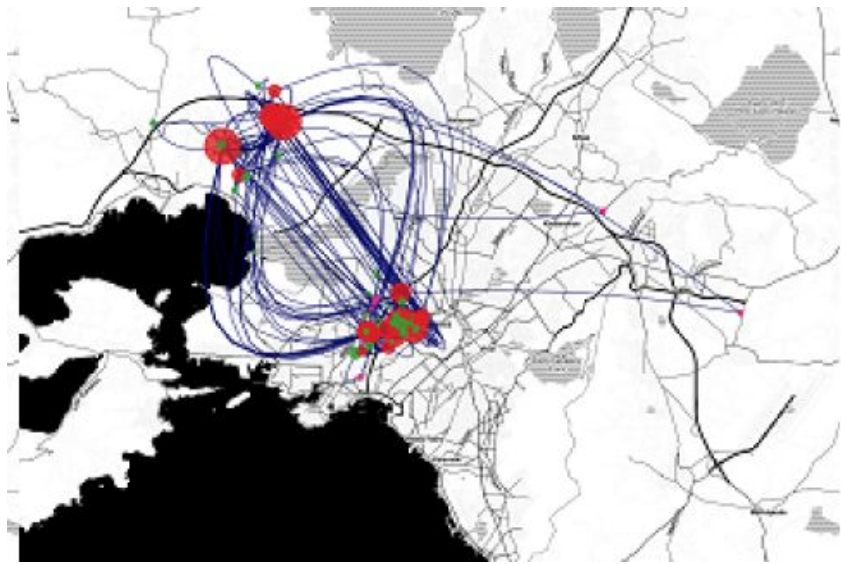
Healthcare Services: Royal Papworth (PAP) pilot

- Reconstructed (synthetic) patient data
- Objectives:
 - Obstructive Sleep Apnoea (OSA) service access,
 - OSA diagnostic test prevalence, etc.



Critical questions

- How much out of this data is **useful** for analytic purposes?
- What kind of analytics is **suitable** for this data?
- Are these datasets **large/complex** enough, for us to call them “Big Data”?



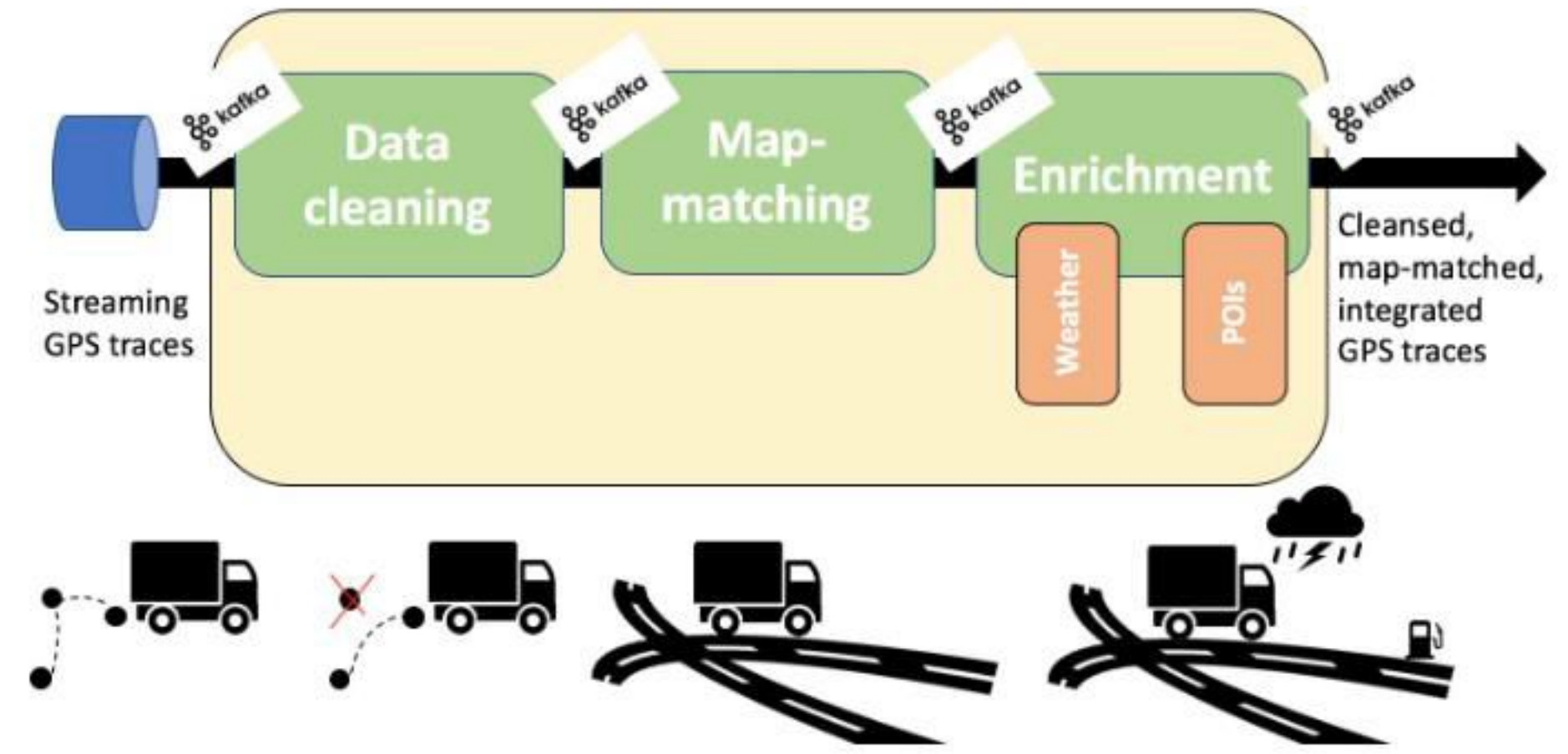
Track&Know architecture -1



Big Data Processing (BDP)

Data collection, curation, and preparation:

- cleansing (noise removal),
- enrichment (weather, POIs, etc.),
- summarization, etc.

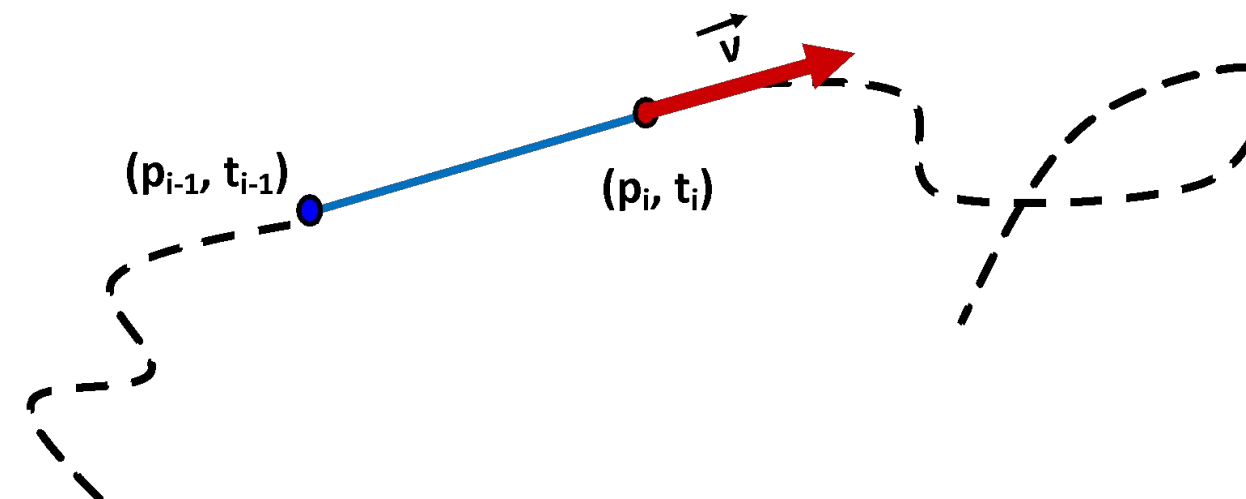
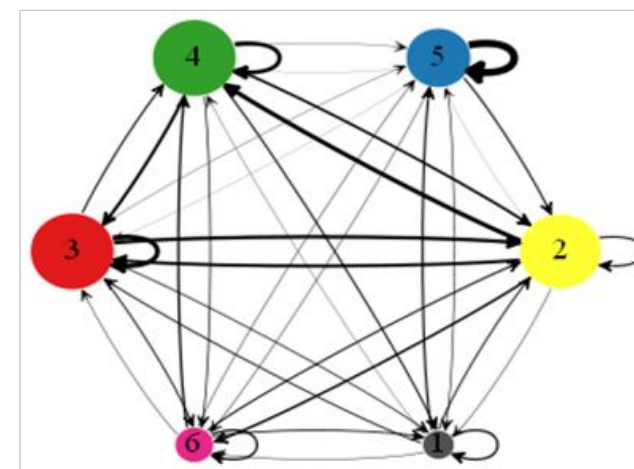
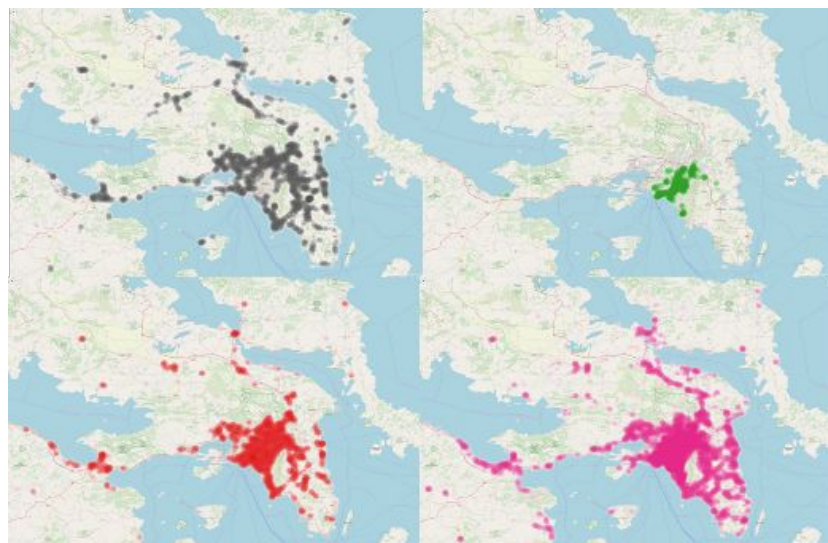


Track&Know architecture -2



Big Data Analytics (BDA)

- Extraction and semantic annotation of Individual Mobility Networks (IMNs)
 - mobility networks, driving 'signatures', etc.
- Future location & trajectory prediction
 - location forecasting using either clustering (network-aware) or deep learning (network-unaware) techniques

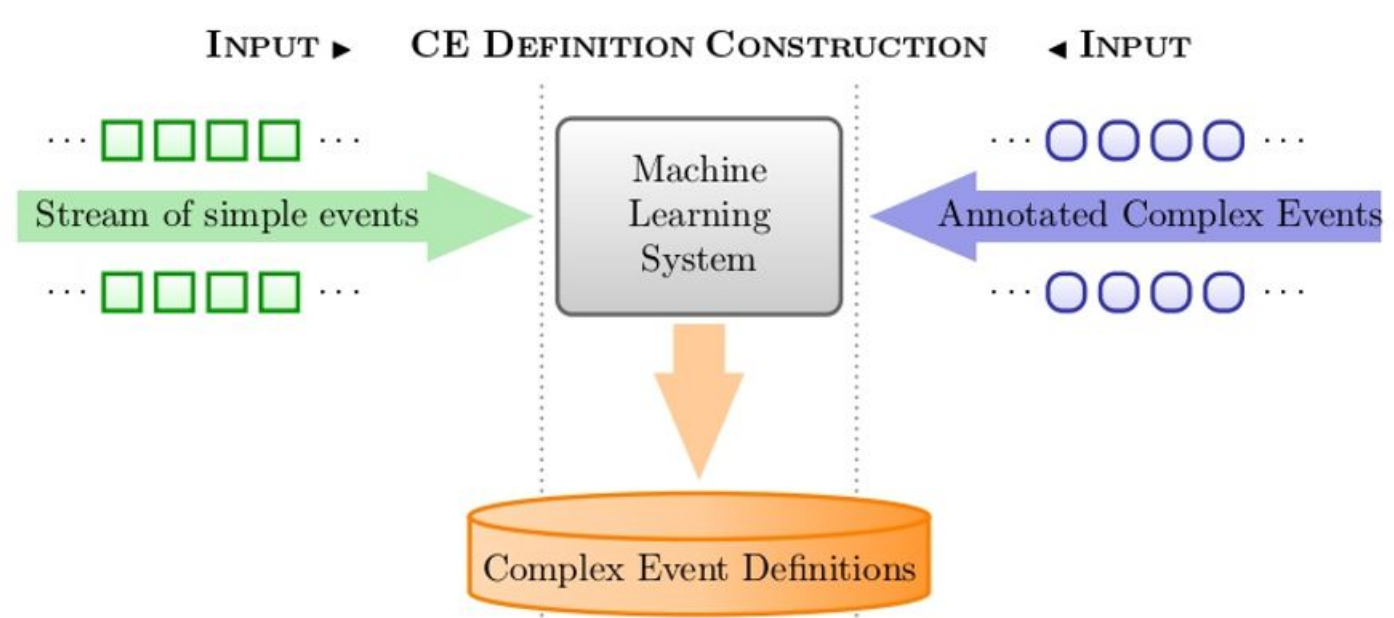
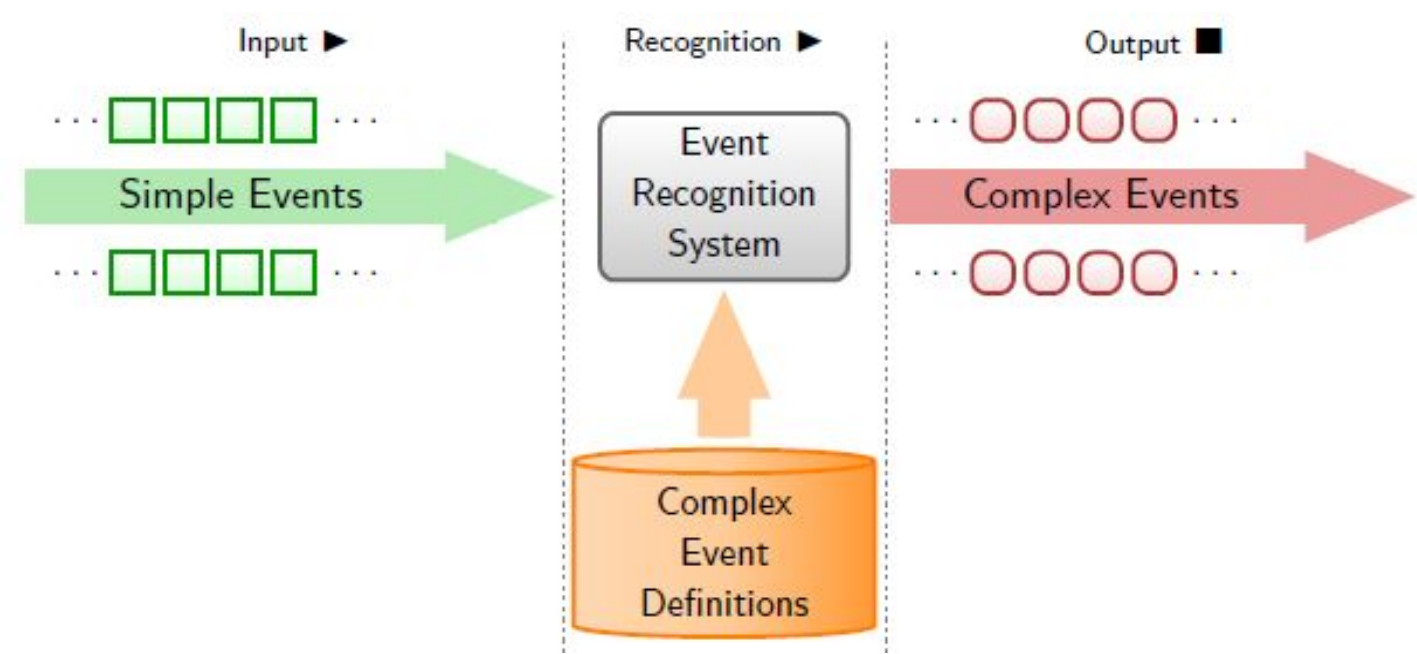


Track&Know architecture -3



Complex Event Recognition (CER)

- Support of various complex events (over-speeding, dangerous driving, etc.); delay-tolerant incremental CER
- Online learning of complex event patterns



Track&Know architecture -4

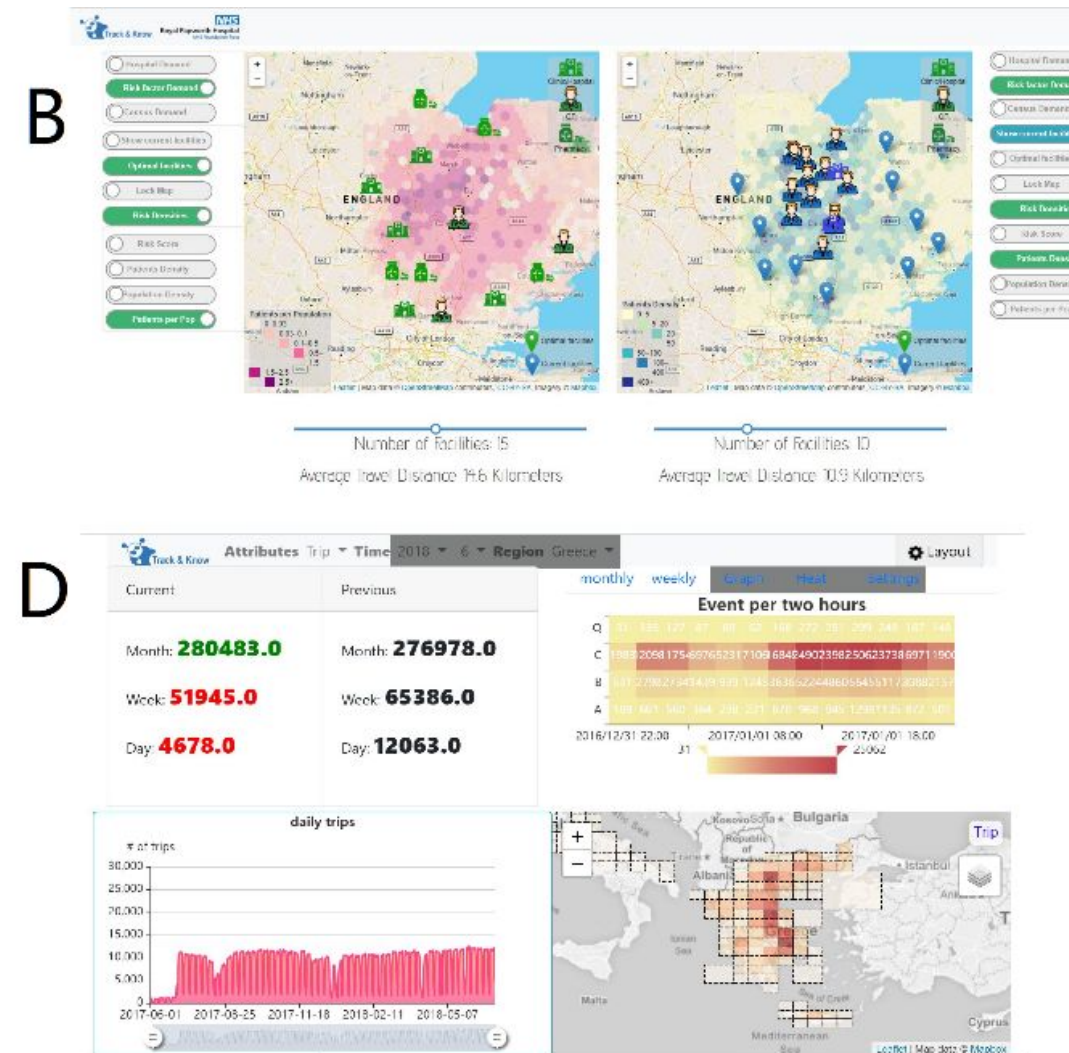
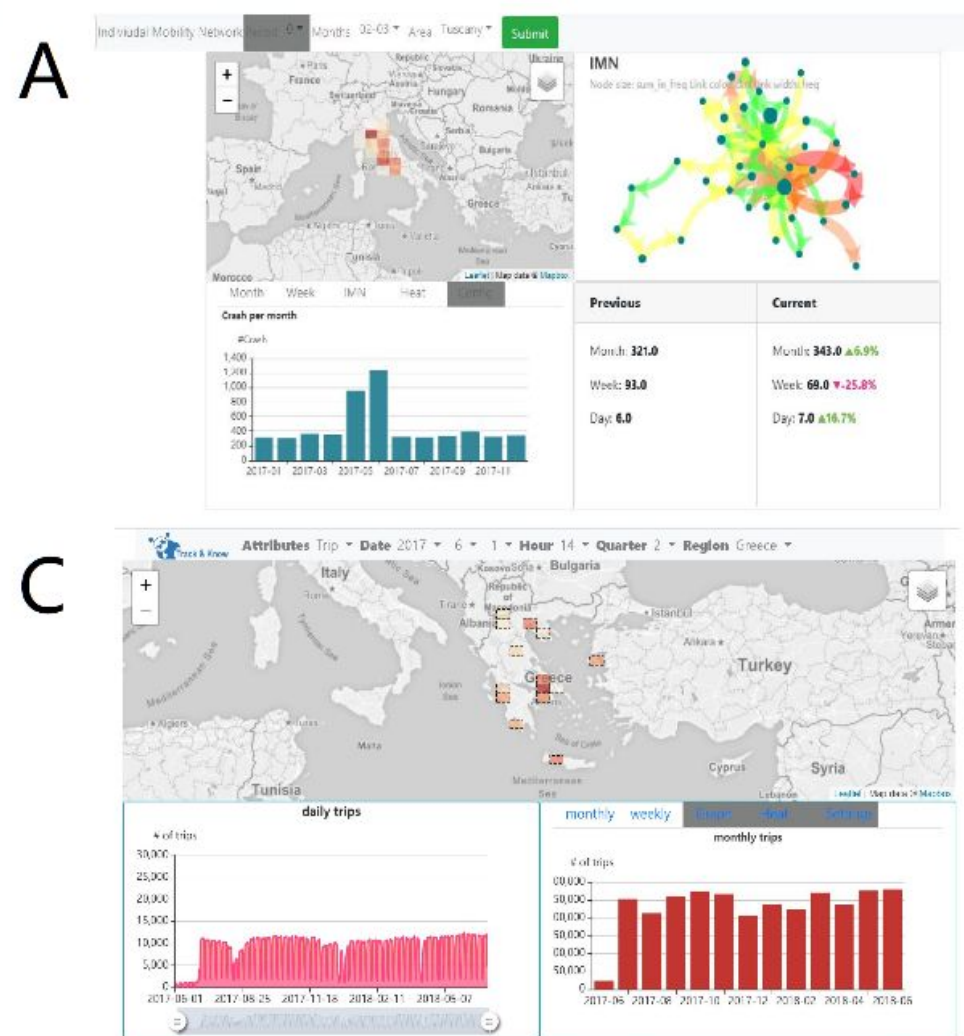


Visual Analytics (VA)

- Visualization of aggregated statistics as well as toolboxes' analytical results, e.g.

dashboards for

- A. SIS pilot;
- B. PAP pilot;
- C. VFI pilot (technicians);
- D. VFI pilot (managers)



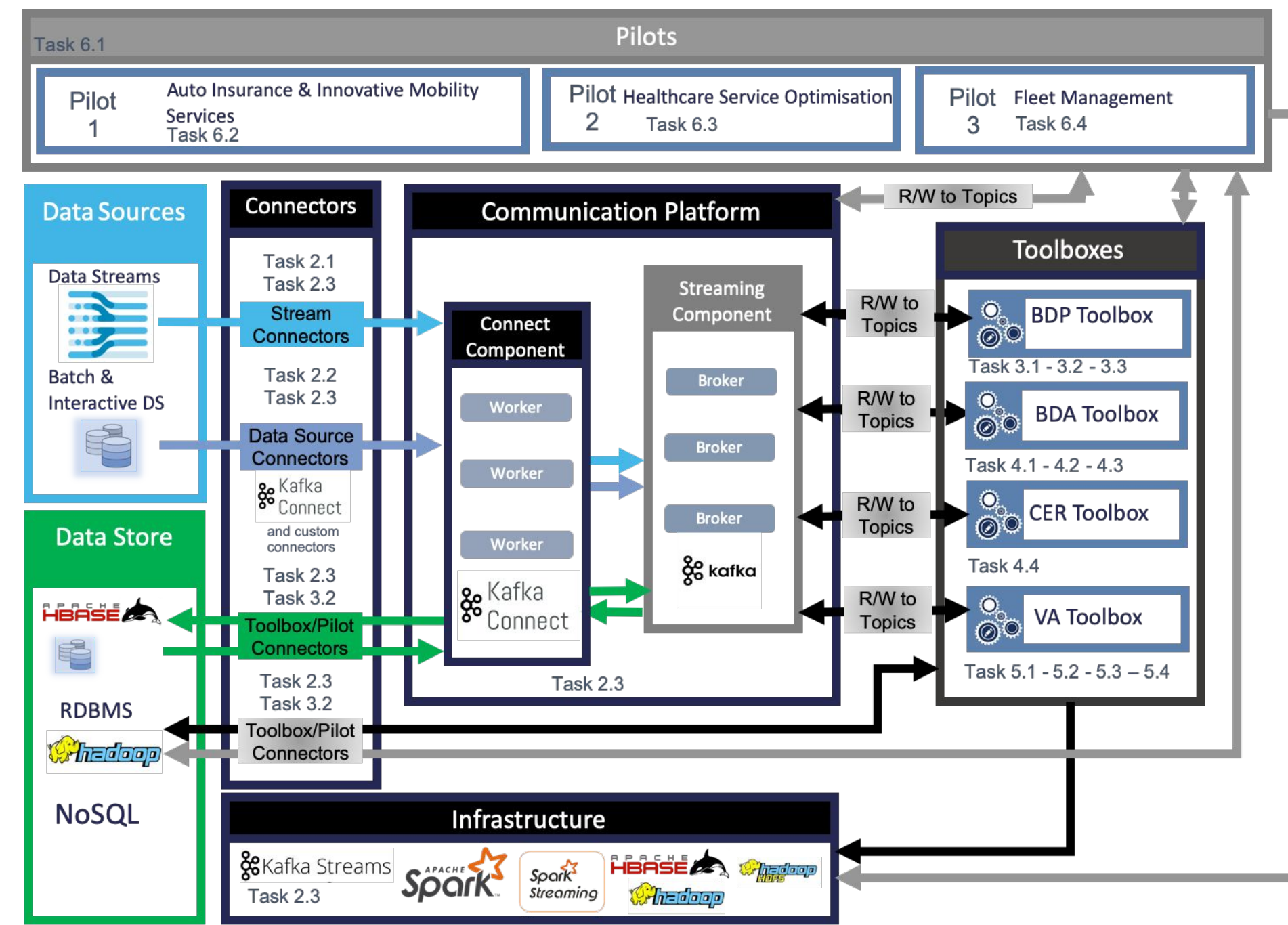
Track&Know architecture -5



all-in-one: Big Mobility Data Integration (BMDI)

• Data from streaming and archival sources as well as results from other toolboxes

- BDP
- BDA
- CER
- VA



Track&Know software



A number of 9 (so far) open s/w packages, publicly available through our Repository:

<https://trackandknowproject.eu> >
Online Observatory

| Toolbox | Component /partner(s) in charge | Open s/w |
|------------|--|------------------|
| BDP | GPS data enrichment (weather, POI, etc.) /UPRC | YES ¹ |
| BDA | Extraction of Individual Locations of Interest /CNR | YES ² |
| | Construction of Individual Mobility Networks /CNR | YES ³ |
| | Semantic annotation of Individual Mobility Networks /CNR | YES ⁴ |
| | Flow prediction /CNR | YES ⁵ |
| | Trip planning & simulation for electric vehicles /CNR | YES ⁶ |
| | Individual long-term event risk prediction /CNR | YES ⁷ |
| CER | Complex Event Recognition /NCSR | YES ⁸ |
| | Event Pattern Learning /NCSR | YES ⁹ |

¹ URL: <https://github.com/DataStories-UniPi/Trajectory-Weather-Integrator>

² URL: https://github.com/mirconanni/TandK_CNR/tree/master/Location_extraction

³ URL: https://github.com/mirconanni/TandK_CNR/tree/master/IMN_extraction

⁴ URL: https://github.com/mirconanni/TandK_CNR/tree/master/Annotated_IMN

⁵ URL: https://github.com/mirconanni/TandK_CNR/tree/master/Flow_prediction

⁶ URL: https://github.com/mirconanni/TandK_CNR/tree/master/EV_simulation

⁷ URL: https://github.com/mirconanni/TandK_CNR/tree/master/Crash_prediction

⁸ URL: <https://github.com/aartikis/RTEC>

⁹ URL: <https://github.com/nkatzz/ORL>



Summary



The field of **Mobility Data Analytics (MDA)** has many success stories to tell on:

- data processing (DBMS extensions, etc.)
- data analytics (cluster analysis, trajectory prediction, etc.)

The new MDA era that emerges is about

- **Context-enriched trajectories**: annotated information about when, where, what, how, why
- **Big mobility data**: voluminous, complex, heterogeneous information about movement of objects

*We are excited that **Track & Know** has contributed to this new era!*



Thank you!



Track & Know project

(Big Data for Mobility Tracking Knowledge Extraction in Urban Areas), H2020 grant #780754, 2018.01 – 2020.12

Data Science Lab. @ Univ. Piraeus

www.datastories.org

**DATA
STORIES**



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

UNIVERSITY OF PIRAEUS



EUROPEAN BIG DATA VALUE FORUM 2020