100% OPEN SOURCE

EU GDPR COMPLIANT

PEN Test Passed

Powered by FIWARE

SNAP4CITY on EUROPEAN OPEN SCIENCE CLOUD MARKETPLACE

Node-RED

SNAP4CITY

KM4CITY

SNAP4 Appliances and Dockers Installations

Be smart in a SNAP!

LIVING LAB

Data Analytics vs Smart Applications

Data Analytic

May 2022, Course
https://www.snap4city.org/577

SCALABLE SMART ANALYTIC APPLICATION BUILDER FOR SENTIENT CITIES

UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO | DISIT

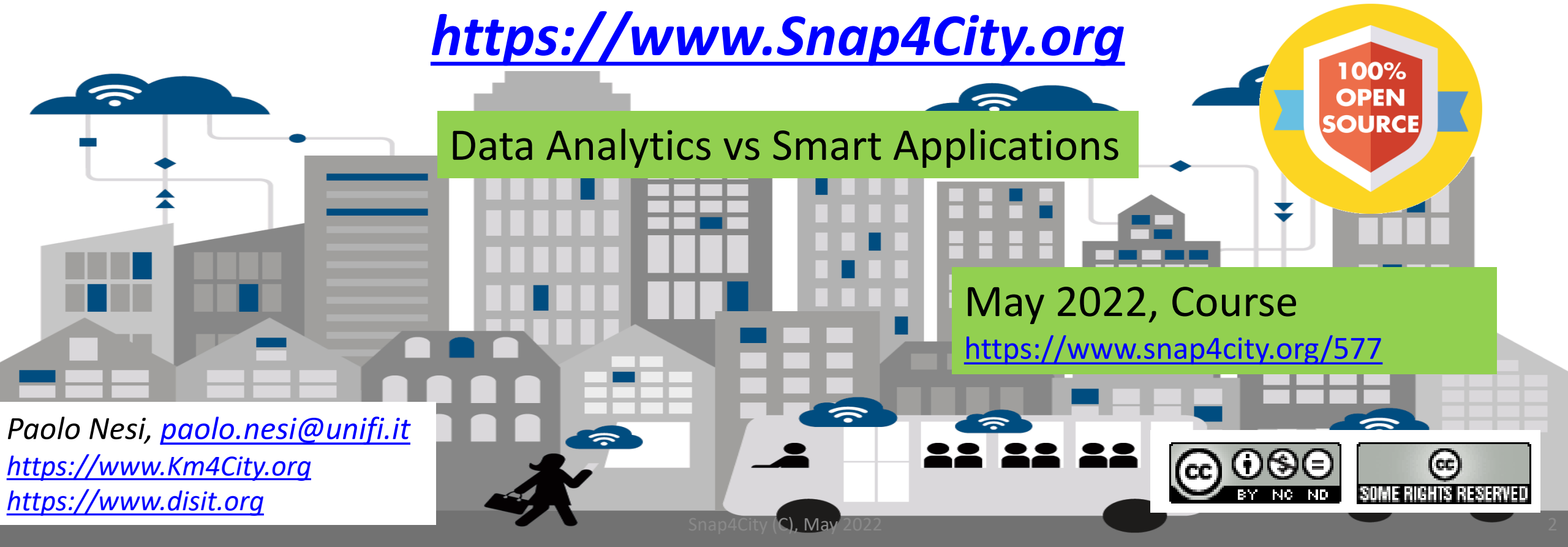
UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB

SNAP4CITY

Powered by KM4CITY

scalable Smart aNalytic APplication builder for sentient Cities: for Living Lab and co-working with Stakeholders

https://www.Snap4City.org



100% OPEN SOURCE

Data Analytics vs Smart Applications

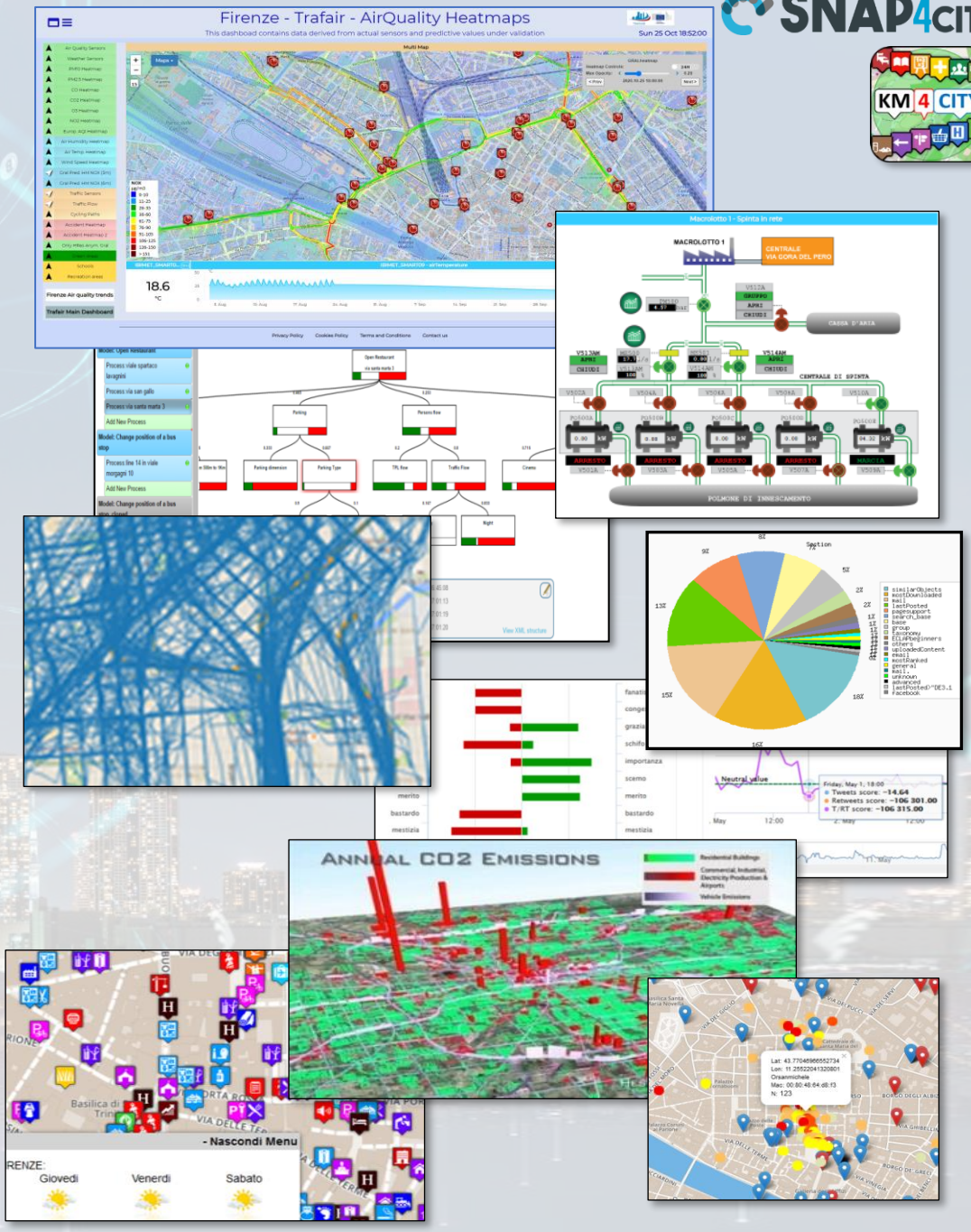
May 2022, Course
https://www.snap4city.org/577

Paolo Nesi, paolo.nesi@unifi.it
https://www.Km4City.org
https://www.disit.org

CC BY NC ND SOME RIGHTS RESERVED

Data Driven Decision Support

- Decision Support system
- Assessment / Strategies
- Data Rendering, visual analytics
- Data Analytics
- Data aggregation, Storage, indexing
- Data Ingestion

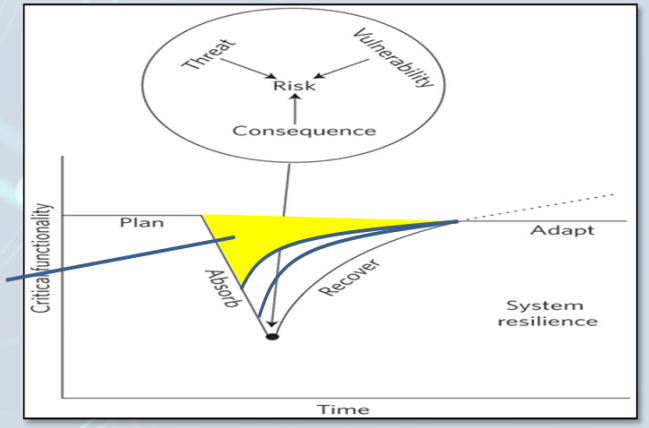


Objectives

- Decision support systems
- Improvement of life quality
- Sustainable Solutions
- Reduction of costs
- Resilience

Prepare
Absorb
Recover
Adapt

damage



Data Analytics

System Modeling and Simulation

Knowledge Models

Scenarious

Short Term Predictions, Anomaly Detections

Long Term Predictions, Anomaly Analysis

Recommendations & Prescriptions

What-if Analysis

Early Warning

Disaster Recovery

Strategies and Plans

Partial graph

Quality of Life

Tools for rapid implementation of sustainable Smart Solutions and Decision Support Systems

www.snap4city.org

100% OPEN SOURCE

Powered by FIWARE

FREE TRIAL

PEN Test Passed

EU GDPR COMPLIANT

SNAP4 Appliances and Dockers Installations

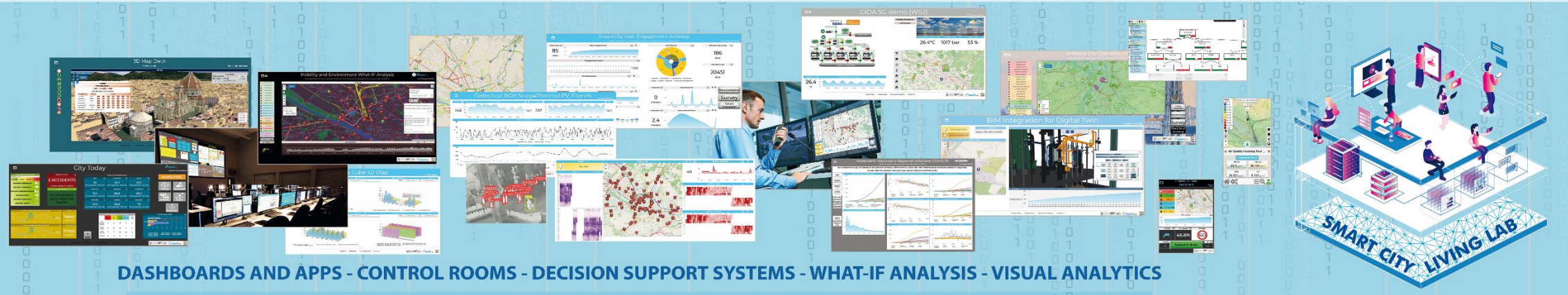
EUROPEAN OPEN SCIENCE CLOUD

Node-RED

JS Foundation

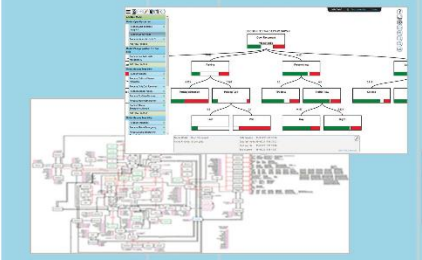
E015 digital ecosystem

NVIDIA

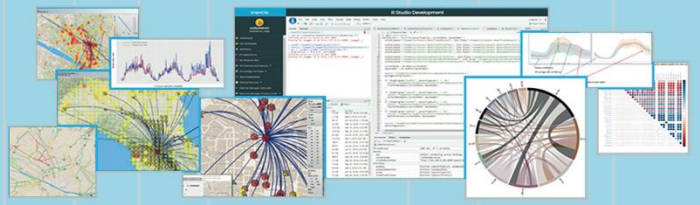


DASHBOARDS AND APPS - CONTROL ROOMS - DECISION SUPPORT SYSTEMS - WHAT-IF ANALYSIS - VISUAL ANALYTICS

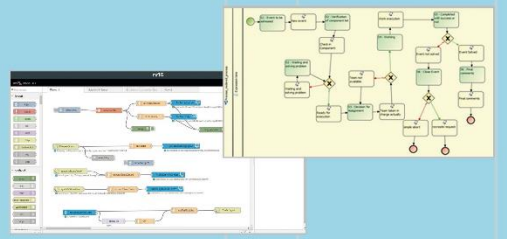
PREDICTION - ANOMALY DETECTION - ENVIRONMENTAL MODEL - 3D MODEL
KPI - SIMULATION - EARLY WARNING - SYNOPTIC - DIGITAL TWIN - VIRTUAL REALITY



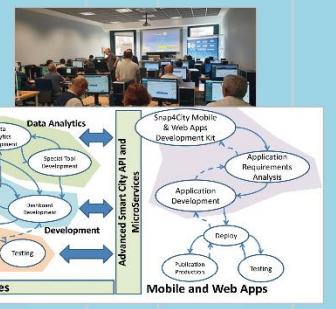
EXPERT SYSTEM
KNOWLEDGE BASE
STORAGE



BIG DATA ANALYTICS
EXPLAINABLE ARTIFICIAL INTELLIGENCE
BUSINESS INTELLIGENCE
MACHINE LEARNING



DATA FLOWS, DATA DRIVEN
WORKFLOWS, MICROSERVICES
PARALLEL DISTRIBUTED PROCESSING



METHODOLOGIES
LIVING LABS
COURSES AND COMMUNITY
DEVELOPMENT TOOLS

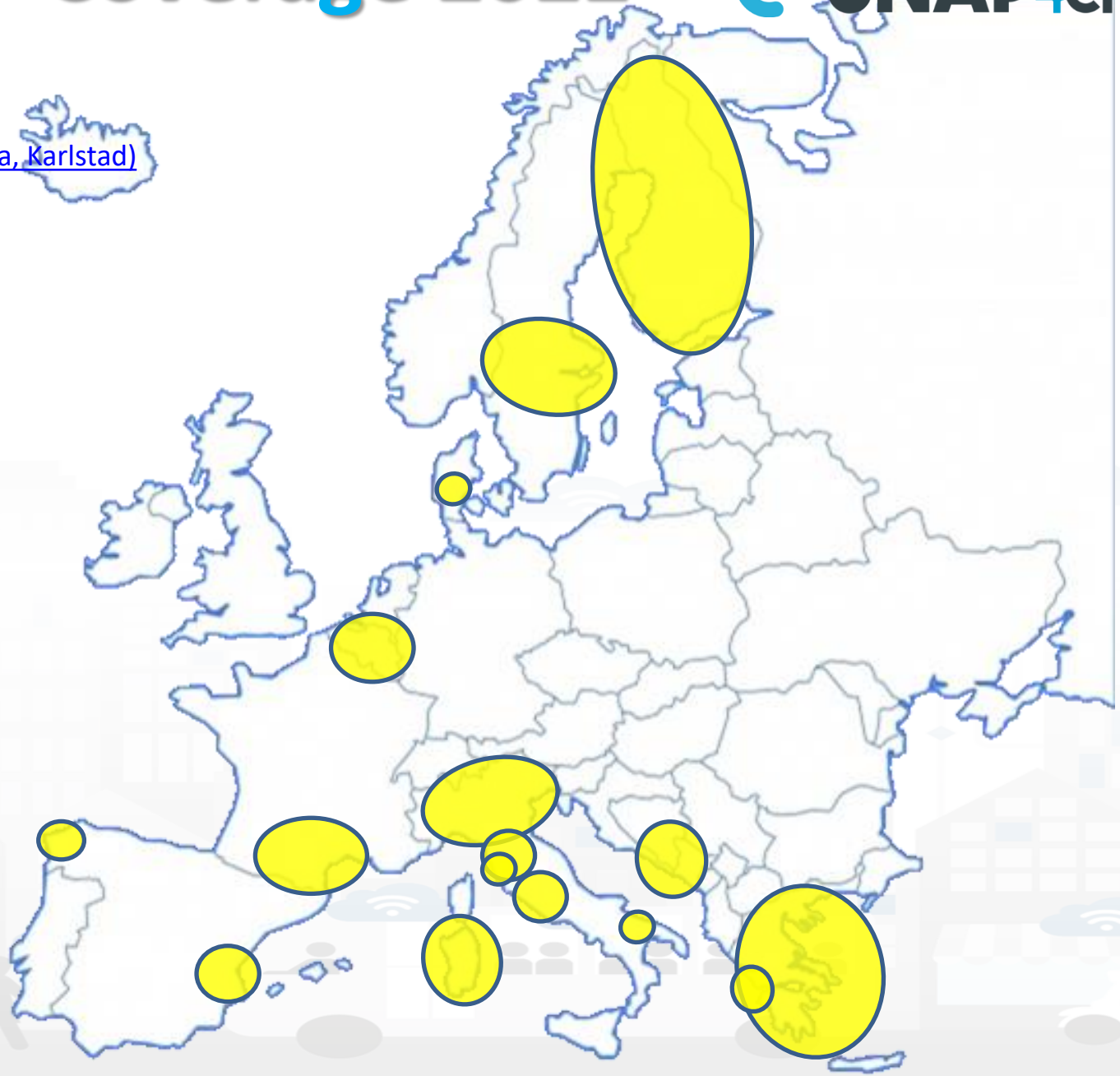
Snap4City/Industry structure

- The **Snap4xxxx** solution is released in Open Source, VM and Docker with fully support of MultiTenant/multiple-Organizations
  - Each Organization may be configured for a separate environment with a set of Maps, Menus, Users, Data, Dashboards, IOT Apps, MicroApplications, Custom Widgets, Models, resources, open data, etc.
- **Https://www.Snap4City.ORG** is the main instance of Snap4xxxx solution managed by DISIT Lab. The main documentation is located and updated on Snap4City.org, GitHUB, dockerHub and Node-Red Library. Snap4City.org is where the last tools are tested and news published.
  - Organizations on Snap4City.org have been created with contracts as for *Platform as a Service*, for testing and for providing *SmartCity as a Service* as well as *Industry 4.0 as a Service*

Coverage 2022

Main Organizations/areas

- Antwerp area (Be)
- Bologna (I)
- Capelon (Sweden: Västerås, Eskilstuna, Karlstad)
- DISIT demo (multiple)
- Dubrovnik, Croatia
- Firenze area (I)
- Garda Lake area (I)
- Greece (Gr)
- Helsinki area (Fin)
- Livorno area (I)
- Lonato del Garda (I)
- Modena (I)
- Mostar, Bosnia-Herzegovina
- Oslo & Padova (Impetus)
- Pisa area (I)
- Pistoia (I)
- Pont du Gard, Occitanie (Fr)
- Prato (I)
- Roma (I)
- Santiago de Compostela (S)
- Sardegna Region (I)
- Siena (I)
- SmartBed (multiple)
- Toscana Region (I), SM
- Valencia (S)
- Venezia area (I)
- WestGreece area (Gr)



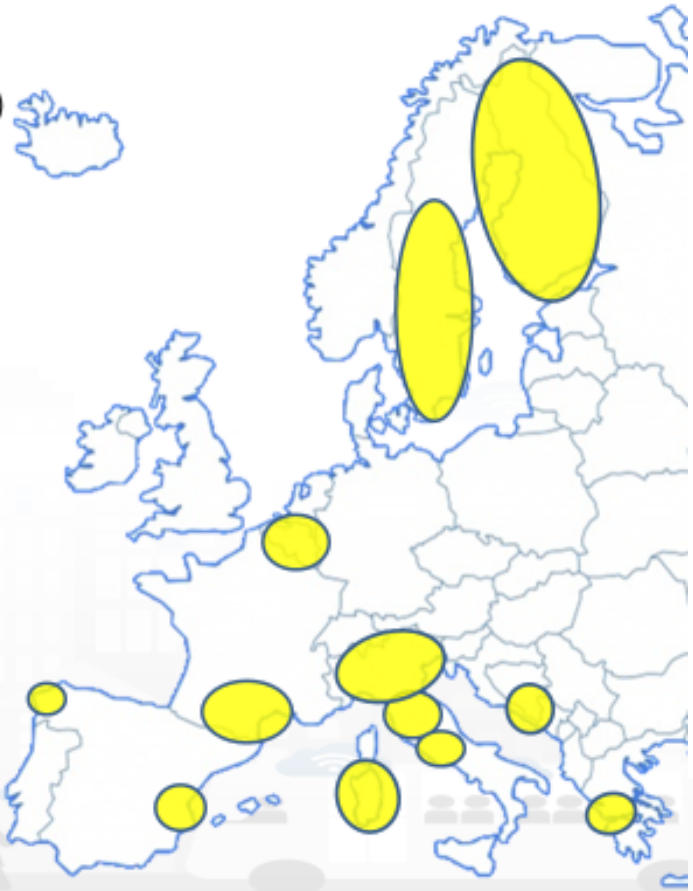
Snap4City/Industry Community

- Most of Organizations on Snap4City.org also correspond to companies or institutions that have an installation of Snap4City tools on their Premise,
  - such as: Pisa, SmartGarda Lake, Snap4, ALTAIR, etc.
- This double way allows them to:
  - test the news,
  - share experiences with other groups,
  - get visibility,
  - work in the collaborative environment, and
  - be better supported by Snap4City.org and DISIT Lab personnel.
- Each instance of Snap4xxxx solution ***can decide to join the federation*** of SmartCity API to exploit shared data.
  - This allows to exploit regional data for city installations applications (web, mobile, dashboards, etc.) without reloading them for example.

Main Organizations/areas

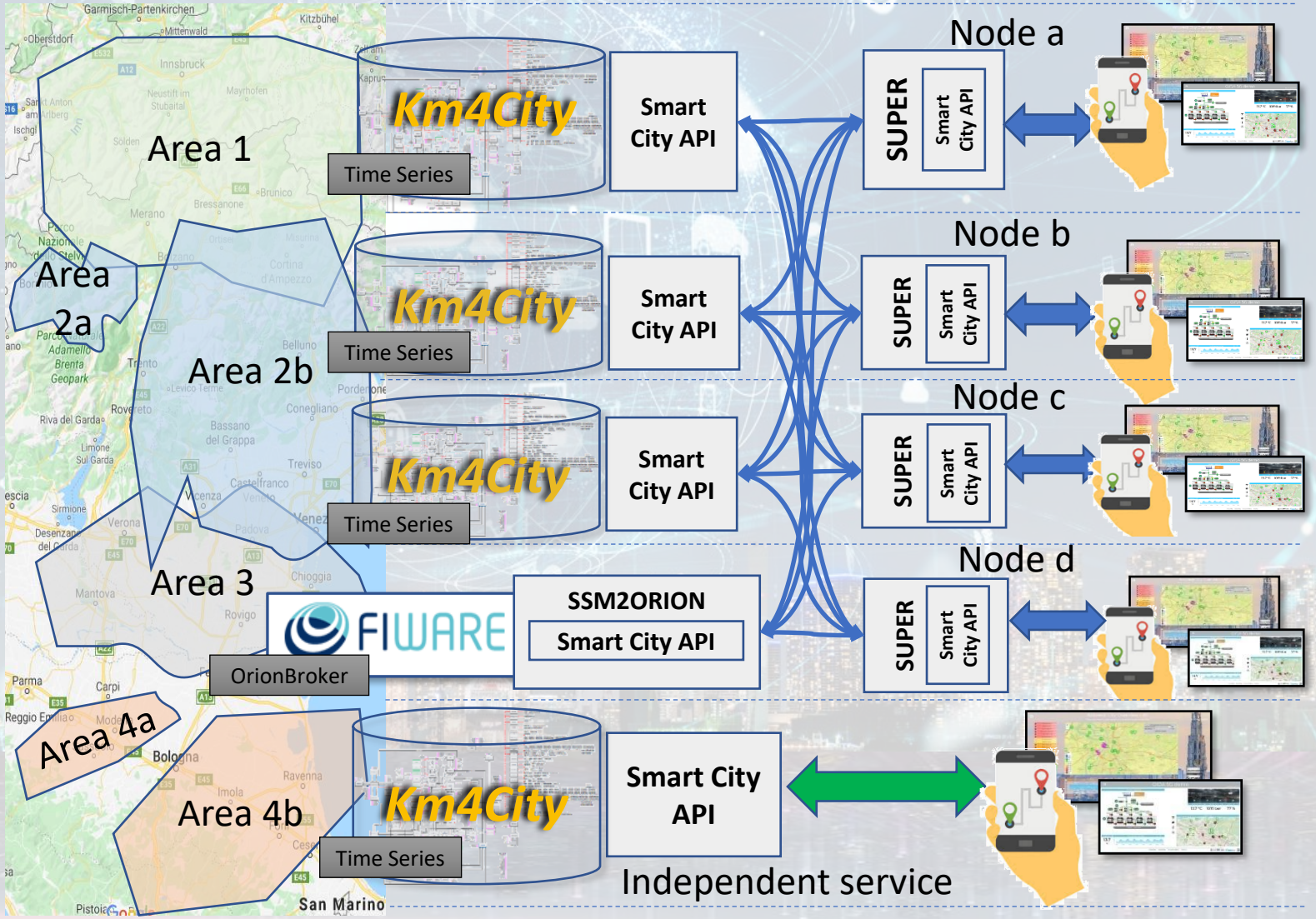
- Antwerp area (Be)
- Capelon (Sweden: Västerås, Eskilstuna, Karlstad)
- DISIT demo (multiple)
- Dubrovnik, Croatia
- Firenze area (I)
- Garda Lake area (I)
- Helsinki area (Fin)
- Livorno area (I)
- Lonato del Garda (I)
- Modena (I)
- Mostar, Bosnia-Herzegovina
- Pisa area (I)
- Pont du Gard, Occitanie (Fr)
- Roma (I)
- Santiago de Compostela (S)
- Sardegna Region (I)
- SmartBed (multiple)
- Toscana Region (I), SM
- Valencia (S)
- Venezia area (I)
- WestGreece area (Gr)

Snap4City (C), October 2020



Federation of Smart City Services

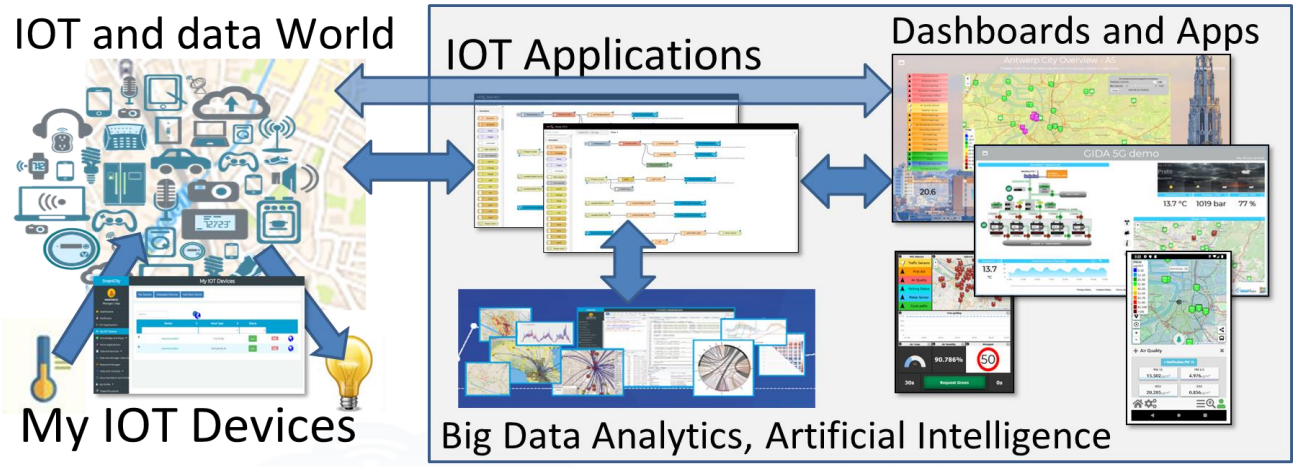
Knowledge Base
Semantic Reasoners



- Km4City **Semantic Reasoner**
- **ServiceMap** interoperability
- **Seamless for multiple Mobile Apps**
- **Smart City API**
- **Super**:
  - distributed access and sharing services
  - Each city control its own data
  - Final user can pass from one city / area to another in seamless manner: without changing the mobile Apps

Free Trial

- Register on WWW.snap4city.org
  - Subscribe on **DISIT Organization**
- **You can:**
  - Access on basic Tools
  - Access to a large volume of Data
  - Create Dashboards
  - Create IOT Applications
  - Connect your IOT Devices
  - Exploit Tutorials and Demonstrations



IF you need to go more in deep you can ask us to pass at the next Role becoming full AreaManager with full rights of development, also for Data Analytics, machine learning, etc.

https://www.snap4city.org/577

On Line Training Material (free of charge)

	1st part (*)	2nd part (*)	3rd part (*)	4th part (*)	5th part (*)	6th part (*)	7th part (*)
what	General	Dashboards	IOT App, IOT Network	Data Analytics	Data Ingestion processes	System and Deploy Install	Smart City API: Web & Mob. App
PDF							
Inter active							
Video1							
Video2							
Video3							
Video4				none		none	none
duration	2:55	3:16	3:41	2:00	2:48	2:35	1:47

General Overview of the full Course 2021

- ***1st part:*** *General Overview*
- ***2nd part:*** *Dashboards Creation and Management*
- ***3rd part:*** *IOT Applications development, IOT Devices, IOT Networks*
- ***4th part:*** *Data Analytics, in R Studio, in Python, how to Exploit and Manage Data Analytics in IOT Applications*
- ***5th part:*** *Data Ingestion, Data Warehouse, Data Gate, IOT Device Data ingestion, IOT App for Data Ingestion, Interoperability, etc.*
- ***6th part:*** *Snap4City Development, Extension, Administration, and Installation*
- ***7th part****: Smart city API (internal and external) Web and Mobile App development tool kit*

A number of the training sections include exercitations

Updated versions on: https://www.snap4city.org/577

See also courses in ITALIANO: https://www.snap4city.org/485

Data Analytics

- **Data Analytics: Examples from Snap4City**
  - Smart parking: Predictions
  - Smart Bike Sharing
  - User Behavior Analysis, via Wi-Fi, OD Matrices
  - User Behavior Analysis via trajectories
  - Recognition of Used Transportation means
  - Traffic Flow Predictions
  - Traffic Flow in/out of the city
  - Traffic Flow Reconstruction, from Traffic Sensors Data
  - Covid-19 vs other data: traffic and environmental
  - Quality of Public Transport Service
  - Origin Destination Matrices (from: Wi-Fi, Mobile Apps, etc.)
  - Modal and Multimodal Routing for Navigation and Travel Planning
  - Etc.

- **Data Analytics for City/Area Management**
  - Predicting Land Sliding
  - Estimating CO2 from Traffic Flow Data
  - DORAM: Demand of Mobility vs Offer of Transportation
  - Environmental Data Analysis and Predictions, early Warning
  - Environmental Sensors Hosted by City Users
  - Long Term Prediction of Annual Mean of NO2 index of EC
  - Anomaly Detection
  - What-IF Analysis….
  - 15MinCityIndex Assessment of Cities
  - Predictive Maintenance on Plant
  - Fashion Retail shop recommendations

- **Engaging City users Towards a Virtuous behavior**
- **Decision Support Systems, Smart DS and Resilience DS**
- **Twitter Vigilance: Social Media Analysis: Early Warning, Predictions**
- **HOW TO: Data Analytics: Enforcing and Exploiting**
  - Real Time Data Analytics: using R Studio Exploitation in IOT Applications

4th part Agenda

GO • **Data Analytics: Examples from Snap4City**

- GO – Smart parking: Predictions
- GO – Smart Bike Sharing
- GO – User Behavior Analysis, via Wi-Fi, OD, trajectories
- GO – Recognition of Used Transportation means
- GO – Traffic Flow Predictions
- GO – Traffic Flow Reconstruction, from Traffic Sensors Data
- GO – Covid-19 vs other data: traffic and environmental
- GO – Quality of Public Transport Service
- GO – Origin Destination Matrices from: Wi-Fi, Mobile Apps, etc.
- GO – Demand of Mobility vs Offer of Transportation
- GO – Modal and Multimodal Routing for Navigation and Travel Planning

• **Data Analytics for Management**

- GO – Environmental Data Analysis and Predictions, early Warning
- GO – Prediction of Air Quality Conditions
- GO – Anomaly Detection
- GO – What-IF Analysis….
- GO – 15MinCityIndex Assessment of Cities
- GO – Predictive Maintenance on Plant
- GO – Fashion Retail shop recommendations

GO • **Data Analytics: Enforcing and Exploiting**

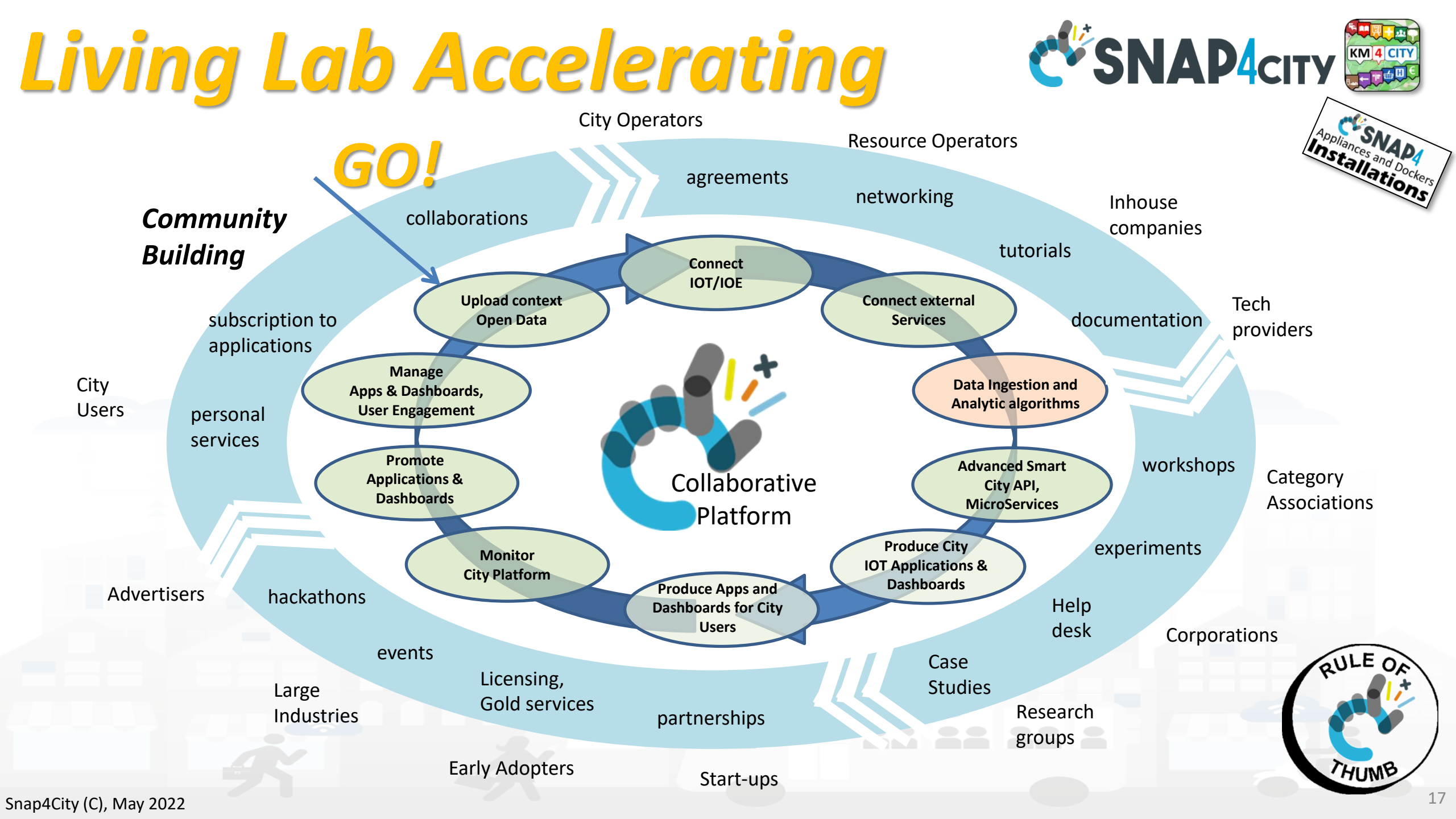
- Real Time Data Analytics: using R Studio Exploitation in IOT Applications

GO • **Engaging City users Towards a Virtuous behavior**

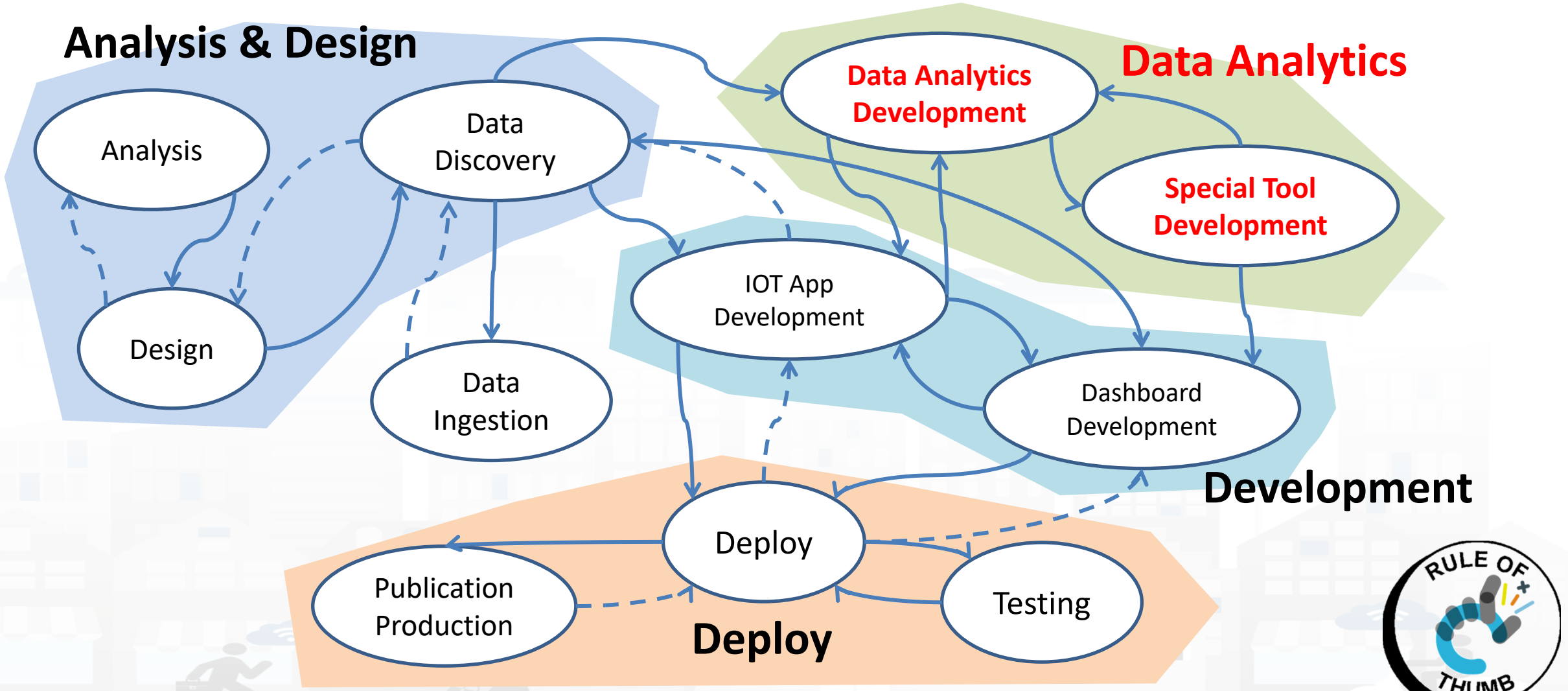
GO • **Decision Support Systems, Smart DS and Resilience DS**

GO • **Twitter Vigilance: Social Media Analysis: Early Warning, Predictions**

Living Lab Accelerating



Development Life Cycle Smart City Services



Levels of Difficulty

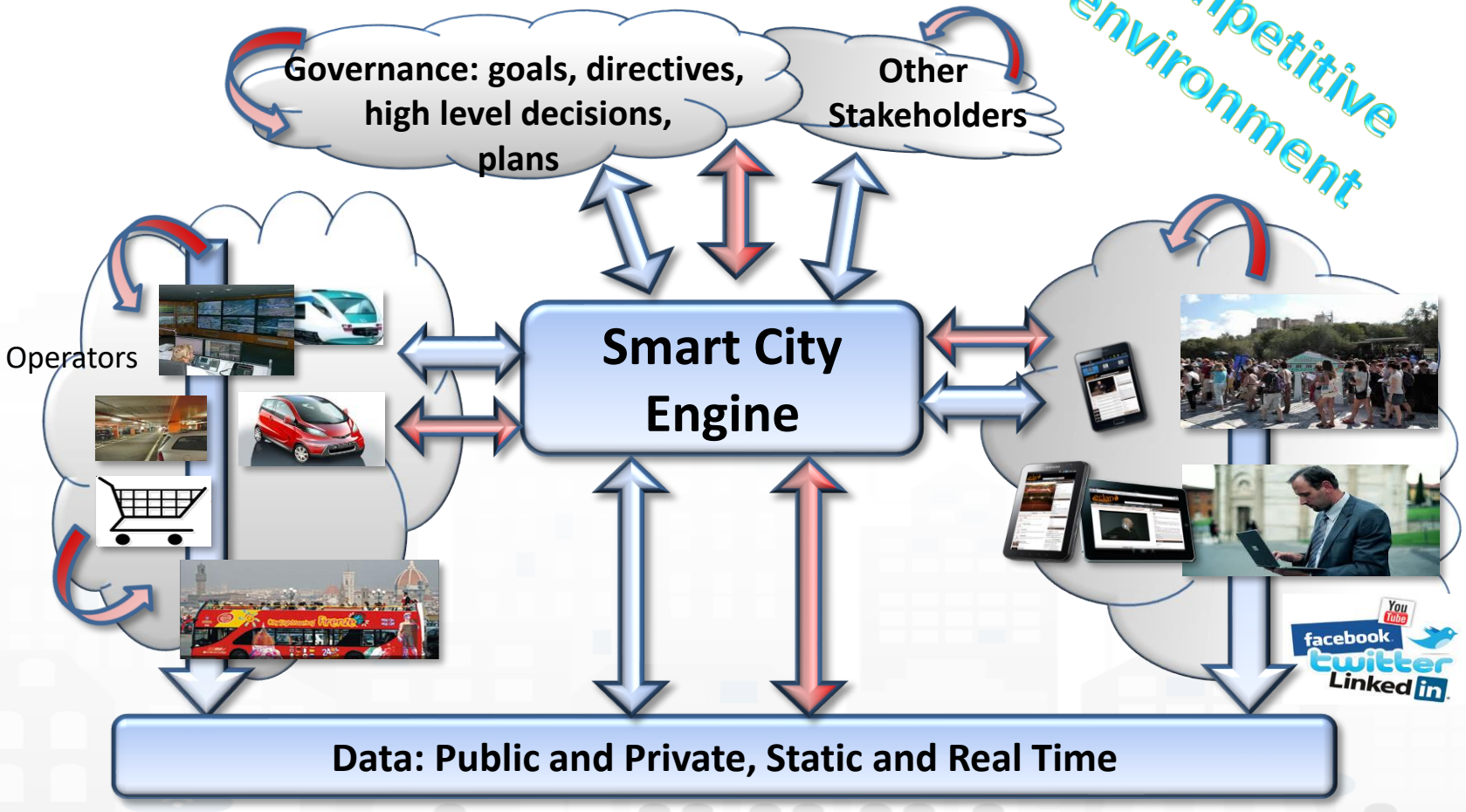
- Easy. non programmer level
- Moderate. Some JavaScript rudiment coding
- Good. JavaScript programming
- Golden. Programming in R Studio
- Professional. Exploiting Smart City API
- Excellent. Developing Full IOT Applications, Dashboard and Mobile Apps

TOP

Exploiting data produced by Snap4City Solutions

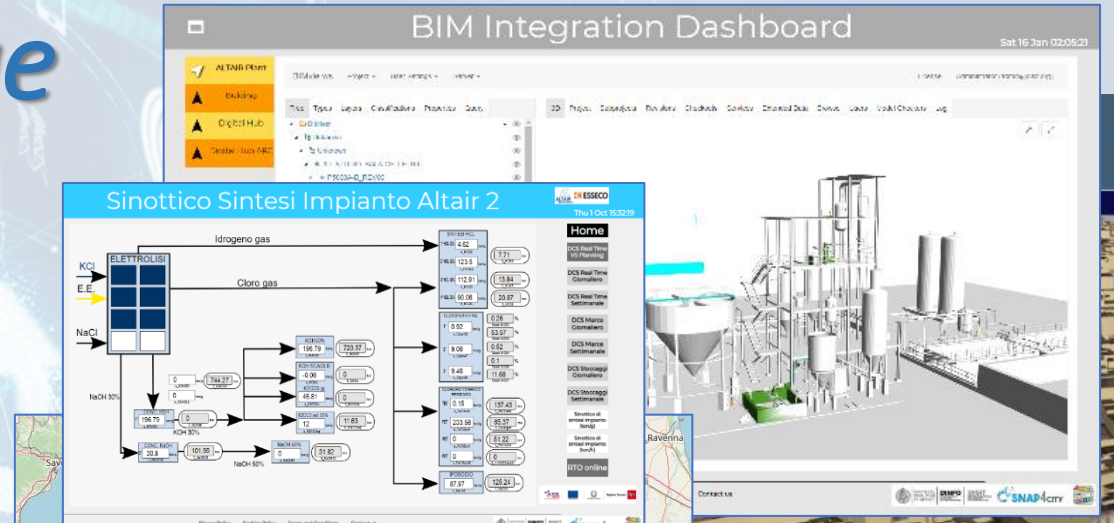
From Strategies to (re-)Actions

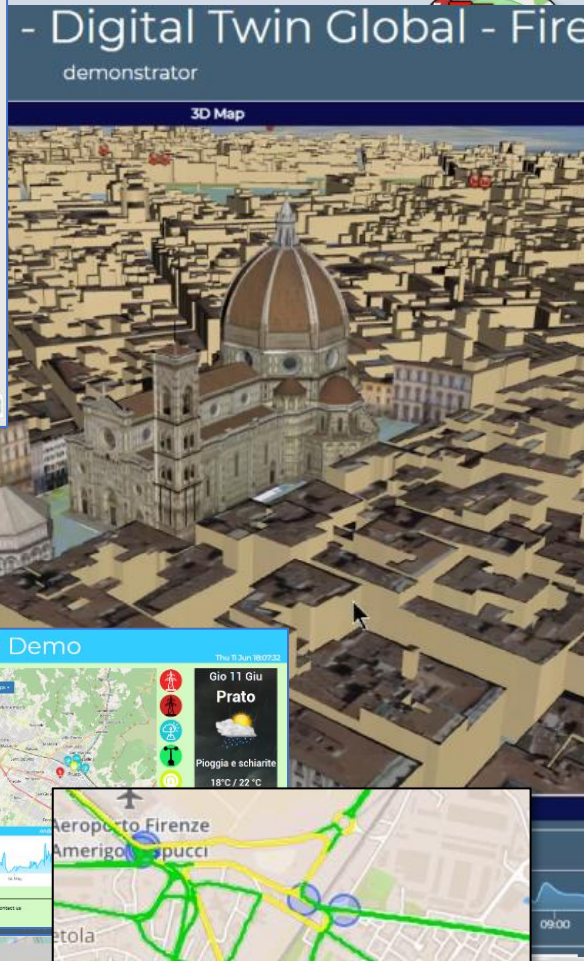
- Analyze
- Alerting, Early Warning
- Support Decision makers
- Plans
- Prescriptions
- Inform
- Suggest
- Engage
- Research

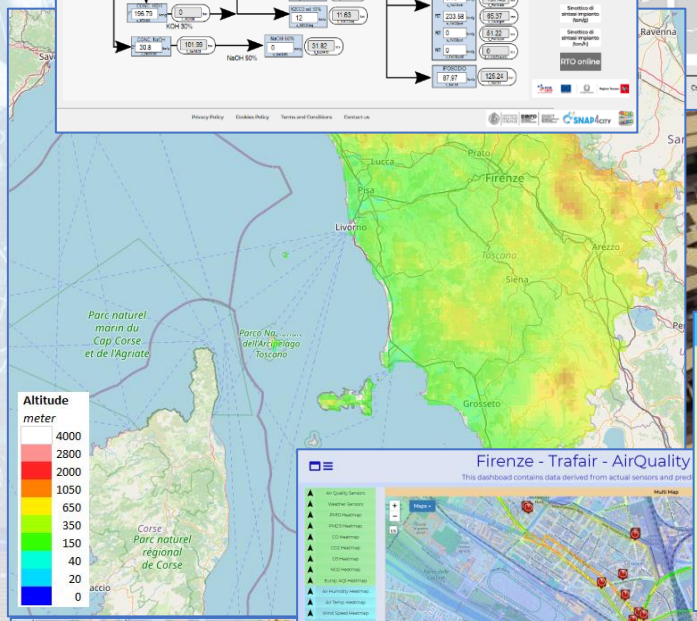


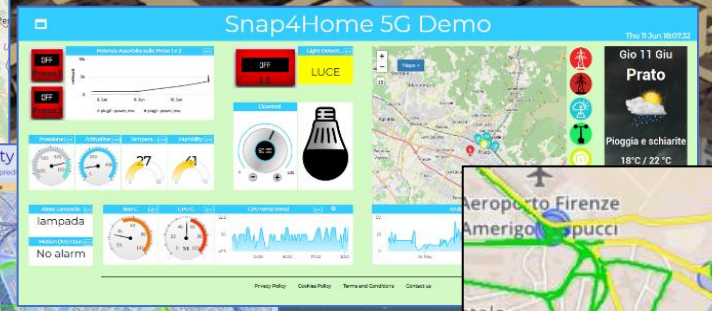
Data Type Coverage

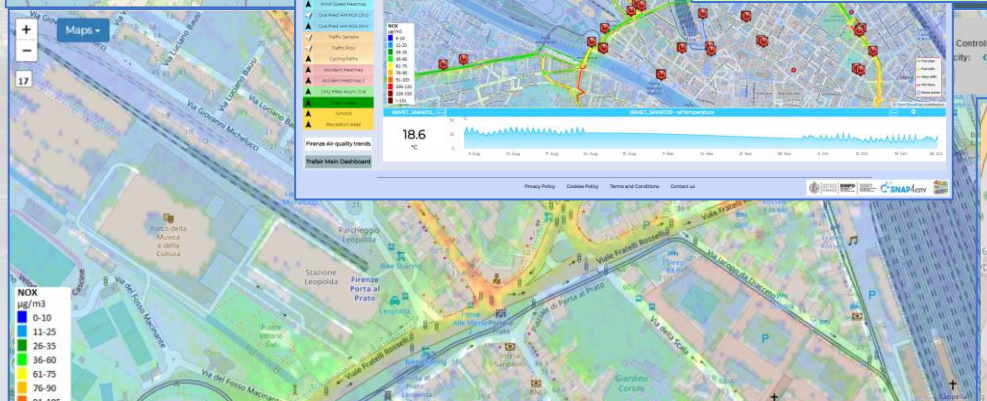
- POI, IOT, shapes,..
- maps, orthomaps, GTFS, GIS WFS/WMS, GeoTiff, ..
- calibrated heatmaps, ..
- traffic flow, typical trends, ..
- trajectories, events, ..
- 3D, BIM, Workflow, ..
- Dynamic icons/pins, ..
- OD Matrices, scenarios, ..
- prediction models, ...
- decision support, ...
- Synoptics, animations, ..
- social media, Routing, ..
- Satellite data, ..
- KPI, personal KPI,..
- etc.

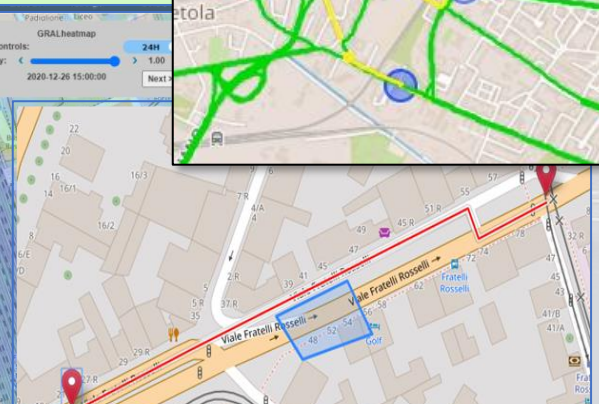






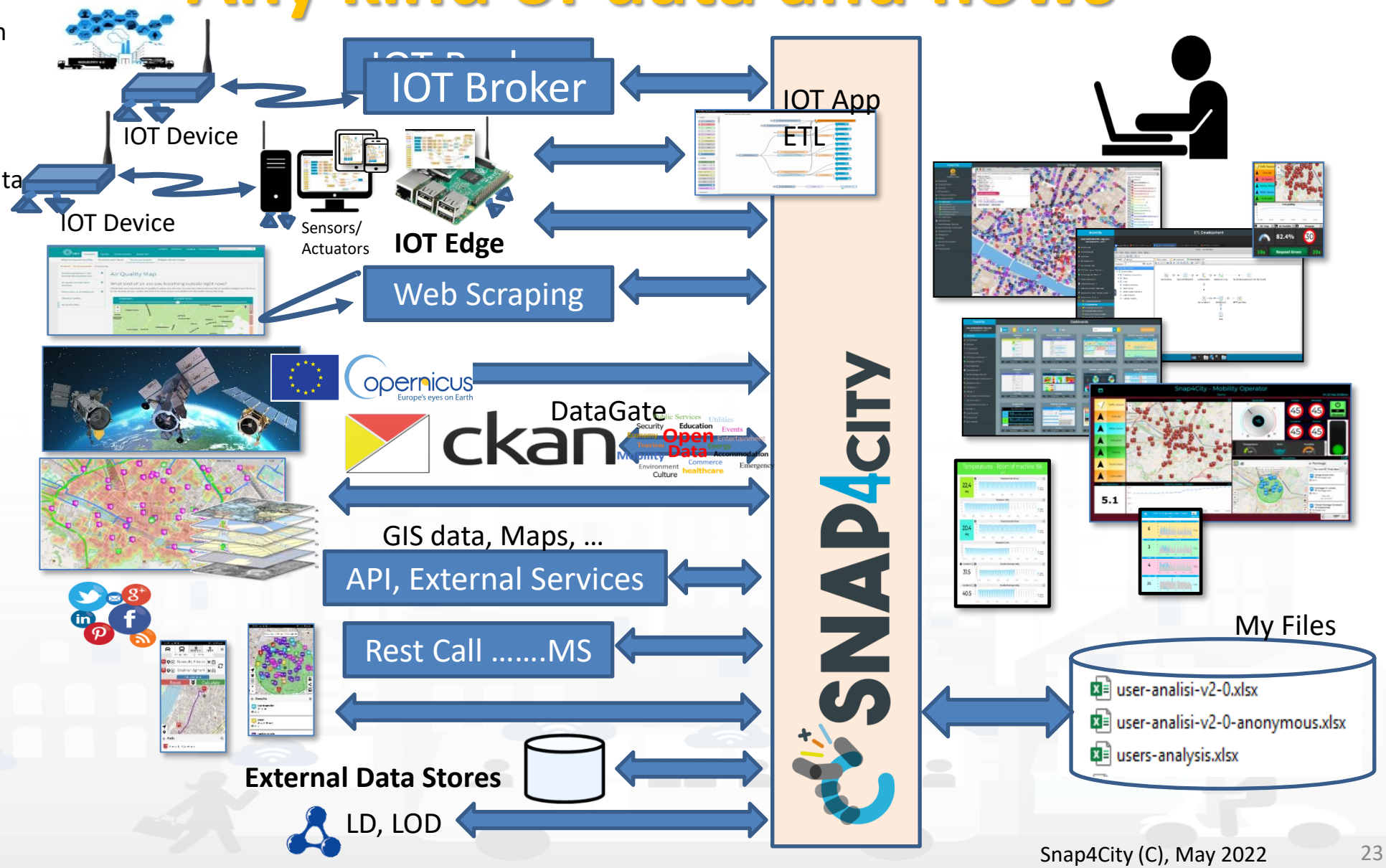






Any kind of data and flows

- **Open Data:**
  - Data gate, federation of Open Data Portals
  - IOT App, ETL proc(PULL)
- **IOT Networks:**
  - IOT Application processes, data driven or PULL
  - IOT Brokers (Push) → IOT Shadow
- **Web Pages:**
  - Web scraping, crawling processes
- **Satellite data**
- **Social media: Twitter, Facebook,..**
  - Twitter Vigilance, IOT App
- **Mobile Apps**
  - Smart City API
- **Files upload: CSV, Excel, etc.**
  - IOT Applications, ETL
- **REST API, WS, FTP, LD, LOD, etc.**
  - IOT Applications, ETL
- **Data base accesses**
  - GIS: WFS, WMS
  - ETL, IOT Application



What we do

- **Data modelling and management**
  - Data discovery and data surrogates/replacements
  - Big data management and provider
  - GIS data management
  - satellite Copernicus data processing for smart city and industry
  - IoT interoperability, edge, fog and cloud
  - Data interoperability, data aggregation and semantic processing
- **AI, Data Analytic, Visual Analytic**
  - AI for: **predictions**, **anomaly detection**, **clustering**, suggestions, simulation, fluid dynamics, classification, recognition, ..
  - **XAI, Explained AI, Trustworthy AI**
  - cognitive reasoning: ontology development, semantic computing
  - modelling and computing KPI
  - What-if analysis by mixing simulation, AI, statistics, semantics
- **Different contexts:**
  - industry, smart city, human behaviour, mobility, environment, terrain sliding
  - E.g. predictions pollutants/aerosol, CO2, NO2, GHG; traffic, parking, etc.

to cope with

- any data, format
- any channel, protocol
- any AI/ML
- any place
- online development
- multi-tenant
- Secure, PENTest
- GDPR, privacy
- → ***low costs***
- → ***easy to evolve***

Big Data Analytics + Artificial Intelligence

- **Short and Long terms predictive models on:**
  - traffic, parking, people flow, maintenance, land sliding, NO2
- **3D Flow prediction:** Pollutant (NOX, NO2, …)
- **Early warning, City Indexes, etc.**
- **AI & XAI:**
  - RF, XGBoost, BRNN, RNN, SVR, DNN, LSTM, CNN-LSTM, Autoencoders, …
  - Clustering: K-means, K-Medoid, …
  - XAI: Shap, variations, ..
- **Modeling, simulation, routing**
  - **Traffic Flow reconstruction**
  - **Constrained Routing**
- **What-IF analysis** (simulation + AI + data)
- **Based on several computational models**:
  - trajectories, OD matrices, Typical Time Trends, etc.

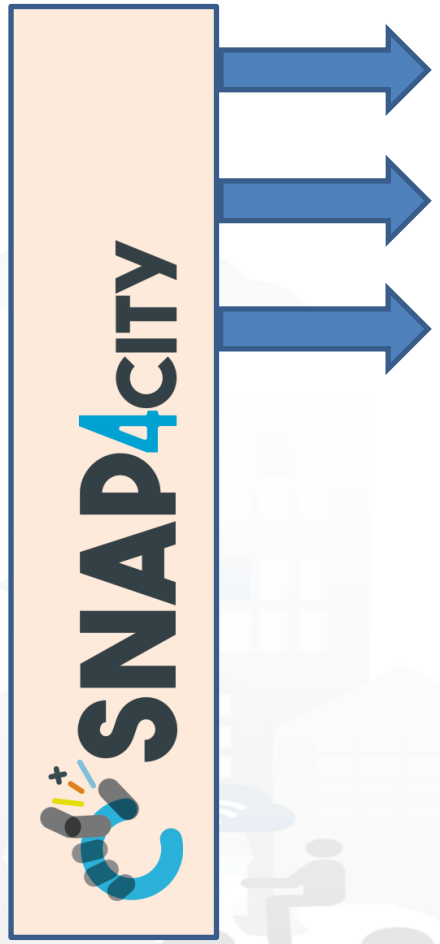
to cope with

- *any data, format*
- *any channel, protocol*
- *any AI/ML*
- *any place*
- *online development*
- *multi-tenant*
- *Secure, PENTest*
- *GDPR, privacy*
- ***→ low costs***
- ***→ easy to evolve***

https://www.snap4city.org/download/video/course2020/da/Snap4City-4th-slot-Data-Analytic-v4-6.pdf

Snap4City data can be accessed /Exported in multiple manners

- **Dashboards and mobile Apps**
- **Downloads in different formats from**
  - dedicated user interface: IoT Data, MyKPI data, Historical data, etc.
  - Tools: Twitter Vigilance user interface
  - API: real time data, historical data, heatmaps, OD matrices, Typical Time trends, predictions, anomaly detections, etc.
- **Programming the massive data export:**
  - Using an IOT App for exporting the data via any kind of channel: WS, FTP, RestCall, Open Data CKAN, etc.
  - Origin destination maps, typical time trends, trajectories, etc.
- **A large range of API:**
  - Smart City API
  - IOT data: IOT Directory, IOT brokers, IoT Protocols
  - KPI: MyKPI
  - GIS data as API: WMS, WFS protocols
  - Linked Open Data from ServiceMap, Knowledge Base
  - Etc…

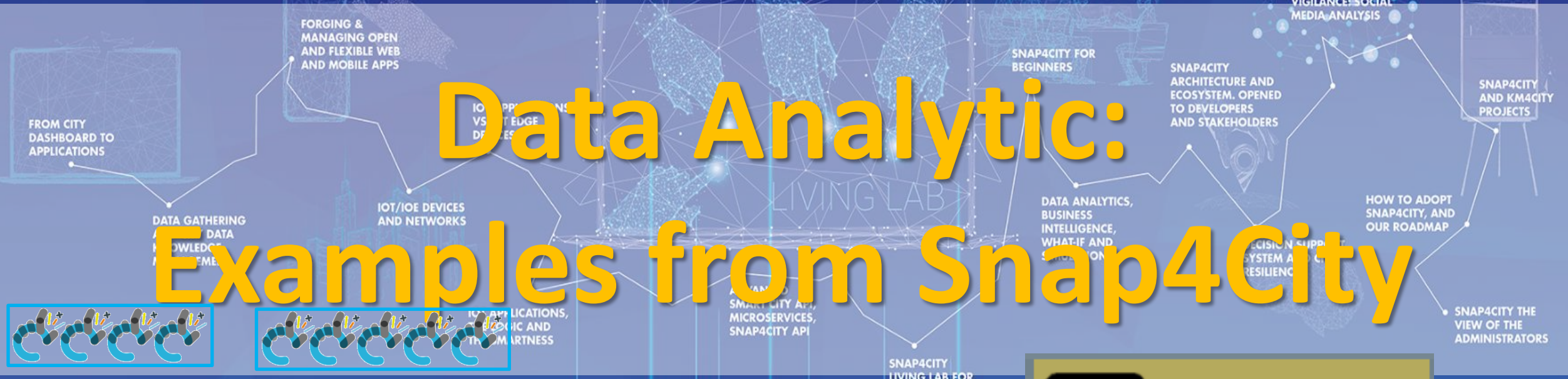


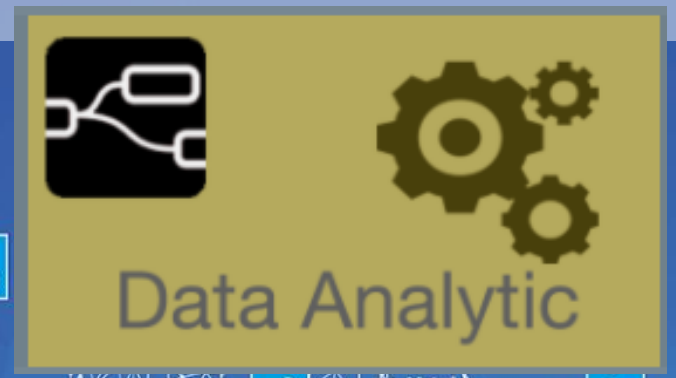
Examples

- **Many Dashboards for all the cities, etc..**
- **Modena and Santiago di Compostela** are receiving in real time traffic flow data which are exploited in other applications
- **Firenze, Pisa, and Livorno** are receiving Heatmaps and NOX predictive heatmaps
- **Antwerp** is periodic downloading data regarding PAXCounters
- **Dubrovnik, Valencia**, etc., accesses to Twitter Vigilance data for further analysis
- Many **Mobile applications** are using the Smart City API, Snap4City and/or Km4City
  - Final used are using: predictions on parking, heatmaps, notifications on heatmaps predictions, routing, typical time trends, etc.
- Etc…

TOP

Data Analytic: Examples from Snap4City





Data vs Smart Services enabling on Snap4City

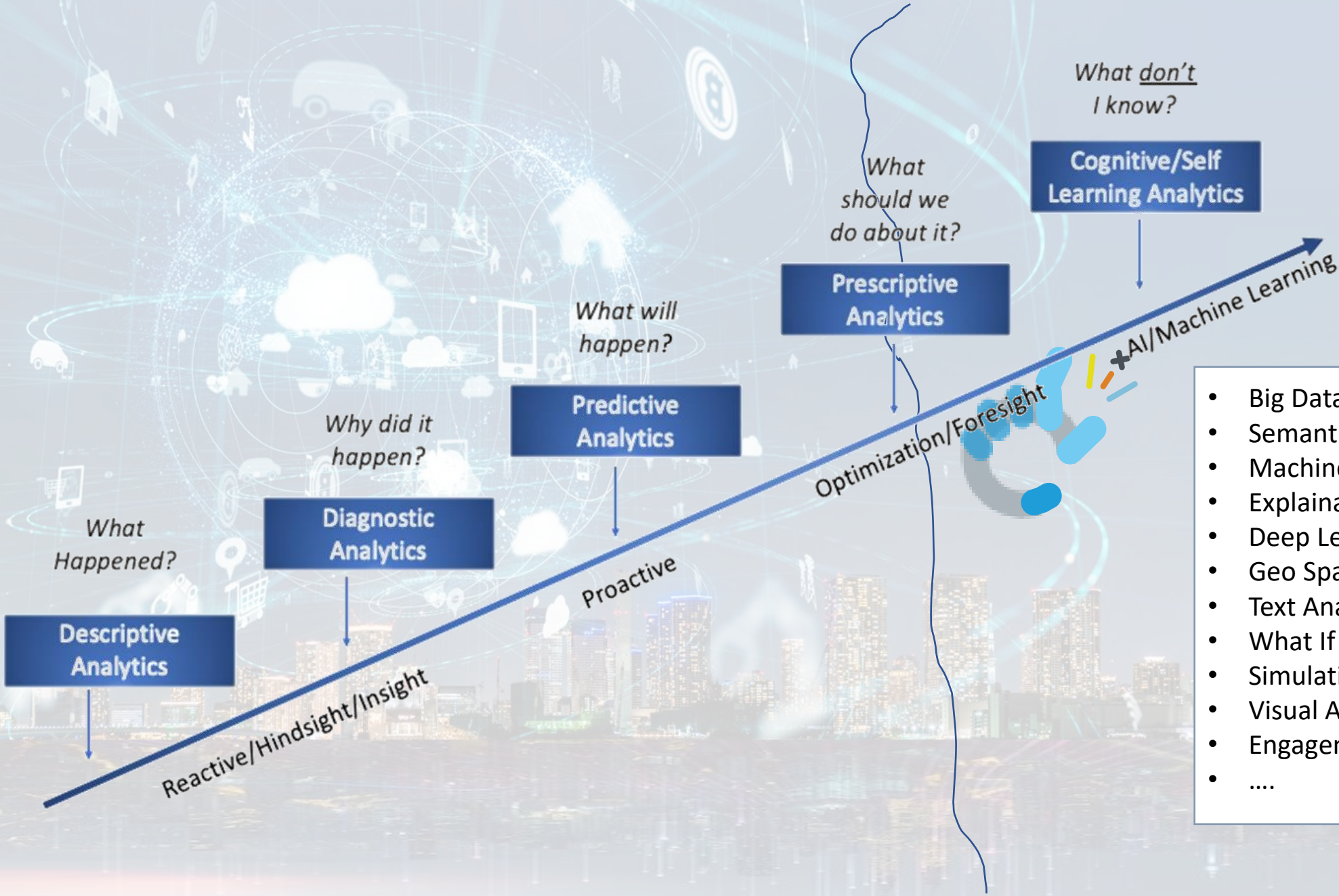
- **Public Transportation and mobility activated services in some where with Snap4City**
  - **Smart parking** (parking locations and real time parking data) … predictions
  - **Smart Fuel pricing** (fuel station locations and real time prices)
  - **Routing** (detailed GIS information, text indexing of streets, POI, etc.)
    - **Quite routing, perfect shopping, etc. etc**. (more data in needed….)
  - **multimodal routing** (detailed GIS information, Public transport time schedule)
  - **Info traffic** (traffic flow sensors, real time Traffic events, their localization, etc.)
  - **Dense info traffic** (traffic flow sensors and traffic flow reconstruction algorithm)
  - **Car/Bike/Scooter Sharing** (position and availability of Cars/Bikes, Scooters) … predictions
  - **Smart Biking** (cycling paths, environmental data) … predictions on bike racks
  - **E-vehicles** (position, status of recharging stations,. …) … predictions vs booking
  - **Smart river crossing** (position and status of Underpass, Ferry) … prediction
  - **Quality of Public Transport** (actual time of arrival at the bus stops, wrt planned time schedule**)**
  - **Early Warning vs Resilience** (combination of several data including mobility, events, Social to perform early warning…)

Data vs Smart Services enabling on Snap4City

- **Social and Users Behaviour**
  - **Smart First Aid** (Location of First AID, real time status of triage)
  - **search for POI and public transport services** (POI geolocalized, spatial queries, along paths)
  - **Social Media Monitoring and acting** (Identif. of dysfunction, quality of service perceived)
  - **Information to Tourists** (Entertainment Events)
  - **Early Warning, prediction of audience** (Twitter data, social media)
  - **Improvement of services for Tourists** (people flow, usage of services)
    (Origin Destination Matrices, trajectories, heatmaps)
    (People Monitoring, via App, Wifi, PAX Counter)
    (Twitter Data, social mea,….)
- **Weather and environment, quality of life**
  - **Weather forecast/condition** (Weather forecast)
  - **Air quality Pollution** (pollution sensors, PM10, PM2.5, NOX, etc.)
  - **Pollination** (Pollination sensors)
  - **Alerting on Air quality for multiple parameters** (Prediction of parameters time slots, notification)
  - **Information Heatmaps for weather and air quality** (air quality sensors, heatmaps, prediction)
  - **Air quality indexes, and forecast** (…….

DATA ANALYTICS

- Resilience
  - Resilience and risk analysis
  - Early warning computation
  - What-if analysis, dynamic routing, origin destination matrices production from a large range of sources
- Mobility and transport
  - Traffic flow reconstruction from sensors and other sources
  - Predictions for: traffic flow, smart parking, smart bike sharing, etc.
  - Analysis of the demand vs offer of mobility according to public transportation and multiple data sources
  - Accidents heatmaps
  - Tracking fleets, people, via devices: OBU, OBD2, mobile apps, etc.
  - Routing and multimodal routing
- Environment and weather
  - NOX, PM10 pollution prediction on the basis of traffic flow, 48 hours
  - Long term prediction of European Commission KPIs on NOX, PM10, etc.
  - Heatmaps production, dense data interpolation
- User and Social
  - People flows prediction and reconstruction, via Wi-Fi, mobile apps, etc.
  - User engagement for sustainable mobility
  - User's behaviour analysis, origin destination matrices, hot places, time schedule, Recency and frequency, permanence, etc.
  - People flow analysis from PAX Counters
  - Social media analysis on specific channel, specific keywords: see Twitter Vigilance, for NLP and Sentiment Analysis, SA
  - Tweet proneness, retweet-ability of tweets, impact guessing
  - Audience prediction to TV channels and physical events
- Generic
  - Data quality assessment, prediction, anomaly detection
  - Maintenance prediction and costs predictions
  - Estimation of KPI and local indexes for: quality of life, 15 minutes, etc.



Sentient and active processes

- Big Data Analytics
- Semantic Computing
- Machine Learning
- Explainable Artificial Intelligence
- Deep Learning
- Geo Spatial Reasoning
- Text Analysis, Sentiment Analysis
- What If Analysis
- Simulations
- Visual Analytics
- Engagement Analysis
- ….

Disappearing Data Analytics

	Antwerp					Helsinki								Where				
SNAP4CITY	City official	ICT official	Developer	Citizen, tourist, visitor	Business owner	City officials	City officials Domain experts	City officials City developers	Third party developers	Citizen	Citizens with respiratory problems	Tourists	Business owners	Mobile	MIcroApplication	Tool, via Portal (ICT Developers)	Dashboards	RULE OF THUMB Main Data Sources
Discovery near to me	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x			POI, OSM
Discovery along a path	x	x	x	x		x		x	x	x	x	x		x	x			POI, OSM
Discovery in an area, shape	x	x	x	x	x	x	x	x	x	x	x	x	x	x		x		POI, OSM
browsing Public Transport	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x			OSM, GTFS
Full Text search	x	x	x	x	x	x		x	x	x	x	x	x	x		x		POI, OSM
Routing: pedestrian				x	x			x	x	x	x	x	x	x	x			OSM
Routing: pedestrian quite				x	x			x	x	x	x	x	x	x	x			OSM
Routing: private vehicles	x		x	x		x		x	x	x	x	x		x	x			OSM
Routing: Multimodal Public Transport				x					x	x	x	x		x	x	x		OSM, GTFS
heatmaps: weather (Temp, Humidity)	x	x		x	x	x	x		x	x	x	x	x	x			x	Sensors data, OSM
heatmaps: environmental variables, PM10, PM2.5, NO2, EAQI	x	x		x	x	x	x		x	x	x	x	x	x			x	Sensors data, OSM
heatmaps: environmental variables, Noise						x	x		x	x	x	x	x	x			x	Sensors data, OSM
heatmaps: safe on bike (Antwerp)	x	x		x	x									x			x	Spec. Portal
heatmaps: Enfuser prediction, PM10, PM2.5, AQI						x	x		x	x	x	x	x	x			x	Enfuser data
heatmaps piking values any place	x	x			x	x	x	x	x				x				x	Computed Heatmps
heatmaps: GRAL prediction, PM10						x	x		x	x	x	x	x	x			x	OSM, Traffic, Weather
Comparsison: Enfuser, Gral, Real Time						x	x										x	Enfuser, Sensors, GRAL
Sensors Data Time Trends, & drill down	x	x	x		x	x	x	x					x			x	x	Sensors data, OSM
Weather Forecast	x	x		x	x	x	x		x	x	x	x	x	x			x	Forecast Service
Origin Destination Matrices	x	x	x		x	x	x	x	x				x				x	Snap4City Mobile App
Typical trajectories	x	x	x	x	x	x	x	x	x				x			x	x	Snap4City Mobile App
Hot Area in the city	x	x	x	x	x	x	x	x	x	x	x	x	x	x		x	x	Snap4City Mobile App
Hot Places in Smart Zone	x	x	x	x	x									x		x	x	Snap4City PAXcounters
Services Suggestions on mobiles				x						x	x	x		x	x			Snap4City Mobile App
Alerts on critical cases: several variables	x			x	x	x	x			x	x		x	x				Sensors data, OSM
The most used services		x		x	x		x			x	x	x	x				x	Snap4City Mobile App
Twitter Trends Daily	x	x	x		x	x	x	x	x				x			x	x	Twitter Vigilance
The auditing of user and living lab		x				x		x								x		Snap4City Portal
Self assessment	x	x	x	x	x	x	x	x	x	x	x	x	x			x		Snap4City Portal
Trajectories reg from mobile PAX Counters	x	x	x			x	x	x							x		x	PAX Counters
Engagement real time assessment	x	x	x			x	x	x									x	Snap4City Mobile App

Design of Data Analytics & smart solution

- Profiling for user kind, they need different
  - Data analytics
  - Devices: mobile app, desktop, dashboards
  - user interface for interaction, intelligence, KPI
  - data representation and analytics, interactivity, business intelligence

From Simple Data Analytic to Complex Tools

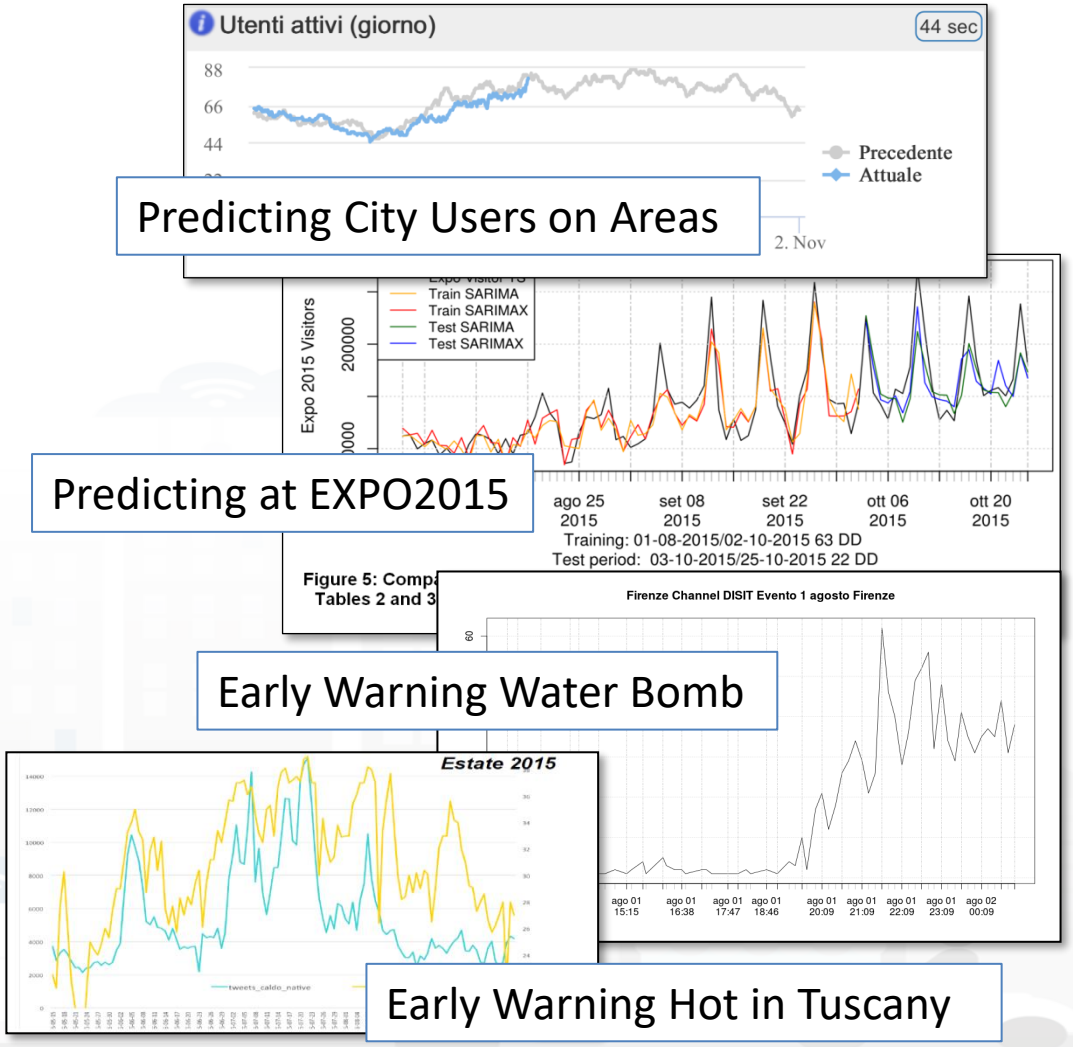
- **Structural:**
  - **Data Ingestion, Quality Control** on data: data mining, anomaly detection, etc.
  - **Typical Time trends**: traffic flow, people flow, sensors data, etc.
  - **Indexing** for fast search and retrieval: Geospatial, textual, temporal, mixt
- **Dynamical:**
  - **Analysis**: heatmap, hot places, distribution, statistical analysis
  - **Predictions** to inform and plan (e.g.: parking, people flow,)
  - **Anomaly detection** for Early Warning, Alerting
- **Special Analytics and Tools** → What-IF Analysis:
  - **Routing** for navigation: modal, multimodal, constrained
  - **Typical Trajectories** of: people flows, vehicles, etc.
  - **Traffic Flow** reconstruction
  - **Origin Destination Matrices:** people and vehicles, …
  - **Simulations**: demand vs offer, etc.

Snap4City and Data Analytic (summary)

- allows to create simple data processing as well as massive computing solutions exploiting statistics, machine learning, operating research, etc. for:
  - predictions, anomaly detection, early warning, OD Matrix construction, simulation, trajectories, typical trends, what-if analysis, smart routing, heatmaps, etc.
- **can be developed** in:
  - R Studio / Tensor Flow, Java, Python, ETL, IOT Applications
  - If HDFS/Hadoop/Hbase/Phoenix is installed: MapReduce, Spark, etc.
- **may be shared** with other colleagues, and organizations via the Resource Manager

Predicting Models for Administrators & City Users

- **Aiming at improving**
  - quality of service, distributing workload
  - early warning
- **Predictions**:
  - Short (15 min, 30 Min), mid Term (1 week), long term (months)
- **Data Analytics**: ML/AI, NLP/SA, Clust., …
  - Traffic Flows → multi-flow reconstruction
  - Parking Status → free slots
  - Environmental Alarms
  - Air Quality parameters and indexes
  - People Flows (Wi-Fi, Twitter) → crowd , #number of people



Data Analytics on Snap4City platform

Swagger

Ontology Schema

LOG.disit.org

Knowledge Base, Km4City

Big Data Store Facility

Smart City API from Knowledge Base and other tools

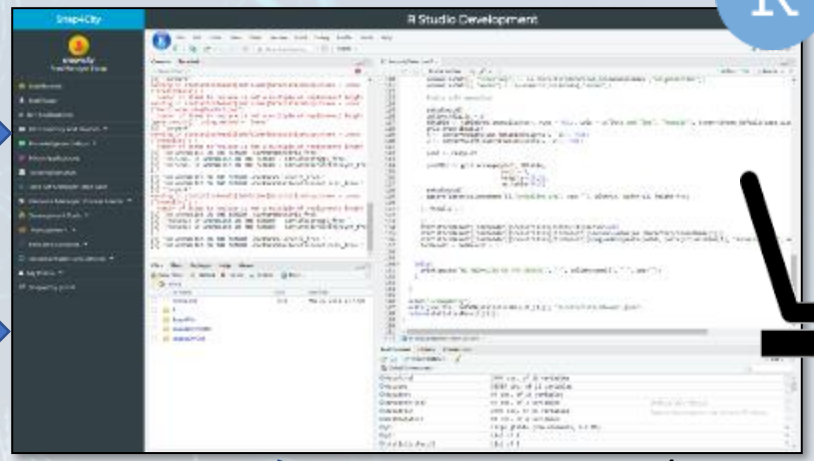
TensorFlow

NVIDIA CUDA

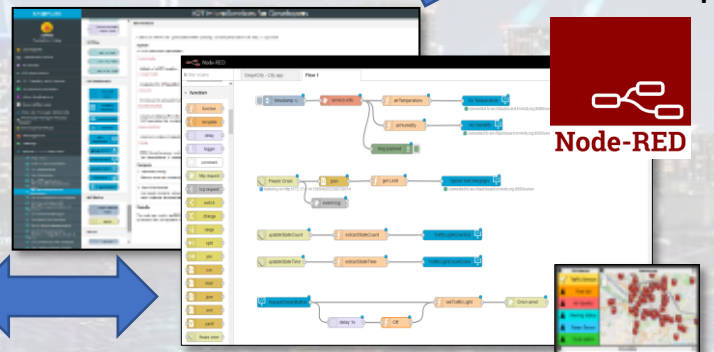
python

jupyter

R Studio



Creating MicroServices



Node-RED

Using them into IOT Applications

Saving / Sharing reusing

Resource Manager

Data Analytics Life Cycle

- **Problem analysis**, business requirements
- **Data Discovery**, Data ingestion, acquisition
- **Data set preparation**, transformation, identification of features
- **Target Assessment Model Definition**
  - Identification of metrics for the assessment, KPI
- **Screening on Models/Techniques**
- **For each Model/Technique or for the selection Model/Technique** perform the
  - ***Model/Technique Development/testing***
- **Best Model selection among those tested**
  - If needed reiterate for different parameters, features, etc.
- **Deploy best Model in production**

Model/Technique Development/testing

- **Planning**
  - How to compute, which language
  - Which environment, which libraries
- **Model building**
- **Training**
- **Technical Assessment**
  - **Validation in testing**
  - **Assessment** on a set of metrics depending on the goals
- **Impact Assessment**
  - **Ethic Assessment** and incidental findings
  - **Global and Local Explanation**
- Optimisation of computation cost for features, if needed reiterate

Some Measures

The $\mathbf{R^2}$ is calculated as follows:

$$\overline{y} = \frac{1}{n}\sum_{i=1}^{n} obs_i \qquad \boldsymbol{R^2} = 1 - \left(\frac{\sum_{i=1}^{n}(obs_i - pred_i)^2}{\sum_{i=1}^{n}(obs_i - \overline{y})^2}\right)$$

The **MAE** is calculated as follows:

$$\boldsymbol{MAE} = \frac{\sum_{i=1}^{n}|obs_i - pred_i|}{n}$$

The **MAPE** is calculated as follows:

$$\boldsymbol{MAPE} = \frac{\sum_{i=1}^{n}\left|\frac{obs_i - pred_i}{obs_i}\right|}{n} * 100$$

The **RMSE** is calculated as the Root square of the Mean Squared Error (**MSE**):

$$\boldsymbol{MSE} = \frac{\sum_{i=1}^{n}(obs_i - pred_i)^2}{n}, \qquad \boldsymbol{RMSE} = \sqrt{MSE}$$

where: $obs_i = observation\ at\ time\ i,\ pred_t = prediction\ at\ time\ t,$
n is the number of the values in the test set.

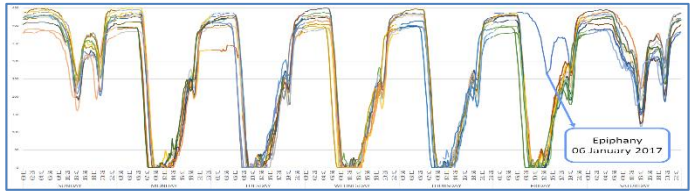
Development in R Studio (self training)

- R Studio Development
- TC7.2 - R Studio for Analytics, exploiting Tensor Flow
- TC7.4 - From R Studio process to MicroService for IOT application, data analytics, machine learning
- TC7.5 - Developing Data Analytics Processes
- US7. Data Analytics and related integration aspects

Smart Parking: predictions

autostrada
FIESOLE
SIENA
AREZZO - FORLI'

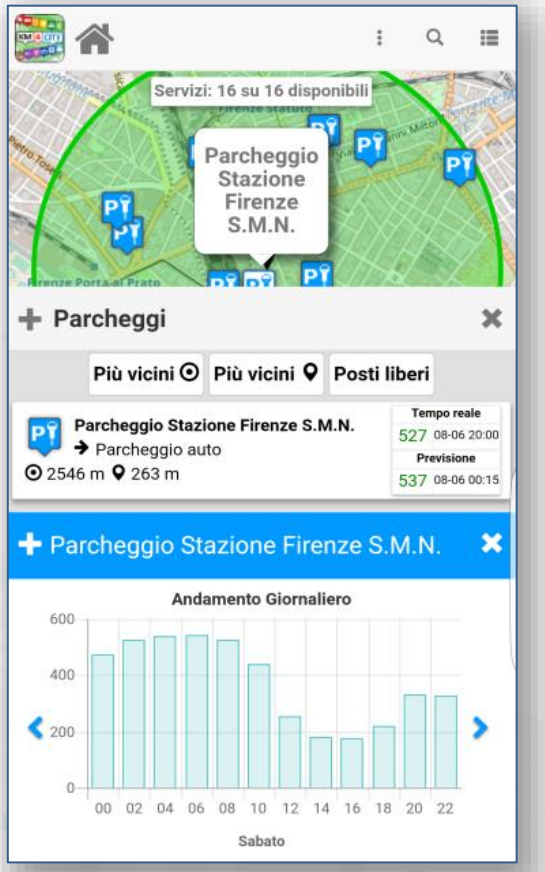
I would arrive to surely Park in 45 Minutes??



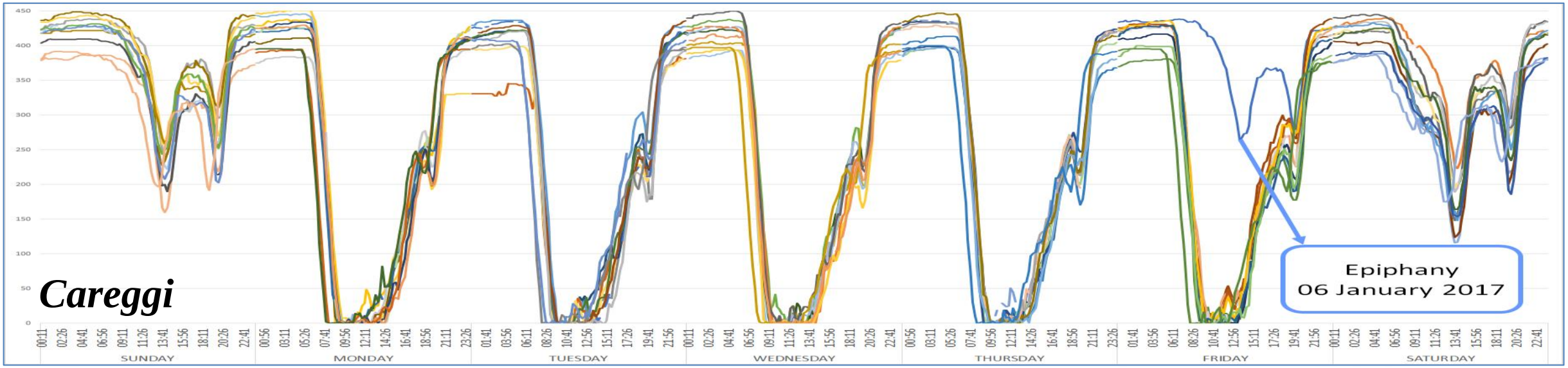
Category	Features	Description of features variable
Baseline features of free slot data	Free parking slots	Real number of available slots recorded every 15 minutes
	Time	Hours and minutes
	Month	Month of the year (1-12)
	Day	Day of the month (1-31)
	Day week	Day of the week (0-6)
	Weekend	0 for working days, 1 else
	Previous observation's difference (POD)	Difference between the number of free spaces at time i and number of free spaces at time (i − 15 minutes) recorded in the previous week
	Subsequent observation's difference (SOD)	Difference between the number of free spaces at time i, and the number of free spaces at time (i + 15 minutes) recorded in the previous week
Weather features	Temperature	City temperature measured one hour earlier than Time (°C)
	Humidity	City humidity measured one hour earlier than Time (%)
	Rainfall	City rainfall measured one hour earlier than Time (mm)
Traffic Sensors features	Average Vehicle Speed	Average speed of vehicles on the road being closest to the parking, over one-hour period (km/h)
	Vehicle Flow	Number of vehicles passing by closest to the parking, over one-hour period
	Average Vehicle Time	Average of distance between vehicles, over one-hour period
	Vehicle Concentration	Number of vehicles per kilometer, over one-hour period

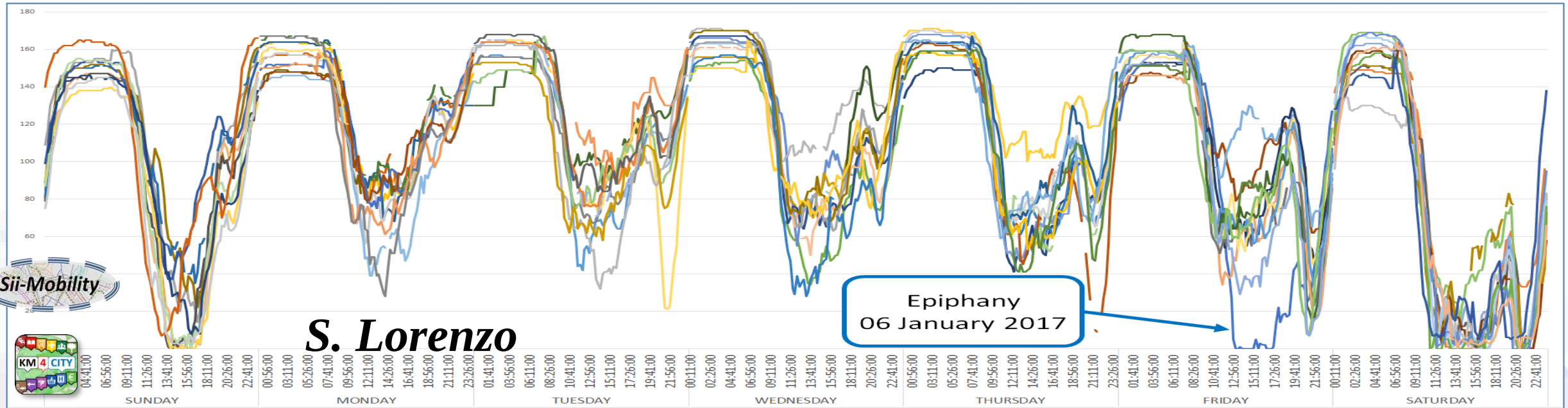
Artificial Intelligence Predictions

97% of precision



Free Parking space trends

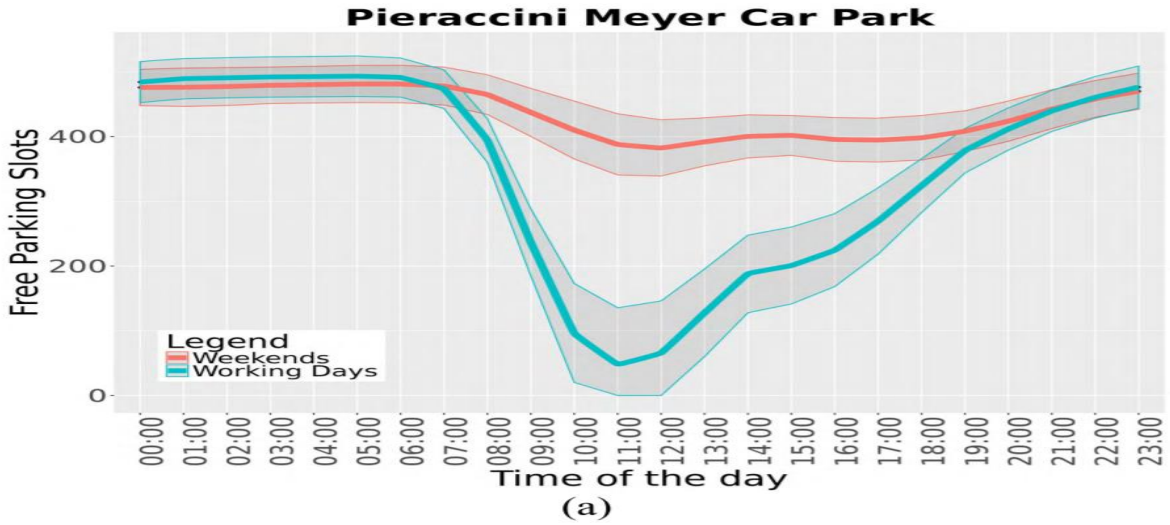


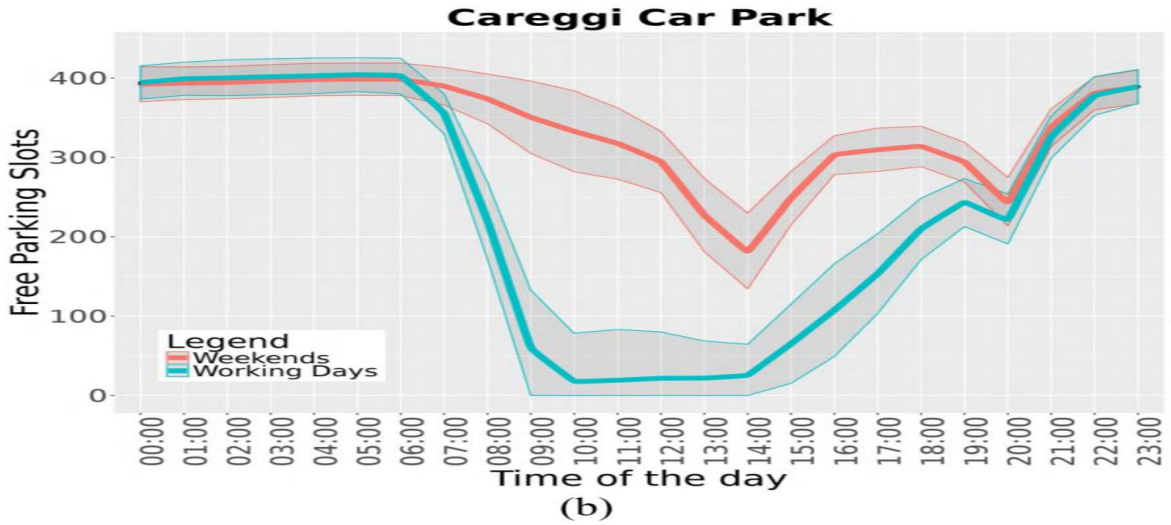


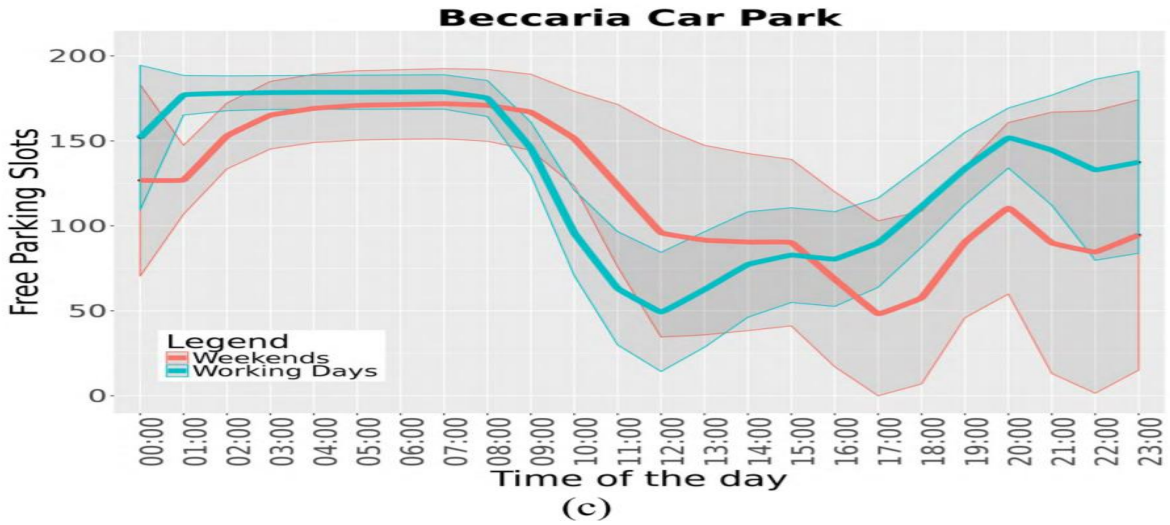
Sii-Mobility

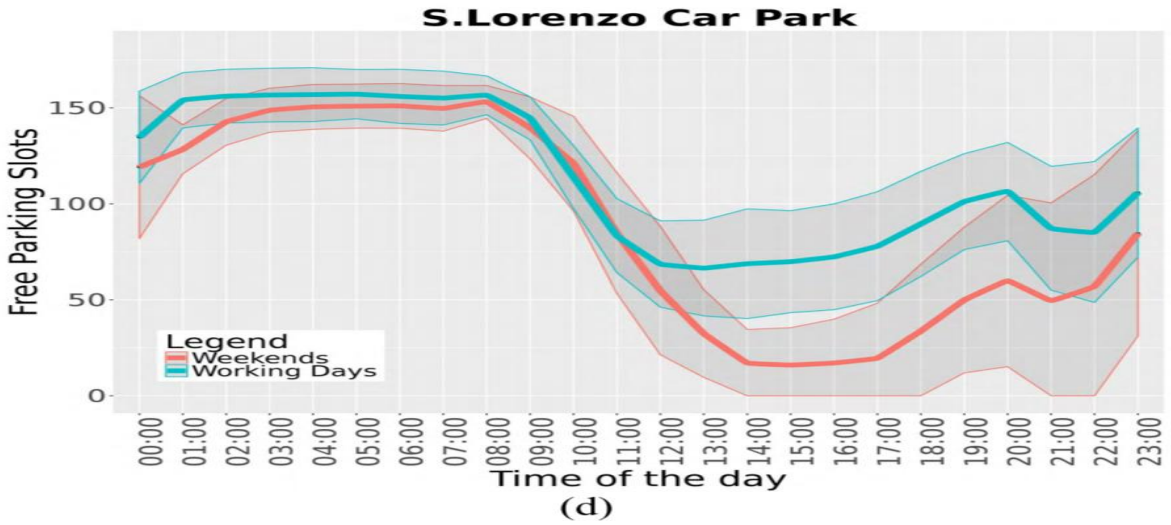
KM 4 CITY

Free Parking space trends







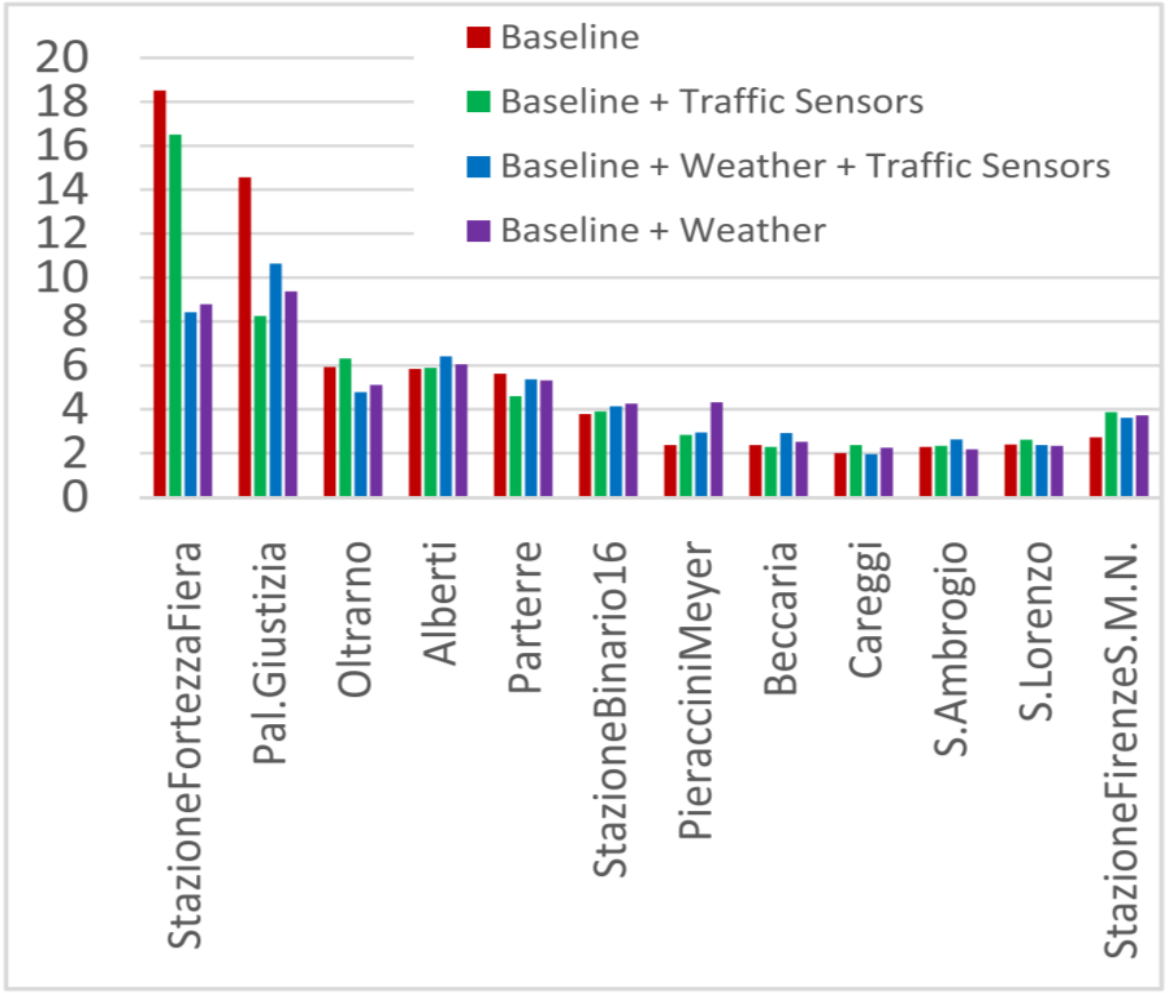


12 parking areas in Florence

Free Parking PREDICTIONS

C. Badii, P. Nesi, I. Paoli, "Predicting available parking slots on critical and regular services exploiting a range of open data", IEEE Access, preprint, 2018, https://ieeexplore.ieee.org/abstract/document/8430514/

Comparison Error	Forecasting Techniques		
	BRANN	SVR	RNN
Careggi car park			
MASE Night	34.85	16.29	20.01
MASE Morning	0.76	1.42	2.82
MASE Afternoon	1.89	4.34	3.66
MASE Evening	1.99	1.51	2.33
MASE	1.87	2.34	3.16
Pieraccini Meyer car park			
MASE Night	6.08	12.83	10.03
MASE Morning	0.86	1.27	4.90
MASE Afternoon	1.87	2.91	6.75
MASE Evening	1.36	1.57	10.23
MASE	1.37	2.06	6.67
S. Lorenzo car park			
MASE Night	10.33	11.81	18.34
MASE Morning	2.13	1.91	3.93
MASE Afternoon	2.70	3.15	2.37
MASE Evening	2.15	3.09	3.82
MASE	2.72	3.21	4.19
Beccaria car park			
MASE Night	9.32	7.80	12.47
MASE Morning	0.95	1.25	4.87
MASE Afternoon	2.49	2.14	2.45
MASE Evening	2.96	4.75	5.91
MASE	2.13	2.67	4.85



ML models

The best selected models for the purpose have been:

- BRNN:
  - Bayesian Regularized Neural Network
- SVR:
  - Support Vector Regression
- ARIMA
  - Autoregressive Integrated Moving Average
- RNN
  - Recurrent neural networks

BRNN: Bayesian Regularized Neural Network

$$y_i = g(x_i) + e_i$$
$$y_i = \sum_{k=1}^{s} w_k g_k \left(b_k + \sum_{j=1}^{p} x_{ij} \beta_j^{|k|} \right) + e_i, \quad i = 1, \ldots, n$$

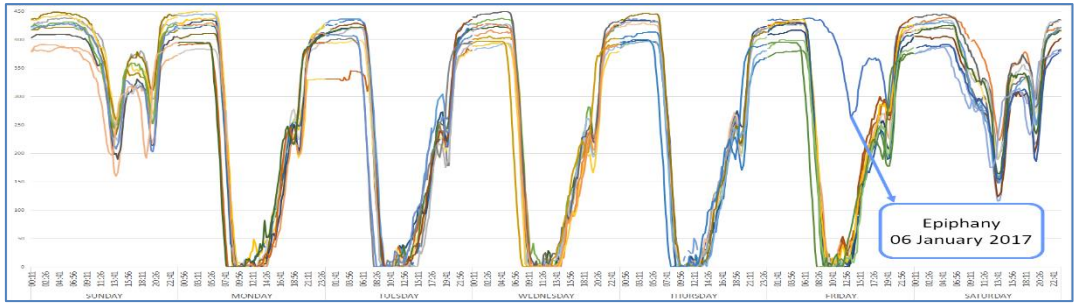
- $e_i \sim N(0, \sigma_e^2)$;
- s is the number of neurons;
- w_k is the weight of the k -th neuron, $k = 1, \ldots, s$;
- b_k is a bias for the k -th neuron, $k = 1, \ldots, s$;
- $\beta_j^{|k|}$ is the weight of the j -th input to the net, $j = 1, \ldots, p$;
- $g_k(\cdot)$ is the activation function: in this case

$$g_k(x) = \frac{e^{2x} - 1}{e^{2x} + 1}$$

The objective function consists of minimizing $F = \alpha E_W + \beta E_D$, where E_W is the sum of squares of network parameters (weight and bias), and E_D is the error (sum of squares), α and β are the objective function parameters.

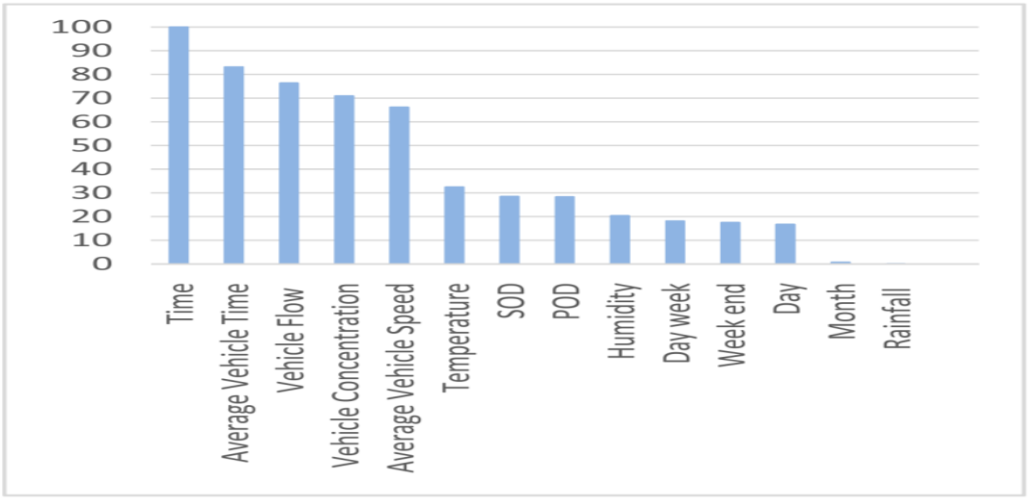
UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB

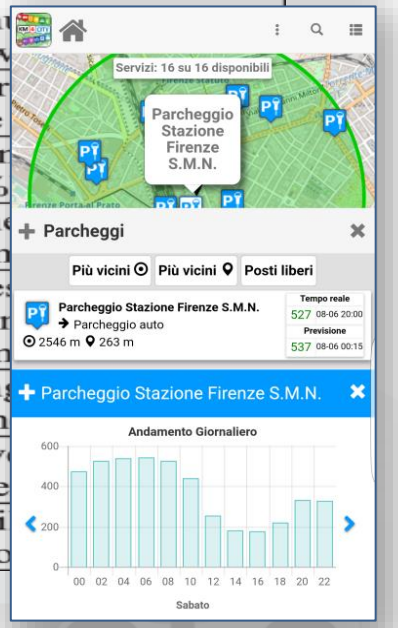
Free Parking PREDICTIONS



Category	Features	Description of features variable
Baseline features of free slot data	Free parking slots	Real number of available slots recorded every 15 minutes
	Time	Hours and minutes
	Month	Month of the year (1-12)
	Day	Day of the month (1-31)
	Day week	Day of the week (0-6)
	Weekend	0 for working days, 1 else
	Previous observation's difference (POD)	Difference between the number of free spaces at time i and number of free spaces at time (i − 15 minutes) recorded in the previous week
	Subsequent observation's difference (SOD)	Difference between the number of free spaces at time i, and the number of free spaces at time (i + 15 min in the previous w
Weather features	Temperature	City temperature measur earlier than Time
	Humidity	City humidity measured or than Time (%
	Rainfall	City rainfall measured on than Time (mn
Traffic Sensors features	Average Vehicle Speed	Average speed of vehicles being closest to the parkir hour period (km
	Vehicle Flow	Number of vehicles passing the parking, over one-h
	Average Vehicle Time	Average of distance betw over one-hour pe
	Vehicle Concentration	Number of vehicles per ki one-hour perio

Relevance of Variables





C. Badii, P. Nesi, I. Paoli, "Predicting available parking slots on critical and regular services exploiting a range of open data", IEEE Access, preprint, 2018
https://ieeexplore.ieee.org/abstract/document/8430514/

Free Parking PREDICTIONS

Performances

Training	Forecasting Techniques			
	BRANN	SVR	RNN	ARIMA
Average Training processing time (sec)	76.3	9.1	598.7	9.2
Re-Training frequency	Daily	Daily	Daily	Hourly
Training period	3 months	3 months	3 months	3 months
Estimation	**BRANN**	**SVR**	**RNN**	**ARIMA**
Average Estimation time (sec)	0.0031	0.0052	0.034	0.0015
Estimation frequency	Hourly	Hourly	Hourly	Hourly
Estimation predicted period	1 hour	1 hour	1 hour	1 hour

Free Parking Predictions

Careggi car park			
Model features	BRNN model results		
	R-squared	RMSE	MASE
Baseline	0.974	24	1.87
Baseline + Weather	0.975	24	1.75
Baseline + Traffic sensors	0.975	24	2.04
Baseline + Weather + Traffic sensors	0.975	24	1.87

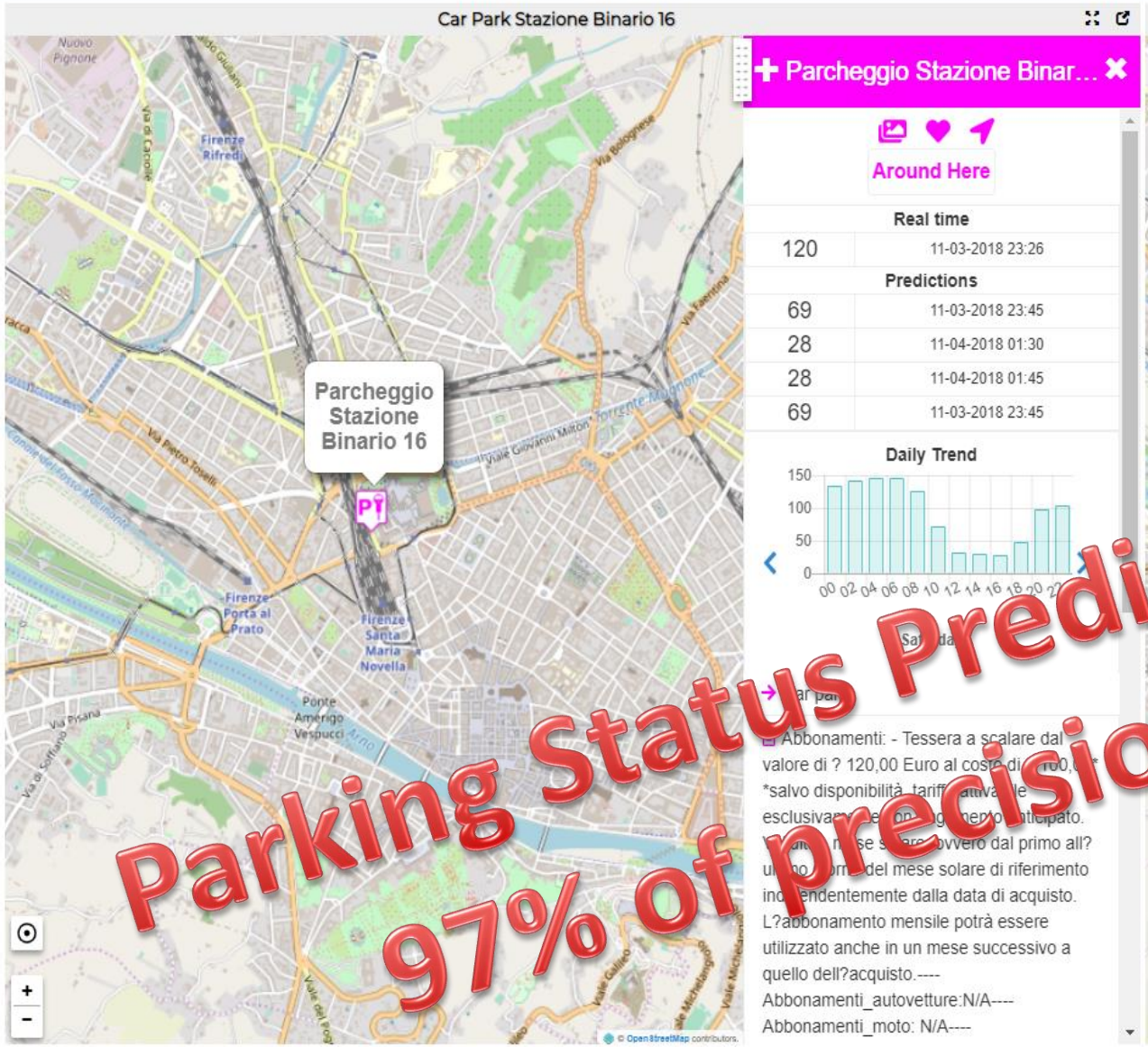
Active on Mobile Apps as:

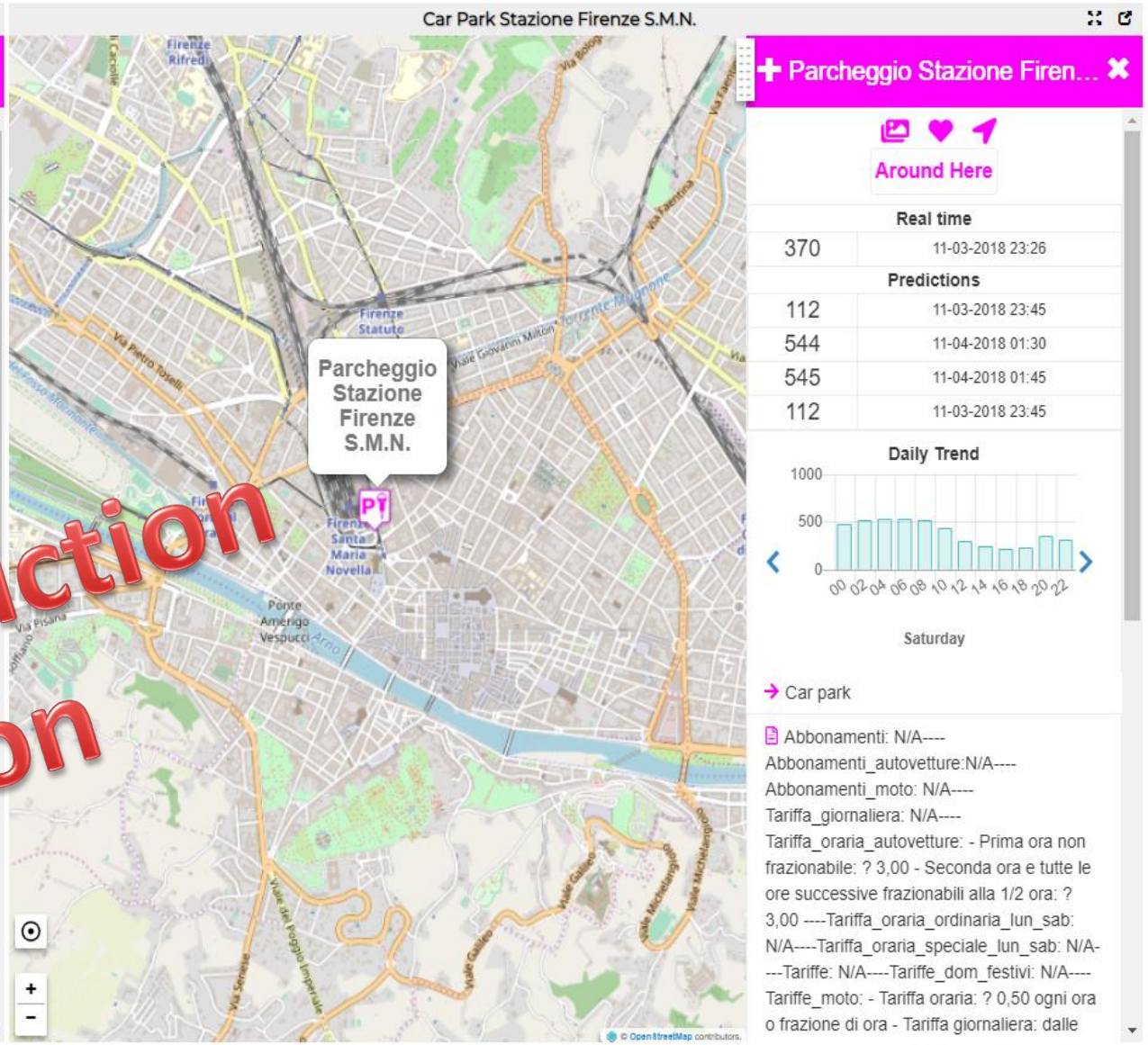
- «Firenze dove cosa»
- «Toscana dove cosa»

Precision: 97,5%

Monitoring Station for Parking

Sat 3 Nov 23:39:55





Parking Status Prediction

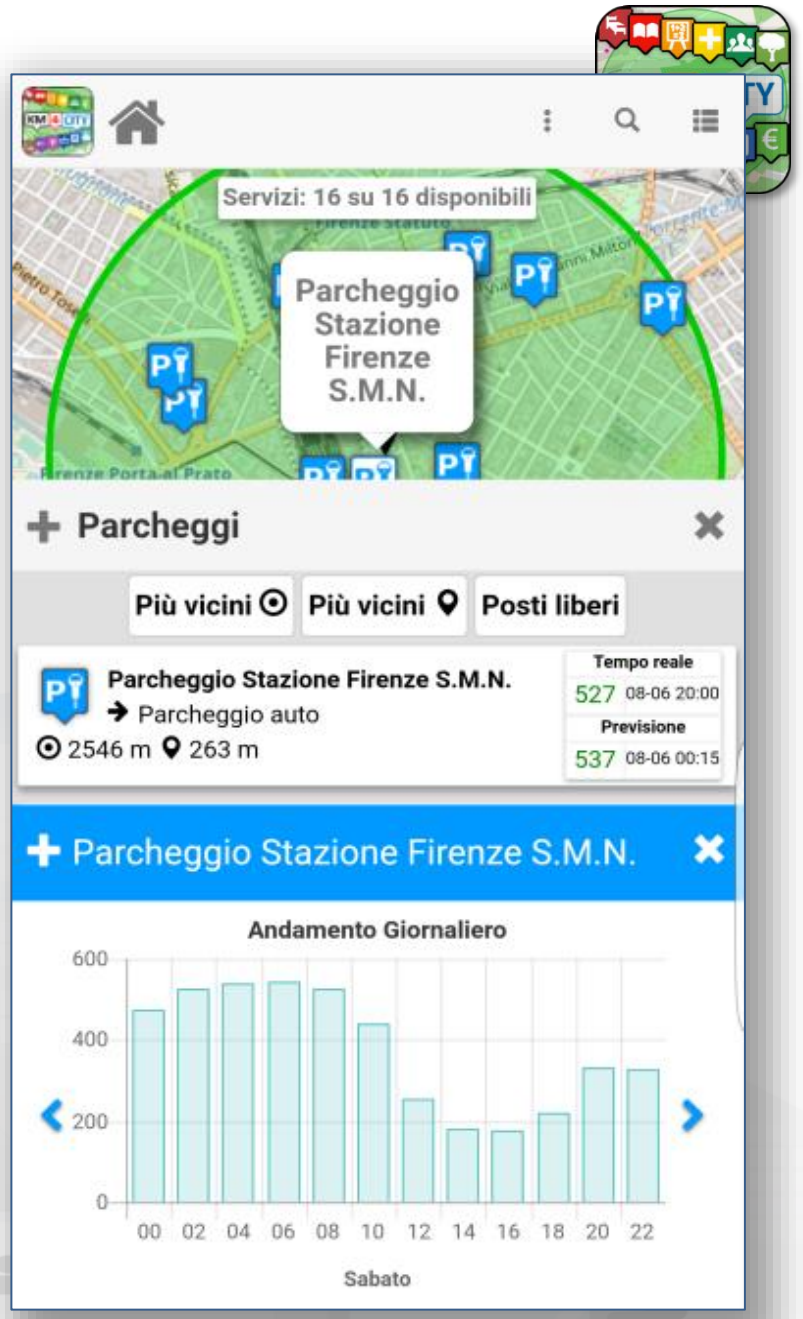
97% of precision

https://www.disit.org/dashboardSmartCity/view/index.php?iddasboard=MjQ2

Predictions on Parking

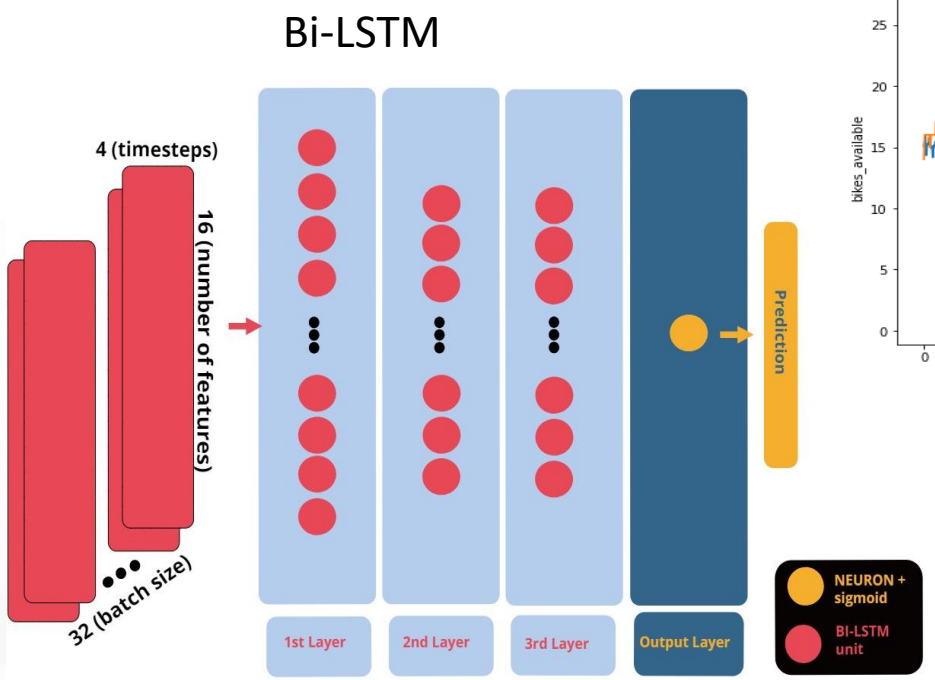
IEEE Access

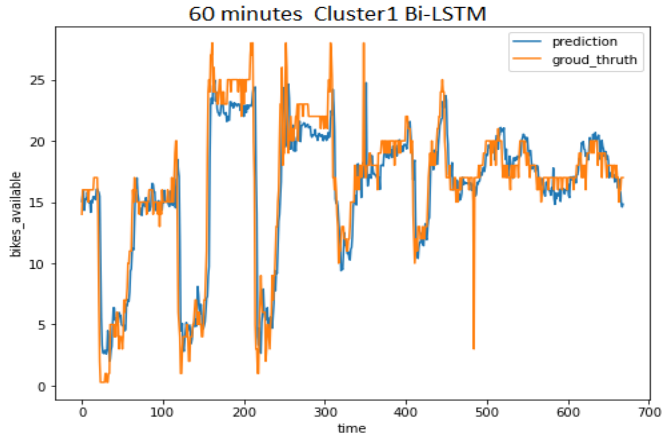
- C. Badii, P. Nesi, I. Paoli, "Predicting available parking slots on critical and regular services exploiting a range of open data", IEEE Access, preprint, 2018, https://ieeexplore.ieee.org/abstract/document/8430514/



Smart Bike Sharing

Deep Learning for Short-Term Prediction of Available Bikes on Bike-Sharing Stations





E. Collini, P. Nesi and G. Pantaleo, "Deep Learning for Short-Term Prediction of Available Bikes on Bike-Sharing Stations," in *IEEE Access*, vol. 9, pp. 124337-124347, 2021, doi: 10.1109/ACCESS.2021.3110794.
https://ieeexplore.ieee.org/abstract/document/9530580

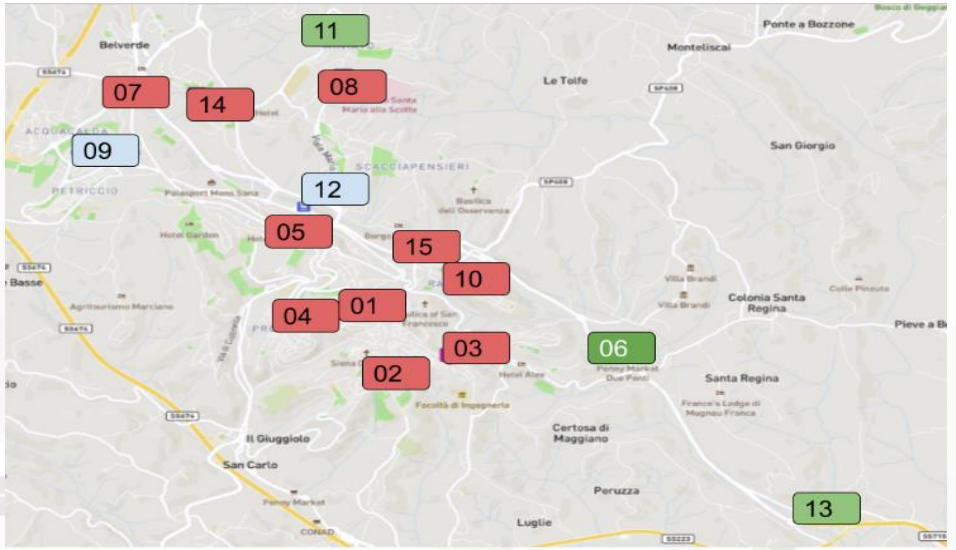
Bike Sharing

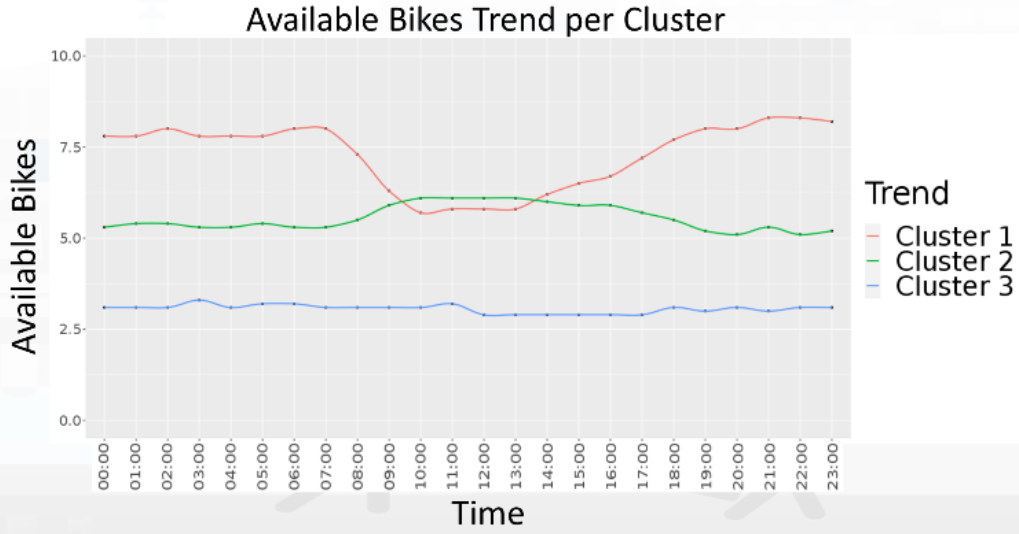
- Pros:
  - Eco-friendly
  - Prevent traffic congestions
  - Reduce the probability of social contacts in public transports
  - Regular bikes or e-bikes
- Problems:
  - Irregular distribution of bikes on racks/areas
  - Difficulty of knowing in advance their status with a certain degree of confidence
    - available bikes at a specific bike-station
    - free slot for leaving the rented bike

→ providing **PREDICTIONS** can be useful to improve quality of service

Clustering

- A **clustering** approach has been applied in order to classify Pisa and Siena stations based on their mean trend H24 of bikes availability
  - This is also correlated to the typical services in the neighbourhoods
- **K-means** clustering method has been applied to identify clusters
  - The optimal number of clusters resulted to be equal to **3**, and it has been identified by using the **Elbow criteria**





Features

Category	Feature	Description
target	AB	Number of available Bikes
Baseline-Historical	Time	The observation time hh-mm-ss
	month	Month of observation {1-12}
	Day Of The Week	Day of the week {1-7}
	Weekend	1 if the observation day is Saturday of Sunday, 0 otherwise
Differences Over Time	dP	the difference between the number of available bikes in the observation day (d) at the time slot t and the number of available bikes during the previous time slot (t-1) of the previous day (d-1)
	dS	the difference between the number of available bikes in the observation day (d) at the time slot t and the number of bikes during the successive time slot (t+1) of the previous day (d-1).
	PwAB	the number of available bikes of the previous week (d-7) in the same time slot (t).
Real-time weather and weather forecast	Temperature	Air temperature at the observation time, in °C
	Max Temperature	Forecast of max temperature of the observation day, in °C
	Min Temperature	Forecast of Min temperature of the observation day, in° C
	Humidity	Humidity of the hour prior to the observation time, in percentage
	Rain	mm of rain registered in the hour prior to the observation time
	Pressure	Pressure at the observation time, in millibar (mb)
	Wind Speed	Average wind speed registered in the hour prior to the observation time, in km/h
	Cloud Cover Percentage	Cloud Cover Percentage at the observation time

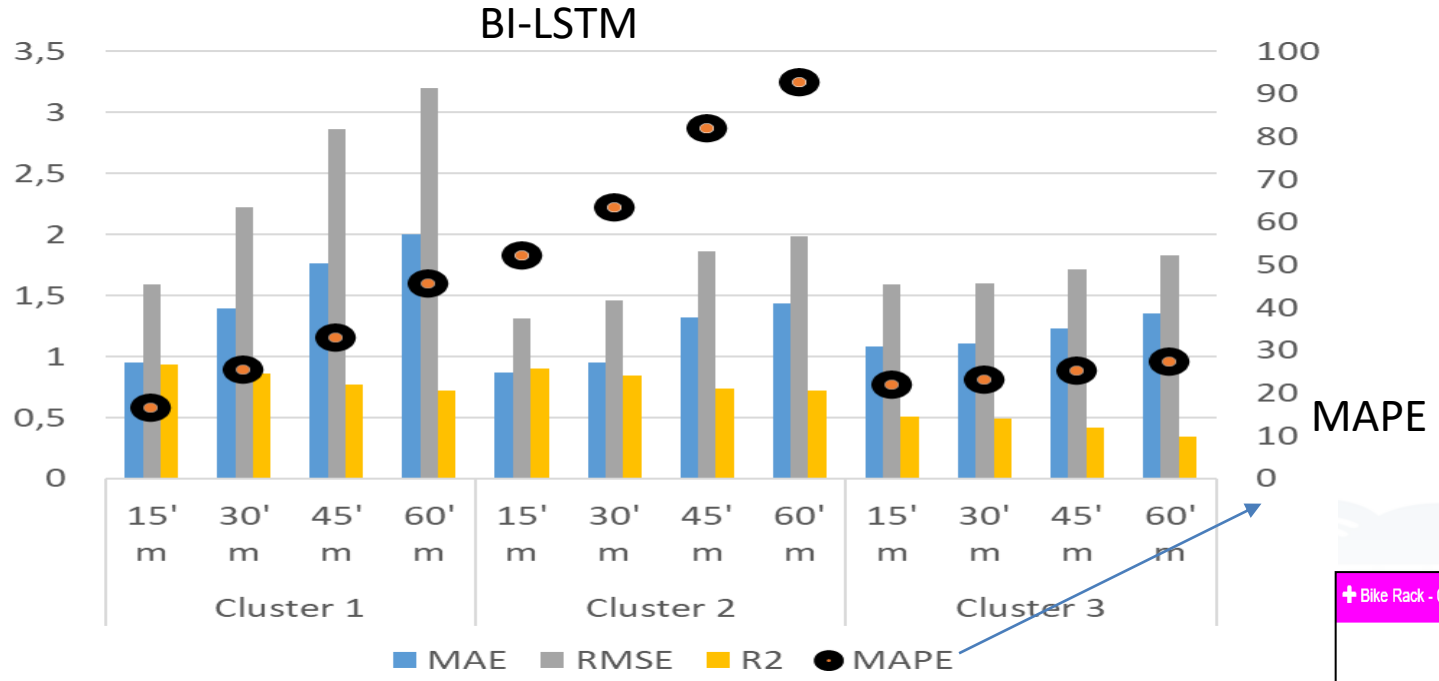
Comparison

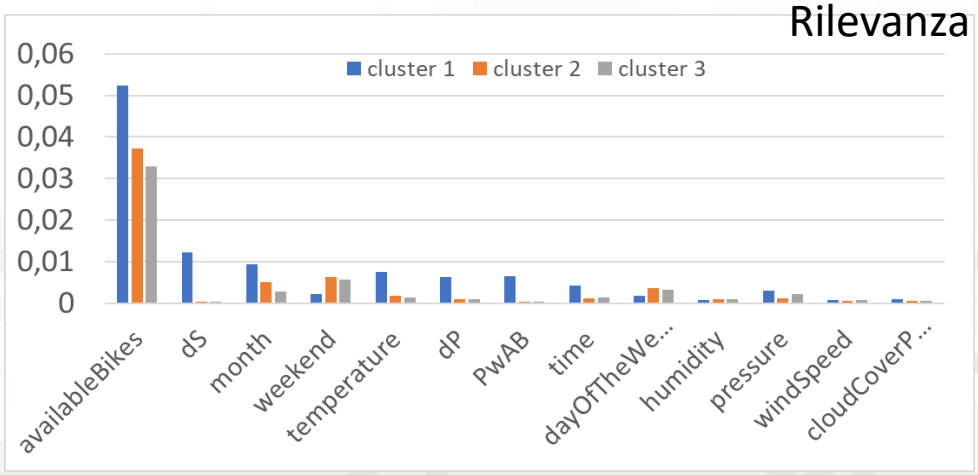
Comparative	Cluster1:				Cluster2:				Cluster3:			
Results	15'	30'	45'	60'	15'	30'	45'	60'	15'	30'	45'	60'
RF	35.16	44.93	53.73	59.57	107.03	146.16	196.55	238.49	30.29	31.60	35.13	36.49
XGBoost	18.75	27.16	40.33	49.09	58.43	83.54	112.46	119.56	28.62	27.30	26.97	29.36
DNN	21.12	28.39	36.01	49.56	109.69	127.23	149.84	178.23	30.29	28.00	27.98	28.68
LSTM	17.68	40.56	44.54	51.16	85.09	120.00	79.30	164.00	22.13	22.91	26.21	25.88
Bi-LSTM	16.46	25.35	33.00	45.53	52.18	63.45	132.00	92.62	21.98	23.00	25.15	27.32

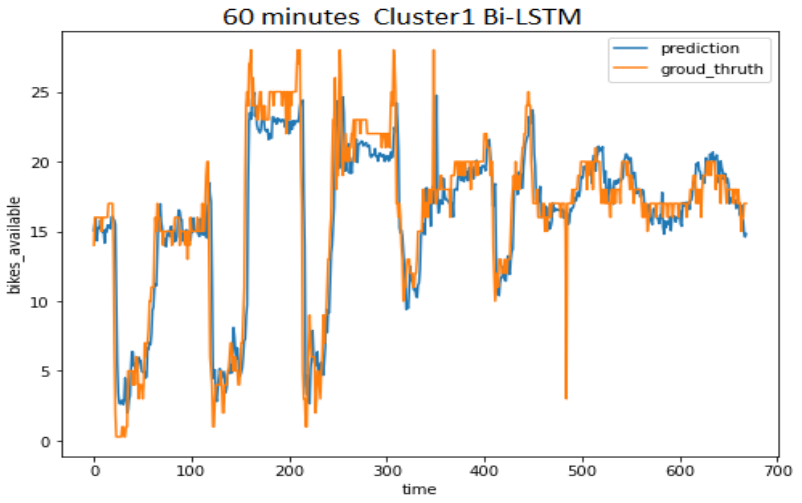
- Among the solutions XGBoost, Deep NN, LSTM, and Bidirectional LST
  - The Bi-LSTM resulted to be the better ranked in terms of MAPE

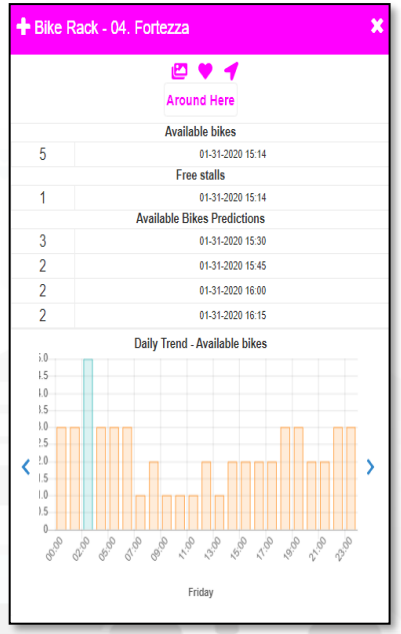
Siena Bike Sharing

- For each Bike Rack, Prediction of the number of
  - available bikes in sharing
  - free slots for leaving the bike



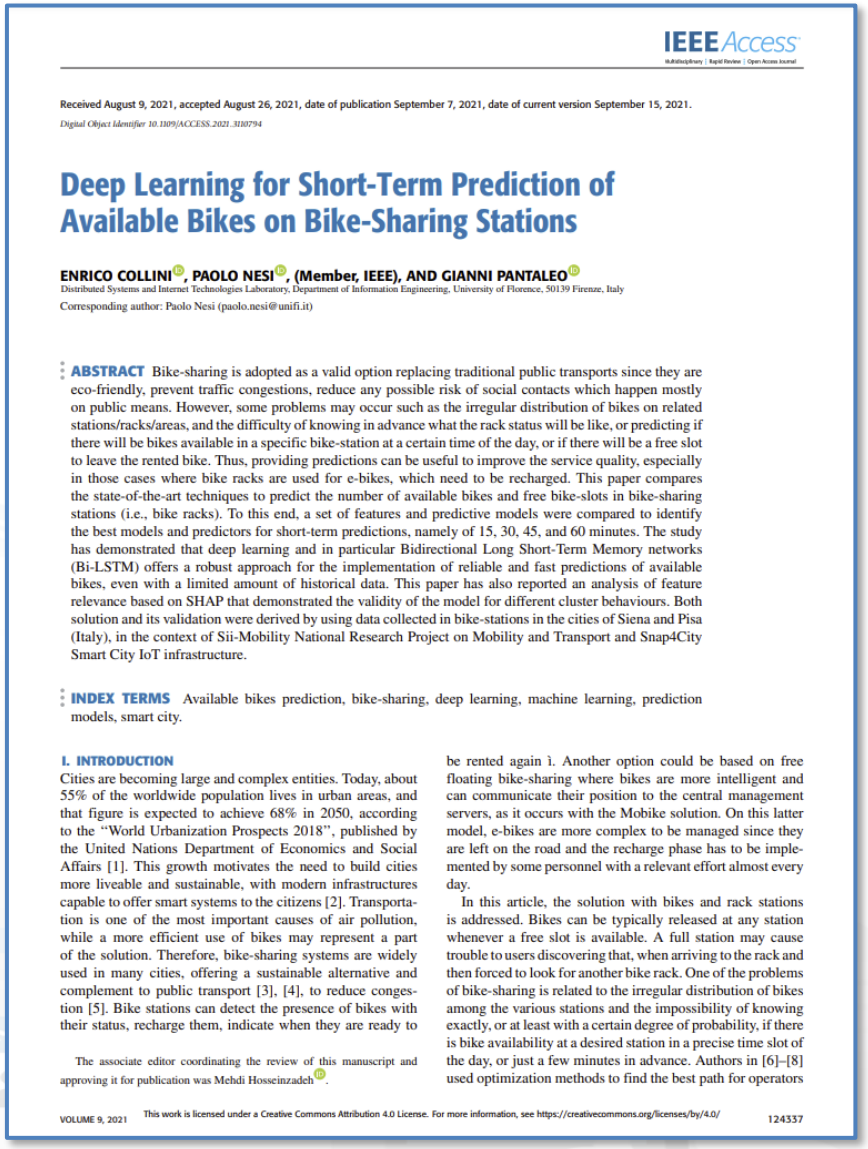






IEEE Access

- E. Collini, P. Nesi and G. Pantaleo, "Deep Learning for Short-Term Prediction of Available Bikes on Bike-Sharing Stations," in *IEEE Access*, vol. 9, pp. 124337-124347, 2021, doi: 10.1109/ACCESS.2021.3110794.
- https://ieeexplore.ieee.org/abstract/document/9530580



IEEE Access

Received August 9, 2021, accepted August 26, 2021, date of publication September 7, 2021, date of current version September 15, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3110794

Deep Learning for Short-Term Prediction of Available Bikes on Bike-Sharing Stations

ENRICO COLLINI, PAOLO NESI, (Member, IEEE), AND GIANNI PANTALEO

Distributed Systems and Internet Technologies Laboratory, Department of Information Engineering, University of Florence, 50139 Firenze, Italy

Corresponding author: Paolo Nesi (paolo.nesi@unifi.it)

ABSTRACT Bike-sharing is adopted as a valid option replacing traditional public transports since they are eco-friendly, prevent traffic congestions, reduce any possible risk of social contacts which happen mostly on public means. However, some problems may occur such as the irregular distribution of bikes on related stations/racks/areas, and the difficulty of knowing in advance what the rack status will be like, or predicting if there will be bikes available in a specific bike-station at a certain time of the day, or if there will be a free slot to leave the rented bike. Thus, providing predictions can be useful to improve the service quality, especially in those cases where bike racks are used for e-bikes, which need to be recharged. This paper compares the state-of-the-art techniques to predict the number of available bikes and free bike-slots in bike-sharing stations (i.e., bike racks). To this end, a set of features and predictive models were compared to identify the best models and predictors for short-term predictions, namely of 15, 30, 45, and 60 minutes. The study has demonstrated that deep learning and in particular Bidirectional Long Short-Term Memory networks (Bi-LSTM) offers a robust approach for the implementation of reliable and fast predictions of available bikes, even with a limited amount of historical data. This paper has also reported an analysis of feature relevance based on SHAP that demonstrated the validity of the model for different cluster behaviours. Both solution and its validation were derived by using data collected in bike-stations in the cities of Siena and Pisa (Italy), in the context of Sii-Mobility National Research Project on Mobility and Transport and Snap4City Smart City IoT infrastructure.

INDEX TERMS Available bikes prediction, bike-sharing, deep learning, machine learning, prediction models, smart city.

I. INTRODUCTION

Cities are becoming large and complex entities. Today, about 55% of the worldwide population lives in urban areas, and that figure is expected to achieve 68% in 2050, according to the "World Urbanization Prospects 2018", published by the United Nations Department of Economics and Social Affairs [1]. This growth motivates the need to build cities more liveable and sustainable, with modern infrastructures capable to offer smart systems to the citizens [2]. Transportation is one of the most important causes of air pollution, while a more efficient use of bikes may represent a part of the solution. Therefore, bike-sharing systems are widely used in many cities, offering a sustainable alternative and complement to public transport [3], [4], to reduce congestion [5]. Bike stations can detect the presence of bikes with their status, recharge them, indicate when they are ready to be rented again ì. Another option could be based on free floating bike-sharing where bikes are more intelligent and can communicate their position to the central management servers, as it occurs with the Mobike solution. On this latter model, e-bikes are more complex to be managed since they are left on the road and the recharge phase has to be implemented by some personnel with a relevant effort almost every day.

In this article, the solution with bikes and rack stations is addressed. Bikes can be typically released at any station whenever a free slot is available. A full station may cause trouble to users discovering that, when arriving to the rack and then forced to look for another bike rack. One of the problems of bike-sharing is related to the irregular distribution of bikes among the various stations and the impossibility of knowing exactly, or at least with a certain degree of probability, if there is bike availability at a desired station in a precise time slot of the day, or just a few minutes in advance. Authors in [6]–[8] used optimization methods to find the best path for operators

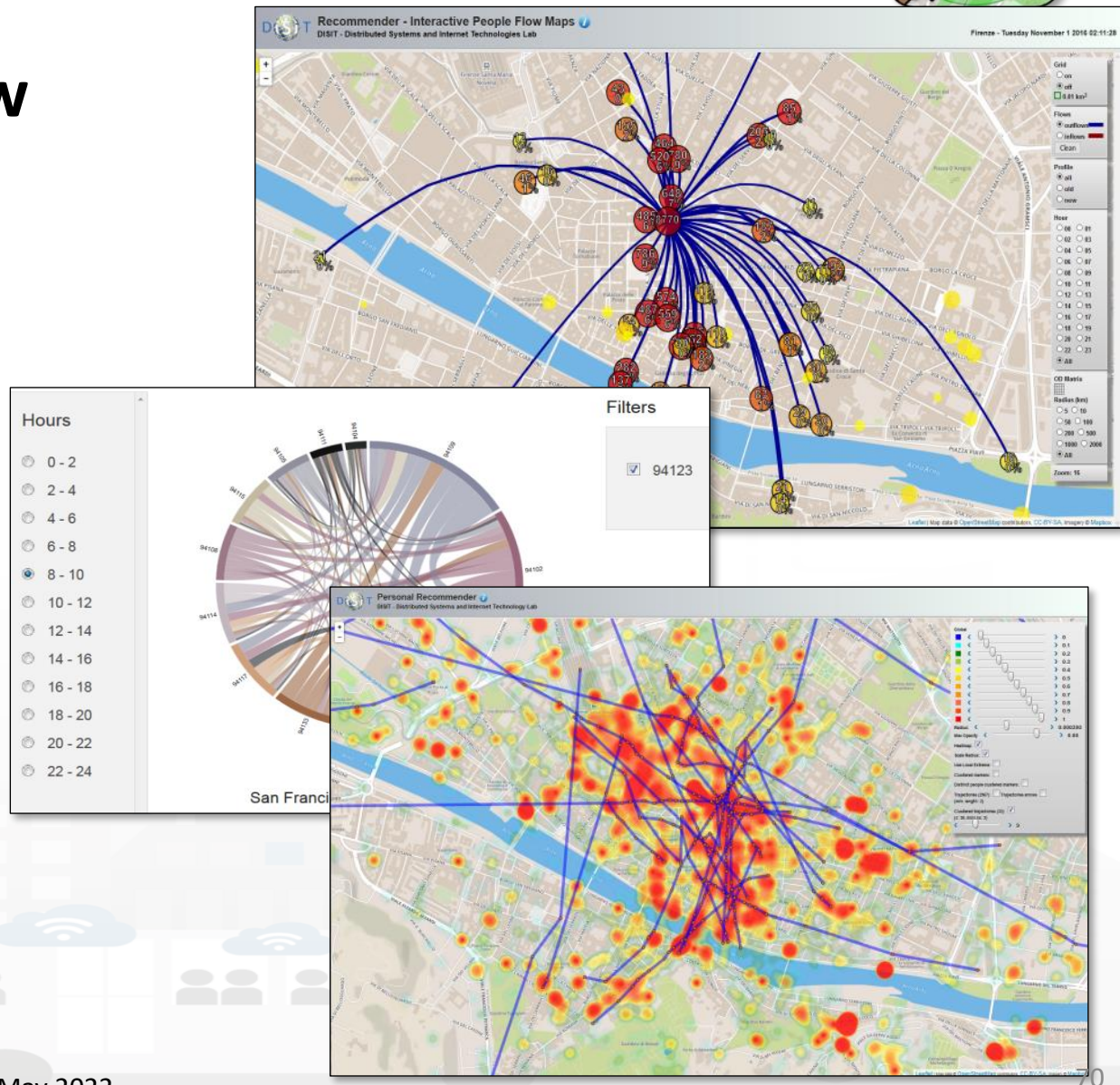
The associate editor coordinating the review of this manuscript and approving it for publication was Mehdi Hosseinzadeh.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

User Behaviour Analysis via Wi-Fi, OD Matrices

User Behaviour Analysis

- **Monitoring movements by traffic flow sensors**
  - Spires and virtual spires
- **Monitoring movements from Mobile Cells**
  - Unsuitable for precise tracking and OD production
- **Monitoring movements from Wi-Fi**
- **Monitoring movements** and much more from
  - mobile Apps
  - OBU, On Board Unit



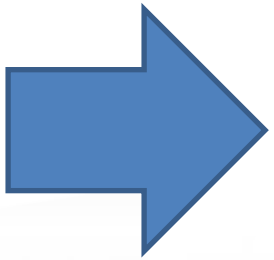
applications

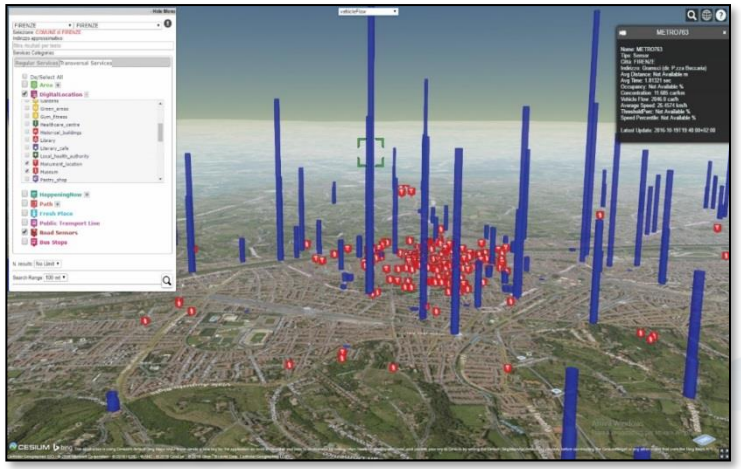
- **Data Results**
  - Hot places, trajectories, OD matrices
  - Typical presences, Typical trajectories
  - Predictions, anomaly detection, etc.
- **Emissions of management services**
  - City Plans and strategies
  - Optimized Routing: private, public services, ambulances
  - Optimization of public mobile services
  - Early warning: respond to emergences
- **Business solutions**
  - Suggestions and marketing
  - Shops placement

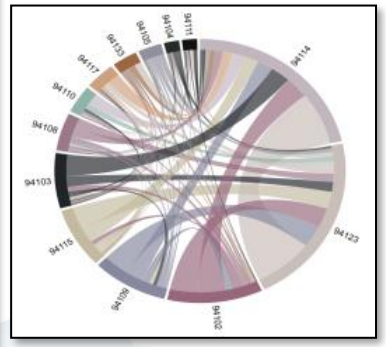
1007
GEST
1016
GEST

Predicting City users movements

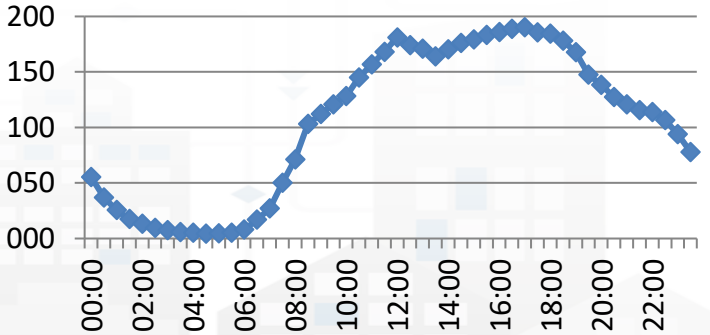
- **Issue:**
  - How they move: vehicles, pedestrian, bike, ferry, metro,
  - Where they go….
- **Impact:**
  - Tuning the services: cleaning, police, control, security
- **Several metrics related to**
  - Knowledge of the city
  - Monitoring traffic and people flow
  - …….





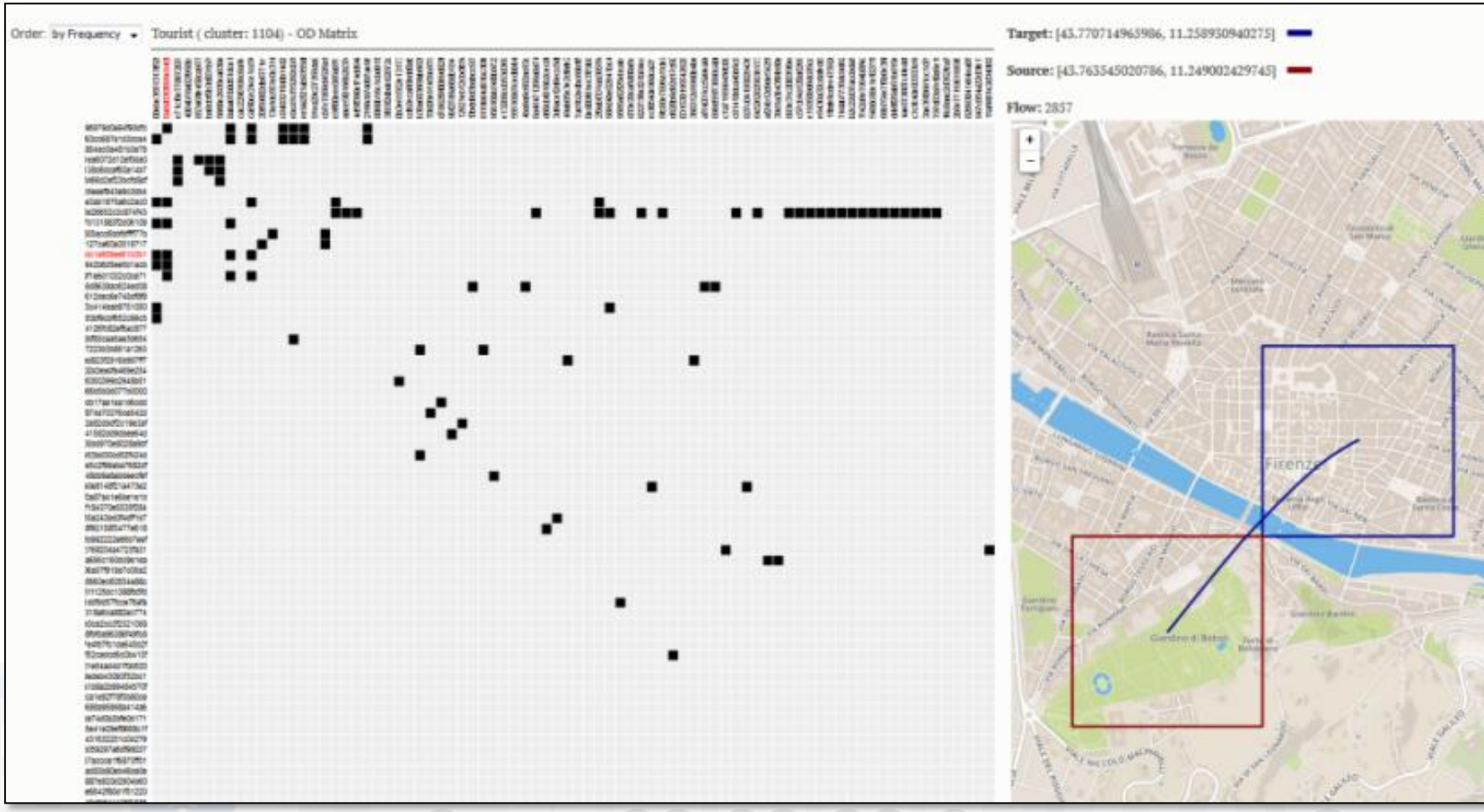


- Daily trends
- OD matrices
- Trajectories
- Prediction models

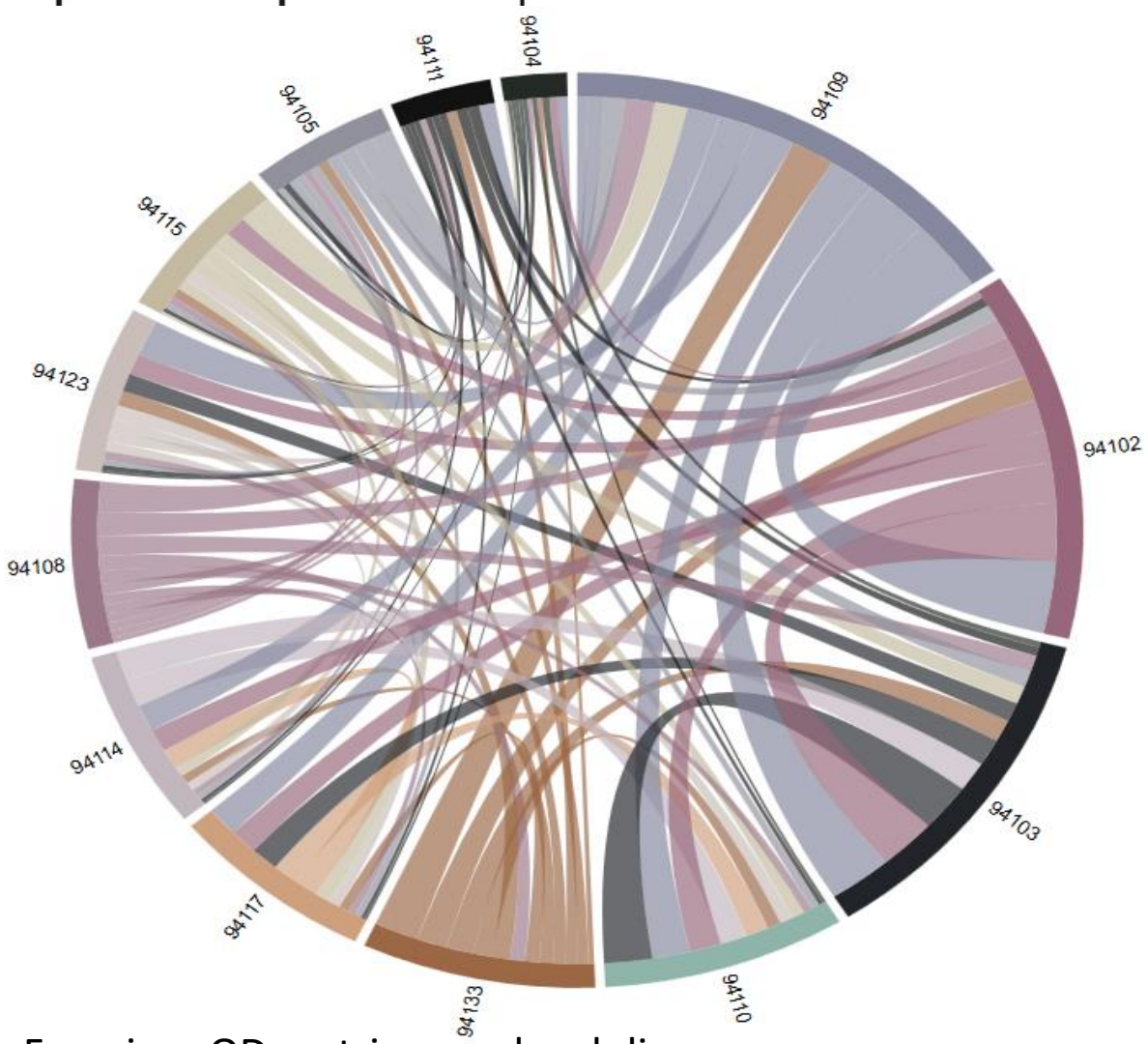


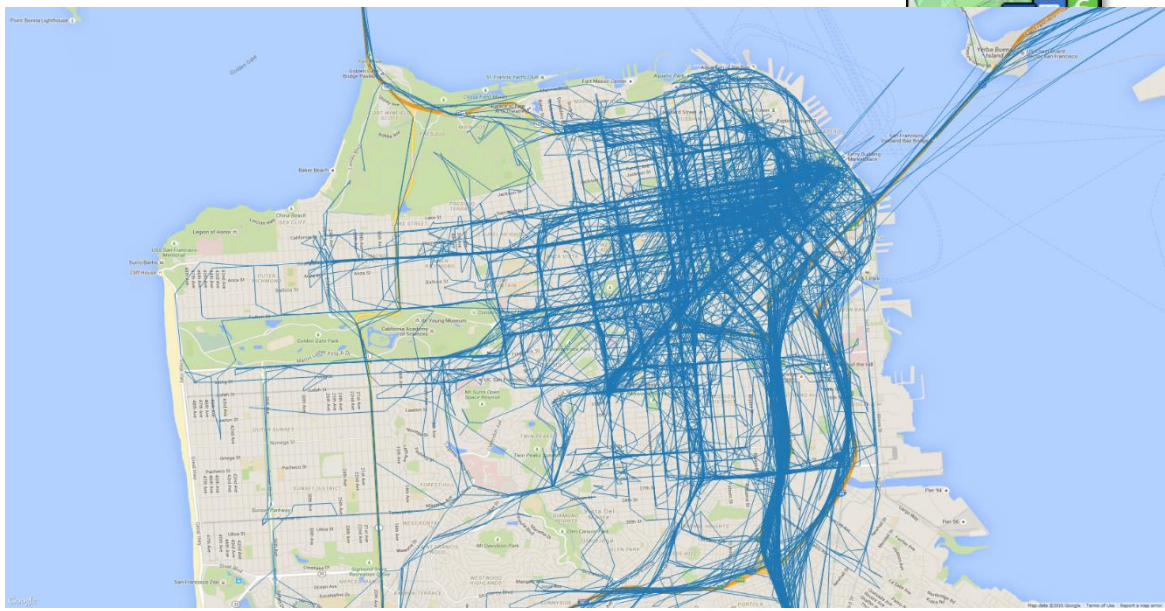
OD Matrices

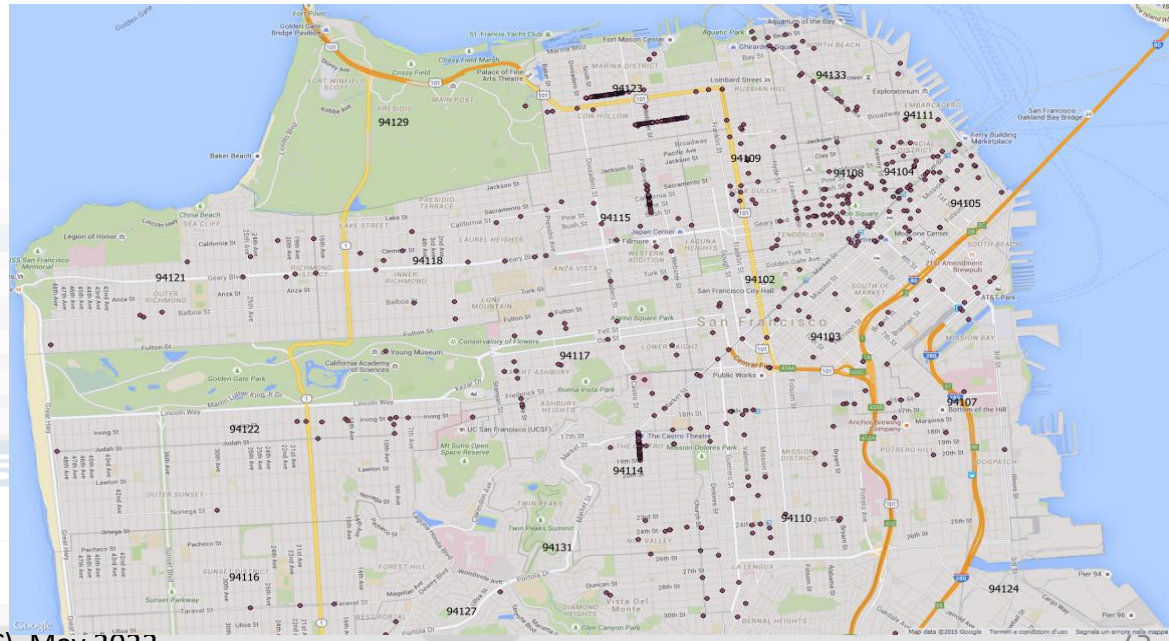
- Origins and destinations
  - Any area of the zone
  - From to
  - To from
- By temporal slice
- By inflow or outflow
- By user profile
- By motivations
- By trevel means
- By extraction technique
- By civic area VS segmented GPS area
- ....



San Francisco Taxi



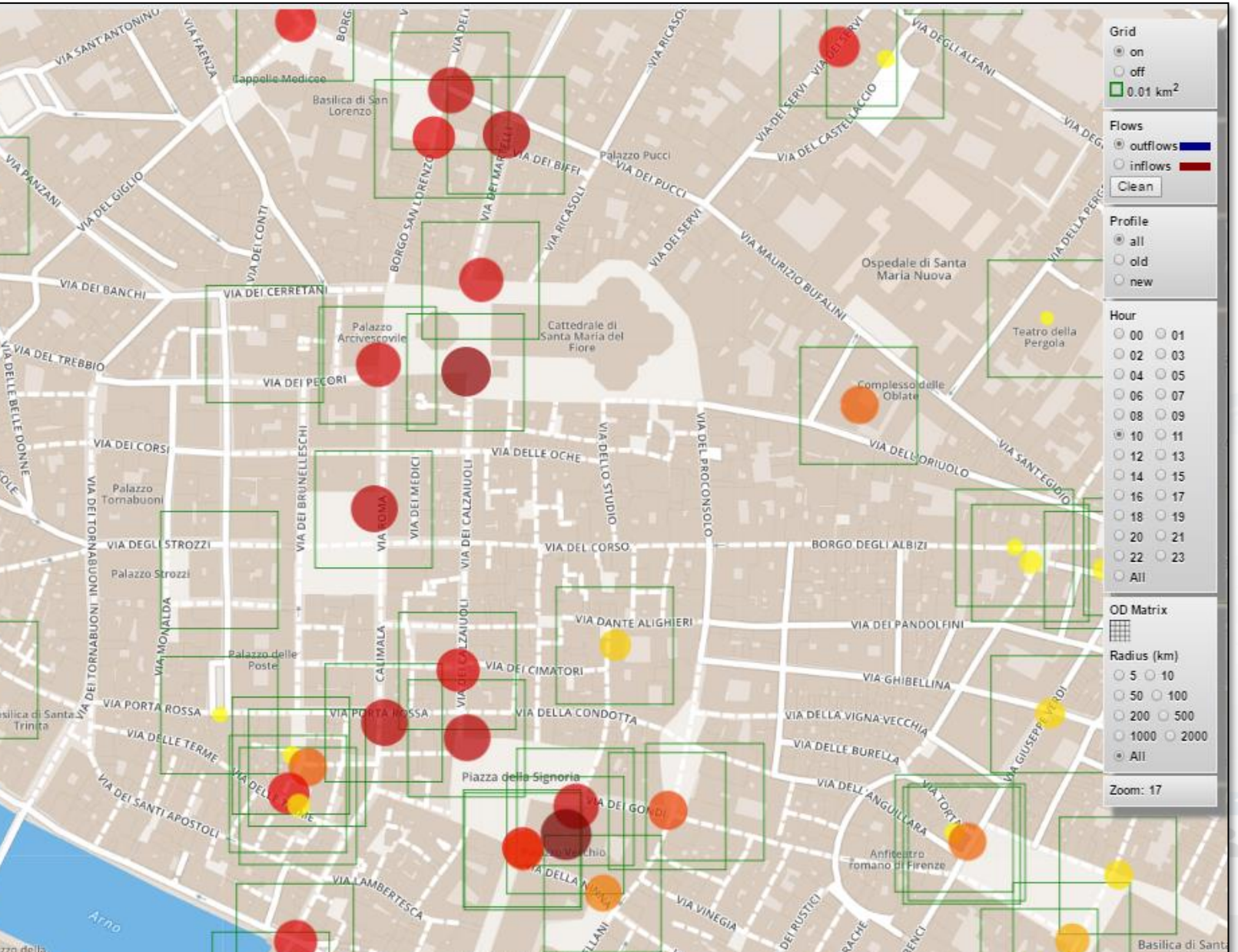




San Francisco OD matrix as a chord diagram

- 13 central ZIP areas of the city (real cab flows),
- ONLY on FLOW: from to

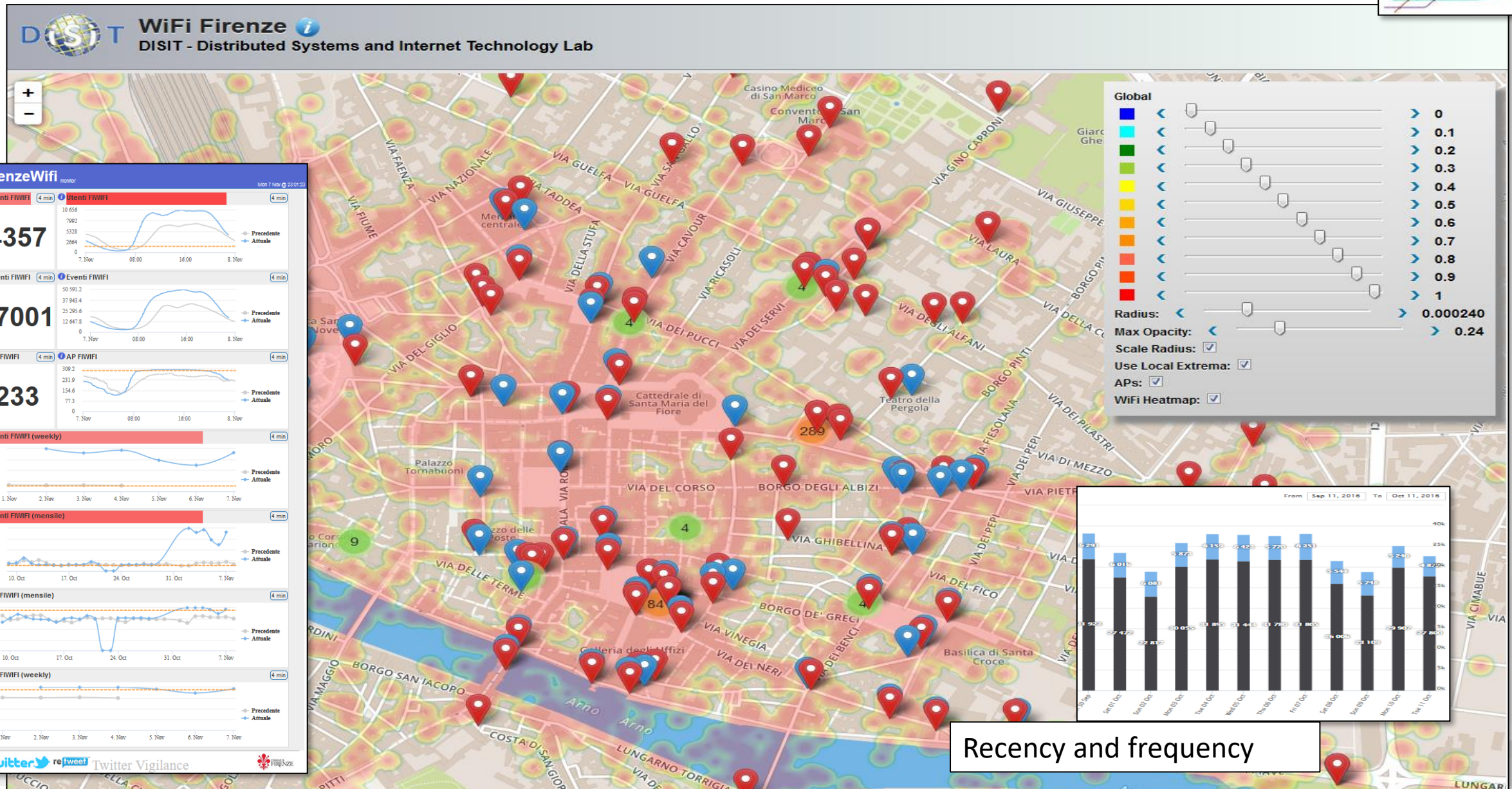
Firenze Wi-Fi



- AP data → heatmap sparse
  - Scattered detection area

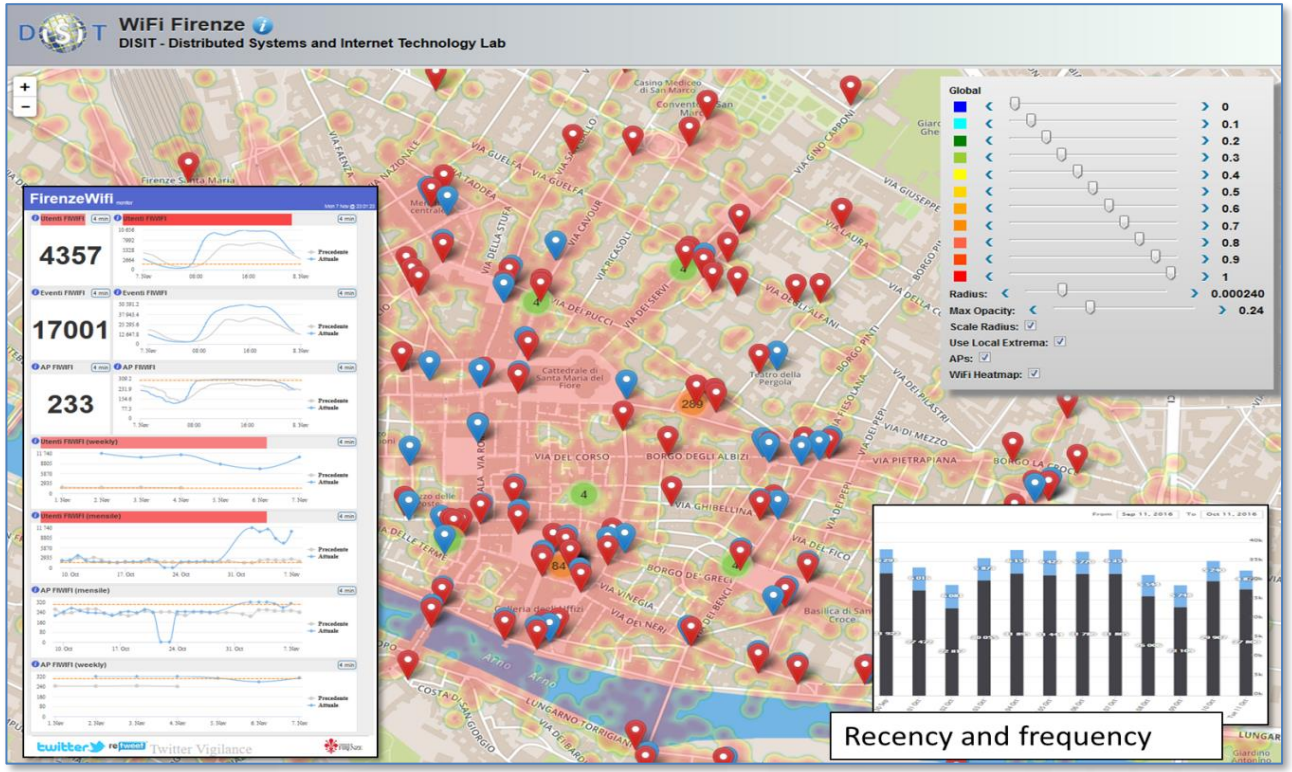
- Inflow/outflow
- New/Old users
- for time slot
- Undetected zones

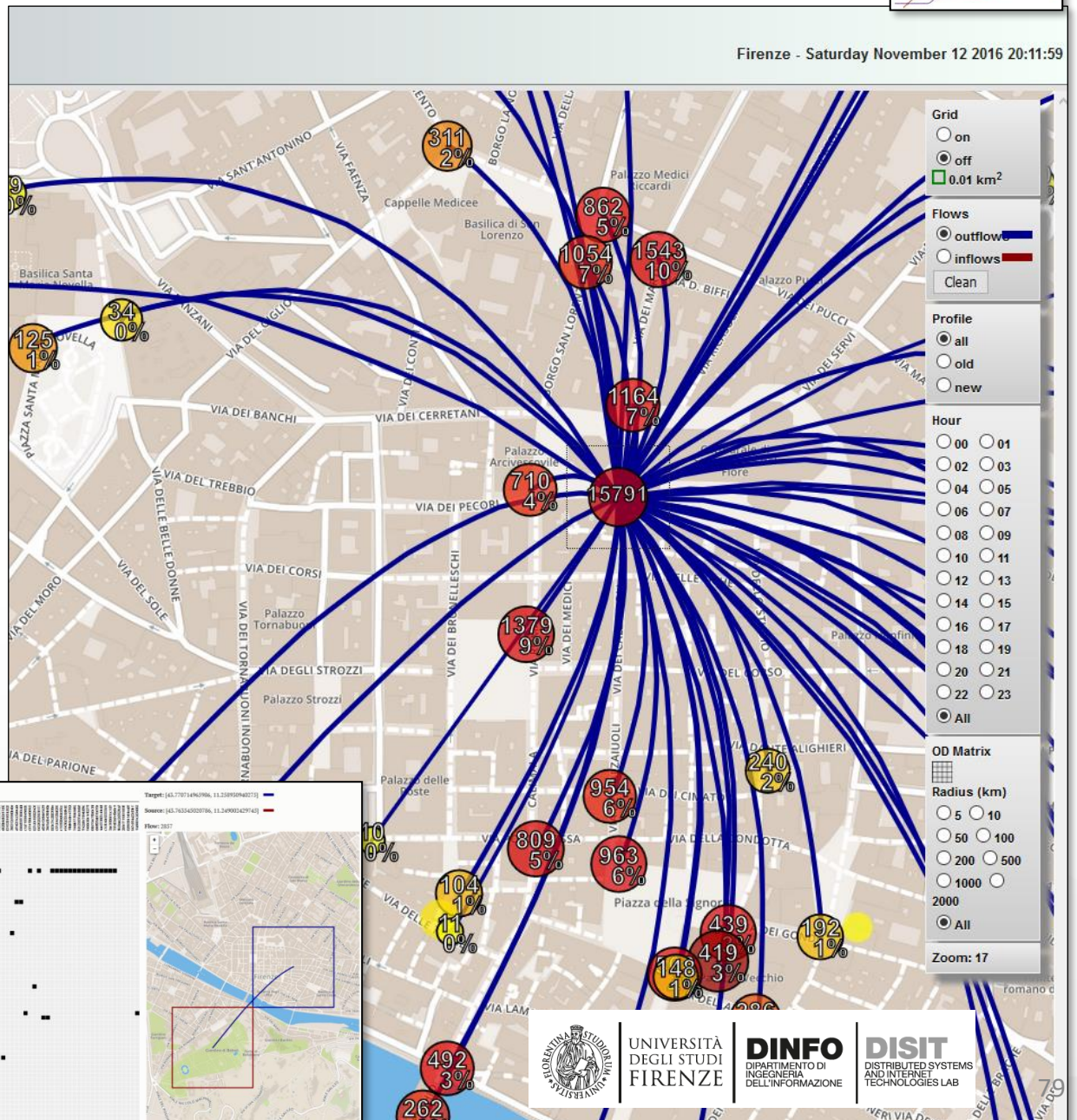
Wi-Fi Monitor Tool

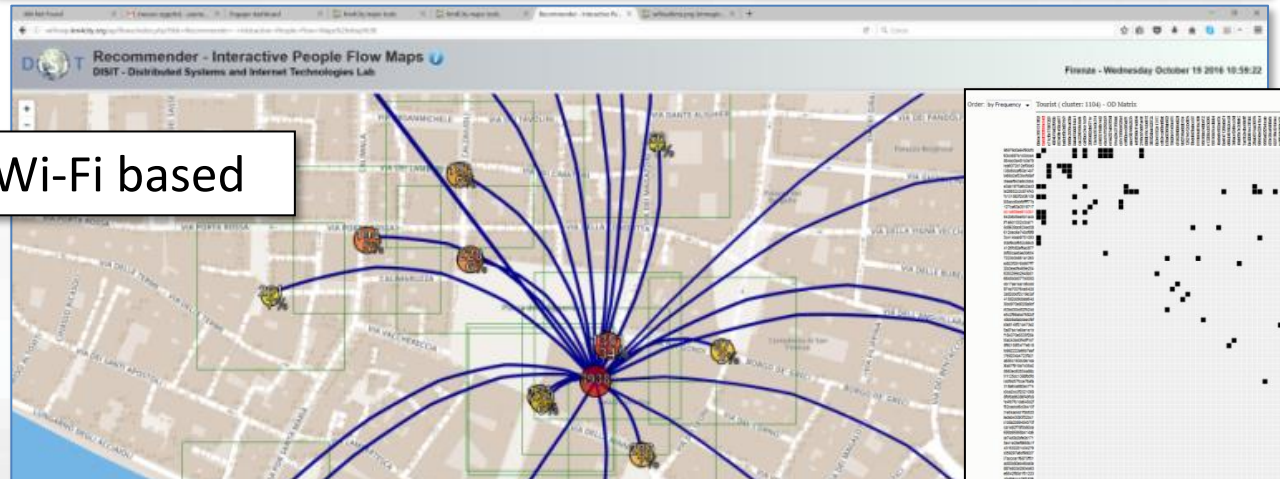


Recency and frequency

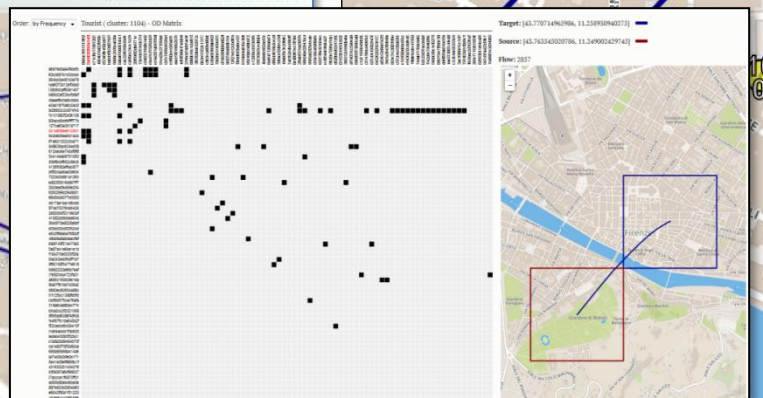
Origin Destination Matrix Estimation



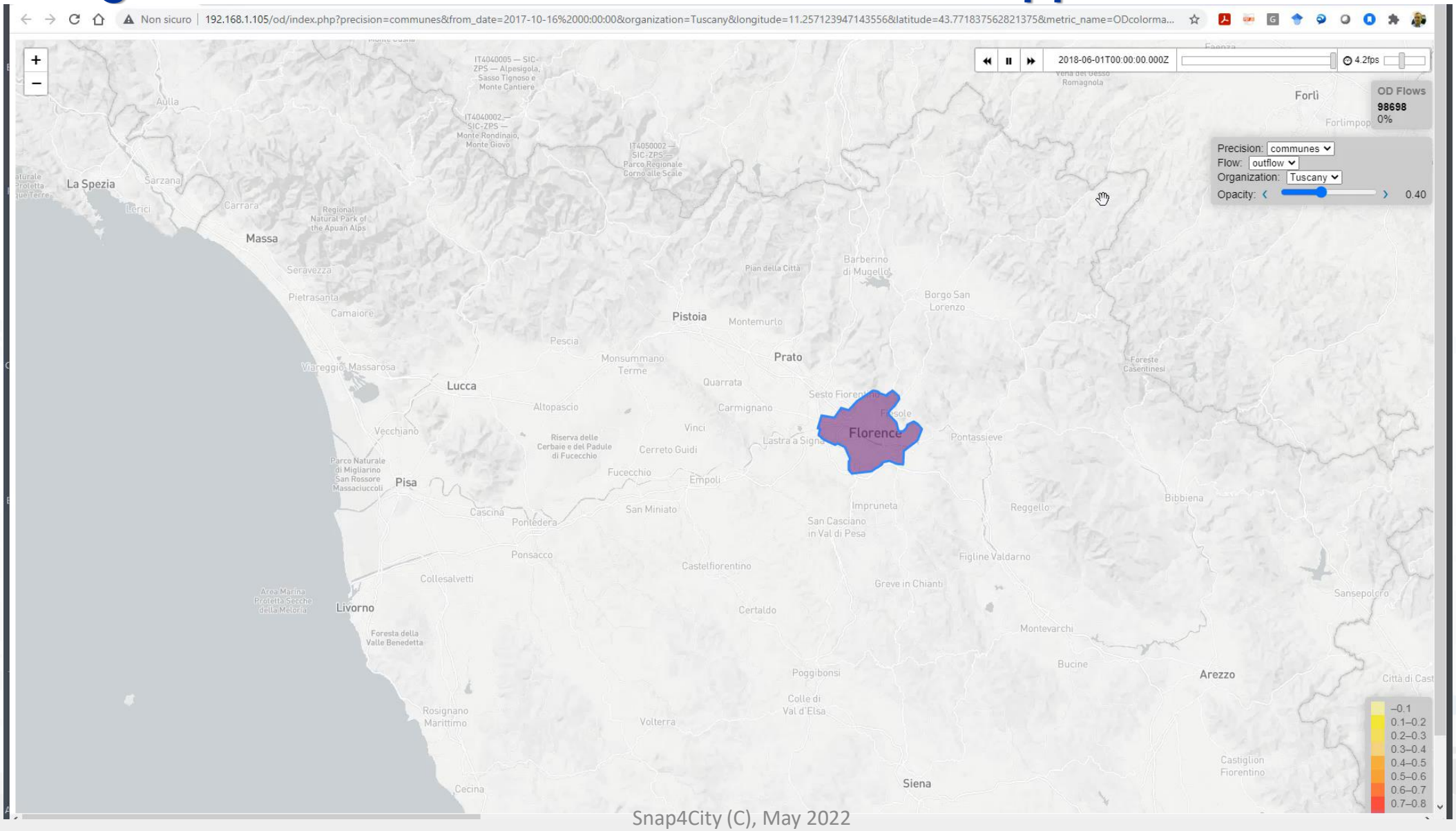




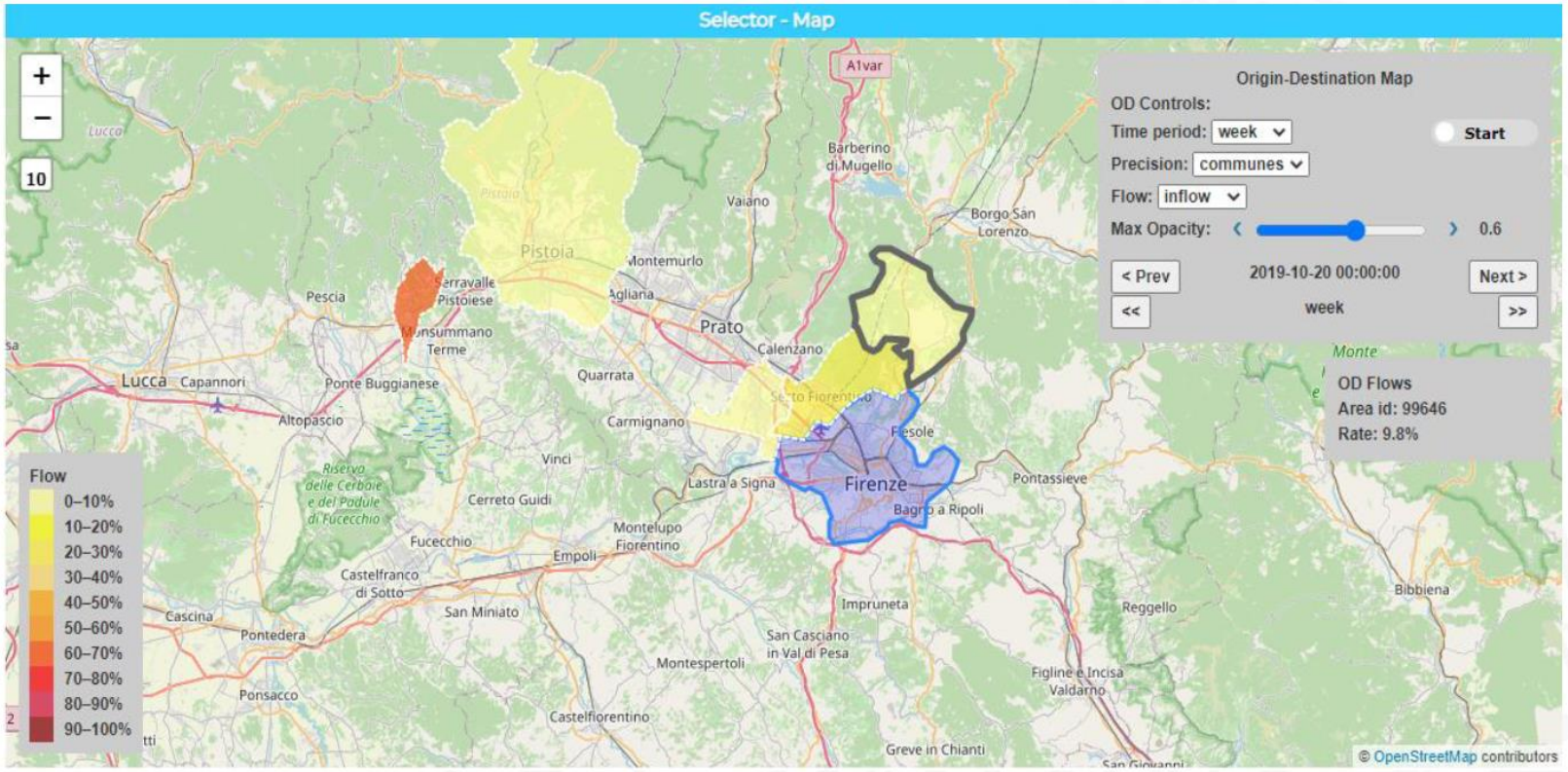
Wi-Fi based

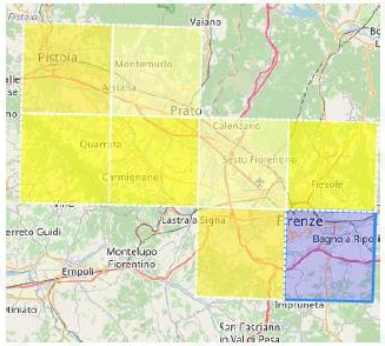


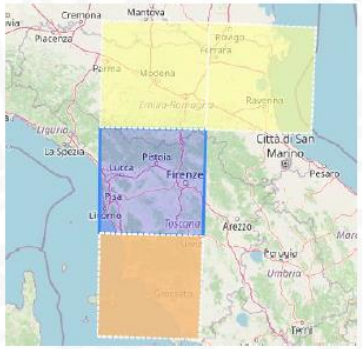
Origin Destination Matrix based on Mobile App data



Different Origin Destination Matrices





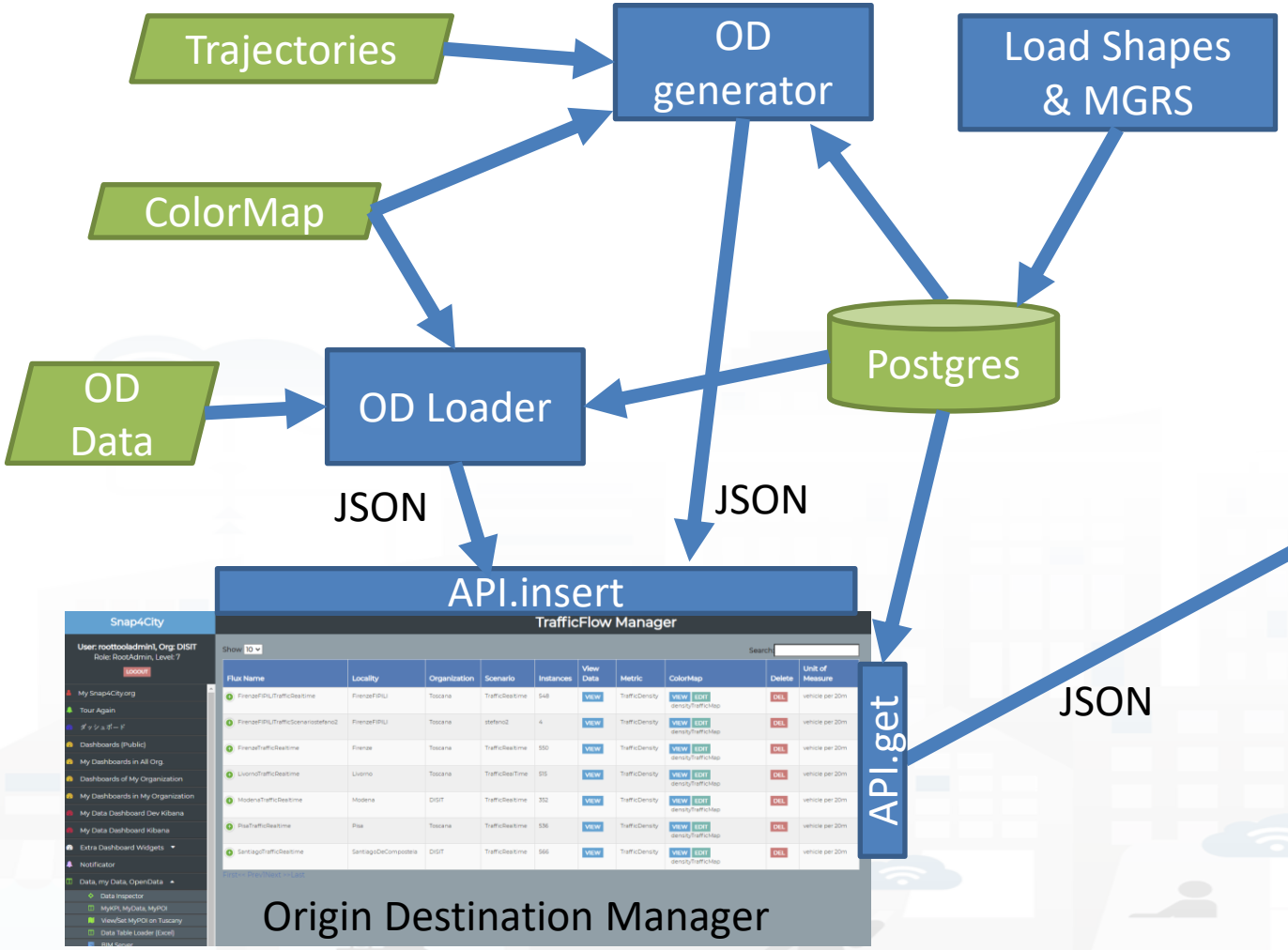


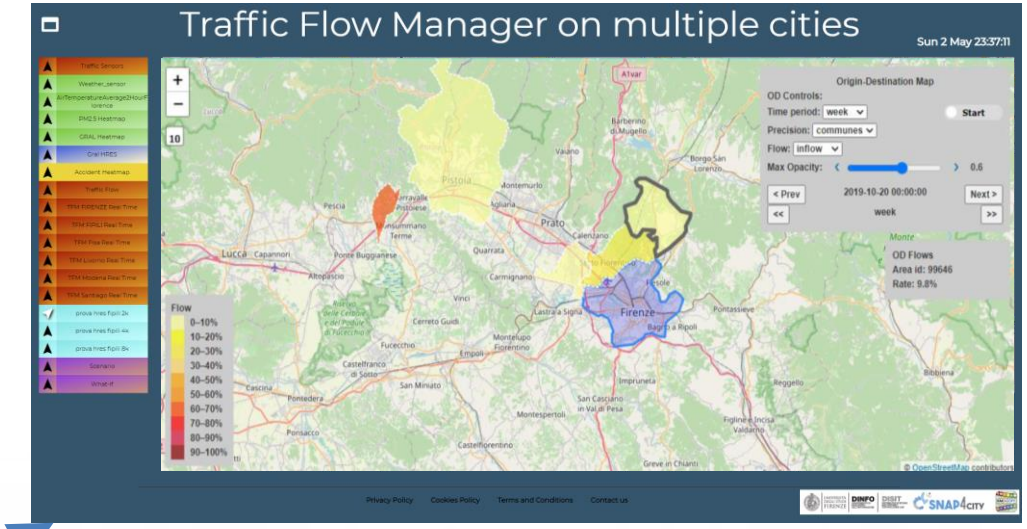
- **Get specific value**
- **Time window**
- **Opacity**
- **Animation**
- **Inflow/outflow**
- **Sequence of OD matrices: next/prev**

Geo shapes as:

- **Shapes**: city, region, territories, etc.
- **MGRS**: 1Km, 10Km, 100Km

How it works: Origin Destination Manager





Postgress loaded for

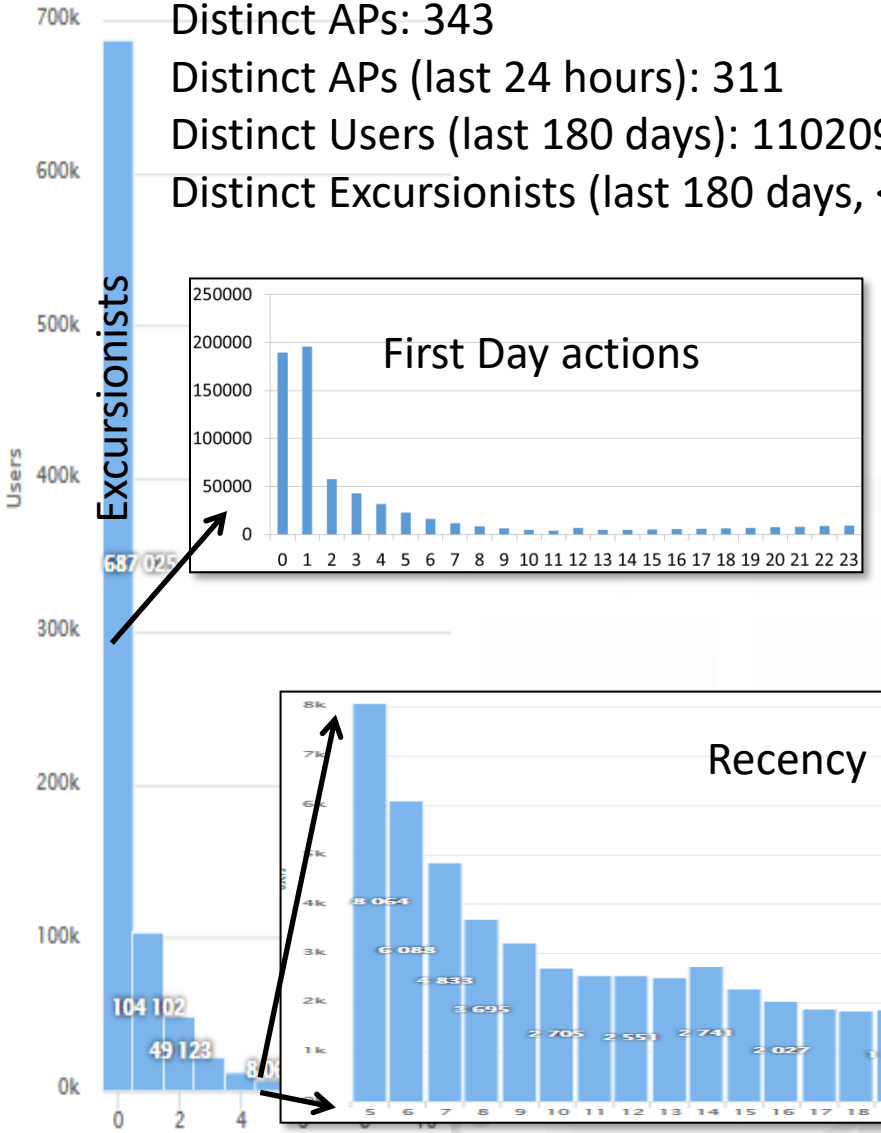
- **Shapes**: city, region, territories, etc.
  - GADM https://gadm.org/ , and ACE
- **Squared** MGRS:
  - 1m, 10m, 100m, **1Km, 10Km, 100Km**

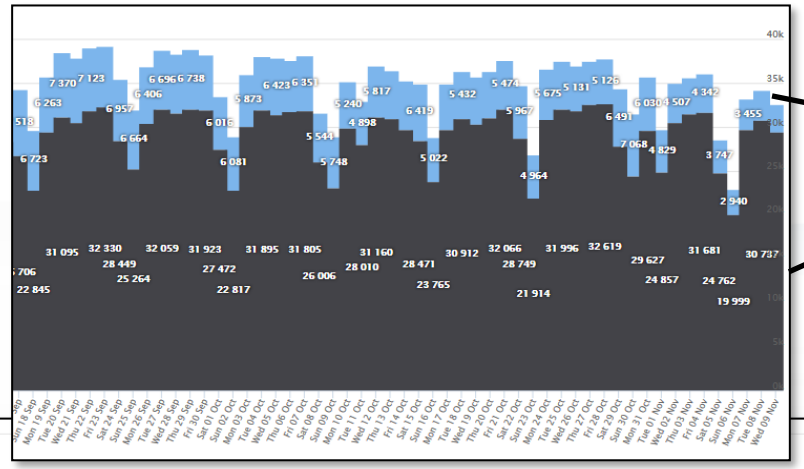
User Behaviour Analysis

Distinct APs: 343
Distinct APs (last 24 hours): 311
Distinct Users (last 180 days): 1102098
Distinct Excursionists (last 180 days, < 24 h): 687025

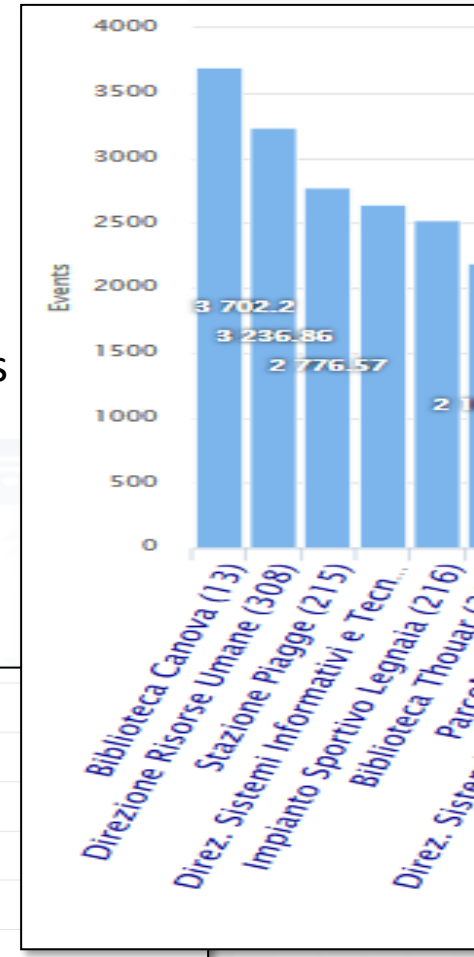
About the 15% of the peoples passing by

Where

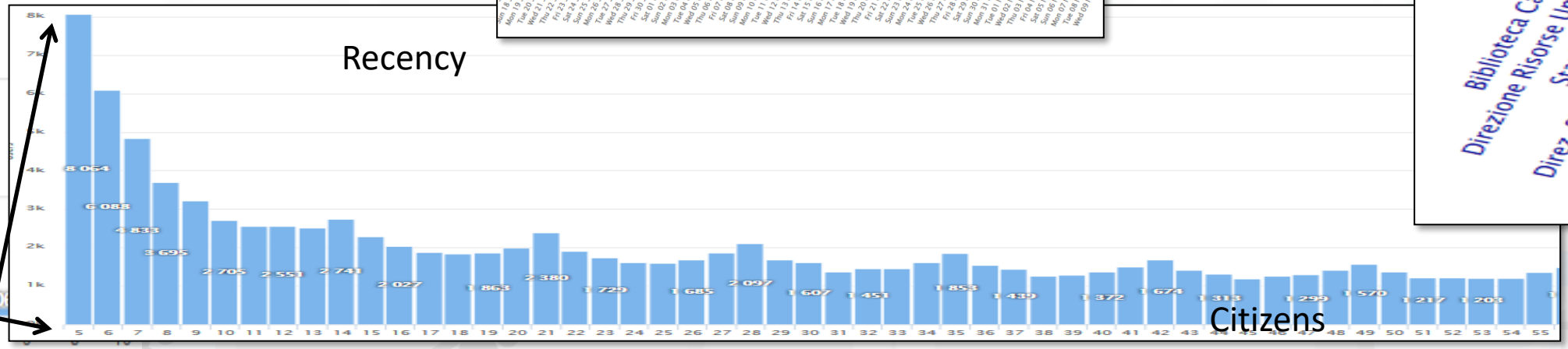




New City Users
VS
Returning



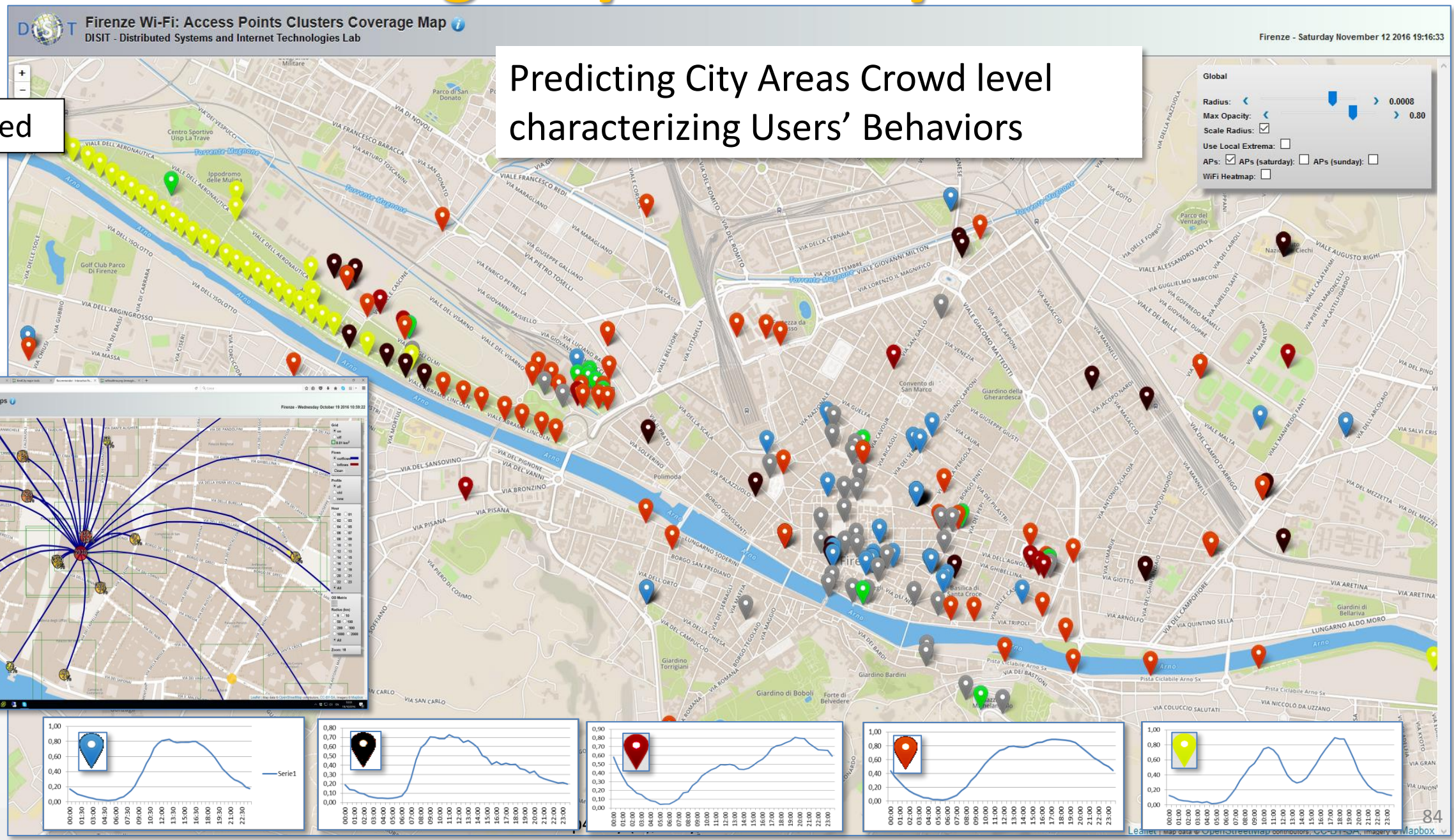
Recency

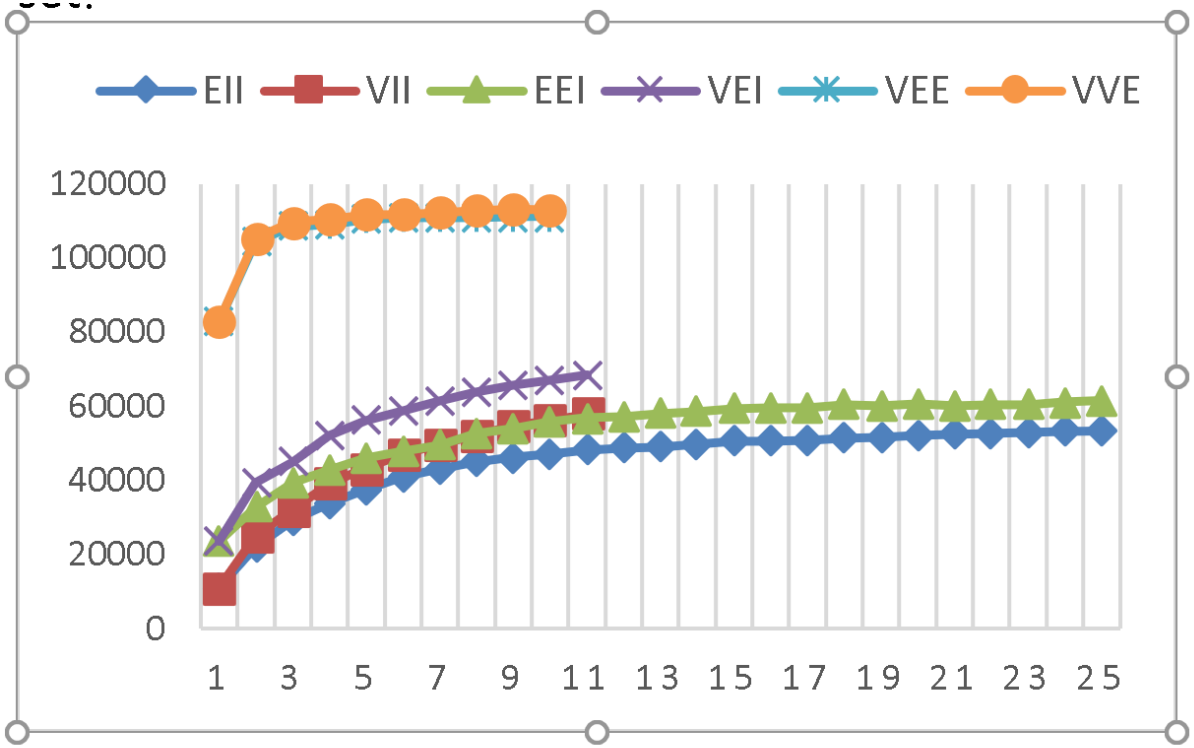


Characterizing City Areas by User Behavior

Wi-Fi based

Predicting City Areas Crowd level characterizing Users' Behaviors





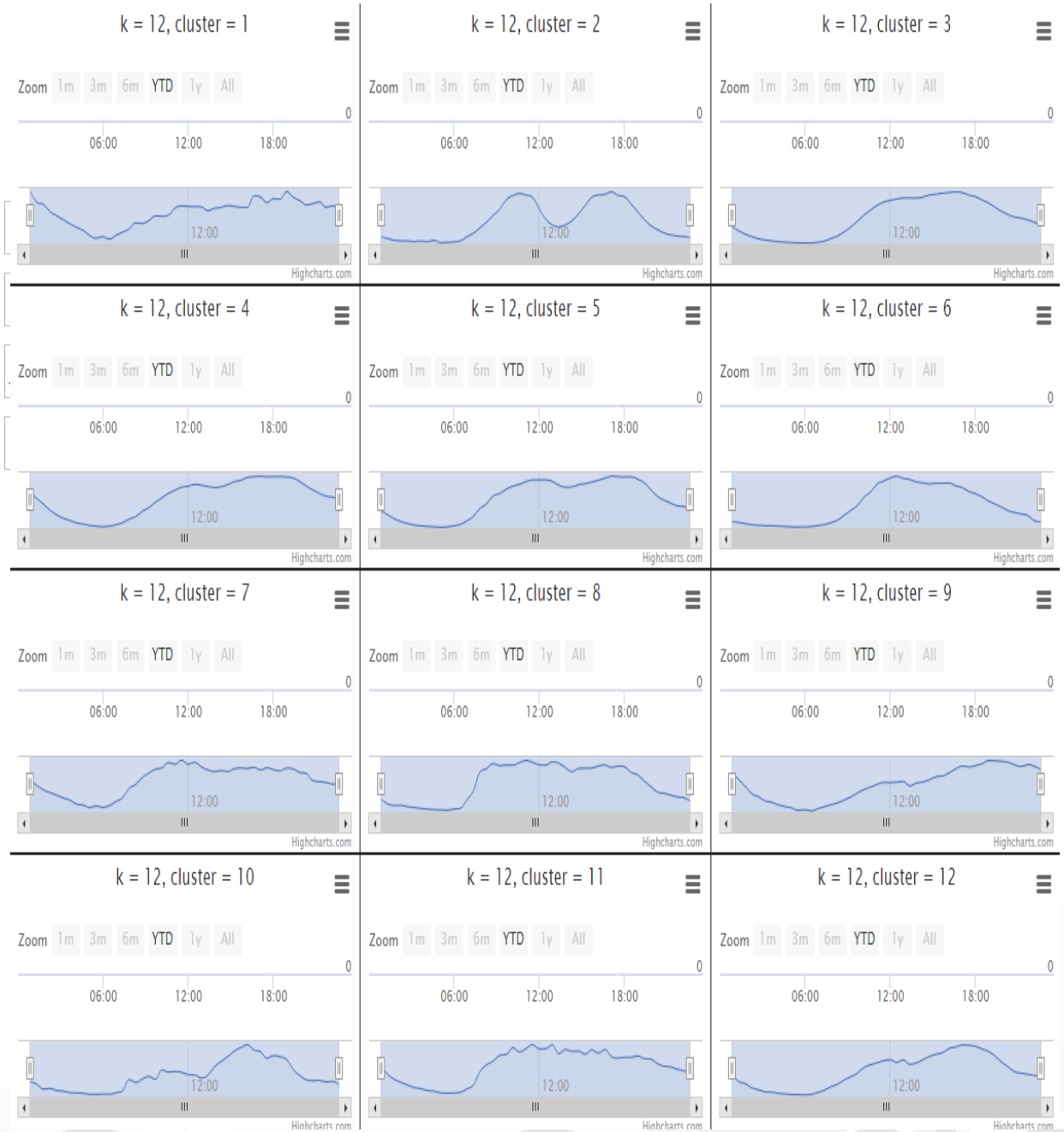
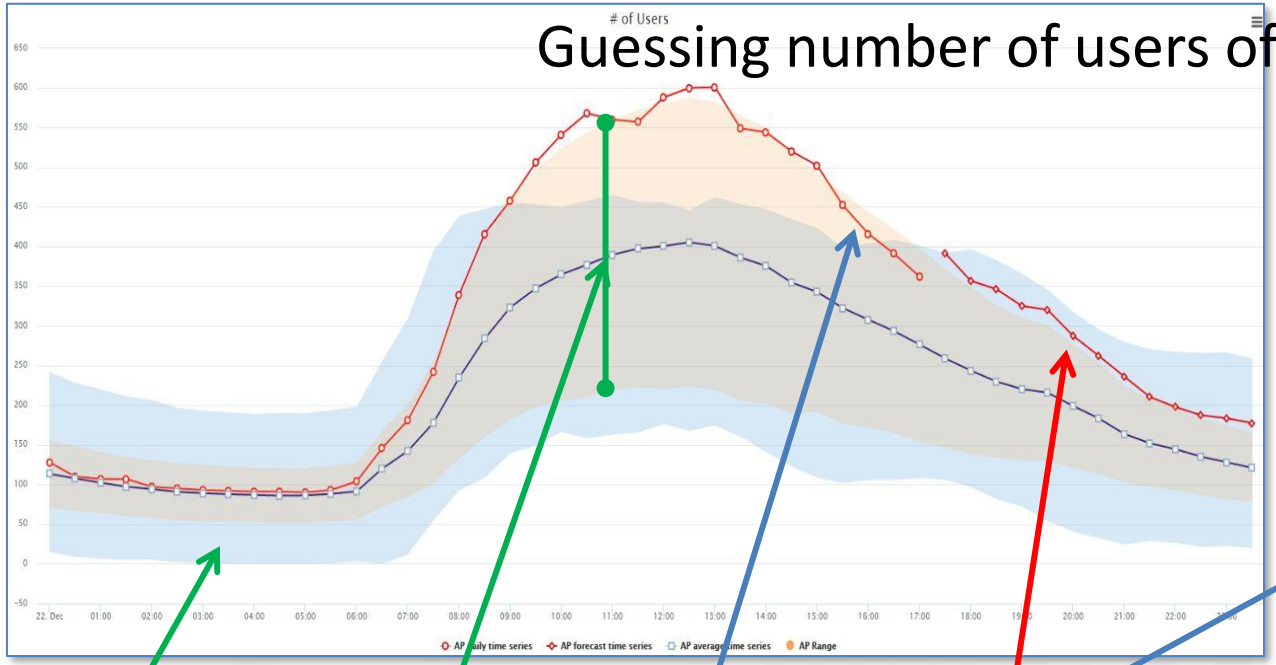
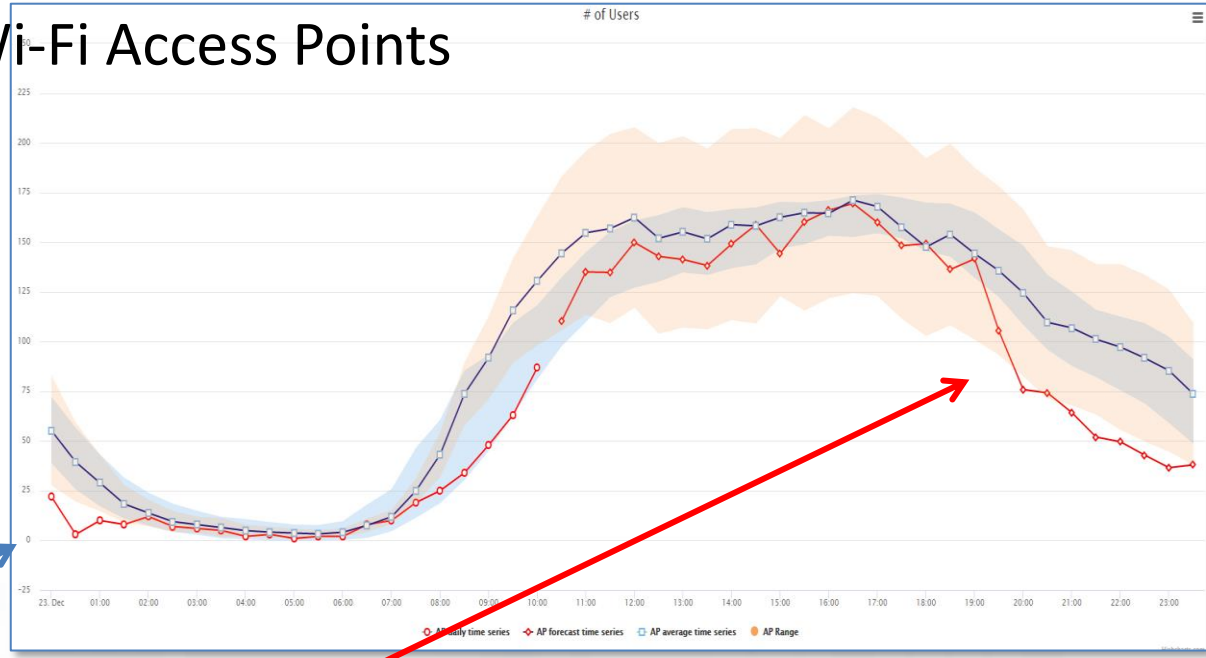


Figure 14 – Average BIC for mixture models vs K number of cluster, higher values are better, the curves are truncated at the best value for K they found.

Prediction and Identification of Anomalies

Guessing number of users of Wi-Fi Access Points





Cluster confidence

AP average and confidence

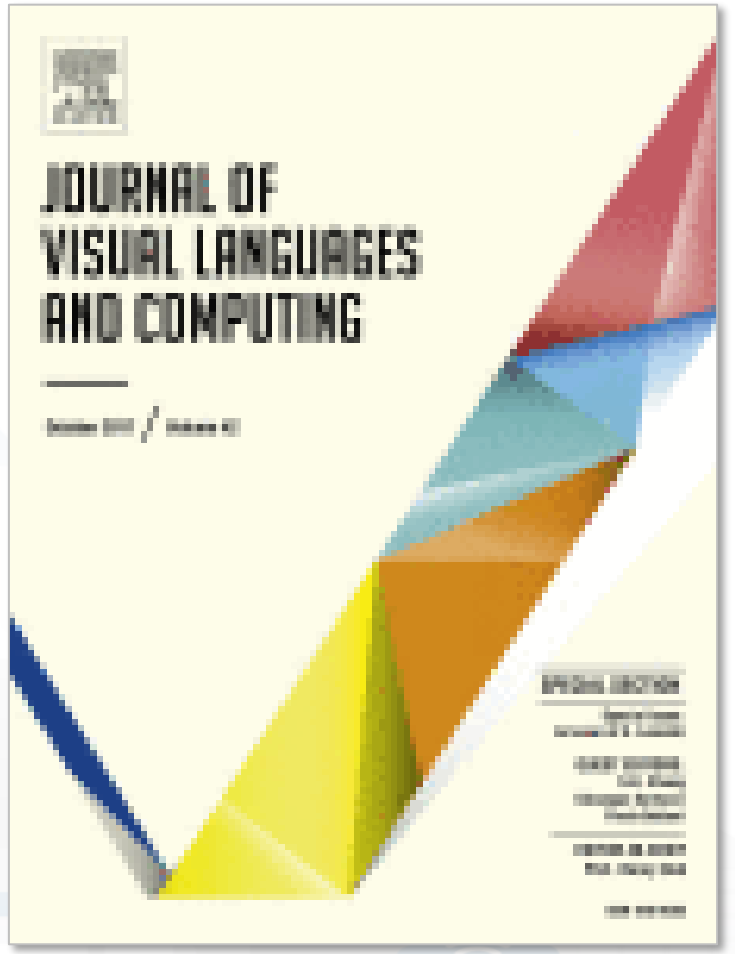
Actual AP trend for today

AP prediction for the next time slot in the day on the basis of past weeks

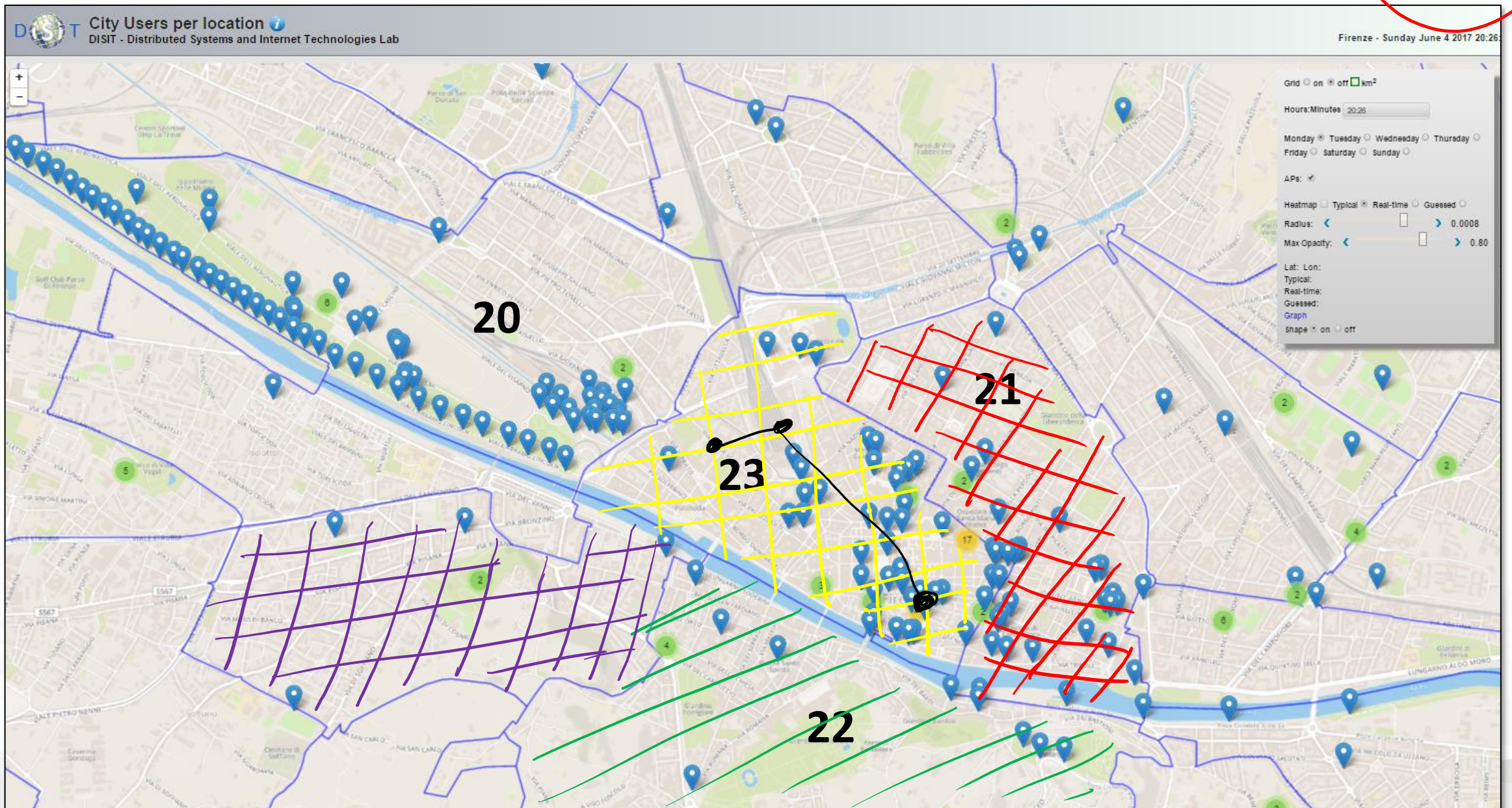
Predictive precision of the 95%

User Behaviour Analysis

- P. Bellini, D. Cenni, P. Nesi, I. Paoli, "Wi-Fi Based City Users' Behaviour Analysis for Smart City", Journal of Visual Language and Computing, Elsevier, 2017. http://www.sciencedirect.com/science/article/pii/S1045926X17300083



Firenze Wi-Fi vs ACE



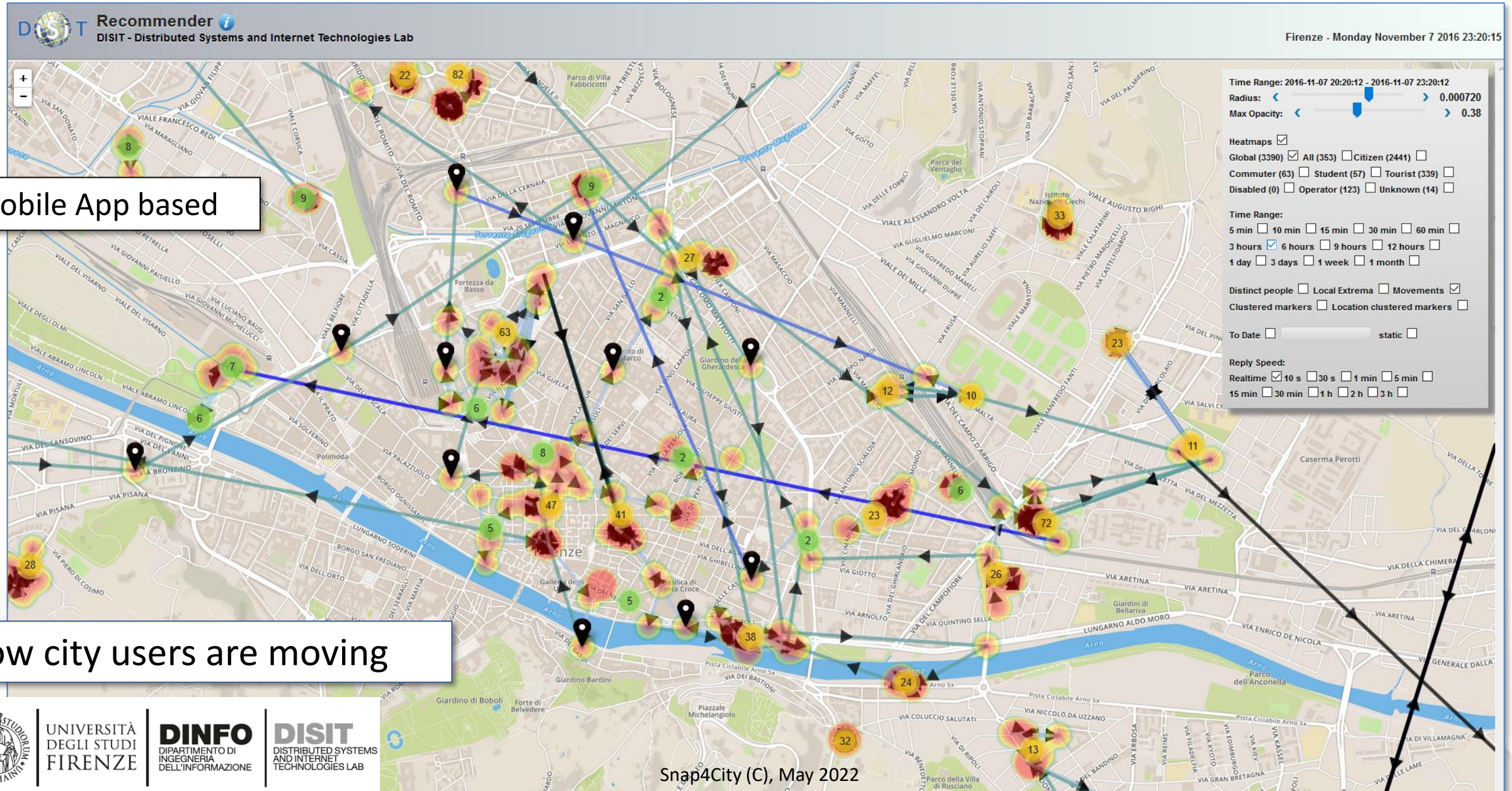
User Behaviour Analysis via Trajectories

Anonymous User Behavior Analysis

Case Study E

Sii-Mobility

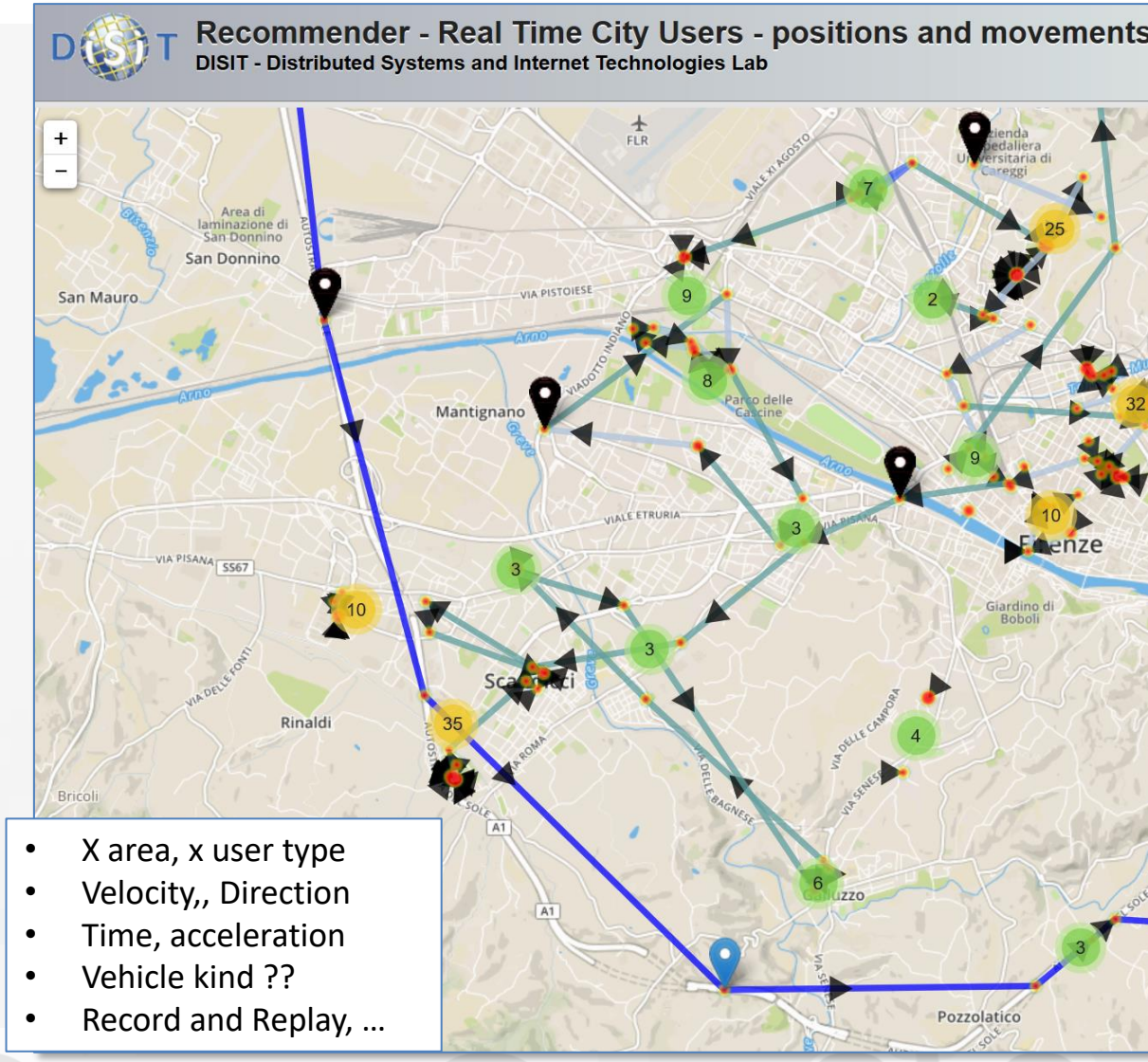
Mobile App based



How city users are moving

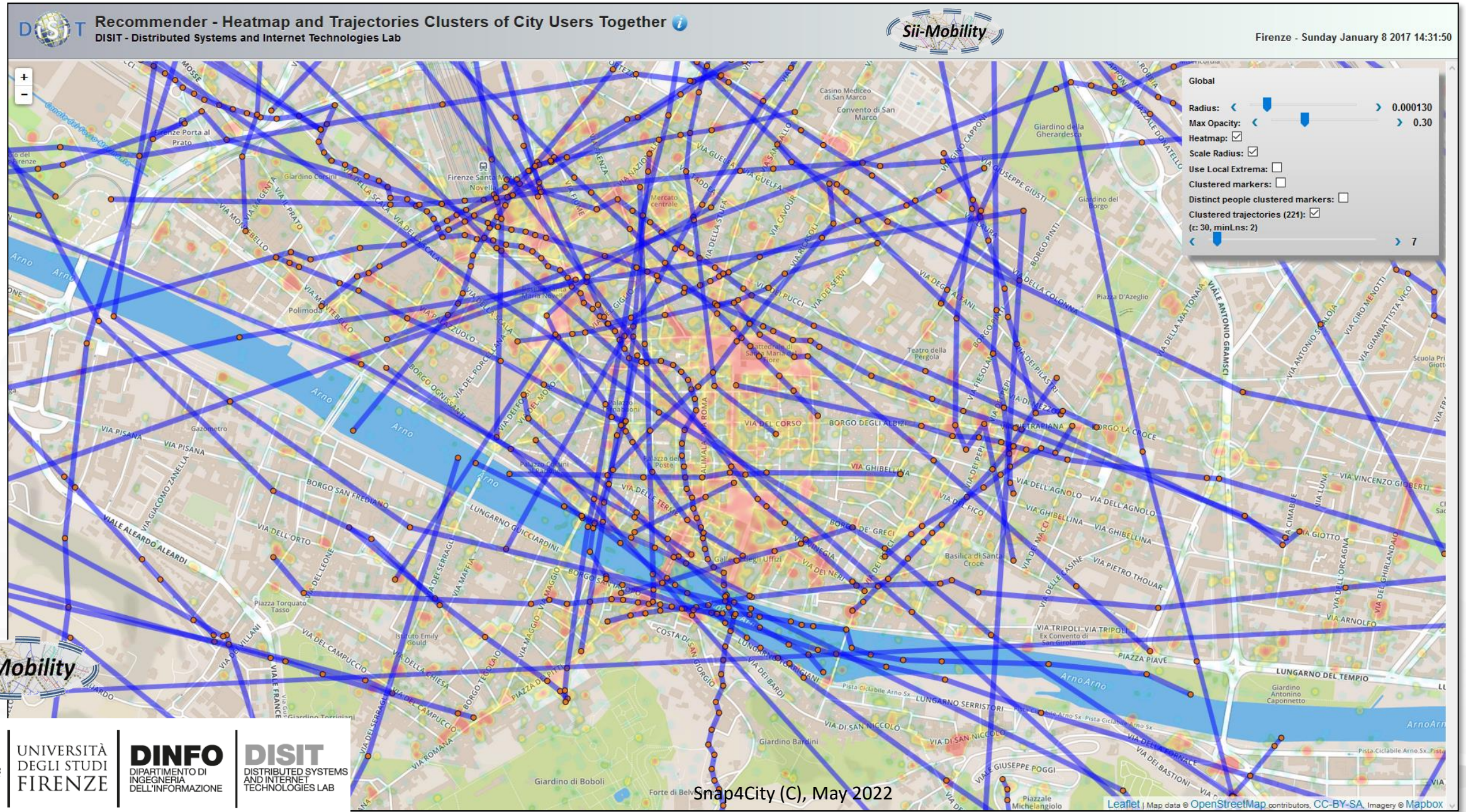
People flow Trajectories from Apps

- **From mobile app:**
  - Resolving GPS location: GPS, cells, wifi-network, ..mixt
  - Noisy, different kind of devices, ..
  - Smart algorithm on devices for location acquisition
  - Anonymized data, terms of use on mobile
- **Issues and Filtering**
  - Gps Accuracy, kind of measure (GPS, mixt)
  - Jump in time, space, velocity
  - General noise (diff. devices)
  - Knowledge of precision map
- **Clustering**: time, space, user kind, etc.



- X area, x user type
- Velocity,, Direction
- Time, acceleration
- Vehicle kind ??
- Record and Replay, …

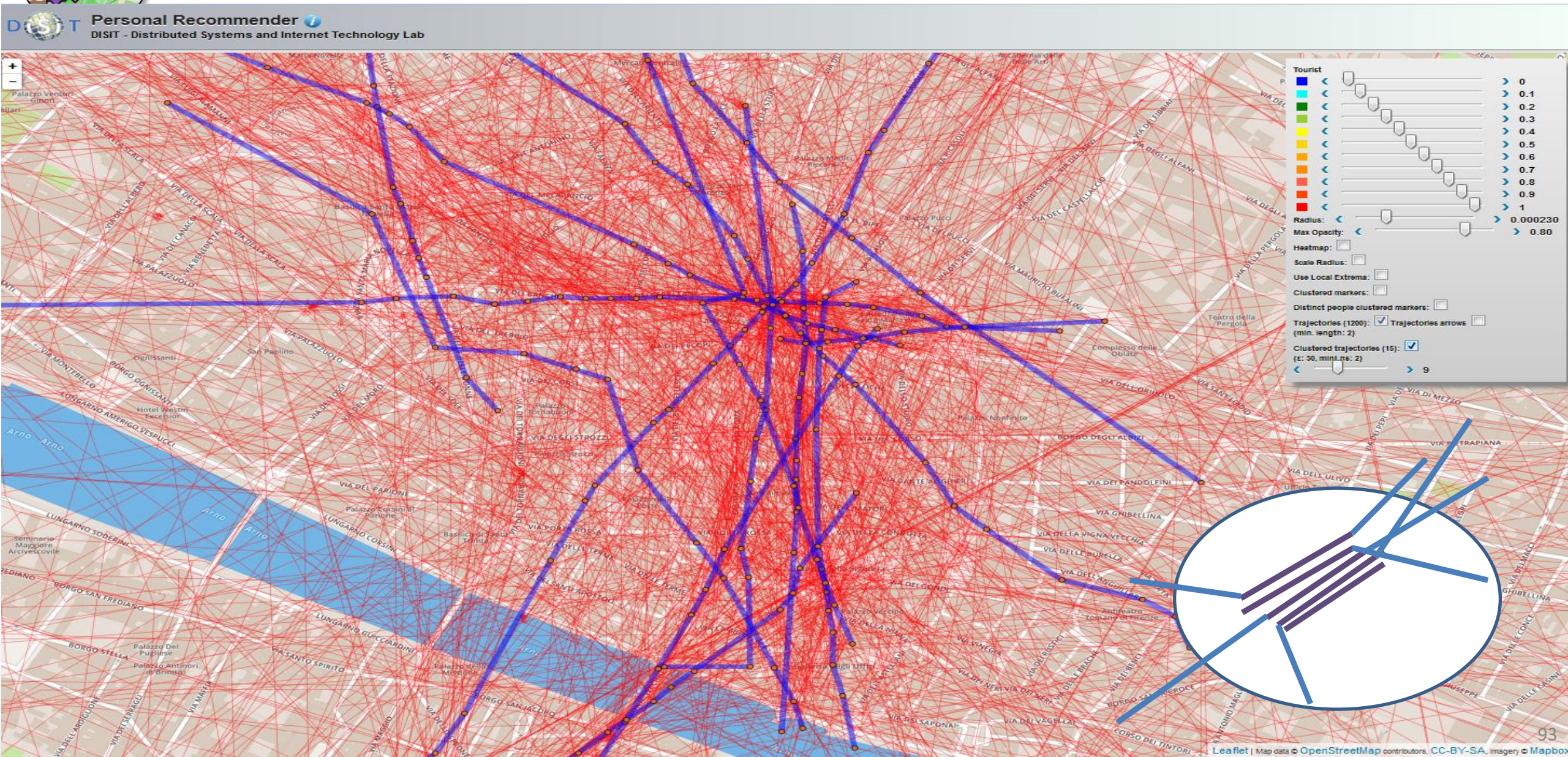
Heat Map from Mobile: users as sensors



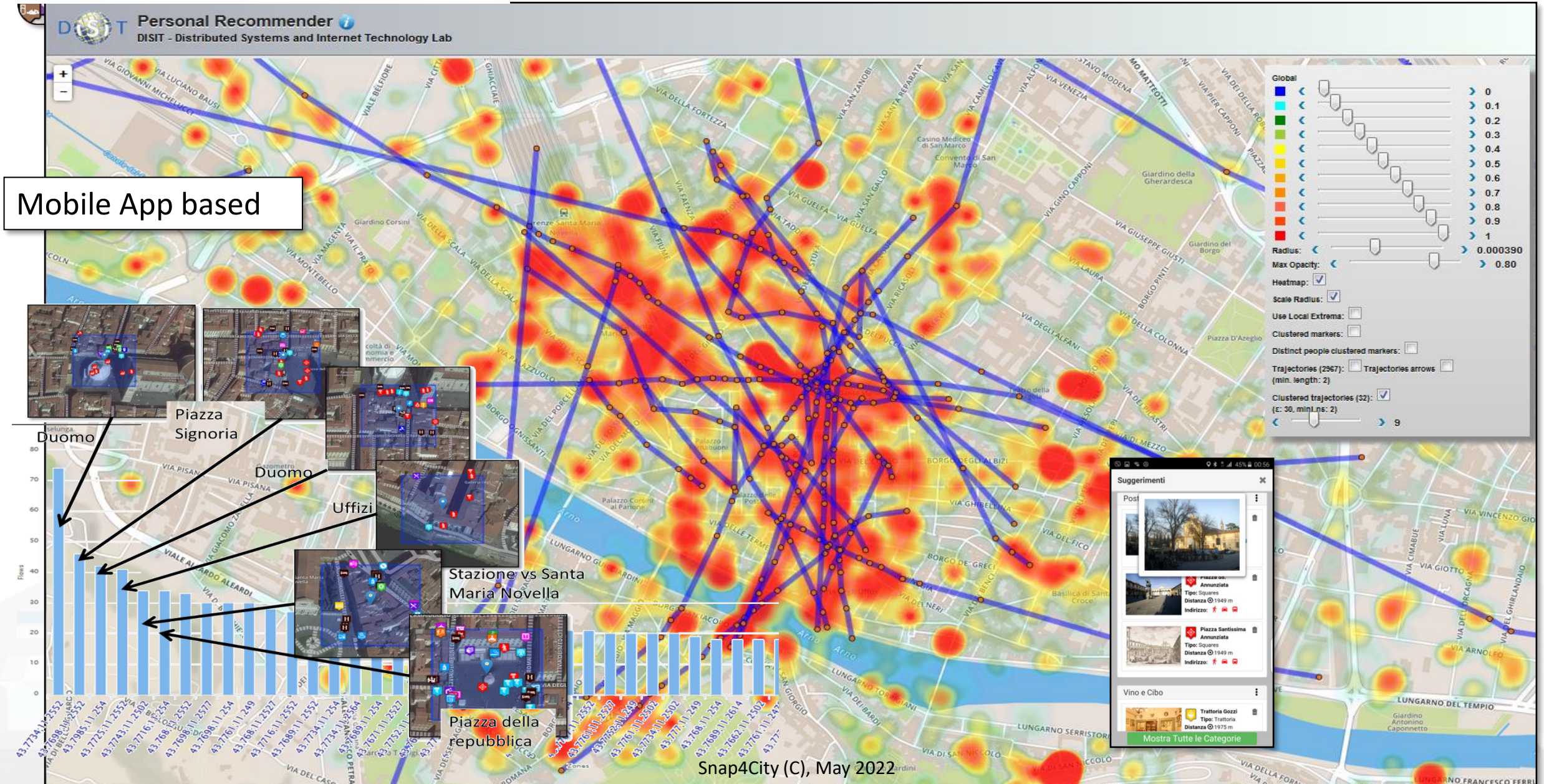
Sii-Mobility

UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB

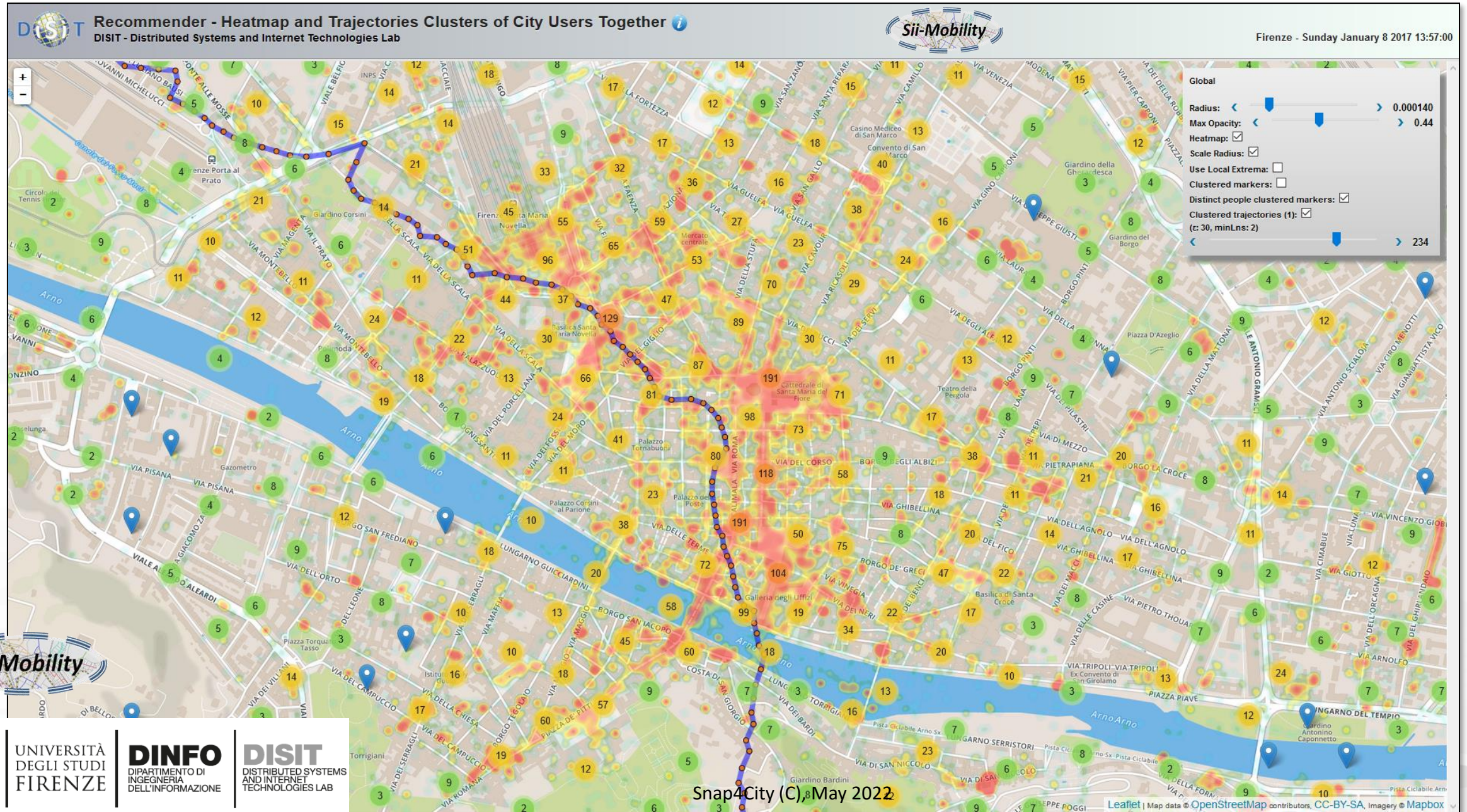
Cluster di Trajectories



User Behavior Analyzer



Heat Map from Mobile: users as sensors

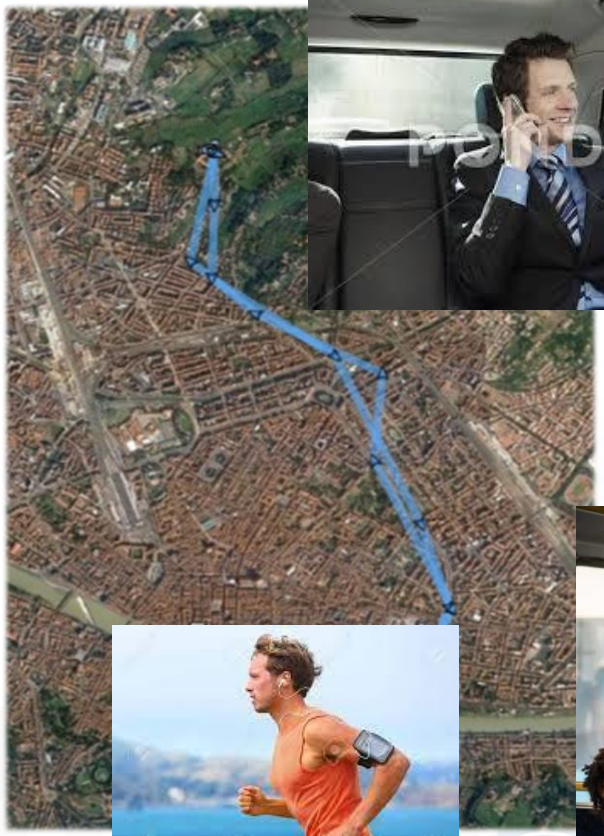


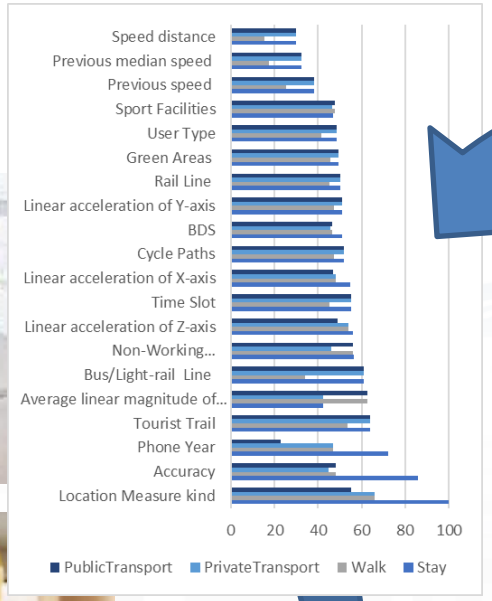
Sii-Mobility

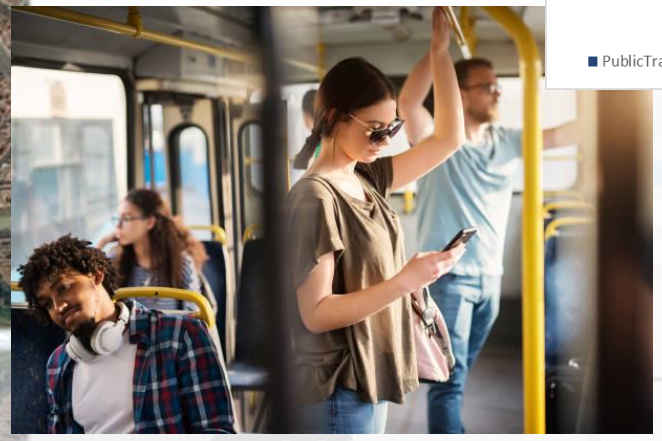
UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB

Recognition of City Users' Transportation means

To propose suggestions and Engage city user I need to know how they are moving

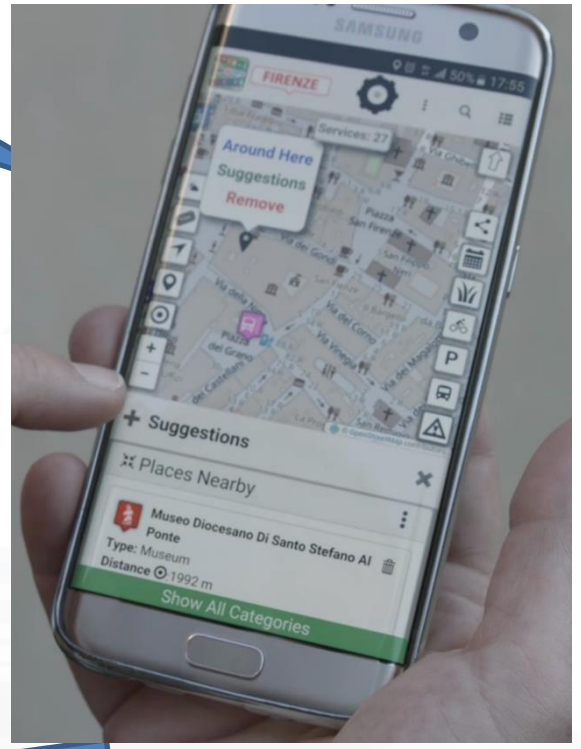






Artificial Intelligence Classification

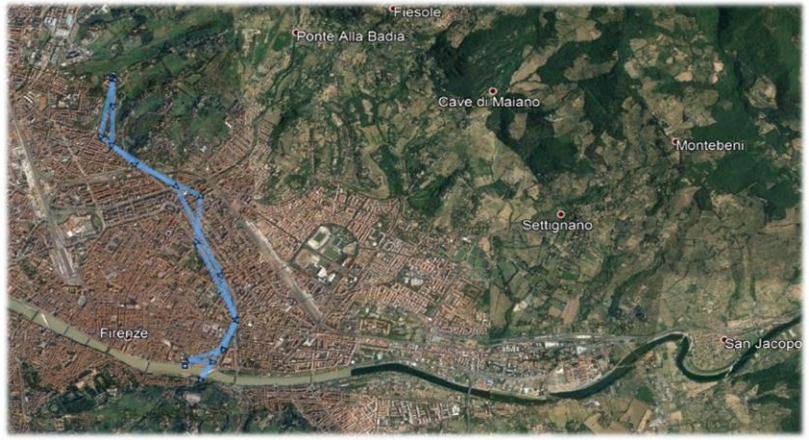
Suggestions



Automated Classification of Users' Transportation Modality in Real Conditions

Variables taken into account:

- ***Day/Time Baseline and GPS***:
- ***Accelerometer***
- ***Proximity***
- ***Temporal window***



Dataset:

- 30K observations
- 25 variables
- 38 different users
- 30 different kinds of devices
- 4 classes (Stationary, Walking, Private Transport, Public Transport)

Four combinations of the different categories of data:

1. Baseline features and distance feature
2. Baseline, distance feature and accelerometer features
3. Baseline, distance feature and temporal window features
4. Baseline, distance, accelerometer, temporal features together

Note that, *each user have used the mean of transport of his/her own preference*.
When the mode of transport is changed, the user was asked to notify the change to the App for creating the learning set and for validation.

Automated Classification of Users' Transportation Modality in Real Conditions

Note that:

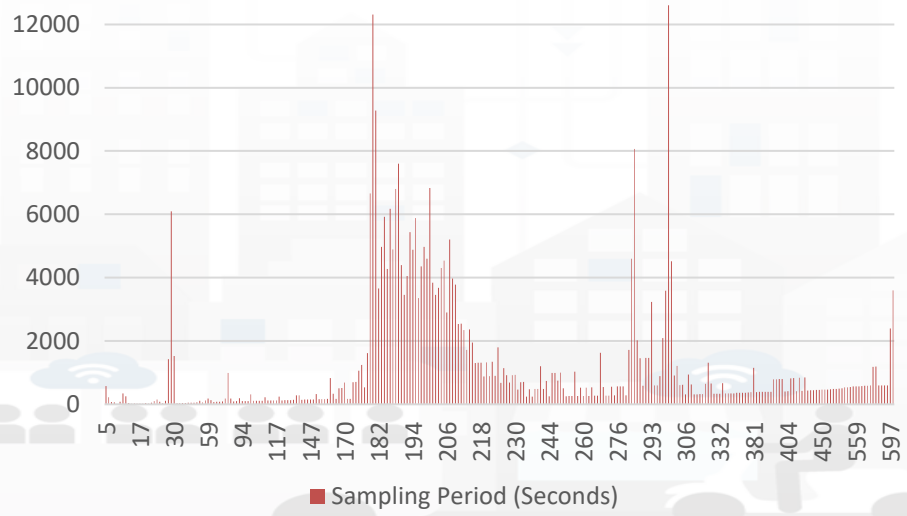
- ***Large discontinuities samples of data*** (from sensors and sporadic communications to the central computation modules)
- Relevant ***differences due to the different kind of mobile phone features in terms of sensors and precision***.

In the state of the art experiments the devices have been asked to keep the application running in foreground to get more precise GPS data, the device in a proper position/orientation during the usage and to use specific devices.

In the proposed solution *no restrictions on the modality of mobile device usage have been imposed.*

- Most of the data was collected in the background because the phones were kept in pocket or bag.
- There is a non-conformity in the Sampling Period frequency distribution of the collected data.

  In details, the frequency average is equal to 180 seconds and the variance is equal to 13240 seconds.



Automated Classification of Users' Transportation Modality in Real Conditions

One-Step machine learning approach:

- *Random Forest* (**RF**)
- *Extremely Randomized Trees* (**Extra-Trees**)
- *Extreme Gradient Boosting procedure (XGBoost)*

Classifier Models	Accuracy	Precision	Recall	F_1 score
Extreme Gradient Boosting	0.947	0.773	0.828	0.800
Random Forest	0.942	0.774	0.869	0.819
Extra-Trees	0.953	0.827	0.869	**0.847**

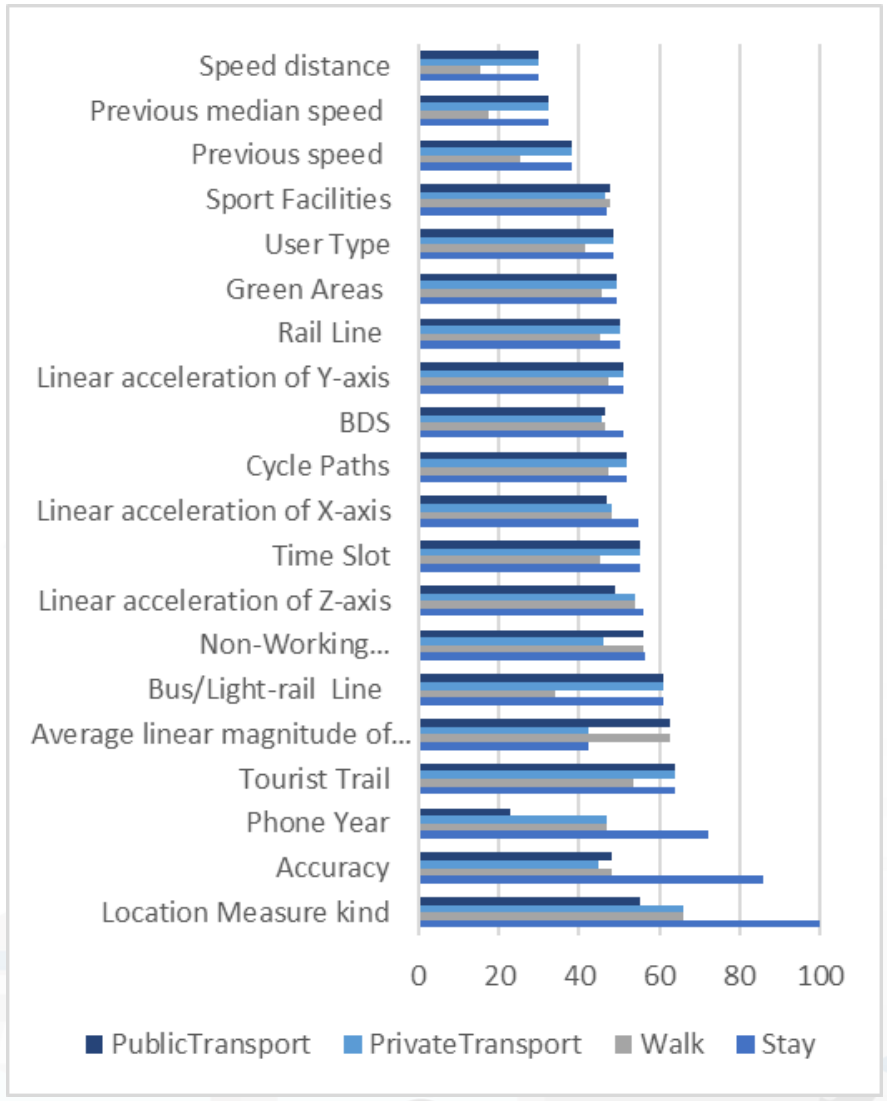
Super Learner Binary Classification Models Combination	Accuracy	Precision	Recall	F_1 score
	0.960	0.865	0.857	**0.861**

Extra Trees Model	Stay	Walk	Private Transport	Public Transport
Sensitivity	0.978	0.731	0.869	0.917
Specificity	0.901	0.988	0.987	0.996
Pos Pred Value	0.977	0.770	0.827	0.936
Neg Pred Value	0.904	0.985	0.990	0.994
Balanced Accuracy	0.940	0.859	0.928	0.956

Super Learner Model	Stay	Walk	Private Transport	Public Transport
Sensitivity	0.990	0.662	0.857	0.927
Specificity	0.892	0.993	0.990	0.996
Pos Pred Value	0.975	0.831	0.865	0.953
Neg Pred Value	0.955	0.982	0.989	0.994
Balanced Accuracy	0.941	0.828	0.924	0.961

- **Super Learner approach:** identification of the multi-class problem into binary classification sub-problems to estimate the risk on future data and select the optimal learner based on the One-Step machine learning approach candidates.
  - Four binary classification models have been constructed:
    1. *stationary vs walking, private transport, public transport*
    2. *walking vs stationary, private transport, public transport*
    3. *private transport vs stationary, walking, public transport*
    4. *public transport vs stationary, walking, private transport*

❖ In **Super Learner**, Binary Classification Models results have been combined **on the highest probability estimation**.



Feature relevance

Model features categories	Extra Tree Model results			
	Accuracy %	Precision %	Recall %	F_1 Score
Baseline and GPS	91.0	68.2	75.1	0.714
Baseline and GPS + proximity	92.4	73.9	69.1	0.715
Baseline and GPS + proximity + Accelerometer	92.6	81.4	74.4	0.777
Baseline and GPS + proximity + Temporal window	94.9	80.5	78.7	0.787
Baseline and GPS + proximity + Accelerometer + Temporal window	95.3	82.7	86.9	0.847

Automated Classification of Users' Transportation Modality in Real Conditions

Two-Steps Hierarchical approach:

combination of the **Extra-Tree** multi-class classification and the **Super learner** algorithm.

- **First Step**: Extra-Tree multi-class classifier *to select* the *two transportation means with higher probability* - 4 different training models.

A **threshold** has been used to decide which class can be considered directly correct at the first step: *if the probability of the class is higher respect the considered threshold* (0.90), *the transportation modality is regarded correct* without proceeding to the second step.

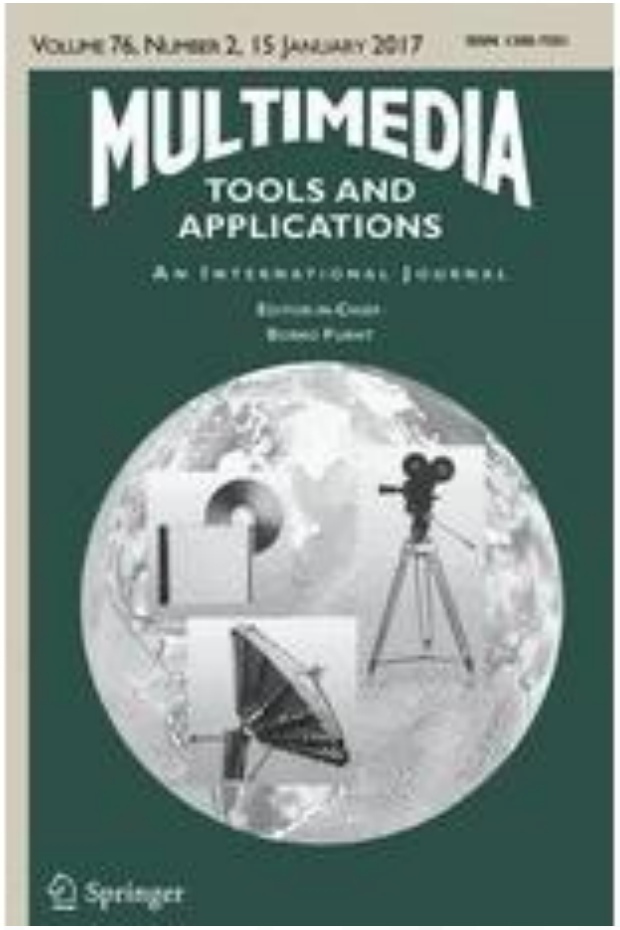
- **Second Step:** Super learner approach to *discriminate between the two transportation means selected in the first step* - 24 different training models

(6 transportation modality pairs combinations per 4 categories combinations)

Two-Steps Hierarchical Approach		Predicted			
		Stay	Walk	Private Transport	Public Transport
Actual	Stay	**0.98**	0.30	0.09	0.03
	Walk	0.01	**0.60**	0.02	0.01
	Private Transport	0.01	0.07	**0.87**	0.07
	Public Transport	0.00	0.03	0.01	**0.89**

Accuracy = **0.940**
Precision= 0.786
Recall = 0.869

- Badii, C., Difino, A., Nesi, P. *et al.* Classification of users' transportation modalities from mobiles in real operating conditions. *Multimed Tools Appl* (2021). https://doi.org/10.1007/s11042-021-10993-y
- https://link.springer.com/article/10.1007/s11042-021-10993-y



Traffic Flow Prediction

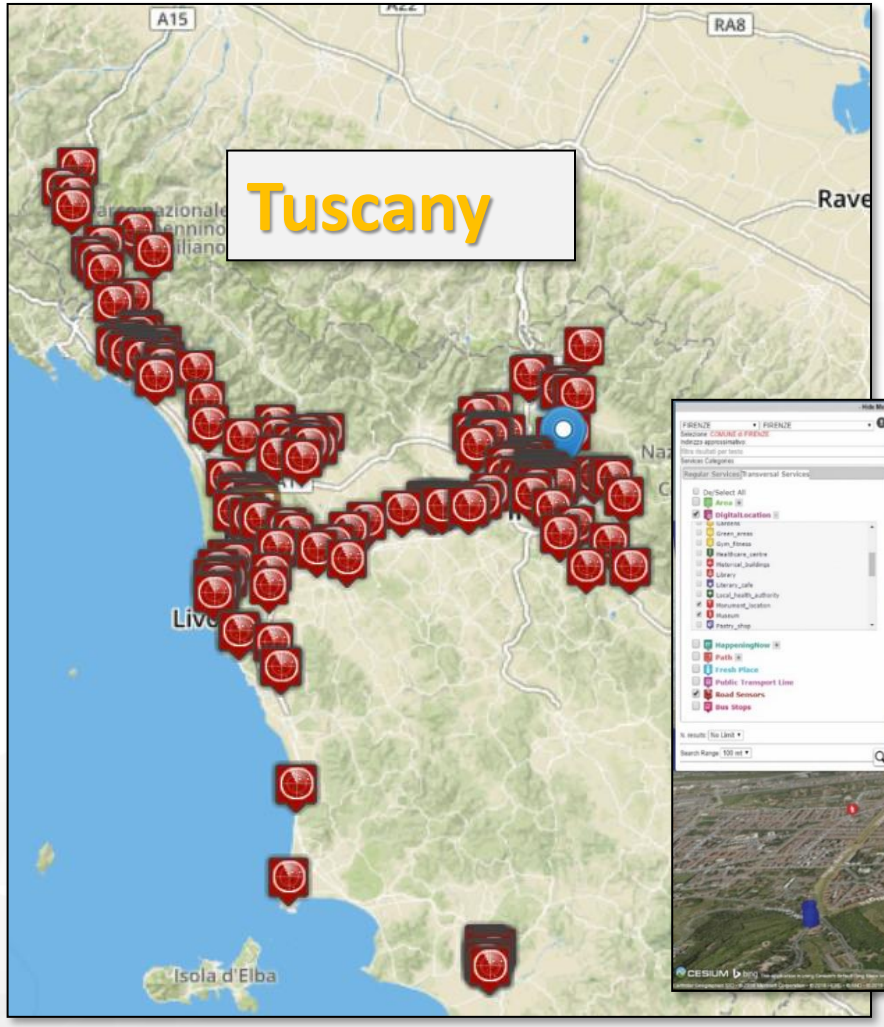
60

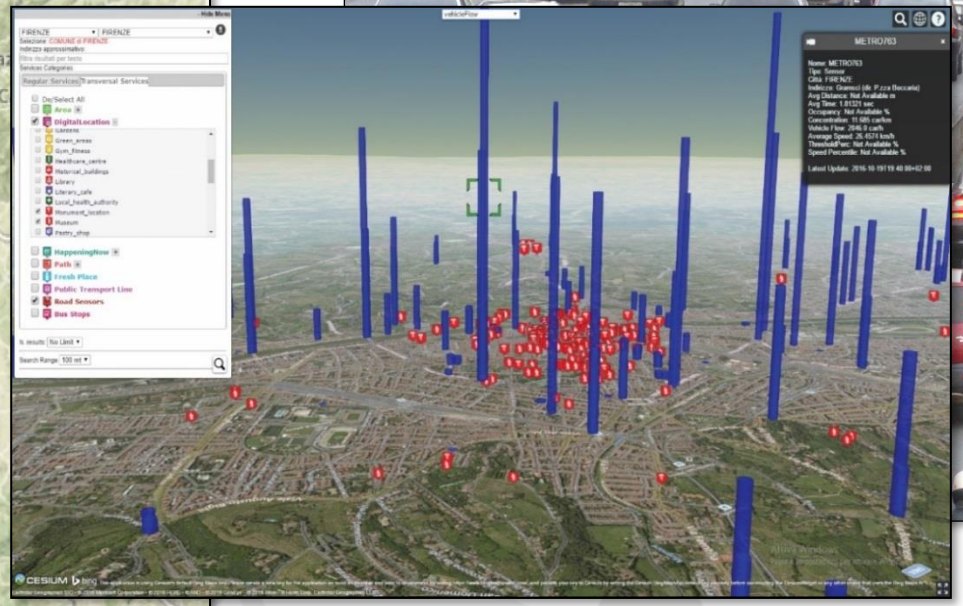
60
35

Traffic Flow Tools

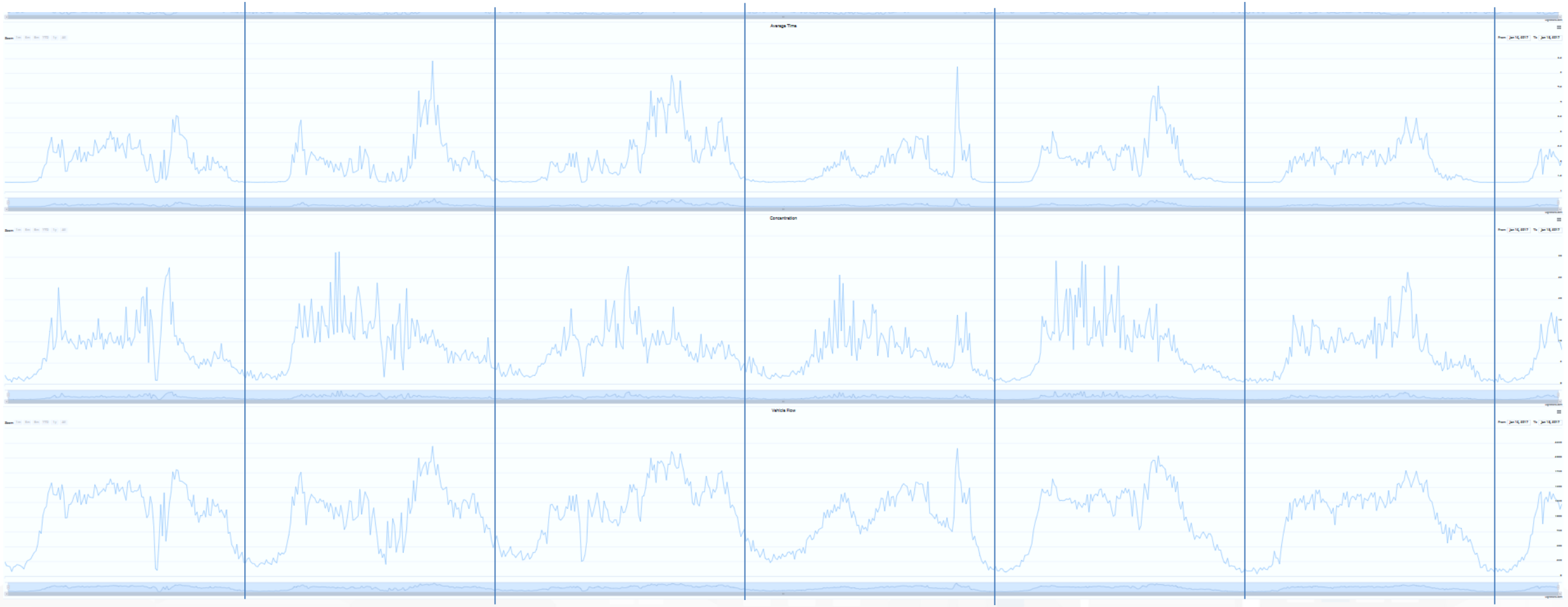
Spire and Virtual Spires (cameras), Bluetooth, ...

Specifically located: along, around, on gates, on x...





Traffic Flow data



- Day by day traffic flow, on the week data from 3 sensors

Firenze - Trafair - AirQuality Heatmaps

This dashboad contains data derived from actual sensors and predictive values under validation

Mon 6 Apr 15:12:27

- Air Quality Sensors
- Weather Sensors
- PM10 Heatmap
- PM2.5 Heatmap
- CO Heatmap
- CO2 Heatmap
- O3 Heatmap
- NO2 Heatmap
- Europ. AQI Heatmap
- Air Humidity Heatmap
- Air Temp. Heatmap
- Wind Speed Heatmap
- Gral Pred. HM NOX (3m)
- Gral Pred. HM NOX (6m)
- Traffic Sensors
- Traffic Flow
- Cycling Paths
- Accident Heatmap
- Accident Heatmap 2
- Only HRes Anym. Gral
- Green Areas
- Schools

Air quality trends

Multi Map

Maps

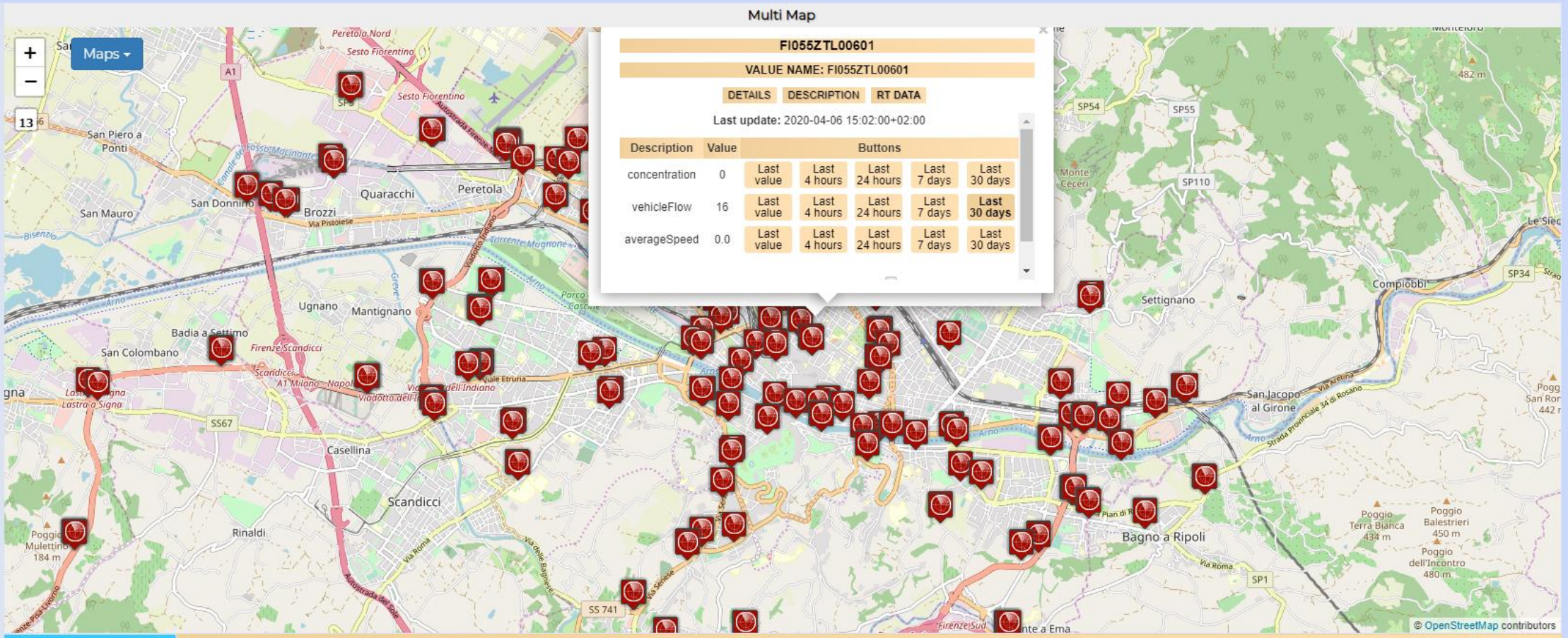
FI055ZTL00601

VALUE NAME: FI055ZTL00601

DETAILS | DESCRIPTION | RT DATA

Last update: 2020-04-06 15:02:00+02:00

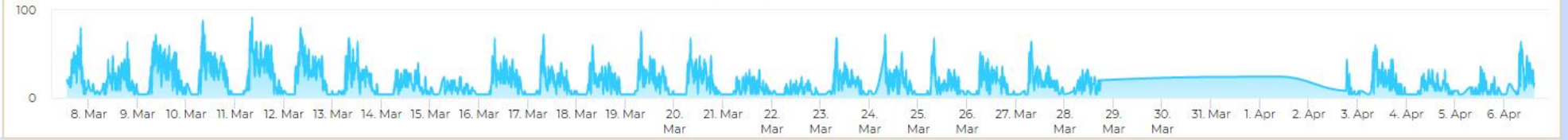
Description	Value	Buttons				
concentration	0	Last value	Last 4 hours	Last 24 hours	Last 7 days	Last 30 days
vehicleFlow	16	Last value	Last 4 hours	Last 24 hours	Last 7 days	**Last 30 days**
averageSpeed	0.0	Last value	Last 4 hours	Last 24 hours	Last 7 days	Last 30 days



Air Temperature 9m

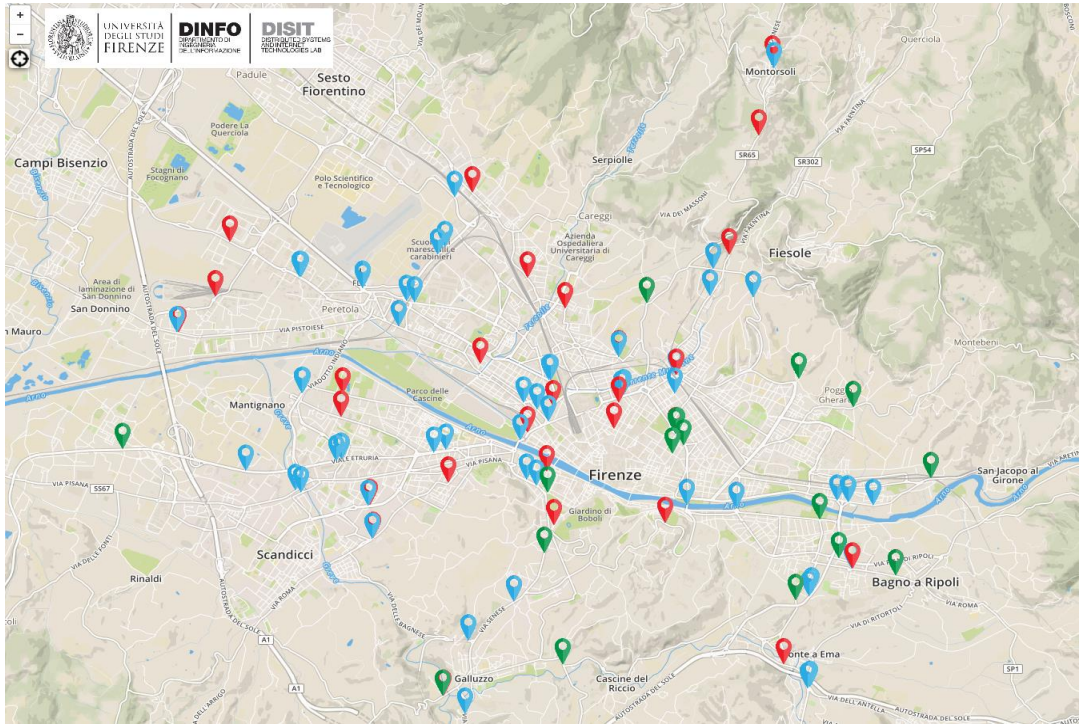
21.8 °C

vehicleFlow - 30 days 9m

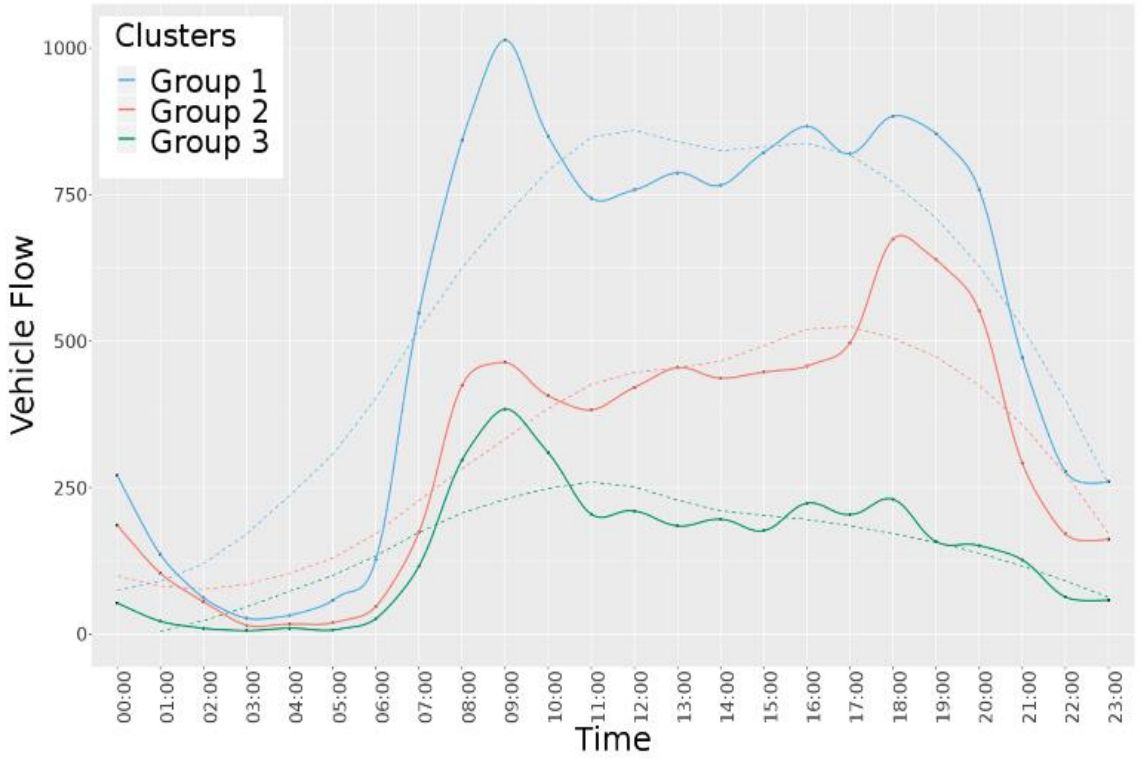


Traffic Data Analysis

Map of the traffic sensors location per cluster in Florence municipality



Hourly median vehicle flow trends per cluster

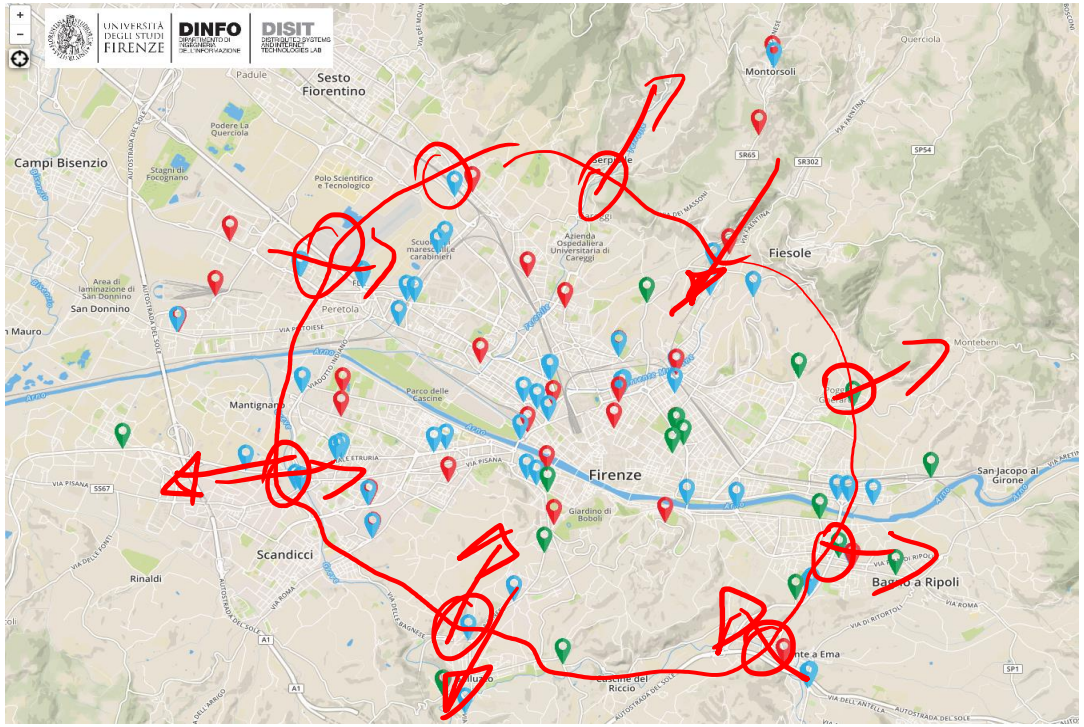


Example of Volume of data

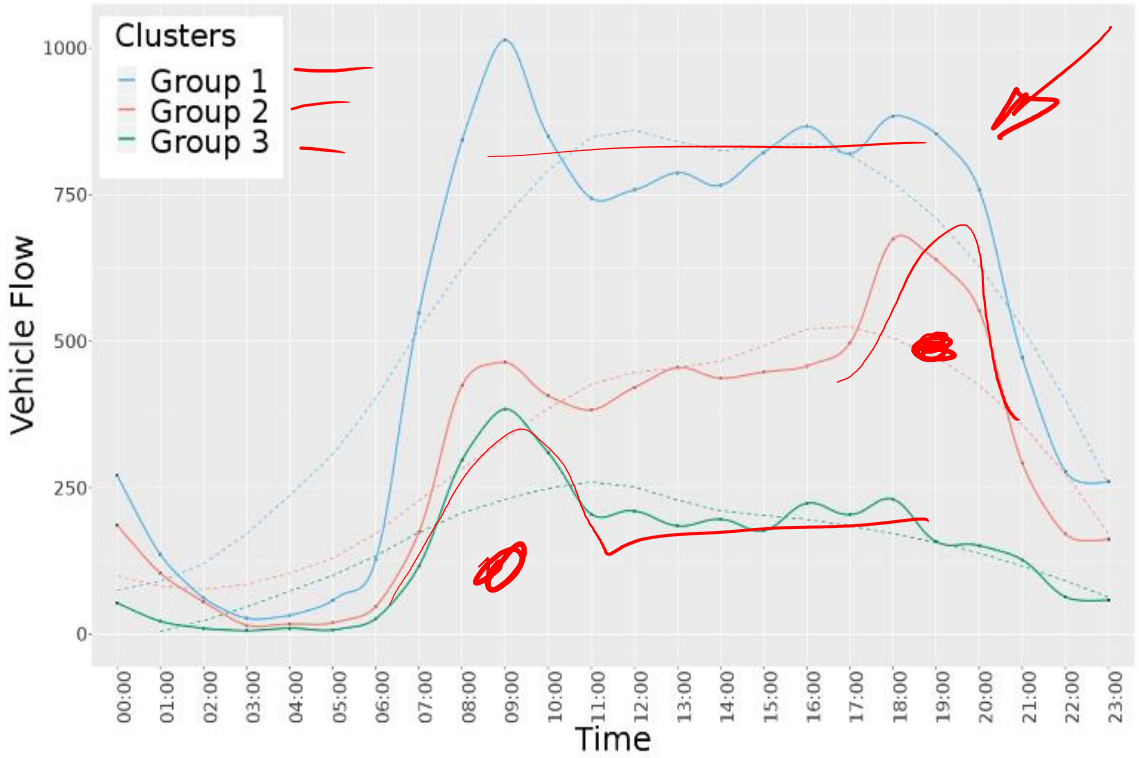
- Variable per sensor: 15
- Byte per sensor/message: 150 Byte
- Sensors: 150
- Days per year: 365
- Hours in the day: 24
- Sample for hour: 6, one each 10 minutes
- Platform factor: number of replicas, indexing, etc…
  - May range from 100 to 2000 Byte per Variable

Traffic Data Analysis

Map of the traffic sensors location per cluster in Florence municipality

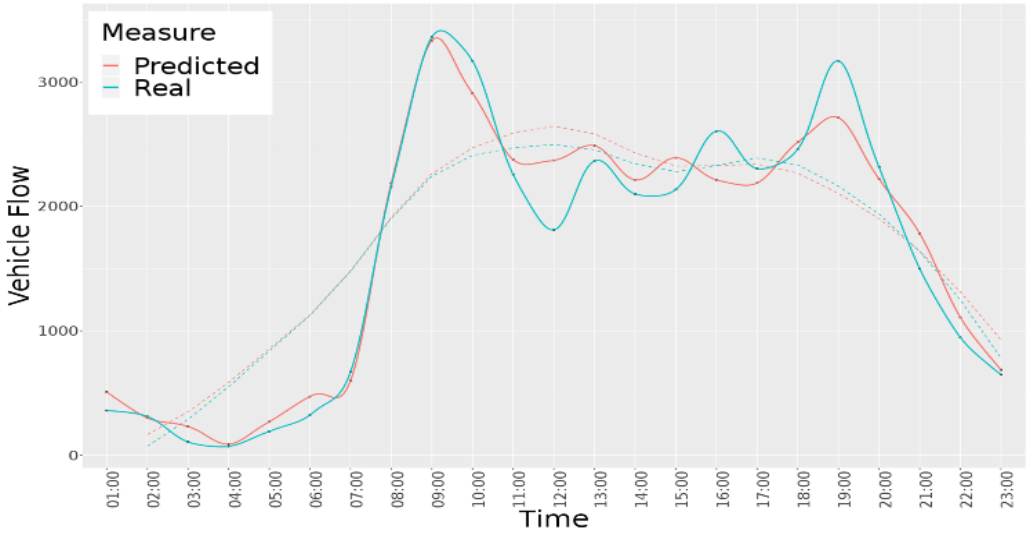


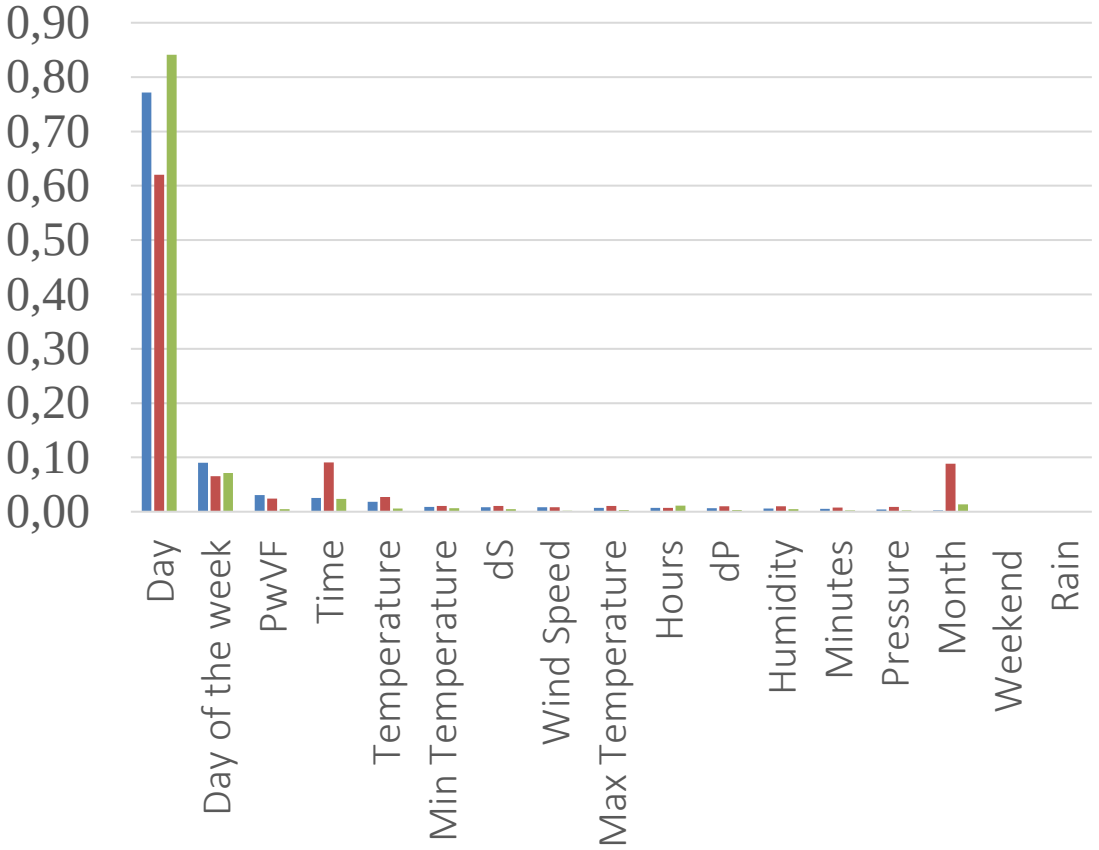
Hourly median vehicle flow trends per cluster



Traffic Flow Predictions

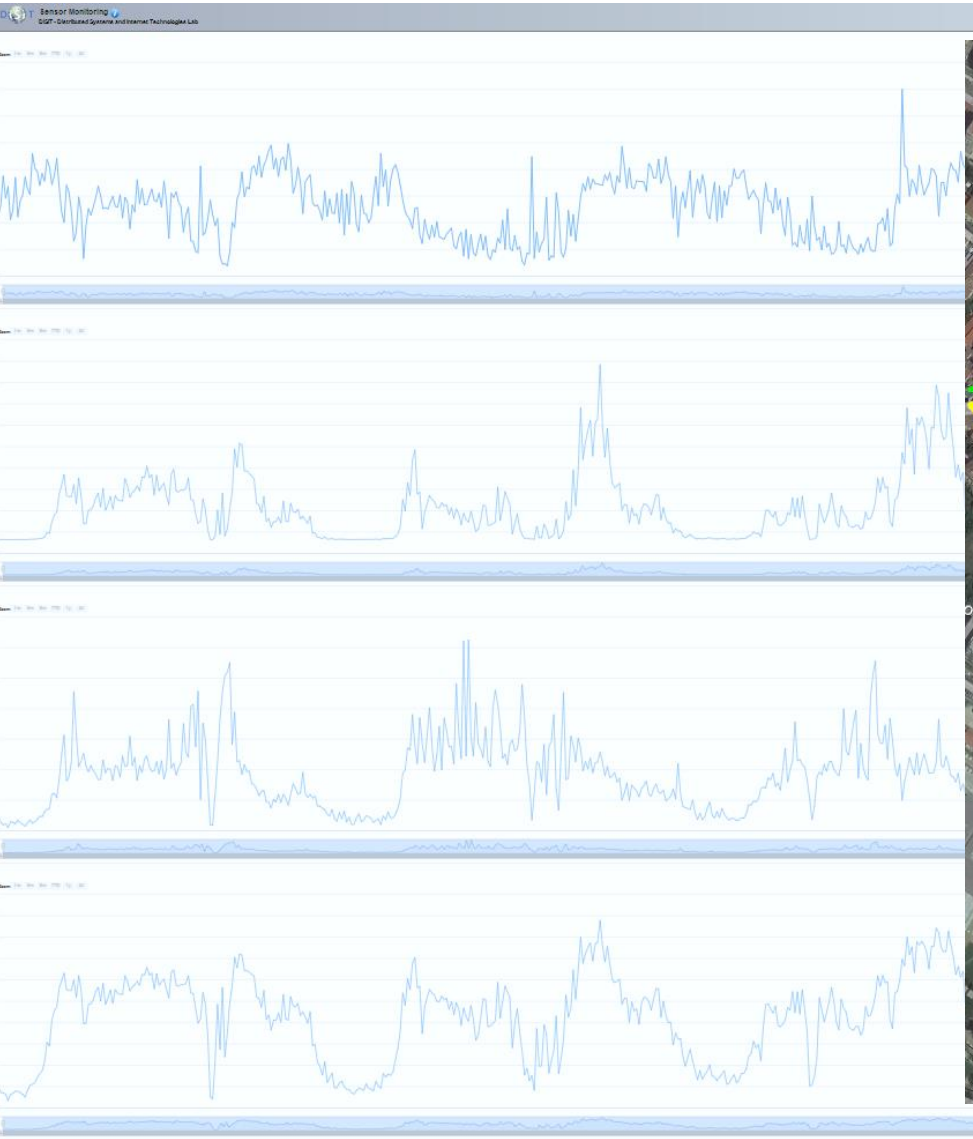
XGBoost Model Results	R^2	RMSE	MASE
Sensors of Group 1	0.95	215	0.89
Sensors of Group 2	0.91	178	0.82
Sensors of Group 3	0.86	127	0.92

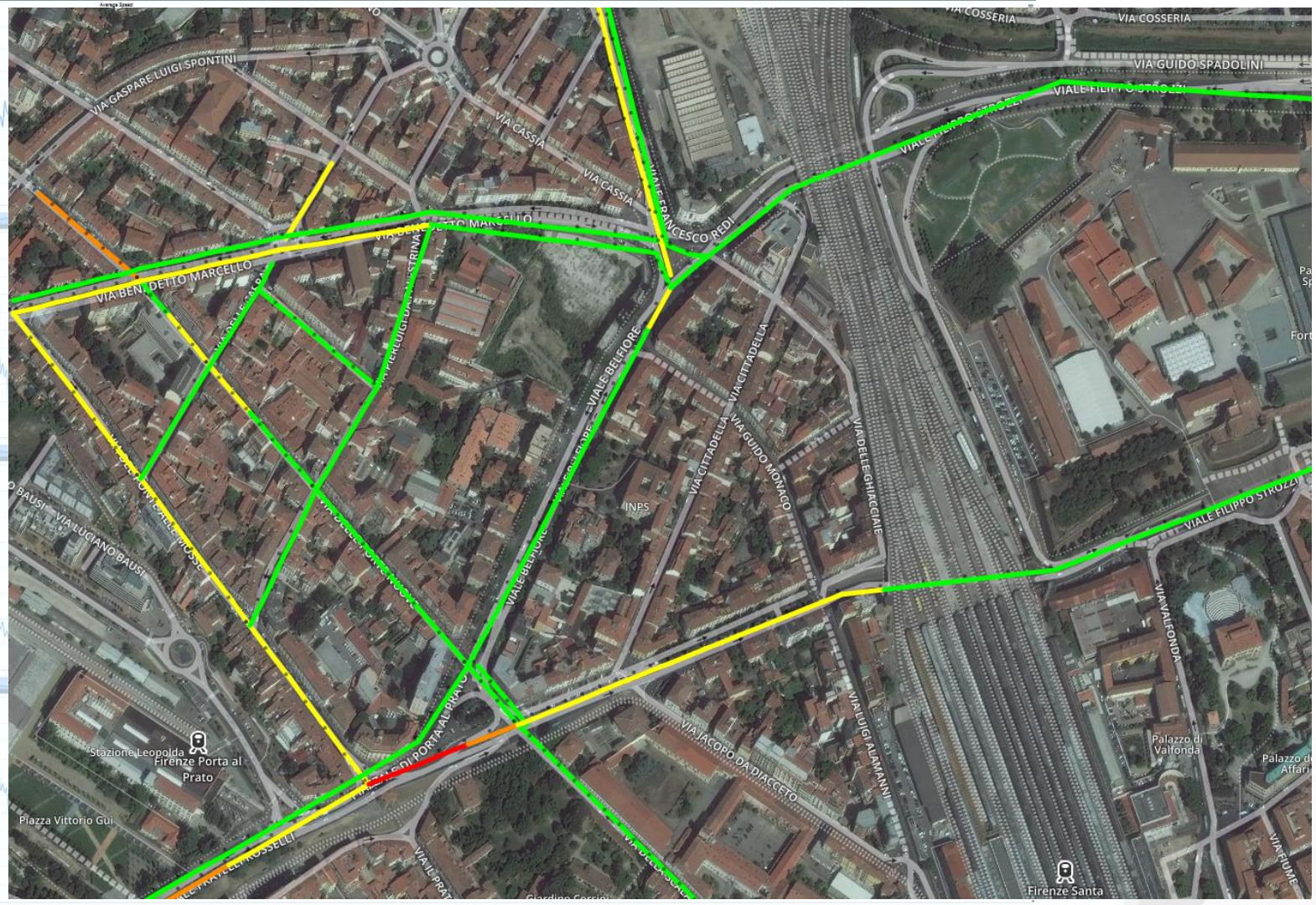




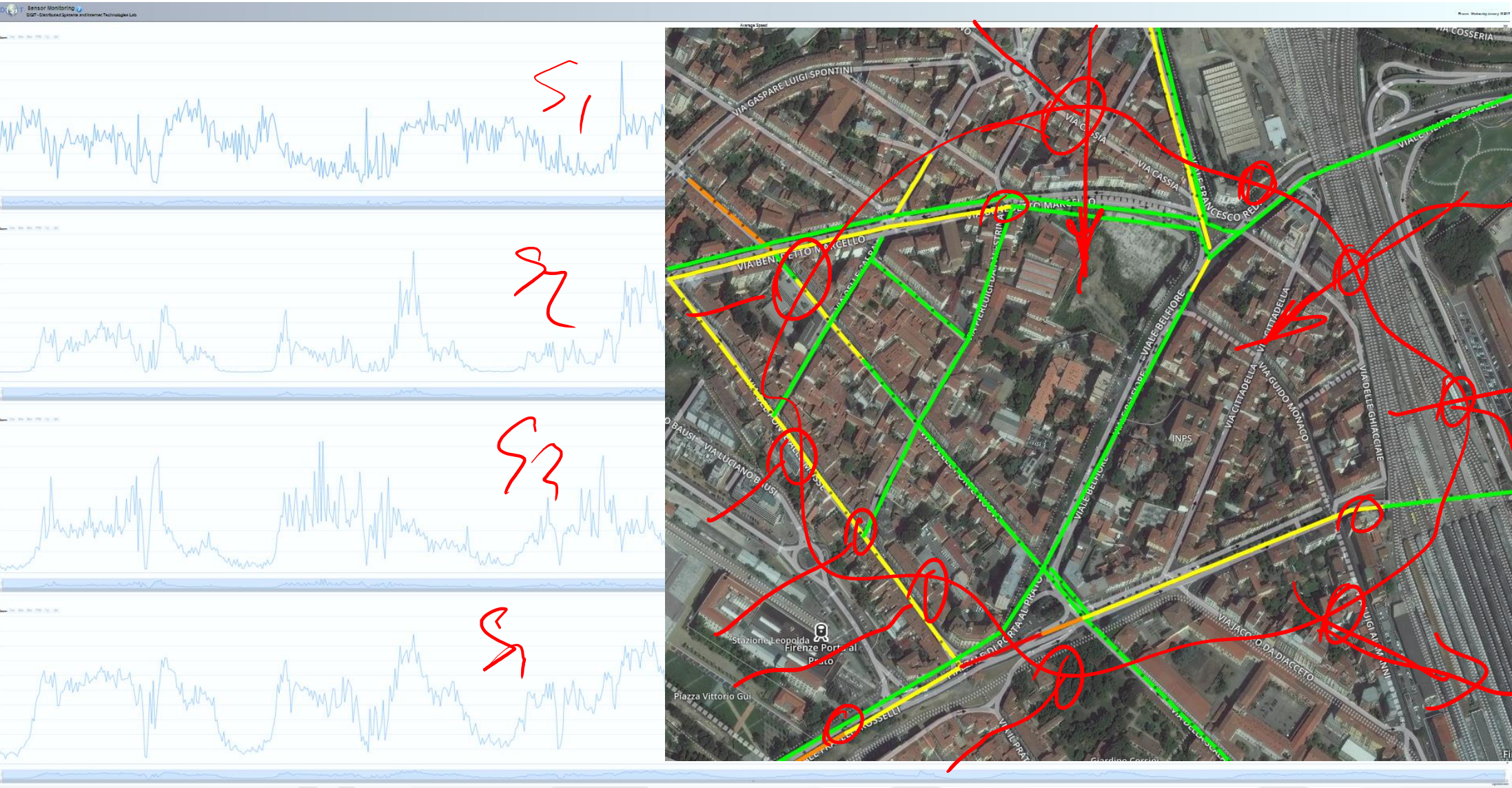
Traffic Flow
In/out of the city

Traffic Flow data





Traffic Flow data



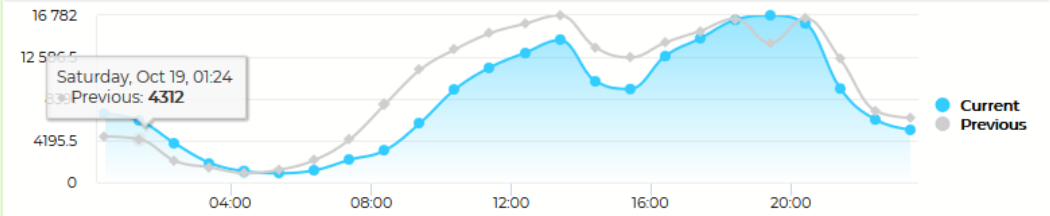
Traffic Flow Monitoring - Firenze

Sun 20 Oct 23:37:24

 # IN FLOW 9m

5302

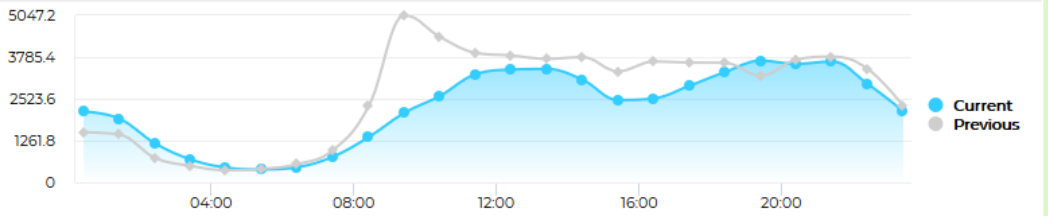
Firenze IN Traffic Flow (number of vehicles) 9m



ZTL in 9m

2149

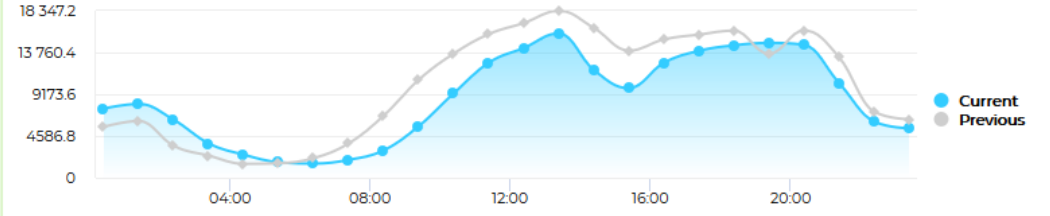
ZTL in Traffic Flow daily trend 9m



 # OUT FLOW 9m

5448

Firenze OUT Traffic Flow (number of vehicles) 9m

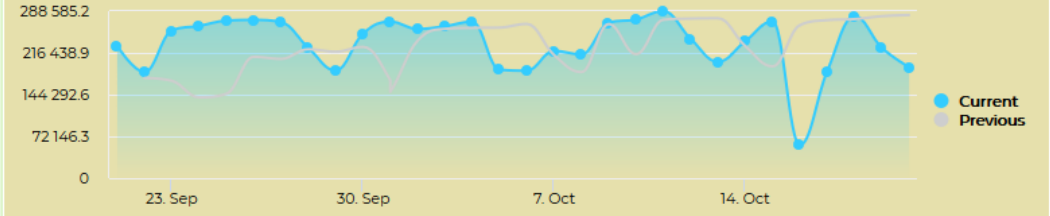


Main

Inc Daily Inp... 9m

191840

Daily Inputs (monthly) (last value is incremental) 9m

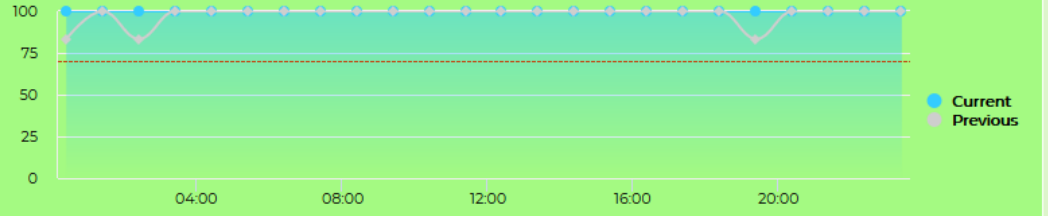


Current Time - ...

20/10/2019

23:37:25

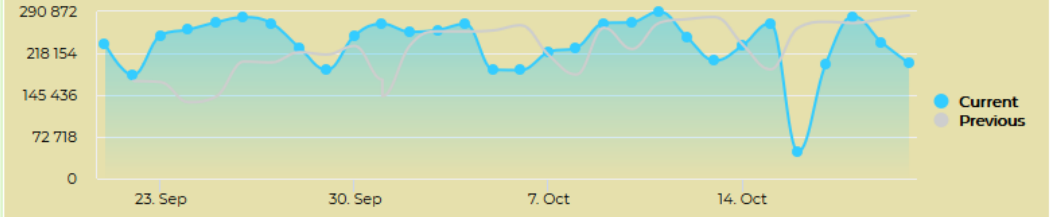
Quality of Measure in terms of percentage of expected measures considered 9m



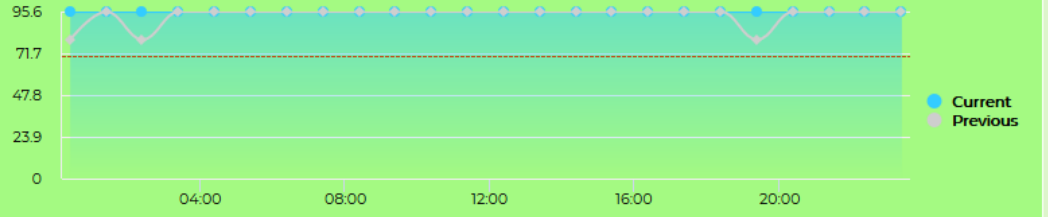
Inc Daily Ou... 9m

201019

Daily Outputs (monthly) (last value is incremental) 9m



Quality of Measure in percentage - ZTL in time trend compare 9m



https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MTc2MQ==

Privacy Policy Cookies Policy Terms and Conditions Contact us

On sunday

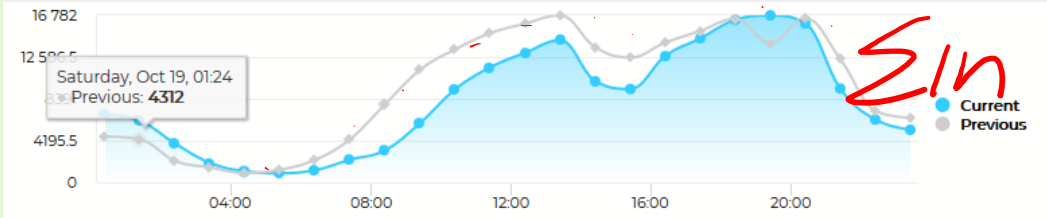
Traffic Flow Monitoring - Firenze

Sun 20 Oct 23:37:24

 # IN FLOW 9m

5302

Firenze IN Traffic Flow (number of vehicles) 9m

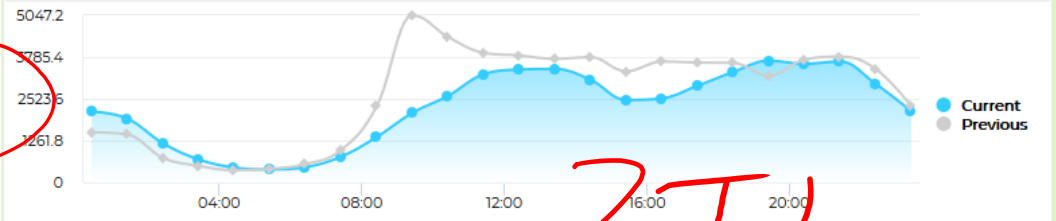


Ein

ZTL in 9m

2149

ZTL in Traffic Flow daily trend 9m

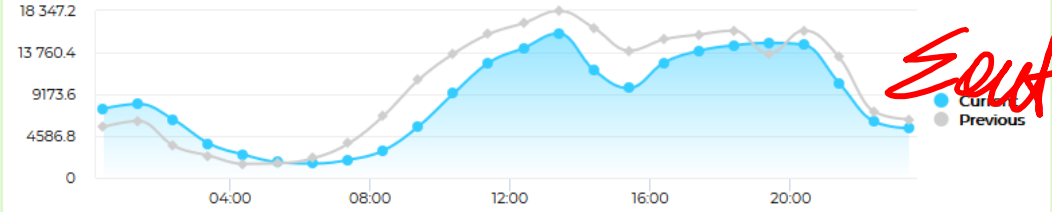


ZTL

 # OUT FLOW 9m

5448

Firenze OUT Traffic Flow (number of vehicles) 9m



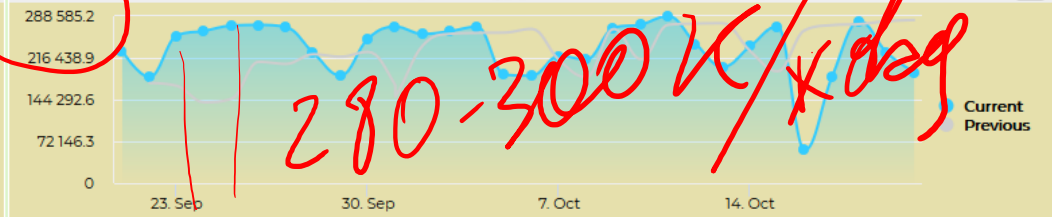
Eout

Main

Inc Daily Inp... 9m

191840

Daily Inputs (monthly) (last value is incremental) 9m



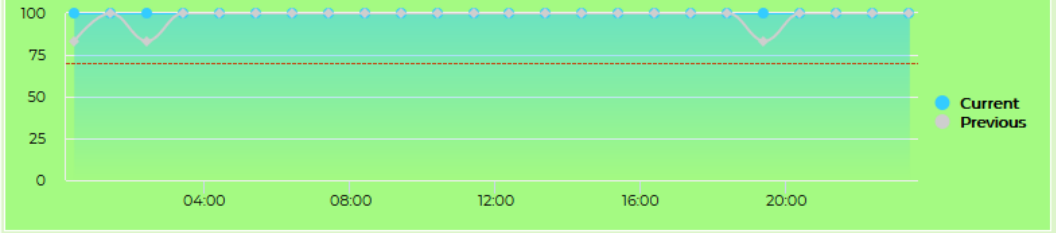
290-300 K/day

Current Time - ...

20/10/2019

23:37:25

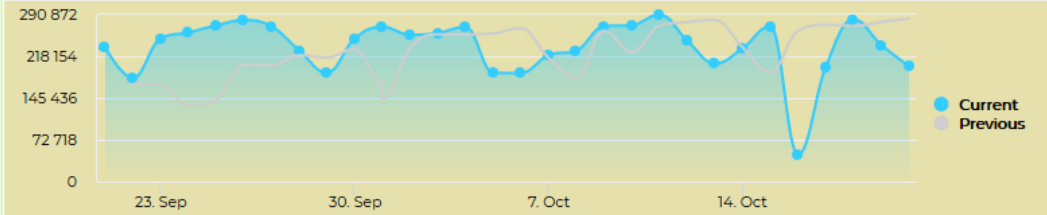
Quality of Measure in terms of percentage of expected measures considered 9m



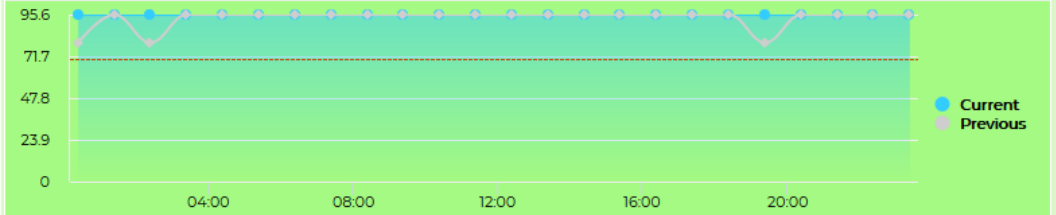
Inc Daily Ou... 9m

201019

Daily Outputs (monthly) (last value is incremental) 9m



Quality of Measure in percentage - ZTL in time trend compare 9m



https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MTc2MQ==

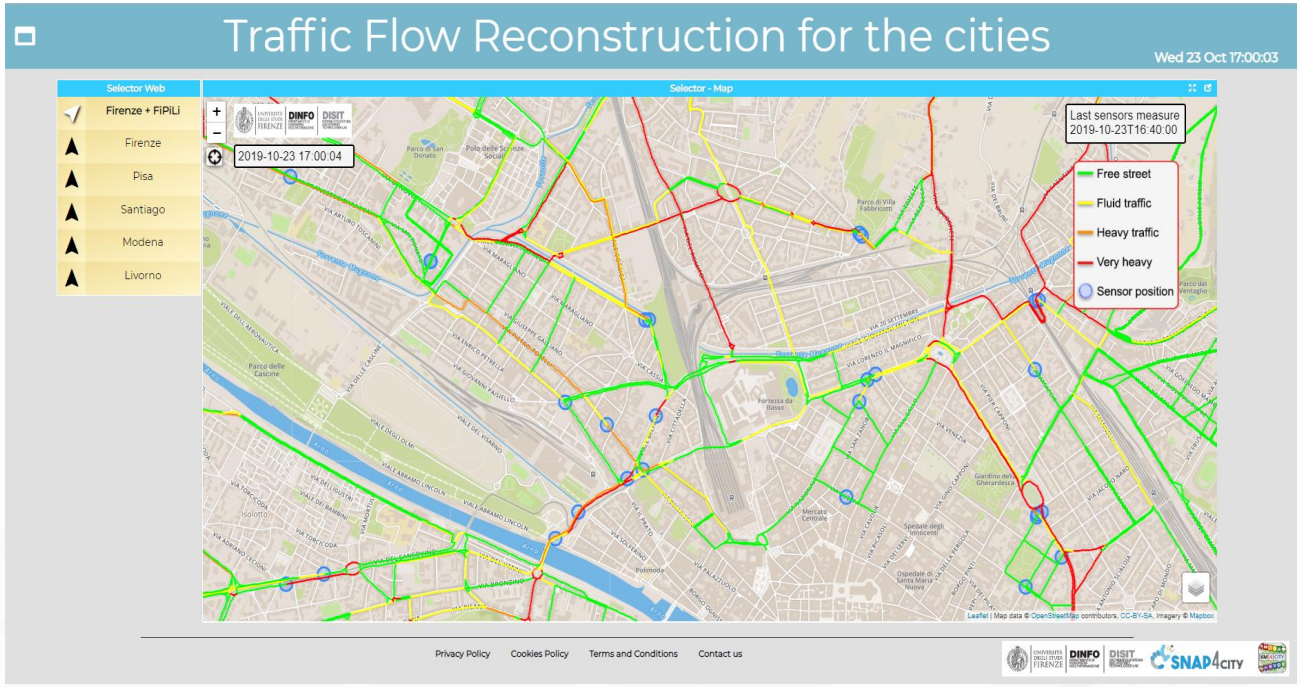
Privacy Policy Cookies Policy Terms and Conditions Contact us

On sunday

Traffic Flow Reconstruction from Traffic Sensors Data

Why Dense Traffic Flow Reconstruction ?

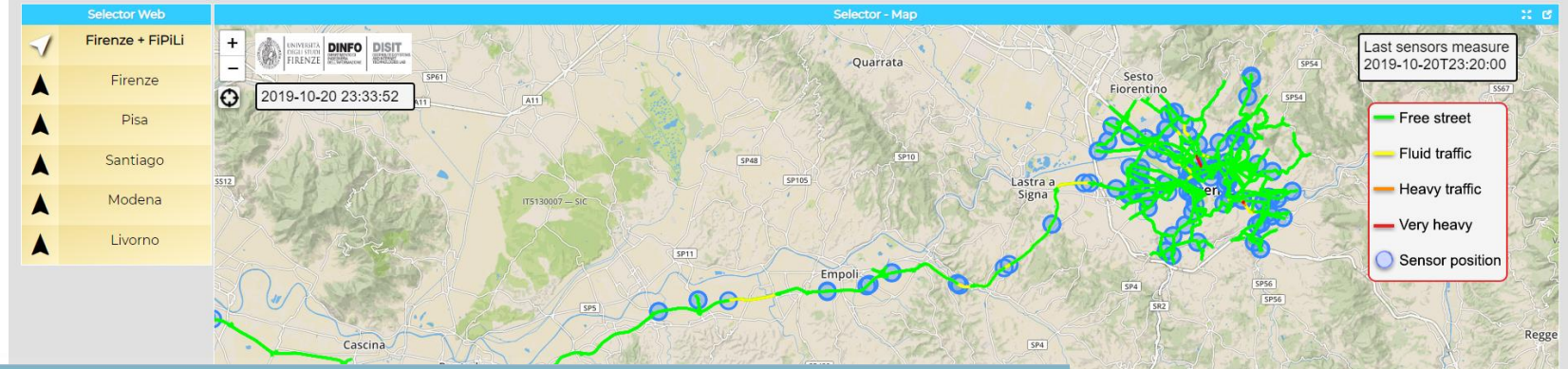
- Making decision on mobility and transport solutions → what if analysis
- Controlling pollution
- Dynamic Routing for Firebrigade, Ambulances, general public
- Planning Public Transportation routing



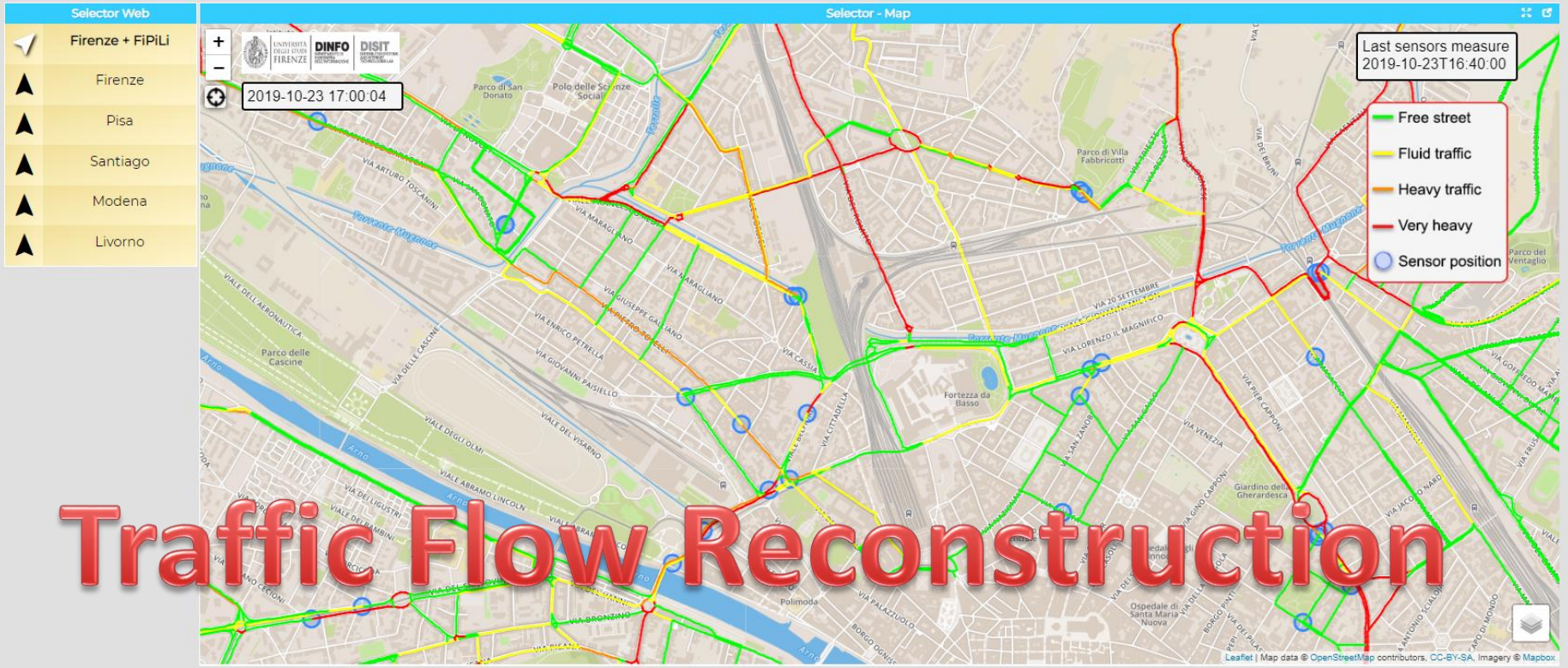
Traffic Flow Reconstruction
2018-02-01T00:10:00
Last sensors measure
2018-02-01T00:10:00
Free street
Fluid traffic
Heavy traffic
Very heavy
Sensor position
Firenze
FiPiLi
Pisa
Modena
real time
Leaflet | Map data © OpenStreetMap contributors, CC-BY-SA, Imagery ¿½ Mapbox

UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB

Traffic Flow Reconstruction for the cities



Traffic Flow Reconstruction for the cities

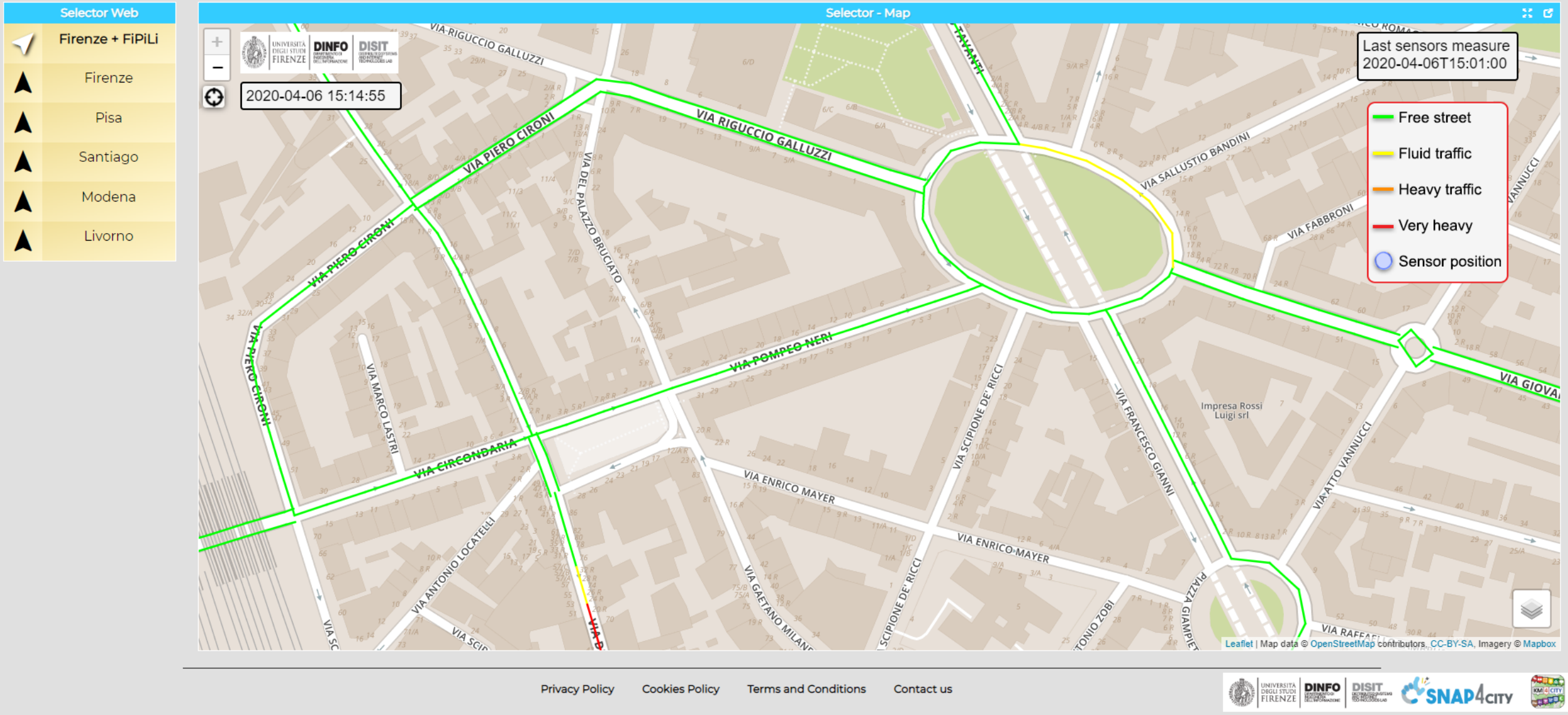


Traffic Flow Reconstruction

https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MTc5NQ==

Traffic Flow Reconstruction for the cities

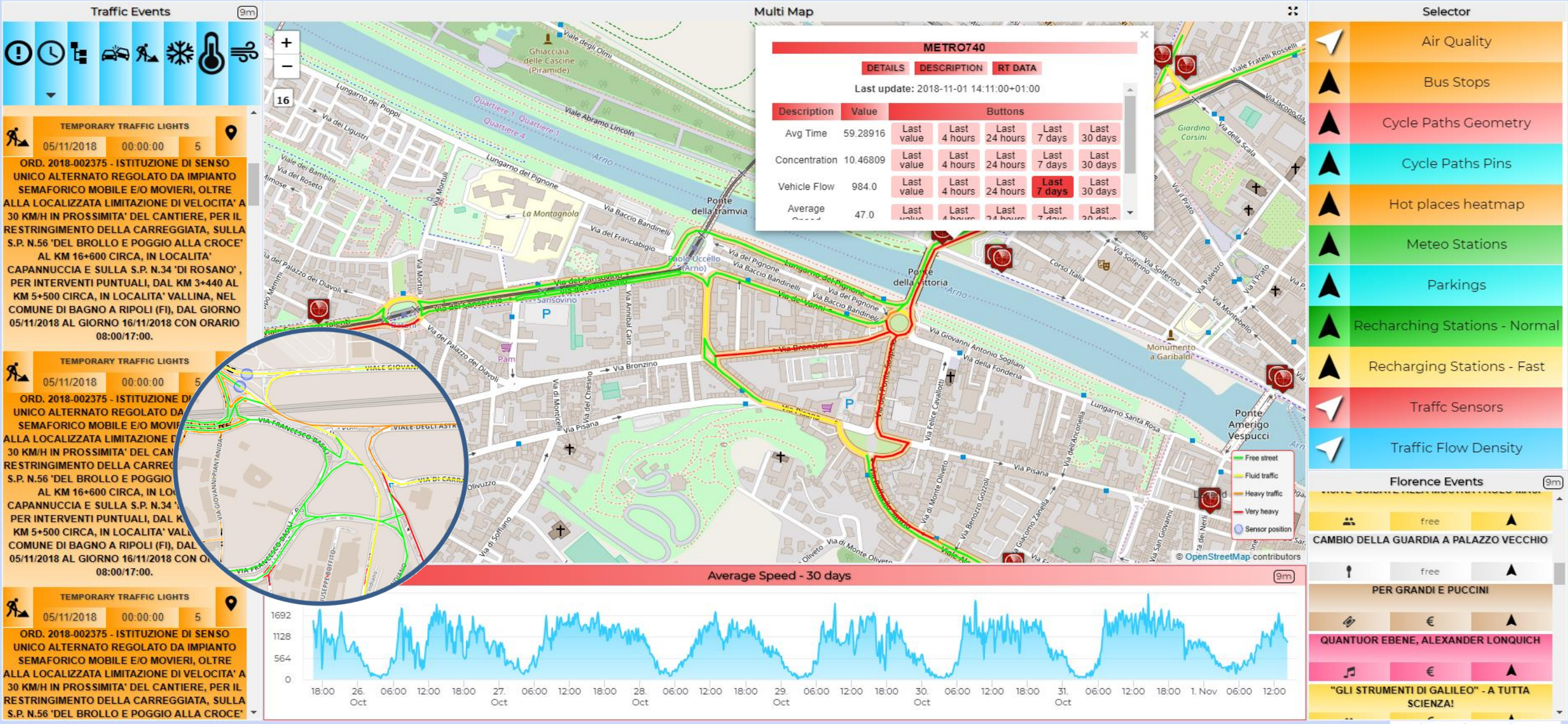
Mon 6 Apr 15:14:55



https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MTc5NQ==

Toscana Traffico

Thu 1 Nov 14:15:47



https://main.snap4city.org/view/index.php?iddasboard=MTE5MQ==

Traffic Flow reconstruction, real time

Mathematical model

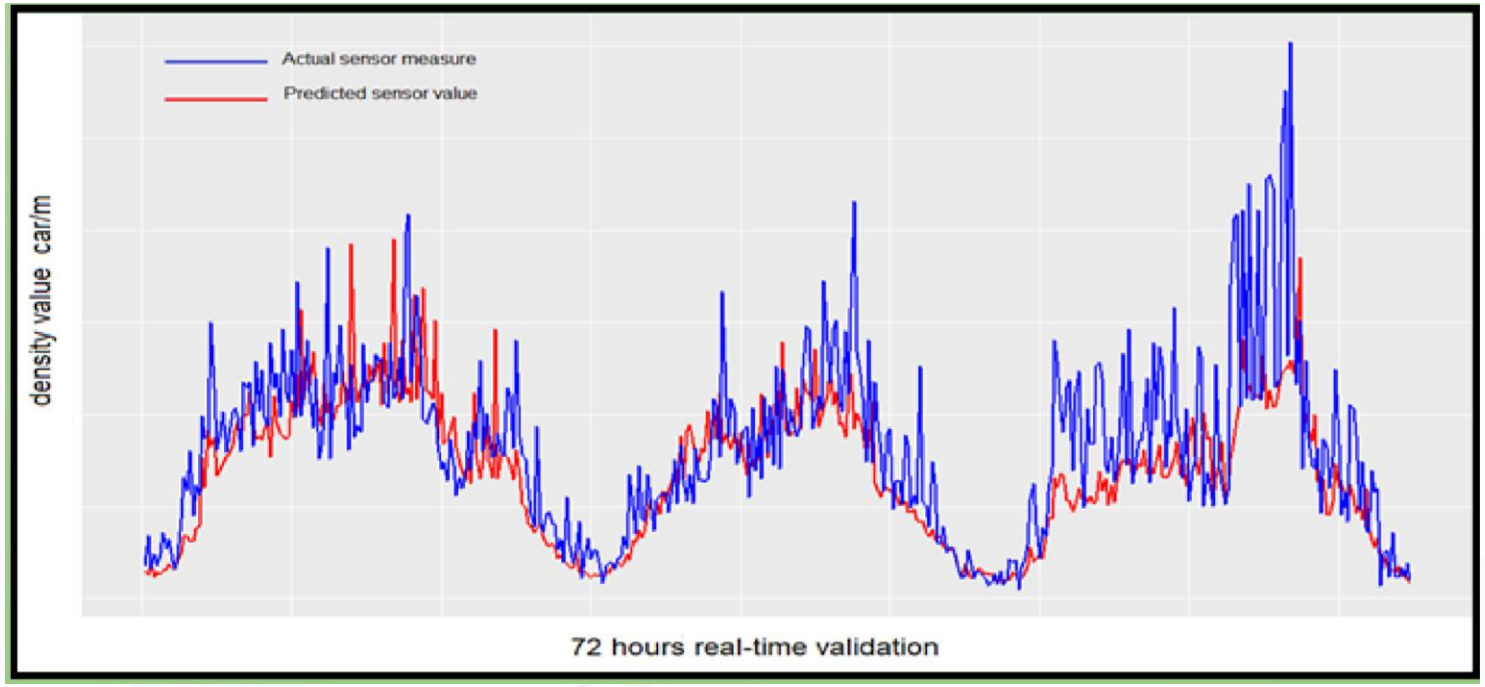
The vehicular traffic flow is propagated in the city graph according to a fluid dynamics model which is based on the conservation law of the vehicles. In a single road, it is described by the following partial differential equation:

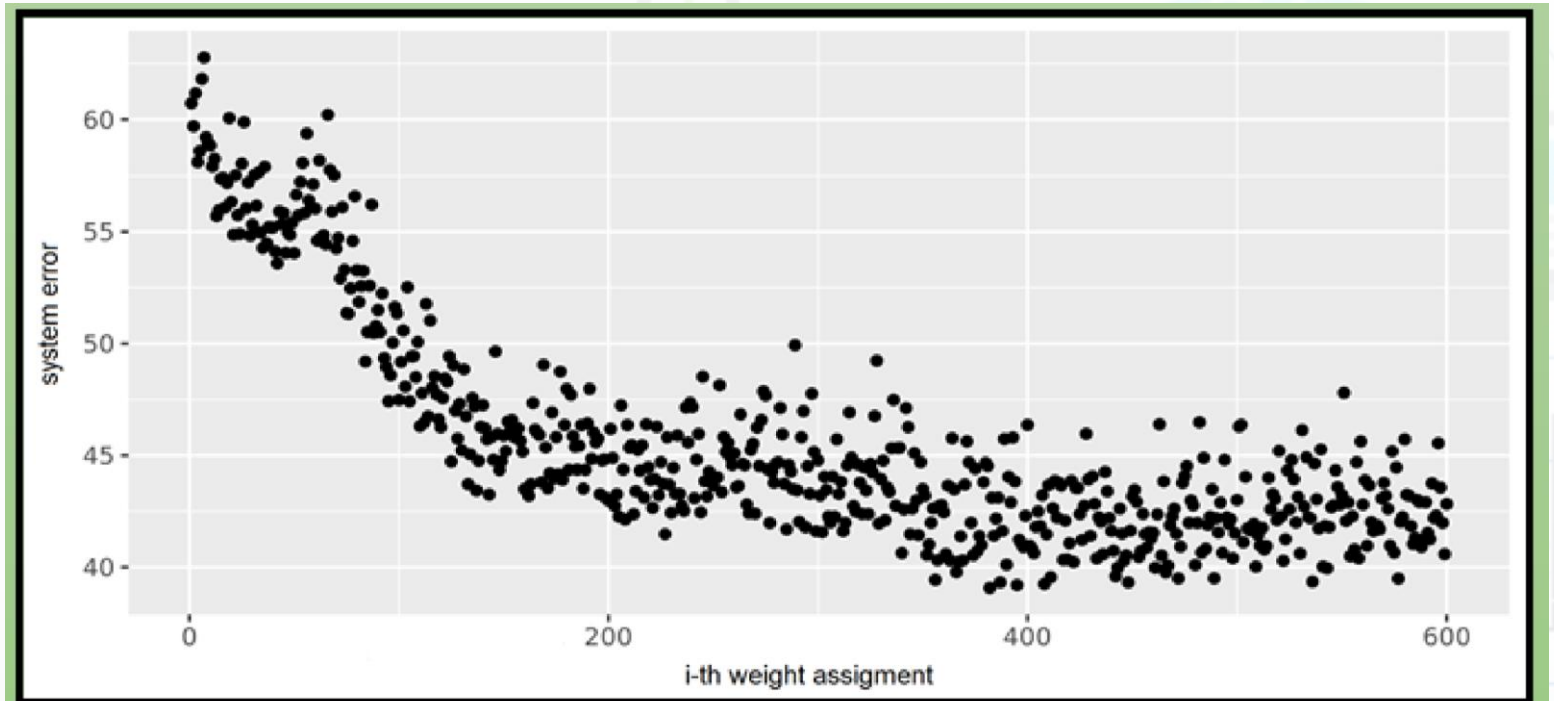
$$\frac{\partial \rho(t,x)}{\partial t} + \frac{\partial f\left(\rho(t,x)\right)}{\partial x} = 0$$

where $\rho(t,x)$ denotes the vehicular density and the function $f(\rho(t,x))$ is the vehicular flux which is defined as the product $\rho(t,x)v(t,x)$, being $v(t,x)$ the local speed of the vehicles.

A discretization scheme in terms of *finite differences* is considered to obtain a numerical solution of the above equation. The traffic flow is then distributed through the junctions in the city.

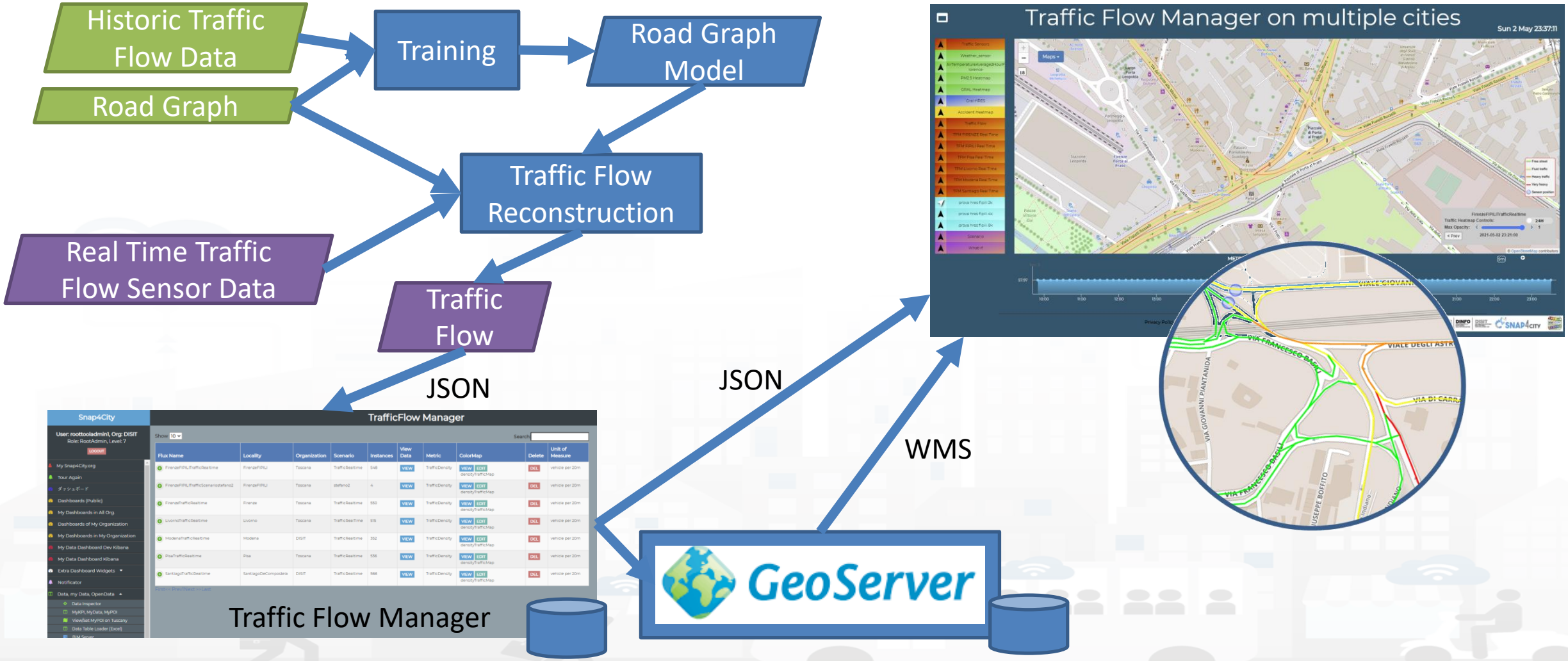
Convergence of learning phase





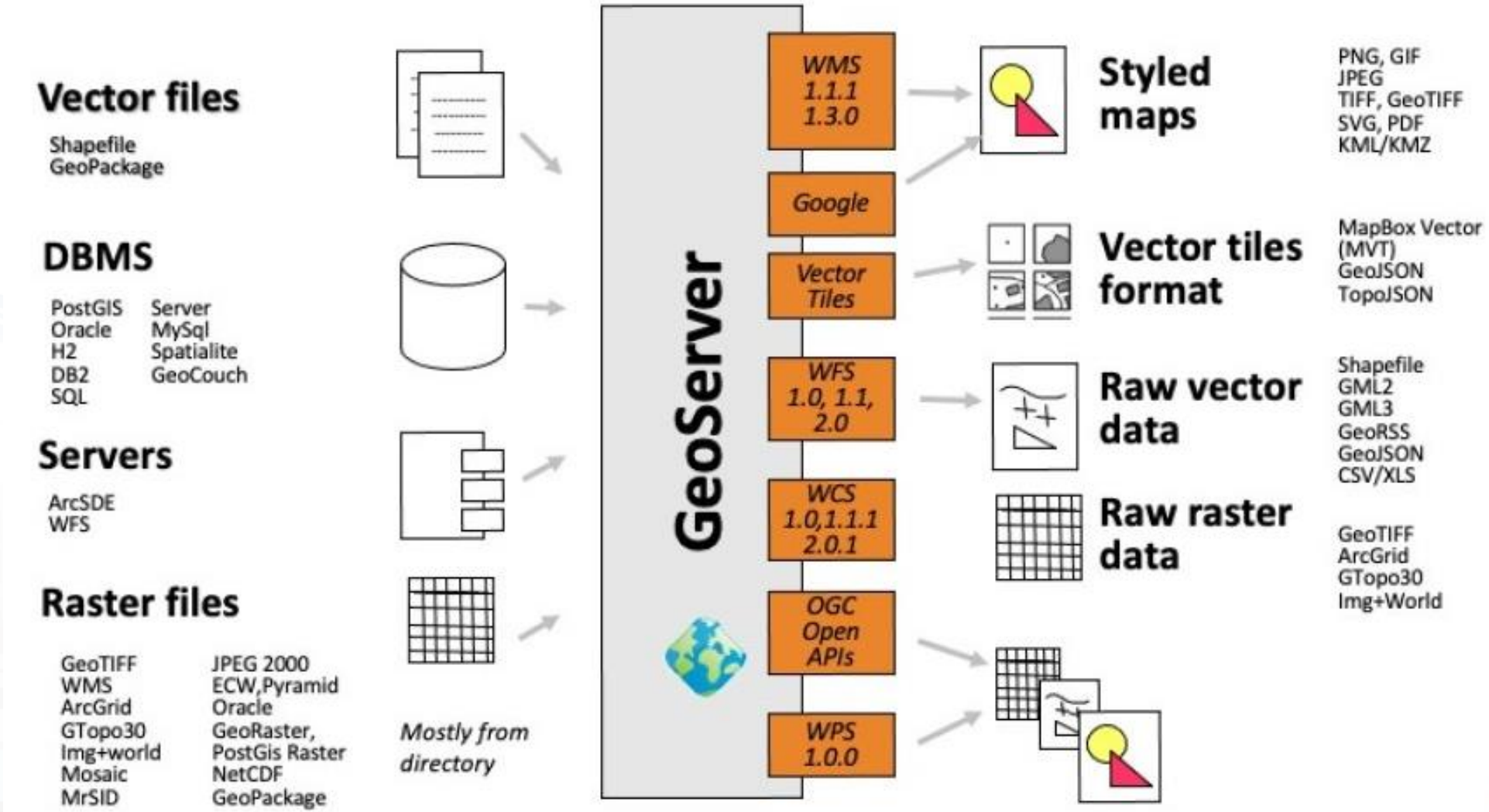
Traffic Flow reconstruction, real time

How it works: Traffic Flow Manager



The service of the GeoServer

GeoServer



ScienceDirect

ELSEVIER

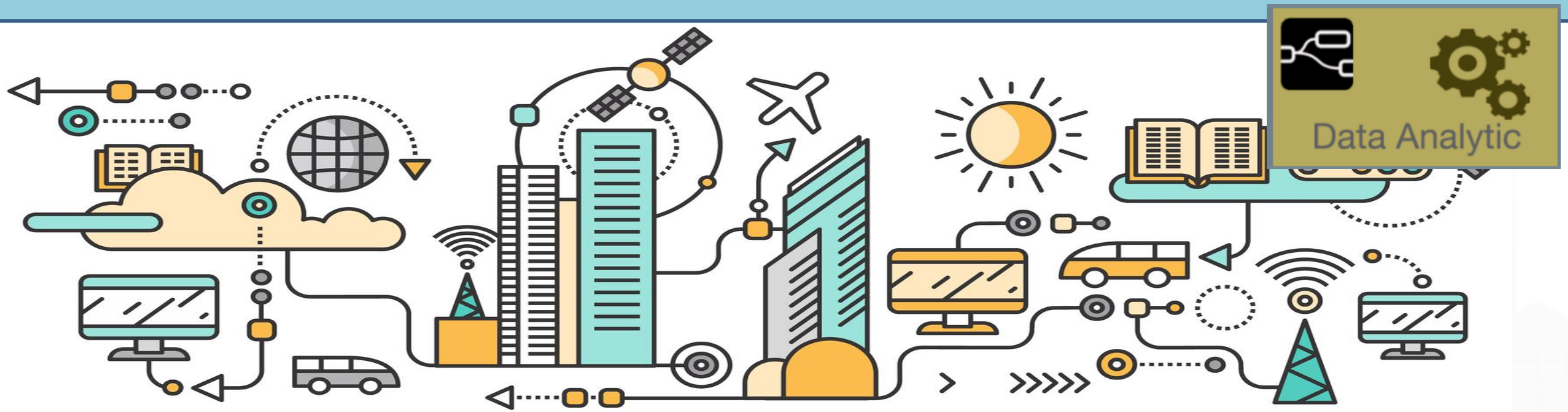
- Stefano Bilotta, Paolo Nesi,
- **Traffic flow reconstruction by solving indeterminacy on traffic distribution at junctions**, Future Generation Computer Systems, Volume 114, 2021, Pages 649-660, ISSN 0167-739X, https://doi.org/10.1016/j.future.2020.08.017.

https://www.sciencedirect.com/science/article/pii/S0167739X20308359

Traffic Flow Reconstruction (self training)

- P. Bellini, S. Bilotta, P. Nesi, M. Paolucci, M. Soderi, "Traffic Flow Reconstruction from Scattered Data", IEEE SMARTCOMP, IEEE international conference on smart computing, 18-20 June, Taormina, Sicily, Italy. 2018
- P. Bellini, S. Bilotta, P. Nesi, M. Paolucci, M. Soderi, "Real-Time Traffic Estimation of Unmonitored Roads", IEEE-DataCom'2018, Athens, 2018

COVID-19 vs other data: traffic and environment



Andamenti Nazionali e Regionali infezione COVID-19

Sulla base dei dati della protezione civile, elaborazioni DISITLab

Sun 19 Apr 19:21:39

per evidenziare gli andamenti di vostro interesse: eliminare le curve che non interessano selezionandole in legenda.
Alcuni dati in passato non sono pervenuti alla protezione civile

486	nuovi_att_positivi
108257	Attualmente_positivi
23660	deceduti
47055	dimessi_guariti

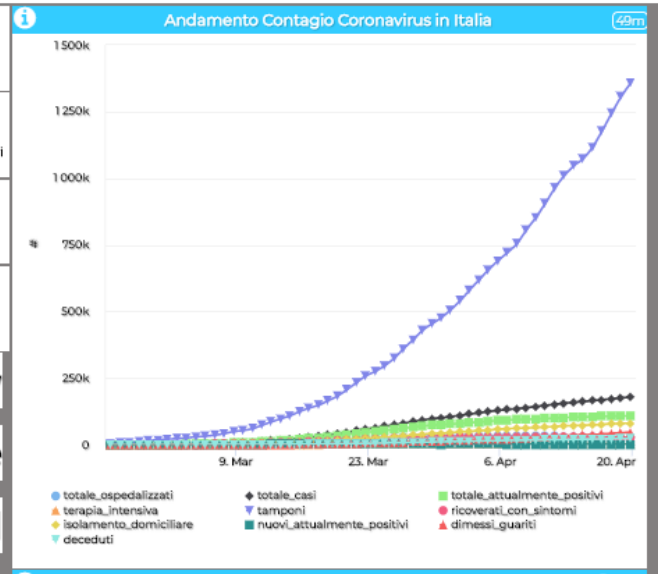
LOG view

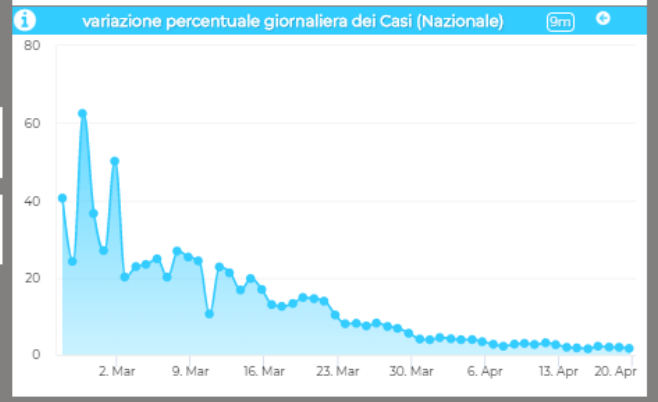
Mobile

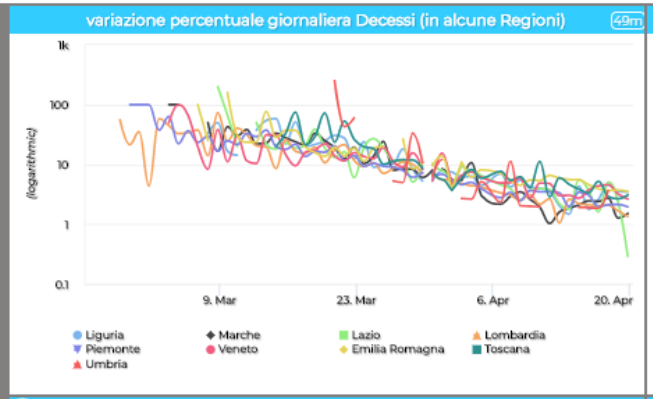
Staked

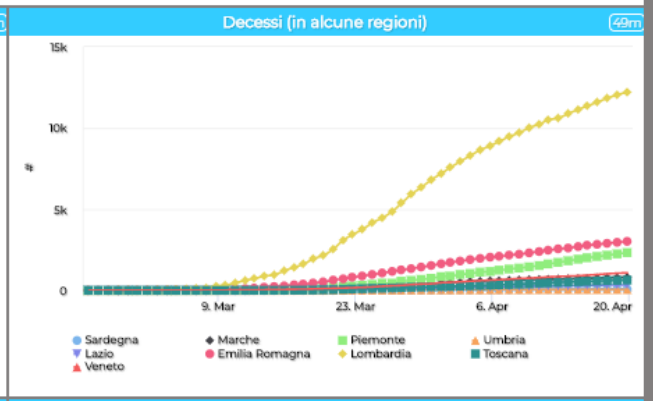
Regione Toscana

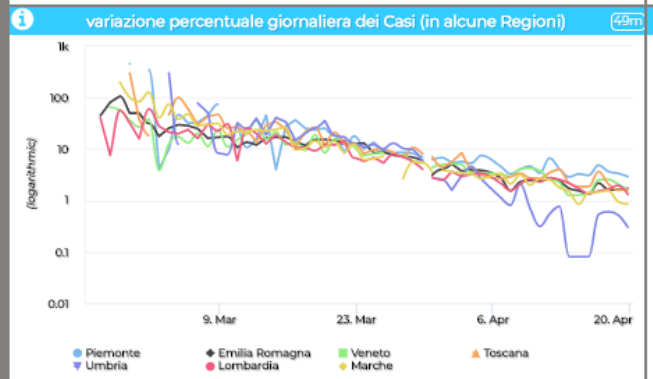
Regione LOG

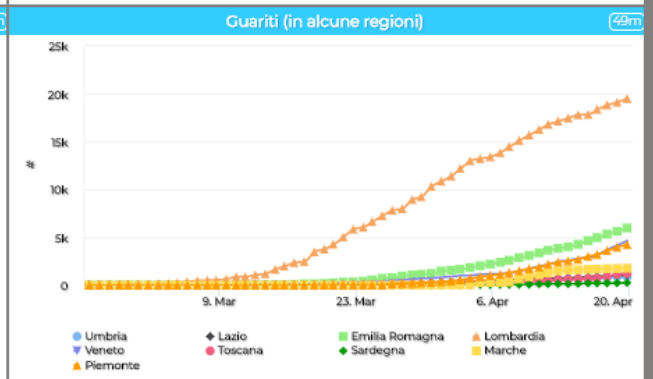


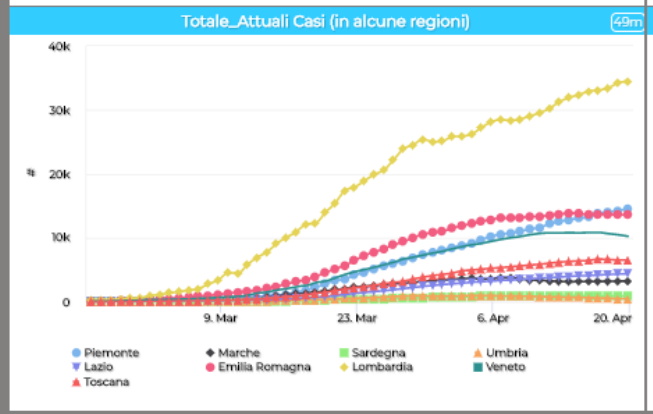


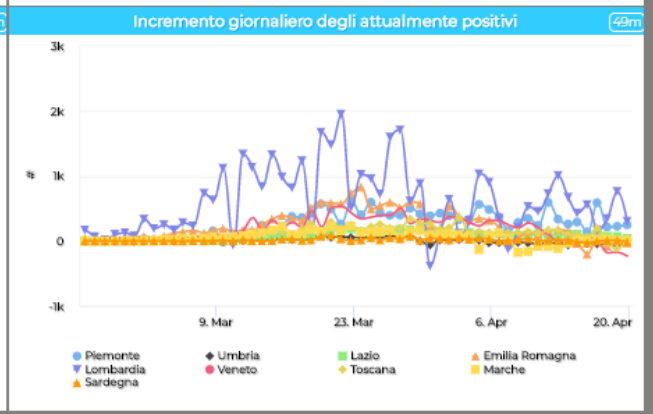












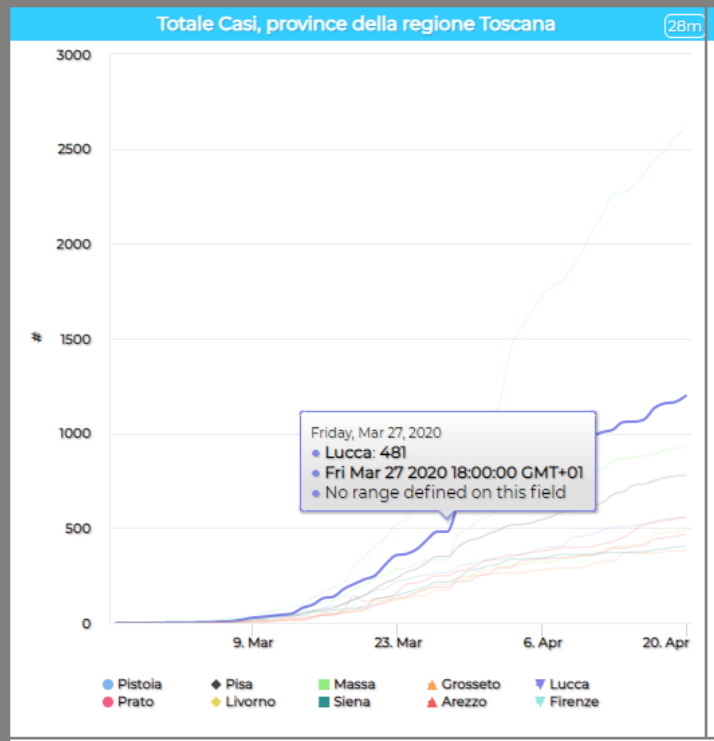
https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MjU2OQ==

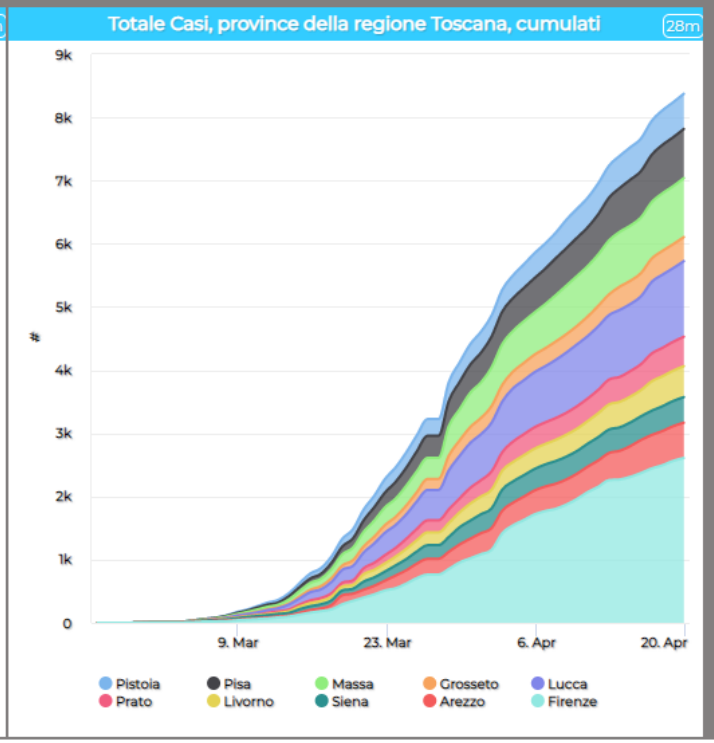
Main IT Provinces vs Tuscany Provinces: COVID-19

Andamento Regione Toscana e Province, COVID-19

Sulla base dei dati della protezione civile, elaborazioni DISITLab

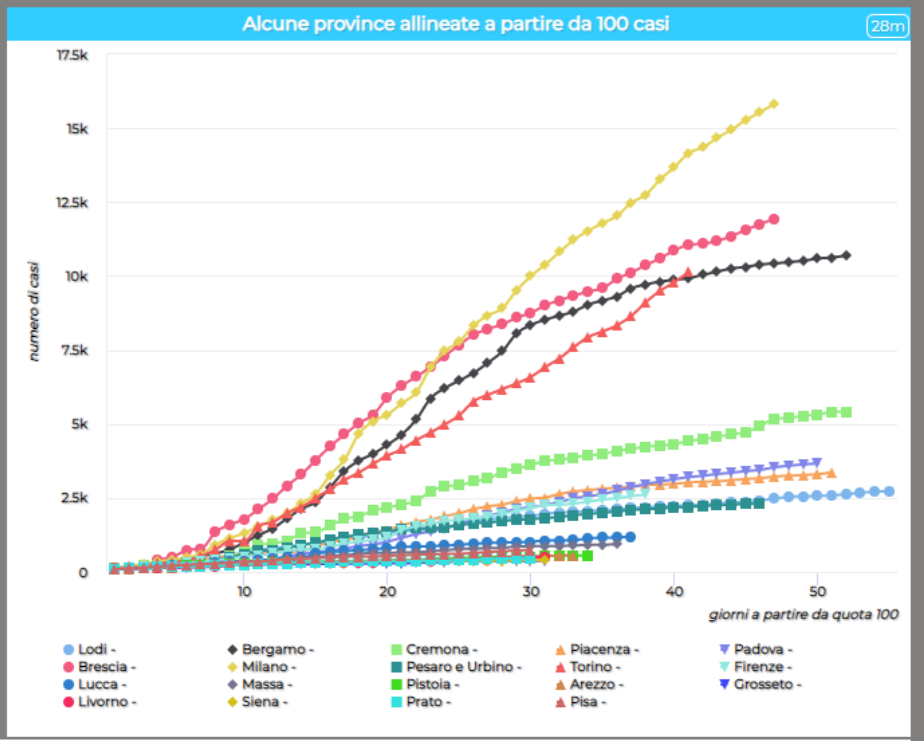
Sun 19 Apr 19:19:56





passa al Nazionale Regionale

Log view



per evidenziare gli andamenti di vostro interesse: eliminare le curve che non interessano selezionandole in legenda.

Alcuni dati in passato non sono pervenuti alla protezione civile

https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MjU3NQ==

Privacy Policy Cookies Policy Terms and Conditions Contact us

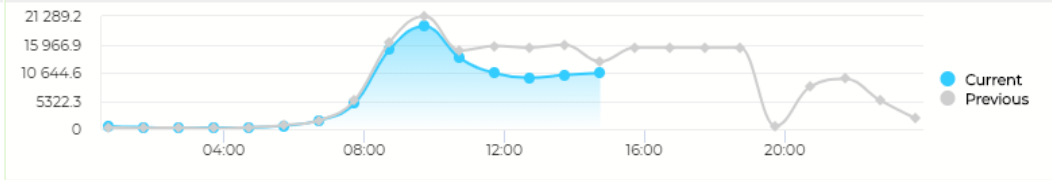
Traffic Flow Monitoring - Firenze - Cloned2

Wed 11 Nov 15:01:32

 # IN FLOW 9m

10549 #ofvehicles

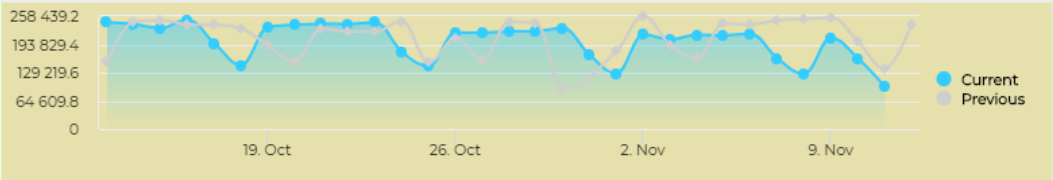
Firenze IN Traffic Flow (number of vehicles) 9m



Inc Daily Inp... 9m

97137 #ofvehicles

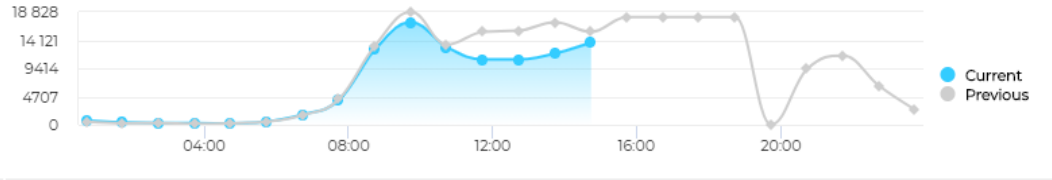
Daily Inputs (monthly) (last value is incremental, real time) 9m



 # OUT FLOW 9m

13720 #ofvehicles

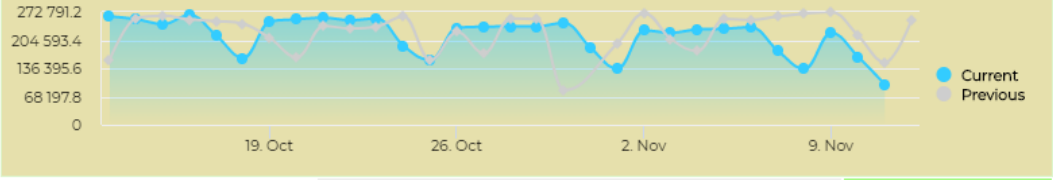
Firenze OUT Traffic Flow (number of vehicles) 9m



Inc Daily Out... 9m

97457 #ofvehicles

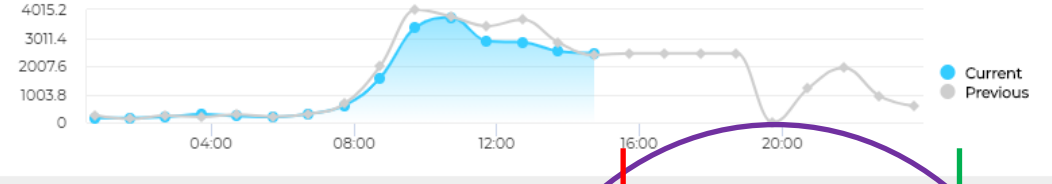
Daily Outputs (monthly) (last value is incremental real time) 9m



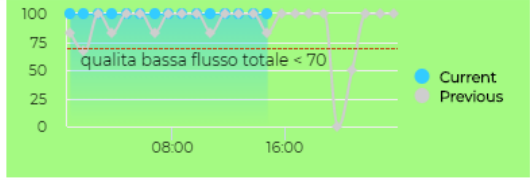
ZTL in 9m

2468 #ofvehicles

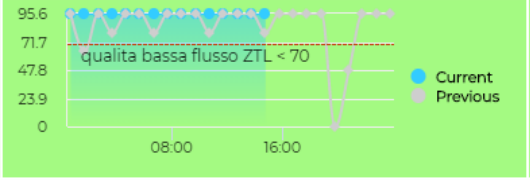
ZTL in Traffic Flow daily trend, entering in ZTL 9m



QoS as perc. of measures taken 9m



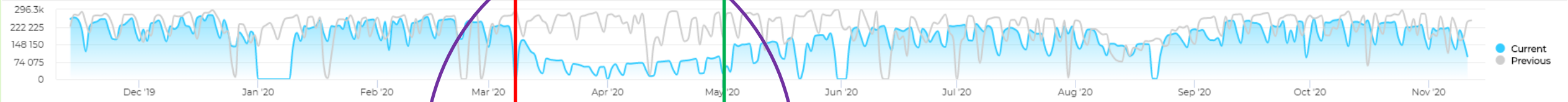
QoS as perc. of measures in ZTL 9m



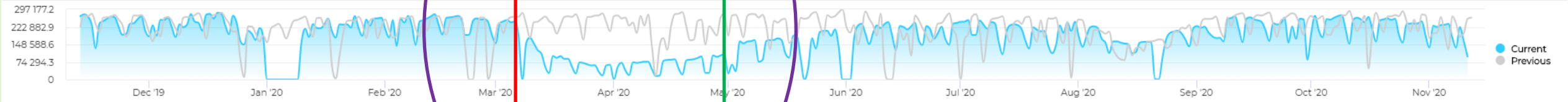
11/11/2020

15:01:33

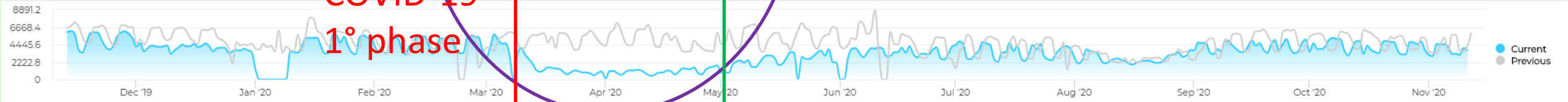
inflow total of the day, yearly 9m



outflow total over the day Yearly 9m



in ZTL yearly compare 9m



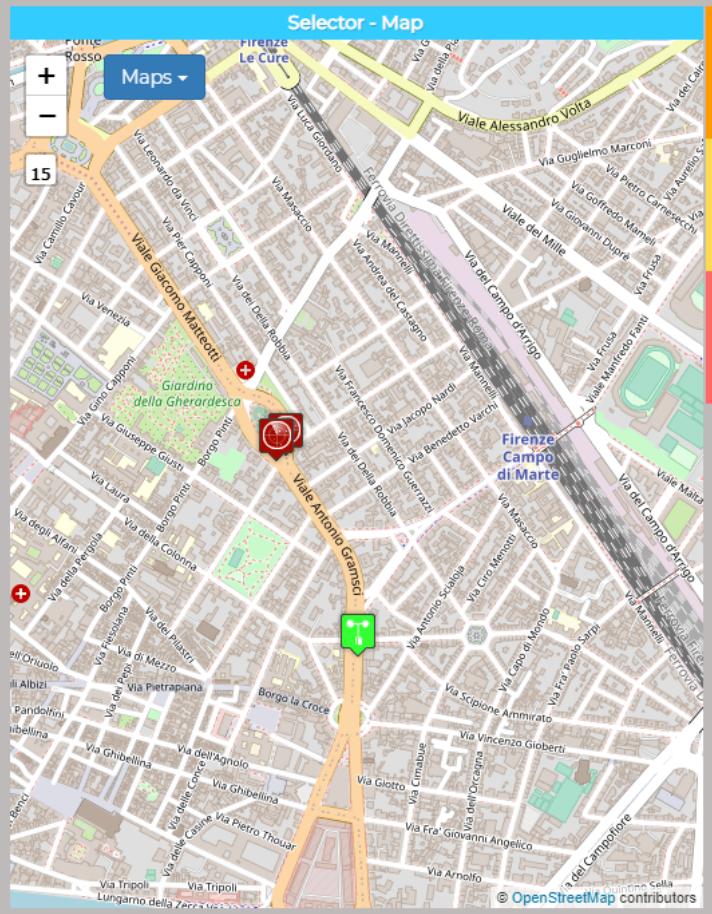
COVID-19

1° phase

NOX reduction for COVID

Monitoraggio Area Gramsci: NO2 vs Traffico

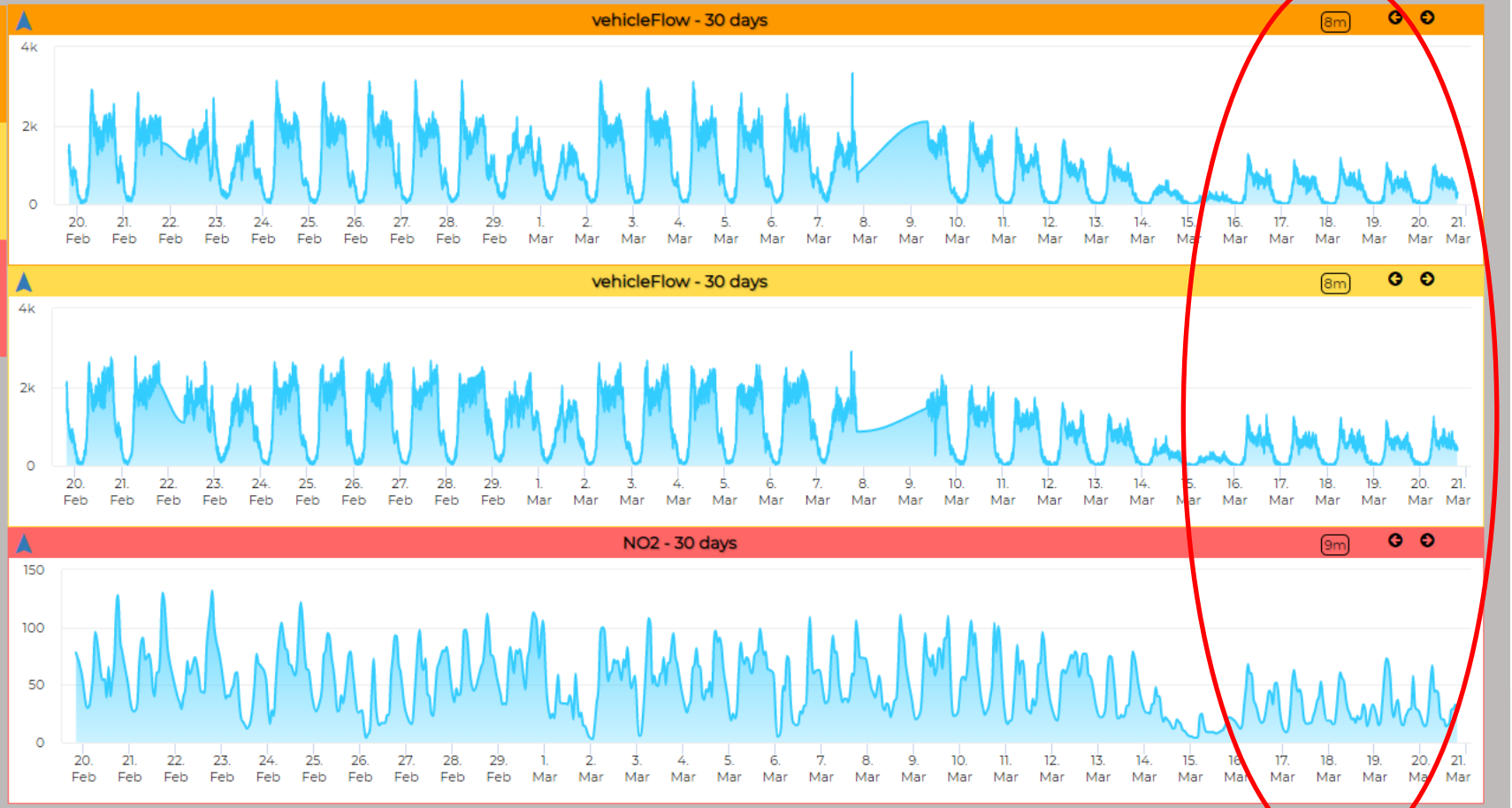
Sun 19 Apr 19:16:42



InOut traffic Firenze

NOX Predictions

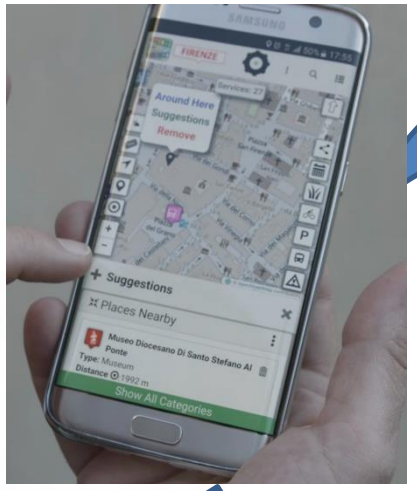
Real Time Traffic



https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MjI4OQ==

Quality of Public Transport

How much confident is the guess for bus arrival



Assessment and prediction

Customer satisfaction

Firenze Oggi

Sun 20 Oct 23:35:33

26976
Totale utenti WIFI

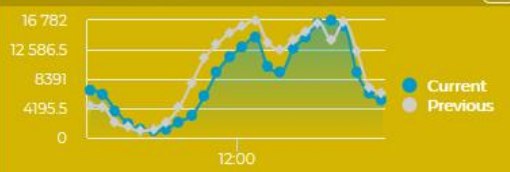
COLONNINE RICARICA... 9m
176 INSTALLATE
71 % ATTIVE
5.1 % IN USO

9m
GENERAL | METEO
MINIMO | BASSO | MEDIO | ALTO
RISCHIO IDRAULICO
RISCHIO TEMPORALI
RISCHIO IDROGEOLOGICO
RISCHIO NEVE
RISCHIO GHIACCIO
RISCHIO VENTO

SITUAZIONE VIABILITA 55s
0 INCIDENTI
0 CHIUSURE AL TRAFFICO (TOT)
0 CHIUSURE PER CANTIERI
0 PROGR. | 0 NON PROG.
0 LIMITAZIONI AL TRAFFICO (TOT)
0 LIMITAZIONI PER CANTIERI
0 NON PROG. | 0 PROGR.
0 TOT. EVENTI SULLA RETE

SMN 9m	BINARIO16 9m	FORTEZZA 9m
21.6 % occupati su 607 posti	43 % occupati su 165 posti	19.2 % occupati su 521 posti
LEOPOLDA 9m	**CALZA 9m**	**S.AMBROGIO 9m**
34 % occupati su 300 posti	39.2 % occupati su 148	21.6 % occupati su 379 posti
PARTERRE 9m	**CAREGGI 9m**	**BECCARIA 9m**
31.1 % occupati su 656 posti	4.4 % occupati su 406 posti	23.3 % occupati su 210 posti

ANALYSIS
Energy | Environment | Mobility | Social | Resilience

FLUSSI INGRESSO CITTA 9m


TOTALE 9m
176560
VEICOLI

FLUSSI INGRESSO ZTL 9m

TOTALE ZTL 9m
47368
VEICOLI

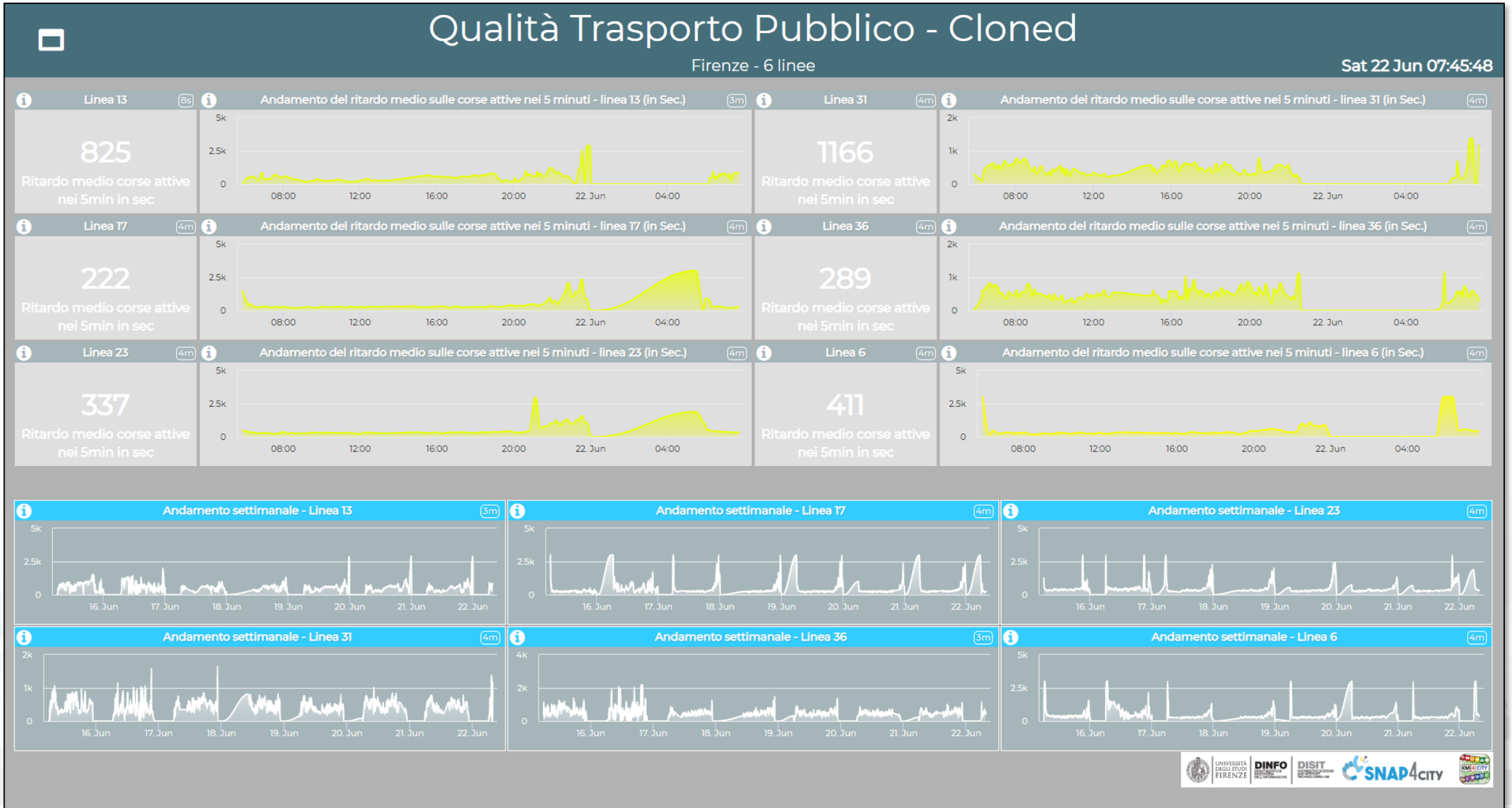
Nati Italiani 119m	Nati stranieri 119m	Deceduti 119m	Matrimoni 119m	Unioni Civili 119m
164 ultimo mese consolidato	57 ultimo mese	399 ultimo mese	18 ultimi 7 giorni	0 ultimi 7 giorni

Segnalazioni ricevute in attesa 119m	In Lavorazio... 119m	Risolte 119m	Chiuse senza risoluzion... 119m
1116 ultimo mese	524	305	285

Manutenzioni Stradali 59m	Verde Pubbl... 59m	Decoro Urbano 59m	Relitti 59m
6 oggi	3	5	0

Attesa media alla fermata

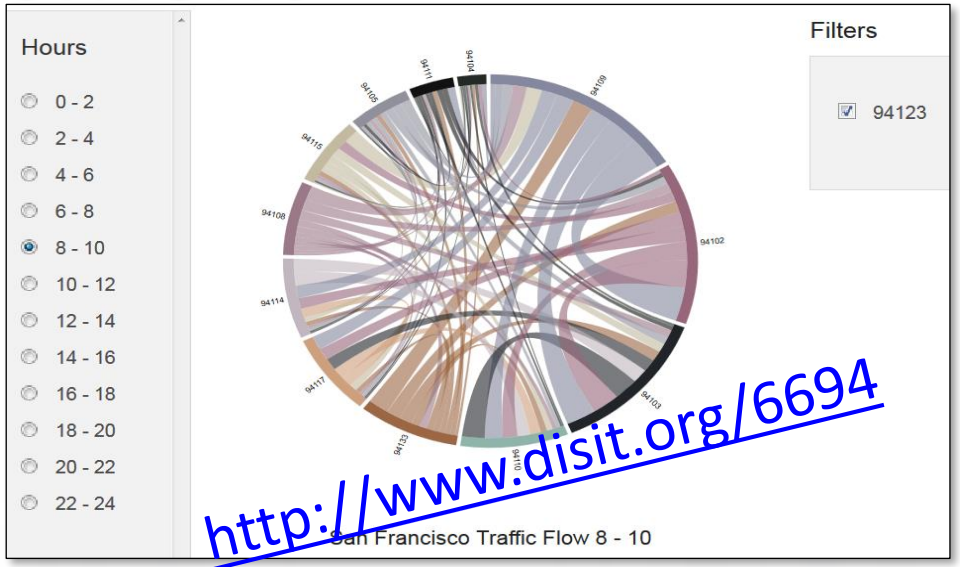
Linea 6 9m	Linea 13 9m
3 min	13 min
Linea 17 9m	**Linea 23 9m**
4 min	5 min
Linea 31 9m	**Linea 36 9m**
19 min	2 min


Qualità Trasporto Pubblico - Cloned
Firenze - 6 linee
Sat 22 Jun 07:45:48
Linea 13
825
Ritardo medio corse attive nei 5min in sec
Andamento del ritardo medio sulle corse attive nei 5 minuti - linea 13 (in Sec.)
Linea 31
1166
Ritardo medio corse attive nei 5min in sec
Andamento del ritardo medio sulle corse attive nei 5 minuti - linea 31 (in Sec.)
Linea 17
222
Ritardo medio corse attive nei 5min in sec
Andamento del ritardo medio sulle corse attive nei 5 minuti - linea 17 (in Sec.)
Linea 36
289
Ritardo medio corse attive nei 5min in sec
Andamento del ritardo medio sulle corse attive nei 5 minuti - linea 36 (in Sec.)
Linea 23
337
Ritardo medio corse attive nei 5min in sec
Andamento del ritardo medio sulle corse attive nei 5 minuti - linea 23 (in Sec.)
Linea 6
411
Ritardo medio corse attive nei 5min in sec
Andamento del ritardo medio sulle corse attive nei 5 minuti - linea 6 (in Sec.)
Andamento settimanale - Linea 13
Andamento settimanale - Linea 17
Andamento settimanale - Linea 23
Andamento settimanale - Linea 31
Andamento settimanale - Linea 36
Andamento settimanale - Linea 6

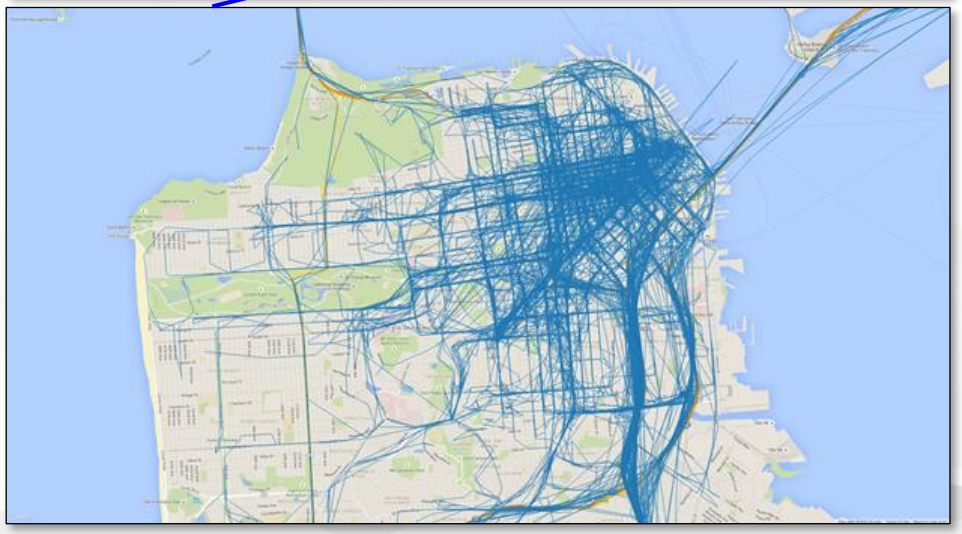
Origin Destination Matrices

Traffic and People Flow Assessment

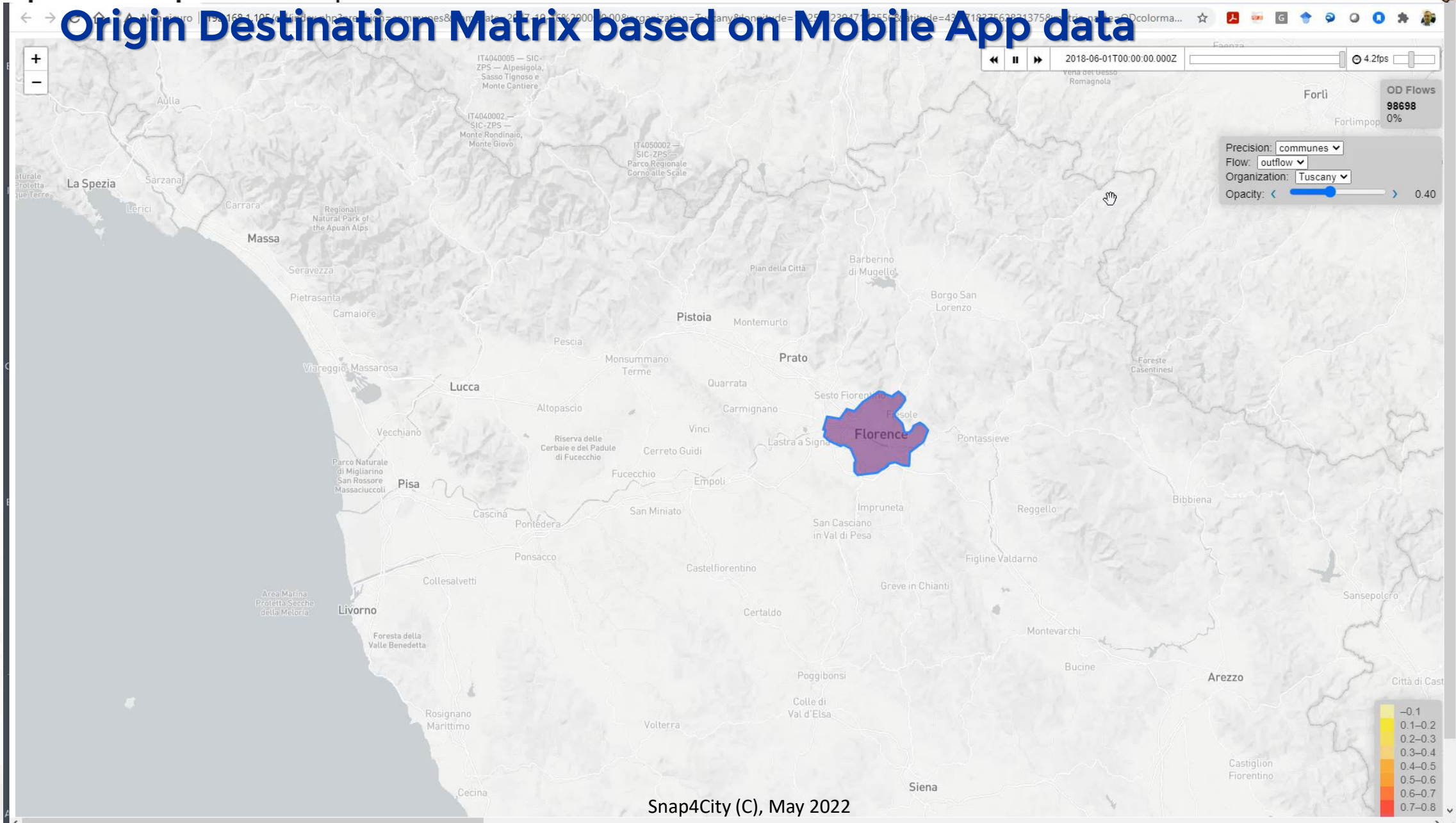
- **Origin Destination Matrix**
  - Specific Sensors, vehicle Kits, mobile App, Wi-Fi Access Points, etc.
  - Data from Taxi in San Francisco
  - Census data
  - Telecom Operator data
- **Assess people and traffic flows to**
  - improve services
  - predict critical conditions on Crit. Infra.
  - take real time decisions and sending messages in push to population
  - Increase city resilience
  - optimize traffic flow
  - take decision of routing



http://www.disit.org/6694

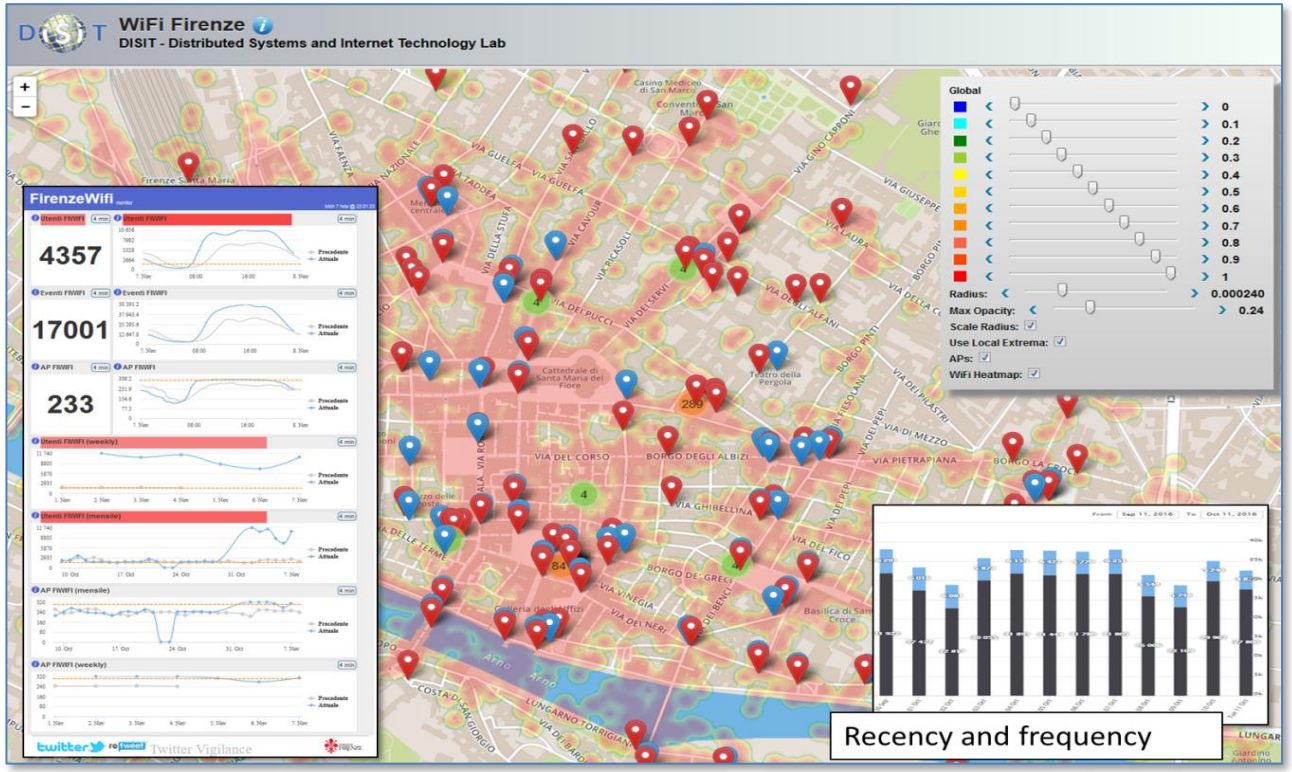


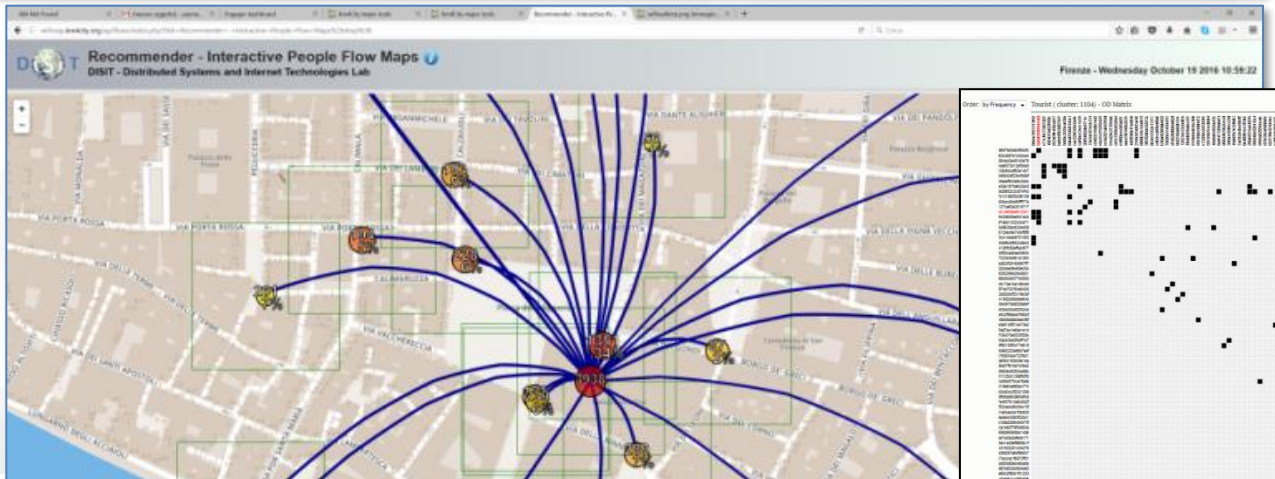
Origin Destination Matrix based on Mobile App data

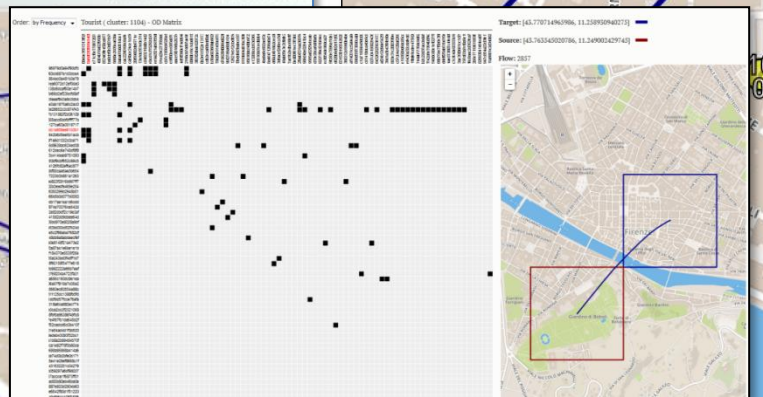


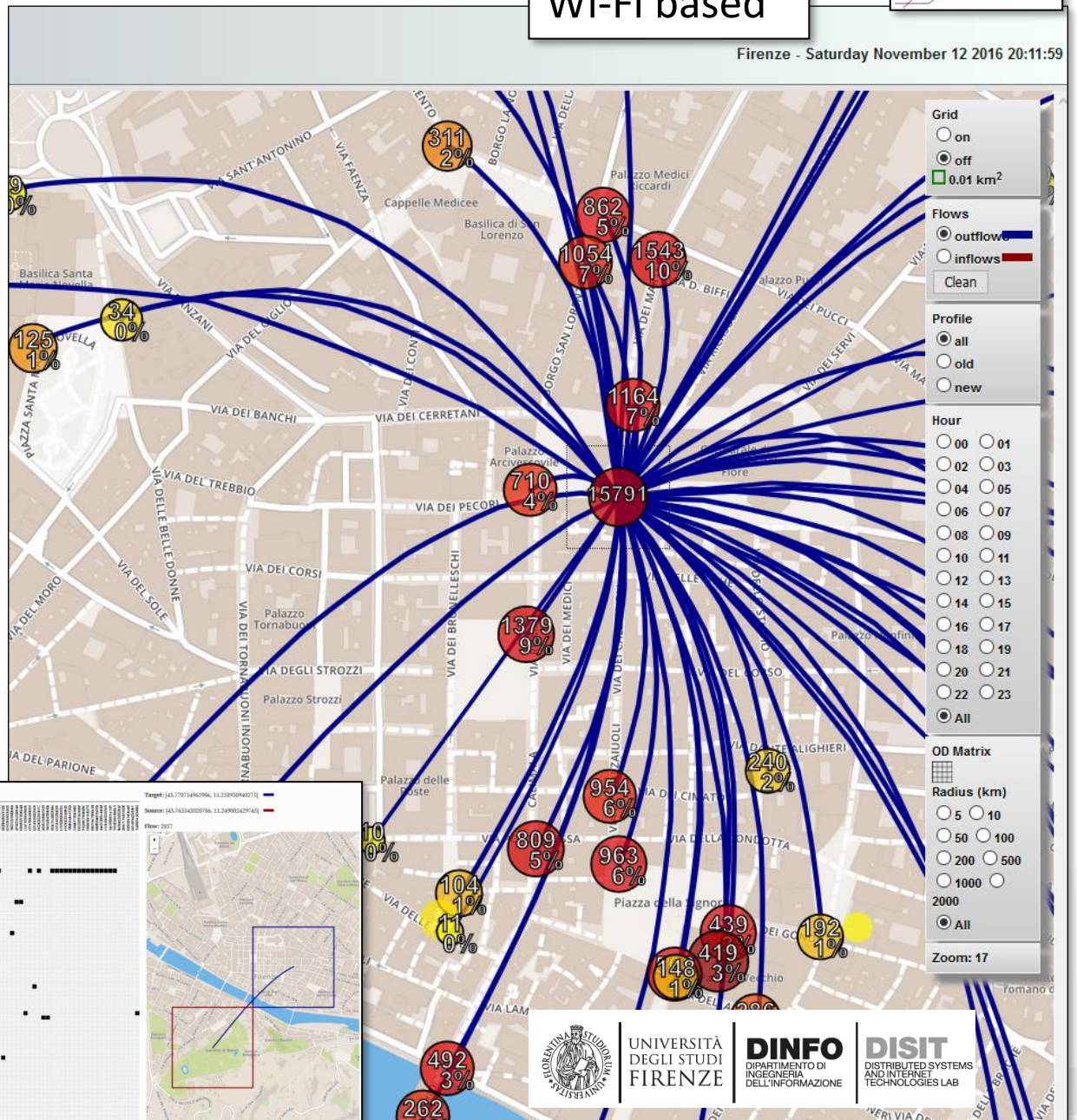
Origin Destination Matrix Estimation

Wi-Fi based







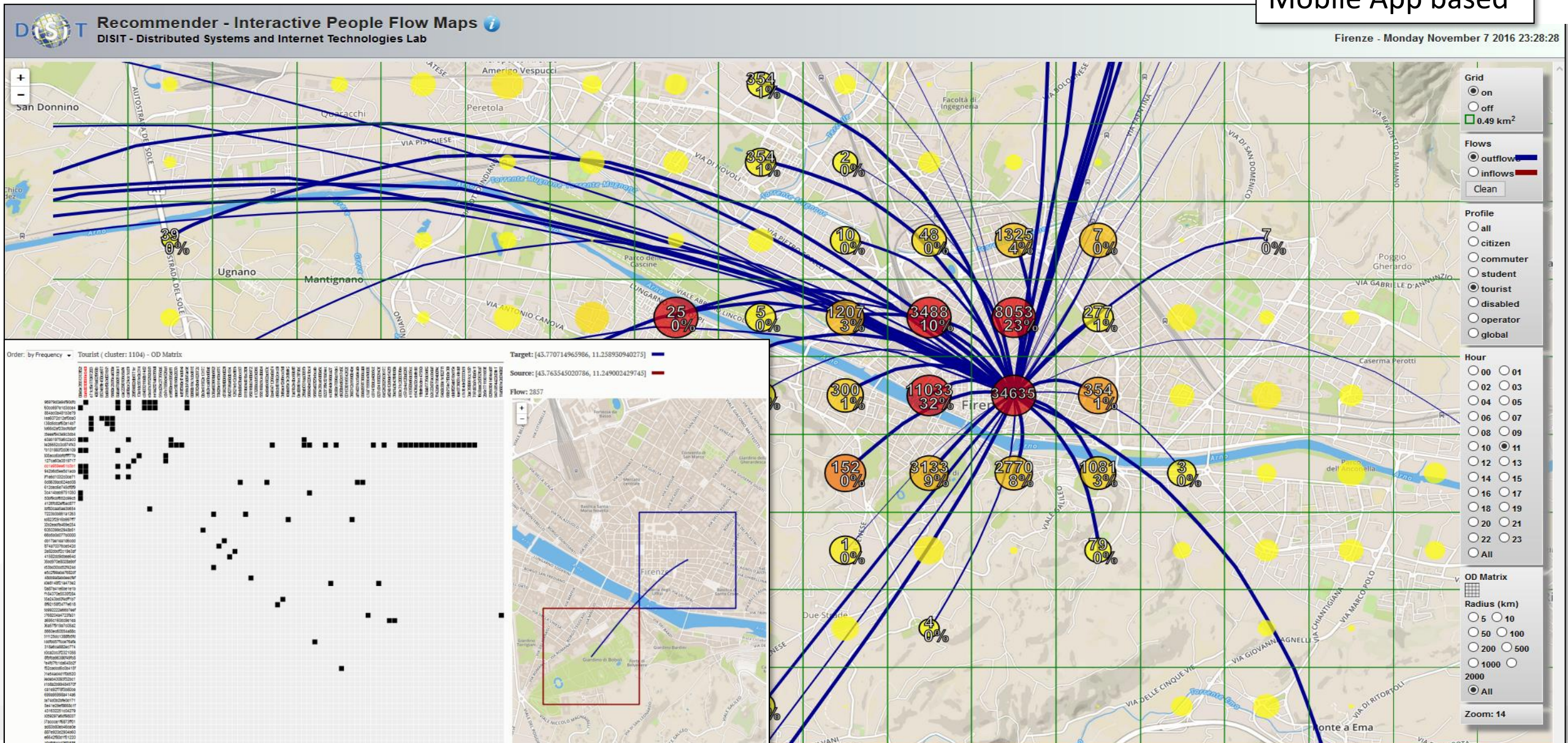


UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB

Scalable multiresolution OD matrix

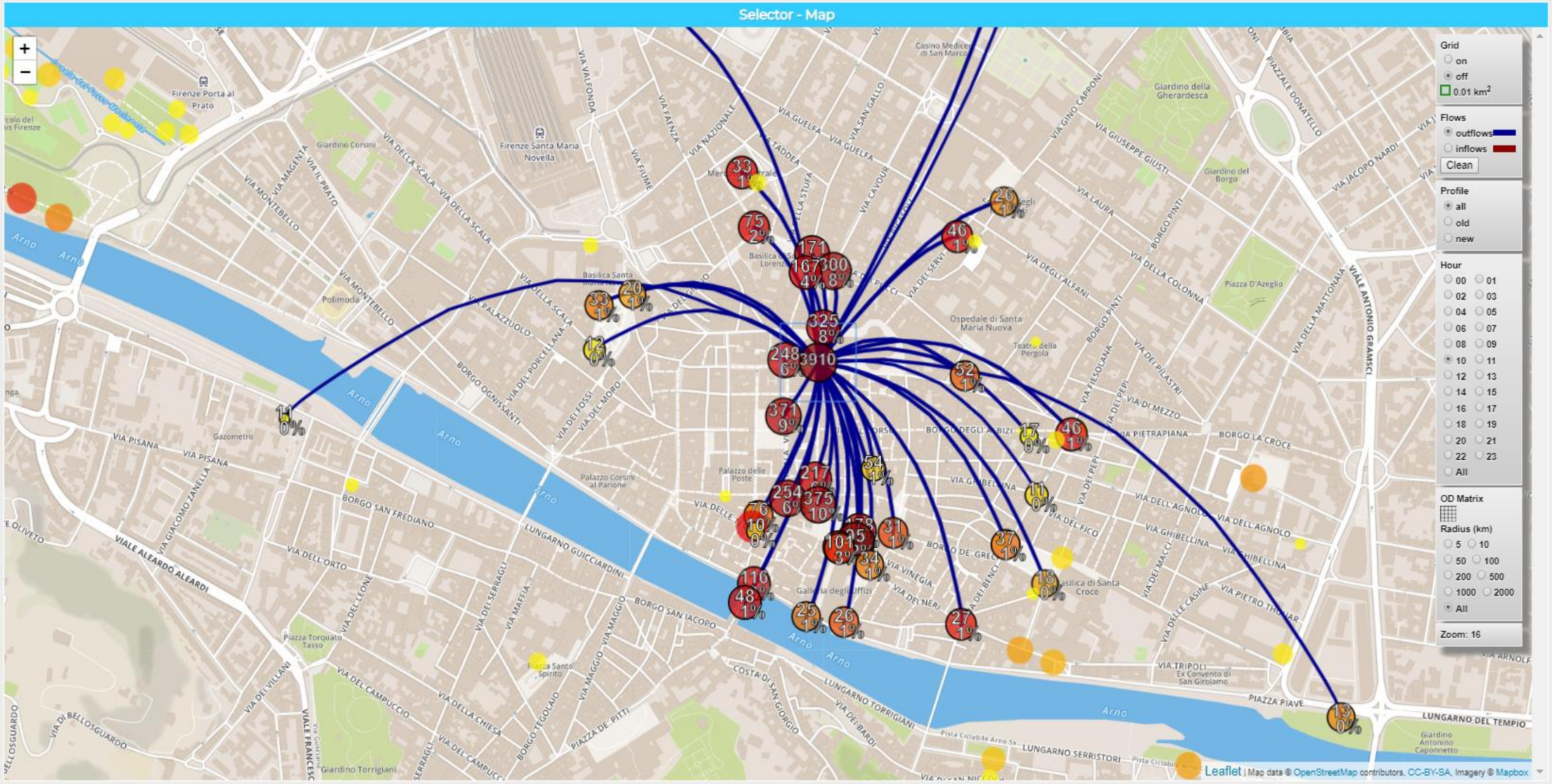
Sii-Mobility

Mobile App based



Life in Toscana: Dashboard

Sun 20 Oct 23:44:05

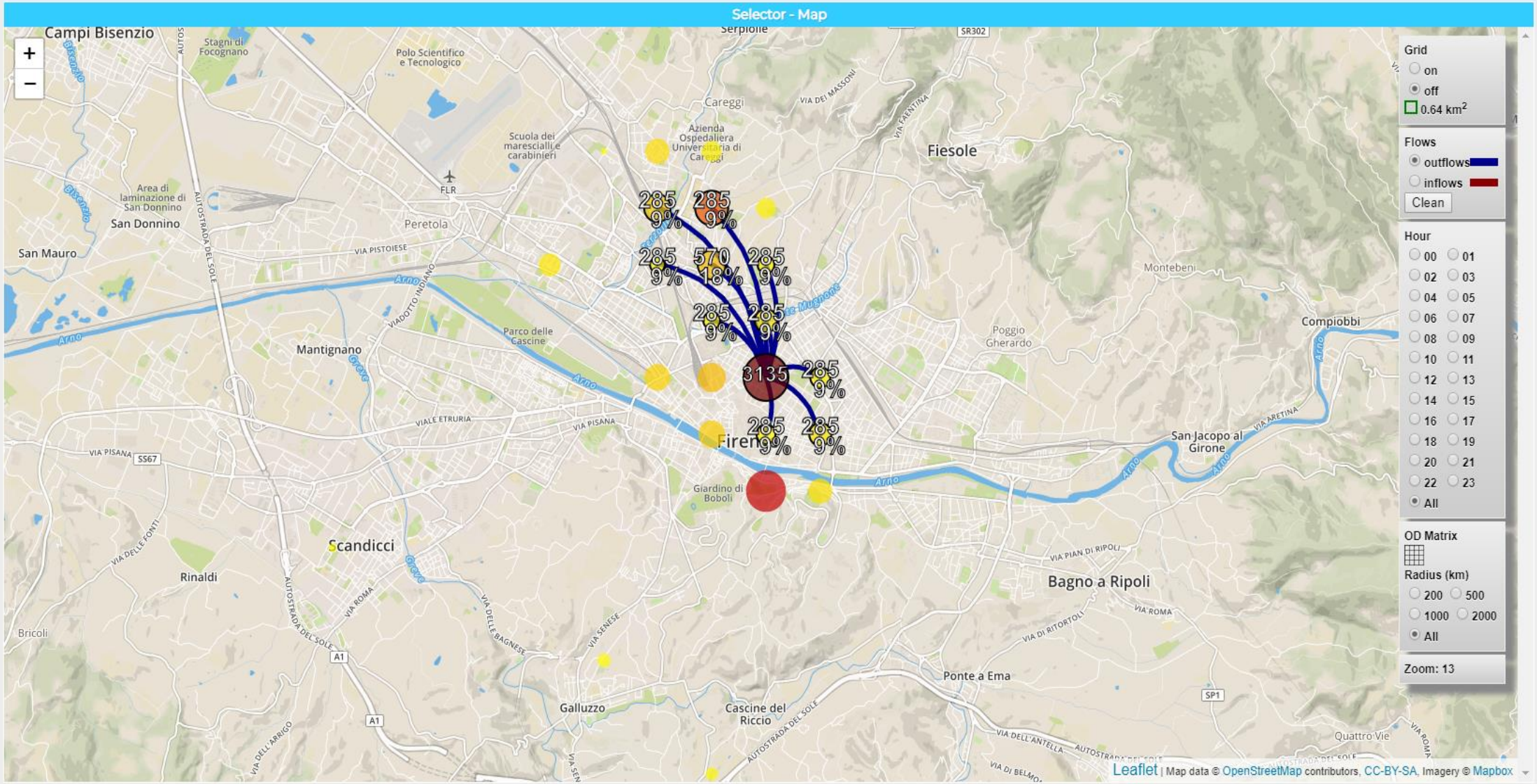


- Line of Transport
- Public Transport
- Travel Plan
- Traffic Flow FIPILI
- Air Quality
- Weather
- Origin Dest. Matrix
- Typical Trajectories
- People Flow WIFI
- People Flow KM App
- Cultural Activities
- Forum Discussion
- CAM
- Ponte Vecchio
- Real Time Busses

Main

Life in Toscana: Dashboard

Sun 20 Oct 23:40:25



- Line of Transport
- Public Transport
- Travel Plan
- Traffic Flow FIPILI
- Air Quality
- Weather
- Origin Dest. Matrix
- Typical Trajectories
- People Flow WIFI
- People Flow KM App
- Cultural Activities
- Forum Discussion
- CAM
- Ponte Vecchio
- Real Time Busses

Main

The Life of Antwerp

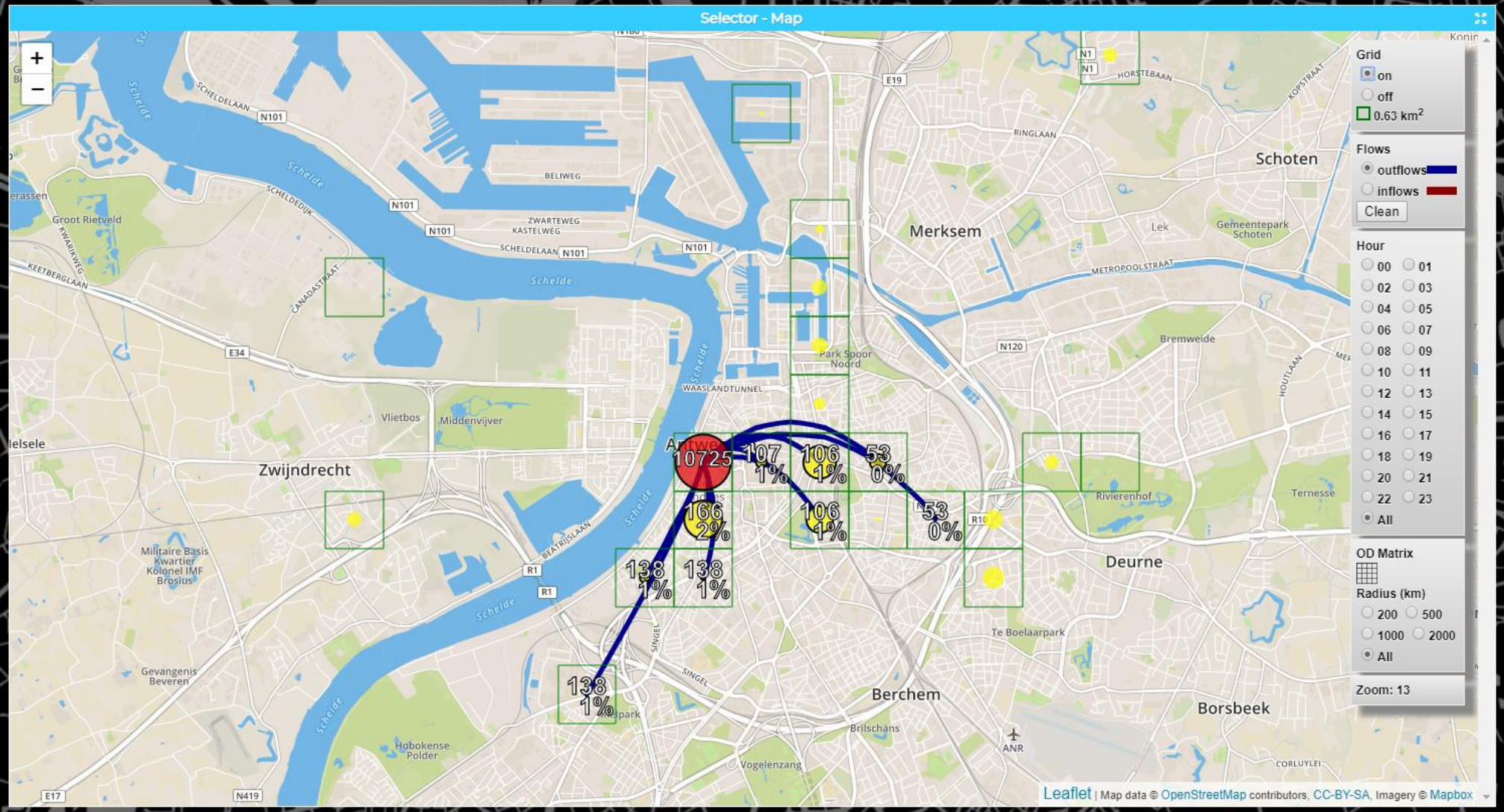
Please note that the data results are not always based on real data.

Sun 20 Oct 23:42:07

- Shopping: POI
- Wine and Food: POI
- Discovery Antwerp
- Point of Interest
- 3D view POI
- Routing on Antwerp
- Line of Transport
- Public Transport
- Safe On Bike
- Tunnel and Ferry
- Twitter Vigilance
- Twitter Vig. Real Time
- Air Quality
- Weather
- Origin Dest. Matrix
- Typical Trajectories
- Forum Discussion

Documentation

Survey



https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MTcwNg==

Modal & Multimodal Routing for Navigation and Travel Planning

Routing

- **Routing**: From XX to YY, Travel means:
  - private as cars, bike, pedestrian, ..
  - Public: busses, tramway, train, etc.
- **Multimodal routing**: public means, pedestrian, etc.
- **Taking** into account:
  - Multiple intermediate points
  - Constraints/preferences:
    - size of roads, pollutant, traffic, obstacle/barriers, noise
    - Limitations on paths per vehicle kind
- **Dynamic Routing** enabling the addition of constraints on the user interface. For example: *barriers*

User: roottooladmin1, Org: DISIT
Role: RootAdmin, Level: 7

- Dashboards
- My Dashboards
- Notificator
- IOT Applications
- My Personal Data
- IOT Directory and Devices
- Knowledge and Maps
  - Service Map
  - Loading WKT on Service Map
  - Creating WKT
  - Service Map 3D
  - Helsinki Service Map
  - Antwerp Service Map
  - My Annotation on Services/Data
  - Mapping Services Data
  - ArcGIS DISIT Service
- Micro Applications
- External Services
- Data Set Manager: Data Gate
- Resource Manager: Process Loader
- Development Tools
- Management
- Settings
- User Management and Auditing
- Help and Contacts
- Documentation and Articles
- My Profile
- Snap4City portal
- Km4City portal
- DISIT Lab portal

- Hide Menu

Public transport | Municipalities | Text Search | Address Search | Events

Select an agency:
- Select an Agency -

Select a line:
- Select a Line -

Select a route:
- Select a Route -

Select a bus stop:
- Select a Stop -

Position of selected Busses

Actual Selection

Coord: 43.7130,10.9272
Address: VIA DI PRATOVECCHIO, 58, EMPOLI

Path from here | Path to here | Search geometry

Path

From: VIA DELLA PERGOLA, 39, FIRENZE
To: VIA DI PRATOVECCHIO, 58, EMPOLI
Route via: public_transport
Start date&time: today now

Search Path

3. Piazza della Santissima Annunziata 49m (14:40:13)
4. Via Cesare Battisti 148m (14:40:52)
5. Piazza San Marco 126m (14:42:47)
6. Via Camillo Cavour 1m (14:44:17)
7. 11 : Arazzieri - Volterrana 7800m (15:06:00)
8. 37 : Galluzzo Via Volterrana - Montespertoli (V. Risorgimento) Sn 26620m (18:52:00)
9. Viale Risorgimento 207m (19:45:00)
10. 32 : Montespertoli (V. Risorgimento) Sn - Viasanzio Fr.157 Sn 17534m (07:16:00)
11. 1 : Via Sanzio Fr.157 - Via Sanzio Fr. Coop Sn 1002m (08:08:00)
12. Via Raffaello Sanzio 45m (08:10:00)
13. nd 33m (08:10:35)

- Hide Menu

Regular Services | Transversal Services

Services Categories
De/Select All
- Accommodation
- Advertising
- AgricultureAndLivestock
- CivilAndEdilEngineering
- CulturalActivity
- EducationAndResearch
- Emergency
- Entertainment
- Environment
- FinancialService
- GovernmentOffice
- HealthCare
- IndustryAndManufacturing
- IoTDevice
- MiningAndQuarrying
- ShoppingAndService
- TourismService
- TransferServiceAndRenting
- UtilitiesAndSupply
- Wholesale
- WineAndFood

Filter:
search text into service
Service providing value type:
select value type
N. results: 100

Search Range 100 mt
Search Area
select...

Multimodal routing

Leaflet | Map data © OpenStreetMap contributors, CC-BY-SA, Imagery © Mapbox

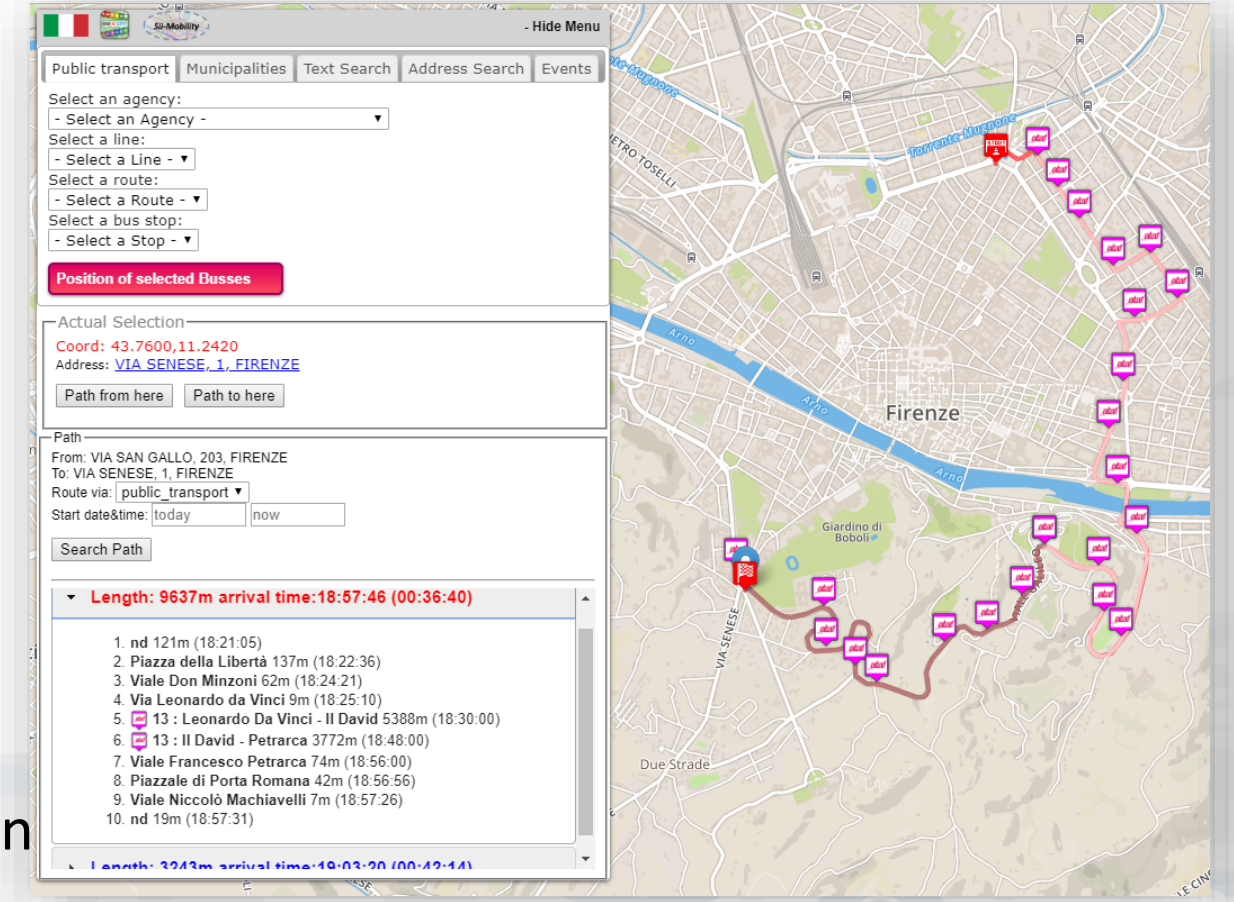
Routing and Multimodal Routing

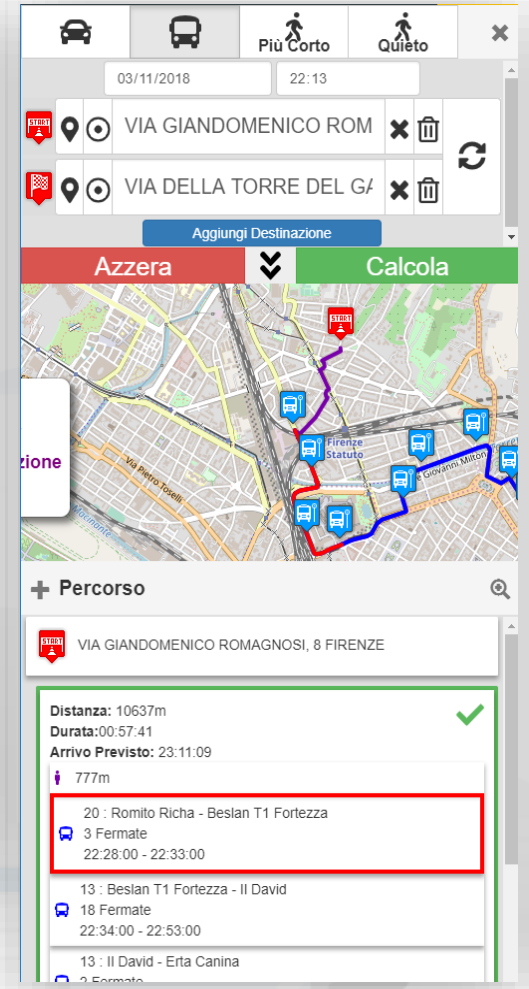
Modes:

- Pedonal, Vehicles
- Public Multimodal
- Multi Point for Delivering
- Constrained: quite, blocked, etc.

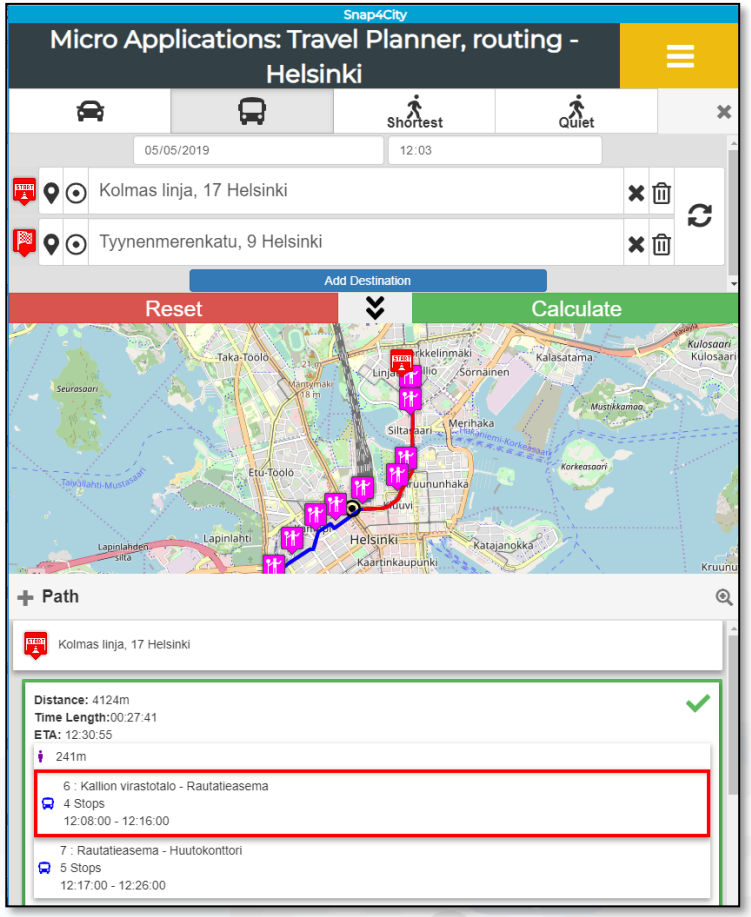
Test it on our:

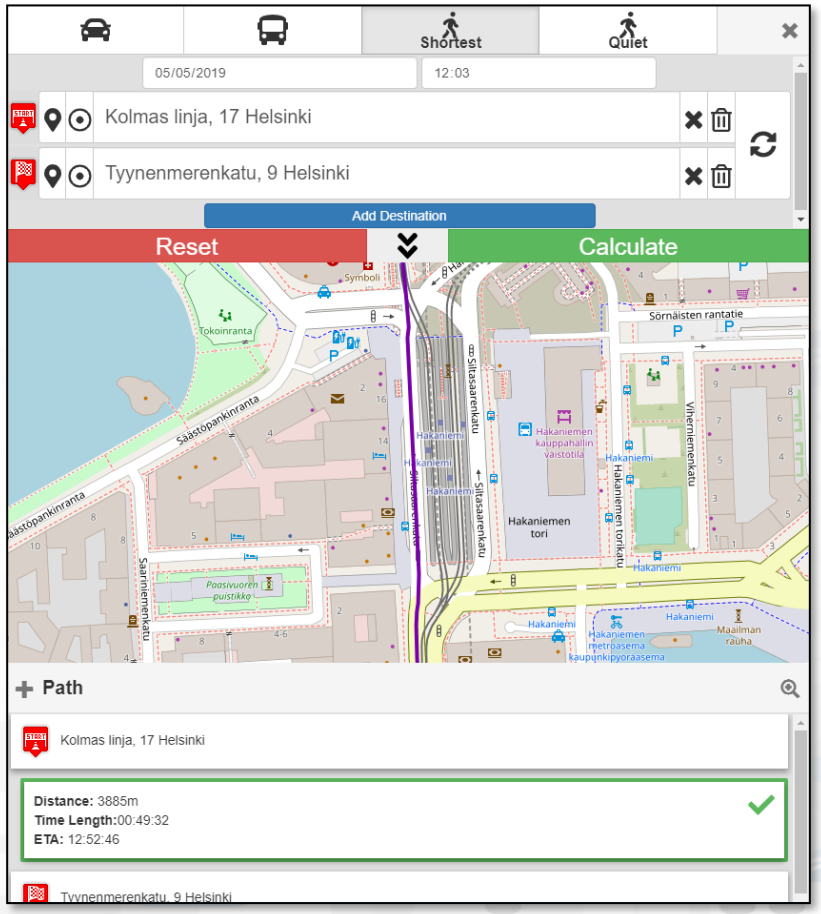
- Mobile Apps
- MicroApplication
- Dashboard
- ServiceMap service on Tuscany in Snap4City

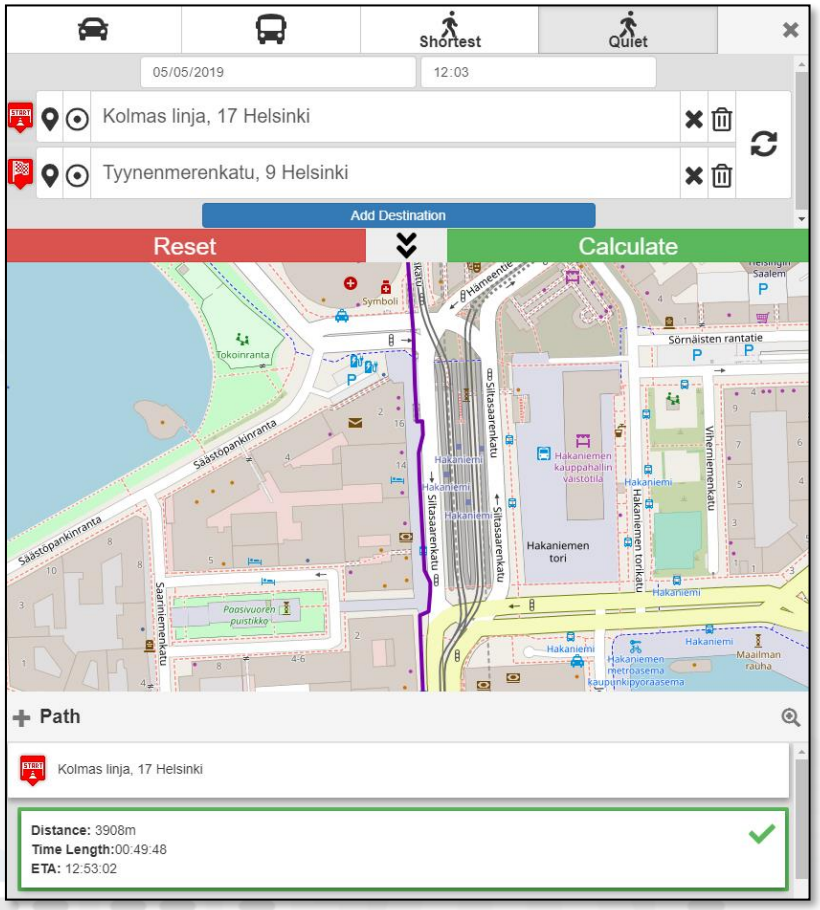




Routing for quite



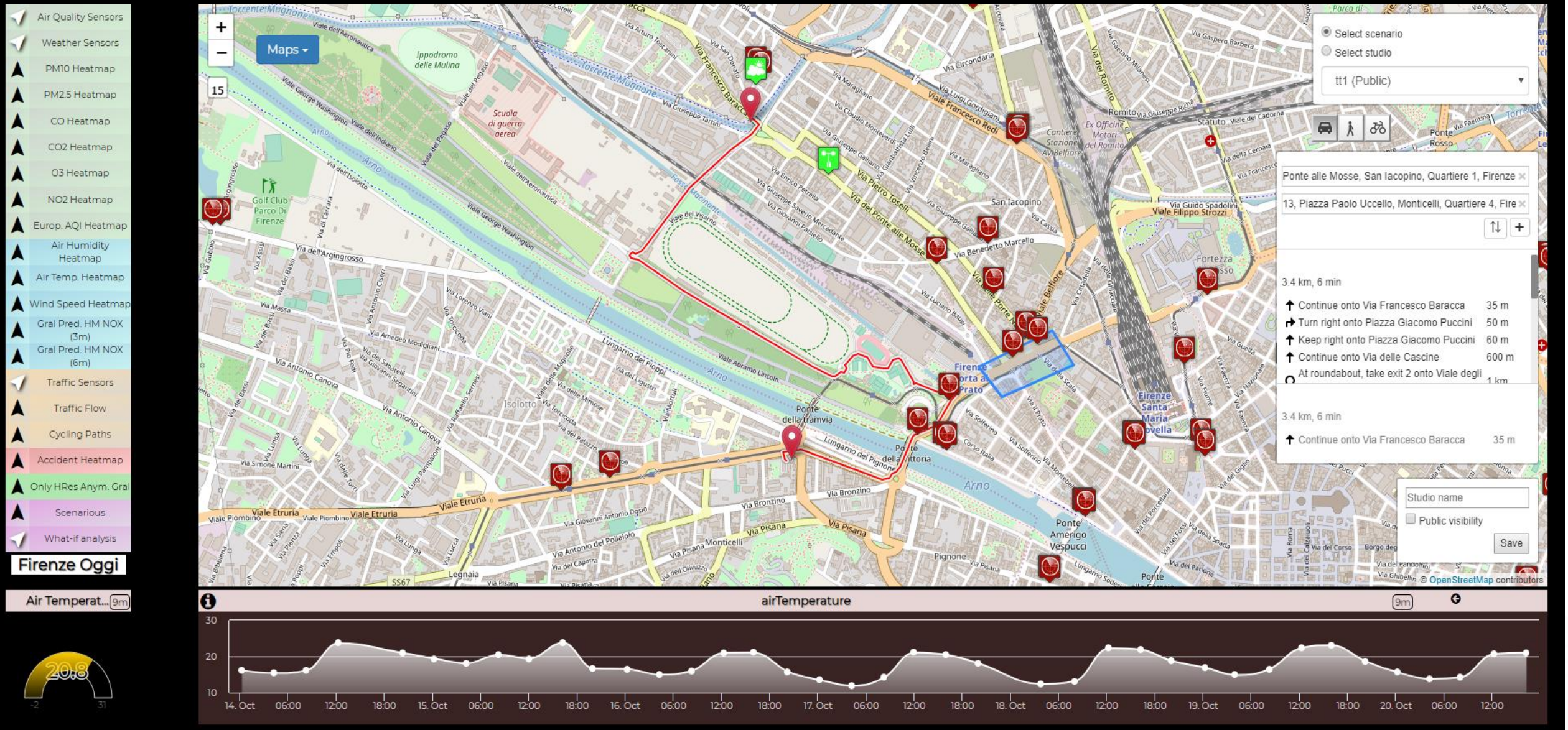




Mobility and Environment What-IF Analysis

This dashboad contains data derived from actual sensors and predictive values under validation

Sun 20 Oct 23:50:38

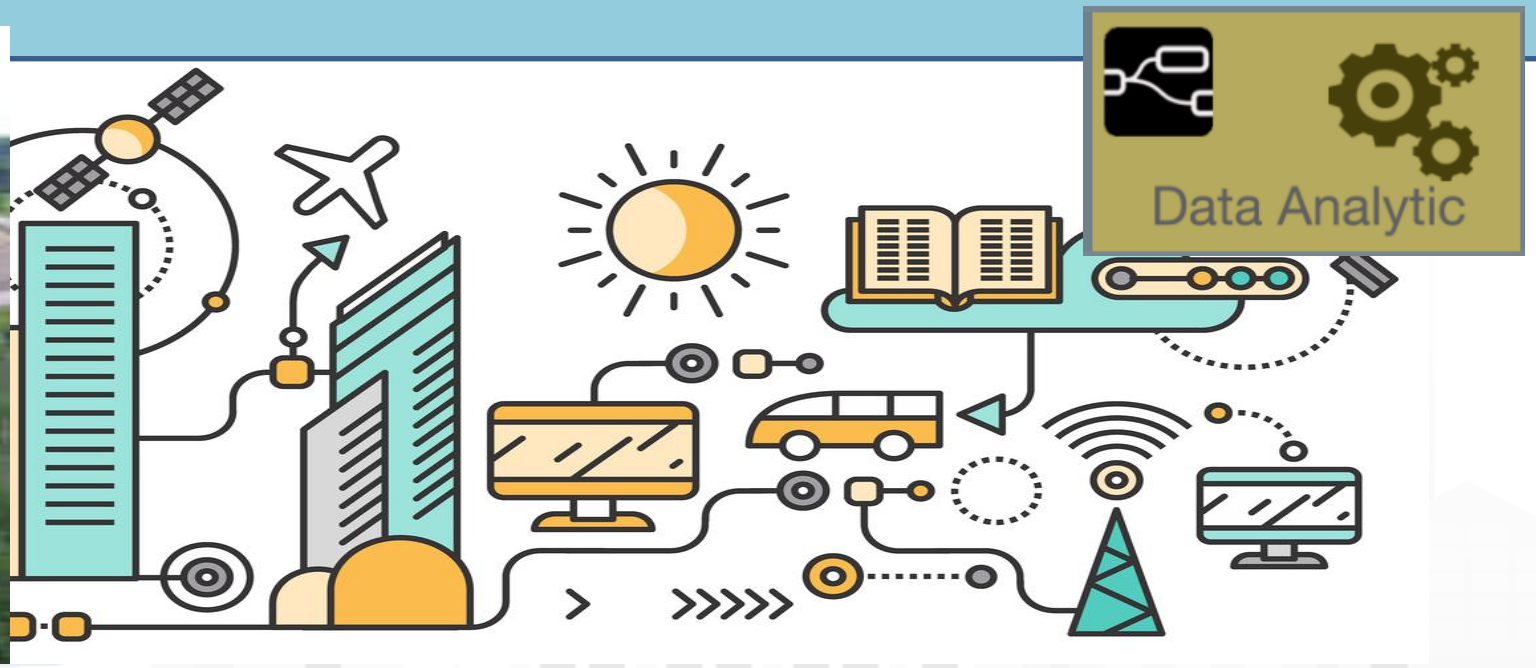


https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MjE5MA==

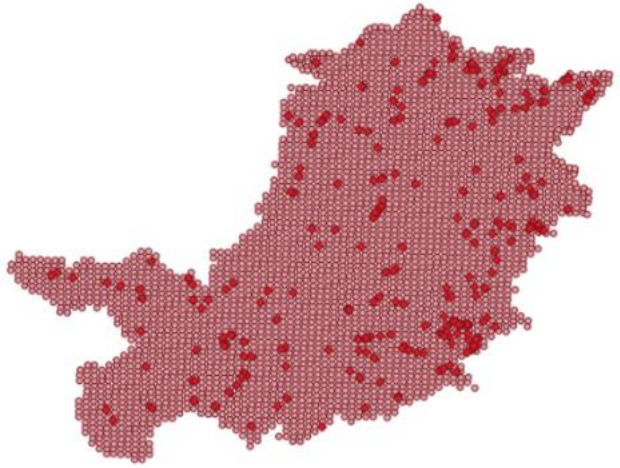
TOP

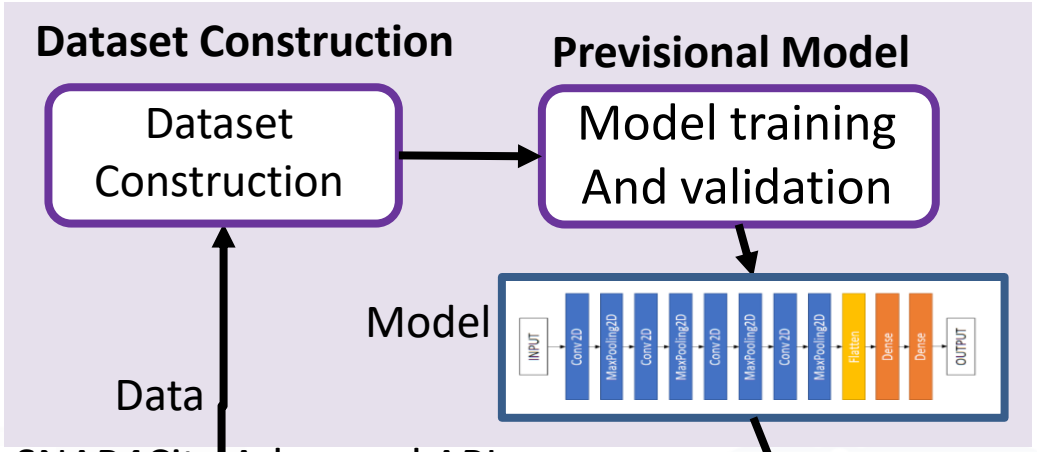
Data Analytic: for Management

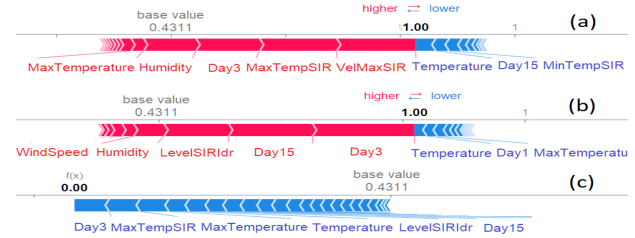
Predicting Land sliding

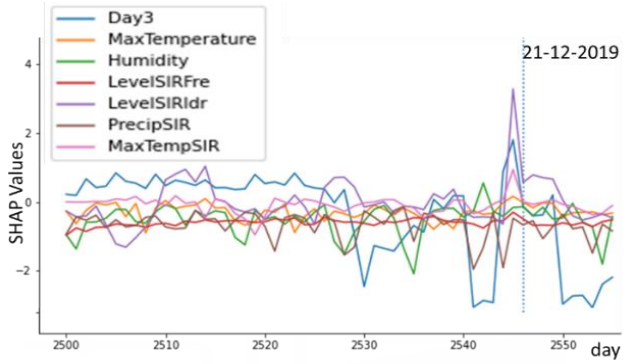


Predicting Land slides









landslide DB

Real Time data from Field: rain, weather, etc.

Ingestion Processes

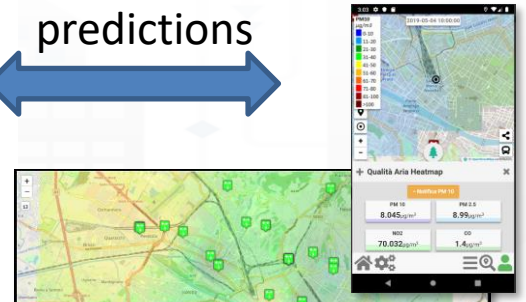
Big Data Storage and KB

Predictions

Model execution Shap Assessment

Data Analytics IOT App Management

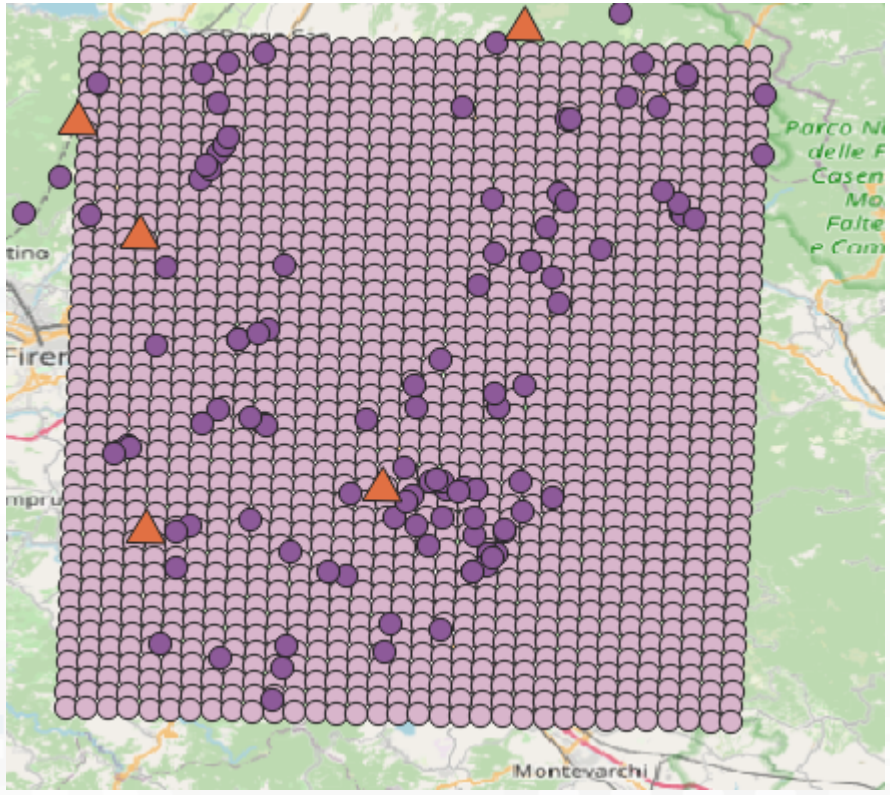
predictions



Snap4City Servers and Tools: Dashboard manager, Heatmap manager, GeoServer, Smart City API.

E. Collini, L. A. I. Palesi, P. Nesi, G. Pantaleo, N. Nocentini and A. Rosi, "Predicting and Understanding Landslide Events with Explainable AI," in *IEEE Access*, doi: 10.1109/ACCESS.2022.3158328.
https://ieeexplore.ieee.org/abstract/document/9732490

Features as Predictors: static + dynamic data

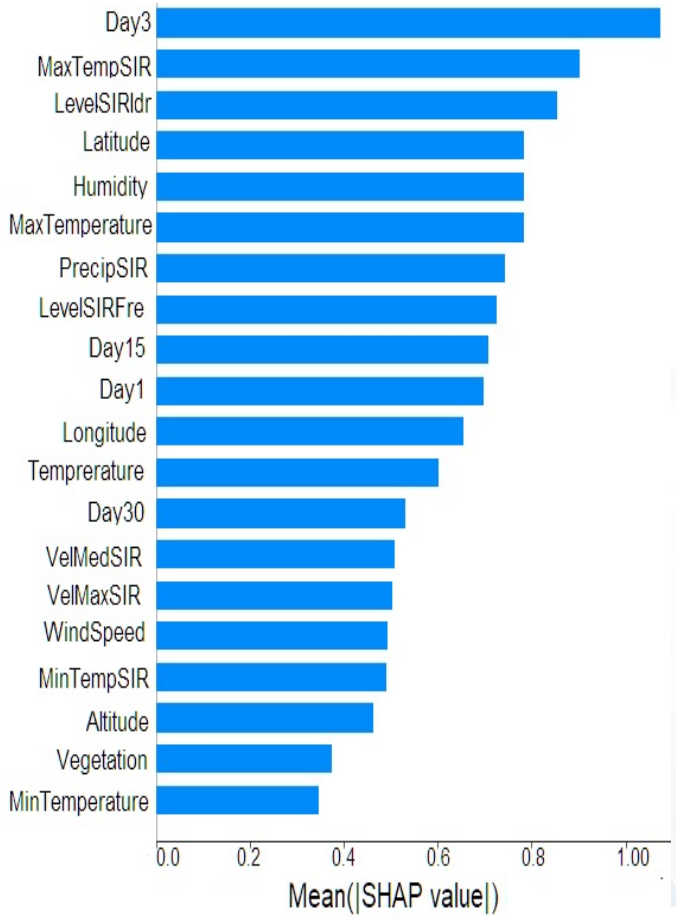


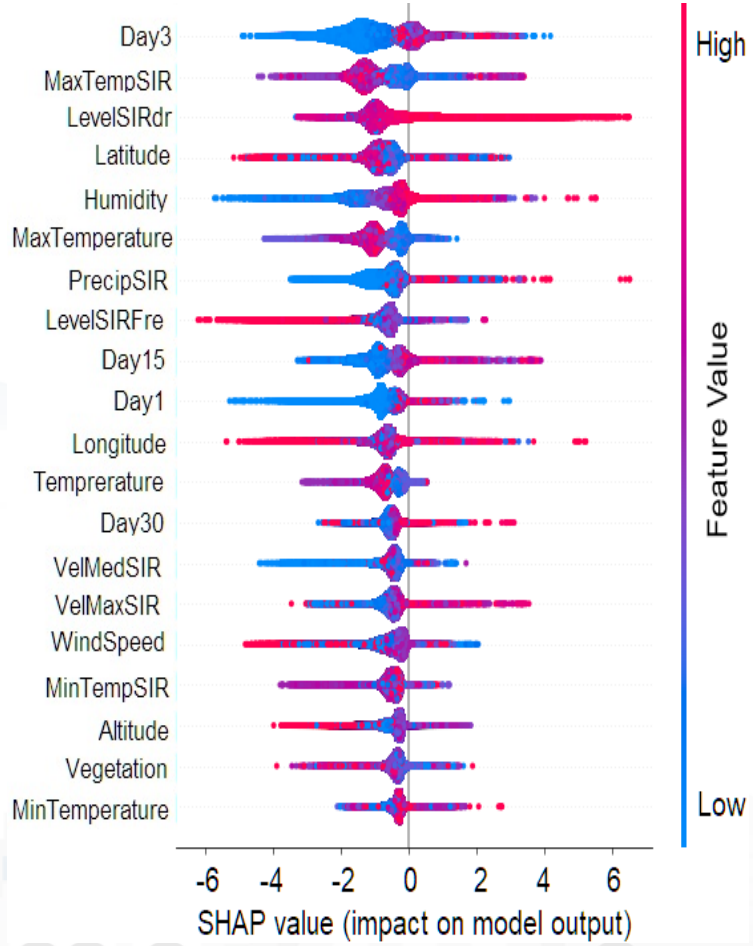
landslide events rain gauges grid

Feature	Description	Unit	Example
Date	Observation date, in the format YYYY-MM-DD	Day	2013-01-14
Latitude	Latitude of the area, EPSG:4326 format	Deg	43.86239
Longitude	Longitude of the area in the EPSG:4326 format	Deg	11.51586
Altitude	Altitude of the area	m	467.204
Slope	Acclivity of the area	%	45.942
Vegetation	Vegetation of the area	%	0.262
Ground	Soil type at the event site (class UCS)		223-Oliveti
Day1	Rainfall on the day before the observation	mm	12.453
Day3	Rainfall on the 3 days preceding the observation	mm	15.072
Day15	Rainfall on the 15 days preceding the observation	mm	16.160
Day30	Rainfall on the 30 days preceding the observation	mm	51.515
Temperature	Mean Temperature on the observation day (IlMeteo.it)	°C	6.965
MinTemperature	Minimum temperature on the observation day (IlMeteo.it)	°C	2.99
MaxTemperature	Maximum temperature on the observation day (IlMeteo.it)	°C	9.942
Humidity	Humidity (average) on the observation day (IlMeteo.it)	%	92.96
WindSpeed	Average wind speed on the observation day (IlMeteo.it)	Km/h	5.991
VelMedSIR	Average wind speed on the observation day (SIR)	m/s	0.9
VelMaxSIR	Maximum wind speed on the day of observation (SIR)	m/s	1.8
LevelSIRFre	phreatimetric data on the observation day (SIR)	m	-4.34
LevelSIRIdr	Water (river) level recorded on the observation day (SIR)	m	0.8
PrecipSIR	Precipitation on the observation day (SIR)	mm	0
MinTempSIR	Minimum temperature on the observation day (SIR)	°C	0.5

Comparing Predictive Model/architectures

Model	XGBoost	RF	CNN	Auto encoder	SIGMA
MAE	0.000173	0.000334	0.000600	0.009218	0.004169
MSE	0.000173	0.000334	0.000259	0.009218	0.004169
RMSE	0.0131	0.0182	0.0160	0.0960	0.064572
Accuracy	0.99	0.99	0.99	0.99	0.99
Sensitivity	0.79	0.36	0.24	0.19	0.06
Specificity	0.99	0.99	0.99	0.99	0.99
TSS	0.78	0.35	0.23	0.18	0.05
PfA	0.01%	0.02%	0.01%	0.11%	0.39%
Precision	0.63	0.35	0.33	0.64	0.003
F1 score	0.70	0.36	0.27	0.29	0.007
MCC	0.70	0.36	0.28	0.35	0.01
OA	2.40	1.72	1.55	1.64	1.02
Kappa	0.70	0.36	0.27	0.29	0.01
AUC	0.89	0.68	0.99	0.92	0.53



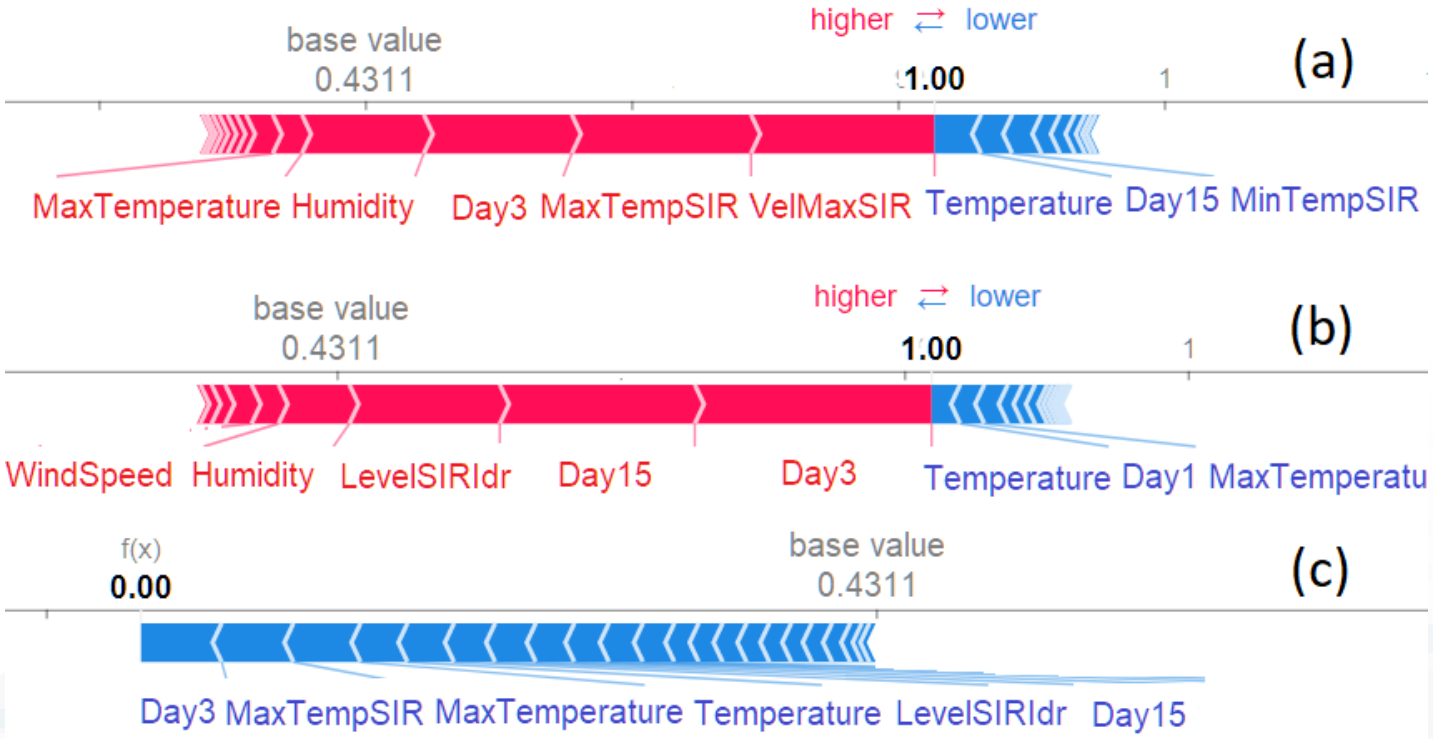


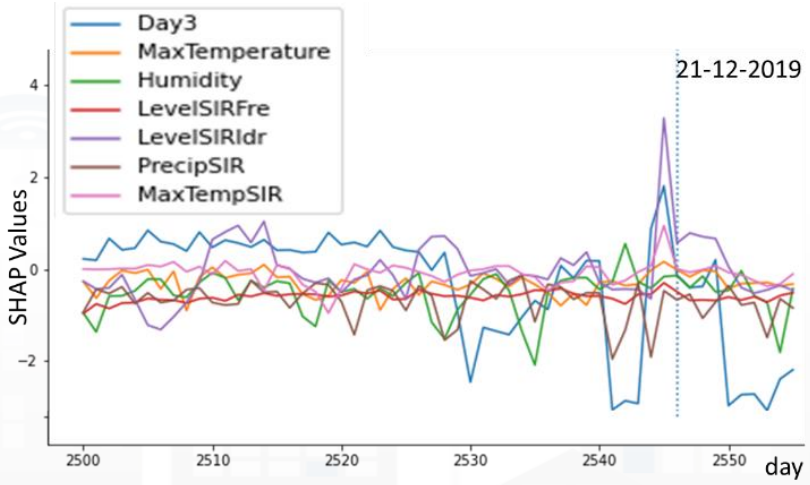
Globl Explainable AI

- Feature relevance

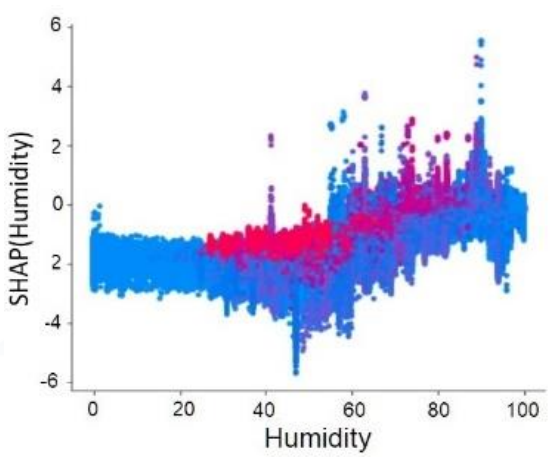
- Red: positive, blue: negeative;
- vs intensity and impact

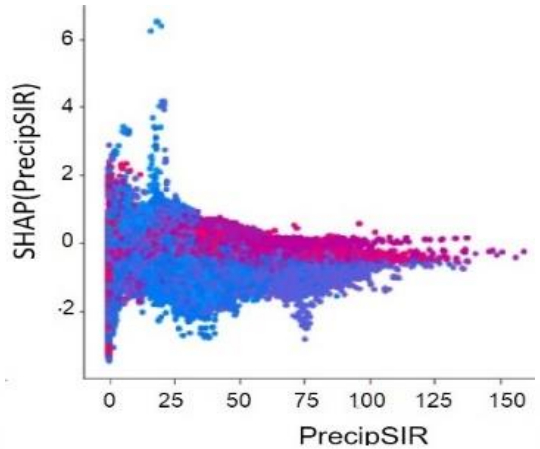
Local Explainable AI: understanding the single events

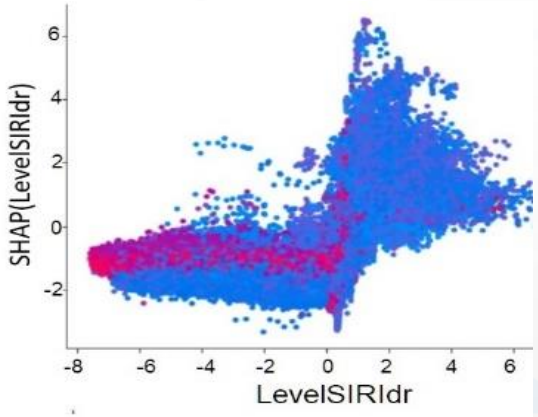


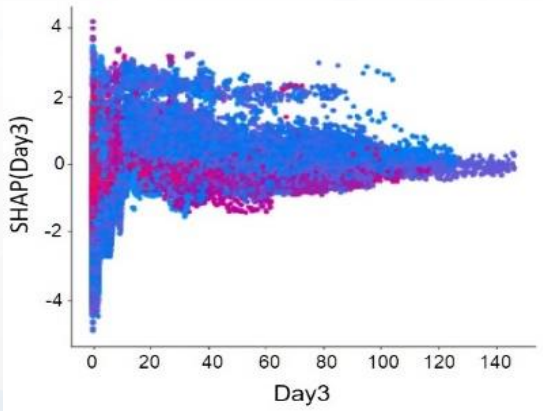


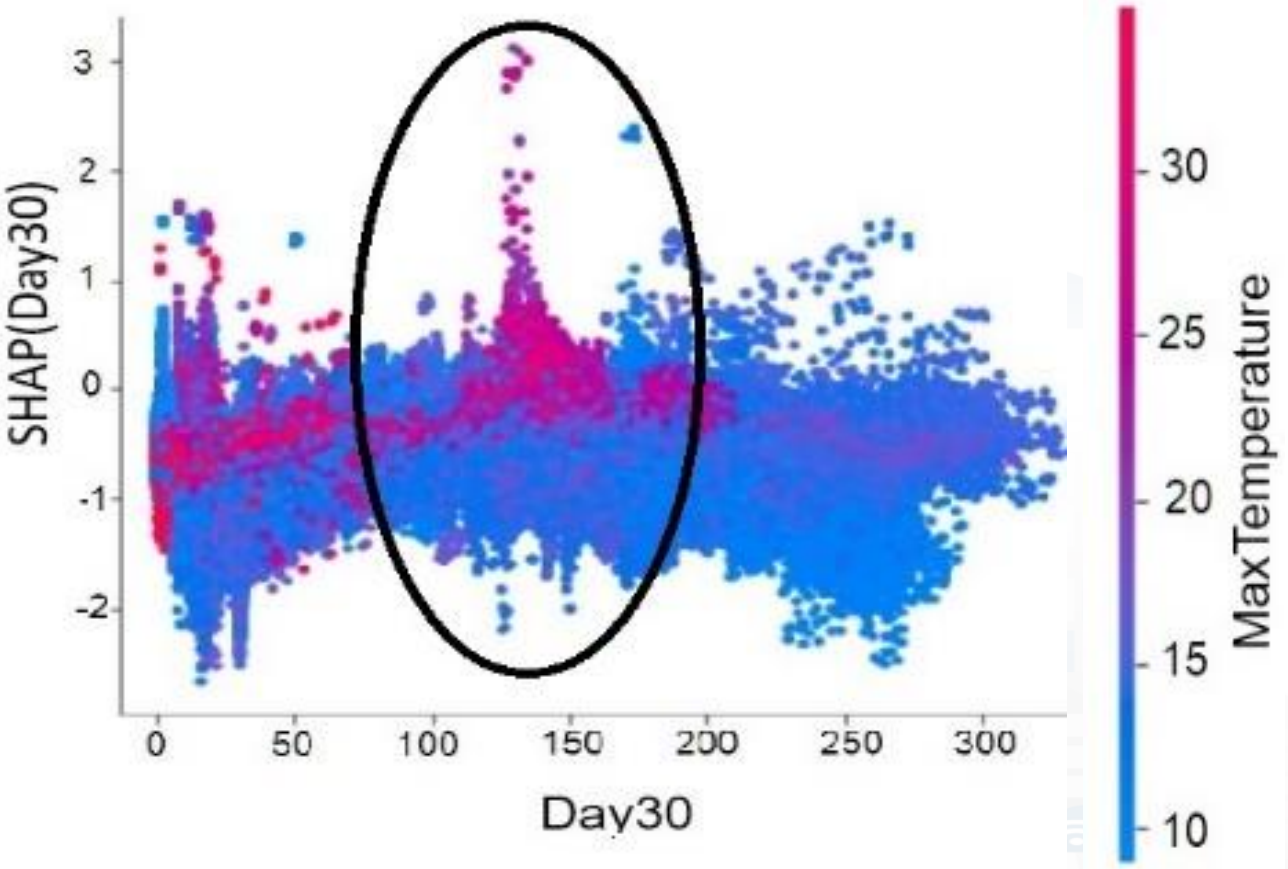
Impact of Features on corresp. SHAP Values vs MaxTemp





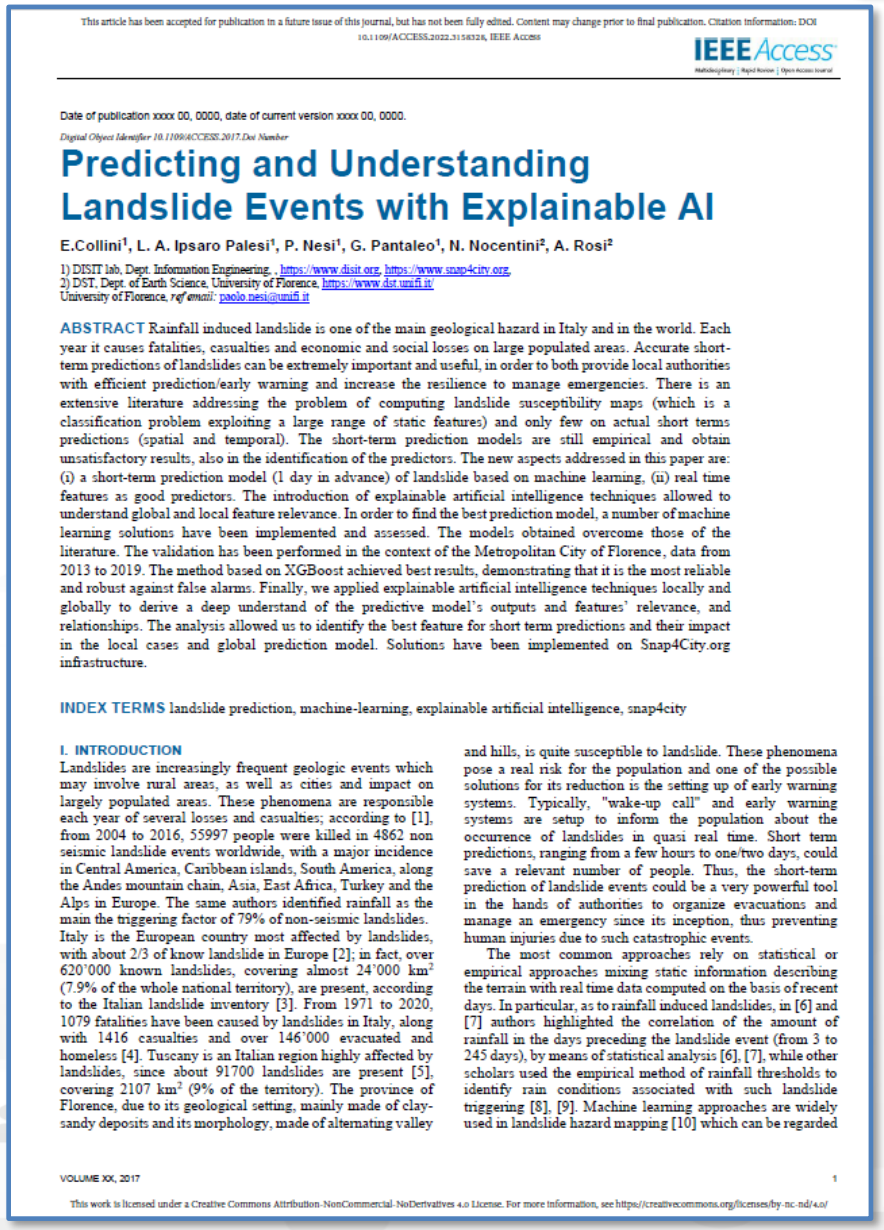






IEEE Access

- E. Collini, L. A. I. Palesi, P. Nesi, G. Pantaleo, N. Nocentini and A. Rosi, "Predicting and Understanding Landslide Events with Explainable AI," in *IEEE Access*, doi: 10.1109/ACCESS.2022.3158328.
- https://ieeexplore.ieee.org/abstract/document/9732490



This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2022.3158328, IEEE Access

IEEE Access

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.Doi Number

Predicting and Understanding Landslide Events with Explainable AI

E.Collini[1], L. A. Ipsaro Palesi[1], P. Nesi[1], G. Pantaleo[1], N. Nocentini[2], A. Rosi[2]

1) DISIT lab, Dept. Information Engineering, , https://www.disit.org, https://www.snap4city.org,
2) DST, Dept. of Earth Science, University of Florence, https://www.dst.unifi.it/
University of Florence, *ref email:* paolo.nesi@unifi.it

ABSTRACT Rainfall induced landslide is one of the main geological hazard in Italy and in the world. Each year it causes fatalities, casualties and economic and social losses on large populated areas. Accurate short-term predictions of landslides can be extremely important and useful, in order to both provide local authorities with efficient prediction/early warning and increase the resilience to manage emergencies. There is an extensive literature addressing the problem of computing landslide susceptibility maps (which is a classification problem exploiting a large range of static features) and only few on actual short terms predictions (spatial and temporal). The short-term prediction models are still empirical and obtain unsatisfactory results, also in the identification of the predictors. The new aspects addressed in this paper are: (i) a short-term prediction model (1 day in advance) of landslide based on machine learning, (ii) real time features as good predictors. The introduction of explainable artificial intelligence techniques allowed to understand global and local feature relevance. In order to find the best prediction model, a number of machine learning solutions have been implemented and assessed. The models obtained overcome those of the literature. The validation has been performed in the context of the Metropolitan City of Florence, data from 2013 to 2019. The method based on XGBoost achieved best results, demonstrating that it is the most reliable and robust against false alarms. Finally, we applied explainable artificial intelligence techniques locally and globally to derive a deep understand of the predictive model's outputs and features' relevance, and relationships. The analysis allowed us to identify the best feature for short term predictions and their impact in the local cases and global prediction model. Solutions have been implemented on Snap4City.org infrastructure.

INDEX TERMS landslide prediction, machine-learning, explainable artificial intelligence, snap4city

I. INTRODUCTION

Landslides are increasingly frequent geologic events which may involve rural areas, as well as cities and impact on largely populated areas. These phenomena are responsible each year of several losses and casualties; according to [1], from 2004 to 2016, 55997 people were killed in 4862 non seismic landslide events worldwide, with a major incidence in Central America, Caribbean islands, South America, along the Andes mountain chain, Asia, East Africa, Turkey and the Alps in Europe. The same authors identified rainfall as the main the triggering factor of 79% of non-seismic landslides. Italy is the European country most affected by landslides, with about 2/3 of know landslide in Europe [2]; in fact, over 620'000 known landslides, covering almost 24'000 km^2 (7.9% of the whole national territory), are present, according to the Italian landslide inventory [3]. From 1971 to 2020, 1079 fatalities have been caused by landslides in Italy, along with 1416 casualties and over 146'000 evacuated and homeless [4]. Tuscany is an Italian region highly affected by landslides, since about 91700 landslides are present [5], covering 2107 km^2 (9% of the territory). The province of Florence, due to its geological setting, mainly made of clay-sandy deposits and its morphology, made of alternating valley and hills, is quite susceptible to landslide. These phenomena pose a real risk for the population and one of the possible solutions for its reduction is the setting up of early warning systems. Typically, "wake-up call" and early warning systems are setup to inform the population about the occurrence of landslides in quasi real time. Short term predictions, ranging from a few hours to one/two days, could save a relevant number of people. Thus, the short-term prediction of landslide events could be a very powerful tool in the hands of authorities to organize evacuations and manage an emergency since its inception, thus preventing human injuries due to such catastrophic events.

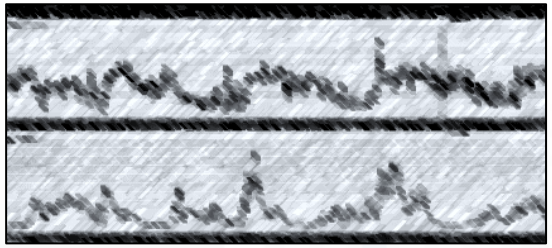
The most common approaches rely on statistical or empirical approaches mixing static information describing the terrain with real time data computed on the basis of recent days. In particular, as to rainfall induced landslides, in [6] and [7] authors highlighted the correlation of the amount of rainfall in the days preceding the landslide event (from 3 to 245 days), by means of statistical analysis [6], [7], while other scholars used the empirical method of rainfall thresholds to identify rain conditions associated with such landslide triggering [8], [9]. Machine learning approaches are widely used in landslide hazard mapping [10] which can be regarded

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

Computing CO2 from traffic Data

Estimating City Local CO2 from Traffic Flow Data

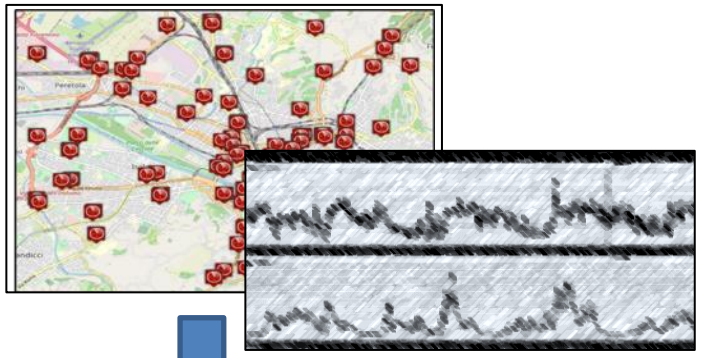
- CO2 sensors are very expensive and thus few



- Traffic Flow is one the main source of CO2
- Most of the cities have many sensors on traffic flow
- **Dense estimation of CO2 into the city** is very useful to know to target the EC limits/KPI

S. Bilotta, P. Nesi, "Estimating CO2 Emissions from IoT Traffic Flow Sensors and Reconstruction", Sensors, MDPI, 2022. https://www.mdpi.com/1424-8220/22/9/3382/

Estimating City Local CO2 from Traffic Flow Data



Computing Traffic Flow into CO2 sensor area

Traffic Flow data

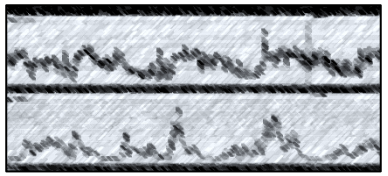
- Traffic Flow is one the main source of CO2
- **Dense estimation of CO2 into the city** is very useful to know to target EC's KPIs

Computing CO2 on the basis of traffic flow data

CO2 estimation

S. Bilotta, P. Nesi, "Estimating CO2 Emissions from IoT Traffic Flow Sensors and Reconstruction", Sensors, MDPI, 2022. https://www.mdpi.com/1424-8220/22/9/3382/

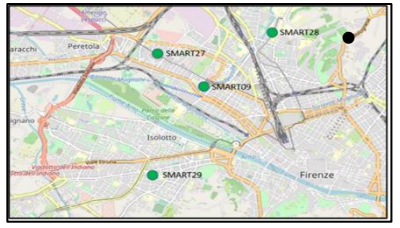
Estimating City Local CO2 from Traffic Flow Data



Traffic Flow data

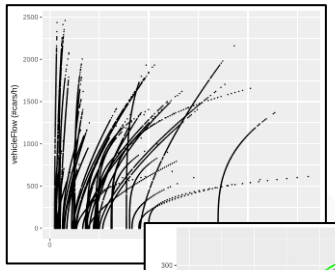
Traffic Flow data Normalization

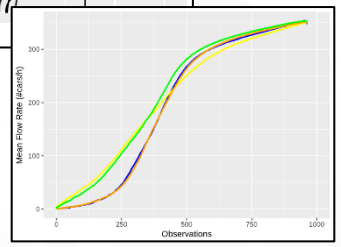
Traffic Flow Reconstruction



CO2 data sensors

CO2 sensors are very expensive and few





Normalized Traffic Flows

Traffic Flow reconstructed

Ref locations

- Traffic Flow is one the main source of CO2
- **Dense estimation of CO2 into the city** is very useful to know

Flow rate analisys, for detecting congested and uncongested situations

Computing Traffic Flow into CO2 sensor area

Detection Criteria

Traffic Flow

CO2 data

Computing Emission Factors, by using CO2 and Traffic Flow relationship

CO2 traffic emissions factors

Traffic Flow data

Computing CO2 on the basis of traffic flow data

CO2 estimation

S. Bilotta, P. Nesi, "Estimating CO2 Emissions from IoT Traffic Flow Sensors and Reconstruction", Sensors, MDPI, 2022. https://www.mdpi.com/1424-8220/22/9/3382/

DORAM: Demand of Mobility vs Offer of Transportation

DORAM

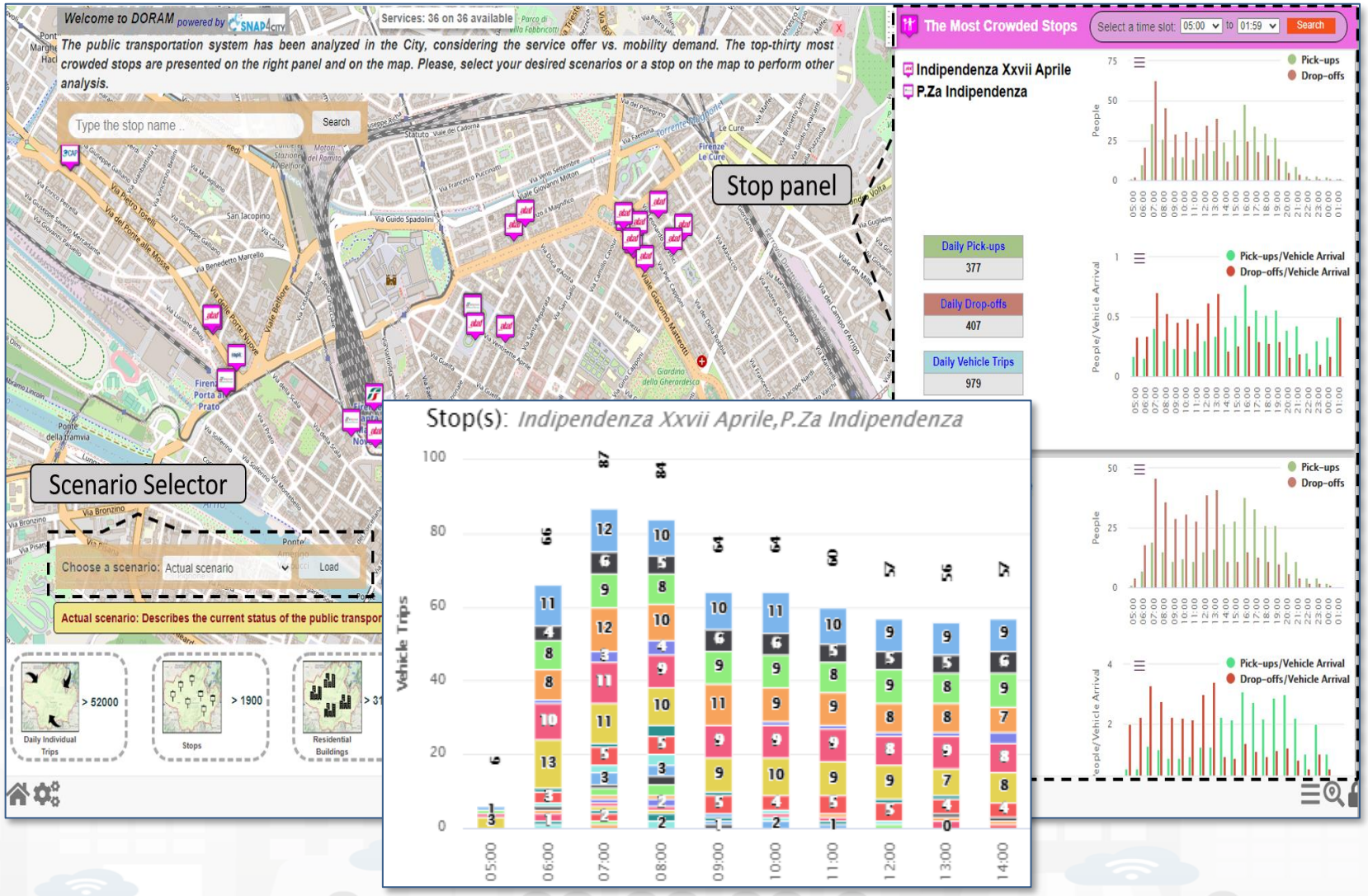
Analysis of

- **Demand** of Mobility
  - Via OD matrices
  - POI, city structure, etc.

With respect to

- **Offert** of Transportation:
  - Public services
  - Private services
  - Multiple agencies
  - GTFS

Critical Busses, busstops, paths, rides, etc.



https://www.snap4city.org/odanalyzer/#b

DORAM

DORAM - Demand OfeR Analyzer for Mobility

Citizen Mobility demand and habits, real time traffic

Service offer (GTFS)

https://www.snap4city.org/odanalyzer/#b

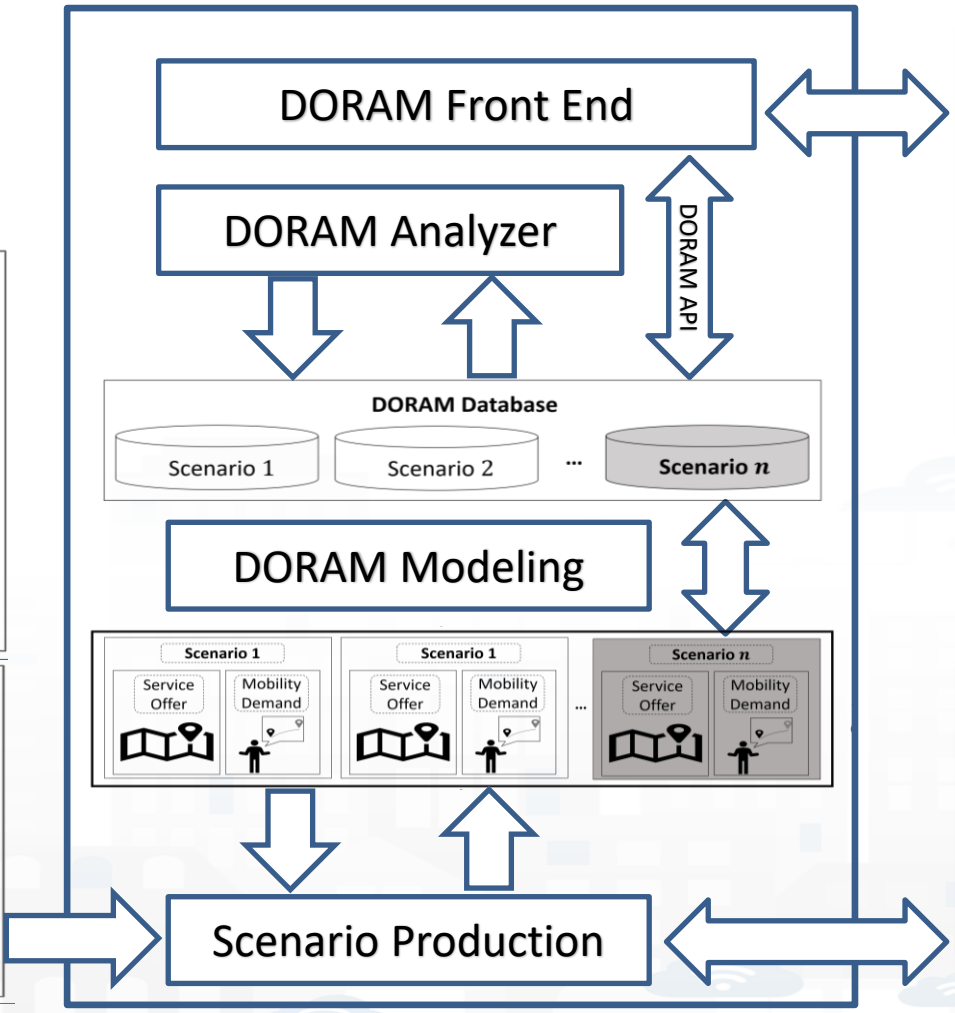
City Mobility Operator(s)

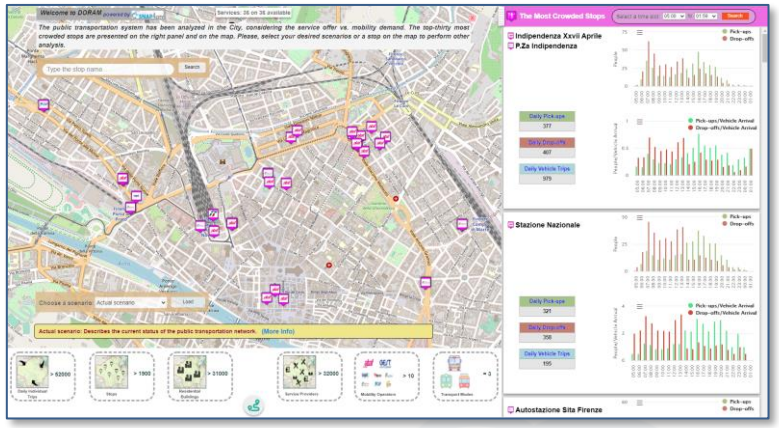
analysis of the offer vs demand (DORAM)

Planned Bus/Tram/Train/ etc. stops/trips and timetables (GTFS)

GTFS variation to improve the efficiency of the service

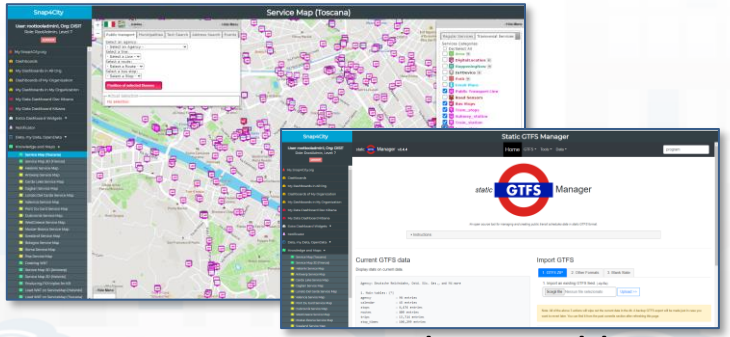
DORAM





DORAM tool

Snap4City tools for City data



GTFS Editor and browser

https://www.snap4city.org/odanalyzer/#b

What can produce the Analysis tool by KPI

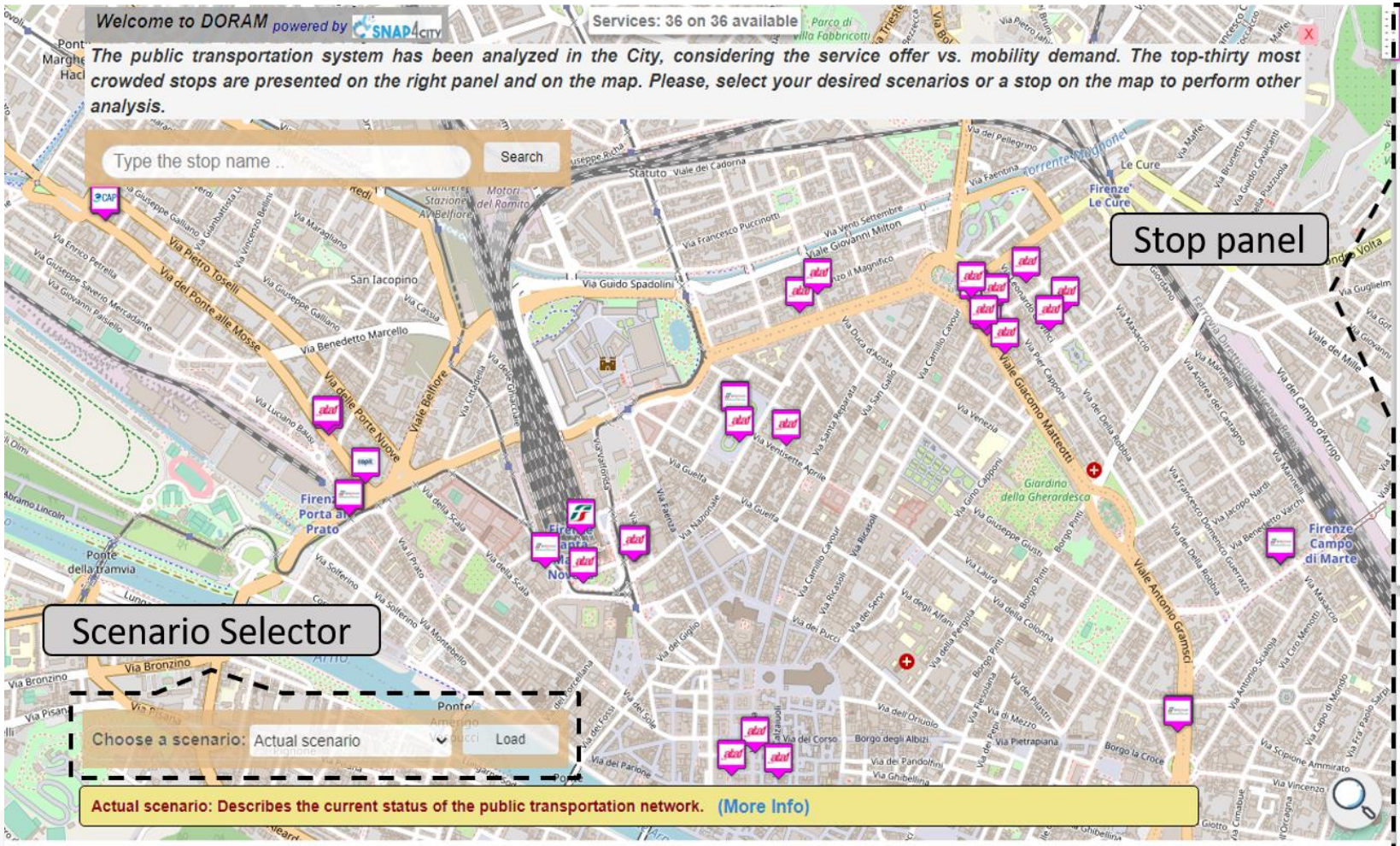
- Identification of critical Bus Stops over time
- Identification of critical courses of bus lines, over day and week
- Effects of changing the position of Bus Stops, courses and line schedules, bus size, etc.
- Effects of changing the contextual conditions:
  - The opening of shopping centers, cinemas, schools, etc..
  - Seize of the buses

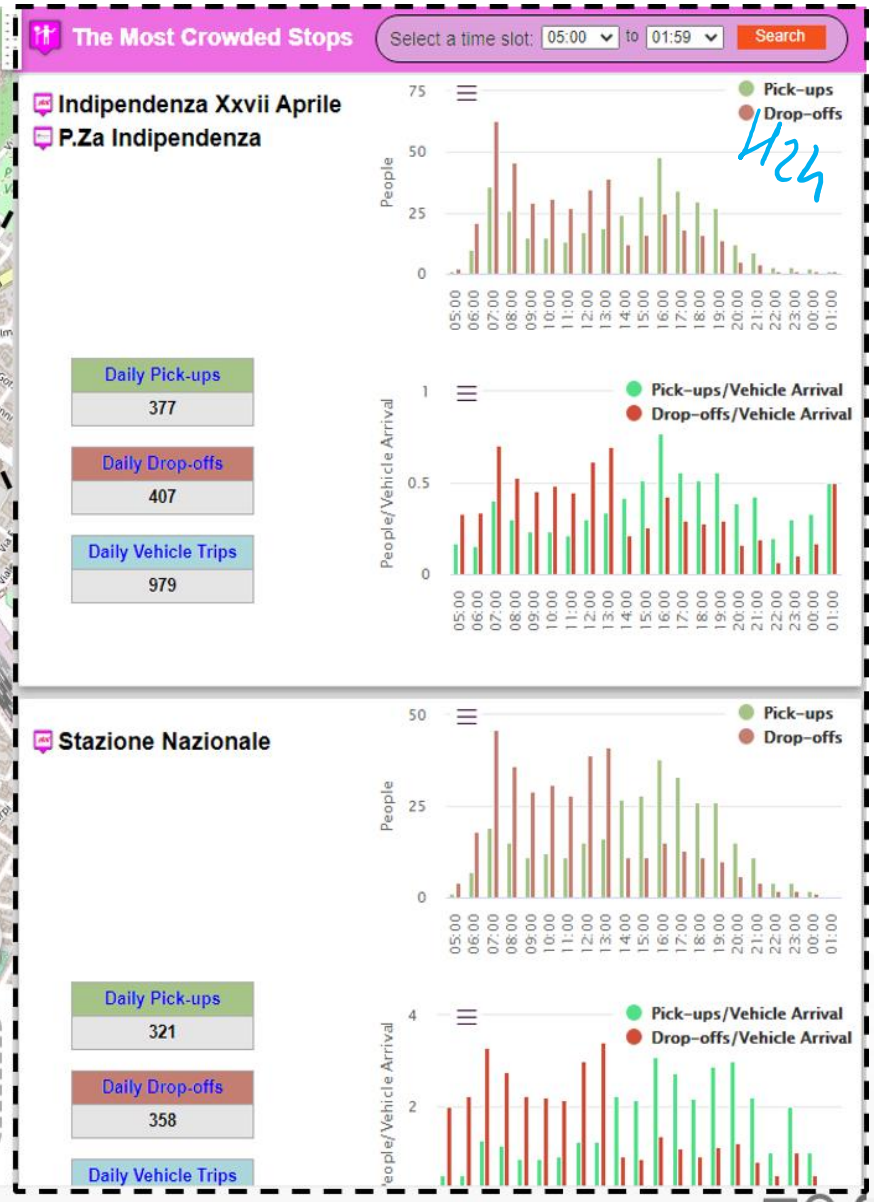
https://www.snap4city.org/odanalyzer/#b

UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB

DORAM

SNAP4CITY | KM4CITY



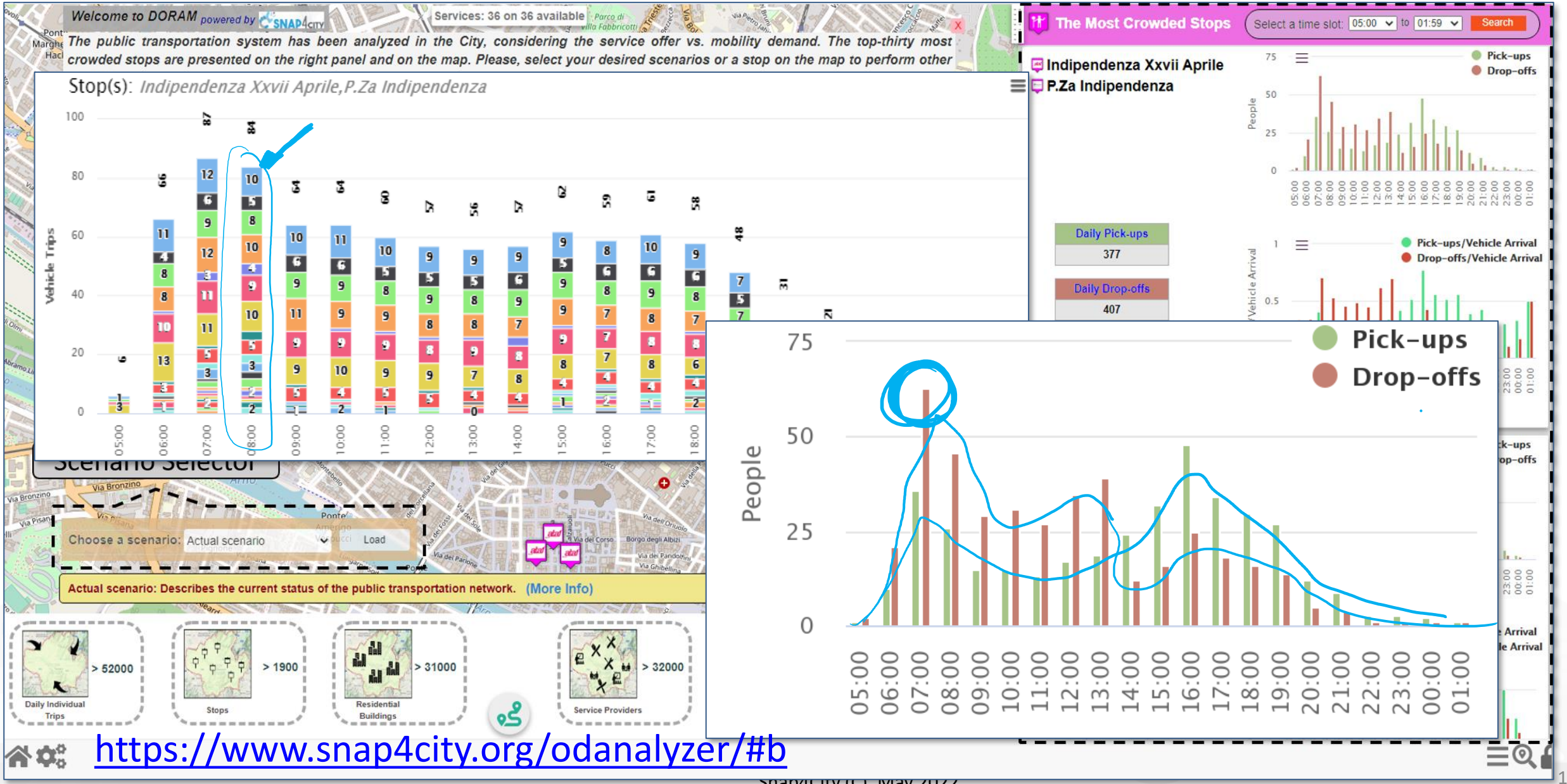


https://www.snap4city.org/odanalyzer/#b

UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB

DORAM

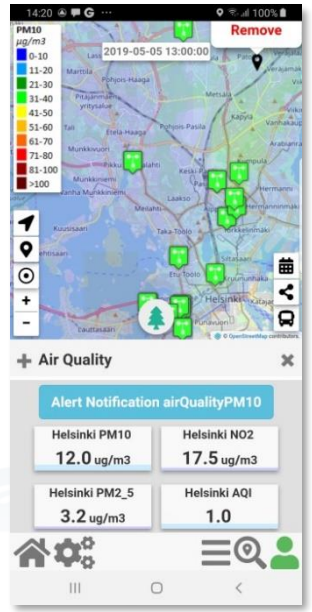
SNAP4CITY | KM 4 CITY



https://www.snap4city.org/odanalyzer/#b

Environmental Data: Predictions, Early Warning

Environmental Heatmaps

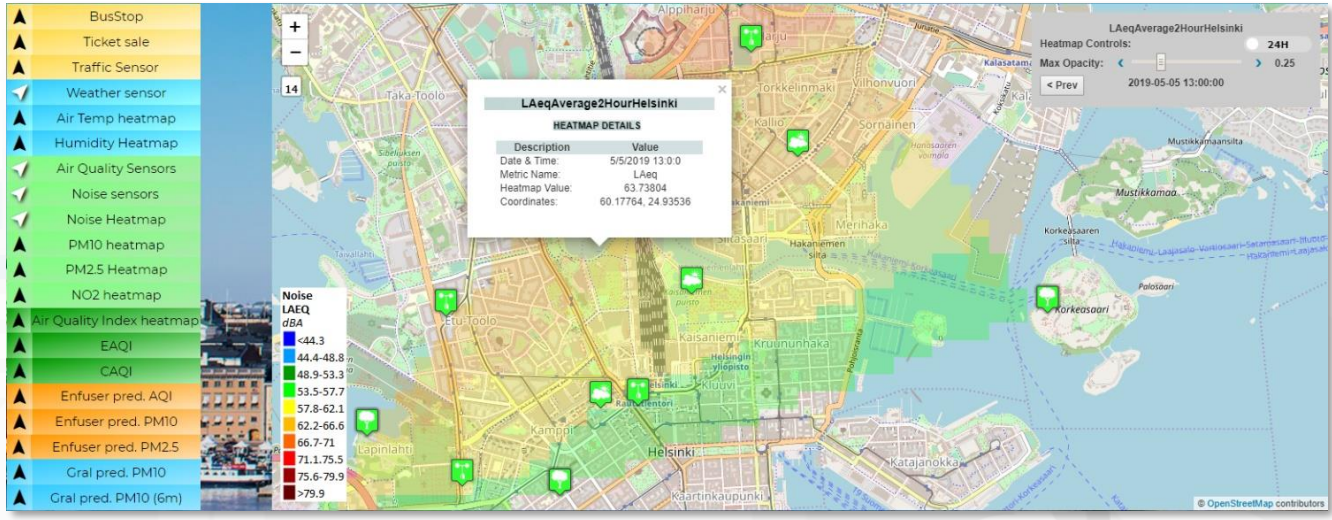


Why ?

- Knowing the Pollutant level an its prediction may increase quality of life
- Making decision about the time slot to walk, run and bring baby at the park
- Be timely informed about the inception of eventual critical conditions
- Deciding where to get an apartment
- …

Calibrated heatmaps based on Interpolated data:

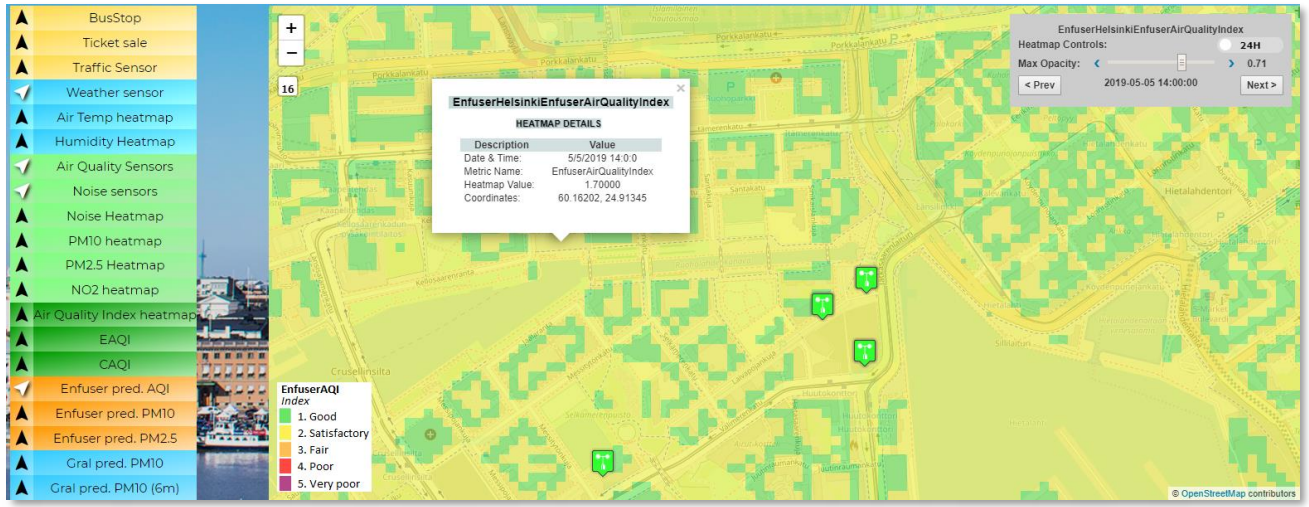
- **Real time** measures (PM_{10}, $PM_{2,5}$, NO_2, SO_2, Noise, NO, O_3, AQI,..)
- **Predictive** measures (ENFUSER, GRAL)
- From **200x200** to **4x4** m
- Hourly concentration
- Any programmed Color map
- Animations over H24
- Picking values in any place
- On Web and Mobile App

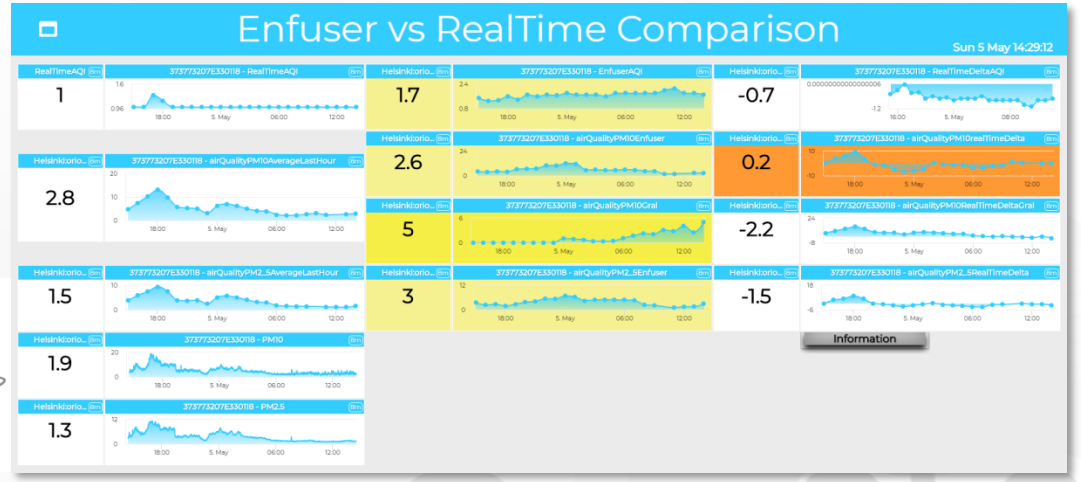


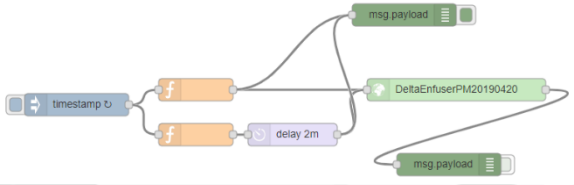
https://www.snap4city.org/snap4cityWebAppHelsinki/

Data Analytics: Enfuser (FMI) predictions

- **Enfuser predictions: AQI, PM10, PM2.5**
  - Data gathering, data processing for Piking
  - Delta Estimation Predictions vs Actual: on 12 points/sensors via R-Studio and IOT App
  - API for accessing data of Heatmaps in real time

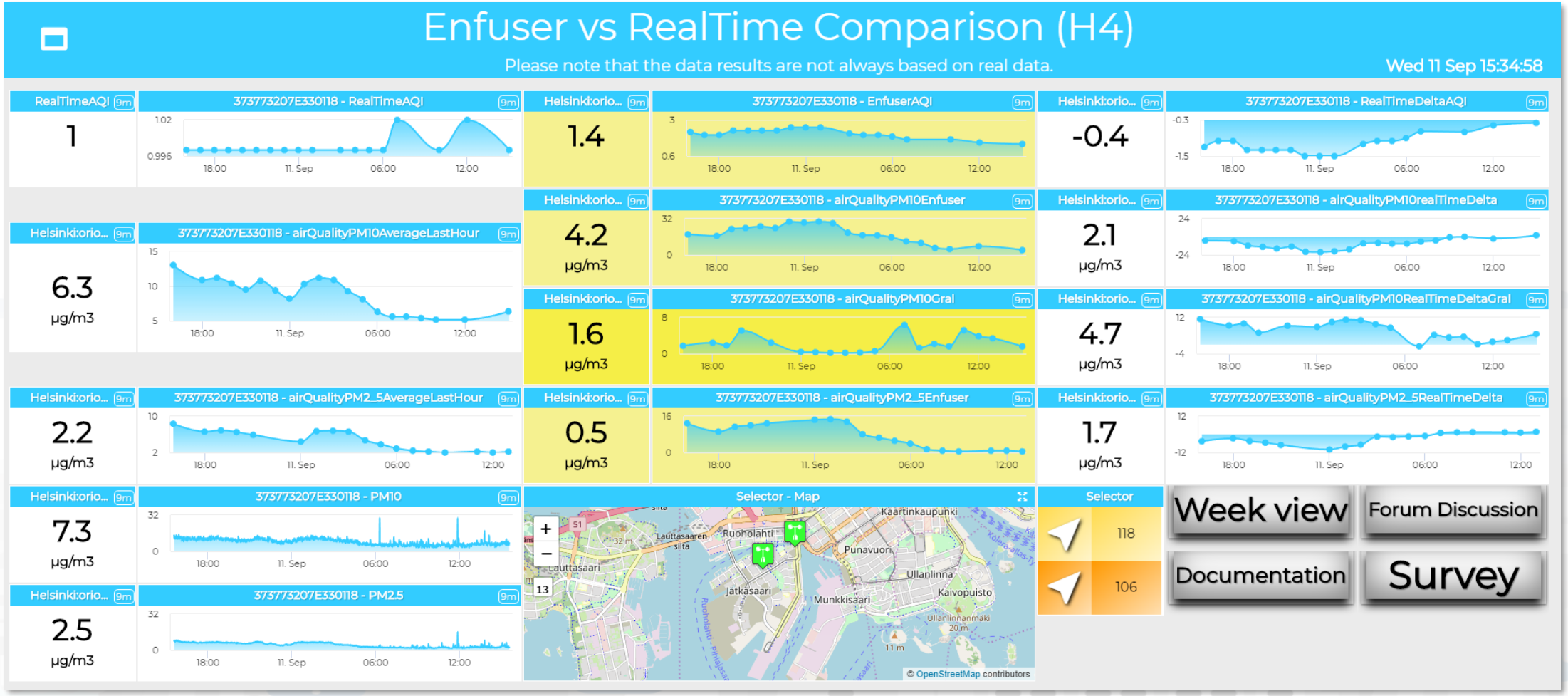






Comparative Dashboard

- ***Delta Estimation Predictions vs Actual*** on 12 points/sensors via R-Studio and IOT App



https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MTczMg==

AQI Indexes estimation via R studio and IOT App

European Air Quality Index **EAQI**

http://airindex.eea.europa.eu/

Pollutant	Index level (based on pollutant concentrations in µg/m3)				
	Good	Fair	Moderate	Poor	Very poor
Particles less than 2.5 µm ($PM_{2.5}$)	0-10	10-20	20-25	25-50	50-800
Particles less than 10 µm (PM_{10})	0-20	20-35	35-50	50-100	100-1200
Nitrogen dioxide (NO_2)	0-40	40-100	100-200	200-400	400-1000
Ozone (O_3)	0-80	80-120	120-180	180-240	240-600
Sulphur dioxide (SO_2)	0-100	100-200	200-350	350-500	500-1250

Measurements of up to five key pollutants supported by modelled data determine the index level that describes *the current air quality situation at each monitoring station*.

The index corresponds to **the poorest level for any of five pollutants according to the following scheme.**

Common Air Quality Index **CAQI**

http://www.airqualitynow.eu

Qualitative name	Index or sub-index	Pollutant (hourly) density in µg/m³			
		NO_2	PM_{10}	O_3	$PM_{2.5}$(optional)
Very low	0–25	0–50	0–25	0–60	0–15
Low	25–50	50–100	25–50	60–120	15–30
Medium	50–75	100–200	50–90	120–180	30–55
High	75–100	200–400	90–180	180–240	55–110
Very high	>100	>400	>180	>240	>110

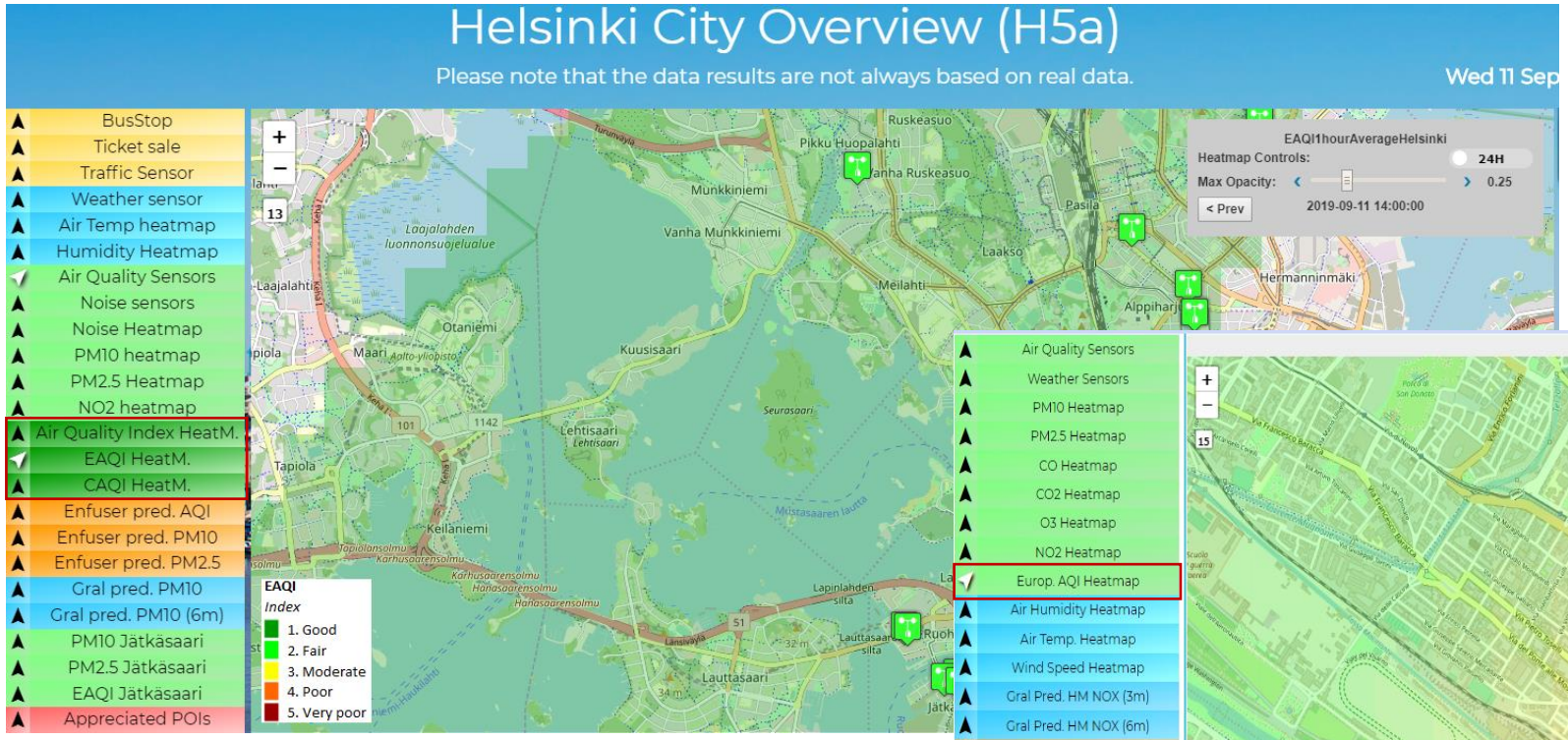
The index is defined away from roads (a "background" index). **CAQI** is computed on the basis of **NO_2, $PM_{2,5}$, PM_{10}** and **O_3**.

Legend of Environmental data:

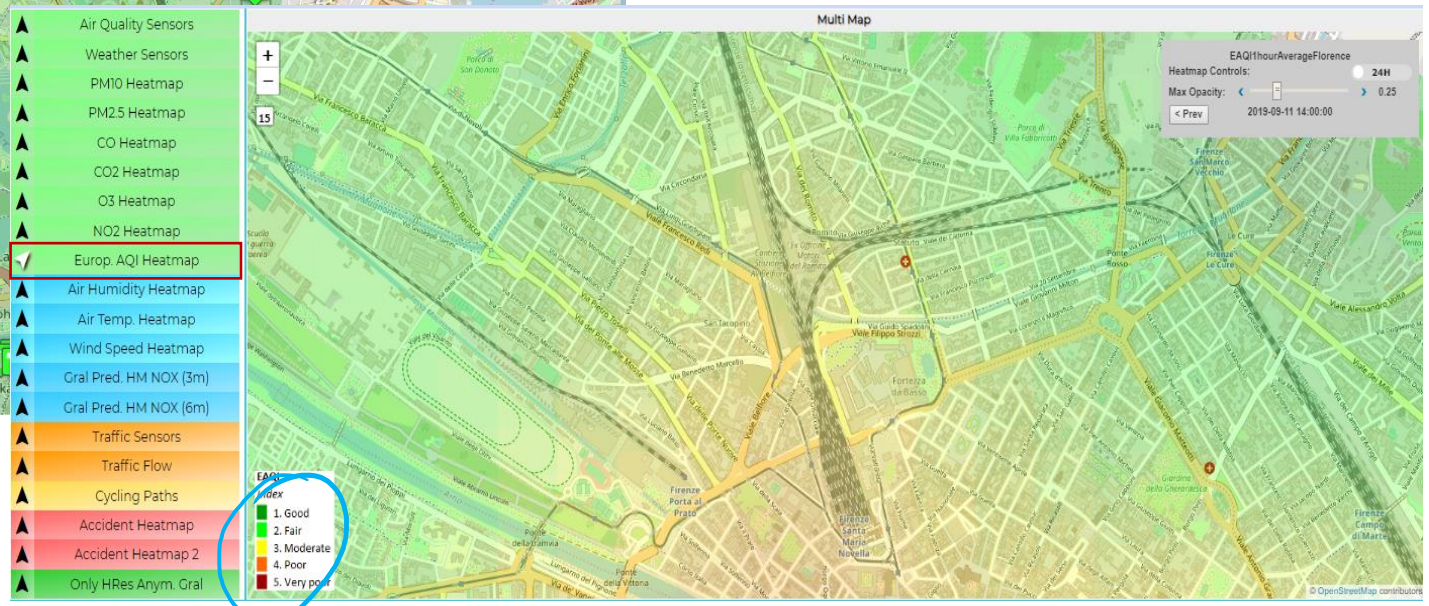
https://www.snap4city.org/435

AQI Indexes estimation Heatmaps

Hourly pollutant concentration



https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MTQwNg==



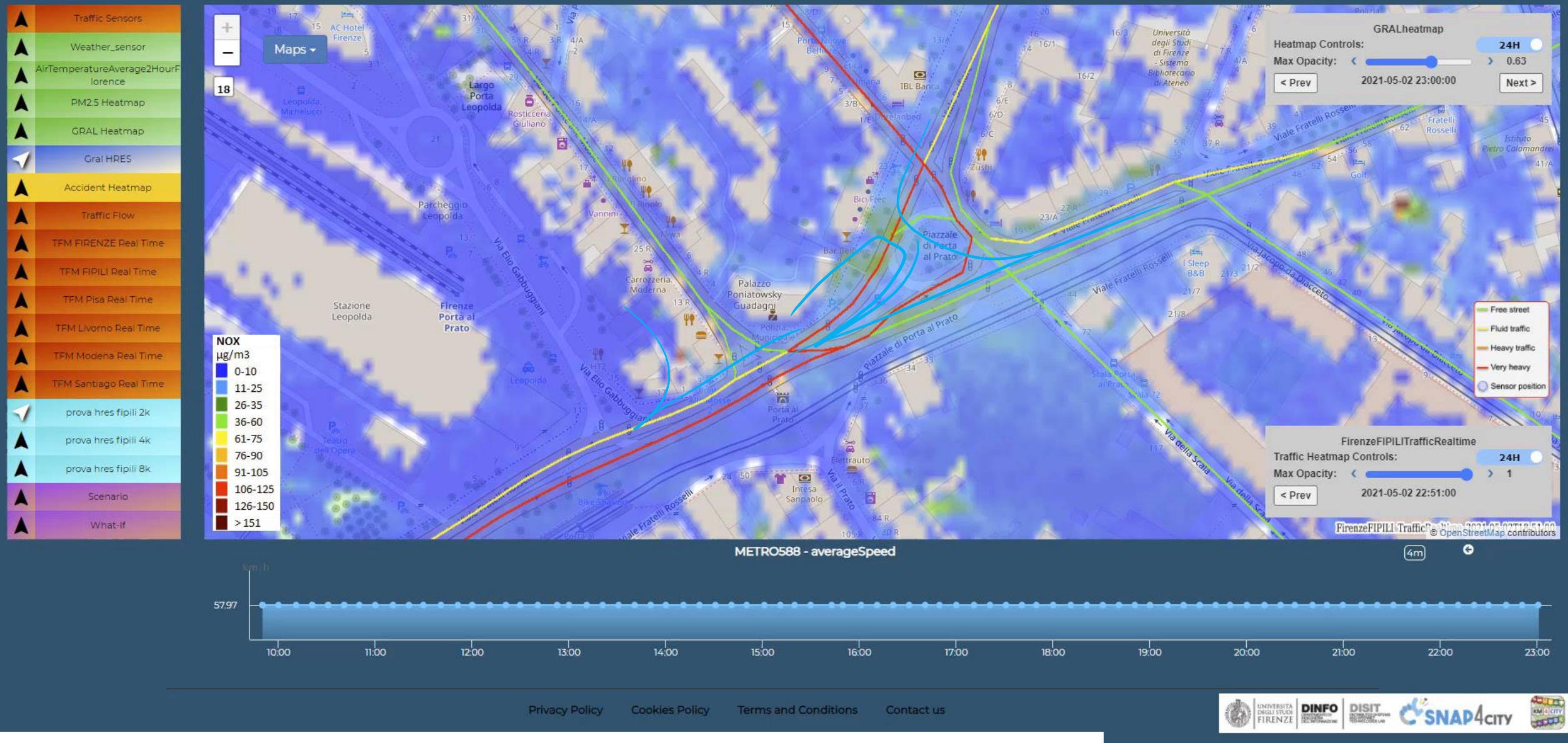
https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MTUzMg==

Environmental Data Predictions: GRAL

- **GRAL predictions: PM10, NOX, ….**
  - **Comparison wrt real time values in actual value of Sensors**
  - Graz Lagrangian Model.
- GRAL model takes into account:
  - pollution sources (for example the vehicles, their distribution on the streets, the about of pollution they produce according to their distribution over time and space, etc.),
  - structure of the city (streets and shape 3D of the buildings),
  - weather forecast (wind intensity and direction), etc.
- GRAL can be applied on NOX, PM10, PM2.5, … or any other particles

Traffic Flow Manager on multiple cities

Sun 2 May 23:16:31

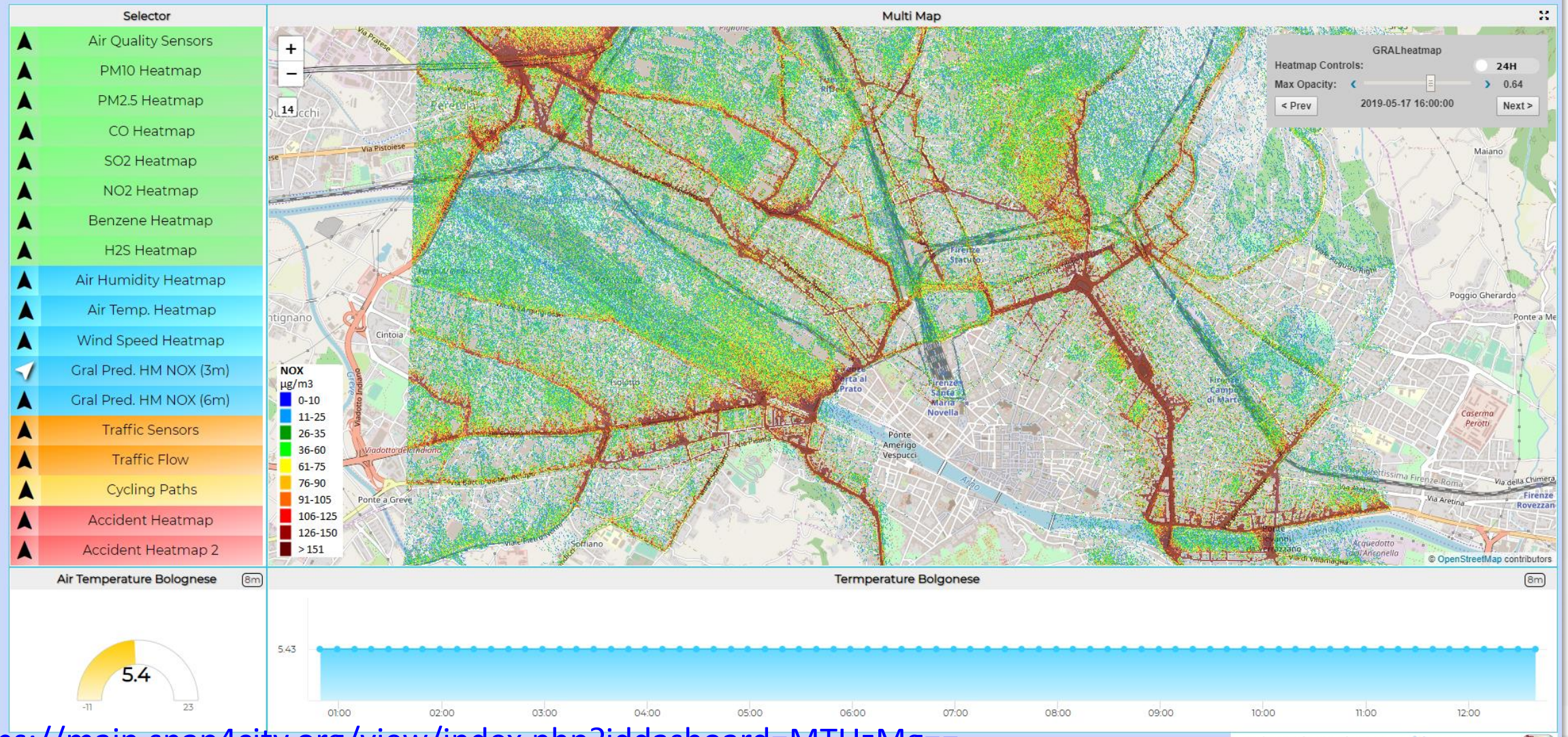


https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MzEyNg==

Heatmap Firenze - trafair

different data

Fri 17 May 12:49:34

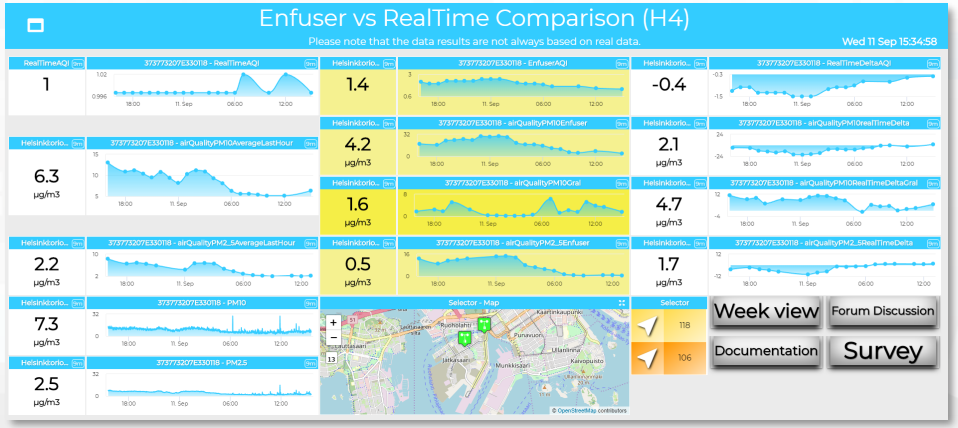


https://main.snap4city.org/view/index.php?iddasboard=MTUzMg==

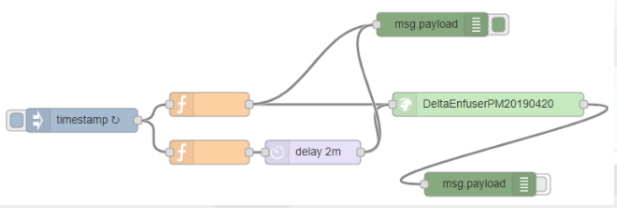
Environmental Data Network and Automated Analysis and Representation

Environmental Predictive Measures:

- **Enfuser pred. AQI:** heatmap of Air Quality Index hourly Enfuser predictions, every 12 mt. the Heatmap Controls widget you can see the forecast.
- **Enfuser pred. PM_{10} :** heatmap of PM_{10} particles hourly Enfuser predictions every 12mt in $\mu g/m_3$.
- **Enfuser pred. $PM_{2,5}$** heatmap of $PM_{2,5}$ particles hourly Enfuser predictions every 12mt in $\mu g/m_3$.
- **Gral pred. PM_{10} (h 3m):** heatmap of PM_{10} particles hourly predictions in µg/m3 measured 3 meters on the ground and computed using Gral model every 4mt.
- **Gral pred. PM_{10} (h 6m)**: heatmap of PM_{10} particles hourly predictions in $\mu g/m_3$ measured 6 meters on the ground and computed using Gral model every 4mt.



- Data gathering, data processing for Piking
- **API** for accessing data of Heatmaps in real time
- ***Delta Estimation Predictions vs Actual***: on 12 points/sensors via R-Studio and IOT App



How it works: NOX predictions

Road Graph

3D City Structure

Traffic Flow Predictions

Weather Forecast: Temp, wind…

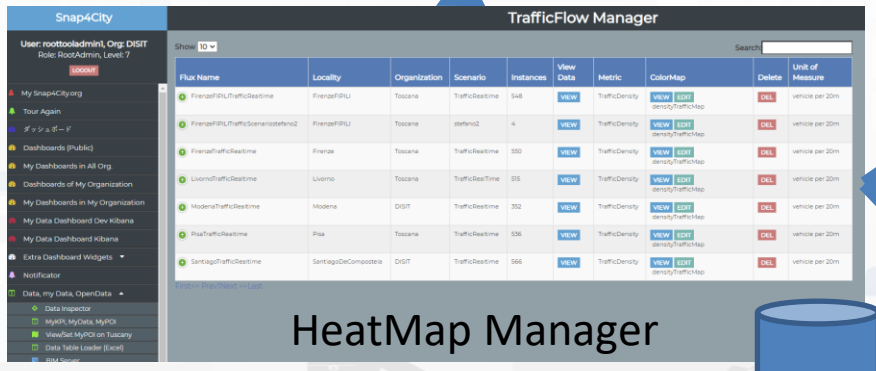
Particles Flow Simulation

Heatmap

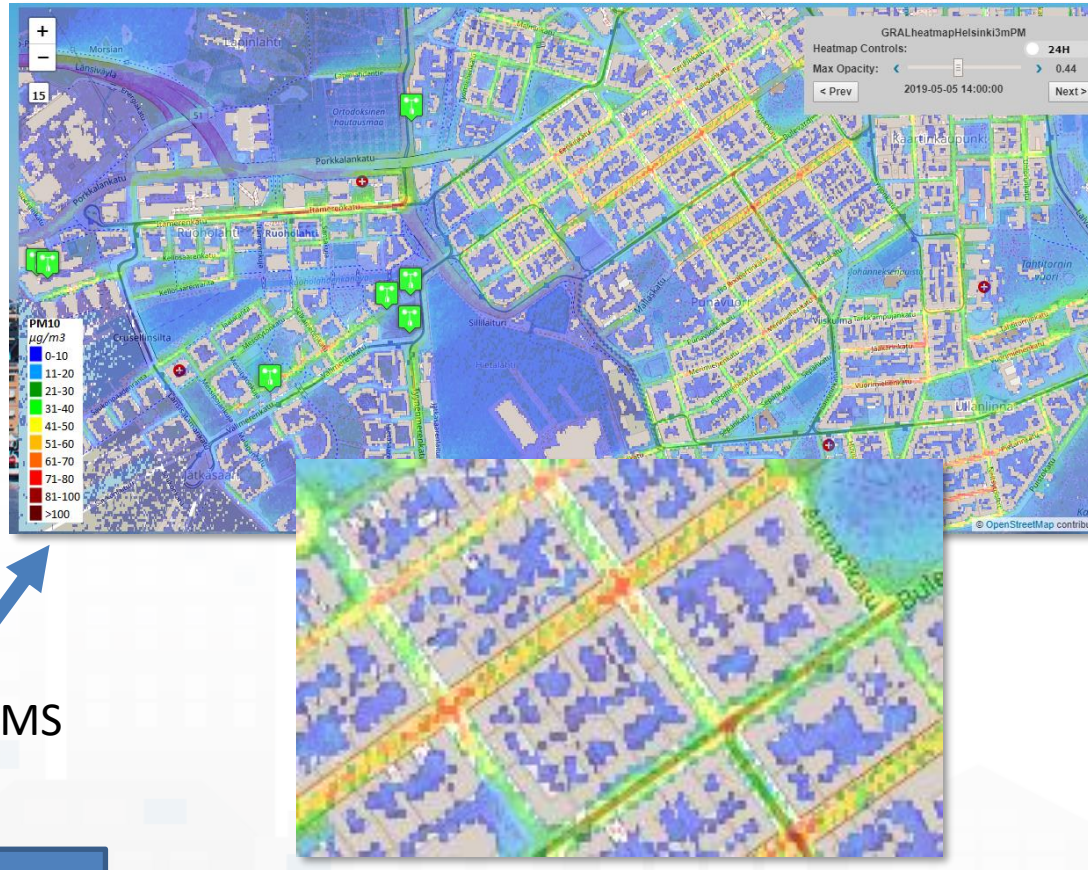
JSON

JSON

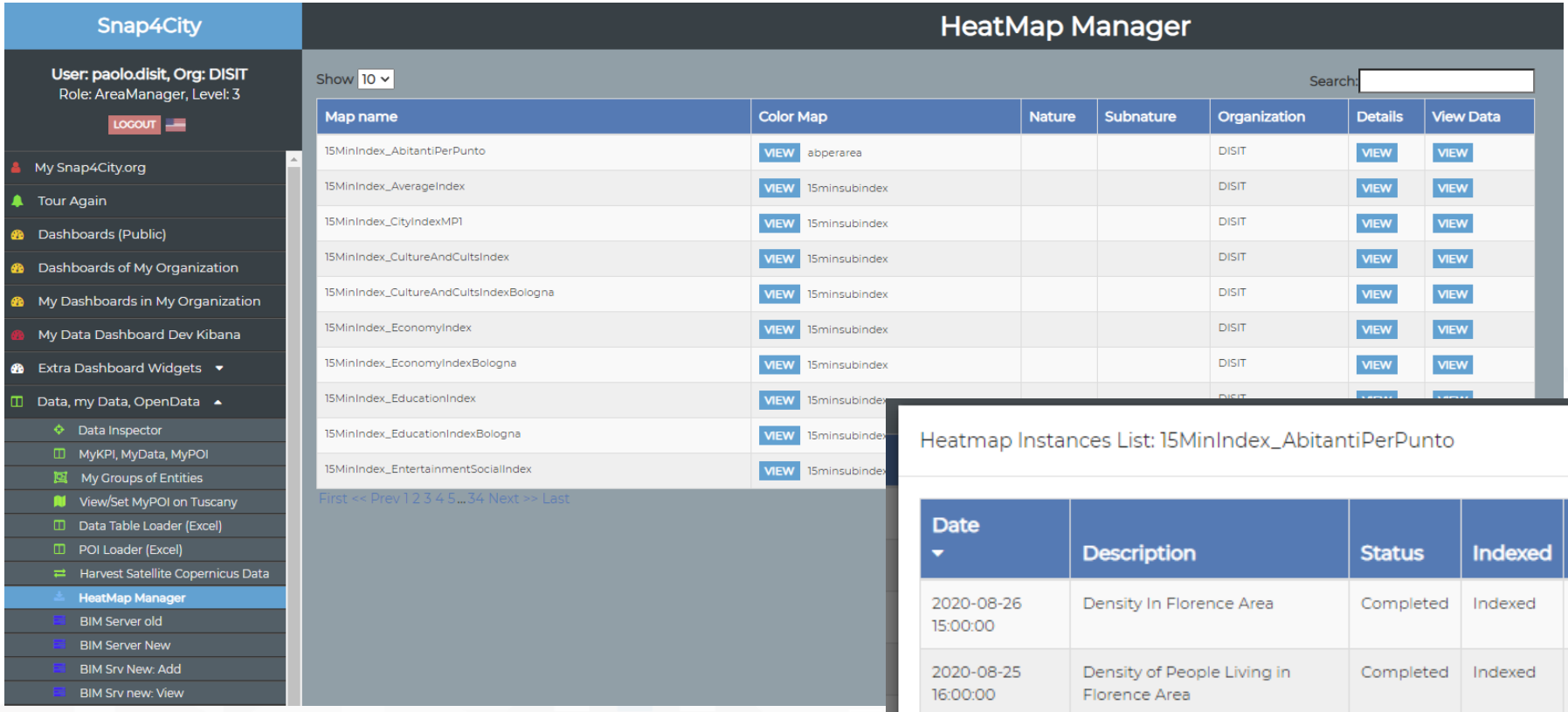
WMS



GeoServer



HeatMap Manager (Area Manager view)

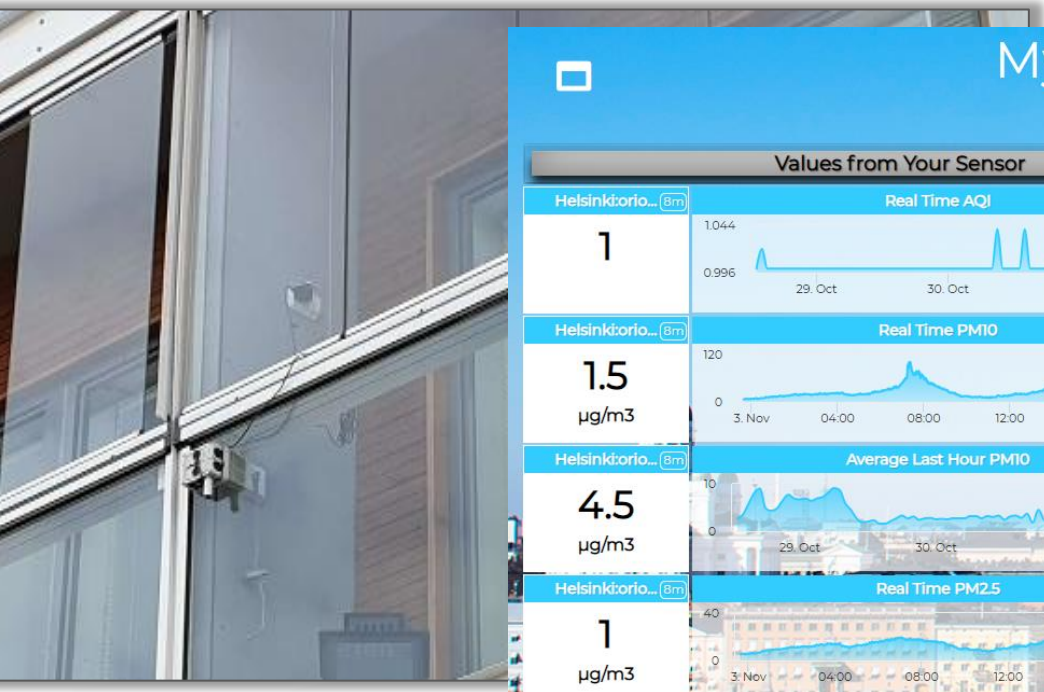


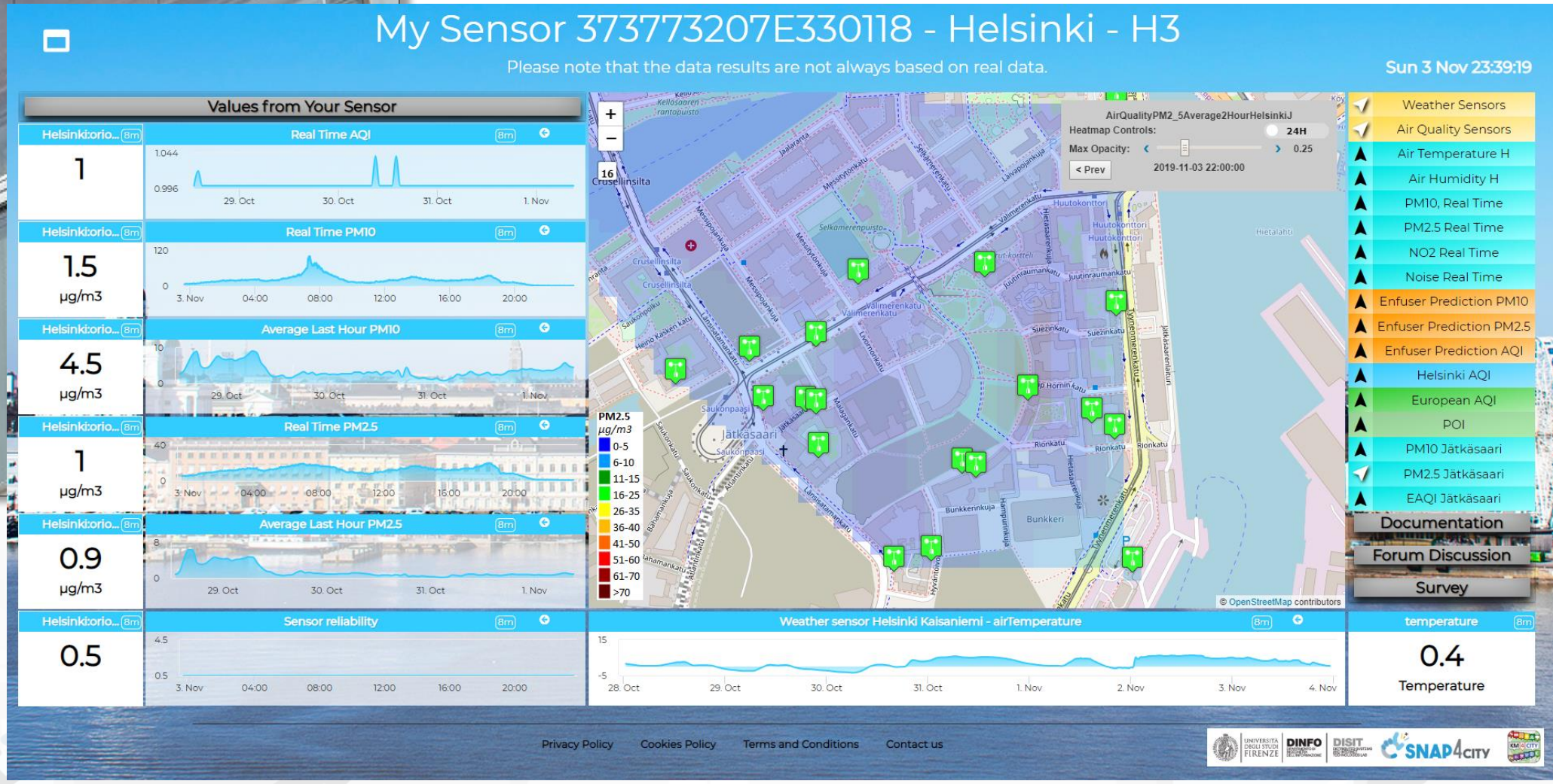
- Sequence of Heatmaps
- Colormap used
- Details

Editing Mode for RootAdmin only

Environmental Sensors Hosted by City Users

Environmental Devices hosted by Citizens





Environmental Data Network and Automated Analysis and Representation

Goal:

- **Real time aggregation, integration, assessment of data** independently on the number of sensors, on their position.
- **Real time analysis and representation** of environmental data automatically in dedicated Dashboards on Snap4City platform.

The **target** has been to:

- Provide ***informative view of the city users*** regarding Environmental data via some mobile App.
- Provide detailed information about the Environmental data to *city officials for decision making*, as ***comparison between predictions and real time*** in specific point of the city.

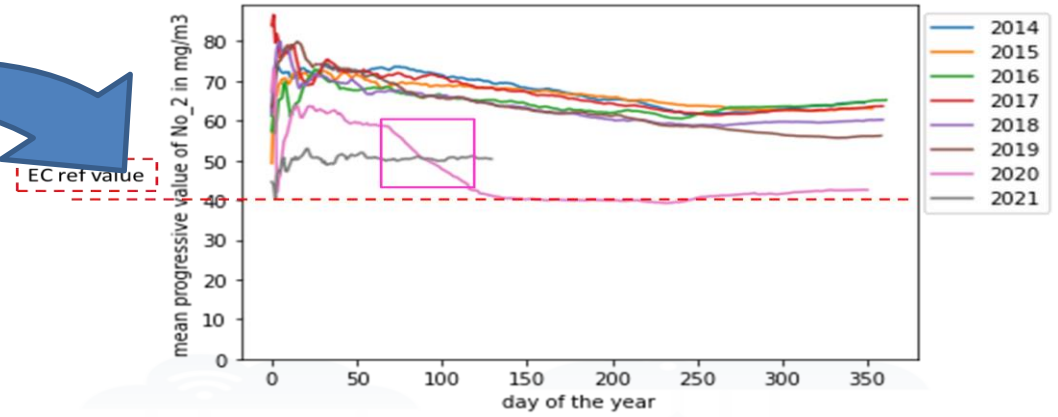
Data have been collected from:

- IOT Brokers included ***IOT Devices hosted by city users***.
- Data Providers.

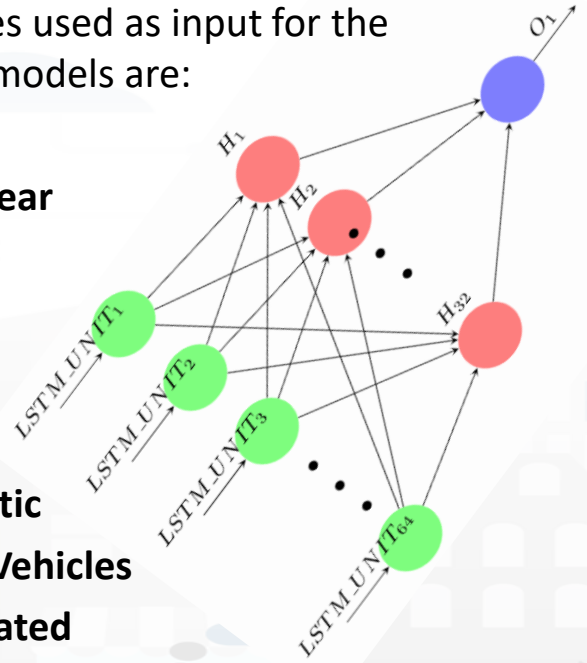
Long Term Prediction of Annual Mean of NO2 index of EC

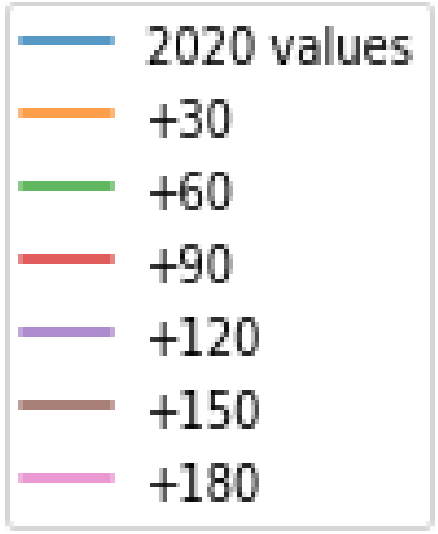
Predicting EC's KPI on NO2 months in advance

Deep Learning Long Terms Predictions of NO2 mean values, From 30 to 180 days in advance



- The features used as input for the predictive models are:
  - **Month**
  - **dayOfTheYear**
  - **NO2**
  - **Tmean**
  - **Humidity**
  - **windMean**
  - **NoxDomestic**
  - **numberOfVehicles**
  - **NO2cumulated**
  - **NO2progresseveMean**
  - **numberOfVehiclesCumulated**



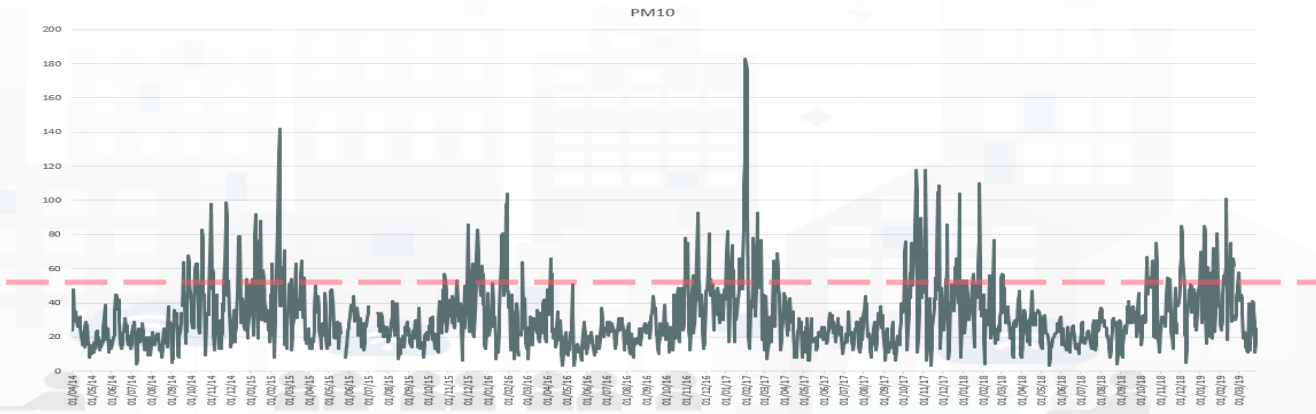


		Air Quality Directive		WHO guidelines	
Pollutant	**Averaging period**	**Objective and legal nature and concentration**	**Comments**	**Concentration**	**Comments**
$PM_{2.5}$	One day			25 µg/m³ (*)	99th percentile (3 days/year)
$PM_{2.5}$	Calendar year	Target value, 25 µg/m³	The target value has become a limit value since 1 January 2015	10 µg/m³	
PM_{10}	One day	Limit value, 50 µg/m³	Not to be exceeded on more than 35 days per year.	50 µg/m³ (*)	99th percentile (3 days/year)
PM_{10}	Calendar year	Limit value, 40 µg/m³ (*)		20 µg/m³	
O_3	Maximum daily 8-hour mean	Target value, 120 µg/m³	Not to be exceeded on more than 25 days per year, averaged over three years	100 µg/m³	
NO_2	One hour	Limit value, 200 µg/m³ (*)	Not to be exceeded more than 18 times a calendar year	200 µg/m³ (*)	
NO_2	Calendar year	Limit value, 40 µg/m³		40 µg/m³	

Predicting Air Quality

- European Air Quality Directive
- Predicting critical days
  - PM10 with an accuracy of more than 90% and precision of 85%;
  - PM2.5 with an accuracy of 90% and precision greater than the 95%.
- Simulating Long terms values
  - For long terms predictions

		Air Quality Directive		WHO guidelines	
Pollutant	**Averaging period**	**Objective and legal nature and concentration**	**Comments**	**Concentration**	**Comments**
$PM_{2.5}$	One day			25 µg/m³ (*)	99th percentile (3 days/year)
$PM_{2.5}$	Calendar year	Target value, 25 µg/m³	The target value has become a limit value since 1 January 2015	10 µg/m³	
PM_{10}	One day	Limit value, 50 µg/m³	Not to be exceeded on more than 35 days per year.	50 µg/m³ (*)	99th percentile (3 days/year)
PM_{10}	Calendar year	Limit value, 40 µg/m³ (*)		20 µg/m³	
O_3	Maximum daily 8-hour mean	Target value, 120 µg/m³	Not to be exceeded on more than 25 days per year, averaged over three years	100 µg/m³	
NO_2	One hour	Limit value, 200 µg/m³ (*)	Not to be exceeded more than 18 times a calendar year	200 µg/m³ (*)	
NO_2	Calendar year	Limit value, 40 µg/m³		40 µg/m³	



Long term NO2 predictions

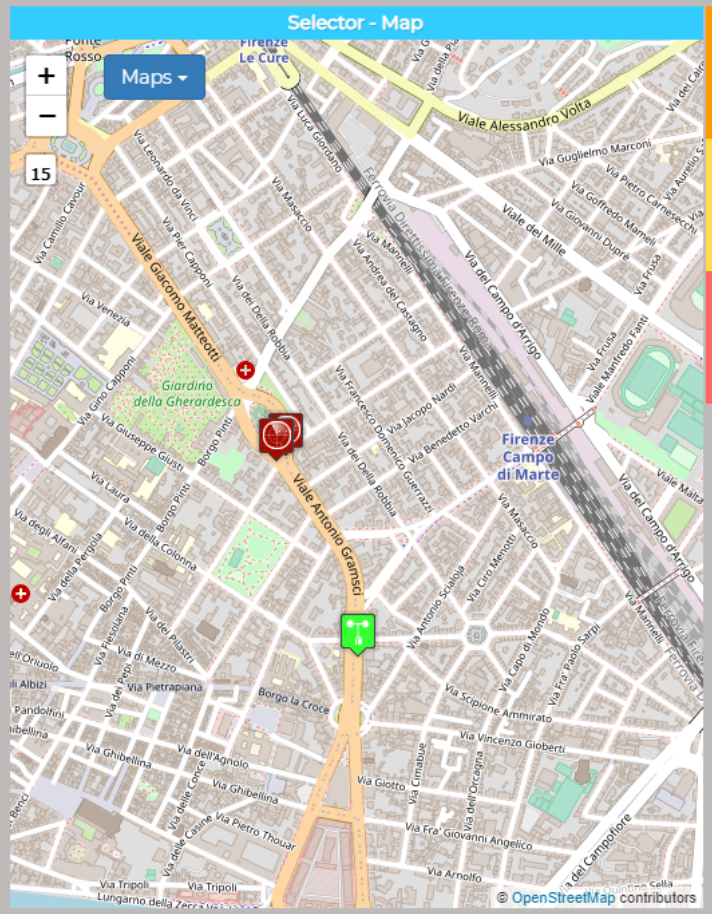
- Threshold value of NO2 mean average has been passed since 2014 in Florence
  - Gramsci Sensors and area
- Aim
  - Analysing the trends
  - Predicting 180 days in advance the value

		Air Quality Directive		WHO guidelines	
Pollutant	Averaging period	Objective and legal nature and concentration	Comments	Concentration	Comments
$PM_{2.5}$	One day			25 µg/m³ (*)	99th percentile (3 days/year)
$PM_{2.5}$	Calendar year	Target value, 25 µg/m³	The target value has become a limit value since 1 January 2015	10 µg/m³	
PM_{10}	One day	Limit value, 50 µg/m³	Not to be exceeded on more than 35 days per year.	50 µg/m³ (*)	99th percentile (3 days/year)
PM_{10}	Calendar year	Limit value, 40 µg/m³ (*)		20 µg/m³	
O_3	Maximum daily 8-hour mean	Target value, 120 µg/m³	Not to be exceeded on more than 25 days per year, averaged over three years	100 µg/m³	
NO_2	One hour	Limit value, 200 µg/m³ (*)	Not to be exceeded more than 18 times a calendar year	200 µg/m³ (*)	
NO_2	Calendar year	Limit value, 40 µg/m³		40 µg/m³	

NOX reduction for COVID

Monitoraggio Area Gramsci: NO2 vs Traffico

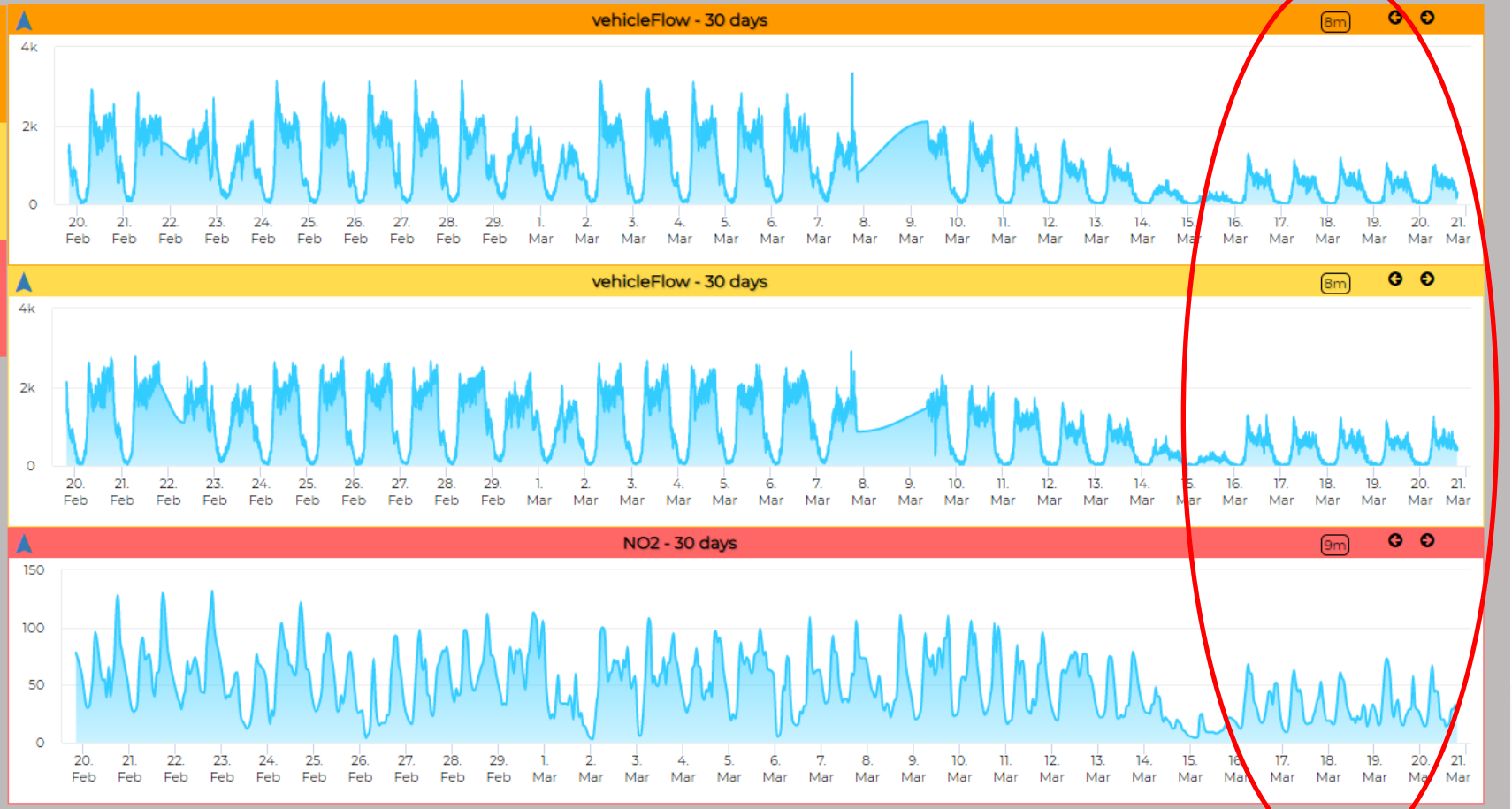
Sun 19 Apr 19:16:42



InOut traffic Firenze

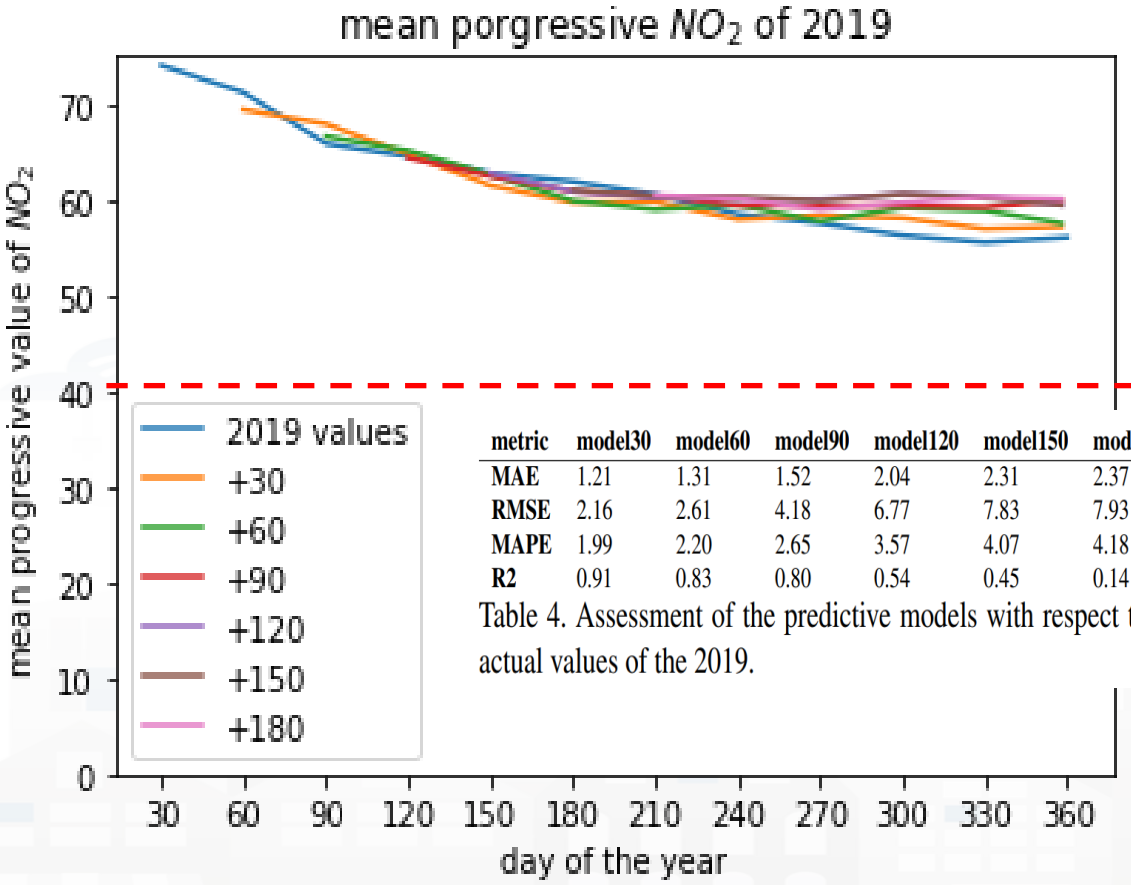
NOX Predictions

Real Time Traffic



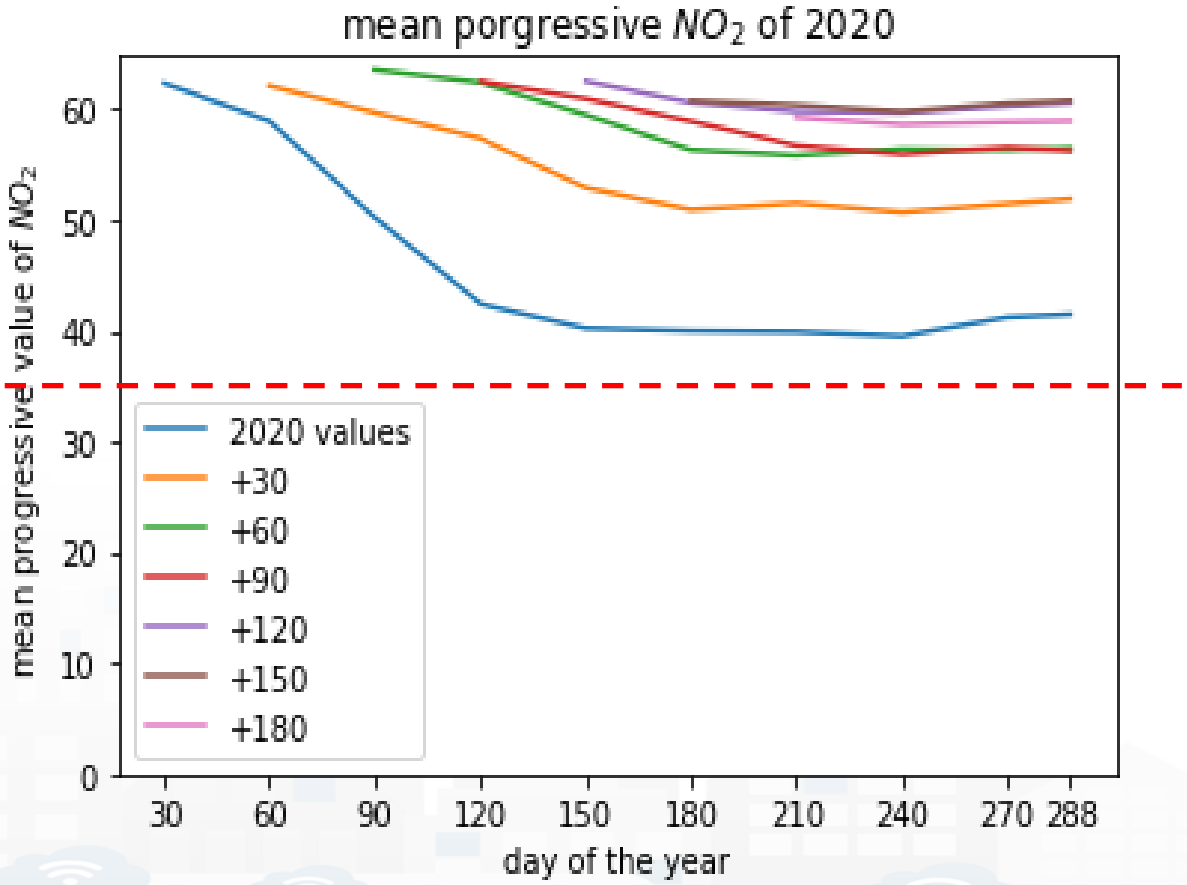
Using data since 2014

Predicting: the 2019 the 2020



metric	model30	model60	model90	model120	model150	model180
MAE	1.21	1.31	1.52	2.04	2.31	2.37
RMSE	2.16	2.61	4.18	6.77	7.83	7.93
MAPE	1.99	2.20	2.65	3.57	4.07	4.18
R2	0.91	0.83	0.80	0.54	0.45	0.14

Table 4. Assessment of the predictive models with respect to the actual values of the 2019.



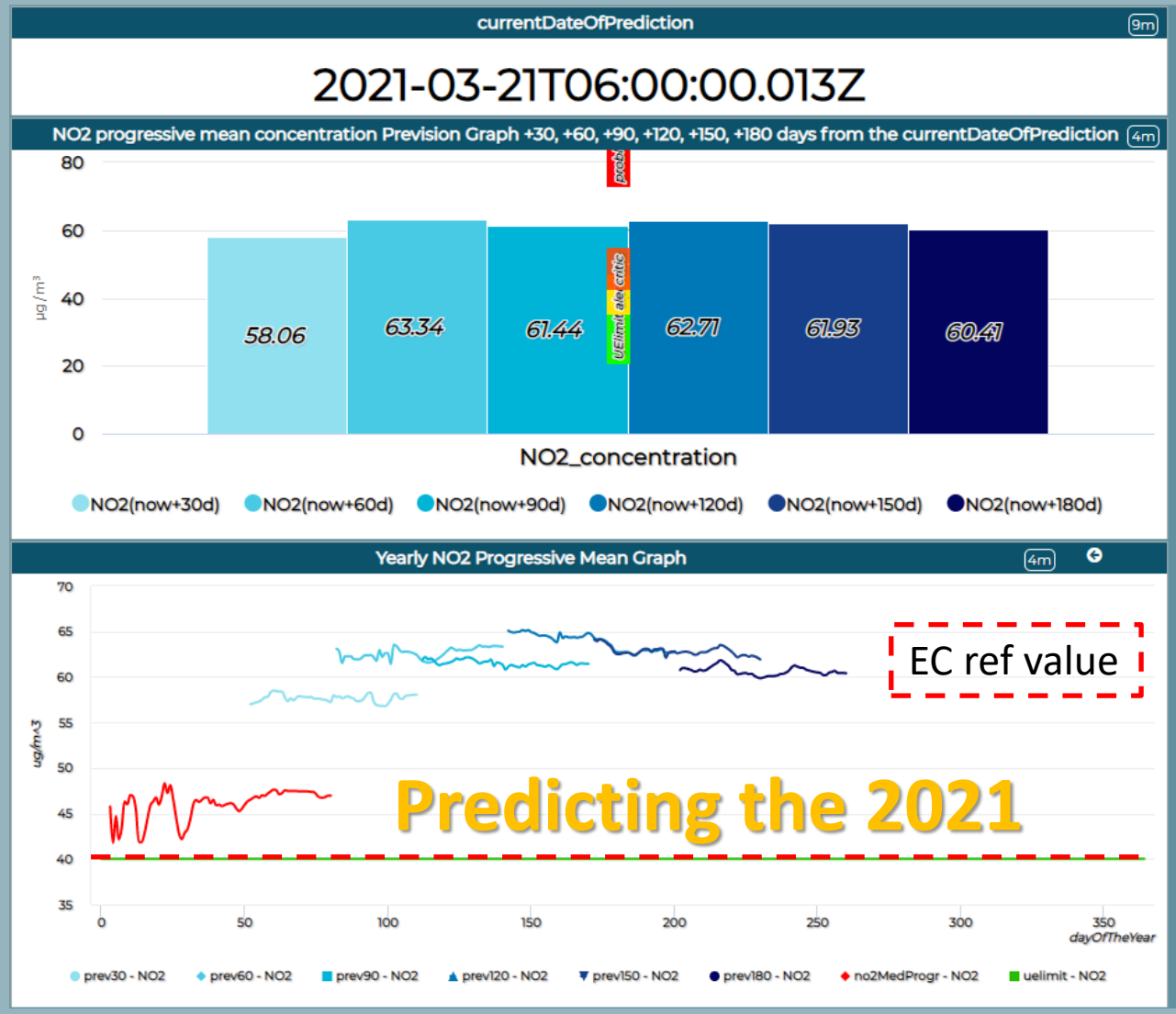
Deep Learning Approach

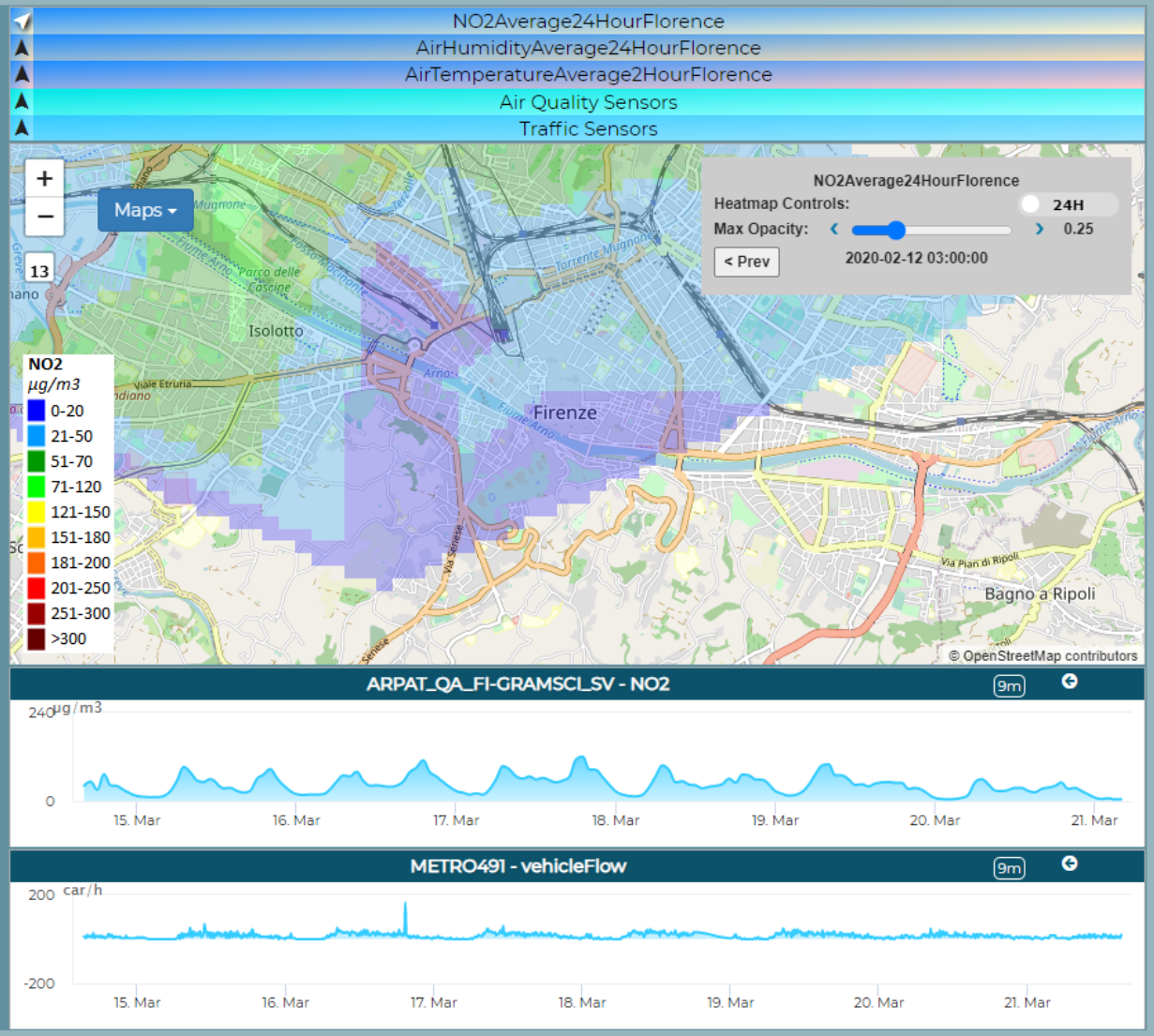
EC ref value

Yearly NO2 Florence Monitoring Dashboard

This dashboard shows the predicted NO2 concentrations from the current date of prediction of 30, 60, 90, 120, 150 and 180 days ahead. The European Union has set the limit on the mean yearly value of NO2 concentration to 40 ug/m3.

Sun 21 Mar 15:18:47





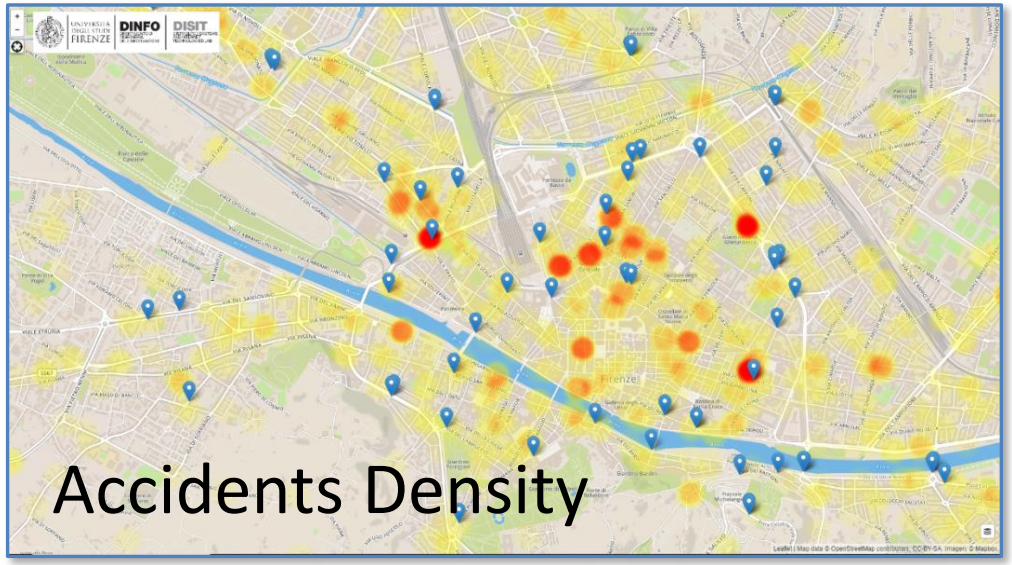
Privacy Policy Cookies Policy Terms and Conditions Contact us

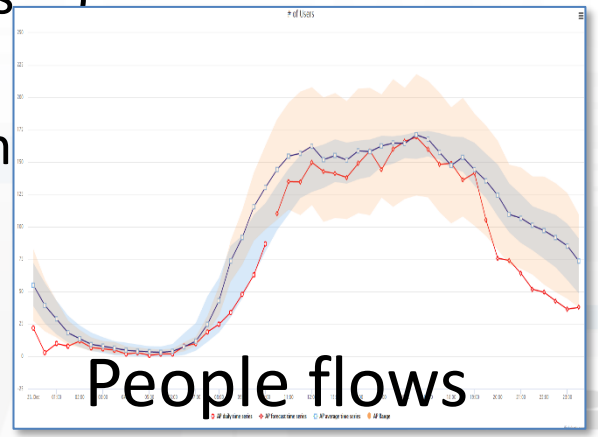
https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MzA2OQ==

Anomaly Detection
Early Warning

Anomaly Detections

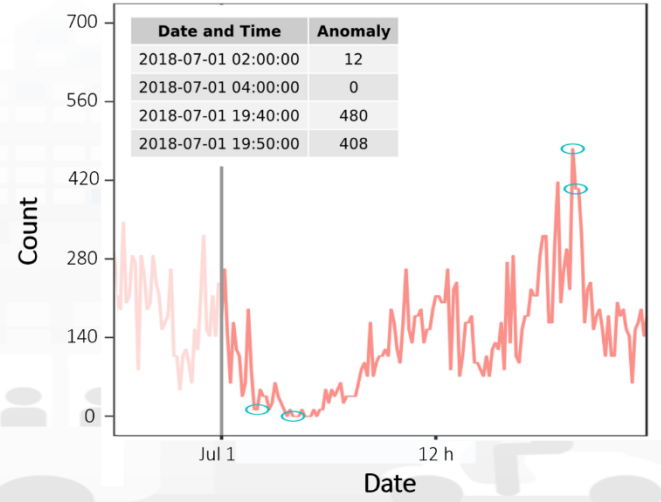
- About the IoT Devices status
  - Eventual problems on IoT Devices, connections, etc.
- About People Flows and Density
  - Early warning of the ineption of critical events
- About traffic flow
  - Early warning on eventual incidents, or on the inception of critical conditions on the traffic (e.g., a reduction in viability, a broken bus…)
- About….
  - Early warning, early detection of problem
- Recurrence analysis
- Causal Analysis





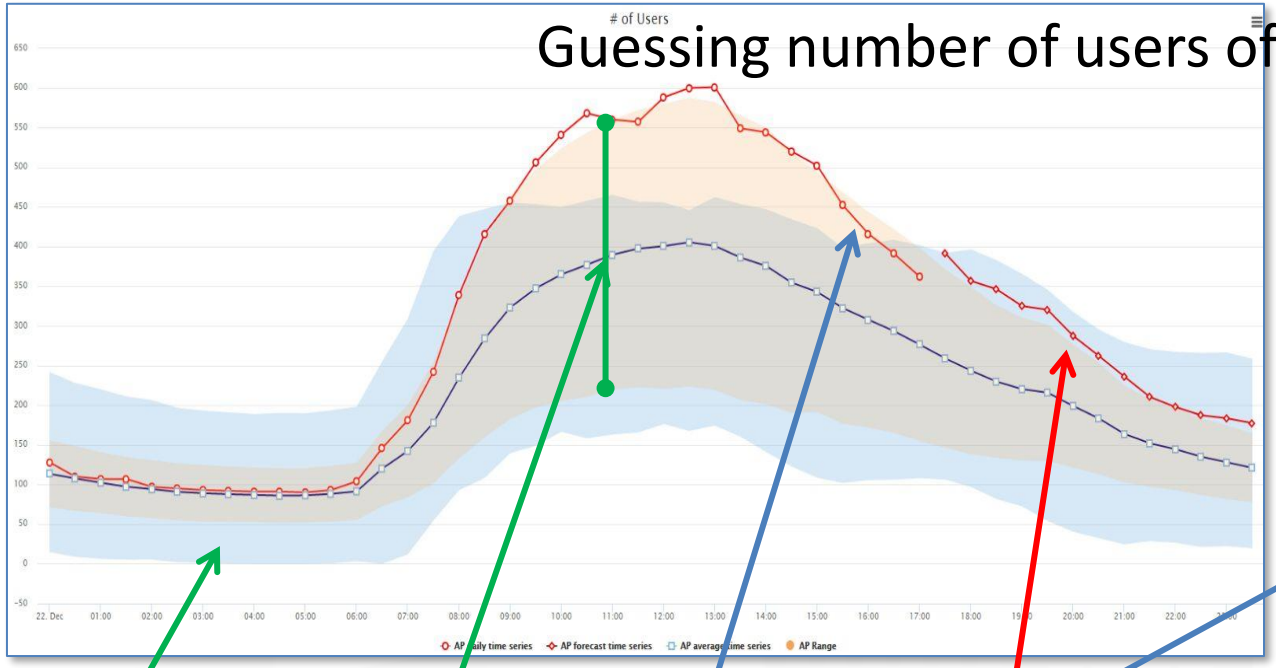
METRO 705

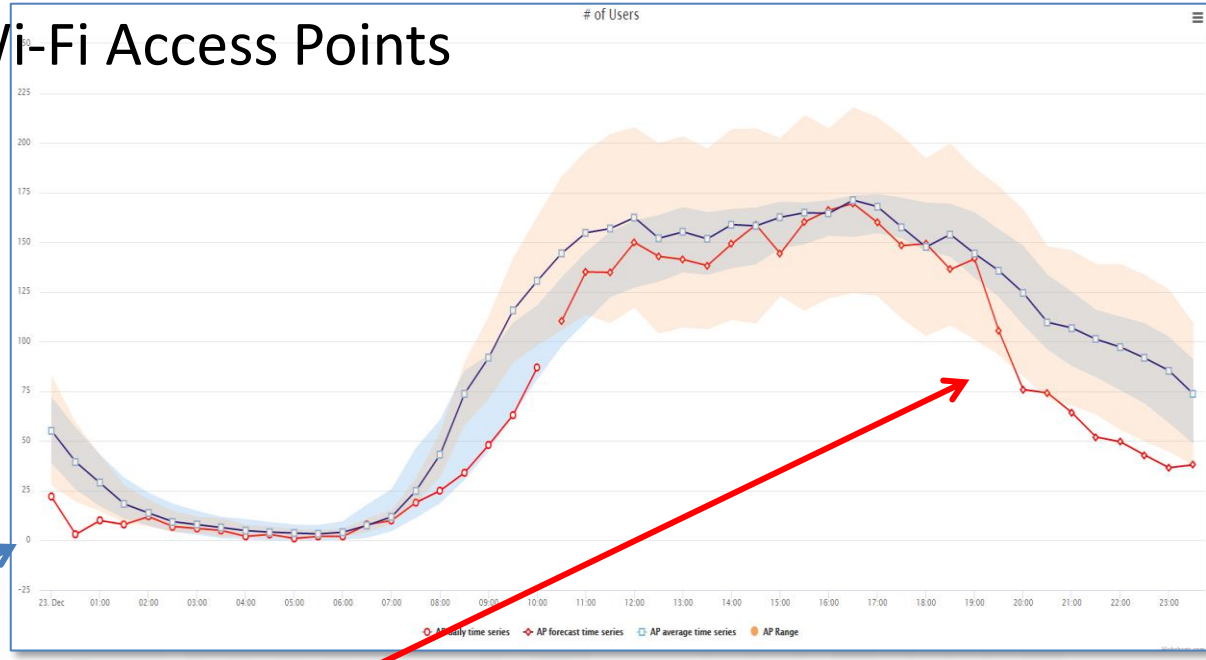
Date and Time	Anomaly
2018-07-01 02:00:00	12
2018-07-01 04:00:00	0
2018-07-01 19:40:00	480
2018-07-01 19:50:00	408



Prediction and Identification of Anomalies

Guessing number of users of Wi-Fi Access Points





Cluster confidence

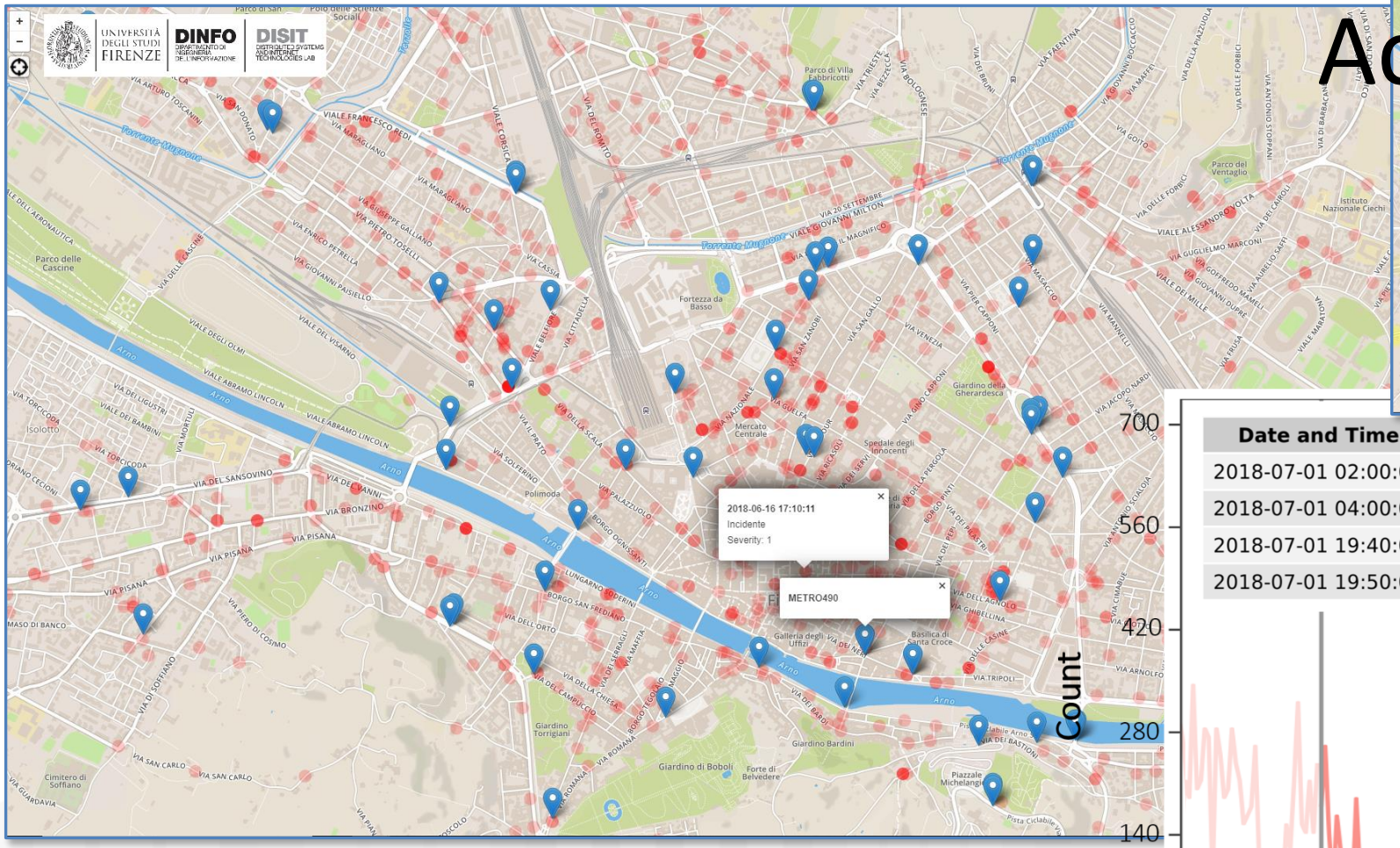
AP average and confidence

Actual AP trend for today

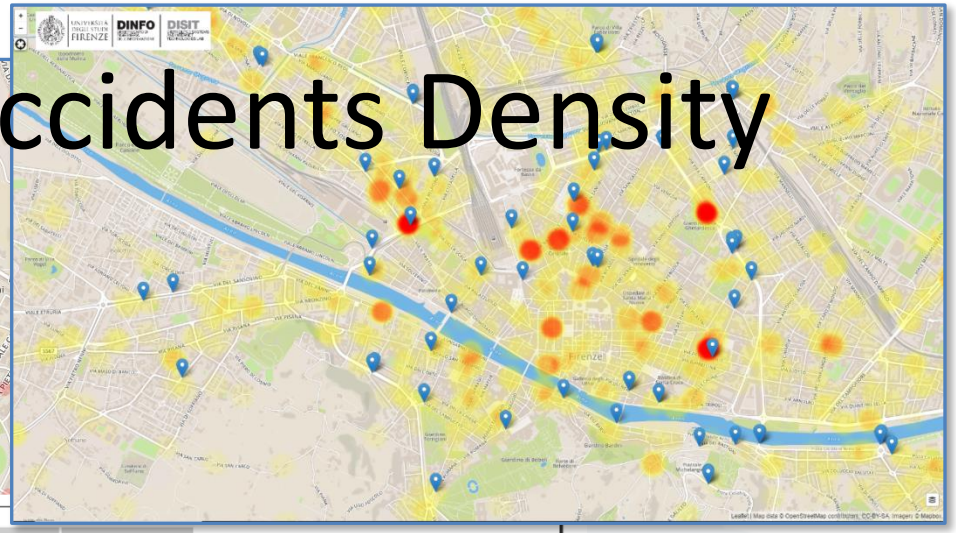
AP prediction for the next time slot in the day on the basis of past weeks

Predictive precision of the 95%

Anomaly Detection

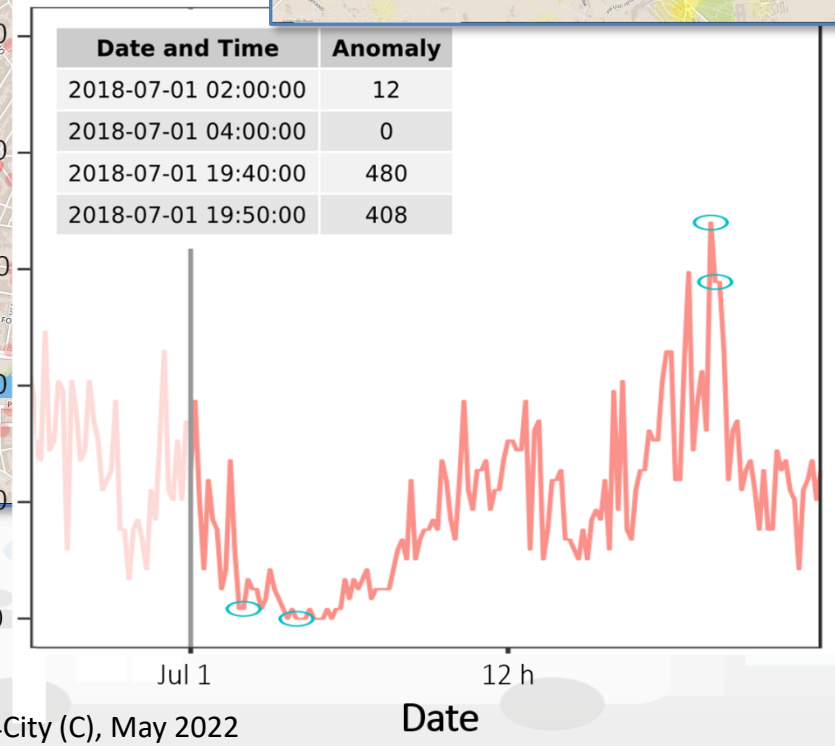


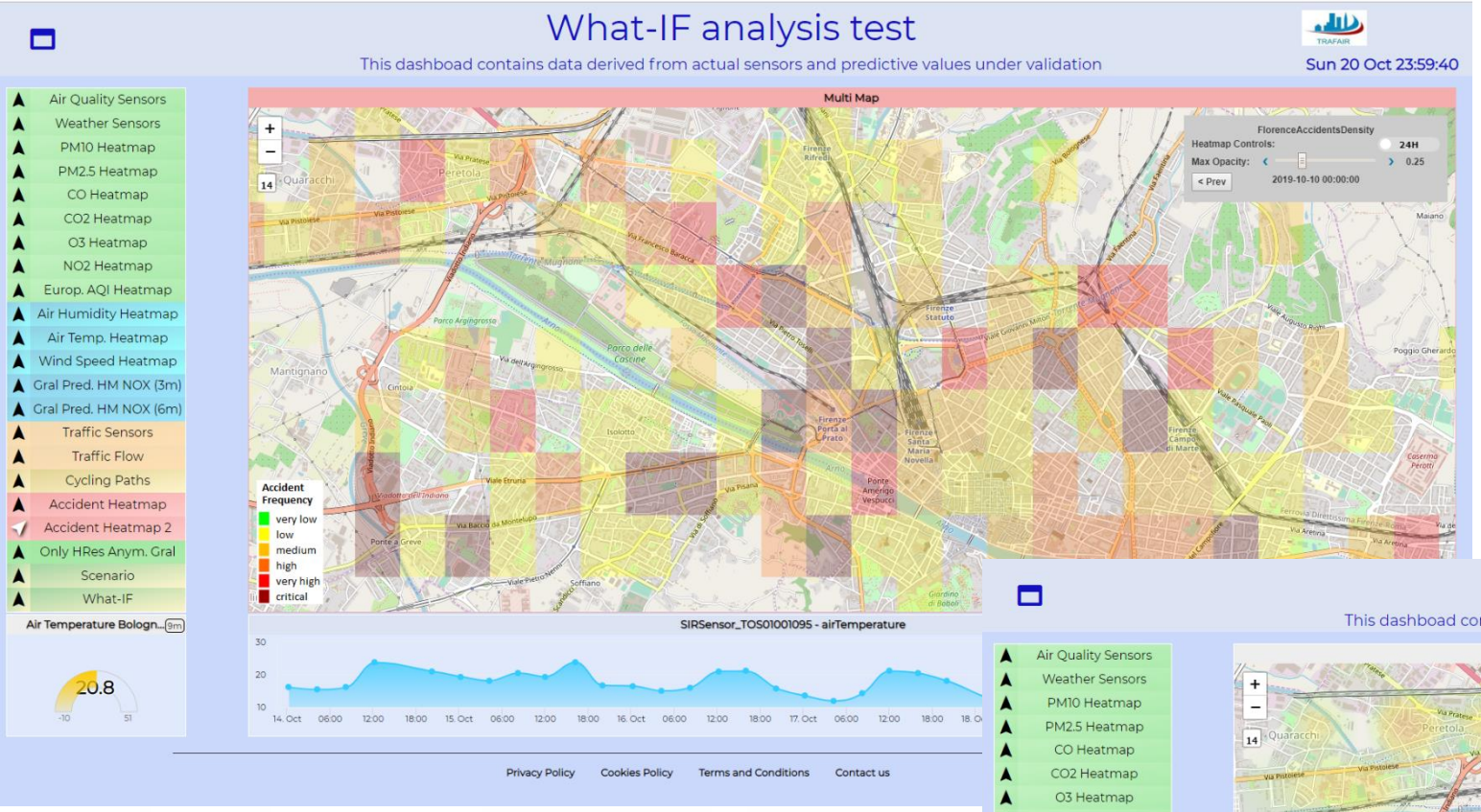
Accidents Density

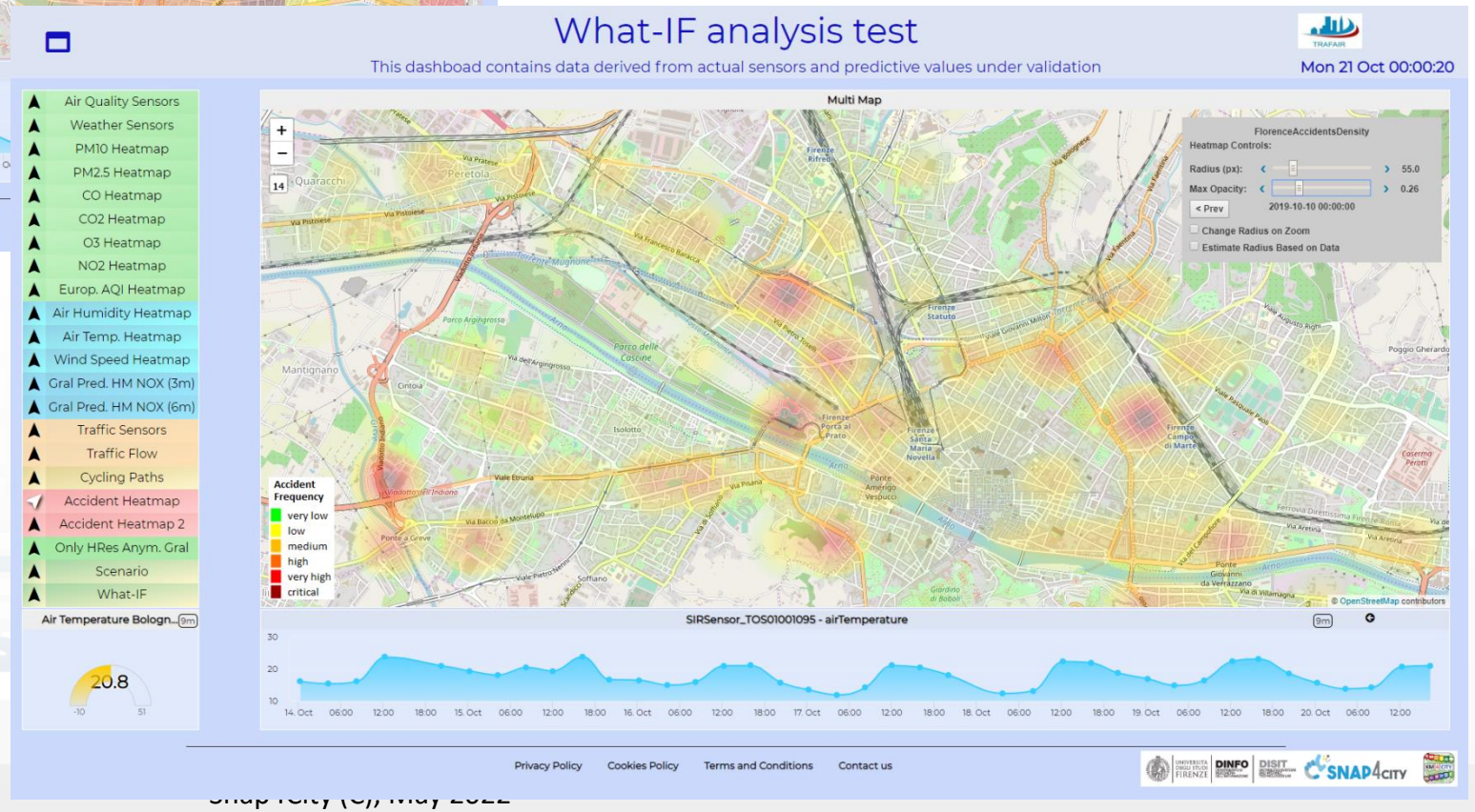


Accidents vs Traffic

Date and Time	Anomaly
2018-07-01 02:00:00	12
2018-07-01 04:00:00	0
2018-07-01 19:40:00	480
2018-07-01 19:50:00	408





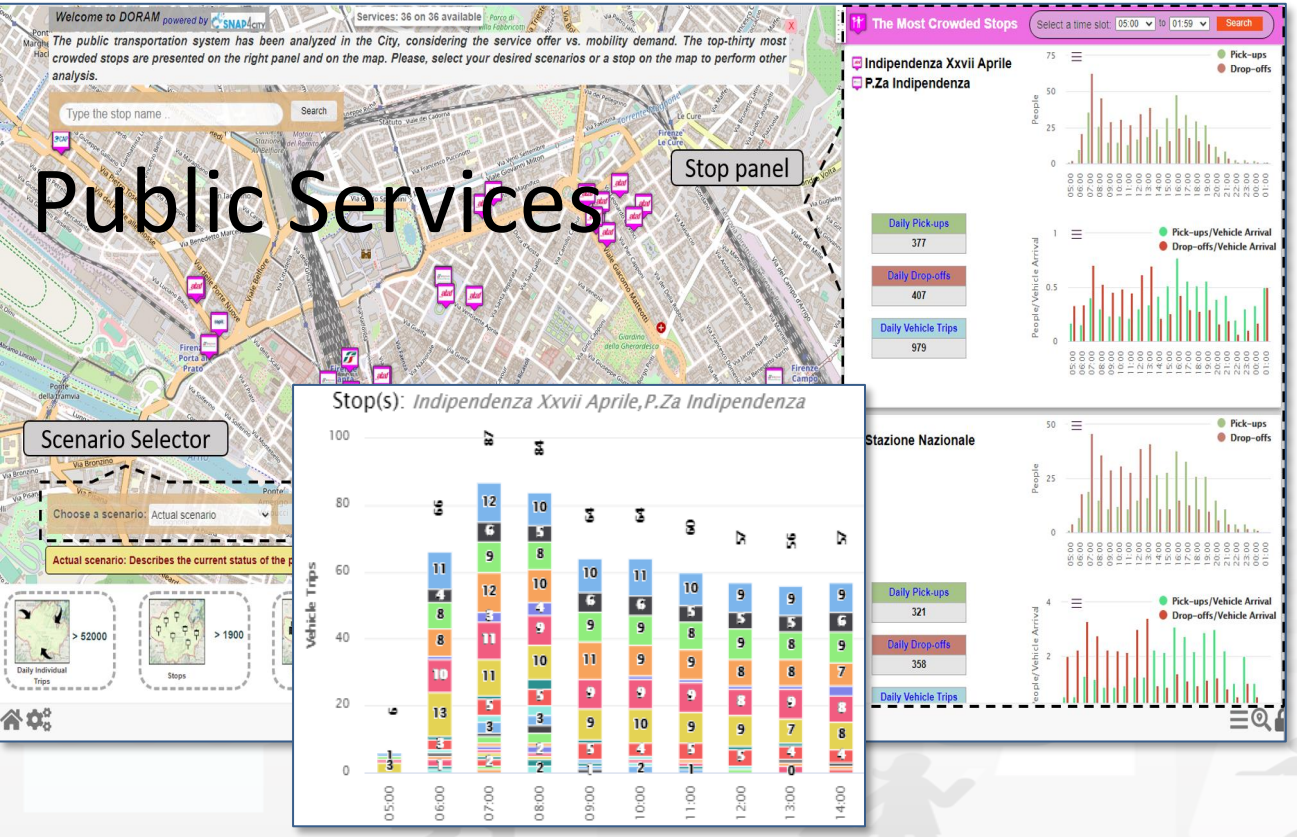


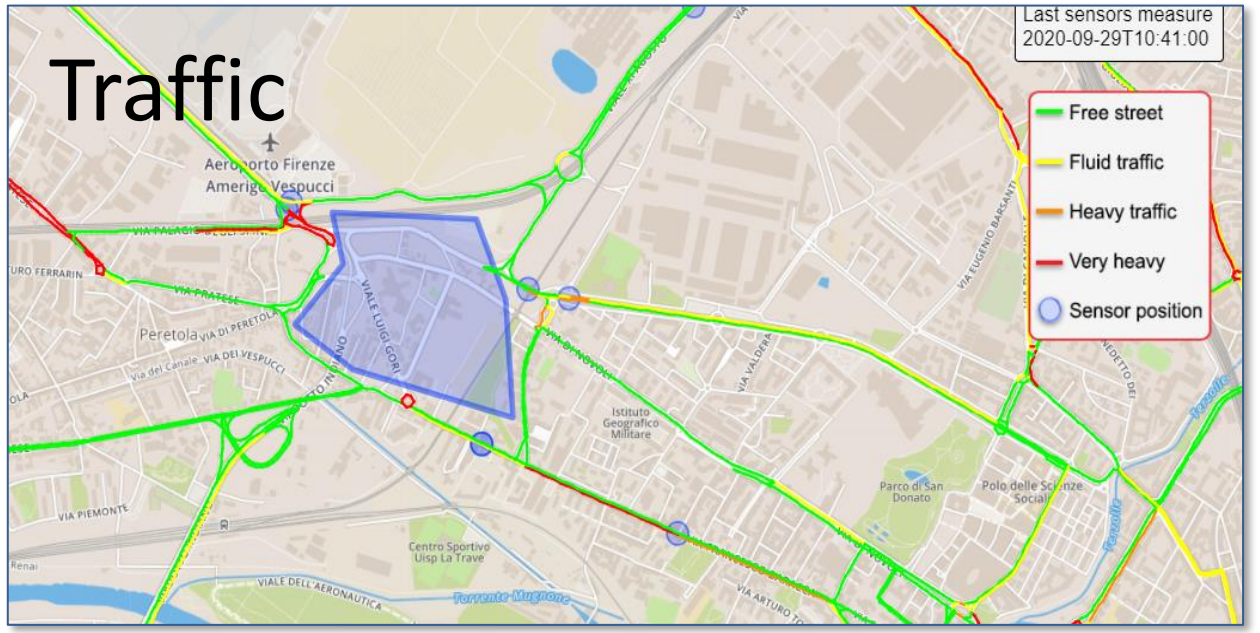
https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MjE4Nw==

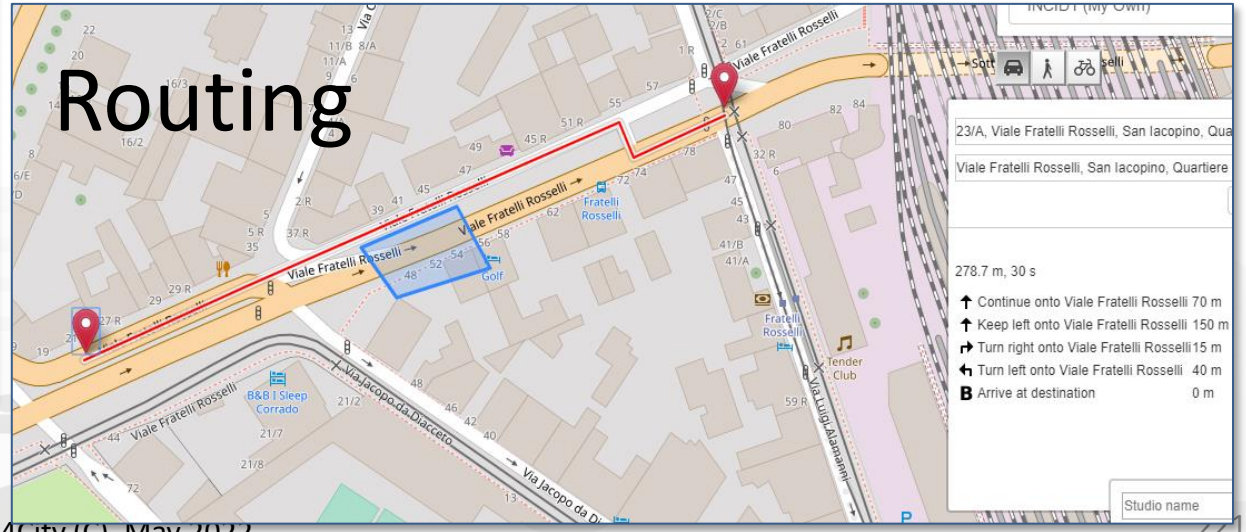
WHAT-IF Analysis

What-if Analysis

- Definition of scenarious impact on
  - Traffic, Pollutant, parking, public transport, private flows, etc.
  - KPI analysis



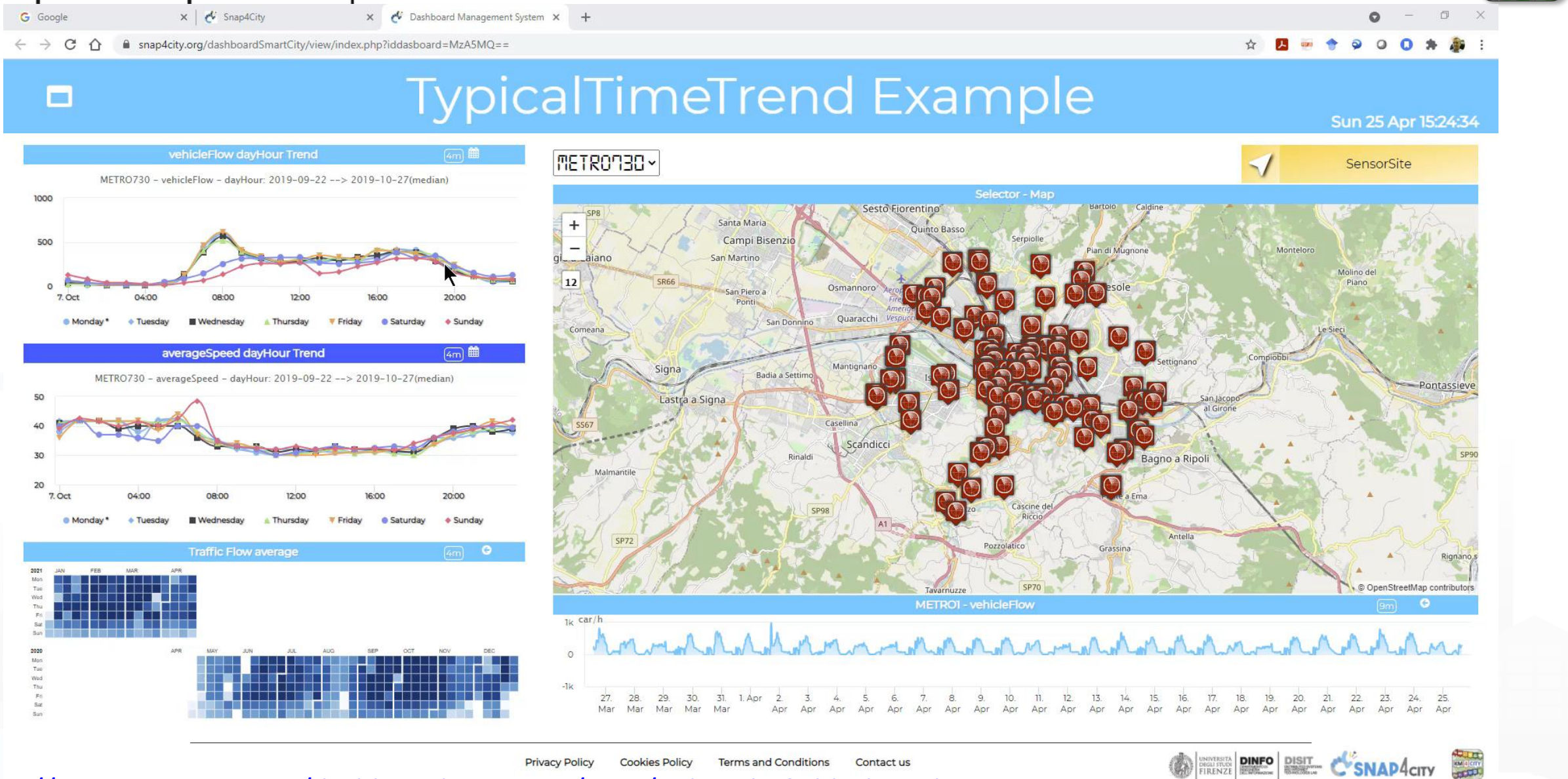




What-If Analysis

Available data and techniques	What happened	What is going on now	What is going to happen	What-If: what is going to happen if a scenario occurs in the future	Which is the best solution
Historical Data, HD	Yes	No	No	No	No
Real Time Data, RTD	No	Yes	No	No	No
HD + RTD + Short term Predictions, STP(.)	Yes	Yes	Yes	No	No
HD + RTD + Analytical Model, AM(.) + Scenario Model, SM(.)	Yes	Yes	Yes	(Yes)	No
HD + RTD + Short and Very Long Term Predictions, SVLTP(.) + AM(.) + SM(.) + Simulation, S(.)	Yes	Yes	Yes	Yes	No
HD + RTD + SVLTP(.) + AM(.) + SM(.) + S(.) + KPI(.) based Decision	Yes	Yes	Yes	Yes	Yes

Typical Time Trends

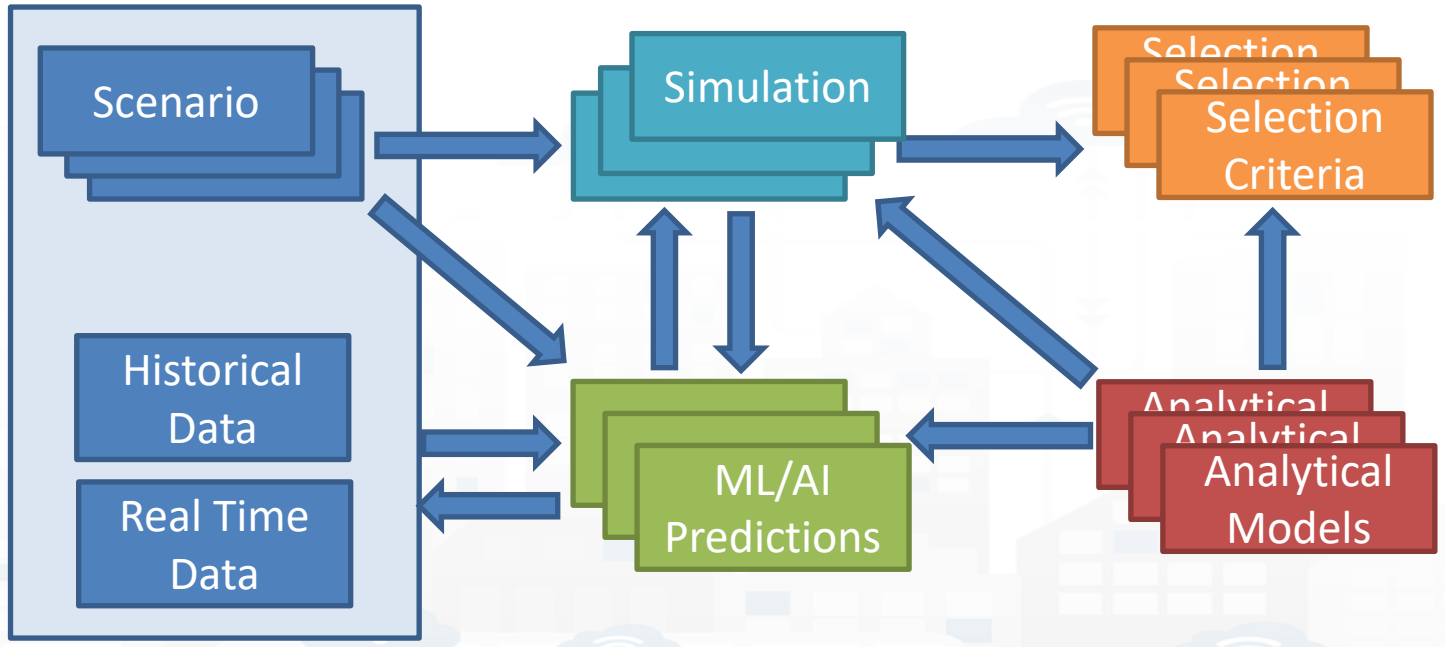


https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MzA5MQ==

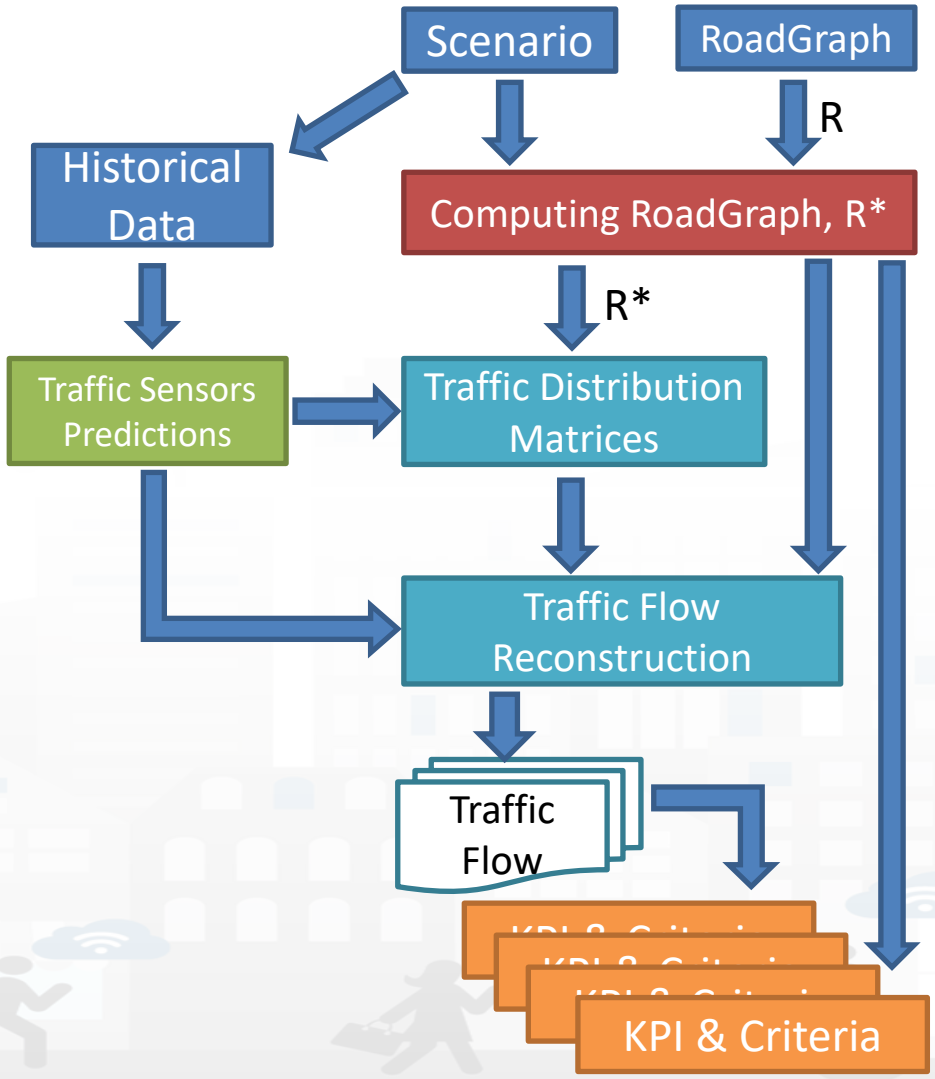
What-if: what is going to happen if …. this and that

- **What is going to happen at:**
  - People, Economy, Society, ..
  - Traffic, Pollutant, Parking, structures
  - Equipment, .. …
- **if certain unexpected events would occur**
  - Scenario definition
  - Guessing future data…
- **Taking into account**
  - Historical Data
  - Real Time Data
  - Contextual data

Decision Support System
KPI, Optimization
Visual Analytic: animations



What-if: Simumation for Traffic Flow



At the same color corresponds the same area:

- Data / information
- Data Analytics
- Selection Criteria
- Predictions / imputation
- Simulation

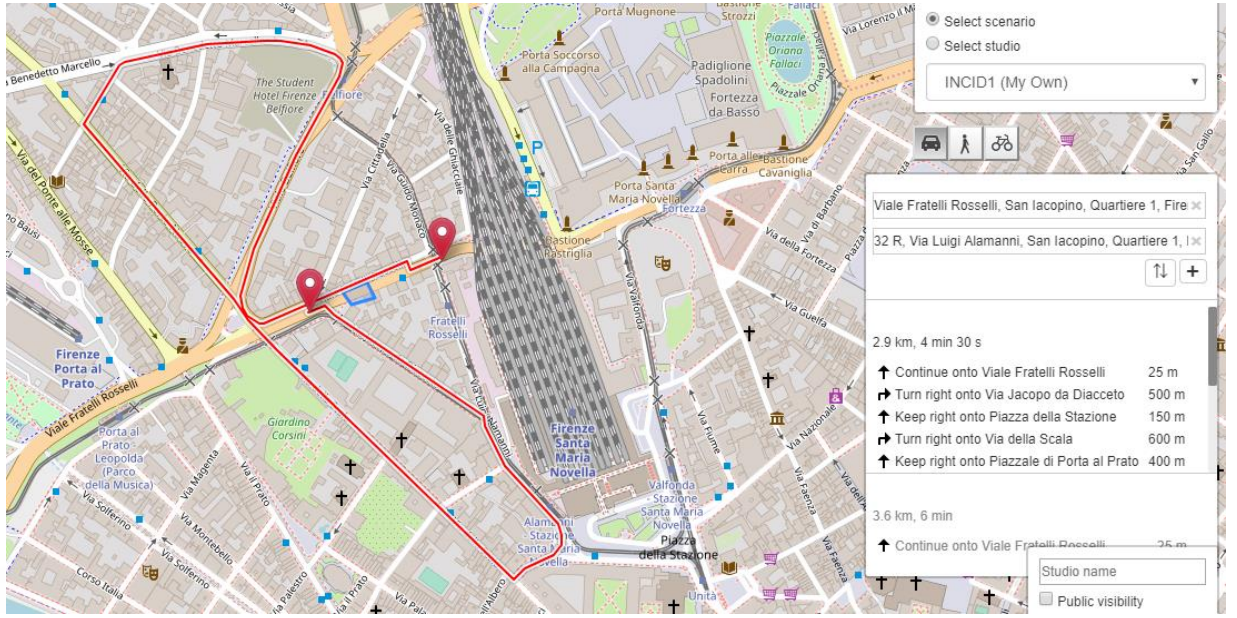
What-If Analysis

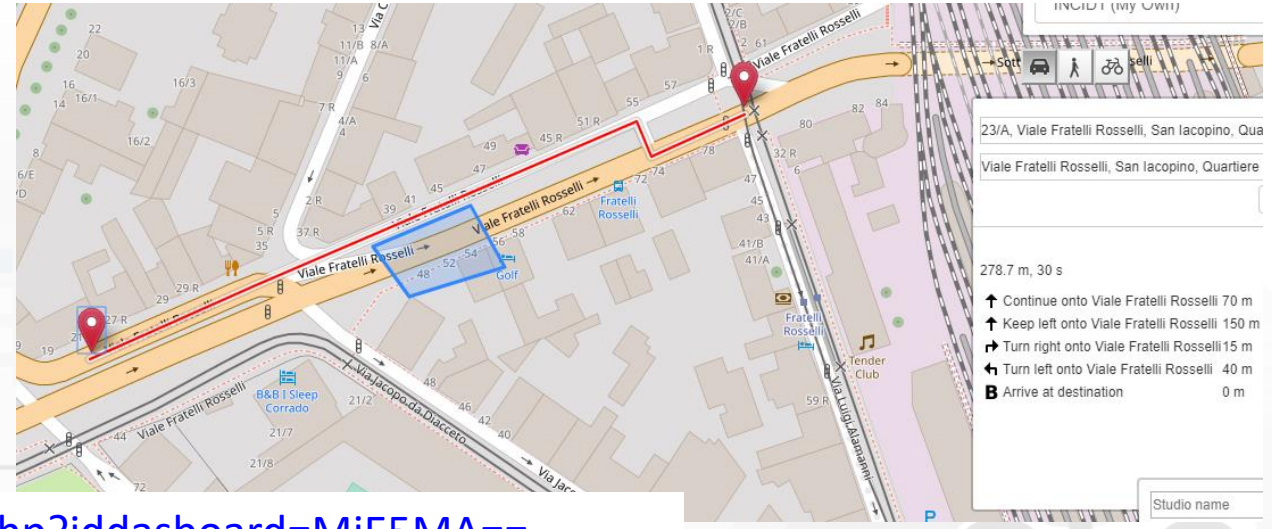
Accidents and elements blocking Points and Shapes taken into account for:

- Routing
- Traffic Flow reconstruction
- Evacuation paths
- Rescue team paths

Assessment on the basis of changes:

- Mobility demand assessment
- Mobility Offer assessment





https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MjE5MA==

Mobility and Environment What-IF Analysis

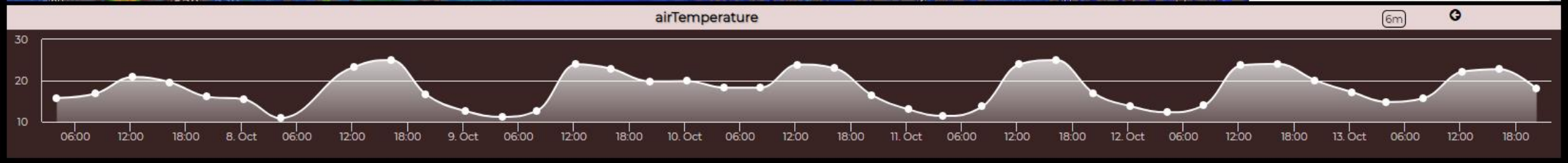
This dashboad contains data derived from actual sensors and predictive values under validation

Mon 14 Oct 00:48:17

- Air Quality Sensors
- Weather Sensors
- PM10 Heatmap
- PM2.5 Heatmap
- CO Heatmap
- CO2 Heatmap
- O3 Heatmap
- NO2 Heatmap
- Europ. AQI Heatmap
- Air Humidity Heatmap
- Air Temp. Heatmap
- Wind Speed Heatmap
- Gral Pred. HM NOX (3m)
- Gral Pred. HM NOX (6m)
- Traffic Sensors
- Traffic Flow
- Cycling Paths
- Accident Heatmap
- Only HRes Anym. Gral
- Scenarious
- What-if analysis

Firenze Oggi

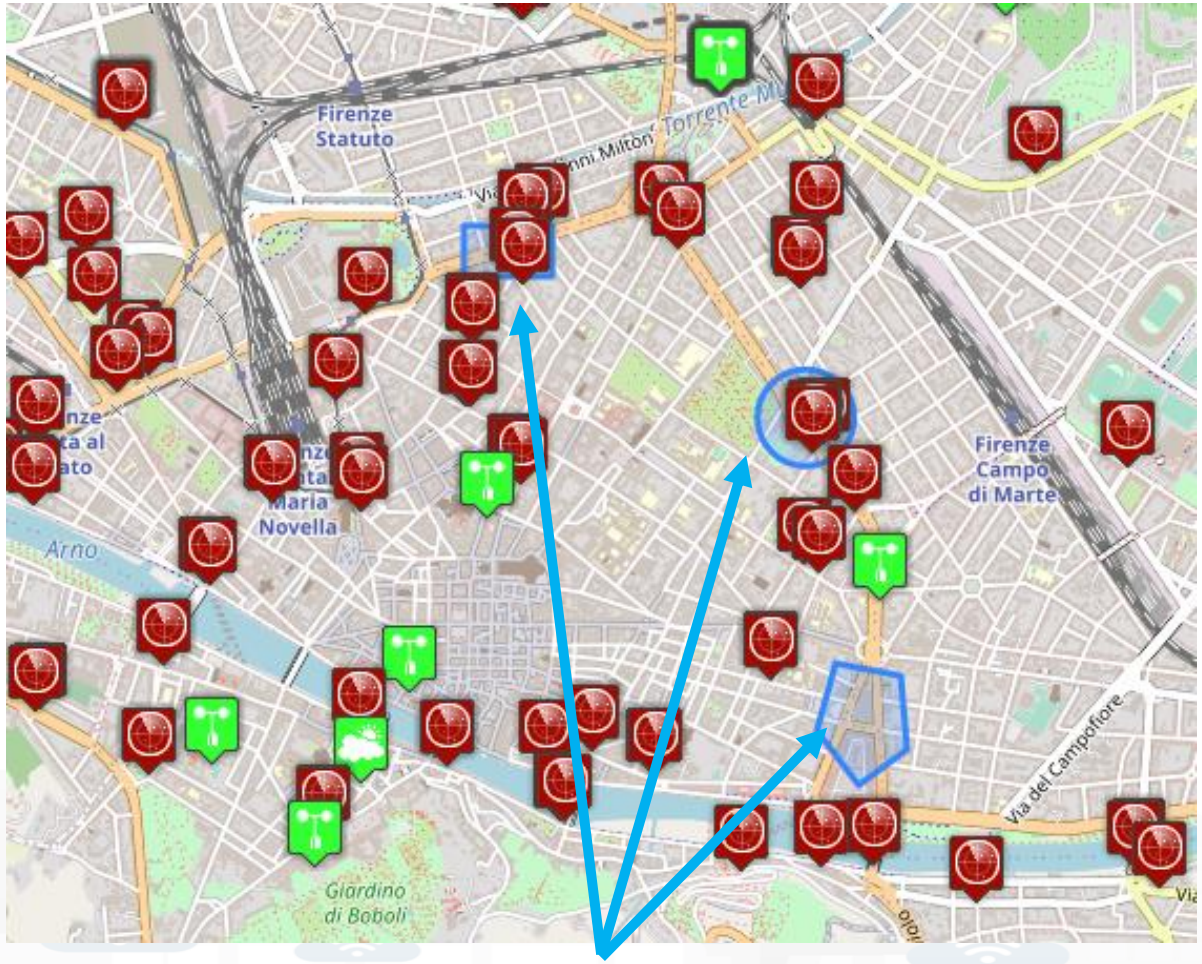
Air Temperat... 6m



https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MjE5MA==

What-If Analysis Concepts

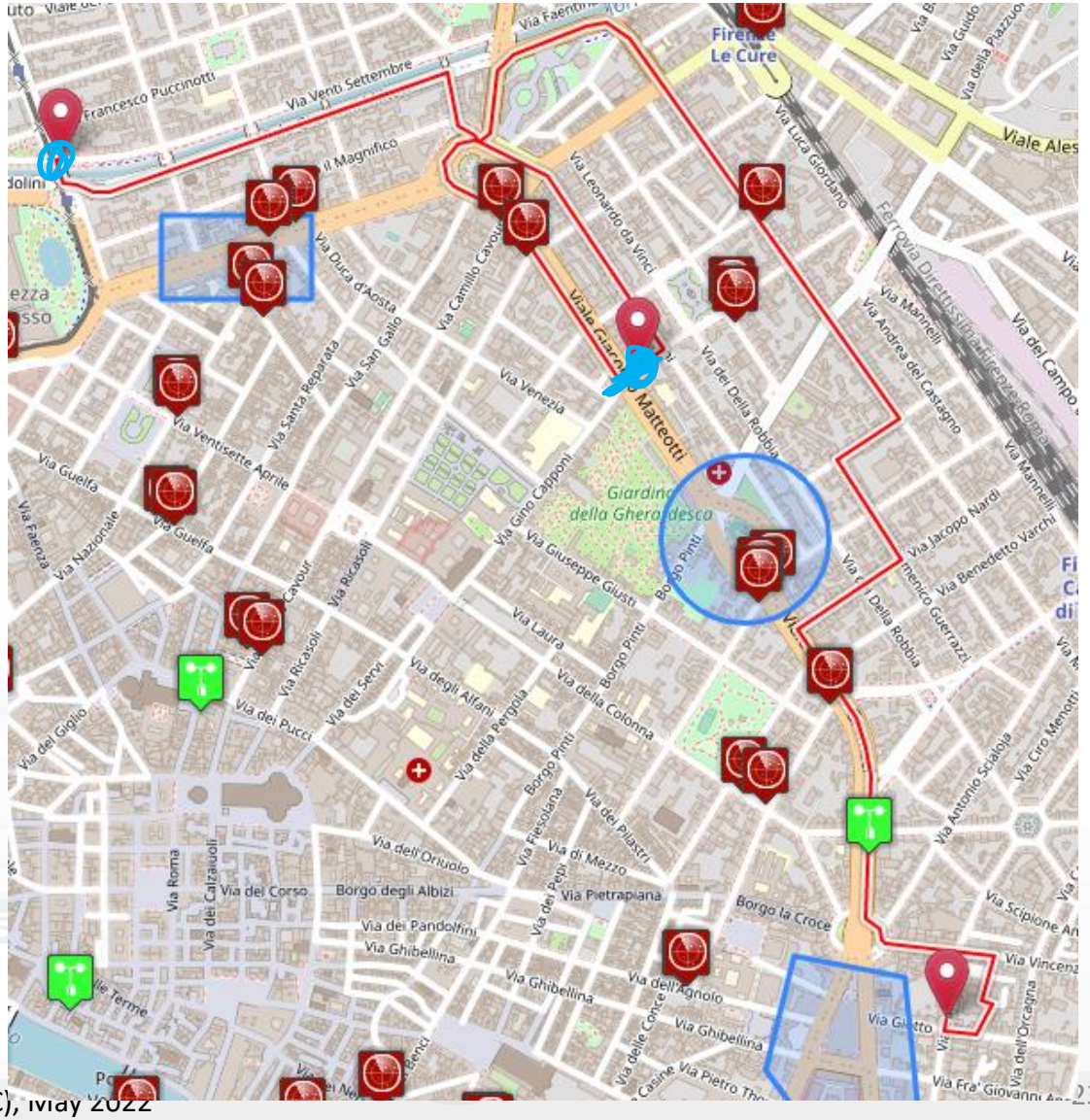
- What is going to happen at Services if certain conditions/cases are going to occur
- Formalize: Conditions/cases, Services
- Scenarios of Cases+Services Vs Solutions are Studios
- You can define, save, load:
  - Scenarios and Studios



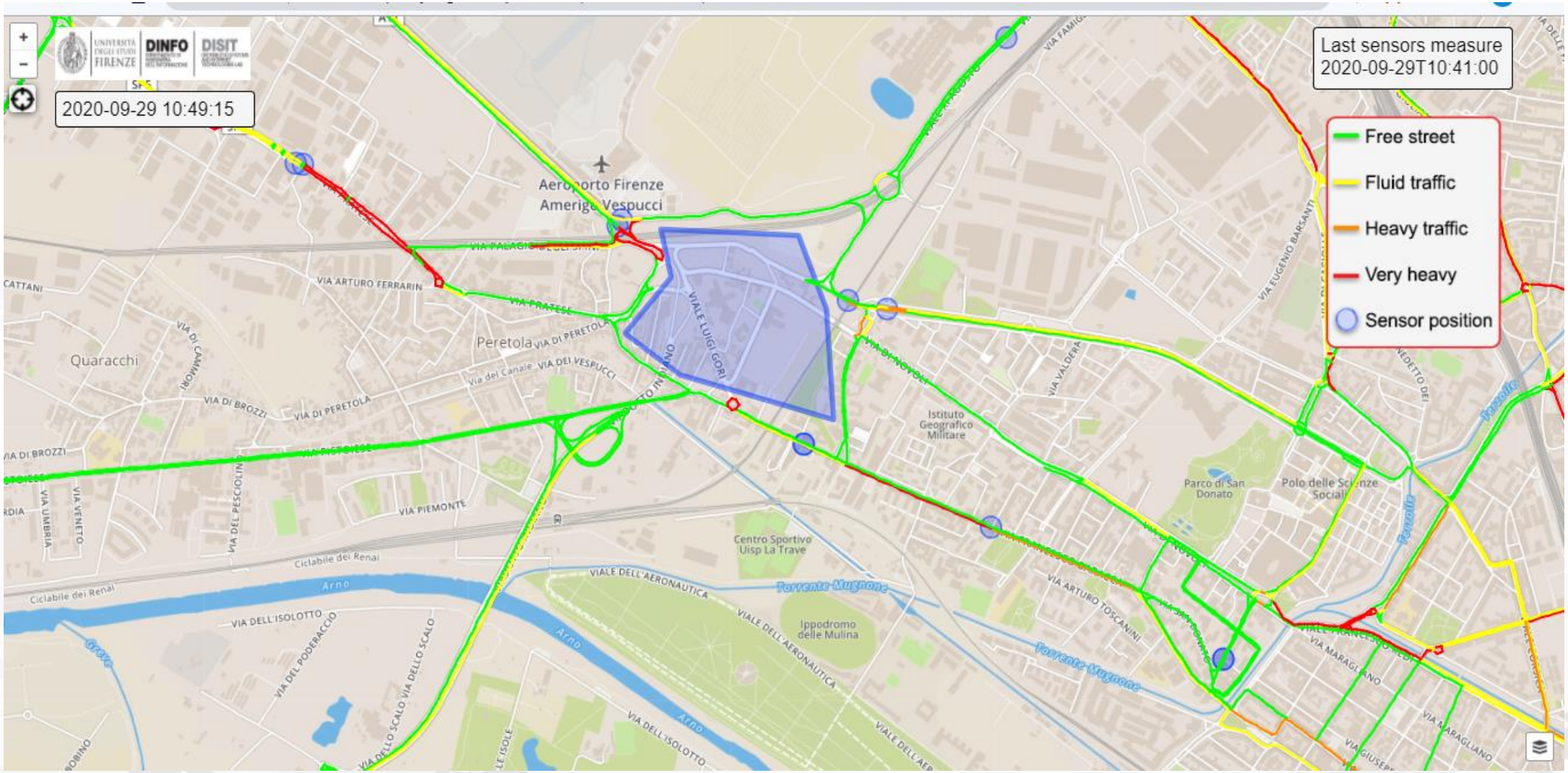
Scenario

Impact on Routing

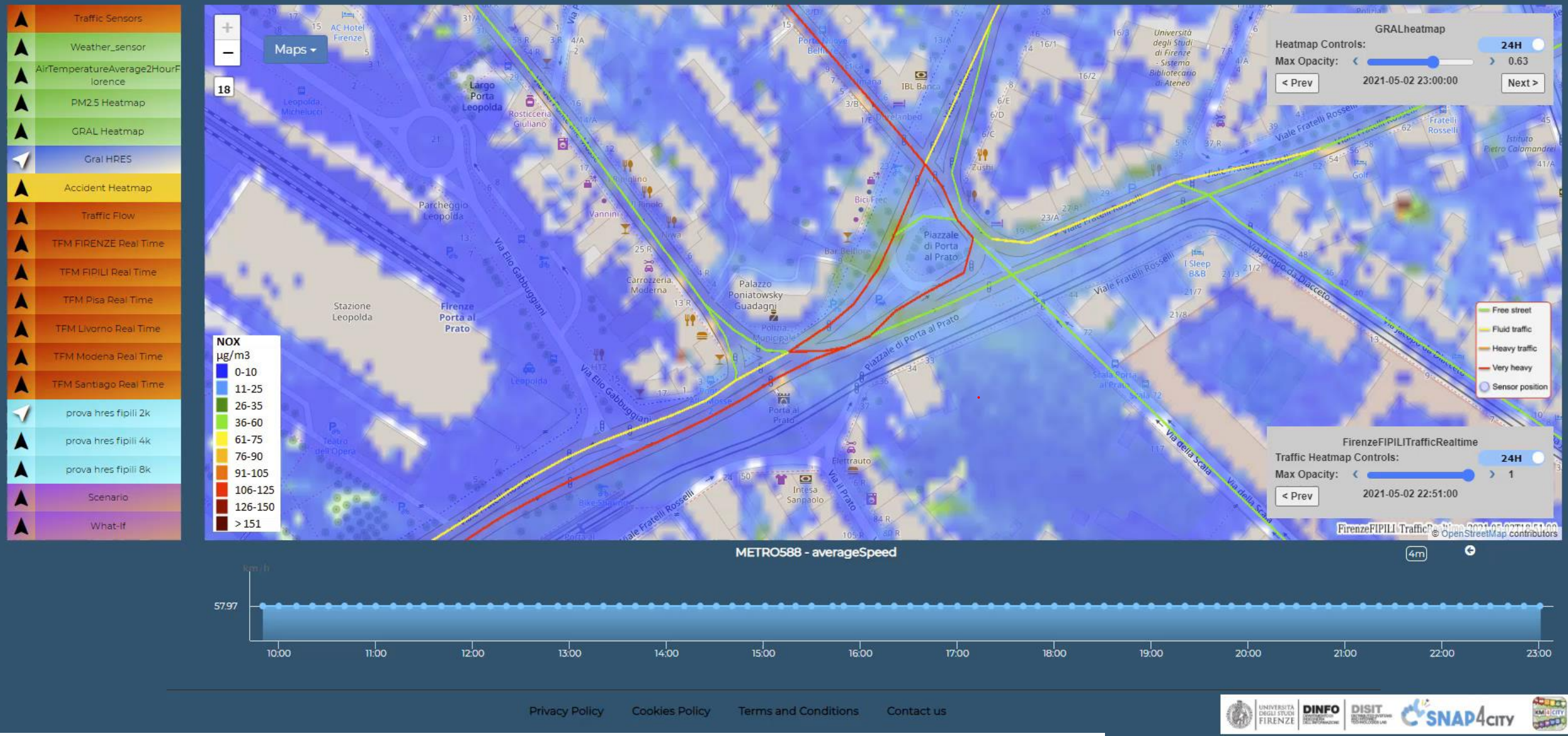
- Scenario with multiple shapes
- Conditional Routing
  - avoiding areas or
  - reducing traffic in those areas
  - Multiple stop points



Computation of Traffic Flow Evolution



Traffic Flow Manager on multiple cities



https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MzEyNg==

RESOLUTE Dashboard 4XHD v5

Wed 21 Jun @ 12:49:14

ESB

Alarms

TEST - ALLUVIONE MUGNONE
MAJOR
07/08/2017 12:45:54

ALLUVIONE ARNO COVERCIANO
CRITICAL
15/08/2017 15:47:31

ALLUVIONE ARNO VIA PETRIOLO
MAJOR
07/08/2017 16:48:11

TEST - SCONTRI ALLO STADIO
CRITICAL
15/08/2017 09:50:04

Traffic events

SEVERE ACCIDENT
SERIOUS ACCIDENT(S)
21/08/2017 11:24:01 5

ROAD WORKS
ROADWORKS
21/08/2017 11:24:10 6

ORD 2017-001302 - SP NÃ°17 Ã??ALTO VALDARNOÃ??IST...
RESTRICTIONS
13/08/2017 00:00:00 5

INTERVENTO URGENTE PUBLIACQUA
CLOSED AHEAD
20/08/2017 10:00:00 10

Evacuation plans

PLAN N.1497263867
IN_PROGRESS
12/06/2017 - 11:37:47 PATHS: 3

PLAN N.2000000000
PROPOSED
16/08/2017 - 09:42:06 PATHS: 2

PLAN N.1497604404
PROPOSED
16/08/2017 - 11:06:47 PATHS: 1

Tram position

TRAM/1
21/08/2017 11:32:27

TRAM/2
21/08/2017 11:32:31

Network Analysis

FM0495
(NORMALIZED) NODE BETWEENNESS
1.0000

FM0496
(NORMALIZED) NODE BETWEENNESS
0.8091
ON MAP

FM0500
(NORMALIZED) NODE BETWEENNESS
0.7878

FM2276
(NORMALIZED) NODE BETWEENNESS
0.7856

FM0497
(NORMALIZED) NODE BETWEENNESS
0.7825

FM0701

Environment

PROTEZIONE CIVILE

NULLO BASSO MEDIO ALTO

RISCHIO IDRAULICO
RISCHIO TEMPORALI
RISCHIO IDROGEOLOGICO
RISCHIO NEVE
RISCHIO GHIACCIO
RISCHIO VENTO
RISCHIO MAREGGIATE

Anemomet... 221

Ground Te... 22 C°

Temperatu... 24 C°

Umidity 20 %

Pluv. Flore... 24 mm

Pluviometer Florence City - last 24h

Hydromete... 20 m.s.z.l.

Hydrometer Arno Nave di Rosano - Last 24h

Hidromete... 24 m.s.z.l.

Hydrometer Arno Florence Uffizi - Last 24h

Hydromete... 24 m.s.z.l.

Hydrometer Mugnone Ponte alle Mosse - Last 24h

Mobility

Bus Lines Status

ULTIMO AGGIORNAMENTO: 2016-08-04 23:59:03

Underpass Viale Mariti Aperto

Underpass Viale Talenti closed

Underpass Fortezza Aperto

Parterre Pa... 76.2% 23.8%

SMN 65.7% 34.3%

Fortezza Fi... 37.2% 62.8%

Mercato Ce... 65.4% 34.6%

Resources

Planned events

E 1966 - 2016 - LA BELLEZZA SALVATA - PROROGATA FINO AL 2
2016-12-01 to 2017-07-02
VIA CAMILLO CAVOUR 3

NE SPLORATI.LE BIBLIOTECHE DELL'UNIVERSITÀ DI FIRENZE IN
2017-02-15 to 2017-06-23
PIAZZA DI SAN LORENZO 9

BIBLIOTECHE DELL'UNIVERSITÀ DI FIRENZE IN MOSTRA - PROR
2017-02-15 to 2017-08-29
PIAZZA DI SAN LORENZO 9

Ambulance... 20 num.

Ambulance... 33

Voluntiers 33 num.

Florence main first aids status

Priority\ Hospitals	Red code	Yellow code	Green code	Blue code	White code
PS SANTA MARIA NUOVA	1	7	14	1	0
PS SAN GIOVANNI DI DIO TORREGALLI	2	10	24	2	0
PS AO CAREGGI	10	27	31	10	2
PS SANTA MARIA ANNUNZIATA	2	9	22	4	0

TRENDS QUOTES

#FIRENZE
#NEWS #FIRENZE
#FIRENZE #RIFIUTI
#UFFIZI
#PONTEVECCHIO
#BARGELLO
#AEROPORTO #FIRENZE
#DUOMO #FIRENZE
#VIAGGI #FIRENZE
#CRONACA #FIRENZE

TRENDS QUOTES

@MUSEOFIRENZE
@COMUNEFI
@NAZIONE_FIRENZE
@PROTCIVCOMUNEFI
@UNI_FIRENZE
@HUBFIRENZE
@NOVEDAFIRENZE
@ILSITODIFIRENZE
@FIRENZETURISMO
@BIBLIO_COMUNEFI

Parking availability (%)

28:2

TV-RT Firenze

Twitter Vigilance Real Time Dashboard
Last crawling: 2017-06-21 12:44:30

Menu On Login

Home › Channel statistics › Channel sentiment analysis

Channel active from 2016-06-27 to 2017-06-21 12:35:00
Data processed from 2016-06-28 08:10:00 to 2017-06-21 12:35:00
NLP SA

Sentiment trends in channel Firenze

Zoom 1H 3H 6H 12H 1D 1W 1M Y

Tweets score Tweets score pos Tweets score neg Retweets score Retweets score pos Retweets score neg T+RT score T+RT score pos T+RT score neg

Hide All

Most Significant Tweets for Sentiment

Last tweets per channel Firenze

@GiorgiaLodi - 2017-06-21 12:13:45
RT @emanuelegeri: Primo giorno d'#Estate: pubblicato aggiornamento luoghi freschi https://t.co/isKeZoeGhw su #opendatafirenze @comunefi

TV Firenze

Twitter Vigilance Dashboard
Last crawling: 2017-06-21 12:44:49

Menu On Login

Home › Channel statistics › Statistics on single Channel

Channel active from 2009-02-27 to today
Data processed from 2015-05-22 to 2017-02-24
NLP SA

Search related to channel Firenze

Zoom 1m 3m 6m YTD 1y All From May 21, 2017 To Jun 21, 2017

#Aeroporto #Firenze #Aeroporto #Firenze #unifi #Aeroporto #Firenze #americani #Firenze #bargello #cronaca #firenze #cultura #Firenze #duomo #firenze #expo2015 #Firenze #Expo2015Firenze #firenze #firenze #abusivi #firenze #allafacciadeidivieti #firenze #bottiglia #Firenze #centrostorico #firenze #docorurbano #firenze #rifiuti #firenze #souvenir #firenze #turisti #coda #firenze #turisti #folla #firenze @UNESCO #firenzecard #Florence #smartcity #LiberiamoFirenze #forociuffonna #lungarnofirenze #meteo #neve #Firenze #news #Firenze #notiziafirenze #ODDITI5 #Firenza #opendata firenze #OpenDataDay #Firenze #opandatafirenze #pontevecchio #previsionimeteo #Firenze

Hide All

Daily mumber of tweets/retweets for channel: Firenze

ESB events

Show Menu Hide Menu

Firenze

Leaflet | Map data © OpenStreetMap contributors, CC-BY-SA, Imagery © Mapbox

Firenze Wi-Fi - Access Points Clusters Coverage Map

Firenze Wi-Fi: Access Points Clusters Coverage Map
DISIT - Distributed Systems and Internet Technologies Lab
Firenze - Wednesday June 21 2017 12:49:12

Show Legend

Leaflet | Map data © OpenStreetMap contributors, CC-BY-SA, Imagery © Mapbox

Firenze Wi-Fi - APs Streaming Realtime

Firenze Wi-Fi: APs Streaming Realtime
DISIT - Distributed Systems and Internet Technologies Lab
Firenze - Wednesday June 21 2017 12:49:12

Show Legend

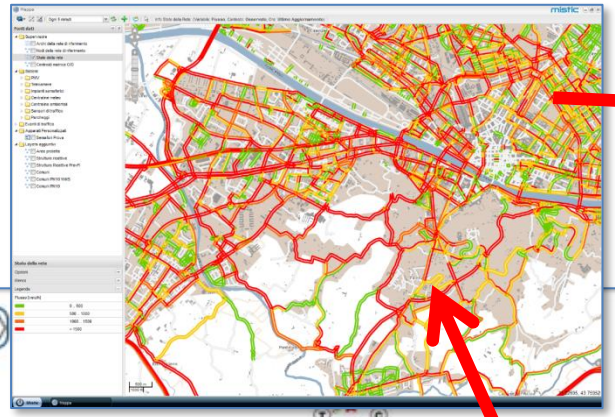
Leaflet | Map data © OpenStreetMap contributors, CC-BY-SA, Imagery © Mapbox

UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB

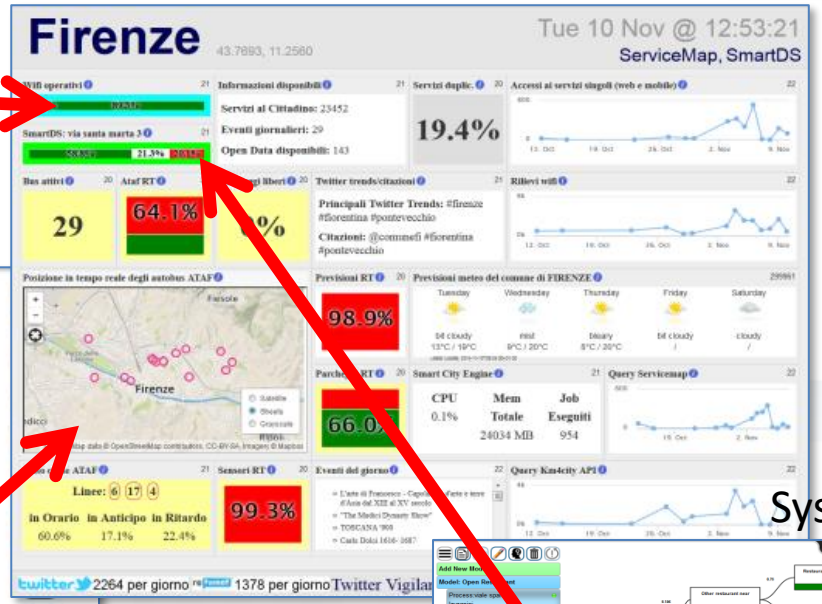
SN

resolute

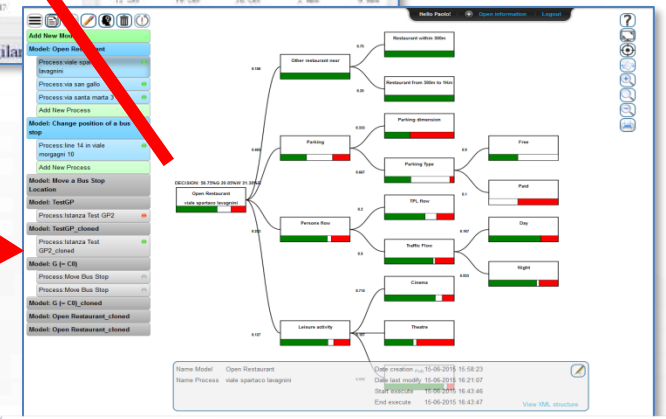
Dashboarding City Resilience

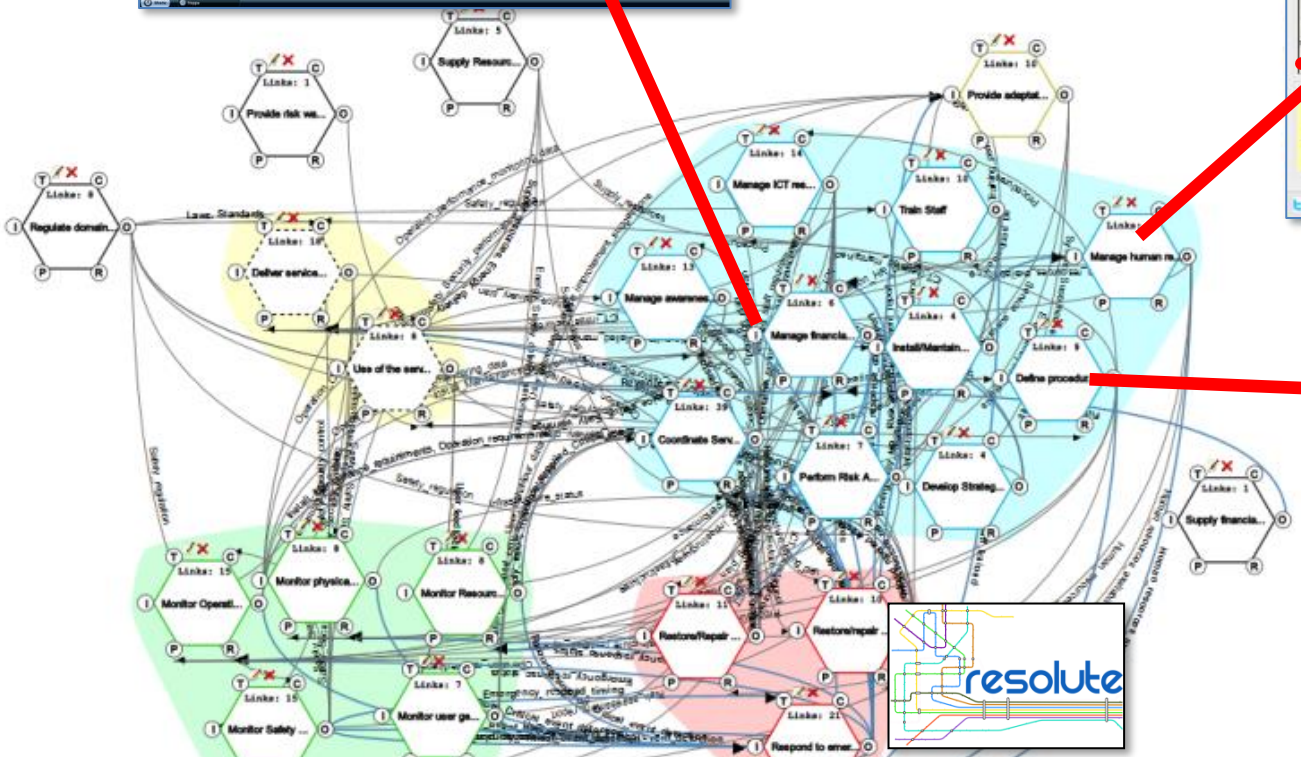


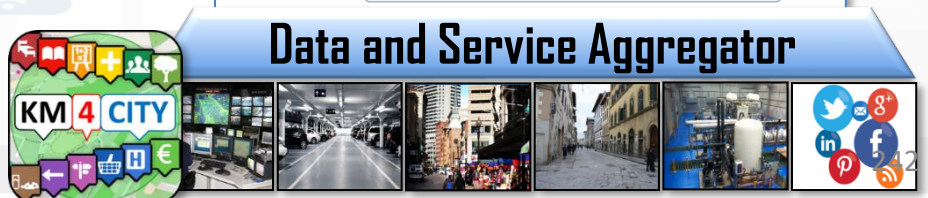
FRAM



System Thinking



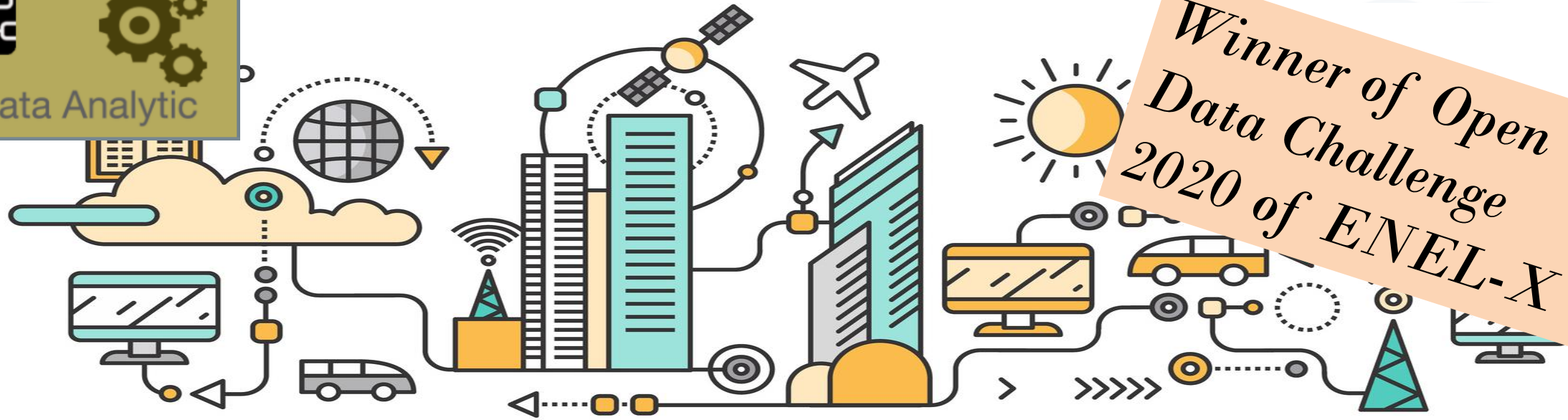




15MinCityIndex

Assessing City 15Min City Compliance

Data Analytic



Winner of Open Data Challenge 2020 of ENEL-X

Concept 15MinIndex

Assessing in each point of the area (city or rural) the capability of providing services ad 15 Min walking distance for the city users

- Several different approaches from early Carlos Moreno concept
- Several different subindexes

Carlos Moreno Functions	Li et al., 2019	15MinCityIndex subindexes
living		Housing viability
	Gov	Govern Services
		Safety Services
		Culture and Cults Services
		Environment Quality
	Roads	Slow Mobility Services
		Fast Mobility Services
	[Medical]	Sport Services
working		Economy/ sustainability
	pension	
commerce	commerce	
	dining	Food Services
healthcare	medical	Health Services
education	edu	Education Services
entertainment	entertainment	Entertainment Services

Technical Challenges

- **The main constraints:**
  - use Open Data
  - maintain the concept of 15 Min distance
- **Problems: conceptual, computational and technical such as**
  - For Who: young, elders, impaired, etc.
    - If not walking: how moving…
    - Effective routing for walking ? walkability
  - For Where, strong differences on: city, rural, industrial area
  - How:
    - Data collection processes and costs
    - Independency of the subindexes
    - Combining measures for computing subindexes
    - Combining subindexes for computing the 15MinCityIndex
    - Computational complexity of **combining data from different sources** and data density, availability, semantics, etc.
    - Definition of the acceptance level according to **objective KPIs**

Lesson Learnt

If a specific data is not accessible another one can be taken as surrogate

The SubIndexes categories

A) environment quality:

- **quality** of the environment in which one is going to live.

B) economy/sustainability:

- assessing **sustainability** by matching the offer vs the demand of work, in first approximation.

C) housing viability:

- assessing **affordability** of having a house of reasonable quality.

D) other: health services, food services, education services, govern services, safety services, culture and cults services, slow mobility services, fast mobility services, entertainment services, sport services.

- assessing the **availability** of services, POI, etc.

Environment SubIndex

- AIM: assess the environmental **quality** of the area for living.
- Good indicators can be:
  - EAQI (European Air Quality Index), or others,
  - specific pollutants (NO2, CO2, O3, PM10, PM2.5, etc.). For most of them, recommended values are provided at the EC level. Please note that, most of the pollutants are also taken into account by the EAQI which is based on PM2.5, PM10, NO2, O2 and SO2, so that, the EAQI supersede them.
  - quality of water,
  - presence of green areas/gardens (which can be recovered from municipality or from Copernicus Satellite data with some limitations and complexity),
  - noise level in most data set related to traffic,

SubIndex (env) = (EAQI + K GreenAreasSM/#Inhabitants)/2

Pollutant	Index level (based on pollutant concentrations in µg/m3)				
	Good	Fair	Moderate	Poor	Very poor
Particles less than 2.5 µm ($PM_{2.5}$)	0-10	10-20	20-25	25-50	50-800
Particles less than 10 µm (PM_{10})	0-20	20-35	35-50	50-100	100-1200
Nitrogen dioxide (NO_2)	0-40	40-100	100-200	200-400	400-1000
Ozone (O_3)	0-80	80-120	120-180	180-240	240-600
Sulphur dioxide (SO_2)	0-100	100-200	200-350	350-500	500-1250

Econony SubIndex

- AIM: assess the economic **sustainability** of the area,
  - demand-vs-offer index,
  - self-sustainability: working places (commercial and industrial) wrt workforce in the area is satisfactory to ensure the economic well-being of the people.
- Good Indicator can be:
  - census data: people living in the area: ages, skill, etc.
  - commercial / economic data
    - POI as shops, services, industries, etc. as POI
    - capability in terms of personnel per POI
    - level of salary of the area (including retired people) for the sustainability of the whole population of the area.

Housing SubIndex

- AIM: assess the **affordability** of houses in the area.
- Indicators can be:
  - services of the buildings, hard to get
  - age of the buildings, possible
  - prices for square meters of the real estate market, which is also related to the affordability of living.
- Consideration:
  - An affordable price with respect to the population living there could simplify the live-ability.

Health SubIndex

- AIM: assessing the **availability** of health services
- Indicators:
  - Presence of: pharmacies, private and public hospitals, labs, studio of doctors, presence of defibrillators, etc. with respect to the number of users.
- Consideration
  - There are some *KPI* provided by the EC.
  - coverage of the pharmacies in terms of potential clients of the area.
    - a pharmacy every 3000 people for Italian govern,
    - 500 hospitals' beds every 100.000 inhabitants,
    - 80 dentists per 100.000 inhabitants,
    - 1.6 hospitals per 100.000 inhabitants, etc.

Slow Mobility SubIdx

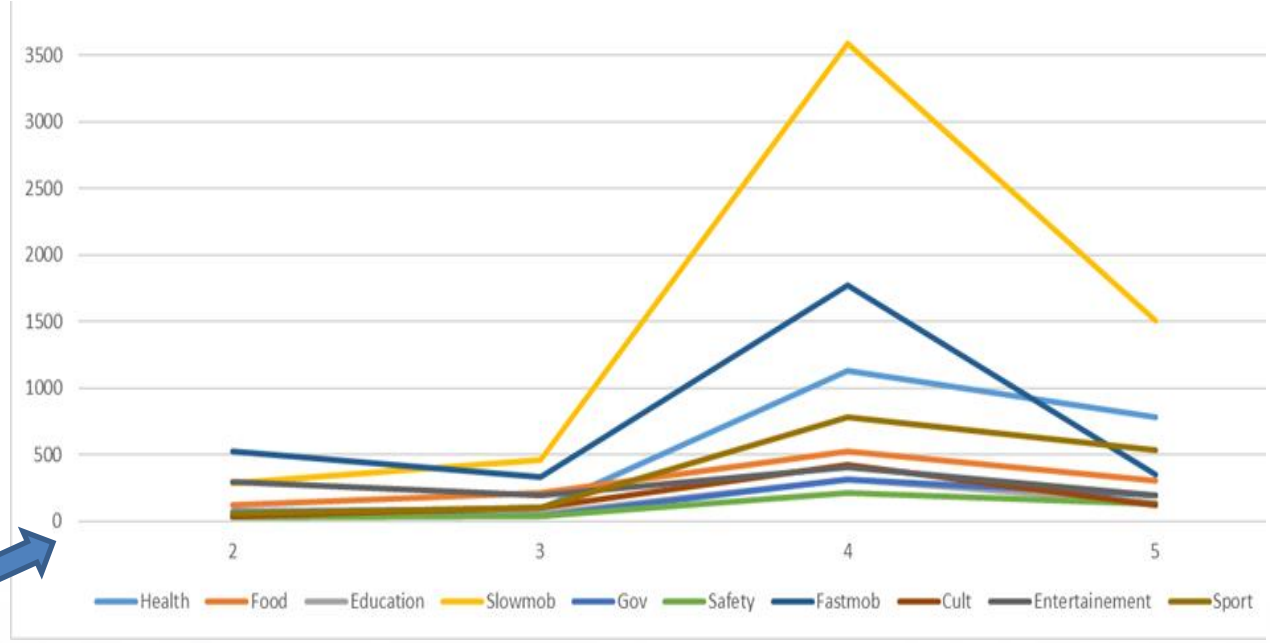
- AIM: assess **presence** of services to facilitate the local mobility.
- Indicators:
  - walkability, presence of sidewalks, number of Km of roads (that can be used for walking and biking), presence and number of cycling paths, availability of bike/cart sharing, number of sharing station for short range travel,
  - critical presence of barriers such as bridges, highways, etc., which can limit the walkability. Eventual usage of routing but computationally too complex.
- Consideration:
  - good services can facilitate the movements in the area and reduce the needs of taking less sustainable travel means.
  - having a good rank on most of the subindexes (taking into account of services in the area) it would reduce the needs of exiting from the local area by using a private or public travel means to access at the not accessible services.

Fast Mobility SubIdx

- AIM: assess the **presence** of services for moving out of the area.
- Indicators:
  - Presence of city busses, area busses and long distances (railways, airports, highway joint), with respect to the number of users of the authorized age to use them.
  - presence of fuel stations,
  - size of bus stops and railways stations,
  - long range recharging stations, etc.,
- Consideration:
  - Open Data, and POI.
  - Public transportation data (bus stops, bus lines, train stops, etc.) can be recovered from GTFS and/or Transmodel files which are accessible from the operators.

The process

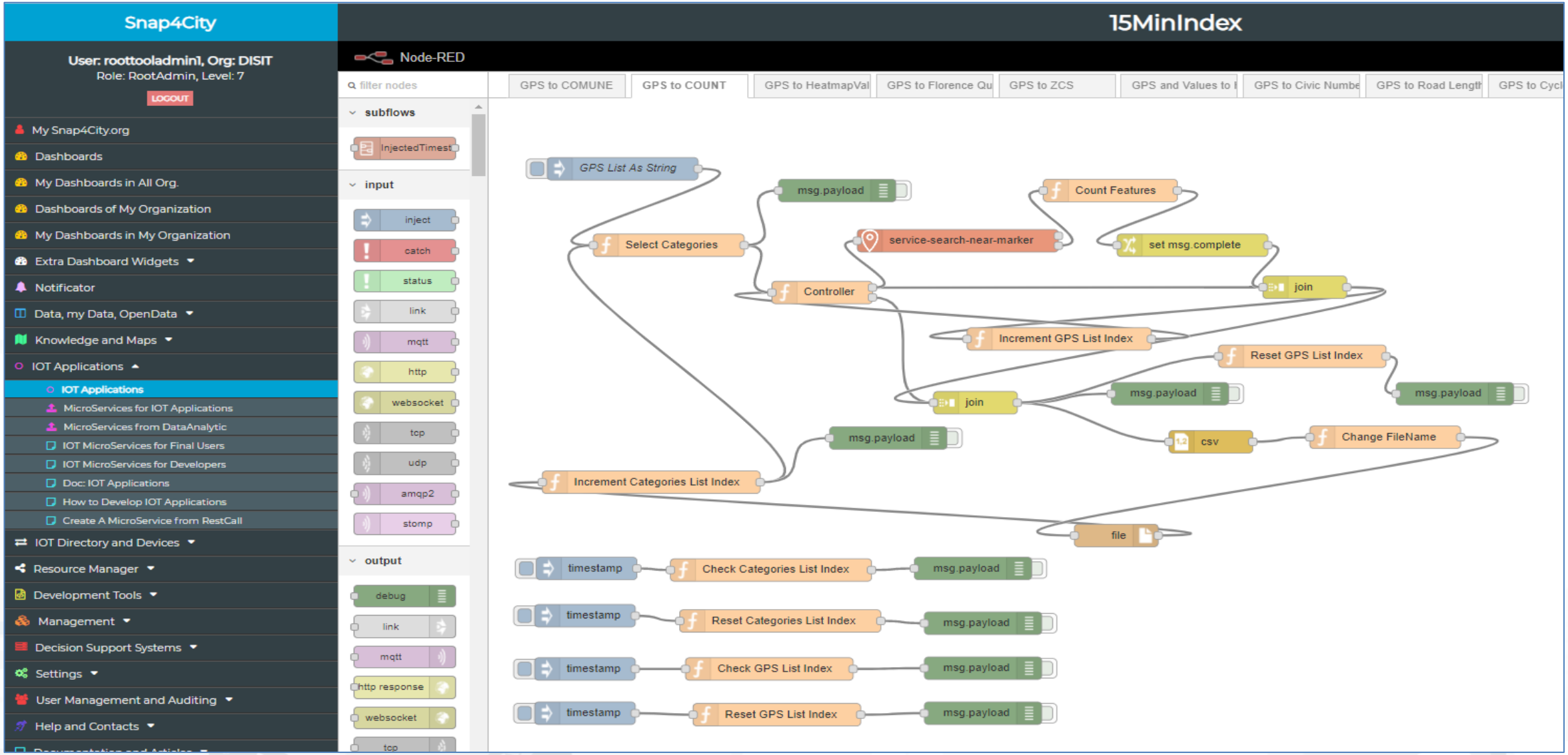
- Creating a grid 700x700 mt
  - Computing number of inhabitants for grid area
- For each Subindex
  - Collecting: each sub-subindex data by using IoT App and saving them on Snap4City
  - Mapping data on the grid Analysing distribution,
    - scaling and centering the distribution according to KPI
  - Saving the Subindex on the HeatMapServer
- Computing the 15MinCityIndex
  - Saving the 15MinCityIndex on the HeatMapServer



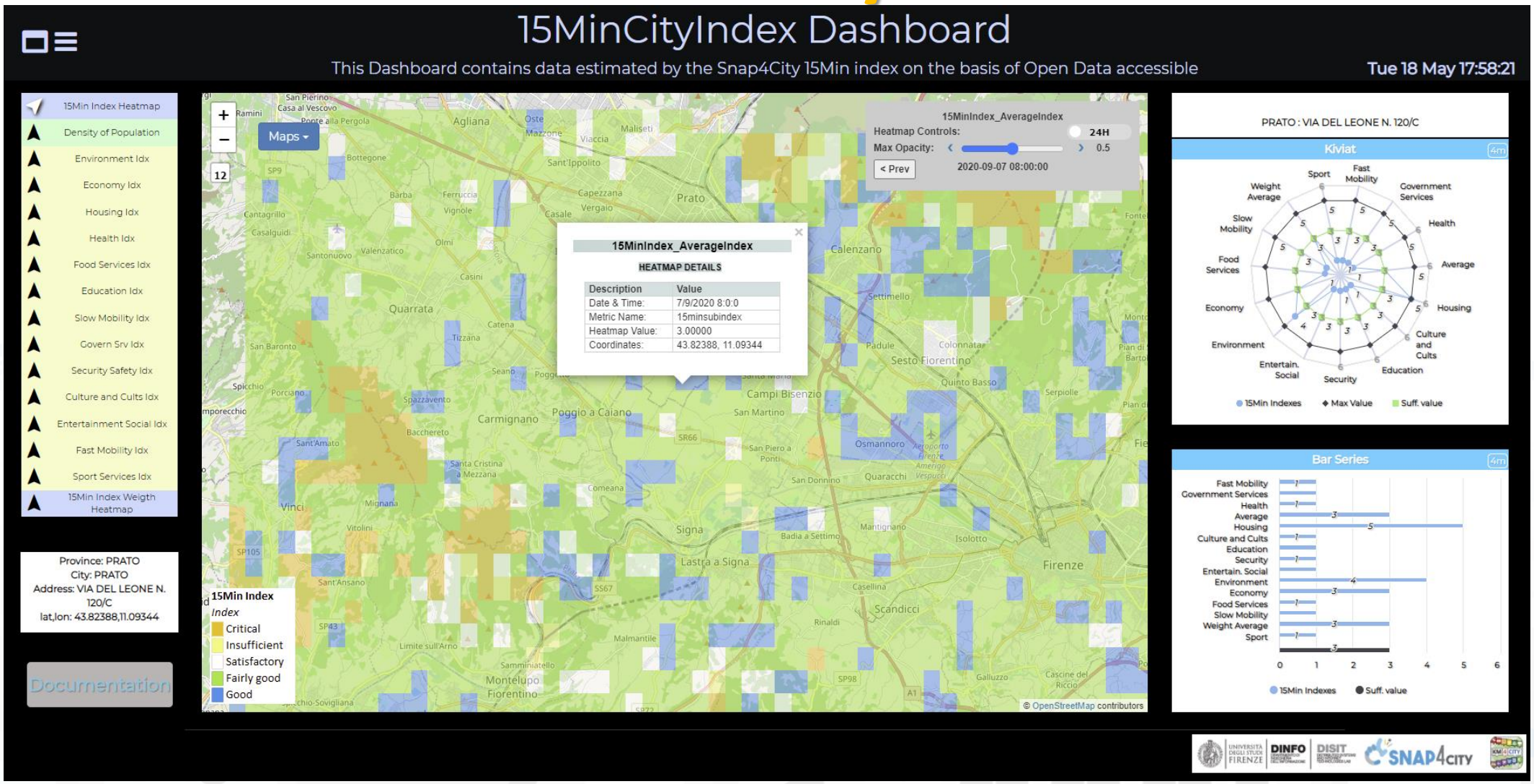
from	to	class	
2000	infinite	orange	1
1500	2000	yellow	2
1000	1500	white	3
20	1000	green	4
0,01	20	blue	5

Index
Critical
Insufficient
Satisfactory
Fairly good
Good

IoT App....


Snap4City
User: roottooladmin1, Org: DISIT
Role: RootAdmin, Level: 7
LOGOUT
My Snap4City.org
Dashboards
My Dashboards in All Org.
Dashboards of My Organization
My Dashboards in My Organization
Extra Dashboard Widgets
Notificator
Data, my Data, OpenData
Knowledge and Maps
IOT Applications
IOT Applications
MicroServices for IOT Applications
MicroServices from DataAnalytic
IOT MicroServices for Final Users
IOT MicroServices for Developers
Doc: IOT Applications
How to Develop IOT Applications
Create A MicroService from RestCall
IOT Directory and Devices
Resource Manager
Development Tools
Management
Decision Support Systems
Settings
User Management and Auditing
Help and Contacts
15MinIndex
Node-RED
filter nodes
subflows
InjectedTimest
input
inject
catch
status
link
mqtt
http
websocket
tcp
udp
amqp2
stomp
output
debug
link
mqtt
http response
websocket
tcp
GPS to COMUNE
GPS to COUNT
GPS to HeatmapVal
GPS to Florence Qu
GPS to ZCS
GPS and Values to
GPS to Civic Numbe
GPS to Road Length
GPS to Cycl
GPS List As String
msg.payload
Count Features
Select Categories
service-search-near-marker
set msg.complete
Controller
join
Increment GPS List Index
Reset GPS List Index
join
msg.payload
msg.payload
msg.payload
csv
Change FileName
Increment Categories List Index
file
timestamp
Check Categories List Index
msg.payload
timestamp
Reset Categories List Index
msg.payload
timestamp
Check GPS List Index
msg.payload
timestamp
Reset GPS List Index
msg.payload

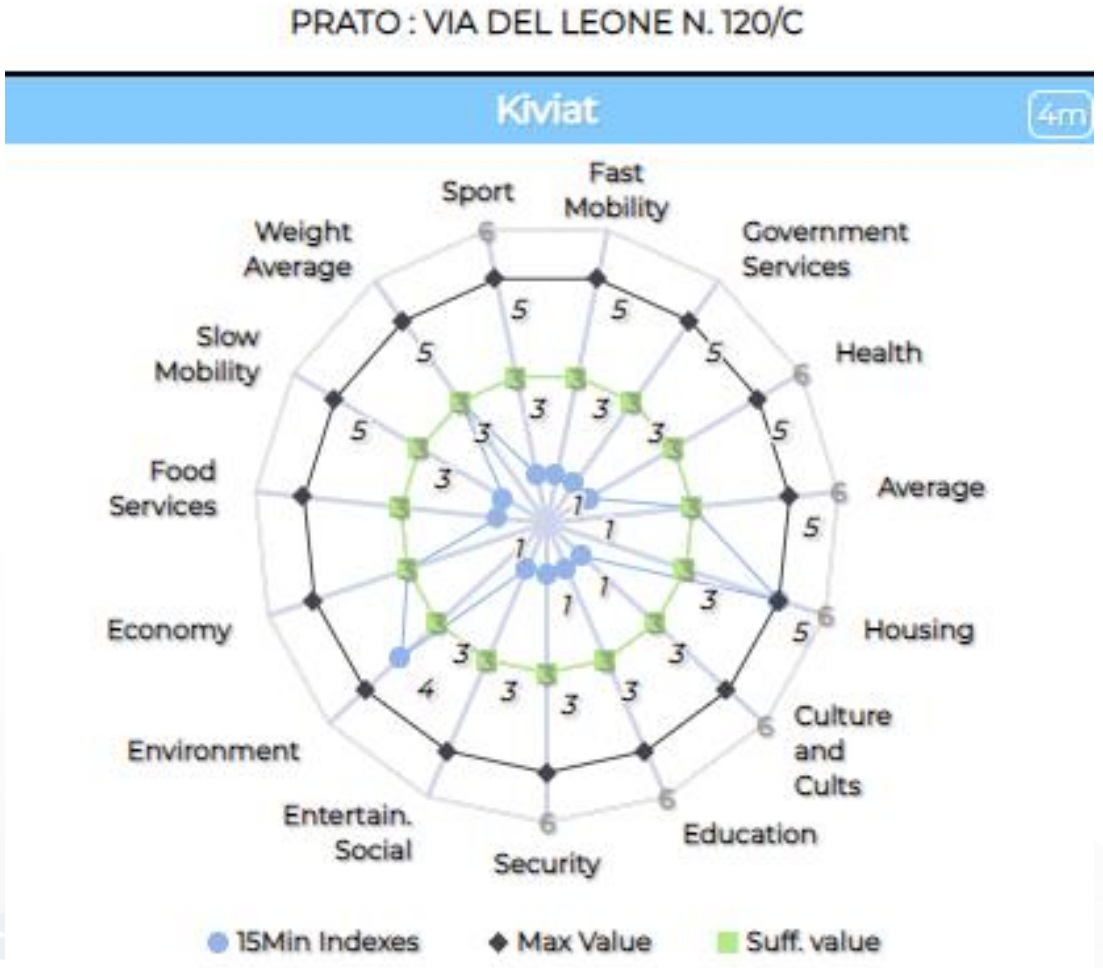
15MinCityIndex



- https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MjkzOA==

15MinCityIndex

- Derived and extended from Carlos Moreno 15 Minute concept
  - But with 13 different subindexes
- Computable on the basis of Open Data
- Validated on major city area
- Extendable to whole national level or large regions
- Accessible as Dashboard to play with
- ***Winner of the ENEL-X Open Data Challenge 2020.***

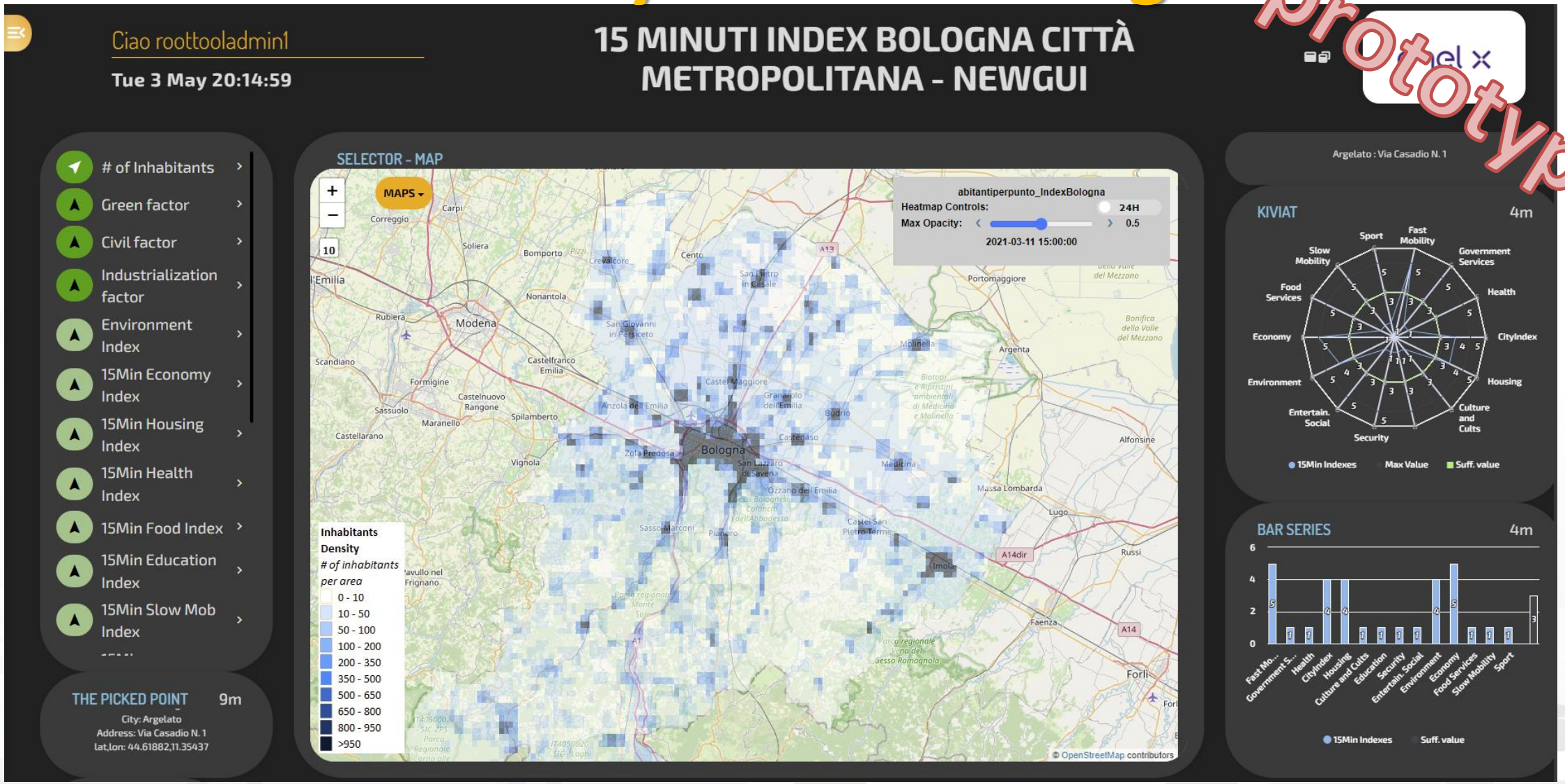


Conclusions

- **The 15MinCityIndex** has been
  - defined in terms of its subindexes, only using Open Data
  - developed taking into account the needs of local services as in the main Lockdown period in march 2020
  - validated in a large are of 800.000 inhabitants on Florence area
- **The Solution** in the presented form has been
  - the winner of the International Open Data Challenge 2020 of ENEL-X
  - developed exploiting data ingestion and computing capabilities of Snap4City infrastructure and tool, making it affordable.
  - Also using the IoT App for data ingestion and business intelligence on dashboard making them interactive

15MinCityIndex on Bologna

prototype



https://www.snap4city.org/dashboardSmartCity/view/Baloon-Dark.php?iddasboard=MzQxMg==

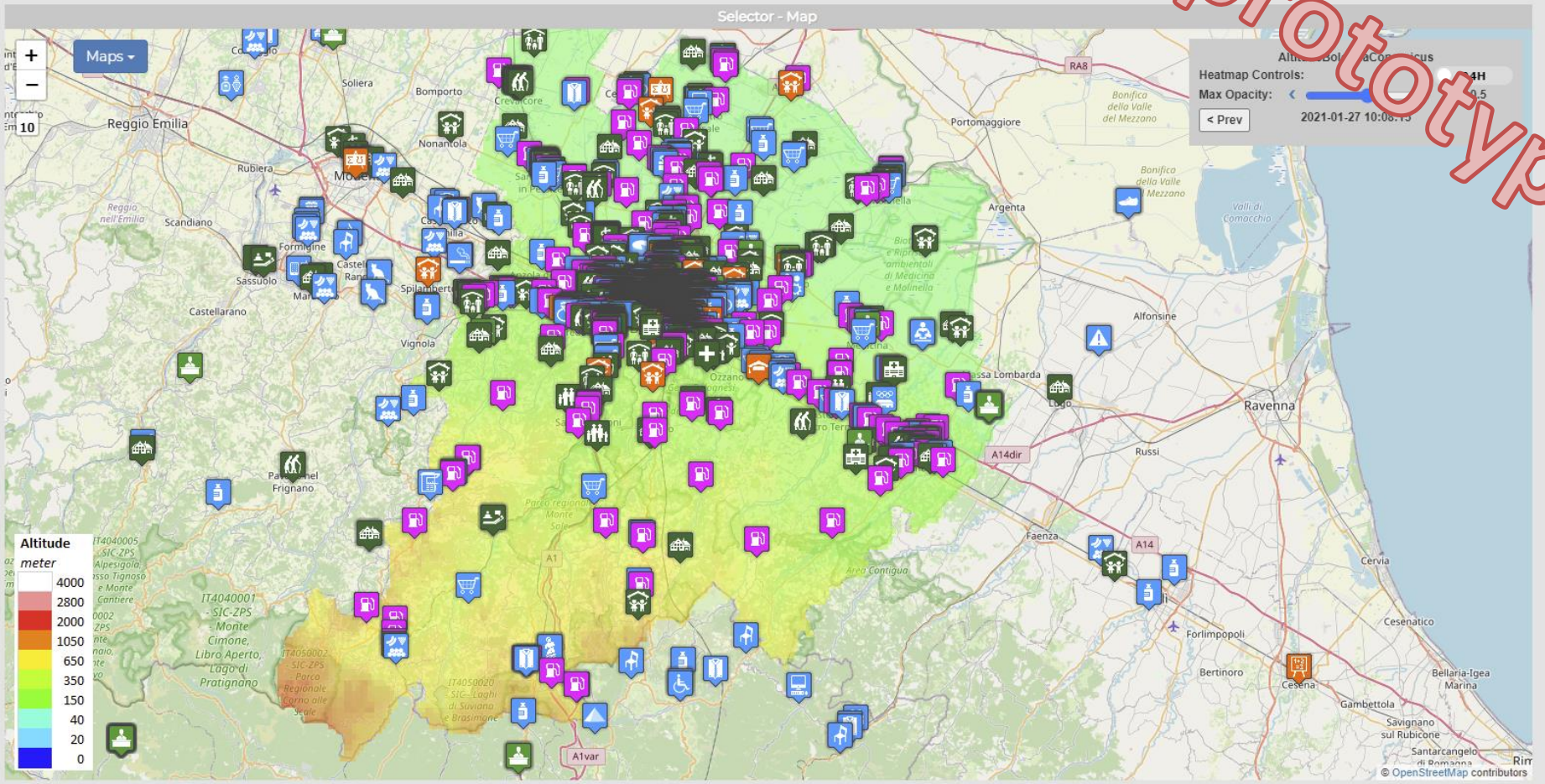
Bologna Metropolitan Area Dashboard

enel x

Sun 19 Sep 11:59:28

prototype

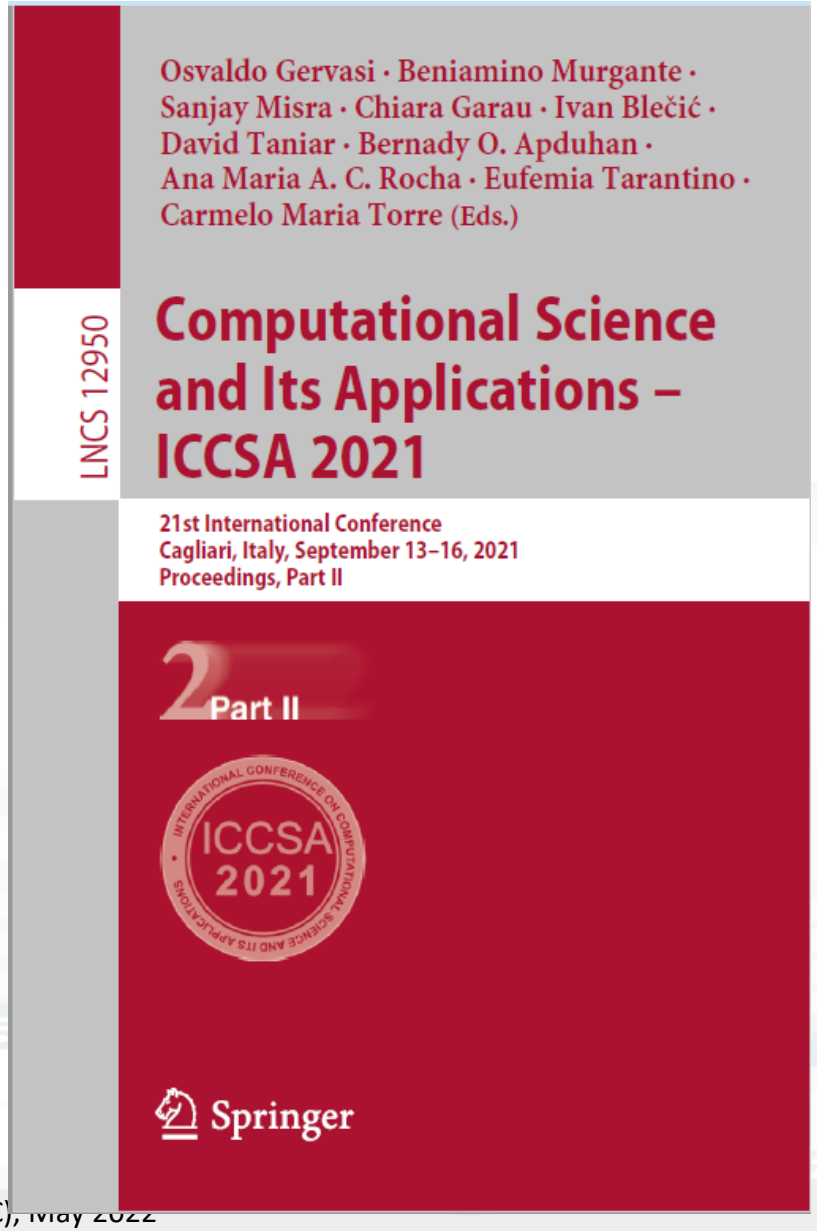
- Train station
- Charging Stations
- Bus Stops
- Fuel stations
- Cultural Activities
- Education
- Entertainment
- Goverment
- Healthcare
- Shopping
- Bike Racks
- Wine and Food
- Emergency Services
- Air Quality Stations
- Air Temperature Heatmap
- Humidity Heatmap
- Global Vegetation Index Heatmap
- Altitude Heatmap
- Fractional Cloud Cover Heatmap
- SciHub CO
- SciHub NO2
- SciHub O3
- SciHub SO2
- # of Inhabitants
- Green factor
- Civil factor
- Industrialization factor
- Environment Index
- 15Min Economy Index
- 15Min Housing Index
- 15Min Health Index
- 15Min Food Index
- 15Min Education Index
- 15Min Slow Mob Index
- 15Min Government Index
- 15Min Safety Index
- 15Min Culture and Cults Index
- 15Min Entertainment Index
- 15Min Fast Mobility Index
- 15Min Sport Index
- CityIndex MP1



Privacy Policy | Cookies Policy | Terms and Conditions | Contact us

https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MzA1OA==

- Bellini, Pierfrancesco, et al. "Long Term Predictions of NO 2 Average Values via Deep Learning." *International Conference on Computational Science and Its Applications*. Springer, Cham, 2021
- https://link.springer.com/chapter/10.1007/978-3-030-87010-2_44



Predictive Maintenance

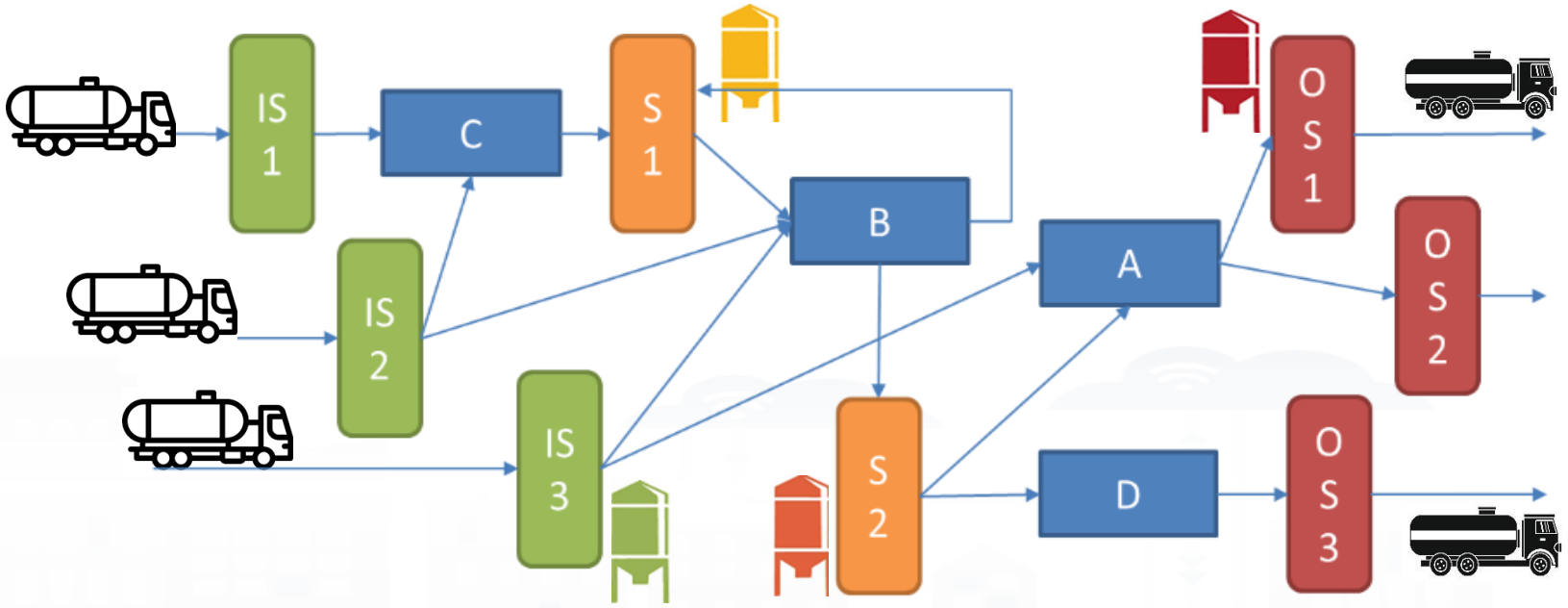
Industry 4.0

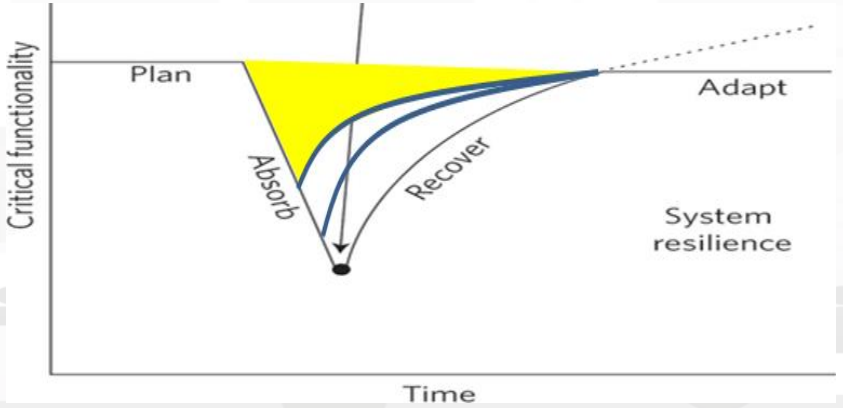
- **ALTAIR SODA-4.0 project**
  - maximize the efficiency and productivity of plants, reducing downtime
  - in order to improve competitiveness in the market

- **Goals and drivers:**
  - Business intelligence tools on maintenance data
  - predictive maintenance approach into the whole control and management systems Predictive models for engagement
  - predict plant failures 60 minutes before it happens
  - Provide indications on the area of failure via XAI

Complex cause-effect realtionships

- **Elements:**
  - Machines: A…C
  - Storage: silos…
  - Flows:…
- **Dependencies**
  - Cascade effects
- **Early warning**
  - Reduction of costs
  - Recovering from failure is more expensive than correcting in advance
  - Possible advanced replan and reschedule: secondary solutions





Workflow

Data acquisition
Production - Maintenance

Feature selection,
Feature engineering

CNN-LSTM Training

Validation AI
Model

Classification
Model

CNN-LSTM
Engine

F/NF

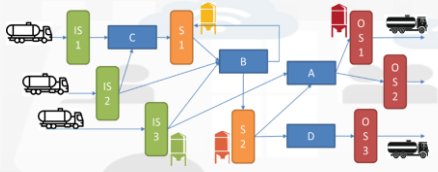
Explainable/XAI
(Shap)

{features (t)}

Business Intelligence for
Maintenance

integration of workflow management system for maintenance with general control systems and data flow

Structure of the plant



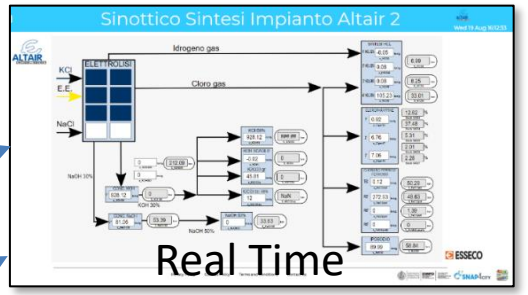
Decision making
on more cases

KPI

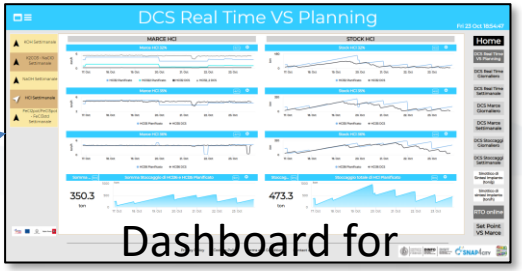
CNN: Convolutionary NN
LSTM: Long Short-Term Memory
(deep learning model)

Solution

Plant Status

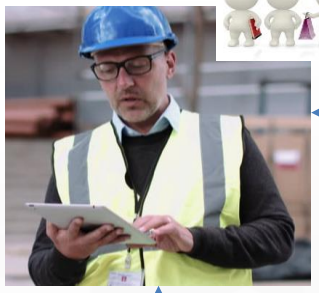


Real Time Production Synoptic



Dashboard for Production Control

Control Supervisor

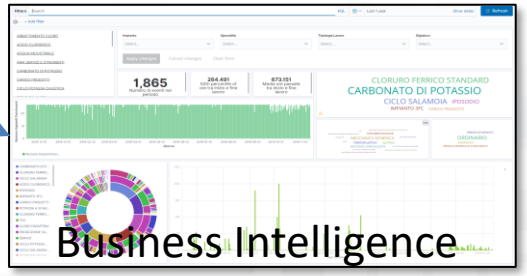


Business Logic

IoT App

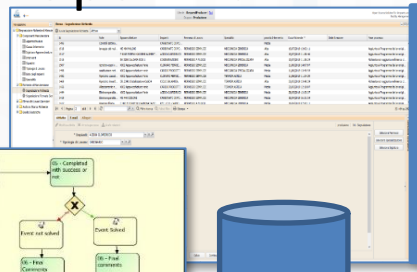
Data Storage

Elastic Search

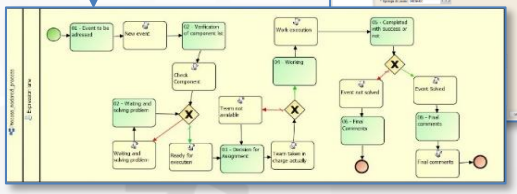


Business Intelligence Maintenance

OpenMaint



API



IoT App

IoT App

Business Logic 2

Predictive Maintenance

Predictive Training

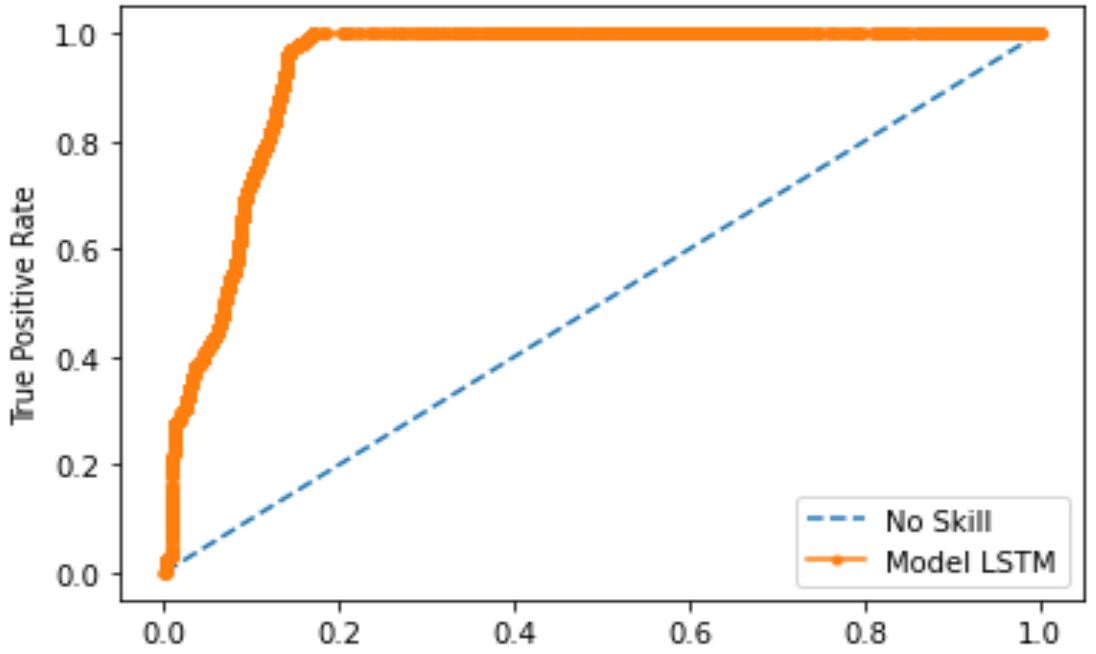
Overview Features

Feature	Plant	Description	Unit of measure
TempreactoreR4001 - TempreactoreR4002 - TempreactorR4003	chlorine paraffins (CPS)	reactor temperature indication	°C
S904A - S904B - S904C	Potable Ferric std	Storage level indication	%
S4304	chlorine paraffins (CPS)	Storage level indication	%
standardFerric Chloride	Potable Ferric std	flow rate measurement and totalization	m3
potFerricChloride	Potable Ferric Chloride	flow rate measurement and totalization	m3
S904E - S904D	Potable Ferric Chloride	Storage level indication	%
QuantNaOHperBatchNaClO - QuantNaOHBatchNaClO_2	NaOH KOH	flow rate measure and totalization	lt – m3
ConversionNaOH - ConversionKOHlinea1	NaOH KOH	electrolysis load adjustment (production)	kA
KOH_1_charge - KOH_2_charge	NaOH KOH	flow rate measure and totalization	m3
S487 - S484 - S5104	NaOH KOH	Storage level indication	%
hypo sodium	sodium hypochlorite	quantity of material produced	m3
S851 - S852 - S854 - S856 - S857	sodium hypochlorite	Storage level indication	%
S871	HCl	Storage level indication	%
RedoxFeCl3Pot	Ferric Chloride std	potential measure redox Ferric Chloride	mV

Preditive capabilities

- Deep Learning: LSTM, CNN-LSTM approached
- Explainable AI: Identification of possible causes of fault

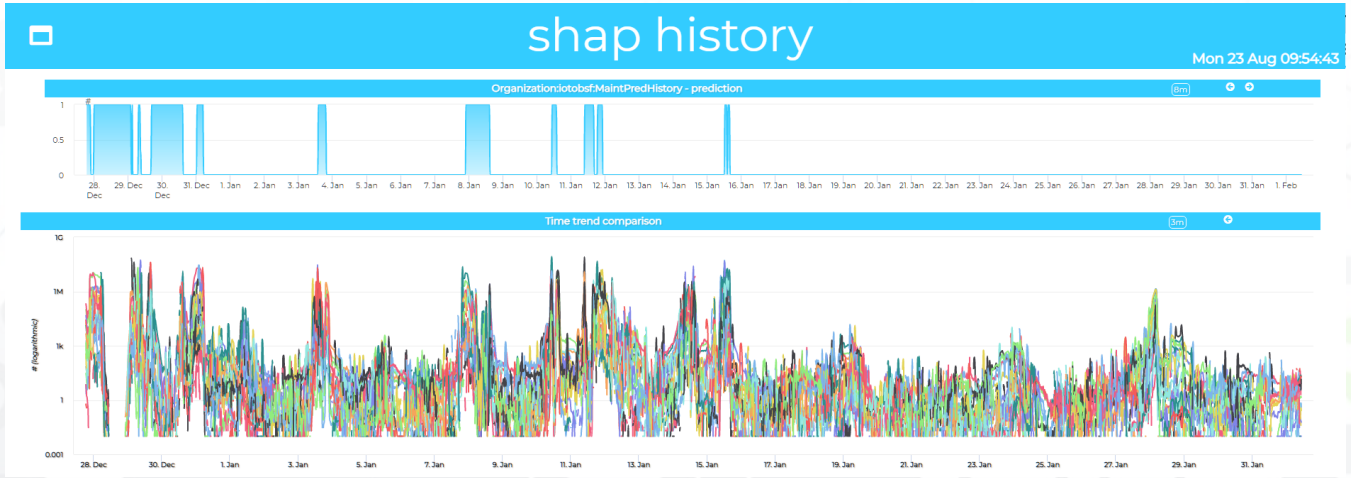
	Precision %	Recall %	F_1 score %
weighted avg	0.90	0.92	0.90



Explainable/XAI - CNN-LSTM (SHAP)

Explanation of prediction generated by model for fault

Explanation of prediction generated by model for normality

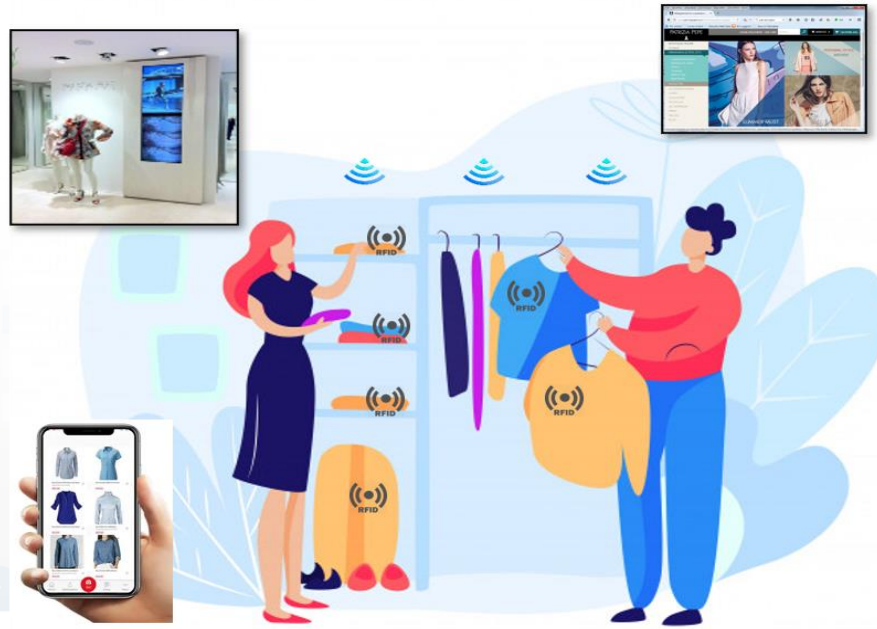


Fashion Retail Recommendations

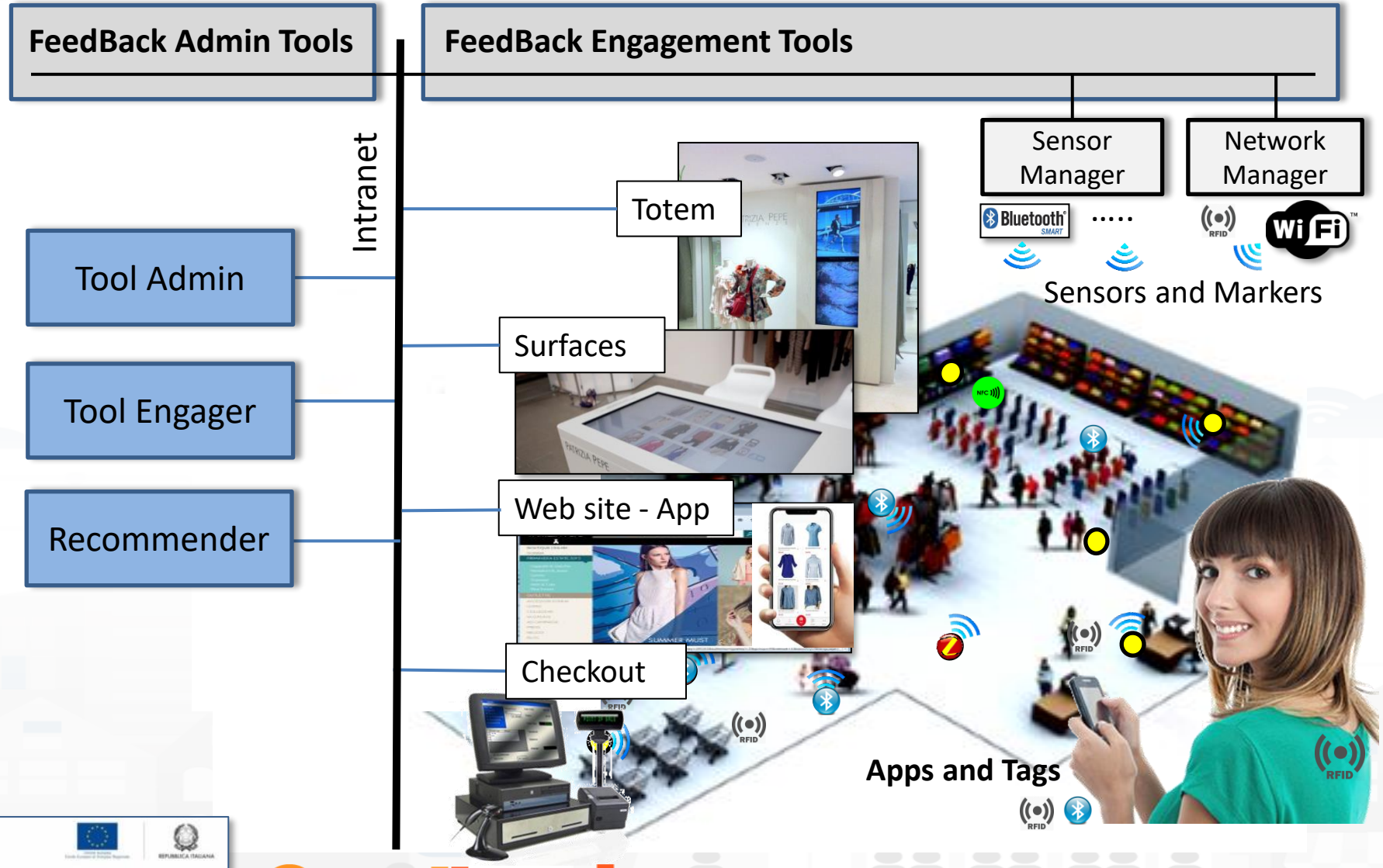
Feedback Pilot

Smart Retail

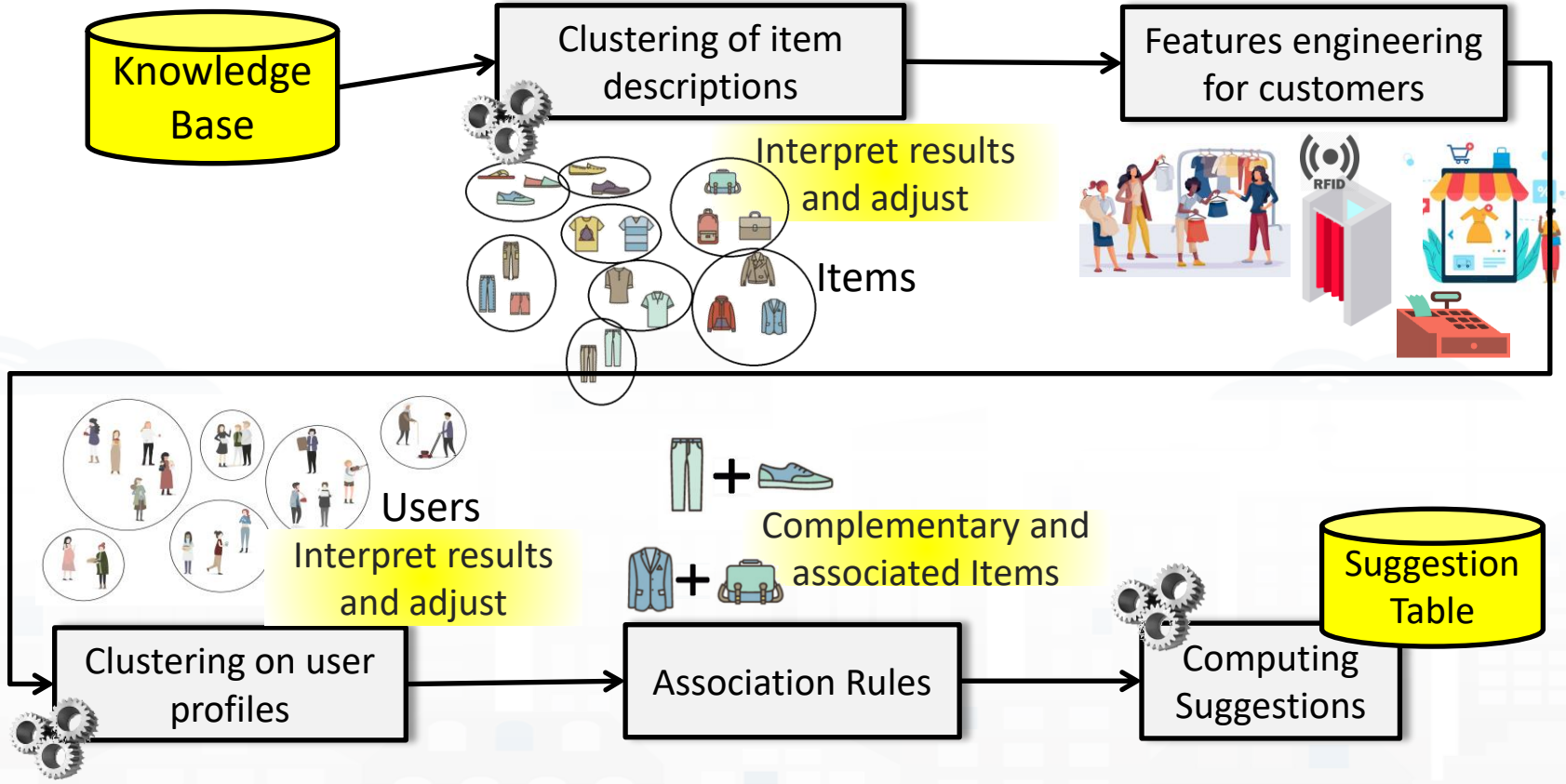
- **Feedback Project:**
  - Flexible Advanced Engagement Exploiting User Profiles and Product/Production Knowledge
  - VAR, PatriziaPepe (Tessilform), DISIT, Effective Knowledge, SICE
  - Keywords: retail, GDO, …
- **Goals and drivers:**
  - adaptive user engagement, customer experience
  - Advanced user profiling, user behaviour analysis
  - IOT and instrumentation
  - Predictive models for engagement
  - Integrated in city customer experience
- **Aiming to solve current State of the Art issues:**
  - Cold start problems in generating recommendations for new users, also addressing seasonality of products and items
  - GDPR compliance



Scenario


FeedBack Admin Tools
FeedBack Engagement Tools
Intranet
Tool Admin
Tool Engager
Recommender
Totem
Surfaces
Web site - App
Checkout
Sensor Manager
Network Manager
Bluetooth SMART
.....
RFID
Wi Fi
Sensors and Markers
Apps and Tags

Workflow


Knowledge Base
Clustering of item descriptions
Features engineering for customers
Interpret results and adjust
Items
RFID
Users
Interpret results and adjust
Complementary and associated Items
Clustering on user profiles
Association Rules
Computing Suggestions
Suggestion Table

Suggestions

customer similarity for each customer cluster the most representative items are suggested;
item similarity: considering the last items purchased by the customer according to the information contained into its profile, and randomly selecting items in the same item clusters;
item complementary: considering items that may complement the last items that have been bought by the customer according to a table of complementary items;
item associated: in order to improve a customer's purchase frequency, we generated suggestions for customers who purchased an item in the last three months;
suggestions for serendipity: randomly selecting items to be suggested from the whole present collection, taking also into account what is available in the physical shop;

Item Cluster	Complementary Item Clusters				
	cluster	support	confidence	lift	count
1	2	0.26486066	0.6069351	1.106003	12935
	7	0.24864345	0.5697729	1.253423	12143
	3	0.24465057	0.5606231	1.213722	11948
	8	0.24336057	0.5576670	1.277549	11885
	4	0.22298667	0.5109797	1.282096	10890
2	3	0.34351004	0.6259701	1.355196	16776
	7	0.32391425	0.5902612	1.298495	15819
	8	0.31392182	0.5720522	1.310504	15331
	4	0.29840080	0.5437687	1.364367	214573
3	2	0.34351004	0.7436830	1.355196	16776
	7	0.30397035	0.6580814	1.447690	14845
	8	0.29868747	0.6466442	1.481385	14587
	4	0.27753548	0.6008511	1.507592	13554
	1	0.24465057	0.5296569	1.213722	11948
4	2	0.29840080	0.7487156	1.364367	214573
	3	0.27753548	0.6963625	1.507592	13554
	7	0.26578209	0.6668722	1.467029	12980
	8	0.27260069	0.6839807	1.566918	13313
	1	0.22298667	0.5594945	1.282096	10890

Item selection
1. Item previously not purchased
2. Confidence recommended item. Confidence established with Market Basket Analysis

Conclusions

- Using the stimulus of the recommendation system, we have increased the customers' attention of the 3.48%
- The solution is also functional in presence of a low number of customers and items
- The solution solved the cold start problems
- GDPR compliant

TOP

Data Analytic: Enforcing and Exploiting

Data Analytics on Snap4City platform

Swagger

Ontology Schema

LOG.disit.org

Knowledge Base, Km4City

Big Data Store Facility

Smart City API from Knowledge Base and other tools

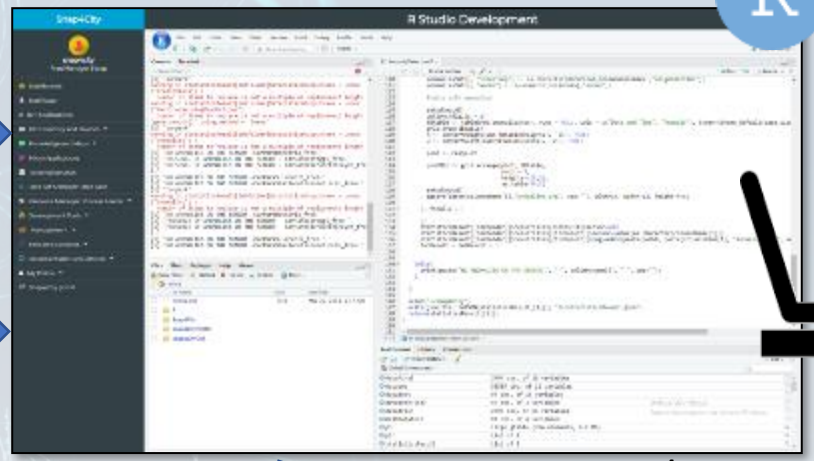
TensorFlow

NVIDIA CUDA

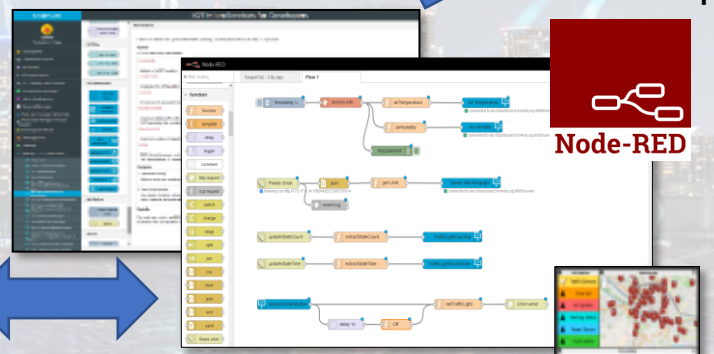
python

jupyter

R Studio



Creating MicroServices



Node-RED

Using them into IOT Applications

Saving / Sharing reusing

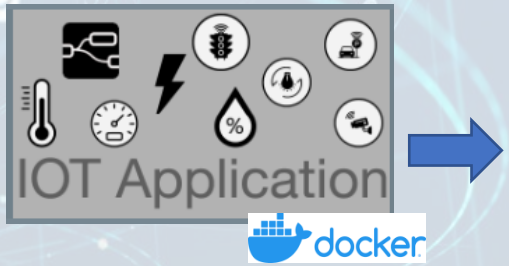
Resource Manager

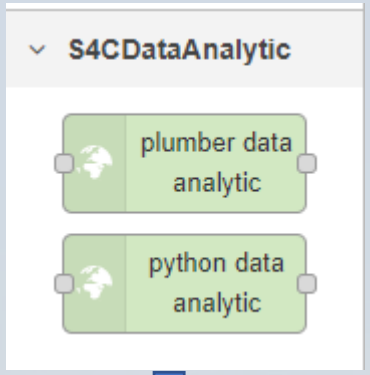
Data Analytic Container

1

Develop .py **or** .r program on (i) Snap4City platform online, or (ii) your Development Machine.
The code has to respect the guidelines provided. For example see:
https://www.snap4city.org/641
https://www.snap4city.org/645

2 Open an advanced IoT App / Node-RED

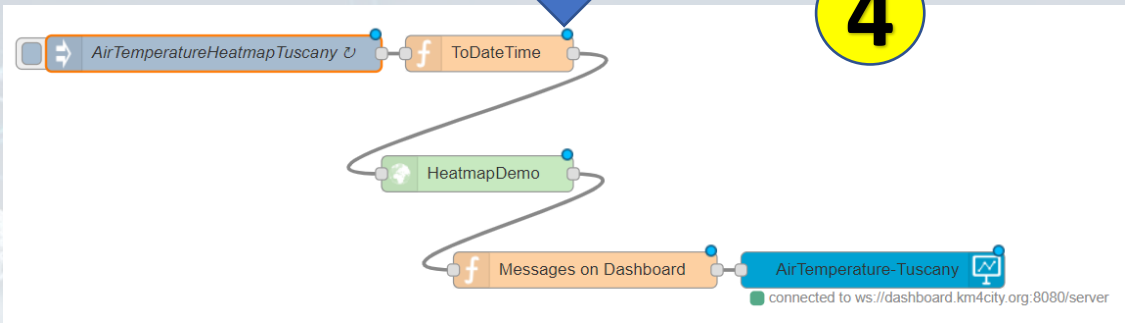




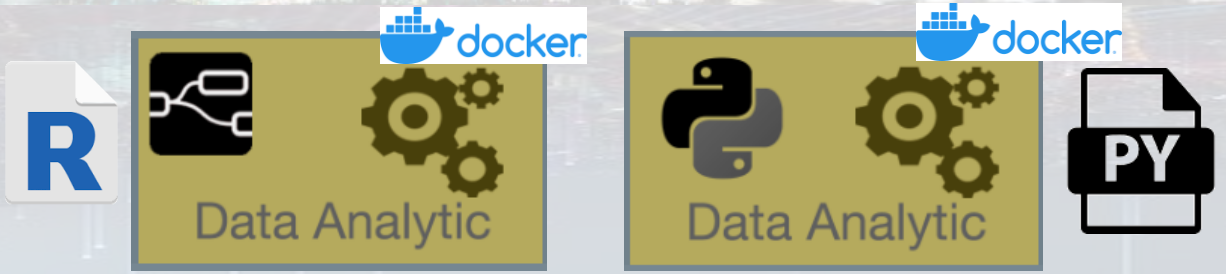
3

Use Snap4City Data Analytic Node, and load in the code you developed

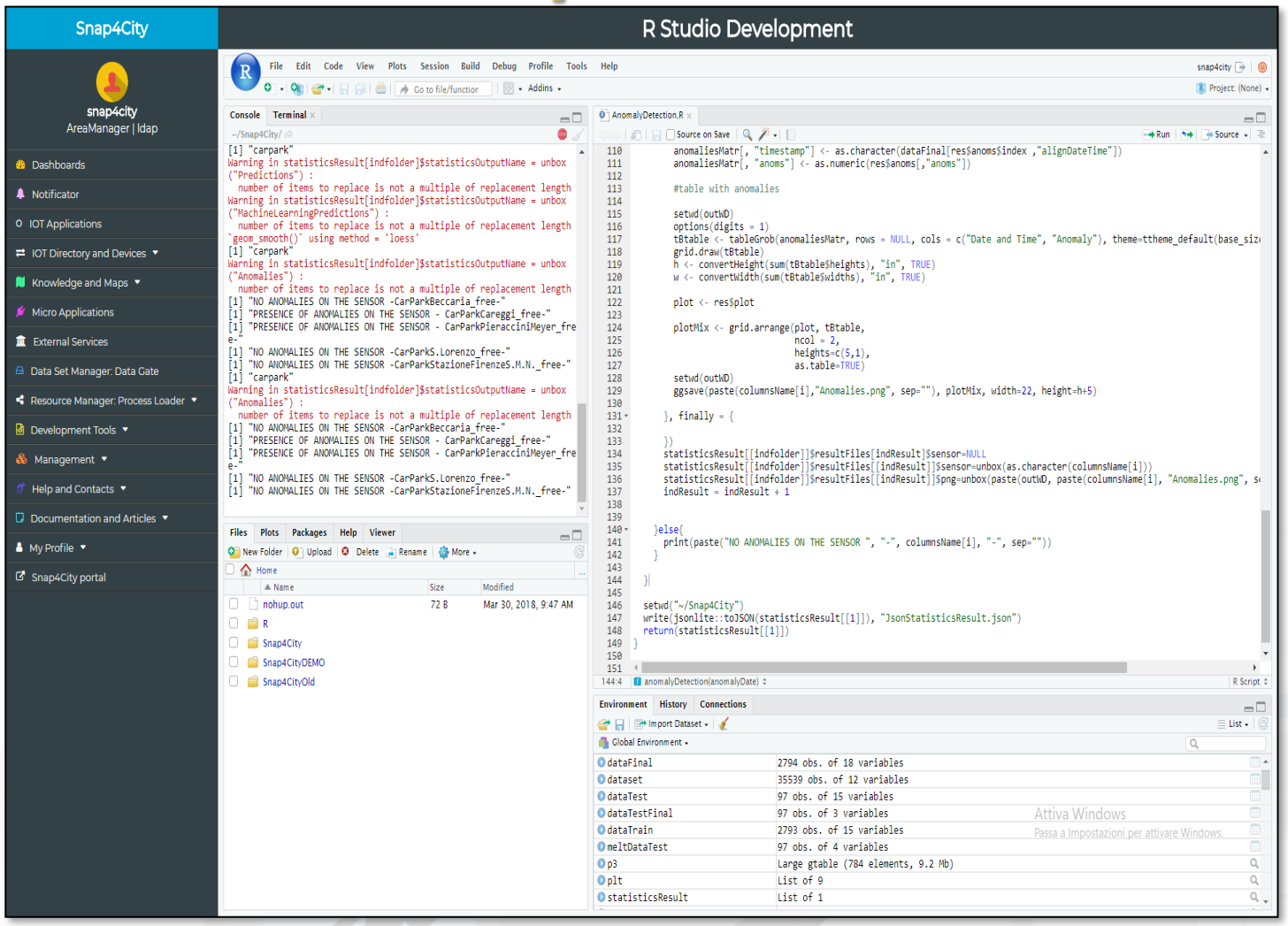
4

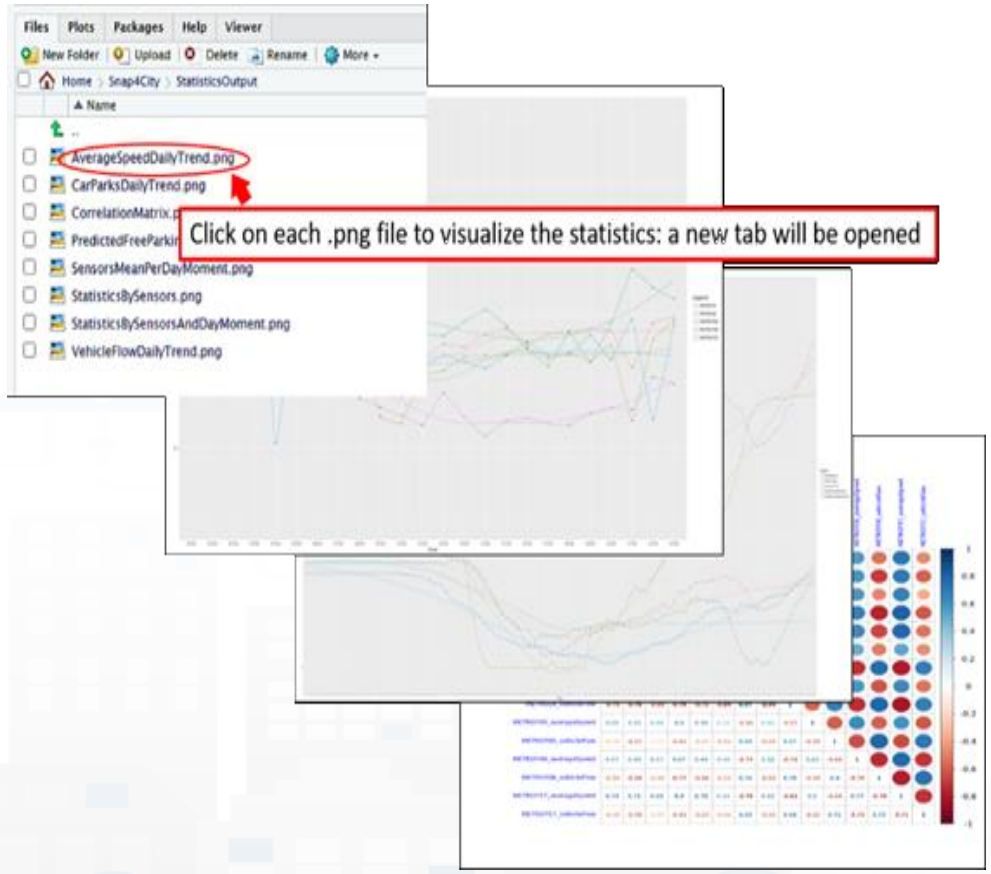


5 Deploy the IoT App → Snap4City Container Manager based on Marathon/Mesos is creating a Container for your Data Analytic code



1 Developer in R Studio + Tensor Flow

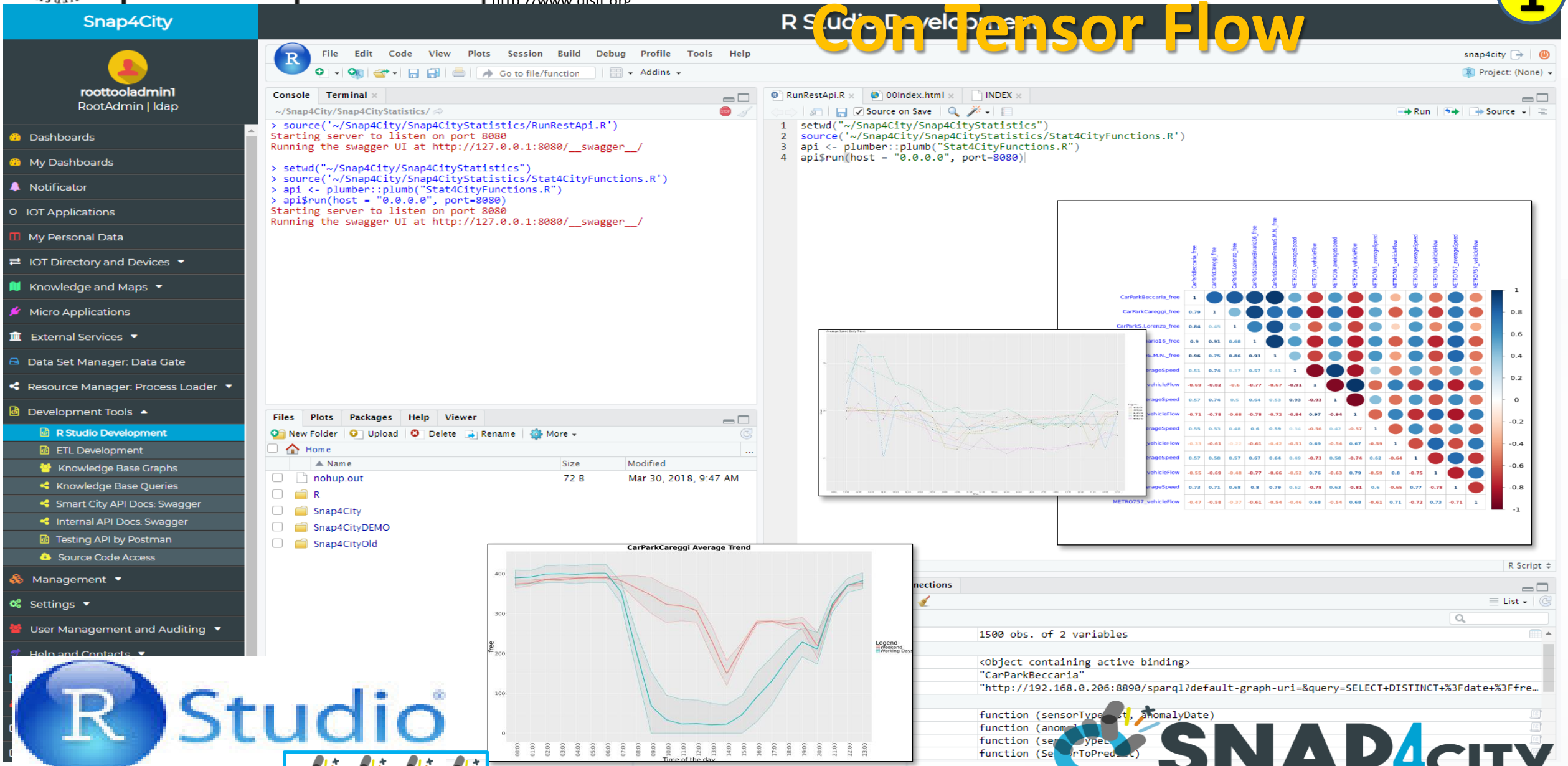




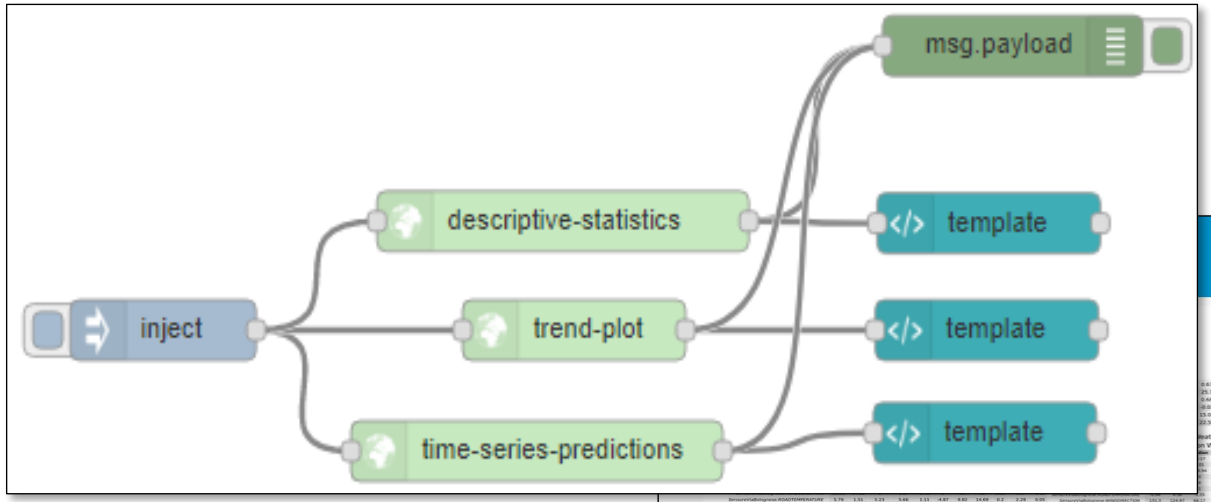
UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB http://www.disit.org

Data Analytics in R Studio Con Tensor Flow

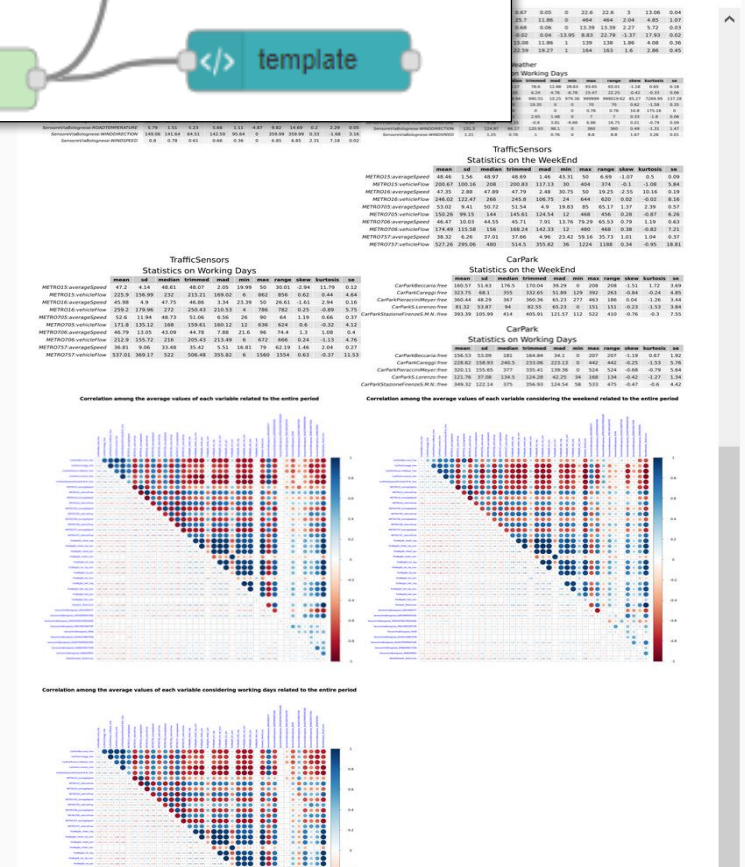
1

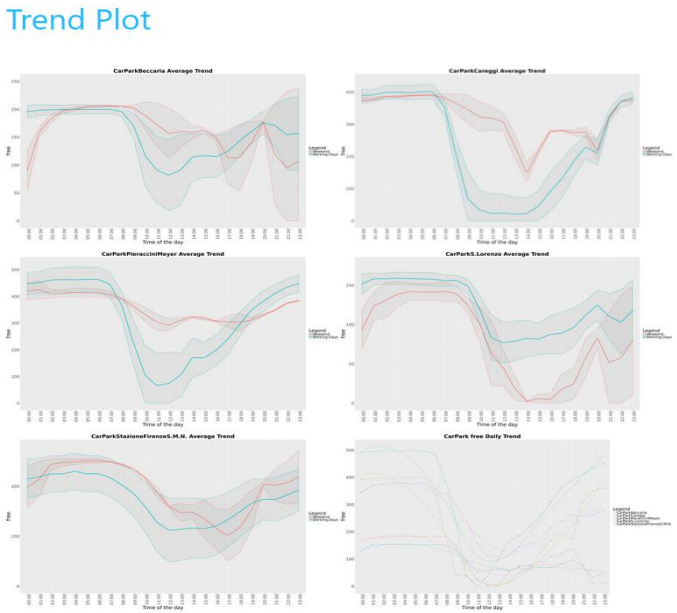


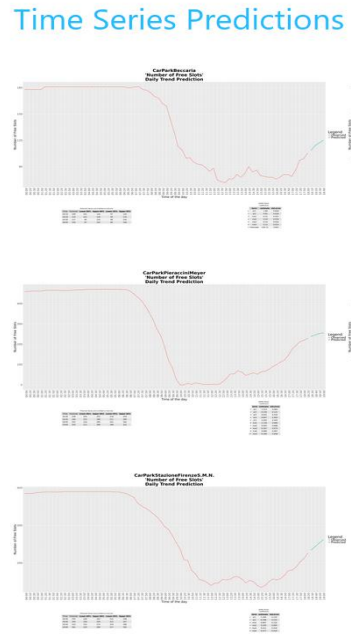
From R studio data analytics to MicroService



R Studio algorithms are automatically transformed into MicroServices for your IOT Applications

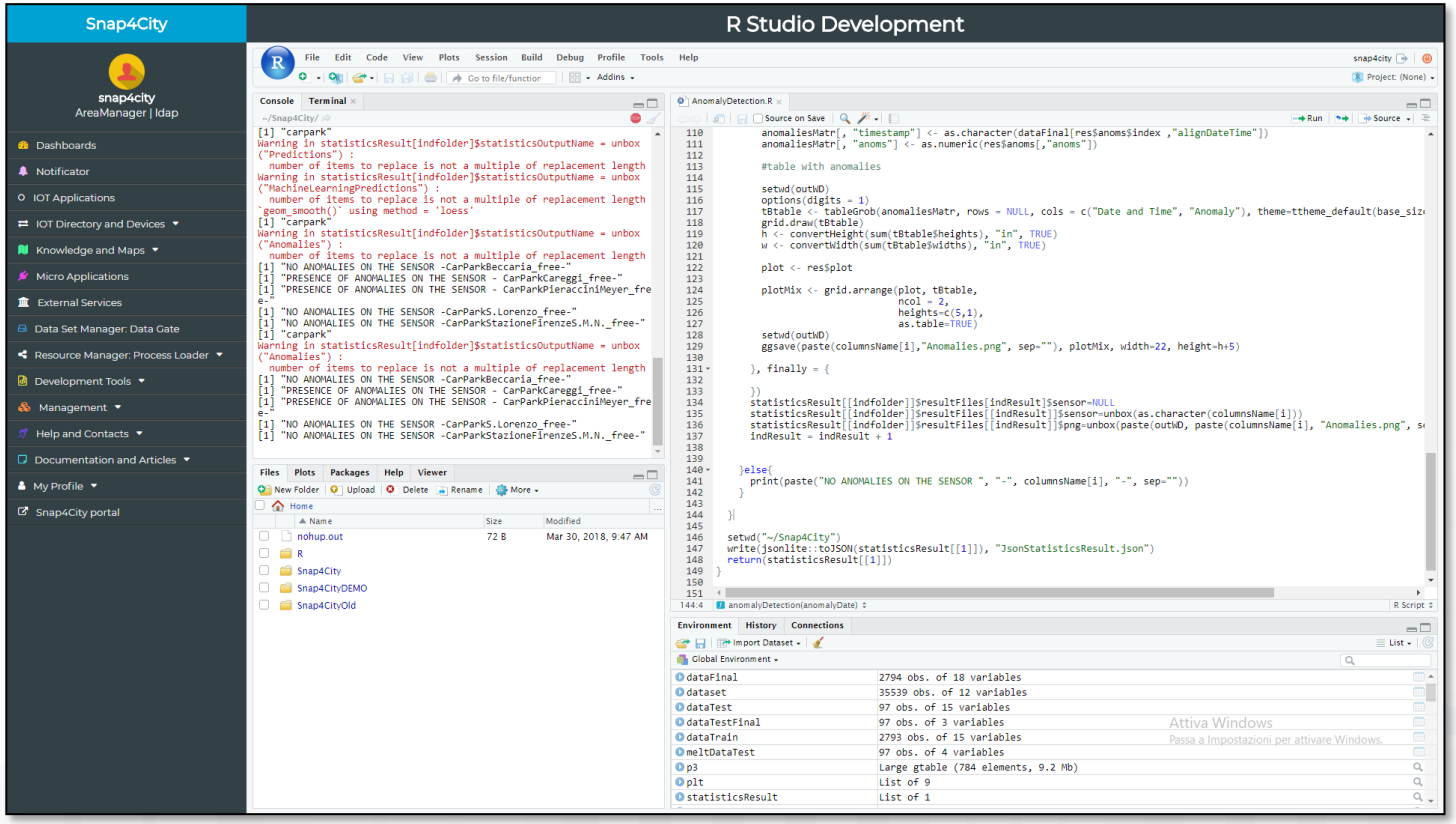


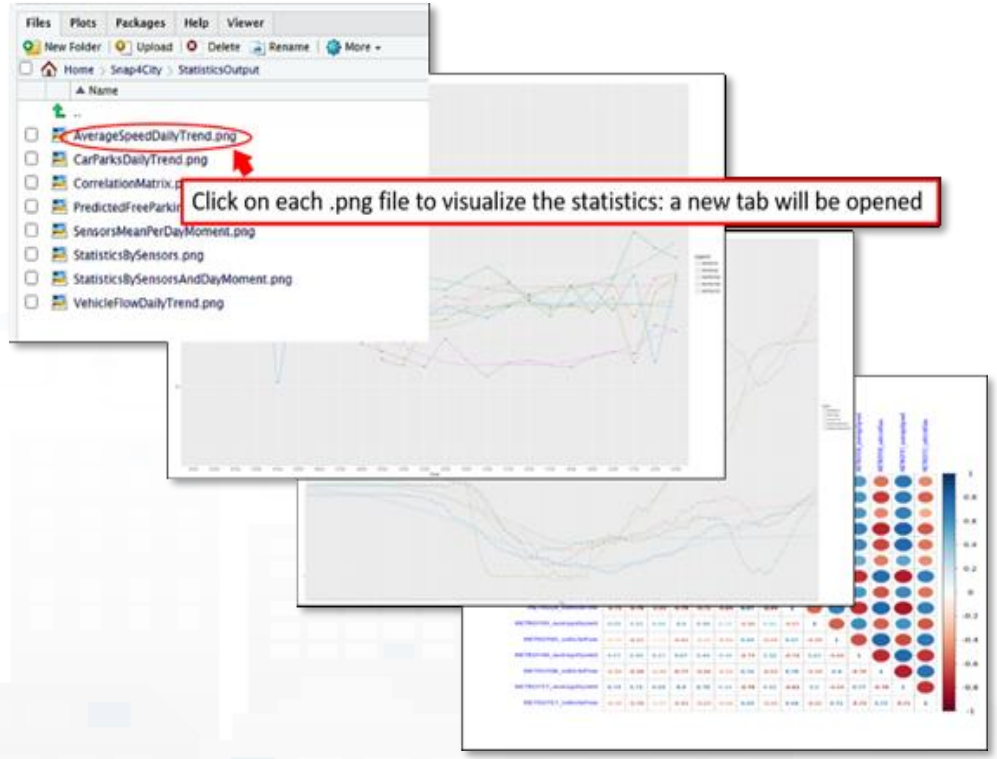


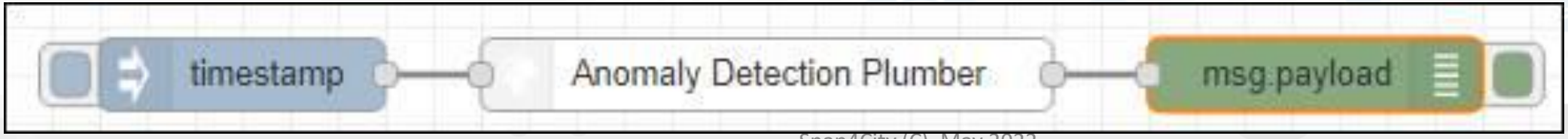


SNAP4CITY

Developing in R Studio and/or Tensor Flow

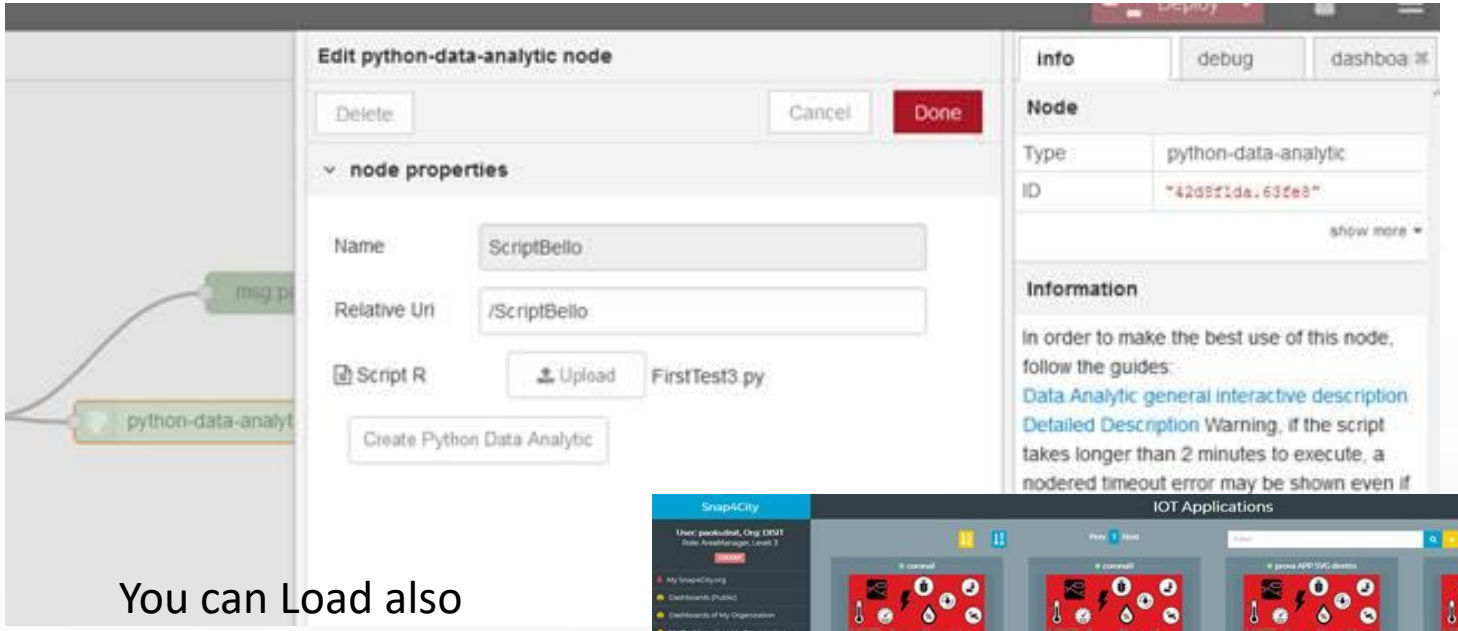






Python process

- Develop Python code exploiting Flask calls
- Test on local for the Call
- Test on Cloud for API
- Deploy via IOT App

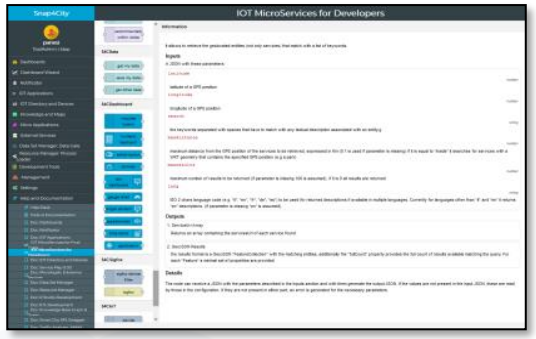


You can Load also ZIP Files including Code and dataModels

IOT Applications Development

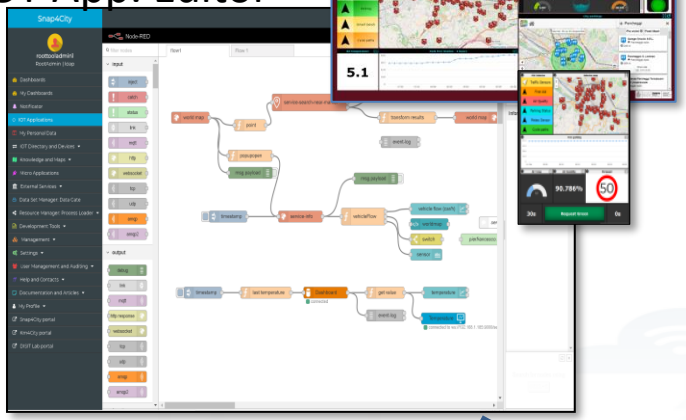
IOT Discovering

MicroServices collections

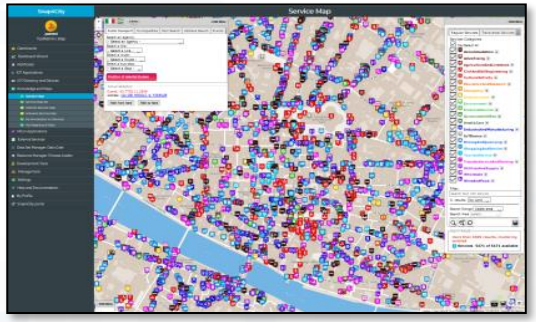


My IOT Applications

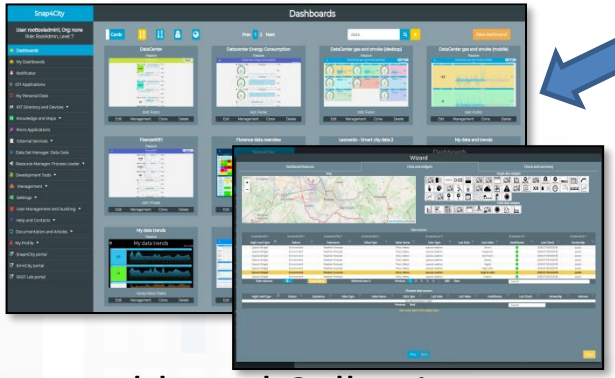
IOT App. Editor



Generating IOT App
With Dashboard



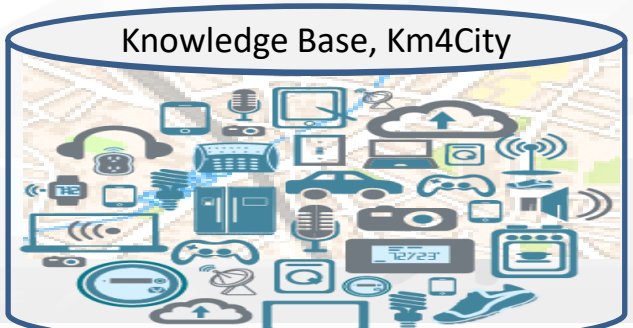
ServiceMap Discovery



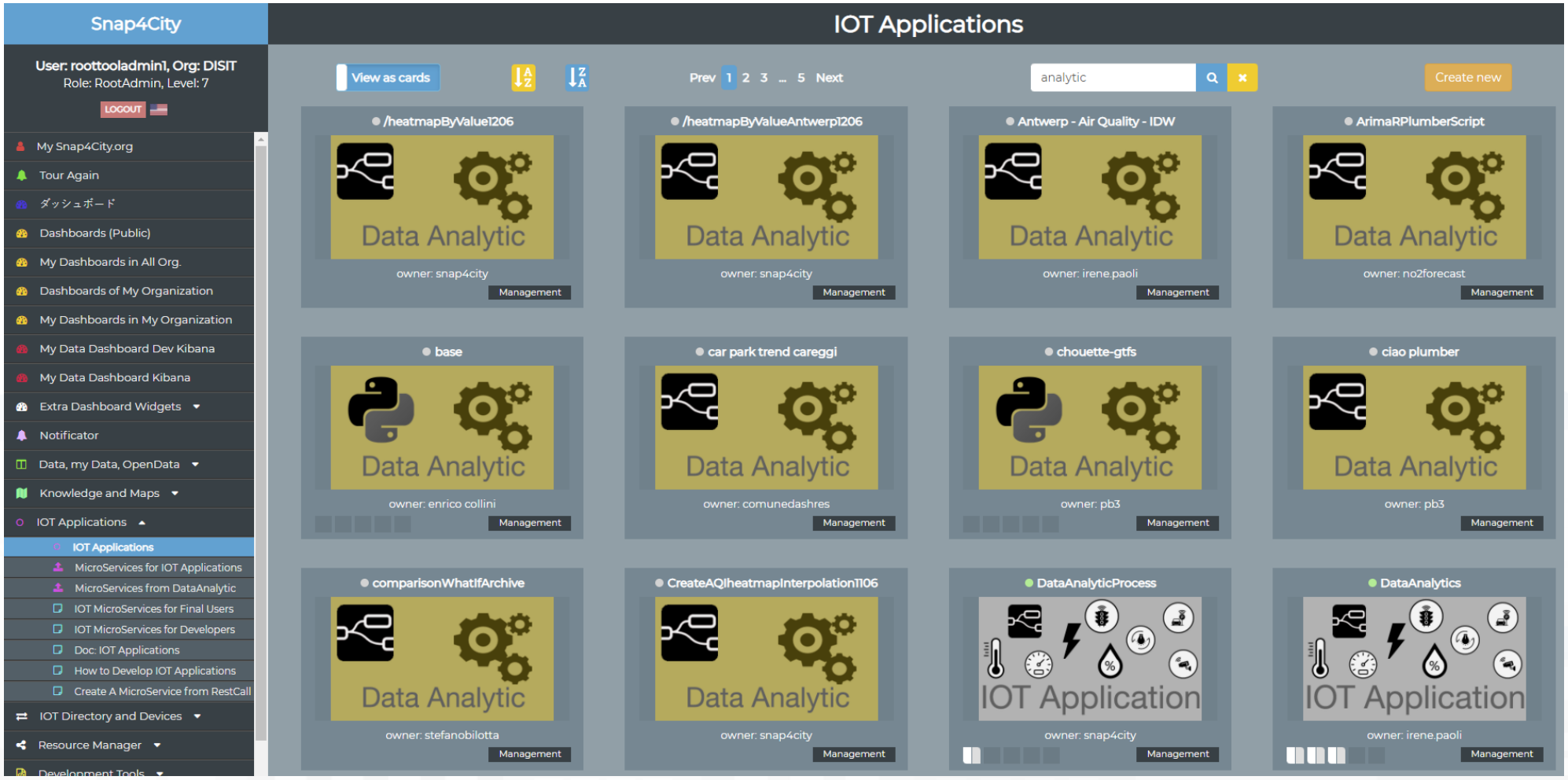
Dashboard Collection,
Editor and Wizard

Sharing/saving
reusing IOT App

Resource Manager



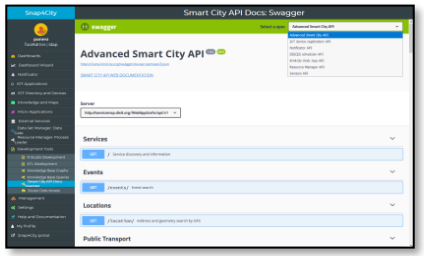
More information



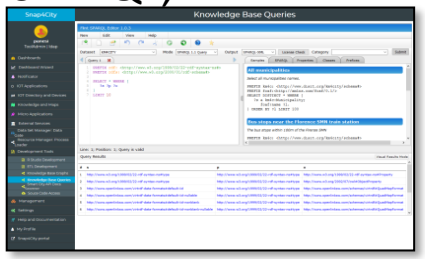
- [HOW TO: develop DataAnalytic in Python and manage them via IOT App]

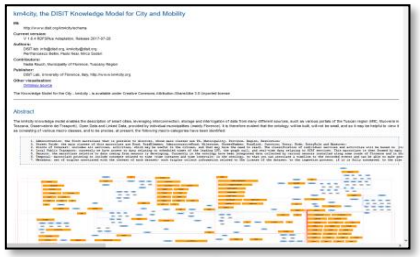
Data Analytics Dev. in Java

Swagger



SPARQL, FLINT



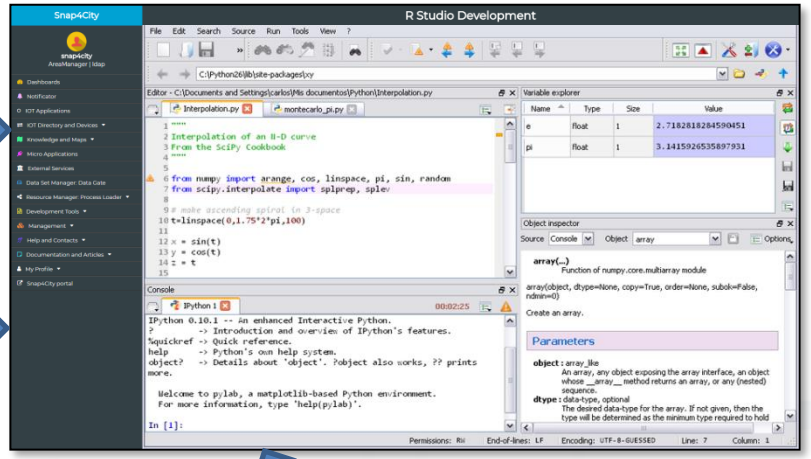


Ontology Schema

LOG.disit.org

Big Data Store Facility

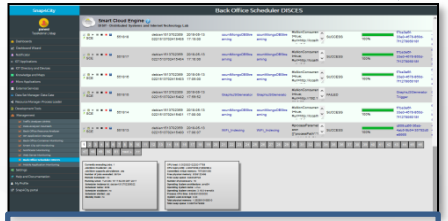
Smart City API from Knowledge Base and other tools



Coding
Testing

Saving /
Sharing
resusing

DISCES scheduler



Monitoring

Resource Manager

Process

Data sources

Distributed Back Office

How to work with R Studio in Snap4City

How to work with R Studio in Snap4City

Data Manipulation using RStudio

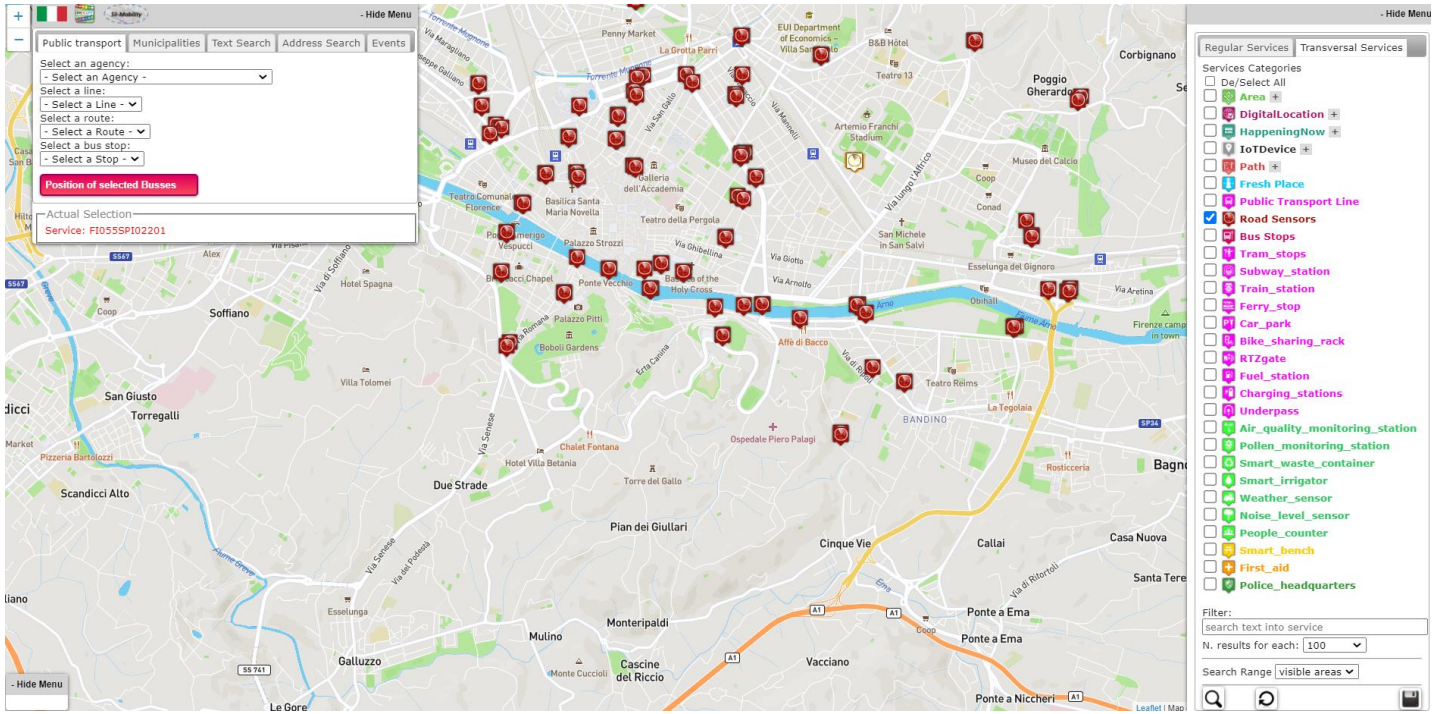
1. How to download Real-Time data using APIs
2. How to manipulate data
3. How to create

4. ARIMA time series prediction

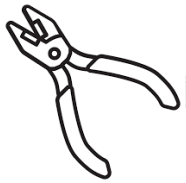
Service Map

1. http://servicemap.km4city.org/WebAppGrafo/
2. Choose a sensor category
3. Select a sensor and save its

sensor_uri

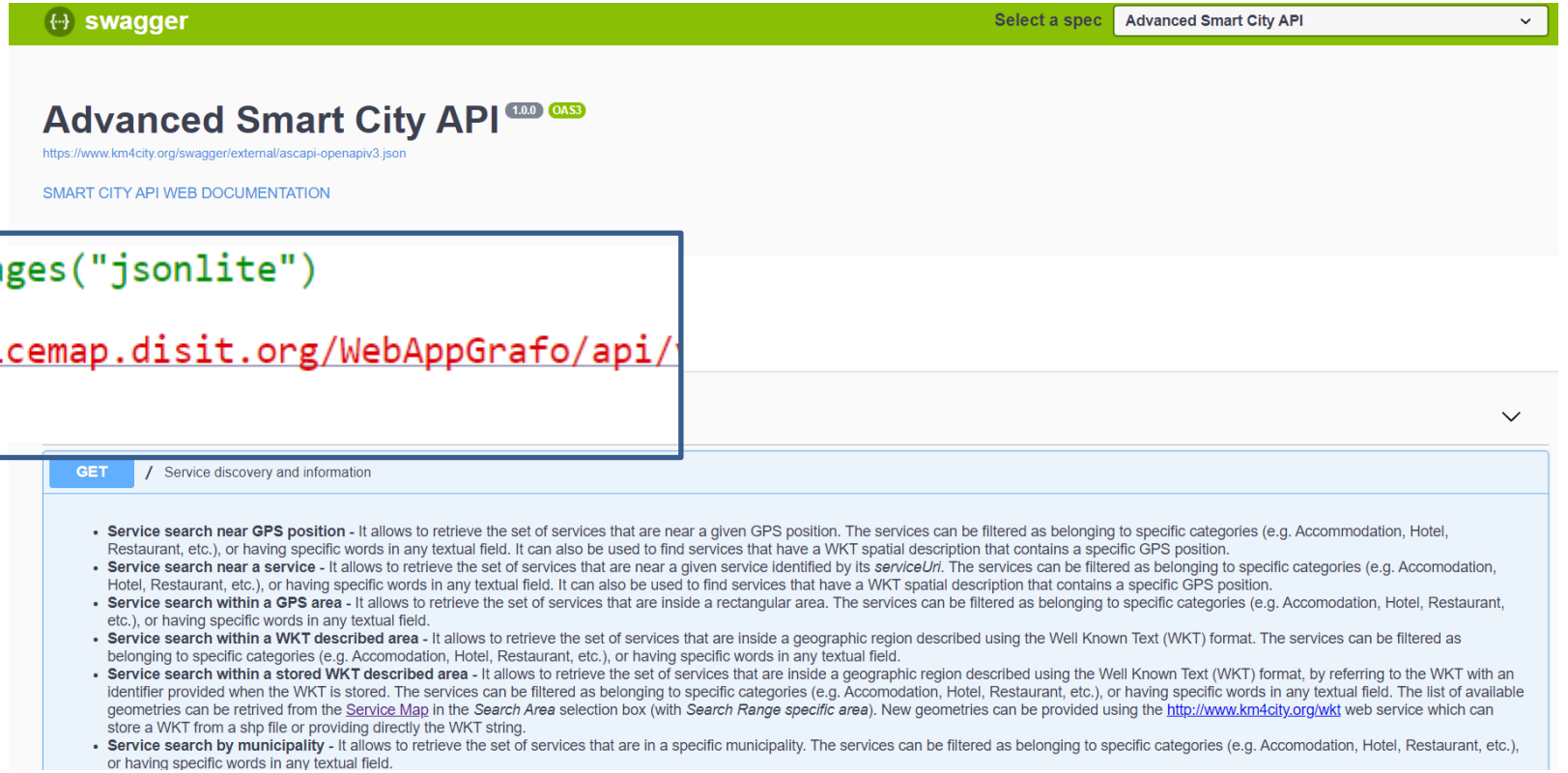


Download the data on RStudio



1. https://www.km4city.org/swagger

   /external/index.html




```
#install.packages("jsonlite")
library("jsonlite")
json_res <- read_json("https://servicemap.disit.org/WebAppGrafo/api/
View(json_res)
```

Data Manipulation using

1. How to download Real-Time data using APIs

```
# 3 # CREATE THE DATASET ##################################################
# | Date | VehicleFlow| Hour | Minutes | DayOfTheWeek | ||||||||
data = matrix(NA, ncol= 5, nrow = data_dim)
colnames(data) = c("Date", "VehicleFlow", "Hour", "Minutes", "DayofTheWeek")
data = as.data.frame(data)
```

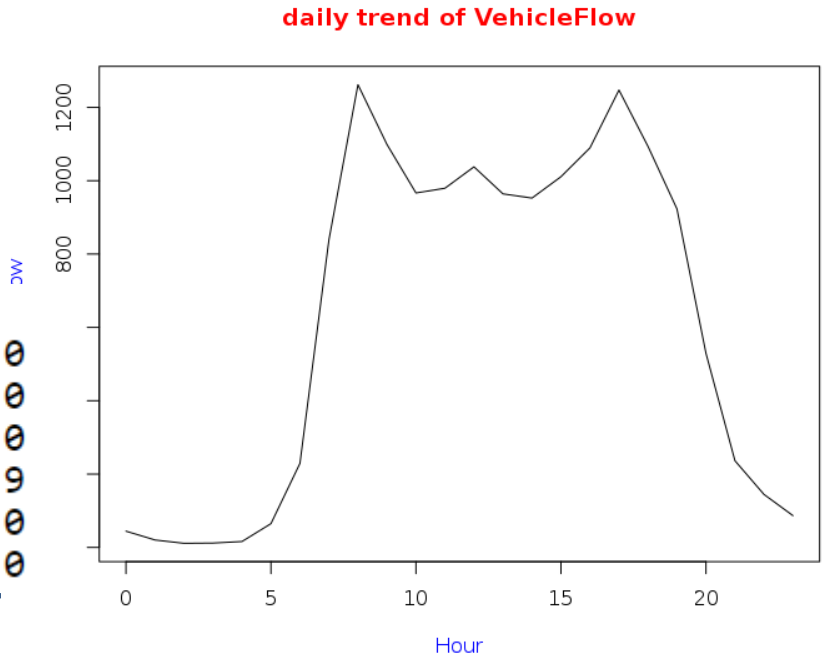
3. How to create

4. ARIMA time series prediction

Data Manipulation using RStudio

1. How to download Real-Time data using APIs
2. How to manipulate data
3. How to create

```
  VehicleFlow
Min.   :   0.0
1st Qu.:  48.0
Median : 552.0
Mean   : 622.9
3rd Qu.:1104.0
Max.   :2520.0
```

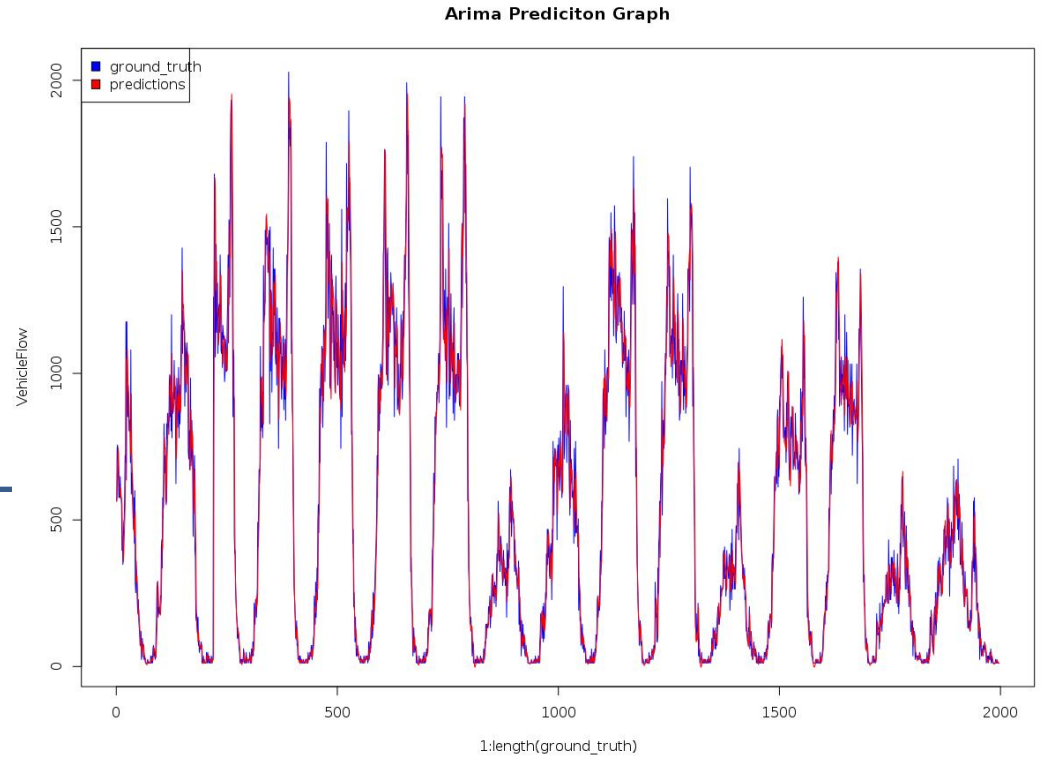


4. ARIMA time series prediction

Data Manipulation using RStudio

1. How to download Real-Time data using APIs
2. How to manipulate data
3. How to create

4. ARIMA time series prediction



DEMO Section 1

Real Time Data Analytics using R Studio for time-series prediction.

Time series predictions IoTApp &

1. How to create a plumberized R script
2. How to create an IoT Application for Rea-Time time-series prediction & visualization
3. How to visualize the predictions using the default Node-Red Dashboard

Plumberized R Script

- https://www.rplumber.io/

- **PLUMBER** is an **R** package that generates a web API from the **R** code you already have

workshop_plumber_script.R | newheatmap.R | example.R*

Source on Save

```
1  #' @get /ARIMApredictions
2  #' @serializer unboxedJSON
3
4  ARIMApredictions <- function(toDay,sensorU
5    print("--------------------------------
6    print("--------- R Plumberized Script --
7    print("--------- timeseries predictions
8    print("---------- example using ARIMA --
```

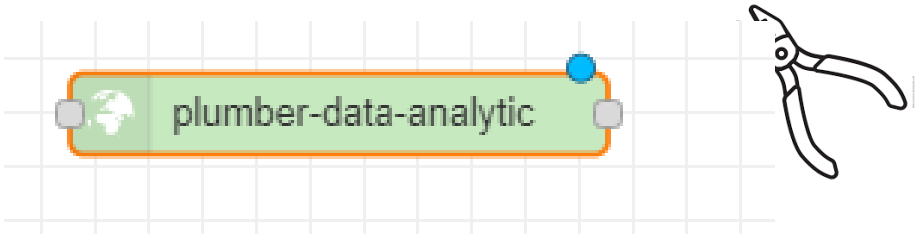
Time series predictions IoTApp & RStudio

1. How to create a plumberized R script

2. How to create an IoT Application for Rea-Time time-series prediction & visualization

3. How to visualize the predictions using the default Node-Red Dashboard

IoTApp plumber-data-analytic



workshop_plumber_script.R | newheatmap.R | example.R*

Source on Save

```
1  #' @get /ARIMApredictions
2  #' @serializer unboxedJSON
3
4  ARIMApredictions <- function(toD
5    print("------------------------
6    print("--------- R Plumberized
7    print("--------- timeseries pr
8    print("---------- example usin
```

Edit plumber-data-analytic node

Delete | Cancel | Done

Properties

Name: ArimaRPlumberScript

Relative Uri: /ARIMApredictions

Script R | Upload | workshop_plumber_script …

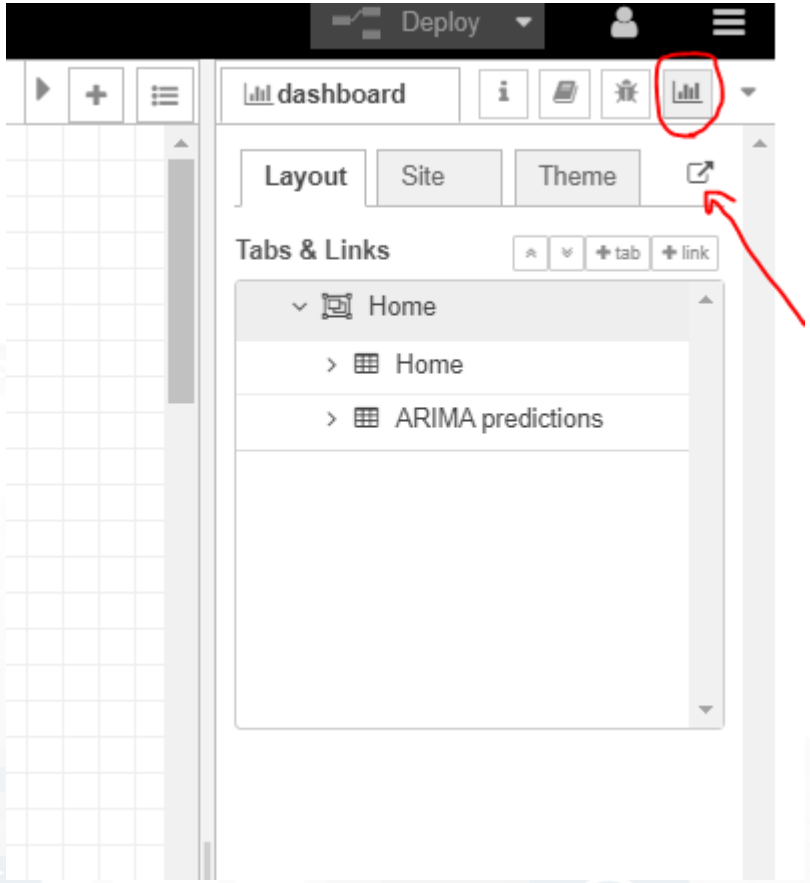
Create Plumber Data Analytic

Time series predictions IoTApp & RStudio

How to visualize the predictions

using the default Node-Red

Dashboard



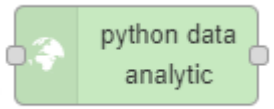
ANGULARJS

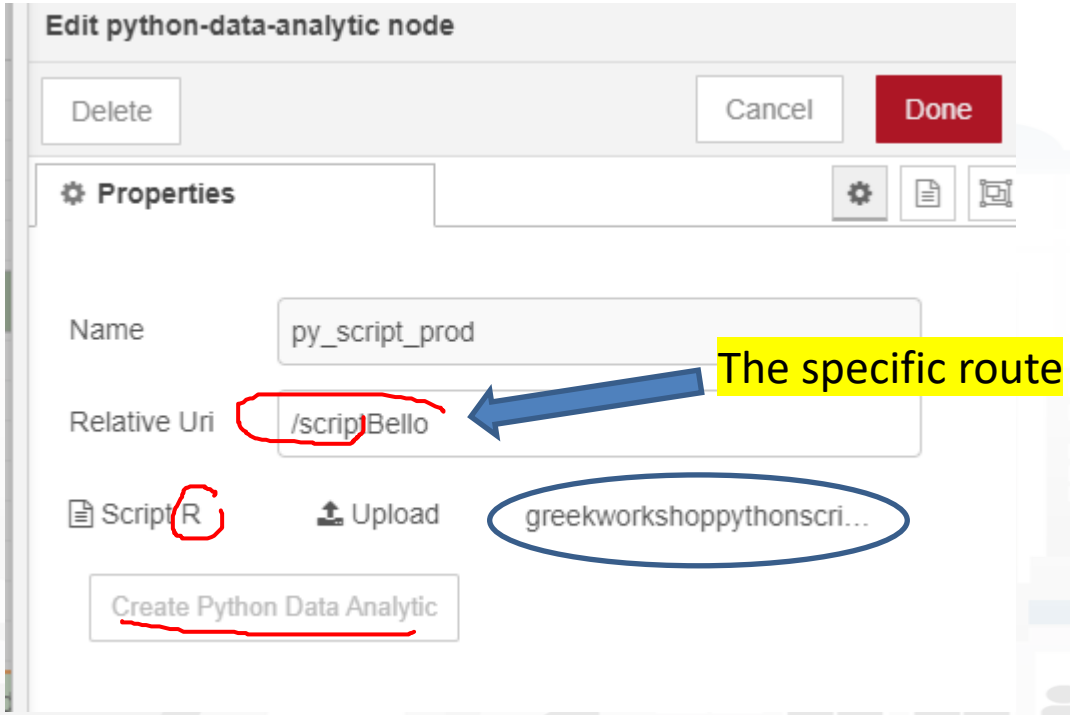
DEMO Section 2

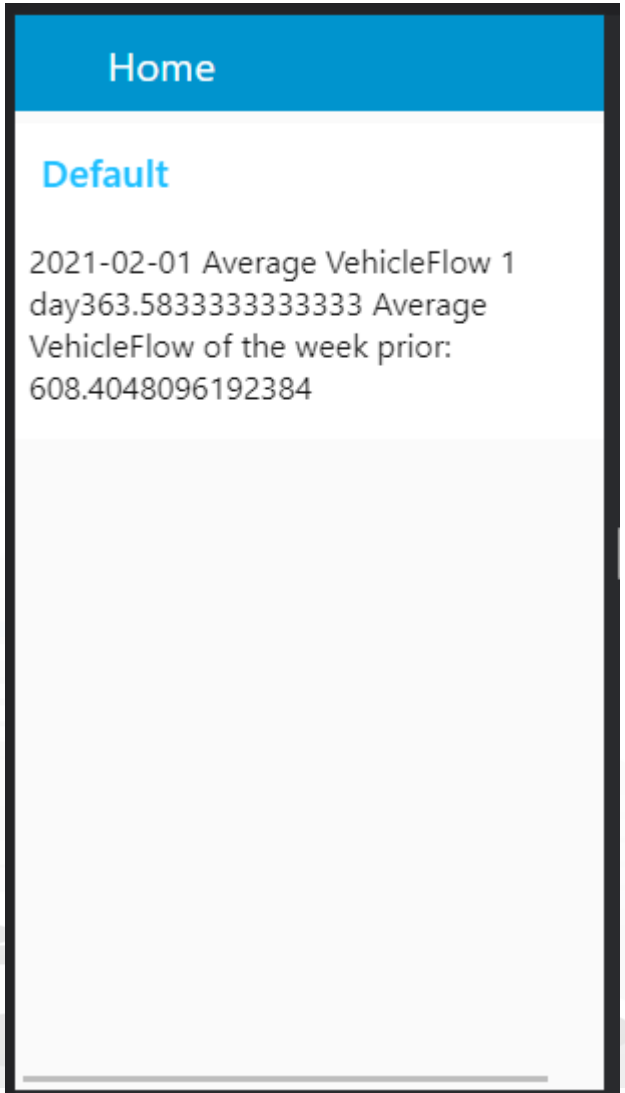
Iot App with Python Data Analytics Script

Average VehicleFlow of METRO768

- 1 DAY PRIOR TO A SPECIFIC DATE
- 1 WEEK PRIOR TO THE SPECIFIC DATE





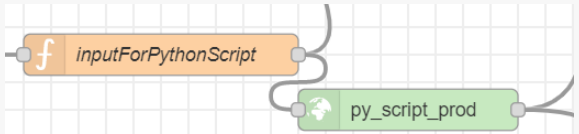


Python Data Analytics Script setup for IoT App

```python
#import json
import flask
import os
import sys
import requests
import datetime
import pandas as pd
from datetime import date
import numpy as np

app = flask.Flask(__name__)
os.environ["FLASK_APP"]= __name__ + ".py"
append_services = []
@app.route('/scriptBello', methods=['GET', 'POST'])

def scriptBello():
    #GET PARAMS IN INPUT (day_date,sensor_uri)
    try:
        if flask.request.method == 'GET':
            day_date = str(flask.request.values.get('day_date'))
            sensor_uri = str(flask.request.values.get('sensor_uri'))
            #print("day_date:", day_date)
            #print("sensor_uri:", sensor_uri)
        else:
            return("Error")

    except Exception as e:
        print ("Error: "+str(e))
        message = "Error: "+str(e);
        return message
```

The specific route

The parameters specified in json format in input to the python_data_analytic block




```
msg.payload = {
    "day_date" : "2021-02-01T00:00:00",
    "sensor_uri" : "http://www.disit.org/km4city/resource/METRO768"
}
return msg;
```

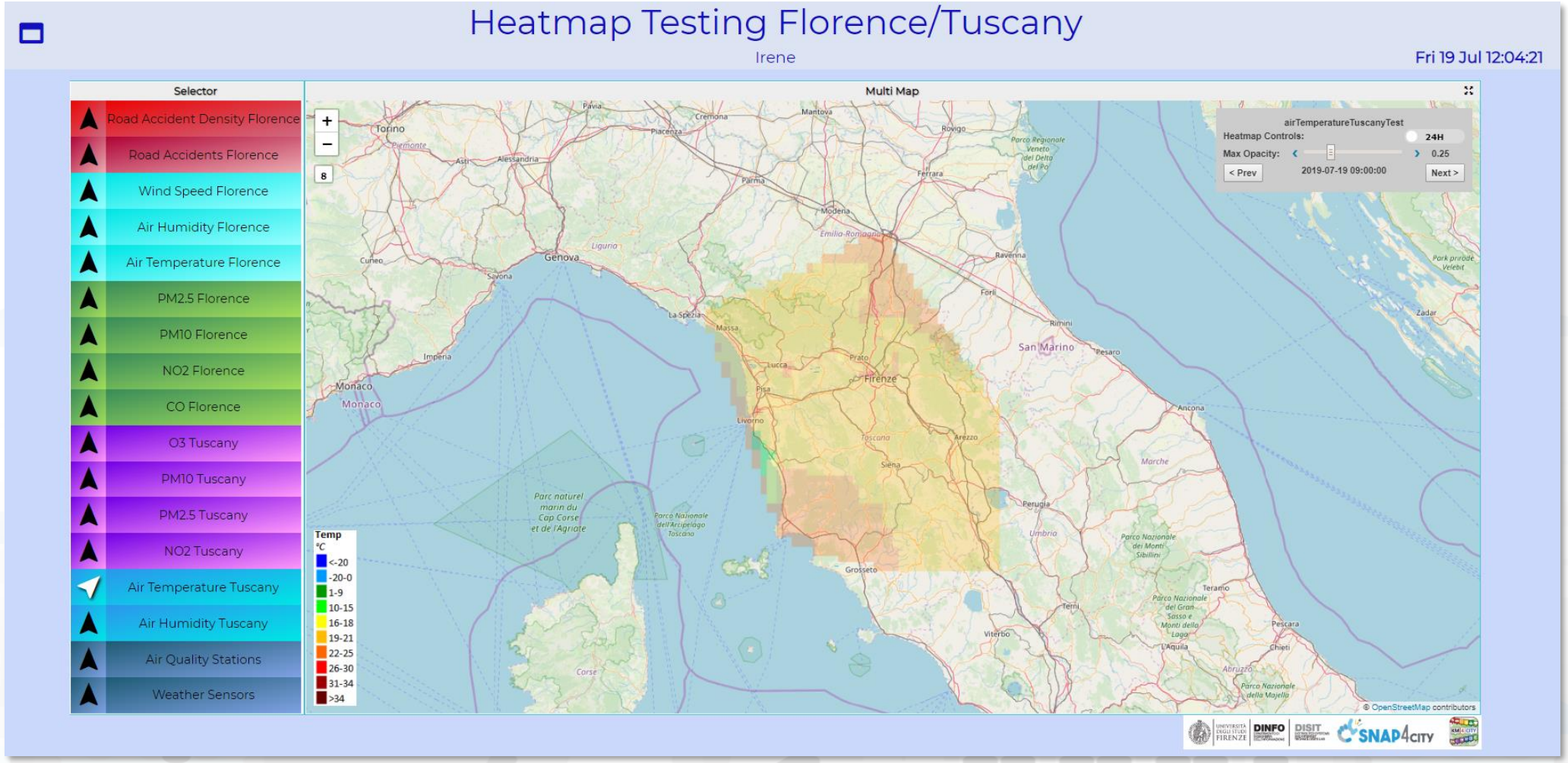
Python Data Analytics Script setup for IoT App

```
    #SCRIPT CODE HERE:
    try:
        # some parsing for the processing of the dates
        date =day_date.split("T")[0]
        year = int(date.split("-")[0])
        month = int(date.split("-")[1])
        day =  int(date.split("-")[2])
        #print(year, month, day)
        dayBefore = str(datetime.datetime(year, month, day) - datetime.timedelta(days=1)).split(" ")[0] + "T00:00:00"
        weekBefore = str(datetime.datetime(year, month, day) - datetime.timedelta(days=7)).split(" ")[0] + "T00:00:00"

        #get data since yesterday
        sensor_data = getDatafromSensorURI(sensor_uri, dayBefore, day_date)
        res_since_yesterday = []
        for i in range(len(sensor_data)):
            res_since_yesterday.append(float(sensor_data[i]['vehicleFlow']['value']))
        #calculating the average value
        average_VehicleFlow_since_yesterday = np.mean(res_since_yesterday)

        #get data of the last week
        sensor_data = getDatafromSensorURI(sensor_uri, weekBefore, day_date)
        res_last_week = []
        for i in range(len(sensor_data)):
            res_last_week.append(float(sensor_data[i]['vehicleFlow']['value']))
        #calculating the average value
        average_VehicleFlow_last_week = np.mean(res_last_week)

        return ("<h1> Average VehicleFlow since yersterday: " + str(average_VehicleFlow_since_yesterday)
                +"<br> Average VehicleFlow of last week: " + str(average_VehicleFlow_last_week)
                + "<br> </h1> <br> Script Execution Completed")
    except Exception as e:
        print ("Error: "+str(e))
        message = "Error: "+str(e);
        return message

if __name__ == "__main__":
    app.run(host="127.0.0.1", port=5000)
```

host = ‘127,0,0,1’
Port = ‘5000’

Real Time Data Analytics and Heatmaps creation using Rstudio and IotApp

Heatmap Visualization



https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MTI2OA==

Real Time Data Analytics and Heatmaps creation using RStudio

1. How to create a *plumberized* R script:

   - How to download Real-Time data using APIs
   - How to save heatmaps using APIs

2. How to create an IOT Application for Real-Time Data Analytics:

   - How to upload the R script and create a plumber instance

3. How to visualize the created heatmap in a dashboard

Real Time Data Analytics using R Studio

How to create a *plumberized* R script – 1

PLUMBER is an **R** package that generates a web API from the **R** code you already have.

- Step 1 - *Plumberize* the code:

```
#' @get /TuscanyHeatmap
#' @serializer unboxedJSON
```

❖ In order to send a response from R to an API client, the object must be *serialized* into some format that the client can understand (JSON format).

Note that, **@get** and **@serializer** annotations must to be put on the top of the code.

Any comments must not be inserted before the annotations or between them and the R function.

Real Time Data Analytics using R Studio

How to create a *plumberized* R script – 2

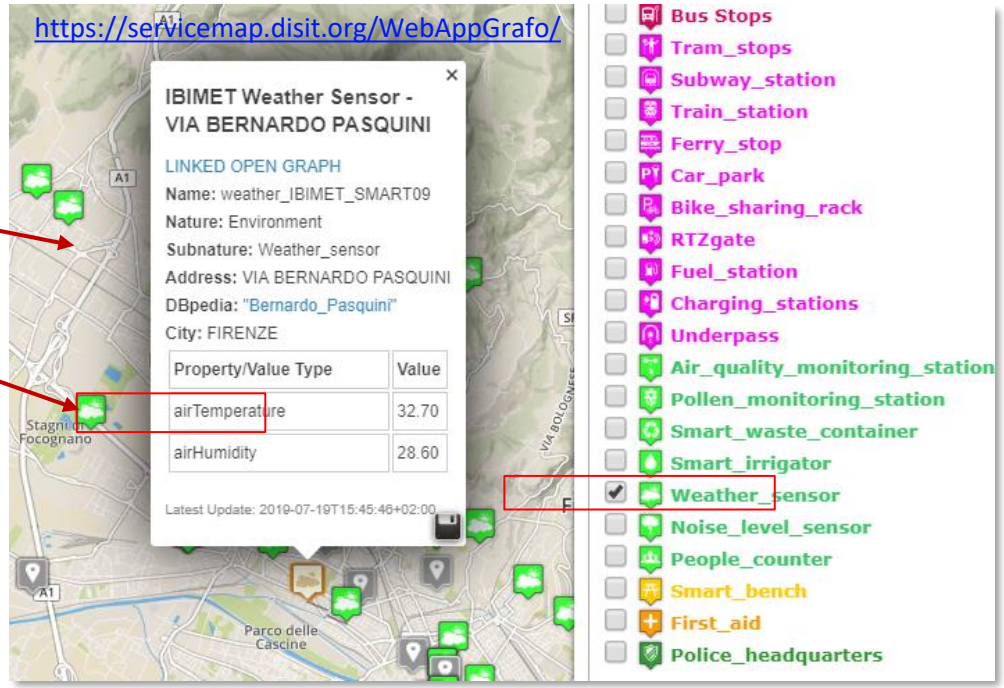
- Step 2 - Create an R function with the same name of the **@get** parameter:

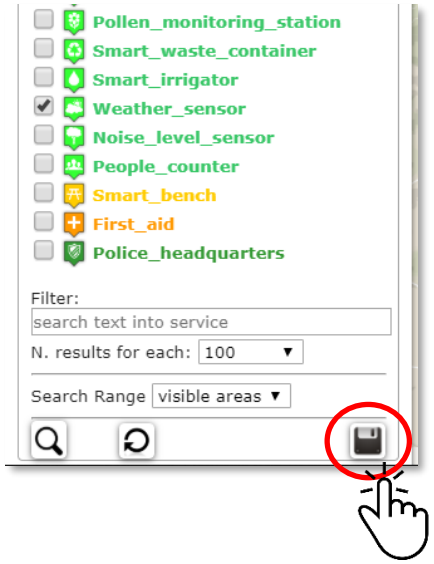
```
TuscanyHeatmap <- function(sensorCategory, varName, fromDateTime, toDateTime, heatmapName){
```

heatmapName = “airTemperatureTuscanyTest”

sensorCategory = “Weather_sensor”
varName = “airTemperature”

toDateTime = “2020-04-14T10:00:00”
fromDateTime = “2-hour”





Real Time Data Analytics using R Studio

How to create a *plumberized* R script – 3

- Step 3 - Upload All Service Uris (sensor stations) from service map in the area of interest:

```
query <- paste("https://servicemap.disit.org/WebAppGrafo/api/v1/?selection=
               42.50247797334869;8.19580078125;44.606112745173 9;13.422546386718 7
               &categories=",sensorCategory,
       "&maxResults=0&maxDists=0.1&format=json",sep="")

sensorCategoryJson <- fromJSON(query) #jsonlite package

suri <- sensorCategoryJson$Services$features$properties$serviceUri #serviceUri
```

Real Time Data Analytics using R Studio

How to create a *plumberized* R script - 3

```
https://servicemap.disit.org/WebAppGrafo/api/v1/?selection=42.6789
73163549 54;9.954032295814045;44.00523270268637;12.063407295814045&
categories=Weather_sensor&maxResults=0&maxDists=0.1&format=json
```

```
"http://www.disit.org/km4city/resource/IBIMET_SMART11"
"http://www.disit.org/km4city/resource/IBIMET_SMART04"
"http://www.disit.org/km4city/resource/IBIMET_SMART13"
"http://www.disit.org/km4city/resource/IBIMET_SMART06"
"http://www.disit.org/km4city/resource/IBIMET_SMART17"
"http://www.disit.org/km4city/resource/IBIMET_SMART33"
"http://www.disit.org/km4city/resource/IBIMET_SMART33"
"http://www.disit.org/km4city/resource/IBIMET_SMART25"
"http://www.disit.org/km4city/resource/IBIMET_SMART24"
"http://www.disit.org/km4city/resource/IBIMET_SMART30"
[…]
```

Real Time Data Analytics using R Studio

How to create a *plumberized* R script - 4

- Step 4 - Upload data related to a specific time interval (fromTime/toTime) for each Service Uri:

```
sensorData <- vector("list", length(suri))
for (i in 1:length(suri)){
  temp=c()
  #api to upload the realtime data
  api <- paste("https://servicemap.disit.org/WebAppGrafo/api/v1/?serviceUri=",
               suri[i],"&fromTime=",fromDateTime,
               "&toTime=",toDateTime,sep="")
  sensorCategoryData <- fromJSON(api)
```

```
https://servicemap.disit.org/WebAppGrafo/api/v1/?serviceU
ri="http://www.disit.org/km4city/resource/IBIMET_SMART11"
&fromTime=2-hour&toTime=2020-04-14T10:00:00
```

Real Time Data Analytics using R Studio

How to create a *plumberized* R script - 5

- Step 5 – Data manipulation and data Interpolation…

… After data manipulation and interpolation we obtain something like this:

long	lat	value
11.24686	42.76616	39.87238
11.30287	42.76616	39.54115
11.35888	42.76616	39.20993
11.41489	42.76616	38.87870
11.47090	42.76616	38.54747
11.52691	42.76616	38.21624
11.58292	42.76616	37.88501

[…]

Interpolated values

Real Time Data Analytics using R Studio

How to save heatmaps using API - 1

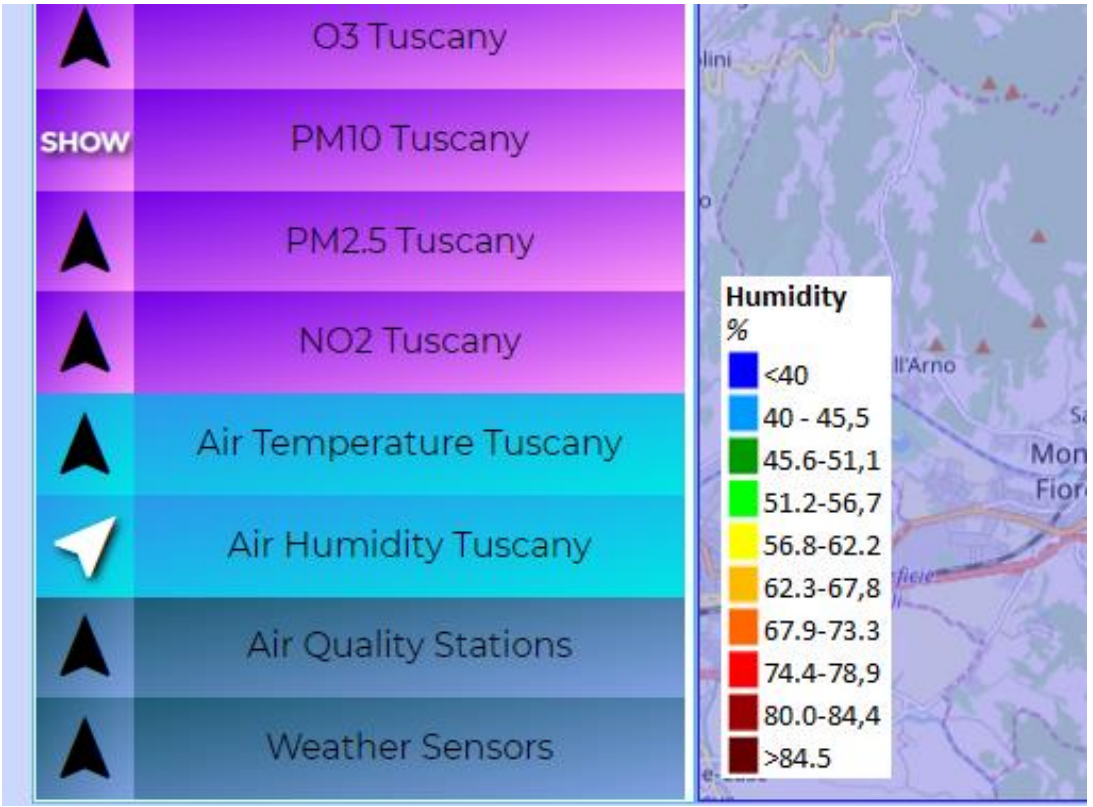
- Step 6 - Create a R list:

```
interpolatedHeatmap=list()
interpolatedHeatmap$attributes=vector("list", dim(interpolatedData)[1])
interpolatedHeatmap$saveStatus=list()

for(i in 1:dim(interpolatedData)[1]) {

  #list
  lat = as.numeric(interpolatedData[i, "lat"])
  long = as.numeric(interpolatedData[i, "long"])
  meanObs = interpolatedData[i, "value"]

  listAttribTemp = list("mapName"=heatmapName,  "metricName"= metricName,
                        "description"= paste("Average from",fromDateTime,"to",toDateTime,sep=" "),
                        "clustered"= 0, "latitude"=lat, "longitude"=long,
                        "value"= meanObs, "date"= paste(toDateTime, "Z", sep=""), "org"="DISIT")

  interpolatedHeatmap$attributes[[i]]=listAttribTemp
```

Real Time Data Analytics using R Studio

How to save heatmaps using API - 2

Note that, the "metricName" identifies the legend for each heatmap and the colour scale to be used.



It corresponds to the *varName* of the R function except for PM10 and PM2.5 measurements:

- "HighDensityPM10"
- "HighDensityPM25"

Real Time Data Analytics using R Studio

How to save heatmaps using API - 3

- Step 7 - Transform the R list in a Json and save heatmap data using API:

```
request_body_json <- toJSON(interpolatedHeatmap$attributes, auto_unbox = TRUE, digits = 10)

resultPOST <- POST(url = "http://snap4city:disit2019@192.168.0.59:8000/insertArray",
                   body = request_body_json,
                   encode = "json", add_headers("Content-Type" = "application/json"))

                                    [...]

        apiFinal <- paste("http://192.168.0.59/setMap.php?mapName=",heatmapName,
                          "&metricName=",metricName,
                          "&date=",paste(toDateTime, "Z", sep=""),
                          "&completed=", completed,  sep="")
      resultPOST <- GET(url = apiFinal)
```

Real Time Data Analytics using R Studio

How to save heatmaps using API - 4

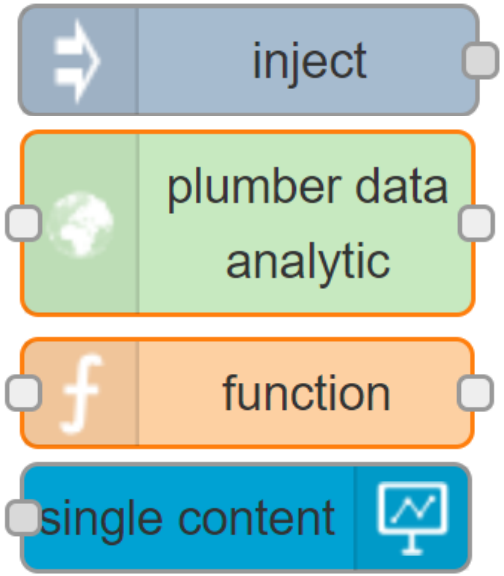
JSON Array Format example

```
[
{
"mapName": “airTemperatureTuscany",
"metricName": "airTemperature",
"description": " Air Temperature heatmap … ",
"clustered": 0,
"latitude": 43.1,
"longitude": 11.1,
"value": 16.5,
"date": "2020-04-14T10:00:00Z"
"org": "DISIT"
}, { […] }]
```

IOT App for Real Time Data Analytics

How to create a Data Analytics IOT Application

What we need:



To insert the R function parameter

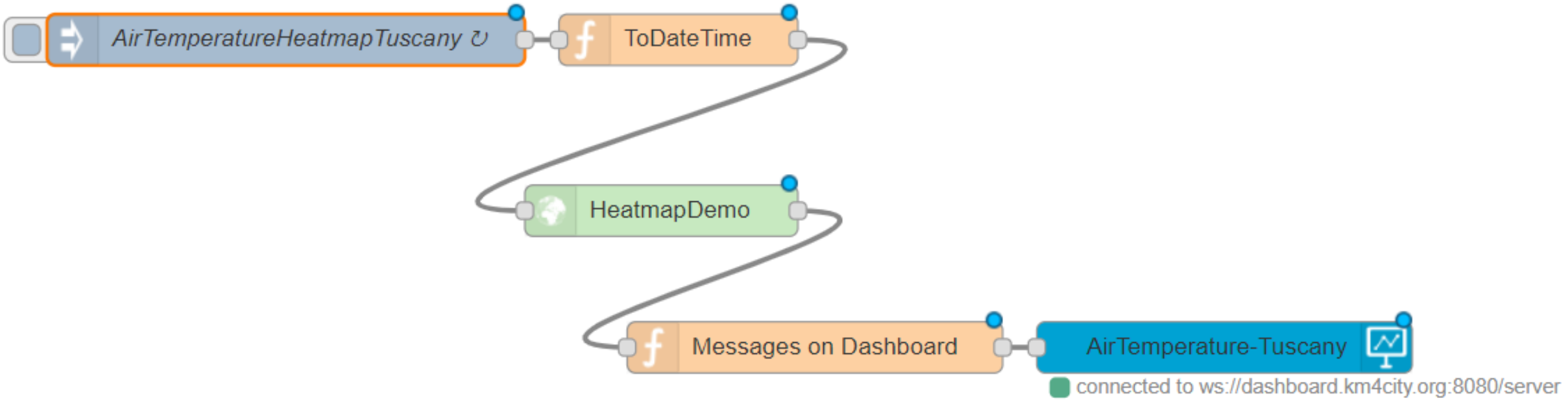
To upload the R script and create a plumber instance

To visualize strings/numbers/html on a dashboard

To execute JavaScript code on output messages

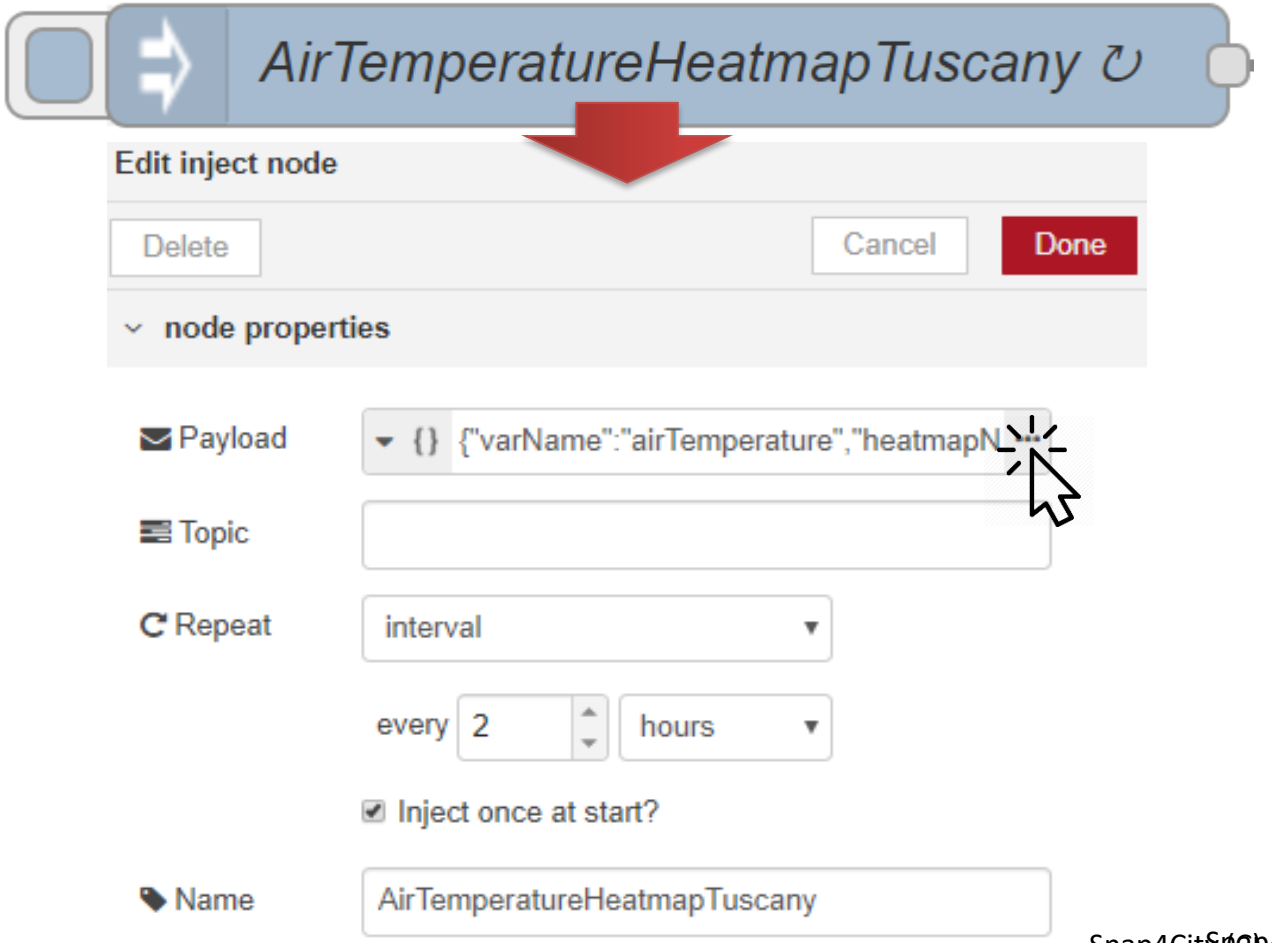
IOT App for Real Time Data Analytics

How to create a Data Analytics IOT Application



IOT App for Real Time Data Analytics
Nodes Configuration – Inject Node

How to configure the **inject** node:



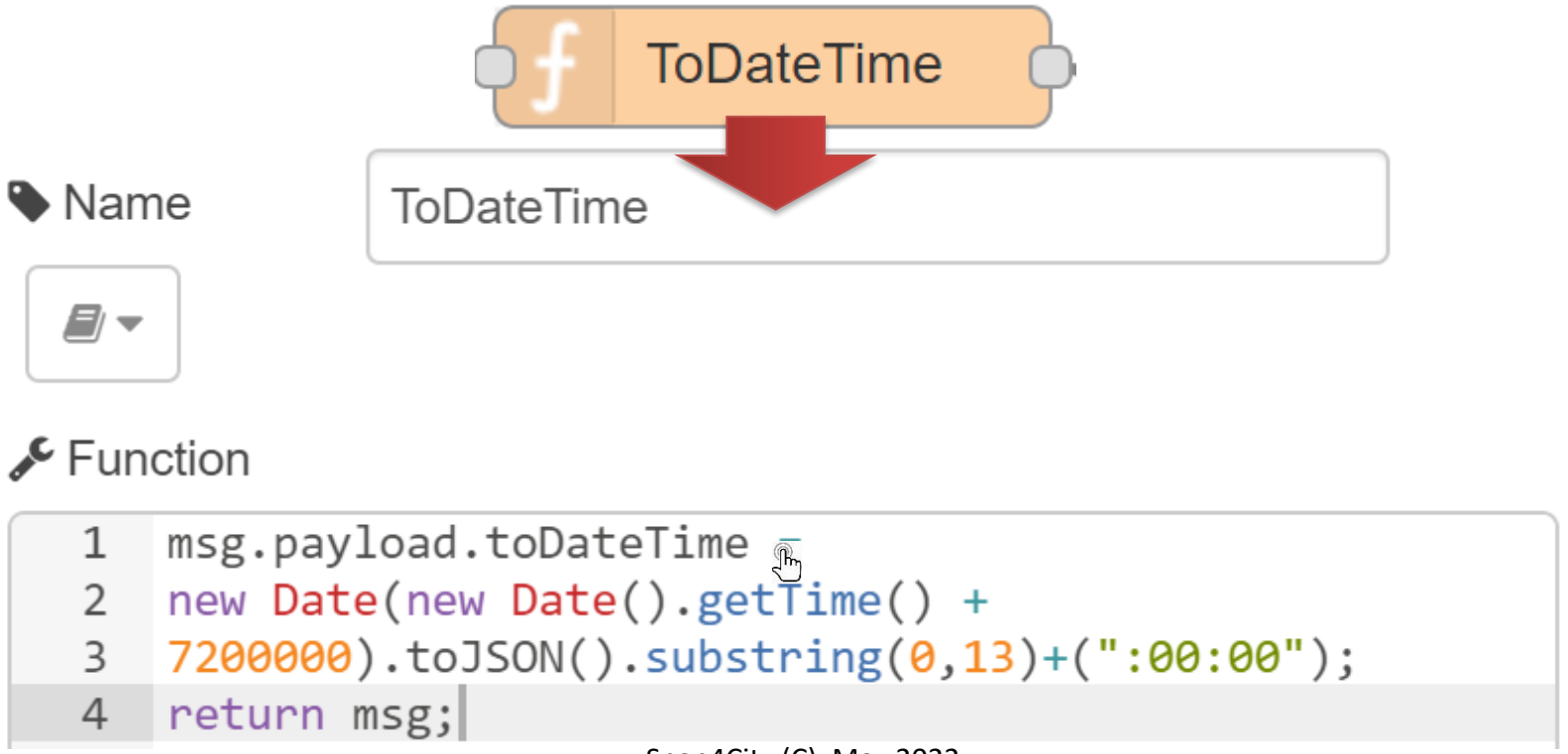
The JSON Format of the Payload property has the same notation of the R function parameters:

```
{   "varName": "airTemperature",
"heatmapName":
"airTemperatureTuscanyTest",
"fromDateTime": "2-hour",
"sensorCategory": "Weather_sensor"
}
```

IOT App for Real Time Data Analytics

Nodes Configuration – Function Node for Date and Time

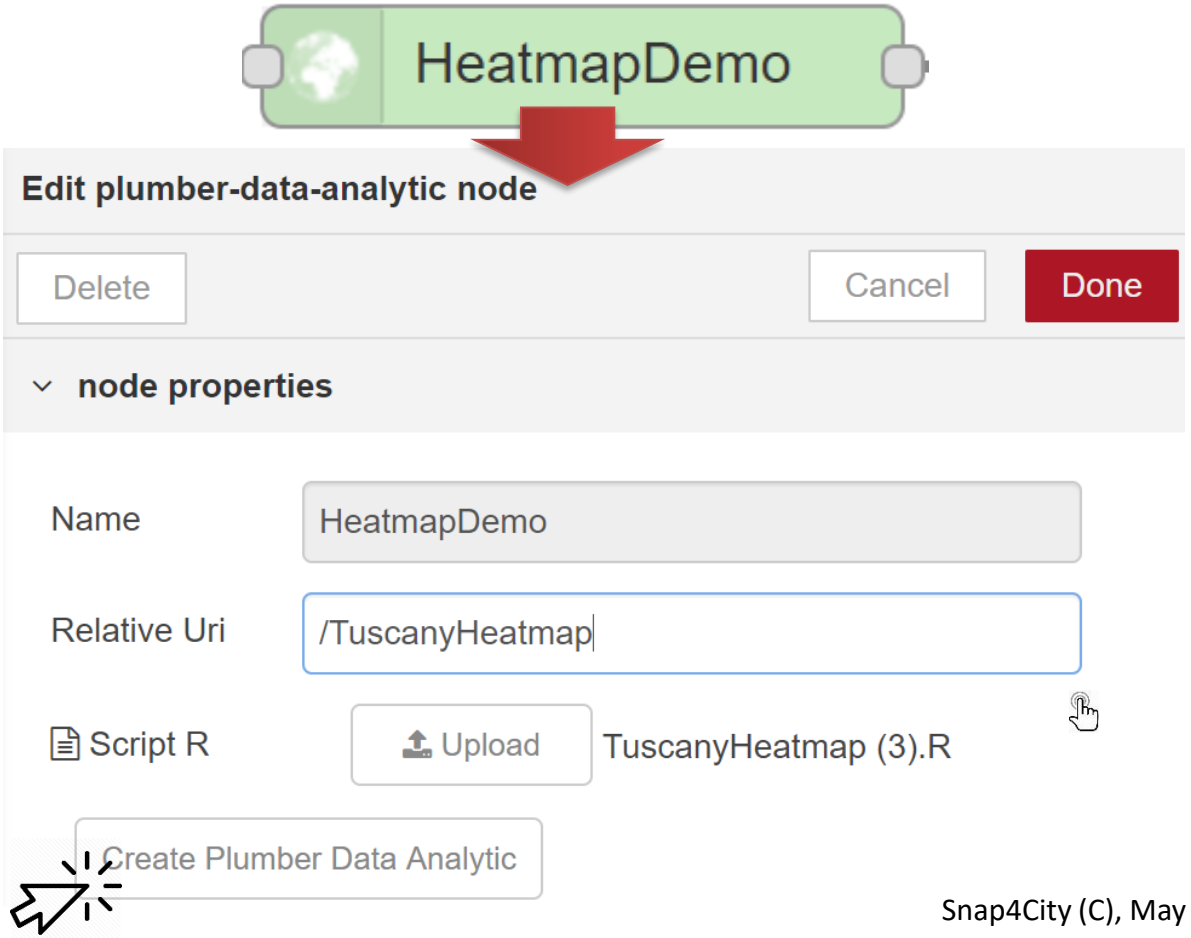
- Before configure the plumber data analytic node is necessary to execute a JavaScript code to dynamically update the date ("toDateTime" parameter):



IOT App for Real Time Data Analytics

Nodes Configuration – Plumber Data Analytic Node

How to configure the **plumber data analytic** node:



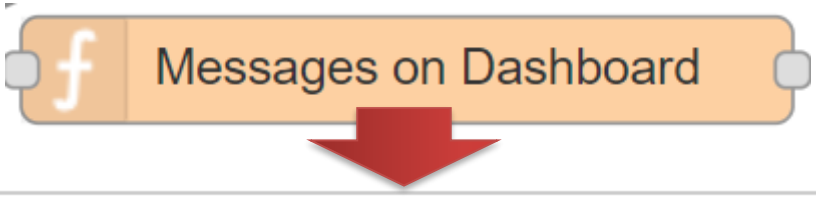
Relative Uri is the same of the R @get annotation:

```
#' @get /TuscanyHeatmap
```

IOT App for Real Time Data Analytics

Nodes Configuration – Function Node for Messages on Dashboard

- Before configure the single content node is necessary to execute a JavaScript code to visualize the status of the heatmap:



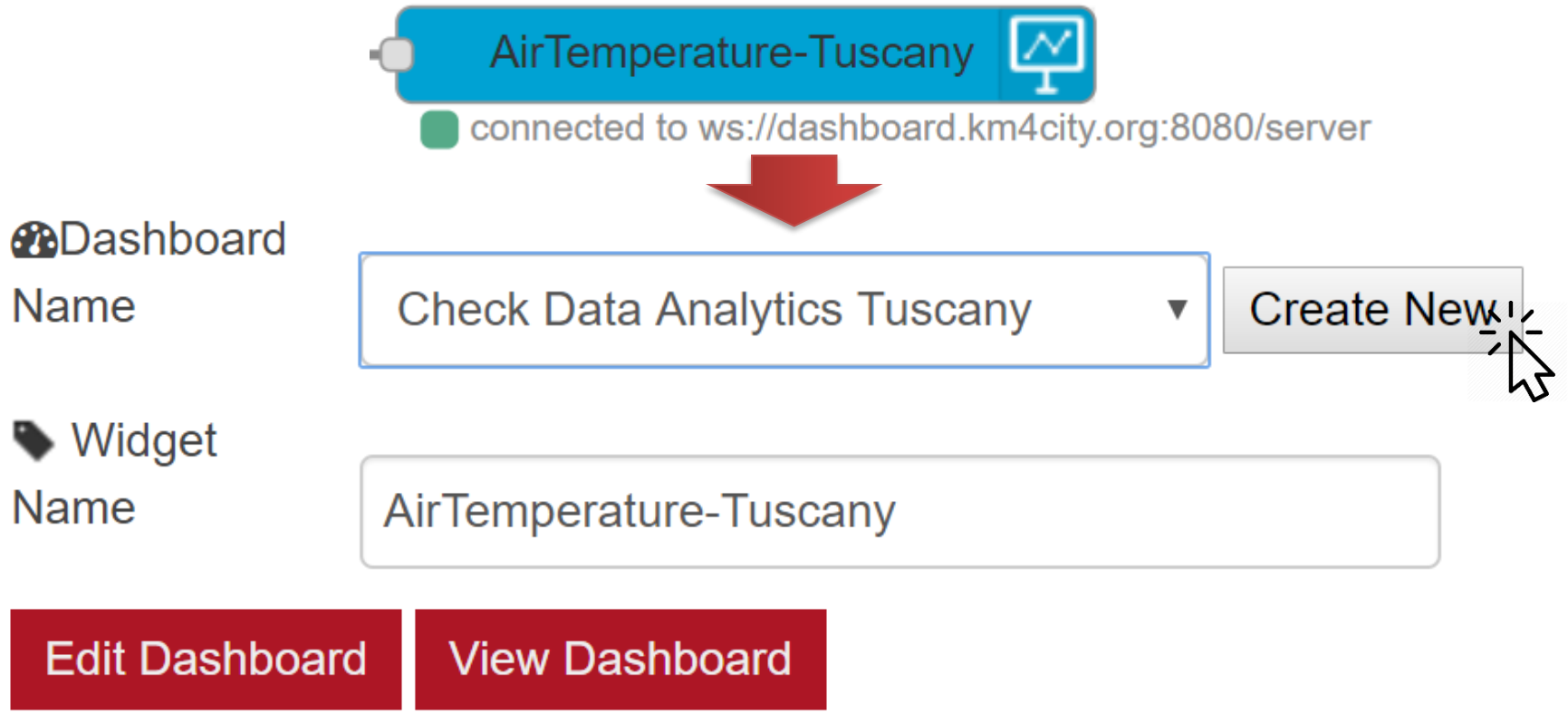
Function

```
1  msg.payload=msg.payload.message+" "+msg.payload.dateTime;
2  if(msg.payload.indexOf("Completed")!= -1){
3      msg.payload ="<span style='color:green;'>"+
4      msg.payload + "</span>"
5  } else  if (msg.payload.indexOf("No Availabe Data") != -1){
6      msg.payload ="<span style='color:orange;'>"+
7      msg.payload + "</span>"
8  }
9  return msg;
```

IOT App for Real Time Data Analytics

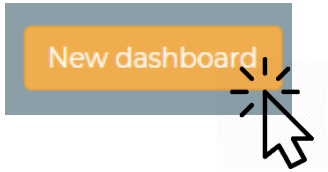
Nodes Configuration – Single Content Node

How to configure the **single content** node:

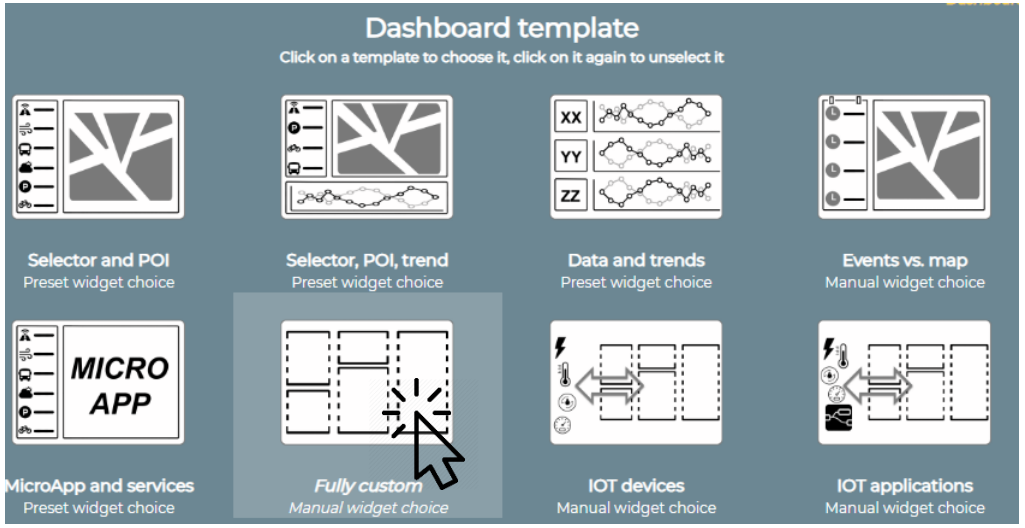


Wizarded Heatmap Visualization

1. Create a New Dashboard from Dashboard (Public)

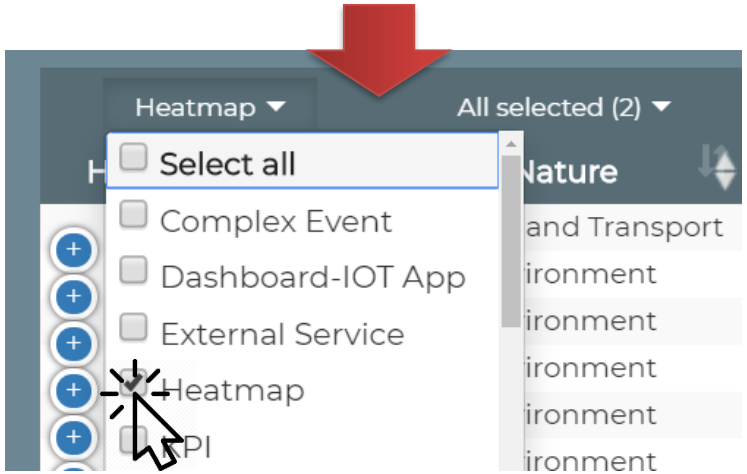


2. Insert a Dashboard Title and select a Dashboard Template

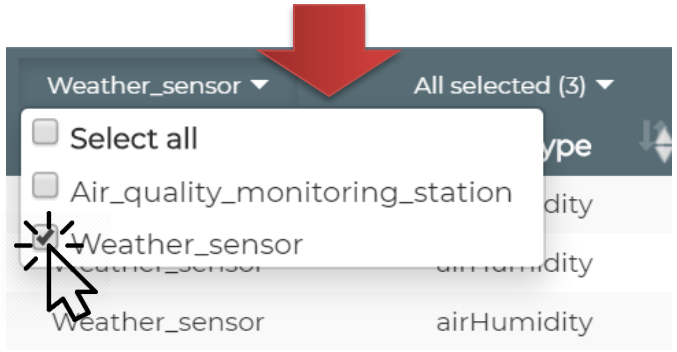


Wizarded Heatmap Visualization

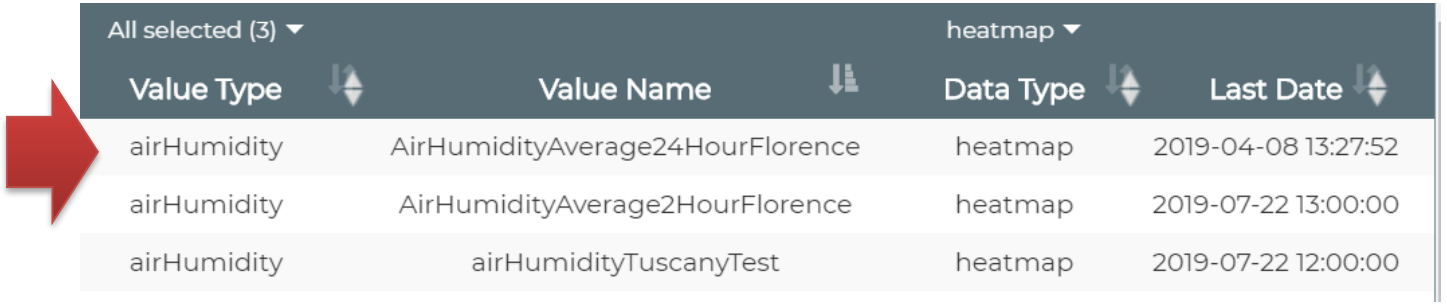
3. Select the Heatmap box as High-Level-Type



4. Select the Sensor Category (Subnature)



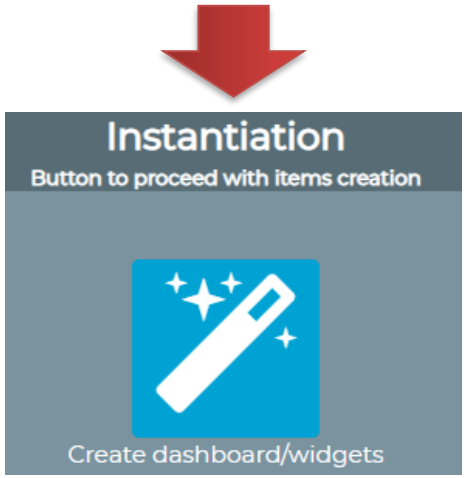
5. Select the measure (Value Type) and the Heatmap Name (Value Name)

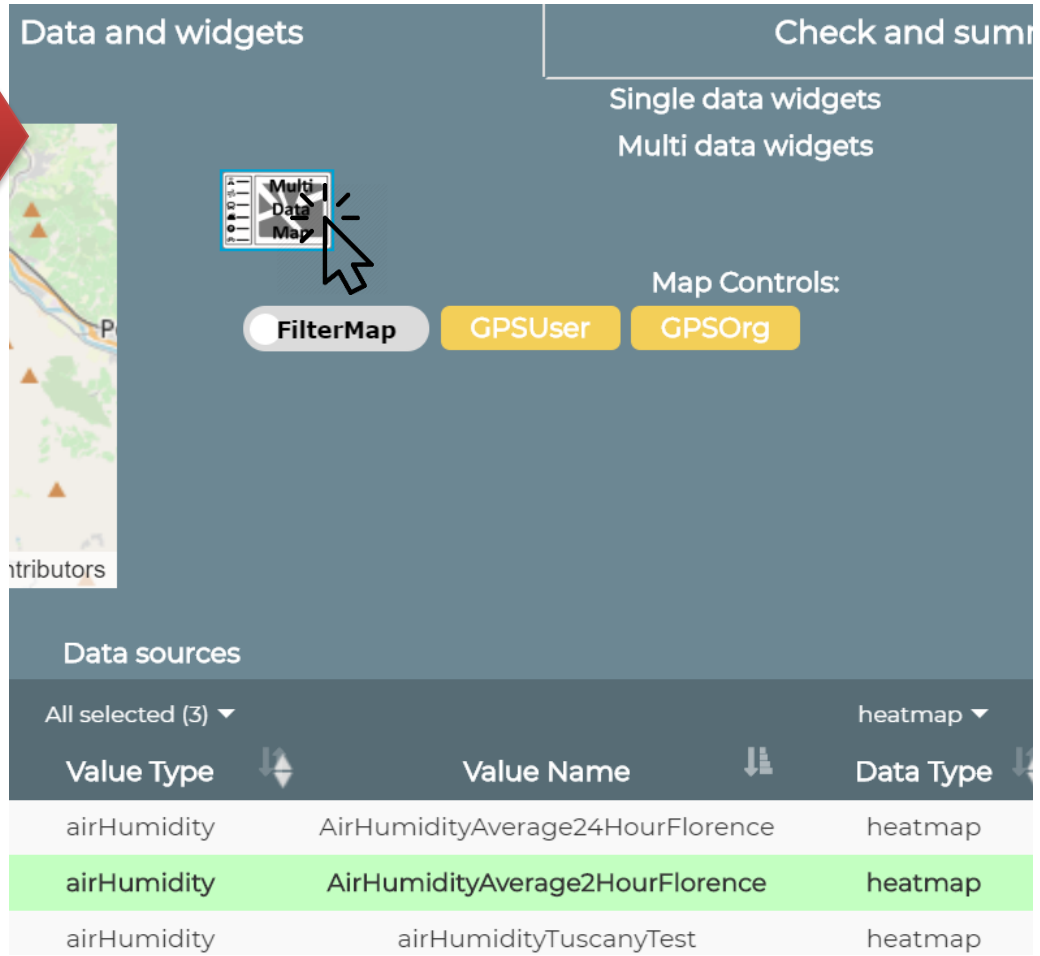


Manually Heatmap Visualization

6. After the Heatmap selection, select the Multi Data Map button and click on next

7. Select the instantiation button to proceed with items creation

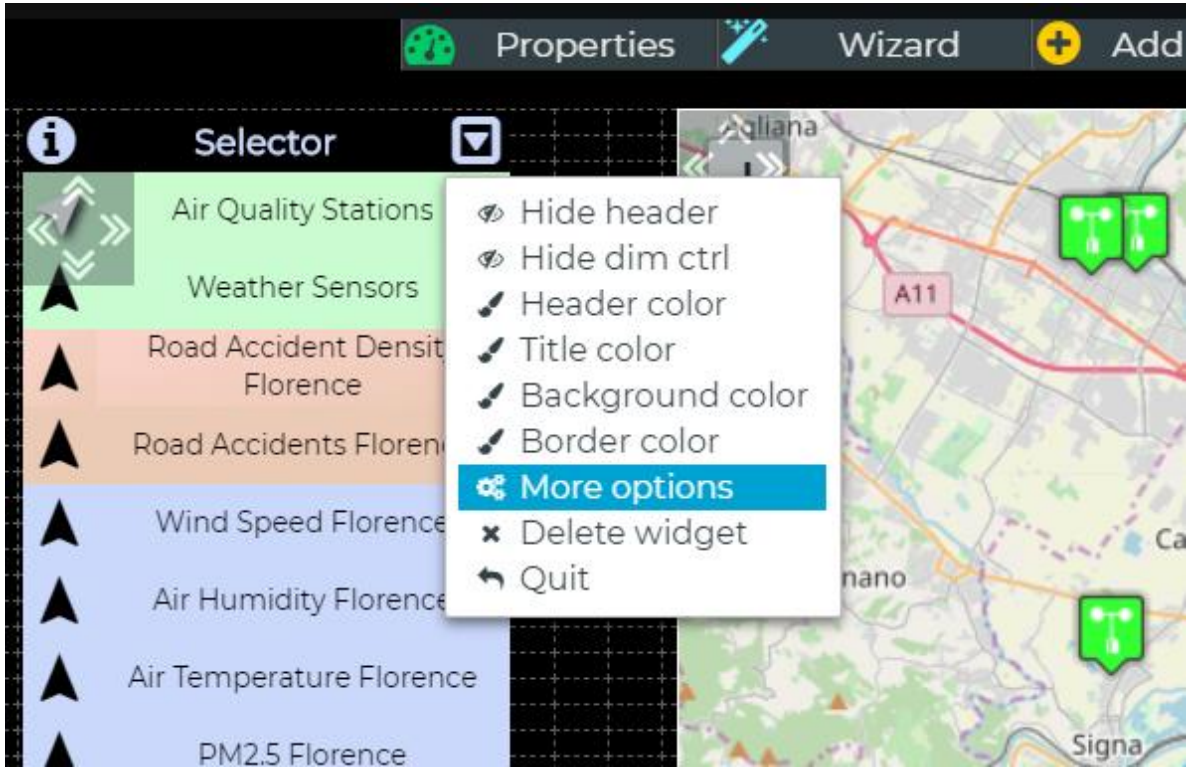




Manually Heatmap Visualization

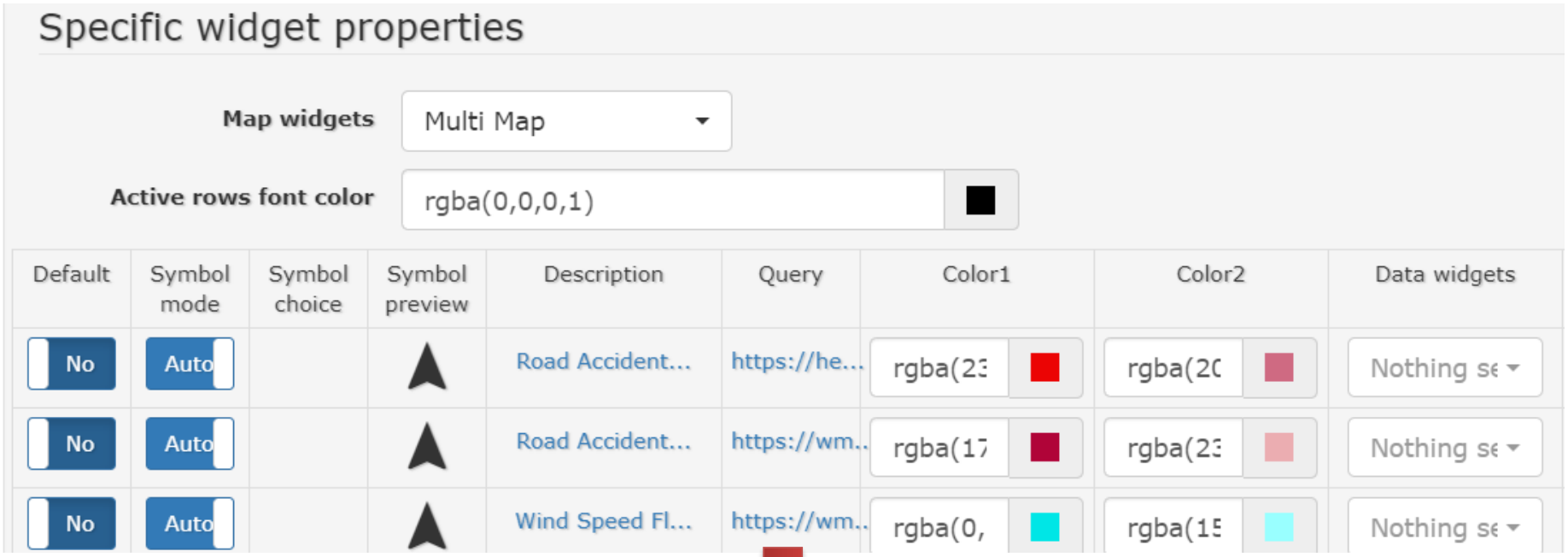
1. Select a Dashboard and click on Edit

2. Select on More Options to modify the widget properties



Manually Heatmap Visualization

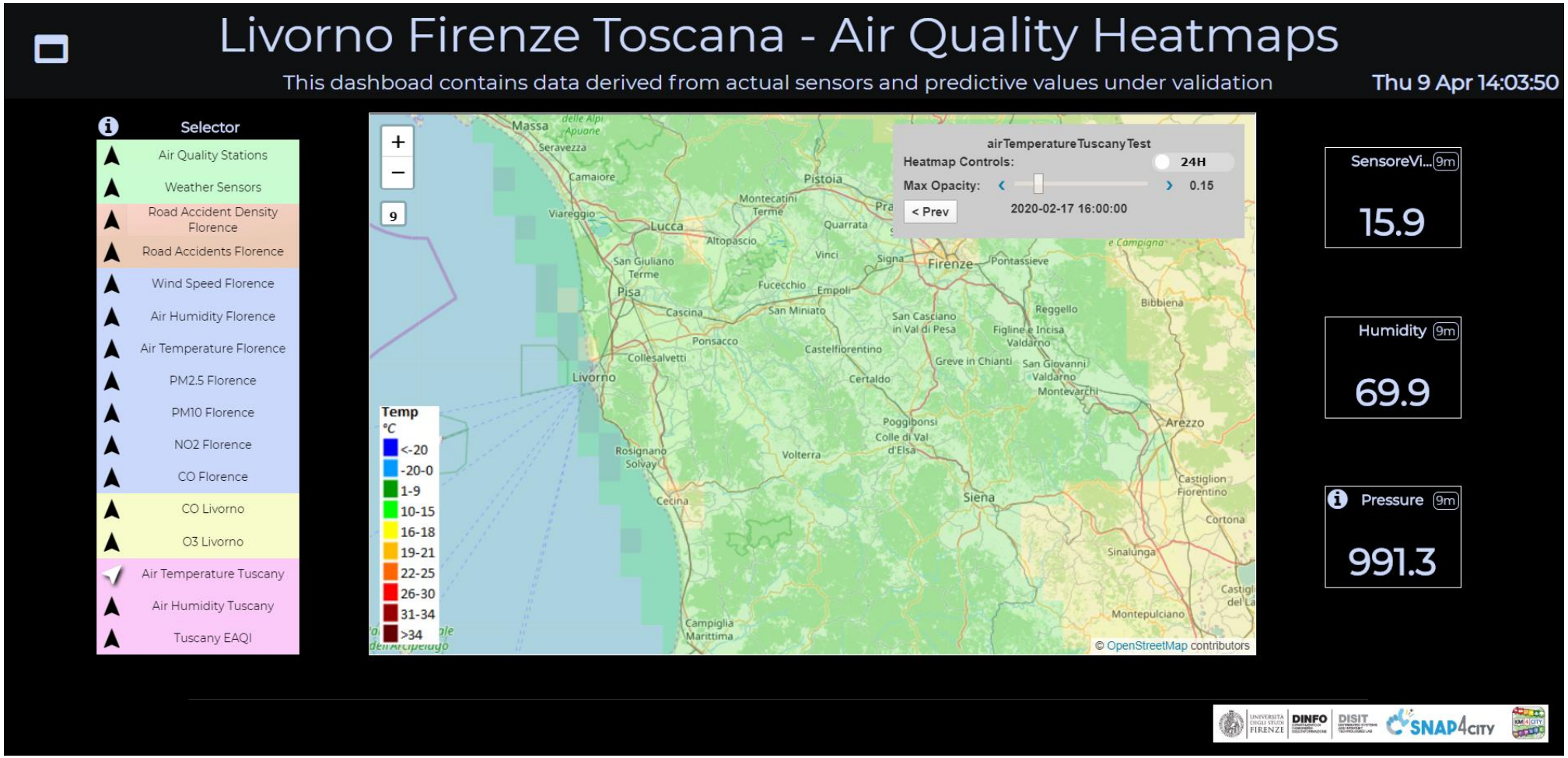
3. Change the Query to visualize the new heatmap



https://wmsserver.snap4city.org/geoserver/Snap4City/wms?service=WMS&layers=heatmapName

https://wmsserver.snap4city.org/geoserver/Snap4City/wms?service=WMS&layers=airTemperatureTuscanyTest

Heatmap Visualization

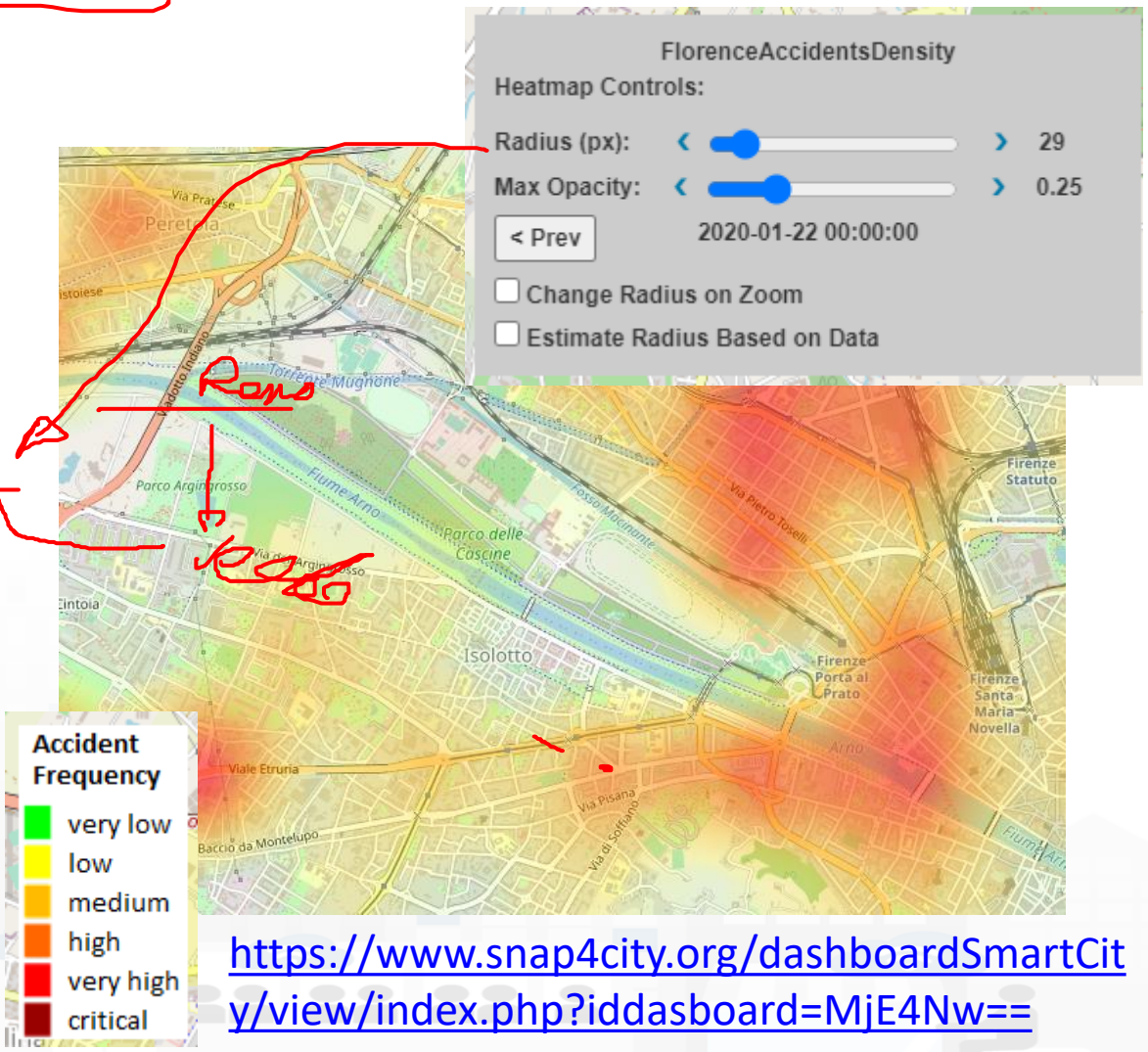


https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MTI2OA==

Different Heatmap Models
A comparison and features

Gaussian vs WMS heatmap settings/calls

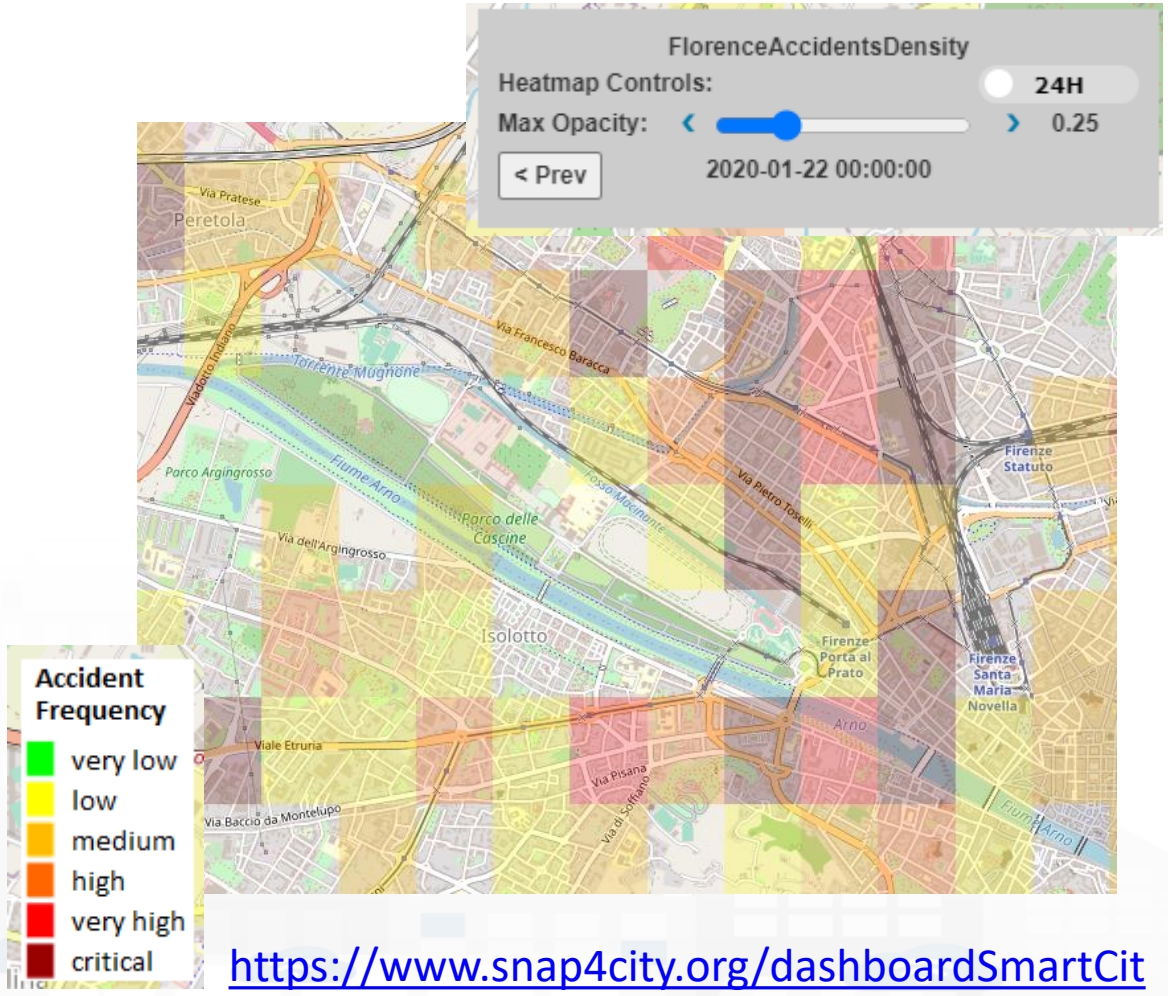
- https://heatmap.snap4city.org/heatmap.php?dataset=15MinIndex_HousingIndex
  - GPS coordinates (points)
  - Metrics to be defined into the Dashboard table
  - Legenda of colormap in PNG to be uploaded if not standard
  - Heatmap non calibrated, created on client. Data provided from Heatmap server, some limitations on the number of points since the heatmap is created on client side
- If data are on Heatmap Server, the data Piker from Heatmap is accessible



https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MjE4Nw==

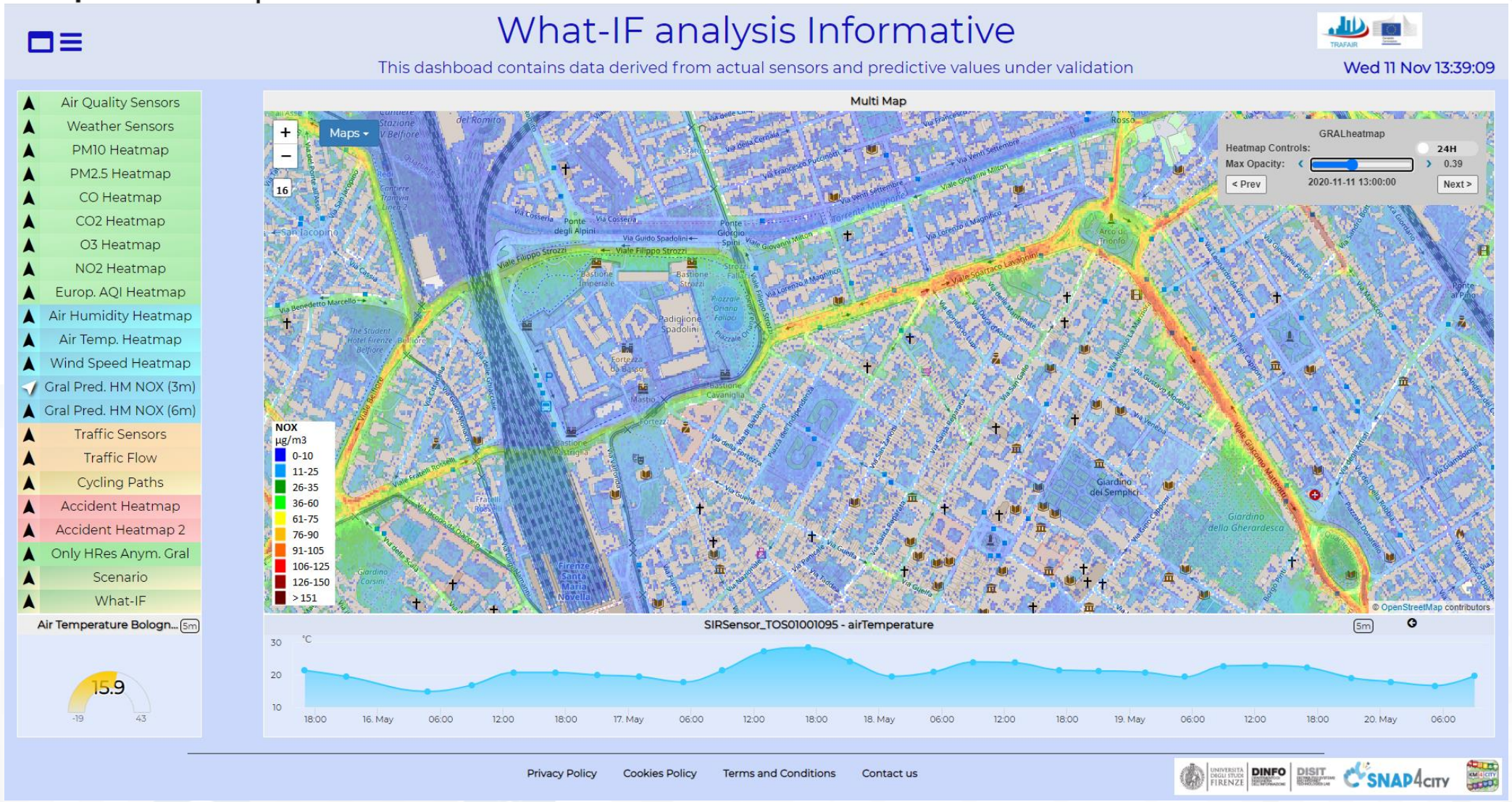
Gaussian vs WMS heatmap settings/calls

- https://wmsserver.snap4city.org/geoserver/Snap4City/wms?service=WMS&layers=15MinIndex_HealthIndex
  - UTM coordinates (points, grid size, etc.)
  - Legenda of Colormap in PNG to be uploaded if not standard
  - Heatmap built as Tiled Images in GeoTIFF and provided from GeoServer
- It is possible to create Heatmap on GeoTiff without loading data on Heatmap Server



https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MjE4Nw==

WMS heatmap



https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MjE4Nw==

R studio Development documentation (self training)

https://www.snap4city.org/dashboardSmartCity/management/iframeApp.php?linkUrl=https%3A%2F%2Fwww.snap4city.org%2Fdrupal%2Fnode%2F25&linkId=25link&pageTitle=Doc:%20R%20Studio%20Development&fromSubmenu=handddocLink

- TC7.1. Exploiting data analytics and machine learning in IOT Applications as MicroService
- TC7.2. R Studio for Analytics, exploiting Tensor Flow
- TC7.3. Download data from AMMA (Application and MicroService Monitor and Analyser), ResDash (Resource Dashboard) and DevDash (Development Dashboard) tools
- TC7.4. From R Studio process to MicroService for IOT application, data analytics, machine learning
- TC7.5. Developing Data Analytics Processes
- TC7.6. How to get data from API into R studio
- TC7.7. How to Save resulting data via API from R studio
- TC7.8. Example of how to CreateLastValuesMean.R
- TC7.9. CreateHourlyAvgTrendPerDay.R
- TC7.10. CreateHeatmap.R
- TC2.31 - Create Data Analytic Flow
- TC2.32 - Make Your Data Analytic Flow Public

GeoTIFF files

How it works: HeatMap Manager

GeoTiff

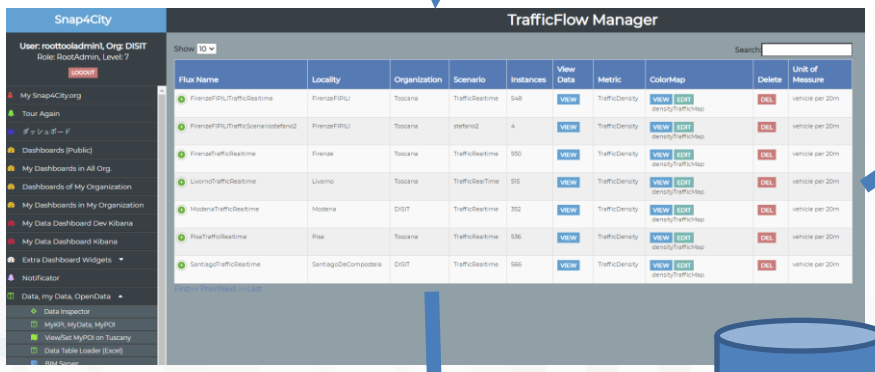
IOT App

Name
ColorMap
Unit of Meaure
…

https://www.km4city.org/swagger/external/index.html?urls.primaryName=Heatmap%20API

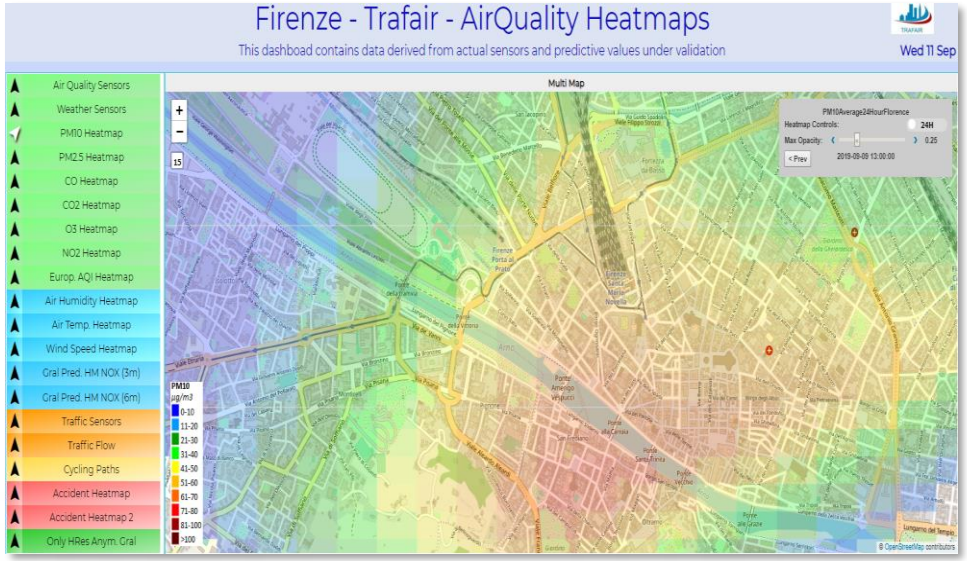
API and MicroServices

Heatmap Manager



JSON

WMS



GeoServer

Heatmap GeoTIF Generation

Helsinki

The Life of Helsinki (H5b)

Please note that the data results are not always based on real data.

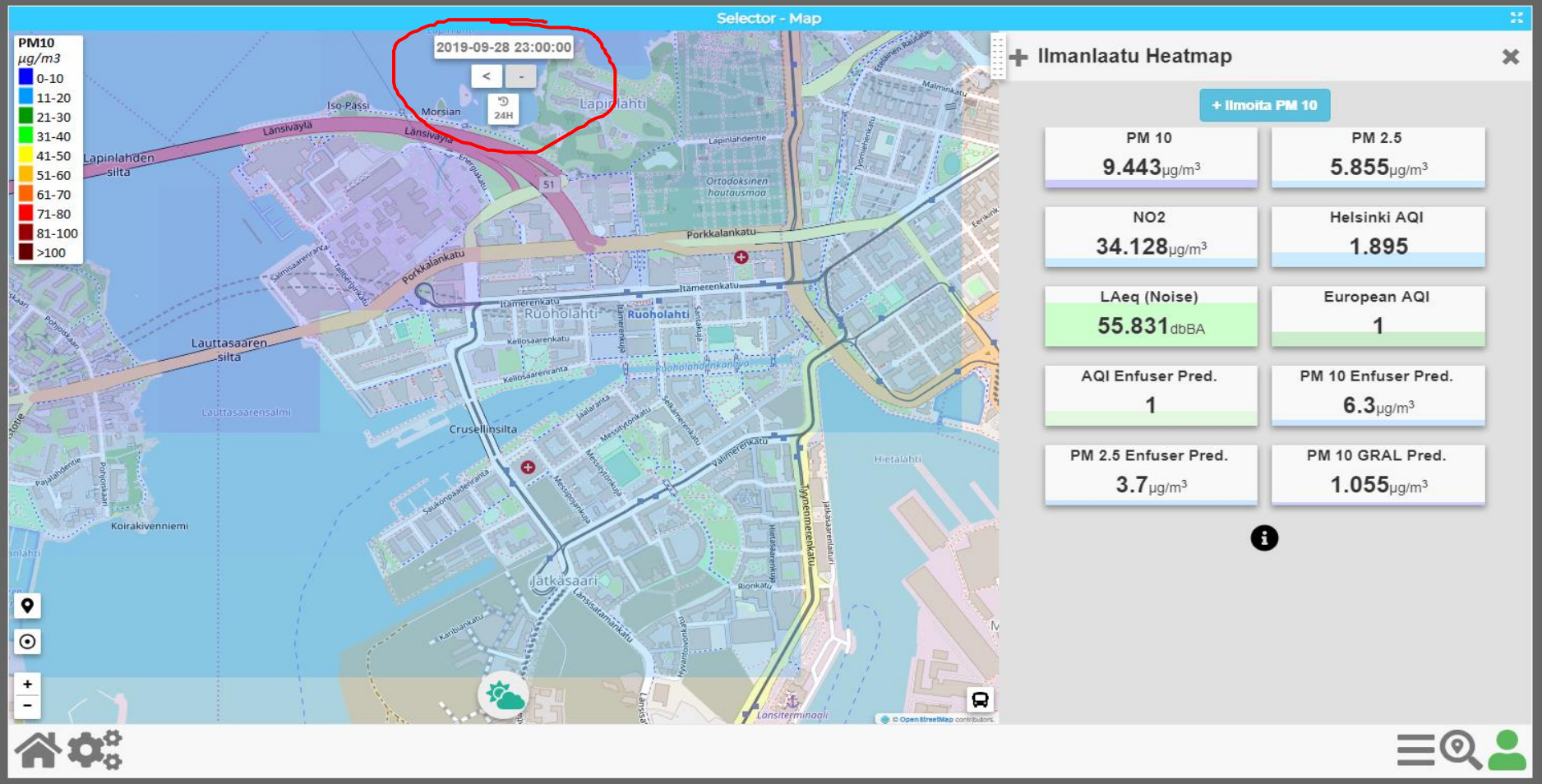
Sun 29 Sep 00:42:50

- Origin Dest. Matrix
- Typical Trajectories
- Twitter Vigilance
- Twitter Vig. Real Time
- Entertainment Events
- Shopping: POI
- Wine and Food: POI
- Discovery Helsinki
- Points of Interest
- 3D view POI
- Routing on Helsinki
- Line of Transport
- Public Transport
- Air Quality
- Air Quality Jätkäsaari
- Weather
- Forum Discussion

Documentation

Survey

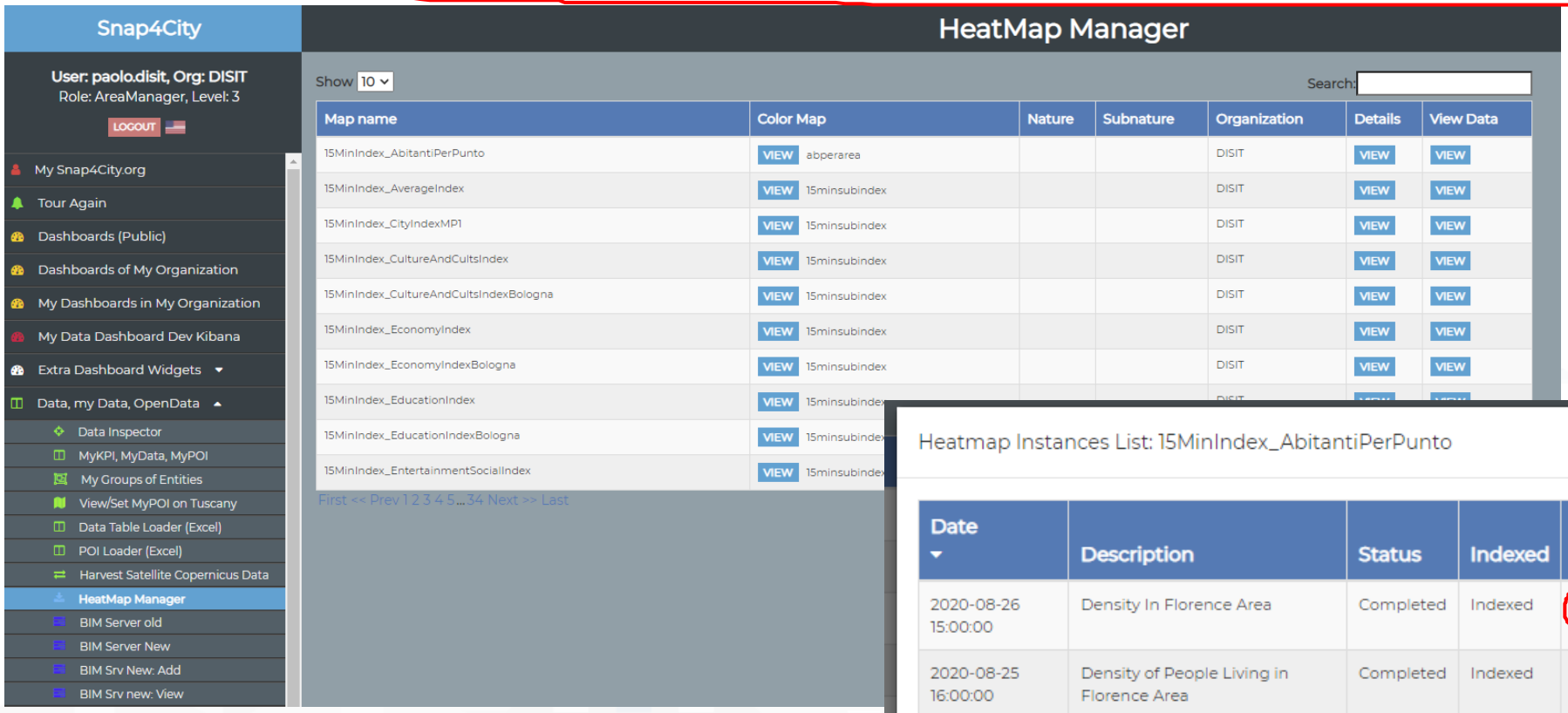
Environment



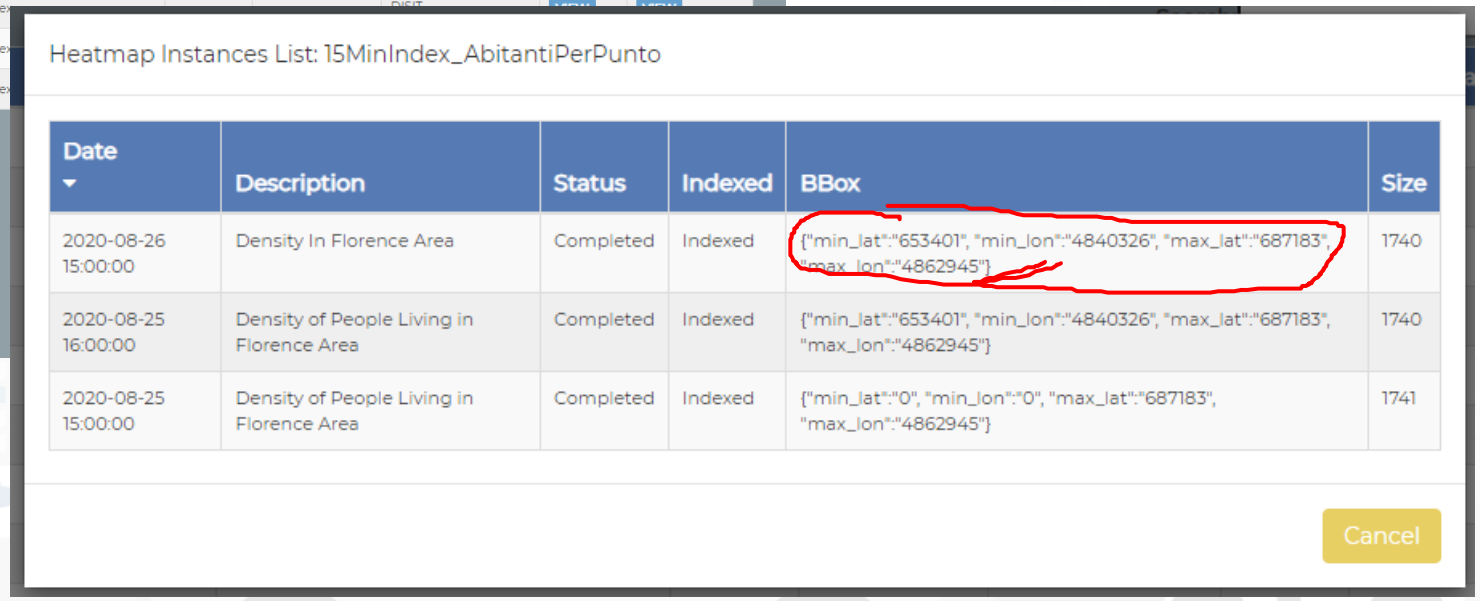
Privacy Policy Cookies Policy Terms and Conditions Contact us

https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MTc1Mg==

HeatMap Manager (Area Manager view)



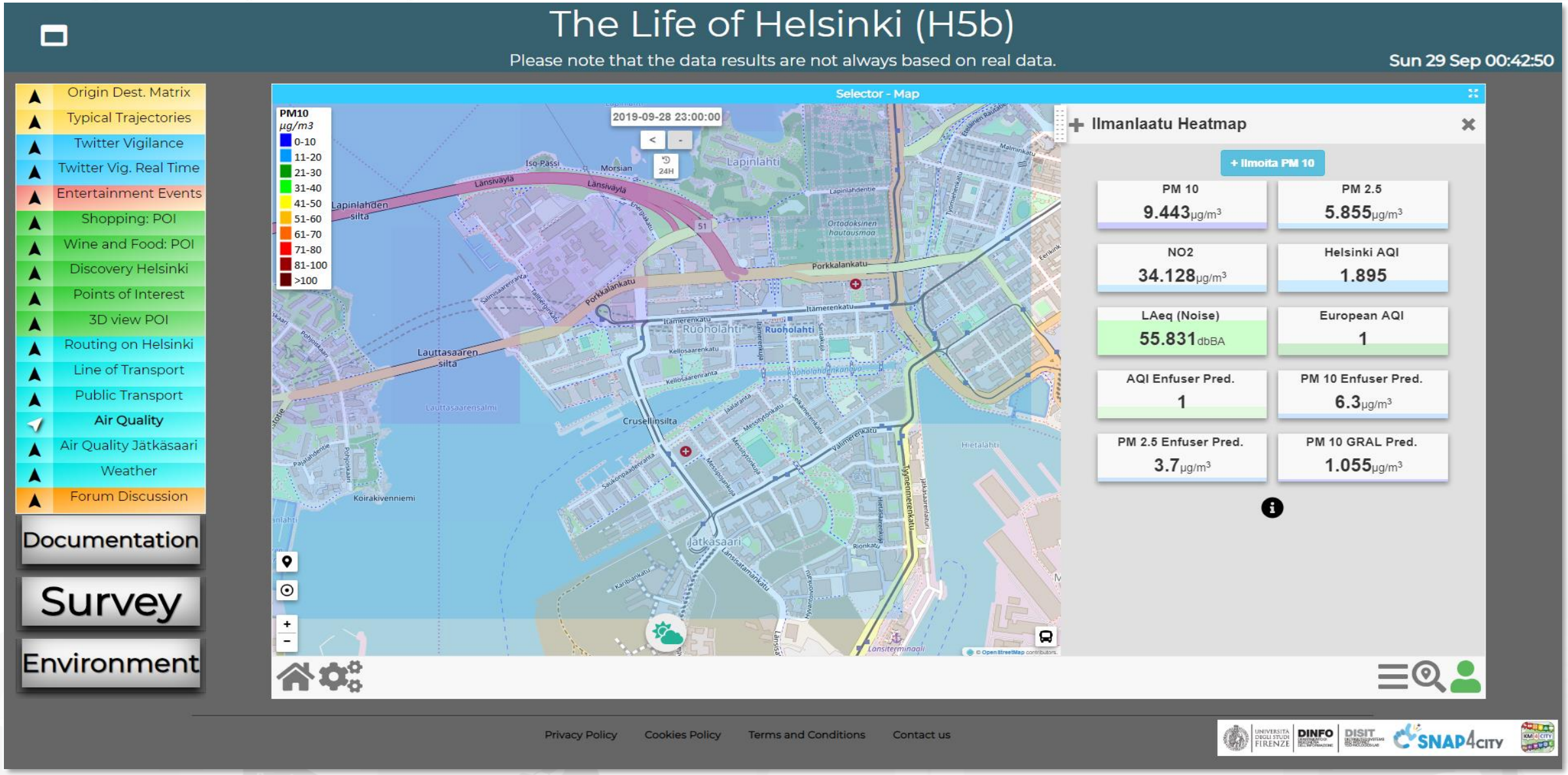
- Sequence of Heatmaps
- Colormap used
- Details



Editing Mode for RootAdmin only

Dynamic Heatmap Exploitation on the Front-End Tools

Helsinki



https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MTc1Mg==

Firenze - Trafair - AirQuality Heatmaps

This dashboad contains data derived from actual sensors and predictive values under validation

Sun 9 Jun 17:41:58

- Air Quality Sensors
- PM10 Heatmap
- PM2.5 Heatmap
- CO Heatmap
- CO2 Heatmap
- SO2 Heatmap
- O3 Heatmap
- NO2 Heatmap
- Benzene Heatmap
- H2S Heatmap
- Air Humidity Heatmap
- Air Temp. Heatmap
- Wind Speed Heatmap
- Gral Pred. HM NOX (3m)
- Gral Pred. HM NOX (6m)
- Traffic Sensors
- Traffic Flow
- Cycling Paths
- Accident Heatmap
- Accident Heatmap 2

Multi Map

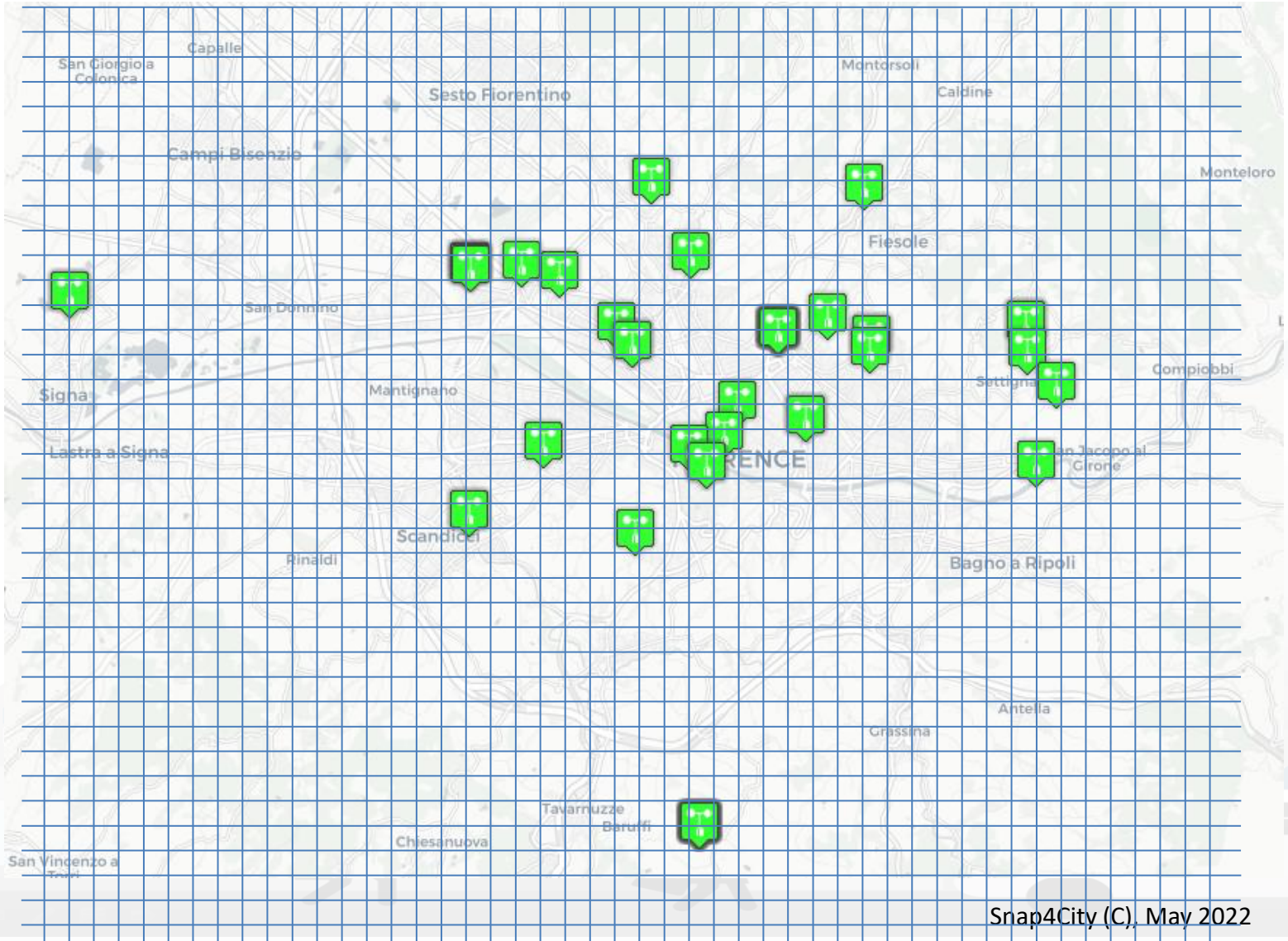
Air Temperature Bolognese 7m

SIRSensor_TOS01001095 - airTemperature 7m

Services

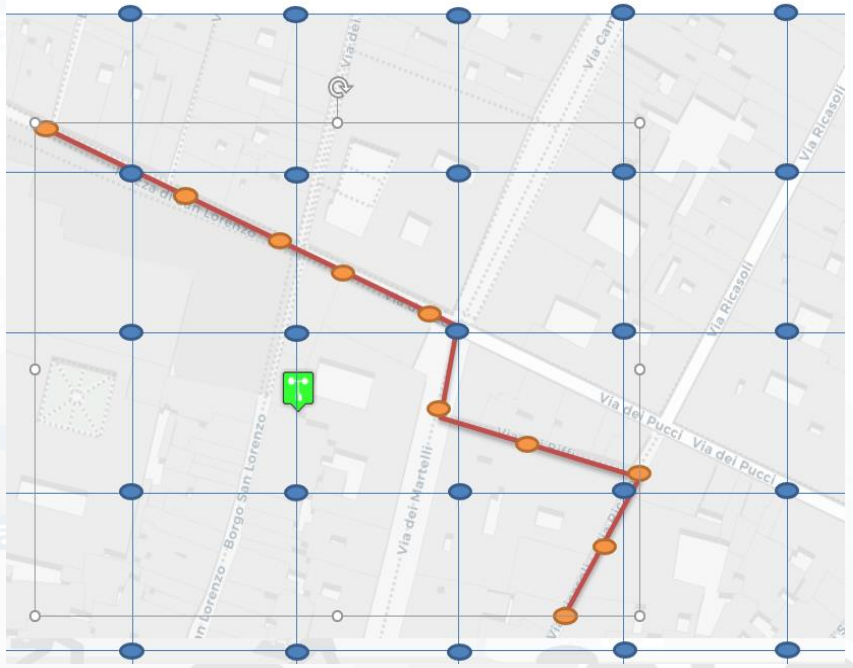
- **Air Quality sensors are**
  - Collected on scattered positions
- **AirQuality Services**
  - AirQuality indicators independent on the sensors' position, in any GPS position of the area
  - **Multiple data:** PM_{10}, $PM_{2.5}$, CO, CO_2, SO_2, O_3, H_2S, NO, NO_2, NO_x, air temperature, air humidity, velocity of wind speed, dew point, etc.
- **Applications**
  - Alerting on specific personal GPS locations
  - Constrained routing for: runners, walking with baby, people with pulmonary problems,
  - Control Room Rendering
  - Mobile Phone Rendering, this means to have thousands of users active at the same time, and a reasonable memory consumption in the server.

The GRID density is never enough



4x4 meters grid is really too expensive

1000x1000 area (small town)
4x4mt * 10 variables * 24 hours per day
→ 3.8 Billions of data



Workflow

Scattered IoT sensor Data of Value {V}

IOT App

Computation of Data Value on Grid

Data Grid {V'} of {V}

Heatmap Data Server

Area

V'' values on GPS

On Demand Fine Interpolation

Heatmap GeoTIFF Generation

GeoTIF of {V'}

Point

Points

IOT App

Alert Assessment

Conditioned Routing

GeoServer

Point

Alerts

Points

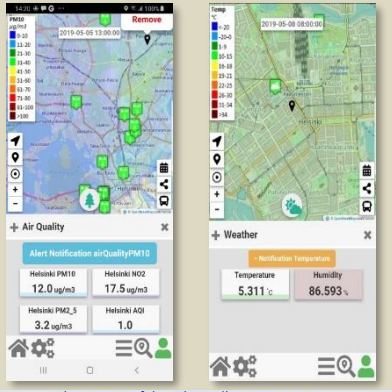
Paths

Area

WMS

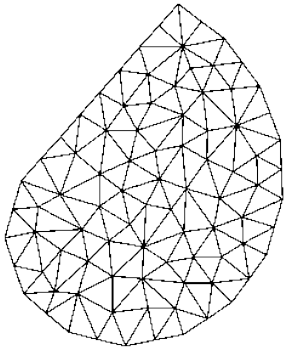
heatmaps

Web and Mobile Apps



Snap4City Heatmap Server

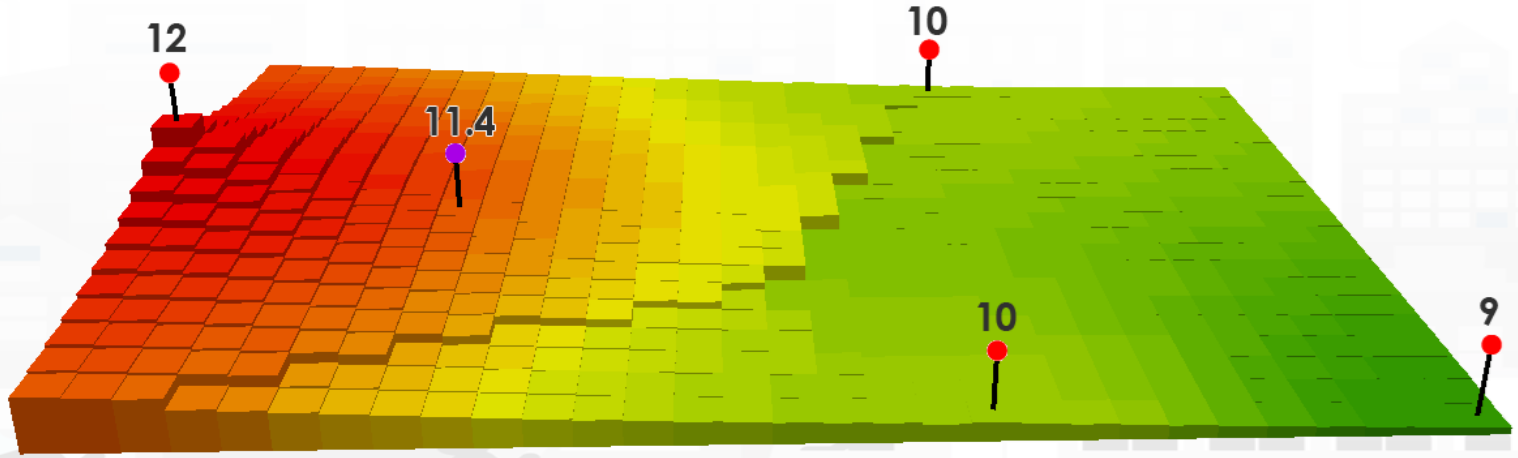
Bivariate Interpolation Method [Akima]



- Steps for irregular data
  - *triangulation* (i.e., partitioning of the area into a number of triangles) of the *x*-*y* plane
  - *selection* of several data points that are closest to each data point (sensor) and are used for estimating the partial derivatives;
  - *organization* of the resulting data with respect to triangle numbers;
  - *estimation* of partial derivatives at each data point;
  - *computation of* the interpolation at each output point.

Inverse Distance Weighting, IDW Method

- It is a deterministic mathematical method widely used in the geoscience.
  - the interpolated value at the location (x, y); z_i is the observed value; d_i is the Euclidean distance between the point i and the interpolated point; and w_i is the weight for the point each point (x_i, y_i) and (x, y)



Validation via Error Estimation

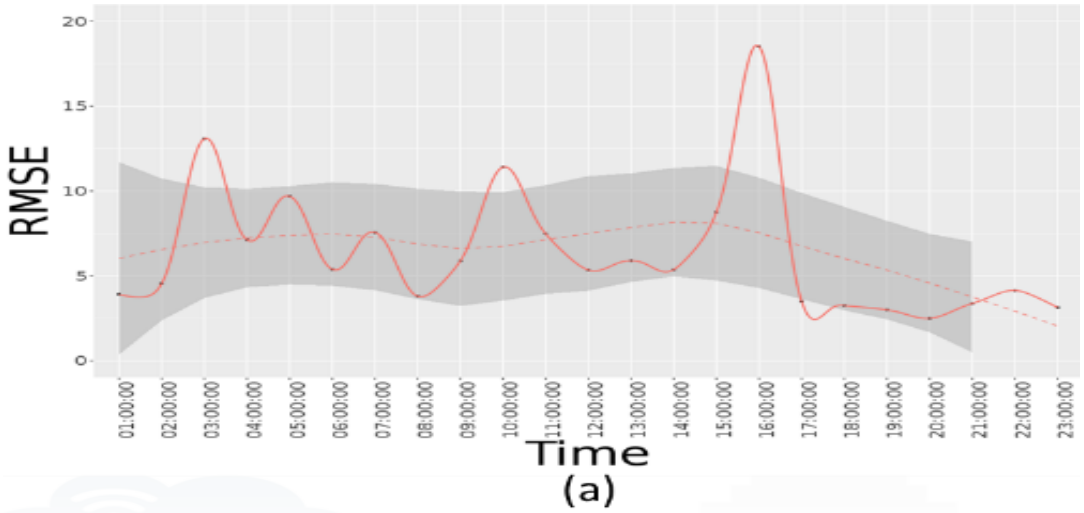
- alternate exclusion of selected air quality sensor in contributing to the model and using the excluded as true value for validation in that point on the basis of the estimation performed exploiting all the others.

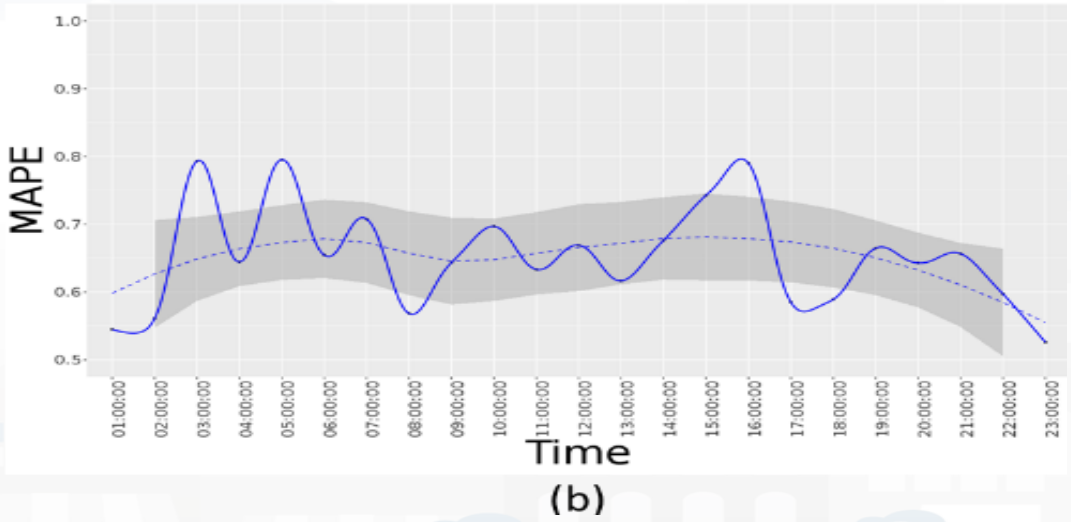
Error Measures	Akima	IDW
MAPE	**0.69**	0.79
RMSE	**8.90**	12.20
MAPE-we	**0.60**	0.95
MAPE-wd	**0.70**	0.93
RMSE-we	**8.60**	10.70
RMSE-wd	**9.70**	17.00

mean absolute percentage error (MAPE)

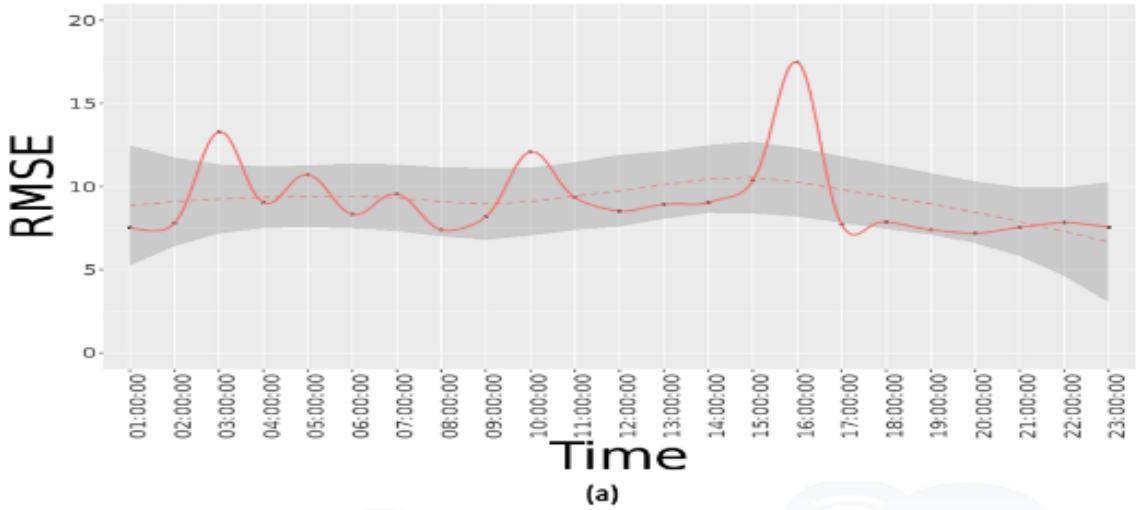
root mean squared error (RMSE)

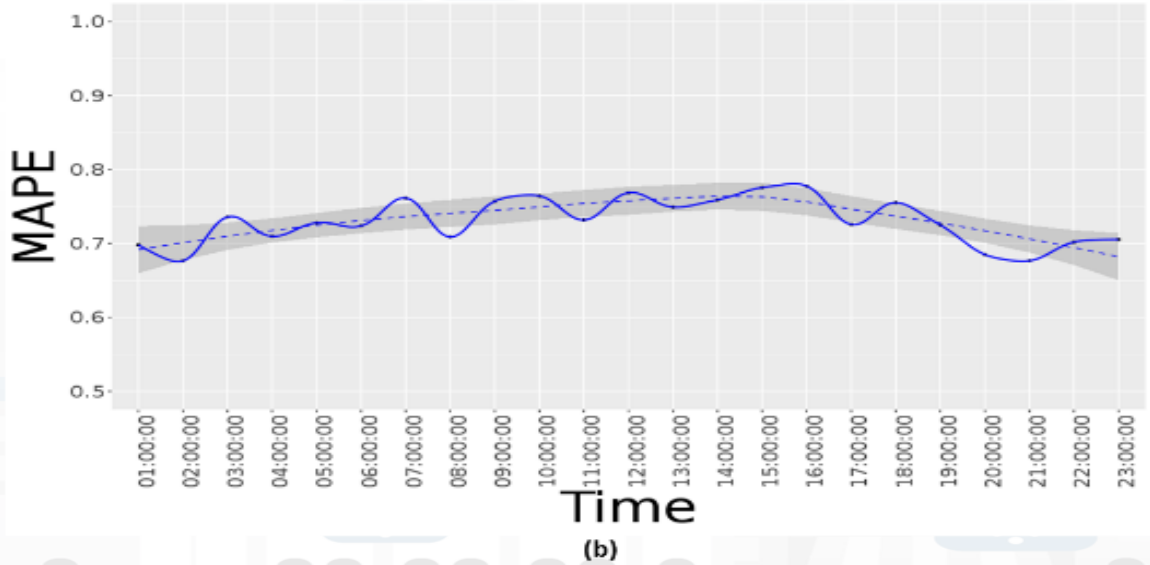
Error Trends





PM_{10}working days RMSE (a) and MAPE (b) per time slots (Akima Method)

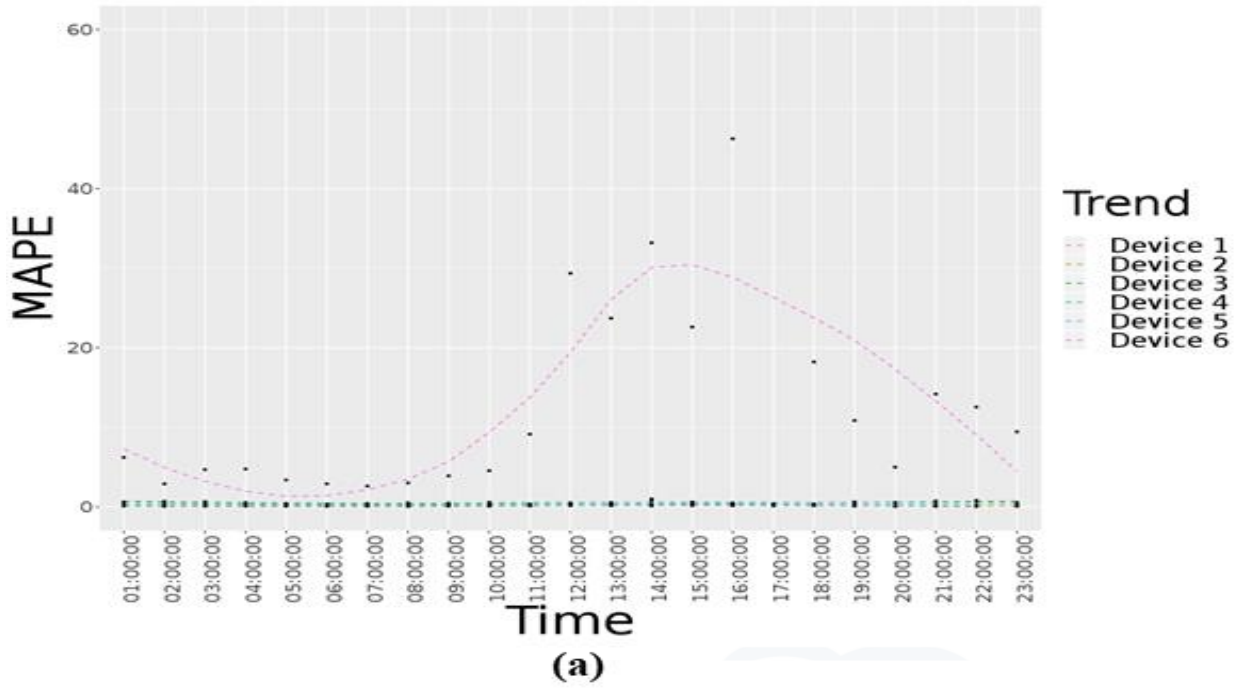




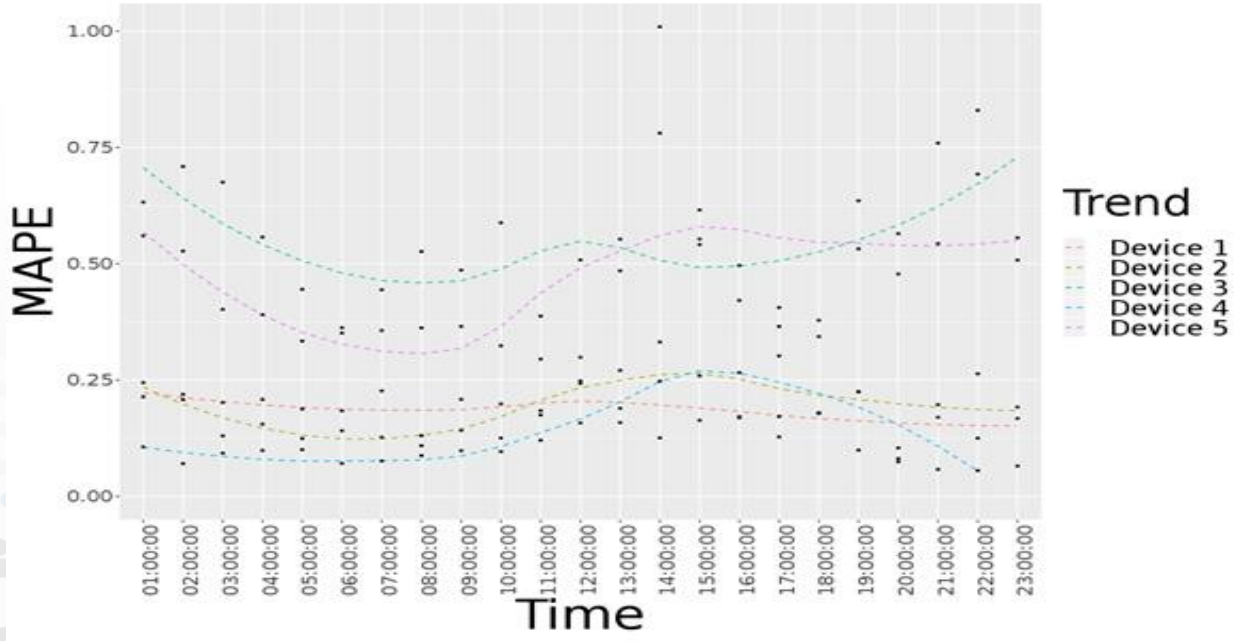
PM_{10}working days RMSE (a) and MAPE (b) per time slots (IDW Method)

Detecting dysfunction on devices using error detection

- Air Quality PM_{10} working days interpolation error trends per hour in terms of mean absolute percentage error for
  - (a) six personal devices including the device with a dysfunction;
  - (b) five personal devices



(a)



(b)

Considerations

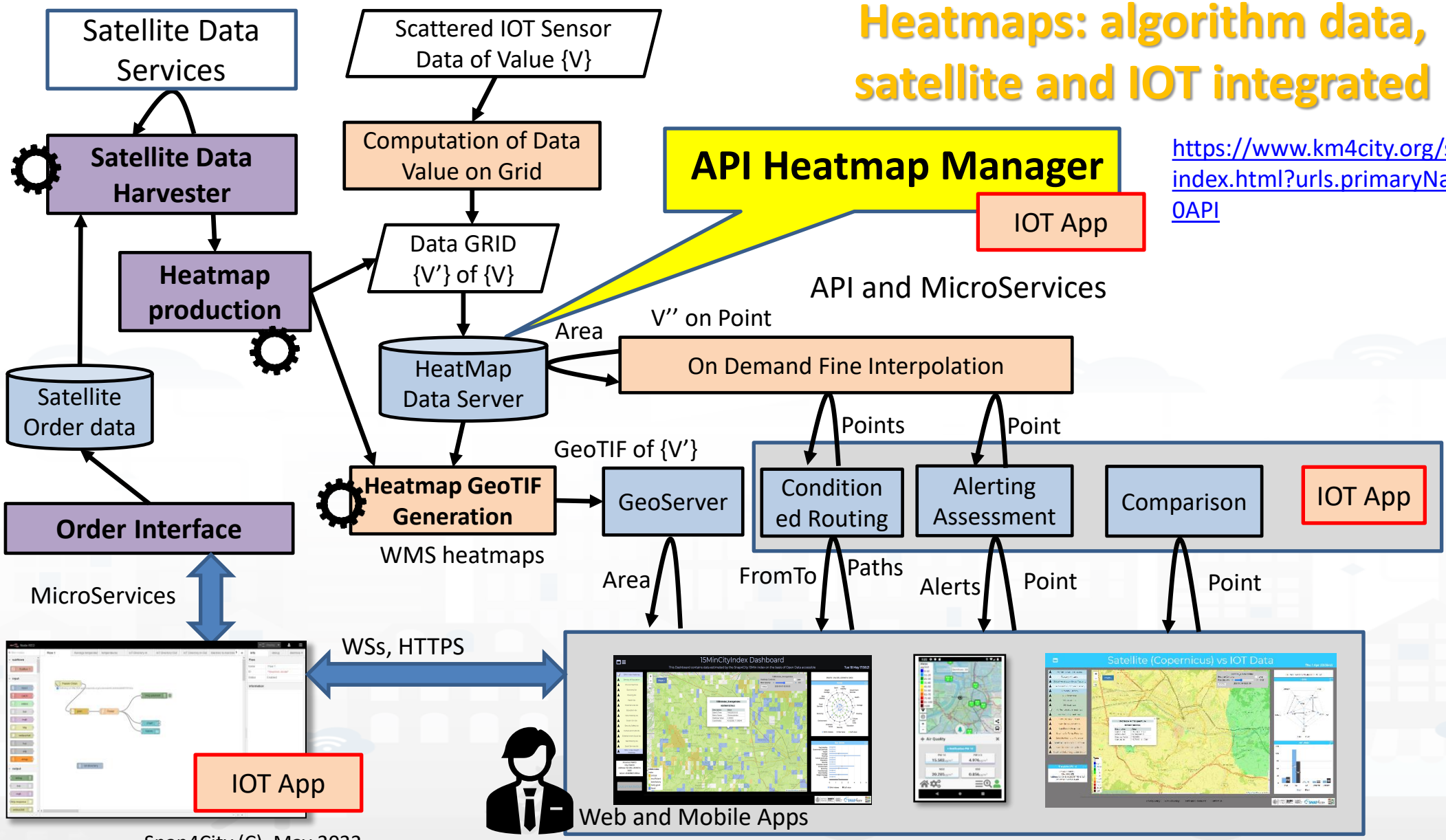
- In order to **satisfy the requirements** reported above:
  - Provide Sensor value in any GPS point, for implementing alerts and other applications (routing), rendering on mobile and control room web pages
- What:
  - Two methods have been implemented
  - A scalable architecture has been defined and implemented to provide these services to several thousands of users
- The selection of the best method has been performed on the basis of and error assessment in which Akima solution has been better ranked.
- The Solution can be also used for detecting eventual dysfunction of specific IOT Devices in the same area, for example for bad positioning, turned off, etc.

UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB

High Density Real-Time Air Quality Derived Services from IoT Networks

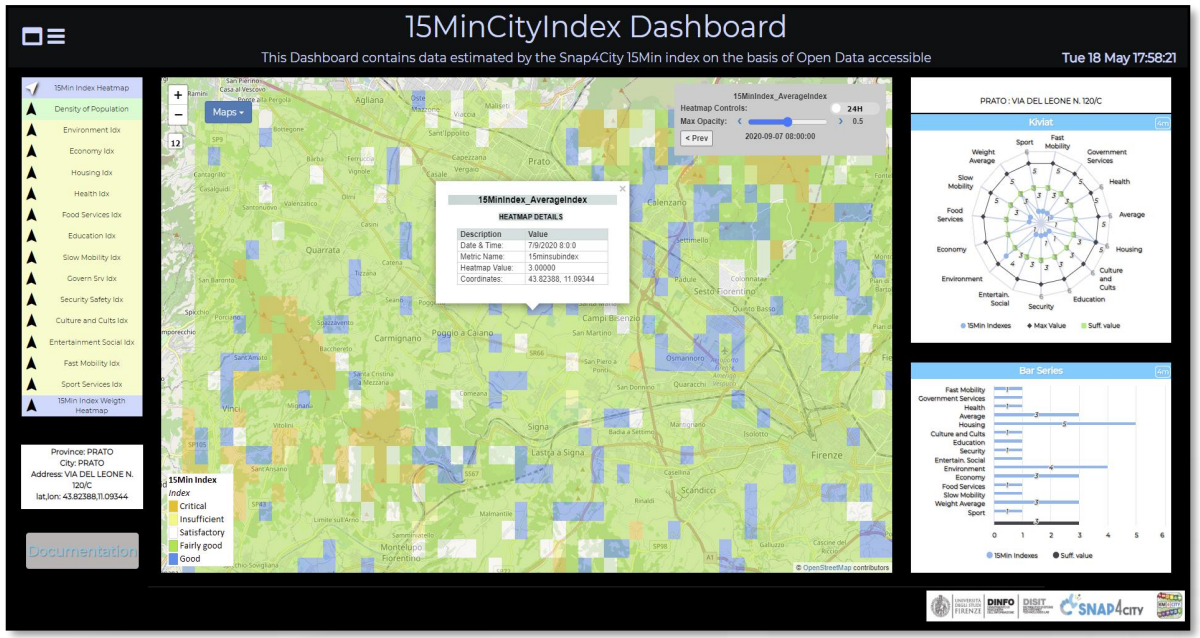
- C. Badii, S. Bilotta, D. Cenni, A. Difino, P. Nesi, I. Paoli, M. Paolucci, Sensors, Vol.20, 2020, N.18, ISSN 1424-8220
- DOI 10.3390/s20185435
- https://www.mdpi.com/1424-8220/20/18/5435/pdf

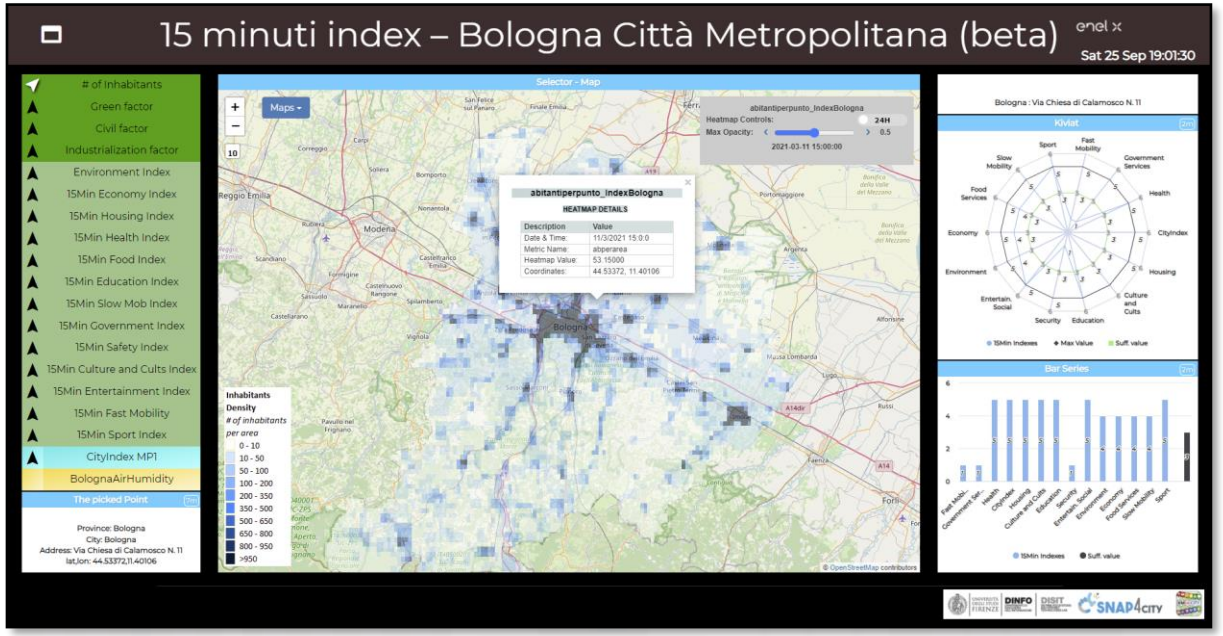
Heatmaps: algorithm data, satellite and IOT integrated



https://www.km4city.org/swagger/external/index.html?urls.primaryName=Heatmap%20API

15MinCityIndex





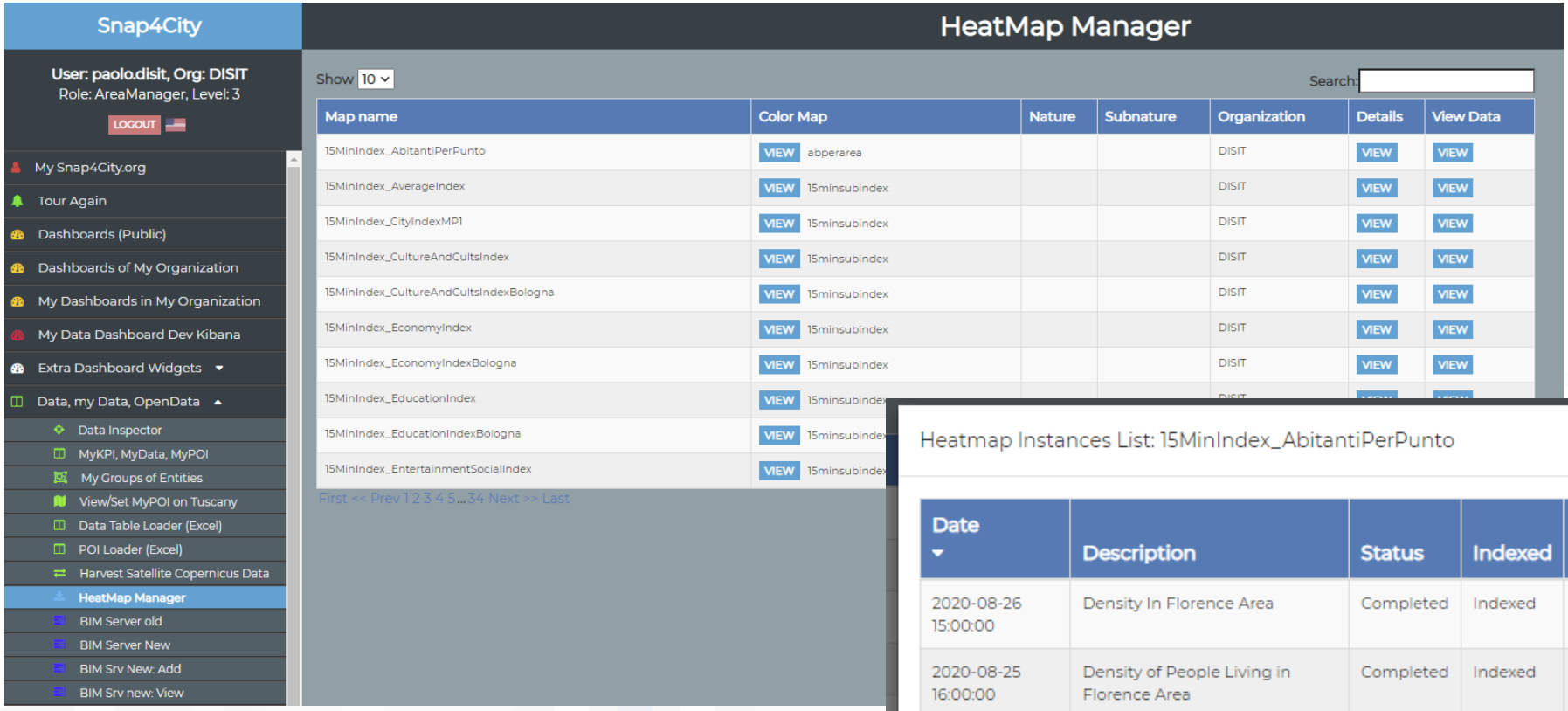
[FLORENCE metro city](https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MjkzOA=)

https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MjkzOA=

[Bologna metro city](https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MzA1OQ==)

https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MzA1OQ==

HeatMap Manager (Area Manager view)



- Sequence of Heatmaps
- Colormap used
- Details

Editing Mode for RootAdmin only

TOP

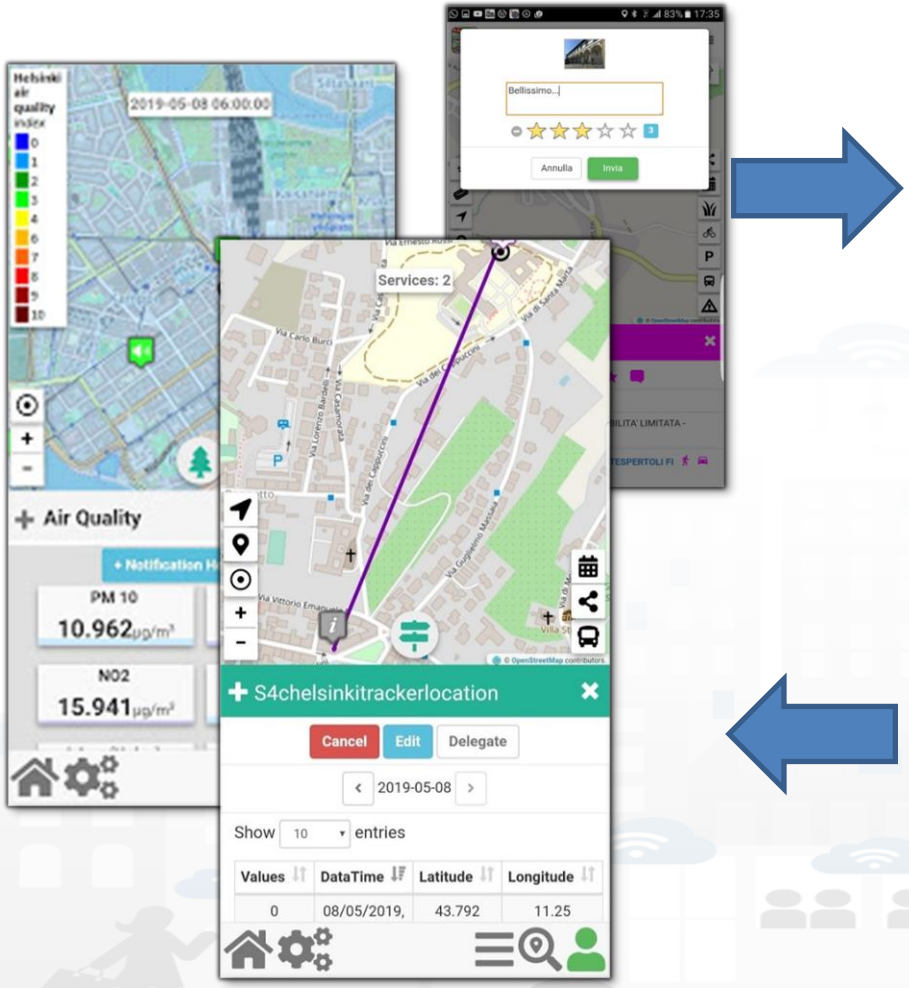
Engaging City Users toward a Virtuous Behaviours (real time)

The App is a Bidirectional Device

- GPS Positions
- Selections on menus
- Views of POI
- Access to Dashboards
- searched information
- Routing
- Ranks, votes
- Comments
- Images
- Subscriptions to notifications
- ….

Produced information

- Accepted ?
- Performed ?
- …



Derived information

- Trajectories
- Hot Places by click and by move
- Origin destination matrices
- Most interested topics
- Most interested POI
- Delegation and relationships
- Accesses to Dashboards
- **Cumulated Scores from Actions**
- Requested information
- Routing performed
- …..

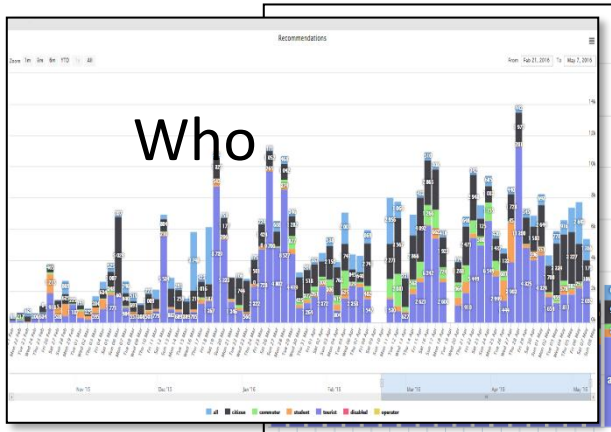
Produced information

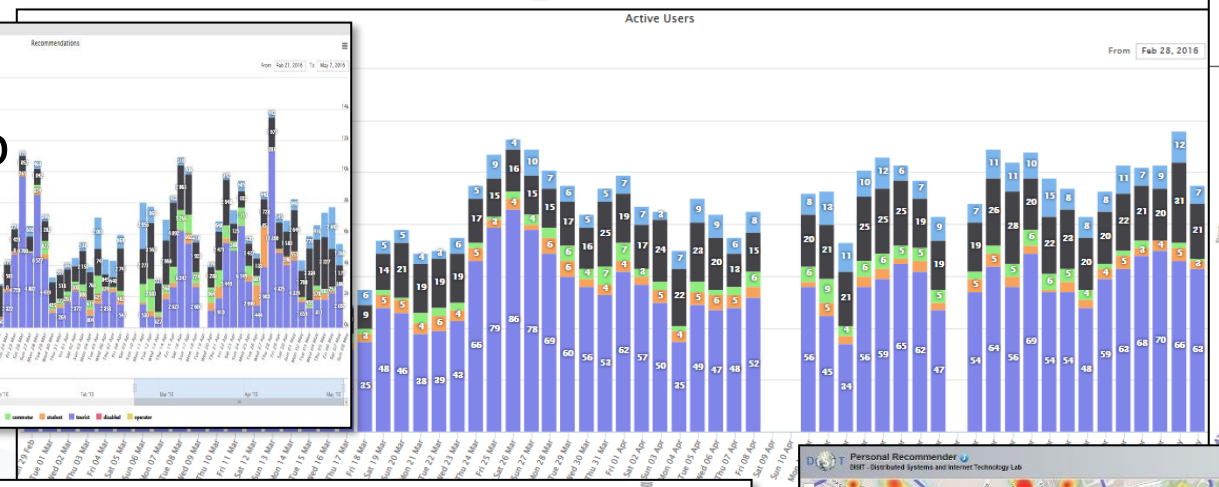
- Suggestions
- Engagements
- Notifications
- …

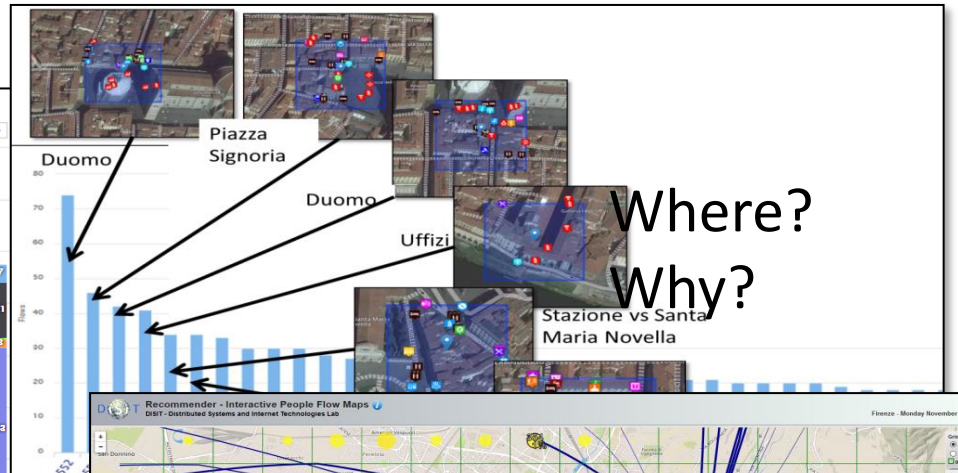
Users

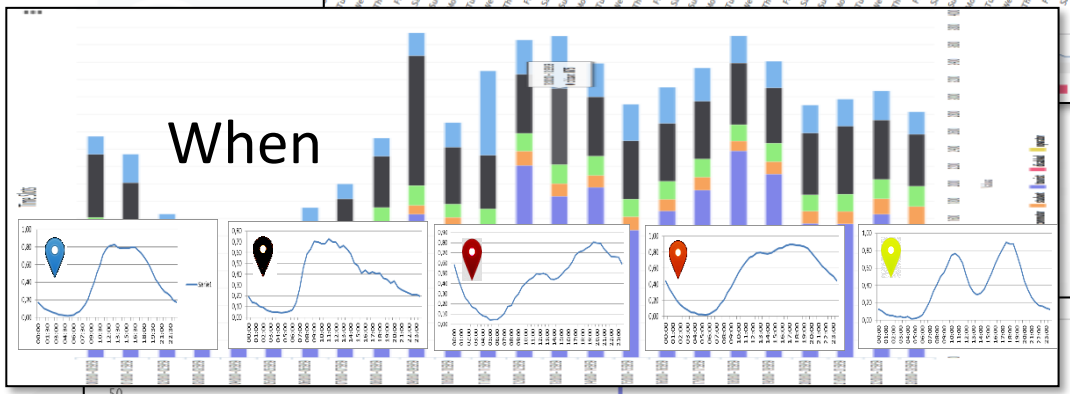
System

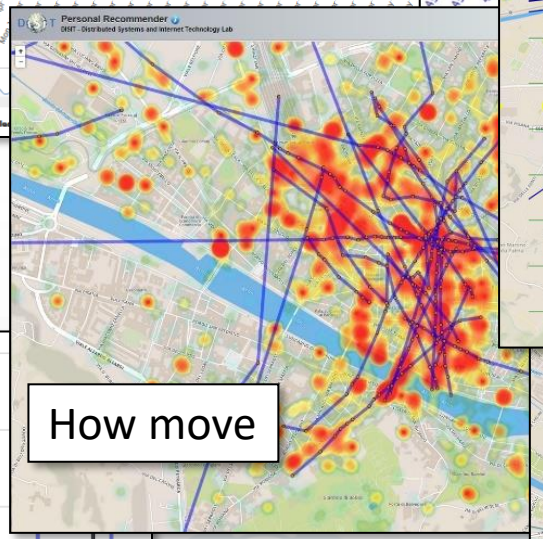
User Behavior Analyser for Collective Profiling

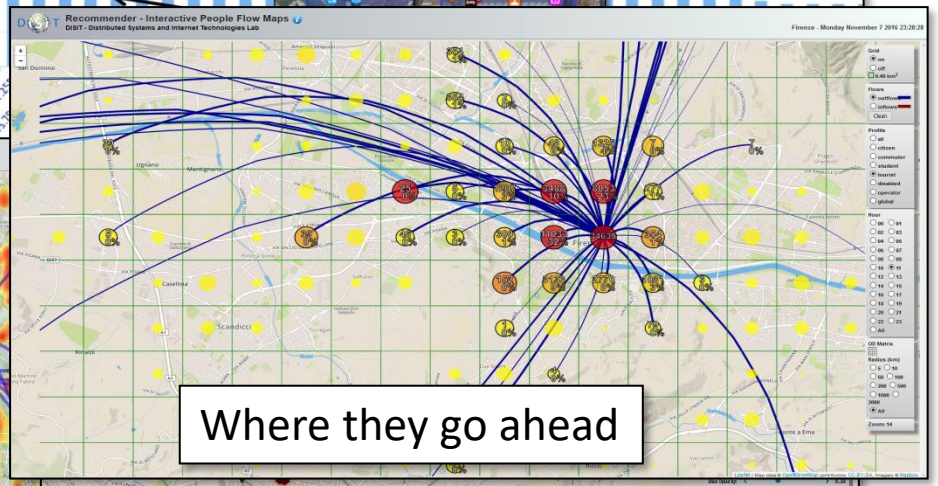


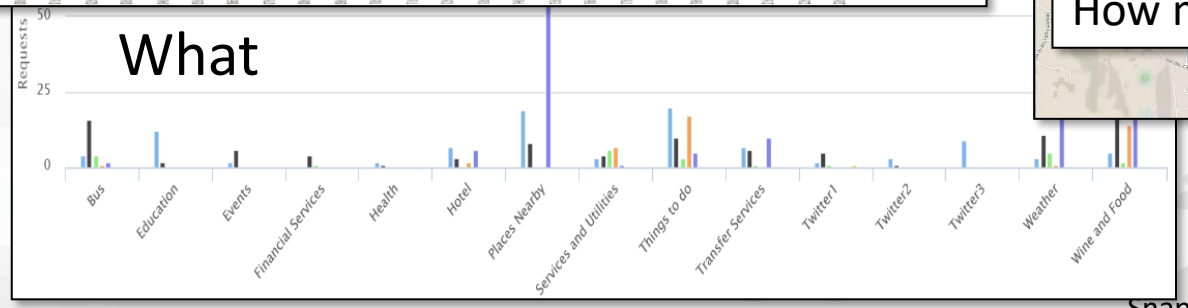


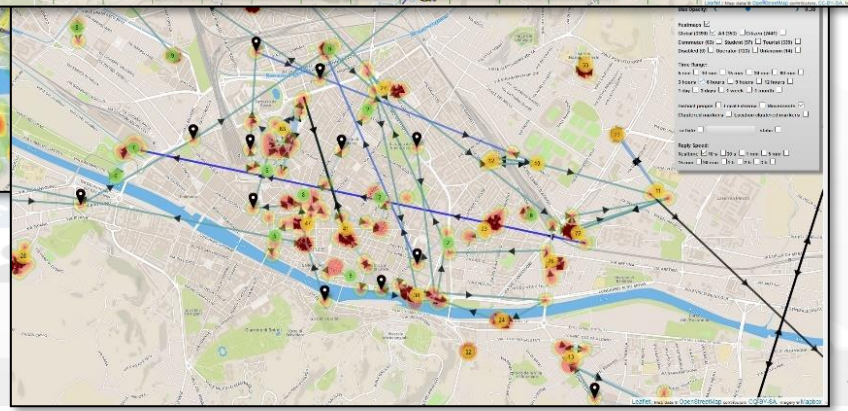






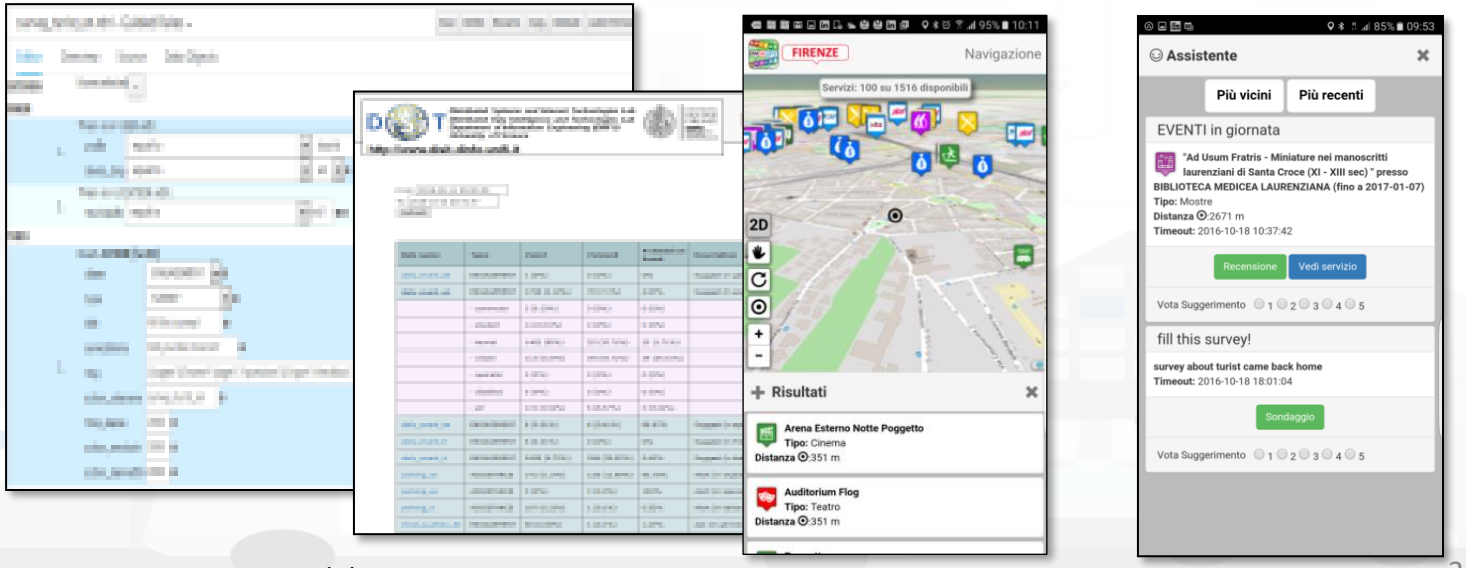




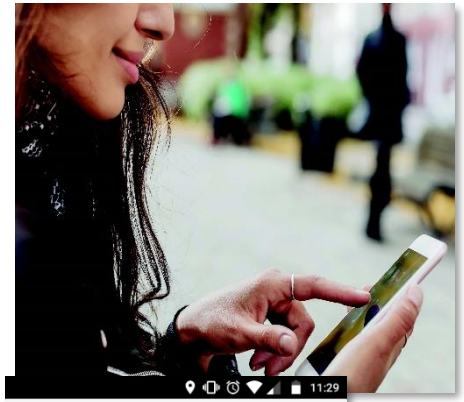


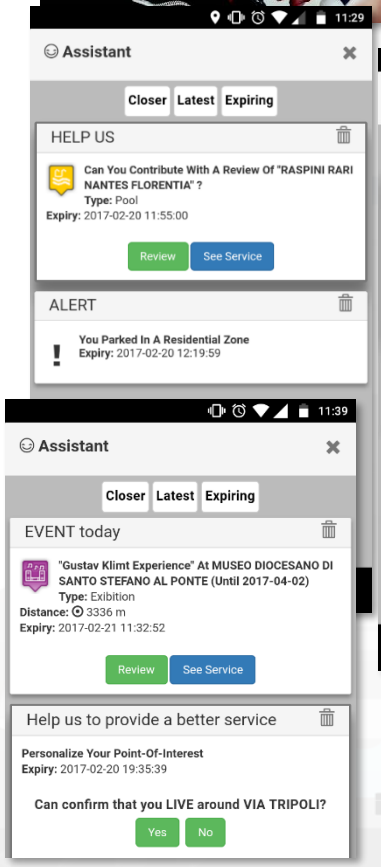
Profiled Engagements to City Users

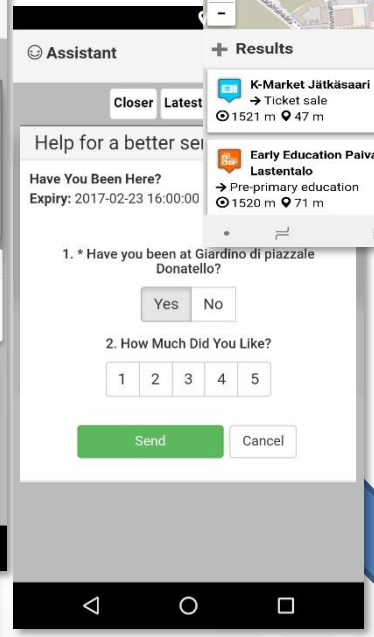
- The users are profiled to learn habits:
  - Personal POI, paths, Mobility habits
- Information and engagements sent to the users are programmed according to the context and user behavior to:
  - Stimulate virtuous habits
  - More sustainable habits
  - More healthy habits, etc.
  - Get feedbacks
  - Provide bonus and prices, …..
  - Send alerts, ….



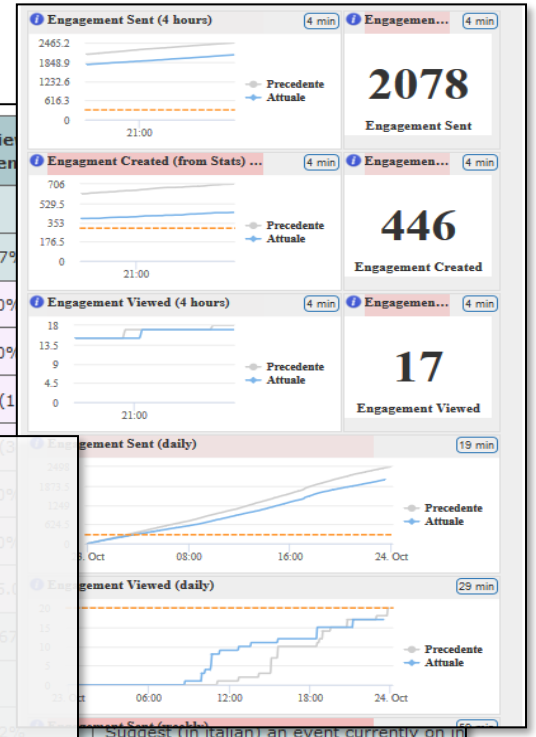
Users' Engagement







Rule name	Type	#sent	#viewed	#vie #sen
daily_event_de	ENGAGEMENT	1 (0%)	0 (0%)	0%
daily_event_en	ENGAGEMENT	1720 (2.12%)	70 (7.1%)	4.07%
	- commuter	5 (0.29%)	0 (0%)	0 (0%
	- student	14 (0.81%)	0 (0%)	0 (0%
	- tourist	1462 (85%)	25 (35.71%)	25 (1



Suggest (in italian) an event currently on in
Alert (in english) if the user parked in a res
Alert (in spanish) if the user parked in a res
Alert (in italian) if the user parked in a resid
Ask (in german) a contribution for a nearby

Inform
- Air Quality forecast is not very nice
- You have parked out of your residential parking zone
- The Road cleaning is this night
- The waste in S.Andreas Road is full

Engage
- Provide a comment, a score, etc.

Stimulate / recommend
- Events in the city, services you may be interested, etc..

Provide Bonus, rewards if needed
- you get a bonus since you parked here
- We suggest: leave the car out of the city, this bonus can be used to by a bus ticket

Rules

User context

City context

Engaging City Users

- **Mobile Applications** can use Advanced Smart City API to collect data about the city usage by the city users via a signed consent
- It can be used for sending engagements to them such as to:
  - **Inform**
    - You have parked out of your residential parking zone
    - The Road cleaning is this night
    - The waste in S.Andreas Road is full
  - **Engage**
    - Please Provide a comment, a score, etc.
  - **Stimulate / recommend**
    - Events in the city, services you may be interested, etc..
  - **Provide Bonus**
    - Since you have parked here you can get 1 Bonus
    - We suggest you to leave the car out of the city, this bonus can be used to buy a bus ticket

UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB
http://www.disit.org

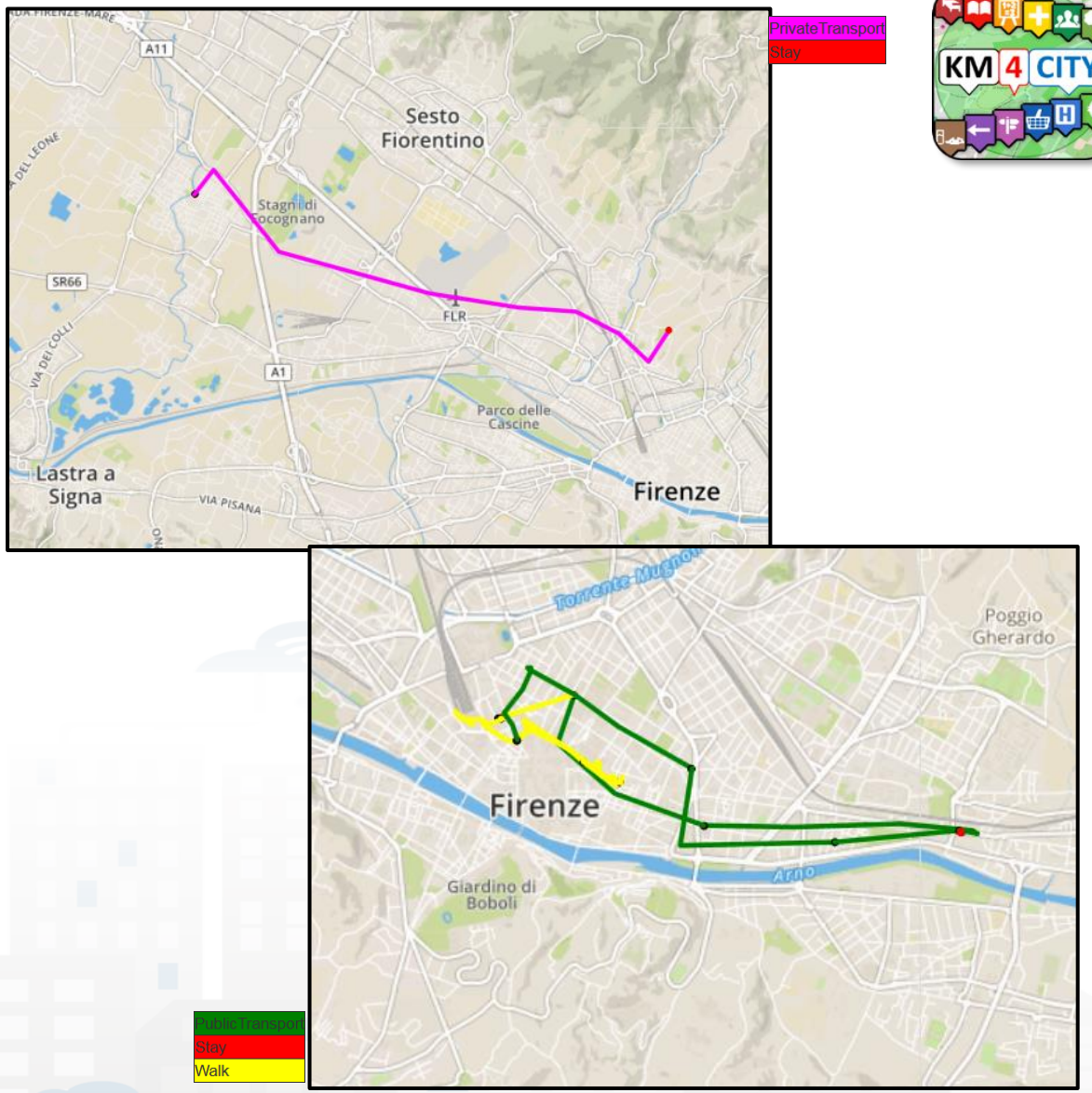
REWARDING'S RULES

- **ASSISTANCE**
- If public transport is detected after bus line suggestion on trajectory usually made on private transport → 10points
  - Why don't you take the bus line 4 in Piazza Marconi to reach your workplace? You save money, you respect the environment and you will be stress free for not worry about parking!
- Once a day, if public transport is detected after suggestion on an alternative bus line availability →3points
  - Why don't you take the bus line 4 that stop just 50 meters far from you? You save money, you respect the environment and you will be stress free for the traffic jam!
- If public transport is detected for at least 30(?) minutes a day → 1point

- **ENGAGEMENT**
- Survey on commuter and their preferred way of mobility → 1point
  - How many minutes you usually commute to go to work? How do you rate the service?
- Feedback on public transport →1point
  - Which current public transport are you using? Are the service in line with your expectation?
- Comments/Photo/Rate or survey on POI (public transport) →1point
- Survey on use of the App after N days or for tourist coming home → 1point
- Feedback on PPOI or mobility → 1point

WALLET / PROFILE

- On homepage
  - How many points have been distributed?
  - How many rewards has been already delivered?
  - How many rewards are still available?
  - How many CO2 has been saved?
  - How many km our users made this week?



CURRENT NUMBERS

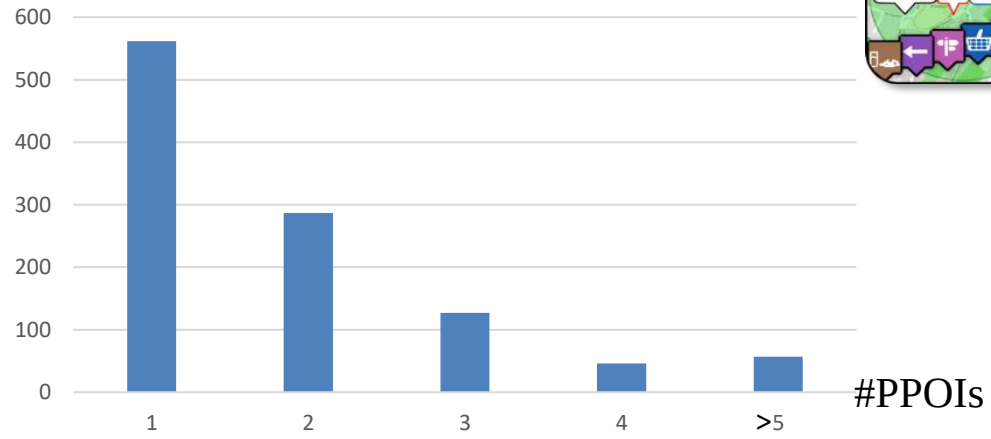
- 50 engagement's rules in 5 languages
  - (surveys, feedbacks, suggestions, assistances)
- From1° September 2016
  - Produced 322.900 engagements on 4270 users
- From 1° July 2017
  - 233 registered users with email (154 via social)

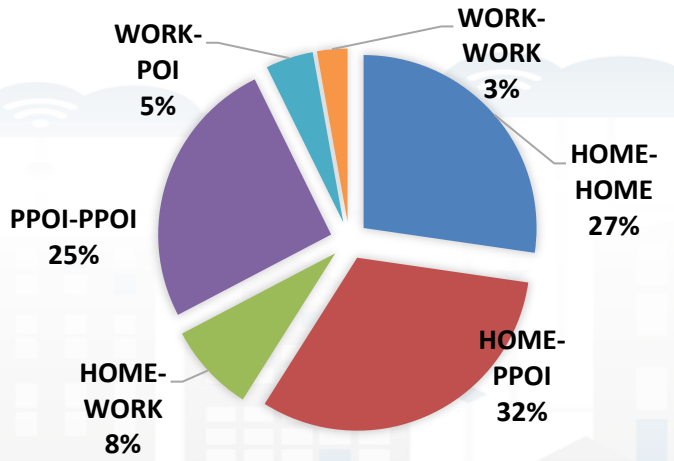
From: 2016-08-01 00:00:00
To: 2017-05-26 11:19:52
Refresh

Rule name	Type	#sent	#viewed	#viewed on #sent	#executed	#executed on #viewed	Description
confirm_ppoi_home_de	ENGAGEMENT	4 (0%)	1 (0.03%)	25%	0 (0%)	0%	Ask (in german) a confirmation of the position of user's HOME
confirm_ppoi_home_en	ENGAGEMENT	90 (0.03%)	0 (0%)	0%	0 (0%)	0%	Ask (in english) a confirmation of the position of user's HOME
confirm_ppoi_home_es	ENGAGEMENT	13 (0.01%)	1 (0.03%)	7.69%	1 (1.04%)	100%	Ask (in spanish) a confirmation of the position of user's HOME
confirm_ppoi_home_fr	ENGAGEMENT	1 (0%)	0 (0%)	0%	0 (0%)	0%	Ask (in french) a confirmation of the position of user's HOME
confirm_ppoi_home_it	ENGAGEMENT	1794 (0.69%)	97 (2.97%)	5.41%	16 (16.67%)	16.49%	Ask (in italian) a confirmation of the position of user's HOME
confirm_ppoi_school_it	ENGAGEMENT	45 (0.02%)	1 (0.03%)	2.22%	1 (1.04%)	100%	Ask (in italian) a confirmation of the position of user's SCHOOL
confirm_ppoi_work_en	ENGAGEMENT	15 (0.01%)	0 (0%)	0%	0 (0%)	0%	Ask (in english) a confirmation of the position of user's WORK
confirm_ppoi_work_it	ENGAGEMENT	368 (0.14%)	37 (1.13%)	10.05%	7 (7.29%)	18.92%	Ask (in italian) a confirmation of the position of user's WORK
daily_event_de	ENGAGEMENT	105 (0.04%)	19 (0.58%)	18.1%	0 (0%)	0%	Suggest (in german) an event currently on in Florence
daily_event_en	ENGAGEMENT	2115 (0.81%)	66 (2.02%)	3.12%	0 (0%)	0%	Suggest (in english) an event currently on in Florence
parking_it	ASSISTANCE	659 (0.25%)	75 (2.3%)	11.38%	0 (0%)	0%	Alert (in italian) if the user parked in a residential parking zone
shoot_a_photo_de	ENGAGEMENT	604 (0.23%)	4 (0.12%)	0.66%	0 (0%)	0%	Ask (in german) a contribution for a nearby point-of-interest
shoot_a_photo_en	ENGAGEMENT	11159 (4.29%)	37 (1.13%)	0.33%	0 (0%)	0%	Ask (in english) a contribution for a nearby point-of-interest
shoot_a_photo_es	ENGAGEMENT	2140 (0.82%)	11 (0.34%)	0.51%	1 (1.04%)	9.09%	Ask (in spanish) a contribution for a nearby point-of-interest
shoot_a_photo_fr	ENGAGEMENT	2880 (1.11%)	4 (0.12%)	0.14%	0 (0%)	0%	Ask (in french) a contribution for a nearby point-of-interest
shoot_a_photo_it	ENGAGEMENT	216479 (83.29%)	976 (29.88%)	0.45%	14 (14.58%)	1.43%	Ask (in italian) a contribution for a nearby point-of-interest
spent_time_en	ENGAGEMENT	53 (0.02%)	14 (0.43%)	26.42%	0 (0%)	0%	Ask (in english) a confirmation of the position of user's POI
spent_time_it	ENGAGEMENT	1192 (0.46%)	175 (5.36%)	14.68%	15 (15.63%)	8.57%	Ask (in italian) a confirmation of the position of user's POI
	- commuter	24 (2.01%)	2 (1.14%)	8.33%	0 (0%)	0%	
	- student	131 (10.99%)	11 (6.29%)	8.4%	2 (13.33%)	18.18%	
	- tourist	238 (19.97%)	59 (33.71%)	24.79%	0 (0%)	0%	
	- citizen	665 (55.79%)	122 (69.71%)	18.35%	10 (66.67%)	8.2%	
	- operator	26 (2.18%)	5 (2.86%)	19.23%	1 (6.67%)	20%	
	- all	155 (13%)	23 (13.14%)	14.84%	2 (13.33%)	8.7%	
survey_turist_de	ENGAGEMENT	179 (0.07%)	15 (0.46%)	8.38%	1 (1.04%)	6.67%	Propose (in german) a survey to turist after they left Florence
survey_turist_en	ENGAGEMENT	966 (0.37%)	13 (0.4%)	1.35%	2 (2.08%)	15.38%	Propose (in english) a survey to turist after they left Florence
survey_turist_es	ENGAGEMENT	115 (0.04%)	2 (0.06%)	1.74%	0 (0%)	0%	Propose (in spanish) a survey to turist after they left Florence
survey_turist_fr	ENGAGEMENT	51 (0.02%)	2 (0.06%)	3.92%	0 (0%)	0%	Propose (in french) a survey to turist after they left Florence
survey_turist_it	ENGAGEMENT	1006 (0.39%)	42 (1.29%)	4.17%	6 (6.25%)	14.29%	Propose (in italian) a survey to turist after they left Florence

CURRENT NUMBERS

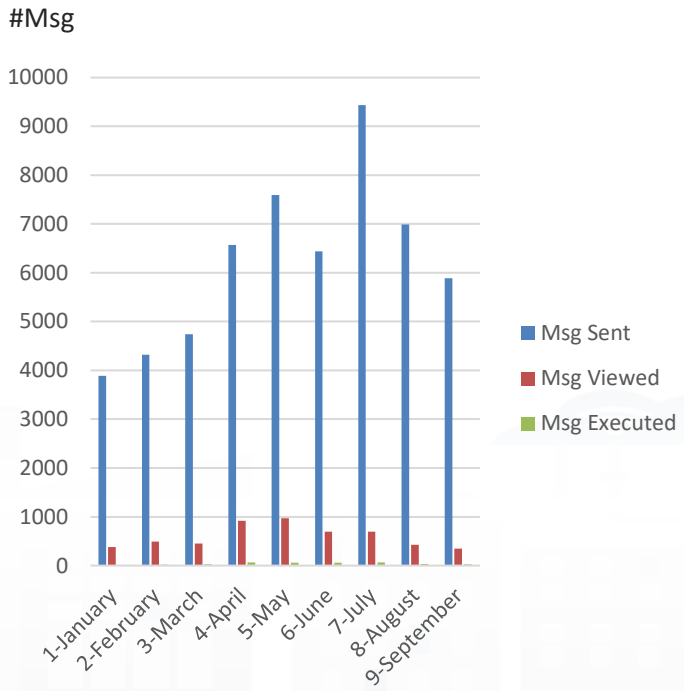
- From 1° September 2016
  - Detected 2108 PPOIs on 1080 users
    - 437 HOME
    - 285 WORK
    - 34 SCHOOL
    - 1350 EXTRA
  - 130 PPOIs are feedbacked
  - 460 survey responses
- From 1° August 2017
  - Built 524 Markov Networks about user's trajectories

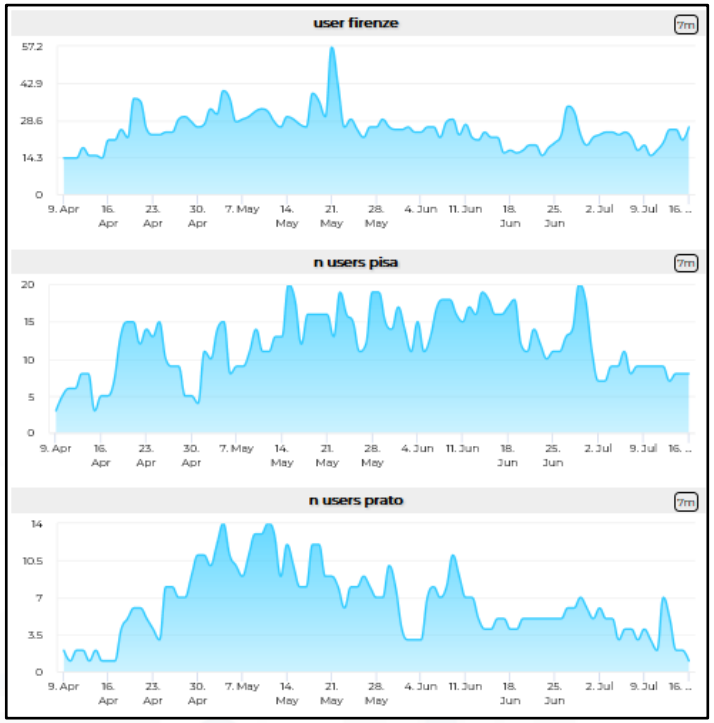


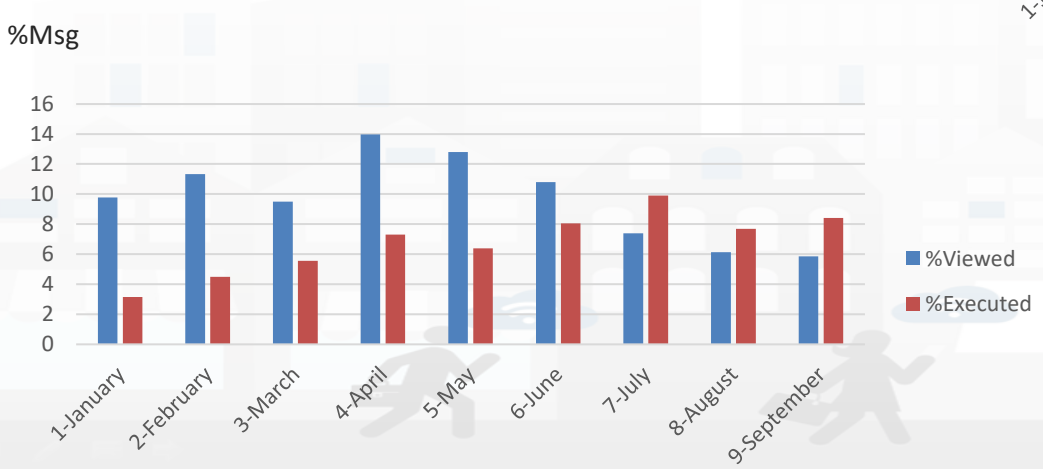


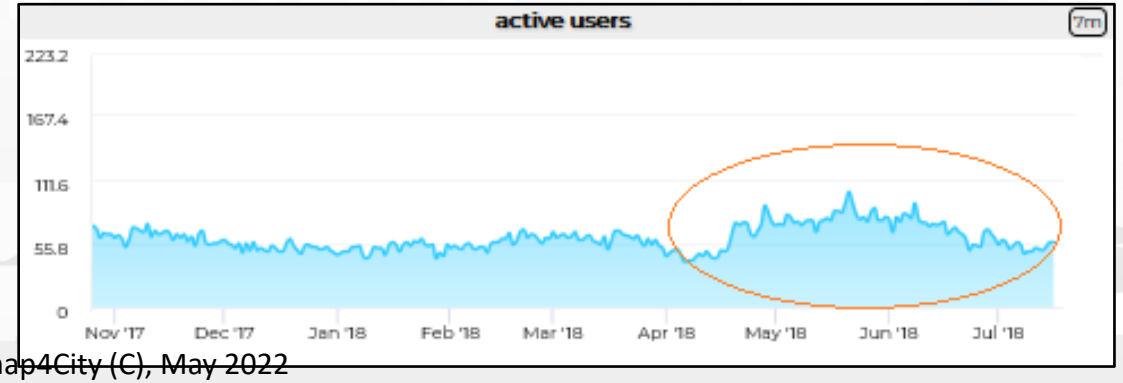
Validation of user Engagement

Months	Msg Sent	Msg Viewed	Msg Executed
1-January	3888	380	12
2-February	4319	489	22
3-March	4739	450	25
4-April	6567	918	67
5-May	7594	972	61
6-June	6437	695	55
7-July	9432	697	69
8-August	6988	429	73
9-September	5885	345	49
Total	55849	5375	433







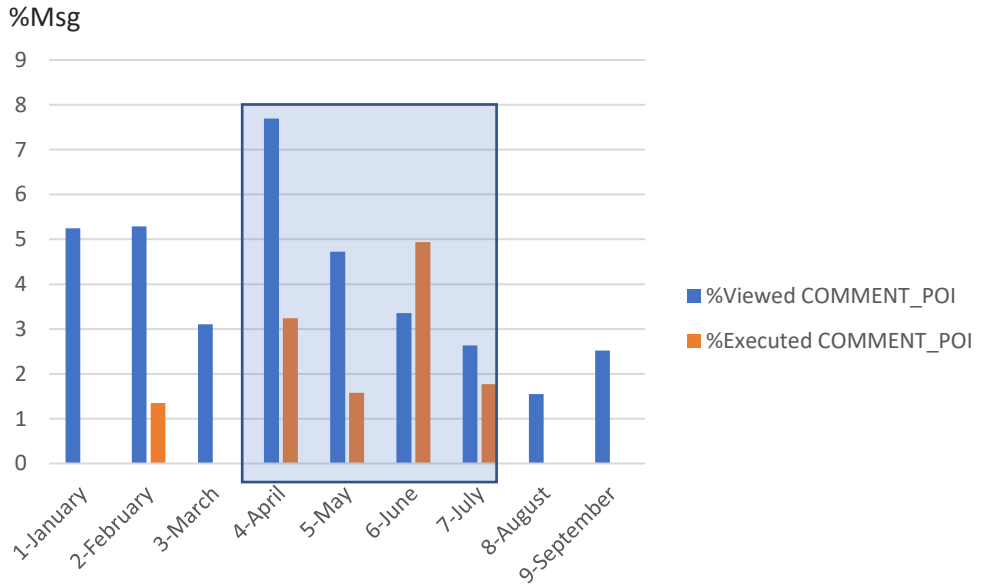


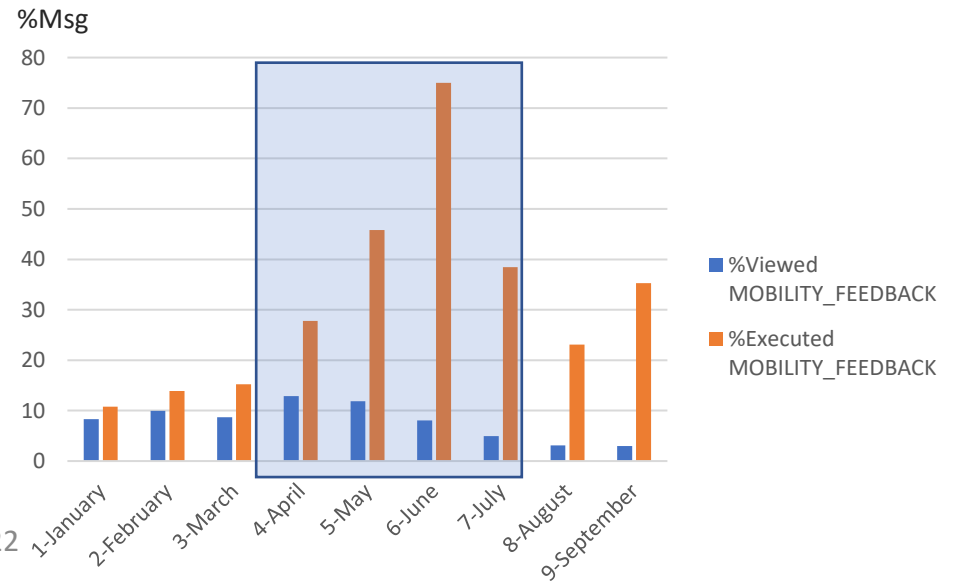
User Behaviour Analysis

VALIDATION

- During the PILOT new rules has been added (30 on a total of 80) and mostly all of them are still online
- COMMENT_POI: requires more user interaction and not very contextualized (POI proximity) → higher rate of sent, lower rate on execution
- MOBILITY_FEEDBACK: requires less user iteration and very contextualized (user in MOBILITY) → normal rate of sent, high rate on execution

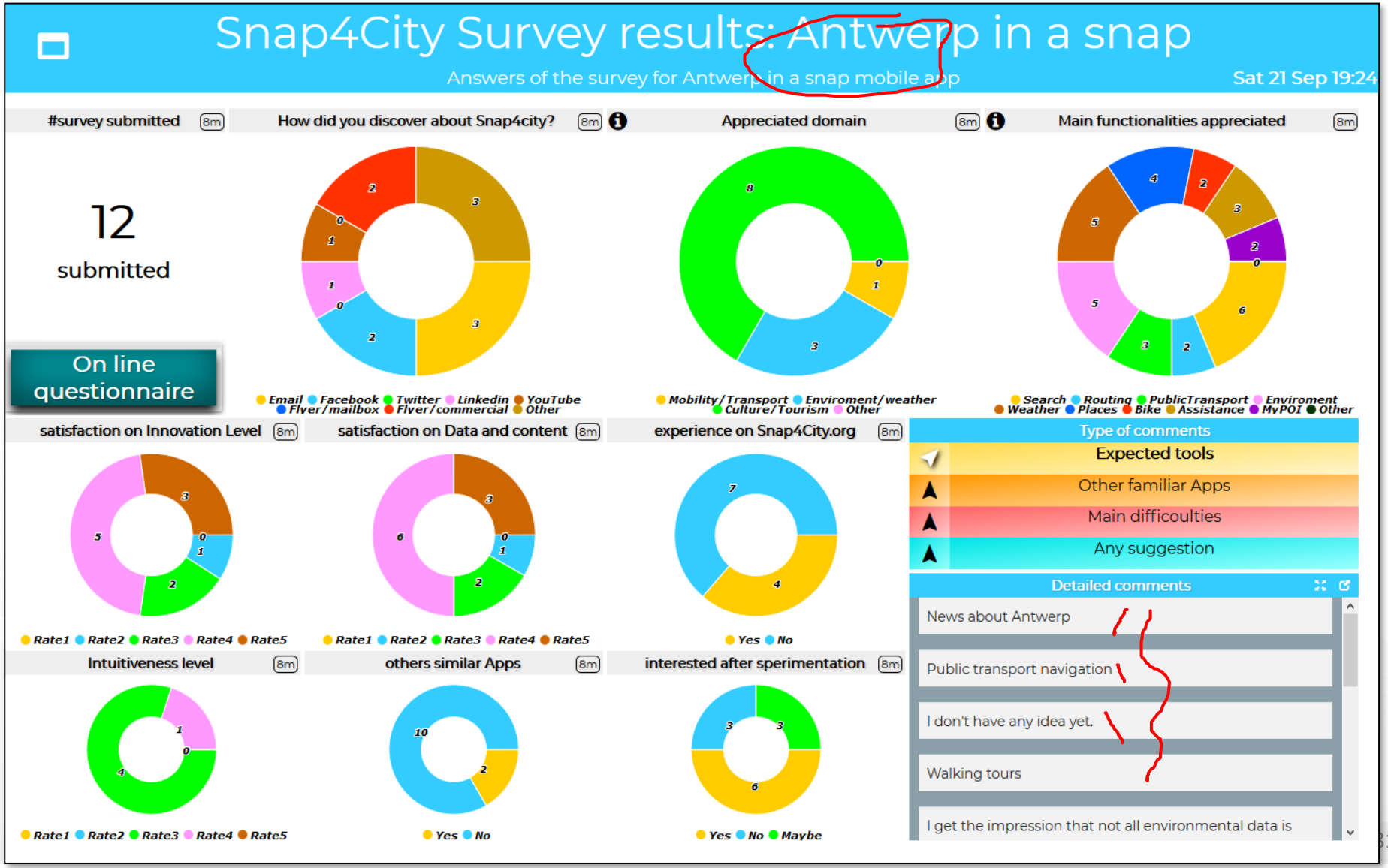
	Msg Sent	Msg Viewed	Msg Executed
COMMENT_POI	**21632**	**804**	**15**
MOBILITY_FEEDBACK	**5378**	**371**	**94**





https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MTc2OQ==

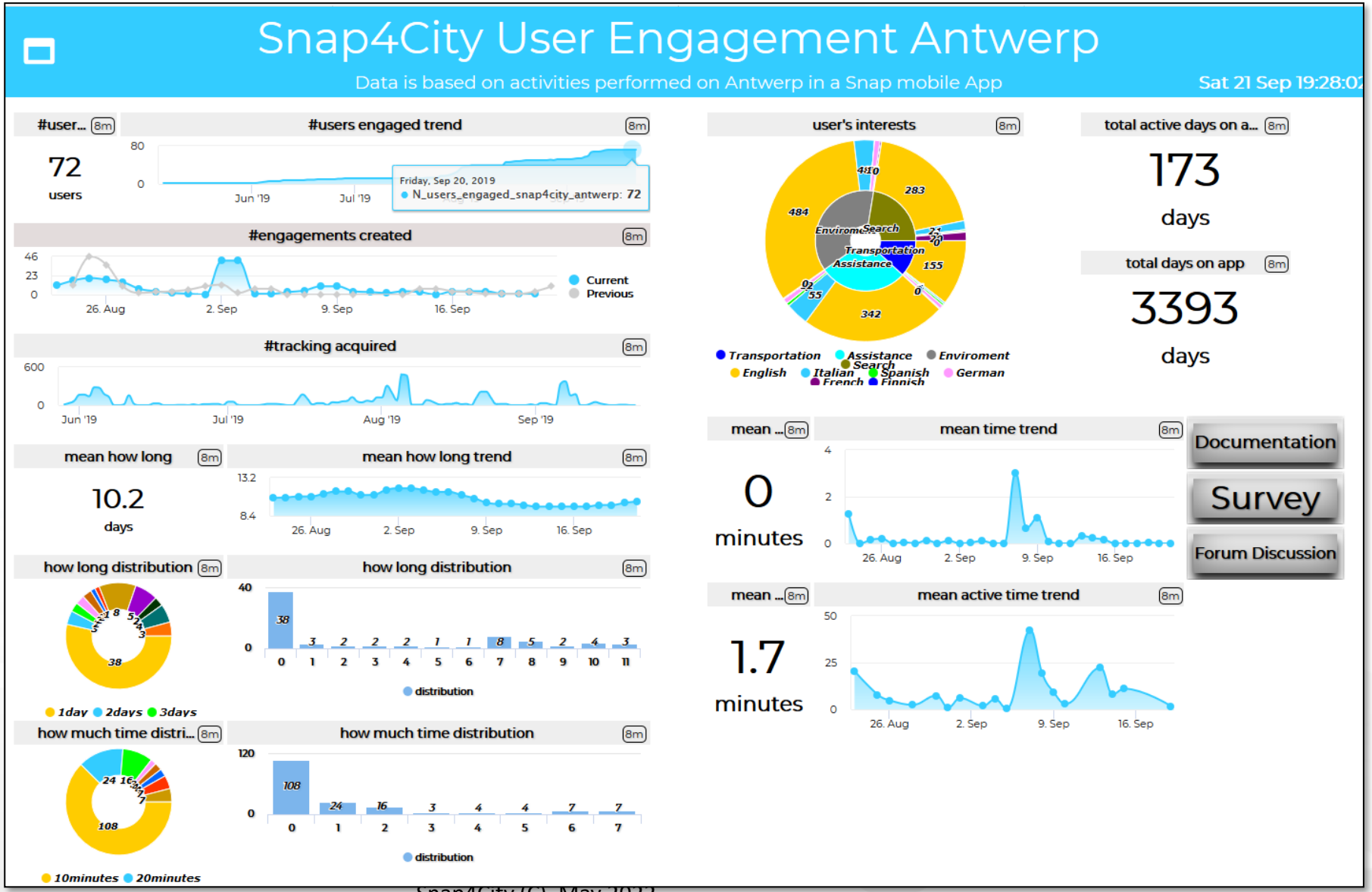
Dashboard created to monitor in real time the answers to the survey provided on the Mobile App directly by the Engagement tool



https://www.snap4city.org/dashboardSmartCity/view/index.php?iddasboard=MTc1OQ==

Dashboard monitoring the Mobile App:

- Collecting the clicks
- Describing the community of users in terms of the profile aspects
- Measuring the time spend, and topics of interest of the users, etc.



TOP

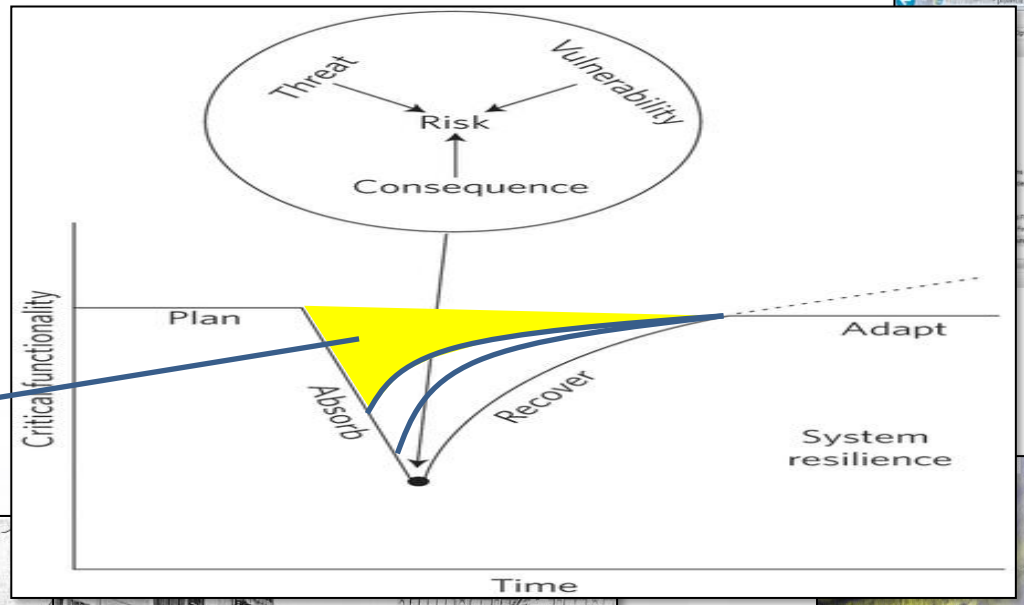
DECISION SUPPORT SYSTEM AND CITY RESILIENCE

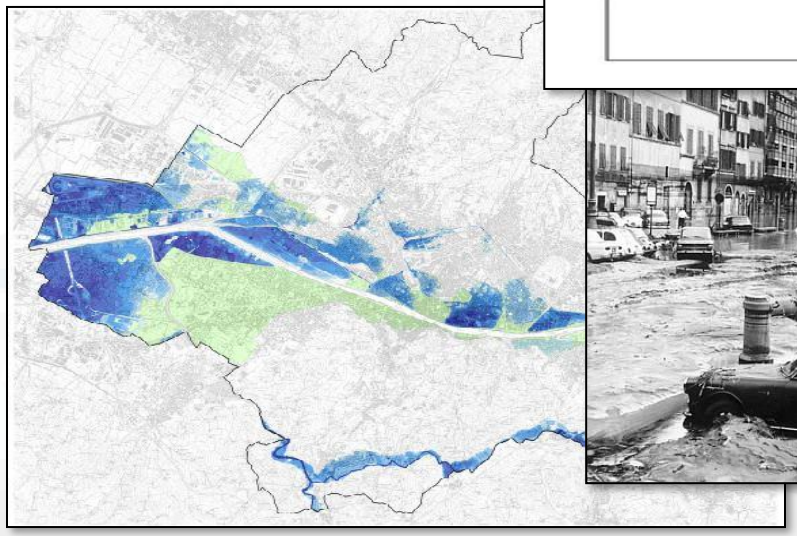
City Resilience

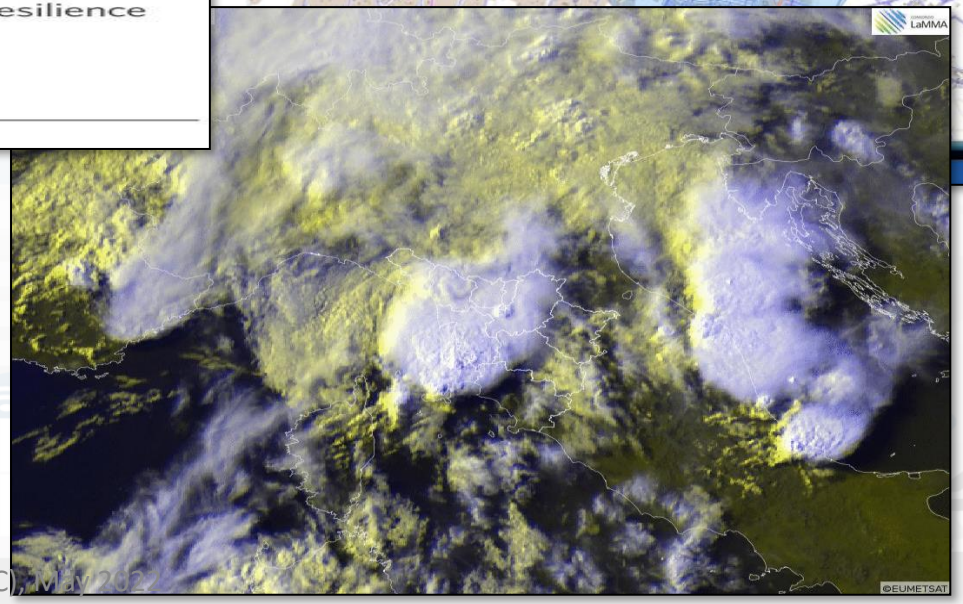
Early Warning, Detection

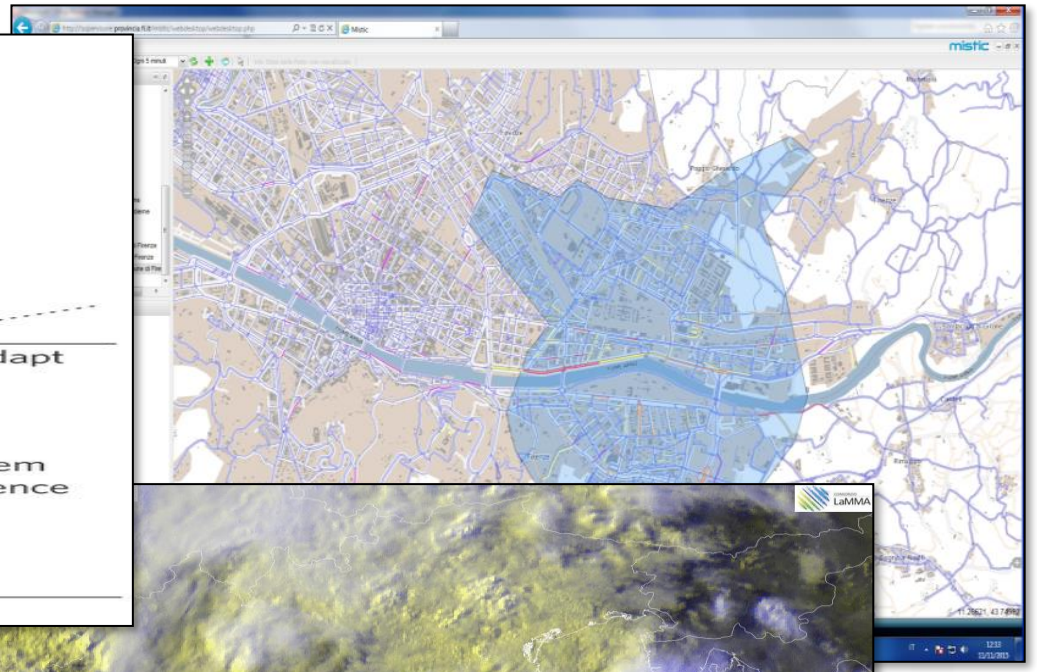
Prepare
Absorb
Recover
Adapt

damage









Early Warning, Detection

Issue:

- Detection of critical condition
- Not easily detected with other means

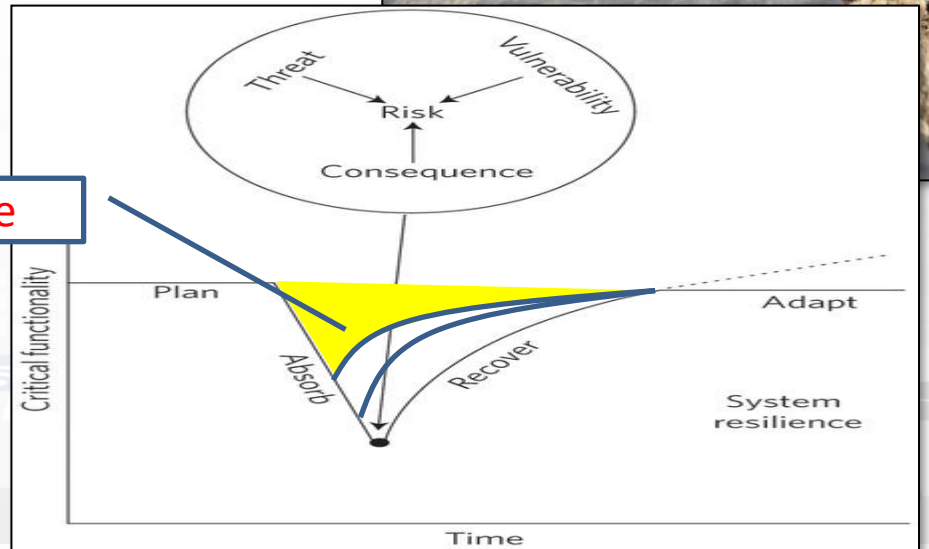
Impact:

- Early warning, faster reaction
- Increased resilience

Several metrics related to:

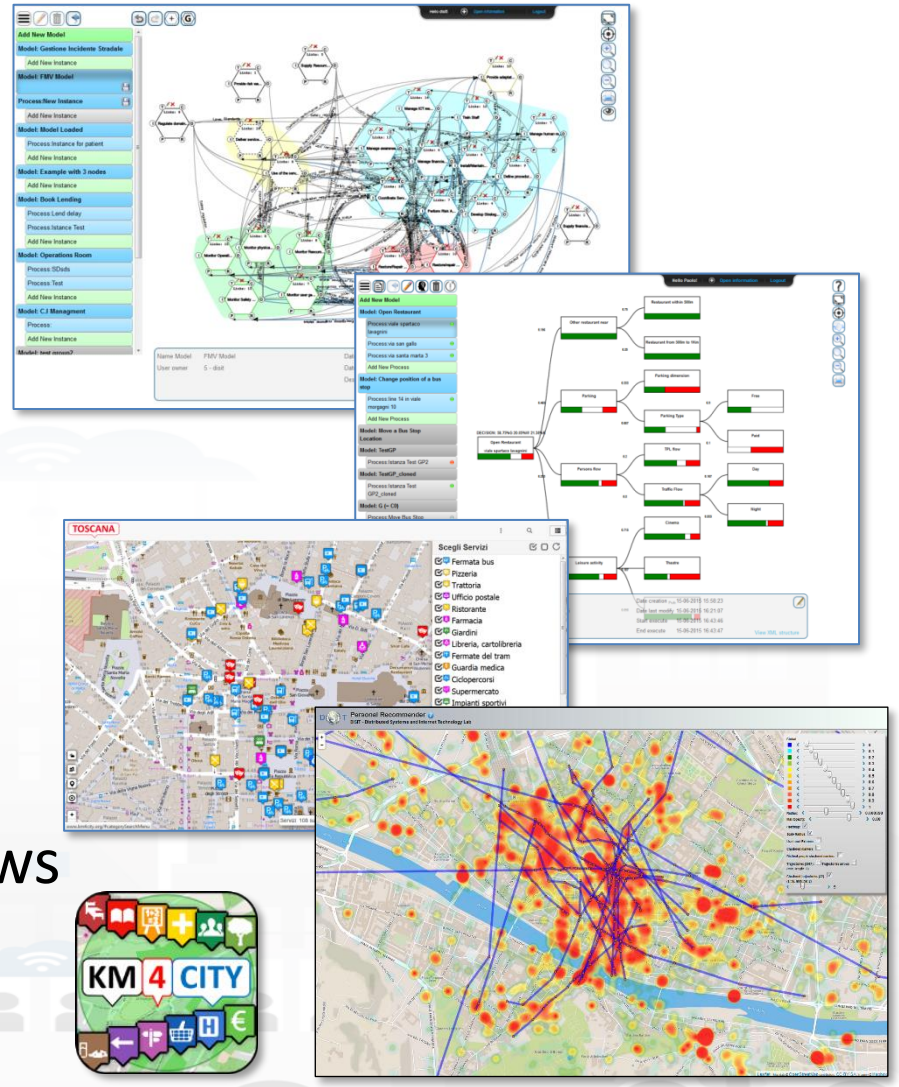
- Volume of retweets
- Sentiment analysis

Prepare
Absorb
Recover
Adapt



Main Approach

- Three main layers
- Complex System modeling: function, processes, resources, time, events, etc.
  - Functional Resonance Analysis Method, FRAM
  - Resilience Analysis Grid, RAG
- Decision Support System, DSS
  - System Thinking, Goal Models
  - Risk analysis
  - UTS/ITS decision supports
- Data, big data access and exploitation
  - Data Analytics, Internet of Things, sensors, flows
  - People flow and behavior
  - Social Media



ERMG: European Resilience Management Guide

http://www.resolute-eu.org/

CRAMSS
Collaborative Resilience Assessment and Management Support System

ANTICIPATING

- European Resilience Management Guidelines
- Game Based Training

MONITORING

- Big Data Platform
- IoT/IoE/Open Data
- Real Time Dashboard
- Resilience Control Room
- Data Analytics
- Early Warnings
- Urban Traffic Manager Data Exchenge

RESPONDING

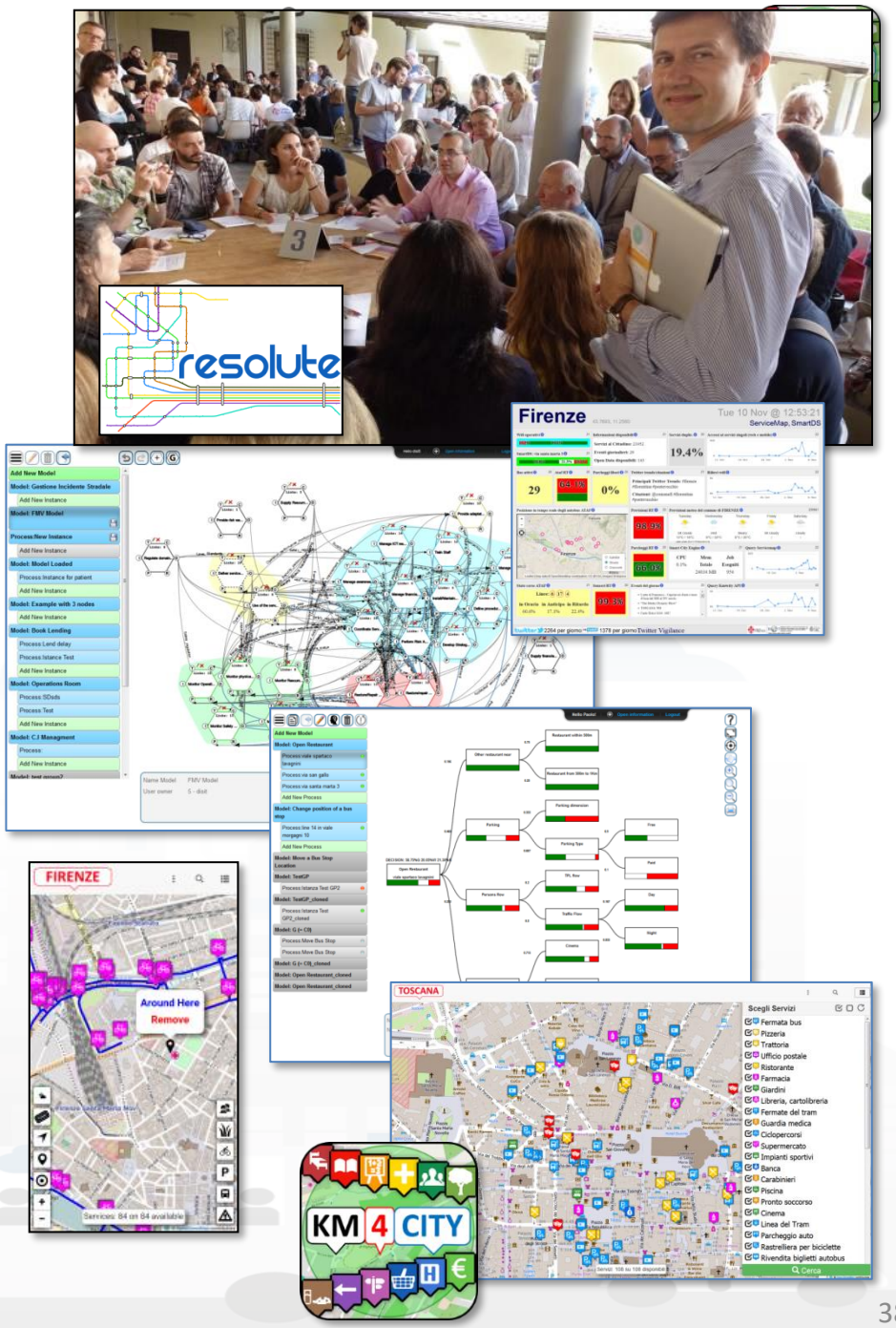
- Smart Decision Support Systems (DSS)
- Evacuation Decision Support
- Smart Intelligent Transport Systems
- Emergency Support Smart App
- Resilience DSS

LEARNING

- Human Behavior Analysis
- Predictive Analytics
- Urban Transport System Dynamic Analysis
- Resilience Quantification
- Network Analysis

Improve city resilience, reducing risks and decision support

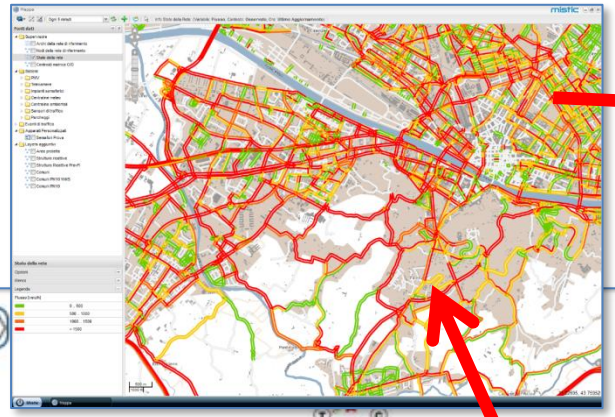
- **assessing** city **resilience** level
- **improving** city **resilience**, providing objective hints
- **improving** city users **awareness** with personal city assistants and participatory tools

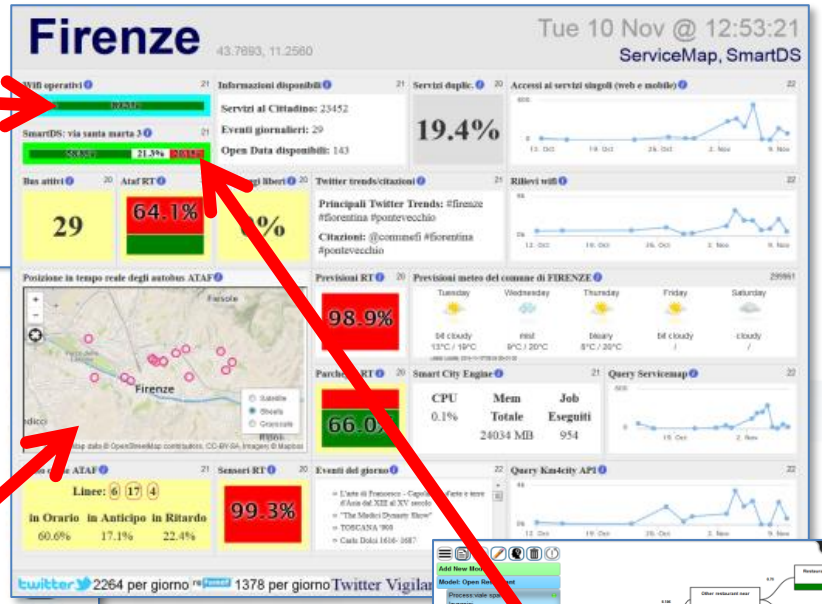


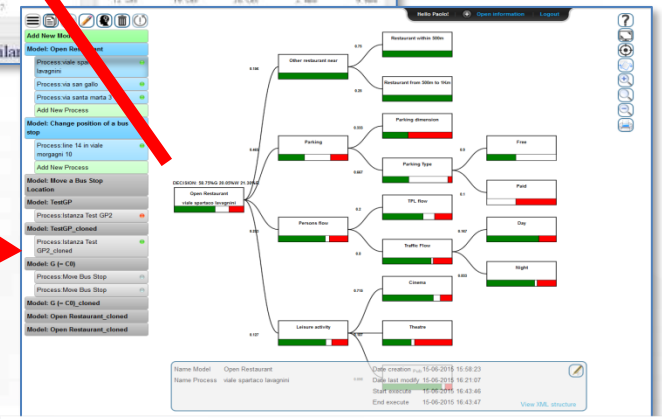
UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB

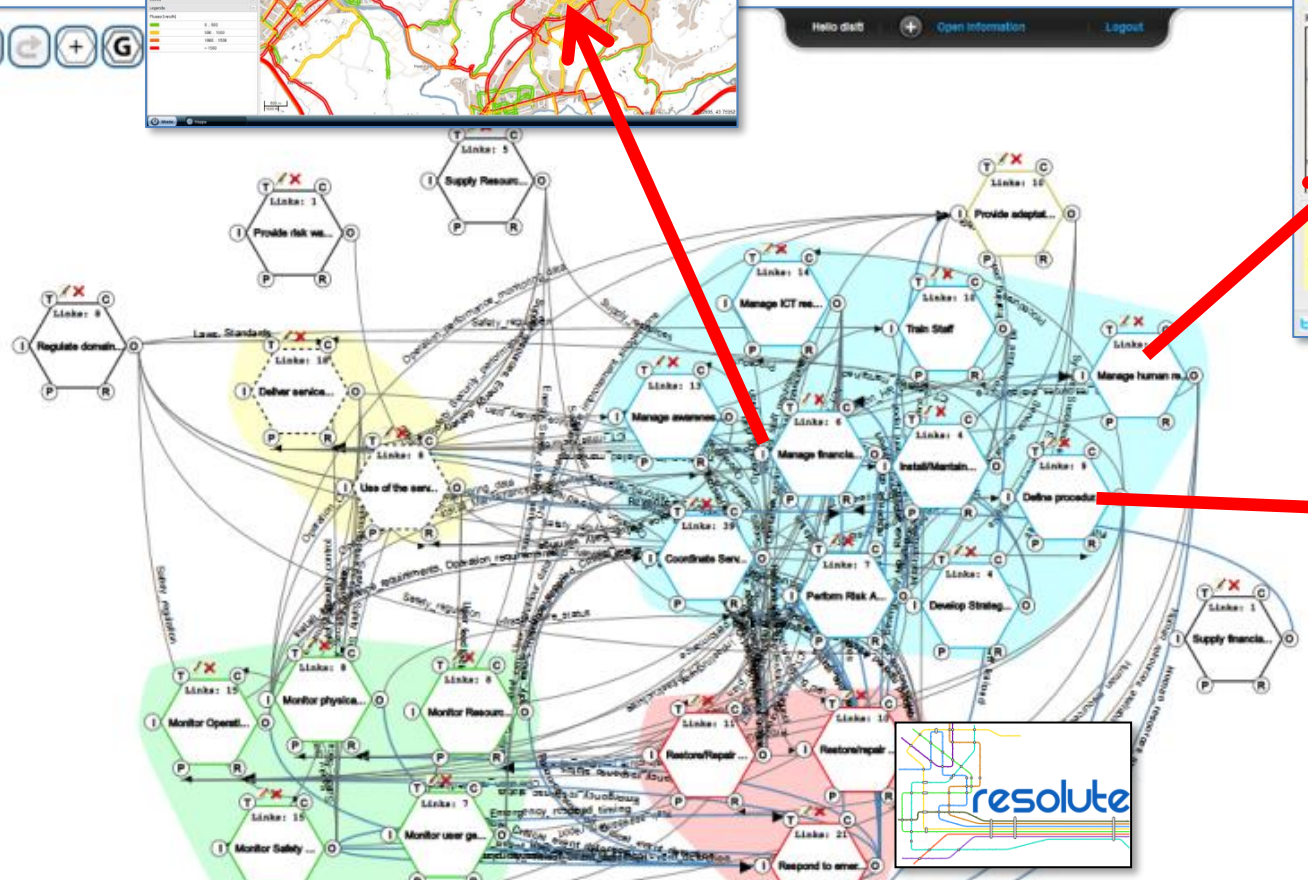
SNAP4CITY | KM 4 CITY

Dashboarding city resilience





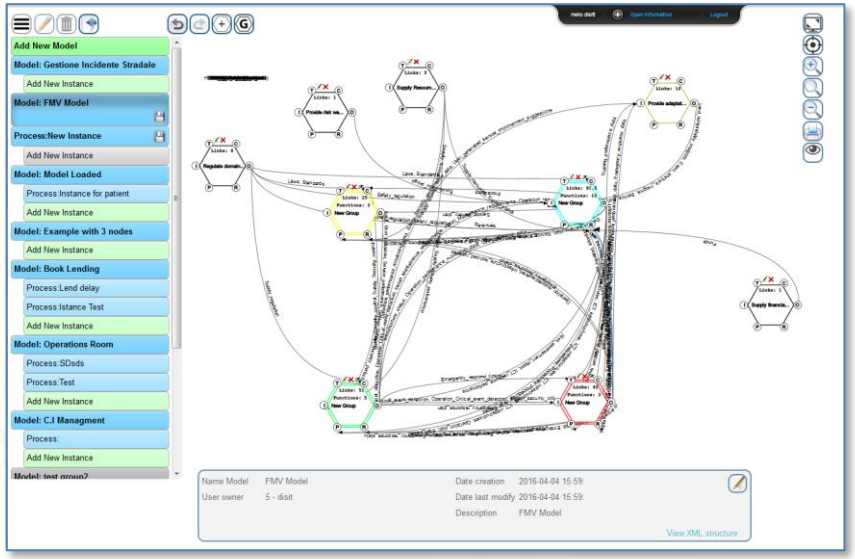


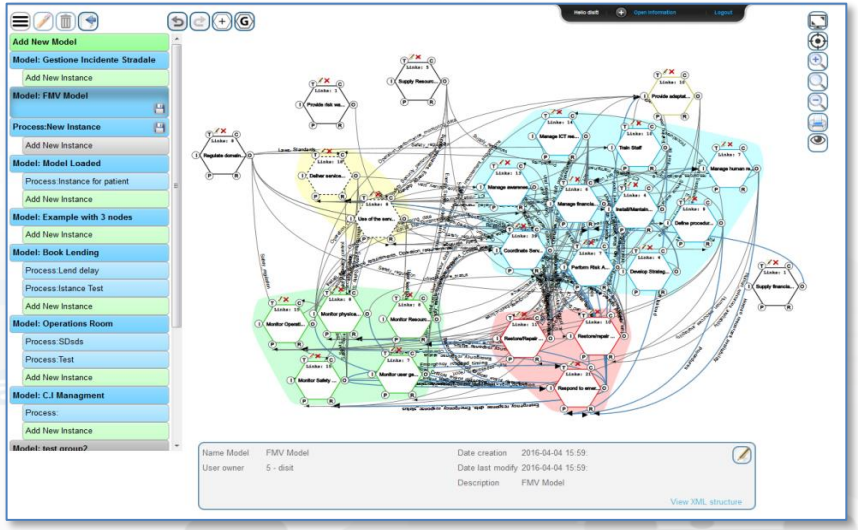


ResilienceDS Tool

http://resilienceds.km4city.org

- **FRAM Model**
  - Macro FRAM processes
  - Metrics for Process complexity assessment
  - Operational Semantic for executing FRAM model
  - Connection with SmartDS
  - Connection with BigData open to multiple sources of data and workgroup results, Km4City
- **Collaborative work, web tool**
- **Open for all**
- **Validated on ERMG: European Guidelines**

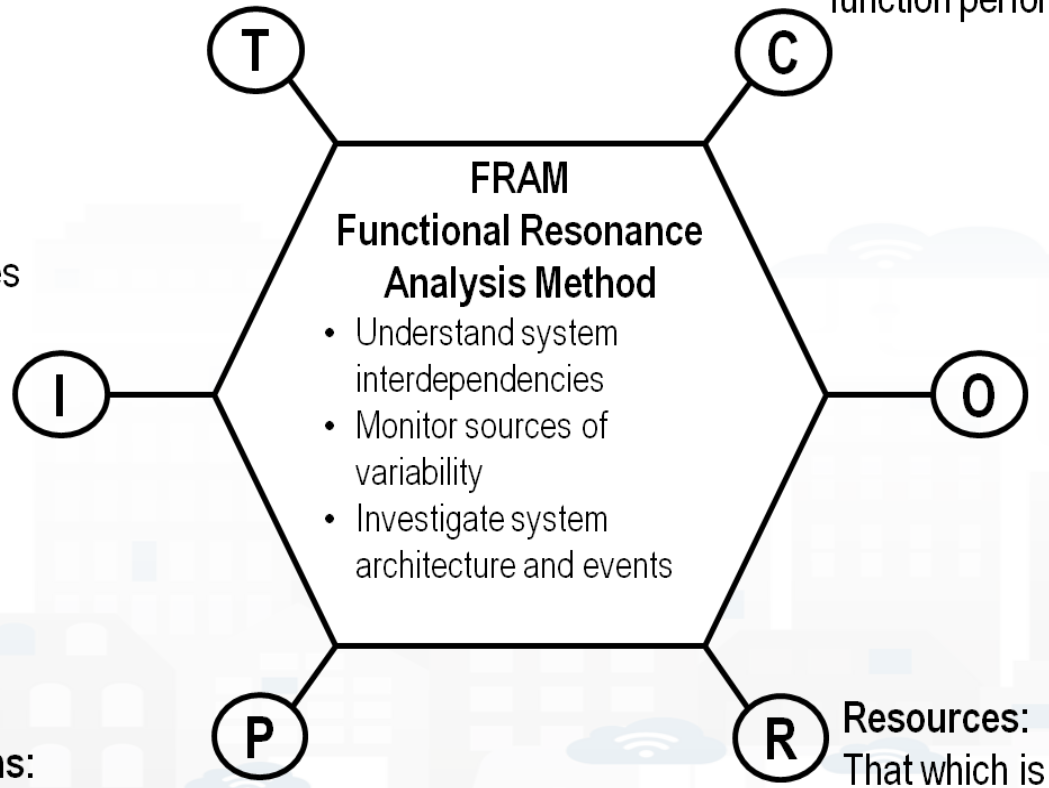




Functional Resonance Analysis Method

Time:
This is simultaneously a resource and a constraint

Controls:
That which supervises or adjust function performance



Input:
That which engages function operation and is used to produce the function output

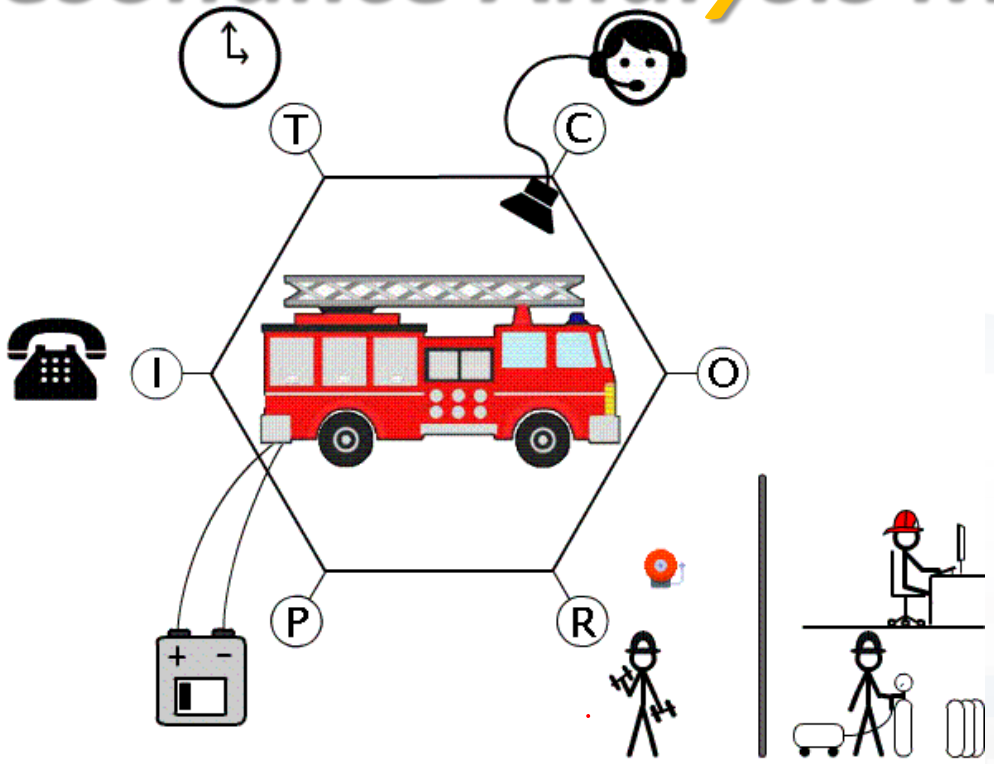
Output:
That which is produced by the function and becomes the input for other functions

Preconditions:
System conditions that must be met before function can be carried out

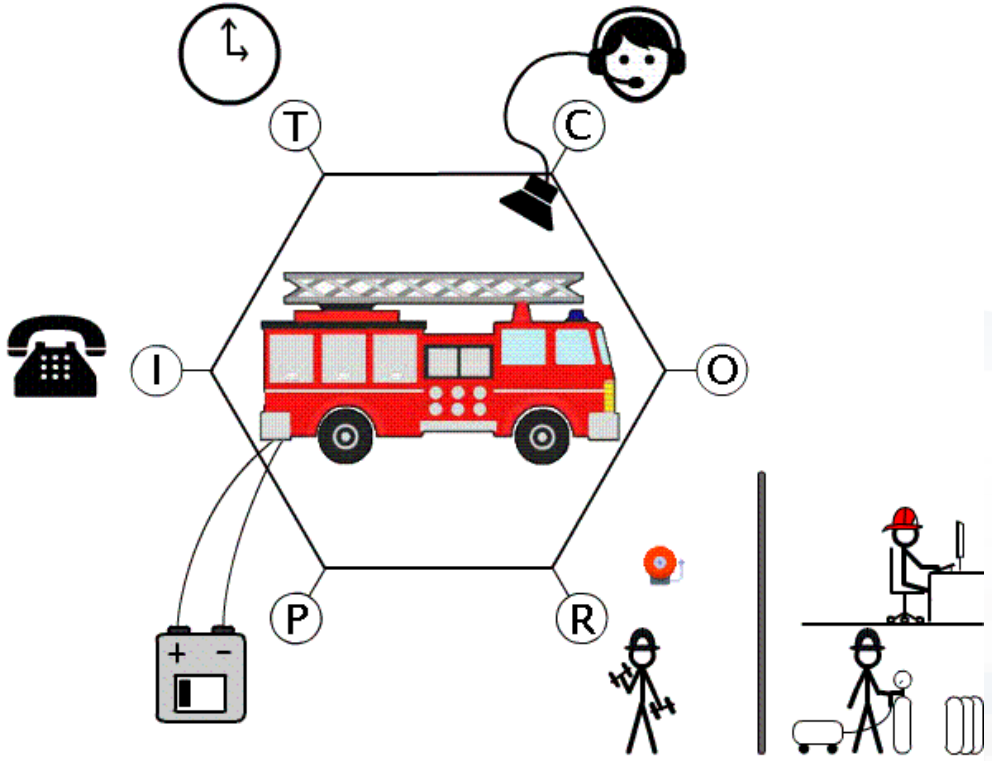
Resources:
That which is needed and/or consumed by the function to process the input

Fram Model: *Functional Resonance Analysis Method*

- Success and failure are equivalent in the sense that they both emerge from performance variability.
- Variability, intended as a way for people to adjust tools and procedures to match operating conditions.
- Emergence of either success or failure is due to unexpected combination of variability from multiple functions.
- The unexpected “amplified” effects of interactions between different sources of variability are at the origin of the phenomenon described by functional resonance.



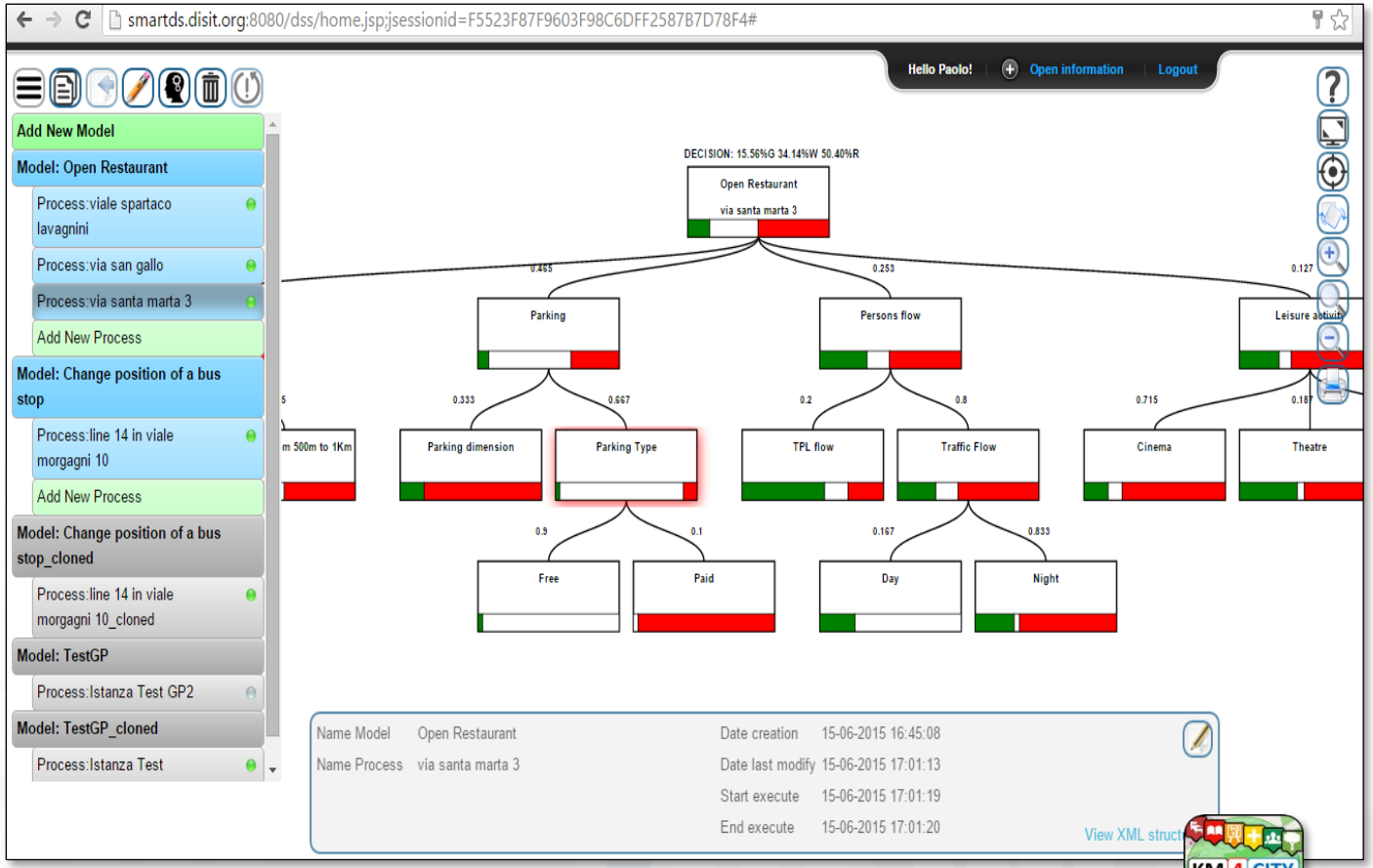
- Success and failure are equivalent in the sense that they both emerge from performance variability.
- Variability, intended as a way for people to adjust tools and procedures to match operating conditions.
- Emergence of either success or failure is due to unexpected combination of variability from multiple functions.
- The unexpected "amplified" effects of interactions between different sources of variability are at the origin of the phenomenon described by functional resonance.



resolute

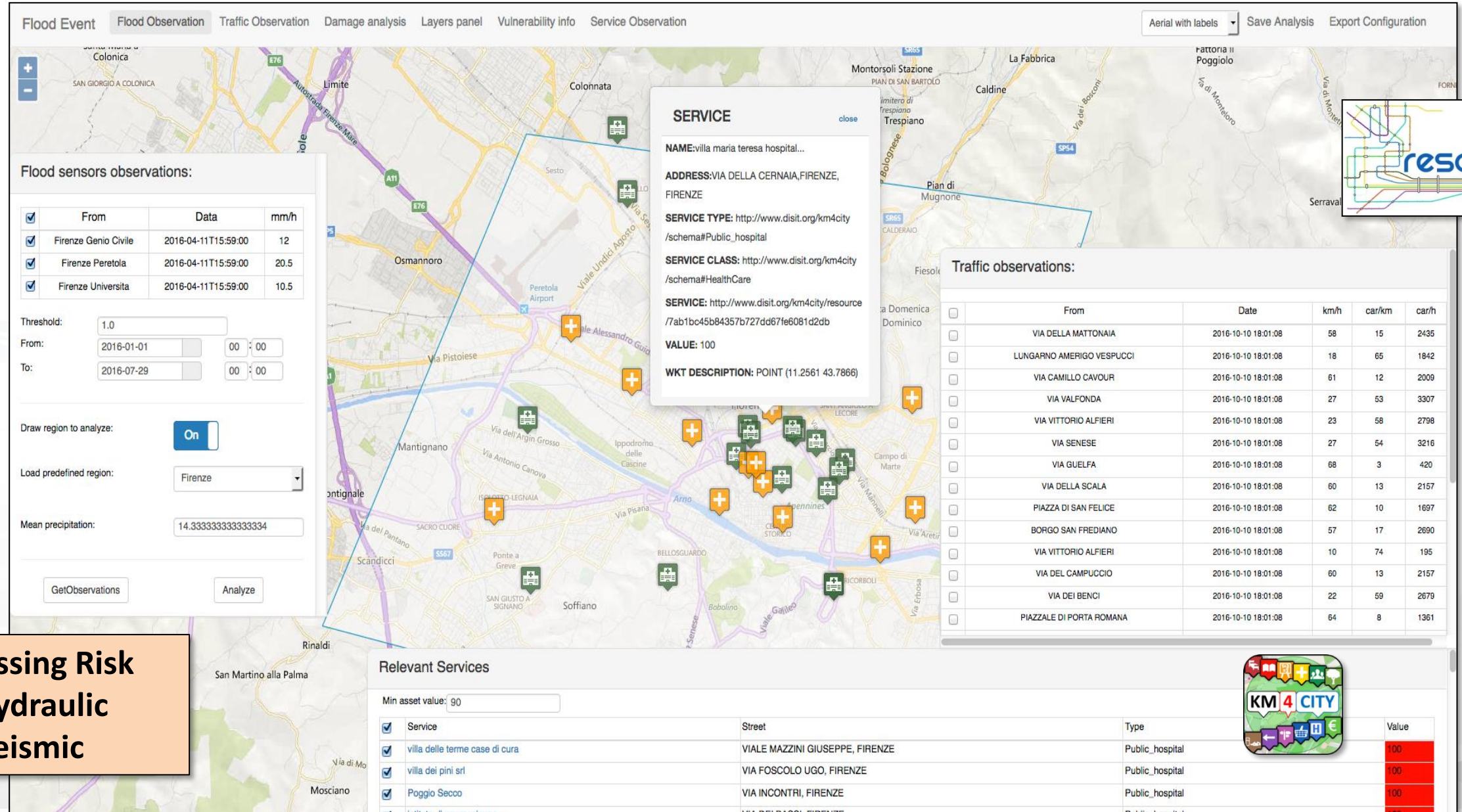
Smart Decision Support , system thinking

- **Smart Decision Support System** based on System Thinking plus
- Actions to city reaction, resilience, smartness, …
- Enforcing Mathematical model for propagation of decision confidence..
- Collaborative work, …
- Processes connected to city data: DB, RDF Store, Twitter, etc.
- Production of alerts/alarms
- Data analytics process
- Twitter Processes
- reuse, copy past, …



http://smartds.km4city.org

Risk Assessment

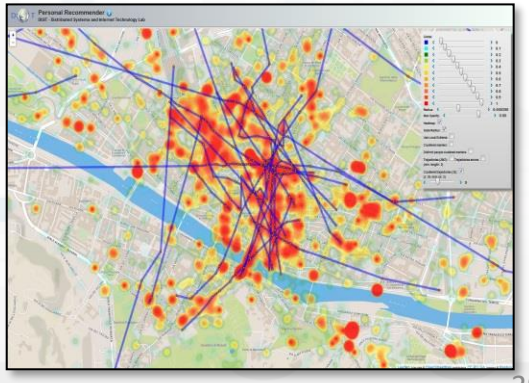


Assessing Risk

- **hydraulic**
- **Seismic**

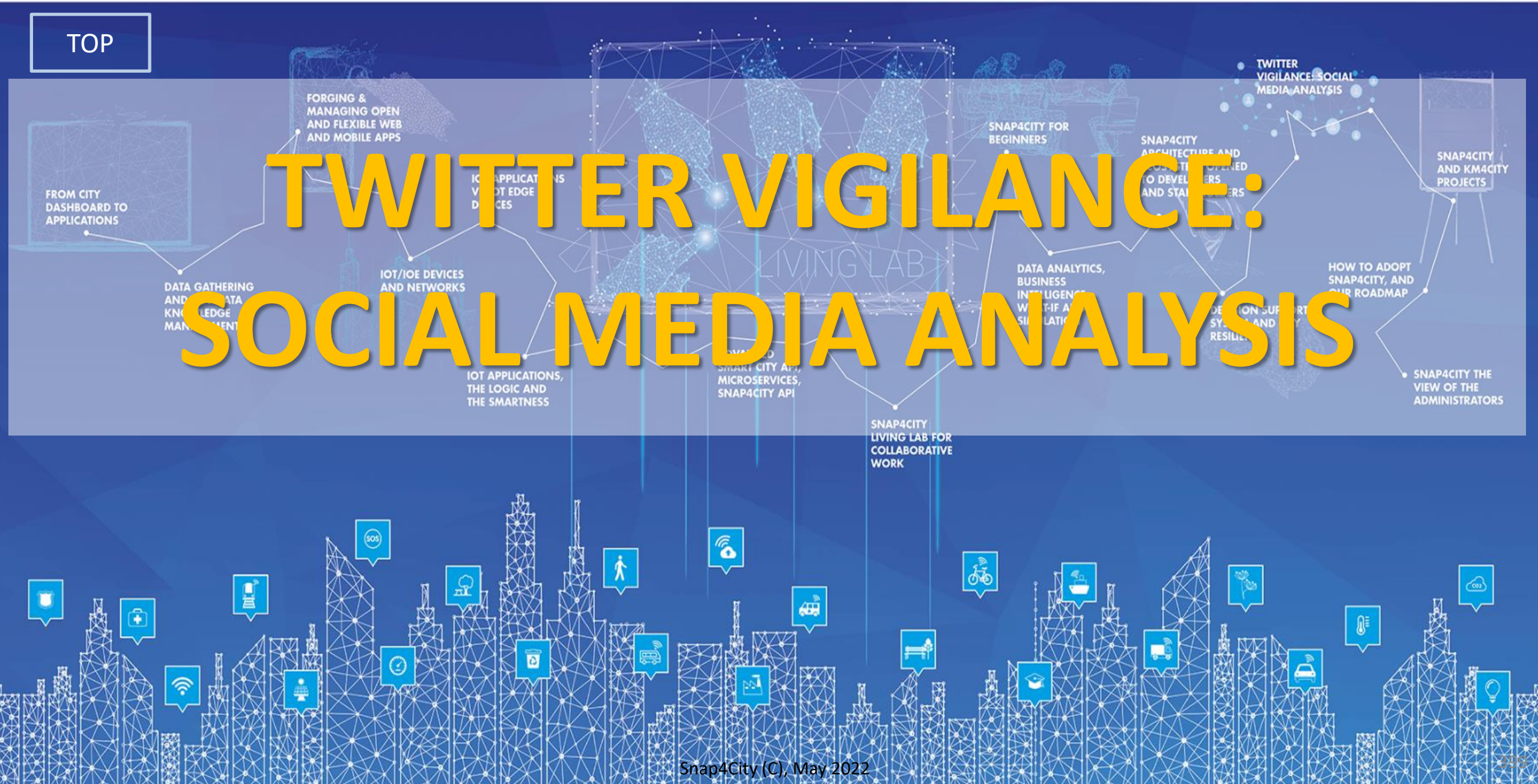
Mobile Emergency

- Personalized menu for Operators
- Providing information and suggestions to citizens
  - Civil Protection Page
  - Twitter Info
  - Geolocalized Info
- Tracking people and operators flows
- Collecting information from citizens
  - Comments
  - Images



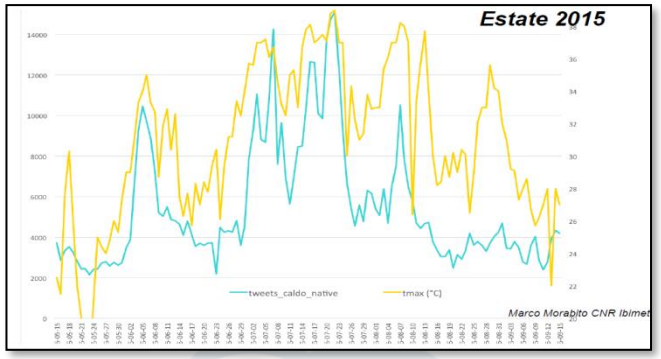
TOP

TWITTER VIGILANCE: SOCIAL MEDIA ANALYSIS



Prediction/Assessment

- Football game results as related to the volume of Tweets
- Number of votes on political elections, via sentiment analysis, SA
- Size and inception of contagious diseases
- marketability of consumer goods
- public health seasonal flu
- box-office revenues for movies
- places to be visited, most visited
- number of people in locations like airports
- audience of TV programmes, political TV shows
- weather forecast information
- Appreciation of services



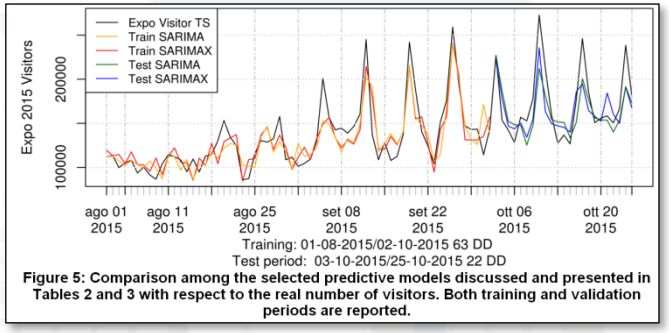
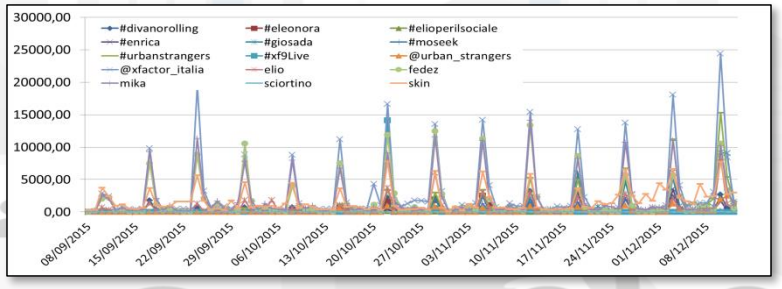
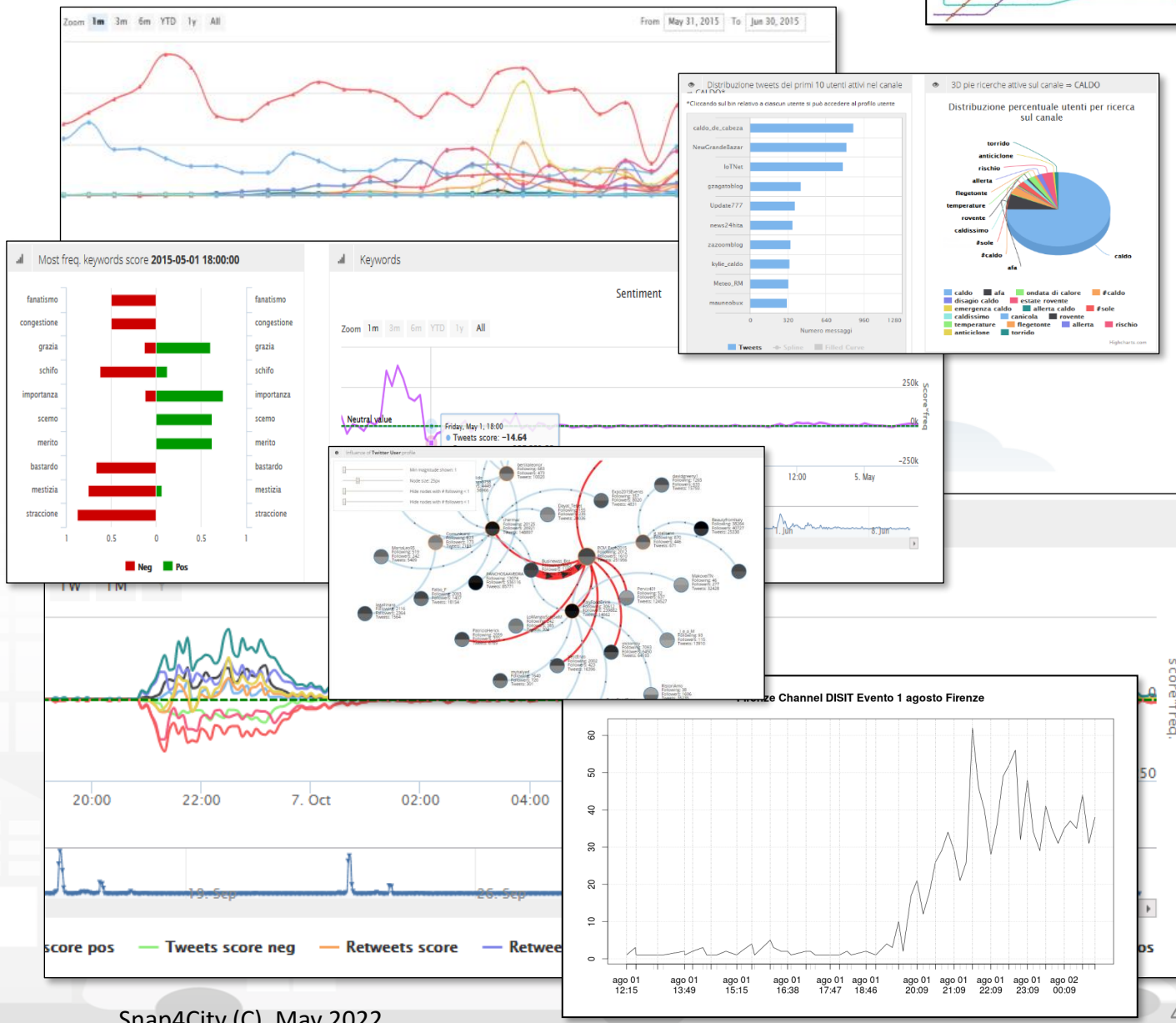


Figure 5: Comparison among the selected predictive models discussed and presented in Tables 2 and 3 with respect to the real number of visitors. Both training and validation periods are reported.



Twitter Vigilance

- http://www.disit.org/tv
- http://www.disit.org/rttv
- Citizens as sensors to
  - Assess sentiment on services, events, …
  - Response of consumers wrt, …
  - Early detection of critical conditions
  - Information channel
  - Opinion leaders
  - Communities
  - Formation
  - Predicting volume of visitors for tuning the services



Twitter Vigilance

Twitter Vigilance in Numbers

- **Used by several users:**
  - UnivFirenze, LAMMA, IBIMET, ARPAT, Master on Big Data, …
- **Active since** Aprile 2015
- **3 platforms for automated**:
  - Daily collection: statistical direct analysis and sentiment analysis
  - Real time collection and statiscal, sentiment analysis
  - Full faceted indexing: thus enabling search on collected tweets
- **All: precomputation of basic metric opening the activities of deep analysis**
- More than 350 million of tweets in the storage: ready on Hadoop cluster
- More than 250 channels
- More than 450 search activities daily
- From 400.000 to 4 Million of tweets per day.

Several Channels

Sii-Mobility

Twitter Vigilance Dashboard

Last crawling: 2015-12-11 01:20:20

Login

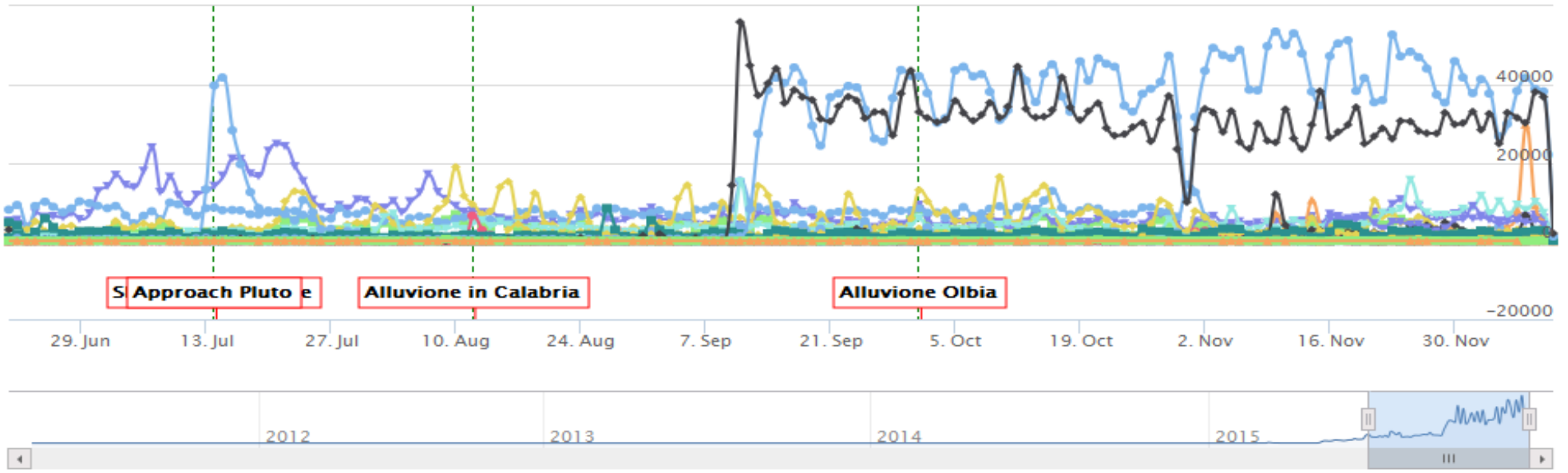
- Search parameters
- Data analysis 4
- Channel statistics
- Search statistics
- Retweet statistics
- Twitter Users statistics
- INFO

Home > Channel statistics

Global view of user channels

Zoom 1m 3m 6m YTD 1y All

From Jun 20, 2015 To Dec 11, 2015



ads aeroporto firenze Allertameteo TOSCANA apretoscana CALDO cambiamenti climatici
Codified Hashtags Allerta ConsumoSuolo Emergenza_acqua Europeana EXPO2015 Firenze
FirenzeICT Giubileo iononrischio LaMMA Maltempo maturità 2015 MeteoUSER mymeteo
NASA New Horizons papafrancesco PA_social_PA protezione civile toscana resilienza rossano
siena smartcity smartcitybigdata2015 SOLO allertameteoTOSCANA tech uber Univ-Firenze
zanzara Events

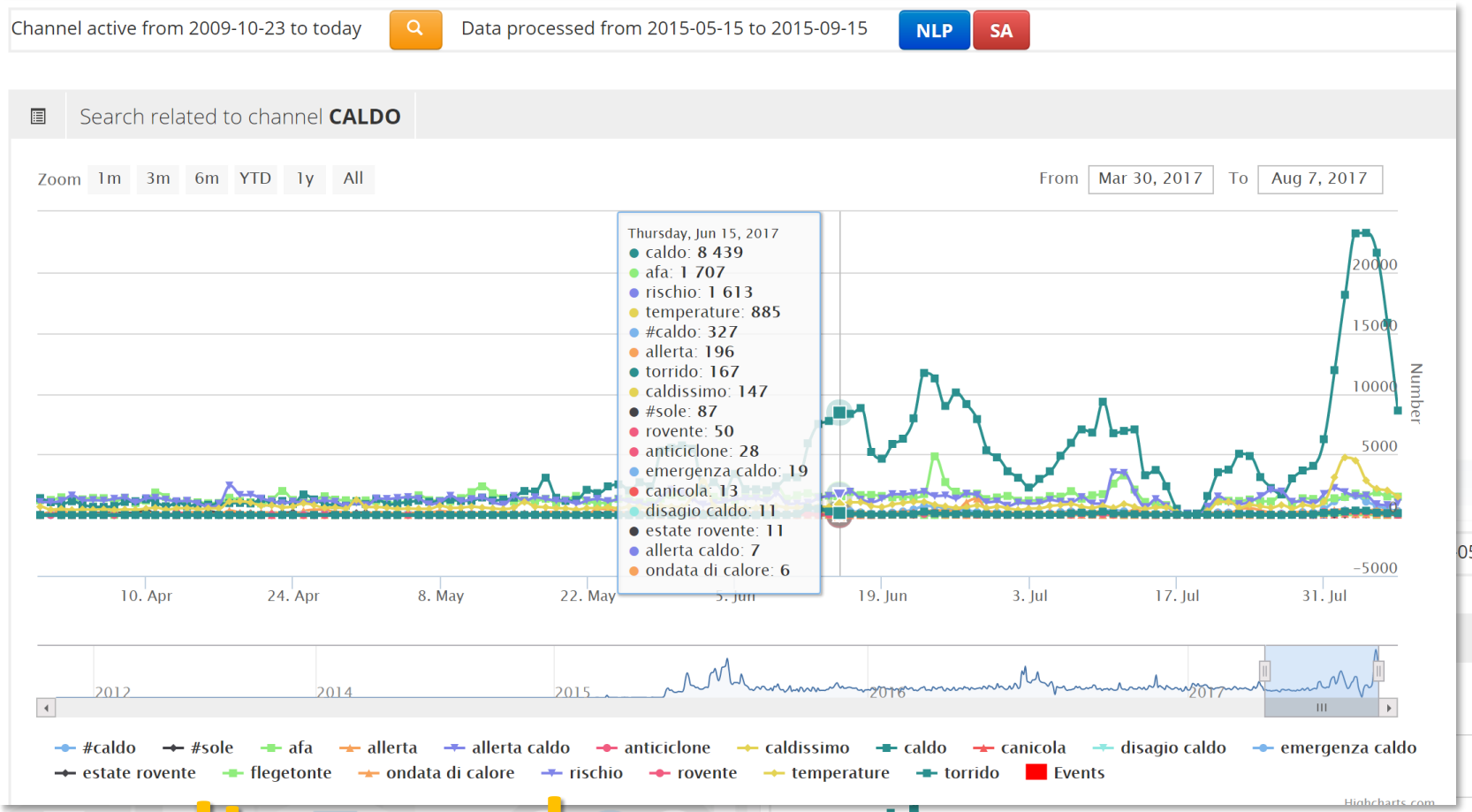
UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB
http://www.disit.org

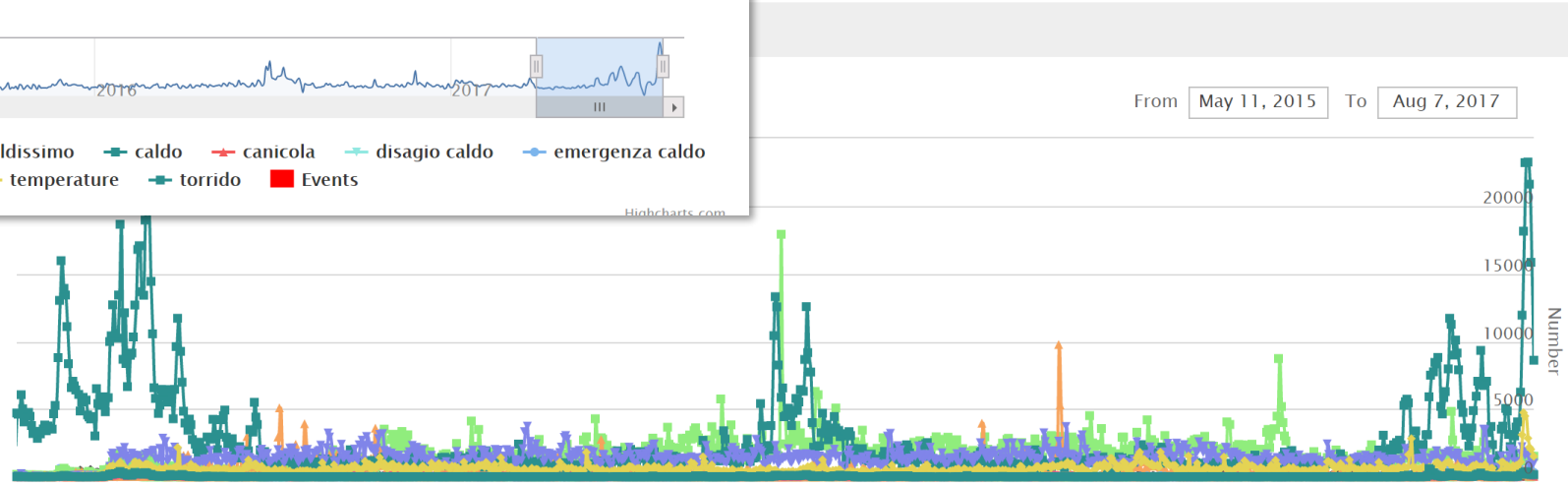
A Channel

Sii-Mobility

Its searches

Twitter Vigilance

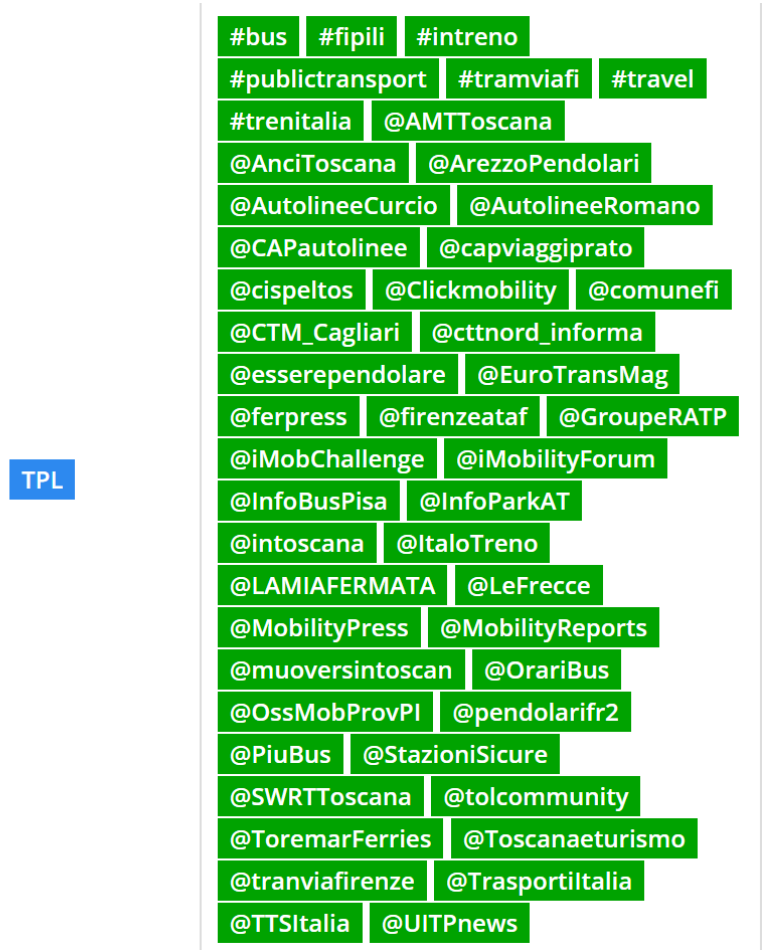




Snap4City (C), May 2022

Twitter Syntax for Searches

- String substring: Caldo
- Hashtag: #Caldo,
- Citations: @CivilProtection, @paolonesi
- From users: From:@paolonesi
- Etc.
- ….ANDed and ORed



Metrics' Kinds

- **Volume Metrics**
  - Number of TW, number of RTW
- **User Metrics**
  - Number of distinct users
  - Number of followers, following
- **NLP and SA metrics**
  - Counting word, adjective, noun, verbs, ….
  - Estimating SA, weighting with SentiWordNet (extended to Italian)
- **High level metrics (compositing all the other metrics)**
  - Addition of metrics..
  - Ratio among metrics, e.g.: num of TW/num of RTW,…
  - Cumulated metrics over time, e.g.: number of TW in the last X days..

- All: (i) per day, per hour, etc. (ii) per channel, per search
- Recently: we added the possibility of using metrics as firing conditions for alerts and bot on Twitter.

Problem addressed

Strong Limitations of the Search API of Twitter

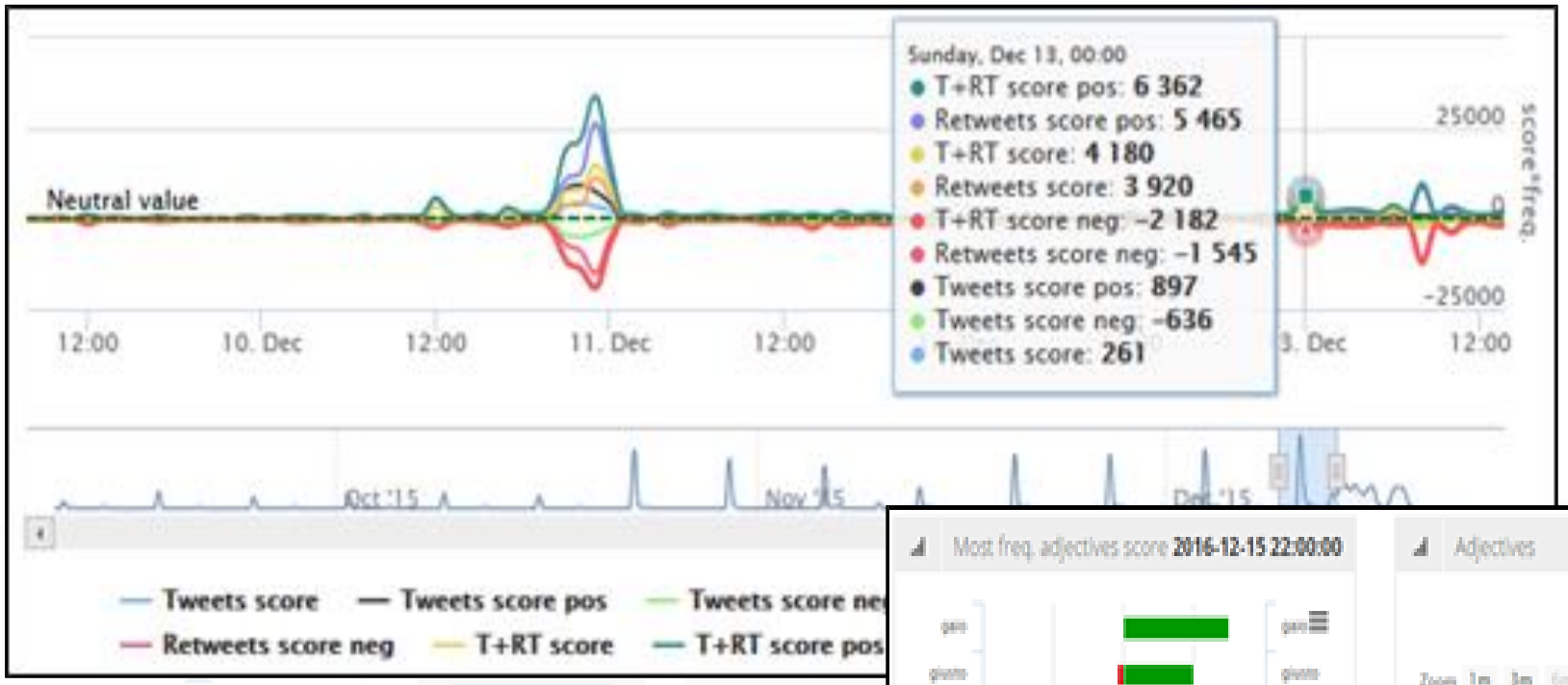
- minimizing the number of searches on the basis of the user requests:
  - different users with their queries request tweets already requested by others
- Recovering of parent Tweets from Orphan reTweets taken in the searching process

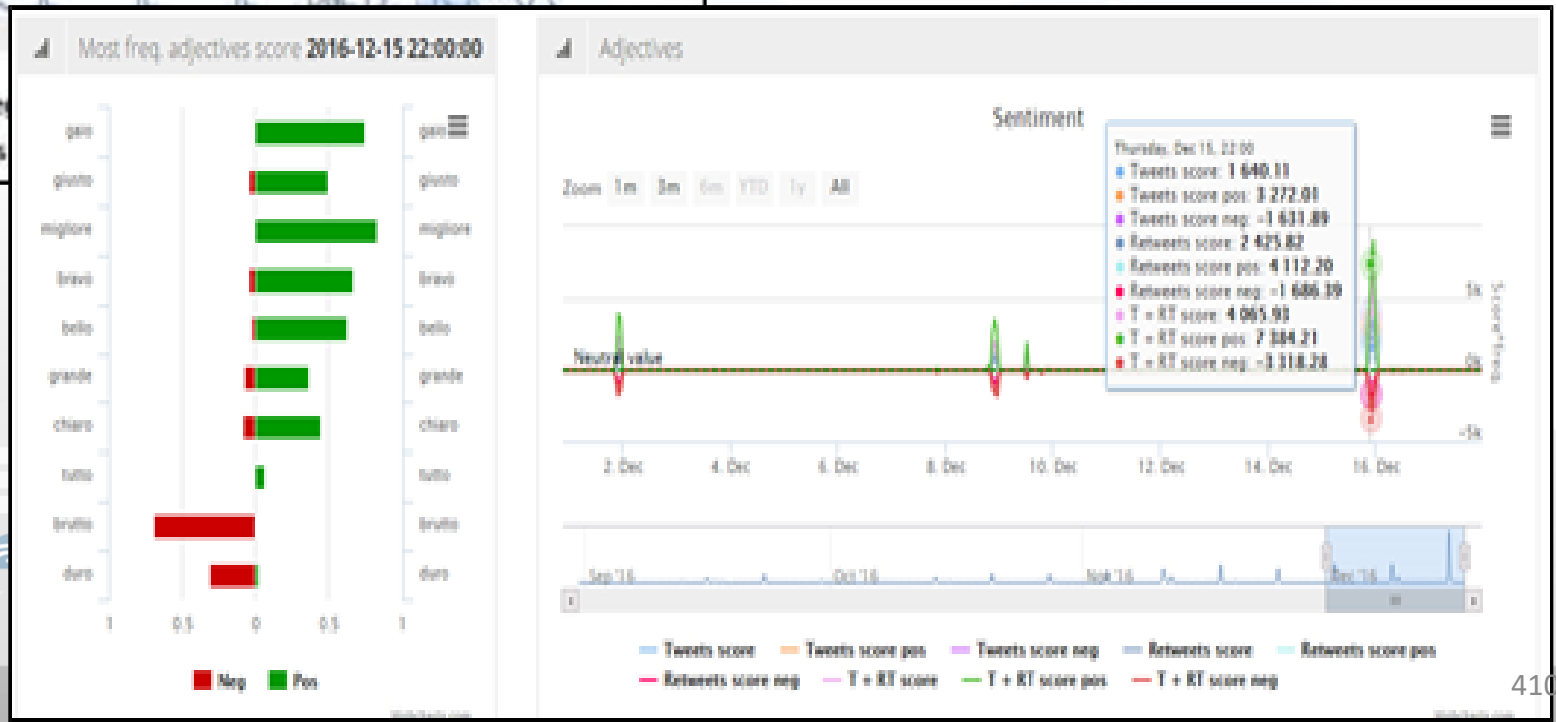
Analytics:

- High performance solution based on HDFS, Hadoop for NLP and SA, exploiting MapReduce programming model
- Estimating the network of influencer
- Computing metrics and prediction in real time.

Sentiment Analysis

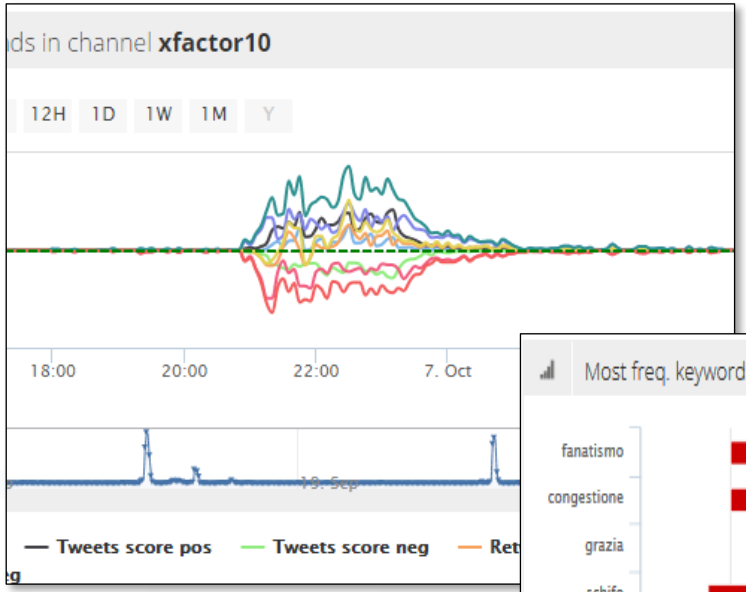
Sentiment Analysis

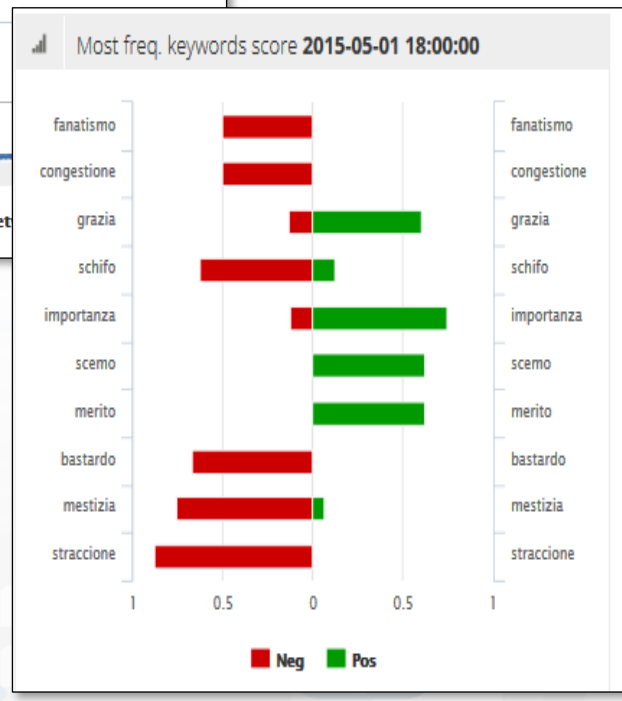


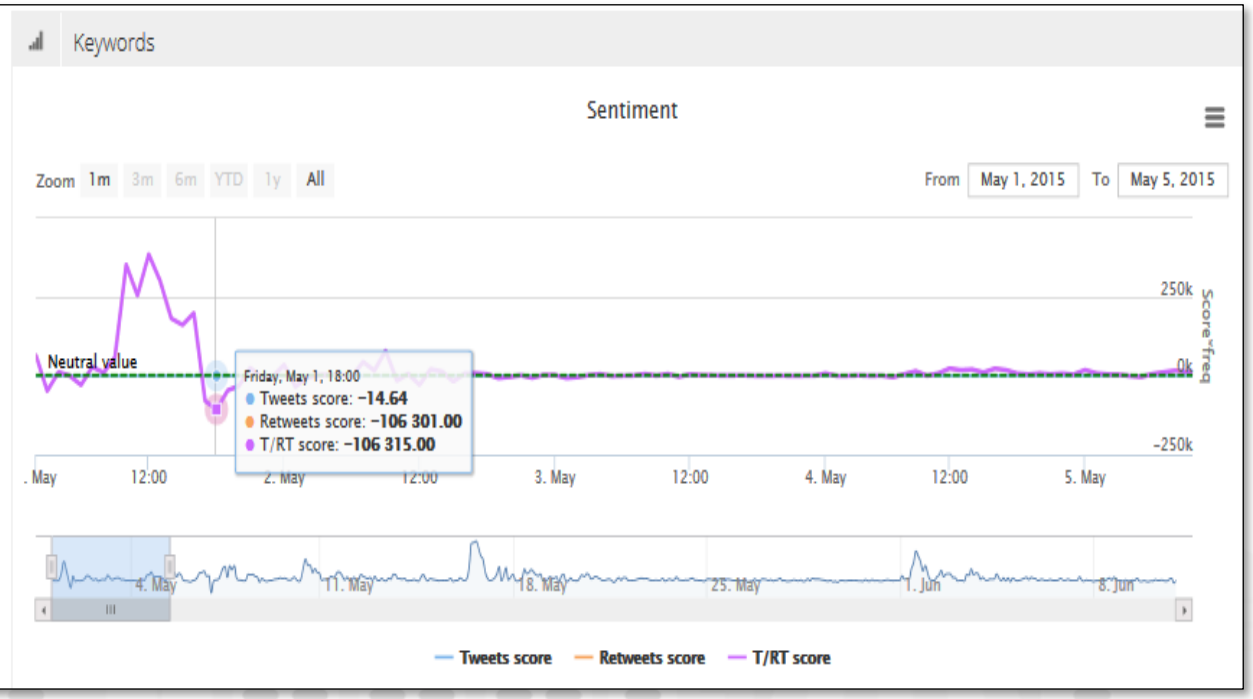


Twitter Vigilance

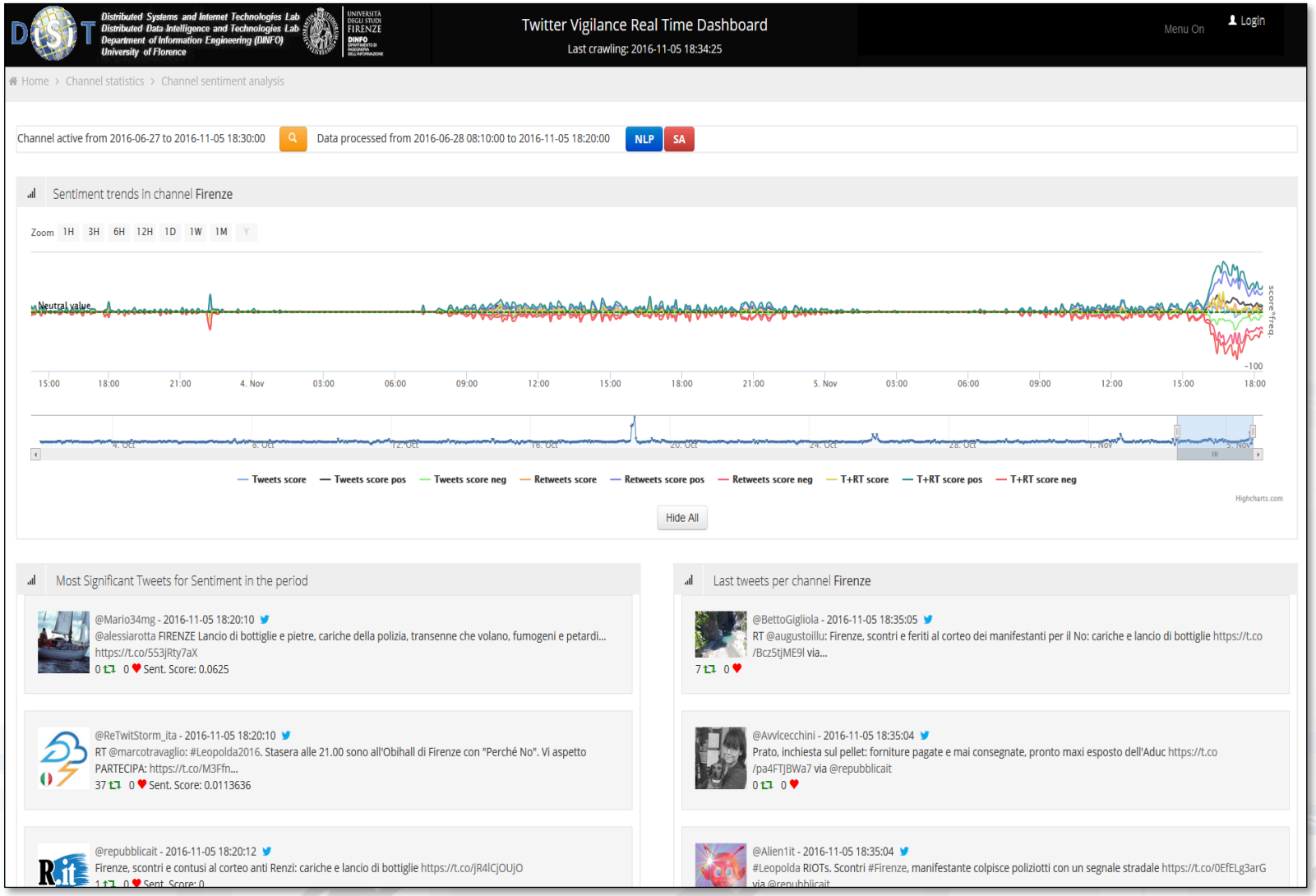
Typical trends on First positive-negative / negative-positive sentiments



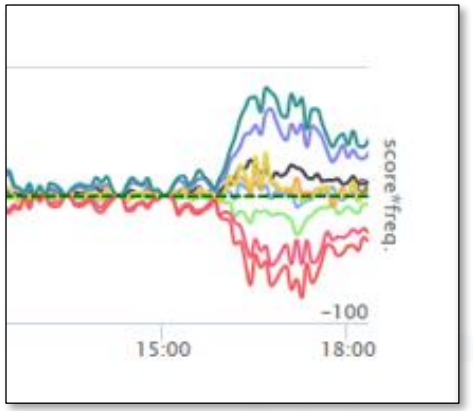


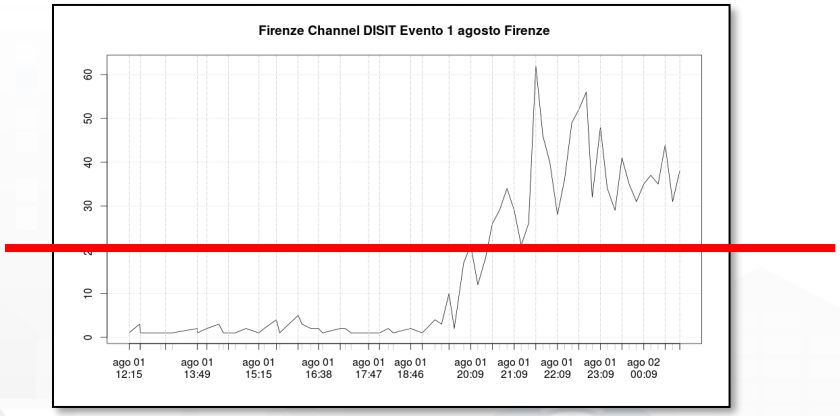


Real Time Twitter Vigilance, Early Warning



Sentiment Analysis





Twitter Vigilance

Twitter Vigilance su Firenze
(sperimentale)

Sat 24 Dec @ 10:37:57

Twitter: Volume di Tweet tramite Twitter Vigilance

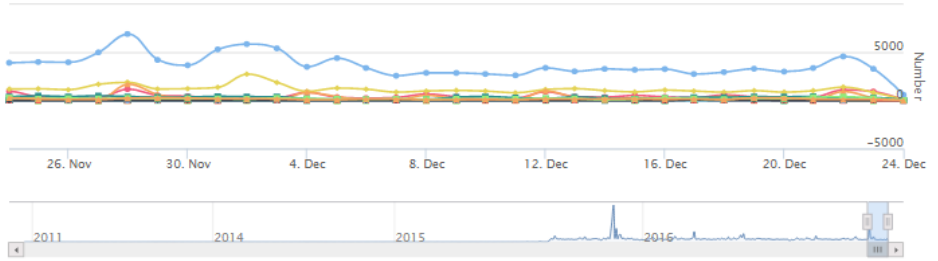
Twitter Vigilance Dashboard
Last crawling: 2016-12-24 12:34:51
Menu On | Login

Home > Channel statistics > Statistics on single Channel

Channel active from 2009-02-27 to today | Data processed from 2015-05-22 to 2016-12-02 | NLP | SA

Search related to channel **Firenze**

Zoom 1m 3m 6m YTD 1y All | From Nov 24, 2016 To Dec 24, 2016



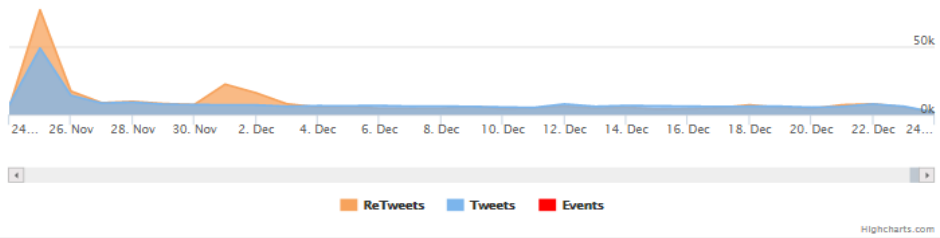
#Accademia #Aereoporto #Firenze #Aeroporto #Firenze #bargello #expo2015 #Firenze #fiorentina
#firenze #Florence #smartcity #fodd #ODDIT15 #Firenze #opendata firenze #pontevecchio
#uffizi @acffiorentina @BandieraiUffizi @Biblio_ComuneFi @bncfirenze @comunefi
@DarioNardella aereoporto @DiscoverTuscany @emergency_fi @firenzedigitale @firenzesulweb
@FirenzeTNews @museofirenze @museonovecento @Nazione_Firenze @NicolodiDaria @novedafirenze
@opentoscana @Pitti_Immagine @PoloMusealeFi @ProtCivComuneFi @prov_Fi_PC @santacroceopera
@STEEPproject @storianaturale @UNI_FIRENZE aereoporto firenze aeroporto firenze Firenze
firenze terremoto firenzesulweb Florence from:@UNI_FIRENZE lungarno medicea palazzopitti
palazzostrozzi ProtCivComuneFi smartcity firenze torrigiani uffizi voragine Events

Highcharts.com

Hide All

Daily mumber of tweets/retweets for channel: **Firenze**

Zoom 1m 3m 6m YTD 1y All | From Nov 24, 2016 To Dec 24, 2016



Highcharts.com

Sentiment Analysis in Tempo Reale su Firenze

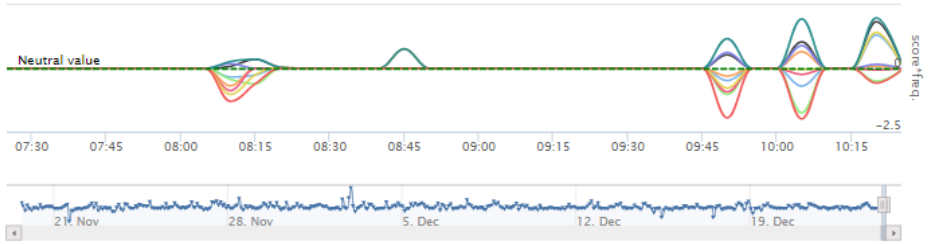
Twitter Vigilance Real Time Dashboard
Last crawling: 2016-12-24 10:32:33
Menu On | Login

Home > Channel statistics > Channel sentiment analysis

Channel active from 2016-06-27 to 2016-12-24 10:30:00 | Data processed from 2016-06-28 08:10:00 to 2016-12-24 10:25:00
NLP | SA

Sentiment trends in channel **Firenze**

Zoom 1H 3H 6H 12H 1D 1W 1M Y



Tweets score | Tweets score pos | Tweets score neg | Retweets score | Retweets score pos
Retweets score neg | T+RT score | T+RT score pos | T+RT score neg

Highcharts.com

Hide All

Most Significant Tweets for Sentiment in the period

@forealziesdayle - 2016-12-24 10:25:39
@Ahno_its_arno meh, I've only had average curry there
0 0 Sent. Score: 0.166667

@esercizistorici - 2016-12-24 10:30:01
Stanotte Concerto di Natale di Peter Guth @ORT_Toscana in collaborazione con Conservatorio 'L.Cherubini' di Firenze... https://t.co/uFeTp4wG1y
0 0 Sent. Score: 0

@infoitinterno - 2016-12-24 10:30:05
Firenze, un piano per un Natale sicuro (055firenze) https://t.co/vpqDAsEspi https://t.co/2Bn5k3zoyv

Last tweets per channel **Firenze**

@FormulaLatina - 2016-12-24 10:30:26
★ ★ ★ CAPODANNO LATINO Woodstock Club Firenze ★ ★ ★ H. 21:00 > > GRAN BUFFET DI SAN SILVESTRO CON BRINDISI E... https://t.co/cM84Gw65sO
0 0

@MORINEMMANUELL2 - 2016-12-24 10:30:14
RT @Navymat: Francesco #Salviati Carità,1543 #Manierisme Galleria degli Uffizi, Firenze @ChevernyM @mariaireneali @pieroBENEDETTO @Giusepp...
6 0

@infoitinterno - 2016-12-24 10:30:12

Twitter Citazioni

TRENDS | QUOTES

@NICOLODIDARIA
@COMUNEFI
@ACFFIORENTINA
@NAZIONE_FIRENZE
@SANTACROCEOPERA
@MUSEONOVECENTO
@FIRENZEDIGITALE

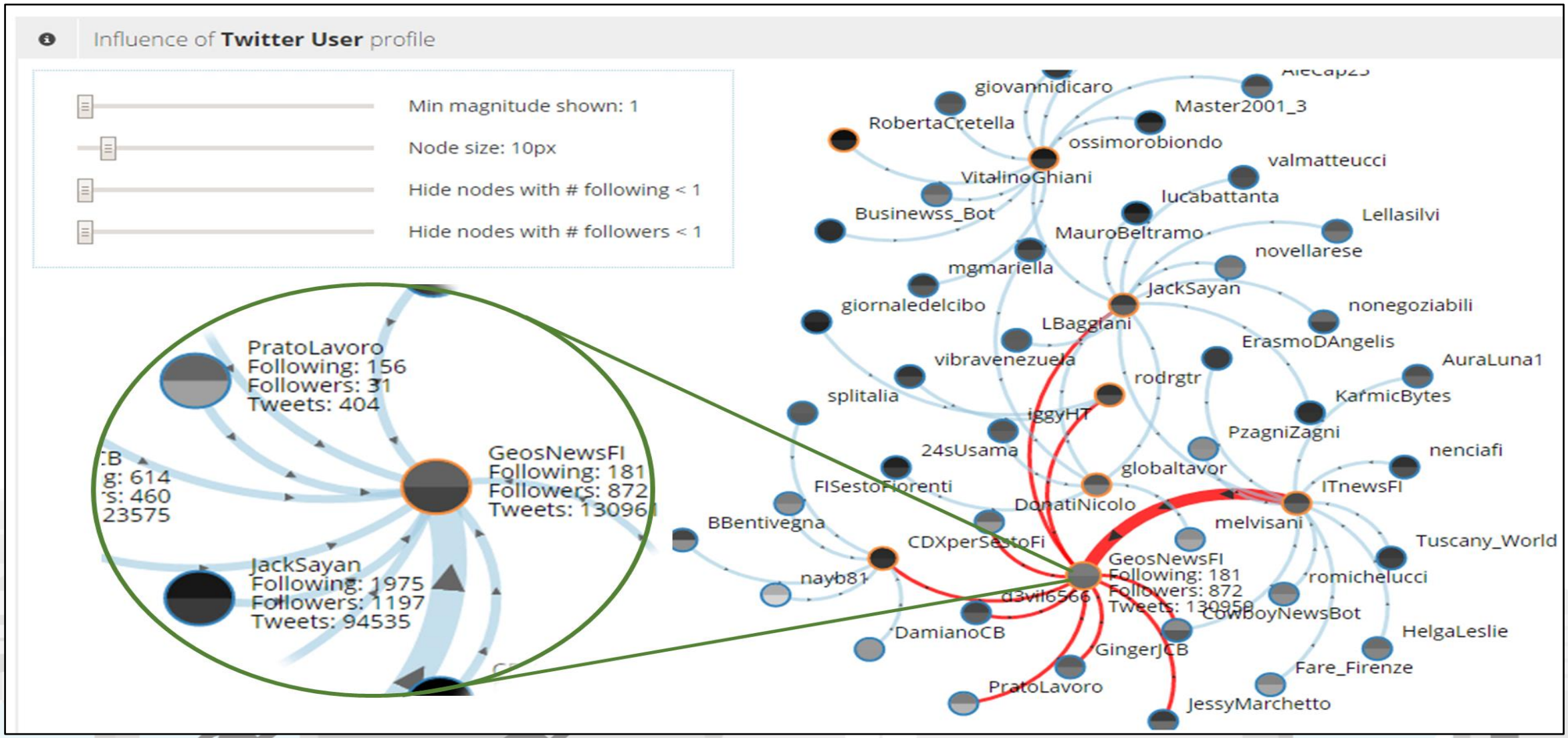
Twitter Hashtag trend

TRENDS | QUOTES

#FIRENZE
#FIORENTINA
#UFFIZI
#PONTEVECCHIO

We suggest to use Chrome browser for better experience

Influence Network

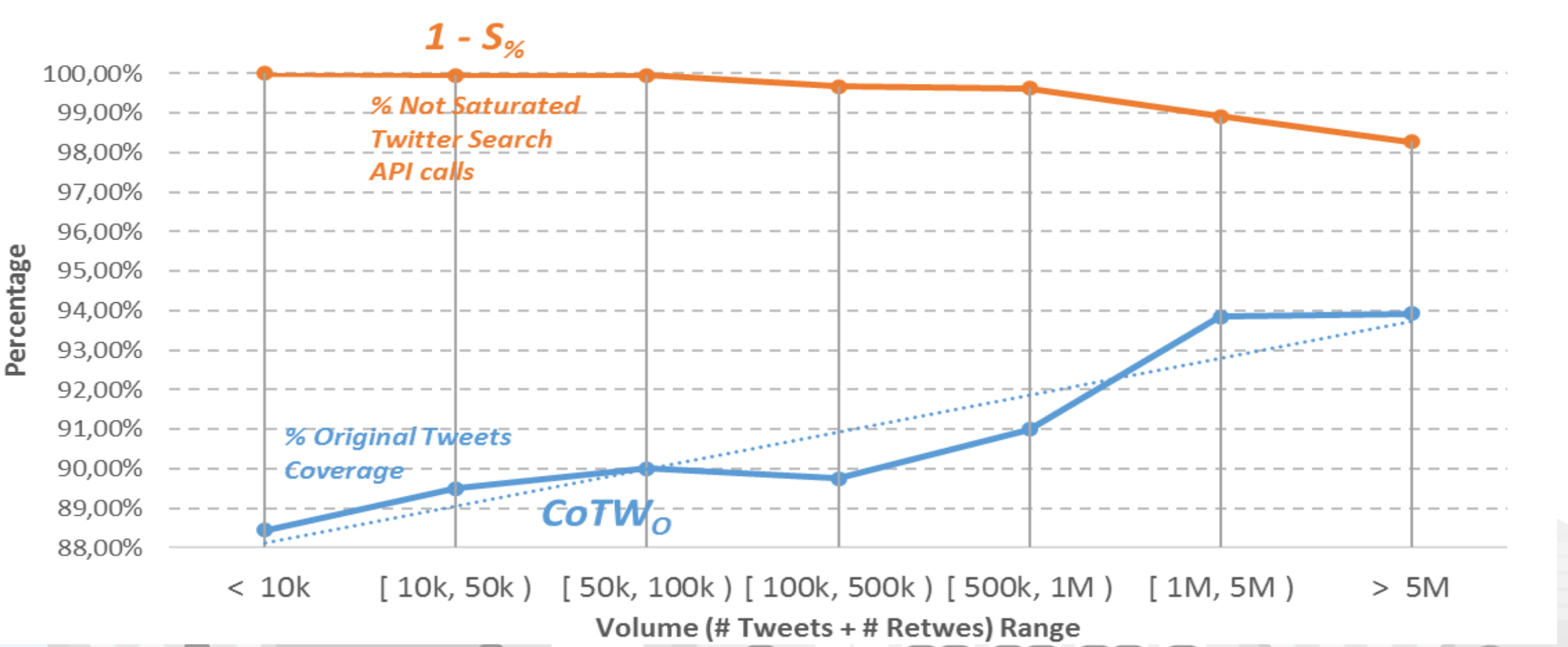


Reliability in collecting tweets

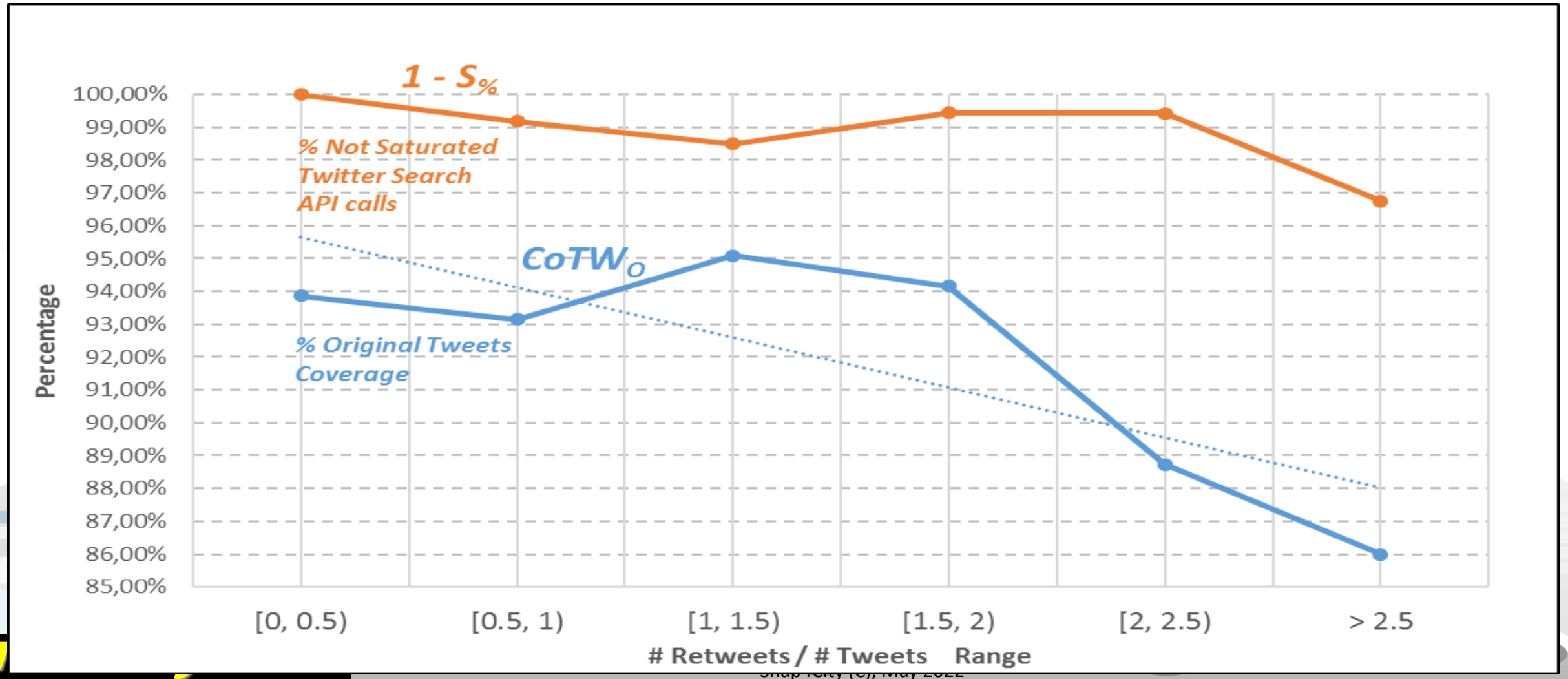
Efficiency in retrieval

Posts Volume (Tweets + Retweets) Range	*# Recovered Original Tweets*	*# Missing Original Tweets*	*% Original Tweets Coverage ($CoTW_O$)*	*# Twitter Search API requests*	*# Saturations on Twitter Search API requests*	*% Saturations on Twitter Search API requests ($S_{\%}$)*	*% Not-Saturated Twitter Search API requests (1-$S_{\%}$)*
< 10k	18571	2033	89,05%	124299	1	0,00%	100,00%
[10k, 50k)	130051	13716	89,45%	399170	100	0,03%	99,97%
[50k, 100k)	96171	10278	89,31%	123804	165	0,13%	99,87%
[100k, 500k)	997833	86755	91,31%	849062	1589	0,19%	99,81%
[500k, 1M)	930646	61632	93,38%	439956	1998	0,45%	99,55%
[1M, 5M)	6454463	439628	93,19%	2787485	31585	1,13%	98,87%
> 5M	14714124	899035	93,89%	4509184	64284	1,43%	98,57%

Original Tweets coverage and Twitter Search API



Dependance on RTW/TW ratio



Tweets as Early Warning

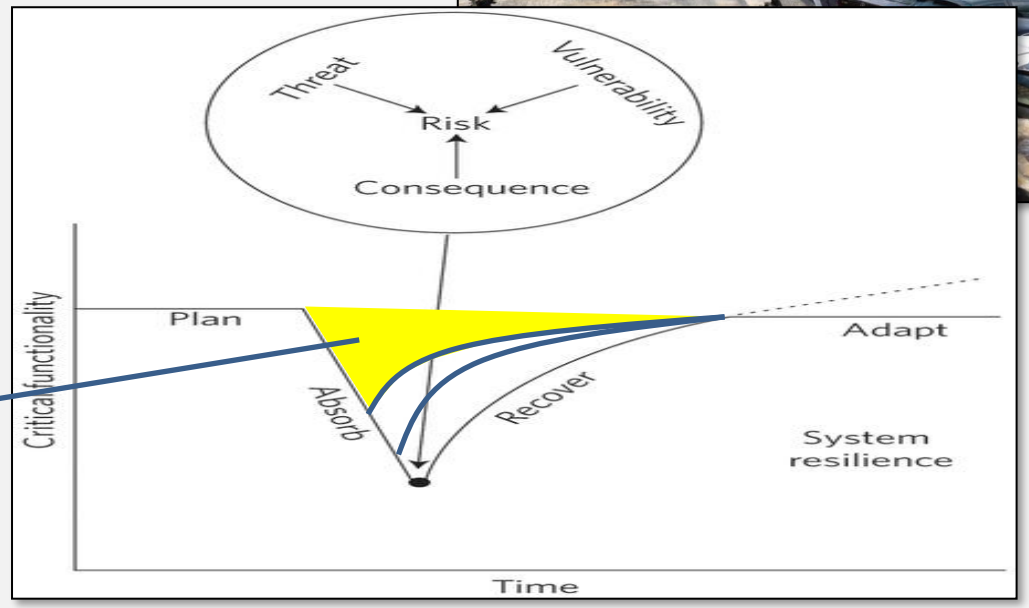
Early warning, detection

resolute

City Resilience

- **Issue:**
  - **Detection of critical condition**
  - **Not easily detected with other means**
- **Impact:**
  - **Early warning, faster reaction**
  - **Increased resilience**
- **Several metrics related to**
  - **Volume of retweets**
  - **Sentiment analysis**

Prepare
Absorb
Recover
Adapt



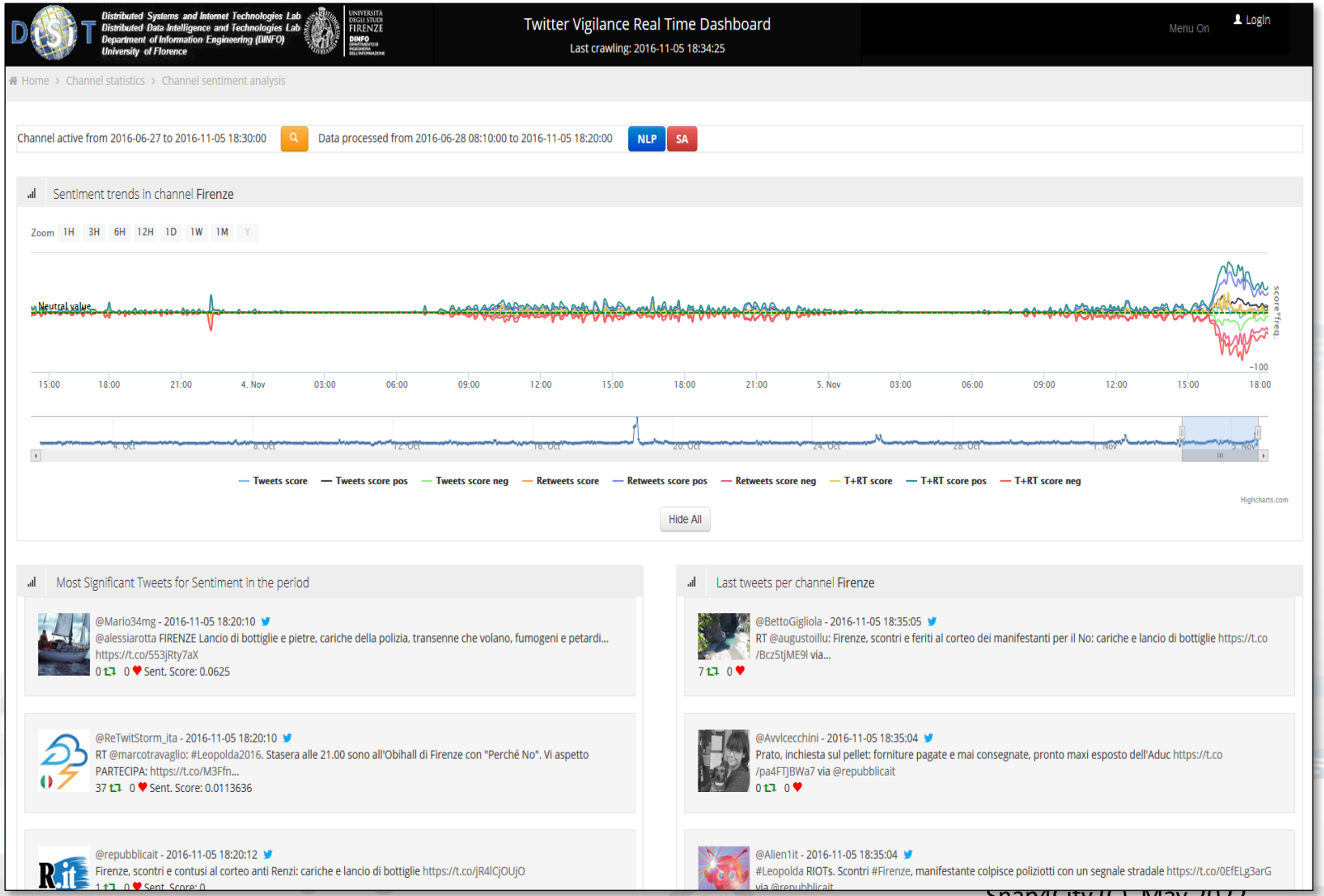
damage

UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB

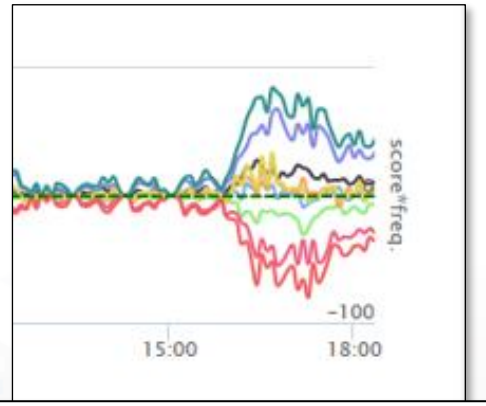
Twitter Vigilance

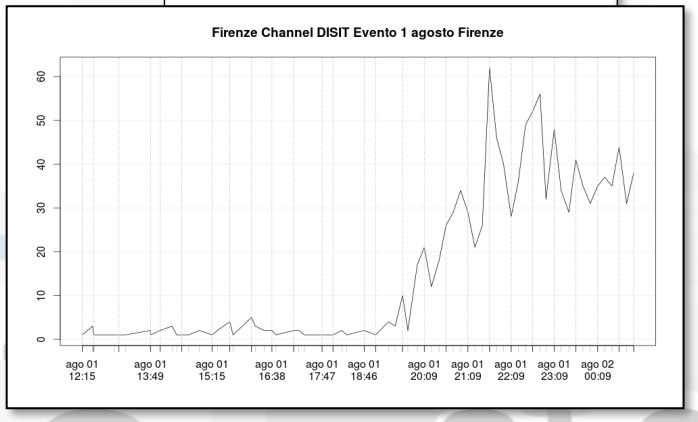
resolute

Twitter Vigilance RT: sentiment analysis



Real time
Early Warning

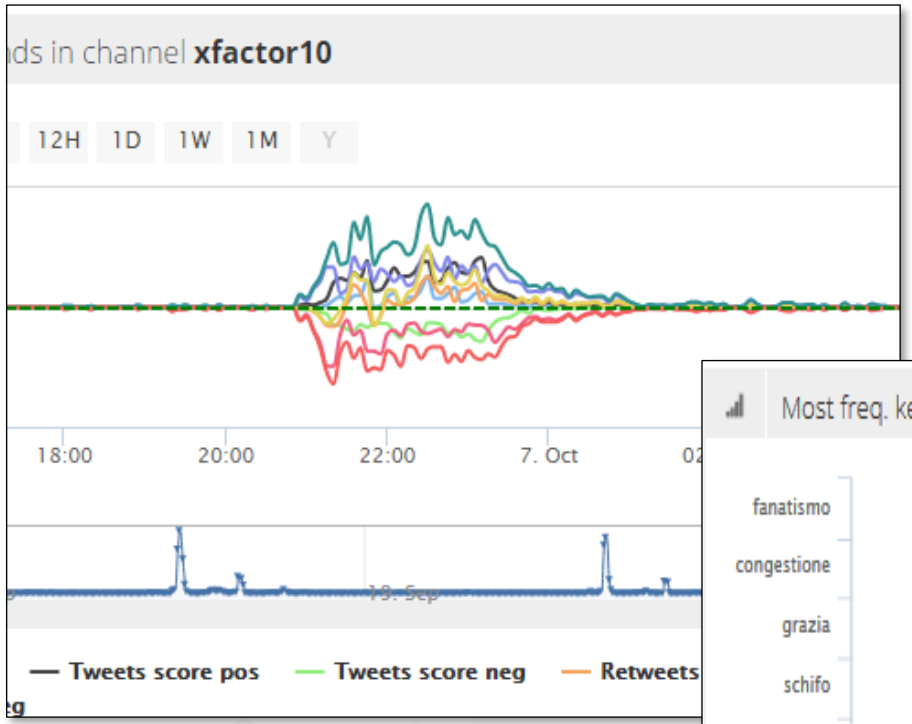


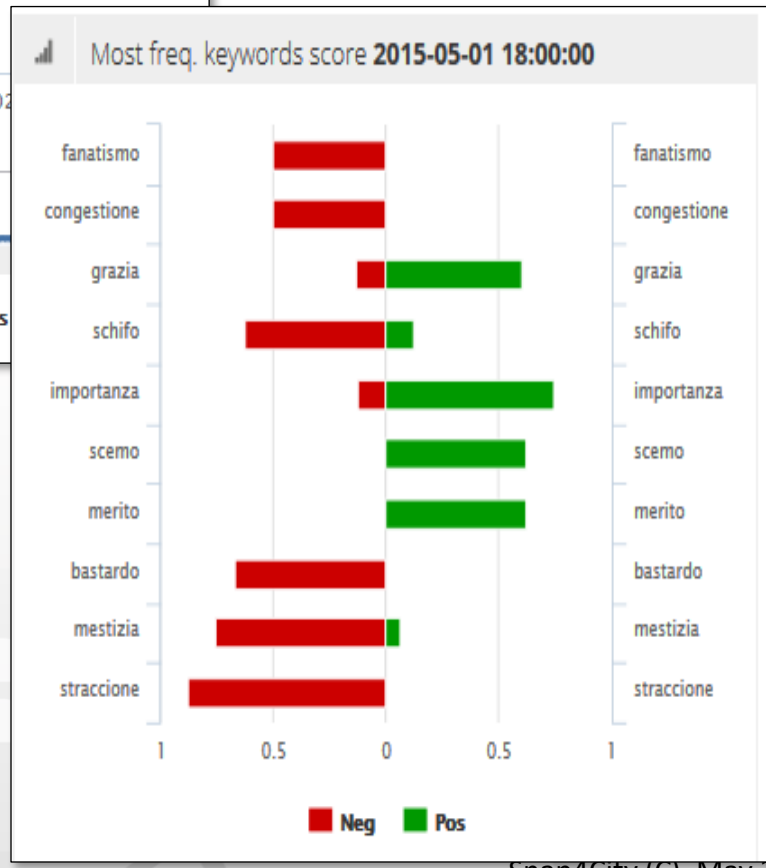


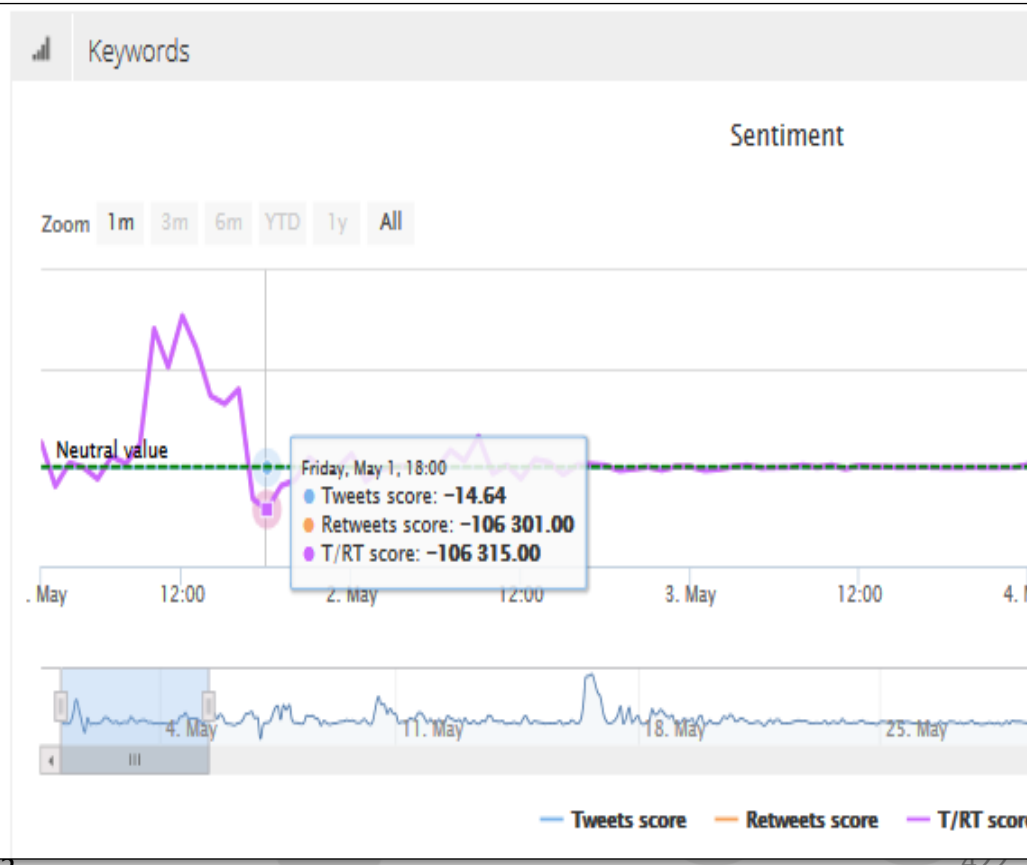
Sentiment Analysis

Languages

- Italian
- English
- French
- Spanish

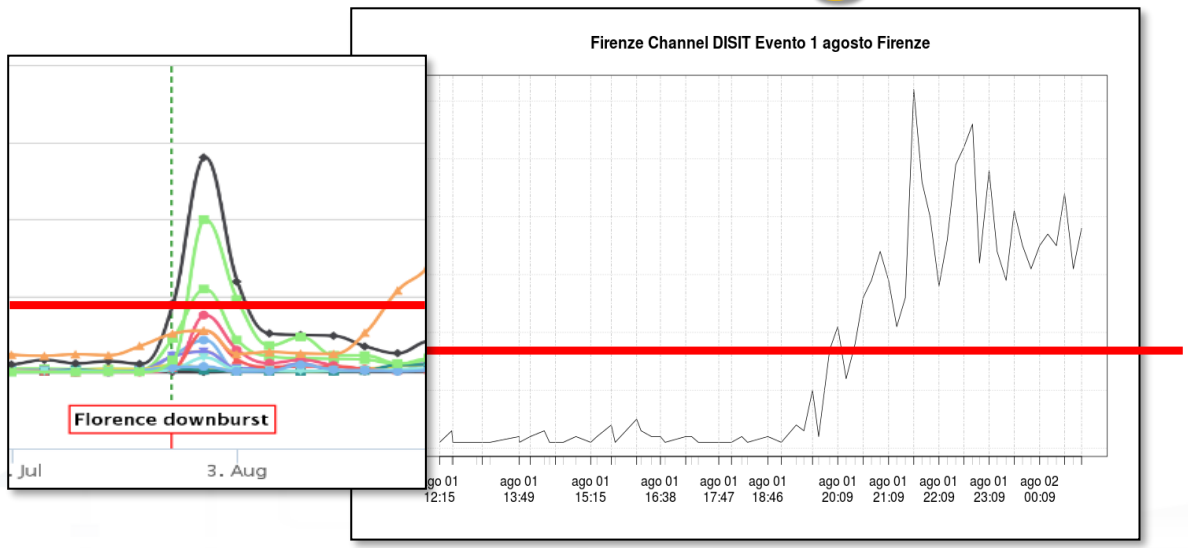


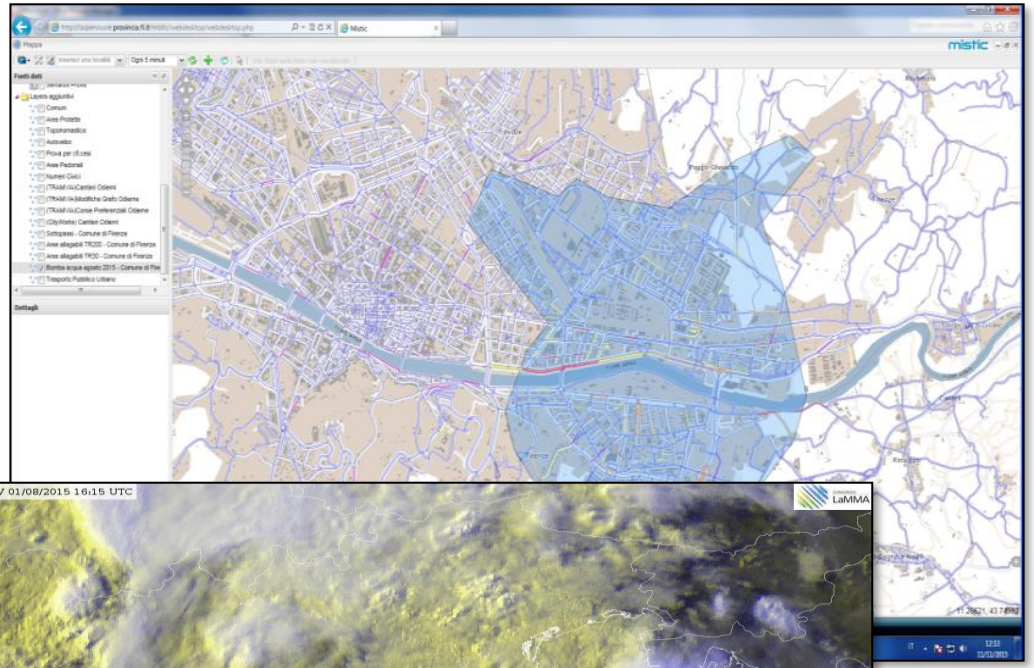


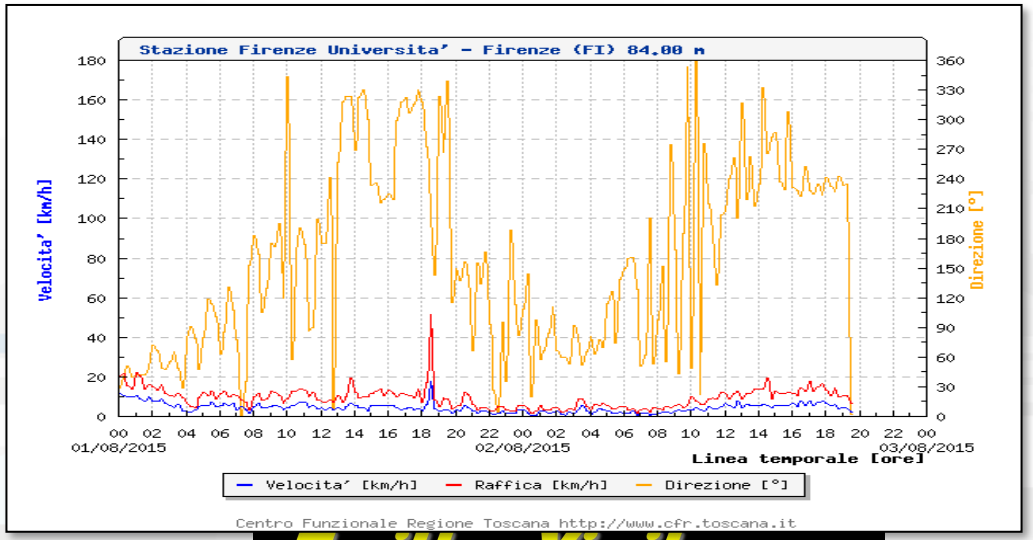


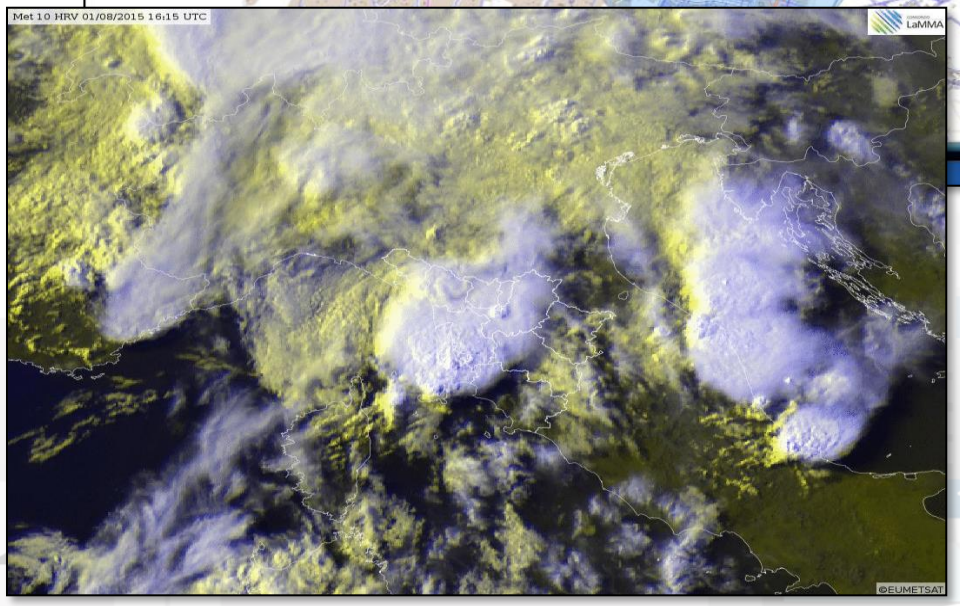
Early Warning
Twitter Vigilance and Water Bomb

resolute





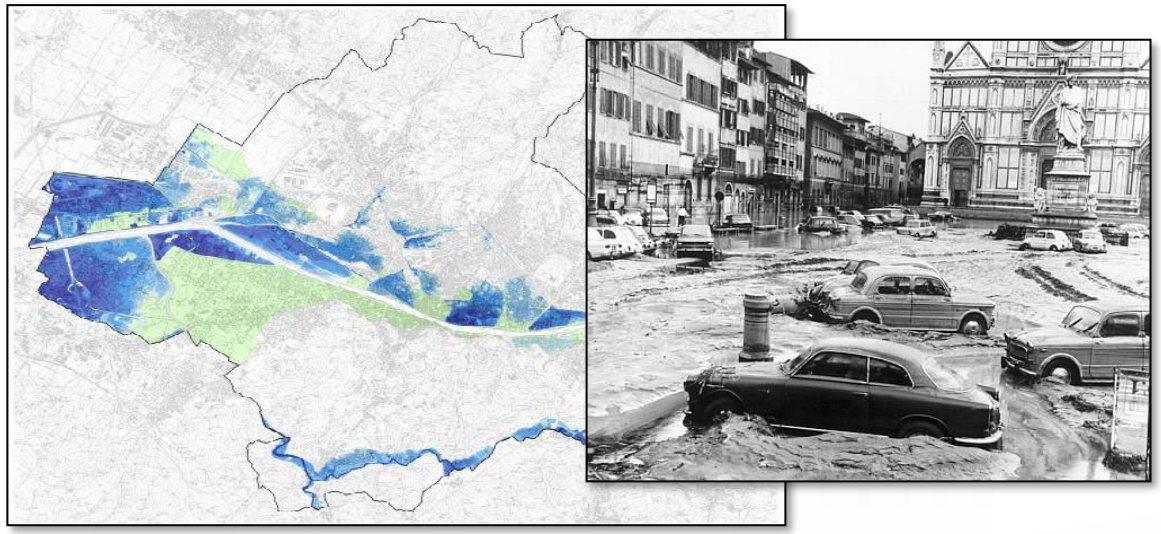




Twitter Vigilance

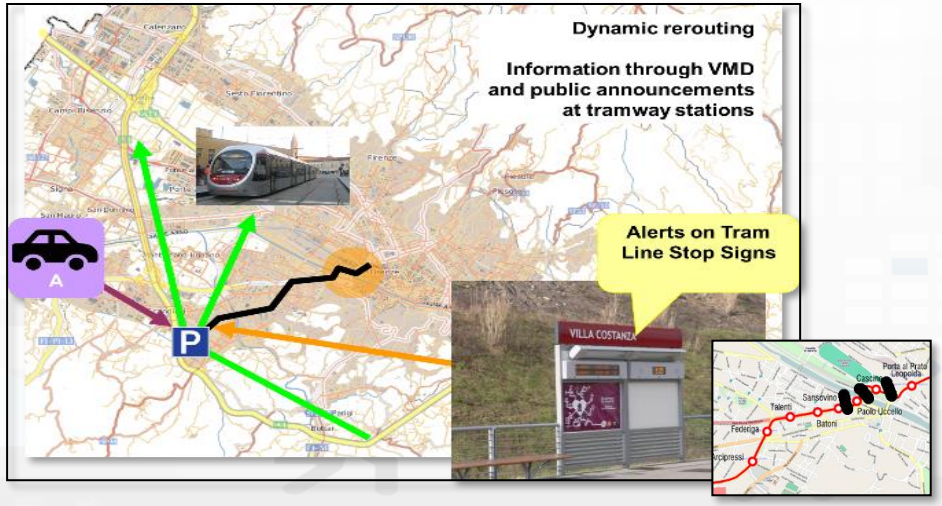
City Resilience ERMG

200 years probability Arno flooding



30 years probability Arno flooding

Arno Flood Impact on Tram Line & Traffic



Water bomb (down burst) in South Florence

Case Study D

Twitter Vigilance su Firenze
(sperimentale)

Sat 24 Dec @ 10:37:57

Twitter: Volume di Tweet tramite Twitter Vigilance

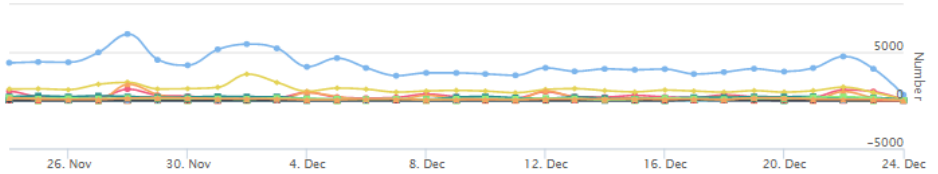
Twitter Vigilance Dashboard
Last crawling: 2016-12-24 12:34:51
Menu On | Login

Home > Channel statistics > Statistics on single Channel

Channel active from 2009-02-27 to today | Data processed from 2015-05-22 to 2016-12-02 | NLP | SA

Search related to channel **Firenze**

Zoom 1m 3m 6m YTD 1y All | From Nov 24, 2016 To Dec 24, 2016



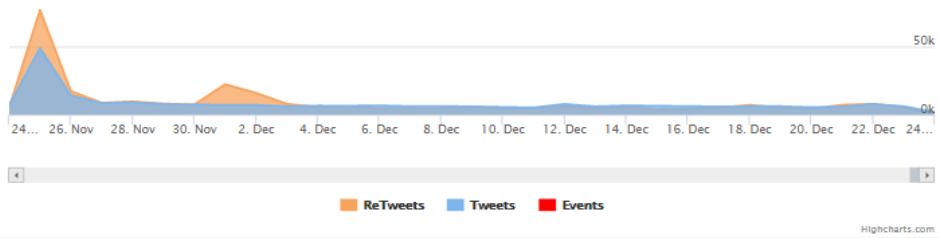
#Accademia #Aereoporto #Firenze #Aeroporto #Firenze #bargello #expo2015 #Firenze #fiorentina
#firenze #Florence #smartcity #fodd #ODDIT15 #Firenze #opendata firenze #pontevecchio
#uffizi @acffiorentina @BandieraiUffizi @Biblio_ComuneFi @bncfirenze @comunefi
@DarioNardella aereoporto @DiscoverTuscany @emergency_fi @firenzedigitale @firenzesulweb
@FirenzeTNews @museofirenze @museonovecento @Nazione_Firenze @NicolodiDaria @novedafirenze
@opentoscana @Pitti_Immagine @PoloMusealeFi @ProtCivComuneFi @prov_Fi_PC @santacroceopera
@STEEPproject @storianaturale @UNI_FIRENZE aereoporto firenze aeroporto firenze Firenze
firenze terremoto firenzesulweb Florence from:@UNI_FIRENZE lungarno medicea palazzopitti
palazzostrozzi ProtCivComuneFi smartcity firenze torrigiani uffizi voragine Events

Highcharts.com

Hide All

Daily mumber of tweets/retweets for channel: **Firenze**

Zoom 1m 3m 6m YTD 1y All | From Nov 24, 2016 To Dec 24, 2016



Highcharts.com

Sentiment Analysis in Tempo Reale su Firenze

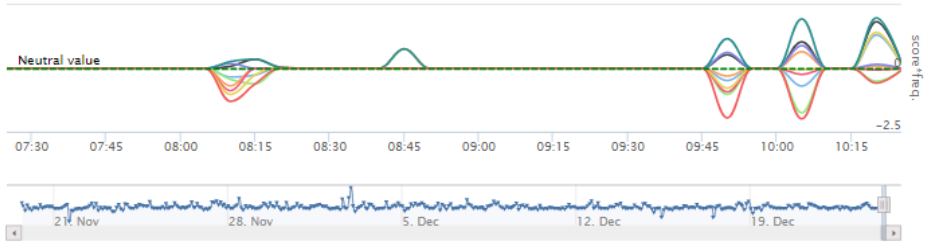
Twitter Vigilance Real Time Dashboard
Last crawling: 2016-12-24 10:32:33
Menu On | Login

Home > Channel statistics > Channel sentiment analysis

Channel active from 2016-06-27 to 2016-12-24 10:30:00 | Data processed from 2016-06-28 08:10:00 to 2016-12-24 10:25:00
NLP | SA

Sentiment trends in channel **Firenze**

Zoom 1H 3H 6H 12H 1D 1W 1M Y



Tweets score | Tweets score pos | Tweets score neg | Retweets score | Retweets score pos
Retweets score neg | T+RT score | T+RT score pos | T+RT score neg

Highcharts.com

Hide All

Most Significant Tweets for Sentiment in the period

@forealziesdayle - 2016-12-24 10:25:39
@Ahno_its_arno meh, I've only had average curry there
0 0 Sent. Score: 0.166667

@esercizistorici - 2016-12-24 10:30:01
Stanotte Concerto di Natale di Peter Guth @ORT_Toscana in collaborazione con Conservatorio 'L.Cherubini' di Firenze... https://t.co/uFeTp4wG1y
0 0 Sent. Score: 0

@infoitinterno - 2016-12-24 10:30:05
Firenze, un piano per un Natale sicuro (055firenze) https://t.co/vpqDAsEspi https://t.co/2Bn5k3zoyv

Last tweets per channel **Firenze**

@FormulaLatina - 2016-12-24 10:30:26
★ ★ ★ CAPODANNO LATINO Woodstock Club Firenze ★ ★ ★ H. 21:00 > > GRAN BUFFET DI SAN SILVESTRO CON BRINDISI E... https://t.co/cM84Gw65sO
0 0

@MORINEMMANUELL2 - 2016-12-24 10:30:14
RT @Navymat: Francesco #Salviati Carità,1543 #Manierisme Galleria degli Uffizi, Firenze @ChevernyM @mariaireneali @pieroBENEDETTO @Giusepp...
6 0

@infoitinterno - 2016-12-24 10:30:12

Twitter Citazioni

TRENDS | QUOTES

@NICOLODIDARIA
@COMUNEFI
@ACFFIORENTINA
@NAZIONE_FIRENZE
@SANTACROCEOPERA
@MUSEONOVECENTO
@FIRENZEDIGITALE

Twitter Hashtag trend

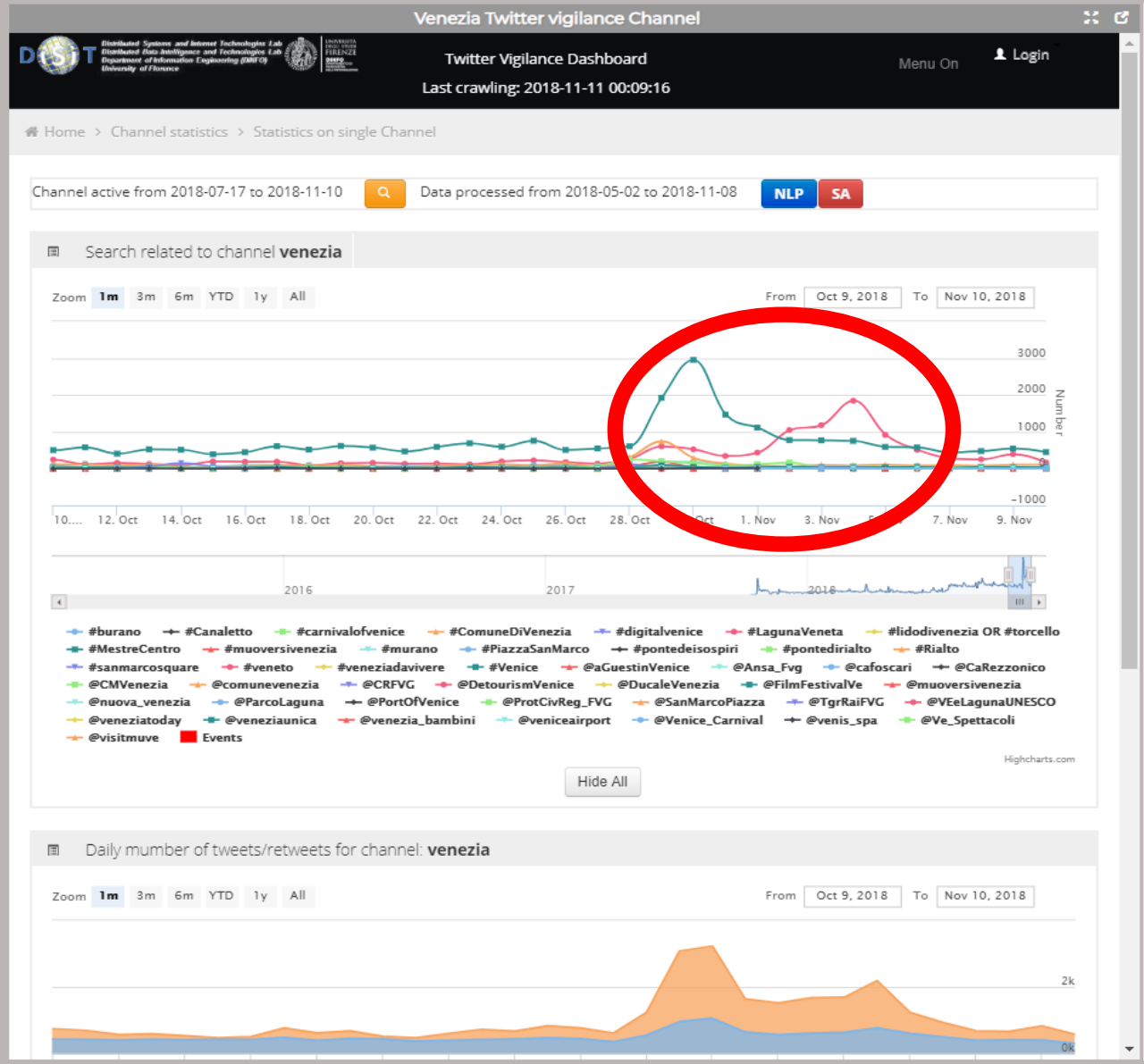
TRENDS | QUOTES

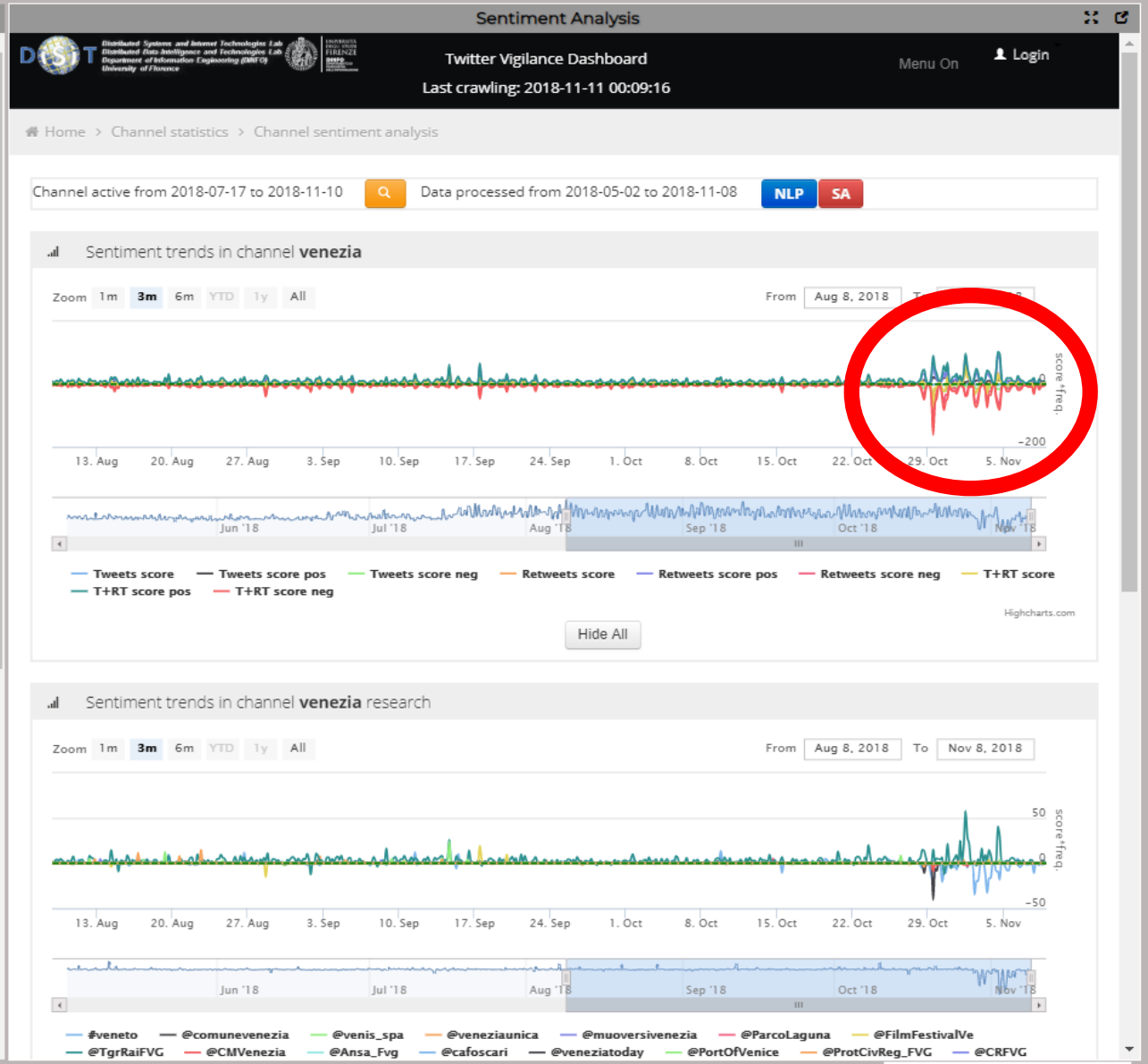
#FIRENZE
#FIORENTINA
#UFFIZI
#PONTEVECCHIO

We suggest to use Chrome browser for better experience

Venezia Social - Twitter Vigilance

Sun 11 Nov 00:09:40



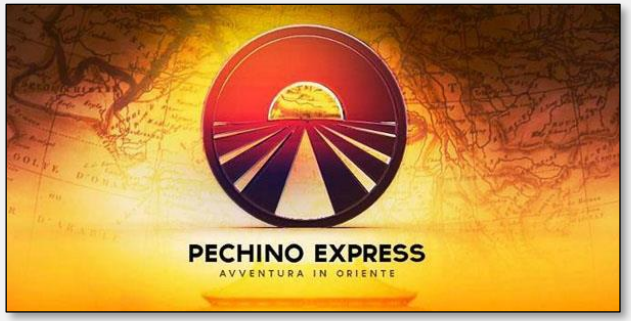


https://main.snap4city.org/view/index.php?iddasboard=MTIxOA==

Reliability Audience on TV programs

Predicting Audience on Social intensive TV show

- **Issue:**
  - **How to predict the number of people following a TV reality show in life**
- **Impact:**
  - **Making Advertising, promotion**
  - **Valorizing advertising**
  - **Adjusting the show**
- **Several metrics related to**
  - **Structure of volume of TW, RTW**
  - **Features of the tweet authors**
  - **Relationships ….**



- Periodic events
- Specific rules
- Strong influence and user engagment
- Audience can vote
- Audience espress appreciation and rejects
- .. Similar to the presence at large and log terms event, such as EXPO2015

Case Study B2

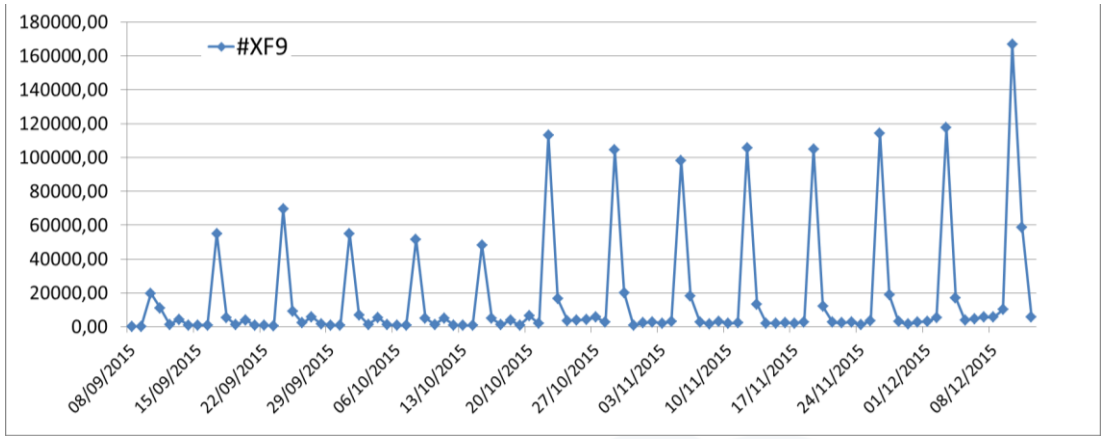
Twitter Metrics

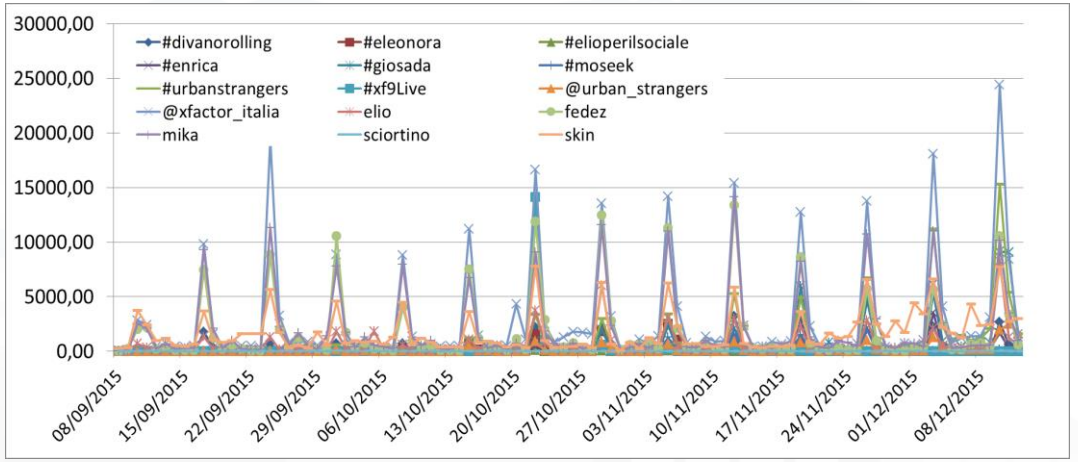
- TW: Number of Tweets per **Search/Channel** (as called Volume) , per day, per hour
- RTW: Number of ReTweets per **Search/Channel**, per day, per hour
- NRT/TW: ratio from ReTweets and Tweets per **Search/Channel**, per day, per hour
- NumSearch: number of Tweets including the Search per **Channel,** per day, per hour
- Sentiment Analysis Score per **Search/Channel**, per day, per hour
- Num of xxxxx

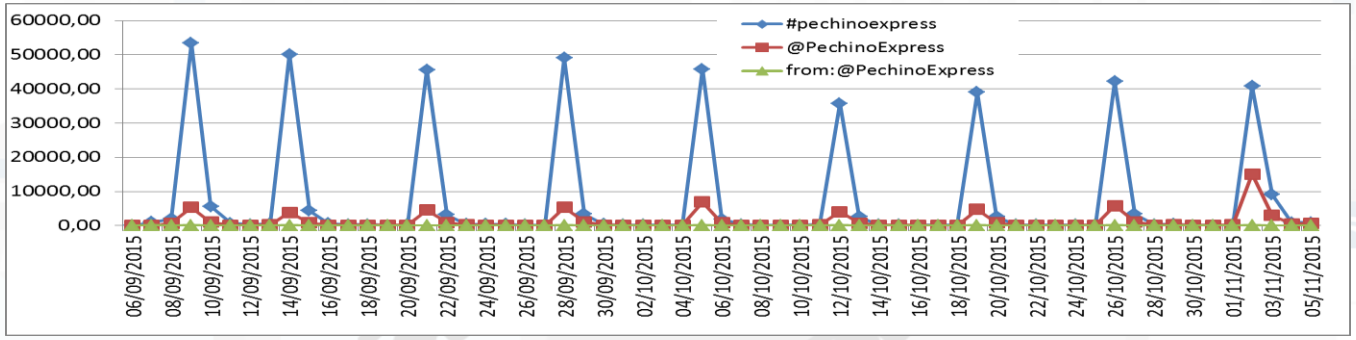
Predicting Audience: X-Factor, PechinoExpress, …

- Trend of TW and RTW for X-Factor 9
  - Several searches
- Similar model for other Social Intensive TV shows
  - See below Pechino Express

$$x_t = \beta_1 z_{1,t} + \beta_2 z_{2,t} + \beta_3 z_{3,t} + \cdots + \beta_k \, z_{k,t} + n$$



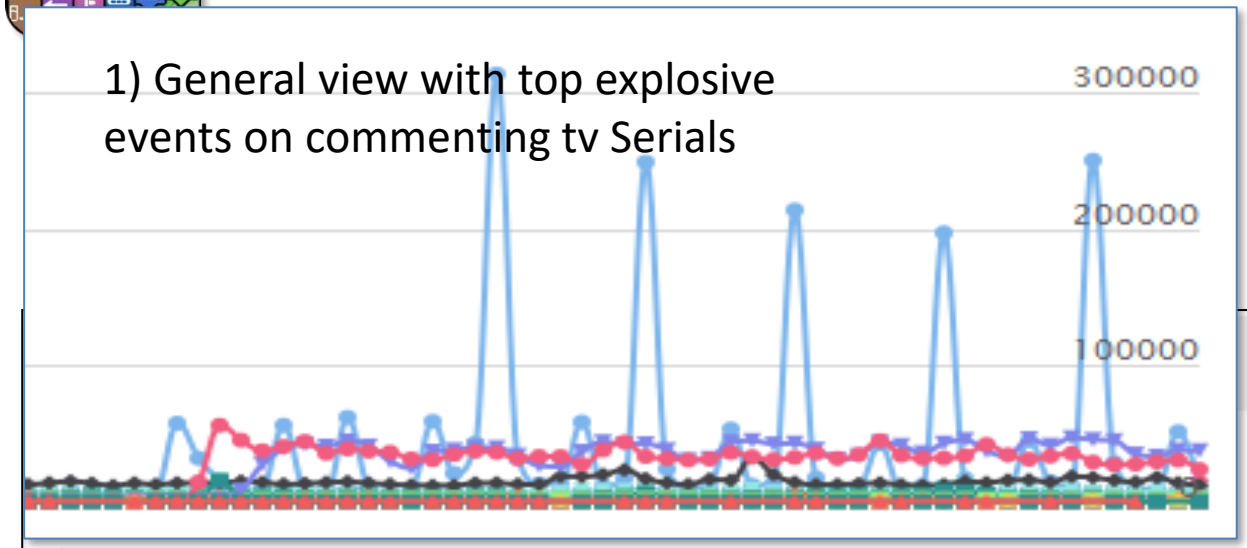


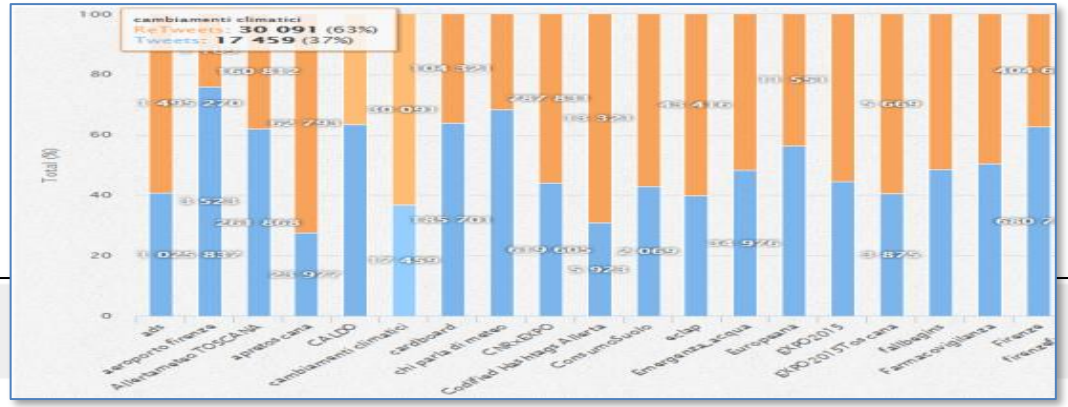


Twitter Vigilance

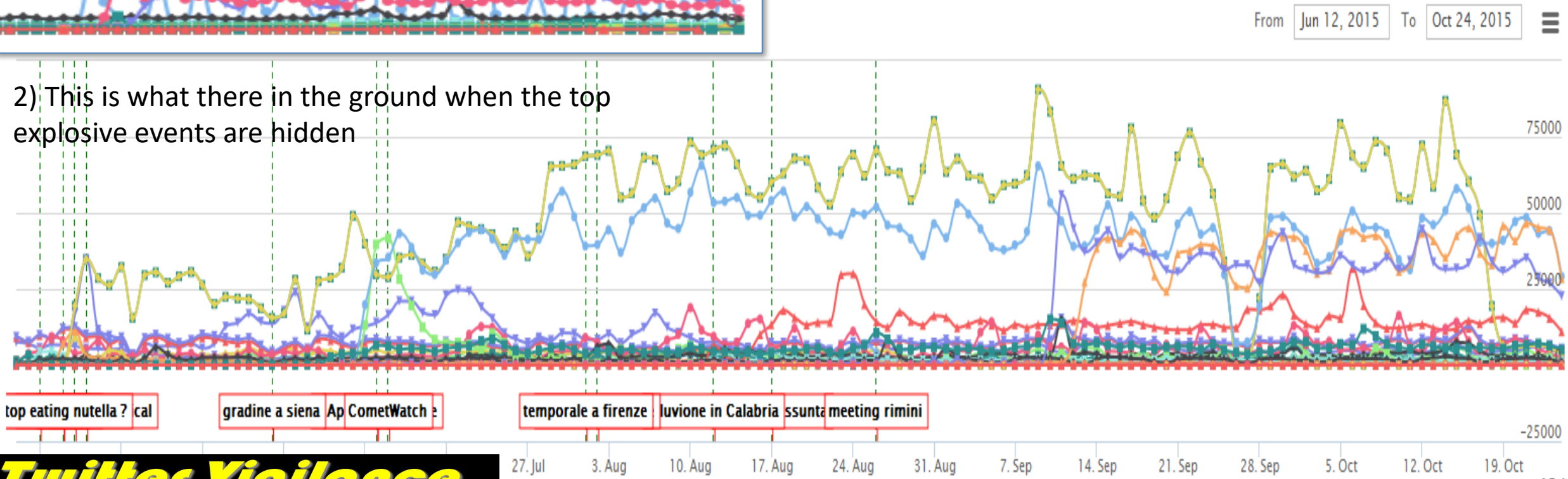
All Channels (private information)

1) General view with top explosive events on commenting tv Serials





2) This is what there in the ground when the top explosive events are hidden

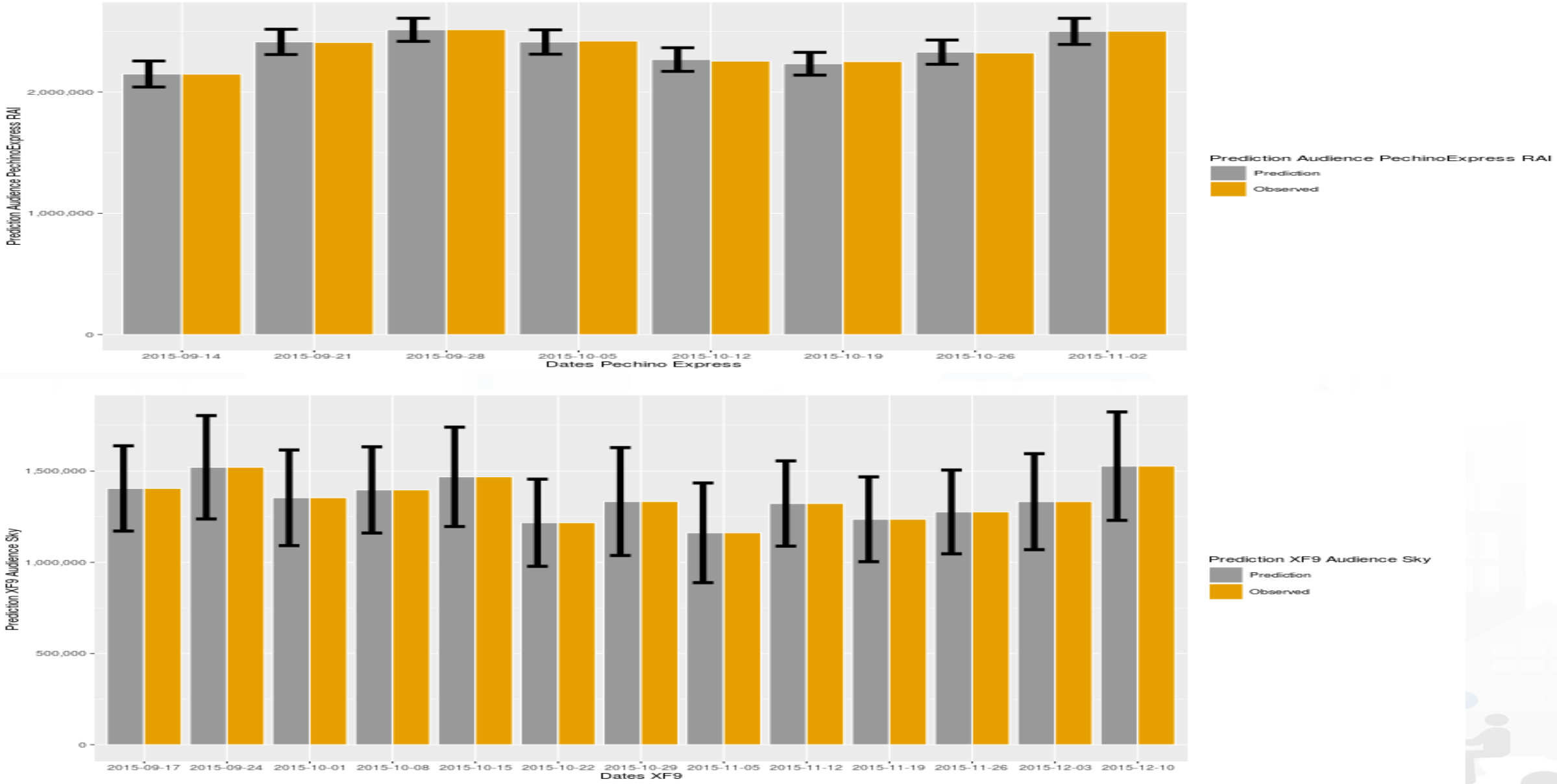


Twitter Vigilance

Details of Predictive Models Validities

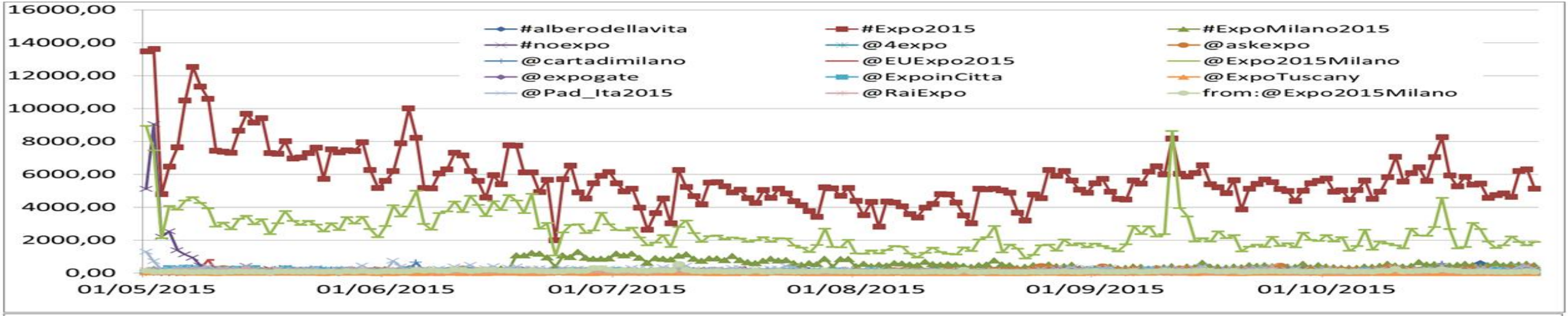
Metrics collected over the 5 days before the event.		X-Factor 9 - Model				Pechino Express - Model			
		Coeff	Std Err	t-val	p-val	Coeff	Std Err	t-val	p-val
Total number of tweets + retweets on main hashtag	β_1	-73.48	58.49	-1.256	0.2494	-954.3	64.69	-14.750	0.0045
Total number of tweets on main hashtag,	β_2	122.7	70.27	1.745	0.1244	4144	284	14.590	0.0046
Ratio between: number of RTW/TW on main hashtag,	β_3	1358851	462704	2.937	0.0218	937920	80946	11.590	0.0073
UnqURetweet	β_4	264.3	153	1.728	0.1277	2175	345.6	6.293	0.0243
FUnqUsers	β_5	-214.9	132.5	-1.622	0.1488	-1640	270.6	-6.061	0.0261
Intercept	n	-762730	627238	-1.216	0.2634	-2560461	401675	-6.374	0.0237
R squared		0.727				0,995			
RMSE		66467				8851			
MAE		55589				6805			
AIC		340				182			
TV broadcasting company		Sky				RAI			
Weeks		13				9			
millions of registered tweets on Twitter Vigilance		1.625				0.455			

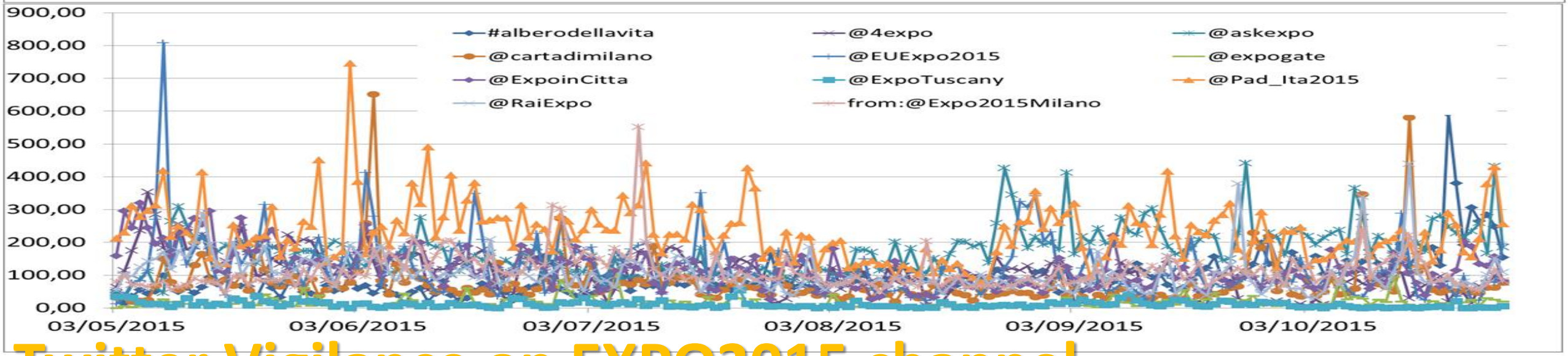
Predicting Confidence



Reliability presences to major events

Predicting EXPO2015





Twitter Vigilance on EXPO2015 channel

Twitter Vigilance

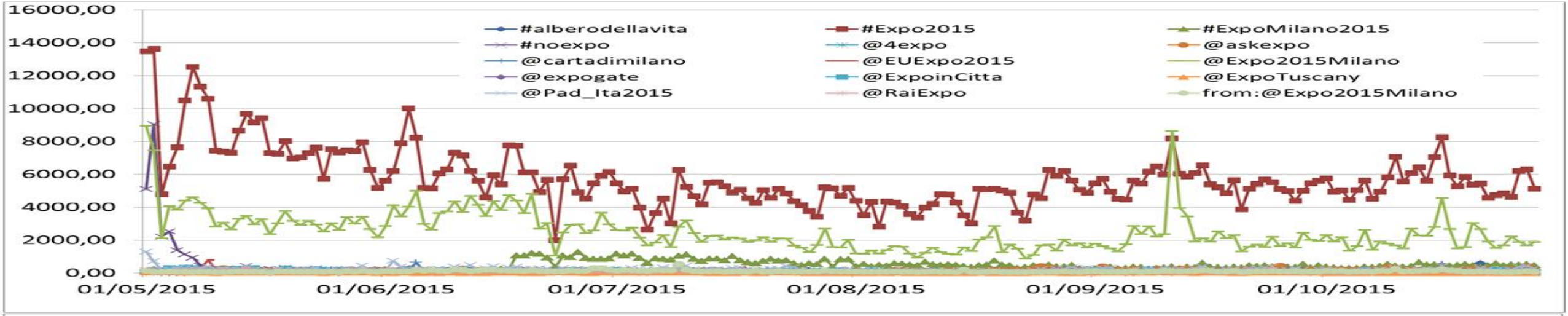
Twitter Metrics

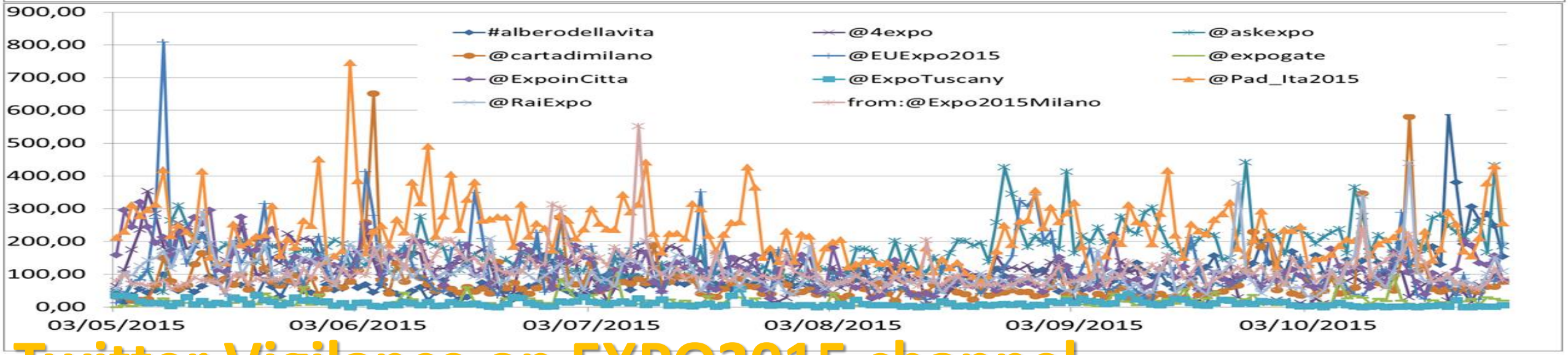
- TW: Number of Tweets per **Search/Channel** (as called Volume) , per day, per hour
- RTW: Number of ReTweets per **Search/Channel**, per day, per hour
- NRT/TW: ratio from ReTweets and Tweets per **Search/Channel**, per day, per hour
- NumSearch: number of Tweets including the Search per **Channel,** per day, per hour
- Sentiment Analysis Score per **Search/Channel**, per day, per hour
- Num of xxxxx

Predicting presences at events

UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB

Predicting EXPO2015



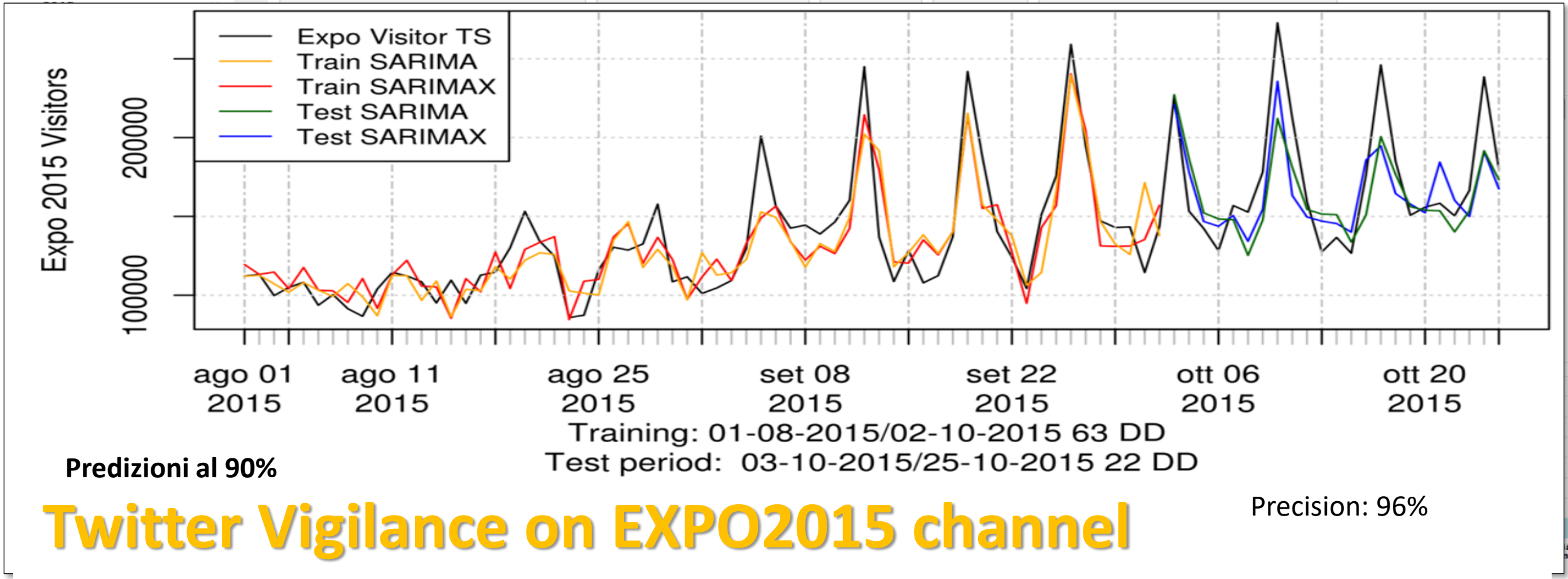


Twitter Vigilance on EXPO2015 channel

Twitter Vigilance

UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB

Twitter Vigilance
monitoring and predictions



Predizioni al 90%

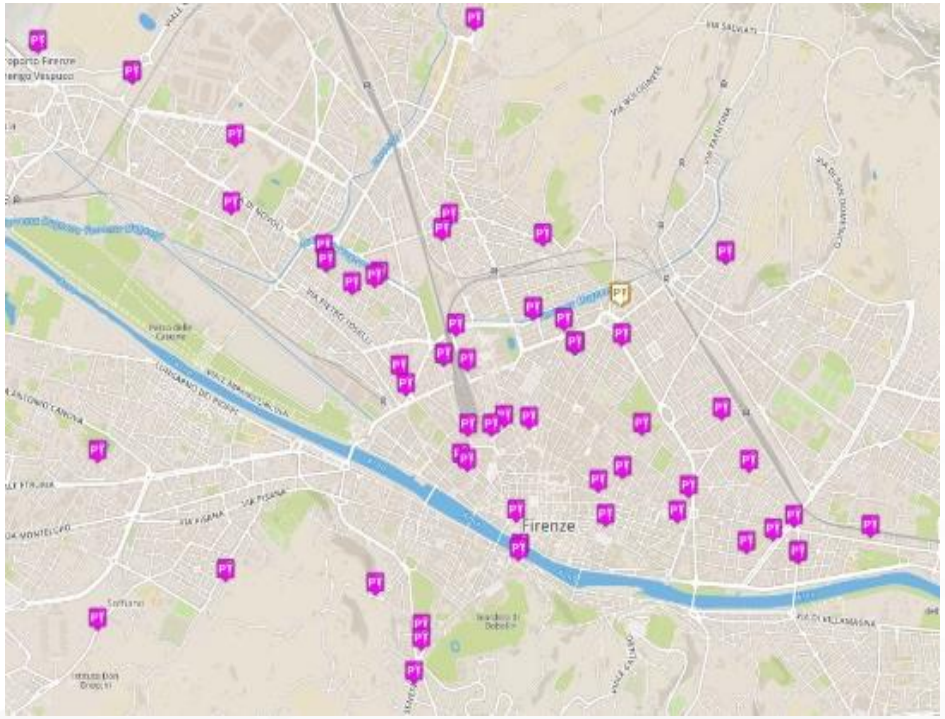
Precision: 96%

Twitter Vigilance on EXPO2015 channel

Predicting volume of visitors for tuning the services

Tweets Volume vs Lockdown

Different Reduction of parking for Covid



main Purpose	% of reduction in Lockdown	Recover in July 2020
Hospitals	55,9	31%
railway stations	84,2	60%
social hubs	85,8	72%
markets	45,1	25%
justice	96,1	82%

2020

Impact on Pollutant on LockDown 2020

	CO2	CO	NO2	PM10
Percentage at which the average value of the pollutant has been reduced (February wrt April, 2020)	41,06	47,54	38,84	61,99
Percentage at which the MAX value of the pollutant has been reduced (February wrt April 2020)	38,89	34,62	63,02	60,21

Event: clashes between police and demonstrators in Florence, during an unauthorized demonstration against the restrictions of Covid-19 government measures (on 30th October 2020 night).

Italy: Protests flare in Florence over restrictions

31 Oct, 2020

Facebook | Twitter | Mi piace 655 | Tweet

Florence mayor condemns violence and vandalism in historic centre in "surreal, terrible and painful night" for Tuscan city.

Protesters clashed with police in central Florence last night during an unauthorised protest over Italy's latest restrictions aimed at curbing the spread of covid-19.

In partnership with la Repubblica FIRENZE.it

Oct. 30 Protest Devastates Florence

10/31/2020

During the unauthorized evening demonstration against recent social lockdown pandemic measures lasting four hours in downtown Florence on the evening of October 30 saw

TOP NEWS STORIES

Oct. 30 Protest Devastates Florence
During the unauthorized evening demonstration against recent social lockdown pandemic measures lasting four hours in downtown Florence [...]

COVID Demonstrations Hit Florence
Unlike the U.S. protesters in Italy reserve public space for peaceful demonstrations which are organized to communicate a [...]

New Anti-Covid Regulations: Dos & Dont's
Given the upswing in Coronavirus infections and consequent hospital admissions since the beginning of October, Italian Prime Minister [...]

New Pandemic Regulations in Italy
In the face of the rise of Coronavirus infections, the Italian government has extended the national state [...]

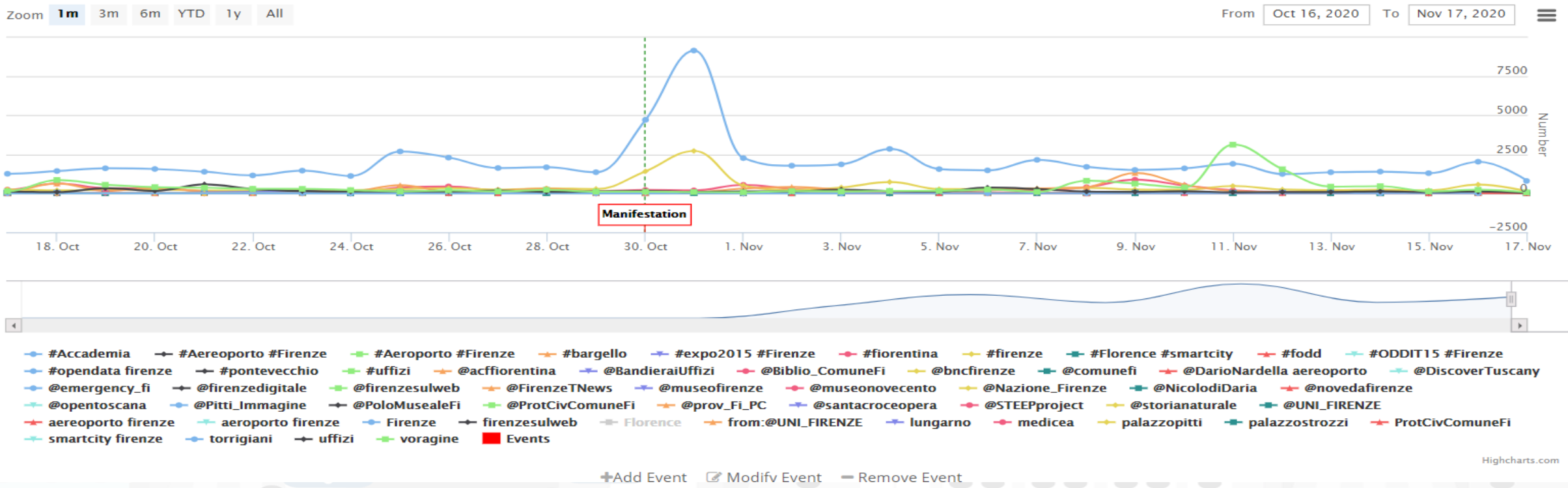
LA COMPAGNIA
Original Language Movies
International Film Festivals

ORCHESTRA DELLA TOSCANA

30 October Manifestation

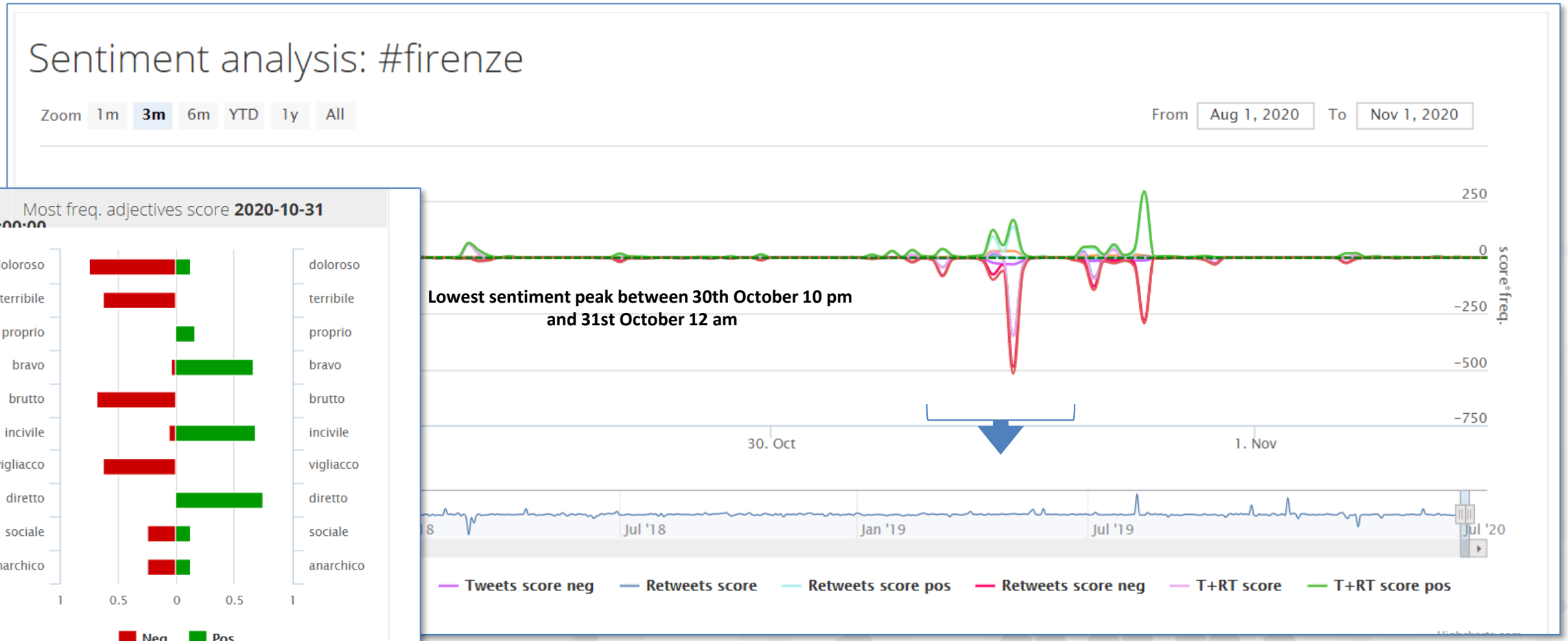
Channel active from 2009-02-27 to today | Data processed from 2015-05-22 to 2020-11-02 | NLP | SA

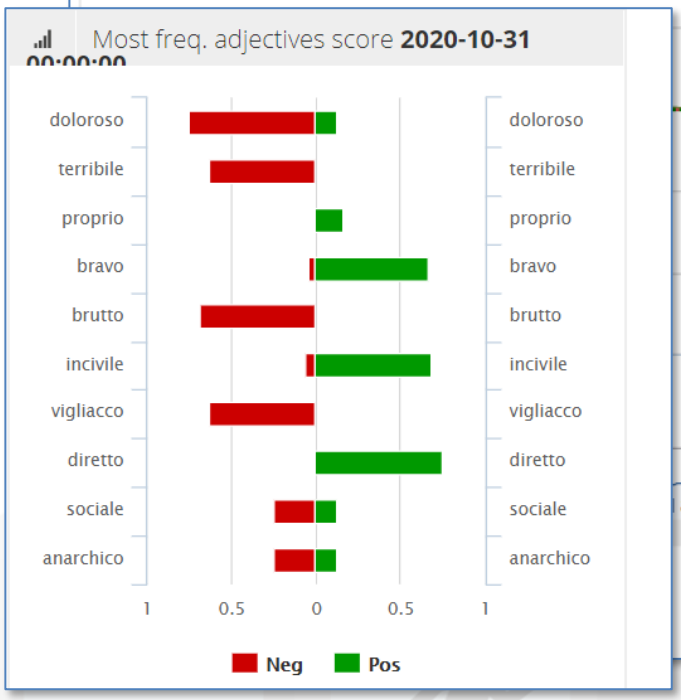
Search related to channel **Firenze**



Twitter Vigilance

Sentiment Analysis for search «Firenze»

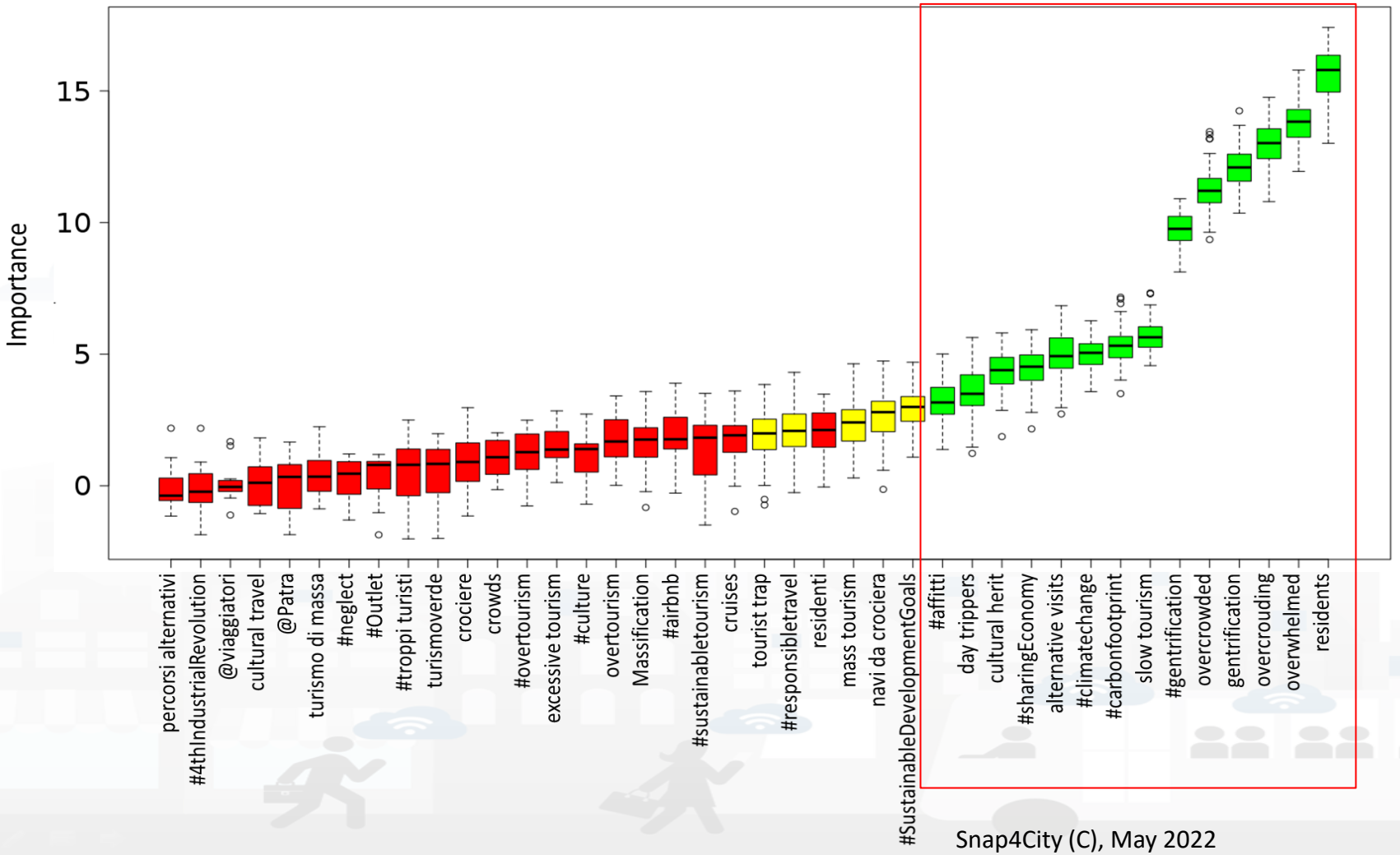




http://disit.org/tv/index.php?p=sa_ricerche&ricerca=436

Twitter Vigilance

Relevant Feature Selection - Italy



Random Forest Model Results expoiting the *confirmed* keys:

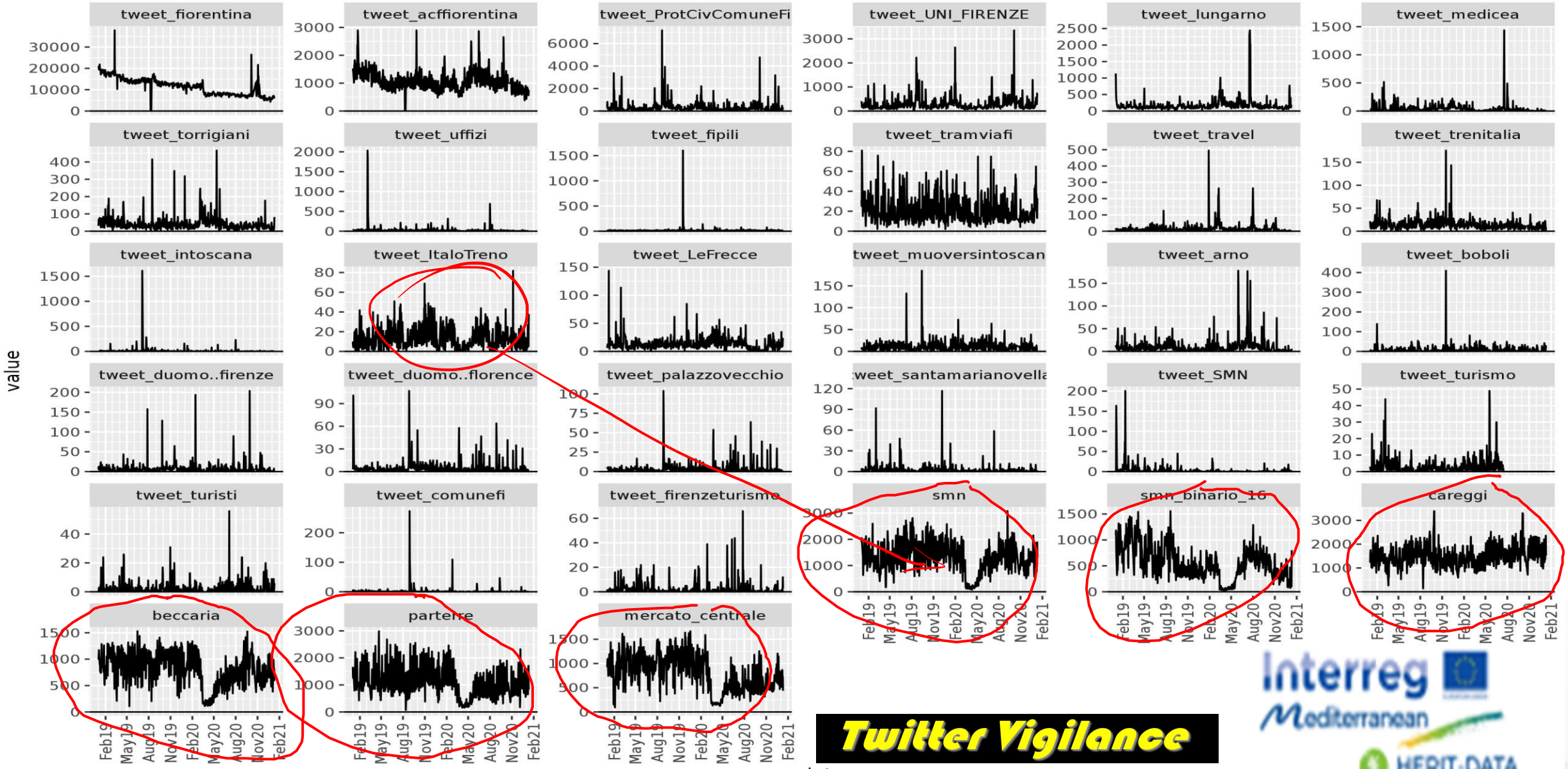
- Accuracy = 0.90
- Kappa index = 0.81

Twitter Vigilance

UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB

Twitter trends

SNAP4CITY | KM4CITY



Twitter Vigilance

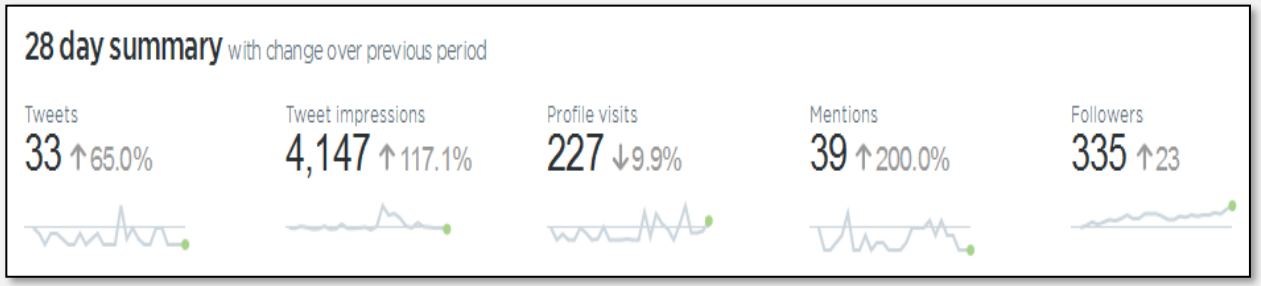
Interreg Mediterranean | HERIT-DATA

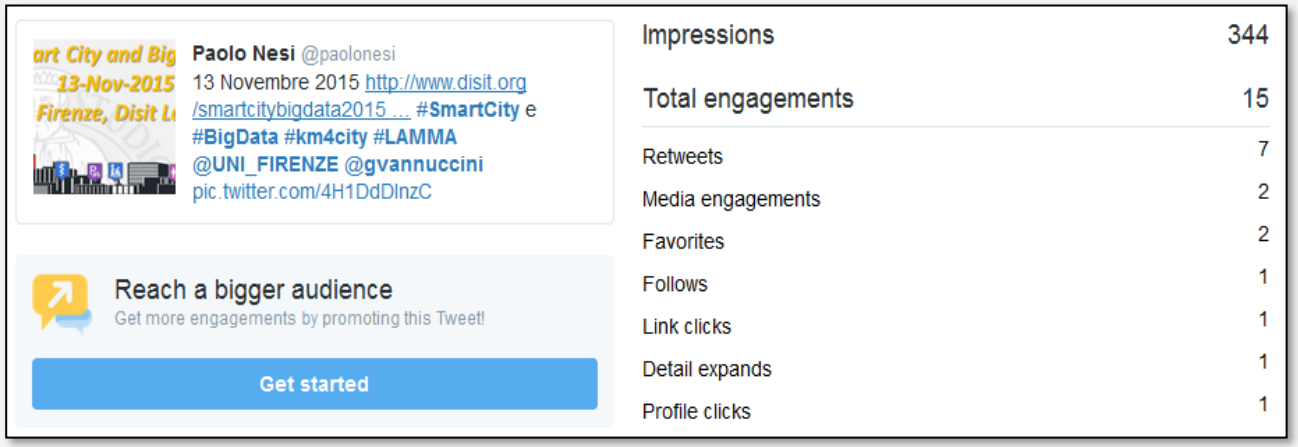
Predicting reTweet Proneness

Predicting the reTweet Proneness

- **Issue:**
  - How to understand if a tweet has a good probability of being retweeted?
- **Impact:**
  - Advertising, promotion, training
- **Several metrics related to**
  - Structure of the tweet
  - Features of the tweetting author
  - Relationships ….

Twitter Analytics





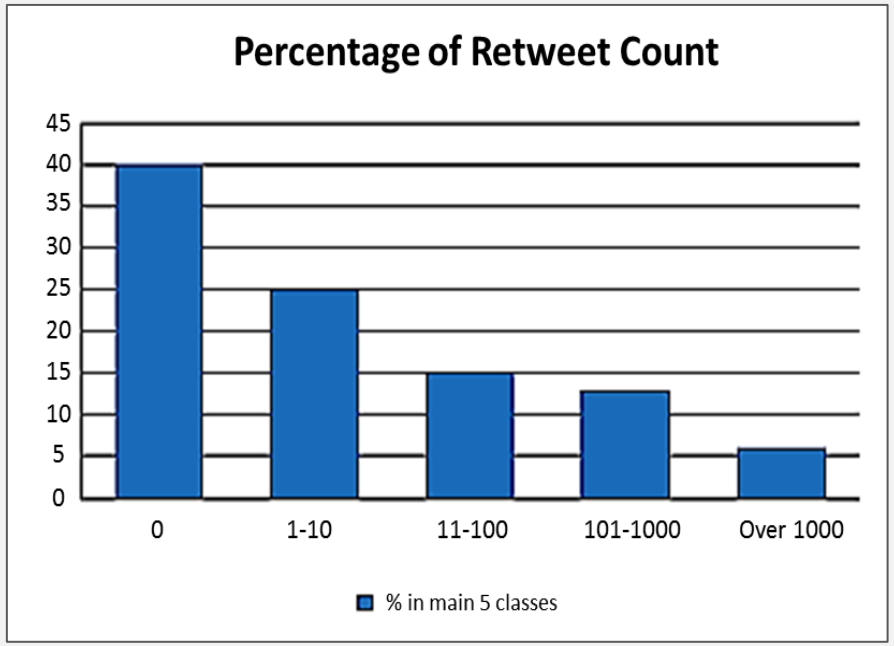
DISIT Lab, Distributed Data Intelligence and Technologies
Distributed Systems and Internet Technologies
Department of Information Engineering (DINFO)
http://www.disit.dinfo.unifi.it

Tweet proneness Metrics

Tweet metrics	
URLs Count	# of URLs in the tweet
Mentions Count	# of mentions/citation of Twitter users in the tweet
Hashtags Count	# of hashtags included in the tweet
Favorites Count	# of favorite obtained by the tweet
Publication Time	Local hour H24 in which the tweet has been published in the day according to the author' local time.
Author of Tweet metrics	
Days Count	# of days since the tweet's author created its Twitter account
Statuses Count	# of tweets made by the tweet's author since the creation of its own account
Author Network metrics	
Followers Count	# of followers the author of the tweet
Followees Count	# of friends the tweet's author is following
Listed Count	# of people added the tweet's author to a list

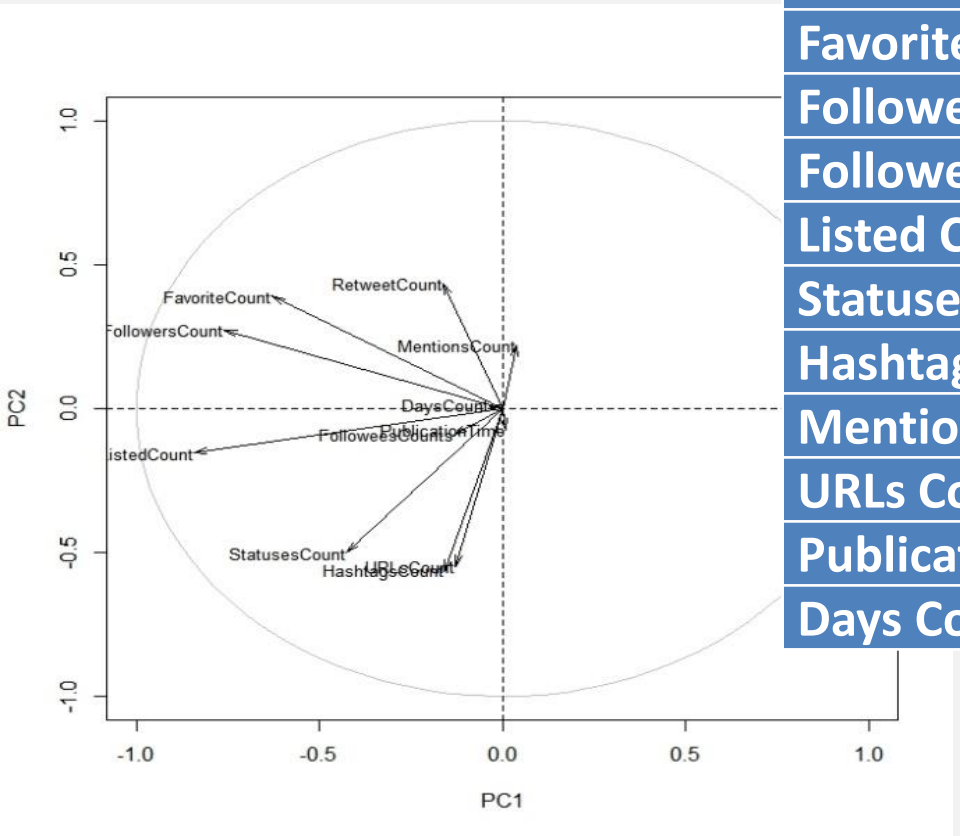
Data sets:

- 100 Million of Tweet
- 500 K
- 100 K



reTweet proneness: assessment

- PCA



Metrics	PC1	PC2	PC3	PC4	PC5
Retweet Count	-0.1623	0.4346	0.1635	-0.0026	-0.1009
Favorites Count	**-0.6294**	0.3908	0.1922	-0.1128	-0.1880
Followers Count	**-0.7599**	0.2736	0.0522	-0.0983	-0.0857
Followees Count	-0.1336	-0.0907	**-0.4627**	-0.2494	0.1182
Listed Count	**-0.8431**	-0.1549	-0.0498	0.1500	0.1871
Statuses Count	-0.4256	**-0.5016**	-0.3781	0.2795	0.2410
Hashtags Count	-0.1585	**-0.5661**	0.4377	-0.0517	0.0309
Mentions Count	0.0394	0.2194	0.0786	-0.1607	**0.7697**
URLs Count	-0.1288	**-0.5483**	0.2539	-0.3388	-0.3248
Publication Time	0.0076	-0.0728	0.3639	**-0.5186**	0.3707
Days Count	-0.0370	0.0070	-0.**5072**	**-0.6604**	-0.1691

reTweet proneness: Classification methods

- Statistic classifications vs machine-learning methods
- 80% of training data set, 20% of testing data sets; 500K data set
- → Recursive partitioning procedure models (RPART), good compromise for Big data problems

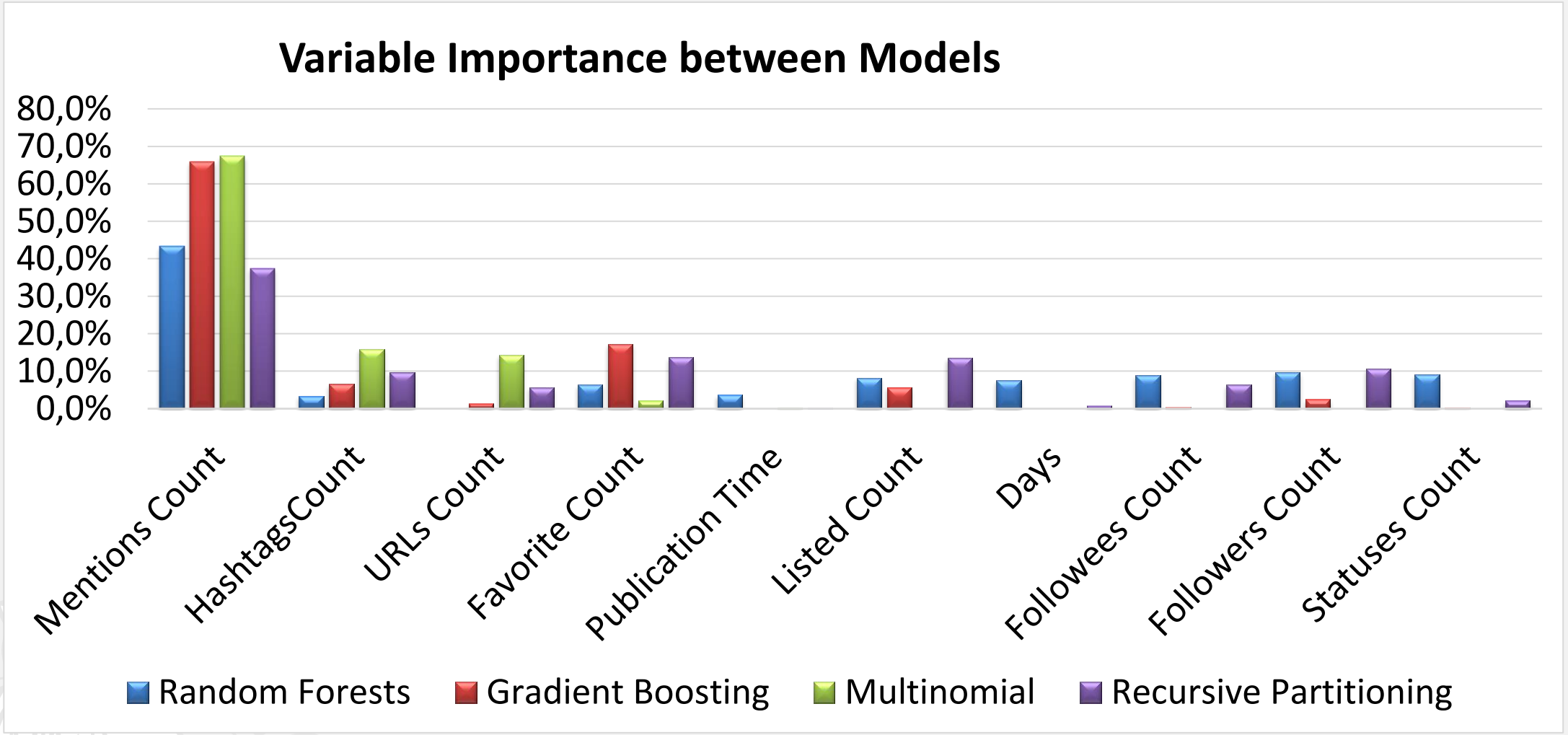
Classifier Models	Accuracy	Precision	Recall	F_1 score	Processing Time in sec.
Recursive Partitioning (Stat)	0.6807	0.8512	0.7767	0.8122	180
Random Forests (ML)	0.6884	0.8601	0.7866	0.8217	198968
Gradient boosting (ML)	0.6796	0.8534	0.7731	0.8113	64448
Multinomial Model (Stat)	0.6411	0.8367	0.7245	0.7765	31576

reTweet proneness (RPART), 100M

Assessment drivers	Degree of Retweeting Classes				
	0	1-100	101-1000	1001-10000	Over 10000
Sensitivity	0.7737	0.8105	0.3142	0.0208	0.0136
Specificity	0.9132	0.6694	0.9199	0.9996	1.0000
Positive Predictive Value	0.8564	0.6256	0.3752	0.7345	0.8488
Negative Predictive Value	0.8579	0.8382	0.8975	0.9485	0.9915
Prevalence	0.4007	0.4053	0.1328	0.0526	0.0086
Detection Rate	0.3100	0.3285	0.0417	0.0011	0.0001
Detection Prevalence	0.3620	0.5251	0.1112	0.0015	0.0001
Balanced Accuracy	**0.8435**	**0.7399**	**0.6170**	0.5102	0.5068

Accuracy	0.6815
Accuracy 95% Confidential Interval (min, max)	(0.6813, 0.6817)
Recall	0.7737
Precision	0.8564
Kappa	0.4922

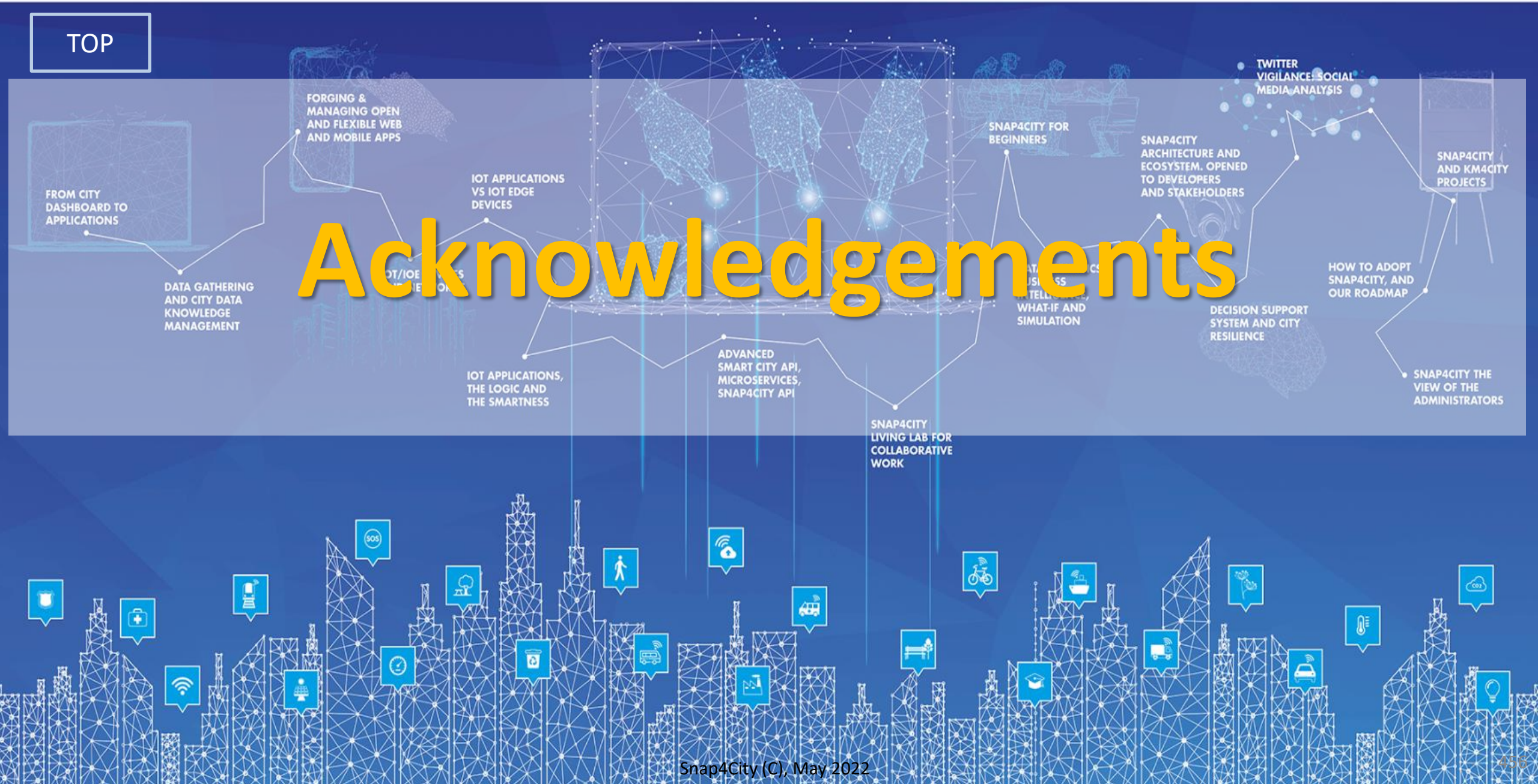
Predictive models VS metrics relevance



Citations and self training

- P. Nesi, G. Pantaleo, I. Paoli, I. Zaza, "Assessing the reTweet Proneness of tweets: predictive models for retweeting", Multimedia Tools and Applications, Springer, 2018. https://link.springer.com/article/10.1007/s11042-018-5865-0
- A. Crisci, V. Grasso, P. Nesi, G. Pantaleo, I. Paoli, I. Zaza, "Predicting TV programme Audience by Using Twitter Based Metrics", Multimedia Tools and Applications, springer. 10.1007/s11042-017-4880-x, 2017 https://link.springer.com/article/10.1007/s11042-017-4880-x
- V. Grasso, A. Crisci, M. Morabito, P. Nesi, G. Pantaleo, "Public crowdsensing of heat waves by social media data", Adv. Sci. Res., 14, 217-226, https://doi.org/10.5194/asr-14-217-2017, 2017, 10.5194/asr-14-217-2017 . http://www.adv-sci-res.net/14/217/2017/
- V. Grasso, A. Crisci, M. Morabito, P. Nesi, G. Pantaleo, I. Zaza, B. Gozzini, "Italian codified hashtags for weather warning on Twitter–who is really using them?." Advances in Science and Research 14 (2017): 63-69. http://www.adv-sci-res.net/14/63/2017/asr-14-63-2017.pdf

TOP



Acknowledgements

DISIT lab roadmap vs model and tools' usage

2013
Km4City Ontology 1.1

- Tuscany, Road Graph
- Mobility
- culture, tourism
- Events
- Parking
- Services
- Linked open graph

2014
- Weather Forecast
- Real Time Wi-Fi
- Entertainment
- LOD

- Twitter Vigilance
- Social Media Analytics, Sentiment Analysis

Km4City 1.4

2015
Km4City 1.5

- Resilience Decision Support
- Smart First Aid
- User Behaviour Analysis, predictions
- Risk Analysis

ckan, resolute

H2020 (2015-18)

2016
(2016-21) H2020 REPLICATE

Km4City 1.6.4

- Smart Energy
- Sustainable Mobility
- Control Room
- Dashboard

Km4City 1.6.2

Sii-Mobility (2016-21)

SII-MOBILITY SCN

- Infomobility
- Mobile App
- Routing
- Multimodality

FIWARE

Km4City 1.5

- Resilience Decision Support
- Smart First Aid
- User Behaviour Analysis, predictions
- Risk Analysis

2017
Sii-Mobility

- Origin-Destination and trajectories
- Traffic Reconstruction
- Offer Analysis
- OBU, smart devices

E0I5 digital ecosystem (Powered by EXPO MILANO 2015)

(2016-19) GHOST SIR

- Sardinia Region Smart City Strategies and plan

GREEN IMPACT C.A.P.A.C.I.T.Y.

GREEN IMPACT
POR FESR 2014-2020

- Industry 4.0
- Critical Plant
- Monitoring

- IOT/IOE, IOT App
- Living Lab
- Maker Support
- IOT Edge
- Smart City IOT
- GDPR,
- Privacy & Security

life weee (2017-20)

- Smart Waste

2018
MOSAiC MObility 4.0 for SmArt (i) City (2018-20)

- Mobility Demand / Offer Analytics and Strategy

Sii-Mobility

- User engagement
- Bike Sharing
- Data Analytics ++
- Social Predictions
- OBD2

IOT/IOE

Km4City 1.6.6

Node-RED

SNAP4CITY SELECT for Cities H2020 (2017-19)

2019
5G (2018-21)

- 5G tech
- Energy
- Industry 4.0
- Synoptics

TRAFAIR CEF (2018-21)

- Traffic and Mobility Impact on Pollution
- NOX predictions

EUROPEAN OPEN SCIENCE CLOUD

LAID Smart Bed

- Smart Health

GREEN FIELD PEAS Soda4.0 Optimization and Automated Decision System 4.0

- Industry 4.0

bee smart city

- Smart Lonato

Reverberi Enetec Gruppo MPES

SNAP4CITY KM4CITY

Winner of Select4Cities PCP

SNAP4CITY roadmap

SODA

2020

JRC EUROPEAN COMMISSION

Interreg Mediterranean HERIT-DATA

- Smart Tourism
- 6 Pilots
- Data Analytics
- Extended platform

Interreg MARITTIMO-IT FR-MARITIME MOBI MART

- Smart Mobility
- PISA, PUMS
- Living lab

2021

JRC EUROPEAN COMMISSION

PC4City (2020-21)
Monitoring Terrain

Winner of Open Data Challenge of enel x

CAPELON

- Smart Light
- Sweden

Km4City 1.6.7

enel x
Contract

Smart Ambulance (2021-22)

Enterprise (2021-22)
Industry 4.0

Almafluida
Industry 4.0
(2021-22)

AMPERE (2021-22)
Industry 4.0

SYN-RG-AI
SmartCity

Smartea

IMPETUS
Security and Risk

ITALFERR
GRUPPO FERROVIE DELLO STATO ITALIANE
Smart City

ASYMMETRICA URBAN TECHNOLOGY
Smart City

enel x
Contract

2022

AXIS COMMUNICATIONS
SmartCity

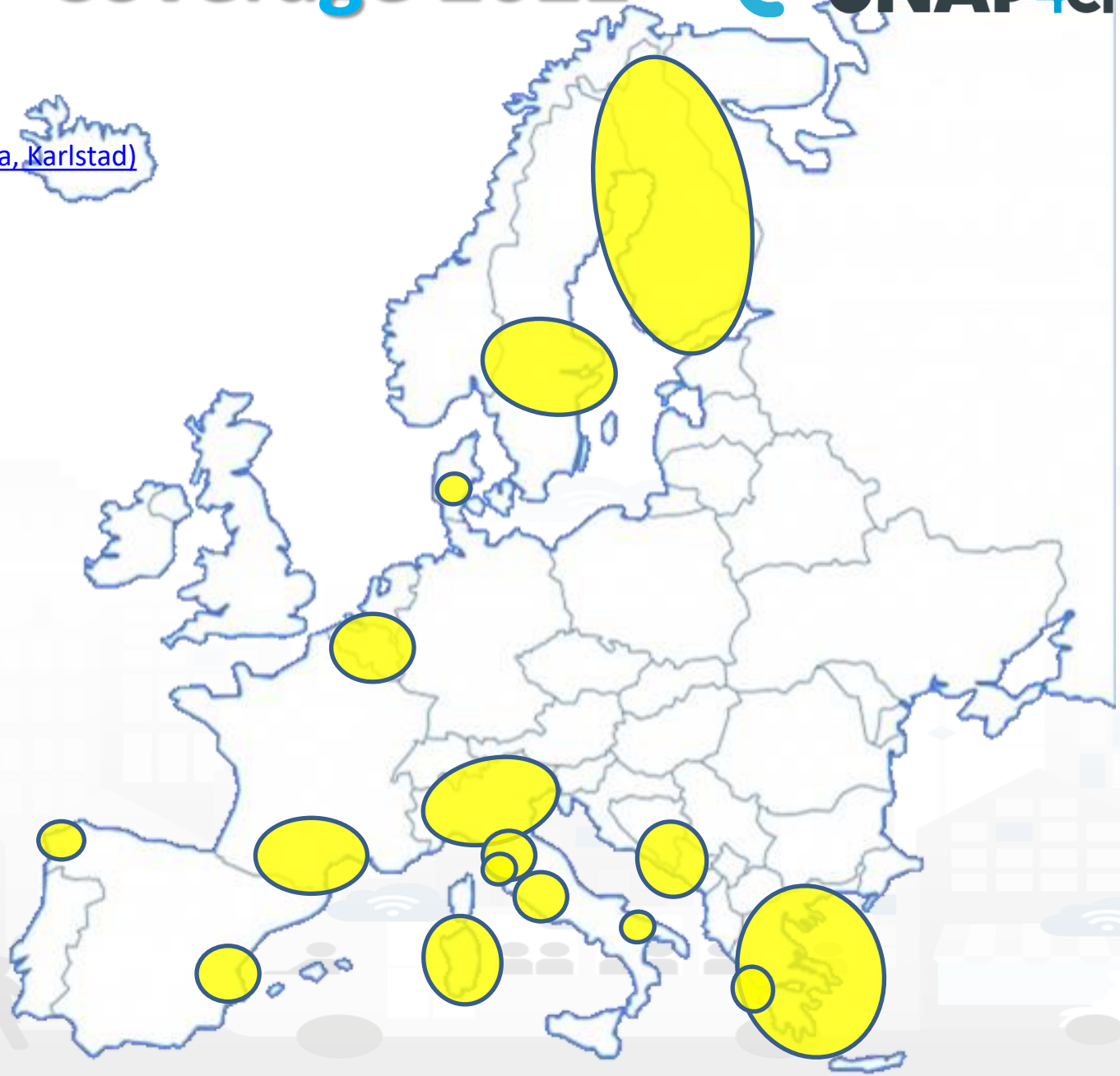
uni.systems
SmartCity

GRUPPO PRETTO
Industry 4.0

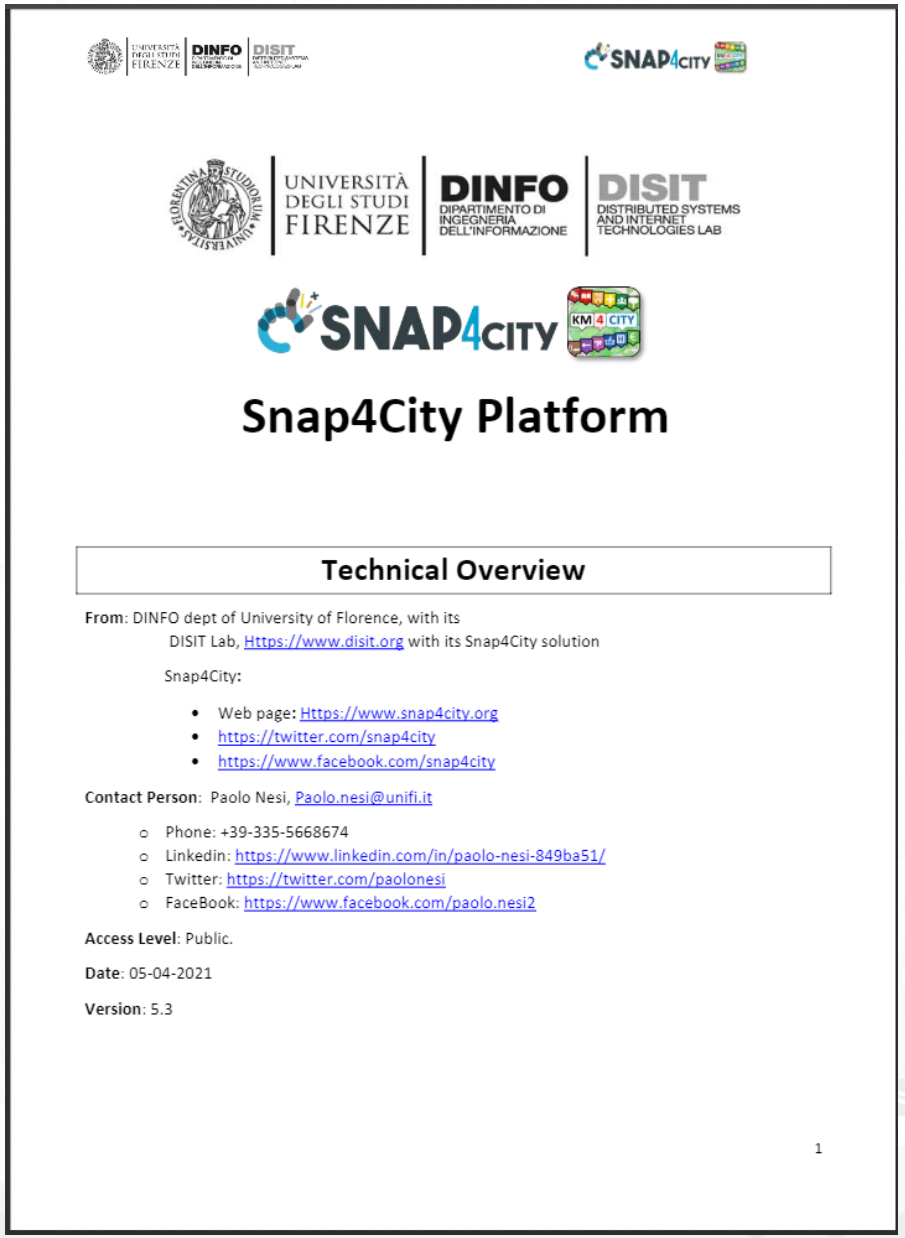
Coverage 2022

Main Organizations/areas

- Antwerp area (Be)
- Bologna (I)
- Capelon (Sweden: Västerås, Eskilstuna, Karlstad)
- DISIT demo (multiple)
- Dubrovnik, Croatia
- Firenze area (I)
- Garda Lake area (I)
- Greece (Gr)
- Helsinki area (Fin)
- Livorno area (I)
- Lonato del Garda (I)
- Modena (I)
- Mostar, Bosnia-Herzegovina
- Oslo & Padova (Impetus)
- Pisa area (I)
- Pistoia (I)
- Pont du Gard, Occitanie (Fr)
- Prato (I)
- Roma (I)
- Santiago de Compostela (S)
- Sardegna Region (I)
- Siena (I)
- SmartBed (multiple)
- Toscana Region (I), SM
- Valencia (S)
- Venezia area (I)
- WestGreece area (Gr)



Overview



Snap4City Platform

Technical Overview

From: DINFO dept of University of Florence, with its
DISIT Lab, Https://www.disit.org with its Snap4City solution

Snap4City:

- Web page: Https://www.snap4city.org
- https://twitter.com/snap4city
- https://www.facebook.com/snap4city

Contact Person: Paolo Nesi, Paolo.nesi@unifi.it

- Phone: +39-335-5668674
- Linkedin: https://www.linkedin.com/in/paolo-nesi-849ba51/
- Twitter: https://twitter.com/paolonesi
- FaceBook: https://www.facebook.com/paolo.nesi2

Access Level: Public.

Date: 05-04-2021

Version: 5.3

1

- https://www.snap4city.org/drupal/sites/default/files/files/Snap4City-PlatformOverview.pdf

UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB

Main running instances (11/21)

- Sii-Mobility → mobility and transport, sustainability
- REPLICATE → ICT, smart City Control room, Energy, IOT
- RESOLUTE → Resilience, ICT, Big Data
- GHOST → Strategies, smart city
- TRAFAIR → Environment & transport
- MOSAIC → mobility and transport
- WEEE Life → Smart waste, environment
- Smart Garda Lake → Castelnuovo del Garda, SMARTEA
- 5G → Industry 4.0 vs SmartCity
- Green Impact → Industry 4.0, Chemical Plant, control and plan
- SmartBed (Laid) → smart health
- Green Field Peas (Soda) → Industry 4.0, Chemical plant
- MobiMart and PISA Agreement → data aggregation, mobility and transport, Living Lab
- Lonato del Garda → smart parking, environment
- Herit Data → tourism, culture and management
- ISPRA JRC → site management and services
- Capelon (Sweden) → smart light solutions
- PC4City → land slide monitoring and predictions
- Italmatic → industry 4.0 production control

Acknowledgements

- Thanks to the European Commission for founding. All slides reporting logo of **Snap4City** https://www.snap4city.org **of Select4Cities H2020** are representing tools and research founded by European Commission for the **Select4Cities** project. **Select4Cities** has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation Programme (grant agreement n° 688196)
- **TRAFAIR is a CEF** project. All slides reporting logo of TRAFAIR project are representing tools and research founded by the EC on CEF programme http://trafair.eu/
- Thanks to the European Commission for founding. All slides reporting logo of **REPLICATE H2020** are representing tools and research founded by European Commission for the REPLICATE project. **REPLICATE** has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation Programme (grant agreement n° 691735).
- Thanks to the European Commission for founding. All slides reporting logo of **RESOLUTE H2020** are representing tools and research founded by European Commission for the RESOLUTE project. **RESOLUTE** has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation Programme (grant agreement n° 653460).
- Thanks to the MIUR for co-founding and to the University of Florence and companies involved. All slides reporting logo of **Sii-Mobility** are representing tools and research founded by MIUR for the Sii-Mobility SCN MIUR project.
- **Km4City** *is an open technology and research line of DISIT Lab exploited by a number of projects. Some of the innovative solutions and research issues developed into projects are also compliant and contributing to the Km4City approach and thus are released as open sources and are interoperable, scalable, modular, standard compliant, etc.*

https://www.snap4city.org/577

On Line Training Material (free of charge)

	1st part (*)	2nd part (*)	3rd part (*)	4th part (*)	5th part (*)	6th part (*)	7th part (*)
what	General	Dashboards	IOT App, IOT Network	Data Analytics	Data Ingestion processes	System and Deploy Install	Smart City API: Web & Mob. App
PDF							
Inter active							
Video1							
Video2							
Video3							
Video4				none		none	none
duration	2:55	3:16	3:41	2:00	2:48	2:35	1:47

TOP

SNAP4CITY

KM 4 CITY

Be smart in a SNAP!

SNAP4
Appliances and Dockers
Installations

CONTACT

DISIT Lab, DINFO: Department of Information Engineering
Università degli Studi di Firenze - School of Engineering

Via S. Marta, 3 - 50139 Firenze, ITALY
https://www.disit.org

www.snap4city.org

Email: snap4city@disit.org

Office: +39-055-2758-515 / 517
Cell: +39-335-566-86-74
Fax.: +39-055-2758570

UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB