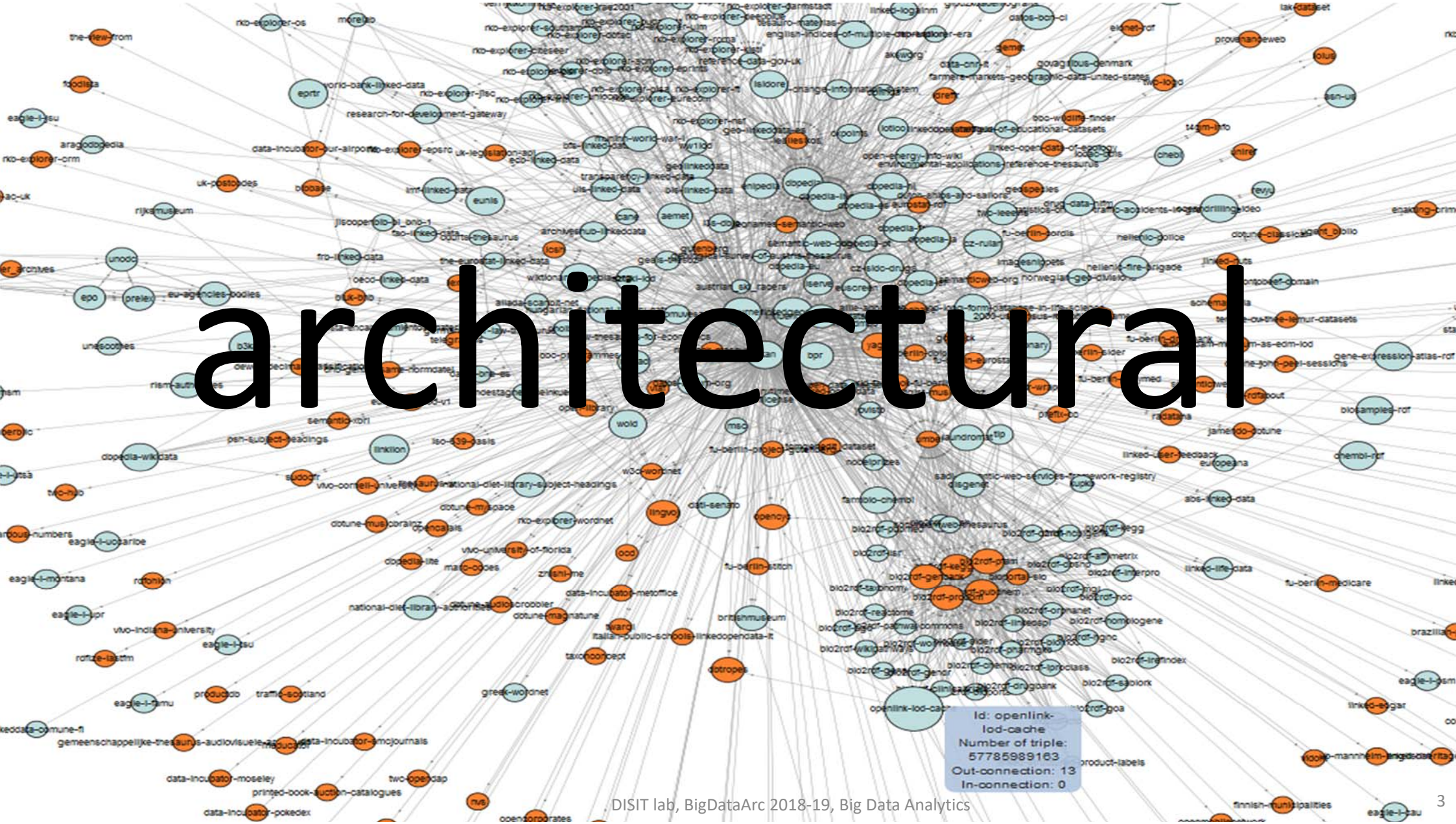
UNIVERSITÀ DEGLI STUDI FIRENZE
DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE
DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB

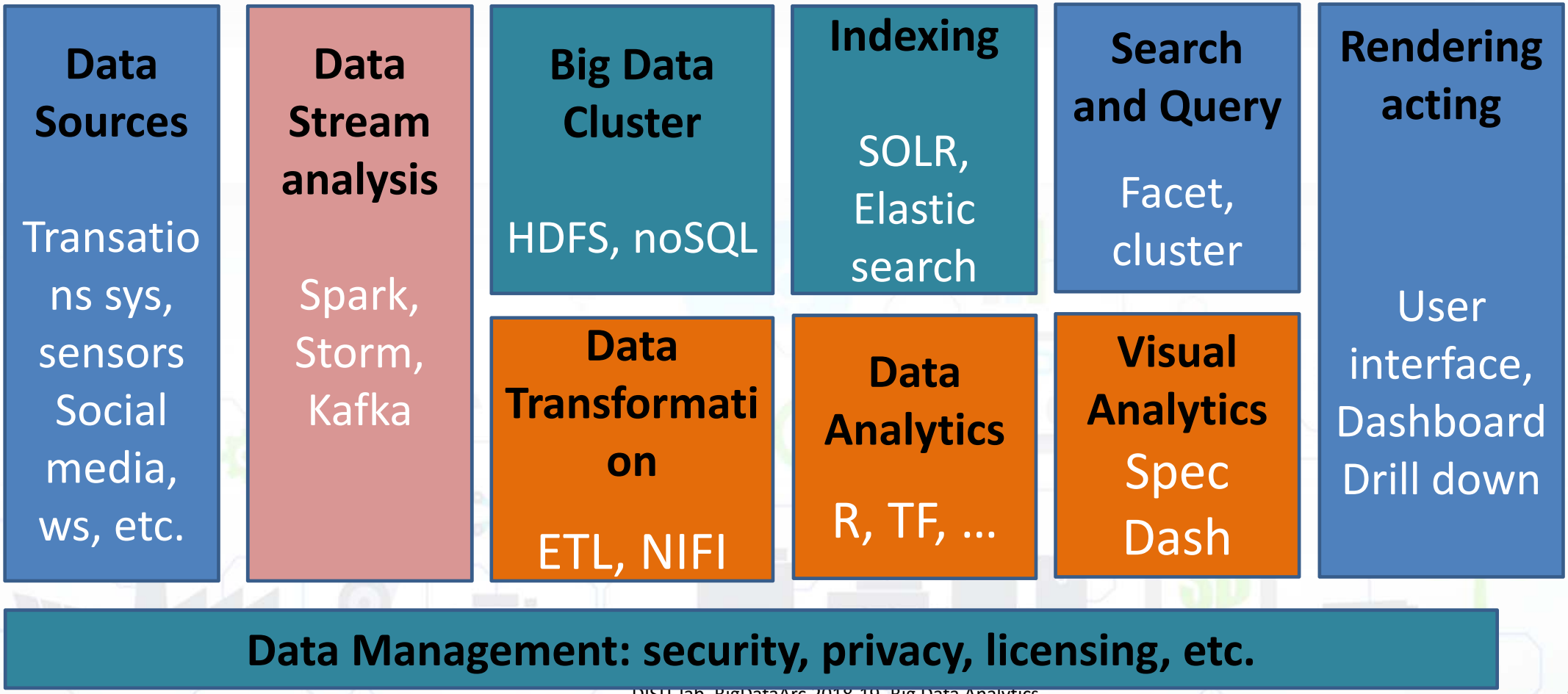
https://www.disit.org/
Paolo Nesi, paolo.nesi@unifi.it

Corso Big Data Architectures
Parte: Big Data Analytics, examples

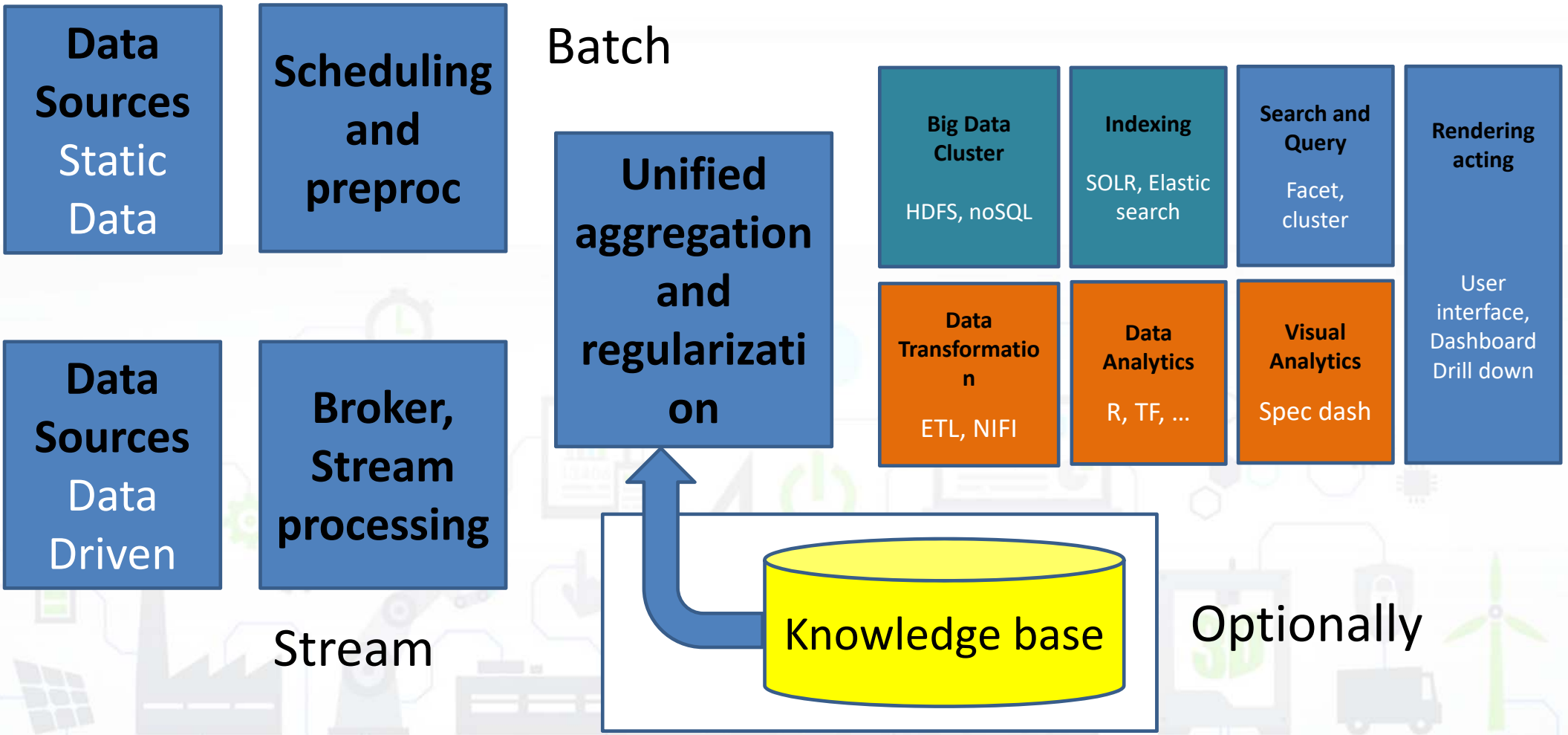
https://www.snap4City.org https://www.Km4City.org


architectural
Id: openlink-lod-cache
Number of triple: 57785989163
Out-connection: 13
In-connection: 0

Architettura di base Big Data, IOT, Industry 4.0



Lambda Architecture



Transport systems
Mobility, parking

Public Services
Govern, events, …

Sensors, IOT
Cameras, ..

Environment,
Water, energy

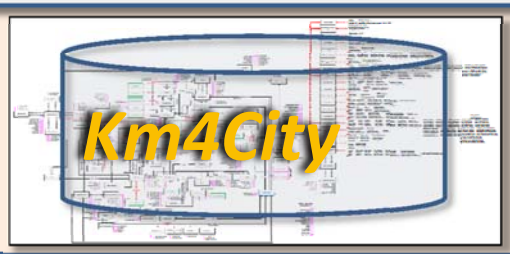
Shops, services,
operators

Social Media
WiFi, network

Static, Slow and Real Time data flows

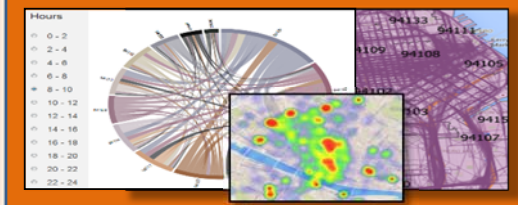
DISCES -- Distributed and parallel architecture on Cloud

Km4City Smart City Engine



User Profiling and Suggestions on Demand

Flow and Origin Destination Matrix
Http://www.disit.org/odsf



Km4City Tools for Developers

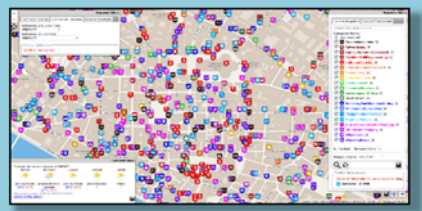
Km4City Smart City API

Tools for City Operators and Decision Makers

Smart City Dashboard
Http://www.disit.org/dash

Smart Decision Support
Http://Smartds.disit.org

Service map browser
Http://servicemap.disit.org



Twitter Vigilance
Http://www.disit.org/tv

Collective User behavior Analyzer

Tools for Final Users

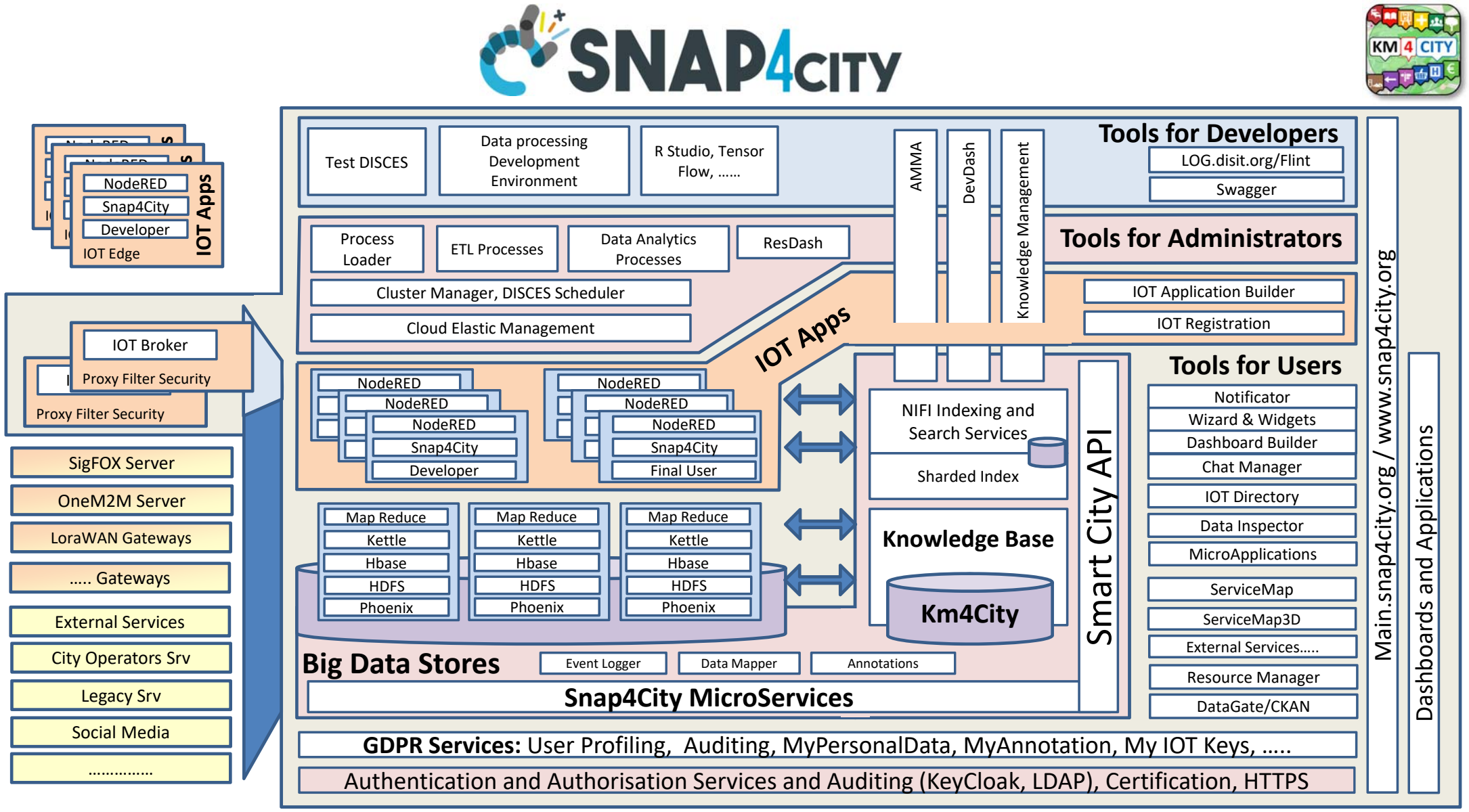
Mobile e Web Apps

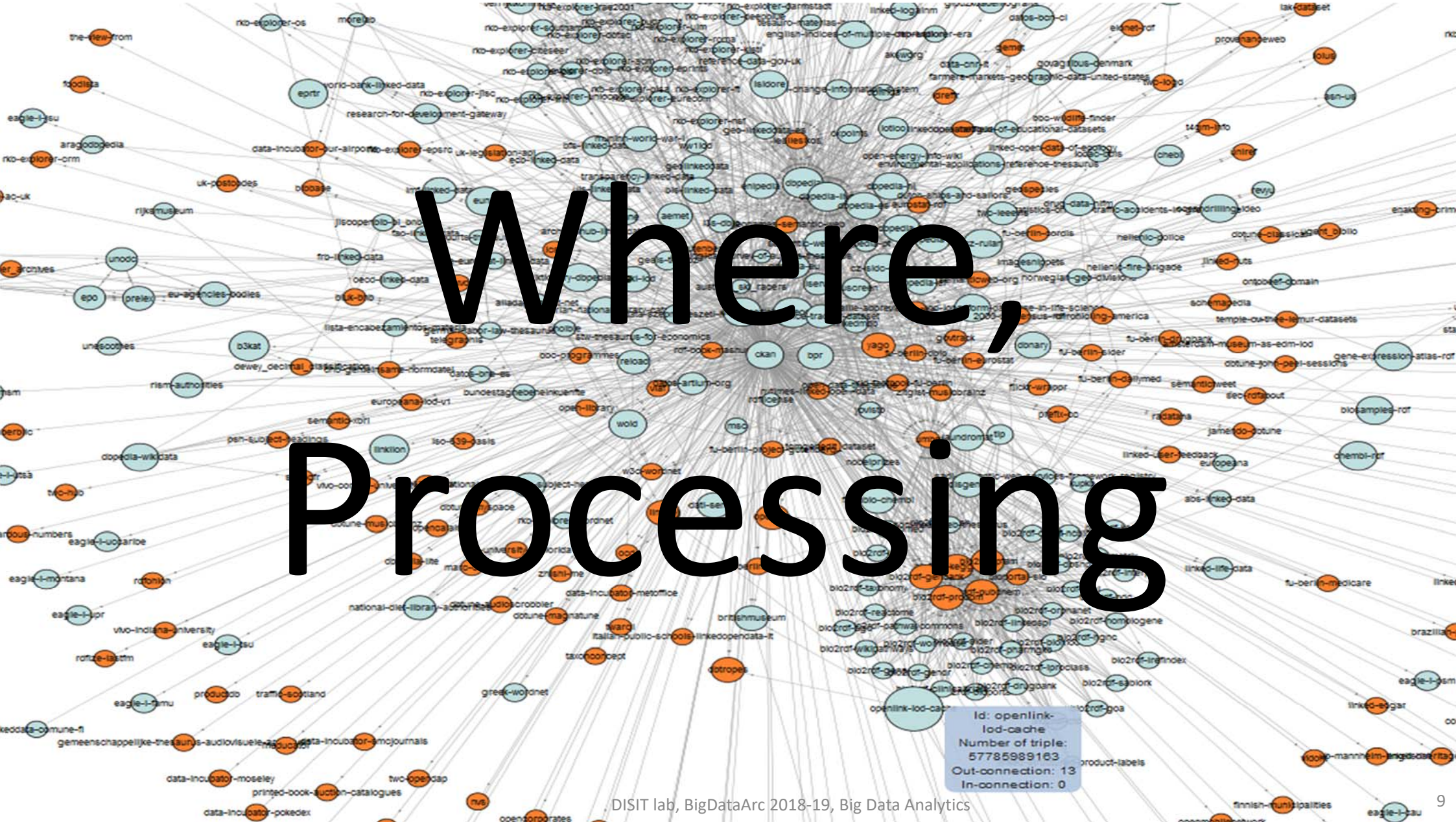
Http://www.km4city.org

A livello di Sistema

non abbiamo il tempo di pulire i dati o regolarizzarli, il sistema deve lavorare con quello che arriva, ed inoltre deve essere:

- in grado di operare H24/7, in HA?
- in grado di reggere il carico delle richieste? è scalabile?
- in grado di lavorare alla massima precisione in predizione?
- in grado di rispondere in tempo reale?
- resiliente: recupera stabilità a fronte di eventi inattesi?
- modulare, è flessibile, è replicabile, è open, è …..
- sicuro?
- In grado di rispettare la Privacy?

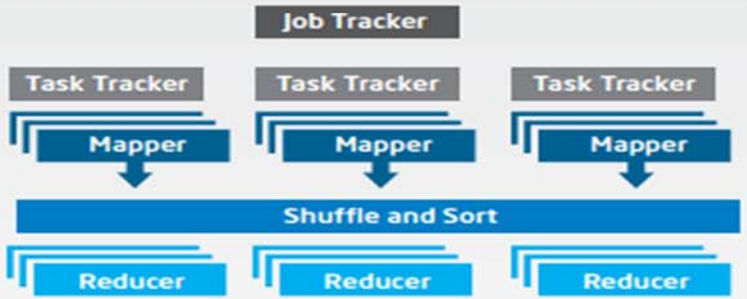

SNAP4CITY
KM 4 CITY
NodeRED
Snap4City
Developer
IOT Edge
IOT Apps
IOT Broker
Proxy Filter Security
Proxy Filter Security
SigFOX Server
OneM2M Server
LoraWAN Gateways
..... Gateways
External Services
City Operators Srv
Legacy Srv
Social Media
.....
Test DISCES
Data processing Development Environment
R Studio, Tensor Flow, ……
AMMA
DevDash
Knowledge Management
Tools for Developers
LOG.disit.org/Flint
Swagger
Process Loader
ETL Processes
Data Analytics Processes
ResDash
Cluster Manager, DISCES Scheduler
Cloud Elastic Management
Tools for Administrators
IOT Application Builder
IOT Registration
IOT Apps
NodeRED
NodeRED
NodeRED
Snap4City
Developer
NodeRED
NodeRED
NodeRED
Snap4City
Final User
NIFI Indexing and Search Services
Sharded Index
Knowledge Base
Km4City
Smart City API
Tools for Users
Notificator
Wizard & Widgets
Dashboard Builder
Chat Manager
IOT Directory
Data Inspector
MicroApplications
ServiceMap
ServiceMap3D
External Services.....
Resource Manager
DataGate/CKAN
Map Reduce
Kettle
Hbase
HDFS
Phoenix
Map Reduce
Kettle
Hbase
HDFS
Phoenix
Map Reduce
Kettle
Hbase
HDFS
Phoenix
Big Data Stores
Event Logger
Data Mapper
Annotations
Snap4City MicroServices
GDPR Services: User Profiling, Auditing, MyPersonalData, MyAnnotation, My IOT Keys, ...
Authentication and Authorisation Services and Auditing (KeyCloak, LDAP), Certification, HTTPS
Main.snap4city.org / www.snap4city.org
Dashboards and Applications


Where,
Processing
Id: openlink-lod-cache
Number of triple: 5785989163
Out-connection: 13
In-connection: 0

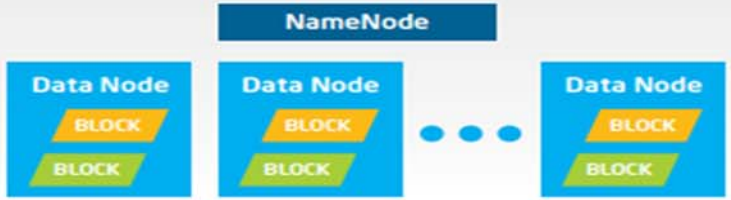
Hadoop and MapReduce

LOGICAL ARCHITECTURE

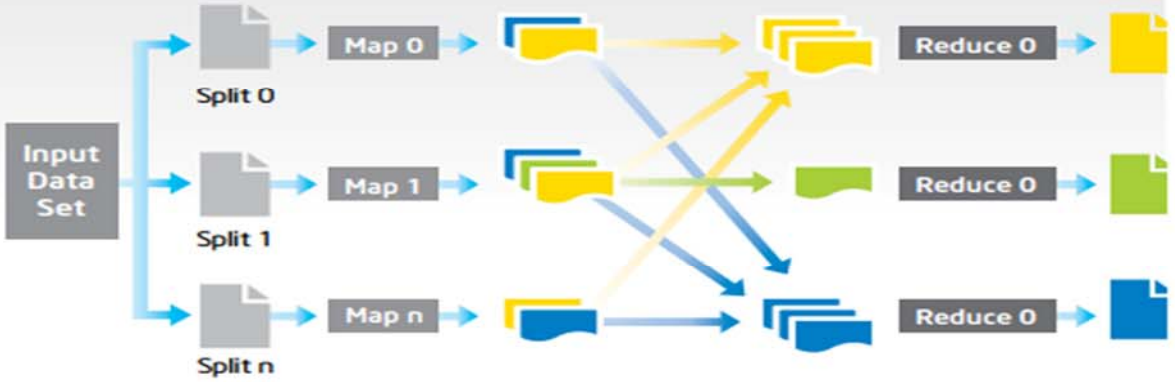
Processing: MapReduce



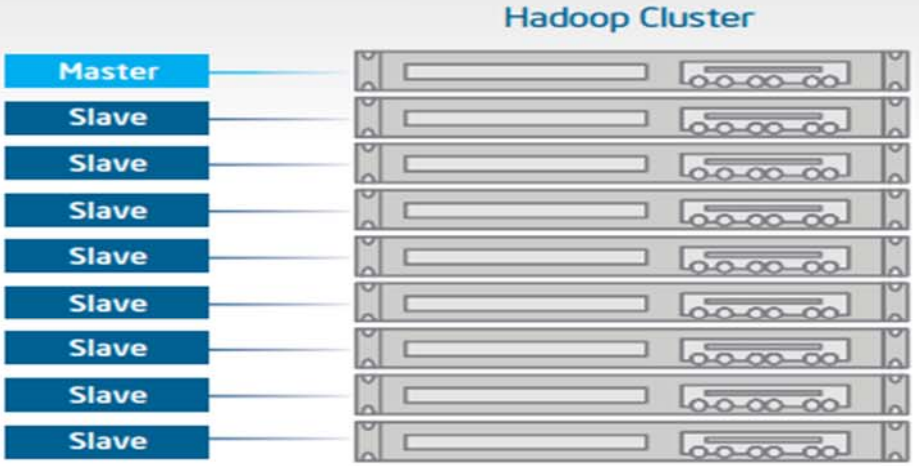
Storage: HDFS



PROCESS FLOW

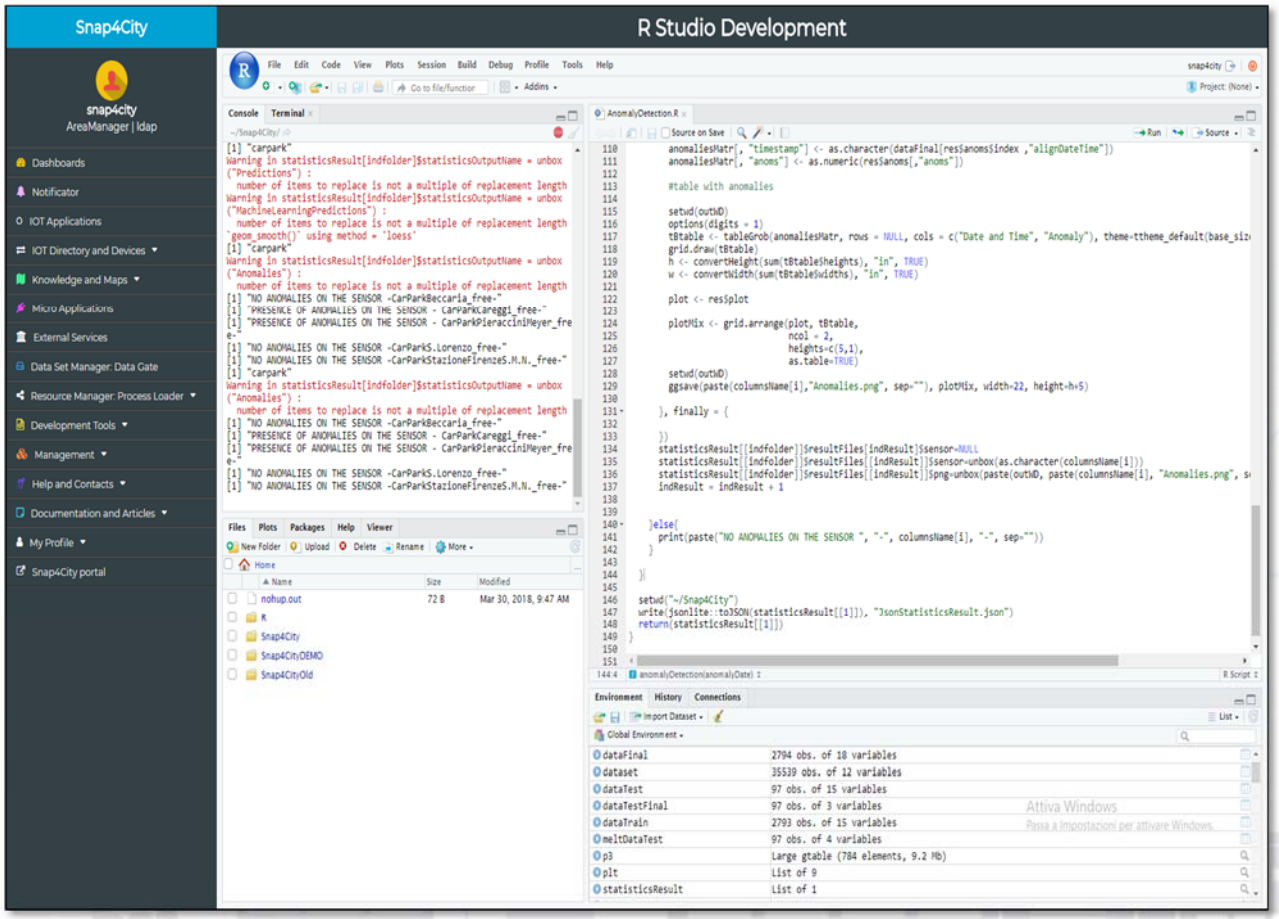


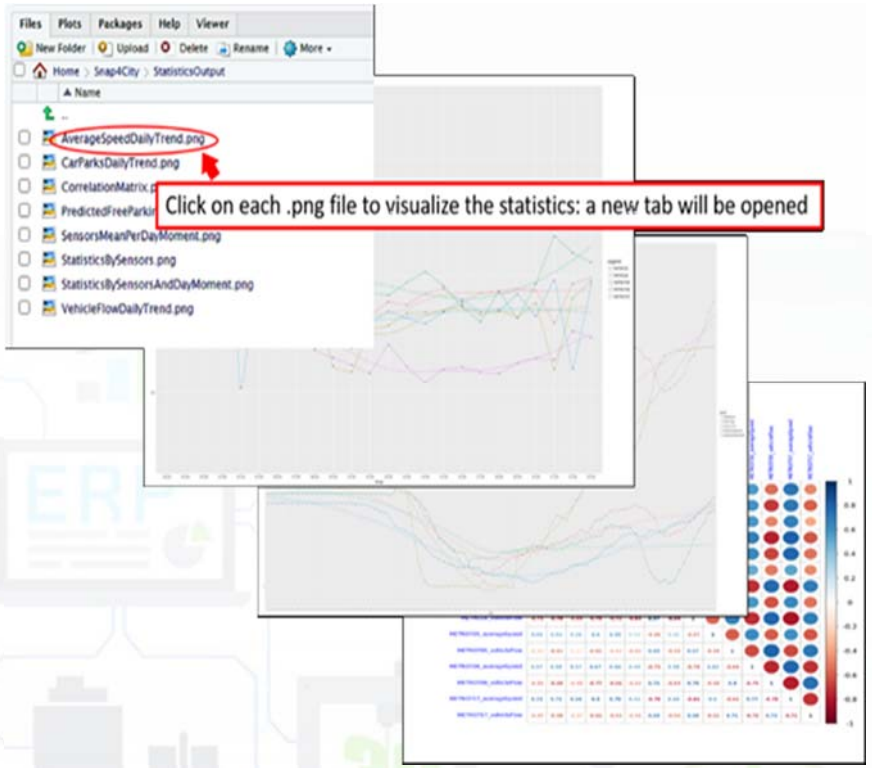
PHYSICAL ARCHITECTURE



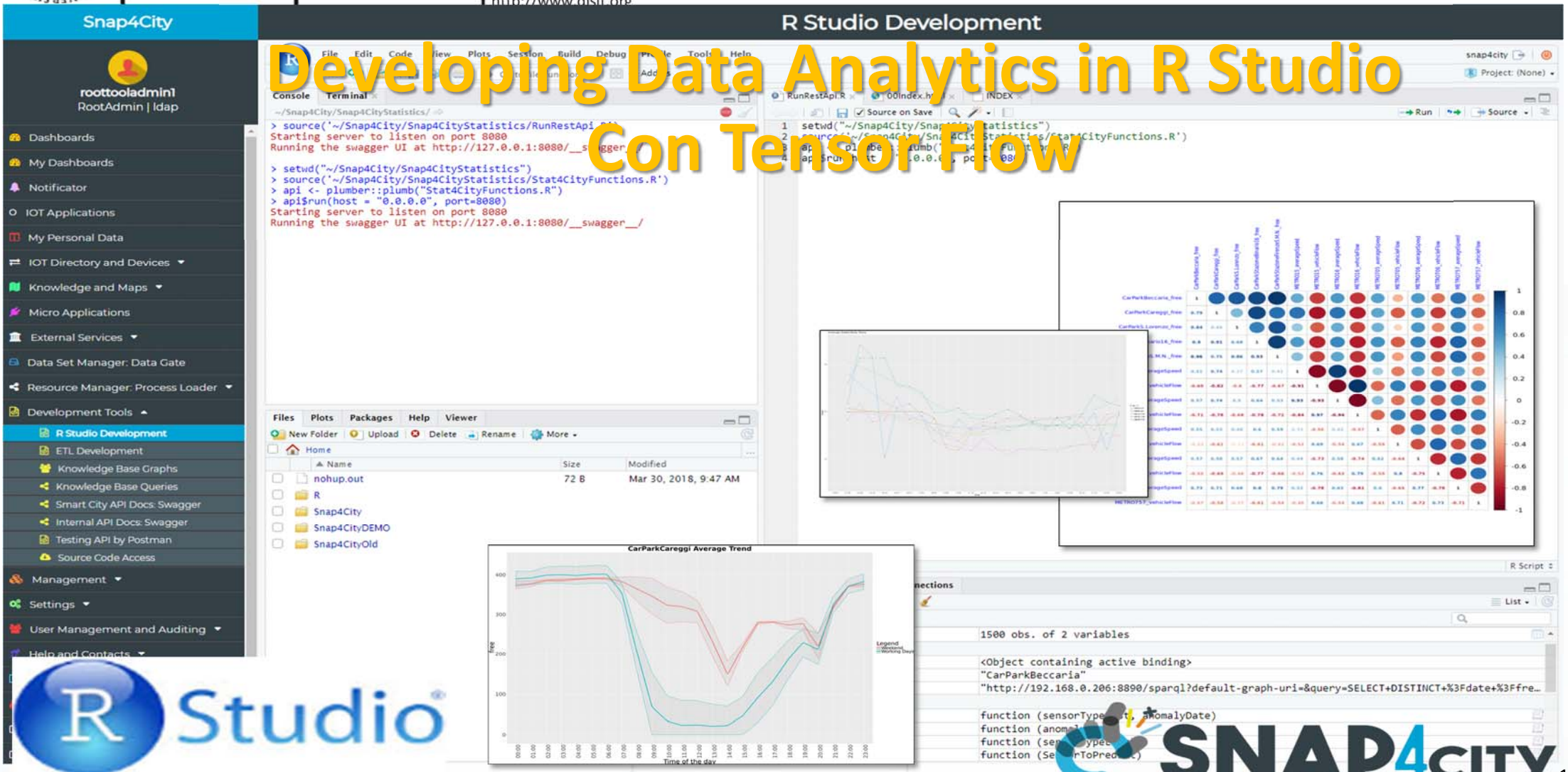
UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB
http://www.disit.org

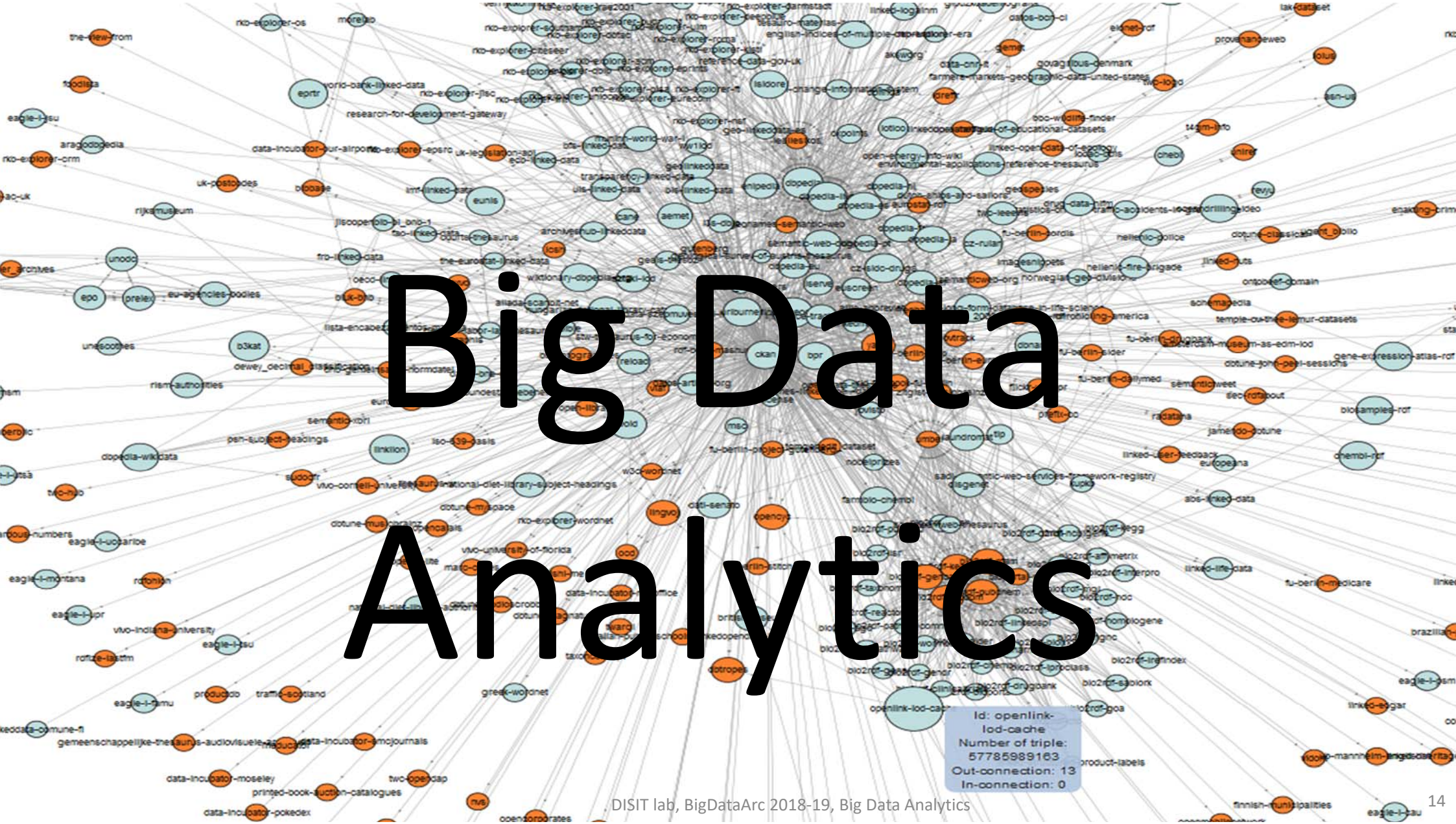
Developer in R Studio + Tensor Flow





Developing Data Analytics in R Studio Con Tensor Flow



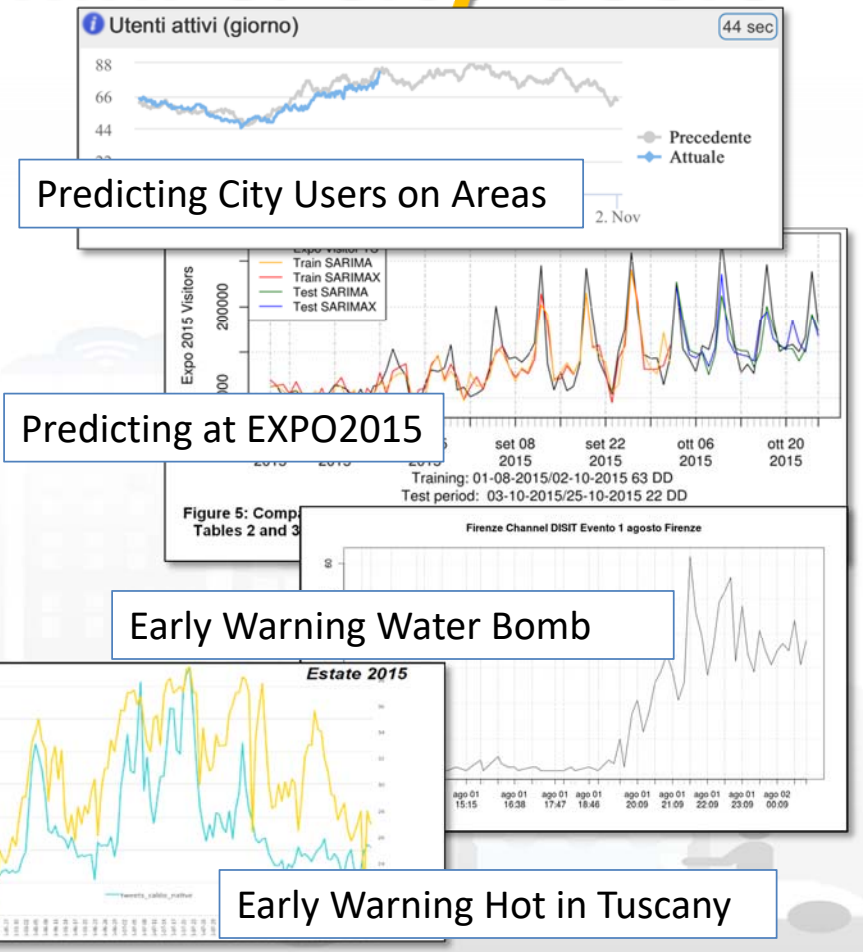

Big Data
Analytics
Id: openlink-lod-cache
Number of triple: 57785989163
Out-connection: 13
In-connection: 0

Big Data analytics Aree Applicative DISIT Lab

- Smart manufacturing
- Personal assistants
- Autonomous engines, semantic reasoners
- Experts systems, decision support systems
- Smart Cloud, elastic computing
- Services and microservices integration
- Mobilità e Trasporti
- Smart City, Innovation Lab, Living Lab
- Servizi alla persona
- Industrie farmaceutiche
- Turismo e Cultura

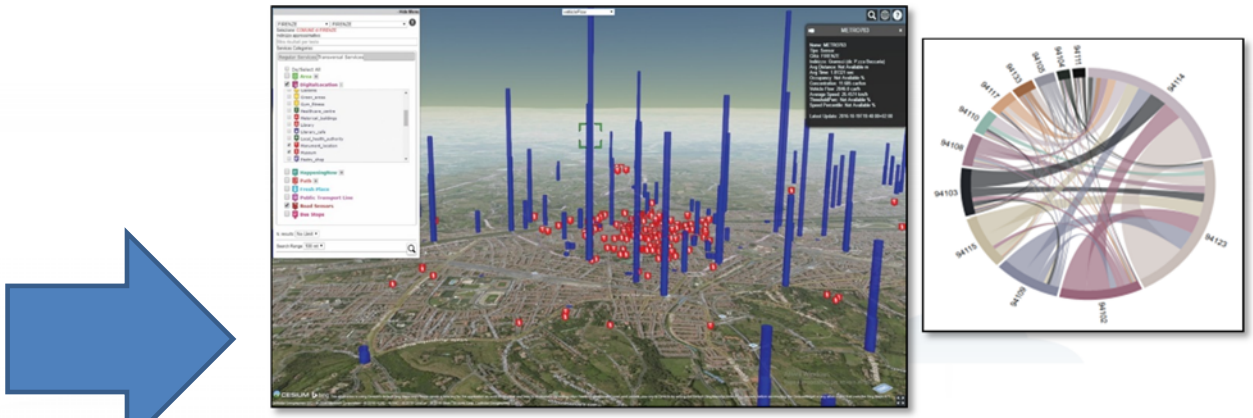
Predicting Models for Admin. & City Users

- Aiming at improving
  - quality of service,
  - distributing workload
  - early warning
- Traffic Flows & People Flows
  → crowd , #number of people
- Parking Status → free slots
- Weather Forecast (LAMMA)

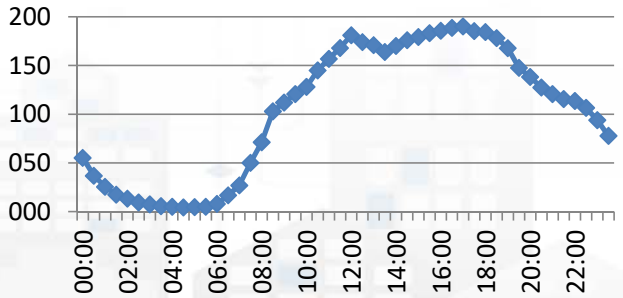


Predicting City users movements

- **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**
  - **…….**

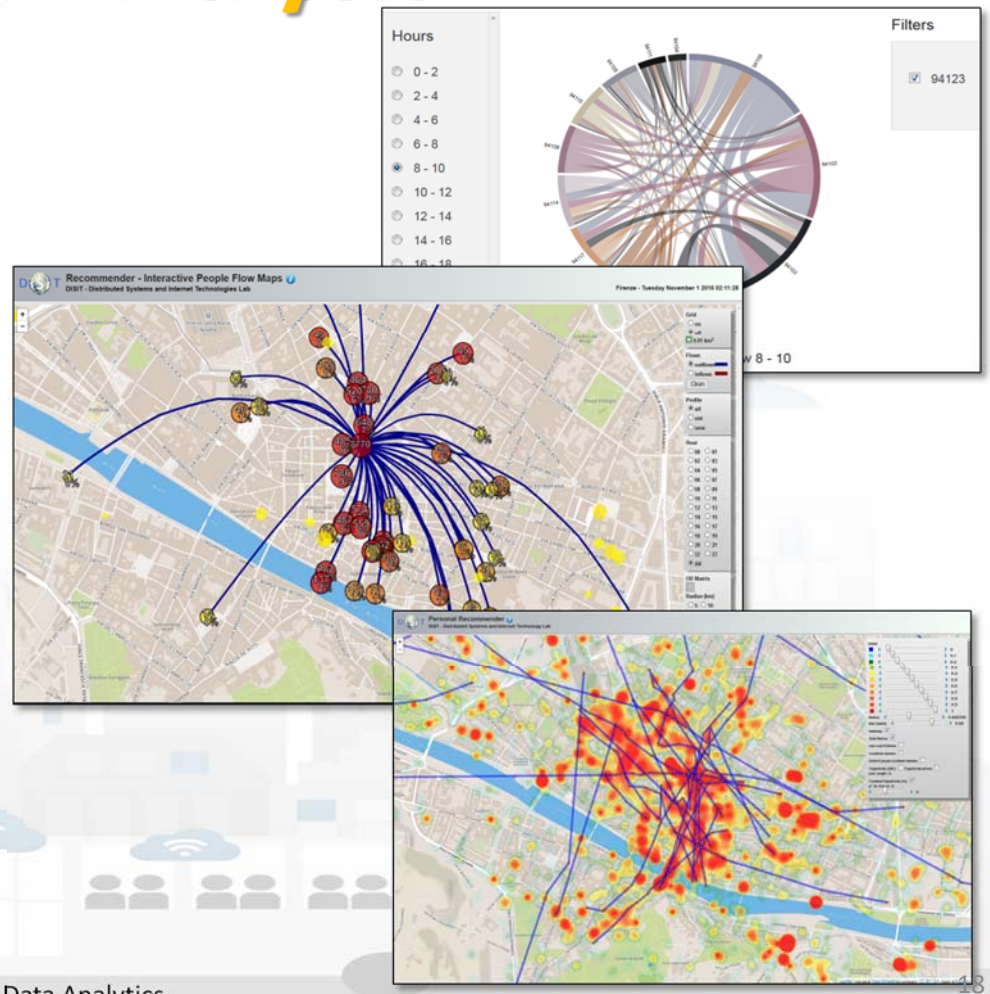


- Daily trends
- OD matrices
- Trajectories
- Prediction models



User Behaviour Analysis

- Monitoring movements of vehicles by
  - traffic flow sensors
  - Spires and virtual spires
- Monitoring users' movements
  - from Mobile Cells
  - Unsuitable for precise tracking and OD production
- Monitoring movements via Wi-Fi
- Monitoring movements and much more from mobile Apps



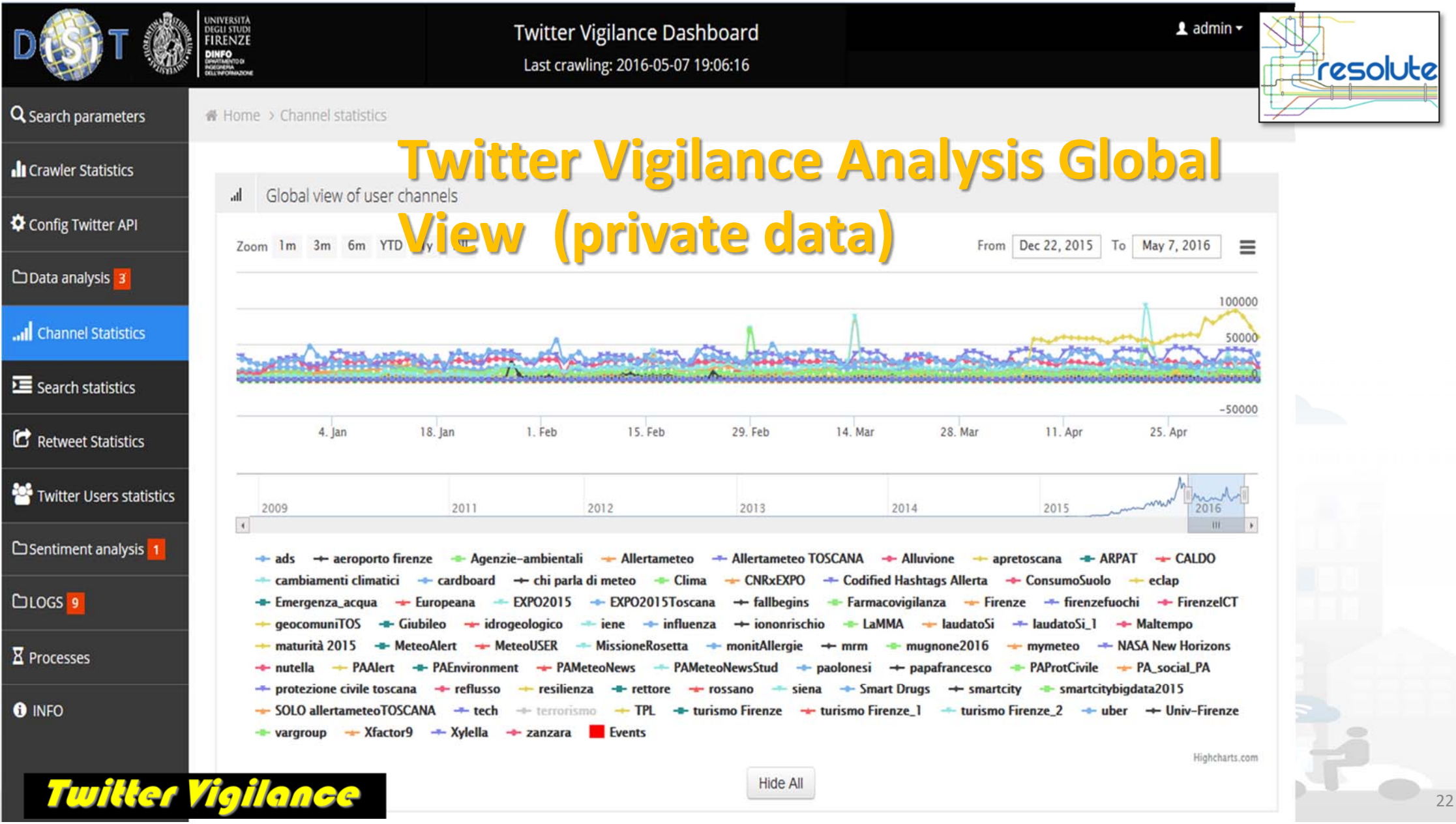
Overview

- **Analisi Dati raccolti via Social Media**
  - Predicting presences
- **Analisi Dati raccolti via Mobile App**
  - Tracce, matrici OD, heatmap
  - Regency e frequency
  - Impatto in uscita da Firenze
- **Analisi Dati raccolti dai Flussi di traffico**
  - Ricostruzione del traffico in punti non misurati
- **Analisi Dati raccolti dai Parcheggi**
  - Predizione dei posti liberi
- **Analisi Dati raccolti via Firenze WiFi**
  - Tracce, matrici OD, heatmap
  - Predicting presences
- **Analisi Dati Dati raccolti via Cellular**
  - Valutazione comparativa TIM-VODA
  - Valutazione comparativa FirenzeWiFi-Tim-Voda

Social Media

Overview

- **Analisi Dati raccolti via Social Media**
  - Predicting presences
- **Analisi Dati raccolti via Mobile App**
  - Tracce, matrici OD, heatmap
  - Regency e frequency
  - Impatto in uscita da Firenze
- **Analisi Dati raccolti dai Flussi di traffico**
  - Ricostruzione del traffico in punti non misurati
- **Analisi Dati raccolti dai Parcheggi**
  - Predizione dei posti liberi
- **Analisi Dati raccolti via Firenze WiFi**
  - Tracce, matrici OD, heatmap
  - Predicting presences
- **Analisi Dati Dati raccolti via Cellular**
  - Valutazione comparativa TIM-VODA
  - Valutazione comparativa FirenzeWiFi-Tim-Voda

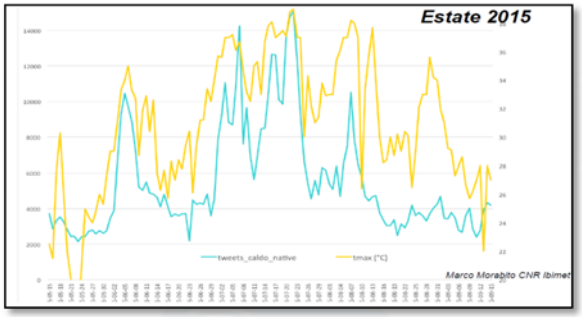


Twitter Vigilance Analysis Global View (private data)

Twitter Vigilance

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



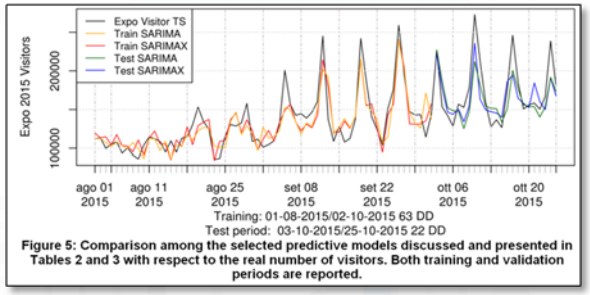
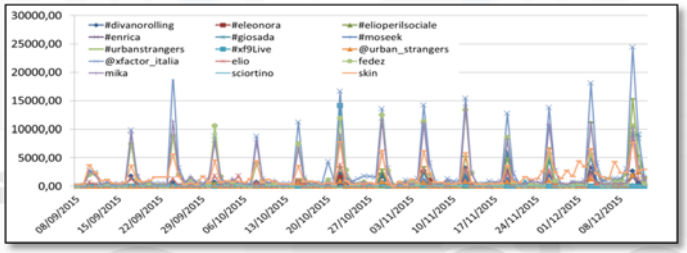
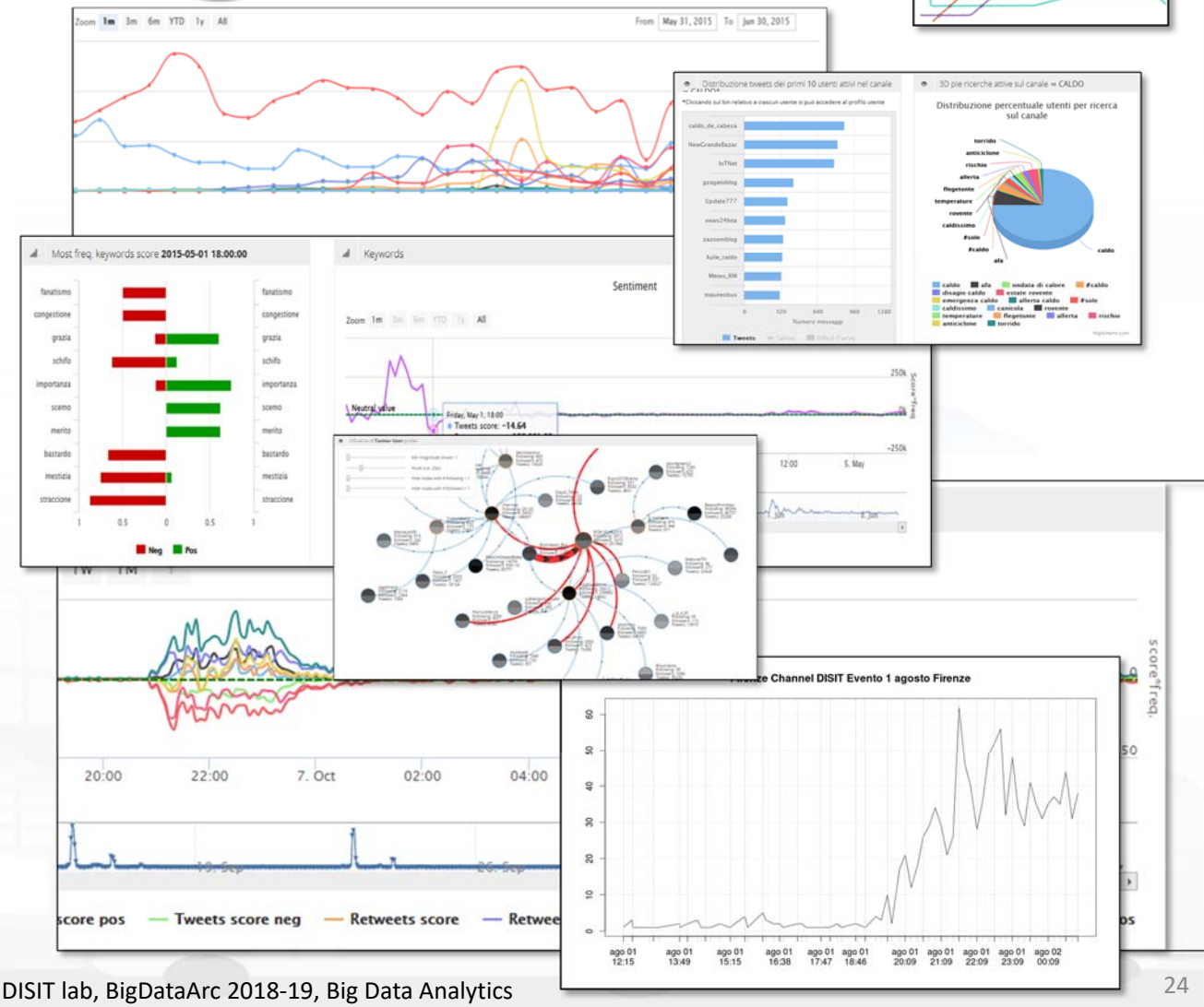


Figure 5: Comparison among the selected predictive models discussed and presented in Tables 2 and 3 with respect to the real number of visitors. Both training and validation periods are reported.



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

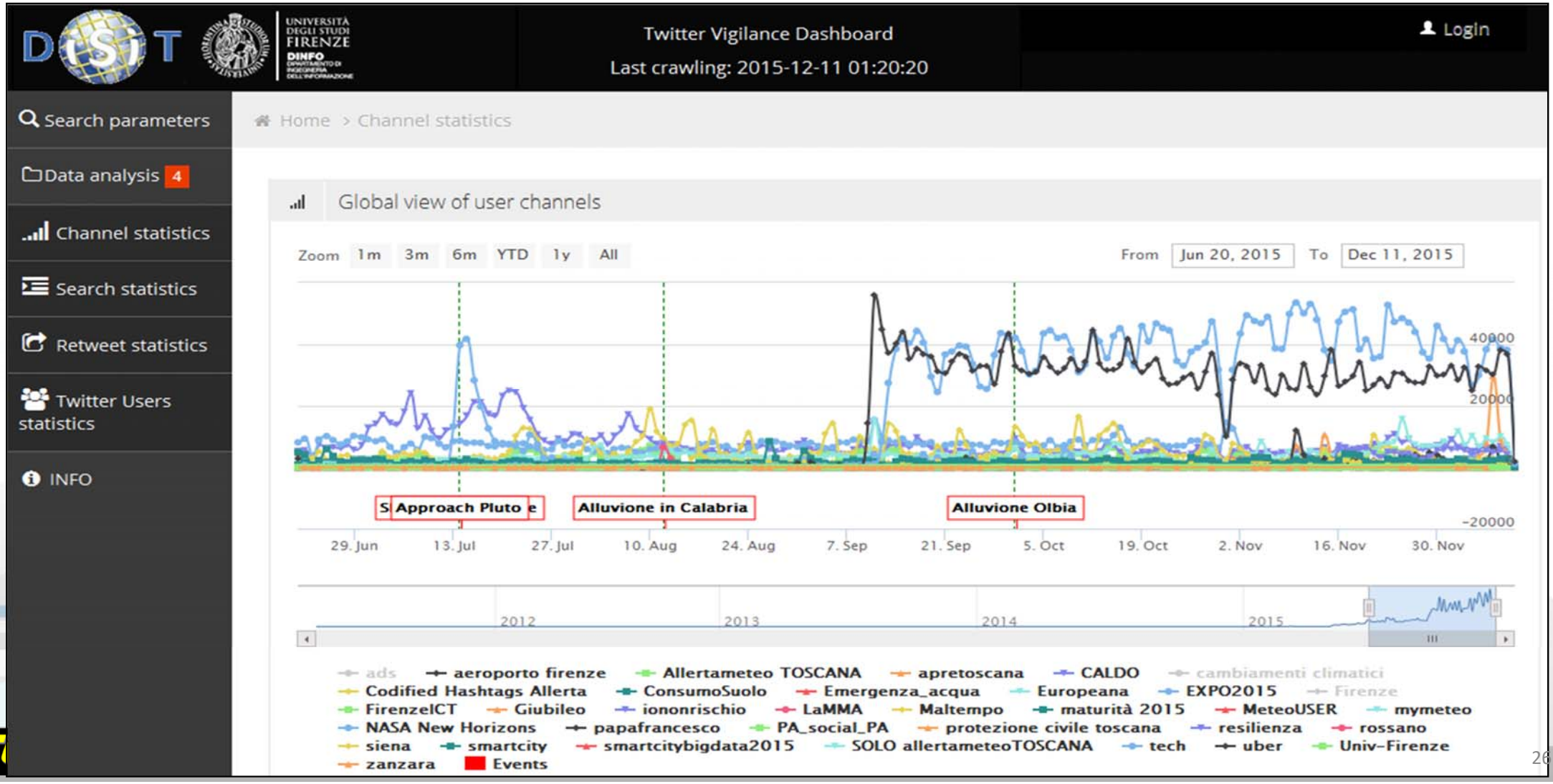
Twitter Vigilance in Numbers

- **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.

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

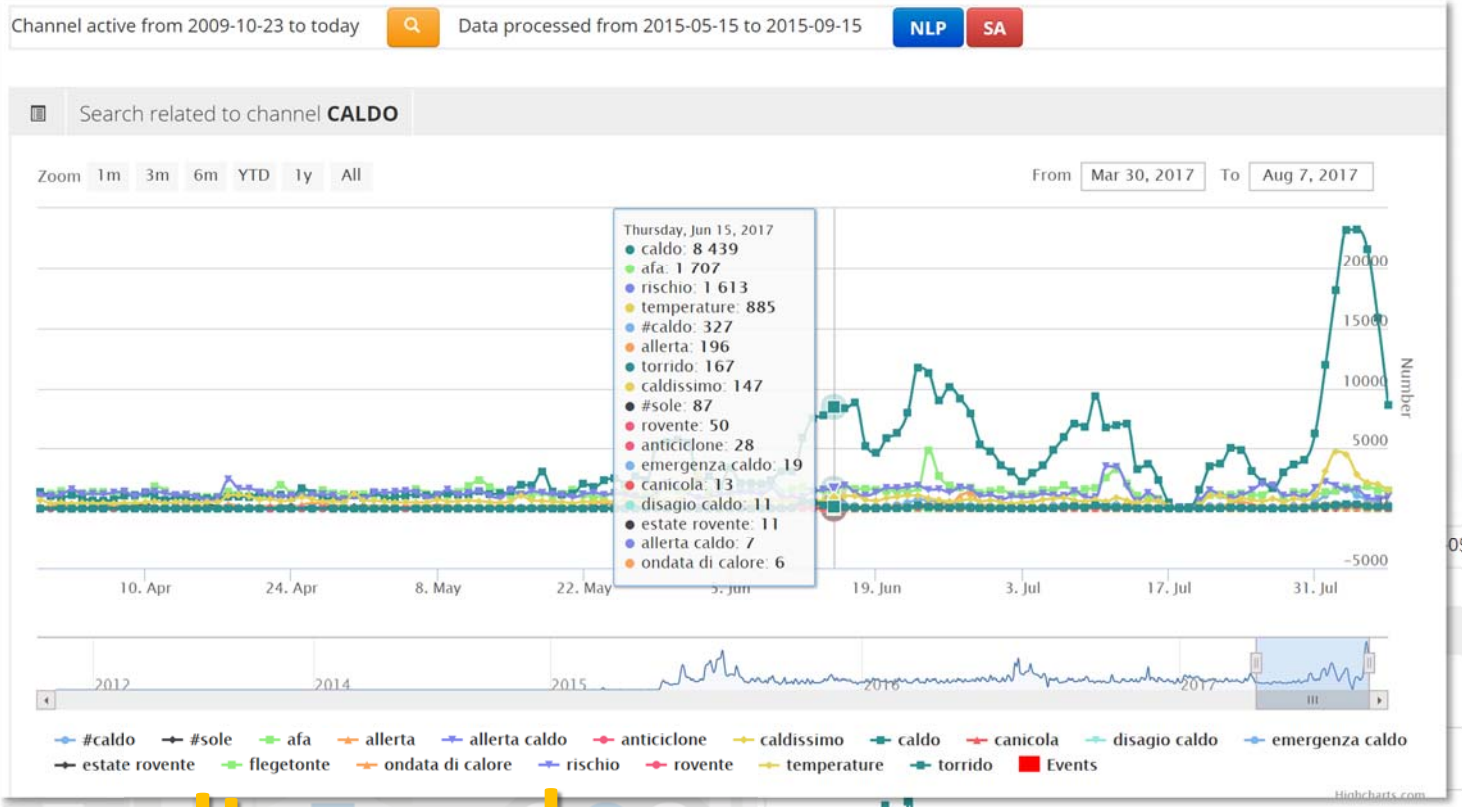
Several Channels

Sii-Mobility

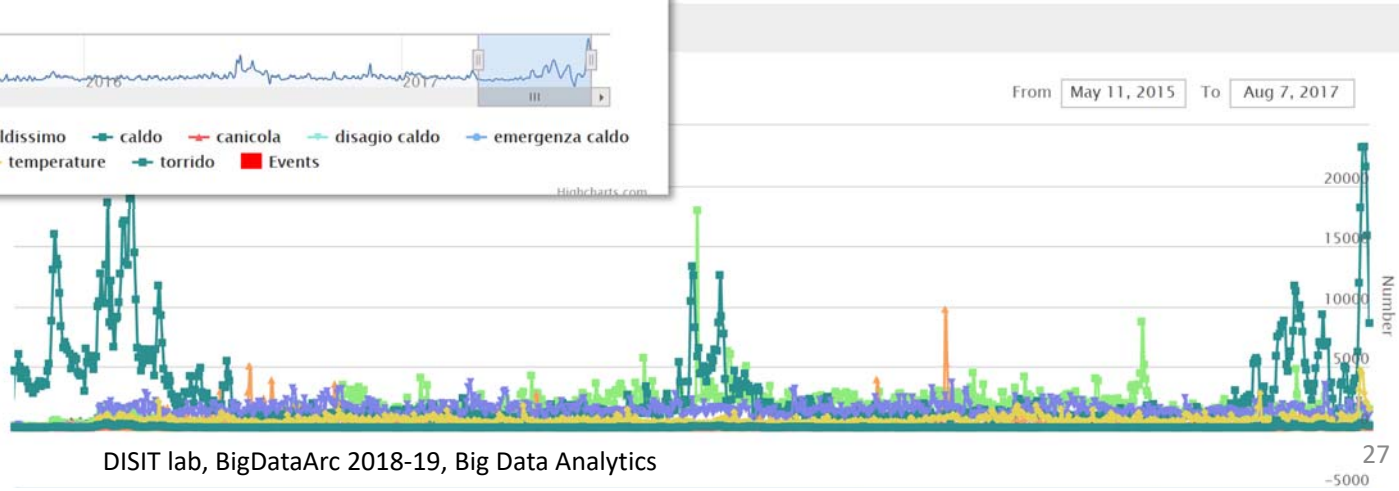


A Channel

Sii-Mobility

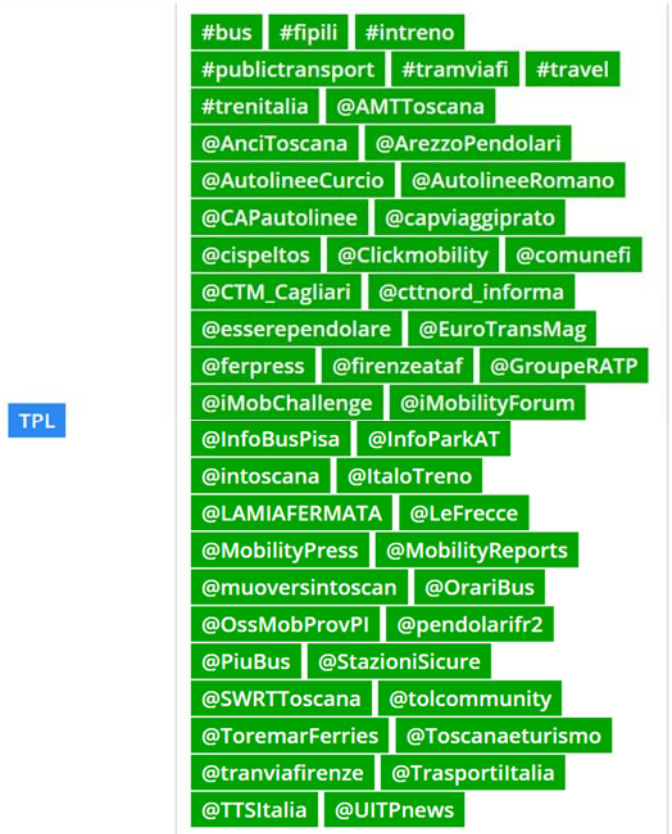


Its searches



Twitter Syntax for Searches

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



TPL

Twitter Vigilance

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.

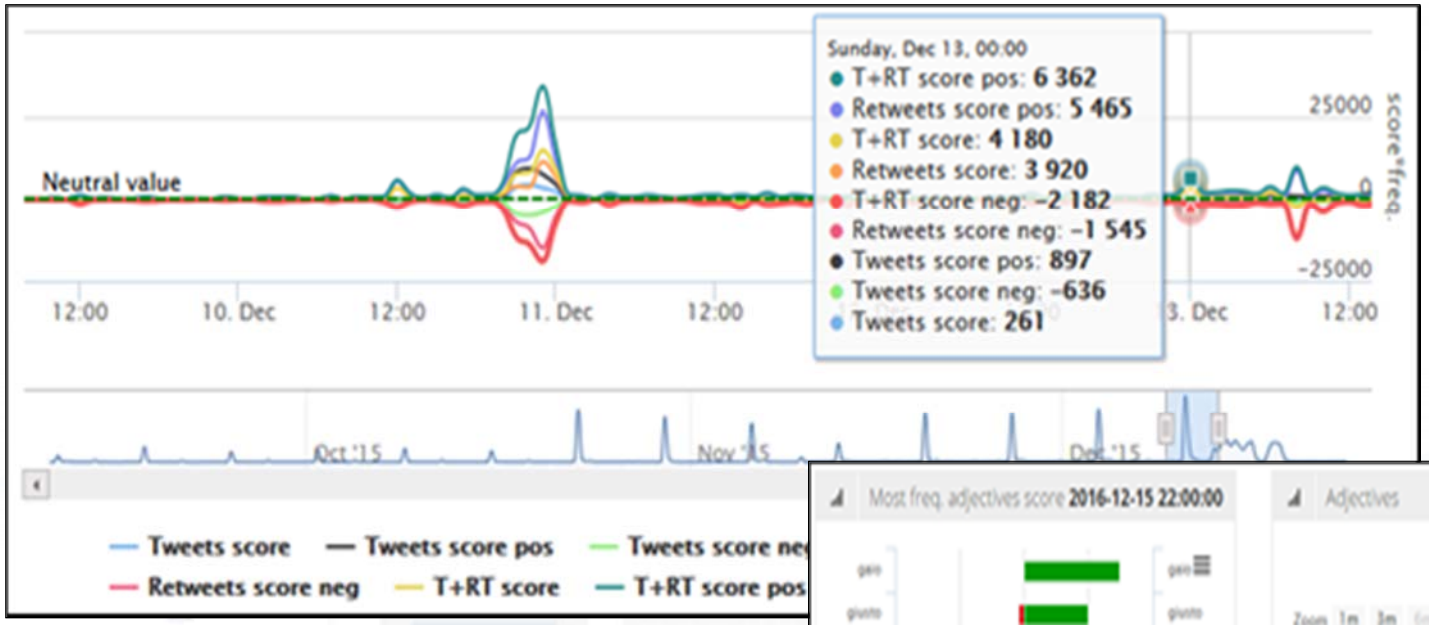
Sentiment Analysis

ARC
INDEX

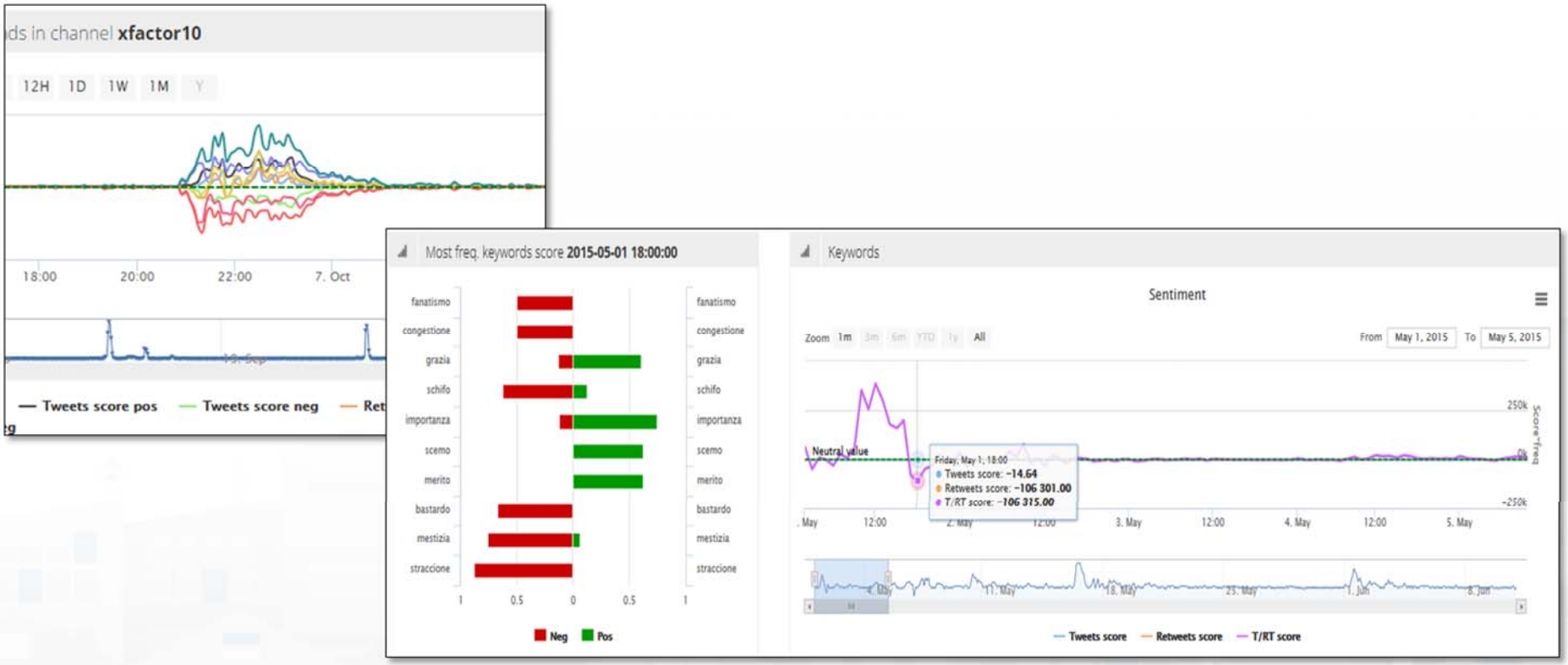
UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB
http://www.disit.org

Sentiment Analysis

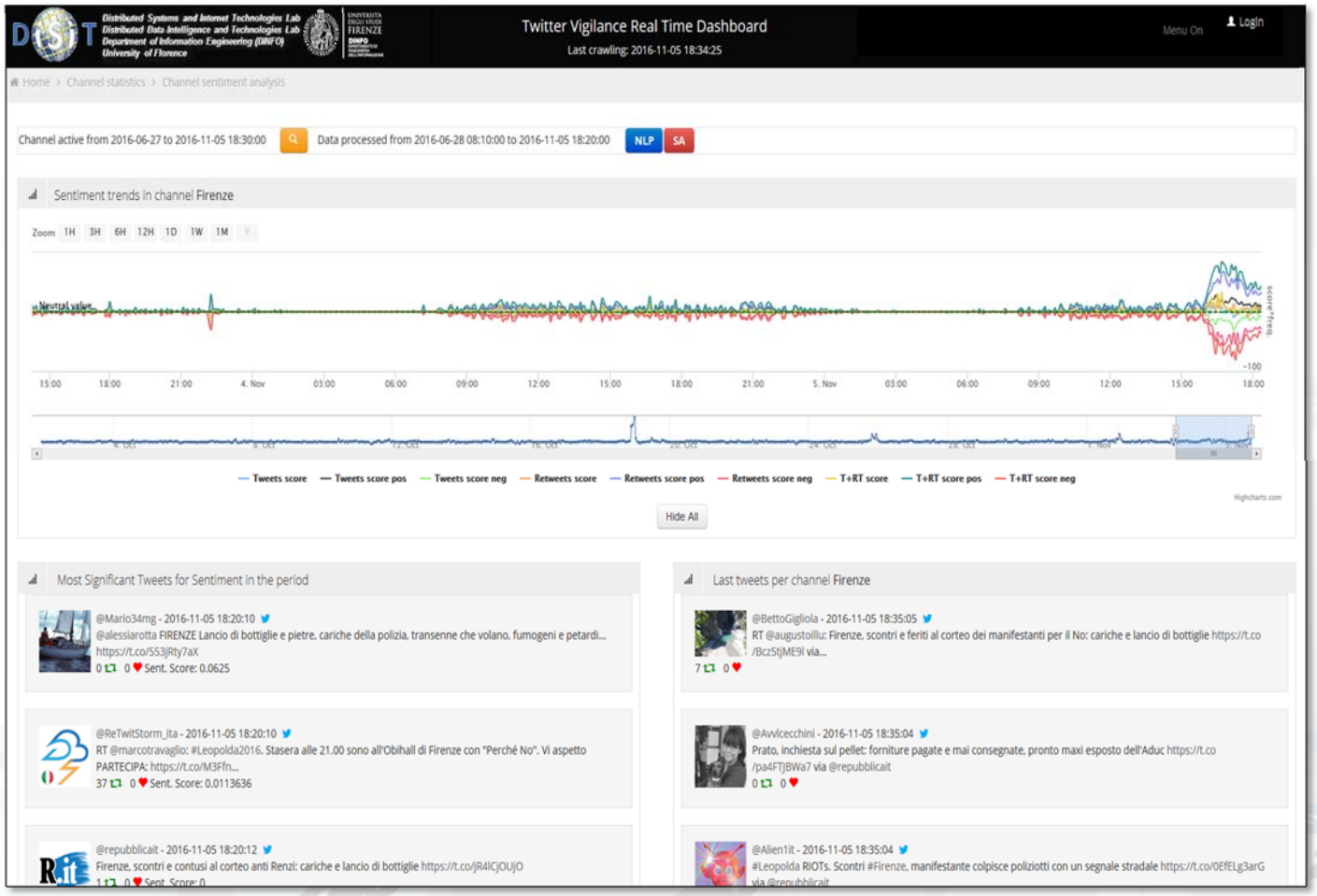
Sii-Mobility



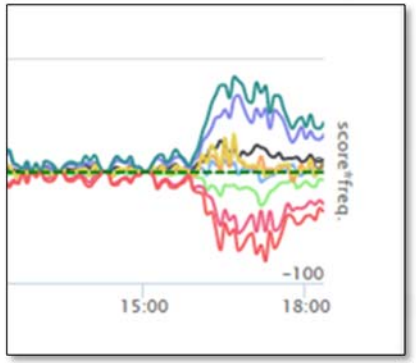
Twitter Vigilance

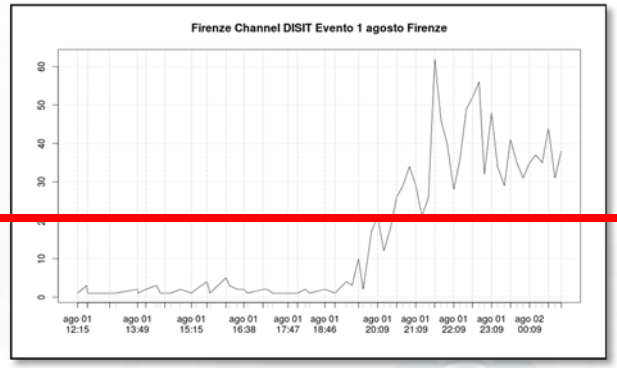

ds in channel xfactor10
12H 1D 1W 1M Y
18:00 20:00 22:00 7. Oct
Tweets score pos
Tweets score neg
Most freq. keywords score 2015-05-01 18:00:00
fanatismo
congestione
grazia
schifo
importanza
scemo
merito
bastardo
mestizia
straccione
1 0.5 0 0.5 1
Neg Pos
Keywords
Sentiment
Zoom 1m 3m 6m YTD 1y All
From May 1, 2015 To May 5, 2015
Neutral value
Friday, May 1, 18:00
Tweets score: -14.64
Retweets score: -106 301.00
T/RT score: -106 315.00
250k
-250k
Score*freq
12:00
3. May
4. May
5. May
Tweets score
Retweets score
T/RT score

Real Time Twitter Vigilance, Early Warning



Sentiment Analysis

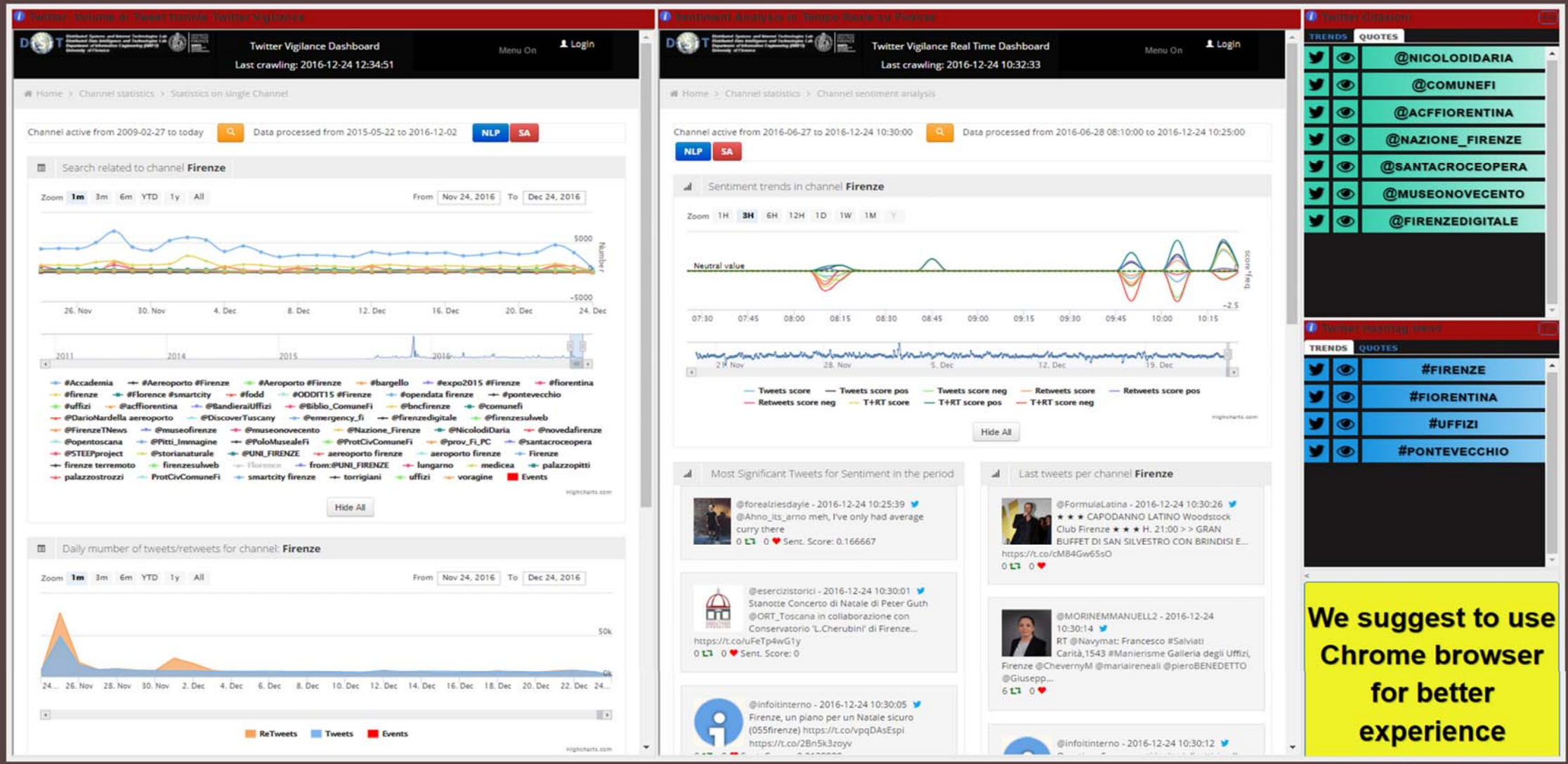




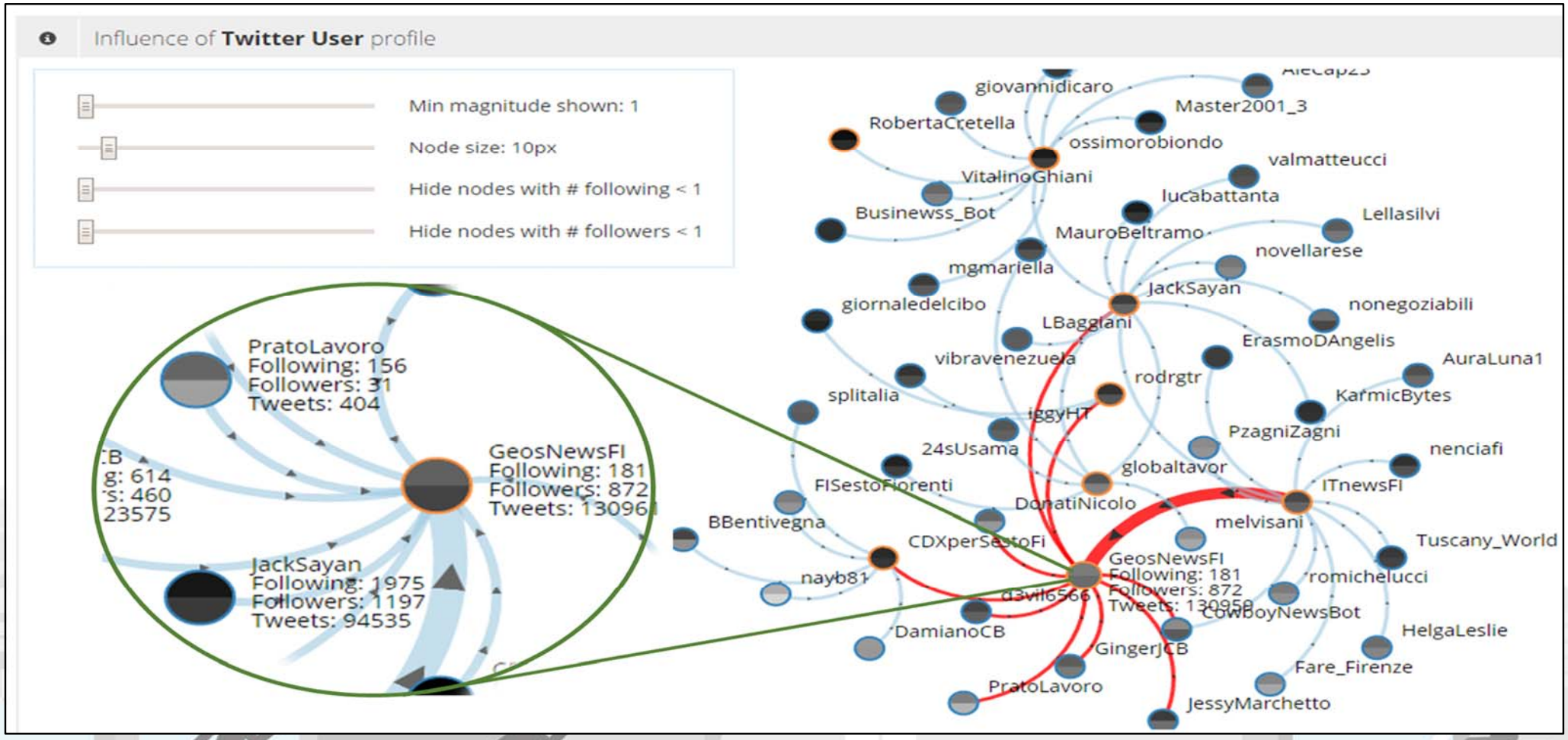
Twitter Vigilance

Twitter Vigilance su Firenze
(sperimentale)

Sat 24 Dec @ 10:37:57



Influence Network



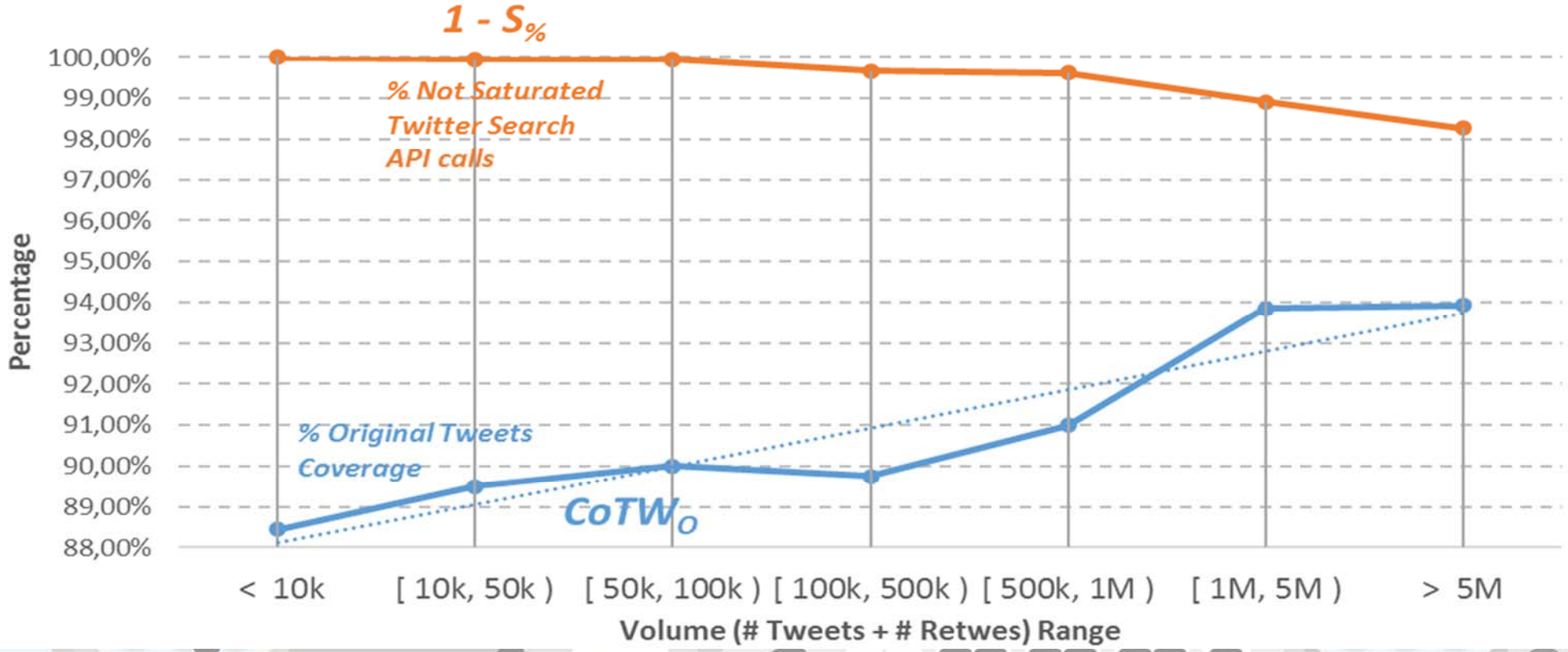
Reliability in collecting tweets

ARC
INDEX

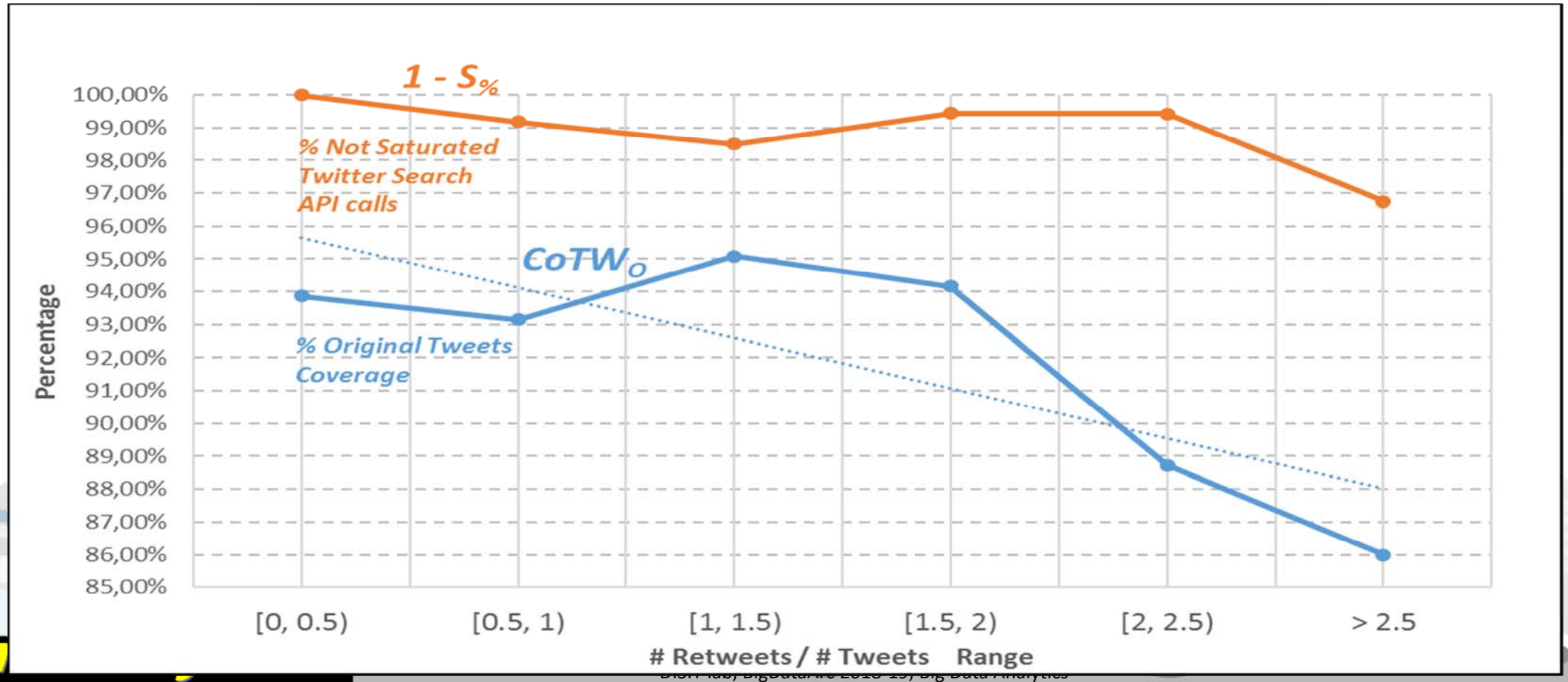
Efficiency in retrieval

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



Predicting Audience

ARC
INDEX

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

Twitter Metrics

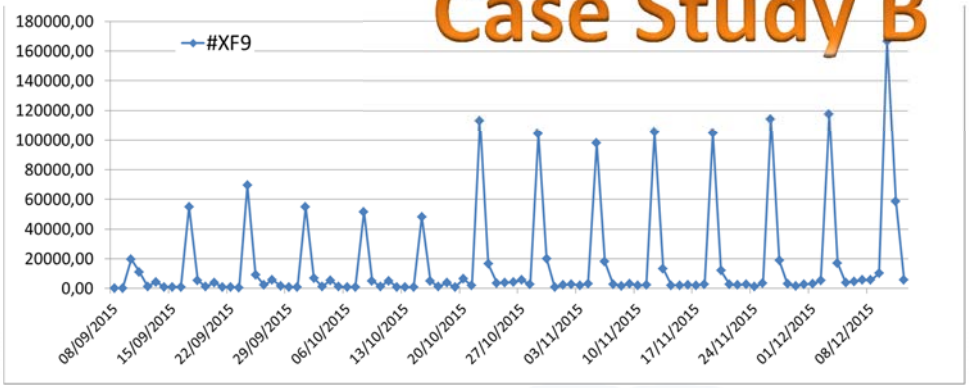
- TW: Number of Tweets per **Search/Channel** (as called Volume) , per day, per hour
- RTW: Number of ReTweets per **Search/Channel**, per day, per hour
- NRT/TW: ratio from ReTweets and Tweets per **Search/Channel**, per day, per hour
- NumSearch: number of Tweets including the Search per **Channel,**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

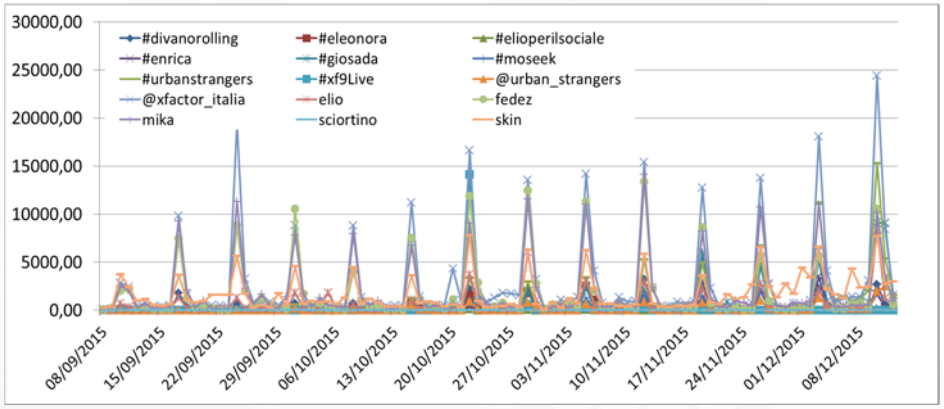
Predicting Audience: X-Factor, PechinoExpress…

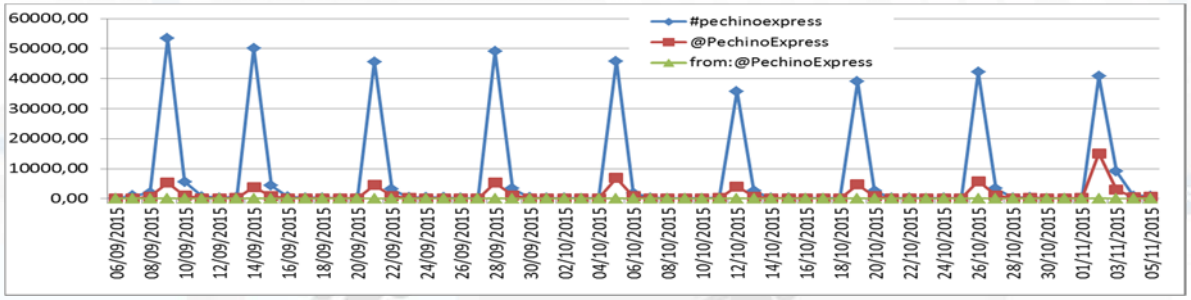
Case Study B

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

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

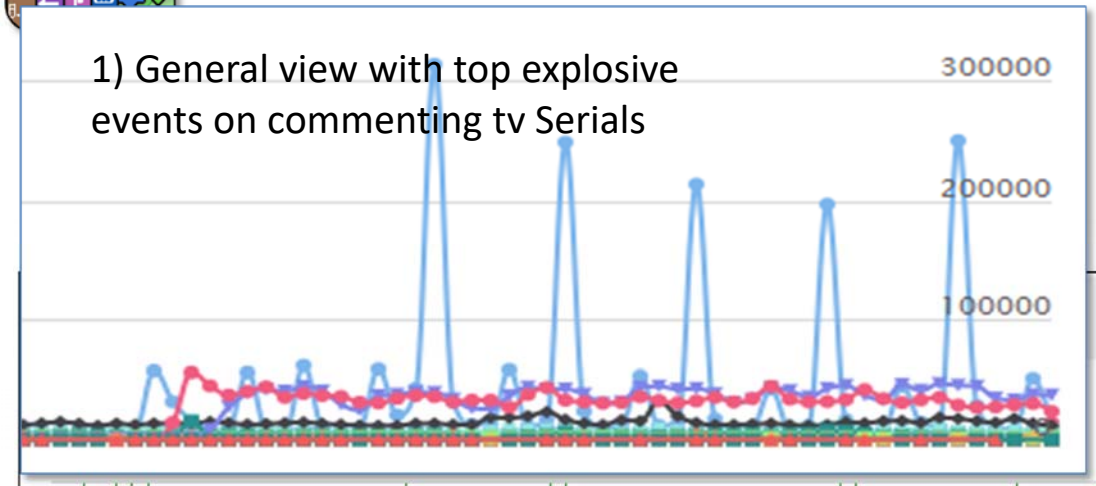


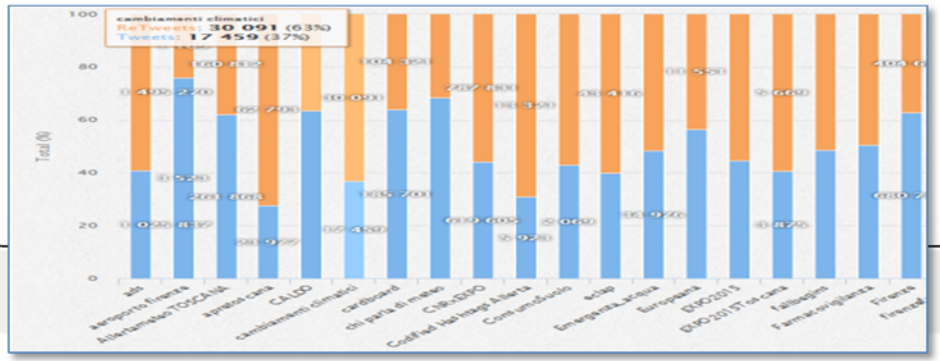




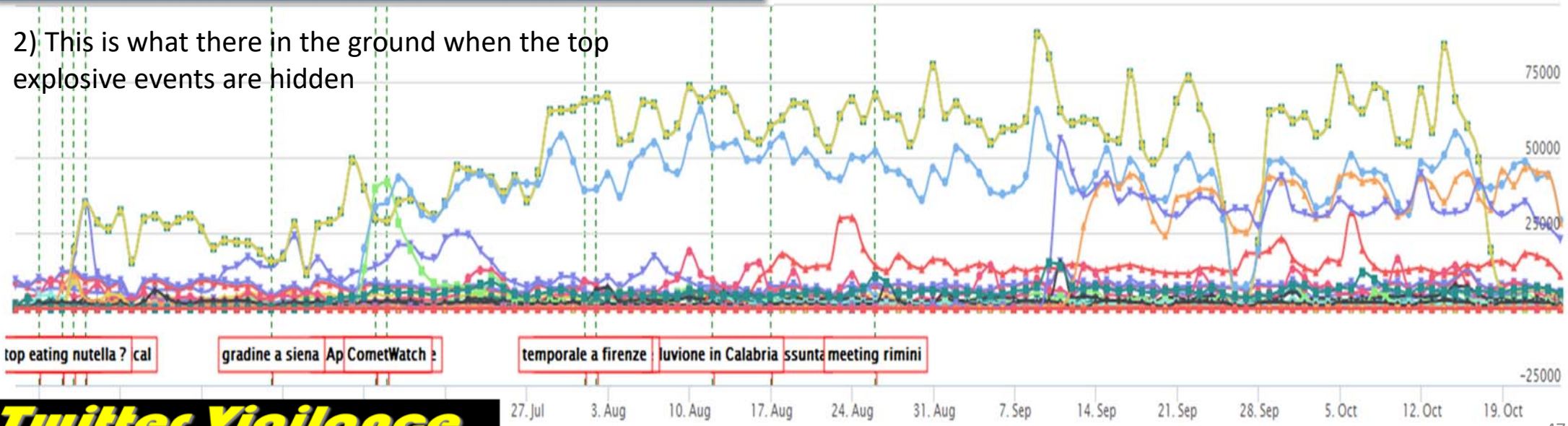
All Channels (private information)

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

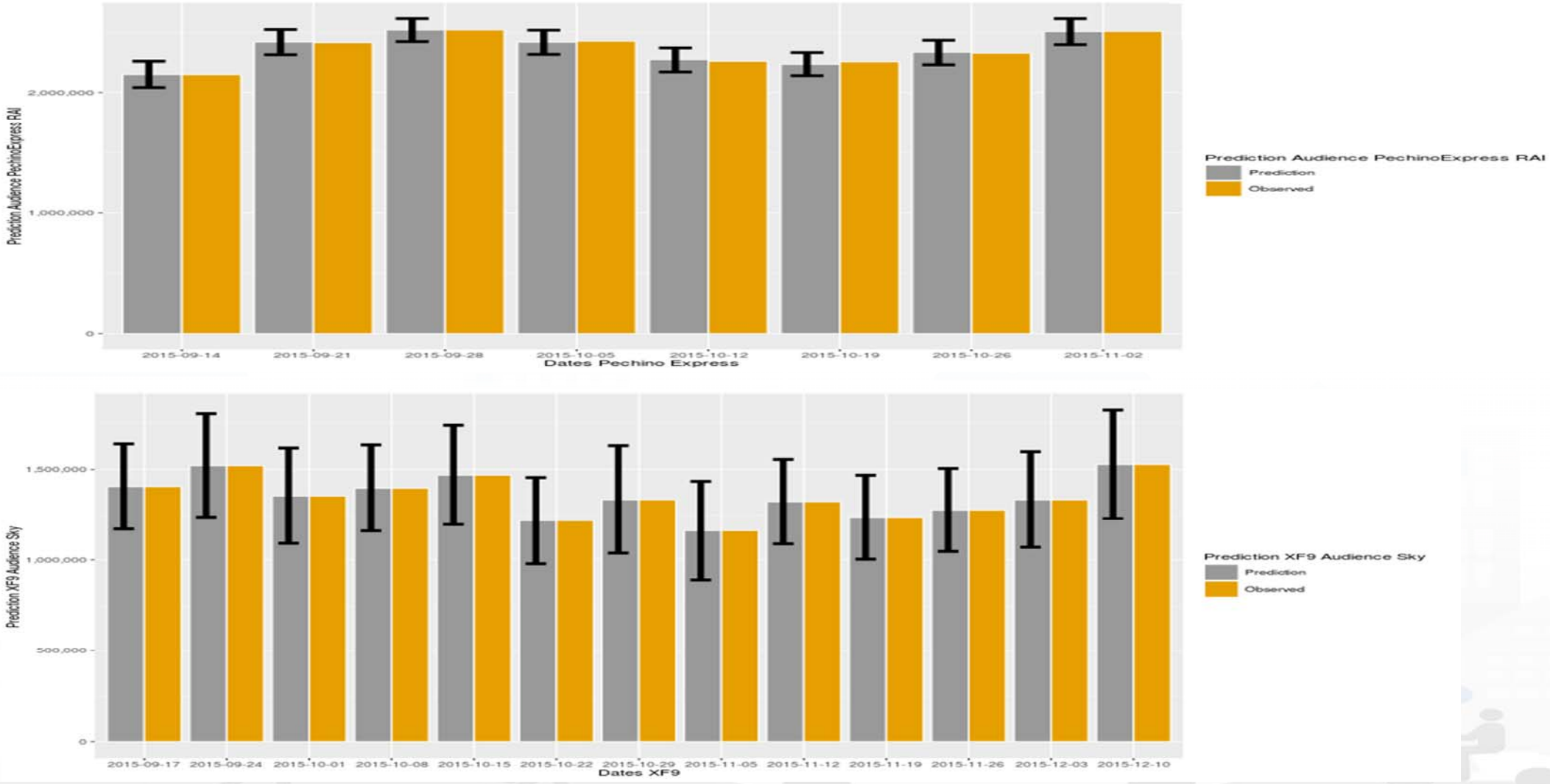
Details of Predictive Models Validities

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			

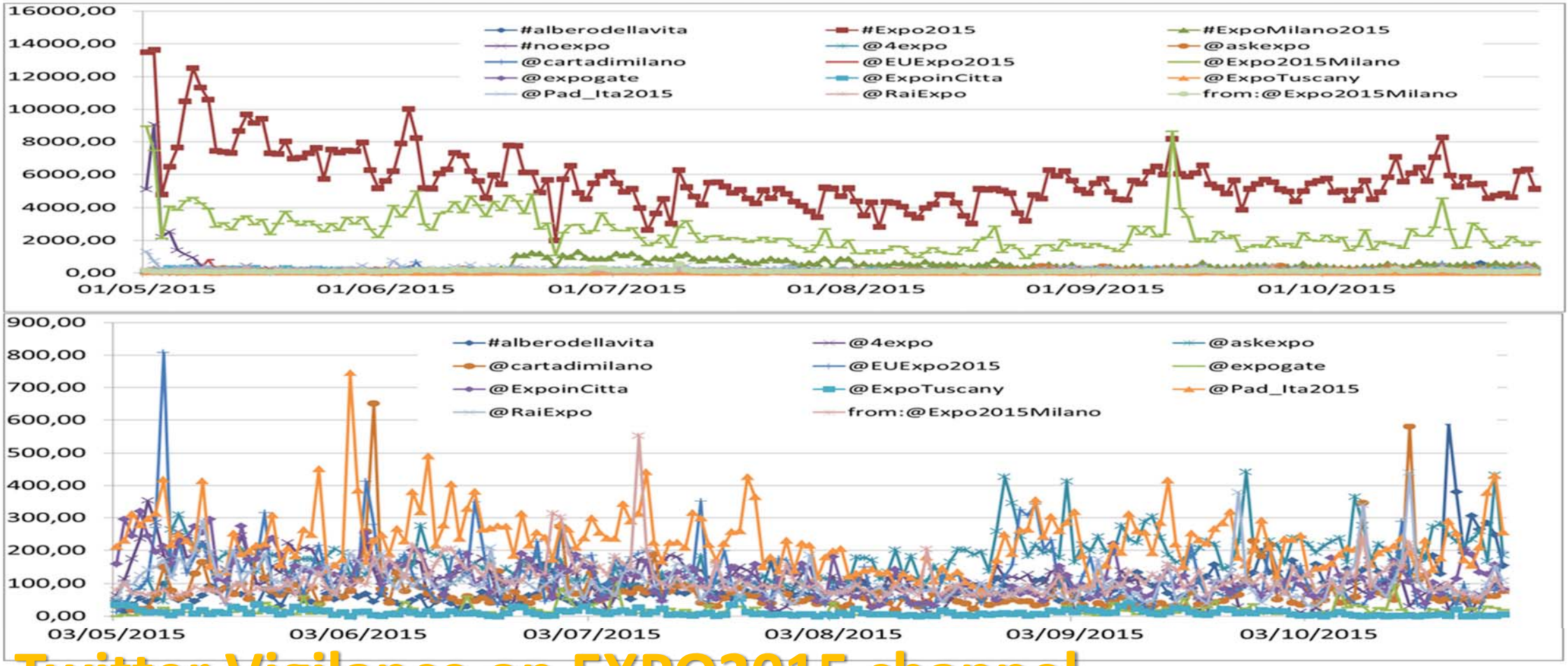
Case Study B

Predicting Confidence

Case Study B



Predicting EXPO2015



Twitter Vigilance on EXPO2015 channel

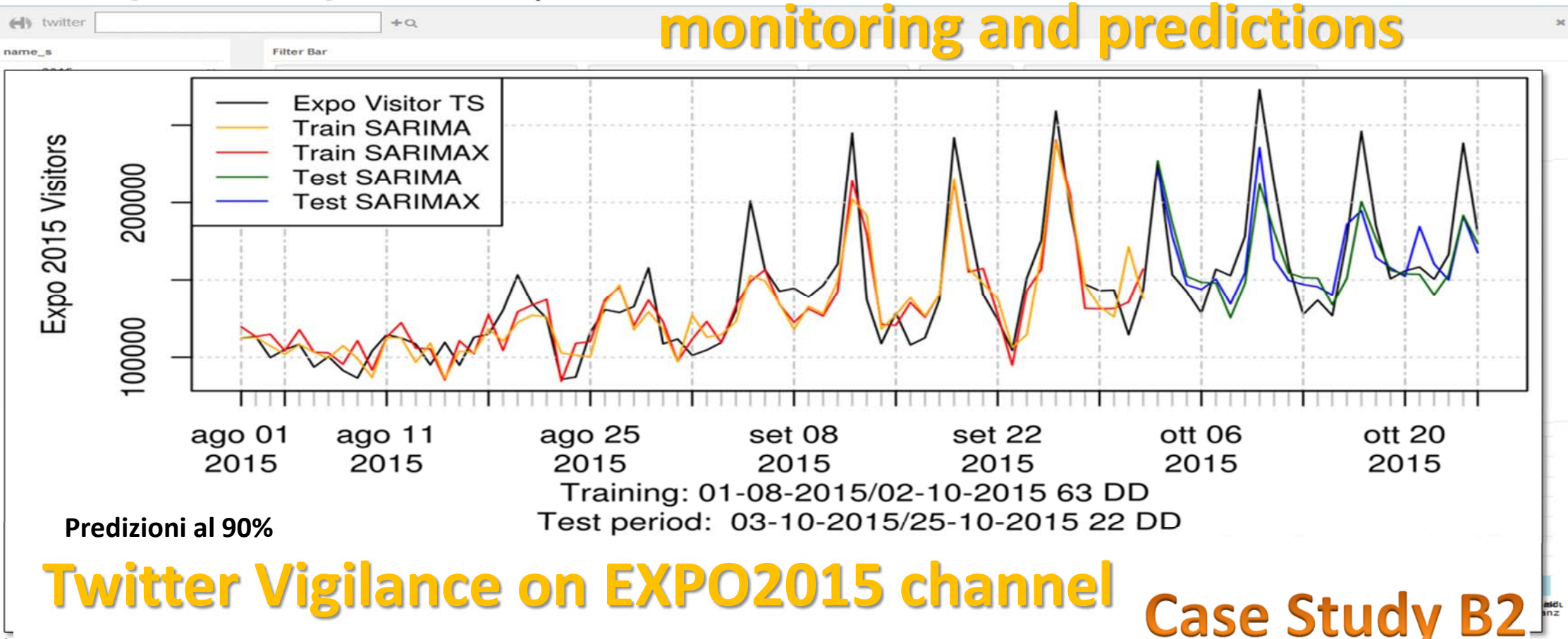
Twitter Vigilance

Predicting presences at events

ARC
INDEX

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%

Twitter Vigilance on EXPO2015 channel

Case Study B2

Predicting volume of visitors for tuning the services

Twitter Vigilance

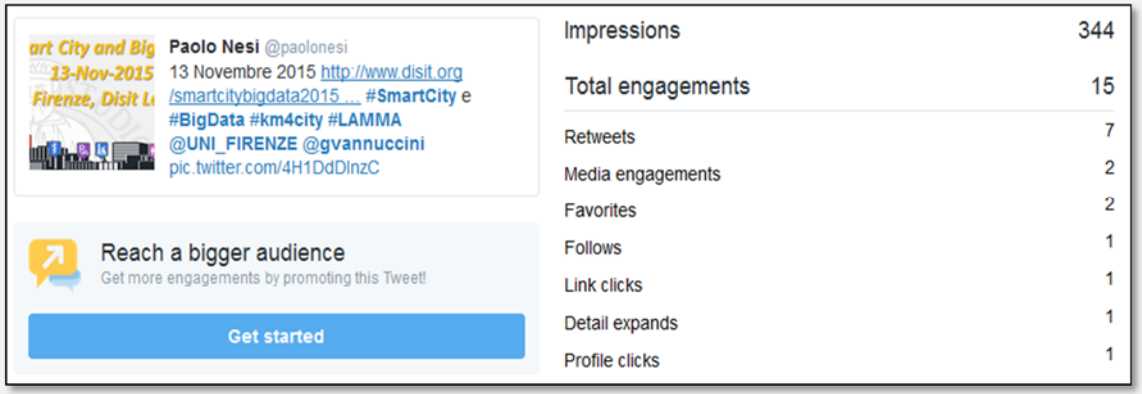
Predicting Twitter Proneness

ARC
INDEX

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



Case Study C

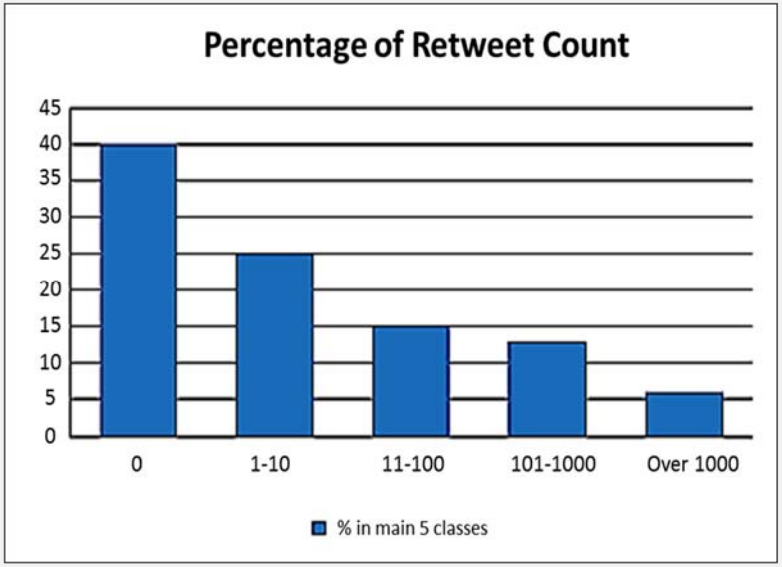
DISIT Lab, Distributed Data Intelligence and Technologies
Distributed Systems and Internet Technologies
Department of Information Engineering (DINFO)
http://www.disit.dinfo.unifi.it

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

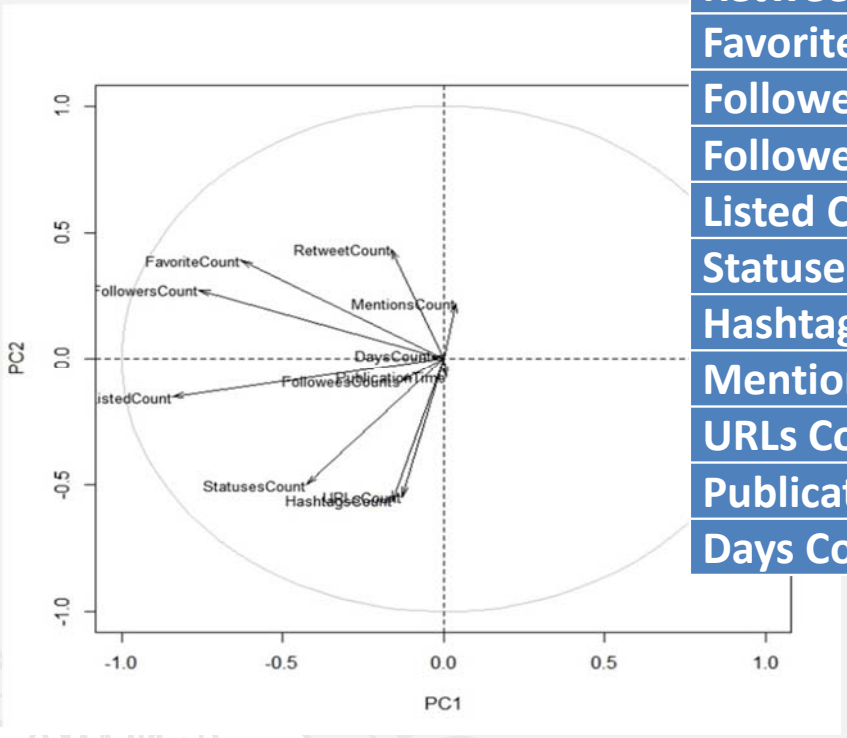
Data sets:

- 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

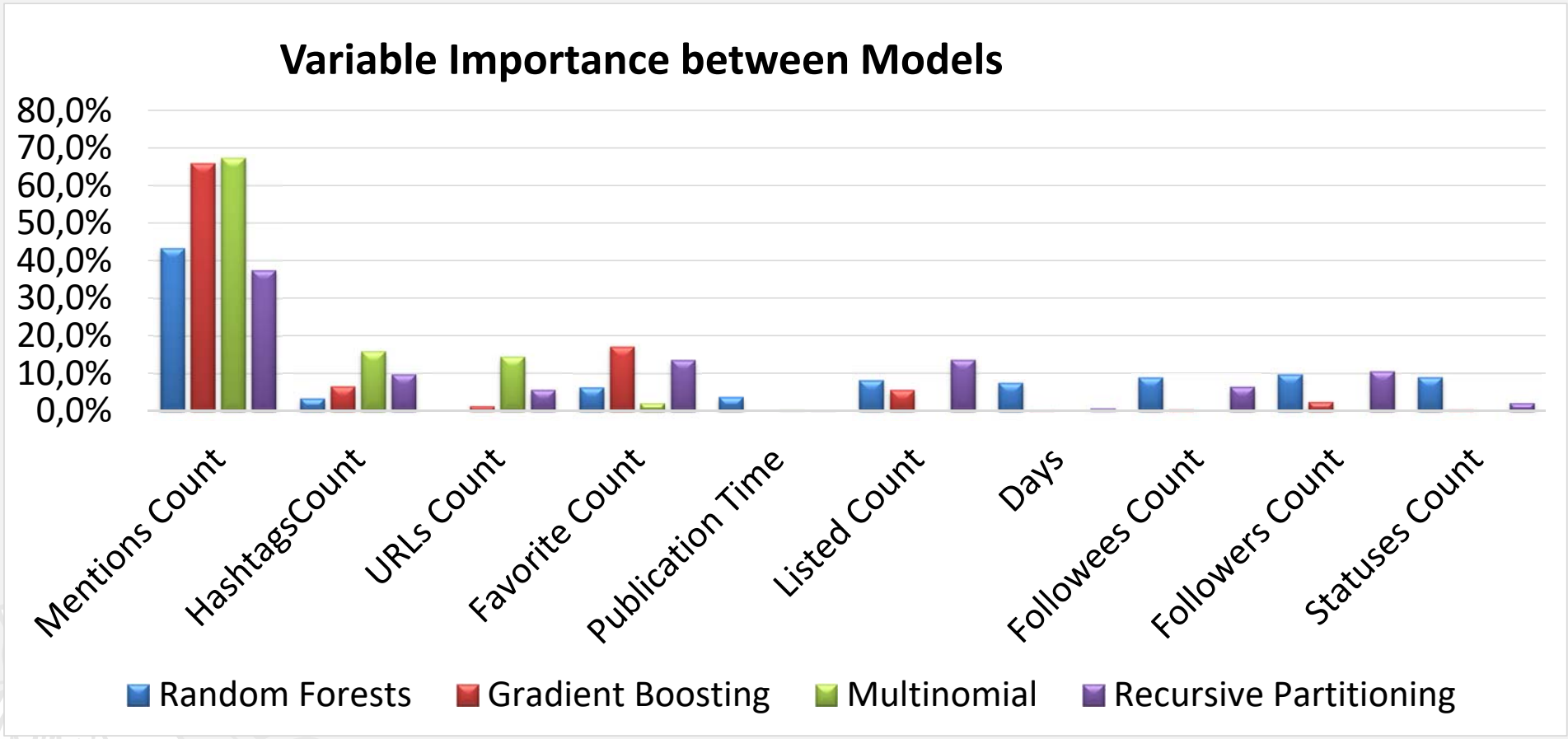
Classifier Models	Accuracy	Precision	Recall	F_1 score	Processing Time in sec.
Recursive Partitioning (Stat)	**0.6807**	**0.8512**	**0.7767**	**0.8122**	**180**
Random Forests (ML)	0.6884	0.8601	0.7866	0.8217	198968
Gradient boosting (ML)	0.6796	0.8534	0.7731	0.8113	64448
Multinomial Model (Stat)	0.6411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



Tweets as Early Warning

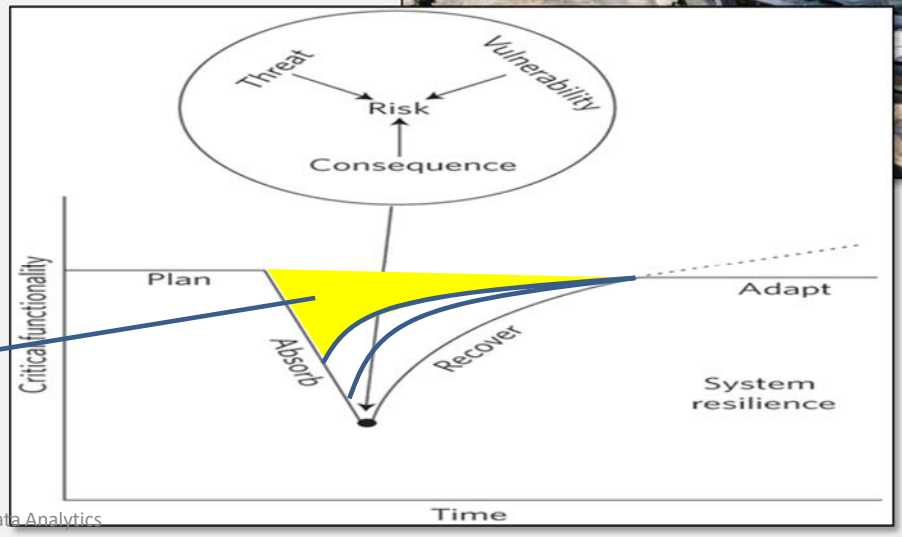
ARC
INDEX

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

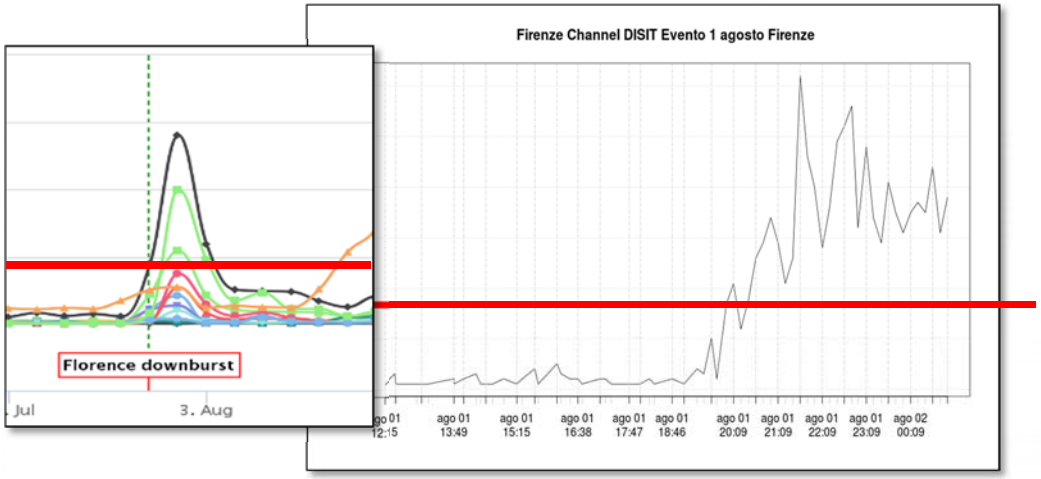
City Resilience

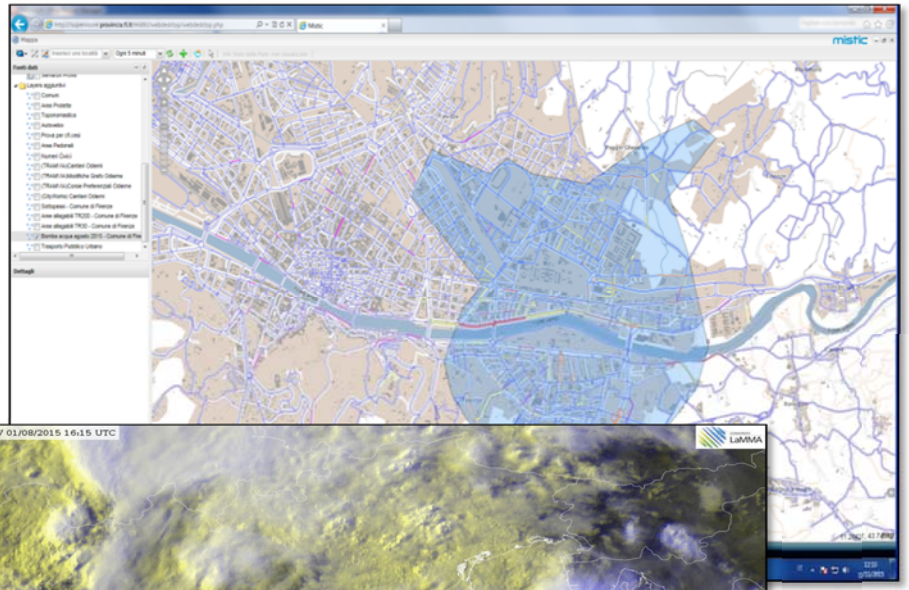


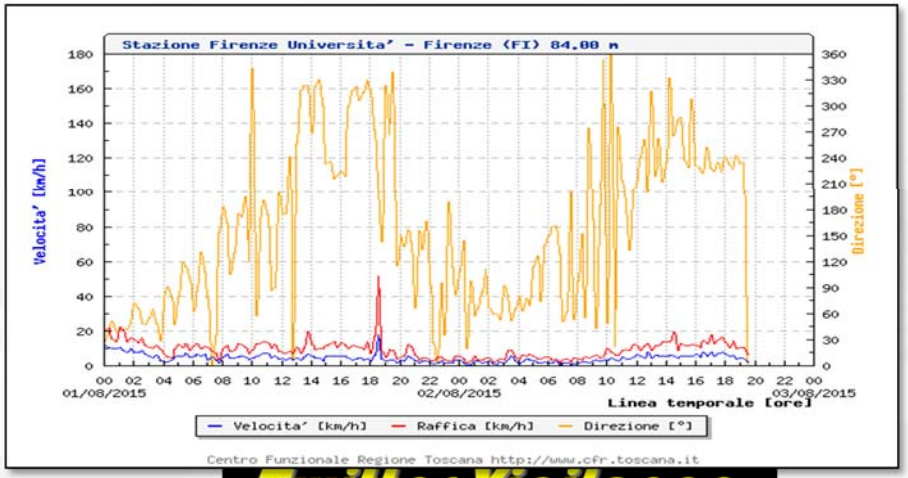
damage

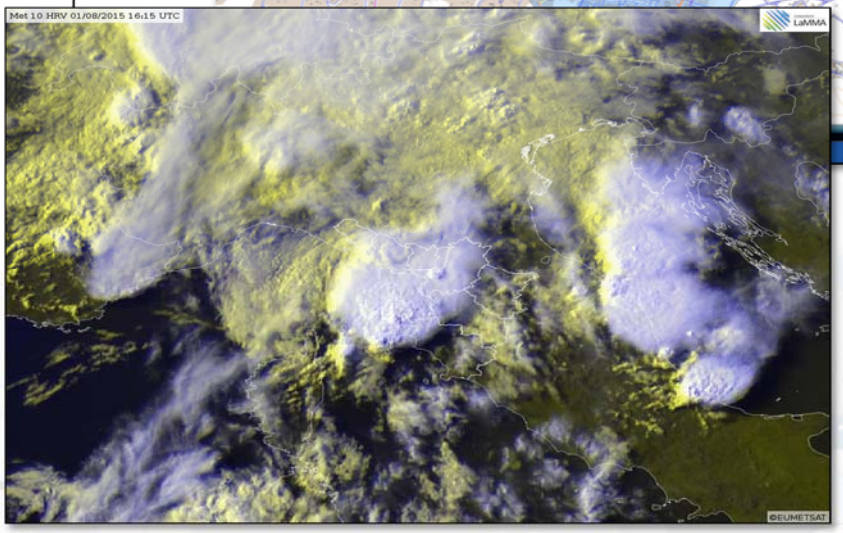
UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB http://www.disit.org

Early Warning
Twitter Vigilance and Water Bomb





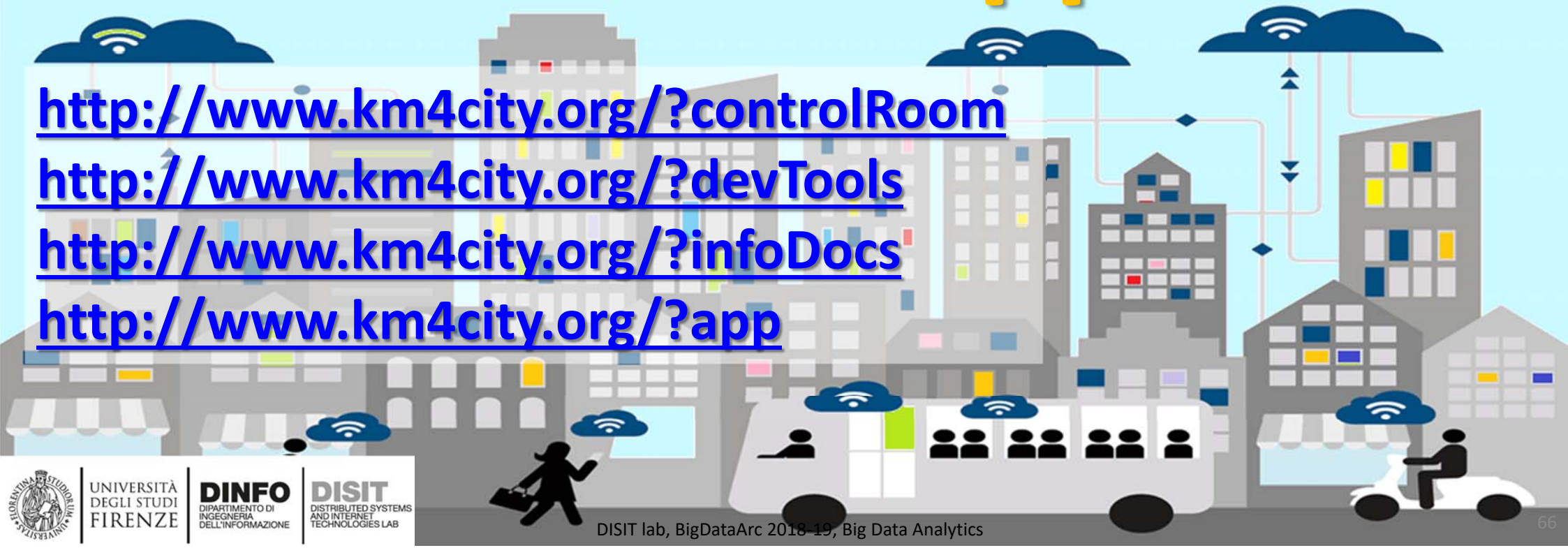




Twitter Vigilance

Case Study D

Monitoring via Mobile App



http://www.km4city.org/?controlRoom
http://www.km4city.org/?devTools
http://www.km4city.org/?infoDocs
http://www.km4city.org/?app

Overview

- **Analisi Dati raccolti via Social Media**
  - Predicting presences
- **Analisi Dati raccolti via Mobile App**
  - Tracce, matrici OD, heatmap
  - Regency e frequency
  - Impatto in uscita da Firenze
- **Analisi Dati raccolti dai Flussi di traffico**
  - Ricostruzione del traffico in punti non misurati
- **Analisi Dati raccolti dai Parcheggi**
  - Predizione dei posti liberi
- **Analisi Dati raccolti via Firenze WiFi**
  - Tracce, matrici OD, heatmap
  - Predicting presences
- **Analisi Dati Dati raccolti via Cellular**
  - Valutazione comparativa TIM-VODA
  - Valutazione comparativa FirenzeWiFi-Tim-Voda

Mobile Computing

Sii-Mobility

- **Smart City Problems:**
  - Reaching the users
  - Understanding the user preferences and behavior
  - Understating how they move, where they go, etc..
- **Solutions:**
  - Monitoring the activities on the mobile device
  - Monitoring the activities of user in the environment
- **Technologies for Solutions:**
  - Assessing the usage of Smart city and services
  - Integrated Indoor/outdoor navigation
    - Routing, multimodal routing
  - Content distribution: e-learning
  - User networking and collaboration
  - OS: iOS, Android, Windows Phone, etc.
  - Tech: IOT, iBeacoms, NFC, QR, ….

Available on the App Store

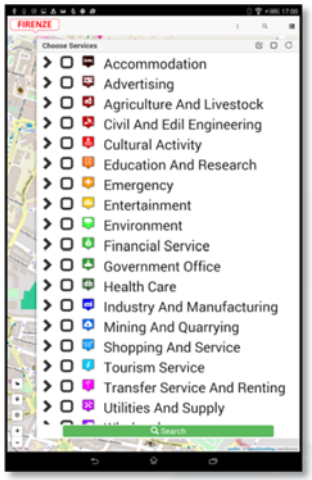
ANDROID APP ON Google play

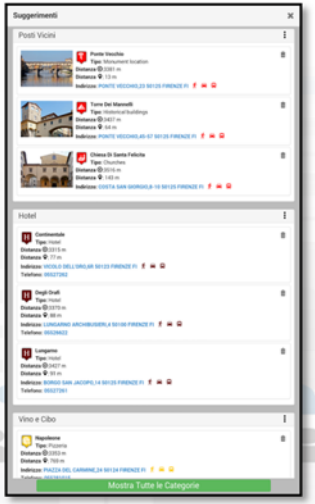
Download from Windows Phone Store

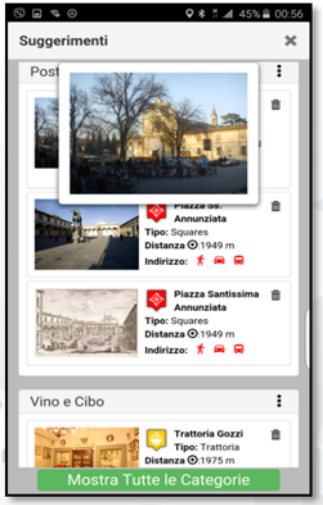
TOSCANA

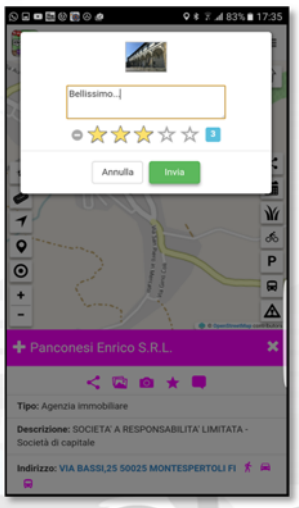
Km4City APP, features 1/3

- **5 languages:** IT, EN, SP, DE, FR
- **city users**: citizens, commuter, students, Tourists, etc..
- **Profiled menu for** POI, different for different City users
- **Personalized main menu**
- **Search Textual**
- **Search for POI**, POI kind, etc..
  - Close to you, close to a point
- **Direct searches**
  - Events, green areas, public transport,
  - Cycling paths, Parking (NEW: triage, fuel station)
  - Etc.
- **POI sharing and contributing**
  - Preferred, Social icon connection
  - Ranking, Comments, images









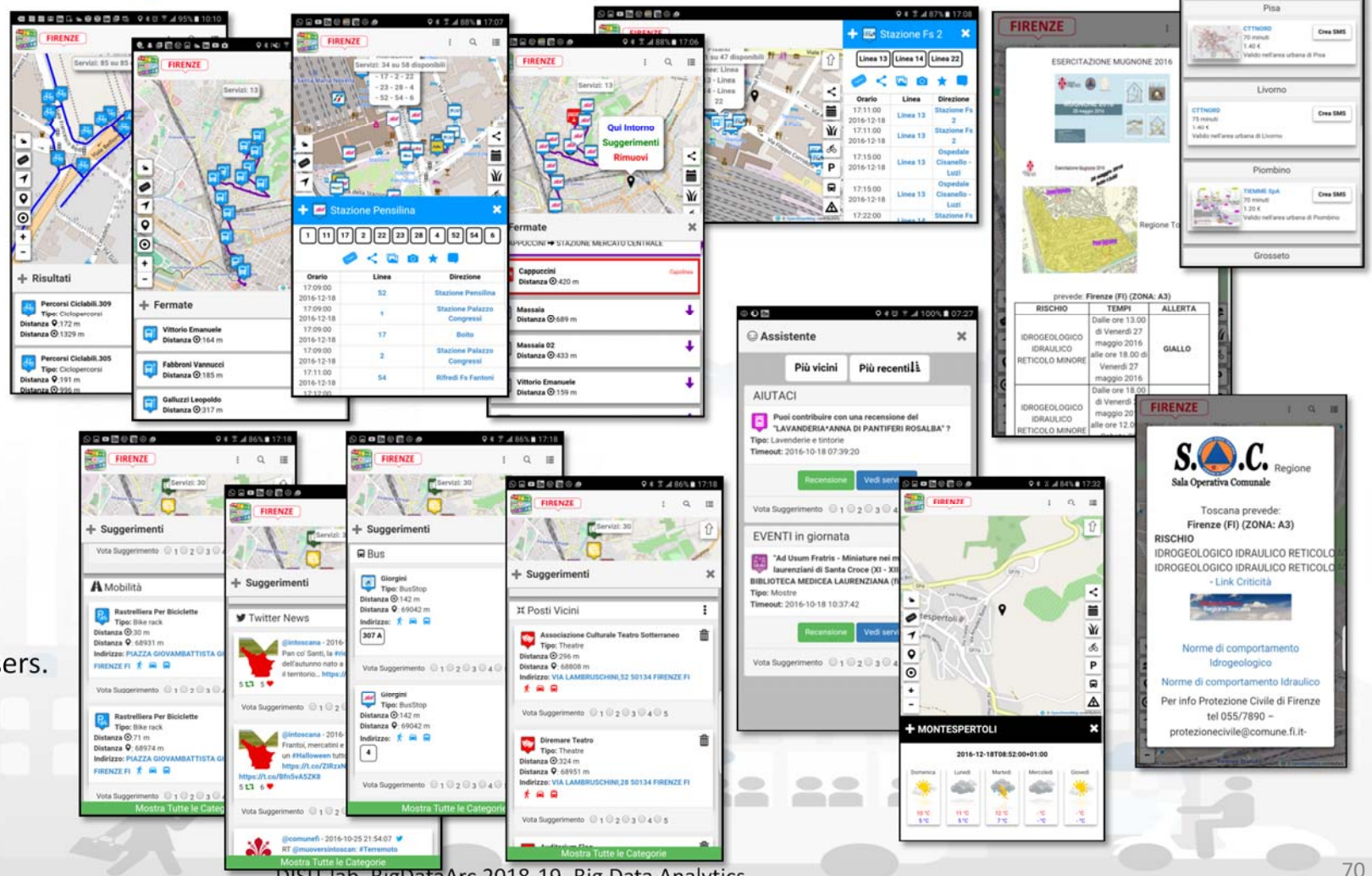
Available on the App Store

ANDROID APP ON Google play

Download from Windows Phone Store

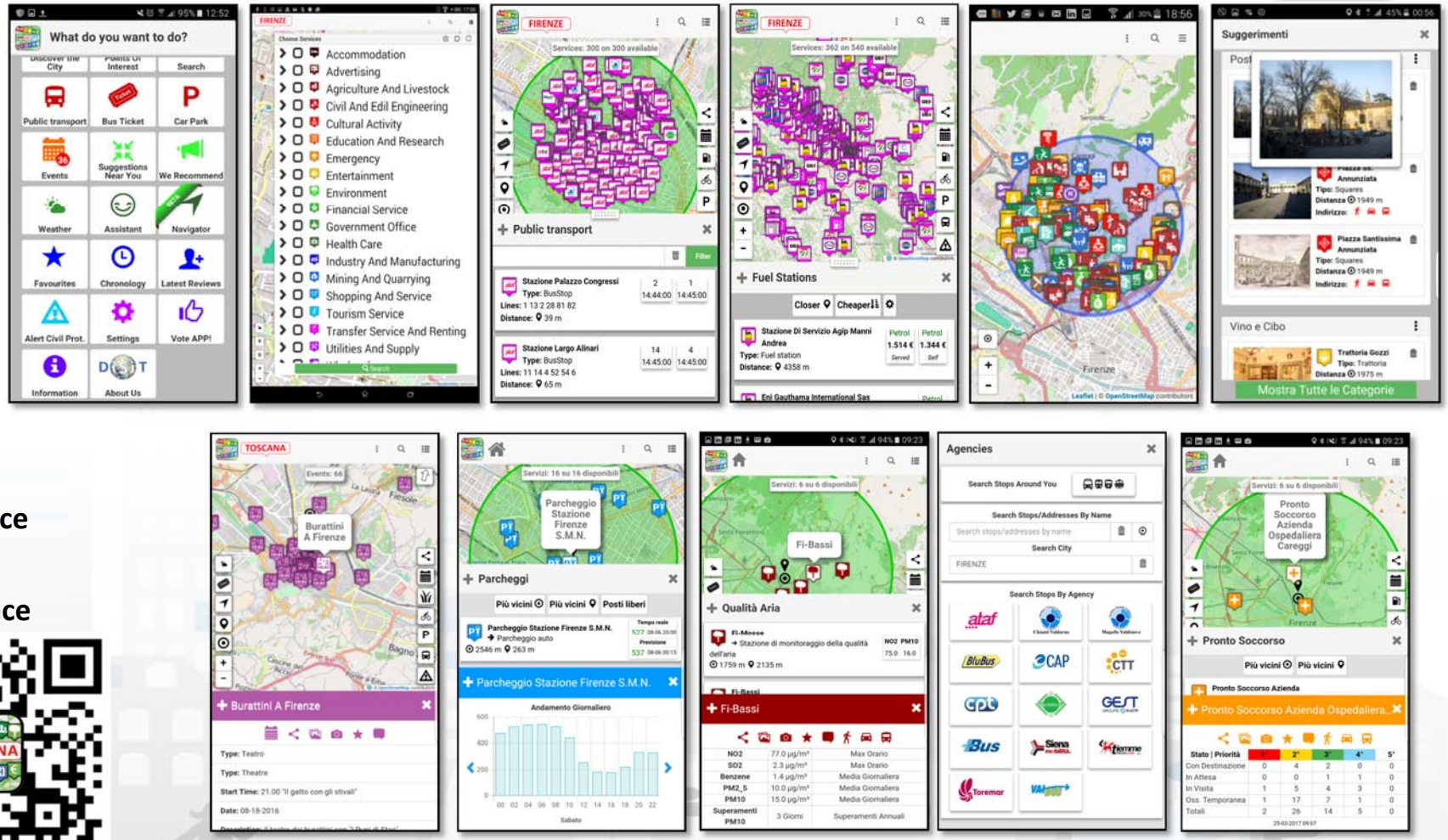
Km4City APP, features 2/3

- **Mobility**
  - Lines, bus stops, schedule
  - Parking status
  - Tickets for Busses
  - Cycling paths
  - Fuel stations
- **Personal Assistant**
  - Info and help
  - Engagement
  - Civil protection, alerts
  - Hospital triage status
- **Suggestions:**
  - Personalized and adaptive: banned & profiles for each users.
  - POI, Twitter hints, Events,
  - Weather forecast
  - ……



Km4City APP

- Smart Parking, in Tuscany
- Smart First Aid in Tuscany
- Smart Public Transportation in Tuscany
- Smart Fuel pricing in Tuscany
- Bike Sharing in Pisa
- Weather condition in Tuscany
- Pollution and Pollination in Tuscany
- Traffic Sensors in Tuscany
- Smart Routing in Tuscany
- Smart Transportation in Florence
  - Events, traffic, …
- Entertainment Events in Florence



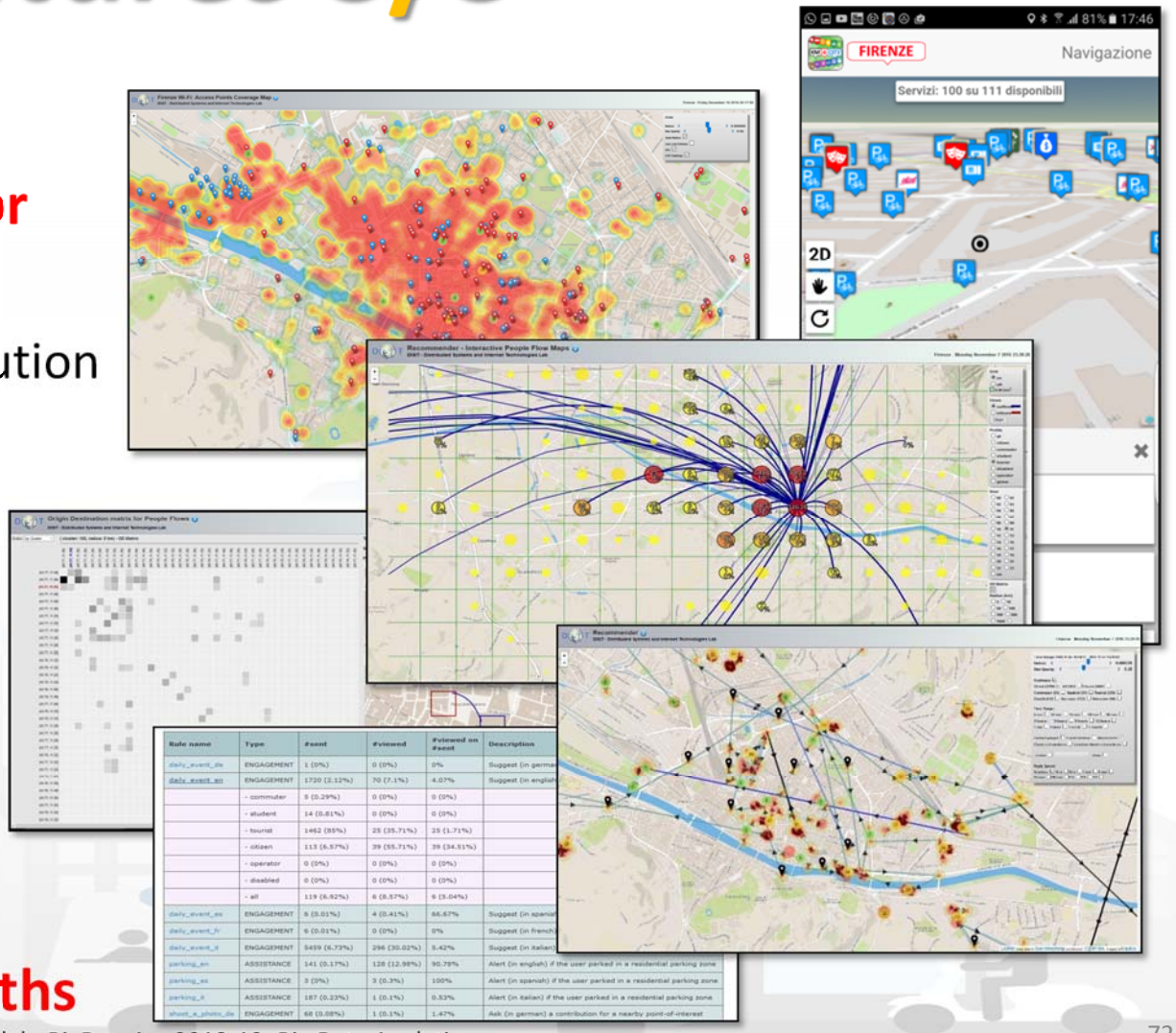
Available on the App Store

ANDROID APP ON Google play

Download from Windows Phone Store

Km4City APP, features 3/3

- **Navigation 3D (BETA)**
- **App as a tool for city user behavior analysis**
  - Measuring Wifi status: power distribution of Free Wi-Fi AP
  - Detection and measure of Beacon
  - Computing User Behavior
    - Fluxes of people via APP, GPS:
    - OD matrix
    - Fluxes out of Tuscany and more
- **Producing Engagements**
- **Producing Multimodal Routing paths**

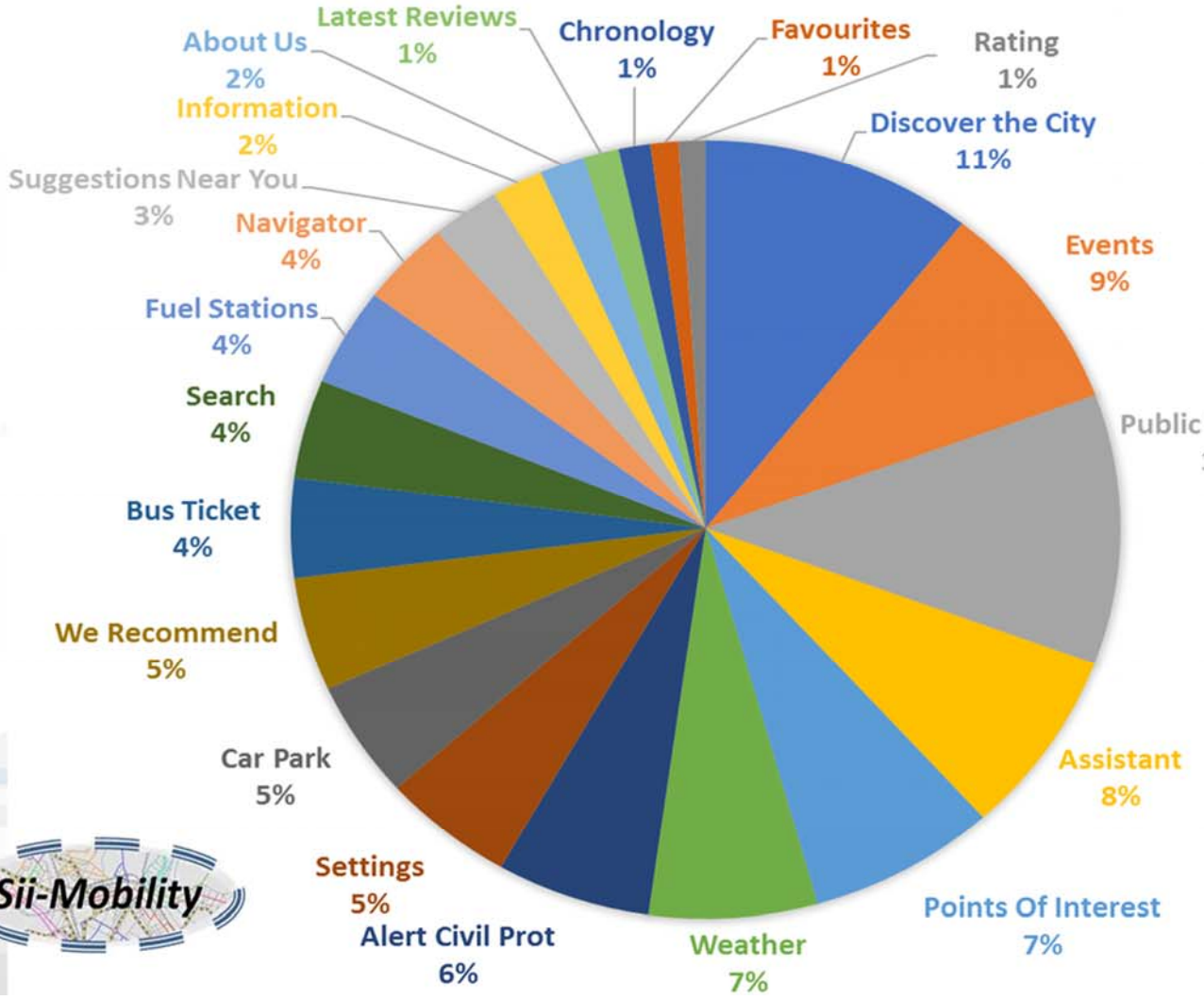


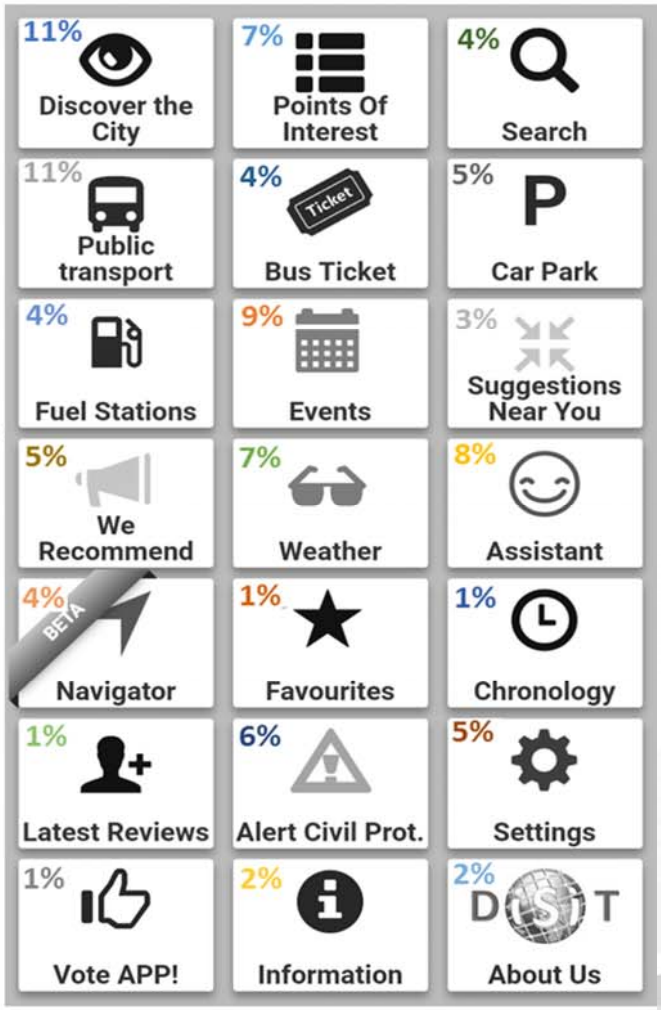
Users' Preferences

ARC
INDEX

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

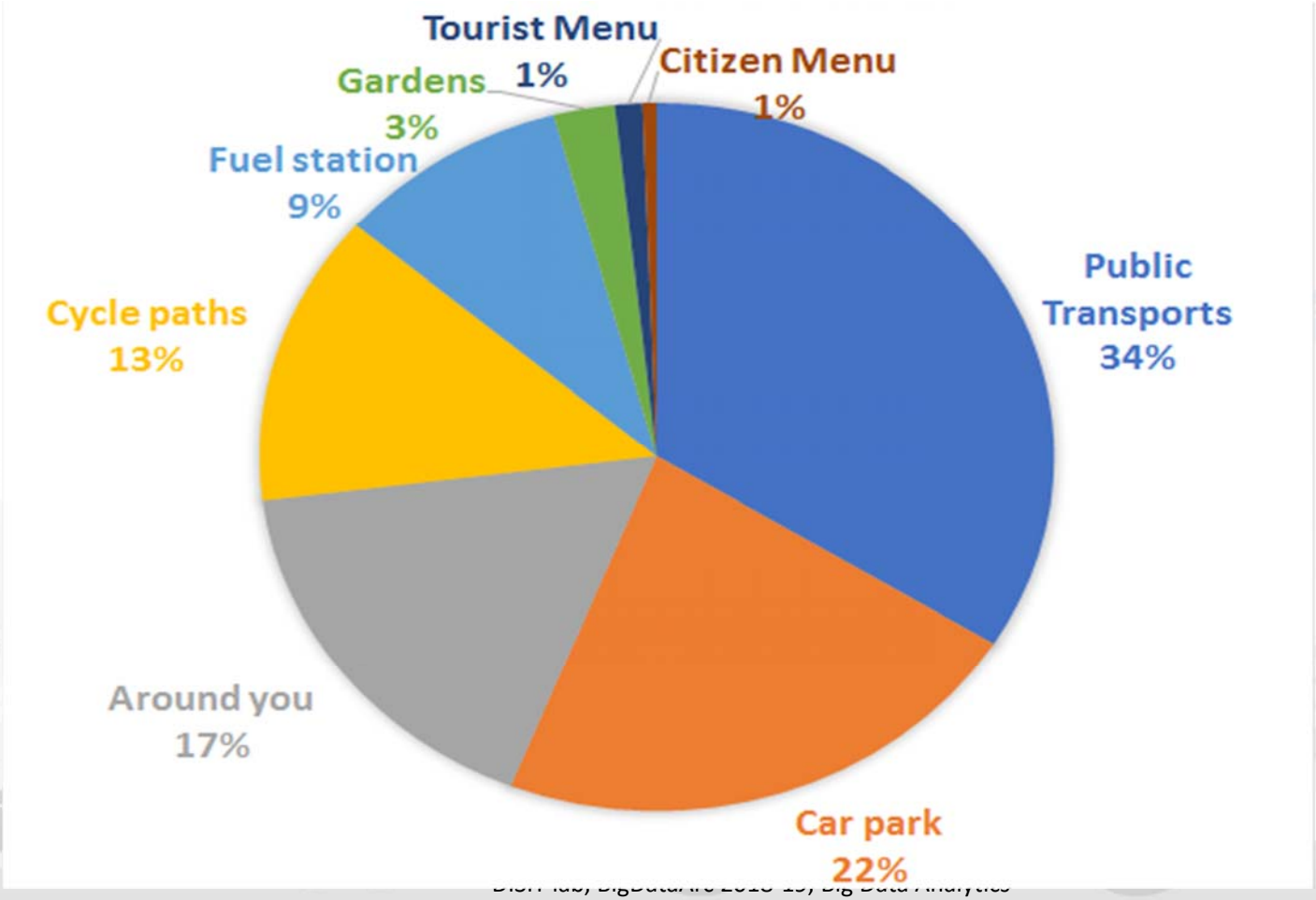
Users' preferences





Sii-Mobility

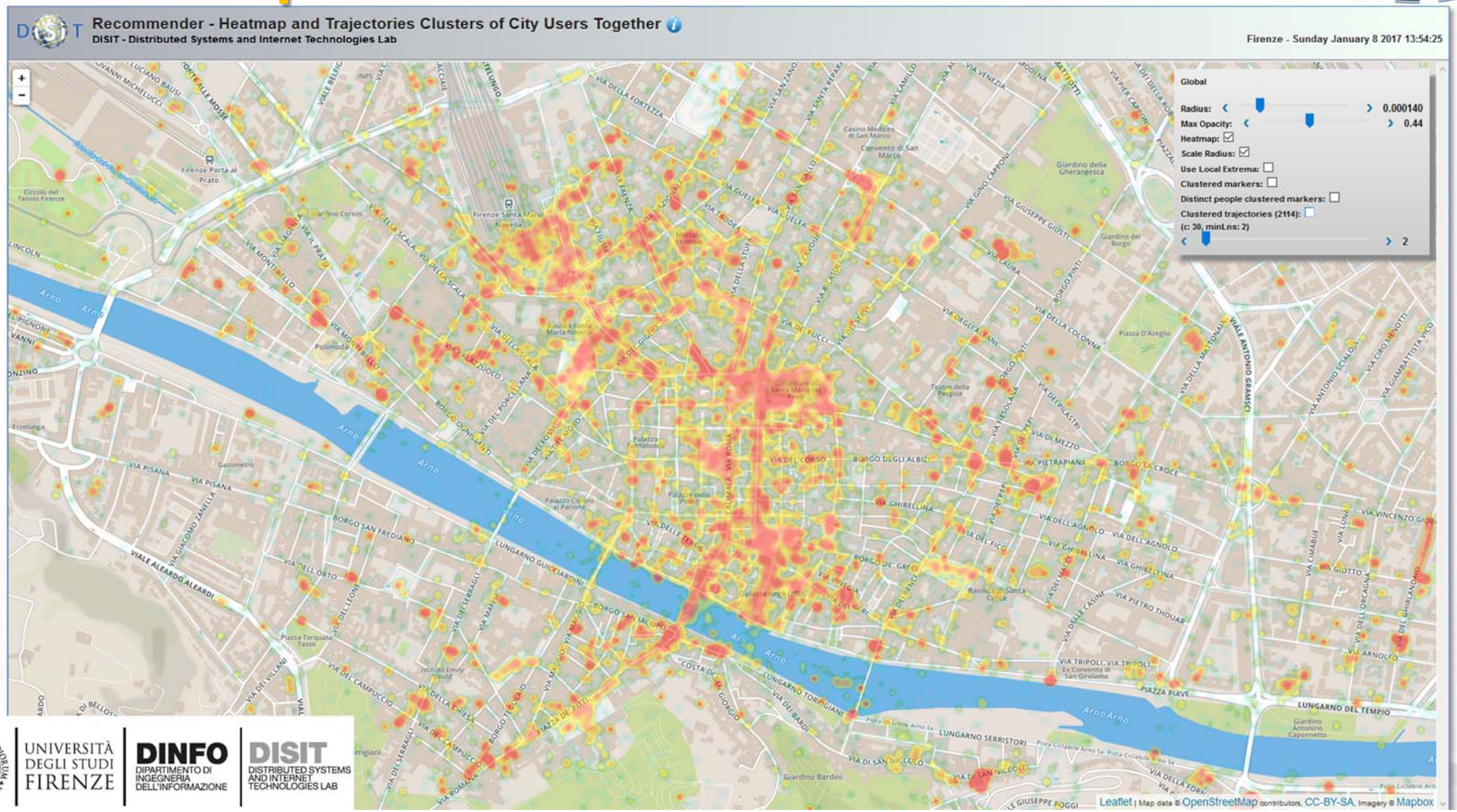
Preferred Users' Cathegories



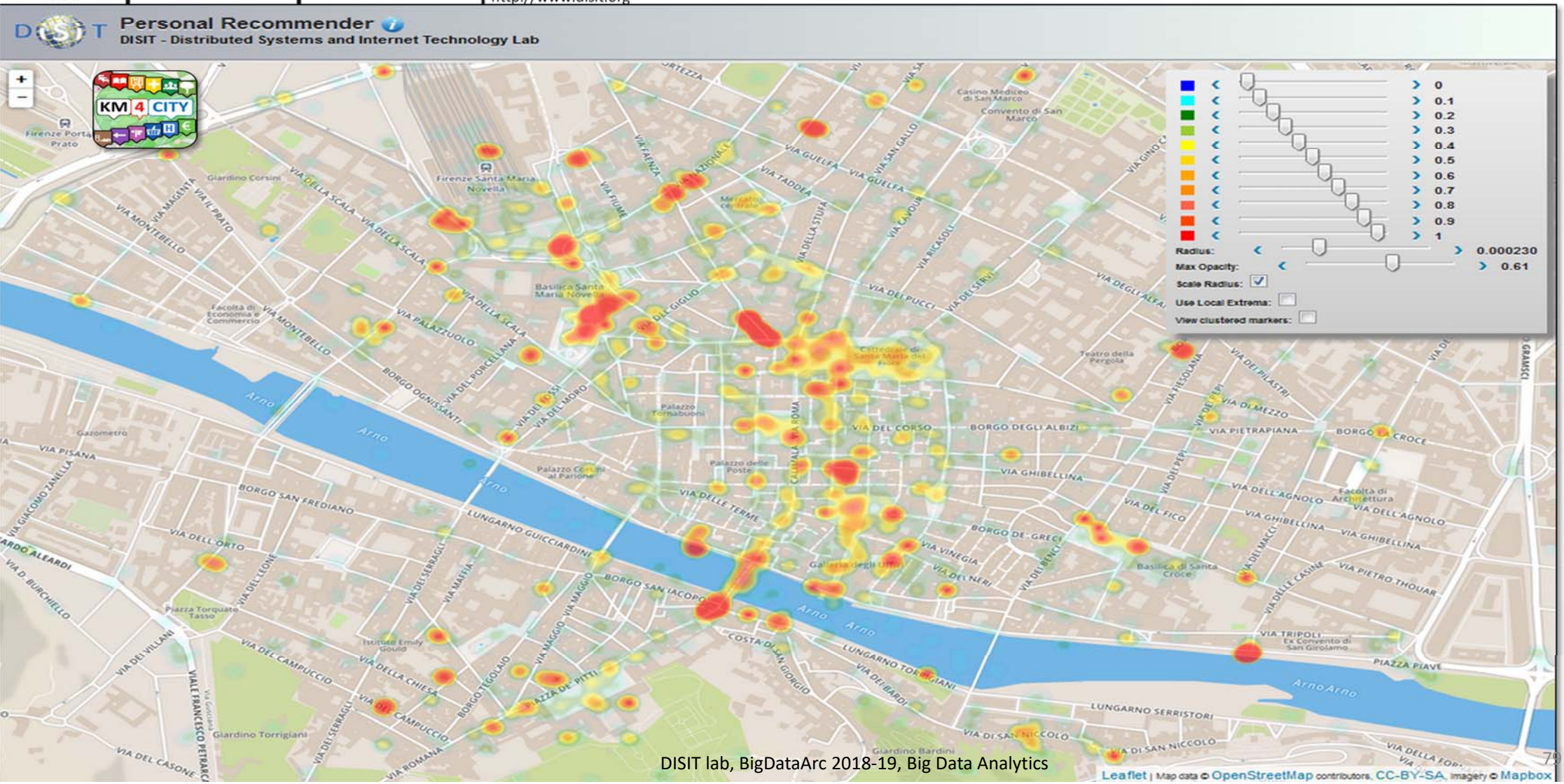
Users' Preferences on locations

ARC
INDEX

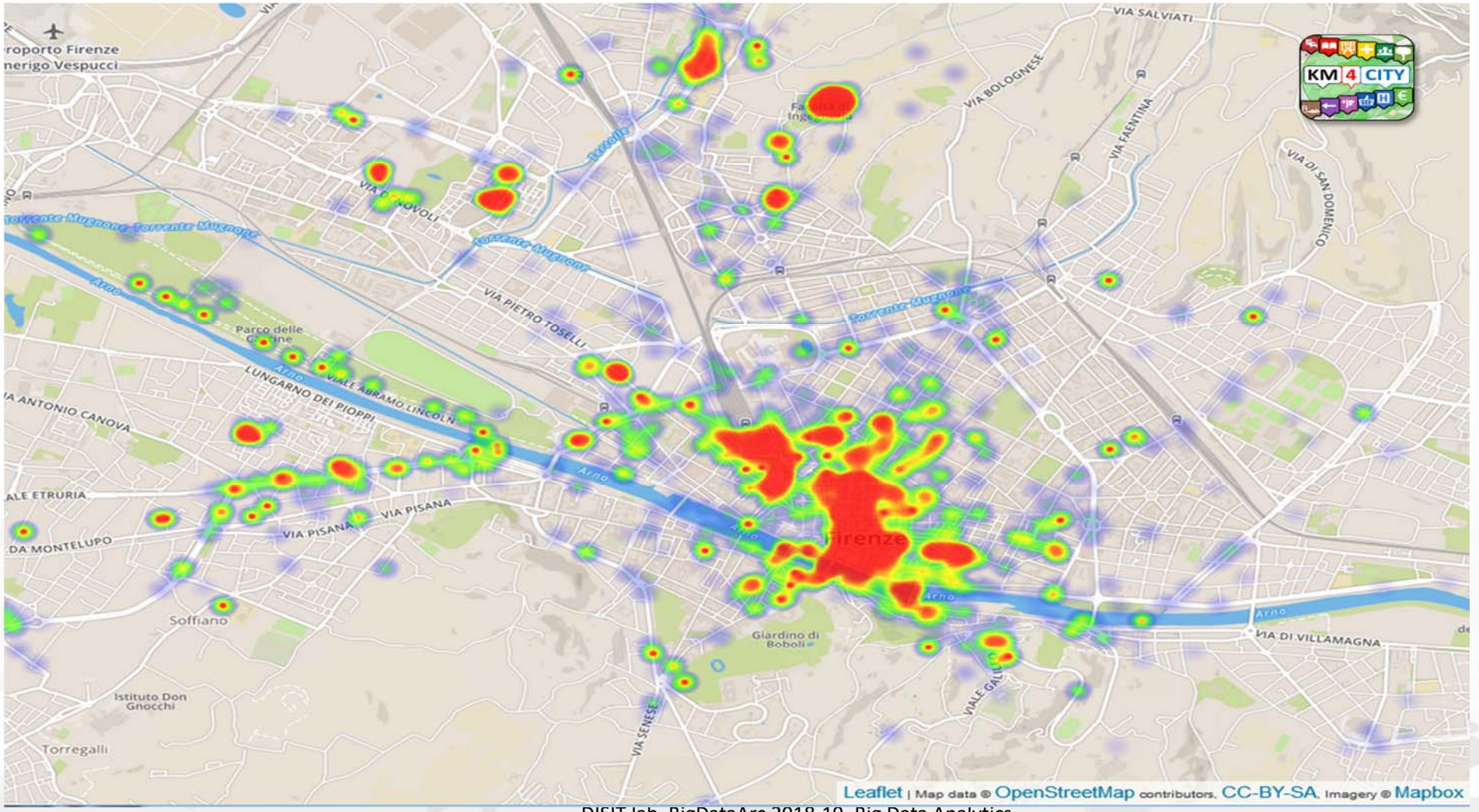
Heat Map from Mobile: users as sensors



Tourists in Florence



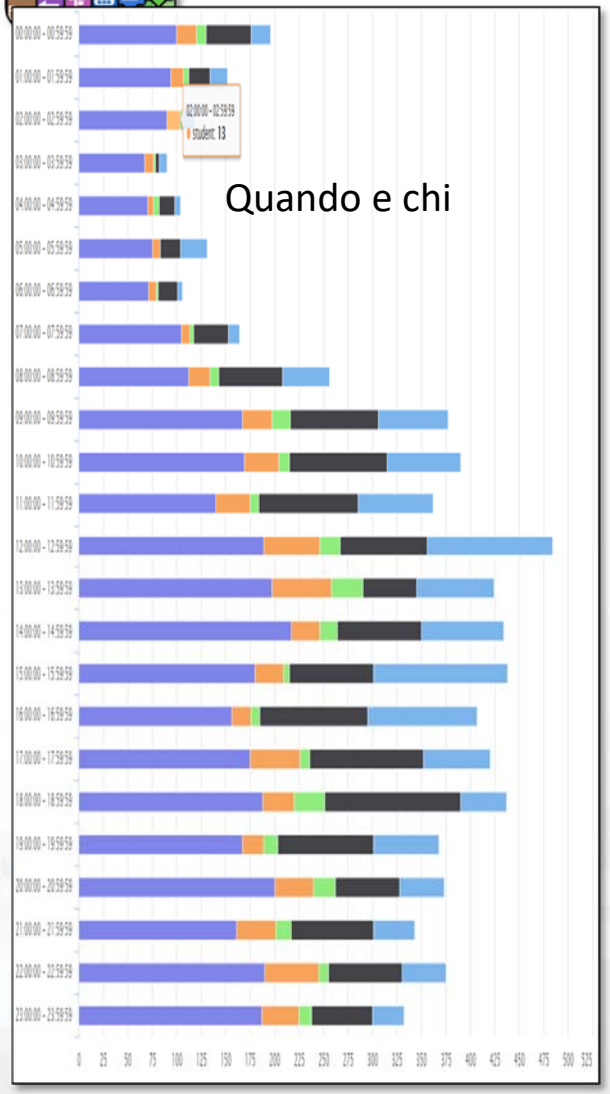
Hot WiFi in Florence

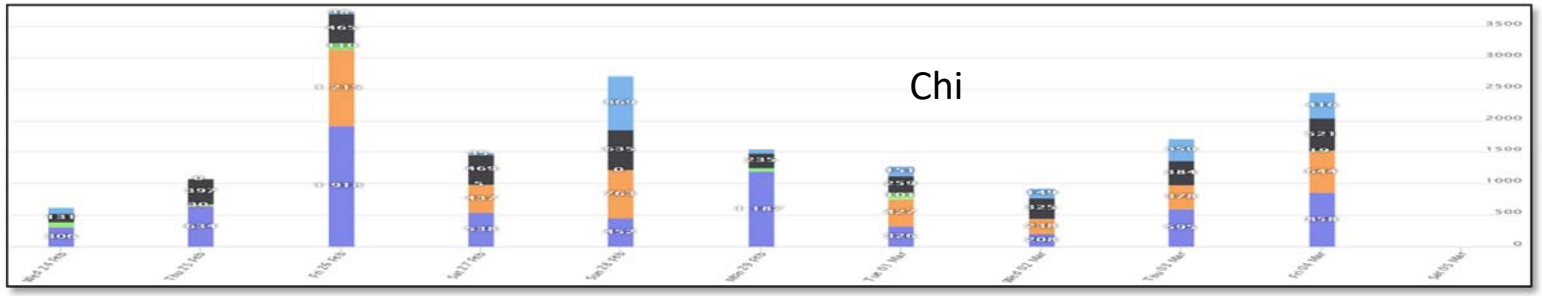


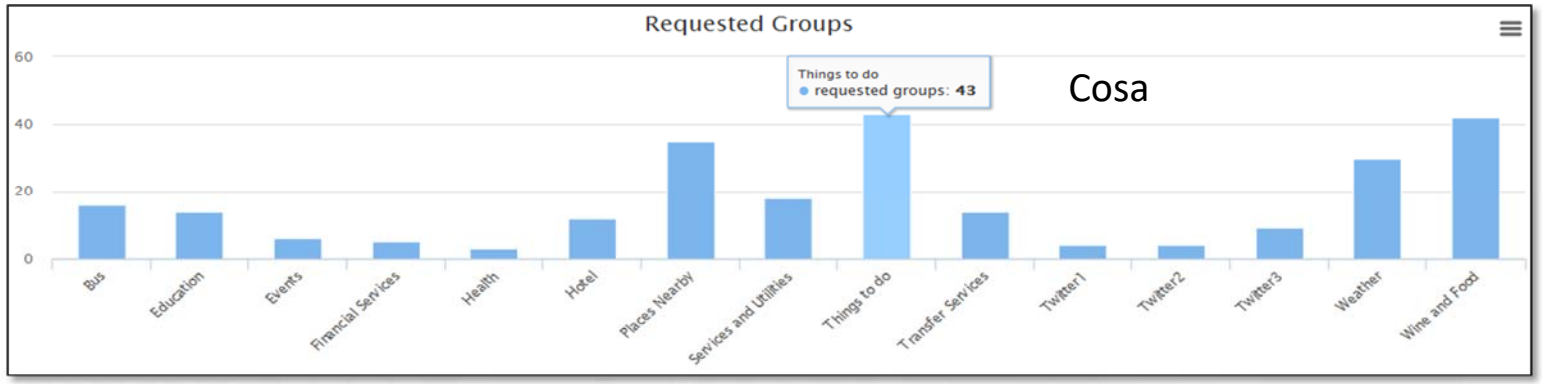
Users' preference and behavior

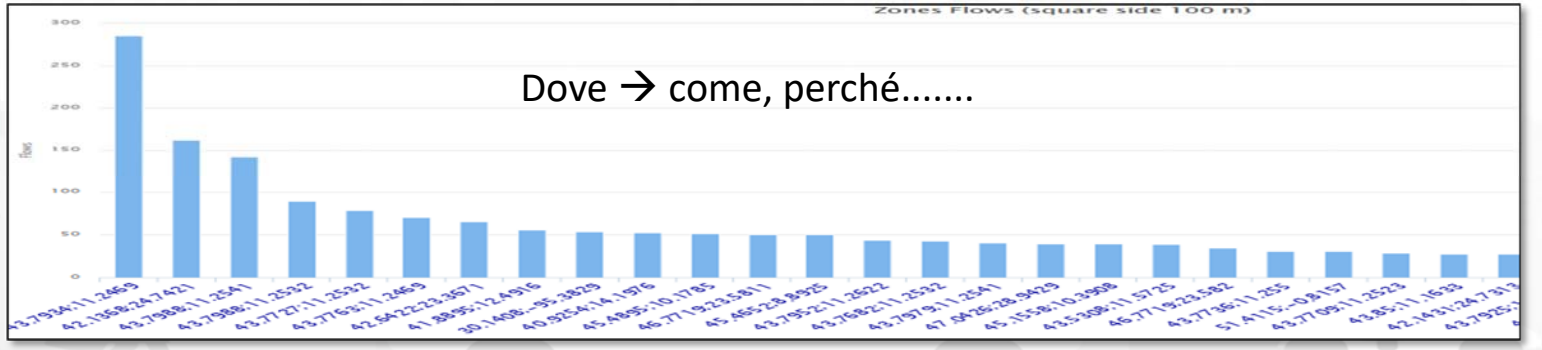
ARC
INDEX

Recommender





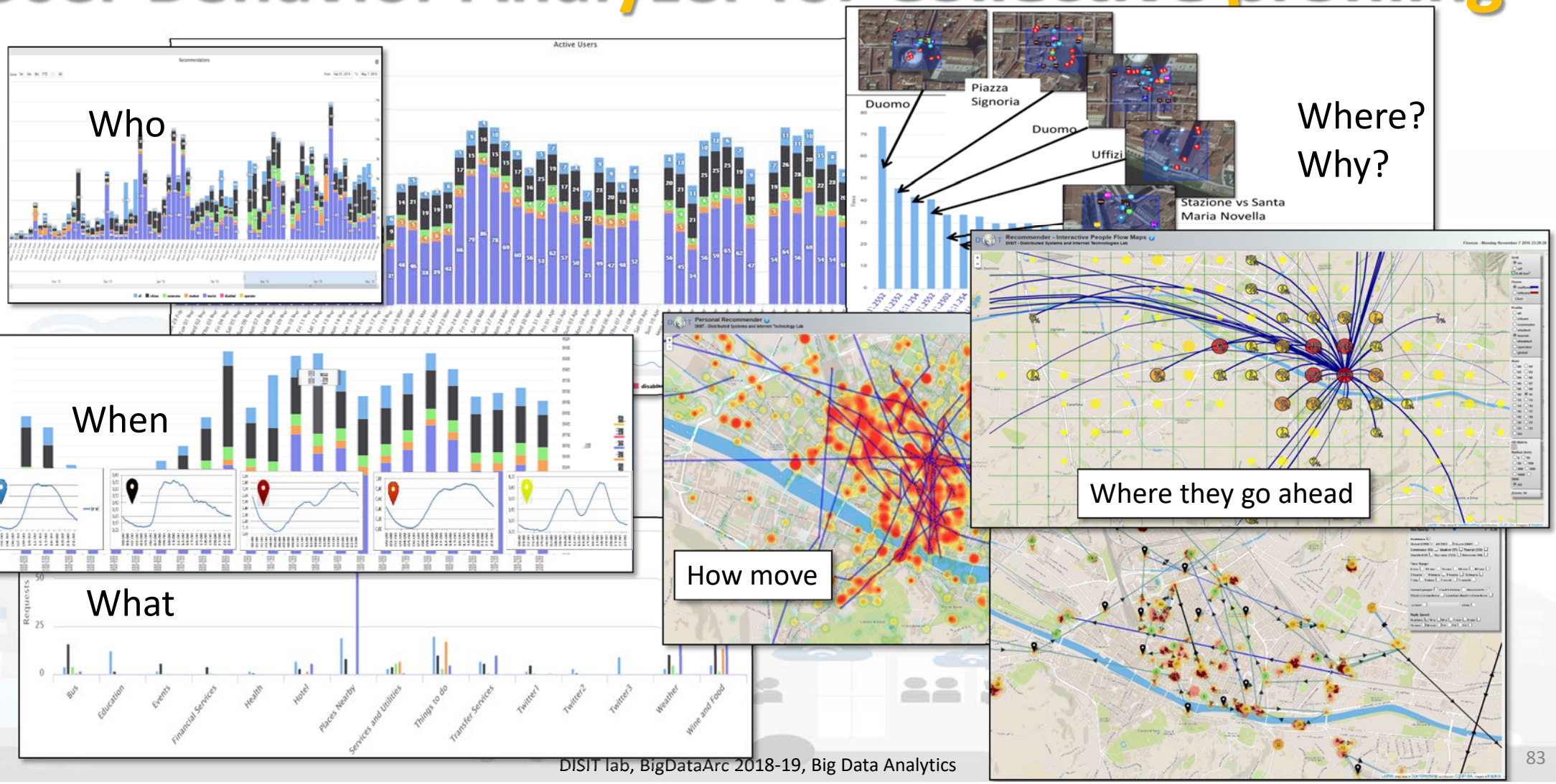




Case Study E

Sii-Mobility

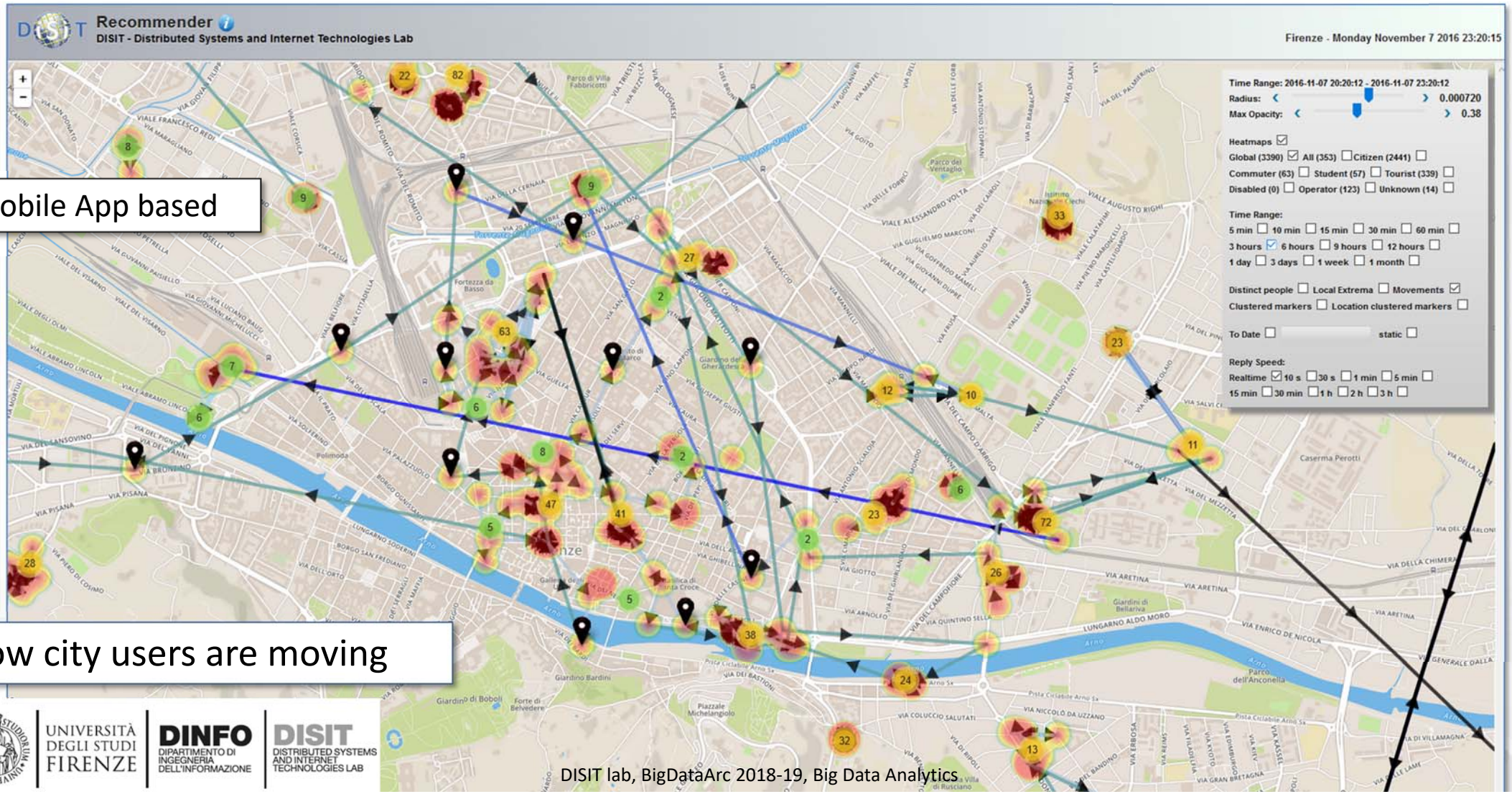
User Behavior Analyzer for Collective profiling



Anonymous User Behavior Analysis

Case Study E

Sii-Mobility

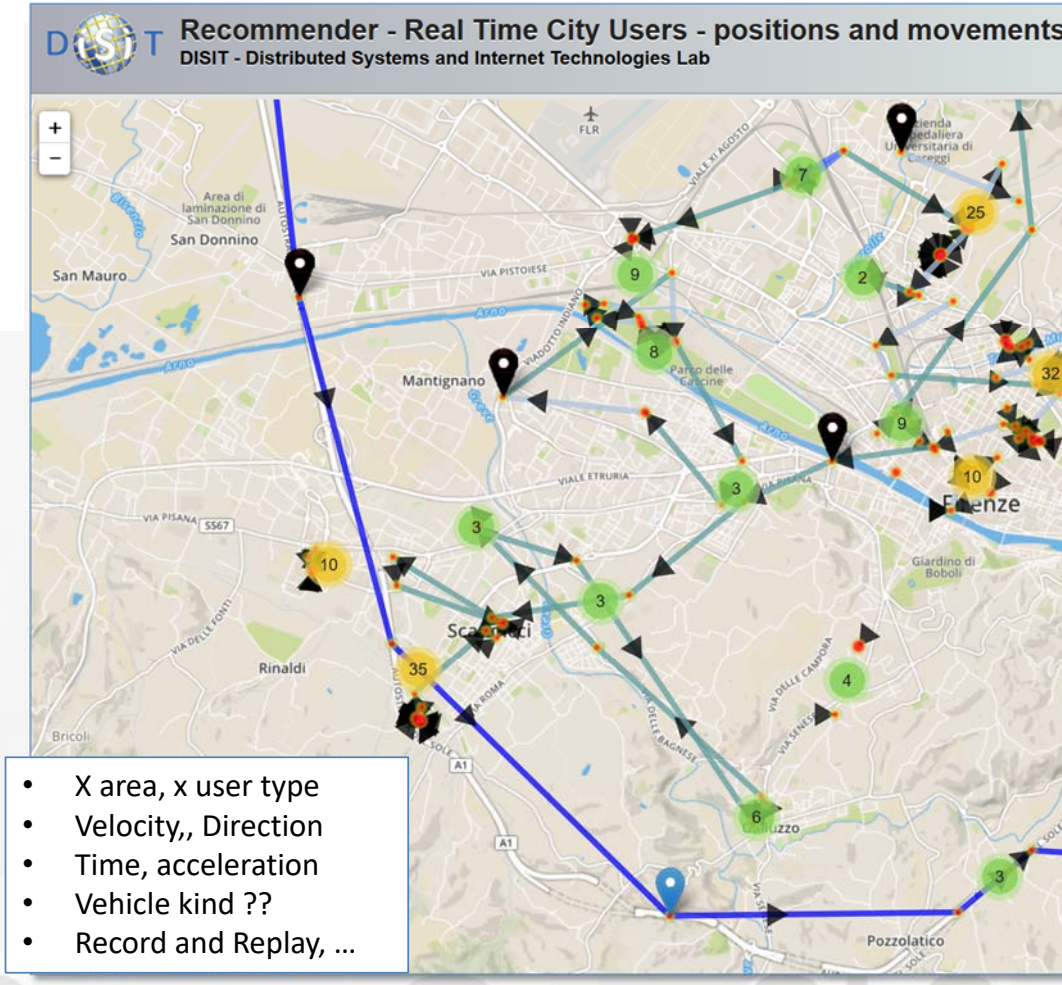


Mobile App based

How city users are moving

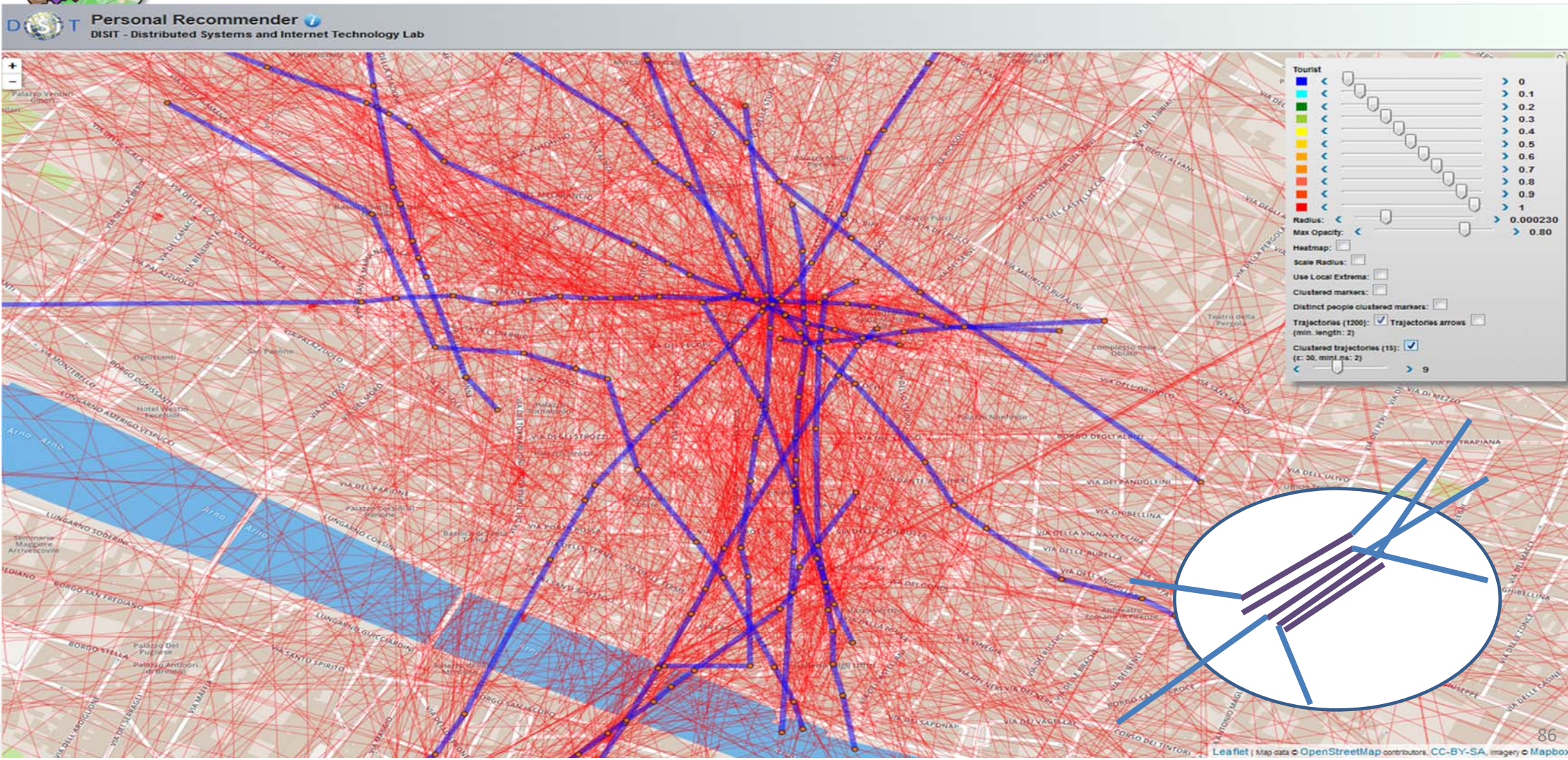
Problems of Trajectories from Apps

- **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.

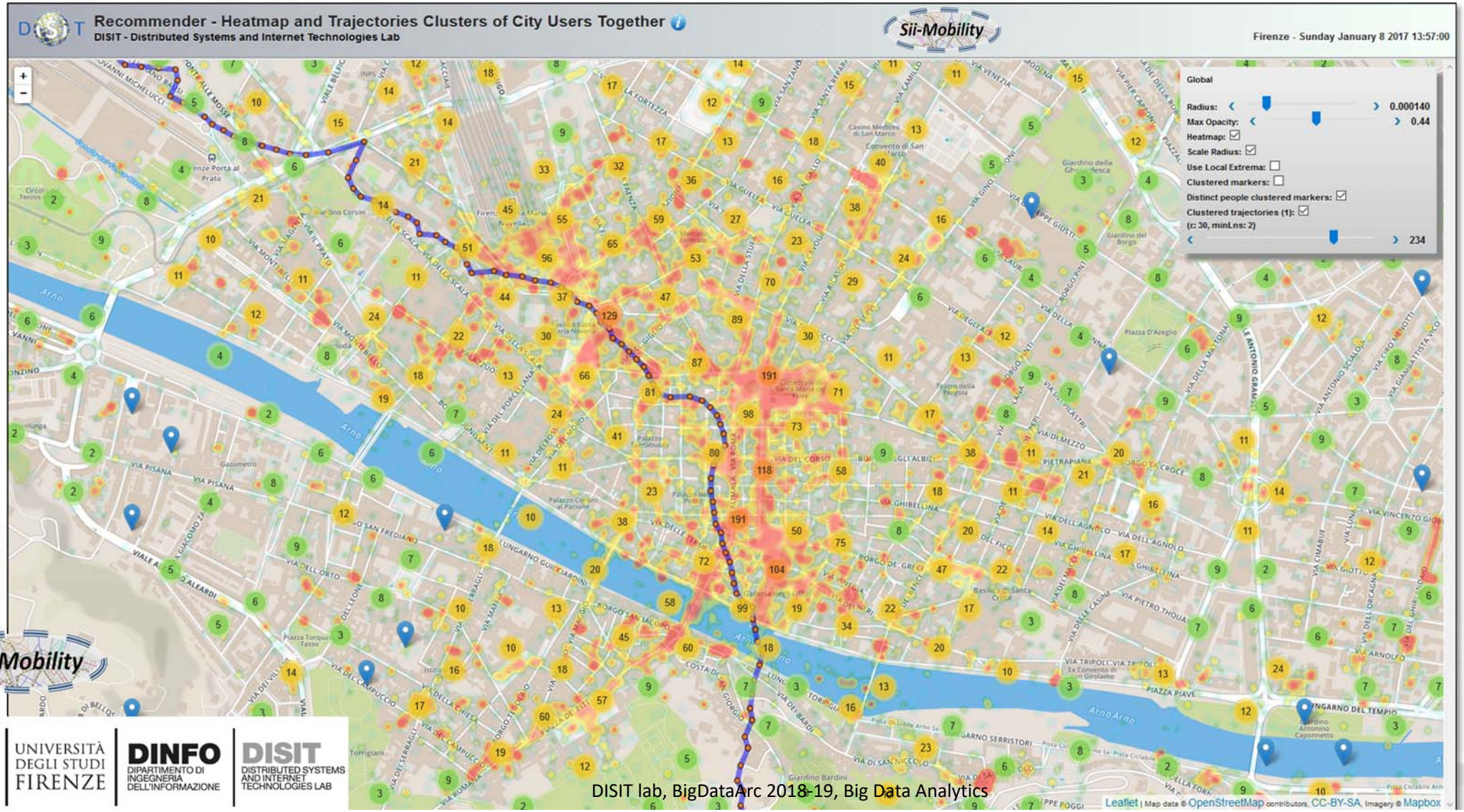


- X area, x user type
- Velocity,, Direction
- Time, acceleration
- Vehicle kind ??
- Record and Replay, …

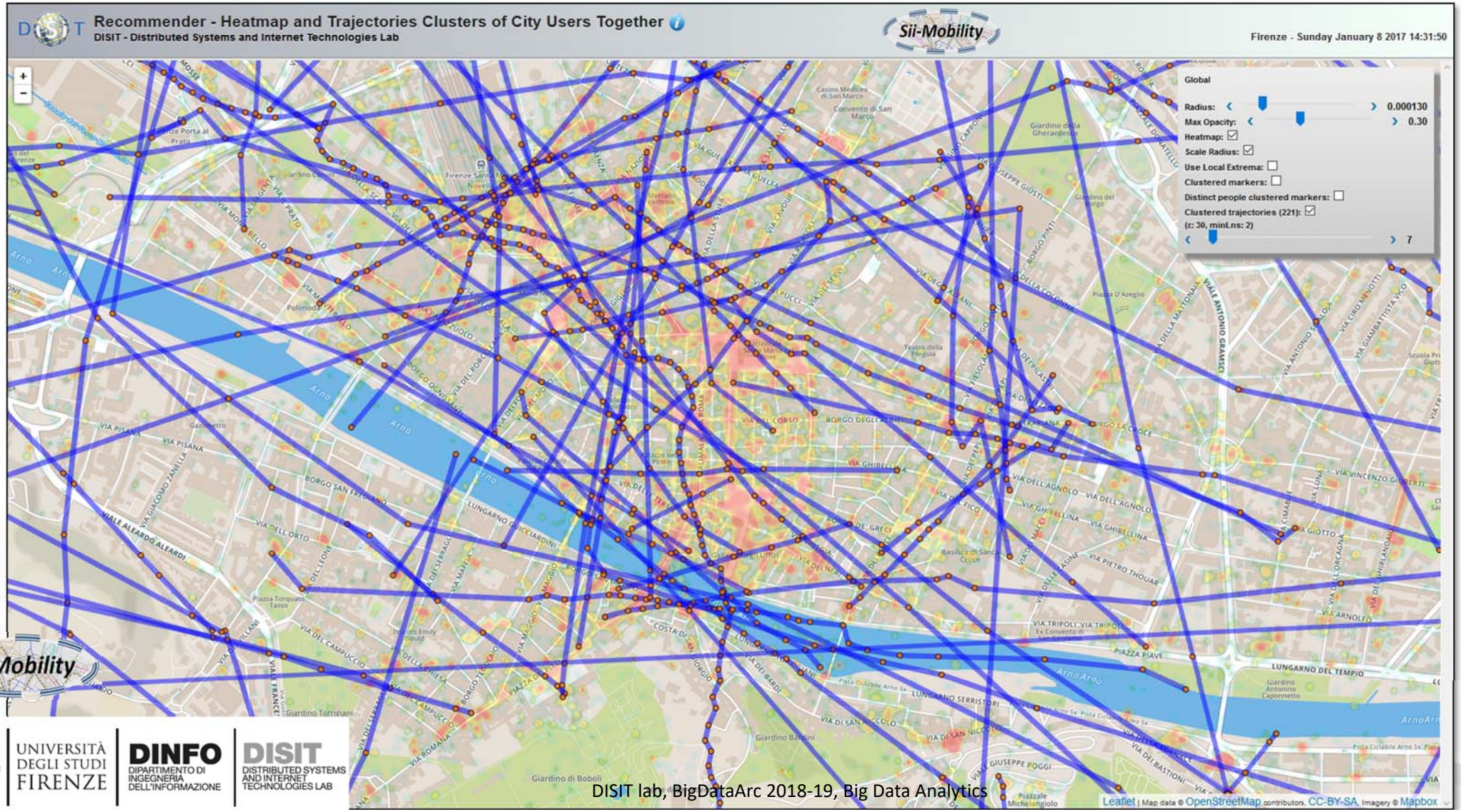
Cluster di Trajectories



Heat Map from Mobile: users as sensors



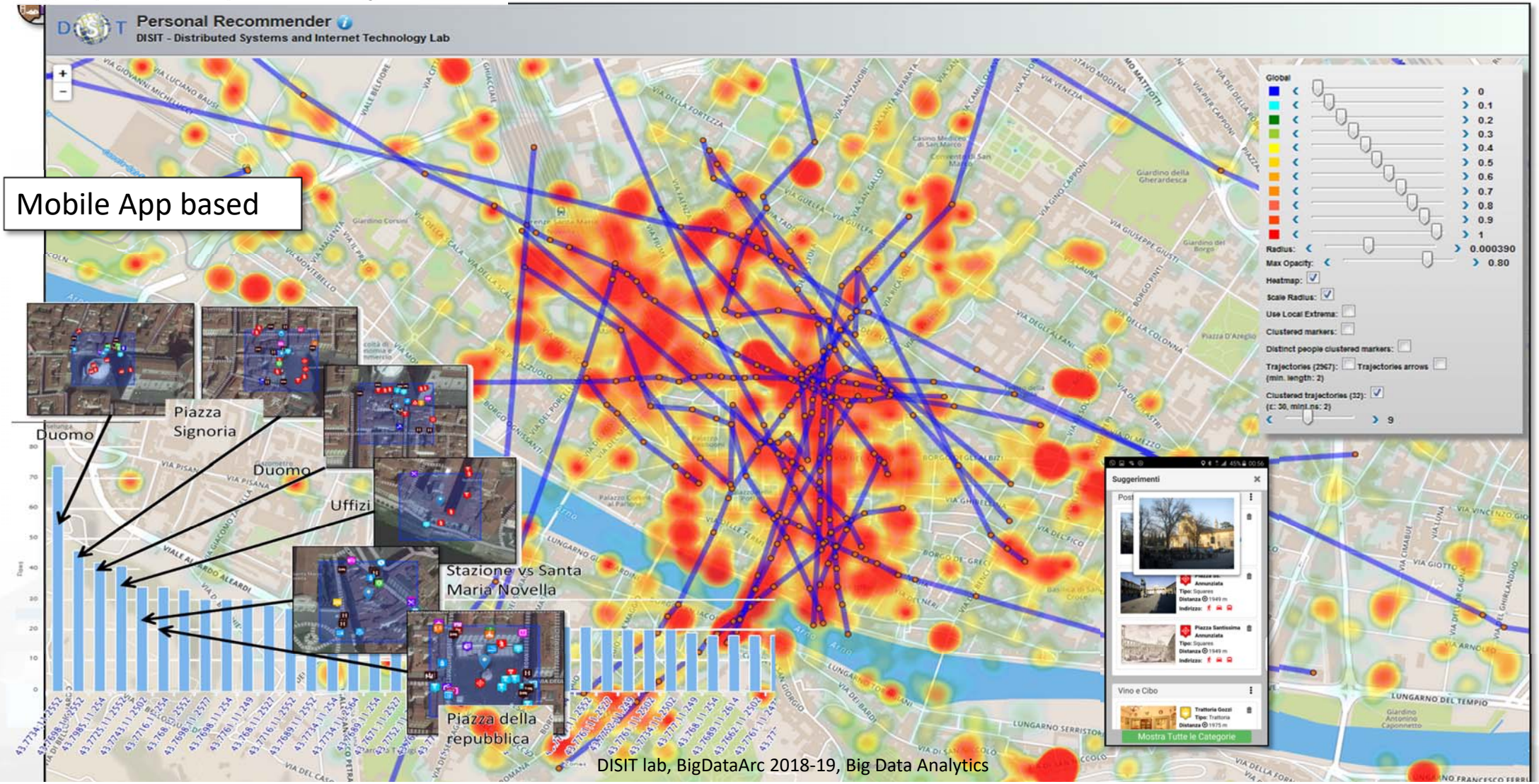
Heat Map from Mobile: users as sensors



Sii-Mobility

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

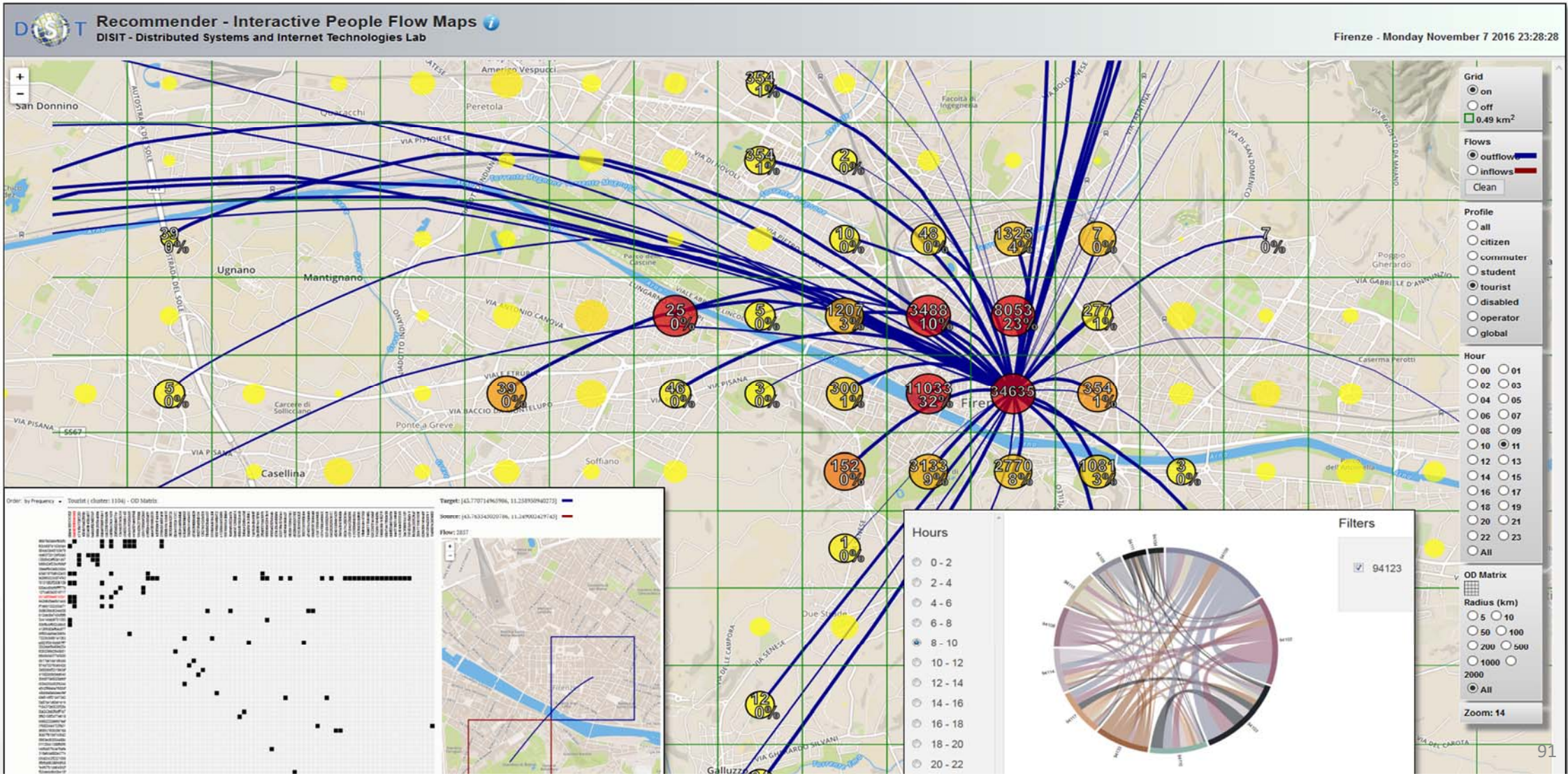
User Behavior Analyzer



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

OD Matrix scalabile

Sii-Mobility



Engaging City Users toward a Virtuous Behaviours (real time)

Strategies Implementation via Engagement

- **Produce value from data enabling** to
  - ***Stimulate virtuous behavior,***
  - ***influence engage City Users!***
  - Increase efficiency in energy consumption
  - Reduce pollution and traffic congestion
  - Improve quality of service, quality of life
- **Create an ecosystem for innovation** and put in action any smart city solutions and services.

User influencing, engaging, monitoring & Follow Up

Sii-Mobility

City & City Operators Strategy Editor

Rule name	Type	#sent	#viewed	#viewed on #sent	Description
daily_event_de	ENGAGEMENT	1 (0%)	0 (0%)	0%	Suggest (in german) an event currently on i
daily_event_en	ENGAGEMENT	1720 (2.12%)	70 (7.1%)	4.07%	Suggest (in english) an event currently on in
	- commuter	5 (0.29%)	0 (0%)	0 (0%)	
	- student	14 (0.81%)	0 (0%)	0 (0%)	
	- tourist	1462 (85%)	25 (35.71%)	25 (1.71%)	
	- citizen	113 (6.57%)			
	- operator	0 (0%)			
	- disabled	0 (0%)			
	- all	119 (6.92%)			
daily_event_es	ENGAGEMENT	6 (0.01%)			
daily_event_fr	ENGAGEMENT	6 (0.01%)			
daily_event_it	ENGAGEMENT	5459 (6.73%)			
parking_en	ASSISTANCE	141 (0.17%)			
parking_es	ASSISTANCE	3 (0%)			
parking_it	ASSISTANCE	187 (0.23%)			
shoot_a_photo_de	ENGAGEMENT	68 (0.08%)			

Any Mobile and Web App

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

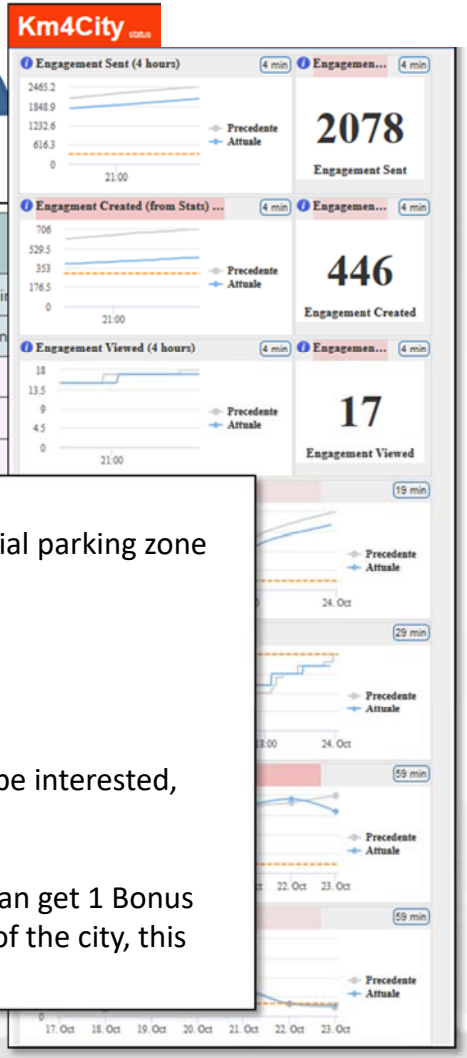
Provide a comment, a score, etc..

Stimulate / recommend

Events in the city, services your may be interested, etc..

Provide Bonus

Since you have parked here you we can get 1 Bonus
We suggest you to leave the car out of the city, this bonus can be used to by a bus ticket



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

http://www.disit.org

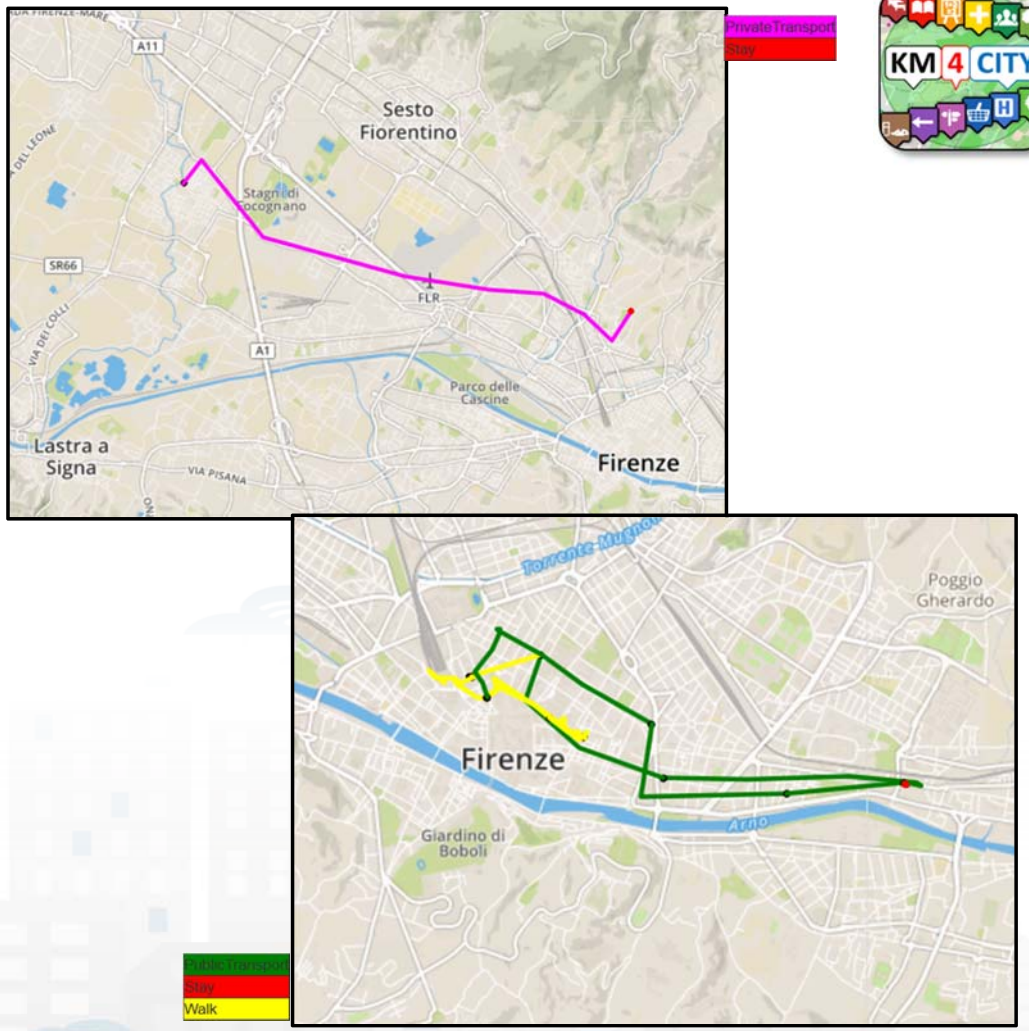
REWARDING'S RULES

- **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

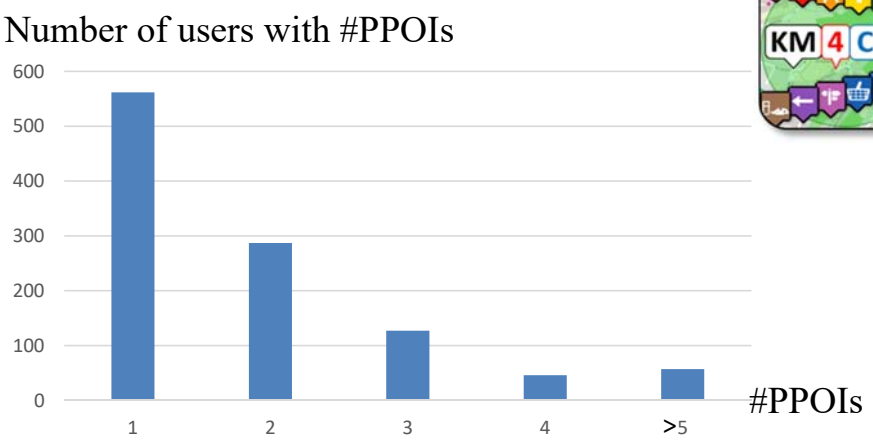
- 50 engagement's rules in 5 languages
  - (surveys, feedbacks, suggestions, assistances)
- From1° September 2016
  - Produced 322.900 engagements on 4270 users
- From 1° July 2017
  - 233 registered users with email (154 via social)

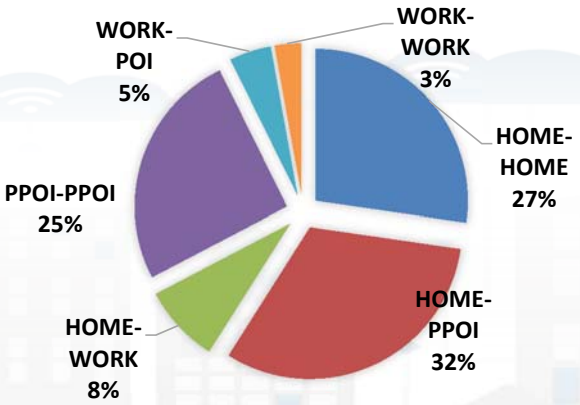
From: 2016-08-01 00:00:00
To: 2017-05-26 11:19:52
Refresh

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





Monitoring Traffic Flow

Sii-Mobility

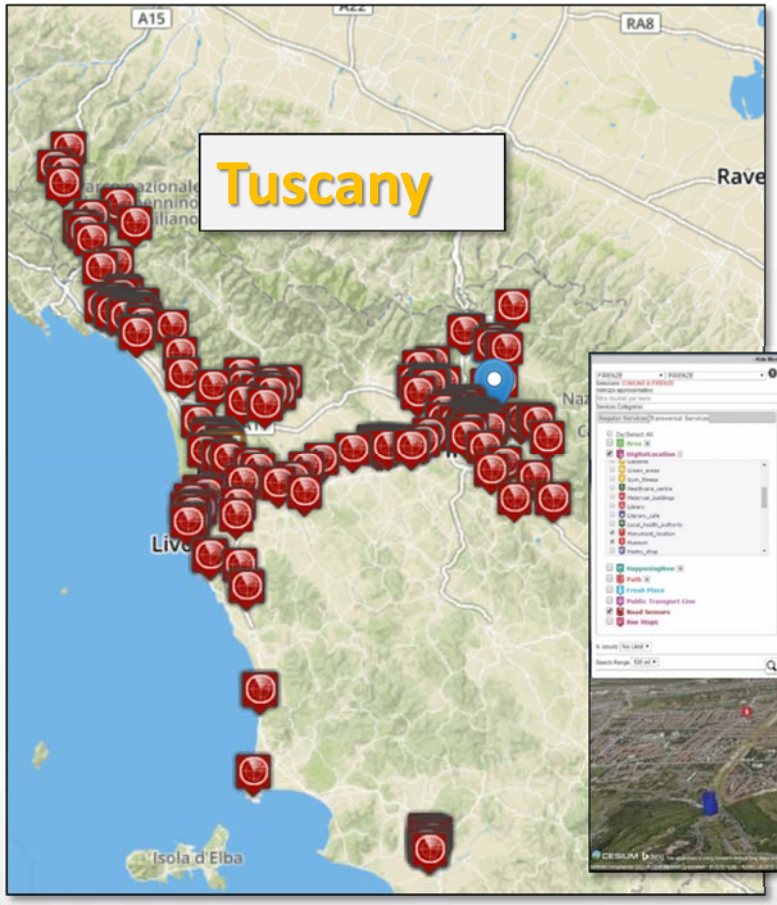
http://www.km4city.org/?controlRoom

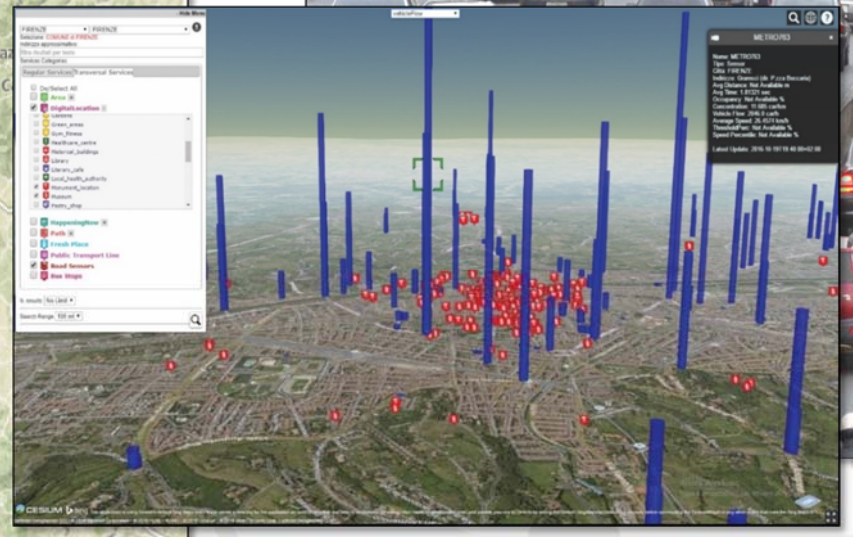
Overview

- **Analisi Dati raccolti via Social Media**
  - Predicting presences
- **Analisi Dati raccolti via Mobile App**
  - Tracce, matrici OD, heatmap
  - Regency e frequency
  - Impatto in uscita da Firenze
- **Analisi Dati raccolti dai Flussi di traffico**
  - Ricostruzione del traffico in punti non misurati
- **Analisi Dati raccolti dai Parcheggi**
  - Predizione dei posti liberi
- **Analisi Dati raccolti via Firenze WiFi**
  - Tracce, matrici OD, heatmap
  - Predicting presences
- **Analisi Dati Dati raccolti via Cellular**
  - Valutazione comparativa TIM-VODA
  - Valutazione comparativa FirenzeWiFi-Tim-Voda

Traffic Flow Tools

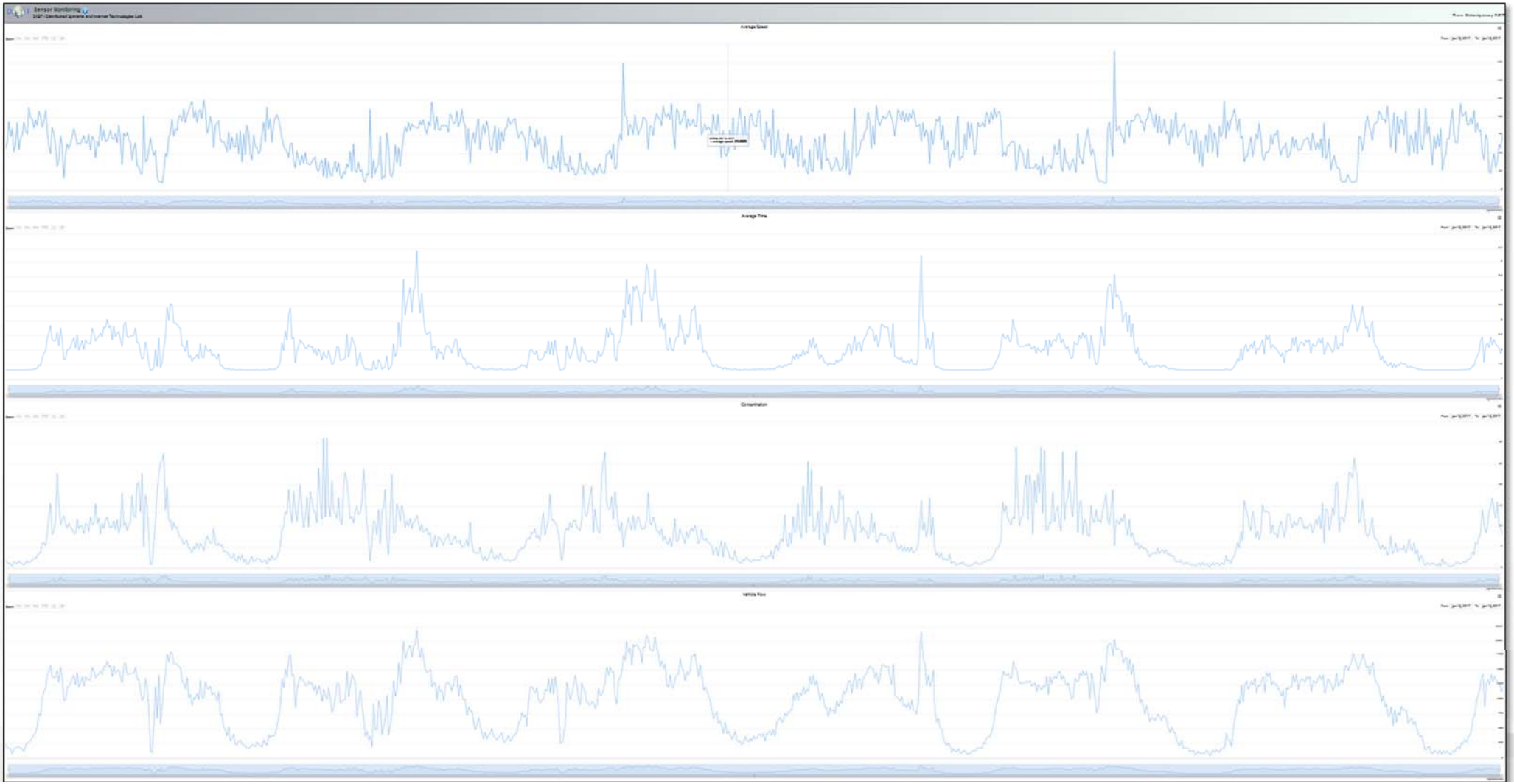
- Spire and Virtual Spires (cameras), Bluetooth, ..
- Specifically located: along, around, on gates, ..





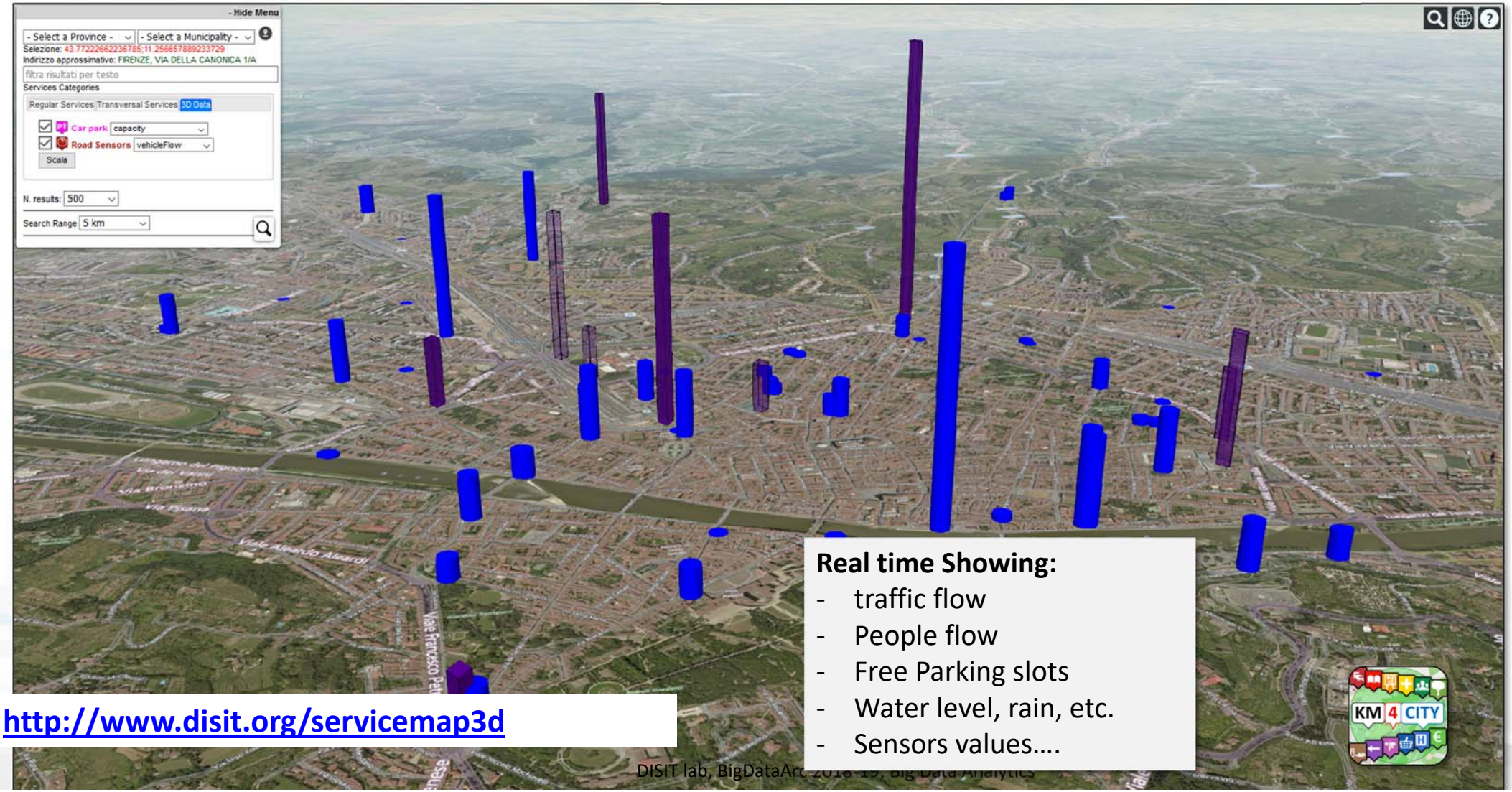
Traffic Sensors, global counting

Traffic Flow data



RealTime Values 3D

Sii-Mobility



Real time Showing:

- traffic flow
- People flow
- Free Parking slots
- Water level, rain, etc.
- Sensors values….

http://www.disit.org/servicemap3d

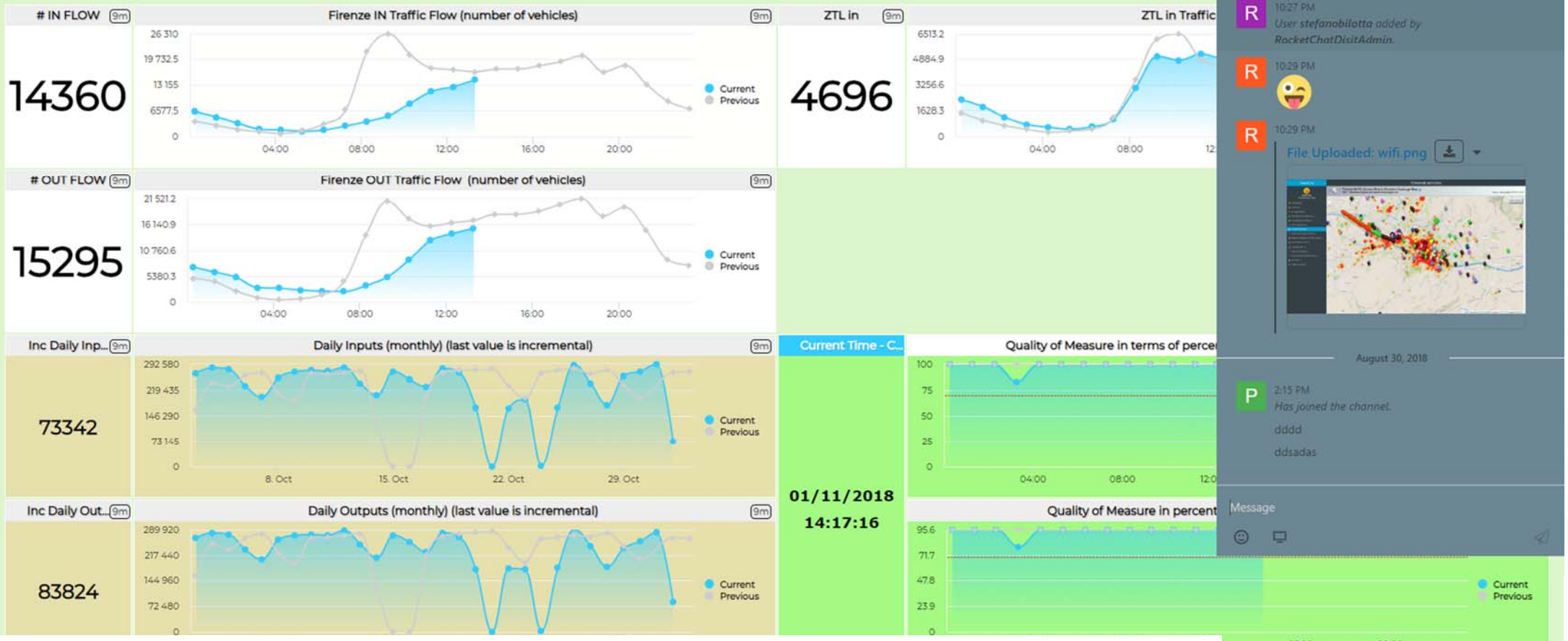
UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB

SNAP4CITY

KM 4 CITY

Total Traffic Count per hour for Firenze

Thu 1 Nov 14:17:16



https://main.snap4city.org/view/index.php?iddasboard=OTg5

Traffic Flow Reconstruction

ARC
INDEX

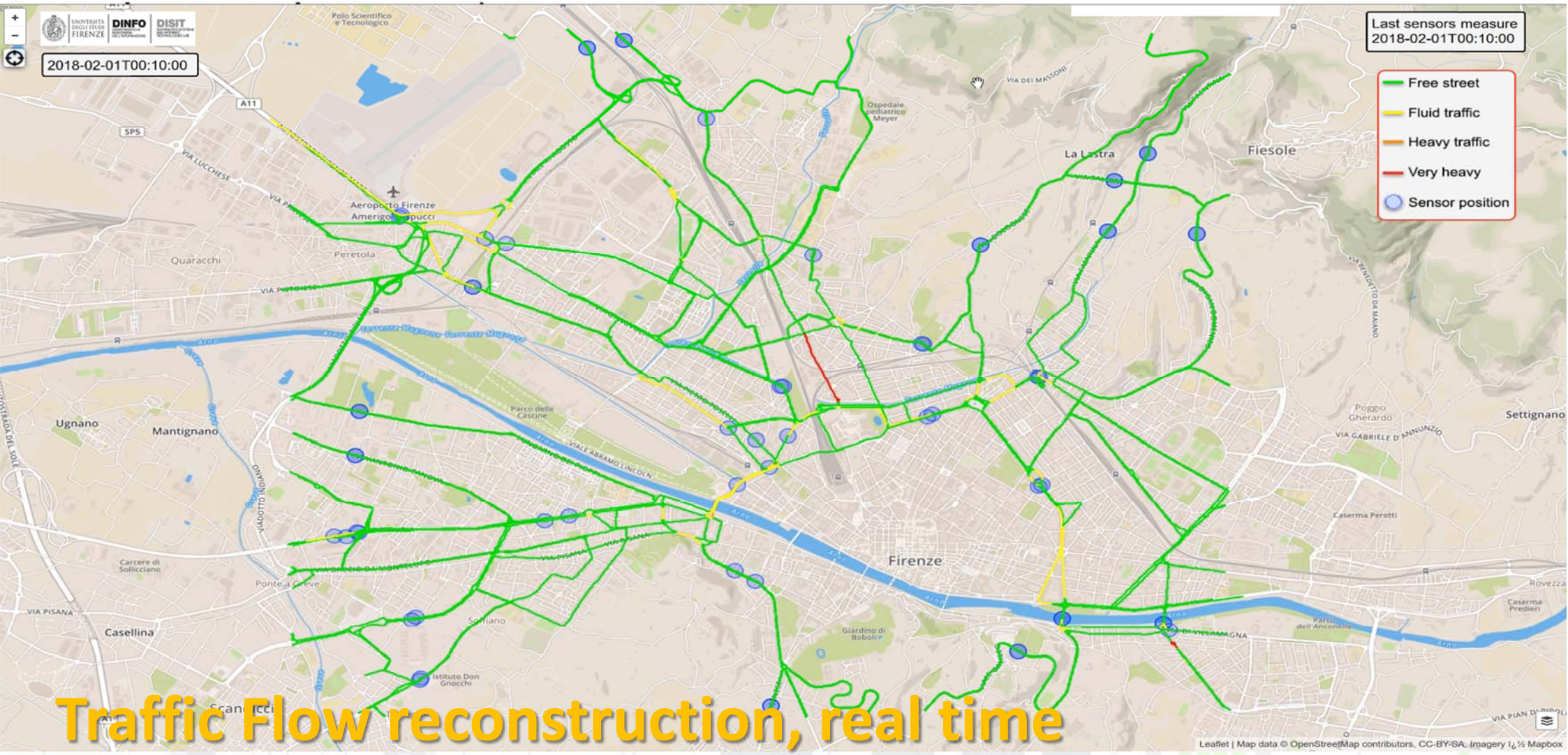
UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB http://www.disit.org

Traffic Flow Reconstruction

KM 4 CITY

Sii-Mobility

http://www.disit.org/siimobilitytrafficflow2/



Traffic Flow reconstruction, real time

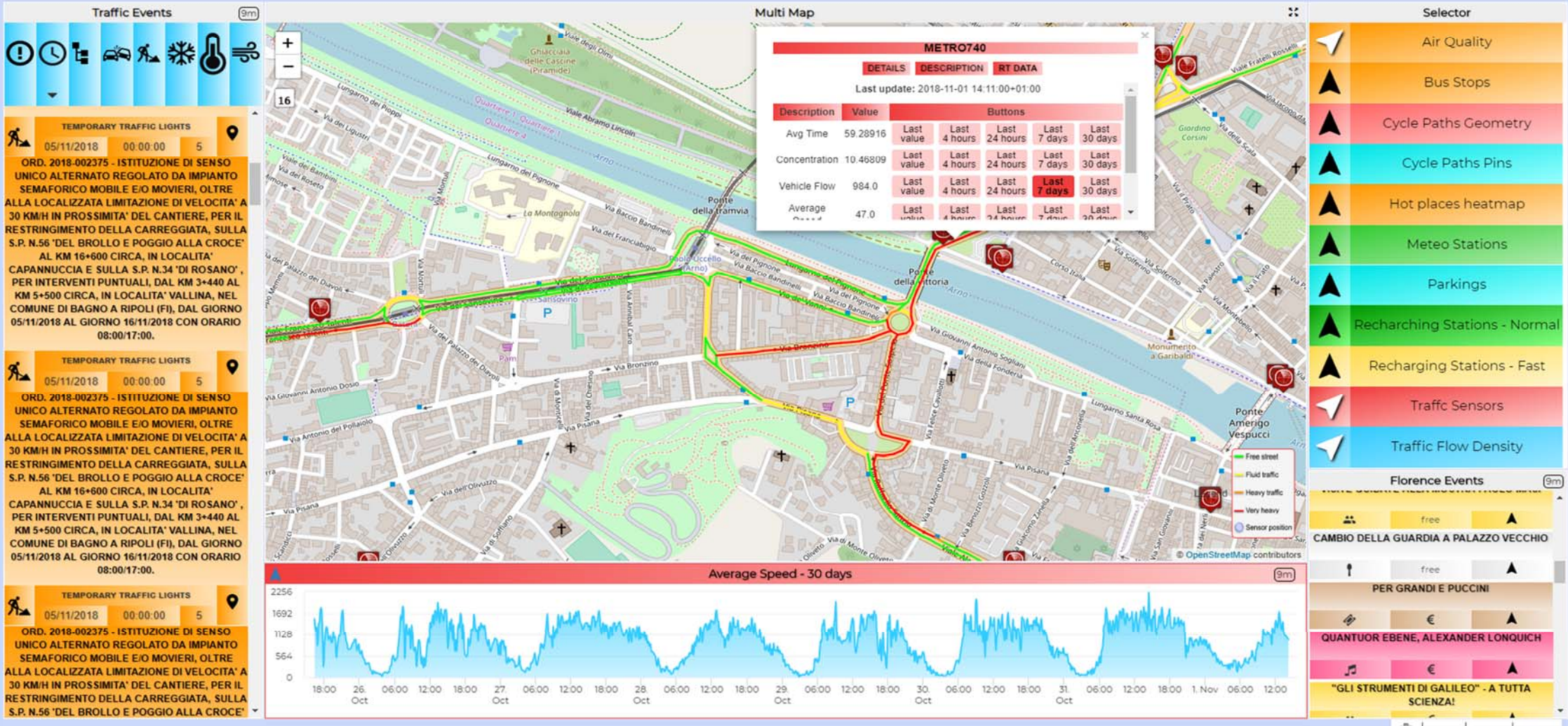
http://firenzetraffic.km4city.org

http://firenzetraffic.km4city.org/newSensors.html

http://firenzetraffic.km4city.org/new.html

Toscana Traffico

Thu 1 Nov 14:15:47



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

Traffic Flow reconstruction, real time

Mathematical model

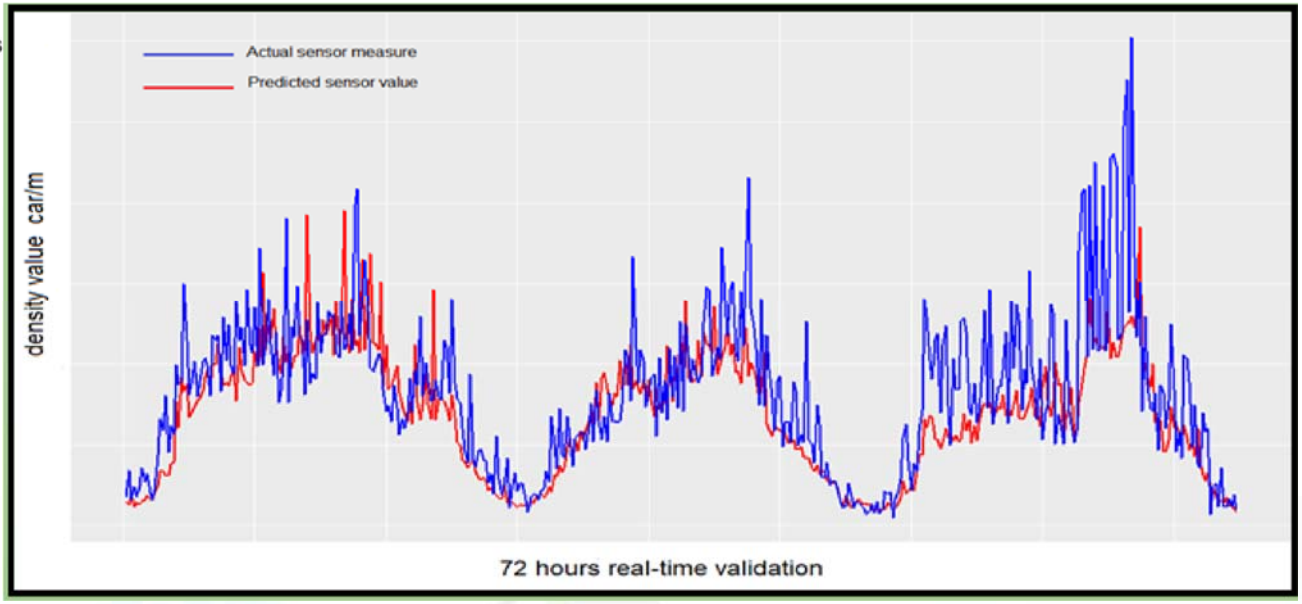
The vehicular traffic flow is propagated in the city graph according to a fluid dynamics model which is based on the conservation law of the vehicles. In a single road, it is described by the following partial differential equation:

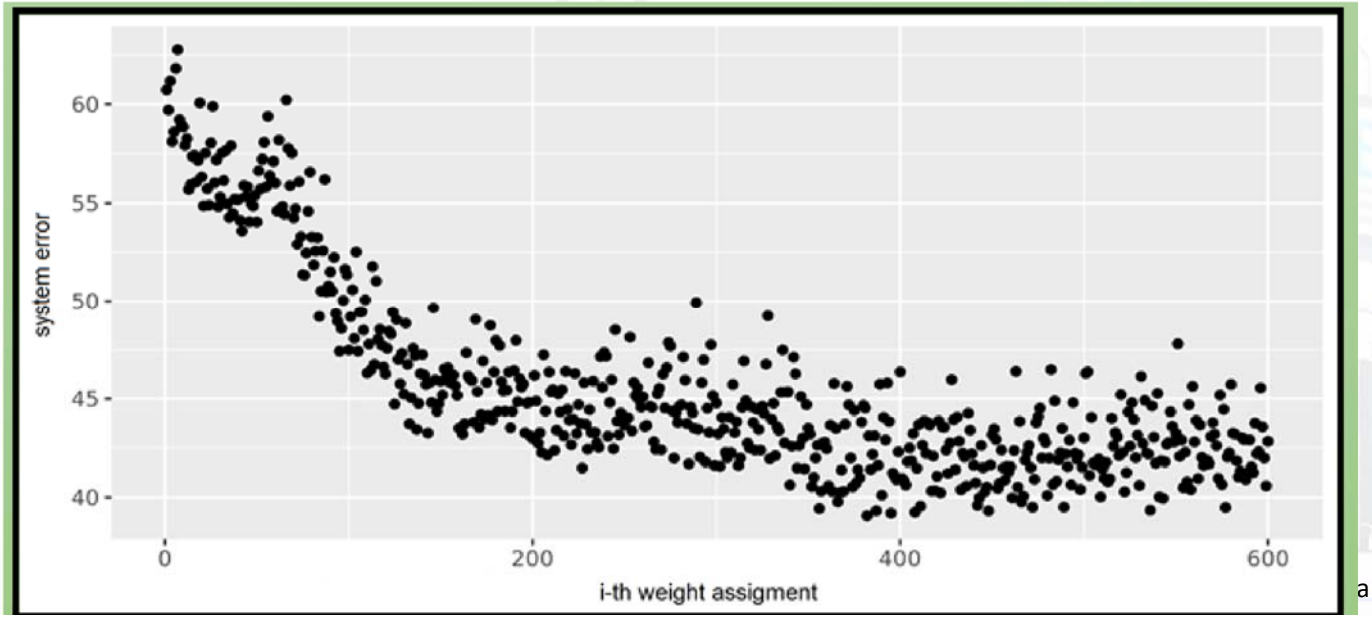
$$\frac{\partial \rho(t,x)}{\partial t} + \frac{\partial f\left(\rho(t,x)\right)}{\partial x} = 0$$

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

Precision: 85%

Monitoring Parking status

Sii-Mobility

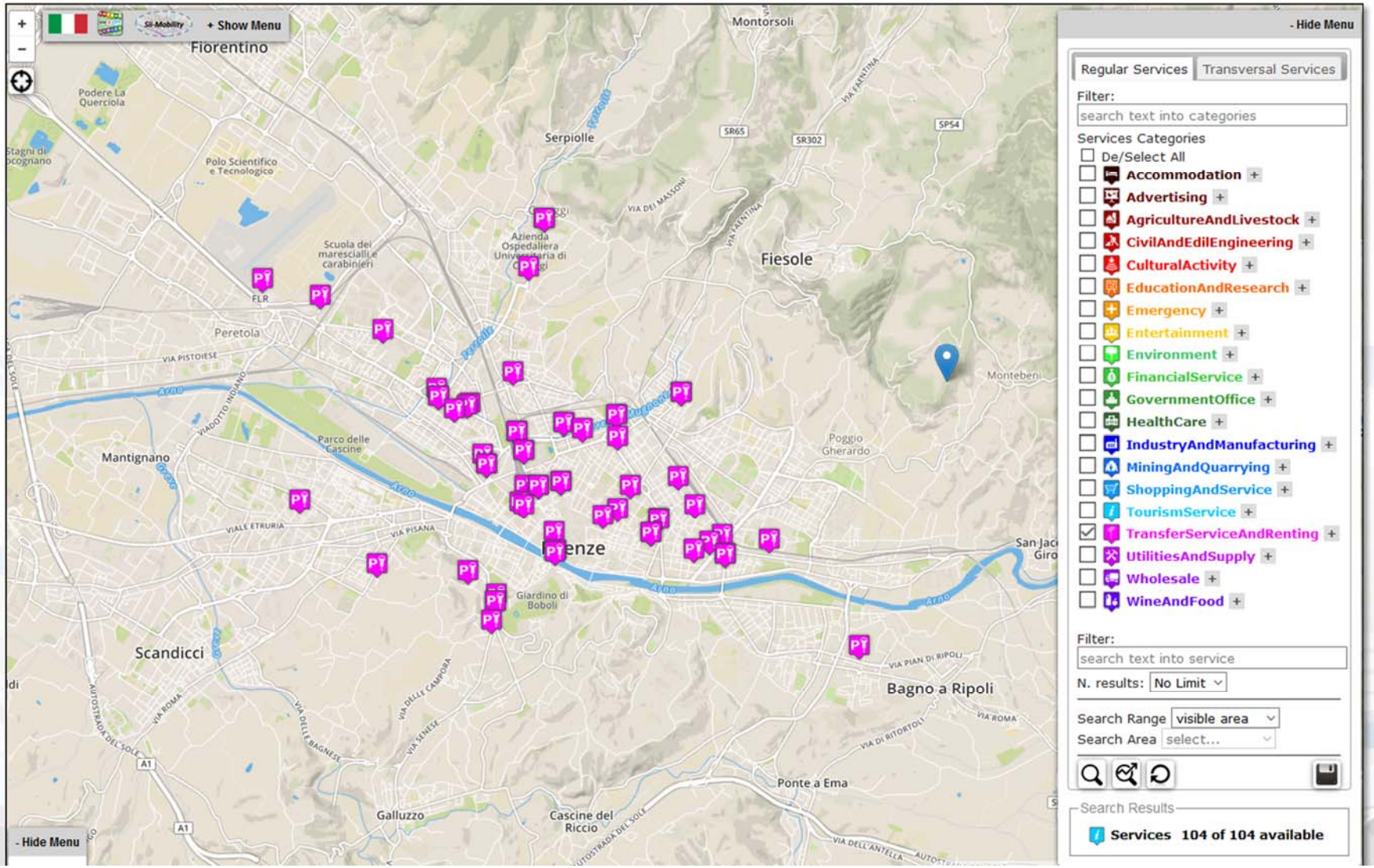
http://www.km4city.org/?controlRoom

Overview

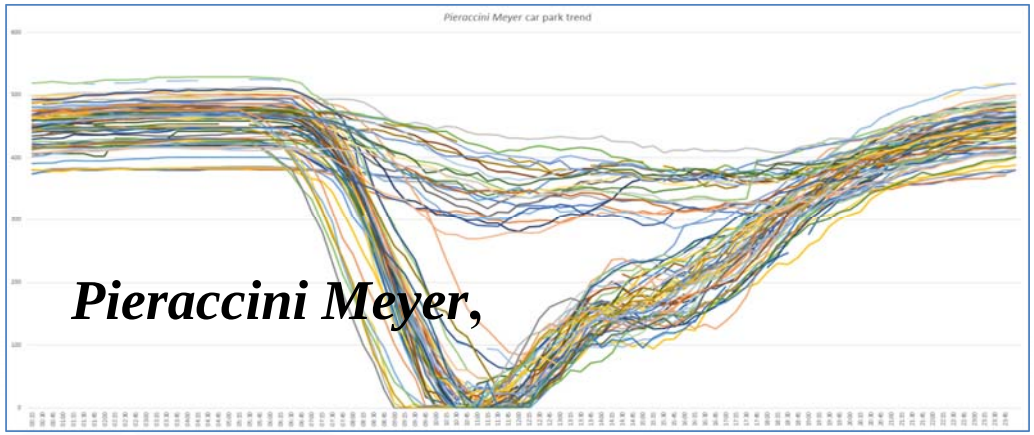
- **Analisi Dati raccolti via Social Media**
  - Predicting presences
- **Analisi Dati raccolti via Mobile App**
  - Tracce, matrici OD, heatmap
  - Regency e frequency
  - Impatto in uscita da Firenze
- **Analisi Dati raccolti dai Flussi di traffico**
  - Ricostruzione del traffico in punti non misurati
- **Analisi Dati raccolti dai Parcheggi**
  - Predizione dei posti liberi
- **Analisi Dati raccolti via Firenze WiFi**
  - Tracce, matrici OD, heatmap
  - Predicting presences
- **Analisi Dati Dati raccolti via Cellular**
  - Valutazione comparativa TIM-VODA
  - Valutazione comparativa FirenzeWiFi-Tim-Voda

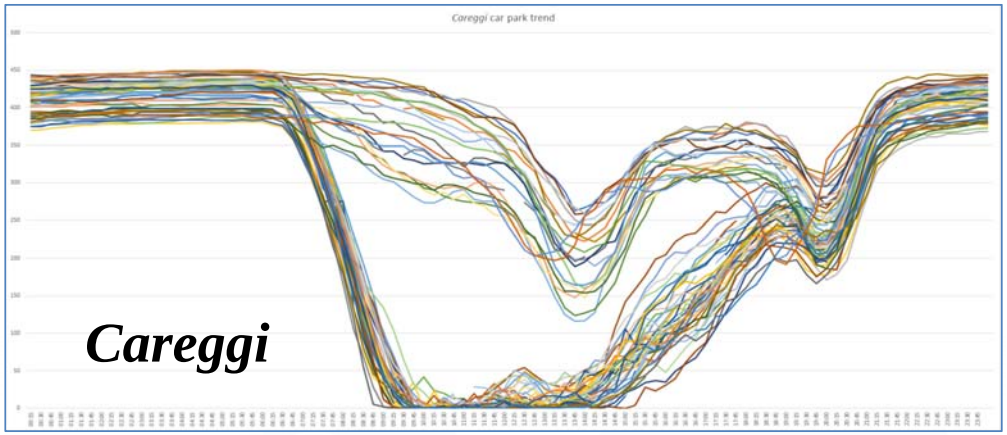
UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB
http://www.disit.org

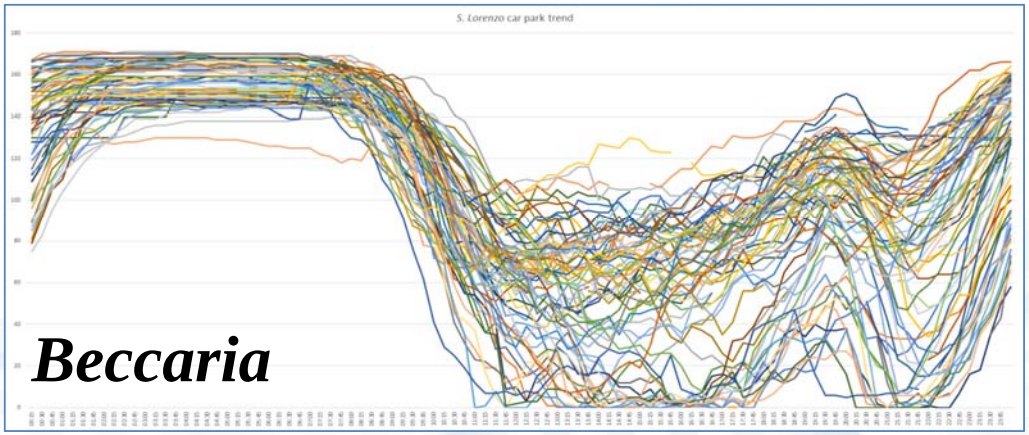
104 Parking in Tuscany

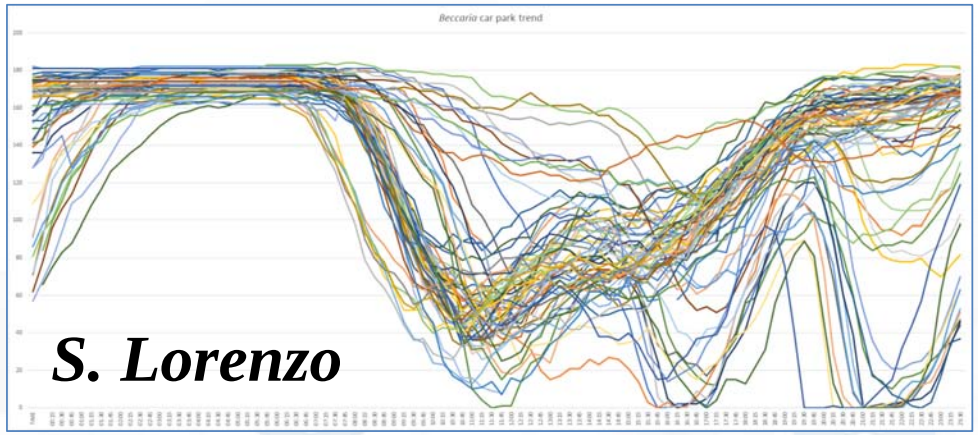


Free Parking space trends



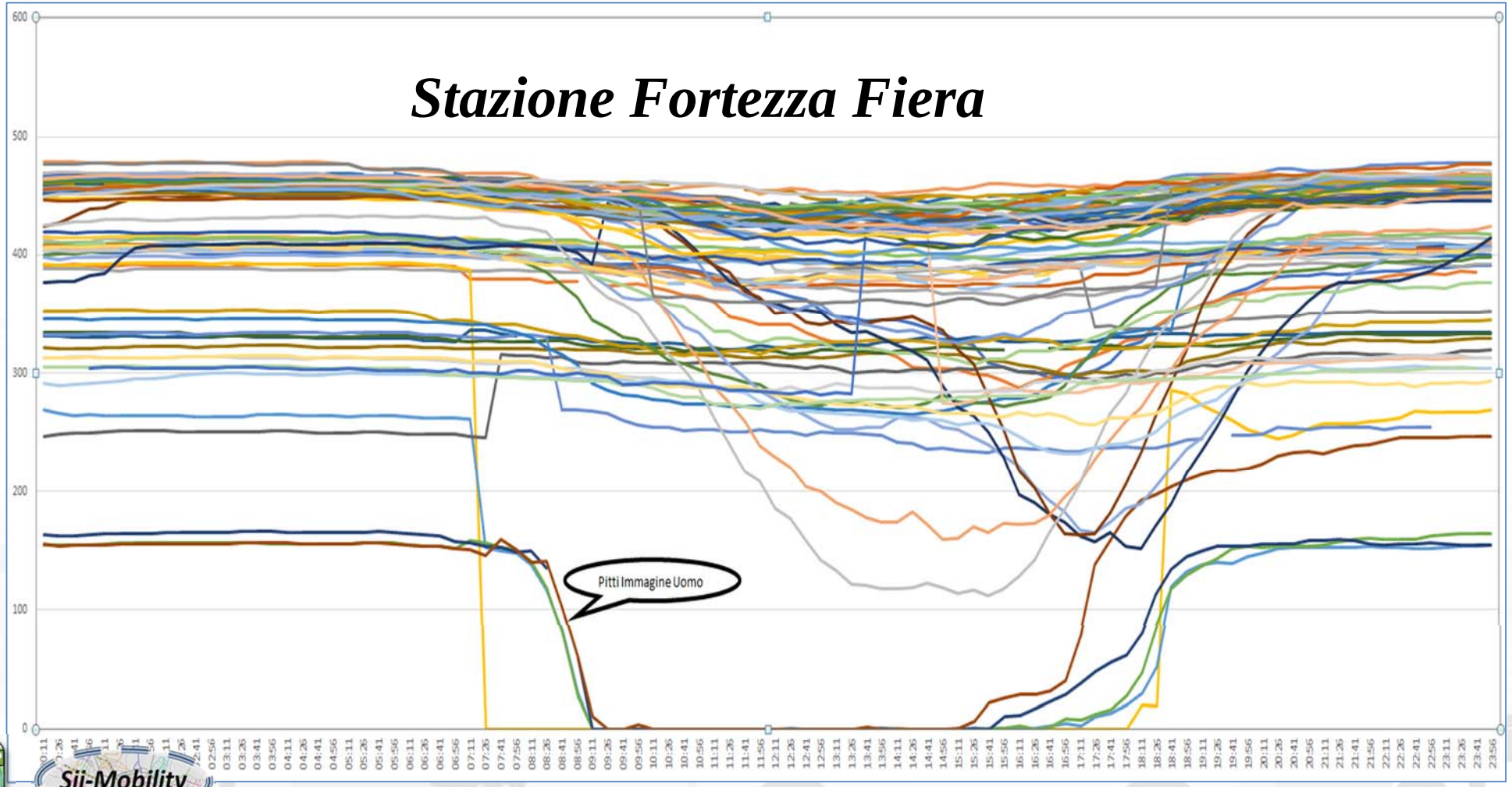




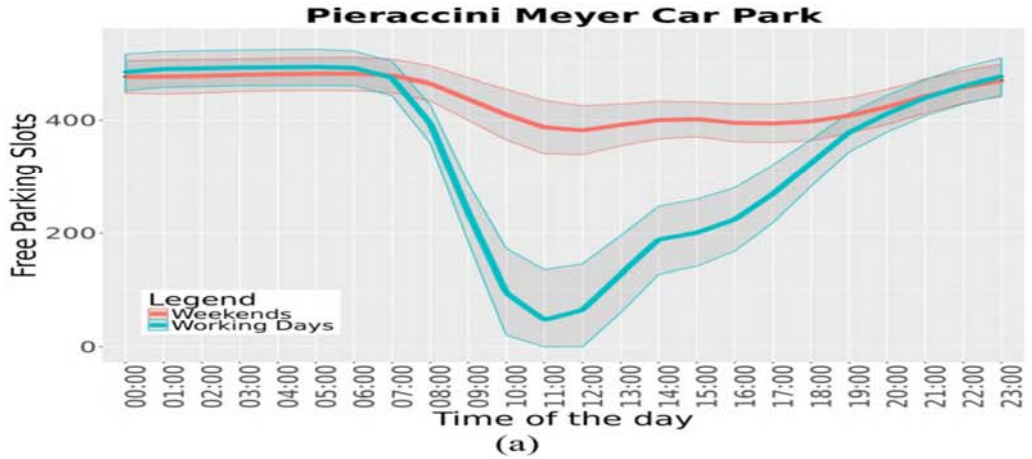


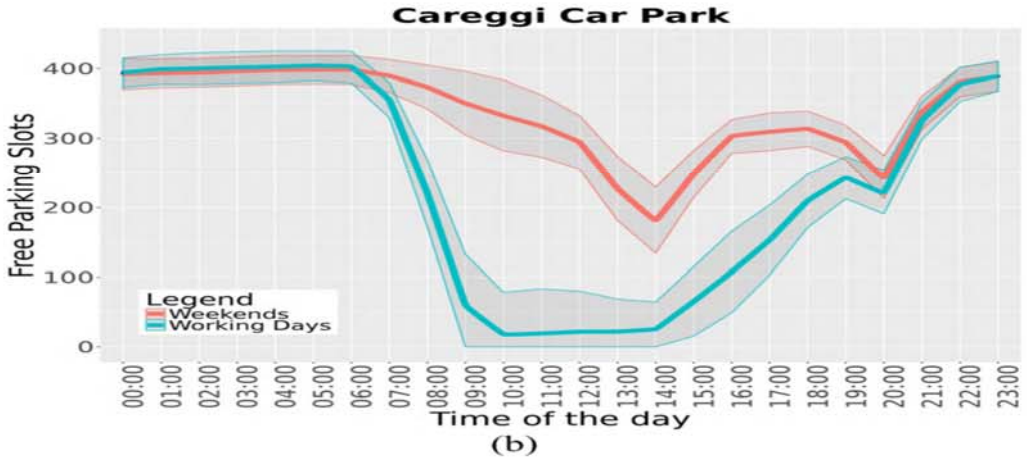
12 parking areas in Florence

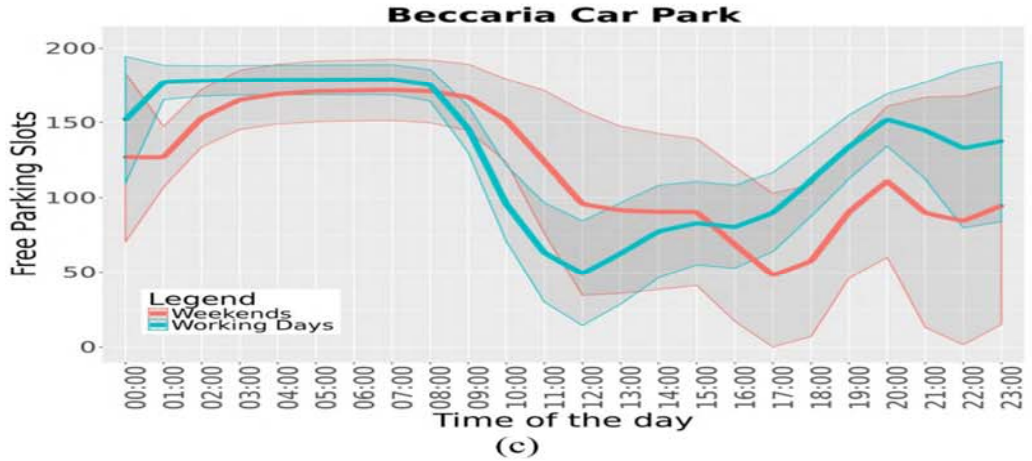
Free Parking space trends

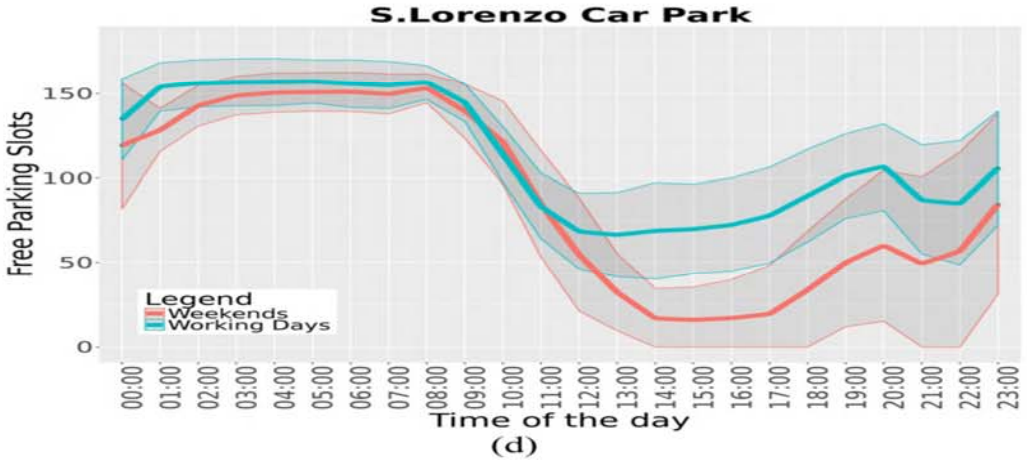


Free Parking space trends





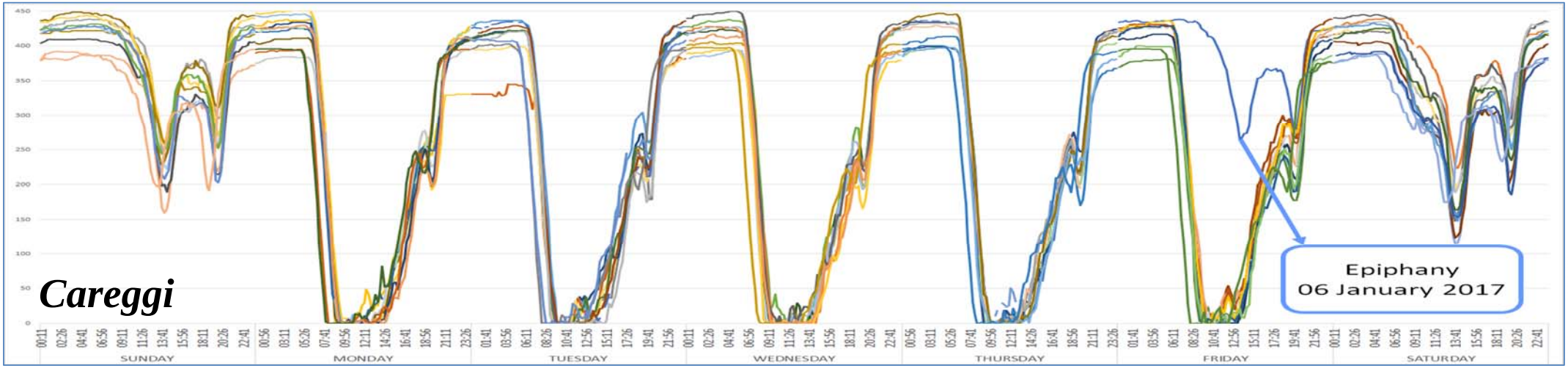


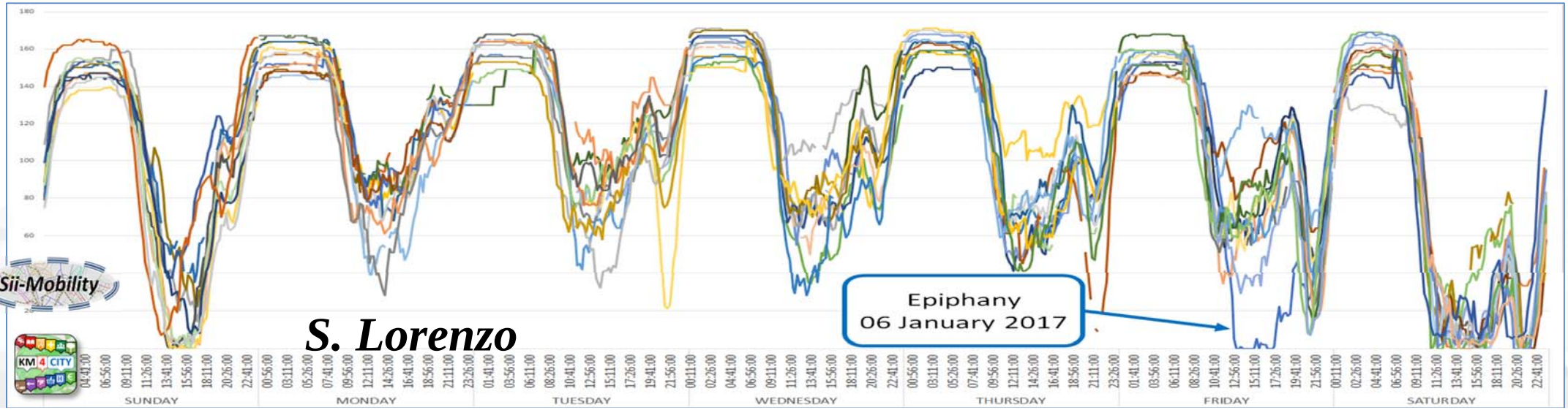


12 parking areas in Florence

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

Free Parking space trends





Free Parking PREDICTIONS

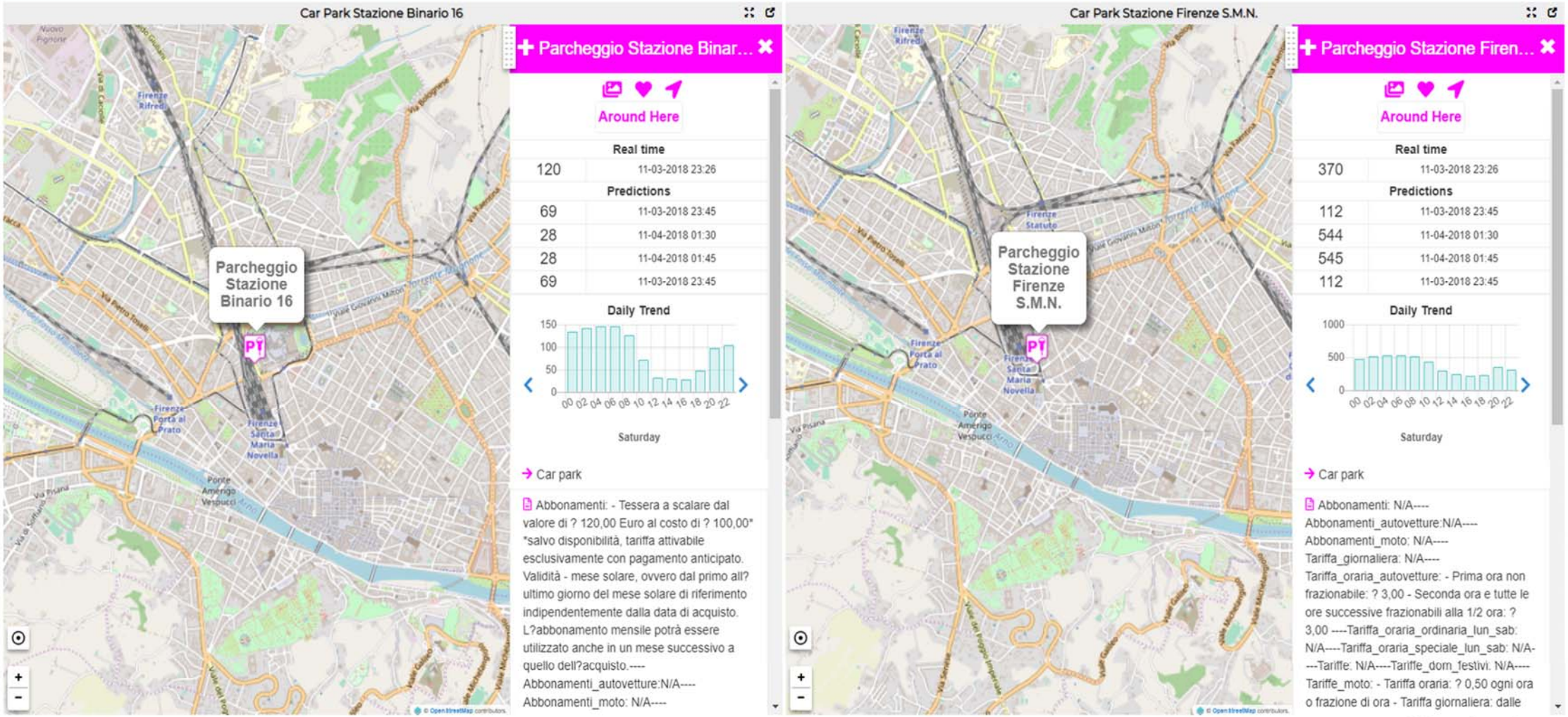
Careggi car park			
Model features	BRNN model results		
	R-squared	RMSE	MASE
Baseline	0.974	24	1.87
Baseline + Weather	0.975	24	1.75
Baseline + Traffic sensors	0.975	24	2.04
Baseline + Weather + Traffic sensors	0.975	24	1.87

- Active on Mobile Apps as:
  - «Firenze dove cosa»
  - «Toscana dove cosa»

Precision: 97,5%

Monitoring Station for Parking

Sat 3 Nov 23:39:55



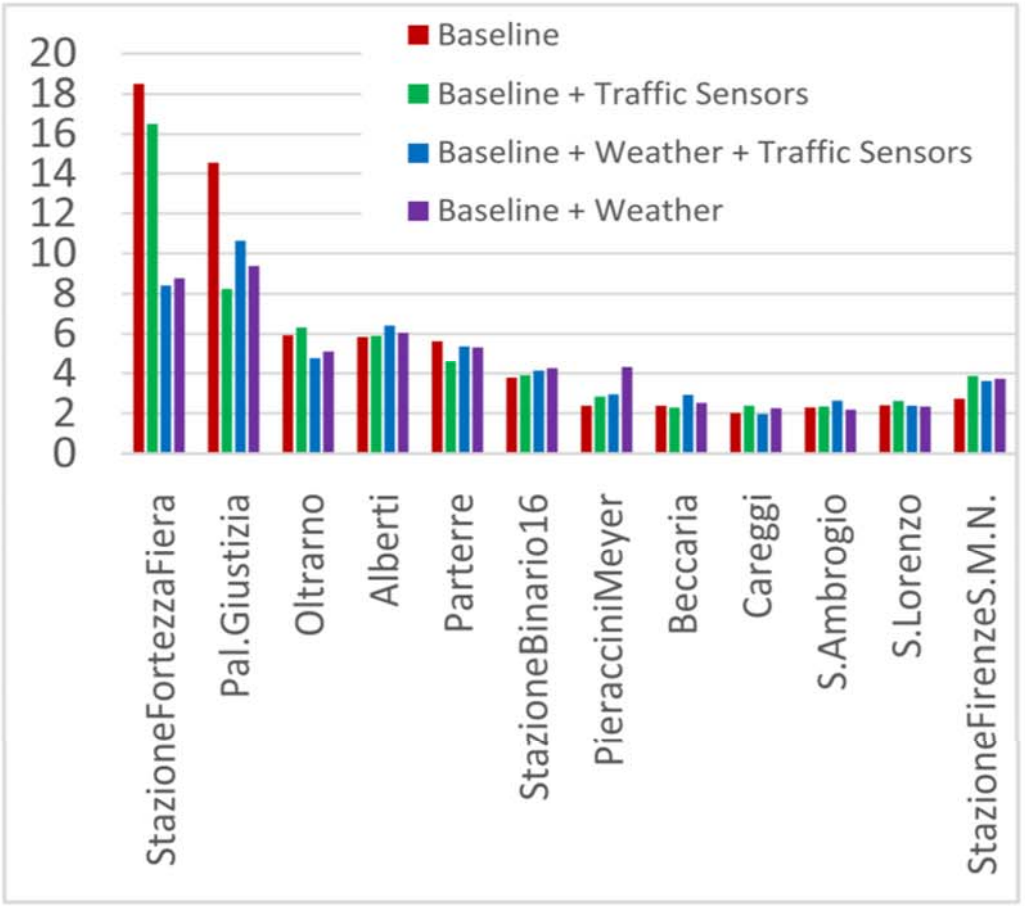
https://www.disit.org/dashboardSmartCity/view/index.php?iddasboard=MjQ2

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

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

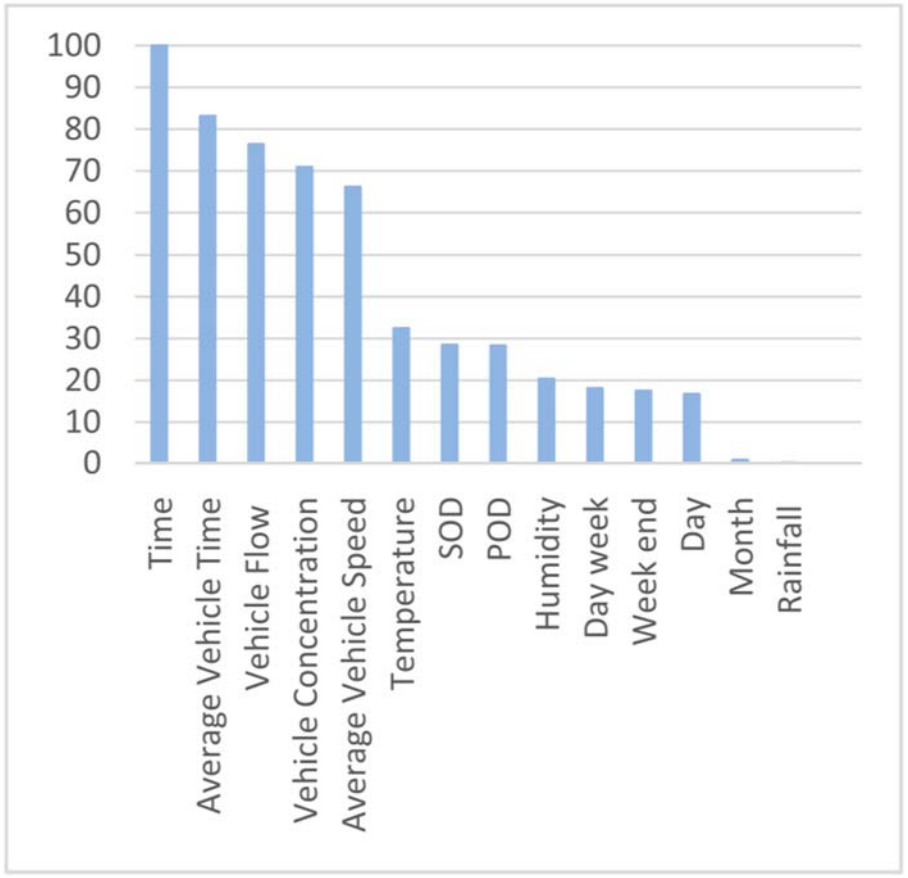


Free Parking PREDICTIONS

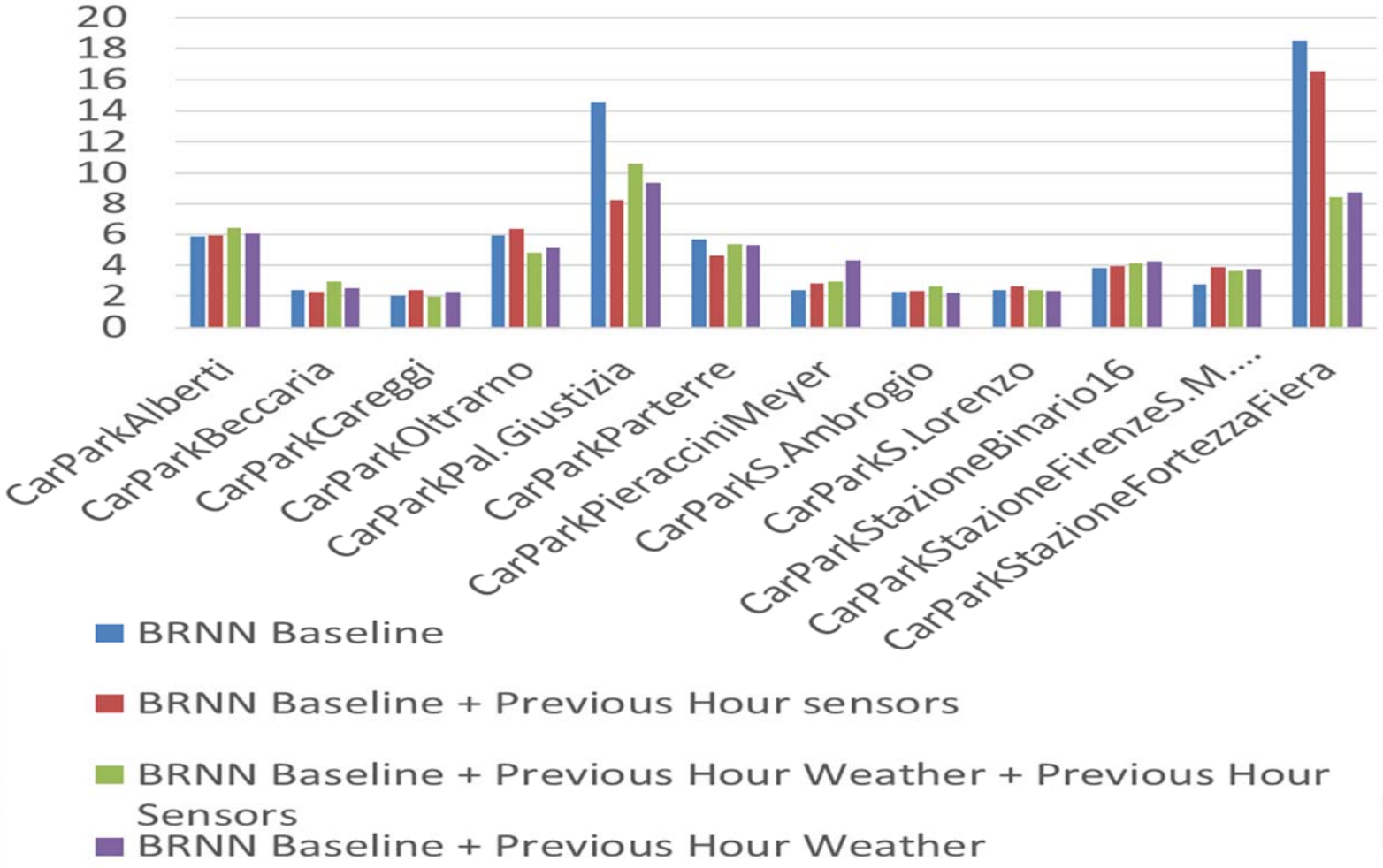
Performances

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

Relevance of Variable



Free Space on Parking lots



Monitoring City users Via Wi-Fi

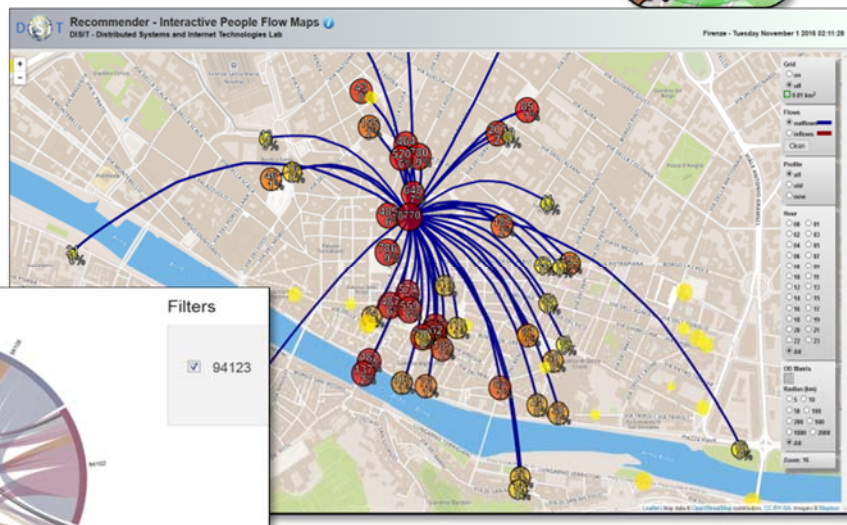
http://www.km4city.org/?controlRoom
http://www.km4city.org/?devTools
http://www.km4city.org/?infoDocs
http://www.km4city.org/?app

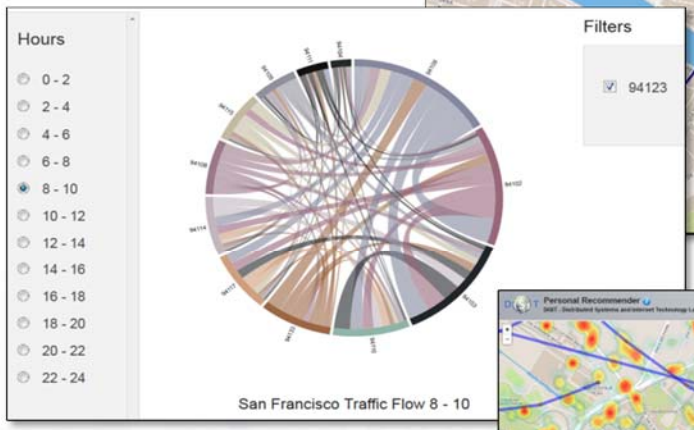
Overview

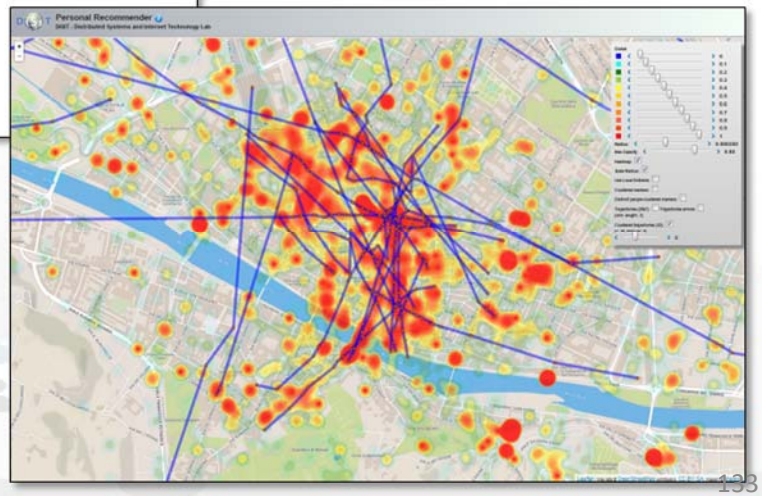
- **Analisi Dati raccolti via Social Media**
  - Predicting presences
- **Analisi Dati raccolti via Mobile App**
  - Tracce, matrici OD, heatmap
  - Regency e frequency
  - Impatto in uscita da Firenze
- **Analisi Dati raccolti dai Flussi di traffico**
  - Ricostruzione del traffico in punti non misurati
- **Analisi Dati raccolti dai Parcheggi**
  - Predizione dei posti liberi
- **Analisi Dati raccolti via Firenze WiFi**
  - Tracce, matrici OD, heatmap
  - Predicting presences
- **Analisi Dati Dati raccolti via Cellular**
  - Valutazione comparativa TIM-VODA
  - Valutazione comparativa FirenzeWiFi-Tim-Voda

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

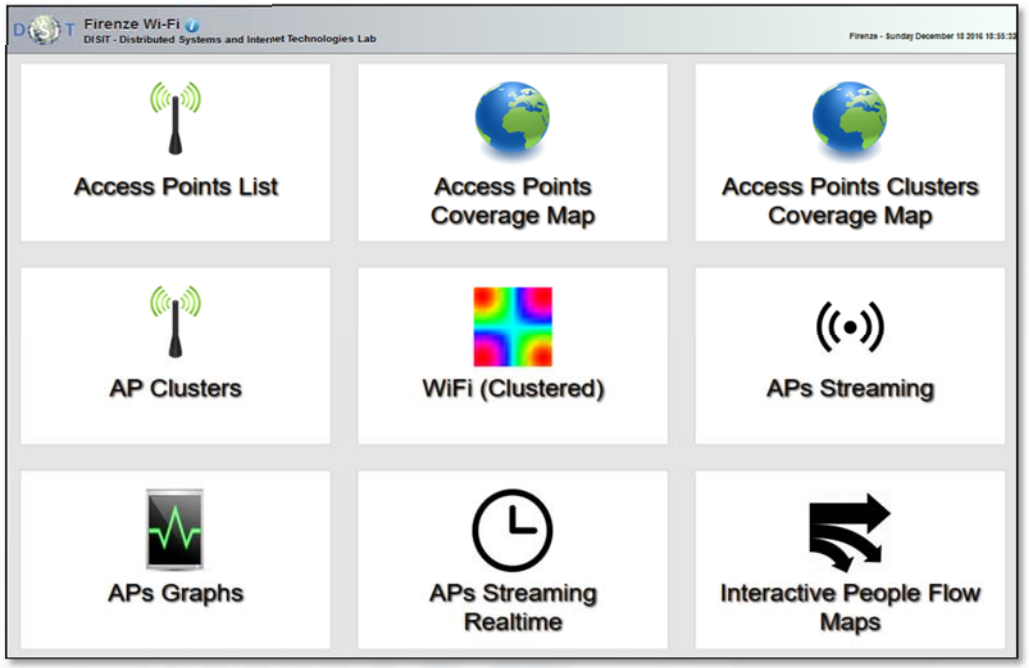






Monitoring City usage via Wi-Fi

- **Instrumenting Wi-Fi**
  - 1500 AP → 345 instrumented
  - Stream from AP to DISIT Lab
  - Real time monitoring → dashboard
- **Data Analytics**
  - heat maps
  - Analysis of user behavior
  - Clustering user behavior
  - Predictive models about user behavior
  - Identification of critical conditions, anomalies

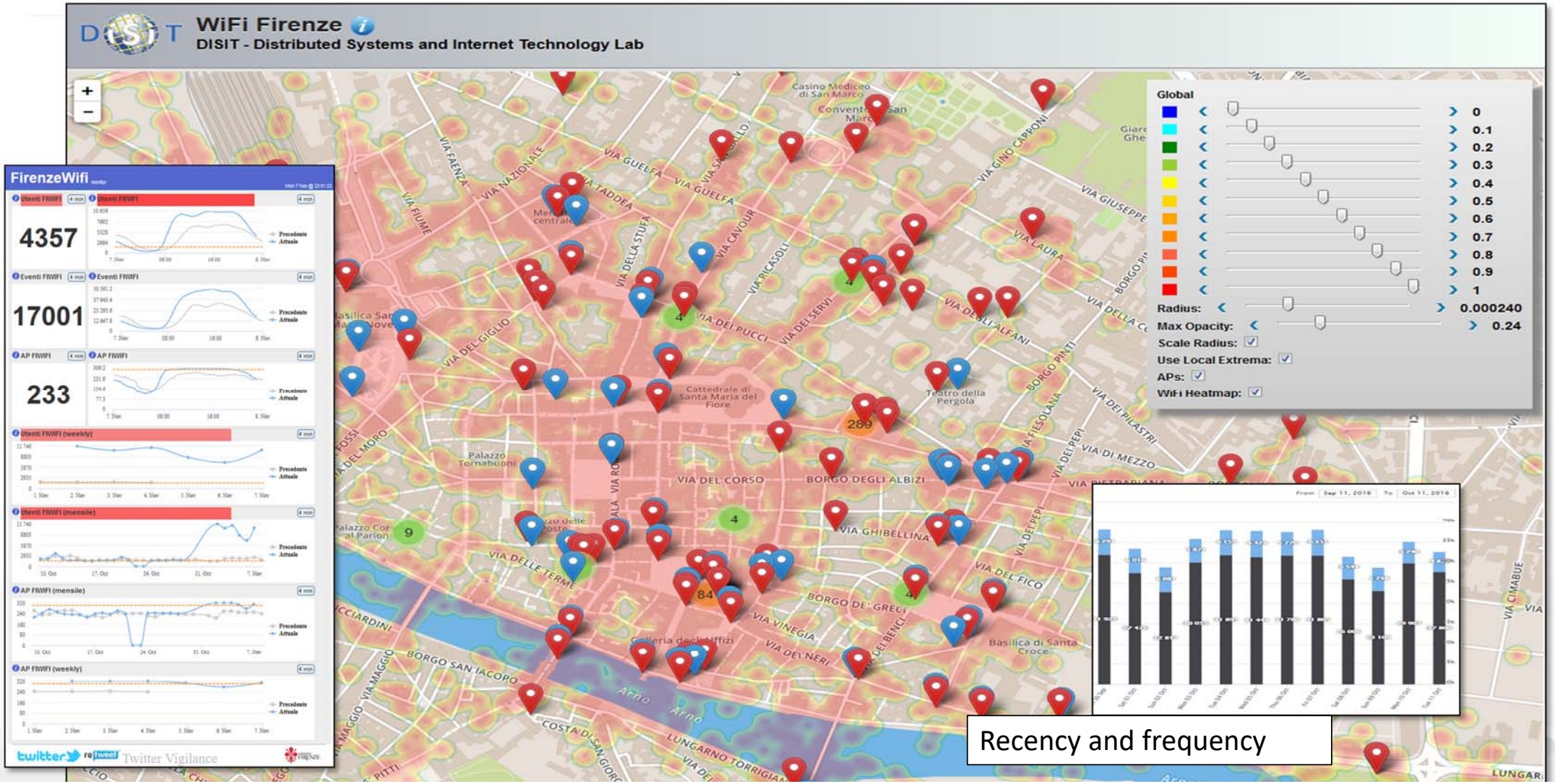


Trajectories from Wi- Fi and Cells

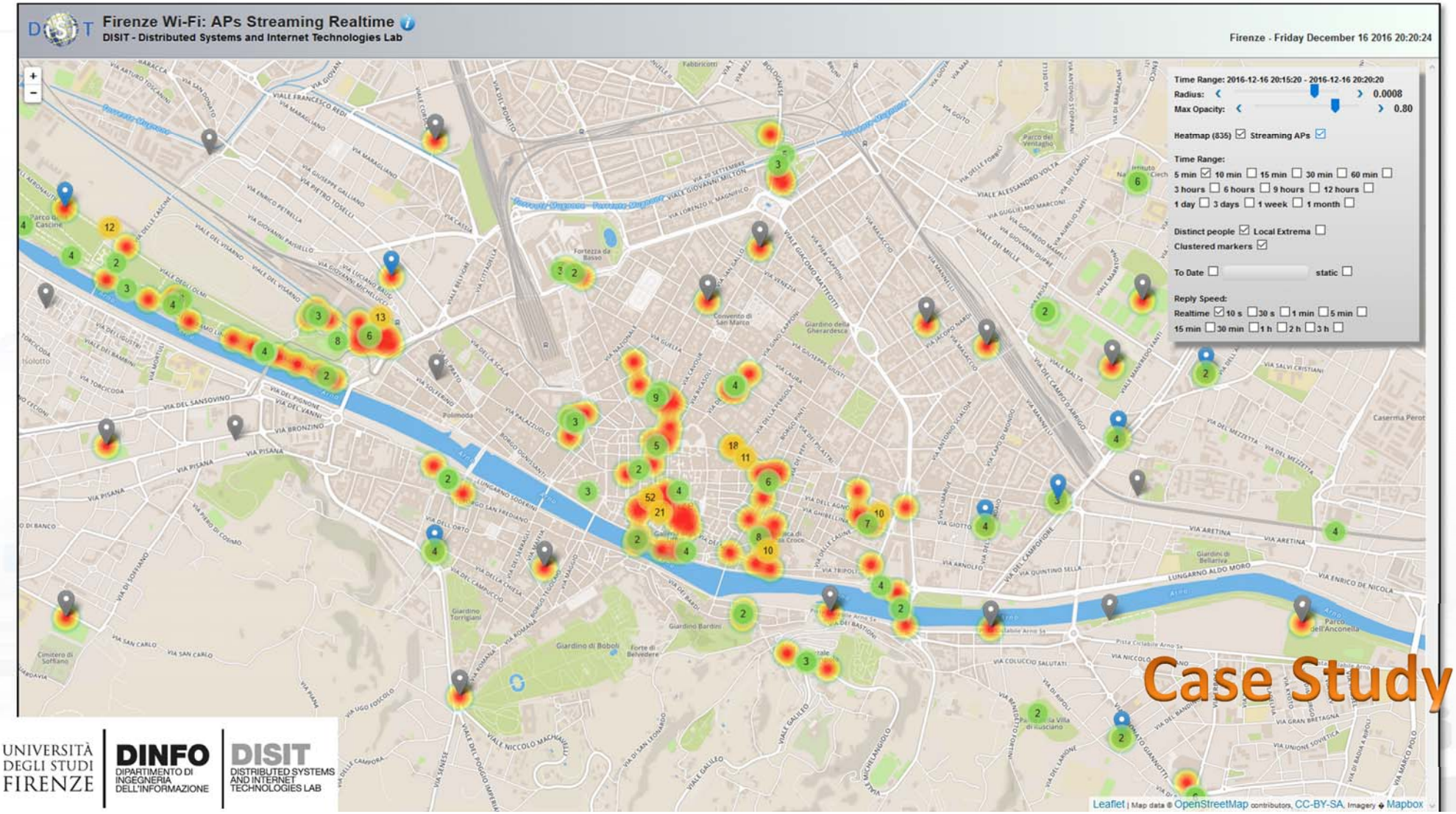
- **From Wi-Fi:**
  - Anonymized data
  - Constrained to be on AP GPS location
  - Change of AP location
  - Different precision
  - Non dense spatial coverage
  - Problems in the case of high speed traffic
- **From Telecom Operators Cells:**
  - Typically not densely measured in time
  - Precise tracking not possible with data provides, only statistical reconstruction
- **Clustering**: time, space, user kind, etc.

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

Wi-Fi Monitor tool



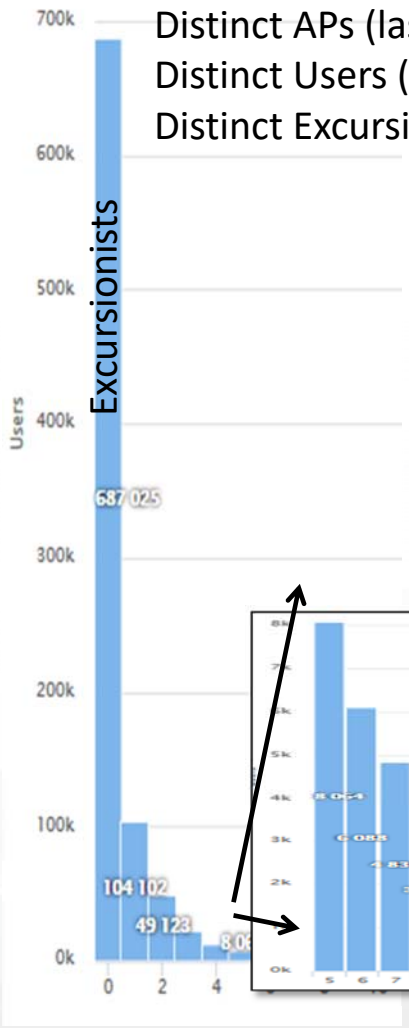
Real Time Monitoring of Wi-Fi network

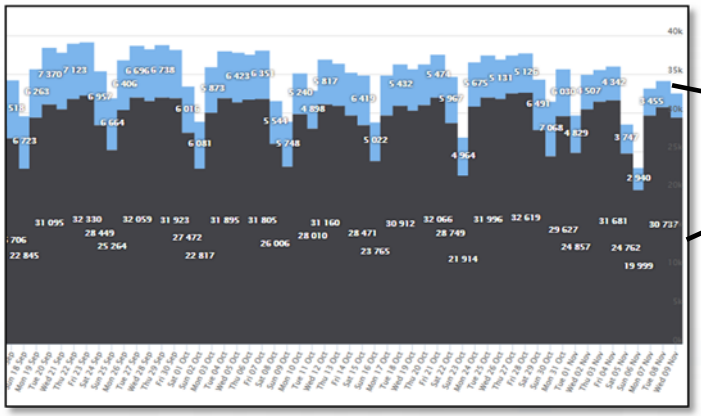


Case Study A

User Behavior Analysis

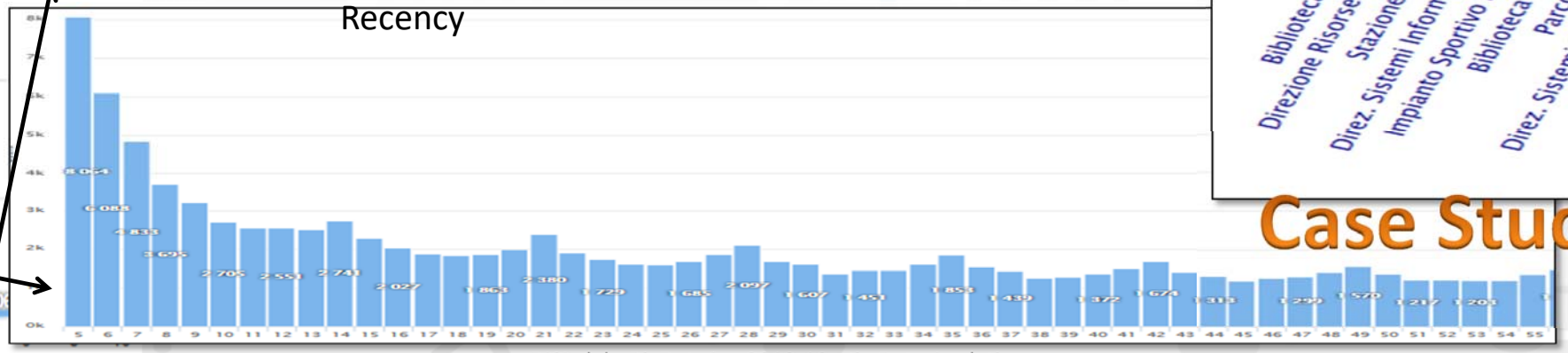
Distinct APs: 343
Distinct APs (last 24 hours): 311
Distinct Users (last 180 days): 1102098
Distinct Excursionists (last 180 days, < 24 h): 687025

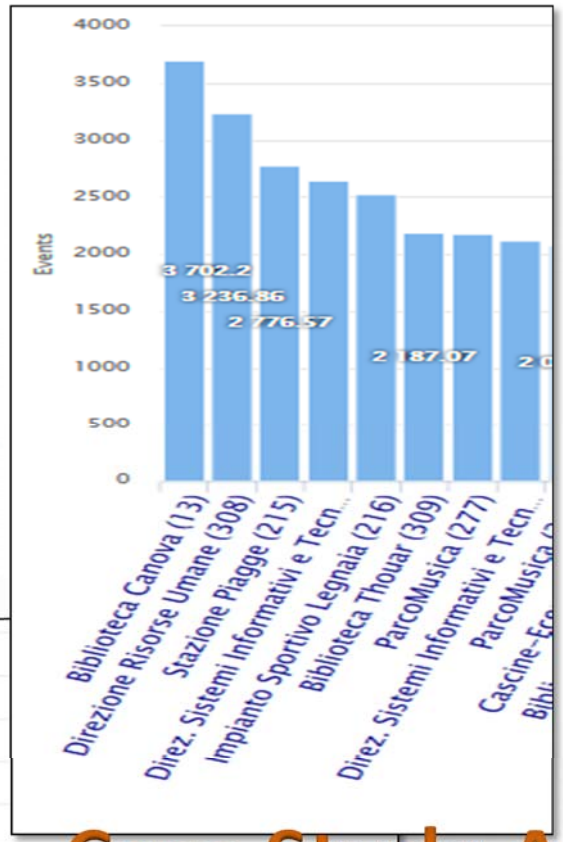




New City Users
VS
Returning

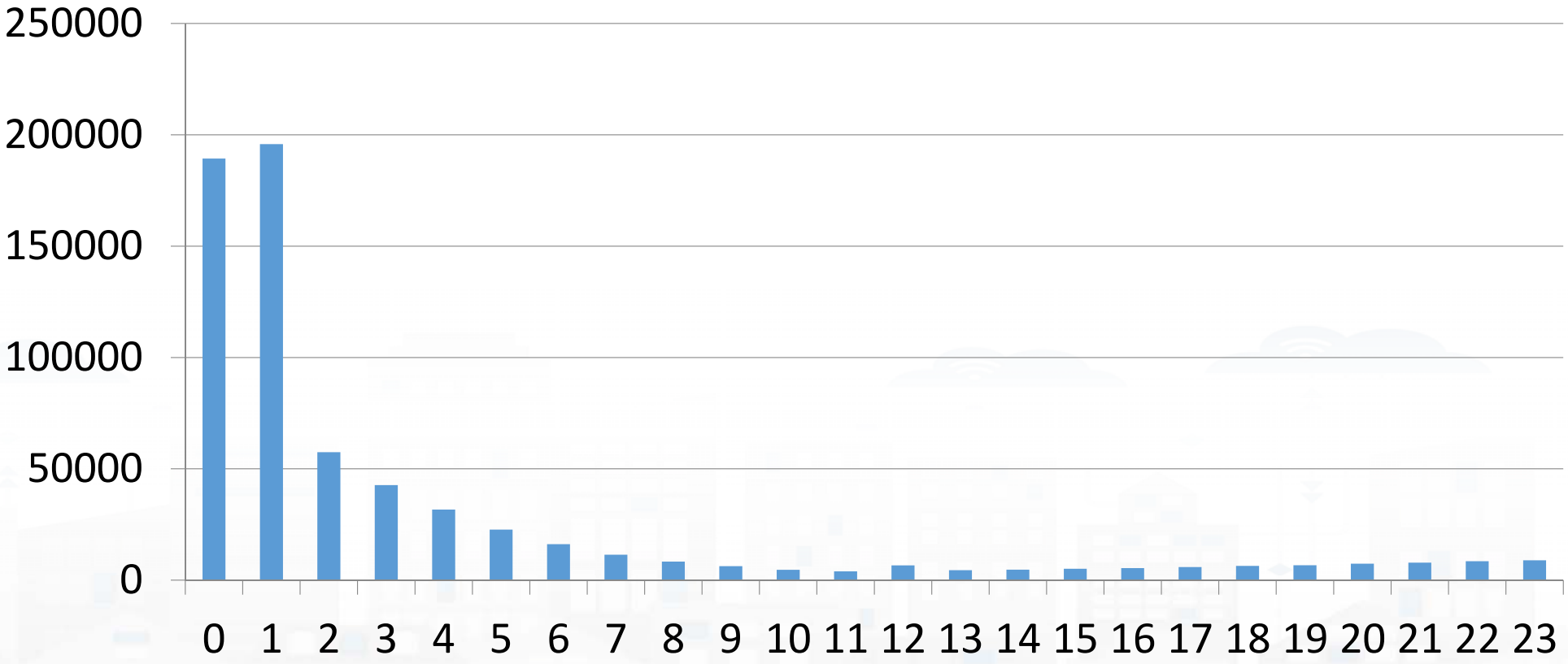
Recency





Case Study A

Distribution in the first 24 hours



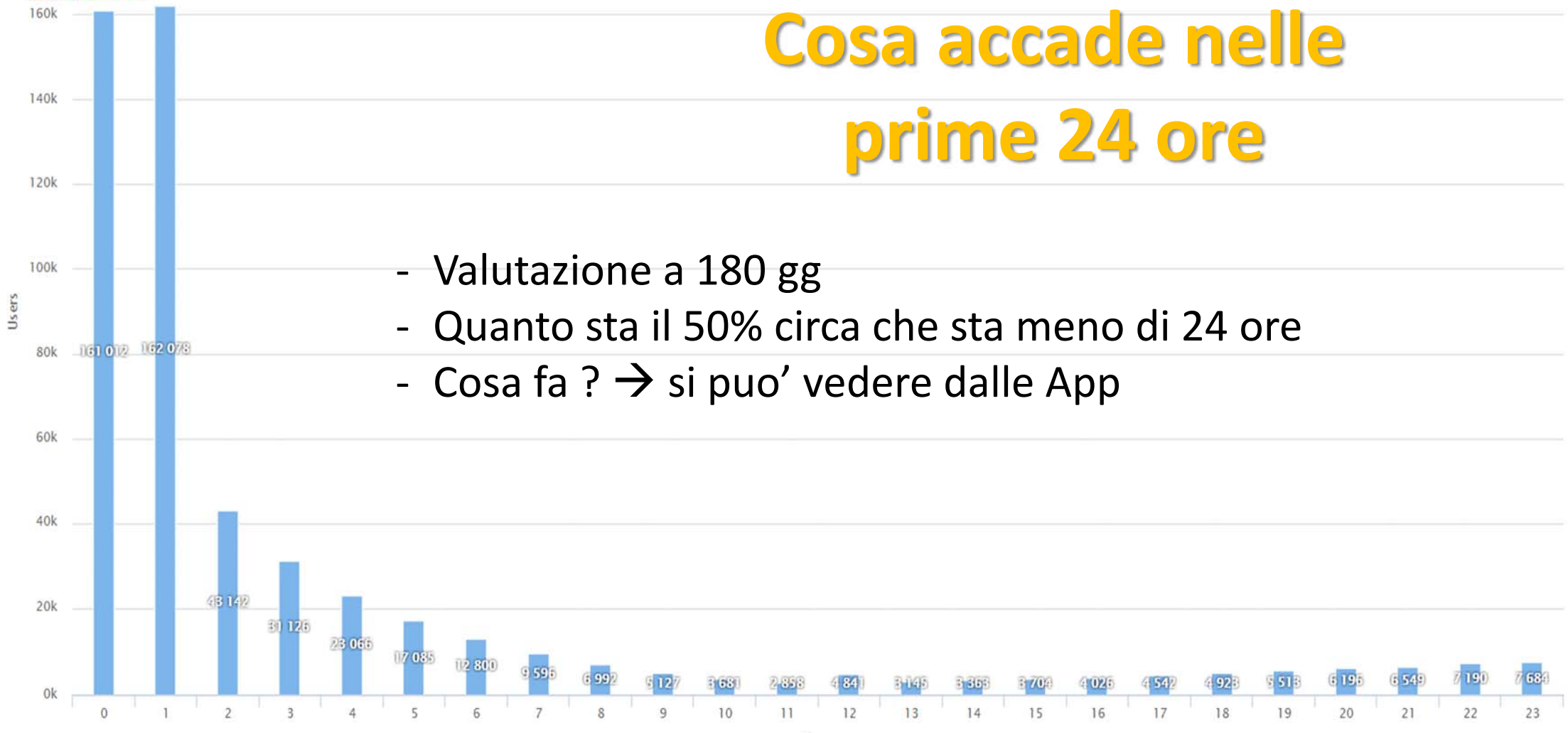
4 means: permanence for more than 3 hours and less than 4 hours

Firenze Wi-Fi

permanenze da 1 a xxx

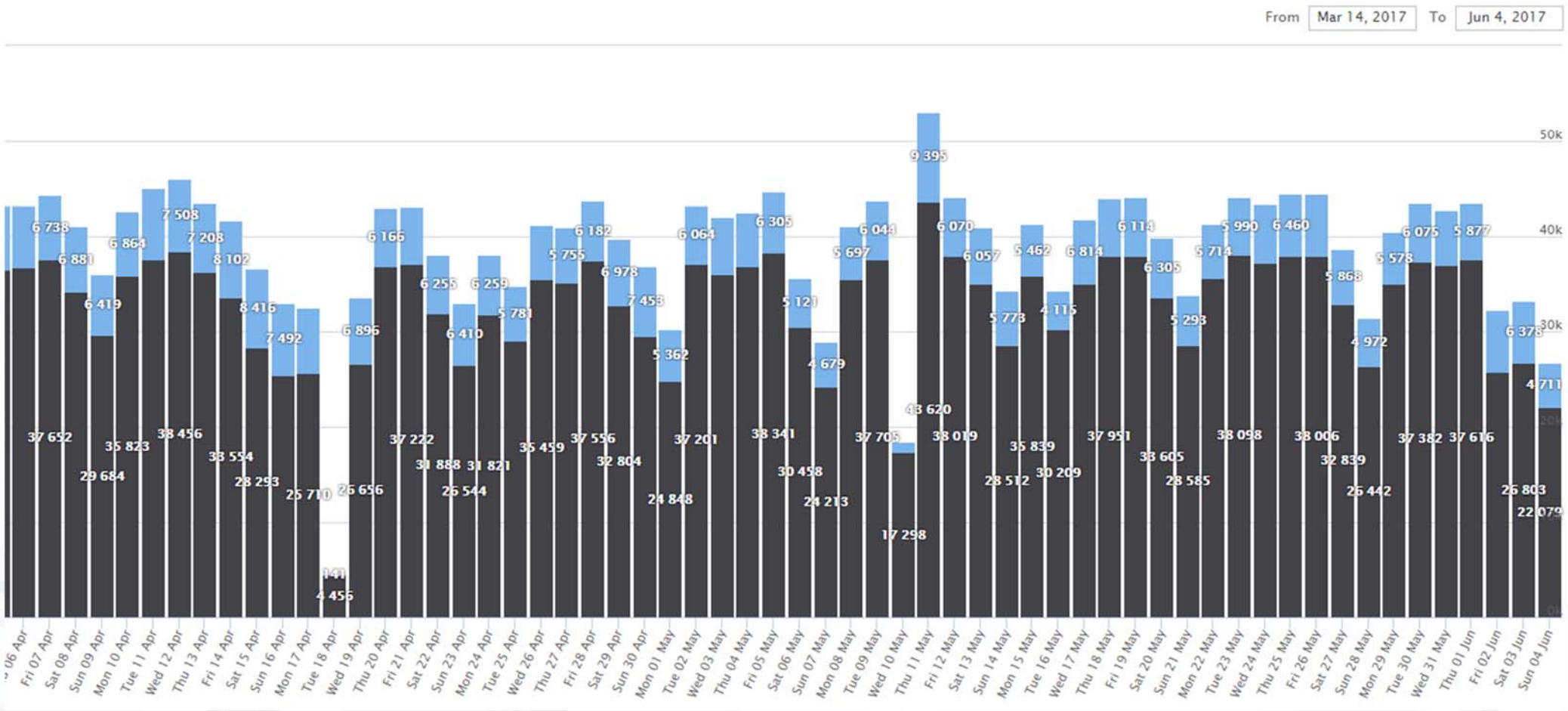
- Valutazione a 180 gg
- 50% circa sta meno di 24 ore

Firenze WiFi
Cosa accade nelle prime 24 ore

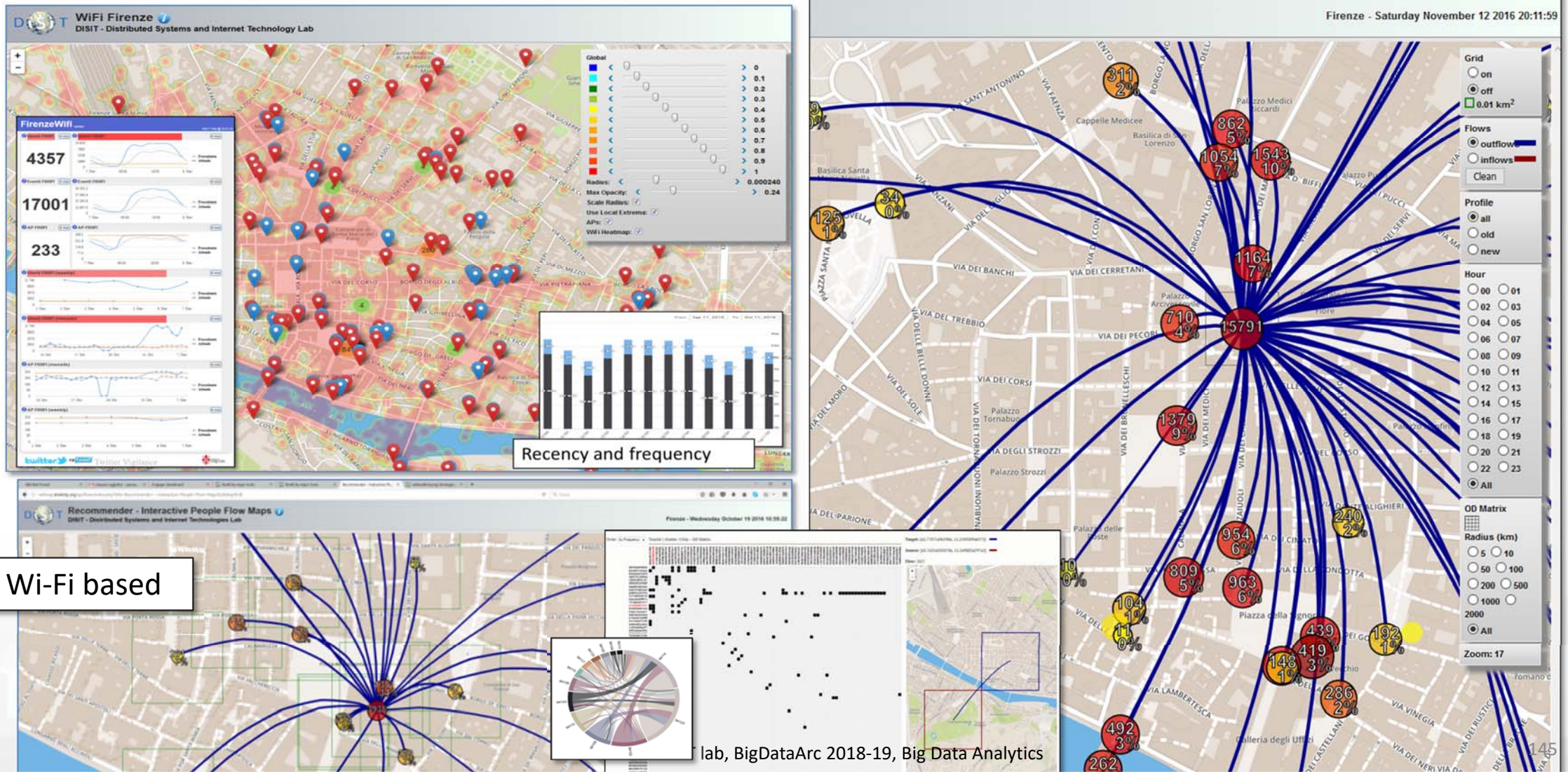


- Valutazione a 180 gg
- Quanto sta il 50% circa che sta meno di 24 ore
- Cosa fa ? → si puo' vedere dalle App

Firenze Wi-Fi
nuovi arrivi a Firenze



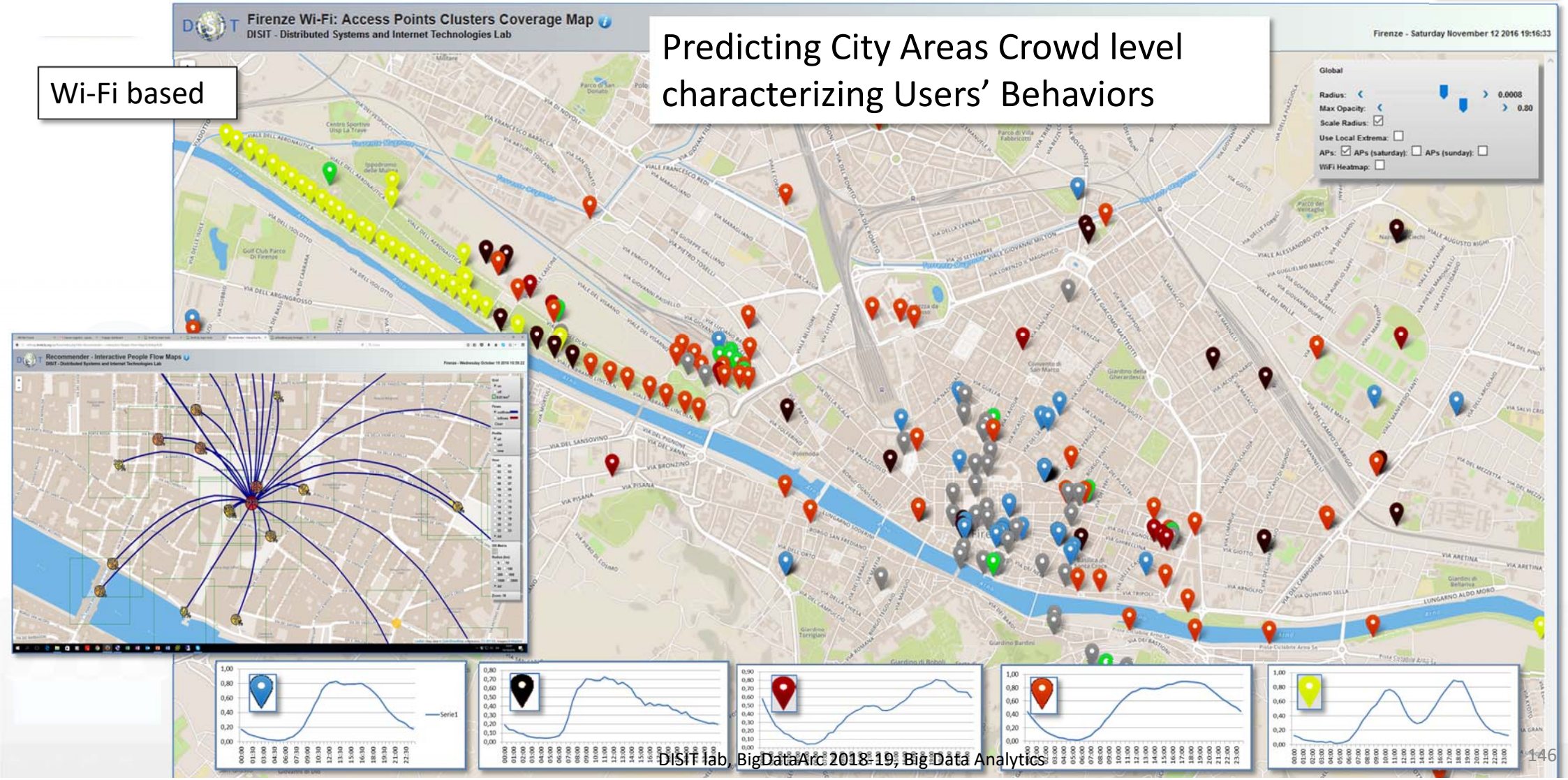
Origin Destination Matrix Estimation

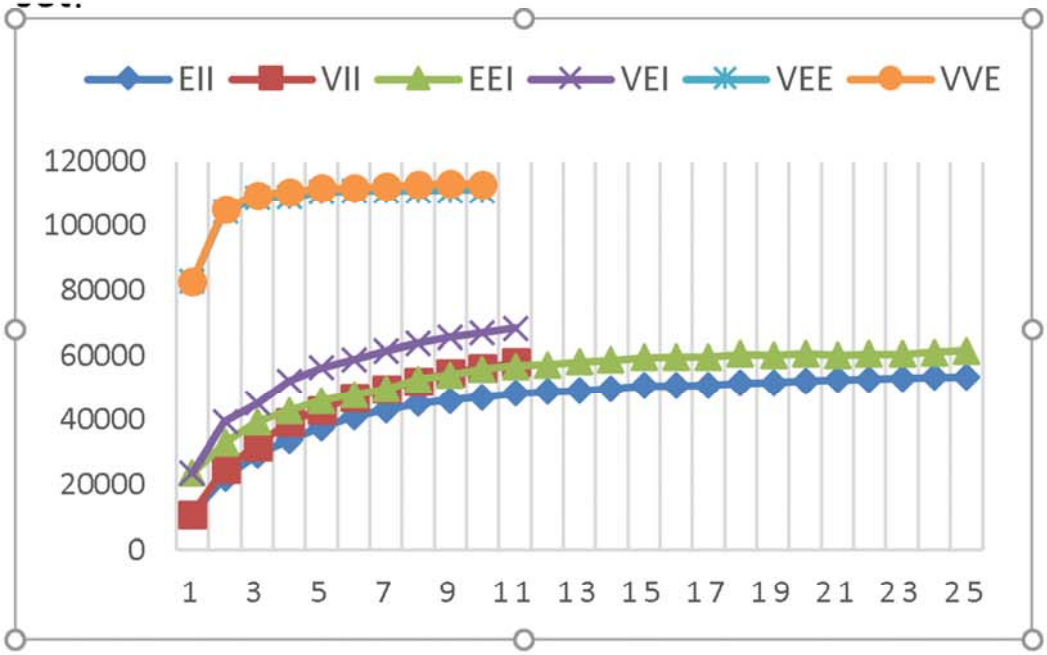


Characterizing City Areas

Wi-Fi based

Predicting City Areas Crowd level characterizing Users' Behaviors





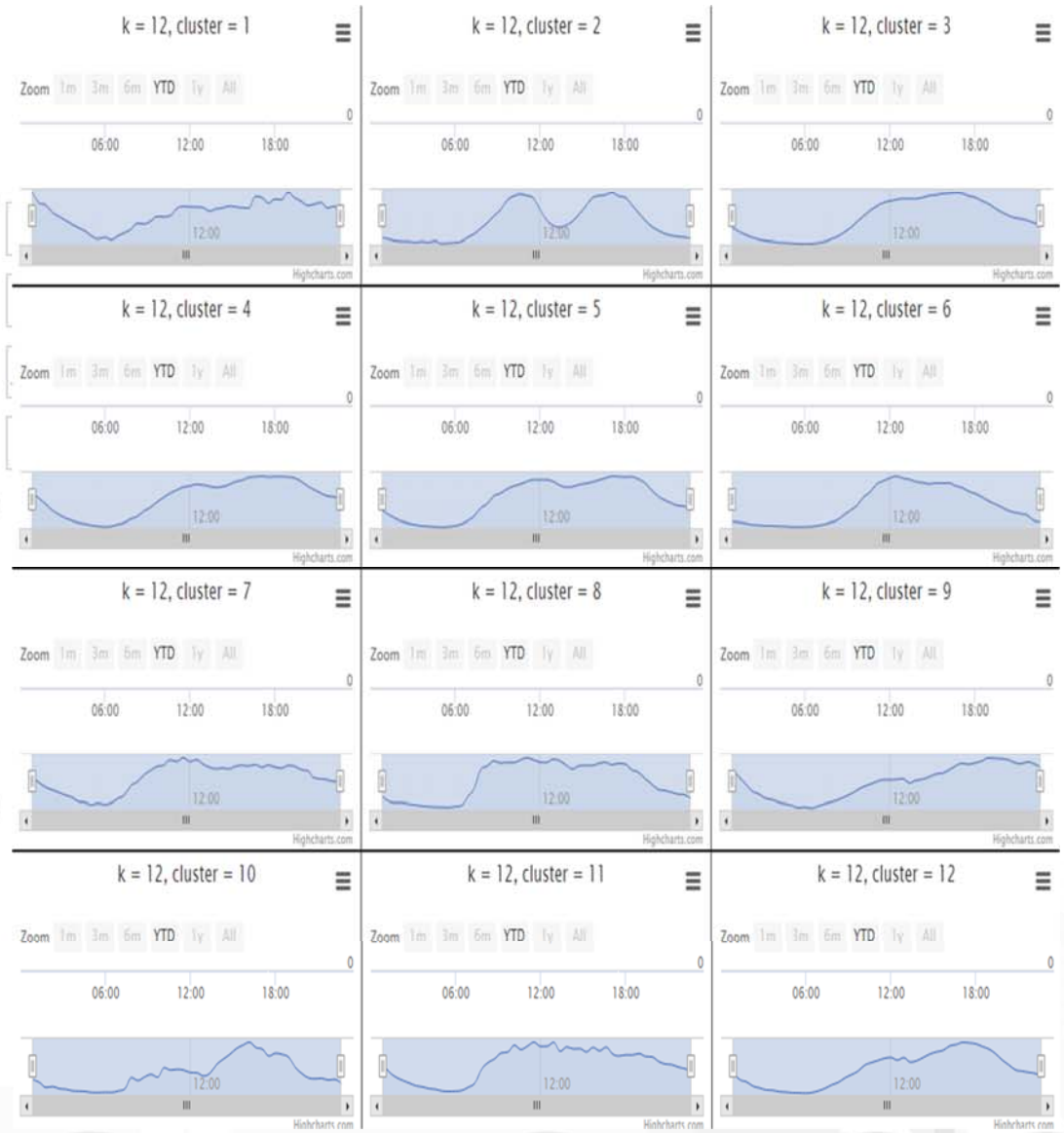
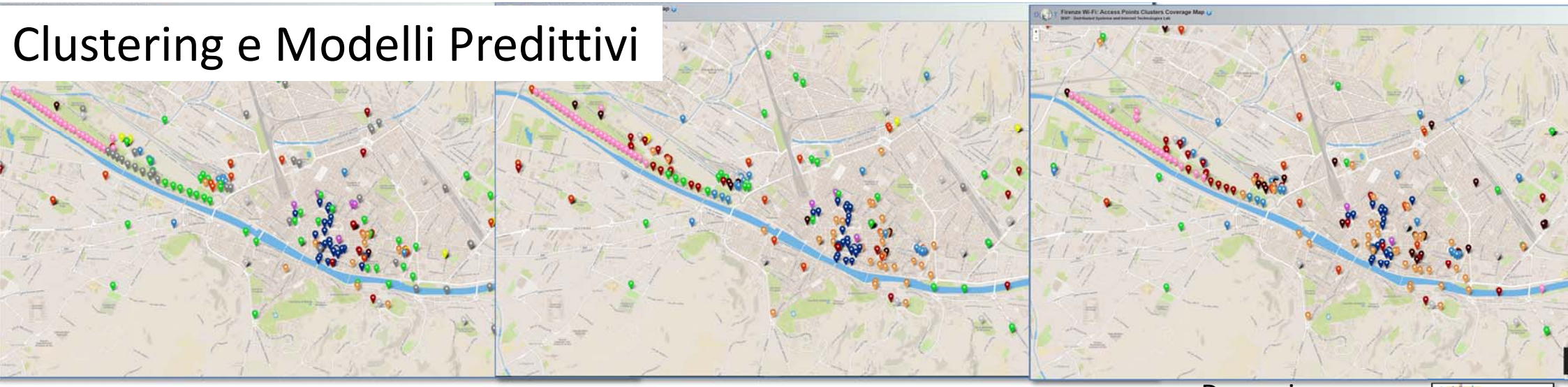


Figure 14 – Average BIC for mixture models vs K number of cluster, higher values are better, the curves are truncated at the best value for K they found.

Clustering e Modelli Predittivi



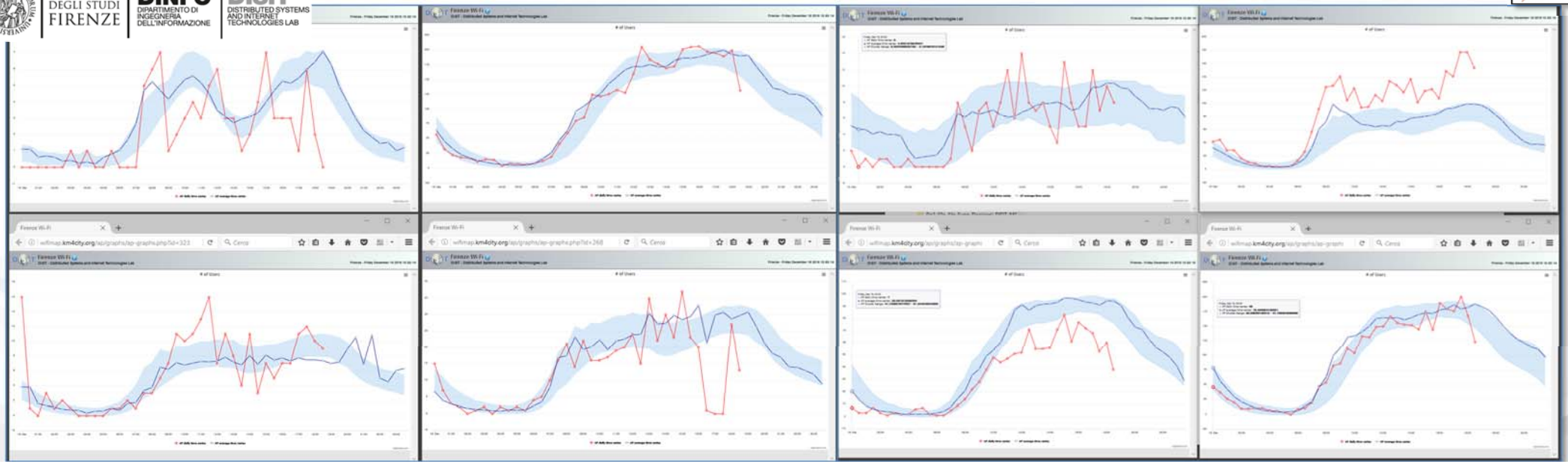
Lunedi-Venerdi

Sabato

Domenica

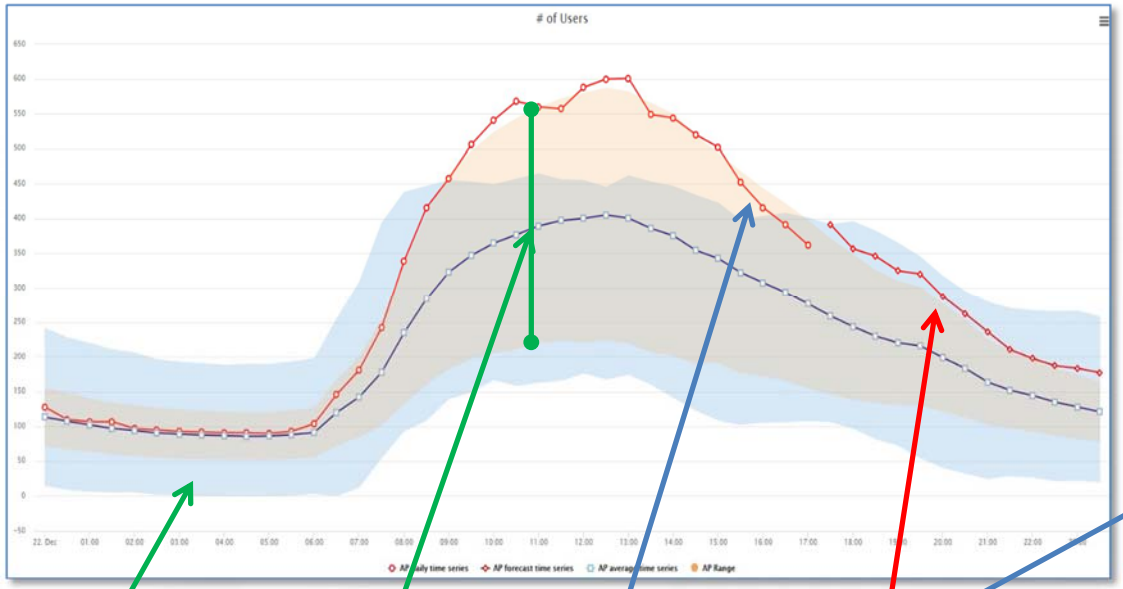
UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB

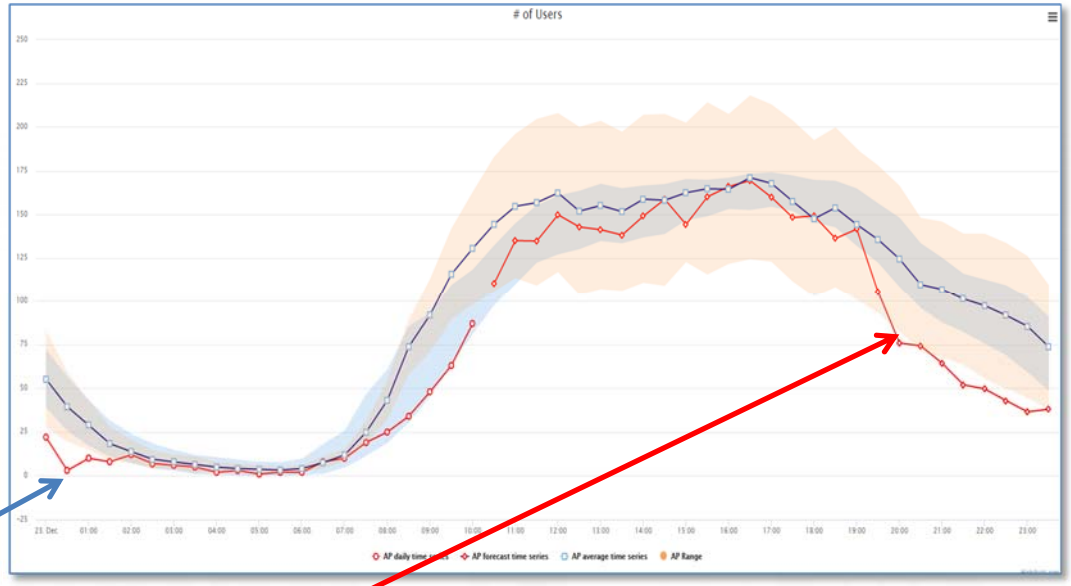
resolute



Predizione e identificazione anomalie

Guessing number of users of Wi-Fi Access Points





Cluster confidence

AP average and confidence

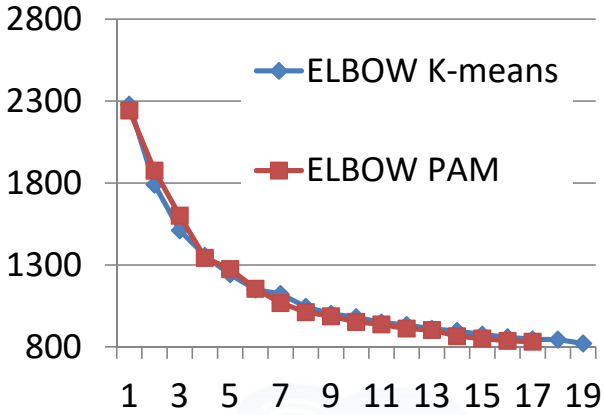
Actual AP trend for today

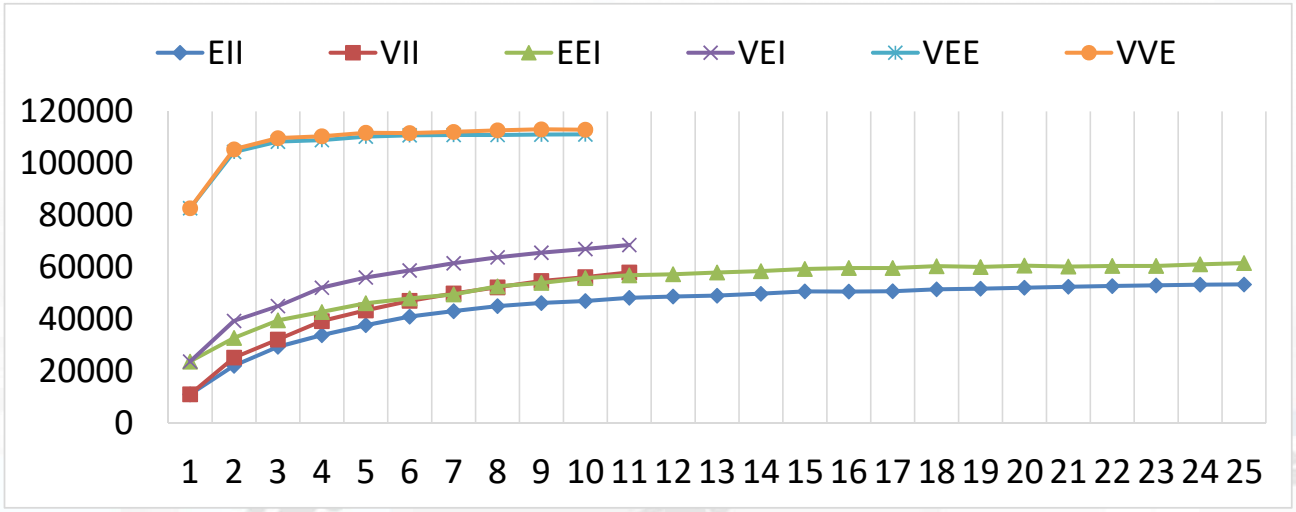
AP prediction for the next time slot in the day on the basis of past weeks

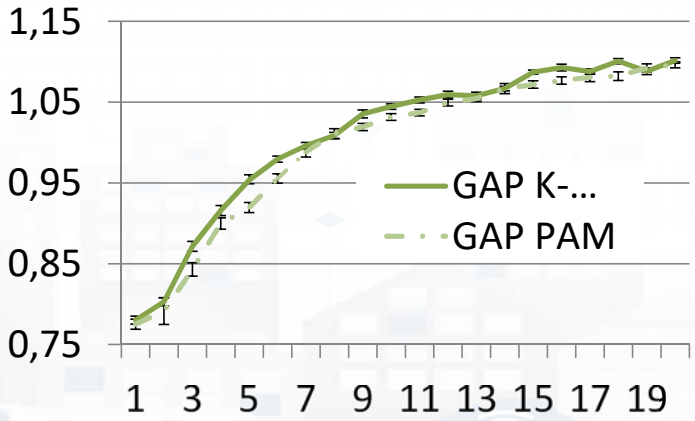
Case Study A

Clustering of City Usage of Wi-Fi

- Approaches: K-Means, K-Medoids (PAM), …..
- Assessment Methods: ELBOW, GAP
- Identification of the most suitable K

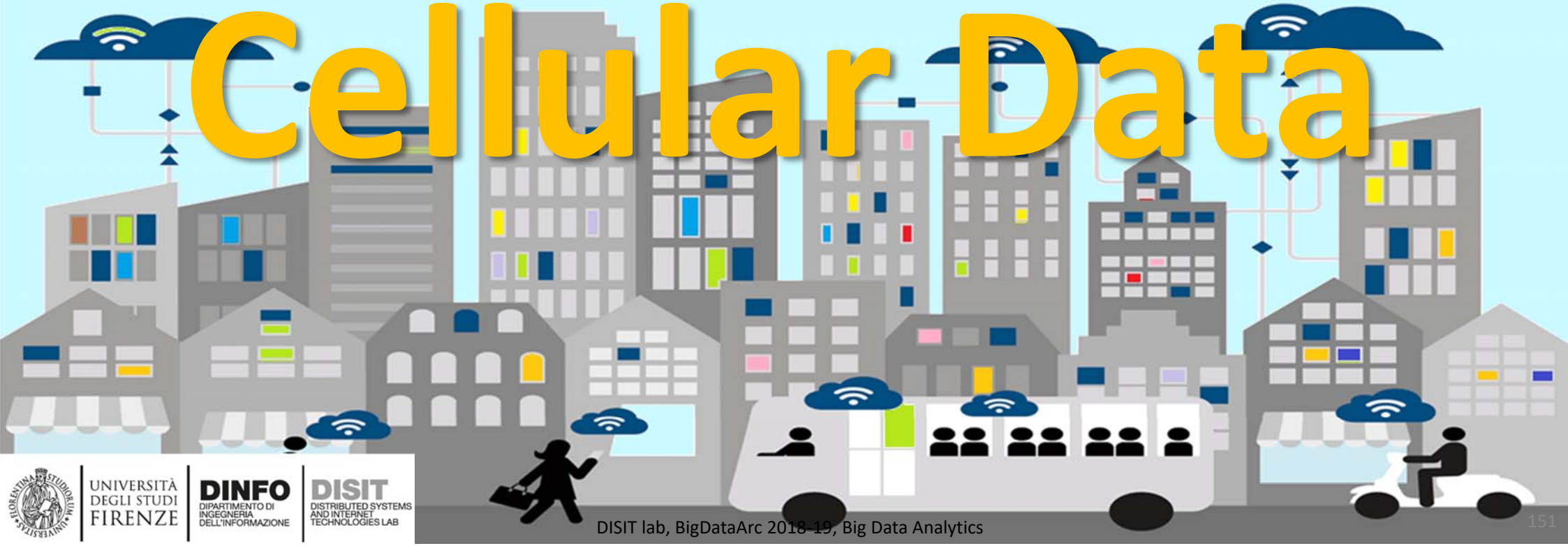






Case Study A

Monitoring via Cellular Data

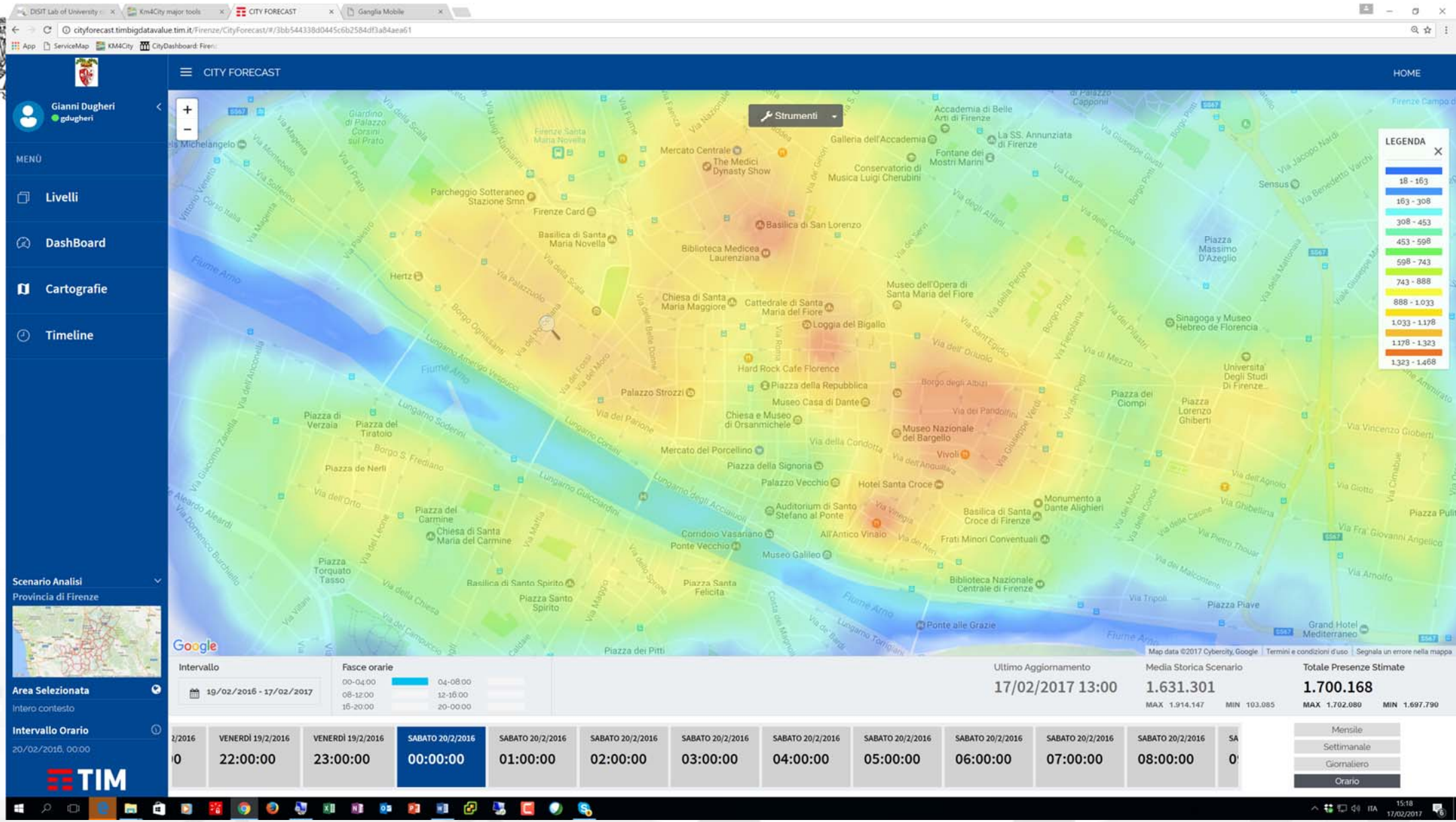


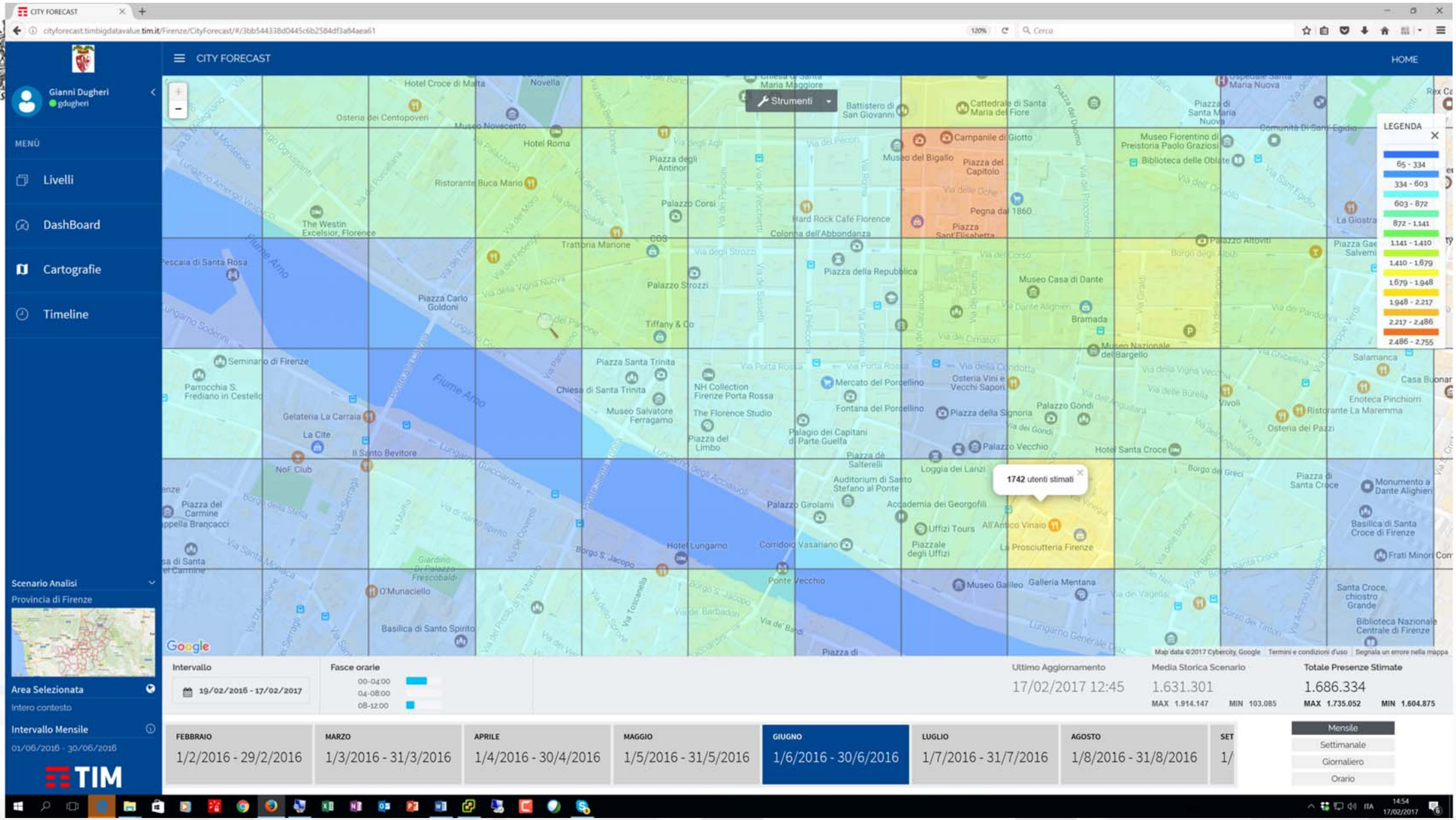
Overview

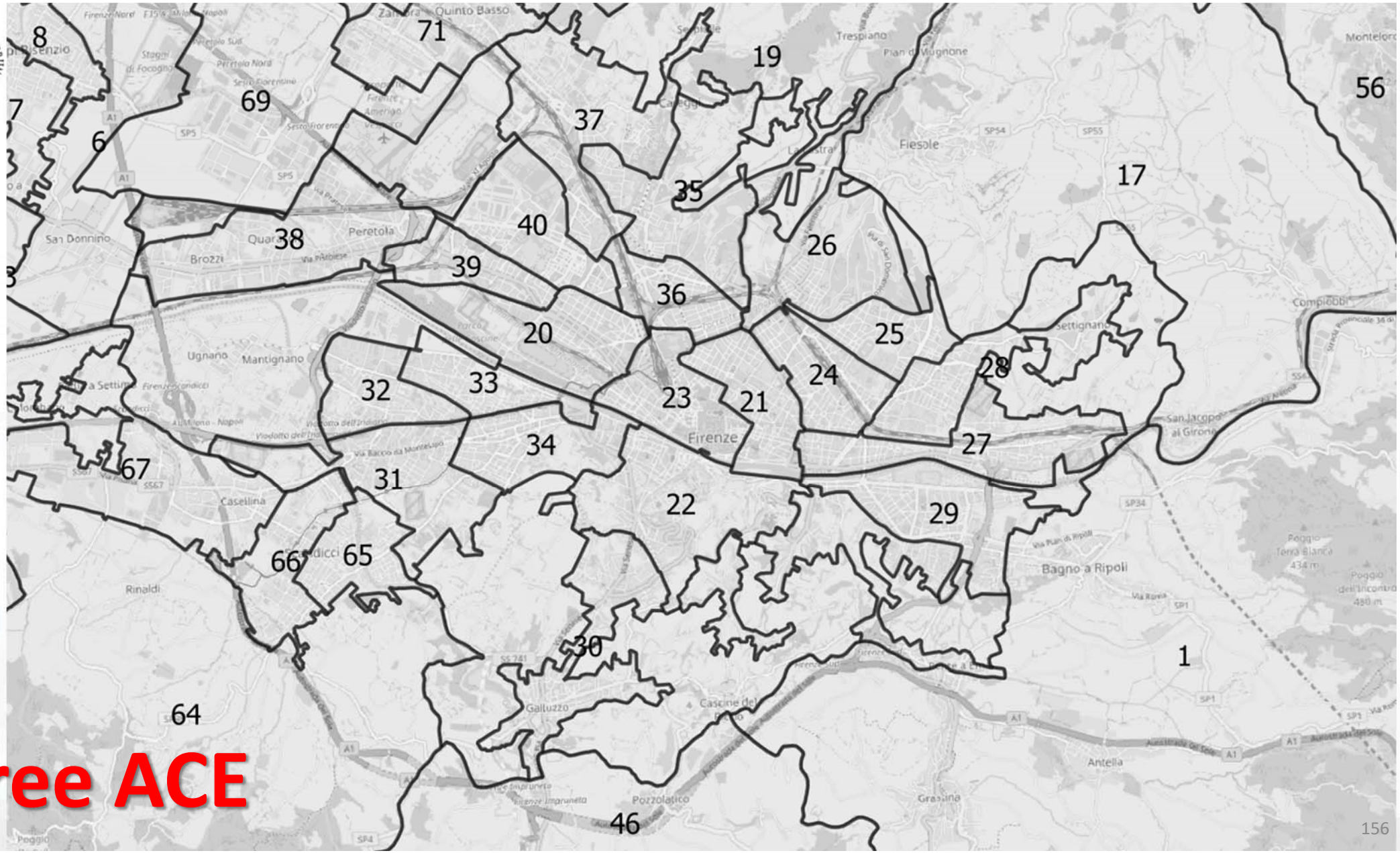
- **Analisi Dati raccolti via Social Media**
  - Predicting presences
- **Analisi Dati raccolti via Mobile App**
  - Tracce, matrici OD, heatmap
  - Regency e frequency
  - Impatto in uscita da Firenze
- **Analisi Dati raccolti dai Flussi di traffico**
  - Ricostruzione del traffico in punti non misurati
- **Analisi Dati raccolti dai Parcheggi**
  - Predizione dei posti liberi
- **Analisi Dati raccolti via Firenze WiFi**
  - Tracce, matrici OD, heatmap
  - Predicting presences
- **Analisi Dati Dati raccolti via Cellular**
  - Valutazione comparativa TIM-VODA
  - Valutazione comparativa FirenzeWiFi-Tim-Voda

Letture possibili

- Andamenti giornalieri
- Andamenti settimanali
- Matrici OD (Voda)
- Heatmap di TIM (250 metri) non paiono precise
- Differenze sulle fasce orarie
- Differenze sul calcolo del numero delle presenze, mediate
- …..

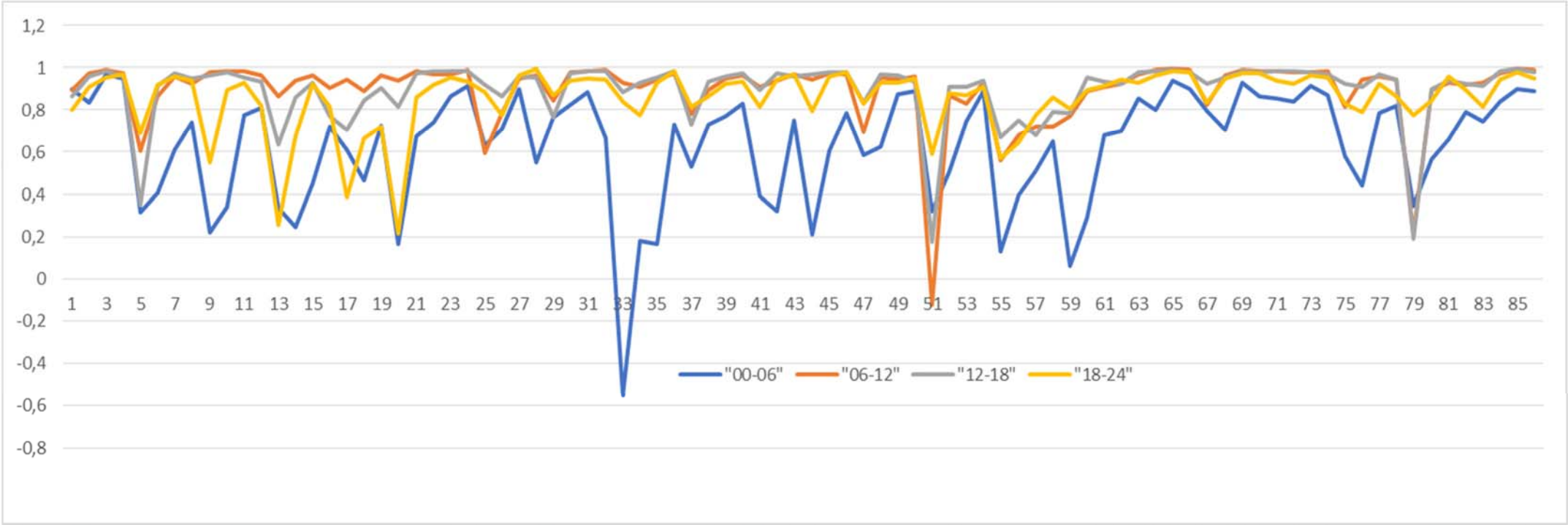

CITY FORECAST
HOME
Gianni Dugheri
gdugheri
MENÙ
Livelli
DashBoard
Cartografie
Timeline
Strumenti
LEGENDA
18 - 163
163 - 308
308 - 453
453 - 598
598 - 743
743 - 888
888 - 1.033
1.033 - 1.178
1.178 - 1.323
1.323 - 1.468
Scenario Analisi
Provincia di Firenze
Area Selezionata
Intero contesto
Intervallo Orario
20/02/2016, 00:00
Intervallo
19/02/2016 - 17/02/2017
Fasce orarie
00-04:00
04-08:00
08-12:00
12-16:00
16-20:00
20-00:00
Ultimo Aggiornamento
17/02/2017 13:00
Media Storica Scenario
1.631.301
MAX 1.914.147
MIN 103.085
Totale Presenze Stimate
1.700.168
MAX 1.702.080
MIN 1.697.790
VENERDÌ 19/2/2016 22:00:00
VENERDÌ 19/2/2016 23:00:00
SABATO 20/2/2016 00:00:00
SABATO 20/2/2016 01:00:00
SABATO 20/2/2016 02:00:00
SABATO 20/2/2016 03:00:00
SABATO 20/2/2016 04:00:00
SABATO 20/2/2016 05:00:00
SABATO 20/2/2016 06:00:00
SABATO 20/2/2016 07:00:00
SABATO 20/2/2016 08:00:00
Mensile
Settimanale
Giornaliero
Orario
TIM


CITY FORECAST
HOME
Gianni Dugheri
gdugheri
MENÙ
Livelli
DashBoard
Cartografie
Timeline
Strumenti
LEGENDA
65 - 334
334 - 603
603 - 872
872 - 1.141
1.141 - 1.410
1.410 - 1.679
1.679 - 1.948
1.948 - 2.217
2.217 - 2.486
2.486 - 2.755
1742 utenti stimati
Scenario Analisi
Provincia di Firenze
Area Selezionata
Intero contesto
Intervallo Mensile
01/06/2016 - 30/06/2016
Intervallo
19/02/2016 - 17/02/2017
Fasce orarie
00-04:00
04-08:00
08-12:00
Ultimo Aggiornamento
17/02/2017 12:45
Media Storica Scenario
1.631.301
MAX 1.914.147
MIN 103.085
Totale Presenze Stimate
1.686.334
MAX 1.735.052
MIN 1.604.875
FEBBRAIO
1/2/2016 - 29/2/2016
MARZO
1/3/2016 - 31/3/2016
APRILE
1/4/2016 - 30/4/2016
MAGGIO
1/5/2016 - 31/5/2016
GIUGNO
1/6/2016 - 30/6/2016
LUGLIO
1/7/2016 - 31/7/2016
AGOSTO
1/8/2016 - 31/8/2016
SET
Mensile
Settimanale
Giornaliero
Orario
TIM


8
71
19
56
69
37
17
6
35
40
38
26
39
36
20
25
24
28
32
33
23
21
27
34
67
31
22
29
66
65
30
1
64
46
Firenze
Fiesole
Bagno a Ripoli
Antella
Peretola
Brozzi
Ugnano
Mantignano
Casellina
Rinaldi
Galluzzo
Pozzolatico
Grassina
Settignano
Compiobbi
San Jacopo al Girone

Aree ACE

Confronto TIM-Vodafone

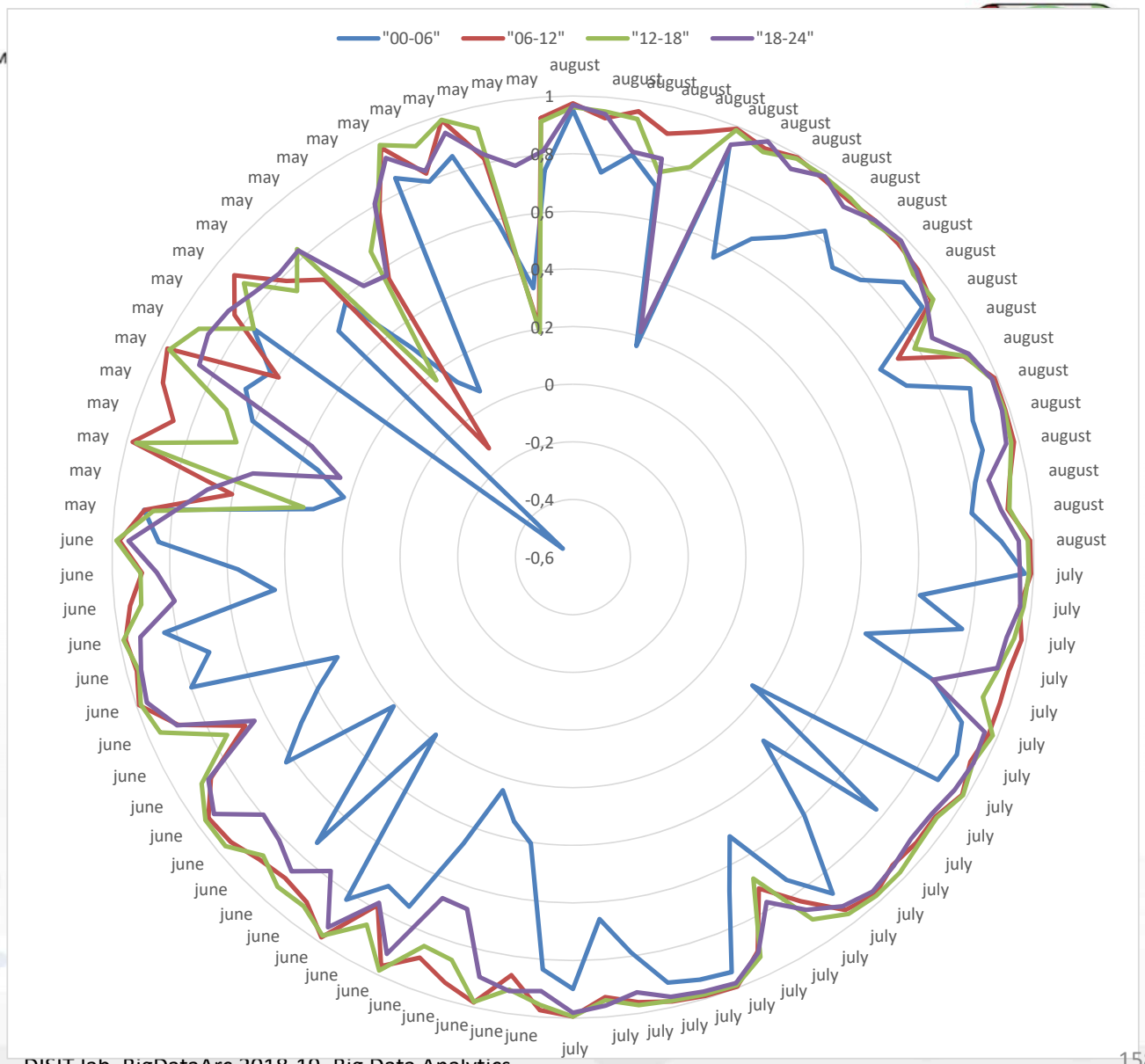


Correlazione	"00-06"	"06-12"	"12-18"	"18-24"
media	0,6350	0,8948	0,8965	0,8576
var	0,0720	0,0271	0,0227	0,0220
mediana	0,7214	0,9471	0,9515	0,9182

-Su tutte le aree ACE
- Probabili differenze su residenti !

Confronto TIM-Vodafone

- Andamento della correlazione nel tempo
- Varie fasce orarie



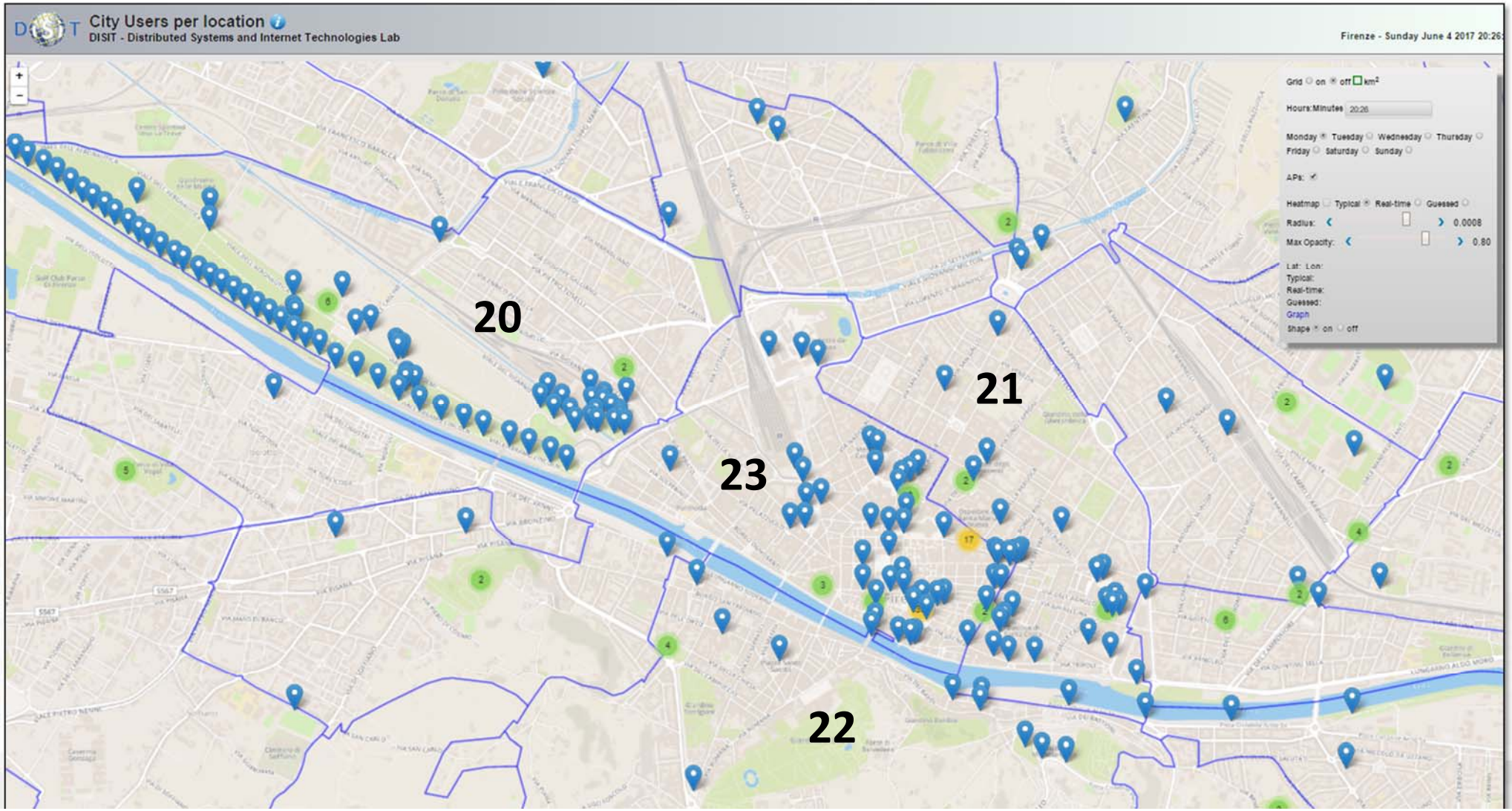
Confronto TIM-Vodafone

- Volumi del numero delle persone per ogni fascia oraria sovrapponibile e per ogni mese
  - Maggio, Giugno, Luglio, Agosto
  - Settembre non e' completo

	Vodafone	tim	Voda/Tim
Media, Numero utenti (non distinti) mese	8.183.017	15.036.991	0,582
varianza	2,27549E+12	4,83645E+12	0,009

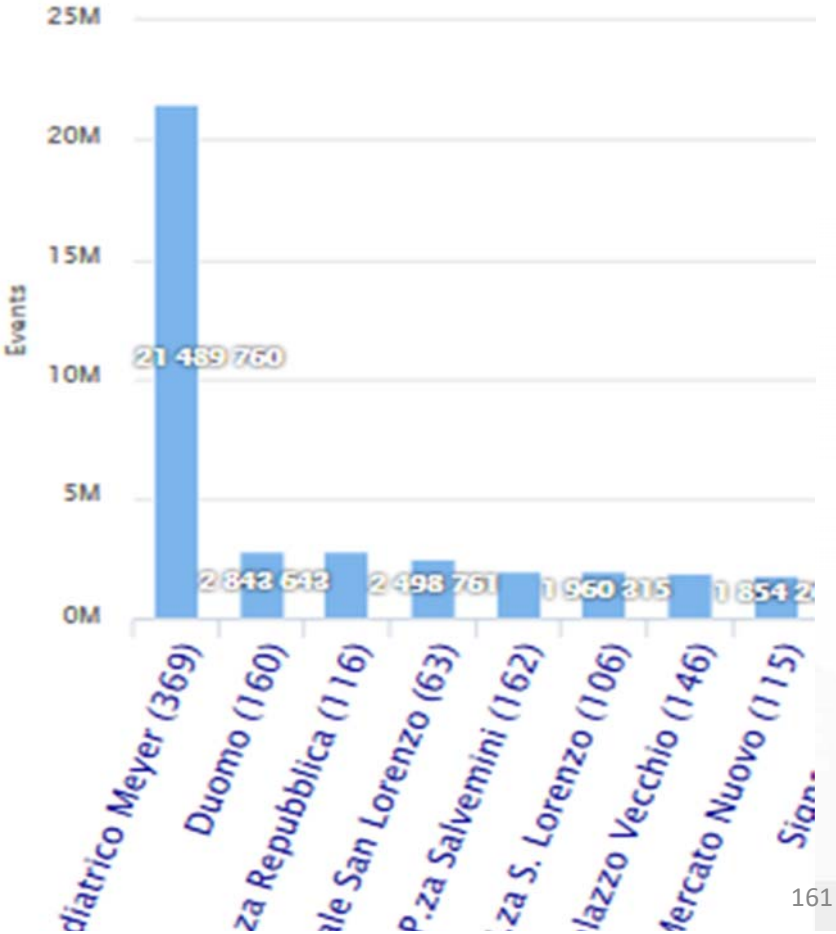
UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB
http://www.disit.org

Firenze Wi-Fi

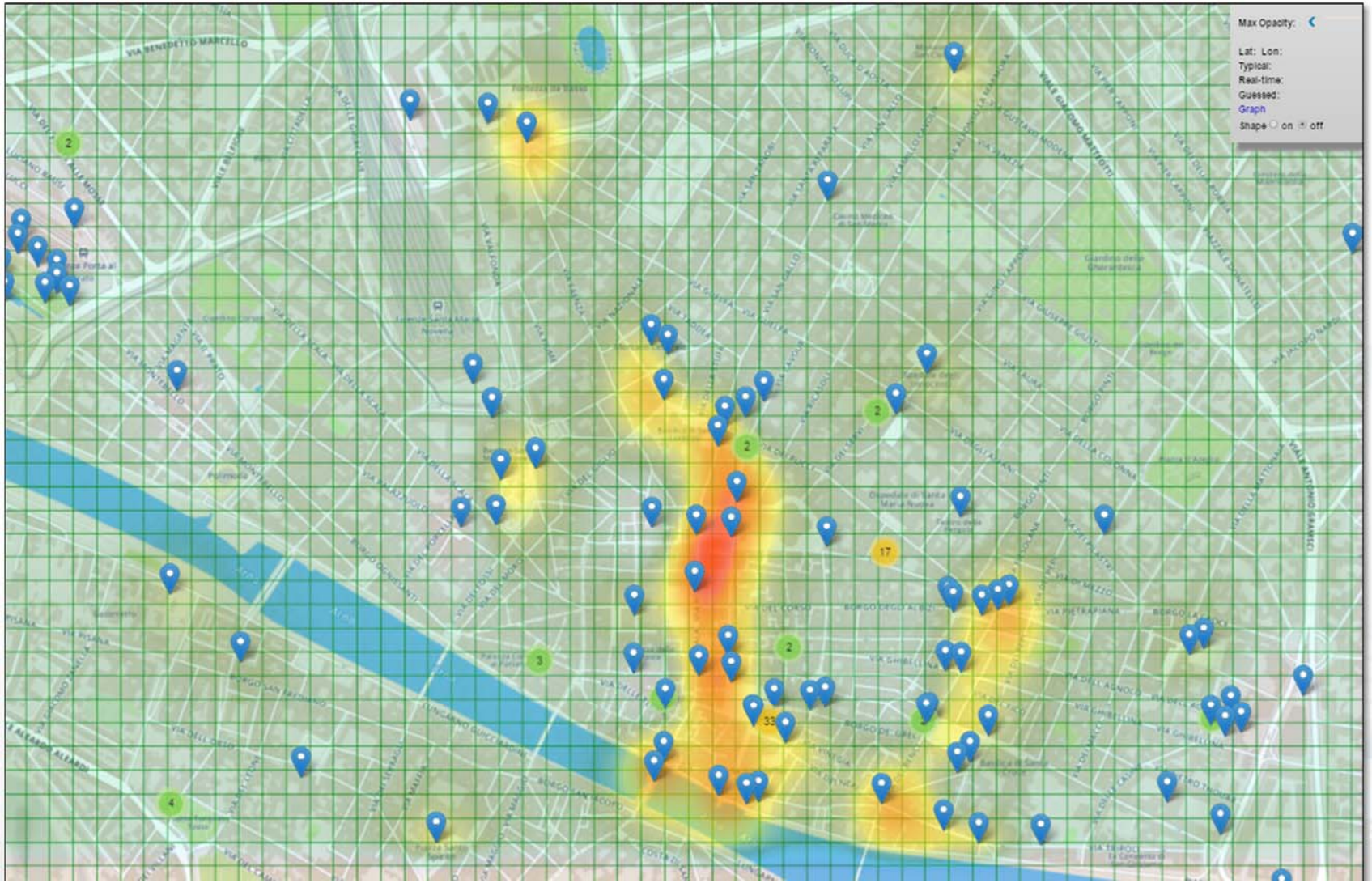


Firenze - Sunday June 4 2017 23:43:57

- Distinct APs: 368
- Distinct APs (last 24 hours): 288
- Distinct Users (last 180 days): 1009929
- Distinct Users (last 180 days, < 24 h): 563200

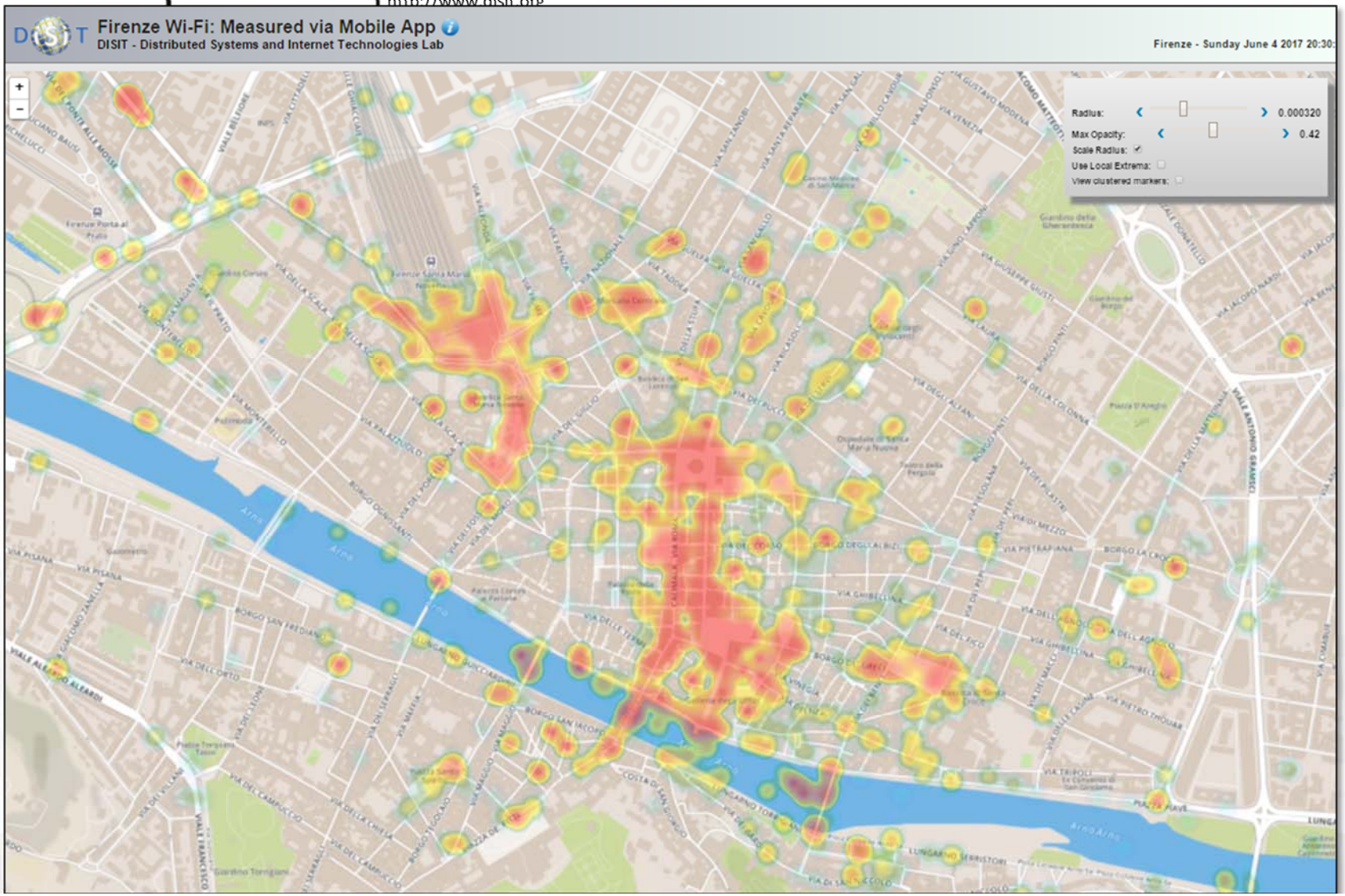


Firenze Wi-Fi



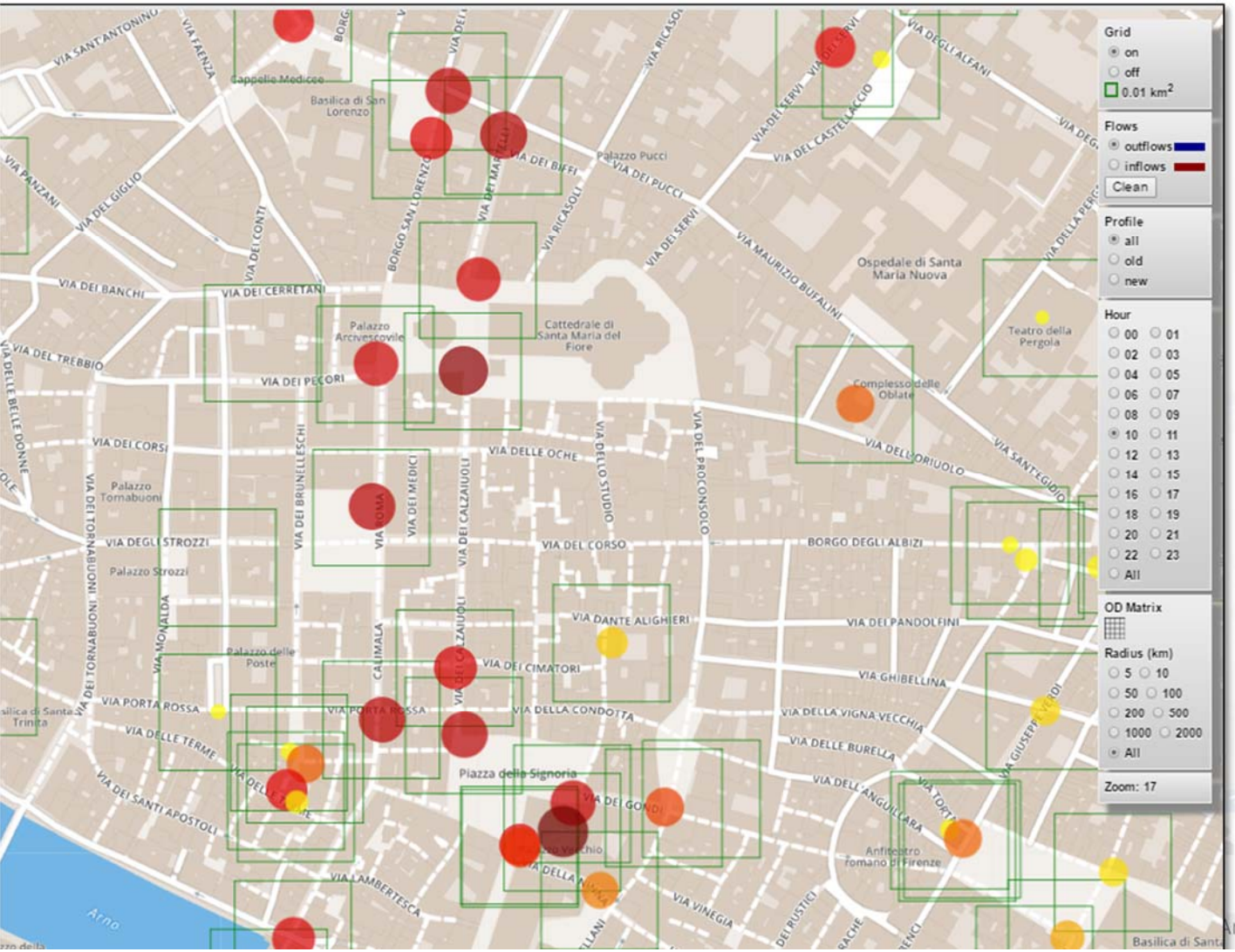
UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB
http://www.disit.org

Mobile App, Km4City



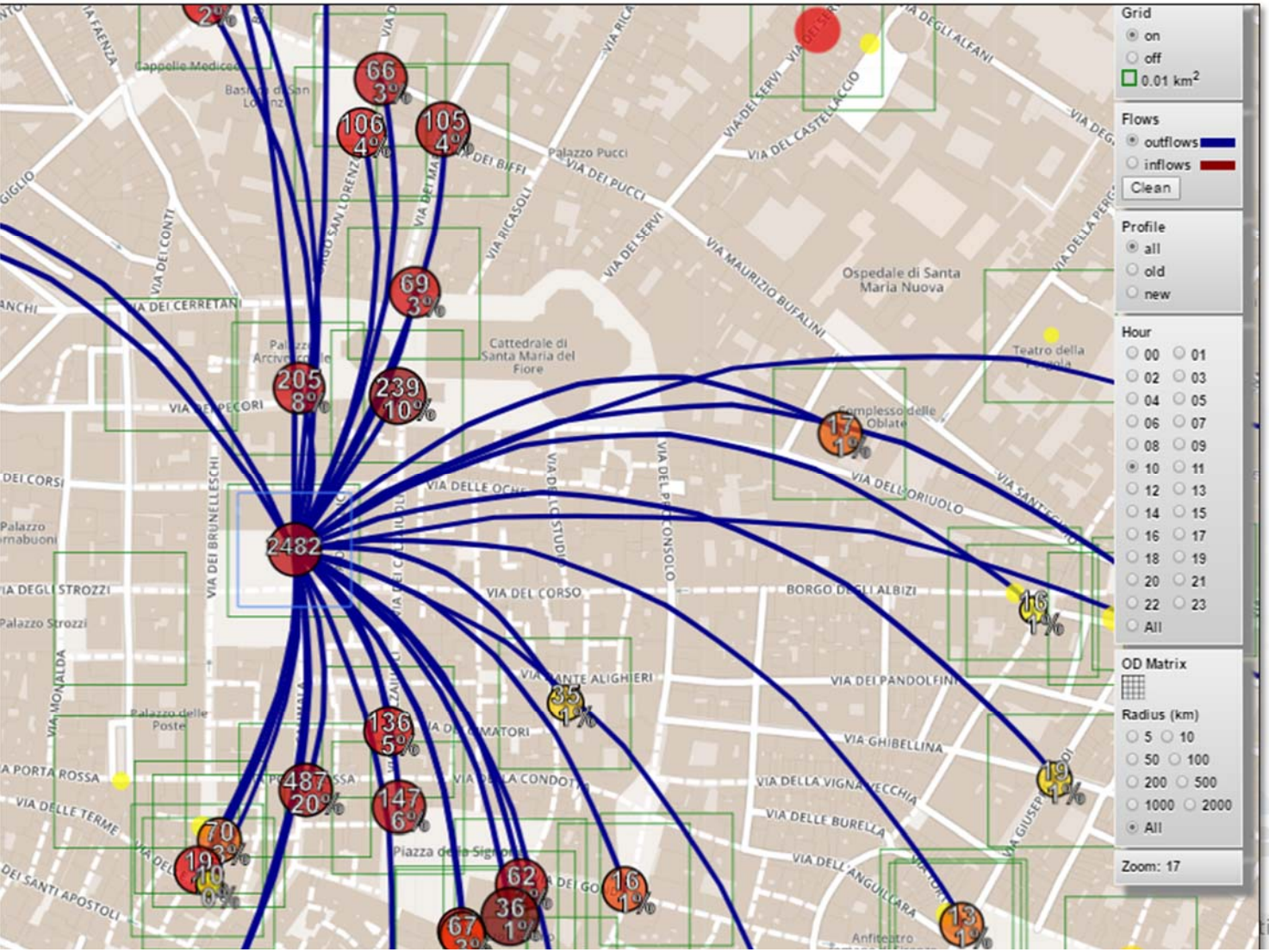
UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB
http://www.disit.org

Firenze Wi-Fi



- AP → heatmap sparsa
- Inflow/outflow
- New/Old users
- per fascia oraria

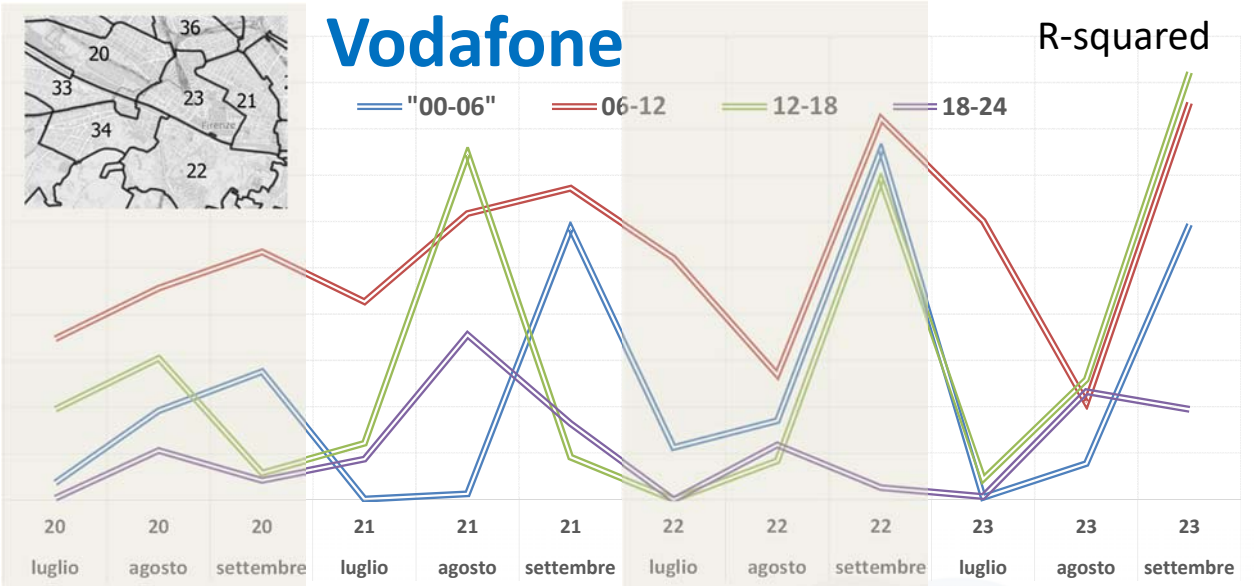
Firenze Wi-Fi

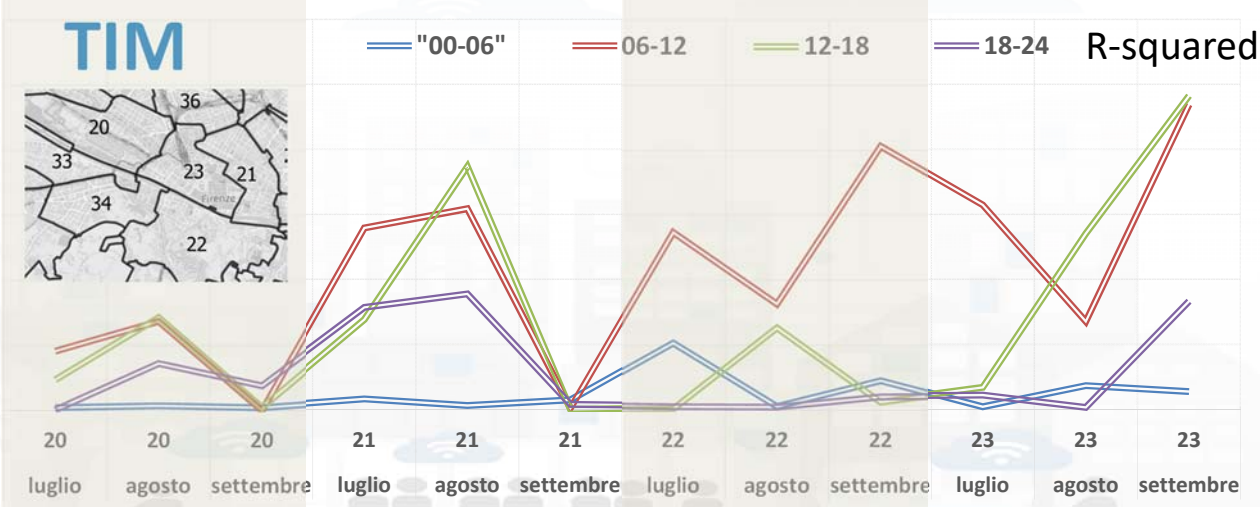


- AP → matrice OD sparsa
- Inflow/outflow
- New/Old users
- Flussi per fascia oraria

Firenze Wi-Fi vs cell

- Correlazione nella fascia 6-12
- Scarsa correlazione se vi sono pochi AP e pochi dati, oltre 350.000 eventi si ha una correlazione elevata
  - Settembre
- In 22 vi sono pochi AP
- Scarsa correlazione per fasce 18-24, 0-6 dove l'incidenza di residenti e' elevata sui cellulari
- Vodafone presenta una maggior correlazione per la presenza di un mino numero di residenti





considerazioni

- Dati telefonici:
  - Catturano sia residenti che turisti senza una forte distinzione
  - Cons: Non permettano di tracciare matrici origine destinazione
  - Cons: Sono ad una risoluzione troppo bassa: 15 minuti → 6 ore
  - Pros: valutano anche fuori dall'area urbana
- Dati Firenze WiFi:
  - Catturano principalmente i turisti e movimenti in strada
  - Permettono di fare matrici OD solo se la rete e' ben instrumentata
  - Forte correlazione con dati delle reti cellulari negli orari centrali
  - Cons: non lavorano fuori dall'area ubana

Other Aspects

User behavior on GDO and Retail

ARC
INDEX

Feedback Project

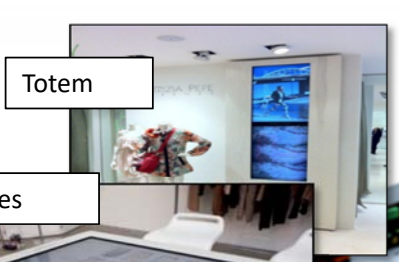
Personal Assistant

Operators

Strategy Editor and feedback

Rule name	Type	#sent	#viewed	#viewed on #sent	Description
daily_event_de	ENGAGEMENT	1 (0%)	0 (0%)	0%	Suggest (in german) an event currently on in Flore
daily_event_en	ENGAGEMENT	1720 (2.12%)	70 (7.1%)	4.07%	Suggest (in english) an event currently on in Florence
	- commuter	5 (0.29%)	0 (0%)	0 (0%)	
	- student	14 (0.81%)	0 (0%)	0 (0%)	
	- tourist	1462 (85%)	25 (35.71%)	25 (1.71%)	
	- citizen	113 (6.57%)	39 (55.71%)	39 (34.51%)	
	- operator	0 (0%)	0 (0%)	0 (0%)	
	- disabled	0 (0%)	0 (0%)	0 (0%)	
	- all	119 (6.92%)	6 (8.57%)	6 (5.04%)	
daily_event_es	ENGAGEMENT	6 (0.01%)	4 (0.41%)	66.67%	Suggest (in spanish) an event currently on in Florence
daily_event_fr	ENGAGEMENT	6 (0.01%)	0 (0%)	0%	Suggest (in french) an event currently on in Florence
daily_event_it	ENGAGEMENT	5459 (6.73%)	296 (30.02%)	5.42%	Suggest (in italian) an event currently on in Florence
parking_en	ASSISTANCE	141 (0.17%)	128 (12.98%)	90.78%	Alert (in english) if the user parked in a residential parking zone
parkit					Perked in a residential parking zone
parkit					arked in a residential parking zone
shoot					on for a nearby point-of-interest

Inform
Engage
Stimulate / recommend
Anomalies Detection
Provide Bonus, incentives



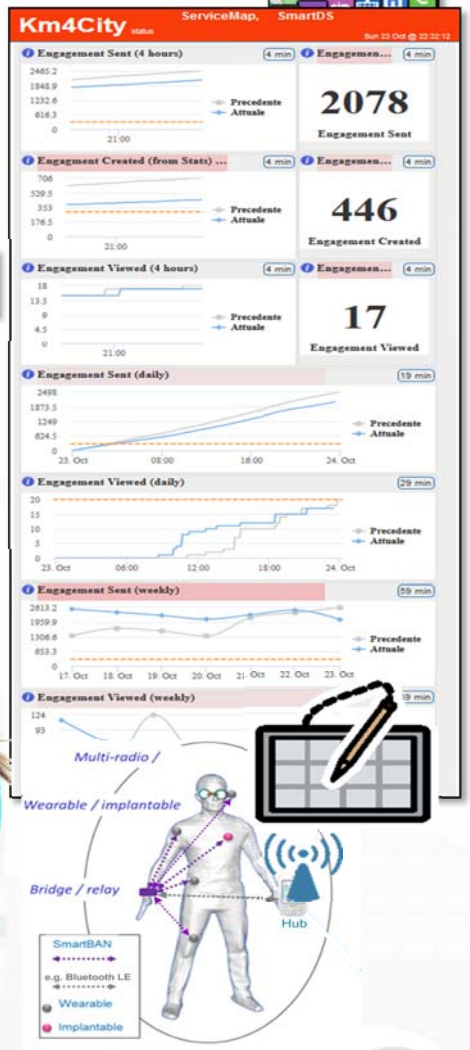
Sensor Manager

Network Manager

Bluetooth ..... Wi Fi

Sensors and Markers

IOT/IOE



Demand and Offer: Origin Destination Matrices

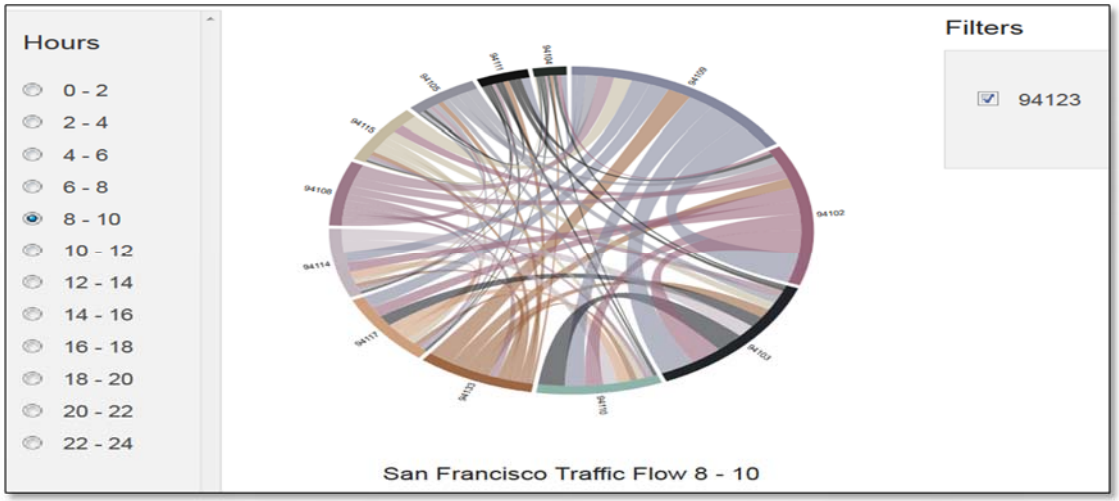
ARC
INDEX

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

Traffic and People Flow Assessment

http://www.disit.org/6694

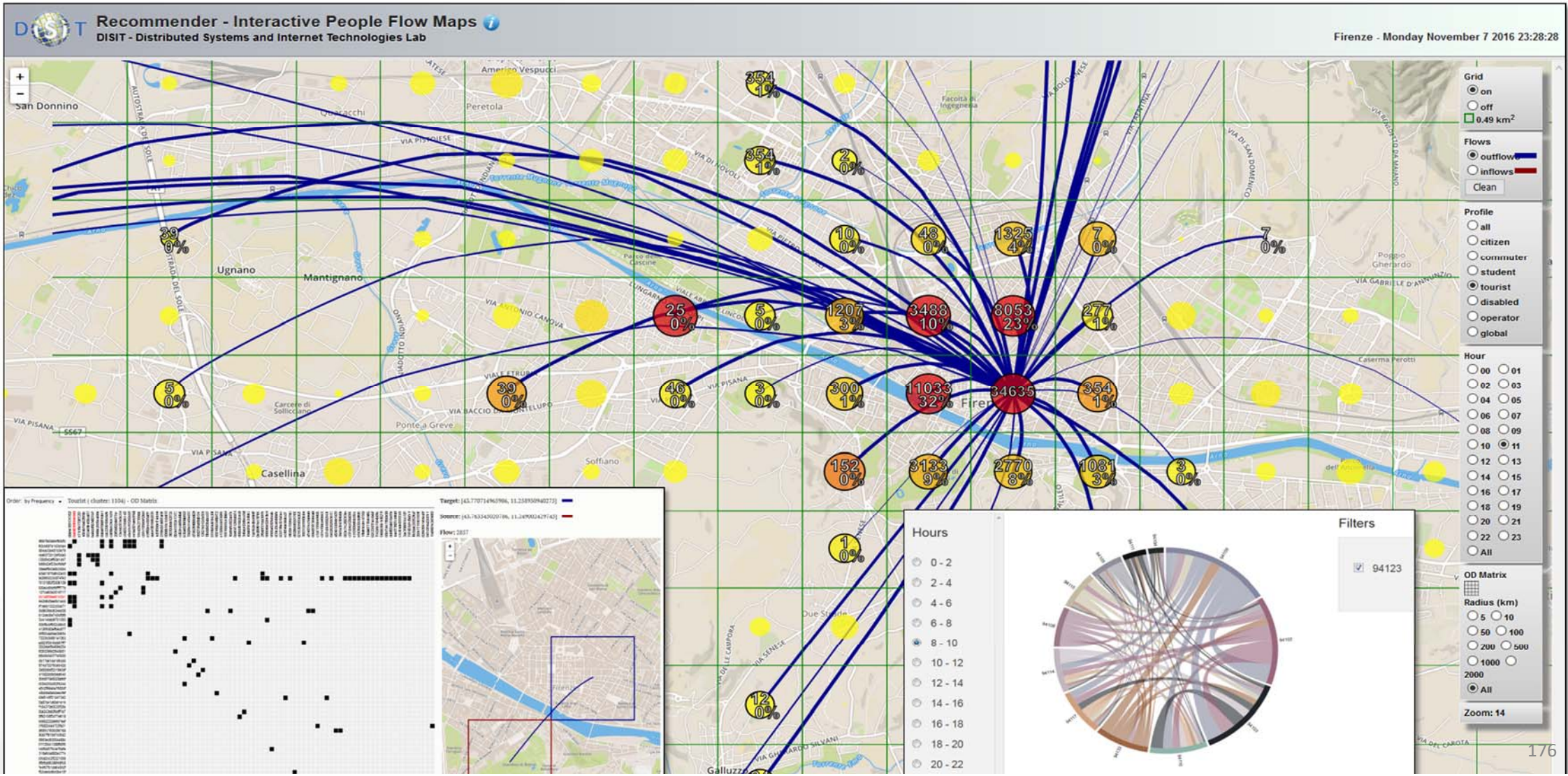
- **Origin Destination Matrix**
  - Specific Sensors, vehicle Kits, mobile App, Wi-Fi Access Points, etc.
- **Assess people and traffic flows to**
  - improve services
  - predict critical conditions on Crit. Infra.
  - take real time decisions and sending messages in push to population
  - Increase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



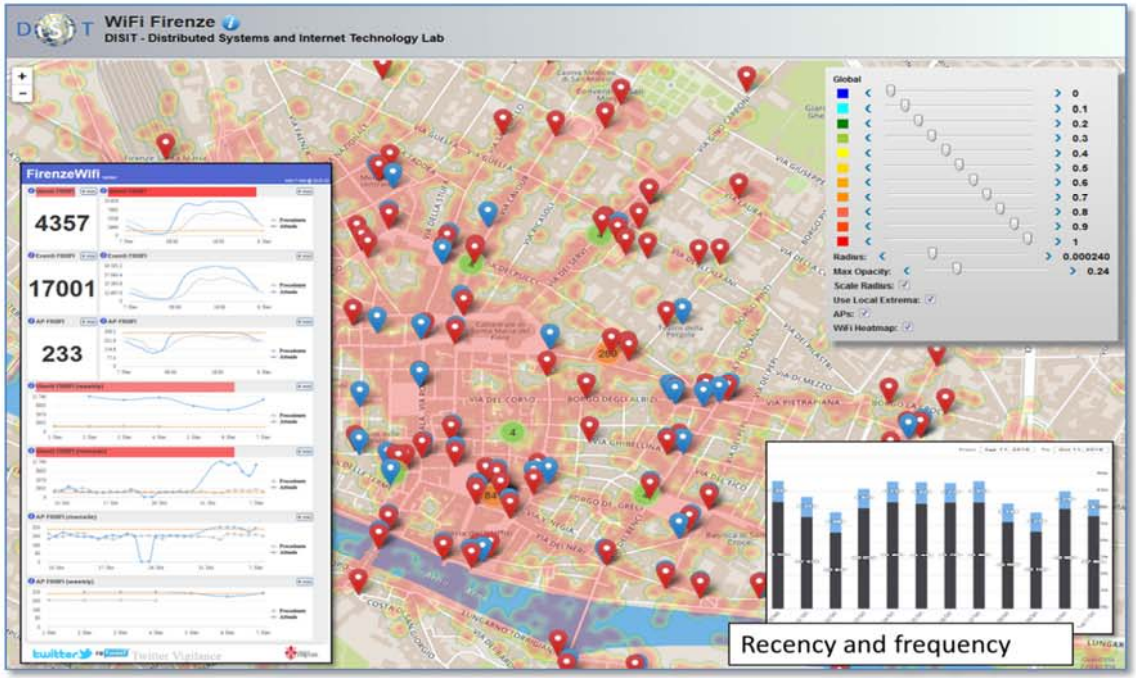
UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB

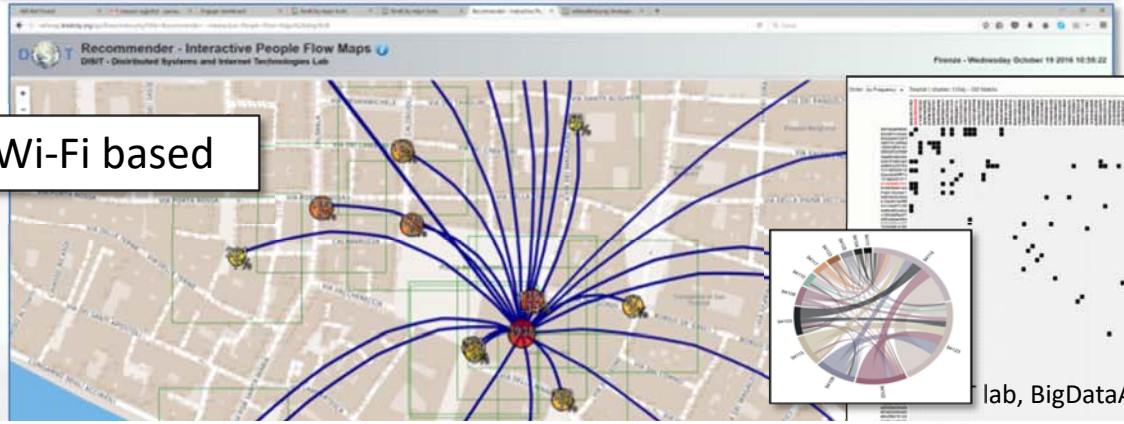
OD Matrix scalabile

Sii-Mobility

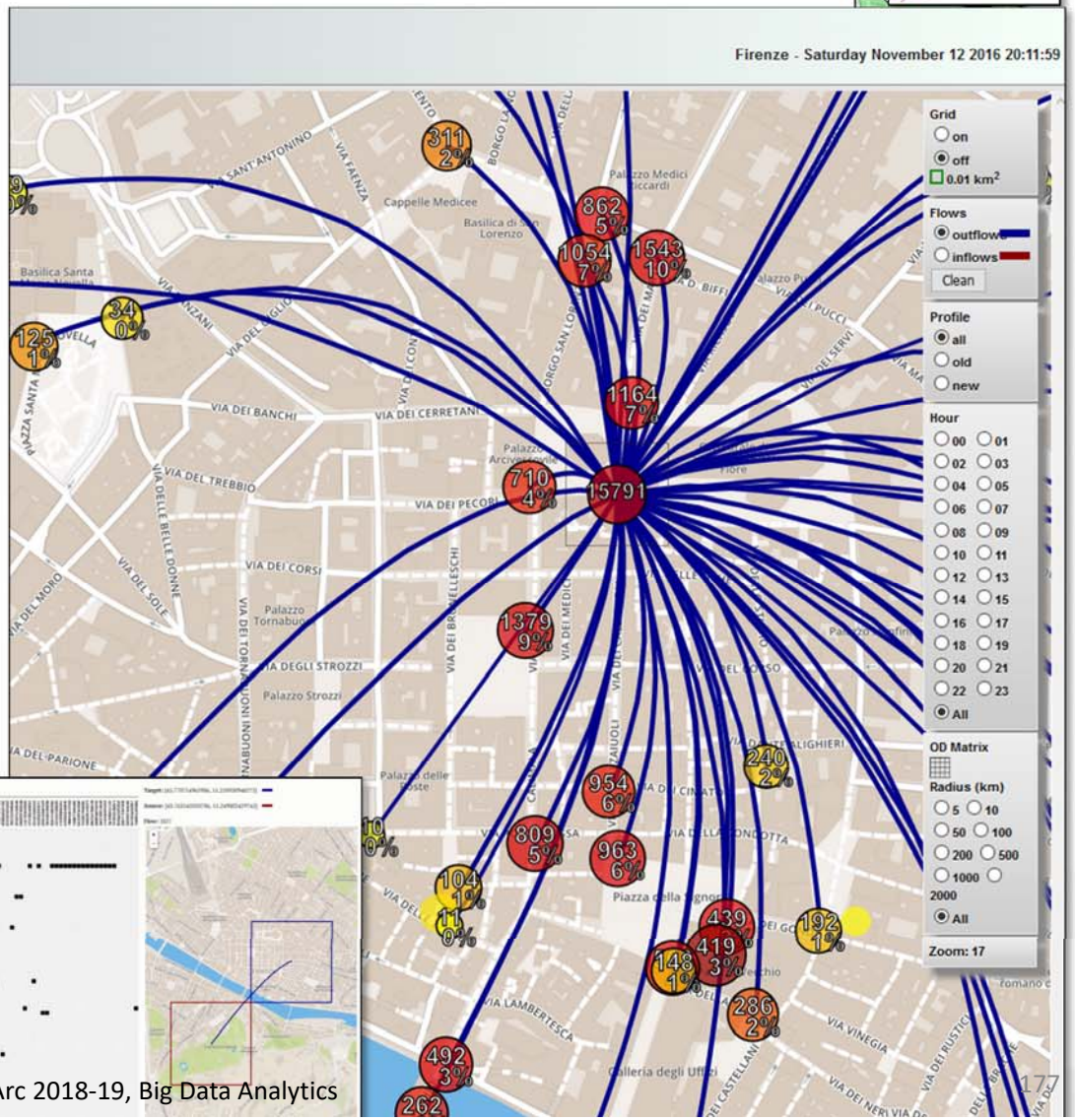


Origin Destination Matrix Estimation





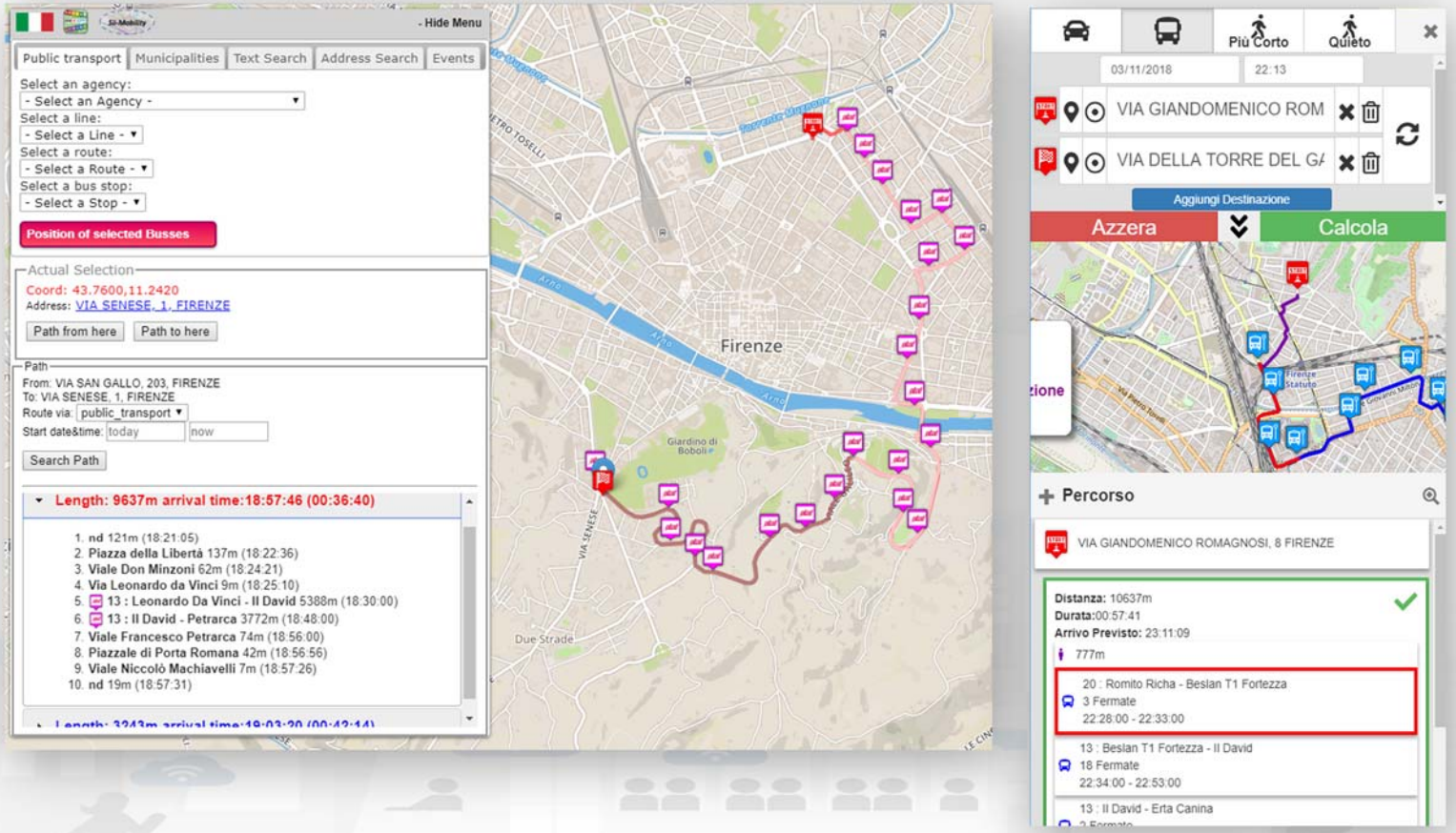
Wi-Fi based



Modal & Multimodal Routing for Navigation and travel planning

Routing and Multimodal Routing

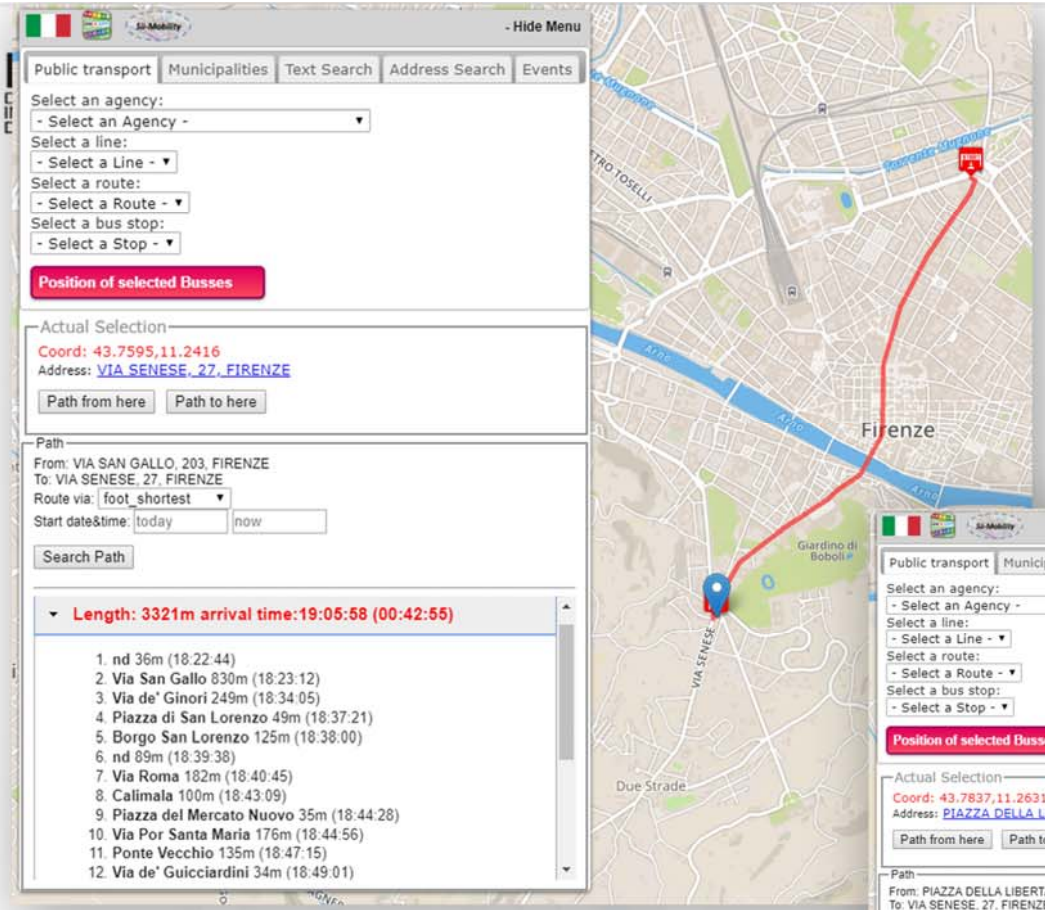
- **Modes**
  - Pedonal, Veichles
  - Public Multimodal
  - Multi Point for Delivering
  - Constrained
- **Test it on our**
  - Mobile Apps
  - MicroApplication
  - Dashboard
  - ServiceMap service on Tuscany in Snap4City

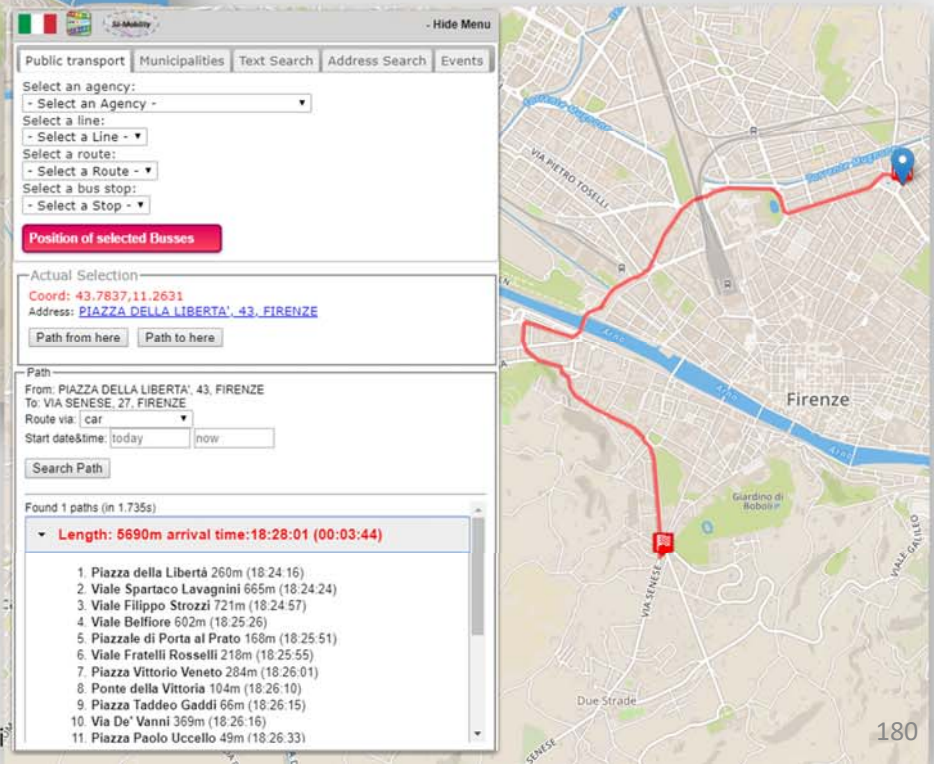


UNIVERSITÀ DEGLI STUDI FIRENZE

Routing

- Pedonal
  - Normal
  - Quite





Problemi aperti per le soluzioni constrained

- Ponti
- Sottopassi
- Passerelle
- Allagamenti
- Riduzione di carreggiata
- Barriere puntiformi e ad shape

Anomaly Detection Early Warning

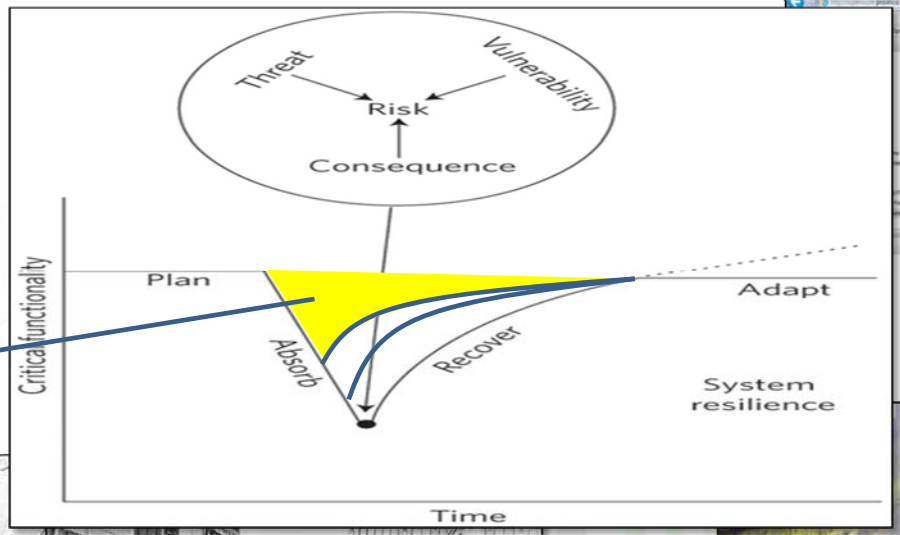
ARC
INDEX

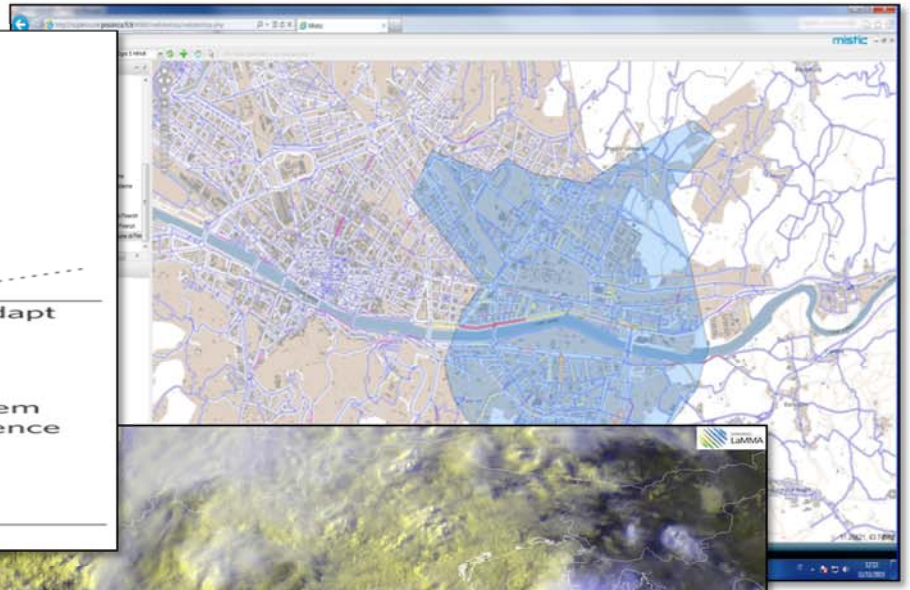
Resilience

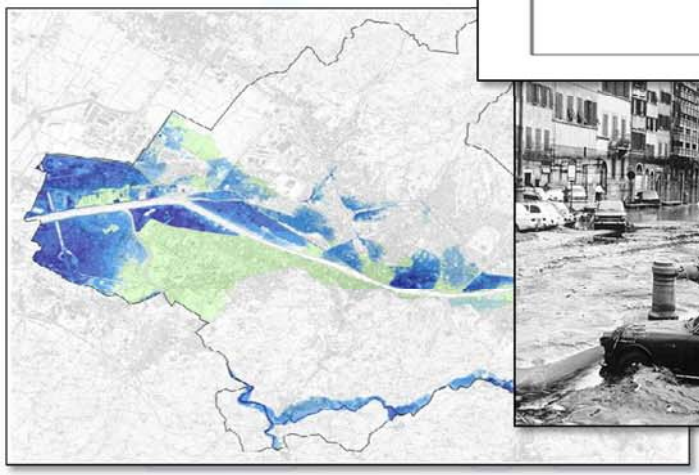
Early warning, detection

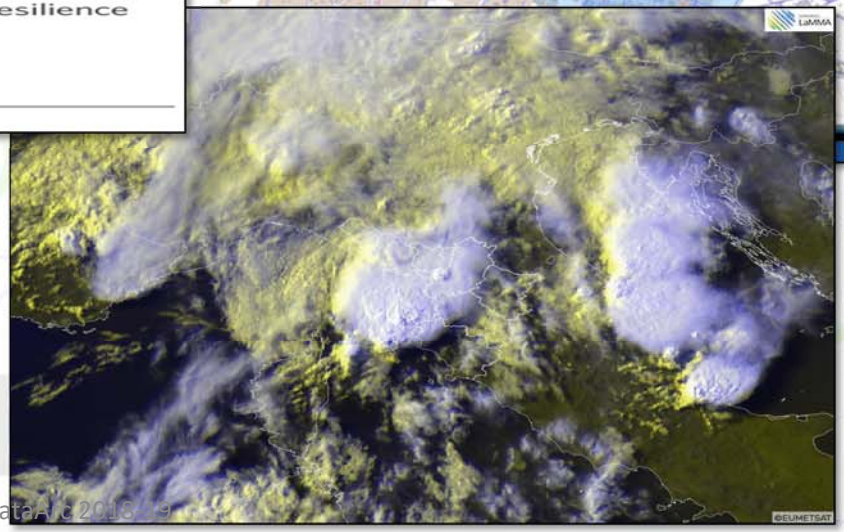
Prepare
Absorb
Recover
Adapt

damage



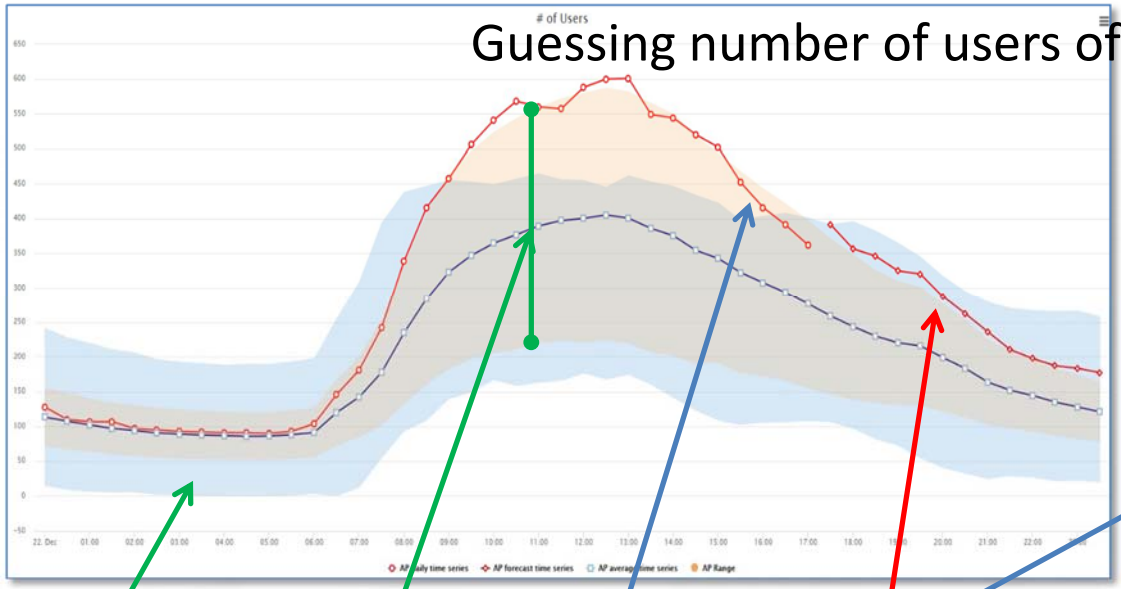


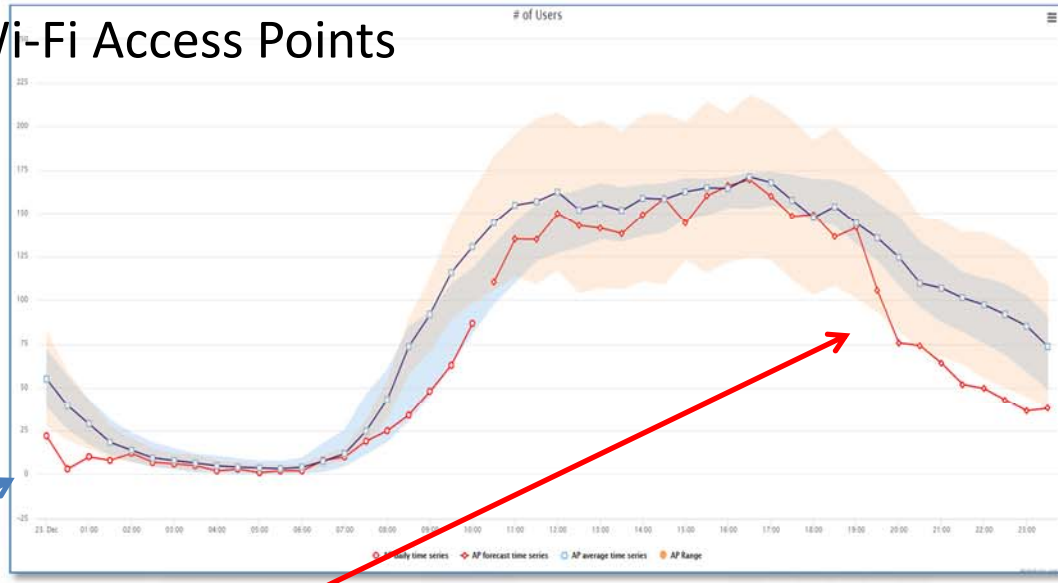




Prediction and identification of anomalies

Guessing number of users of Wi-Fi Access Points





Cluster confidence

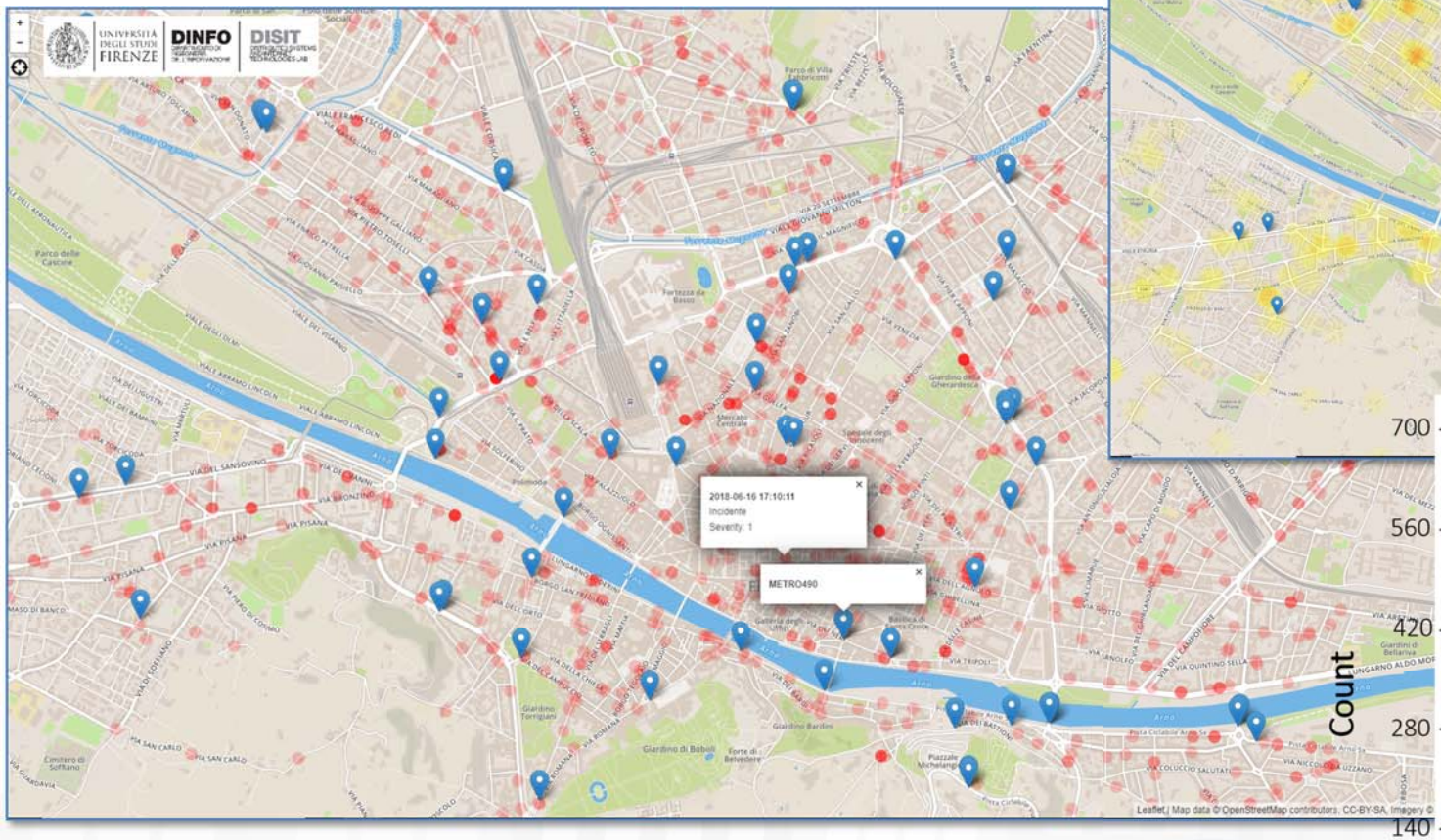
AP average and confidence

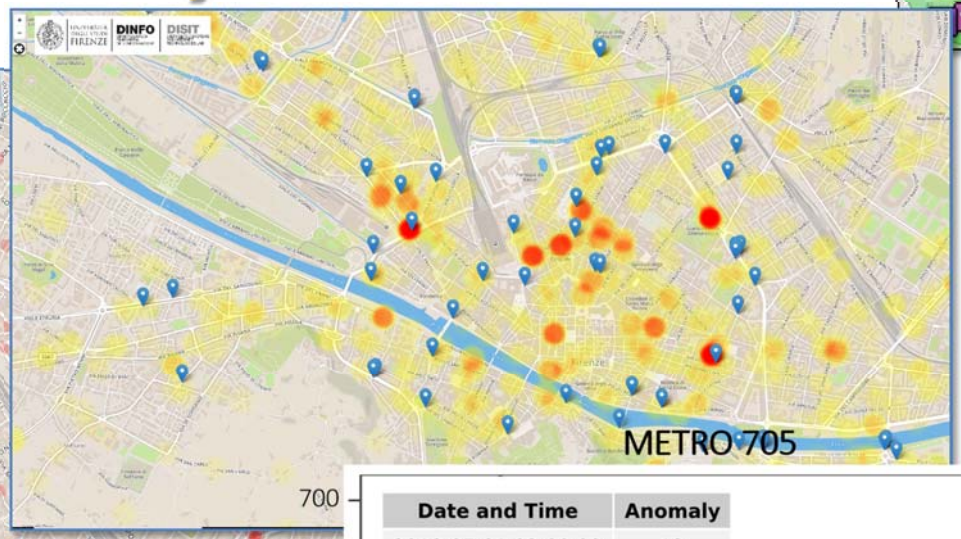
Actual AP trend for today

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

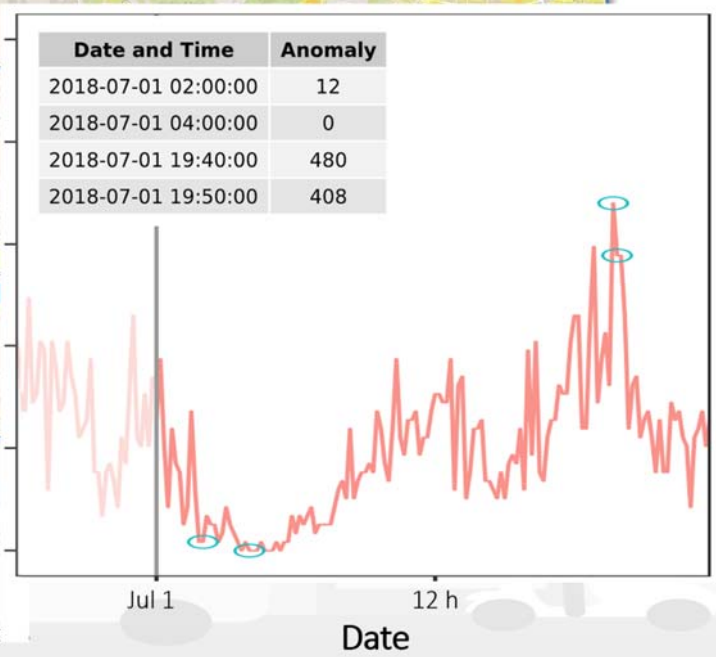
UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB http://www.disit.org

Anomaly Detection





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

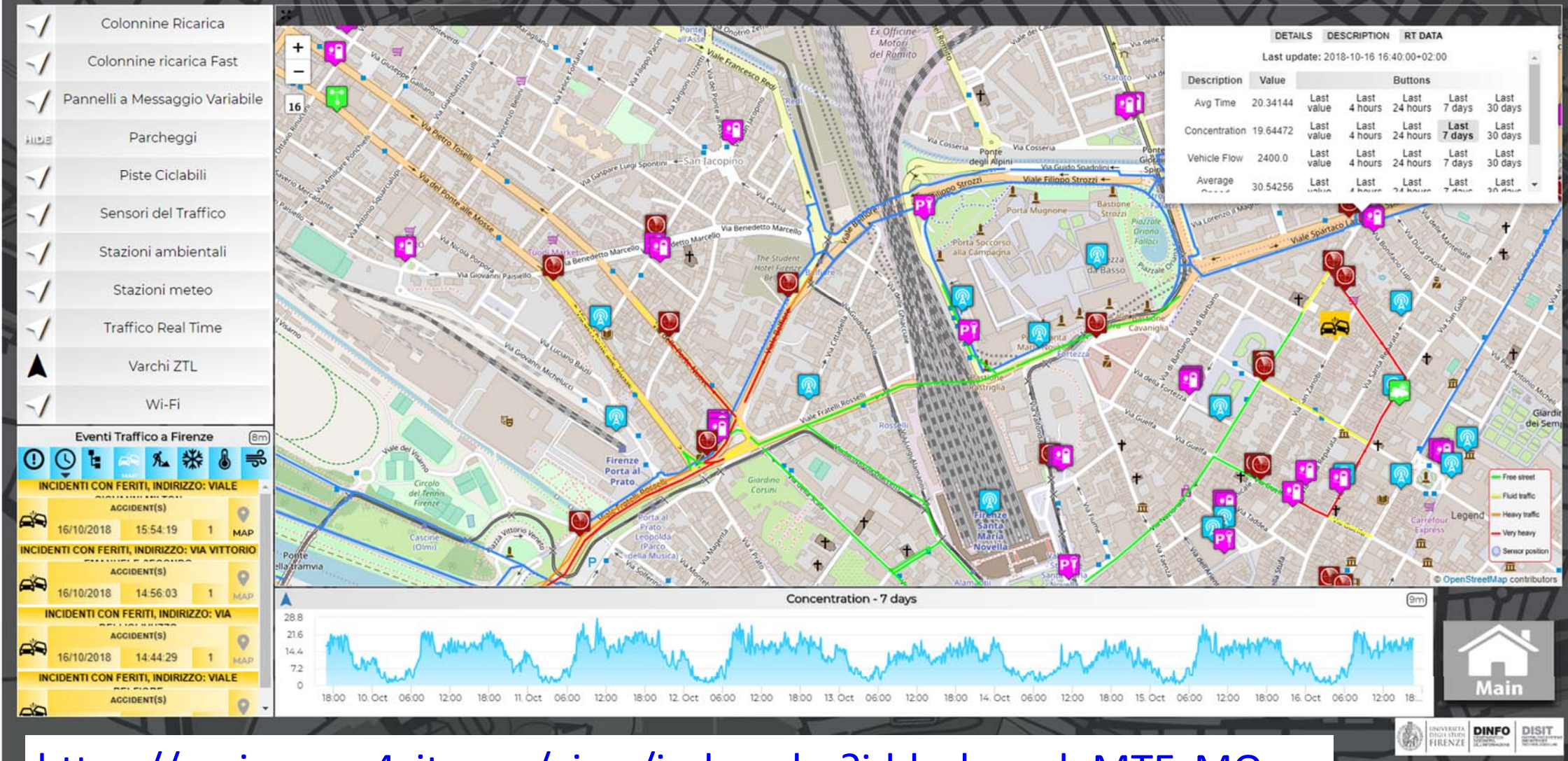


Accidents vs Traffic

WHAT-IF analysis, Reasoning and Simulation

MAPPA FIRENZE

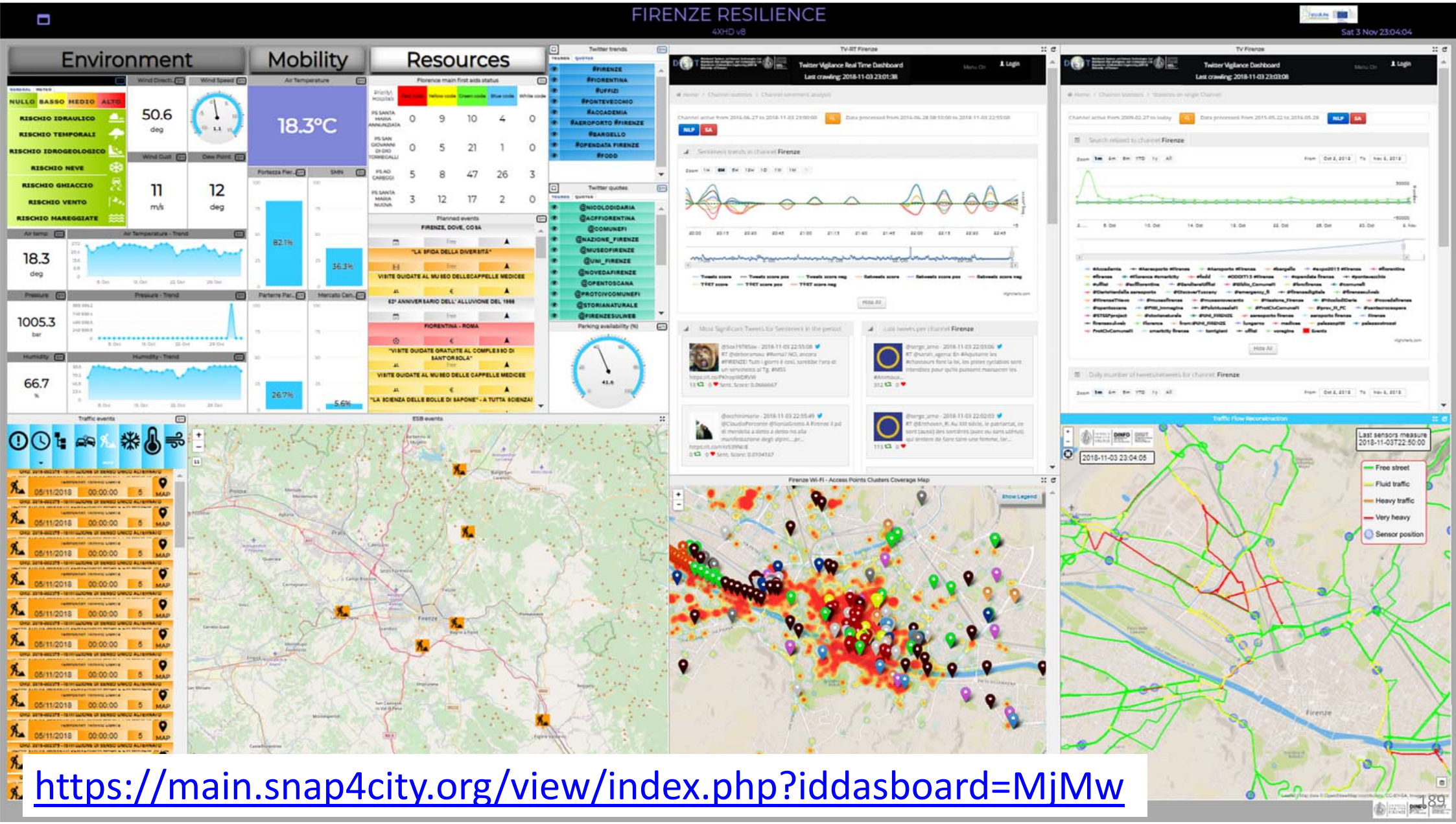
Tue 16 Oct 16:46:17



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

Multi-Widget Map

https://main.snap4city.org/view/index.php?iddasboard=OTc2

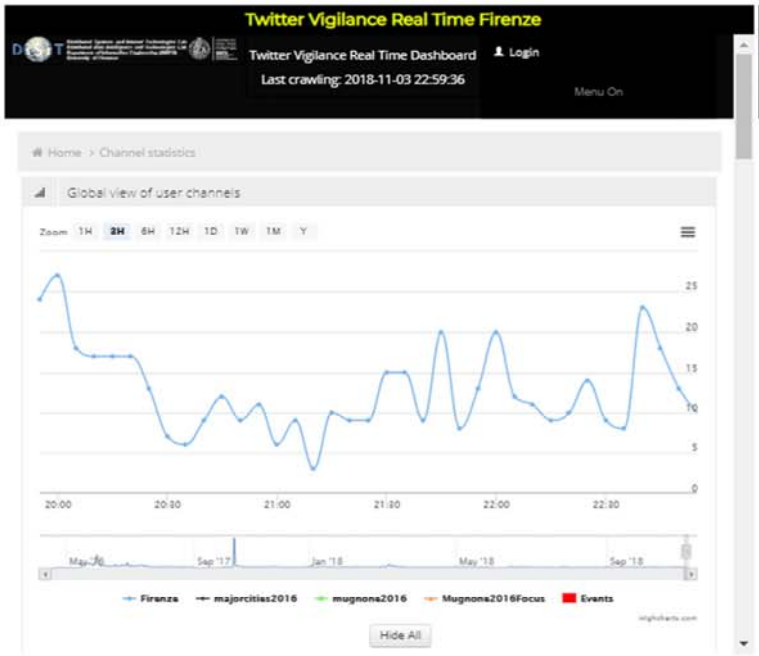


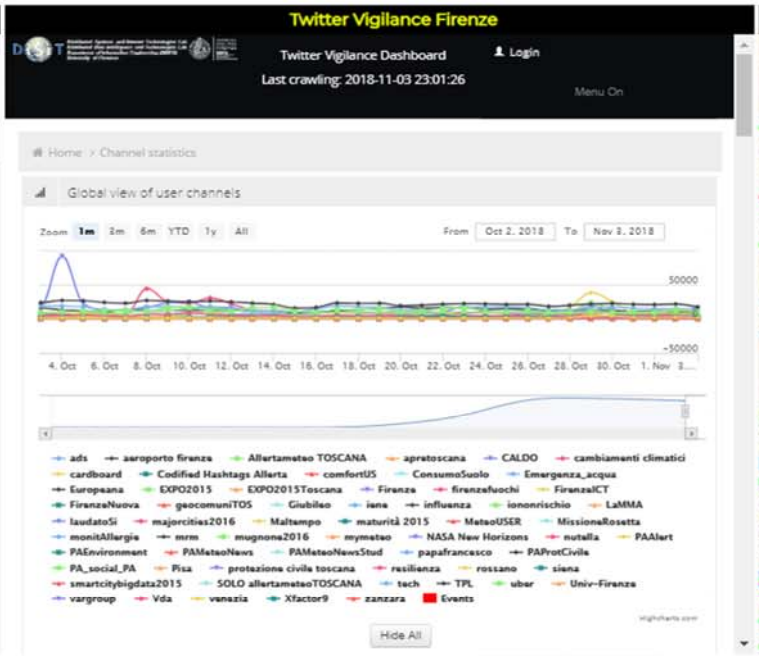
https://main.snap4city.org/view/index.php?iddasboard=MjMw

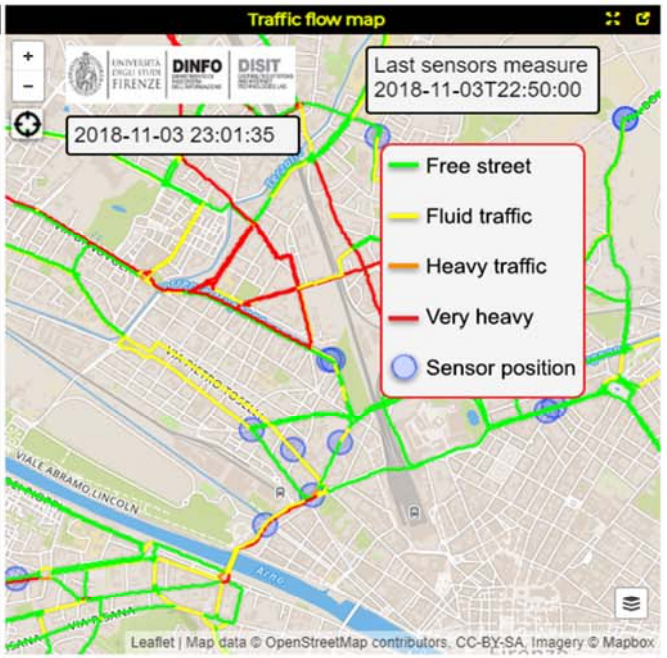
Firenze Social and Social Media new

Demo cams

Sat 3 Nov 23:01:35







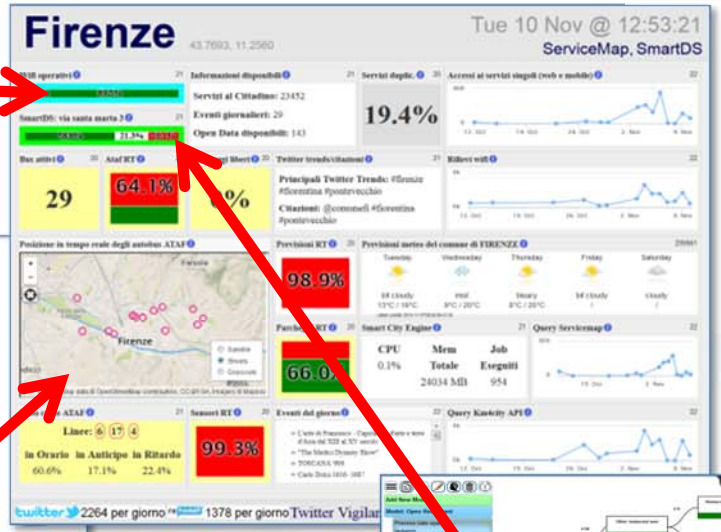
https://dashboard.km4city.org/view/index.php?iddasboard=MjY3

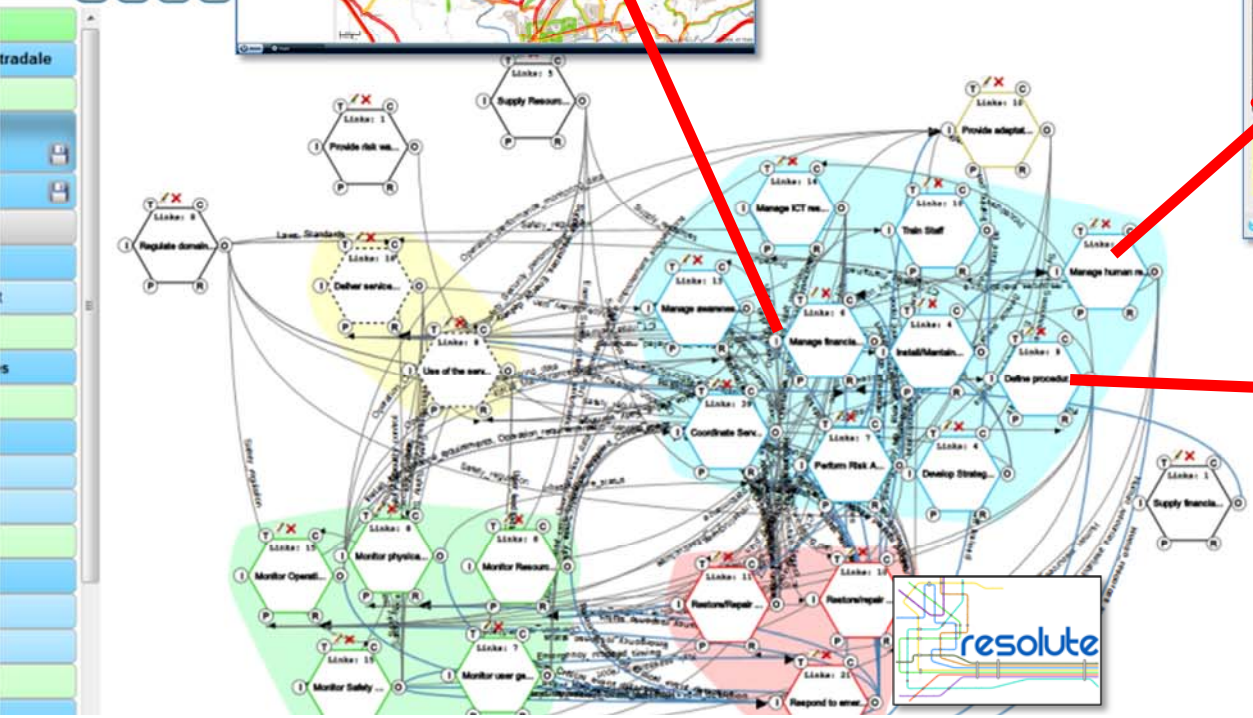
UNIVERSITÀ DEGLI STUDI FIRENZE | DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE | DISIT DISTRIBUTED SYSTEMS AND INTERNET TECHNOLOGIES LAB

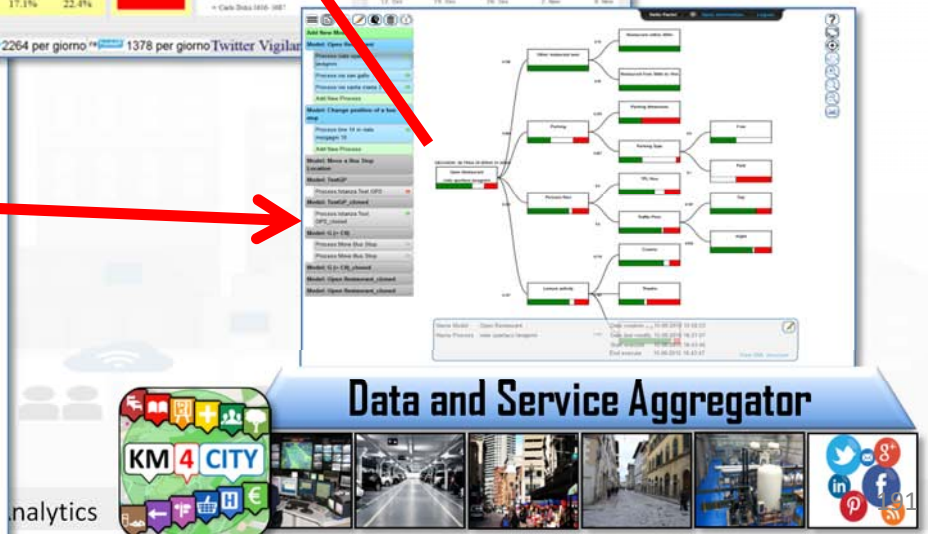
SNAP

resolute

Dashboarding city resilience





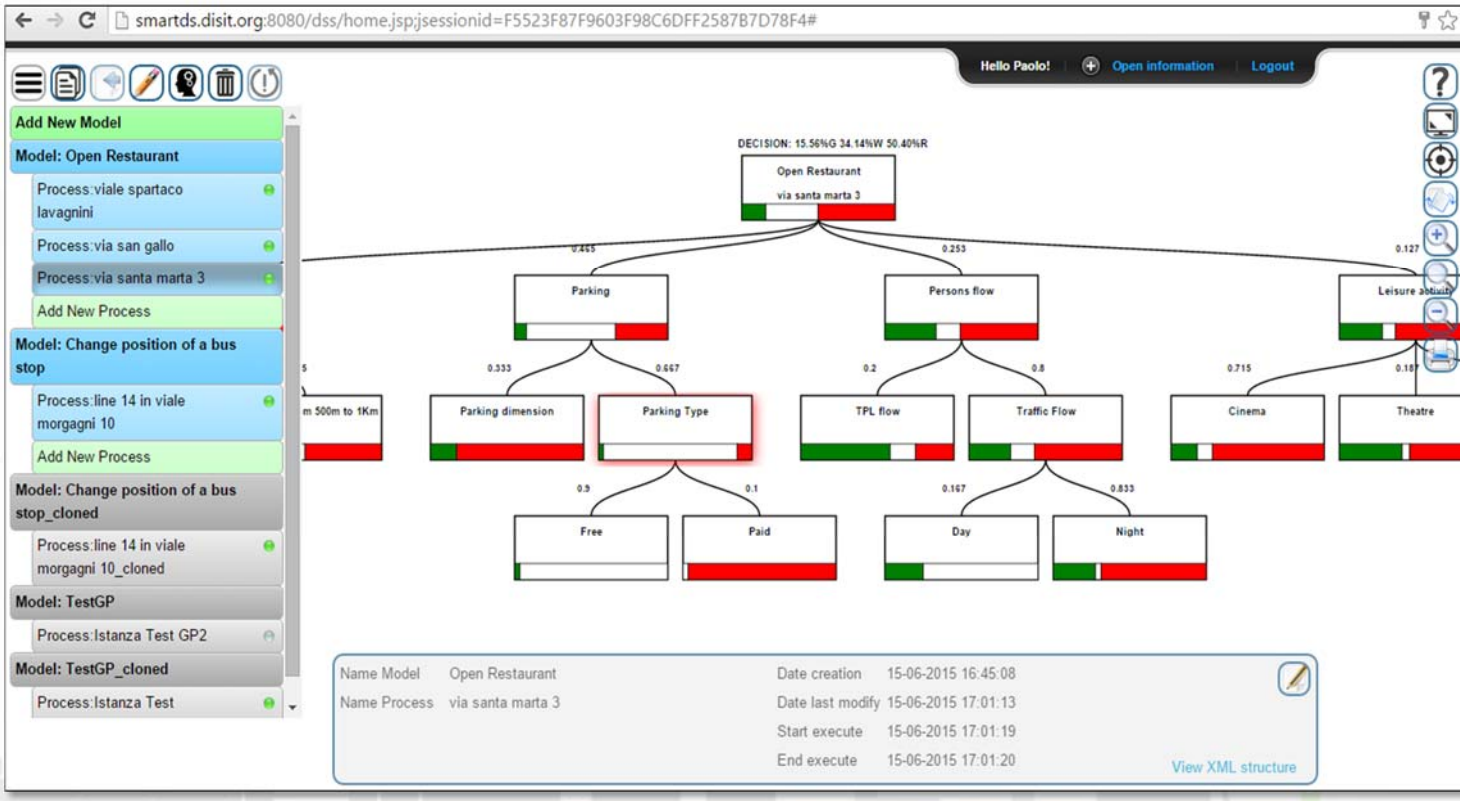


Smart City Decision Support

- ***Smart Decision Support System*** *based on System Thinking plus*
- Actions to city reaction, resilience, smartness..

Enforcing

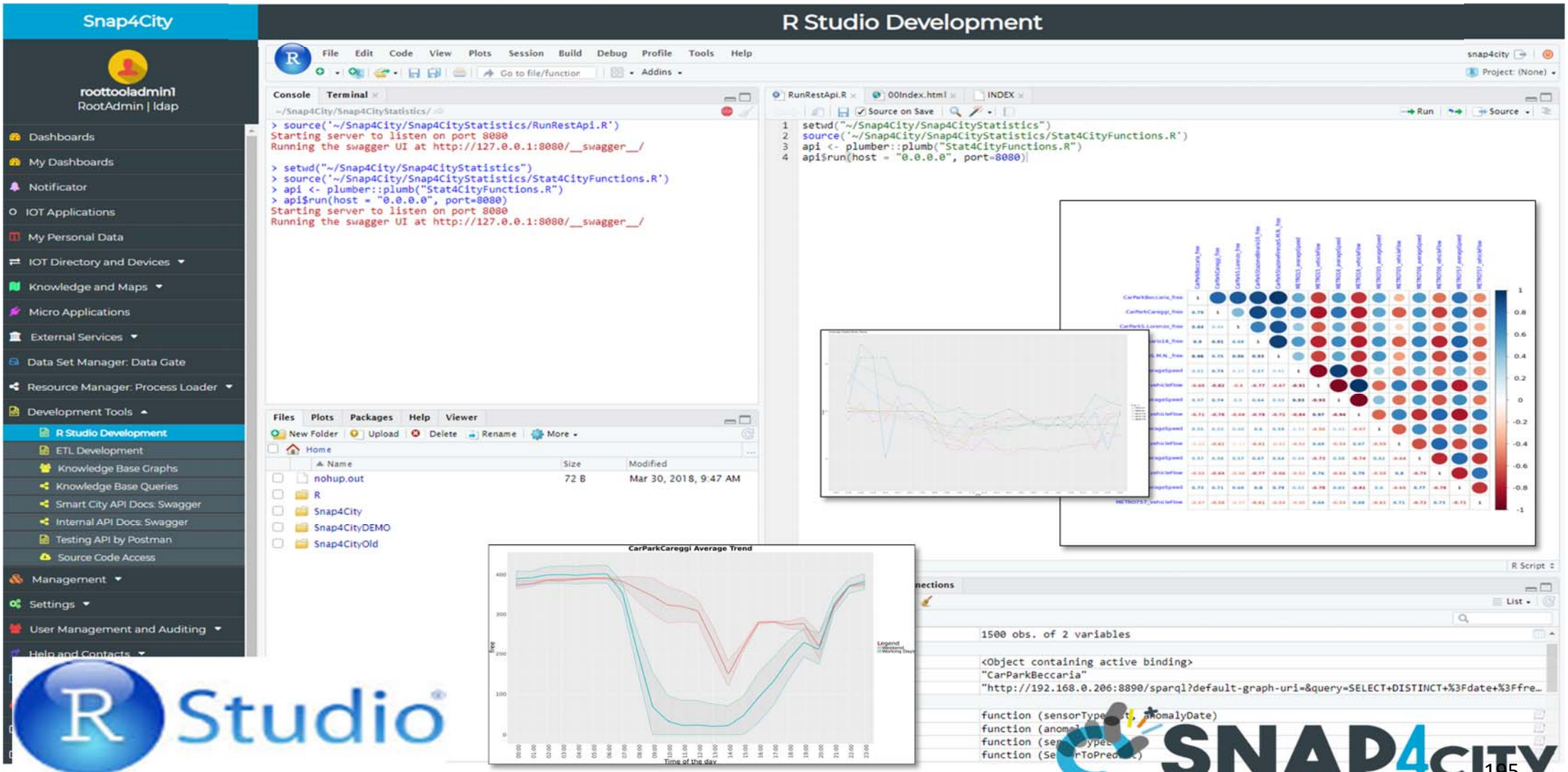
- Mathematical model for propagation of decision confidence..
- Collaborative work…,
- Processes connected to city data: DB, RDF Store, Twitter, etc.
- Production of alerts/alarms
- Data analytics process
- Twitter Processes
- reuse, copy past, …



Data Analytics in R as Development Tool for IOT Apps

ARC
INDEX

Developing Data Analytics in R Studio



END