



# *Big Data Architectures*

<https://www.snap4City.org>

<https://www.Km4City.org>

Parte: 0



# Agenda



- Laboratorio DISIT
- Tematiche del corso
- Struttura del corso
- Infrastruttura del DISIT Lab
- Modalità dell'esame





UNIVERSITÀ  
DEGLI STUDI  
FIRENZE

**DINFO**  
DIPARTIMENTO DI  
INGEGNERIA  
DELL'INFORMAZIONE

**DISIT**  
DISTRIBUTED SYSTEMS  
AND INTERNET  
TECHNOLOGIES LAB

## Distributed Data Intelligence and Technologies Lab Distributed Systems and Internet Technologies Lab

*Paolo Nesi*

Department of Information Engineering

University of Florence

Via S. Marta 3, 50139, Firenze, Italy

tel: +39-055-2758515, fax: +39-055-2758570

<http://www.disit.dinfo.unifi.it/>

[paolo.nesi@unifi.it](mailto:paolo.nesi@unifi.it) , <http://www.disit.dinfo.unifi.it/nesi/>



UNIVERSITÀ  
DEGLI STUDI  
FIRENZE

**DINFO**  
DIPARTIMENTO DI  
INGEGNERIA  
DELL'INFORMAZIONE

**DISIT**  
DISTRIBUTED SYSTEMS  
AND INTERNET  
TECHNOLOGIES LAB

**DISIT Lab**

- Researchers: 20
- Current Active Projects: 18
- Project in the last 4 years: 34
- Research Budget (last 2 years): 1.5M€
- Foreseen Research Budget (next 2 years): 2.2M€
- SpinOff: 1





**DISIT** Distributed Systems and Internet Technologies Lab  
Distributed Data Intelligence and Technologies Lab  
Department of Information Engineering (DINFO)  
University of Florence

<http://www.disit.dinfo.unifi.it>

UNIVERSITÀ  
DEGLI STUDI  
FIRENZE  
DINFO  
DIPARTIMENTO DI  
INGEGNERIA  
DELL'INFORMAZIONE

qualsiasi tipo   deep search

HOME ABOUT RESEARCH INNOVATION CORSI E TESI COME FARE EVENTI MIO PROFILO

root Uscire

Mostra Modifica Log Translate Devel

## DISIT LAB OVERVIEW

<http://www.disit.dinfo.unifi.it>

**DISIT lab and research group** is active since 1994. It is one of the most active ICT labs of the University of Florence, metropolitan Tuscany area. DISIT successfully developed a relevant number of International and National research, development and innovation projects. DISIT provides an infrastructure for R&D and dissemination activities. In the last 10 years DISIT has coordinated a number of more than 120 projects, mainly at European level. DISIT has covered the role of partner, and also coordinating scientific and technical WP and performing activities of dissemination and assessment. DISIT has received a relevant number of awards and is directly involved into top level international conferences, advisory boards, and committees.

**DISIT research areas:** big data, artificial intelligence, natural language

**CONTENUTI**

- Ultime Attività
- In primo piano
- Più visti
- Most Viewed (last 500)
- Most Viewed All (last 500)
- Ultimi caricati
- Più votati
- Mie collezioni pubblicate
- Miei contenuti
- Carica un nuovo contenuto

**ROOT**

- Gruppi
- Cerca Utenti
- Contenuti ed attività non lette relative ai tuoi gruppi
- Crea la matrice di tassonomia
- Forum
- Invite a colleague
- Issues
- Keyword cloud
- Messaggi e Sottoscrizioni
- Mio MatchMaking
- My issues
- News Blog
- Salva informazioni del cluster
- Workflow summary



UNIVERSITÀ  
DEGLI STUDI  
FIRENZE

**DINFO**  
DIPARTIMENTO DI  
INGEGNERIA  
DELL'INFORMAZIONE

**DISIT**  
DISTRIBUTED SYSTEMS  
AND INTERNET  
TECHNOLOGIES LAB  
<http://www.disit.org>

# Con chi lavoriamo



First in Traffic Solutions.



**THALES**

**BUSITALIA**  
GRUPPO FERROVIE DELLO STATO ITALIANE

**BBC**



**TIM**

**LEONARDO**

**ataf**



**BMG**



**CN**  
COSTRUZIONI NOVICROM



**ALSTOM**

**COMPUTER  
GROSS**

**PHILIPS tiscali:**

e-distribuzione



**Fraunhofer**

Consiglio Nazionale  
delle Ricerche

**cini**

**cniit**

**CERTH**  
CENTRE FOR  
RESEARCH & TECHNOLOGY  
HELLAS

consorzio nazionale  
interuniversitario  
per le telecomunicazioni



**Camera di Commercio  
Firenze**

**GIUNTI EDITORE**



CITTÀ METROPOLITANA  
DI FIRENZE



**REGIONE  
TOSCANA**



Agenzia regionale  
per la protezione ambientale  
della Toscana

National  
Technical  
University of  
Athens

**EPFL**  
ÉCOLE POLYTECHNIQUE  
FÉDÉRALE DE LAUSANNE

UNIVERSITÀ  
DEGLI STUDI  
BICOCCA



**UNIVERSITÀ  
DEGLI STUDI  
DI CAGLIARI**



**APRE TOSCANA**  
AGENZIA PER LA PROMOZIONE  
DELLA RICERCA EUROPEA

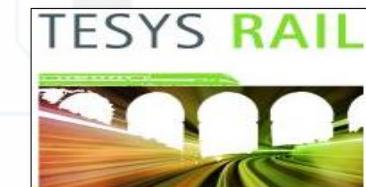
**CONCORDIA DISCORSO**  
ACCADEMIA NAZIONALE  
DI SANTA CECILIA

**CONSORZIO  
LaMMA**



# DISIT projects

- **Snap4City: IOT/IOE smart city** <https://www.snap4city.org>
- **Trafair: CEF project con varie Citta'**
- **Mosaic: Mobility and transport model**
- **Smart City of Florence Metropolitan Area**
- **Km4City:** <http://www.km4city.org>
- **REPLICATE H2020, SCC1, EC flagship**
  - <http://replicate-project.eu/>
- **Sii-Mobility SCN MIUR:** <http://www.sii-mobility.org>
- **Feedback: reail and GDO Big Data analytics**
- **5G with 3-Wind Open Fiber Estra**
- **Coll@bora Social Innovation, MIUR:**
  - <http://www.disit.org/5479>
- **RESOLUTE H2020, EC:**
  - <http://www.resolute-eu.org>
- **TRACE-IT, RAISSS, TESYSRAIL, ....**
- **Mobile Emergency:**
  - <http://www.disit.org/5404>



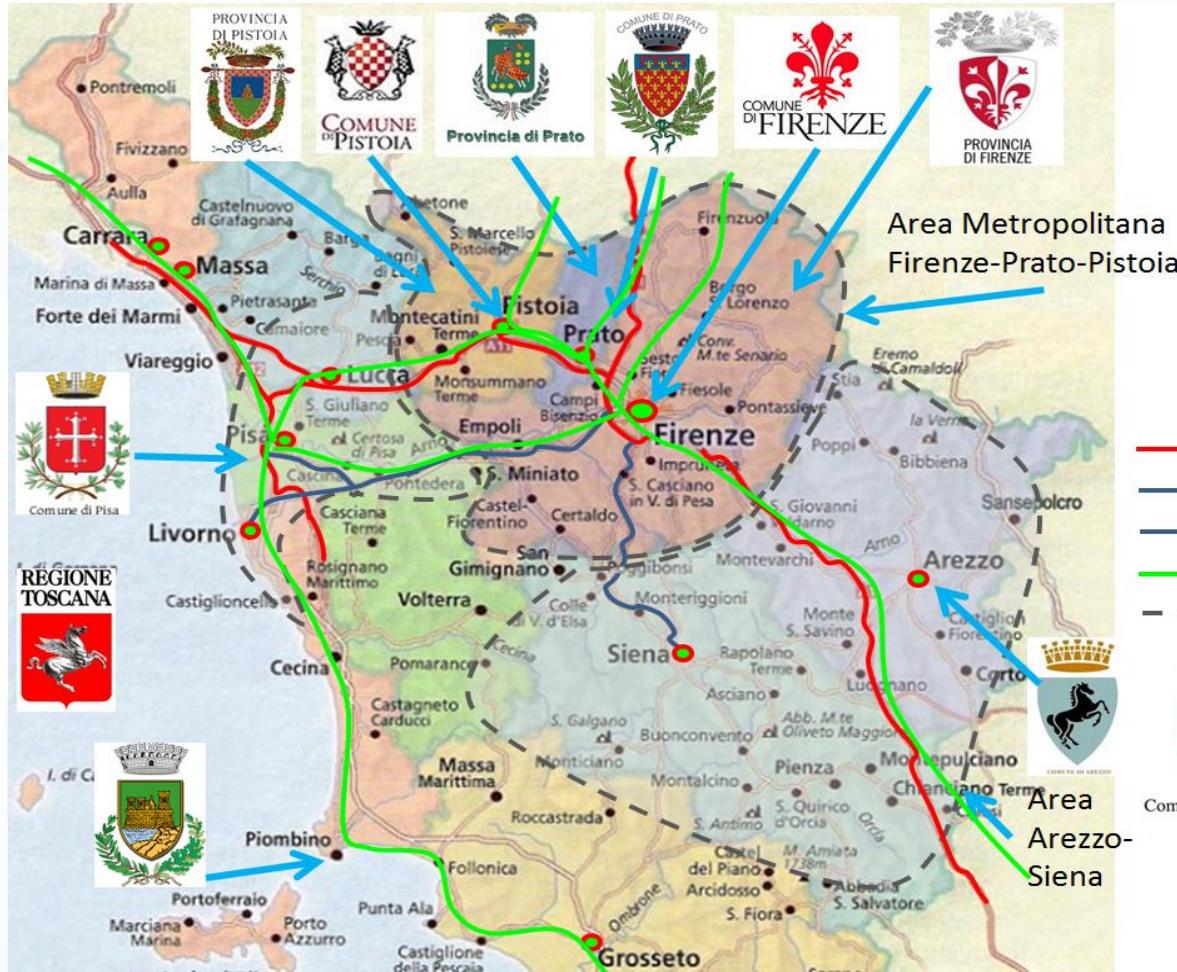
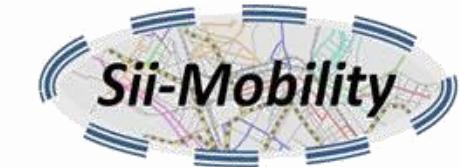
**Trace-IT**  
**RAISSS**



# Sii-Mobility

□ <http://www.Sii-Mobility.org>

- Experiments and validation in Tuscany
- Integration with present central station and subsystems
- DISIT lab, Università di Firenze, is the tech-scientific coordinator

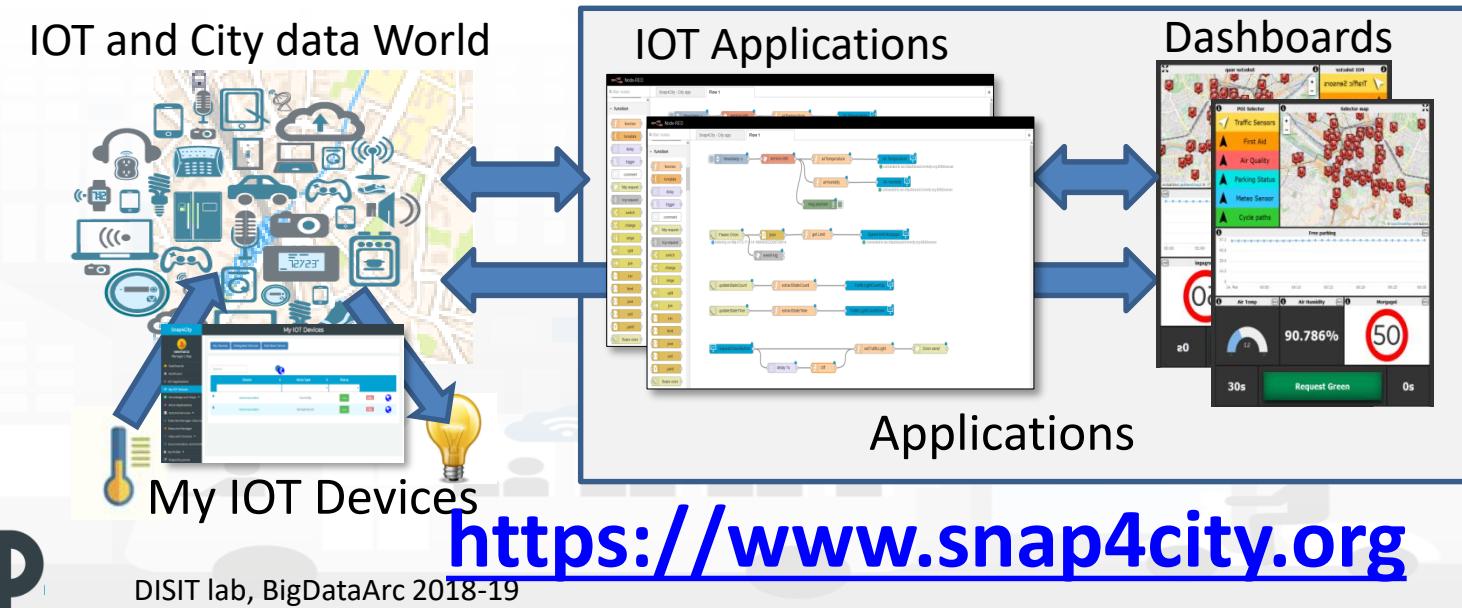


- Autostrade
- SS Fi-Pi-Li
- SS Fi-Si
- Ferrovie (primarie)
- - Aree

*ECM; Swarco Mizar;  
Inventi In20; Geoin;  
QuestIT; Softec; T.I.M.E.;  
LiberoLogico; MIDRA  
(autostrade, motorola);  
ATAF; Tiemme; CTT  
Nord; BUSITALIA;  
A.T.A.M.; Effective  
Knowledge; eWings;  
Argos Engineering; Elfi;  
Calamai & Agresti;  
Project; Negentis*



- enabling large-scale co-creation IOT/IOE applications for Helsinki, Antwerp:
  - Open source, standardized, data-driven, service-oriented, user-centric platform, robust, scalable, easy to use solution, co-creation of mixt data driven, stream and batch processing
- extending the powerful semantic reasoner of Km4City <https://www.km4city.org>, with IOT/IOE, GDPR, and city dashboards.
- validated in multiple devices (PC, Android, Raspberry, IOT Button,..), and domains: mobility and transport, tourism, health, welfare, social
- The innovation on semantic reasoning, IOT interoperability, microservices, automated dashboard production, .. thus
- *smart city solutions in a*



<https://www.snap4city.org>



Horizon 2020  
European Union Funding  
for Research & Innovation

REnaissance of PLaces  
with Innovative Citizenship  
And TEchnology

<http://replicate-project.eu/>

- demonstrate Smart City technologies in energy, transport and ICT in districts in:
  - San Sebastian, Florence and Bristol,
  - follower cities of Essen, Nilufer and Lausanne
- Cities are the customer: considering local specificities
- Solutions must be replicable, interoperable and scalable.
  - Integrated Infrastructure: deployment of ICT architecture, from internet of things to applications
  - Low energy districts
  - Urban mobility: sustainable and smart urban services

- 1 (coordinator) FOMENTO DE SAN SEBASTIAN FSS SPAIN
- 2 AYUNTAMIENTO DE SAN SEBASTIAN SAN SEBASTIAN SPAIN
- 3 COMUNE DI FLORENCE FLORENCE ITALY
- 4 BRISTOL COUNCIL BRISTOL UNITED KINGDOM
- 5 STADT ESSEN ESSEN GERMANY
- 6 NILUFER BELEDIYESI NILUFER TURKEY
- 7 VILLE DE LAUSANNE LAUSANNE SWITZERLAND
- 8 IKUSI ANGEL IGLESIAS, S.A. IKUSI SPAIN
- 9 ENDESA ENERGÍA, S.A. ENDESA SPAIN
- 10 EUROHELP CONSULTING, S.L. EUROHELP SPAIN
- 11 ILUMINACION INTELIGENTE LUIX, S.L. LUIX SPAIN
- 12 FUNDACION TECNALIA RESEARCH & INNOVATION TECNALIA SPAIN
- 13 EUSKALTEL, S.A. EUSKALTEL SPAIN
- 14 COMPAÑÍA DEL TRANVÍA DE SAN SEBASTIÁN DBUS SPAIN
- 15 CONSIGLIO NAZIONALE DELLE RICERCHE CNR ITALY
- 16 ENEL DISTRIBUZIONE, SPA ENEL ITALY
- 17 MATHEMA, SRL MATHEMA ITALY
- 18 SPES CONSULTING SPES ITALY
- 19 TELECOM ITALIA, SPA TELECOM ITALY
- 20 UNIVERSITA DEGLI STUDI DI FLORENCE UNIFI ITALY:  
*DINFO.DSIT Lab and DIEF*
- 21 THALES ITALIA, SPA THALES ITALY
- 22 ZABALA INNOVATION CONSULTING ZABALA SPAIN
- 23 TECHNOMAR TECHNOMAR GERMANY
- 24 UNIVERSITY OF BRISTOL UOB UNITED KINGDOM
- 25 UNIVERSITY OF OXFORD UOXF UNITED KINGDOM
- 26 BRISTOL IS OPEN, LTD BIO UNITED KINGDOM
- 27 ZEETTA NETWORKS ZEETTA UNITED KINGDOM
- 28 KNOWLE WEST MEDIA CENTRE, LGB KWMC UNITED KINGDOM
- 29 TOSHIBA RESEARCH EUROPE, LTD TREL UNITED KINGDOM
- 30 ROUTE MONKEY, LTD ROUTE MONKEY UNITED KINGDOM
- 31 ESOTERIX SYSTMES, LTD ESOTERIX UNITED KINGDOM
- 32 NEC LABORATORIES EUROPE, LTD NEC UNITED KINGDOM
- 33 COMMONWHEELS CAR CLUB CIC CO-WHEELS UNITED KINGDOM
- 34 UNIVERSITY OF THE WEST OF ENGLAND UWE UNITED KINGDOM
- 35 ESADE BUSINESS SCHOOL ESADE SPAIN
- 36 SISTELEC SOLUCIONES DE TELECOMUNICACION, S.L. SISTELEC SPAIN



Horizon 2020  
European Union Funding  
for Research & Innovation

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

- Develop European Resilience Management Guidelines (ERMG)
  - Develop a conceptual framework for creating/maintaining Urban Transport Systems
- Enhance resilience through improved support of human decision making processes, particularly by training professionals and civil users on the ERMG and the RESOLUTE system
- Operationalize and validate the ERMG by implementing the RESOLUTE Collaborative Resilience Assessment and Management Support Systems (CRAMSS) for Urban Transport Systems addressing Road and Urban Rail Infrastructures
  - Pilots in Florence and Athens
- Adoption of the ERMG at EU and Associated Countries level

University of Florence: DISIT lab DINFO (Proj coordinator), DISIA and DST	UNIFI	IT
THALES	THALES	IT
ATTIKOMetro	ATTIKO	GR
Comune di Firenze	CDF	IT
Centre for Research and Technology Hellas	CERTH	GR
Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V.	FHG	DE
HUMANIST	HUMANIST	FR
SWARCO Mizar	SWMIZ	IT
Associação para o Desenvolvimento da Investigação no Instituto Superior de Gestão	ADI-ISG	PT
Consorzio Milano Ricerche	CMR	IT

# Roadmap



2013

## Km4City 1.1

- Tuscany Map
- Services
- AVM
- Sensors
- Parking
- Cultural Heritage
- Enrichment cities
- Event in the city
- Digital Locations
- Fresh places

- <http://servicemap.km4city.org>
- <http://log.disit.org>
- <http://www.disit.org/fodd>
- <http://www.disit.org/tv> Twitter Vigilance
- <http://smartds.km4city.org>

- Weather
- Cultural Heritage
- Energy recharge pillar
- Wi-Fi
- Events in the city

2015



RESilience management guidelines and Operationalization applied to Urban Transport Environment

## Km4City 1.5

- SmartDS
- Km4City App

## Km4City 1.4

- Embed
- More API
- iBeacon

API



<http://www.sii-mobility.org>

2016

## REPLICATE H2020 2016-2021 - Started



**REPLICATE**  
Renaissance of Places with Innovative Citizenship And Technology  
INNOVATION ACTIONS

# Roadmap



12/2017  
WEEE  
2017-2020

6/2017

## ALTAIR



**SELECT**  
for Cities

waste

- Territorial areas and paths
- Health, Bike sharing
- Statistics, Energy, ICT, ...
- E-vehicles

- Risk analysis
- Environmental, water
- Data Licensing models
- Energy Meters
- Fi-Ware compliant



- More Sensors, IOT, IoT
- Dashboard Builder
- Territorial areas and paths
- User Engagement
- Mobility and transport
- Resilience Decision Support

GHOST SIR  
2016-2019 - Started



GHOST  
the smart city a governance-centred approach to Smart Urban

IOT/IOT

- Monitoring
- Smart City IOT integration
- Living Lab

**5G**

- Smart City vs IOT, Industria 4.0

**MOSAIC**

6/2018

**Trafair CEF**

**SMARTCITY**  
EXPO WORLD CONGRESS

11/2018

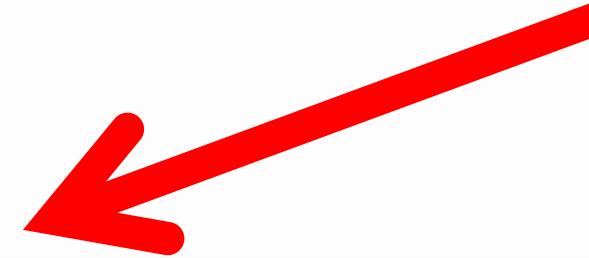
**ICT 2018**  
IMAGINE DIGITAL CONNECT EUROPE

12/2018

2021

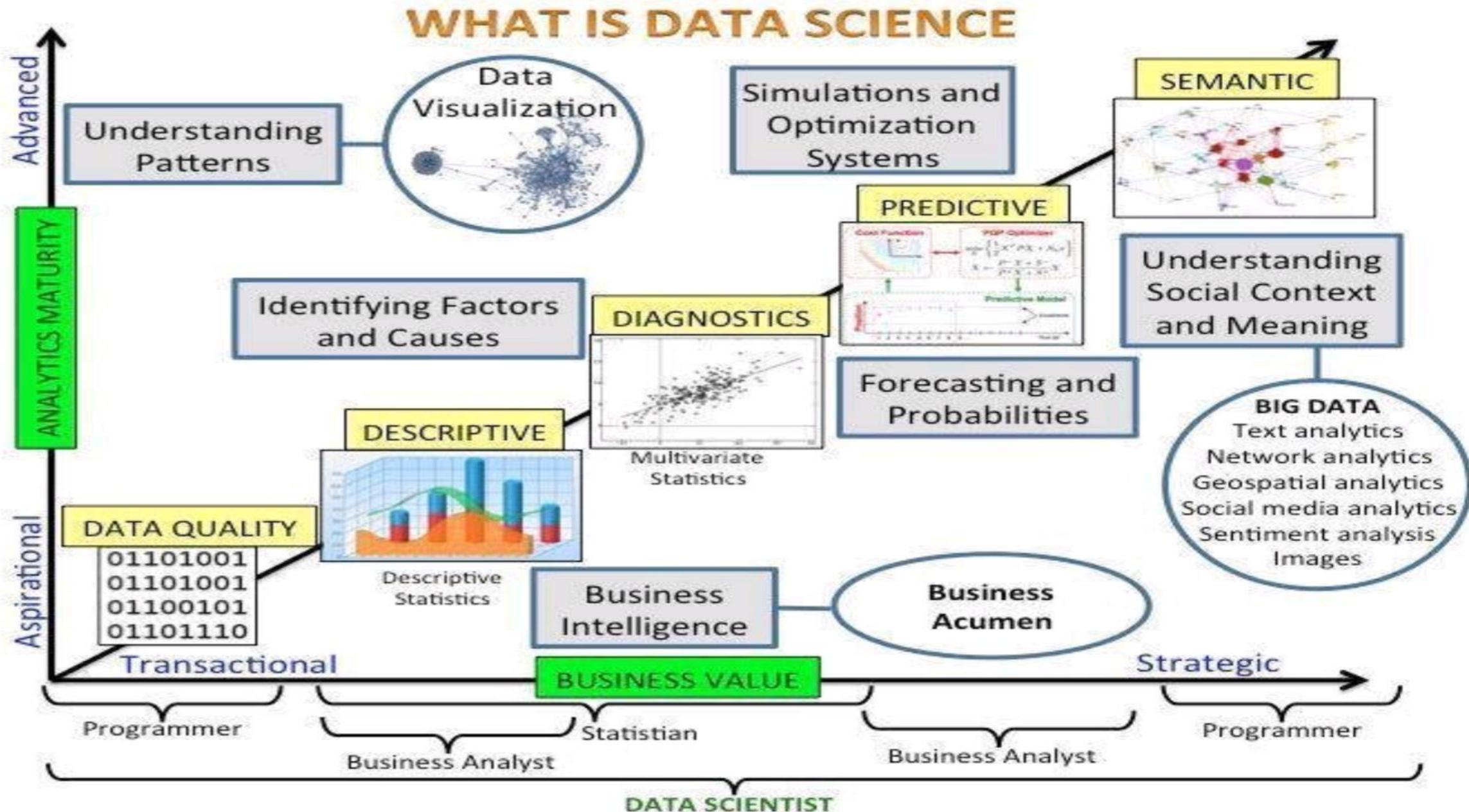
# Agenda

- Laboratorio DISIT
- Tematiche del corso
- Struttura del corso
- Infrastruttura del DISIT Lab
- Modalità dell'esame

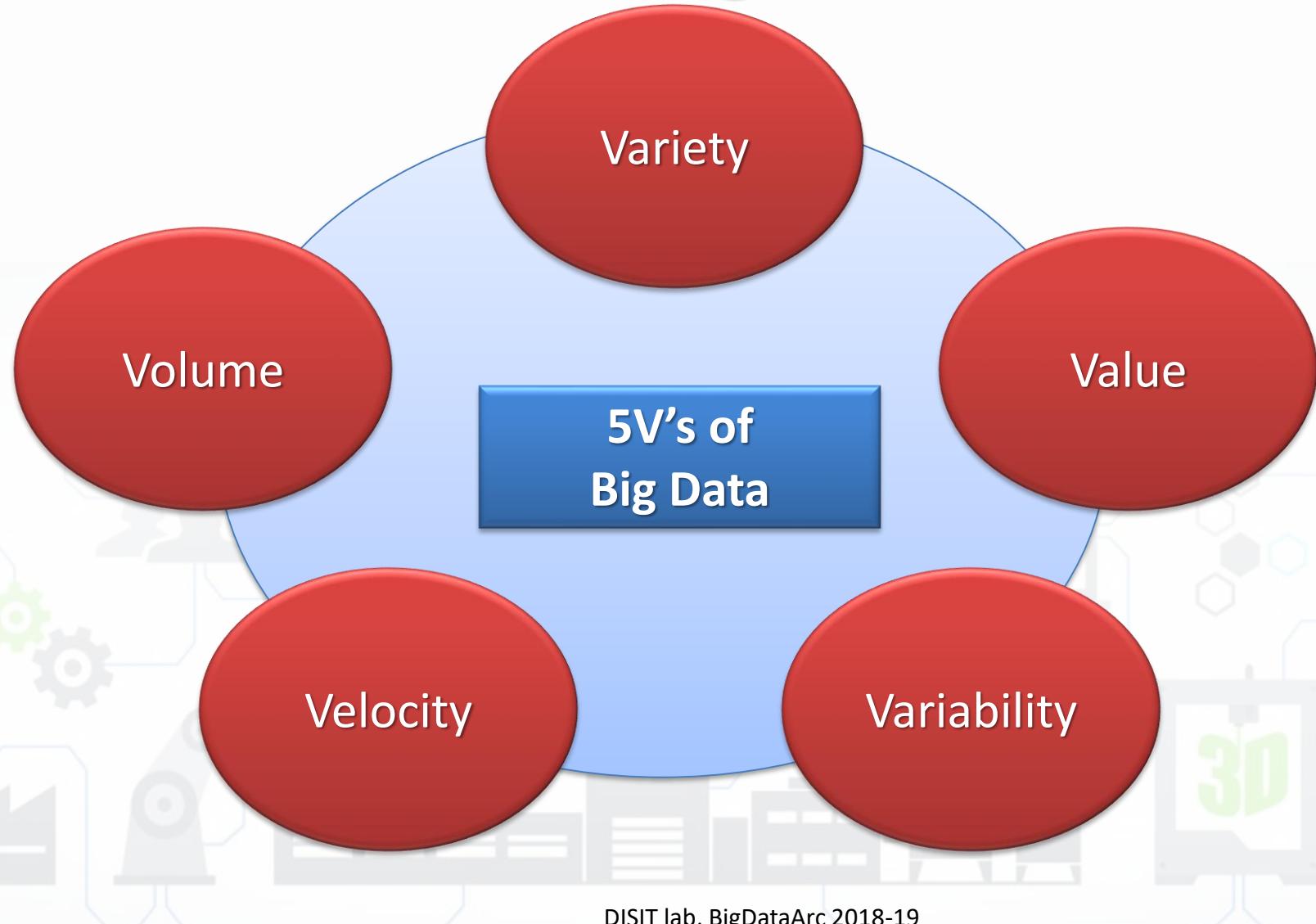


DISIT lab, BigDataArc 2018-19

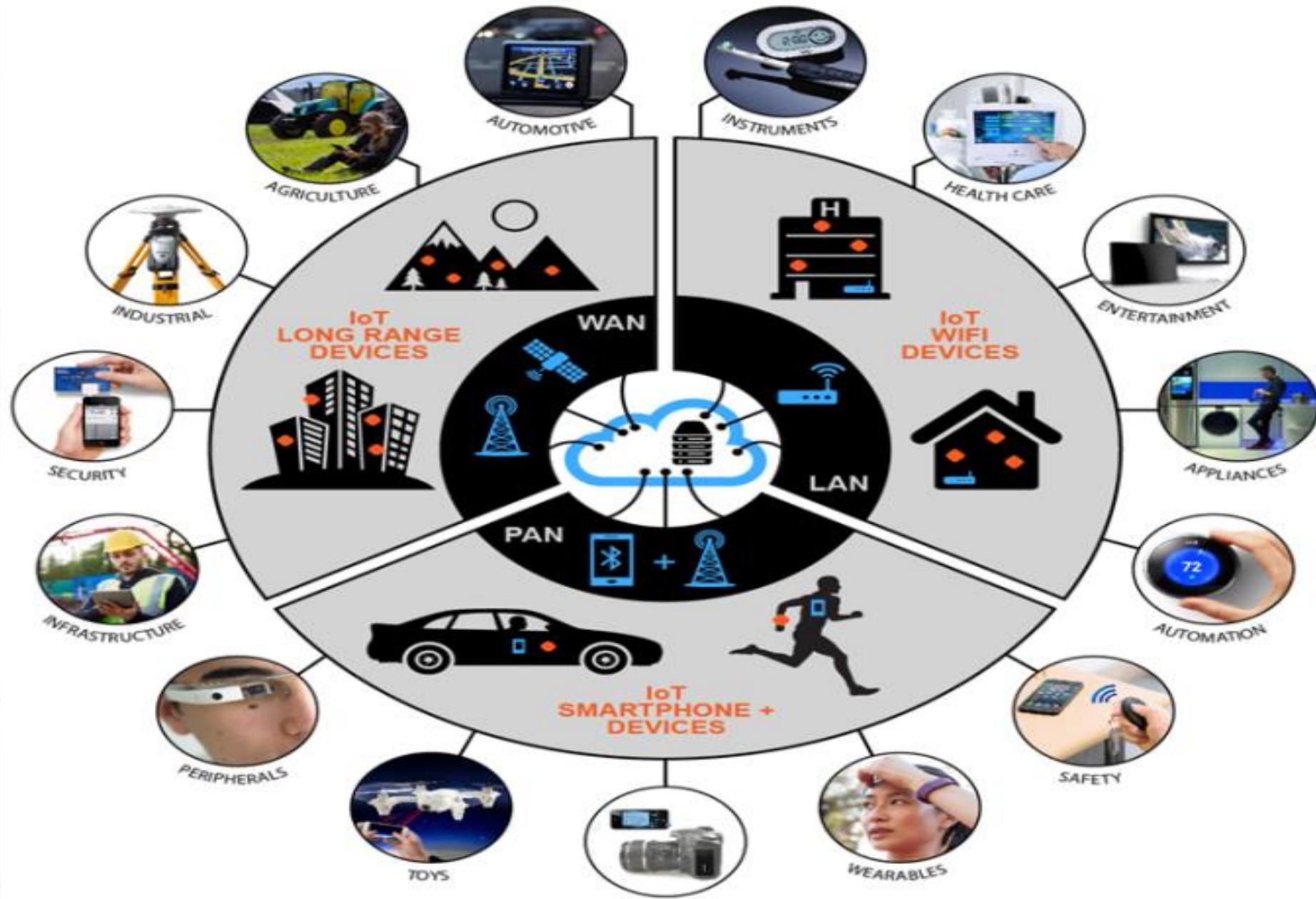
**Id:** openlink-  
**Iod-cache**  
**Number of triple:**  
57785989163  
**Out-connection:** 13  
**In-connection:** 0



# 5V of Big Data

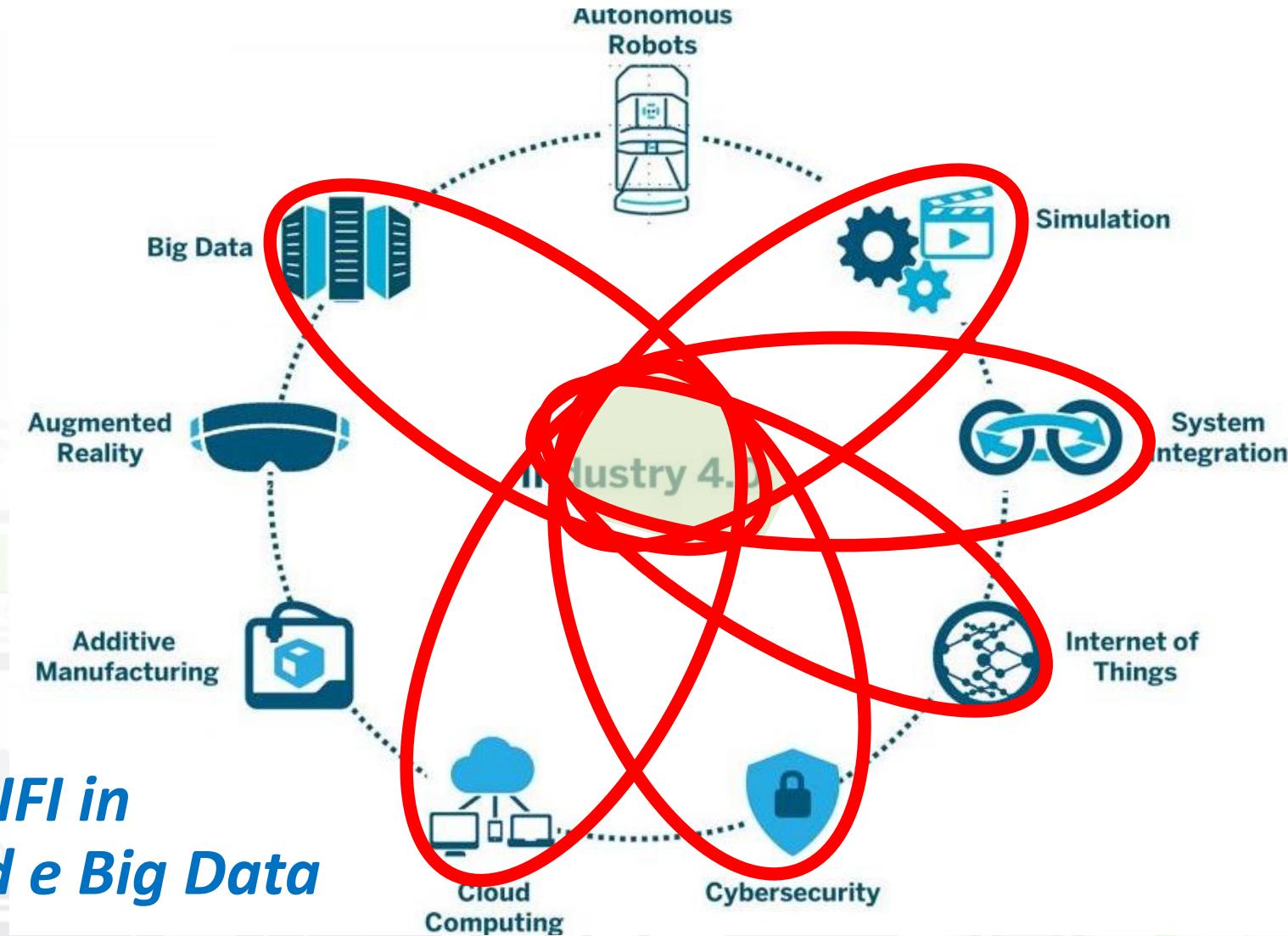


# BIG Data & Analytics



# Industria 4.0 vs DISIT Lab

- Big Data
- Cloud Computing
- Cybersecurity
- IOT/IOE
- System Integration
- Simulation
- +
- Data Analytics
- *P. Nesi è referente per UNIFI in Regione Toscana per Cloud e Big Data*



# *DISIT: Competences*

- **Technologies:**
  - **Big Data and Analytics:** data management, user analysis, user engagement, prediction, predictive maintenance, early detection, anomaly detection, data intelligence, ...
  - **Data Mining:** artificial intelligence, machine learning, natural language processing, semantic computing, semantic reasoner, expert systems, statistic analysis, ..
  - **IOT/IOE:** internet of things/everything, brokers, microservices, ..
  - **Cloud:** smart cloud, cloud simulation, optimization, containers, ..
  - **Mobile Computing:** mobile application, user behavior analysis, ..
  - **NLP and Sentiment Analysis:** response vigilance, interaction, answering, Personal Assistant, NLP, SA, ..
- See projects on: <http://www.disit.org/5501>

# Application Fields

Increasing investments in Big Data can lead to interesting discoveries in **science, medicine**, benefits and gains in the **ICT sector** and in **business** contexts, new services and opportunities for digital **citizens** and **web users**.

- Mobility and transport
- Energy and Govern
- Data Analysis – Scientific Research
- Industry 4.0
- Healthcare and Medicine
- Social Network – Internet Service – Web Data
- Educational
- Financial/Business
- Security

The image shows a large, dense network graph centered around the word "architettura". The word is written in a large, bold, black sans-serif font. The network consists of numerous small, semi-transparent nodes of various colors (light blue, orange, yellow, green) connected by thin white lines representing edges. These nodes represent different entities or concepts, many of which are labels for datasets or projects, such as "eurostat-linked-data", "wikidata", "austrian-sk\_racers", "eu-screan", "yago", "bpr", "tu-berlin-elder", and "hellenic-fire-brigade". The background is white, and the overall effect is a complex web of connections radiating from the central word.

```
openlink  
od-cache  
ber of triple:  
785989163  
onnection: 13  
onnection: 0
```

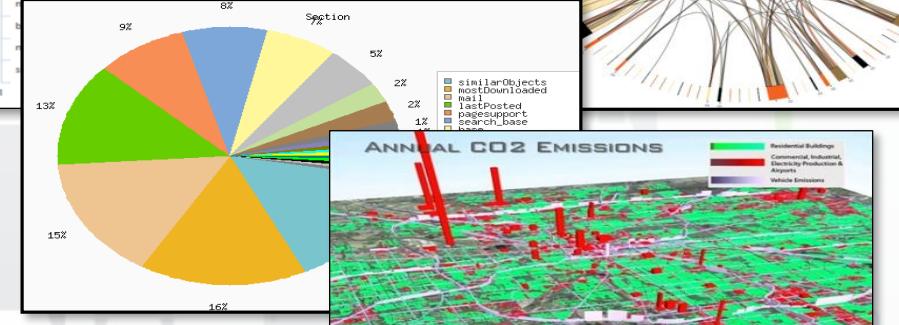
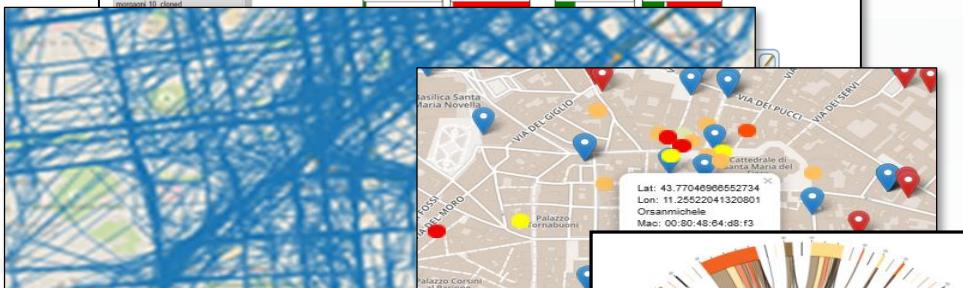
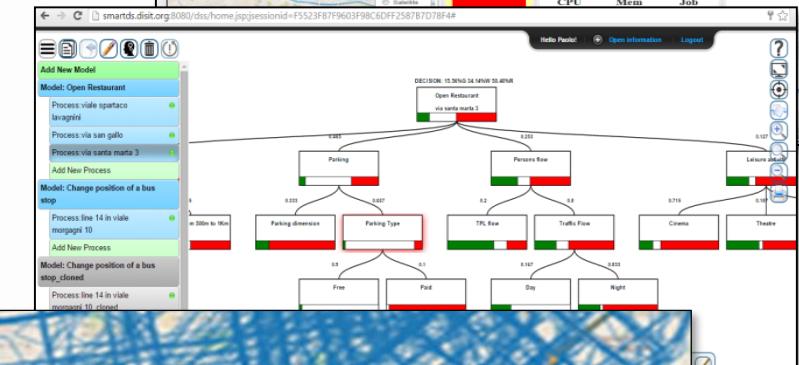
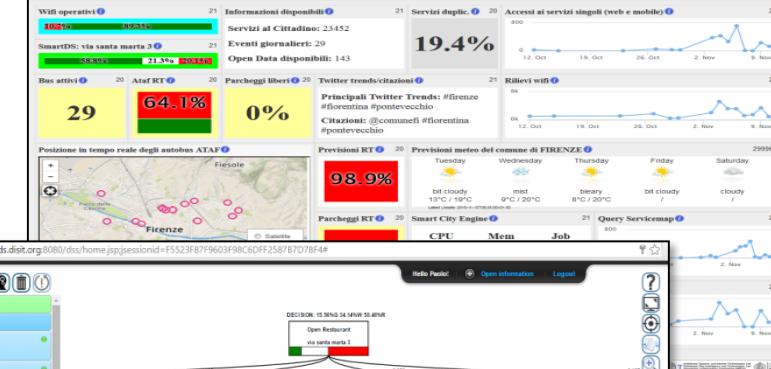
# Decisioni supportate dai dati

## periodiche ed in tempo reale

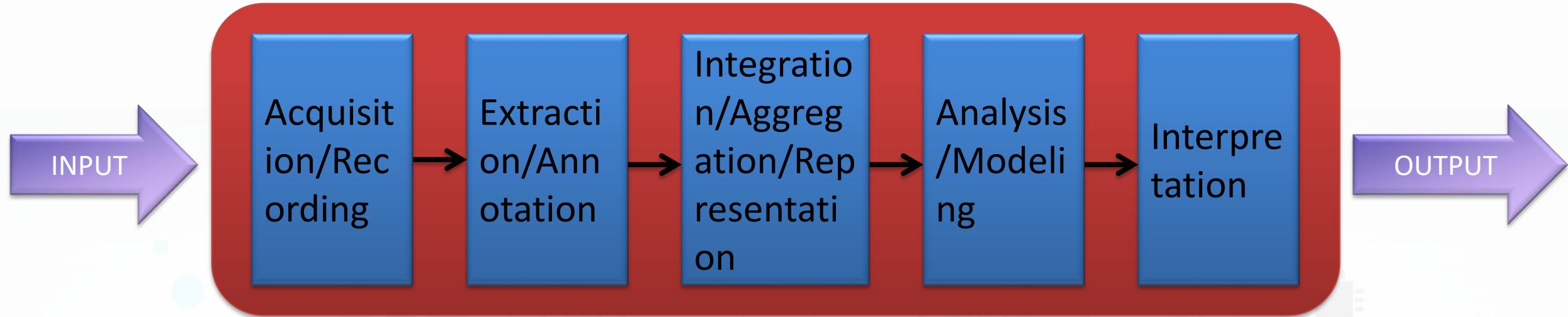
- Condivisione e Integrazione Dati  
multidominio: *semantica e bigdata*
- Dati → Smart City Engine → Control Room
- analisi: monitoraggio, flussi e comportamenti, sondaggi, mining, correlazioni, cause – effetti, etc.
  - Per il miglioramento di servizi correnti
  - Per reagire ad eventi, incremento della resilienza,
  - Per la creazione servizi innovativi

Firenze

43.7693, 11.2560



# Pipeline



# 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?

# Architettura di base Big Data, IOT, Industry 4.0

**Data Sources**

Transactions sys,  
sensors  
Social media,  
ws, etc.

**Data Stream analysis**

Spark,  
Storm,  
Kafka

**Big Data Cluster**

HDFS, noSQL

**Data Transformation**  
ETL, NIFI

**Indexing**

SOLR,  
Elastic search

**Data Analytics**

R, TF, ...

**Search and Query**

Facet,  
cluster

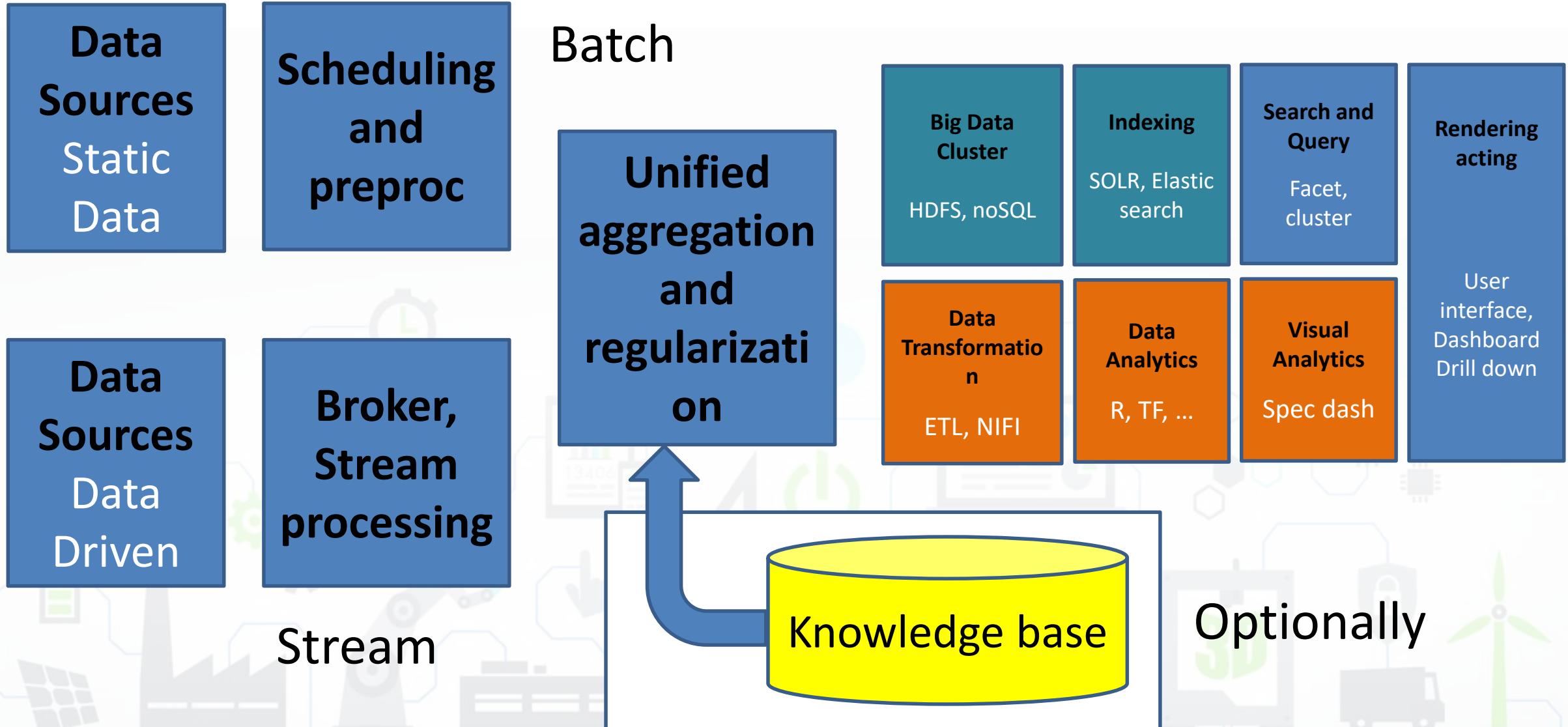
**Visual Analytics**  
Spec Dash

**Rendering acting**

User interface,  
Dashboard  
Drill down

**Data Management: security, privacy, licensing, etc.**

# Lambda Architecture



Transport systems  
Mobility, parking



Public Services  
Govern, events, ...



Sensors, IOT  
Cameras, ..



Environment,  
Water, energy



Shops, services,  
operators



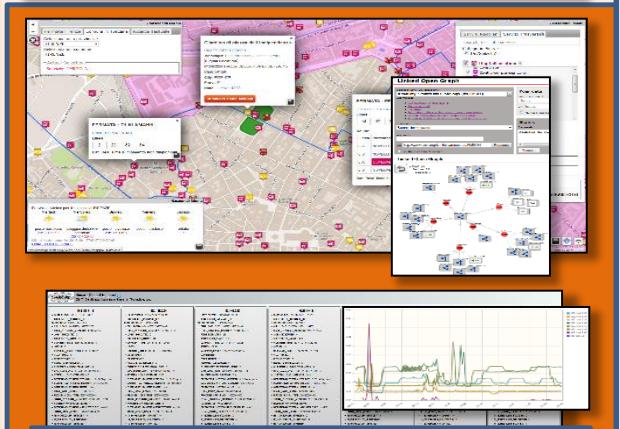
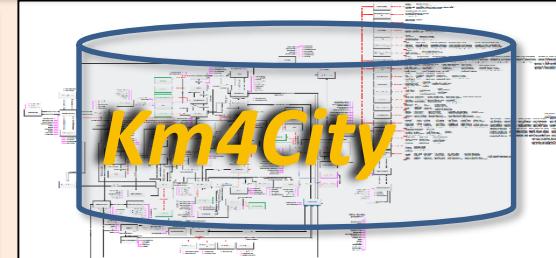
Social Media  
WiFi, network



## Static, Slow and Real Time data flows

DISCES -- Distributed and parallel architecture on Cloud

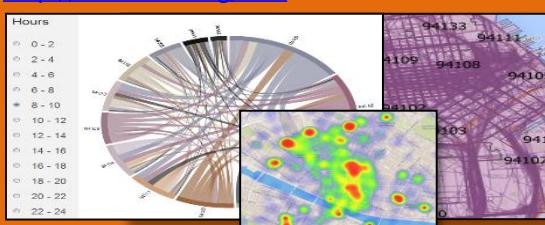
## Km4City Smart City Engine



### User Profiling and Suggestions on Demand

#### Flow and Origin Destination Matrix

[Http://www.disit.org/odsf](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](http://www.disit.org/dash)



### Service map browser

[Http://servicemap.disit.org](http://servicemap.disit.org)



### Collective User behavior Analyzer



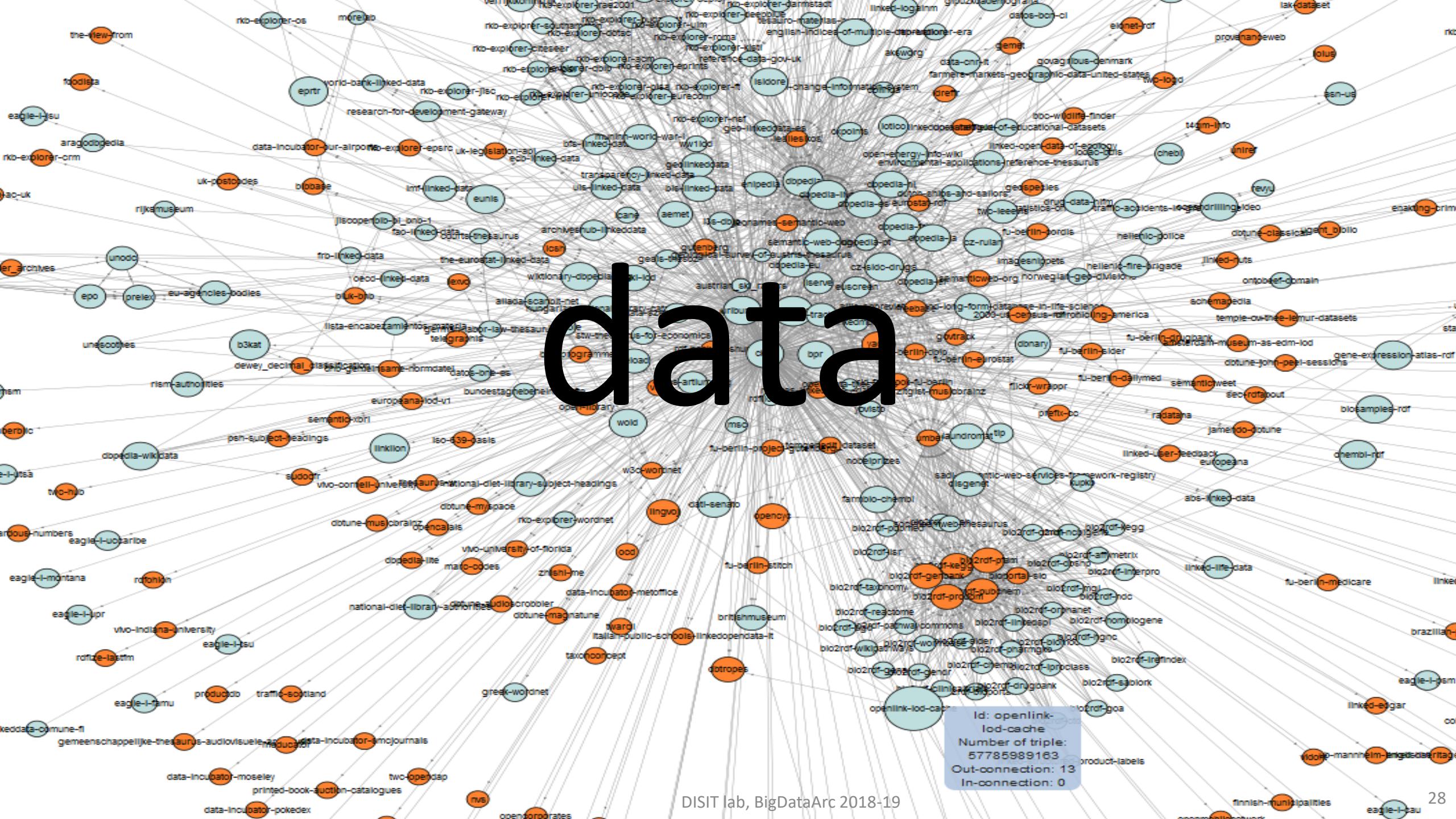
## Tools for Final Users

### Mobile e Web Apps

[Http://www.km4city.org](http://www.km4city.org)



# data



## Privati Statici

- Movimenti personali non pubblicati
- Relazioni personali non pubblicate

- comportamenti social media
- contributi consumi

## Privati Tempo reale

- Codice fiscale
- Foto non condivise
- Aspetti legali
- Cartella clinica
- ..

- Traffico personale
- Posizione mezzi,
- Parcheggi
- Posizione taxi
- Posizione CarSharing ...

## Pubblici statici (open data)

statistiche: incidenti, censimenti, votazioni

- Statistiche accessi alla ZTL
- Strutture pubbliche UNIFI

*posizione dei punti  
di interesse*

- Musei
- Strutture della città
- Servizi attivi

- Info traffico
- video camere
- Info Meteo
- Info Ambiente
- Code ai musei pubblici
- Terremoti
- Parcheggi

- Stato accessi alla ZTL
- Stato dei servizi

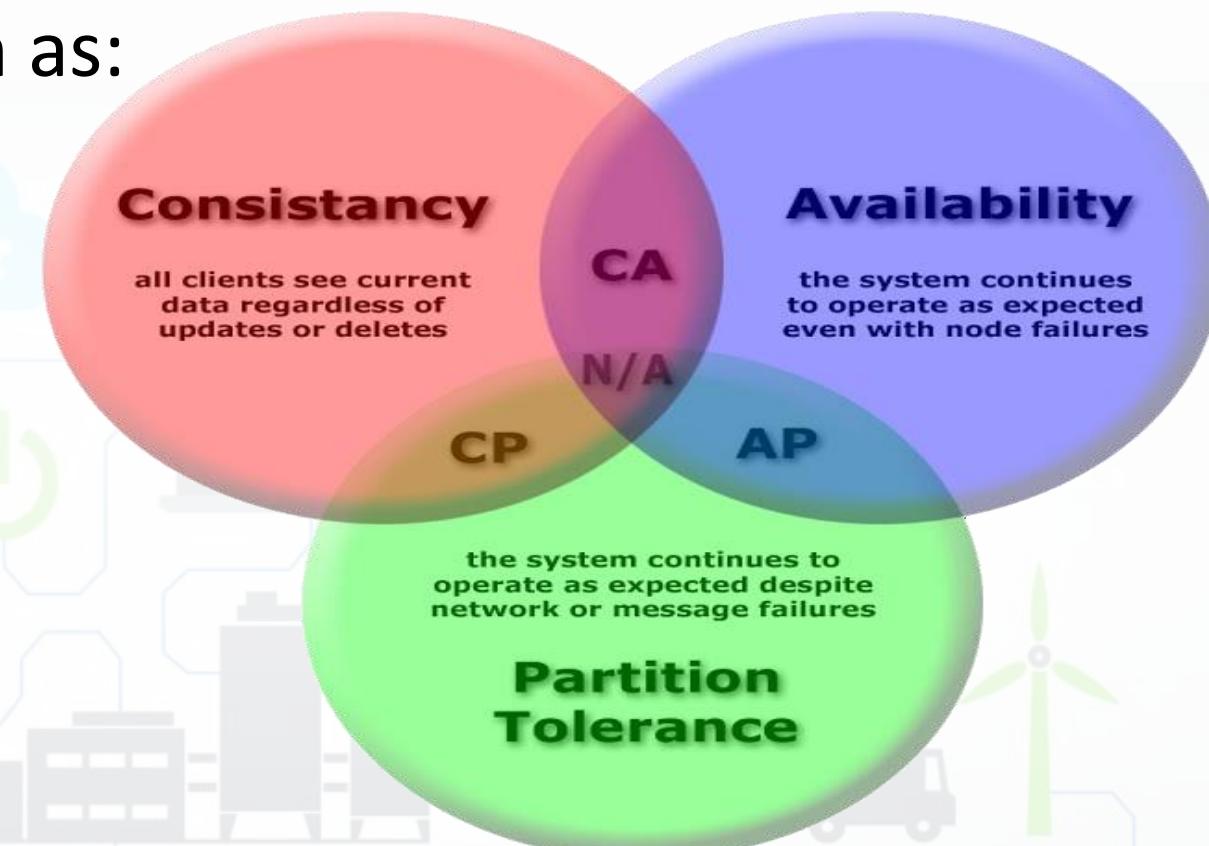
## Pubblici Tempo reale (open data)

# Complessità del Dato

- **Formati diversi, valori sparsi e discontinui anche in stream**
  - Data healthiness, integrity, etc.
- Tecniche: Data Lake per la normalizzazione del dato
  - → big data graveyards based on HDFS
- **Formati e dati non riferibili in modo preciso alla stessa semantica** delle entità in gioco: temperatur\*e\*, coordinat\*e\*, misure dei sensori in generale, ... → molto dipendenti dal contesto
  - Modelli ontologici → ontologie → knowledge base → expert systems, per ricerche in chiave semantica, Riconciliazione semantica, completamento, contestualizzazione, ...

# CAP theorem

- The **CAP theorem** (Consistency - Availability - Partition tolerance) is essential to **understand the behavior of distributed SW systems**, and **how to design the architecture** in order to meet stringent requirements, such as:
  - High **performance**.
  - Continued **availability**.
  - **Geographically distributed** systems.
- Working on billions and trillions of day, **scalability** became a key concept.



# Modeling data store

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

# Types of NoSQL Database

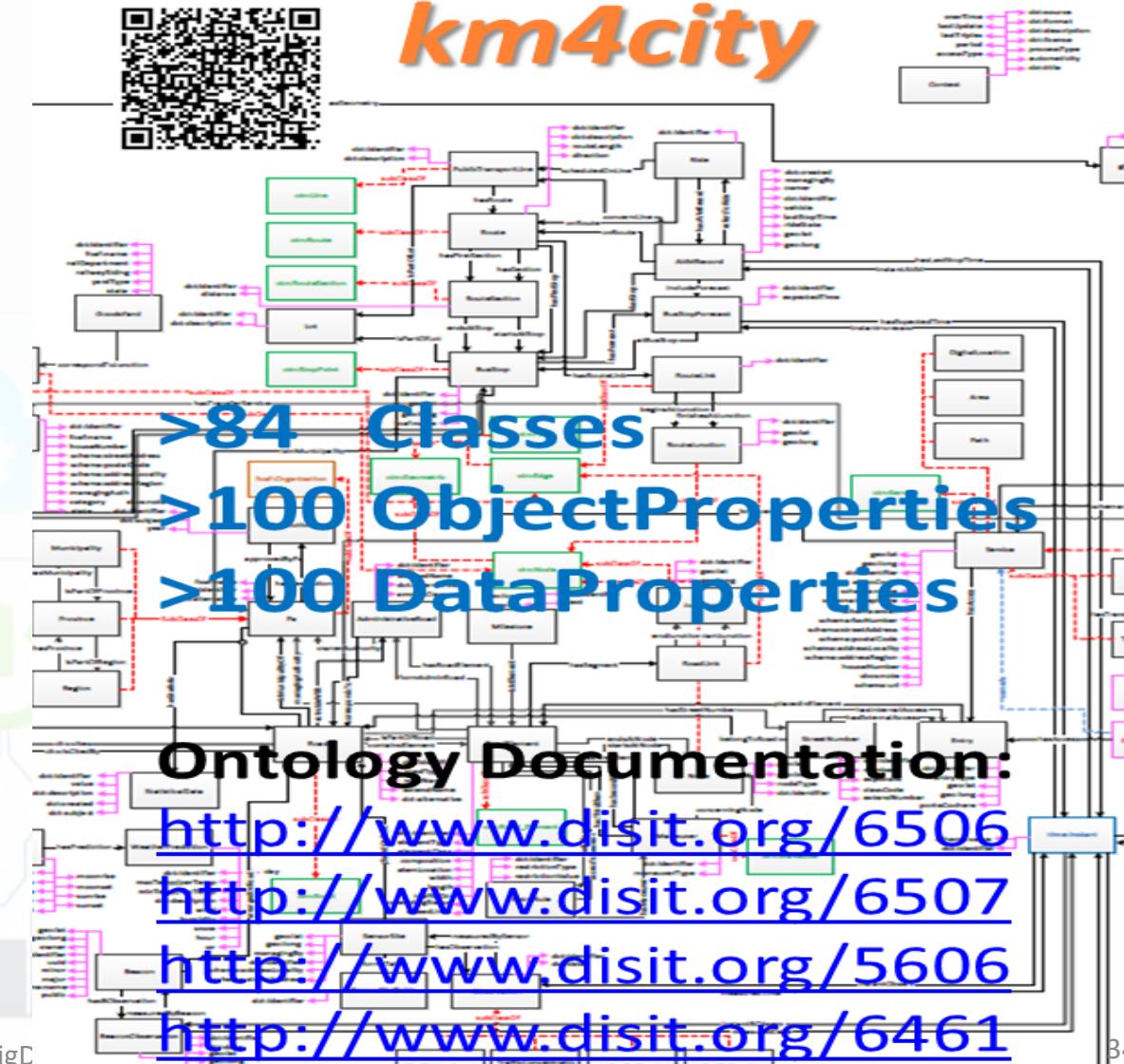
- Key-Value DB
- Col- Family/Big Table DB
- Document DB
- XML DB
- Object DB
- Multivalue DB
- ACID NoSQL
- Graph DB

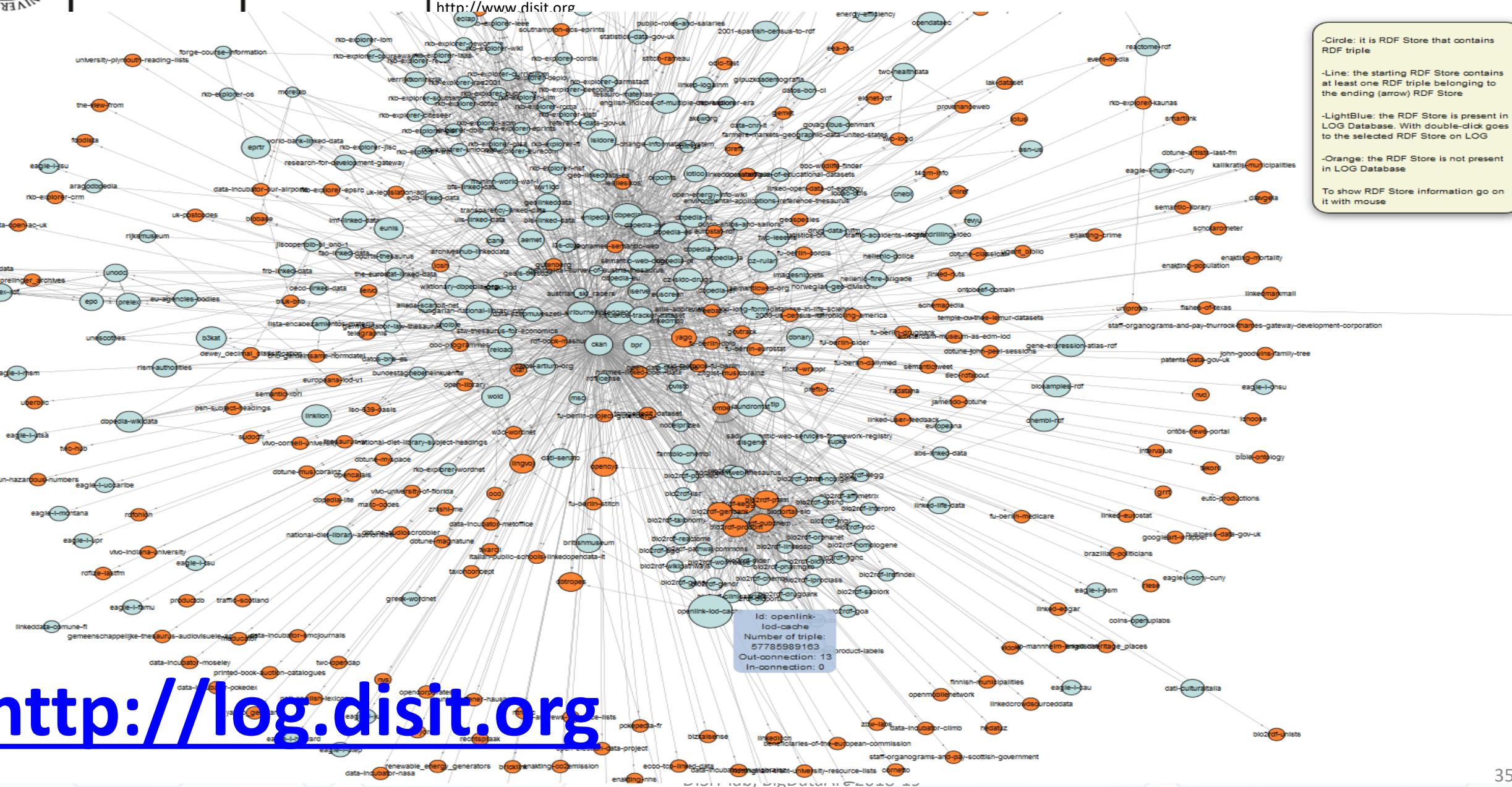


# I Dati

- **Collezioneamento dati statici, quasi statici e real time, stream**
  - **Dati open:** geo localizzati, servizi, statistiche, censimenti, etc.
  - **Dati privati degli operatori:** con licenze limitate per non permettere di fare profitto ad altri operatori sulla base dei loro dati
  - **Dati personali delle persone:** profili, comportamenti tramite APP, IOT, sensori, web, etc.
- **Integrazione dati per renderli *semanticamente interoperabili*, ed operare deduzioni (time, space... )**
  - I tradizionali **collettori di open data** danno visioni statistiche ma **non sono adatti a produrre servizi integrati**
  - **Integrazione con modelli semanticci unificanti come Km4City**

## Smart-city Ontology km4city





**Linked Open Graph****SiiMobility (by DISIT)**

Examples:

- [VIA GIACOMO MATTEOTTI](#)
- [Bagno a ripoli](#)
- [Florence](#)

Choose a class:

Search for keyword

keyword:

uri: **Your data**

sparql endpoint: (optional)

uri: **Type of relations**

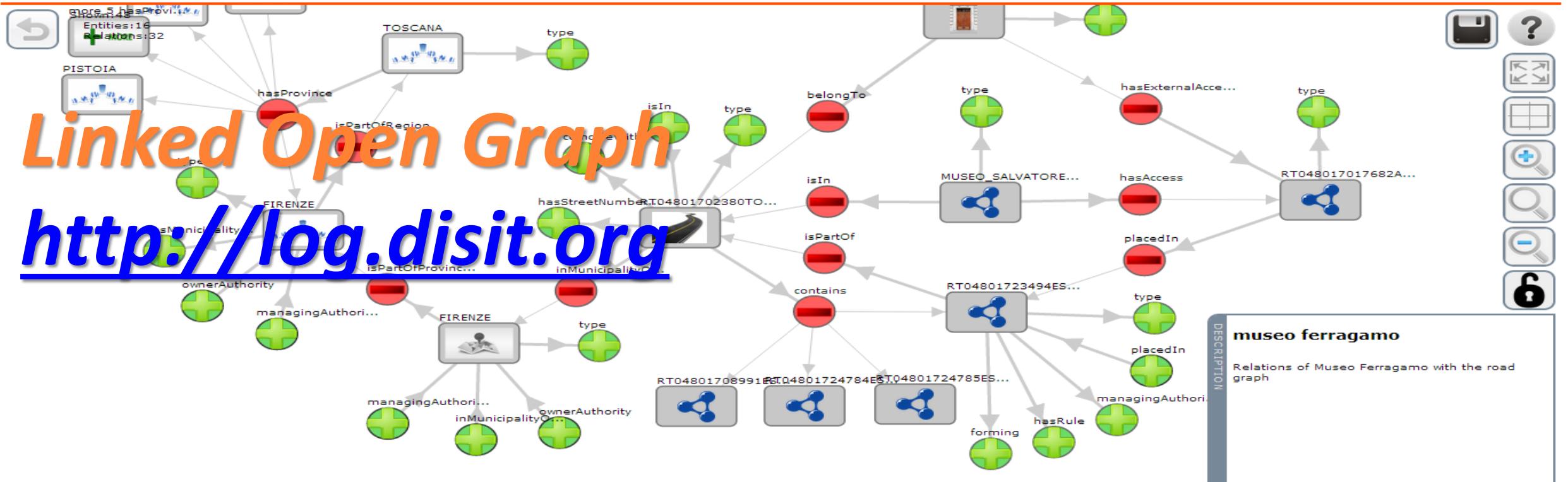
Select all

Deselect all

Invert

 Hide all inverse

- |   |  |
|---|--|
| <input checked="" type="checkbox"/> belongTo          | <input checked="" type="checkbox"/> coincideWith     |
| <input checked="" type="checkbox"/> contains          | <input type="checkbox"/> depiction                   |
| <input type="checkbox"/> ends                         | <input checked="" type="checkbox"/> forming          |
| <input type="checkbox"/> has                          | <input checked="" type="checkbox"/> hasAccess        |
| <input checked="" type="checkbox"/> hasExternalAccess | <input checked="" type="checkbox"/> hasMunicipality  |
| <input checked="" type="checkbox"/> hasProvince       | <input checked="" type="checkbox"/> hasRule          |
| <input checked="" type="checkbox"/> hasStreetNumber   | <input checked="" type="checkbox"/> inMunicipalityOf |
| <input checked="" type="checkbox"/> isIn              | <input checked="" type="checkbox"/> isPartOf         |
| <input checked="" type="checkbox"/> isPartOfProvince  | <input checked="" type="checkbox"/> isPartOfRegion   |
| <input checked="" type="checkbox"/> managingAuthority | <input checked="" type="checkbox"/> ownerAuthority   |
| <input checked="" type="checkbox"/> placedIn          | <input type="checkbox"/> sameAs                      |
| <input checked="" type="checkbox"/> seeAlso           | <input type="checkbox"/> starts                      |

**Linked Open Graph**



# Smart City

<http://www.km4city.org>

**Present data Tuscany Region April 2017**

**Road Graph (Tuscany region)**

- 132,923 Roads
- 389,711 Road Elements
- 318,160 Road Nodes
- 1,508,207 Street Numbers

**Info on:** points, paths, areas, etc.

**Services (20 cat, 512 cat.)**

- 16 Pub. Transport Operators
- 21.280 Bus stops & 1081 bus lines

**Dynamic/real-time in Tuscany Region**

- Real time bus lines: 144 updates X day X line
- 1081 Pub Lines: 1-2 updates per day, time and path
- 210 parking status: 76 updates X day X sensor
- 796 traffic Sensors: 288 updates X day X sensor
- 285 weather area: 2 updates X day X area
- 12 hospital Triage status: 96 updates X day X FA
- 1600 Fuel stations: 1 update X day X station
- 22 Environmental data: 20 updates X day X sensor
- Florence events: about 60 new events X day
- Wi-Fi: > 400.000 measures X day
- App mobiles: > 50.000 measures X day
- more than 40.000 distinct users X day
- From 600.000 to 4.5 M Tweets X day
- .....+ many IOT are coming .....

**Servizi Regolari | Servizi Trasversali**

search text into service

Categorie Servizi

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

N. risultati: Nessun Limite

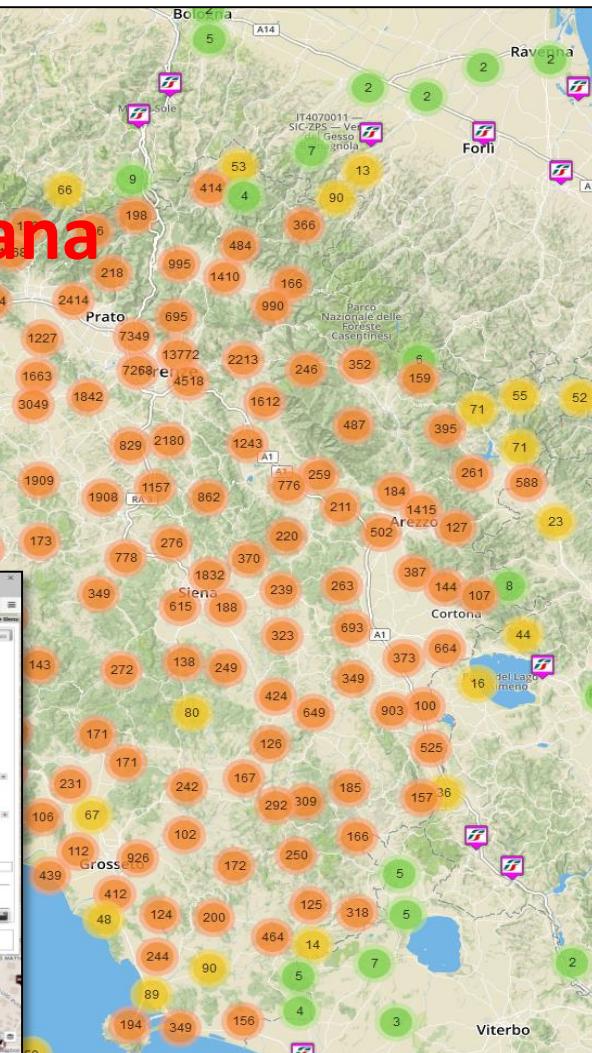
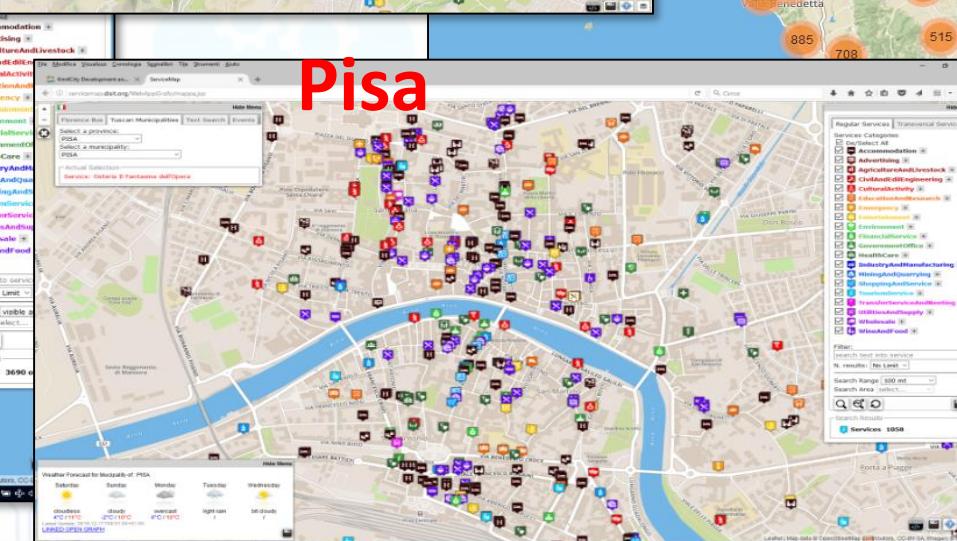
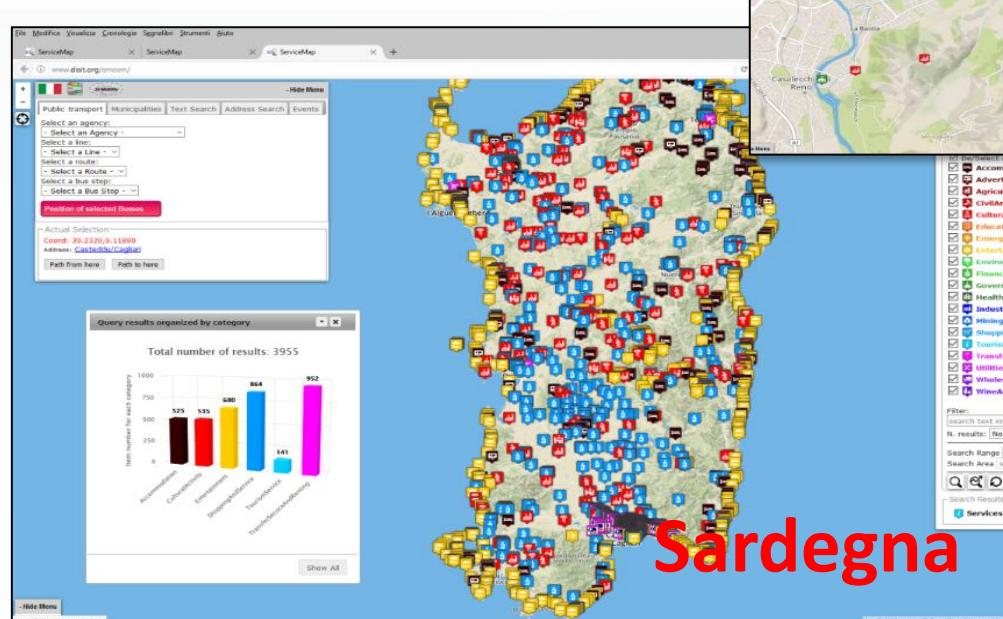
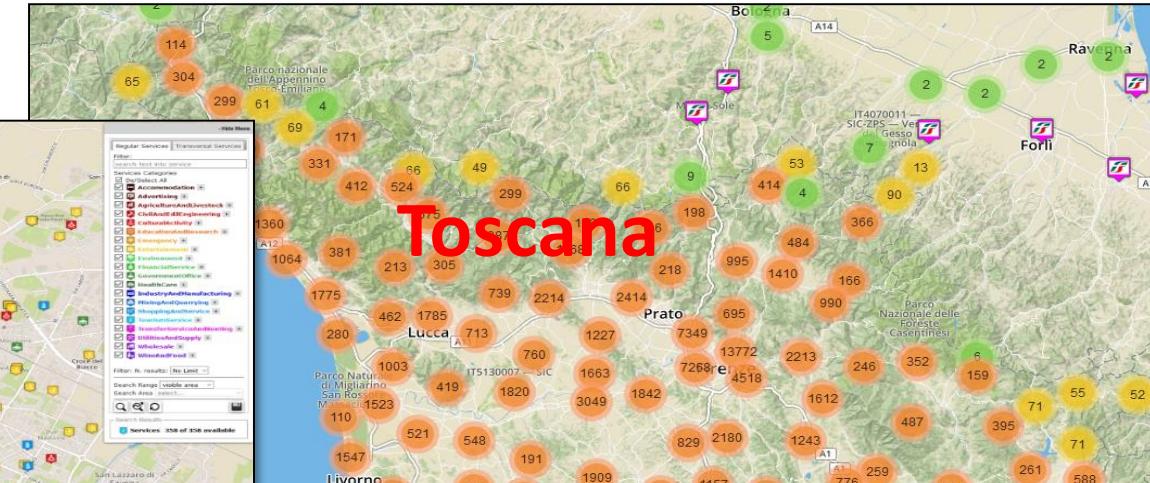
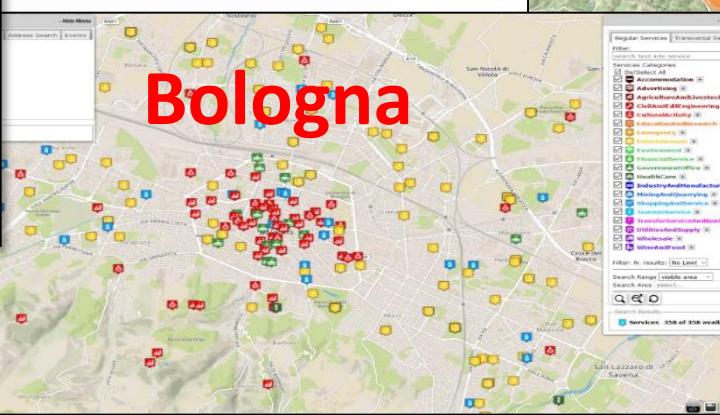
Raggio ricerca: 100 metri

Risultati della ricerca

più di 4000 risultati, attivato clustering

Services 16858

# Coverage: examples



- Search all services in the area

**- Nascondi Menu**

Fermate Firenze Comuni in Toscana Ricerca Testuale

Selezione una provincia:  
FIRENZE

Selezione un comune:  
FIRENZE

Actual Selection  
COMUNE di FIRENZE

**KM 4 CITY**

Previsioni meteo FIRENZE:  
Giovedì: poco nuvoloso 23°C / 27°C  
Venerdì: poco nuvoloso 20°C / 33°C  
Sabato: velato /

<https://servicemap.km4city.org>

**What is enabling and providing smart services**

- Smart Parking, in Tuscany
- Smart First Aid in Tuscany
- Smart search for POI and public transport srv.
- Public Transportation in Tuscany
- Routing in Tuscany, simple and multimodal
- Social Media Monitoring and acting
- Traffic events and Resilience in Florence
- Bike Sharing in Pisa and Siena
- Recharge stations for e-vehicles
- Entertainment Events in Florence
- Traffic Sensors in Tuscany
- IOT/IOE sensors and actuators
- Weather forecast/condition in Tuscany
- Pollution and Pollination in Tuscany
- People Monitoring, in Tuscany via App
- ..People Monitoring Assessment in the City, in Florence via Wi-Fi

**- Nascondi Menu**

Servizi Regolari Servizi Trasversali

search text into service

Categorie Servizi

De>Select All

Accommodation +

Advertising +

AgricultureAndLivestock +

CivilAndEdilEngineering +

CulturalActivity +

EducationAndResearch +

Emergency +

Entertainment +

Environment +

FinancialService +

GovernmentOffice +

HealthCare +

IndustryAndManufacturing +

MiningAndQuarrying +

ShoppingAndService +

TourismService +

TransferServiceAndRenting +

UtilitiesAndSupply +

Wholesale +

WineAndFood +

N. risultati: Nessun Limite

Raggio ricerca 100 metri

Risultati della ricerca

più di 4000 risultati, attivato clustering

Services 16858

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

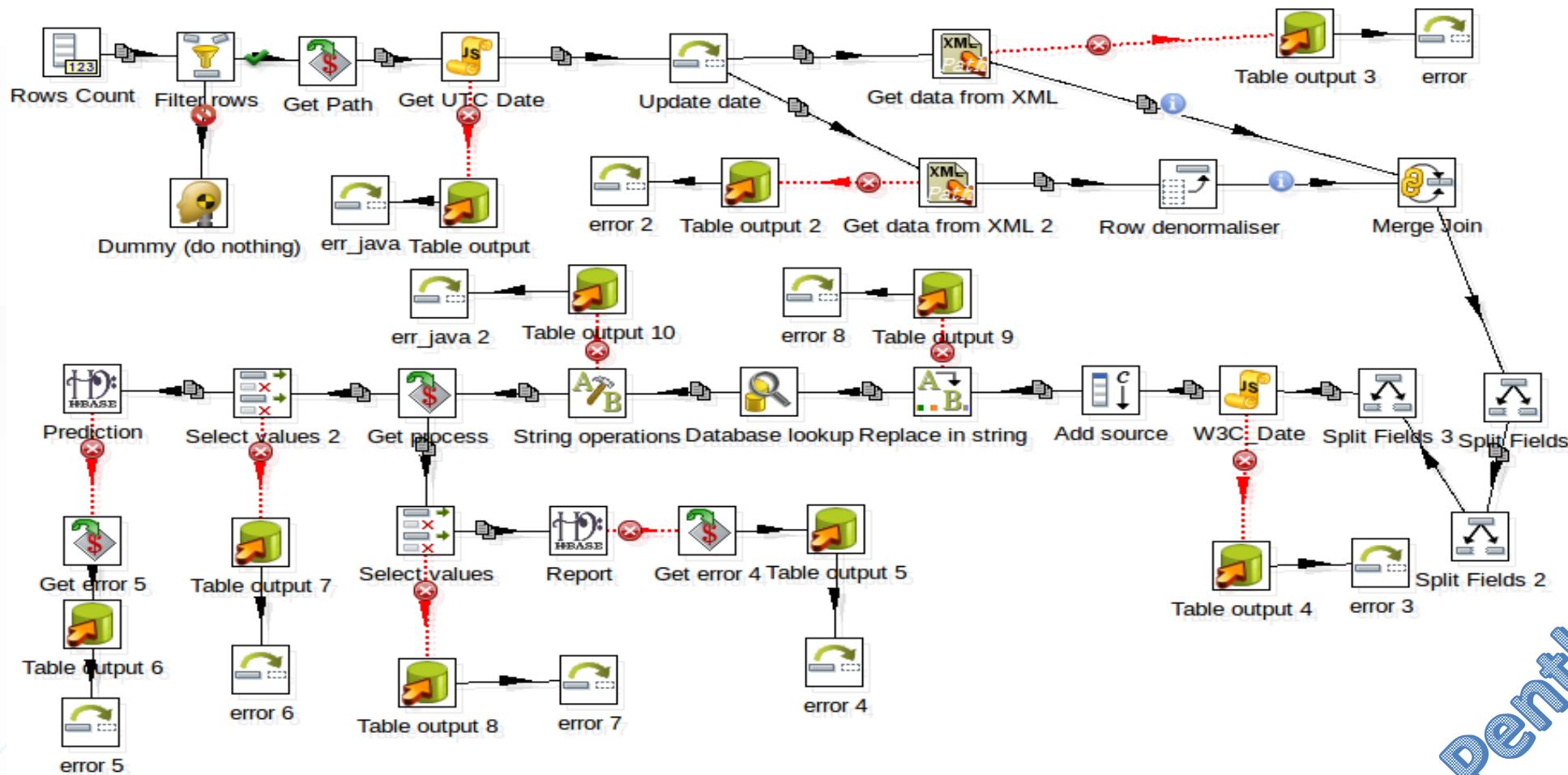
Embed

39

# gathering

lod-cache  
umber of triple:  
57785989163  
ut-connection: 13

# Example of ETL



## Batch Processing

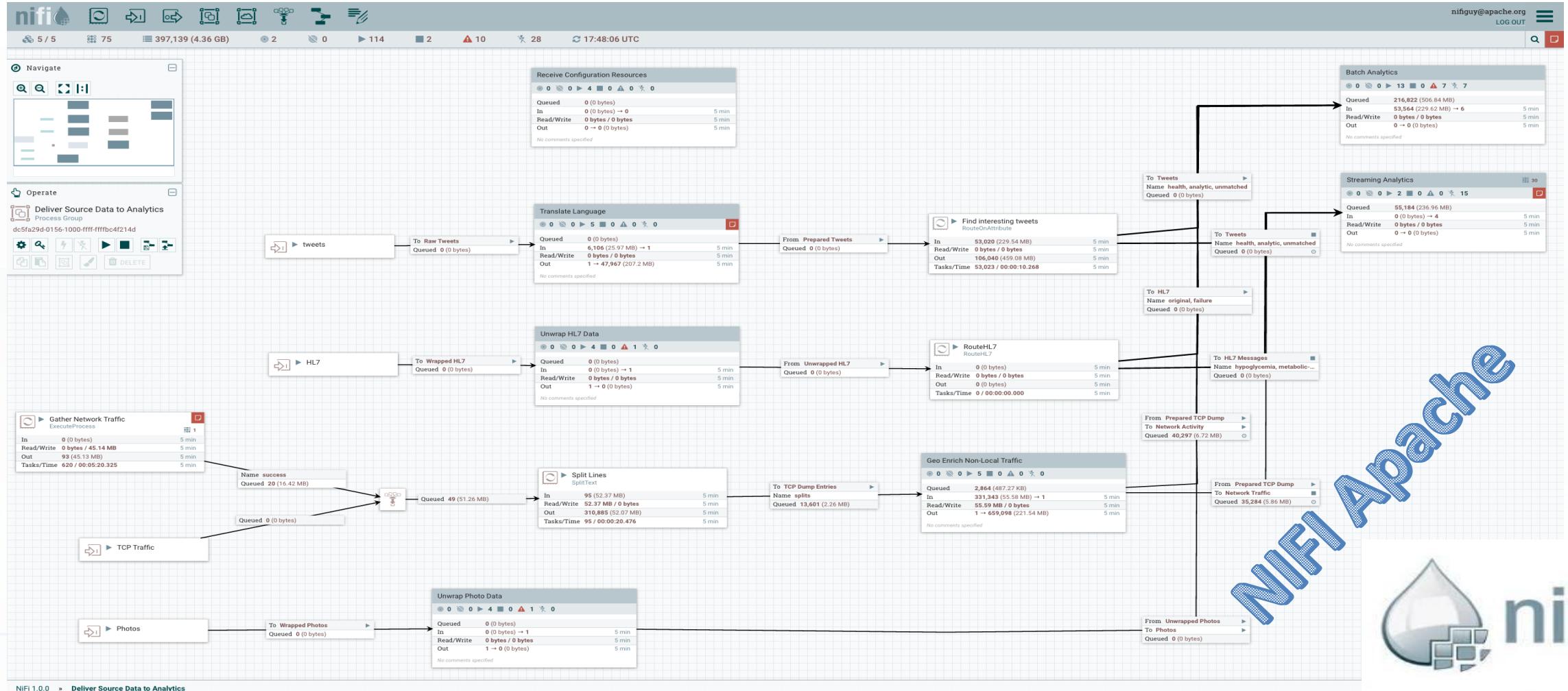
ETL Pentaho Kettle



kettle™  
pentaho data integration



# Example of NIFI



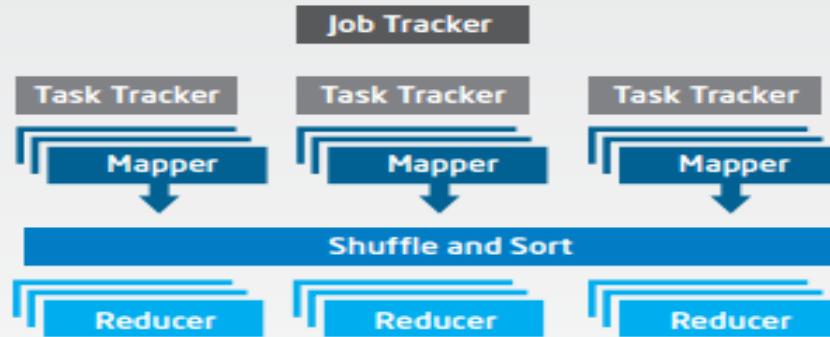
## Batch Processing

# Processing

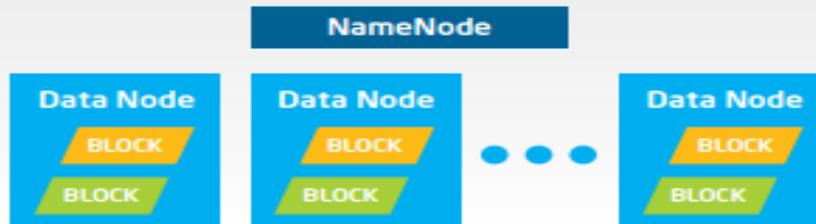
**Id:** opentlink-  
**Iod-cache**  
**Number of triple:**  
57785989163  
**Out-connection:** 13  
**In-connection:** 0

### LOGICAL ARCHITECTURE

#### Processing: MapReduce

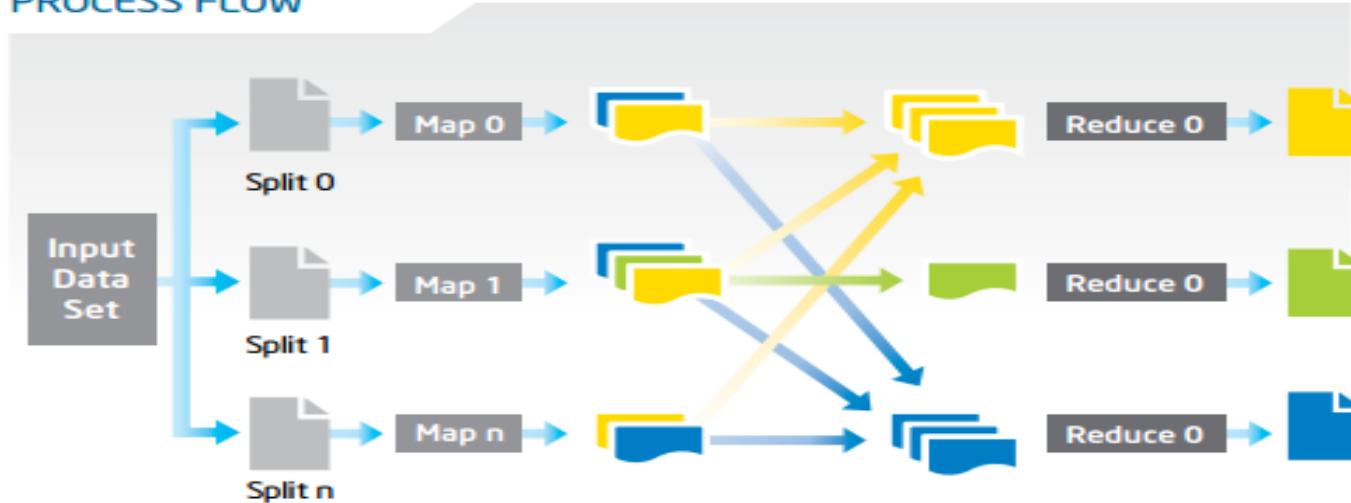


#### Storage: HDFS



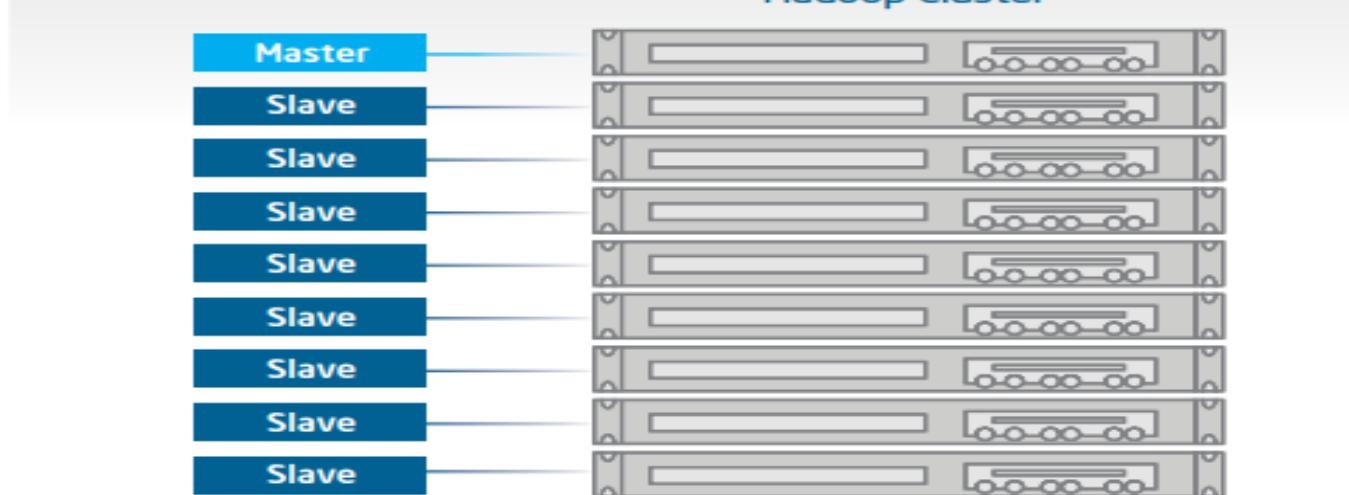
# Hadoop and MapReduce

### PROCESS FLOW



### PHYSICAL ARCHITECTURE

#### Hadoop Cluster





# Developer in R Studio + Tensor Flow

The screenshot shows the Snap4City platform interface with the R Studio Development window open. The R Studio window displays an R script named 'AnomalyDetection.R' which processes sensor data to detect anomalies. The script uses the tidyverse library for data manipulation and ggplot2 for visualization. It reads data from a SQLite database, performs statistical analysis, and generates a plot for anomalies. The RStudio environment tab shows variables like 'dataFinal', 'dataset', and 'statisticsResult'. Below the R Studio window is a file browser showing local files like 'nohp.out' and various 'Snap4City' folder structures.

```
[1] "carpark"
Warning in statisticsResult[indfolder]$statisticsOutputName = unbox ("Predictions") :
  number of items to replace is not a multiple of replacement length
Warning in statisticsResult[indfolder]$statisticsOutputName = unbox ("MachineLearningPredictions") :
  number of items to replace is not a multiple of replacement length
`geom_smooth()` using method = 'loess'
[1] "carpark"
Warning in statisticsResult[indfolder]$statisticsOutputName = unbox ("Anomalies") :
  number of items to replace is not a multiple of replacement length
[1] "NO ANOMALIES ON THE SENSOR - CarParkBecattaria_free."
[1] "PRESENCE OF ANOMALIES ON THE SENSOR - CarParkCareggi_free."
[1] "PRESENCE OF ANOMALIES ON THE SENSOR - CarParkPieracciniMeyer_free."
[1] "NO ANOMALIES ON THE SENSOR - CarParkS.Lorenzo_free."
[1] "NO ANOMALIES ON THE SENSOR - CarParkStazioneFirenzeS.M.N._free."
[1] "carpark"
Warning in statisticsResult[indfolder]$statisticsOutputName = unbox ("Anomalies") :
  number of items to replace is not a multiple of replacement length
[1] "NO ANOMALIES ON THE SENSOR - CarParkBecattaria_free."
[1] "PRESENCE OF ANOMALIES ON THE SENSOR - CarParkCareggi_free."
[1] "PRESENCE OF ANOMALIES ON THE SENSOR - CarParkPieracciniMeyer_free."
[1] "NO ANOMALIES ON THE SENSOR - CarParkS.Lorenzo_free."
[1] "NO ANOMALIES ON THE SENSOR - CarParkStazioneFirenzeS.M.N._free."
[1] "carpark"

anomaliesMat[, "timestamp"] <- as.character(dataFinal$resAnoms$index , "alignDateTime")
anomaliesMat[, "anoms"] <- as.numeric(res$anoms[, "anoms"])

#table with anomalies

setwd(outWD)
options(digits = 1)
tBTable <- tableGrob(anomaliesMat, rows = NULL, cols = c("Date and Time", "Anomaly"), theme=ttheme_default(base_size=10))
grid.draw(tBTable)
h <- convertHeight(sum(tBTable$heights), "in", TRUE)
w <- convertWidth(sum(tBTable$widths), "in", TRUE)

plot <- res$pplot

plotMix <- grid.arrange(plot, tBTable,
                        ncol = 2,
                        heights=c(5,1),
                        as.table=TRUE)

setwd(outWD)
ggsave(paste(columnsName[i], "Anomalies.png", sep=""), plotMix, width=22, height=h*5)

}, finally = {
})

statisticsResult[[indfolder]]$resultFiles[[indResult]]$sensor=NULL
statisticsResult[[indfolder]]$resultFiles[[indResult]]$sensor=unbox(as.character(columnsName[i]))
statisticsResult[[indfolder]]$resultFiles[[indResult]]$png=unbox(paste(outWD, paste(columnsName[i], "Anomalies.png", sep=""), sep=""))
inhdResult = inhdResult + 1

} else{
  print(paste("NO ANOMALIES ON THE SENSOR ", "-", columnsName[i], "-", sep=""))
}

}

setwd("~/Snap4City")
write(jsonlite::toJSON(statisticsResult[[1]]), "JsonStatisticsResult.json")
return(statisticsResult[[1]])
}

144.4 anomalyDetection(anomalyDate) : R Script
```

The screenshot shows a dashboard interface with the following components:

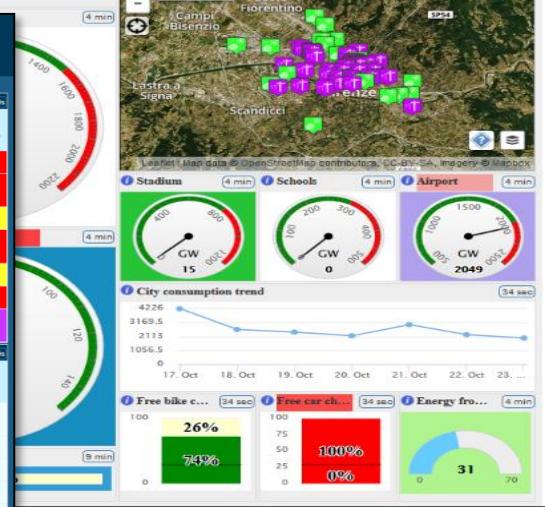
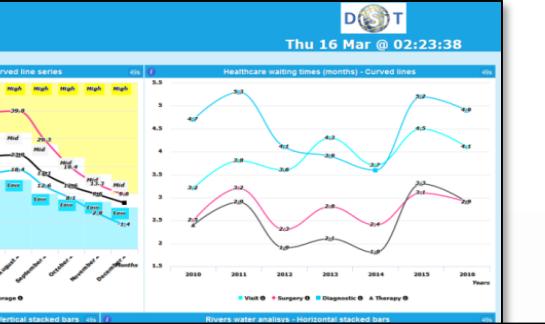
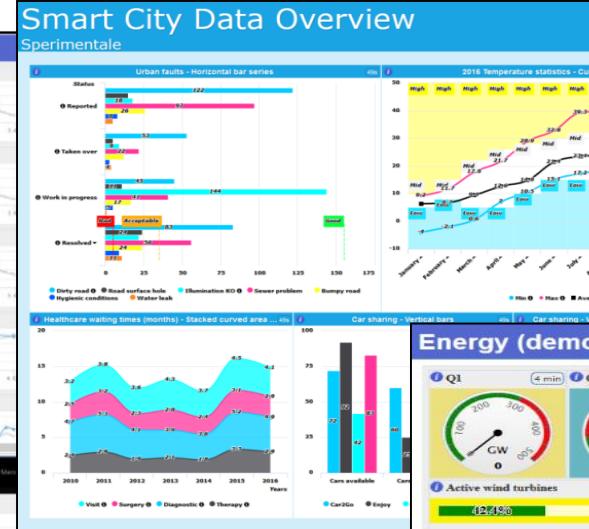
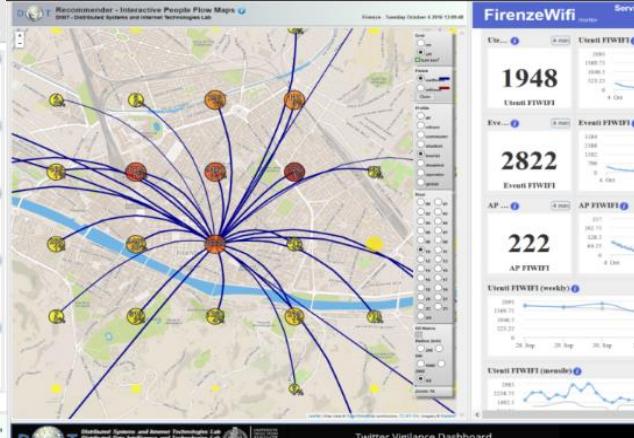
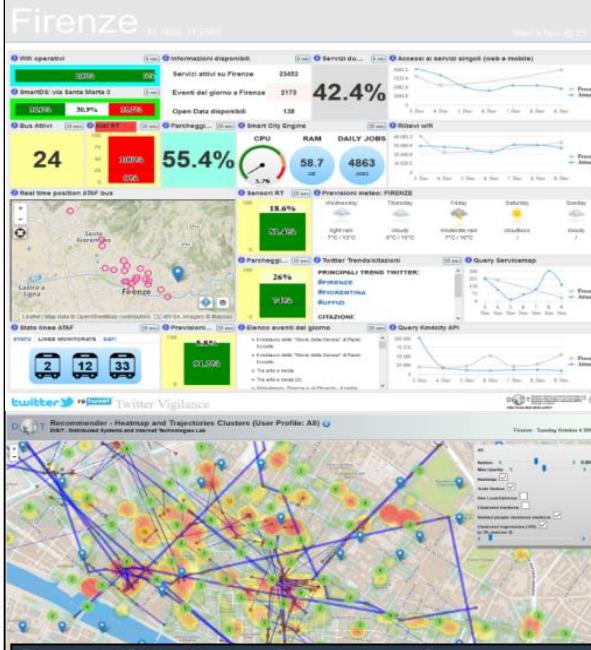
- File List:** A sidebar on the left lists several files as .png files:
  - AverageSpeedDailyTrend.png
  - CarParksDailyTrend.png
  - CorrelationMatrix.p
  - PredictedFreeParking.png
  - SensorsMeanPerDayMoment.png
  - StatisticsBySensors.png
  - StatisticsBySensorsAndDayMoment.png
  - VehicleFlowDailyTrend.png
- Line Chart:** A large chart area displays multiple line graphs representing data trends over time.
- Correlation Matrix:** A heatmap showing the correlation between various sensors and parking metrics.
- 3D Visualization:** A 3D bar chart with a green '3D' logo in front of it.

A red box highlights the first item in the file list, "AverageSpeedDailyTrend.png", with a red arrow pointing to it. A red border surrounds the entire file list area. A red callout box with the text "Click on each .png file to visualize the statistics: a new tab will be opened" is positioned above the file list.

# Rendering visual analytics

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

# Dashboard vs Business Intelligence



## Smart City Data Overview 2



## Florence data overview

A table based overview over city main data

Air Quality Index					Weather stations					Citizens satisfaction index							
Substances / Quarters	OZn	PM2+	PM10+	CO2+	NO2+	Data / Station	Wind speed (km/h) +	Direction	Temperature (°C) +	Humidity (%) +	Rain today (mm) +	Pressure (mbar) +	Criteria / Services	Quality (%) +	Cost (%) +	Availability time (%) +	Emergency handling (%) +
Q1	120	41	165	36	4	Sesto Fiorentino	50	N	12	0	922	Water	92	67	95	42	
Q2	33	25	66	123	45	Livorno	65	NE	17	0	876	Public transportation	36	29	27	31	
Q3	225	153	342	193	217	Grosseto	78	E	22	0	1022	Public safety	77	64	58	62	
Q4	174	221	87	122	93	Vada	42	S	6	34	895	Roads management	28	42	27	25	
Q5	79	87	23	27	65	Follonica	102	N	7.2	23	913	Healthcare	72	64	23	56	
						Giglio	97	O	3	19	957	Welfare	43	51	38	36	
												Public administration	58	16	18	22	

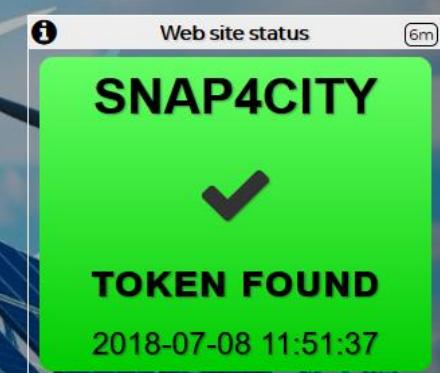
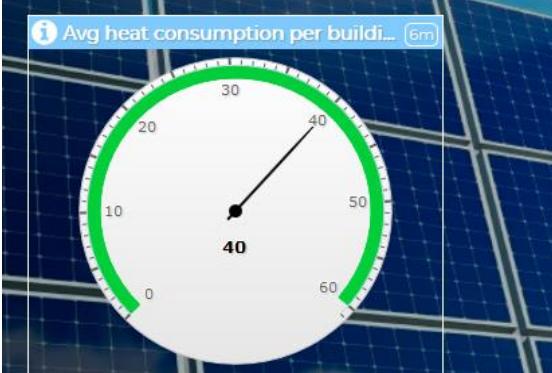
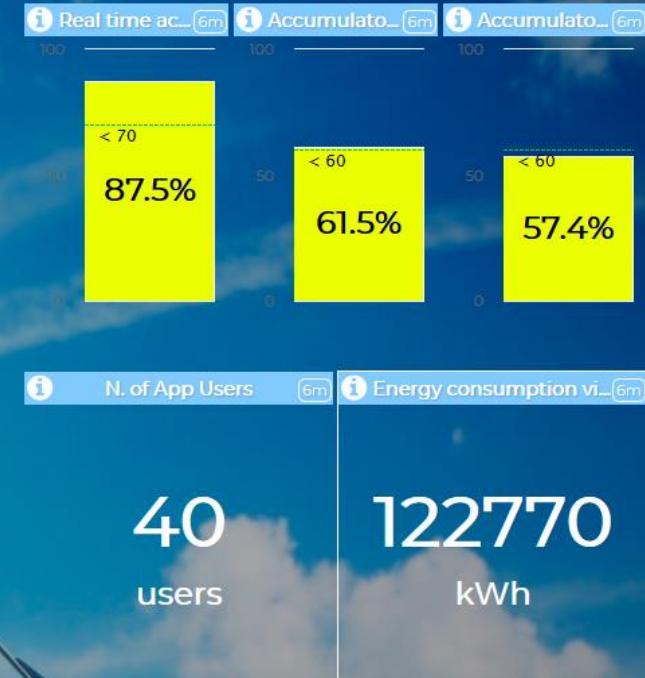
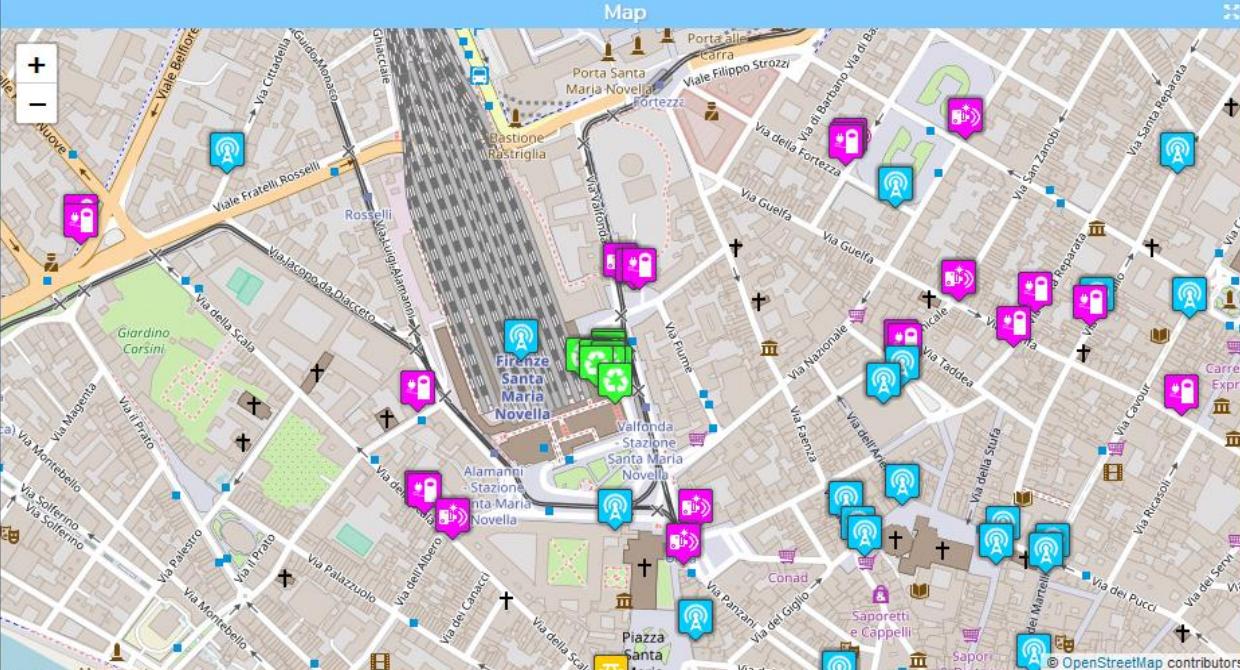
Tourists flow			Florence events 2017 overview												
Categories / Vehicle	Total arrivals	Overnights	Day trippers	Fields / Categories	Free	Paid	Winter	Spring	Summer	Autumn					
Airplane	56	36	20	Classical music, opera, ballet	7	23	6	10	4	10					
Train	122	81	41	Exhibitions	4	16	3	7	6	4					
Car	215	133	82	Guided tours	60	140	15	100	50	35					
Bus	157	110	47	Film festival	0	0	0	0	0	0					
Cruise	0	0	0	Markets, fairs	7	7	2	6	2	4					
Boat	0	0	0	Readings, conferences	35	15	10	22	9	9					
Total	550	360	190	Contemporary music	30	42	8	25	30	9					
				Sport	20	192	55	104	27	26					

# Energy Dash

Pilot dashboard



Sun 8 Jul 11:59:06



# Data Kinds vs Widgets: Special Widgets

- Complex Event
- External Service
- MicroApplication
- **Special Widget**
- POI, Point of Interest
- KPI, Key Performance Indicator
- Sensor
- Sensor-Actuator
- My Personal Data
- Dashboard-IOT App



Florence main first aids status					
Priority Hospitals	Red code	Yellow code	Green code	Blue code	White code
PS AO CAREGGI	7	9	43	25	0
PS SAN GIOVANNI DI DIO TORREGALLI	1	6	20	6	0
PS SANTA MARIA ANNUNZIATA	0	8	10	8	1
PS SANTA MARIA NUOVA	1	6	17	5	0

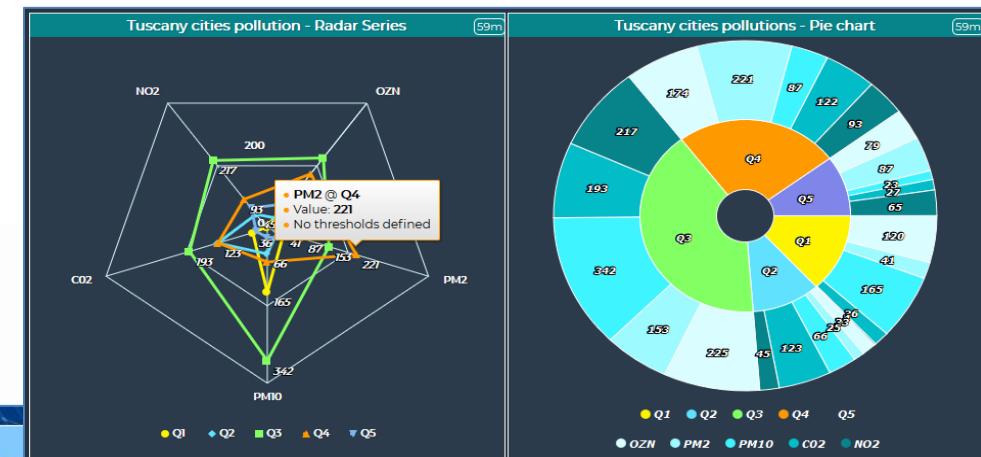
# Data Kinds vs Widgets: KPI

- Complex Event
- External Service
- MicroApplication
- Special Widget
- POI, Point of Interest
- **KPI, Key Performance Indicator**
- Sensor
- Sensor-Actuator
- My Personal Data
- Dashboard-IOT App

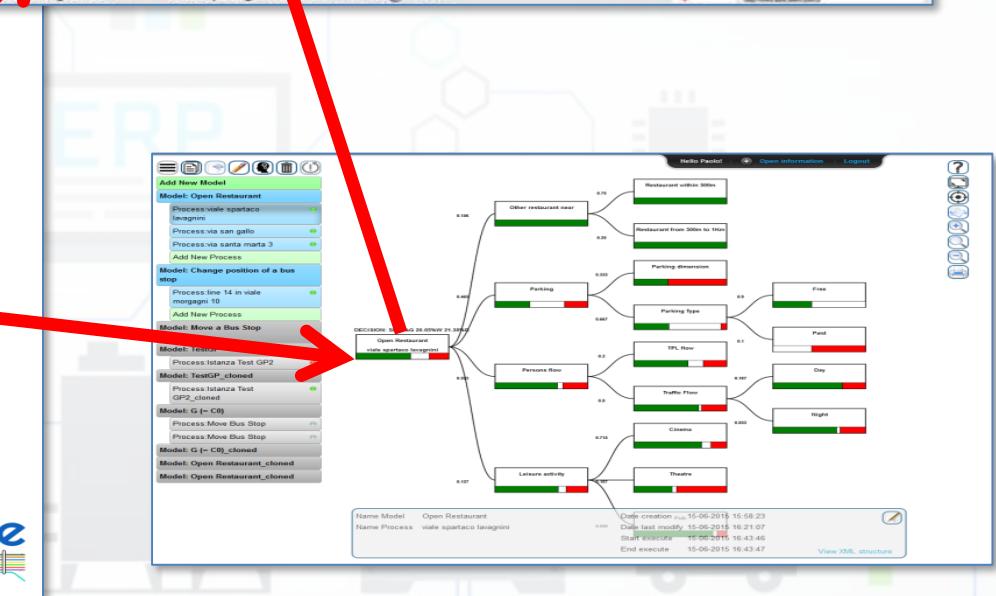
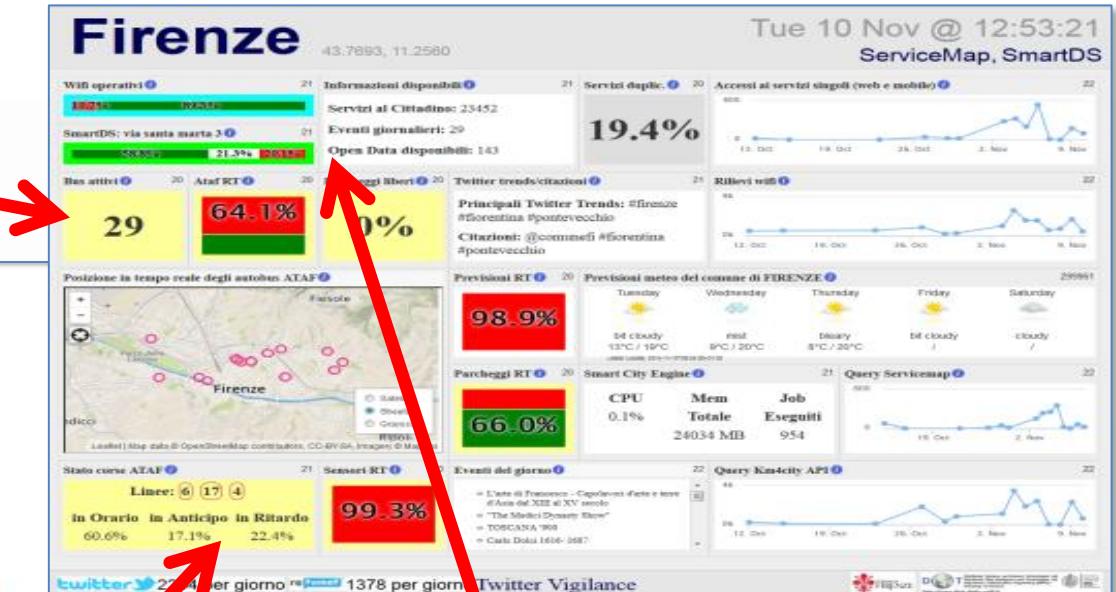
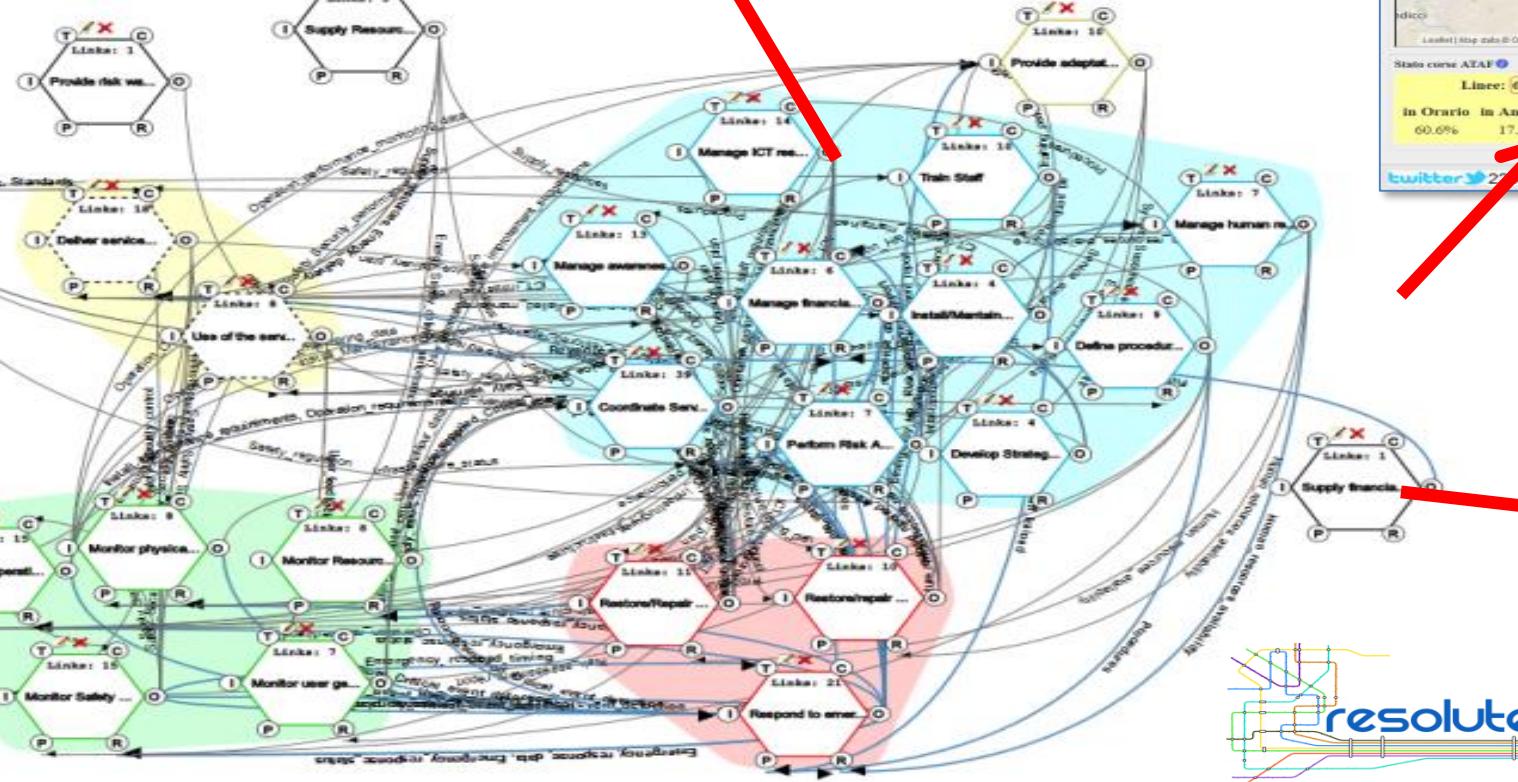
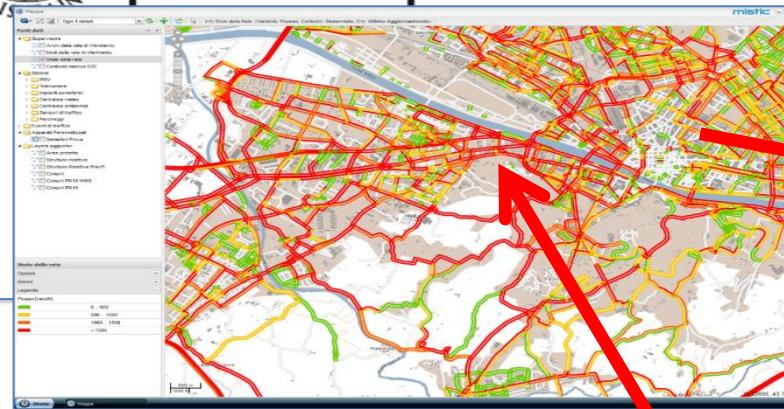
Real Time  
Event Driven  
Historical Data



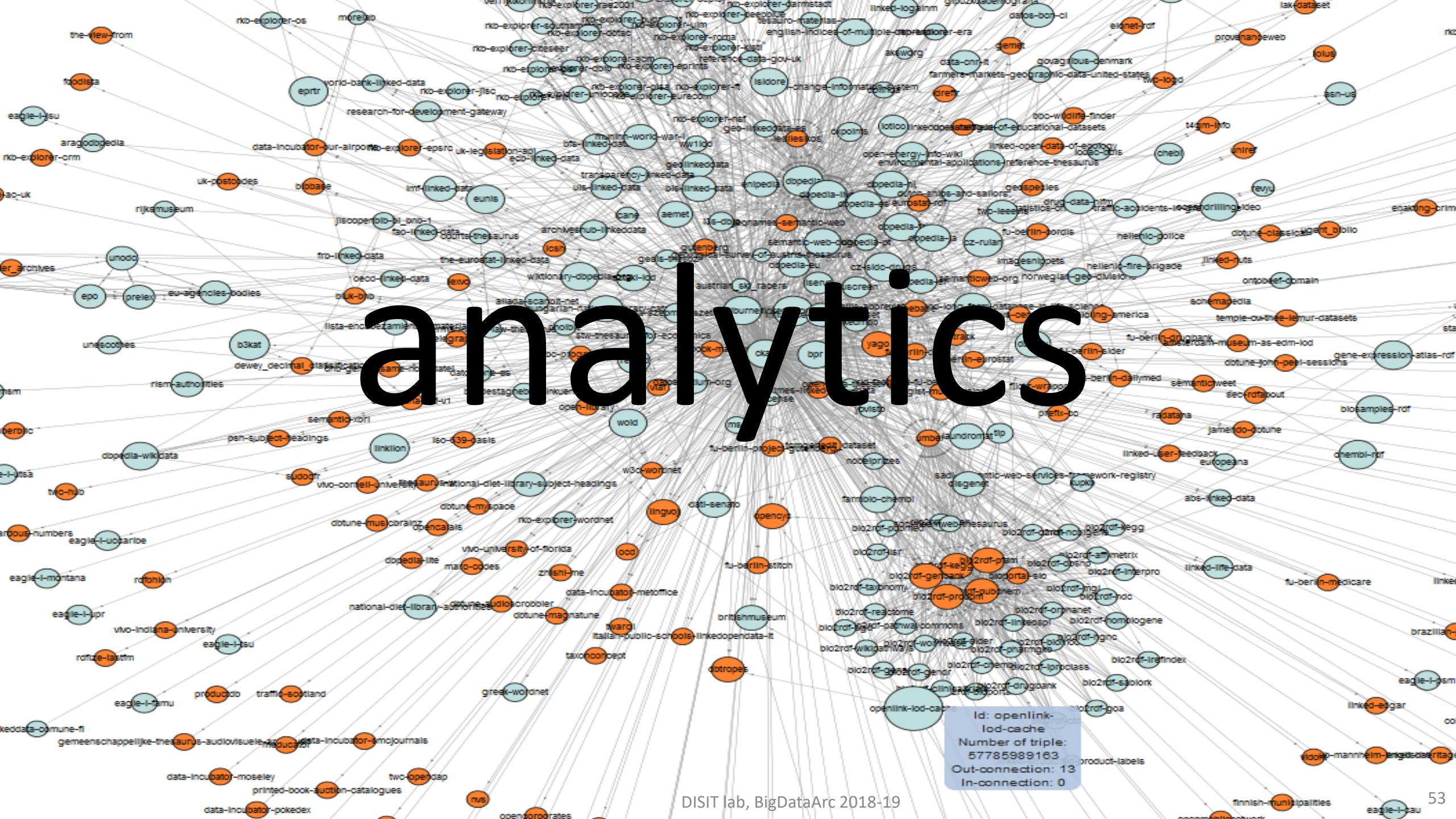
Weather stations						
Data / Station	Wind speed (km/h)	Direction	Temperature (°C)	Humidity (%)	Rain today (mm)	Pressure (mbar)
Sesto Fiorentino	50	N	7	32	0	922
Livorno	65	NE	5.5	37	0	876
Grosseto	78	E	4	22	0	1022
Vada	42	S	6	0	34	895
Follonica	102	N	7.2	23	0	913
Giglio	97	0	3	19	0	957



# Dashboarding city resilience



# analytics



# *Big Data analytics Area*

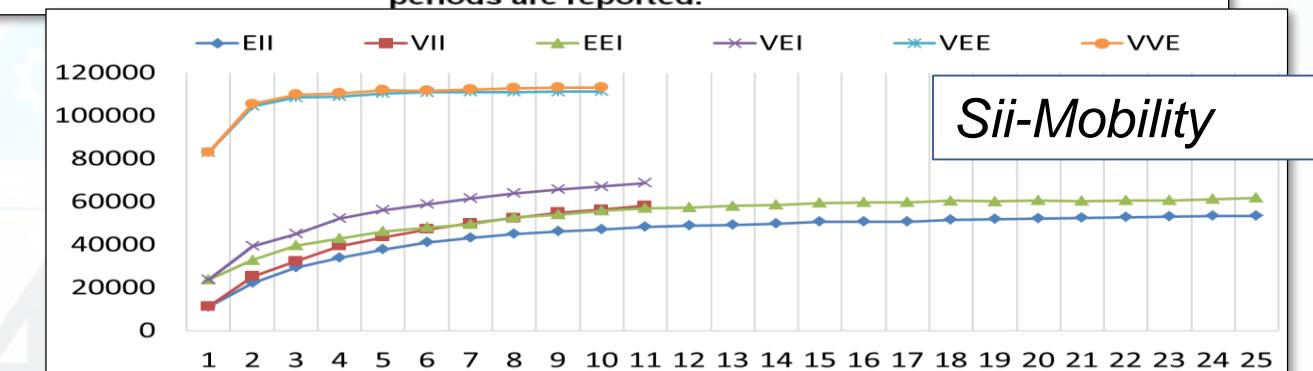
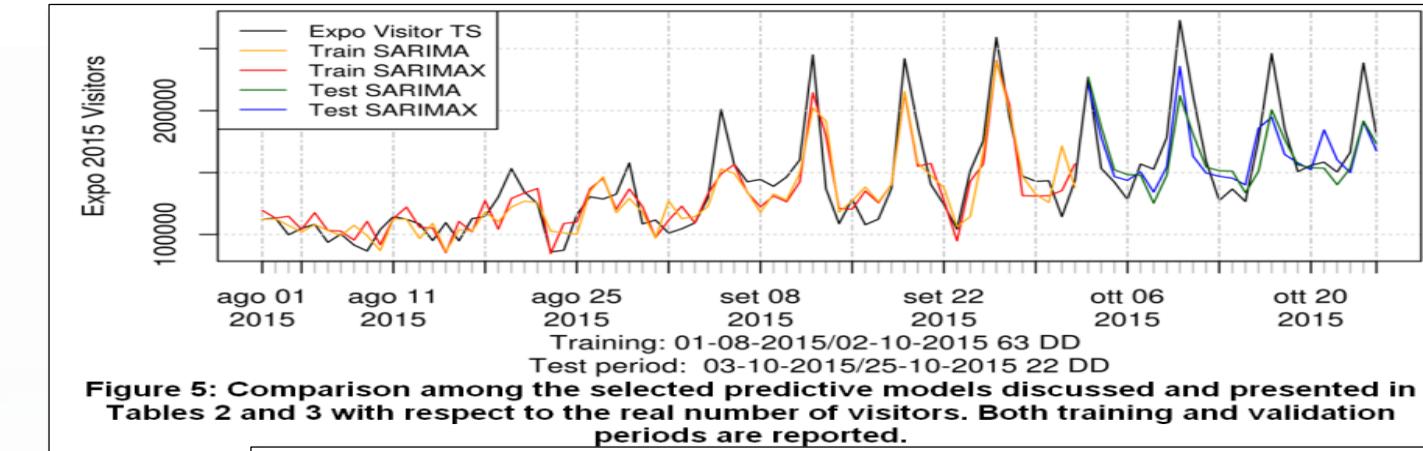
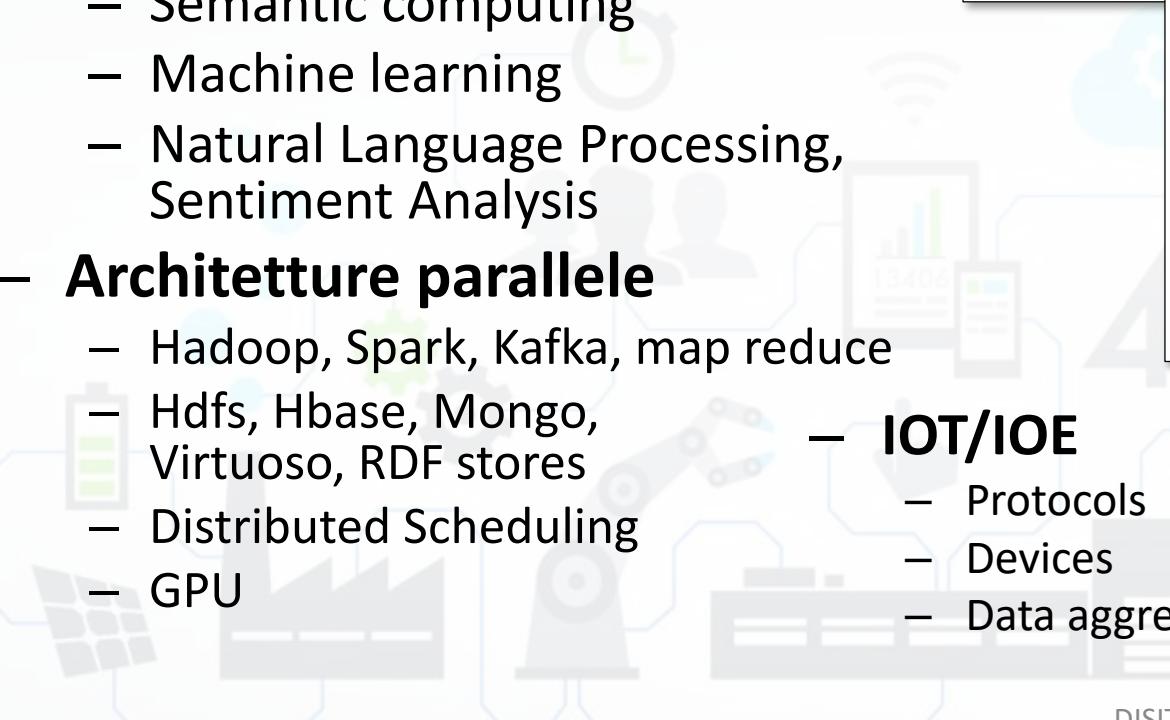
## *Applicative DISIT Lab*

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

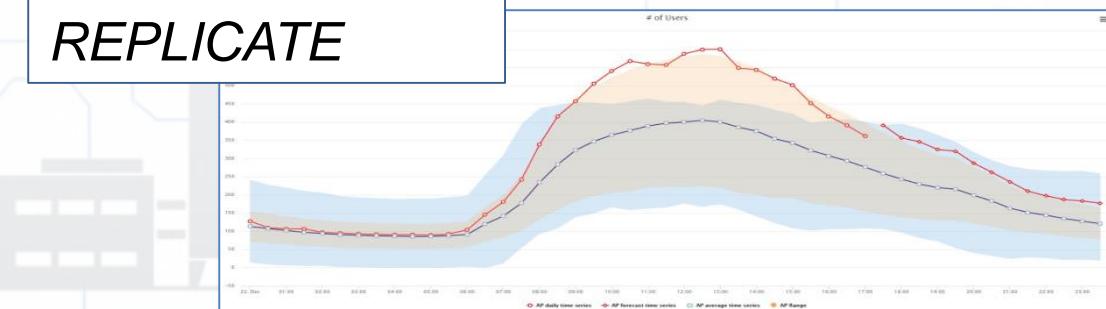
- Solutions for
  - Predicting models
  - Early detection
  - Anomaly detection

- Data Analytics
  - Data mining, Clustering
  - Semantic computing
  - Machine learning
  - Natural Language Processing, Sentiment Analysis

- Architetture parallele
  - Hadoop, Spark, Kafka, map reduce
  - Hdfs, Hbase, Mongo, Virtuoso, RDF stores
  - Distributed Scheduling
  - GPU

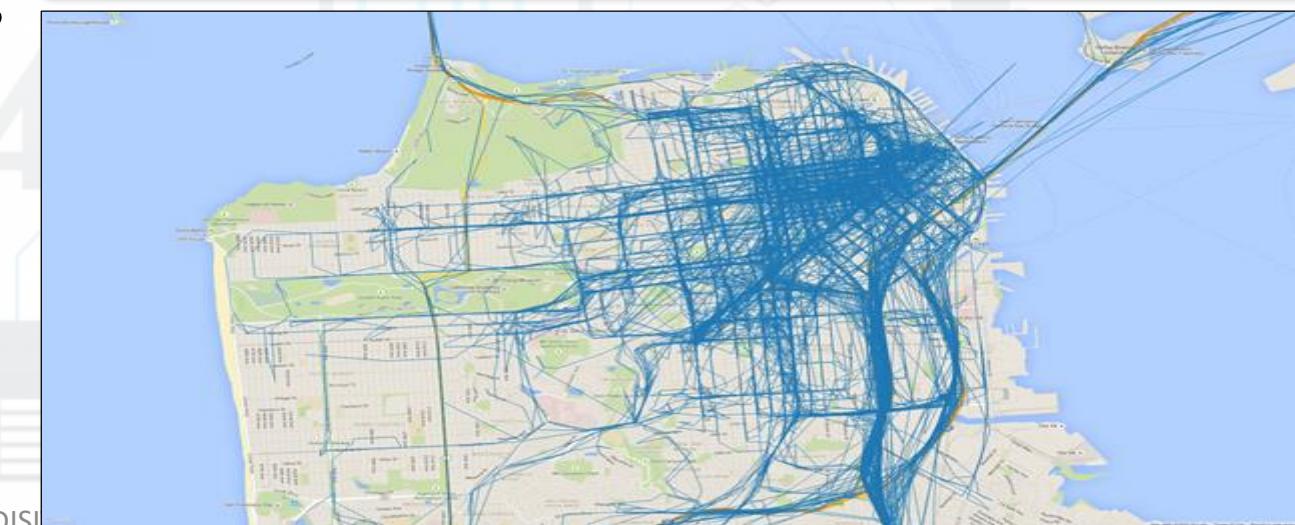
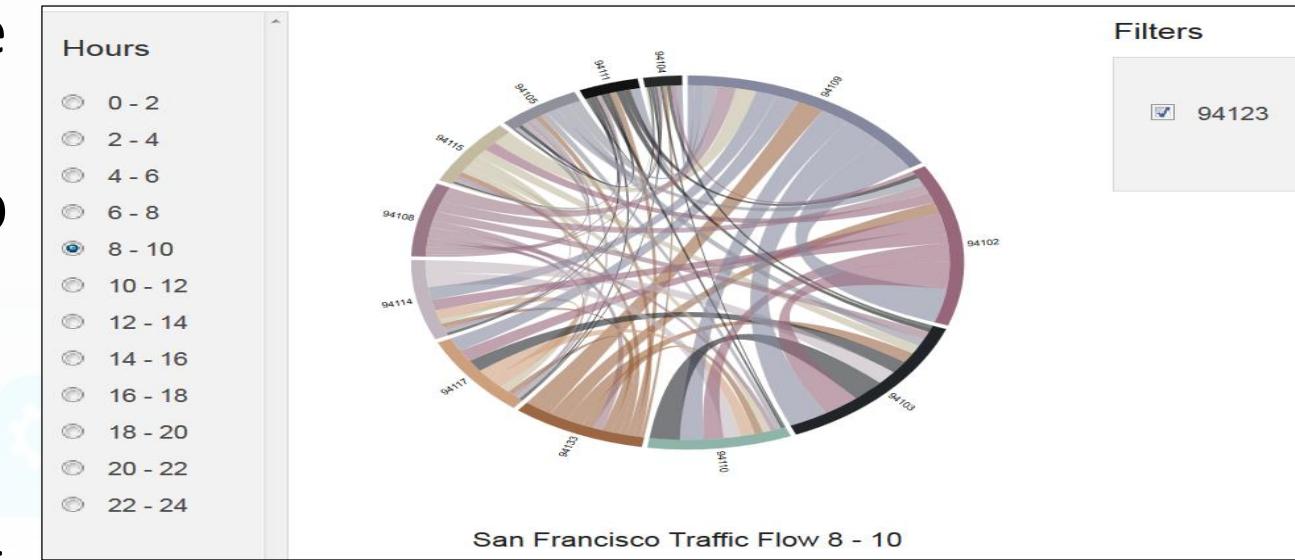


- IOT/IOE
  - Protocols
  - Devices
  - Data aggregation



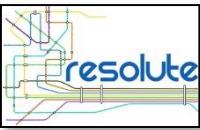
# Traffic and People Flow Assessment

<http://www.disit.org/6694>

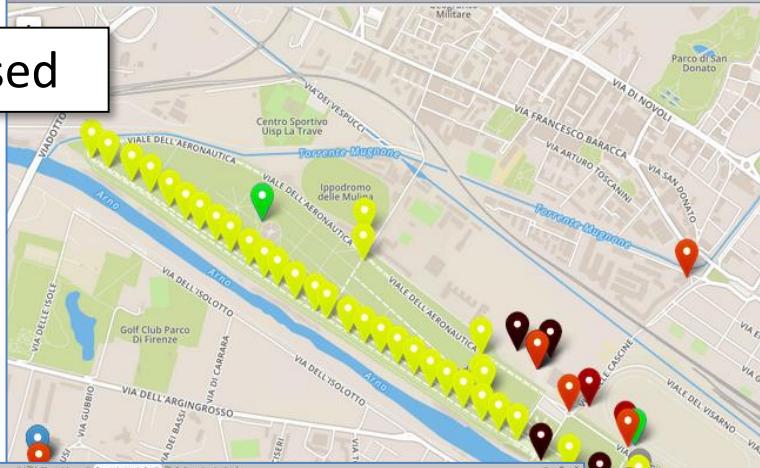


# Characterizing City Areas

Wi-Fi based

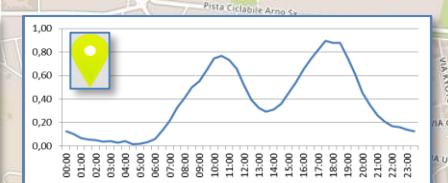
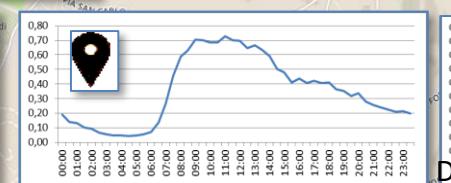
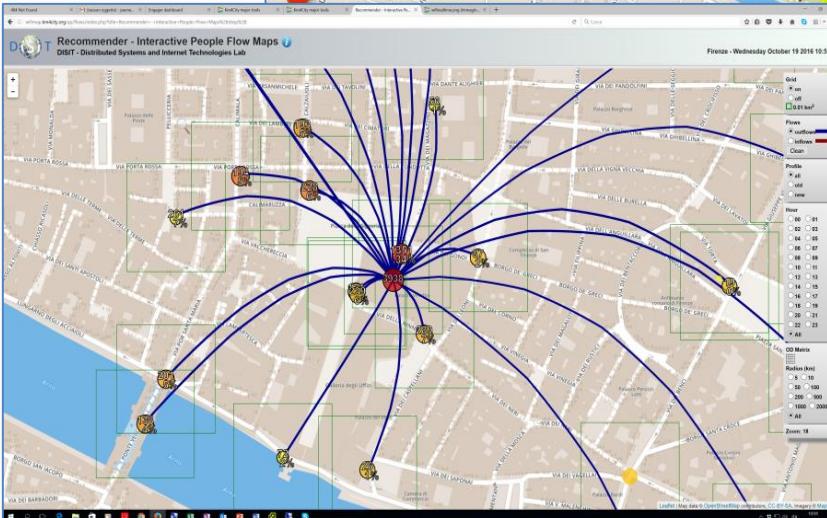
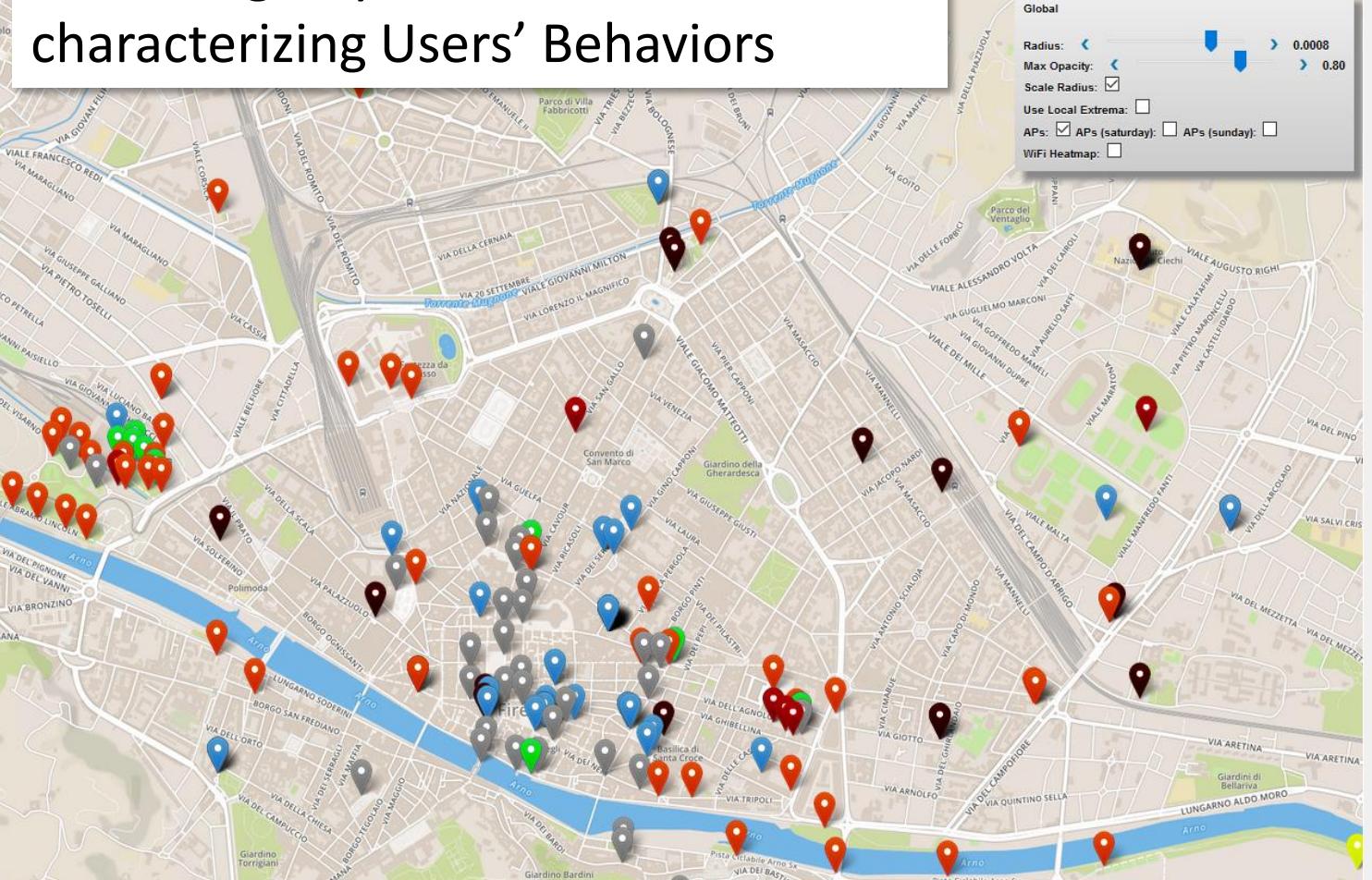


**Firenze Wi-Fi: Access Points Clusters Coverage Map**   
DISIT - Distributed Systems and Internet Technologies Lab

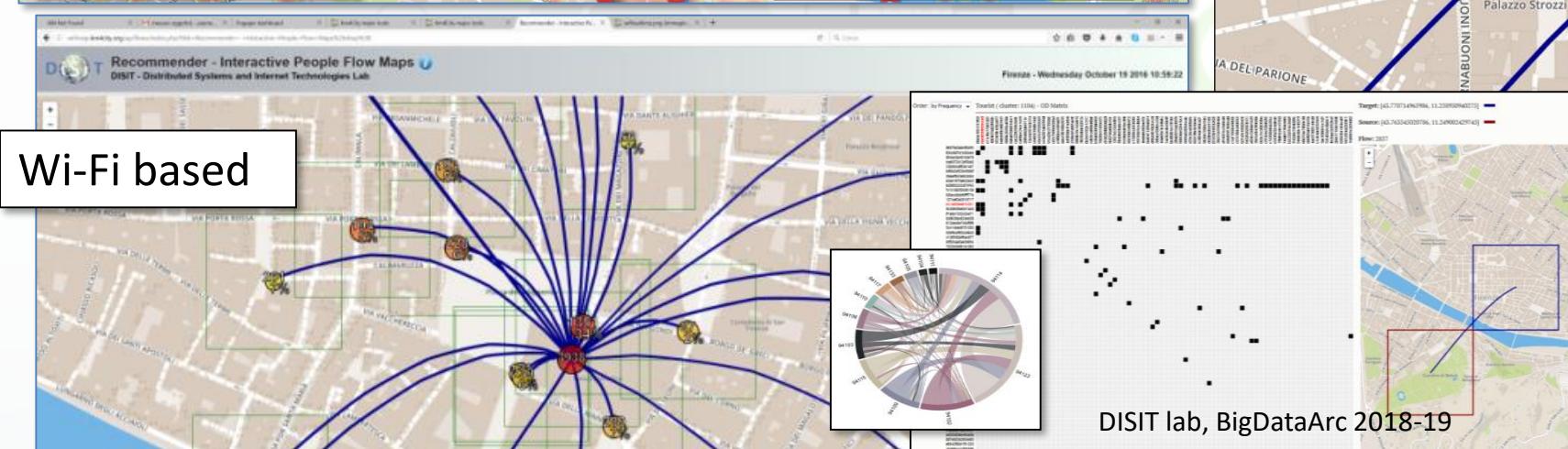
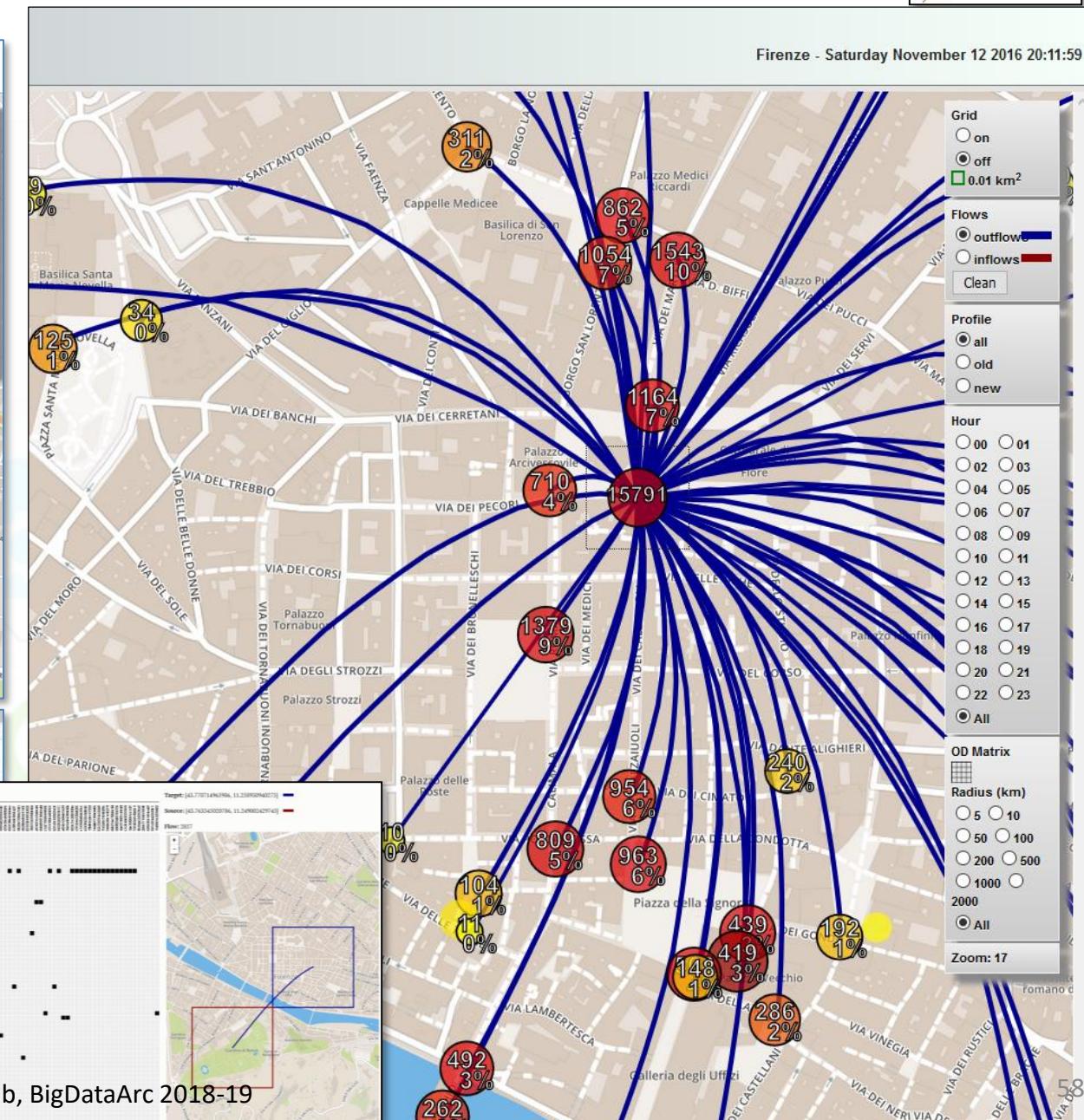
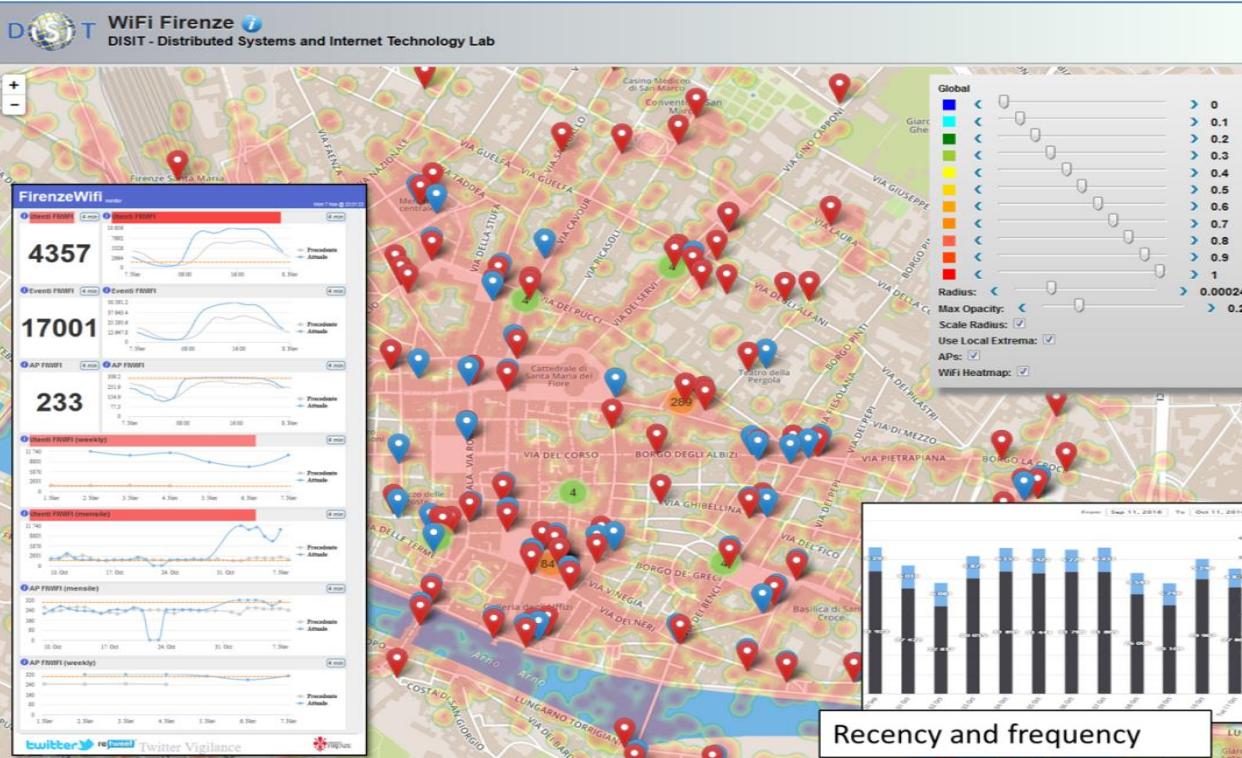


Predicting City Areas Crowd level  
characterizing Users' Behaviors

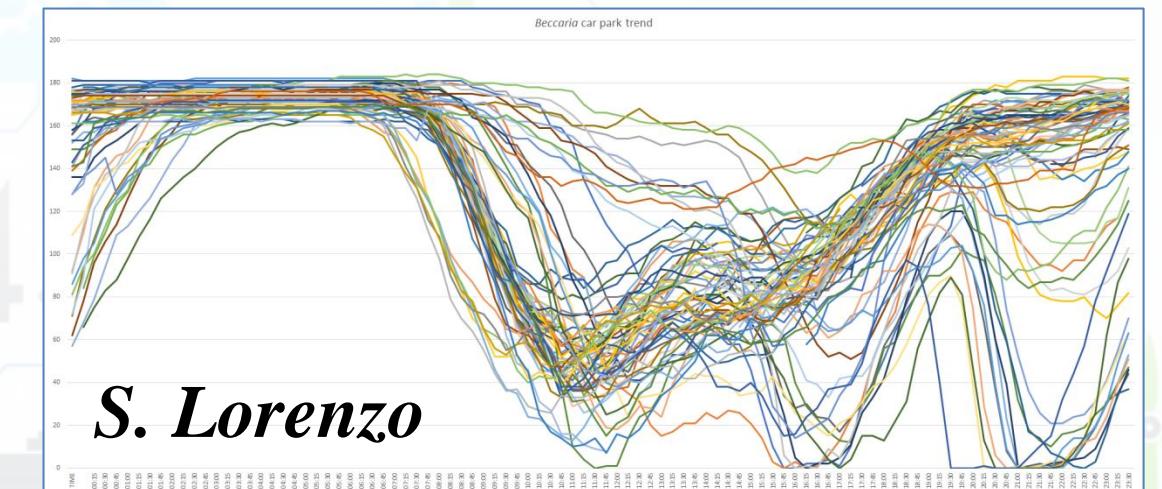
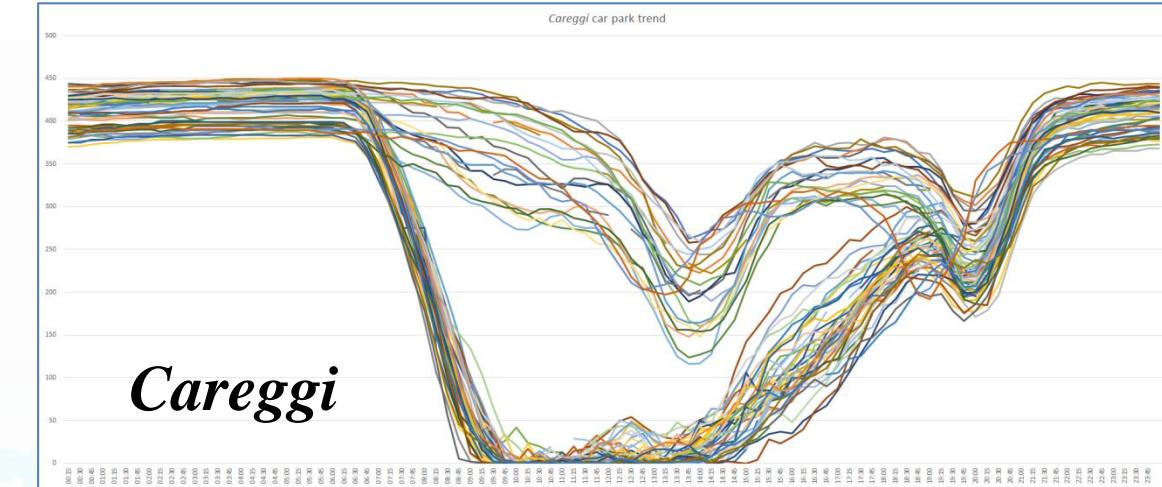
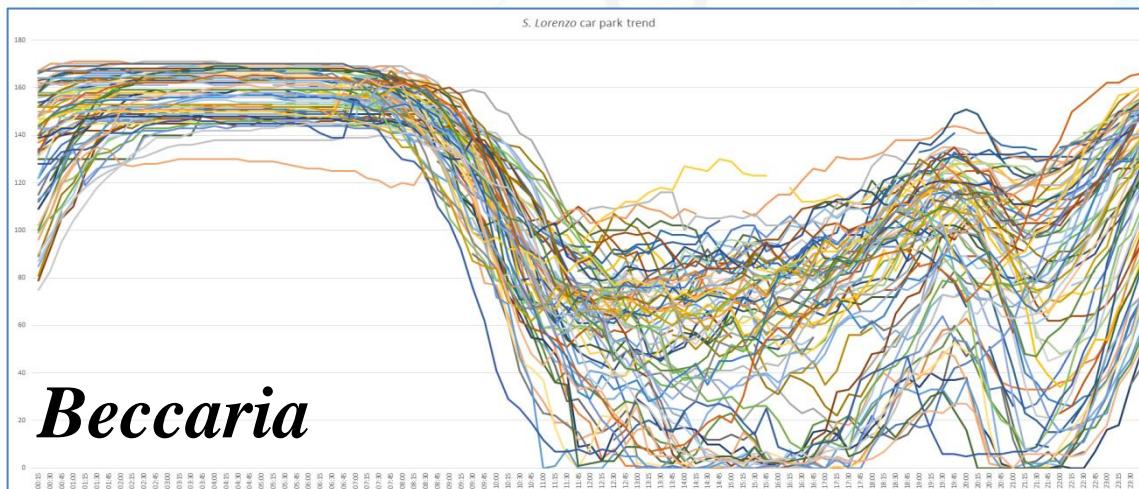
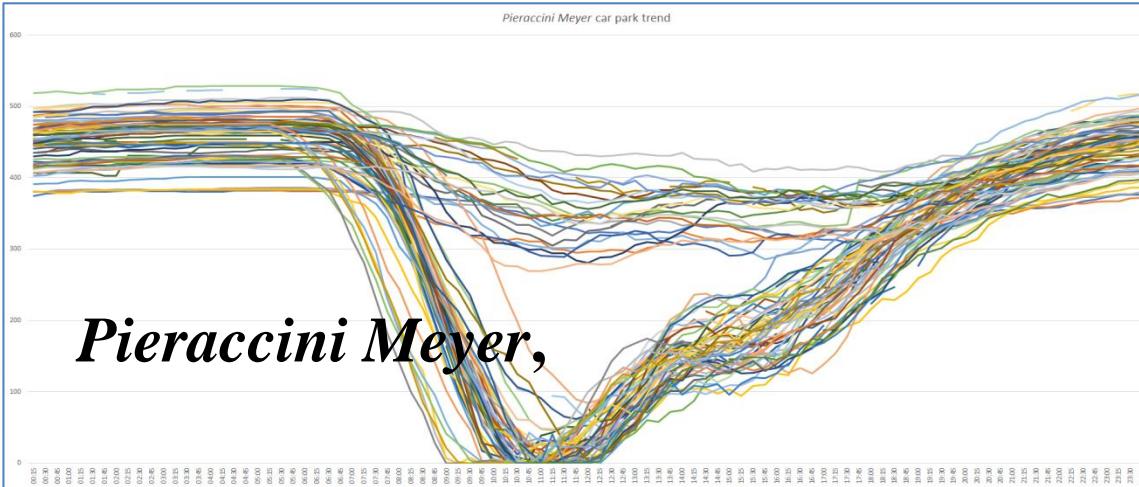
Firenze - Saturday November 12 2016 19:16:33



# destination Matrix Estimation



# Free Parking space trends



12 parking areas in Florence



- Develop European Resilience Management Guidelines (ERMG)
  - Develop a conceptual framework for creating/maintaining Urban Transport Systems
- Enhance resilience through improved support of human decision making processes, particularly by training professionals and civil users on the ERMG and the RESOLUTE system
- Operationalize and validate the ERMG by implementing the RESOLUTE Collaborative Resilience Assessment and Management Support Systems (CRAMSS) for Urban Transport Systems addressing Road and Urban Rail Infrastructures
  - Pilots in Florence and Athens
- Adoption of the ERMG at EU and Associated Countries level



Horizon 2020  
European Union Funding  
for Research & Innovation

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

University of Florence: DISIT lab DINFO (Proj coordinator), DISIA and DST	UNIFI	IT
THALES	THALES	IT
ATTIKOMetro	ATTIKO	GR
Comune di Firenze	CDF	IT
Centre for Research and Technology Hellas	CERTH	GR
Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V.	FHG	DE
HUMANIST	HUMANIST	FR
SWARCO Mizar	SWMIZ	IT
Associação para o Desenvolvimento da Investigação no Instituto Superior de Gestão	ADI-ISG	PT
Consorzio Milano Ricerche	CMR	IT

# Free Parking PREDICTIONS

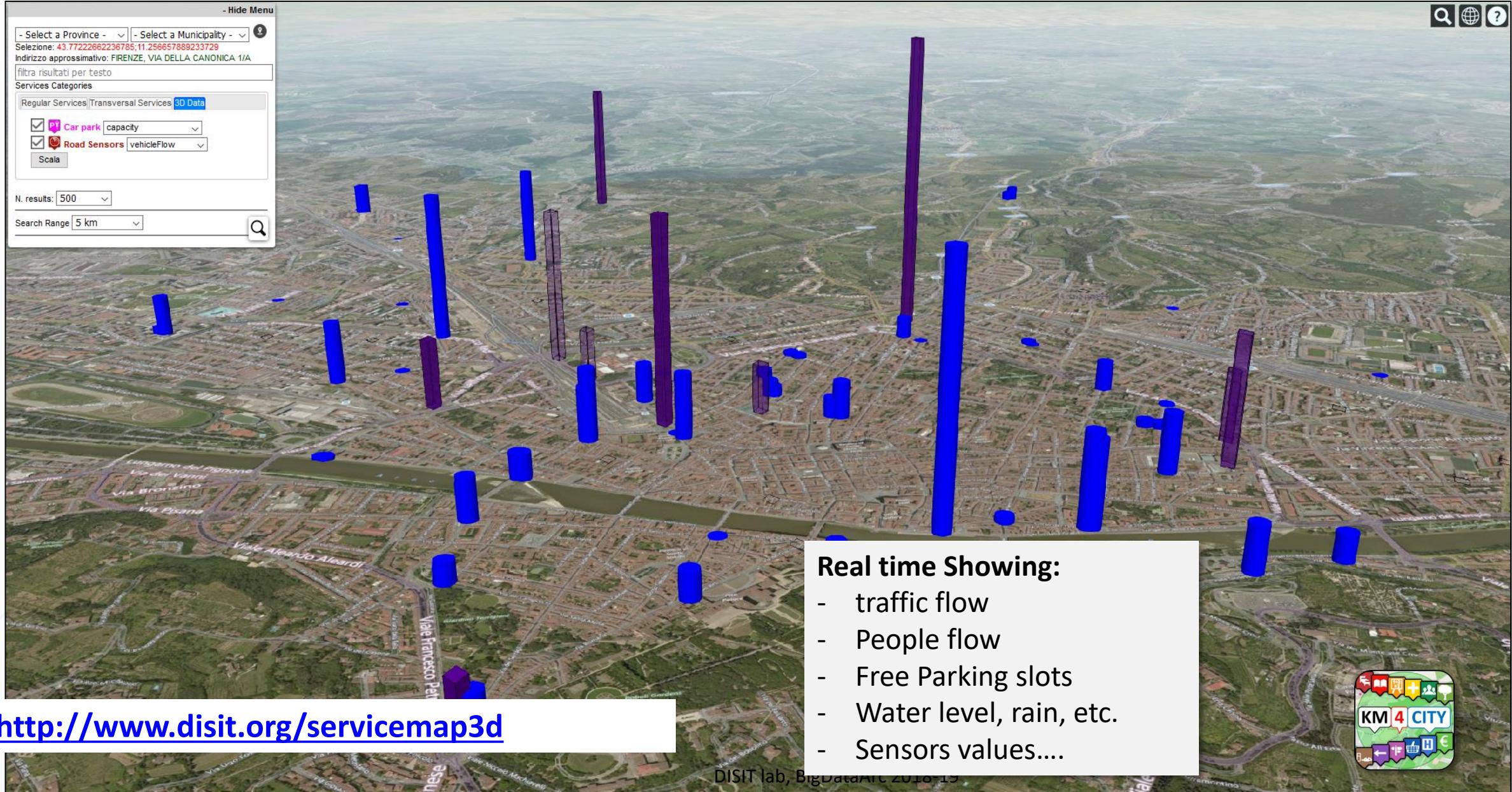
## Careggi car park

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

# RealTime Values 3D



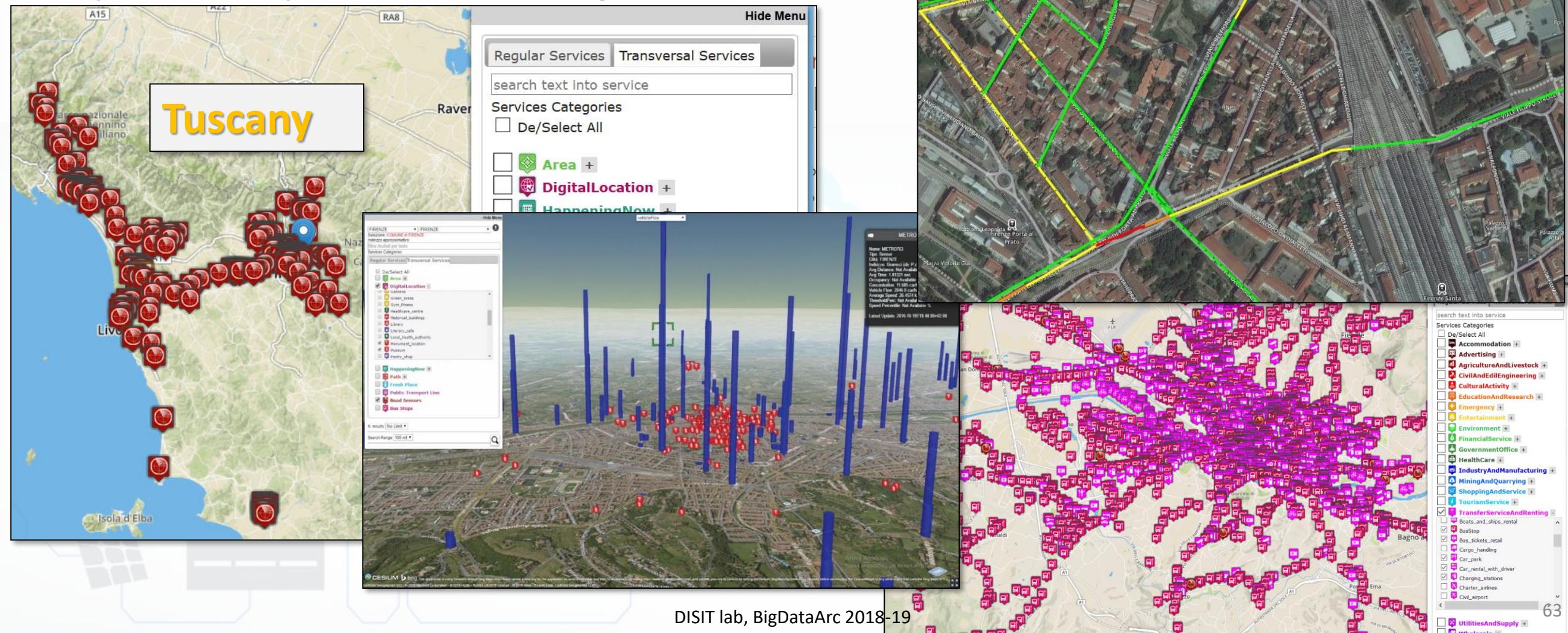
The image shows a 3D aerial map of Florence, Italy, with various data overlays. On the left, a sidebar provides search and filter options: 'Select a Province' (Tuscany), 'Select a Municipality' (Firenze), address (via della Canonica 1/A), results per page (500), and a search range of 5 km. Services categories include Regular Services, Transversal Services, and 3D Data. Under 3D Data, 'Car park capacity' (blue cylinders) and 'Road Sensors vehicleFlow' (purple cylinders) are selected. A legend at the bottom right identifies these symbols. A large white box on the right contains the text 'Real time Showing:' followed by a bulleted list of data types.

<http://www.disit.org/servicemap3d>

**Real time Showing:**

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

- Spire and Virtual Spires (cameras), Bluet
- Specifically located: along, around, ..



2018-02-01T00:10:00

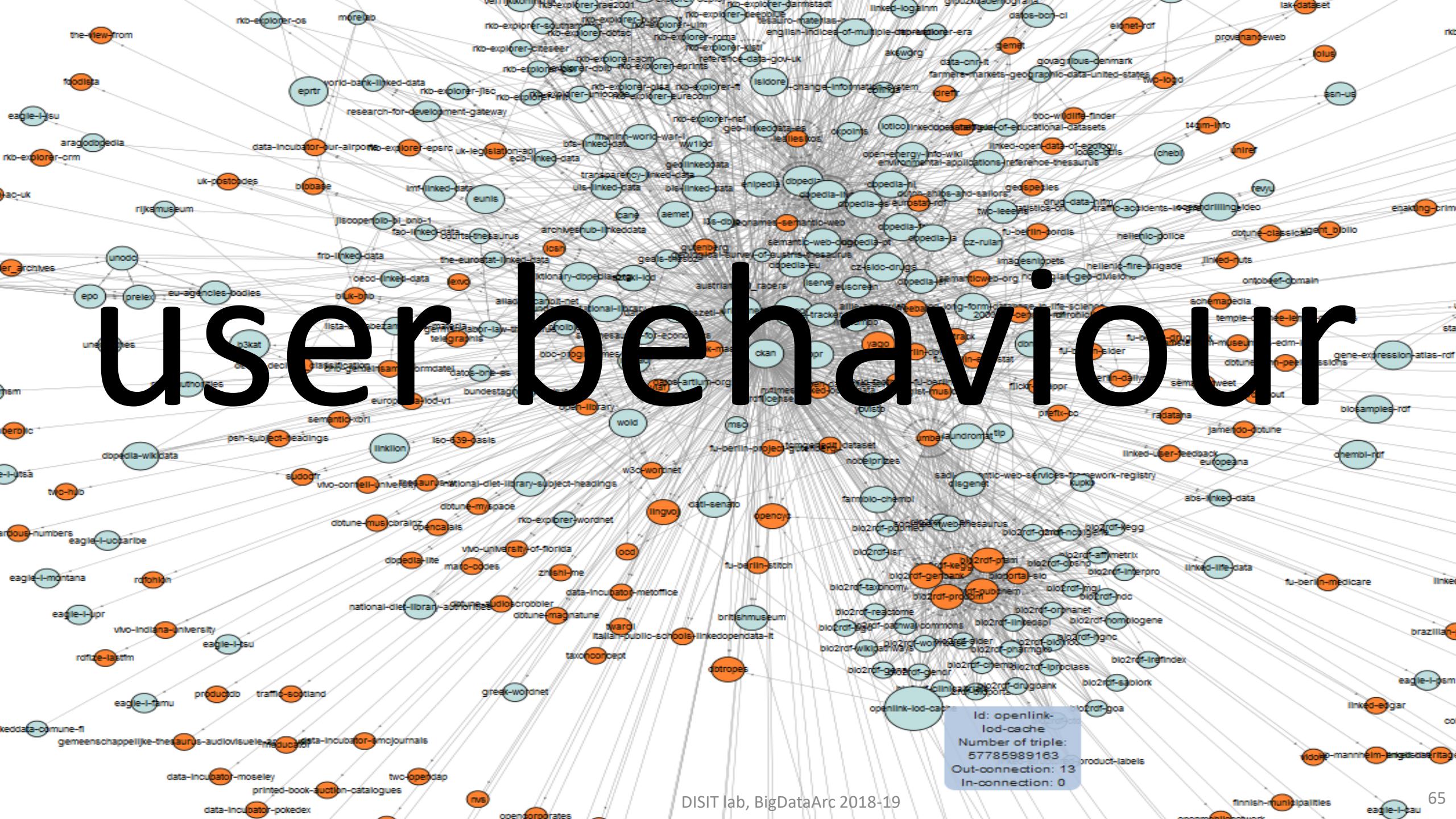
Last sensors measure  
2018-02-01T00:10:00

- Free street
- Fluid traffic
- Heavy traffic
- Very heavy
- Sensor position

<http://firenzetraffic.km4city.org>

# Traffic Flow reconstruction, real time

# user behaviour

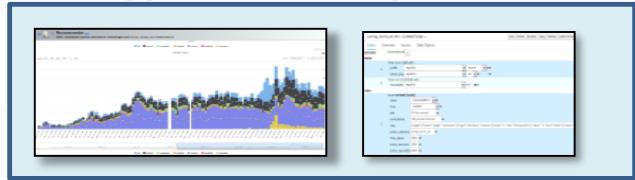


# Feedback Project

# Personal Assistant

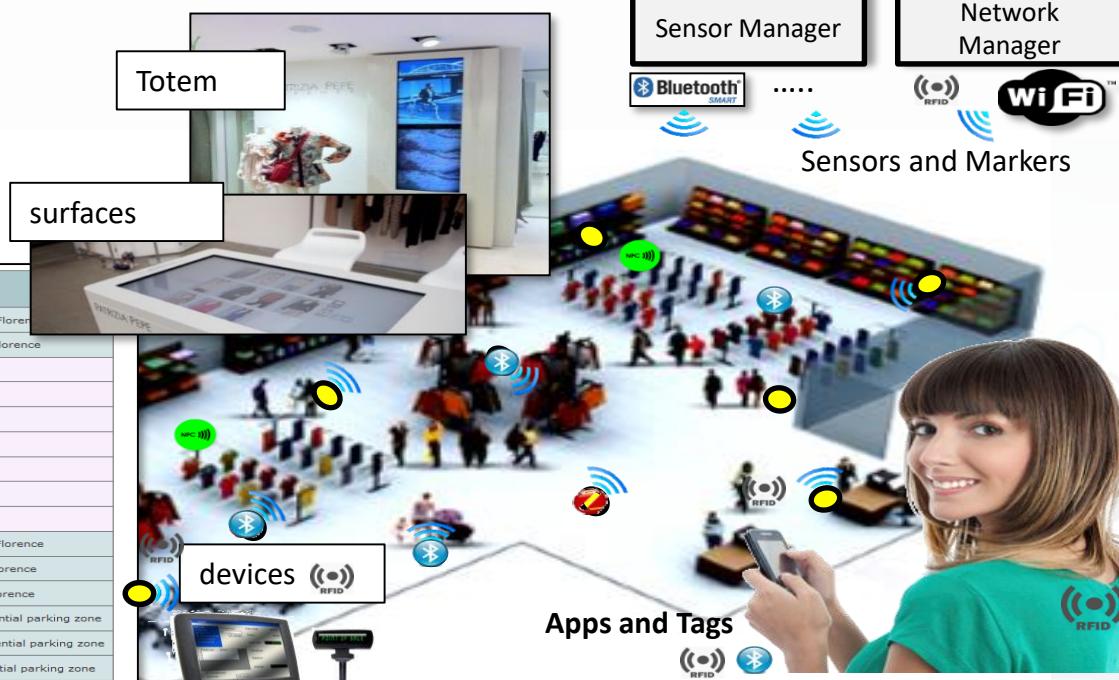
## Operators

### Strategy Editor and feedback

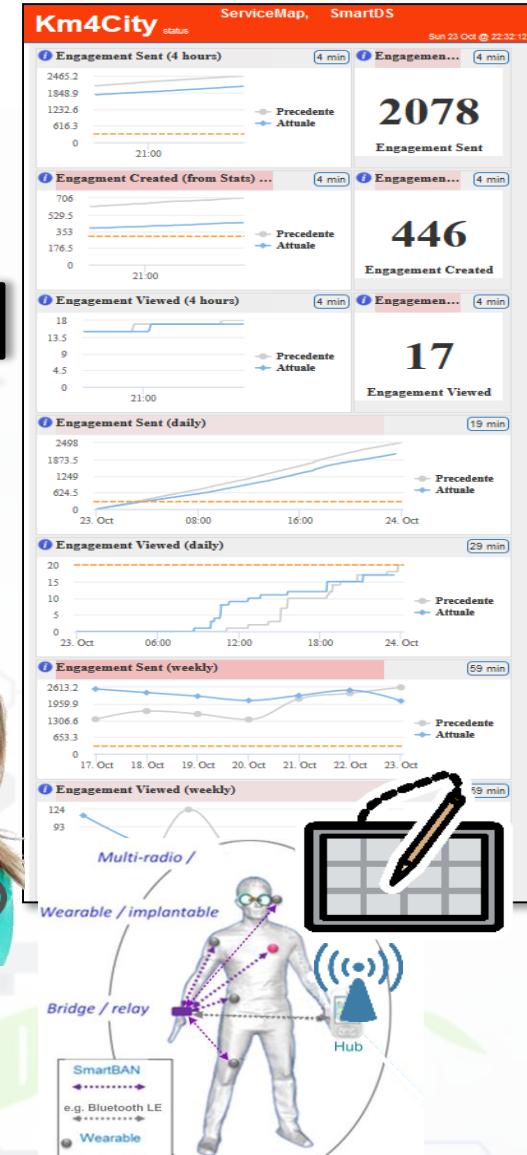


Rule name	Type	#sent	#viewed	#viewed on #sent	Description
daily_event_de	ENGAGEMENT	1 (0%)	0 (0%)	0%	Suggest (in german) an event currently on in Florence
daily_event_en	ENGAGEMENT	1720 (2.12%)	70 (7.1%)	4.07%	Suggest (in english) an event currently on in Florence
- commuter		5 (0.29%)	0 (0%)	0 (0%)	
- 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

Inform  
Engage  
Stimulate / recommend  
Anomalies Detection  
Provide Bonus, incentives



## IOT/IOE





# Toscana dove cosa, ....



- Tutta la Toscana
- Personalizzabile
- Profilata per tipo di utente
- Trasporto pubblico
- Traffico, percorsi, navigazione
- Parcheggi liberi
- Costi benzina
- Suggerimenti
- Assistenza
- Protezione civile
- Meteo
- Biglietti bus
- Punti di Interesse
- Contributi degli utenti

A large blue watermark with the URL <http://www.km4city.org/app> is overlaid diagonally across the center of the collage.

# Km4CityMobile App: all stores



KM4CITY

ponte vecchio

- Torre Dei Mannelli
- Fontanella Ponte Vecchio
- La Strega Noccia - Gelateria Artigianale
- Ponte Vecchio Liner C3 - D
- Ponte Vecchio

web application

<http://www.km4city.org>

KM4CITY

Accommodation +

CulturalActivity +

Education +

Emergency +

Entertainment +

Environment & Ag +

FinancialService +

ATM

Bank

Financial Institute

Insurance

Government Office

Healt Care

Shopping

Tourism Service

Transfer Service

WineAndFood +

Choose

Accommodation +

CulturalActivity +

Education +

Emergency +

Entertainment +

Environment & Ag +

FinancialService +

ATM

Bank

Financial Institute

Insurance

Government Office

Healt Care

Shopping

Tourism Service

Transfer Service

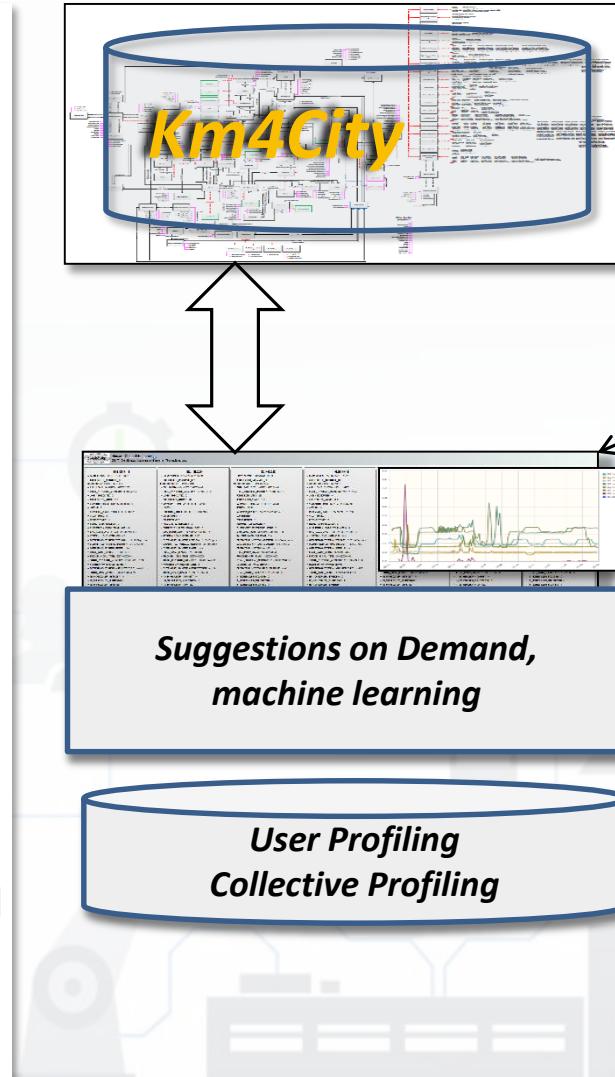
Bus Stops

Go

# Proximity Suggestion Architecture

Categorie Servizi

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



**Km4City SmartCity API**

Proximity  
search

Suggestion request

FIRENZE

Cose da fare  
Tocca per vedere nuovi suggerimenti

Mobilità  
Tocca per vedere nuovi suggerimenti

Suggestions

Weather FIRENZE

Martedì	Mercoledì	Giovedì	Venerdì	Sabato
20°C 13°C	17°C 14°C	17°C 14°C		

Events

La Casa Delle Eccellenze  
Type: Other events  
Description: Il meglio della produzione fiorentina in vendita nel cuore di Firenze. Un temporary mall all'interno del bel complesso barocco dell'ex Tribunale di San Firenze, nella Sala della Musica. Partecipano numerose aziende di vari settori (moda, artigianato, enogastronomia, casa, meccanica e innovazione) con una rotazione.

Show All Categories

Suggerimenti

Posti Vicini

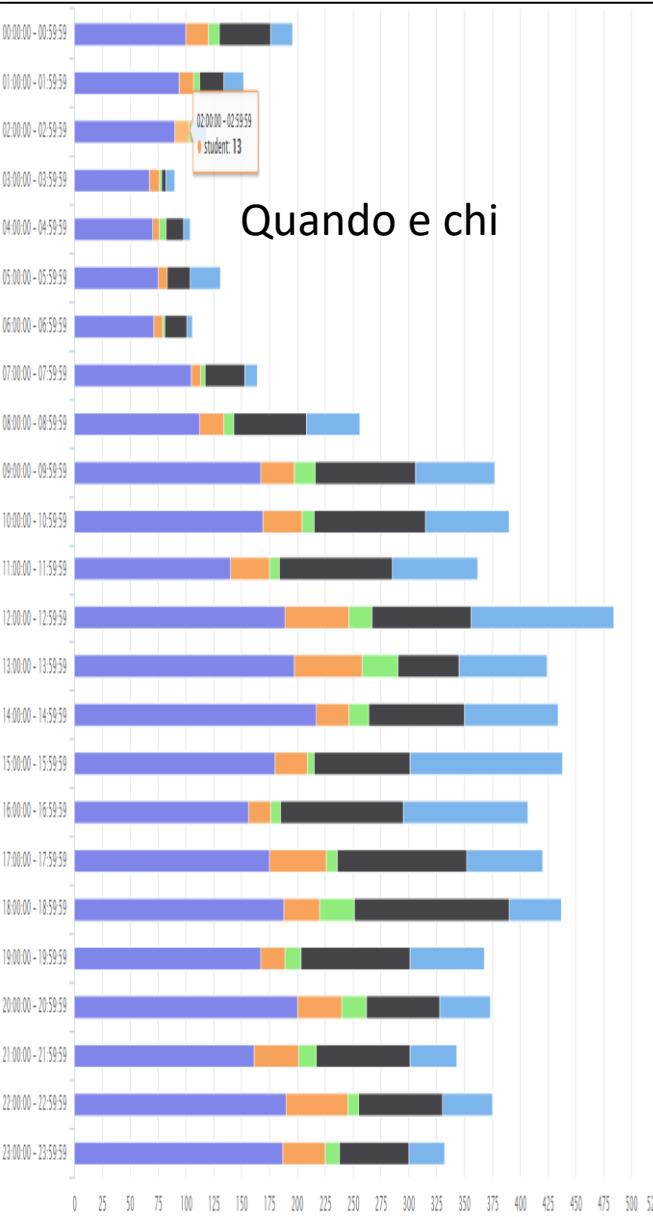
Chiostro Dello Scalzo  
Tip: Monument location  
Distanza: 0.1678 m  
Indirizzo: VIA CAMILLO BENSO CAURO, 54/58 50121 FIRENZE FI

Orto Botanico ("Giardino Dei Semplici")  
Tip: Museum  
Distanza: 0.1700 m  
Indirizzo: VIA PIER ANTONIO MICHELI, 3 50121 FIRENZE FI

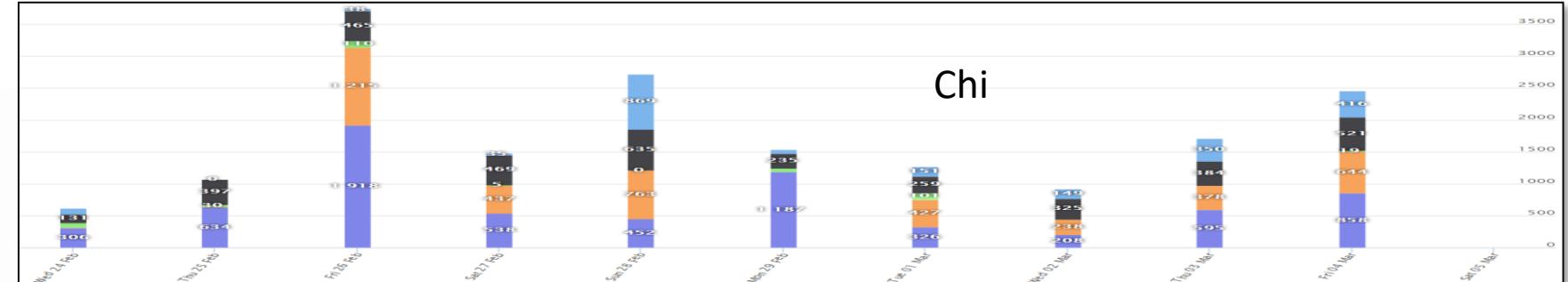
Museo Di  
Mostra Tutte le Categorie

# Recommender

Quando e chi



Chi

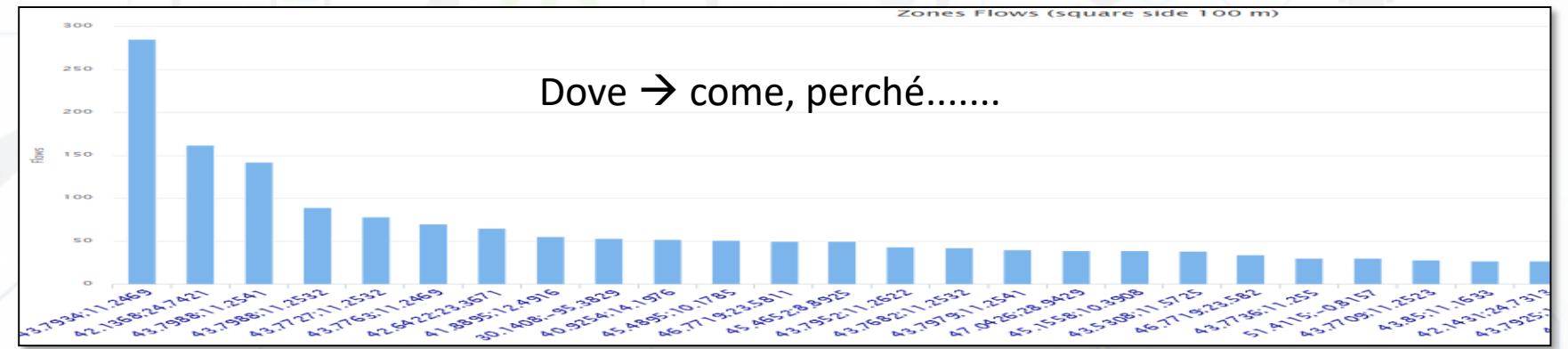


Requested Groups

Cosa



Dove → come, perché.....



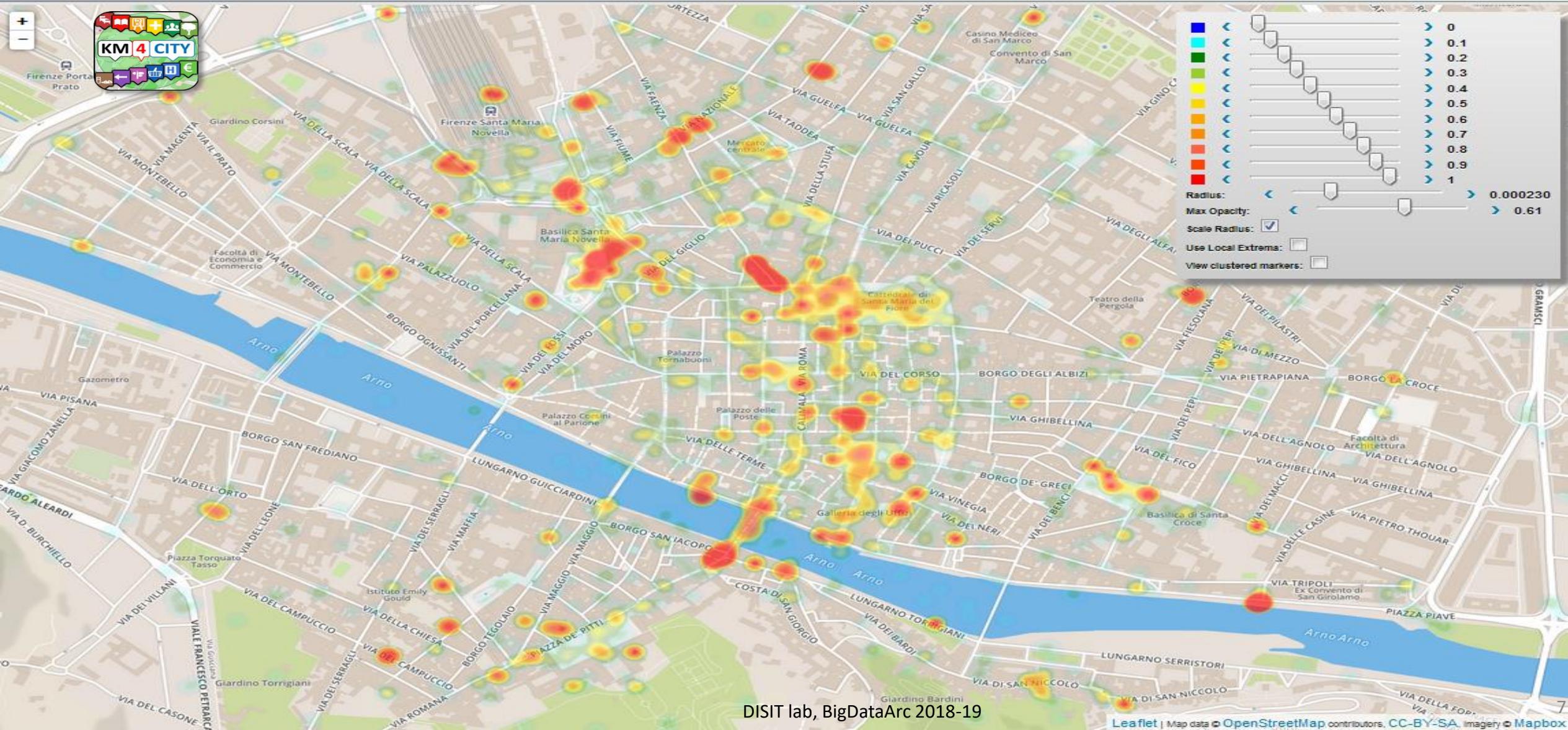


# Tourists in Florence



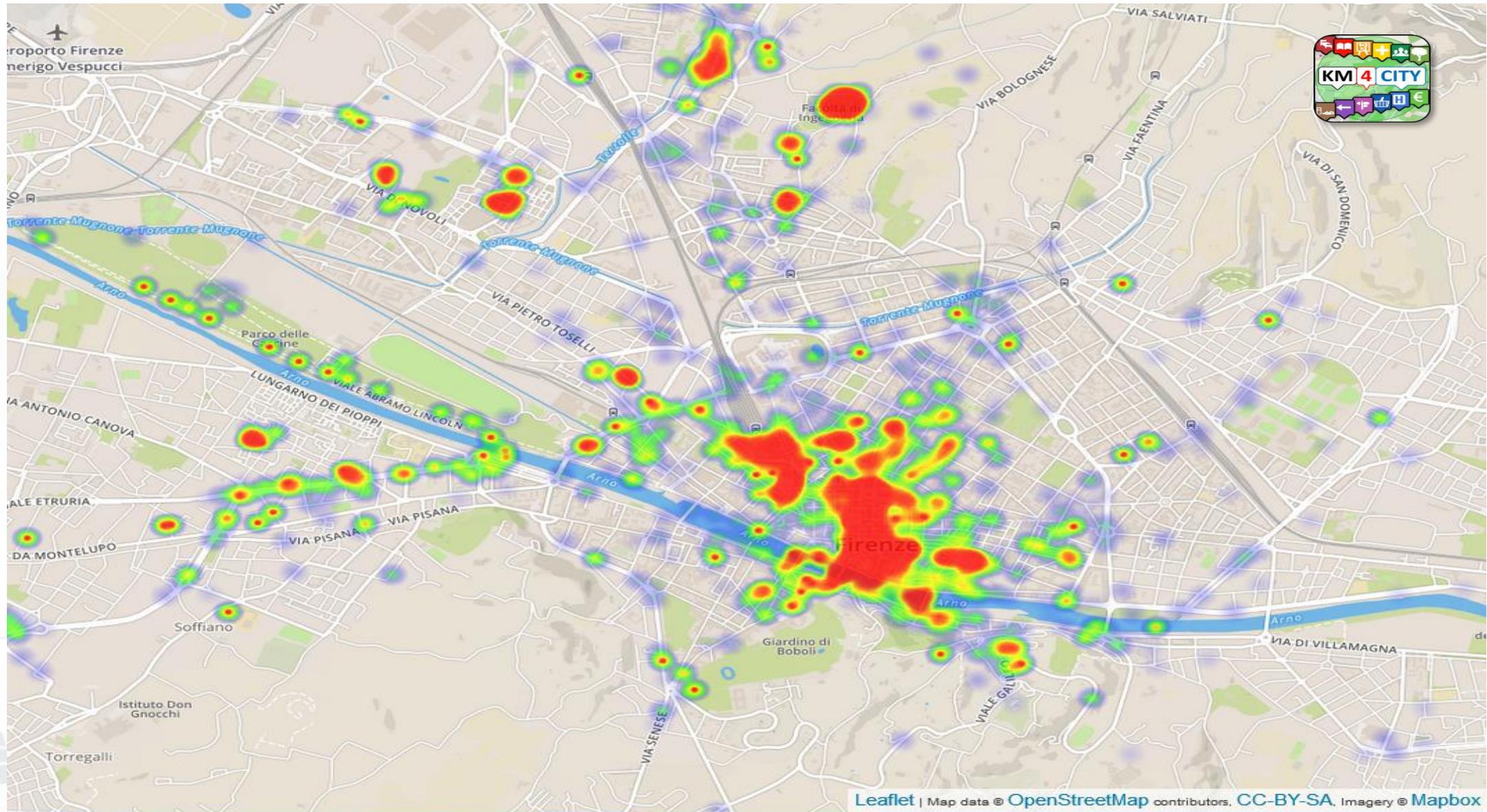
Personal Recommender

DISIT - Distributed Systems and Internet Technology Lab

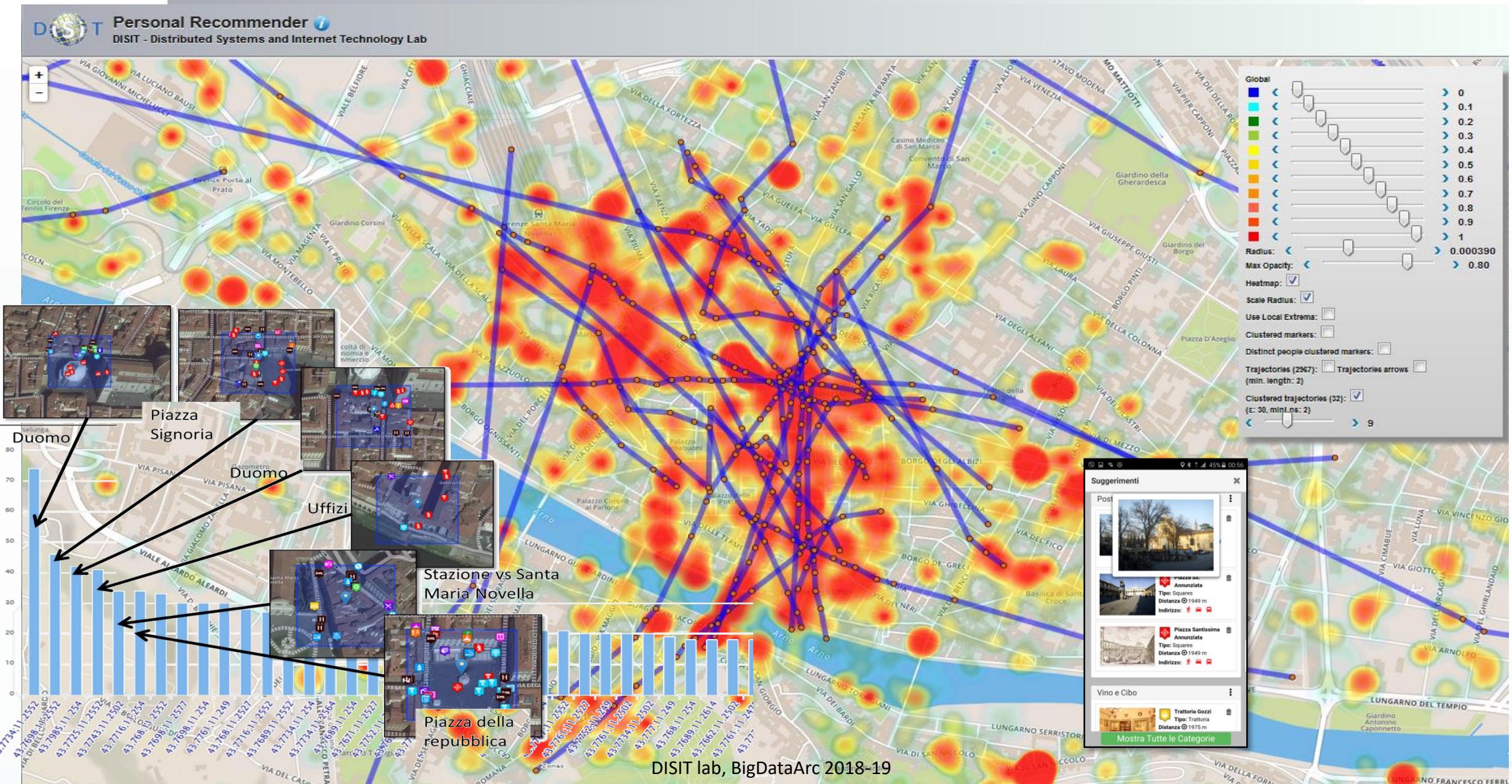




# Hot WiFi in Florence



# User Behavior Analyzer

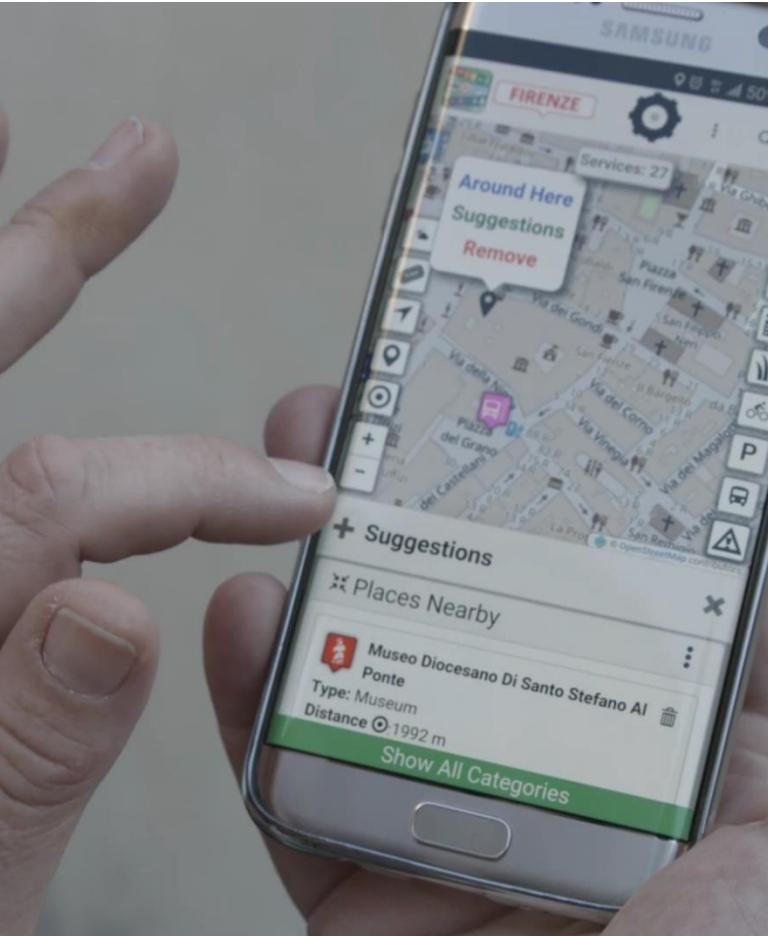




UNIVERSITÀ  
DEGLI STUDI  
FIRENZE

**DINFO**  
DIPARTIMENTO DI  
INGEGNERIA  
DELL'INFORMAZIONE

**DISIT**  
DISTRIBUTED SYSTEMS  
AND INTERNET  
TECHNOLOGIES LAB  
<http://www.disit.org>



**Sii smart.  
Sii-Mobility!**  
Scarica, viaggia, vinci!



Dal 15 aprile al 15 luglio scegliere il trasporto pubblico ti premia!  
Scarica l'app "Toscana dove, cosa", guadagna punti viaggiando in autobus e vinci tanti fantastici premi.  
Per maggiori informazioni visita il sito [info.sii-mobility.org](http://info.sii-mobility.org)



UNIVERSITÀ  
DEGLI STUDI  
FIRENZE

**DINFO**  
DIPARTIMENTO DI  
INGEGNERIA  
DELL'INFORMAZIONE

**DISIT**  
DISTRIBUTED SYSTEMS  
AND INTERNET  
TECHNOLOGIES LAB

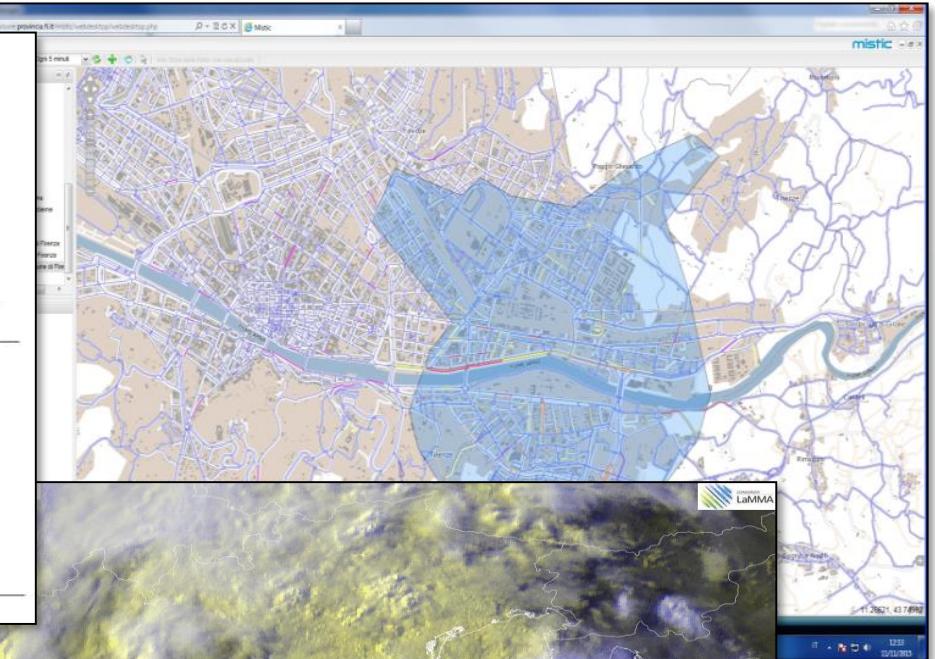
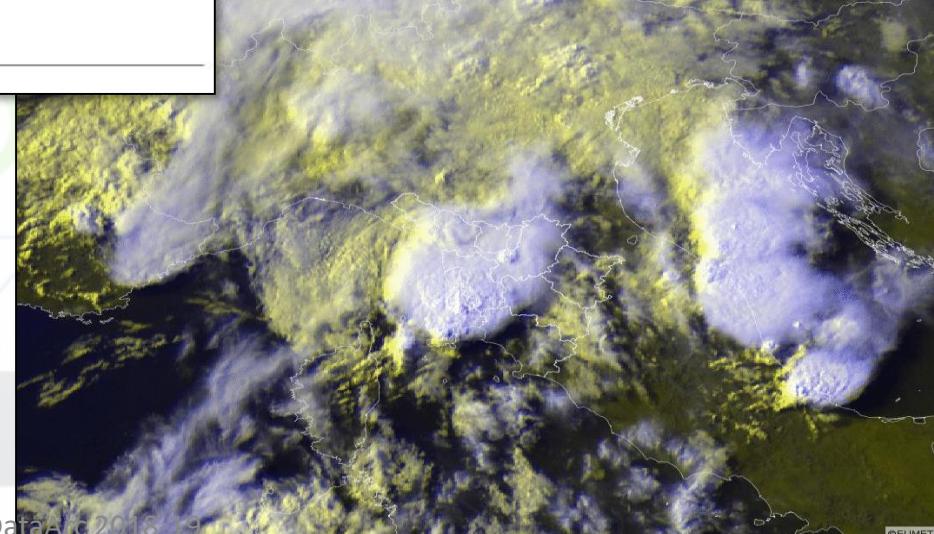
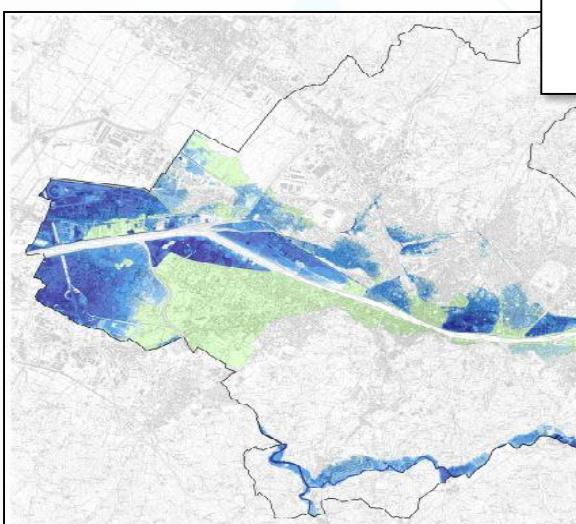
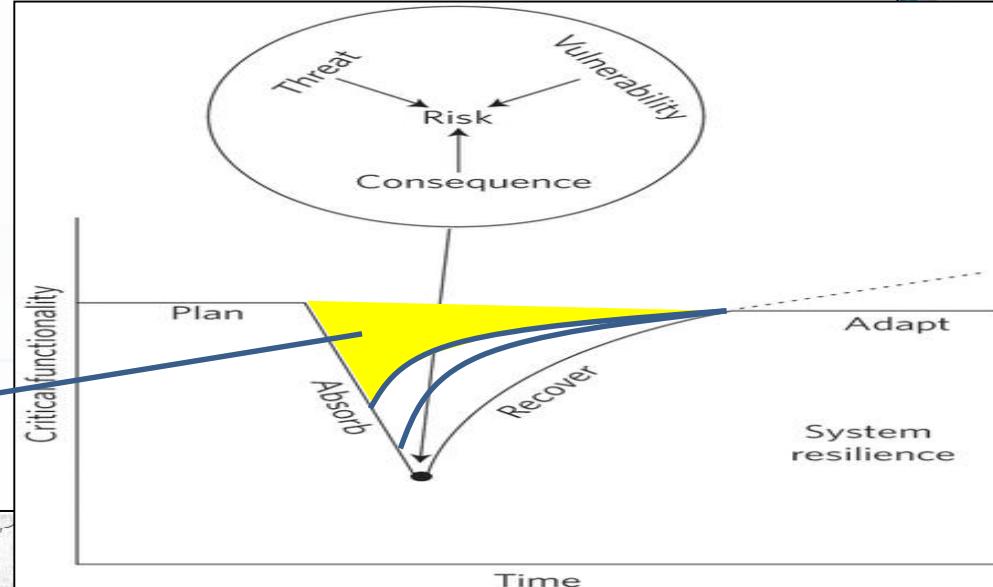
# early warning anomaly detection

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

# Early warning, detection

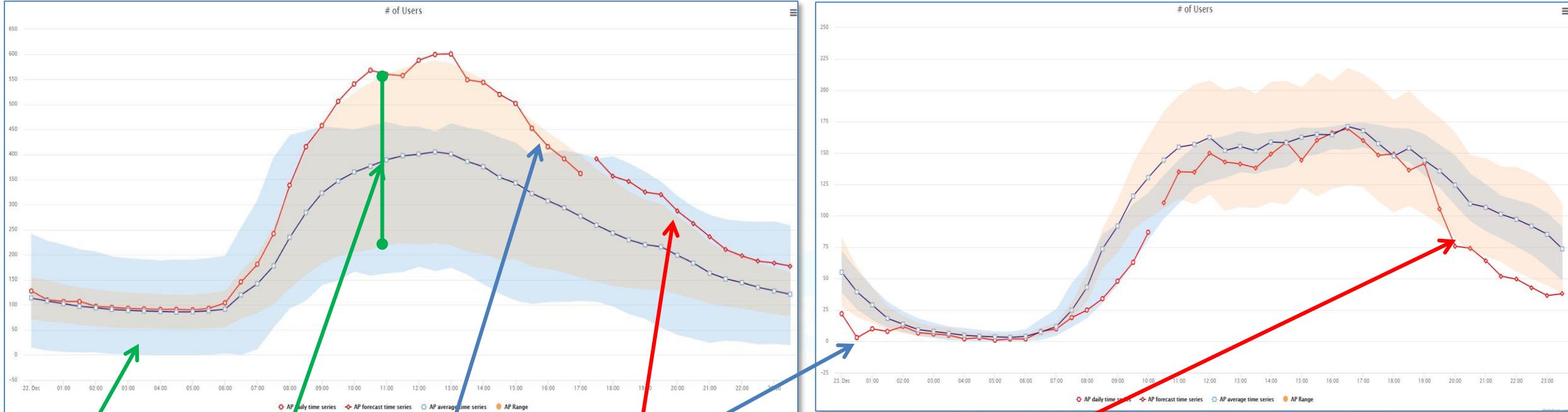
**P**repare  
**A**sorb  
**R**ecover  
**A**dapt

damage



# Prediction and identification of anomalies

## Guessing number of users of Wi-Fi Access Points



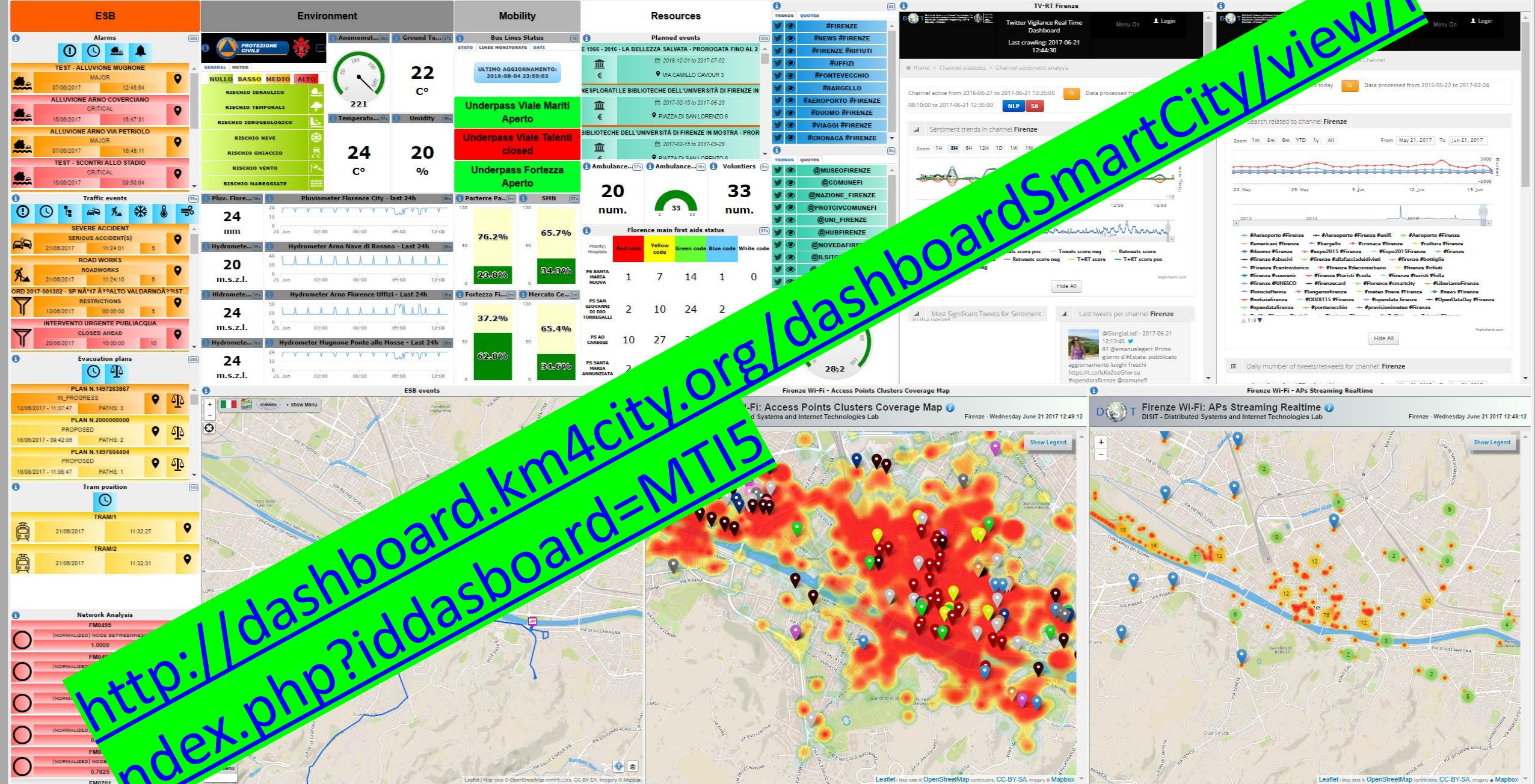
Cluster confidence

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



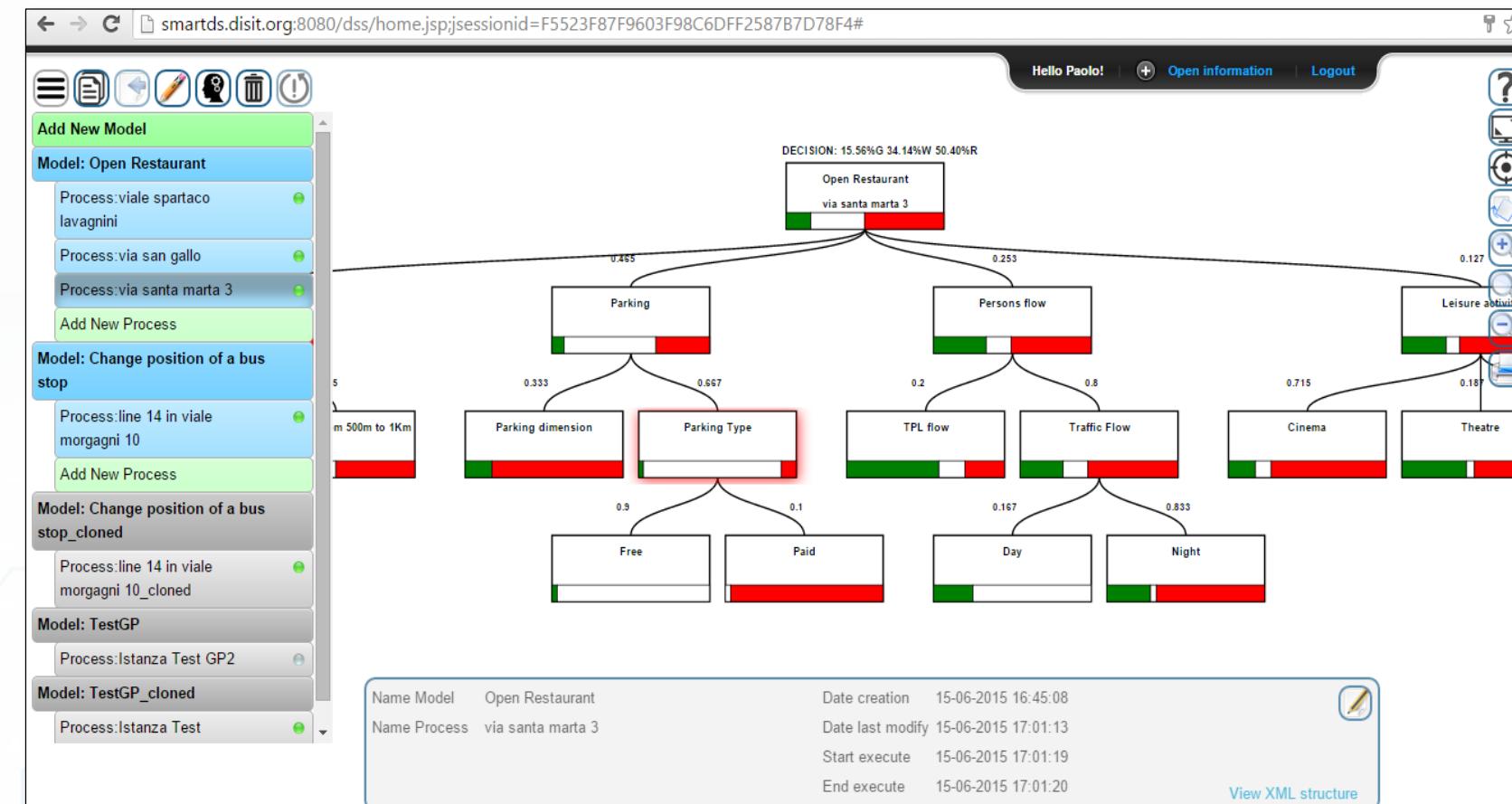


# Smart City Decision Support

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

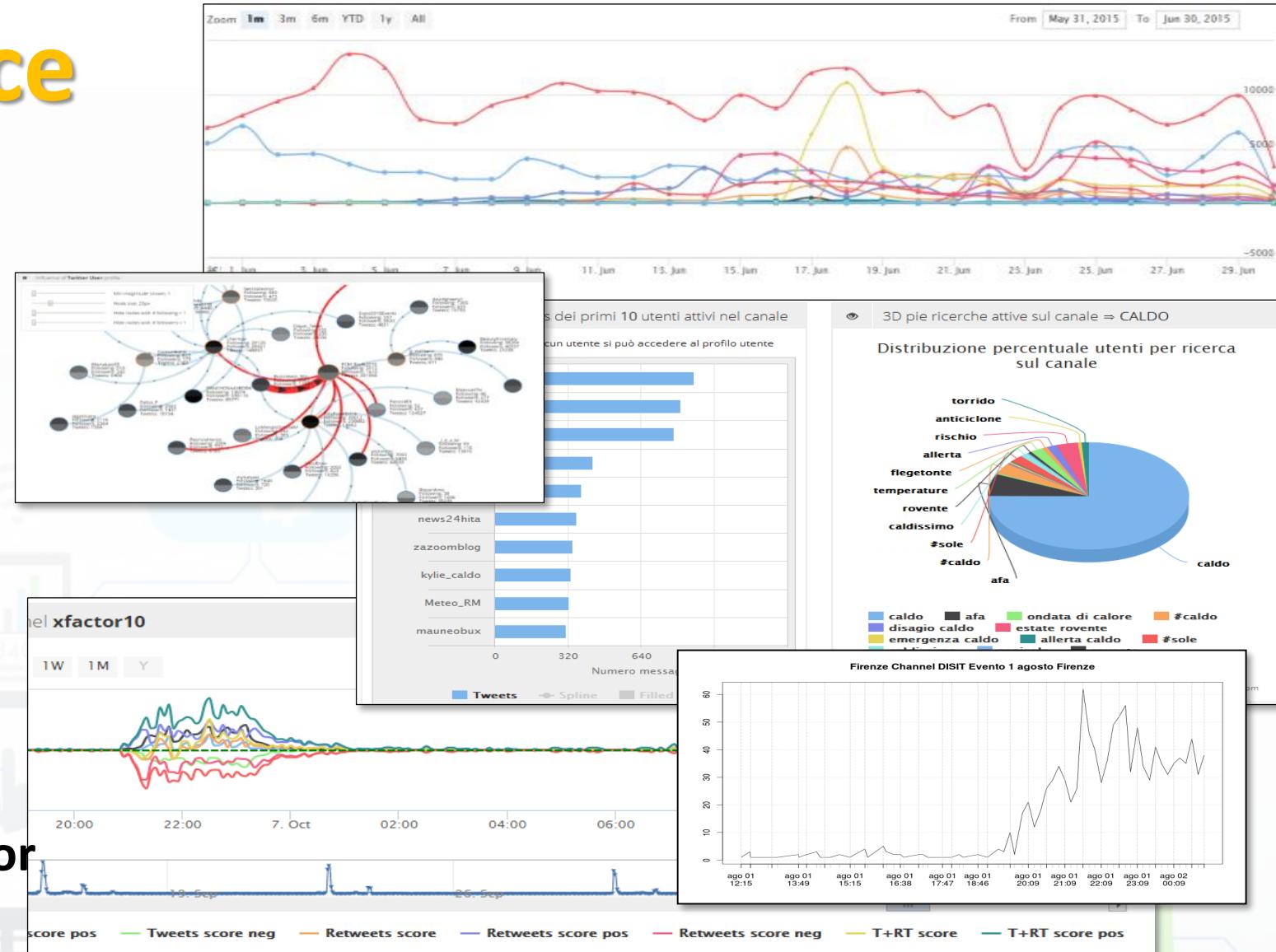
## Enforcing

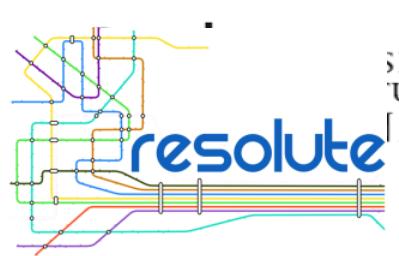
- Mathematical model for propagation of decision confidence..
- Collaborative work...,
- Processes connected to city data: DB, RDF Store, Twitter, etc.
- Production of alerts/alarms
- Data analytics process
- Twitter Processes
- reuse, copy past, ...



# Twitter Vigilance

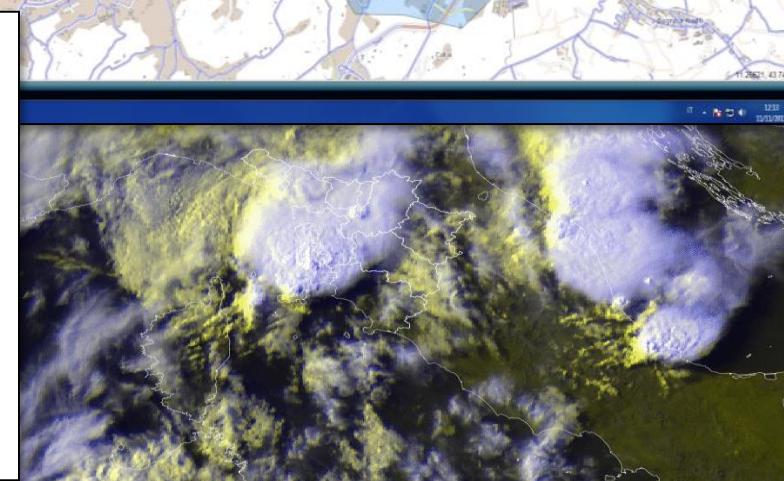
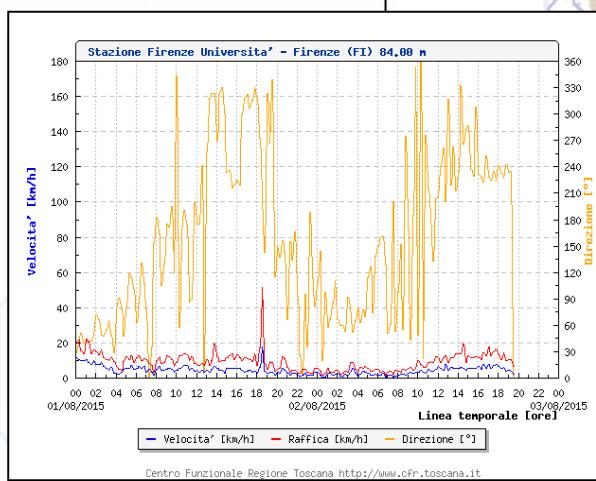
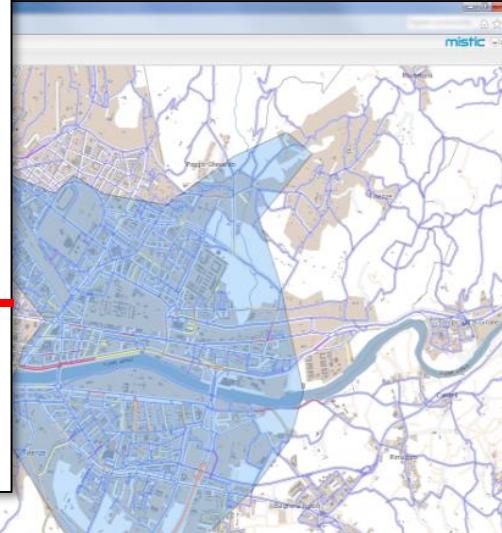
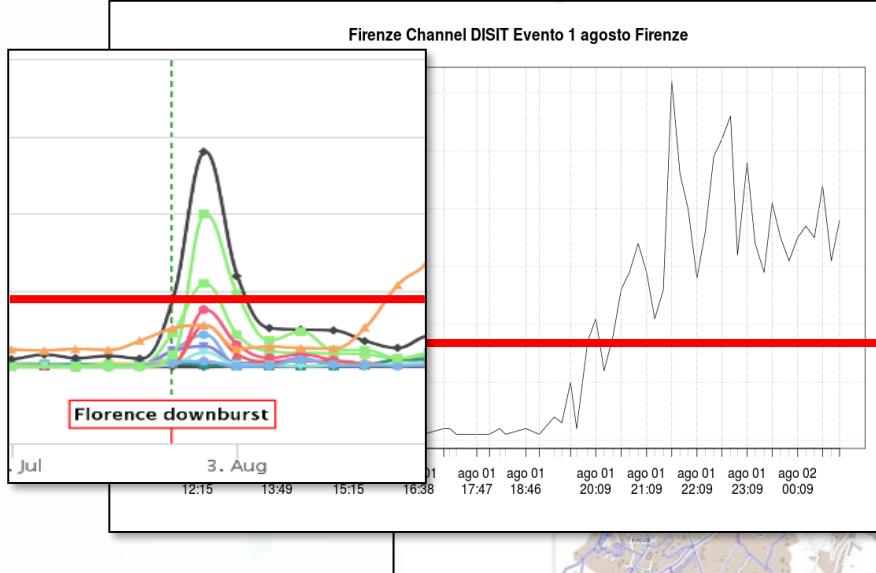
- <http://www.disit.org/tv>
- Citizens as sensors to
  - Assess sentiment on services, events, ...
  - Response of consumers wrt...
  - **Early detection** of critical conditions
  - Information channel
  - Opinion leaders
  - Communities
  - Formation
  - **Predicting volume of visitors for tuning the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



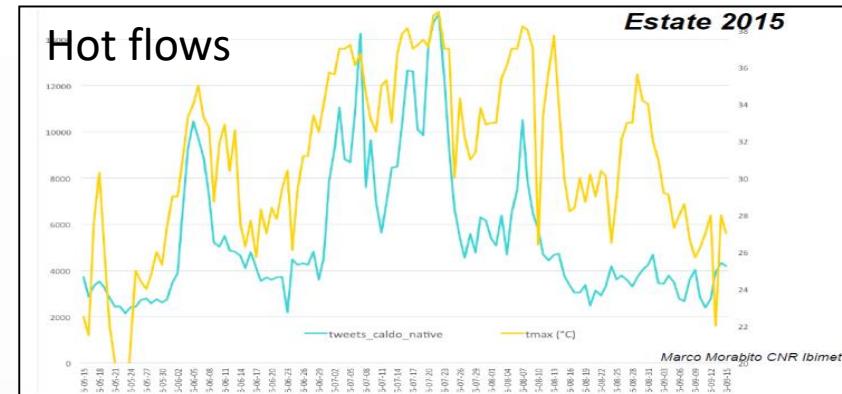


# Twitter Vigilance

## Early Warning



## Predictive models



## Attendance at long lasting events: EXPO2015

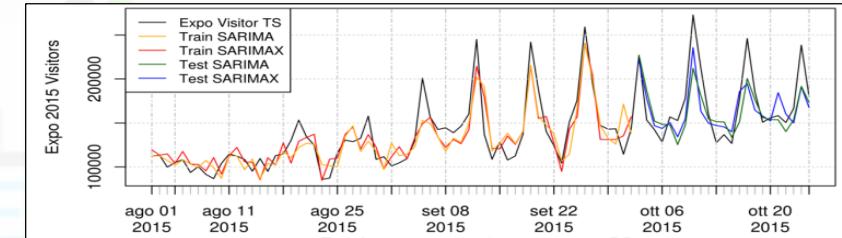
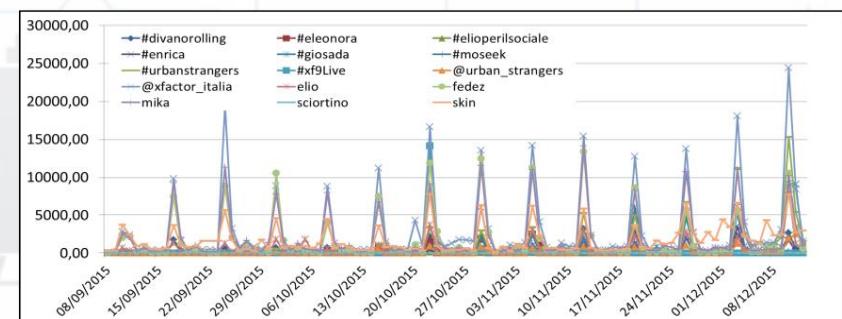
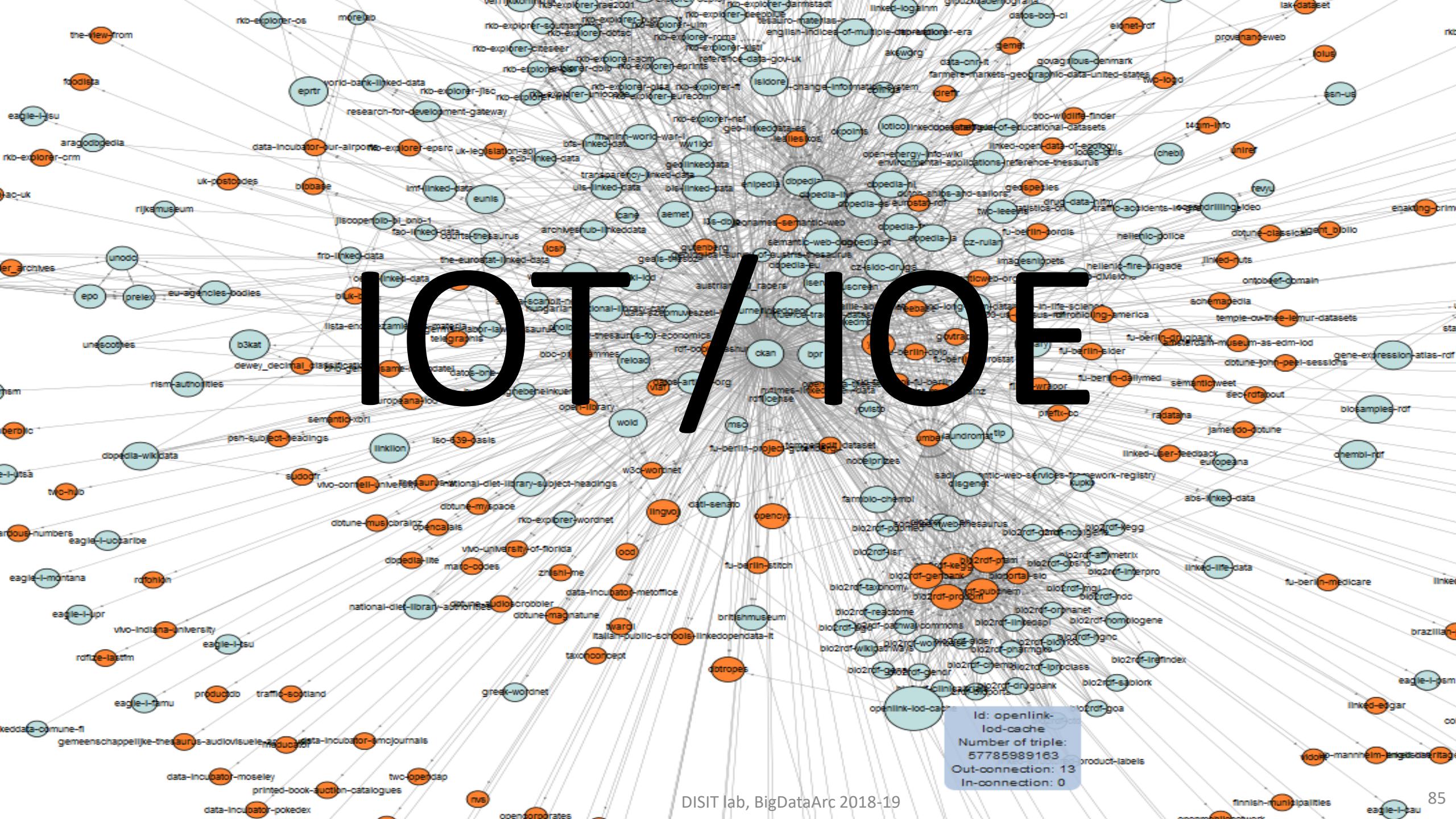


Figure 5: Comparison among the selected predictive models discussed and presented in Tables 2 and 3 with respect to the real number of visitors. Both training and validation periods are reported.

## Attendance at recurrent events: TV, football

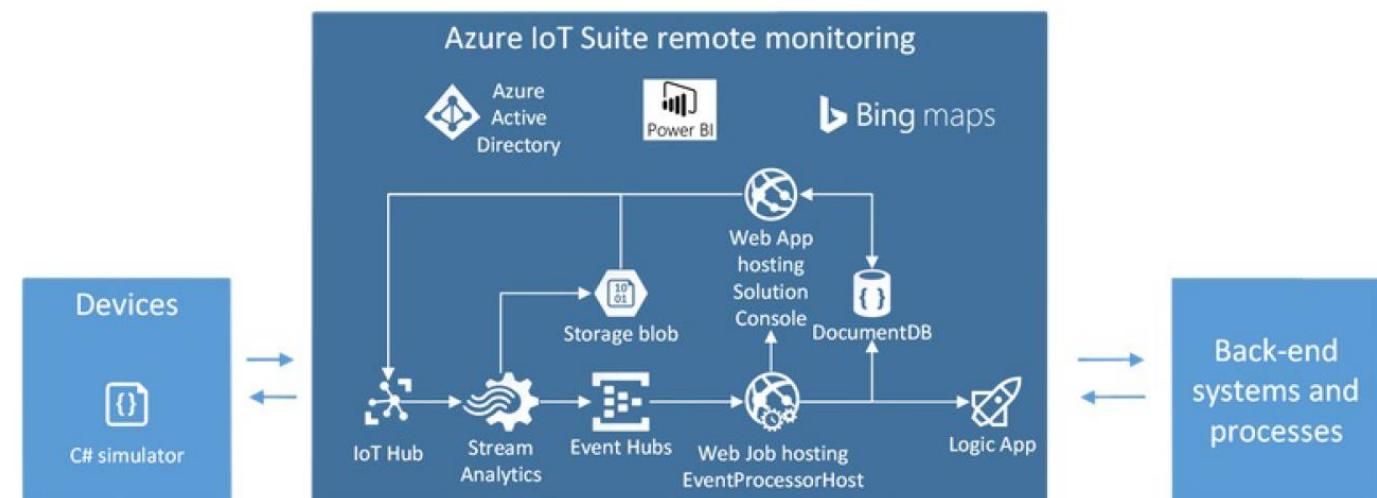
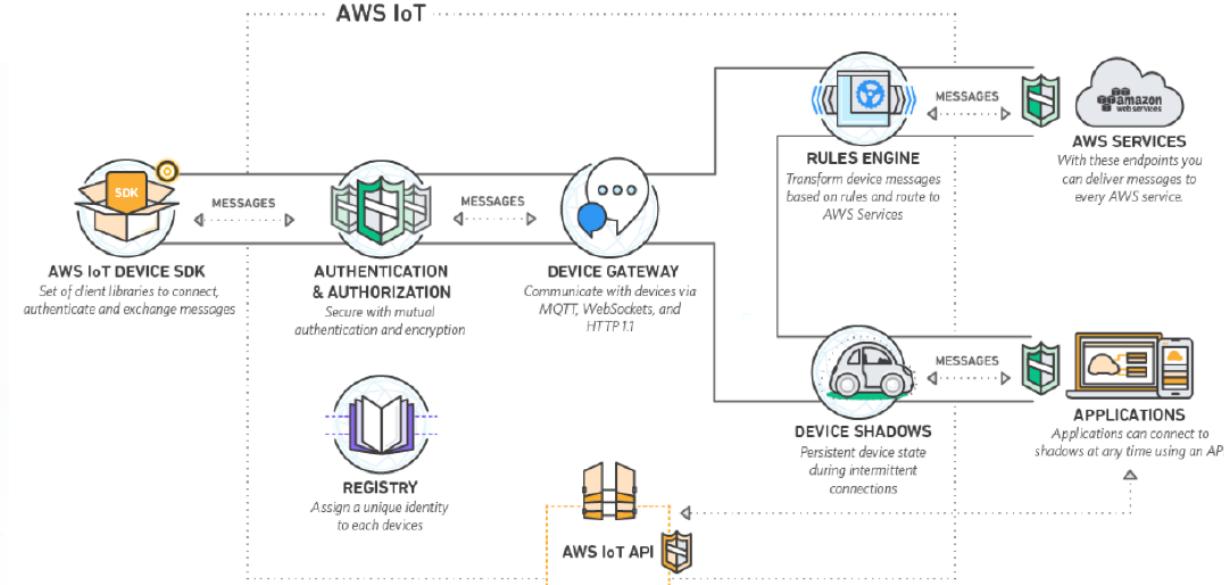


# IOT / IOE



# Big Data vs IOT/IOE

- Store everything
- Shadow (Data Indexing and storage)
- Access to external services ??
- Access to external context of the territory ??
- IOT Applications??
- Data Traffic Monitoring ??
- Privacy vs GDPR ??





- In una certa misura minor problemi di volume nei casi specifici
- Maggiori problemi di real time, event driven
- Maggiori problemi di non uniformità
  - degli stream data in ingresso
  - Formati e protocolli diversi per devices e pacchetti di comunicazione dati
- Problemi di tracciabilità del dato
- Problemi di licensing del dato
- Problemi diffusi di security nelle soluzioni ICT: comunicazione, storage, accesso e monitoraggio, etc.



# IOT/IOE Protocols

## Communication Patterns



### Discovery

Discover, register and "thrust" new devices on the network



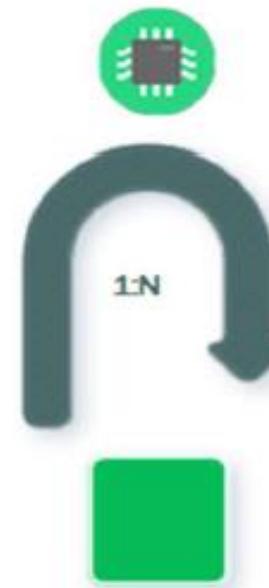
### Telemetry

Information Flows From device to another system for conveying status changes in the device



### Inquiries

Requests from devices looking to gather required information or asking to initiate activities



### Commands

Commands from other systems to a device or a group of devices to perform specific activities



### Notifications

Information flows from other systems to a device or a group for conveying status changes in the world

- MQTT
- HTTP(s)
- AMQP
- COAP
- NGSI
- OneM2M
- WebSocket
- Etc.
- Etc.

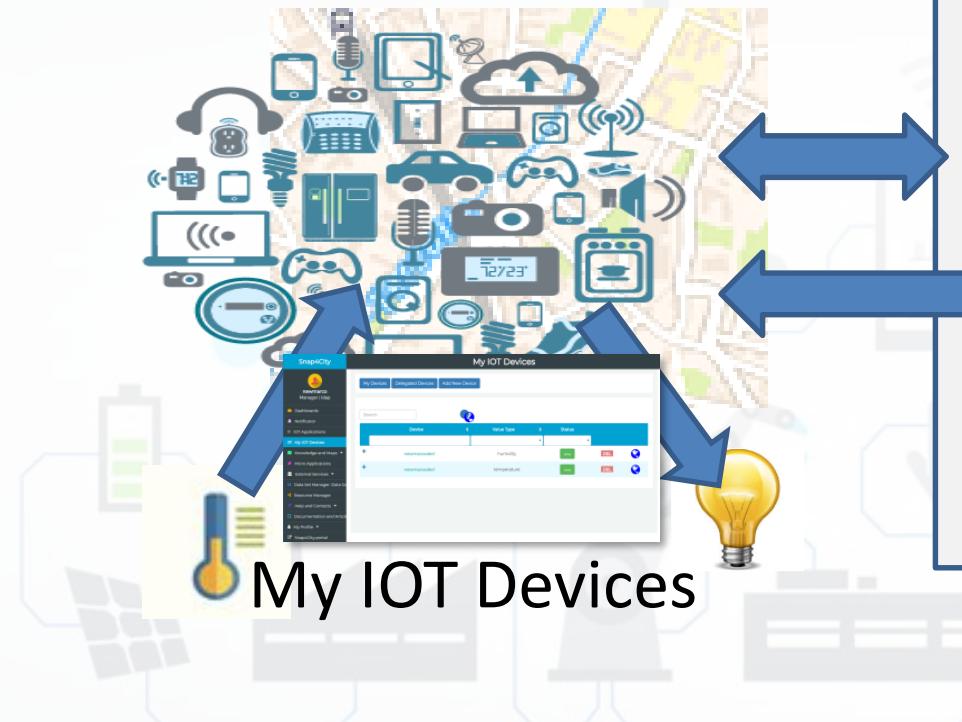


	Open Source end-to-end	Scalability IoT	Execution scalability	Visual Programming end-to-end applications	Advanced Smart City API, MicroServices	Multi Domain Semantic Platform	External services via API	Standard based Modules and IoT, Open Devices	Integrated Community management	Resource Sharing	Referral data management	Security end-to-end	Dashboard 24/7	Flexible and easy dashboard creation	Multi-protocol on IoT
<b><i>Snap4City</i></b>	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
<b><i>KAA</i></b>	Y	Y	Y <sup>1</sup>	N	Y	N	Y	N/Y	Y	N	--	Y	Y	N	Y
<b><i>AWS</i></b>	N	Y	Y	N	N	N	Y	Y	N	Y	Y	Y	Y	(Y)	Limited
<b><i>Azure IOT</i></b>	N	Y	Y	(Y)	N	N	Y	Y	(Y)	Y	Y	Y	Y	(Y)	Limited
<b><i>IOT IGNITE</i></b>	Y	Y <sup>2</sup>	N	Y	N	N	Y	N	N	N	--	N	Y	(Y)	MQTT
<b><i>PTC</i></b>	N	Y	(Y)	Y	N	N	Y	Y	N	N	--	Y	Y	(Y)	Y
<b><i>ThingWorx</i></b>	Y	N	N	N	N	Y	--	Y	N	N	--	N	N	N	Y
<b><i>BEZIRK</i></b>	Y	N	N	N	N	Y	--	Y	N	N	--	N	N	N	Y
<b><i>Bosch IoT Suite</i></b>	N	Y	(Y)	Y	Y	N	Y	Y	N	N	Y	Y	Y	(Y)	Y
<b><i>FIWARE</i></b>	Y	(Y)	N	N	Y	N	N	Y	N	N	N	N	Y	N	Y
<b><i>CISCO Jasper</i></b>	N	Y	N	N	N	N	Y	N	--	--	Y	--	Y	--	N <sup>3</sup>
<b><i>IBM Watson IoT</i></b>	(N)	Y	(Y)	Y	Y	Y	Y	Y	N	Y	(y)	Y	Y	Y	Y
<b><i>Siemens MindSphere</i></b>	N	Y	--	Y	N	N	N	Y	N	N	Y	N	Y	N	Y
<b><i>Carriots</i></b>	N	Y	--	N	N	N	Y	--	N	N	--	N	Y	Y	MQTT
<b><i>Thingsboard</i></b>	Y	Y <sup>4</sup>	N	N	N	N	N	N	N	N	--	Y	Y	Y	(MQTT, CoAP, http)
<b><i>IOT eclipse.org</i></b>	Y	Y	N	N	N	N	Y	Y	N	N	N	N	N	N	Y
<b><i>Google IOT</i></b>	N	Y	Y	N	N	N	Y	N	N	N	N	Y	N	N	MQTT, HTTP

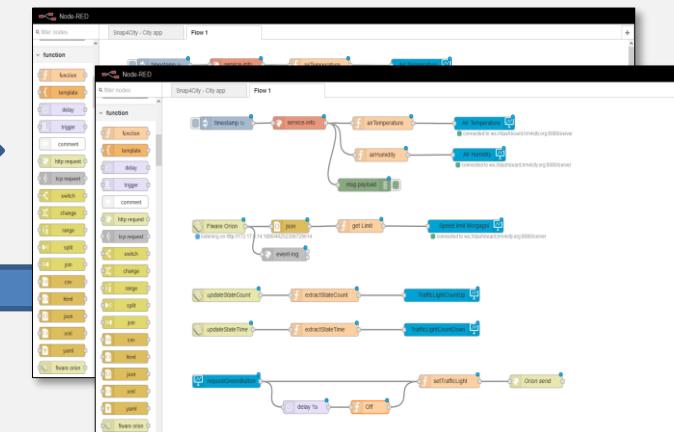
# Level 4 users: dashboard with intelligence App

- Dashboards with IOT Applications for enforcing smart and intelligence into them.

IOT and City data World



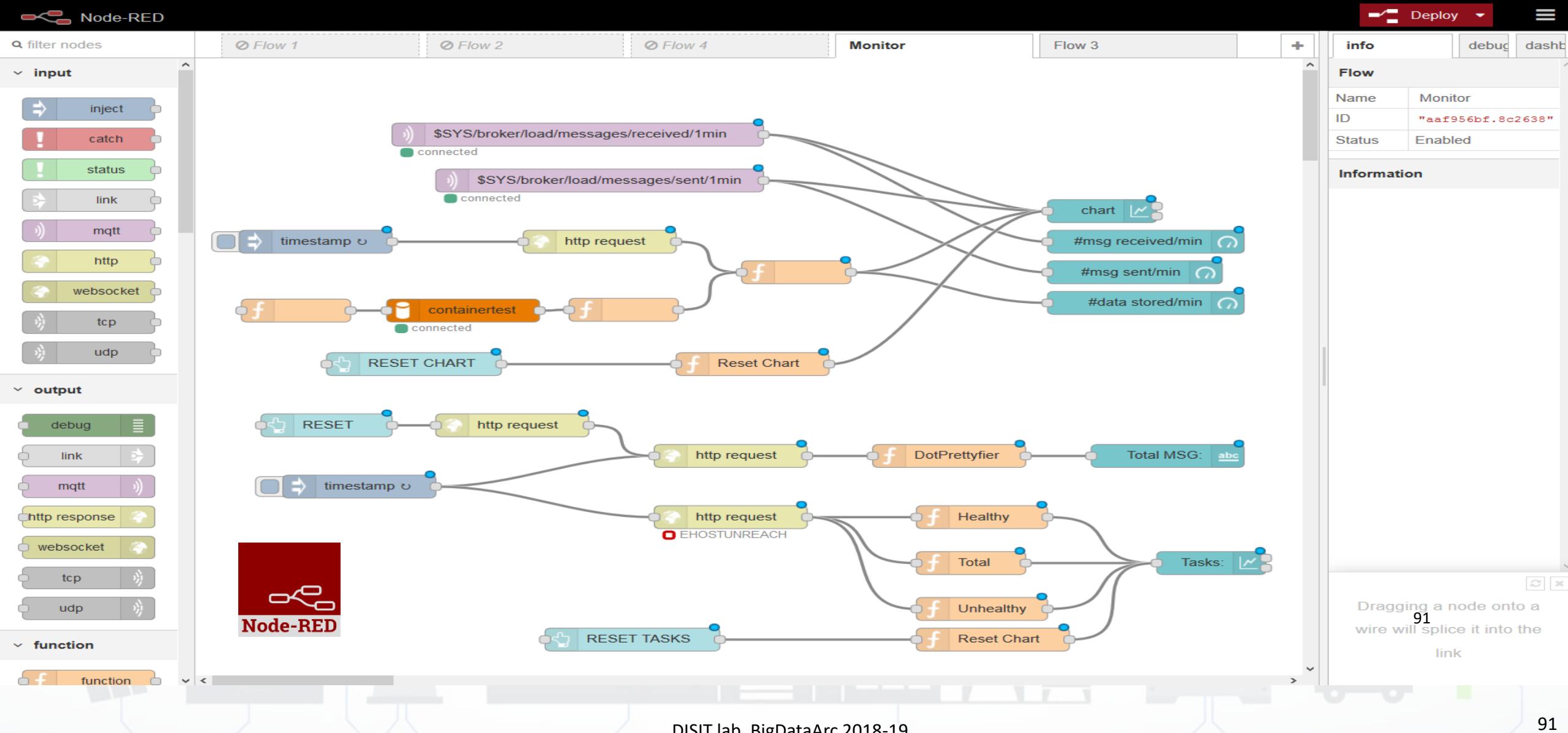
IOT Applications



Dashboards

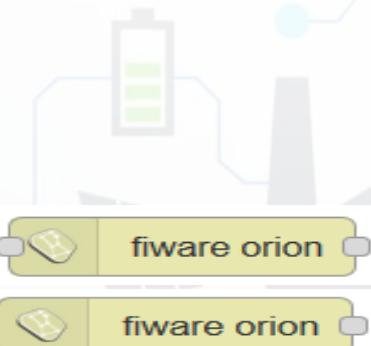
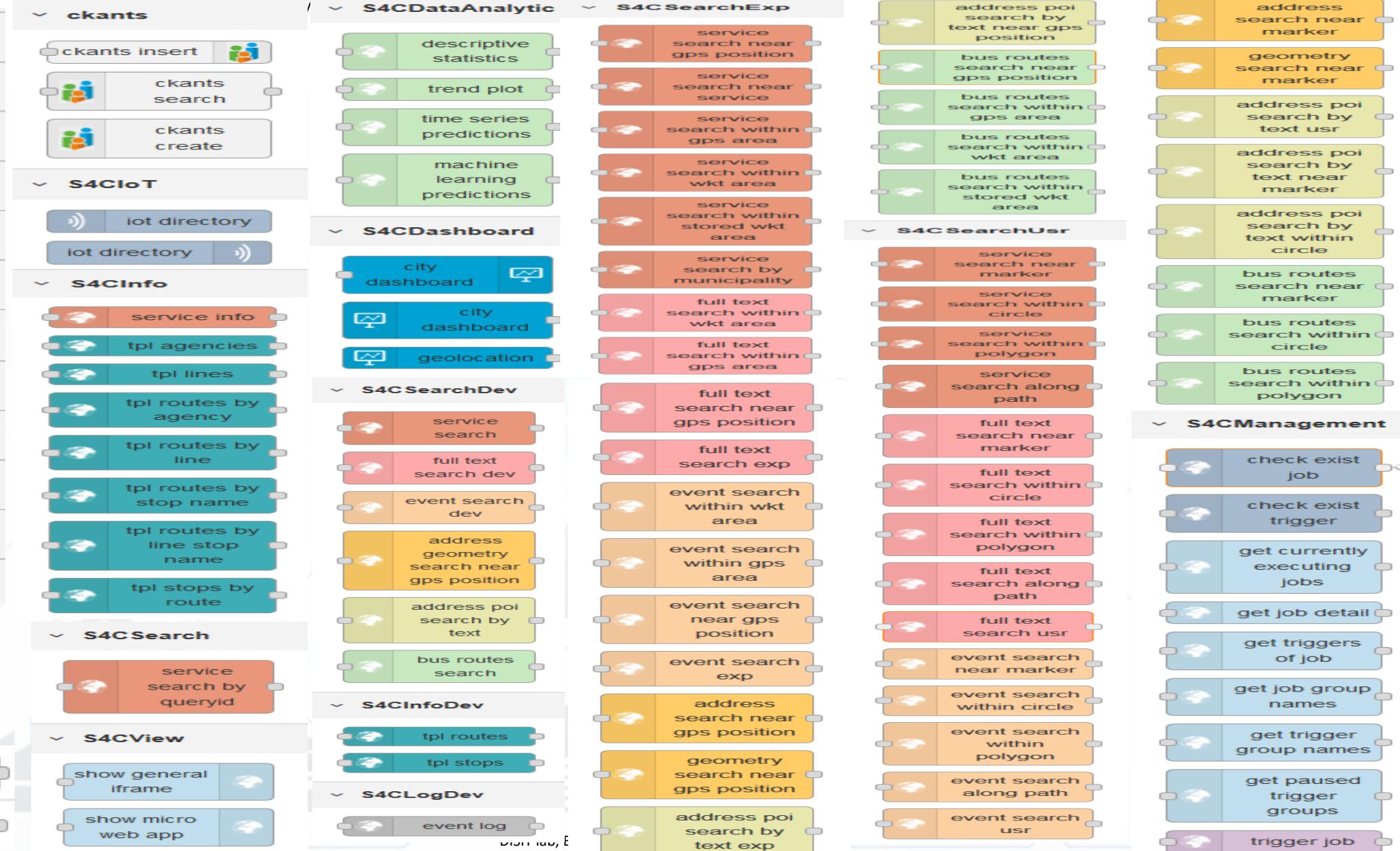
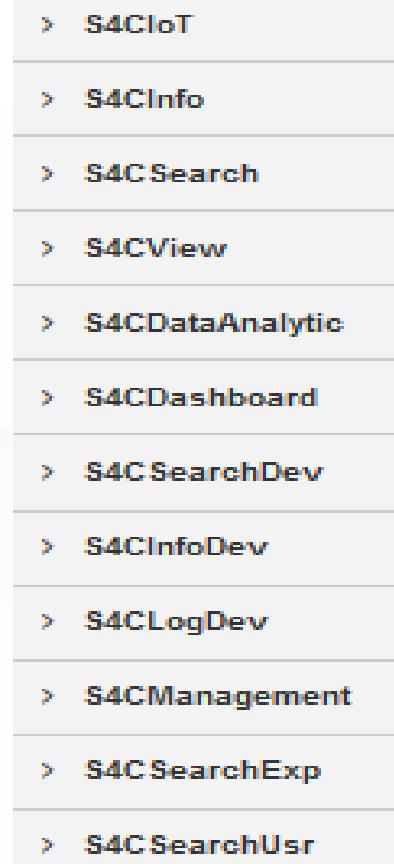


Applications





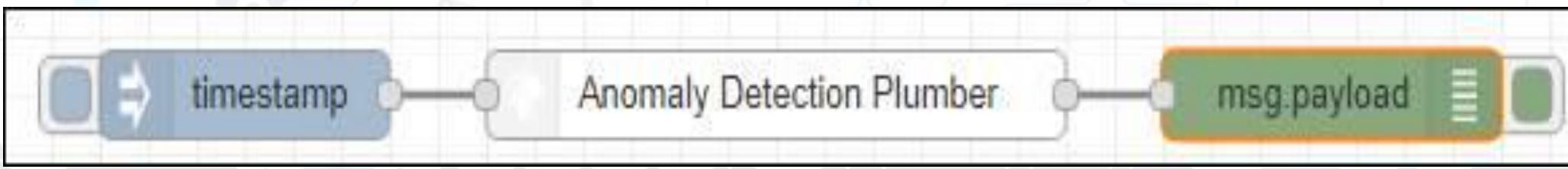
# MicroServices





# Developer in R Studio + Tensor Flow

The screenshot shows a user interface for a data analysis tool. At the top, there's a navigation bar with 'Files', 'Plots', 'Packages', 'Help', and 'Viewer' tabs. Below the navigation bar is a toolbar with icons for 'New Folder', 'Upload', 'Delete', 'Rename', and 'More'. The main area has a breadcrumb path: 'Home > Snap4City > StatisticsOutput'. A sidebar on the left lists files with checkboxes next to them, including 'AverageSpeedDailyTrend.png' which is circled in red with a red arrow pointing to it. A large red box highlights the text: 'Click on each .png file to visualize the statistics: a new tab will be opened'. To the right of the sidebar are several plots: a line chart showing speed trends over time, a scatter plot with a color scale, and a correlation matrix heatmap.





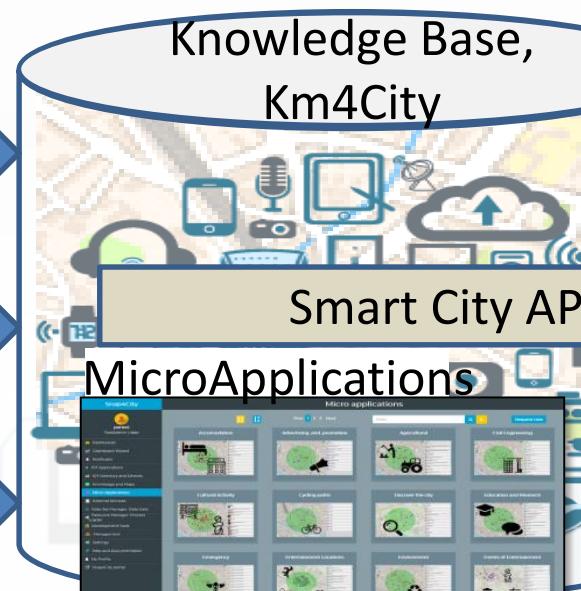
## IOT Directory

## Ontology

## SPARQL, FLINT

## LOG.disit.org

## Knowledge Base, Km4City



ServiceMap

ServiceMap3D

Smart City API from Knowledge Base and other tools

## MicroApplications

## Swagger

## MicroServices

## Back Office Processes

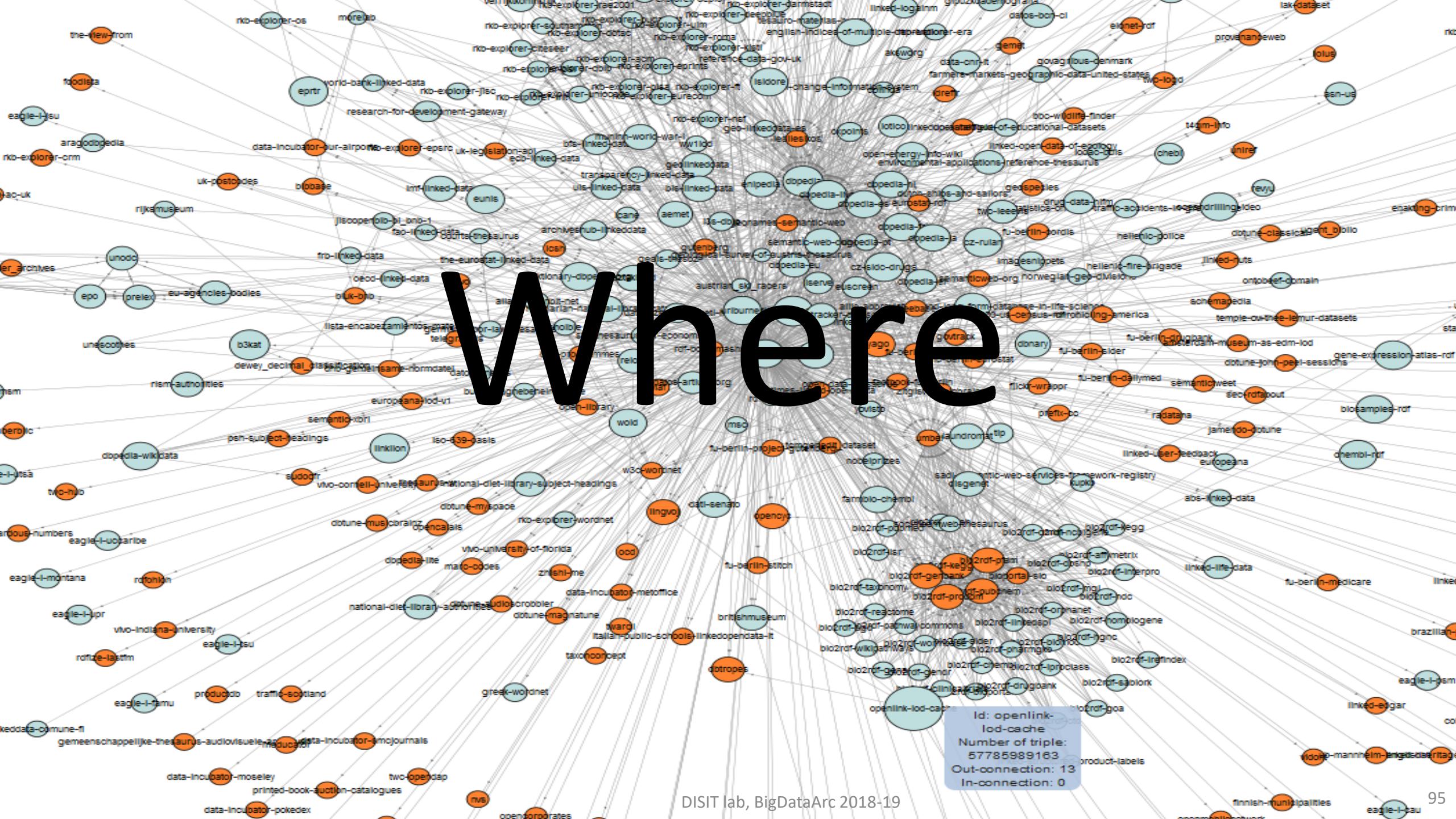
DISCES and back office management tools

## Resource Manager

## Web and Mobile

## IOT Applications

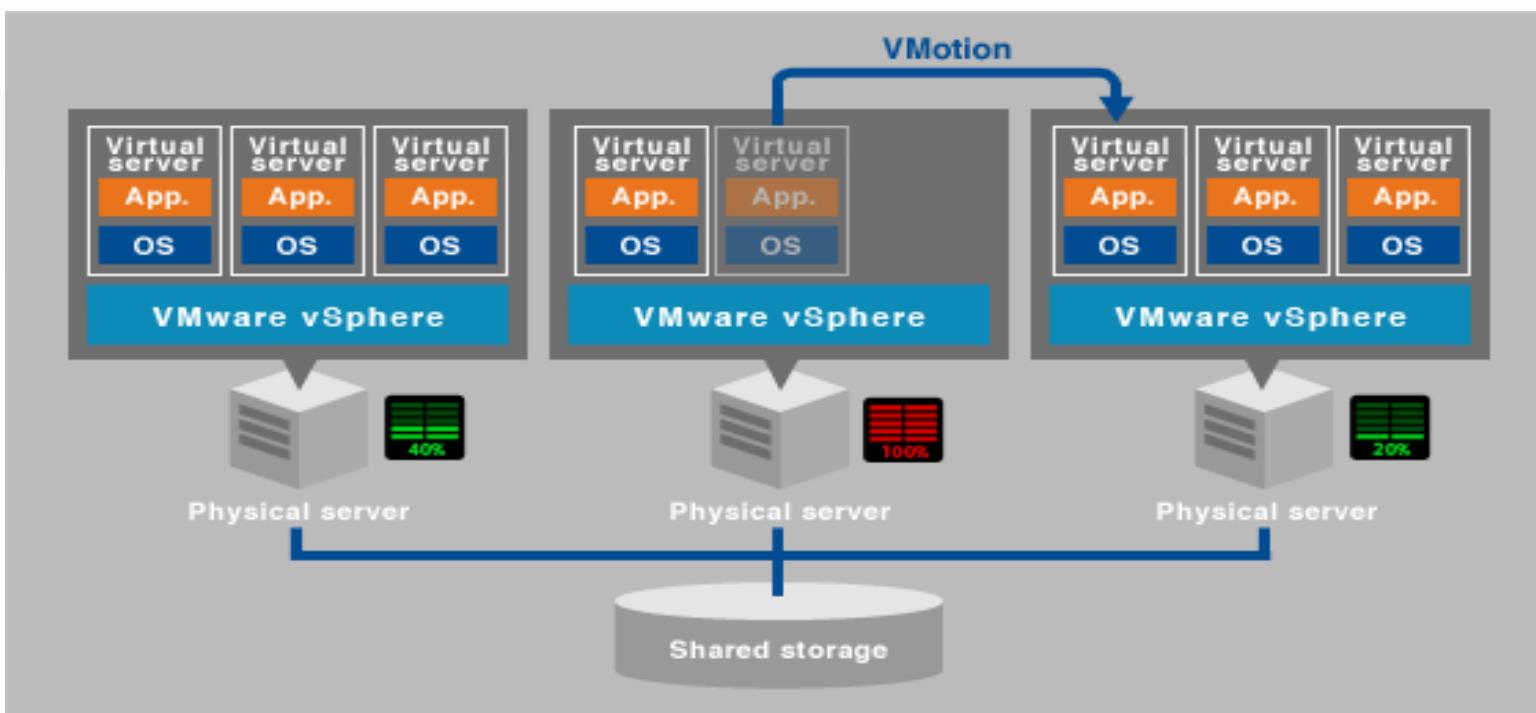
# Where





# Virtualizzazione e Cloud

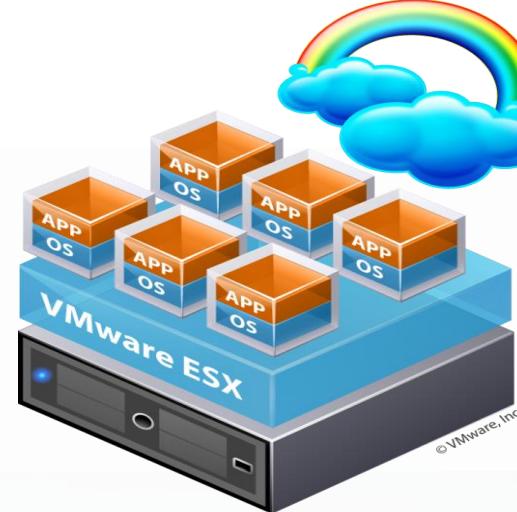
- HA: High Availability,  
DRS: Distributed Resource Scheduler





# Cloud computing e Virtualizzazione

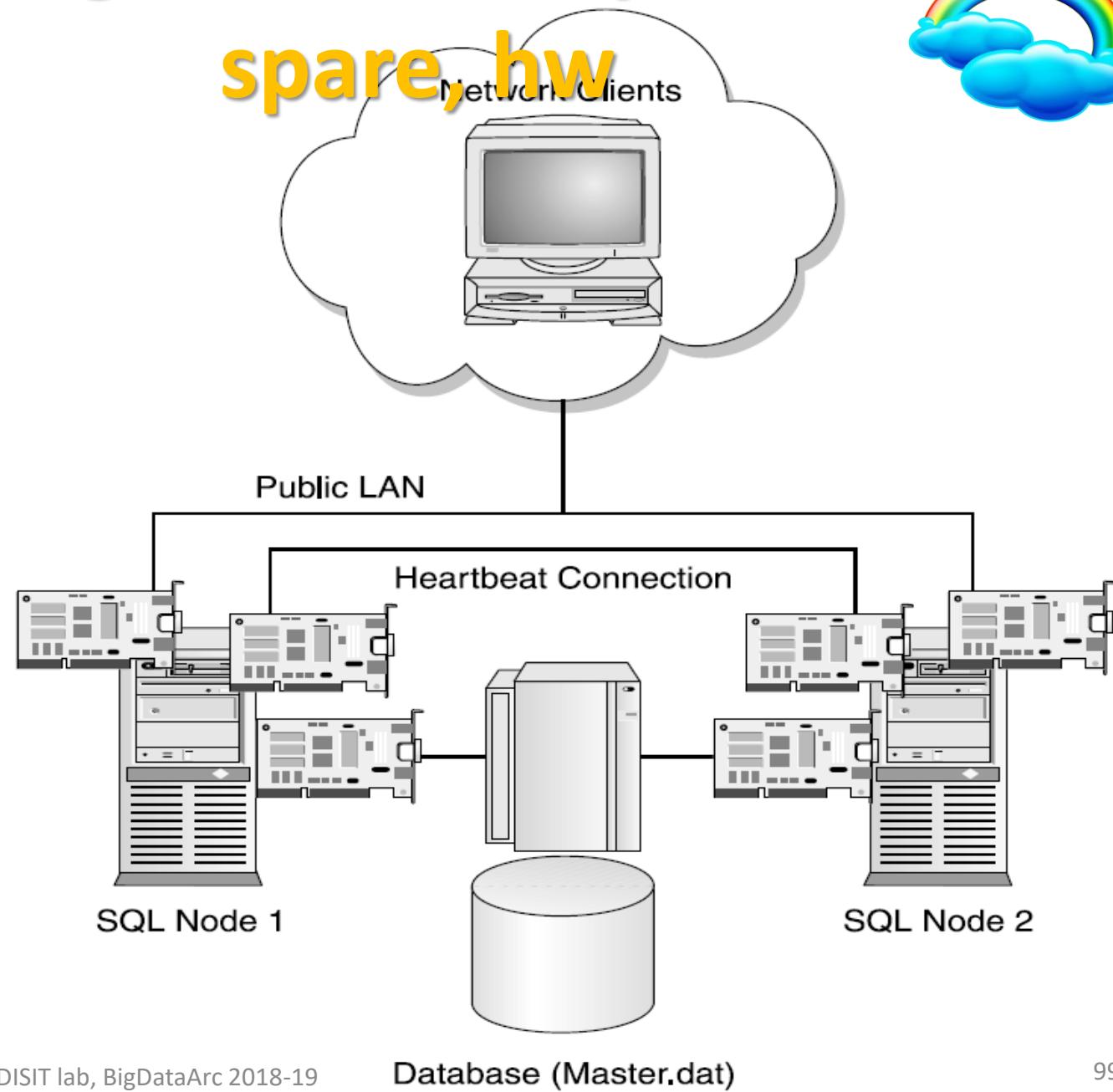
- **Motivations for Cloud computing and Virtualization**
- **Virtual Machine concepts:** emulation, para-virtualization, snapshots
- **Cloud Computing,** cloud vs grid, goals of cloud computing
- **High Availability,** Workload Balancing
- **vSphere Infrastructure,** Vmotion, Power Management, Resource Scheduling, Fault Tolerance
- **Security on the Cloud**
- **Conversions among VM and physical machines**
  - Covnversions: P2V, V2V
- **vCenter, datacenters and cluster management**
  - Performance analysis for the cloud
- **Comparison among Cloud computing solutions**
- **ICARO project**



# High Availability: Hot spare, hw



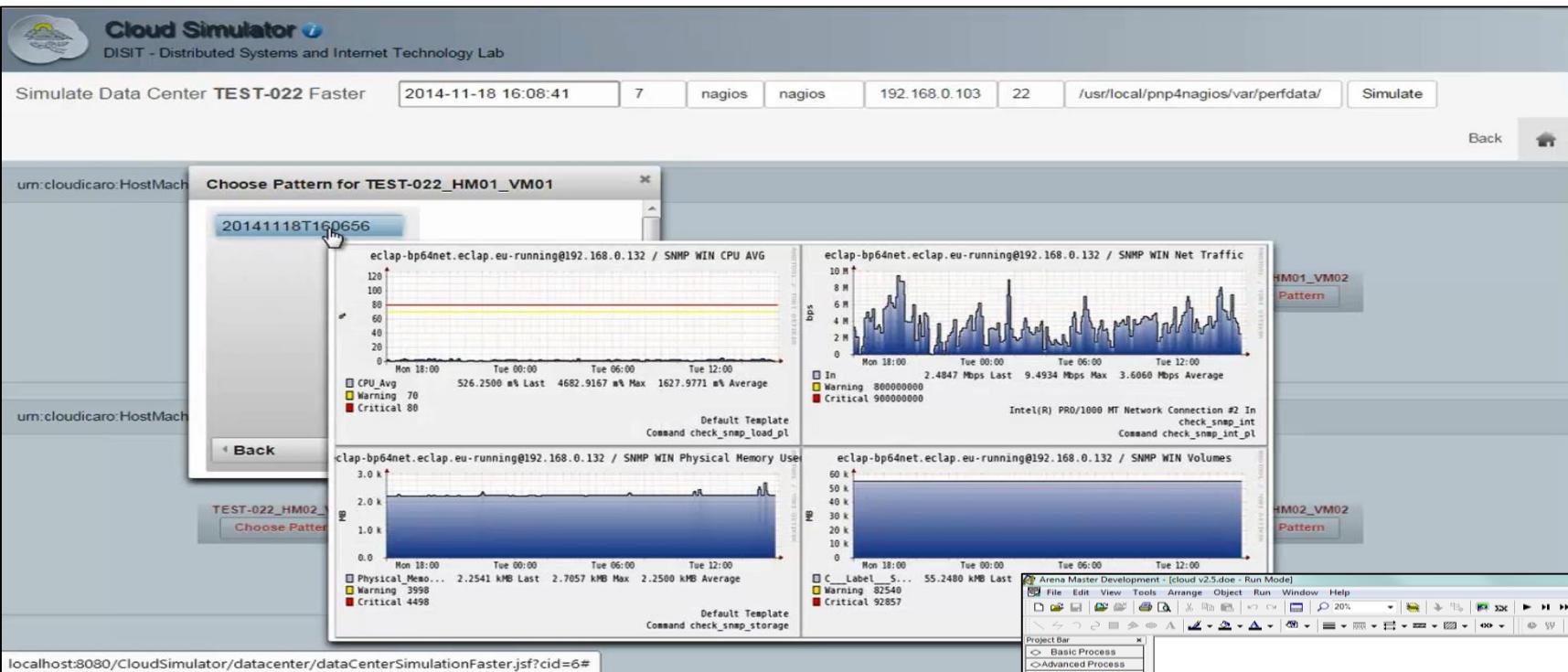
- Three separate networks cards
  - Front end
  - Heartbeat
  - Database NAS/SAN
- UPS/APC solutions with
  - 2 UPS, each of which with network card
- NAS/SAN
  - Raid 5 or 6, 60
  - Fiber connection



# SmartCloud Engine



# Cloud Simulator

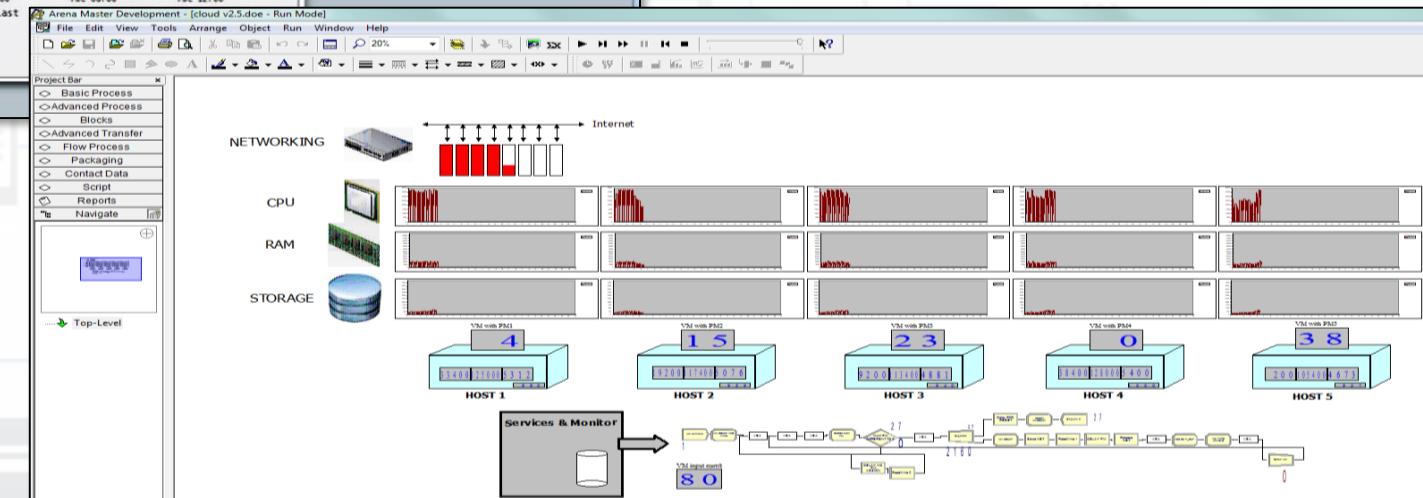


Identification of optimal configurations  
allocations on the basis of effective  
workload

To simulate complex cloud  
configurations



<http://www.cloudicaro.it>



# Agenda

- Laboratorio DISIT
- Tematiche del corso
- Struttura del corso
- Infrastruttura del DISIT Lab
- Modalità dell'esame



# Big Data Architectures

- Overview del costo e del Lab (queste slide)
- Hadoop, MapReduce, Spark
- Batch and Data Stream processing: NIFI, Kafka, ETL, ...
- Indexing and Search: SOLR, SOLR Sharded, Elastic Search
- Cloud: virtualizzazione, HA, DRS, FT, architetture
- Container: Marathon, Mesos, Docker
- Big Data Storage Confronto: Hbase, Mongo
- IOT Architecture: AWS, Azure IOT, Google IOT, Snap4City IOT/IOE
- Visual Analytics: Business intelligence, Dashboard

# Modello del Corso

- Tipicamente per ogni argomento sono presentati:
  - Requisiti e motivazioni dello sviluppo dell'argomento
    - Punto di vista dell'utente e del gestore
  - Stato dell'arte
    - Basi teoriche e tecnologiche, Eventuali standard
    - Prodotti di mercato (leader), pro e contro
  - Recenti Innovazioni e tendenze
  - Confronti fra le varie tecnologie e nuove soluzioni, pro e contro
  - Dettagli progettuali
  - Aspetti prestazionali e di scalabilità
- Seminari di altri studenti e/o esperti, ....

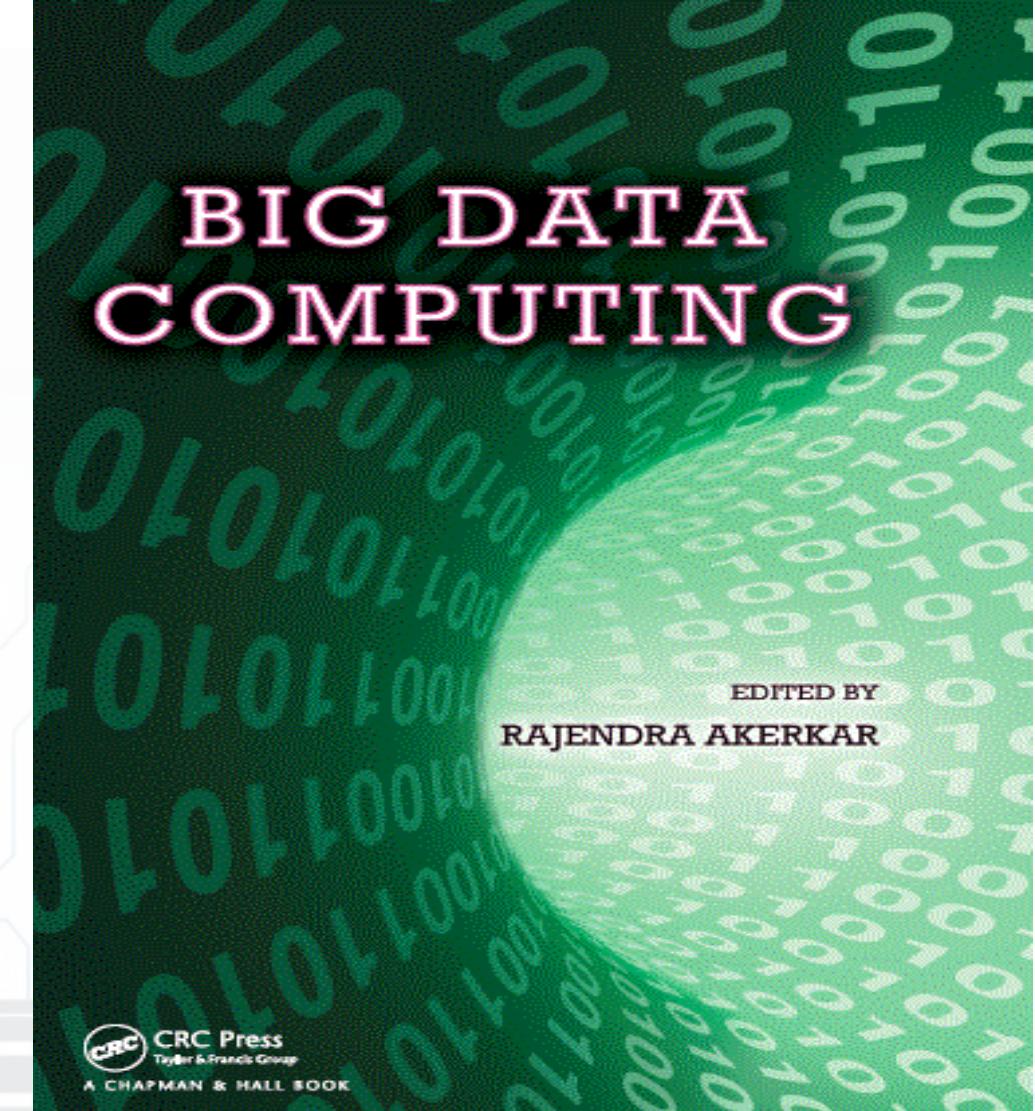
# Ricevimento ed esame

- **Ricevimento per la didattica frontale**
  - In ufficio: Via S. Marta
  - Tutti i Venerdì dalle ore 11:00 alle 13:00
- **Ricevimento per elaborati**
  - *Ogni giorno, dalle 8:00 alle 20:00, inviate una email*
- **Nuova Modalita' per il superamento dell'esame**
- **Eventuali stage e tesi**



P. Bellini, M. Di Claudio, P. Nesi, N. Rauch, "Taxonomy and Review of Big Data Solutions Navigation", in "Big Data Computing", Ed. Rajendra Akerkar, Western Norway Research Institute, Norway, Chapman and Hall/CRC press, ISBN 978-1-46-657837-1, eBook: 978-1-46-657838-8, **july 2013**, in press.

<http://www.tmrfindia.org/bigdata.html>

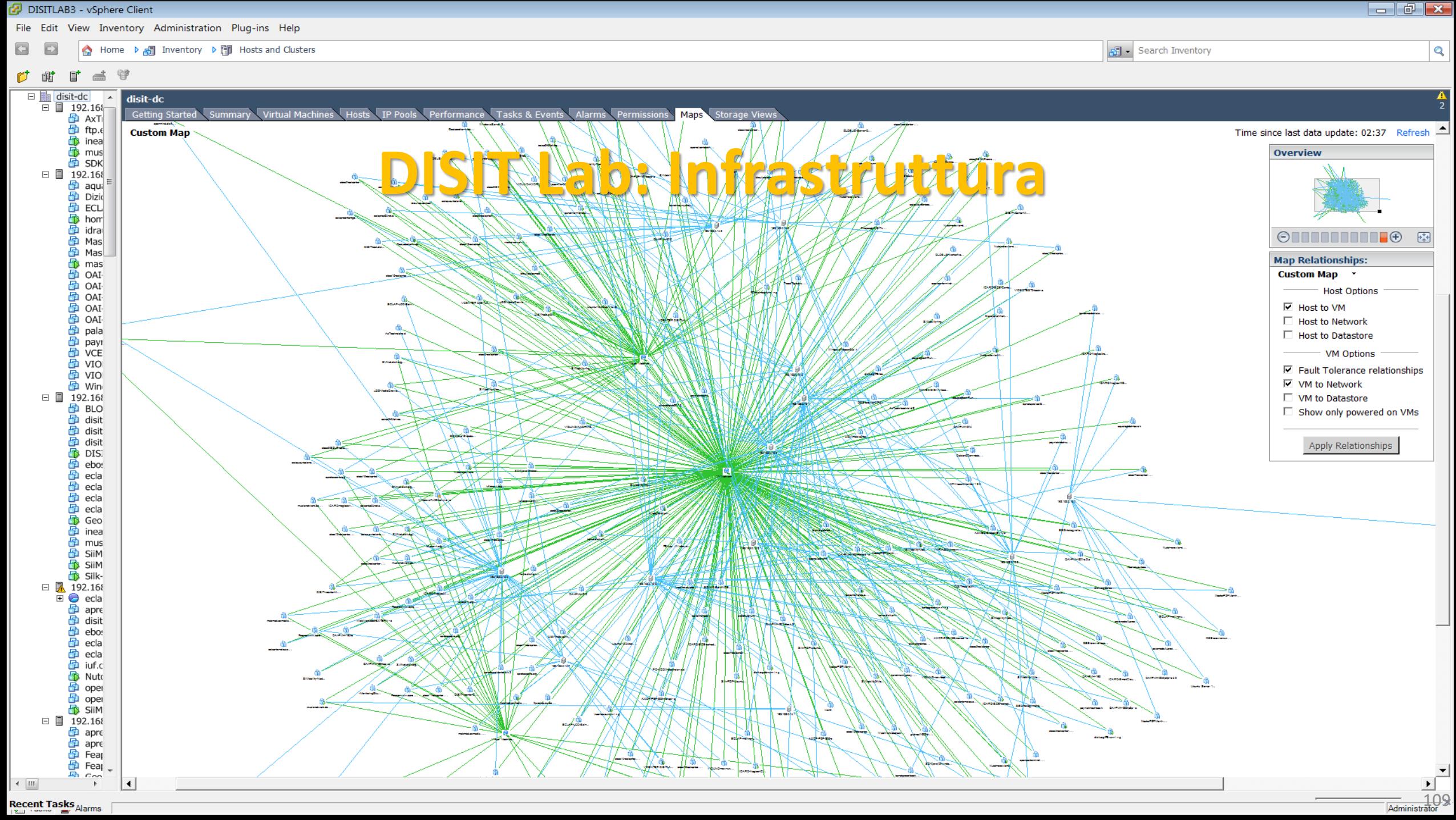


# Agenda

- Laboratorio DISIT
- Tematiche del corso
- Struttura del corso
- Infrastruttura del DISIT Lab
- Modalità dell'esame



- Research group since 1994
- Cloud and data center with >700 TByte storage in raid 50/60,
  - >800 CPU cores, 3800 GPU cores, >4 Tbyte RAM
  - Managing several infrastructure
- IoT center: reference center
- Open Data and Linked Open Data center
  - Integration of more than 800 different Open Data
  - LOD for global linked data <https://LOG.disit.org>
- Nodo CINI per: Big data, Smart City



# Corsi di docenti DISIT

- **Triennale**

- Sistemi Distribuiti – Prof. Paolo Nesi
- Sistemi Operativi – Pierfrancesco Bellini
- Fondamenti di Informatica per Ingegneria Gestionale – Michela Paolucci
- Fondamenti di Informatica per Infermieristica – Gianni Pantaleo

- **Magistrale**

- Big Data Architectures – Prof. Paolo Nesi (Big Data, Architecture, Cloud, IoT)
- Security and Knowledge Engineering – Prof. Pierfrancesco Bellini (Knowledge Engineering, Web Security, Natural Language Processing)

- **Altri corsi:**

- Data Intelligence – Corso di Intelligence e Sicurezza Nazionale – Prof. Paolo Nesi
- Master in Big Data-MABIDA: architetture, Big Data, Knowledge engineering, Natural Language Processing, cloud, etc.

# Agenda

- Laboratorio DISIT
- Tematiche del corso
- Struttura del corso
- Infrastruttura del DISIT Lab
- Modalità dell'esame



# Elaborati

- **Gli elaborati (singoli o coppie di studenti) possono essere di tipo**
  - (A) con sviluppo di software, oppure di tipo
  - (B) compilativi che non implicano lo sviluppo di software (per esempio: confronti fra prodotti, progettazione su carta di soluzioni, valutazione delle prestazioni di prodotti e soluzioni, etc.).
  - Possono essere o meno completati con successo raggiungendo o meno gli obiettivi proposti.
  - Il voto viene stimato sulla base del lavoro svolto su base qualitativa e quantitativa
- **Lo studente può**
  - chiedere la sostituzione dell'elaborato e/o del tutor di laboratorio tramite email al docente.
  - decidere di interrompere l'elaborato in ogni momento chiedendo la valutazione e consegnando la relazione breve di alcune pagine.

# Acknowledgement

- Thanks to the European Commission for funding. All slides reporting logo of **RESOLUTE H2020** are representing tools and research founded by European Commission for the RESOLUTE project. **RESOLUTE** has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement n° 653460).
- Thanks to the European Commission for funding. All slides reporting logo of **REPLICATE H2020** are representing tools and research founded by European Commission for the REPLICATE project. **REPLICATE** has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement n° 691735).
- Thanks to the MIUR for co-founding and to the University of Florence and companies involved. All slides reporting logo of **Sii-Mobility** are representing tools and research founded by MIUR for the Sii-Mobility SCN MIUR project.
- Thanks to the European Commission for funding. All slides reporting logo of **Snap4City** <https://www.snap4city.org> of **Select4Cities H2020** are representing tools and research founded by European Commission for the **Select4Cities** project. **Select4Cities** has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement n° 688196)
- **Km4City** is an open technology exploited by those projects and line of research of DISIT Lab. Some of the innovative solutions and research issues developed into the above mentioned projects are also compliant and contributing to the Km4City approach and thus are contributing to the open Km4City model of DISIT lab.



MINISTERO DELL'ISTRUZIONE DELL'UNIVERSITÀ E DELLA RICERCA

