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

Rights Enforcement and Licensing Understanding for RDF Stores Aggregating Open and Private Data Sets

P. Bellini, L. Bertocci, F. Betti, P. Nesi

Distributed Systems and Internet Technology, DISIT Lab, Univ. of Florence, Italy

http://www.disit.org/, http://www.km4city.org

Paolo Nesi, paolo.nesi@unifi.it

Problem Addressed

- **RDF stores are noSQL (suitable for bigdata) store and are**
  - Used for building knowledge base, KB: by using ontologies and data/triples, making reasoning via SPARQL
  - used to aggregate static and real time data sets; open and/or with many different kinds of licenses
  - based on triples/quadruple, where the 4th element is typically used for identifying the data set
  - used for Cloud, Smart City, Railways, etc.
  - **unsuitable to cope with multiple data sets having multiple licenses with different constraints: location, time, roles..**
- **Developers need powerful tools** for realizing applications:
  - Getting Calls from visual environments
  - Discovering suitable query to obtain desired data
  - A the end, the RDF store has to be to perform actions on data ONLY to who has the right to do that,
    - but not always in development phases, e.g., rights depends on data sets provider, locations, roles, etc.

In the context of Smart City

- **RDF stores are used** for realizing knowledge base RDF endpoints, connected to Smart City API, via REST Call of different kinds ...
  - See IEEE SMARTCOMP paper
- **Accessing** data read/write, exploiting inference, etc..
  - City operators/users having different rights, roles, etc. in different contexts
  - Data with different licenses for different operator and GPS coordinates, and time..
  - ...

Licenses and More..

- **Licenses**: MPEG-21, ODRL, XACML, Xrml, etc..
  - Suitable for media, Unsuitable for data
- **Data Licenses**: CC, ODC, OGL, IODL
  - Mainly open data and declinations
  - **Permissions**: derivative, commercialize, derivative...
  - **Restrictions/duties**: attribution, notice, ...
- **Getting Composing Data Set → Licences Composition is needed**
  - **See www.disit.org/6877 extension**
- Formal models to **grant rights**
- Techniques for **right enforcement/verification**
  - **Almost missing on RDF stores**

Composing Licenses

DISIT extension

	Second License												
First License	CC0	CC-PDM	CC-BY-ND	CC-BY-NC-ND	CC-BY	CC-BY-SA	CC-BY-NC	CC-BY-NC-SA	ODC-PDDL	ODC-BY	ODC-ODbL	OGL 2.0	OS OpenData
CC0	No restrictions	No restrictions	-	-	CC-BY	CC-BY-SA	CC-BY-NC	CC-BY-NC-SA	No restrictions	ODC-BY	ODC-ODbL	OGL 2.0	OS OpenData
CC-PDM	No restrictions	No restrictions	-	-	CC-BY	CC-BY-SA	CC-BY-NC	CC-BY-NC-SA	No restrictions	ODC-BY	ODC-ODbL	OGL 2.0	OS OpenData
CC-BY-ND	-	-	-	-	-	-	-	-	-	-	-	-	-
CC-BY-NC-ND	-	-	-	-	-	-	-	-	-	-	-	-	-
CC-BY	CC-BY	CC-BY	-	-	CC-BY	CC-BY-SA	CC-BY-NC	CC-BY-NC-SA	CC-BY	CC-BY	ODC-ODbL	CC-BY	OS OpenData
CC-BY-SA	CC-BY-SA	CC-BY-SA	-	-	CC-BY-SA	CC-BY-SA	-	-	CC-BY-SA	CC-BY-SA	ODC-ODbL	CC-BY-SA	CC-BY-SA
CC-BY-NC	CC-BY-NC	CC-BY-NC	-	-	CC-BY-NC	-	CC-BY-NC	CC-BY-NC-SA	CC-BY-NC	CC-BY-NC	-	CC-BY-NC	-
CC-BY-NC-SA	CC-BY-NC-SA	CC-BY-NC-SA	-	-	CC-BY-NC-SA	-	CC-BY-NC-SA	CC-BY-NC-SA	CC-BY-NC-SA	CC-BY-NC-SA	-	CC-BY-NC-SA	-
ODC-PDDL	No restrictions	No restrictions	-	-	CC-BY	CC-BY-SA	CC-BY-NC	CC-BY-NC-SA	No restrictions	ODC-BY	ODC-ODbL	OGL 2.0	OS OpenData
ODC-BY	ODC-BY	ODC-BY	-	-	ODC-BY	CC-BY-SA	CC-BY-NC	CC-BY-NC-SA	ODC-BY	ODC-BY	ODC-ODbL	ODC-BY	OS OpenData
ODC-ODbL	ODC-ODbL	ODC-ODbL	-	-	ODC-ODbL	ODC-ODbL	-	-	ODC-ODbL	ODC-ODbL	ODC-ODbL	ODC-ODbL	ODC-ODbL
OGL 2.0	OGL 2.0	OGL 2.0	-	-	CC-BY	CC-BY-SA	CC-BY-NC	CC-BY-NC-SA	OGL 2.0	ODC-BY	ODC-ODbL	OGL 2.0	OS OpenData
OS OpenData	OS OpenData	OS OpenData	-	-	OS OpenData	CC-BY-SA	-	-	OS OpenData	OS OpenData	ODC-ODbL	OS OpenData	OS OpenData

RDF state of the art

- **Fuseky-Jena, GraphDB** support access control to the whole repository, not at level of data set/graph.
  - **Jena** provides API to write JAVA processes for filtering triples.
- **ORACLE** support access control to users at level of triple and model, not on graphs
- **Virtuoso and Stardog** allow to formalize simple licenses (as read/write permissions) at level of data set (RDF graph), and associate them to users.
  - an user performing a SPARQL query get back only triples for which is authorized **without any explanation** about filtered triples, and thus about potentially accessible data set with a different user profiles and licenses.

How is possible

- **SPARQL analyzer**
  - Read SPARQL query and rewrite to ask at the RDF store the Union of all the Graphs involved in the query
  - Several different constructs are addressed

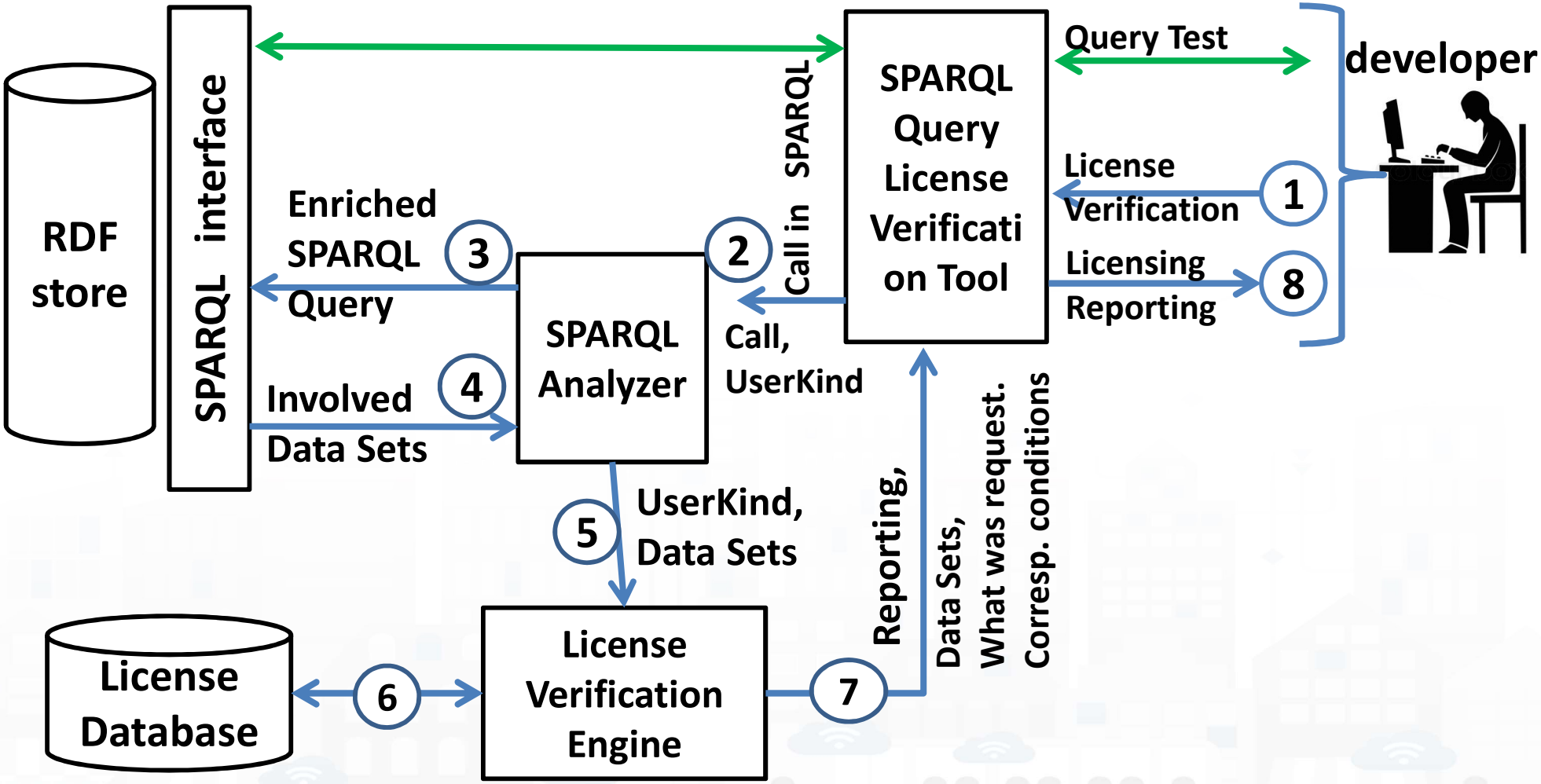
$$Graphs(Q) = \bigcup_{i=1}^{n} Graphs(Q.subQ_i) \cup G(Q)$$

- **License Verification Engine**
  - To compute the Duties and Permissions according to the category for which the query has been requested.

$$allow(Q,c,p) = \bigwedge_{g \in ResultGraphs(Q)} allow(g,c,p)$$

$$require(Q,c,d) = \bigvee_{g \in ResultGraphs(Q)} require(g,c,d)$$

SPARQL Query License Verification Tool



License Combination example

dataset/graph description	license	SD	Duties		permisions				user categories				
		sharealike	attribution	notice	derivative	commecialize	redistribute	reproduce	Citizen	Tourist	Police	Civil protection	Firefighters
DigitalLocation	CC-By-NC-SA	✔	✔	✔	✔	✘	✔	✔	✔	✔	✔	✔	✔
Eneergy Cabins	protected	✘	✘	✘	✘	✘	✘	✔	✘	✘	✔	✔	✔
Commercial firms	CC-By	✘	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
Graph street	CC-By	✘	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
Services on the city	CC-By-NC	✘	✔	✔	✔	✘	✔	✔	✔	✔	✔	✔	✔
Renting bikes	CC-By	✘	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
Taxi	CC-By	✘	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
Enogastronomy	CC-By	✘	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔

Two Main Cases

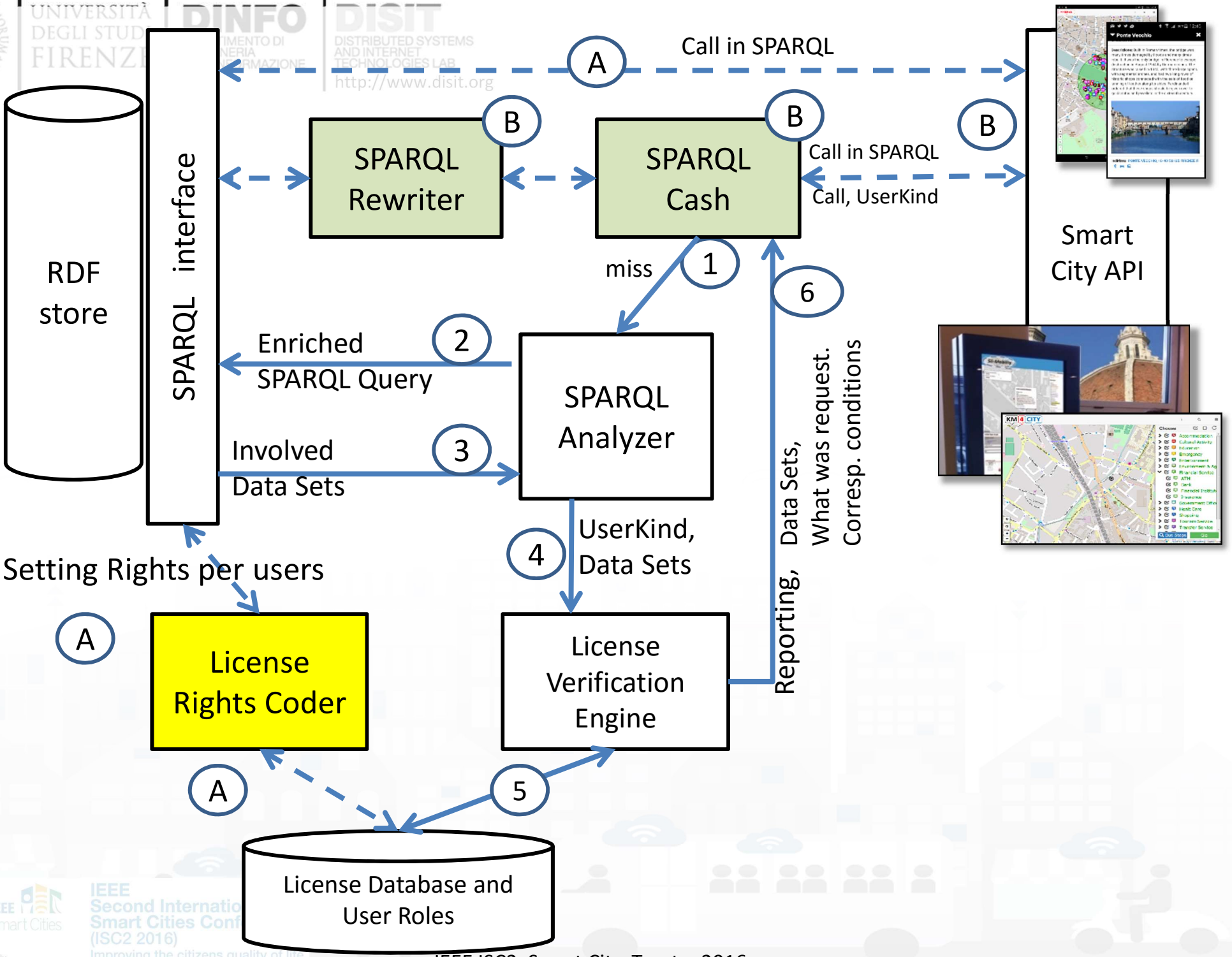
The solution allows:

- Developers to test and validate queries
- Cope with all kinds of RDF Stores

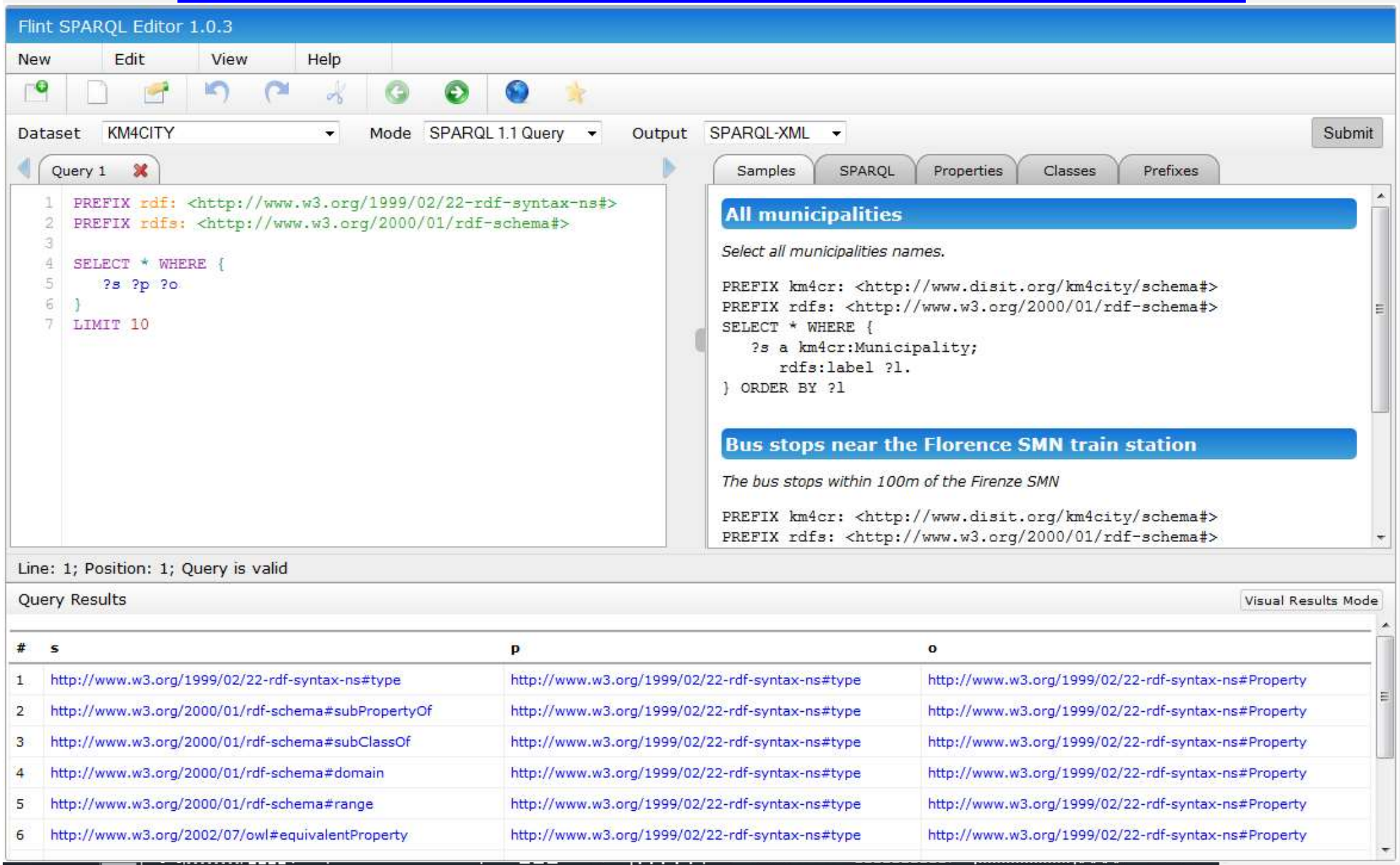
RDF Stores can be classified in two cases, those:

A. *provide* Triple filtering according to rights
   - Set up the correct rights into the RDF store via the **License Rights Coder**

B. *does not provide* triple filtering support,
   - provide support for **rights enforcement via query rewriting avoiding triple filtering**


A
Call in SPARQL
B
SPARQL Rewriter
B
SPARQL Cash
B
Call in SPARQL
Call, UserKind
RDF store
SPARQL interface
miss
1
6
Smart City API
Enriched SPARQL Query
2
SPARQL Analyzer
Involved Data Sets
3
UserKind, Data Sets
4
Reporting, Data Sets, What was request. Corresp. conditions
Setting Rights per users
A
License Rights Coder
License Verification Engine
A
5
License Database and User Roles

http://log.disit.org/spqlquery/



#	s	p	o
1	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://www.w3.org/1999/02/22-rdf-syntax-ns#Property
2	http://www.w3.org/2000/01/rdf-schema#subPropertyOf	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://www.w3.org/1999/02/22-rdf-syntax-ns#Property
3	http://www.w3.org/2000/01/rdf-schema#subClassOf	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://www.w3.org/1999/02/22-rdf-syntax-ns#Property
4	http://www.w3.org/2000/01/rdf-schema#domain	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://www.w3.org/1999/02/22-rdf-syntax-ns#Property
5	http://www.w3.org/2000/01/rdf-schema#range	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://www.w3.org/1999/02/22-rdf-syntax-ns#Property
6	http://www.w3.org/2002/07/owl#equivalentProperty	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://www.w3.org/1999/02/22-rdf-syntax-ns#Property

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

www.km4city.org

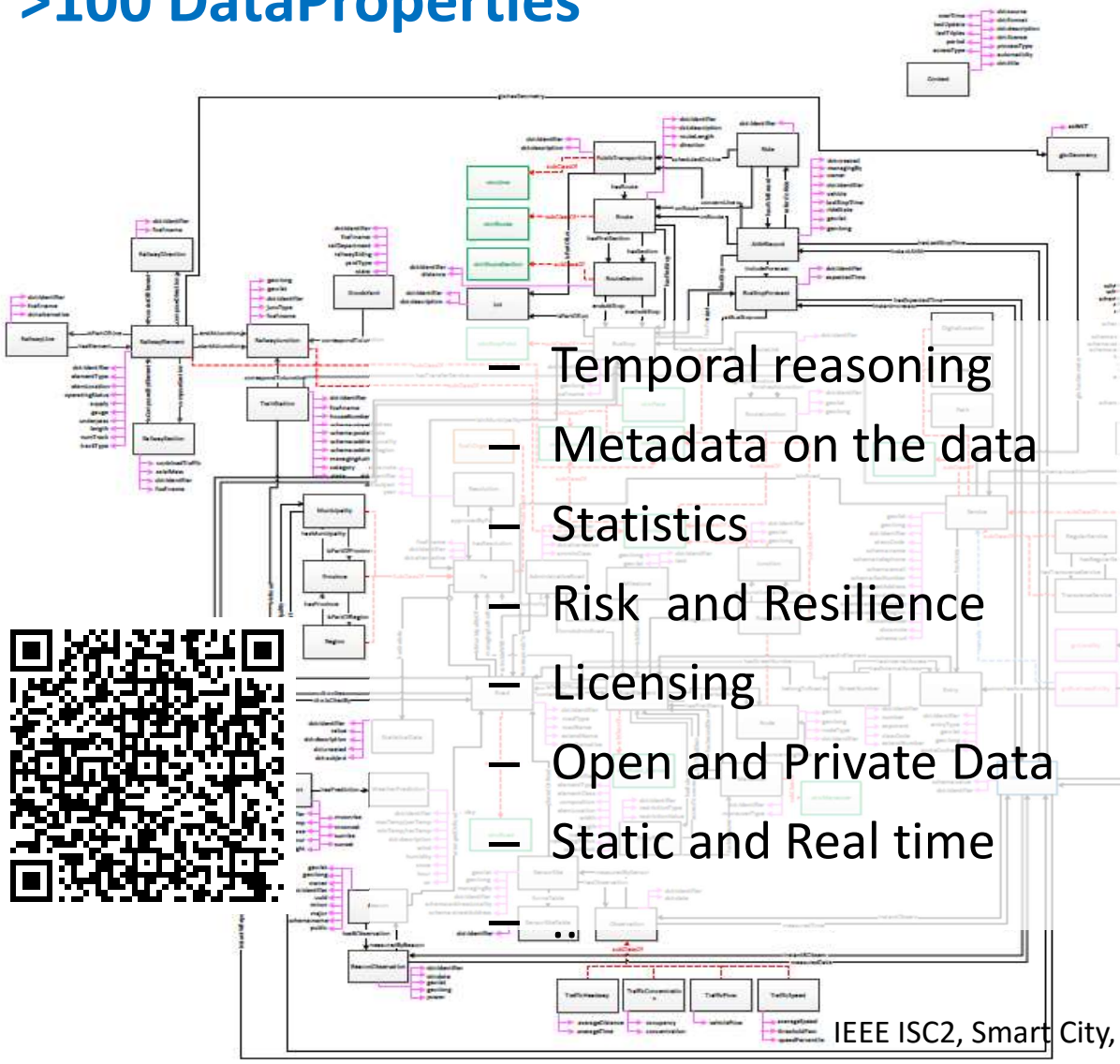
Technical info on: http://www.disit.org/km4city

km4City Ontology

>84 Classes

>100 ObjectProperties

>100 DataProperties



- to cover different aspects:
  - Street-Guide
  - Mobility and transport
  - Points of interest
  - Sensors, IOT, ..
  - Energy
  - Administration
  - Citations from strings
  - ..

- Temporal reasoning
- Metadata on the data
- Statistics
- Risk and Resilience
- Licensing
- Open and Private Data
- Static and Real time
- ..

Ontology Documentation:

http://www.disit.org/6506
http://www.disit.org/6507
http://www.disit.org/5606
http://www.disit.org/6461

Transport systems
Mobility, parking

Public Services
Govern, events,

Sensors, IOT
Cameras, ..

Environment,
Water, energy

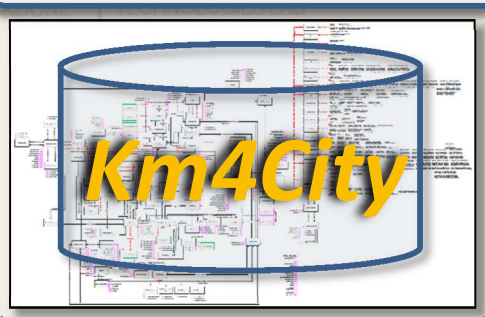
Shops, services,
operators

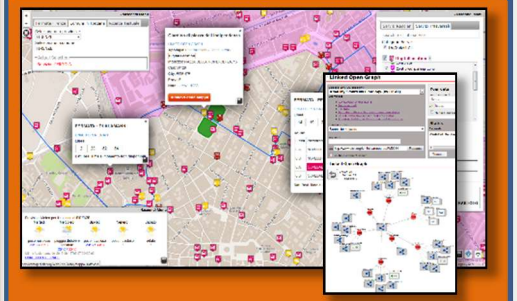
Social Media
WiFi, network

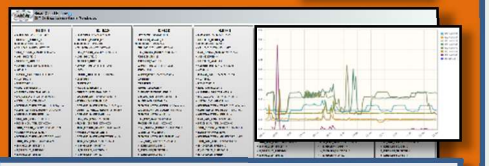
Static, Slow and Real Time data flows

DISCES -- Distributed and parallel architecture on Cloud

Km4City Smart City Engine

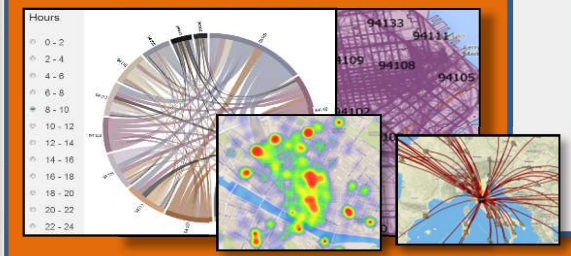






User Profiling and Suggestion Engine

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

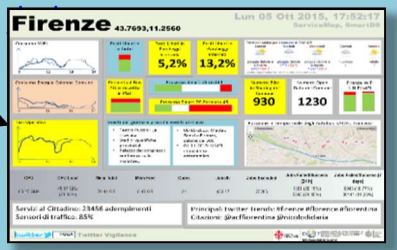


Km4City Tools for Developers

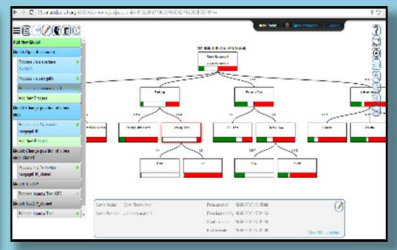
Km4City Smart City API

Tools for City Operators and Decision Makers

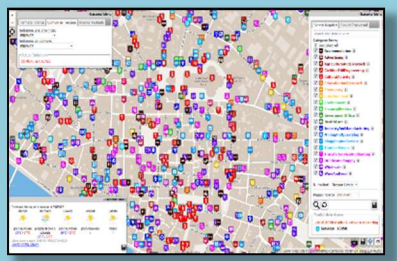
Smart City Dashboard
Http://dashboard.km4city.org/



Smart Decision Support
Http://Smartds.km4city.org

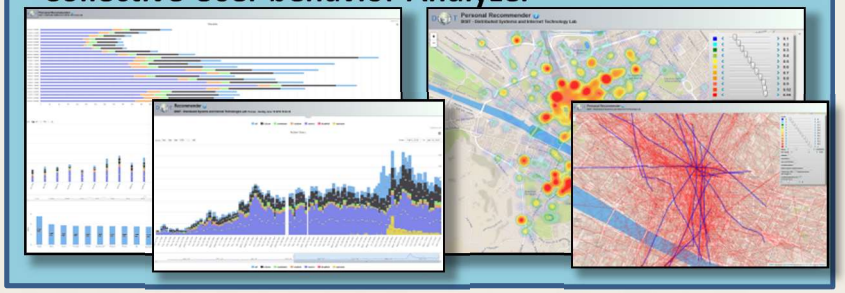


Service map browser
Http://servicemap.km4city.org



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

Collective User behavior Analyzer

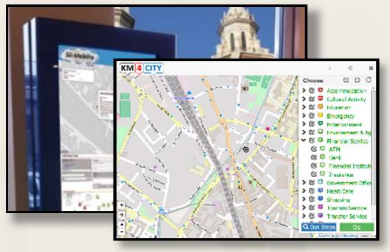


Tools for Final Users

Mobile e Web Apps

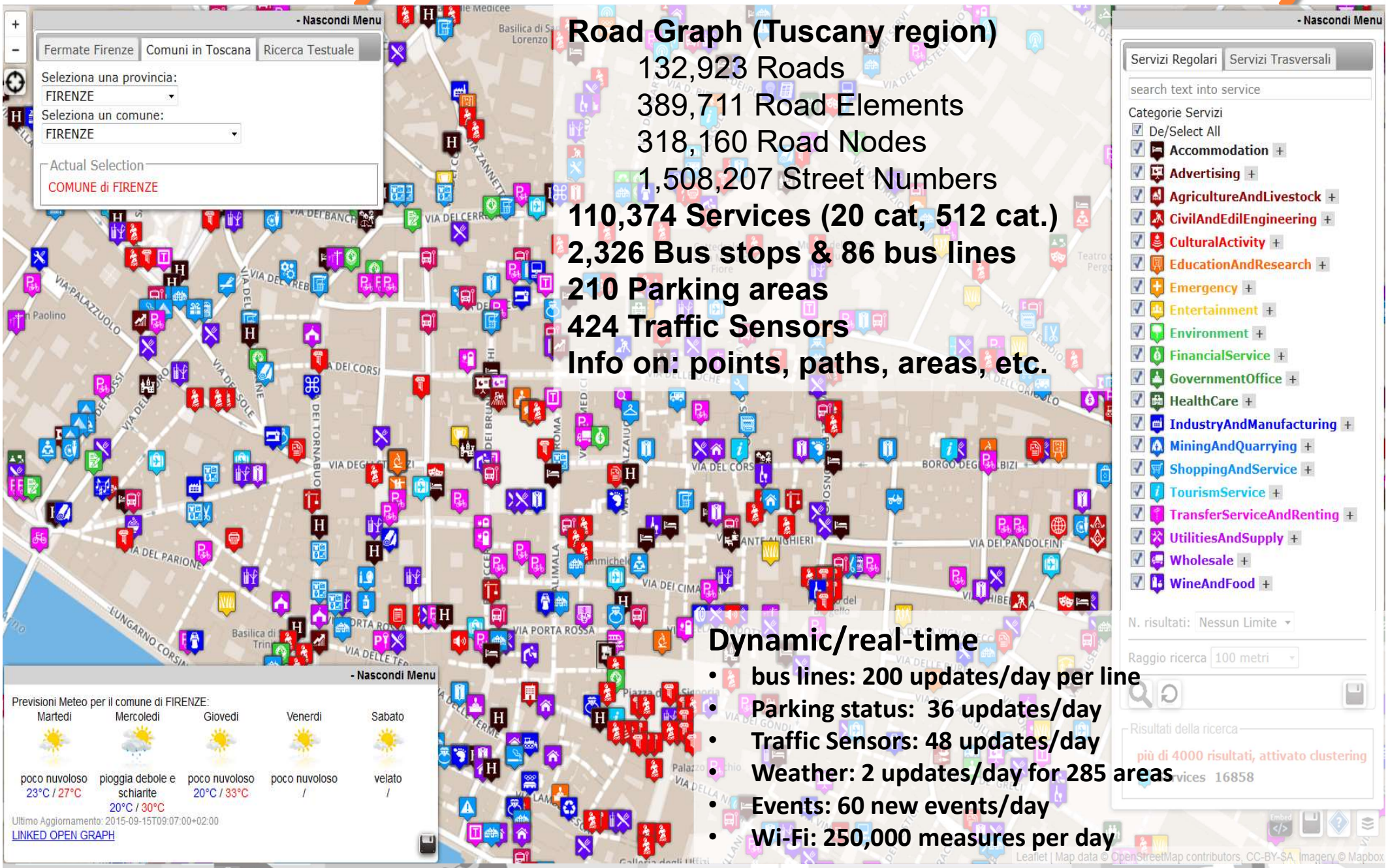
Http://www.km4city.org/app

Http://www.km4city.org



www.km4city.org

Km4City on Firenze & Tuscany



Road Graph (Tuscany region)

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

110,374 Services (20 cat, 512 cat.)

2,326 Bus stops & 86 bus lines

210 Parking areas

424 Traffic Sensors

Info on: points, paths, areas, etc.

Dynamic/real-time

- **bus lines: 200 updates/day per line**
- **Parking status: 36 updates/day**
- **Traffic Sensors: 48 updates/day**
- **Weather: 2 updates/day for 285 areas**
- **Events: 60 new events/day**
- **Wi-Fi: 250,000 measures per day**

UNIVERSITÀ DEGLI STUDI FIRENZE

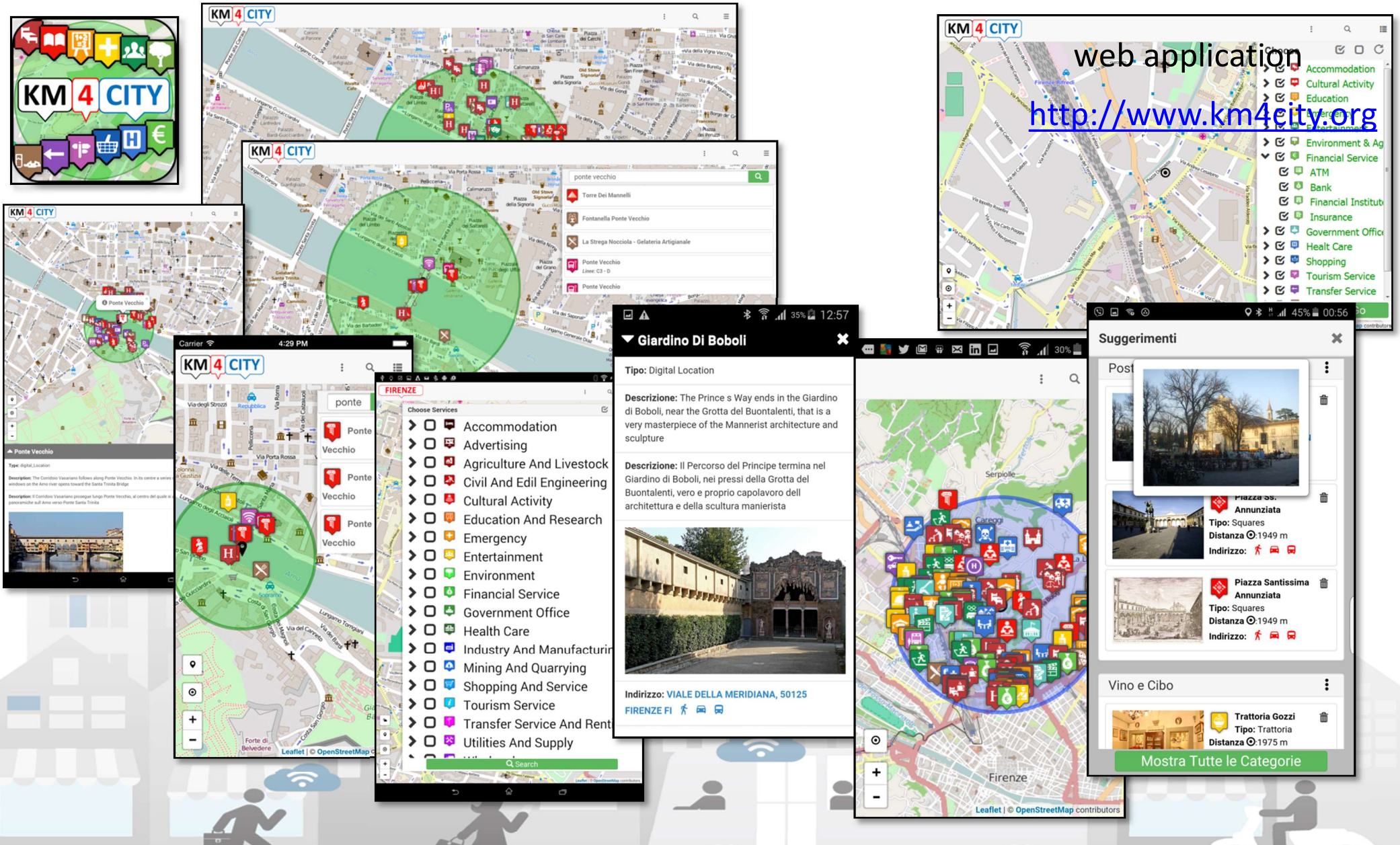
DINFO DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

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

Km4CityMobile App

web application

http://www.km4city.org



Projects based on KM4City

- Sii-Mobility, http://www.sii-mobility.org

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

RESilience management guidelines and Operationalization appLied to Urban Transport Environment

Horizon 2020 European Union Funding for Research & Innovation

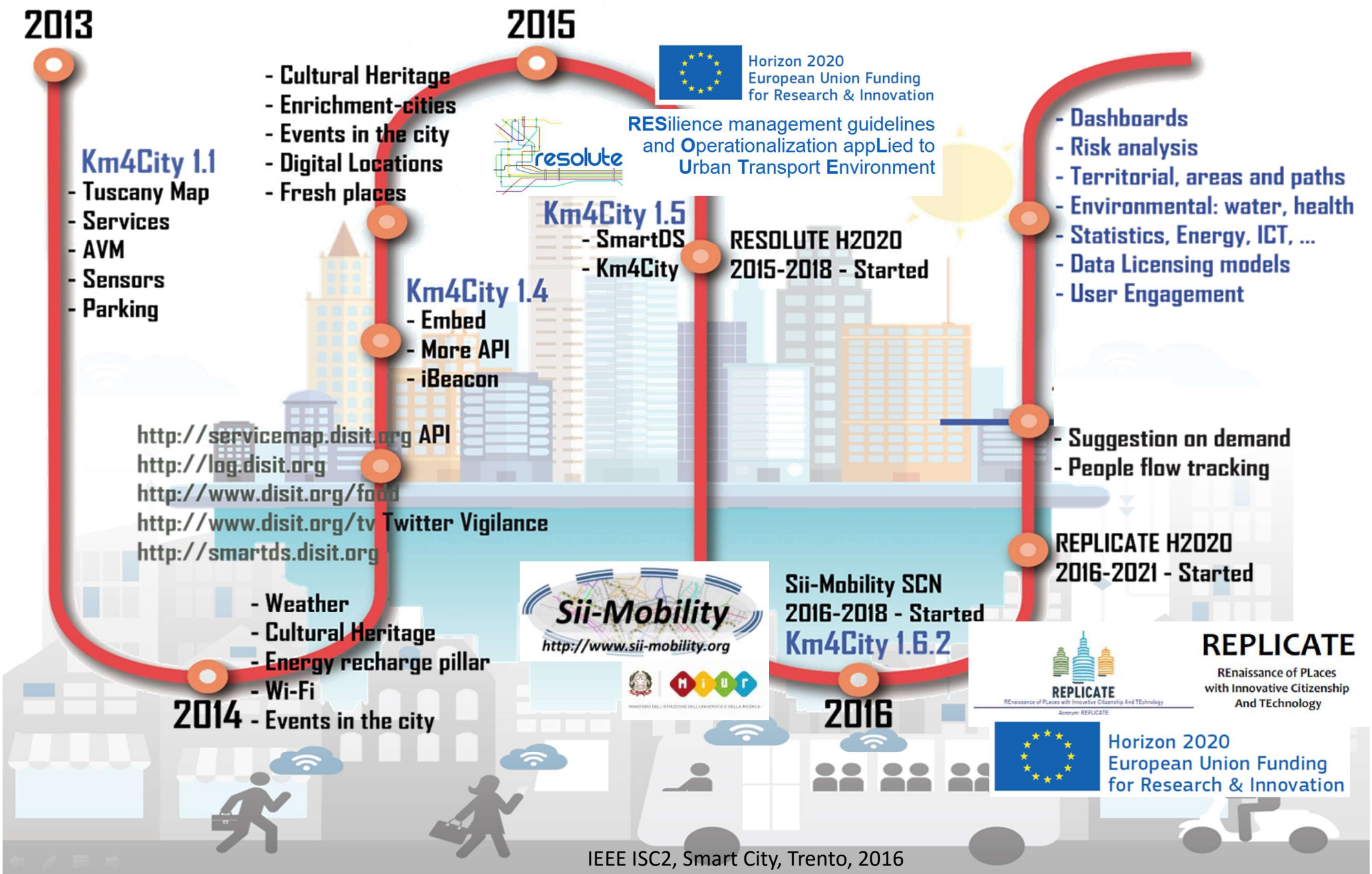
- Replicate, http://www.disit.org/6778

REPLICATE
REnaissance of PLaces with Innovative Citizenship And TEchnology

Horizon 2020 European Union Funding for Research & Innovation

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

km4City Roadmap



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

Km4City

Smart City Ecosystem

http://www.disit.org/km4city

Paolo Nesi, paolo.nesi@unifi.it

Smart City EcoSystem

https://github.com/disit

- **Decision Makers, Public administrator tools:**
  - ✓ Smart City Dashboards, http://dashboard.km4city.org → Dashboard Builder
  - ✓ Resilience Decision Support, http://resilienceds.km4city.org
  - ✓ Smart decision support system, http://smartds.km4city.org
  - ✓ Twitter Vigilance, http://www.disit.org/tv
  - ✓ Recommender and User Behavior Analyzer, http://recommender.km4city.org
  - ✓ WiFi monitor, http://wifimap.km4city.org
  - ✓ ServiceMap Server, http://servicemap.km4city.org
  - ✓ Traffic and People Flow Assessment http://www.disit.org/6694
- **Final Users tools:**
  - ✓ Km4City mobile applications, http://www.km4city.org/app
  - ✓ Km4City web application, http://www.km4city.org
  - — Open Source Mobile Application, FODD http://www.disit.org/6595
- **Developers tools: http://www.disit.org/km4city**
  - – ServiceMap Server, plus API, http://servicemap.km4city.org
  - ✓ Smart City API , http://www.disit.org/6597
  - ✓ Km4City Ontology, http://www.disit.org/km4city
  - – SPARQL query tool and licenser, http://log.disit.org/sparql_query_frontend/
  - – LOG LOD browser, http://log.disit.org
- **Back Office tools:**
  - – Data Ingestion Manager, DIM, http://www.disit.org/6732
  - – Distributed Smart City Engine, SCE, Scheduler, DISCES http://www.disit.org/6515
  - – RDF Indexer Manager, RIM, http://www.disit.org/6708
  - – RDF store enricher with dbPedia
- **Adopted on projects and real scenarios**
  - – **Sii-Mobility SCN MIUR, http://www.sii-mobility.org**
  - – **RESOLUTE H2020, http://www.resolute-eu.org**
  - – **REPLICATE H2020, http://www.zabala.co.uk/en/projects/replicate**