Km4City: Smart City Model and Tools for City Knowledge Exploitation

Claudio Badii, Pierfrancesco Bellini, Ivan Bruno, Mariano DiClaudio,
Giacomo Martelli, Paolo Nesi, Nadia Rauch
DISIT Lab, Distributed [Systems and internet | Data Intelligence and] Technologies Lab
Dep. of Information Engineering (DINFO), University of Florence, Italy,
Fax: 0039-055-2758570, tel: 0039-3355668674
http://www.disit.dinfo.unifi.it, http://www.disit.org, http://servicemap.disit.org,
http://www.disit.org/km4city

Extended abstract of the presentation

Despite the huge work performed by Public Administrations (PAs) on producing and publishing open data they are not typically semantically interoperable each other and neither with the many private data available in the city. Open data coming from PAs contains information about the city on multiple domains (such as data on the population, risk, votes, admin, energy, cultural heritage, etc.), location of point of interests, POIs, on the territory (including: museums, tourism attractions, commercial, restaurants, shops, hotels, charging points, waste points, bike paths, etc.), major GOV services (tax office, tourist office, hospitals, hosting houses, etc.), ambient data, weather status and forecast, changes in traffic rules for maintenance interventions, etc. Most of these data are produced by different entities, in different moments and time, and not with the same vocabulary and standards. In addition, private data coming from mobility and transport such as those created by Intelligent Transportation Systems, ITS, for bus management, and solutions for managing and controlling parking areas, car and bike sharing, car flow in general, delivering services, accesses on Restricted Traffic Zone, RTZ, people flows, etc. Other aspects are related to risk assessment and resilience of the city with respect to disasters and/or accidents may occur: water flooding, terrorists' attacks, weather, energy, etc.

Both Open and Private Data may include real time data (traffic flow, people flow, position of private and public vehicles, delivering services, events, etc.), railway and train status with respect to the arrival, parking status, ambient sensors (pollution, earthquake), and TV cameras streams for security and flow assessment, status of the emergency services in the hospitals, etc.

PAs and city operators (energy, mobility, etc.) have large difficulties in elaborating and aggregating data to provide new services, even if they could have a strong relevance in improving the citizens' quality of life and services. Therefore, *our cities are not so smart as they could be* by exploiting a semantically interoperable knowledge base on the available data [1]. This condition is also present in highly active cities on open data publications and on smart city services. When data models are analyzed and then processed to become semantically interoperable, they can be used to create an integrated knowledge base that can be feed by corresponding data instances (with static, quasi-static and real time data). Therefore, variability, complexity, variety, and size of these data make the data process of ingestion, aggregation, to enable their exploitation a "Big Data" problem as addressed in [2]. One of the issues is the scarce (or non-existing) semantics interoperability among the data sets. Therefore, common models and tools are mandatory to reach the needed quality level for providing valid services, as numerically proof in [1].

Presently, there are many ontologies for modeling smart city aspects as derived from the EC project READY4SmartCities FP7 CSA, http://smartcity.linkeddata.es/ . On the other hand, only a few of them are available, open, and cover a large range of domains. These features are only endowed by Km4City ontology developed by DISIT lab in Florence [1]. Km4City is a very comprehensive and open ontology for smart cities according to READY4SmartCities EC project covering domains of: weather, cultural heritage, smart sensors, public structures, city parking, services, transportation, events, geographic locations, etc. It has in addition an interesting roadmap in which other domains are going to be added with Sii-Mobility SCN and RESOLUTE H2020 projects

working on it. On the other hand, the availability of a well-defined ontology with a large coverage is not enough to cope with the Smart City modeling and data integration, since the data have to be ingested and made semantic interoperable each other and with the model to enable the set up and delivering of "Smart Services" to citizens. A full set of tools are needed to data gathering, data processing, decision support, providing API, etc.

The proposed presentation is going to expose the integrated solutions around Km4City model which has been set up by the DISIT lab in Florence (http://www.disit.org/6056) and adopted in some EC and national smart city projects (Sii-Mobility Smart City MIUR project, RESOLUTE H2020, Km4City service and tools in place in the Florence Area, with many industrial partners as Thales, Swarco, ECM, etc.). The solution is based on Km4City model that is capable to model a large set of the above data kind and provides support for inference and reasoning, on time and space, on public and private data, on static and real time data. In more details, the solution developed is open and accessible for city providing models and tools for its adoption and exploitation, also enabling the full customization. It includes a set of tools:

- Service map: http://servicemap.disit.org is a tool for PA administrators and for developers. For the PA administrators provide access to several kinds of geospatial queries in Florence and in the whole Tuscany region, taking as a results static and real time data, geo-localized. The ServiceMap is also a tool for developers, which can be used to understand the usage of API to access at the Km4City services http://www.disit.org/6597 . The ServiceMap facility allows the visual creation of queries on the city, and may send to the connected user via email the SPARQL code of the visual queries performed, and in addition also a simple Query ID. The Query ID can be used to pose the query without writing it, from any Mobile and Web applications without the needs of learning complex ontological and SPARQL models;
- Linked Open Graph for browsing LOD RDF Stores model including the Km4City Smart City model in Florence and Tuscany and thus for learning how to formulate SPARQL queries http://LOG.disit.org. See for example a view of Florence http://log.disit.org/service/?graph=0f50fffc5bcfc205de5a19b606b61310
- Demonstrative mobile application exploiting ServiceMap API, also presented at the Florence Open Data Day and accessible as open source via: http://www.disit.org/6595 .
- Km4City ontology model and documentation [1], http://smartcity.linkeddata.es/, document http://www.disit.org/5606 , http://www.disit.org/km4city/schema/
- Parallel and distributed architecture based on ETL, scheduler, HBase, Hadoop, for massive big data ingestion (static, quasi static, and real time data), reconciliation, data enrichment (for connecting Km4City URI to dbPedia, geonames, etc. [3]) and for making decision: [1], slide http://www.disit.org/6566 with thousands of accesses on SlideShare. Several examples are accessible about the ETL transformation for data ingestion, quality improvement, conversion in triples, reconciliation in SILK, [1], etc. This engine is also adopted in other Smart City Projects as SMST national cluster.

[1] P. Bellini, M. Benigni, R. Billero, P. Nesi and N. Rauch, "Km4City Ontology Bulding vs Data Harvesting and Cleaning for Smart-city Services", International Journal of Visual Language and Computing, Elsevier, http://dx.doi.org/10.1016/j.jvlc.2014.10.023

[2] Bellini P., Di Claudio M., Nesi P., Rauch N., "Tassonomy and Review of Big Data Solutions Navigation", Big Data Computing To Be Published 26th July 2013 by Chapman and Hall/CRC

[3] P. Nesi, G. Pantaleo and M. Tenti, "Ge(o)Lo(cator): Geographic Information Extraction from Unstructured Text Data and Web Documents", SMAP 2014, 9th International Workshop on Semantic and Social Media Adaptation and Personalization, November 6-7, 2014, Corfu/Kerkyra, Greece. technically co-sponsored by the IEEE Computational Intelligence Society and technically supported by the IEEE Semantic Web Task Force.

[4] P. Bellini, P. Nesi and N. Rauch, "Smart City data via LOD/LOG Service", LOD2014, Workshop Linked Open Data: where are we?, organized by W3C Italy and CNR, Rome, 2014.