Open Government Data: Fostering Innovation

Ivan Bedini, Feroz Farazi, David Leoni, Juan Pane, Ivan Tankoyeu, Stefano Leucci

I. Bedini, Trento RISE, Via Sommarive 18 Trento (ITA), i.bedini@trentorise.eu, +390461312346

F. Farazi, University of Trento, Department of Information Engineering and Computer Science, Via Sommarive, 9 Trento (ITA), farazi@disi.unitn.it, +390461283938

D. Leoni, Universtity of Trento, Department of Information Engineering and Computer Science, Via Sommarive, 9, Trento (ITA), david.leoni@disi.unitn.it

J. Pane, National University of Asunción, Polytechnic Faculty, Campus Universitario, San Lorenzo (PY), jpane@pol.una.py

I. Tankoyeu, Universtity of Trento, Department of Information Engineering and Computer Science, Via Sommarive, 9, Trento (ITA), tankoyeu@disi.unitn.it

S. Leucci, University of Trento, Department of Information Engineering and Computer Science, Via Sommarive, 9, Trento (ITA), stefano.leucci@unitn.it

Abstract: The provision of public information contributes to the enrichment and enhancement of the data produced by the government as part of its activities, and the transformation of heterogeneous data into information and knowledge. This process of opening changes the operational mode of public administrations, leveraging the data management, encouraging savings and especially in promoting the development of services in subsidiary and collaborative form between public and private entities. The demand for new services also promotes renewed entrepreneurship centred on responding to new social and territorial needs through new technologies. In this sense we speak of Open Data as an enabling infrastructure for the development of innovation and as an instrument to the development and diffusion of Innovation and Communications Technology (ICT) in the public system as well as creating space for innovation for businesses, particularly SMEs, based on the exploitation of information assets of the territory. The Open Data Trentino Project has initiated and fosters the process of opening of public information and develops as a natural consequence of this process of openness, the creation of innovative services for and with the citizens. In this paper we present how our project acts on long-chain, from raw data till reusable meaningful and scalable knowledge base that leads to the production of data reuse through the implementation of services that will enhance and transform the data into information capable of responding to the specific questions of efficiency and innovation.

1. Introduction

The process of opening the data in Public Administrations (PA) as a process of enhancement of public information implies a radical change to the data approach and work inside the PA. This

---

[1] http://disi.unitn.it/~knowdive/

takes time, an improvement of the data management methodology, creation of operational tools and providing a reliable space for sharing. Governments of various countries and administrative divisions thereof worldwide have nowadays been starting to release a huge quantity of datasets in the context of the Open Government Data (OGD) movement (Ubaldi, 2013). The movement became a shared policy after the European Directive no. 2003/98/CE (Public Sector Information Directive), improved by the Directive no. 2013/37/UE. The first Directive has been transposed in the Italian regulatory system with the Legislative Decree no. 36/2006 and the second one still needs to be implemented. The main purpose of the PSI Directive is to enable the so called "data re-use", that means "the use by persons or legal entities of documents held by public sector bodies, for commercial or non-commercial purposes other than the initial purpose within the public task for which the documents were produced" (Directive 2003/98/EC).

In the context of the Autonomous Province of Trento (PAT), a large, diverse and interesting collection of datasets are already published as OGD. New datasets are slowly becoming available and the existing ones are updated whenever needed for the purposes such as correcting mistakes and adding new data horizontally (as instances) or vertically (as properties). The data catalogue is also linked with the website of the Department of Innovation of the PA[2]. This department governs the process of opening new data and the dissemination of the so called data culture. This is an important result to reach both inside and outside of the PAT. People started understanding the value of publishing high-quality data and the power in the reuse of them. Linking data will highly foster the value of sharing data. It will also lead to a new kind of data-centric public bodies that will empower citizens and generate innovative services.

Immense numbers of government datasets could open up new opportunities for application developers and trigger disruptive business models to come (Ferro & Osella, 2013; Manyika e al., 2013; Vickery, 2008). While quantity of such datasets is considered as satisfactory enough, quality (e.g., correctness and vertical completeness) is yet to be improved (Bohm et al., 2012). Moreover, loosely coupled nature of data is posing challenge in developing applications on top of them. Therefore, there is a pressing need to leverage this data before putting them in action. To overcome the issues and fulfil the demand, we made the following contributions in this paper:

    I.   The description of the adopted opening procedure as an integral part of the change management in a public administration.

    II.   The implementation of a methodology for generating entity types (Maltese, 2012) out of published datasets to model data as entities (Giunchiglia, 2012) for facilitating an integrated, combined and extensible representation.

    III.   The implementation of a procedure and the corresponding tool for dealing with unforeseen data (along with known ones) about an entity, taking into account the semantics.

    IV.   Description of our experience in handling Open Big Data for building life style changing unprecedented (in the region) applications.

The paper is structured as follows: in Section 2 we present the process that has been put in place within the PAT and the choices made. In Section 3 we describe the entity type methodology that we have adopted, helping creating integrated entities. Section 4 shows the automatic creation of entities matching dataset schemas to the entity types. Section 5 provides a brief description of open big data approach we are building. In Section 6, we present some applications developed on top of entities and open big data. Section 7 concludes the paper.

---

[2] http://www.innovazione.provincia.tn.it/opendata (in Italian)

2. Open Data Trentino Overview

The Open Data Trentino[3] project was created under the push of the local government of opening their public information as expressed in the guidelines for the reuse of public data official document (DGP, 2012). From the legal point of view, the process started with the adoption of the Provincial Law no. 16/2012 (LP 16/2012), on the promotion of the information society and the digital administration and for the distribution of free software and open data formats.

The process continued by adapting and improving, for the local administration context, the state of the art of existing European good practices in matter of Public Sector Information (PSI) [4]. The result was the drafting of the Open Data Guidelines (Resolution 2858/2012).

The data hunting and publication process followed is a step by step, day by day, federated approach by involving the local authorities since the beginning by asking to every provincial department to open at least one dataset. Although the direct engagement of any department is time consuming it has proved of being successful and brought the desired side effect of creating enough awareness to the whole public administration and to spread the change paradigm. Many aspects are involved in this process: from the data cleaning to the data modelling, from the privacy issues to the intellectual property rights, from the dissemination aspects to the process design. At the same time, we have focused at the creation of the Data as a Culture, by acting in several dissemination and educative actions internally to the institution and with a broader scope at the national and international level.

The Dati Trentino portal is built over CKAN[5], an open source data management system started by the Open Knowledge Foundation and maintained by the CKAN community itself. CKAN is specifically designed to allow programmatic access, finding and retrieval of dataset metadata through web APIs to which we are currently contributing. For instance we developed dedicated clients to access datasets in two different programming languages. The first one, *Ckan Api client*[6], is an open source Python based library used to add automatically harvested datasets into a standard CKAN platform version 2.2. The second one, called Jackan[7], is an open source Java based lightweight library with built-in support for provenance tracking to easily access the catalogue data from Java.

As of October 2014, the government of the PAT published about 860 datasets in a catalogue made available under the link http://dati.trentino.it more than a year ago, with an open license for ensuring free and unlimited use and reuse of data, representing the engagement of about 60 provincial departments. The catalogue is clustered into 13 broad categories, each consisting of a number of datasets, which are represented as one or more resources that are easily accessible and downloadable often in CSV and/or JSON format, and occasionally in XML. Each dataset can be mapped to DCAT[8], a vocabulary designed for describing catalogues and datasets thereof for increasing interoperability. Just to cite a few of the published resources of high importance, there are provincial budget and cadastre.

As shown by the result published by Dati.Gov.It[9], as we write, the Province of Trento is the most performing catalogue in Italy in terms of number of published dataset, which confirm that the

---

[3] http://www.provincia.tn.it/progetto_open_data (in Italian)

[4] "Libro bianco per il riutilizzo dell'informazione del settore pubblico", EVPSI Project (2012), http://www.evpsi.org/evpsifiles/bianco_beta.pdf;

[5] CKAN Platform, http://ckan.org

[6] Ckan API Client, git.io/ckan-api-client

[7] Jackan, https://github.com/opendatatrentino/jackan

[8] http://www.w3.org/TR/vocab-dcat/

[9] http://www.dati.gov.it/content/infografica

approach of involving directly the people working in the PA with the creation of a concrete data culture is a successful approach.

3. Entity Type Generation

This Section depicts an entity centric approach for modelling OGD. It describes the generation of entity types according to the data published in government data catalogues. An entity type (also called as eType) is the class of an entity that has the right amount of attributes and relations for forming the foundation of creating entities with their non-trivial details intended for an application (Giunchiglia, 2012; Maltese, 2012; Farazi, 2008). Some examples of entity types are person, location, organization and facility. An entity is a real world physical or abstract or digital thing that can be referred with a name (Giunchiglia, 2012). For example, Dante Alighieri (person), Trento (location), University of Trento (organization) and Trento Railway Station (facility) are entities.

In the context of this work, we have been dealing with the catalogue published by the PAT. As shown in Figure 1, we divided the entity type development in three macro-phases - datasets survey, attributes survey and producing entity types - each of them with different macro-steps.
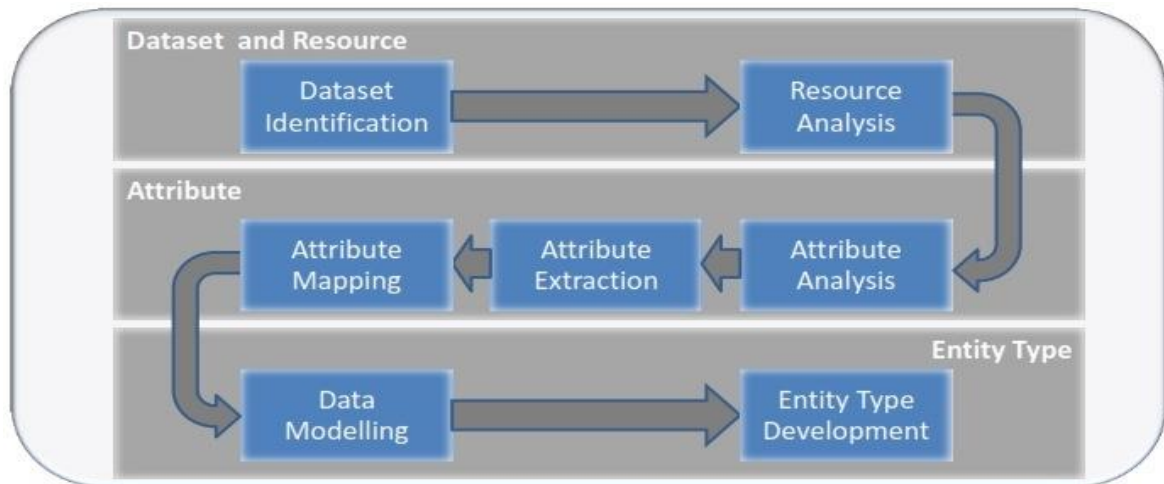

Figure 1: Modelling Open Government Data as Entities

Modelling starts with the *dataset identification* step which relies on the scenario or task at hand. For example, our scenario involves points of interest which include datasets representing among others refreshment facility (such as restaurant, pizzeria, bar, etc.), recreational facility (such as ski lift, sports ground, museum, etc.) and transportation facility (such as bus stop, railway station, cable car stop, etc.). In the *resource analysis* step a rigorous study on the structure and content of the corresponding file(s) is performed to understand the relevant resources for singling them out.

The *attribute analysis* proceeds through examining the attributes of the already selected resources. With attributes, it means column headers in the CSV files, properties of objects in the JSON files and sub-tags under repeated object tags in XML. Attribute values are also analyzed in terms of availability (e.g., always, frequently, sometimes and never present) and quality (e.g., complete data with no or occasional syntax error, partial data with or without error and relational data with or without disambiguated reference) of the data.

In the *attribute extraction* step, we differentiate between the kinds of attributes according to the data encoded in them. Some attributes are used for encoding data and some others for managing data (e.g., identifiers internal to the resource used as primary keys). In this step, we also merge and split attributes, if necessary, according to the data. For example, nameEn (name in English) and nameIt (name in Italian) can be merged into name, while opening hours can be split into opening time and closing time.

The *attribute mapping* step incorporates disambiguation of the attributes proceeds through linking them to the right concepts in the knowledge base (KB). We identify two kinds of mapping, *direct mapping* and *indirect mapping*. Finding the name of the attribute attached to the intended concept in the KB and linking to it is called direct mapping. Not always a direct mapping is present for an attribute. Finding the right sense of a term through its synonyms is called *indirect mapping*. Note that synonyms can be suggested by user or retrieved from a KB.

The *data modelling* step leads to understanding an entity type from the attributes extracted in the previous step and finding it (if exists) or a suitable parent (if created newly) in the already existing *entity type lattice*, a lightweight ontology (see (Giunchiglia, 2009)) formed with the concepts of the entity types. In the *entity type development* step, we produce a specification of an entity type defining all possible attributes, their data types (e.g. string, float) and meta attributes such as permanence (e.g. temporary, permanent), presence (e.g. mandatory, optional) and category (e.g. temporal, physical).

While producing entities of a given entity type, mandatory and optional attributes are filled in with data, which are semantified and disambiguated wherever applicable. In fact data for an entity can come from multiple resources. Through semantification, we facilitate the integration of loosely coupled data. In the case of unavailability of a mandatory attribute in the possible resources, we signal it to the data provider as *pro-sumers*, see (Charalabidis, 2014) and do not allow the creation of the corresponding entity unless all necessary data are present. This is how we can improve the vertical completeness.

4. Open Data Rise

Open Data Rise[10] (ODR) is an open source web application for data curation of OGD. It allows to easily fix errors in source data and also to enrich names by linking them to their precise meaning in a process called *semantification*. The framework employs the entity-centric model (Giunchiglia, 2012) and extends OpenRefine [11], an open source tool for cleansing messy data. The semantification pipeline implemented in ODR is the result of the collaborative effort research work developed in the KnowDive team of the University of Trento and Trento Rise. In detail, it consists of the six steps depicted in Figure 2:
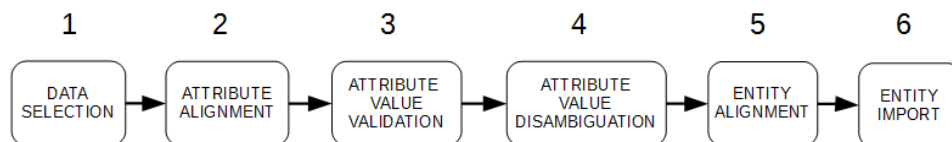

Figure 2: ODR Semantification Pipeline

During selection step (see screenshot in Figure 3) the framework allows the user to select a dataset from any CKAN compliant repository. For catalogs previously analyzed by Ckanalyze[12], a component we developed, repository statistics are also shown.

On the *attribute alignment* step the framework matches columns in the source data set to a predefined set of entity types. A simplified example is presented in Figure 4. This step includes finding the appropriate entity type for a given data set and mapping its attributes with column headers in the dataset.

---

[10] https://github.com/opendatatrentino/OpenDataRise

[11] http://openrefine.org
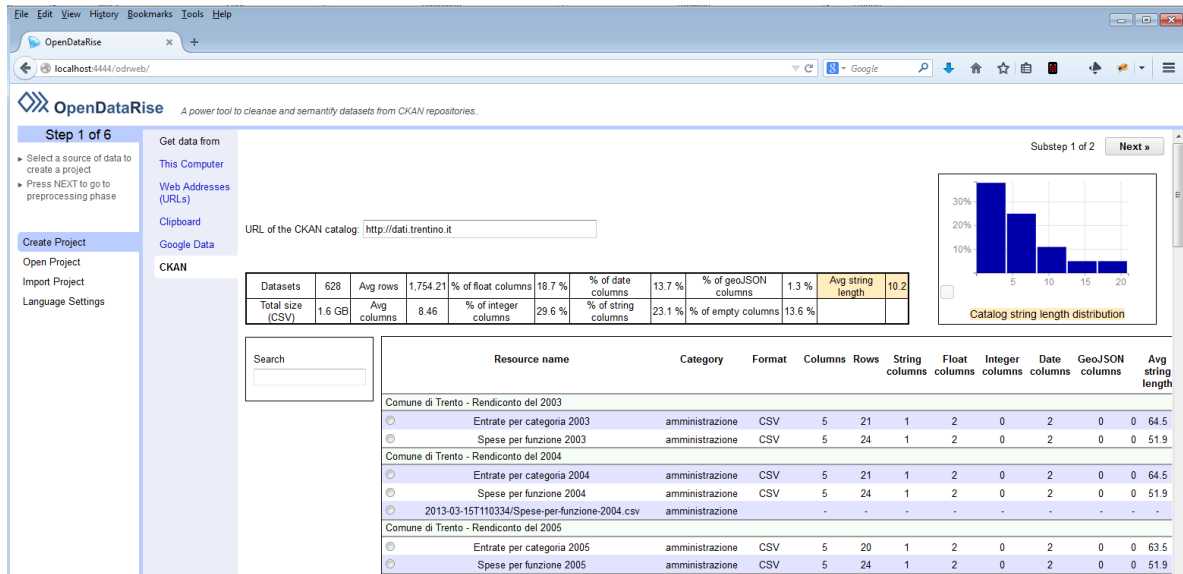
[12] https://github.com/opendatatrentino/CKANalyze

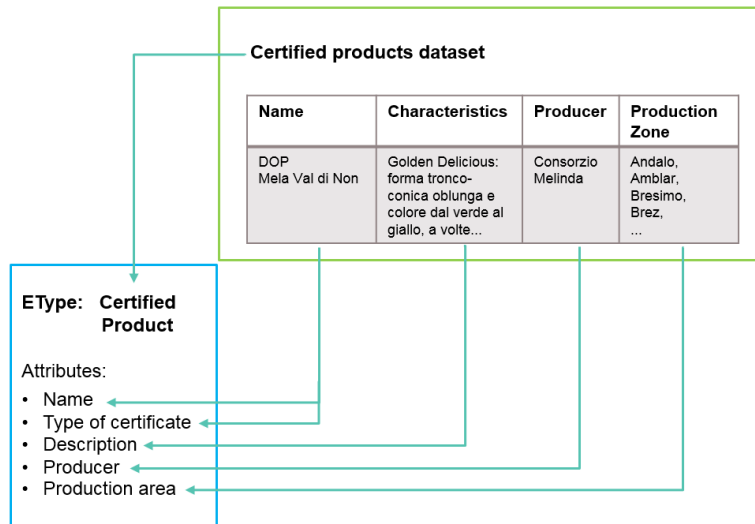Figure 3: Dataset Selection step of ODR pipeline

Figure 4: An example of a simplified schema matching

Once dataset schema has been determined, during *attribute value validation* step the user can adapt the dataset to the schema, exploiting *OpenRefine* data cleansing capabilities.

Successive *attribute value disambiguation* step employs Natural Language Processing techniques for enriching dataset content by linking names to known entities (such as *Dante Alighieri*, *Florence)* and words to concepts (such as *male*, *city).* In Figure 5 we see a screenshot of long text that has been automatically enriched. OpenDataRise will show in red elements that still require manual intervention from the user.

Within *entity alignment* step the framework considers rows in the dataset as entities, i.e. real instances. The goal of this step is to schedule changes to entity storage to be committed in the next step. Such changes can be either update of existing matching entities or creation of new entities with values from the source dataset.

Figure 5: Example of semantic enriching in ODR

In the last *entity import* step, the user can indicate the license of entities to import and other metadata to publish to CKAN. Updates and insertions are then committed to entity storage and a new semantified resource is published on CKAN. The resource will contain the provided metadata and a reference to imported entities in the entity storage.

5. Open Big Data

Nowadays big volumes of data are processed at an increasing rate, creating additional hidden information. This information represents a central aspect in the definition of Big Data that can be defined as Value. In fact, according to some industry analysts, dealing with Big Data means facing the following aspects: Volume (huge amount of data generated or data intensity that must be ingested, analyzed, and managed to make decisions based on complete data analysis), Velocity (the speed at which data must be processed), Variety (the different types and sources of data that must be analyzed and the complexity of each and the whole), Variability (intended as the inherent "fuzziness" of data, in terms of its meaning or context) and indeed, last but not least, Value. Since the public sector is increasing the quantity of data available to the public through many open data initiatives, we expect that in the near future also collected data by these initiatives will thrive by the adoption of BigData technologies to gather useful information from published data.

As part of OGD initiative of the PAT, we then focus on the problem of data explosion and the consequent need of having fast and scalable solutions for storage and analysis. We estimate the trend for growth will be up to hundred times per year, easily reaching the order of TB of data in few years from now. For instance the Trentino portal already have sensors based datasets, such as weather, traffic sensors, real time energy consumption and few others that, they alone already provide few GB of data per day if collected. Due to this very nature, they pose challenge in using traditional relational database management systems to handle them and at the same time appear as a problem to be dealt with the Big Data technologies such as Apache Hadoop[13], Hive[14], Pig[15] and NoSQL databases. A kind of big data generated by various actors including government

---

[13] http://hadoop.apache.org

[14] https://hive.apache.org

[15] https://pig.apache.org

agencies and companies and published as open data is called *open big data*. The interest here is focused on producing a Big Data platform directly integrated with the Open Data portal and able to provide useful analytics, historical analysis in reasonable time.

6.   Applications

Entities generated from open data as well as open big data exploitation resulted to the development of applications which often appear as innovations to the citizens. This is because they offer services that are either novel or come up with better results in comparison with the contemporary ones. Moreover, open data propel innovations that help devising novel applications and services (Chan, 2013).

In line with this consideration, as shown in Figure 6 and Figure 7, we developed an application running on top of entities helps finding points of interest including restaurant, pizzeria and bar with opening hours and bus stop, cable car stop and railway station with timetable and ski lift, ski rental and ski school with timetable.
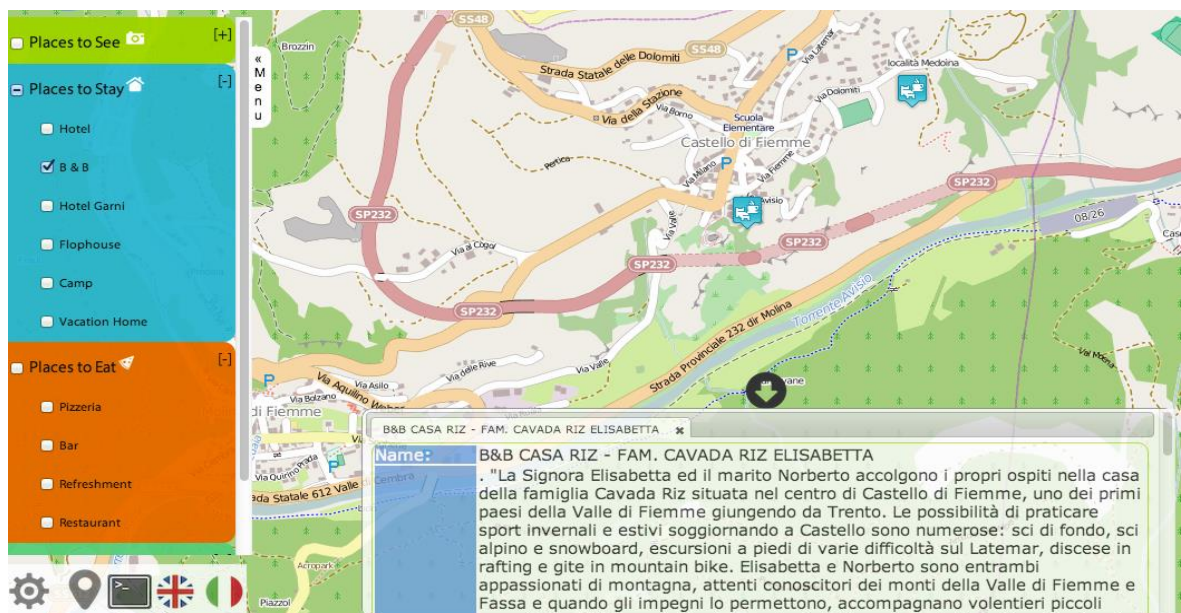
Figure 6: Faceted navigation for finding points of interest

Figure 6 sketches faceted navigation. For example, when user is in a ski lift location, selecting a point of interest category or a subcategory shows the corresponding results on the map. Figure 7 shows semantic navigation. We call it semantic navigation as it exploits semantic relations like part-of while exploration based search proceeds. To provide an example, when user is in a museum, it can show, within a radius of, e.g. 500 m or 1 km, all kinds of points of interest also located in the proximity of the museum itself. Finally, the bottom widget shows attributes and general information about the selected entity.

Figure 7: Semantic navigation for finding points of interest

OpenBus[16] shown in Figure 8 is an application developed using open big data. This application approximates real time (geographical) positions of the city buses in all the routes in Trento, based on published bus schedule and 2 minutes temporal snapshots of their registered geo-location. This last dataset, as we write, is not publicly available yet because of technical reasons.
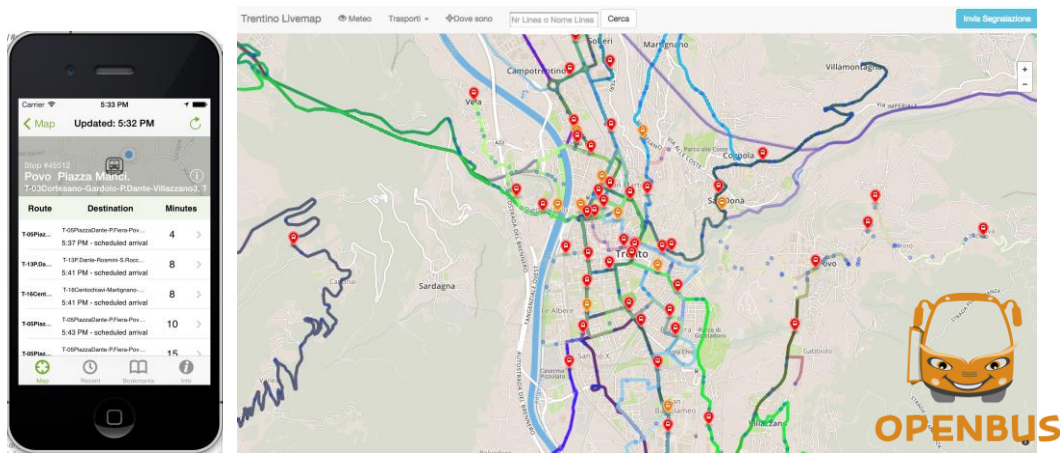
Figure 8: Open bus mobile app and livemap view of the real time simulation

It employs big data analytics tools and techniques in collecting, storing and analysing data generated by on-board sensors to predict delays on the basis of the current position of the buses. The API providing access to the real-time prediction and the associated mobile app are based on the open source OneBusAway platform[17].

---

[16] http://livemap-bigdata.trentorise.eu (prototypal version)

[17] http://onebusaway.org

7. Conclusion

In this paper we depicted the approach we implemented in the publication of Open Data in the Province of Trento. The result that we have already obtained is promising and we show that the complement of the *data culture* along with a direct technical and legal support to PA employees allow a faster diffusion and sustainability in the process of opening data within the public administration. Furthermore, we proposed an approach for generating entity leveraging open government data grounded on entity based infrastructure that by design facilitates consistency in data representation and management. As such enabling the re-use of public sector information. We also address the forthcoming data explosion issue by investigating and integrating since the beginning BigData Technologies. The entity centric data representation and the infrastructure as a whole can be considered as an input to the W3C Data on the Web Best Practices Working Group to provide guidance to data publishers. Our future work involves the creation of a broader data ecosystem. Communities, universities and other actors will be able to link their data using Open Data Rise. Moreover, we will implement crowdsourcing techniques for improving the quality of the existing data.

References

Bohm, C., Freitag, M., Heise, A., Lehmann, C., Mascher, A., Naumann, F., Ercegovac, V., Hernandez, M. A., Haase, P., Schmidt, M. (2012). Gov-WILD: integrating open government data for transparency. In Proceedings of WWW 2012, pages 321–324.

Chan, C.M.L. (2013). From Open Data to Open Innovation Strategies: Creating E-Services Using Open Government Data. Hawaii IEEE International Conference on System Sciences.

Charalabidis, Y., Loukis, E., and Alexopoulos, C. (2014). Evaluating Second Generation Open Government Data Infrastructures Using Value Models. Hawaii IEEE International Conference on System Sciences.

Delibera Giunta Provinciale 2858/2012. (DGP, 2012)
http://www.innovazione.provincia.tn.it/binary/pat_innovazione/notizie/Lineeguida_21dicembre_def.1356705195.pdf

Directive 2003/98/EC (2003). Directive 2003/98/EC of 17 November 2003 on the re-use of public sector information. OJ L, 345, 90. European Parliament and Council. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:345:0090:0096:EN:PDF

Farazi, F. (2010). Faceted Lightweight Ontologies: A Formalization and Some Experiments. PhD Thesis, supervised by Prof. Fausto Giunchiglia, International Doctoral School in ICT, University of Trento, Italy.

Ferro E., Osella M. (2013). Eight Business Model Archetypes for PSI Re-Use, Open Data on the Web Workshop, Google Campus, London.

Giunchiglia, F., Zaihrayeu, I. (2009). Lightweight Ontologies. Encyclopedia of Database Systems.

Giunchiglia, F., Maltese, V., Dutta, B. (2012) Domains and context: first steps towards managing diversity in knowledge. J Web Semantics (special issue on Reasoning with Context in the Semantic Web).

LP 16/2012. Provisions for the promotion of the information society and the digital administration and for the distribution of free software and open data formats.
http://www.innovazione.provincia.tn.it/documentazione/in_trentino/pagina185.html

Maltese, V., Kharkevich, U., Radu, A.L., Semertzidis, T., Lazaridis, M., Ratailidis, D., Drosou, A., and Georgescu, M. (2012). Deliverable D4.1 Space and Time Entity Repository, CUbRIK Project.

Manyika, J., Chui M., Groves, P., Farrell, D., Van Kuiken, S., Doshi, E. A. (2013). Open data: Unlocking innovation and performance with liquid information. McKinsey Global Institute Report.

Resolution of the Provincial Government no. 2858/2012 - Guidelines for the use and dissemination of public information:
http://www.innovazione.provincia.tn.it/notizie/pagina482.html

Ubaldi, B. (2013). Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives, OECD Working Papers on Public Governance, No. 22, OECD Publishing.

Vickery, G. (2008). Review of recent studies on PSI re-use and related market developments. Information Economics Paris. Final Version.

About the Authors

Ivan Bedini is Team Leader for the Open Big Data collaborative project at Trento RISE. As team leader Ivan is responsible of the development and maintenance of innovative solutions for the collection, integration and analysis of Open Big Data linked with the Open Data portal of Trentino. Previously Ivan has been Member of Technical Staff in Bell Labs Research Ireland, where he has been active contributor of the Semantic Data Access (SDA), and for the BigData analytics project. Prior to joining Bell Labs in 2010, he was researcher at Orange Labs France in the Enterprise Applications and the Trust & Secure Transactions research departments. During ten years at Orange Labs he held different roles as developer, technical expert and project manager, and specialized in the domains of Business Process Management, e-business, knowledge engineering and Semantic Technologies. He has also four years of contribution to standardization bodies as member of the UN/CEFACT Information Content Management Group and member of OASIS ebXML Registry/Repository Committee. He received his PhD from the University of Versailles, France, in 2009. He is expert in Information Extraction, Data Integration, Semantic Technologies, BigData technologies and distributed computing.

Feroz Farazi is a postdoctoral research fellow at the University of Trento (UNITN) in Italy. Currently he is working in the Open Data Project (http://www.provincia.tn.it/progetto_open_data/) funded by the Autonomous Province of Trento. He obtained a PhD degree in Computer Science from UNITN in 2010. His research interests include Knowledge Management, Ontologies, the Semantic Web and Data Integration. He has authored several publications including a book entitled "Faceted Lightweight Ontologies: A Formalization and Some Experiments". He has collaborated in the development of GeoWordNet knowledge base at UNITN from 2009 to 2010. He has also collaborated in the Semantic Geo-Catalogue project funded by the Autonomous Province of Trento. He served as an assistant professor from 2003 to 2006 in the department of Computer Science and Engineering (CSE) at the University of Chittagong in Bangladesh. He also served as a lecturer in the department of CSE at Shahjalal University of Science and Technology, Bangladesh from 2001 to 2003.

David Leoni is a Software Engineer in the KnowDive research group at DISI department, University of Trento. He is the core developer of OpenDataRise, an open source tool for data curation. With a background in privacy preserving data mining, his interests currently span semantic web, teaching creative computing and visual programming languages.

Juan Pane is a Lecturer and researcher for Open Data at the National University of Asunción, Paraguay. He is also a researcher of the Latin American Open Data Initiative (ILDA) and a consultant for the Democracy and Governance Program for the Paraguayan Government funded by USAID and implemented by CEAMSO, Paraguay. He received his PhD in Computer Science from the University of Trento in March 2012, where he later stayed as a post-doc research fellow (2012-2013). He led the development of Multilingual Open Source tools for Open Data Integration such as OpenDataRise and CKAnalize for semi-automatically extracting semantics from open datasets for the Open Data Trentino (Italy) project (dati.trentino.it). He contributed to the EU funded OpenKnowledge project in the area of semantic matching and to the Insemtives project in the areas of semantic annotations models and semantic disambiguation in Folksonomies, all these aimed at data integration. His current research and development activities focus in ICT based Innovations for Open Data and Open Government initiatives by applying Semantic technologies and on designing automatic tools for Software Quality Assurance. He has authored several scientific publications and journal articles.

Ivan Tankoyeu is a postdoctoral associate at the University of Trento. Currently he is working in the Open Data Project funded by the Autonomous Province of Trento. He obtained his Ph.D. from University of Trento, Italy in 2013, having awarded a M.Sc. in Computer Science from the Belorussian State University in 2007. His main research interests include data and knowledge management, event based media indexing, event mining and exploitation from spatio-temporal data, context based analysis. He has authored peer-reviewed scientific publications in the top conferences of the Multimedia research field (ACM MM/ICMR). He has served as the technical program committee on ACM MM Workshop HuEvent'14 and ACM ICMR'11. Previously he has worked on the EU FP7 Project CubRIK and EU FP7 Project Glocal.

Stefano Leucci is a Ph.D. Candidate at the ICT Doctoral School of the University of Trento, where he is member of the KnowDive research group (http://disi.unitn.it/~knowdive). Due to his background in Law, he focuses his work on the inter-disciplinary aspects of data management and low. He is currently a Fellow of the Nexa Center for Internet and Society of the Polytechnic of Turin and he works as legal advisor in the Open and Big Data Project of the Autonomous Province of Trento and TrentoRise.