

Big Data and Open Data for a Smart City

Andrea Molinari
Department of Industrial Engineering
University of Trento, Trento(Italy)
andrea.molinari@unitn.it

Vincenzo Maltese
Department of Information Engineering and
Computer Science, University of Trento,
Trento(Italy)
vincenzo.maltese@unitn.it

Lorenzino Vaccari
European Commission
DG Joint Research Centre
Digital Earth and Reference Data Unit
Ispra(Italy)
lorenzino.vaccari@jrc.ec.europa.eu

Andrea Almi
Security & Information Systems Division
Selex ES – Finmeccanica Group
Genova (Italy)
Andrea.Alm@selex-es.com

Eleonora Bassi
Trento Rise
Trento(Italy)
eleonora.bassi@trentorise.eu

Abstract— In this paper, we present a general overview of the perspectives and issues of Big Data and Open Data in a Smart City, with specific implications for the municipality of Trento, Italy. We start by presenting the current state of the art of Big Data and Open Data in Trento and continue by proposing a line of development for these two topics that could positively impact the everyday life of its citizens. We will place particular emphasis on the results that emerged from the discussion we had during the working group meeting that the municipality organized on this topic. The challenges posed by four enablers of Big Data and Open Data projects and initiatives are described: cultural enablers, organizational enablers, governance enablers and technological enablers.

Keywords: big data, open data, smart city

I. INTRODUCTION

Currently, citizens of several cities worldwide live under a so-called “data deluge.” Data generated by people and sensors and later processed and stored by computers have reached incredible levels and not only are assuming a crucial role inside the information systems of organizations but also represent an incredible source of opportunities for even the individual citizen. As highlighted by prestigious economic journals, the data deluge is completely transforming the way we do business, the way government manages the “res publica,” the way research and science are looking at the future, and especially the way our everyday life will be related to computers [1], [8].

Data have gone from relatively small-scale databases relegated inside information systems to superabundant data made available from many different data sources, with this heterogeneity mainly constituted by data coming from structured institutional and enterprise datasets, unstructured social media, sensors and people themselves. All these new perspectives of data management bring new, relevant benefits but also significant headaches, especially for public administrators and IT personnel who are on the forefront of this data deluge and who have to provide professional, prompt and efficient responses to these challenges.

The consequences of this phenomenon involve many stakeholders, such as citizens, politicians, IT personnel, decision makers, and companies, and so the entire society is

involved. For IT specialists, a new buzzword has been created to highlight the new discipline (and respective issues): Big Data (see, for instance, [2],[3],[4],[5]). As a direct consequence of the data deluge, this word summarizes many different aspects, but one in particular is a challenge for the IT sector: in Big Data settings, it becomes difficult to manage data using traditional applications. Therefore, we need new tools and techniques to manage Information Systems. This is clearly an issue for researchers and IT producers, but it is an issue mostly for end users and Big Data producers. Consider the following simple facts that stress this condition:

- a Boeing 737, during an ordinary intercontinental flight, produces approximately 200 terabytes of data, mainly through sensors;
- the current volume of data accumulated from many different data sources (from weather sensors to log files, from information systems to e-commerce, from medical applications to telephone calls, from video uploaded on the Internet to air traffic control records) is approximately 2.5 quintillion bytes [3].
- 90% of the data in the world today were produced within the past two years.

Big Data and the related area of Big Data Analytics have such large applications and datasets to manage (from terabytes to exabytes) with a very high level of variety and complexity (from sensors to social media data) that they require new advanced data storage and management tools and new analysis and data visualization technologies. Dealing with these issues is not just a technological problem. It requires new interdisciplinary and multidisciplinary competences, ranging from economics to management sciences and law studies. This has given rise to the data scientist profile, which has been defined as the sexiest job of the 21st century [11].

Big Data are often characterized by V’s, such as Volume, Velocity and Variety (this is what people in the Semantic Web also call *heterogeneity* or *diversity* [12]). Other authors have noted that additional Vs exist concerning IT, business and data scientists: Veracity, Validity and Volatility. Even if

disputed, there is no doubt about the need for users to have clear ideas when facing these problems. Cities are, of course, great producers of Big Data, and they have to constantly address data collection, integration and interpretation. Decision makers realize that appropriate management of this data deluge can constitute an incredibly powerful tool for making better decisions and increasing the appeal and effectiveness of political actions.

Moreover, some of this data must be given back to the original owners, the citizens, because that data have a public origin and public destination; the Open (Government) Data is another idea of managing data and is here to stay (see, for instance, [13]). Open Data is particularly engaging to many public administrations of the world, under the strong pressure of stakeholders who consider (open) data a form of democracy and an opportunity for business. In this sense, any “smart” city has to address these two sides of the same coin: on one side, the enormous amount of data that can help all stakeholders to provide and accept better decision processes; on the other side, the need for these data to be available to the same stakeholders (citizens, mainly) in order to provide added value to these data sources through this openness.

This paper will analyze the position of Trento as a smart city with respect to the many aspects of Big Data and Open Data. The paper is organized as follows. Section II describes the state of the art concerning Big Data and Open Data in Trento. Section III addresses the current challenges facing the four pillars we have identified during our working group. Section IV concludes the paper, summarizing our findings and our recommendations for the municipality of Trento.

II. BIG DATA AND OPEN DATA IN TRENTO

The process of integrating Big Data and Open Data with the Information systems of Trento has taken off only very recently. Several projects have started in this direction, especially in the area of Open Data (rather than Big Data). This is because, at the moment, the data sources are mainly datasets coming from traditional database applications, and their size is still manageable. Some initiatives could soon lead Trento to jump into Big Data configuration, specifically sensors acquiring environmental and traffic conditions as well as fiscal data [17] collected in order to support tax evasion detection.

The *Information System Service* of the City of Trento today manages most of the databases used for internal applications. Such applications have been designed to be functional to the city structures themselves. Data from databases is periodically extracted and analyzed by the *Statistics Office* and the *Office of Planning and Management Control* to produce their yearly reports.

Although this process has thus far been proven to serve its purpose, it has two main drawbacks. The first pertains to the frequency of the release of such data (only yearly), which implies that actions are delayed. As shown by the census of internal processes of the Administration, the second pertains to the dependency (especially from a technical point of view) on specific Structures of the City managing the databases and on specific people holding the know-how (a subset of

their employees). This particularly pertains to the *Information Systems Service* and depends on the *Works of Primary Urbanization Service* (e.g., for data concerning traffic lights). These two factors may clearly lead to a loss of efficiency.

The objective is therefore to restructure the internal operational processes currently in place for data management to share knowledge, to empower the open data culture among the administrative structures, and to move from internal management to shared information management. We expect that this will increase overall efficiency and enable new services and new processes.

Table 1 provides a snapshot of the current situation in the municipality of Trento pertaining to the number of published datasets, the modalities by which it makes them available, the areas covered and the types of licenses employed to distribute them.

Currently, the publication of datasets as Open Data is an ad-hoc and manual process that is costly, is difficult to handle and does not guarantee an automatic update of exposed data. A virtuous exception is represented by map datasets, exposed as a service via links updated directly by the cartographic systems, with no need for manual operations.

An alternative solution is under experimentation within the project ComunWEB [14] of the Consortium of the Municipalities of Trentino, to which the City of Trento is contributing. Through this approach, the automatic update of the data content can be guaranteed, at least for the structured information published on the website of the Municipality of Trento (or, more precisely, the information coming from the underlying Content Management System). As soon as this system is put in place, information of interest posted on the website (e.g., events, determinations, and offices) by various stakeholders will be automatically available in real-time as Open Data. In fact, it will provide an API to automatically expose structured information as Open Data in JSON format via a RESTful service.

Table 1 – Published datasets from the municipality of Trento

| Indicator | Status |
|---------------------------------|--|
| Datasets published as Open Data | 63 |
| Open Data Policies | Only ad-hoc isolated actions have been taken thus far |
| Open Data website | No dedicated website has been put in place for the municipality. Open Data are published on a section of its institutional website and in the catalogue of the Province of Trento (dati.trentino.it) |
| Main categories | Cartography, Statistics, Environment, Financial data |
| License | CC-0 / CC-BY |

In addition, in collaboration with the Province of Trento and within the Open Data Trentino project [15], the Municipality of Trento has declared its intention to set up an automatic harvester to automatically publish information about its Open Data datasets on the official Open Data catalogue of the Trentino region (dati.trentino.it) that collects data from the entire territory. The first steps towards Open Data adoption in the territory date back to 2012 with deliberation n. 195/2012 of the province, which formally decided to adopt the CC0 (public domain) license to release 161 of its geographical core datasets [10]. This was one of the important factors that contributed to the innovation of public services. We can mention, for instance, the semantic geo-catalogue (also cited in the Italian Digital Agenda) of the province that facilitates the effective search for maps using their metadata [9].

The direction towards a larger adoption of the Open Data spirit has been identified. The City of Trento plans to enhance its smart dimension, proposing the city as a catalyst of activities focused on citizens' needs and as a territorial laboratory for actions in technological and social innovation aiming for the improvement of the overall quality of life of its population. The strategy to reach this goal is to base the realization of services on the potential offered by Open Data as the primary enabling infrastructure.

This includes the design and activation of a flow of processes in which data are exploited from their creation, ensuring a high level of quality and trustworthiness. In order to allow easier data integration between data and content generated by other local entities and individuals operating in the same area, the process will be governed at cultural, organizational, and technological levels.

We envision that the new way of managing public data will lead to greater simplification and better efficiency within the administration itself, while defining a systematical improvement in the generation and use of data in connection with the services that the city will decide to activate, thus reducing costs and promoting the sustainability of the process.

Therefore, the City of Trento will adopt an organizational and technological data infrastructure that is enhanced and made functional by technologies for Big Data Integration and Data Analytics, which will enable different types of services, and designed flexibly and incrementally to meet the following objectives:

- To improve the quality of data processing inside the municipality itself so that data are more easily accessible and usable for the pursuit of institutional public tasks of different structures;
- To provide citizens with services based on municipal and territorial data that will improve their quality of life and facilitate access to information;
- To provide high-quality and easy-to-reuse data to citizens and companies.

III. BIG DATA AND OPEN DATA CHALLENGES: THE ENABLERS

In the definition of the project vision and plan, we identified four different enablers that can allow a smart city to take advantage of Big Data and Open Data benefits: cultural, organizational, governance and technological enablers. In the next sections, we will discuss various details pertaining to these four enablers.

3.1 Cultural Enablers

Open Data is an opportunity to change. It is clear that the administration of a city that has the ambition of becoming "smart" should overcome the traditional culture of being the privileged "owner" of public data. Data must instead be seen as a common good for every citizen and should be therefore presented without any resistance from stakeholders; this is what people sometimes define as "frictionless data."

Another complementary core element is represented by the virtuous circle that should be created across the public administration, its citizens, enterprises and back. A public administration should valorize believers, active citizenship and civic hackers. It should promote education and the involvement of both producers and consumers of Big Data and Open Data.

Therefore, Open Data can be seen both as an opportunity and as an instrument that the public administration can use to foster change. In fact, making data publicly available to the entire society implies, on the one hand, the adoption of principles of subsidiarity and transparency as a pervasive element of the public administration and, on the other hand, a change in the way in which it operates its public services.

The ultimate goal is to make public information available to the entire society as a means to create social innovation, regional development, and value-added services and applications; foster openness and transparency; and stimulate discussions and participation.

Fundamental in this respect are the activities that the public administration should consider putting in place for the engagement of the various participants in the society. Tim Davies came up with a 5-star classification of engagement aspects for Open Data [16]. Engagement with Open Data should meet the following:

★ **Be demand driven:** choices about the released data should be based on community needs and demands. This also includes answering the data requests of users.

★★ **Consider putting data into context:** clear metadata must be provided, including a description of the data quality, the frequency of updates, the data formats, the lineage and links to any additional document that describes how to address the data (manuals, publications, reports, deliverables, analysis of the data, etc.).

★★★ **Support conversations pertaining to the data:** users should be able to comment on datasets or create structured conversations around data with other users. Data publishers,

metadata catalogue maintainers and individual “data owners” should be able to be contacted directly if needed.

★★★★ **Build capacity, skills and networks:** links to tools for people who would like to work with the published datasets should be provided, along with a “How to Guidance” and Open Data analysis tools, so that people can increase their capacity and skills for interpreting and using data in the manner they need to. Additionally, skill-building sessions on using data in particular ways and sponsoring or engaging with capacity building to help the community work with open data should be put in place.

★★★★★ **Stimulate collaboration on data as a common resource:** feedback loops to improve the quality of datasets and collaborations with communities to create new data resources, lend support to build and sustain useful tools and services that work with the published data, and work with other organizations to connect published data sources should be adopted.

Implementing this is not expected to be easy. In fact, it requires a revision of the internal processes of the public administration that should be able to identify the data suppliers and the data sources, select the most valuable datasets, document them with high quality metadata and finally publish them as Open Data.

The creation of such an integrated publication workflow goes far beyond a mere technological solution. Orthogonal to these fundamental activities, a deeper “cultural change” inside and outside the institutions must also be adopted. It includes not only the pure publication (which is, by the way, the first part of the Open Data management lifecycle) but also the evaluation, adoption, implementation and integration of tools able to integrate, analyze and visualize Open Data as well as activities aimed at the training of people so that they can use such tools and interpret the data. These activities aim to facilitate the publication, management and sustainability of the Open Data publication and reuse process and include (i) the definition and implementation of a plan for the creation and maintenance of a community around the theme of the Open Data that is able to sustain itself in the long term; (ii) the integration of activities with the national and international Open Data projects and initiatives; (iii) the support to facilitate the consumption of the data by the different stakeholders (e.g., by organizing thematic schools of data, data journalism seminars, courses, and meetings with local enterprises and startups); (iv) the creation and sustenance of specific skills and technologies needed to support the Open Data community; and (v) the dissemination through participation in various local, national and international events.

In this respect, the municipality of Trento has actually initiated some relevant activities. Members from all its departments are invited twice a year to attend a course on Open Data principles and usefulness. Particular emphasis is given to the opportunities and legal aspects to be considered

with regard to making the institutional datasets openly available. Moreover, a number of meetings with each single office are organized when particular issues and problems related to datasets in a specific domain (e.g., tourism locations, bicycle paths, traffic jams, etc.) have to be solved.

Feedback from the various stakeholders is constantly collected and monitored, and for each stakeholder, customized answers and information are provided. Here, the main issue is to explain to the final users (and, in particular, to citizens) the importance and the meaning of the term “Open Data,” which is often not easily understandable by them and by the data providers themselves. Usually, the following issues are discussed:

- the difference between data and documents;
- the difference between a human readable document (e.g., published in a PDF format) and a machine readable dataset (e.g., published in a CSV format); both can contain data, but only the latter can be used directly by software;
- the meaning and the importance of the data format and in particular of an Open Data format (e.g., the CSV format);
- how to publish (providers) and how to find/use (final users) a datasets in the Open Data portal [6].

Informing citizens about new services and apps is also a priority for the municipality of Trento, which usually publishes news about Open Data on its institutional portal. This has been done already for the GIS domain via a dedicated webpage. The webpage currently exposes two mobile apps related to the topographic map of Trento and the map of the cemeteries (currently in preparation). A public announcement was made to inform the citizens of the availability and characteristics of these two apps. This could be extended to other domains.

3.2 Organizational Enablers

From an organizational point of view, the discussion of smart cities can be focused on three main aspects: (a) the reengineering of the processes, (b) the incentive mechanisms to be put in place and (c) the necessity of involving the territory.

Citizen-centric BPM Reengineering. Making data “open” requires a significant review of city operational models and processes, aimed to subvert current information access rules. Smart and “citizen-centric” city model design implies that stakeholders will put in place a global mind-change: information sharing is an essential ingredient to destroy obstructions existing in a context based on the “silo-centric” information management approach. Thus, a review of the information management processes is required, one that sees the citizen as the key element of the model. Citizens are those who produce information as input for an Open Data management system (through their implicit and explicit behaviors and their habits and customs), but they are also

those who benefit from the availability of Open Data. In a nutshell, they are both the main supplier and the main consumer of managed information. The design of Open Data creation, publication, and update has to be driven by citizens' requirements and has to be focused on satisfying their expectations. Additionally, feedback produced by citizens should be considered the main input for data upgrades and information system evolution.

Incentivize different actors. The enabling of a smart community through information sharing cannot neglect the involvement of different actors, including the local government, industries, citizen associations, university and research centers, financial institutions, and so on. In order for them to become an active part of Big Data and Open Data projects, it is necessary to highlight the contribution required of each stakeholder and, as a consequence, the related outcomes. This approach requires, first of all, a clear knowledge of stakeholders' interests in making information open and the outcomes that each of them intend to pursue by taking part in the project and sharing "their" information. Those interests must coexist and be balanced in order to guarantee project sustainability at all, provide expected outcomes to stakeholders and heal citizens' "information thirst."

Involving the Territory: from Smart City to Smart Land. A smart city must cross geographical boundaries, i.e., citizen involvement has to be extended to both big cities and small villages. Big Data and Open Data projects have to foster citizens' virtual encounters independently from their physical presence in the territory. At the same time, a big change is needed inside the public sector, i.e., cooperation beyond roles and responsibilities and the partition of local government in order to contribute to citizen common well-being.

3.3 Governance Enablers

A smart city is a data-enabled city, i.e., a city in which data generated by various stakeholders and systems are collected, integrated, explained and allowed to become shared knowledge that can be used and reflected upon to foster increased quality of life, sustainability and efficiency.

The protection of citizens' rights, and indeed the strengthening of them, is an inescapable part of this vision, which is expressed both in terms of public policies (both in a protective and proactive sense) and in terms of industry declinations.

However, it is clear that a data-enabled city requires a reinterpretation of each phase of data treatment, circulation and possible usages in the light of legal rules. The areas of law that are called into question are potentially all-encompassing. In fact, both the data and the rules that guide it, define it, and relate the human and environmental components have an important infrastructural value.

This implies (a) identifying the best rules to enable data circulation, (b) restructuring data management processes such that they are oriented to citizens' rights, (c) curating data and metadata quality to incorporate the rules that affect data processing and usage, and finally (d) adopting Open Data and open source policies.

Rules to enable data circulation need to give particular attention to licensing, access policies for documents and public records, and privacy. To this end, we need clear rules of data ownership, for IPRs and policies oriented to data sharing in general and to Open Data in particular.

Among these, the presence of rewarding policies (with obligations and incentives) that apply to the opening of data in public calls and tenders have crucial value for a transparent and efficient management of the services of the smart city. This is especially true if they are integrated within policies for citizen participation and clear engagement rules such that the city can use data coming directly from the monitoring of calls and tenders as part of the data system that it turn makes services more efficient. Attention should be given to both privacy and data ownership aspects across different public sectors.

It is crucial that the rules of the smart city are in line with the new PSI directive [7] for the reuse of public sector information, which provides a common legal framework for a European market for government-held data (public sector information). The directive is built around two key pillars of the internal market: transparency and fair competition. It extends its scope to museums, libraries and archives (with a set of exceptions and less restrictive rules w.r.t. other subjects). It is based on existing national regulations concerning access rights, privacy and data protection. The general principle for reuse has been revisited, as it makes clear that all the documents falling within its scope (i.e., public by law) must be re-usable for commercial and non-commercial purposes.

Our hope is that the directive becomes increasingly oriented towards real domains and data-centric application services and considers also what is done (in terms of standards, processes and policies) in the private sector to incentivize Open Data. In this way, the offer of services to citizens will be as uniform as possible in the rules and diversified and customized in the specificity of the offer.

The restructuring of data management processes (together with Open Government strategies) needs to be oriented to citizens' rights. This should apply to public authorities, to companies and to any other entity (public or private) that provides services to the city. In this respect, for designing new information systems, we recommend privacy-by-design and participatory design approaches.

Rules should also guarantee that an ethical governance of the city is pursued [8]. For instance, there should be a balance between the necessity of guaranteeing public and citizen safety on the one hand and the privacy and the

individual freedom of the citizens on the other hand. There is a clear tension between transparency and privacy. The first requires accountability mechanisms to be put in place. The second requires access control and privacy policies to be enforced, such as privacy-by design mechanisms.

Data and metadata curation mechanisms need to incorporate, as much as possible, the rules that affect data processing and usage. An example is represented by the possibility of incorporating spatial and temporal attributes that define data validity, as well as authority and ownership information,

This need is not only a consequence of the choice to orient the processes towards guaranteeing citizens' rights (the point above) but also meets requirements for sustainability and better efficiency of the city infrastructure and administration. It also increases the chances for a correct data circulation, even across national borders governed by different legal systems.

Open Data and open source policies are also important aspects to be considered. In order to define the rules for the smart city, it is crucial to take into account the Public Sector Information laws. Open Data and PSI regulations should focus on specific domains, promoting procedural, technological and organizational standards, in dialogue with the Open Data policy for the private sector and standards from industry. Moreover, open-source policies and practices should be encouraged, adopting them as a bottom-up strategy of innovation and growth of specialized knowledge in the service of the city and in continuous review by citizens and in general by users.

3.4 Technological Enablers

At the core of the smart city action of the City of Trento, two main technological enablers have been identified as strategic towards reaching the goal and objectives described in Section II:

1. **An infrastructure to automatically publish data of the municipality of Trento as Open Data**, which is able to collect data from the various civic stakeholders and make them available for free to foster transparency and economic growth;
2. **The Trento City Dashboard**, providing an access point offering various services to citizens that can directly (without further processing) consult data integrated from various providers and data sources;

Important factors that will have to be addressed as part of the technological solutions above include the following:

- **Data sources:** data will come from different, potentially very heterogeneous, sources; sources will include the various departments of the city administration, other local institutions or even companies; some data will

surely come from sensors (e.g., traffic and weather information) or directly from citizens via crowdsourcing techniques (e.g., citizens' comments or reports about faults of services and infrastructures);

- **Data formats:** sources can be structured, semi-structured, or even unstructured; particular effort will have to be devoted to the harmonization of the formats used for data collection and exchange;
- **Data quality:** attention will have to be paid to data quality in terms of accuracy, adherence to standards, data certification and mechanisms to measure specific key performance indicators (KPI);
- **Data traceability:** in order for data to be trusted, accountability mechanisms to track their provenance and assess data authority will be set up; a way of achieving this is to ensure that appropriate metadata "travel" together with data via appropriate standards (for instance, DCAT);
- **Data integration:** the definition of shared data models and integration platforms can surely favor the process of incremental data integration. The ongoing activity around the ComunWeb platform towards harmonizing data from the various local administrations of the province constitutes an example in this direction;
- **Enabling technologies:** given the current global trends, there is already a quite broad landscape of existing technologies supporting Open and Big Data; we will rely on synergies with projects that have already experimented on or have been improving such technologies; in this respect, we will collaborate with the Province of Trento in the context of the Open Data Trentino project.

Regarding the Open Data Infrastructure, public administrations produce, manage and collect data as part of their everyday activities. Though part of it cannot be clearly circulated (e.g., because of IPR or privacy constraints), a significant part of it can actually be freely distributed such that citizens and companies can access and reuse it. It is, in fact, well recognized that simply making data freely available not only allows transparency to be met but also can give the data a great economic value, create new opportunities in the territory and stimulate innovation growth.

For instance, data regarding the public transportation system and infrastructures (e.g., roads, bicycle tracks, traffic information), food and accommodation facilities (e.g., restaurants, hotels), and public events (e.g., festivals, concerts) can be of clear utility to tourists, families and citizens in general.

The Province of Trento already provides Open Data through its institutional websites. The Open Data catalogue dati.trentino.it has been implemented and is managed by a dedicated interdisciplinary group of experts, covering themes from technological to organizational (e.g., legal).

The catalogue is based on the CKAN open source software, which is widely used for such purposes.

Though what the city currently releases as Open Data already converges into this portal, this is done unsystematically and with significant manual effort. Therefore, the next step will consist of putting in place a strategy and an infrastructure to automatically produce, maintain and publish Open Data. In this respect, one important source of data will be the central ComunWeb portal, which already aggregates and organizes data from various local administrations of the province. A common agreed schema for such data is currently under definition and is expected to favor data integration and interoperability.

Regarding the second point above, the goal is to define and implement a system, called the Trento City Dashboard, that allows citizens to have direct feedback on the improvement and integration of municipal services. This instrument will exploit the potential of management and improvement of data quality developed in the field of Open and Big Data, enabling its reuse.

The system will be developed in collaboration with Trento RISE (<http://www.trentorise.eu/it>), a core partner of the ICT network of the European Institute of Innovation and Technology whose goal is to foster local innovation, sustainability and quality of life through a competitive and dynamic knowledge economy.

The dashboard will become the main access point to the services offered by the municipal administration for the different parties involved, with specific methods of interaction and views:

- Internal for officials;
- External for citizens;
- Increasingly integrated with systems and services provided by various public institutions of the territorial system, and firstly with the open data catalogue of the Trentino region.

The ambition is to build a scalable system infrastructure able to progressively integrate internal (to the municipality) and external data sources. The dashboard will introduce a new organizational model as an instrument that is expected to facilitate the construction of an open dialogue with the community and easy access to data already collected and processed and therefore directly “digestible” for citizens and other institutions.

In essence, the aim is to create a technological and organizational infrastructure able to generate cleaned and reliable data that are strategic and functional for the objectives of transparency, efficiency and innovation pursued by the city.

Therefore, our immediate next steps will include the definition of the back-end technological infrastructure, aimed at making systemic all the aspects (including the cultural, organizational and governance aspects) pertaining

to data production and integration within the administration. This initial phase will gradually bring to fruition the integration of the different existing services (and possibly new services) offered by the administration, taking into account citizens’ needs.

IV. CONCLUSIONS

In this paper, we have depicted the current situation of the municipality of Trento with respect to Big Data and Open Data and provided some guidelines that will allow the current practices to evolve from their traditional approach to a smarter one closer to citizens’ needs. Trento has already activated many initiatives in this direction, especially in the field of Open Data, but other initiatives are ready to start, so a clear identification of research and implementation lines is clearly needed. In order to provide the Trento Municipality with adequate guidelines as a Smart City willing to increase quality of life and intensify citizen engagement, the workshop organized in Trento has highlighted four different “enablers” that, in our opinion, can facilitate this long and complex process of becoming a smart city. In the paper, cultural, organizational, governance and technological enablers have been identified, analyzed and specified. What emerged from all four enablers is the need for a great multi-disciplinary and inter-disciplinary approach to the problem. A smart city will be “smart” only if the problem is managed by keeping this multi-faceted approach in mind, crossing the findings and the backgrounds of the different areas, and listening to citizens’ needs and ideas. From this perspective, we believe that Trento possesses most, if not all, of the competences required given the quantity and value of the research centers, universities and startups in the area, where investments in innovation are constantly on the agenda of the local administration.

ACKNOWLEDGEMENTS

This paper reports the outcomes of the discussions that occurred during the “Big Data and Open Data” working group held during the Trento Smart City event on the 10th and 11th of December 2014. We would like to thank all the people who participated in the working group: Claudio Covelli and Giacomo Fioroni from Trento municipality, Maurizio Napolitano from FBK, Francesca Gleria from the Autonomous Province of Trento, Paolo Bertolini from Trentino Network, Matteo Brunati from ePSI and SpazioDati, Alberto Corò from Padua municipality, Luca Paolazzi from TrentoRise, Enrico Gennari from the University of Urbino, and Stefano Leucci from the University of Trento.

REFERENCES

- [1] The data deluge: Businesses, governments and society are only starting to tap its vast potential, Feb 25th 2010.
- [2] Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., and Byers, A. H. 2011. “Big Data: The Next Frontier for Innovation, Competition, and

- Productivity,” McKinsey Global Institute (http://www.mckinsey.com/insights/mgi/research/technology_and_innovation/big_data_the_next_frontier_for_innovation; accessed August 4, 2012).
- [3] Mervis J. 2012, U.S. SCIENCE POLICY: Agencies Rally to Tackle Big Data, *Science*, vol.336, no.6077, pp. 22.
 - [4] Labrinidis A., Jagadish, H. (2012), Challenges and Opportunities with Big Data, In Proc. of the VLDB Endowment, 5(12):2032-2033.
 - [5] IBM 2012, What is big data: Bring big data to the enterprise, <http://www-01.ibm.com/software/data/bigdata/>, IBM.
 - [6] <http://www.comune.trento.it/Comune/Organizzazione-comunale/Amministrazione-trasparente/Altri-contenuti/Open-data>
 - [7] <http://ec.europa.eu/digital-agenda/en/european-legislation-reuse-public-sector-information>
 - [8] Miorandi, D., Maltese, V., Rovatsos, M., Nijholt, A., Stewart, J., Social Collective Intelligence: Combining the Powers of Humans and Machines to Build a Smarter Society, Springer, 2014.
 - [9] Shvaiko, P., Farazi, F., Maltese, V., Ivanyukovich, A., Rizzi, V., Ferrari, D., & Ucelli, G. (2012). Trentino government linked open geo-data: a case study. In *The Semantic Web – ISWC 2012* (pp. 196-211). Springer Berlin Heidelberg.
 - [10] Farazi, F., Maltese, V., Dutta, B., Ivanyukovich, A., & Rizzi, V. (2013). A semantic geo-catalogue for a local administration. *Artificial intelligence review*, 40(2), 193-212.
 - [11] Davenport, Thomas H., and D. J. Patil. *Data scientist*. Harvard business review 90 (2012): 70-76.
 - [12] Giunchiglia, F., Maltese, V. Dutta, B. (2012). Domains and context: first steps towards managing diversity in knowledge. *Web Semantics: Science, Services and Agents on the World Wide Web*: 53-63.
 - [13] Shadbolt, N., O'Hara, K., Berners-Lee, T., Gibbins, N., Glaser, H., & Hall, W. (2012). Linked open government data: Lessons from data.gov.uk. *IEEE Intelligent Systems*, 27(3), 16-24.
 - [14] <http://www.innovazione.comunitrentini.tn.it/Progetti/ComunWEB/Il-progetto-ComunWEB>
 - [15] Bedini, I., Farazi, F., Pane, J., Tankoyeu, I., Leoni, D., Leucci, S. (2014). Open Government Data: Fostering innovation. *eJournal of eDemocracy & Open Government*, 6(1).
 - [16] Davies, T. (2012). Supporting open data use through active engagement. In *W3C Using Open Data Workshop*.
 - [17] Bouquet P., Molinari A. (2013) A global Entity Name System (ENS) for data ecosystems. *Proceedings of the VLDB Endowment*, Volume 6, Issue 11, August 2013, Pages 1182-118