Mapping a City's Activity. A Project of Volunteered Geographic Information Using Mobile Mapping Collection

Giuseppe Borruso and Viola Defend

Abstract The work done deals with the concept of Volunteered Geographic Information and is based on the use of a mobile mapping collection tool to retrieve geographical data from an urban fieldwork. The research has been carried on during the academic year 2014/2015, involving students from the course of Geography of Networks within the post graduate degrees in 'Economics' and 'Business' of the University of Trieste (Italy). The intent was testing the potential of crowdsourcing in retrieving data using a bottom up approach, relying on a set of trained and aware 'urban sensors' as data collectors. This allowed us to derive first-hand geographical data concerning a particular topic and analyze its spatial distribution by means of Geographical Information Systems and spatial analytical tools. The topic studied was represented by the urban 'movida', the analysis of the areas of the city that are more or less active during the days and during the week. This major aim was also coupled with an ancillary one, as the coverage of Wi-Fi hotspots and networks over the urban area of Trieste. It is known that many Italian cities still do not allow a very wide coverage of wireless networks to access the Internet. The city of Trieste, on the border between Italy and Slovenia, is suited with a certain degree of coverage, particularly in main roads and squares, thanks to free Wi-Fi coverage managed by the municipality and an academic network of Eduroam system, quite spread over European and world cities hosting universities and research centers.

Keywords Volunteered geographic information · Citizens as sensors · Smart cities · GIS · Mobile data collection

---

The paper derives from the joint reflections of the two authors. However, Viola Defend realized paragraphs 2, and 5, while Giuseppe Borruso wrote paragraphs 1, 3, 4 and 6.

G. Borruso (✉) · V. Defend
Department of Economics, Business, Mathematics and Statistics,
University of Trieste, Via Tigor 22, 34124 Trieste, Italy
e-mail: giuseppe.borruso@econ.units.it; giuseppe.borruso@deams.units.it

1 Introduction

The current debate on Smart Cities is running on the impact of the ICT (Information and Communication) in fostering sustainability of our urban environment and, in a broader sense, in improving the quality of our lives (Toppeta 2010; Schuurman et al. 2012). The debate is often split in two main fronts On one side we find particularly hardware and software vendors, promoting ICT out-of-the-box solutions, aimed at managing and controlling cities through big operation centers, mainly hosted by local authorities, in terms of traffic, security and other form of 'smart', top down remote control of the different aspects of a city (IBM 2011). On the other side, many authors follow a bottom up approach in which inputs and instances come from citizens and city users as active participants in highlighting issues and possible solutions at urban level, helped by new technologies and tools. The debate is between top down—techy—solutions and bottom up approaches integrated into an urban planning process (Townsend 2013; Murgante and Borruso 2013, 2014, 2015).

In such sense, citizens are becoming more and more data producers, both unawares and knowingly. In the former case, citizens can be unaware of the amount of data they create, by being connected to smartphones with geolocation functionality on and registered to the most popular apps and social networks and media, making it possible to private companies as mobile phone and services providers to rely on a wealth of data from which deriving spatial patterns of movements and occurrences of some phenomena. In the latter case, citizens can be aware 'urban sensors' and deliberately collect data for some purpose, using the mobile devices nearly everybody of us to-date have access to and few skills for data collection (Goodchild 2007).

Scope of the present paper is to present how citizens can be involved in projects dealing with the collection of data with a spatial reference on urban issues, therefore playing a role of 'urban sensors' and provide scholars with original data to work on. The paper is focused on urban issues and on the process of data acquisition, visualization, analysis and representation made possible by combining the citizen science approach of citizens as urban sensors, the implementation of mobile data collection tools by means of low cost solutions, and the opportunities offered by free and open source GIS platforms to perform analysis and cartographic visualization and representation.

The paper is organized as follows. In paragraph 2 the concepts of Volunteered Geographic Information, Neogeography and Citizen Science are briefly presented, included an overview of the cartographic implications of this process. Paragraph 3 is more focused on the role of users in collecting data in a field work, acting as sensors in a bottom up approach in a smart city scenario. In paragraph 4 the methodology adopted is followed, including the workflow made, that concerns the setup of the project, the data collection process and the successive analysis and visualization of the results made through GIS packages (online and desktop) and visualization software. The different tools developed and the software, as well as the analysis on the data after their collection and the visualization chosen are presented. Paragraph 5 deals with the case study: the choice of the mobile data kit to collect data, the volunteers and sensors to engage, open issues. Conclusions and future developments are can be found in paragraph 5.

2 VGI, Neogeography, Public Participation, Citizen Science

The rationale behind the project presented deals with a particular case of Citizen Science where geographical element is essential: Volunteered Geographic Information. It configures not only as a mere data collection plan, but a way to increase awareness and engagement of participants in the community through new technologies (Publications Office of the European Union 2014). This phenomenon of Volunteered Geographical Information (VGI), coniated by Goodchild (2007) sees, as a particular case of Neogeography (Turner 2006, 2007; Graham 2009; Warf and Sui 2010; Eisnor 2006; Elwood 2006, 2008; Fischer 2008a, b, 2009; Ghosh and Dasgupta 2015), the involvement of citizens in collecting geographical data as a particular case of user-generated content (Neis and Zielstra 2014; Hudson-Smith and Crook 2008).

Together with these two characteristics—that are the engagement of private citizens without formal qualification and the creation of geographic information—in Goodchild's definition of VGI we can find other main points:

- A voluntary contribution to the projects;
- A level of uncertainty in the accuracy of collected data;
- It's being a "dramatic innovation" affecting GIS and geography in general.

The scientific data collection in the last century has become an expensive process due to the increased costs related to the high-end equipment and technologies. However, when referring to data acquisition aimed at geographic information retrieval, analysis and representation, nowadays low cost solutions are available, allowing practically everyone to make geographic information.

Mapping procedures and the way spatial information are made and transmitted, have been widely changing over the last years. Communication systems are upgrading quickly, adapting to new technological discoveries. These facts, merged with the new mobile internet and computing devices like personal computers, allowed the switch from printed maps to web and digital one. New kinds of mapping technologies flexible and dynamic way of presenting spatial data to a mobile user based on his context and his profile (Reichenbacher 2001).

The need of a mobile cartography was born from the need of adaptation and customization. New kinds of mapping technologies present in a flexible and dynamic way spatial data to a mobile user based on his context and his profile

(Reichenbacher 2001). That regards many aspects of geographic data processing and transformation into information—and possibly knowledge (Longley et al. 2011).

Data collection is made simpler, faster and low cost to-date. First-hand geographical data can be easily collected by means of portable devices fitted with GNSS—GPS receivers. The GNSS receivers installed inside our portable digital devices are in fact powerful and allow reaching a decent level of precision, spanning from 2–3 m in good conditions to a worse case of 10 or more meters in adverse multipath situations (Pesyna et al. 2015; GPS.gov 2015), typical of cities, where 'urban canyons' effect can cause distortions to the signals received by the satellites. That allows virtually everyone to become an amateur geographical data gatherer or a map-maker and participate to mapping local or global projects. Examples can be found in OpenStreetMap (http://openstreetmap.org) and Wikimapia (http://wikimapia.org), where maps are made completely by users, as results of a 'democratization' of GIS (Goodchild 2007). Other examples include the use of citizens over specific projects over a certain amount of time or the use of finalized user groups to target some specific issues (e.g. former Nokia, Here Maps, maps.here.com). This is because citizens are the stakeholders par excellence and represent the part that actually lives the urban environment and should—for that—be engaged to collect and use shared geographic information.

More in general, talking about citizen science and the involvement of citizens in scientific projects, the role played by citizens can be different depending on the task considered, both in cases it regards data acquisition and in cases it deals with data analysis. Projects concerning a various range of topics have been made, and participants are asked more observation skills than costly equipments. Three main types of public participation in scientific research can be identified, providing different models (Bonney 2009):

- Contributory;
- Collaborative;
- Co-created.

The first one reckons on scientists as designers of projects and public contribute to data collection. Collaborative projects see instead the involvement of citizens also in various phases, from planning, analyzing and diffusion of results. The extreme level of public participation is in the third model, the co-creation of projects between public and scientists in most of the steps of scientific process (Shirk et al. 2012).

Easiest ones relies on the computing power of computer, as in the case of SETI@home (e.g. http://setiathome.berkeley.edu), while in others cases an analysis of data is required to participants (e.g. Planet 4 Terrain, http://planetfour.org). In the last cases, scientists may give some basic or more specific information in scientific methodology: this way, an impressive number of 'researchers' can be a part of the data collection, with a higher level of training. Sensors can give a little help, but most of the work lays on the analysis and interpretation of all the data collected.

In the case of SETI@Home the need was the mere power of the personal computer, but in more recent projects the need switches to the sensing of a human and his capacity of analysis. In the panorama of Citizen Science there are also projects interely designed by citizens: a model of 'science by the people' or Participatory Action Research' (Cooper 2015) and kinds of 'do it yourself science' (European Union 2014).

Citizen science, Volunteered Geographic Information and neogeography have some things in common: people willing to be included in the scientific research, no formal skills—or a very limited training over the cartographic and data acquisition process—, a passion for a topic and the dissemination of these activities. It is interesting to underline that there is an exchange of knowledge between the two sides; that depends according to the effort needed.

The only requirement is an Internet access or a smartphone app: mobile connectivity allows also more closeness between scientists and citizens. The motivations can vary over places and types of projects, but the initial participation is due to personal interest, self-promotion, self-efficacy and social responsibility (Rotman et al. 2014).

On the other side, many firms and companies tend to refuse using data collected through VGI because of the difficulty of evaluating quality, and so does a large part of academic environment, questioning information and source credibility due to the involvement of the public and not only experts. However, many researches (Haklay 2010; Girres and Touya 2010; Cipeluch et al. 2010) show that these data have a relatively high quality:, even if not collected by specialists: for this reason, a lot of participatory projects are considered interesting, mostly if there is a strong need toto involve a wider community, like in the case of Geo-Wiki Project (e.g. http://geo-wiki.org) in which the aims are the geospatial cover validation and improvement of global maps (Fritz et al. 2009; Perger et al. 2014). The fact that these data are collected mostly by non-experts or little trained and qualified volunteers lead to mine for some their credibility (Haklay 2010; Flanagin and Metzger 2008). This is not true in every case: from surveys conducted on OSM users there is evidence that more than 50 % of respondent was not entirely "new" to the world of GIS, geography, geomatics, urban planning, informatics (Neis and Zielstra 2014).

Applications of this particular case of user generated content are rising in many field, from the public utility projects on environmental or social issues, to humanitarian ones. These latter motivations are particularly interesting as many free and open source projects and platforms have been developed. This fact, coupled with the widespread use of smartphones and tablets and Internet connection, allows easier and faster data collection and sharing.

The ways of interaction and participation of citizens in the urban scenario are, in some ways, filled by a technological layer. A 'smart citizen' should not only be involved in the proposals of administrations, but also wish to be a part of a system to improve the city. In this scenario stands our project about the urban movida and the Wi-Fi coverage in the urban area or Trieste (Italy), thanks to the help of a particular category of citizens: students.

3 The Role of Users/Citizens in Collecting Mobile Data

In the recent literature concerning Smart Cities, a significant element is put onto the importance of citizens as the sensors who actively participate to a city's life and activities related to its good operation. As Murgante and Borruso recall (2013, 2014, 2015), citizens as 'urban sensors' can be put among the three 'pillars' sustaining a (smart) city. According to these authors, in fact, a Smart city can be thought as an urban system sustained by three main elements (the pillars) that includes citizens as urban sensors, data, in terms of both open and big data, and a network, or a set of connections among places and people. These elements are organized by an urban governance setting a limited set of rules concerning the interaction in the urban environment. Citizens and urban users can participate publicly to a city's life and build and realize their own services and activities, meeting needs they often know better than the final decision and policy makers. New technologies are of course strengthening such a process, thanks to the possibilities mostly given by the Internet, social networks and portable devices. However, the enhancement given as citizens by such technologies in terms of extension of our capability of sensing the city around us (Batty 2013), must face the still limited use that we do of them. Still to-date, the vast majority of people uses smartphones and portable devices for playing games, social networking, web browsing and professional communications (corporate emails), often blurring private and professional uses but still limiting the use of more advanced functions. In reminding this, still Murgante and Borruso (2014, p. 747) point out how "Our smartness as citizens should therefore be that of using the potential of such devices to exploit our interaction with the city to monitor it and highlight both positive and negative aspects and help its better management".

An important element in setting up a project of urban citizen science is therefore represented by the kind of users that are involved, acting as real sensors of the phenomenon under exam, and on the level of knowledge, awareness and (technical) skills they have to show. As Haklay (2013) points out, different can be the categories of citizens involved.

The 'classic' citizen science involves amateurs in traditional scientific activities; a second category is made by amateurs engaged by the science community in measurement and analysis to set action plans. A third flavor is represented by the citizen cyberscience, where users rely on the capabilities offered mainly by portable devices as smartphones and tablets, now fitted with positioning systems as GPS—GNSS receivers and network connection (Wi-Fi, Bluetooth, GSM—3/4G). Still with Haklay we can find volunteered computing, i.e. citizens download data, perform analysis locally and then send them back to a server; volunteered thinking, where citizens perform classification works, and participatory sensing. This latter involving applications centered on mobile phones capabilities.

In the present research, we relied on citizens as urban sensors in the 'citizen cyberscience' category, as a set of users was trained to install a mobile data collection app on their smartphones and then use it for an urban fieldwork, therefore with a certain degree of autonomy and participation to the project aims and ideas.

4 The Methodology

In this research the aim was to examine an urban phenomenon from data collection to visualization, to analysis to the interpretation of results, as the presence of people in different times of the day and places of the city, to try understanding the patterns of 'movida' in the city of Trieste and also the presence of free Wi-Fi coverage in such areas.

After deciding the topic of our research, we first started with the quest for an adequate mobile application to collect data that needed to be free and easy to use and to set up and allowing a set of different questionnaires to be loaded. That drew us to focus our attention on GeoODK system of data collection and aggregation (Ghosh and Dasgupta 2015; GeoODK Tutorial http://geoodk.com). It is an environment for collecting data via mobile devices and storing them onto a remote server allowing performing of field data surveys. GeoODK Collect is an app, developed for Android operating systems devices, capable of hosting forms and questionnaires, and suitable for collecting a wide set of data, including georeferenced ones. Such an application, whose forms for data collection are fully customizable, can be linked to a remote server where georeferenced data can be stored and processed. GeoODK is based on the ODK—Open Data Kit—environment for data collection, gathering, managing and analysis, with more fitted-for-purposes functionalities from the geographical point of view.

GeoODK can be fully implemented by means of questionnaire and survey forms. That was made possible using a tool named XLSForm, that allows creating forms using a MS Excel spreadsheet and then converting them as XForms, a format usable in platforms for mobile data collection as the ODK—Open Data Kit environment (http://opendatakit.org).

In parallel with the use of an app (GeoODK) and the realization of a form for collecting data, a web-form was also implemented, both to allow users of other devices to collect data (i.e., iOS and Windows phone), and a digital entry of data from a desktop or laptop pc, an useful feature also as for editing already inserted data. The survey form created using XLSForm was then made available also through a web-form as Enketo Smart Paper, a web form usable on portable devices and desktop or laptop pcs to collect data. Enketo Smart Paper is a part of the ecosystem of ODK and OpenRosa community, developed and implemented as a standardized open-source form format (http://enketo.org).

A successive step implied the definition of a set of users to be involved in the data collection campaign and their training. Our initial idea was to involve common citizens in the project, hopefully students and young people in general, potentially more prone to using apps and similar technologies on portable devices and to be involved in project dealing with a matter of interest for them, as the areas of free time and leisure (pubs and bars) and the presence of free Wi-Fi hotspots, both public and private. To do so, a dissemination campaign has been carried on among authors' acquaintances in the city, students of some university courses (Geography of Networks in the post graduate degrees in Economics and Business) and local groups on popular social networks. We realized a training by means of videos and slideshows both presented in seminars and broadcasted through social networks. An important aspect was in fact considered also exploring the potentiality of crowd-sourcing and retrieve useful data for public utility purposes.

Further steps involved the analysis and visualization of the data obtained. As an aggregation server we used a cloud based Ona, as that provided by Ona, a platform provided by a social company that has been developed to host data collected from fieldwork by means of forms and questionnaires on apps on mobile devices or through web forms (http://ona.io/home).

The platform allows also visualization, cartographic representation, editing of the data collected, as well as the realization of statistics and reports. It allows also exporting the collected data in popular formats, as spreadsheets and geographical formats. There are free services but also more complex solutions available through payment of a fee. Different levels of privacy can be set on the data server, so that each project can be shared with mobile data collectors at different levels of security. Although a certain level of visualization and reporting is possible through Ona, we needed to test different modes of cartographic representation and further analysis. This was done by means of GIS packages, as QGIS (http://qgis.org) and also visualization software as CartoDB (http://cartodb.com). In particular scatterplots were used to represent the dataset, as well as 'heat maps', also in their dynamic version.

The analysis on the spatial distribution of point data was done using basic point pattern analysis techniques as scatterplots, and using the more refined Kernel Density Estimation (KDE), a spatial function that allows the transformation of point events in space in a continuous density function over the a region, resulting a pseudo—3D surface that shows a concentration of point features.

The Kernel Density Estimation function creates an estimate of point events' distribution in space, within a searching radius, according to the distance to the point where the intensity is being estimated (Bailey and Gatrell 1995).

$$\hat{\lambda}(s) = \sum_{i=1}^{n} \frac{1}{\tau^2} k\left(\frac{s - s_i}{\tau}\right)$$

$\hat{\lambda}(s)$ is an estimate of the intensity of the events' spatial distribution, measured at location s; s_i is the ith events, k (.) is the kernel function and τ is the bandwidth, or searching radius. Changes in the size of the searching radius allow obtaining more or less smoothed surfaces: wider bandwidths oversmoothe the estimate including distant events, while a narrower ones overemphasize local variations in the events' distribution, this allowing the analysis of the phenomenon at different scales

(Levine 2004). One of the most desirable advantages of using this function, particularly in terms of the representational aspects, is given by the fact that it spreads all over the study region, therefore assigning to each point (or cells, as the study region is 'gridded' for an easier computation) an estimate of the density value (i.e., events per square kilometer) or as probability estimates.[1]

5 The Case Study

5.1 The Choice of the Mobile Data Collection Kit—GeoODK App

As the project was developed as an applicative part of an academic course, one of the most important element was the search for low cost and low skill requiring solutions, in order to be used for a research involving students and with a minimum effort in programming and customizing IT component, the aim being on focusing on the social aspects of the project and therefore leaving us time and resources to concentrate over the kind of data to be collected. We took into consideration many mobile applications, but most of them were limited in terms of number of submitted surveys per day or devices per project.

After several trials (Table 1) we approached the Open Data Kit (ODK) environment, appreciating its structure and application. The ODK Collect app was therefore tested as a tool for mobile data collection, given the high possibilities of customizing its content according to different projects. We decided to use an advanced version of ODK, as GeoODK, open source platform created by the University of Maryland and International Institute of Applied System Analysis that is composed of two main parts. On one side there is the mobile application, GeoODK Collect, that runs on Android smartphones and can be also used in offline mode. On the other side we have an "aggregate" web platform (examples are Formhub, ODK Aggregate and Ona). This web platform, based on html pages, aggregates data, lets you visualize on the screen the map of data, modify or delete data or export data in CSV, KML or XLS files. In addition, for those who do not have a smartphone or have a different operative system running (such as iPhone-users or Windows-users) and cannot use GeoODK Collect, the web system allows to retrieve data using Enketo Smart Paper, a web form of the survey (e.g. http://enketo.org).

The key element, and the element bringing added value to any data collection project, is represented by type of questionnaire or form used for collecting the data.

[1]For a more in depth overview of the Kernel Density Estimation and its characteristics see Diggle (1985), Levine (2004), O'Sullivan and Unwin (2003), Silverman (1986).

Table 1 Comparison between mobile data collection tools tested for the research project

Pros	Cons	Characteristics
ODK collect		
Free	Runs only on Android devices, no iOS, nor Nokia/Windows phone	Open Data Kit (ODK) is based on an architecture of an app for smartphone, a collection server and a set of trained individuals to work on the results
Easy to customize		http://opendatakit.org/use/collect
Allows creation of a form/questionnaire using an excel spreadsheet		
Allows multiple data types and formats		
Store point data		
GeoODK Collect		
Free	Runs only on Android devices, no iOS, nor Nokia/Windows phone	Geo Open Data Kit (GeoODK) is based on an architecture of an app for smartphone, a collection server and a set of trained individuals to work on the results
Easy to customize		Based on ODK Collect architecture
Allows creation of a form/questionnaire using an excel spreadsheet		http://geoodk.com
Allows multiple data types and formats		
Store geographical data as points, lines and polygons		
Kobo Collect		
Free	Runs only on Android devices, no iOS, nor Nokia/Windows phone	Kobo Collect is based on an architecture of an app for smartphone, a collection server and a set of trained individuals to work on the results
Easy to customize		Based on ODK Collect architecture
Allows creation of a form/questionnaire using an excel spreadsheet		http://kobotoolbox.org
Allows multiple data types and formats		
Store point data		
GIS Cloud—Mobile Data Collection		
Easy to customize	Free version limited to one device	Integrated with GIS Cloud platform
Allows creation of a form/questionnaire using an easy interface		http://giscloud.com/apps/mobile-data-collection
Allows multiple data types and formats	Fee for multiple devices	
Store point data		
Integrated with the GISCloud development environment		
EpiCollect		
Free; Easy to customize	Seems to be not adequately supported or developed at present	Web and mobile app for the generation of forms (questionnaires) and freely hosted project websites for data collection
Runs on Android and iOS operating systems		http://epicollect.net

Source Mobile data collection tools site

Table 2 Structure as structured in our project with the full list of questions

Type	Name	Label
Datetime	data_ora	Data e ora della rilevazione
Geopoint	Geolocalizzazione	La tua posizione GPS
Text	nome_evento	Nome dell'evento a cui partecipi o locale in cui ti trovi
Text	indirizzo_evento	Ci puoi dire l'indirizzo o la zona dell'evento?
Image	immagine_evento	Carica una tua foto dell'evento!

An example of the form, contained in the first worksheet of the Excel file "survey"

In this case, the form or survey form could be realized setting up different kind of data to be stored and registered, spanning from text, to numbers and to multi-media contents.

The design of the form is made through an Excel file composed by two standardized worksheets: one is 'survey' (Table 2) and the other is 'choices' (Table 3). In the first worksheet there is the list of questions: for everyone it is to explicit the kind of input required: text, media (photographs, audio, video) numbers, multiple choice questions, GPS or Polygon coordinates, date, time and barcode.

In the other worksheet, one has to specify the options for multiple-choice questions. One, with little more practice, can also add rows to create cascading selects, grouping and nesting questions, skipping questions add hints (http://geoodk.com/xlsform_format.php).

After loading the file in the aggregation server (http://Ona.io), the form is converted to Xform. To start collecting data through GeoODK Collect one has to

Table 3 Second worksheet, "choices", as structured in our project with the list of multiple choices questions

List name	Name	Label
si_no	Si	Si
si_no	No	No
affluenza_evento	Scarsa	Scarsa
affluenza_evento	Discreta	Discreta
affluenza_evento	Buona	Buona
affluenza_evento	Ottima	Ottima
nome_wifi	Eduroam	Eduroam
nome_wifi	Triestefreespots	TriesteFreeSpoTS
nome_wifi	Retelocale	La rete wifi del locale in cui mi trovo
nome_wifi	Nonrilevate	Non rilevo alcun free wifi
intensita_segnale	Scarsissima	1
intensita_segnale	Scarsa	2
intensita_segnale	Media	3
intensita_segnale	Discreta	4
intensita_segnale	Buona	5
intensita_segnale	no_reti	Non rilevo alcun free wifi

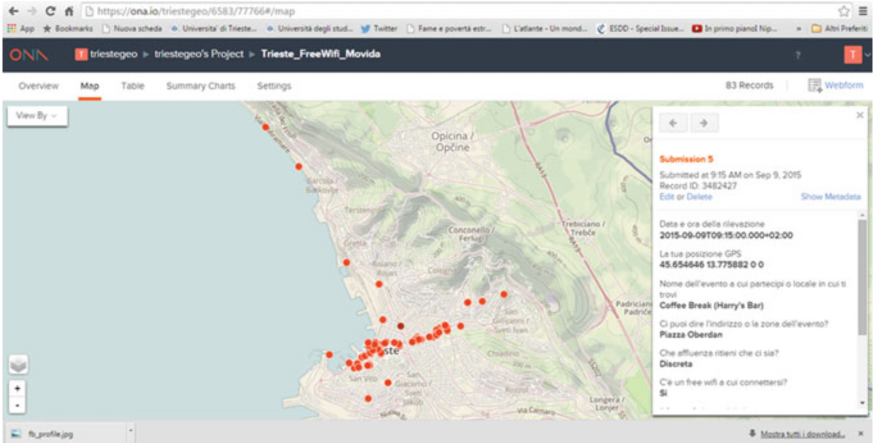

Fig. 1 First analysis of the data collected, available on http://ona.io/home

set the app settings, just for the first time: the procedure requires inserting the URL of the aggregation platform and then downloading the needed form. It is interesting to note that Ona.io allows not only visualizing the data both as a table and as points on a map (Fig. 1), but also to have a first analysis of the data collected and export them in multiple formats for further elaboration.

The use of GeoODK Collect is quite intuitive and does not require particular effort. As portrayed in Fig. 2 the main screen of the app is quite easy to use and present a limited set of options for collecting, editing, mapping and sending data to the aggregation server. Among the possible and interesting characteristics of the

Fig. 2 Main screen of GeoODK collect app and map visualization

GeoODK collect app, is the capability of capturing point events and correct them, in their geographical and attribute data, on the portable device, other than on the server side of the aggregation kit. This can prove to be very useful in cases of difficulties in getting a good GPS location in urban contexts and a need to correct the data already during the campaign.

5.2 *The Choice of the Volunteers and Sensors. Students from Post Graduated Courses*

In starting our project we had different options available to start the mobile data collection campaign and acquire a number of first-hand data. We had the idea to involve citizens and students in particular and incentive them to download an app and 'check in'—as in many popular social networks—in the different places of the city where they go, checking for the number of people located there—and therefore estimating their number—and also for the coverage of the Wi-Fi network, public or private. A second option was that of relying on a small, selected group of sensors citizens that, after a short training on the device and application's characteristic, would go and collect data in the urban environment.

Initially we tried to follow the first path, advertising the project via social networks and media, particularly within the University of Trieste students' community. However, such option did not prove to be particularly effective, with a limited number of people downloading the application and following the instructions on how to perform 'urban citizen science' tasks.

Our choice therefore felt on the second category of sensors, as that of aware and trained people over topics as 'smart cities', 'urban geography', 'GIS and mobile data collection'.

During the course of "Geography of networks", within the post graduate degrees in 'Economics' and 'Business' of the University of Trieste (Italy), students were given the opportunity to interact with people of the local administration to give some ideas to improve the city and city services. As this study of the "city's activity" emerged as an applicative example of this collaboration, we decided to involve the student of this course, suitably trained by the authors through a small presentation to use the mobile application or the web form, as well as on basic GIS and Cartography, as part of the topics covered in class.

Students were a suitable choice to collect these kind of data—urban movida and Wi-Fi, especially in a city, Trieste, where there is almost a 30 % of people over 65 years old (Adnkronos 2013) in a population of around 210,000 inhabitants, and the University students' community alone counts for nearly 20,000 people (including residential and non-residential students). This project wanted to be a way to show the places of activity of the city and have the opportunity to provide some sort of evidence to extend the free Wi-Fi coverage.

In addition to the engagement of this group of students, an advertising campaign on Facebook was made, together with the creation of a Facebook page "Data collection about movida and free Wi-Fi in Trieste". The data collection campaign took place during the summer period and involved volunteers in mapping places in the city of Trieste, stating the presence of public or private Wi-Fi network, as well as he level of affluence in the considered venues. These two elements could be interesting for further studies, in order to map places where people actually move in different days and times of the year, as well as understanding the level of coverage of the Wi-Fi network. These elements were useful both for experimenting visualization techniques and to provide local authorities with a set of information on where to address investments in Wi-Fi development and coverage according to the presence of users.

5.3 *Visualization*

After exporting data in *.csv or *.xls formats, they can be uploaded in various types of both online and desktop version of GIS software. We used CartoDB (http://cartodb.com) and ArcGis in its online and desktop versions, as well as QGIS.

A simple scatterplot of the data shows only a point in correspondence of every survey (Fig. 3): a more interesting map is made with heat map (Fig. 4)—realized by a density estimation—, which shows—as explained in the methodology—with a different colour the presence of overlying data.

This basic visualization gives us an initial idea of the spatial distribution of the locations mapped. An initial observation refers to the presence of locations particularly in the city center of Trieste, in proximity of the 'Piazza Unità d'Italia' and the 'Rive'—major corridors along the sea—and in some main pedestrian streets in the inner city.

Using a heat map (Fig. 4) on CartoDB there is evidence of a concentration of data collected in two main parts of the 'urban movida': Piazza Unità d'Italia, the main square of the city, and Viale XX Settembre, a street where many bars and restaurants are located and open until late hours. These reasons make them

Fig. 3 A scatterplot of the data collected. The survey data in *orange* on the *left* in a CartoDB elaboration; in *blue*, on the *right*, in a QGIS elaboration

Fig. 4 Heat map of data on CartoDB. Data are concentrated in some parts of the centre of the city

frequented during evenings and nights and this is also compatible with the presence of high number of locations registered by the urban sensors involved in the project. The heat map has been realized according to a bandwidth of nearly 400 m, generally considered a good approximation for urban areas analysis, as it generally approximates a 5 min walk, a distance generally easily walkable by people.

Using the torque heat map instead, one can have a temporized map that shows a dynamic data population over time: in Fig. 5 a screenshot of four different moments of the data acquisition campaign are portrayed.

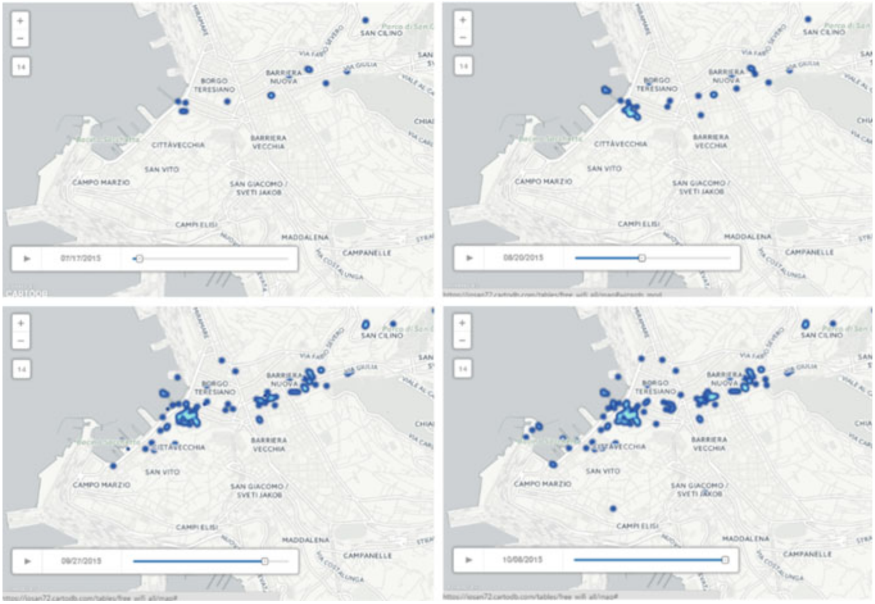

Fig. 5 Torque map of data population over time (*Top left*, 17 July 2015; *Top right*, 20 August 2015; *Bottom left*, 27 September 2015; 8 October 2015). CartoDB elaboration

The dynamic torque map gives us an idea of the data acquisition process that started during the summer time and had its peak at the end of September 2015 and beginning of October, with the mapping of a set of events that took place in the city and that helped in registering the affluence of people and coverage.

5.4 Analysis

The analysed data involved a total of 225 point events: 150 georeferenced data from our project about urban movida were merged with 75 referring to two particular events, Trieste Next 2015, conference and exhibition that took place on 25–27 September 2015, and Fuoriregata, a series of events held on 2–11 October 2015 as a 'backstage' of the more popular sailing race 'Barcolana' and the related activities on the sea shore.

As said before, we were interested about presence of people at places on a 'normal' week and in case of special events organized by the local administration. Data showed that in the 71.43 % of cases there was a discrete or optimal presence (3 or 4 on a scale of 4) to an event or in the bar/restaurant where the survey was taken and only a 9.82 % of poor affluence.

Users were asked to give information about the availability of free Wi-Fi, marking if public or private. In Fig. 6 red dots indicates the absence of a free Wi-Fi: considering that the places near Barriera Nuova are very crowded, this can point out the need of new hotspot in this area; more can be done so behind Piazza Unità

Fig. 6 Users reported *blue dots* indicating availability of free Wi-Fi. *Red dots* points with no free Wi-Fi. CartoDB elaboration

d'Italia. It is evident as some popular places among city-users are still not covered by free—public or private—Wi-Fi network.

5.5 Open Issues

We surely can say that the GeoODK solution we tested in our project was adequate because of its easiness to use and customize. It uses a XML standard working on many platforms, app and servers are free and the app allows to export data in multiple geo data formats.

On the other side, as in many similar projects of Citizen Science, we had a general problem in effectively engage people in the project. In our case, given the tight period of collection, we used a server side "ready-to-use": at first we started with Formhub aggregator (http://formhub.org) but due to problems of overcapacity and maintenance of the site we had to switch to another solution, Ona.io (http://ona.io/home). We had to download the file with already collected data from the Formhub server, but we could not merge it to the already started project on Ona.io. In any case this operation was possible in post-processing and elaboration in QGIS and CartoDB.

This solution was considered interesting by local administration and associations: the University and the team of Fuoriregata asked us to customize the app creating a survey form to retrieve data. During Trieste Next 2015 conference and exhibition, an academic and scientific event connected to the university, GeoODK was used to present two activities: one aimed at creating a digital map of the provenience of the visitors, the other to collect data about urban accessibility. In addition, the research group of the University of Trieste was asked to customize the application to retrieve data about the affluence to the activities and events proposed at Trieste Next. The same thing was asked by the team of FuoriRegata (www.fuoriregata.org), which was about to start a series of events "land side" preceding the well-known Barcolana, an international regatta.

After the activity of mapping with the urban accessibility (presented at Trieste Next), the research group is collaborating with LabAc, a Laboratory of Accessibility that involves University, Province and Municipality of Trieste (Italy), as well as associations of people with disabilities. The project is aimed at mapping walkways and pedestrian nodes in the area in the City of Trieste, where people with disabilities can move freely.

6 Conclusions and Future Developments

In this research, we applied some of the concepts related to Citizen Science and Volunteered Geographic Information to analyze and represent an urban phenomenon and to examine the opportunities given by geographical mobile data collection.

On one side, the attention can be drawn on the use of portable devices as smartphones and tablets and ad hoc apps to collect data “on the move” and on specific locations in a given territory. This can be done relying on the standard tools installed in such devices, as a GNSS receiver, wireless network connection systems, camera, video and audio recorders as well as QR code reader and other tools. Such elements can be integrated into apps that allow collecting georeferenced data of different sources relying on them and on the data input by the user. Our research was aimed at experimenting such integration and the use of free or low cost apps capable of transforming normal devices as mobile data collection ones. For that, we explored the ODK environment and particularly the GeoODK Collect app customized for collecting data and map them both on the device and on the server. The app and the overall environment, including a form input set up and a data collection server were considered very useful as allowing a high level of customization of the form or questionnaires for the data entry and collection, and also very promising for further development. This particularly in term of the kind of geographical data that can be collected, as the ODK collect apps (as KoBo collect and ODK Collect) can only store point events, while GeoODK allows to collect, edit, manage and store polyline and polygon data, de facto allowing a high level of variety on the data type. Also the combination of a standard device and free of charge app and environment allow users to focus just on the content and structure of the data they need to collect, therefore with a limited need for a training.

From another point of view we need to focus on the choice of the ‘citizen scientists’ and therefore on the category of users to be involved in a data collection campaign. A first lesson learnt deals with the difficulty in involving a wide range of people in a project, concerning mapping the most popular parts of the city during events or regular days according to the users’ experience and knowledge, as well as observing the availability of the free availability of a Wi-Fi network in the urban environment. A campaign just based on a limited social network advertising and a ‘word of mouth’, although supported by detailed instructions on how to install the app and customize it in order to collect easily urban data, did not prove to be particularly effective. On the contrary, it was more effective creating a selected group of knowledgeable users, in this case among the students of the course of “Geography of Networks” in the Postgraduate Degree in Economics and Business, where topics as ‘smart cities’, ‘VGI’ and basics of GIS and urban geography were introduced. Students were prepared about the topics throughout the term and during the final part were involved in the research project involving data collection. That allowed relying on a narrower but more focused network of participants in the data collection process.

A third set of conclusions deal with the results of the research area. The results portrayed some levels of clustering of ‘urban activity’ in some areas of the city, as the city centre in the area around the main square of the city (“Piazza Unità d’Italia”) and a major axis along a main walkable road (“Viale XX Settembre”). Both are parts of the city with a high presence of activities at night and during the free time (“movida”), in the second area particularly frequented by students. In addition, some other areas along the coastline were highlighted. The research was also useful in helping highlighting some hotspots in correspondence of some major events happened during the period considered—as the Science and University fair "Trieste Next" and the "Barcolana" sailing race.

We also experimented different modes of cartographic representation and further analysis, using both standard GIS packages for the analysis done—QGIS—and some visualization software as CartoDB for dynamic mapping. Also, a relevance for cartographic matters is given by the level of control on the accuracy of data collected that can be performed relying on trained user groups.

From the lessons learnt from the research, some plans for future research activities can be done. We plan to extend the capabilities of GeoODK and of the aggregation component of the server side. It is in fact our aim to explore the possibilities offered by the app of collecting polyline and polygon data other than the 'standard' point data as we experiment in such research. Also working on the server side is an aim of future research and development. Relying on platforms as Formhub or Ona.io were practical and did not imply so many computational skills, but we faced in some cases instability of the platforms and a slow reaction time in some moments of the day in loading data. In addition, many of this cloud data storage server are not suitable for managing polyline and polygon datasets, so the idea is to implement a local ODK aggregate server in order to have access and control over the data before further elaborations and analysis. The plan is also to set up an adequate campaign of promotion and communication of the future projects involving a bottom up mobile data collection process, in order to allow potential users and data collectors to be and feel more involved in the different parts of the project. This way we want to allow them to participate to the different steps of the project and observing the progresses done. Use of social media and networks as well as other forms of dissemination—working groups and conferences—will be carried on in order to inform about the project, the use of the apps and the state of the research. Also the use of expert users will be enhanced, in this case involving people—as students of the courses more active on the 'smart city' and mobile data collection issues—from the beginning and therefore relying on an audience capable of evolving, training, understanding the issues at stake and learning the functionality of the system as well as the importance of the research activity.

References

Adnkronos (2013, October 18). ANZIANI: A TRIESTE OVER 65 OLTRE 28 %, IN ITALIA ACCADRA' NEL 2030. December 12, 2015 da Conferenza delle Regioni e delle Province autonome: http://www.regioni.it/sociale/2013/10/18/anziani-a-trieste-over-65-oltre-28-in-italia-accadra-nel-2030-318995/

Bailey TC, Gatrell AC (1995) Interactive spatial data analysis. Longman Harlow

Batty B (2013) The new science of cities. The MIT Press

Cipeluch B, Jacob R, Winstanley A, Mooney P, (2010) Comparison of the accuracy of OpenStreetMap for Ireland with Google maps and bing maps. In: Tate NJ, Fisher PF (eds) Proceedings, ninth international symposium on spatial accuracy assessment in natural resources and environmental sciences, accuracy 2010, Leicester, UK, July 20–23, pp 337–340

Cooper C (2015, September 30). Science of the people, by the people and for the people. Tratto il giorno 12 13, 2015 da Scientific American. http://blogs.scientificamerican.com/guest-blog/science-of-the-people-by-the-people-and-for-the-people/

Diggle PJ (1985) A kernel method for smoothing point process data. Appl Stat J Roy Stat Soc Ser C 153:349–362

Eisnor D (2006) Neogeography. http://www.platial.com, accessed 23rd April 2008

Elwood S (2006) Critical issues in participatory GIS: deconstructions, reconstructions, and new research directions. Trans GIS 10(5):693–708

Elwood S (2008) Geographic information science: new geovisualization technologies—emerging question and linkages with GIScience research. Progress Hum Geogr 33(2):256–263

European Union (2014) From citizen science to do it yourself science. JRC Science and Policy Reports, Tratto il giorno 12 13, 2015. Publications Office of the European Union, Luxembourg

Fischer F (2008a) Collaborative mapping—how wikinomics is manifest in the geo-information economy. GeoInformatics 11(2):28–31

Fischer F (2008b) We are excited about maps—so are They!—TeleAtlas feeds and makes use of neogeographers. GeoInformatics 11(7):10–13

Fischer F (2009) Donate your geo-data!—rethinking the geo-information economy with neogeography. GeoInformatics 12(5):12–14

Flanagin A, Metzger M (2008) The credibility of volunteered geographic information. GeoJournal 72:137–148

Fritz S, McCallum I, Schill C, Perger C, Grillmayer R, Achard F, Obersteiner M (2009, August 3) MDPI. Tratto il giorno 12 10, 2015 da Geo-Wiki.Org: The use of crowdsourcing to improve global. www.mdpi.com/journal/remotesensing

GeoODK Tutorials http://geoodk.com/gisday/tutorials.html

Ghosh M, Dasgupta S (2015) How to use open data kit (ODK): a brief tutorial, global change Programme-Jadavpur University working paper # GCP/JU/15/02 http://www.juglobalchangeprogram.org

Girres J-F, Touya G (2010) Quality assessment of the French OpenStreetMap dataset. Trans GIS 14(4):435–459

Goodchild M (2007) Citizens as sensors: the world of volunteered geography. GeoJournal 69(4):211–221

Graham M (2009) NeoGeography and Web 2.0: concepts, tools and applications. J Locat Based Serv Spec Issue NeoGeogr, 18–145

GPS.gov http://www.gps.gov/systems/gps/performance/accuracy/, accessed 2015

Haklay M (2010) How good is OpenStreetMap information: a comparative study of OpenStreetMap and Ordnance survey dataset for London and the rest of Engrland. Environ Plan B 37:682–703

Hudson-Smith A, e Crooks A (2008) The renaissance of geographic information: neogeography, gaming and second life. “UCL Working Papers Series”, 142

IBM’s Smarter Cities Challenge (2011) Rio de Janeiro Summary report. IBM Corporation

Levine N (2004) CrimeStat III, A Spatial statistics program for the analysis of crime incident locations. Ned Levine & Associates, Houston, TX, and the National Institute of Justice, Washington, DC

Murgante B, Borruso G (2013) Cities and smartness: a critical analysis of opportunities and risks. Lecture notes in computer science vol 7973, pp 630–642. Springer-Verlag, Berlin Heidelberg

Murgante B, Borruso G (2014) Smart city or Smurfs City. In: Murgante et al. (eds) ICCSA 2014, Part II, LNCS 8580, pp 738–749. Springer International, Switzerland

Murgante B, Borruso G (2015) Smart cities in a smart world. In: Rassia STh, Pardalos PM (eds) Future city architecture for optimal living, springer optimization and its applications, vol 102, pp 13–35. Springer International Publishing, Switzerland

Neis P, Zielstra D (2014, January 27). Recent developments and future trends in volunteered geographic information. Tratto il giorno December 12, 2015 da MDPI. www.mdpi.com/journal/futureinternet

O’ Sullivan D, Unwin DJ (2003) Geographic information analysis. Wiley, Chichester

Perger C, LeDrew E, Fritz S, See L (2014 September 29). Geography Geo-Wiki in the classroom: using crowdsourcing. Tratto il giorno December 15, 2015 da MDPI. www.mdpi.com/journal/futureinternet

Pesyna KM, Heath RW, Humphreys TE (2015) Accuracy in the Palm of Your Hand, GPS World, 2 February

Reichenbacher T (2001) The world in your pocket. Towards a mobile cartography. In: Conference paper, international cartographic conference

Rotman D, Hammock J, Preece J, Hansen D, Boston C, Bowser A, He Y (2014) Motivations affecting initial and long-term participation in citizen science projects in three countries. In: iConference, Tratto il giorno 2 2015 (p. 110–124)

Schuurman D, De Marez L (2012) Structuring user involvement in panel-based Living Labs. Technol Innov Manage Rev 31–38

Silverman BW (1986) Density estimation for statistics and data analysis. Chapman Hall London

Toppeta D (2010) The smart city vision: how innovation and ICT can build smart, “liveable”, sustainable cities. Think! Report 005/2010

Townsend AM (2013) Smart Cities. Big data, civic hackers, and the quest for a new utopia. Norton & Company, New York

Turner AJ (2006) Introduction to neogeography. O’ Reilly Media, Sebastopol, USA

Turner AJ (2007) Neogeography and GIS, GISDay 2007. University of Kansas, 16 Nov 2007

Warf B, e Sui D (2010) From GIS to neogeography: ontological implications and theories of truth. Ann GIS 16(4):197–209